



Proceedings of the **63rd**

International Conference of
Scandinavian Simulation Society

SIMS 2022



Credit photo front page: Thomas Høstad/NTNU

This template originates from [LaTeXTemplates.com](https://www.LaTeXTemplates.com) and is based on the original version at:
https://github.com/maximelucas/AMCOS_booklet

Contents

About	4
Title	4
SIMS	4
Hosting Institutions	4
Editors	4
Published by	5
Preface	6
Welcome to SIMS 2022	6
Norwegian organizing committee	7
Timetable	8

About

Title

Proceedings of the 63rd International Conference of Scandinavian Simulation Society, SIMS 2022, Trondheim, Norway, September 20-21, 2022

SIMS

SIMS is the Scandinavian Simulation Society with members from the five Nordic countries Denmark, Finland, Norway, Sweden and Iceland. The SIMS history goes back to 1959.

SIMS practical matters are taken care of by the SIMS board. It consists of two representatives from Denmark, Finland, Norway, Sweden, and Iceland.

SIMS annual meeting takes place at the annual SIMS conference or in connection to international simulation conferences arranged in the Nordic countries.

www.scansims.org

Hosting Institutions

NTNU Norwegian University of Science and Technology

OsloMet Oslo Metropolitan University

USN University of Southeast Norway

UiS University of Stavanger

Editors

Lars O. Nord, Tiina Komulainen, Corinna Netzer, Gaurav Mirlekar, Berthe Dongmo-Engeland, Lars Eriksson

© 2022 Scandinavian Simulation Society and the Authors. All rights reserved.

Published by

Scandinavian Simulation Society & Linköping University Electronic Press, Linköping, Sweden

ISBN: 978-91-7929-545-5

Series info: Linköping Electronic Conference Proceedings Nr. 192

ISSN: 1650-3686

eISSN: 1650-3740

DOI: <https://doi.org/10.3384/ecp192>

License: Creative Commons Attribution 4.0 International

License URL: <https://creativecommons.org/licenses/by/4.0/>

Year of Publication: 2022

Preface

Welcome to SIMS 2022

The Norwegian University of Science and Technology (NTNU) in collaboration with Oslo Metropolitan University (OsloMet), University of Southeast Norway (USN), University of Stavanger (UiS) and the Scandinavian Simulation Society were proud to arrange the 63rd International Conference on Simulation and Modelling (SIMS 2022) in Trondheim, Norway.

About 500 kilometers north of Oslo lies a beautiful city called Trondheim in Trøndelag county. This municipality has approximately 200,000 inhabitants. This makes Trondheim Norway's third largest municipality. The Nidelva river flows through the city from south to north, and empties into the Trondheim fjord. Trondheim has a wide range of rich outdoor and cultural offer. There are many historical tourist attractions, such as Nidaros Cathedral, Munkholmen and Kristiansen Fortress. In addition, there are many natural attractions, such as Ringve Bay, Korsvika, Ilabekken, Ringve Botanical Garden and Burma Klippen. You can also be a part of the city's student and technology life as country's largest state university, NTNU, has its main campus in Trondheim.

The Scandinavian Simulation Society consists of members from five Nordic countries: Denmark, Finland, Norway, Sweden, and Iceland. The goal of SIMS is to develop further the science and practice of modelling and simulation in all application areas and to be a forum for information interchange between professionals and non-professionals in the Nordic countries.

The SIMS 2022 conference covered broad aspects of recent research and development work in modeling, simulation and optimization in engineering applications. The scientific program included technical sessions with submitted papers. Ph.D. students were especially encouraged to contribute with papers according to the conference themes. Conference Themes included:

1. Modeling and simulation for design, planning, optimization, control, and monitoring
2. Tools for modeling and simulation, numerical methods for simulation, novel techniques
3. Visualization of modeling and simulation results

Application areas included:

- A. Renewable energy systems: bioenergy and biofuels, geothermal, hydro, solar, thermal, wave, tidal, and wind energy
- B. Hydrogen technologies: production, storage and transportation, hydrogen value chain

C. Energy systems: electric power, energy storage, fuel cells, heat pumps, industrial plants, energy use in buildings, power plants

D. Transportation: automotive, hybrid and electrical vehicles, marine, infrastructure

E. Industrial processes: carbon capture and storage, chemical processing, hydrogen production, oil and gas, and water treatment, cyber-physical systems, biosystems and medical systems

A key component of the SIMS 2022 conference was the opportunity to socialize and make new connections. This was the first time for the conference to take place physically after COVID-19 pandemic. The conference offered several possibilities for networking including a conference dinner at the restaurant AiSuma, a social mini-golf event at Trondheim Camping, a conference study tour to various laboratory facilities at NTNU, and last but not least several coffee and lunch breaks.

On behalf of the organizers, we wish to thank all participants, authors, keynote speakers, session chairs and presenters for their contribution to this conference! We also want to acknowledge the support we have received from the conference board, the program committee, and the SIMS board. Finally, we hope that you will find the proceedings to be a valuable resource in your professional, research, and educational activities whether you are a student, academic, researcher, or a practicing professional.

Lars O. Nord, Tiina Komulainen, Corinna Netzer, Gaurav Mirlekar, Berthe Dongmo-Engeland, Lars Eriksson

Norwegian organizing committee

Ambrose Ugwu

Berthe Dongmo-Engeland

Britt Moldestad

Corinna Netzer

Debbie Koreman van den Bergh

Gaurav Mirlekar

Haoran Li

Kristian Thorsen

Lars Ivar Hatledal

Lars O. Nord

Marcin Pilarczyk

Nils-Olav Skeie

Tiina Komulainen

Vadim Engelson

© 2022 Scandinavian Simulation Society. All rights reserved.

Timetable

KL: Keynote Lecture, IT: Invited Industry Talk, PS: Parallel Sessions, LT: Lab Tour

Monday, 19 of September

18:30–19:00	On-site registration at Trondheim Camping
19:00	Social event: Indoor mini-golf at Trondheim Camping Olav Tryggvasons gt. 5

Tuesday, 20 of September

8:30–9:00	Registration, Radisson Blu Royal Garden Hotel			
9:00–9:20	Welcome remarks			
9:20–10:00	KL	Terese Løvås NTNU, Trondheim	How can numerical modelling support exploring ammonia as a carbon-free fuel?	
10:10–11:10	PS	District Heating, Cooling & Resources	CO ₂ Capture I	Medical Applications
11:10–11:30	Coffee break			
11:30–12:30	PS	Sustainable Fuel Production & Purification I	CO ₂ Capture II	Control Engineering I
12:30–14:00	Lunch break			
14:00–14:30	IT	Even Solbraa, Equinor	Development and user experience for an online simulator for natural gas dehydration processes	
14:30–15:00	IT	Adriaen Verheyleweghen, Cybernetica	Industrial physics-based non-linear model predictive control	
15:00–15:20	Coffee break			
15:20–16:40	PS	Sustainable Fuel Production & Purification II	Electric Power	Control Engineering II Solver & Method Development
16:40	End of day 1			
16:45–17:45	SIMS General Meeting			
19:00	Conference Dinner - AiSuma, Kjøpmannsgata 57			

Wednesday, 21 of September

9:00-10:30	LT	Lab tour at Gløshaugen Campus		
10:30-11:10	Walk Gløshaugen Campus to Radisson Blu Royal Garden Hotel			
11:10-11:30	Coffee break			
11:30-12:30	PS	Oil Production	Multiphase Modelling I	Data-Driven Modelling I
12:30-14:00	Lunch break			
14:00-15:00	PS	Buildings	Multiphase Modelling II	Data-Driven Modelling II
15:00-15:20	Coffee break			
15:20-16:40	PS	Fluidized Beds	Multiphase Modelling III	Heat to Power Applications
16:40-17:00	Best Paper Awards & Closing			
17:00	End of Conference			

Development of a simulation tool for design and off-design performance assessment of offshore combined heat and power cycles

Mohammad A. Motamed^{a*}, Lars O. Nord^a

^a NTNU - The Norwegian University of Science and Technology, Department of Energy and Process Engineering, NO-7491, Trondheim, Norway

*corresponding author: mohammad.a.motamed@ntnu.no

Abstract

Ambitious targets for reducing carbon dioxide (CO₂) emissions are set by Norwegian authorities to address the concerns about global warming. Emission reductions in the offshore heat and power sector can play a role in reaching these targets. Parts of the efforts in industry and academia to reduce offshore emissions are concerned with introducing new design configurations or proposing novel operational strategies for the combined heat and power cycles. Therefore, there is a desire to have a fast and reliable design and assessment tool to be used in the early design stage. Here, a generalized design and performance simulation tool is developed presenting a design point and off-design simulation of the offshore heat and power cycles. It helps the designer provide a fast and accurate thermodynamic assessment of proposed design solutions. The tool has a graphical user interface to facilitate working with the tool with a minimum level of effort and background knowledge from the user. Five part-load control strategies are included in the tool. The tool is verified with available data in the open literature and the results are shown to be in good agreement with the reference data. A combined heat and power cycle is designed and simulated at part-loads as a case study. The cycle includes a gas turbine, a process heat extraction unit, and an organic Rankine bottoming cycle. The simulated performance of the design case in various control strategies is compared showing a 2.5% emission reduction relative to the baseline control strategy.

Keywords: Process simulation, Variable area nozzle turbine, Sliding pressure, Offshore heat and power, Organic Rankine cycle, Carbon emission

1 Introduction

Norway and Iceland have targeted to cut CO₂ emissions by at least 40% relative to the level from 1990 [1]. Norway has pushed the target further up to 55% under the Paris agreement [1]. Oil and gas extraction activities have the highest share of total CO₂ equivalent emissions in Norway. About 27% of total emissions in 2020 originated from oil and gas installations [2]. A potential solution to reduce CO₂ emission in offshore oil and gas installations is producing extra power from the recovered waste heat of gas turbines (GTs). It was shown in [3] that smaller size gas turbines have the opportunity for higher power recovery from the waste heat per unit of the installed gas turbine power size.

Organic Rankine cycles (ORC) have shown competency for low-footprint power production as the bottoming cycle in offshore installations. They are compact and can operate autonomously with lower operating and maintenance costs relative to steam bottoming cycles [4]. Offshore combined heat and power cycles can be accompanied by renewable

energy sources for carbon-reduced power production. Intermittent availability of renewable energies puts gas turbines and the bottoming cycles in part-load operation for most of their lifetime. Therefore, several efforts are seen in the open literature to further optimize ORCs in the off-design part-load operation. The improvements include the development of new operational strategies, component performance upgrades, finding suitable working fluids, and proposing layout design solutions. It was studied in [3] how different ORC configurations with a heat transfer interloop and recuperators can influence the system performance. An optimization on determining the most appropriate organic working fluid among 39 different candidates is carried out in [5]. The study showed that the optimal ratio of fluid critical temperature to the cycle evaporation temperature lies in the range of 0.93 to 1.02. Manente et al., presented an off-design simulation model to optimize the control strategy in an ORC [6]. They showed that ambient temperature has a great

influence on the cycle performance in air-cooled systems. A thorough insight into how ORC components upgrade can influence the cycle performance is presented in [7].

With the ongoing improvements in the industry and academia on ORC performance, it is desired to have a fast and accurate design and simulation tool for performance assessment of the system in the early stages. A knowledge gap is identified in simulation tools covering different part-load control strategies in a gas turbine – ORC combined cycles. A tool that can provide users with the flexibility to select the configuration and operational strategy. Here an in-house tool is developed to design and simulate an ORC cycle with different part-load control strategies. The tool has three main featuring sections. The design section provides a design tool that enables the designer to have a fast assessment of the ORC in the design point. The parametric study section determines the performance behavior of a designed candidate in the design choice range. The simulation section evaluates the key performance indicators under different control strategies at off-design operation. An arbitrary working fluid is allowed in the tool. Therefore, users can choose among about 100 known working fluids in the library. A calculation of the required area and volume of the heat exchangers is provided to help the designer have a good estimation of the design case footprint.

2 Method

The design and simulation algorithm, scientific background, and mathematical formulations are presented. The calculations are based on basic thermodynamic and fluid dynamic principles. The open-source CoolProp package is used as the thermodynamic library to determine the thermodynamic properties of the fluids in each thermodynamic state. CoolProp is a comprehensive and free thermodynamic database that provides a fast and accurate estimation of thermodynamic properties for a wide range of organic fluids [8]. A list of available substances as the working fluid is available on the CoolProp reference list.

The cycle performance is estimated by identifying the thermodynamic states in the intercomponent stations in the cycle. It is known that two independent thermodynamic properties are sufficient to determine the thermodynamic state of a single-phase substance uniquely [9]. Mass flow rate is the third parameter to determine component performance. Therefore, three parameters at each station are necessary and enough to determine a component's performance. Each station is then identified by setting two thermodynamic properties and the mass flow rate passing through that intercomponent point. A pinch point temperature difference (PPTD) approach is taken to design temperatures in the heat exchangers.

2.1 System layout configuration

A single spool gas turbine is used as the topping cycle. In a single spool gas turbine, the compressor (COM), the turbine (TUR1), and the electric generator (GEN1) are mounted on the same shaft. Energy is added to the cycle by burning the fuel in the combustion chamber (CC). Process heat is extracted from the gas turbine exhaust at heat exchangers (Hex) before the bottoming cycle waste heat recovery unit. An intermediate heat transfer oil loop is placed between hot gas flow and the ORC to avoid direct contact of oxygen-rich exhaust gas and the organic fluid. A cascade layout is chosen for the combined cycle to avoid power capacity disturbance upon process heat demand change [4]. The exhaust heat then is retracted in a superheater (SUP), an evaporator (EVA), and an economizer (ECO) placed in series to provide heat to the ORC cycle. The ORC cycle is a simple cycle that consists of a variable frequency drive pump, an economizer, an evaporator, a superheater, a throttle valve (in a throttling scenario), and a turbine expander (TUR2), and a condenser (CND). A cooling loop supplied by the sea freshwater is located downstream to absorb the ORC rejected heat. A second electric generator (GEN2) is coupled to the ORC turbine to convert the shaft mechanical power to electric power. Cycle configuration and intercomponent station nomenclatures are illustrated in Fig. 1.

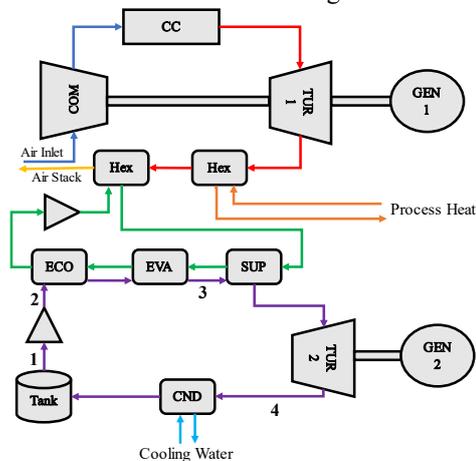


Figure 1: Gas turbine ORC layout configuration. (The abbreviations are explained in the Nomenclature.)

Table 1: Design point input/output list

Input Data	Output Results
Working fluid name	ORC power
Process Heat demand	Working fluid T_c
Superheating	ORC pressure ratio
$EGT, \dot{m}_{GT}, \eta_{GT}$	η_{ORC}, η_{CC}
$\eta_{pump}, \eta_{turbine}$	\dot{m}_{ORC}
T_3, T_1, T_{cw}	T_{stack}
$PPTD_{ECO, IOL, CND}$	$\epsilon_{EVA}, \epsilon_{ECO}$
$\Delta P_{ECO, EVA, CND}$	$PPTD_{EVA}$

2.2 ORC design tool

The design module in the tool gets the input parameters and provides the user with output results based on the design calculations. Tab. 1 includes the list of required input data and output results. This section includes the sequence of design calculations.

2.2.1 Process heat

Hot gas leaving the gas turbine undergoes a constant pressure heat transfer to supply the process heat required by the platform. The amount of needed heat, the exhaust gas mass flow rate, and the exhaust gas temperature are set by the user as inputs. Discharge air properties are calculated based on the conservation of energy law.

2.2.2 Intermediate oil heat exchanger

The flue gas then passes through an intermediate oil heat exchanger to transfer the energy from hot air to the ORC working fluid. A temperature drop equal to the allowed PPTD is imposed to determine the oil temperature. Air pressure drops in the heat exchangers are neglected as they do not influence the ORC performance.

2.2.3 ORC heat transfer

Heat transfer to the organic fluid is carried out in three main heat exchangers: an economizer, an evaporator, and a superheater (Fig. 1). Evaporation and condensing pressure are determined according to the saturation temperature in the evaporator and condenser. P_1, P_2, P_3, P_4 are then calculated based on the pressure loss values provided by the user as inputs. The thermodynamic cycle is designed to have the working fluid in the liquid saturation phase at the economizer-evaporator interface. Evaporator discharge temperature is elevated by the degree of superheating to set the turbine inlet temperature in the superheated gas region. The cooling water mass flow rate through the condenser is designed for having the allowable PPTD at the condenser's hot side.

2.2.4 ORC pump and turbine

Knowing the inlet conditions of the pump and turbine, the discharge thermodynamic properties are calculated based on the isentropic efficiency concept. The pump or turbine discharge enthalpy is determined by knowing the pressure ratio across the component, the isentropic efficiency, and the inlet thermodynamic conditions.

$$h_{out} = h_{in} + (h_{out,s} - h_{in})\eta_s^n \quad (1)$$

Where n is 1 for a turbine and -1 for a pump.

2.2.5 Stack

Flue gas temperature is determined by setting the temperature difference at the economizer cold side to the allowed PPTD value. The organic fluid mass flow rate is adjusted to reach the desired temperature difference at the economizer cold side. Heat transfer effectiveness of the evaporator and the economizer

are calculated by knowing the air temperature on both sides of the heat exchangers.

2.2.6 ORC performance

Cycle power output is simply the difference between power generated by the turbine and power consumed by the pump. The turbine/pump power is determined based on energy conservation law and knowing the mass flow rate, inlet, and discharge conditions of the component. Mechanical efficiency and electrical generator efficiency are set to unity in this work but can be adjusted in the code. Cycle thermal efficiency is defined as the ratio of power output to energy input in the cycle. Where delivered energy is calculated from the enthalpy difference between the superheater hot side and the economizer cold side. Finally, the combined cycle efficiency is defined as the ratio of ORC and gas turbine power output divided by the energy input to the system.

$$\eta_{cc} = \eta_{GT} + (1 - \eta_{GT}) \frac{\dot{W}_{ORC}}{\dot{m}_{GT}(h_{EGT} - h_{amb})} \quad (2)$$

2.2.7 Heat exchanger size estimation

The heat exchanger's footprint is estimated based on the required effective area for the heat transfer. A generic shell and tube configuration is assumed for the heat exchangers. The generic model proposed in [10], allows for fast design and acceptably accurate performance estimation in the early design stage. The heat exchanger effective area calculation procedure is adapted from [11], [12]. Heat transfer effectiveness is calculated according to the heat exchanger temperatures.

$$\epsilon = \frac{T_{hot,in} - T_{hot,out}}{T_{hot,in} - T_{cold,in}} \quad (3)$$

$$C_{min/max} = \min/\max[C_{cold}, C_{hot}] \quad (4)$$

$$CR = \frac{C_{min}}{C_{max}} \quad (5)$$

$$UA = \frac{C_{min}}{CR - 1} \log\left(\frac{1 - \epsilon}{1 - \epsilon CR}\right) \quad (6)$$

The overall heat transfer coefficient is calculated based on the method suggested in [10], [12] and the Nusselt number, Reynolds number, Prandtl number, and friction factor inside the heat exchanger tubes.

$$Nu = \frac{(Re_D - 10^3)Pr \frac{C_f}{2}}{1.0 + 12.7 \sqrt{\frac{C_f}{2}} (Pr^{2/3} - 1)} \quad (7)$$

$$\alpha = \frac{Nuk}{D} \quad (8)$$

$$U = [\alpha_{hot}^{-1} + \alpha_{cold}^{-1}]^{-1} \quad (9)$$

By knowing the overall heat transfer coefficient, the required effective area is calculated from equations 6 and 9. Afterwards, the width and volume of heat exchangers are estimated according to the effective area needed. It is assumed for the heat exchangers to

have 1000 pipes with 30 mm diameter and 90-micron surface roughness. These assumptions can be changed by the user. The on-design calculation is repeated for various design points upon user request in the parametric study mode. The designed thermodynamic cycle is graphically generated in a temperature-entropy diagram and is shown to the user.

2.3 Off-design Simulation

Plant off-design simulation includes performance analysis of the combined cycle in 30%-100% of gas turbine power loads. Gas turbine discharge temperature, mass flow rate, and thermal efficiency are the inputs to the off-design calculations. ORC power output, ORC thermal efficiency, and combined cycle efficiency at the power load range are the output results from the off-design analysis tool. The simulation is carried out for five different ORC control strategies. The simulated control strategies are sliding pressure, throttling valve control, partial admission turbine control, variable area nozzle (VAN) turbine control, and cooling water flow rate control strategies. The five implemented part-load control strategies in the tool are introduced here while a more detailed explanation of their operation principle and background can be reached in [13].

2.3.1 Control strategies

A controller with a sliding pressure strategy uses a variable speed pump that adjusts the cycle flow rate by manipulating the pump's rotational speed. The evaporation pressure slides to match the cycle with reduced heat available to the ORC in part-loads.

A VAN turbine has pivoted vanes as stator blades in the turbine stationary part. In this strategy, the evaporation pressure is kept as high as possible in the part-loads by adjusting the turbine vanes setting angle and modifying the turbine performance without blockage losses [14]. Partial admission turbine control logic has the same strategy as with variable area nozzle except that the mass flow rate admitted to the turbine is regulated by changing the turbine inlet annulus area [15]. The VAN turbine control strategy shows higher part-load cycle efficiency than the partial admission turbine strategy due to less aerodynamic pressure loss in the turbine inlet passage throughflow [16].

The throttling part-load control strategy uses a throttle valve placed at the evaporator discharge to regulate the pressure of the flow entering the turbine. The turbine inlet pressure and mass flow rate are reduced simultaneously in part-load to adjust the ORC power output according to the waste heat available from the gas turbine.

In the cases where a gap exists between the ORC condensing temperature and the supply cooling water temperature in the design point operation, a cooling flow adjustment can be used to allow more efficient power regulation in part-load. Despite the previously mentioned control strategies, the cooling

flow control logic can accompany various control logics to further boost the part-load ORC efficiency. The off-design simulator is an optimizer that finds the optimal thermodynamic cycle in each off-design operational condition. The optimization target is set to be the plant combined-cycle thermal efficiency but can be easily changed to ORC power output, ORC thermal efficiency, or any other desired figure of merits. The optimizer undergoes a simple plane search between all possible manipulating parameters to find the optimal operating point in each off-design condition. In each part-load condition, an ORC cycle is established including the thermodynamic state and the mass flow rate at all intercomponent stations, and the corresponding setting variables in the controller. The number of manipulating parameters is three in the cooling flow control strategy, two in variable area nozzle and partial turbine control strategies, and one in the throttling and sliding pressure control strategy. The cycle mass flow rate is a manipulating parameter in all mentioned control strategies. The cycle pressure ratio is the second setting parameter in the control logic with more than one degree of freedom. Condensing temperature is the third manipulating parameter used in the cooling flow rate control strategy.

2.3.2 Heat exchangers performance

Off-design pressure drop and heat transfer effectiveness in the heat exchangers deviates from the design values in part-load operation and are simulated based on the method presented in [13]. A β parameter is introduced which accounts for the change in the heat transfer coefficient of a heat exchanger [17]. The shift in heat transfer coefficient is calculated according to the change in flow Nusselt number and the fluid conductivity.

$$\beta = \left(\frac{Nu}{Nu_{dp}}\right)\left(\frac{k}{k_{dp}}\right) \quad (10)$$

$$\frac{U}{U_{dp}} = \frac{2\beta_{cold}\beta_{hot}}{\beta_{cold} + \beta_{hot}} \quad (11)$$

2.3.3 Turbine off-design performance

Turbine off-design performance prediction is a calculation of turbine pressure ratio and isentropic efficiency for given mass flow rate, inlet flow conditions, and rotational speed. To accomplish the analysis in given off-design working conditions, a performance map is introduced for each turbine geometry. It is a graphical diagram that illustrates the quantitative relation between four non-dimensional parameters determining the turbine performance [18].

A generalized turbine performance map is used since very little geometrical information is available in the early design stage. The turbine performance map and the pump performance map are adapted from [19] and [20], respectively. The performance maps are normalized and scaled to the design mass flow rate and pressure ratio values. This approach offers flexibility to designers for locating the design

point locating in the performance map. In this work, a choke margin is defined to parametrize the location of the turbine design point in the design speed line of the turbine performance map. Choke margin is defined as the ratio of the mass flow rate difference between the design point and maximum mass flow rate in the design speed line over the difference between maximum and minimum mass flow rates in the design speed line.

A generalized turbine efficiency model is presented in [21] for radial turbines and is used here. Turbine performance change due to adjusting the setting angle in the variable nozzle vanes is modelled in [19] for the range of 20% to 144% of vanes opening angle. The required turbine vane opening angle in each part-load condition is determined based on the desired mass flow rate and pressure ratio through the turbine. Afterwards, the variable area nozzle turbine isentropic efficiency is calculated using the specific speed parameter, the setting value for the vanes opening angle and the model presented in [19]. Specific speed is a well-known indicator for turbomachines which accounts for a combination of mass flow rate and pressure rise.

$$Ns = \frac{NQ^{1/2}}{\Delta h_s^{3/4}} \quad (12)$$

Partial admission turbine efficiency at off-design conditions is simulated based on the method presented in [13]. The throttle valve performance placed before the turbine is predicted according to the method presented in [13] where a constant enthalpy pressure reduction is considered in the valve.

2.3.4 Off-design solver algorithm

The algorithm of the off-design solver is introduced here. A design point calculation is carried out before starting an off-design analysis. Therefore, design data required in the off-design analysis would be available to the solver. The off-design optimization is a plane search over all possible sets of manipulating parameters and then picking up the optimal one by comparing them to all other sets of target outputs. An n-dimensional array of manipulating parameters is set representing the possible options of the controller setting in different control strategies. Where n is the controller degree of freedom in each control strategy. Tab. 2 represents the list of manipulating parameters for the studied control strategies.

Table 2: manipulating variables as the controller setting

Control strategy	Manipulating parameters
Sliding pressure	\dot{m}_{pump}
Throttling valve	\dot{m}_{pump}
Partial admission turbine	$\dot{m}_{pump}; PR_{turb}$
VAN turbine	$\dot{m}_{pump}; PR_{turb}$
Cooling flow	$\dot{m}_{pump}; PR_{turb}; T_{CND}$

Evaporation and condensing temperature/pressure are determined using the pressure ratio value and assuming no change in the condensing temperature; except for the cooling flow rate control strategy where the condensing temperature is set by the controller. The pump discharge pressure is calculated based on the modified pressure loss in the heat exchangers. The pump outlet temperature is identified using the pump isentropic efficiency resulting from the pump performance map. The economizer hot/cold side temperatures are calculated according to the modified heat transfer coefficient of the heat exchangers. The implicit equations for heat exchanger's inlet/outlet temperatures require an iterative approach to solve. The trial-and-error method is used for this iterative solver. After determining the heat exchanger's temperatures, turbine inlet conditions are calculated accordingly. With all thermodynamic states and mass flow rates known, ORC performance (power output and thermal efficiency) is calculated.

2.4 Software framework

The simulation process is implemented as an in-house code in MATLAB [22]. A user interface is integrated into the code to facilitate using the tool for the users. The inputs can be fed into the tool both through a graphical table interface and through a data file. The user has the option to save and load both input data and results for a more convenient operation with the tool. Gas turbine off-design input data is delivered to the tool in the form of an Excel spreadsheet.

2.5 Tool verification

The validity of the tool is assessed by checking the output results against available data in the open literature. Simulation results are compared with the information presented in [23]. The verification is carried out in the rated power and 50% power load. In the reference work, A Solar Centaur 50 gas turbine is working as the topping cycle and an ORC operates as the bottoming cycle. Two different organic fluids are covered in the verification process to check the dependency of the results on the working fluid. The verification data are tabulated in Tab. 3 showing approximately 1% relative error in the results. Therefore, it could be inferred that the simulation tool results are in good agreement with openly published data in the literature.

2.6 Case study

The developed tool is named ORCSIM and is used to design and simulate a combined cycle. SGT800 gas turbine is considered as the topping cycle on the platform. The gas turbine load is controlled using compressor variable guide vanes to prevent drastic decay of exhaust heat temperature at part-load. 8 MW heat is extracted at the gas turbine discharge to provide the heat demand on the platform. Cyclopentene is the selected working fluid for the ORC since it has suitable pressure values in the range of operating temperatures.

Table 3 Simulation tool verification

Parameters	Verification cases		
	Case 1	Case 2	Case 3
gas turbine load [%]	100	100	50
ORC working fluid	MDM2	Toluene	Toluene
EGT [°C]	520	520	358
\dot{m}_{GT} [kg/s]	19.2	19.2	19.2
η_{GT} [% LHV]	28.9	28.9	23.8
η_{ORC} , open literature [%]	17.1	27.8	26.7
η_{ORC} , current study [%]	17.2	27.9	27.0
relative error [%]	0.5	0.3	1.1

Design assumptions and input data to the design case are adopted from the design case in [13] except for the PPTD in the heat exchangers which is set to 15°C for a lower footprint on the platform. A parametric study is conducted to find the optimal design point based on higher power capacity and lower footprint. Afterwards, an off-design simulation is conducted on the designed combined cycle to assess the cycle performance in part-load with the presented control strategies.

3 Results and Discussion

Fig. 2 and Fig. 3 show how the power capacity of the ORC and total heat exchangers volume per MW vary with different design temperatures in the subjected design case, respectively. The design evaporation and condensing temperature are selected from the parametric study to be 200°C and 50°C, respectively. The designed cycle is shown to have the ability to provide 3.4 MW of power with 18.1% thermal efficiency at the design point. The required total volume and total effective heat transfer area of the heat exchangers are estimated to be 1100 m³ and 4200 m² for the design case, respectively.

The simulations showed near-constant thermal efficiency at off-design loads with VAN turbine control logic and cooling flow control logic. However partial admission turbine control strategy, sliding pressure strategy, and throttling control strategy experienced higher efficiency loss relative to the VAN turbine control strategy. The plant with VAN turbine as the ORC expander showed to have a 1.25 percentage point higher combined cycle efficiency at 50% gas turbine load. It will result in 2.5% less CO₂ emission at part-load operation.

4 Summary and Conclusions

An in-house design and simulation tool was developed to facilitate the design procedure and an early performance assessment of GT-ORC combined cycles. The tool offers a graphical user interface for a more convenient design experience. Background scientific principles, mathematical formulations, and the coding algorithm were explained. A sample case was designed, and the part-load performance was discussed according to the results from the simulation tool. Five control strategies for off-design power demands were

studied. VAN turbine control strategy showing higher thermal efficiency can be a potential solution for reducing carbon emission on the offshore oil and gas platforms as it reduced the CO₂ emission by 2.5% at part-load operation.

ORC part-load thermal efficiency of the subjected design case is illustrated in Fig. 4. VAN turbine control strategy outperformed other studied control strategies by higher thermal efficiency at part-load.

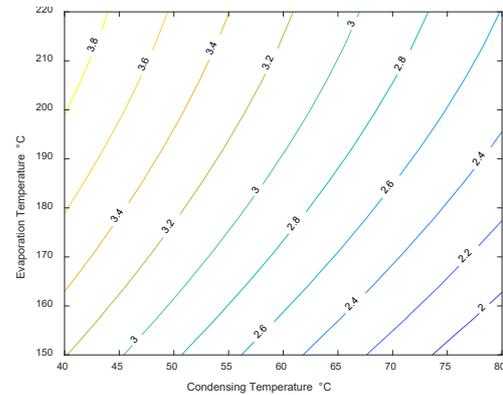


Figure 2: ORC power output design study [MW]

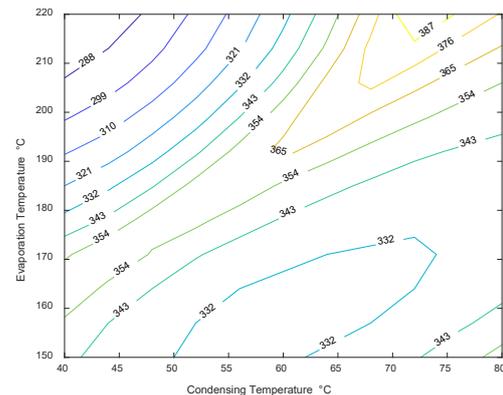


Figure 3 Heat recovery unit specific volume [lit/kW]

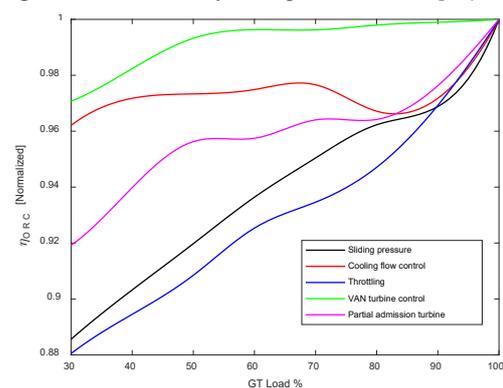


Figure 4: ORC part-load efficiency

Acknowledgement

This publication has been produced with support from the LowEmission Research Centre (www.lowemission.no), performed under the Norwegian research program PETROSENTER. The authors acknowledge the industry partners in LowEmission for their contributions and the Research Council of Norway (296207).

Nomenclature

A	heat exchanger effective area (m^2)
C	heat capacity (J/K)
CR	heat exchanger heat capacity ratio
c_f	skin friction factor
D	diameter (m)
h	enthalpy (J/kg)
k	thermal conductivity (W/mK)
\dot{m}	mass flow rate (kg/s)
n	exponent in the efficiency formula
N	rotational speed (rad/s)
Nu	Nusselt number
Ns	turbine specific speed
P	pressure (Pa)
Pr	Prandtl number
Q	volume flow rate
Re	Reynolds number
T	temperature (K)
U	overall heat transfer coefficient (W/m^2K)
\dot{W}	power (W)

Greek letters

α	convective heat transfer coefficient
β	heat transfer ratio coefficient
η	efficiency
ϵ	heat transfer effectiveness

Abbreviations

CC	combustion chamber
CND	condenser
COM	compressor
ECO	economizer
EGT	gas turbine exhaust temperature
EVA	evaporator
GEN	generator
GT	gas turbine
Hex	heat exchanger
ORC	organic Rankine cycle
PPTD	pinch point temperature difference
PR	pressure ratio
SUP	superheater
TUR	turbine

Subscripts

1	ORC pump inlet
2	economizer inlet
3	Evaporator outlet
4	Turbine outlet
amb	ambient
c	critical
cc	combined cycle
$cold$	heat exchanger cold side
cw	cooling water supply
D	based on diameter
dp	design point

hot	heat exchanger hot side
in	inlet
IOL	intermediate oil loop
max	maximum
min	minimum
out	outlet
$pump$	related to pump
s	isentropic
$turbine$	related to turbine

References

- [1] NPD, "Exploration Resource report 2020," pp. 1–77, 2020, [Online]. Available: <https://www.npd.no/en/facts/publications/reports2/resource-report/>.
- [2] "Emissions to air," *Statistics Norway*, 2021. <https://www.ssb.no/en/natur-og-miljo/forurensning-og-klima/statistikk/utslipp-til-luft>.
- [3] R. K. Bhargava, M. Bianchi, L. Branchini, A. De Pascale, and V. Orlandini, "Organic rankine cycle system for effective energy recovery in offshore applications: A parametric investigation with different power rating gas turbines," *Proc. ASME Turbo Expo*, vol. 3, 2015, doi: 10.1115/GT2015-42292.
- [4] H. Nami, I. S. Ertesvåg, R. Agromayor, L. Riboldi, and L. O. Nord, "Gas turbine exhaust gas heat recovery by organic Rankine cycles (ORC) for offshore combined heat and power applications - Energy and exergy analysis," *Energy*, vol. 165, pp. 1060–1071, 2018, doi: 10.1016/j.energy.2018.10.034.
- [5] I. Encabo Cáceres, R. Agromayor, and L. O. Nord, "Thermodynamic Optimization Of An Organic Rankine Cycle For Power Generation From A Low Temperature Geothermal Heat Source," *Proc. 58th Conf. Simul. Model. (SIMS 58) Reykjavik, Iceland, Sept. 25th – 27th, 2017*, vol. 138, pp. 251–262, 2017, doi: 10.3384/ecp17138251.
- [6] G. Manente, A. Toffolo, A. Lazzaretto, and M. Paci, "An Organic Rankine Cycle off-design model for the search of the optimal control strategy," *Energy*, vol. 58, pp. 97–106, 2013, doi: 10.1016/j.energy.2012.12.035.
- [7] M. Astolfi, "Technical options for organic rankine cycle systems," in *Organic Rankine Cycle (ORC) Power Systems: Technologies and Applications*, no. 1, 2017, pp. 67–89.
- [8] I. H. Bell, J. Wronski, S. Quoilin, and V. Lemort, "Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp," *Ind. Eng. Chem. Res.*, vol. 53, no. 6, pp. 2498–2508, 2014.

- [9] M. J. Moran, H. N. Shapiro, D. D. Boettner, and M. B. Bailey, *Fundamentals of Engineering Thermodynamics, 8th Edition*. Wiley, 2014.
- [10] B. A. L. Hagen, M. Nikolaisen, and T. Andresen, "A novel methodology for Rankine cycle analysis with generic heat exchanger models," *Appl. Therm. Eng.*, vol. 165, no. March 2019, p. 114566, 2020, doi: 10.1016/j.applthermaleng.2019.114566.
- [11] P. L. Dhar, *Thermal System Design and Simulation*. 2016.
- [12] F. Ceglia, E. Marrasso, C. Roselli, and M. Sasso, "Effect of layout and working fluid on heat transfer of polymeric shell and tube heat exchangers for small size geothermal ORC via 1-D numerical analysis," *Geothermics*, vol. 95, no. February, p. 102118, 2021, doi: 10.1016/j.geothermics.2021.102118.
- [13] M. A. Motamed and L. O. Nord, "Assessment of organic rankine cycle part-load performance as gas turbine bottoming cycle with variable area nozzle turbine technology," *Energies*, vol. 14, no. 23, 2021, doi: 10.3390/en14237916.
- [14] C. Liu and T. Gao, "Off-design performance analysis of basic ORC, ORC using zeotropic mixtures and composition-adjustable ORC under optimal control strategy," *Energy*, vol. 171, pp. 95–108, 2019, doi: 10.1016/j.energy.2018.12.195.
- [15] D. Casartelli, M. Binotti, P. Silva, E. Macchi, E. Roccaro, and T. Passera, "Power Block Off-design Control Strategies for Indirect Solar ORC Cycles," *Energy Procedia*, vol. 69, pp. 1220–1230, 2015, doi: 10.1016/j.egypro.2015.03.166.
- [16] M. A. Motamed and L. O. Nord, "IMPROVING THE OFF-DESIGN EFFICIENCY OF ORGANIC RANKINE BOTTOMING CYCLE BY VARIABLE AREA NOZZLE TURBINE TECHNOLOGY," 2021, doi: 10.14459/2021mp1632935.
- [17] N. M. Phu and N. T. M. Trinh, "Modelling and experimental validation for off-design performance of the helical heat exchanger with LMTD correction taken into account," *J. Mech. Sci. Technol.*, vol. 30, no. 7, pp. 3357–3364, 2016, doi: 10.1007/s12206-016-0645-0.
- [18] H. I. H. Saravanamuttoo, G. F. C. Rogers, and H. Cohen, *Gas turbine theory*. Pearson Education, 2001.
- [19] P. L. Meitner and A. J. Glassman, "Off-Design Performance Loss Model for Radial Turbines with Pivoting, Variable-Area Stators.," NATIONAL AERONAUTICS AND SPACE ADMINISTRATION CLEVELAND OH LEWIS RESEARCH CENTER, 1980.
- [20] D. Hu, S. Li, Y. Zheng, J. Wang, and Y. Dai, "Preliminary design and off-design performance analysis of an Organic Rankine Cycle for geothermal sources," *Energy Convers. Manag.*, vol. 96, pp. 175–187, 2015, doi: 10.1016/j.enconman.2015.02.078.
- [21] B. A. L. Hagen, R. Agromayor, and P. Nekså, "Equation-oriented methods for design optimization and performance analysis of radial inflow turbines," *Energy*, vol. 237, p. 121596, 2021, doi: 10.1016/j.energy.2021.121596.
- [22] "MATLAB version R2020b." The MathWorks Inc, 2020.
- [23] J. M. Muñoz De Escalona, D. Sánchez, R. Chacartegui, and T. Sánchez, "Part-load analysis of gas turbine & ORC combined cycles," *Appl. Therm. Eng.*, vol. 36, no. 1, pp. 63–72, 2012, doi: 10.1016/j.applthermaleng.2011.11.068.

Modelling particle degradation and intermediate dynamics in a dispersed activated sludge microcosm

Kristin T. Ravndal^{a,*}, Roald Kommedal^a

^a Department of Chemistry, Bioscience & Environmental Engineering, University of Stavanger, 4036 Stavanger, Norway
kristin.t.ravndal@uis.no

Abstract

Municipal wastewater consists of a large fraction of particulate organic matter. During biological wastewater treatment these particles undergo extracellular depolymerisation before products are taken up by bacteria ($MW < 0.6$ kDa). Particle degradation and intermediate formation dynamics is important in process analysis of wastewater treatment as the transport regime differ. This work aims to develop a model for particle degradation that includes intermediate dynamics as observed in experimental work. A model for particle degradation including intermediate dynamics, bacterial growth and endogenous respiration is proposed. Particle hydrolysis was modelled using the particle breakup model. Depolymerisation products were separated into five different size groups: colloids; high, medium and low molecular weight (HMW, MMW and LMW) polymers; and one fraction for oligomers and monomers (S_B). Depolymerisation of colloids, HMW and MMW polymers was modelled using first order kinetics. LMW polymer degradation was modelled using Michaelis-Menten kinetics, while growth was based on traditional Monod kinetics and endogenous respiration followed ASM3. The proposed model was implemented in AQUASIM for a batch reactor system, and parameter estimation by LSE fitting to experimental data on particulate starch degradation over 117 days in a dispersed biomass microcosm was performed. Validation of the model against experimental data gave a very good fit to the PBM. The intermediate dynamics seen in the experimental data was also qualitatively demonstrated by the model, with accumulation of HMW, MMW and LMW polymers in the bulk liquid. However, the accumulation of monomers and oligomers in the bulk liquid could not be reproduced in the suspended growth model proposed. Hence, a structured biomass model (biofilm) is suggested for future work.

1. Introduction

Wastewater consists of a large fraction of particulate organic matter (POM) (Levine et al., 1991; Ravndal et al., 2018). During biological wastewater treatment these particles must undergo extracellular depolymerisation before products can be taken up by bacteria ($M_w < 0.6$ kDa) (Decad and Nikaido, 1976; White et al., 2012). Organic matter (OM) proceeds through a range of colloidal and polymeric intermediates during this process. The dominant mechanisms for this degradation are hydrolytic and lytic depolymerisation, and theoretically this allows for any sub-polymeric intermediates to be formed. Hence, both particle degradation and intermediate formation are important in process analysis of wastewater treatment.

Over the years, several different approaches have been proposed for modelling of particle degradation in wastewater treatment processes (Hauduc et al., 2013; Morgenroth et al., 2002; Vavilin et al., 2008). These models include one step, parallel and sequential hydrolysis, in addition to direct growth using adsorbed substrate and different types of surface related kinetics. For surface related kinetics two different models have been proposed, these are the shrinking particle model (SPM; Sanders et al.,

2000) and the particle breakup model (PBM; Dimock and Morgenroth, 2006). Hydrolysis is dependent on available surface area in both models. However, in the SPM particles shrink gradually as they are degraded, thus available surface area decreases. While in the PBM particles break up as they are degraded, leading to an increase in available surface area, and therefore increasing substrate availability over time. Hence, in the PBM surface area to volume ratio is considered, whereas surface area is used in the SPM. An open question of the PBM is that an increase in particle porosity and increased particle colonization could also lead to increased substrate availability over time (Dimock and Morgenroth, 2006). Hence the same dynamics would be observed as when particles physically break up by hydrolysis.

In previous work by the authors an experimental study was performed to look at particle degradation in activated sludge microcosms (Ravndal and Kommedal, 2017). In this study starch was used as a model particle substrate. Starch is a good model substrate as it has a known structure, at the same time as much of the complexity expected for unidentified particles in wastewater are represented in starch granules. Starch is a natural component in

wastewater, and starch degrading bacteria are commonly found in activated sludge systems (Xia et al., 2008). In the experimental test starch particles were colonised by bacteria, leading to both increase in particle porosity and particle breakup (Ravndal and Kommedal, 2017). Polymeric, oligomeric and monomeric intermediates formed during particle degradation, however, not all intermediate sizes were formed to the same extent. Based on the data a conceptual model was proposed for chemical oxygen demand (COD) flow during starch depolymerisation. This conceptual model assumed the surface of the particle to be a hotspot for extracellular hydrolytic activity. The different intermediates were grouped based on size and included in the conceptual model.

In this paper the aim is to develop a mathematical model for particle degradation that includes intermediate dynamics as observed in the experimental work by Ravndal and Kommedal (2017). This is done by developing a model including intermediate dynamics, bacterial growth, and endogenous respiration, and fitting the model to the experimental data.

2. Methodology

The model was developed based on experimental data previously published in Ravndal and Kommedal (2017). In this experiment starch degradation was followed over 117 days in batch tests inoculated with flocculated and dispersed activated sludge biomass. During the experiment the following was monitored: Oxygen utilisation rate (OUR), particle number and size (volume and surface area), polymer concentration and size (molar mass), and oligomer and monomer type and concentrations.

2.1. Model development

The experimental data in Ravndal and Kommedal (2017) supported the PBM proposed by Dimock and Morgenroth (2006). Hence, this model was chosen for particle degradation in the proposed model (Tab. 1). The PBM is based on surface area to volume ratio (f_{av}). f_{av} was included as a state variable in the model and change in f_{av} was directly coupled to the hydrolysis process with the constant c_{av} . Colloids (C_B) and polymeric intermediates are formed during particle degradation. Theoretically all intermediate sizes of polymers can be formed during particle degradation, however, for simplicity the polymers in the model were grouped into three groups based on size ranges; high molecular weight (HMW, $S_{pol,HMW}$), medium molecular weight (MMW, $S_{pol,MMW}$) and low molecular weight (LMW, $S_{pol,LMW}$) polymers. Oligomers and monomers were grouped into one state variable, S_B , representing readily biodegradable substrate small enough to be taken up by bacteria.

To minimize model parametrisation, degradation of colloids, HMW and MMW polymers was implemented using first order biomass independent kinetics (Tab. 1). LMW polymer degradation was modelled using Michaelis-Menten kinetics (Tab. 1). Most Michaelis-Menten constants (K_m) for hydrolytic enzymes are high (Technical University of Braunschweig, 2022), hence, in activated sludge where enzyme concentrations are high, simplification to first order kinetics can normally be assumed. However, LMW polymers are a common product of depolymerisation of starch, colloids and larger polymeric substrates (Robyt, 2009), hence large substrate concentrations can also be achieved.

Table 1: Process matrix for starch degradation. Particle hydrolysis was based on the PBM (Dimock and Morgenroth, 2006), colloids, HMW and MMW polymer degradation followed first order biomass independent kinetics, LMW polymer degradation followed Michaelis-Menten kinetics, growth was based on traditional Monod kinetics, and endogenous respiration was based on ASM3 (Henze et al., 2000). Nomenclature based on Corominas et al. (2010).

Process	S_{O_2}	S_B	$S_{pol,LMW}$	$S_{pol,MMW}$	$S_{pol,HMW}$	C_B	$f_{av,X}$	X_B	X_U	X_{OHO}	Rate
Particle hydrolysis		$f_{SB,X}$	$f_{LMW,X}$	$f_{MMW,X}$	$f_{HMW,X}$	$f_{CB,X}$	c_{av}	-1			$q_X f_{av,X} X_B$
Colloid hydrolysis		$f_{SB,C}$	$f_{LMW,C}$	$f_{MMW,C}$	$f_{HMW,C}$	-1					$q_{C_B} C_B$
HMW polymer hydrolysis		$f_{SB,HMW}$	$f_{LMW,HMW}$	$f_{MMW,HMW}$	-1						$q_{S_{pol,HMW}} S_{pol,HMW}$
MMW polymer hydrolysis		$f_{SB,MMW}$	$f_{LMW,MMW}$	-1							$q_{S_{pol,MMW}} S_{pol,MMW}$
LMW polymer hydrolysis		1	-1								$q_{S_{pol,LMW}} \frac{S_{pol,LMW}}{K_{m,LMW} + S_{pol,LMW}} X_{OHO}$
Growth on S_B	$-\frac{1 - Y_{OHO}}{Y_{OHO}}$	$-\frac{1}{Y_{OHO}}$								1	$\mu_{OHO,Max} \left(\frac{S_B}{K_{S,OHO} + S_B} \right) X_{OHO}$
Endogenous respiration	$-(1 - f_{XU})$							f_{XU}		-1	$b_r X_{OHO}$

In addition, enzyme affinity increases with a decreasing polymer size. Hence, Michaelis-Menten kinetics was used for LMW polymer degradation. For particle, colloid, HMW, MMW and LMW polymers smaller intermediate sizes are formed, and fractions of each intermediate formed is included in the process matrix in Tab. 1. Traditional Monod kinetics was used for growth on S_B , and endogenous respiration from ASM3 was implemented for active biomass decay (Henze et al., 2000).

2.2. Model implementation

The model was implemented in AQUASIM (Reichert, 1994) using a simple mixed reactor compartment for the test bottle liquid phase. This was coupled by a diffusive link to a mixed reactor gas phase compartment representing the headspace in the test bottles. Parameter estimation was performed by LSE fitting the model to experimental data from the dispersed activated sludge microcosms published in (Ravndal and Kommedal, 2017). The following experimental data were used: Total particle volume (X_{vol}), HMW, MMW and LMW polymer concentrations, S_B concentration, and OUR. Total particle volume was defined as the sum of starch particle volume ($X_{B,vol}$), biomass volume ($X_{OHO,vol}$) and inert particle volume ($X_{U,vol}$) (eq. 1).

$$X_{vol} = X_{B,vol} + X_{OHO,vol} + X_{U,vol} \quad (1)$$

Particle volume was related to COD concentrations using density and theoretical oxygen demand (ThOD) of the different particle types. In addition, starch particle swelling was observed in the early phase of the experiment and had to be accounted for. A swelling constant (f_{swell}) for starch particles was estimated based on measured initial volume increase. f_{swell} was set to increase linearly from 1 at time 0 to 1.53 after 5.93 d, and then kept constant throughout the experiment. Starch particle volume ($X_{B,vol}$) can be related to starch COD concentration (X_B) using eq. 2. Where ρ_{XB} , the density of potato starch granules, is $1.54 \cdot 10^6 \text{ g m}^{-3}$, (BNID103206, Milo et al., 2010), and the ThOD of starch ($f_{v_{XB}}$) is $1.19 \text{ g COD (g } X_B)^{-1}$.

$$X_{B,vol} = \frac{X_B}{f_{v_{XB}} \cdot \rho_{XB}} f_{swell} \quad (2)$$

Volume of microbial biomass ($X_{OHO,vol}$) was estimated based on literature data for *Escherichia coli* cells. Average density (ρ_{XOHO}) of an *E.coli* cell is 1.094 g mL^{-1} (BNID106306, Milo et al., 2010). The ThOD of bacterial biomass ($f_{v_{XB}}$) is $1.42 \text{ g COD (g } X_{OHO})^{-1}$. The volume of bacterial biomass was related to biomass COD concentration (X_{OHO}) using eq. 3.

$$X_{OHO,vol} = \frac{X_{OHO}}{f_{v_{XOHO}} \cdot \rho_{XOHO}} \quad (3)$$

Inert particle volume ($X_{U,vol}$) was estimated with the same ThOD and density as for microbial biomass. ThOD of polymeric intermediates was the same as for starch particles. Monomer and oligomer concentrations were related to individual specific ThODs and summarised in one state variable, S_B .

3. Results and discussion

3.1 Parameter estimation

Estimated stoichiometric parameters are summarised in Tab. 2. The experimental data showed that the MMW polymer concentration started to increase at the same time as the system shifted from a starch particle dominated system to a HMW polymer dominated system (Ravndal and Kommedal, 2017). Hence, MMW polymers where most likely degradation products of HMW polymer depolymerisation, and f_{MMW_X} and f_{MMW_C} was estimated as zero (Tab. 2).

Starch granules consist of a mix of amylopectin (70-80 %) and amylose (20-30 %) (Dona et al., 2010). Amylopectin is the largest of the two polymers with a molecular weight of 10^4 - 10^6 kDa, it is highly branched containing 5 % α -1,6 branches and water soluble (Ball et al., 1996; Shannon et al., 2009). Amylose is a smaller polymer with molecular weight of 100-1000 kDa, it is essentially linear with less than 1 % α -1,6 branches and has variable solubility depending on branching where linear amylose is essentially insoluble in water (Ball et al., 1996; Mukerjea and Robyt, 2010). During particle degradation these polymers will be released from the starch granule. Due to the large size of amylopectin and the amylose being mostly insoluble, these polymers are expected to behave as colloids initially. Hence, f_{HMW_X} was also estimated as zero (Tab. 2). Both the LMW polymer fraction and the grouped S_B state variable contains several known degradation products from enzymatic degradation of starch (Robyt, 2009). All hydrolysis processes therefore led to fractions of these two variables (Tab. 2). The fractions increased, the closer in size the initial variable was. Because parameter estimation in Aquasim did not ensure that the sum of all fractions was one, it was chosen to manually test different numbers for the fractions and the numbers giving the best fit with the experimental data was chosen.

Kinetic parameters for the extracellular degradation processes (Tab. 3) were estimated using LSE parameter estimation in Aquasim. The estimated hydrolysis rate constants increased with decreasing substrate size. This is expected, as faster degradation is expected for smaller substrate sizes.

Table 2: Stoichiometric parameters estimated for the starch particle degradation model.

Parameter	Definition	Value [1]	Reference
$f_{CB,X}$	Fraction of colloids formed in particle hydrolysis	0.45	Estimated
$f_{HMW,X}$	Fraction of HMW polymers formed in particle hydrolysis	0	Estimated
$f_{MMW,X}$	Fraction of MMW polymers formed in particle hydrolysis	0	Estimated
$f_{LMW,X}$	Fraction of LMW polymers formed in particle hydrolysis	0.2	Estimated
$f_{SB,X}$	Fraction of monomers and oligomers (S_B) formed in particle hydrolysis	0.35	Estimated
$f_{HMW,C}$	Fraction of HMW polymers formed in colloid hydrolysis	0.15	Estimated
$f_{MMW,C}$	Fraction of MMW polymers formed in colloid hydrolysis	0	Estimated
$f_{LMW,C}$	Fraction of LMW polymers formed in colloid hydrolysis	0.45	Estimated
$f_{SB,C}$	Fraction of monomers and oligomers (S_B) formed in colloid hydrolysis	0.4	Estimated
$f_{MMW,HMW}$	Fraction of MMW polymers formed in hydrolysis of HMW polymers	0.5	Estimated
$f_{LMW,HMW}$	Fraction of LMW polymers formed in hydrolysis of HMW polymers	0.25	Estimated
$f_{SB,HMW}$	Fraction of monomers and oligomers (S_B) formed in hydrolysis of HMW polymers	0.25	Estimated
$f_{LMW,MMW}$	Fraction of LMW polymers formed in hydrolysis of MMW polymers	0.75	Estimated
$f_{SB,MMW}$	Fraction of monomers and oligomers (S_B) formed in hydrolysis of MMW polymers	0.25	Estimated
Y_{OHO}	Aerobic yield of ordinary heterotrophic organisms	0.5	(Sykes, 1975)
f_{XU}	Production of inert (X_U) in endogenous decay	0.2	(Henze et al., 2000)

Table 3: Kinetic parameters estimated for the starch particle degradation model.

Parameter	Definition	Value	Reference
q_x	Modified hydrolysis constant for particles in PBM	$1.25 \cdot 10^{-9} \text{ m d}^{-1}$	Estimated
c_{av}	PBM constant relating particle breakup to hydrolysis rate	$1.54 \cdot 10^5 \text{ m}^2 \text{ g}^{-1}$	Estimated
q_{CB}	Hydrolysis rate constant for colloid degradation	0.05 d^{-1}	Estimated
$q_{Spot,HMW}$	Hydrolysis rate constant for degradation of HMW polymers	0.08 d^{-1}	Estimated
$q_{Spot,MMW}$	Hydrolysis rate constant for degradation of MMW polymers	0.09 d^{-1}	Estimated
$q_{Spot,LMW}$	Hydrolysis rate constant for degradation of LMW polymers	0.3 d^{-1}	Estimated
$K_{m,LMW}$	Michaelis-Menten constant for degradation of LMW polymers	93 g COD m^{-3}	Estimated
$\mu_{OHO,Max}$	Maximum growth rate for heterotrophic organisms	8 d^{-1}	Estimated
$K_{S,OHO}$	Saturation constant for substrate S_B	4 g COD m^{-3}	Estimated
b_h	Aerobic endogenous respiration rate of X_{OHO}	0.15 d^{-1}	(Henze et al., 2000)

The proposed model contains many different stoichiometric and kinetic parameters (Tab. 2 and 3). Some of these were estimated manually, some estimated using parameter estimation in Aquasim and some are based on literature data. To test for parameter identifiability a sensitivity analysis could be done. This analysis is not included here, but it is suggested to include this in future work on an improved version of the model.

3.2 Are the model able to reproduce the experimental data?

The model output is compared to experimental data for OUR, particle volume and concentrations of different intermediates in Fig. 1 and Fig. 2. In the experiment, the OUR increased fast between day 2 and 4 before it stabilised at a level of $1.8 \text{ mg L}^{-1} \text{ h}^{-1}$ until day 36, before a steady decrease occurred (Fig. 1). The model was not able to reproduce the initial increase and the stable phase, however, the decrease after 36 days was reproduced by the model.

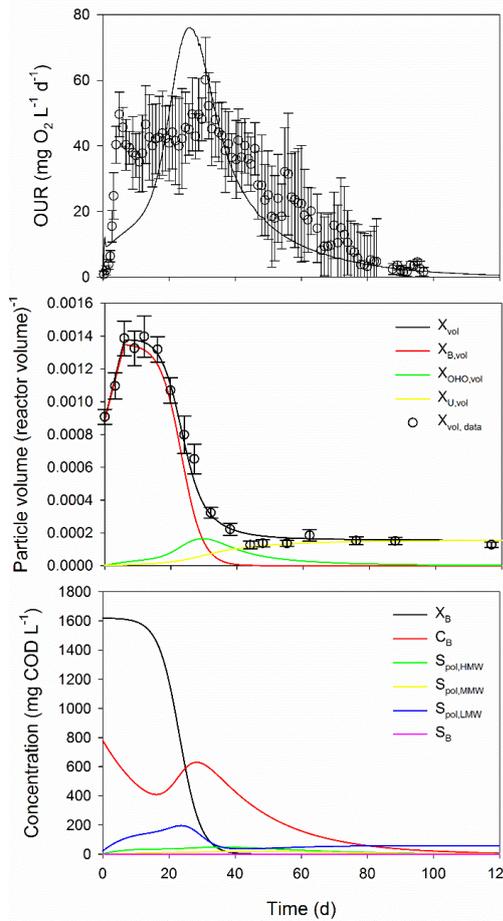


Figure 1: Experimental data (point \pm standard error bars) and modelled data (lines) for OUR, particle volume and substrate concentrations in dispersed biomass tests.

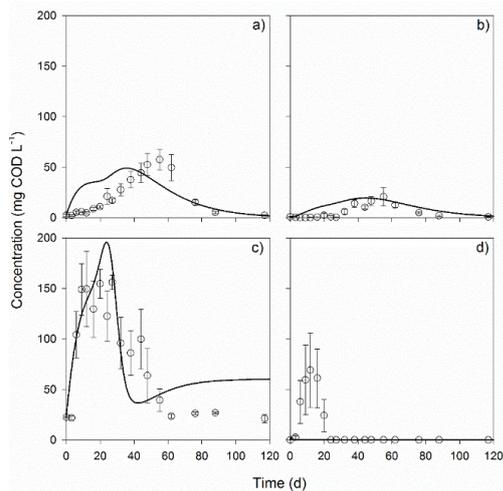


Figure 2: a) HMW polymers, b) MMW polymers, c) LMW polymers and d) monomers and oligomers (S_B). Points (\pm standard error bars) show experimental data for dispersed biomass tests, lines show modelled data.

Total particle volume fitted the PBM model very well, with growth and decay modelled after ASM3 (Fig. 1). This supported the conclusion in Ravndal and Kommedal (2017) of the data supporting the PBM proposed by Dimock and Morgenroth (2006). LMW, MMW and HMW polymers were produced and accumulated in the bulk liquid in the experiment (Ravndal and Kommedal, 2017). In the model the initial LMW peak had a good fit to the experimental data, while a faster accumulation of MMW and HMW polymers was seen compared to the experimental dataset (Fig. 2).

Colloids was included as a state variable in the model. However, we did not have access to detection methods covering the colloidal size range in the experiment. This made it challenging to identify all parameters in the model. As HMW polymers are in the size range immediately below colloids, a better fit of HMW polymer production would be expected if the experimental dataset had included colloids concentrations.

S_B accumulated in the bulk liquid in the experiment, however, this data was not reproduced by the model (Fig. 2). Apart from a tiny peak reaching a maximum of 3 mg L^{-1} of S_B after 0.37 d, no accumulation of S_B was seen in the modelled data. Hence, after the start up, the uptake of S_B by the biomass in the model was faster than the hydrolysis processes.

3.3 Model limitations and suggestions for improvements

The lack of S_B accumulating in the bulk liquid in the modelled data compared to the experimental data (Fig. 2) showed that the model did not account for diffusion limitations in the system. The model was a simple suspended growth model, however, the experiment showed that in reality the biomass colonized the surface of the starch particles (Ravndal and Kommedal, 2017). Hence, a small biofilm was formed by the colonizing biomass leading to diffusion limitations for any intermediate in the bulk liquid. To allow for S_B accumulation due to diffusion limitations, a heterogeneous biomass structure must be considered.

A model based on activated sludge flocs as small bio-aggregates with diffusion limitations and sorption dynamics, would greatly improve modelling of intermediate dynamics compared to the simplified model presented in this work. Diffusion limitations and sorption dynamics would allow for accumulation of all intermediates, including monomers and oligomers, in the bulk liquid. Hence, the initial increase of monomers and oligomers as seen in the experimental data, would be present in the modelled data. The delayed accumulation of MMW and HMW polymeric intermediates in the experimental data compared to the simplified model output, would also be recreated in a bio-aggregate model. However, including these factors in the

model will increase model complexity, thus, complicate model identification and validation. Model complexity must be balanced to the need for detailed understanding of degradation dynamics and acceptable uncertainty.

To validate a bio-aggregate model, data on bio-aggregate volume, depth and surface are needed. In the experimental work used to validate the model in this paper, this data is lacking as the particle size measurements is dominated by the starch particle fraction. To further improve modelling of intermediate dynamics during particle degradation it is suggested to combine this type of experimental work with work on a bio-aggregate model.

4. Conclusions

A mathematical model for particle degradation, including intermediate dynamics, was proposed in this work.

- The PBM fitted experimental data for particulate substrate well.
- Accumulation of LMW, MMW and HMW polymeric fractions in the bulk liquid was qualitatively replicated by the model, with the initial peak of LMW polymers giving the best fit to the experimental data.
- The model was unable to replicate data showing oligomers and monomers to accumulate in the system. It is suggested to include a bio-aggregate model in further work to account for diffusion limitations in the system due to the biomass growing as a colonizing biofilm on the starch particles.

Acknowledgment

Internal funding at the University of Stavanger, Department of Chemistry, Bioscience & Environmental Engineering, financed this work.

References

Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., Colonna, P. and Preiss, J. (1996) 'From glycogen to amylopectin: A model for the biogenesis of the plant starch granule', *Cell*, Cell Press, 86(3), pp. 349-352. doi: 10.1016/s0092-8674(00)80107-5

Corominas, L., Rieger, L., Takács, I., Ekama, G., Hauduc, H., Vanrolleghem, P.A., Oehmen, A., Gernaey, K.V., van Loosdrecht, M.C.M. and Comeau, Y. (2010) 'New framework for standardized notation in wastewater treatment modelling', *Water Science and Technology*, IWA Publishing, 61(4), pp. 841-857. doi: 10.2166/wst.2010.912

Decad, G.M. and Nikaido, H. (1976) 'Outer membrane of gram-negative bacteria. XII. Molecular-sieving function of cell wall', *Journal of Bacteriology*, ASM Journals, 128(1), pp. 325-336. doi: 10.1128/jb.128.1.325-336.1976.

Dimock, R. and Morgenroth, E. (2006) 'The influence of particle size on microbial hydrolysis of protein particles in activated sludge', *Water Research*, Elsevier, 40(10), pp. 2064-2074. doi: 10.1016/j.watres.2006.03.011

Dona, A.C., Pages, G., Gilbert, R.G. and Kuchel, P.W. (2010) 'Digestion of starch: In vivo and in vitro kinetic models used to characterise oligosaccharide or glucose release', *Carbohydrate Polymers*, Elsevier, 80(3), pp. 599-617. doi: 10.1016/j.carbpol.2010.01.002

Hauduc, H., Rieger, L., Oehmen, A., van Loosdrecht, M.C.M., Comeau, Y., Heduit, A., Vanrolleghem, P.A. and Gillot, S. (2013) 'Critical review of activated sludge modeling: State of process knowledge, modeling concepts, and limitations', *Biotechnology and Bioengineering*, Wiley Periodicals, 110(1), pp. 24-46. doi: 10.1002/bit.24624

Henze, M., Gujer, W., Mino, T. and van Loosdrecht, M. (2000) *Activated sludge models ASM1, ASM2, ASM2d and ASM3*. IWA Publishing.

Levine, A.D., Tchobanoglous, G. and Asano, T. (1991) 'Size distributions of particulate contaminants in wastewater and their impact on treatability', *Water Research*, Elsevier, 25(8), pp. 911-922. doi: 10.1016/0043-1354(91)90138-g

Milo, R., Jorgensen, P., Moran, U., Weber, G. and Springer, M. (2010) 'BioNumbers—the database of key numbers in molecular and cell biology', *Nucleic Acids Research*, Oxford Academic, 38(suppl_1), pp. D750-D753. doi: 10.1093/nar/gkp889

Morgenroth, E., Kommedal, R. and Harremoes, P. (2002) 'Processes and modeling of hydrolysis of particulate organic matter in aerobic wastewater treatment - a review', *Water Science and Technology*, IWA Publishing, 45(6), pp. 25-40. doi: 10.2166/wst.2002.0091.

Mukerjea, R. and Robyt, J.F. (2010) 'Isolation, structure, and characterization of the putative soluble amyloses from potato, wheat, and rice starches', *Carbohydrate Research*, Elsevier, 345(3), pp. 449-451. doi: https://doi.org/10.1016/j.carres.2009.11.021

Ravndal, K.T. and Kommedal, R. (2017) 'Starch degradation and intermediate dynamics in flocculated and dispersed microcosms', *Water Science and Technology*, IWA Publishing, 76(11), pp. 2928-2940. doi: 10.2166/wst.2017.467

Ravndal, K.T., Opsahl, E., Bagi, A. and Kommedal, R. (2018) 'Wastewater characterisation by combining size fractionation, chemical composition and biodegradability', *Water Research*, Elsevier, 131, pp. 151-160. doi: https://doi.org/10.1016/j.watres.2017.12.034

Reichert, P. (1994) 'AQUASIM - a tool for simulation and data-analysis of aquatic systems', *Water Science and Technology*, IWA Publishing, 30(2), pp. 21-30. doi: 10.2166/wst.1994.0025

Robyt, J.F. (2009) *Starch: Chemistry and technology*. BeMiller, J. and Whistler, R. (eds), pp. 237-292, Academic Press.

Sanders, W.T.M., Geerink, M., Zeeman, G. and Lettinga, G. (2000) 'Anaerobic hydrolysis kinetics of particulate substrates', *Water Science and Technology*, IWA Publishing, 41(3), pp. 17-24. doi: 10.2166/wst.2000.0051

Shannon, J.C., Garwood, D.L. and Boyer, C.D. (2009) *Starch: Chemistry and Technology*. BeMiller, J. and Whistler, R. (eds), pp. 23-82, Academic Press.

Sykes, R.M. (1975) 'Theoretical Heterotrophic Yields', *Journal (Water Pollution Control Federation)*, Wiley, 47(3), pp. 591-600. <http://www.jstor.org/stable/25038666>.

Technical University of Braunschweig. (2022) BRENDA: The comprehensive enzyme information system, Available at: <http://www.brenda-enzymes.org/index.php>. (Accessed: 30.05.22)

Vavilin, V.A., Fernandez, B., Palatsi, J. and Flotats, X. (2008) 'Hydrolysis kinetics in anaerobic degradation of particulate organic material: An overview', *Waste Management*, Elsevier, 28(6), pp. 939-951. doi: 10.1016/j.wasman.2007.03.028

White, D., Drummond, J. and Fuqua, C. (2012) *The physiology and biochemistry of prokaryotes*. 4th edn. Oxford university press.

Xia, Y., Kong, Y.H. and Nielsen, P.H. (2008) 'In situ detection of starch-hydrolyzing microorganisms in activated sludge', *Fems Microbiology Ecology*, Oxford Academic, 66(2), pp. 462-471. doi: 10.1111/j.1574-6941.2008.00559.x

A Ray Tracer for optimizing solar concentrating systems: The case of discretized Compound Parabolic Concentrator

Casiana Blasius Lwiwa ^{a,*}, Ole Jørgen Nydal ^a

^a Norwegian University of Science and Technology, 7491, Trondheim, Norway
casiana.b.lwiwa@ntnu.no

Abstract

Concentrating solar systems use reflective surfaces to concentrate sunlight onto a small area, where it is absorbed and converted to heat. Many classes of concentrating collectors such as Compound Parabolic Concentrators (CPCs), parabolic dish and parabolic trough are available, each with different concentrating ratio and maximum absorber temperature, depending on the type of applications. A simplified 3D Compound Parabolic Concentrator (CPC) with 2 rings and 4 sectors has been designed. The designed CPC is cost effective as it requires only 8 mirrors to cover the reflector surface. It does not require sun tracking, but have capability to accept incoming solar radiation over a relatively wide range of angles. For further capturing of the solar radiations, tilting of the CPC during a day can be made a few times. This study aims to model the discretized 3D CPC using the ray tracing, to optimize the CPC for achieving optimal interceptions on a 0.2 m diameter cylindrical absorber, placed inside the CPC. The ray tracing methodology is presented together with the results of the interceptions on the cylindrical absorber using the discretized CPC. Results show the effect of tilting the discretized CPC is not very strong as the interception values are slightly reduced and the curves a little bit not symmetric around the normal sun angles.

1. Introduction

The concentrating solar thermal systems use mirrors or lenses to concentrate and collect solar radiations onto an absorber in order to produce high temperatures needed for various applications. Compound Parabolic Concentrators (CPC) falls under the class of non imaging concentrators which can be promising options for thermal solar energy applications. The important features of the CPC is that it does not require tracking of the sun and the capability of collecting solar radiations over a wide range of angles [1]. The general principles of a CPC are reported in [1, 2]. CPC is made up of two parabolic mirror segments with closely located focal points and their axes inclined to each other, such that rays incident within the acceptance angle of the CPC are collected at the lower exit opening. Fig. 1 gives the schematic diagram of the 2D CPC.

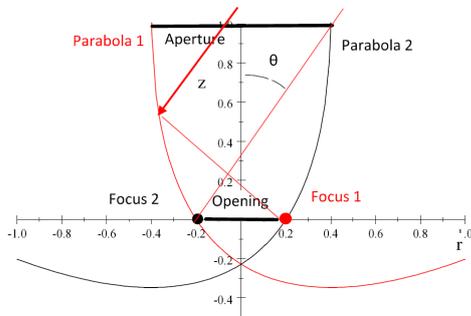


Figure 1: The schematic diagram of the 2D CPC

The 2D CPC can have an ideal performance where all

incoming rays within the acceptance angle will pass through the lower opening. The concentration ratio for a 2D CPC is given by [1, 3]:

$$C = \frac{A_{ap}}{A_{ab}} = \frac{1}{\sin\theta} \quad (1)$$

where; C is the concentration ratio, A_{ap} is the aperture area, A_{ab} is the absorber area and θ is the acceptance half angle.

The CPC equation can be obtained from the rotation and shift of parabolas. Let a coordinate system in the focus point of a parabola be (z, r) . (Z, R) is the rotated coordinate system, where the rotation is around the focal point and where the parabola has been shifted distance a . Then, the parabola with focal point in the origo is given by:

$$z = \frac{r^2}{4f} - f \quad (2)$$

To represent both parabolas in one coordinate system, denoted (R, Z) , a rotation and translation is made. The rotation angle is ϕ and the displacement along the R axis is a as shown in Eqs. 3-8.

Rotation

$$r = R\cos(\phi) + Z\sin(\phi) \quad (3)$$

$$z = -R\sin(\phi) + Z\cos(\phi) \quad (4)$$

For compact notation, we define:

$$s = \sin(\phi) \quad (5)$$

$$c = \cos(\phi) \quad (6)$$

$$r = Rc + Zs \quad (7)$$

$$z = -Rs + Zc \quad (8)$$

The inverted transformations are:

$$R = cr - sz \quad (9)$$

$$Z = cz + rs \quad (10)$$

Inserting into the equation for the parabola in Eq. 2, for the rotation gives:

$$-Rs + Zs = \frac{(Rs + Zs)^2}{4f} - f \quad (11)$$

Solving for Z, we obtain two solutions:

$$Z_1 = \frac{1}{s^2} \left(2f \sqrt{\frac{1}{f}(f - Rs)(c^2 + s^2)} + 2cf - Rcs \right) \quad (12)$$

$$Z_2 = -\frac{1}{s^2} \left(2f \sqrt{\frac{1}{f}(f - Rs)(c^2 + s^2)} - 2cf + Rcs \right) \quad (13)$$

The solutions for Z_1 and Z_2 are shown on Fig. 2. Inspection shows that of the two solutions, Z_2 is the relevant for the CPC (black curve in the figure).

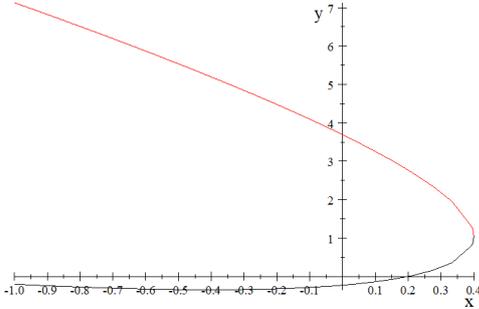


Figure 2: The schematic diagram of the CPC from rotating the parabolas

CPC can be generated by rotating positive and shifting left which gives the right side of the CPC and mirror on the other side gives the left side of the CPC. For a 3D CPC, we consider the right side and revolve around z axis. Considering only the rotation, where the origo is the focal point for both, we obtain Fig. 3.

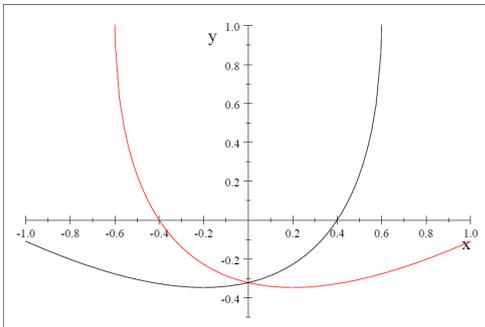


Figure 3: The schematic diagram of the CPC from rotating the parabolas; the origo is the focal point for both parabolas

The parabola should be shifted such that the focal point is at the rim of the other. The black will then be shifted to the left such that the focal point (now at the origo) reaches

the red curve. By symmetry this should be the solution to $Z = 0$. We call this shift b . In the figure it is $b = 0.4$. But as we keep the origo, we shift the other way with a similar amount. So, we can only shift half the distance given by the $Z = 0$ solution. The real shift is what enters the CPC equation, with the variable a . So, b can be obtained from $Z = 0$ solution and then $a = \frac{b}{2}$.

The solution for the only rotated, $Z_2 = 0$ is given by:

$$0 = -\frac{1}{s^2} \left(2f \sqrt{\frac{1}{f}(f - Rs)} - 2cf + Rcs \right) \quad (14)$$

The solutions for the shift b becomes:

$$b = -\frac{2f}{c^2 s} \left(\sqrt{-c^2 + 1} - c^2 + 1 \right) \quad (15)$$

$$b = \frac{2f}{c^2 s} \left(\sqrt{-c^2 + 1} + c^2 - 1 \right) \quad (16)$$

As c is less than one, the first solution is the one on the negative side. We are looking for the other from Eq. 16.

Simplifying we get:

$$f = \frac{b}{2}(1 + s) = a(1 + s) \quad (17)$$

a is now half the lower opening

Thus, the CPC equation becomes:

$$Z = -\frac{1}{s^2} \left(\sqrt{A + BR} + C - csR \right) \quad (18)$$

where:

$$A = a^2(s + 1)$$

$$B = -as(s - 1)$$

$$C = -ac(2 - s)$$

The amount of solar radiations intercepted by the CPC is directly related to the width of the entrance aperture and the acceptance angle [4], where the aperture allows the concentration of the solar radiation at the absorber. The direction and angle of the solar irradiation can be obtained by using the geometrical coordinates and time which works as a seasonally altered clock.

The aperture size of the absorber affects ray interceptions with the absorber such that, the larger the aperture size, the higher the interceptions on the absorber. Interception ratio is the ratio of the sun rays hitting the reflector and reflecting on the absorber. It shows the performance of a solar concentrator to enable the determination of the optical losses. Interception ratio depends on the geometry of the optical systems and the sun angles, the solar tracking errors and varies for smooth and tiled surfaces [5, 6].

Fig. 4 is a setup of the discretized CPC, a cylindrical absorber and the vertical sun. The designed CPC is strongly simplified but cost effective since only 8 mirrors are needed to cover the reflector surface.

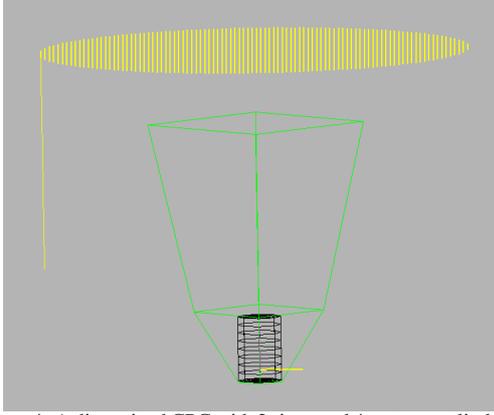


Figure 4: A discretized CPC with 2 rings and 4 sectors, cylindrical absorber and vertical sun

2. Ray tracing methodology

Ray tracing technique is used for the optical analysis of the solar concentrating systems. It encompasses the tracing of the individual paths of a large number of incident solar rays passing through the optical systems to determine the distribution pattern on the surface of interest. For analysis of thermal systems, ray tracing can give the distribution and intensities at the absorber [5].

Various software are available for ray tracing [7]. An in-house ray tracer algorithm (Tracelt) is implemented in C++ with Qt and the Open GL [5]. The Qt library provides the basis for the user interface, and the OpenGL library for the 3D graphical visualization. A model view of the data is included, where panels can be selected for translation, rotation or deletion. A screen capture of the program is shown in Fig. 5.

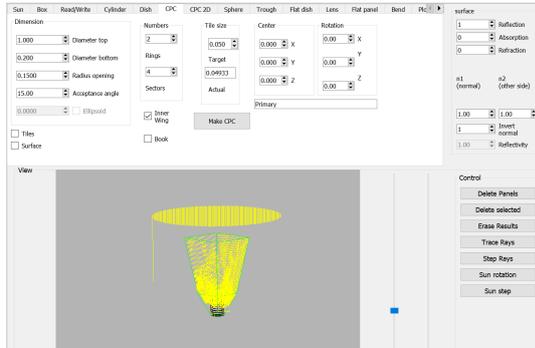


Figure 5: A screenshot of the Tracelt program: The discretized CPC, cylindrical absorber and the vertical sun

Panels are the basic elements in Tracelt, and they can be absorbers, reflectors or refractors. An array of sun rays (origin and direction) of individual rays are made. For each ray, rays intersection with the panels are computed. Ray tracing is performed after the reflections and absorptions system are made. For the nearest intersection, if the reflection side of the panel, the reflected rays (origin and direction) are made and repeated (recursive function), otherwise, the rays are terminated (absorber or backside).

In this study, the reflectivity is assumed to be 100%, mirror imperfections are ignored. A high density of sun rays leads to longer computation times (minutes). In the current case, a sun grid size of 2 cm can be sufficient to give low simulation times and acceptable number of rays for the visualization. The rays can be visualized as lines, points or as colors on the surfaces. Computational loops with

varying sun angles or geometrical parameters have been included.

After the ray tracing is completed, the data generated by the Tracelt are saved in a text file and can be re initiated afterwards. The data can be post processed with other software. In our case, MATLAB was used for better graphical presentations as compared to what is implemented in Tracelt program. Fig. 6 shows the flow chart for the description of the ray tracer algorithm.

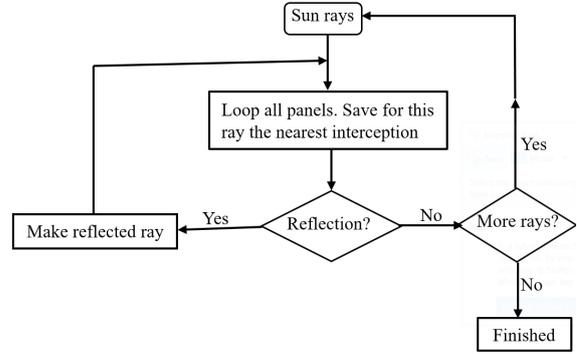


Figure 6: Flow chart of the ray tracer algorithm for simulating the performance of solar concentrating systems

3. Ray tracing sequence

The basic equation encountered in a ray tracer is to determine the intersections between a line and a surface. Subsequently, the reflected ray from the incoming ray and the surface normal vector at the intersection point are computed. Consider a ray, with point of intersection at $P(x, y, z)$ such that:

$$\vec{P} = \vec{S} + u\vec{d} \quad (19)$$

$u\vec{d}$ is a line that intersects \vec{P} and \vec{d} is a unit vector.

A surface, can be cylinder, sphere, flat plate etc. and can be described by an algebraic relation. For a CPC, this relation is given in Eq. 18.

Assuming the starting point of the sun to be $S(s_x, s_y, s_z)$. The components of parametric equation for each ray can be described by:

$$\begin{cases} x = s_x + u \cdot d_x \\ y = s_y + u \cdot d_y \\ z = s_z + u \cdot d_z \end{cases} \quad (20)$$

where (s_x, s_y, s_z) are initial points and (d_x, d_y, d_z) are the direction vector.

Inserting the P components of Eq. 20 into the shape of the surface $f(x, y, z) = 0$ gives an equation for u and thus the intersection point, $f(u)=0$.

To formulate reflections in a ray tracing technique, the normal vector (\vec{n}), calculated at the point of intersection (P), has the reflected ray and direction (\vec{d}) given by:

$$\vec{r} = \vec{d} - 2(\vec{d} \cdot \vec{n})\vec{n} \quad (21)$$

For a 2D CPC, the solution for u can be solved analytically, but for a 3D CPC, the solution for u can be obtained using numerical methods such as bisection scheme [8].

4. Examples of ray tracing for analysis of solar concentrators

Studies with Tracelt include references [4, 5, 9, 10] as briefly described below:

4.1. Optimization of the size of surface tile on a CPC

Tracelt was used to perform a sensitivity analysis on the dimensions of mirror tiles on the CPC surface. The results show that uniform tile sizes up to 15 cm gives quite similar interception ratios (less than 10% changes) compared with the smooth surface. Larger and non-uniform tiles discretized in terms of rings and sectors can give similar results [4].

4.2. Simulation of a light guide efficiency

The tracelt has been applied for evaluation of a light guide to provide the energy transfer from a solar concentrator to an absorber. Using a tracelt ray tracer, results shows energy losses through a light guide can be associated with the number of internal reflections in the guide, and the number of back scattered rays through the tube inlet [9].

4.3. Double reflector system

The optical performance of the double reflector system was investigated by means of a ray tracer. Results shows that increasing the absorber dish from 0.1 to 0.2 m makes the system more tolerant for tracking accuracy and for low solar angles. A large absorber dish also reduces the risk of superheating the top plate which has conduction based heat transfer to a salt storage [5].

4.4. Evaluation of the SK14 solar concentrator as a solar fryer

SK14 is a solar dish reflector which is made for cooking in a pot positioned at the focal point. [10] used Tracelt ray tracing to evaluate the SK14 solar concentrator to fry Ethiopian injera bread. Ray tracing reveals that the reflected rays spread more evenly on the pan, but also gives higher losses compared with an ideally smooth reflector.

The novelty of this study is to analyze the efficiency of this simplified system where the intention is to accumulate heat for cooking on the cylindrical absorber.

5. Simulation parameters

A simplified CPC (15 degrees acceptance angle) with 2 rings and 4 sectors has been applied in the simulations. The surface of the reflector is covered with 8 mirrors. The CPC has aperture diameter of 1.0 m and exit diameter opening of 0.15 m. The cylindrical absorber has 20 cm diameter and 30 cm length.

6. Results

The sun angle-interception values were compared for the fixed and tilted CPC at 5, 10 and 15 degrees. Fig. 7 shows the effects of tilting the discretized CPC as compared to the fixed CPC. In all cases, the absorber is fixed. The sun angles are chosen to exceed the maximum acceptance angle (15 degrees) to see the effect of using the discretized CPC beyond the design regions. The effect of tilting the discretized CPC does not affect much the interception values, as they are slightly reduced and the curves a little bit not symmetric around the normal sun angles.

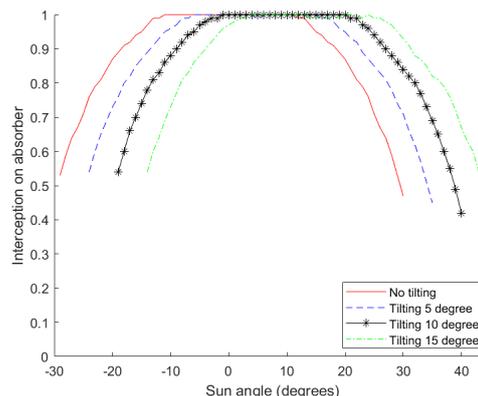


Figure 7: Effects of tilting the discretized 3D CPC to the interceptions on the absorber

7. Summary and Discussions

The ray tracing methodology for optimizing the solar concentrating systems have been presented, together with the results of the interceptions on the cylindrical absorber using the discretized CPC. Results shows when the CPC is tilted while the absorber is kept stationary, the interception values are slightly reduced and the curves somewhat not symmetric around the normal sun angles.

The CPC can be built with 8 mirrors and still give high interceptions on the absorber.

Acknowledgment

The authors extend their gratitude to the Norwegian University of Science and Technology, NTNU Energy strategic Area for financial support to the study.

References

- [1] A. Rabl, "Comparison of solar concentrators," *Solar energy*, vol. 18, no. 2, pp. 93–111, 1976.
- [2] R. Winston and H. Hinterberger, "Principles of cylindrical concentrators for solar energy," *Solar Energy*, vol. 17, no. 4, pp. 255–258, 1975.
- [3] A. Rabl, "Optical and thermal properties of compound parabolic concentrators," *Solar energy*, vol. 18, no. 6, pp. 497–511, 1976.
- [4] C. Lwiwa and O. Nydal, "3d ray tracing for optimizing the size of surface tiles on a compound parabolic concentrator," Southern African Sustainable Energy Conference, South Africa, 17-19 November, 2021., 2022.
- [5] O. J. Nydal, "Ray tracing for optimization of a double reflector system for direct illumination of a heat storage," *Energy Procedia*, vol. 57, pp. 2211–2220, 2014.
- [6] K. Nyeinga, D. Okello, and O. J. Nydal, "A ray tracer model for analysis of solar concentrating systems," *Journal of Energy in Southern Africa*, vol. 30, no. 1, pp. 8–20, 2019.
- [7] D. Jafrancesco, J. P. Cardoso, A. Mutuberria, E. Leonardi, I. Les, P. Sansoni, F. Francini, and D. Fontani, "Optical simulation of a central receiver system: Comparison of different software tools," *Renewable and Sustainable Energy Reviews*, vol. 94, pp. 792–803, 2018.
- [8] G.-L. Dai, X.-L. Xia, C. Sun, and H.-C. Zhang, "Numerical investigation of the solar concentrating characteristics of 3d cpc and cpc-dc," *Solar Energy*, vol. 85, no. 11, pp. 2833–2842, 2011.
- [9] O. Nydal, "Ray tracing for simulation of a light guide efficiency," 3rd Southern African Solar Energy Conference, South Africa, 11-13 May, 2015., 2015.
- [10] O. J. Nydal and A. Tesfay, "Ray tracing for evaluation of the sk14 solar concentrator as a solar fryer;" in *Conference Proceedings, ISES Solar World Congress, Daegu, Korea, 8–12 November 2015*, 2015.

Developing Voltage Droop/Compensation Controller for a Hydro Power Controller in Modelica

Luxshan Manoranjan^a, Dietmar Winkler^{b,*}

^a University of South-Eastern Norway, Porsgrunn, Norway, luxshan@hotmail.no

^b Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway, Porsgrunn, Norway
dietmar.winkler@usn.no

Abstract

With the introduction of unregulated renewable energy such as wind, solar and tidal power, the operation of the electrical grid has become more and more challenging. The more dynamic production pattern requires more advanced control algorithms in order to maintain an acceptable voltage quality which is within the limits given by the electrical network regulators. Better tooling and improved simulation of different operation scenarios is required.

This paper presents the development of voltage droop/compensation controller as used in a typical hydro power controller. The controllers has been implemented using the Modelica language [4] and are according to the Norwegian Energy Regulatory Authority (NERA). Having the controller available in Modelica makes it possible to integrate them with hydro power system models build with the use of OpenHPL [7]. The behaviour of the controller have been tested against a verified generator model of the OpenIPSL [1].

1. Introduction

The electrical power demand is still increasing, and it leads to pushing the society to find a renewable source to produce electricity. Therefore the development of existing and new hydropower stations is still increasing. The development of hydropower plants focuses not only on larger hydropower plants but also on small-scale hydropower plants in order to utilise as much resource from nature.

A hydropower plant consists of several components such as a valve, turbine, generator, turbine-regulator, excitation system, switch gear, etc. The generator converts mechanical energy into electrical energy. A generator needs an excitation system to provide field current to the field winding in order to induce the voltage in the generator terminals. An excitation system contains mainly an exciter that produces field current and an excitation control system that consists of an Automatic Voltage Regulator (AVR), controllers, and protective limiters.

In order to keep the voltage quality within the limits of what the electrical network regulators allow, the government has developed requirements to the hydropower stations to adequate operation of power plants.

The legal requirement for excitation systems in Norway is given in the National Guide for Functional Requirements in the Power System, NVF 2020, [6]. It is a guideline for the power system administrators to build, maintain or operate their system in order to fulfil the functional requirements set by the Norwegian Energy Regulatory Authority (NERA). The NVF 2020 contains requirements for the Norwegian grids, production power plants, High Voltage Direct Current (HVDC), consumers, protections, and measuring equipment. The production power plant part in the NVF 2020 describes requirements for

synchronous power plants and power parks. Several requirements are described under synchronous power plants, such as turbine regulation, excitation system, maximum reactive power, etc. Where the excitation system section describes the requirement for excitation system response time, VAR/PF control or regulation, voltage droop/compensation control, limiters, Power System Stabiliser, and reset functionality. This paper focuses on modelling and simulation of the voltage droop/compensation control function.

2. Theory

The voltage droop/compensation controller is an additional function that is required by NERA. The purpose of this controller is to maintain constant generator terminal voltage concerning additional measurement signals from the generator, such as reactive, active current, and frequency. The voltage droop/compensation controller influences the voltage reference in the AVR to obtain the desired terminal voltage output. This controller consists of four control functions, reactive current droop, reactive current compensation, active current compensation, and frequency droop, the characteristics of each control function are described below.

2.1. Reactive Current Droop Control

Reactive current droop control is one of the functionality implemented to stabilise the distribution of reactive load between two or more generators on the same busbar. Alternatively, to reduce the reactive load changes at a small generator that is connected to an unstable grid with high voltage variations. This control function has a negative droop that reduces the terminal voltage as a function of increasing reactive current (see Fig. 1, which gives the same effect as an inductor connected in series with the generator.

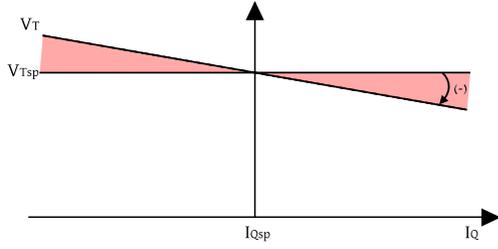


Figure 1: Characteristic curve of reactive current droop control function. V_T : Generator terminal voltage, I_Q : Generator terminal reactive current output, V_{Tsp} : Generator terminal voltage setpoint, I_{Qsp} : Reactive current setpoint.

2.2. Reactive Current Compensation Control

This control function is used to compensate for the voltage drop due to reactive components as transformers or transmission lines in the grid. The reactive current compensation control function is quite the opposite of the reactive current droop control function, as depicted in Fig. 2. This control function has a positive droop, meaning the terminal voltage increases for increasing reactive current.

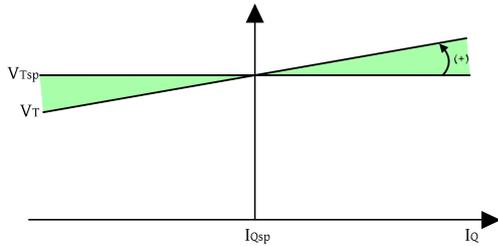


Figure 2: Characteristic curve of reactive current compensation control function. V_T : Generator terminal voltage, I_Q : Generator terminal reactive current output, V_{Tsp} : Generator terminal voltage setpoint, I_{Qsp} : Reactive current setpoint .

2.3. Active Current Compensation Control

The active current compensation control function is used to compensate for voltage drop over transformers or transmission lines due to active power consumption. This control function increases the terminal voltage as a function of increasing active current, see Fig. 3.

2.4. Frequency Droop Control

This control function can be used to help the turbine regulator to stabilise the frequency at the local grid. The frequency droop control function increases or decreases the terminal voltage as a function of increasing or decreasing frequency within a limited span (see Fig. 4). As a consequence, active power consumption in the resistive load increases if the generator runs at a higher speed. This control is only active if the circuit breaker is closed.

3. Modelling of Controller

The voltage droop/compensation controller varies the generator terminal voltage considering the changes in active and reactive current and frequency. The controller's output interacts with the summing point to modify the voltage reference signal V_{REF} , and consequently, the

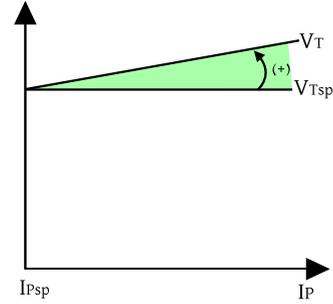


Figure 3: Characteristic curve of active current compensation control function. V_T : Generator terminal voltage, I_P : Generator terminal active current output, V_{Tsp} : Generator terminal voltage setpoint, I_{Psp} : Active current setpoint.

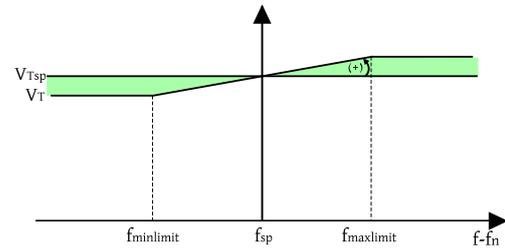


Figure 4: Characteristic curve of frequency droop control function. V_T : Generator terminal voltage, f : Actual frequency of the generator, V_{Tsp} : Generator terminal voltage setpoint, f_{sp} : Frequency setpoint.

generator terminal voltage will be regulated. The voltage droop/compensation controller consists of four control functions or three controllers. The modelling of the controllers is described below.

3.1. Reactive Current Droop and Compensation Controller

This controller inherent a combination of reactive current droop and reactive current compensation control functions. The controller is fundamentally modelled based on the formula given in (1) [5]. This formula is used to calculate the new generator terminal voltage setpoint considering the droop/compensation value and actual generator reactive current output. The E_{rr} obtained by subtracting the measured terminal voltage V_T from the calculated generator terminal voltage value V_{Tcal} and then the E_{rr} is applied through the PID controller to the output. Additionally, the droop (regulation) value should be given in percent, and most importantly, it should be a negative value “-” for the droop control function and a positive value “+” for the compensation control function. Fig. 5 illustrates the block diagram of the reactive current droop/compensation controller. The Boolean signal $I_{Qcontroller}$ should be “true” in order to activate the output of the controller, else the output will be zero.

$$V_{Tcal} = \left(\left(\frac{I_Q - I_{Qsp}}{I_{Qn}} \cdot \frac{R_{IQ}}{100} \right) + 1 \right) \cdot V_{Tsp} \quad (1)$$

where

R_{IQ} :	The droop (regulation) value	[%]
I_Q :	The actual generator reactive current output	[pu]
I_{Qsp} :	The generator reactive current setpoint	[pu]
I_{Qn} :	Generator nominal reactive current	[pu]
V_{Tcal} :	The calculated new generator terminal voltage setpoint	[pu]
V_{Tsp} :	The generator terminal voltage setpoint	[pu]

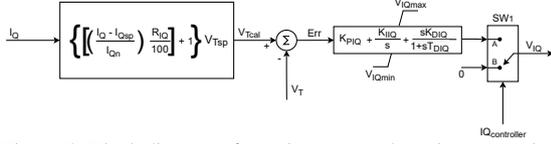


Figure 5: Block diagram of reactive current droop/compensation controller

3.2. Active Current Compensation Controller

The active current compensation control function model uses the formula given in (2) to calculate the new generator terminal voltage setpoint [35]. This controller regulates the voltage only if the actual active current I_P is higher than the active current setpoint I_{Psp} . Meaning, if the I_P is less than the I_{Psp} the active compensation controller will not react. Further, the E_{rr} is calculated by subtracting the V_T from V_{Tcal} then applied to the output through the PID controller. The regulation value R_{IP} should be a positive value to obtain the compensation function. Otherwise, the controller will behave on the contrary. The block diagram of the active current compensation controller is presented in Fig. 6. The Boolean signal $IP_{controller}$ should be “true” to change the position in the switch SW_1 to enable the output of the controller, otherwise the it will be zero.

$$V_{Tcal} = \left(\left(\frac{I_P - I_{Psp}}{I_{Pn}} \cdot \frac{R_{IP}}{100} \right) + 1 \right) \cdot V_{Tsp} \quad (2)$$

where

R_{IP} :	The droop (regulation) value	[%]
I_P :	The actual generator active current output	[pu]
I_{Psp} :	The generator active current setpoint	[pu]
I_{Pn} :	Generator nominal active current	[pu]
V_{Tcal} :	The calculated new generator terminal voltage setpoint	[pu]
V_{Tsp} :	The generator terminal voltage setpoint	[pu]

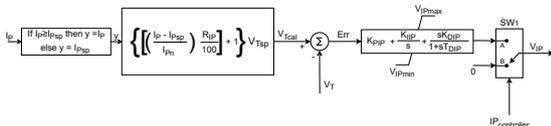


Figure 6: Block diagram of active current compensation controller

3.3. Frequency Droop Controller

This controller model is modelled based on (3) to determine the new generator terminal voltage setpoint. This controller behaves similarly to the latter controllers, where the E_{rr} is obtained by comparing the V_T and V_{Tcal} , then applying this through the PID controller to the

output. The droop (regulation) value should be a positive value to obtain the compensation function. Besides, this controller has an additional function limiting the voltage support when the frequency exceeds the maximum and minimum limit, $f_{maxlimit}$ and $f_{minlimit}$, respectively. This means that the frequency droop controller will not increase or decrease the V_T when the frequency exceeds the latter limits. The Boolean signal $f_{controller}$ should be “true”, and the circuit breaker should be closed in order to activate the output of the controller, else the output is zero. The block diagram of the frequency droop controller is depicted in Fig. 7.

$$V_{Tcal} = \left(\left(\frac{f - f_{sp}}{f_n} \cdot \frac{R_f}{100} \right) + 1 \right) \cdot V_{Tsp} \quad (3)$$

where

R_f :	The droop (regulation) value	[%]
f :	The actual frequency	[pu]
f_{sp} :	The frequency setpoint	[pu]
f_n :	Nominal frequency	[pu]
V_{Tcal} :	The calculated new generator terminal voltage setpoint	[pu]
V_{Tsp} :	The generator terminal voltage setpoint	[pu]

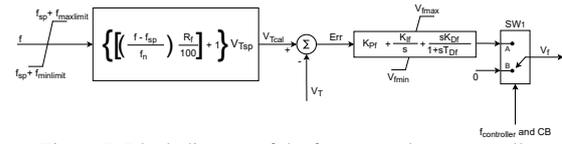


Figure 7: Block diagram of the frequency droop controller

3.4. Final Combined Controller

All three controllers mentioned in Section , , and are added together into a voltage droop/compensation controller model as shown in Fig. 8). The checkboxes III, V, and VII in Fig. 9 shall be selected in order to enable the outputs of the reactive current droop/compensation, active current compensation, and frequency droop controllers, respectively. These checkboxes are associated with the switches, $SW1_{IQ}$, $SW1_{IP}$, and $SW1_f$. Thus, the controllers can either be used alone or in combination with others to regulate the voltage. All the setpoint values can be chosen as constant or variable setpoints by choosing the checkboxes indicated with II, IV, VI, and VIII in Fig. 9. When the checkboxes II, IV, VI, and VIII are checked, variable setpoint inputs such as terminal voltage setpoint V_{Tvsp} , reactive current setpoint I_{Qvsp} , active current setpoint I_{Pvsp} , and frequency setpoint f_{vsp} will be enabled to connect, respectively. Simultaneously, when those are activated, the associated constant setpoints will be disabled. Moreover, conditional connections are visibly indicated with dashed lines in Fig. 8. Note that the controller’s parameters are placed in individual tabs as indicated with (I) in Fig. 9, while the common terminal voltage setpoint options are placed in the tab called “General”. Further, in active compensation and frequency droop control functions, an absolute block is used to assure that a given negative droop (regulation) value (R_{IP} and R_f) does not change the characteristics of the control functions. Be aware of the named parameters in the controller’s block diagram and the model because they are changed due to modelling purposes.

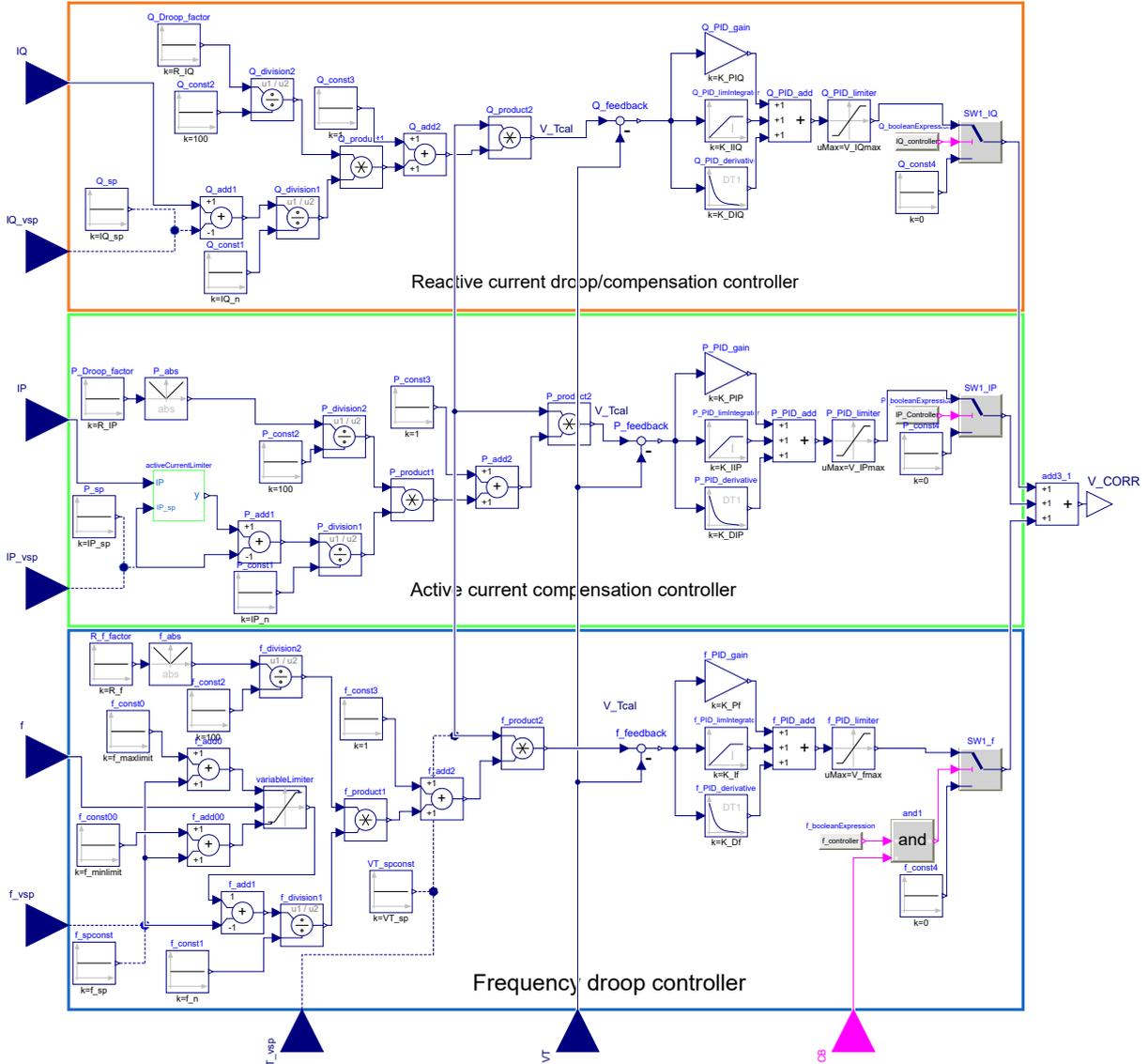


Figure 8: Implementation of voltage droop/compensation controller in Modelica

4. Simulation Results

This section presents simulation results of voltage droop/compensation controller. A test setup of the voltage droop/compensation controller is portrayed in Fig. 10. The test setup is created using a GENSL generator, transmission line, infinite grid, and excitation system type *ST7C* from the OpenIPSL version 2.0.0 [1], as shown in Fig. 10. The system power base and frequency for all the components are set to 10 MVA and 50 Hz, accordingly. The generator is initialised, as presented in Table 1, during the various simulations. Also the voltage setpoint V_{Tsp} , reactive current setpoint I_{Qsp} , active current setpoint I_{Psp} , and frequency setpoint f_{sp} are varied to examine the controller. The simulation is performed individually for each controller by changing the controller's latter setpoints at 1200 s and the voltage setpoint at 2200 s.

4.1. Reactive Current Droop and Compensation Controller

There are performed two tests with this controller,

Table 1: Initialisation of GENSL generator for simulation

Name	Description	Value	Units
P_0	Initial active power	2	MW
Q_0	Initial reactive power	1	Mvar
v_0	Initial voltage magnitude	1	pu
$angle_0$	Initial voltage angle	0	°
ω	Initial generator speed	0	pu

first with the droop function and the second with the compensation function. The generator reactive current setpoint I_{Qsp} is changed from zero to 0.5 pu at 1200 s, and the voltage setpoint V_{Tsp} is changed from 1 to 1.05 pu at 2200 s. Initially, the controller starts to influence the AVR to reduce the terminal voltage equal to the predefined voltage setpoint, as shown in Fig. 11. Please note that when the AVR influences, the field voltage applied to the generator will be affected. Consequently, the reactive power or current output is affected to obtain the desired

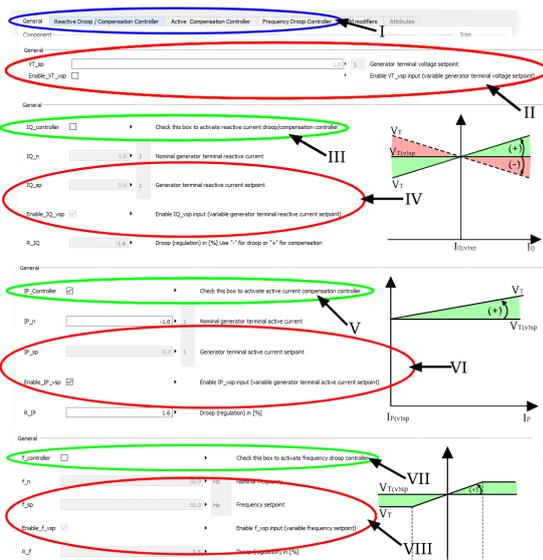


Figure 9: Implemented user interfaces in Modelica for voltage droop/compensation controller. I: Tabs for each controller. II: Checkbox to enable the variable generator terminal voltage setpoint. III: Checkbox to enable Reactive current droop/compensation controller. IV: Checkbox to enable the variable generator terminal reactive current setpoint. V: Checkbox to enable Active current compensation controller. VI: Checkbox to enable the variable generator terminal active current setpoint. VII: Checkbox to enable the Frequency droop controller. VIII: Checkbox to enable the variable frequency setpoint.

terminal voltage. Hence, the terminal voltage is stabilised at 1 pu before I_{Qsp} changes. When I_{Qsp} changes, the reactive current increases, hence the voltage increases. The stabilised voltage ends up at 1.00748 pu, which is similar to the calculated value. Simultaneously, when the V_{Tsp} increases, there is a significant change in terminal voltage due to an increase in the reactive current. The deviation between the calculated value and the simulated value is found to be about $1.97 \cdot 10^{-6}$, which is reasonable. Note that the steady-state terminal voltage after a setpoint change can be calculated using (1) to verify the results. Since the initial terminal voltage is higher than the preset setpoint, the controller reduces the terminal voltage, as shown in Fig. 12. The voltage is finally stabilised at 0.9999 pu before any setpoint changes, which is corresponds to manually calculated terminal voltage using (1). After the increase in I_{Qsp} at 1800 s, the reactive current and the terminal voltage are decreased to roughly -0.0489 pu and 0.9912 pu, respectively. Whereas change in V_{Tsp} is causing the voltage to rise again to 1.0453 pu, as expected.

4.2. Active Current Compensation Controller

The results from the active current compensation controller simulation are presented in Fig. 13. Where the generator active current setpoint I_{Psp} is changed from 0.5 to 0 pu at 1200 s, and the voltage setpoint V_{Tsp} is changed from 1 to 1.05 pu at 2200 s. At the initial stage, when the setpoint is at 0.5 pu, the generator active current output I_P is at 0.1996 pu, thus the controller does not react on I_P . It will rather consider the setpoint as an actual active current output, and it compensates for it because the controller does not operate for any I_P below the setpoint. Thus, the terminal voltage is reduced to 1 pu by regulating the generator's reactive power or current. Later, the setpoint

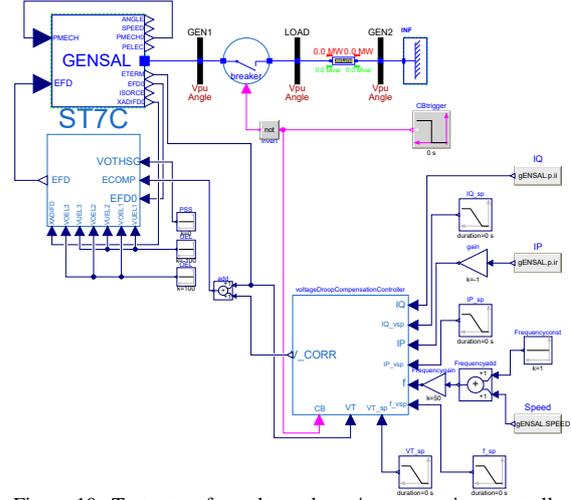


Figure 10: Test setup for voltage droop/compensation controller model

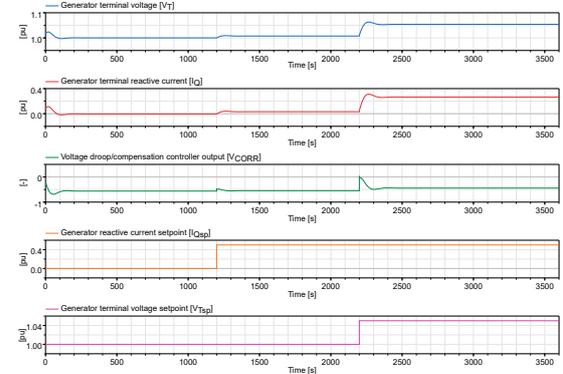


Figure 11: Performance of voltage droop/compensation controller when the reactive current droop function is activated.

reduces to zero, then the controller starts to compensate for the actual active current output. Hence, I_P is higher than the setpoint, the terminal voltage increased and stabilised at 1.0048 pu, as expected. When the voltage setpoint increased, as a result, the terminal voltage increased to 1.0498 pu, which is equal to the manually calculated value, where the manually calculated value is acquired by using (2).

4.3. Frequency Droop Controller

Fig. 14 illustrates the simulation results of the voltage droop/compensation controller using the frequency droop function. Since in the beginning, the nominal frequency and the frequency setpoint is at 50 Hz, and terminal voltage is higher than V_{Tsp} , the controller tries to reduce the voltage to 1 pu. Afterwards, when the frequency setpoint is reduced to 48 Hz, consequently the voltage is increased to 1.001 pu as expected. And, when the voltage setpoint V_{Tsp} is increased to 1.05 pu at 2200 s, the terminal voltage rises again and stabilises at roughly 1.05 pu as desired.

5. Discussion

This paper aims to model voltage droop/compensation controller in the Modelica modelling language. Fundamentally, the controller is modelled based on

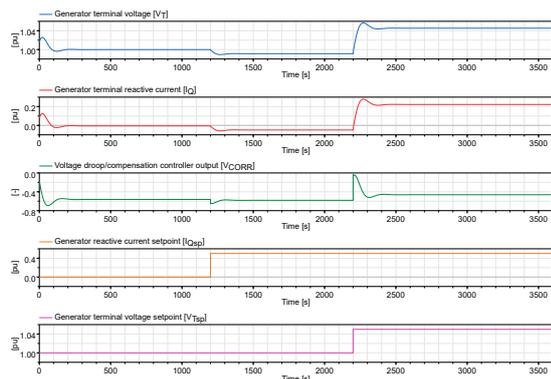


Figure 12: Performance of voltage droop/compensation controller when the reactive current compensation function is activated.

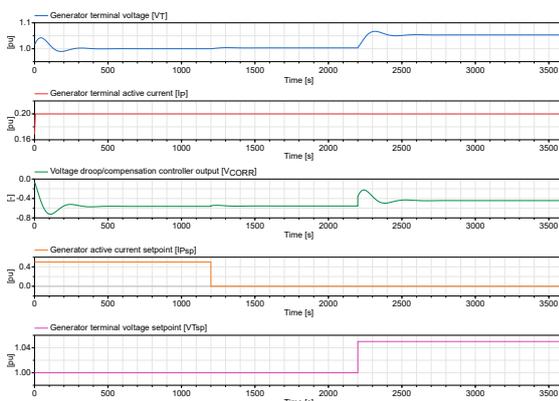


Figure 13: Performance of voltage droop/compensation controller when the active current compensation function is activated.

the requirements in NVF 2020; however, the models have been modified slightly for modelling purposes. The excitation system, type ST7C, is obtained from the OpenIPSL library.

Since the primary focus of this paper is to model the voltage droop/compensation controller, the test setup modelling is kept simple as possible to analyse the model performance.

The overall behaviour of the model was reasonable to compare to the theoretical behaviour. However, the simulation results should be compared with the real controller to verify the model performance. The controller models have a switch at the output to disable the control function. These switches may cause a sudden increase in the control signal, consequently, an overshoot in terminal voltage output may occur. It can be eliminated by adding a self-reset function for the PID controller to reset when the output is re-enabled.

6. Conclusions

In this paper, the voltage droop/compensation controller is mainly object-oriented modelled in Modelica modelling language using Dymola software. The models are fundamentally modelled with reference to requirements in the National Guide for Functional Requirements in the Power System, NVF 2020, [6].

The voltage droop/compensation controller modelled separately from scratch, except the AVR, obtained from the external library OpenIPSL. Later, the model was simulated separately and then compared to these

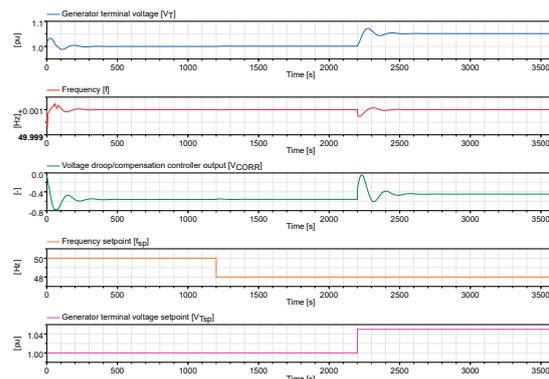


Figure 14: Performance of voltage droop/compensation controller when the frequency droop function is activated.

controller's theoretical behaviour.

In conclusion, the model performed as desired but still need proper tuning and further development to enhance the performance. For the future it is planned to run further tests with real power plant data in order to improve and verify the behaviour of the limiter models.

Acknowledgement

This paper is the result of the Master's Thesis of Luxshan Manoranjan, with the title "Modeling the Excitation Control System of a Hydropower Controller in Modelica" [2]. The thesis also covered the modelling of limiters which has been published in [3].

References

- [1] ALSETLab, Maxime Baudette, Marcelo Castro, Tin Rabuzin, Jan Lavenius, Tetiana Bogodorova, and Luigi Vanfretti. OpenIPSL: Open-Instance Power System Library. 7:34–36. ISSN 23527110. doi: 10.1016/j.softx.2018.01.002.
- [2] Luxshan Manoranjan. Modeling the Excitation Control System of a Hydropower Controller in Modelica. URL <https://hdl.handle.net/11250/2773845>.
- [3] Luxshan Manoranjan and Dietmar Winkler. Developing Protective Limiters for a Hydro Power Controller in Modelica. In *Proceedings of the 14th International Modelica Conference*. doi: 10.3384/ecp21181617.
- [4] Modelica Association. Modelica a unified object-oriented language for systems modeling. Language specification version 3.5. URL <https://specification.modelica.org>.
- [5] Pieter Schavemaker and Lou Van der Sluis. *Electrical Power System Essentials*. John Wiley. ISBN 978-0-470-98768-1.
- [6] Statnett SF. National Guide for Functional Requirements in the Power System (Nasjonal veileder for funksjonskrav i kraftsystemet).
- [7] University of South-Eastern Norway, TMCC. OpenHPL. URL <https://openhpl.opensimhub.org>.

Control strategies minimizing wastewater overflow in Oslo

M. Halid Seyhan^a, Tiina M. Komulainen^{a,*}

^a *Oslo Metropolitan University, Institute for Mechanical, Electrical and Chemical Engineering,*

**tiina.komulainen@oslomet.no*

Abstract

As climate change intensifies storms, larger rainwater volumes load the sewage network systems above the design capacity and escalate the risk of combined sewer wastewater overflow to natural waterways. Accordingly, the control challenge is to prevent the combined sewer overflow by adjusting the manipulated variables, such as pumps and gates, in the sewage network system. The aim of this study is to (1) compare traditional and predictive control strategies to four different storm scenarios, and to (2) quantify the preventive effect of these control strategies on wastewater overflow. The case study is applied in the Oslo combined sewage network system. Control strategies applying feedforward strategy minimize the overflow within the constraints of the sewer infrastructure. Compared to no control, applying feedforward-feedback control strategy decreases overflow to natural water ways by 21- 88% in different rain scenarios. Compared to feedback control strategy, the feedforward-feedback strategy can decrease overflow by 3--9%.

1. Introduction

As climate change intensifies storms, larger rainwater volumes load the sewage network systems above the design capacity and escalate the risk of combined sewer wastewater overflow to natural waterways. Accordingly, the control challenge is to prevent the combined sewer overflow by adjusting the manipulated variables in the sewage network system.

Most urban water systems are controlled by passive control, rule based local control or manual operation in order to minimize overflow to natural waterways and optimize energy consumption for pumping [1]. In manually controlled combined sewer systems, heavy rain can easily lead to overflows if preventive manual actions are not taken prior to and during the rain event. In order to minimize the combined sewer overflow, research in different real time control algorithms has gained strong interest during the past decade. Real time control uses sensors and controllers, and an automation system with supervisory control and data acquisition. The controllers convert real-time measurements into operational decisions by rules and algorithms of varying complexity [1].

Most of the advanced control applications to combined sewer systems have been implemented in simulation environment. For example, Schuetze and Alex [2] have used Simba model of the Astlingen sewer network in Germany to quantify the combined sewer overflow volume of the during a year using different control strategies. Their findings confirm that MPC would provide the minimal volume close

between the base case and the theoretical optimum for the Astlingen sewer system [2]. A model-based approach on estimation and model predictive control of the wastewater levels in the Oslo sewer network tunnel basins has been proposed by [3].

So far, only a few industrial implementations of real-time control have been presented, due to insufficient instrumentation of the sewer network. Model predictive control has been applied to utilize the water storage capacity of the sewer network to minimize costs of pumping during varying electricity prices in Denmark [4].

In this study model-based control is applied to the Oslo combined sewage network system. Oslo metropolitan area with its surrounding municipalities host more than one million citizens and the region receives approximately 1010 mm precipitation (rain) each year. Despite continuous improvements on the infrastructure of the sewer system, storm events exert a considerable pressure on the city's urban drainage system and combined sewer overflows to the Oslo fjord occasionally. The overflow of diluted, untreated wastewater-rainwater mixture affect negatively the marine life and water quality at beaches along the Oslo Fjord. The aim was to (1) compare traditional and predictive control strategies to four different storm scenarios, and to (2) quantify the preventive effect of these control strategies on wastewater overflow. Research question: Which control strategies can minimize overflow of diluted wastewater during heavy rain?

2. Materials and Methods

2.1. Materials

This study uses a high-fidelity MIKE simulator of the Oslo urban drainage system for the data collection. DHI has developed and calibrated the dynamic simulator as part of the Future City Flow EU-project. Simplified modeling, control strategy development and testing were done in Matlab/Simulink.

2.2. Methods

Modeling: High-fidelity MIKE model is used to collect data of historical rain scenarios in the Oslo sewage network. The MIKE+ data was used to develop and calibrate the simplified model consisting of ordinary differential equations and algebraic equations. The simplified model was implemented in Matlab/Simulink.

Control: The ordinary differential equations were linearized and Laplace-transformed to transfer function models. The transfer function models were used for parametrization and tuning the MPC and the PID controllers. The control algorithms were tested in Matlab/Simulink environment using the ordinary differential equations. ODE23t solver was used for all scenarios and controllers. Due to late data access, control algorithm testing in MIKE+ software was not possible.

2.3. Symbols

Symbols used in equations are given in Table 1.

Table 1: List of symbols

Symbol	Description	Unit
A	Catchment area	m ²
B	Magazine bottom area	m ²
F	Flowrate	m ³ /min
i	Catchment i	-
I	Imperviousness	%
j	Pipeline j	-
m	Magazine m	-
PE	Person equivalent	-
R	Rain fall	mm/min
t_D	Transport delay in pipeline or magazine	min
θ	Catchment area delay	min
		-

2.4. System description

The combined sewer network in Oslo and the surrounding municipalities stretches over 108 catchment areas. In this study, based on geography, these areas were combined into 9 catchments with estimated average characterization given in Table 2. The characterization of the catchments is based on data from the MIKE+ model. The water from the catchment areas is led to a network sewer tunnels that transport the water to two separate water

resource recovery facilities, Veas in the west side of Oslo and Bekkelaget in the east side of Oslo.

Table 2: Catchments areas in the Oslo metropolitan sewer network with tunnel inlet point, estimated person equivalent PE, area A, imperviousness I and delay θ.

Inlet	Catchment	PE [10 ³]	A [10 ⁶ m ²]	I [%]	θ [min]
I1	Festning	150	77	11,7	28,6
I2	Østensjø	125	77	9,77	28,6
I2	Furuset	150	115	7,81	42,9
I3	Torshov	150	115	7,81	42,9
I4	Nordstrand	125	77	9,77	28,6
I2	Lillestrøm+	181	415,8	2,62	154
I5	Bærum	135	192	4,22	71
I6	AskerLier	130	192	4,06	71
I5	SkiOppgård	65	192	2,03	71

2.5. Data collection and case scenarios

Four historical rain events were chosen for this study, case 1 and 2 with heavy storm, case 3 with extremely heavy storm and case 4 with moderate rain. The length of each case scenario is four days (96h) in order to let the combined sewer system settle back to the dry conditions after the rain event. The periods selected have dry weather during the first day, rain during the second day and dry weather during the third and fourth day, as presented in Figure 1. A summary of the rain scenarios in the Oslo combined sewer network is given in Table 3.

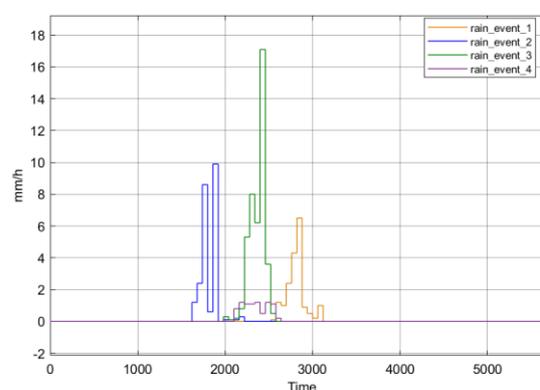


Figure 1: Rain fall during case scenarios.

Table 3: Characterization of case scenarios with length of 96 hours.

Case	Rain start time [dd:mm:yy]	Rain fall [mm]	Rain duration [h]
1	11.06.19 19:00	18	9
2	29.08.19 04:00	22,7	5
3	09.08.17 10:00	41,5	7
4	09.05.16 12.00	8,2	8

3. Modeling

The modelling of the Oslo sewer network consists of four sub-models. The first sub-model is between rainfall and sewer network inlet. The second sub-model provides continuous municipal wastewater flow with sinusoidal diurnal pattern to the inlet of the sewer network. The third sub-model consist of the main tunnels in the sewer network. The sewer tunnel network is approximated as 13 links (transport delay) and 5 magazines with storage volume and a final control element (pump or gate). The fourth sub-model is the overflow accumulator in the magazines. In the simplified Matlab/Simulink model (Figure 2), every catchment area has only one inlet to a tunnel in the sewer network. The tunnel network leads water to the water resource recovery facilities. Parts of the tunnels are used to store a fixed volume of water during heavy rains. The storage capacity in the tunnels is modeled as magazines with finite volumes. The water flow through the magazines is controlled with a pump or a gate. If the maximum level limit in the magazine is reached, overflow is directed to natural waterways.

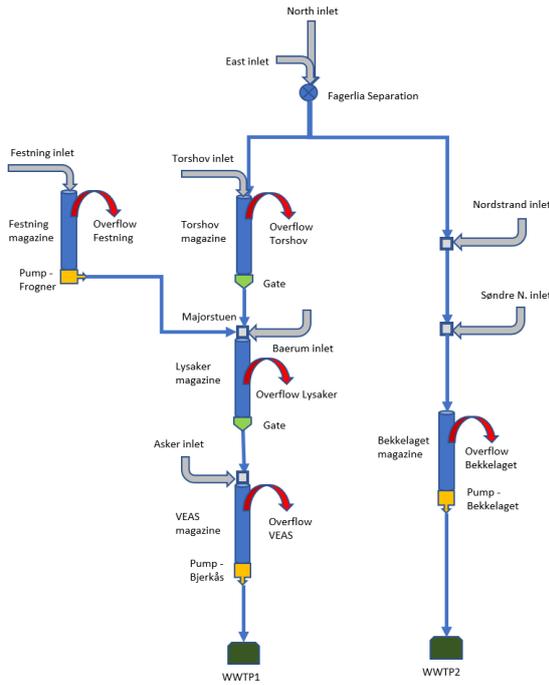


Figure 2: Simplified model of the combined sewage network in Oslo area.

The inlet boundaries of the model are the flowrates from the six catchment areas. The flowrate from catchment area i to the tunnel inlet $F_{inlet,i}$ is estimated as a sum of the municipal wastewater flow $F_{mun,i}$ and the rain induced water flow $F_{rain,i}$.

$$F_{inlet,i} = F_{mun,i} + F_{rain,i} \quad (1)$$

The municipal wastewater flow is calculated based on population (person equivalent) in the area, person equivalent wastewater flow and the diurnal pattern (

24h day-night variation). The person equivalent wastewater flow F_{PE} is assumed $0.0002005 \text{ m}^3/\text{min}$. The diurnal pattern is estimated as sinus curve with amplitude of 1, bias of 1 and angular frequency of $2\pi/(1440\text{min})$, with zero value at 05:00 and maximum value of 2 at 17:00.

$$F_{mun,i} = F_{PE} \cdot PE_i \cdot \left(1 + \sin\left(\frac{2\pi}{1440}t\right)\right) \quad (2)$$

The rain induced water flow from catchment area i is estimated as a product of rain intensity R_i , catchment area A_i and catchment imperviousness I_i , and delay θ between rain fall and rain induced flow at the tunnel inlet.

$$F_{rain,i} = R_i(t + \theta) \cdot 10^{-3} \cdot A_i \cdot I_i \quad (3)$$

The wastewater flow from tunnel inlets is gathered to the sewer tunnels. The tunnels are modeled as links and magazines. A link is modeled as a plug flow through the tunnel, a pure transport delay, given in Table 4. The flowrate through the link j is the flowrate in with a transport delay t_{Dj} :

$$F_{outlet,j} = F_{inlet,j}(t - t_{D,j}) \quad (4)$$

Table 4: Parameters of the tunnel/link models with inlet, outlet, length L and transport delay t_D .

Link	from	to	L [km]	t_D [min]
L1	Festning inlet	Festning magazine - Majorstuen	4,4	44
L2	East inlet (Østensjø)	Fagerlia separation	3,5	35
L3	North inlet (Furuset)	Fagerlia separation	7,1	70
L4	Fagerlia separation	Torshov magazine - Majorstuen	7,3	72
L5	Torshov inlet	Torshov magazine - Majorstuen	3,0	30
L6	Majorstuen	Vaekerø	5,4	54
L7	Vaekerø	Lysaker magazine	7,6	75
L8	Baerum inlet	Lysaker magazine	6,0	59
L9	Lysaker magazine	VEAS magazine - WWTP1	15,0	149
L10	Asker inlet	VEAS magazine - WWTP1	8,0	79
L11	Fagerlia separation	Bekkelaget magazine - WWTP2	5,5	55
L12	Nordstrand inlet	Bekkelaget magazine - WWTP2	1,0	10
L13	Søndre N. inlet	Bekkelaget magazine - WWTP2	7,5	74

The Oslo sewer network model includes five magazines which are modelled as a pipeline with transport delay t_{Dm} and a vertical tank volume in the end. The parametrization of the magazines is given in Table 5. The plug flow through the pipeline part of the magazine feeds to the tank at the end of the magazine. The flow rates from different links (L) is summed together at the inlet of the magazine pipeline and plug flow with transport delay t_{Dm} is assumed along the magazine pipeline:

$$F_{inlet,m} = \sum F_{outlet,j}(t - t_{D,m}) \quad (5)$$

Water height in the vertical tank at the end of the magazine is modeled with a simple tank. The magazine tank has inlet at the top and outlet at the bottom with a pump at the exit line. The maximum water height in the vertical tank model is about 3 meters and the minimum water level is 0,1 m (to avoid numerical instabilities). The nominal dry weather level in the magazines is half of the maximum, about 1,5 m. The water height in the magazine is given as:

$$\frac{dh_m(t)}{dt} = \frac{1}{B} (F_{inlet,m}(t) - F_{outlet,m}(t)) \quad (6)$$

Each magazine has an overflow model that applies if the high-level limit is reached. The volume of the overflow water is calculated as accumulated sum.

Table 5: Parameters of the magazine models with max water height (H), bottom area (B) and nominal transport delay (t_D).

ID	Magazine	Link	H _{max} [m]	B [10 ³ m ²]	t _D [min]
M1	Festning	L1	3	10,3	44
M2	Torshov	L4	3	17,2	72
M3	Lysaker	L7	3,3 5	19,9	75
M4	Bekkelaget	L1 1	3	12,9	55
M5	Veas	L9	3,3 5	39,4	14 9

The time delay t_D for each link and magazine with length L_j is calculated based on estimated speed of flow through the tunnel. This estimate is based on the minimum flow velocity $V_{fmin}=84.05$ m/min, and the flow velocity difference $V_{fdif}=33.6$ m/min between maximum flow velocity (117,7 m/min) and minimum flow velocity. The flow velocity difference is weighted with the hydrostatic ratio in the magazine, the water level $H(t)$ in the magazine j divided by maximal water level H_{max} .

$$t_{D,j} = \frac{L_j}{V_{f,min} + V_{f,dif} \cdot (H(t)_j / H_{max})} \quad (7)$$

The final control element in each magazine is either a pump station (Frognerparken, Bjerke, Bekkelaget) or a gate (Torshov, Lysaker). It is assumed that the time dynamics of these are negligible, and thus they are represented with adjustable gain of 0-100% of the maximum outlet flowrate given in Table 6. The Fagerlia weir has position between 0,3-0,7 the nominal value is 0,5. The weir does not restrict the flowrate.

Table 6: Final control elements of the sewer network with minimum, nominal and maximum flowrates.

ID	Placement	F _{min} [m ³ /min]	F _{nom} [m ³ /min]	F _{max} [m ³ /min]
M1	Festning (pump)	6	30	150
M2	Torshov (gate)	0	152	343
M3	Lysaker (gate)	0	266	638
M4	Bekkelaget (pump)	17	84	440
M5	Veas (pump)	32	159	660
F6	Fagerlia (weir)	0	-	-

In addition, wastewater from the catchment areas in North of Oslo can be divided at Fagerlia separation using a flow separation weir between west sewer system leading to Veas (WWTP1) and east sewer system leading to Bekkelaget (WWTP2).

The outlet boundary conditions of the model are the two water resource recovery plants, Veas at west and Bekkelaget at east, with parameters given in Table 7. The Bekkelaget capacity was increased by 50% in October 2021, but in the simulation model for all scenarios it is assumed that this capacity applies already in 2016.

Table 7: Water resource recovery facility with estimated capacities.

Name	P.E capacity	Normal flow in [m ³ /min]	Max flow in [m ³ /min]
Veas	793000	159	660
Bekkelaget	418000	83,8	440

4. Control strategy development and testing

4.1 Control strategy development

The control goal is to avoid overflow in the magazines. Three different control strategies were developed based on linearized transfer function models of the governing nonlinear equations. The five (5) controlled variables are the wastewater levels in the five magazines. The manipulated

variables are three (3) pump stations, two (2) gates, and separation weir (1). The disturbance variables of the system are the six (6) lumped tunnel inlet flowrates.

Transfer functions

For the controller development purposes, the ordinary differential equations of the system between the controlled variables and manipulated variables were linearized and Laplace transformed. The first order transfer functions between the controlled variables and the manipulated variables are integrators with gain K_p , presented in Table 8. The transfer functions in the magazine tanks were estimated with a delay τ_d of zero as the final control element is placed right at the outlet of the tank model. It is assumed that the Fagerlia flow separation weir has negligible time dynamics compared to the pumps and gates.

Table 8: Transfer functions between controlled variables (levels) and manipulated variables (pumps and gates).

TF	MV	CV	K_p [-]
TF1	Festning pump P_F	H_F	$-1/B_1$
TF2	Torshov gate G_T	H_T	$-172/B_2$
TF3	Lysaker gate G_L	H_L	$-319/B_3$
TF4	Bekkelaget pump P_B	H_B	$-1/B_4$
TF5	Veas pump P_V	H_V	$-1/B_5$

4.2 SISO control strategy development

As each magazine has one controlled variable (level) and one manipulated variable (pump station or weir), these were paired for the single-input single-output control strategy. Fagerlia separation weir has its own independent controller for allocating the northern inflow between the west and east sewer networks

The PI-controller parameters for the level controllers in the magazines were calculated based on the Skogestad IMC tuning [5] rules for integrating system. The transfer functions are given in Table 8. The tuning parameter τ_c was chosen as 1 minute because the transfer functions did not include a delay. The controller gain at magazine j was calculated as:

$$K_{c,j} = \frac{1}{K_{p,j}(\tau_{c,j} + \tau_{d,j})} \quad (8)$$

The controller integral time at magazine j was calculated as:

$$T_{i,j} = 4 \cdot (\tau_{c,j} + \tau_{d,j}) \quad (9)$$

The PI-controller parameters are given in Table 9.

Table 9: PI-control parameters for the final control elements at magazines.

ID	Description	TF	Kc	Ti [min]	Rate limiter [1/min]
M1	Festning (pump)	1	-10 367	4	$\pm 4,8$
M2	Torshov (gate)	2	- 100	4	$\pm 0,033$
M3	Lysaker (gate)	3	-63	4	$\pm 0,033$
M4	Bekkelaget (pump)	4	-39 466	4	± 21
M5	Veas (pump)	5	-12 959	4	± 14

Proportional control with the following relationship between water height at Torshov magazine $H_T(t)$ and Bekkelaget magazine $H_B(t)$ was applied to the flow separation weir at Fagerlia:

$$W_1(t) = 0,5 + 0,2 \left(\frac{H_T(t)}{H_{T,max}} + \frac{H_B(t)}{H_{B,max}} \right) \quad (10)$$

4.3 MIMO control strategy development

Linear model predictive control was chosen as the multi input-multi output control algorithm due to its popularity in similar applications. The linear MPC tuning rules were adopted from [6] and [1]. The MPC controller was parametrized using the transfer functions between the controlled variables (levels) and the manipulated variables (pumps, gates and weir) given in Table 8. The sampling time of the system was chosen 5 min. As the transfer functions of the integrating systems (magazines in Table 8) do not have time constants, a model horizon N of 120 minutes was chosen. According to [6], the control horizon M was chosen to be half of N , 60 minutes. The collective horizon approach with equal control and prediction horizons, presented in [1] was the most used approach found in the literature review and adopted in this study. The prediction horizon P was therefore 60 minutes. The tuning parameter Q , weighting the importance of the five controlled variables, was chosen as identity matrix, because keeping the level under high limit is as important in every magazine. The tuning parameter R , penalizing changes in the manipulated variables was set to zero. The rate limiters of the PI-controllers (Table 9) were implemented also as separate blocks for the control signals (outputs) of the model predictive controller.

4.4 Feedforward control strategy development

A feedforward control strategy was designed to maximize the available storage volume in the main tunnels of the combined sewer system. The strategy uses the weather forecast and lowers the water levels in the combined sewer tunnels to minimum before the anticipated rain event. In this study, the

feedforward strategy was implemented as real-time optimization of controlled variable setpoints using rain forecast data. The nominal setpoints of the magazines water levels were 50% of the maximum level. Twelve hours prior to the forecasted rain event, the setpoints were lowered to 0,2 m for 28 hours, and then, lifted back up to nominal value.

4.5. Control strategy testing

The control results are illustrated for case 4. The precipitation data is presented in Figure 3. The combined sewer inlet flowrates from the six lumped catchment areas are presented in Figure 4. The changes in the manipulated variables are shown in Figure 5 and the setpoint tracking of the controlled variables in Figure 6.

The results for total overflow with the difference control strategies compared to no control is given in Table 10. Applying control strategies decrease the overflow as can be seen from Table 11 comparing overflow reduction between no-control strategy and the other control strategies. The maximal overflow reduction is about 0,6 million m³, if the water levels in the combined sewer tunnels (magazines) are at its lowest when the rain starts.

Implementation of PI and MPC control strategies decrease the overflow by 40 % in heavy rain cases 1 and 2, and up to 80% for moderate rain case 4, whereas in extreme heavy rain case 3 the limitation of the buffer/storage capacity of the combined sewer system only allow overflow decrease of 16 %.

Both PI and MPC control strategies with forecast (lowering storage setpoint 12 hours prior to rain) improves the results compared to no-control result by 45 % in heavy rain cases 1 and 2, 88% in moderate rain case 4 and by 21 % in extremely heavy rain case 3.

Table 10: Total overflow in Mm³ for case scenarios with different control strategies.

Case	No control	PI	PI+F	MPC	MPC+F
1	1,13	0,62	0,58	0,62	0,58
2	1,49	0,96	0,88	0,95	0,89
3	2,78	2,34	2,20	2,34	2,21
4	0,45	0,095	0,055	0,094	0,056

Table 11: Difference [Mm³] in total overflow between no-control strategy and other control strategies.

Case	PI	PI+F	MPC	MPC+F
1	0,51	0,55	0,51	0,55
2	0,54	0,61	0,55	0,60
3	0,44	0,59	0,44	0,58
4	0,36	0,40	0,36	0,40

Table 12: Difference [%] in total overflow between no-control strategy and other control strategies.

Case	PI	PI+F	MPC	MPC+F
1	45	49	45	49
2	36	41	37	40
3	16	21	16	21
4	79	88	79	88

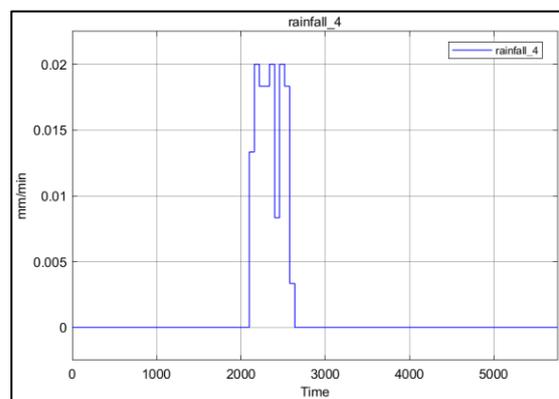


Figure 3: Rain event 4 with real precipitation measurement from Blindern weather station in Oslo.

5. Summary and Discussions

For all scenarios, real-time optimization using forecast data have significantly improved the performance of both PI and MPC control strategies. PI control with forecast has performed slightly better than MPC with forecast. Additionally, PI with forecast had lower pump loads than MPC with forecast.

Thus, the answer our research question is: Control strategies applying feedforward strategy will minimize the overflow within the constraints of the infrastructure. Compared to no control, applying feedforward control strategy that uses rain prediction, overflow to natural water ways can be decreased up to 88% in moderate rain scenario, up to 49% in heavy rain scenarios and up to 21% in extreme storm scenario. Compared to control strategy without feedforward strategy, the feedforward strategy can minimize overflow by 3--9%.

Due to late data access during the project, some assumptions in the combined sewer network model need to be updated. As our project partners have pointed out, some of the catchment areas (Lillestrøm, Lier, Ski and Oppegård), should be omitted as these deliver water to other water resource recovery facilities than Bekkelaget and Veas. As these four catchment areas produce only small flowrates to the current combined sewer model, it is assumed that the effect to the final results is minimal.

Future work is suggested on alternative SISO and MIMO controller parametrizations, new case studies

on different rain events, and handling of uncertainties of rain events using advanced control strategies such as stochastic MPC using the Matlab model. Implementation of the best control algorithms should be implemented in the MIKE+ model and tested with different rain scenarios.

Acknowledgments

Contributions on problem statement and comments on research results from Hilde Johansen/ Veas water resource recovery facility, Lars-Olov Orre/ Oslo Water and Sewer Department and Christian Almestad/ DHI are greatly acknowledged. The authors are thankful for MIKE+ course arranged by Berislav Tomicic/DHI and support with data collection by Ryan Walter Murray/ DHI.

References

- [1] N. S. V. Lund, A. K. V. Falk, M. Borup, H. Madsen and P. S. Mikkelsen, "Model predictive control of urban drainage systems: A review and perspective towards smart real-time water management," *Critical Reviews in Environmental Science and technology*, vol. 48, no. 3, pp. 279-339, 2018.
- [2] M. Schuetze and J. Alex, "Real time control of sewer systems -a convenient approach to MPC in practice," in *15th International Conference on Urban Drainage*, Online, 25.-29.10.2021.
- [3] Y. Wang and F. A. Haugen, "Averaging Level Control for Urban Drainage System," in *62nd International Conference of Scandinavian Simulation Society (SIMS2021)*, Online, 21-23.9.2021.
- [4] P. A. Stentoft, L. Vezzaro, P. S. Mikkelsen, M. Grum, T. Munk-Nielsen, P. Tychsen, H. Madsen and R. Halvgaard, "Integrated model predictive control of water resource facilities and sewer systems in a smart grid: example of full-scale implementation in Kolding," *Water Science & Technology*, vol. 81, no. 8, pp. 1766-1777, 2020.
- [5] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *Journal of Process Control*, vol. 13, pp. 291-209, 2003.
- [6] D. E. Seborg, T. F. Edgar, D. A. Mellichamp and F. J. I. Doyle, *Process Dynamics and Control*, 4th EMEA ed., Hoboken: Wiley, 2017, p. 512.

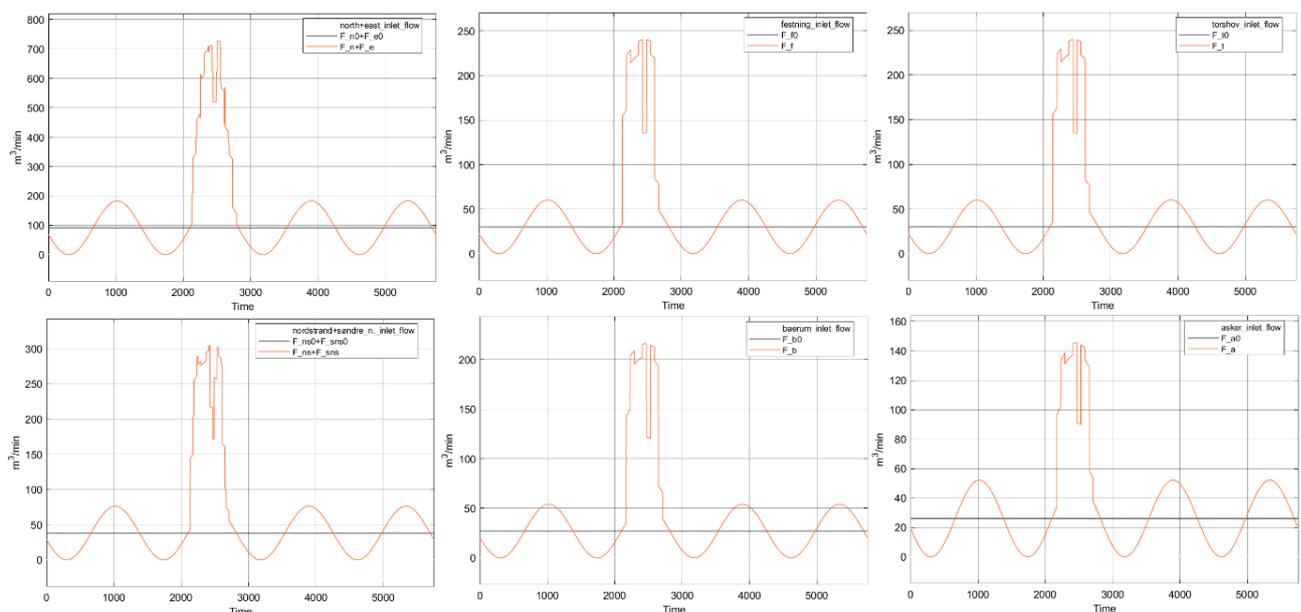


Figure 4: Rain event 4 with combined sewer inlet flowrates I1 North+East; I2:Festning; I3: Torshov; I4: Nordstrand; I5:Bærum I6: Asker. The orange line presents the simulated sewer system inlet flowrate (combined municipal wastewater flow with diurnal pattern and rain event) around the average daily value of municipal wastewater flow indicated with black line.

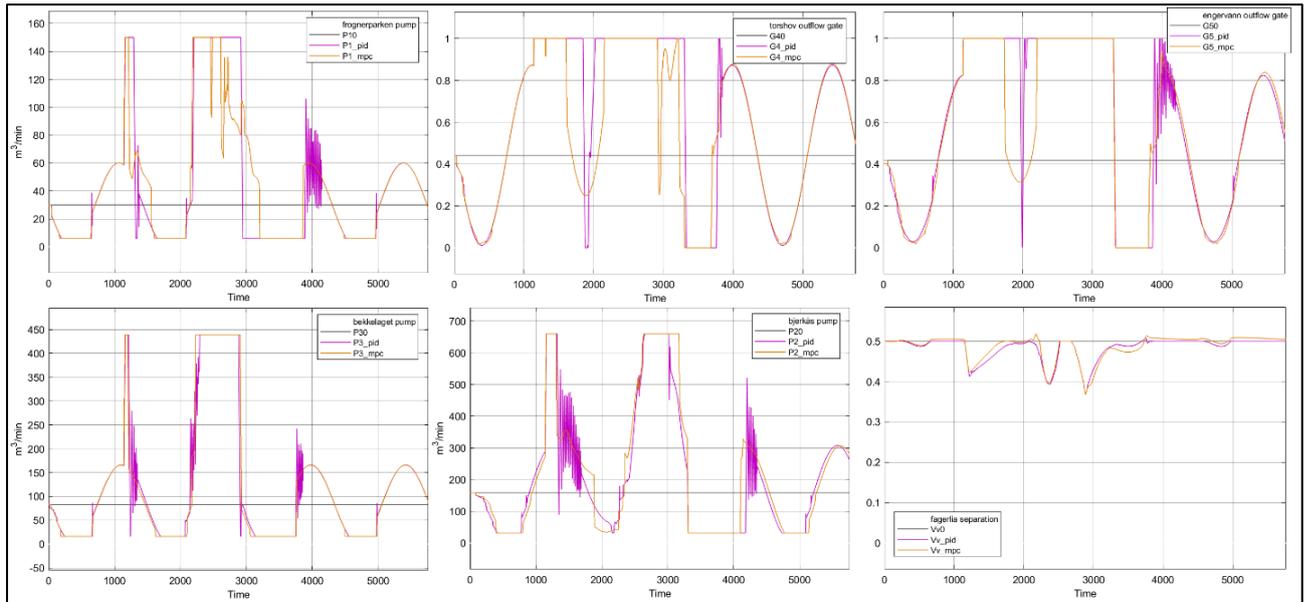


Figure 5: Rain event 4 with changes in manipulated variables in Frognerparken magazine, Torshov magazine, Lysaker magazine, Bekkelaget magazine, Veas magazine and Fagerlia separation using the feedforward strategy combined with PI-controllers (magenta), MPC algorithms (yellow). The nominal MV value is indicated with black line.

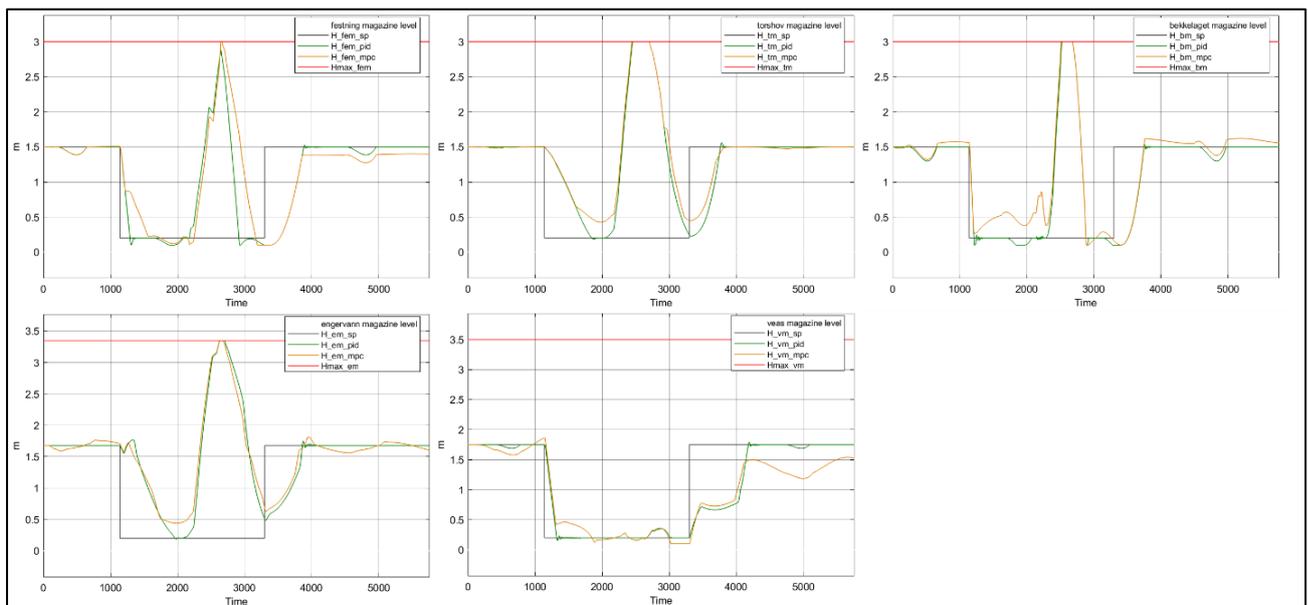


Figure 6: Rain event 4, controlled variables (levels) in Frognerparken magazine, Torshov magazine, Lysaker magazine, Bekkelaget magazine and Veas magazine using the feedforward strategy combined with PI-controllers (magenta), MPC algorithms (yellow). The setpoint is indicated with black line and the high-limit with red line. If the high level is exceeded, wastewater flows over to natural waterways.

Simulation of the Effect of Local Electric Potential and Substrate Concentration on CO₂ Reduction via Microbial Electrosynthesis

Vafa Ahmadi*, Carlos Dinamarca

Department of Process, Energy and Environmental Technology, University of Southeastern Norway
Vafa.Ahmadi@usn.no

Abstract

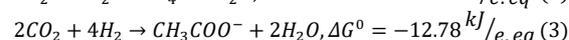
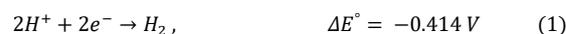
Integrating anaerobic digestion into electrochemical reactors is an advanced technology for biomethane recovery. Imposing low electric potential between electrodes, supplies CO₂, electrons, and hydronium ions from anodic oxidation of organic and/or inorganic compounds. Then, autotrophic methanogens on the cathode produce methane from CO₂ and H⁺ by electron uptake from the cathode. However, in mixed microbial environments, acetogens produce acetate as well. These reactions can take place via two different mechanisms, DIET (direct interspecies electron transfer) or IMET (indirect mediated electron transfer). This work investigates CO₂ conversion to acetate and methane in an electrochemical biofilm reactor comparing the efficiency of CO₂ reduction via DIET and IMET mechanisms at hydrogen evolving potentials from -0.3 to -0.7 vs SHE. The other goal is to prove the importance of mass balance in CO₂ reduction at applied voltages. Simulations are done in AQUASIM version 2.1. Simulation results depicted that higher H⁺ concentration at -0.7 V vs SHE can reduce more CO₂ in DIET with less current generation compared to IMET. This shows DIET the more efficient mechanism. Methane production is dominant in IMET model, however higher current is needed for CO₂ fixation in this mechanism. Also, biomass concentration, acetate and methane production, substrate concentration, biofilm thickness, biomass distribution in biofilm, and current density over time in both mechanisms are investigated at variant voltages and substrate concentrations. Simulations showed that at high CO₂ levels in both mechanisms CO₂ conversion cannot reach maximum if the voltage is not high enough to supply H⁺.

Keywords: biofilm reactor, cathode, DIET, IMET, CO₂ conversion, methane, acetate

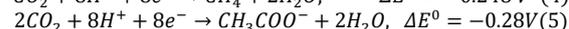
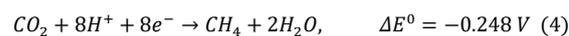
1. Introduction

Microbial electrosynthesis systems (MES) can solve the limitations of anaerobic digestion which have gained attention as power to gas technology; PtG in recent years (Nelabhotla et al., 2021). Biogas normally contains 50-70% CH₄, 50-30% CO₂ and other trace elements. Microbial communities can be stimulated to higher biogas production by a slight increase in the redox potential of the microbial environment. Applying low electric potentials between electrodes to execute electron transfer from anodic to cathodic biofilm has several advantages. Low voltage in the range of microbial redox activities triggers microorganisms to produce more biogas for a longer period and convert more CO₂ to methane. likewise, CH₄ content in biogas reaches up to 90-98%. This could happen when autotrophic microbes namely methanogens and mediator-producing microbes as acetogens contribute to CO₂ reduction via consumption of the available hydrogen. In these systems, CO₂ and hydrogen could be products of anodic dissociation of organic compounds such as complex carbohydrates or

inorganic substances such as ammonium (Nelabhotla & Dinamarca, 2018; Sivalingam et al., 2020). Autotrophic methanogenesis and acetogenesis (equations 1 to 3) could take place via indirect mediated electron transfer (IMET)¹:



Or via equations 4 and 5 corresponding to direct interspecies electron transfer (DIET) (Nelabhotla & Dinamarca, 2019):



In a microbial electrosynthesis system, both mechanisms are possible depending on the microbial species types. According to equation 1, IMET needs higher applied potentials to happen

¹ Voltage values in the paper are stated versus standard hydrogen electrode (vs, SHE)

compared to DIET. This makes IMET less energy efficient since higher potentials are necessary to supply electrons and H^+ for H_2 formation. However, if the inflow to the MES reactor is fully digested, it has low volatile fatty acids (VFAs), but high amounts of hardly degradable organic compounds and soluble CO_2 . Such systems are dependent on hydrogen evolving potentials to provide H^+ for CO_2 fixation. It is important the potentials must be lower than the voltage required for water electrolysis in single chamber MES to avoid oxygen formation that is toxic for anaerobes. Electroactive hydrogen producing bacteria can produce hydrogen via IMET (Gharbi, et al., 2022). Hydrogenotrophic methanogens are of this kind which can produce hydrogen and consume it again for CO_2 fixation to methane (Berghuis et al., 2019). Tremblay et al. (Tremblay et al., 2019), reported that in an enriched medium with microbiome *S. Ovata* known as a hydrogen producing species, a gradual increase in the cathodic voltage from -0.3 to -0.7 vs, SHE, increased microbial H_2 evolution in the MES system. In another work, acetate production happened parallel to methane and hydrogen gas formation in a mixed microbial broth of chemolithoautotrophs at -0.9 V vs SHE on cathode (Bajracharya et al., 2015). Although theoretical reactions in equations 1 to 5 show a lower voltage, the local potential reported for hydrogen gas evolution, acetate and CH_4 formation is higher than the theoretical values. The reason is to overcome potential losses in the MES reactor which depend on factors such as the feed, microbial medium, electrode material, electrode surface area and the reactor volume. This study simulates autotrophic CO_2 conversion to CH_4 and acetate via DIET and IMET mechanisms which depend highly on the microbial community in the reactor. Not much particular work is available to compare DIET and IMET mechanisms. Nevertheless, in a modelling performed by Storck et al, DIET is suggested more advantageous for microbes because of fewer thermodynamic barriers (Storck et al., 2016). It is difficult to control microbial communities, however if the more efficient mechanism is found, the process operation can be controlled to increase the microbial species that demand less energy for CO_2 capture i.e, lower electron flow requirement for the process.

2. Methodology

The simulation is based on a single chamber biofilm reactor. Anode and cathode are in the same microbial medium. The reactor is fully mixed and has continuous inflow and discharge. Figure 1 shows the scheme of the model.

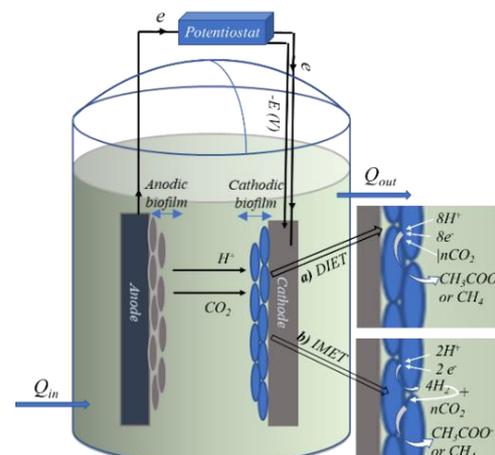


Figure 1. A schematic overview of a single chamber anaerobic MES reactor. The external electric potential on the cathode leads to bio anodic oxidation of organic or inorganic components and generation of CO_2/H^+ that transfer in the liquid toward the cathodic biofilm, and electrons which flow through the wire connection from anode to cathode. Cathodic biofilm ingests the substrates together with electrons to produce acetate ($n=2$) or methane ($n=1$) via two possible mechanisms: a) DIET or b) IMET.

2.1. Model assumptions and expressions

Assumptions of the model are as follows:

- The reactor operates at 1 atm and pH 7.
- Physiochemical gas/liquid mass transfer is not included. Diffusion and convection are assumed between liquid and biofilm.
- Effect of electrode material is not incorporated.
- Only cathodic reactions and cathodic biofilm is included in the simulated model.
- The source of H^+ for cathode is limited. H^+ values in the simulation are taken from experimental work based on anodic oxidation of organic compounds corresponding to a certain applied potential. However, anodic reactions are not included in the simulations.
- CO_2 is the carbon source for cathode which is soluble and in equilibrium with HCO_3^- .
- The active biomass is attached to cathode. Collaboration of the detached biofilm and suspended media to redox reactions is ignored.
- Initial biomass fraction is equal for all species. In DIET model, biomass consists of acetogens and methanogens. In IMET model, hydrogenotrophs are added to the other two species.
- The detachment velocity in the simulation (Reichert, 1998) is assumed as an indicator of bacterial decay (K_d) and loss of biofilm due to operating conditions.

In IMET model, hydrogenotrophic microbes produce hydrogen molecule at cathode. So, the electron acceptor in Monod expression is H^+ ion. Cathode performs as electron donor in the Nernst

term (equation 6). Autotrophic methanogens and acetogens take part in H₂ consumption (equations 7 and 8) which is expressed by multiplicative Monod equation. Cathodic reactions are the opposite of anode. therefore, the Nernst term is positive to represent cathode where electrons are taken from cathode and the voltage is negative (Flowers & Theopold, 2019; Metcalf et al., 2014; Torres et al., 2007).

$$\frac{d[X_{H_2}]}{dt} = X_{H_2} \cdot \left(\mu_{X_{H_2}}^{max} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} \cdot \frac{1}{1 + \exp\left[\frac{2F}{RT}(E_{app})\right]} - k_{d,H_2} \right) \quad (6)$$

$$\frac{d[X_{CH_4}]}{dt} = X_{CH_4} \cdot \left(\mu_{X_{CH_4}}^{max} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H_2}}{K_{H_2} + S_{H_2}} - k_{d,CH_4} \right) \quad (7)$$

$$\frac{d[X_{ac}]}{dt} = X_{ac} \cdot \left(\mu_{X_{ac}}^{max} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H_2}}{K_{H_2} + S_{H_2}} - k_{d,ac} \right) \quad (8)$$

In DIET, microbes take electrons directly from cathode and consume H⁺ and CO₂ from the bulk liquid. H⁺ and CO₂ are both the limiting electron acceptor substrates. Therefore, Nernst-Monod equations in DIET can be expressed as in equations 9 and 10.

Product concentration in IMET and DIET is calculated according to equations 11 and 12.

$$\frac{d[X_{CH_4}]}{dt} = X_{CH_4} \cdot \left(\mu_{X_{CH_4}}^{max} \cdot \frac{1}{1 + \exp\left[\frac{8F}{RT}(E_{app})\right]} \cdot \frac{S_{CO_2}}{K_{CO_2} + S_{CO_2}} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} - k_{d,CH_4} \right) \quad (9)$$

$$\frac{d[X_{ac}]}{dt} = X_{ac} \cdot \left(\mu_{X_{ac}}^{max} \cdot \frac{1}{1 + \exp\left[\frac{8F}{RT}(E_{app})\right]} \cdot \frac{S_{CO_2}}{K_{CO_2, HCO_3^-} + S_{CO_2}} \cdot \frac{S_{H^+}}{K_{H^+} + S_{H^+}} - k_{d,ac} \right) \quad (10)$$

$$\frac{d[S_{CH_4}]}{dt} = \left(\frac{d[X_{CH_4}]}{dt} \right) / Y_{CH_4} \quad (11)$$

$$\frac{d[S_{ac}]}{dt} = \left(\frac{d[X_{ac}]}{dt} \right) / Y_{ac} \quad (12)$$

Change in *j* (current density, A·m⁻²) over time correlates to electroactive biomass concentration, obtained by equation 13 (Torres et al., 2008).

$$\frac{dj}{dt} = \frac{d[X_i]}{dt} \cdot \gamma \cdot L_f \cdot (f_s^0 - 1) \quad (13)$$

Table 1. Model parameters for the simulation

Parameter	Symbol	Value	Unit	Ref
Specific respiration rate of microorganisms	b_X	0.05	d ⁻¹	[a]
Diffusivity of acetate	D_{ac}	$1.54 \cdot 10^{-5}$	m ² · d ⁻¹	[b]
Diffusivity of methane	D_{CH_4}	$1.296 \cdot 10^{-4}$	m ² · d ⁻¹	[b]
Diffusivity of CO ₂	D_{CO_2}	$1.658 \cdot 10^{-4}$	m ² · d ⁻¹	[b]
Diffusivity of H ₂	D_{H_2}	$3.88 \cdot 10^{-5}$	m ² · d ⁻¹	[b]
Diffusivity of hydrogen ion (H ⁺)	D_{H^+}	$8.04 \cdot 10^{-5}$	m ² · d ⁻¹	[b]
Diffusivity of biomass	D_X	$1 \cdot 10^{-7}$	m ² · d ⁻¹	[c]
Biomass density	ρ	222	mol · m ⁻³	[c]
Half saturation concentration of CO ₂	K_{CO_2}	3.8	mol · m ⁻³	[d]
Half saturation concentration of H ₂	K_{H_2}	$8 \cdot 10^{-4}$	mol · m ⁻³	[b]
Max growth rate of methanogens	$\mu_{X_{CH_4}}^{max}$	2.28	d ⁻¹	[f]
Max growth rate of acetogens	$\mu_{X_{ac}}^{max}$	1.008	d ⁻¹	[e]
Max growth rate of Hydrogenotrophs	$\mu_{X_{H_2}}^{max}$	2.2	d ⁻¹	Assumed*1
Acetogenic growth yield	Y_{CH_4}	$6.8 \cdot 10^{-3}$	-	[f]
Methanogenic growth yield	Y_{ac}	$6 \cdot 10^{-3}$	-	[e]
Hydrogenotrophic growth yield	Y_{H_2}	$6.4 \cdot 10^{-3}$	-	Assumed
Boundary layer resistance	LL	$1 \cdot 10^{-4}$	M	[c]
Applied potential on cathode	E_{app}	Variant	V	[g]
Biofilm reactor volume	V	0.1	m ³	[c]
Cathodic biofilm surface area	A	10	m ²	[c]
Energy production of acetogenic cells	$f_{S,ac}^0$	0.067	-	Calculated*2 [h]
Energy production of methanogenic cells	f_{S,CH_4}^0	0.07	-	Calculated [h]
Stoichiometric coefficient of biomass in acetogenesis	-	0.0232	-	Calculated [h]
Stoichiometric coefficient of biomass in methanogenesis	-	0.028	-	Calculated [h]

[a] (Kazemi et al., 2015), [b] (Picioreanu et al., 2010), [c] (Reichert, 1998), [d] (Cabau-Peinado et al., 2021), [e] (Vandecasteele, 2016), [f] (Muñoz-Tamayo et al., 2019), [g] (Tremblay et al., 2019), [h] (Rittmann & McCarty, 2001).

*1. Different values are reported for hydrogenotrophic microbes. Here, the growth rate and yield are assumed based on hydrogenotrophic methanogens. (Berghuis et al., 2019; De Silva Muñoz et al., 2010).

*2. The biomass formula is C₅H₇O₂N (Metcalf et al., 2014).

Where, γ is the equal electron production, that for acetate and methane is 8, and for hydrogen is 2.

2.2. Simulation parameters, inputs, and outline

Simulation parameters are given in Table 1. The choice of cathodic voltage and relevant H^+ concentration was established from experimental work by Tremblay et al. where H_2 evolution depends on cathodic electric potential at ambient temperature $25^\circ C$ (Tremblay et al., 2019). H^+ concentration from the experiment is scaled up to the reactor volume in the simulations. H^+ concentration depends highly on voltage, so at each step, H^+ concentration varies at the same time with voltage change at constant CO_2 concentration. The objective is to calculate the reduced CO_2 at each voltage and H^+ concentration. Also, picture the importance of H^+/H_2 concentration in CO_2 conversion efficiency. All varying steps is executed for both DIET and IMET mechanism to find the more efficient mechanism based on the defined model parameters and inputs. Furthermore, the effect of CO_2 concentration is simulated at three different levels at constant H^+ concentration and constant voltage to prove that controlling CO_2 concentration according to the available H^+ at the corresponding voltage is important to have the highest CO_2 reduction efficiency. Table 2 shows the simulation inputs.

Table 2. Substrate concentration and voltage inputs to the cathodic biofilm for each step.

Step	E_{app} (V vs SHE)	S_{H^+} ($mol \cdot m^{-3}$)	S_{CO_2} ($mol \cdot m^{-3}$)
1	-0.3	60	25
2	-0.4	70	25
3	-0.5	80	25
4	-0.6	120	25
5	-0.7	160	25
6	-0.7	160	10, 25, 50

According to the inflow and the reactor volume, simulation is done at 1 day HRT for 400 days for step 1 to 5. For step 6, CO_2 concentration increases every 100 days for total 300 days.

Simulations include the change in state variables including concentration of acetate (S_{ac}), methane (S_{CH_4}), CO_2 (S_{CO_2}), H^+ (S_{H^+}), hydrogen (S_{H_2}) and biomass (hydrogen producing microbes (X_{H_2}), methanogens (X_{CH_4}) and acetogens (X_{ac})), current density (j), biofilm thickness (L_f), and the distribution of biomass through the biofilm in IMET and DIET mechanisms at different cathodic voltages and substrate concentration.

3. Results and discussion

3.1. Effect of voltage at constant CO_2 level

Figures 2 and 3 compare the simulated methane and acetate concentration in DIET and IMET models. In DIET (Figure 2), both acetate and methane are formed from day one. However, acetate concentration decreases over time at all hydrogen concentrations while methane production increases until day 300. Then acetate production reaches $0.02 mol \cdot m^{-3} \cdot d^{-1}$ and methane production reaches steady state at $19.95 mol \cdot m^{-3} \cdot d^{-1}$.

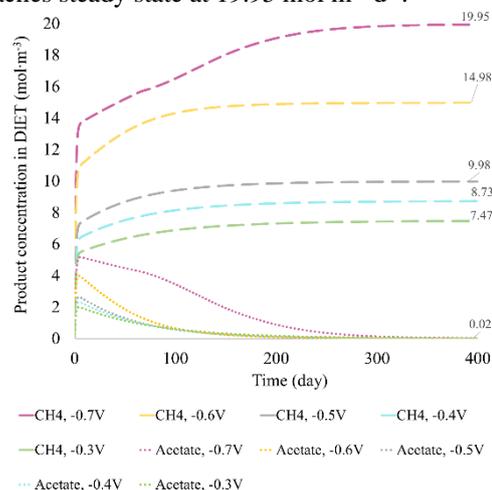


Figure 2. Acetate and methane concentration in DIET model at different potentials.

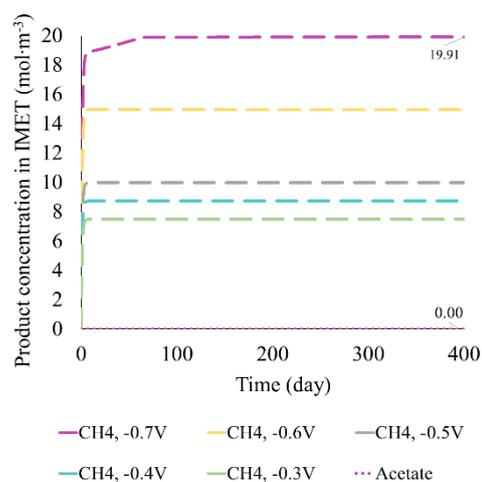


Figure 3: Acetate and methane concentration in IMET model at different potentials.

In IMET (Figure 3) acetogens fail to contribute to hydrogen uptake. Since the third species (hydrogenotrophs) contributes to the IMET model. In this case, diffusion of two CO_2 molecules toward acetogens for acetate formation will be limited. In IMET model, methane attains steady state in the first week at all voltages from -0.3 to $-0.6 V$. However, at $-0.7 V$, it takes 90 days for methane production to reach steady at $19.91 mol \cdot m^{-3} \cdot d^{-1}$. Experimental studies also reported that methane production can be dominated by hydrogenotrophic methanogens via IMET (Gharbi et al., 2022).

However, in DIET model at steady state, the total amount of both products is 1.2% higher than in IMET model. Voltage increment has a positive impact on product formation in both mechanisms especially at -0.7 V which corresponds to higher concentration of H^+ ions, so more CO_2 can be reduced.

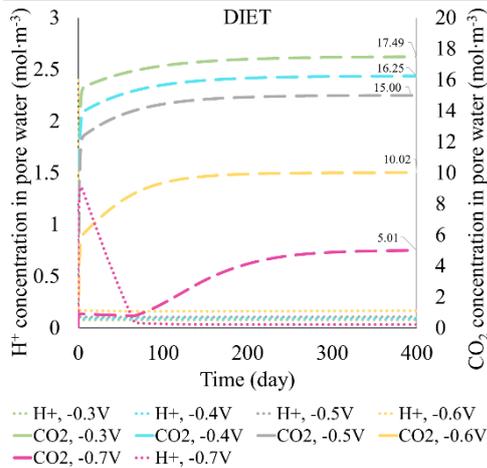


Figure 4. Unconsumed concentration of substrates in pore water in DIET model at different potentials.

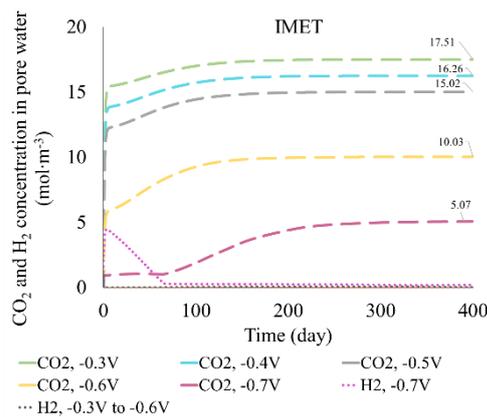


Figure 5: Unconsumed substrate concentration in pore water in IMET model at different potentials.

Figures 4 and 5 show the unconsumed substrate concentration in pore water. In DIET model (Figure 4), CO_2 consumption is slightly higher than IMET (80% for DIET and 79.7% for IMET). In both figures, higher voltage, which supplies more H^+ concentration, increases CO_2 consumption. Also, the highest unconsumed CO_2 in pore water refers to -0.3 V. All H^+ in DIET and all H_2 in IMET are consumed. This can give an idea of a real scenario that increasing CO_2 levels in an MES reactor without supplying enough H^+ , cannot increase product formation. The simulation shows low CO_2 conversion efficiency at -0.3V in both mechanisms. The conversion increases gradually by increasing voltage to -0.7 V due to higher H^+ supply. However, at -0.7 V in the simulation, it takes

longer for the system to stabilize the hydrogen consumption and CO_2 reduction.

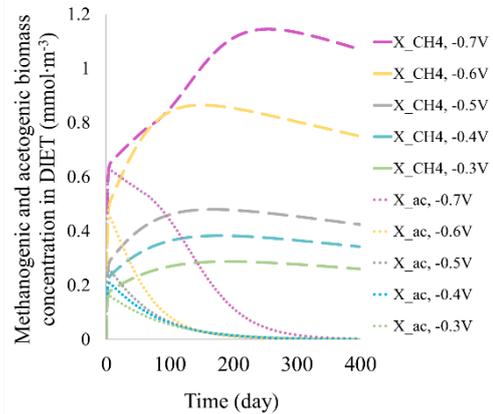


Figure 6. Electroactive biomass concentration in DIET model at different potentials.

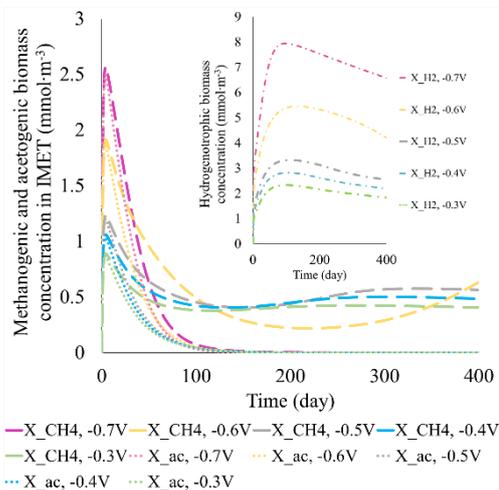


Figure 7. Biomass concentration in IMET model at different potentials.

Figures 6 and 7 show biomass concentration on the cathodic biofilm which in DIET model (Figure 6) consists of electroactive acetogens and methanogens. In IMET model (Figure 7) biomass concentration comprises acetogens, methanogens and electroactive hydrogenotrophs. In DIET, acetogenic and methanogenic biomass concentration change relatively. Finally, methanogens become dominant while acetogenic growth reaches close to zero. Methanogenic biomass attains maximum until day 300 then it follows a slow decreasing trend until day 400. That is due to the biofilm detachment velocity in the simulation. Increasing voltage in the simulation shows a positive effect on microbial growth and biomass concentration especially at -0.6 and -0.7 V. Figure 7 shows in IMET model, the dominant biomass is hydrogenotrophs with a concentration of 3.2 times higher than the total acetogenic and methanogenic biomass after 90 days. There is a

decline in acetogenic and methanogenic concentration while electroactive hydrogenotrophic biomass concentration reaches maximum $8 \text{ mmol}\cdot\text{m}^{-3}$ until day 90. Despite the electroactive biomass, acetogenic concentration attains zero at day 100, and methanogenic biomass concentration becomes stable between $0.5\text{-}1 \text{ mmol}\cdot\text{m}^{-3}$. The decline in biomass may be according to microbial maintenance and detachment velocity parameters included in the simulated biofilm model.

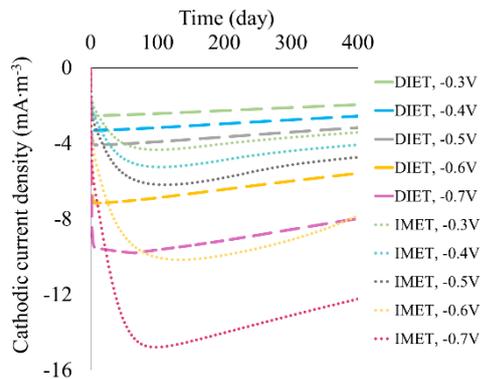


Figure 8. Generated current density in DIET and IMET models at different potentials.

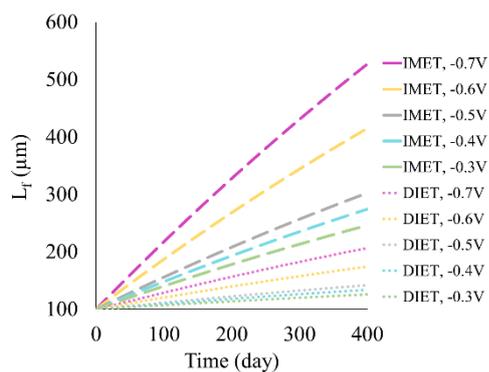


Figure 9. Biofilm thickness in DIET and IMET models at different potentials.

Figure 8 shows current density obtained from DIET and IMET models which is presented as $j\cdot(L_f)^{-1}$ in $\text{mA}\cdot\text{m}^{-3}$ related to electroactive biomass concentration. In DIET model, current density is calculated based on the total current consumed by electroactive acetogens and methanogens. In IMET, it is obtained according to electroactive hydrogenotrophs. In this simulation, the consumed current in DIET is lower than IMET (which corresponds to the total electron flow from anode to cathode). It rises by increasing voltage because of the growing biomass concentration. Considering -0.7 V at steady state, the current density is -7.8 in DIET and $-12.22 \text{ mA}\cdot\text{m}^{-3}$ in IMET model. According to the simulations, DIET can fix 80% of CO_2 at 36% lower required current density. CO_2 conversion efficiency in the simulation with respect to energy consumption is higher in DIET model

(regarding current density and consumed electrons). IMET model shows higher current density. The reason may be due to consuming more electrons and substrates for microbial growth and maintenance.

Figure 9 shows the simulated biofilm thickness. The five lower lines depict DIET, and the five upper lines correspond to IMET mechanisms. In the simulation, biofilm thickness increases by increasing the voltage due to higher biomass formation. However, L_f in DIET is 2.5 times thinner than IMET. Thinner biofilm in DIET at the same voltage can fix marginally higher CO_2 than IMET. In IMET, biofilm is 60% thicker than in DIET which is due to existence of hydrogenotrophs. According to the assumed parameters in the simulation, hydrogenotrophs can grow faster because they grow only on one substrate. So, they become the abundant biomass in IMET and increase the biofilm thickness.

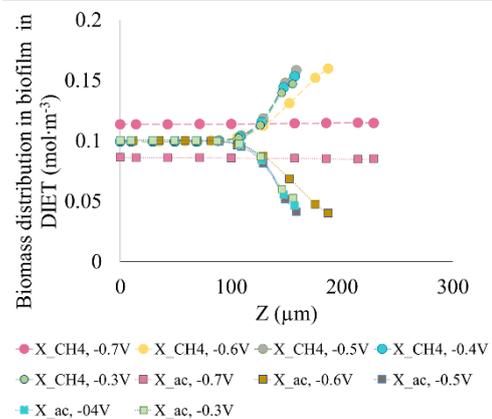


Figure 10. Acetogenic and methanogenic biomass distribution in biofilm in DIET model at different potentials.

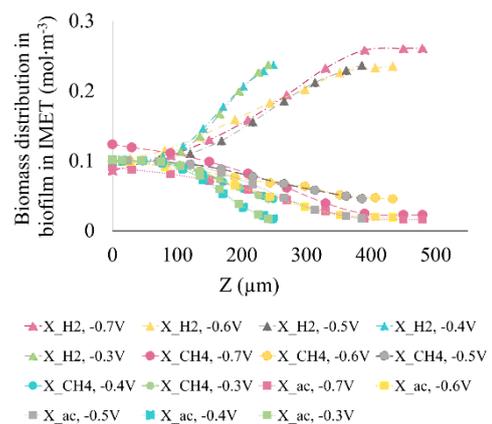


Figure 11. Hydrogenotrophic, acetogenic, and methanogenic biomass distribution through the biofilm in IMET model at different potentials.

Figures 10 and 11 show the simulated biomass distribution after 400 days in DIET and IMET. In

both models, biomass distribution has a direct relation with biofilm thickness. In DIET (Figure 10), distribution of both acetogens and methanogens is equal at the substratum ($Z=0$). The abundance of methanogens becomes higher in outer layers close to the pore water. In contrast, acetogenic community decrease at outer layers. That could be the reason for acetogenic decline in the simulation. The argument could be as the biofilm thickness is low at preliminary stages, acetate can be produced easier since the substrate diffusion is not limited by biofilm thickness. As the biofilm becomes thicker over time, methanogens will be dominant in outer layers which can uptake substrates from pore water faster than acetogens. Therefore, the dominant product is methane in the simulation after 400 days.

In the IMET model (Figure 11) most species in outer layers are hydrogenotrophs which consume all the H^+ to produce H_2 . Compared to that, the abundance of acetogens and methanogens is almost equal at the substratum to the outer layers. However, acetogens grow by consuming two CO_2 molecules. If CO_2 reaches inner layers, methanogen which grow on one CO_2 molecule can consume it easier, so acetogens cannot produce acetate in such simulated conditions.

3.1. Effect of CO_2 concentration

Figures 12 and 13 resulted from run 6, indicate the effect of CO_2 concentration on product formation at a constant voltage of -0.7 V and constant H^+ concentration for both DIET and IMET mechanisms for 100 days each. In DIET model (Figure 12), the whole CO_2 ($10 \text{ mol}\cdot\text{m}^{-3}$) is consumed in the presence of H^+ . Following the stoichiometry of these biological reactions, 62% of H^+ remains unconsumed in pore water. If this happens in a reality, surplus H^+ may decrease the pH, or H^+ may be consumed in other possible reactions which are not desired. Increasing CO_2 concentration in the simulations, leads to full hydrogen consumption while CO_2 consumption efficiency reaches 80%. By further increase in CO_2 concentration, all the hydrogen will be consumed, while 60% of CO_2 remains unconsumed in pore water. If this is a case in real experiments, unused CO_2 may go out with methane due to gas/liquid CO_2 equilibrium and decrease the final methane concentration in biogas (Metcalf et al., 2014). The same trend can be seen in the simulated IMET (Figure 13) that at low CO_2 concentration, H_2 is not fully consumed. If such scenario happens, biogas may contain 82% H_2 and 18% CH_4 which is not desired. Reversely, at high CO_2 levels in the simulation, microorganisms consume all H_2 , but in real cases, excess CO_2 may appear in biogas and

decrease the biogas methane content according to biogas production theory (Metcalf et al., 2014).

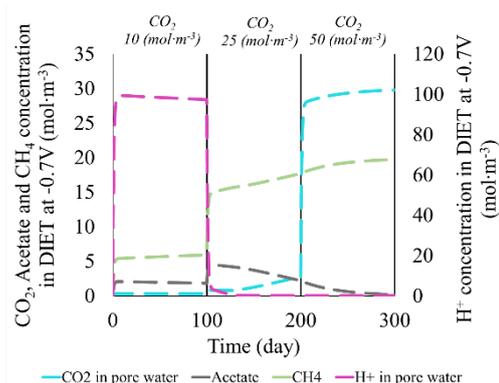


Figure 12. Effect of CO_2 concentration on acetate and CH_4 production in DIET model at $-0.7V$.

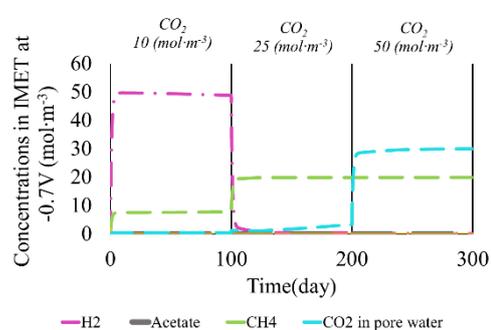


Figure 13. Effect of CO_2 concentration on acetate and CH_4 production in IMET model at $-0.7V$.

4. Summary and conclusions

This work simulates CO_2 reduction to acetate and methane in a continuous flow MES reactor via IMET and DIET mechanisms at H^+ evolving potentials (-0.3 to -0.7 vs SHE) at constant CO_2 concentration, then at constant H^+ concentration and constant voltage. Simulation results show that higher voltage which could provide higher H^+ concentration could convert more CO_2 to methane and acetate in DIET model with 80% CO_2 fixation. Voltage increment could enhance product formation in both DIET and IMET models. Product formation in DIET model is calculated at 36% lower current density which shows it the more energy efficient mechanism compared to IMET model. Moreover, simulations show that controlling CO_2 concentration is another criterion of importance in limited source of H^+ supply.

DIET model implies to be more efficient than IMET model in CO_2 fixation. This means 1.2% higher product formation in DIET at steady state with 36% lower generated current. That means less energy requirement than in IMET mechanism. Not much particular work is available to compare DIET and IMET mechanisms in bioelectrochemical CO_2 fixation. The overall result of this simulation aligns with Stork et al. that DIET needs less energy in

form of electrons for CO₂ fixation. However, the simulations were done assuming simple conditions. Real MES reactors have complex microbial communities which may affect acetogenesis and methanogenesis. Also, intracellular, and extracellular limitations of diverse types of acetogens and methanogens, and their substrate uptake capacity is not studied in these simulations. The yield and maximum growth rate of the assumed microbes in the simulations are found in literature which might overestimate or underestimate the production in DIET or IMET models. Also, detachment velocity and diffusion phenomena are the only limitations included in the simulations. Moreover, there is a non-zero initial value for microbial concentration on the biofilm, so the product formation starts from the first day without a lag phase. However, in real reactors, the lag phase of microbial adaptation to start CO₂ fixation is longer (Metcalf et al., 2014). In these simulations, since H⁺ is available from the first day, microbes start CO₂ reduction from the first day. This simulation studied the cathodic reactions for both mechanisms. However, electron generation source which depends on the available compounds for anode is an important measure of investigation at the next stage. The number of transferred electrons depends on the molecular compounds accessible to anodic biofilm for degradation at a certain voltage. Thus, the model can be completed by including anodic reactions.

References

- Bajracharya, S., ter Heijne, A., Dominguez Benetton, X., Vanbroekhoven, K., Buisman, C. J. N., Strik, D. P. B. T. B., & Pant, D. (2015). Carbon dioxide reduction by mixed and pure cultures in microbial electrosynthesis using an assembly of graphite felt and stainless steel as a cathode. *Bioresour. Technol.*, *195*, 14-24. <https://doi.org/https://doi.org/10.1016/j.biortech.2015.05.081>
- Berghuis, B. A., Yu, F. B., Schulz, F., Blainey, P. C., Woyke, T., & Quake, S. R. (2019). Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens. *Proceedings of the National Academy of Sciences*, *116*(11), 5037-5044. <https://doi.org/doi:10.1073/pnas.1815631116>
- Cabau-Peinado, O., Straathof, A. J. J., & Jourdin, L. (2021). A General Model for Biofilm-Driven Microbial Electrosynthesis of Carboxylates From CO₂ [Original Research]. *Frontiers in Microbiology*, *12*. <https://doi.org/10.3389/fmicb.2021.669218>
- De Silva Muñoz, L., Bergel, A., Féron, D., & Basséguy, R. (2010). Hydrogen production by electrolysis of a phosphate solution on a stainless steel cathode. *International Journal of Hydrogen Energy*, *35*(16), 8561-8568. <https://doi.org/https://doi.org/10.1016/j.ijhydene.2010.05.101>
- Flowers, P., & Theopold, K. (2019). [eTextbook] *Chemistry-2e*. OpenStax.
- Gharbi, R., Vidales, A. G., Omanovic, S., & Tartakovsky, B. (2022). Mathematical model of a microbial electrosynthesis cell for the conversion of carbon dioxide into methane and acetate. *Journal of CO₂ Utilization*, *59*, 101956.
- Kazemi, M., Biri, D., & Rismani-Yazdi, H. (2015). Modelling bio-electrosynthesis in a reverse microbial fuel cell to produce acetate from CO₂ and H₂O. *Physical Chemistry Chemical Physics*, *17*(19), 12561-12574.
- Metcalf, Eddy, Abu-Orf, M., Bowden, G., Burton, F. L., Pfrang, W., Stensel, H. D., Tchobanoglous, G., Tsuchihashi, R., & AECOM. (2014). *Wastewater engineering: treatment and resource recovery*. McGraw Hill Education.
- Muñoz-Tamayo, R., Popova, M., Tillier, M., Morgavi, D. P., Morel, J.-P., Fonty, G., & Morel-Desrosiers, N. (2019). Hydrogenotrophic methanogens of the mammalian gut: Functionally similar, thermodynamically different—A modelling approach. *PLoS one*, *14*(12), e0226243.
- Nelabhotla, A. B. T., & Dinamarca, C. (2018). Electrochemically mediated CO₂ reduction for bio-methane production: a review. *Reviews in Environmental Science and Bio-Technology*, *17*(3), 531-551.
- Nelabhotla, A. B. T., & Dinamarca, C. (2019). Bioelectrochemical CO₂ reduction to methane: MES integration in biogas production processes. *Applied Sciences*, *9*(6), 1056.
- Nelabhotla, A. B. T., Pant, D., & Dinamarca, C. (2021). Power-to-gas for methanation. In *Emerging Technologies and Biological Systems for Biogas Upgrading* (pp. 187-221). Elsevier.
- Picioreanu, C., van Loosdrecht, M. C. M., Curtis, T. P., & Scott, K. (2010). Model based evaluation of the effect of pH and electrode geometry on microbial fuel cell performance. *Bioelectrochemistry*, *78*(1), 8-24. <https://doi.org/https://doi.org/10.1016/j.bioelechem.2009.04.009>
- Reichert, P. (1998). AQUASIM 2.0-Tutorial. *Swiss Federal Institute for Environmental Science and Technology (EAWAG): Dübendorf, Switzerland*.
- Rittmann, B. E., & McCarty, P. L. (2001). *Environmental biotechnology: principles and applications*. McGraw-Hill Education.
- Sivalingam, V., Dinamarca, C., Samarakoon, G., Winkler, D., & Bakke, R. (2020). Ammonium as a Carbon-Free Electron and Proton Source in Microbial Electrosynthesis Processes. *Sustainability*, *12*(8), 3081. <https://www.mdpi.com/2071-1050/12/8/3081>
- Storck, T., Virdis, B., & Batstone, D. J. (2016). Modelling extracellular limitations for mediated versus direct interspecies electron transfer. *The ISME Journal*, *10*(3), 621-631. <https://doi.org/10.1038/ismej.2015.139>
- Torres, C. I., Kato Marcus, A., & Rittmann, B. E. (2007). Kinetics of consumption of fermentation products by anode-respiring bacteria. *Applied microbiology and biotechnology*, *77*(3), 689-697.
- Torres, C. I., Marcus, A. K., Parameswaran, P., & Rittmann, B. E. (2008). Kinetic Experiments for Evaluating the Nernst–Monod Model for Anode-Respiring Bacteria (ARB) in a Biofilm Anode. *Environmental Science & Technology*, *42*(17), 6593-6597. <https://doi.org/10.1021/es800970w>
- Tremblay, P.-L., Faraghiparapari, N., & Zhang, T. (2019). Accelerated H₂ Evolution during Microbial Electrosynthesis with *Sporomusa ovata*. *Catalysts*, *9*(2), 166. <https://www.mdpi.com/2073-4344/9/2/166>
- Vandecasteele, J. (2016). Experimental and modelling study of pure-culture syngas fermentation for biofuels production. *MSc Universiteit Gent*, 356.

Experimental PDE Solver in Julia – Comparison of Flux Limiting Schemes

Amir Farzin^{a,*}, Zahir Barahmand^b, Bernt Lie^a

^a *Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway,*

^b *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway.*
{amir.farzin, zahir.barahmand, Bernt.Lie}@usn.com

Abstract

Finite Volume Methods (FVM) are high quality methods for solving conservative/hyperbolic partial differential equations (PDEs). A popular class of high-resolution methods utilize a nonlinear combination of low order methods and high order methods via flux limiting functions. Another class of high-resolution methods is the class of weighted essentially non-oscillatory (WENO) schemes. Here, the focus is on flux limiting schemes. An experimental finite volume (FV) semi-discrete solver for systems of hyperbolic PDEs has been implemented in Julia, utilizing Julia's DifferentialEquations.jl package for handling the time marching. A first order upwind formulation is used for the low order method, and a central second order formulation is used for the high order method. The PDE can be provided either in flux form, or in quasi-linear form. In the former case, automatic differentiation (AD) package ForwardDiff.jl is used to compute the Jacobians of the flux vector. Package LinearAlgebra.jl is used to compute the eigenspace of the Jacobians. The implementation allows for up to 3 internal/external coordinates. More than a dozen flux limiting functions are given, with the possibility of the users to write their own flux limiters. The implementation allows for user provided spatial discretization points, and source terms in the PDE. In this paper, we will compare various flux limiting schemes for PDEs with analytic solutions, and will also compare flux limiting schemes for a simple granulation model (layering). Possible extensions of the experimental implementation include: (i) higher order methods, (ii) more extensive support for boundary conditions, (iii) improved support for source terms.

1. Introduction

A partial differential equation (PDE) is a mathematical equation having two or more independent variables, an unknown function (depending on those variables), and partial derivatives of the unknown function with respect to the independent variables [1]. Solving a PDE leads to a function that solves the equation or, in other words, converts it into an identity when it is replaced into the equation. Although some variants of PDEs have analytical solutions, in general numerical methods are used to solve PDEs.

There are different types of numerical methods for solving PDEs such as finite elements, finite volumes, and finite difference. In this paper, the finite volumes method (FVM) is considered as the basic for converting PDE problem into a set of ordinary differential equations (ODEs). Then, the ODEs are solved using ODE solver. There are several standard ODE solvers available in almost every programming language. Usually, this approach is called semi-discretization.

FVMs are high quality methods for solving conservative/hyperbolic PDEs. To achieve high-order accurate numerical approximation of PDEs, especially in presence of shocks or discontinuities, a group of FVM related schemes called high-resolution schemes are vastly used. Among the methods, flux limiter and WENO¹ schemes are shown effective in solving difficult-to-solve PDE problems [2, 3].

Flux limiters (or slope limiters) schemes utilize a nonlinear combination of low order methods and high order methods via flux limiting functions. One simple but effective combination of methods is to use upwind as the low (first) order and centered difference as the high (second) order schemes (called MUSCL² scheme) [2]. MUSCL scheme is the core approach in this paper and more than a dozen of flux limiter functions are utilized to solve the PDEs.

This paper focus on the introduction of an open-source solver for Julia programming language. This

¹ Weighted Essentially Non-Oscillatory

² Monotonic Upstream-centered Scheme

solver package is available for the interested reader via the following GitHub link:

https://github.com/amirfarzin/FVM_PDEsolver.jl

In addition, to evaluate the performance of the developed package, several PDE problems are solved using different flux limiter functions. Therefore, as the result, a fair comparison of the functions is also presented in this paper.

The rest of this paper is structured as follows: the second section starts with an introduction to the class of PDEs the solver is design to deal with. Then, the MUSCL scheme equations and flux limiter functions are reviewed. The third section contains a brief documentation for the package. The PDE problems used for evaluation and their solution are presented in section four. In addition, the flux limiter functions are compared based on their performance in solving those PDEs. Finally, the paper ends with discussion in section five.

2. PDE problem definitions

As mentioned, PDEs are equations involving at least two independent variables, one or more dependent variables and their derivatives which are functions the independent variables. The first assumption here, is that time is always an independent variable in all PDEs. Therefore, we denote the variables as follows:

- $z_1, z_2, \dots, z_{n_z}, t$ are independent variables where $n_z \geq 1$. With this notation, the PDE involves $n_z + 1$ independent variables. In vector form, $\mathbf{z} = [z_1, z_2, \dots, z_{n_z}]^T \in \mathbb{R}^{n_z}$ is called the vector of spatial (i.e., internal/external coordination) independent variables. Commonly, the time variable t is in the range $[0, \infty)$ (i.e., $t \in \mathbb{R}^+$). And the spatial variables z_1, z_2, \dots, z_{n_z} are defined in a domain denoted by $\Omega \subset \mathbb{R}^{n_z}$ (i.e., $\mathbf{z} \in \Omega$).
- q_1, q_2, \dots, q_{n_q} are dependent variables where $q_i = q_i(z_1, z_2, \dots, z_{n_z}, t) = q_i(\mathbf{z}, t)$ and $n_q \geq 1$. In vector form, $\mathbf{q} = \mathbf{q}(\mathbf{z}, t) = [q_1, q_2, \dots, q_{n_q}]^T : \mathbb{R}^{n_z+1} \rightarrow \mathbb{R}^{n_q}$ is the vector of unknown functions or dependent variables.

With above definitions, the general form of a PDE problem is as follows:

$$\mathbf{F}(\mathbf{z}, t, \mathbf{q}, D\mathbf{q}, D^2\mathbf{q}, \dots, D^m\mathbf{q}) = \mathbf{0} \quad (1)$$

where $D^\alpha \mathbf{q}$ denotes the tensor of all partial derivatives of order $0 \leq \alpha \leq m$, and $m \geq 1$ is the highest order of partial derivatives in the system. System of Eq. **Error! Reference source not found.**, in general perspective, consists of n_f equations. And the number of unknown variables is n_q .

- if $n_q = n_f$ then the PDE problem is called determined,
- if $n_q < n_f$ or $n_q > n_f$ then the PDE problem is called over-determined or under-determined, respectively.

Without losing generality, it can be assumed that the system of PDEs is determined (i.e., $n_q = n_f$).

The classification of PDE problems, helps easier referring to what type of PDE someone is dealing with. In addition, it is essential to know the class of PDE while using the textbooks and papers, and selecting the solution methodologies. The PDE problems can be classified in four ways:

- **Order of PDE:** The highest order of partial derivatives exists in any of the system equations (i.e., m).
- **Number of equations:** if the number of equations is one (i.e., $n_q = 1$), then the PDE is called scalar. Otherwise, the equations represent a system of PDEs.
- **Linearity:** A PDE system expressed by Eq. (1) is non-linear unless it fits into one of the following groups: (i) quasi-linear³, (ii) semi-linear⁴, or (iii) linear⁵.
- **Homogeneity:** A PDE system represented by Eq. (1) is called homogenous if $\mathbf{v} \in \mathbb{R}^{n_q}$ is a solution to the system then $a\mathbf{v}$ is also a solution for any scalars a .

The package is developed to solve conservation law equation which appears in several physical phenomena such as electromagnetism, fluid dynamics, heat transfer, etc. However, only semi-linear first order system of PDEs with up to three spatial dimensions is considered. For increasing the readability of the notations, let us replace z_1, z_2 , and z_3 with x, y , and z , respectively. Note that, x, y , and z are not necessarily 3-D Cartesian axes, but can represent internal coordination as well.

The differential form of the conservation law is as follows:

$$\partial_t \mathbf{q} + \vec{\nabla} \cdot \vec{\mathcal{F}}(\mathbf{q}) = \sigma(\mathbf{q}) \quad (2)$$

where $\vec{\nabla} = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}]$ is the divergence vector and $\vec{\mathcal{F}} = [\mathbf{f}, \mathbf{g}, \mathbf{h}]$ is the field vector. Equivalently, the Eq. (2) can be written as:

³ A PDE system of order m is called **quasi-linear** if it is linear in terms of the highest order (i.e., m) partial derivatives.

⁴ A quasi-linear PDE system of order m is called **semi-linear** if the coefficients of the highest order (i.e., m) partial derivatives are only functions of the independent variables.

⁵ A PDE is called **linear** if the coefficients of all dependent variables and their derivatives are only functions of the independent variables.

$$\partial_t \mathbf{q} + \partial_x \mathbf{f}(\mathbf{q}) + \partial_y \mathbf{g}(\mathbf{q}) + \partial_z \mathbf{h}(\mathbf{q}) = \boldsymbol{\sigma}(\mathbf{q}) \quad (3)$$

In above equations, $\boldsymbol{\sigma}(\mathbf{q})$ is called the source term. If $\boldsymbol{\sigma}(\mathbf{q}) = \mathbf{0}$, Eq. (3) represent a homogenous system of PDEs. In above equations, \mathbf{f} , \mathbf{g} , and \mathbf{h} are arbitrary functions for field vectors in x , y , and z directions respectively.

If the flux functions are differentiable with respect to conserved variable \mathbf{q} , using the chain rule, the Eq. (3) can be written as follows:

$$\partial_t \mathbf{q} + \mathbf{J}_f \partial_x \mathbf{q} + \mathbf{J}_g \partial_y \mathbf{q} + \mathbf{J}_h \partial_z \mathbf{q} = \boldsymbol{\sigma} \quad (4)$$

The matrices $\mathbf{J}_f = \frac{\partial \mathbf{f}}{\partial \mathbf{q}}$, $\mathbf{J}_g = \frac{\partial \mathbf{g}}{\partial \mathbf{q}}$, and $\mathbf{J}_h = \frac{\partial \mathbf{h}}{\partial \mathbf{q}}$ are the Jacobians of \mathbf{f} , \mathbf{g} , and \mathbf{h} . If all the eigenvalues of the Jacobian matrices are real with linearly independent eigenvectors, the system of PDEs is called hyperbolic [4]. Note that this paper and the solver package only consider hyperbolic PDEs. The solver is design to handle the PDEs expressed either by Eq. (3) or Eq. (4).

3. Methodology

In this section, for the simplicity of equations, the methodology is discussed for 1-D homogenous conservation law equation as follows:

$$\partial_t \mathbf{q} + \partial_x \mathbf{f}(\mathbf{q}) = \mathbf{0} \quad (5)$$

At the end this section, the method is extended for solving PDE problems of Eq. (3) and (4).

3.1. Grid

To solve Eq. (5), the general idea is to convert it into a set of ODE problems where we need the values of $\partial_t \mathbf{q}$ at each time step. As \mathbf{q} is a function of time and space, the values of $\partial_t \mathbf{q}$ at each time step should be found for every point in the space. Here, the space domain is simply bounded to min/max values (i.e., $x \in [x^{\min}, x^{\max}]$).

Numerically, working with a continuous space domain is impossible. Therefore, it is necessary to discretize the space first. The developed package uses a non-uniform rectangular grid up to three dimensions for space discretization. It is also possible to extend it for more dimensions or improve it with triangular cells. A small part of the x axis is shown in Fig 1.

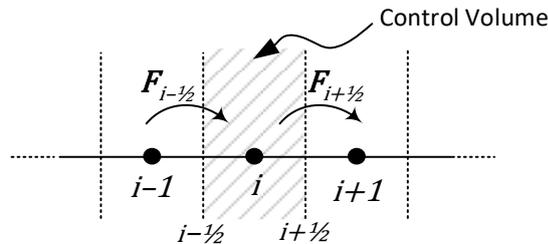


Figure 1: Grid-point cluster for one-dimensional problem

The advantage of non-uniform grid is the variable cell sizes. This feature allows larger cells in non-challenging and smaller cells in critical areas.

Regarding the grid, the following notation is used in the rest of this paper:

- Integer subscripts such as $i - 1$, i , and $i + 1$ are used for quantities at cell centers. For example, x_i and \mathbf{q}_i are the coordination and \mathbf{q} vector at the center point of i th cell.
- Subscripts such as $i - \frac{1}{2}$ and $i + \frac{1}{2}$ are used for referring to the quantities at the cell boundaries. For example, as shown in Fig. 1, $\mathbf{F}_{i-\frac{1}{2}}$ is the flux vector from $(i - 1)$ th cell into i th cell at the boundary between them.
- As mentioned, the grid is assumed to be non-uniform which means the lengths of cells can be different. Therefore, the length of the i th cell is denoted by $\Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$. In addition, $\Delta x_{i-\frac{1}{2}} = x_i - x_{i-1}$ is the distance between center points of $(i - 1)$ th and i th cells.

3.2. Finite volume method (FVM)

The FVM is a popular and efficient numerical approach for solving PDE problem. In the FVM, the problem domain (i.e., only space in semi-discretization strategy) is broken into grid cells. Then, the total integral of \mathbf{q} over each grid cell is approximated [5]. If the problem presents a conservation equation (e.g., Eq. (5)), the value of the integral is equal to the net flow into the control volume:

$$\partial_t \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{q} dx = \mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}} \quad (6)$$

where $\mathbf{F}_{i-\frac{1}{2}}$ is some approximation to the flux along the left boundary of the i th cell:

$$\mathbf{F}_{i-\frac{1}{2}} \approx \mathbf{f}(\mathbf{q})|_{x=x_{i-\frac{1}{2}}} \quad (7)$$

In addition, the cells (called control volumes) are small enough to assume that \mathbf{q} is constant in each cell. Let us say the average value of \mathbf{q} in i th cell is equal to its value at the center of that control volume (i.e., \mathbf{q}_i). Replacing \mathbf{q} with \mathbf{q}_i in Eq. (6) gives:

$$\begin{aligned} \partial_t \mathbf{q}_i \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx &= \mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}} \\ \xrightarrow{\text{yields}} \partial_t \mathbf{q}_i &= \frac{\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}}{\Delta x_i} \end{aligned} \quad (8)$$

Hence, we only need to find an accurate approximation of fluxes at the boundaries to solve the PDE problem using the FVM. The difference

between various schemes is how they estimate $F_{i-\frac{1}{2}}$ at each time step.

3.3. High-resolution flux limiter schemes

One drawback of linear methods like upwind is predicted by Godunov's theorem which states that no linear convection scheme of second-order accuracy or higher can be monotonic [6]. This limit would be significant while working with shock waves with sharp gradients. To tackle the problem, non-linear discretization methods are used.

One category of such non-linear approaches is flux limiters (FL) which uses a limiter function for the flux in the boundaries. A unified formulation for flux limiters is presented in [7] which is used here with some manipulations.

As mentioned in previous section, to solve Eq. (5), we need to find an approximation for fluxes at the boundaries (i.e., $F_{i-\frac{1}{2}}$ and $F_{i+\frac{1}{2}}$). A flux limiter scheme gives the following formulae for calculating the boundary fluxes [8]:

$$\begin{cases} F_{i+\frac{1}{2}}^{FL} = F_{i+\frac{1}{2}}^{low} - \psi(r_i) \left(F_{i+\frac{1}{2}}^{low} - F_{i+\frac{1}{2}}^{high} \right) \\ F_{i-\frac{1}{2}}^{FL} = F_{i-\frac{1}{2}}^{low} - \psi(r_{i-1}) \left(F_{i-\frac{1}{2}}^{low} - F_{i-\frac{1}{2}}^{high} \right) \end{cases} \quad (9)$$

Here, $\Psi(r)$ is called the flux limiter function, F^{low} is the low-resolution flux (i.e., the flux calculated based on a low order scheme) and F^{high} is the high-resolution flux (i.e., the flux calculated based on a high order scheme). The value of r at point i is given by:

$$r_i = \left(\frac{\partial q}{\partial x} \right)_{i-\frac{1}{2}} / \left(\frac{\partial q}{\partial x} \right)_{i+\frac{1}{2}} \quad (10)$$

If the grid is uniform, Eq. (10) reduces to:

$$r_i = \frac{q_i - q_{i-1}}{q_{i+1} - q_i} \quad (11)$$

The package use Eq. (10) which provides the capability of non-uniform grids.

Several methods can be used as the low/high-resolution schemes. Our implementation uses the first order upwind (UW) and central difference (CD) as low-resolution and high-resolution schemes, respectively:

$$\begin{cases} F_{i-\frac{1}{2}}^{low} = F^{UW} \left(q_{i-\frac{1}{2}} \right) \\ F_{i-\frac{1}{2}}^{high} = F^{CD} \left(q_{i-\frac{1}{2}} \right) \end{cases} \quad (12)$$

One promising issue about the UW-CD combination is that it allows a generalization method called κ -scheme to model linear schemes as flux limiters [9].

The UW scheme considers the flow direction when determining the flux value at a boundary [10]: the

flow at a boundary is calculated based on the value of conserved variable in the upstream cell (i.e., the cell that corresponds the flow:

$$\begin{cases} F^{UW+} \left(q_{i-\frac{1}{2}} \right) = f(q_{i-1}) \\ F^{UW+} \left(q_{i+\frac{1}{2}} \right) = f(q_i) \end{cases} \quad (13)$$

$$\begin{cases} F^{UW-} \left(q_{i-\frac{1}{2}} \right) = f(q_i) \\ F^{UW-} \left(q_{i+\frac{1}{2}} \right) = f(q_{i+1}) \end{cases} \quad (14)$$

The superscripts + and - are used for distinguishing between positive and negative flux direction.

In contrast, the CD does not consider the flux direction and simply uses the average value of adjacent cells to calculate the flux at a boundary:

$$\begin{cases} F^{CD} \left(q_{i-\frac{1}{2}} \right) = f \left(\frac{q_{i-1} + q_i}{2} \right) \\ F^{CD} \left(q_{i+\frac{1}{2}} \right) = f \left(\frac{q_i + q_{i+1}}{2} \right) \end{cases} \quad (15)$$

For hyperbolic problems the information propagates with finite speed as determined by the eigenvalues of the flux Jacobian [5]. Therefore, the positive eigenvalues of A corresponds to positive fluxes (i.e., left to right) and vice versa [11]. This suggests to separate the positive and negative fluxes using the concept of eigenvalues. Decomposition of the Jacobian matrix (i.e., A) in terms of eigenvalue matrix ($\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{n_q})$) and right eigenvector matrix (V) yields:

$$AV = VA \quad (16)$$

As the PDE problem is assumed hyperbolic, the eigenvector matrix V is nonsingular, so that:

$$A = V\Lambda V^{-1} \quad (17)$$

Let us define the positive and negative eigenvalues as follows:

$$\lambda_k^+ \triangleq \max(\lambda_k, 0) = \frac{1}{2}(\lambda_k + |\lambda_k|) \quad (18)$$

$$\lambda_k^- \triangleq \min(\lambda_k, 0) = \frac{1}{2}(\lambda_k - |\lambda_k|) \quad (19)$$

for $1 \leq k \leq n_q$. Now, using above values, positive and negative eigenvectors are constructed as follows:

$$\Lambda^+ \triangleq \begin{bmatrix} \lambda_1^+ & 0 & \dots & 0 \\ 0 & \lambda_2^+ & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n_q}^+ \end{bmatrix} \quad (20)$$

$$\mathbf{A}^- \triangleq \begin{bmatrix} \lambda_1^- & 0 & \dots & 0 \\ 0 & \lambda_2^- & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n_q}^- \end{bmatrix} \quad (21)$$

Finally, the positive and negative Jacobians are defined as:

$$\mathbf{A}^+ \triangleq \mathbf{V}\mathbf{A}^+ \mathbf{V}^{-1}, \quad \mathbf{A}^- \triangleq \mathbf{V}\mathbf{A}^- \mathbf{V}^{-1} \quad (22)$$

It can easily be shown that:

$$\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^- \quad (23)$$

Now, to put all above equation together, we need to change the right-hand side of Eq. (8) as follows:

$$\frac{\mathbf{F}_{i-\frac{1}{2}} - \mathbf{F}_{i+\frac{1}{2}}}{\Delta x_i} = - \left(\frac{\partial \mathbf{f}(\mathbf{q})}{\partial x} \right)_i = -(\mathbf{A} \partial_x \mathbf{q})_i$$

$\xrightarrow{\text{yields}} \quad \partial_t \mathbf{q}_i = -(\mathbf{A} \partial_x \mathbf{q})_i \quad (24)$

Replacing Jacobian from Eq. (23) in Eq. (24) yields:

$$\partial_t \mathbf{q}_i = -\mathbf{A}_i^+ \partial_x \mathbf{q}_i^+ - \mathbf{A}_i^- \partial_x \mathbf{q}_i^- \quad (25)$$

where $\partial_x \mathbf{q}^+$ and $\partial_x \mathbf{q}^-$ are defined as:

$$\partial_x \mathbf{q}_i^+ = \frac{\mathbf{q}_{i+\frac{1}{2}}^+ - \mathbf{q}_{i-\frac{1}{2}}^+}{\Delta x_i} \quad (26)$$

$$\partial_x \mathbf{q}_i^- = \frac{\mathbf{q}_{i+\frac{1}{2}}^- - \mathbf{q}_{i-\frac{1}{2}}^-}{\Delta x_i} \quad (27)$$

Using UW-CD combination the values of $\mathbf{q}_{i\pm\frac{1}{2}}^{-/+}$ are as follows:

$$\mathbf{q}_{i+\frac{1}{2}}^- = \mathbf{q}_i + \psi(r_i) \left(\frac{\mathbf{q}_{i+1} - \mathbf{q}_i}{2} \right) \quad (28)$$

$$\mathbf{q}_{i-\frac{1}{2}}^- = \mathbf{q}_{i-1} + \psi(r_{i-1}) \left(\frac{\mathbf{q}_i - \mathbf{q}_{i-1}}{2} \right) \quad (29)$$

$$\mathbf{q}_{i+\frac{1}{2}}^+ = \mathbf{q}_{i+1} - \psi(r_i) \left(\frac{\mathbf{q}_{i+1} - \mathbf{q}_i}{2} \right) \quad (30)$$

$$\mathbf{q}_{i-\frac{1}{2}}^+ = \mathbf{q}_i - \psi(r_{i-1}) \left(\frac{\mathbf{q}_i - \mathbf{q}_{i-1}}{2} \right) \quad (31)$$

3.4. Generalization

In previous subsection, the methodology derived for solving Eq. (5). However, solving PDE problem of Eq. (3) is straightforward. As we need the values of $\partial_t \mathbf{q}$ at each time step, we can apply the same procedure explained for $\partial_x \mathbf{f}(\mathbf{q})$, on $\partial_y \mathbf{g}(\mathbf{q})$ and $\partial_z \mathbf{h}(\mathbf{q})$. In addition, in semi-linear problems, the source term depends only on the values of \mathbf{q} and x . Therefore, at each time step, the value of the source term can easily be calculated. This means, the methodology is applicable for solving Eq. (5)

without any further manipulation. The complete semi-discretization formula (i.e., generalization of Eq. (25)) is as follows:

$$\begin{aligned} \partial_t \mathbf{q}_{ijk} = & \sigma(\mathbf{q}_{ijk}) - \mathbf{A}_{ijk}^+ \partial_x \mathbf{q}_{ijk}^+ - \mathbf{A}_{ijk}^- \partial_x \mathbf{q}_{ijk}^- \\ & - \mathbf{B}_{ijk}^+ \partial_y \mathbf{q}_{ijk}^+ - \mathbf{B}_{ijk}^- \partial_y \mathbf{q}_{ijk}^- \\ & - \mathbf{C}_{ijk}^+ \partial_z \mathbf{q}_{ijk}^+ - \mathbf{C}_{ijk}^- \partial_z \mathbf{q}_{ijk}^- \end{aligned} \quad (32)$$

where \mathbf{B} and \mathbf{C} are Jacobians of flux functions \mathbf{B} and \mathbf{C} , respectively. In addition, in case of problems given as Eq. (4), the Jacobians are already given which reduce a discretization step for computing them numerically.

3.5. Flux limiter functions

Several limiter functions can be found in books and scientific papers. The interested reader may refer to [7], which reviews a wide range of limiter functions in a unified approach. A list of the limiter functions implemented in the package given in Table 1. For each limiter function, its syntax in the package and the formula for $\Psi(r)$ is provided. The first six schemes are linear; while the rest of them are non-linear schemes.

Note that, the developed package has the capability to accept any user-defined limiter function.

Table 1: List of limiter functions

Syntax	Name	Flux limiter formula
uw1	First-order upwind	$\Psi(r) = 0$
uw2	Second order upwind	$\Psi(r) = 1$
uw3	Cubic upwind	$\Psi(r) = \frac{2}{3}r + \frac{1}{3}$
uw4	Quadratic upwind	$\Psi(r) = \frac{3}{4}r + \frac{1}{4}$
scd	Second-order central difference	$\Psi(r) = r$
fr	Fromm	$\Psi(r) = \frac{1}{2}r + \frac{1}{2}$
kn	Koren	$\Psi(r) = \max \left[0, \min \left(2r, \min \left(\frac{1+2r}{3}, 2 \right) \right) \right]$
sb	Superbee	$\Psi(r) = \max[0, \min(2r, 1), \min(r, 2)]$
mm	Minmod	$\Psi(r) = \max[0, \min(r, 1)]$
mu	MUSCL	$\Psi(r) = \max \left[0, \min \left(2r, \frac{r+1}{2}, 2 \right) \right]$
ha	Harmonic	$\Psi(r) = \frac{r+ r }{r+1}$
va1	van Albada 1	$\Psi(r) = \frac{r(r+1)}{r^2+1}$
va2	van Albada 2	$\Psi(r) = \frac{2r}{r^2+1}$
vl	van Leer	$\Psi(r) = \frac{r+ r }{1+ r }$
op	OSPRE	$\Psi(r) = \frac{3r(r+1)}{2(r^2+r+1)}$
hc	HCUS	$\Psi(r) = \frac{1.5(r+ r)}{r+2}$
hq	HQUICK	$\Psi(r) = \frac{2(r+ r)}{r+3}$
cm	CHARM	$\Psi(r) = \max \left[0, \frac{r(3r+1)}{(r+1)^2} \right]$
mc	Monotonized central	$\Psi(r) = \max[0, \min(2r, 0.5(r+1), 2)]$
sm	Smart	$\Psi(r) = \max[0, \min(2r, (0.75r+0.25), 4)]$
um	UMIST	$\Psi(r) = \max[0, \min(2r, (0.75r+0.25), (0.25r+0.75), 2)]$

4. Developed package

The package “FVM_PDEsolver.jl” is developed for Julia programming language to solve first-order semi-linear hyperbolic PDE systems using high-resolution flux limiters. It is capable of handling PDEs given as Eq. (3) and Eq. (4). The package uses the semi-discretization method and calculates the right-hand side of Eq. (32). Then, it uses ODE solvers from “DifferentialEquations.jl” to find the final solution of the PDE system. If the equation is in the form of Eq. (4), then the Jacobians of flux functions are calculated numerically using “ForwardDiff.jl” package. To calculate the eigenvalues and eigenvectors, “LinearAlgebra.jl” package is used.

4.1. Algorithm

The algorithm is exactly what was discussed as the flux limiter scheme in the previous section. Here, a detailed description of the algorithm is presented:

Step 0: Initialize the grid values for $t = 0$: the current values of \mathbf{q} vector is set to the initial values provided by the user.

Step 1: Calculate $\frac{\partial \mathbf{q}}{\partial t}$ for each cell in current time step based on the current values of \mathbf{q} are as follows (expressed for 1-D case⁶):

- 1 Calculate $\left(\frac{\partial \mathbf{q}}{\partial x}\right)_{i+\frac{1}{2}} = \frac{q_{i+1} - q_i}{x_{i+1} - x_i}$ at all cells boundaries
- 2 Calculate $r_i = \left(\frac{\partial \mathbf{q}}{\partial x}\right)_{i-\frac{1}{2}} / \left(\frac{\partial \mathbf{q}}{\partial x}\right)_{i+\frac{1}{2}}$ for each cell.
- 3 Calculate $q_{i\mp\frac{1}{2}}^{+/-}$ using Eq. 28 to 31 at all cell boundaries.
- 4 Calculate $\partial_x q_i^+ = \frac{q_{i+\frac{1}{2}}^+ - q_{i-\frac{1}{2}}^+}{\Delta x}$ and $\partial_x q_i^- = \frac{q_{i+\frac{1}{2}}^- - q_{i-\frac{1}{2}}^-}{\Delta x}$ for each cell.
- 5 Calculate the Jacobian matrices for each cell:
- 5-1 If the equation is in form **Error! Reference source not found.** then by the given function (i.e., $A_i = J_f(q_i)$).
- 5-2 If the equation is in form **Error! Reference source not found.** then by using numerical method (i.e., $A_i = \left. \frac{\partial f(q)}{\partial q} \right|_{q=q_i}$).
- 6 Calculate the eigenvalues and eigenvectors matrices numerically for each cell (i.e., A_i and V_i).
- 7 Calculate positive/negative eigenvalue matrices for each cell (i.e., $A_i^+ = \frac{1}{2}(A_i + |A_i|)$ and $A_i^- = \frac{1}{2}(A_i - |A_i|)$ where absolute sign denotes the elementwise absolute value).
- 8 Calculate positive and negative Jacobian matrices for each cell (e.g., $A_i^+ = V_i \Lambda_i^+ V_i^{-1}$ and $A_i^- = V_i \Lambda_i^- V_i^{-1}$).
- 9 Calculate $S_i = S(q_i)$ for each cell.
- 10 Calculate $\frac{\partial q_i}{\partial t} = -(A_i^+ \partial_x q_i^+ + A_i^- \partial_x q_i^-) + S_i$

Step 2: Use the ODE solver to solve $\frac{\partial \mathbf{q}}{\partial t} = \dots$ find \mathbf{q} for the next time step.

Step 3: Set the current values of \mathbf{q} to the values found in step 1.

Step 4: Re-do step 1 to step 3 until reaching t_{end} .

5. Evaluation of package performance

⁶ For 3-D case, the index i should be changed to ijk and the steps 1-8 should be done for y and z axes as well. Step 10 also calculates Eq. (32).

In this section, some example PDE problems are defined and solved using the package.

5.1. Moving wave with constant velocity

A simple step like wave which propagate with constant speed in a 1-D space ($x \in [0,1]$) is defined by the following PDE:

$$\frac{\partial q}{\partial t} + a \frac{\partial q}{\partial x} = 0 \quad (33)$$

with initial condition:

$$q(t = 0, x) = \begin{cases} 1.0 & x \leq 0.5 \\ 0.0 & x > 0.5 \end{cases}$$

where $a = 0.5$ is the propagation speed in x direction. As $a > 0$, the exact solution at each time step is defined by shifting the wave position by at in x direction. Assume we want to find the solution at $t = 0.2$. To solve the problem a uniform grid with 0.01 intervals is used. Then, the problem is solved by the package and the results using uw1, uw2, scd, kn, and sb schemes are shown in Figure 2.

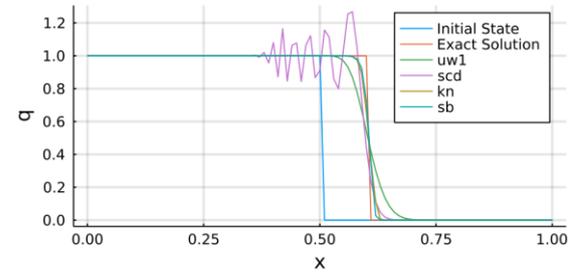


Figure 2: Solution of moving wave with constant velocity problem using different schemes

As expected, the worst scheme is the scd; while the best is sb in this example. The same problem is also solved using all nonlinear schemes available in the package. For better illustration, the values in range $[0.5, 0.7]$ are plotted in Figure 3.

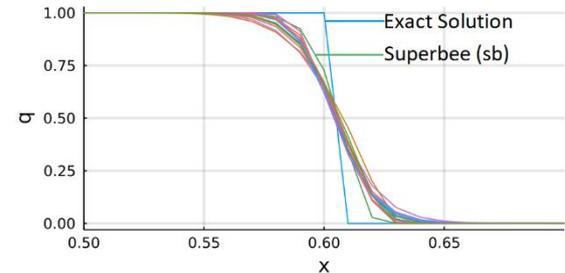


Figure 3: Solution of moving wave with constant velocity problem using all available nonlinear flux limiters

Although the plot in Figure 3 is a little nasty, it shows all nonlinear flux limiter functions have similar performances on this particular problem. However, the sb method shows a slightly better performance than others.

5.2. Euler problem: rarefaction-shock case

The Euler equations are a collection of nonlinear hyperbolic conservation rules that regulate the dynamics of compressible fluids while ignoring the effects of body forces and viscous stress [12]. The 1-D Euler equation is expressed by Eq. (5); where \mathbf{q} and $\mathbf{f}(\mathbf{q})$ are given as follows:

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho u \\ \rho e \end{bmatrix}, \quad \mathbf{f}(\mathbf{q}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho u h \end{bmatrix} \quad (34)$$

where: $h = e + \frac{p}{\rho}$, $p = \rho(\gamma - 1)\left(e - \frac{1}{2}u^2\right)$

In above equations, ρ is the density, p is the pressure, u is the velocity in x direction, e is the internal energy, h is the static enthalpy, and γ is the ratio of specific heats which is a constant scalar. Although $\mathbf{f}(\mathbf{q})$ is not expressed explicitly in terms of \mathbf{q} , it is possible to reformulate it doing some maths:

$$\mathbf{f}(\mathbf{q}) = \begin{bmatrix} q_2 \\ \frac{q_2^2}{q_1} + q_1(\gamma - 1)\left(\frac{q_3}{q_1} - \frac{1}{2}\left(\frac{q_2}{q_1}\right)^2\right) \\ q_2\left(\frac{q_3}{q_1} + (\gamma - 1)\left(\frac{q_3}{q_1} - \frac{1}{2}\left(\frac{q_2}{q_1}\right)^2\right)\right) \end{bmatrix} \quad (35)$$

Assume the domain $x \in [0,1]$ and following initial conditions:

$$\rho(x, t = 0) = \begin{cases} \rho_L & 0.0 \leq x < x_0 \\ \rho_R & x_0 < x \leq 1.0 \end{cases}$$

$$u(x, t = 0) = \begin{cases} u_L & 0.0 \leq x < x_0 \\ u_R & x_0 < x \leq 1.0 \end{cases}$$

$$p(x, t = 0) = \begin{cases} p_L & 0.0 \leq x < x_0 \\ p_R & x_0 < x \leq 1.0 \end{cases}$$

Using the following parameters, the PDE problem represents a rarefaction-shock case:

$$x_0 = 0.5, \quad \Gamma = 1.4,$$

$$\rho_L = 1.0, \quad \rho_R = 0.125,$$

$$u_L = 0.0, \quad u_R = 0.0,$$

$$p_L = 1.0, \quad p_R = 0.1$$

It is desired to find the solution at $t = 0.2$. The exact analytical solution is given in [13]. To solve the problem a uniform grid with 0.01 intervals is used. Figure 4 and Figure 5 shows the solution using uw1 and kn schemes, respectively.

5.3. Shallow water equation: supercritical case

The 1-D shallow water equation is as follows [14]:

$$h_t + (hu)_x = 0$$

$$(hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x = -ghB_x \quad (36)$$

Here, g is the gravitational constant, h is the water depth, u is the mean velocity in x direction, and $B(x)$ is the waterway bottom topography. Although

the source term in Eq. (36) has a differential term, as the topography is not changing, the values of B_x are constant and can be pre-calculated.

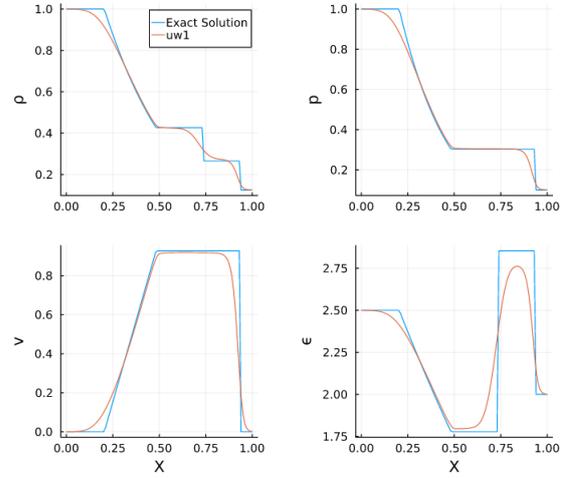


Figure 4: Solution of Euler problems: rarefaction-shock case using uw1 scheme

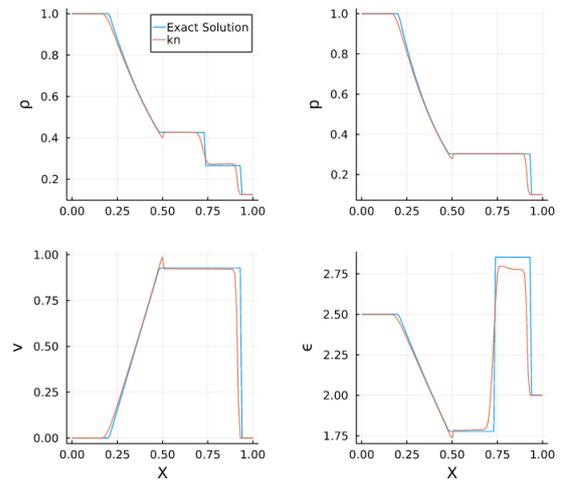


Figure 5: Solution of Euler problems: rarefaction-shock case using kn scheme

To match Eq. (36) with Eq. (3), the following equations should be considered:

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} h \\ hu \end{bmatrix}$$

$$\mathbf{f}(\mathbf{q}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \end{bmatrix} = \begin{bmatrix} \frac{q_2}{q_1} \\ \frac{q_2^2}{q_1} + \frac{1}{2}gq_1^2 \end{bmatrix} \quad (37)$$

$$\boldsymbol{\sigma}(\mathbf{q}) = \begin{bmatrix} 0 \\ -ghB_x \end{bmatrix} = \begin{bmatrix} 0 \\ -gq_1B_x \end{bmatrix}$$

In the definition of $\boldsymbol{\sigma}(\mathbf{q})$, g is a constant scalar and B_x is a known function of x . So, the Eq. (36) is semi-linear and the package is capable of solving it. Note that, in some contexts the Eq. (36) is referred to as the dam-break wave equation.

Assume the problem domain $x \in [-10,10]$ and the following initial conditions:

$$h(x, t = 0) = \begin{cases} h_L & -10 \leq x < 0 \\ h_R & 0 < x \leq 10 \end{cases}$$

$$u(x, t = 0) = \begin{cases} u_L & -10 \leq x < 0 \\ u_R & 0 < x \leq 10 \end{cases}$$

$$B(x) = \begin{cases} B_L & -10 \leq x < 0 \\ B_R & 0 < x \leq 10 \end{cases}$$

The following parameter setting makes the supercritical case:

$$g = 9.8, \quad h_L = 4.0, \quad h_R = 1.0,$$

$$u_L = -10.0, \quad u_R = -6.0, \quad B_L = 0.0, \quad B_R = 1.0$$

It is desired to find the solution at $t = 0.2$. Although the exact solution is not available, a solution to this problem can be found in [14]. To solve the problem a uniform grid with 0.05 intervals is used. Figure 6 shows the solution using kn schemes.

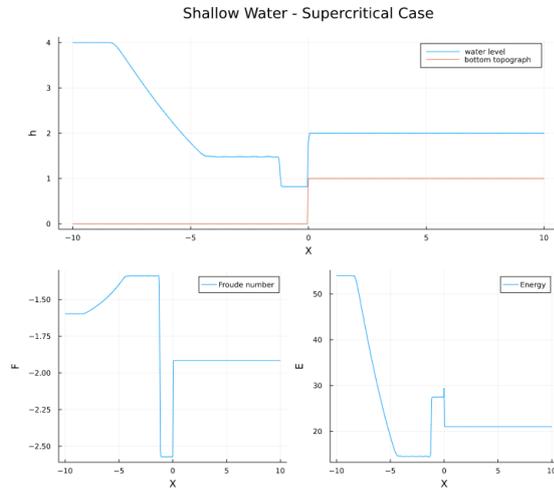


Figure 6: Solution of Shallow water equation: supercritical case using kn scheme

Comparing the result with those presented in [14] approve the performance of the package.

5.4. Granulation process: growth by layering

The initiative of this work was the granulation process modeling and finding the numerical solution for the growth by layering equation. The equation according to [15] is as follows:

$$\frac{\partial}{\partial t} n_b(V_p, t) + G(t) \frac{\partial}{\partial V_p} n_b(V_p, t) = 0 \quad (38)$$

where, V_p is the internal coordination of particles volumes, $G(t)$ is the growth rate, and n_b is the particle size distribution function. Eq. (38) represents a square wave moving with speed $G(t)$. Therefore, the for $G(t) = cte$. the exact solution is available.

As the Jacobian is given outside of the derivative term, this equation is in the form of Eq. (4). Assume the problem domain $V_p \in [0, 400]$ and the following initial conditions:

$$n_b(V_p, t = 0) = \begin{cases} 10 \times 10^4 & 15 \leq x < 50 \\ 0 & \text{otherwise} \end{cases}$$

Given $G(t) = 1$, the solution at $t = 150$ is desired. To solve the problem a uniform grid with $\Delta V_p = 5$ is used. Figure 7 shows the solution using uw1, scd, and kn schemes.

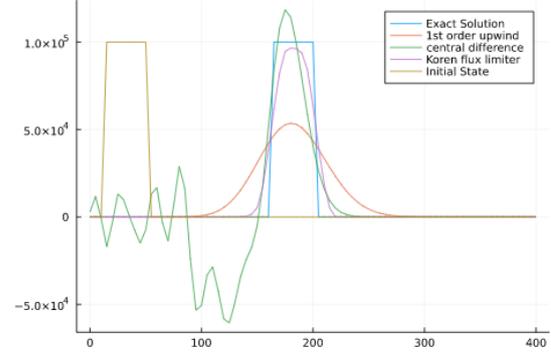


Figure 7: Solution of granulation process: growth by layering problem using different schemes

Comparing the results with the same schemes used in [15], approve the performance of the package. One difference between our results and those given in [15] is that they limited the values of n_b to be always non-negative. This is not done in our solution which caused the scd scheme find negative values for n_b . This issue will be resolved in further improvements of the package.

6. Summary and Discussions

The PDE problems are challenging to solve. High-resolution schemes provide high-order accurate numerical approximation of PDEs, especially in presence of shocks or discontinuities. In this paper, flux limiter schemes, which are among the high-resolution methods, was reviewed in detail. The derived equations made the basis for development of a package named “FVM_PDEsolver.jl” in Julia programming language. The algorithm of solving mechanism is explained and its performance is shown in several examples. The package is an open-source program available in GitHub. The authors aim to improve the package in many aspects in future. Some ideas for improvement are: triangular grids, limiting the quantities’ values, adding more schemes, etc. One interesting observation in this paper was the high accuracy of Superbee flux limiter which was the best among all available limiters in the package.

References

- [1] A. D. Polyanin, W. E. Schiesser, and A. I. Zhurov, “Partial differential equation,” *Scholarpedia*, vol. 3, no. 10, pp. 4605, 2008.
- [2] B. Van Leer, “Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov’s method,” *Journal of*

computational Physics, vol. 32, no. 1, pp. 101-136, 1979.

[3] C.-W. Shu, “Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws,” *Advanced numerical approximation of nonlinear hyperbolic equations*, pp. 325-432, 1998.

[4] C. Hirsch, *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics*, pp. 105-140: Elsevier, 2007.

[5] R. J. LeVeque, *Finite volume methods for hyperbolic problems*: Cambridge university press, 2002.

[6] S. Godunov, and I. Bohachevsky, “Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics,” *Matematičeskij sbornik*, vol. 47, no. 3, pp. 271-306, 1959.

[7] N. P. Waterson, and H. Deconinck, “Design principles for bounded higher-order convection schemes—a unified approach,” *Journal of Computational Physics*, vol. 224, no. 1, pp. 182-207, 2007.

[8] J. Pietrzak, “The Use of TVD Limiters for Forward-in-Time Upstream-Biased Advection Schemes in Ocean Modeling,” *Monthly Weather Review*, vol. 126, no. 3, pp. 812-830, 1998.

[9] P. L. Roe, “Approximate Riemann solvers, parameter vectors, and difference schemes,” *Journal of computational physics*, vol. 43, no. 2, pp. 357-372, 1981.

[10] H. K. Versteeg, and W. Malalasekera, *An introduction to computational fluid dynamics: the finite volume method*, 2nd ed., pp. 134-178: Pearson education, 2007.

[11] C. Hirsch, *Numerical computation of internal and external flows: Computational Methods for Inviscid and Viscous Flows*, pp. 408-492: Wiley, 1990.

[12] E. F. Toro, *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*: Springer Science & Business Media, 2013.

[13] F. Lora-Clavijo, J. Cruz-Pérez, F. Siddhartha Guzmán, and J. Gonzalez, “Exact solution of the 1D riemann problem in Newtonian and relativistic hydrodynamics,” *Revista mexicana de física E*, vol. 59, no. 1, pp. 28-50, 2013.

[14] S. Jin, and X. Wen, “An efficient method for computing hyperbolic systems with geometrical source terms having concentrations,” *Journal of Computational Mathematics*, pp. 230-249, 2004.

[15] L. Vesjolaja, B. Glemmestad, and B. Lie, “Solving the population balance equation for granulation processes: particle layering and agglomeration,” 2021.

Comparison and Application of Multi-Rate Methods for Real-Time Simulations of Production Systems

Lars Klingel^{a,*}, Valentin Kamm^a, Alexander Verl^a,

^a*Institute for Control Engineering of Machine Tools and Manufacturing Units (ISW)*

University of Stuttgart

Seidenstraße 36, 70174 Stuttgart, Germany

**lars.klingel@isw.uni-stuttgart.de*

Abstract

A distributed simulation makes it possible to couple simulation tools and lays the foundation for the usage of multicore capabilities to decrease the calculation time. In consequence the simulation is partitioned on multiple simulation tasks. If simulation tasks with different integration step sizes are used, the configuration is called multi-rate simulation. In real-time simulations the tasks are calculated parallelly, which means that fast tasks do not wait for the simulation result of slower tasks. A sequential approach where fast tasks wait for slower tasks would slow down the overall simulation and therefore tear the real-time requirements. In a real-time multi-rate approach, signal processing of the coupling signals between the tasks is required. For this signal processing, multi-rate methods are used. Easy multi-rate methods lead to stepped signals in the faster task, because the slower task does not provide a new calculated signal at each timestep of the faster task. In this work further methods are investigated in an industrial real-time simulation environment. The analysis contains continuous and discontinuous as well as energy conserving methods. It is shown how these different methods perform for various kinds of signals. The methods are compared and evaluated on signals with different characteristics, which allows a recommendation for the choice of a method in a specific simulation scenario. The application of the multi-rate methods is shown on an example virtual commissioning simulation of an industrial robot. It shows that the right choice of a multi-rate method has a big impact on the overall simulation result.

1. Introduction

The increasing digitization in the life cycle of production systems can be summarized under the term digital factory, which contains models, methods and tools with the aim of planning, evaluation and ongoing improvement of the real factory [1]. A method of the digital factory which can help to reduce the development time of production systems and increase their quality is virtual commissioning (VC).

VC can help to detect errors and validate the software in the engineering phase of a production system [2]. VC always contains a simulation model of the production system which interacts with the control system of the machine or plant. VC can be performed using different configurations which differ in the degree of realisation of the control system. In the early phase of the control development model-in-the-loop simulations are used, where the modeled control logic is tested. If the control logic is already available in a programming language for control systems, the code can be run on an emulated controller. The most common and realistic configuration is the hardware-in-the-loop simulation. In a hardware-in-the-loop simulation, the simulation model is connected to the real control system of the machine or plant through a fieldbus. With this structure real-time requirements arise for the simulation because the simulation must be calculated in the same cycle time as the deterministic fieldbus communication and the control system. Pritschow and Röck present an approach for

the architecture of a hardware-in-the-loop simulation tool which meets the real-time requirement by calculating the simulation on a real-time operating system on one processor core [3].

By using only one processor core on a real-time operating system, two main restrictions are imposed: It is not possible to a) calculate complex simulation models due to the limited performance and b) no other simulation models or tools can be coupled. To counter this problem a real-time co-simulation has been introduced by Scheifele [4]. In [4] a real-time capable coupling and synchronisation between simulation tasks is described. A block diagram based model can be partitioned on several simulation tasks which are calculated in a jacobi sequence. That means the tasks are calculated parallelly and only exchange signals at dedicated coupling times [5] in order to maintain the real-time capability. In Figure 1 the real-time co-simulation is shown with an example block diagram, which is partitioned on two tasks. The approach is not limited to a specific number of tasks and can be extended as needed.

In a co-simulation it often occurs that the simulation tasks are calculated with different cycle times, which is called a multi-rate simulation. This is exemplary shown in Figure 1 where the upper task is calculated with the cycle time h_1 and is three times as fast as the lower task with cycle time h_2 . This can be the case if a computationally intense model, like a physical simulation is integrated into the

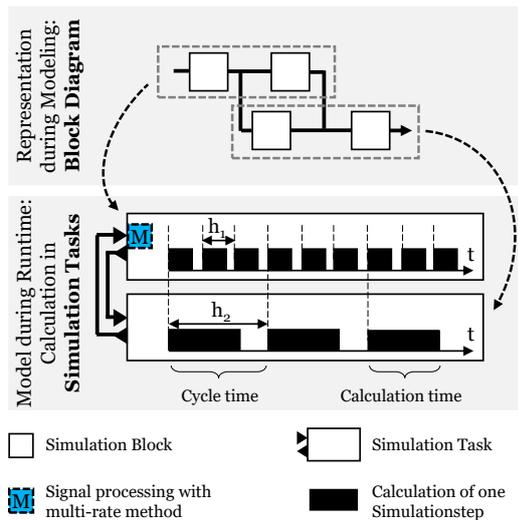


Figure 1: A block diagram based simulation model is partitioned on two simulation tasks and calculated in a jacobi sequence in real-time.

simulation model. In a VC it is also possible that the model's calculation time is too slow for a coupling with the fieldbus. In this case the fieldbus emulation can be calculated with a shorter cycle time than the rest of the model. No matter why the model is partitioned on tasks with different cycle times, a slow simulation task cannot provide a new calculated signal at each time step of a faster task. At this point multi-rate methods come into play [6]. In Figure 1 the position where multi-rate methods can be integrated into a real-time co-simulation is shown with a blue box.

In this work, the authors show how multi-rate methods integrated into a real-time industrial simulation environment can improve partitioned simulations by processing the coupling signals. This work is organized as follows: After introducing related work, a concept for integrating multi-rate methods is shown, as well as the implementation of the different methods. In the fourth section, the validation of the multi-rate methods is shown. A use case of an industrial robot simulation and the conclusion as well as an outlook on the future complete this paper.

2. Related Work

In this section an overview of the current state of the art concerning multi-rate methods is given. Basically, there are three approaches for the modification of coupling signals in a distributed multi-rate simulation [7]:

- Standard method: Zero-order-hold
- Application-specific methods
- General methods: Interpolation and extrapolation

Zero-order-hold is the simplest variant, in which the input of the fast system is held until the slow system provides a new value. This approach is also the standard for a real-time co-simulation. The implementation of a zero-order-hold method is simple but the method is subject to jumps in the signal course. The zero-order-hold method shall be used in the following sections as a reference for the evaluation of the investigated multi-rate methods.

Application-specific methods use behavioral models to predict the course of certain signals. In [8] a kalman

filter was used as a multi-rate method for a physics-based material flow simulation. As an example, a cuboid running on a narrowing conveyor belt is simulated. By applying the smoothing filter and boundary conditions, good results are obtained for continuous signals. An automated integration into other coupling signals is described as difficult. In general, application specific methods like a kalman filter are usually not transferable to other simulations models.

The general multi-rate methods can be divided into extrapolation and interpolation. In the case of a real-time co-simulation, due to the parallel operation of the subsystems in a jacobi sequence, no future information of the signals are available. For this reason, interpolation cannot be used. Instead, an extrapolation is necessary, where previous coupling signals are used for the prediction of the further behaviour of the signal. In [9] the hardware-in-the-loop simulation of a conveyor belt is investigated. The real control hardware is coupled to the simulation computer through a fieldbus. The simulation is divided into real-time capable models and the slower physics-based material flow model. The calculation time of a simulation step is due to the complexity of the material flow model up to 40 ms. To enable communication with the control hardware, signals must be provided for the real-time capable solver in 1 ms. The current position and velocity data of the slow model are used to extrapolate new position data for the fast solver. By using the method, the error between the reference signal and the processed signal with the multi-rate method is significantly reduced compared to the standard zero-order-hold method.

Different basic extrapolation models and their convergence behaviors have been investigated in [10]. The multi-rate methods are then successfully applied to the reduction of the simulation time in different areas of simulated vehicle dynamics. Polynomial extrapolation with and without the use of derivative values as well as continuous extrapolation are used. In [11], the overall thermal system of a bus is simulated. By partitioning it into four subsystems, a co-simulation architecture was created. A smoothed extrapolation method was used as a multi-rate method to optimize the system behavior and the simulation speed. An extension of the extrapolation methods by an error correction analogous to control engineering concepts is presented in [12]. Thereby a nearly energy-conserving multi-rate method is obtained. By simulating a driver assistance system, the effect of the developed coupling method is demonstrated.

There are multiple multi-rate methods available [6, 10–12]. Some already have been implemented into a real-time VC simulation [8, 9] but only specific simulation scenarios have been considered. Therefore it is necessary to implement and compare several general methods into a real-time simulation system which can be used for different virtual commissioning simulation scenarios. In the next chapter a concept and implementation of general multi-rate methods is shown.

3. Concept and Implementation

In this section the concept for the integration of the multi-rate methods into an industrial simulation environment is shown. Afterwards different multi-rate implementations are shown which can be classified into:

- Discontinuous multi-rate methods
- Continuous multi-rate methods
- Energy conserving multi-rate methods

3.1. Concept for the Integration of Multi-Rate Methods into Real-Time Simulation Environments

For the integration of multi-rate methods into an industrial simulation environment the approach from [9] is used. In Figure 1 this concept is shown for the example of two tasks which exchange signals. The multi-rate methods are inserted into the fast task to process signals from the slow task. In the following the step size of the slow subsystem is referred to as macro step size and the step size of the fast subsystem is referred to as micro step size. In the next subsection different multi-rate methods are shown. Those methods are implemented to be usable as black boxes which can be parameterized and used for any model.

3.2. Implementation of Multi-Rate Methods

In this section the implemented multi-rate methods are described and visualized on a sine model.

3.2.1. Discontinuous Multi-Rate Methods

For the first set of multi-rate methods a discontinuous approach is chosen where the input is directly carried over to the output at every coupling time (macro step size). In between these communication steps an extrapolation method is executed to approximate the expected output. The most simple attempt for this is to use a polynomial. An efficient way to calculate the polynomials $p(t)$ is Newton's method:

$$p(t) = \sum_{j=0}^n c_j N_j(t). \quad (1)$$

The Newton basis function $N_j(t)$ is given by:

$$N_j(t) = \prod_{i=0}^{j-1} (t - t_i), \quad j = 1, \dots, n. \quad (2)$$

The initial condition is $N_0(t) = 1$. The coefficients $c_j = f[t_0 \dots t_j]$ are calculated using the method of divided differences which is shown for a polynomial of second order in Table 1 in form of a Horner-scheme.

Table 1: Divided differences for a polynomial of second order.

t_j	f_j	$f[t_j, t_{j+1}]$	$f[t_j, t_{j+1}, t_{j+2}]$
t_0	$f[t_0]$		
t_1	$f[t_1]$	$\frac{f[t_1] - f[t_0]}{t_1 - t_0} = f[t_0, t_1]$	
t_2	$f[t_2]$	$\frac{f[t_2] - f[t_1]}{t_2 - t_1} = f[t_1, t_2]$	$\frac{f[t_1, t_2] - f[t_0, t_1]}{t_2 - t_0} = f[t_0, t_1, t_2]$

The order of the extrapolation is determined by the number of used data points n and can be adapted to each problem. Figure 2 shows the theoretical evaluation of the extrapolation using a polynomial of third order compared to the standard zero-order-hold. The blue signal is the reference signal which should be accurately predicted by the multi-rate methods. The only information, the multi-rate methods have about the reference signal is signal value on the past coupling times (shown with a dotted line).

The approximated function is $f(t) = \sin(2\pi t)$ and the step size for the slow subsystem is 40 ms.

If additionally the derivation of the input value is known for every coupling time, the method according to Hermite can be used. Every coupling time there can be used two values for the calculation of the polynomial. Especially in the beginning of the simulation, when fewer values are known than needed for the desired order of extrapolation, the hermite variant leads to more accurate performance.

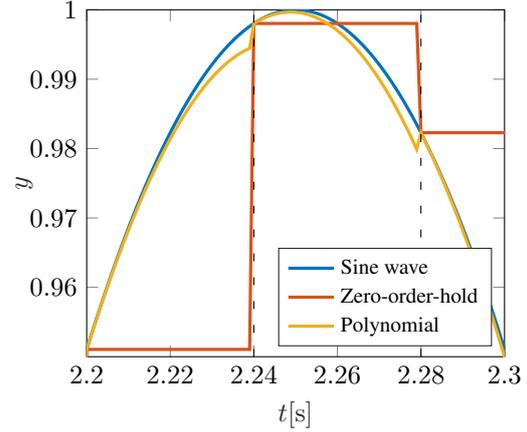


Figure 2: Polynomial of third order compared to the standard zero-order-hold for approximation of a sine. Step size for the slow subsystem is 40 ms.

3.2.2. Continuous Multi-Rate Methods

As can be seen in Figure 2 the discontinuous methods result in small jumps at the coupling time. To avoid this, extrapolation methods are introduced which are continuous and differentiable. For a first approach a combination of extrapolation with interpolation is integrated which has been introduced by Knorr [10]. At every coupling time only the expected value for the next coupling time is extrapolated and saved in an array. The array is then used for interpolation to approximate the value for every step of the fast subsystem. The polynomials needed for the interpolation are calculated according to section 3.2.1. This approach is referred to as 'integrated extrapolation'.

A second approach based on Kossels work [11] is based on making the transition between the previously calculated extrapolation function $f_p(t)$ and the new function $f_n(t)$ calculated at the current coupling time using smoothing. The output from the multi-rate method is defined as follows:

$$f(t) = \begin{cases} g(x(t)) \cdot f_p(t) + (1 - g(x(t))) \cdot f_n(t) & , t_i \leq t < t_s \\ f_n(t) & , \text{else} \end{cases} \quad (3)$$

Switching from the previous to the new extrapolation polynomial happens half way to the next coupling time:

$$t_s = t_i + \frac{h_s}{2}. \quad (4)$$

Here t_s is the time of the switchover and t_i is the time stamp of the previous coupling time. The variable $x(t)$ has to be normalized to guarantee the continuous behavior of equation (3):

$$x(t) = \frac{t - t_i}{t_s - t_i} \quad (5)$$

For the smoothing function $g(x(t))$ then follows that $g(x=0) = 1$ and $g(x=1) = 0$. Additionally to ensure differentiability:

$$g^{(i)}(x=0) = g^{(i)}(x=1) = 0, \quad i = 1, \dots, n \quad (6)$$

If $n = 2$ meaning double differentiability the system of equations is solved to give the following smoothing function:

$$g(x) = -6x^5 + 15x^4 - 10x^3 + 1 \quad (7)$$

Using the smoothing function and the polynomials calculated at every coupling time, equation (3) is used as a multi-rate method to approximate values for each micro step. Figure 3 shows the result for both continuous methods for the same sine as before with the discontinuous approach using the same parameters. As can be seen, the jumps at the coupling times disappear at the expense of some accuracy, which means that the processed signal does not have the exact same value as the reference signal at the coupling times.

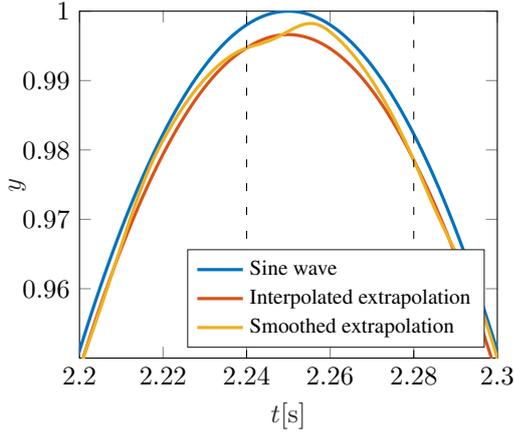


Figure 3: Comparison of the continuous multi-rate methods for approximation of the sine. Degree of extrapolation: $n_E = 3$; degree of interpolation (only interpolated extrapolation): $n = 3$; macro step size: $h_s = 40$ ms.

3.2.3. Energy conserving Multi-Rate Methods

A closed loop approach to multi-rate methods resulting in nearly energy conserving behavior was shown by Benedikt [12]. Following his example the discontinuous as well as the continuous multi-rate methods are enhanced using control loops. The schematic overview of the approach for the discontinuous method is shown in Figure 4.

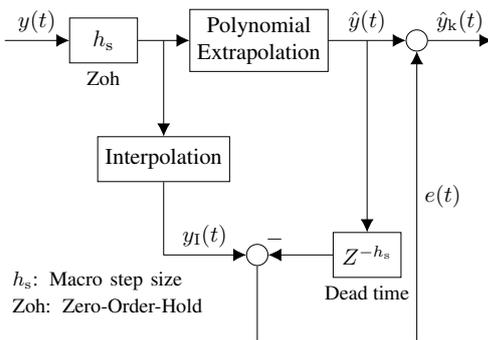


Figure 4: Schematic overview of the integrated closed loop for the extrapolation using polynomials resulting in almost energy conserving behavior.

The slow system generates new output values every h_s seconds which is shown using the zero-order-hold (Zoh). These values are used for extrapolation to generate the estimated signal $\hat{y}(t)$ which after a dead time of the macro step size can be compared to the much more accurate interpolated signal $y_I(t)$:

$$e(t) = y_I(t) - \hat{y}(t) \quad (8)$$

The identified error $e(t)$ is added to the estimated signal to generate the corrected values $\hat{y}_k(t)$ at each micro step. If the macro step size is low enough compared to the frequency of the input signal the closed loop can lead to a significantly improved accuracy as shown in Figure 5.

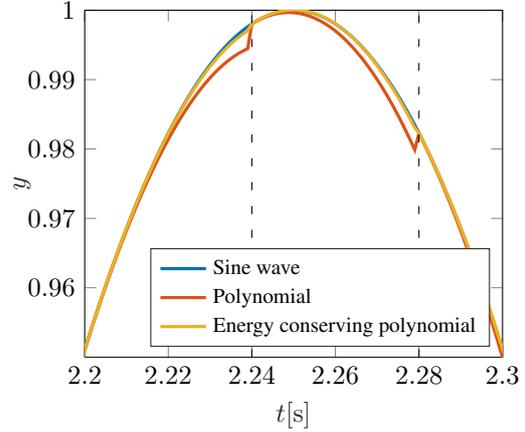


Figure 5: Approximation of the sine using polynomials with and without closed loop control.

Using the energy conserving multi-rate methods while still getting continuous behavior requires a different approach. As a basis the interpolated extrapolation method from section 3.2.2 is used. The schematic overview is shown in Figure 6.

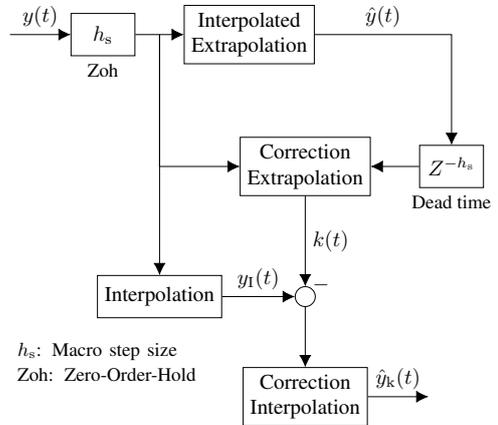


Figure 6: Schematic block diagram for the energy conserving continuous multi-rate method.

Similar to the discontinuous control loop the extrapolated value can be compared to the real value delayed by a dead time of the slow systems step size h_s . The error is added to the following estimation which is then in turn used for the continuous interpolation. In the block diagram this is marked as *correction extrapolation* resulting in the corrected signal $k(t)$. Although this simple approach increases the accuracy of the approximation, it does not provide energy conserving behavior. Therefore the delayed but more accurate interpolation between the actual input values is again computed to be compared to the multi-rate methods estimated interpolation. For every

macro step the individual errors added up to result in the error total E :

$$E(T_i) = \sum_{i=0}^N (y_I(t_i) - k(t_i)) \quad (9)$$

Here N is the number of micro steps:

$$N = \frac{h_s}{h_f} - 1 \quad (10)$$

T_i are the coupling times at the distance of the macro step size h_s whereas t_i are the micro steps in the faster cycle time of h_f . Adding the error total E at the next macro step to the corrected signal $k(t)$ then ensures energy conservation. To avoid losing continuity this summation is done using a quadratic distribution function d which has to be zero at the boundaries of the macro steps:

$$d(t_i) = \frac{E(T_{i-1})}{\sum_{i=0}^N (i(i-N))} \cdot i(i-N) \quad (11)$$

Here i is the counter of the micro steps which is reset after every coupling time. With this procedure, shown in the figure as *correction extrapolation*, the corrected output is obtained:

$$\hat{y}_k(t) = k(t) + v(t) \quad (12)$$

In Figure 7 the result of the correction is compared to the interpolated extrapolation without energy conserving behaviour. As can be seen, the correction leads to an improved accuracy at least for this particular example. Generally the improvement depends on the cycle time of the subsystem compared to the frequency of the input signal as explained for the discontinuous method. For the chosen example signal, the sine wave, that means that the frequency of the sine wave needs to be higher than the cycle time of the macro and micro step size, otherwise the signal may swing up. This phenomenon is based on the Nyquist-Shannon-theorem.

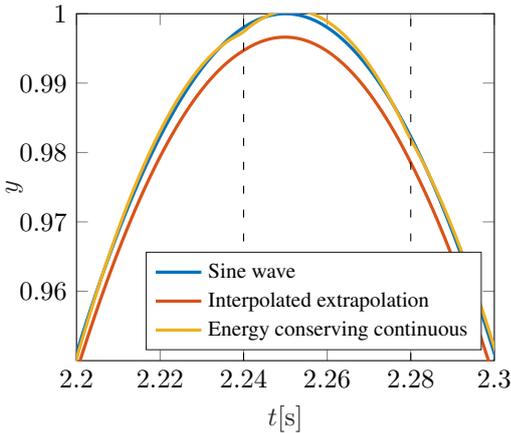


Figure 7: Approximation of the sine with and without control loops using continuous multi-rate methods.

4. Validation of the Multi-Rate Methods

For the validation of the implemented multi-rate methods, three input signals are examined. The first function $y_1(t)$ is the previously used sine wave as a periodic function. Second is an aperiodic function $y_2(t)$ resembling a

damped harmonic oscillator with a displacement at $t = 0$. Third is a function $y_3(t)$ with a point of discontinuity:

$$y_1(t) = f(t) = \sin(2\pi t) \quad (13)$$

$$y_2(t) = e^{(-\frac{3t}{2})} \cos(3t) \quad (14)$$

$$y_3(t) = \begin{cases} \frac{1}{3}(4t - t^2) & , 0 \leq t < 3 \\ 1 & , t \geq 3 \end{cases} \quad (15)$$

The three functions are shown in Figure 8.

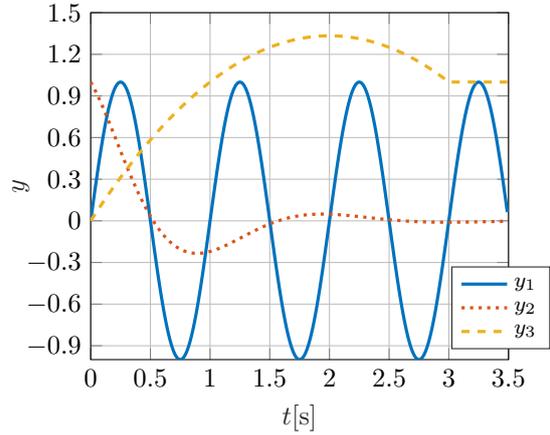


Figure 8: Input signals used to validate the multi-rate methods.

Using the validation functions in equations (13) to (15), the implemented multi-rate methods are evaluated by different criteria while using the validation functions as the input signal for the slow subsystem. The step size as well as the order of extrapolation of the multi-rate methods are varied to investigate different behaviour of the multi-rate methods. The variation of those parameters is shown in Table 2. For the validation three different combinations of step size and order of extrapolation are used, which is shown as variant A1, A2 and A3. For each variant the behavior of the multi-rate methods for the three different validation functions is recorded.

Table 2: Recordings for the theoretical validation of the multi-rate methods. The index after the dot corresponds to the respective input signal.

	A1{ y_1, y_2, y_3 }	A2{ y_1, y_2, y_3 }	A3{ y_1, y_2, y_3 }
h_s (ms)	40	40	75
n_E	3	4	3

The first and most important criterion to validate the multi-rate methods is the achievable accuracy. For this the mean squared error is calculated:

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2 \quad (16)$$

Where n is the number of considered time steps, y_i is the actual value at the time step i and \hat{y}_i is the value estimated by the multi-rate method. The results for the mean squared error of the nine recordings are listed in Table 3.

As additional criteria, continuity and differentiability are evaluated. Finally, the response time of the method is evaluated with respect to abrupt changes in the input

Table 3: Mean squared error of the multi-rate methods: Zero-Order-Hold (Zoh), discontinuous methods polynomial extrapolation (Pol), hermite extrapolation (Her), interpolated extrapolation (Int), smoothed extrapolation (Smo), energy conserving discontinuous method (Ecd) and energy conserving continuous method (Ecc).

	Zoh	Pol	Her	Int	Smo	Ecd	Ecc
A1 y_1	$1.01 \cdot 10^{-2}$	$1.54 \cdot 10^{-6}$	$2.57 \cdot 10^{-6}$	$7.64 \cdot 10^{-6}$	$4.34 \cdot 10^{-6}$	$9.57 \cdot 10^{-8}$	$2.72 \cdot 10^{-7}$
A1 y_2	$1.79 \cdot 10^{-4}$	$7.79 \cdot 10^{-10}$	$7.73 \cdot 10^{-9}$	$3.88 \cdot 10^{-9}$	$2.19 \cdot 10^{-9}$	$1.41 \cdot 10^{-11}$	$9.57 \cdot 10^{-11}$
A1 y_3	$1.45 \cdot 10^{-4}$	$1.29 \cdot 10^{-5}$	$1.31 \cdot 10^{-5}$	$5.14 \cdot 10^{-5}$	$2.77 \cdot 10^{-5}$	$3.95 \cdot 10^{-5}$	$2.45 \cdot 10^{-4}$
A2 y_1	$1.01 \cdot 10^{-2}$	$8.38 \cdot 10^{-8}$	$1.39 \cdot 10^{-5}$	$4.54 \cdot 10^{-7}$	$2.56 \cdot 10^{-7}$	$5.75 \cdot 10^{-9}$	$3.79 \cdot 10^{-8}$
A2 y_2	$1.79 \cdot 10^{-4}$	$6.17 \cdot 10^{-12}$	$4.77 \cdot 10^{-8}$	$3.32 \cdot 10^{-11}$	$1.89 \cdot 10^{-11}$	$2.08 \cdot 10^{-13}$	$2.51 \cdot 10^{-9}$
A2 y_3	$1.45 \cdot 10^{-4}$	$3.77 \cdot 10^{-5}$	$1.81 \cdot 10^{-5}$	$2.83 \cdot 10^{-4}$	$9.11 \cdot 10^{-5}$	$1.24 \cdot 10^{-4}$	$1.10 \cdot 10^{-3}$
A3 y_1	$3.56 \cdot 10^{-2}$	$2.31 \cdot 10^{-4}$	$2.34 \cdot 10^{-5}$	$1.20 \cdot 10^{-3}$	$6.39 \cdot 10^{-4}$	$4.62 \cdot 10^{-5}$	$3.52 \cdot 10^{-4}$
A3 y_2	$6.19 \cdot 10^{-4}$	$1.34 \cdot 10^{-7}$	$3.97 \cdot 10^{-8}$	$1.50 \cdot 10^{-6}$	$3.98 \cdot 10^{-7}$	$6.58 \cdot 10^{-9}$	$3.40 \cdot 10^{-6}$
A3 y_3	$5.10 \cdot 10^{-4}$	$8.68 \cdot 10^{-5}$	$8.84 \cdot 10^{-5}$	$3.39 \cdot 10^{-4}$	$1.83 \cdot 10^{-4}$	$2.67 \cdot 10^{-4}$	$1.60 \cdot 10^{-3}$

signal. An example for such a change would be the point of discontinuity in $y_3(t)$. All the multi-rate methods as well as the standard zero-order-hold are graded using a harvey balls table. An empty circle means the method can not meet the criteria. The more filled the circle is, the better is the performance of the method for the specific criteria. It should be noted that the evaluation in Table 4 is qualitatively and should serve as a decision-making aid for the selection of a multi-rate method. All the chosen criteria depend on the parameters of extrapolation order as well as the macro step size.

Table 4: Classification of the multi-rate methods regarding the output behavior for different input signals.

	Achievable Accuracy	Continuity	Differentiability	Reaction time
Zoh	○	○	○	●
Pol	●	◐	◐	●
Her	◐	◐	◐	●
Int	◐	●	●	◐
Smo	◐	●	●	◐
Ecd	●	◐	◐	◐
Ecc	●	●	●	○

The table can be used to choose the proper multi-rate method for given signals in a simulation scenario. For example, the highest accuracy can be achieved using the energy conserving methods. However if the input signal tends to change abruptly another method with a better reaction time should be chosen since the energy conserving methods tend to overshoot. Additional considerations are necessary if continuity and differentiability are necessary which come at the cost of accuracy.

5. Application Use Case in Industrial Robotics

In this section the presented concept is applied to a real VC use case. VC simulations which contain real programmable logic controllers (PLC) and computerized numerical controllers (CNC) for motion are normally calculated in a cycle time of one or a few milliseconds on a real-time operating system [3]. In comprehensive production systems with robots, the robot tasks are

triggered and observed by the PLC. To validate the interaction of all controls in a production system, virtual robot controls are integrated into a VC simulation. Those virtual controls are also used for the programming of the robots and are normally running on a non-real-time operating system with a higher cycle time than the PLC, CNC and the virtual production system. This is why multi-rate methods come into play at this use case. As an application example an industrial robot handling task is chosen. In the simulated scene the robot is performing a handling task, which is the most common task for industrial robots [13]. The simulation setup, which is used for the application example in this work is shown in Figure 9.

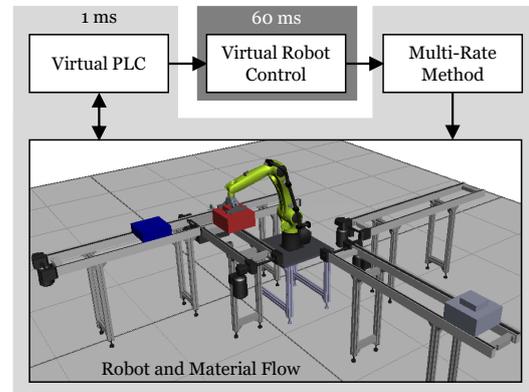


Figure 9: The example use case: A robot is performing a handling task.

For this example the robot control is running with a cycle time of 60 ms and the rest of the simulation setup is running with a cycle time of 1 ms. For the orchestration of the robot tasks a simple PLC is implemented into the simulation tool. The multi-rate methods are inserted between the virtual robot control and the simulation model of the robot. The robot itself performs a pick and place task where it interacts with parts, which are modeled physically. The robot, the material flow and the virtual PLC are modeled with the real-time simulation environment ISG-virtuos while the multi-rate methods are integrated with custom C++ blocks. As a robot control

a virtual CNC kernel is used and integrated into the simulation environment.

The robot movement is very dynamic, this is why a hermite-method is chosen as a multi-rate method for this example. According to Table 4 the hermite extrapolation has a quick reaction time which is necessary to prevent overshoots in the signal path. In Figure 10 the signal of the second axis of the robot is shown. The blue signal is the output signal of the robot control and the input to the fast task with a zero-order-hold method which has jumps at the macro step size. The red signal is the processed signal after the hermite-method has been applied to the input signal.

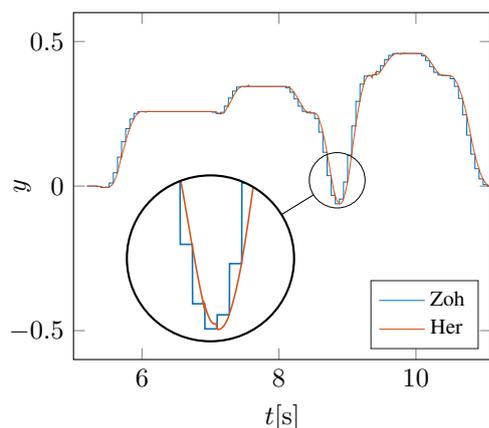


Figure 10: The input angle of the second axis of the robot with zero-order-hold (Zoh) and the hermite-method (Her).

It shows how the hermite-method can smooth the signal which helps to prevent collisions of the gripper with the transported parts, which can be seen in Figure 11. In the top picture the physics engine could not resolve the collision anymore, which is evident on the picture by the penetration of the workpiece and the gripper. With the help of the hermite-method in the lower picture it is possible to resolve the collision in the fast subsystem. Furthermore the multi-rate methods help that the robot movement looks smooth for the human eye.

6. Conclusion and Future Prospects

With the help of multi-rate methods it is possible to process coupling signals in distributed real-time simulations and therefore improve the overall simulation outcome. The focus of this work was the investigation of general multi-rate methods. Different methods have been implemented and compared. With Table 4 a basis for the choice of a proper method in a given simulation scenario is provided. To demonstrate the methods on a real use-case of VC, a robot handling task was chosen. It shows that the hermite-method is a good choice for this use case and can help to smooth the signals which are integrated in the fast subsystem. In future investigations the methods could be considered in a simulation setup which is distributed on different hardware where latency and fluctuating signals have a larger impact.

Acknowledgment

This work was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation DFG) within the project 438646126.

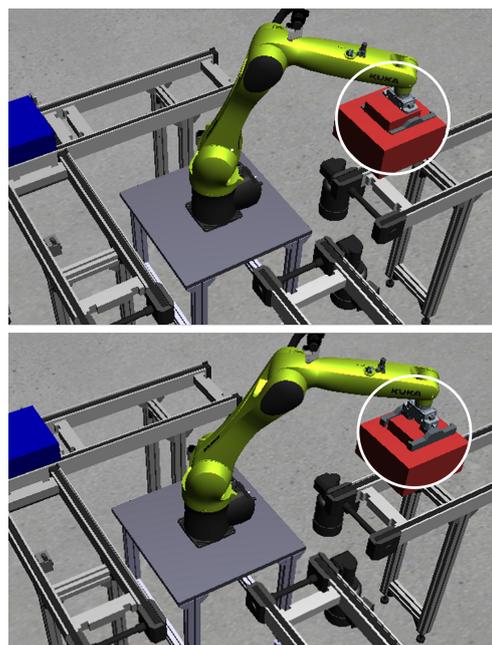


Figure 11: Comparison of zero-order-hold (top) and the hermite-method (bottom) at the application example where the coupling signals of all robot axis are processed by each method.

References

- [1] VDI, 4499 *Digital factory fundamentals*, 2008.
- [2] VDI, 3693 *Virtual commissioning model types and glossary*, 2008.
- [3] G. Pritschow and S. Röck, "Hardware-in-the-Loop" *Simulation of Machine Tools*, CIRP Annals, 2004.
- [4] C. Scheifele, O. Riedel, and A. Verl, "Real-time co-simulation for the virtual commissioning of production systems," CIRP ICME, 2018.
- [5] C. Gomes, C. Thule, D. Broman, P. G. Larsen, and H. Vangheluwe, *Co-simulation: State of the art*, 2017.
- [6] C. W. Gear and D. R. Wells, *Multirate linear multistep methods*, 1984.
- [7] C. Scheifele, *Plattform zur Echtzeit-Co-Simulation für die virtuelle Inbetriebnahme*. PhD thesis, 2019.
- [8] A. Kienzlen, C. Scheifele, and A. Verl, "Predicting coupling signals in a material flow real-time co-simulation with a kalman filter," CIRP ICME, 2020.
- [9] C. Scheifele and A. Verl, "Hardware-in-the-loop simulation for machines based on a multi-rate approach," EUROSIM and SIMS, 2016.
- [10] S. Knorr, *Multirate-Verfahren in der Co-Simulation gekoppelter dynamischer Systeme mit Anwendung in der Fahrzeugdynamik*, 2002.
- [11] R. Kossel, *Hybride Simulation thermischer Systeme am Beispiel eines Reisebusses*. PhD thesis, 2012.
- [12] M. Benedikt, *Eine Kopplungsmethode für die nicht-iterative Co-Simulation*. PhD thesis, 2012.
- [13] IFR, *World Robotics 2020 Industrial Robots*, 2020.

Application of Autonomous Inflow Control Valve for Enhanced Bitumen Recovery by Steam Assisted Gravity Drainage

Soheila Taghavi^{a,b}, Farzan Farsi Madan^a, Ramesh Timsina^a, Britt M. E. Moldestad^a

^a Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway,

^bInflowControl AS, Norway

corresponding author: soheila.t.hosnaroudi@usn.no

Abstract

Steam assisted gravity drainage is a thermal method for enhanced bitumen recovery. In this method, steam is injected to bitumen and heavy oil to reduce the viscosity and make the oil mobile. However, early breakthrough of steam in some parts of the well results in loss of the required amount of steam in contact with the cold bitumen, and poor distribution of the steam chamber. This limits the oil production and increases the SAGD operation cost. Autonomous inflow control valve (AICV) is able to prevent the steam breakthrough and restrict the production of steam. The objective of this paper is to investigate the performances of AICV and passive inflow control device (ICD) in a SAGD production well. This is achieved by developing a dynamic wellbore-reservoir model in the OLGA-ROCX simulator. Reservoir and fluid properties have been specified in ROCX, and the wellbore model has been developed in OLGA. Coupling OLGA and ROCX enable the user to simulate the fluid production from the reservoir into the well. The simulation results demonstrate the significant benefit of AICV in steam to oil ratio (SOR) reduction compared to ICD. Indeed, the simulation results show that utilizing AICV in the SAGD production wells will reduce the steam production by 88% after 300 days of production. From environmental aspect, reduction in the steam to oil ratio by utilizing AICV will reduce the energy demand for steam generation. This will eventually improve the economics of SAGD projects. Also, reduction in the steam and energy demand will consequently contribute to lower the intensity of greenhouse gas (GHG) emissions.

1.Introduction

Steam assisted gravity drainage is a thermal recovery method based on gravity drainage for extraction of bitumen and heavy oil. More than 80% of the world's annual heavy oil production is by means of deploying this technology [1]. As the bitumen and heavy oil viscosity are as high as 10^6 cP, the mobility is very low. Thus, the viscosity must be reduced drastically to make the bitumen mobile and extractable. Therefore, the SAGD method is used, where the oil is heated to temperatures around 200°C and higher. At this temperature range, the oil viscosity is below 20 cP (see Figure 1) which implies that the oil is mobile and is able to flow towards the production well by gravity.

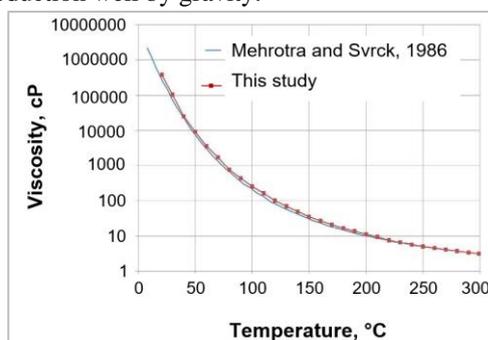


Figure 1: Viscosity of Athabasca bitumen sample versus temperature. [2]

The SAGD process is shown in Figure 2. Steam is injected continuously from the steam injection well which is located about 4-6m above the production well. As steam is injected, it forms a growing steam chamber with uniform temperature, called a depletion chamber. The continuous injected steam flows to the interface and condenses in contact with the cold bitumen. As a result, the latent steam energy is released leading to the higher oil temperature, lower oil viscosity, and consequently greater oil mobility. The low viscous mobile oil and condensate flow continuously from the edge of the steam chamber towards the production well.

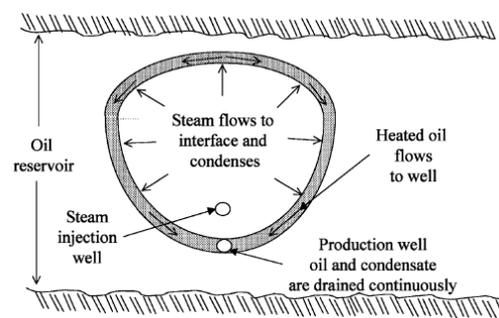


Figure 2: SAGD process. [3]

One of the key parameters of an efficient SAGD operation is attaining an even steam distribution

along the injection well. This can be achieved by deploying inflow control devices (ICDs) which balance the steam outflow to the reservoir. The role of ICD installation on the injection well is of great importance specially in the early phases of steam chamber growth, since it encourages more uniform steam development. [4]

One of the challenges in the SAGD wells is steam and water breakthrough in some parts of the well. This reduces the heavy oil/bitumen production and will consequently increase the SAGD operation cost. ICDs initially and autonomous inflow control devices (AICDs) latterly have been used to overcome this challenge. The newest generation of AICD is autonomous inflow control valve (AICV). AICV is able to delay the onset of steam and water breakthrough and ensure an even influx of oil along the well. In addition, in case of breakthrough of unwanted fluids like steam and/or water, AICV restricts the production of these fluids significantly. The ratio of steam injection to oil production (SOR) is of great importance in the SAGD process. From both environmentally and economically aspects, it is crucial to implement technologies which contribute to decrease the SOR.

The aim of this paper is to examine the impact of the ICD and AICV technology on reducing SOR and consequently improving the SAGD economics.

2. Inflow control technologies; ICD and AICV

Inflow control technologies such as ICDs and AICVs were introduced to the oil industry in order to overcome the early water and gas breakthrough challenges associated with heel-toe effect in horizontal wells. The heel-toe effect refers to the variations of the inflow rate of the fluid along the well, from toe to heel, due to the frictional pressure losses [5]. In addition, these technologies promote a balance drainage of long horizontal wells, and in general increase the oil production and recovery.

The following sub-sections present the functionality and performance curves of passive and autonomous inflow control device.

2.1. ICD

Figure 3 shows a pipe section with nozzle-type ICD completion. Fluid from the reservoir (red arrows) flows through the sand-screen, traverses along the annulus, and enters the production tubing through the ICD.

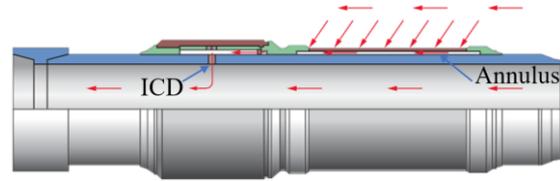


Figure 3: Nozzle-type ICD mounted on the pipe.[6]

The governing equation of the nozzle-type ICD is as follows [7]:

$$\Delta P = \frac{8\rho Q^2}{d^4\pi^2n^2C_D^2} \quad (1)$$

Where ΔP is the pressure drop through the nozzle, ρ is the fluid density, Q is the volumetric flow rate of the fluid through the nozzle, d is the diameter of the nozzle, n is the number of tested nozzles, and C_D is the discharge coefficient. C_D is mostly a function of the Reynolds number (Re) [7]. It can be interpreted from the equation (1) that the pressure drop through the nozzle is mainly dependent on the fluid density. The performance curve of a nozzle type ICD for water, oil and gas is shown in Figure 4. A nozzle type ICD with 4.75 mm diameter is used in the simulations. By performance curve, it means that the pressure drop through the device is plotted as a function of the volume flow rate of the fluid.

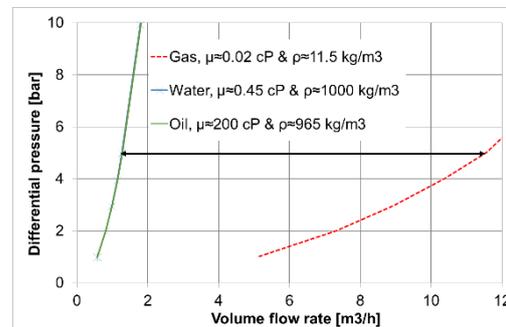


Figure 4: ICD performance curves.[8]

As it is illustrated in Figure 4, at constant pressure drop, the volume flow rate of oil and water differ slightly as the density differences are minor (1000 kg/m^3 for water vs 965 kg/m^3 for oil), while the volume flow rate of gas is much higher due to the much lower density (about 11.5 kg/m^3). This indicates that when gas breakthrough occurs, ICD will not restrict the gas production significantly.

2.1. AICV

Figure 5 shows a pipe section with sand screen and AICV completion.

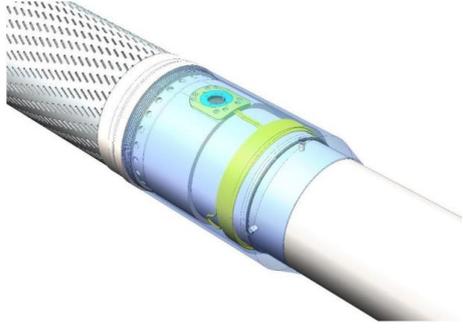


Figure 5: AICV mounted in a base-pipe with sand screen.[9]

The mathematical model describing the performance of the AICV can be described as:

$$\Delta P_{Tot} = \left(\frac{\rho_{mix}}{\rho_{cal}} \right) \cdot \left(\frac{\mu_{cal}}{\mu_{mix}} \right)^y \cdot a_{AICV} \cdot Q^x \quad (2)$$

where ΔP_{Tot} is the differential pressure across the AICV, ρ_{cal} and μ_{cal} are the calibration fluid density and viscosity, and ρ_{mix} and μ_{mix} are the mixture fluid density and viscosity respectively. The parameter a_{AICV} is a valve characteristic given by the ICD strength, Q is the volumetric mixture flow rate, and x and y are constants. [10]

It can be interpreted from equation (2) that the pressure drop through the AICV is much more viscosity dependent than density dependent. The concept and principle of AICV is described in detail in earlier scientific works [11, 12].

The performance curve of AICV for water, oil and gas is shown in Figure 6.

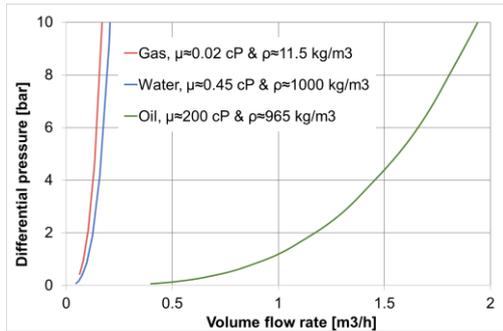


Figure 6: AICV performance curves.[8]

As it is illustrated in Figure 4 and Figure 6, at constant pressure drop the volume flow rates of oil with 200 cP viscosity through the ICD and the AICV are almost the same, while the gas flow rate is significantly higher through the ICD than the AICV. The location of performance curves of the gas and water for AICV have changed compared to the performance curves through ICD. The gas and water curve are now located far away to the left side of the oil curve. This indicates that when gas and water breakthrough occur, AICV will restrict the gas and

water production significantly while maintaining a high oil production.

3. OLGA-ROCX set up

In this study, OLGA-ROCX is used to describe and compare the behavior of ICD and AICV in the SAGD reservoir, and to illustrate the impact of AICV completion on increased oil recovery. OLGA-ROCX is an integrated transient well/reservoir model. The reservoir model and the wellbore model are coupled in an implicit way [13]. Reservoir and fluid properties are specified in ROCX, and the wellbore model is developed in OLGA. Coupling OLGA and ROCX enable the user to simulate the fluid production from the reservoir into the well. NETool was used to simulate the AICV performance in a SAGD reservoir in previous author's work [8]. NETool is a static one-dimensional steady state simulation tool that shows the instantaneous inflow profile along the well, while OLGA/ROCX is a robust transient simulation tool to perform integrated well-reservoir simulations.

3.1. Reservoir model, ROCX

The black oil model which simulates the multiphase fluid transport in porous media is selected in ROCX. Input data to the model are grid dimensions, fluid, and reservoir properties. Reservoir properties such as permeabilities and porosities of the porous medium, and in addition thermal properties of the rock and fluids are among the input data. Initial conditions such as fluid saturations and temperature together with the boundary conditions at the well and at the outer near well boundary are needed. [14]

The boundary conditions of the reservoir grid elements are matched with inflow points of the components placed in the wellbore model. This will define the flow from the reservoir model. So, the pressure boundary for the reservoir model is provided by the wellbore model while the reservoir model provides the flow and the fluid temperatures into the pipeline[14]. The shape of the reservoir drainage area is considered to be rectangular, and the dimensions are given in Table 1.

Table 1: The dimensions of the drainage area.

Direction	Length (m)	Number of blocks	Block size (m)
X	1000	10	100
Y	117	15	20,20,5,5,3,3,2,1,2,3,3,5,5,20,20
Z	40	10	4

As the fluid properties varies significantly around the well and in the Y-Z plane, a finer mesh is considered in the grid setup to achieve more accurate results. The size of the blocks varies along the y direction while a uniform mesh along the z and x-direction is defined. Finer mesh along the x-direction will have insignificant impact on the overall flow rate [15]. The well length is 1000 m containing 10 segments with a length of 100 m each. One equivalent ICD/AICV is placed in each segment.

Since the purpose of this study is to evaluate the ICD/AICV performance in case of steam breakthrough, the well is located as near as possible to the bottom of the drainage area to delay the probable steam breakthrough. The schematic of the drainage area geometry and the well location is shown in Figure 7.

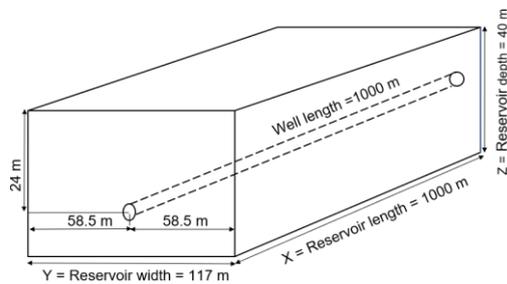


Figure 7: The schematic of the drainage area geometry.

The grid in three dimensions is shown in Figure 8.

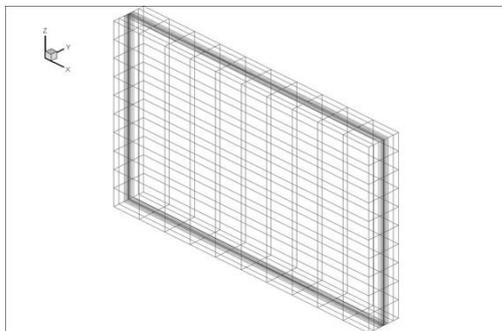


Figure 8: 3-D view of grid.

As seen from the figure, the mesh in the y-direction and towards the well located in the middle of the drainage area is finer than in the z and x-direction.

3.1.1 Fluid and reservoir properties

The black oil model which estimates the pressure volume temperature (PVT) relations is selected in ROCX. The basic modeling assumption is that the gas may dissolve in the liquid oil phase, but no oil will dissolve in the gaseous phase. This implies that the composition of the gaseous phase is assumed the same at all pressure and temperatures [12],[16]. In other words, the black oil model assumes that the oil

components will always be in the liquid phase despite any changes in the conditions [15].

The reservoir pressure at initial conditions is 27 bar and it is assumed to be constant. The fluid properties used for the simulation are listed in Table 2. The oil viscosity is measured at 180°C at atmospheric pressure [2].

Table 2: Fluid properties as input in ROCX.

Property	Value
Oil viscosity [cP]	15.50
Oil specific gravity [-]	0.90
Gas specific gravity [-]	0.64
Gas oil ratio (GOR) [Sm ³ /Sm ³]	150

A gas cap is placed at the top of the reservoir in the boundary conditions section. Hence a gas feed in addition to oil feed are defined. The feed streams are presented in Table 3.

Table 3: Feed streams.

Feed stream	Fraction type	Fraction
Oil	GOR	150
Gas	OGR	0.99

The reservoir porosity is assumed to be constant 0.3 throughout the reservoir. Different permeabilities are specified for each block in order to simulate a heterogenous reservoir. It is assumed that the area close to the toe section of the well has a higher permeability in all directions. The permeability of the heterogenous reservoir in both x and y-direction, varies from 3000 mD in relatively low permeable zones to 6000 mD in relatively high permeable zones. The vertical permeability is specified in each block of the reservoir, and it varies from 300 to 600 mD for relatively low permeable and relatively high permeable zones respectively. The vertical permeability profile of the heterogenous reservoir is illustrated in Figure 9

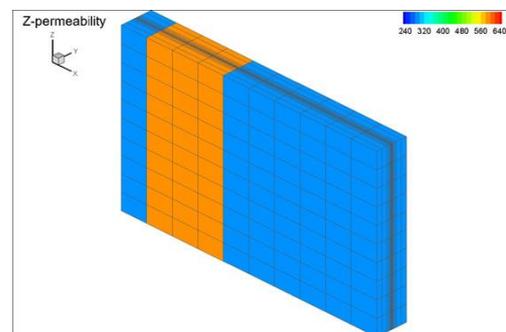


Figure 9: Vertical permeability profile.

Generally, it is challenging to obtain information about relative permeability for different fields. Data for relative permeabilities are set manually in table

form in Rocx. The ‘‘Stone II’’ model with exponent 2 is used to evaluate the oil phase relative permeability while the Corey correlation with exponent 1.5 is used to estimate the gas phase relative permeability. The relative permeability curves for oil and gas are presented in Figure 10.

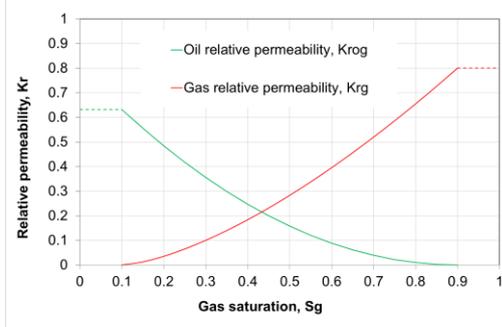


Figure 10: Relative permeability curves for the SAGD reservoir.

3.1.2 Initial and boundary conditions

Initially, it is assumed that the fraction of the black oil feed is equal to one. The initial oil and gas saturation of in the reservoir are set to 0.9 and 0.1. Pressure and temperature of the reservoir are 27 bar and 180°C, respectively.

3.2. Wellbore model, OLGA

In OLGA, separate continuity equations are applied for the gas, for the oil (or condensate) and water liquids and also for oil (or condensate) and water droplets. Three momentum equations are also used: one for each of the continuous liquid phases (oil/condensate and water) and one for the combination of gas with liquid droplets. One mixture energy equation is also applied. Totally, seven conservation equations and one equation of state to be solved: the seven conservation equations are three for mass, three for momentum, and one for energy, while the equation of state is for pressure.[17]

3.2.1 Mass Transport Equations:

$$\partial_t m_i + \partial_z(m_i U_i) = \sum_j \Psi_{ji} + G_i \quad (3)$$

where m_i and U_i denote mass field (gas, oil in liquid layers, water in liquid layers, oil droplets in gas layer, and water droplets in gas layer) and velocity respectively. In addition, ∂_t denotes differentiation in time, ∂_z denotes spatial differentiation, Ψ_{ji} denotes the rate of mass transfer between the j -th and i -th mass field, that is, dispersions, droplet deposition and entrainment, and phase transitions, and G_i denotes any mass source/sink.

3.2.2 Momentum Balance Equations

$$\partial_t(m_i U_i) + \partial_z(m_i U_i^2) = m_i \cdot g \cdot \cos(\varphi) + p_i + G_i U_i + \sum_j (\Psi_{ji}^+ U_i - \Psi_{ji}^- U_i) + \sum_j F_{ji}^l (U_j - U_i) - F_i^w U_i \quad (4)$$

where ∂_t denotes differentiation in time, ∂_z denotes spatial differentiation. g is the acceleration of gravity, φ is the pipe angle relative to the gravitational vector, P_i is the pressure force, $G_i U_i$ is the momentum contribution corresponding to the mass source/sink G_i . Also, F_{ji}^l are friction forces between the i -th and j -th mass field, and F^w denotes the wall friction. Ψ_{ji} denotes momentum contributions corresponding to the mass transfer between the j -th and i -th mass field. In the equation (4), Ψ_{ji}^+ accounts for a net contribution from mass field i to j while Ψ_{ji}^- accounts for a net contribution from mass field j to i .

3.2.3 Energy balance equation

$$\partial_t(m_i E_i) + \partial_z(m_i U_i H_i) = \delta_i + Q_i + \sum_j T_{ij} E_j \quad (5)$$

where E_i denotes the field energy, H_i denotes the field enthalpy, S denotes enthalpy source/sink, Q is the heat flux through the pipe wall, and T_{ij} models the energy transfer between fields.

3.3 OLGA set-up

In this work, a basic case is selected to generate the wellbore model in OLGA. The model consists of two pipes: pipeline and flowpath. The flowpath represents the production tubing, and the pipeline represents the annulus and the inflow from annulus to the well. On the pipeline, the flow components such as inflow controls and packers are placed. The characteristics and dimensions of the two pipes are listed in Table 4.

Table 4: The characteristics of pipeline and flowpath.

Pipe	Length (m)	Diameter (mm)	Roughness (mm)
Pipeline	1000	222	0.028
Flowpath	1000	114	0.050

Figure 11 shows the set-up in OLGA for one production zone which consists of two sections. This set-up was developed and proposed for the first time in 2012 [12] and results were presented in a scientific paper [11].

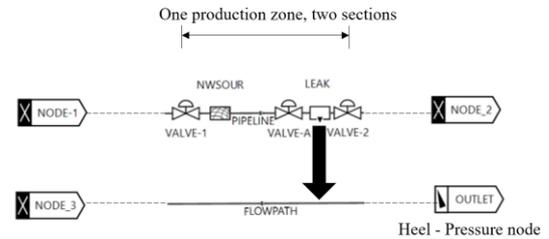


Figure 11: Set-up of a single production zone with inflow controller in OLGA.

The pipeline (PIPELINE) includes a near-well source (NWSOUR) which connects the ROCX file

as input data. The near-well source is the flow source from the reservoir to the annulus. In addition, the pipeline consists of one inflow controller (VALVE-A), two packers (VALVE-1 and 2 with zero opening), and leak (LEAK) which outflows the flow through the inflow control to the production well (FLOWPATH) and towards the heel (OUTLET). Each production zone is divided into two sections and is isolated by packers. The near-well source and the leak is placed in each section and the packers, and the inflow controllers are the section boundaries. Each pipe is divided into 10 production zones each 100 m long which implies that there are totally 20 sections. 10 nozzle type ICDs with a diameter of 15 mm is distributed in the 10 production zones. The flow area of this ICD in one production zone (100 m) corresponds to the flow area of 10*4.75 mm ICDs. Usually, in the industry, the ICDs are installed in approximately 10-11 m long joints. So, the flow area of one equivalent ICD in a 100 m production zone is approximately the same as the flow area of 10 ICDs with 4.75 mm diameter which are installed in 10-11 meters long joints.

A comprehensive and step by step workflow for modelling of the rate controlled production (RCP) valves in the OLGA simulator was proposed in a scientific paper [18]. In this new method, a controller is used to check the RCP valve based on the characteristics of the RCP valve and the reservoir fluid mixture. This method can also be used for AICVs. Set-up of a single production zone with AICV in OLGA is illustrated in Figure 12 .

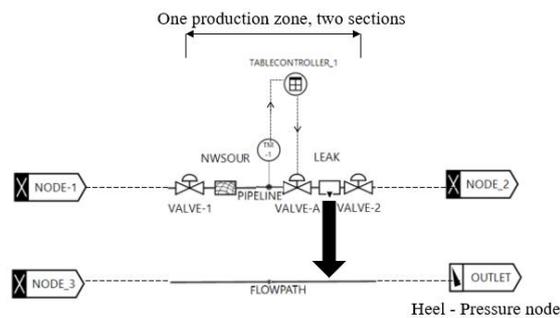


Figure 12: Set-up of a single production zone for AICV in OLGA.

Based on the single-phase (Figure 6) and multi-phase gas/oil performance of the AICV, a table controller (TABLECONTROLLER-1) is created. This table controller gets the measured gas volume fraction (GVF) data from the transmitter (TM-1) and provides corresponding control signals for chocking the AICV. The multiphase gas/oil behavior of the AICV for SAGD conditions was presented earlier in a scientific paper [8] . According to the experimental results provided in that paper, the AICV gradually

opens when the oil/gas mixture flows through the valve. However, the AICV restricts the gas flow when the GVF is getting higher, until pure gas flows through the valve and the valve is almost closed.

The control signal table in the OLGA simulator for controlling the AICV, consists of independent and dependent variables. In this case, the percentage of the valve opening is a function of the GVF. Indeed, the valve opening is getting less and less by increasing the GVF.

4. Results and discussion

In this chapter, the obtained simulation results which are conducted for two cases are presented. The simulation cases are as follows:

1. Heterogeneous reservoir with ICDs
2. Heterogeneous reservoir with AICVs

The gas density and viscosity in the simulations performed by OLGA/ROCX, are set to 11.5 kg/m³ and 0.02 cP respectively. In the simulations and the experiments, the gas represents steam.

In order to study the performance of ICD and AICV in a specific well production period, the accumulated oil and gas for AICV and ICD completions are compared. Figure 13 illustrates the accumulated oil and gas produced from the well after 300 days of production.

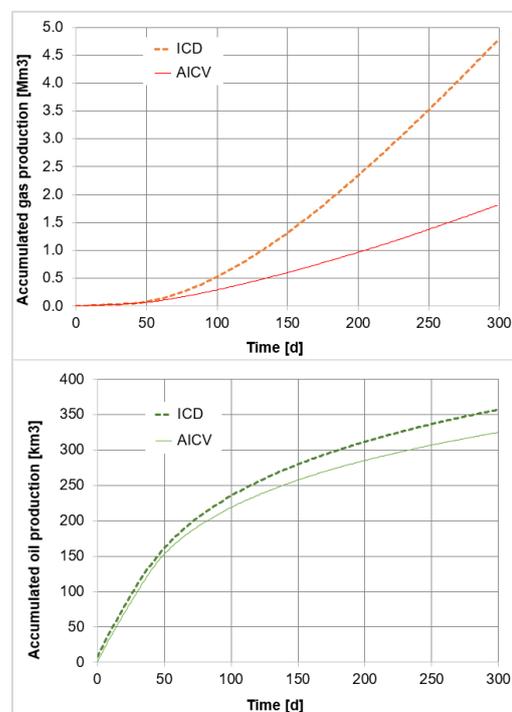


Figure 13: Accumulated oil and gas from the well completed with AICVs and ICDs.

As can be seen in this figure, after 300 days of production, the accumulated oil in the AICV and ICD cases differs marginally. However, due to the better performance of AICV in both single and multiphase flow regions, the amount of accumulated gas drops significantly from 4.8 Mm³ to 1.9 Mm³ after 300 days of production.

When the gas breakthrough occurs, AICV starts to choke the gas production gradually. Indeed, AICV chokes the gas production consistently by increasing GOR. This behavior, which is based on experimental data [8], was implemented in the Table Control module in OLGA.

Figure 14 Shows the comparison of oil and gas production rates for AICV and ICD completion for 300 days of production. The oil production rate for both cases reach its maximum at the start of the production. The oil production decreases slightly as the gas production increases simultaneously. Gas production grows suddenly and rapidly at 35th day of production, which implies that gas breakthrough has occurred. At the time of gas breakthrough, gas enters the well rapidly due to its high mobility. This restricts the production of oil significantly, and as a result, the oil production rate drops drastically. However, oil production is continued at an acceptable level until the end of the production time.

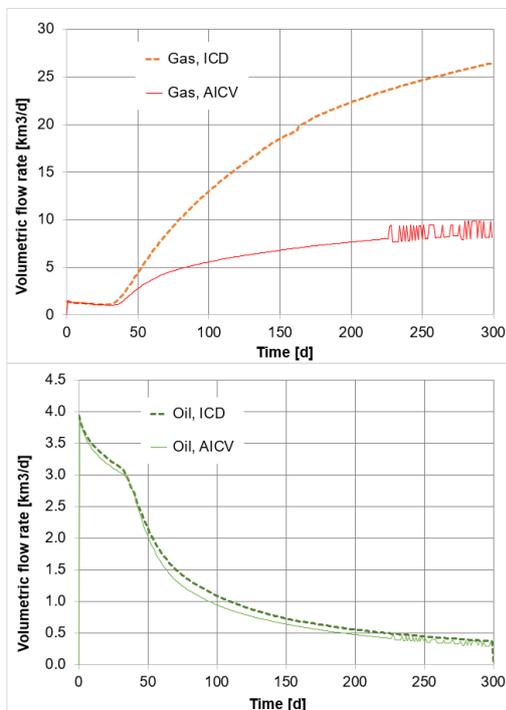


Figure 14: Volumetric flow rate of oil and gas for the well completed with AICVs and ICDs.

As can be seen from the figure, the volumetric oil flow rate of ICD is on average slightly higher than

the volumetric oil flow rate of AICV during the whole period of production. However, the gas breakthrough occurs a few days later for the AICV case than for the ICD case. Also, the development of gas breakthrough is much faster for the ICD case compared to the AICV case. It can be concluded from the figure that the well completed with AICVs reduces the gas production by approximately 88% compared to using ICDs after 300 days of production.

Figure 15 shows the GOR at standard conditions as a function of accumulated oil production. This figure illustrates how the GOR varies with accumulated oil. Usually in the wells, the total allowable gas production is limited, since the total gas processing capacity is an active constraint [19]. This highlights the importance of developing new inflow control technologies which guarantee a higher maximum oil production while meeting the GOR constraint.

As it can be seen from the figure, the accumulated oil at a specific GOR, for example 600, for the AICV case is 15% more than the accumulated oil for the ICD case.

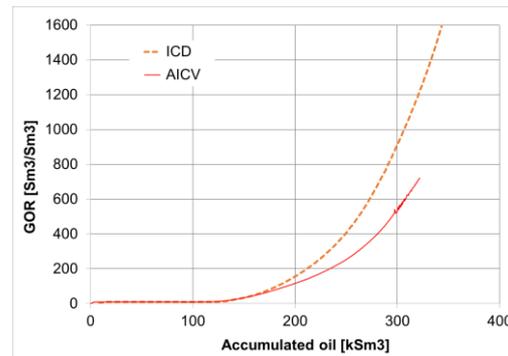


Figure 15: Accumulated oil production versus gas oil ratio for AICV and ICD.

4. Conclusions

The performances of AICV and ICD in a SAGD production well are investigated. This is achieved by developing a dynamic wellbore-reservoir model in the OLGA-ROCX simulator. Reservoir and fluid properties are specified in ROCX, and the wellbore model is developed in OLGA. Coupling OLGA and ROCX enable the user to simulate the fluid production from the reservoir into the well.

The simulation results demonstrate the significant benefit of AICV in SOR reduction compared to ICD. Indeed, simulation results show that utilizing AICV in the SAGD production wells will reduce the gas (steam) production by 88% after 300 days of production. Reduction in SOR, will improve the overall SAGD operation performance. This will also result in more cost-effective oil production.

From environmental aspect, reduction in the steam to oil ratio by utilizing AICV, will reduce the energy demand for steam generation. This will improve the economics of SAGD projects. Also, reduction in the steam and energy demand will consequently contribute to lower the intensity of greenhouse gas (GHG) emissions.

Acknowledgment

The author would like to thank the university of South-Eastern Norway for providing the necessary software arrangements for this work.

References

- [1] Z. Xu, S. Li, B. Li, D. Chen, Z. Liu, and Z. Li, "A review of development methods and EOR technologies for carbonate reservoirs," *Petroleum Science*, pp. 1-24, 2020.
- [2] A. Jahanbani, T. Jelmert, J. Kleppe, M. Ashrafi, Y. Souraki, and O. Torsæter, "Investigation of the Applicability of Thermal Well Test Analysis in Steam Injection Wells for Athabasca Heavy Oil," 06/04 2012, doi: 10.2118/154182-MS.
- [3] R. Butler, "SAGD Comes of AGE!," *Journal of Canadian Petroleum Technology*, vol. 37, no. 07, 1998, doi: 10.2118/98-07-da.
- [4] B. Least, S. Greci, R. Huffer, R. V. Rajan, and H. Golbeck, "Steam Flow Tests for Comparing Performance of Nozzle, Tube, and Fluidic Diode Autonomous ICDs in SAGD Wells," in *SPE Heavy Oil Conference-Canada*, 2014, vol. Day 3 Thu, June 12, 2014, D031S013R005, doi: 10.2118/170083-ms. [Online]. Available: <https://doi.org/10.2118/170083-MS>
- [5] V. M. Birchenko, K. M. Muradov, and D. R. Davies, "Reduction of the horizontal well's heel-toe effect with inflow control devices," *Journal of Petroleum Science and Engineering*, vol. 75, no. 1, pp. 244-250, 2010/12/01/ 2010, doi: <https://doi.org/10.1016/j.petrol.2010.11.013>.
- [6] T. Jokela *et al.*, "Inflow Control Devices — Raising Profiles," 2010.
- [7] J. E. Lauritzen and I. B. Martiniussen, "Single and Multi-phase Flow Loop Testing Results for Industry Standard Inflow Control Devices," in *SPE Offshore Europe Oil and Gas Conference and Exhibition*, 2011, vol. All Days, SPE-146347-MS, doi: 10.2118/146347-ms. [Online]. Available: <https://doi.org/10.2118/146347-MS>
- [8] S. Taghavi, H. Aakre, and B. M. E. Moldestad, "Performance Analysis of Autonomous Inflow Control Valve in a SAGD Late Life Process with Non-Condensable Gases," in *SPE Canadian Energy Technology Conference*, 2022, vol. Day 1 Wed, March 16, 2022, D011S011R002, doi: 10.2118/208915-ms. [Online]. Available: <https://doi.org/10.2118/208915-MS>
- [9] H. Aakre, V. Mathiesen, and B. Moldestad, "Performance of CO₂ flooding in a heterogeneous oil reservoir using autonomous inflow control," *Journal of Petroleum Science and Engineering*, vol. 167, pp. 654-663, 2018/08/01/ 2018, doi: <https://doi.org/10.1016/j.petrol.2018.04.008>.
- [10] S. Taghavi and A. Ghaderi, "On Uncertainty Analysis of the Rate Controlled Production (RCP) Model," in *First SIMS EUROSIM Conference on Modelling and Simulation*, Finland, 2021: Linköping Electronic Conference Proceedings, 2021.
- [11] H. Aakre, B. Halvorsen, B. Werswick, and V. Mathiesen, "Smart Well With Autonomous Inflow Control Valve Technology," in *SPE Middle East Oil and Gas Show and Conference*, 2013, vol. All Days, SPE-164348-MS, doi: 10.2118/164348-ms. [Online]. Available: <https://doi.org/10.2118/164348-MS>
- [12] H. Aakre, n. o. m. f. Høgskolen i Sørøst-Norge Fakultet for teknologi, and S.-N. Høgskolen i, "The impact of autonomous inflow control valve on increased oil production and recovery," no. 32, University College of Southeast Norway, Faculty of Technology, Natural Sciences and Maritime Sciences, Kongsberg, 2017.
- [13] G. Chupin, B. Hu, T. Haugset, J. Sagen, and M. Claudel, "Integrated Wellbore/Reservoir Model Predicts Flow Transients in Liquid-Loaded Gas Wells," in *SPE Annual Technical Conference and Exhibition*, 2007, vol. All Days, SPE-110461-MS, doi: 10.2118/110461-ms. [Online]. Available: <https://doi.org/10.2118/110461-MS>
- [14] J. Sagen, T. Sira, A. Ek, S. Selberg, M. Chaib, and H. Eidsmoen, "A Coupled Dynamic Reservoir and Pipeline Model – Development and Initial Experience," in *13th International Conference on Multiphase Production Technology*, 2007, vol. All Days, BHR-2007-G1.
- [15] N. C. I. F. Ramesh Timsina, Britt M. E. Moldestad, "Modeling and simulation of light oil production using inflow control devices," in *58th Conference on Simulation and Modelling (SIMS 58) Reykjavik, Iceland*, 2017: Linköping Electronic Conference Proceedings doi: <http://dx.doi.org/10.3384/ecp17138180>.
- [16] "Black-Oil Simulations." wikitext. https://dwsim.org/wiki/index.php?title=Black-Oil_Simulations (accessed).
- [17] K. H. Bendiksen, D. Malnes, R. Moe, and S. Nuland, "The Dynamic Two-Fluid Model OLGA: Theory and Application," *SPE Production Engineering*, vol. 6, no. 02, pp. 171-180, 1991, doi: 10.2118/19451-pa.
- [18] A. Moradi and B. M. E. Moldestad, "A Proposed Method for Simulation of Rate-Controlled Production Valves for Reduced Water Cut," *SPE Production & Operations*, vol. 36, no. 03, pp. 669-684, 2021, doi: 10.2118/205377-pa.
- [19] A. Mjaavatten, R. Aasheim, S. Saelid, and O. Groenning, "A Model for Gas Coning and Rate-Dependent Gas/Oil Ratio in an Oil-Rim Reservoir," in *SPE Russian Oil and Gas Technical Conference and Exhibition*, 2006, vol. All Days, SPE-102390-MS, doi: 10.2118/102390-ms. [Online]. Available: <https://doi.org/10.2118/102390-MS>

Process Simulation and Automated Cost Optimization of CO₂ Capture Using Aspen HYSYS

Lars Erik Øi*, Pouya Rahmani, Sumudu Karunaratne, Solomon Aromada

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway
lars.oi@usn.no

Abstract

A standard process for CO₂ capture based on absorption in mono ethanolamine (MEA) has been simulated and cost estimated with an equilibrium-based model in Aspen HYSYSTM. The aim has been to calculate cost optimum process parameters and evaluate the possibility of automated cost optimization using a spreadsheet facility. An Excel spreadsheet is used for dimensioning and cost estimation of the specified process. New in this work is that Visual Basic for Application (VBA) was used to automatically update installation factors for next iteration based on cost calculations in previous iteration. The equipment cost was obtained from the Aspen In-plant Cost EstimatorTM, and an enhanced detailed factor (EDF) method was used to estimate the total investment cost. The optimum process was found as the process with minimum calculated total cost. The optimum process parameters for the standard process were calculated to 15 m absorber packing height, 9 °C minimum approach temperature and 2.2 m/s superficial gas velocity through the absorber. With this approach, iterative cost estimation and optimization of CO₂ absorption and desorption processes can be performed automatically.

Keywords: Carbon capture, Aspen HYSYS, simulation, cost estimation, optimization

1. Introduction

1.1. Aim

One aim of this work was to calculate cost optimum process parameters for a traditional amine based CO₂ capture process. Another aim was to evaluate the possibility of automated cost optimization using a spreadsheet facility. An Excel spreadsheet is traditionally used for dimensioning and cost estimation of the specified process with optimization performed by minimizing the total cost calculated in the spreadsheet. New in this work is that Visual Basic for Application (VBA) was used to automatically update installation factors for next iteration based on cost calculations in previous iteration.

1.2. Literature

There is a large number of papers on cost estimation of CO₂ capture plants (Rao and Rubin, 2002; Rubin *et al.*, 2013; van der Spek *et al.*, 2019). Some of these are based on a combination of process simulation and cost estimation (Abu-Zahra *et al.*, 2007; Amrollahi *et al.*, 2012; Nwaoha *et al.*, 2018). This work is a further development of earlier works at the Telemark University College and the University of South-Eastern Norway (USN). The projects have been focused on process simulation, equipment dimensioning, capital and operating cost estimation, and cost optimization of CO₂ capture processes using the process simulation tool Aspen

HYSYS. Some of the previous works are Kallevik (2010), Øi (2012), Park and Øi (2017), Aromada and Øi (2017), and Øi *et al.* (2021).

The cost estimation part has been based on the Enhanced Detailed Factor (EDF) method (Ali *et al.*, 2019; Aromada *et al.*, 2021). While this method has several advantages, the time required to implement the detailed installation factors in capital cost estimation is a drawback. This becomes cumbersome when there is a need to run several process simulations by varying a process parameter followed by cost estimation for each iteration.

Recently, the focus has been on automatic process simulation combined with cost estimation for fast cost optimization of CO₂ capture processes in Aspen HYSYS (Haukås, 2020; Øi *et al.*, 2021). The Iterative Detailed Factor (IDF) scheme was then developed (Aromada *et al.*, 2022a). The IDF scheme was applied for several minimum temperature approach cost optimization studies in (Aromada *et al.*, 2022a). However, there was yet a need for manual observation for implementing any change required in the detailed installation factors and subfactors whenever process parameters are varied for subsequent simulation of the CO₂ capture process. Therefore, there is a need to make the entire process simulations, equipment dimensioning and cost estimation automatic, without requiring any manual input as done in the IDF scheme mode of implementation (Aromada *et al.*, 2022a). This was

accomplished in this work by linking Aspen HYSYS simulation spreadsheets with Microsoft Excel by a VBA code. This has been discussed by Sharma and Rangaiah (2016).

The Aspen HYSYS library was activated in Microsoft Excel from the developer tab > visual basic > tools> and preference (Rangaiah, 2016; Rahmani, 2021). The Aspen HYSYS root needs to be inserted into a Microsoft Excel sheet and it should be updated for each model. The Aspen HYSYS application should be closed during the process. The VBA script was developed for coupling the Aspen HYSYS and Microsoft Excel spreadsheets. The code imports equipment prices from the Aspen HYSYS spreadsheet into an Excel spreadsheet for cost estimation and optimization. At the same time, the right equipment units' installation factors/sub-factors are automatically imported from the Microsoft Excel sheet into the Aspen HYSYS spreadsheet. By this, the EDF method is simply and automatically implemented very fast without errors in the selection of the detailed installation factors/subfactors. The code and more details are documented in (Rahmani, 2021).

With this new approach of involving a VBA code, human errors in selecting EDF method installation factors and subfactors for different equipment are eliminated. And most importantly, each time a new process simulation is performed, the costs are automatically available without requiring any form of manual inputs.

With this work, process simulation based CO₂ capture process parameter cost optimization studies and sensitivity analysis can be conducted quickly and obtain reasonably accurate results. This paper documents cost optimization studies conducted with this new approach and comparison with other works. This work is based on the Master thesis work of Rahmani (2021).

1.3. Process Description

Prior to the CO₂ capture process, the flue gas is cooled in a direct contact cooler (DCC) with circulating water before it is sent to the absorption column. The amine with absorbed CO₂ from the bottom of the absorption column is pumped through the rich/lean heat exchanger with the temperature after the heat exchanger specified. The hot amine solution is entering the desorption column which separates the feed into the CO₂ product at the top and hot regenerated amine at the bottom. The regenerated amine is pumped to a higher pressure in a pump, then passes through the lean/rich heat exchanger and is further cooled in the lean cooler before it again enters the absorption column.

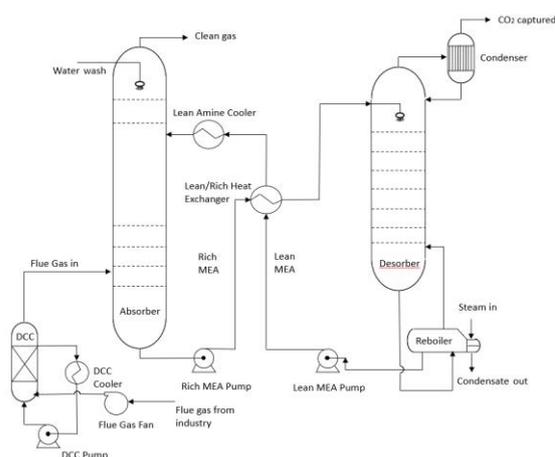


Figure 1: Process flow diagram of a standard amine-based CO₂ capture process (Aromada *et al.*, 2020).

2. Specifications and simulations

2.1. Specifications and simulation of standard CO₂ capture process

Specifications for the base case are given in Tab. 1. In the base case, 85 % CO₂ removal efficiency and a minimum approach temperature of 10 °C was achieved in the lean/rich heat exchanger. Aspen HYSYS version 11 was used with the Acid Gas package as the recommended equilibrium model by Aspen HYSYS.

The calculation sequence is similar to earlier works (Øi and Haukås, 2021). The calculation strategy is based on a sequential modular approach (Kisala *et al.*, 1987; Ishii and Otto, 2008). The calculation starts with the flue gas inlet stream and a guessed amount of the lean amine stream. After calculation of the DCC and the absorption column, the amine/amine heat exchanger is calculated based on a guessed (or specified) temperature in the stream from the desorber. The temperature can be adjusted in an adjust block to obtain a specified minimum temperature approach. After calculation of the desorber, the lean amine pump and the amine/amine heat exchanger the lean amine cooler is calculated to give an updated lean amine amount and composition.

The updated lean amine amount and composition is checked in a recycle block with the amine stream from the last iteration.

Adjust and Recycle operations in the flowsheet are used to get an automated simulation model. An Adjust block is adjusting the minimum approach temperature in the lean/rich heat exchanger by varying the temperature on the hot side outlet. A Recycle block is adjusting the removal efficiency by varying the lean amine mass flow. The Aspen HYSYS flowsheet is shown in Fig. 2.

Water was added to the process (water make-up). The make-up water can be calculated by a material balance.

Table 1: Aspen HYSYS model parameters and specifications for the base case alternative.

Parameter	Value
Inlet flue gas temperature [°C]	110/40.0
Inlet flue gas pressure [kPa]	101/120
Inlet flue gas flow rate [kmol/h]	10910
CO ₂ content in inlet gas [mole %]	3.3
Water content in inlet gas [mole %]	6.9
Lean amine temperature [°C]	40.0
Lean amine pressure [kPa]	101.0
Lean amine rate [kg/h]	132100
MEA content in lean amine [mass %]	22.5
CO ₂ content in lean amine [mass %]	3.5
Number of stages in absorber [-]	10
Murphree efficiency in absorber [m ⁻¹]	0.25
Rich amine pump pressure [kPa]	200.0
Rich amine temp. out of HEX [°C]	103.6
Number of stages in desorber [-]	6
Murphree efficiency in desorber [m ⁻¹]	0.5
Reflux ratio in stripper [-]	0.3
Reboiler temperature [°C]	120.0
Lean amine pump pressure [kPa]	500.0

2.2. Parameter variations

10 stages, 85 % removal efficiency and 10 °C as minimum approach temperature were specified in the base case simulation. For parametric studies, the packing height and minimum approach temperature were varied. The gas velocity through the absorber column and the pressure drop were also varied.

The Case study function in Aspen HYSYS was used to perform a series of calculations automatically. An important restriction is that when using the Case study function, it is not possible to perform other adjustments for each new parameter value.

2.3. Process convergence and stability

A Recycle block and an Adjust function were used in the Flowsheet calculation. The Modified Hysim Inside-Out algorithm with adaptive damping was used according to a recommendation by Øi (2012). This is the algorithm to solve the material, enthalpy and equilibrium equations in a column simulation. Stable convergence is especially important when running a Case study in Aspen HYSYS for the purpose of optimization.

3. Cost estimation procedures and assumptions

3.1. Equipment dimensioning and assumption

The equipment was dimensioned using the mass and energy balances obtained from the process simulations as done in previous works (Øi *et al.*, 2021).

The absorber was specified to have 10 stages and the desorber 6 stages (Aromada and Øi, 2017). Each stage of both columns was assumed to be 1 meter (Aromada *et al.*, 2020). A constant stage (Murphree) efficiencies of 0.15 and 0.5 were specified for the absorber and the desorber respectively (Aromada *et al.*, 2022b). Structured packing was specified for both the absorber and desorber. Superficial gas velocity of 2.5 m/s was applied to estimate the diameter of the absorber (Øi *et al.*, 2020). For the desorber, the desorber was evaluated using a superficial gas velocity of 1 m/s (Park and Øi, 2017). The tangent-to-tangent shell height of the absorber was specified to be 40 meter (Aromada *et al.*, 2022a). A tangent-to-tangent shell height of 15 m was used for the desorber. The height in both columns were necessary to account for distributors, water wash packing, demister, gas inlet, outlet and sump.

The sizing of the direct contact cooler and the flash tank were based on Souders Brown's equation with k-parameter 0.15 and 0.075 respectively. The heat exchange equipment units were dimensioned based on the heat exchange areas calculated from the heat duties. The overall heat transfer coefficient of 500 W/(m²K), 800 W/(m²K), 800 W/(m²K), and 1000 W/(m²K) were specified for the lean/rich heat exchanger, lean amine cooler, reboiler and condenser respectively (Aromada *et al.*, 2020).

The pumps were specified as centrifugal pumps with 75 % adiabatic efficiency. They were sized based on flow rate and duty.

3.2. Capital cost estimation methods

The capital costs were estimated using the Enhanced Detailed Factor (EDF) method (Ali *et al.*, 2019; Aromada *et al.*, 2021). The total capital cost is the sum of all the equipment installed costs. The EDF updated installation factors for 2020 by Nils Eldrup was used (Aromada *et al.*, 2021).

A traditional factor method for cost estimation is based on a table of factors multiplying the purchased cost of each type of equipment unit. In a detailed factor method, the total factor for each type is the sum of contributions from e.g. installation, electrical, instrumentation, administration etc. In the EDF method, these detailed factors are also dependent on the size and cost of the purchased equipment, so that the factors may change from one iteration to the next in an optimization procedure.

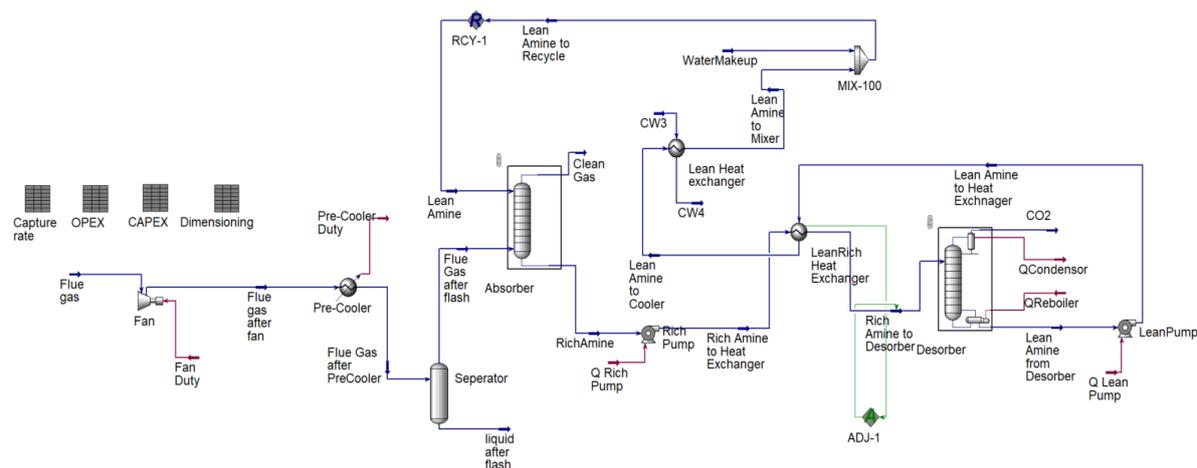


Figure 2: Aspen HYSYS flow-sheet of the base case simulation (Pouya, 2021).

The cost of each equipment was obtained from Aspen In-plant Cost Estimator (v.10), based on the sizes estimated. The cost currency is in Euro (€) and the cost year is 2016. The default location for Europe, Rotterdam was selected in this study. Stainless steel (SS316) was specified for all equipment. Welded equipment has a material factor of 1.75. The seamlessly manufactured equipment, the pumps and fan have a material factor of 1.3 (Øi *et al.*, 2021).

Since this work is aimed at automatic cost estimation, the capital cost is initially estimated based on equipment costs obtained from Aspen In-Plant Cost Estimator. Subsequent equipment cost is then estimated automatically based on the Power law. The Power law is based on the assumption that the cost ratio of two sizes of a unit is proportional to the dimension ratio raised to a power factor (typically 0.65). The installation factors and subfactors are also automatically implemented in each simulation iteration by the aid of the VBA code which connects Aspen HYSYS spreadsheets with Microsoft Excel spreadsheet.

The economic assumptions used for the capital cost estimation are summarized in Tab. 2 (Rahmani, 2021).

Table 2: Cost calculation specifications (Rahmani, 2021).

Parameter	Value
Plant lifetime	20 years
Discount rate	8.5 %
Maintenance cost	4 % of installed cost
Electricity price	0.06 Euro/kWh
Steam price	0.015 Euro/kWh
Annual operational time	8000 hours
Location	Rotterdam

Currency exchange rate 2016	9.21
Cost index 2016	103.6
Cost index September 2020	111.3

3.3. Operating cost estimation and assumption

The annual operating cost in this work was limited to cost for consumption of steam, electricity, cooling water and an annual maintenance cost as done in (Aromada and Øi, 2017). The electricity consumption was based on the pump duties obtained from Aspen HYSYS. Similarly, the steam consumption was based on the reboiler steam duty in kW. The annual hours of operation were assumed to be 8000 hours/year. The annual maintenance cost was specified as 4 %.

3.4. Net present value

The cost metric in this work for cost optimization is negative net present value (NPV) as done in (Haukås, 2020). The NPV is the sum of investment cost and the operation cost for each year in the plant lifetime. The spreadsheet unit in Aspen HYSYS was used to calculate the detailed cost estimation of CAPEX, OPEX and NPV (net present value). For the case of optimizing the temperature difference in the main heat exchanger, the calculation could be performed effectively by using the Case Study option in Aspen HYSYS. The optimum solution can then be found by the simulation giving the lowest NPV as shown in Fig. 3. For the case of optimizing the number of absorber stages, each calculation was performed independently by specifying the number of stages in each calculation.

4. Results and Discussion

4.1. Base case cost results

For the base case, the total cost (or negative NPV) was estimated to 401 mill. EURO. This is the sum of CAPEX and OPEX for a Plant lifetime of 20 years.

4.2. Optimization of minimum ΔT approach

The minimum temperature approach optimization for the process is shown in Fig. 3. The absorber packing height was 15 m in these optimizations. The optimum value at 9 K can be found as the one with minimum (negative) NPV. The cost optimum minimum temperature approach has been calculated in literature to values typically between 10 and 15 K. The differences are due to different ratios between cost of heat exchangers and cost of heat.

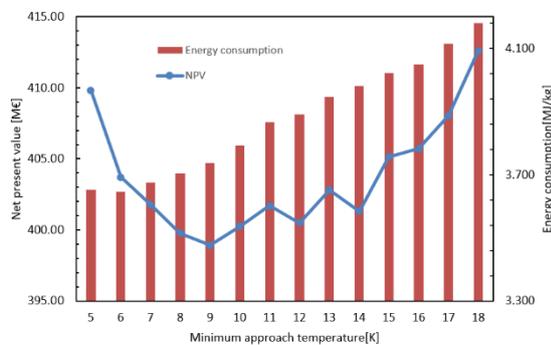


Figure 3: NPV and energy consumption as a function of ΔT_{\min} with 85% capture rate, $E_M = 0.25$, 20 years calculation period, and 8.5% interest rate.

4.3. Optimum absorber height

The results from the optimization of the absorber packing height are given in Fig. 4. It shows an optimum for 15 stages equivalent to 15 meter packing height. This is similar to results in earlier work (Kallevik, 2010; Aromada and Øi, 2017; Øi *et al.*, 2020).

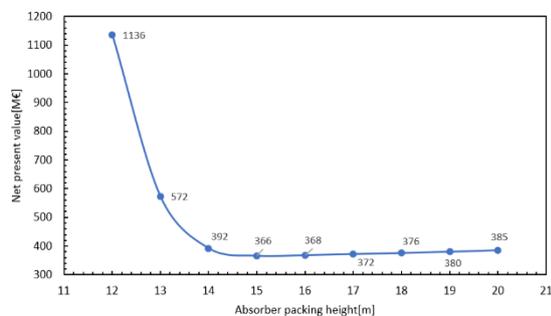


Figure 4: NPV as a function of absorber packing height with removal efficiency 85%, $E_M=0.15$, 20 years calculation period and 8.5% interest rate.

The minimum NPV when optimizing the absorber packing height is 366 mill. EURO. This is a cost reduction of 9 % compared to the base case.

4.4. Optimization of gas velocity

When the superficial velocity through the absorption column is increased, the cross section of the absorber decreases and reduces the cost, while the pressure drop increases and increases the cost. The result when varying the superficial velocity is given in Figure 5.

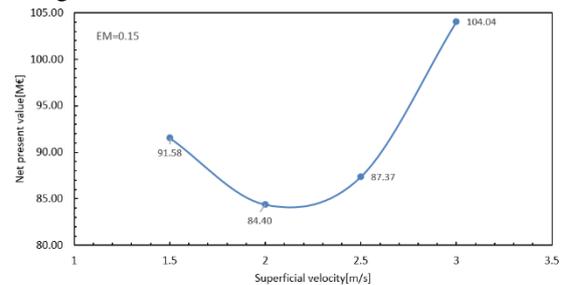


Figure 5: NPV as a function of superficial velocity with 85% capture rate, 20 years calculation period, and 8.5% interest rate and constant packing volume.

Fig. 5 shows an optimum for a gas velocity of about 2.2 m/s. This is a value close to values from Park and Øi (2017). There are not found other references in the open literature showing similar results to compare with.

5. Automation

5.1. Automation approaches

Automation of the simulations has been investigated, and results have been compared with manual simulations. Some of the input data should be changed in the simulations manually, which is time-consuming. Connecting Excel and Aspen HYSYS to transfer the data is the first step toward automating the process. In order to make the connection, one possibility is to use an Aspen simulation workbook and programming in Visual Basic. In addition, defining a case study in the Aspen HYSYS can be useful for automating the simulations.

The Aspen simulation workbook is an Excel feature that can be activated through Excel's settings. The Aspen HYSYS simulation model should be linked to Excel, under the simulation tab in the Aspen simulation workbook. Variables in the Aspen HYSYS simulation can be copied to the Aspen simulation workbook. In the scenario table, all of the input data are collected once, and the simulation runs one at a time.

ΔT_{\min} is considered as input in the lean-rich heat exchanger. The capture rate and NPV are considered as outputs. In order to fix the capture efficiency at about 85%, a controller is added to the simulation.

Visual Basic for application (VBA) programming language in Excel is another method for automating the process and cost estimation in Aspen HYSYS. In Aspen HYSYS, it can be activated in Excel from the developer tab, visual basic, tools and preference. One of the most time-consuming steps in cost estimations is determining the correct installation factor e.g. from a table. With this approach VBA was used to automatically update installation factors for next iteration based on cost calculations in previous iteration.

5.2. Automatic optimization of ΔT_{min}

In Fig. 6, sensitivity analysis is performed comparing manual calculations and a case study including updated cost factors using the Visual Basic for Application approach as explained in subsection 5.1.

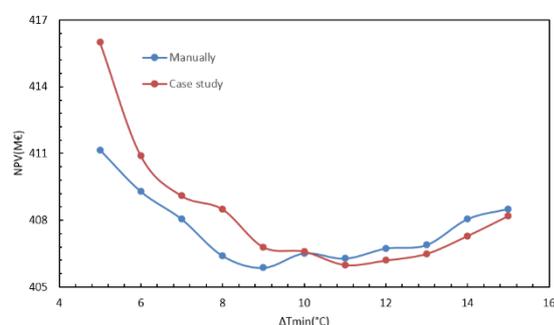


Figure 6: NPV as a function of ΔT_{min} with removal efficiency 85%, $E_M=0.25$, 20 years calculation period and 8.5% interest rate for case study (automatic) and manually, by using Aspen HYSYS model for $\Delta T_{min}=10^\circ\text{C}$.

The results are not exact equal for manual and automatic calculations. The accuracy is however reasonable, the difference in NPV is less than 1%. The calculated optimum ΔT_{min} where 9 and 11 K, respectively. The study shows that it is possible to calculate reasonable optimums automatically.

5.3. Automatic optimization of column height and gas velocity

Automatic optimization of column height (number of absorber stages) is still a challenge. In Aspen HYSYS, the number of stages is specified in the simulation input and can only be changed manually. This is also the case for pressure drop in the absorber column which can be used for gas velocity optimizations. This limitation is not necessarily a restriction in the future versions or in other tools, so this is an interesting challenge for further work.

5.4. Automatic optimization of other processes

The automatic optimization method in this work is specific for a process simulation in Aspen HYSYS. In principle, this approach could be used for any process using any process simulation tool. However, the specific challenges using Aspen HYSYS is

related to the limitations in Aspen HYSYS regarding possibilities for varying some specified parameters. Examples are the number of stages in a column and the pressure drop for a column stage which have to be specified manually.

6. Conclusion

The aim has been to calculate CO_2 capture cost optimum process parameters and evaluate the possibility of automated cost optimization using a spreadsheet facility. The adjust and recycle blocks are used to automate the energy and material balance for a specified simulation. An Excel spreadsheet is used for dimensioning and cost estimation of the specified process. New in this work is that Visual Basic for Application (VBA) was used to automatically update installation factors for next iteration based on cost calculations in previous iteration.

Equipment cost was obtained from Aspen In-plant Cost Estimator™, and an enhanced detailed factor method was used to estimate the total investment cost. Parametric studies of the absorber packing height and the minimum approach temperature in the main heat exchanger were performed at 85% capture efficiency. The cost optimum process parameters for the standard process were calculated to 15 m absorber packing height, 9 °C minimum approach temperature and 2 to 2.2 m/s superficial gas velocity through the absorber.

With this approach, iterative cost estimation and optimization of CO_2 absorption and desorption processes can be performed automatically. Automatic optimization of some parameters like the number of column stages is a challenge because they must be specified manually.

References

- Abu-Zahra, M. R. *et al.* (2007) 'CO₂ capture from power plants: Part II. A parametric study of the economical performance based on mono-ethanolamine.' *International journal of greenhouse gas control*, 1(2), pp. 135-142.
- Ali, H. (2019) *Techno-economic analysis of CO₂ capture concepts*. PhD Thesis, University of South-Eastern Norway.
- Amrollahi, Z. *et al.* (2012) 'Optimized process configurations of post-combustion CO₂ capture for natural-gas-fired power plant—Power plant efficiency analysis.' *International Journal of Greenhouse Gas Control*, 8, pp. 1-11.
- Aromada, S. A. and Øi, L.E. (2017) 'Energy and economic analysis of improved absorption configurations for CO₂ capture.' *Energy Procedia*, 114: pp. 1342-1351.
- Aromada, S. A. *et al.* (2020) 'Simulation and Cost Optimization of different Heat Exchangers for CO₂ Capture', Linköping Electronic Conference Proceedings, SIMS 61, pp. 22-24. doi:10.3384/ecp20176318
- Aromada, S. A. *et al.* (2021) 'Capital cost estimation of CO₂ capture plant using Enhanced Detailed Factor (EDF) method: Installation factors and plant construction characteristic factors' *International Journal of Greenhouse Gas Control*. 110, pp.103394.

- Aromada, S. A. *et al.* (2022a) 'Simulation-based Cost Optimization tool for CO₂ Absorption processes: Iterative Detailed Factor (IDF) Scheme', *Scandinavian Simulation Society*, pp. 301-308.
- Aromada, S. A. *et al.* (2022b) 'Cost and Emissions Reduction in CO₂ Capture Plant Dependent on Heat Exchanger Type and Different Process Configurations: Optimum Temperature Approach Analysis', *Energies*, 15(2), pp. 425.
- Haukås, A. L. (2020) *Process simulation and cost optimization of CO₂ capture Using Aspen HYSYS*. Master's Thesis, University of South-Eastern Norway, Porsgrunn.
- Ishii, Y. and Otto, F. D. (2008) 'Novel and fundamental strategies for equation-oriented process plowsheeting Part I: A basic algorithm for inter-linked, multicolumn separation processes' *Computers and Chemical Engineering*, 32, pp. 1842-1860.
- Kallevik, O. B. (2010) *Cost estimation of CO₂ removal in HYSYS*. Master's Thesis, Telemark University College, Porsgrunn.
- Kisala, T. P. *et al.* (1987) 'Sequential modular and simultaneous modular strategies for process flowsheet optimization', *Computers and Chemical Engineering*, 11(6) pp. 567-579.
- Nwaoha, C. *et al.* (2018), 'Techno-economic analysis of CO₂ capture from a 1.2 million MTPA cement plant using AMP-PZ-MEA blend', *International Journal of Greenhouse Gas Control*, 78 pp. 400-412.
- Park, K. and Øi, L. E. (2017) 'Optimization of gas velocity and pressure drop in CO₂ absorption column', Linköping Electronic Conference Proceedings SIMS 58, pp. 292-297. doi: [10.3384/ecp17138292](https://doi.org/10.3384/ecp17138292)
- Rahmani, P. (2021) *Process simulation and automated cost optimization of CO₂ capture using Aspen HYSYS*. Master's Thesis, University of South-Eastern Norway, Porsgrunn.
- Rao, A. B. and Rubin, E. S. (2002) 'A technical, economic, and environmental assessment of amine-based CO₂ capture technology for power plant greenhouse gas control', *Environmental science & technology*, 36(20) pp. 4467-4475.
- Rubin, E. S. *et al.* (2013) 'A proposed methodology for CO₂ capture and storage cost estimates', *International Journal of Greenhouse Gas Control*, 17, pp. 488-503.
- Sharma, S. and Rangaiah, G. P. (2016) 'Mathematical Modeling, Simulation and Optimization for Process Design. Chemical Process Retrofitting and Revamping, Techniques for Retrofitting and Revamping', G. P. Rangaiah, Ed., 1. ed. ProQuest Ebook Central: John Wiley & Sons, Incorporated, pp. 97-127.
- Van der Spek, M. *et al.* (2019) 'Best practices and recent advances in CCS cost engineering and economic analysis', *International Journal of Greenhouse Gas Control*, 83, pp. 91-104.
- Øi, L. E. (2021) *Removal of CO₂ from exhaust gas*. PhD Thesis, Telemark University College, Porsgrunn.
- Øi, L. E. *et al.* (2021) 'Automated Cost Optimization of CO₂ Capture Using Aspen HYSYS', Linköping Electronic Conference Proceedings SIMS 62, pp. 293-300. doi: [10.3384/ecp21185293](https://doi.org/10.3384/ecp21185293)

Machine learning techniques for modeling chemical absorption in CO₂ capture process

Huilan Zheng^a, Gaurav Mirlekar^{a,*}, Lars O. Nord^a

^a NTNU - Norwegian University of Science and Technology, Dept. of Energy and Process Engineering, Trondheim, Norway
gaurav.mirlekar@ntnu.no

Abstract

Post-combustion carbon capture (PCC) technologies play an important role in the reduction of CO₂ emissions to address climate challenges. This process is usually simulated in process simulation software based on first-principle models, which calculate physical properties directly from basic physical quantities such as mass and temperature. Using first-principle models usually requires a long computation time, which makes optimization and control difficult. In this study, machine learning algorithms, such as eXtreme Gradient Boosting (XGBoost) and Support Vector Regression (SVR), are investigated as potential alternative modeling approaches. XGBoost is an ensemble algorithm that is based on the decision tree and optimized by gradient boosting. SVR fits the best line within a predefined or threshold error value. These two algorithms are used to build models to predict the CO₂ capture rate (CR) and specific reboiler duty (SRD) in a monoethanolamine-based PCC process. By using the XGBoost, the verification result shows R² (a statistical measure that represents the fitness of the model) in predicting CR is 91.7% and in predicting SRD is 80.8%, while by using SVR the R² in predicting CR and SRD is 87.9% and 87.2% individually. In addition, XGBoost and SVR take 0.022 seconds and 0.317 seconds respectively to predict CR and SRD of 1318 cases, while the first-principal process simulation model needs 3.15 seconds to calculate 1 case. The data-driven models built using the XGBoost algorithm are employed for further optimization, which aims to find an operating point to have a higher CR and lower SRD. Particle swarm optimization (PSO), a stochastic optimization technique based on the movement and intelligence of swarms, is implemented for the optimization. The CR and SRD for optimal operating conditions are 72.2% and 4.3 MJ/kg each. The computations are faster with the data-driven models incorporated in the optimization technique. Thus, the application of machine learning techniques in carbon capture technologies is demonstrated successfully.

Keywords: Post-combustion carbon capture; Machine learning; Optimization

1. Introduction

The world has experienced global warming due to greenhouse gas emissions. The temperature difference between global mean surface temperature in 2020 and the pre-industrial baseline (1850-1900) reaches 1.2°C [1]. *Net zero emissions* is proposed to avoid the worst climate impacts. It refers to the balance between the amount of produced greenhouse gas and the amount removed from the atmosphere. Carbon capture and storage (CCS) is one of the technologies to achieve net zero emissions. CCS is the process of capturing CO₂ before it enters the atmosphere, transporting and storing it for centuries or millennia. Three technology routes are usually discussed in the CO₂ capture: pre-combustion capture, oxy-fuel combustion capture, and post-combustion capture (PCC).

PCC is a process to capture CO₂ from flue gas generated after burning the coal, oil, and gas before transportation, and it is the most commonly used carbon capture technology. Therefore, this study focused on the PCC process. The most advanced technology used in PCC technologies is chemical absorption followed by the thermal-stripping route. In the absorption process, CO₂ from the gas stream is captured by an absorbent solvent. While in the stripping process, pure CO₂ is released and the absorbent solvent is regenerated [2]. The typical solvents for absorption processes are amines, such as monoethanolamine (MEA) and diethanolamine (DEA).

In the literature, the PCC process is widely studied with

the assistance of simulations by researchers to analyze the process behavior. In particular, several methods have been applied to improve prediction performance. Xiaobo Luo et al. developed an accurate rate-based steady-state model for the MEA-based carbon capture process and validated it against thermodynamic and physical properties calculations over a wide range of pressures, temperatures, and CO₂ loadings [3]. A case study was then performed to capture CO₂ from a 250 MWe combined cycle gas turbine (CCGT) power plant to achieve shorter packing height and lower specific duty. Rohan Dutta et al. used reduced stage efficiencies in an equilibrium-based absorber model to predict operating conditions within an accepted range [4]. This modification reduces the computational time for simulation. Typically, these processes are simulated by using mass and energy balance equations to calculate physical properties such as mass flow rate and temperature in a computer software environment, e.g., Aspen HYSYS®, Aspen Plus® or gPROMS® [5][6][7]. However, the process simulation models based on the first principle require considerable computational time for solving complex equations and thus pose challenges in the implementation of advanced optimization and control techniques. To overcome the challenges brought by the first-principle models, the application of machine learning techniques has also been investigated for various CO₂ capture processes. For example, the effectiveness of the machine learning techniques in recognizing high-performing metal-organic framework materials

for CO₂ capture has been proved in the past [8]. In addition, the applicability of these advanced methods can also be seen in ionic liquid-based CO₂ capture for prediction of structure-property relationships between molecular structures of cations and anions and their CO₂ solubilities in comparison with the quantum chemistry based COSMOtherm predictions [9]. Furthermore, improvement in absorption and regeneration-based carbon capture processes and opportunities for machine learning methods are investigated in the literature [10]. Specifically, the opportunities associated with reinforcement learning to get the optimal operational parameters by using data from software simulation and pilot plants are highlighted. Abdelhamid Shalaby, AliElkamel et al. developed machine learning approaches to predict the outputs of the PCC process simulation in gPROMS[®] using Matérn Gaussian process regression (GPR), rational quadratic GPR, squared exponential GPR models, and feed-forward artificial neural network model [11]. These approaches were able to forecast the system's energy requirement, capture rate, and the purity of the condenser outlet stream with artificial neural network showing higher accuracy. Fei Li, Jie Zhang et al. used gPROMS[®] to simulate the PCC process and collect data and applied the bootstrap aggregated extreme learning machine and bootstrap aggregated neural networks to predict capture rate [12]. In this case, the BA-ELM was demonstrated as a powerful tool due to smaller mean-square error (MSE) and less computational time.

The challenge of long simulation times for first-principles models remains. During the optimization, the optimal value is usually settled after many searches, if the search time can be reduced, the optimization efficiency could be improved. Therefore, a time-efficient model is highly desirable. Although Abdelhamid Shalaby, AliElkamel et al., Fei Li, Jie Zhang et al. have demonstrated the application of machine learning in the CO₂ capture process, the application of eXtreme Gradient Boosting (XGBoost) and support vector regression (SVR) in steady-state PCC process simulation is not studied yet. This study aims to fill the research gap.

In the past, the application of data-driven models for the optimization of the energy system is studied [13][14]. In particular, the Autoregressive model with exogenous inputs (ARX) method is used for deriving the simplified dynamic model and employed in the biologically inspired optimal control strategy (BIOCS) for implementation on a subsystem of a CO₂ capture process associated with an integrated gasification combined cycle (IGCC) power plant [14]. In this study, a systematic methodology is proposed to implement computationally efficient machine learning techniques in CO₂ capture process. XGBoost and SVR techniques are employed using data samples generated by process simulation to summarize the characteristics of the data sets and establish data-driven models. The main tasks are to predict CO₂ capture rate (CR) and specific reboiler duty (SRD) in the PCC process. Consequently, the developed data-driven models are incorporated into an optimization. In the present paper, particle swarm optimization (PSO) which iteratively improves an alternative solution for a given measure of quality is used as an optimization routine. It is a heuristic global optimization method extensively employed in mathematics and computer science for solving problems more quickly when classic methods are too slow or for finding an approximate solution when classic methods fail to find any exact solution. It has

been used in engineering optimization such as, CCS cost and revenue optimization and the optimization of CO₂ solubility predicting model [15][16]. The optimization goal for this study is to find operating conditions with a relatively high CR and low SRD, which would be beneficial for improving energy efficiency during the process operation.

The paper is organized as follows: the methodology is described in section 2; section 3 discusses the results; and in section 4, the summary and conclusions are presented. The development of the PCC process simulation model, the theory of XGBoost and SVR algorithms, and the principle of PSO are introduced in the next section.

2. Methodology

In this section, the developed methodology for the implementation of machine learning techniques in the PCC process is described. The proposed framework of the work is shown in Figure 1. A steady-state process model is built in Aspen HYSYS[®] to simulate the PCC process. The data collected from this model is used as the raw data to build a data-driven model. After the data collection, the XGBoost and SVR algorithms are applied to build data-driven models. Coefficient of determination (R²) and MSE are assessment indicators to measure model performance. Besides, a new data set is created to validate the model. The proposed data-driven models are developed by the process illustrated in Figure 1. Then the PSO is adopted to search for the optimal operating conditions in optimization and control.

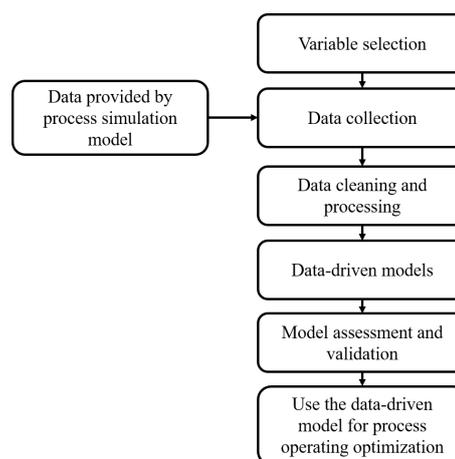
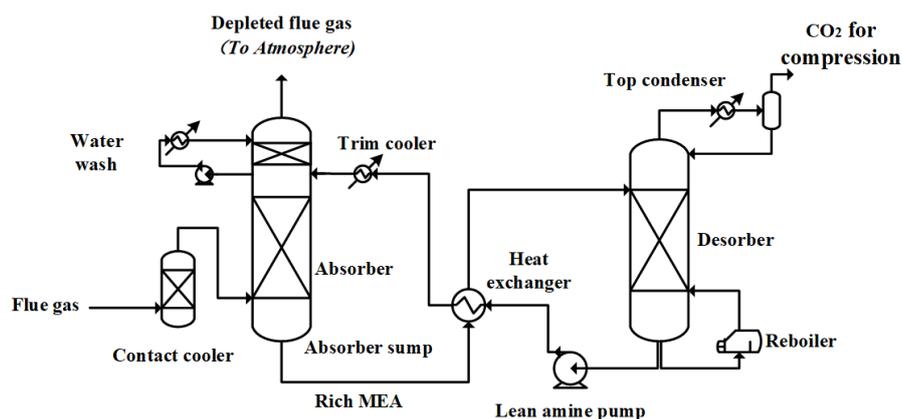


Figure 1: Methodology to implement machine learning technique in PCC process.

2.1. PCC process simulation model

The process flow diagram of the PCC steady-state simulation used for demonstration purposes is shown in Figure 2. This process is a two-step regenerative process, one is the absorption chemical process with solvent, and the other one is the desorption of CO₂ from the solvent and generation of the lean solvent. The flue gas enters the absorber from the bottom and encounters the lean amine which is going down in the column. Depleted gas and rich amine leave the absorber from the top and bottom. Rich amine is pumped, heated up by lean amine from the stripper, then enters the regenerator. In the regenerator, the stripping vapor goes up and condenses at the cooler and condenser, which are on the top of the column. The condensate containing the regenerated solvent is recycled back to the regenerator column. The CO₂ separated is sent

Figure 2: Schematic of chemical absorption process for post-combustion CO₂ capture

to the compressor and further processed for transportation or storage. Lean amine is recycled back to the absorber via a heat exchanger and a cooler.

Based on the selected independent variables, the control variable's value is changed in a step-wise manner while the rest variables are held constant. The corresponding output data information is collected and stored. A base case is defined with main parameters which are listed as Table 1 and Table 2 show.

Table 1: Main parameters in base case

Stream	Flow rate (kg/s)	Temperature (K)	Pressure (bar)
Lean amine	0.642	313.7	1.703
Flue gas	0.158	332.4	1.033

Table 2: Main parameters in base case

Stream	Value
CO ₂ molar fraction in flue gas (-)	0.1666
Lean amine loading (-)	0.2814
Rich amine loading (-)	0.4879
Capture rate (-)	0.7627
Reboiler duty (MJ/kg)	5.9580

For the lean amine stream, the compositions are H₂O, CO₂ and MEA. The main compositions in flue gas are H₂O, CO₂ and N₂. Mass fractions of lean amine stream and molar fraction of compositions in the flue gas are listed in Table 3.

Table 3: Lean amine composition in base case

Compositions	Lean amine mass fraction	Flue gas molar fraction
H ₂ O	0.6334	0.0325
CO ₂	0.0618	0.1666
MEA	0.3048	-
N ₂	-	0.8009

A set of ranges for the selected variables is defined to create the data sets. As the system encounters the converged problem when the flue gas flow rate increases to 0.2623 kg/s or the lean amine flow rate decrease to 0.4520 kg/s, these two values are set to be the upper limit of the flue gas flow rate and the lower limit of lean amine flow rate. The initial objective is to collect around 1000 data samples when each variable changes, therefore, the step is set at 0.0001 for flue gas flow rate and 0.001 for lean amine flow rate. The numerical changes of flue gas and lean amine flow rate are done automatically via the script written in Python 3.8 in connection to the Aspen HYSYS® process simulation file.

Table 4: Variables range

	Unit	Lower limit	Upper limit
Flue gas flow rate	kg/s	0.1081	0.2623
Lean amine flow rate	kg/s	0.4520	2.0040
CO ₂ molar fraction in flue gas	-	0.1666	0.2264
Lean loading	-	0.0278	0.5242

The steps and number of collected data samples for each variable are listed in Table 5.

Table 5: Number of collected samples

	Step	Total sample
Flue gas flow rate	0.0001	1543
Lean amine flow rate	0.01	1553
CO ₂ molar fraction in flue gas	around 0.0004	150
Lean loading	around 0.0005	1147
Total	-	4393

Challenges encountered during data collection:

1. The values in the composition worksheet can not be changed automatically by the Python script that is connected to Aspen HYSYS®, no other scripts are found to realize the automatic filling of composition parameters.
2. While changing the CO₂ molar fraction, the total molar fractions of the flue gas stream do not sum up to 1.
3. When the molar fractions of different compositions are changed, Aspen HYSYS® would perform normalization, the step size of CO₂ molar fraction would not be exact 0.0004.

Proposed solutions to overcome these challenges:

1. Use the Python script to control the mouse and keyboard to achieve automatic filling. The mouse is set to click on a fixed position on the desktop, so the value of CO₂ and N₂ molar fraction is naturally written into the Aspen HYSYS® worksheet. But the worksheet menu of Aspen HYSYS® is not always in a fixed position when it pops up, this program fails sometimes.
2. The sum of CO₂ and N₂ molar fraction is assumed to be constant. When CO₂ molar fraction is changed, the N₂ molar fraction is changed to a value to make the total molar fraction of the flue gas stream to be 1.

- The step size of CO₂ molar fraction is set to be approximately 0.0004.

Due to these challenges and technical limits, only 150 cases are generated for the change of CO₂ molar fraction in the flue gas. Lean amine loading is manipulated by changing the CO₂ molar flow rate in the lean amine stream. The method is the same as the mouse and keyboard control in changing CO₂ molar fraction in the flue gas stream. Consequently, the step is also an approximation. By controlling and changing values of different variables, 4393 cases are simulated to produce the raw data.

2.2. Machine learning techniques

2.2.1. eXtreme Gradient Boosting (XGBoost)

Ensemble machine learning technique is used to combine several base models for building one optimal predictive model. XGBoost stands for "eXtreme Gradient Boosting", an ensemble algorithm that is based on the decision tree and optimized by gradient boosting [17].

The following introduction starts from the decision tree. The evolution route is shown in Figure 3.



Figure 3: Schematic of evolution from decision tree to XGBoost.

Decision tree mainly consists of the root node, decision nodes, and leaf nodes. The root node is the start point. The decision node is a judgment condition, a question like "lean amine loading is larger than 0.2 or not", "flue gas flow rate is larger than 0.18kg/s or smaller than 0.1kg/s", different answers lead to the different nodes in the next level. It may enter another decision node to cross a new judgment, or go to the leaf node, which is the end of the prediction process. The leaf node is the final prediction result.

Bagging, stands for "Bootstrap aggregating", is an ensemble learning method. When bagging is used, a certain amount of samples can be randomly selected from data sets with replacement. The average or most voted result is the final result. A sample can be chosen more than one time in bagging. It is a good way to reduce variance within noisy data sets.

Random forest is a bagging-based algorithm. In a decision tree, some characteristics are chosen to be the judgment condition in a decision node. However, in a random forest, characteristics can be randomly picked to form different decision trees. All trees gather together to become a random forest, the prediction result of the random forest comes from the average or most voted of all tree results.

Boosting is a set of ensemble algorithms that can help convert weak learners (refer to models) to strong learners. The boosting would train a basic learner from training data, then focus more on the wrong prediction samples, and correct errors from the first learner to generate the second learner. Repeat this "correct" work step by step, until a strong learner is trained.

Gradient boosting is one kind of boosting. As the loss function represents the unreliability of the model, gradient

boosting adjusts the model towards a gradient descent direction of the loss function from the previous model.

XGBoost is a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting framework. It takes the bootstrap sample 1 to build model 1, then takes the bootstrap sample 2 to build model 2, which is more advanced than model 1. Then iterate to get the final model and result. It minimizes a regularized objective function:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (1)$$

where $\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$.

$L(\phi)$ is the objective in this minimization problem. It is the summation of convex loss function represented as $\sum_i l(\hat{y}_i, y_i)$ and regularization item denoted as $\sum_k \Omega(f_k)$. Here, y_i and \hat{y}_i stands for targeted value and predicted value respectively. λ and γ are hyperparameter constants. T is the number of leaf nodes. w is the predicted value of the leaf node.

2.2.2. Support vector regression (SVR)

An SVR model gives users the freedom to decide how much error is accepted. A line or a hyperplane (when the data has higher dimensions) is searched to fit the data. The goal is to minimize the coefficients, the l_2 -norm of the coefficient vector, rather than the squared error which is usually used in linear regression [18]. The constraint would be:

$$|Y_i - \alpha X_i| \leq \varepsilon \quad (2)$$

The goal is to minimize:

$$\text{MIN} \frac{1}{2} \|\alpha\|^2 \quad (3)$$

Here, Y_i stands for the targeted value, X_i represents the feature value, α denotes coefficients. In SVR models, kernel functions can be used to transform input data to the required form of processing data. There are different kinds of kernels: Gaussian Kernel, Radial Basis Function, Sigmoid Kernel, and Polynomial Kernel.

2.3. Optimization algorithm

2.3.1. Particle swarm optimization (PSO)

PSO is a search optimization technique inspired by the migration behavior of birds. Assume there are N particles in a swarm. These particles are subject to random uniform initialization, and they have random positions and velocities in a D-dimensional search space. These particles move at a certain speed to find the best position inside the whole space. For each particle, the new velocity is updated based on its own historical experience and the group experience. Assume the D-dimensional position vector of the i -th particle is:

$$S_i = (S_{i1}, S_{i2}, S_{i3}, \dots, S_{iN}), i = 1, 2, \dots, N \quad (4)$$

The velocity vector of i -th particle is:

$$V_i = (V_{i1}, V_{i2}, V_{i3}, \dots, V_{iN}), i = 1, 2, \dots, N \quad (5)$$

The equation to update the position and velocity of each particle is:

$$v_{id}^{k+1} = mv_{id}^k + c_1 r_1 (p_{id}^k - s_{id}^k) + c_2 r_2 (p_{gd}^k - s_{id}^k) \quad (6)$$

$$s_{id}^{k+1} = s_{id}^k + r v_{id}^{k+1} \quad (7)$$

This speed update equation of the PSO algorithm is the sum of three parts. The v_{id} and s_{id} are the velocity and position of the i -th particle in the d -th dimension, k and $k+1$ represent the current and next iterations. First part $m v_{id}^k$ is the exploration, m is an inertia factor of fixed value. Second part $c_1 r_1 (p_{id}^k - s_{id}^k)$ is self-learning, and the third part $c_2 r_2 (p_{gd}^k - s_{id}^k)$ is group learning. c_1 and c_2 are learning factors, r_1 and r_2 are random numbers within the range $[0, 1]$. p_{id} and p_{gd} are the best positions searched by the i -th particle and the whole group so far.

In the next section, the data analysis result, the prediction performance of built models, and the optimization result are shown and discussed.

3. Results

3.1. Data analysis and model development

The purpose of data analysis is to initially have an overview understanding of the data sets by observing the connections between variables. The Pearson correlation coefficient, which is a measure of linear correlation between two sets of data, is used in data analysis. Given paired data set $(Z_{11}, Z_{21}), (Z_{12}, Z_{22}), \dots, (Z_{1n}, Z_{2n})$, the formula of Pearson correlation coefficient R_{Z_1, Z_2} is:

$$R_{Z_1, Z_2} = \frac{\sum_{i=1}^n (Z_{1i} - \bar{Z}_1)(Z_{2i} - \bar{Z}_2)}{\sqrt{\sum_{i=1}^n (Z_{1i} - \bar{Z}_1)^2} \sqrt{\sum_{i=1}^n (Z_{2i} - \bar{Z}_2)^2}} \quad (8)$$

where \bar{Z}_1 and \bar{Z}_2 are the average values of variables Z_1 and Z_2 . Figure 4 shows Pearson correlation coefficients between different variables in this study.

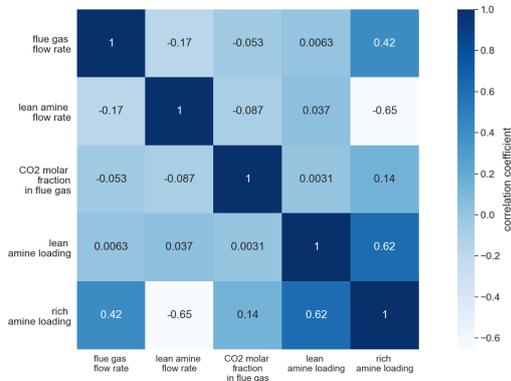


Figure 4: Plot of data correlation coefficients

The number on the grid is the correlation coefficient of the corresponding horizontal and vertical axis variable. The darkness or grid color shade is directly proportional to the correlation coefficient. For example, the grids which are intersected by lean amine flow rate and the flue gas flow rate are light blue, and the value of the correlation coefficient is -0.17. This indicates that they are slightly inverse related. The correlation coefficient between rich amine loading and lean amine loading is 0.62, which shows they might have positive relationships. Thus, rich amine loading is not included as an input variable. There may be some non-linear relations between different variables that are not included in the correlation coefficient. The purpose of the data analysis is to get a general understanding of data sets, thus, non-linear relations are not considered here.

The relation between dependent variables and independent variables can be further investigated in Figure 5. In the following section, *To atmosphere*, *CO₂ for compression*, *Rich MEA* refer to gas stream out from absorber, captured CO₂ stream out from desorber, loaded amine stream out from absorber in Aspen HYSYS[®] simulation model as shown in Figure 2.

Figure 5-(a) shows the relationship between the capture rate and the flue gas flow rate of 1543 data samples. The highest point represents a case in which the flue gas flow rate is 0.1081kg/s and the capture rate is 0.8738, the lowest point represents a case the flue gas flow rate is 0.2623kg/s and the capture rate is 0.5557. Capture rate decreases as flue gas flow rate increases. As the flue gas mass flow rate increases, more CO₂ enter the absorber in a time unit. Due to the absorbing solvent limitations and increased inflow of CO₂, the capture rate is decreased and more CO₂ is released into the atmosphere through *To atmosphere* stream. Consequently, the capture rate becomes smaller. When the flue gas flow rate is close to 0.2623, the capture rate value starts fluctuating, this could be caused by the model instability. In Figure 5-(b), the SRD decreases as the flue gas flow rate increases. The reasons behind this are the increase of the CO₂ in *CO₂ for compression* stream and the decrease of the reboiler duty in the regenerator. SRD is calculated by the reboiler duty divide mass of CO₂ in *CO₂ for compression*, therefore, SRD decreases. When the lean amine flow rate increases, the CO₂ enter the rich stream increases. Thus, as Figure 5-(c) shows, the capture rate increases with the lean amine flow rate. In Figure 5-(d), when the lean amine flow rate is increased, the CO₂ in the *To atmosphere* stream becomes less and CO₂ in the *rich amine* stream increased, which increases the CR. Reboiler duty increases since the stream entering the regenerator include more MEA. Accordingly, the SRD increases. The increase of CO₂ molar fraction in flue gas results in the increase in CO₂ in the stream that is released into the atmosphere and *rich amine* stream, but the increase in *To atmosphere* is larger while the increase in *rich amine* stream is smaller, consequently, the CR is lower. The reboiler duty of the regenerator is lower, resulting in a lower SRD. Lean amine loading is changed by adjusting the mass flow rate of CO₂ in lean amine. When increases lean amine loading, CO₂ in the *To atmosphere* stream increases, and the capture rate decreases. Reboiler duty also decreases and accordingly SRD is lower.

It is observed that SRD has fluctuations quite often. It is caused by the unstable reboiler duty. Data smoothing is suggested to deal with these instabilities. There are many ways to smooth data, such as simple exponential, moving average, exponential moving average, and Holt-Winters smoothing. The models developed by using smoothed didn't show an expected improvement. Therefore, to keep the original information of data, all the following models are developed based on original data without smoothing.

3.2. Modeling and validation

All data sets are divided into training data sets and test data sets with a division ratio of 7:3. After building models based on training data sets, the predicted values based on independent variables in test data sets are compared with the value of dependent variables in test data sets. Cross-validation score, MSE, RMSE, R² are calculated to assess model performance. Using XGBoost to predict capture rate had a cross-validation average score of 0.9995. The R² was 0.9996 and the MSE was 0.0000. The difference

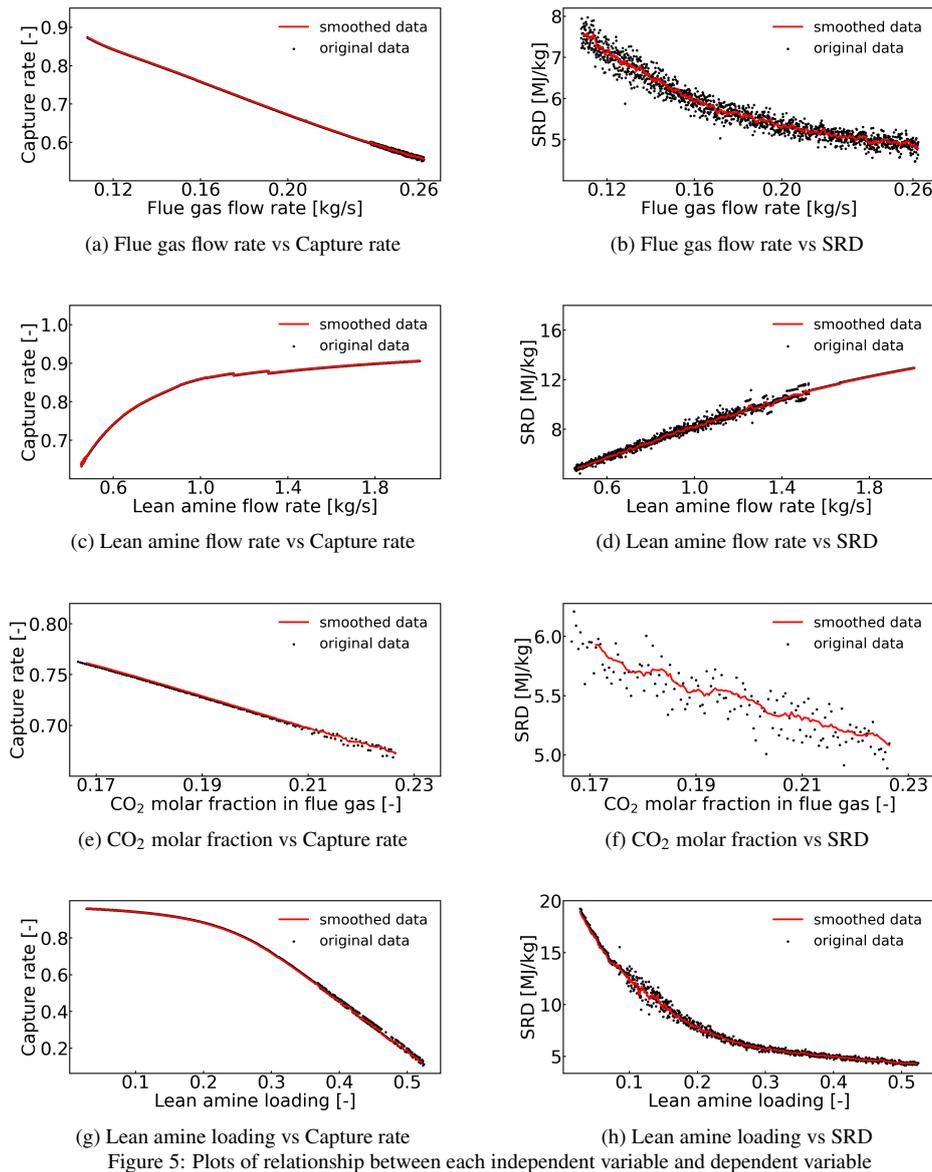


Figure 5: Plots of relationship between each independent variable and dependent variable

between predicted CR and tested CR is shown in Figure 6-(a). Most points were distributed around line $y=0$, which indicated that there are no huge deviations in prediction, and the model has a good prediction accuracy. It is shown that the accuracy of the XGBoost model is quite high. To validate this accuracy, cross-validation was used to avoid overfitting. And the SVR algorithm was applied to build the other model. If SVR shows a different accuracy level, the qualities of the models can be compared. The XGBoost and SVR algorithms are employed to predict capture rate and SRD respectively, therefore, 4 models are built. The performances of each model can be seen from Figure 6.

In each subplot is the scatter plot of the difference between the predicted value and the real value. It can be observed that when using XGBoost to predict SRD, the errors are smaller. SVR may not be a suitable method to predict CR. The error-represented points are randomly distributed in the different positions instead of scattering around line $y=0$.

3.3. Results and verification

Table 6 summarize the performance of different models:

Table 6: Summary table		
XGBoost		
assessment	predict CR	predict SRD
cross-validation score	0.9995	0.9934
R-squared(R²)	0.9996	0.9930
MSE	0.0000	0.0595
RMSE	0.0032	0.2439
SVR		
assessment	predict CR	predict SRD
cross-validation score	0.8573	0.9620
R-squared(R²)	0.8418	0.9622
MSE	0.0041	0.3231
RMSE	0.0643	0.5684

A verification data set of 78 data samples are generated to verify the model. This data set is different from training data sets and test data sets but within the same range as shown in Table 4. Note that the verification of the developed model can be improved with higher number

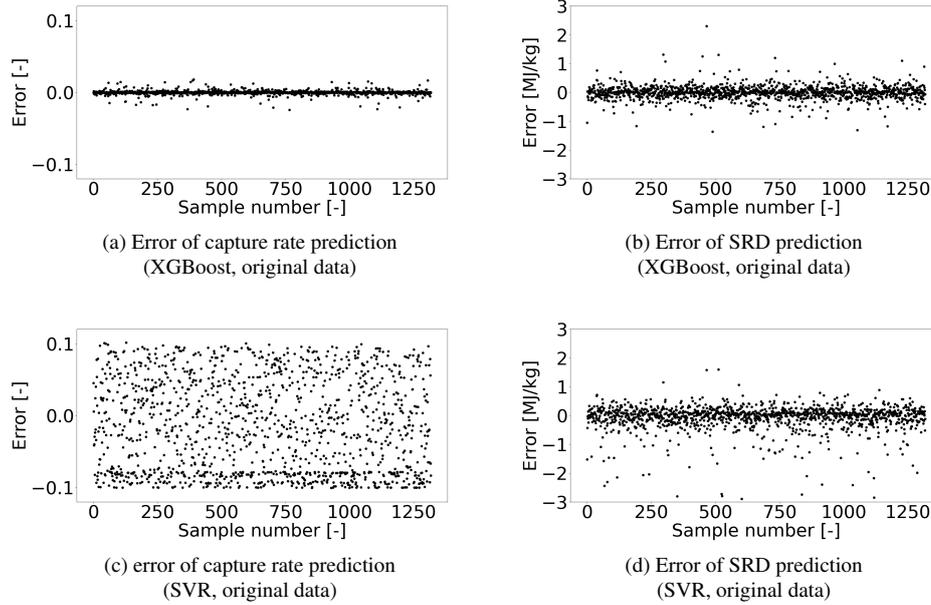


Figure 6: Plots of errors of model prediction: (a)XGBoost model to predict CR; (b)XGBoost model to predict SRD; (c)SVR model to predict CR; (d)SVR model to predict SRD.

of data samples. However, the number of data samples mentioned above are restricted due to the operational range in which this model is developed. The verification results of models are shown in Table 7 :

Table 7: Verification result

XGBoost		
assessment	predict CR	predict SRD
R-squared(R²)	0.9170	0.8077
MSE	0.0029	0.4043
RMSE	0.0541	0.6358
SVR		
assessment	predict CR	predict SRD
R-squared(R²)	0.8793	0.8716
MSE	0.0042	0.2700
RMSE	0.0653	0.5196

XGBoost models are efficient as they showed verification accuracies of 91.7% to predict CR and 82.82% to predict SRD. These high values of R-squared indicated that models have generalization ability. The accuracy of the developed model is close to the accuracy of the simulation within the range of process variables in which this model is developed. The precision of the developed model in all cases or outside the range is a subject of investigation. The model is available when the input variables are within the ranges shown in Table 4. For the values of the variables outside the ranges, the model's availability is not investigated. Some observations can be discussed: (1) In a general machine learning modeling process, the variables selection is a step after the data collection. During the data correlation and visualization step, data should be visualized to explore the trends or relations between various variables. Then the variables selection is implemented. In this order variables that are completely irrelevant or replaceable can be avoided. However, no data are pre-provided in this study, and it is hard to generate large data sets as all data are generated by Aspen HYSYS[®]. Thus the variables are chosen mainly based on some research results and previous experience, then data collection is executed.

(2) Some curves in Figure 5 have fluctuations (Figure

5-(b)) and breakpoints(Figure 5-(c)). That is because the steady-state simulation is calculated based on first-principle models which often include several differential and algebraic equations. The model is unstable sometimes, consequently, the same inputs may lead to a bit different outputs.

(3) Although the number of raw data samples is more than 4000, the steps between samples are quite small, which may not differentiate much. More data samples could help increase the performance of the model or lead to higher prediction accuracy in verification.

(4) XGBoost algorithm has been shown to have high accuracies in a lot of applications. In this study, training and test data sets are synthetic data generated by the steady-state simulation. This simulation has a certain mathematic expression, making the R-squared of the model 99% possible. The other reasons for this high accuracy lie in the complexity of the prediction task, the amount of training data sets, and the possibility of overfitting.

3.4. Optimization

In this section, optimization of the process operation conditions is discussed. From the operational point of view, the goal is to maximize the CR and minimize the SRD as much as possible. However, the general rule is that the smaller SRD is usually associated with lower CR. Therefore, it's important to make a trade-off. To show this in mathematical form, the objective goal can be viewed as:

$$f(x) = \frac{SRD(x)}{SRD_{upperlimit}} - \frac{CR(x)}{CR_{upperlimit}} \quad (9)$$

where the SRD(x) and CR(x) are the SRD and CR predicted by the XGBoost model based on the vector x , which represents the operation conditions. These two outputs are then scaled, as the CR and SRD are in different ranges and could not be evaluated at the same level without scaling.

The process simulation model needs 3.15s to predict CR and SVR for 1 case. However, it only takes 0.022s and

0.317s for XGBoost and SVR model to predict CR and SVR of 1318 cases. The computation time is drastically reduced, which would be beneficial for optimization implementation. With a swarm of 10 particles, after 1500 times iteration, the minimal value of fitness $f(x)$ is -0.5299. The operating conditions are: flue gas flow rate of 0.18 kg/s, the lean amine flow rate of 0.46 kg/s, CO₂ molar fraction in the flue gas of 0.2002, lean amine loading of 0.3085, and corresponding CR and SRD is 72.2% and 4.3 MJ/kg respectively. Compared with the base case which has a CR of 76.3% and SRD of 5.9 MJ/kg, although the CR was 4.1% lower, the SRD decreased by 1.7 MJ/kg, which accounts for 28.2% of 5.9 MJ/kg.

4. Conclusions

In this paper, the application of machine learning techniques such as XGBoost and SVR for the PCC process model simulation was demonstrated successfully. The energy efficiency indicators and parameters associated with the PCC process model were identified, then machine learning algorithms were applied to build models to predict CR and SRD. The models were used in optimization, and adequate operation conditions are characterized. The data-driven models showed high accuracy in predicting the capture rate and energy requirement in the reboiler of the PCC model. The XGBoost model had the accuracy of 91.7% and 80.8% for predicting CR and SRD based on the validation data sets. The SVR model showed 87.9% and 87.2% in CR and SRD prediction. And the calculation time of 1318 cases for the XGBoost model and SVR model was 0.022 seconds and 0.317 seconds. Compared with the first-principle-based process model, which needed 3.15 seconds to calculate the parameters of 1 case, the data-driven models showed improved performance in the time-efficient aspect. The goal of developing time-efficient models by machine learning techniques was achieved. Integrated the data-driven model within optimization, the ideal operating condition was flue gas flow rate as 0.18 kg/s, lean amine flow rate as 0.46 kg/s, CO₂ molar fraction in flue gas as 0.20, lean amine loading as 0.31. The corresponding CR and SRD were 72.2% and 4.3 MJ/kg individually. The CR decreases 4.1% lower and SRD dwindles 1.7 MJ/kg compared with the base case. Thus, the machine learning techniques were demonstrated useful in process optimization and advanced control methods where faster model predictions are necessary.

Acknowledgment

The authors gratefully acknowledge the Department of Energy and Process Engineering at NTNU for funding support and facilitating with computer resources and software package. The authors would also like to thank NTNU Ph.D. candidate Valentin Formont for help during the work.

References

- [1] W. M. Organization, "2020 was one of three warmest years on record."
- [2] H. F. Svendsen, E. T. Hessen, and T. Mejdell, "Carbon dioxide capture by absorption, challenges and possibilities," *Chemical Engineering Journal*, vol. 171, no. 3, pp. 718–724, 2011. doi:10.1016/j.cej.2011.01.014.
- [3] X. Luo and M. Wang, "Improving prediction accuracy of a rate-based model of an MEA-based carbon capture process for large-scale commercial deployment," *Engineering*, vol. 3, no. 2, pp. 232–243, 2017. doi:10.1016/J.ENG.2017.02.001.

- [4] R. Dutta, L. O. Nord, and O. Bolland, "Applicability and validation of use of equilibrium-based absorber models with reduced stage efficiency for dynamic simulation of post-combustion CO₂ capture processes," *Energy Procedia*, vol. 114, pp. 1424–1433, 2017. doi:10.1016/j.egypro.2017.05.048.
- [5] "Aspen HYSYS®(version 10) [software]," <https://www.aspentech.com/en>, 2021.
- [6] "Aspen Plus®(version 12) [software]," <https://www.aspentech.com/en>, 2021.
- [7] "gPROMS®(version 2) [software]," <https://www.psenterprise.com/products/gproms>, 2021.
- [8] M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji, and T. K. Woo, "Rapid and accurate machine learning recognition of high performing metal organic frameworks for CO₂ capture," *The journal of physical chemistry letters*, vol. 5, no. 17, pp. 3056–3060, 2014. doi:10.1021/jz501331m.
- [9] V. Venkatraman and B. K. Alsberg, "Predicting CO₂ capture of ionic liquids using machine learning," *Journal of CO₂ Utilization*, vol. 21, pp. 162–168, 2017. doi:10.1016/j.jcou.2017.06.012.
- [10] M. Rahimi, S. M. Moosavi, B. Smit, and T. A. Hatton, "Toward smart carbon capture with machine learning," *Cell Reports Physical Science*, p. 100396, 2021. doi:10.1016/j.xcrp.2021.100396.
- [11] A. Shalaby, A. Elkamel, P. L. Douglas, Q. Zhu, and Q. P. Zheng, "A machine learning approach for modeling and optimization of a CO₂ post-combustion capture unit," *Energy*, vol. 215, p. 119113, 2021. doi:10.1016/j.energy.2020.119113.
- [12] F. Li, J. Zhang, E. Oko, and M. Wang, "Modelling of a post-combustion CO₂ capture process using extreme learning machine," *International Journal of Coal Science & Technology*, vol. 4, no. 1, pp. 33–40, 2017. doi:10.1007/s40789-017-0158-1.
- [13] G. Mirlekar, B. Gebreslassieb, U. Diwekar, and F. V. Lima, "Biomimetic model-based advanced control strategy integrated with multi-agent optimization for nonlinear chemical processes," *Chemical Engineering Research and Design*, vol. 140, pp. 229–240, 2018. doi:10.1016/j.cherd.2018.10.005.
- [14] G. Mirlekar, G. Al-Sinbol, M. Perhinschi, and F. V. Lima, "A biologically-inspired approach for adaptive control of advanced energy systems," *Computers Chemical Engineering*, vol. 117, pp. 378–390, 2018. doi:10.1016/j.compchemeng.2018.07.002.
- [15] H. Saboori and R. Hemmati, "Considering carbon capture and storage in electricity generation expansion planning," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1371–1378, 2016. doi:10.1109/TSTE.2016.2547911.
- [16] N. A. Menad, A. Hemmati-Sarapardeh, A. Varamesh, and S. Shamshirband, "Predicting solubility of co₂ in brine by advanced machine learning systems: Application to carbon capture and sequestration," *Journal of CO₂ Utilization*, vol. 33, pp. 83–95, 2019. doi:10.1016/j.jcou.2019.05.009.
- [17] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016. doi:10.1145/2939672.2939785.
- [18] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004. doi:10.1023/B:STCO.0000035301.49549.88.

Comparison of Machine Learning Approaches for Spectroscopy Applications

Ioanna Aslanidou ^{a,*}, Jerol Soibam ^a

^a *Mälardalen University, Box 883, 72123, Västerås, Sweden*

* *ioanna.aslanidou@mdu.se*

Abstract

In energy production the characterization of the fuel is a key aspect for modelling and optimizing the operation of a power plant. Near-infrared spectroscopy is a well-established method for characterization of different fuels and is widely used both in laboratory environments and in power plants for real-time results. It can provide a fast and accurate estimate of key parameters of the fuel, which for the case of biomass can include moisture content, heating value, and ash content. These instruments provide a chemical fingerprint of the samples and require a calibration model to relate that to the parameters of interest.

A near-infrared spectrometer can provide point data whereas a hyperspectral imaging camera allows the simultaneous acquisition of spatial and spectral information from an object. As a result, an installation above a conveyor belt can provide a distribution of the spectral data on a plane. This results in a large amount of data that is difficult to handle with traditional statistical analysis. Furthermore, storage of the data becomes a key issue, therefore a model to predict the parameters of interest should be able to be updated continuously in an automated way. This makes hyperspectral imaging data a prime candidate for the application of machine learning techniques. This paper discusses the modelling approach for hyperspectral imaging, focusing on data analysis and assessment of machine learning approaches for the development of calibration models.

Keywords: machine learning, near-infrared spectroscopy, hyperspectral imaging; calibration models

1. Introduction

In the energy and process industry the characterization of the feedstock material is a key aspect for modelling and optimizing the operation of a process or power plant. The operating parameters are continuously adjusted in order to provide output that meets certain specifications, which can be the quality of the end product or the power output. These are dependent on the quality of the feedstock, as a difference in its key parameters will result in different requirements for its processing. Detailed knowledge of the properties of the fuel can be used to improve the operation of the plant using feed-forward control approaches. A sensor that can provide this information is one of the foundations for a learning system that can support optimal operation and decision-making.

Near-infrared (NIR) spectroscopy is a well-established method for characterization of different fuels and is widely used both in laboratory environments and in power plants for real-time results. The method itself is based on the excitation and vibration of the molecules, which in turn provides the chemical information for the material. This needs to be correlated to the parameters of interest

of the fuel, which is typically done using statistical analysis. Near-infrared spectroscopy is one of the technologies used for the development of smart sensors in the learning system described in Rahman et al. (2021). Near-infrared spectroscopy has been widely used in the literature (Tsuchikawa et al., 2003; Skvaril et al., 2017) as it can provide a fast and accurate estimate of key parameters of the fuel, which for the case of biomass can include moisture content, heating value, and ash content.

Near-infrared spectroscopy can only provide single point measurements. This information is well suited for homogeneous mixtures, where it can provide a good estimate for the parameters of interest of the entire batch. In recent years, hyperspectral imaging (HSI), which combines spectral information and conventional imaging, is also increasingly used for fuel characterisation. This allows the collection of spatial data for the near-infrared spectrum and can be applied in real environments (e.g. above a conveyor belt) to provide real-time information about the fuel. The simultaneous acquisition of spectral and visual information without the need for synchronization is another advantage, which also makes the use in real environments more realistic. In hyperspectral imaging, the instrument

can acquire images of the sample, as well as spectral information for each pixel, providing a hypercube of data. This can provide multiple opportunities, as discussed by Mäkelä and Geladi (2017), who used HSI to distinguish different materials (from different feedstocks or prepared under different temperatures) and evaluate their homogeneity. Another application of HSI in biomass characterization is for pelleting of biomass feedstocks where the spatial resolution allows the classification of images to assess the efficiency of the mixing of different biomass streams (Gillespie et al., 2016).

Regardless of the method or instrument used to acquire spectral information for the samples, a model is required to correlate that information to the parameters of interest and provide quantitative information about the parameters of interest. Linear regression techniques are considered the standard for quantitative characterization, with Partial Least Squares Regression (PLSR) being the most commonly used approach (Skvaril et al., 2017). Non-linear methods, such as Artificial Neural Networks (ANNs), have been shown to improve the results but are more demanding in terms of computational power. Advances in computing power have allowed machine learning techniques to be used to extract information from spectral data, and recent literature presents results from different applications.

Machine learning in combination with IR spectroscopy has been widely used for classification purposes. Mancini et al. (2020) used NIR to study the supply chain for biomass pellets and applied different classification algorithms to predict pellet quality. A similar approach was used by Tiitta et al. (2020) who employed electric impedance spectroscopy to classify wood chips of different origin, which can then allow the derivation of more accurate models for moisture content. Pitak et al. (2021) focused on the biomass pellet production process, using machine learning for wavelength selection and PLS regression for their calibration model. Tao et al. (2020) obtained IR spectra of biomass and waste with an attenuated total reflectance (ATR) and used ML for classification and characterization, employing regression techniques. Ahmed et al. (2018) applied different methods for the characterization of biomass wood chips using NIR, namely ANN, Gaussian Process Regression (GPR), Support Vector Regression (SVR) and traditional PLSR, with GPR showing the best results.

The use of hyperspectral imaging results in a larger amount of data than what is obtained with NIR and storage of the data becomes a key issue. The calibration model for hyperspectral imaging should be able to be updated continuously in an automated way, which makes hyperspectral imaging data a prime candidate for the application of machine learning techniques. Gewali et al. (n.d.) present a review of the literature in the use of ML for HSI, primarily for analysis of hyperspectral images captured from earth observing satellites and aircraft. They looked into techniques used for classification of images based on land cover, concluding that

deep learning is a promising approach. For the estimation of physical/chemical parameters related to agriculture, Bayesian methods were considered to be more suitable due to their flexibility, ability to handle uncertainty, and capacity to perform well with limited data.

This paper discusses the use of hyperspectral imaging and machine learning for biomass characterization. The focus is on data analysis and assessment of machine learning (ML) approaches for the development of calibration models. A comparison of different ML approaches for HSI for the prediction of biomass properties is not available in the literature, and neither is a comparison of ML with a conventional model to assess when the use of ML is beneficial in such applications. The contribution of the paper therefore lies in discussing the suitability of different methods depending on the purpose of the analysis and the type of data available, as a first step towards a thorough study of the use of ML for HSI data analysis in such applications.

2. Methodology

A set of biomass samples was analysed with a hyperspectral imaging camera and different machine learning techniques were used to create a calibration model. The following sections discuss the methods used to acquire the data, pre-processing techniques, and methods used to build calibration models, followed by validation and testing of their predictive capabilities.

2.1. Sample preparation and data acquisition

A set of 100 biomass fuel samples were used in this study. The spectral data of the samples was obtained with a push-broom line scanning hyperspectral imaging Specim FX17e camera (Specim Spectral Imaging Ltd, Finland). The camera is equipped with an InGaAs based NIR detector with spectral range of 900-1700 nm, 224 spectral bands, and 640 pixels over the cross-track field of view (FOV). The samples were illuminated with six halogen light sources of 150W and moved on a laboratory scanning table (20cm \times 40cm) at a velocity of approximately 90mm/s. The acquisition was done under constant ambient conditions, at frame rate of 300fps and an exposure time of 5ms, in order to acquire images with the correct aspect ratio. The setup for data acquisition is shown in Figure 1.

Reflectance calibration was carried out to correct for background response of the instrument. The dark reference image (D) was acquired by closing the shutter of the camera lens and white reference image (W) was obtained from a 99% reflectance ceramic tile surface. The reflectance value R was calculated from the measured signal (S) on a pixel-by-pixel basis, as shown in equation 1, where i is the pixel index.

$$R_i = \frac{S_i - D_i}{W - D_i} \quad (1)$$

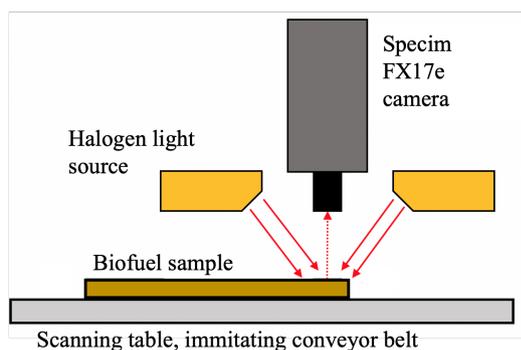


Figure 1: The experimental setup

A hyperspectral image of a biomass sample is shown in Figure 2. A set of spectral data is acquired for each pixel of the image. The spectra obtained from the camera are shown in Figure 3. It should be noted that the noisy parts of the spectra at the lower and higher wavelengths that contain no useful information have been removed in this figure and for all data before pre-treatment. The reference to no pre-treatment in the rest of this paper refers to data in which the noisy parts have been removed.

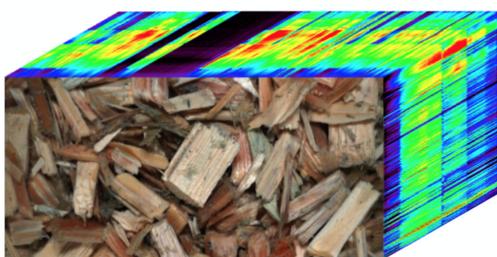


Figure 2: Hyperspectral image of a biomass sample

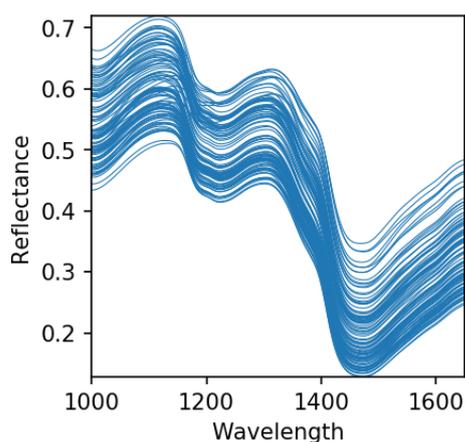


Figure 3: Spectra with no pre-treatment

The moisture content in the samples was determined according to the European standard EN ISO 18134. The samples were oven dried for 20 hours in 105°C and weighed before and after the process. The resulting moisture content range was from 31.0% to 55.8%.

2.2. Data Preprocessing

The spectral data was pretreated to enhance the differences among the samples in order to provide a better calibration model. Noisy parts of the spectrum were removed as they contain no useful information and can instead confound the model. Two different pre-processing techniques were applied: Savitzky-Golay first derivative (SG1) and Standard Normal Variate (SNV), which were shown to perform best in similar samples analysed with NIR spectroscopy Ahmed et al. (2018). The results were also compared to those obtained without any pre-treatment of the data. The pre-treated data was also scaled in accordance with the requirements of the data analysis method, using either the mean and standard deviation or the range of the dataset to obtain a range from 0 to 1.

2.3. Methods for data analysis

In this paper SVR and ANN were compared with PLSR. The different techniques were implemented in Python using the Scikit-learn module (Pedregosa et al., 2011).

SVR is an extension of the Support Vector Machines (SVM) method for classification problems to solve regression problems. It can allow the user to determine the maximum error that is acceptable in the model and find an appropriate hyperplane to fit the data. Hyperparameters C , γ , and ϵ were adjusted in order to obtain a model that can provide the best prediction. Three different kernel functions (linear, polynomial, radial basis function - RBF) were tested to allow the separation of the data and allow for a better model to be obtained. The hyperparameters were tuned using a grid search with K -fold cross-validation for all kernel functions.

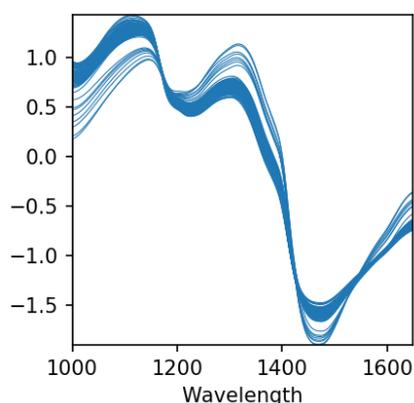
ANNs are widely used for regression problems and can provide good results when the underlying relationship between the different parameters is non-linear. A network with two hidden layers was used, with 128 nodes in the first layer and 32 nodes in the second layer. A third hidden layer was not found to improve the results, which were also similar for 32, 64, and 128 neurons in the hidden layers. The optimal learning rate was selected based on the cross-validation results and the epoch with the lowest error was selected using a callback function.

PLSR is considered the standard approach for spectroscopy applications and performs well when the underlying relationship is linear. Principal component analysis was performed to select the number of components that gives the highest prediction accuracy.

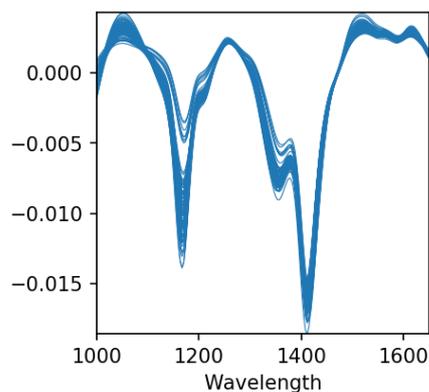
In this work the collected data was split into a training set of 80 samples and a testing set of 20 samples. K -fold cross-validation was also performed during the training to improve the prediction capability of the models and avoid over-fitting. The optimal setting for the cross-validation was found to be 15 folds using 10% of the training dataset.

3. Results and Discussion

The results of the pre-treatment are shown in Figure 4(a) for the Standard Normal Variate and in Figure 4(b) for the Savitzky-Golay first derivative. The setup for SG1 aimed to ensure enough information was retained in the spectra to provide a good model and was evaluated based on the cross-validation results for PLSR. As seen in figures 3 and 4, there are two clear dips in the spectra: one at 1180nm and one at 1430nm. These areas contain much useful information about the chemical composition of the samples. The wavelength of 1180nm is the fingerprint of the C-H stretching overtone, whereas the wavelength of 1430nm corresponds to the O-H overtone. It is this differentiation in the spectra of the different samples that can be coupled to the moisture content and be used to create a robust model.



(a) Spectra with SNV pre-treatment



(b) Spectra with SG1 pre-treatment

Figure 4: Spectra with pretreatment

The pre-treated spectra of the training set were used to build the calibration models. K -fold cross validation was employed to increase the predictive capability of the models. The models were then evaluated on an unseen test set. The results for the PLSR, SVR, and ANN regression for the different pre-treatment approaches are summarized in Table 1. The evaluation metrics used are the goodness of fit measure for linear regression R^2 and the root mean square error, $RMSE$, both for

the prediction of the unseen test set. As seen in the results, both the the Savitzky-Golay 1st derivative (SG1) and the SNV pre-treatment methods provide an improvement in terms of fit and error for all modelling approaches.

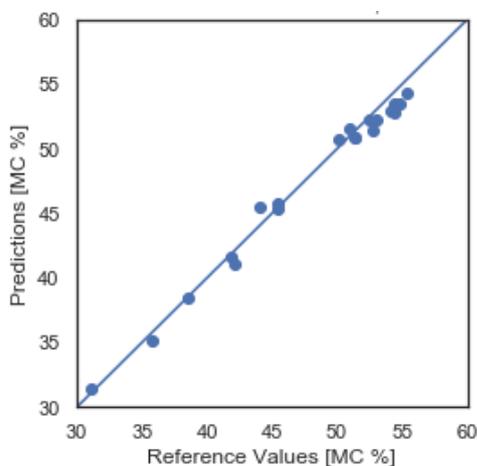
Table 1: Results of cross-validated models for different pre-treatment methods evaluated on the test set

PLSR		
<i>Pre-treatment</i>	R^2	$RMSE$
None	0.977	1.087
SG1	0.975	1.196
SNV	0.984	0.772
SVR		
<i>Pre-treatment</i>	R^2	$RMSE$
None	0.919	3.893
SG1	0.980	0.952
SNV	0.968	1.530
ANN		
<i>Pre-treatment</i>	R^2	$RMSE$
None	0.949	2.433
SG1	0.969	1.49
SNV	0.973	1.312

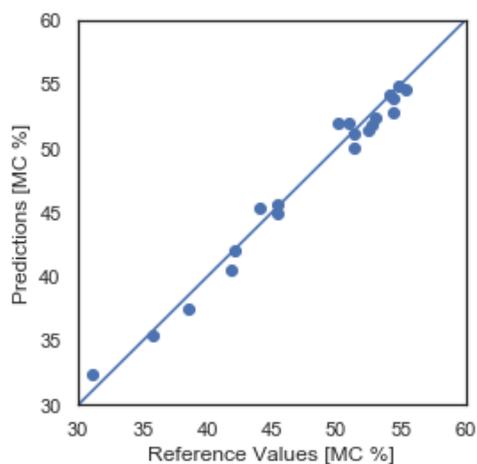
In the case of PLSR the difference between the untreated data and the pre-treated data is very small, which can be attributed to the cleaning of the data and the removal of the noisy parts of the spectra before the pre-treatment. In fact, SG1 pre-treatment performs worse than the untreated data. This is most likely due to the fact that there is little noise in the spectra, rendering the treated and untreated data very similar. It is possible that the window selected for the derivative was slightly larger than the optimal. The SG1 setup was evaluated based on the cross-validation set, and it appears that the results on the unseen test set point out that a different setup would be optimal for SG1. However, this is not possible to know before the models are tested.

In the case of SVR and ANN, the difference between the untreated and pre-treated data is much larger. This is due to the fact that the untreated data was not normalized for these cases, resulting in notably worse models than when the data was pre-treated. For SVR, the best models for each of the different pre-treatments were selected. The models with a linear or polynomial kernel were the best in all cases, whereas the models with the RBF kernel were sometimes slightly overfitted, despite the cross-validation. For ANN, the pre-treatment did not affect model selection as much, and the best models were not as good as those built with SVR. Nonetheless, the difference was not very large, despite the relatively small dataset.

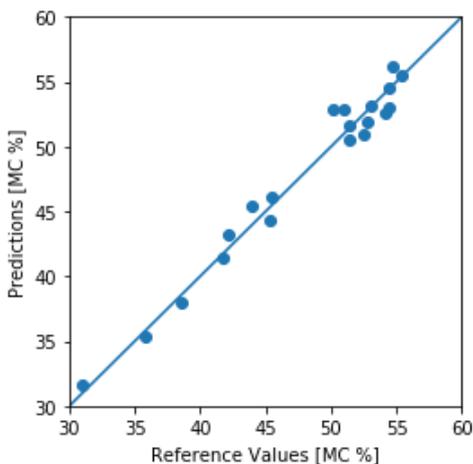
The predicted values of the unseen test set are plotted against the reference values in Figure 5 for the best models, selected based on the R^2 and $RMSE$ values.



(a) PLSR



(b) SVR



(c) ANN

Figure 5: Predicted vs reference moisture content (MC) of the unseen test set

These are the SNV pretreatment for PLSR (Figure 5(a)), SG1 pretreatment with a polynomial kernel for SVR (Figure 5(b)), and SNV pre-treatment for ANN (Figure 5(c)). It can be seen that the predictions of both the SVR and the ANN models are worse than those of the PLSR especially for the higher moisture content, and this could be addressed in future work.

4. Conclusion

This paper compared different methods to develop calibration models for hyperspectral imaging applied to the characterization of biomass samples. The HSI technique can provide good predictions of the moisture content in woody biomass fuel. The two different preprocessing techniques improved the results for SVR and ANN, with standard normal variate performing better than Savitzky-Golay first derivative in the best models developed with two out of the three methods. Overall, PLSR provided the best results, particularly with SNV pre-treatment of the data. Nonetheless, both SVR with SG1 pre-treatment and ANN with SNV pre-treatment were able to deliver accurate and robust models. It is also worth noting that the differences between SNV and SG1 pre-treatment were not very large for many of the models, and the simplicity in the setup of SNV compared to SG1 should be taken into consideration. When taking into account the time requirements to train and tune a model, PLSR and SVR are the best options for this application.

It was noted that in a number of models the results from the evaluation of the models based on their performance in cross-validation did not agree with those of the evaluation based on the test set, with the two pointing at different model setup. Model selection has to be based on the performance on the test set, but since the models will be tuned only with the cross-validation, it is important to be aware that the best model might still be somewhat overfitted with this process. This also points out the usefulness of a more extensive dataset. Nonetheless, the differences are very small and all models are providing acceptable predictions for a real application where this information can be used for the control of a power plant.

Further analysis is required to assess whether different machine learning methods, such as GPR, can be employed to develop a model with an even better predictive capability. GPR is based on a Bayesian approach and as such can also provide the uncertainty range of the predictions. Taking into account the uncertainty of the measurements can provide more information on the suitability of the different techniques for the development of more realistic, probabilistic models. An expansion of the dataset, taking advantage of the range of measurements obtained with the hyperspectral imaging camera, can help increase the accuracy of the models and understand whether the difference in performance is due to the size of the dataset.

Acknowledgment

The authors would like to thank Mikael Karlsson, Konstantinos Kyprianidis, and Robert Tryzell for their technical input and Lotta Lejdborg and Per Örvind for their support.

This work is part of the IFAISTOS and DYNOP projects. IFAISTOS has received funding in the framework of the joint programming initiative ERA-Net Smart Energy Systems' focus initiatives Smart Grids Plus and Integrated, Regional Energy Systems, with support from the European Union's Horizon 2020 research and innovation programme under grant agreements No 646039 and 775970. DYNOP was funded by the Swedish Knowledge Foundation (KKS).

Nomenclature

ANN	Artificial Neural Network
GPR	Gaussian Process Regression
HSI	HyperSpectral Imaging
MC	Moisture Content
ML	Machine Learning
NIR	Near InfraRed
PLSR	Partial Least Squares Regression
RBF	Radial Basis Function
SNV	Standard Normal Variate
SVM	Support Vector Machines
SVR	Support Vector Regression
SG1	Savitzky-Golay first derivative

References

- Ahmed, M., Andersson, P., Andersson, T., Tomas Aparicio, E., Baaz, H., Barua, S., Bergström, A., Bengtsson, D., Skvaril, J. and Zambrano, J. (2018), Real-time biomass characterization in energy conversion processes using near infrared spectroscopy - a machine learning approach.
- Gewali, U. B., Monteiro, S. T. and Saber, E. (n.d.), 'Machine learning based hyperspectral image analysis: A survey'. doi: 10.48550/ARXIV.1802.08701.
- Gillespie, G. D., Farrelly, D. J., Everard, C. D. and McDonnell, K. P. (2016), 'The use of near infrared hyperspectral imaging for the prediction of processing parameters associated with the pelleting of biomass feedstocks', *Fuel Processing Technology* **152**, 343–349. doi: 10.1016/j.fuproc.2016.06.026.
- Mancini, M., Mircoli, A., Potena, D., Diamantini, C., Duca, D. and Toscano, G. (2020), 'Prediction of pellet quality through machine learning techniques and near-infrared spectroscopy', *Computers Industrial Engineering* **147**, 106566. doi: 10.1016/j.cie.2020.106566.
- Mäkelä, M. and Geladi, P. (2017), 'Hyperspectral imaging to determine the properties and homogeneity of renewable carbon materials', *ChemSusChem* **10**. doi: 10.1002/cssc.201700777.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011), 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research* **12**, 2825–2830.
- Pitak, L., Laloon, K., Wongpichet, S., Sirisomboon, P. and Posom, J. (2021), 'Machine learning-based prediction of selected parameters of commercial biomass pellets using line scan near infrared-hyperspectral image', *Processes* **9**(2). doi: 10.3390/pr9020316.
- Rahman, M., Fentaye, A. D., Zaccaria, V., Aslanidou, I., Dahlquist, E. and Kyprianidis, K. (2021), A framework for learning system for complex industrial processes, in K. Kyprianidis and E. Dahlquist, eds, 'AI and Learning Systems', IntechOpen, Rijeka, chapter 2. doi: 10.5772/intechopen.92899.
- Skvaril, J., Kyprianidis, K. G. and Dahlquist, E. (2017), 'Applications of near-infrared spectroscopy (nirs) in biomass energy conversion processes: A review', *Applied Spectroscopy Reviews* **52**(8), 675–728. doi: 10.1080/05704928.2017.1289471.
- Tao, J., Liang, R., Li, J., Yan, B., Chen, G., Cheng, Z., Li, W., Lin, F. and Hou, L. (2020), 'Fast characterization of biomass and waste by infrared spectra and machine learning models', *Journal of Hazardous Materials* **387**, 121723. doi: 10.1016/j.jhazmat.2019.121723.
- Tiitta, M., Tiitta, V., Heikkinen, J., Lappalainen, R. and Tomppo, L. (2020), 'Classification of wood chips using electrical impedance spectroscopy and machine learning', *Sensors* **20**. doi: 10.3390/s20041076.
- Tsuchikawa, S., Inoue, K., Noma, J. and Hayashi, K. (2003), 'Application of near-infrared spectroscopy to wood discrimination', *Journal of Wood Science* **49**, 0029–0035. doi: 10.1007/s100860300005.

Parameter and State Estimation for an Oil Production Model using Julia

Zhe Ban, Carlos Pfeiffer, Bernt Lie

*Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway, 3918 Porsgrunn, Norway
{zhe.ban, carlos.pfeiffer, Bernt.Lie}@usn.no*

Abstract

Dynamic models of industrial processes play an instrumental role in the operation of such processes from smart sensors, data reconciliation, to advanced control. For good performance, a precise model is normally required. The issue of improving models has received considerable critical attention. In this work, we consider the estimation of model parameters and initial states of a gas lifting oil well model, followed by filtering of its states. By utilizing information from both first-principle model and data, the results are presented to show the estimated values and their uncertainties. Julia is the main programming language used in this study. This research study provided an opportunity to advance the understanding of the optimization and estimation for the oil well operation.

1. Introduction

As a common statistic method, Bayesian inference has been used in many scientific fields and industries to provide estimates for unknown quantities considering uncertainty. The application of the Bayesian approach can be classified as batch and sequential according to the property of the data set. In the case of the estimation related to batch datasets, all data are collected at once before processing. Markov chain Monte Carlo (MCMC) algorithms, as stochastic simulation methods, are commonly used for the batch case to solve model uncertainty problems. In the sequential application, the data arrive sequentially, so the estimation or the data evolution is required in real-time. The relevant algorithms for Bayesian sequential updating of probability distributions include the Kalman Filter (KF) [1], the Particle Filter (PF) [2], the Ensemble Kalman Filter (EnKF) [3].

With the information of prior and likelihood, the posterior distribution can be calculated, which helps gain more insight into the estimated quantities. In some cases, the posteriors are simple and can be presented or approximated as tractable and common distributions. However, the posterior distribution may be complex for sophisticated models, for example, the distribution is multimodal or high dimensional. The MCMC algorithm provides a solution where the distribution of the accepted samples converges to the true posterior distribution in the long run [4]. Compared with many deterministic approximation methods, the MCMC algorithm is conceptually easy to adopt for complex systems.

Although the MCMC algorithm has less requirement for the system, the algorithm is computationally slow. Therefore, it was not commonly applied to oil well systems in early research. Benefiting from the development of computer technology, a parallel computing scheme can be used for the MCMC algorithm.

For the dynamically evolving datasets, KF and PF have been used in the past to solve data reconciliation and data assimilation problems. Derived KF approaches, algorithms such as the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and the EnKF, solve the problem where the model is nonlinear. Models and algorithms are formed in these methods to recursively update and estimate quantities.

By linearizing the model using differentiation, the covariance matrices are propagated assuming Gaussian distribution in EKF. However, the analytical computation is not feasible for non-Gaussian models or models which can not be succinctly linearized. In these cases, numerical strategies can be applied to estimate the system states.

Another approach to solving nonlinear problems is PF. As a nonparametric approach, PF estimates the belief by sampling the model output and assigning weights to these samples. Resampling schemes are used to estimate the posterior distribution in each iteration. The high weighted particles are used to design copies and low weighted particles are rejected [5, 6].

In the EnKF method, stochastic models are used to calculate the probability distribution and time evolution of the states. During the initialization, particles are generated around the initial states and these particles are designed according to the prior probability distribution. The term ‘ensemble members’ will be used in this paper to refer to these particles in the EnKF algorithm. All ensemble members are propagated and updated during the iterations. The uncertainty is predicted by calculating the error covariances. Because only a few ensemble members play a vital role eventually in the integration and the weights of most ensemble members become very small, the accuracy of estimation is related to the ensemble size [7]. Typically, the number of ensemble members should be in the range of 50-200 for computing mean and covariance, and considerably larger for computing higher order statistics. However, the large ensemble size also

causes high computational cost. Compared with EKF, EnKF is easy to be implemented as it does not demand differentiation. For the sequential filters, EnKF requires substantially less computational cost than PF.

In the oil well area, MCMC and EnKF have been used in some research work. One study by Maraggi et al. [8] chose a Bayesian approach to estimate two parameters of oil reservoirs, which was simplified as a single dimensionless equation. Following an adaptive MCMC algorithm, posterior predictive checks were adopted to examine the inferences. Uncertainty of the estimated ultimate recovery was addressed in the study. The convergence of the chains, acceptance ratio, posterior distributions, correlation between posterior parameters, the reliability plot and posterior predictive checks were presented to evaluate the approach.

Kang et al. [9] applied a similar method to an electric submersible pump system. With given prior, Bayesian inference and the MCMC methods were adopted for parameter estimation. After estimating five parameters, validation and cross-validation were deployed using two sets of data to examine the model. Dynamic and steady-state uncertainty of the model were obtained via probability density functions using uncertainty assessment [10]. Autocorrelation was used to evaluate the samples, and sensitivity analysis was employed for capturing the region of convergence of the likelihood function. However, the research study did not take into account the samples process in much detail. The convergence of multiple chains was not clearly shown to provide reliable parameter estimation.

In terms of the application of the EnKF to the oil well field, a recent study [11] proposed to use principal component analysis for selecting the initial models for EnKF, so that fewer ensemble members were needed to predict production performances of the channel reservoirs. Meanwhile, the accuracy of the prediction is better than the prediction using one model when the ensemble size is the same. With a smaller ensemble size, computational time decreased significantly. Compared with the original EnKF, this work with model selection scheme provided a solution for filtering data of a complex model with less time. However, the selection of representative initial models demands empirical information of the system, for example, permeability distribution in this work. The selection is the part of the work which belongs to prior design.

With coupled machine learning and EnKF, a data-driven method was proposed to estimate the properties of the reservoir using pressure transient data [12]. Prior to commencing parameter estimation, the random forest method was adopted to classify the model using the discrete linear segment slopes in transition parts. After deciding which model should be used, the grid search method was conducted to estimate three hyper-parameters. With partitioned subspace, a decision tree was designed to address the optimal partitions. The accuracy of the optimization was validated by using cross validation. Once the model and parameters were decided, EnKF was applied to predict the pressure transients of the water drive gas reservoir well.

The purpose of this investigation is to estimate parameters and initial states in the process of a gas lifting oil well model using a first-principle model and batch data, and then sequentially estimate the states online. The dominant noise is Gaussian white noise on the measurements. MCMC was employed to approximate the posterior distributions of parameters and initial states in the first

stage. Once the estimates and uncertainty were extracted, EnKF was carried out to estimate states at each sampling time. This study provides a solution to advance the understanding of the uncertainty in a complex system in real-time.

The overall structure of the study is as follows. Section 2 presents the methodology which is adopted in this work. The third section introduces the gas lifting oil well model which is used in this work and explains the procedure of the simulation. Section 4 validates the performance of the method, Section 5 provides a discussion of the results, while conclusions are given in Section 6.

2. Parameter Estimation and States Filtering

In this work, the estimation include two parts: parameters and initial state estimation, and state estimation. We used the Metropolis-Hastings (MH) Algorithm to estimate the distribution of parameters and initial states based on Bayesian analysis. Once we have found the distribution of the parameters and initial states, we draw samples from part of the accepted list for the initialization of the EnKF.

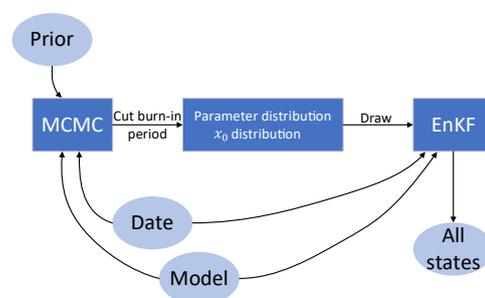


Figure 1: Outline of the estimation process.

2.1. Parameter and Initial State Estimation

Among the MCMC algorithms, the MH algorithm is one of the algorithms which has least requirement of the posterior and is simple to apply [13]. We use MH here to estimate the parameters and initial states by accepting or rejecting proposed samples.

A strong relationship between the accuracy of the MCMC estimation and the chain length has been reported in the literature [14]. Similar to most MCMC algorithms, the step length of MH is a pivotal parameter in the algorithm and impacts the efficiency of the algorithm. A series of experiments were run and step length for each quantity was selected after comparing the results, so that proposed samples can reach large parameter space and can also converge to an optimised value.

For each MH iteration, differential equations in the oil well model need to be solved. Therefore, it takes a long time to run more than 20 chains with a large number of iterations, for example, 20 chain with 1000 iterations for the outputs which contain 3600 samples takes more than six hours with the processor Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz. To increase the computational efficiency, the number of chains and the number of samples for each chain were chosen carefully by trial and error.

After running the MH algorithm, we have a list of accepted samples. These samples form a chain. We can check how the chain explores the parameter space by plotting these samples. The distribution of these samples can be used to approximate the distribution of the posterior of the parameters.

Algorithm 1: Ensemble Kalman Filter algorithm**Initialization:**

Draw samples for initial states, $x_{0|0}^i$, from Q , $i \in \{1, \dots, n\}$

Calculate the mean of the samples $\hat{x}_{0|0} = \frac{1}{n} \sum_{i=1}^n x_{0|0}^i$

Give uncertainty

$$X_{0|0} = \frac{1}{n-1} \sum_{i=1}^n (x_{0|0}^i - \hat{x}_{0|0})(x_{0|0}^i - \hat{x}_{0|0})^T$$

for $k = 1, 2, \dots, N$ **do**

Propagation:

$$x_{k|k-1}^i = f(x_{k-1|k-1}^i, u_{k-1}, e_{k-1}^i)$$

$$\hat{x}_{k|k-1} = \frac{1}{n} \sum_{i=1}^n x_{k|k-1}^i$$

Update:

$$y_{k|k-1}^i = h(x_{k|k-1}^i, u_{k-1}, v_{k-1}^i)$$

$$\hat{y}_{k|k-1} = \frac{1}{n} \sum_{i=1}^n y_{k|k-1}^i$$

$$Z_{k|k-1} =$$

$$\frac{1}{n-1} \sum_{i=1}^n (x_{k|k-1}^i - \hat{x}_{k|k-1})(y_{k|k-1}^i - \hat{y}_{k|k-1})^T$$

$$\varepsilon_{k|k-1} = \frac{1}{n-1} \sum_{i=1}^n (y_{k|k-1}^i - \hat{y}_{k|k-1})(y_{k|k-1}^i - \hat{y}_{k|k-1})^T$$

$$K_k = Z_{k|k-1} \varepsilon_{k|k-1}^{-1}$$

$$x_{k|k}^i = x_{k|k-1}^i + K_k (y_k - y_{k|k-1}^i)$$

$$\hat{x}_{k|k} = \frac{1}{n} \sum_{i=1}^n x_{k|k}^i$$

$$X_{k|k} = \frac{1}{n-1} \sum_{i=1}^n (x_{k|k}^i - \hat{x}_{k|k})(x_{k|k}^i - \hat{x}_{k|k})^T$$

end

Finally, a subset Q is prepared to include the estimation information for the initialization of the estimation of the state. According to the posterior distribution, Q is designed to cover the most likely estimation by ignoring the burn-in, which is the beginning part of the chain before the chain converges to a certain value.

2.2. State Estimation

The aim of this section is to estimate the model states, given a first-principle model and a set of measurements with unknown uncertainties. The data assimilation method, EnKF, was applied as in Algorithm 1 [7].

The distribution of states at each sampling time were approximated via a number of ensemble members x_k^i . The superscript shows the number of the ensemble member and the subscript presents the sampling time. Initial ensemble members are drawn from the subset Q to include the prior knowledge. In the EnKF algorithm, previous studies evaluating the algorithm results observed that the ensemble size is important [7]. The ensemble size that we used in the algorithm is n . The covariance of the states at sampling time k is X_k . The experiment contains N time steps.

In the propagation stage, the states are forecasted based on the last states and the dynamic system model. During the update stage, the estimated outputs are expressed based on the dynamic system model. e and v are the process noise and measurement noise with zero means. y_k is the measurement vector which contains noise. The means and covariance of the states are updated at the end of each time step. The states of the oil well model can be filtered in real time.

3. Simulation Study

3.1. Gas Lifting Oil Well Model

The gas lifting oil well model used in this work is based on previous work [15, 16]. The input of the model is the

valve opening of the gas lift choke valve, u . As one of the parameters to be estimated, the water cut, WC, is the volume of water produced compared to the volume of total liquids produced from an oil well. Other parameters in the model to be estimated are gas-to-oil ratio, GOR, and the productivity index, PI [$\frac{\text{kg/hr}}{\text{bar}}$]. PI is a mathematical means of expressing the ability of a reservoir to deliver fluids to the wellbore. All these parameters are dimensionless.

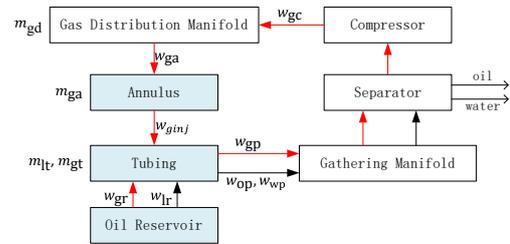


Figure 2: Flow chart of an oil field: the blue blocks are the components of an oil well. The red arrows show the gas flow and black arrows present the liquid phase flows, which can contain oil and water. The mass and flow rate are depicted beside the corresponding components where these states occur in the gas lifted oil well.

The flow chart in Fig. 2 shows the change of flows and masses in the gas lifting oil well. m denotes accumulated mass, and w denotes mass flow rate. A dot on top of a mass implies the time derivative of mass. A subscript indicates the phase (g: gas, o: oil) and location of the various variables. The output vector is $y = \{w_{ga}, w_{gp}, w_{op}, w_{wp}, P_{wf}, P_{wh}, P_a\}$. Flow measurements are impacted by bubbles and are not as reliable as other data such as temperature and pressure. Pressure transmitters are used to measure the bottom hole pressure and well flow pressure (P_{wf}), the pressure in the tubing upstream the production choke valve (P_{wh}), and

the pressure in the annulus downstream the lift gas choke valve (P_a). We assume the pressure of the gas in the gas distribution pipeline (P_c) is a constant as 200 [bar]. In the production process, temperature sensors detect the temperature in the gas distribution pipeline, the annulus and the tubing. Because of the small difference between these temperatures, we assume the temperature is constant everywhere and all these temperatures are assumed to be equal to T [K].

Considering the principle of the gas lifted oil well, its model is designed based on the mass balance of three states: the mass of gas in the annulus m_{ga} , the mass in the tubing above injection point m_{gt} , and the mass of liquid in the tubing above injection m_{lt} . After time derivative, the mass balance is mainly presented as Eq. (1):

$$\begin{aligned} \dot{m}_{ga} &= w_{ga} - w_{ginj}, \\ \dot{m}_{gt} &= w_{ginj} + w_{gr} - w_{gp}, \\ \dot{m}_{lt} &= w_{lr} - w_{lp}, \end{aligned} \quad (1)$$

, where w_{ga} is the flow rate of the gas through the gas lift choke valve which is injected into the annulus. The flow rates of the lift gas from the annulus and reservoir to the tubing are w_{ginj} and w_{gr} respectively. The flow rate of produced gas through the production choke valve is presented as w_{gp} . w_{lr} and w_{lp} are the liquid phase flow from the reservoir into the well and through the production choke valve, respectively. The state vector is $x = \{m_{ga}, m_{gt}, m_{lt}\}$.

3.2. Simulation Setup

There has been an increased interest in the programming language Julia [17] in recent years. As a free language, it has shown rapid advances in the field of numerical computing. In this work, we adopted Julia as the main language, especially for the state estimation part. The parameter and initial state estimation part is based on the previous work [16]. The Julia packages *Plots*, *StatsPlots*, *Distributions*, *DifferentialEquations*, *Noise*, *Random* were used in the this work.

The work began by generating data from the oil well model with true parameters and initial states, and then Gaussian noise with zero means was added to the outputs. The oil well model does not contain process noise. The noise in the measurements only include measurement noise. The true parameters are $WC = 0.18$, $PI = 2.4 \times 10^4$, $GOR = 0.15$. The true initial states of m_{ga} , m_{gt} and m_{lt} are $m_1 = 8650$, $m_2 = 3306$, $m_3 = 18250$ respectively. The true values were only revealed after the estimation for evaluation. The information of noise is not accessible during the experiments. The simulation will then go on to estimate the parameters and the initial states, as well as the uncertainty of the model. Once the distribution of these quantities are found, the states will be estimated.

The prior of the parameters and initial states were set as uniform distribution in Eq. (2). The GOR and WC parameters represent ratios, so they are between 0 and 1.

$$P(\theta, m_0 | u_i, y_{e(i,j)}) = \begin{cases} B & PI \in [10^4, 10^5], GOR \in [0, 1], WC \in [0, 1], \\ & m_1 \in [4000, 10000], m_2 \in [1000, 5000], \\ & m_3 \in [1 \times 10^4, 2.5 \times 10^4] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

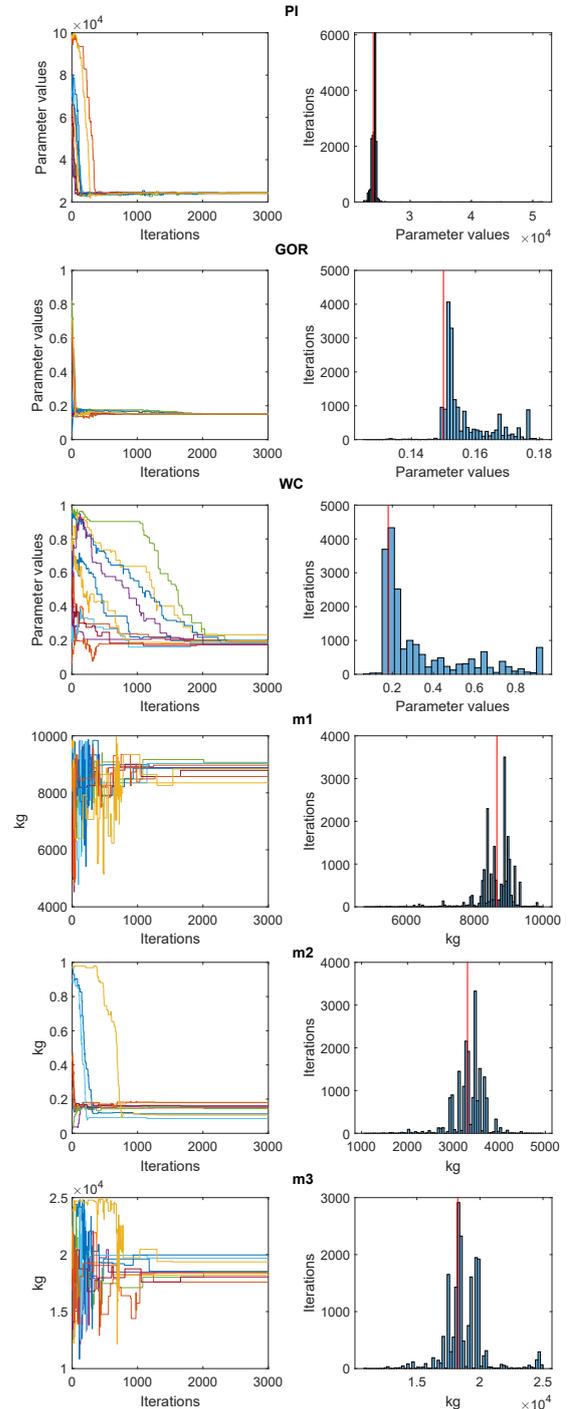


Figure 3: Estimation of the parameters and the initial states. The top three rows of plots show the parameter estimates and the remaining plots present the initial value estimates. The red vertical lines in the right columns of plots show the true values of the parameters and initial states.

In this work, we run 10 chains, and each chain contains 3000 iterations to identify the parameters and initial states using data with 40 samples. We assume the measurements are independent and contain white noises which are normally distributed with zero means, but with different variances. For every individual measurement, the variance is assumed to be constant during the experiments. The process error, e_{k-1}^i , used at the propagation stage in EnKF was tuned by trial and error. The measurement noise, v_{k-1}^i , used at the update stage was calculated from the data.

4. Estimate validation

The main aim of this section is to present and analyse the results of experiments in various ways. To gain insights into every parameter in the estimation process, we present plots of individual quantities distributions and chains. The interactions between every pair parameters are shown in scatter plots and contour plots. The estimated states are compared with the true states.

4.1. The Distribution of Quantities and Chains

In order to test the influence of the random initial value of the MCMC algorithm and check the explored range of the estimated distribution, we present the plots of the chains and the distribution for each parameter and initial value, shown in Fig.3. The plots include the estimations of ten chains of the MCMC algorithm with 3000 iterations.

The left column of plots show trace plots with various initial values, which provides information on the converging speed of the chains. The vertical axis is the estimated value, and the horizontal axis is the iteration number. The burn-in periods of these chains are generally less than 2000. Compared with other chains, WC and m_1 took longer time to converge to a certain range. We cut the burn-in periods and used the last 1000 samples as the subset Q .

According to these plots, all chains converge to around the true values after exploring the whole prior range. The end of the chains overlap each other and fairly smoothly drift around the optimal estimated value, namely the estimate uncertainty decreased as the iterations progressed. All the chains for the same parameters converge to a similar region and mix well, which indicates convergence is achieved.

The right column of the figures illustrates the distribution of each individual estimate. The histograms of the quantities are drawn based on the accepted lists of all chains. According to the distribution, the most likely estimates of these quantities are around the true values. The MCMC algorithm was able to identify the true values of the quantities.

4.2. Pair plots of the Distribution of Quantities

Scatter plots and contour plots are shown in Fig.4 to present the relationship between each pair of the parameters and initial states. According to the top two rows of plots, the main distribution of PI is between 2.35×10^4 to 2.50×10^4 . The first and third rows of plots show that GOR distributes around $[0.14, 0.20]$. The samples of WC spread between 0.2 to 0.8. The initial state estimates are not as converged as the parameter estimates. Most of m_1 samples are between 8000 and 10000. m_2 is estimated within $[3000, 4000]$, and the value of m_3 is locate in $[1.6, 2] \times 10^4$. These plots confirm the estimates in Fig.3. The contour plots in the second column demonstrate more details in the converging area and show that some of the posterior distributions are multimodal.

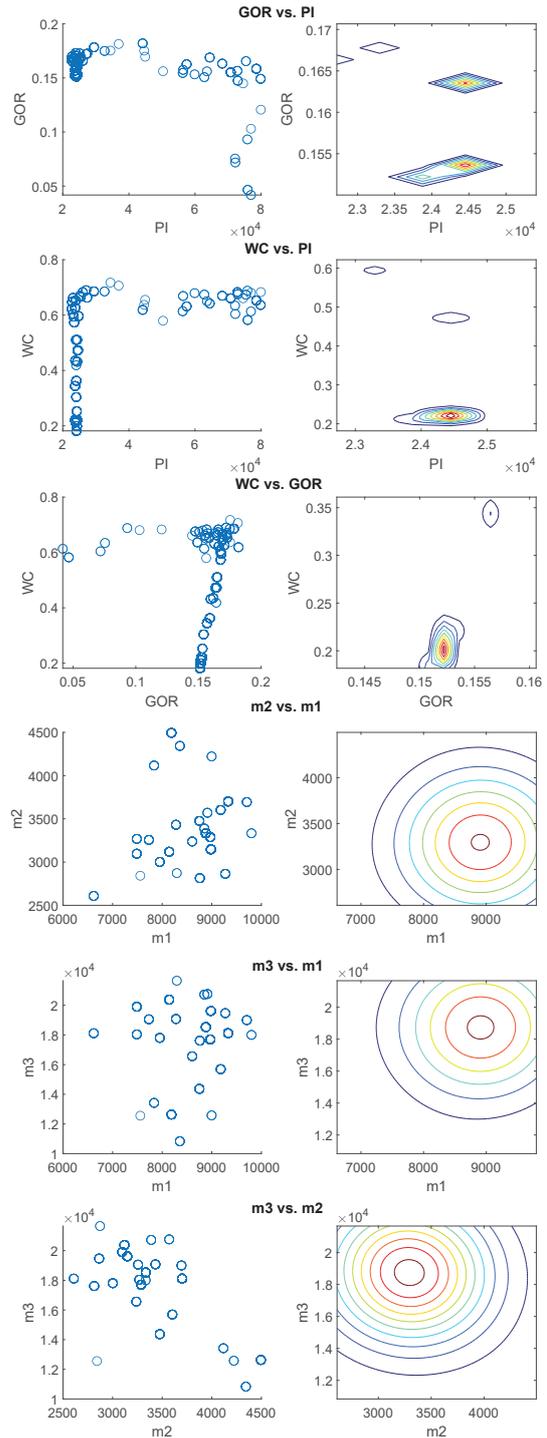


Figure 4: Estimation of parameters and initial states with 2D plots in scatter plots and contour plots. The top three rows of plots show the parameter estimates and the remaining plots present the initial value estimates. The contour plots are drawn in the range with minimum and maximum values. The colors of the contour plots show the density of estimated values. The red curves shows the estimation at high density, while the blue curves correspond to the low density contour. For the estimation of parameters, the converge parts were zoomed in to present the contour clearly.

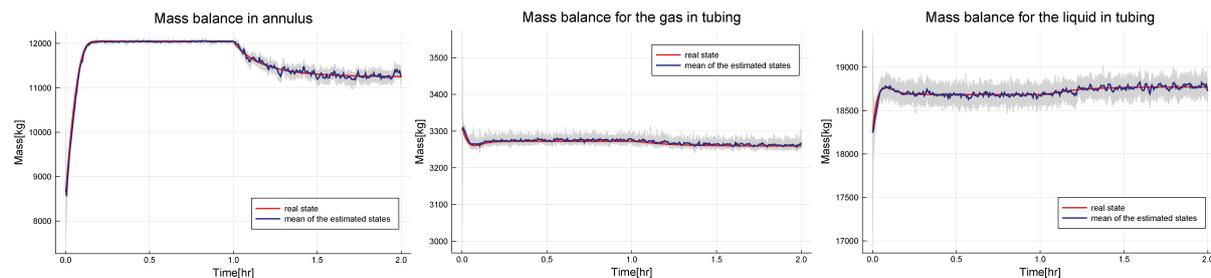


Figure 5: Estimation of states. The grey shadows are the estimated states which include 100 estimates for each sampling time. The dark blue lines are the mean of these ensemble values. The true values of these states are presented as red lines.

4.3. State estimation

We compare the estimated states and the true states in this subsection. Figure 5 shows the estimation result. With 1000 samples from the accepted list Q , the EnKF algorithm is initialised with 100 ensemble members. The time evolution of these ensemble members are shown as grey shadows in the plots. The starting points of these grey shadows indicate the initialisation range of ensemble. For example, the initial ensemble members of mass in the annulus is distributed within the range (7000, 10000) and the initial ensemble members of mass balance for gas and liquid in tubing are in the ranges (2950, 3550) and (16800, 19400), respectively. The true states were collected from model simulation without any noise.

According to the result, the mean of all estimated states are close to the true states. All the ensemble values converge quickly towards the true states, though some of the initial values are distributed with a considerable deviation from the true initial states. Because the estimation of the states is based on the model and the measurements with noise, the estimates are less smooth than the true state.

The covariance matrix shows the uncertainty of the estimation. The uncertainty can also be checked through the width of the grey shadow. The uncertainty of the estimates of the mass for the gas and liquid in the tubing does not change significantly after the input changing at $t = 1$, while the uncertainty of the estimates of the mass for the gas in tubing increases significantly. The increase of uncertainty results from the propagation of the error between the estimated output and the measurement. Compared with the last two states, the first state changes more noticeable after the input changes.

5. Discussion

Due to the lack of real data, it is not possible to exam the method here to a real gas lifting oil well system. In real life, the model mismatch might lead to undesirable estimation. Besides, in real oil well systems, most states and measurements are coupling. We simplified the problem and assumed the noises of these quantities are independent, so that the measurement error in the EnKF algorithm could be easily found by calculating the variance of the measurement during steady state. The assumption might lead to some mismatch in the estimation.

During the data filtering stage, the initial ensemble contains the prior information. However, it is not possible to add constraints to the time evolution of these ensemble. The estimated states are calculated based on the measurement and the last estimated states. For example, the mass should be positive all the time according to physics knowledge, but the EnKF algorithm can not guarantee that the estimates are always positive. The shortage of the EnKF algorithm might lead to some

improper estimation.

Another potential problem is that this research study does not engage with the circumstance where the parameters change over time. In reality, models are always imperfect representations of a system, and it may be necessary to allow for parameters to change over time to achieve the best possible model fit.

6. Conclusions

The purpose of this study was to investigate the feasibility of estimating parameters and states of a gas lifting oil field. In the first stage, the posterior distributions of the parameters and the initial states were identified using a MCMC algorithm. Then, we drew samples from a subset of the posterior distribution for the initialization of the EnKF and estimated the states. The analysis of experimental results undertaken here has confirmed that the proposed solution was able to find the true values of the unknown quantities with uncertainty during estimation. The study contributes to our understanding of the uncertainty in the estimation of the quantities of a system in real-time. This understanding helps to improve the gas lifting oil well model and provide estimations of the process states in the system, why might benefit the operation and advanced control in real life. Further research could be conducted to improve the algorithm for the estimation of changing parameters.

Acknowledgment

This work was supported by The Research Council of Norway and Equinor ASA through Research Council project “308817 - Digital wells for optimal production and drainage” (DigiWell).

References

- [1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” 1960.
- [2] A. Doucet, S. Godsill, and C. Andrieu, “On sequential monte carlo sampling methods for bayesian filtering,” *Statistics and computing*, vol. 10, no. 3, pp. 197–208, 2000.
- [3] G. Evensen *et al.*, *Data assimilation: the ensemble Kalman filter*, vol. 2. Springer, 2009.
- [4] S. Brooks *et al.*, “Handbook of markov chain monte carlo. boca raton: Crc,” 2011.
- [5] W.-s. Chen, B. R. Bakshi, P. K. Goel, and S. Ungarala, “Bayesian estimation via sequential monte carlo sampling: unconstrained nonlinear dynamic systems,” *Industrial & engineering chemistry research*, vol. 43, no. 14, pp. 4012–4025, 2004.
- [6] A. Doucet, N. De Freitas, N. J. Gordon, *et al.*, *Sequential Monte Carlo methods in practice*, vol. 1. Springer, 2001.
- [7] G. Evensen and P. J. Van Leeuwen, “Assimilation of geosat altimeter data for the agulhas current using the ensemble kalman filter with a quasigeostrophic model,” *Monthly Weather Review*, vol. 124, no. 1, pp. 85–96, 1996.

- [8] L. M. R. Maraggi, L. W. Lake, and M. P. Walsh, "A bayesian framework for addressing the uncertainty in production forecasts of tight oil reservoirs using a physics-based two-phase flow model," in *SPE/AAPG/SEG Latin America Unconventional Resources Technology Conference*, OnePetro, 2020.
- [9] E. Costa, O. de Abreu, T. d. O. Silva, M. Ribeiro, and L. Schnitman, "A bayesian approach to the dynamic modeling of esp-lifted oil well systems: An experimental validation on an esp prototype," *Journal of Petroleum Science and Engineering*, vol. 205, p. 108880, 2021.
- [10] JCGM, "Evaluation of measurement data—guide to the expression of uncertainty in measurement," *JCGM*, vol. 100, no. 2008, pp. 1–116, 2008.
- [11] B. Kang, H. Jung, H. Jeong, and J. Choe, "Characterization of three-dimensional channel reservoirs using ensemble kalman filter assisted by principal component analysis," *Petroleum Science*, vol. 17, no. 1, pp. 182–195, 2020.
- [12] L. Xue, S. Gu, L. Mi, L. Zhao, Y. Liu, and Q. Liao, "An automated data-driven pressure transient analysis of water-drive gas reservoir through the coupled machine learning and ensemble kalman filter method," *Journal of Petroleum Science and Engineering*, vol. 208, p. 109492, 2022.
- [13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [14] J. Kruschke, "Doing bayesian data analysis: A tutorial with r, jags, and stan," 2014.
- [15] R. Sharma, K. Fjalestad, and B. Glemmestad, "Modeling and control of gas lifted oil field with five oil wells," in *52nd International Conference of Scandinavian Simulation Society, SIMS*, pp. 29–30, 2011.
- [16] Z. Ban, A. Ghaderi, N. Janatian, and C. F. Pfeiffer, "Parameter Estimation for a Gas Lifting Oil Well Model Using Bayes' Rule and the Metropolis–Hastings Algorithm," *Modeling, Identification and Control*, vol. 43, no. 2, pp. 39–53, 2022.
- [17] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.

A Comparative Model-Analysis on Sulphide Bio-oxidation with Different Electron Acceptors

Vibeke B. Karlsen, Gamunu Samarakoon, Carlos Dinamarca

Department of Process, Energy, and Environmental Technology, University of South-Eastern Norway

Carlos.Dinamarca@usn.no

Abstract

Sulphide (H_2S , HS^- and S^{2-}) is an undesired by-product of biogas production processes. This modelling work in Aquasim was carried out to study three parallel processes related to sulphide in AD environments: 1) H_2S liquid-gas mass transfer; 2) Acid-base equilibrium; and 3) Sulphide oxidation with three different electron acceptors; nitrate, oxygen, and a biotic anode with a given potential. Multiplicative Monod (biotic processes) and Nernst-Monod kinetics (bioelectrochemical process) provide the basis for the sulphide bio-oxidation processes. At the current stage, the model can be used to study sulphide bio-oxidation and the effect of relevant parameters, including initial biomass concentration, uptake rates, temperature, and pH. The model can be improved further by implementing anaerobic microbial processes as competing reactions. With the proposed improvements, the model can be a useful tool for calculating the chemical dosage or electrode potential required for sulphide removal. These calculations can be based on both the concentration of $\text{H}_2\text{S}(\text{g})$ in the headspace (ppm) often available at full-scale plants and the concentration of sulphide ($\text{HS}^-(\text{liq})$) in effluent streams from the plants.

Keywords: Sulphide oxidation, liquid-gas mass transfer, nitrate, oxygen, bioelectrochemistry

1. Introduction

Biogas produced through anaerobic digestion (AD) contains 30-50 % CO_2 , 50-70 % CH_4 and trace gases, including H_2S . Depending on the substrate composition, normal $\text{H}_2\text{S}(\text{g})$ - concentrations are 0-10 000 ppm (Angelidaki *et al.*, 2018). Sulphide is toxic, odorous, and corrosive, even at low concentrations. Different techniques can remove sulphide from liquid and gaseous streams, including physicochemical and biological methods. Biological desulphurisation processes exploit the microbial ability to oxidise sulphide with oxygen or nitrate as electron acceptors. Bioelectrochemical systems (BESs) with a bioanode working as the electron acceptor have been examined as an alternative (Sun *et al.*, 2009). In these systems, microorganisms work as catalysts at the electrodes in oxidation or reduction reactions. A solid anodic surface can function as an electron acceptor for the electrons generated through sulphide oxidation. The electrons are transported to the cathode, where reduction reactions such as CH_4 production from CO_2 occur. This bioelectrochemical technique can provide a chemical-free and environmentally friendly solution for sulphide removal if a renewable energy source is used to supply electricity.

The processes related to sulphide in AD environments are complex and usually involve the following main processes: 1) Sulphate bioreduction to sulphide; 2) Liquid-gas mass transfer of H_2S ; 3) Acid-base equilibrium of $\text{H}_2\text{S}/\text{HS}^-/\text{S}^{2-}$; and 4) Chemical and biological sulphide oxidation to remove sulphide. Barrera *et al.* (2015) implemented the first three processes as an extension of ADM1. In our modelling work, the main focus was sulphide removal through bio-oxidation (process 4). The

primary goal of the model was to develop a simple simulation tool for studying sulphide bio-oxidation with different electron acceptors and estimating $\text{H}_2\text{S}(\text{g})$ based on liquid-gas mass transfer. The secondary target was estimating the required chemical dosage to reduce the sulphide concentration to a specified target concentration. The bio-oxidation processes were implemented with three different electron acceptors; nitrate, oxygen (biotic processes) and a bioanode (bioelectrochemical process). At this modelling stage, the aim was to study these oxidation processes and identify the most critical parameters for further model development.

2. Methodology

The model was developed using Aquasim as a simulation tool. The reactor space was defined as a modified mixed compartment. The following assumptions were made: 1) Sulphide was included as a loading rate. The model's primary focus is sulphide bio-oxidation processes. Therefore, sulphate bioreduction to sulphide was not included; 2) The microorganisms are assumed to be chemolithoautotrophs, capable of gaining energy by oxidising inorganic sulphur-containing compounds. One microbial group is included for each of the sulphide oxidising processes instead of specific microbial species; X_{eet} for electroactive microorganisms, X_{sob} for microorganisms with oxygen as the electron acceptor, and X_{snb} for microorganisms with nitrate as the electron acceptor; 3) For all the microbial groups, first-order kinetics for decay were assumed valid, as proposed by Batstone *et al.*, (2002); 4) Due to the slower kinetics compared to biotic oxidation, chemical

oxidation is not included; 5) Biotic oxidation of H_2S is slower compared to oxidation of HS^- (Hvitved-Jacobsen *et al.*, 2013), thereby requiring separate rate kinetics. For simplicity, only oxidation of HS^- was included in the model; 6) It is assumed that the microbial culture is adapted to the substrate. The growth kinetics only depends on the concentrations of the electron donor and acceptor (substrate dependence); 7) Only sulphur and sulphate are included as oxidation products. Other sulphur intermediate products are assumed to have faster kinetics; 8) A Nernstian expression can be used to model the transfer of electrons from the microorganisms to the conductive biofilm on the anode surface. This is applicable because the transfer is assumed to be reversible and rapid, as adapted from a modelling study by Marcus *et al.* (2007). 9) It is assumed that there is no proton accumulation in the reactor.

2.1. The reactor vessel and operational parameters

A virtual full-scale biogas reactor with a total volume of 2000 m³ was defined in the model (Figure 1). A representative initial biogas composition of 60-70 % CH_4 , 30-40 % CO_2 , and 100-10000 ppm H_2S was considered. The pH range for AD reactors is often 6.8 to 8.0. In the model, this range is important for estimating $H_2S(g)$ in biogas. To estimate the H_2S -generation with psychrophilic, mesophilic, and thermophilic microorganisms, a representative temperature range of 15-55 °C was chosen. The pH and temperature ranges were used in the acid-base equilibrium- and air-water equilibrium calculations (3.1.2). In the Aquasim simulations, a representative temperature of 308 K and pH of 7.2 were used if not stated otherwise. Other relevant parameters used in the model are listed in Table 1.

Table 1: Parameters related to the modelled AD reactor.

Parameter	Value
Total reactor size [m ³]	2000
Constant bulk liquid [m ³]	1600
Headspace [m ³]	400
Inflow [m ³ d ⁻¹]	75
HRT [d]	21

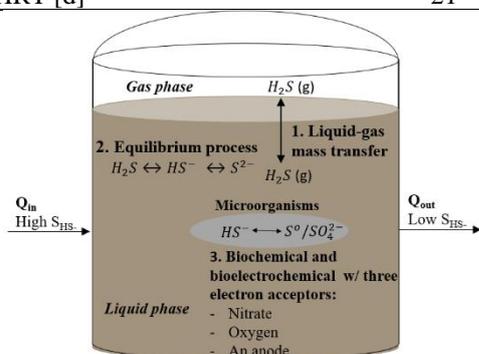


Figure 1: Illustration of the sulphide related processes implemented in the model.

2.2. Liquid-Gas transfer

The H_2S liquid-gas mass transfer was included based on the following expression used by Barrera *et al.* (2015):

$$\rho = k_{La,H_2S} \cdot (S_{H_2S} - K_{H_{H_2S}} \cdot P_{H_2S}) \quad (1)$$

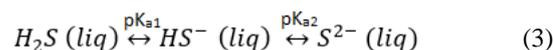
Where k_{La} : gas transfer coefficient, S_{H_2S} : concentration of H_2S , P_{H_2S} : partial pressure of H_2S , and $K_{H_{H_2S}}$: temperature-dependent Henry's law constant calculated with the following equation (Eq. 2):

$$K_{H_{H_2S}} = k_{H^0_{H_2S}} \cdot R \cdot T \cdot \exp\left(\frac{\Delta H^0_{kH_{H_2S}}}{R \cdot 100} \cdot \left(\frac{1}{298} - \frac{1}{T}\right)\right) \quad (2)$$

Where $k_{H^0_{H_2S}}$: Henry's constant at standard conditions and $\Delta H^0_{H_2S}$: enthalpy of reaction for $H_2S(g)$ to $H_2S(liq)$.

2.3. Liquid-liquid transfer /acid-base equilibrium

Depending on the pH in the liquid, sulphide can be present as H_2S , HS^- , or S^{2-} (Eq. 3).



The negative logarithm of the first dissociation constant (pK_{a1}) is close to 7, whereas a pK_{a2} from 14 is reported in the literature (Barrera *et al.*, 2015). Therefore, the concentration of S^{2-} is negligible within the pH range relevant for AD. Only the first dissociation step (H_2S/HS^-) was included in the model. The pK_{a1} -value is calculated as a function of temperature, based on Eq. 4 implemented by Broderius *et al.* (1977).

$$pK_a = 3.122 + 1132/T \quad (4)$$

Eq. 5 represents the expression incorporated for sulphide acid-base equilibrium in the model, based on the acid-base equilibrium included in ADM1 by Batstone *et al.* (2002).

$$\rho_{H_2S/HS^-} = K_{a,H_2S} \cdot S_{H_2S} - (S_{HS^-} \cdot (S_{H^+} + K_{a,H_2S})) \quad (5)$$

Where ρ_{H_2S/HS^-} : kinetic rate equation for the acid-base equilibrium, S_{H_2S} : concentration of H_2S , S_{HS^-} : concentration of HS^- , S_{H^+} : concentration of protons, and K_{a,H_2S} : acidity constant of H_2S with temperature correction.

2.4. Sulphide oxidation

The model was developed to predict biological and bioelectrochemical sulphide removal. The microorganisms oxidise sulphide with three different electron acceptors; 1) Nitrate; 2) Oxygen; and 3) A bioanode with a given potential (Fig. 2).

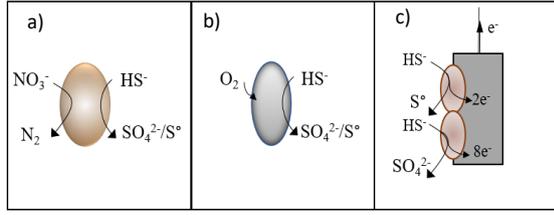


Figure 2: The biotic sulphide oxidation processes implemented in the model, with a) nitrate, b) oxygen, and c) an anode as the electron acceptors.

In this model, the basis of the growth kinetics is the multiplicative Monod established by Bae and Rittman (1996). The kinetics display dependence on both the electron donor and acceptor concentrations. The resulting formula is used to express the uptake rate in this model (Eq. 6):

$$\rho_i = k_m^o \cdot X_i \cdot \frac{S_d}{K_{s,d} + S_d} \cdot \frac{S_a}{K_{s,a} + S_a} \quad (6)$$

Where k_m^o : maximum uptake rate, X_i : concentration of active microorganisms, S_a and S_d are the concentration of the electron acceptor and donor, and $K_{s,a}$ and $K_{s,d}$: corresponding half-saturation constants. Sulphide, S_{HS^-} , is the electron donor (S_d), whereas nitrate, oxygen and the bioanode are the electron acceptors (S_a). By modifying Eq. 6, the following expressions for nitrate (Eq. 7) and oxygen (Eq. 8) were implemented in the model.

$$\rho_1 = k_m^o \cdot X_{snb} \cdot \frac{S_{HS^-}}{K_{s,HS^-} + S_{HS^-}} \cdot \frac{S_{NO_3^-}}{K_{s,NO_3^-} + S_{NO_3^-}} \quad (7)$$

$$\rho_2 = k_m^o \cdot X_{sob} \cdot \frac{S_{HS^-}}{K_{s,HS^-} + S_{HS^-}} \cdot \frac{S_{O_2}}{K_{s,O_2} + S_{O_2}} \quad (8)$$

To describe the kinetics of a biotic anode, a modification of Eq. 6 is required because an electrical potential controls the bioelectrochemical reaction rate. This is accomplished by implementing a Nernst-type equation. Marcus *et al.* (2007) developed a dual-limitation Nernst-Monod kinetic expression. This expression was modified to apply for bioelectrochemical sulphide oxidation (Eq. 9).

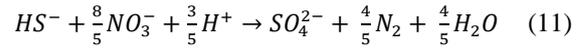
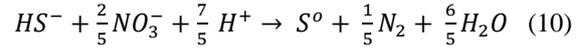
$$\rho_3 = k_{m_eet}^o \cdot X_{eet} \cdot \frac{S_{HS^-}}{K_{s,HS^-} + S_{HS^-}} \cdot \frac{1}{1 + \exp\left(\frac{nF}{RT} \eta\right)} \quad (9)$$

Where F : Faradays constant, n : number of electrons transferred to the anode, η : the local potential ($E_{\text{anode}} - E_{KA}$), E_{anode} : anode potential. E_{KA} corresponds to $K_{s,a}$ and can be determined experimentally (Markus *et al.*, 2007, Samarakoon *et al.*, 2019). In this model, E_{KA} is the reference point and was set to 0. Therefore, η is considered as the anode potential.

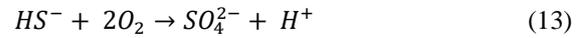
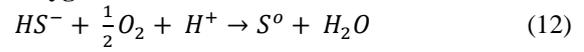
2.5 Stoichiometry

The stoichiometry of the relevant reactions depends on several factors, including the type of microorganisms, the ratio of electron donor to electron acceptor, system design and operational parameters. The reactions are incorporated as 100 % conversion to sulphur or 100 % sulphate. The following six stoichiometric equations were included in the model:

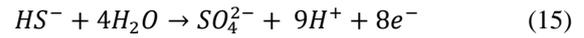
Nitrate:



Oxygen:



Anode:



2.6 Parameters

Sulphide (S_{HS^-}), nitrate ($S_{NO_3^-}$), and oxygen (S_{O_2}) were included as both initial and influent concentrations. The operational parameters are presented in Table 1. The maximum sulphide uptake rate, k_m^o , is equal to μ_{max}/Y_i , where μ_{max} is the maximum microbial growth rate (Batstone *et al.* 2002). This rate depends on the microbial group, and the values vary in the literature. Other relevant parameters were obtained from the literature (Tab. 2).

2.7 Case study

A case study was performed to compare bio-oxidation of sulphide with the three different electron acceptors, and to estimate the dosage and time required to reduce the concentration to a predefined target concentration. The following base conditions were established: 1) Both the initial and inflow sulphide concentrations, S_{HS^-} , are 1.5 mM; 2) Nitrate and oxygen initial concentrations were based on the stoichiometric equations 10 and 12 with 100 % oxidation to sulphur, corresponding to 0.6 mM nitrate and 0.75 mM oxygen, and varied according to the scenarios described in 3.3; 3) Sulphide target concentration was set at 0.1 mM, which according to Figure 4a ensures a P_{H_2S} below 500 ppm with pH 7.2 and 35 °C temperature. The reactor pH was constant at 7.2. An initial active biomass of 0.1 mM was established in all three cases and the k_m^o was assumed to be 7.5 mole S mole⁻¹ X d⁻¹.

Table 2: Parameters and constants related to the processes in Figure 1 implemented in Aquasim.

Parameter	Description	Value	Unit	Reference
$K_{s,sulphide,1}$	Half-saturation constant for sulphide oxidation, with nitrate	$4.6 \cdot 10^{-4}$	mole L ⁻¹	Wanga <i>et al.</i> , 2010
$K_{s,nitrate}$	Half-saturation constant for nitrate	$1.0 \cdot 10^{-5}$	mole L ⁻¹	Wanga <i>et al.</i> , 2010
$K_{s,sulphide,2}$	Half-saturation constant for sulphide oxidation, with oxygen	$1.0 \cdot 10^{-6}$	mole L ⁻¹	Pokorna-Krayzelova <i>et al.</i> , 2018
$K_{s,oxygen}$	Half-saturation constant for oxygen	$1.0 \cdot 10^{-4}$	mole L ⁻¹	Pokorna-Krayzelova <i>et al.</i> , 2018
$K_{s,sulphide,3}$	Half-saturation constant for sulphide oxidation with an anode	$7.0 \cdot 10^{-6}$	mole L ⁻¹	Assumed
Y_i	Yield of biomass on uptake of sulphide. Used for all three types of microbes	0.03	mole S_X mole S ⁻¹	Assumed
F	Faraday's constant	96485	C mole e ⁻¹	-
R	Ideal gas constant	8.314	J mole ⁻¹ K ⁻¹	-
T	Temperature	288-328	K	-
η	Local potential	-0.1 to +0.3	V	-
k_{H^0,H_2S}	Henry's constant at standard conditions	$9.86 \cdot 10^{-2}$	M bar ⁻¹	Sander, 1999
k_{H^0,N_2}	Henry's constant at standard conditions	$6.42 \cdot 10^{-4}$	M bar ⁻¹	Sander, 1999
k_{H^0,CH_4}	Henry's constant at standard conditions	$1.38 \cdot 10^{-3}$	M bar ⁻¹	Sander, 1999
k_{H^0,CO_2}	Henry's constant at standard conditions	$3.55 \cdot 10^{-2}$	M bar ⁻¹	Sander, 1999
$\Delta H^0_{Ka,H_2S}$	Enthalpy of reaction $HS^- + H^+ \rightarrow H_2S$	21670	J mole ⁻¹	Batstone <i>et al.</i> , 2002
$\Delta H^0_{Ka,CO_2}$	Enthalpy of reaction $CO_2 \rightarrow HCO_3^-$	7646	J mole ⁻¹	Batstone <i>et al.</i> , 2002
$\Delta H^0_{KH,H_2S}$	Enthalpy of reaction $H_2S(liq) \rightarrow H_2S(g)$	-17459	J mole ⁻¹	Sander, 1999
$\Delta H^0_{KH,CO_2}$	Enthalpy of reaction $CO_2(liq) \rightarrow CO_2(g)$	-19410	J mole ⁻¹	Batstone <i>et al.</i> , 2002
$\Delta H^0_{KH,CH_4}$	Enthalpy of reaction $CH_4(liq) \rightarrow CH_4(g)$	-14240	J mole ⁻¹	Batstone <i>et al.</i> , 2002
$\Delta H^0_{KH,N_2}$	Enthalpy of reaction $N_2(liq) \rightarrow N_2(g)$	-10808	J mole ⁻¹	Sander, 1999
k_{La}	Mass flux coefficient	200	d ⁻¹	Batstone <i>et al.</i> , 2002
k_{dec}	Decay rate of sulphide oxidising microbes	0.048	d ⁻¹	Sun <i>et al.</i> , 2017

3. Results and discussion

3.1. Identification of key parameters

For the model's practical application, it is important to identify the key parameters which have the highest impact on the three most central state variables; 1) The H₂S-concentration in the headspace, S_{H_2S} ; 2) The concentration of sulphide in the liquid, S_{HS^-} ; and 3) The concentration of active biomass, $X_{sob/snb/eet}$, as a function of growth and decay. This was accomplished through both a sensitivity analysis and dissociation- and air-water equilibrium- calculations. The sensitivity analysis was performed with the sensitivity function in Aquasim. The dissociation- and equilibrium calculations are additional tools for developing the model and not a direct part of the model in Aquasim.

3.1.1 Sensitivity analysis

The sensitivity analysis indicates that the

concentration of hydrogen sulphide, S_{H_2S} , in the headspace displays the highest sensitivity towards Henry's law constant, K_{H,H_2S} , followed by the mass transfer coefficient, kL_a . K_{H,H_2S} is dependent on T , whereas kL_a vary depending on different properties, including T , mixing degree, and liquid properties (Yongsiri *et al.*, 2004). With multiplicative Monod kinetics, S_{HS^-} displayed the highest sensitivity towards the maximum uptake rate, k_m^o , and low sensitivity towards $K_{s,nitrate}$, $K_{s,sulphide}$, and T . The k_m^o -value also exhibited the highest impact on S_{HS^-} with Nernst- Monod kinetics, while S_{HS^-} displayed low sensitivity towards the anode potential (further discussed in chapter 3.2). The concentration of biomass, X_i , presented the highest sensitivity towards Y_i and k_m^o -value as expected, in addition to being highly dependent on the substrate concentrations. The pH

was not included in the sensitivity analysis due to the dependence on the concentration of protons, which is a state variable.

The sensitivity analysis revealed that the model is sensitive to k_m^o . This maximum specific uptake rate depends on the microbial species (Barrera *et al.*, 2015). To study the parameter's effect on the sulphide oxidation rate, simulations were performed with k_m^o -values corresponding to a μ_{max} in the range of 0.075 d⁻¹ to 1.5 d⁻¹ (Fig. 3). Too low uptake rates (Fig. 3: $k_m^o = 2.5$ and 5 mole S mole⁻¹ X d⁻¹) cause inefficient sulphide removal, sulphide accumulation, and potentially high volatilisation of H₂S(g) with both multiplicative Monod- and Nernst-Monod kinetics. The effect of the uptake rate is more prominent with a lower concentration of active biomass (results not included).

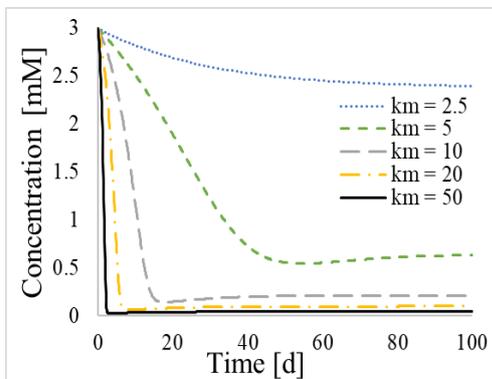


Figure 3: Sulphide oxidation with different uptake rates, here with nitrate as the electron acceptor. Initial biomass: 0.01 mM. Nitrate dosed as both initial and continuous supply (see discussion in 3.3).

The simulation results illustrate that access to a microbial group with an efficient uptake rate for sulphide increases the system's efficiency. In practice, this can be achieved with specialised and adapted microbial cultures or by increasing the concentration of active biomass.

3.1.2. Acid-base equilibrium and air-water equilibrium calculations

The main purpose of the acid-base equilibrium and air-water equilibrium calculations was to study the effect of T and pH, which are two important operational parameters. The dissociation of sulphide (Eq. 3) is highly dependent on pH and temperature, as demonstrated in Figure 4. At pH 6.8 and T = 35 °C, the H₂S/HS⁻ - ratio is 50/50. At a constant pH, a decrease in T causes an increase in the proportion of H₂S. In contrast, increasing the pH to 8 reduces the proportion of H₂S(liq) to less than 10 % for the whole temperature range.

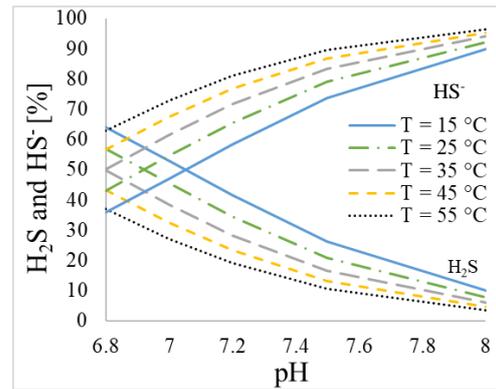


Figure 4: Equilibrium concentration of H₂S and HS⁻ as a function of pH and T

The partial pressure of H₂S(g), P_{H_2S} , was estimated with air-water equilibrium calculations for different total sulphide concentrations in the liquid phase. The calculation method was modified from Hvitved-Jacobsen *et al.* (2013) to account for dissociation, pH- and temperature dependence. The results illustrate that within the operational ranges for AD processes, the pH has a more significant impact on P_{H_2S} compared to temperature (Fig. 5a and b). As an example, the P_{H_2S} at T = 15 °C is 77 % of the P_{H_2S} at T = 55 °C, whereas an increase in pH from 6.8 to 8 reduces the P_{H_2S} to 12 % of the P_{H_2S} at pH 6.8. Despite the higher proportion of sulphide present as H₂S in the liquid phase at lower temperatures, an increase in temperature causes an increase in partial pressure of H₂S in the gas phase. This can be attributed to the increase in Henry's law constant, K_{H,H_2S} , with an increase in temperature.

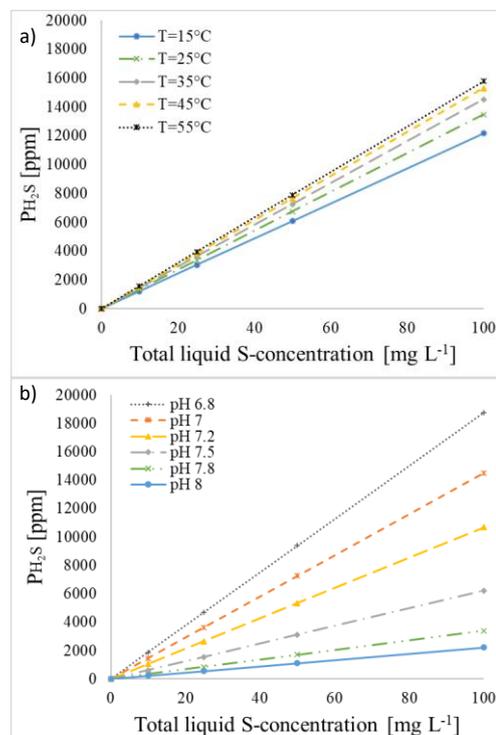


Figure 5: Partial pressure of H₂S estimated with air-water equilibrium calculations, as a function of a) T and b) pH.

Figures 5a and b illustrate that even low total sulphide concentrations in the liquid can cause high H_2S concentrations in biogas. The values should be considered maximum levels as the calculation assumes that all the H_2S in the liquid phase can volatilise. Due to different physical phenomena, including sulphide precipitation and biological or chemical oxidation, the actual values will be lower (Hvitved-Jacobsen *et al.*, 2013). However, due to the concerns and regulations related to releasing $\text{H}_2\text{S}(\text{g})$, this simple method helps predicting the potential concentration of H_2S in the biogas.

3.2. Bioelectrochemical oxidation and the Nernst-Monod term

The anode potential and the number of electrons transferred from the microorganisms to the anode surface are important parameters in BESs. To study the impact of the anode potential, simulations were performed with an increase in the local potential from -0.1 to $+0.3$ V (stepsize = 0.05 V) (Fig. 6). By convention, n is set to 1, as yield and stoichiometric parameters can be defined per electron. The uptake rate of sulphide increases with an increase in the potential up to a certain threshold (here: 0.1 - 0.15 V). With a local potential of 0.1 V or higher, the Nernst-Monod term is close to 1. Consequently, increasing the potential further would not improve the oxidation rate with the current model implementation. With a potential of 0.1 V or higher, Nernst-Monod kinetics resembles single Monod kinetics, as the uptake rate mainly depends on electron donor consumption (substrate consumption).

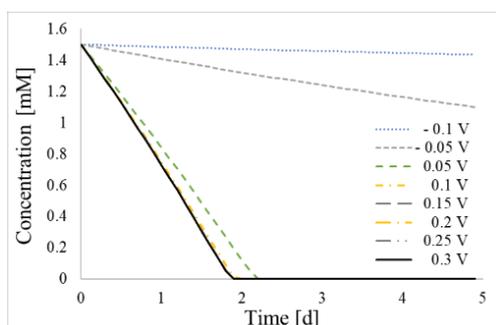


Figure 6: Bioelectrochemical sulphide oxidation rate with different local potentials. Initial biomass: 0.1 mM.

3.3 Case study: practical application of the model

The case study illustrates the differences between the three bio-oxidation processes. The simulation results show that the target concentration is reached within 3.9 days with nitrate, 2.4 days with oxygen, and 1.7 days with an anode potential of 0.1 V (Figures 7a, b, and c) under the given conditions.

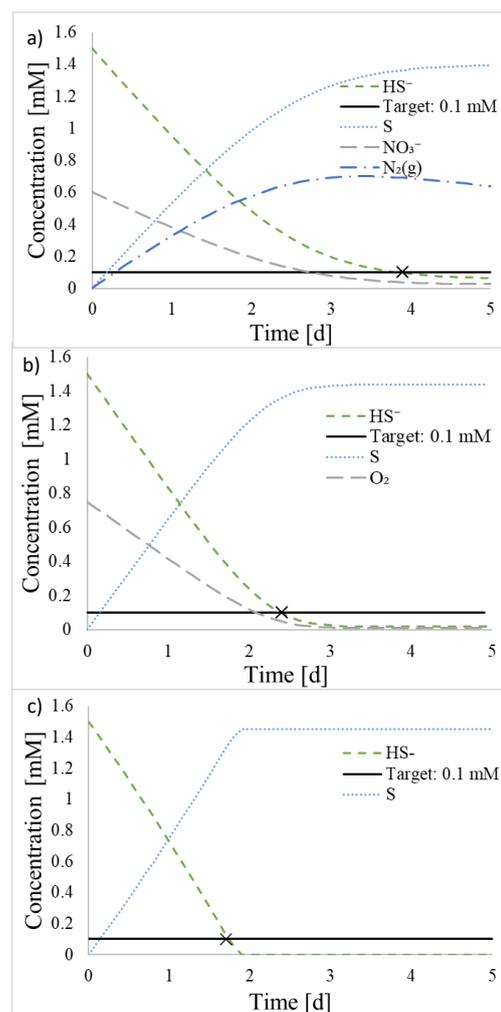


Figure 7: Estimation of required time until the target concentration is reached with a) nitrate, b) oxygen, and c) anode as the electron acceptors.

The estimated nitrate dosage required to ensure a sulphide concentration lower than 0.1 mM in continuous operation is an initial addition of 55.55 kg and a continuous supply of 2.60 kg d^{-1} with the assumed conditions. The corresponding oxygen dosage is an initial supply of 35.84 kg and a continuous supply of 1.68 kg d^{-1} . However, supplying high initial concentrations is unrealistic, as the microorganisms can be negatively affected by too high dosages of nitrate or oxygen. Therefore, simulations were repeated without supplying the initial dosage of nitrate or oxygen. The results show that a steady-state concentration of sulphide below the target concentration of 0.1 mM can be reached in 65.0 days and 76.1 days with a continuous supply of 1.68 kg d^{-1} oxygen and 2.60 kg d^{-1} nitrate, respectively (Fig. 8a and b). However, the required time will depend on different parameters, including the dosage, the microbial uptake rate (see Fig. 3), and the microbial density. By increasing the oxygen

dosage and nitrate dosage by a factor of 1.5, the target concentration is reached in 21.8 and 22.7 days, respectively (results not included). This is a considerable reduction in the required time to reach the target concentration, and it illustrates that the supplied dosage has a significant impact on the simulation results.

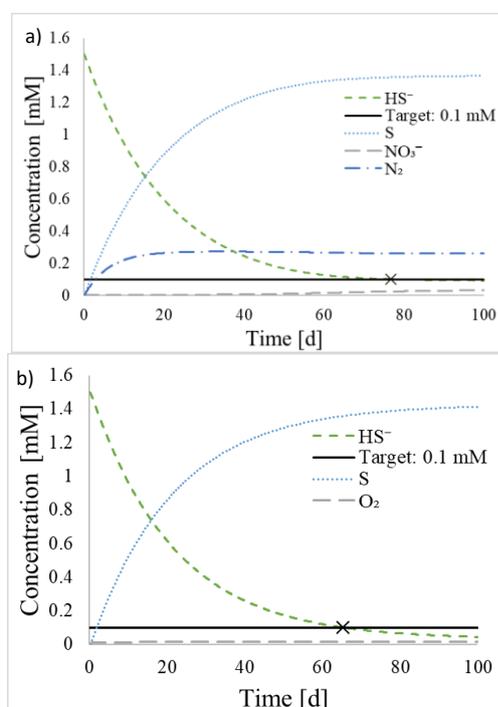


Figure 8: Sulphide removal in continuous operation mode with a) nitrate and b) oxygen.

An increased oxidation rate can be obtained by increasing the initial biomass concentration, using a microbial group with a higher maximum uptake rate, or gradually increasing the nitrate or oxygen dosage. Repeated simulations showed that by increasing the initial biomass from 0.1 to 1 mM, the target concentration of 0.1 mM sulphide was reached in less than 60 days, with nitrate or oxygen as the electron acceptors (results not included). Despite the faster sulphide removal obtained, the oxidation rate was restricted by the supply of the electron acceptor. Alternatively, by increasing the nitrate dosage stepwise every fifth day by 0.06 mM (10 % of the initial dosage), the target sulphide concentration was reached in 29.7 days. The corresponding number of days with oxygen (an increase of 0.075 mM every fifth day) was 28.7 days (results not included).

This case study illustrates the main advantage of developing a simple simulation tool for studying bio-oxidation processes and reactor operations. By adjusting the parameters in the model according to the reactor operation, different strategies could be evaluated. Simulation results can be obtained quickly without negatively affecting the reactor or

the AD processes. Obtaining the same results experimentally would be costly, time-consuming, and potentially lead to reactor failure. An additional advantage is that the calculations can be based on both the concentration of $\text{H}_2\text{S}(\text{g})$ in the headspace or $\text{HS}^-(\text{liq})$ in the effluent streams from plants.

3.4 Evaluation of the model and further model development

The overall goal of the model is to develop a simple tool for estimation of $\text{H}_2\text{S}(\text{g})$ in biogas and to estimate the time, chemical dosage, or anode potential required to reduce $\text{HS}^-(\text{liq})$, $\text{H}_2\text{S}(\text{liq})$ and $\text{H}_2\text{S}(\text{g})$ to acceptable levels. At this modelling stage, a study of the substrate uptake rates and identification of the most sensitive parameters were accomplished. This initial modelling work provides the framework for further development of the model. However, certain model limitations have been identified at the current stage. To improve the model, other anaerobic microbial processes and competing reactions should be implemented. These processes can affect the oxidation rates through competition for substrates and inhibition. The proposed improvements would be more realistic, as the oxidation processes at the current modelling stage are only affected by the substrate concentrations, anode potential, and the specific microbial activity.

To improve the model implementation of bioelectrochemical sulphide oxidation, incorporation of different stoichiometry and other bioanode-related processes would be valuable. This includes biofilm thickness, mass transfer limitations, competing reactions at the bioanode surface, and electron transfer limitations (due to different electron sinks) (Pham *et al.*, 2009). A further study of anode potential implementation can improve the model, as the anode potential has a lower impact on the simulated bioelectrochemical sulphide oxidation rate than expected. Additionally, this model only considers the bioanodic reactions. Sulphide oxidation can contribute positively to biocathodic processes such as methane production by generating electrons and protons in a BES (Jiang *et al.*, 2014; Dykstra *et al.*, 2021). Modelling a complete bioelectrochemical reactor would require implementation of both the oxidation and the reduction processes. In this modelling work, the goal was to study sulphide removal. Therefore, only the bioanodic process was studied.

Lastly, different kinetic constants and parameters can be found in the literature related to the sulphide routes and the specific microbial groups. The variation will affect the simulation results. Therefore, calibration and validation of the model

with data from real plants will improve the model further.

4. Conclusion

The model provides a simple tool in Aquasim for studying H₂S liquid-gas mass transfer and sulphide bio-oxidation with three different electron acceptors: 1) Nitrate; 2) Oxygen; and 3) A bioanode with a given potential. Multiplicative Monod kinetics (nitrate and oxygen) and Nernst-Monod kinetics (bioanode) provide the framework for the respective biotic sulphide oxidation processes.

A sensitivity analysis revealed that the model is most sensitive toward the maximum microbial uptake rate, k_m^o . Low k_m^o -values can result in inefficient sulphide removal, accumulation of sulphide, and high concentrations of H₂S(g). A local potential of 0.1-0.15 V is defined as the plateau potential with Nernst-Monod kinetics in the current study, as further increasing the potential did not improve the sulphide oxidation rate.

At the current stage, the model can be used to study the defined processes. However, certain model limitations have been identified. Therefore, the model needs further improvements to function as a simulation tool for studying sulphide-related processes in AD and calculate the required dosage and oxidation time in full-scale processes.

References

- Angelidaki, I., Treu, L., Tsapekos, P., Luo, G., Campanaro, S., Wenzel, H., and Kougias, P. G. (2018) 'Biogas upgrading and utilization: Current status and perspectives', *Biotechnology Advances*, 36, pp. 452-466. doi:10.1016/j.biotechadv.2018.01.011
- Bae, W. and Rittman, B. E. (1996) 'A structured model of dual-limitation kinetics', *Biotechnology and Bioengineering*, 49, pp. 683-689. doi:10.1002/(SICI)1097-0290(19960320)49:6<683::AID-BIT10>3.0.CO;2-7
- Barrera, E. L., Spanjers, H., Solon, K., Amerlinck, Y., Nopens, I. and Dewulf, J. (2015) 'Modeling the anaerobic digestion of cane-molasses vinasse: Extension of the Anaerobic Digestion Model No. 1 (ADM1) with sulfate reduction for a very high strength and sulfate rich wastewater', *Water Research*, 71, pp. 42-54. doi:10.1016/j.watres.2014.12.026.
- Batstone, D. J. et al., (2002) *Anaerobic Digestion Model No. 1 (ADM1)*, IWA Publishing, UK
- Broderius, S. J. and Smith, L. L. (1977) 'Direct determination and calculation of aqueous sulfide', *Analytical Chemistry*, 49, pp. 424-428. doi:10.1021/ac50011a024
- Dykstra, C. M. and Pavlostathis, S. G. (2021) 'Hydrogen sulfide affects the performance of a methanogenic bioelectrochemical system used for biogas upgrading', *Water Research*, 200. doi:10.1016/j.watres.2021.117268
- Hvitved-Jacobsen, T., Vollertsen, J., and Nielsen, A. H. (2013) *Sewer Processes: Microbial and chemical process engineering of sewer networks*, 2nd edition, Taylor and Francis Group.
- Jiang Y., Su M. and Li D. (2014) 'Removal of sulfide and production of methane from carbon dioxide in microbial fuel cells-microbial electrolysis cell (MFCs-MEC) coupled system'. *Applied Biochemical Biotechnology*, 172, pp.2720-2731. doi:10.1007/s12010-013-0718-9
- Marcus, A. K., Torres, C. I., and Rittman, B. E. (2007) 'Conduction-based Modeling of the Biofilm Anode of a Microbial Fuel Cell', *Biotechnology and Bioengineering*, 98, pp. 1171-1182. doi:10.1002/bit.21533.
- Pham, T. H., Aelterman, P., and Verstraete, W. (2009) 'Bioanode performance in bioelectrochemical systems: recent improvements and prospects' *Trends in Biotechnology*, 27, pp. 168-178. doi:10.1016/j.tibtech.2008.11.005
- Pokorna-Krayzelova, L., Veimelkova, D., Selan, I., Jenicek, P., Volcke, E. I. P., Bartacek, J. (2018) 'Final products and kinetics of biochemical and chemical sulfide oxidation under microaerobic conditions', *Water Science*, 78, pp. 1916-1924. doi:10.2166/wst.2018.485.
- Samarakoon, G., Nelabhotla, A. B., Dinamarca, C., Winkler, D., and Bakke, R. (2019) 'Modelling Bio-electrochemical CO₂ Reduction to Methane'. In Proceedings of the 10th Trondheim Conference on Carbon Capture, Transport and Storage, Trondheim, Norway, 17.-19. June 2019'.
- Sander, R. (1999) *Compilation of Henry's law constants for inorganic and organic species of potential importance in environmental chemistry*, 3rd edition.
- Sun M., Mu Z.-X., Chen Y.-P., Sheng G.-P., Liu X.-W., Chen Y.-Z., Zhao Y., Wang H.-L., Yu H.-Q., Wei L. and Ma F. (2009) 'Microbe-assisted Sulfide Oxidation in the anode of a microbial fuel cell'. *Environmental Science and Technology*, 43;3372-3377. doi:10.1021/es802809m
- Sun, J., Dai, X., Peng, L., and Ni, B.-J. (2017) 'Sulfide removal and sulfur production in a membrane aerated biofilm reactor: Model evaluation', *Chemical Engineering Journal*, 309, pp. 454-462. doi:10.1016/j.cej.2016.09.146.
- Yongsiri, C., Vollertsen, J., Rasmussen, M., and Hvitved-Jacobsen, T., (2004) 'Air-Water Transfer of Hydrogen Sulfide: An Approach for Application in Sewer Networks', *Water Environment and Research*, 76, pp. 81-88. doi:10.2175/106143004X141618
- Wanga, A., Liu, C., Rena, N., Hana, H., and Lee, D. (2010) 'Simultaneous removal of sulfide, nitrate and acetate: Kinetic modelling', *Journal of Hazardous Materials*, 178, pp. 35-41. doi:10.1016/j.jhazmat.2010.01.039.

Optimal operation of grid-connected hydropower plants through voltage control methods

Emil G. Melfald ^{a,*}, Thomas Øyvang ^a,

^a University of South-Eastern Norway, Department of Electrical Engineering, IT and Cybernetics, Campus Porsgrunn

Emil.G.Melfald@usn.no

Thomas.Oyvang@usn.no

Abstract

The ongoing decarbonization, and the rapid increase in renewable penetration in the electric grid, will demand enhanced flexible operational schemes of the conventional hydropower plants. This paper explores a grid-connected hydropower plant's best-efficiency operating conditions when meeting the renewable energy transition. Power loss models combined with various voltage control methods are investigated for achieving optimal operation. This simulation study is carried out on a static Single Machine Infinite Bus (SMIB) environment to perform voltage control comparisons. Simulations show that the plant and grid power loss models can be utilized in an optimal controller setting to increase the accumulated average efficiency (AAE). However, optimal controllers had slow prediction times, and therefore a Reinforcement Learning (RL) method, A2C, has been trained to learn an optimal control policy that maximizes system efficiency. The RL agent supersedes the optimal control techniques with up to 40 times faster prediction times.

1 Introduction

In an energy mix with intermittent renewable energy sources such as wind and solar, the flexible operation of reliable reservoir-based hydropower plants can be the backbone for energy balance and increased stability (Pandey et al., 2021)(Abadie et al., 2020). Synchronous machines have inherent stability characteristics in the form of, e.g., rotational inertia and damper bars. Stability characteristics are further enhanced by controllers for frequency and voltage, e.g., the governor (GOV) and automatic voltage regulator (AVR) (*IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*, 2016). Together they provide essential system services (constant frequency and voltage) for the operation of the electricity grid.

The Transmission System Operator (TSO) is responsible for the operational security of supply and power transfer capability while reducing investment costs and transmission losses (NVE, 2019). The active power dispatch is correlated to market demand by the TSO, while the voltage levels are considered more as free control variables in the power system operation. Through traditional centralized optimization techniques (Hasan et al., 2020), e.g., Optimal Power Flow (OPF) (Wang et al., 2017), grid losses and generation costs can be minimized with respect to active power production and terminal voltages (Qiu et al., 2009).

There has also been done much research on decentralized and distributed control schemes (Molzahn et al., 2017). This technique solves optimization problems by considering control of smaller grid sections instead of the complete system. Another approach for optimal control is reinforcement learning (RL). RL is a machine learning approach tailored to learning control policies to maximize an objective function, often referred to as a reward function (Buşoniu et al., 2018). Moreover, RL

has been implemented in recent literature, e.g., for finding feasible grid voltage setpoints (Wang et al., 2020)(Duan et al., 2020), and energy management systems (Chen et al., 2021).

The use of power loss models of grid (Kundur, 1994) and synchronous machines (Bortoni et al., 2020) has the potential for more efficient voltage control.

This paper presents a simulation study comparing standard decentralized voltage control methods against optimal power system controllers. Controller schemes for efficiency maximization will be formulated and simulated in a grid-connected single machine infinite bus (SMIB) environment. Power losses are considered for the waterway, turbine, generator, and grid. Eight different controller methods under steady-state operation are considered. Four traditional primary setpoint control strategies of the AVR will be investigated. Three optimization-based controllers will be implemented and tested for the minimization of (i) generation losses, (ii) transmission losses, and (iii) overall losses. Finally, the A2C (Mnih and et al., 2016) reinforcement learning algorithm will be trained to maximize the system efficiency (plant and grid). The computational speed of RL compared to optimal controllers will be evaluated.

The paper is organized as follows. Section 2 introduces the power loss modeling for hydropower plants, while Section 3 introduces the different voltage controllers under study. Section 4 presents the study case and simulation results.

2 Power and energy loss modeling

In hydropower systems, the potential and kinetic energy in the water is converted first into mechanical energy before being converted to electrical energy. The electrical energy is then transported in cables, lines, and transformers,

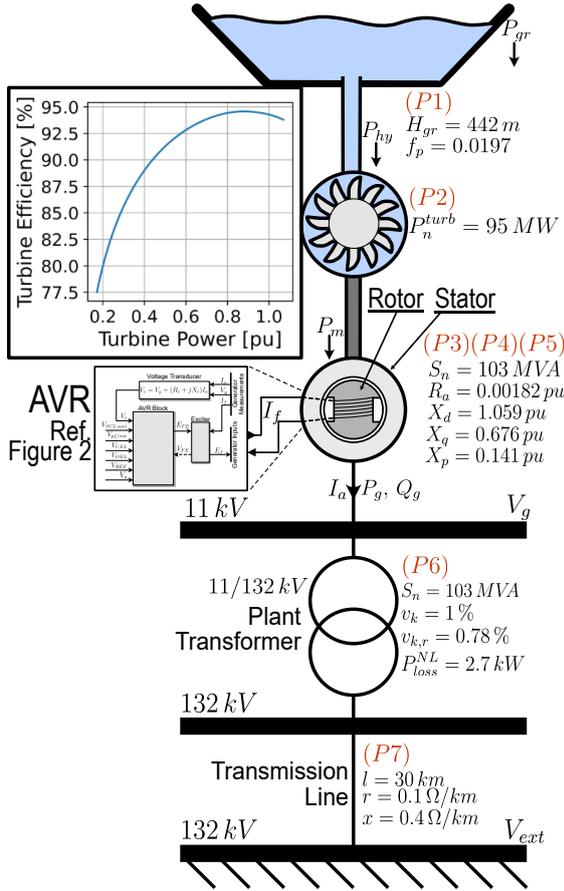


Figure 1: An overview of the SMIB topology and parameters. Power loss P1 to P7 are indicated close to the respective component.

which have some power losses. The power losses in a single machine infinite bus (SMIB) environment are proposed to be divided into seven different parts, e.g., (P1) to (P7). These losses are presented from top to bottom in Figure 1. The hydraulic losses (P1) represent water conduit friction losses from water movement. Turbine losses (P2) represent the power loss in the transition from pressure and kinetic energy to mechanical energy on the rotor shaft. The generator losses are divided into three separate categories: stator loss (P3), rotor loss (P4), and constant losses (P5). *Stator loss* (P3) is the ohmic losses $R_a I_a^2$ in the armature coils and stray loss. *Rotor loss* (P4) is the ohmic loss $R_f I_f^2$ from the field and excitation currents in the excitation circuit. *Constant losses* (P5) are friction and windage losses, core loss, and bearing losses. As long as the grid-connected machine's voltage and speed are constant, the constant losses will not change. Moreover, the transformer losses (P6) consist of constant and load-dependent losses. Finally, the transmission line losses (P7) are the ohmic losses in the transmission line. (Chapman, 2012)

2.1 Synchronous generator loss modeling

The calculation framework that was established in Bortoni et al. (Bortoni et al., 2020) is used for estimating the synchronous machine losses. This method extrapolates the machine losses by using known relations between armature current I_a , field current I_f , terminal voltage

V_g , and nominal losses noted with *. *Stator loss* (P3) is extrapolated from nominal armature and stray losses, P_a^* and P_s^* respectively, as shown in Eq. 1. Armature current I_a should be known beforehand and is calculated from Eq. 2, implying that P_g , Q_g , and V_g are known.

$$P_{loss}^{stator} = (P_a^* + P_s^*) \left(\frac{I_a}{I_a^*} \right)^2 \quad (1)$$

$$I_a \angle -\varphi = \frac{P_g - jQ_g}{V_g} \quad (2)$$

Rotor loss (P4) is extrapolated from nominal field, brush, and excitation losses, P_f^* , P_{br}^* , and P_{ex}^* respectively, and depends on the field current I_f as shown in Eq. 3. For estimating the field current at any value of P_g , Q_g , and V_g , the generator's no-load and short-circuit characteristic curves are required. These curves are usually provided by the manufacturer or measured during commissioning of the machine. The no-load characteristics establish a relation between the internal emf E_g and field current I_f , usually linear. The short-circuit characteristics relates field and armature currents and capture saturation in the core. The saturation can be represented as the Potier reactance (X_p) and Potier emf (E_p) (Kundur, 1994). To calculate the field current I_f , the rotor angle δ and internal emf E_g is calculated with Eq. 4 (Kundur, 1994) and Eq. 5 (Chapman, 2012), respectively. The Potier voltage angle θ and emf E_p is calculated similarly through Eq. 6 and 7. Field current in the rotor circuit can then be estimated using Eq. 8. The parameters b_v , k , C_m , and m are obtained through curve fitting Eq. 8 to measured values of I_f in the short-circuit characteristic of the machine (Karikezi et al., 2021).

$$P_{loss}^{rotor} = (P_f^* + P_{br}^*) \left(\frac{I_f}{I_f^*} \right)^2 + P_{ex}^* \left(\frac{I_f}{I_f^*} \right) \quad (3)$$

$$\delta = \tan^{-1} \left(\frac{I_a (X_q \cos(\varphi) - R_a \sin(\varphi))}{V_g + I_a (X_q \sin(\varphi) + R_a \cos(\varphi))} \right) \quad (4)$$

$$E_g = V_g \cos(\delta) + I_a X_d \sin(\delta + \varphi) + R_a I_a \cos(\delta + \varphi) \quad (5)$$

$$\theta = \tan^{-1} \left(\frac{I_a (X_p \cos(\varphi) - R_a \sin(\varphi))}{V_g + I_a (X_p \sin(\varphi) + R_a \cos(\varphi))} \right) \quad (6)$$

$$E_p = V_g \cos(\theta) + I_a X_p \sin(\theta + \varphi) + R_a I_a \cos(\theta + \varphi) \quad (7)$$

$$I_f = \frac{E_g - E_p}{b_v} + k(E_p + C_m E_p^m) \quad (8)$$

Constant losses (P5) is extrapolated from the nominal core, bearing, friction and windage losses, noted P_c^* , P_b , and P_{wf}^* , respectively, where the core losses is dependent on the terminal voltage V_g , as shown in Eq. 9.

$$P_{loss}^{const} = P_c^* \left(\frac{V_g}{V_g^*} \right)^2 + P_b^* + P_{wf}^* \quad (9)$$

Nominal loss and saturation parameters for this paper's study case are displayed in Table 1 together with a short description of the parameters. The presented values are utilized for the study case calculations described in Section 4.

2.2 Hydraulic and turbine loss modeling

Hydraulic losses in the hydropower plant are separated into two parts, e.g., turbine losses and waterway losses, (P1) and (P2) in Figure 1 respectively. (P2) can be

Table 1: General Data for Loss Calculations of 103 MVA synchronous generator, turbine, and water way

Symbol	Description	Value
b_v	Saturation parameter	1.0 [-]
C_m	Saturation parameter	0.16 [-]
m	Saturation parameter	7 [-]
k	Saturation parameter	1.0308 [-]
I_a^*	Nominal Armature Current	5406.1 [A]
I_f^*	Nominal Field Current	1065 [A]
V_g^*	Nominal Generator Terminal Voltage	1 [pu]
P_a^*	Nominal Armature Losses	187.46 [kW]
P_s^*	Nominal Stray Losses	89.16 [kW]
P_f^*	Nominal Field Current Losses	173.65 [kW]
P_{br}^*	Nominal Brush Losses	2.13 [kW]
P_{ex}^*	Nominal Exciter Losses	15.88 [kW]
P_c^*	Nominal Core Losses	211.92 [kW]
P_b^*	Bearing Losses	240.9 [kW]
P_{wf}^*	Windage and Friction Losses	172.92 [kW]
q_w^*	Nominal Water Flow Rate	21.843 [m ³ /s]
H_{gr}	Gross water head	442 [m]
f_p	Water way friction coefficient	0.0197 [-]

estimated from Eq. 10. The mechanical rotor power P_m is the generator output power P_g plus generator rotor, stator, and constant losses. The hydraulic power P_{hy} is the input power to the turbine before the conversion between kinetic/pressure water energy to mechanical energy and is calculated from Eq. 11.

$$P_{loss}^{turb} = P_{hy} - P_m = P_m \left(\frac{1}{\eta_{turb}} - 1 \right) \quad (10)$$

$$P_{hy} = \frac{P_m}{\eta_{turb}} \quad (11)$$

Calculating (P1) is done by solving Eq 12 concerning q_w . The gross power P_{gr} is the potential power from the water, assuming no losses, and is shown in Eq. 13. P_{loss}^H is estimated by the friction loss in the water way, shown in Eq. 14. ρ is water density in [kg/m³], g is the acceleration of gravity in [m/s²], H_{gr} is the gross water head in [m], H_{loss} is the head loss from friction [m], and q_w is the water flow rate [m³/s]. Inserting Eq. 11, 13, and 14 into Eq. 12, yields a third degree polynomial for q_w , shown in Eq. 15. All values are known except for q_w , and this equation can therefore be solved. Assuming $q_w \in [0, q_w^*]$ only one solution to q_w is valid and used for the loss calculation.

$$P_{gr} = P_{loss}^H + P_{hy} \quad (12)$$

$$P_{gr} = \rho g H_{gr} q_w \quad (13)$$

$$P_{loss}^H = \rho g H_{loss} q_w = \rho g f_p q_w^3 \quad (14)$$

$$\rho g (H_{gr} q_w - f_p q_w^3) - \frac{P_m}{\eta_{turb}} = 0 \quad (15)$$

2.3 Efficiency Calculations and AAE

The power loss calculation framework for the hydropower plant and transmission system has the following workflow. Firstly, the Active power P_g is defined. Then, the applied voltage controller determines the reactive power dispatch Q_g , and a power flow calculation is executed. The generator stator, rotor, and constant losses (P3, P4, P5) are calculated. Moreover, turbine loss (P2) and waterway loss (P1) are then calculated. From P_g and Q_g , the transformer loss P_{loss}^T (P6) and transmission loss P_{loss}^l (P7) are calculated through the Pandapowers power flow calculation (Thurner et al., 2018). For convenience, grid

losses P_{loss}^{grid} and plant losses P_{loss}^{plant} is defined in Eq. 16 and 17, respectively.

$$P_{loss}^{grid} = P_{loss}^T + P_{loss}^l \quad (16)$$

$$P_{loss}^{plant} = P_{loss}^{stator} + P_{loss}^{rotor} + P_{loss}^{const} + P_{loss}^{turb} + P_{loss}^H \quad (17)$$

Overall system efficiency calculation is shown in Eq. 18. The numerator represents the active power reaching the external grid, while the denominator is the sum of active power production and losses from (P1) to (P7).

$$\eta = \frac{P_g - P_{loss}^{grid}}{P_g + P_{loss}^{grid} + P_{loss}^{plant}} \quad (18)$$

There are different ways of estimating the efficiency of a time series data set. One method is to average the efficiencies over the data. Another method proposed by (Karikezi et al., 2021) is the AAE (Accumulated Average Efficiency). In a data set of varying active power production, the AAE is a more accurate representation of the energy losses of the system compared to averaging over the efficiencies. Eq. 19 shows how AAE is calculated, assuming evenly spaced data points. The sum in the numerator represents the total energy production reaching the external grid (customers) in the data set. The sum in the denominator represents the total energy production and all losses.

$$\eta_{AAE} = \frac{\sum_{i=1}^N [P_g^{(i)} - P_{loss}^{grid(i)}]}{\sum_{i=1}^N [P_{loss}^{grid(i)} + P_{loss}^{plant(i)} + P_g^{(i)}]} \quad (19)$$

3 Voltage control methods

The rotor induces an internal emf $E_g \angle \delta$, determining the terminal voltage V_g and how much reactive power production/consumption the synchronous generator (SG) has. An approximate relation between the internal voltage and reactive power production is shown in Eq. 20 (Kundur, 1994).

$$Q_g = \frac{E_g^2 - E_g V_g \cos \delta}{X_d} \quad (20)$$

The excitation system of the SG, depicted in Figure 2 (*IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*, 2016), determines both the terminal voltage and reactive power production. Transducers measure V_g , P_g , and Q_g on the generator terminal. These signals go through a voltage compensation block that calculates V_c . In this paper, the automatic voltage regulator AVR has two operating modes, indicated by switch SW1 and SW2 in the AVR block. If SW1 is closed, the AVR is in a voltage setpoint tracking mode. Two different controller modes are determined from this AVR state: constant voltage controller (C1) and voltage compensated controller (C2). Constant voltage control (C1) bypasses the voltage compensation block in Figure 2, effectively setting controller parameters R_c and X_c to 0. This mode controls the generator exciter such that $V_g = V_{REF}$, where V_{REF} is a voltage reference. Controller (C2) includes the voltage compensation block and is a setpoint controller that forces the resulting compensated voltage $V_c = V_{REF}$.

When SW2 is closed (and SW1 is open) the AVR is in reactive power setpoint mode. This AVR state can define

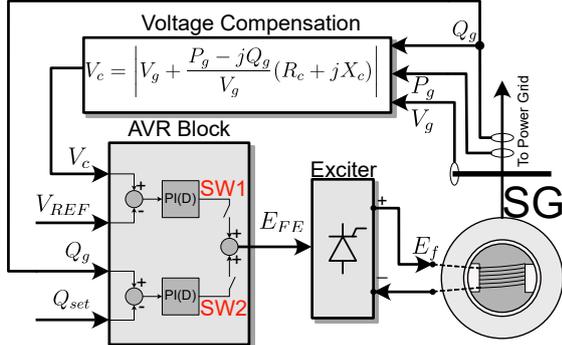


Figure 2: Block diagram of the AVR, exciter, synchronous machine, and voltage transducer (*IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*, 2016).

two controller modes: constant Q_g controller (C3) and constant power factor controller (C4). Controller (C3) adjusts the excitation current such that the reactive power production reaches some setpoint value $Q_g = Q_{set}$. Controller (C4) updates Q_{set} continuously according to the active power production, such that $Q_{set} = P_g \tan \varphi_{set}$, where φ_{set} is the setpoint phase angle. Controllers (C1) to (C4) are standard control modes for the AVR. No information on generator or grid losses is explicitly used in these controllers. Therefore, three optimal controllers and one RL controller are implemented, which produces optimal values of Q_{set} to the AVR, according to some objective function. The first optimal controller is an optimal plant efficiency control (C5) to minimize plant losses. The second optimal controller is the optimal grid efficiency (C6), which minimizes grid losses. The third optimal controller is an optimal system controller (C7), which minimizes the sum of plant and grid losses. In addition, a reinforcement learning (RL)-based controller (C8) is implemented to minimize system losses. All controllers, numbered (C1) to (C8) is listed in Table 2 with descriptions of controller objectives and tuned controller parameters.

3.1 Implementation of voltage controllers C1 to C7

Controller (C1) is equal to (C2), with the only difference being that parameters R_c and X_c are zero for (C1) and not for (C2). Therefore, the implementation for these two controllers is equal. In this paper, the generator bus is modeled as a PQ bus. This means P_g and Q_g have to be specified in the power flow calculation. Converting a voltage setpoint V_{REF} to a reactive power Q_g can be done through Eq. 21 - 23 and Algorithm 1. The voltage V_c is adjusted towards a voltage setpoint V_{REF} and is calculated as the generator terminal voltage plus the voltage drop over a fictitious compensator impedance $Z_c = R_c + jX_c$ (*IEEE Recommended Practice for Excitation System Models for Power System Stability Studies*, 2016). The compensated voltage V_c is expressed in Eq. 21, where it is assumed that terminal voltage V_g has an angle of zero.

$$V_c \angle -\theta_c = V_g + \frac{1}{V_g} (P_g - jQ_g)(R_c + jX_c) \quad (21)$$

Separating Eq. 21 into real and imaginary parts results in Eq. 22 and 23. Assuming steady-state operation, V_c must be equal to the voltage reference, $V_c = V_{REF}$. Eq. 21 and

Table 2: Controller overview with a description of the controllers. Optimized controller parameters Θ^{opt} is displayed for controllers (C1) to (C4). Controller objectives are shown for optimal controllers (C5) to (C7), and the reward function \mathcal{R} for the RL controller (C8).

Controller Name	Description	Tuned Controller Parameters/ Controller Objective
Const V (C1)	Setpoint control of V_g	$\Theta^{opt} \leftarrow V_{REF} = 1.02 \text{ pu}$
V-comp (C2)	Setpoint control of V_c	$\Theta^{opt} \leftarrow \begin{cases} V_{REF} = 1.00 \text{ pu}, \\ R_c = -0.03 \text{ pu}, \\ X_c = 0.043 \text{ pu} \end{cases}$
Const Q (C3)	Setpoint control of Q_g	$\Theta^{opt} \leftarrow Q_g^{set} = 7.01 \text{ Mvar}$
Const φ (C4)	Setpoint control of φ	$\Theta^{opt} \leftarrow \varphi = 0.002 \text{ rad}$
Opt Plant (C5)	Minimizes plant losses controlling Q_{set}	$Q_{set} = \min_{Q_{set}} (P_{loss}^{plant})$
Opt Grid (C6)	Minimizes grid losses controlling Q_{set}	$Q_{set} = \min_{Q_{set}} (P_{loss}^{grid})$
Opt Sys (C7)	Minimizes plant + grid losses controlling Q_{set}	$Q_{set} = \min_{Q_{set}} (P_{loss}^{grid} + P_{loss}^{plant})$
RL control (C8)	RL control predicting Q_{set} for optimal system efficiency.	$\mathcal{R} = \left(\frac{P_g}{P_g + P_{loss}^{gen} + P_{loss}^{grid}} \right)^2$

22 is a system of two equations with the two unknowns being the terminal voltage V_g and compensated voltage angle θ_c . Solving this system is done iteratively by using an initial power flow as a guess for V_g , which will be adjusted such that V_c converges to the reference voltage V_{REF} . Algorithm 1 shows this procedure.

$$V_g^2 - V_g V_c \cos(\theta_c) + P_g R_c + Q_g X_c = 0 \quad (22)$$

$$V_g V_c \sin(\theta_c) + P_g X_c - Q_g R_c = 0 \quad (23)$$

Algorithm 1: Constant V controller (C2 and C4) at PQ bus.

Data: $Q_{set}^{(1)} = 0 \text{ pu}$, $Q_{set}^{(2)} = 0.1 \text{ pu}$
Do a power flow with $Q_{set}^{(1)}$ and $Q_{set}^{(2)}$ to obtain $V_c^{(1)}$ and $V_c^{(2)}$ by solving Eq. 22 and 23
while $|Q_{set}^{(2)} - Q_{set}^{(1)}| > 1e-3$ **do**
 $\frac{dQ_{set}}{dV_c} \leftarrow \frac{Q_{set}^{(2)} - Q_{set}^{(1)}}{V_c^{(2)} - V_c^{(1)}}$
 $Q_{set}^{new} \leftarrow Q_{set}^{(2)} + \frac{dQ_{set}}{dV_c} (V_{REF} - V_c^{(2)})$
 $V_c^{new} \leftarrow$ Power flow with $Q_g = Q_{set}^{new}$
 $V_c^{(1)} \leftarrow V_c^{(2)}$, $V_c^{(2)} \leftarrow V_c^{new}$
 $Q_{set}^{(1)} \leftarrow Q_{set}^{(2)}$, $Q_{set}^{(2)} \leftarrow Q_{set}^{new}$
end
Result: Set $Q_{set} = Q_{set}^{(2)}$

Constant reactive power (C3) and constant power factor controller (C4) have, in essence, similar controller mechanisms, where a reactive power setpoint from the generator terminal Q_{set} is determined from the controllers and applied to the PQ bus directly.

The optimal controllers (C5, C6, and C7) utilize a minimization algorithm to achieve the controller objectives presented for each controller in Table 2. The

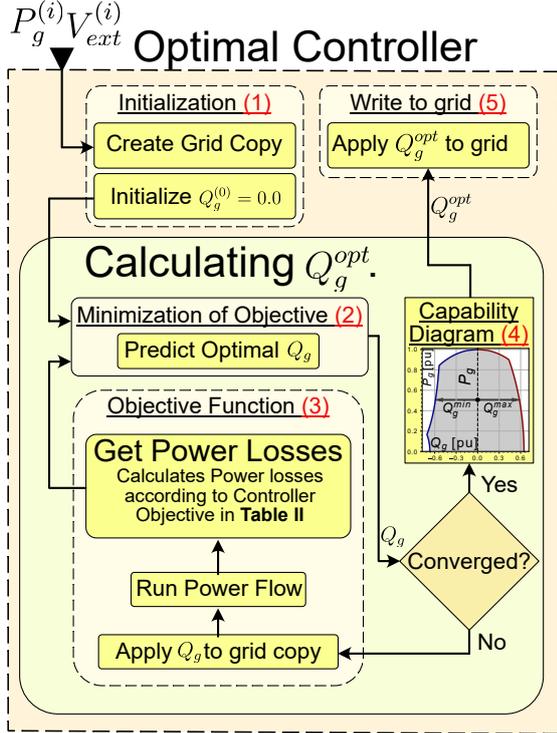


Figure 3: The figure presents the workflow diagram of the voltage controller C8 for calculating the optimal efficiency operation point Q_g^{opt} through step (1) to (4).

decision variable for the optimization is the reactive power setpoint Q_{set} .

The workflow diagram for the optimal controllers is illustrated in Figure 3, where the workflow steps are noted chronologically from step (1) to (5). The workflow diagram is explained in the sequel: In a time series simulation, an active power $P_g^{(i)}$ and external grid voltage $V_{ext}^{(i)}$ is given to the controller. The controller is initialized (1) by creating a grid copy, and the initial value for Q_g is specified at 0.0. A grid copy is made because the controller modifies the grid during the minimization of losses, and these modifications are not wanted in the original grid. The initial condition, a grid copy, and the objective function (3) is sent to the *minimize_scalar* function (2) from (Virtanen and et al., 2020), and utilizes the *brent* method. Step (2) and (3) is done iteratively to calculate an optimal value of Q_g^{opt} . When the minimization has converged, the optimal reactive power is bounded between a minimum and maximum allowed value, dictated by the capability diagram (4) (*IEEE Standard for Salient-Pole 50 Hz and 60 Hz Synchronous Generators and Generator/Motors for Hydraulic Turbine Applications Rated 5 MVA and Above*, 2006). The resulting reactive power value is then applied to the grid (5).

3.2 Reinforcement Learning Controller C8

Controller (C8) is a reactive power setpoint control based on the reinforcement learning method A2C (Mnih and et al., 2016). The utilized algorithm is implemented in StableBaselines 3 (Raffin et al., 2021), and the training environment for the agent is defined in the Gym framework (Brockman et al., 2016). The training environment is defined as a one-step deterministic environment, where the states $S = \{V_{ext}, P_g\}$, and the

Table 3: Tuned hyperparameters of the A2C Algorithm in SB3 on the Power System Loss Environment.

Hyperparameter	Value
Policy Neural Network Size	[8, 8]
Value Neural Network Size	[8, 8]
Steps before updating	32
Discount Factor γ	1.0
Learning Rate	0.01
Value function coefficient for the loss calculation	0.6

action is $\mathcal{A} = \{Q_g\}$. During training, an episode is initialized by sampling state values P_g and V_{ext} from a uniform distribution, where $P_g [pu] \in [0.2, 1.0]$ and $V_{ext} [pu] \in [0.9, 1.1]$. The random sampling causes no two training episodes to be equal, and the agent should therefore learn to generalize for all allowed values of P_g and V_{ext} . Eq. 24 defines the reward function \mathcal{R} that is utilized to find an optimal system control policy for Q_{set} .

$$\mathcal{R}(S, \mathcal{A}) = \left(\frac{P_g}{P_g + P_{loss}^{gen} + P_{loss}^{grid}} \right)^2 \quad (24)$$

After a grid search, the highest resulting reward determined the A2C hyperparameters, with the results shown in Table 3. Hyperparameters not displayed in the table have default values from the SB3 library. Both the policy and value neural network sizes are relatively small and shallow. These small networks most likely caused faster training compared to larger networks. However, a higher learning rate of 0.1 did not yield improved results. The discount factor was not part of the search and was arbitrarily set to 1.0 as it doesn't affect the training in one-step environments.

3.3 Parameter tuning of controllers C1 to C4.

As shown in Table 2, each controller (C1) to (C4) have different controller parameters noted as Θ . To find optimal parameter setting for each controller, parameters Θ is defined as a decision variable in an optimization problem, as shown in Eq. 25. The objective function is to minimize overall system power losses for each controller as shown in Eq. 26. The system losses is defined in Eq. 27. Nine operating points close to the nominal (rated) operation is used in the objective function to achieve optimal controller parameters Θ_{opt} . Three values of external grid voltage is chosen, where $V_{ext} [pu] = \{0.99, 1.0, 1.01\}$. For each external grid voltage level, three active power levels are chosen, where $P_g [pu] = \{0.9, 0.95, 1.0\}$. These operating points is represented as the two summations in Eq. 26. The optimization results of Θ for each controller is shown in Table 2 as Θ^{opt} .

$$\Theta_{opt} = \min_{\Theta} J(\Theta) \quad (25)$$

$$J(\Theta) = \sum_{i=1}^3 \sum_{j=1}^3 [P_{loss}(P_g^{(i)}, V_{ext}^{(j)}; \Theta)] \quad (26)$$

$$P_{loss}(P_g^{(i)}, V_{ext}^{(j)}; \Theta) = P_{loss}^{plant} + P_{loss}^{grid} \quad (27)$$

4 Simulation results of a 103 MVA study case

A SMIB (Single Machine Infinite Bus) testbed with the presented controllers was established on a grid-connected 103 MVA hydro generator. The system overview is shown in Figure 1, and system parameters are listed in Table 1.

4.1 Controller efficiency characteristics

Figure ?? shows the controller efficiency characteristics for a range of active power and external voltage values for all controllers. Efficiencies are estimated by Eq. 18 for a range of external voltage (x-axis) and active power (y-axis) combinations. The P_g and V_{ext} axis each has 30 evenly spaced data points on which the contours are based. A reactive power dispatch is calculated from the voltage controllers at each point, followed by a power flow and efficiency calculation. In addition, the yellow “x” in each subplot in Figure 4 indicates the Best Efficiency Point (BEP). BEP is the point where the values of P_g and V_{ext} yield the highest system efficiency for a given controller. Moreover, the efficiency value for each BEP is shown in Table 4.

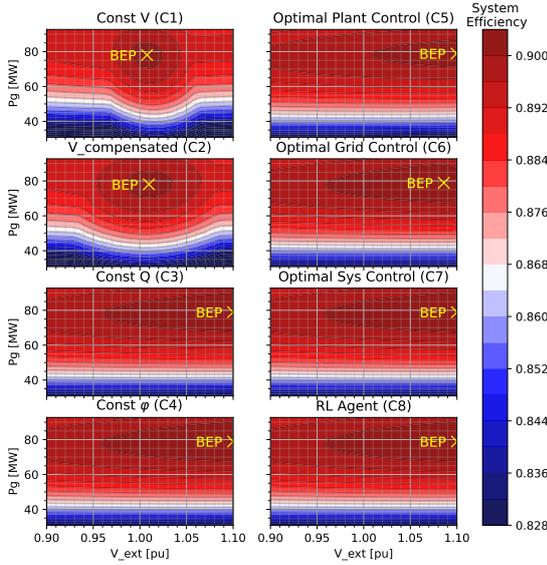


Figure 4: Efficiency characteristic plots for the eight presented controllers. Moreover, the BEP point for each controller is highlighted.

4.2 Controller performance indicators and data set

Controller performance is evaluated through two indicators. These are i) the energy losses in the grid, plant, and system and ii) the voltage variation at the generator bus. These indicators will be evaluated by simulating the SMIB environment with an artificially made data set, which is displayed in Figure 5. In total, there are 8760 values of P_g and V_{ext} . The data set is available in (Melfald, n.d.), and was generated by sampling the active power and external voltage values from normal distributions. When making the data set, the average active power value was set to 0.7 pu with a standard deviation of 0.3, $P_g \sim \mathcal{N}(0.7, 0.3^2)$. Sampled values was bounded between 0.3 and 0.9, such that $P_g \in [0.3, 0.9]$. The average external voltage value was set to 1.0 pu with a standard deviation of 0.02 with no bounds, $V_{ext} \sim \mathcal{N}(1.0, 0.02^2)$. It is assumed to be a one-hour time-step per data point.

4.3 Energy losses and efficiencies from simulation

The energy losses, and terminal generator voltage, are obtained from the simulation of the 8760 data points. Results from the simulation are shown in Table 4. The grid, plant, and system losses are calculated from accumulated power losses. A system

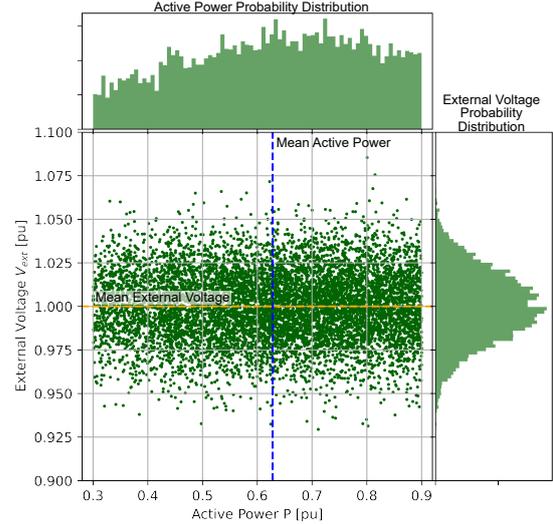


Figure 5: Data Points used for efficiency comparison between voltage controllers. (Melfald, n.d.)

energy loss comparison is made with ΔP_{add}^{sys} . This variable is obtained by subtracting the system losses of any controller with the system losses of controller (C7). This represents additional system losses generated by not utilizing controller (C7). Similarly, the additional grid losses ΔP_{add}^{grid} is comparing the grid losses when using (C6) against the other controllers. Additional plant losses ΔP_{add}^{plant} are also compared between controller (C5) and the rest of the controllers.

4.4 Voltage variation comparisons

The voltage variation on the generator terminal V_g for different controllers is shown and compared in the box plot in Figure 6 (a). The box plot shows four properties of the voltage variation in the study case for each controller. First, the green horizontal line in each box represents the median value of the voltages. The box’s upper and lower vertical boundaries represent the middle 50 percent of the voltage points, assuming a sorted data set. There is no box for (C1) because the voltage variation is negligible for large parts of the data set. The lines stretching out vertically from the boxes are called “whiskers” and show the boundaries between the data and outliers, represented as small circles outside the whiskers. The lower outliers are defined as data points with voltage values lower than 0.01 % of the data set. Upper outlier data points are values above 99.9 % of the data set. In addition, arrows to the left of all boxes and whiskers are displayed with a number. The arrows and number represent the maximum voltage variation [%] in the positive and negative direction from the median voltage in the simulation.

The voltage variation in Figure 6 (a) shows that controllers (C3, C4, C6, C7, C8) have similar generator voltage variation throughout the simulation. This similarity indicates that the reactive power dispatch is similar for the controllers, which is shown with the reactive power distributions in Figure 6 (b). On average, the optimal plant controller (C5) has lower terminal voltage values during the simulation. The main reason for this is that the optimal plant operation for the generator is close to 0.2 reactive power consumption, as described by (Karikezi et al., 2021). This reactive power consumption lowers the terminal voltage compared to the other controllers with a higher reactive power dispatch.

Table 4: Result Data from power flow simulations of the 8760 data points.

	Const V C1	V-comp C2	Const Q C3	Const φ C4	Opt Plant C5	Opt Grid C6	Opt Sys C7	RL Agent C8
AAE [%]	87.6	87.812	87.914	87.936	87.754	87.932	87.937	87.936
Total System Losses [MWh]	70251	69048	68469	68346	69377	68371	68341	68346
Total Grid Losses [MWh]	10925	9983	9410	9405	10614	9378	9404	9394
Total Plant Losses [MWh]	59326	59064	59058	58941	58762	58992	58937	58951
ΔP_{add}^{sys} [MWh]	1910.3	706.5	127.5	5.1	1035.7	29.6	0.0	4.5
ΔP_{add}^{grid} [MWh]	1547.1	605.5	32.4	26.8	1236.1	0.0	25.8	16.1
ΔP_{add}^{plant} [MWh]	563.7	301.5	295.6	178.8	0.0	230.1	174.6	188.9
BEP Efficiency [%]	90.134	90.135	90.309	90.324	90.144	90.300	90.325	90.324
Calculation Speed [steps/s]	12.9	12.7	25.5	32.1	5.6	1.3	0.8	34.2

The constant voltage controllers (C1, C2) have the lowest voltage variation and, consequently, the highest reactive power variation. When the external grid voltage V_{ext} changes, controller (C1) and (C2) force the terminal voltage towards the voltage setpoint by adjusting the reactive power dispatch. The reactive power dispatch is proportional to the voltage difference between V_{ext} and controller voltage setpoint V_{REF} . It is therefore clear from Figure 6 (c) that controllers (C1) and (C2) operate less efficiently compared to the other controllers.

5 Discussion

Figure 4 shows that the active power has the most significant impact on the system efficiency, with Best Efficiency Points (BEPs) residing at around 80 MW for all controllers. However, there is a significant difference between the BEPs on the V_{ext} axis. All controllers except (C1) and (C2) prefer higher V_{ext} , likely because of the lower currents leading to lower ohmic losses.

Controller (C1) and (C2) have, on average, higher reactive power dispatch because of the additional reactive compensation needed for constant voltage control, indicated in Figure 6 (a) and (b). Therefore, the BEPs of these controllers reside where V_{ext} is close to V_{REF} . With similar generator and external voltage values, the reactive power dispatch is close to 0. Lower reactive power flow has higher efficiencies according to Figure 6 (c).

The additional system losses ΔP_{add}^{sys} in Table 4 indicates that the optimal grid controller (C6) has almost the same performance as the system-wide optimal control (C7). This indicates that the grid is most sensitive to the reactive power flow. The optimal plant control (C5) is overall inefficient. An explanation for this is that the reactive power dispatch of around -0.2 pu, which negatively affects the grid efficiency. The saved losses in the plant are less than the additional grid losses at this reactive power dispatch level, which is clear from the efficiency curves in Figure 6 (c).

Both constant voltage controllers (C1 and C2) have the lowest voltage deviation but at the cost of lower AAE compared to the other controllers. The constant V control (C1) has the best performance for limiting the voltage variation and the worst efficiency. The compensated V controller (C2) showed some efficiency improvements, but at the cost of having a higher voltage variation.

Table 4 also shows that the difference between the optimal system controller (C7) and RL controller (C8) is relatively small, indicating that the RL agent learned an approximate optimal control policy. The main difference between (C7) and (C8) is the calculation speed. The RL agent is a factor about 40 times faster than (C7). Such a prediction speed is achieved because the RL agent does not require

an iterative solver to predict Q_g , and predicts Q_g directly based on the system states P_g and V_{ext} .

6 Conclusions

This paper has studied eight different voltage control methods for the best efficiency operation of a grid-connected hydropower plant. The results are promising and show that improved efficiency for the overall system, plant, and grid efficiency is available as more knowledge on the power losses of the system is provided. Moreover, results indicate that overall system efficiency increases for both the optimal controllers (C5) and (C7), and the RL controller (C8) compared to the traditional control methods (C1-C4). Combining a hydropower plant loss model with traditional optimal power flow methods on the grid side improves the accumulated average efficiency (AAE) by reducing total power losses and increasing revenue for plant owners and TSOs. The RL algorithm A2C has shown to be capable of learning an approximate optimal system control policy for maximizing system efficiency. The main difference between (C7) and (C8) is that (C8) has a significantly faster computational time. RL can therefore be considered viable for online operation with time frames of seconds. However, simulations show that voltage variation on the generator bus is higher when utilizing controllers for increased system efficiency compared to voltage setpoint controllers (C1) and (C2). Future work should focus on expanding the power system size. In addition, voltage restrictions as part of the reward/objective functions should be implemented. Moreover, an in-depth dynamic studies on system stability criteria must be implemented in further studies.

Acknowledgment

We gratefully acknowledge the support from The Research Council of Norway and industry partners through Research Council project "ref: 326673" (SysOpt).

References

- Abadie, L. M., Chamorro, J. M., Huclin, S. and van de Ven, D. J. (2020), 'On flexible hydropower and security of supply: Spain beyond 2020', *Energy* **203**, 117869. doi: 10.1016/J.ENERGY.2020.117869.
- Bortoni, E. C., Uemori, M. K. I., Araujo, B. T., Bernardes, J. V., Rocha E., J. J. and Siniscalchi, R. T. (2020), Accurate methodology to obtain efficiency mapping of synchronous machines, in '2020 IEEE Power Energy Society General Meeting (PESGM)', pp. 1–5. doi: 10.1109/PESGM41954.2020.9281622.
- Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J. and Zaremba, W. (2016), 'Openai gym'.
 Buşoniu, L., de Bruin, T., Tolić, D., Kober, J. and Palunko, I. (2018), 'Reinforcement learning for control: Performance,

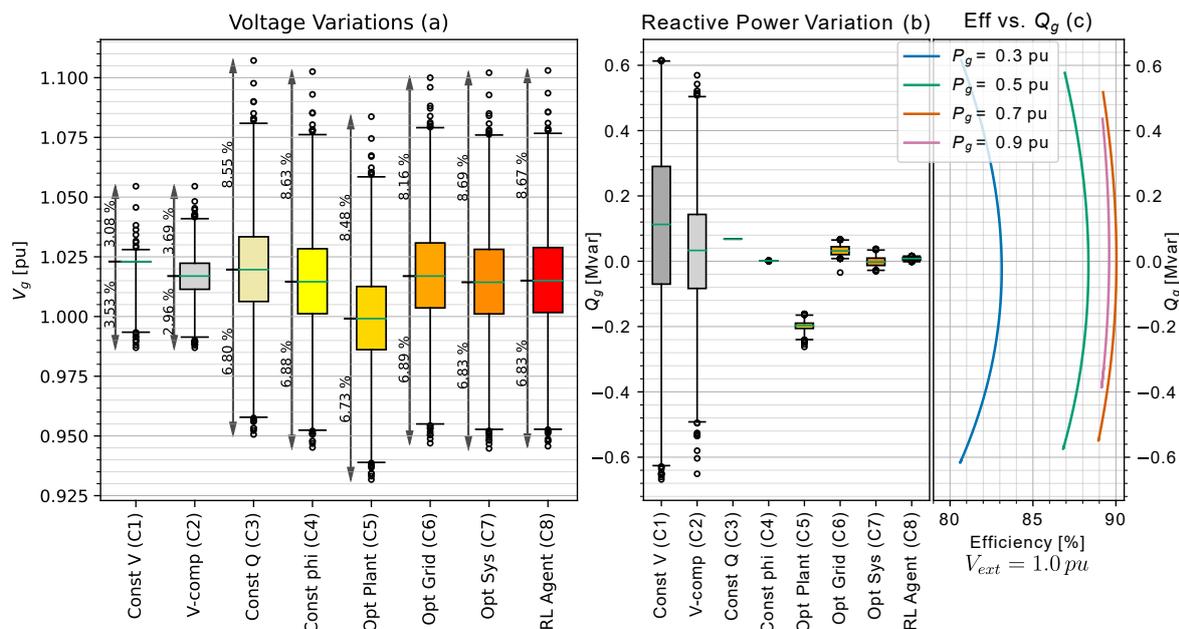


Figure 6: Fig. (a) shows the voltage variation Box & Whiskers plot from simulation of the 8760 data points in the study case, with indications of the % variation from the median value. Fig. (b) shows the reactive power dispatch distributions for the controllers and corresponds to the voltage variation distributions. Fig. (c) illustrates the system efficiency over reactive power values for four different active power levels, namely $P_g = 0.3, 0.5, 0.7,$ and 0.9 pu. The external grid voltage is set to 1.0 pu for the shown efficiency curves.

- stability, and deep approximators', *Annual Reviews in Control* **46**, 8–28. doi: 10.1016/J.ARCONTROL.2018.09.005.
- Chapman, S. J. (2012), *Electric Machinery Fundamentals*, 5 edn, New York: McGraw-Hill.
- Chen, Y., Peng, X., Xu, X. and Wu, H. (2021), 'Deep Reinforcement Learning based Applications in Smart Power Systems', *Journal of Physics: Conference Series* **1881**(2), 361–371. doi: 10.1088/1742-6596/1881/2/022051.
- Duan, J., Shi, D., Diao, R., Li, H., Wang, Z., Zhang, B., Bian, D. and Yi, Z. (2020), 'Deep-Reinforcement-Learning-Based Autonomous Voltage Control for Power Grid Operations', *IEEE Transactions on Power Systems* **35**(1), 814–817. doi: 10.1109/TPWRS.2019.2941134.
- Hasan, F., Kargarian, A. and Mohammadi, A. (2020), *A Survey on Applications of Machine Learning for Optimal Power Flow*. doi: 10.1109/TPEC48276.2020.9042547.
- IEEE Recommended Practice for Excitation System Models for Power System Stability Studies* (2016), *IEEE Std 421.5-2016 (Revision of IEEE Std 421.5-2005)* pp. 1–207. doi: 10.1109/IEEESTD.2016.7553421.
- IEEE Standard for Salient-Pole 50 Hz and 60 Hz Synchronous Generators and Generator/Motors for Hydraulic Turbine Applications Rated 5 MVA and Above* (2006), *IEEE Std C50.12-2005 (Previously designated as ANSI C50.12-1982)* pp. 1–45. doi: 10.1109/IEEESTD.2006.99082.
- Karikezi, Y. C., Nøland, J. K. and Øyvang, T. (2021), 'The Energy Transition's Impact on the Accumulated Average Efficiency of Large Hydrogenerators', *IEEE Transactions on Energy Conversion* pp. 1–10.
- Kundur, P. (1994), *Power system stability and control*, New York: McGraw-Hill.
- Melfald, E. G. (n.d.), 'Opensimhub/sysopt-wp1-hydropowercontrol: 2022-09-sims2022 (2022-09-sims2022)'. doi: 10.5281/zenodo.6597305.
- URL:** <https://doi.org/10.5281/zenodo.6597305>
- Mnih, V. and et al. (2016), 'Asynchronous methods for deep reinforcement learning'. doi: 10.48550/ARXIV.1602.01783.
- Molzahn, D. K., Dörfler, F., Sandberg, H., Low, S. H., Chakrabarti, S., Baldick, R. and Lavaei, J. (2017), 'A survey of distributed optimization and control algorithms for electric power systems', *IEEE Transactions on Smart Grid* **8**(6), 2941–2962. doi: 10.1109/TSG.2017.2720471.
- NVE (2019), 'System operation in the Norwegian power system'.
URL: <https://2021.nve.no/norwegian-energy-regulatory-authority/system-operation-in-the-norwegian-power-system/>
- Pandey, M., Winkler, D., Sharma, R. and Lie, B. (2021), 'Using mpc to balance intermittent wind and solar power with hydro power in microgrids', *Energies* **14**(4). doi: 10.3390/en14040874.
- Qiu, Z., Deconinck, G. and Belmans, R. (2009), A literature survey of optimal power flow problems in the electricity market context, in '2009 IEEE/PES Power Systems Conference and Exposition', pp. 1–6. doi: 10.1109/PSCE.2009.4840099.
- Raffin, A., Hill, A., Gleave, A., Kanervisto, A., Ernestus, M. and Dormann, N. (2021), 'Stable-baselines3: Reliable reinforcement learning implementations', *Journal of Machine Learning Research* **22**(268), 1–8.
URL: <http://jmlr.org/papers/v22/20-1364.html>
- Thurner, L., Scheidler, A., Schäfer, F., Menke, J., Dollichon, J., Meier, F., Meinecke, S. and Braun, M. (2018), 'pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems', *IEEE Transactions on Power Systems* **33**(6), 6510–6521. doi: 10.1109/TPWRS.2018.2829021.
- Virtanen, P. and et al. (2020), 'SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python', *Nature Methods* **17**, 261–272. doi: 10.1038/s41592-019-0686-2.
- Wang, S., Duan, J., Shi, D., Xu, C., Li, H., Diao, R. and Wang, Z. (2020), 'A Data-Driven Multi-Agent Autonomous Voltage Control Framework Using Deep Reinforcement Learning', *IEEE Transactions on Power Systems* **35**(6), 4644–4654. doi: 10.1109/TPWRS.2020.2990179.
- Wang, Y., Wang, S. and Wu, L. (2017), 'Distributed optimization approaches for emerging power systems operation: A review', *Electric Power Systems Research* **144**, 127–135. doi: 10.1016/J.EPSR.2016.11.025.

Process simulation of CO₂ absorption data fitted to performance efficiency at TCM Mongstad

Lars Erik Øi^{a,*}, Njål Sæter^a, Neda Razi^b, Sumudu Karunaratne^a

^a*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway*
lars.oi@usn.no

Abstract

This work's first aim is to fit the simulated results to pilot plant data from Test Centre Mongstad (TCM) for both high and low CO₂ exhaust gas inlet concentrations. The next aim is to evaluate whether fitted parameters for one scenario (a set of experimental data under specified conditions) give reasonable predictions under other conditions. The scenarios at TCM have been simulated in both a rate-based model in Aspen Plus (RateSep) and an equilibrium-based model in Aspen HYSYS. The rate-based model's performance data were fitted by only changing the liquid hold-up factor (multiplied with the hold-up estimated by the Bravo/Rocha/Fair correlations) to obtain the experimental CO₂ removal efficiency. In the equilibrium-based model, a Murphree efficiency (E_M) was specified for 24 or 18 stages (meter of packing) to fit the CO₂ removal efficiency and the temperature profile from performance data. The specified E_M profiles were then used to fit performance data for other scenarios by adjusting only an E_M -factor, multiplying all the E_M values in an E_M profile. The performance (CO₂ removal and temperature profile) was reasonably simulated for each given scenario for all the models. It is shown in this work that the use of the liquid hold-up factor (and not the interfacial area factor) is convenient to fit the rate-based model to the pilot plant data. Using fitted parameters at low CO₂ inlet concentration to predict performance at high concentration needed an adjustment factor (liquid hold-up factor or E_M -factor) to obtain correct CO₂ removal predictions. A liquid hold-up factor of 0.72 and an E_M -factor of 1.72 fitted to performance data for high CO₂ concentration at TCM gave reasonable predictions compared to performance data for high CO₂ concentration from the Esbjerg pilot plant.

Keywords: CO₂, amine, absorption, simulation

1. Introduction

The CO₂ Technology Centre Mongstad (TCM) is the world's largest test facility for CO₂ capture technology. To test CO₂ absorption into amine-based solvents, there is an absorption column with a rectangular cross-section equivalent to a packing diameter of 3 meters, and a packing height up to 24 meters. There have been performed performance tests of CO₂ absorption from flue gas into 30 wt.% monoethanolamine (MEA) in 2013 (Thimsen *et al.*, 2014; Hamborg *et al.*, 2014) and in 2015 (Gjernes *et al.*, 2017; Faramarzi *et al.*, 2017). Figure 1 shows the principle of the amine-based CO₂ absorption and the desorption facility at TCM. In this work, the evaluated parameters are especially the total CO₂ capture rate (in % of incoming CO₂) in the absorption section and the temperature profile from top to bottom.

In earlier work (Sæter, 2016; Øi *et al.*, 2018; Fagerheim, 2019; Øi and Fagerheim, 2020), the equilibrium models (in Aspen Plus and Aspen HYSYS) were fitted to one specific scenario by adjusting the Murphree efficiency (E_M) for each stage, and the rate-based model (in Aspen Plus) was fitted to another scenario by changing the interfacial

area factor. Razi *et al.* (2013) evaluated different correlations in rate-based simulations in Aspen Plus for predicting performance from the Esbjerg pilot plant. Some sets of correlations gave very good predictions, and some sets gave poor predictions of CO₂ removal and temperature profiles.

In rate-based models, it is standard to use the interfacial area factor to fit the model to performance data. In this work, the hold-up factor (multiplied with the hold-up estimated by in-built correlations) was used for the same purpose. One of the estimation methods in Aspen Plus for liquid hold-up is the Bravo/Rocha/Fair model (Rocha, 1992).

This work is based on the Master Thesis of Njål Sæter (2021) and also on earlier work at the University of South-Eastern Norway.

The first aim of this work is to compare results from simulations with performance data from TCM for CO₂ absorption into 30 wt.% MEA using both rate-based and equilibrium models. The second aim, which is specific to this work, is to test whether fitted parameters for one scenario (a set of experimental data at specified conditions) give reasonable predictions under other conditions. Especially it is evaluated what adjustment is necessary for using

fitted parameters for low CO₂ inlet concentrations to predict performance at high CO₂ inlet concentration.

2. Process simulation models for CO₂

Process simulation tools available for CO₂ absorption processes contain models for vapour/liquid equilibrium calculations and efficient solvers. The rate-based tools also include models for chemical, heat transfer and mass transfer kinetics. The commercially available programs Aspen Plus and Aspen HYSYS are used here. Both Aspen Plus and the new Aspen HYSYS acid gas model use an Electrolyte-NRTL equilibrium model based on Austgen *et al.* (1989). Other tools are described in Øi and Fagerheim (2020).

Aspen Plus has a rate-based model to describe the reactive absorption processes. The rate-based model is based on MERSHQ (material balances, energy balances, rate of mass and heat transfer, summation of composition, hydraulic equations for pressure drop and equilibrium) equations which are used to determine molar and energy fluxes transfer across the vapour-liquid interfaces.

The CO₂ capture plant at the University of Texas in Austin was modelled with ASPEN rates, a second-generation rate-based multistage separation unit operation model in ASPEN Plus. The parameters in this file are mainly based on the work of Zhang *et al.* (2009), who fitted Aspen Plus simulations to experimental runs at a CO₂ absorption pilot plant. Different rate-based models have been developed for TCM in the Master Thesis works of Desvignes (2015), Sætre (2016) and Fagerheim (2019).

Equilibrium based models assuming equilibrium at each stage can be extended by using a Murphree efficiency (the ratio of the change in mole fraction from a stage to the next divided by the change assuming equilibrium). An advantage of using Murphree efficiencies compared to rate-based simulations is that it is simpler and fewer parameters need to be specified. In the Master Thesis work of Zhu (2015) and Sætre (2016), a Murphree efficiency for each stage (meter of packing) was estimated for one set (scenario) of TCM data (Hamborg *et al.*, 2014). A good agreement between the measured and simulated temperature profile was obtained using different fitted Murphree efficiencies for each stage. Luo *et al.* (2009) tested Aspen RadFrac, ProTreat, ProMax, Aspen RateSep, CHEMASIM from BASF and CO2SIM from SINTEF/NTNU by comparing with pilot plant data. The result was that all models could fit the CO₂ capture rate, but the temperature and concentration profiles were not well predicted. When comparing Aspen HYSYS and Aspen Plus, Øi (2012) claimed that there were only slight differences between the tested equilibrium models. A rate-based and equilibrium-based model with estimated Murphree efficiencies gave similar results. In the work by Øi *et al.* (2018), different models were compared for 4 scenarios from TCM at low CO₂ input concentrations (3-4 vol-%). 5 scenarios were compared. The results from these comparisons showed that the equilibrium and rate-based models performed equally well in both fitting performance data and in predicting performance at changed conditions.

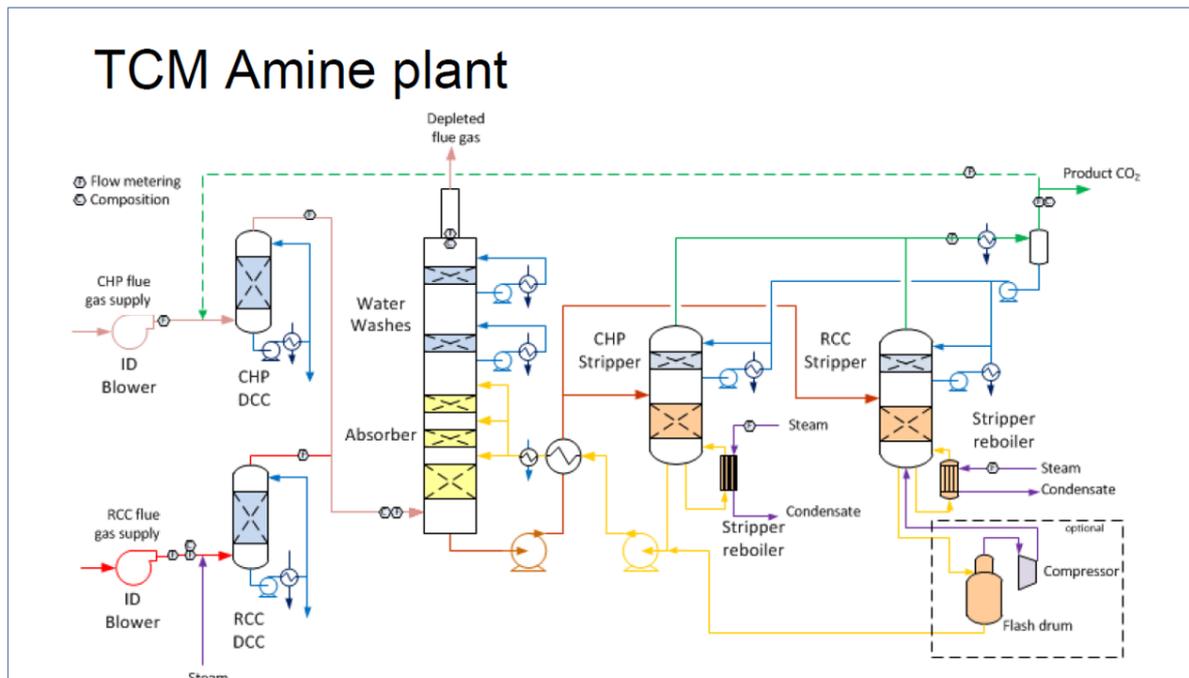


Figure 1: Simplified process diagram of the amine-based CO₂ capture plant at TCM (Hamborg *et al.*, 2014)

3. Data and specifications

3.1. Performance data from TCM

Performance data with low CO₂ inlet concentration (3-4 vol-%) has been taken from 5 sets of conditions (scenarios) at TCM. They are from campaigns in 2013 and 2015 for approximately 30 wt.% MEA in water. 24 meters of packing height (the maximum available) were used in these scenarios.

The data (mainly conditions of the inlet gas stream and the inlet amine stream to the absorption section) for the 5 scenarios are listed in Tab. 1. The 5 scenarios which have been selected in this work are named H14 and 6w from 2013 (Hamborg *et al.*, 2014), 2B5 and Goal1 from 2015 (Gjernes *et al.*, 2017) and F17 from Faramarzi (2017). The names have been used internally at TCM, except for the H14 and F17 scenarios, which are named by first

author's initial and year. The different scenarios cover typical conditions. 17F is at a low liquid to gas conditions, Goal1 and H14 are at low temperature, 2B5 is standard conditions and 6w is at a high liquid to gas ratio.

Performance data with high CO₂ inlet concentration for 6 scenarios are specified in Tab. 2 (Sæter, 2021). The original data are taken from Shah *et al.* (2018). These data are from a test campaign with high CO₂ inlet concentration (13.5 vol-%) from a Residue Fluidized Catalytic Cracker (RFCC) at Mongstad. The data are from a campaign called SRD because the purpose was to evaluate the specific reboiler duty under different conditions. The main differences between the scenarios are mainly due to different liquid to gas ratios.

Table 1: Input data for simulations of TCM conditions with low CO₂ inlet concentrations (3-4 vol-%).

Key inputs	Case				
	17F	Goal1	H14	2B5	6w
Lean amine loading (mole CO ₂ /moleMEA)	0.20	0.20	0.23	0.23	0.25
Lean amine flowrate (kg/hr)	57434	44391	54900	49485	54915
MEA weight% (without CO ₂)	31.0	32.3	30.0	31.6	30.4
Flue gas flowrate (kg/hr)	72389	57157	57300	57193	56788
Flue gas pressure (bara)	1.0313	1.0313	1.0313	1.0313	1.0313
Lean amine pressure (bara)	1.0630	1.0630	1.0630	1.0630	1.0630
Flue gas temperature (°C)	29.8	25.0	25.0	28.2	25.0
Lean amine temperature (°C)	37.0	28.6	36.5	36.8	36.9
CO ₂ removal (%)					
Test result	83.5	90.1	90	87.3	79

Table 2: Key input data and test results for the TCM SRD test cases (high CO₂ inlet concentration).

Key inputs	Case					
	6c	6a	8a	5c	3	4
Lean amine loading (mole CO ₂ /moleMEA)	0.16	0.19	0.199	0.204	0.251	0.273
Lean amine flowrate (kg/hr)	99670	114873	120360	116455	136867	160821
Flue gas flowrate (5mr/hr)	33908	33900	33934	33918	33699	33874
Flue gas flowrate used (for simplicity) (kg/hr)	43500	43500	43500	43500	43250	43500
Mass ratio lean amine / flue gas	2.29	2.64	2.77	2.68	3.16	3.70
Measured CO ₂ removal efficiency (%)	88.3	87.3	87.4	87.3	88.1	85.9

Table 3: Key input data and test results for Campaign 4 test cases with high CO₂ inlet concentration (13.5 vol-%)

Key Inputs	Case			
	1A-1	1C	1D	2B
Lean amine loading (mole CO ₂ /moleMEA)	0.215	0.29	0.318	0.266
Lean amine flowrate (kg/hr)	120100	200500	200600	165600
Flue gas flowrate used (for simplicity) (kg/hr)	43500	43500	43500	43500
Measured CO ₂ removal efficiency (%)	90.1	89.7	78.7	89.4

The temperature was measured at different locations at the same column height for each meter of packing, and the different locations were named A, B, C and D. For the H14 and F17 scenarios, a mean value for

each meter of packing was specified in the temperature profile (Hamborg, 2014; Faramarzi 2017).

Data for Campaign 4 from TCM are given in Tab. 3. The original data are from Fosbøl *et al.* (2019). The campaign had the aim of studying lean vapor compression (LVC), but only standard regeneration cases were used as scenarios in this study.

3.2. Specifications for the Equilibrium Based Simulation Tools

The Acid Gas model, which is the recommended equilibrium model in the last Aspen HYSYS version, was used. This is now based on the Electrolyte-NRTL model (Austgen *et al.*, 1989), while earlier versions of Aspen HYSYS used other models. The work of Fagerheim (2019) used the

Kent Eisenberg (1976) model. When using Aspen Plus, the Electrolyte-NRTL (Non-Random-Two-Liquid) model was used.

Table 4: Murphree efficiency profiles used in this work.

Stage	Zhu	Zhu_M	Zhu_Adjusted	0.1 (18 meter)
1	0.23	0.1805	0.310	0.1
2	0.2192	0.1720	0.296	0.1
3	0.2085	0.1636	0.281	0.1
4	0.1977	0.1551	0.267	0.1
5	0.1869	0.1466	0.252	0.1
6	0.18	0.1412	0.243	0.1
7	0.1762	0.1382	0.238	0.1
8	0.1546	0.1213	0.209	0.1
9	0.1438	0.1128	0.194	0.1
10	0.1331	0.1044	0.180	0.1
11	0.1223	0.0960	0.165	0.1
12	0.1115	0.0875	0.150	0.1
13	0.1007	0.0790	0.136	0.1
14	0.09	0.0706	0.121	0.1
15	0.01	0.0078	0.013	0.1
16	0.01	0.0078	0.013	0.1
17	0.01	0.0078	0.013	0.1
18	0.01	0.0078	0.013	0.1
19	0.01			
20	0.01			
21	0.01			
22	0.01			
23	0.01			
24	0.01			

In the Master Thesis work of Zhu (2015), a Murphree efficiency for each of the 24 stages (meter of packing) was estimated for the TCM data set (Hamborg *et al.*, 2014). The simplest approach for fitting the E_M profile to the temperature profile was a constant E_M for every stage (Zhu, 2015). Fagerheim (2019) fitted several E_M profiles to the performance data. The Zhu profile has $E_M = 0.01$ on the 10 lowest stages. These profiles were specified in both the Aspen HYSYS and Aspen Plus simulation tools. The different E_M profiles used in this work are presented in Tab. 4.

The Zhu profile is from Zhu (2015). The Zhu_M profile is an adjusted Zhu profile developed by Sæter (2021). The Zhu_Adjusted profile is developed to fit both CO₂ capture efficiency and the temperature profile data from the SRD case 4.

3.3. Specifications for the Rate-Based Tool

The specifications in the rate-based Aspen Plus simulation tool at TCM have been developed for several years, and different versions have been used (Desvignes, 2015; Sætre, 2016). Most of these specifications are based on the work by Zhang *et al.* (2009). Detailed documentation of the rate-based

model can be found in the Aspen Plus program documentation. The interfacial area factor was kept constant (as 1.0) in this work. The hold-up factor was varied. The main specifications for the rate based model is given in Tab. 5.

Table 5: Main input to rate-based model (Sæter, 2021).

Common input			
Calculation type	Rate-based		
Fluid properties	RK-ENERTL		
Reaction ID	MEA		
Mass transfer coeff method	BRF-1992 (Bravo-Rocha-Fair)		
Interfacial area method	BRF-1992 (Bravo-Rocha-Fair)		
Holdup method	BRF-1992 (Bravo-Rocha-Fair)		
Number of stages	50		
Film Liquid phase	Discrxn		
No. of discr. points liquid	5		
Film Vapor phase	Film		
Reaction conduction factor	0.9		
Flow model	VPLUG		
Interfacial area factor	1		
Heat transfer coeff method	Chilton and Colburn		
Campaign specific input			
	TCM CHP flue gas	TCM CHP flue gas	Esbjerg cases
Packed height [m]	24	18	17
Packed diameter	3	3	1.1
Packing type	Koch metal 2x	Koch metal 2x	Sulzer Mellapak 2X
Liquid hold up factor	1.0 & tuning	0.72 & tuning	0.72 & tuning
Film discretization ratio	1	2	2

4. Results and discussion

4.1. General Results

The results shown for each model in the scenario figures is the temperature profile. The model parameters (in the case of the rate-based model, the hold-up factor (one parameter)) is adjusted to achieve the specified capture rate. In the case of using an E_M profile, all the E_M values were multiplied with an E_M -factor (one parameter).

4.2. Scenario H14 with equilibrium model

The results from Øi and Fagerheim (2020) are based on the Kent Eisenberg equilibrium model. This work is based on the electrolyte-NRTL model. The temperature profiles for the H14 data in Tab. 2 were calculated in Øi and Fagerheim by fitting the E_M -factor in the simulations to achieve the experimental CO₂ removal. The resulting temperature profiles using the same E_M profiles and E_M -factors are shown in Fig. 2.

The verification shows that the models in this work give approximately the same results as Øi and Fagerheim (2020) using different equilibrium models.

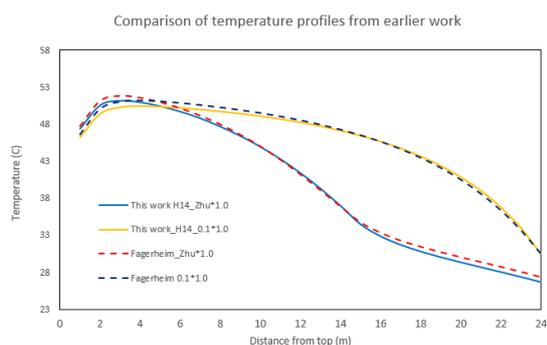


Figure 2: Comparison of measured temperatures for plant data scenario H14 and equilibrium based simulated temperature profiles.

4.3. Scenario H14 with rate-based model

The results in Fig. 3 verifies the simulation of the rate-based model. The interfacial area factor is 1.0, and for this model the hold-up factor is not tuned (equals 1.0). The difference can be explained by Øi and Fagerheim (2020) using the BRF-1985 model and in this work, the BRF-1992 (Bravo *et al*, 1992) is used.

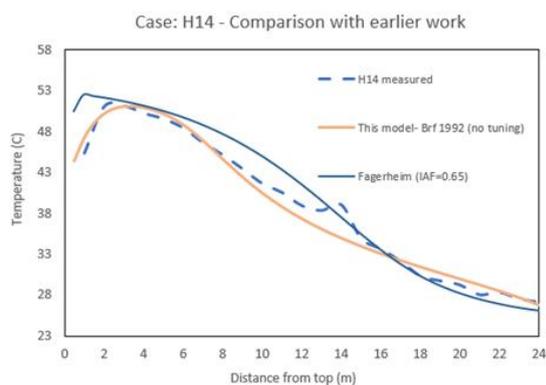


Figure 3: Comparison of measured temperatures for plant data scenario H14 and rate-based simulated temperature profiles.

4.4. Scenario 2B5

The rate-based simulations were fitted to the data from Scenario 2B5. The scenario is for standard conditions at low CO₂ input concentration.

Fig. 4 shows the tuned rate-based model fitted to the CO₂ capture rate for the 2B5 conditions. The interfacial area factor is kept at 1.0, while the hold-up factor is adjusted to 1.6.

When fitting the 5 low CO₂ inlet concentration scenarios, the rate-based adjusted factors (either interfacial area factor or hold-up factor) had to be adjusted more than the E_M-factor for the equilibrium-based models. This was also a general result in Øi and Fagerheim (2020). Especially the H14 and the 6w scenarios are difficult to fit in the rate-based model by adjusting the interfacial area factor.

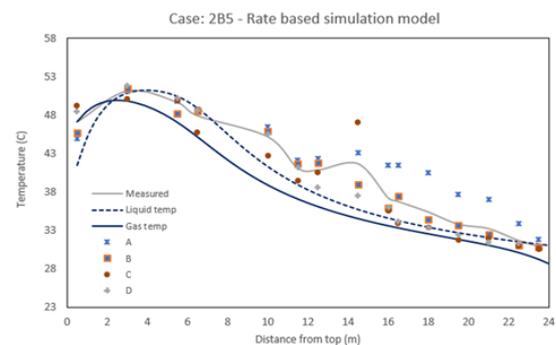


Figure 4: Comparison of measured temperatures for plant data scenario 2B5 and rate-based simulated temperature profiles.

4.5. Scenario SRD Case 4

The 6 scenarios for high CO₂ inlet concentration scenarios specified in Tab. 2 were fitted in the rate-based model by adjusting the hold-up factor and in the equilibrium model by adjusting the E_M-factor. The hold-up factor was adjusted to values between 0.3 and 1.3. The E_M-factor was adjusted to values between 1.33 and 2.05.

All the rate-based scenarios were fitted using 30 wt-% MEA in the input data. Scenario 8a and 3 probably had slightly different MEA concentrations, 28.0 and 30.2 respectively (Sæter, 2021). This resulted in hold-up factors 0.6 and 0.84 with a mean value of 0.72. The value of 0.72 is selected as the hold-up factor for later predictions.

Fig. 5 shows measured and calculated temperatures for Case 4. (CO₂ removal equals test results for all calculations)

The fit is very good when the adjustment factors are used to fit the experimental data. The temperature profiles for rate-based and equilibrium simulations are very similar. In this case, it was tried to adjust the interfacial area factor to fit the CO₂ removal rate, and this was not achievable. Because this was the case also for some of the 5 cases for low CO₂ inlet concentration, adjusting of the hold-up factor was used in this work.

4.6. Scenarios from Campaign 4

The 4 scenarios for high CO₂ inlet concentration specified in Tab. 3 were fitted in the rate-based model by adjusting the hold-up factor and in the equilibrium model by adjusting the E_M-factor. The hold-up factor was adjusted to values between 0.2 and 0.85. The E_M-factor was adjusted to values between 1.72 and 3.4.

For comparison, the highest (Case 1C) and lowest (Case 1A-1) temperature profiles from the rate-based model calculations are included with dotted lines in Fig. 6, together with the results from the HYSYS E_M model. The comparison shows that the two models agree very well for all four cases.

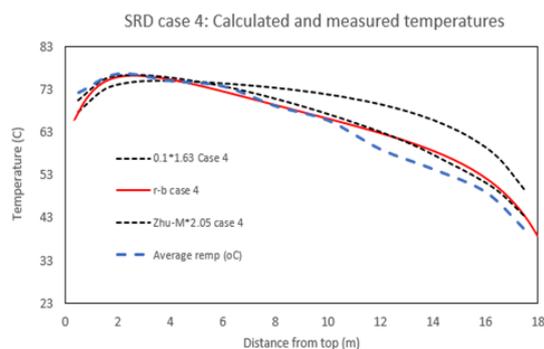


Figure 5: Calculated (equilibrium based and rate-based) and measured temperatures for SRD case 4.

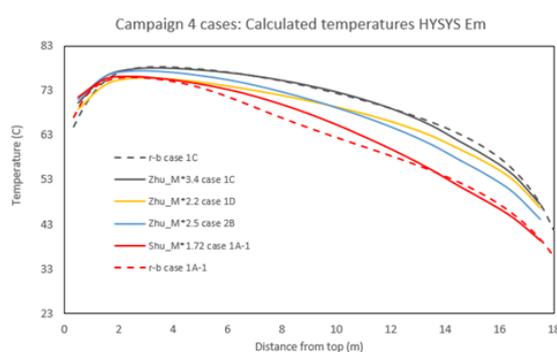


Figure 6: Calculated temperatures for the selected Campaign 4 test cases with the HYSYS E_M model using the Zhu-M profile as a basis. The rate-based model calculations for cases 1C and 1A-1 are included for comparison.

The Zhu E_M profile with an E_M-factor of 1.72 is found to give a good fit to CO₂ removal efficiency as indicated in Fig. 6. The value of 1.72 is selected as the E_M-factor for later predictions in this work.

4.7. Esbjerg cases (4 Cases)

Experimental data were compared to rate-based and equilibrium-based simulations. The purpose of this subsection is to find out whether rate-based and equilibrium-based models fitted to TCM conditions are able to predict CO₂ removal rate and temperatures at Esbjerg conditions.

Table 6: Key input data from Esbjerg test cases and CO₂ removal results.

Key inputs	Case			
	E-1	E-2	E-3	E-4
Lean amine loading (mol CO ₂ / mol MEA)	0.29	0.258	0.222	0.181
Lean amine flowrate (m ³ /hr)	24000	21000	18000	15000
Flue gas flowrate (Nm ³ /hr)	4952	4975	4999	4999
Mass flow ratio lean amine / flue gas	3.78	3.30	2.83	2.36
CO₂ removal (%)				
Test result	88	90	88	87
Rate-based (lhuf= 0.72)	86.7	88.7	88.8	86.7

For the rate-based model with hold-up factor 0.72, the predicted CO₂ removal rates are given in Tab. 6. The CO₂ removal predictions for all 4 cases are very good.

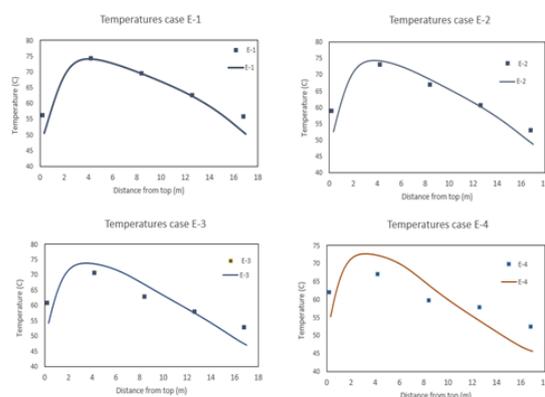


Figure 7: Rate based calculated temperatures compared with measurements for the Esbjerg cases. (Liquid hold-up factor 0.72 for all calculations).

Rate-based simulation results of the temperature profiles are presented in Fig. 7. The predictions of the temperature profiles are very good for case E-1 and E-2 but not so good for E-3 and E-4.

For the equilibrium-based model with the “Adjusted Zhu model”, the predicted CO₂ removal rates were 86.1, 89.5, 91.8 and 93.4 and should be compared with the experimental values in Tab. 6. The predictions for E-1 and E-2 are very good. The predictions are not especially good for E-3 and E-4. Predictions of the temperature profiles for case E-1 and E-2 are given in Fig. 8 and 9. The differences in the top and bottom can be explained by the definitions of simulated temperatures and also the measurement location at top and bottom conditions. As for the rate-based models, the predictions are very good. The temperature profiles from the equilibrium-based models for E-3 and E-4 (not shown here) are very similar to the rate-based temperature profiles.

For the cases E-3 and E-4, the efficiency was overpredicted (or the E_M values were overpredicted). This overprediction can be explained by stating that Case 3 and 4 were from cases with low amine rate and high rich amine concentrations. At these conditions it is expected that the Murphree efficiencies will be reduced. The rate-based model probably has this efficiency decrease as an integrated part of the model.

4.8. General discussion

The results from this and earlier work show that it is possible to fit both CO₂ removal rate and the temperature profile using either a rate-based or an equilibrium-based model. A model based on other conditions can in most cases be fitted to new conditions by adjusting only one parameter. This

can be the interfacial area factor or the liquid hold-up factor in a rate-based model or an E_M -factor (adjusting all the stage efficiencies in an E_M profile) for an equilibrium-based model.

When trying to use a model fitted at a low CO_2 inlet concentration to predict CO_2 removal rate at a high CO_2 inlet concentration, both the rate-based and the equilibrium-based models had to be heavily adjusted.

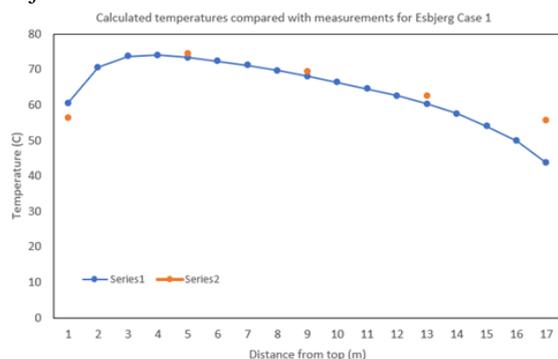


Figure 8: Calculated temperatures compared with measurements for the Esbjerg Case 1 (E_M -factor 1.72).

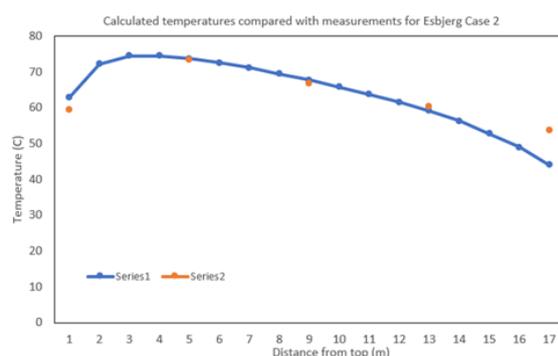


Figure 9. Calculated temperatures compared with measurements for the Esbjerg Case 2 (E_M -factor 1.72).

The rate-based model fitted at TCM conditions for high CO_2 inlet concentration was able to predict the Esbjerg conditions very well. For some cases, also the equilibrium-based model was able to predict the Esbjerg conditions well. When the lean amine flowrate and the lean amine loading changed much, the predictions based on the equilibrium-based model were not good. The dependence of lean amine loading and lean amine flowrate was better predicted by the rate-based model. These dependencies are probably modelled reasonably in the rate-based model and not at all in the equilibrium-based model in this work.

In Øi *et al.* (2018) and in Øi and Fagerheim (2020) it was found that at their conditions (low inlet CO_2 concentration) the equilibrium-based models and rate-based models performed about equally well. This work indicates that for predicting performance at very different conditions, a rate-based model performs better. As stated in Øi and Fagerheim (2020), when the knowledge of the factors used in

rate-based simulations becomes better known, the rate-based models can probably be made more predictive.

5. Conclusions

Performance data at TCM have been simulated in both a rate-based model in Aspen Plus (RateSep) and an equilibrium-based model in Aspen HYSYS. The rate-based model's performance data were fitted by only changing the liquid hold-up factor to obtain the experimental CO_2 removal efficiency. In the equilibrium-based model, a Murphree efficiency was specified to fit the CO_2 removal efficiency and the temperature profile. The specified E_M profiles were then used to fit performance data for other scenarios by adjusting only an E_M -factor. The performance (CO_2 removal and temperature profile) was reasonably simulated for each given scenario for all the models. It is shown in this work that the use of the liquid hold-up factor (and not the interfacial area factor) is convenient to fit the rate-based model to performance data. Using fitted parameters at low concentration to predict performance at high inlet CO_2 -concentration conditions needed an adjustment factor (liquid hold-up factor or E_M -factor) to obtain correct CO_2 removal predictions. A liquid hold-up factor of 0.72 and an E_M -factor of 1.72 fitted to performance data for high CO_2 concentration at TCM gave reasonable predictions compared to performance data for high CO_2 concentration from the Esbjerg pilot plant.

This work indicate that it is not expected that models fitted to performance data can be used to predict performance at very different conditions. It is however showed that adjusting only one parameter in either a rate-based or an equilibrium based model can give a good fit.

References

- Austgen, D. M. *et al.* (1989) 'Model of Vapor-Liquid Equilibria for Aqueous Acid Gas-Alkanolamine Systems Using the Electrolyte-NRTL Equation', *Industrial & Engineering Chemistry Research*, 28, pp.1060-73.
- Desvignes, C. (2015) *Simulation of post-combustion CO₂ capture process with amines at CO₂ Technology Centre Mongstad*. Master's Thesis, CPE Lyon, France.
- Fagerheim, S. (2019) *Process simulation of CO₂ absorption at TCM Mongstad*. Master's Thesis. University of South-Eastern Norway.
- Faramarzi, L. *et al.* (2017) 'Results from MEA testing at the CO₂ Technology Centre Mongstad: Verification of baseline results in 2015', *Energy Procedia*, 114, pp. 1128-1145.
- Fosbøl, P. L. *et al.* (2019) 'Results of the fourth Technology Centre Mongstad campaign: LVC testing', *International Journal of Greenhouse Gas Control*, 89, pp. 52-64.
- Gjernes, E. *et al.* (2017) 'Results from 30 wt% MEA performance testing at the CO₂ Technology Centre Mongstad', *Energy Procedia*, 114, pp. 1146-1157.
- Hamborg, E. S. *et al.* (2014) 'Results from MEA testing at the CO₂ Technology Centre Mongstad. Part II: Verification of baseline results', *Energy Procedia*, 63, pp. 5994-6011.
- R. L. Kent, R. L. and Eisenberg, B. (1976) 'Better data for Amine Treating', *Hydrocarbon Processing*, 55(2), pp. 87-90.

- Luo, X. *et al.* (2009) 'Comparison and validation of simulation codes against sixteen sets of data from four different pilot plants', *Energy Procedia*, 1, pp. 1249-56.
- Razi, N. *et al.* (1989) 'Validation of mass transfer correlations for CO₂ absorption with MEA using pilot data', *International Journal of Greenhouse Gas Control*, 19, pp. 478-491.
- Rocha, J.A. *et al.* (1992) 'Distillation Columns Containing Structured Packings: A comprehensive Model for their Performance. 1. Hydraulic Models', *Industrial & Engineering Chemistry Research*, 32, pp. 641-651.
- Shah, M. I. *et al.* (2018) 'CO₂ capture from RFCC flue gas with 30 wt% MEA at Technology Centre Mongstad, process optimization and performance comparison', *GHGT-14*.
- Sæter, N. (2021) *Process simulation of CO₂ absorption data fitted to performance efficiency at TCM Mongstad*. Master's Thesis. University of South-Eastern Norway.
- Sætre, K. A. (2016) *Evaluation of process simulation tools at TCM*. Master's Thesis. University College of South-Eastern Norway.
- Thimsen, D. *et al.* (2014) 'Results from MEA testing at the CO₂ Technology Centre Mongstad, Part I: Post-Combustion CO₂ capture testing methodology', *Energy Procedia*, 63, pp. 5938-5958.
- Zhang, Y. *et al.* (2009) 'Rate-based Process Modelling Study of CO₂ Capture with Aqueous Monoethanolamine Solution', *Industrial & Engineering Chemistry Research*, 48, pp. 9233-46.
- Zhu, Y. (2015) *Simulation of CO₂ capture at Mongstad using Aspen HYSYS*. Master's Thesis. Telemark University College, Norway.
- Øi, L. E. (2012) 'Comparison of Aspen HYSYS and Aspen Plus simulation of CO₂ absorption into MEA from atmospheric gas', *Energy Procedia*, 23, pp. 360-369.
- Øi, L. E. (2018) 'Comparison of Simulation Tools to Fit and Predict Performance data of CO₂ Absorption into Monoethanol Amine', *Linköping Electronic Conference Proceedings SIMS 59*, pp. 230-235. doi: <http://dx.doi.org/10.3384/ecp18153230>
- Øi, L. E. and Fagerheim, S. (2020) 'Simulation of CO₂ Absorption at TCM Mongstad for Performance Data Fitting and Prediction', *Linköping Electronic Conference Proceedings SIMS 61*, pp. 332-337. doi: <https://doi.org/10.3384/ecp20176332>

Reducing the CO₂ emissions of a Gas Cracker by Reforming Fuel Gas

Emmy Kristine Rustad, Klaus-Joachim Jens, Lars Erik Øi*

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway

lars.oi@usn.no

Abstract

Ethylene is the most produced organic substance in the world and is a pillar in the chemical process industry. The production pathway is energy intensive and has a corresponding high carbon footprint. This work explores, simulates, and presents the possibility of reforming the fuel gas of a steam cracking furnace. Currently, methane (CH₄) and hydrogen (H₂) from the cracking furnace outlet makes up the fuel source. If H₂ produced by reformation of the current fuel can cover the energy demand of the steam cracking furnace, it can drastically reduce the CO₂ emissions.

Keywords: Decarbonizing, ethylene, steam cracking, reforming, fuel substitution, hydrogen production, ATR, Aspen HYSYS

1. Introduction

Cracking of hydrocarbons is the dominating ethylene production method. The highly endothermic reaction requires high temperatures to produce the desired products. Combustion of fossil fuel to produce the heat adds to the emissions and energy demand, resulting in 1-2 tons of CO₂ for every ton of produced ethylene. The combustion of fossil fuels to supply the steam cracking process, resulted in an emission of 200 million tons of CO₂ in 2000 and 300 million tons CO₂ in 2019 (Ren *et al.*, 2008; Amghizar *et al.*, 2020). With the Paris agreement from 2015, and the national goals of Norway towards a greener society, the exploration of different CO₂ reduction methods has become necessary.

Decarbonizing and/or reduction of the energy consumption in the industrial sector is a necessity to reach the climate goals of Norway. The possibilities are numerous in the refinery sector and can be divided into seven categories, namely fuel substitution, feedstock substitution, process intensification and optimization, recycling, product solution, energy recovery and Carbon Capture and Storage (CCS) or re-use (Negri and Lighthart, 2021). Industrial GreenTech (IGT) has a vision of making the industrial region in Grenland climate neutral by 2040 and have mapped out the current emission status and suggested some reduction possibilities. One of the suggested methods for evaluation is replacing the fuel to the crackers at INEOS. If this fuel gas is reformed to pure hydrogen, it can potentially reduce the current emissions of CO₂ (Aas *et al.*, 2020).

A steam cracker with ethane as feedstock, such as INEOS, has a product stream consisting mainly of ethylene, unconverted ethane, hydrogen, methane and some amount ethyne, propane, propene, propadiene, butane, butene, butadiene, pygas, and fuel oil (Ullmann's Encyclopedia of Industrial Chemistry, 1988). INEOS has ethylene as their most important product and the separated methane and hydrogen is burned in the cracker as fuel. The outlet of the combusted fuel is primarily H₂O and CO₂ and is usually utilized for heating or steam export before being emitted to the atmosphere.

When producing hydrogen from hydrocarbon sources, the production pathway is most commonly divided into 3 sections. First a section to produce syngas, short for synthesis gas, which in this context refers to a mixture of H₂ and CO. The syngas can originate from both natural gas, heavier hydrocarbons or coal. Depending on application and the following process, the syngas is sent for processing to second section. This processing can be water-gas-shift (WGS) reaction, CO removal, or other treatments to obtain the desired composition and/or CO/H₂ ratios. The third section relates to reach the specifications of the H₂, regarding purity and content of different contaminations (Moulijn *et al.*, 2013).

Mature methods for H₂ production originating from hydrocarbons is Steam Methane Reforming (SMR), Partial Oxidation (POX), Catalytic Partial Oxidation (CPOX) and Auto Thermal Reforming (ATR). Among more advanced, novel and/or promising methods for H₂ production originating from hydrocarbons is methane pyrolysis, integrated membrane reactors, integrated sorption-enhanced

systems, chemical looping variations and electric reforming (Wismann *et al.*, 2019; Basile *et al.*, 2015; Voldsund *et al.*, 2016).

2. Process Description

The reforming method utilized in this work is ATR, which is a highly developed and well-tested technology. ATR utilize O₂ and the production of the O₂ contributes to a higher production cost of the hydrogen. Even so, it is regarded as a more attractive option due to the planned construction of water electrolysis at Herøya Industry Park, close to the INEOS production plant (INOVYN, 2021; Krohn-Fagervoll, 2020). This opens the opportunity to purchase O₂ at a lower cost.

ATR is a combination of SMR and POX, where the reactor consists of a burner, a combustion section, and a catalyst bed section. The system is simulated in Aspen HYSYS v12 with various configurations, and partly maximized for hydrogen production. An ATR system along with the primary occurring reactions is illustrated in Figure 1. The methane is fed into the top section along with steam and oxygen. The reaction between CH₄ and O₂ is exothermic. The surplus heat is utilized in the endothermic section in the catalytic bed. This combination of reactions is making the reactor self-supplied with energy (Nielsen and Christiansen, 2011).

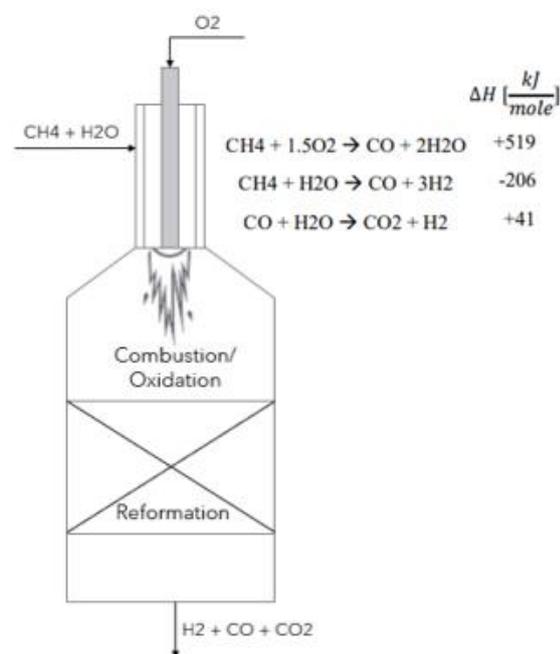


Figure 1: Illustration of an ATR based on a figure from Rostrup-Nielsen and Christiansen(2011). From Rustad (2021)

Compared to a fired furnace (SMR) the ATR is a system with reduced size and complexity. With no external heat supply required, the fuel cost to the reformer will disappear and lead to a reduction in the

CO₂ emission. Another advantage is that the ATR needs less water than an SMR. This is because the high temperature in an ATR leads to a high methane conversion. Additionally, oxygen helps with the prohibition of soot formation (Baltrusaitis and Luyben, 2015). There is a balance between the inlet oxygen flow rate and the temperature in the ATR, where more oxygen will lead to higher temperature. However, the temperature is limited by material and construction cost.

3. Simulations

The specifications and input to the simulation is based on several previous studies and literature (Moulijn *et al.*, 2013; Nielsen and Christiansen, 2011; Jakobsen, 2016; Soltani *et al.*, 2014; Chen *et al.*, 2010). Additionally, some input has been based on assumptions and simplifications.

All the simulations use the Peng Robinson equation of state. The pressure-drop over every heat exchanger is set to 10 kPa and there is assumed not to be a pressure drop in the reactors. There is also assumed no heat loss in the system and the adiabatic efficiency of the compressor is 75 %. All the reactors are simulated as Gibbs reactors (minimizing the Gibbs free energy). There is also assumed to be no build-up in the system. The initial separation step is suggested to be a membrane that separates out 95 mol% of the inlet hydrogen prior to the reforming system to a purity of 100 %. This is simulated as a component splitter. The component splitter is also used for simulating the CO₂ separation and the H₂ purification. The component splitters are merely to exemplify the possibility of different separation steps.

3.1. Burner comparison

The main goal of this project is to investigate if there are enough energy to cover the requirement of a steam cracker with reformed fuel. The natural starting point is to determine how much hydrogen is the minimum to keep the same production rate of ethylene. The current energy supply is combustion of the methane and the hydrogen that are products of the cracker. The amount of CH₄ and H₂ that is burned is the total amount of these components exiting the cracker. A typical amount is 5.5 wt% CH₄ and 4.0 wt% H₂ in the exit gas of the cracking furnace (Mathisen, 2021). This is in the same range as a standard steam cracker for ethylene production with ethane as feed (Ullmann, 1988) and is therefore assumed to be comparable with the actual composition at INEOS. The comparison is illustrated in Figure 2.

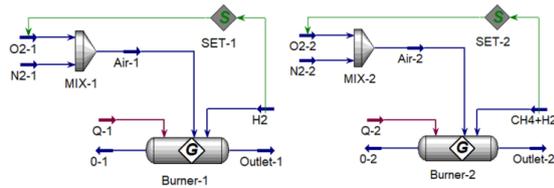


Figure 2: Flowsheet from Aspen HYSYS simulation of burners. Left is Burner 1 and right is Burner 2. From Rustad (2021)

An inlet flow of 150 ton ethane per hour will result in a flow of 514.2 kmole/h CH₄ and 2976.3 kmole/h H₂ on a molar basis. To simplify and to make the comparison on the same terms is two burners simulated in Aspen HYSYS. Burner 1 has pure hydrogen (100 mole%) as fuel and Burner 2 has a mix of hydrogen (14.73 mole%) and methane (85.27 mole%) as fuel. Both are combusted in air with a flow rate that gives 100 % conversion of the fuel, and both burners have the same inlet and outlet pressure and temperature. Burner 1 has a flow rate of 423 kmole/h, corresponding to the flow of hydrogen that enters the burner after reformation in Case 0 (base case, presented in 3.2). Burner 2 has a flow rate of 3491 kmole/h, corresponding to the flow of CH₄ and H₂ that exits the steam cracker (and is the inflow to the reforming process). Burner 1 and Burner 2 has a heat flow of -6.935×10^8 kJ/h and

7.478×10^8 kJ/h, respectively. The negative sign indicating that heat is going out.

3.2. Case 0 Base case

A snapshot of the base case simulation from Aspen HYSYS is presented in Fig. 3. The feed to the system is “Inlet flow, CH₄ and H₂” with a flow rate of 514.2 kmole/h CH₄ and 2976.3 kmole/h H₂. This is fed into a membrane where 95 mol% of the H₂ is separated out and sent directly to the burner. The retentate side of the membrane is mixed with steam with a ST/C ratio of 1.62 and compressed to 2300 kPa. The compressed flow is heated to 650 °C before entering the ATR. O₂ is added to the ATR with a O₂/C ratio of 0.6, and the occurring reactions are highly exothermic, eliminating the need of external energy. The ATR outlet temperature is 1050 °C and the stream is partly utilized to heat the inlet flow to the ATR. The outlet of the ATR is cooled before being fed to two WGS reactors. One high temperature water-gas-shift (HT-WGS) and one low temperature water-gas-shift (LT-WGS) with the inlet temperature at 300 °C and 175 °C, respectively. The WGS reaction is slightly exothermic, and the process stream must be cooled both before and after the reactors.

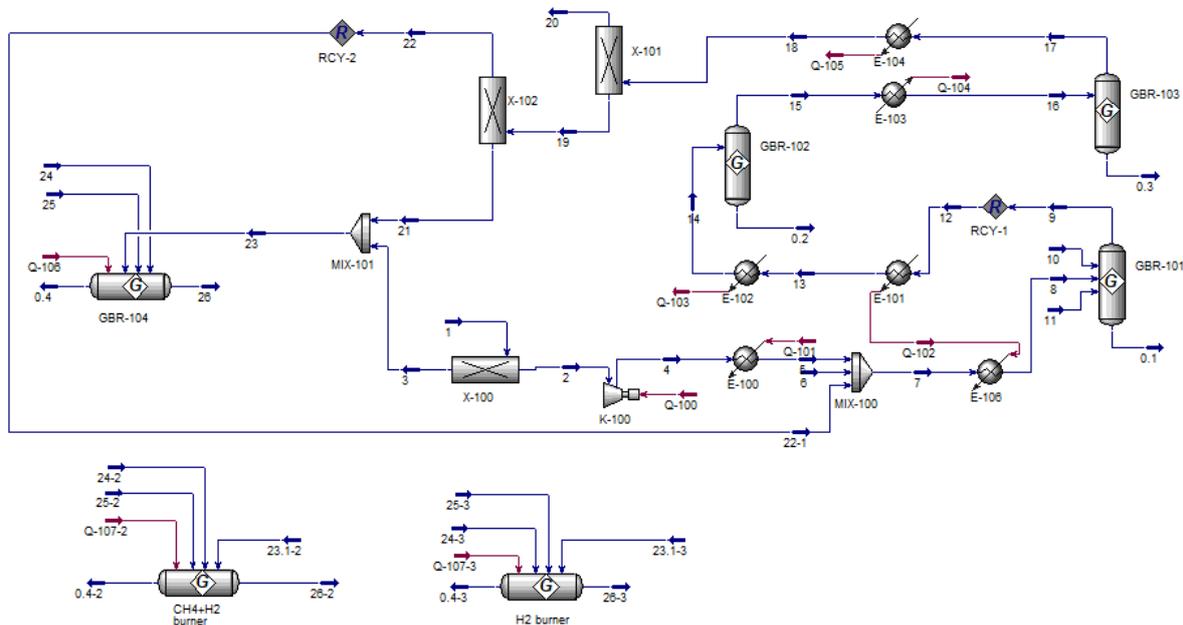


Figure 3: Flowsheet for Base Case. From Rustad (2021)

After the WGS reactors there are two separation processes simulated as component splitters. The first separates out 90 mol% of the CO₂ to a purity of 99.55 mol% and the second one separates out 90 mol% of the H₂ to a purity of 96.8 mol%. The two flows of hydrogen are sent to a burner. The combustion receives stoichiometric amount of air, and the outlet stream of the burner is mainly water

and nitrogen. The energy released from the combustion process (“Burner” in Figure 3) is representing the energy supplied to the steam cracking furnace

3.3. Case 1 Recycle stream

Case 1 has all the equal input parameters as Case 0 with one exception, being the gas exiting the H₂

purification unit is recycled back and enters along with the steam prior to compression. The stream consists of approximately 60 mol% H₂O, 27 mol% H₂, 1 mol% CO₂ and minor amounts of CO and unconverted CH₄. The flow amounts to 624.4 kmole/h. A snapshot of the simulation is presented in Figure 4. Recycling of this stream leads to a few observations. One being the reduced flow of inlet water, from 835 to 450.6 kmole/h, while remaining the ST/C ratio of 1.62. This is due to the content of water in the recycle. This will also result in a lower temperature at the ATR outlet which generally relates to lower conversion of methane. This can be seen from the increased flow of methane, from approximately 3 to 12 kmole/h. This can be compensated by increasing the inlet temperature of the inlet flow(s) or by increasing the flow of oxygen.

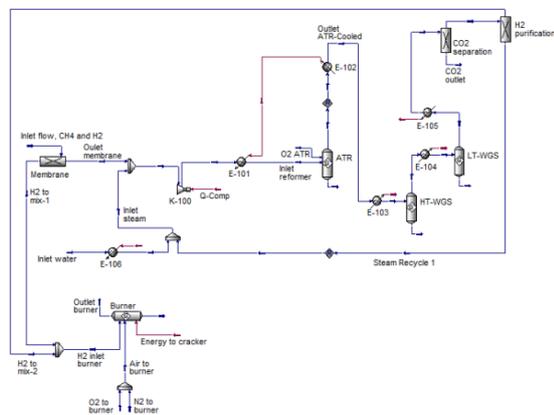


Figure 4: Case 1 Implementing a recycle. From Rustad (2021)

An optimization of the process is not completed at this stage for this case. The process flow in the system from the compressor and for all following the units is increased, and the result is a need for equipment and vessels with higher capacity and therefore higher cost. However, valuable H₂ will be fed back into the system causing the overall hydrogen production to increase.

3.4. Case 2 Additional inlet stream

Neither Case 0 nor Case 1 fulfilled the required amount of hydrogen to satisfy the desired energy supply. There are multiple approaches to increase the amount of hydrogen to the burner. The chosen approach was to have an additional inflow of ethane to the reforming system. The ethane is assumed to be available on site or easy to purchase since this is the raw material to the cracker. More specifically, the ethane is added between the membrane and the compressor, together with the recycle stream. This can be seen from Figure 5. The ethane flow was adjusted until the desired flow of hydrogen to the burner was reached. Note that this both increases the steam and oxygen consumption, as discussed in the

following chapters. If all ratios, temperatures, and pressures from the base case are kept constant, the inflow of ethane is adjusted to approximately 40 kmole/h.

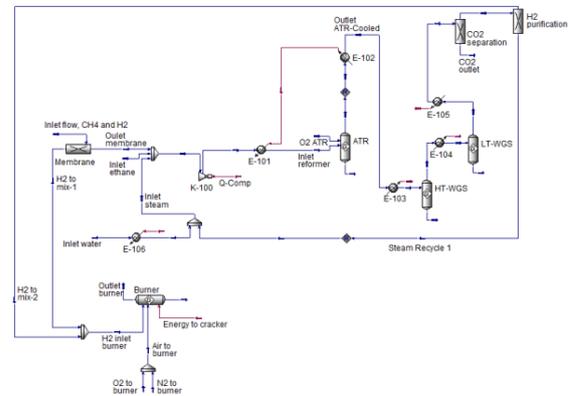


Figure 5: Case 2 Adding ethane to the process. From Rustad (2021)

3.5. Case 3 Adding CO₂

The main advantage of the ATR process is its self-supply of thermal energy. However, there are some challenges related to this reforming method. The temperature in the reactor may exceed the limits of the material and/or lead to total combustion. In the previous cases (Case 0, 1, and 2) this temperature was altered by selecting an appropriate ST/C and O₂/C ratio. More steam or less oxygen leads to a lower temperature and vice versa. The optimum is a tradeoff between conversion, consumption and cost, but that is not pursued here. Reforming by CO₂ is referred to as dry reforming and has received some attention in literature, both encouraging and constructive. It is regarded as an unfavorable option compared to other methods available (Oyama *et al.*, 2012). It is in this work chosen to add a CO₂ flow as an example.

3.6. Case 4 Replacing HT-and LT WGS with MT-WGS

Replacing two reactors of high and low temperature with one medium temperature reactor has obvious advantages, considering the reduction in number of units. Regarding hydrogen production, an MT-WGS reactor is an attractive solution, because it potentially reduces the required steam to the process. It is more challenging to develop a catalyst which functions adequately in the whole interval, being active in the lower region and stable in the higher region. However, the advances in catalyst technology have made the MT-WGS the preferred solution when producing hydrogen. (Moulijn *et al.*, 2013)

The MT-WGS reactor is illustrated in Figure 6 and has an inlet temperature of 225 °C. The outlet temperature increases to 366 °C. This is an increase

of 141 °C compared with the increase in HT and LT in Case 0 of 121.8 and 45.3 °C, respectively.

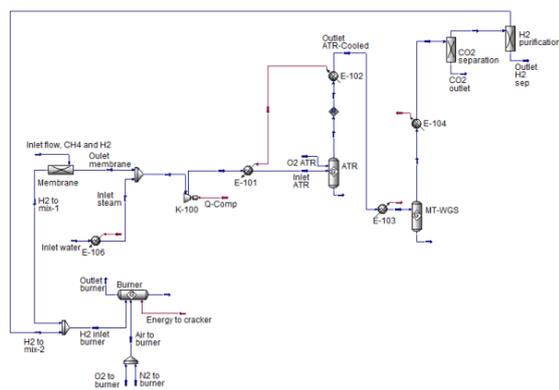


Figure 6: Case 4 Replacing HT- and LT WGS with MT-WGS. From Rustad (2021)

4. Comparisons

Using Case studies in Aspen HYSYS is a method that can be used to analyze the effect of a parameter variation in a process. This is done by choosing an independent variable which is varied within a chosen interval, and with a chosen step size. The dependent variables of interest are selected and there is only one independent variable that varies at the time. The input and results from Case studies of Case 2 are presented.

The ST/C ratio was varied from 0.5 to 5.0, corresponding to an inlet flow of water in the range of 300 to 3000 kmole/h. The result was that increased ratio led to a decrease in the outlet temperature of the ATR. Not surprisingly did it also lead to an increase in the energy stream to the heat exchanger prior to the ATR (E-102), because of a higher flow. The outlet molar composition flow of H₂, H₂O and CO₂ increased, while for CO it decreased. The flow of CH₄ and C₂H₆ only changed slightly, while the O₂ remained at zero.

The O₂/C ratio was varied from 0.2 to 1.0, corresponding to 120 to 600 kmole/h of O₂. This ratio should be one of the last parameters to be optimized due to the fact that it depends on the pressure, temperature and the ST/C ratio. (de Souza, 2015) Too much oxygen could also lead to total combustion which will result in a lower hydrogen production and temperature above material restrictions. The increase in the O₂/C ratio gave an increase in the outlet temperature of the ATR (725 to 1936 °C). The energy stream to E-102 remained constant because the identical flow was risen to the identical temperature throughout the case study. The outlet molar composition changes after all the CH₄ is consumed. This happened at an O₂ flow of 375 kmole/h (which is a O₂/C ratio of 0.635). Until that point is there an increase in CO and a decrease in CO₂. Hydrogen reaches its highest fraction when the flow of O₂ is at 315 kmole/h (which is a O₂/C ratio

of 0.53). The composition of H₂O decreases until it turns before the flow of O₂ reached 300 kmole/h.

The case study for the pressure in the ATR was conducted by reusing the ATR in the Case 2 and complete this study without the upstream and downstream part. The case study involves a pressure from 2000 to 10000 kPa. The result was that an increase in pressure gave an increase in temperature. The outlet molar composition has less H₂ and CO and more H₂O. CH₄ increased slightly and CO₂ decreased slightly.

The case study for the temperature of oxygen to the ATR was in the interval of 20 to 750 °C. The increase in temperature of O₂ also increased the outlet temperature of the ATR (from 1050 to 1161 °C). This caused the outlet molar composition to have more CO and less CO₂, a slight increase in the H₂O content and decrease in H₂, while CH₄ did not change substantially, and the oxygen and ethane remained unchanged.

The case study for the inlet temperature to the HT-WGS was conducted between 250 and 450 °C. Not surprisingly did the outlet temperature increase along with the inlet temperature. It also showed that outlet CO and H₂O molar composition increased, and CO₂ and H₂ decreased. CH₄ (inert), C₂H₆ (inert) and O₂ remained constant. The molar H₂/CO ratio in the outlet got lower when the temperature got higher with 13.3 at 300 °C and 10.7 at 350 °C.

The same trend lines can be observed when increasing the inlet temperature for the LT-WGS in the range from 150 to 250 °C. The H₂/CO ratio decreased from 14.1 at 175 °C to 95.2 at 200 °C.

The case study for varying the flow of oxygen (and nitrogen) was varying from 1000 to 3000 kmole/h. Heat is released and is representing the energy supply to the ethane cracking furnace, meaning that the highest absolute value of the heat flow is desired. This value is at 2201 kmole/h of oxygen, in the case study this is exactly stoichiometrically. Before the amount reaches the stoichiometric amount there is a descending flow of H₂ and CO. When the flow surpasses the stoichiometric amount of O₂, the H₂ and CO are stable at zero, while the molar composition of CO₂ and H₂O decreases and oxygen increases.

The case study varying the temperature into the ATR varied in the range of 600 to 800 °C. Increased temperature leads naturally to a higher heat flow to the heat exchanger before the reformer. Keeping all other variables constant, it is observed that increased temperature leads to increased CO and H₂O composition in the ATR outlet, and a decrease for H₂ and CO₂. This might appear unfavorable but the increase in the CO results in higher possible conversion in the following WGS reactors. The overall molar flow of the hydrogen increases.

Tab.1 presents a summary of the effects on outlet flow of CO₂ and the overall H₂ production. The

base case simulation had a molar flow of 524 kmole/h of CO₂ and 4450 kmole/h of H₂. The numbers presented in the table are in a range, representing the lowest and highest values achieved within the limits used in the corresponding case study and again is the recycle not attached. The pressure in the ATR was analyzed by reusing the

reactor in Aspen HYSYS, eliminating the possibility to see the CO₂ and H₂ molar flow rates directly. The interval given in the table is therefore the minimum and maximum found when manually testing various pressures in the interval of 100 to 5000 kPa.

Table 1: Summary of CO₂ outlet flow and H₂ production for the various case studies. From Rustad (2021)

Variables	CO ₂ outlet flow [kmole/h]	H ₂ production [kmole/h]	Note
ST/C-ratio	361.3-535.7	4258-4614	Higher ratio results in higher CO ₂ and H ₂ molar flows
O ₂ /C-ratio	235.1-534.1	3747-4464	Higher ratios result in higher CO ₂ . H ₂ production peaks when O ₂ /C is 0.57
Inlet temp. HT- and LT-WGS	HT: 526.0-516.3 LT: 504.5-528.0	HT: 4444-4451 LT: 4434-4453	Increasing temperature results in degreasing flows of CO ₂ and H ₂
Inlet temp. MT-WGS	386.9-419.9	4227-4252	Increasing temperature results in degreasing flows of CO ₂ and H ₂
Pressure ATR	522.7-525.8	4432-4457	Increasing pressure results in degreasing flow of CO ₂ and H ₂
Inlet temp. ATR	523.0-525.0	4445-4455	Increasing temperature results in increase in flow of CO ₂ and H ₂
Inlet temp. of O ₂ to ATR	524.0-524.9	4450-4454	Increasing temperature results in slightly increase in flow of CO ₂ and H ₂
Flow of CO ₂ to ATR	420.4-581.4	4182-4222	Increasing flow results in increase in flow of CO ₂ and decrease in flow of H ₂

As can be seen from Tab.1 hydrogen production is favored at high ST/C ratio, O₂/C ratio of 0.57, low inlet temperatures to the WGS reactors, low pressure, low flow of CO₂ and high temperature in the ATR. A 'perfect' system has minimum amount of CO₂ produced and maximum amount of H₂ produced.

4.1. Evaluation and discussion

The most important output of each simulation is the energy requirement for the ethane steam cracking furnace. Case 2 has enough energy to meet this requirement. However, it is highly advantageous that the process operates economically, in a sense that consumption of ethane and oxygen is at minimum, as this is purchased. The increase in flow throughout the system will also be a determining factor when sizing the equipment, which will affect the investment cost. So, there will be a tradeoff

between efficiency, amount produced, operational cost and capital cost.

The selection of optimal operation conditions in an overall sense is not possible to complete properly when the cost is not a part of the equation and only one parameter is evaluated at the time. Nevertheless, there are some parameters that are more advantageous to promote than others. CH₄ and ethane conversion, H₂ and CO₂ production, as well as the possibility of steam export. This system will by no means be completely optimized, neither for production or conversion, nor for cost. This is a feasibility study and an investigation of the possibility to implement a reforming process for a cracking production plant, such as INEOS. The focus will therefore be on high hydrogen production. But there are some advantages and disadvantages related to all the cases described and the main ones are summarized in Table 2.

From Tab.1 and Tab.2 it is observable that some choices for further work are more appealing than others. Case 2 and 4 are the ones that overall comes out on the winning side of the selection. Both may accomplish a satisfactory production of hydrogen, both consists of well tested technologies that are ready for implementation, and both have potential for further optimization along with their possible

steam export. The hydrogen production is likely to increase when optimal conditions are chosen.

To summarize what favors conversion of hydrocarbons and production of hydrogen, the following points as seen from the simulations, case studies, and literature research. The trend lines are in agreement with previous studies (Nielsen and Christiansen, 2011; Baltrusaitis and Luyben, 2015; Rashid, 2019; Jakobsen, 2016; Souza *et al.*, 2015).

Table 2: Main advantages and disadvantages for the simulated cases (Case 0 - 4). From Rustad (2021)

Cases	Main advantages	Main disadvantages
Case 0	Lowest complexity	Does not produce enough H2 Low utilization of resources
Case 1	Increased utilization of resources Higher H2 production (less lost)	Does not produce enough H2
Case 2	Produce enough hydrogen	Higher flow throughout the system
Case 3	Produce enough hydrogen Easier temperature control	Lower conversion
Case 4	One less reactor and heat exchanger	Lower H2 production (but can be compensated with ethane inflow) Higher flow throughout the system

Low pressure. It should be as low as possible and is generally limited by the downstream purification steps, the volumetric flow through the system (which affect the dimensions), and carbon formation.

High temperatures, both in the inlet and outlet of the ATR are generally limited by the material and the risk of total combustion.

High ST/C ratio. The steam contributes to higher hydrogen production, as well as reduction of coke formation. The downside is increased energy requirement (because the temperature drops with an increase of steam) and the flow through the system. O₂/C ratio gave varying results and highly affects the outlet temperature. Keeping a ST/C ratio of 1.62 was the highest fraction of H₂ in the outlet when the O₂/C ratio was 0.53. The optimum ratio is affected by the inlet temperature, ST/C ratio and pressure, and should therefore be the last parameter to be optimized. This is not evaluated further but a ST/C ratio in the range 0.5-0.6 should be adequate for this purpose.

Low inlet temperature to the WGS reactors should be as low as possible within the limits of the catalyst and the dew point of the inlet gas. The energy out of the burner (to the cracking furnace) is maximized when the inlet is combusted stoichiometrically.

4.2. Case 5 Partly maximized for H₂ production

Case 5 has its starting point from Case 2 and the goal is to increase the H₂ production. The snapshot from Aspen HYSYS is identical to Case 2 (Figure 5). Since the CO₂ separation and H₂ purification is not decided, and absorption with MDEA and PSA is considered as the state-of-the-art technology the

pressure is not reduced further. The reduction would improve the production rate but since there is likely to be a pressure requirement down the process line it is kept the same. The ST/C ratio is increased to 2, which is the highest value within the normal operating condition of an ATR. The inlet temperature to the ATR is risen from 650 to 750 °C, and not higher due to assumed material and corrosion restrictions. Lastly is a case study to find the optimal O₂/C ratio for this system and the peak in hydrogen production is when the flow of O₂ is 310 kmole/h, which corresponds to a ratio of 0.52, slightly lower than the one observed for Case 2 (0.53). It should be noted that the amount and the composition in the recycle steam changes when parameter changes, causing some small variations in the ratios. Next is that the inlet temperature to the HT- and LT-WGS reactors are kept at the (normal) minimum of 300 and 175 °C, respectively. It is not unlikely that these temperatures can be even lower in the near future, if the catalyst development continues. The hydrogen produced in this system amounts to 4659 kmole/h H₂, exceeding the amount that in section 6.1 was stated as the minimum (4565 kmole/h) if the same ethylene production rate were to be kept. When burned stoichiometrically the energy supply to the cracker is 7.615*10⁸ kJ/h. This is above what is necessary and can be reduced by reducing the ethane feed again.

If the ethane is reduced to 22 kmole/h (keeping temperature and ratios the same) the hydrogen to the burner will be 4571 kmole/h. This is a sufficient amount and the energy from the burner to the cracker will be 7.485*10⁸ kJ/h. An initial calculation of steam export of case 5 was found to be 21 MW. Case

5 is an attractive solution. It is recommended to develop this alternative further.

5. Conclusion

The main goal of this work was to investigate if fuel replacement by reforming the current fuel to only hydrogen can cover the energy demand in a steam cracking furnace for ethylene production.

The selected method for hydrogen production was an autothermal reforming process (ATR) with integrated pre-combustion CO₂ capture. The process is simulated in Aspen HYSYS and partly maximized for the hydrogen production. Five different configurations and nine case studies were examined for optimization.

The result was a system with a pressure of 2290 kPa (in the ATR), Steam/Carbon-ratio of 2, O₂/C-ratio of 0.52, and an inlet temperature to the ATR of 750 °C. The inlet temperature to the high and low-temperature water-gas-shift reactors was 300 °C and 175 °C, respectively. The inlet flow was 514 kmole/h methane and 2976 kmole/h hydrogen, where 95 mol-% of the hydrogen is separated prior to the reformation process.

To cover the gap between the simulated and desired flow of hydrogen, an inlet flow of ethane was added to the process. By adding only 22 kmole/h of ethane, a sufficient amount of hydrogen flow was reached. Potential steam export was found to be 21 MW, excluding the heat required for a CO₂ and H₂ separation unit.

Fuel replacement in industrial furnaces can have a significant impact towards decarbonization of ethylene production. Reforming traditional fuels containing hydrocarbons to hydrogen is shown to have potential.

Acknowledgements

The presented work is based on the Master thesis of Emmy Kristine L. Rustad. We would like to express our gratitude to INEOS Rafnes AS for being collaborative and helpful.

References

- Aas, K.L. *et al.* (2020) *Veikart for en klimapositiv industriregion*. (SINTEF, Periti, USN) Retrieved 20210220, <https://industrialgreentech.com>
- Amghizar, I. *et al.* (2020) 'Sustainable innovations in steam cracking: CO₂ neutral olefin production', *Reaction Chemistry & Engineering*, 5, pp. 239-257.
- Baltrusaitis, J. and Luyben, W.L. (2015) 'Methane conversion to syngas for gas-to-liquids (GTL): is sustainable CO₂ reuse via dry methane reforming (DMR) cost competitive with SMR and ATR processes?', *ACS Sustainable Chemistry & Engineering*, 3, pp. 2100-2111.
- Basile, A. *et al.* (2015) 'In *Compendium of Hydrogen Energy: Hydrogen Production and Purification. Vol. 1 Hydrogen and Production*. Woodhead Publishing; Cambridge, United Kingdom.

- Rashid, M. F. A. B. (2019) 'Optimization study of Fischer-Tropsch gas-to-liquid (GTL) process using Aspen HYSYS', *BSc thesis*. Universitat Politècnica de Catalunya, Spain.
- Chen, W.H. *et al.* (2010) 'Thermodynamic analysis of hydrogen production from methane via autothermal reforming and partial oxidation followed by water gas shift reaction', *International Journal of Hydrogen Energy*, 35, pp. 11787-11797.
- Souza, T. L. *et al.* (2015) 'Thermodynamic Analysis of Autothermal Reforming of Methane via Entropy Maximization: Hydrogen Production', *Blucher Chemical Engineering Proceedings*, 1, pp. 15633-15646.
- INOVYN (2021) 'INOVYN launches clean hydrogen project to support', www.ineos.com. Retrieved 20210316, <https://www.ineos.com/businesses/inovyn/news/inovyn-launches-clean-hydrogen-project-to-support-decarbonisation-in-norway>.
- Jakobsen, D. (2016) 'Concepts for large scale hydrogen production', *MSc thesis*. NTNU, Norway.
- Krohn-Fagervoll, S. (2020) 'Get ready for industrial halls for NEL, a game changer', <https://eng.heroya-industripark.no>. Retrieved 20200304.
- Mathisen, M. W. (2021). 'Outlet composition of the cracking furnace', *Personal Communication*.
- Moulijn, J.A. *et al.* (2013) *Chemical process technology*. John Wiley & Sons, Chichester, United Kingdom.
- Negri, A. and Ligthart, T. (2021) 'Decarbonisation options for the Dutch polyolefins industry', PBL Netherlands Environmental Assessment Agency; TNO; The Hague, The Netherlands.
- Oyama, S.T. (2012) 'Dry reforming of methane has no future for hydrogen production: Comparison with steam reforming at high pressure in standard and membrane reactors', *International Journal of Hydrogen Energy*, 37, pp. 10444-10450.
- Ren, T. (2008) 'Steam cracking and methane to olefins: Energy use, CO₂ emissions and production costs', *Energy*, 33, pp. 817-833.
- Nielsen, J and Christiansen, L.J. (2011) 'Concepts in syngas manufacture' *Catalytic Science Series*. Vol 10; Hutchings, G.J.; Series Ed., Imperial College Press.
- Rustad, E.K. (2021) 'Reducing the CO₂ emissions of a gas cracker by reforming fuel gas', *Master Thesis*. University of South-Eastern Norway.
- Soltani, R.M. (2014) 'Assessment of CO₂ capture options from various points in steam methane reforming for hydrogen production', *International Journal of Hydrogen Energy*, 39, pp. 20266-20275.
- Ullmann's Encyclopedia of Industrial Chemistry (1988) 'Ethanalamines and Propanolamines to Fibers, Synthetics'; Vol. A 10, Wiley-VCH; Weinheim, Germany.
- Voldsund, M. (2016) 'Hydrogen production with CO₂ capture', *International Journal of Hydrogen Energy*, 41, pp. 4969-4992.
- Wismann, S.T. (2019) 'Electrified methane reforming: A compact approach to greener industrial hydrogen production', *Science*, 364, pp. 756-759.

Dynamic modeling of diafiltration system for a biorefinery application

Markku Ohenoja*, Pekka Uusitalo, Hanna Valkama, Buddhika Rathanyake, Riitta L. Keiski, Mika Ruusunen

Environmental and Chemical Engineering Research Unit, University of Oulu, Finland
markku.ohenoja@oulu.fi

Abstract

Application of membrane technologies in biorefinery processes has been studied for some time. The heterogenous nature of biorefinery streams, however, results in unideal performance of membrane systems and considerable fouling of membranes, which is decreasing the efficiency of separation. As a part of BioSPRINT project, this study focuses on application of separating monomeric sugars from the hemicelluloses fraction of lignocellulosic biomass, where pressure-driven nanofiltration with several diafiltration stages has been proposed for the separation task. Diafiltration is required to overcome the decreased separation efficiency when the retentate concentrations and viscosity increases. A lumped parameter dynamical model of the diafiltration plant is applied. The key model parameters are identified from experimental data from a laboratory membrane unit to reflect the considered biorefinery process. The model is then simulated to study the sensitivity of the uncertain model parameters (related to membrane fouling, solute concentrations, viscosity, and mass transfer coefficients) to the diafiltration plant performance (product purity, operation time). The model is implemented in the MATLAB[®]/Simulink environment. The simulation results are expected to identify potential sources of scale-up challenges in biorefinery-related membrane applications. The developed dynamic model also allows to investigate different operational strategies of diafiltration plants in the future.

1. Introduction

In order to mitigate the global challenges related to the environmental impact and climate change, the bio-based economy drives to substitute fossil-based fuels and chemicals with their more sustainable, bio-based counterparts (European Commission and Directorate-General for Research and Innovation, 2017). Chemical industries are energy-intensive and mostly relied on fossil-based petrochemical feedstocks; the use of biomass as a feedstock for intermediate chemicals production can be seen as a major shift toward more sustainable production (Fiorentino *et al.*, 2019). From the potential raw materials, lignocellulosic biomass has a great potential having a good availability and being a non-food biomass.

In the BioSPRINT project,¹ the hemicellulosic fraction of lignocellulosic biomass is valorized into furan-based polymers. The hemicelluloses are heterogenous polysaccharides comprising a range of sugar monomers, such as xylose, glucose, arabinose, mannose *etc.* However, in biomass side streams from different biorefineries (*e.g.*, pulping, steam explosion or fermentation), the hemicelluloses are in a water-solute mixture also

containing impurities such as short-chain acids, degradation products, phenolic compounds, and soluble inorganic species. In catalytic conversion, these can have considerable interactions with the catalyst. Thus, efficient purification of hemicellulosic sugars is needed prior to the catalytic conversion steps to produce bio-based chemicals.

Membrane technologies have been proposed for different kinds of biorefinery applications (Abels *et al.*, 2013). Applicability studies of nanofiltration (NF), reverse osmosis (RO) and ultrafiltration (UF) membranes for lignocellulosic hydrolysates includes examples such as fractionating hemicelluloses from pulp mill process waters (Krawczyk, 2013), concentration and purification of lignin from pulping liquors (Jönsson *et al.*, 2008), and development of purification cascades for the Organosolv process (Nitzsche *et al.*, 2022). Some of the results have also been transferred to the pre-industrial scale (Nguyen *et al.*, 2016). Diafiltration is a membrane purification technology, where the retentate (stream that the membrane rejects) is diluted to overcome the decreased separation efficiency when the retentate viscosity increases, and thus enables further removal of the impurities to the permeate (stream

¹ <https://www.biosprint-project.eu/>

that passes the membrane) (Doran, 2013; Nguyen *et al.*, 2016).

Process modeling typically plays an important role in process design and scale-up. In the case of membrane systems, the scale-up is governed by permeate flux values, which affect the separation efficiency, purity of product stream and required membrane surface area (sizing). The membrane fouling, which is accumulation of material to membrane surfaces and pores thus decreasing the permeate flux, is another important factor in process design. Theoretical analysis cannot offer tools to predict all of the aspects (Doran, 2013). For instance, (Nguyen *et al.*, 2016) stated that the pilot-scale membranes may have 30–45% lower water permeability than observed in laboratory membrane units. Thus, the mathematical modeling of diafiltration processes has focused, for example, on studying intermittent diluent addition scenarios in different levels of details (Wang *et al.*, 2008; Tan and Franzreb, 2022), prediction of fouling dynamics (López-Murillo *et al.*, 2021), or integrated process and control design (Hunter, 2000; Yee *et al.*, 2012; Saltık *et al.*, 2017).

This study takes the first steps related to the integrated process and control design of a diafiltration plant processing hemicelluloses from lignocellulosic biomass. A dynamical mass-balance based model of the diafiltration plant is implemented, the key model parameters are identified from laboratory membrane experiments and the scaled-up purification step of the biorefinery process is simulated. The simulations in this study focus on assessing the uncertainty of laboratory data derived model parameters on a selected design criterion.

The process model is presented in Section 2 together with a short description of acquisition of the experimental data. Section 3 focuses on the parameter estimation of the model and the development of a scaled-up simulator. The sensitivity of the selected model parameters is studied in Section 4, followed by the summary in Section 5.

2. Material and methods

2.1. Simulation model

A lumped parameter dynamical model of the diafiltration plant originated from (Hunter, 2000) is applied. The model comprises retentate mass balances of component j (see Eq. 1) over membrane stage i given in volumetric flows Q and concentrations C as weight fractions. Subscripts F , DF , P and R stand for feed, diafiltration, permeate and retentate, respectively. Fig. 1 illustrates the stage i with inputs, outputs and state variables given in Eq. 2 and Eq. 3.

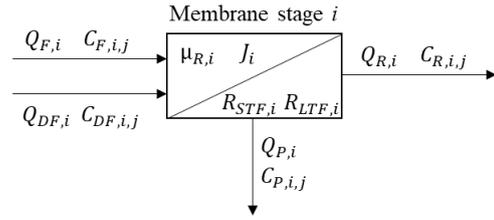


Figure 1: Model boundaries and variables for a single membrane stage.

The equation set for the case, where the feed flow is treated as a degree of freedom, is as follows:

$$\rho_{R,i} V_i \frac{dC_{R,i,j}}{dt} = \rho_{F,i} Q_{F,i} C_{F,i,j} + \rho_{DF,i} Q_{DF,i} C_{DF,i,j} - \rho_{P,i} Q_{P,i} (1 - R_j) C_{R,i,j} - (\rho_{F,i} Q_{F,i} + \rho_{DF,i} Q_{DF,i} - \rho_{P,i} Q_{P,i}) C_{R,i,j} \quad (1)$$

$$\frac{dR_{STF,i}}{dt} = 10^{-13} \frac{k_1 J_i C_{R,i,foul} \mu_{R,i}}{\rho_{R,i}} - k_2 R_{STF,i} \quad (2)$$

$$\frac{dR_{LTF,i}}{dt} = k_3 R_{STF,i} \quad (3)$$

$$J_i = \frac{\Delta P_{TMP} - \Delta \pi_i}{\mu_{P,i} R_{tot,i}} \quad (4)$$

$$R_{tot,i} = R_{P,i} + 10^{13} (R_{STF,i} + R_{LTF,i}) \quad (5)$$

$$R_{P,i} = \max \left\{ \frac{\Delta P_{TMP} - \Delta \pi_i}{\mu_{P,i} K_{m,i} \ln \left(\frac{C_{gel,i}}{C_{R,i,foul}} \right)} - R_m, 0 \right\} \quad (6)$$

Where $\rho_{R,i}$ is the retentate stream density in stage i , V_i is the stage volume, $\rho_{F,i}$, $\rho_{DF,i}$, $\rho_{P,i}$ are the stream densities and R_j is the retention coefficient for component j . In Eq. 2, the short-term fouling ($R_{STF,i}$) is given as a function of permeate flux J_i , concentration of fouling components ($C_{R,i,foul}$), dynamic viscosity of the retentate stream ($\mu_{R,i}$) and density ($\rho_{R,i}$). The two parameters k_1 and k_2 determine the rate of change. The long-term fouling (Eq. 3) is dependent on the short-term fouling via the parameter k_3 .

The total permeate flux (J_i) in Eq. 4 is dependent on the trans-membrane pressure (ΔP_{TMP}), the osmotic pressure ($\Delta \pi_i$), the permeate stream viscosity ($\mu_{P,i}$), and the total membrane resistance ($R_{tot,i}$, Eq. 5), which is a sum of the above fouling terms together with the concentration polarization resistance (R_p , Eq. 6). R_p is given as a function of ΔP_{TMP} , $\mu_{P,i}$, mass-transfer coefficient ($K_{m,i}$), threshold for fouling components concentration

($C_{gel,i}$), and constant membrane resistance (R_m). However, in this study a negligible osmotic pressure is assumed due to the low feed flow and working in narrow volumetric reduction ratio. Finally, by multiplying the permeate flux with the membrane cross-sectional area (A_i), the permeate flow is attained. In addition, the model implementation involves a perfectly mixed buffer tank for the purified stream (Eq. 7) and first-order dynamics for the diafiltration addition based on the permeate flux of the previous stage (Eq. 8). In the simulations, the inlet mass flow ($\dot{m}_{B,j,in}$) to the buffer tank is the final retentate stream and no outflow from the buffer tank is assumed. Thus, Eq. 7 simplifies into an integral.

$$\frac{d\dot{m}_{B,j}}{dt} = \dot{m}_{B,j,in} - \dot{m}_{B,j,out} \quad (7)$$

$$G(s) = \frac{Q_{DF,i+1}}{Q_{P,i}} = \frac{1}{0.95s + 1} \quad (8)$$

Other assumptions related to the diafiltration simulation model are as follows:

- No hydraulic losses along the flow direction.
- Stream consists of monomeric sugars, impurity components (small-chain acids and degradation products), phenolic compounds (lignin) and inorganics mixed to water.
- Only sugar and phenolic concentrations contribute to fouling ($C_{R,i,foul}$).
- The water content of the feed and permeate streams are high, thus the density and dynamic viscosity are equal to water at 25°C.
- The changes in the dynamic viscosity of retentate streams ($\mu_{R,i}$) are assumed to follow the relation for aqueous solution of glucose given in (Converti *et al.*, 1999).
- Retention coefficients R_j do not change in the studied range of flows, concentrations, and operation conditions.
- The model parameters k_1 , k_2 , k_3 , K_m , R_m and C_{gel} are similar for all stages i .

The stage model was implemented as a Matlab[®] function and extended to the diafiltration plant model in Simulink where several stage models can be connected to each other and solved simultaneously. In order to improve the numerical performance of the simulator, the exponential factors were used in Eq. 2 and Eq. 5. This allowed to balance the amplitudes of the derivatives during the numerical integration, where in total 52 ordinary differential equations were solved simultaneously. The initial states were zero, corresponding a start-up of a membrane system initially filled with water only. The 'ode45' variable-step solver was used in these simulations.

2.2 Experimental data

The laboratory-scale batch membrane filtration unit CELFA P-28 (CM-Celfa Membrantechnik AG, Switzerland) was used in the experiments with a commercially available thin film composite NF membrane. The experimental conditions correspond to the cross-flow velocity of 0.63 m/s and trans-membrane pressure of 19 kPa. The permeate flux values were recorded during the batch trial with a duration of 120 minutes, when the system was tested with a lignocellulosic hydrolysate comprising different sugar monomers, short-chain acids and degradation products, soluble phenols, and inorganic compounds. The unit was also run with pure water before and after the experiments to evaluate the flux degradation and calculating the model parameter R_m analytically. The average values of the permeate flux in four consecutive experiments were used in the parameter estimation in Section 3. The experiments were conducted in a concentration mode, but due to the low flux values, and the unavailability of reliable and frequent concentration measurements, it was assumed that the feed concentrations in the experiments are constant.

3. Parameter estimation and scale-up

The three fouling parameters k_1 , k_2 , k_3 and the mass-transfer coefficient K_m were estimated based on the flux behavior in a lab-scale unit. The model identification assumed a spiral-wound membrane leaf with similar cross-flow velocity as in the laboratory experiments. According to (Roy *et al.*, 2015), the leaf area of 1 m² and the channel height of 0.7 mm was chosen for a single leaf. These results in a feed flowrate of 1.58 m³/h (0.00044 m³/s).

The optimization of the parameters was conducted with the Nelder-Mead Simplex Method using Matlab[®] function 'fminsearch'. Fig. 2 shows the observed and modelled permeate flux given as a volumetric flow per unit area.

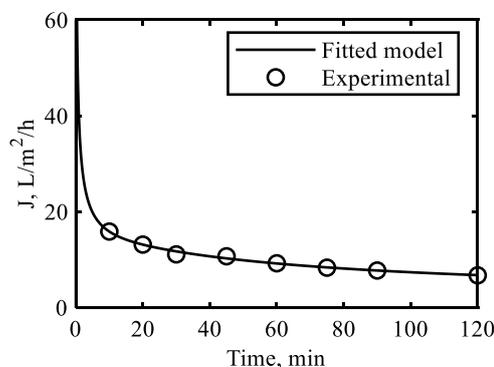


Figure 2: Permeate flux.

The estimated parameters are given in Tab. 1. Other model parameters are listed in Tab. 2. The lumped retention coefficients for sugars, impurities and inorganics are selected based on the experimental results and the retention coefficient of phenols is assumed to follow the sugar retention for simplicity.

Finally, the system was scaled-up for the feed flow value of 20,000 m³/h. The total membrane area was selected to be near the minimal feed flow rate of 60 L/h for each leaf to minimize the effect of osmotic pressure in NF (Roy *et al.*, 2015). Seven diafiltration stages with identical membrane areas and cross-membrane pressures were assumed to the scaled-up plant configuration.

Table 1: Estimated model parameters.

$K_m \cdot 10^{-5}$ [m/s]	$k_1 \cdot 10^{24}$ [s/m ⁴]	$k_2 \cdot 10^{-3}$ [1/s]	$k_3 \cdot 10^{-4}$ [1/s]
7.54	1.74	1.81	5.36

Table 2: Other model parameters.

Parameter	Value	Unit	Note
A	330	m ²	
ΔP_{TMP}	19	kPa	
ρ_P	997.05	g/L	Equal to water in 25°C
ρ_R	997.05	g/L	Equal to water in 25°C
C_{gel}	0.15	-	Estimated from (Doran, 2013)
μ_P	889.1	mPas	Equal to water in 25°C
R_m	$3 \cdot 10^{13}$	1/m	
R_{sugars}	0.93	-	Retention coefficients
$R_{impurities}$	0.22	-	
$R_{phenols}$	0.93	-	
$R_{inorganics}$	0.67	-	
$Q_{F,1}$	0.0056	m ³ /s	Plant feed flow
C_{sugars}	80	g/L	Feed concentrations
$C_{impurities}$	20	g/L	
$C_{phenols}$	7.04	g/L	
$C_{inorganics}$	0.97	g/L	

4. Sensitivity analysis

Even the continuous membrane systems are actually operated in a batchwise-continuous manner, as the static operation cannot be achieved due to the progressive membrane fouling, requiring periodic production breaks for cleaning (Hunter, 2000). Thus, it is assumed that in the simulated biorefinery case, two operational boundaries hold; (1) The sugar purity in the buffer tank need always to be over 80% to be viable in downstream conversion steps, and (2) the maximum running time of diafiltration plant is 8 hours before service (cleaning).

By simulating the system with the feed concentrations shown in Tab. 2, these thresholds

are barely met. Fig. 3 shows the buffer purity as a function time, closing to value of 80% after 8 hours (480 minutes). The model was then simulated to study the sensitivity of uncertain model parameters (related to membrane fouling, solute concentrations, viscosity, and mass transfer coefficients) to the diafiltration plant performance by recording the operation time, where the purity threshold was violated.

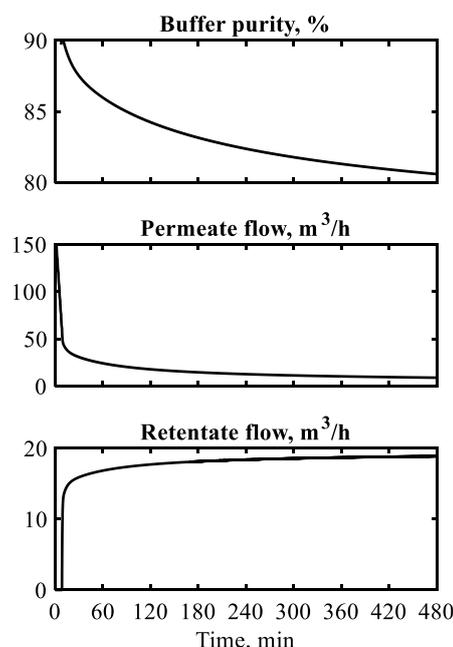


Figure 2: Base case simulation of a diafiltration plant. From top to bottom, sugar purity in buffer tank, total permeate flow, and product flow to the buffer tank, respectively.

In total, 33 simulations were performed by varying model parameters one-at-a-time in different extent. Tab. 3 presents qualitatively the direction of change (in product purity) and quantitatively the operational time, where the above-mentioned purity threshold was not met in the simulation. It can be seen that most of the parameters have only a small effect on this criterion; Even changes $\pm 50\%$ in K_m , R_m or C_{gel} show no effect on the selected criterion. Less surprisingly, the three fouling parameters (k_1 , k_2 , k_3) show more sensitivity to the selected criterion. ρ_R was only simulated with slightly higher values (up to +5%) as it can be expected that the stream density cannot exceed substantially higher values due to low permeate fluxes and diafiltration water addition. μ_R shows a moderate effect only after an increment of over 30% in its value. This kind of an increment in a stream dynamic viscosity is not expected unless the stream contains longer hydrocarbons (such as oligomeric sugars, alcohols), which is not the case with the studied stream.

Table 3: Direction of change in product purity and the time of threshold violation (in minutes). The target operation time for the diafiltration plant was 480 minutes.

Parameter	-50%	-25%	-15%	+5%	+15%	+25%	+30%	+50%
ρ_R				↑				
μ_R				↓	↓	↓	466	
R_m	↓		↓			↑		↑
K_m	↓	↓						
C_{gel}	↓	↓	↓		↑	↑		
k_1			↑			↓	466	401
k_2	299	454	↓			↑		
k_3		↑				↓	471	
$C_{phenols}$	↑	↑	↑		↓	463		336

The solute feed concentration changes were simulated by adjusting the ratio of impurities and phenolic compounds, and this way keeping the initial stream purity at constant value. As the retention coefficient of phenols is equal to that of sugars, the higher amount of phenols in the feed stream make the purification less efficient. This can also be seen in the results presented in Table 3, where the increment of phenol concentration quite easily results in a violation of the purity threshold. Thus, in the case of inhomogeneous diafiltration feedstock composition, the dimensioning should be made for the worst-case scenario, *i.e.*, for the high phenolic concentration. Otherwise, too optimistic performance of the membrane system is expected, having a considerable effect on the operational and cleaning cycles.

K_m , R_m or C_{gel} are all related to the fixed membrane resistance and the concentration polarization fouling in the applied model (Eq. 6). It has been reported that the concentration polarization is the major fouling mechanism in these kinds of real hydrolysate streams, contributing 71–82% to the total resistance (Nitzsche *et al.*, 2021). Hence, the negligible sensitivity in these simulation results indicate that the applied model cannot adequately represent the concentration polarization behavior and the fouling is dominated by the short-term and long-term fouling instead (Eq. 2 and Eq. 3).

According to the simulation results (Tab. 3), the fouling parameters related to the short-term and long-term fouling (k_1 , k_2 , k_3) require quite significant changes before the purity threshold is compromised. By looking at Eq. 2, it can be observed that the parameter k_1 should be the most dominant as it initializes the dynamic fouling effects. Indeed, around 30% larger value for k_1 results in a shortened operation time. The effect of parameter k_2 is opposite; it is a negative gain term for the short-term fouling (see Eq. 2) and thus a smaller value accelerates the fouling, whereas a larger value mitigates the fouling dynamics. Finally, parameter k_3 basically describes the shift from short-term fouling to long-term fouling and it

takes of around +30% change in the parameter value to deteriorate the diafiltration process outside of the performance criteria.

Fig. 3 presents the product purity as a function of time for the base case (thick dash-dotted line) and for the simulations with changes in parameters k_1 , k_2 , k_3 (thin gray lines). Obviously, three of the parameter combinations provide improved performance (as also qualitatively shown in Tab. 3), where the purity target is still clearly exceeded at time 480 minutes. For these cases, the total permeate flow at the end of simulation is slightly over 10 m³/h, whereas for the base case, the total permeate flow was 8.93 m³/h. Hence, even the modest changes in the flux values correspond to the increased availability of the membrane plant, decreased membrane area requirements, or increased throughput.

On the other hand, the worst simulated case (k_2 –50%) shows the total permeate flow of 6.3 m³/h at the end of simulation. At the time of product purity violation ($t = 299$ minutes), the permeate flux in this case was around 8 m³/h. It is known that the flux degradation of 30–45% can be expected during the scale-up of membranes (Nguyen *et al.*, 2016). In the simulated cases, the total permeate flow between the best and the worst cases was around –37%. Hence, the parameter ranges of the fouling parameters used may reflect the uncertainty related to the scale-up.

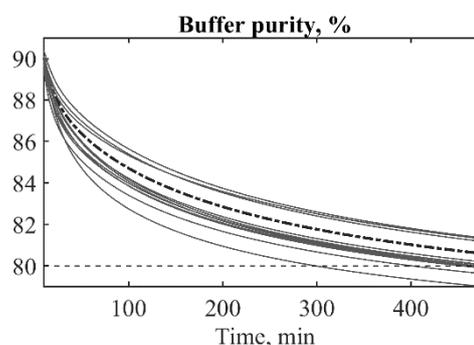


Figure 3: Envelope of diafiltration plant product purity in the simulated fouling scenarios.

5. Summary

This study presented the development of a dynamic simulation model describing a diafiltration system for hemicellulose purification in a biorefinery process. The key model parameters related to the membrane fouling were estimated based on experimental data from a laboratory membrane unit. The model was then scaled-up to reflect an industrial scale system.

Since the model is governed by high number of model parameters, this study focused on assessing the sensitivity of selected model parameters to operational aspect of the diafiltration plant. The simulation results showed that the expected range of uncertainty in many of the studied parameters had a sufficiently small effect on the process (dynamic) performance. This was unexpected especially for the parameters related to the concentration polarization behavior. However, the fouling parameters, together with the feed stream concentrations, showed an important effect on the product purity and operational time of the diafiltration plant.

The overall feasibility of such a diafiltration plant depends not only on the factors studied in these simulations. For example, recovery of valuable compounds can limit the process design; in the studied case, the monomeric sugars are not rejected completely ($R_{sugars} = 0.93$), therefore causing a loss of sugars to the permeate streams. The model should be accompanied with permeate mass balances to explore these in detail. These would also allow to implement calculation of transmembrane pressure losses due to osmotic pressure (see e.g. (Nguyen *et al.*, 2016)) and this way probably enhance the modeling of the concentration polarization behavior.

Evaluation of economic performance of a diafiltration plant requires more detailed specifications on membrane module configurations and cleaning intervals. For instance, tubular modules instead of spiral-wound modules are more preferable if frequent cleaning is required, but they have lower packing density (Doran, 2013). In addition, energy costs need to be assessed. They are related to pumping costs of streams and diafiltration water and the need to maintain transmembrane pressure throughout the system.

The developed dynamic model allows to study different process configurations (number of stages, recycle flows, diafiltration strategies) and can be utilized in integrated process and control design in the future. It should be noted that the solver selection for this simulation model requires extra care for example in controllability studies, where new membrane modules or diafiltration feeds are started during the operation, thus creating discontinuities to the simulation.

Acknowledgements

This project has received funding from the Bio-based Industries Joint Undertaking (JU) under the European Union's Horizon 2020 research and innovation program under grant agreement no. 887226. The JU receives support from the European Union's Horizon 2020 research and innovation program and the Bio-based Industries Consortium.

References

- Abels, C., Carstensen, F. and Wessling, M. (2013) 'Membrane processes in biorefinery applications', *Journal of Membrane Science*, 444, pp. 285–317. doi: 10.1016/j.memsci.2013.05.030.
- Converti, A., Zilli, M., Arni, S., Di Felice, R. and Del Borghi, M. (1999) 'Estimation of viscosity of highly viscous fermentation media containing one or more solutes', *Biochemical Engineering Journal*, 4(1), pp. 81–85. doi: 10.1016/S1369-703X(99)00028-5.
- Doran, P.M. (2013) 'Chapter 11 - Unit Operations', in P.M. Doran (ed.) *Bioprocess Engineering Principles*. 2nd edn. London: Academic Press, pp. 445–595. doi: 10.1016/B978-0-12-220851-5.00011-3.
- European Commission and Directorate-General for Research and Innovation (2017) *Review of the 2012 European Bioeconomy Strategy*. <https://data.europa.eu/doi/10.2777/086770> (Accessed: 4 July 2022).
- Fiorentino, G., Zucaro, A. and Ulgiati, S. (2019) 'Towards an energy efficient chemistry. Switching from fossil to bio-based products in a life cycle perspective', *Energy*, 170, pp. 720–729. doi: 10.1016/j.energy.2018.12.206.
- Hunter, T.J. (2000) 'Dynamics and control of multistage membrane plants'. Doctoral thesis. University of Canterbury. Available at: <https://doi.org/10.26021/2835>.
- Jönsson, A.-S., Nordin, A.-K. and Wallberg, O. (2008) 'Concentration and purification of lignin in hardwood kraft pulping liquor by ultrafiltration and nanofiltration', *Chemical Engineering Research and Design*, 86(11), pp. 1271–1280. doi: 10.1016/j.cherd.2008.06.003.
- Krawczyk, H. (2013) *Separation of Biomass Components by Membrane Filtration - Process Development for Hemicellulose Recovery*. Doctoral Thesis. Lund University. Available at: <https://portal.research.lu.se/ws/files/5705805/3990689.pdf>.
- López-Murillo, L.H., Grisales-Díaz, V.H., Pinelo, M. and Prado-Rubio, O.A. (2021) 'Ultrafiltration intensification by dynamic operation: Insights from hybrid modeling', *Chemical Engineering and Processing - Process Intensification*, 169, p. 108618. doi: 10.1016/j.cep.2021.108618.
- Nguyen, D.T.N.N., Lameloise, M.-L., Guiga, W., Lewandowski, R., Bouix, M. and Fargues, C. (2016) 'Optimization and modeling of diafiltration process for the detoxification of ligno-cellulosic hydrolysates - Study at pre-industrial scale', *Journal of Membrane Science*, 512, pp. 111–121. doi: 10.1016/j.memsci.2016.04.008.
- Nitzsche, R., Köchermann, J., Gröngröft, A. and Kraume, M. (2021) 'Nanofiltration of Organosolv Hemicellulose Hydrolyzate: Influence of Hydrothermal Pretreatment and Membrane Characteristics on Filtration Performance and Fouling', *Industrial & Engineering Chemistry Research*, 60(2), pp. 916–930. doi: 10.1021/acs.iecr.0c03256.
- Nitzsche, R., Etzold, H., Verges, M., Gröngröft, A. and Kraume, M. (2022) 'Demonstration and Assessment of Purification Cascades for the Separation and Valorization of Hemicellulose from Organosolv Beechwood Hydrolysates', *Membranes*, 12(1), p. 82. doi: 10.3390/membranes12010082.
- Roy, Y., Sharqawy, M.H. and Lienhard, J.H. (2015) 'Modeling of flat-sheet and spiral-wound nanofiltration configurations and

its application in seawater nanofiltration', *Journal of Membrane Science*, 493, pp. 360–372. doi: 10.1016/j.memsci.2015.06.030.

Saltik, M.B., Özkan, L., Jacobs, M. and van der Padt, A. (2017) 'Dynamic modeling of ultrafiltration membranes for whey separation processes', *Computers & Chemical Engineering*, 99, pp. 280–295. doi: 10.1016/j.compchemeng.2017.01.035.

Tan, R. and Franzreb, M. (2022) 'Simulation-based evaluation of single pass continuous diafiltration with alternating permeate flow direction', *Separation and Purification Technology*, 282, p. 119987. doi: 10.1016/j.seppur.2021.119987.

Wang, L., Yang, G., Xing, W. and Xu, N. (2008) 'Mathematic model of the yield for diafiltration processes', *Separation and Purification Technology*, 59(2), pp. 206–213. doi: 10.1016/j.seppur.2007.06.007.

Yee, K.W.K., Bao, J. and Wiley, D.E. (2012) 'Dynamic operability analysis of an industrial membrane separation process', *Chemical Engineering Science*, 71, pp. 85–96. doi: 10.1016/j.ces.2011.11.046.

Simulation-Based Life Cycle Assessment for Office Building Façade: A Case Study of the Leadenhall Building in London

Tahmineh Akbarinejad Khameneh ^{a,*}, Zahir Barahmand ^b, Gamunu Samarakoon ^b

^a Department of Civil and Environmental Engineering, Norwegian University of Science and Technology,

^b Department of Process, Energy and Environmental Technology, University of South-Eastern Norway

Tahmineh.akbarinejad@ntnu.no

Abstract

High-rise glazing systems are among the most important components affecting energy efficiency. Through the lens of Life Cycle Assessment, glass has always been an unlikely material for large buildings due to its considerable energy consumption throughout the pre-use and post-use phases. Moreover, the use of high-tech materials has a negative impact on the environment. Therefore, the present study aims to assess a comparative life cycle of four different glazing system technologies (BIPV, smart glass, low-E, and double glazing) representing the most used commercial high potential glazing systems. The next step has been optimized for the Leadenhall iconic tower as the case study. In this analysis, energy simulation is combined with life cycle assessment to investigate the environmental impacts. ZEB-COM tool, Rhino®, and Grasshopper® have been used to calculate emissions, 3D modeling, and energy modeling, respectively. The results reveal that BIPV achieved 37% of total energy-saving and stood first. A hybrid solution (two glazing systems) has been proposed to eliminate negative aspects and increase livability. Although it can generate almost 30% less energy than the complete BIPV installation, with a specific design by the authors, it can cover aesthetic concerns in this system and compensates for 27% of the total energy demand of the Leadenhall project.

Keywords: life cycle assessment, glazing system, comparative analysis, Leadenhall building, high-rise,

1 Introduction

Climate change has become one of the most crucial global concerns (O'Neill et al., 2021). The construction industry is a big energy user with significant environmental consequences which are not negligible (Stegou-Sagia et al., 2007). Saving energy, particularly in buildings, is worthy of attention (Hee et al., 2015). In developed countries, residential and commercial building energy use accounts for 20–40% of total final energy consumption in the country (Pérez-Lombard et al., 2008). These energies are primarily utilized for residential space heating, cooling, and lighting for commercial buildings. In a building, the energy efficiency is influenced by its envelope, particularly its windows (Lee et al., 2013). Buelow-Huebe (2001) reported that a window is responsible for 20–40% of a building's lost energy. The minimum size of a building's window is required to restrict heat gain or loss (Buelow-Huebe, 2001).

On the other hand, a window enables natural light to enter a building. Furthermore, studies have proved the health benefits of natural sunlight and a view of the outdoors (Chang & Chen, 2005; Hee et al., 2015). The net energy gain from glazing and

windows in buildings is determined by thermal and total solar energy transmittance. Therefore, picking optimal glazing systems or windows for a specific case regarding energy efficiency is challenging (Nielsen et al., 2001).

On the one hand, industrial development and increased housing and construction requirements led to skyscrapers' use (Ahmad et al., 2017). On the other hand, with global warming and the depletion of fossil fuels, interest in zero-energy buildings is increasing (Bravo-Hidalgo & Baez-Hernandez, 2019). In the next 20 years, the global market for zero-energy buildings is expected to overgrow, worth nearly \$1.3 billion in 2035 (Cao et al., 2016). Advanced glazing systems have become an urgent requirement for high-rise office buildings to minimize energy consumption and adapt to external environmental conditions (He et al., 2019).

Using glazing technologies as building facades are becoming more popular (Rezaei et al., 2017). In a building, thermal comfort, light comfort, and skin health are all intimately tied to windows (Edlich et al., 2004). Furthermore, glazing systems provide acoustic comfort, vision, ventilation (Park & Kim, 2015), and photoprotection (Tuchinda et al., 2006).

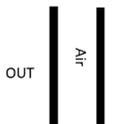
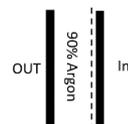
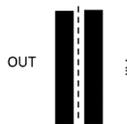
Scenarios	Basic	Passive	Active	Active
Type	Clear Double Glazing system	Low E-class	Smart Glass-Low Emittance	BIPV Semi-transparent
Description	The Clear double-glazing system (CDGS) is the conventional glazing system in office buildings, which has few environmental advantages and low initial cost.	It is coated by layers of selective coatings which reflect the infrared radiation without any significant loss in visible light transmission.	electrochromic glazing (ECGS) uses a low voltage (typically under 5 V) to control its visible light and heat transmittance properties.	BIPV provides sun shading and generates electricity by the semitransparent photovoltaic films.
Cross-Section				
Specification	6mm clear glass pane - 16mm Sealed air - 6mm clear glass pane	6mm (Dual-coated) glass - 16mm 90% argon cavity - (Low-E) coating -6mm glass	6mm green glass, 0.89mm Sentry Glass, 2.2mm Sage Glass, 10mm air space w/90% Argon fill, 6mm clear glass	(6+3+6) mm amorphous silicon photovoltaic glass with a transparency 20% + 12mm air chamber + (4+4) mm clear glass

Figure 1: Specifications and descriptions for the alternatives

Because of these critical roles, windows must be designed and selected based on various factors, notably energy efficiency and visual comfort.

There are different approaches to classifying glazing systems. The glazing system, for example, can be classified into conventional glazing systems, advanced glazing materials and coatings, and smart (intelligent) technologies (Rezaei et al., 2017).

LCA is a widely used tool for comparing the environmental impacts of products or systems to assist decision-makers in selecting the most sustainable alternative (Elkhatay et al., 2020). Therefore, this study uses Life Cycle Assessment (LCA) methodology to investigate and compare the total environmental impacts of the four glazing systems in a commercial high-rise. For a sustainable building design, the starting point is a global "view" of the building's efficiency associated with all phases of a product's life cycle, including sourcing raw materials, processing these resources, producing, distributing, using, maintaining, and repairing the product, reselling or recycling them and disposal (Hernandez et al., 2019).

Through a comparative analysis, the present study aims to make a simulation-based life cycle assessment of four different glazing system technologies: conventional double glazing as the base case, passive Low-E glazing, Building-Integrated Photovoltaic (BIPV), and electrochromic smart glazing. These technologies are among the most worldwide used commercial high potential glazing systems. As the case study, these assessments are investigated for London's iconic commercial high-rise, the Leadenhall building. In this integrated approach, energy simulation is combined with life cycle assessment to investigate the environmental impacts.

The LCA results will determine which system has the lowest energy use and environmental impacts to

support building designers and decision-makers in choosing the most environmentally friendly glazing system for their office buildings.

2 Methodology

The life cycle assessment methodology is generally broken down into four steps (Hernandez et al., 2019): goals and scope definition, Life Cycle Inventory (LCI) analysis, Life Cycle Impact Assessment (LCIA), and interpretation of the results. This comparative LCA analysis is accomplished according to the ISO14040:2006 guidelines and framework (2006a, 2006b), and (IEA) guidelines related to PV and BIPV, and the Norwegian Zero Emission Building (ZEB).

2.1 Goals and scope definitions

This comparative LCA study aims to find and contrast the environmental impacts of four different glazing systems to determine which glazing system's life cycle has the lowest energy use and environmental impacts for the case of a commercial high-rise in London. The Leadenhall building, known as Cheesegrater (Booth, 2014), launched in 2014, has 225 meters in height. As a result, it will aid in developing solutions to reduce the environmental impact of glazing systems. The life cycle impact data and analysis are supposed to assist architects and decision-makers in prioritizing the most environmentally friendly glazing system for high-rise office buildings.

The LCA evaluated the environmental impacts of the glazing system over its pre-use phase (manufacturing and construction) and use phase. ZEB Tool provided the LCI data associated with material extraction and manufacturing steps, and the Grasshopper® with Honeybee® plug-in calculated the BIPV energy produced and overall building's energy demand in the used-phase in Rhino®.

The scenarios are chosen by comparing the three major glazing system categories: passive, active, and BIPV, with a conventional double glazing system as the base case. Each class represents the most widely used commercial glazing technology for office building facades worldwide.

2.1.1 Functional unit

Defining a Functional Unit (FU) is crucial for developing and modeling a product system in Life Cycle Assessment (LCA). (Arzoumanidis et al. (2019) defined the functional unit as “a quantified description of the function of a product that serves as the reference basis for all calculations regarding impact assessment.”

The present study considered the functional unit of the case study (the Leadenhall building) and four glazing systems alternatives (Fig. 1) as followings.

- **Case study:** the Leadenhall building in London. This 48-story building has 225 meters in height, 57,000 m² gross area, and a 75,000 m² façade area with a curtain wall glazing system. The service life of this project is considered 60-years with eight working hours per day (Young et al., 2013).
- **Current façade** (base case): double glazed clear insulating glass is assumed for the base case. The declared unit for processed glass is 1 m² of glass. Each 1 m² of double-pane insulating glass is 15.6 kg. The thickness of each pane is 3 mm (see Fig. 1).
- **Alternative 1:** passive Low-E coating. The declared unit for processed glass is one m² of glass. Each 1 m² pane weighs 7.5 kg, assuming a glass density of 2,500 kg/m³ and thickness of 3 mm. For a double unit, the specific weight is 15 kg (see Fig. 1).
- **Alternative 2:** electrochromic insulating glass unit. The declared unit for each with 1 m² of triple-pane glass is 17 kg. The thickness of each pane is provided in Fig. 3.
- **Alternative 3:** semi-transparent BIPV. The functional unit for LCA of BIPVs is 2 m² and 19 kg. Due to the lack of compiled data from the manufacturing companies, the U.S. Energy Information Administration (EIA) was used.

Furthermore, the degradation rate was assumed to be 0.7% per year. Transportation of all alternatives was considered 1000 km to the site (for example, Germany) by long trucks and 300 km for machinery on-site.

2.1.2 System boundaries

LCA's system boundary has been considered from cradle to site, and replacement contains the product stage and construction process stage, and use stage (A1-A5, B4, and B6). All alternatives are assumed

to have the same production process as clear glass. Material extraction, manufacture, transportation, and on-site installation are part of the glazing system's pre-use phase. Only the operational processes were addressed during the usage phase. In this work, the service life of all alternatives is considered 60 years. Acidification, eutrophication, Global Warming Potential (GWP), Non-Renewable Energy (NRE), Ozone layer Depletion (OD), and terrestrial acidification and/or nitrification are the six impact categories studied because of their significance where (Elkhatay et al., 2020) define for each as the following.

2.1.3 Use-phase energy inputs

Grasshopper® plug-in is used to calculate BIPV's power generation. The simulations showed that the possible annual power generation is about 3,654 MWh/yr with a maximum efficiency of 19%. Alternatively, the energy generation in the building's use phase can be estimated by simple calculations. In the first step, finding approximate sunny hours in London is necessary. The monthly average daily sunny hours for five consecutive years were used to achieve this approximation. The average sunny hour is 4.4 hours per day. In calculations, 4 hours of sunlight per day is assumed. Considering 250 watt LG panel, each panel's nominal daily power generation can be calculated as 1 kWh/day (250 watts x 4.4 hours). Considering 75,000 m² usable façade area to install 2x1 meter panels, the annual possible power generation with 17% efficiency is about 2,556 MWh per year which is considerably lower than the results from the Grasshopper® simulation. In the LCA calculation, the power calculated by Grasshopper® has been employed. On the other hand, the Grasshopper-Honeybee® plug-in was utilized to estimate the average annual energy demand of the Leadenhall building.

2.1.4 Assumptions and limitations

The assumptions used in this comparative LCA analysis were chosen to simplify assessing environmental impacts and directly compare the three alternatives with the base case (clear double glazing system). The lifespan of buildings and transportation distances were considered to be constant. It is also assumed that all glazing systems are manufactured with the same process and location. As mentioned earlier, only the outer skin of the double façade was studied with the same framing in all alternatives. The dimensions of all glazing systems are considered the same. The materials with no effect on the LCA were not considered. Tab. 1 summarizes each of the scope definitions and assumptions.

Table 1: Summary of scope definition and assumptions

Functional unit	48-story (225 m) commercial high-rise in London with 57,000 m ² heated floor, 75,000 m ² façade area, and 60 years of service life.
System boundary	Pre-use and use stages of the life cycle and associated transportation, including A1-A5, B4, and B6 Environmental inputs: natural resources and energies Environmental outputs: emissions
Impact categories	Acidification, eutrophication, global warming potential, nonrenewable energy, ozone layer depletion, and terrestrial acidification/nitrification
Tools	ZEB Tool, Rhino®, Grasshopper® (Honeybee® and Ladybug® plug-ins)
Major assumptions	Assuming the same framing for all alternatives. Leadenhall building has a double-skin façade. Therefore, it was assumed that in all alternatives, the inner façade is the same, and only the glass of the outer façade is changing. The degradation rate was assumed to be 0.7% per year. BIPV panel dimension 2x1 meter with an efficiency of 17 %. Energy savings of Low-E and smart glass in the use phase were considered 5%, and 10%, respectively.
Distances/methods of transportation	1,000 km to the site for all glazing systems by long trucks. 300 km for machinery on-site.

Table 2: Sources of used EPD data

Alternatives	GWP (kgCO _{2eq} /m ²)	Sources
Double glazing	39.6	ASTM-EPD149
Low-E	13.7	ASTM-EPD149
Smart glass	7.43E+01	4787287780.101.1 Sage Glass
BIPV	2.79E+02	S-P-01817 Userhuus

2.2 Inventory Analysis

This inventory analysis phase analyzes and quantifies the environmental inputs and outputs associated with glazing system scenarios, considering the FU system limits and assumptions. The input energy, raw materials, output emissions, and solid waste for each stage of the life cycle of the scenario were analyzed. ZEB Tool's database was considered the primary source. ZEB Tool delivers extensive global and regional datasets that allow flow models to reflect integrated supply chains accurately and account for differences in the nation's glazing system.

2.2.1 Material inputs

The life cycle of glazing systems was assessed, starting with the extraction of raw materials (the "cradle"), followed by the use phase. Because float glass is the primary component of every Insulating Glass Unit (IGU), all glazing systems begin with the same production process. Then, specific manufacturing processes for each glazing system are used to create the IGU, mounted on a building façade at the end. Due to the comparative objective, the materials without effect on the LCA were not considered.

2.2.2 EPDs and Databases for LCI

The alternative EPDs listed in Tab. 2 were manually inserted into the ZEB Tool database. The functional unit and scope for all alternatives are 1 m² and A1-A3, respectively.



Figure 2: The Leadenhall Building (right photo: (Esper, 2007), and the left photo: (Howarth, 2014))

2.3 The case study

The Leadenhall Building (Fig. 2) is an iconic skyscraper in London that is 225 meters (738 ft) tall and 48 stories. It was launched in July 2014 and was

designed by Rogers Stirk Harbour and Partners. Since this project's sightline is never in shadows, the morphology of Leadenhall could be a proper case study for this paper's purpose (Krolikowski et al., 2017). The general information about the building is elucidated in Tab. 3.

Table 3: The Leadenhall Building's overall information

Alternative name:	Cheesegrater
Location:	London, UK
Building function:	Office
Height:	225 m (48 stories)
Heated floor area:	57,000 m ²
Floors below ground:	4
Gross internal area:	84,424 m ²
Façade area	75,000 m ²

3 Results and discussion

In the present study, ZEB Tool is used to assess environmental impacts. Demand for lighting and equipment was set based on projected realistic use for a normalized operation period. The results show a correspondence between calculated and measured energy for the double-glazing, smart glass, and Low-E. However, BIPV performed differently by delivering 3,654 MWh/yr. However, there is considerable uncertainty in the estimated power generation. Furthermore, some factors have not been fully verified and may not be reliable for concrete conclusions. To optimize the energy use and delivered energy, several adjustments have already been made, such as:

- The BIPV market reports a variety of efficiencies between 5-19%. This study has chosen a relatively high-performance panel to get a more acceptable and feasible result.
- Considering the amount of energy saved during the operation of the ZEB Tool for Low-E and Smart Glass, this energy has been considered the on-site power production in the simulation.

Fig. 3 illustrates the comparison of GWP for each alternative, and by far, the highest belongs to BIPV with more than three times greater than others. At the same time, other alternatives with different specifications have almost identical impacts. However, there is a positive correlation between the technologies used in glazing systems and the total GWPs. Fig. 4 provides a phase-wise comparison of emissions in all the alternatives. Most other options' emissions come from A1-A3 (production) and B4 (replacement), and emissions in other phases are negligible.

Emissions related to material extraction and production, including related materials, are illustrated in Fig. 5. As seen, the CO₂ emission of the BIPV is significantly higher than other alternatives. Because BIPV has a higher specific weight and more components than other alternatives. The same

information for the construction phase is provided in Fig. 6, indicating that the carbon emission of transportation in all alternatives is higher than installation in the construction phase. This emission showed a correlation with the weight of their materials.

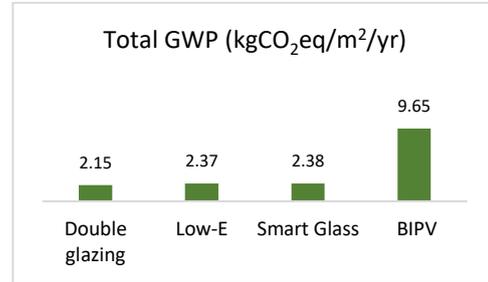


Figure 3: Total GWP for different alternatives

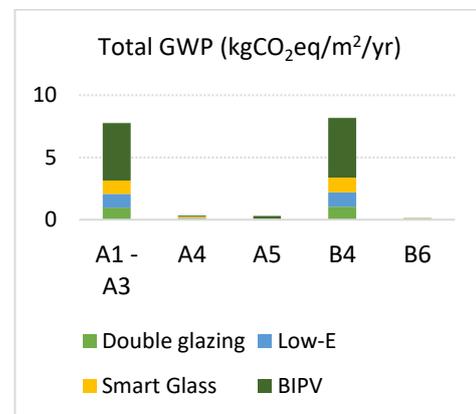


Figure 4: Total GWP per Step for all alternatives

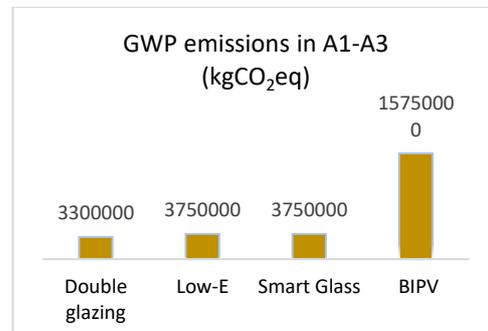


Figure 5: Production step's GWP CO₂e emissions

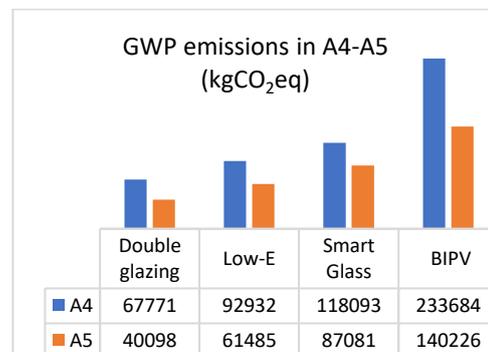


Figure 6: Construction step's GWP CO₂e emissions

3.1 Energy Demand

According to initial energy analyses in Grasshopper® with Honeybee® plug-in, the Leadenhall Energy Usage Intensity (EUI) calculated 62.5 Wh per square meter. The current building's average annual energy consumption is 9816000 (62.5 (Wh) x 8 (working hours) x 340 (working day)) kWh/yr for all different purposes such as lighting, cooling, heating, and electricity for electrical equipment. It is assumed that the total energy demand includes 18% for lighting, 18% for heating, 4 % for cooling, and 60% for electricity. Using the Grasshopper-Ladybug® plug-in, the daily radiation (kWh/m²/day) in a vertical surface (façade) has been calculated. As discussed, the Ladybug® plug-in result has shown that about 37,500 panels can provide 37.2% of the building's annual energy. It should be noted that each panel's surface area modeled in Rhino® is 2 square meters (2x1 panels). As presented in Fig. 7, in a building with smart glass (electrochromic), the energy demand is the lowest because it allows occupants to benefit from natural light without suffering from glare or heat. In contrast, BIPV consumes the highest energy due to less radiance and natural light, forcing occupants to have more lighting and heating energy. On the other hand, Low-E positively affects energy demand compared to fully transparent double glazing systems. As discussed earlier, only BIPV has this advantage in energy production and produces a considerable amount of energy (37% of its energy demand) on site. The negative values for smart glass and Low-E represent the energy saving effect in the building. Smart glazing technology performs more efficiently than other alternatives.

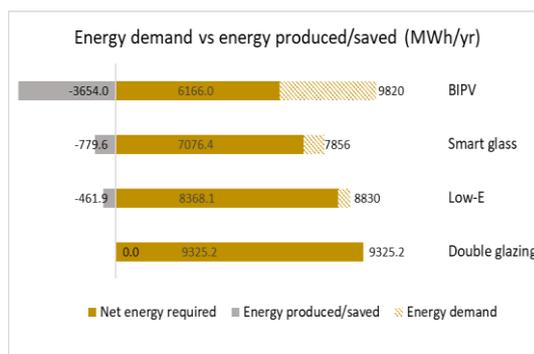


Figure 7: Leadenhall building's comparison of energy demand and production/saving in different alternatives

Fig. 7 proves that a conventional glazing system needs the highest energy compared to other alternatives. Because not only does it have the second-highest energy demand, but also it has no savings or energy production. BIPV shows the best performance in total energy demand due to its significant electric energy generation, and the building's net energy for one year of operation is the lowest.

More detail, the proportion of electricity from the building's energy demand in all alternatives is the highest (more than 50% of their total demand). Because the electrical consumption of the equipment is independent of the glazing type, BIPV passes the lowest sunlight among all alternatives. Therefore, as one of the main disadvantages, this scenario needs considerably higher energy demand for lighting. While in other scenarios, there is a minor deviation. According to the result, around 40% saving in the annual energy demand of the Leadenhall building can be achieved. Moreover, due to a lower U-value than a regular glazing system, the BIPV leads to substantial energy savings and reduced pollutant sources (Olivieri et al., 2014). However, other negative aspects include a lack of aesthetic added value, radiation loss, and a darker view.

Therefore, in the continuation of this study, the LCA of the proposed scenario (hybrid scenario) using a multi-parameter optimization process to decrease the negative aspects of BIPV from an architectural point of view is studied.

3.2 Hybrid Scenario

There is a possibility to optimize the BIPV façade and reduce the disadvantages. Some BIPVs were replaced by clear double glazing in the proposed hybrid scenario. As seen in Fig. 8, dividing each floor's glazing into three sections, the middle section of the BIPV scenario can be replaced by a clear double glazing window. Therefore, the view at eye level is preserved in the hybrid scenario. In most office spaces, desks mostly block the lower section, which causes a minor problem for people. This LCA and energy demand approach shows almost 30 % less carbon emission in the product and use stage. It also produced 70% of the energy, generating full BIPV glazing. Fig. 9 compares the GWP of the hybrid scenario with other alternatives.

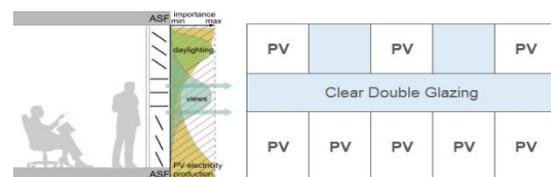


Figure 8: left: the conceptual cross-section of BIPV and clear glazing glass in eye-line borrowed from (Nagy et al., 2016) , and right: suggested hybrid scenario configuration arrangement.

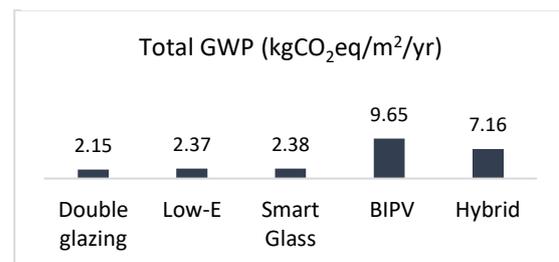


Figure 9: Total GWP for different alternatives

4 Conclusion

The present study investigates the LCA of four scenarios representing the most widely used commercial high-potential glazing systems. The conventional double glazing system showed the lowest GWP in the LCA study. The energy analysis was carried out to understand how these glazing systems' energy savings in the use phase could offset the initial energy consumption in the pre-use step for raw materials extraction, transportation, production, and installation. Due to the generated electricity, the BIPV achieved a 37% saving in required total energy, which puts it first as the most energy-saving system. A hybrid glass construction system has been proposed due to some negative aspects of using BIPVs from an architectural point of view. Although it can generate 30% less energy than the full BIPV scenario, it covers aesthetic concerns in this system and compensates for 27% of the total energy demand of the project. However, Low-E and smart glass consumed more energy in the pre-use phase than the conventional glazing system; their energy savings in the use phase (5%, 10%) could offset these initial consumptions to achieve lower total energy consumption values. This study had limitations due to the lack of technical data from suppliers and software databases. As a future study, it is suggested that applying different climates, considering economic aspects, extending boundaries, and assessing the environmental impacts under uncertainty would be appropriate starting points.

References

Ahmad, T., Aibinu, A., & Thaheem, M. J. (2017). The Effects of High-rise Residential Construction on Sustainability of Housing Systems. *Procedia Engineering*, 180, 1695–1704. <https://doi.org/10.1016/j.proeng.2017.04.332>

Arasteh, D. (1995). Advances in Window Technology: 1973-1993. *Advances in Solar Energy, An Annual Review of Research and Development*.

Arzoumanidis, I., D'Eusano, M., Raggi, A., & Petti, L. (2019). *Functional Unit Definition Criteria in Life Cycle Assessment and Social Life Cycle Assessment: A Discussion* (pp. 1–10). https://doi.org/10.1007/978-3-030-01508-4_1

Asif, M. (2019). An empirical study on life cycle assessment of double-glazed aluminium-clad timber windows. *International Journal of Building Pathology and Adaptation, ahead-of-print*. <https://doi.org/10.1108/IJBPA-01-2019-0001>

Azari, R. (2014). Integrated energy and environmental life cycle assessment of office building envelopes. *Energy and Buildings*, 82, 156–162. <https://doi.org/10.1016/j.enbuild.2014.06.041>

Baetens, R., Jelle, B. P., & Gustavsen, A. (2010). Properties, requirements and possibilities of smart windows for dynamic daylight and solar energy control in buildings: A state-of-the-art review. *Solar Energy Materials and Solar Cells*, 94(2), 87–105. <https://doi.org/10.1016/j.solmat.2009.08.021>

Booth, R. (2014, August 13). Inside the Cheesegrater – London's latest skyscraper. *The Guardian*. <https://www.theguardian.com/business/2014/aug/13/london-office-evolution-lloyds-leadenhall-cheesegrater>

Bravo-Hidalgo, D., & Baez-Hernandez, A. (2019). Technologies of zero energy buildings. A review. *Ingeniería y*

Competitividad, 21(2). <https://www.redalyc.org/journal/2913/291362343004/html/>

Buelow-Huebe, H. (2001). *Energy-efficient window systems. Effects on energy use and daylight in buildings*. <https://www.osti.gov/etdweb/biblio/20226371>

Buratti, C., & Moretti, E. (2012). Experimental performance evaluation of aerogel glazing systems. *Applied Energy*, 97, 430–437. Scopus. <https://doi.org/10.1016/j.apenergy.2011.12.055>

Cao, X., Dai, X., & Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings*, 128, 198–213. <https://doi.org/10.1016/j.enbuild.2016.06.089>

Chang, C.-Y., & Chen, P.-K. (2005). Human Response to Window Views and Indoor Plants in the Workplace. *HortScience*, 40(5), 1354–1359. <https://doi.org/10.21273/HORTSCI.40.5.1354>

Cui, Q., & Shao, F. (2022). Comparison of life cycle assessment for laminating and glazing processes based on simapro. *Pigment & Resin Technology, ahead-of-print*(ahead-of-print). <https://doi.org/10.1108/PRT-09-2021-0116>

Cupelli, D., Nicoletta, F. P., Manfredi, S., De Filipo, G., & Chidichimo, G. (2009). Electrically switchable chromogenic materials for external glazing. *Solar Energy Materials and Solar Cells*, 93(3), 329–333. Scopus. <https://doi.org/10.1016/j.solmat.2008.11.010>

Deng, Y., Li, S.-Q., Yang, Q., Luo, Z.-W., & Xie, H.-L. (2021). High-Efficiency Responsive Smart Windows Fabricated by Carbon Nanotubes Modified by Liquid Crystalline Polymers. *Crystals*, 11(4), 440. <https://doi.org/10.3390/cryst11040440>

Edlich, R., Winters, K. L., Cox, M. J., Becker, D. G., Horowitz, J. H., Nichter, L. S., Britt, L. D., Iii, W. B. L., & Edlic, E. C. (2004). Use of UV-Protective Windows and Window Films to Aid in the Prevention of Skin Cancer. *Journal of Long-Term Effects of Medical Implants*, 14(5). <https://doi.org/10.1615/JLongTermEffMedImplants.v14.i5.70>

Elkhatay, Y. O., Ibrahim, M. G., Tokimatsu, K., & Ali, A. A. M. (2020). A comparative life cycle assessment of three high-performance glazing systems for office buildings in a hot desert climate zone. *Clean Technologies and Environmental Policy*, 22(7), 1499–1515. <https://doi.org/10.1007/s10098-020-01891-2>

Ferrara, M., Castaldo, A., Esposito, S., D'Angelo, A., Guglielmo, A., & Antonaia, A. (2016). AlN–Ag based low-emission sputtered coatings for high visible transmittance window. *Surface and Coatings Technology*, 295, 2–7. <https://doi.org/10.1016/j.surfcoat.2015.12.015>

Ghosh, A., Norton, B., & Duffy, A. (2015). Measured overall heat transfer coefficient of a suspended particle device switchable glazing. *Applied Energy*, 159, 362–369. Scopus. <https://doi.org/10.1016/j.apenergy.2015.09.019>

He, Q., Hossain, M., Ng, S., & Skitmore, M. (2019). Energy-Efficient Window Retrofit for High-Rise Residential Buildings in Different Climatic Zones of China. *Sustainability*. <https://doi.org/10.3390/su11226473>

Hee, W. J., Alghoul, M. A., Bakhtyar, B., Elayeb, O., Shameri, M. A., Alrubaih, M. S., & Sopian, K. (2015). The role of window glazing on daylighting and energy saving in buildings. *Renewable and Sustainable Energy Reviews*, 42(C), 323–343.

Hernandez, P., Oregi, X., Longo, S., & Cellura, M. (2019). Chapter 4—Life-Cycle Assessment of Buildings. In F. Asdrubali & U. Desideri (Eds.), *Handbook of Energy Efficiency in Buildings* (pp. 207–261). Butterworth-Heinemann. <https://doi.org/10.1016/B978-0-12-812817-6.00010-3>

Horup, L., Reymann, M., Rørbech, J., Ryberg, M., & Birkved, M. (2019). Partially dynamic life cycle assessment of windows indicates potential thermal over-optimization. *IOP Conference Series: Earth and Environmental Science*, 323, 012152. <https://doi.org/10.1088/1755-1315/323/1/012152>

- IEA – International Energy Agency. (n.d.). IEA. Retrieved 1 March 2022, from <https://www.iea.org>
- ISO. (2006a). *14040:2006 Environmental management-life cycle assessment-principles and framework*. International Standards Organization.
- ISO. (2006b). *14044:2006 Environmental management-life cycle assessment-requirements and guidelines*. International Standards Organization.
- Jonsson, A., & Roos, A. (2010). Visual and energy performance of switchable windows with antireflection coatings. *Solar Energy*, *84*(8), 1370–1375. Scopus. <https://doi.org/10.1016/j.solener.2010.04.016>
- Kenisarin, M., & Mahkamov, K. (2016). Passive thermal control in residential buildings using phase change materials. *Renewable and Sustainable Energy Reviews*, *55*, 371–398. Scopus. <https://doi.org/10.1016/j.rser.2015.10.128>
- Krolkowski, D., ELEY, D., Gramazio, F., Kohler, M., & Langenberg, S. (2017). THE LEADENHALL BUILDING: DESIGN FOR FABRICATION-DIGITAL WORKFLOW AND DOWNSTREAM FABRICATION SYSTEM. In *Fabricate 2014* (DGO-Digital original, pp. 68–75). UCL Press. <https://doi.org/10.2307/j.ctt1tp3c5w.12>
- Lee, J. W., Jung, H. J., Park, J. Y., Lee, J. B., & Yoon, Y. (2013). Optimization of building window system in Asian regions by analyzing solar heat gain and daylighting elements. *Renewable Energy*, *50*, 522–531. <https://doi.org/10.1016/j.renene.2012.07.029>
- Li, D. H. W., Lam, T. N. T., Chan, W. W. H., & Mak, A. H. L. (2009). Energy and cost analysis of semi-transparent photovoltaic in office buildings. *Applied Energy*, *86*(5), 722–729. Scopus. <https://doi.org/10.1016/j.apenergy.2008.08.009>
- Li, Z., Zhang, W., Xie, L., Wang, W., Tian, H., Chen, M., & Li, J. (2021). Life cycle assessment of semi-transparent photovoltaic window applied on building. *Journal of Cleaner Production*, *295*, 126403. <https://doi.org/10.1016/j.jclepro.2021.126403>
- Luong, D. L., Nguyen, Q. T., Pham, A. D., Truong, Q. C., & Duong, M. Q. (2020). Building a Decision-Making Support Framework for Installing Solar Panels on Vertical Glazing Façades of the Building Based on the Life Cycle Assessment and Environmental Benefit Analysis. *Energies*, *13*(9), 2376. <https://doi.org/10.3390/en13092376>
- Lyu, Y., & Chow, T. (2020). Economic, energy and environmental life cycle assessment of a liquid flow window in different climates. *Building Simulation*, *13*. <https://doi.org/10.1007/s12273-020-0636-z>
- Megange, P., Ngae, P., Feiz, A.-A., Melhaoui, A., Chpoun, A., & Le, T.-P. (2018). Dynamic Life Cycle Assessment of a Double Glazing Bay. *6th International Renewable and Sustainable Energy Conference (IRSEC 2018)*, 1–5. <https://doi.org/10.1109/IRSEC.2018.8702852>
- Midtdal, K., & Jelle, B. P. (2013). Self-cleaning glazing products: A state-of-the-art review and future research pathways. *Solar Energy Materials and Solar Cells*, *109*, 126–141. Scopus. <https://doi.org/10.1016/j.solmat.2012.09.034>
- Minne, E., Wingrove, K., & Crittenden, J. C. (2015). Influence of climate on the environmental and economic life cycle assessments of window options in the United States. *Energy and Buildings*, *102*, 293–306. <https://doi.org/10.1016/j.enbuild.2015.05.039>
- Nielsen, T., Duer, K., & Svendsen, S. (2001). Energy performance of glazings and windows. *Solar Energy*, *69*, 137–143. [https://doi.org/10.1016/S0038-092X\(01\)00062-7](https://doi.org/10.1016/S0038-092X(01)00062-7)
- Olivieri, L., Caamaño-Martín, E., Moralejo-Vázquez, F. J., Martín-Chivelet, N., Olivieri, F., & Neila-Gonzalez, F. J. (2014). Energy saving potential of semi-transparent photovoltaic elements for building integration. *Energy*, *76*, 572–583. <https://doi.org/10.1016/j.energy.2014.08.054>
- O'Neill, R., Window, A., Kenway, S., & Dargusch, P. (2021). Integrated operational and life-cycle modelling of energy, carbon and cost for building façades. *Journal of Cleaner Production*, *286*, 125370. <https://doi.org/10.1016/j.jclepro.2020.125370>
- Park, H. K., & Kim, H. (2015). Acoustic insulation performance of improved airtight windows. *Construction and Building Materials*, *93*, 542–550. Scopus. <https://doi.org/10.1016/j.conbuildmat.2015.05.058>
- Parkin, I. P., & Manning, T. D. (2006). Intelligent thermochromic windows. *Journal of Chemical Education*, *83*(3), 393–400. Scopus. <https://doi.org/10.1021/ed083p393>
- Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings*, *40*(3), 394–398. <https://doi.org/10.1016/j.enbuild.2007.03.007>
- Rezaei, S. D., Shannigrahi, S., & Ramakrishna, S. (2017). A review of conventional, advanced, and smart glazing technologies and materials for improving indoor environment. *Solar Energy Materials and Solar Cells*, *159*, 26–51. <https://doi.org/10.1016/j.solmat.2016.08.026>
- Rosencrantz, T., Bülow-Hübe, H., Karlsson, B., & Roos, A. (2005). Increased solar energy and daylight utilisation using anti-reflective coatings in energy-efficient windows. *Solar Energy Materials and Solar Cells*, *89*(2–3), 249–260. Scopus. <https://doi.org/10.1016/j.solmat.2004.12.007>
- Saadatian, S., Simões, N., & Freire, F. (2021). Integrated environmental, energy and cost life-cycle analysis of windows: Optimal selection of components. *Building and Environment*, *188*, 107516. <https://doi.org/10.1016/j.buildenv.2020.107516>
- Souviron, J., van Moeseke, G., & Khan, A. Z. (2019). Analysing the environmental impact of windows: A review. *Building and Environment*, *161*, 106268. <https://doi.org/10.1016/j.buildenv.2019.106268>
- Stegou-Sagia, A., Antonopoulos, K., Angelopoulou, C., & Kotsiovelos, G. (2007). The impact of glazing on energy consumption and comfort. *Energy Conversion and Management*, *48*(11), 2844–2852.
- Taghizade, K., Heidari, A., & Noorzai, E. (2019). Environmental Impact Profiles for Glazing Systems: Strategies for Early Design Process. *Journal of Architectural Engineering*, 25.
- Teixeira, H., Gomes, M. G., Moret Rodrigues, A., & Pereira, J. (2020). Thermal and visual comfort, energy use and environmental performance of glazing systems with solar control films. *Building and Environment*, *168*, 106474. <https://doi.org/10.1016/j.buildenv.2019.106474>
- Tuchinda, C., Srivannaboon, S., & Lim, H. W. (2006). Photoprotection by window glass, automobile glass, and sunglasses. *Journal of the American Academy of Dermatology*, *54*(5), 845–854. Scopus. <https://doi.org/10.1016/j.jaad.2005.11.1082>
- Tushar, Q., Bhuiyan, M. A., & Zhang, G. (2022). Energy simulation and modeling for window system: A comparative study of life cycle assessment and life cycle costing. *Journal of Cleaner Production*, *330*, 129936. <https://doi.org/10.1016/j.jclepro.2021.129936>
- Urbikain, M. K., & Sala, J. M. (2009). Analysis of different models to estimate energy savings related to windows in residential buildings. *Energy and Buildings*, *41*(6), 687–695. <https://doi.org/10.1016/j.enbuild.2009.01.007>
- Wittwer, V., Datz, M., Ell, J., Georg, A., Graf, W., & Walze, G. (2004). Gasochromic windows. *Solar Energy Materials and Solar Cells*, *84*(1–4), 305–314. Scopus. <https://doi.org/10.1016/j.solmat.2004.01.040>
- Young, A., Harbour, R. S., Annereau, N., Butler, A., & Smith, B. (2013). Case Study: The Leadenhall Building, London. *CTBUH Journal*.

Density and viscosity correlations of aqueous solvents for amine-based CO₂ capture

Sumudu Karunaratne*, John Ikechukwu Okoro, Saroj Neupane, Lars Erik Øi

Faculty of Technology, Natural Sciences and Maritime Studies, University of South-Eastern Norway, Norway,
sumudu.karunaratne@usn.no

Abstract

Density and viscosity data are essential in designing process equipment and process simulations in the amine-based CO₂ capture process. In literature, semi-empirical and empirical correlations for density and viscosity were fitted to measured data available in the literature and the goodness of fit by calculating Average Absolute Relative Deviations (AARDs %) were examined. The correlations based on excess properties give insights on intermolecular interactions and packing efficiency in multicomponent liquid mixtures.

In this study, correlations for density, excess density (ρ^E , $\ln(\rho^E)$) viscosity and excess viscosity (η^E , $\ln(\eta^E)$) were examined for binary aqueous mixtures (Monoethylethanolamine) EMEA (1) + H₂O (2) and (Monoethylethanolamine) EMEA (1) + (Diethylethanol amine) DEEA (2) to represent the measured density and viscosity. The Redlich and Kister type polynomials were used to fit the excess properties and the goodness of fit was determined by calculating AARD%. The fitted experimental data at different mole fractions and temperatures were able to acquire a good accuracy. Highest deviation for density correlations of EMEA + DEEA mixtures was observed with AARD 0.135 % and AMD 2.98 kg·m⁻³. For the EMEA + H₂O mixtures, the highest deviation was observed with AARD 0.23 % and AMD 8.38 kg·m⁻³. Viscosity correlations showed a highest deviation for the EMEA + DEEA mixtures in which AARD was 5.2 % and AMD was 1.2 mPa·s. For the EMEA + H₂O mixtures, the highest deviation observed with AARD was 1.1% and AMD was 0.75 mPa·s.

McAllister's kinematic viscosity model is a semi-empirical model based on Eyring's theory for viscosity. The kinematic viscosity data were fitted to McAllister's three-body model to investigate whether the suggested intermolecular interactions in the model are capable of describing the nature of the binary mixtures. The results indicated a good agreement between data and model with accuracies $R^2 = 0.99$ and AARD 2% for EMEA + DEEA mixtures and AARD 6% for EMEA + H₂O mixtures.

Keywords: correlations, density, viscosity

1. Introduction

Amine-based CO₂ capture is a matured technology that has been adopted to remove CO₂ and H₂S from natural gas. The aqueous solvents of MEA (Monoethanolamine), MDEA (Methyldiethanolamine), and DEA (Diethanolamine) are proven solvents for gas purification. There are advantages and disadvantages of using these amines, and none of them are perfect to use in post-combustion CO₂ capture. Solvents with characteristics such as higher CO₂ loading capacity and lower heat of reaction enable the applicability of using this technology for post-combustion CO₂ capture. As a result, new amines and blends of existing amines are continuously examined to find a solvent that can give higher CO₂ capture efficiency at a low rate of use of energy. The physicochemical properties are measured to support further research in CO₂ capture.

2. Literature

Density and viscosity data are essential in designing process equipment and performing process simulations in the amine-based CO₂ capture process. Semi-empirical and empirical correlations for density and viscosity were fitted to measured data that are available in literature and the goodness of the fit was examined by calculating relative deviation. The correlations based on excess properties give insights into intermolecular interactions and packing efficiency in multicomponent liquid mixtures. In a binary liquid mixture, excess molar volume V^E rises mainly due to the difference in intermolecular interaction between unlike molecules compared to the intermolecular interaction present in pure components, and differences in size and shapes of molecules that affect the packing efficiency in the mixture (Mahajan and Mirgane 2013). The positive

deviation for V^E are responsible for having weak dipole-dipole interactions among the unlike molecules while negative deviations are responsible for charge transfer, H-bonds and other complex forming interactions. The negative V^E can also be a result of the fitting of component molecules of the mixture due to the structural differences of size and shape. The viscosity deviation or excess viscosity η^E calculated using [measured] mixture and pure liquid viscosities can provide similar information as V^E for the intermolecular interactions (Fort and Moore 1966). There, positive η^E indicates stronger intermolecular interactions like H-bonds while negative η^E indicates weak dipolar interactions.

Eyring's viscosity model, which is based on Eyring's theory of absolute reaction rate provides a mechanism to explain the viscosity of liquids (Eyring 1936). The excess free energy of activation for viscous flow ΔG^{E*} defined for mixtures provides information of the nature of intermolecular interactions similar to V^E and η^E (Meyer *et al.*, 1971) in which positive ΔG^{E*} indicates stronger intermolecular interactions like H-bonds while negative ΔG^{E*} indicates weak dipolar interactions. McAllister (1960) developed a model to predict kinematic viscosity using Eyring's approach for binary mixtures. There different types of molecular interactions were considered in the model development. The relevant parameters in model can be estimated by regression.

Correlations based on Redlich-Kister type polynomial (Redlich and Kister 1948) is a commonly used approach to represent measured density and viscosity for binary mixtures. The calculated excess properties such as excess molar volume V^E and excess viscosity η^E are fitted to a Redlich-Kister type polynomial and relevant parameters are found by minimizing the deviation between predictions and data. Karunarathne *et al.* (2020) reviewed the developed correlations for MEA (Monoethanolamine) + H₂O + CO₂ mixture and discussed the accuracies of predictions by comparing different data sets.

This study discusses several approaches to develop correlations for density and viscosity of Monoethylethanolamine EMEA (1) + Diethylethanol amine DEEA (2) and Monoethylethanolamine EMEA (1) + H₂O (2) mixtures. The correlations were evaluated for the accuracy of data fit by calculating different statistical quantities such as average absolute relative deviation AARD and absolute maximum deviation AMD. Measured densities and viscosities for EMEA + DEEA and EMEA + H₂O mixtures for this study were taken from published data by Chen *et al.* (2016).

3. Methodology

Several approaches have been considered for the fitting of measured data to density and viscosity correlations. For the density, [four corrections] were considered as shown from (1) to (6). Equation (1) is an analogous correlation proposed for density based on the viscosity correlation proposed by Heric and Brewer (1967). The parameter A_i is further correlated to the temperature of the mixture through a linear relation of $A_i = a_{i1} + a_{i2}T$. Equation (2) was proposed by Aronu *et al.* (2012) to represent the density of aqueous amino acid salt and amine amino acid salt solutions. Equations (3) and (4) were based on the excess density ρ^E in which excess properties are represented by Redlich and Kister type polynomials with temperature dependent parameters. Equations (5) and (6) considered natural logarithm of excess density $\ln(\rho^E)$ and Redlich and Kister type polynomials with temperature dependent parameters were adopted to fit the calculated $\ln(\rho^E)$ as described in (6).

$$\ln(\rho) = x_1 \ln(\rho_1) + x_2 \ln(\rho_2) + x_1 \ln(M_1) + x_2 \ln(M_2) - \ln(x_1 M_1 + x_2 M_2) + x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (1)$$

$$\rho = \left(A_0 + \frac{A_1 x_2}{T} \right) \exp \left(\frac{A_2}{T^2} + \frac{A_3 x_1}{T} + A_4 \left(\frac{x_1}{T} \right)^2 \right) \quad (2)$$

$$\rho^E = \rho - \sum_{i=1}^n x_i \rho_i \quad (3)$$

$$\rho^E = x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (4)$$

$$\ln(\rho^E) = \ln(\rho) - \sum_{i=1}^n x_i \ln(\rho_i) \quad (5)$$

$$\ln(\rho^E) = x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (6)$$

Four correlations for dynamic viscosity have been considered as illustrated from (7) to (12). Equation (7) is proposed for viscosity by Heric and Brewer (1967). A correlation developed by Aronu *et al.* (2012) in (8) was adopted as it does not require any pure component viscosities to calculate mixture viscosity. For (9) and (10), excess viscosity or viscosity deviation η^E was calculated from pure component viscosities and Redlich and Kister type polynomials with temperature dependent parameters were proposed to represent η^E . The natural logarithm of excess viscosity $\ln(\eta^E)$ and Redlich and Kister type polynomials with temperature

dependent parameters were adopted to fit the calculated $\ln(\eta^E)$ as described in (11) and (12). The kinematic viscosities of EMEA (1) + DEEA (2) and EMEA (1) + H₂O (2) mixtures were calculated using measured dynamic viscosities and densities. McAllister (1960) model was fitted to the range of mole fractions of x_1 from 0 to 1 and temperatures from 293.15 K to 333.15 K. Equations from (13) to (17) described the McAllister model in which the enthalpies and entropies were found using regression.

$$\ln(\eta) = x_1 \ln(\eta_1) + x_2 \ln(\eta_2) + x_1 \ln(M_1) + x_2 \ln(M_2) - \ln(x_1 M_1 + x_2 M_2) + x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (7)$$

$$\eta = \left(1 + \frac{A_0 x_1}{T} + \frac{A_1 x_1 x_2^2}{T^2} \right) \exp \left(A_2 + \frac{A_4}{T} + \frac{A_5}{T^3} + \frac{A_6 x_1 x_2}{T^4} \right) \quad (8)$$

$$\eta^E = \eta - \sum_{i=1}^n x_i \eta_i \quad (9)$$

$$\eta^E = x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (10)$$

$$\ln(\eta^E) = \ln(\eta) - \sum_{i=1}^n x_i \ln(\eta_i) \quad (11)$$

$$\ln(\eta^E) = x_1 x_2 \left[\sum_{i=0}^n A_i (x_1 - x_2)^i \right] \quad (12)$$

$$\ln(v) = x_1^3 \ln(v_1) + 3x_1^2 x_2 \ln(v_{12}) + 3x_1 x_2^2 \ln(v_{21}) + x_2^3 \ln(v_2) - \ln \left(x_1 + x_2 \frac{M_2}{M_1} \right) + 3x_1^2 x_2 \ln \left[\frac{\left(2 + \frac{M_2}{M_1} \right)}{3} \right] + 3x_1 x_2^2 \ln \left[\frac{\left(1 + \frac{2M_2}{M_1} \right)}{3} \right] + x_2^3 \ln \left(\frac{M_2}{M_1} \right) \quad (13)$$

$$v_1 = \frac{hN}{M_1} e^{-\Delta S_1^*/R} e^{\Delta H_1^*/RT} \quad (14)$$

$$v_{12} = \frac{hN}{M_{12}} e^{-\Delta S_{12}^*/R} e^{\Delta H_{12}^*/RT} \quad (15)$$

$$v_{21} = \frac{hN}{M_{21}} e^{-\Delta S_{21}^*/R} e^{\Delta H_{21}^*/RT} \quad (16)$$

$$v_2 = \frac{hN}{M_2} e^{-\Delta S_2^*/R} e^{\Delta H_2^*/RT} \quad (17)$$

$$AARD = \frac{100\%}{D} \sum_{i=1}^D \left| \frac{A_i^m - A_i^c}{A_i^m} \right| \quad (18)$$

$$AMD = \text{MAX} |A_i^m - A_i^c| \quad (19)$$

4. Results and Discussion

4.1. Density correlations

The density correlations from (1) to (6) have been fitted for the measured densities for EMEA+DEEA and EMEA + H₂O binary mixtures. Tab. 1 summarized the calculated AARD (%) and AMD (kg·m⁻³) for the different correlations. Fig. 1 and 2 illustrate the accuracy of fitting (1) and (2) into the measured densities and both correlations showed acceptable accuracies of EMEA + DEEA and EMEA + H₂O mixtures. For EMEA + H₂O, the deviations are relatively higher compared to EMEA + DEEA mixtures. The density correlation from Aronu *et al.* (2012) showed a highest AMD of 2.98 kg·m⁻³ at $x_1 = 0$ and temperature 333.15 K for EMEA + DEEA and AMD of 8.38 kg·m⁻³ at $x_1 = 0$ and temperature 293.15 K for EMEA + H₂O mixtures respectively. The correlations (3,4) and (5,6) were fitted and parameters were found at each temperature level. This led to higher fitting accuracies as listed in Tab. 1. The estimated parameters for the correlations are given in Tab. 2 and 3.

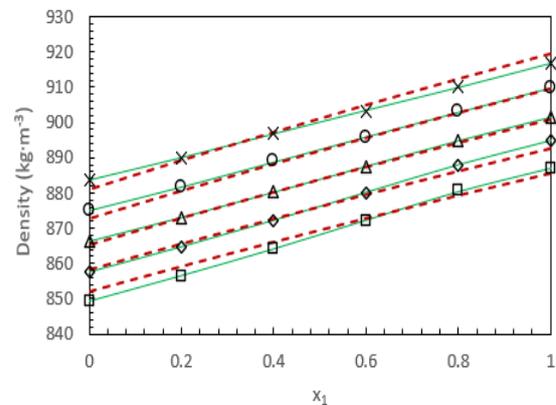


Figure 1: Density of EMEA + DEEA mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◇'; 333.15 K, '□'. Correlation: Equation (1) '—', Equation (2) '---'.

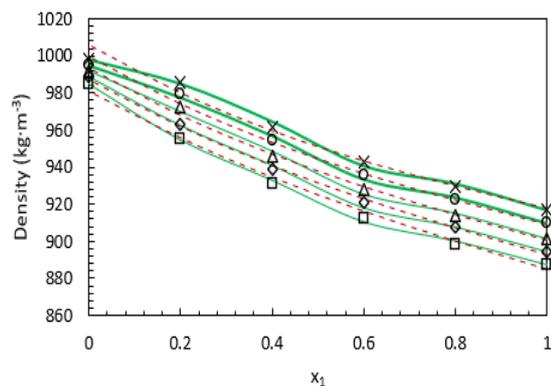


Figure 2: Density of EMEA + H₂O mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◊'; 333.15 K, '□'. Correlation: Equation (1) '—', Equation (2) '---'.

4.2. Viscosity correlations

Viscosity correlations showed from (7) to (12) were fitted into the measured viscosities of EMEA + DEEA and EMEA + H₂O mixtures. The accuracy of the fit was examined using calculated AARD (%) and AMD (mPa·s) as listed in Tab. 4. A comparison between correlation (7) and (8) for the viscosities of EMEA + DEEA mixtures is illustrated in Fig. 3. Correlation (7) was able to represent measured data with acceptable fitting accuracies as given in Tab. 4. Correlation (8) showed a relatively high AARD (%) of 5.2. The deviation between measured data and (8) is high as demonstrated in Fig. 4 and AMD is 1.2 mPa·s at $x_1 = 1$ and temperature 293.15 K. For the EMEA + H₂O mixtures, (7) gives an acceptable accuracy for the data fit as shown in Tab. 4. The correlation (8) was not appropriate for EMEA + H₂O mixtures as it gave a higher AARD (%). For the correlations (9,10) and (11,12), measured viscosities were fitted and parameters were found at each temperature level. As a result, higher accuracies were obtained as given in Tab. 4. This is a result of

the simplification of the complexity of the correlations and at the same time correlation had to be fitted into a relatively small data set compared to the previous scenarios.

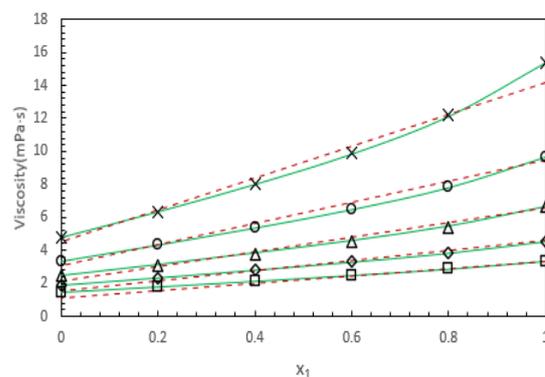


Figure 3: Viscosity of EMEA + DEEA mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◊'; 333.15 K, '□'. Correlation: Equation (7) '—', Equation (8) '---'.

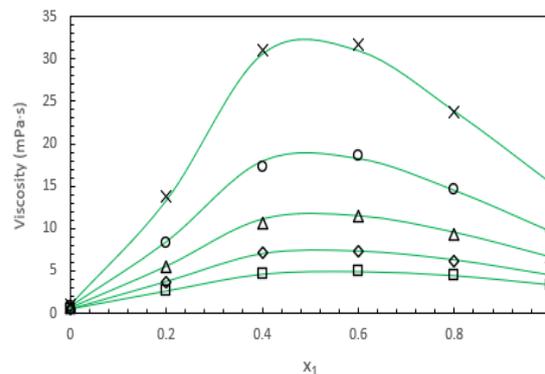


Figure 4: Viscosity of EMEA + H₂O mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◊'; 333.15 K, '□'. Correlation: Equation (7) '—'.

Table 1: Calculated AARD (%) and AMD ($\text{kg}\cdot\text{m}^{-3}$) for different density correlations and mixtures.

Correlation	AARD (%)		AMD ($\text{kg}\cdot\text{m}^{-3}$)	
	EMEA + DEEA	EMEA + H ₂ O	EMEA + DEEA	EMEA + H ₂ O
Equation (1)	0.012	0.127	0.37	3.03
Equation (2)	0.135	0.23	2.98	8.38
Equation (3) & (4)	1.85×10^{-6}	1.04×10^{-11}	1.2×10^{-6}	8.2×10^{-11}
Equation (5) & (6)	0.09	8.3×10^{-15}	9.1×10^{-5}	1.3×10^{-16}

4.3. McAllister model for kinematic viscosity

McAllister model was fitted to calculated kinematic viscosities from measured dynamic viscosity and density data for EMEA + DEEA and EMEA + H₂O mixtures for the mole fraction range from 0 to 1 and temperature from 293.15 K to 333.15 K. Fig. 5

compares the fitted McAllister model with kinematic viscosity data for EMEA + DEEA mixtures. The accuracy of the fitting is acceptable as it showed AARD (%) of 1.9 and AMD of $4.3 \times 10^{-7} \text{ m}^2\cdot\text{s}^{-1}$ at $x_1 = 1$ and temperature of 293.15 K. For the EMEA + H₂O mixtures, the deviations are higher

as AARD (%) of 5.9 and AMD of $2.8 \times 10^{-6} \text{ m}^2 \cdot \text{s}^{-1}$ at $x_1 = 0.8$ and temperature of 293.15 K. Fig. 6 compares the fitted McAllister model with kinematic viscosity data for EMEA + H₂O mixtures and local maximums for kinematic viscosities at different temperature levels were observed around mole fraction $x_1 = 0.5$.

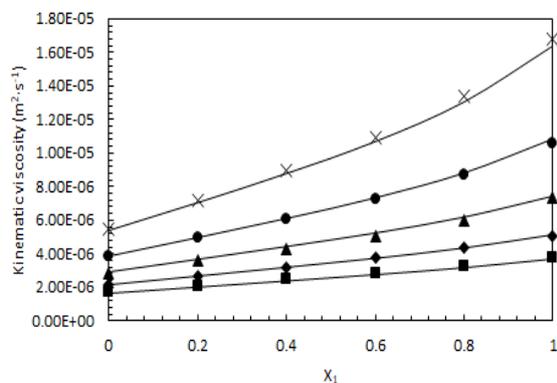


Figure 5: Kinematic viscosity of EMEA + DEEA mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◇'; 333.15 K, '□'. Correlation: Equation (13) '___'.

Tab. 7 lists the estimated enthalpies and entropies for the viscous flow from the McAllister model from

(13) to (17). This approach enables to fit the model to the complete temperature interval rather than fitting the model and estimating parameters at different temperature levels. Then this model can easily be implemented in a computer program to estimate kinematic viscosities for unmeasured mole fractions and temperatures.

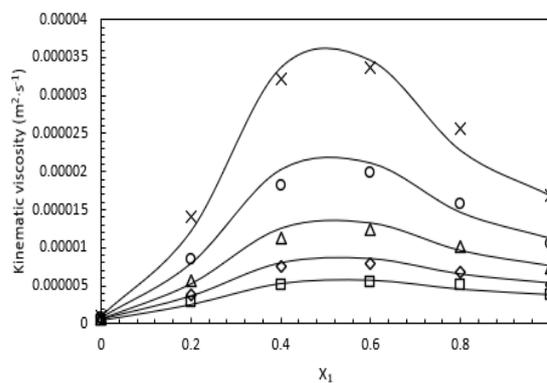


Figure 6: Kinematic viscosity of EMEA + H₂O mixtures. Measured: 293.15 K, 'x'; 303.15 K, 'o'; 313.15 K, 'Δ'; 323.15 K, '◇'; 333.15 K, '□'. Correlation: Equation (13) '___'.

Table 2: Estimated parameters for (1) and (2) for EMEA + DEEA and EMEA + H₂O mixtures.

Parameters	Equation (1)	
	EMEA + DEEA	EMEA + H ₂ O
A_0	$a_{0,1} = 0.0323$ $a_{0,2} = 1.80 \times 10^{-5}$	$a_{0,1} = 1.433$ $a_{0,2} = -9.982 \times 10^{-4}$
A_1	$a_{1,1} = -0.0472$ $a_{1,2} = 1.734 \times 10^{-4}$	$a_{1,1} = -0.9877$ $a_{1,2} = 12.1 \times 10^{-4}$
A_2	$a_{2,1} = -0.0741$ $a_{2,2} = 2.41 \times 10^{-4}$	$a_{2,1} = 0.8215$ $a_{2,2} = -14.45 \times 10^{-4}$
Parameters	Equation (2)	
	EMEA + DEEA	EMEA + H ₂ O
A_0	756.4	712.5
A_1	2334	1.08×10^5
A_2	1.22×10^4	-6221
A_3	17.04	58.13
A_4	-413.2	1.10×10^4

Table 3: Estimated parameters for (3-4) and (5-6) for EMEA + DEEA and EMEA + H₂O mixtures.

T/K	EMEA + DEEA for Equation (3-4)					
	A_0	A_1	A_2	A_3	A_4	A_5
293.15	-0.4543	0.9649	-4.3077	-220.741	1.0024	618.5851
303.15	-0.6092	-5.3242	-0.4388	23.6887	1.0971	0.9997
313.15	0.5983	3.9766	0.6330	-0.4727	0.8590	0.3681
323.15	-0.0575	3.6604	1.3813	2.1979	1.4297	1.0157
333.15	-0.1798	1.4632	4.3446	15.4171	3.7869	2.1474

T/K	EMEA + H ₂ O Equation (3-4)					
	A ₀	A ₁	A ₂	A ₃	A ₄	A ₅
293.15	-23.478	-21.8423	59.9364	-41.852	24.3164	-15.812
303.15	-31.131	-15.3706	53.4441	-37.049	21.8259	-13.9841
313.15	-43.1324	2.0342	51.4202	-55.204	21.1892	-21.4963
323.15	-52.0017	9.7800	23.5946	-9.9850	10.1867	-3.5205
333.15	-61.211	4.3467	14.7591	10.9855	-2.7181	3.4607
T/K	EMEA + DEEA for Equation (5-6)					
	A ₀	A ₁	A ₂	A ₃	A ₄	A ₅
293.15	0.01534	0.0020	-0.4258	-0.2184	1.0525	0.6052
303.15	0.0154	0.0020	-0.4236	-0.2155	1.0584	0.6132
313.15	0.0163	0.0020	-0.4106	-0.2134	1.0289	0.6116
323.15	0.0158	0.0021	-0.4152	-0.2110	1.0435	0.6116
333.15	0.01581	0.0021	-0.4130	-0.2067	1.0486	0.6236
T/K	EMEA + H ₂ O Equation (5-6)					
	A ₀	A ₁	A ₂	A ₃	A ₄	A ₅
293.15	-0.0210	-0.0163	0.0611	-0.2523	0.0247	0.5137
303.15	-0.0288	-0.0100	0.0550	-0.2478	0.0224	0.5154
313.15	-0.0412	0.0077	0.0054	-0.2674	0.0216	0.5107
323.15	-0.0505	0.0152	0.0253	-0.2121	0.01213	0.5031
333.15	-0.1783	-0.0094	3.2896	0.3112	-8.1848	-0.7458

Table 4: Calculated AARD (%) and AMD (mPa·s) for different viscosity correlations and mixtures.

Correlation	AARD (%)		AMD (mPa·s)	
	EMEA + DEEA	EMEA + H ₂ O	EMEA + DEEA	EMEA + H ₂ O
Equation (7)	0.4	1.1	0.1	0.75
Equation (8)	5.2	-	1.2	-
Equation (9) & (10)	1×10 ⁻⁶	8×10 ⁻¹²	2.5×10 ⁻⁶	1.6×10 ⁻¹⁰
Equation (11) & (12)	1×10 ⁻⁴	6×10 ⁻¹²	6.6×10 ⁻⁷	1.5×10 ⁻¹¹

Table 5: Estimated parameters for (7) and (8) for EMEA + DEEA and EMEA + H₂O mixtures.

Parameters	Equation (7)	
	EMEA + DEEA	EMEA + H ₂ O
A ₀	a _{0,1} = -0.03826	a _{0,1} = 31.37
	a _{0,2} = 7.24 × 10 ⁻⁴	a _{0,2} = -0.0742
A ₁	a _{1,1} = -1.243	a _{1,1} = -21.55
	a _{1,2} = 33.23 × 10 ⁻⁴	a _{1,2} = 0.0527
A ₂	a _{2,1} = 0.2905	a _{2,1} = 7.403
	a _{2,2} = -10.58 × 10 ⁻⁴	a _{2,2} = -0.0149
Parameters	Equation (8)	
	EMEA + DEEA	
A ₀	633.1	
A ₁	0.3369	
A ₂	-16.81	
A ₃	3343	
A ₄	0.724	
A ₅	0.0539	

Table 6: Estimated parameters for (9-10) and (11-12) for EMEA + DEEA and EMEA + H₂O mixtures.

T/K	EMEA + DEEA for Equation (9-10)			
	A ₀	A ₁	A ₂	A ₃
293.15	-4.7227	-3.0143	-1.2044	0.8790
303.15	-2.2813	-1.5235	-0.7812	0.0651
313.15	-1.6719	-1.1849	-0.9115	-0.5859
323.15	-0.5898	-0.3581	-0.3581	0.0977
333.15	-0.4102	-0.4362	-0.1628	0.4883
T/K	EMEA + H ₂ O Equation (9-10)			
	A ₀	A ₁	A ₂	A ₃
293.15	100.3320	-26.2695	-94.2383	92.6758
303.15	54.7930	-7.0247	-44.3034	32.3893
313.15	31.8516	-6.0026	-23.8932	20.3776
323.15	19.8906	-6.8229	-12.8906	20.5729
333.15	12.4609	-2.7344	-5.2734	8.9844
T/K	EMEA + DEEA for Equation (11-12)			
	A ₀	A ₁	A ₂	A ₃
293.15	0.1449	-0.2919	-0.0011	0.1276
303.15	0.1630	-0.2461	-0.02023	0.04032
313.15	0.0878	-0.2493	-0.0837	-0.0785
323.15	0.1951	-0.1457	-0.0730	0.0682
333.15	0.1650	-0.2096	-0.0276	0.2274
T/K	EMEA + H ₂ O Equation (11-12)			
	A ₀	A ₁	A ₂	A ₃
293.15	8.5296	-5.4086	2.8146	-0.7541
303.15	7.6668	-4.4286	2.6108	-1.2722
313.15	6.8749	-4.1453	2.2607	-0.9154
323.15	6.2346	-4.0936	2.0644	0.4892
333.15	5.5350	-3.3344	2.3431	-0.3038

Table 7: Estimated enthalpies and entropies for the viscous flow.

Mixture	$\Delta H^*/\text{kJ}\cdot\text{mol}^{-1}$	$\Delta S^*/\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$
EMEA + DEEA	$\Delta H_1^* = 30.25$	$\Delta S_1^* = 35.0$
	$\Delta H_{12}^* = 30.0$	$\Delta S_{12}^* = 26.5$
	$\Delta H_{21}^* = 26.4$	$\Delta S_{21}^* = 25.3$
	$\Delta H_2^* = 23.45$	$\Delta S_2^* = 18.8$
EMEA + H ₂ O	$\Delta H_1^* = 30.35$	$\Delta S_1^* = 35.04$
	$\Delta H_{12}^* = 31.6$	$\Delta S_{12}^* = 40.8$
	$\Delta H_{21}^* = 54.5$	$\Delta S_{21}^* = 97.3$
	$\Delta H_2^* = 15.4$	$\Delta S_2^* = 20.6$

5. Conclusions

This work presents developed correlations to represent measured densities and viscosities of EMEA + DEEA and EMEA + H₂O mixtures. Further, the study discusses the applicability of McAllister's kinematic viscosity model for the kinematic viscosities of EMEA + DEEA and EMEA + H₂O mixtures. Developed correlations showed acceptable accuracies in data fitting in which for the density, correlations from (3) to (6) showed a very low deviation from measured data. This is mainly because the data were fitted into the correlation at different temperature levels and the number of data points for the fitting was small compared to the fitting for (1) and (2). The highest deviation for density correlations for the EMEA + DEEA

mixtures was observed for (2) in which AARD (%) was 0.135 and AMD was 2.98 kg·m⁻³. For the EMEA + H₂O mixtures, the highest deviation was observed for (2) in which AARD (%) was 0.23 and AMD was 8.38 kg·m⁻³.

Proposed correlations for the viscosity of EMEA + DEEA and EMEA + H₂O mixtures showed acceptable accuracies for the fitting and correlations can be used in engineering applications. Equations (7) and (8) have parameters to correlate the effect of temperature on viscosity of the mixtures that do not have in (9) to (12). Equations from (9) to (12) were fitted at different temperature levels into the correction. Consequently, that led to high accuracies in data fitting. Viscosity correlations indicated the highest deviation for the EMEA + DEEA mixtures

for (8), where AARD (%) was 5.2 and AMD was 1.2 mPa·s. The McAllister model was able to represent kinematic viscosities for both EMEA + DEEA and EMEA + H₂O mixtures with acceptable accuracies. The calculated enthalpies and entropies for the viscous flow enable the use of the model at different temperatures and mole fractions.

In future work, suggestions are to develop new semi-empirical and empirical correlations for measured densities and viscosities to reduce the complexity of the corrections and acquire higher accuracies.

Nomenclature

Letters	Description	Units
A_i	Parameters	
A_i^m	Measured property	$kg \cdot m^{-3}$ or $Pa \cdot s$
A_i^c	Calculated property	$kg \cdot m^{-3}$ or $Pa \cdot s$
h	Planck's constant	$m^2 \cdot kg \cdot s^{-1}$
H	Enthalpy of viscous flow	$kJ \cdot mol^{-1}$
M	Molecular weight	$kg \cdot mol^{-1}$
N	Avogadro's number	mol^{-1}
R	Universal gas constant	$J \cdot mol^{-1} \cdot K^{-1}$
S	Entropy of viscous flow	$kJ \cdot mol^{-1} \cdot K^{-1}$
T	Temperature	K
x_i	Mole fraction	-
ρ	Density	$kg \cdot m^{-3}$
ρ^E	Excess density	$kg \cdot m^{-3}$
η	Dynamic viscosity	$Pa \cdot s$
η^E	Excess viscosity	$Pa \cdot s$
ν	Kinematic viscosity	$m^2 \cdot s^{-1}$

References

- Aronu, U. E. *et al.* (2012) 'Density, viscosity, and N₂O solubility of aqueous amino acid salt and amine amino acid salt solutions', *J. Chem. Thermodynamics*. 45, pp. 90-99.
- Chen, S. *et al.* (2016) 'Density and viscosity of monoethylethanolamine + H₂O and monoethylethanolamine + diethylethanolamine solutions for CO₂ capture', *Thermochimica Acta*. 642, pp. 52-58. doi:10.1016/j.tca.2016.08.021.
- Eyring, H. (1936) 'Viscosity, plasticity, and diffusion as example of absolute reaction rates', *Journal of chemical physics*. 4, pp. 283-291.
- Fort, R. J. and Moore, W. R. (1966) 'Viscosities of binary liquid mixtures' *Transactions of the faraday society*. 62, 1112-1119.
- Heric, E. L. and Brewer, J. G. (1967) 'Viscosity of some binary liquid nonelectrolyte mixtures' *J. Chem. Eng. Data*. 12(04), pp. 574-583.
- Karunaratne, S.S. *et al.* (2020) 'Physical Properties of MEA + Water + CO₂ Mixtures in Postcombustion CO₂ Capture: A review of correlations and experimental studies', *Journal of Engineering*. 7051368. doi:10.1155/2020/7051368.
- Mahajan, A. R. and Mirgane, S. R. (2013) 'Excess molar volumes and viscosities for the binary mixtures of n-Octane, n-Decane, n-Dodecane, and n-Tetradecane with Octan-2-ol at 298.15 K', *Journal of Thermodynamics*. pp. 1-11.

doi:10.1155/2013/571918.

McAllister, R. A. (1960) 'The viscosity of liquid mixtures' *A.I.Ch.E. Journal*. 6, pp. 427-431.

Meyer, R. *et al.* (1971) 'Thermodynamic and physicochemical properties of binary solvent', *Journal de Chimie Physique et de Physico-Chimie Biologique*. 68, pp. 406-412.

Redlich, O. and Kister, A. T. (1948) 'Algebraic representation of thermodynamic properties and the classification of solutions' *Ind. Eng. Chem.* 40(2), pp. 345-348.

Mathematical Models for Physicochemical Properties of Different Amine-based Solvents in Post combustion CO₂ Capture

Sumudu Karunaratne^{a,*}, Jeanette Larsen^b, Lars Erik Øi^a

^aDepartment of Process, Energy and Environmental Technology, University of South-Eastern Norway

^bEquinor ASA, Hammerfest LNG, Norway

sumudu.karunaratne@usn.no

Abstract

In order to reduce global CO₂ emissions, CO₂ capture based on absorption in an amine/water mixture is an established method. To develop such processes, correct physicochemical properties like densities and viscosities are important.

The first objective of this work is to explore mathematical correlations for fitting viscosity data for aqueous Monoethanolamine (MEA) and Methyldiethanolamine (MDEA). A second objective is to evaluate the prediction of viscosity based on parameters independent of viscosity measurements.

13 developed correlations have been evaluated by comparing the maximum deviation of fitted models to the measured property, and by determining the average absolute relative deviation (AARD%). Python 3.6, MATLAB R2020b and Excel were used as the tools for regression.

The results indicated that viscosity for aqueous amines was better correlated by Eyring's viscosity model based on NRTL (Non-Random-Two Liquid model) rather than a Redlich-Kister correlation. The achieved AARD% of aqueous MEA were 2.39 for Redlich-Kister, 1.87 for Eyring-NRTL and 1.88 for the segment-based Eyring-NRTL model. The same trend was achieved for aqueous MDEA with AARD% of 3.04, 2.23 and 1.88 for different approaches.

The possibility of using data from vapor/liquid equilibrium parameters to predict viscosity in MEA/water and MDEA/water was evaluated. Using parameters in the equilibrium model NRTL from the simulation program Aspen HYSYS in a model from Karunaratne indicated that it is possible to predict viscosity reasonably well without experimental viscosity data.

Keywords: density, viscosity, Eyring's viscosity model

1. Introduction

Physicochemical properties of liquid amine mixtures have been studied as they are highly important in designing process equipment like absorption and desorption columns, heat exchangers and reboilers in the post-combustion CO₂ capture process. Viscosity is a highly influential physical property on both hydrodynamics and the mass transfer phenomenon in the absorption column. Viscosity appears in most of the mass transfer correlations for random and structures packings and heat transfer correlations for the heat exchangers. The measured viscosity data are fitted to semi-empirical and empirical mathematical expressions to calculate values for unmeasured conditions.

Eyring's viscosity model provides a theoretical background for viscosity by considering the motions of molecules from one place to another vacant position. The energy that needs to be overcome for such a motion is known as the free energy of activation for viscous flow ΔG^* (potential-energy

barrier). Eyring developed the following equation for liquid viscosity (Bird *et al.*, 2002; Eyring, 1936):

$$\eta = \frac{hN_A}{V} \exp\left(\frac{\Delta G^*}{RT}\right) \quad (1)$$

Where, η , V , h , ΔG^* , N_A , R , T are dynamic viscosity, molar volume, Planck's constant, free energy of activation for viscous flow, Avogadro's number, universal gas constant and temperature respectively.

Wu (1986) discussed an overview of the evolution of models for predicting the viscosity of liquid mixtures and the connection between the excess free energy of activation for viscous flows ΔG^{E*} and Gibbs free energy of mixing ΔG . Martins *et al.* (2000) discussed the applicability of the UNIQUAC model along with Eyring's viscosity model for multicomponent liquid mixtures. In addition to the size and shape parameters for each pure component, the model requires only two adjustable parameters

per binary system that can be determined by measured viscosities (Martins *et al.*, 2001). Cao *et al.* (1992) developed a model for viscosity of pure liquids and liquid mixtures based on statistical thermodynamics, local compositions, and Eyring's absolute rate theory. A further developed model was discussed by Cao *et al.* (1993) in which the molecular size was introduced to the model. The new model can predict both viscosity and activity coefficients of liquid mixtures and it was named as viscosity-thermodynamic model (UNIMOD). Weirong and Lempe (2006) combined cubic equations of states with Eyring's viscosity model to calculate viscosities in liquid mixtures. The different mixing rules of van Laar-type (VL) and Redlich-Kister-type (RK) were adopted for the comparison. Novak (2003) presented a model based on Eyring viscosity model and the NRTL model for describing deviations from ideal mixtures. The viscosity model was presented and evaluated using data from polymer-solvent systems. An extended work from Novak *et al.* (2004) transformed the original component-based Eyring-NRTL viscosity model into a segment-based Eyring-NRTL viscosity model for polymer-solvent and polymer-polymer systems that provided a more physically realistic model for large molecules. Matin *et al.* (2013) examined the use of an Electrolyte-NRTL model for the prediction of the viscosity of CO₂-loaded aqueous amine solutions. The study revealed that the Gibbs free energy of mixing as the appropriate thermodynamic quantity to replace the activation energy term for viscous in Eyring's viscosity model for liquid mixtures.

This work discussed several approaches to represent viscosities of binary aqueous amine mixtures of Monoethanolamine (MEA) and Methyldiethanolamine (MDEA) that were measured at different amine mole fractions and temperatures. The measured viscosities are available in the literature from our previous works (Karunaratne *et al.*, 2020(a); Karunaratne *et al.*, 2020(b)). There, the applicability of Eyring's viscosity model along with Redlich and Kister type polynomial for the excess free energy for viscous flow have been examined to represent viscosities.

In this study, Redlich-Kister, Eyring-NRTL, segment-based Eyring-NRTL and Eyring's viscosity model were adopted to represent measured viscosities. Different tools in Python 3.6, MATLAB R2020b and Excel were used as the tools for regression and the results were compared. This work confirmed the accuracies reported in the literature by performing the curve fitting through different tools in the Python program. Further, it discusses the applicability of NRTL (non-random two liquid) model based on vapor-liquid equilibrium (VLE) data. The proposed parameters for the NRTL model in Aspen Plus were used along with the correlation

discussed in (Karunaratne and Øi, 2019). Finally, the segment-based Eyring-NRTL model was adopted to represent mixture viscosities and relevant fitting parameters were estimated by minimizing the error between correlated and measured viscosities.

2. Methodology

Eyring's viscosity model shown in (1) is applied for the binary mixtures to represent the viscosities. Then a new term called excess free energy of activation for viscous flows ΔG^{E*} is defined as in (2).

$$\ln(\eta V) = \sum_{i=1}^{i=2} x_i \ln(\eta_i V_i^0) + \frac{\Delta G^{E*}}{RT} \quad (2)$$

A Redlich and Kister type polynomial can be fitted to represent ΔG^{E*} as shown in (3).

$$\frac{\Delta G^{E*}}{RT} = x_1 x_2 \sum_{i=1}^{i=2} C_i (1 - 2x_2)^i \quad (3)$$

$$C_i = a_i + b_i(T)$$

The NRTL model provides an expression for the excess Gibbs free energy of mixing ΔG^E as shown in (4). The model parameters are found by fitting the calculated partial pressures through the model into the measured partial pressures in experiments.

$$\frac{\Delta G^E}{RT} = x_1 x_2 \left(\frac{G_{21} \tau_{21}}{x_1 + x_2 G_{21}} + \frac{G_{12} \tau_{12}}{x_1 G_{12} + x_2} \right) \quad (4)$$

$$G_{ij} = \exp(-\alpha_{ij} \cdot \tau_{ij}) \quad (5)$$

$$\tau_{ij} = a_{ij} + \frac{b_{ij}}{T} + e_{ij} \ln T + f_{ij} T \quad (6)$$

$$\alpha_{ij} = c_{ij} + d_{ij}(T - 273.15K) \quad (7)$$

In this study, model parameters for the MEA + H₂O and MDEA + H₂O were taken from Aspen Plus. The correlation proposed in (8) was adopted to calculate ΔG^{E*} and use in (2) to recalculate mixture viscosities at different amine mole fractions and temperatures. A detailed discussion of developing (8) can be found in (Karunaratne and Øi, 2019) in which the parameters for ΔG^E were taken from (Schmidt *et al.*, 2007) for MEA + H₂O mixtures.

$$\frac{-\Delta G^E}{\Delta G^{E*}} = a + b \cdot x_1 \cdot T + c \cdot T^2 \quad (8)$$

The segment-based Eyring-NRTL model proposed by Novak *et al.* (2004) was adopted to fit the calculated ΔG^{E*} from Eyring's viscosity model for

MEA + H₂O and MDEA + H₂O mixtures. The modified segment-based Eyring-NRTL model to fit ΔG^{E*} is shown from (9) to (10).

$$\frac{\Delta G^{E*}}{RT} = \tilde{x}_1 \tilde{x}_2 \left(\frac{G_{21} \tau_{21}}{\tilde{x}_1 + \tilde{x}_2 G_{21}} + \frac{G_{12} \tau_{12}}{\tilde{x}_1 G_{12} + \tilde{x}_2} \right) \quad (9)$$

* a

$$\tilde{x}_1 = \frac{r_1 x_1}{r_1 x_1 + r_2 x_2}, \quad \tilde{x}_2 = 1 - \tilde{x}_1 \quad (10)$$

$$G_{12} = \exp(-\alpha_{12} \tau_{12}), \quad G_{21} = \exp(-\alpha_{12} \tau_{21}) \quad (11)$$

$$\tau_{12} = a_{12} + \frac{b_{12}}{T}, \quad \tau_{21} = a_{21} + \frac{b_{21}}{T} \quad (12)$$

The proposed models for the viscosities have been evaluated through average absolute relative deviation and absolute maximum deviation as given in (13) and (14).

$$AARD = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{A_i^m - A_i^c}{A_i^m} \right| \quad (13)$$

$$AMD = MAX |A_i^m - A_i^c| \quad (14)$$

3. Results and Discussion

3.1. Redlich-Kister polynomials

The previous work performed by (Karunaratne *et al.*, 2020(a)) to adopt Redlich and Kister type polynomial to represent calculated ΔG^{E*} from Eyring's viscosity model for MEA + H₂O mixtures as given in (3) was validated. Here, the density correlation proposed in (Karunaratne *et al.*, 2020(a)) for MEA + H₂O mixtures was used to calculate the mixture molar volumes rather than using measured densities. The study reveals AARD of 2.39% and an AMD of 0.84 mPa·s. A similar approach was taken to represent measured viscosities of MDEA + H₂O mixtures.

Table 1: Parameters for Redlich and Kister polynomial for excess free energy of activation for viscous flow of MDEA + H₂O mixtures.

Coefficients	
a ₀	30.45
a ₁	-20.4
a ₂	5.081
b ₀	-0.0684
b ₁	0.0444
b ₂	-0.00394
R ²	0.998

Tab. 1 illustrates the estimated parameters for the Redlich and Kister type polynomial. The calculated AARD and AMD are 3% and 7.2 mPa·s respectively

and that is acceptable. Fig. 1 illustrates the goodness of the data fit between calculated and measured viscosities at different amine concentrations and temperatures for MDEA + H₂O mixtures.

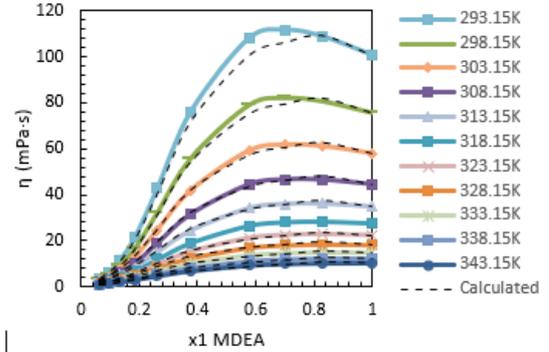


Figure 1: Comparison of calculated viscosity by Redlich-Kister polynomial for ΔG^{E*} with measured viscosity for MDEA + H₂O mixtures.

3.2. Eyring-NRTL model

The applicability of NRTL model for ΔG^E to represent ΔG^{E*} was discussed for MEA + H₂O and MDEA + H₂O mixtures. Tab. 2 lists the model parameters of NRTL model from (4) to (7) provided by Aspen Plus and the estimated parameters for correlation (8) are given in Tab. 3. The viscosities were calculated by (2) using pure liquid densities, viscosities and calculated ΔG^{E*} from (8). The calculated AARD and AMD of 1.9% and 1.5 mPa·s for MEA + H₂O mixtures and 3% and 7.2 mPa·s for MDEA + H₂O indicated that this approach could represent viscosity data with acceptable accuracy.

Table 2: Parameters for the NRTL model from Aspen Plus.

Parameters	MEA+H ₂ O	MDEA+H ₂ O
A ₁₂	-0.0352	4.75322
A ₂₁	1.1605	-1.79134
B ₁₂	-438.061	159.444
B ₂₁	-110.329	-716.787
C ₁₂ = C ₂₁	0.3	0.1
D _{ij} , E _{ij} , F _{ij}	0	0

Table 3: Estimated parameters for correlation (8).

Coefficients	MEA+H ₂ O	MDEA+H ₂ O
a	0.2157	0.1089
b	0.0001957	0.0004605
c	-4.674 × 10 ⁻⁰⁷	-7.107 × 10 ⁻⁰⁷
R ²	0.8441	0.9989

The deviations are relatively high at lower temperatures as shown in Fig. 2 to 5. With the increase in temperature, the real mixtures gradually

come to a state of ideal mixtures due to the weakening of intermolecular interactions among unlike molecules.

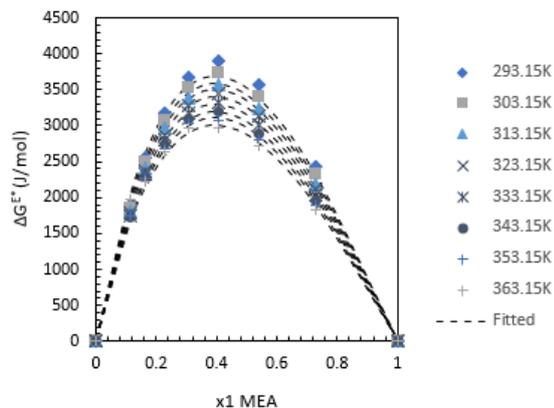


Figure 2: Comparison of calculated ΔG^{E*} using NRTL model and correlation (8) with calculated ΔG^{E*} from measured viscosities and densities for MEA + H₂O mixtures.

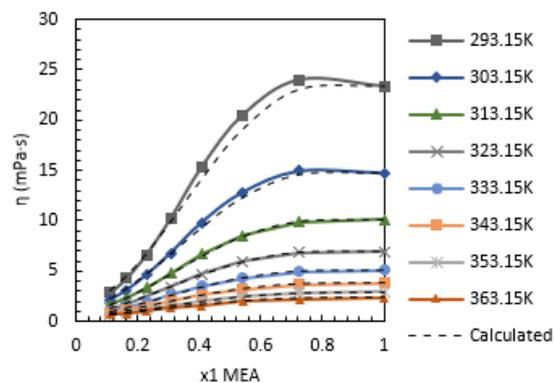


Figure 3: Comparison of calculated viscosities using NRTL model and correlation (8) with measured viscosities for MEA + H₂O mixtures.

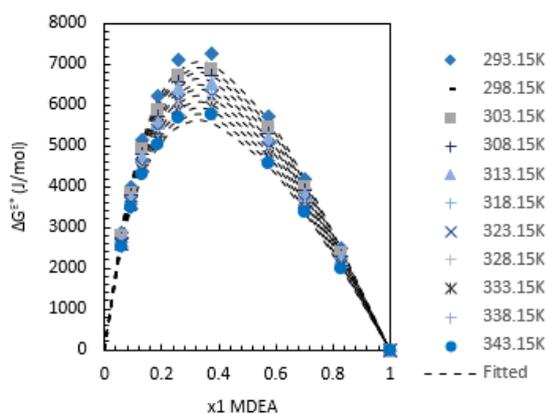


Figure 4: Comparison of calculated ΔG^{E*} using NRTL model and correlation (8) with calculated ΔG^{E*} from measured viscosities and densities for MDEA + H₂O mixtures.

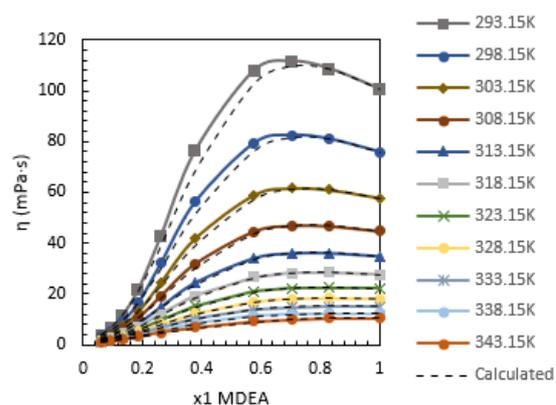


Figure 5: Comparison of calculated viscosities using NRTL model and correlation (8) with measured viscosities for MDEA + H₂O mixtures.

3.3. Segment-based Eyring-NRTL model

In the fitting of the segment-based model for MEA+H₂O mixtures, a problem with the relation was noticed. The small value of α_{12} in the parameter overview in Tab. 4 essentially renders G_{21} and G_{12} redundant, as the (11) reduces to the value of 1.

$$G_{12} = \exp(0 \cdot \tau_{12}) = 1$$

$$G_{21} = \exp(0 \cdot \tau_{21}) = 1$$

The results imply that (9) could be reduced to a function without G_{21} and G_{12} , which would exclude the nonrandomness parameter (α) from the correlation.

$$\frac{\Delta G^{E*}}{RT} = \tilde{x}_1 \tilde{x}_2 \left(\frac{\tau_{21}}{\tilde{x}_1 + \tilde{x}_2} + \frac{\tau_{12}}{\tilde{x}_1 + \tilde{x}_2} \right) * a \quad (15)$$

Table 4: Estimated parameters for Segment-based Eyring-NRTL correlation for excess free energy of activation for viscous flow of MEA + H₂O mixtures.

Parameters	MEA + H ₂ O
a	-8.173
a ₁₂	0.2936
a ₂₁	0.3271
α ₁₂	4.56 × 10 ⁻¹¹
b ₁₂	-403.6
b ₂₁	1.712
r	1.52
R ²	0.995

The formula appeared to work better for Aqueous MEA if the entire formula was multiplied by -1, and if parameters α_{12} , b_{12} and r were set to contain positive numbers. Resulting in the parameters in Table 5 for (9) to (12).

Table 5: Estimated parameters for Segment-based Eyring-NRTL model correlation (improved) for MEA + H₂O mixtures.

Parameters	MEA + H ₂ O
a	-1
a ₁₂	-14.89
a ₂₁	16.71
α ₁₂	0.041
b ₁₂	4827
b ₂₁	-6891
r	1.675
R ²	0.998

The alternative correlation gave a high coefficient of determination (R²) with the value of 0.998, meaning that the fit for excess free energy of activation for viscous flow for MEA + H₂O mixtures became slightly better than with the Redlich-Kister correlation. The goodness of the fit for both ΔG^{E*} and measured viscosity can be seen in Fig. 6 and 7 respectively. Further, calculated AARD and AMD were 1.2 % and 1.04 mPa·s respectively.

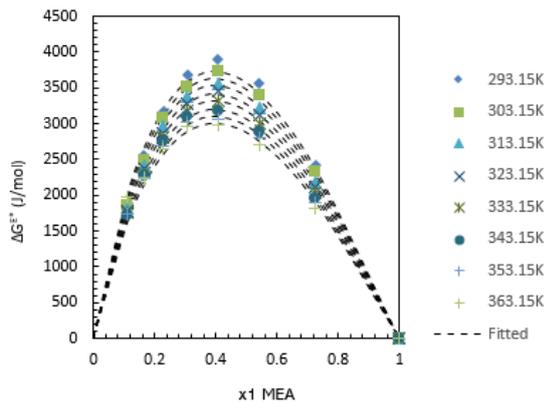


Figure 6: Estimated excess free energy of activation from the Segment-based Eyring NRTL model for aqueous MEA.

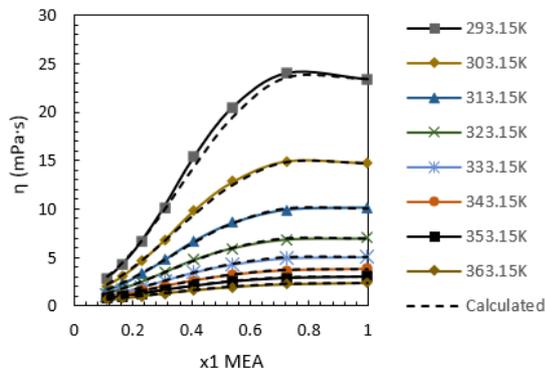


Figure 7: Comparison of measured and calculated viscosity by the segment-based Eyring-NRTL model for aqueous MEA.

The viscosities of MDEA + H₂O mixtures were found through fitted segment-based Eyring-NRTL model for calculated ΔG^{E*} . Relevant density data to calculate molar volumes were found from the fitted density data into a Redlich and Kister type polynomial discussed in Chapter 5.4.1 (Larsen, 2021). The calculated ΔG^{E*} from segment-based Eyring-NRTL model shown in Figure 8 considered estimated parameters from Tab. 6. The result was used to portray the viscosity in Figure 9, where the calculated values gave an AARD of 1.9% and an AMD of 8.1 mPa·s.

Table 6: Estimated parameters for segment-based Eyring-NRTL model for excess free energy of activation for viscous flow of MDEA + H₂O mixtures.

Parameters	MDEA + H ₂ O
a	1
a ₁₂	-2.504
a ₂₁	-3.817
α ₁₂	-0.101
b ₁₂	1596
b ₂₁	2937
r	2.349
R ²	0.999

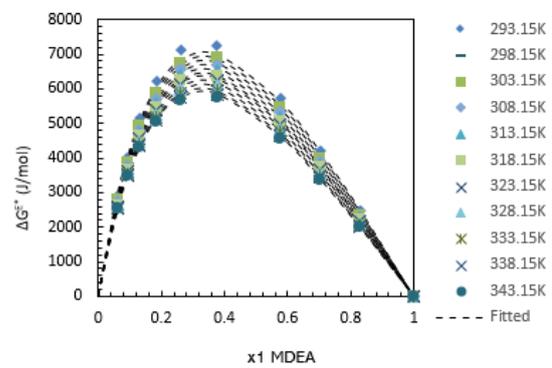


Figure 8: Fitted excess free energy of activation from the Segment-based Eyring-NRTL model for aqueous MDEA.

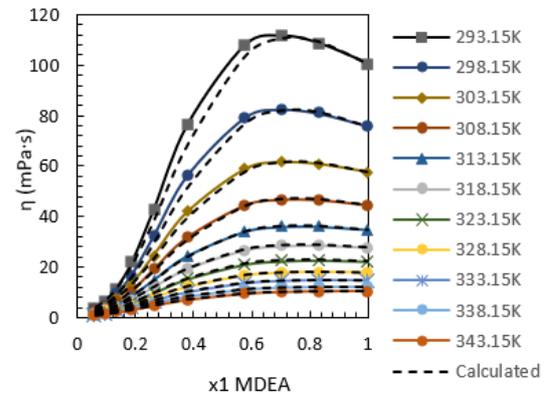


Figure 9: Comparison of calculated and measured viscosity of aqueous MDEA through the Segment-based Eyring-NRTL model.

4. Conclusion

Physicochemical properties are important in designing amine-based post combustion CO₂ capture processes as it can affect the mass transfer of CO₂ and the packing height of absorption/desorption columns. In this work, it was also found that viscosity seems to have a greater impact than density in the CO₂ capture process.

For the models for viscosity for aqueous amines, the applied methods of Redlich-Kister, Eyring-NRTL and segment-based Eyring-NRTL for ΔG^{E*} gave an AARD% of 2.39, 1.87 and 1.88 for MEA + H₂O mixtures respectively. Further, For MDEA + H₂O mixtures, these approaches were able to represent viscosities with an AARD% of 3.04, 2.23 and 1.88 respectively. The results revealed that the performed calculations connected to the NRTL model worked better for both mixtures. Eyring's model was connected to a correlation proposed in the literature and it showed the feasibility to apply the NRTL model and estimated parameters from VLE data in Aspen Plus to represent measured viscosities. In the segment-based Eyring-NRTL model, the estimation of all parameters was made possible by a correction factor. As future work, this study could extend to examining the applicability of VLE models like UNIQUAC and UNIFAC along with Eyring's viscosity model to make predictions for viscosities of binary aqueous amine mixtures.

Nomenclature

Letters	Description
A_i^m	Measured property
A_i^c	Calculated property
ΔG	Gibbs free energy of mixing
ΔG^E	Excess Gibbs free energy of mixing
ΔG^*	Free energy of activation for viscous flow
ΔG^{E*}	Excess free energy of activation for viscous flow
h	Planck's constant
N_A	Avogadro's number
r	Average number of segments
R	Universal gas constant
T	Temperature
V	Molar volume
x_i	Mole fraction
\tilde{x}_i	Segment of moles
η	Dynamic viscosity
α	Non-randomness factor
τ	Interaction energy between different species

References

- Bird, R. B. *et al.* (2002) *Transport Phenomena*. 2nd edn. John Wiley & Sons, Inc.
- Cao, W. *et al.* (1992) (a) 'Statistical thermodynamic model for viscosity of pure liquids and liquid mixtures' *Industrial & Engineering Chemistry Research*. 31(11), pp. 2603-2619. doi:10.1021/ie00011a027.
- Cao, W. *et al.* (1993) (b) 'Simultaneous correlation of viscosity and vapor-liquid equilibrium data' *Industrial & Engineering Chemistry Research*. 32(9), pp.2077-2087. doi:10.1021/ie00021a033.
- Eyring, H. (1936) 'Viscosity, Plasticity, and Diffusion as example of absolute reaction rates' *Journal of chemical physics*. 4, pp.283-291.
- Karunaratne, S. S. *et al.* (2020) (a) 'Density, viscosity and free energy of activation for viscous flow of monoethanol amine (1) + H₂O (2) + CO₂ (3) mixtures' *Fluids*. 5:13. doi:10.3390/fluids5010013.
- Karunaratne, S. S. *et al.* (2020) (b) 'Density, viscosity, and excess properties of MDEA + H₂O, DMEA + H₂O, and DEEA + H₂O mixtures' *Applied Sciences*. 10:3196. doi:10.3390/app10093196.
- Karunaratne, S. S. and Øi, L. E. (2019) 'Applicability of NRTL model for prediction of the viscosity of alkanolamine + water mixtures' *SIMS 60*. 13-16 August, Västerås, Sweden.
- Larsen, J. (2021) 'Mathematical models for the physicochemical properties of different amine-based solvents in post combustion capture' *Thesis (Master)*. University of South-Eastern Norway.
- Martins, R. J. *et al.* (2000) 'Excess Gibbs free energy model for calculating the viscosity of binary liquid mixtures' *Ind. Eng. Chem. Res.* 39, pp. 849-854.
- Martins, R. J. *et al.* (2001) 'Calculation of viscosity of ternary and quaternary liquid mixtures' *Industrial & Engineering Chemistry Research*. 40(4), pp. 1271-1275. doi:10.1021/ie000618h.
- Matin, N. S. *et al.* (2013) 'Application of electrolyte-NRTL model for prediction of the viscosity of carbon dioxide loaded aqueous amine solutions' *Ind. Eng. Chem. Res.* 52, pp.16979-16984.
- Novak, L. T. *et al.* (2003) 'Modeling the viscosity of liquid mixtures: polymer-solvent systems' *Industrial & Engineering Chemistry Research*, 42(8), pp. 1824-1826. doi:10.1021/ie030051f.
- Novak, L. T. *et al.* (2004) Segment-Based Eyring-NRTL viscosity model for mixtures containing polymers. *Ind. Eng. Chem. Res.* 43, pp. 6231-6237.
- Schmidt, K. A. G. *et al.* (2007) 'Use of the NRTL equation for simultaneous correlation of vapour-liquid equilibria and excess enthalpy' *Journal of Thermal Analysis and Calorimetry*. 89, pp. 61-72.
- Weirong, J. and Lempe, D. A. (2006) 'Calculation of viscosities of liquid mixtures using Eyring's theory in combination with cubic equation of state' *Chinese J. Chem. Eng.* 14(6), pp. 770-779.
- Wu, D. T. (1986) 'Prediction of viscosities of liquid mixtures by a group contribution method' *Fluid Phase Equilibria*. 30, pp. 149-156. doi: 0.1016/0378-3812(86)80050-4.

Himalayan Run-Off River Power Generation Modelling for Power Security in Evolving Weather Conditions

Swaechchha Dahal ^{a,b,*}, Thomas Øyvang ^b, Gunne John Hegglid ^b, Shailendra Kumar Jha ^a
Bhupendra Bimal Chhetri ^a

^a Department of Electrical and Electronics Engineering, Kathmandu University, Nepal, ^b Department of Electrical Engineering, Information Technology and Cybernetics, University of South Eastern Norway, Norway, sdahall13@gmail.com

Abstract

Extreme black-swan occurrences like earthquakes, glacial lake outbursts, flash floods, landslides, etc. are important concerns in Himalayan countries like Nepal, which are highly susceptible, geologically active, and exquisitely fragile. Nepal generates 97 percent of its electricity from hydropower, where 56.08 percent of it is coming from seasonal run-off-river (RoR) hydro plants. Landslides and mudflows are common in the monsoon, and low discharge is common in the winter season. These RoR plants must be able to withstand high-impact events like earthquakes and lengthy droughts in order for the Nepalese grid to remain secure. This study gives a presentation and overview of previously occurred natural hazards in Nepal related to hydropower plants. In particular, the 2014 Sunkoshi landslide and the 2021 Melamchi flood are evaluated as extreme events and their impacts on hydropower plant has been studied. In addition, an in-depth investigation on a ROR plant is carried out. Moreover, the water discharge and extreme rainfall peaks in time series data is evaluated using an ARIMA-based model. This paper shows the feasibility of predicting the energy produced by a run-off river hydropower plant. The purpose is to forecast discharge and hence the ROR power generation with the aim to facilitate the hydropower operators for their availability declaration which will again help in the overall energy planning. The results are discussed together with performance metrics, and indicates that the implemented technique is promising. These predictions can be further used for planning and estimating the power generation on a more complex level.

1. Introduction

1.1. Background

The electrical power system supports a variety of important infrastructures in today's world, including transportation, communication, health, and education. Countries, on the other hand, are subjected to severe exigencies and natural disasters that directly or indirectly damage electricity systems. The rising expense of power outages caused by natural disasters or climate fluctuations, as well as the devastating impact on different fields of security and personnel security, cannot be overlooked. Power outages over extended periods of time, substantial and important equipment failures in the system, cascading failures, load-shedding, and even system blackouts are all possible outcomes. For the power system to run smoothly, it is important to keep the system in balance while keeping its security and economic limits.

Extreme weather events, as well as climatic differences, are becoming more common in many countries throughout the world. As a result, today's electricity system must be resilient in this area. Power system engineers face a difficult task in designing a resilient power system that can resist climate change and extreme occurrences. After all, weather is stochastic, unexpected, and difficult to anticipate. Figure 1 depicts an overview of climate influences on the electricity system based on a number of papers. The globe is currently confronted with the issues of climate change and global warming; the primary fo-

cus has been on clean energy and the decarbonization of global energy systems. The movement toward a flexible electrical system based on renewables is gaining traction (Mitchell, 2016). Hydropower is the world's greatest renewable energy source, and it contributes significantly to global power system balance and regulation (Yang et al., 2018). Hydropower is not only needed for electricity; it also helps to balance the intermittent renewable energy supply. The relevance of hydropower to the reliability, sustainability, and economy of energy systems is addressed in Europe's 2050 Energy Strategy (Roadmap, 2011). The backbone for connecting multiple renewable energy systems is more flexible in a hydro-dominated power system. However there are still challenges with a power system that are heavily dependant on hydropower mainly due to natural events, for instance, scheduling the right amount of reserve capacity, the possibility of frequency oscillations in the system, and the overall quality, security, and dependability of the power system.

1.2. Nepalese Power System and the impacts of natural disaster

The Nepalese grid, known as the Integrated Nepal Power System (INPS), is one of the hydro-dominated electric grid systems in the Himalayan region. Currently, Nepal generates 97% of total electricity from hydropower; 56.08% of it from Run-Off-River (RoR) plants. These RoR plants are subject to a large discharge variation between wet and dry seasons. Also, the catchment area faces land-

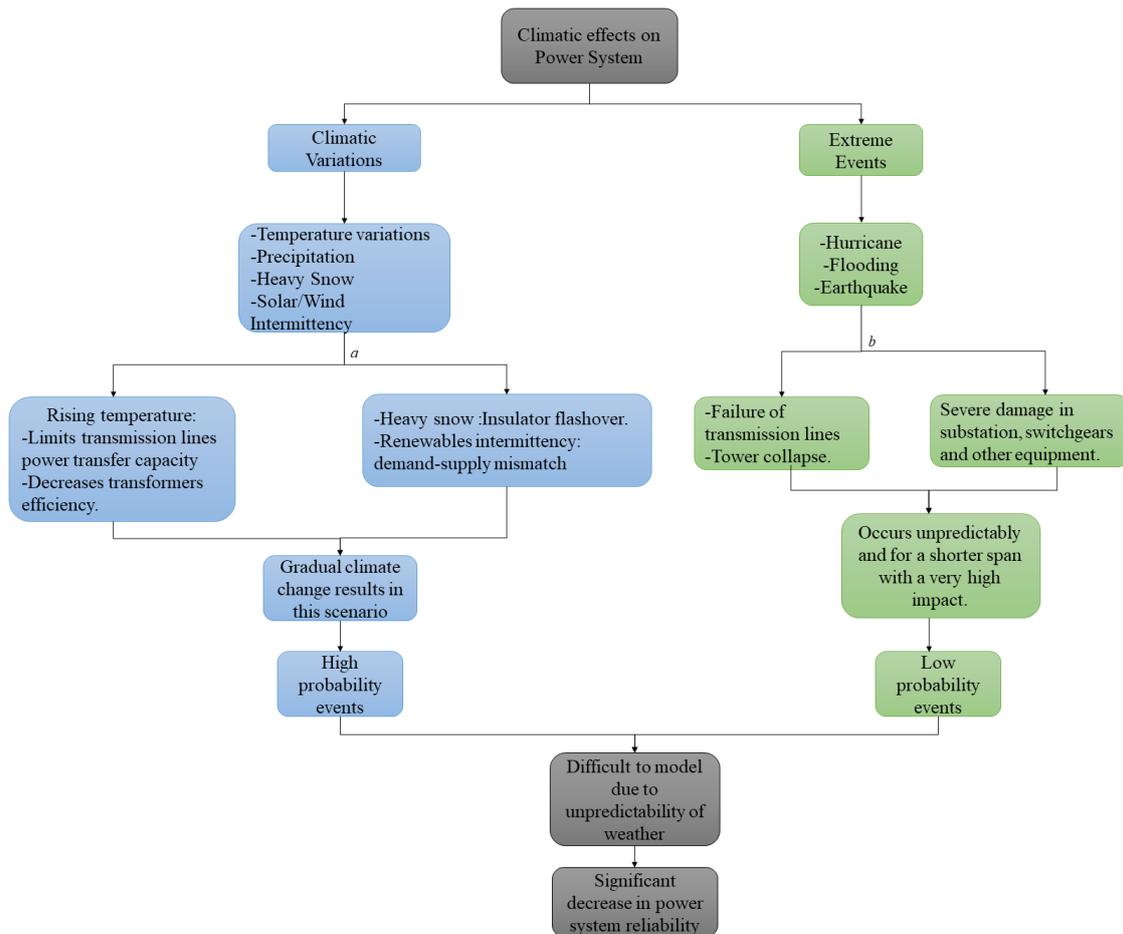


Figure 1: The figure presents the impact of climatic effects on the electric power system distinguished between climate variations and extreme events.

a represents climatic variation and their effects on power system

b represents effects of extreme events on power system

slides and heavy mudflow conditions in monsoon, and prolonged periods of low discharge in the winter. Typical generation capacity of a RoR plant is roughly 100% for 6 wet months, when it is most vulnerable to landslides and floods, 80% for 2 months, 50-30% for 4 months of winter when any form of energy is highly precious to the livelihood. Therefore, the power system operational security of the Nepalese grid is heavily dependent on how these RoR plants respond to winter demands. RoR should safeguard itself during the monsoon floods and landslides, and be resilient to other high impact events such as earthquakes and prolonged drought. Moreover, concerns with the Nepalese grid have posed a threat to INPS's reliability and security. Until 2017, Nepal was subjected to up to 18 hours of load shedding during the dry season. Although load shedding has been officially resolved, there is still a problem with INPS' power supply quality and reliability. In INPS, the existing transmission lines are mostly working at full capacity. In all existing high voltage lines, there is no sufficient provision for n-1 contingency. As a result of this and inappropriate contingency analysis, the INPS reliability is compromised.

River flow and hydropower are inextricably linked, and rainfall has a considerable impact on both. Because rainfall is influenced by a multitude of factors, hydropower

generation is highly seasonal. Nepal is also at risk from earthquakes, flooding, landslides, and a variety of other natural disasters. In Table 1, the effects of a few natural risks on Nepal's power generators and system over the last decade are presented. It is clear from this table that the Nepalese power infrastructure is extremely sensitive to weather and natural disasters.

1.3. Contribution of the work

Having a good grasp of the trends and behavior of river discharge is critical for an effective power/energy management, especially in the hydropower-dominated power system. The amount of rain, temperature, and other environmental factors influence how much water is discharged. Reliably predicting hydropower generation under certain operational conditions is essential to making the most of hydropower's advantages as a clean and inexpensive energy source. Economic and societal gains can be gained through its use. It is also possible to perform some preventative measures if the serious condition is anticipated. Consequently, this research focuses on analyzing the river discharge and hydroelectric output of one of the hydropower facilities in Nepal. The following points will serve as a reminder of the study's contributions during the discussion:

- Extreme events including floods, landslides, and

Table 1: Natural hazards in Nepal in the last decade.

Event	Major Effect
Flash Flood-Seti River-2012	Damage to infrastructure and livelihood including water supply systems and electric poles (Ojha, 2018)
Landslide-Jure-2014	Damage of more than 1.5 million USD in Sunkoshi HPP (Liu et al., 2020)
Earthquake-Gorkha- 2015	Severe damage to infrastructure and livelihood including significant damage in Sunkoshi HPP (Liu et al., 2020)
Glacial Lake Outburst Flood (GLOF)- Bhotekoshi- 2016	Damage to infrastructure and livelihood including damage in intake dam of Upper Bhotekoshi HPP (Action, 2018)
Terai Flood-2017	The financial loss of 584.7 million USD including damages in the energy sector as well (<i>First Tornado in Nepal – March 2019</i> , 2019)
Tornado-Bara/Parsa-2019	Damage to infrastructure and many livelihoods leaving affected places without electricity and communication (Service, 2020)
Landslide-Sindhupalchok-2020	Loss of many lives and houses and other infrastructure (Samiti, 2020)
Flash flood-Melamchi-2021	Swept houses, bridges, and severe damage to infrastructure and livelihoods including hydropower plants (Debnath and Mourshed, 2018)

earthquakes have all been researched in relation to the discharge statistics of the rivers.

- It is the goal of this work to analyze weather, discharge, and extreme events connected to time series data, and to estimate their effects on power generation using a modeling approach. The Auto-Regressive Integrated Moving Average modeling approach is used for the forecasting of time series data on discharge and power generation.

1.4. Structure of the paper

This paper opens by describing the setting in which the study is being conducted, as well as the reasons for doing so. Section 2 presents the assumptions and strategies utilized to tackle these problems. In the third section, the study's results are presented, and the conclusions taken from them are discussed in Section 4.

2. Methodology

For the investigation of underlying forecasting methods, a systematic review of published literature's was conducted. A number of literature's have shown that linear time series models, such as the Auto-Regressive Integrated Moving-Average (ARIMA) model, are among the most popular statistical models used to forecast time series based on historical data. It owes its reliability for this sort of data (Debnath and Mourshed, 2018). In (El Desouky and Elkateb, 2000), an Artificial Neural Networks (ANN) and ARIMA modeling were used to forecast electric load where both strategies were implemented to reduce forecasting mistakes, and they found that both techniques produce lower errors as compared to similar predictions obtained using the established time-series method. (Erdogdu, 2007) implemented ARIMA modeling technique for studying energy demand in Turkey and obtained results with very small error. Similarly, (Wang et al., 2011) employed a seasonal decomposition approach with vector regression to forecast hydropower usage in China and demonstrated that the method they utilized for time series forecasting was accurate. Likewise, (Cassiano et al., 2013) established an ARIMA model for forecasting effluent flow in a hydroelectric facility in Brazil by merging hierarchical clustering and Principal Component Analysis (PCA).

Forecasting, in general, refers to making future predictions based on the analysis of current and historical trends, with three primary components: input variables (historical and

current data), prediction methods (trend analysis), and output variables (future predictions), as illustrated in Figure 2.

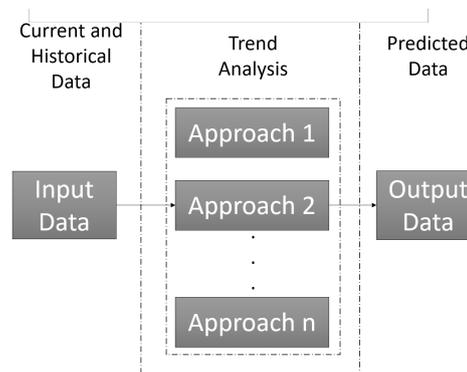


Figure 2: General forecasting model structure.

In this study, the prediction method used is the ARIMA modeling approach. The input variable is the river discharge (cubic meter per second/cumecs), and the output variable is the forecasted average discharge values. The ARIMA modeling technique is one of the efficient time-series forecasting model for hydropower generation forecast along with the prediction in the river discharge (Polprasert et al., 2021). Hence, the study has considered the implementation of ARIMA model in case of hydropower in Nepal which has not been done in the previous literature. Nepal is a developing country with huge potential in hydropower generation and it is very important to ensure a balance between the generation and consumption in order to make the grid secure. This study aims in contributing to this fact by analyzing ARIMA time-series modeling approach for investigating the behavior of Runoff River (RoR) hydropower plant subject to seasonal river discharge and its power generation. This will aid in further study of various other hydropower plants in case of Nepal to ensure power security in INPS.

The ARIMA model has the format of ARIMA (p, d, q), where p is the Auto-Regressive term, d is the order of differencing required for the data to make it stationary and q is the Moving-Average term. The established model is identified through the data of sample autocorrelation function (ACF) and partial autocorrelation function (PACF). Box and Jenkins presented a complete stepwise approach

for analysis and forecasting of a time series using ARIMA models in 1976 (Box et al., 2015). Because of the popularity of its methodology, ARIMA models are frequently referred to as Box-Jenkins models (Jamil, 2020).

The ARIMA model is a type of Box-Jenkins series analysis. The prediction value is viewed as a function defined by the time sequence. Integrated autoregressive and moving average models are used to model the data. These are regression models that include delays in the dependent variable as well as delays in the error term.

The ARIMA (p, d, q) model is decomposed into three parts (Polprasert et al., 2021).

The part (p) is AR part which represents the dependent variable regressed on its own lagged values as shown in equation 1.

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_t X_{t-p} + \epsilon_1 \quad (1)$$

where,

α_0 = constant

X_t = interpreted variable

$\alpha_1, \dots, \alpha_t$ = coefficients or AR model parameters

X_{t-p} = pre-stage data

t = periodic time

ϵ_1 = error term

Similarly, the part (q) is MA part which shows that the regression error is a linear combination of error terms whose values occurred simultaneously and at various times in the past as shown in equation 2.

$$X_t = u_t + \beta_1 u_{t-1} + \dots + \beta_q u_{t-q} \quad (2)$$

where,

β_1 = coefficient of MA model which is the weight

X_t = interpreted variable

t = periodic time

u_{t-q} = error term

Finally, the part (d) is the I part that indicates that the data values have been replaced with the difference between their current values and the previous values.

A general ARIMA model can be written as shown in equation 3.

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_t X_{t-p} + \epsilon_1 + u_t + \beta_1 u_{t-1} + \dots + \beta_q u_{t-q} \quad (3)$$

Before building ARIMA model, it is necessary to check the stationarity of the data. The first step is to analyze the data-time plot to examine whether it is covariance stationary. The time series is subjected to a unit root test to determine its stationarity. The unit root test regresses the time series X_t on its lag value X_{t-1} , and the results reveal the data's nature in respect of stationarity (*First Toronado in Nepal – March 2019*, 2019). If the data is non-stationary, after basic first-order or second-order differencing, the ARIMA model is still suitable to nonstationary time series. At this stage, the level of differentiation i.e., the number of times the data is differenced for its stationary is (d). Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (Kwiatkowski et al., 1992) test was used in the research as a unit root test for testing a null hypothesis that the observable time series is stationary around a deterministic trend. The KPSS test is a linear regression-based statistical test for determining if a time series is stationary. The test's null hypothesis is that the time series is stationary, whereas the alternative hypothesis is that the time series has a unit root. (Kwiatkowski et al., 1992) provides a detailed mathematical formulation behind the KPSS test.

The study is carried out using EXPLORATORY tool (Exploratory, Inc., n.d.). The software builds an ARIMA time series model and performs forecast based on input time series data. The model parameters p,d,q are selected on the basis of ACF and PACF plots. The effectiveness of the model is observed from three performance metrics: Root Mean Square Error (RMSE) which gives the root of mean of squares between the forecasted and actual value, Mean Absolute Error (MAE) which is the mean of absolute differences between forecasted and actual value and Mean Absolute Percentage Error (MAPE) which is the mean of absolute differences in percentage of actual value (Exploratory, Inc., n.d.).

3. Results and Discussion

The efficacy of the presented approach has been tested on a RoR Hydropower Project test case. The salient features of the project is shown in table Table 2.

Table 2: Salient features of the test case project

Particulars	Features
Location	Pyuthan District, Central / West Nepal
Intake River	Jhimruk Khola
River Training	2 km Canal, Gabion Mattresses
Dam	300m Curvilinear with desilting basin and intake, 10,000m ³ Concrete
Tunnels	Headrace 1100m x 8.5m ² , fully lined Inclined Shaft (45 degree) 280m x 9/3.5m ² , fully lined (Steel, Concrete)
Powerhouse	Semi-underground, Steel Trusses, 8m x 20m x 18m
Rated power	12,6 MW
Head	201,5 m
Flowrate	12 m ³ /s (4m ³ /s each turbine)
Turbines	3 units, Francis, Kvaerner
Speed	1000 rpm

The two years daily average discharge and power generation data collected are presented in Figure 3a. The data shows the daily variations of average discharge from 2017 to 2019. It is evident that the average discharge shows a seasonality trend with high discharge during wet seasons and low discharge in dry months. Moreover, there is an abundant water supply, typically from July to October.

It can be seen that the hydropower plant can not generate more than the rated maximum even at the times of maximum discharge. However, there is a substantial difference in the values of maximum and minimum energy production, which indicates the significant dependence of hydropower generation on discharge, water quality and operational constraints.

Ideally, in a hydropower plant, the abundance water suggests high hydropower generation. Therefore, a flat hydropower generation plot with maximum rated value could be expected during the times of high discharge. But, contrary to the ideal assumption, it can be observed that the hydropower generation seems to decrease when the river discharge spikes as shown in Figure 3b which is a zoomed in view of Figure 3a. This is because the selected site and the rivers of Nepal in general also face the challenges of high sediment issues which results in shut down/rescheduling of the generation units in order to prevent the sediment erosion in a turbine to ensure its durability when the river discharge increases during monsoon season. The generation plot is almost flat post monsoon season as seen in Figure 3b as the sediment is less during that period.

To understand the correlation between variables and the trend in decrease/increase of data, the raw data along with its trend line is illustrated in the following Figure 4.

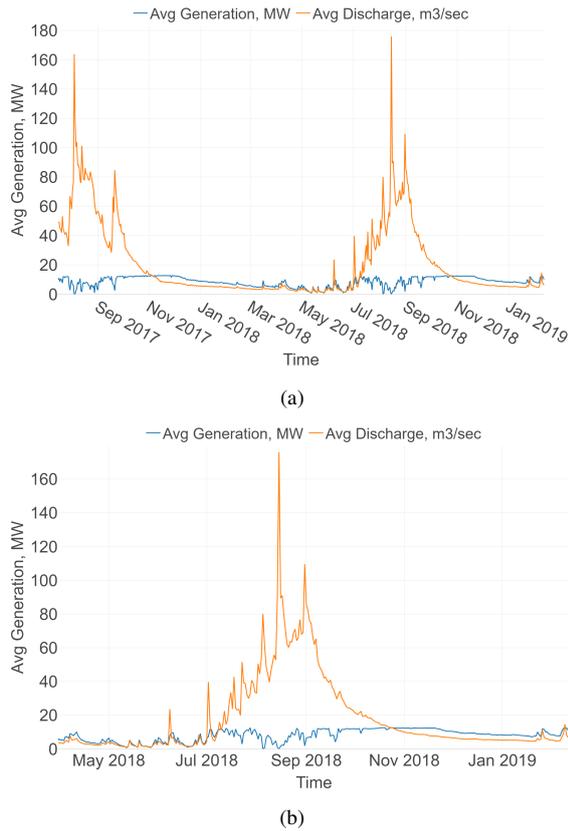


Figure 3: The figure presents the average discharge and average generation plot. (a) represents the overall plot of the input data for two years (b) represents a zoomed version of (a) from May 2018 for a better analysis of the plot

The trend change lines represent the points in a year where the river discharge is either changed to increasing or decreasing flow. The major rate of change in the trend line is observed around September, where the discharge sharply decreases from that month onward.

The predicted average discharge from ARIMA model as compared to the actual average discharge is shown in Figure 5a.

There was a considerable increase in average discharge around August 2017 as shown in Figure 5b, resulting in a significant difference between predicted and actual values. However, the model fit improves significantly after that, as the predicted and observed values nearly perfectly overlap.

The performance metrics RMSE, MAE and MAPE from the study was obtained as shown in Table 3. In general, a lower value yields a better prediction. It is evident from

Table 3: Performance Metrics

Metric	Value
RMSE	8.04
MAE	2.8
MAPE	0.12

the results that the RMSE value (squared error) is significant larger than the approaches for average error estimations. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. The RMSE is most useful when large errors are par-

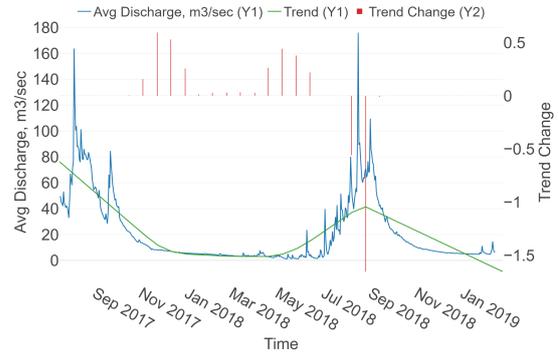


Figure 4: Change in the Average Discharge Trend.

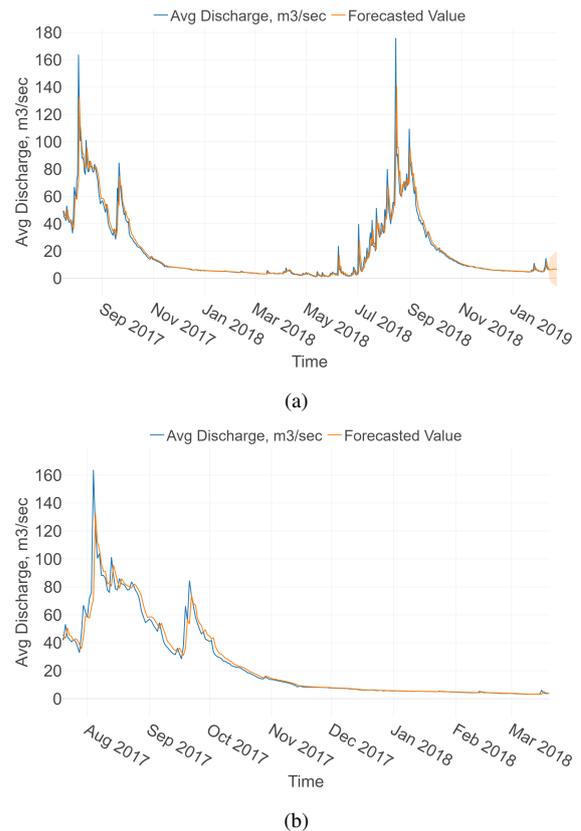


Figure 5: The figure presents the ARIMA model for average discharge. (a) represents the overall plot of the model for entire input data (b) represents a zoomed version of (a) for a better analysis of the model

ticularly undesirable. From the figures it can be seen that the difference between actual and forecasted values at the discharge peaks (rainfall) are significant. Hence, all three methods should be used together to diagnose the variation in the errors in a set of the forecasts. As the ROR plant has large seasonal variation and different magnitude, a mix of these methods should be considered in the forecasting evaluation. An ideal model should have the lowest error metrics. However, the results show a good fit of the developed model as compared to the previous literature in the relevant domain (Jamil, 2020), (Polprasert et al., 2021).

Figure 6 shows the short-term forecasting result obtained from the ARIMA model for average discharge. The forecasting for immediate 10 days has been obtained from the model. In the historical data, average discharge plot during January-March is almost flat which is reflected in the forecasted data for February as well. As the input data for only two years has been used in the modeling, the forecasted data tends to follow the trend shown by the historical data. The model do not seem to forecast the sudden events such as increment in discharge and the peak it causes. A minimum sample size of 50 is required for a reliable statistical analysis of a time series using the ARIMA method, which indicates that data for at least 50 years should be available (Box and Tiao, 1975). The intent of this work is to analyze and forecast annual generation based on discharge and other operation or maintenance variables; which is expected to show an annual trend but not the monthly or weekly or daily trend. As the data for large number of years was not available, the model accuracy is compromised. The model would have been more effective if larger volume of historical data was available as input for the model.

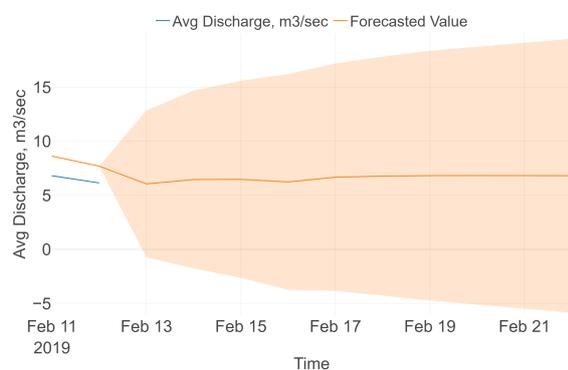


Figure 6: The figure represents the forecasting part of average discharge with ARIMA modeling as seen in bottom right part of Figure 5(a).

Proper energy planning is a crucial factor to ensure power security. There are several cases of power cut and unreliable power supply due to mismatch in the energy production capacity and its consumption. The scenario gets worse because of evolving climatic variations which has affected the power systems all over the world. Nepal, being a rich country in hydropower generation potential, is also a victim of seasonality leading to frequent power cuts and low power quality in various parts of the country. Several studies are being conducted to estimate the power system parameters subject to changing climatic conditions but the study in case of Nepal is very limited. Therefore, as a small step in ensuring power security for INPS, this study has applied ARIMA forecasting model for varying seasonal discharge to make it easy for the electricity regulators to conduct proper energy planning amidst the fluctuating river discharge.

The study, till this point, is done for average river discharge by fitting an ARIMA model. River discharge is a natural process which can be accurately modeled. Therefore, preliminary investigation to study and forecast has been obtained in this study. Further, the implications of weather events on hydropower production, as mentioned in Section 1, is yet to be addressed. Correlating weather events and actual hydropower generation and its modeling is a complex process as generation is affected by operational constraints such as operator's instructions as well. This is yet to be incorporated in the model. Although the aim is to incorporate extreme events such as GLOF, landslide, flooding, etc., the modeling of these events in relation to a hydropower production is a complex methodology which is yet to be addressed.

The major challenge during the study was the availability of the data from hydropower in Nepal. There was lack of proper record maintenance and also the power plants were reluctant to provide data. The study was performed with the daily discharge and generation data for two years. Results from modelling could be improved if more observations were available e.g., either hourly data or if data for a larger period of time was available.

4. Conclusions

This paper presents a study that investigates the historical and future trends of energy produced by a hydropower plant of Nepal, taking into account its historical average discharge conditions. Forecasting studies play a significant role in resource planning and management in the future. The gathered historical data was statistically analyzed, and ARIMA modeling was used to forecast Nepal's future average discharge impacting hydroelectricity generation. From the research investigation, it is concluded that the proposed model can be utilized to better understand the trend in discharge and generation of a ROR hydropower plant and as a reference for energy planning. However, when utilizing the model to anticipate electricity generation from green energy, which is based on natural resources, additional care must be taken because environmental circumstances and climate fluctuation might have a significant impact.

The major findings of the paper can be summarized as:

- It is evident that ROR hydropower plants have large natural yearly variations following the rainfall seasons. Future climate variations may further impact this variation, demanding good statistical models for discharge and power generation forecasting.
- An effective forecasting model demands more observations e.g., several years of discharge data to give a satisfying prediction.
- The ARIMA modeling approach is an effective method for predicting river discharge. The model follows the river discharge trend in mostly of the study and gives a short-term forecasting result. However, seasonal rainfall peaks are rather challenging to track.

Future work will focus to enhance the modelling framework and environment. More sophisticated data-driven models e.g., Long short-term memory (LSTM), and other artificial neural network and deep learning will be explored. A novel architecture such as Temporal Fusion Transformer (TFT) which is also being used in time-series forecasting and gives globally important variables, sustaining temporal patterns and significant events in the dataset is of interest. The proposed approach can be an improved decision tool for power producers for planning

and revenue. In addition, government officials in the energy industry can make informed energy production decisions and develop a long-term strategic plan to keep up with economic growth in Nepal.

References

- Action, P. (2018), 'Nepal flood 2017 wake up call for effective preparedness and response'.
URL: <https://reliefweb.int/sites/reliefweb.int/files>
- Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015), *Time series analysis: forecasting and control*, John Wiley & Sons.
- Box, G. E. and Tiao, G. C. (1975), 'Intervention analysis with applications to economic and environmental problems', *Journal of the American Statistical Association* **70**(349), 70–79.
- Cassiano, K., Junior, L., Souza, R., Menezes, M., Pessanha, J. and Souza, R. (2013), 'Hydroelectric energy forecast', *International Journal of Energy and Statistics* **01**. doi: 10.1142/S2335680413500142.
- Debnath, K. B. and Mourshed, M. (2018), 'Forecasting methods in energy planning models', *Renewable and Sustainable Energy Reviews* **88**, 297–325. doi: <https://doi.org/10.1016/j.rser.2018.02.002>.
URL: <https://www.sciencedirect.com/science/article/pii/S1364032118300200>
- El Desouky, A. and Elkateb, M. (2000), 'Hybrid adaptive techniques for electric-load forecast using ann and arima', *Generation, Transmission and Distribution, IEE Proceedings-* **147**, 213 – 217. doi: 10.1049/ip-gtd:20000521.
- Erdogdu, E. (2007), 'Electricity demand analysis using cointegration and arima modelling: A case study of turkey', *Energy Policy* **35**(2), 1129–1146. doi: <https://doi.org/10.1016/j.enpol.2006.02.013>.
URL: <https://www.sciencedirect.com/science/article/pii/S030142150600098X>
- Exploratory, Inc. (n.d.), 'Exploratory'.
URL: <https://exploratory.io>
- First Tornado in Nepal – March 2019* (2019).
URL: <https://www.inf.org.uk/first-tornado-in-nepal/>
- Jamil, R. (2020), 'Hydroelectricity consumption forecast for pakistan using arima modeling and supply-demand analysis for the year 2030', *Renewable Energy* **154**, 1–10. doi: <https://doi.org/10.1016/j.renene.2020.02.117>.
URL: <https://www.sciencedirect.com/science/article/pii/S0960148120303128>
- Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of econometrics* **54**(1-3), 159–178.
- Liu, M., Chen, N., Zhang, Y. and Deng, M. (2020), 'Glacial lake inventory and lake outburst flood/debris flow hazard assessment after the gorkha earthquake in the bhote koshi basin', *Water* **12**(2). doi: 10.3390/w12020464.
URL: <https://www.mdpi.com/2073-4441/12/2/464>
- Mitchell, C. (2016), 'Momentum is increasing towards a flexible electricity system based on renewables', *Nature Energy* **1**(2), 15030. doi: 10.1038/nenergy.2015.30.
URL: <https://doi.org/10.1038/nenergy.2015.30>
- Ojha, S. (2018), 'Rehabilitation of sunkoshi small hydropower plant (2.6 mk), nepal after sequences of natural disaster', *INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY AND SCIENCES (IJETS)*. doi: <https://doi.org/10.15282/ijets.v5i2.1398>.
- Polprasert, J., Nguyen, V. A. H. and Charoensook, S. N. (2021), Forecasting models for hydropower production using arima method, in '2021 9th International Electrical Engineering Congress (iEECON)', IEEE, pp. 197–200.
- Roadmap, E. (2011), '2050, impact assessment and scenario analysis', *European Commission, Brussel* **15**. doi: 10.2833/10759.
URL: <https://ec.europa.eu/energy/sites/ener/files/documents>
- Samiti, R. S. (2020), 'Flood sweeps houses, bridges in melamchi'.
URL: <https://thehimalayantimes.com/nepal/flood-sweeps-houses-bridges-in-melamchi>
- Service, H. N. (2020), '75 killed, 40 missing in sindhupalchowk floods, landslides'.
URL: <https://thehimalayantimes.com/nepal/75-killed-40-missing-in-sindhupalchowk-floods-landslides>
- Wang, S., Yu, L., Tang, L. and Wang, S. (2011), 'A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in china', *Energy* **36**(11), 6542–6554. doi: <https://doi.org/10.1016/j.energy.2011.09.010>.
URL: <https://www.sciencedirect.com/science/article/pii/S0360544211006074>
- Yang, W., Norrlund, P., Saarinen, L., Witt, A., Smith, B., Yang, J. and Lundin, U. (2018), 'Burden on hydropower units for short-term balancing of renewable power systems', *Nature Communications* **9**(1), 2633. doi: 10.1038/s41467-018-05060-4.
URL: <https://doi.org/10.1038/s41467-018-05060-4>

LSTM-based PSS Design for Modern Power Systems

Khaled Aleikish ^a, Thomas Øyvang ^b,

^{a,b} *Department of Electrical engineering, Information Technology and Cybernetics, University of South-Eastern Norway, Norway,*
Khaled.Aleikish@usn.no, Thomas.Oyvang@usn.no

Abstract

With the ever-increasing incorporation of wind and solar power in the electric power system, enhanced performance of classical bulk hydropower plants for robust operation of the power system is required. This current energy transition may cause a rapid increase in undesirable low-frequency oscillations (LFOs) in modern power system operations. A power system stabilizer (PSS) located at hydropower plants is one solution to damp such oscillations. This paper presents a new method based on Long Short-Term Memory (LSTM) neural networks for sine-wave phase shifting to possibly enhance PSS damping. The proposed controller considers the PSS input and the rotor speed deviation as a damped sinusoidal signal, simplifying PSS control and real-time optimization of PSSs parameters. Results show that the proposed LSTM architecture is able to learn multiple damped sine waves with different frequencies and decay rates. Therefore, the proposed controller can operate on the entire range of LFOs, unlike simple feedforward neural network (FNN) controllers, which can only learn and function on a single LFO frequency.

1. Introduction

The shift towards a more sustainable energy system demands increased stability properties from the hydropower fleet. Power system stabilizers (PSS), formerly known as supplementary excitation control systems, were developed to enhance the damping of low-frequency oscillations (LFOs) and increase power transfer limits [4, 5, 9–11]. In Norway, it is obligatory to install PSS on synchronous generators with a capacity of 30 MW and above (type D) [16]. The first PSSs were installed in the Nordic power system around 1970 to reduce power oscillations and increase power exchange on the interconnections between Norway and Sweden. PSS design and tuning were re-visited through the 1990s when the system loading and demand for power exchange increased, becoming one of the most cost-effective solutions for enhancing power system stability [12]. Traditionally, PSSs are only tuned and validated during the commissioning of the machine. These start-up settings of the PSS have the intention to dampen a wide range of low-frequency oscillations in the grid system during operation [18]. However, as the power system develops and expands with more intermittent energy sources such as wind and solar, new challenges are introduced to the operation of the power system. High-impedance weak grid systems and reductions in short-circuit power capabilities will transform the grid characteristics and may adversely affect the damping performance of the online PSS operation.

Over the past couple of decades, advancements in machine learning algorithms and computing power have enabled researchers to explore automatic calibration of PSS parameters to changing grid conditions [2, 3, 15]. Moreover, [6] proposed two methods to enhance the small-signal stability of a single-machine infinite bus (SMIB) system. Firstly, a particle swarm optimization (PSO)

algorithm was used to determine optimal parameters for a conventional power system stabilizer (CPSS) [8]. CPSS is a simplified version of the PSS1A type PSS [1]. The PSO algorithm optimizes the CPSS parameters for a specific value of the external (Thévenin) impedance connecting the synchronous machine to the infinite bus. However, PSO algorithms are computationally expensive and potentially slow at finding solutions. Hence, a simple feedforward neural network (FNN) was used to map a range of external impedance values to the corresponding set of optimized parameters by the PSO algorithm. The end design is an auto-tuning system that automatically adjusts CPSS parameters in response to changes in the external network impedance.

Secondly, a model-free approach to PSS design was proposed in [6]. A simple FNN-based PSS, called the sine shifting neural network (SSNN) controller, was developed without relying on complex electrical machine theory. Unlike the first method, which augmented the CPSS with an auto-tuning system based on artificial neural networks, the SSNN controller completely replaces the CPSS.

This paper proposes replacing the simple FNN architecture of the SSNN controller with a more complex neural network architecture to improve the stability performance of the PSS when subjected to a wide range of LFO in the electricity grid. Specifically, the proposed approach, which is called the Sine Shifting LSTM (SSLSTM) controller, uses a Long Short-Term Memory (LSTM) neural network, which is a type of recurrent neural network (RNN) architecture, to expand the operational range of the SSNN controller. While the SSNN controller can only function correctly over a single LFO frequency, the proposed SSLSTM controller can operate effectively over a wide range of LFO frequencies (0.1 to 2.5 Hz) with minimal performance impact. Moreover, a detailed discussion on best-practice for

picking training sets and training options is included. The paper is organized as follows. Section 2 briefly describes the excitation control system of synchronous generators and CPSS's role in the control loop. Section 3 describes the SSNN controller, while section 4 describes LSTM networks. Section 5 describes the proposed controller. Simulation and results are presented in Section 6, and results are discussed in Section 7. Finally, conclusions and future work are given in Section 8.

2. Excitation systems

A typical excitation control system consists of an exciter, automatic voltage regulator (AVR), and a PSS. The AVR regulates the generator terminal voltage by controlling the amount of current supplied to the generator field winding by the exciter, while the PSS is a feed-forward supplementary control device. The primary function of the PSS is to damp generator rotor oscillations (LFO's) and enhance both steady-state stability and transient stability. A well-tuned excitation system provides benefits such as improved oscillation damping, relay coordination and enhanced first-swing transient stability. Fig. 1 shows a block diagram of a grid-connected synchronous generator's excitation control system. In the figure, the PSS output V_{PSS} is an auxiliary control signal passed to the AVR. The AVR input can be expressed as:

$$\Delta V_{\Sigma} = \Delta V + V_{PSS}. \quad (1)$$

Here, the voltage error ΔV is expressed as:

$$\Delta V = V_{ref} - V_g. \quad (2)$$

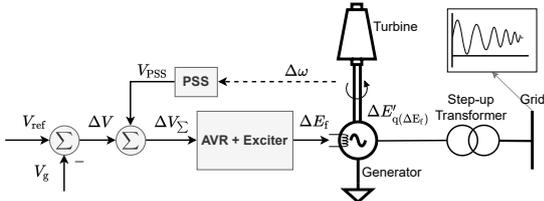


Figure 1: Block diagram of a grid-connected synchronous generator's excitation control system. In the figure, V_{PSS} is the supplementary control signal, ΔV is the voltage error, $\Delta E'_{q(\Delta E_f)}$ is the quadrature component of the transient emf, $\Delta \omega$ is the speed deviation, V_{ref} is the reference voltage, and V_g is the generator terminal voltage [12].

Fig. 2 shows a block diagram of a CPSS. In this figure, the CPSS is made up of four parts: (a) amplifier gain, (b) a signal washout high-pass filter, (c) lead elements for phase compensation, and (d) a limiter. In addition, some CPSS also include a signal sensor and a low-pass filter stage, which is not shown in the figure. A common signal used as input in the CPSS is the speed deviation $\Delta \omega$. In most studies the amplifier gain K_{PSS} and the time constants of the phase compensation stage are typically tuned to damp the LFOs. The other parts of the CPSS ensure that it functions as intended and does not disrupt the AVR action.

Classical tuning and performance evaluation of the PSS are typically done through phases compensation, root locus, and time domain analysis [9–11]. In phase compensation, which is the most widely used approach, the stabilizer is tuned to compensate for the phase lags through the generator, the excitation system, and the power system in such a way that torque changes are in phase with speed changes.

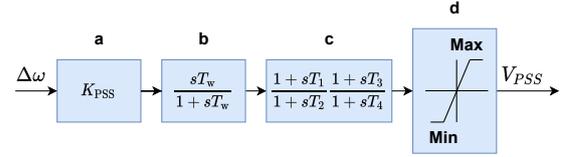


Figure 2: Block diagram of a conventional PSS (CPSS) [6, 12].

3. Neural Network PSS design

In [6] a Sine Shifting Neural Network (SSNN) Controller was developed. This approach has been evaluated in this paper. The SSNN approach is based on the assumption that the rotor speed deviation $\Delta \omega$ can be considered a damped sinusoidal signal. However, the neural network used to build the controller was trained on a sinusoidal signal without taking damping into account. That is, the controller input, the speed deviation $\Delta \omega$ was expressed as:

$$\Delta \omega \approx A_s \sin(\omega_s t). \quad (3)$$

Here, A_s and ω_s are the amplitude and the frequency of the sinusoidal signal, respectively. Also, the controller output was expressed as an identical, phase-shifted sine wave

$$\hat{\Delta \omega} \approx A_s \sin(\omega_s t + \beta). \quad (4)$$

where β is a control variable that represents the desired phase shift. Ideally, the controller would require only two inputs: the speed deviation $\Delta \omega$ and the desired phase shift β . However, since the SSNN controller was designed using FNN, which is a simple neural network architecture without internal memory, additional inputs were required. The additional inputs are the speed deviation values at the two previous time steps: $\Delta \omega|_{t=t-1}$ and $\Delta \omega|_{t=t-2}$. Algorithm 1 shows the pseudo-code for generating the SSNN training data. In essence, the

Algorithm 1 Pseudo-code to generate the training data set for the SSNN controller [6]. An FNN-based SSNN requires four inputs.

```

1: for Every amplitude  $A_i$  in  $A$ 's range do
2:   for Every phase shift  $\beta_j$  in  $\beta$ 's range do
3:     for each time step  $t_k$  in one period do
4:       SSNN input 1 =  $\beta_j$ 
5:       SSNN input 2 =  $A_i \sin(\omega t_k)$ 
6:       SSNN input 3 =  $A_i \sin(\omega t_{k-1})$ 
7:       SSNN input 4 =  $A_i \sin(\omega t_{k-2})$ 
8:       SSNN output =  $A_i \sin(\omega t_k + \beta_j)$ 
9:     end for
10:   end for
11: end for

```

additional inputs serve as external memory cells to the SSNN. However, information stored in these memory cells is lost as new data is measured, which limits the functionality of the SSNN. Furthermore, the performance assessment in [6] showed that under the FNN architecture, the SSNN performs well only at the frequency at which it was trained. In addition, a slight attenuation or gain was observed in the amplitude when $\beta \neq 0$. The results indicate that a controller based on FNN can only output the correct amplitude for undamped sine waves. This is evident from the amplitude of the oscillations

in the results; until the amplitude decays, the controller outputs the correct amplitude. For damped sine waves, the controller outputs the correct amplitude only when $\beta = 0$. Moreover, according to [6], it is practically impossible to train an FNN to differentiate between sine waves of different frequencies; training SSNN for more than one frequency under the FNN architecture results in a network that performs as if it were trained on the average of all the frequencies. Consequently, the results in [6] show the controller's performance deteriorates rapidly at all frequencies except for the one for which it was trained. To correct the amplitude and the phase drift, [6] describes several approaches. One approach assumes that the frequency of oscillation is known at the time of the disturbance and proposes training several SSNNs at different frequencies and enabling the one that was trained for the current frequency of the oscillation. However, this approach of multiple SSNNs is valid only if the oscillation frequency can be determined in real-time and relatively fast to avoid degrading the PSS's initial performance.

In this work, the frequency of oscillations is considered to be unknown, but it is assumed to be in the range of 0.1 – 5 Hz. If the oscillation frequency is unknown, [6] proposes using a more complex neural network to correct the amplitude and the phase drift, namely RNN and the LSTM architecture.

4. Long Short-Term Memory network

For a better understanding on the underlying architecture in the proposed method this section describes the Long Short-Term Memory (LSTM) network. LSTM is an advanced type of RNN that is capable of learning long-term dependencies between time steps of time-series data or any other type of sequential data [7]. Fig. 3 shows a block diagram of an LSTM cell (left) and LSTM layer (right). A single LSTM layer can contain N LSTM cells, where N depends on the length of the longest sequence of interest. In addition, it is also possible to stack several LSTM layers in a single neural network architecture to create deeper LSTM networks.

Furthermore, each LSTM cell consists of three gates: (a) a Forget gate, which controls what information should be discarded from the old cell state c_{t-1} , (b) an Update gate, which controls the flow of new information into the new cell state c_t , and (c) an Output gate, which controls the value of the next hidden state h_t (also called the output state). At time step t , the cell state c_t and the hidden state h_t are expressed as:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot \sigma_{\tanh}(c_t). \quad (6)$$

Here, the symbol \odot denotes the Hadamard product (element-wise multiplication). Moreover, the functions f_t , i_t and o_t are given by:

$$f_t = \sigma_{\text{sigmoid}}(W_f x_t + R_f h_{t-1} + b_f) \quad (7)$$

$$g_t = \sigma_{\tanh}(W_g x_t + R_g h_{t-1} + b_g) \quad (8)$$

$$i_t = \sigma_{\text{sigmoid}}(W_i x_t + R_i h_{t-1} + b_i) \quad (9)$$

$$o_t = \sigma_{\text{sigmoid}}(W_o x_t + R_o h_{t-1} + b_o). \quad (10)$$

Here, the matrices W , R , and b represents the learnable parameters (weights) of the LSTM cell. To view and analyze the learnable parameters in MATLAB, the neural network can be imported to the Deep Network Designer app, which can analyze the network parameters. For example, Tab. 1 shows the analysis results of an LSTM

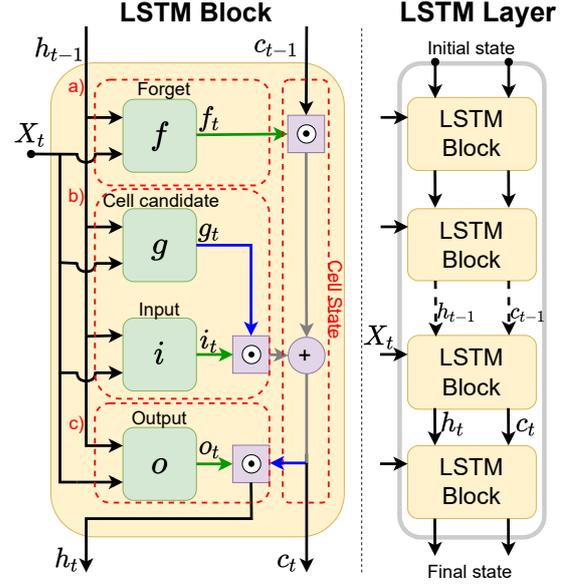


Figure 3: An LSTM cell (left) and LSTM layer (right). In MATLAB, the default activation function for f_t , i_t and o_t is the sigmoid function (represented by green lines), while the hyperbolic tangent function (\tanh is represented by blue lines) is used for g_t [13, 17].

Table 1: Analysis results of an LSTM network with 128 hidden units using MATLAB's Deep Network Designer application. The total learnable parameters in the network is 67.2×10^3 .

Type	Properties
1 Sequence input	–
2 LSTM	Weights: 1×128 Bias: 1×1 InputWeights: 512×2
3 Fully Connected	RecurrentWeights: 512×128 Bias: 512×1
4 Regression Output	–

network with 128 hidden units, where the number of hidden units refers to the size of the hidden state h_t .

Tab. 2 summarizes the hyperparameters of the LSTM layer. In MATLAB, **InputSize** is automatically set at training time. It is typically set to the number of features in the data set. **NumHiddenUnits** (the hidden size) is a hyperparameter that determines how much information the hidden state can store from previous time steps. This hyperparameter highly influence SSLSTM prediction accuracy. **OutputMode** determines if the layer outputs the complete sequence or the last time step of the sequence. Since SSLSTM phase-shift one point at a time, this parameter is set to 'last'. The last two hyperparameters are the activation functions used in Equations 6 - 10.

Table 2: Hyperparameters of the LSTM layer.

Hyperparameter	Value
1 InputSize	auto
2 NumHiddenUnits	128
3 OutputMode	last
4 StateActivationFunction	\tanh
5 GateActivationFunction	sigmoid

5. Sine Shifting LSTM Controller

In this work, a Sine Shifting controller based on LSTM networks is developed to phase shift the speed deviation $\Delta\omega$ to some optimal value. In the CPSS structure shown in Fig. 2 the required phase shift is obtained using two lead/lag stages. The time constants of these stages are tuned to produce a control signal that induces positive damping in the synchronous machine. Moreover, the proposed SSLSTM controller like the SSNN controller uses a single control parameter, β , to obtain the required phase shift. However, in contrast to the SSNN controller which can only function properly on a single LFO frequency, SSLSTM can function on the entire range of LFO frequencies.

The ability of LSTMs to predict discrete sine functions was studied by [14]. However, in this work it is focused on phase-shifting rather than forecasting future values. The objective is to develop a simple controller with only two control parameters, K_{PSS} to control the gain, and β to control the phase. It should be pointed out that the optimal values of K_{PSS} and β are outside the scope of this work and left for future research.

5.1. Generating the training data sets

Training, validation, and testing data sets are required to develop the machine learning model. In this study, different techniques were evaluated to generate the data sets: (a) an *Expanding Window*, (b) a *Sliding Window*, (c) a *Sliding Data*, and (d) An *Expanding-Sliding Window*. Depending on the technique, training time and model performance may be adversely affected. Moreover, in these techniques each predictor has a dimension of $2 \times W$, where W is the window width (A sequence contains at least two points, therefore the minimum value of W is 2). However, the window length is fixed to 2 (the number of features). Unlike the SSNN (Eq. 3 and Eq. 4), the SSLSTM features $e^{-\lambda t} A_s \sin(\omega_s t)$ and β include the Decay constant λ . Also, the targets, $e^{-\lambda t} A_s \sin(\omega_s t + \beta)$, have a fixed dimension of 1×1 .

Fig. 4 shows the *Expanding Window* method. In this method, the data sets consist of sequences of varying width. This method guarantees that targets are generated for all time steps $\{t_1, t_2, \dots, t_{end}\}$. However, this method will add unnecessary information to all sequences beyond S_N . This implies that any target generated after the first period will have a predictor pair that contains unnecessary data points since one sine wave period contains all the necessary information to learn λ and ω . Hence, a better solution is to discard these redundant data points to save memory space, reduce training time, and enhance training performance.

In Fig. 5 the *Sliding Window* method is presented. In this method, the data sets consist of sequences of fixed width. This method guarantees that enough information is contained in all predictors when the window width is greater or equal to N . However, if the window width is greater or equal to N , then no targets are generated for time steps that comes before N . Moreover, Fig. 6 shows a comparison between the data sets generated by the *Expanding Window* method and the *Sliding Window* method. In the figure, no targets were created for t_1 and t_2 in the *Sliding Window* method. To create targets for t_1 and t_2 , either the window width can be reduced from 4 to 2, or the time steps prior to t_0 can be filled with zeros (pre-padding the data array with zeros).

Fig. 7 shows the *Sliding Data* method, which is a slight modification to the *Sliding Window* method. In this method, as in the *Sliding Window* method, the data sets

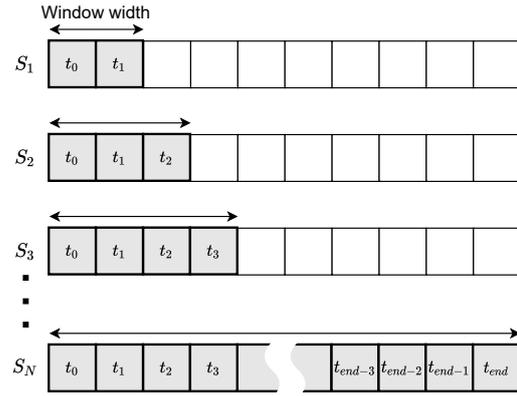


Figure 4: (a) The Expanding Window method. In this method, sequences have a variable length. Also, at some time t_n , S_n contains all previous data points.

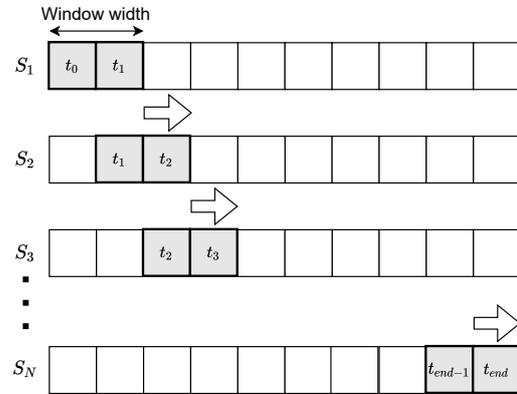


Figure 5: (b) The Sliding Window method. In this method, the window width is fixed. To illustrate, the window width is set to two.

consist of sequences of fixed width. However, the data points slide into the window. This slight modification guarantees that targets are generated for all time steps, as in the *Expanding Window* method. However, this method pre-pads the window/predictors with zeros, which results in an inefficient memory allocation for all targets with predictors length less than the window width. This is illustrated in Fig. 7 for the first three sequences $\{S_1, S_2, S_3\}$.

Fig. 8 shows the *Expanding-Sliding Window* method. This method combines the *Expanding Window* and *Sliding Window* methods to ensure memory-efficient generation of targets for all time steps. In this study, this approach proved to be the most effective method for generating the data sets and was selected to develop the SSLSTM PSS. Also, it is possible to add redundant data for the first quarter of the first period using the *Expanding Window* method to speed up and improve model training for early predictions. To illustrate this, Fig. 9 shows a histogram of a training data set generated by the *Expanding-Sliding Window* method for 2.2, 3.2, and 4.2 Hz sine waves. In addition, the figure illustrates how the *Expanding Window* method can be used to add redundant data to the training data set. However, it is advised not to add too much redundant data or the model can overfit.

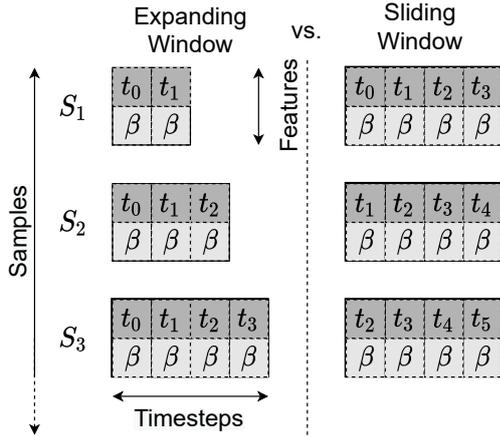


Figure 6: A comparison between data sets generated by the Expanding Window and the Sliding Window methods. Each sample consist of two features and W data points. Where W is the window width.

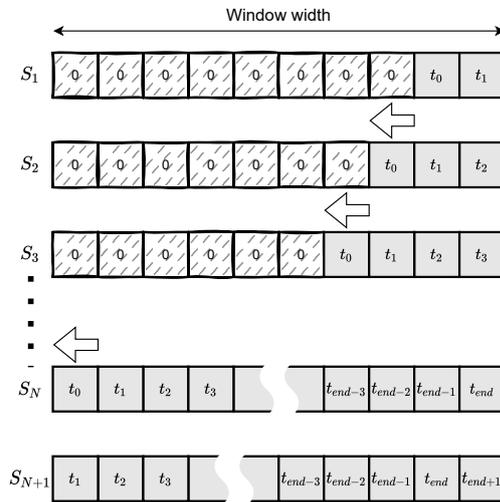


Figure 7: (c) The Sliding Data method. In this method, the data points slide into a window with fixed width. Although all sequences has same number of data points, some might contain zeros (pre-padding the sequences with zeros before training).

5.2. Neural network architecture of SSLSTM

The proposed neural network architecture of SSLSTM is shown in Fig. 10 using Deep Network Designer application in MATLAB. The network layers were described in Tab. 1. In addition, **zscore** normalization option was used in the **sequenceInputLayer** layer. Also, the **Output mode** was set to 'last' in the LSTM layer to perform sequence-to-one regression.

5.3. Training options

This section describes MATLAB's training options used in this work. Moreover, Tab. 3 lists the training options used in this study. Apart from the training options shown in the table, no other changes were made to the default training options. Throughout this study, the Adam optimizer was used for training the network. In MATLAB, except for **solverName** and **Plots**, all other options in the table are considered optional arguments. These arguments can be categorized based on their function into eight groups:

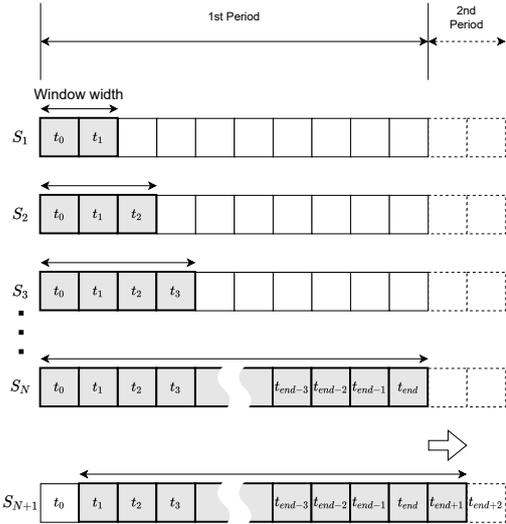


Figure 8: (d) The Expanding-Sliding window method. In this method, the Expanding window and the sliding window methods are combined to improve the model performance.

(a) Plots and Display options, (b) Mini-Batch options, (c) Validation options, (d) Solver options, (e) Gradient Clipping options, (f) Sequence options, (g) Hardware options, and (h) Checkpoints options.

Table 3: Training Options.

	Name	Value
1	solverName	adam
2	Verbose	0
3	Plots	training-progress
4	MaxEpochs	100
5	MiniBatchSize	2048
6	Shuffle	every-epoch
7	InitialLearnRate	0.01
8	LearnRateSchedule	piecewise
9	LearnRateDropPeriod	1
10	LearnRateDropFactor	0.95
11	GradientThreshold	1
12	SequenceLength	longest
13	SequencePaddingDirection	left

First, from the Plots and Display options, the option **Plots** was set to "training-progress" to visualize the training progress since it is easier to monitor the training progress by the accuracy and loss plots of the validation and training sets. The downside to this is that after training the network for a long time, it can become difficult to tell if the metrics are improving or not since the plots do not automatically scale to the last few training iterations. In this case, it is best to use **Verbose** to monitor the training progress, which displays the training progress metrics in the command window.

Secondly, under the Mini-Batch options, the number of Epochs was fixed to a relatively large value when comparing models with different hyperparameters. In this work, the value of **MaxEpochs** typically range from 50 to 500, ensuring that the model does not get stuck in a local minimum. Moreover, the value of **MiniBatchSize** is a trade off between training speed and accuracy. In this work, the max value of **MiniBatchSize** was limited by the GPU memory, while the minimum value was limited by the available training time. Setting **MiniBatchSize**

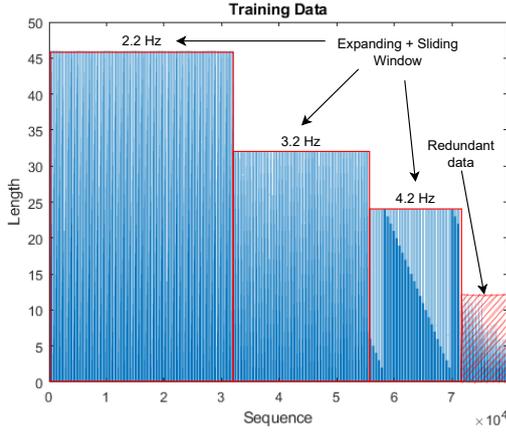


Figure 9: A histogram of a training data set generated by the Expanding-Sliding Window method for 2.2, 3.2, and 4.2 Hz. The redundant data generated by the Expanding window method does not contain any new information, but it helps improve model training speed on early targets by exposing the model to the data more than once during one training epoch.

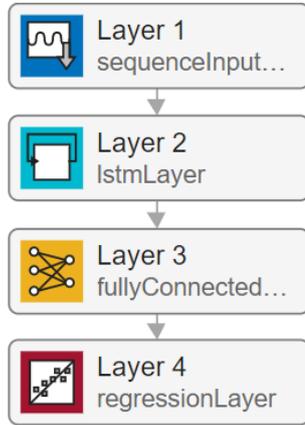


Figure 10: The proposed architecture of SSLSTM [13]. The layers are described in Tab. 1.

to a small value increases the training time and the regularization effect of mini-batches. Thus, the model generalizes better, resulting in a lower validation's Root Mean Square Error (RMSE). In addition, it was observed that shuffling the training data after each epoch helps to reduce the validation RMSE. Thus, the option **Shuffle** was set to "every-epoch".

Thirdly, for the Solver options the **InitialLearnRate**, **LearnRateDropPeriod**, and **LearnRateDropFactor** were set with the options **MaxEpochs** and taken **Shuffle** into account to give the best training performance. The strategy adopted was to decrease the learning rate for every few epochs as training data is shuffled. Also, for Gradient Clipping options, **GradientThreshold** is set to 1 instead of "inf" (default) to improve the training stability at high learning rate. This helps prevent gradient explosions and speeds up the training process.

Finally, from the Sequence options, **SequencePaddingDirection** was set to "left". This is because any padding done in the final time steps of the sequence can negatively impact the training process in a Sequence-to-one regression. For **SequenceLength**, the value depends on the method used for generating the

data. When the *Expanding Window* method was used, the sequence length was set to "longest". In this case, the sequences in each mini-batch are padded to the length of the mini-batch's longest sequence. However, when the *Sliding Window* method is used, the sequence length is set to the length of the longest possible sequence; it was set to the length of one period of the lowest frequency of interest. For example, if the lowest frequency considered is 0.1 Hz, and the step size is 10^{-3} , then **SequenceLength** is set to 10^4 .

6. Results and validation

In order to demonstrate SSLSTM's ability to learn multiple frequencies and decay rates, four simulation tests were conducted for $\beta \in \{0^\circ, 45^\circ, 90^\circ, 180^\circ\}$. In these tests, the SSLSTM was trained on the data range in Tab. 4 with a step size of 0.5×10^{-2} . Fig. 11, 12, 13, and 14 show the results of these tests. In the figures, SSLSTM's predictions match the ground truth (the ideal expected output) over a wide range of β , while SSNN's predictions only match when $\beta = 0$.

Table 4: Data resolution and range. ω : frequency, λ : decay rate, A : sine amplitude, β : phase-shift.

	Lower bound	Upper bound	Resolution
ω	2.2 Hz	4.2 Hz	3
λ	0	1	10
A	1×10^{-5}	0.1	10
β	0°	180°	9

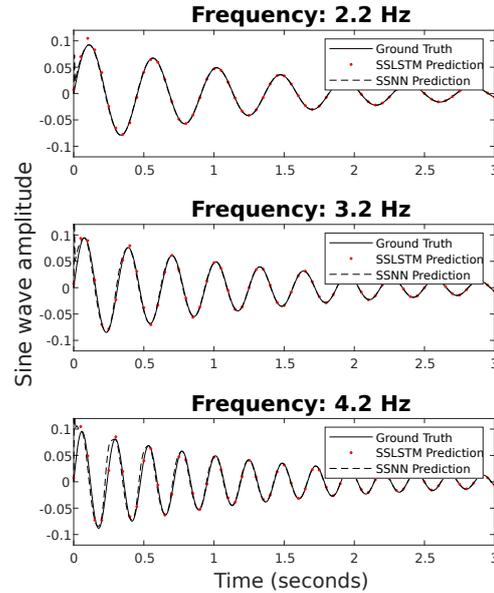
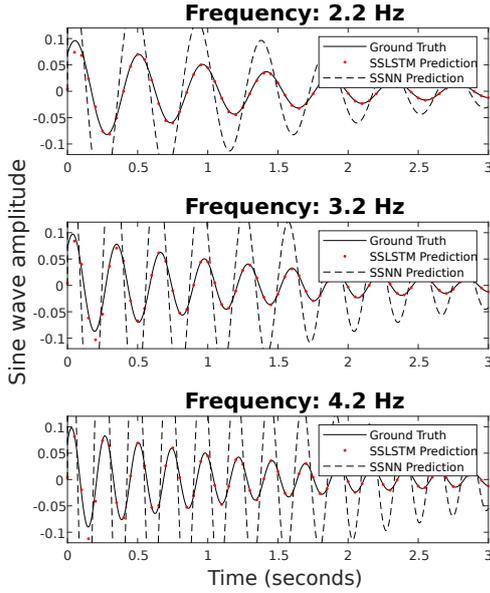
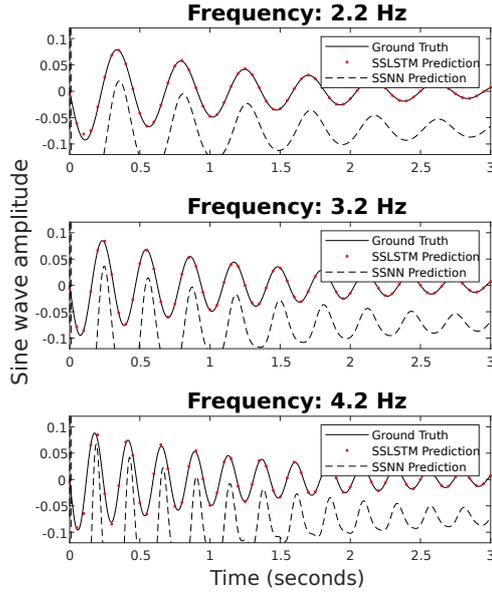
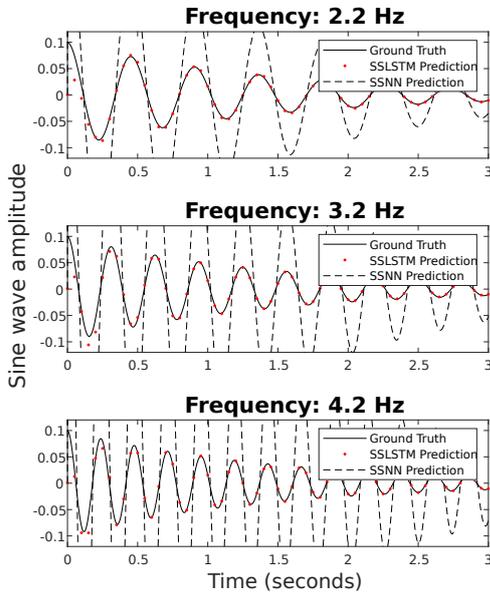


Figure 11: SSLSTM performance with $\beta = 0^\circ$.

The equations describing the dynamics of a synchronous generator during transient operation are quite stiff e.g., fast changing differential equations. Hence, often small step-sizes of Ode solvers are desired. In [6] the solver (**ode23tb variable-step**) step size was set to 3×10^{-5} . In this work, it was not possible to generate training data with a step size of 3×10^{-5} due to limited hardware resources. To validate and compare the performance of SSLSTM and SSNN, both models were trained with a step size of 10^{-3} . Fig. 15 shows the simulation results for SSLSTM and SSNN when a disturbance occurs after 4s of simulation time.

Figure 12: SSLSTM performance with $\beta = 45^\circ$.Figure 14: SSLSTM performance with $\beta = 180^\circ$.Figure 13: SSLSTM performance with $\beta = 90^\circ$.

7. Discussion

In this work, the neural network architecture and the training options were tuned to improve the model performance. These are summarized in Tab. 1, Tab. 2, and Tab. 3. Nevertheless, training LSTM models is computationally expensive compared to FNN models since the predictors of LSTM models are sequences. In this work, to train and tune the models in a reasonable amount of time, the training data was generated with the smallest possible step size and limited to the range of interest. Although it is possible to reduce the amount of data generated by increasing the step size to 10^{-2} and instead train the model on more frequencies and decay rates, it would decrease the controller performance on LFOs in the upper range. For example, a control action each 0.01s maybe sufficient to damp 0.1 Hz LFOs, but it might not be sufficient to damp 3 Hz LFOs effectively. A step size of 10^{-3} achieves the best possible model

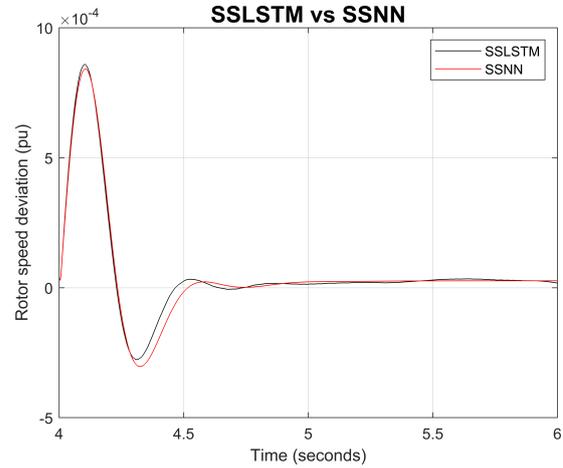


Figure 15: A comparison between SSLSTM and SSNN in damping the dominant mode of oscillation (3.2 Hz) of the machine under study. The disturbance occurs after 4s of simulation time.

accuracy and generalization capability in the range of interest while keeping model training and tuning time to a minimum. Also, besides the data resolution and step size, it is also possible to generate more observations by generating more periods per sine wave. When generating targets for the current period, it makes sense to discard data points from the previous period to save memory space. Thus, the *Sliding Window* method is preferred to generate predictor sequences for targets after the second period. It is also possible to increase the sliding window width to include more information in each predictor. While this would increase the model accuracy at the cost of increased memory usage, it will not improve the model generalization capability. Thus, this approach was not preferred in this work. It is also possible to improve the model accuracy at the cost of training time by increasing the number of epochs and reducing the mini-batch size.

In Fig. 11, 12, 13, and 14, SSLSTM's ability to learn multiple frequencies and decay rate was demonstrated. Moreover, In Fig. 15 SSLSTM and SSNN

performance was compared. Both controllers showed similar performance, which is to be expected since both controllers were trained on the LFO's frequency. However, SSLSTM shows a slightly worse performance, this is likely due to SSLSTM's lower prediction accuracy compared to SSNN's accuracy. Although SSLSTM's prediction accuracy is lower, it has a better generalization capability; unlike SSNN, it can phase shift sine waves with different frequencies and decay rates correctly.

8. Conclusion

In this paper, LSTM networks were used to develop a new sine-wave phase shifter for stability enhancements of electric power systems through the PSS. Simulation results show promising results and the main findings are:

- LSTM networks are capable of learning and tracking sine waves with multiple frequencies and decay rates.
- SSLSTM outperforms the SSNN controller at all frequencies except for the one SSNN was trained to phase shift.
- Training LSTM networks to learn periodic signals with a wide range of frequencies entails selecting the smallest step size to sample the highest frequency of interest while avoiding increasing the computational load significantly during model training. Training LSTM to learn long sequences, such as a 0.1 Hz LFO with a high sample rate, requires significant computing power.
- Carefully selecting the best method to generate the training data set can significantly improve the model performance.

Suggestions for future work:

- Develop an online adaptive PSS by combining SSLSTM and auto-tuning algorithms to adjust β and K_{PSS} during operation.
- Combine LSTM with other neural networks to improve prediction accuracy and reduce the amount of data required to train the LSTM network.

References

- [1] IEEE recommended practice for excitation system models for power system stability studies. *IEEE Std 421.5-2016 (Revision of IEEE Std 421.5-2005)*, pages 1–207, 2016. doi: 10.1109/IEEESTD.2016.7553421.
- [2] Zeyad Assi Obaid, L.M. Cipcigan, and Mazin T. Muhssin. Power system oscillations and control: Classifications and pss's design methods: A review. *Renewable and Sustainable Energy Reviews*, 79:839–849, 2017. ISSN 1364-0321. doi: <https://doi.org/10.1016/j.rser.2017.05.103>. URL <https://www.sciencedirect.com/science/article/pii/S1364032117307499>.
- [3] Ahmed A. Ba-muqabel and Mohammad A. Abido. Review of conventional power system stabilizer design methods. In *2006 IEEE GCC Conference (GCC)*, pages 1–7, 2006. doi: 10.1109/IEEEGCC.2006.5686203.
- [4] C.Y. Chung, K.W. Wang, C.T. Tse, X.Y. Bian, and A.K. David. Probabilistic eigenvalue sensitivity analysis and pss design in multimachine systems. *IEEE Transactions on Power Systems*, 18(4):1439–1445, 2003. doi: 10.1109/TPWRS.2003.818709.
- [5] Francisco P. Demello and Charles Concordia. Concepts of synchronous machine stability as affected by excitation control. *IEEE Transactions on Power Apparatus and Systems*, PAS-88(4):316–329, 1969. doi: 10.1109/TPAS.1969.292452.
- [6] Thomas Grong. Adaptive neural network-based PSS designs for modern power systems. Master's thesis, NTNU, 2020.
- [7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. ISSN 0899-7667.
- [8] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, 1995. doi: 10.1109/ICNN.1995.488968.
- [9] E. V. Larsen and D. A. Swann. Applying power system stabilizers part i: General concepts. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(6):3017–3024, 1981. doi: 10.1109/TPAS.1981.316355.
- [10] E.V. Larsen and D.A. Swann. Applying power system stabilizers part ii: Performance objectives and tuning concepts. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(6):3025–3033, 1981. doi: 10.1109/TPAS.1981.316410.
- [11] E.V. Larsen and D.A. Swann. Applying power system stabilizers part iii: Practical considerations. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(6):3034–3046, 1981. doi: 10.1109/TPAS.1981.316411.
- [12] Jan Machowski, Janusz W. Bialek, Jim Bumby, and Dr Jim Bumby. *Power System Dynamics : Stability and Control*. John Wiley & Sons, Incorporated, New York, UNITED KINGDOM, 2008. URL <http://ebookcentral.proquest.com/lib/ucsn-ebooks/detail.action?docID=406510>.
- [13] MATLAB. *version 9.11.0.1884302 (R2022a)*. The MathWorks Inc., Natick, Massachusetts, 2022. URL <https://www.mathworks.com/>.
- [14] Patricia Melin, Oscar Castillo, and Janusz Kacprzyk. *Nature-Inspired Design of Hybrid Intelligent Systems*, volume 667. Springer, 2017. ISBN 9783319836508.
- [15] G Shahgholian. REVIEW OF POWER SYSTEM STABILIZER: APPLICATION, MODELING, ANALYSIS AND CONTROL STRATEGY. *IJTP Journal*, 2013.
- [16] Statnett. Nasjonal veileder for funksjonskrav i kraftsystemet. Technical report, 2021. URL <http://www.statnett.no/Fosweb>.
- [17] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7):1235–1270, 2019. ISSN 0899-7667.
- [18] Junbo Zhang, C. Y. Chung, and Yingduo Han. A novel modal decomposition control and its application to pss design for damping interarea oscillations in power systems. *IEEE Transactions on Power Systems*, 27(4):2015–2025, 2012. doi: 10.1109/TPWRS.2012.2188820.

Unsteady flow analysis of hinged and sliding door openings

Fredrik Kihlberg*, Øystein Formo Hermansen, Arnab Chaudhuri

Department of Civil Engineering and Energy Technology

OsloMet – Oslo Metropolitan University, Oslo, Norway

*fre.kihlberg@gmail.com

Abstract

Existing thermal condition and indoor air quality have a big impact on our work performance, comfort, and health in an indoor environment. Apart from many other parameters, door motions and human movements play crucial role in mass and thermal exchange affecting safety and/or energy management issues in various situations. An isolation room in a hospital setup, for instance, helps to protect patients and staff against the risk of infection by airborne pathogens. Another example is cold storage room facilities, where temperature and moisture control are the key parameters for an optimal operation and energy usage. In this study, we present a transient flow analysis of door motions in indoor environment. The flow physics is resolved by solving 3D compressible RANS (Reynolds-averaged Navier-Stokes) equations together with the energy and species transport equations and two-equation turbulence models utilizing an overset mesh strategy to address the rigid body motion of doors in a relevant fluid domain involving air and sulfur hexafluoride (SF_6). Simulations are performed for three different types of doors, namely a hinged door, a two-way sliding door, and a sliding door considering door opening and closing phases. Transient flow-field data through the door opening area have been processed and a comparative analysis is performed considering the mass flux of the constituents, normal velocity, cumulative mass exchange through the different doors.

1. Introduction

We spend the majority of the time in indoor environment and our work performance, comfort and health are heavily impacted by the indoor temperature and air quality. In certain special environments, such as hospital isolation rooms and clean rooms, it is particularly important to understand the process of pollutant transmission. In healthcare settings, isolation rooms are used to contain infectious patients or to protect vulnerable patients from infection. On the other hand, in cold storage room facilities, for instance, the temperature and moisture control are the key parameters for an optimal operation and energy usage. These are widely applicable for industries like food, pharmaceutical, photographic and more. Multiple coupled interactions, involving heat-mass-momentum transfer and phase change of constituent components play important roles during the operating condition of a cold storage room in such applications. Among several other factors, infiltration of hot and moist air through open doors become crucial for the infiltration load and the performance of buildings. Often, a big difference in temperature exists between the adjacent rooms and this may alter due to the door opening. In these scenarios, understanding of air flow, mass/heat exchange by the opening of a door is beneficial for control strategies. The effect of door motion is not only relevant for special rooms. It is also very relevant for example in educational and office buildings. It can be realized that the analysis of the combined effect of human motion along with and the door openings are necessary often for predicting realistic situations. In their study, Tang et al. [1] emphasized that there is likely to be some leakage across the doorway to a lesser or greater degree as a human moves through

the door at a reasonable walking speed affecting the flow physics. Shao et al. [2] and Kalliomäki et al. [3] have drawn similar conclusions in their experimental work. In the following paragraph, we provide a brief account of both experimental and numerical studies of the literature regarding this field.

Several studies show that door opening motion generates a notable air exchange and airborne contaminant transfer across a doorway. This effect is especially notable for a hinged door [1, 3–6]. Also, the effect is larger when a person passes through a doorway. The amount of contaminant transfer is influenced by door hold open time and temperature difference. Tang et al. [7] investigated a clinical situation where a severe case of adult chickenpox was managed in a negative pressure isolation room, with no adjacent anteroom. Previous studies have also shown that the hinged doors allow more mass exchange than sliding doors. In the similar context, Eames et al. [8] have estimated motion and diffusion of a contaminant in an isolation room, in the absence of differential pressure. Several authors performed experimental studies, often with smoke as tracer fluid or particle generators to simulate transmission, e.g. small scale model of adjacent rooms [9], connected rooms [10], clean room with double hinged door [2]. In this regard, Hathway et al. [11] performed a field and small scale study of rooms with hinged type doors in office and hospital settings. They have used water as the tracer fluid in their small scale model showing that there exists a linear relationship between time doors are held open and air volume flux. On the other hand, the experimental works presented in [3, 12–14] considered full scale models and realistic setups. In the context of

the recent pandemic situation, Bhattacharya et al. [15] investigated positively pressurized room with a physical model. The goal of their study was to examine the possibility of the COVID19 virus, contaminating clean areas due to the door opening and closing motion. Apart from experimental approach, Computational fluid dynamics (CFD) simulations are widely used to resolve detailed, complex flow dynamics associated with the indoor environment. Numerical approach via CFD solutions are increasingly used nowadays in design, optimization of ventilation systems, and the prediction of air movement in ventilated spaces. Several studies [4, 16–22] reported CFD analysis of the effect of the door motion in various scenarios. Among these [18, 19, 22] analyzed cold storage facilities and Zhang et al. [21] studied refrigerated vehicles via numerical approach. Note that the experimental work of Tang et al. [1] revealed that sliding doors (single or double) are advantages over the more conventional hinged-door towards general infection control purposes. However, systemic numerical studies addressing a comparison of different type of door motions are not abundant.

In this work, we revisit and explore the effect of door opening motion considering three different types of door by numerical approach. To address this, we solve 3D compressible RANS equations together with the energy and species transport equations and two-equation turbulence models are used to account the effects of turbulence. An overset mesh strategy is utilized to resolve the rigid body motion of doors in a relevant fluid domain involving air and SF₆ as working fluids. The article is organized as follows. First, we present the governing equations and numerical setup for the different cases in section 2. The description of the problem setup is provided in section 3, followed by results and discussions in section 4. Finally, the conclusions are drawn in section 5.

2. Method

2.1. Governing equations

The multi-component gas system consisting of air and SF₆ is governed by 3D compressible Navier-Stokes system of equations together with mass, species conservation and energy conservation equations. For any conserved property ($\rho\phi$), the general transport equation can be expressed as in the following standard notation [23]:

$$\frac{\partial \rho\phi}{\partial t} + \vec{\nabla} \cdot \rho\phi\vec{V} = \vec{\nabla} \cdot \Gamma\vec{\nabla}\phi + S_\phi, \quad (1)$$

where ρ is the density, \vec{V} is the velocity and Γ is the diffusion coefficient. The first and second terms in the left hand side are known as the temporal term and the advection term respectively. The right hand side consists of the diffusion term and the generation term (S_ϕ). We solve the unsteady Reynolds averaged Navier-Stokes (RANS) formulation and Realizable $\kappa - \epsilon$ model is used for turbulence. The multi-component gas mixture is assumed to be Newtonian fluid and follow ideal gas law. The finite volume method (FVM) based commercial CFD tool StarCCM+ is used to perform the simulations. An overset mesh strategy is employed to address the rigid body motion of the doors in the fluid domain. A background mesh (throughout the computational domain) and overset mesh (section of domain, including door and

a surrounding region of choice) are generated for this purpose. The time varying rigid body rotation/translation motion is applied according to the desired door motion by setting the motion parameters associated with the door geometry. The overall setup of the different test cases is described in section 3. The brief description of the numerical schemes used in the present study is as follows. Second order upwind schemes for convection, hybrid Gauss least-squared gradient method based 2nd order schemes for diffusion and Venkatakrishnan limiter function are chosen in solver setup. Second order implicit scheme (Newmark method) is used for time integration with a time step $\Delta t = 0.005$ s. Algebraic Multi-Grid (AMG) techniques are also invoked with the setup mentioned above. In Star-CCM+, under physics continua, the following models were chosen: coupled energy, coupled flow, coupled species, gradients, ideal gas, implicit unsteady, k-epsilon turbulence, multi-component gas (with air and SF₆ as gas components), non-reacting, overset conservation, realizable k-epsilon two-layer, Reynolds-averaged Navier-Stokes, solution interpolation, three dimensional, turbulent, two-layer all y+ wall treatment, and wall distance. Standard values were used for all of these, excluding multi-component gas composition. The gas mixture composition is assigned as desired and is given in section 3.

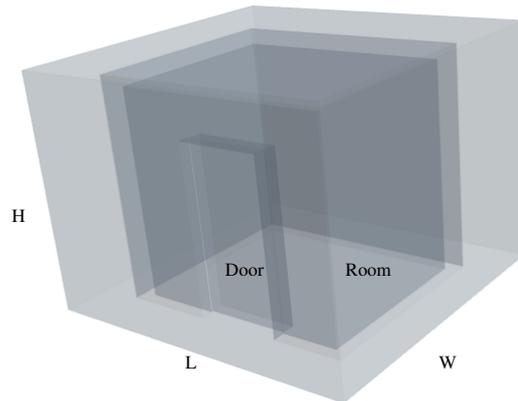
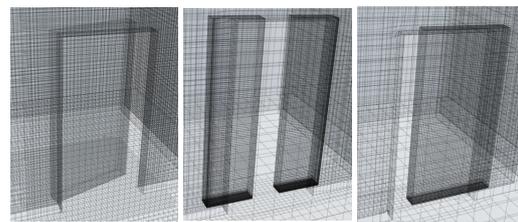


Figure 1: 3D view of the computational domain.



(a) hinged (b) two-way sliding (c) sliding
Figure 2: Partially opened doors.

Table 1: Dimensions of the model

	W (m)	L (m)	H (m)
Boundary	3.60	3.80	2.75
Outer wall	2.405	2.85	2.6
Inner wall	2.105	2.55	2.45
Two-way sliding door	0.15	0.5	1.98
Single door	0.15	1.00	1.98

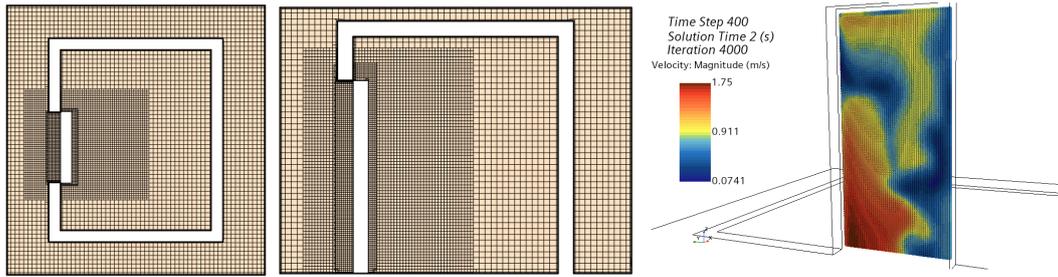


Figure 3: left: mid x-y plane (top view), middle: mid x-z plane (side view), right: presentation grid (y-z plane) at $t = 2.0s$

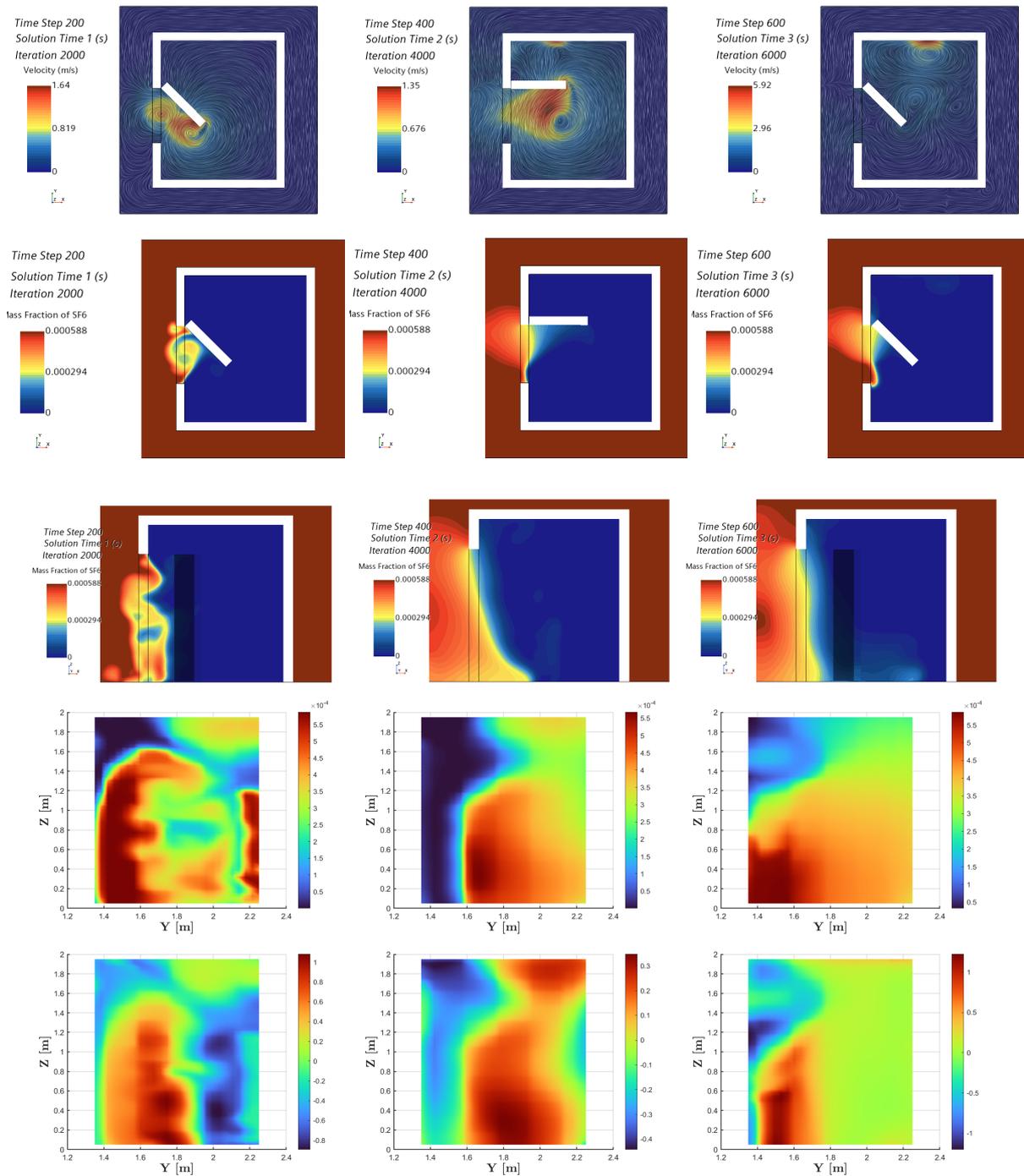


Figure 4: Flow visualization of Case-1, left: $t = 1s$, middle: $t = 2s$, right: $t = 3s$.

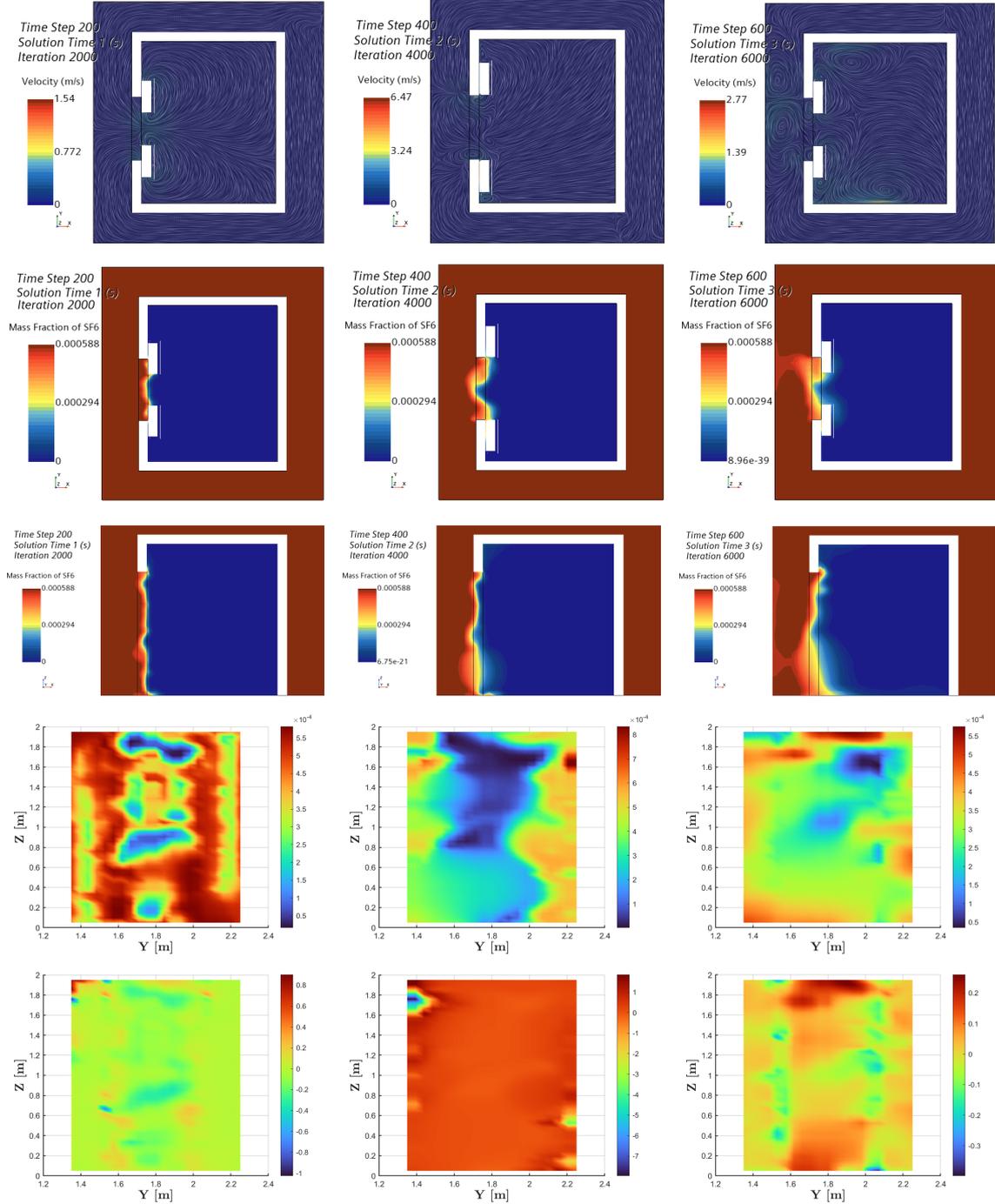


Figure 5: Flow visualization of Case-2, left: $t = 1\text{s}$, middle: $t = 2\text{s}$, right: $t = 3\text{s}$.

3. Problem setup

A room with a door and a simplified connected outer corridor type area is made as computational domain (see the schematic in figure 1). Table 1 summarizes the Length (L), Width (W) and Height (H) of different components of the computational domain. All walls of the room are considered adiabatic. The initial condition of the fluid domain is set as: temperature $T = 300\text{K}$, velocity field $\vec{V} = 0$, pressure $P = 101325\text{Pa}$, turbulent kinetic energy $k = 1.0e^{-3} \text{ J/kg}$ and turbulent dissipation rate

$\epsilon = 1.0e^{-6} \text{ m}^2/\text{s}^3$. The mass fraction of the SF₆ in the outer region is set as $Y_{\text{SF}_6} = 0.00058$. This corresponds to the concentration of SF₆ to be 706.2 mg/m³. The room is considered as filled with pure air and devoid of any SF₆ at the initial state. The initial condition of mass fraction of SF₆ is similar to that reported by Chang et al. [4].

In this work, we consider three different type of doors, namely hinged, two-way sliding (also known as double sliding) and sliding. Figure 2 shows orientation of the partially opened doors for all three cases. The door

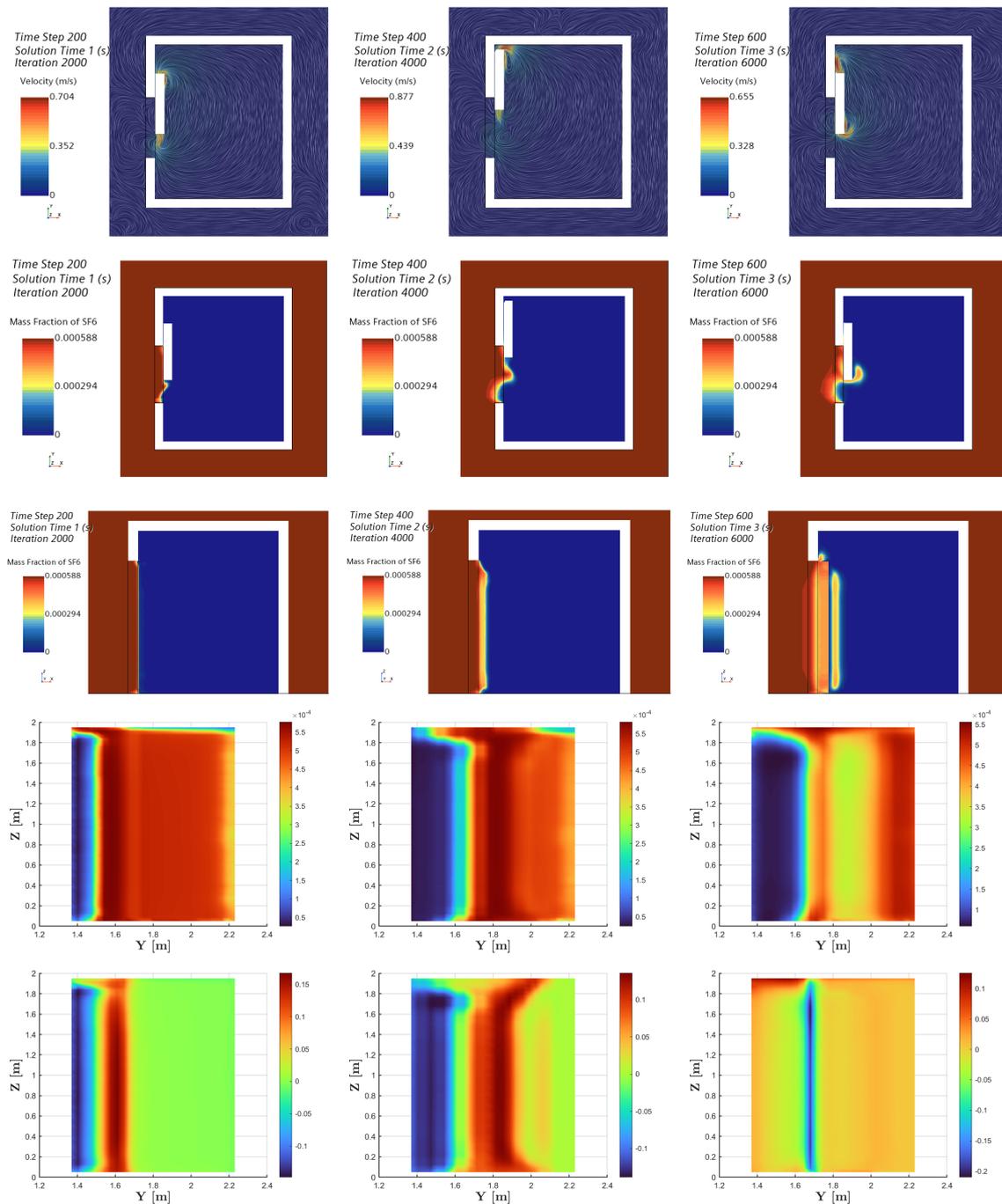


Figure 6: Flow visualization of Case-3, left: $t = 1$ s, middle: $t = 2$ s, right: $t = 3$ s.

opening-closing motions are assigned to different types of doors as follows. For hinged door, we set the opening with -0.785 radian/s during $0 < t \leq 2$ s and closing with $+0.785$ radian/s when $2s < t \leq 4$ s. In case of the two-way sliding door, the opening is set to ± 0.25 m/s during the initial 2s and the closing is set for next 2s with ∓ 0.25 m/s. Similarly, the door motion of sliding door is set as opening with $+0.5$ m/s during initial 2s and closing with -0.5 m/s when $2s < t \leq 4$ s. Suitable user defined field functions are assigned to apply these door motions and to set initial mass fraction distribution in the computational domain. Simulations are executed till the final physical time reaches 4 second for all cases.

The mesh for the simulations consists of a background mesh and an overset mesh around the moving door. Additionally, we use refined mesh near the entrance of the door to resolve the flow physics. Note that the refinement zone varies depending upon the type of the door. The background mesh size was set to 20cm, the overset mesh size was set to 2cm, and the refinement mesh size was set to 4cm. The background mesh and the overset mesh region had two prism layers, while the refinement mesh region had five prism layers. For this, we have used the "Trimmer" mesh model of StarCCM+. Figure 3 illustrates the mesh setup for the hinged type door. The total number of cells (background cells together

with the cells in the overset region) are 673221, 400539, and 529740 for hinged door, sliding door and two-way sliding door respectively. A suitable “derived part” called presentation grid is generated at the location of door to record time varying solution data and analyze the mass exchange through the door opening. This is shown in figure 3 as an example with hinged door. The presentation grid contains 20×40 data points for post processing giving a resolution of $\Delta y = \Delta z \approx 0.05\text{m}$.

4. Results and discussions

We define Case-1 as hinged door, Case-2 as two-way sliding door and Case-3 as sliding door henceforward. We will present the results of each case with line integral convolution (LIC) of \vec{V} and contour plots of other field variables mainly in three different planes. As shown in figure 3, these are the mid x-y plane (top view), mid x-z plane (side view) and presentation grid (y-z plane). Figures 4, 5 and 6 are for Case-1, Case-2 and Case-3 respectively. Note that these plots are presented systemically as follows: topmost row with LIC of \vec{V} (top view), second row from the top with Y_{SF_6} (top view), third row from the top with contours of Y_{SF_6} (side view), second row from the bottom with contours Y_{SF_6} (presentation grid i.e., door frame) and bottom row with contours of u-component of \vec{V} (presentation grid i.e., door frame).

4.1. Flow evolution and general characteristics for different types of doors

The time evolution of door opening and closing for Case-1 is depicted in figure 4. First, it can be clearly seen from the LIC of \vec{V} that quite large vortex structures exist during the flow evolution. The hinged door makes an angle of $\approx 45^\circ$ at 1s during the opening phase, $\approx 90^\circ$ fully open at 2s and in the closing phases at 3s it again makes an angle of $\approx 45^\circ$. The complex, highly three dimensional shear/mixing layers are clearly seen from the distribution of mass fraction of SF_6 in x-y, x-z and y-z planes revealing a substantial extent of mixing of SF_6 from outside into the room. The u-component of the velocity field is normal to the door opening plane. Note that a positive u-component is associated with the normal velocity field that is entering to the room. On the other hand, a negative u-component is associated with the normal velocity field that is leaving the room. The contours of u-component match well with the contours of Y_{SF_6} .

Figure 5 illustrates the flow development of Case-2. Note that, similar to the hinged door, two-way sliding door completely opens at 2s and closes at 4s. The door remains approximately half open at 1s and 3s during the opening phase and during the closing phase respectively. Noticeably, the vortex structures are entirely different than a hinged door. Also, it is clear that the penetration of mixing due the door induced movements are relatively less in this case compared to Case-1. Nevertheless, the contours of Y_{SF_6} and u-component of the \vec{V} at the door frame reveal the existing complex three dimensional nature of the mixing process. One could realize that the shear-layer instability induced by the door motion in Case-2 is different than Case-1.

The results for Case-3 are depicted in figure 6. Note that, due to the existing dimension of the room the

sliding door is not completely open ($\approx 80\%$ open). The mixing behavior shows some two-dimensional nature in the middle region of the door opening. These can be seen from the contour structures of Y_{SF_6} and u-component of the velocity (see y-z planes). The three dimensional nature is visible in the upper and lower parts of the door opening area. This corroborates the mixing layer behavior visible at x-y plane (second row from the top of the figure 6). The contours of u-component of the velocity (positive and negative patches) are in accordance with the contours of Y_{SF_6} .

4.2. Quantitative comparison of different doors

In order to quantify the mass of SF_6 exchanged through the door opening area one need to estimate first the surface averaged flow rate of SF_6 ,

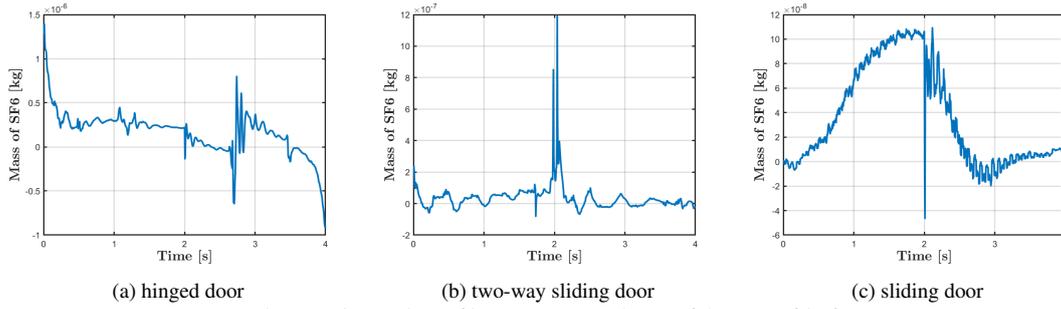
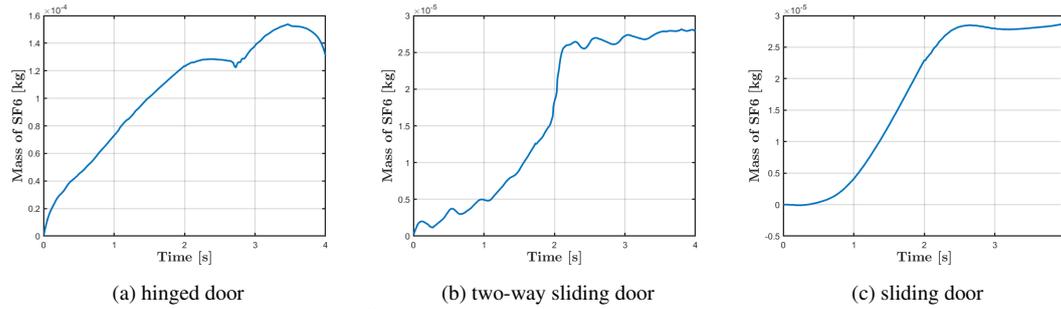
$$\dot{M}_{\text{SF}_6} = \iint_A (\rho Y_{\text{SF}_6} \vec{V} \cdot \vec{n} dA) \quad (2)$$

where the A represent the door opening area. The instantaneous mass of SF_6 exchanged can thus be calculated by $M_{\text{SF}_6}(t) = \dot{M}_{\text{SF}_6} dt$. The transient data stored via presentation grid mentioned before is used to compute this. We compare $M_{\text{SF}_6}(t)$ and the cumulative exchange of mass of SF_6 exchanged through the door opening for all cases. Figure 7 shows the instantaneous exchange of mass of SF_6 through the door opening for each case. The comparison of the cumulative exchange of mass of SF_6 through the door opening is presented in figure 8.

Table 2: Comparison of mass exchange at $t = 4\text{s}$.

	Cumulative mass of SF_6 exchanged (kg)
Case-1	$\approx 1.30e^{-4}$
Case-2	$\approx 2.79e^{-5}$
Case-3	$\approx 2.87e^{-5}$

It can be seen from figure 7a that during the door opening phase of Case-1, the mass of SF_6 through the door remains mostly positive meaning the net influx into the room. On the other hand, during the closing phase of 2s to 4s the mass exchange changes sign after $\approx 3.5\text{s}$. This passes through an oscillatory behavior around 2.7-2.8s. In figure 8a, the cumulative exchange of mass of SF_6 reflects these behaviors. The opening phase of the door opening shows almost linear growth, and a local saturation of the mass exchange occurs between 2-3s. The cumulative mass exchange drops to $1.30e^{-4}$ kg at 4s after a global maximum of $\approx 1.52e^{-4}$ kg. Case-2 and Case-3 show local spikes at the instance of change the door motion from opening to closing direction. Similar spikes (not shown for brevity) are observed of the average u-component of the velocity field through the doors at those instances for Case-2 (0.51 m/s) and Case-3 (-0.044 m/s) accordingly. Our preliminary mesh dependency study considering 4cm, 3cm and 2cm refined mesh sizes in the door interaction zone reproduces this behavior for Case-3. However, note that a detailed mesh sensitivity analysis is not performed taking the complete computational domain. Nevertheless, the behavior of the instantaneous mass exchange of Case-2 is different than Case-3. During the door opening phase, we found a relatively smooth non-linear growth in the cumulative mass exchange for Case-3. At the final stage (4s), the values for both Case-2 and

Figure 7: Comparison of instantaneous exchange of the mass of SF₆.Figure 8: Comparison of cumulative exchange of mass of SF₆.

case-3 yielded comparable magnitudes. However, these are an order of magnitude lower than Case-1. Table 2 summarizes the cumulative mass exchange of SF₆ through the different doors. Table 3 shows the CPUh required for each simulation.

Table 3: Computing time	
Central Processing Unit hours (CPUh)	
Case-1	≈ 145
Case-2	≈ 66
Case-3	≈ 68

5. Conclusions

In this study, we presented a systematic analysis of mass exchange through the door opening and closing scenarios for three different types of doors. The complex unsteady flow physics involves multi-component gaseous system with rigid body motion of the doors. We solved 3D compressible RANS equations together with the energy and species transport equations. The effects of turbulence are resolved by two-equation turbulence models. An overset mesh strategy is employed to address the rigid body motion of doors in a relevant fluid domain consisting of air and SF₆ s working fluids. Simulations are made until the physical time reaches 4 second. The initial 2s is meant for door opening while door closing phase is performed during the last 2s. The detailed flow dynamics of each test cases have been illustrated with the contours of flow-field variable. The preliminary findings of the comparison reveal that under the similar condition the overall cumulative exchange of mass is SF₆ is about an order of magnitude less for both two-way sliding door and sliding door when compared with hinged door (≈ 4.6 times less). Matching with previous findings the case with hinged door is found be associated with most complex 3D

mixing layers with maximum mass exchange. With the present setup, we observed similar mass exchange for two-way sliding and sliding door.

The authors would like to emphasise that the present study lacks comparison and validation of these preliminary findings with experimental data. We will consider the limitations of the present study in our subsequent work. The future work will also consider systematic numerical analysis of thermal effects, ventilation strategies, human motion in realistic indoor environment with different scenarios, different door and room size ratios.

Acknowledgment

This work is compiled and expanded from a part of the Master thesis work of the first author. The authors greatly acknowledge the commercial computer program StarCCM+, used in this study. The results are obtained by the licensed version 2022.1 17.02.007 of StarCCM+. The work utilized standalone Windows machines as well as the super-computing facility of Fram computer cluster (UiT Arctic University of Norway) under the project NN8005K at Notur, UNINETT Sigma2 AS (National infrastructure for scientific computational science in Norway).

References

- [1] J. W. Tang, A. Nicolle, J. Pantelic, C. A. Klettner, R. Su, P. Kalliomäki, P. Saarinen, H. Koskela, K. Reijula, P. Mustakallio, D. K W Cheong, C. Sekhar, and K. Wai Tham, "Different types of door-opening motions as contributing factors to containment failures in hospital isolation rooms," *PLoS One*, vol. 8, no. 6, p. e66663, 2013.
- [2] X. Shao, K. Hashimoto, L. Fang, A. K. Melikov, K. G. Naydenov, and C. Rasmussen, "Experimental study of airborne particle transmission through the doorway of a cleanroom due to the movement of a person," *Building and Environment*, vol. 183, p. 107205, 2020.
- [3] P. Kalliomäki, P. Saarinen, J. W. Tang, and H. Koskela, "Airflow patterns through single hinged and sliding doors

- in hospital isolation rooms – effect of ventilation, flow differential and passage,” *Building and Environment*, vol. 107, pp. 154–168, 2016.
- [4] L. Chang, X. Zhang, S. Wang, and J. Gao, “Control room contaminant inleakage produced by door opening and closing: Dynamic simulation and experiments,” *Building and Environment*, vol. 98, pp. 11–20, 2016.
- [5] I. G. Papakonstantis, E. A. Hathway, and W. Brevis, “An experimental study of the flow induced by the motion of a hinged door separating two rooms,” *Building and Environment*, vol. 131, pp. 220–230, 2018.
- [6] P. Kalliomäki, P. Saarinen, J. W. Tang, and H. Koskela, “Airflow patterns through single hinged and sliding doors in hospital isolation rooms,” *International Journal of Ventilation*, vol. 14, no. 2, pp. 111–126, 2015.
- [7] J. W. Tang, I. Eames, Y. Li, Y. Taha, P. Wilson, G. Bellingan, K. Ward, and J. Breuer, “Door-opening motion can potentially lead to a transient breakdown in negative-pressure isolation conditions: the importance of vorticity and buoyancy airflows,” *Journal of Hospital Infection*, vol. 61, no. 4, pp. 283–286, 2005.
- [8] I. Eames, D. Shoaib, C. Klettner, and V. Taban, “Movement of airborne contaminants in a hospital isolation room,” *Journal of the Royal Society Interface*, vol. 6, no. suppl_6, pp. S757–S766, 2009.
- [9] L. Fontana and A. Quintino, “Experimental analysis of the transport of airborne contaminants between adjacent rooms at different pressure due to the door opening,” *Building and Environment*, vol. 81, pp. 81–91, 2014.
- [10] J. Hendiger, M. Chludzińska, and P. Zietek, “Influence of the pressure difference and door swing on heavy contaminants migration between rooms,” *PLoS One*, vol. 11, no. 5, pp. e0155159–e0155159, 2016.
- [11] A. Hathway, I. Papakonstantis, A. Bruce-Konuah, and W. Brevis, “Experimental and modelling investigations of air exchange and infection transfer due to hinged-door motion in office and hospital settings,” *The International Journal of Ventilation*, vol. 14, no. 2, pp. 127–140, 2015.
- [12] P. Kalliomäki, K. Hagström, H. Itkonen, I. Grönvall, and H. Koskela, “Effectiveness of directional airflow in reducing containment failures in hospital isolation rooms generated by door opening,” *Building and Environment*, vol. 158, pp. 83–93, 2019.
- [13] E. S. Mousavi, R. Jafarifiroozabadi, S. Bayramzadeh, A. Joseph, and D. San, “An observational study of door motion in operating rooms,” *Building and Environment*, vol. 144, pp. 502–507, 2018.
- [14] A. Bhattacharya, A. R. Metcalf, A. M. Nafchi, and E. S. Mousavi, “Particle dispersion in a cleanroom - effects of pressurization, door opening and traffic flow,” *Building Research and Information: The International Journal of Research, Development and Demonstration*, vol. 49, no. 3, pp. 294–307, 2021.
- [15] A. Bhattacharya, A. Ghahramani, and E. Mousavi, “The effect of door opening on air-mixing in a positively pressurized room: Implications for operating room air management during the covid outbreak,” *Journal of Building Engineering*, vol. 44, p. 102900, 2021.
- [16] L. Wang, L. Zhang, and G. Lian, “A cfd simulation of 3d air flow and temperature variation in refrigeration cabinet,” *Procedia Engineering*, vol. 102, pp. 1599–1611, 2015.
- [17] S. Lee, B. Park, and T. Kurabuchi, “Numerical evaluation of influence of door opening on interzonal air exchange,” *Building and Environment*, vol. 102, pp. 230–242, 2016.
- [18] J. Gonçalves, J. Costa, and A. Lopes, “Analysis of the air infiltration through the doorway of a refrigerated room using different approaches,” *Applied Thermal Engineering*, vol. 159, p. 113927, 2019.
- [19] S. Tian, Y. Gao, S. Shao, H. Xu, and C. Tian, “Numerical investigation on the buoyancy-driven infiltration airflow through the opening of the cold store,” *Applied Thermal Engineering*, vol. 121, pp. 701–711, 2017.
- [20] B. Zhou, L. Ding, F. Li, K. Xue, P. V. Nielsen, and Y. Xu, “Influence of opening and closing process of sliding door on interface airflow characteristic in operating room,” *Building and Environment*, vol. 144, pp. 459–473, 2018.
- [21] X. Zhang, J. Han, J. Qian, Y. Wang, L. Wang, and X. Yang, “Computational fluid dynamic study of thermal effects of open doors of refrigerated vehicles,” *Journal of Food Process Engineering*, vol. 41, no. 3, pp. e12662–n/a, 2018.
- [22] A. Ayarmal, O. Melhus, and A. Chaudhuri, “The effects of fan and door opening on a cold storage room: a numerical study,” *Linköping Electronic Conference Proceedings*, 2018.
- [23] H. K. Versteeg and W. Malalasekera, *An introduction to computational fluid dynamics: the finite volume method*. Pearson Education, 2007.

Evaluation of Complex Spray Behaviors of Sprinkler Spray Using FDS

Rajath Ramachandran, Joachim Lundberg*

*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway.
joachim.lundberg@usn.no*

Abstract

It takes money, time, and energy to set up an experimental grid to measure the effectiveness of fire suppression parameters. Therefore, computational fluid dynamics (CFD) is an alternative in all fields as consequences modeling. Fire Dynamic Simulator (FDS) is a CFD software developed by the National Institute of Standards in Technology (NIST) to model and generate the results for the spray models. FDS uses Large Eddy Simulation (LES) to represent turbulence. The current study utilizes FDS to investigate the extinguishing efficiency of sprinkler spray on a general fire. The study focuses on analyzing the effectiveness of suppression parameters using complex (polydisperse) in contrast to the simplified (monodisperse) representation of the spray with- and without a fire scenario of a 2560 kW propane fire. Measurements were taken by digitally enabled Phase Doppler Particle Analyzer (PDPA) to measure the fire suppression parameters such as number concentration, droplet size distribution (DSD) & velocity distribution. The measurements were taken 1.5 m downstream of the sprinkler. The suppression parameters are compared with monodisperse and polydisperse with and without fire. Thus, the suppression parameters have been compared to measure the effect.

1. Introduction

In domestic and industrial firefighting, the halon was widely used as a fire suppression agent for fixed installations. The advantage was the ability to extinguish fires without destroying the equipment or surroundings of the fire. However, this agent was banned due to its unfriendly nature to the environment. Thus, water was introduced as a fire suppression agent. Water mist and deluge (sprinkler) systems are vital and economical to use in industry, offshore sites, domestic buildings, and other general commercial structures. The effectiveness of these fire suppression systems is the central topic of discussion in the field.

The effectiveness of the fire suppression systems can be done by conducting experiments. However, conducting experiments demands funds, time, risk, and workforce. Alternatively, the experiments can be replaced by Computational Fluid Dynamics (CFD) using Fire Dynamic Simulator (FDS) by NIST. FDS was designed to simulate thermally driven flows within buildings and uses the most straightforward rectilinear numerical grid. FDS is a Large Eddy Simulation (LES) model and prefers uniform meshing. In FDS, the water droplets are represented by Lagrangian particles.

Many research scholars digitally enabled PDPA to quantify the effectiveness of the suppression parameters. For example, DesJardin et al. (2000) studied the effect of water spray suppression on

large-scale pool fires by modeling in VULCAN based on KAMELEON-Fire code. VULCAN uses RANS model along with the $k-\epsilon$ turbulence model. The water mist spray was modeled in a 3.04 m closed entity investigating the temperature. The main findings indicate a strong sensitivity of fire suppression to initial droplet size, where injection of large droplets may cause an increase in overall temperature. Myers and Marshall (2016) and Liu et al. (2022) developed an Euler-Lagrangian representation of sprinkler spray using CFD with the results of Ren et al. (2011) and Zhou et al. (2012). Further, Liu et al. (2022) modeled hot air jets using FDS. Sheppard (2002) conducted full-scale experiments on characterizing the suppression parameters of spray like water droplet size, droplet velocity, and water flux using Particle Image Velocimetry (PIV) and Phase Doppler Anemometry (PDA) which is also known as Phase Doppler Particle Analyzer (PDPA). Bourque and Svirsky (2013) verified the inputs from Sheppard (2002) using FDS version 6. The main findings were water flux, droplet size, and velocity, which showed a decent result upon comparison. Sæbø and Wighus (2009) studied the droplet size distribution on a high-velocity sprinkler and medium-velocity nozzle. The full-scale experiment was conducted, and PDA and imaging techniques were used to check the Volume Median Diameter (VMD). The main findings showed that the imaging technique could

capture the large droplets in contrast to the PDA technique. Lundberg (2015) conducted the full-scale experimental characterization study of a sprinkler spray at 2.0 bar (g), 5.0 bar (g), and 8.0 bar (g). The research provided experimental data on droplet size and velocity distribution using shadow-image technique.

The current study utilizes FDS to investigate the effect of detailed water spray characteristics in contrast to simplified monodisperse droplet size distribution for a fire water sprinkler spray. FDS uses a combination of log-normal and Rosin-Rammler droplet size distribution as default. However, monodisperse can be selected. Using monodisperse droplet size distribution, the complexity of the simulation decrease, but the physical behavior might not get preserved. The current study uses input data from an experimental characterization study of a sprinkler spray performed by Lundberg (2021). The operating pressure is 2.0 bars with a Sauter mean diameter (SMD) of 419 μm . The mean diameter is assumed to give similar behaviour as the volume median diameter in the droplet model in FDS.

The behavior is presented with and without a fire to benchmark the effect of polydisperse in contrast to monodisperse droplet size distribution in FDS.

The fire is a propane fire with a Heat Release Rate Per Unit Area (HRRPUA) of 4000 KW/m^2 with a 0.64 m^2 fire vent area.

2. Methods

A medium velocity deluge nozzle (Tyco MV34-110) with a K-factor of 58.8 $\text{L}/(\text{min} \times \sqrt{\text{bar}})$ is used in the current study to represent the sprinkler spray. A line 1.5 meters perpendicular to the nozzle is used as a reference location for the behavior of the sprays. The Sauter mean diameter of Lundberg (2021) is used as Volume median diameter (VMD) and input to FDS as DIAMETER. Figure 1 shows the simulation matrix of the study for different scenarios.

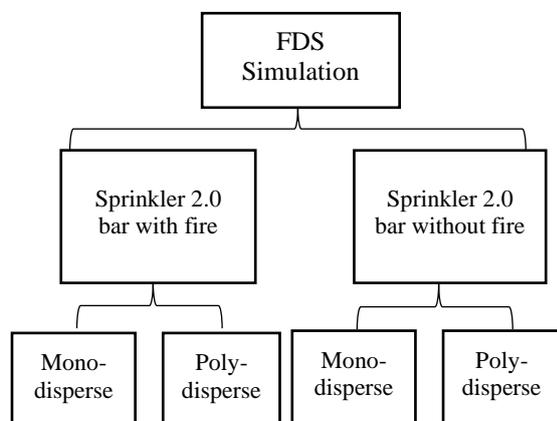


Figure 1: Simulation matrix

FDS utilizes a digital PDPA technique to quantify the suppression factors at various levels down the sprinkler spray. A rectangular closed entity of 6.0 m, 6.0 m, and 4.0 m comprises the simulation domain. In FDS the code was written as follows, &MESH IJK= 60, 60, 40, XB=-3, 3, -3, 3, -2, 2.

The computational domain was discretized uniformly of 10 cm cell size. The cell size was selected based on the FDS user guide (McGrattan et al., 2021). with a fire area of 0.64 m^2 with a heat release rate of 4000 KW/m^2 .

2.1. Simulation Parameters

The simulations were performed using an editable ".fds" input file in a command prompt in windows. The simulation time of 100 seconds was written in the line &TIME T_END=100. The simulation time was chosen based on result convergence and iterations.

The water spray was inserted into the FDS with an offset of 0.30m through Lagrangian particles through the line &SPEC ID = WATER VAPOUR.OFFSET. The offset distance preserved the momentum of the spray.

The water spray from the Tyco nozzle has a 110-degree full cone shape giving 55 degrees half cone angle. The half cone angle was inserted in FDS by the line SPRAY_ANGLE=0, 55.

Since this study does not consider radiation's effect, the RADIATION is set FALSE to speed up the calculation.

Since the Lagrangian method represents the spray, the number of numerical droplets has to be specified. The number for this study was 40000 by the line PARTICLES_PER_SECOND=40000. The number of droplets/particles was determined by recommendations from the user's support of FDS to get sufficient resolution of the distributions without dramatically increasing the simulation time.

To represent the spray as a monodisperse spray, an extra line where added: MONODISPERSE=.TRUE.

The polydisperse spray was represented by the default model for droplet size distribution, explained later in this work. The model input is an average droplet diameter and the shape parameters of the distribution (γ and σ). The average diameter was set to the Sauter mean diameter from Lundberg (2021), and the shape parameter was set to default ($\text{GAMMA}_D=2.4$). σ is just a function of γ ($\sigma = 1.15/\gamma$).

The spray was considered uniformly distributed at the offset for all the simulations (SPRAY_PATTERN_SHAPE=UNIFORM).

The primary main input data used in the simulation code is mentioned in Table 1.

Table 1: Primary input data used in the simulation code

OFFSET	0.30
K_FACTOR	58.8
OPERATING_PRESSURE	2.0 bar
SPRAY_ANGLE	0,55
PARTICLES_PER_SECOND	40000
SPRAY_PATTERN_SHAPE	UNIFORM
GAMMA_D	2.4

2.2. Numerical Model

The FDS uses governing equation upon mass, species, and energy transport. The domain was created computationally and discretized into multiple cells or control volume with a general variable ϕ . The airflow, including the thermal distribution, is simulated by solving one set of the coupled state conservation of mass, momentum, and energy (McGrattan et al., 2021.).

2.2.1. Conservation of Mass

The mass transport equation is solved using the basic predictor-corrector scheme. Conservation of mass states the rate of mass storage due to a change in density in the control volume, balanced by the net rate of mass inflow by convection (Bittern, 2004). Equation 1 is the mass conservation equation for lumped species (air, fuel, and products).

$$\frac{\partial \rho}{\partial t} (\rho Z_\alpha) + \nabla \cdot (\rho Z_\alpha \mathbf{u}) = \nabla \cdot (\rho D_\alpha \nabla Z_\alpha) + \dot{m}_\alpha''' + \dot{m}_{b,\alpha}''' \quad (1)$$

On the right-hand side is the addition of mass from evaporating droplets or other sub-grid scale particles. The bulk density can be found by summing all the species $\rho = \sum (\rho Z_\alpha)_\alpha$. Obtaining the summation equation.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = \dot{m}_b''' \quad (2)$$

$\frac{\partial \rho}{\partial t}$	$\nabla \cdot (\rho \mathbf{u})$	\mathbf{u}
Change in density for time	Mass convection	Vector describing velocity in u, v & w directions

2.2.2. Momentum Transport Equation

The momentum equation is derived from Newton's second law summing up all the forces acting on the control volume. The turbulence is modeled by LES, where Smagorinsky model the sub-grid scales.

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla \rho = \mathbf{f} + \rho \mathbf{g} + \nabla \cdot \boldsymbol{\tau}_{\text{turb}} \quad (3)$$

$\rho \frac{\partial \mathbf{u}}{\partial t}$	$\rho (\mathbf{u} \cdot \nabla)$	$\nabla \rho$	\mathbf{f}	$\rho \mathbf{g}$	$\nabla \cdot \boldsymbol{\tau}_{\text{turb}}$
Momentum forces	Inertia forces	Change in pressure	External force vector $\tau_{ij} = \mu(2S_{ij} - \frac{2}{3}(\nabla \cdot \bar{\mathbf{u}}))$	Pressure & Gravity	Filtered turbulence, sub-grid scale Reynolds stress

2.2.3. Conservation of Energy

The energy equation evaluates the energy accumulation due to internal heat and kinetic energy and energy fluxes associated with convection, conduction, radiation, the interdiffusion of species, and the work done on the gases by viscous stresses and body forces (Bittern, 2004).

$$\frac{\partial}{\partial t} (\rho h) + \nabla \cdot \rho h \mathbf{u} - \frac{\partial \rho}{\partial t} + \mathbf{u} \cdot \nabla \rho = q''' - \nabla \cdot \mathbf{q}_r + \nabla \cdot \mathbf{k} \nabla T + \nabla \cdot \sum_1 h_1 (\rho D)_1 \nabla Y_1 \quad (4)$$

Left Hand Side (L.H.S)	Right Hand Side (R.H.S)
$\frac{\partial}{\partial t} (\rho h) + \nabla \cdot \rho h \mathbf{u} - \frac{\partial \rho}{\partial t} + \mathbf{u} \cdot \nabla \rho$	$q''' - \nabla \cdot \mathbf{q}_r + \nabla \cdot \mathbf{k} \nabla T + \nabla \cdot \sum_1 h_1 (\rho D)_1 \nabla Y_1$
The net rate of accumulation	Energy gain or loss term to this accumulation on the left-hand side

2.2.4. Heat Release Rate

The heat release per unit volume is defined as summing the species' mass production rate times the respective heat of formations:

$$\dot{q}'''_F \equiv \sum_\alpha \dot{m}_\alpha''' \Delta h_{f,\alpha}^0 \quad (5)$$

2.2.5. Droplet Transport Model

FDS uses the Lagrangian approach for the droplet transport model. The velocity and position of a droplet are obtained from the conservation of momentum. The trajectory and function of each droplet satisfy the following equation:

$$\frac{d}{dt} (m \mathbf{U}) = m \bar{\mathbf{g}} - \frac{1}{2} \rho C_d \pi r^2 \mathbf{U}^2 \quad (6)$$

where \mathbf{U} is the relative motion of the droplet to the ambient gas. The drag coefficient C_d , is the function of the Reynolds number based on the droplet terminal velocity, which is represented by

$$C_d = \begin{cases} 24/Re & Re < 1 \\ 24(0.85 + 0.15Re^{0.687})/Re & 1 < Re < 1000 \\ 0.44 & Re > 1000 \end{cases} \quad (7)$$

Reynolds number of a droplet is represented by

$$Re = \frac{\rho U D}{\mu} \quad (8)$$

2.2.6. Droplet Size Distribution Model

FDS uses numerical droplets representing a collection of droplets to calculate the distribution pattern. The droplet size distribution is expressed as cumulative volume fraction (CVF) and characterized by a combination of log-normal and Rosin-Rammler distribution.

$$F(d) = \frac{1}{\sqrt{2\pi}} \int_0^d \frac{1}{\sigma \cdot d} \cdot e^{-\frac{[\ln(\frac{d}{d_m})]^2}{2\sigma^2}} dD \quad (d \leq \text{VMD}) \quad (9)$$

$$1 - e^{-0.693 \left(\frac{d}{d_m}\right)^\lambda} \quad (d > \text{VMD})$$

where d is the generic droplet diameter and d_m is the volume median droplet diameter. γ and σ are empirical constants for curve fitting of distribution patterns.

The median droplet diameter is experimentally determined by sprinkler and nozzle orifice diameter, operating pressure, and geometry.

2.2.7. PDPA Model in FDS

Detailed suppression parameters are taken from spray parameters using digital Phase Doppler Particle Analysis to provide droplet size distribution, number concentration, velocity, and water flux distribution. FDS provides the output quantity that is available for PDPA. PDPA device output at time t is computed as a time integral (McGrattan et al., 2021). Figure 2 describes the output quantities of PDPA, which have been highlighted.

QUANTITY	ϕ	f	Unit
'DIAMETER' (default)	1	f_1	μm
'ENTHALPY'	$(4/3)\rho_l r_i^3 (c_{p,i}(T_i) - c_{p,i}(T_m)T_m)$	f_2	kJ/m^3
'PARTICLE FLUX X'	$(4/3)\rho_l r_i^3 u_i$	f_2	$\text{kg/m}^2\text{s}$
'PARTICLE FLUX Y'	$(4/3)\rho_l r_i^3 v_i$	f_2	$\text{kg}/(\text{m}^2 \cdot \text{s})$
'PARTICLE FLUX Z'	$(4/3)\rho_l r_i^3 w_i$	f_2	$\text{kg}/(\text{m}^2 \cdot \text{s})$
'U-VELOCITY'	u_i	f_1	m/s
'V-VELOCITY'	v_i	f_1	m/s
'W-VELOCITY'	w_i	f_1	m/s
'VELOCITY'	$(u_i^2 + v_i^2 + w_i^2)^{1/2}$	f_1	m/s
'TEMPERATURE'	T_i	f_1	$^\circ\text{C}$
'MASS CONCENTRATION'	$(4/3)\rho_l r_i^3$	f_2	kg/m^3
'NUMBER CONCENTRATION'	1	f_2	

Figure 2: Output quantities for PDPA (McGrattan et al., 2021).

2.3. Geometry

Figure 3 and Figure 4 show the SmokeView image of the simulations with a monodisperse and polydisperse representation of the droplet size distribution, with and without fire. The geometry was 6 m in length, 6 m in width, and 4 m in room

height, and all the boundary was set to open boundaries. The propane fire source for the fire simulations was appointed to a 0.8 m \times 0.8 m area in the center of the lower side of the simulation domain. The cell size was kept as 10 cm cubes and uniform in all the dimensions, as shown in Figure 5. The PDPA was aligned at 1.5 m down from the sprinkler.

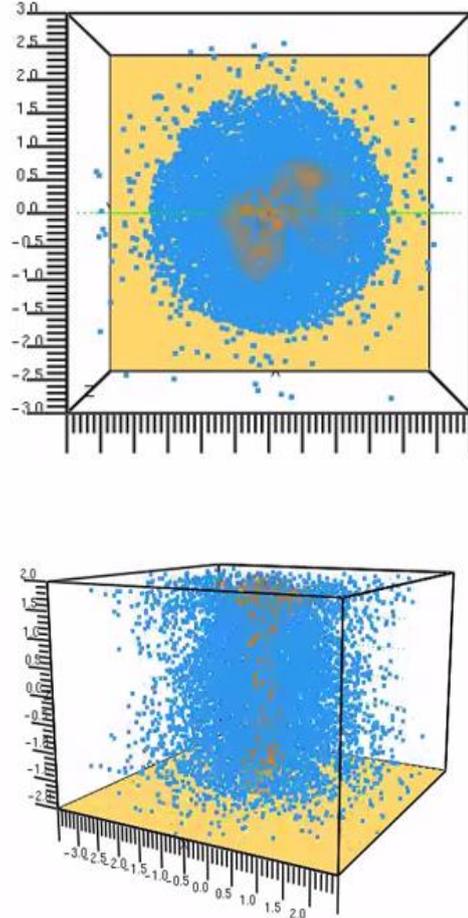


Figure 3: SmokeView monodisperse (top) v polydisperse (bottom) with fire.

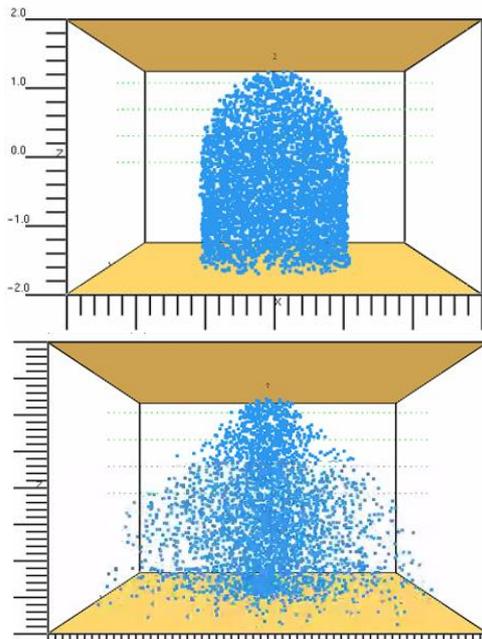


Figure 4: SmokeView monodisperse (top) v polydisperse (bottom) without fire.

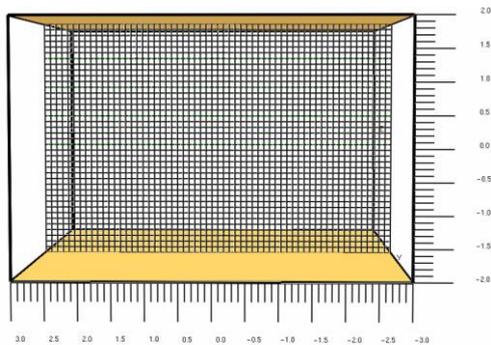


Figure 5: SmokeView of geometry with 10 cm mesh size.

3. Results and Discussion

3.2. Droplet Velocity

The droplet velocity is an important characteristic used to determine the suppression parameter. Fire plumes have a positive upward velocity where the droplets can follow the fire gasses given insufficient downwards momentum. This may result in the spray unable to penetrate the fire plume and reach the base fire. In this present study, there is a comparison study made between monodisperse and polydisperse with and without fire scenarios. The velocity distribution is measured radially using PDPA, wherein Figure 6 seems there is no fire, and the droplets have a velocity of a maximum of 9.0 m/s for monodisperse and 5.0 m/s for polydisperse spray. Since droplets may be collapsing in the gas phase, the droplet diameter tends to break and loses its velocity when reaching 1.5 m down the sprinkler nozzle. For the polydisperse spray scenario, the velocity peak at the center and both ends at

approximately 5.0 m/s. Compared to experimental results at 1 bar and 500 microns (Bourque and Svirsky, 2013), the 400 microns with 2.0 bar seems reasonable to the results obtained from the FDS.

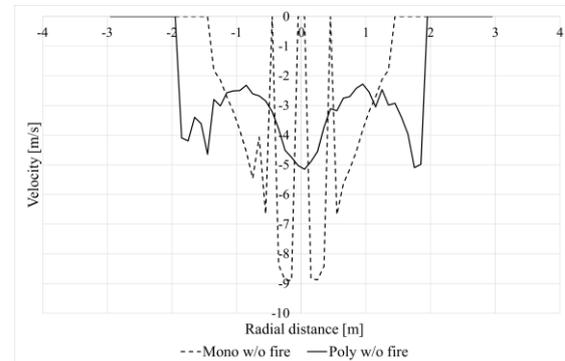


Figure 6: Velocity distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler, without fire.

In a fire scenario, the droplets tend to move upwards since the fire plume has an upward velocity as per Figure 7, measured at 1.5 m down the sprinkler nozzle. The monodisperse and polydisperse spray behave similarly since the spray can not oppose the velocity of the fire downwardly. The polydisperse spray penetrates further in the fire plume but is slower in the outer regions of the spray.

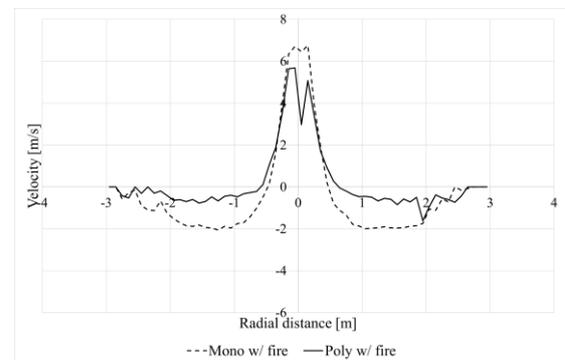


Figure 7: Velocity distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler, with fire.

3.2. Droplet Size Distribution

By default, the FDS uses Rosin-Rammler and log-normal distribution. The PDPA considers the measurement of diameter as SMD. While there is no fire, Figure 8 gives the same output as 419 microns radially -1.5 m to +1.5 m, while there is a varied droplet size for the polydisperse spray. While observing the fire scenario, as per Figure 9, the center, about 400 microns for Sauter mean diameter, is recorded for polydisperse spray. However, the distribution gets minimal at distances further from

the center and peaks at the ends. The effect is because droplets evaporate due to the fire plume. On the other hand, the monodisperse spray records reduced droplet size to 215 microns at the center due to the fire plume. But it has approximately 400 microns from -1 m to +1 m and spreads out radially from -2 m to +2 m. Thus, the bigger the droplet size, the more the probability of fire suppression as it can penetrate the fire plume.

3.3. Number of Droplets

Sprinkler sprays are composed of a large number of droplets. The number of drops without fire in Figure 10 shows 2.5E6 for the monodisperse spray, which radially spread from -1.5 m to +1.5 m. While comparing, the number of drops for the polydisperse spray is measured as 15E6 at the center. Figure 11 shows the impact of fire where the polydisperse spray reduces at the center, opposite the fire plume. Henceforth, the number of droplets on the center is close to 500 for monodisperse and 2E6 for polydisperse spray.

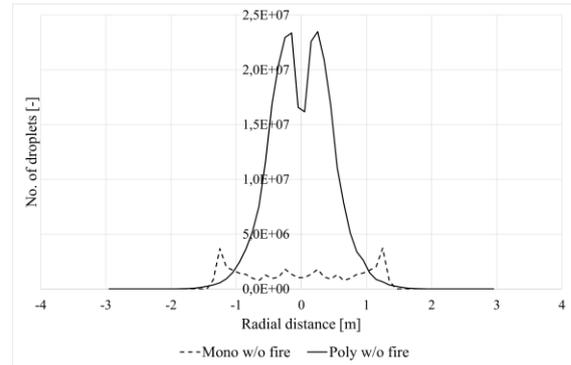


Figure 10: Droplet size distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler without fire.

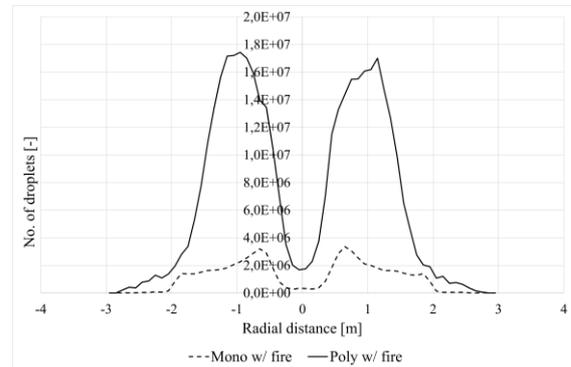


Figure 11: Droplet size distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler with fire.

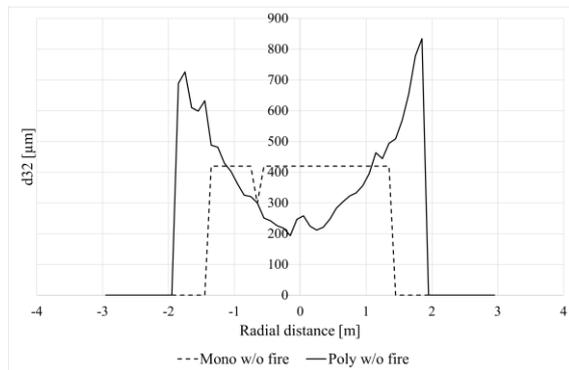


Figure 8: Droplet Size Distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler without fire.

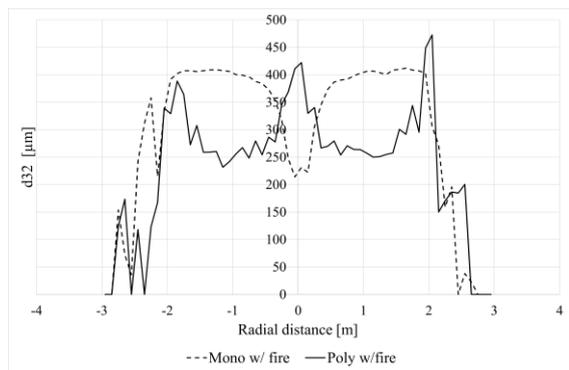


Figure 9: Droplet Size Distribution, monodisperse and polydisperse at 2.0 bar 1.5 m down the sprinkler with fire.

3.4. Cumulative Distribution Function

Figure 12 is data from the output of FDS. As expected, the CVF (cumulative volume fraction) for 0.5 reads 419 µm since the FDS diameter input is for the median volume fraction. However, the CNF (cumulative number fraction) at 419 microns is 0.92.

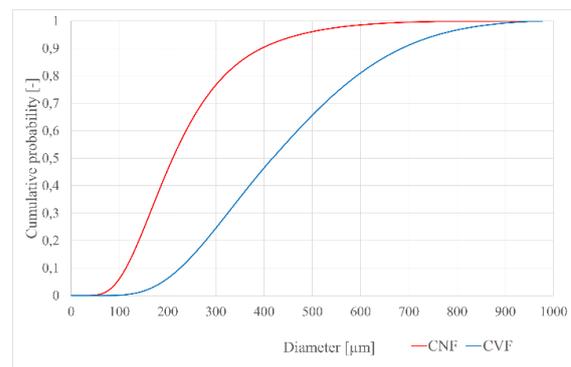


Figure 12: Cumulative distribution function.

3.5. Accumulated Mass Per Unit Area (AMPUA)

Figure 13 depicts the AMPUA by time for a monodisperse and polydisperse spray with and without fire. While observing Figure 13 at 60 seconds of simulation (one minute), the total mass accumulated with fire and without fire for

monodisperse and polydisperse is shown in Table 1, indicating that evaporation has taken place, and the polydisperse distribution has a higher evaporation rate.

Table 2: Accumulated mass flow for one minute with mono- and polydisperse distribution with and without fire.

Monodisperse spray without fire	83.1 kg/min
Polydisperse spray without fire	83.3 kg/min
Monodisperse spray with fire	69.3 kg/min
Polydisperse spray with fire	61.7 kg/min

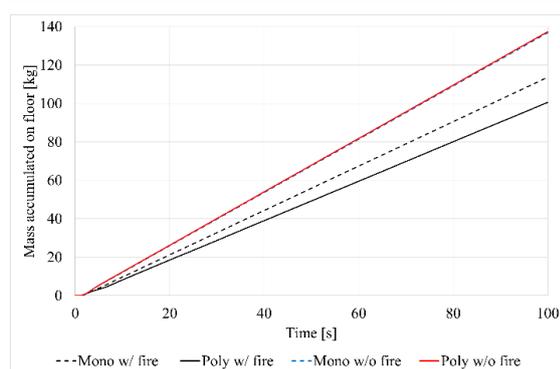


Figure 13: Accumulated mass for mono- and polydisperse distributions with and without fire.

4. Conclusion

A comparison study is performed using the Fire Dynamic Simulator, FDS by NIST. A simple rectangular mesh domain is used to simulate the behavior of a sprinkler spray with and without a propane fire. The purpose of the work is to compare the behavior of monodisperse and polydisperse sprinkler spray. The investigated suppression parameters are droplet velocity, droplet size distribution, and the number of droplets. 2.0 bar is used to supply water pressure and 40,000 Lagrangian droplets per second. The velocity model predicts that polydisperse spray differs from monodisperse spray in a fire. For droplet size distribution, monodisperse spray and polydisperse spray gives a similar output without fire and with fire, a various droplet distribution due to the evaporation and breakage of droplets. Finally, considering the number of droplets, a distribution exists in the measurements without fire, but with fire, it has a massive disturbance at the center, which proves the impact of fire on water droplets. Thus, this study can give the end user an idea for predicting fire suppression.

References

Bittem, A. (2004). 'Analysis of FDS Predicted Sprinkler Activation Times with Experiments (No. Fire Engineering Research Report 04/8)'. Department of Civil Engineering University of Canterbury, Christchurch, New Zealand.

Heskestad G. (2002). 'Scaling the interaction of water sprays and flames'. *Fire Safety Journal* 37:535–548. doi:10.1016/S0379-7112(02)00012-7.

Liu Y., Beji T., Thielens M., Tang Z., Fang Z. and Merci B. (2022). 'Numerical analysis of a water mist spray: The importance of various numerical and physical parameters, including the drag force'. *Fire Safety Journal* 127, 103515. doi:10.1016/j.firesaf.2021.103515.

Lundberg J.(2021). 'Characterization of a medium velocity deluge nozzle for offshore installations'. *Journal of Loss Prevention in the Process Industries*, 71, 104510. doi:10.1016/j.jlp.2021.104510.

Lundberg J. (2015). 'Image-based sizing techniques for fire water droplets'. PhD thesis, Telemark University College, Porsgrunn, Norway.

Bourque M. J. and Svirsky T. A.(2013). 'Computational modeling of fire sprinkler spray characteristics using the fire dynamics simulator'. Bachelor Thesis, Worcester Polytechnic Institute, MA, USA.

McGrattan K., Hostikka S., McDermott R., Floyd J., Weinschenk C. and Overholt K.(2010). 'Fire Dynamics Simulator Technical Reference Guide Volume 3: Validation'. NIST Special Publication 1018-5, Gaithersburg, MD, USA.

McGrattan K., Hostikka S., McDermott R., Floyd J., Weinschenk C. and Overholt K.(2010). 'Fire dynamics simulator (version 5). User's guide'. NIST Special Publication 1019-5, Gaithersburg, MD, USA.

Myers T. M. and Marshall A. W.(2016). 'A description of the initial fire sprinkler spray'. *Fire Safety Journal* 84(1–7). doi:10.1016/j.firesaf.2016.05.004.

Sæbø A. O. and Wighus R.(2015). 'Droplet sizes from deluge nozzles'. Report no A15 107453:1 (replaces restricted report NBL F09117, 2009), Trondheim, Norway.

DesJardin P. E., Gritzo L. A. and Tieszen S. R.(2000). Modeling the effect of water spray suppression on large-scale pool fires. Proceedings of Halon Options Technical Working Conference May 2-4, pages 262-250, University of New Mexico, Albuquerque, NM, USA.

Ren N., Baum H. R. and Marshall A. W.(2011). 'A comprehensive methodology for characterizing sprinkler sprays'. Proceedings of the Combustion Institute 33, 2547–2554. doi:10.1016/j.proci.2010.06.107.

Sheppard D. T.(2002). 'Spray Characteristics of Fire Sprinklers'. National Institute of Standards and Reporting NIST GCR 02-838, Gaithersburg, MD, USA.

Zhou X., D'Aniello S. P., and Yu H.-Z.(2012). 'Spray characterization measurements of a pendent fire sprinkler'. *Fire Safety Journal, Large Outdoor Fires* 54, 36–48. doi:10.1016/j.firesaf.2012.07.007.

Data-driven approaches for modelling of sub-critical coal-fired boiler

Valentin Formont ^{a,*}, Vidar T. Skjervold ^a, Lars O. Nord ^a

^a NTNU - Norwegian University of Science and Technology, Department of Energy and Process Engineering, Trondheim, Norway

*valentin.formont@ntnu.no

Abstract

Due to increasing shares of renewable electricity sources in the grid, thermal power plants need to operate in a more flexible manner in the future. This will involve more frequent startups, shutdowns, and load changes. A central part of a thermal power plant analysed in this study is the coal-fired boiler. In a previous study, a first-principle model of a sub-critical coal-fired boiler has been developed and validated with operational data from a Polish power plant. Based on this model, this work aims to develop a computationally efficient and sufficiently accurate data-driven model that is easy to implement in new software. A selection of multi-output algorithms was first compared using nonoptimised parameters, with very few adaptations to the data set. Then, each algorithm had undergone three different optimisation routines to tune the hyper-parameters. The results of the nonoptimised models were compared with the optimised ones, and then compared to the reference first-principle model using the average Mean Absolute Percentage Error as a score. The methods used comprise six base learners and three algorithms using ensemble methods. The optimisation routines were based on the Powell conjugate direction method, Bayesian optimisation and evolutionary algorithm. All the data-driven models had shown a lower percentage error than the first principle model, and optimisation had resulted in improved prediction capacity for every base learner, but not for ensemble method-based algorithms.

1. Introduction

It is expected that thermal power plants will need to operate more flexibly in the future due to the increased share of renewable energy sources in the grid. This will lead to more frequent startups, shutdowns and load changes. In order to aid this transition, there is a need for models that can describe the transient behavior of such power plants. As a result of the thermal inertia of e.g. boiler walls and heat exchanger surfaces, the boiler has a large influence on the dynamic behavior of a power plant. This unit is therefore an important part of a power plant model, and the aim of this work is to develop a computationally efficient and sufficiently accurate boiler model that is easily implemented in new software.

In the literature, several examples of data-driven modelling approaches applied to coal-fired power plants are available. Based on operational data of around 60 variables in a brown coal-fired power plant, Smrekar et al. developed two artificial neural network (ANN) models [1]. The aim of the work was to examine the feasibility of such models for coal-fired power plants. In Chandrasekharan et al. [2], separate models for the economizer, drum and superheater were developed by using a statistical approach based on the response surface methodology and design of experiments. The data used for model development was collected from a 210 MW power plant over a 2-3 hour period. In Zhu et al. [3], a local model network (LMN) was used to model the boiler-turbine unit of a sub-critical coal-fired power plant and applied as the prediction model in a non-linear model predictive controller. The model was validated with data from a 500 MW unit in China. With

emphasis on cyclic operation, Navarkar et al. [4] built an ANN model of a coal-fired steam generator based on 10 years of operational data from a power plant in the USA. Manaf & Abbas [5] and Oko et al. [6] both applied non-linear autoregressive with exogenous input (NLARX) models to coal-fired power plants. In the former, a complete power plant divided into six sub-systems was modelled by using data from a 660 MW power plant. In the latter, the NLARX model was used to predict drum pressure and level in a sub-critical unit. The data used for model development were generated by a validated first-principle model. As indicated by the literature review, a comparison of several different static data-driven modelling approaches for a single coal-fired boiler has not yet been presented. This knowledge gap will be addressed by this work.

Using data-driven approaches has several advantages:

- It uses real data to make predictions. It makes a change from the first-principle models paradigm that simplifications must be made in order to be able to perform computations;
- It is a very active field of research and there is a big community of researchers developing and enhancing the different methods;
- Once a model is trained, it is very fast, within seconds, to get a prediction;
- The model can be updated by getting more data after it is trained, so it can have a better predictive ability.

2. Methodology

Measurements at different stages of a sub-critical coal-fired boiler at a Polish power plant have been performed, constituting the data set for this paper. The same observations are computed using a first principle model, which constitutes the reference data set. Different regressors have been chosen to compare to the reference model : Ordinary Least Square, Elastic Net, Support Vector Machine, Stochastic Gradient Descent, Nearest Neighbors, Decision Trees, Random Forest, AdaBoost and XGBoost. The first six model structures are called 'base learners' and the other are regressors based on ensemble methods. The ensemble methods use several weak learners to improve prediction capacity. These regressors allow the use of hyper-parameters, i.e. parameters independent from the observations, and the value of these are usually to be determined. First, the regressors are fit to the observations without prior hyper-parameter optimisation. The Mean Absolute Percentage Error (MAPE) of each model is kept for comparison. Then, using different optimisation routines the same methods are used combined with hyper-parameter optimisation. As a result, for each algorithm, one nonoptimised and three optimised models are created.

2.1. Data from industrial boiler

The data from the industrial boiler comprise of 480 observations including 5 inputs and 7 outputs each taken at a 1-minute interval. A sketch of the coal-fired boiler considered in this work is shown in Figure 1.

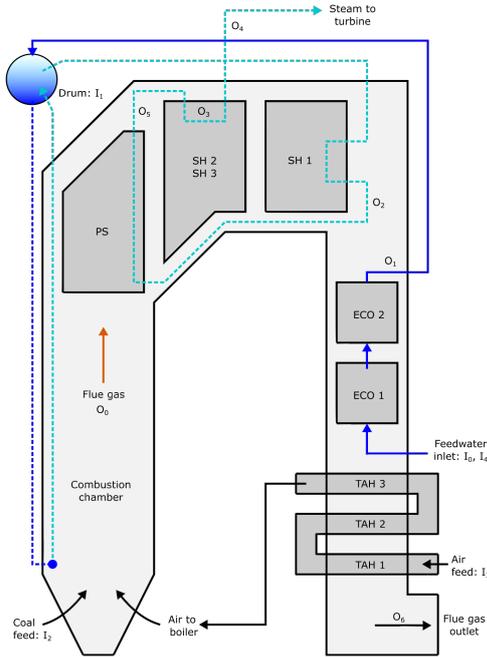


Figure 1: Sub-critical coal-fired boiler considered in this work. TAH stands for tubed air heater, ECO stands for economizer, SH stands for superheater and PS stands for platen superheater. Inputs and outputs, denoted by I and O, are described in Table 1.

The different variables are shown in Table 1. It is important to note that, for this paper, the time is not included in the set of inputs. The reason is that including time in the inputs shifts the analysis into a time-series analysis, and the target of this paper is to develop a multi-output model solely dependent on the state of the

boiler. The time-series will be considered for future work.

Table 1: Monitored variables

Inputs	Code	Unit	Bounds
Time	None	min	[0, 480]
feedwater T	I_0	°C	[225, 236]
Pressure in the steam drum	I_1	MPa	[10, 11]
Fuel mass flow rate	I_2	kg/s	[6, 8]
Air flow rate	I_3	kg/s	[54, 65]
feedwater mass flow rate	I_4	t/s	[0, 51]
Outputs			
Flue gas T combustion chamber	O_0	°C	[977, 1033]
Water T from ECO2	O_1	°C	[307, 319]
Steam T I-st stage SH outlet	O_2	°C	[366, 387]
Steam T II-nd stage SH outlet	O_3	°C	[479, 497]
Steam T III-rd stage SH outlet	O_4	°C	[530, 544]
Steam T PS outlet	O_5	°C	[448, 475]
Flue gas T after TAH1	O_6	°C	[154, 173]

First, the Pearson correlation is calculated for each input variable. Given x_1 and x_2 two different variables in a data set comprising of n observations, it is possible to calculate the strength of association between these two variables by computing the Pearson's coefficient (Equation 1).

$$r = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{\sqrt{\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2 \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2}} \quad (1)$$

For readability, the results of the correlation analysis are displayed in a heatmap in Figure 2. The input names are coded for visualisation purposes, and their code can be retrieved from Table 1. The data-driven models usually perform poorly when the data are not on the same scale, thus it is needed to have a step of feature scaling, Géron [7]. Standardisation offers a robust way of scaling the data by transforming the entire data properties so they follow an unknown distribution with $X \sim \mathcal{D}(0, 1)$. The formula used for standard scaling is given in Equation 2.

$$x_{i, \text{scaled}} = \frac{x_i - \mu}{\sigma} \quad (2)$$

with μ the mean of the observations and σ the associated standard deviation.

To assess the performance of the models, it is common to split the data set into a training set and a test set, Tan et al. [8]. The reason behind this first split is to keep an untouched data set. The test set is not used to train the model, nor to perform hyper-parameter tuning. After the model has been fit to the training set, it is rated using a score function calculated on the test set.

2.2. Data-driven modelling techniques

The regressors used to get predictive models are categorized into two groups: base learners and ensemble methods. The main difference between a base learner and one from ensemble method is that the ensemble method uses several base learners to make predictions, whereas a base learner is a single model. To assess the performance of each model the Mean Absolute Percentage Error (Equation 3) is computed. This choice of score function is justified by the noisy behaviour of the observations. The squared error metrics such as Mean Squared Error or Root Mean Squared Error penalize outliers by squaring the error, making them very sensitive to noise.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (3)$$

with y_i the real value and \hat{y} the predicted value of the i_{th} observation.

As the model should be a multi-output regressor, there are actually 7 different MAPE per model. The error is computed for the 7 outputs, and then averaged to produce an average MAPE. Calculating the MAPE on the entire training set is not robust enough, because the model only sees one set of data, and there is no possibility to know how the model would behave with other samples from the same experiment. This is why the performance on the training set is assessed by implementing cross-validation. The cross-validation is performed along with K-fold sampling. Given a data set with n observations, the K-fold sampling method splits the data set randomly into k subsets. $k - 1$ subsets are used for training the model and one subset acts as a test set. This is repeated until all the subsets acted once as a test set. Then, the outcome of the cross-validation is the mean of the k model scores. Performing cross-validation allows getting a robust estimation of both the bias and the variance of a model. The models will be fit to measurements of a real power plant, and the reference to outperform is a data set coming from the first principle model. The MAPE of the reference model compared to the real data model is 0.716%. All the hyper-parameter initial values are displayed in Table 2. Except where specified, these values are the values recommended by default by the different libraries used. Sci-kit learn is used for every model except XGBoost, which comes from the homonym library [9].

Table 2: Regressors' Hyper-parameters initial values

Algorithm	Hyper-parameter	Value
Elastic Net	λ	1
	α	0.5
SVR	Kernel	RBF
	σ	$\frac{1}{n_{\text{predictors}} \cdot \text{var}(X)}$
	L_2 penalty	1
	ϵ	0.1
SGD	Loss	huber
	iterations	10000
	ϵ	0.1
	η	adaptive
KNN	k	3
	weights	distance
	data structure	ball tree
	s_{leaf}	30
Decision Tree	p	2
	criterion	absolute error
	d_{max}	\emptyset
	ss_{min}	2
	sl_{min}	1
Random Forest	N	$n_{\text{predictors}}$
	criterion	absolute error
AdaBoost	N	100
	η	50
XGBoost	η	1
	N	100
	η	0.3
	d_{max}	6
	sb_s	1
	cols	1
	γ	0
	cw_{min}	1
	α	0
	λ	1

2.2.1. Ordinary Least Square Regression

Given a set of n observations $\{x_i, y_i\}_{i=1}^n$, each observation i comprises a column vector x_i of p predictors such as $x_i = [x_{0i}, x_{1i}, x_{2i}, \dots, x_{pi}]^T$ and a response y . The simplest form of linear regression assumes that y is a function of x_i such as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \quad (4)$$

with ε_i accounting for the variations not explained by the regression model (bias) and β a column vector of unknown coefficients with the shape $p \times 1$. Finding these coefficients consists of optimizing a score P . Using the Mean Squared Error (MSE) as a score function, it is possible to estimate the unknown coefficients $\hat{\beta}$ using the matrix form of the *normal equation* displayed in Equation 5, Goodfellow et al. [10]

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (5)$$

$$\text{with } X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p,1} & x_{p,2} & \dots & x_{p,n} \end{bmatrix}$$

The *normal equation* is obtained by minimizing the MSE in Equation 6

$$\hat{\beta} = \arg \min_{\beta} \frac{1}{2n} \|Y - X\beta\|^2 \quad (6)$$

It is one of the simplest regression techniques, but falls short when the underlying data structure is complex, Zou and Hastie [11].

2.2.2. Elastic Net Regularisation

The Elastic Net Regularisation is a combination of two different regularisation techniques, the Least Absolute Shrinkage and Selection Operator (LASSO) known as L_1 regularisation and the Tikhonov or L_2 regularisation. These regularisation methods add a penalty function to the loss function in Equation 6. The penalty is based on the L_1 and L_2 norms. In Equation 7, it is possible to identify on the right-hand side the first term as the cost function of the Ordinary Least Square (OLS). The second term and the third term represent respectively the L_1 and L_2 penalties.

$$\hat{\beta}_{\text{net}} = \arg \min_{\beta} \frac{1}{2n} \|Y - X\beta\|^2 + \lambda \alpha \|\beta\|_1 + \frac{1}{2} \lambda (1 - \alpha) \|\beta\|^2 \quad (7)$$

where λ is a parameter controlling the strength of regularisation, and α is strictly between 0 and 1. If $\alpha = 1$, the elastic net is the same as LASSO and as α tends toward 0, it gets closer to ridge regression. The Elastic Net is the same as an Ordinary Least Squares when $\lambda = 0$.

2.2.3. Support Vector Regression

Support Vector Machine (SVM) was initially used for classification, transforming a binary classification problem into a convex optimisation. The main principle behind SVM is to find a function $f(x)$ within a maximum deviation ε compared to the real response y . The function should be as flat as possible. To illustrate this method, $f(x)$ is given linear, and can be written as in Equation 8, Vapnik [12].

$$f(x) = \langle w, x \rangle + b \quad (8)$$

with w a normal vector defining the hyperplane $\langle w, x \rangle = b$

For this case, the problem lies in minimizing the loss function represented in Equation 9.

$$\arg \min_w \frac{1}{2} \|w\|^2 \quad (9)$$

with the following constraints to respect

$$C = \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon \\ \langle w, x_i \rangle + b - y_i \leq \epsilon \end{cases} \quad (10)$$

SVM for regression is shortened to SVR in this study. SVR combined with feature transformation allows modelling of complex data distribution. The underlying idea is that if a data structure is not linear in its original space, it might be linear in another space. The function used to change the space is a kernel function. Its general form is written

$$K(x, u) = \phi(x) \cdot \phi(u) \quad (11)$$

with x and u two independent vectors belonging to the same space. In the result section, the Radial Basis Function (RBF) is used as a kernel

$$K_{\text{RBF}}(x, u) = e^{-\frac{\|x-u\|^2}{2\sigma^2}} \quad (12)$$

σ is a free parameter called the length scale of the kernel. Its value is

$$\sigma = \frac{1}{n_{\text{predictors}} \cdot \text{var}(X)} \quad (13)$$

In addition, it is possible to add an elastic net regularisation to the cost function. Smola and Schölkopf [13] made a tutorial about SVR, which details the full scope of this method.

2.2.4. Stochastic Gradient Descent Regression

Stochastic Gradient Descent is initially an iterative method used in optimisation to find an extremum. It inherits from Batch Gradient Descent (BGD), with the main difference that BGD computes the exact gradient of the loss function based on every single observation. On the other side, SGD uses an approximated gradient based on one point at a time. Considering an arbitrary loss function $J(\theta)$, it often takes the form of a sum which can be seen as

$$J(\theta) = \frac{1}{n} \sum_{i=1}^n f_i(\theta) \quad (14)$$

θ is obtained by iteration

$$\theta := \theta - \eta \nabla f_i(\theta) \quad (15)$$

with $f_i(\theta)$ a prediction of the i_{th} observation and η the learning rate of the gradient descent. The parameter η influences the distance between two steps, if it is too large the algorithm might never find an optimal solution. On the other hand, if the learning rate is too small, the algorithm will be very slow to converge. A common method is to use an adaptative learning rate which is set high then decreases as the optimisation converges to a minimum. SGD Regression uses this optimisation routine to fit a linear regression. For instance, if the squared error is used as a loss function, then performing a regression with SGD-based regression is equivalent to an OLS. The Huber loss (Equation 16) is a more robust loss function.

It penalizes less the outliers by being quadratic when the error is smaller than ϵ and linear otherwise, Huber [14].

$$J_h(\theta) = \begin{cases} \frac{1}{2}(y - f(\theta))^2 & \text{for } y - f(\theta) \leq \epsilon \\ \epsilon \cdot (|y - f(\theta)| - \frac{1}{2}\epsilon) & \text{for } y - f(\theta) > \epsilon \end{cases} \quad (16)$$

This method is very suitable for large-scale applications, but can be slower than BGD for convex functions with a single minimum, Bottou and Bousquet [15].

2.2.5. Nearest Neighbors Regression

Nearest Neighbors is an algorithm dealing with both classification and regression. The intuition behind this algorithm is that if a group of observations is close to each other, their responses should be close too. Given a query point q , a k nearest neighbors search consists of finding the k nearest observations to the query point. In its naivest implementation, the euclidean distance is computed using a brute force approach, meaning that the distance is calculated for each pair of points in the entire set of observations. It is then possible to generalize this process to any distance using the Minkowski distance (Equation 17)

$$D_{\text{mink}}(x_1, x_2) = \left(\sum_{i=0}^n (|x_{1i} - x_{2i}|^p) \right)^{\frac{1}{p}} \quad (17)$$

This is very efficient for a small number of observations, but can reach life span calculation time very quickly as the size of the data set gets bigger. Instead of computing the proximity of every single pair of points with a brute force approach, modifying the data structure before applying a nearest neighbors search has led to more reasonable results for higher dimensional or bigger data sets. One of the data structures speeding up the nearest neighbors search is the Kd-tree data structure. For n observations with d dimensions, the i_{th} coordinate is split using the median value as a separator. It results in two partitions of the observations, L_1 and G_1 , respectively the values lower and greater than the median. Then the process is repeated for the coordinate $i + 1$ and iterates once for each dimension. This process goes on until the leaves of the Kd-tree contain a maximum of s_{leaf} observations. An improvement of the Kd-Tree is to use round-shaped leaves instead of rectangles to create clusters. This is called a ball tree data structure. Andrew Moore [16] detailed how this data structure is created. As a result of this new data structure, the nearest neighbors search on the query point is compared only to a small subset of observations, lying within the same leaf or surrounding leaves. For regression, the label assigned to an observation is based on the weighted mean of the labels of its nearest neighbors, Pedregosa et al. [17]. The weights are inversely proportional to the distance between the query point and the neighbors. The determination of the optimal number of neighbors k is made by comparing the error of KNN models with different k values.

2.2.6. Decision Trees

A decision tree is initially a classification algorithm. It splits the observations into different leaves, just like Kd-tree. It comprises an initial node, including all the observations, internal nodes splitting the observation according to decision rules and leaf nodes, the final leaves of the tree. The decision criteria to split a node for a Decision Tree in the case of regression is usually the MSE, but will be set to MAE for this paper. The observations

are split such that the MSE within a leaf is minimized, and simple models can be fit to the different leaves. This partition continues until there is no possible point to split, or a termination criterion is reached. Common parameters to tune include d_{\max} the maximum depth of the tree, ss_{\min} the minimum number of observations that a node must contain to split, sl_{\min} the minimum number of observations that a leaf must contain to be considered as a leaf node and f_{\max} the number of predictors to consider for the best split. Note that the maximum depth of the tree can be set to none, so that the tree expands until all leaves are pure. The main advantage of this algorithm, which is also its main drawback, is its simplicity. It is very sensitive to the data structure, and adding few observations can change drastically the structure of the tree. Thus, this model is associated with high variance and does not compare well to other more stable algorithms.

2.2.7. Random Forest

In 2001, Breiman [18] formally proposed an extension of the Decision Tree algorithm. The intuition behind this improvement is to use several subsets based on the original database to create different decision trees and then average their predictions. First, the observations are bagged N times, creating N random samples with replacement and N decision trees are fit to the different random samples. Then the final prediction for an unknown observation x is the average of the predictions of all the decision trees, such as

$$R(x) = \frac{1}{N} \sum_{i=1}^N F_i(x) \quad (18)$$

where F_i is the output of the i_{th} decision tree from the Random Forest.

This method reduces readability and has a small increase in the bias, but improves generalisation. Some hyper-parameters are unique to the forest, but many are shared between the trees and the forest. For instance, the number of trees is unique to the forest, but the maximum depth or the minimum samples split are parameters for the trees inside the forest.

2.2.8. AdaBoost

Combining several weak learners into a strong learner is called boosting, and the flagship of this method is the Adaptive Boosting algorithm, AdaBoost. The AdaBoost algorithm is very close to a random forest, but instead of having a uniform average of the decision trees, AdaBoost keeps weights associated with every single weak learner. Initially, all the weights are equal, but after each iteration, the weak learners with the worst performance get an increased weight. As a result, the algorithm will focus on minimizing the error on the weak learners with high weights. AdaBoost for regression, called AdaBoost.R2 is detailed by Drucker [19].

2.2.9. XGBoost

eXtra Gradient Boosting is one of the best implementations of gradient boosted trees. It has actually dominated machine-learning competition for a long time, and is recognized as one of the most accurate modeling methods. The main difference between XGboost and the other gradient boosting techniques lies in the optimisation of the algorithm structure. For instance, it treats calculations of base learners in parallel.

Another example of improvement is to use the d_{\max} as a stopping criterion instead of a score to compute, reducing the computation time. The hyper-parameters related to XGBoost are N the number of trees; η the learning rate or shrinkage factor to apply on the weights; sb_s the proportion of observations to shuffle and use from the training set, $cols$ the fraction of predictors to use within each tree (feature reduction), γ the minimum loss reduction to split a leaf node, wc_{\min} the minimum weight instance a child must carry in order to exist, and α and λ respectively the L_1 and L_2 regularisation factors. Other enhancements have been implemented. See Chen and Guestrin [9] for further details about Extreme Gradient Boosting.

2.3. Optimisation methods

All the methods evoked in the last sections include parameters independent from the data. These parameters are called hyper-parameters, and the optimisation of these values is still an active research topic. The naive approach to optimize them is the brute force approach. Because it is not very realistic to do an exhaustive brute force search, a greedy search can be performed on manually specified values to narrow the search space. This method is called the Grid Search approach, but it requires to have an intuition about the parameters optimal values. The most promising way to find the best parameters is to use more complex optimisation methods. For this paper, the cost function is the result of the cross validation on the training set (Equation 19). It can be calculated given θ , an unknown vector of hyper-parameters, and $F(\theta)$, a model already fit to a data set.

$$J(\theta) = \frac{1}{k} \sum_{i=1}^k \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - F_i^k(\theta)}{y_i} \right| \times 100 \quad (19)$$

Three different optimisation techniques will be used and compared on different data-driven models.

2.3.1. Powell Method

The Powell Method as stated by Powell in 1964 [20] is a gradient-free optimisation method derived from the conjugate direction method. It assumes that by combining several one-dimensional optimisations, it is possible to find the optimum of complex multidimensional functions. The optimisation starts with an initial point x_0 and a set of N guess vectors either manually given or parallel to each axis. A bi-directional search is performed along each vector, and a search vector is created based on the result of the search, which gives another search point. This process goes on until a criterion is met. The implementation of the Powell Method is made with SciPy [21].

2.3.2. Bayesian optimisation

The Bayesian optimisation comes from Bayesian Statistics. In Bayesian statistics, the inference of a model is a probabilistic approach based on a set of prior/posterior beliefs about an event. Given the event of fitting a group of observations X into a specific model, the prior probability distribution $P(\theta)$ is the initial belief about the distribution of the parameters θ . In the case of linear regression, the event would be to have the right distribution of β for X . After collecting new observations, an updated probability distribution is made based on the prior and the new observations: the posterior probability distribution, $P(\theta|X)$

$$P(\theta|X) = \frac{P(X|\theta)P(\theta)}{P(X)} \quad (20)$$

with $P(X|\theta)$ the likelihood of X and $P(X)$ the marginal likelihood given by

$$P(X) = \int P(X|\theta)P(\theta)d\theta \quad (21)$$

The Bayesian optimisation method assumes that the unknown function is generated with a Gaussian Process. A Gaussian process is a stochastic process defined by its mean $\mu(x)$ and its covariance $k(x, x')$.

$$\begin{aligned} \mu(x) &= \mathbb{E}[f(x)], \\ k(x, x') &= \text{cov}(x, x') \end{aligned} \quad (22)$$

such that

$$f(x) \sim \text{GP}(\mu(x), k(x, x')) \quad (23)$$

For a detailed explanation of how Gaussian Processes work, see Rasmussen and Williams [22]. After generating a prior over a Gaussian process, the posterior, also called acquisition function, is generated. Three main acquisition functions are mainly used in Bayesian optimisation; the Probability of Improvement, the Expected improvement and the Upper/Lower Confidence Bound. For this paper, the Upper Confidence Bound (UCB) is used as the acquisition function. Once the UCB function is evaluated, the next points for optimisation are chosen where the value of UCB is the highest, ie. where the uncertainty about the function estimation is at its highest. In the result section, 100 iterations were made with 5 initial random guess points. A practical guide for Bayesian optimisation has been written by Snoek et al. [23]. The Bayesian optimisation is implemented using the bayes_opt library [24].

2.3.3. Genetic Algorithm

The Genetic algorithm is a population-based metaheuristic optimisation method mimicking biological evolution. A first population p of chromosomes is randomly generated. Each chromosome comprises n_{genes} genes. Each gene has one random value attributed. The population comprises n_{chr} chromosomes. The cost function $J(\theta)$ is calculated for each chromosome, and a fitness function guides the evolution of the next population generation. The fitness function is a positive gain function. In this paper, the fitness function is the inverse of the cost function. To generate the next population $p + 1$, n_{parents} chromosomes are coupled together, and as a result, a child chromosome is created. Stronger chromosomes, with better fitness scores, are more likely to mate. The child is made of genes from the parents. This mating process is repeated until the next population is full. However, to avoid losing the best solutions in the mating process, k_{elit} genes with the best fitness function output will be kept in the population $p + 1$. At last, random mutations can happen during the process, changing spontaneously the value of one gene. These mutations have a probability of p_{mutation} to happen for each chromosome. The implementation of the Genetic Algorithm is made using PyGAD [25] and the values of the hyper-parameters are displayed in Table 3. A detailed review of the Genetic Algorithm has been made by Katoch et al. [26].

3. Results and discussion

The first step before modeling is to perform the correlation analysis of the predictors. As shown in Figure 2, the feedwater temperature has a mild correlation coefficient with all the other inputs. These coefficients lie within an interval between [0.38,0.52]. However, the pressure in the

Table 3: Genetic Algorithm Parameters

Parameter	Value
$n_{\text{generation}}$	20
n_{parents}	4
n_{chr}	8
n_{genes}	$n_{\text{predictors}}$
k_{elit}	1
p_{mutation}	0.1

drum; the fuel mass flow rate; air flow rate and feedwater mass flow rate are all highly correlated with each other [0.94,1]. It can be a possible lead for future improvement, because dropping highly correlated predictors within a model is usually associated with a low bias drop but a faster execution.

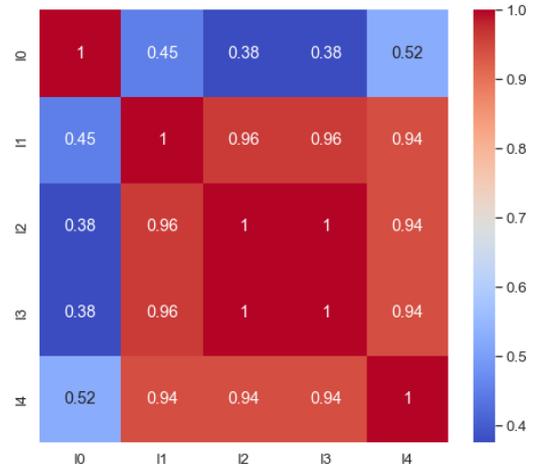


Figure 2: Pearson's correlation coefficients heatmap

For this paper, all the inputs, with the exception of time, are included in the models. Then the data set is split into a training set and a test set containing respectively 70% and 30% of the observations. The data set is randomly split into two subsets, therefore all the base learners do not result in the same outcome if the split happens in a different way. This is especially true for ensemble methods, which add even more randomness using bagging and other stochastic methods. However, cross-validation and several runs of the algorithms ensure that the results of this paper come with a minimized variance. As mentioned in the Methodology section, the reference model to outperform has a MAPE of 0.716%. Nonoptimised values are used to assess 'default' models. Some minor adaptations have been made. For instance, KNN uses the best K from the elbow method, and the SGD regressor minimizes the Huber loss. Table 4 shows the comparison of the average MAPE for the different models mentioned in the Methodology section, the hyper-parameter values associated with each model are in Table 2. Overall, every algorithm outperformed the reference model, by 15% for the worst performing model and 57% for the best performing model. The models do have a good generalisation capacity on the test set, ie. the test error is smaller than the training error. The two best algorithms are issued from ensemble methods. It follows a pattern, which is the more complex the algorithm is, the better the model prediction capacity will be, except for AdaBoost. Indeed very simple algorithms like OLS or Elastic Net are the worst-performing, whereas XGBoost and Random

Forest have the best prediction scores.

Table 4: Nonoptimised prediction models

Model	Training MAPE	Test MAPE
OLS	.472	.429
Elastic Net	.634	.609
SVR	.391	.361
SGD	.473	.437
K-nearest neighbors	.340	.324
Decision Tree	.393	.370
Random Forest	.328	.310
AdaBoost	.397	.390
XGBoost	.349	.317

Then every model, with the exception of OLS, goes through 3 different optimisation methods. The result of the optimisation is shown in Table 5 and the hyper-parameters values for the optimisation are shown in Table 6. In general, there is a net improvement after optimisation of the hyper-parameters. The best improvement was obtained on the worst-performing model. After optimisation, the worst-performing algorithm outperformed the reference by 40% and the best by 58%. It is important to notice that the strength of regularisation λ is set to 0 after optimisation, so regularisation is not necessary on the observations. Fine-tuning the parameters led to a net improvement of the prediction capacity. However, the Random Forest and the XGBoost did not benefit from the optimisation. These regressors are generally optimised by default to general cases and it is possible that the bounds of hyper-parameters need to be shrunk, especially for XGBoost regularisation parameters. Out of eight regressors, only two did not benefit from optimisation. The Powell methods did not find the best solution in any cases; Bayesian optimisation 4 solutions; and the Genetic Algorithm 2. The results do not mean that Bayesian optimisation is a better optimizer, only that for this specific 'random state', it found better solutions. Another interesting finding is that the genetic algorithm seems to prioritize solutions close to the given bounds of the hyper-parameter values.

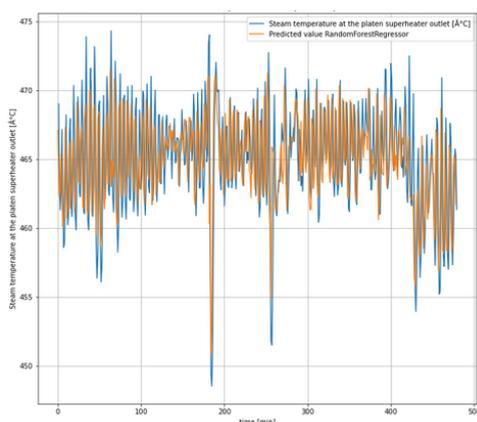


Figure 3: Univariate Time Series representation of a Random Forest prediction of the Steam Temperature at the platen superheater outlet

These results can be criticized in several ways. Data-driven approaches come with their unique drawbacks. Compared to first-principle models, these algorithms are usually considered as black-box functions. They differ from the original 'laws' and must be represented with

Table 5: Nonoptimised and optimised predictions' MAPE on test set. NO stands for Not optimised, PWL for Powell method, BO for Bayesian Optimization and GA for Genetic Algorithm

Model	NO	PWL	BO	GA	Maximum gain
Elastic Net	.634	.428	.428	.428	29.77
SVR	.361	.301	.300	.299	17.04
SGD	.437	.905	.396	.401	9.38
K-nearest neighbors	.324	.307	.319	.304	6.14
Decision Tree	.370	.614	.327	.346	11.22
Random Forest	.310	.322	.315	.314	-1.51
AdaBoost	.390	.374	.353	.361	9.41
XGBoost	.317	.991	.565	.389	-22.75

Table 6: Regressors' Hyper-parameters optimised values. NO stands for Not optimised, PWL for Powell method, BO for Bayesian Optimization and GA for Genetic Algorithm

Model	Param.	Bounds	PWL	BO	GA
Elastic Net	λ	[0;1]	.01	.00	.00
	α	[0;1]	.94	.00	.00
SVR	σ	[0;1]	.74	.98	1.00
	λ	[0;10]	10.00	9.52	10.00
	ϵ	[0;5]	$1e^{-2}$	$1e^{-3}$	$1e^{-3}$
SGD	λ	[0;10]	6.18	.00	.00
	η_{ini}	$[1e^{-7};2]$	1.24	2	2
	α	[0;1]	.69	1	.00
	ϵ	[0;1]	.76	.36	1
KNN	k	[2;50]	4	5	2
	s_{leaf}	[1;100]	99	26	1
	p	[1;5]	4	4	5
Decision Tree	d_{max}	[2;50]	38	6	50
	ss_{min}	[0;1]	.11	.00	.00
	sl_{min}	[0;1]	.17	.00	.00
	f_{max}	[0;1]	.71	1.00	1.00
Random Forest	N	[10;500]	197	434	500
	d_{max}	[2;50]	31	18	50
	ss_{min}	[0;1]	$1e^{-3}$.00	.00
AdaBoost	N	[10;500]	207	126	10
	η	$[1e^{-7};1.5]$.55	$2e^{-3}$	$1e^{-7}$
XGBoost	N	[10;500]	210	494	500
	η	$[1e^{-3};2]$	1.98	1.61	2
	d_{max}	[2;50]	49	5	50
	sb_s	$[1e^{-3};1]$.36	.67	1
	cols	$[1e^{-7};1]$	1	.30	1
	γ	$[1;1e^4]$	$1e^4$	2666	1
	cw_{min}	[1;100]	62	57	1
	α	[0;100]	0	2	0
	λ	[0;100]	18	46	100

alternative forms. Another main issue is the bias-variance trade-off. Models depend on the observations they are trained on, and their capacity of generalisation is what determines which algorithm is good. If a model strongly captures the behaviour of a training set, it will probably perform poorly to predict unknown observations. This is a situation of low-bias/high-variance. Finding the right optimum is a complicated task. Also, Data-driven approaches are often stochastic by nature, and the score associated with each model should be interpreted as one value out of an unknown distribution of solutions. The models are static, as opposed to time-series. It is possible that taking away the time factor from the equation induces a higher bias. Using multivariate time-series forecasting methods is promising and some methods are already selected for future work, such as SARIMAX, Neural

Network, SVR and KNN with time. Regarding the optimisation, the optimisation ran on a restricted number of runs. Powell method did run until convergence was met, but Bayesian Optimisation and Genetic algorithm had limited iterations. It would be possible to change the termination criterion so that the optimization runs until there are no significant improvements on the loss function. This method could give better results, but there is a trade-off with the execution time. Figure 3 shows the plant observations compared to the entire set of observations predicted by a Random Forest. The observation from the power plant in blue seems a lot like random noise, and the model captured this noise. This is the most extreme example from the predictions, but it depicts probable overfitting. This could be confirmed/informed by predicting a totally new data set from the powerplant. The main solution to tackle the noise issue is to use time-series. Indeed the time-series focus on trend, seasonality and noise while static methods do not interpret noise. If static regressors are used, applying filters on the responses, especially low pass filters, might lead to a better generalization. Filtering data is a 'risky' process, because some important information might be lost in the process.

4. Conclusions and further work

A first-principle model with high prediction capacity and several data-driven models were compared to operational data from a coal-fired boiler. At first, the data-driven models were fit to the observation data with few hyper-parameter adaptations. Then regressors were fit to the data with optimised parameters using the same algorithms. The optimizations of these parameters have been performed with three different methods: Powell Conjugate Direction, Bayesian Optimisation and Genetic Algorithm. In general, the simpler models benefited more from the optimisation than ensemble methods. Data-Driven models outperformed the first principle model by at least 15%. Random Forest and XGBoost yielded the best results with a reduction of the error of 56% and 57% for nonoptimised models. After optimization, Support Vector Machine outperformed both with a 58% error reduction. The high accuracy of this model needs to be confirmed with other unknown observations, especially because the models can capture the random noise for outputs. The 'static' approach should then be compared to the dynamic one with the use of time-series analysis. Finally, artificial neural network structures should be added to the model list for both static and dynamic approaches.

Acknowledgement

We acknowledge financial support from the Polish-Norwegian Research Program for funding the InnCapPlant project (Grant NOR/POLNORCCS/0015/2019-00), and the NTNU Department of Energy and Process Engineering. We express our gratitude to Monika Rerak at the Cracow University of Technology for providing the data, which is essential to this work. A special thanks to Pr. Adil Rasheed for his precious advice for future work.

References

- [1] J. Smrekar, M. Assadi, *et al.*, "Development of artificial neural network model for a coal-fired boiler using real plant data," *Energy*, vol. 34, no. 2, pp. 144–152, 2009.
- [2] S. Chandrasekharan, R. C. Panda, and B. N. Swaminathan, "Statistical modeling of an integrated boiler for coal fired thermal power plant," *Heliyon*, vol. 3, no. 6, p. e00322, 2017.
- [3] H. Zhu, G. Zhao, *et al.*, "Nonlinear predictive control for a boiler-turbine unit based on a local model network and immune genetic algorithm," *Sustainability (Switzerland)*, vol. 11, no. 18, 2019.
- [4] A. Navarkar, V. R. Hasti, *et al.*, "A data-driven model for thermodynamic properties of a steam generator under cycling operation," *Energy*, vol. 211, p. 118973, 2020.
- [5] N. Abdul Manaf and A. Abbas, "Dynamic modelling and simulation of clean coal power generation," *Journal of Physics: Conference Series*, vol. 1447, no. 1, 2020.
- [6] E. Oko, M. Wang, and J. Zhang, "Neural network approach for predicting drum pressure and level in coal-fired subcritical power plant," *Fuel*, vol. 151, pp. 139–145, 2015.
- [7] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow*. O'Reilly, 2019.
- [8] J. Tan, J. Yang, *et al.*, "A critical look at the current train/test split in machine learning," 2021.
- [9] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, aug 2016.
- [10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [11] H. Zou and T. Hastie, "Regression shrinkage and selection via the elastic net, with applications to microarrays," 01 2004.
- [12] V. Vapnik, *The Nature of Statistical Learning Theory*. Q325.7.V37, Springer-Verlag, 2nd ed., 1999.
- [13] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, Aug. 2004.
- [14] P. J. Huber, "Robust Estimation of a Location Parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73 – 101, 1964.
- [15] L. Bottou and O. Bousquet, "The tradeoffs of large scale learning," in *Advances in Neural Information Processing Systems 20 (NIPS 2007)* (J. Platt, D. Koller, Y. Singer, and S. Roweis, eds.), pp. 161–168, NIPS Foundation (<http://books.nips.cc>), 2008.
- [16] A. Moore, "The anchors hierarchy: Using the triangle inequality to survive high dimensional data," *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence*, 01 2013.
- [17] F. Pedregosa, D. Cournapeau, *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [18] L. Breiman, "Random Forests," *Machine learning*, 2001.
- [19] H. Drucker, "Improving Regressors using Boosting Techniques," p. 9, 1997.
- [20] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The Computer Journal*, vol. 7, pp. 155–162, 01 1964.
- [21] E. Jones, T. Oliphant, *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [22] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. Adaptive computation and machine learning, Cambridge, Mass: MIT Press, 2006. OCLC: ocm61285753.
- [23] J. Snoek, L. Larochelle, *et al.*, "Practical Bayesian Optimization of Machine Learning Algorithms," Aug. 2012. Number: arXiv:1206.2944 arXiv:1206.2944 [cs, stat].
- [24] F. Nogueira, "Bayesian Optimization: Open source constrained global optimization tool for Python," 2014–.
- [25] A. F. Gad, "Pygad: An intuitive genetic algorithm python library," 2021.
- [26] S. Katoch, S. S. Chauhan, and V. Kumar, "A review on genetic algorithm: past, present, and future," *Multimedia Tools and Applications*, vol. 80, pp. 8091–8126, Feb. 2021.

Spray Drop Size Characterization in an External-Mixing Bluff-Body Atomizer based on Acoustics and Multivariate Analysis

Raghav Sikka^a, Maths Halstensen^b, Joachim Lundberg^a

^a*Department of Process, Energy and Environmental Technology, University of South-Eastern Norway*

Raghav.Sikka@usn.no

^b*Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway*

Maths.Halstensen@usn.no

Abstract

Air-assist atomizers have been widely used in various applications such as the aerospace industry, internal combustion engines, molten metal, food processing, etc. The mean drop size for these atomizers was obtained through the Shadowgraph imaging technique. This study aims to assess the feasibility of the acoustic chemometrics approach for classifying the atomizer types and predicting the mean drop size, such as Sauter mean diameter (SMD), for a two-phase spray atomizer employed. The droplet size measurements were carried out at three radial locations and one axial location for various air and liquid (water) flow rates. The acoustic signals were recorded through two different sensors: accelerometers and microphones. The main objective of this work is to implement prediction models for the mean drop sizes (SMD) measured at various locations. The model prediction is based on the dimensionless number B, whose unique values correspond to different two-phase flow working conditions. This analysis will further cater to the question that whether the acoustics chemometrics approach, including Principal Component Analysis (PCA) and Partial Least Squares Regression (PLS-R), is suitable for extracting valuable information such as predicting mean drop size (SMD) in two-phase flows through recorded acoustic signals.

Keywords: Acoustic Chemometrics, Multivariate Regression, Principal Component Analysis, Twin-fluid Spray, Mean droplet size

1. Introduction

Multi-phase flows can be found in various industrial applications ranging from fuel sprays in IC engines to petroleum pipelines. The most occurred case is two-phase flows in which gas and liquid interact to form various flow patterns generating vibrations and flow-based noise. Flow-based vibrations are classified into four types, in which acoustic resonance (flow-induced pulsations) and turbulent-induced excitation (FIV) are common in internal two-phase flows (Pettigrew and Taylor, 2016). Flow classification utilizing vibrations signals is present in literature (Miwa, Mori, and Hibiki, 2015). Flow rates and the vibration signals recorded using the PAT approach (sensors) showed a strong correlation (Evans, Blotter and Stephens, 2004). The peak frequency of these induced vibrations is proportionate to the flow parameters, such as void fraction (Ortiz-Vidal, Mureithi and Rodriguez, 2017). The time and frequency spectrum of the force fluctuations in two-phase flow were analyzed through flow-induced vibrations (Liu *et al.*, 2012). Though there are studies related to vibration-based analysis for flow classification, it lacks study on the effect of flow-induced

vibrations (FIV) on the flow parameters such as local void fraction, interfacial area, and particle size distribution.

The single-phase flow (air only) can significantly affect the flow-induced vibrations and the acoustic noise. When expanded to high speeds, the jets produce shock-associated noise, further exacerbating flow-induced vibrations (Tam, 1998). The two-phase flow study constituting both vibration study for flow-induced vibrations caused due to internal flows and acoustic analysis from acoustic energy emitted from gas-liquid coaxial flows followed by two-phase mixing is still not considered. There have been attempts to measure the local two-phase flow parameters such as void fraction, Sauter mean diameter with flow-induced vibrations study (Hibiki and Ishii, 1998), and acoustic emission method (Guo *et al.* 2014). But the combination of techniques is still not attempted to date. This work applies a united approach, including accelerometers for acquiring flow-induced vibrations and microphones for obtaining acoustic signals. The novelty of this method lies in correlating the flow parameter, i.e., Sauter mean diameter (SMD), with acquired acoustic data and

parameter prediction using PLS-R. The data fusion with both techniques (described in the methods section) is done in the study to predict the flow parameter such as mean droplet size (SMD). The dual advantage of this approach is that vibrations, an inherent part of fluid flows in piping in industrial plants, can be used for both flow parameter prediction and dynamic stress analysis to estimate fatigue or structural damage.

2. Materials and Methods

The study was conducted utilizing the CMOS-based high-speed camera (Photron SA-Z), Nikon macro lens (80-200 mm), and Questar long-distance microscope (QM1). The rig setup built in-house at the process and energy department laboratory was used for the experiments. The different atomizers configuration with varying cone distances (L_c) and similar orifice (throat) diameters ($d = 3.0$ mm) were attached to the lance mounted at the Bosch Rexroth traverse system. Fig. 1 shows the airflow patterns for three atomizers.

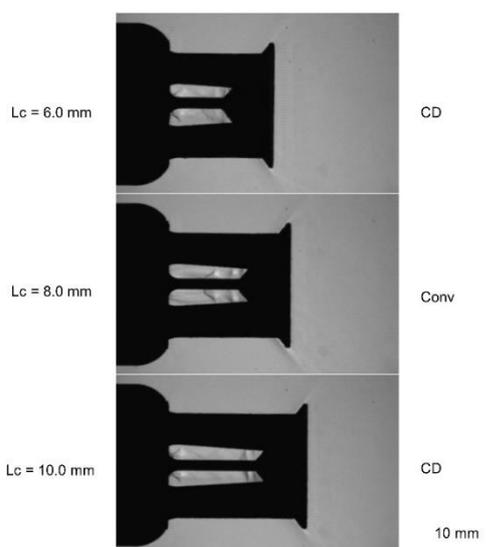


Figure 1: Airflow patterns in atomizers with three different cone distances (L_c).

Yokogawa Rotamass and Endress Hauser Promass 83 (Coriolis type) flowmeter used water and air rate measurements. In the test rig (Figure 2), the compressor with 7.0 bar (g) capacity was employed for the air supply, and the water flow supply pump by Froster AS company was used. The water flow rate employed was 100, 200 and 300 kg/h, and the airflow rate employed are 20, 30 and 40 kg/h. Therefore, nine experiments were done at specific air-to-liquid mass ratios (ALR) and Weber number (We) based on liquid sheet velocity. The range of dimensionless number corresponds to ALR

and the We number is mentioned in (Sikka, Halstensen and Lundberg, 2022)

ALR is defined as:

$$ALR = \frac{m_{air}}{m_{liquid}} \quad (1)$$

where mass flow rate in kg/hr.

Weber number is defined as:

$$We = \frac{\rho U^2 t}{\sigma} \quad (2)$$

A new dimensionless number (B) (depicted in Table 1) was employed, which is defined as:

$$B = We \cdot ALR \quad (3)$$

The experiments were conducted at STP such that fluid physical properties are assumed to be standard values. The spray formation in all three different cone distance (L_c) atomizers is illustrated in Fig. 2. The combined experimental setup for the spray imaging/drop size measurements and acoustic emission study is depicted in Fig. 3.

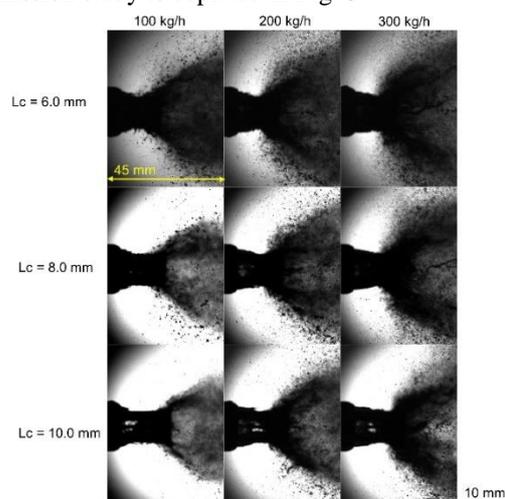


Figure 2: Images showing the spray pattern for 20 kg/h airflow rate at various liquid flow rates and the image size scaled (yellow line).

The acoustic arrangement employed two Piezoelectric types, 4518 accelerometers from Bruel & Kjaer, Denmark, to collect noise/vibration data. Acoustic readings were procured using two electret condenser type Veco Vansonic PVM-6052-5P382 omnidirectional microphones (mounted on an arc at 300 mm from the spray centerline) with a sensitivity of -38 dB and signal-to-noise ratio of 58 dB. The microphone frequency ranges from 50 Hz to 16 KHz. The microphones were mounted along the arc at $\theta = 90^\circ$ and 150° from perpendicular to the nozzle axis at $R = 100D$, termed "far-field" measurements (Wong et al., 2020). The vital point is that the acoustic data were recorded in a non-anechoic

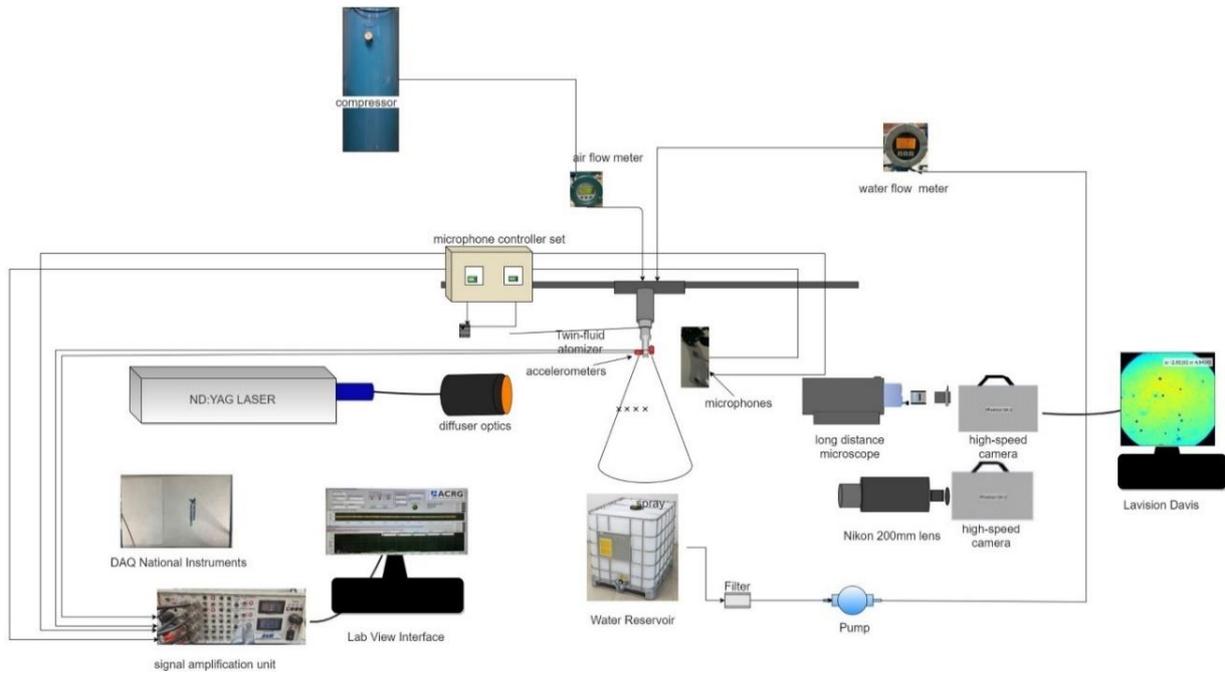


Figure 3: Schematic of the experimental setup for shadowgraphy technique along with the acoustic chemometric.

chamber, affecting the signal through stray noise. A data acquisition device (DAQ) from National Instruments (USB-6363), a signal amplification module, and a personal laptop with an in-house LabView interface was employed for signal acquisition. LabVIEW-based in-house created interface (Halstensen et al., 2019) was used for the acoustic chemometrics signal collection and signal conditioning. The signal processing was carried out on the acquired signal of 8192 recorded samples. The time-series signal was multiplied by a window (Blackman Harris) to avoid spectral leakage in the acoustic spectrum. This signal is finally transformed into the frequency domain using Discrete Fourier Transform. The Discrete Fourier Transform transforms a sequence of N complex numbers $\{x_n\} := x_0, x_1, \dots, x_{N-1}$ into another sequence of complex numbers, $\{X_k\} := X_0, X_1, \dots, X_{N-1}$, which is defined by equation (4):

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \quad k = 0, \dots, N-1 \quad (4)$$

A more advanced and efficient form of the DFT is the Fast Fourier Transform (FFT) (Ifeachor and Jervis, 1993), implemented for fast real-time calculations

2.1. Drop Size Measurements

The laser-based shadowgraphy method measured the mean drop size with a CMOS high-speed camera (Photron SA-Z model). The spray was illuminated by dual-cavity ND: YAG Laser (Photonics industries DM60-532 DH model) at 532 nm (green

light). The uniform speckle-free light background was achieved with diffuser optics. Questar's long-distance microscope (QM1) provides a field of view (FOV) of 8.445 mm x 8.445 mm. The ParticleMaster software package incorporated in Davis 10.1 version (LaVision) is used for droplet sizing. The calibration plate provides a depth of field (DOF) of $\sim 17:1$. The minimum pixel used for particle

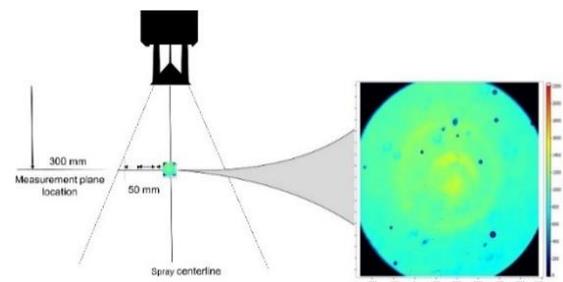


Figure 4: Schematic of the drop size measurement locations for shadowgraphy

detection is 3 pixels (in the area). The images were recorded at four locations depicted in Fig. 4 (line marked) – each 50 mm apart at the radial axis at 300 mm downstream from the outlet. Though, 500 images give convergence for mean droplet sizes. However, 1000 images were recorded for each measurement location which mitigates the measurement uncertainty ($<1\%$). Laser intensity (in the current (A)) was set adequately to provide uniform background in proportion to the droplet

density of the shadowgraph picture. Note that the camera pixel resolution allowed for drop size measurements of 16–2000 μm . The mean drop size (SMD) for different radial locations is shown in Fig. 5.

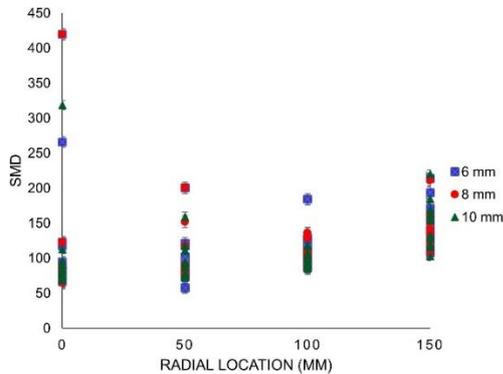


Figure 5: Mean drop size (SMD) for radial locations for all atomizers.

3. Results and Discussions

The acoustics spectrum is hard to analyze due to its multi-dimensional nature. Principal Component Analysis (PCA) is a dimensionality reduction technique that simplifies the analysis and reveals hidden patterns/structures. PCA projects the data into the new orthogonal plane, whose first principal component (PC1) is aligned in such a way that maximizes the variance. The new mean-centred plane is given by scores (T) and loadings (P) and residual (E). It is given by

$$X = T P^T + E \quad (5)$$

The Nonlinear Iterative Partial Least Squares (NIPALS) algorithm developed (Wold, Esbensen and Geladi, 1987) was used for its many advantages. The method is unsupervised due to its independent Singular Value Decomposition (SVD) on the data.

The scores plot (Fig. 6) depicts how the acoustic spectrum was segregated as colour clusters for three different atomizers based on the tests carried out at various fluid flow rates. For the 6mm (L_c) atomizer, the maximum variance is in the first principal component (PC1) direction. Whereas for the 8 mm (L_c) atomizer, cluster points scattered in the PC2 direction show that PC2 contains valuable information. The loading plot (Fig. 7) shows that the information is there in all the frequencies for accelerometers, reflecting the PCA classification model capability. In contrast, for microphones, frequencies recorded are from a narrow spectrum.

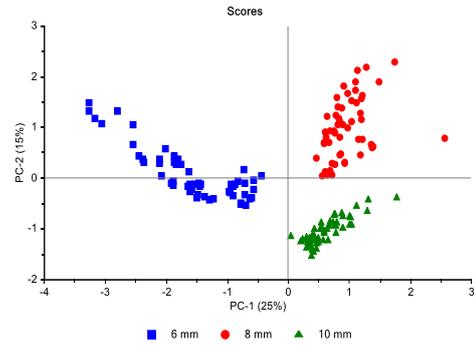


Figure 6: Score plot t1-t2 for all three atomizers.

The prediction model builds upon the regression-based method. Partial Least Squares regression (PLS-R) is a supervised method used to calibrate the predicting models, as explained in the PLS tutorial (Geladi and Kowalski, 1986).

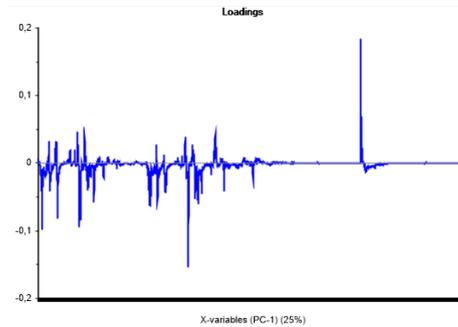


Figure 7: loadings plot for all the sensors.

PLS-R is a more advanced version of other regression techniques like MLR, PCR, etc. The robustness lies in the fact that model parameters vary little when new calibration samples are taken from the population. It builds on two-variable blocks, X and Y, representing training data. The NIPALS algorithm is used for PLS-R modelling (Halstensen, 2020). The X data matrix contains the frequency spectra in our study, and Y is a vector containing the mean drop size (SMD) values for a particular radial location. The regression model for mean drop size prediction is based on both accelerometers and microphone data.

The acoustic spectra used to calibrate the PLS-R model was a 162 x 8192 matrix, each sensor containing 162 frequency spectra. Each spectrum has 2048 frequencies ranging from 0 to 200 KHz for each sensor. The test set validation (50% data) was performed for alternate data values in the column. The root mean squared error of prediction (RMSEP) value, RMSEP (%) the slope, and the correlation coefficient (R^2) (Pearson) are commonly used in evaluating the different prediction models.

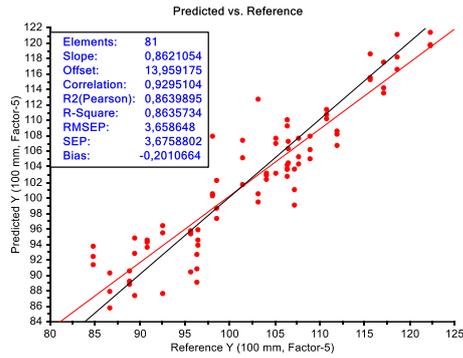


Figure 8: Predicted Vs. Reference (B) value. The target line (black) and regression line (red) are indicated.

The RMSEP is defined as

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_{i,predicted} - y_{i,reference})^2}{n}} \quad (3)$$

Where i = sample index number, n = total samples, RMSEP= Root Mean Squared Error of Prediction. The slope of 0.86 matches well with the target slope. RMSEP value comes out to be 3.65 with a correlation coefficient (R^2) value of 0.86 for a 100 mm location (Fig. 8). The loading weights linked the X matrix block to the Y-matrix through weights based on acoustic data (Fig. 9).

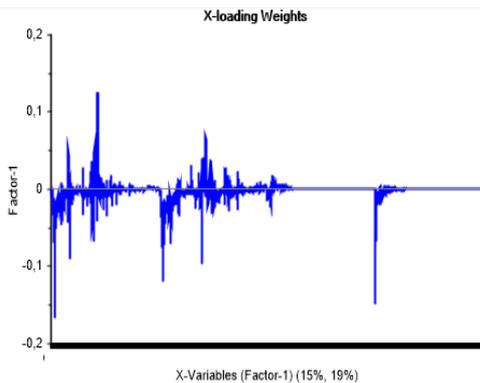


Figure 9: Loading weight plot for all sensors.

Table 1. Prediction models parameters

SMD Location	Prediction parameters (5 factors)			
	Slope	RMSEP	RMSEP (%)	R^2 Pearson
0 mm	0.74	7.22	12.2	0.75
50 mm	0.84	6.93	9.71	0.82
100 mm	0.86	3.65	9.60	0.86
150 mm	0.83	12.50	10.4	0.84

The prediction model parameters for all locations are given in Table 1.

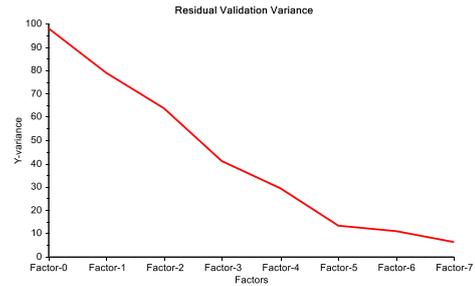


Figure 10: Residual validation variance plot.

Based on the residual validation variance plot (Fig. 10), the number of factors optimal for model prediction is 5, as Y-variance reduces drastically until 5 factors, then slightly decrease with more factors involved.

4. Summary

A feasibility study was conducted for the non-intrusive method using acoustic by applying multivariate data analysis techniques. The frequency data were recorded through accelerometers and microphones. The Principal Component Analysis (PCA) model reveals the clusters belonging to twin-fluid atomizers with the maximum variance in the first principal component (PC1) direction and first principal component (PC2) for the 6.0 mm cone distance (L_c) atomizer and 8.0 mm cone distance (L_c) atomizer, respectively. Prediction models based on the mean drop size (SMD) were fabricated using the Partial Least Squares regression (PLS-R) method. The prediction model works best for the 100 mm radial location as depicted by a low RMSEP (%) value of 9.60 and a high correlation coefficient (R^2) value of 0.86 when validated by test set validation.

References

- Evans, R. P., Blotter, J. D. and Stephens, A. G. (2004) 'Flow rate measurements using flow-induced pipe vibration', *Journal of Fluids Engineering, Transactions of the ASME*, 126(2), pp. 280–285. doi: 10.1115/1.1667882.
- Geladi, P. and Kowalski, B. R. (1986) 'PARTIAL LEAST-SQUARES REGRESSION: A TUTORIAL', *Analytica Chimica Acta*, (185), pp. 1–17.
- Guo, M. *et al.* (2014) 'On-line measurement of the size distribution of particles in a gas-solid two-phase flow through acoustic sensing and advanced signal analysis', *Flow Measurement and Instrumentation*, 40, pp. 169–177. doi: 10.1016/j.flowmeasinst.2014.08.001.
- Halstensen, M. (2020) 'Classification of Gases and Estimation of Gas Flow Rate Based on Unsupervised and Supervised Learning Respectively', (September), pp. 22–24. doi: 10.3384/ecp20176451.
- Hibiki, T. and Ishii, M. (1998) 'Effect of flow-induced vibration on local flow parameters of two-phase flow', *Nuclear Engineering and Design*, 185(2–3), pp. 113–125. doi: 10.1016/s0029-5493(98)00241-6.

- Hu, Y. *et al.* (2014) 'Online continuous measurement of the size distribution of pneumatically conveyed particles by acoustic emission methods', *Flow Measurement and Instrumentation*, 40, pp. 163–168. doi: 10.1016/j.flowmeasinst.2014.07.002.
- Ifeachor, E. C. and Jervis, B. W. (1993) 'Digital signal Processing'.
- Liu, Y. *et al.* (2012) 'Experimental study of internal two-phase flow induced fluctuating force on a 90° elbow', *Chemical Engineering Science*, 76(2012), pp. 173–187. doi: 10.1016/j.ces.2012.04.021.
- Miwa, S., Mori, M. and Hibiki, T. (2015) 'Two-phase flow induced vibration in piping systems', *Progress in Nuclear Energy*, 78(2015), pp. 270–284. doi: 10.1016/j.pnucene.2014.10.003.
- Ortiz-Vidal, L. E., Mureithi, N. W. and Rodriguez, O. M. H. (2017) 'Vibration response of a pipe subjected to two-phase flow: Analytical formulations and experiments', *Nuclear Engineering and Design*, 313, pp. 214–224. doi: 10.1016/j.nucengdes.2016.12.020.
- Pettigrew, M. J. and Taylor, C. E. (2016) 'Two-Phase Flow-Induced Vibration : An Overview', 116(August 1994).
- Sikka, R., Halstensen, M. and Lundberg, J. (2022) 'Characterization of the Flow (Breakup) Regimes in a Twin-Fluid Atomizer based on Nozzle Vibrations and Multivariate Analysis', in *62nd International Conference of Scandinavian Simulation Society, SIMS 2021, September 21-23, Virtual Conference, Finland*, pp. 22–27. doi: 10.3384/ecp2118522.
- Tam, C. K. W. (1998) 'Jet noise: Since 1952', *Theoretical and Computational Fluid Dynamics*, 10(1–4), pp. 393–405. doi: 10.1007/s001620050072.
- Wold, S., Esbensen, K. and Geladi, P. (1987) 'Principal Component Analysis', *Chemometrics and intelligent laboratory systems*, 2(1–3), pp. 37–52. Available at: <http://files.isec.pt/DOCUMENTOS/SERVICOS/BIBLIO/Documentos de acesso remoto/Principal components analysis.pdf>.
- Wong, M. H. *et al.* (2020) 'Azimuthal decomposition of the radiated noise from supersonic shock-containing jets', *The Journal of the Acoustical Society of America*, 148(4), pp. 2015–2027. doi: 10.1121/10.0002166.

Evaluating the acceptability and accuracy of Phasepy as a Python framework to calculate the interfacial properties and phase equilibrium

Javad Tavakolifaradonbe, Sumudu Karunaratne*, Knut Vågsæther

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway
h.javadt@yahoo.com, sumudu.karunaratne@usn.no, Knut.Vagsather@usn.no

Abstract

Phasepy is a scientifically defined open-source package in python for computational thermodynamics. Phasepy indeed calculates the interfacial properties and fluid phases equilibrium using an equation of state. In addition, Phasepy enables the scientists to optimize the relevant parameters to the equilibrium of multicomponent vapor-liquid, liquid-liquid, or vapor-liquid-liquid mixtures. The Phasepy can model the equilibrium in the continuous approach (combining a cubic equation of state and a mixing rule) or the discontinuous approach (using a virial equation and an activity coefficient model). So, this study is to develop a code in a continuous approach using a combination of Soave-Redlich-Kwong (SRK) or Peng Robinson (PR) as the equation of state and quadratic mixing rule (QMR) and modified-Huron-Vidal mixing rule (MHV) as the mixing Rule. Although the algorithm of the developed model is new, it is tried to utilize the predefined function of Phasepy to calculate fluid phase equilibrium and interfacial properties. In fact, the five well-performed previous experimental studies are modeled using Phasepy, and in the following, the outputs of the developed models are compared with the relevant experimental results. The bubble point features, dew point features, liquid and gas composition, and density of multicomponent mixtures are considered parameters in this extended study to evaluate the accuracy of the Phasepy function based on experimental results.

1. Introduction

To study the homogenous and the inhomogeneous behavior of fluids it is required to know about interfacial properties and fluid phase equilibria. These two physical properties enable scientists to design a process and an operation. From one side, the phase equilibrium discloses the physical limitations of a separation process and reveals the required stages for reaching equilibrium. On the other side, interfacial properties determine the efficiency and the size of equipment [1].

To study the feasibility of a process, firstly the most appropriate thermodynamic model for experimental data should be selected, and even sometimes operations need to be optimized. Therefore, it is required to study equilibrium thermodynamics [2]. Studies have shown that finding a perfect model and algorithms for computing phase equilibria are the most important concern of industries, to the extent that, industries prefer to not invest time and money in developing a new algorithm for phase equilibrium calculations even when it is required [3]. So, we see oil and gas industries are still using the classical equation of states (EoS). Indeed, developing a new algorithm based on contemporary EoS requires industries to provide a dissimilar set of parameters. In other words, the necessity for developing a more precise molecular-based model would increase computational expenditures [3], [4].

Prausnitz et al. [5] and Michelsen and Mollerup [6] discussed the fundamental computation for calculating phase equilibria in isothermal isobaric two-phase flash, liquid-liquid equilibrium, heteroazeotropic, bubble points, and dew points. Derived solution methods are based on minimizing Gibbs free energy of the system or criteria for isofugacity. They applied a combination procedure of Newton methods, successive substitution, and second-order minimizations.

Commercial solutions like Aspen Plus, gPORMS, etc. are being developed to simulate the process and calculate phase equilibria. Although these simulators consist of some advantages like different thermodynamic models, an extended database on phase equilibria, and stable manners, some cons should be mentioned also, including license fees, and the limitation of manipulating prefabricated processes. In addition, these commercial software does not use square gradient theory (SGT) for interfacial descriptions. As a theoretical approach, SGT in connection with a continuous EoS specifies the binodal boundary. Moreover, SGT makes a connection between homogenous phase and transition zone and as a result characterizes interfacial tension, Gibbs energy, and density profile. To deal with this type of problems, scientists are nominating homemade programs as a reliable solution [7]–[10].

This study is to use a Python-based module, namely Phasepy, to calculate interfacial properties and phase equilibrium computation and evaluate how much the defined module is applicable for calculating thermodynamic properties in multicomponent mixtures. Then firstly a short description of Phasepy is provided and then it is tried to evaluate the accuracy of calculations based on the previously performed experimental studies.

2. Theory and Methodology

Although more than 200 EoS were published by 1949, Redlich and Kwong made effort to deal with the limitations and revive the van der Waals EoS for high- and low-density fluids. They proposed following EoS [11], [12]:

$$P = \frac{RT}{V-b} - \frac{a_c \alpha(T)}{V(V+b)} \quad (1)$$

Where:

$$\alpha(T) = \frac{a}{T^{0.5}} \quad (2)$$

$$a_c = \frac{\Omega_a R^2 T_c^{2.5}}{P_c} \quad (3)$$

$$b = \frac{\Omega_b R T_c}{P_c} \quad (4)$$

$$\Omega_a = 0.4278 \quad \text{and} \quad \Omega_b = 0.0867$$

Although RK EoS does not possess a considerable background theoretically, this model provides acceptable results. In the following, Soave presented a new version of RK EoS by keeping the RK volume functionality and redefining α as a function of reduced temperature and acentric factor [12].

$$P = \frac{RT}{V-b} - \frac{a_c \alpha(T_r, \omega)}{V(V+b)} \quad (5)$$

$$a_c = \frac{0.42747 R^2 T_c^{2.5}}{P_c} \quad (6)$$

$$b = \frac{0.08664 R T_c}{P_c} \quad (7)$$

$$\alpha(T_r, \omega) = [1 + (0.480 + 1.574\omega - 0.176\omega^2)(1 - T_r^{0.5})]^2 \quad (8)$$

The SRK now is the most popular EoS in hydrocarbon-related industries. Then scientists tried to define new temperature model $\alpha(T_r, \omega)$ and modify volume dependency of pressure-related terms. Peng and Robinson calculated $\alpha(T_r, \omega)$ again and modified SRK EoS. PR EoS is:

$$P = \frac{RT}{V-b} - \frac{a_c \alpha(T_r, \omega)}{V(V+b)+b(V-b)} \quad (9)$$

$$a_c = \frac{0.45724 R^2 T_c^{2.5}}{P_c} \quad (10)$$

$$b = \frac{0.07780 R T_c}{P_c} \quad (11)$$

$$\alpha(T_r, \omega) = [1 + (0.37464 + 1.54226 - 0.26992\omega^2)(1 - T_r^{0.5})]^2 \quad (12)$$

Theoretically, mixing rules have been developed to connect multicomponent mixture parameters to pure fluid parameters [12]. Mostly, classical van der Waals mixing rules are applied as:

$$a = \sum \sum x_i x_j a_{ij} \quad (13)$$

$$b = \sum \sum x_i x_j b_{ij} \quad (14)$$

$$c = \sum \sum x_i x_j c_{ij} \quad (15)$$

The volume parameters, b_{ij} and c_{ij} , are calculated with arithmetic mean and for force parameter, a_{ij} , the geometric mean is utilized. Therefore, phase equilibrium can be correlated more accurately [12].

$$a_{ij} = \sqrt{a_i a_j} (1 - k_{ij}) \quad (16)$$

$$b_{ij} = \frac{1}{2} (b_i + b_j) (1 - \beta_{ij}) \quad (17)$$

$$c_{ij} = \frac{1}{2} (c_i + c_j) (1 - \delta_{ij}) \quad (18)$$

Although these modifications can keep the concentration-related affinity of the parameters, a better modification is required for complex cases like supercritical fluid processes. k_{ij} , β_{ij} , and δ_{ij} are known as the interaction coefficients between components i and j . Moreover, regression analysis of real data or predictive correlations is applied to calculate these interaction coefficients. Some other studies have tried to understand the logic behind the binary interaction coefficient and pure species properties [11].

Phasepy is an open-source, and scientific package based on Python which has been developed to calculate interfacial properties and phase equilibrium. Object-oriented style of Phasepy enables users to apply small codes to calculate thermodynamics' properties. By using Phasepy it is possible to model a pure fluid or a multicomponent mixture fluid based on EoS. Then stability of phase equilibrium computation is evaluated. Finally, selecting continuous approach provides a possibility to investigate interfacial behavior based on SGT [10].

Firstly, the pure components and their properties, including critical temperature [K], critical pressure [bar], critical compressibility factor, critical volume [cm³/mol], and acentric factor should be defined. In this way, the *component* function enables the user to define components separately and then the *mixture* and the *add_component* functions define and develop the fluid mixture. The following code script shows how to define a three-component mixture of Nitrogen, Carbon dioxide, and Methane.

```
N2 = component(name = 'nitrogen', Tc
               = 126.2, Pc = 34, Zc
               = 0.289, Vc = 89.2, w
               = 0.038, GC = {'H2O': 1})
CO2 = component(name = 'carbon
                -dioxide', Tc = 126.2, Pc
                = 34, Zc = 0.289, Vc = 89.2, w
                = 0.038, GC = 'H2O': 1)
CH4 = component(name = 'methan',
                Tc = 190.6, Pc = 45.99,
                Zc = 0.286, Vc = 98.6, w
                = 0.012, GC = {'CH2': 1, 'CH2'
                : 1})
```

```
mix = mixture(N2, CO2)
```

```
mix.add_component(CH4)
```

Now the mixture is ready, and a model should be selected to compute phase equilibria and interfacial

Table 1: Required information for continuous or discontinuous modeling of a multicomponent mixture using Phasepy [10].

Type of modeling	Models	Component definition	Interaction
Discontinuous model $\phi - \gamma$	ideal gas, Abott Van Ness, NRTL, Wilson, UNIFAC	Critical temperature Critical pressure Critical compressibility	Activity coefficient models need specific interaction parameters
Continuous model $\phi - \phi$	VdW, PR, PR, PRSV, RK, RKS,	Critical temperature Critical pressure Acentric factor	Specific activity coefficient models

properties. Phasepy has provided two approaches to model a mixture. Throughout the $\phi - \gamma$ approach, firstly, a virial expansion is applied to model the vapor phase deviation (ϕ). Then, an activity coefficient model like NRTL, modified-UNIFAC, wilson, or Redlich-Kister, is utilized to explain liquid phase deviation (γ).

mix. NRTL(alpha, g, g1)

*model = virialgamma(mix, virialmodel
= 'Abbott', actmodel = 'nrtl')*

The second approach uses a classical cubic EoS and a mixing rule to model the system. EoS functions in Phasepy account for van der Waals (vdW), Redlich-Knowg (RK), Redlich-Knowg-Soave (RKS), Peng-Robinson (PR), Peng-Robinson-Stryjek-Vera EoS (PRSV), and the Pàeloux et al. Moreover, as a mixing rule it is possible to apply quadratic mixing rule (QMR), modified-Huron-Vidal mixing rule (MHV), and Wong-Sandler mixing rule (WS) [10]. In essence, by selecting the preferred EoS and the appropriate mixing rule, it would be possible to calculate interfacial properties and phase equilibrium. For instance, based on Peng Robinson EoS, NRTL for activity coefficient, and MHV mixing rule we can write the following code scripts: *eos = preos(mix, 'mhv1_nrtl')*

For more information about the EoS, mixing rules, and activity coefficient following references are suggested, [13], [14].

Tab. 1. presents the required information according to the type of modeling, including continuous or discontinuous using Phasepy.

3. Phasepy Two-phase equilibrium calculation

To calculate the isothermal isobaric flash composition of a two-phase mixture, Phasepy needs to revamp the phase compositions and solve Rachford-Rice mass balance continually. Therefore, the developer applied accelerated successive substitution (ASS) [15] and Halley's method to deal with the mentioned difficulties respectively.

Based on Rachford-Rice mass balance:

$$\sum_{i=1}^c \frac{z_i(K_i - 1)}{1 + \psi(K_i - 1)} = 0 \quad (19)$$

$$K_i = \frac{x_i^\alpha}{x_i^\beta} = \frac{\hat{\phi}_i^\alpha}{\hat{\phi}_i^\beta} \quad (20)$$

α and β represent two phases and ψ depicts the fraction of phase β . In the case of not convergency especially in higher pressures, the algorithm is

altered with a second-order procedure to find the minimum Gibbs free energy [10].

$$\sum_{i=1}^c (F_i^\alpha \ln \hat{f}_i^\alpha + F_i^\beta \ln \hat{f}_i^\beta) \quad (21)$$

Where F and \hat{f} refer to the number of the mole and effective fugacity respectively, and i is the species index. Apart from the flash calculation, Phasepy possesses the ability to calculate saturation points. Computing bubble points and dew points in a two-phase vapor-liquid-equilibrium (VLE) is based on the suggested method by Mollerup [6]. In this way throughout an inner loop phase compositions are updated using ASS, and the outer loop utilizes the quasi-Newton method to recalculate the pressure or the temperature. in the case of slow convergence, SciPy optimization routines are applied to solve equations 4, and 5 based on the iteration factor K , equilibrium constant.

$$f_i = \ln K_i + \ln \hat{\phi}_i^v(y, T, P) - \ln \hat{\phi}_i^l(y, T, P) \quad (22)$$

$$f_{c+1} = \sum_{i=1}^c (y_i - x_i) \quad (23)$$

The author of Phasepy presented Fig. 1 as the algorithm of Phasepy in computing bubble points or dew points. It is worth saying that the model requires the user to guess an initial value for saturation point, though the algorithm has ability enough for tolerating the initial guesses with large errors.

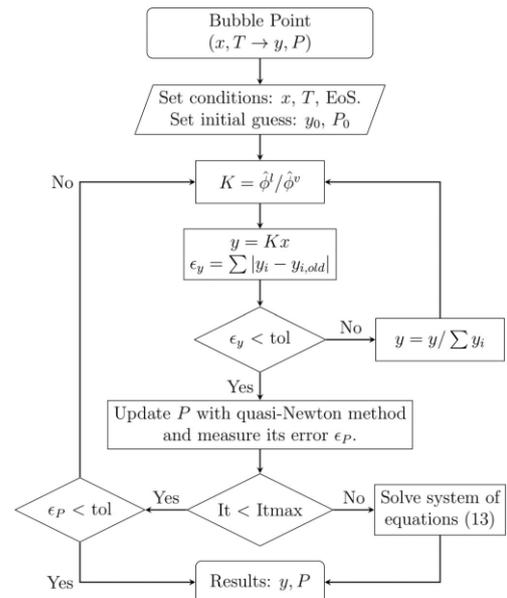


Figure 1: proposed algorithm for calculating the saturation points in Phasepy [10].

The following code scripts show how the saturation points or flash points can be calculated.

```
flash(x_guess, y_guess, equilibrium, Z, T, P, model)
bubbleTy(y_guess, T_guess, X, P, model)
bubblePy(y_guess, P_guess, X, T, model)
dewPx(x_guess, P_guess, y, T, model)
dewTx(x_guess, T_guess, y, P, model)
```

Where:

- x_guess (array): mole fraction of phase 1 (initial guess)
- y_guess (array): mole fraction of phase 2 (initial guess)
- T_guess (float): equilibrium temperature [K] (initial guess)
- P_guess (float): equilibrium pressure [bar] (initial guess)
- $equilibrium$ (string): Two-phase system, including 'LL' (liquid-liquid) or 'LV' (liquid-vapor)
- Z (array): Overall mole fractions of components
- T (float): temperature [K]
- P (float): Pressure [bar]
- $model$ (object): prepared model based on EoS and mixing rule (eos)

4. Results and Discussion

To evaluate how much the developed model is reliable to predict phase equilibrium three two-component mixtures at two pressures and two four-component mixtures are studied and outputs are compared with the experimental results.

In this way, average absolute relative deviation (AARD) and absolute maximum deviation (AMD) are studied.

$$AARD = \frac{1}{n} \sum_{i=1}^n \frac{|Exp_i - Cal_i|}{Exp_i} \times 100 \quad (24)$$

$$AMD = \max(|Exp_i - Cal_i|) \quad (25)$$

$i = 1, 2, \dots, n$

Knudsen et. al. [16] investigated the most appropriate mixing rule, including the Huron-Vidal rule, the MHV-model, the Schwarzenuber - Galivel-Solasttoug - Renon rule, and the density-dependent local composition rule, for SRK EoS. The authors showed that the Huron-Vidal rule and modified Huron-Vidal rule can be used as the best mixing rule. On the other hand, Pedersen et. al. [17] also suggested that when the system is only containing hydrocarbons and sour gas simple QMR can be adequate. Therefore, for binary mixtures, the Modified Huron Vidal mixing rule and Quadratic Mixing Rule are compared while SRK is used as the EoS, and in four-components mixtures, QMR is utilized as the mixing rule and two EoS, including SRK and PR, are compared.

Marlus et. al. [18] studied binary mixtures of Benzene + Cyclohexane, Benzene + Chlorobenzene, and Cyclohexane + Chlorobenzene at the pressure of

101.5 and 40 kPa. T-x,y diagrams for the mixtures are drawn in Fig. 2, 3, 4, and 5.

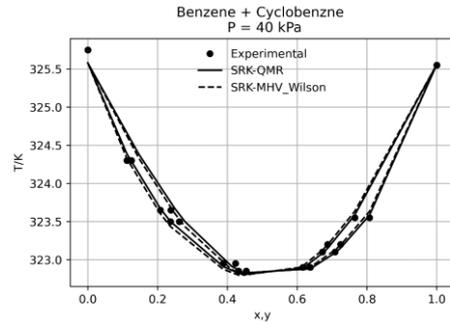


Figure 2: T-x,y diagram for Benzene (1) + Cyclohexane (2) at 40 kPa

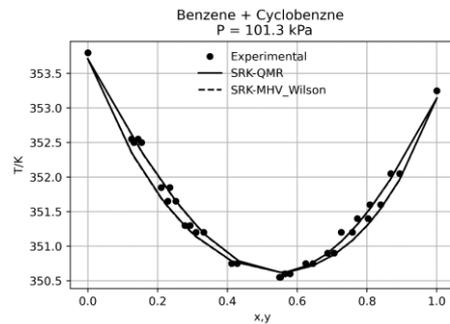


Figure 3: T-x,y diagram for Benzene (1) + Cyclohexane (2) at 101.3 kPa

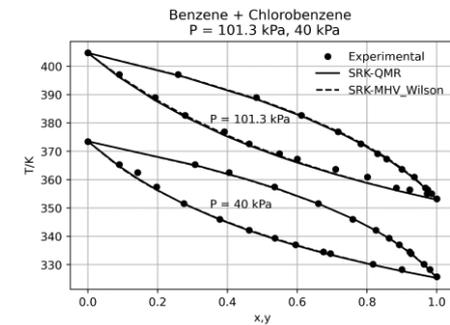


Figure 4: T-x,y diagram for Benzene (1) + Chlorobenzene (2) at $P = 40$ kPa, and $P = 101.3$ kPa

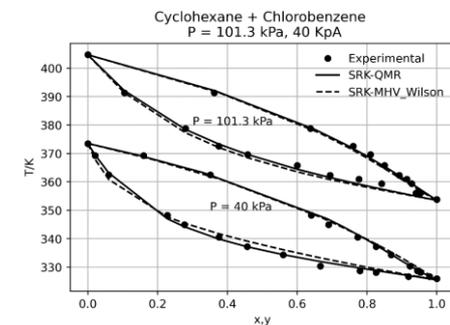


Figure 5: T-x,y diagram for Cyclohexane (1) + Chlorobenzene (2) at $P = 40$ kPa, and $P = 101.3$ kPa

Thomas et. al. [19] experimentally studies the phase equilibrium of LNG. The authors indeed performed experiments on two different four-component

mixtures, and then tried to model the behaviors using PR EoS. This study modeled these two mixtures using the Phasepy package. P-x-y diagrams are shown in Fig. 6 and 7.

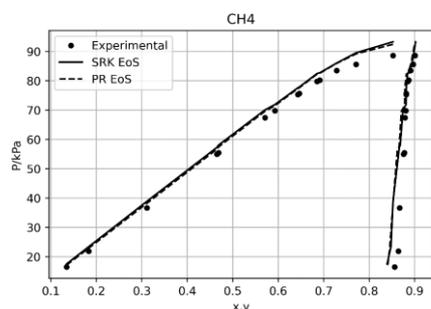


Figure 6: P-x,y diagram for CH4 in CH4, C2H6, C3H8, and n-C4H10 mixture

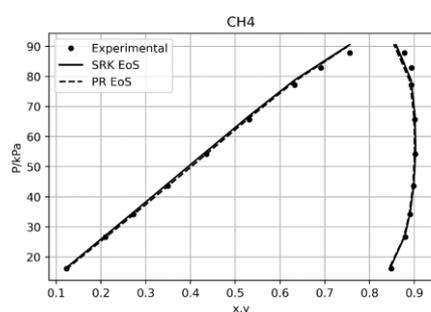


Figure 7: P-x,y diagram for CH4 in CH4, C2H6, C3H8, and i-C4H10 mixture

For all modelled mixture, AARD and AMD are calculated and shown in Table. 2. Therefore, it is possible to compare the consistency of EoS and mixing rules for different mixtures.

Table 2: calculated AARD and AMD for different mixtures

Mixture	P [kPa]	EoS	Mix. rule	AARD (%)	AMD
Benzene + Cyclohexane	40	SRK	QMR	2.87	0.020
			MHV	3.09	0.021
Benzene + Cyclohexane	101.3	SRK	QMR	2.49	0.019
			MHV	2.26	0.019
Benzene + Chlorobenzene	40	SRK	QMR	1.42	0.043
			MHV	1.55	0.046
Benzene + Chlorobenzene	101.3	SRK	QMR	1.73	0.019
			MHV	1.43	0.021
Cyclohexane + Chlorobenzene	40	SRK	QMR	2.93	0.045
			MHV	2.88	0.030
Cyclohexane + Chlorobenzene	101.3	SRK	QMR	1.77	0.045
			MHV	2.06	0.043
CH4, C2H6, C3H8, and n-C4H10		SRK	QMR	0.79	0.029
			PR	0.83	0.032

CH4, C2H6, C3H8, and i-C4H10	SRK	QMR	0.30	0.019
	PR		0.53	0.023

Based on Tab. 2, at the pressure of 40 kPa, SRK-QMR is the better model to predict the equilibrium of the Benzene (1) + Cyclohexane (2) mixture. However, at pressure 101.3 kPa, SRK-MHV-Wilson is the better model.

In the case of the Benzene + Chlorobenzene mixture, Fig. 4, in lower pressure, SRK-QMR is fitted and experimental results are perfectly predicted. However, at Pressure 101.3 kPa SRK-MHV_Wilson is the better model to predict the system.

For Cyclohexane + Chlorobenzene mixture as shown in Fig. 5, in lower pressure, SRK-MHV_Wilson is the better model to be fitted with the experimental results. Moreover, at Pressure 101.3 kPa SRK-QMR and SRK-MHV_Wilson are more closed, but, SRK-QMR is the better model.

Results show that PR and SRK EoS both, are an adequate choice for modeling, but in CH4, C2H6, C3H8, and i-C4H10 mixture, and, for CH4 in CH4, C2H6, C3H8, and n-C4H10 mixture SRK is the better EoS for modeling.

5. Conclusion

Knowing the interfacial properties and fluid phase equilibria are required to study the fluid behavior. In addition, thermodynamic models enable scientists to investigate the feasibility of an operation and process. However, finding an appropriate thermodynamic model and algorithm for computing the phase equilibria and interfacial properties is challenging. Although these days some commercial software has been developed to solve these difficulties, there are still some restrictions, including licenses fee, limitations in manipulating the defined process and lack of square gradient theory (SGT) for interfacial descriptions. Therefore, scientists have been provoked to apply alternative solutions like homemade programs.

Phasepy is a scientifically defined open-source package in python for computational thermodynamics. This package has been developed based on the most popular and reliable theories to calculate the interfacial properties and phase equilibria. The simplicity and accuracy of Phasepy enable studies to compare different EoS, mixing rules, etc.

This study tries to evaluate how much the developed methods based on Phasepy are able to predict the behavior of multicomponent mixtures. In this way, four binary mixtures and two four-components mixtures are modeled. Then the developed models are validated based on the performed experimental results. The results depict that Phasepy could be a solution for thermodynamic modeling, if an

appropriate EoS, mixing rule, and activity coefficient model were selected.

References

- [1] J. D. Seader and Ernest J. Henley, *Separation-Process-Principles-with-Applications-Using-Process-Simulators*, vol. 1, no. 4. 2006.
- [2] P. C. Wankat, *Separation Process Engineering*, vol. 49, no. 11. 2004.
- [3] E. Hendriks *et al.*, “Industrial requirements for thermodynamics and transport properties,” *Industrial and Engineering Chemistry Research*, vol. 49, no. 22, 2010, doi: 10.1021/ie101231b.
- [4] “Applied Thermodynamics of Fluids,” *Chemistry International -- Newsmagazine for IUPAC*, vol. 33, no. 2, 2014, doi: 10.1515/ci.2011.33.2.27a.
- [5] J. Prausnitz, T. Anderson, E. Grens, C. Eckert, R. Hsieh, and J. O’Connell, “Computer calculations for multicomponent vapor-liquid and liquid-liquid equilibrium,” 1980, doi: 10.1016/0378-3812(81)80016-7.
- [6] J. Mollerup and M. L. Michelsen, *Thermodynamic Models: Fundamentals & Computational Aspects*. 2004.
- [7] A. Mejía, E. A. Müller, and G. Chaparro Maldonado, “SGTPy: A Python Code for Calculating the Interfacial Properties of Fluids Based on the Square Gradient Theory Using the SAFT-VR Mie Equation of State,” *Journal of Chemical Information and Modeling*, vol. 61, no. 3, 2021, doi: 10.1021/acs.jcim.0c01324.
- [8] J. M. Garrido, A. Mejía, M. M. Piñeiro, F. J. Blas, and E. A. Müller, “Interfacial tensions of industrial fluids from a molecular-based square gradient theory,” *AIChE Journal*, vol. 62, no. 5, 2016, doi: 10.1002/aic.15190.
- [9] “Statistical mechanics of phases, interfaces, and thin films,” *Choice Reviews Online*, vol. 33, no. 09, 1996, doi: 10.5860/choice.33-5162.
- [10] G. Chaparro and A. Mejía, “Phasepy: A Python based framework for fluid phase equilibria and interfacial properties computation,” *Journal of Computational Chemistry*, vol. 41, no. 29, 2020, doi: 10.1002/jcc.26405.
- [11] J. O. Valderrama, “The state of the cubic equations of state,” *Industrial and Engineering Chemistry Research*, vol. 42, no. 8. 2003. doi: 10.1021/ie020447b.
- [12] J. M. Smith, H. C. van Ness, M. M. Abbot, and M. T. . Swihart, *Introduction to Chemical Engineering Thermodynamics Eight Edition*. 2018.
- [13] J. M. Smith, H. C. van Ness, M. M. Abbot, and M. T. . Swihart, *Introduction to Chemical Engineering Thermodynamics Eight Edition*. 2018.
- [14] J. O. Valderrama, “The state of the cubic equations of state,” *Industrial and Engineering Chemistry Research*, vol. 42, no. 8. 2003. doi: 10.1021/ie020447b.
- [15] C. M. Crowe and M. Nishio, “Convergence promotion in the simulation of chemical processes—the general dominant eigenvalue method,” *AIChE Journal*, vol. 21, no. 3, 1975, doi: 10.1002/aic.690210314.
- [16] K. Knudsen, E. H. Stenby, and A. Fredenslund, “A comprehensive comparison of mixing rules for calculation of phase equilibria in complex systems,” *Fluid Phase Equilibria*, vol. 82, no. C, 1993, doi: 10.1016/0378-3812(93)87159-X.
- [17] S. Valiollahi, B. Kavianpour, S. Raeissi, and M. Moshfeghian, “A new Peng-Robinson modification to enhance dew point estimations of natural gases,” *Journal of Natural Gas Science and Engineering*, vol. 34, 2016, doi: 10.1016/j.jngse.2016.07.049.
- [18] M. P. Rolemberg and M. A. Krahenbuhl, “Vapor-liquid equilibria of binary and ternary mixtures of benzene, cyclohexane, and chlorobenzene at 40.0 kPa and 101.3 kPa,” *Journal of Chemical and Engineering Data*, vol. 46, no. 2, 2001, doi: 10.1021/je0000591.
- [19] T. J. Hughes *et al.*, “High pressure multi-component vapor-liquid equilibrium data and model predictions for the LNG industry,” *Journal of Chemical Thermodynamics*, vol. 113, 2017, doi: 10.1016/j.jct.2017.05.023.

Design optimization of small-scale ORC cycles for fluctuating heat source

Donghoi Kim^{a,*}, Rubén M. Montañés^a, Luca Riboldi^a, Lars O. Nord^b, Jan Spale^c, Vaclav Novotny^c

^a SINTEF Energy Research, Sem Sælands vei 11, Trondheim 7034, Norway

^bNTNU - The Norwegian University of Science and Technology, Department of Energy and Process Engineering, NO-7491, Trondheim, Norway

^cCzech Technical University in Prague, Faculty of Mechanical Engineering, Technická 4, Praha 6, 16607, Czechia,

donghoi.kim@sintef.no

Abstract

Organic Rankine cycles (ORC) are efficient technologies for waste heat recovery (WHR) at low to mid temperatures. For the design of ORC power cycles, several thermodynamic parameters should be considered. A challenge related to small scale (<50 kW) ORC cycles is to define the optimal process given frequent variability in a heat source. Many relevant applications require robust ORC systems to perform under varying heat source loads. This is an area where the body of knowledge must be further developed.

In this work, the design of small-scale ORC cycles with varying heat source conditions is addressed by means of system modelling, simulation, and optimization. A framework is presented that consists of multi-scale optimization for the design of small-scale ORC systems considering seasonal and hourly heat source variations. The framework is developed as a flexible tool allowing to include fit-for-purpose models of key elements of the cycle, such as expander and heat exchanger, to suitably simulate off-design performance.

The optimization framework has been tested on a case study representing a woodchips-fired micro-cogeneration unit via ORC. The case study is representative of an existing unit operating at the Czech Technical University (CTU) in Prague. The results indicate that the tool delivers an ORC design that has a 5 % larger accumulated power production with the hourly variation of the heat source during one year than the original ORC solely optimized at the design heat source condition. The optimal ORC system also shows a 33 % smaller nominal capacity and size of heat exchangers than the ORC at the reference design, indicating a potential reduction in the capital cost.

1. Introduction

For low temperature heat sources, organic Rankine cycles demonstrated to be an advantageous technology (Macchi and Astolfi, 2016). When considering ORC design, several thermodynamic parameters should be considered, as well as a several organic working fluids and process configurations. Optimization could be performed both on thermodynamic and techno-economic parameters (Colonna *et al.*, 2015). A main challenge related to small scale (<50kW) ORC cycles is to define the optimal process design given a specific application when there is frequent variability in heat source. Many relevant applications will require robust ORC systems to perform under varying heat source (Petrollese and Cocco, 2019). This is an area where the body of knowledge is less developed.

This paper presents a tool under development for the design optimization of ORC systems accounting for off-design operations. The primary focus will be on small scale distributed energy systems. The objective of this work is to describe the methodology at the basis of the tool and present the first implementation, highlighting the expected benefits by analyzing a case study. In the future, the tool will be further developed along different lines. The off-design models will be refined to include more advanced approaches, potentially validated on experimental data. The framework will also include multi-scale optimization opportunities, where the cycle components design will be optimized together with the process. For instance, the aerodynamic design of the expander will be embedded into the optimization framework. In view of future developments, the tool will strive to enable a high degree of flexibility. The goal is to ensure the

possibility to accommodate modules and layers in a simple manner, therefore allowing analyses at the requested level of complexity.

The paper is structured as follows. Section 2 describes the methodology used for the study, including the modelling work and optimization framework. The section concludes with an overview of the case study and boundary conditions for the analysis. Section 3 presents the results obtained from the given case study. Section 4 outlines the main findings.

2. Methodology

2.1. Tool for optimal design of ORC systems for small scale distributed energy systems

This study presents the first version of a tool for multi-scale design optimization of ORC systems. The tool is developed as a flexible platform for simulation and optimization. In this first version, the tool includes a model for simulation of ORCs as well as off-design models for predicting part load performances of key elements of the cycle. An optimization framework integrates the various models allowing to identify a design that results in the optimal performance considering all expected operating conditions. Future versions of the tool will allow for incorporating more advanced off-design models, better control structures and modules for multi-scale optimization. However, those aspects are not part of this study.

The approach used for design optimization is depicted in Figure 1. Given an objective function \bar{z} , a design is defined in terms of a set of selected independent variables (see Table 1). The simulation at such design point provides the process thermodynamic values of the cycle at design – in terms of mass flow rates, temperature and pressure levels and also sizes of the heat exchangers. That information is used for the off-design simulations that will be carried out for each operating conditions deemed relevant to describe the operation of the ORC. Each simulation returns the value of the objective function at the specific conditions. The overall objective function is the weighed sum of the specific values obtained from the off-design simulations. For this study, the main objective function was selected to be the accumulated produced electric power over the time horizon of operation. Such objective function is more suitable for systems where power is the main output, while our case study will be based on a combined heat and power (CHP) unit. However, the focus of this first

implementation is to prove the methodology. More detailed considerations on relevant objective functions will be made in future studies.

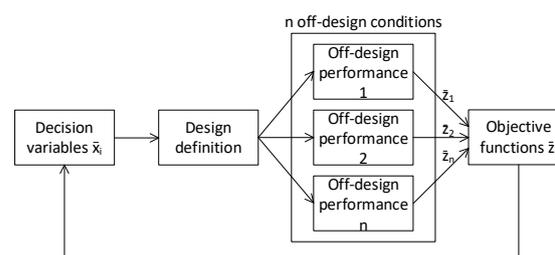


Figure 1. Flowchart of the optimization framework

2.1.1. Rankine Lab code

The Rankine Lab tool¹ (Agromayor and Nord, 2017) has been used as model basis for design optimization. It is an open-source tool in MATLAB that can be used to analyze and optimize Rankine cycles. It utilizes a gradient-based optimizer (SQP), and several cycle configurations are possible to analyze. In this design optimization tool, CoolProp, which is also open-source, is applied for thermodynamic property calculations during the ORC simulation. Various working fluids can be chosen within CoolProp and several parameters can be tuned to perform a simulation. For working fluids that are not supported by CoolProp, RefProp can be linked to the Rankine Lab tool via the CoolProp interface.

The ORC considered is shown in Figure 2 (note that the recuperator is originally included in the Rankine Lab tool, but it is not activated for this analysis). The working fluid selected is MM (hexamethyldisiloxane). For the design optimization, 5 decision variables and 8 inequality constraints are defined as shown in Table 1 and Table 2. The upper and lower bounds of the variables are also reported in Table 1.

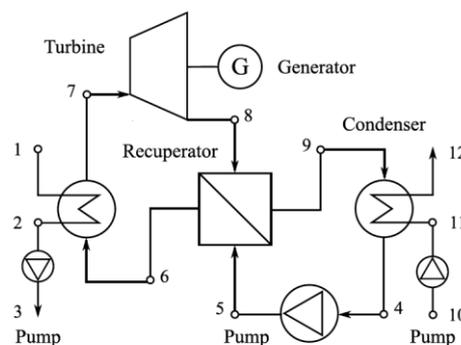


Figure 2. Process flowsheet of the reference ORC.

¹ For the open code and technical documentation, refer to <https://github.com/RoberAgro/RankineLab>

Table 1. Optimization variables and related bounds.

Independent variables	Lower bound	Upper bound
T ₃	T _{flue,min}	T _{flue,supply}
p ₄	p _{cycle,min}	p _{cycle,max}
h ₄	h _{sat} (T ₀)	h(T ₁ , p→0).
p ₇	p _{triple}	p _{cycle,max}
h ₇	h _{sat} (T ₀)	h(1000 K, p→0)

Table 2. Nonlinear optimization constraints.

Independent variables	Value
ΔT within the evaporator*	ΔT _{evap} ≥ ΔT _{evap,min}
ΔT within the condenser*	ΔT _{cond} ≥ ΔT _{cond,min}
Subcooling at pump inlet	ΔT _{sub,min} ≤ T _{sat} (p ₄) - T ₄ ≤ ΔT _{sub,max}
Superheating at expander inlet	ΔT _{sup,min} ≤ T ₇ - T _{sat} (p ₇) ≤ ΔT _{sup,max}
Vapor quality within expander	q _{exp} ≥ q _{exp,min}
Expander capacity	W _{exp} ≤ W _{exp,max}
Expander inlet temperature	T ₇ ≤ T _{exp,in,max}
Expander inlet pressure	p ₇ ≤ p _{exp,in,max}
Expander outlet pressure	T ₈ ≥ T _{exp,out,min}

*This constraint is evaluated at each discretization step of the heat exchanger

2.1.2. Code for off-design:

The off-design model was based on the work in (Riboldi and Nord, 2018) and adapted to an ORC. The inputs to the off-design model are the design parameters (see those parameters with subscript *d* in the following equations) obtained by the solution of Rankine Lab at the design tested.

The heat recovery unit (HRU) is modelled through the relation from Incropera et al. (Incropera et al., 2007), where the off-design heat transfer coefficient is calculated as:

$$UA = UA_d \left(\frac{\dot{m}}{\dot{m}_d} \right)^\gamma \quad (1)$$

where *U* is the overall heat transfer coefficient, *A* is the heat transfer area, \dot{m} is a mass flow rate and γ is the exponent of the Reynolds number in the heat transfer correlation. γ was set equal to 0.6. Assuming a shell and tube heat exchanger configuration, it is assumed that the heat transfer inside the tubes is the dominating factor (Orlandini et al., 2016). The condenser is more simply modelled as a fixed pressure component. The performance of the expander at off-design conditions is model based on the two following equations:

$$C_S = \frac{\dot{m} \sqrt{T_{in}}}{\sqrt{p_{in}^2 - p_{out}^2}} \quad (2)$$

$$\frac{\eta_T}{\eta_{T,d}} = 2 \sqrt{\frac{\Delta h_{T,is,d}}{\Delta h_{T,is}}} - \frac{\Delta h_{T,is,d}}{\Delta h_{T,is}} \quad (3)$$

where *C_S* is the constant flow coefficient (a constant and determined at design conditions), \dot{m} is the mass flow rate, *T_{in}* is the turbine inlet temperature, *p_{in}* is the turbine inlet pressure and *p_{out}* is the turbine outlet pressure, η_T is the isentropic efficiency of the turbine at off-design and $\Delta h_{T,is}$ is the isentropic enthalpy difference due to the expansion in the turbine. The former is the Stodola's cone law and determines the mass flow rate of the cycle as a function of inlet pressure, outlet pressure and the fluid density at the turbine inlet. The latter is the relation proposed by Schobeiri (Schobeiri, 2005) to predict the isentropic efficiency at part-load. The efficiency of the generator is calculated as follows (Haglund and Elmegaard, 2009):

$$\eta_{gen} = \frac{load \cdot \eta_{gen,d}}{load \cdot \eta_{gen,d} + (1 - \eta_{gen,d})[(1 - F_{CU}) + F_{CU}load^2]} \quad (4)$$

The efficiency of the pumps at off-design is defined as a function of the volumetric flow rate, according to the relation developed by Veres (Veres, 1994):

$$\frac{\eta_{pump}}{\eta_{pump,d}} = -0.029265 \left(\frac{\dot{V}}{\dot{V}_d} \right)^3 - 0.14086 \left(\frac{\dot{V}}{\dot{V}_d} \right)^3 + 0.3096 \left(\frac{\dot{V}}{\dot{V}_d} \right)^2 + 0.86387 \quad (5)$$

where η_{pump} is the isentropic efficiency of the pump and \dot{V} is the volumetric flow rate.

The pressure drops (Δp) are modelled with a quadratic dependence from the mass flow rate (Lecompte et al., 2013):

$$\Delta p = \Delta p_d \left(\frac{\dot{m}}{\dot{m}_d} \right)^2 \quad (6)$$

where Δp is the pressure drop and \dot{m} is the mass flow rate.

To cope with the different input conditions to the ORC, a simple control scheme was implemented. The pressure was left varying according to a sliding pressure control mode. This is a common control scheme in ORC applications (Imran et al., 2020). The maximum temperature entering the expander is also kept within a maximum threshold while the expander outlet pressure was maintained at the design value.

2.2 Case Study description

The reference case study is biomass-fired micro-generation of heat and power via ORC. The case

study is representative of an existing unit operating at the Czech Technical University (CTU) in Prague for supplying heat (design $120 \text{ kW}_{\text{th}}$) and electricity (design $6.2 \text{ kW}_{\text{el}}$) to the university research center. The working fluid is MM (hexamethyldisiloxane), while the expander is a rotary vane expander. The unit is woodchips-fired.

2.3 Scenario and boundary conditions

Actual operating data has been provided and used as the basis for the analysis. In this study, only the variations of the flue gas flow rate were considered, and the cycle is simulated to provide the maximum power output. This is a simplification as the actual ORC is designed to provide both heat and backup power. However, the goal of this study was to demonstrate the methodology rather than to simulate a real system.

Figure 3 shows the variation of flue gas flow rate (i.e., the heat source for the ORC) over one year. Significant fluctuations can be noticed, resulting in the ORC operating frequently far from its design. Such design was selected as that at 0.078 kg/s of flue gas flow rate and is indicated as 100 % flow rate. Figure 4 shows the distribution of flue gas flow rates throughout one year.

The low load operation will result in the ORC system at the design condition being oversized most of the time. Thus, in this work, the ORC system is optimized to find the optimal system capacity and operating conditions that give the largest power production throughout the year. The capacity of the ORC system is varied by changing the design flue gas flow rate. During this optimization, the excess amount of the flue gas over 115 % of its design value is assumed to be not utilized in the ORC system.

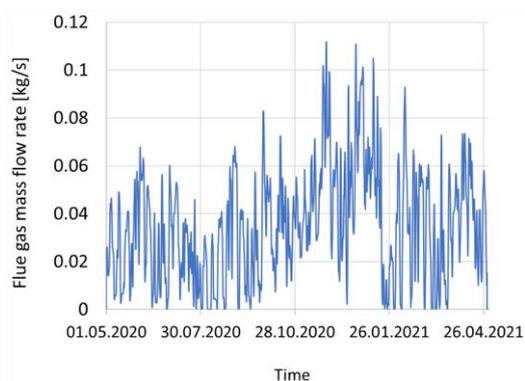


Figure 3. Hourly variation of the heat source (flue gas mass flow rate) over a year period at the Czech Technical University (CTU) campus.

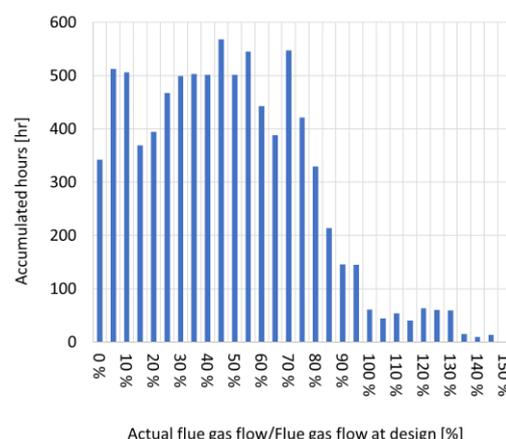


Figure 4. Distribution of the flue gas flow rate at the Czech Technical University (CTU) campus (accumulated hours for one year of operation).

The constraint values in Table 2 can be varied depending on the assumptions made to analyze an ORC system. For example, the practical conditions include limitations to the cycle parameters and operating conditions that are given by the utilization of state-of-the-art equipment.

In this work, however, the ORC is constrained to reflect the actual rig conditions considering the physical limitations of process equipment and working fluids. One example is the cycle is restricted to be operated at sub-critical conditions in order to prevent any mechanical issues. To avoid the thermal degradation of MM, the maximum cycle temperature (expander inlet temperature) is also limited to $190 \text{ }^\circ\text{C}$ considering an extra temperature margin.

Table 3. Design basis and constraints applied in the MM based ORC.

Design parameters and constraints	Rig conditions
$T_{\text{flue, supply}}$	$1442 \text{ }^\circ\text{C}$
$T_{\text{flue, min}}$	$132 \text{ }^\circ\text{C}$
$p_{\text{cycle, max}}$	10 bara
$p_{\text{cycle, min}}$	0.01 bara
$\Delta T_{\text{evap, min}}$	$50 \text{ }^\circ\text{C}$
$\Delta T_{\text{cond, min}}$	$2 \text{ }^\circ\text{C}$
$\Delta T_{\text{sub, min}}$	$9.6 \text{ }^\circ\text{C}$
$\Delta T_{\text{sup, min}}$	$10 \text{ }^\circ\text{C}$
$q_{\text{exp, min}}$	1
$W_{\text{exp, max}}$	15.5 kW
$T_{\text{exp, in, max}}$	$190 \text{ }^\circ\text{C}$
$p_{\text{exp, in, max}}$	8 bara
$p_{\text{exp, out, min}}$	0.2 bara

It is worth noting that the heated cooling water from the condenser is utilized as hot water at the CTU campus. Thus, the cooling water supply (T_{10}) and return temperature (T_{12}) are set to 58 °C and 80 °C to meet the hot water specification. Other design conditions and constraint values are listed in Table 3.

3. Results

Figure 5 presents the performance of two different designs of the MM based ORC system. One is the ORC system optimized for the design flue gas flow rate (referred to as design ORC). Based on the ORC design, the off-design performance is estimated to calculate the accumulated power production per year. The other is the ORC system optimized to maximize the accumulated power production throughout the year while varying the capacity (referred to as optimal ORC). The temperature-enthalpy diagram of the optimal cycle is also presented in Figure 6.

The results indicate that the optimal ORC system has a 5 % larger accumulated power production throughout the year compared to the ORC system at the design. As shown in Figure 5, the design ORC achieves a lower energy efficiency below 80 % load compared to the optimal ORC. Considering the frequent low load operation, such lower efficiency results in smaller accumulated power production as well. Although the ORC at the design outperforms at a larger flue gas flow rate of over 80 %, the fraction of the occurrence is not significant (see Figure 4), thus having a marginal impact on the accumulated power production.

Table 4 introduces the operating conditions of the two different designs. It is worth noting that the optimal ORC has a smaller capacity than the ORC at the design. Based on the flow rate of the working fluid and the expander power output, the capacity of the optimal system is estimated to be around 68 % of the ORC at the design. The lower design capacity of the ORC system allows the cycle to achieve a higher energy efficiency at a lower load where the system is operated most of the time span. The reduced capacity of the ORC will also decrease the capital cost and improve the economic feasibility of the small-scale ORC system, which is one of potential issues regarding the deployment (Tocci *et al.*, 2017).

Although the two different design solutions have a relatively large performance difference, some of the operating conditions at design point are close to identical as indicated in Table 4. The main changes are observed on the size of the heat exchangers (evaporator and condenser). One of the possible

reasons will be the limited feasible region that is caused by the severe constraints considering the rig setup, the low limit of the maximum cycle pressure and temperature due to the working fluid characteristics, and the warm supply and return temperatures of the cooling water.

Relaxation of such constraints and design basis will allow this tool to have a wider search space to identify improved operating conditions considering the off-design performance of the ORC system throughout a year. Applying the tool to other case studies and heat sources might also result into a wider search space.

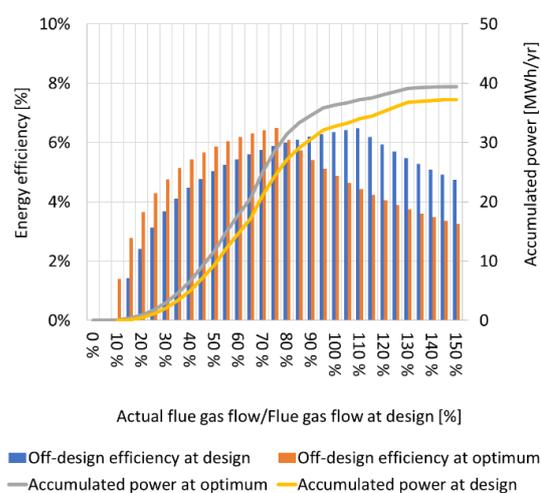


Figure 5. Energy efficiency and the accumulated power output of the MM based ORC system with a hourly variation of the flue gas flow rate for a full year period (Energy efficiency = $W_{net}/\Delta Q_{flue,max}$).

Table 4. Key parameters of the original design and the optimum design accounting for off-design performance (the operating conditions that are similar between the two designs are presented in italic).

Parameter	Unit	Design	Optimum
$W_{accumulated}$	MWh/yr	37.26	39.44
mMM	kg/s	0.37	0.25
mcw	kg/s	1.56	1.04
W_{net}	kW	11.05	7.40
W_{exp}	kW	11.54	7.73
$T_{exp,in}$	°C	<i>190</i>	<i>190</i>
$P_{exp,in}$	bara	<i>6.42</i>	<i>6.42</i>
$T_{exp,out}$	°C	<i>164.51</i>	<i>164.51</i>
$P_{exp,out}$	bara	<i>0.42</i>	<i>0.42</i>
UA_{evap}	kW/C	0.39	0.26
$\Delta T_{min,evap}$	°C	69.16	69.16
UA_{cond}	kW/C	17.59	11.79
$\Delta T_{min,cond}$	°C	2.00	2.00

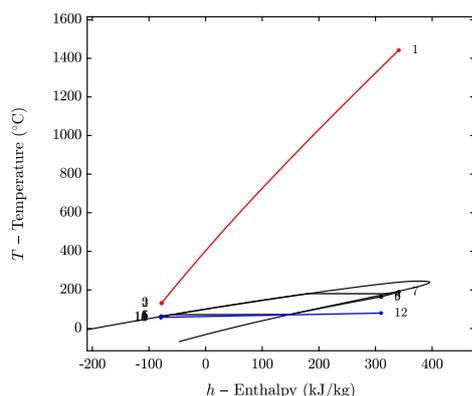


Figure 6. Temperature-enthalpy diagram of the optimal ORC.

4. Summary and Discussions

This work introduces an open-source-based ORC design and optimization tool that can reflect off-design performance. The optimization tool is tested by a case study on an ORC unit operating at the Czech Technical University (CTU) in Prague for supplying heat and electricity to the university research center. Actual heat source variations are considered for the analysis. The results indicate the optimal size of the system is 33 % smaller than the original ORC design, which is optimized without part-load performance estimation. The accumulated power production is also increased with the optimal ORC design by 5 %. Further improvements in the annual accumulated power production with the optimization framework could potentially be achieved with relaxed constraints and design basis, representing practical conditions.

As a next step, the framework will be further tested with different working fluids, heat sources and heat sinks. The tool will be updated with improved process unit and off-design expander models based on experimental data, allowing the framework to be capable of robust and practical multi-scale optimization.

Acknowledgment

This work was supported by EEA / Norway Grants and Technology Agency of the Czech Republic under the KAPPA Program, project no. TO01000160: Optimised expanders for small-scale distributed energy system.

References

Agromayor, R. and Nord, L. O. (2017) 'Fluid selection and thermodynamic optimization of organic Rankine cycles for waste heat recovery applications', in *Energy Procedia*. doi: 10.1016/j.egypro.2017.09.180.
 Colonna, P. *et al.* (2015) 'Organic Rankine Cycle Power Systems: From the Concept to Current Technology,

Applications, and an Outlook to the Future', *Journal of Engineering for Gas Turbines and Power*. doi: 10.1115/1.4029884.

Haglund, F. and Elmegaard, B. (2009) 'Methodologies for predicting the part-load performance of aero-derivative gas turbines', *Energy*, 34(10), pp. 1484–1492.

Imran, M. *et al.* (2020) 'Dynamic modeling and control strategies of organic Rankine cycle systems: Methods and challenges', *Applied Energy*. doi: 10.1016/j.apenergy.2020.115537.

Incropera, F. P. *et al.* (2007) *Fundamentals of Heat and Mass Transfer, Water*. John Wiley & Sons.

Lecompte, S. *et al.* (2013) 'Part load based thermo-economic optimization of the Organic Rankine Cycle (ORC) applied to a combined heat and power (CHP) system', *Applied Energy*, 111, pp. 871–881.

Macchi, E. and Astolfi, M. (2016) *Organic Rankine Cycle (ORC) Power Systems: Technologies and Applications, Organic Rankine Cycle (ORC) Power Systems: Technologies and Applications*.

Orlandini, V. *et al.* (2016) 'Dynamic performance of a novel offshore power system integrated with a wind farm', *Energy*, 109, pp. 236–247.

Petrollese, M. and Cocco, D. (2019) 'Robust optimization for the preliminary design of solar organic Rankine cycle (ORC) systems', *Energy Conversion and Management*. doi: 10.1016/j.enconman.2019.01.060.

Riboldi, L. and Nord, L. (2018) 'Offshore Power Plants Integrating a Wind Farm: Design Optimisation and Techno-Economic Assessment Based on Surrogate Modelling', *Processes*. doi: 10.3390/pr6120249.

Schobeiri, M. (2005) *Turbomachinery flow physics and dynamic performance, Turbomachinery Flow Physics and Dynamic Performance*. Berlin: Springer Berlin.

Tocci, L. *et al.* (2017) 'Small scale Organic Rankine Cycle (ORC): A techno-economic review', *Energies*. doi: 10.3390/en10040413.

Veres, J. P. (1994) 'Centrifugal and axial pump design and off-design performance prediction', *NASA Technical Memorandum 106745*, pp. 1–24.

Effect of temperatures on anaerobic granulated biofilm modelling

Anissa Sukma Safitri^a and Roald Kommedal^a

^a *Institute of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, 4036 Stavanger, Norway*
Email: anissa.s.safitri@uis.no and roald.kommedal@uis.no

Abstract

Anaerobic granulated biomass-based treatment is a sustainable alternative for municipal wastewater treatment. Each granule in the system is comprised of a complex community of anaerobic microorganisms embedded in a biofilm matrix. The aim of this work was to implement a biofilm model for simulation of biogas production and COD removal as observed in an experimental up-flow anaerobic sludge blanket (UASB) reactor system. Additionally, selected scenario simulations were carried out to assess the effect of temperatures (25, 16, and 12 °C) on granulated anaerobic reactor performance at different organic loading rates. The two main model components used are: Dynamic biochemical and physicochemical conversion processes (Anaerobic Digestion Model No. 1) and diffusive mass transfer within the granule (biofilm). The model was implemented in AQUASIM 2.1. Simulations gave insight into non-observables, especially intragranular biomass distribution and substrate profiles, which help our understanding of granule formation and evolution. Results reflected observed effluent COD concentrations and methane production rates at variable temperatures and reactor loadings. Simulations also confirmed observed steady-state reductions in COD removal efficiencies and methane fraction in biogas at increasing organic loading rate. Model simulations also showed intra-granular alkaline pH depth profiles with increasing organic loading rate which may explain calcium-based mineral core formation. The biomass composition and active regions in granules were not significantly affected by organic loading rate. At steady state, organic substrates especially monosaccharides and volatile fatty acids were predicted to degrade approximately within the outer 100 µm. In general, the model can be used as a tool to predict and simulate anaerobic granulated biofilm system performances in UASB reactor.

Keywords: Anaerobic granule; Biofilm modelling; Municipal wastewater

1. Introduction

Mathematical bioprocess modelling is a recognized tool for fundamental bioprocess understanding, data analysis and hypothesis testing, for design and optimization of wastewater treatment processes (Henze et al., 2008). Considerable efforts has been dedicated for the development of mathematical models for anaerobic granular technology, recently summarized by Baeten et al. (2019). Two main approaches are used for anaerobic granular modelling: The intragranular transport models (biofilm models) and the suspended biomass liquid phase models using apparent kinetics. Granular biofilm models are used for any redox system, while the apparent kinetic models are more commonly used for anaerobic systems (Baeten et al., 2019). In biofilm reactors, substrate transport from the bulk liquid into the microbial matrix is normally controlled by diffusion. Wanner & Gujer (1986) identified several beneficial objectives attainable by using biofilm models: Understanding the

mechanisms fundamental to how a biofilm forms or performs; integration of different mechanisms occurring at different spatial and temporal scales; pre-model the system to generate expected results; and evaluating novel process designs.

The Anaerobic Digestion Model no. 1 (ADM1) is recognized as the standard model for conversion of organic substrates to biogas in open wastewater and biosludge systems (Batstone et al., 2002). Bioconversion, physicochemical reactions, and interface mass transfer are combined into a model comprising the most important state variables and processes. In addition, bioconversion parameters under typical anaerobic meso- to thermophilic conditions are suggested.

Biofilm models of anaerobic granules under mass transfer limitations and single as well as multiple limiting substrates are available in the literatures (Batstone et al., 2004; Buffière et al., 1995; Doloman et al., 2020; Feldman et al., 2017; Flora et

al., 1995; Odriozola et al., 2016; Sun et al., 2016). Granulated biofilm models have been implemented on several platforms, such as AQUASIM (Batstone et al., 2004) and MATLAB (Odriozola et al., 2016). Extension of the strictly suspended/homogenous ADM1 into a biofilm setting require diffusive transport for all dissolved components, attachment/detachment mechanisms for particulates and possibly moderation of kinetic and/or stoichiometric coefficients for the growth-related conversions. As for the latter, the work of Bakke et al. (1984) suggest that conversion kinetics are not significantly moderated by a biofilm phenotypic growth state, but possibly the stoichiometry due to EPS formation and degradation (Kommedal et al., 2001). Physio-chemical mass transfer parameters are available in the physical chemical reference literature.

The aim of this work was to propose a biofilm model relevant for an anaerobic granule typical of an up-flow anaerobic sludge blanket (UASB) reactor system. The proposed model was evaluated by simulation of experimental observations in a laboratory scale UASB reactor receiving strong municipal wastewater. Additionally, analyses were carried out to assess the effect of temperatures (25, 16, and 12 °C) on granulated anaerobic reactor performance at different organic loadings and compare these to experimental results.

2. Methodology

2.1. Experimental set-up and results

A long-term operation of up-flow anaerobic sludge blanket (UASB) system treating real municipal wastewater from IVAR IKS, Norway, at decreasing temperatures (25, 16, 12, 8.5, 5.5, and 2.5 °C) and variable organic loading rates (OLR) from 1.0 $\text{gCOD}\cdot\text{l}^{-1}\cdot\text{d}^{-1}$ up to 15.2 $\text{gCOD}\cdot\text{l}^{-1}\cdot\text{d}^{-1}$ was investigated over 1025 days (Figure 1). Experiments were performed in two parallel in-house designed laboratory-scale UASB reactors, which were operated continuously with hydraulic retention time (HRT) of 16.7 h down to 1.1 h. The wastewater may be characterized as a municipal wastewater with significant contributions from agricultural and food industries. The dissolved COD concentrations of inlet wastewater during UASB reactor operation fluctuated in the range 439 - 1473 $\text{mgCOD}_{\text{dissolved}}\cdot\text{l}^{-1}$ with the mean concentration being 741 ± 7 $\text{mgCOD}_{\text{dissolved}}\cdot\text{l}^{-1}$ (\pm standard error).

Stable COD removal efficiencies of 50 - 70 % were achieved at 25 °C down to 8.5 °C with loading up to approximately 15.2 $\text{gCOD}\cdot\text{l}^{-1}\cdot\text{d}^{-1}$. COD removal efficiencies were reduced at temperatures below 8.5 °C, but significant methane formation was observed even at 2.5 °C at reduced loading (up to 5 $\text{gCOD}\cdot\text{l}^{-1}\cdot\text{d}^{-1}$).

More than 90% of COD removed was converted to methane. The overall COD balance closed at above 90% of the inlet COD at all operating temperatures and OLRs. Temperature affected the reactor performances, microbial community structure, and the degradation pathway of organic matter with acetoclastic methanogen and methylotrophic played significant roles (Safitri et al., 2022).

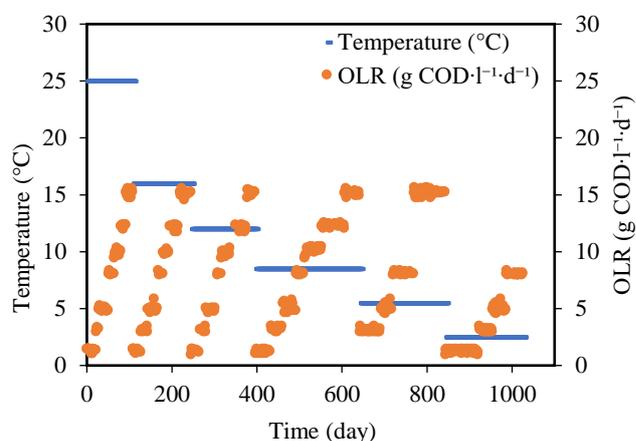


Figure 1 The UASB reactors were operated continuously over 1025 days by the stepwise increase of OLR at decreasing temperatures.

2.2. Model implementation

The biomass in UASB systems is in the form of compact granules that contain a complex community of microorganisms embedded in the extracellular polymeric substances (EPS) matrix, i.e., biofilm. This presented model used to predict the behavior of a granule biofilm representative of the UASB reactor described in section 2.1. For simplification, scenario analyses were carried out to assess the effect of temperatures (25, 16, and 12 °C) on reactor performances, as shown in Figure 2.

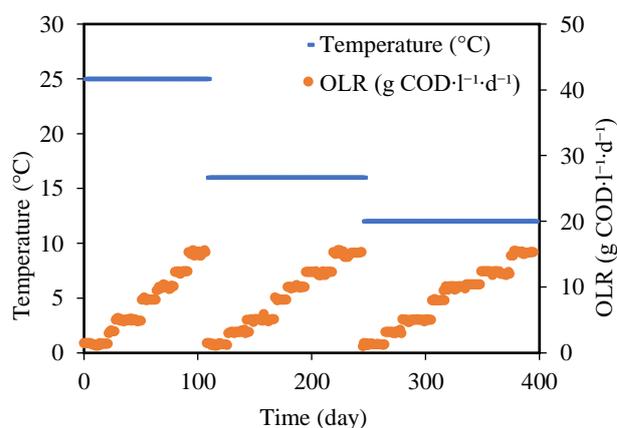


Figure 2 Simulation scenario in biofilm modelling by the stepwise increase of OLR at decreasing temperatures.

The two main model structures implemented are: Biochemical and physico-chemical conversion

processes as described in the Anaerobic Digestion Model No. 1 (ADM1) by Batstone et al. (2002) and diffusive mass transfer model for granulated biomass (the biofilm compartment) based on Wanner and Gujer (1986). The biochemical conversions included: a. Disintegration of particulates to biopolymers (polysaccharides, proteins, and lipids); b. hydrolysis of biopolymers to sugars, amino acids, and long-chain fatty acids (LCFA); c. acidogenesis from sugars and amino acids to volatile fatty acids (VFAs) and hydrogen; d. acetogenesis of LCFA and VFAs to acetate; and e. separate methanogenesis steps from acetate and hydrogen/CO₂ (Hulshoff Pol et al., 2004). The physico-chemical equations describe ion association and dissociation, and gas-liquid mass transfer. Inhibition kinetics have been integrated in relevant biochemical process.

The diffusion limited biogeochemical model was implemented in AQUASIM 2.1 (Reichert, 1994). The ADM1 conversions were implemented in the biofilm compartment using 20 vertical grid points evenly distributed over the biofilm depth. The bulk liquid is modelled as a mixed reactor, with liquid borne in- and out-fluxes and bulk liquid biochemical reactions equal to the biofilm matrix reactions. An additional mixed compartment was implemented to represent the gas phase, connected by a diffusive link to simulate the gaseous transfer of methane, carbon dioxide, and hydrogen. Figure 3 presents the schematic of the biofilm compartment implementation.

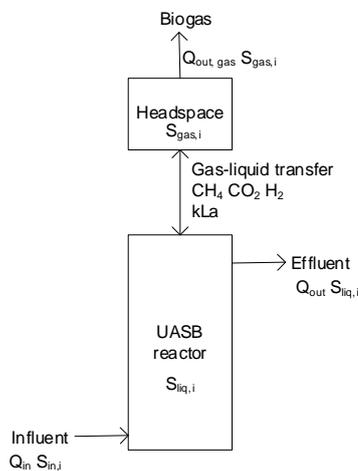


Figure 3 Schematic representation of anaerobic granulated biofilm implementation into ADM1

Acid-base reactions in the ADM1 model can either be implemented as a combination of differential and algebraic sets of equations (DAE) or by fast time dependent differential equations (DE). The standard (commonly available) ADM1 simulators implemented for CSTR (using AQUASIM 2.1) use

the DAE approach. However, herein the acid-base biofilm model is implemented by solving individual acids and conjugated bases separately, as dynamic state variables. All ionic species were implemented as differential variables and a pH model construction followed a step-by-step procedure based on Hofmann et al. (2008). The following Equation 1 for calculating hydrogen ion concentration was used:

$$S_{H^+} = -\frac{\theta}{2} + \frac{1}{2}\sqrt{\theta^2 + 4K_w} \quad (\text{Equation 1})$$

where θ is the net charge in the system resulting from all acid-bases considered in the model:

$$\theta = S_{CAT} + S_{NH_4^+} + S_{H^+} - \frac{S_{AC^-}}{64} - \frac{S_{PRO^-}}{112} - \frac{S_{BU^-}}{160} - \frac{S_{VA^-}}{208} - S_{HCO_3^-} - \frac{K_w}{S_{H^+}} + -S_{AN}$$

As stated by the Stokes-Einstein equation, the diffusion coefficient in water (D_{aq}) depends on temperature, both directly and through the effect of temperature (T) on the solution viscosity (μ) (Cussler, 1984). This temperature dependence of aqueous diffusion coefficients can be calculated through the relationship in Equation 2.

$$\frac{D_{aq} \mu}{T} = constant \quad (\text{Equation 2})$$

The value of the effective diffusion coefficient in the biofilm ($D_{aq,e}$) will be reduced compared to the diffusion coefficient in water due to the presence of microbial cells, extracellular polymers, and abiotic particles or gas bubbles that are trapped in the biofilm (Stewart, 2003). This reduction is described by the ratio $D_{aq,e}/D_{aq}$ (Equation 3). The diffusion coefficients in water (D_{aq}) at 25 °C were taken from Cussler (1984) and Stewart (1998). Effective diffusivities ($D_{aq,e}$) used in current biofilm model (Table 1) were calculated using Equation 2 and 3 with correction to 16 and 12 °C. The solution viscosity (μ) of water at different temperatures were adapted from Coulson & Richardson (1999).

Table 1 Effective diffusivities used in current biofilm model

Variables	$D_{aq,e}/D_{aq}^a$	$D_{aq,e}$ at different temperature ($\times 10^{-5} \text{ m}^2 \cdot \text{d}^{-1}$)		
		25 °C	16 °C	12 °C
Effective diffusivity:				
Amino acids	0.40	2.63	2.05	1.81
Acetate	0.21	2.20	1.71	1.51
Butyrate	0.34	2.56	1.99	1.76
Methane	0.40	5.15	4.02	3.55
CO ₂	0.40	6.64	5.18	4.58
LCFA	0.20	1.62	1.27	1.12
Hydrogen	0.60	23.33	18.20	16.10
Cations/anions	0.49	5.00	3.90	3.45
Ammonia	0.75	10.63	8.29	7.33
Propionate	0.30	2.75	2.14	1.90
Soluble inerts	0.10	0.86	0.67	0.59
Sugar/glucose	0.30	1.74	1.35	1.20
Valerate	0.30	0.50	0.39	0.35

^aReferences: Cussler (1984) and Stewart (1998)

Modified kinetics of the particulate first order disintegration constant (k_{dis}), first order hydrolysis constant (k_{hyd}) and maximum uptake rates (k_m) at different temperatures were retrieved and estimated from literatures (Bergland et al., 2015; Donoso-Bravo et al., 2009; Lohani et al., 2018; Rebac et al., 1995). The temperature compensated kinetic values at 25, 16 and 12 °C were presented in Table 2, taking 35 °C as the reference condition (Batstone et al., 2004).

Table 2 Kinetic parameters k_{dis} , k_{hyd} , k_m with temperature

Process	Temperature (°C)			
	35 ^a	25 ^b	16 ^b	12 ^b
Disintegration (k_{dis} , d ⁻¹)	0.5	0.24	0.07	0.06
First order hydrolysis (k_{hyd} , d ⁻¹):				
Carbohydrate	106	51	15	13
Protein	2.7	1.3	0.4	0.3
Lipid	0.40	0.19	0.06	0.05
Maximum uptake rates (k_m , d ⁻¹):				
Sugars	150	32	24	24
Amino acids	250	53	40	40
Fatty acids	30	6.3	4.8	4.8
Butyrate	100	67	36	23
Propionate	65.0	45.5	18.9	10.4
Acetoclastic methanogens	40.0	27.6	8.0	6.0
Hydrogenotrophic methanogens	175	120.8	35.0	26.3

References:

^aRetrieved from Batstone et al. (2004)

^bEstimated from literatures (Bergland et al., 2015; Donoso-Bravo et al., 2009; Lohani et al., 2018; Rebac et al., 1995)

Physico-chemical equilibria are modeled based on the law of mass action for aqueous substances and by Henry's law for the solubility of a gas. Table 3 presents temperature dependent physico-chemical processes parameters used in biofilm modelling. Detailed calculations and model source file are available on request from the authors.

Table 3 Temperature dependent physico-chemical processes parameters

Description	Temperature compensation	References
Acidity constants (K_a):		
CO ₂	$10^{-6.35} \exp\left(\frac{7646}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Lide (2003)
H ₂ O	$10^{-13.995} \exp\left(\frac{55900}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Lide (2003)
NH ₄ ⁺	$10^{-9.25} \exp\left(\frac{51965}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Lide (2003)
Henry's law constants (K_H):		
CH ₄	$0.00140 \cdot R \cdot T \cdot \exp\left(\frac{-14240}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Batstone et al. (2002)
CO ₂	$0.03400 \cdot R \cdot T \cdot \exp\left(\frac{-7646}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Batstone et al. (2002)
H ₂	$0.00078 \cdot R \cdot T \cdot \exp\left(\frac{-4180}{100 \cdot R} \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Lide (2003)
Pressure of water	$0.0313 \cdot \exp\left(5290 \cdot \left(\frac{1}{T_{std}} - \frac{1}{T}\right)\right)$	Rosen and Jeppsson (2006)

2.3. Simulation set-up

The simulated reactor model consisted of a liquid (0.8 l) and a gas phase (0.2 l). The biofilm compartment corresponded to experimental data from this work (Safitri et al., 2022) and consisted of approximately 300 ml of granular sludge of an average uniform spherical granules with a diameter of 2 mm. The number of granules ($n_{sp} \approx 21500$) was calculated based on the total granule volume. Granules were assumed to have no dispersive solid transport (rigid biofilm matrix) and no suspended solids within biofilm matrix pores (pore volume contains only liquid phase). External mass transfer limitation was for simplicity reasons neglected (no diffusion limitations in the stagnant surface layer). Based on the above assumptions, this gave a total reactor biofilm surface area of 1.08 m². Detachment of biomass (Equation 4) is based on the non-linear biofilm thickness dependency proposed by Stewart et al. (1996):

$$r_{det} = k_{det} \cdot L_f^2 \quad (\text{Equation 4})$$

where k_{det} is an empirical detachment coefficient, (here: 0.024 kg·m⁻²·d⁻¹), and L_f is the simulated biofilm thickness (m). Biomass density within the granules was set at 180 kgCOD·m⁻³, a typical anaerobic granules density based on Batstone and Keller (2001).

Initial conditions were defined for the biofilm matrix and the bulk liquid volume as follows: All modelled microorganisms were considered to have equal initial bulk phase concentrations of 0.05% v/v (9 kgCOD·m⁻³) equal to the biomass fractions in the biofilm matrix. The initial biofilm thickness was set at 0.03 mm and bulk phase initial biomass, initial VFAs and pH was chosen among typical observed values at 10⁻⁵ kgCOD·m⁻³, 10⁻⁶ kgCOD·m⁻³ and 10⁻⁷ kmol·m⁻³, respectively. Approximate steady-state pore and bulk liquid concentrations were used as initial state variables. Simulation time was limited to 400 days, with a time resolution of 0.1 d and a numerical maximum time step limited to 4000 d.

2.4. Input characteristics

The COD influent to the UASB reactor is defined in the model as presented in Table 4 and were assumed to be primarily polysaccharides, proteins, and fats, taking into consideration that IVAR Grødalund wastewater treatment plant receives wastewater from food, animal, and dairy industries. A feed bicarbonate alkalinity of 0.01 M and inorganic nitrogen of 0.007 M were used which were in line with analysis of the wastewater applied in the UASB laboratory experiment.

Table 4 COD input used for simulations

Description	Fraction	Value	Unit
Amino acid	0.002	0.003	kg COD·m ⁻³
Acetic acid	0.08	0.1	kg COD·m ⁻³
Butyrate	0.08	0.1	kg COD·m ⁻³
Propionate	0.08	0.1	kg COD·m ⁻³
Valerate	0.08	0.1	kg COD·m ⁻³
LCFA	0.14	0.175	kg COD·m ⁻³
Sugar	0.1	0.075	kg COD·m ⁻³
Soluble inert	0.07	0.088	kg COD·m ⁻³
Carbohydrate	0.1	0.125	kg COD·m ⁻³
Protein	0.1	0.15	kg COD·m ⁻³
Lipid	0.1	0.15	kg COD·m ⁻³
Particulate inert	0.06	0.075	kg COD·m ⁻³
Composite/complex	0.005	0.006	kg COD·m ⁻³
Input total COD	1.00	1.26	kg COD·m ⁻³

3. Results and Discussions

Figure 4 presents simulated and observed COD effluent concentrations of the UASB reactor at 25, 16 and 12 °C. Simulated dissolved COD effluent concentration was close to that observed results throughout the whole test period, even though there were slightly differences especially at 12 °C. This might be interpreted as a to strong temperature compensation in the ADM1 model, or a change/adaptation in community structure towards low mesophilic bacteria.

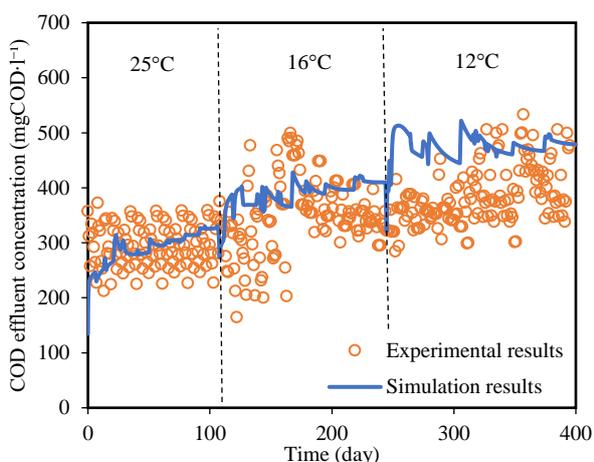


Figure 4 Simulated (blue dots) and measured (orange dots) COD effluent concentration profile in UASB reactor

Both simulation and experimental results show decreasing dissolved COD removal efficiencies at steady-state conditions with decreasing temperatures, from approximately 65 - 70% at 25 °C to around 57 - 66% at 12 °C. The relatively small fraction of particles degraded in our simulations (10 - 24% removal efficiency) indicate hydrolysis to be rate-limiting.

In simulations, methane fractions in the biogas at steady-state conditions decreased with increasing OLR with the range of 80 - 90% (data not shown), a trend that could not be significantly observed in

experimental data (Safitri et al., 2022). The dissolved methane concentrations in the effluent (86, 110, and 117 mgCOD·l⁻¹) were mimicked by the simulation result and are in line with theoretical value of dissolved methane at 25, 16, and 12 °C, respectively (Liu et al., 2014).

Figure 5 shows simulated and observed methane production profile in the UASB reactor at different temperatures and loadings. At each temperature, methane production increased with the increasing OLR, proportional to the amount of organic matter removed in the UASB reactors. Methane production at 12 °C was comparable in all OLRs to 16 °C, indicating that the reduction in operating temperature to 12 °C did not negatively affect methane production. This implies that the biomass has compensated for the temperature reduction by adaptation, and the model is overcompensating for the temperature reduction as both methane production and effluent COD reduction is larger than simulated, we assume an early kinetics to be overcompensated, like a to low hydrolysis coefficient.

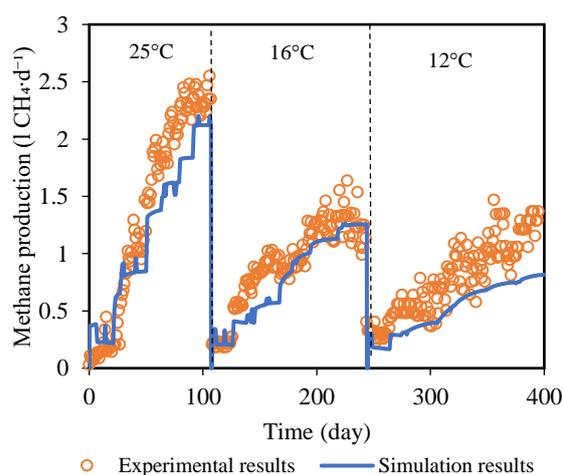


Figure 5 Simulated (blue dots) and measured (orange dots) methane production profile in UASB reactor

Considerable pH profile variations (approximately 7 - 8.3) through the depth of the granule were simulated as shown in Figure 6. The interior increase in pH inside the granules suggest calcium precipitation in the granule core. Amorphous CaPO₄, CaCO₃ and apatite precipitation (Ca₅(PO₄)₃OH) is known to stimulate granule formation, and could be exploited for phosphorous recovery (Cunha et al., 2018). Furthermore, simulated pH range in the bulk phase was about 7 which was comparable to observed pH (data not shown).

Simulated composition of granular sludge active biomass fractions for the experimental UASB

reactor is presented in Figure 7. Based on these results, methane production was mainly performed by acetoclastic methanogens at 25 °C, which dominated compared to hydrogenotrophic methanogens for all selected temperatures. However, hydrogenotrophic methanogens appeared more dominant at decreasing temperature at 12 °C.

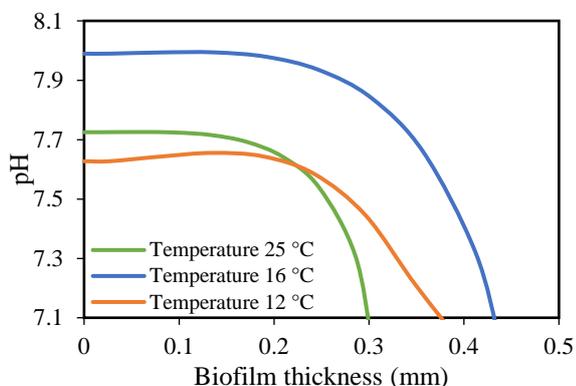


Figure 6 Simulated pH distribution profile along the granule in UASB reactor during steady-state conditions at different temperatures and OLR 15 gCOD·l⁻¹·d⁻¹.

Based on the experimental study, temperature affected the microbial community structure and the degradation pathway of organic matter with acetoclastic and methylotrophic methanogens played significant roles (Safitri et al., 2022). The original ADM1 model did not include methylotrophic methanogens which reduces the methyl-groups of methylated compounds to methane with H₂ as electron donor (Söllinger & Urich, 2019). Therefore, the simulated values are only to be qualitatively interpreted as they are not structurally comparable to experimental community analysis results.

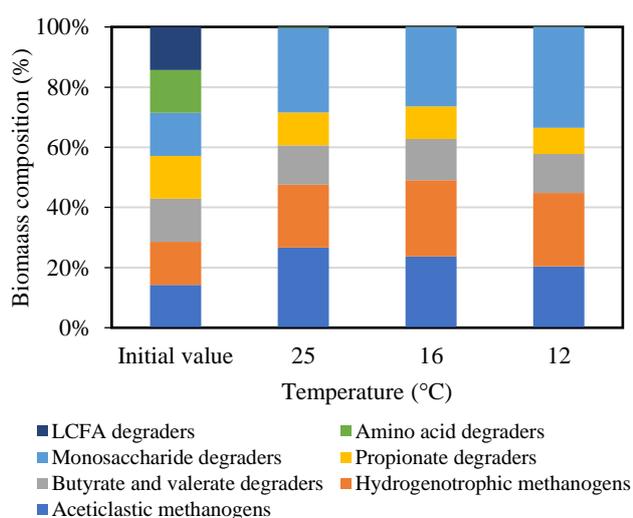


Figure 7 Simulated active biomass composition of the granular sludge of UASB reactor during steady-state conditions at different temperatures and OLR 15 gCOD·l⁻¹·d⁻¹.

According to simulation results, LCFA degraders only accounted for trace quantities of the active biomass. This could be explained by too low LCFA concentration to sustain biomass. There are significant decreases of amino acid degraders. However, more than 90% of the influent amino acid and LCFA was converted in the reactor at all temperatures. There is no significant temperature effect on bacterial distribution profile along the granule in UASB reactor during steady-state conditions. The sugar degraders had the highest concentration on the outer layer of granular sludge followed by butyrate and valerate degraders, hydrogenotrophic and acetoclastic methanogens.

The high amount of carbohydrates in the wastewater, supported these bacterial groups and resulted in a high methane concentration in the produced biogas. In the granules, the acetate degrading biomass peaked approximately 100 μm behind the biofilm-bulk boundary. Acetate concentrations were at their maximum at the biofilm boundary. The delayed front was possibly due to the faster growth of the other organisms, and a consequent high availability (not concentration) of acetate. Monosaccharide and VFA substrates are predicted to degrade approximately within the outer 100 - 200 μm. The granules had generally lower intermediate substrate concentrations than the bulk phase, indicating no net diffusive force out of the biofilm and intermediate reactions are not overloaded.

4. Conclusions

The proposed ADM1-biofilm model reflected the key effluent/bulk phase state variables as observed in the experimental UASB reactor system fed municipal wastewater. Furthermore, the model reflected the effect of reduced temperatures on overall COD conversion and biogas reactor performances at variable organic loading rates. The temperature compensation model used herein is too strong at 12 °C, maybe due to biomass adaptations. Available data are insufficient to validate the intra-granular state variables, but simulation results indicate ideal acid-base chemistry for inorganic granule core formation/growth.

Acknowledgment

The authors gratefully acknowledge the Norwegian Ministry of Education and Research through the university grant program (Norwegian Department of Education) and by the Foundation Stiftelsen Signe-Marie (<https://www.stiftelsensignemarie.no/>), for their support of this research project. We would like to acknowledge the wastewater treatment plant,

IVAR IKS, Norway, for practical support, wastewater logistics and operational advice.

References

- Baeten, J. E., Batstone, D. J., Schraa, O. J., van Loosdrecht, M. C. M., & Volcke, E. I. P. (2019). Modelling anaerobic, aerobic and partial nitrification-anammox granular sludge reactors—A review. *Water Research*, *149*, 322–341. <https://doi.org/10.1016/j.watres.2018.11.026>
- Bakke, R., Trulear, M. G., Robinson, J. A., & Characklis, W. G. (1984). Activity of *Pseudomonas aeruginosa* in biofilms: Steady state. *Biotechnology and Bioengineering*, *26*(12), 1418–1424. <https://doi.org/10.1002/bit.260261204>
- Batstone, D. J., & Keller, J. (2001). Variation of bulk properties of anaerobic granules with wastewater type. *Water Research*, *35*(7), 1723–1729. [https://doi.org/10.1016/S0043-1354\(00\)00446-2](https://doi.org/10.1016/S0043-1354(00)00446-2)
- Batstone, D. J., Keller, J., Angelidaki, I., Kalyuzhnyi, S. V., Pavlostathis, S. G., Rozzi, A., Sanders, W. T. M., Siegrist, H., & Vavilin, V. A. (2002). The IWA Anaerobic Digestion Model No 1 (ADM1). *Water Science and Technology*, *45*(10), 65–73. <https://doi.org/10.2166/wst.2002.0292>
- Batstone, D. J., Keller, J., & Blackall, L. L. (2004). The influence of substrate kinetics on the microbial community structure in granular anaerobic biomass. *Water Research*, *38*(6), 1390–1404. <https://doi.org/10.1016/j.watres.2003.12.003>
- Bergland, W. H., Dinamarca, C., & Bakke, R. (2015, October 7). *Temperature Effects in Anaerobic Digestion Modeling*. 56th Conference on Simulation and Modelling (SIMS 56), Linköping University, Sweden. <http://dx.doi.org/10.3384/ecp15119261>
- Buffière, P., Steyer, J.-P., Fonade, C., & Moletta, R. (1995). Comprehensive modeling of methanogenic biofilms in fluidized bed systems: Mass transfer limitations and multisubstrate aspects. *Biotechnology and Bioengineering*, *48*(6), 725–736. <https://doi.org/10.1002/bit.260480622>
- Coulson, J. M., & Richardson, J. F. (1999). *Chemical Engineering: Fluid flow, heat transfer and mass transfer* (Vol. 1). Pergamon Press.
- Cunha, J. R., Tervahauta, T., van der Weijden, R. D., Temmink, H., Hernández Leal, L., Zeeman, G., & Buisman, C. J. N. (2018). The Effect of Bioinduced Increased pH on the Enrichment of Calcium Phosphate in Granules during Anaerobic Treatment of Black Water. *Environmental Science & Technology*, *52*(22), 13144–13154. <https://doi.org/10.1021/acs.est.8b03502>
- Cussler, E. L. (1984). *Diffusion, mass transfer in fluid systems*. Cambridge University Press.
- Doloman, A., Mahajan, A., Pererva, Y., Flann, N. S., & Miller, C. D. (2020). A Model for Bioaugmented Anaerobic Granulation. *Frontiers in Microbiology*, *11*, 2239. <https://doi.org/10.3389/fmicb.2020.56682>
- Donoso-Bravo, A., Retamal, C., Carballa, M., Ruiz-Filippi, G., & Chamy, R. (2009). Influence of temperature on the hydrolysis, acidogenesis and methanogenesis in mesophilic anaerobic digestion: Parameter identification and modeling application. *Water Science and Technology*, *60*(1), 9–17. <https://doi.org/10.2166/wst.2009.316>
- Feldman, H., Flores-Alsina, X., Ramin, P., Kjellberg, K., Jeppsson, U., Batstone, D. J., & Gernaey, K. V. (2017). Modelling an industrial anaerobic granular reactor using a multi-scale approach. *Water Research*, *126*, 488–500. <https://doi.org/10.1016/j.watres.2017.09.033>
- Flora, J. R. V., Suidan, M. T., Biswas, P., & Sayles, G. D. (1995). A modeling study of anaerobic biofilm systems: I. Detailed biofilm modeling. *Biotechnology and Bioengineering*, *46*(1), 43–53. <https://doi.org/10.1002/bit.260460107>
- Henze, M., van Loosdrecht, M. C. M., Ekama, G. A., & Brdjanovic, D. (2008). *Biological Wastewater Treatment: Principles, Modelling and Design*. IWA Publishing. <https://doi.org/10.2166/9781780401867>
- Hofmann, A. F., Meysman, F. J. R., Soetaert, K., & Middelburg, J. J. (2008). A step-by-step procedure for pH model construction in aquatic systems. *Biogeosciences*, *5*(1), 227–251. <https://doi.org/10.5194/bg-5-227-2008>
- Hulshoff Pol, L. W., de Castro Lopes, S. I., Lettinga, G., & Lens, P. N. L. (2004). Anaerobic sludge granulation. *Water Research*, *38*(6), 1376–1389. <https://doi.org/10.1016/j.watres.2003.12.002>
- Kommedal, R., Bakke, R., Dockery, J., & Stoodley, P. (2001). Modelling production of extracellular polymeric substances in a

- Pseudomonas aeruginosa* chemostat culture. *Water Science and Technology*, 43(6), 129–134. <https://doi.org/10.2166/wst.2001.0357>
- Liu, Z., Yin, H., Dang, Z., & Liu, Y. (2014). Dissolved methane: A hurdle for anaerobic treatment of municipal wastewater. *Environmental Science & Technology*, 48(2), 889–890. <https://doi.org/10.1021/es405553j>
- Lohani, S. P., Wang, S., Bergland, W. H., Khanal, S. N., & Bakke, R. (2018). Modeling temperature effects in anaerobic digestion of domestic wastewater. *Water-Energy Nexus*, 1(1), 56–60. <https://doi.org/10.1016/j.wen.2018.07.001>
- Odrizola, M., López, I., & Borzacconi, L. (2016). Modeling granule development and reactor performance on anaerobic granular sludge reactors. *Journal of Environmental Chemical Engineering*, 4(2), 1615–1628. <https://doi.org/10.1016/j.jece.2016.01.040>
- Rebac, S., Ruskova, J., Gerbens, S., van Lier, J. B., Stams, A. J. M., & Lettinga, G. (1995). High-rate anaerobic treatment of wastewater under psychrophilic conditions. *Journal of Fermentation and Bioengineering*, 80(5), 499–506. [https://doi.org/10.1016/0922-338X\(96\)80926-3](https://doi.org/10.1016/0922-338X(96)80926-3)
- Reichert, P. (1994). AQUASIM – A Tool for Simulation and Data Analysis of Aquatic Systems. *Water Science and Technology*, 30(2), 21–30. <https://doi.org/10.2166/wst.1994.0025>
- Safitri, A. S., Michelle Kaster, K., & Kommedal, R. (2022). Effect of low temperature and municipal wastewater organic loading on anaerobic granule reactor performance. *Bioresource Technology*, 360, 127616. <https://doi.org/10.1016/j.biortech.2022.127616>
- Söllinger, A., & Urich, T. (2019). Methylotrophic methanogens everywhere—Physiology and ecology of novel players in global methane cycling. *Biochemical Society Transactions*, 47(6), 1895–1907. <https://doi.org/10.1042/BST20180565>
- Stewart, P. S. (1998). A review of experimental measurements of effective diffusive permeabilities and effective diffusion coefficients in biofilms. *Biotechnology and Bioengineering*, 59(3), 261–272. [https://doi.org/10.1002/\(SICI\)1097-0290\(19980805\)59:3<261::AID-BIT1>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-0290(19980805)59:3<261::AID-BIT1>3.0.CO;2-9)
- Stewart, P. S., Hamilton, M. A., Goldstein, B. R., & Schneider, B. T. (1996). Modeling biocide action against biofilms. *Biotechnology and Bioengineering*, 49(4), 445–455. [https://doi.org/10.1002/\(SICI\)1097-0290\(19960220\)49:4<445::AID-BIT12>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1097-0290(19960220)49:4<445::AID-BIT12>3.0.CO;2-9)
- Stewart Philip S. (2003). Diffusion in Biofilms. *Journal of Bacteriology*, 185(5), 1485–1491. <https://doi.org/10.1128/JB.185.5.1485-1491.2003>
- Sun, J., Dai, X., Wang, Q., Pan, Y., & Ni, B.-J. (2016). Modelling Methane Production and Sulfate Reduction in Anaerobic Granular Sludge Reactor with Ethanol as Electron Donor. *Scientific Reports*, 6(1), 35312. <https://doi.org/10.1038/srep35312>
- Wanner, O., & Gujer, W. (1986). A multispecies biofilm model. *Biotechnology and Bioengineering*, 28(3), 314–328. <https://doi.org/10.1002/bit.260280304>

Part load performance of PEM fuel cell and electrolyser stacks in hybrid energy system for offshore application

Marcin Pilarczyk^{a,*}, Luca Riboldi^a, Lars O. Nord^{a,**}

^a *Norwegian University of Science and Technology, Department of Energy and Process Engineering, Norway*

*marcin.pilarczyk@ntnu.no; **lars.nord@ntnu.no

Abstract

Global warming and associated climate change are ongoing processes worldwide, behind which anthropogenic greenhouse gas emissions play the main role. Renewable energy sources integrated in hybrid energy systems (HES) are fundamental solutions to meet energy demands in a sustainable manner. Proton exchange membrane (PEM) electrolysers and fuel cell stacks can be used as integral components of HES in several applications, for instance supplying energy offshore. Due to the integration with irregular renewable energy sources and the variability of energy demands, those stacks will be frequently operated at part-load conditions. The novelty of this work lies in the incorporation of part-load performance in the models of high-capacity fuel cell and electrolysers stacks. A zero-dimensional approach for steady-state behavior was applied to calculate the polarisation and performance curves of the system. The determined curves were implemented in an already developed online tool for analyzing HES. The computational results from the tool show a great ability of the PEM systems for decreasing the carbon intensity of an offshore facility and increase the wind energy integration within the HES.

1. Introduction

Throughout last decades, global warming has been observed and will continue to cause further long-term changes in the climate system [1]. This fact poses a cross-generational challenge for decreasing the greenhouse gas emissions (GHG) worldwide. Therefore, various research frontiers encompasses novel technologies and energy carriers to limit GHG or even completely decarbonize different industrial branches. One of the perspective energy carriers is hydrogen, which is considered as a key instrument for reaching net-zero greenhouse gas emission by 2050 by European continent according to European Union's Green Deal strategy [2]. Hydrogen-based technologies comprise various technologies for heat and power production and energy storage [3]. One of the emerging industrial applications are hydrogen fuel cells (FC) and electrolysers (ELY) stacks incorporated in hybrid energy systems (HES). Many of the proposed hydrogen-based HES systems are reaching commercial maturity [4, 5] while other novel layouts are at earlier stage of commercialization [6]. Whenever energy intensity sectors need to pursue ambitious climate policies, like offshore sector in Norway, the hydrogen-based solutions tied to renewables has the potential to be an effective solution for decarbonizing all related activities and ensure more sustainable production [7].

To address this, a hybrid energy system for stable power and heat supply in offshore oil&gas

installation (HES-OFF) was proposed and investigated [8]. HES-OFF integrates offshore wind power, onsite gas turbines (GTs) and an energy storage system based on PEM fuel cell and electrolyser stacks. A previous paper [9] assessed the performance of the concept, considering the power grid stability and optimal design obtained from the optimization framework developed. The analysis proved possible larger integration of wind capacity in the electric grid without violating the 2% grid frequency maximum allowable variation. This enables CO₂ emissions reduction of up to 36% compared to the equivalent standard GT-based system. The optimum capacity of ELY and FC stacks for reference case were defined as 5.7 and 3.0 MW, respectively, providing sense of desired PEM system capacities in future HES systems offshore. Computationally efficient online tool, named HES-OFF app, allows investigation of various designs of the HES-OFF concept [10].

Based on the screening of available technologies for fuel cell and electrolysers, the PEM technology was chosen for HES-OFF, due to the best compactness and operation ability during transients [11]. Since the stack systems are easily scalable and the anticipated capacity falls in the scale of MWs, the question arises about the realistic estimation of the part-load performance of high-capacity PEM systems. Most of the research activities are related to specific small-capacity stacks, which were mathematically modelled and experimentally tested [12, 13]. The case-specific models tuned for

representation of the given stack require a detailed knowledge about the cell architecture, which limits the possibility of reusing model in more generic conceptual simulations. On the other hand, some researchers present more universal models, which avoid specification of the cell design. This approach can be accomplished by means of 0-D modelling. Campanari et al. [14] developed a lumped zero-dimensional model of the stacks to represent 2 MW PEM FC system. The model was developed in Aspen Custom Modeler® what allowed to analyse the off-design operation, including variations in the fuel cell stoichiometry, operating temperature as well as the influence of cell performance decay. That facilitated tuning and successful validation of the model. However, part-load performance of the MW-scale PEM FC system has not been reported. Paper [15] presents a method for reproducing the main conversion performance and dynamic features of ELY and FC PEM systems. Both investigated systems had 100 kW of power capacity and were simplified for the integration into an optimisation procedure by means of two approximation methods. PEM fuel cell and electrolysers stack models are also available in commercial simulating software libraries such as MATLAB Simscape [16] and ThermoFlow [17]. However, such library models are dedicated for analyses performed within given simulation environment and not all details can be extracted to determine the required part-load performance.

The performed literature review shows various approaches towards modelling of FC and ELY systems at different levels of fidelity. The scope of this paper was to develop suitable approaches to model large-scale FC and ELY systems to be used for the analysis of the HES-OFF concept. The information provided by manufacturers is normally limited to a single operation point, typically full load. Therefore, a well-established, 0 - dimensional approach to calculate the part-load performance of high-capacity PEM systems was developed and embedded in the overall modelling framework in order to accomplish the planned simulation campaign for HES-OFF concept investigation.

The organization of the paper is as follows. A methodology section provides necessary details for introducing the developed models. The results obtained by the implementation of those models in HES-OFF app are then presented. Finally, a short conclusion section outlines the main contributions of this work.

2. Methodology

Modelling of fuel cell and electrolyser systems is a challenging task. Gao et al. [18] outline a general classification of the FC/ELY models based on five sub-categories: spatial dimension, temporal

behaviour, the types of equations applied, modelled area and phenomena. Before diving into the selected approach, an introduction to the HES-OFF concept is presented to provide the context for this work. The computational tool [10] developed for the analysis of the HES-OFF concept encompasses an analysis of the HES over its entire lifetime. A 1-hour time discretization is applied in order to include the irregularity of the wind power. With this time interval, the stack systems can be considered as operating in steady-state mode since the dynamics of PEM FC stacks is characterized by order of minutes [15, 19] and PEM ELY by order of seconds [20]. In addition, the purpose-built models of FC/ELY stacks will represent stacks at the component-level, meaning that no physical or spatial approach will be used. Therefore, the incorporated equations do not represent any phenomena in spatial dimension within the stack (such as gas diffusion direction through membrane), but only scalar variables including cell voltage and cell power [11, 21]. According to the previously listed sub-categories, the models of both fuel cell and electrolysers stacks presented in this work, are characterized by means of a zero-dimensional first-principle approach, where analytical and semi-empirical equations for modelling of static behaviour of single cell/stack are used and the electrochemical domain is the phenomenon being investigated. The most significant equations for operational curves calculations are outlined in the following subsections.

2.1. High-capacity PEM system architecture

Commonly available information about high-capacity PEM systems point to their modular designs. The single stack, as a basic component, is installed into modules, then modules are organized into the final system. Exemplary, NEDSTACK 2 MW_e PEM fuel cell system contains 6 modules with 4 stacks and each stack comprises 1046 cells [22]. This approach is common for MW-scale systems [23]. The schematic of the PEM modular system is depicted in Figure 1.

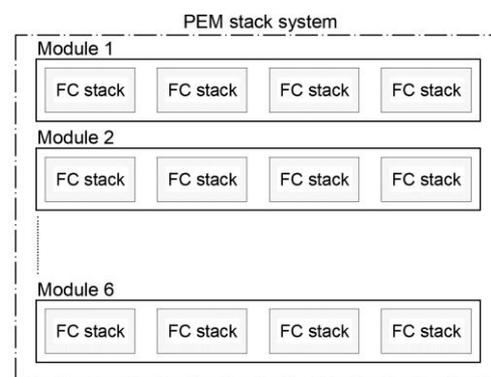


Figure 1: Organisation of fuel cell stacks in modules.

Exact information about the design and configuration of stacks and associated balance of plant (BoP) components are not publicly available, since all manufacturers are cautious about sharing sensitive details, and as for now, not many commercially installed systems are in regular use. Modular design can simplify architecture and reduce cost by sharing common BoP between individual stacks within modules. On the other hand, the control strategy affects the operability of the system and development of an appropriate modelling approach is not a trivial task. In the case of irregular operation, the overarching question is how to load individual stacks defining which ones should be subjected to load changes, hence to more intense cell degradation. At the same time, other stacks can run at relatively steady-state conditions. Another crucial aspect is the weight and footprint of the system. This is the main limitation for designers when it comes to offshore applications. This aspect is thoroughly discussed in [24].

Consideration of complex control strategies and system configuration along with the BoP is out of the scope of this work. The following approach is implemented in analyses accomplished by HES-OFF tool. The total number of required stacks are determined from the desired capacity defined by user. Then for given value of power need determined from the power balance of HES-OFF system, the load is distributed equally to all stacks. In the case of very small amount of power supplied to, or drawn from the system, only one stack is used to provide load as close as possible to the design load conditions of singular stack. To the best authors' knowledge, this approach is a good representation of real systems, especially when the degradation process is not included in the analyses.

2.2. State-of-the-art PEM systems

Currently, the fuel cell manufacturers offer customized scalable systems up to a range of several MW of output power. As mentioned previously, each large-scale PEM system uses a basic singular FC stack, which then is stacked to provide higher power outputs desired by the customer. The performances of some of the commercially available FC stacks are listed in Table 1. The performance is reported in terms of electrical energy output per unit of hydrogen supplied to the FC.

Table 1: PEM fuel cell stacks characteristics.

Stack/ system	Capacity [kW _e]	Perf. @full load	
		[MJ/kgH ₂]	[kWh/ Nm ³ H ₂]
PowerCell S3 stack [25]	125	56.24 ^a	1.288 ^a
Nedstack	500	61.02	1.291
MT-FCPP-	626 ^b	56.34 ^b	1.108 ^b

500 system [26]			
Hydrogenics 1 MW_e system [27]	1 000	58.54	1.461

^acalculated by the developed model

^bpeak power

Practically, to replicate accurately a specific stack performance is not possible due to lack of information. More uncertainties come from the given nominal power of the stack system by manufacturer. For instance, the total nominal power of 500 kW_e Nedstack stack is accomplished by 60 singular smaller stacks (13.6 kW_e each) [26]. Theoretically, the maximum power is $60 \cdot 13.6 \text{ kW}_e = 816 \text{ kW}_e$, what is 62.3% more than nominal 500 kW_e. This opens possibility to claim that stacks are loaded to values closer to their turndown ratios to get higher overall performance. The peak power of this system is limited to 626 kW_e. Thus, the prerequisites behind each specification of particular system, known for manufacturer, adds complexity to tune up and validate the model.

A list of commercial high-capacity PEM electrolyser systems is shown in Table 2. The performance is reported as hydrogen produced per unit of electrical energy supplied.

Table 2: PEM electrolyser stacks characteristics.

Stack/ system	Capacity [kW _e]	Perf. @full load	
		[kgH ₂ /MJ]	[Nm ³ H ₂ / kWh]
NEL MC 100 unit [20]	500	181.44	4.53
H-TEC system ME 450/1400 [28]	1000	192.26	4.80
Hydrogenics HyLYZER 300-30 [29]	1500	176.24- 192.26	4.40-4.80

The performances outlined in Table 2 have been extracted from data sheets and represent rough estimations of expected performances.

For this work, the PowerCell S3 stack for FC system and NEL MC 100 unit for ELY system have been selected. Further literature review has been carried out on those stacks to identify appropriate tuning parameters.

Authors assumed the following simplifications and assumptions for the models development:

- Balance of Plant components are not considered. Thus, there is no limitations for scaling-up the systems' capacities.
- The power consumption for BoP is not considered (power consumption oscillates typically in the range of dozens of kW per 1 MW of the stack capacity).

- The degradation of the stacks is neglected, meaning that the stacks performances at Beginning of Life remain constant until End of Life performance.
- The waste heat produced by stacks is not harvested.

2.3. Fuel cell stack modelling

To compare and assess the performance of the fuel cell stack one needs to know how the operating voltage looks like as a function of current drawn from the fuel cell. To allow comparison of various fuel cells the current density is introduced giving the following correlation:

$$i = I/A_{cell} \quad (1)$$

Where i [A/cm²] is the current density, I [A] the total electrical current produced by fuel cell, A_{cell} [cm²] the surface area of the electrode/electrolyte interface where the fuel cell reactions take place.

The current/voltage density curve (i - V) is called polarisation curve and represents the steady-state performance of singular fuel cell or entire stack. The cell i - V characteristics is an inherent part of PEM cell modelling for building a steady state and dynamic models of the FC/ELY stacks [30].

Having the i - V curve one can easily calculate the power curve of the stack, what in turn allow the estimation of the part-load performance of the stack. The common approach for determination of polarisation curve is presented in Spiegel's book [31]. The MATLAB scripts from this book are adopted. For the sake of clarity and simplicity, not all used correlations are outlined, but only those that are were adopted from other literature references and thus are crucial for understanding the applied approach. For more details, the reader should refer to Spiegel's book or to the other literature references such as [32], where all relevant theory is comprehensively presented.

In order to calculate the open circuit voltage and associated voltage losses the following assumptions are considered:

- Ideal and uniformly distributed reactant gases.
- Constant pressure and temperature in the FC/ELY gas flow channels.
- The hydrogen fuel is humidified and the oxidant is humidified air.
- The FC stack is operated at temperature below 100°C and the reaction product is in liquid phase.
- Parameters for individual cells can be lumped together to represent a fuel cell stack.
- The electrolyte is not electrically conductive and impermeable to gases (no fuel crossovers and internal current associated losses).
- A 100% Faradaic efficiency is assumed.

The polarisation curve is determined calculating the operational voltage output of fuel cell and associated voltage losses.

The actual voltage of fuel cell can be calculated using the following equation:

$$V_{out} = E_{Nernst} + V_{act} + V_{ohmic} + V_{conc} \quad (2)$$

E_{Nernst} is given by the Nernst's equation and determines the ideal voltage that a fuel cell can deliver at the given conditions. According to [31], the equation for PEM fuel cell is expressed as:

$$E_{Nernst} = -\frac{G_{f,liq}}{nF} - \frac{RT}{nF} \ln \left(\frac{p_{H_2O}}{p_{H_2} p_{O_2}^{1/2}} \right) \quad (3)$$

Where $G_{f,liq} = -228170$ J/mol is the Gibbs energy of water in liquid form in standard state, $T = 60^\circ\text{C}$ the operating temperature of fuel cell stack, n the number of electrons transferred per mole of reactant (for hydrogen oxidation $n = 2$), $F = 96485$ Coulombs the Faraday's constant, $R = 8.31446$ J/(mol·K) the ideal gas constant and p_{H_2O} , p_{H_2} , p_{O_2} the partial pressures of water, hydrogen and oxygen, respectively, calculated by the correlations presented in [31].

The voltage losses from Eq. (2) can be determined by the following equations. The activation losses are estimated using the Tafel equation:

$$V_{act} = -\frac{RT}{nF\alpha} \log \left(\frac{i}{i_0} \right) \quad (4)$$

Where $\alpha = 0.5$ is the nondimensional charge transfer coefficient and i_0 the exchange-current density i.e., the current density at zero overpotential. i_0 is assumed to be $6.7 \cdot 10^{-5}$ A/cm² in accordance with the value for a low temperature PEM fuel cell published in [33]. The assumed i_0 value complies with the Tafel's equation assumptions that ($i \gg i_0$). The ohmic losses are estimated using Ohm's law:

$$V_{ohmic} = -ir_{ohmic} \quad (5)$$

Where r_{ohmic} is the area specific ohmic resistance, which is constant for a given fuel cell. In this paper $r_{ohmic} = 0.15$ Ohm · cm² is assumed, a figure that falls in the range of values seen in the literature [32, 33]. Based upon tuning the adopted value gave the best results.

The concentration losses can be calculated using the following equation [34]:

$$V_{conc} = \frac{RT}{nF} \ln \left(1 - \frac{i}{i_L} \right) \quad (6)$$

Where i_L is the limiting current density, which is assumed to be 0.74 A/cm².

The output power of the stack can be determined from the following equation:

$$P_{FCS} = V_{out} \cdot i \cdot A_{FCS} \quad (7)$$

Where A_{FCS} is the total active area of the stack. According to the datasheet [25], the stack comprises 455 cells to provide 125 kW of power output. An assumed value of i_L in Eq. (6) and active area of single cell of 425 cm² allowed determining the power curve (Fig. 2) representing the PowerCell S3 stack. The assumed parameters return a power density of the stack of 646 mW/cm², which is in line with the range of 500-2500 mW/cm² of modern PEM fuel cell systems [32].

2.4. Electrolyser stack modelling

Similarly, as for the fuel cell stack, the model equations representing the electrolyser stack system are programmed in a MATLAB script. The methodology for calculating the polarisation and power curves and performance remain the same as those outlined for the fuel cell stack. The PEM electrolyser works opposite to the fuel cell, relying on the same mechanisms and phenomena. As it is seen in Tribioli et al. [35] the ELY curves can be calculated from Eq. (2), in which the voltage losses are multiplied by reversed coefficient (i.e. -1). This is the simplest approach towards modelling the electrolyser stack and is used for reversible stacks only (i.e. a stack working both in ELY and FC mode, as analysed in [35]). In the case of standalone ELY stack the correlations have to capture the electrochemical phenomenon during electrolyses and for this purpose the following correlation were chosen for the electrolyser modelling.

The actual operating voltage of the ELY system is determined using Eq. (2). The terms of the equation are determined by correlations presented in [36]. The Nernst equation for ELY system is determined by the following formula:

$$E_{Nernst} = 1.229 - 8.5 \cdot 10^{-4}(T - T_a) + 4.3085 \cdot 10^{-5} T \ln \left(\frac{p_{H_2} p_{O_2}^{1/2}}{p_{H_2O}} \right) \quad (8)$$

The activation over potential losses is defined by:

$$V_{act} = \left(\frac{\alpha_A + \alpha_C}{\alpha_A \alpha_C} \right) \frac{RT}{nF} \log \left(\frac{i}{i_o} \right) \quad (9)$$

Where $\alpha_A = 0.5$ and $\alpha_C = 1.0$ are the charge transfer coefficients of the anode and cathode, respectively. In Eqs. (8)-(9), T stands for operating temperature of the ELY stack and its value is assumed to be 70°C according to [37]. The exchange current density i_o is determined by the following correlation [38]:

$$i_o = 1.08 \cdot 10^{-17} \exp(0.086T) \quad (10)$$

The correlation for the ohmic potential is determined by Eq. (5), where $r_{ohmic} = 0.238 \text{ Ohm} \cdot \text{cm}^2$, which

is representative of the value seen in the Nafion membranes at moderate thicknesses according to [37].

The last term for determination of the operational voltage of the electrolysers is the concentration over potential losses, which is calculated by the following correlation [39]:

$$V_{conc} = i \left(\beta_1 \frac{i}{i_L} \right)^{\beta_2} \quad (11)$$

Where $\beta_2 = 2$ is a constant according to [40] and β_1 is defined by the following correlation:

$$\beta_1 = \begin{cases} (8.66 \cdot 10^{-5}T - 0.068)P_x - 1.6 \cdot 10^{-4}T + 0.54, & (P_x > 2) \\ (7.16 \cdot 10^{-4}T - 0.622)P_x - 1.45 \cdot 10^{-3}T + 1.68, & (P_x < 2) \end{cases} \quad (12)$$

Where P_x is expressed as:

$$P_x = \frac{P_{O_2}}{0.1173 \cdot 101325} + \frac{P_{sat}}{101325} \quad (13)$$

The limiting current density i_L is assumed to be 20 A/cm² [40]. Partial pressure and saturation pressure is calculated in the same manner as for FC and the correlation are available in the literature [31]. The pressure of the produced hydrogen is assumed 30 barg according to the data sheet [20].

In order to calculate the polarisation and power curves, and to determine the part-load performance, a number cells equal to 1695 and an area of single cell equal to 70 cm² are selected so to achieve the power output of 500 kW. These values are seen in the state-of-the-art ELY stacks.

Equations (1)-(13) are programmed in the MATLAB script to calculate the polarisation and power curves (Fig. 4). Subsequently, the part load performance of the stack can be estimated.

3. Results and discussion

3.1. Part-load performance – fuel cell stack

Fuel cell and electrolyser manufacturers generally provide scarce information about the performance of their systems. A direct comparison between calculated polarisation and power curves with their counterparts in datasheets become a cumbersome task, partly because of possible lack of compatibility between the operation parameters, partly because of unknown information such as active area or limiting current density. Therefore, the described methodology was used for determination of the operational curves and part-load performance. Some parameters were assumed according to the literature, while some others were tuned to match expected performance. Figure 2 depicts the determined i - V curve and the power curve as a function of current density.

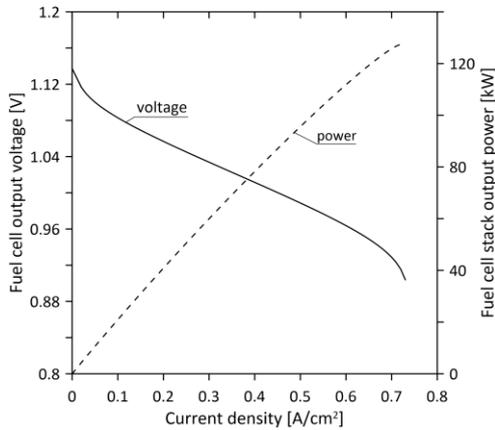


Figure 2: Polarisation curve ('voltage') and power curve of the modelled generic 125 kW fuel cell stack.

The curves appear to be in line with those seen in the literature for similar systems, for instance in [25, 33]. Having power curve as a function of current density i , one can easily determine the actual usage/production of hydrogen by FCS/ELY stack using the Faraday's second law of electrolysis given by Eq. (14).

$$\Delta \dot{N}_{H_2,reacted} = \frac{iA}{nF} \quad (14)$$

Where $\dot{N}_{H_2,reacted}$ [mol/s] represents the molar flow rate of the hydrogen reacted (consumed or produced by the stack system) and A [cm²] is the total active area of the stack.

Once the hydrogen consumption is determined, one can calculate the part-load performance of the stack as a function of the output power.

Fig. 3 shows that the performance at full load (125 kW) is 56.24 MJ/kg_{H₂}, which converts to 1.288 kWh/Nm³. Such value is in line start-of-the-art stacks, such as those listed in Table 1.

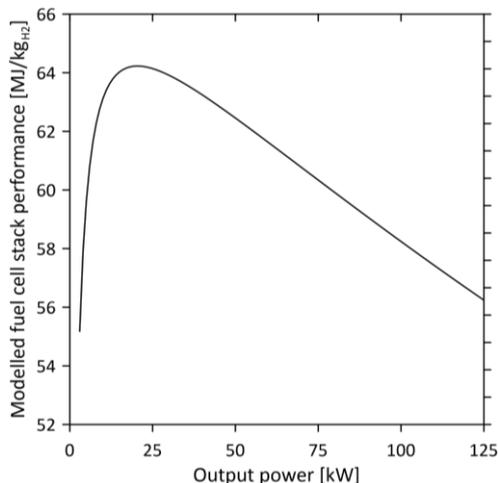


Figure 3: Part-load performance of the modelled 125 kW fuel cell stack.

The performance increases while decreasing the load achieving the maximum at around 20 kW. The fuel cell stack is an electrochemical device so its efficiency is not limited by Carnot cycle, as it is for heat engines. Under those part load conditions the voltage losses are smaller, thus the overall performance is higher. The performance rapidly drops for loads below 20 kW, which is explainable by the activation losses.

3.2. Part-load performance – electrolyser

PEM electrolysers are operated typically at higher current densities. The calculated polarisation and power curves are presented in Fig. 4. The modelled voltage and power curves show a rather linear characteristic. Using the Faraday's law (Eq. (14)), it is possible to calculate the part-load performance of the electrolyser stack. The ELY performance is depicted in Fig. 5.

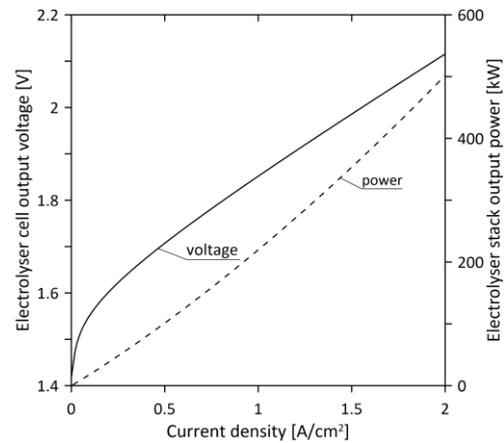


Figure 4: Polarisation curve ('voltage') and power curve of the modelled generic 500 kW electrolyser stack.

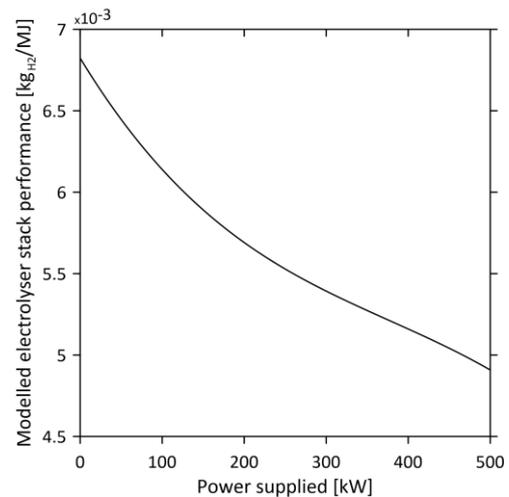


Figure 5: Part-load performance of the modelled 500 kW electrolyser stack.

Similarly, to fuel cells, the electrolyser shows a performance increase as the load drops.

3.3. Results from HES-OFF app

Analysis of the hybrid energy systems requires estimation of performance of their vital components under various loads. Once, the FC and ELY part-load performances are determined and implemented in the HES-OFF app, it is possible to investigate various layouts and configurations of the hybrid system. The default HES-OFF schematic is depicted in Fig. 6.

The offshore installation can be specified by heat and power demand in three distinctive stages of life. The adopted values for this study are specified in Table 4.

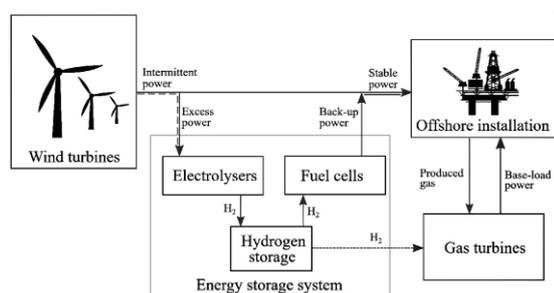


Figure 6: Schematic of the HES-OFF concept.

Table 4: The offshore installation heat and power demands adopted in analyses.

Stage of life	Offshore installation demands	
	Heat [MW]	Power [MW]
Peak years	12.0	35.5
Mid years	10.0	34.2
Tail years	7.0	32.9

The HES-OFF app allows to define own values for parameters and choose predefined components' models used then in the analyses or just assume the predefined default inputs. To demonstrate the HES-OFF performance the following capacities and components have been assumed: single gas turbine (LM2500+G4, rated power 32 MW), NREL wind turbine model, initial level of hydrogen in the storage 80%, wind farm rated power 30 MW, ELY rated power 4 MW, FC rated power 5 MW, hydrogen storage capacity 23 000 kg. The remaining entries were specified according to default values predefined in the app. The calculations were performed for 1 year of time for each stage of life. The results presented in Table 5 outline the environmental gains accomplished by HES-OFF system in comparison to the case where two gas turbines parallel to the same wind farm provide heat and power to the offshore facility.

Table 5: HES-OFF environmental performance with respect to two gas turbines (GT) tied to wind turbines (WT) for a total 30 MW wind farm.

Stage of life	HES-OFF concept vs. GT + WT system	
	CO ₂ emission reduction [%]	Extra wind energy integration [%]
Peak	25.7	3.5
Mid	27.0	3.8
Tail	22.8	2.7

As it can be seen from Table 5, the integration of hydrogen ELY and FC system along with hydrogen storage allows reducing the carbon intensity by around 23-27%. A key advantage is the possibility to use a single GT with the HES-OFF system. In addition, the HES-OFF concept allows a better integration of high-capacity wind farm, leading to a higher wind energy usage by 2.7 to 3.8% compared to a similar system without energy storage. Other meaningful results are related to the energy delivered or supplied by the PEM systems. The exact numbers are listed in Table 6.

Table 6: PEM ELY and FC system energy integrated in HES-OFF system.

Stage of life	FC energy [GWh _e]	ELY energy [GWh _e]
Peak	2.66	14.47
Mid	1.17	8.11
Tail	0.24	4.53

Table 6 shows that ELY system transform more electrical energy than FC system. The ELY system absorbs surplus power from the wind farm, allowing not to dissipate it. The surplus power is converted into hydrogen that is stored and later either used in the FC stack or in the GT. FC stack system plays as a backup power supply when the GT power and wind power at given instance cannot meet the platform power demand. The difference between the ELY and FC energy numbers are due to the roundtrip efficiency and to the utilization of part of the hydrogen in the GT. For the sake of clarity not all results obtained from HES-OFF app are outlined in this work. Readers are advised to explore independently the capabilities of the online tool to get more insight into HES-OFF.

4. Summary and conclusion

Conceptual analyses of new innovative hybrid energy systems require sufficiently accurate components' models, which address the simulation requirements and produce reasonable outcomes. Typically, due to varying heat and power demand of the offshore facility and intermittent nature of wind energy the components are subjected to part-load working conditions. This yields a general need of having insights into part-load performances. Especially, for novel concepts the deep knowledge

about the system architecture and exploitation data are not widely available. Therefore, a well-established methodology was incorporated to determine the expected part-load performances of the high capacity PEM fuel cell and electrolyser systems. Thanks to this approach, the knowledge gap for HES-OFF concept has been filled, allowing to accomplish more realistic simulations outcomes.

Funding

This research was funded by the Research Council of Norway under the program PETROMAKS2, grant number 281986, project “Innovative Hybrid Energy System for Stable Power and Heat Supply in Offshore Oil and Gas Installation (HES-OFF).”

References

- [1] IPCC. *Special Report: Global Warming of 1.5°C*; Cambridge University Press: Cambridge, UK, 2021.
- [2] The European Green Deal. Document 1 „*Communication From The Commission To The European Parliament, The European Council, The Council, The European Economic And Social Committee And The Committee Of The Regions*”.
- [3] A. Maroufmashat, M. Fowler, “Transition of Future Energy System Infrastructure; through Power-to-Gas Pathways”, *Energies*, vol. 10, no. 8, 2017.
- [4] Kongsberg [webpage online](#).
- [5] Handbook for hydrogen-fuelled vessels. DNV & MarHySafe JFP Phase 1, 1st edition June 2021.
- [6] A. L. Facci, M. Lauricella, S. Succi, V. Villani, G. Falcucci, “Optimized Modeling and Design of a PCM-Enhanced H₂ Storage”, *Energies*, vol. 14, no 6, 2021.
- [7] L. Riboldi, M. Pilarczyk, L.O. Nord, “The Impact of Process Heat on the Decarbonisation Potential of Offshore Installations by Hybrid Energy Systems”, *Energies*, vol. 14, no. 23, 2021.
- [8] L. Riboldi, E. F. Alves, M. Pilarczyk, E. Tedeschi, L. O. Nord, “Innovative Hybrid Energy System for stable Power and Heat Supply in offshore oil & gas Installation (HES-OFF): System Design and Grid Stability”, *Computer Aided Process Engineering*, Elsevier, vol. 48, pp. 211-216, 2020.
- [9] L. Riboldi, E. F. Alves, M. Pilarczyk, E. Tedeschi, L. O. Nord, “Optimal Design of a Hybrid Energy System for the Supply of Clean and Stable Energy to Offshore Installations”, *Frontiers in Energy Research*, vol. 8, 2020.
- [10] HES-OFF app. Online access: <https://hes-off.herokuapp.com/>
- [11] F. Barbir. “PEM Fuel Cells: Theory and Practice”. 2nd ed., *Academic Press*, 2013.
- [12] M. H. Nehrir and C. Wang, “Modeling and Control of Fuel Cells: Distributed Generation Applications”, 2009.
- [13] K. Ondrejčka, V. Ferencey, M. Stromko, “Modeling of the air-cooled PEM fuel cell”, *IFAC-PapersOnLine*, vol. 52, pp. 98-105, 2019.
- [14] S. Campanari, G. Guandalini, J. Coolegem, J. ten Have, P. Hayes, A. H. Pichel, “Modeling, Development, and Testing of a 2 MW Polymeric Electrolyte Membrane Fuel Cell Plant Fueled With Hydrogen From a Chlor-Alkali Industry”, *Journal of Electrochemical Energy Conversion and Storage*, vol. 16, no. 4, 2019.
- [15] P. Gabrielli, B. Flamm, A. Eichler, M. Gazzani, J. Lygeros, M. Mazzotti, “Modeling for optimal operation of PEM fuel cells and electrolyzers”, *IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)* pp. 1-7, 2016.
- [16] MATLAB, Simulink – Simscape, 2018. 9.7.0.1190202 (R2019b), Natick, Massachusetts: The MathWorks Inc.
- [17] Thermoflow Version 29, 2021. Thermoflow Inc.
- [18] Gao, F., Blunier, B., Miraoui, A. “*Proton Exchange Membrane Fuel Cells Modeling*”. John Wiley & Sons Inc, 2012.
- [19] PowerCell press release, [no. 829664](#), Gothenburg, Sweden.
- [20] [NEL M Series Containerized Hydrogen Generation Systems](#).
- [21] F. Barbir, A. Basile, N. Veziroglu. “Compendium of Hydrogen Energy Volume 3: Hydrogen Energy Conversion”, *Elsevier*, 2016.
- [22] J. Coolegem. “Nedstack fuel cell technology” presentation during ICAE2016 conference, Beijing, 10.10.2016.
- [23] P. Ekdunge. “H2 and fuel cells in maritime application”, 15 June 2017, Valencia.
- [24] F. B. Hultin. MSc thesis: “[High-Capacity PEM Fuel Cells and Electrolysers in Hybrid Energy System for Offshore Power Generation](#)” NTNU – open 2020.
- [25] PowerCell Sweden AB, 2020. PowerCell S3 datasheet No. 190430.
- [26] PemGen MT-FCPI-500 [Nedstack datasheet](#).
- [27] [Hydrogenics. Fuel cell megawatt power generation platform](#).
- [28] H-TEC SYSTEMS GmbH, 2020. [H-TEC Series-ME: ME 450/1400](#).
- [29] Hydrogenics [HyLYZER@ 300-30](#).
- [30] B. Gou, W. Na, B. Diong. “Fuel Cells: Dynamic Modeling and Control with Power Electronics Applications” Second Edition. *CRC Press Taylor & Francis Group*, 2016.
- [31] C. Spiegel. “PEM Fuel Cell Modeling and Simulation Using MATLAB®”. *Elsevier*, 2008.
- [32] R. O’Hayre, S. Cha, W.G. Colella, F. Prinz. “Fuel Cell Fundamentals”. 3rd ed., *Wiley*, 2016.
- [33] A. Dicks, D.A.J. Rand. “Fuel Cell Systems Explained”. 3rd ed., *Wiley*, 2018.
- [34] P. Gabrielli, Flamm B., A. Eichler, M. Gazzani, J. Lygeros, M. Mazzotti. “Modeling for Optimal Operation of PEM Fuel Cells and Electrolyzers” in *Proceedings of 16th International Conference on Environment and Electrical Engineering*, 2016.
- [35] L. Tribioli, R. Cozzolino, L. Evangelisti, G. Bella, “Energy Management of an Off-Grid Hybrid Power Plant with Multiple Energy Storage Systems”. *Energies*, vol. 9, no. 661, 2016.
- [36] H.C. Zhang, S. H. Su, G. X. Lin, J. C. Chen, “Efficiency Calculation and Configuration Design of a PEM Electrolyzer System for Hydrogen Production”, *International Journal of Electrochemical Science*, vol. 7, Issue 5, pp. 4143-4157, 2012.
- [37] A. Godula-Jopek, “Hydrogen Production by Electrolysis” ©Wiley-VCH Verlag GmbH & Co., Germany, 2015.
- [38] S.O. Mert, I. Dincer, Z. Ozcelik, “Exergoeconomic analysis of a vehicular PEM fuel cell system”, *Journal of Power Sources*, vol. 165, Issue 1, pp. 244-252, 2007.
- [39] J. Nieminen, I. Dincer, G. Naterer, “Comparative performance analysis of PEM and solid oxide steam electrolyzers”, *International Journal of Hydrogen Energy*, vol. 35, Issue 20, pp. 10842-10850, 2010.
- [40] A.T. Marshall, S. Sunde, M. Tsyppkin, R. Tunold, “Performance of a PEM water electrolysis cell using IrxRuyTazO2 electrocatalysts for the oxygen evolution electrode”, *International Journal of Hydrogen Energy*, vol. 32, Issue 13, pp. 2320-2324, 2007.

Modeling of the two-phase flow during depressurization of liquified CO₂ in a pipe

Osama M Ibrahim^{a,*}, Prasanna Welahettige^a, Knut Vågsæther^a, Bernt Lie^a

^a *The University of South-Eastern Norway, Faculty of Technology, Natural Sciences, and Maritime Sciences,*

Kjølnes Ring 56, 3918 Porsgrunn, Norway

Osama.K.Ibrahim@usn.no

Abstract

Modeling transient CO₂ two-phase flow in a pipe is essential in studying depressurization mechanisms resulting from liquified CO₂ accidental release. Generated data from such models predict the released flow characteristics and possible propagating fractures. Accordingly, they provide valuable input for risk prevention in designing and safely operating CO₂ transport pipelines. CO₂ depressurization simulation involves fluid-mechanical and thermodynamic models, primarily expressed by hyperbolic partial differential equations and an equation of state (EOS). Besides, these models are solved with appropriate numerical methods. This paper deals with a drift flux - homogeneous equilibrium model (HEM) construction utilizing central-upwind-weighted essentially non-oscillatory (WENO) numerical schemes to describe two-phase flow during CO₂ decompression in a pipe. Rapid transition in mass, momentum, and energy at the interface between liquid and vapor phases is assumed in the flow HEM. Thus, the two-phase flow is in a thermal, mechanical, and chemical equilibrium. The thermodynamic properties are calculated by applying Span-Wagner EOS. The high-resolution second-order central, central-upwind, and third-order weighted essentially non-oscillatory (WENO) schemes have been executed with the HEM, and they effectively captured rapid phase transition. The central-upwind and essentially non-oscillatory (ENO) schemes' stencils can be appropriate for constructing a higher-order accuracy central-upwind-WENO scheme. This structured scheme uses a smoothness indicator as an alternative to the limiter function. Besides, the variables in the cell interface are determined by WENO reconstruction, while central-upwind is used to compute the flux properties.

1. Introduction

As long as fossil fuel is the predominant energy source, carbon capture and storage (CCS) technology should be critical to reducing greenhouse gas. CCS is projected to contribute more than 20% to CO₂ emissions reduction by 2050, as reported by the International Energy Agency (IEA) [1]. That means above nine gigatons should be captured and stored every year. Safe and consistent carbon dioxide (CO₂) transport from capturing plants to storage sites is vital in the CCS process. On- and offshore pipelines represent a favorable option for transporting large quantities of CO₂. However, this process raises growing safety concerns as CO₂ is transported in a high pressurized dense phase and involves a significant hazard of pipe failure. Pipelines can endanger running-ductile fracture from expansion and explosive evaporation waves following the rapid phase transition during liquid CO₂ depressurization [2]. Subsequently, the pipe may expose to destructive damage due to the fast reduction in temperature during the evaporation process. Therefore, estimating the thermodynamic

state properties during the decompression process is crucial for the appropriate design, safe operation, and maintenance of CO₂ transport pipes. For this purpose, transient flow models are required, whereby fluid-mechanical and thermodynamic models are coupled with appropriate numerical methods to resolve these models. These models should be validated and corrected by experimental data.

The CO₂ pipe's mechanical failure potentially results in a release of two-phase flow, following isentropic expansion due to pressure decrease. The intensity and propagation velocity of downstream two-phase flow depends on the upstream expansion waves' velocity traveling into superheated liquid. Thus, estimating the change in velocities along the pipe is essential to predict the release behavior and the pipe fracture development. Several models have been used to predict state properties during pure or CO₂-rich mixtures depressurization. Elshahomi et al. [3] have predicted CO₂ state properties during condensation phase transition using ANSYS Fluent, CFD procedure. The results had good agreement

with previous shock tube experimental data. Munkejord et al. [4, 5] have applied two models to study the influence of impurities and the initial temperature on wave velocity during CO₂-rich mixture decompression. The homogeneous equilibrium model (HEM) and the two-fluid model (TFM) were formulated by implementing the multi-stage centered scheme (MUSTA) simultaneously with Soave-Redlich-Kwong and Peng-Robinson EOS to calculate the thermodynamical properties. The models' calculation findings were contrasted with shock tube test results, which demonstrated minimal enhancement in TFM predictions against HEM. Botros et al. [6] have performed shock tube experiments on pure CO₂ and compared the results with estimations from models built on three equations of state: GERG-2008, Peng-Robinson, and BWRs. They concluded that only GERG-2008 corresponded with experimental data. However, the inconsistency increased with temperature reduction. Morin et al. [7] have presented an approximate Riemann solver of Roe for the pipe's CO₂ depressurization. Simulation results demonstrated that the central scheme underestimated the pressure pulse's maximum amplitude. While the high-resolution Roe scheme accurately described the pulse. A homogeneous relaxation model (HRM) with cubic EOS was introduced by Brown et al. [8] to describe the released flow resulting from dense phase CO₂ pipe rupturing. The model was resolved by conjugating a semi-discrete Finite Volume Method and HLL (Harten-Lax-van Leer) approximate Riemann solver. It also showed reasonable agreement with the real data from rupturing the CO₂ pipeline. In this paper a drift flux - (HEM) is built utilizing central-upwind-weighted essentially non-oscillatory (WENO) numerical schemes to illustrate two-phase flow following CO₂ decompression in a pipe.

2. Methodology

This model incorporates the drift-flux model (homogeneous equilibrium), Span-Wagner Equation of State, and centered-upwind-WENO scheme, including the Runge-Kutta fourth-order method to describe the transients' characteristics during liquid CO₂ decompression. The model illustrates one component (CO₂) with two-phase flow inside a horizontal pipe and is assumed to be in thermal, mechanical, and chemical equilibrium. i.e., the gas and liquid phases have the same temperature (T), pressure (P), and chemical potential (μ).

$$T_l = T_g = T; \quad P_l = P_g = P; \quad \mu_l = \mu_g = \mu \quad (1)$$

2.1. Fluid dynamics in the pipe flow model

Unlike the Drift Flux Model (DFM) in the HEM model, the gas and liquid velocity are identical (no slip). In addition, the model is one-dimensional due to flow in one direction along the pipe. The flow variables are averaged over the pipe's cross-section. The governing equations are in the form of hyperbolic partial differential equations similar to Euler's equations for compressible inviscid flow of a single fluid and can be described as

$$\frac{\partial}{\partial t} U + \frac{\partial}{\partial x} F = 0 \quad (2)$$

where U is the vector that contains the conserved variables and $F(U)$ is the flux function:

$$U = [\rho, \rho v, E]; \quad F = [\rho v, \rho v^2 + P, v(E + P)] \quad (3)$$

Correlated closure laws for the HEM constitute the two-phase mixture density ρ , the total energy E , the corresponding mixture's specific internal energy e_{mix} , and the volume fraction relation as follows

$$\left. \begin{aligned} \rho &= \alpha_g \rho_g + \alpha_l \rho_l; \\ E &= \rho(e + 0.5v^2); \\ e &= (\alpha_g \rho_g e_g + \alpha_l \rho_l e_l) / \rho; \\ \alpha_g + \alpha_l &= 1 \end{aligned} \right\} \quad (4)$$

where ρ denotes density, v - velocity, E - the total energy, e - specific internal energy, P - pressure, α_i - volume fraction of the phase i , and subscripts g and l are gas and liquid phases, respectively. For simplicity, the source terms are not included, i.e., the heat transfer through the pipe wall, wall friction, and the gravitational forces are neglected. Equation (2) can be written in a quasi-linear form as :

$$U_t + A(U)U_x = 0 \quad (5)$$

where $A(U) = \frac{\partial F}{\partial U}$ is the Jacobian matrix of the system.

2.2. Thermodynamics

A function $P(e, \rho)$ that relates the pressure to the internal energy and density is required to complete the model. This work uses the Span-Wagner Equation of State (SW EOS). It is considered the most accurate EOS for CO₂ and is based on Helmholtz's free energy (HFE) as a function of temperature, volume, and the number of moles: $A(T, V, N)$. The molar properties of the HFE and the volume are

$$a = \frac{A}{N}; \quad v = \frac{V}{N} \quad (6)$$

It implies that $a(T, \nu)$. The relation of a to the other thermodynamic properties can be expressed in terms of partial derivatives of the independent variable (T, ν) . The relation to the pressure, internal energy, and entropy are as follows:

$$\left. \begin{aligned} P &= -\frac{\partial a}{\partial \nu}; \\ e &= a + Ts = a - T\frac{\partial a}{\partial T}; \\ s &= -\frac{\partial a}{\partial T}; \end{aligned} \right\} (7)$$

As the phase transition occurs following adiabatic pressure reduction, The molar entropy is conserved. The temperature is unknown, and it is required to solve one or more linear equations involving e and ν to determine its value. SW EOS demonstrates explicit expressions for HFE's appropriate partial derivatives, allowing for numerically solving for the temperature using Newton's method. The initial molar volume and internal energy (ν_0, e_0) can be determined as a function of the known initial saturated temperature and pressure (T_0, P_0) . Then the equation involving molar internal energy as a function of temperature and molar volume should be solved by newton's method iteration

$$\begin{aligned} G(X) &= [e(X) - e_0] = 0; \\ X_{i+1} &= X_i - J^{-1} \cdot G(X) \end{aligned} \quad (8)$$

Here X signifies the vector $(T, \nu)^T$, and $J = \frac{\partial e}{\partial T}$ is the Jacobian of G . Then the thermodynamic state is determined as a function of T and ν (Mjaavatten, 2022).

The speed of sound in the mixture is given as:

$$c_m = \sqrt{\frac{\nu^2}{M_W} \cdot \left(\frac{\partial a^2}{\partial \nu^2} - \frac{\partial a^2}{\partial \nu \partial T} \right)} \quad (9)$$

In addition, $\rho = \frac{M_W}{\nu}$; where M_W is the molecular weight (for CO₂, 44.01g/mol).

2.3. Numerical method

Equation (2) can be solved using the finite-volume scheme, wherein the domain is discretized into subdomains (control volumes). The semi-discrete formula (spatial operator discretization) can be obtained by Eq. (2) integration:

$$L(U) = -\frac{1}{\Delta x} \left(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} \right) \quad (10)$$

where $L(U)$ - discretization of the spatial operator; $F_{i+\frac{1}{2}}$ - the flux between cells i and $i+1$. Time iteration for the semi-discrete formula is executed using the total variation diminishing (TVD) Runge-Kutta method (RKM). According to Gottlieb et al.

(1998), this method keeps TVD properties satisfying the Courant-Friedrichs-Lewy (CFL) condition. The time iterations by fourth-order TVD RCM are expressed as (Shu and Osher, 1988) :

$$U^{(1)} = U^{(0)} + \frac{1}{2} \Delta t L(U); \quad (11.1)$$

$$U^{(2)} = \frac{1}{2} U^{(0)} + \frac{1}{2} U^{(1)} - \frac{1}{4} \Delta t L(U^{(0)}) + \frac{1}{2} \Delta t L(U^{(1)}); \quad (11.2)$$

$$U^{(3)} = \frac{1}{9} U^{(0)} + \frac{2}{9} U^{(1)} + \frac{2}{9} U^{(2)} - \frac{1}{9} \Delta t L(U^{(0)}) - \frac{1}{3} \Delta t L(U^{(1)}) + \Delta t L(U^{(2)}); \quad (11.3)$$

$$U^{(4)} = \frac{1}{3} U^{(1)} + \frac{1}{3} U^{(2)} + \frac{1}{3} U^{(3)} + \frac{1}{6} \Delta t L(U^{(1)}) + \frac{1}{6} \Delta t L(U^{(3)}) \quad (11.4)$$

where $U^{(0)}$ and $U^{(4)}$ are properties at the time n and $n+1$, respectively. Considering the CFL limitation, the time step is specified as:

$$\Delta t \leq c \cdot (\Delta x / \max\{\lambda_i\} \forall i) \quad (12)$$

where c is the CFL number and λ_i are the eigenvalues of the Jacobian matrix: $A(U) = \frac{\partial F}{\partial U}$.

It is challenging to discretize precisely the cell interface flux function. i.e., The flux $F_{i+\frac{1}{2}}$ is challenging to determine at the cell interface given the values of U_i, U_{i+1}, U_i^n , and U_i^{n+1} . Many numerical approaches have been suggested to compute $F_{i+\frac{1}{2}}$. This work applies the centered-upwind-WENO scheme (CU-WENO), developed by (Welahettige et al., 2022), for the HEM to calculate the flux functions by the central-upwind scheme while the cell interface properties by WENO reconstruction. This formulation seeks to integrate the central-upwind features into the WENO scheme, which could improve the results' accuracy and stability.

The cell interface flux is computed according to the central-upwind expression presented by Kurganov et al. (2001). The scheme specifies the local speeds on the side of wave propagation's directions (discontinuity propagation speeds). Their values are equivalent to the maximum and minimum of the Jacobian matrix's eigenvalues. The positive and negative discontinuity propagation speeds are defined as

$$a_{i+\frac{1}{2}}^{n,+} = \max \left\{ \begin{aligned} &\max \left\{ \lambda \left(U_{i+\frac{1}{2}}^{n,-} \right) \right\}, \\ &\max \left\{ \lambda \left(U_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \end{aligned} \right\} \quad (13.1)$$

$$a_{i+\frac{1}{2}}^{n,-} = \min \left\{ \begin{array}{l} \min \left\{ \lambda \left(U_{i+\frac{1}{2}}^{n,-} \right) \right\}, \\ \min \left\{ \lambda \left(U_{i+\frac{1}{2}}^{n,+} \right), 0 \right\} \end{array} \right\} \quad (13.2)$$

where λ denotes the eigenvalues of HEM:

$$\lambda = \{v + c_m, v, v - c_m\} \quad (13.3)$$

Then, for the central-upwind scheme, the high-resolution numerical flux is expressed as:

$$F(U)_{i+\frac{1}{2}}^n = \frac{a_{i+\frac{1}{2}}^{n,+} \cdot F \left(U_{i+\frac{1}{2}}^{n,-} \right) - a_{i+\frac{1}{2}}^{n,-} \cdot F \left(U_{i+\frac{1}{2}}^{n,+} \right)}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} - \frac{a_{i+\frac{1}{2}}^{n,+} \cdot a_{i+\frac{1}{2}}^{n,-}}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} \cdot \left(U_{i+\frac{1}{2}}^{n,+} - U_{i+\frac{1}{2}}^{n,-} \right) \quad (14)$$

The cell interface properties are calculated by applying a fifth-order WENO reconstruction based on a third-order ENO stencils approach; for (r)-order ENO, (2r-1)-order WENO can be constructed. The cell averages calculate the cell interface properties in each stencil. So, the negative direction value at the cell interface $i + \frac{1}{2}$ is given as:

$$U_{i+\frac{1}{2}}^- = \omega_0 U_{i+\frac{1}{2}}^{-,0} + \omega_1 U_{i+\frac{1}{2}}^{-,1} + \omega_2 U_{i+\frac{1}{2}}^{-,2} \quad (15)$$

Here, ω is a weight function. The sum of weight functions is one ($\omega_0 + \omega_1 + \omega_2 = 1$), and they are calculated as:

$$\begin{aligned} \omega_0 &= \frac{\alpha_0}{\alpha_0 + \alpha_1 + \alpha_2}; \omega_1 = \frac{\alpha_1}{\alpha_0 + \alpha_1 + \alpha_2}; \\ \omega_2 &= \frac{\alpha_2}{\alpha_0 + \alpha_1 + \alpha_2} \end{aligned} \quad (16)$$

Where α_i are given as:

$$\begin{aligned} \alpha_0 &= \frac{1}{10 \cdot (\epsilon + \beta_0)^2}; \alpha_1 = \frac{6}{10 \cdot (\epsilon + \beta_1)^2}; \\ \alpha_2 &= \frac{3}{10 \cdot (\epsilon + \beta_2)^2} \end{aligned} \quad (17)$$

Where setting $\epsilon = 10^{-6}$ aims to avoid division by zero. β - is a smoothness indicator that involves (r-1)-order polynomial function

$$\beta_0 = \frac{13}{12} (U_{i-2} - 2U_{i-1} + U_i)^2 + \frac{1}{4} (U_{i-2} - 4U_{i-1} + 3U_i)^2; \quad (18.1)$$

$$\beta_1 = \frac{13}{12} (U_{i-1} - 2U_i + U_{i+1})^2 + \frac{1}{4} (U_{i-1} - U_{i+1})^2; \quad (18.2)$$

$$\beta_2 = \frac{13}{12} (U_i - 2U_{i+1} + U_{i+2})^2 + \frac{1}{4} (3U_i - 4U_{i+1} + U_{i+2})^2 \quad (18.3)$$

The positive direction value $U_{i+\frac{1}{2}}^+$ at the cell interface $i + \frac{1}{2}$ can be calculated by symmetry.

3. Results and discussions

Numerical simulations analyzing the transient pipe flow behavior during the evaporation of the CO₂ dense phase are presented. The CU-WENO numerical schemes resolve hyperbolic conservation equations. The fluid dynamics part is combined with SW EOS to determine the variations in thermodynamic properties. The release incident involves the decompression of liquified CO₂ in a pipe length of 100 m. The initial conditions are: pressure (P_0) = 7.27 MPa, temperature (T_0) = 303.5 K and velocity (v_0) = 0 m/s. The simulation sets out at $x = 50$ m, $t = 0$ s after the pipe end ruptures, and the liquid CO₂ depressurizes to a pressure of 3.67 MPa. As a result, a rarefaction wave travels through the liquid, causing its expansion to superheated liquid. Then, the evaporation wave moves through the superheated liquid, leading to its evaporation and generating an expanded two-phase mixture propagating toward the pipe's end with high velocity.

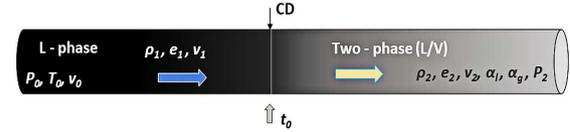
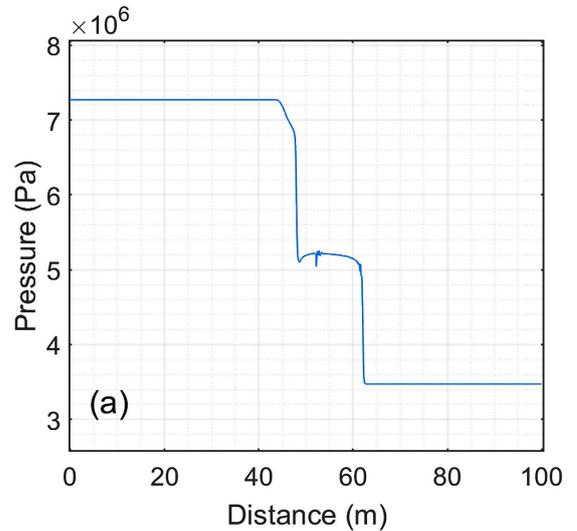


Figure 1: Schematic drawing shows CO₂ depressurization in a pipe.

Figure 2 shows the pressure, mixture velocity, and vapor/liquid volume fraction variations along the pipe's length. The plots demonstrate calculations for 1000 cells and 90 ms after the pipe ruptures. The passage of the rarefaction wave begins with the gradual liquid expansion corresponding to a pressure decrease from about 7.3 to 6.9 MPa and then a steep decline to almost 5.1 MPa (Figure 2(a)). The expanded superheated liquid remains in a metastable state at nearly constant pressure (slightly convex route), passing about 12.6 m before the pressure decreases sharply, indicating rapid evaporation.



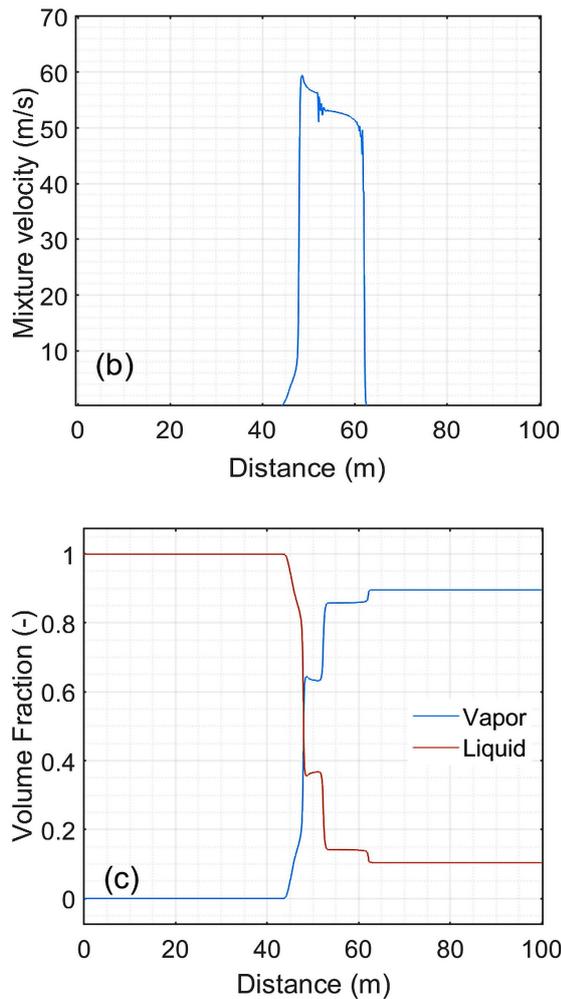


Figure 2: Change in pressure (a), mixture velocity (b), and vapor/liquid volume fraction (c) along the pipe's length during 90 ms of liquified CO₂ depressurization.

In addition, the rapid pressure drop results in the propagation of expanded fluid with increased velocity up to 59.3 m/s. Subsequently, the velocity decreases with small fluctuations before its abrupt drop once the evaporation wavefront has passed.

The vapor volume fraction (VVF) has risen in steps following the change in the thermodynamic state described by the pressure route. As seen in Figure 2 (c), complete evaporation has not been attained, and the third (last) jump in VVF seems to be related to the contact discontinuity wave.

Particular difficulties were encountered in the model implementation when depressurization was chosen to lower pressure near the atmospheric conditions. The temperature and subsequent calculated pressure showed unphysical behavior and may have been affected by the simplifying assumptions used in the model. In particular, the two-phase mixture velocity calculations can be vulnerable in contrast to the two-fluid model. The small rapid fluctuations seen in pressure and velocity profiles at about 52 and 61 m

(Figure 2 (a) and (b)) are likely caused by numerical computation of the temperature as a function of molar internal energy and volume. The intermittent fluctuations sites corresponded with the beginning of the rapid change in the internal energy (see Figure 4 (c)), which may produce transient instability in temperature calculations.

Computations have been implemented on grid cells between 200 and 2000 to evaluate the convergence of the CU-WENO numerical schemes. An appropriate numerical method should not create additional unreal oscillations as the cells' number increases. Figure 3 illustrates the convergence results for mixture density (a) and pressure (b) after 60 ms, implementing various cells' numbers on the CU-WENO schemes. As seen in Figure 3, the solution converges well as the grid is refined, i.e., as the size of the cells shrinks.

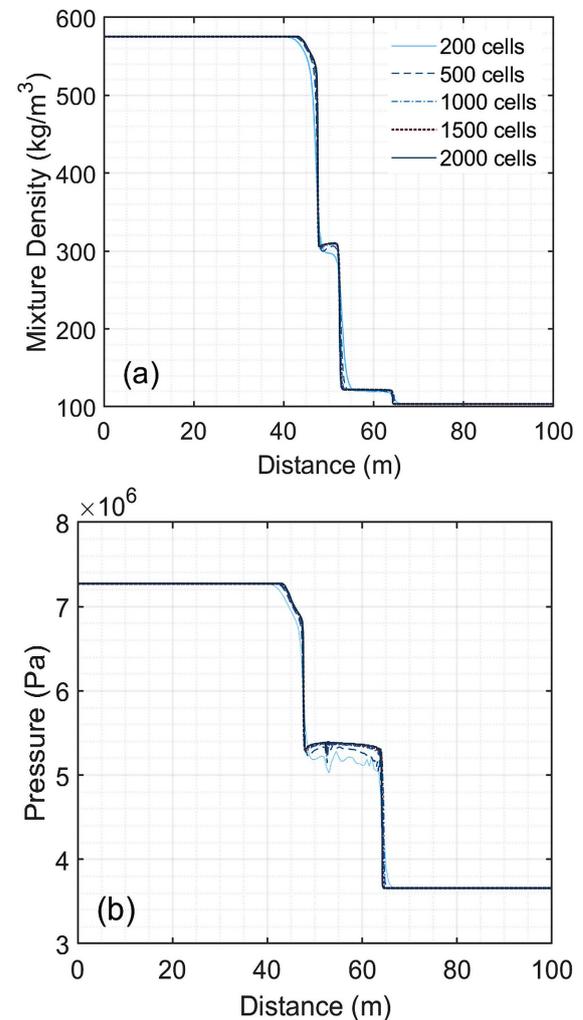


Figure 3: Grid refinement results for CU-WENO numerical schemes' convergence after 60 ms of CO₂ depressurization. (a) for mixture density, and (b) for the pressure.

There are no extra oscillations observed in either plot. Instead, the observed small fluctuations on the

pressure lines in Figure 3 (b) are flattened due to the grid refinement. The diffusivity of results reduces with grid refinement, and a grid size of 1500 can be recommended for the simulation.

Figure 4 shows vapor volume fraction (a), mixture density (b), and mixture internal energy (c) profiles along the pipe's length at different time stamps of 30, 50, 70, and 100 ms after the rupture. The waves' behavior can be described by tracing the changes in property profiles. The initial pressure gradient promotes the generation of three waves: shock and contact discontinuity, traveling in the right direction. Besides, an expansion wave moves in the opposite left direction.

In Figure 4 (a), the second and third sharp increases in VVF signifies the evaporation wavefront and the contact discontinuity, respectively. Additionally, the evaporation wavefront propagates towards the left as time runs, while the shock and contact surface propagate in the right direction.

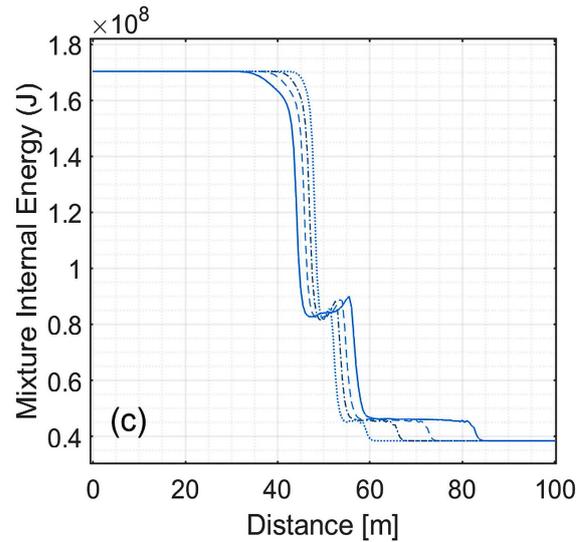
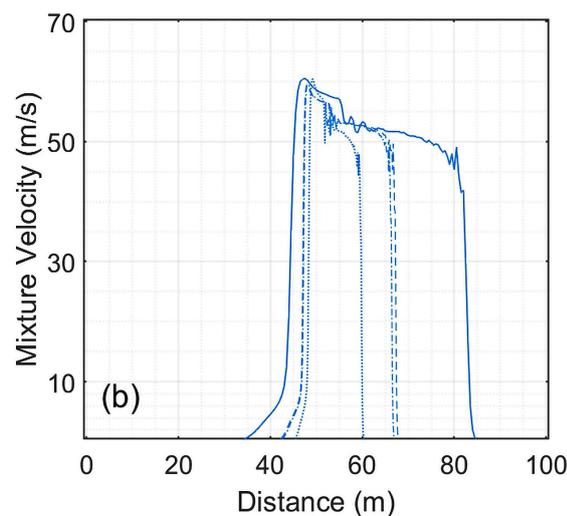
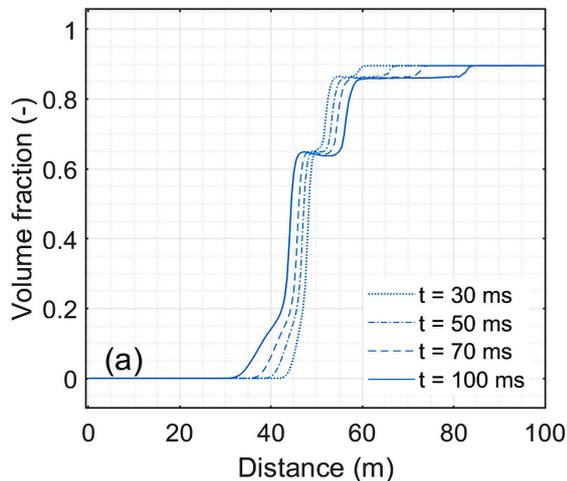


Figure 4: The behavior of vapor volume fraction (a), mixture density (b), and mixture internal energy (c) along the pipe's length during various time intervals and for 1000 grid cells.

The velocity profile in Figure 4 (b) demonstrates the two-phase mixture movement along the pipe towards the vented end. At the same time, the velocity peak increases while it moves in the reverse direction. However, the pressure wave behavior affects the two-phase mixture velocity after the evaporation starts due to the expansion waves' sequence propagation with different speeds (see Figure 4 (b) and Figure 2 (b)).

Figure 4 (c) shows that the mixture's internal energy rises in the metastable state following the rarefaction wave. This trend could be attributed to the liquid state transition from saturated to superheated, whereby the liquid stores the gained energy during its expansion as long as it remains in the metastable state. Then the energy is released during the phase change process providing the required latent heat of vaporization.

4. Conclusions

Then the energy A homogeneous equilibrium model (HEM) for two-phase flow during CO_2 depressurization in a pipe has been developed. The model considers equilibrium in pressure, temperature, and chemical potential, using SW EOS to calculate fluid properties in different thermodynamic states. Additionally, the method integrates two numerical schemes to use the CU-WENO scheme, wherein the numerical flux is calculated from central-upwind flux and the cell interface values from the third-order WENO reconstruction. The aim is to fit the central-upwind features in the WENO scheme, which can enhance accuracy and stability. The model is able to trace the

pressure, two-phase mixture density, and vapor volume fraction patterns during the rapid phase transition. In addition, the simulation results demonstrate the model's ability to predict the rarefaction and evaporation waves' dynamic characteristics and the convergence of the CU-WENO scheme. The numerical method well captured the moving discontinuity without blurs, and they have a similar pattern as in works done, for instance, by Munkejord et al. (2010) and Morin et al. (2010). However, the CU-WENO scheme illustrated the shock and expansion waves with less sharp edge transitions compared to the mentioned studies.

It is important to mention that the presented model does not consider the slip between phases during the compression and describes it as a uniform mixture pattern inside the pipe. However, the pipe rupture could result in a stratified flow pattern whereby the phase slip velocities should be considered. Additionally, the choice of initial pressure is dictated by the depressurization model's characteristics, as good predictions have been achieved close to the critical point. Further study may consider the mentioned remarks in developing pipe CO₂ accidental release models.

References

- Aursand E. *et al.* (2016) 'Fracture propagation control in CO₂ pipelines: Validation of a coupled fluid-structure model', *Engineering Structures*, vol. 123, pp. 192-212, doi: 10.1016/j.engstruct.2016.05.012.
- Botros K. *et al.* (2015) 'Measurements of Decompression Wave Speed in Pure Carbon Dioxide and Comparison With Predictions by Equation of State' *Journal of Pressure Vessel Technology*, vol. 138, no. 3, doi: 10.1115/1.4031941.
- Brown S. *et al.* (2013) 'A homogeneous relaxation flow model for the full bore rupture of dense phase CO₂ pipelines', *International Journal of Greenhouse Gas Control*, vol. 17, pp. 349-356, doi: 10.1016/j.ijggc.2013.05.020.
- Elshahomi A. *et al.* (2015) 'Decompression wave speed in CO₂ mixtures: CFD modelling with the GERG-2008 equation of state', *Applied Energy*, vol. 140, pp. 20-32, doi: 10.1016/j.apenergy.2014.11.054.
- Gottlieb S. and Shu C.W. (1998) 'Total variation diminishing Runge-Kutta schemes' *Mathematics of Computation*, vol. 67, no. 221, pp. 73-85, doi: 10.1090/S0025-5718-98-00913-2.
- Kurganov A., Noelle S., and Petrova G.(2001) 'Semidiscrete Central-Upwind Schemes for Hyperbolic Conservation Laws and Hamilton--Jacobi Equations SIAM' *Journal on Scientific Computing*, vol. 23, no. 3, pp. 707-740, doi: 10.1137/S1064827500373413.
- Mjaavatten A. (2022) 'thermodynamic models and tools for H₂O, H₂, CO₂, and Air', <https://github.com/aremj/thermodynamics/releases/tag/v2.1>, GitHub.
- Morin A. *et al.* (2010) 'Numerical Resolution of CO₂ Transport Dynamics', *Proceedings of the 2009 SIAM Conference on "Mathematics for Industry" The Art of Mathematics for Industry (MI)*, doi:10.1137/1.9781611973303.13
- Munkejord S. *et al.* (2010) 'Thermo- and fluid-dynamical

modelling of two-phase multi-component carbon dioxide mixtures' *International Journal of Greenhouse Gas Control*, vol. 4, no. 4, pp. 589-596, doi: 10.1016/j.ijggc.2010.02.003.

Munkejord S. T. and Hammer M. (2015) 'Depressurization of CO₂-rich mixtures in pipes: Two-phase flow modelling and comparison with experiments' *International Journal of Greenhouse Gas Control*, vol. 37, pp. 398-411, doi: 10.1016/j.ijggc.2015.03.029.

Shu C.-W. and Osher S. (1988) 'Efficient implementation of essentially non-oscillatory shock-capturing schemes' *Journal of Computational Physics*, vol. 77, no. 2, pp. 439-471, doi: 10.1016/0021-9991(88)90177-5.

Welahettige P. *et al.* (2022) 'Development of central-upwind-WENO scheme for two-phase 1-D drift flux model in pipe flow', Submitted to the 63rd Conference on Simulation and Modelling (63rd SIMS),.

Simulation of Flow in the Human Upper Airways Modeled as a Piping System Using the Hydraulic Diameter

Thor Gudmund Weisz ^{a,*}, Bernhard Müller ^b, Reidar Kristoffersen ^c,

^{a,b,c} Department of Energy and Process Engineering, NTNU, Trondheim, Norway

*thor.weisz@gmail.com

Abstract

Obstructive sleep apnea (OSA) is a medical condition characterized by repetitive obstructions in the human upper airways during sleep. Recent estimates from the United States show that the condition impacts 15% to 20% of the adult population. OSA treatment can be subdivided into surgical and non-surgical approaches. Non-surgical approaches such as continuous positive airway pressure (CPAP) devices have the highest success rates when used correctly. However, these approaches have low patient compliance due to the invasive nature of the devices during sleep, leaving surgery as a viable alternative for many. Predicting the outcome of OSA surgery is difficult due to the complex nature of both the airways and the surgeries themselves. CFD modeling of the airways is a helpful way to gain valuable insights into the flow structures and the impact of individual surgeries on the airways. However, CFD is not a viable approach for each patient-specific case due to its time-consuming nature. A pragmatic model has been created to predict the outcome of OSA surgery on a patient-specific basis to produce valid surgical estimates fast to be used by non-CFD engineers. The model transforms the human upper airways into a piping system by applying the hydraulic diameter equation on geometries created from CT scans. This paper aims to validate the use of the hydraulic diameter given by $D_h = 4 \cdot \frac{A}{P_e}$, where A is the cross-sectional area and P_e is the wetted perimeter, on the complex geometries of the nasal cavity and to provide a novel equation for the hydraulic diameter in the nasal cavity. The proposed hydraulic diameter equation is given by $D_h = C_{D_h} \cdot \frac{A}{P_e}$ where C_{D_h} is the hydraulic diameter coefficient. Airflow has been simulated through a simplified geometry using CFD to validate the hydraulic diameter and find an updated equation. Pragmatic model simulations using the hydraulic diameter have been compared to the results from CFD simulations to assess the pragmatic model's accuracy. The results showed that the original hydraulic diameter did not give entirely accurate results and that the novel equation using $C_{D_h} = 3.71$ gave the pragmatic model better accuracy for the validation cases. Tuning the parameter C_{D_h} for flow in an OSA patient's upper airways, the pragmatic model succeeded in quite accurately reproducing the area-averaged pressure in the patient's upper airways.

1. Introduction

Obstructive sleep apnea (OSA) is a sleep disorder characterized by repeated collapses of the upper airways during sleep. These collapses obstruct airflow, leading to loss of oxygen intake and a build-up of CO₂, which can cause daytime symptoms such as drowsiness or loss of functioning while seemingly getting enough sleep (Punjabi, 2008). It may cause a stroke or a heart attack in more severe cases. There are several surgical and non-surgical treatment options that help to alleviate or remove OSA entirely. The most common treatment is the non-surgical use of a sleeping mask which provides a constant pressurized flow through the airways, removing the possibility of airway collapse. Of these devices, the most common one is the continuous positive airway pressure device (CPAP) which is highly effective when properly used. However, patient compliance is a prominent issue with such devices (Sawyer et al., 2011). These factors leave surgical treatment as a viable option in many cases, although it is more invasive in the short term. Since OSA first was described in the middle of the 10th century (Gharibeh and Mehra, 2010), there have been many medical advances in surgical treatment options for OSA. Although significant advances have been made in OSA

treatment, the outcome of surgery is still not entirely predictable (Kezirian et al., 2004; Friedman et al., 2005). There have been marginal improvements in some cases, and in more severe cases, OSA has worsened after surgery.

1.1. Fluid Mechanics in OSA

Computational fluid dynamics (CFD) and widely available computed tomography (CT) image segmentation tools have made it possible to create patient-specific 3D models for airway analyses rapidly (Kim et al., 2013). Fig. 2 shows a 3D model of the human upper airways of an OSA patient prior to surgery. With appropriate verification and validation, CFD becomes a great and trustworthy tool that makes model testing much faster than its experimental counterpart. Even though computational capacity has seen exponential growth, detailed CFD simulations are still computationally expensive. Using CFD software to acquire accurate and trustworthy results requires an experienced engineer, making it a less viable choice for medical doctors to use in patient-specific cases. The pragmatic simulator developed through Weisz' specialization project (Weisz, 2021) and further improved through the same author's master's thesis (Weisz, 2022) is a proposed method for combining the accuracy of CFD

with low computational cost and user-friendliness.

1.2. Pragmatic Model

The pragmatic model is a proposed 1D flow simulator that takes in cross-sectional data from the upper airways of a patient with OSA and converts it into a piping system. Fig. 3 represents the upper airways of an OSA patient, while Fig. 4 is a visual representation of the upper airways as a piping system, used for the pragmatic simulations. The area-averaged gauge pressure is calculated using the Bernoulli equation with losses from cross-section to cross-section (Cengel and Cimbala, 2010). The geometrical variations in the geometry lead to additional pressure losses, which are modeled using known relations for pipes and included in the Bernoulli equation with losses. A doctor can perform the pragmatic calculations and the results can give insights into the current state of the patient's airways, and further help determine what type of surgery to perform.

1.3. Hydraulic Diameter

In the pragmatic model described, the unorthodox transformation of the human upper airway cross-sectional geometry to a pipe using the hydraulic diameter has been made. Fig. 1 shows a representation of this transformation for a cross-section in the nasal cavity which is further expanded to include the complete upper airways in Fig. 4. Since this is not a common approach, validation material is challenging to find. As the airway from the nasopharynx and down has a less complex shape and only one passage, the hydraulic diameter assumption is assumed to be accurate for this anatomical region. Although a validation using the entire geometry would be beneficial, it would yield patient-specific results and might not apply to all OSA patients. It would also be difficult to validate these results since experimental data for velocity and pressure in OSA patients' upper airways are unavailable. This led to the proposal of two simplified test cases used for simulations with the CFD tool Ansys Fluent (ANSYS, 2021a). The results led to a redefinition of the hydraulic diameter for this specific case after comparison with results from the pragmatic simulation.

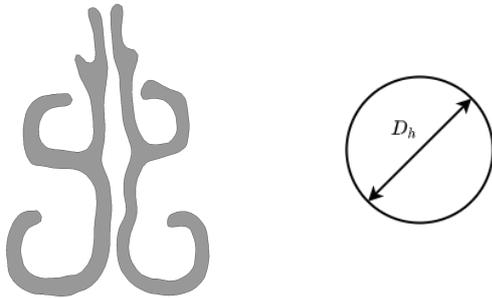


Figure 1: Representation of the conversion from a cross-section in the nasal cavity to a circle with the hydraulic diameter D_h .

2. Theoretical Background

In this paper, the flow simulations from the pragmatic model described in this section are validated using simulations with the commercial CFD software Ansys Fluent. The two simulators have different sets of governing equations solved through the simulations. Both sets of governing equations will be presented in this section of the paper.

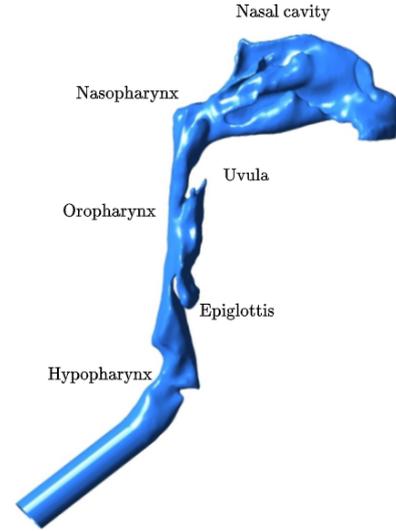


Figure 2: 3D model of the human upper airways adapted from Jordal's master's thesis (Jordal et al., 2017).

2.1. Governing Equations for the Pragmatic Model

The governing equation for the pragmatic model is the Bernoulli equation with losses used to calculate the pressure through the human upper airways. The Bernoulli equation with losses between arbitrary points 1 to 2 along a streamline is given as (Cengel and Cimbala, 2010):

$$\frac{p_1}{\rho g} + \alpha_1 \frac{V_1^2}{2g} + z_1 = \frac{p_2}{\rho g} + \alpha_2 \frac{V_2^2}{2g} + z_2 + h_L \quad (1)$$

where p_1 and p_2 are the pressures and V_1 and V_2 are the velocities at point 1 and 2, respectively. z_1 and z_2 correspond to the height of each point, while α_1 and α_2 are the kinetic energy correction factors. ρ is the density of the fluid and g is the gravitational acceleration. h_L is the loss term which accounts for irreversible losses in the equation. Through solving (1) for the pressure at point 2 and substituting in an index notation, the pressure is given as

$$p_i = p_{i-1} + \rho \frac{\alpha_{i-1} V_{i-1}^2 - \alpha_i V_i^2}{2} + \rho g (z_{i-1} - z_i) - \rho g h_{L,i}, \quad (2)$$

where $i-1$ and i refer to cross-sections along the same streamline. This pressure corresponds to the gauge pressure in the pragmatic model, as the reference pressure is assumed to be atmospheric.

2.2. Losses

The loss term in equation (1) is comprised of both minor and major losses, where major losses are frictional losses and minor losses are caused by losses from geometrical changes in a pipe (Cengel and Cimbala, 2010, p. 364). The minor losses in the pragmatic simulator are given by

$$h_{L,minor,i} = \sum_{j=1}^m K_{L,j} \frac{V_j^2}{2g}, \quad (3)$$

where j is a geometric component causing a minor loss in section i of the airways and $K_{L,j}$ is its minor loss coefficient. The major loss is given by

$$h_{L,major,i} = f_i \frac{L_i}{D_{h,i}} \frac{V_i^2}{2g}, \quad (4)$$

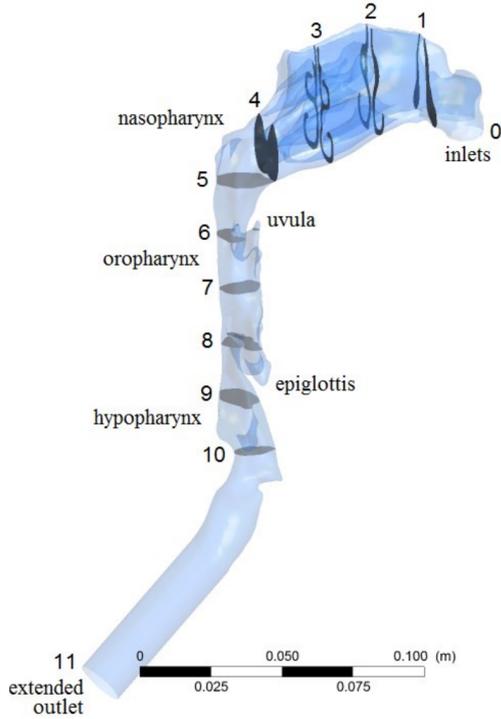


Figure 3: Locations of cutplanes used for the extraction of the area-averaged gauge pressure in the CFD investigation of the human upper airways (Aasgrav, 2017).

where i is the section of the pipe between cross-sections $i - 1$ and i . L_i is the length of the section, and $D_{h,i}$ is its hydraulic diameter.

2.3. Diffuser Effect

The minor losses in Eq. (3) include losses that occur due to gradual expansions. Results from the pragmatic simulations show that additional loss modeling may be required in these regions. These additional losses due to flow separation are modeled through a pressure-recovery coefficient given as (White, 2008, p. 398)

$$C_p = \frac{p_e - p_t}{p_{0t} - p_t}, \quad (5)$$

where p_e and p_t are the pressure at the exit and throat of the diverging nozzle respectively and p_{0t} is the stagnation pressure at the throat. To include this as an additional loss it can be solved for p_e after obtaining a value for C_p and adding it to Eq. (2).

2.4. Hydraulic Diameter

The hydraulic diameter is given by the equation

$$D_h = 4 \cdot \frac{A}{P_e}, \quad (6)$$

where A is the cross-sectional area of the geometry and P_e is the wetted perimeter. In the investigation, a variation of the hydraulic diameter is proposed, which is given by

$$D_h = C_{D_h} \cdot \frac{A}{P_e}, \quad (7)$$

where C_{D_h} is the "hydraulic diameter coefficient" which replaces the constant 4 in the original equation.

2.5. Volumetric Flow Rate

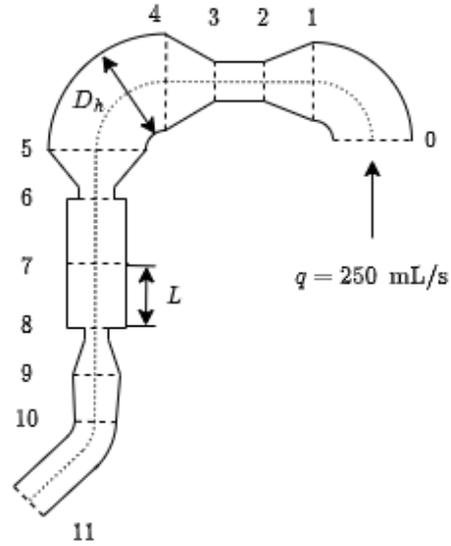


Figure 4: Piping representation of the human upper airways used for simulations using the pragmatic model. The numbered locations correspond to the cutplanes in Fig. 3.

The velocity is one of the input variables in Eq. (2) and therefore needs to be calculated. Since the flow rate is known, the velocity can be calculated using the incompressible volumetric flow rate

$$Q = V_i A_i = \text{constant}. \quad (8)$$

In this equation V_i is the velocity and A_i is the cross-sectional area of cross-section i .

2.6. Governing Equations for CFD

In the commercial CFD software Ansys Fluent (ANSYS, 2021a) the Navier Stokes equations are solved on a discretized mesh using the finite volume method. The continuity equation and the momentum equation, which make up the Navier-Stokes equations, are solved for each cell. Since the Mach number is much lower than 0.3, the incompressible variants of the equations have been used. The incompressible continuity equation is given by

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (9)$$

where u_i is the velocity component in the x_i - direction where $i = 1, 2, 3$. The incompressible momentum equation is given by

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = f_i - \frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j}. \quad (10)$$

f_i is the i^{th} component of the gravitational acceleration, where $i = 1, 2, 3$. ν is the kinematic viscosity, which is a constant for this case. Einstein summation is assumed for equations (9) and (10).

3. Methodology

As mentioned in the introduction, converting two ducts into one using the hydraulic diameter is not a common approach. To the knowledge of the authors, this conversion has not been used for flow calculation using the proposed method. Therefore, the procedure is explained in detail in the following section, along with justifications for the choices made.

3.1. Numerical Setup

A simple numerical case was chosen to validate the use of the hydraulic diameter on cross-sections from the human upper airways. For simplicity, a rectangular duct was chosen as the foundation of the simplified geometry. With a duct as the basis a wall was introduced to the geometry giving the duct a divided geometry, further mimicking the geometry in the nasal cavity. The separating wall is the simple geometry's counterpart to the septum, the cartilage which divides the nasal passage in the airways (Matthias, 2007). One of the ducts was created more narrow than the other because of the prevalence of deviated septums in OSA cases. A deviated septum is a deformation of the cartilage and bone wall separating the nasal passages, which impacts the many functions of the nasal cavity (Fettman et al., 2009). Therefore, the variation in geometry between the two passages in the numerical geometry was implemented to generalize the test case. Fig. 5 shows the setup and the dimensions of the different passages. The length of the wall was chosen to allow the flow to develop while not necessarily becoming fully developed, as the flow in the nasal cavity does not become fully developed due to the short entry length and its varying geometry. Two test cases were chosen since differences will occur in the various patient-specific upper airways. Both of the test cases have the same overall dimensions apart from the leading and trailing ends of the wall. One numerical case has a wedge at the leading and trailing ends of the wall, while the other has flat ends.

3.1.1. Wall with Wedged Ends

In the human upper airways, flow separation is likely to occur at various stages because of the complexity of the geometry. However, for the simple generalizable case presented in this paper, investigating a non-separated flow is of interest, possibly yielding a better base case for future comparison. To avoid flow separation leading into the region with two passages, a 10° wedge is placed in front of the wall. The walls on either side are flat, which effectively leads to an angle of 5° for either passage. For a circular diffuser, flow separation has been found to occur in the range $1000 \leq Re \leq 4000$ (Sparrow et al., 2009) giving a comparable case downstream of the wall. The Reynolds number of the flow in both of the single duct sections, prior to and post the separated passages, was $Re = 2500$. Although the test case is not circular, it was chosen as a case with a lower chance of separation than its wedge-less counterpart. The dimensions of the computational domain have been taken from the height and width of a nasal cavity cross-section at its largest point. The left figure in Fig. 1 shows the largest cross-section. The largest section of the airways was measured using the 3D geometry from a patient who showed great improvement in OSA post-surgery. The same geometry used (Aasgrav, 2017).

3.1.2. Wall with Flat Ends

An alternative numerical setup to the wedged setup presented in the section above was tested. This alternative setup was proposed to capture flow separation and recirculation which are likely to occur in the complex human upper airways (Martonen et al., 2002). The alternative setup chosen is identical to Fig. 5 apart from the leading and trailing ends of the separating wall, which in the alternative case are flat. The alternative setup is shown in Fig. 6. The flat leading and trailing ends of the separating wall are hypothesized to cause flow separation and recirculation, similar to the flow structures found in a backward-facing step (Lee and Mateescu, 1998). This is

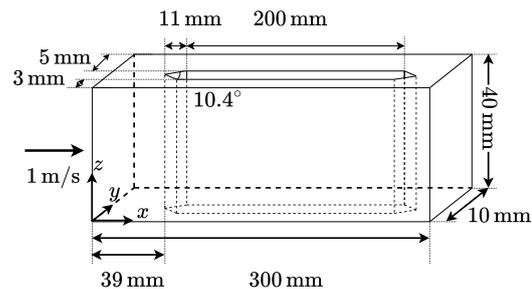


Figure 5: Numerical setup with 10° wedges at both ends of the separating wall.

meant to aid in the validation of the hydraulic diameter by providing more test data more closely resembling actual human upper airway flow structures.

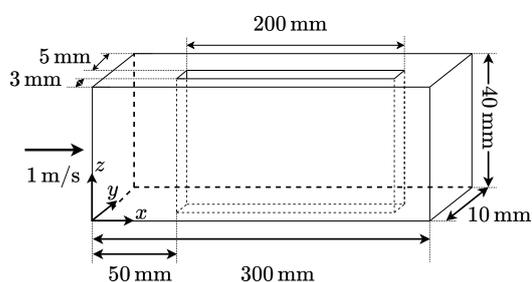


Figure 6: Numerical setup where the leading and trailing ends of the separating wall are flat.

3.2. Inlet and Outlet Boundary Conditions

For both test cases, the boundary conditions at the inlet and outlet, respectively, were the same. In both cases air with a kinematic viscosity of $\nu = 1.6 \cdot 10^{-5}$ was used as the fluid. A fully developed laminar flow was given as the inlet condition. The fully developed flow was found by extending the numerical domain prior to the inlet by an entry length of 800 mm, with a uniform flow of $V = 1$ m/s at the inlet of the extended domain. The entry length was found using the equation for a nondimensional hydrodynamic entry length for a laminar flow (Cengel and Cimbala, 2010, p. 342). The inlet velocity was found by calculating the Reynolds number using the actual human upper airway geometry and is based on a flow rate of 250 mL/s (Aasgrav, 2017). The outlet condition specified the gauge pressure and is set to 0 Pa at the end of the flow domain. The outer and internal separating walls were treated with no-slip boundary conditions.

3.3. Grid Generation

The mesh for both cases was created using Ansys Meshing (ANSYS, 2021b), a part of the Ansys simulation environment, where all of the simulations were carried out. With the help of the mesh creator and Ansys Workbench (ANSYS, 2021c), several different grid configurations were tested to find a grid that would provide a grid-independent solution. The result from the pragmatic model simulations was the area-averaged gauge pressure. This was a natural quantity to check when performing the mesh independence study. The grid was created by predetermining the number of subdivisions along the outer horizontal edges, the horizontal edges along the separating wall, and the outer vertical edges. To create the mesh,

the number of division along the outer edges and along the wall's edges in the x -direction was set to 300, thus giving the numerical domain 300 cells in the x -direction. 8 cells were used in the y -direction and 40 cells were used in the z -direction, both created by setting the number of divisions along the outer edges in their respective directions to 8 and 40 respectively. In Ansys Meshing (ANSYS, 2021b) the behaviour of these sections was set to "hard" to create a structured mesh mostly containing hexahedral cells. This meshing scheme made it possible to generate meshes at different scales with the same proportions in a controllable way. Different resolutions for the mesh were created and simulated while controlling the selected parameter, the gauge pressure. The meshes with a slightly coarser resolution gave similar pressure values for the control plane. Therefore, the mesh with a higher number of cells was chosen for further simulations since the control parameter remained reasonably constant. The mesh independence study was only performed on the numerical domain with a 10° wedge. The same grid generation technique and resolution were applied to the case with flat wall ends.

3.4. Tuning the Hydraulic Diameter

Altering the hydraulic diameter given by Eq. (6) is proposed to match the results from the CFD investigations more accurately. A novel method to tune this parameter was to define a hydraulic diameter coefficient, C_{D_h} . The hydraulic diameter coefficient replaces the constant, 4, in Eq. (6) to obtain a new equation given by (7) allowing the possibility to find a coefficient that better represents the present case. The method involved running the CFD simulations described in this paper along with the pragmatic simulations and comparing the results from both. The pragmatic simulations' input were the flow rate, cross-sectional area and the wetted perimeter from evenly spaced cross-sections along the numerical geometries. The pragmatic model was modified only to include the frictional losses introduced through the major loss term in Eq. (4) when validating the hydraulic diameter. The hydraulic diameter is present in both the Reynolds number and the loss term itself, leading to a negative squared inverse correlation between the pressure and the hydraulic diameter. This correlation implies that a decrease in the hydraulic diameter leads to a decrease in the pressure slope. Pragmatic simulations were run for hydraulic diameter coefficients in the range $3 \leq C_{D_h} \leq 4.2$ to compare the CFD simulations and the pragmatic simulations. The residual sum of squares (RSS) (Wikipedia contributors, 2022) was calculated for each pragmatic simulation. The equation for RSS is given by

$$RSS = \sum_{i=1}^n (y_i - f(x_i))^2, \quad (11)$$

where y_i is the i^{th} component of the value to be predicted (the area-averaged gauge pressure from CFD simulation) and $f(x_i)$ is the i^{th} component of the predicted value (the area-averaged gauge pressure from the pragmatic model). The RSS value was calculated using pragmatic and CFD gauge pressures from the middle of the wall to the end of the walled section, i.e. $150 \text{ mm} \leq x \leq 200 \text{ mm}$ in Fig. 5 and Fig. 6. This was done to reduce the impact of inaccuracies in the error estimation in the entrance regions of the split geometries. In these regions, deviations are expected due to the simple nature of the pragmatic simulations. The hydraulic diameter coefficients could then be obtained by minimizing the error in this region.

3.5. Minor Losses and Diffuser Effects

With the pragmatic model using loss relations known from piping systems, the accuracy of these relations is important. An effect seen in the results of the pragmatic simulations is a nonphysical pressure recovery where the human upper airway geometry has an expansion. The relations for expansions used in the pragmatic model did not accurately model the pressure difference seen in the CFD simulations (Weisz, 2022). Therefore, it was proposed that the effects of flow separation were more significant than initially thought. To further investigate this increased flow separation and model it accurately, the effects were included in the validation simulations presented in the present paper. This was implemented at the end of the walled section, where the nasal passages coincide. The Bernoulli equation with losses Eq. (1) takes flow separation into account though minor losses given by Eq. (3). However, the pressure recovery found through the pragmatic flow simulations was too large. This occurs when the two passages in the geometry coincide downstream of the separating wall and there is an increase in the cross-sectional area for both passages. This effect can be taken into account through added diffuser losses for these sections. These losses were calculated using Eq. (5) by using the CFD gauge pressure to calculate the pressure recovery coefficient C_p , using Eq. (5). The exit pressure, p_e , in Eq. (5) was the CFD gauge pressure from the cross-section downstream of the expansion, and the throat pressure, p_t , in Eq. (5) was the gauge pressure from the cross-section upstream of the same expansion. The calculated C_p values for the expanding sections were included in the pragmatic model by solving Eq. (5) for the exit pressure p_e using the area-averaged gauge pressure from the pragmatic model, p_{i-1} , as the throat pressure, p_t . This was used as p_i in the pragmatic model instead of using Eq. (2) to calculate p_i for this section, leading to a larger, more physically accurate pressure loss in the relevant region.

3.6. Verification of the Numerical Code

A crucial part of any CFD simulation is verifying the accuracy of the code used for simulations. In the present case, this step was done through a simplification of the numerical domain. The 3D domain was simplified to a 2D domain with the same height as the 3D domain, 40 mm, and long enough for the flow to become fully developed. The fully developed flow was found by using a domain which was 7000 mm in length, giving the flow the opportunity to become fully developed. The grid for the verification case was created using Ansys Meshing (ANSYS, 2021b), where a structured grid with rectangular cells was created. The mesh had 1000 cells in the flow direction, where a bias which decreased the cell size from the inlet to the outlet was included to achieve similar cell dimensions as in the 3D case towards the end of the domain. 40 cells were used in the y -direction with even spacings. The inlet velocity was chosen to be $V = 1 \text{ m/s}$ to achieve a Reynolds number of $Re = 2500$ using the height of the domain as the length scale and $\nu = 1.6 \cdot 10^{-5}$ as the kinematic viscosity. The simulations were carried out using Ansys Fluent (ANSYS, 2021a), using the same settings as the full 3D simulations. The results from the verification were compared to the analytical solution of a plane Poiseuille flow (Cengel and Cimbala, 2010, p. 468). Fig. 7 compares the two solutions and shows a high degree of accuracy in the numerical simulations compared to the analytical solution. Fig. 7 also shows the development of the flow, at the locations $x = 0.05, 2$ and 6.9 m , where x is

the downstream distance from the inlet located at $x = 0$ m. The flow profiles show an expected development with the flow reaching its fully developed state at $x = 6.9$ m. Since these results correspond to their analytical counterpart, the solver was considered an accurate enough standard for further numerical investigation.

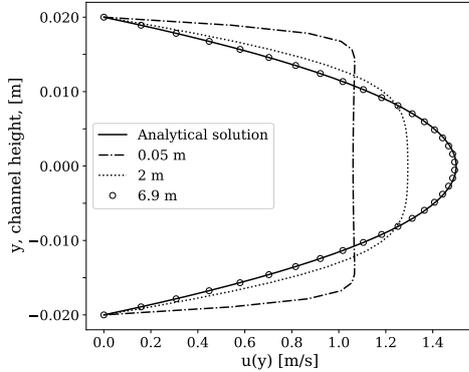


Figure 7: Fluid flow profiles at the locations 0.05, 2 and 6.9 m downstream from the inlet. The flow profile reaches the analytical solution close to the end of the simulation domain, $x = 6.9$ m. The analytical solution is given for a fully developed plane Poiseuille flow.

3.7. Pragmatic Simulations

The pragmatic simulations were run using input data found using the numerical geometries. The area-averaged gauge pressure was extracted when the CFD simulations were completed. The pressure was extracted at evenly spaced cross-sections along the length of the numerical domain. The pragmatic model's input for each location is the cross-sectional area, the perimeter and the flow rate for the given simulation. The cross-sectional area and the perimeter of each cross-section were calculated based on the numerical geometries' dimensions shown in Fig. 5 and Fig. 6, the extracted values are shown in Tab. 1. The flow rate was calculated using the area-averaged input velocity $V = 1$ m/s. The flow rate could then be converted into the velocity at the various cross-sections in the wall-separated region using equation (8). The pragmatic simulations were then carried out and the area-averaged gauge pressure from both the pragmatic simulations and the CFD simulations could be compared.

Table 1: Cross-sectional area and perimeter calculated using dimensions from Fig. 5 and Fig. 6, where the dimensions at $x = 45$ mm only apply to the wedged case.

x [mm]	A [mm ²]	P_e [mm]
0	400	100
45	356.36	177.82
150	320	176

4. Results and Discussion

The results from the validation simulations are to be applied to the pragmatic model to accurately represent the flow through predicting the area-averaged gauge pressure. The area-averaged gauge pressure from the pragmatic simulations and the CFD simulations are compared to assess the accuracy of the standard hydraulic diameter Eq. (6) and find an accurate fit for the hydraulic diameter coefficient C_{D_h} in Eq. (7). Fig. 8 and 9

show the results from both simulations along with the initial pragmatic results using Eq. (6), which are the uppermost curves. The initial results revealed deviations from the CFD simulations, most notably the slope of the area-averaged gauge pressure. The slope of the pragmatic curve was initially too flat and was altered through varying the hydraulic diameter coefficient C_{D_h} . In the Bernoulli equation with losses (1) the hydraulic diameter D_h is inversely correlated to the pressure through its representation in the loss term h_L , given by Eq. (4). However, as the term is negatively signed, the pressure gradient and the hydraulic diameter become correlated, leading to predictable changes when tuning the coefficient C_{D_h} . Changing the hydraulic diameter coefficient and introducing losses from the pressure-recovery coefficient gave more accurate results when compared to the CFD simulations. The effects on the pragmatic model from both of the numerical validation cases are presented and discussed in the following subsections.

4.1. Wall with Wedged Ends

Fig. 8 shows the results from both the CFD simulations and the pragmatic simulations in the case where the wedge is present. The results show an expected decline in the area-averaged gauge pressure through the geometry. The pragmatic simulations have a linearly decreasing pressure in the middle section, where the separating wall is located, but with varying slopes. The CFD simulation has a less linear shape in the section with the separating wall. However, it has a linear trend further downstream. Compared to the CFD simulations, it has a slight additional pressure loss as the duct is split up, and a slight pressure gain at the opposite side. The initial pragmatic results have similar trends but with inflated loss values as well as pressure gain. Using the error minimization approach described in the methodology section, $C_{D_h} = 3.73$ was found as the optimal hydraulic diameter coefficient. A visualization of the optimization is shown in Fig. 10. The pressure-recovery coefficient was calculated by applying pressure values from the CFD simulations to Eq. (5). This showed an improvement in accuracy downstream of the walled section.

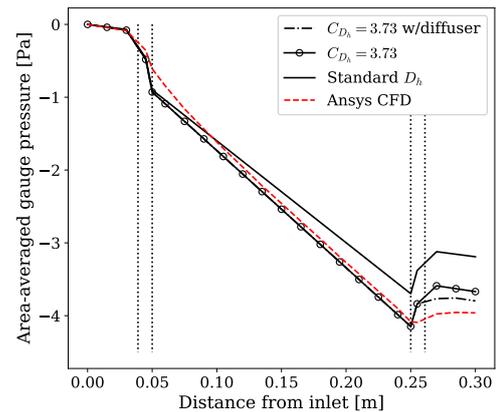


Figure 8: Results from Ansys Fluent and corresponding pragmatic simulations for simulations with a 10° wedge at both ends of the separating wall. The vertical dotted lines indicate the wedge locations.

4.2. Wall with Flat Ends

The alternative flow situation used to study the hydraulic diameter is also investigated. For this case, the separating

wall has flat ends instead of wedges. The aim of this was to study the effects of tuning the hydraulic diameter coefficient, C_{D_h} , and comparing it to the wedged case to find out how the hydraulic diameter coefficient would deviate from the initial results. This was to study the reliability and generalizability of the results obtained in the wedged case. The results from these simulations are displayed in Fig. 9. Minimizing the RSS value for these simulations gave $C_{D_h} = 3.69$ as the hydraulic diameter coefficient with the highest accuracy. Similarly to the case with 10° wedges, applying the increased pressure recovery coefficient to the pragmatic simulations gave a result closer to the CFD simulations. Due to the nature of the geometry with its sudden geometrical changes, the pragmatic model has sudden pressure changes at the beginning of the walled section and at the end, with a pressure loss and a pressure gain respectively. The same procedure for finding the pressure recovery coefficient was used. In the case with flat ends, this gave a highly accurate result. This implies that there is more flow separation in the non-wedged case. However, this increased accuracy may be due to the sudden changes, which make the pressure recovery coefficient from the CFD simulations easier to find. The increased accuracy using the pressure recovery coefficient in the wedge versus the non-wedged case is thus inconclusive. However, the use of a pressure recovery coefficient in general is promising.

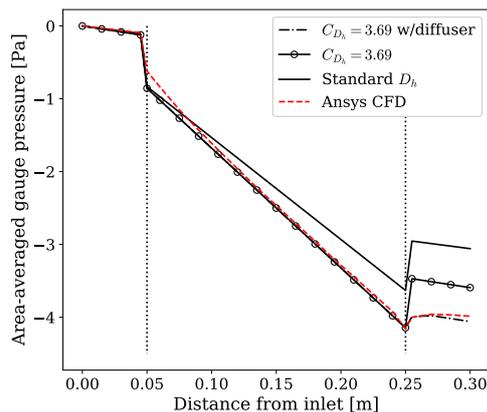


Figure 9: Results from Ansys Fluent and corresponding pragmatic simulations for the simulations with flat ends of the separating wall.

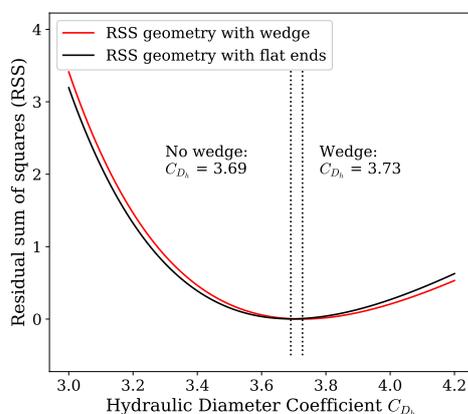


Figure 10: Calculations of the RSS values for simulations with and without a 10° wedge at the ends of the wall.

4.3. OSA Patient's Upper Airways Simulations

The aim of finding an improved version of the hydraulic diameter is to implement the improved version in the pragmatic model and achieve better results when simulating the flow the actual upper airways of an OSA patient. The results of implementing the averaged value of the two hydraulic diameter coefficients into the pragmatic model are shown in Fig. 11. In this figure the cutplanes correspond to the cutplanes from Fig. 3, which indicate the locations the results from the CFD simulations are taken from. The cross-sectional area and perimeter used as input for the pragmatic model have been extracted from the same 3D model at the numbered locations. The results with $C_{D_h} = 3.71$ in Eq. (7) only show a marginal improvement compared to using Eq. (6). The alternative hydraulic diameter coefficient, C_{D_h} , was applied to the first four cutplanes prior to the coinciding of the nasal passages indicated by the dashed and dotted line in Fig. 11, the standard hydraulic diameter $C_{D_h} = 4$ was used downstream of this. Further analyses showed that a more accurate coefficient for the human upper airways is given by $C_{D_h} = 1.80$. One of the reasons for this deviation is that the simple nature of the geometries analysed in this paper are better modeled by the a value closer to the original hydraulic diameter. The hydraulic diameter is meant for square ducts and other simple geometries, thus increasing the complexity of the geometry requires decreasing the hydraulic diameter.

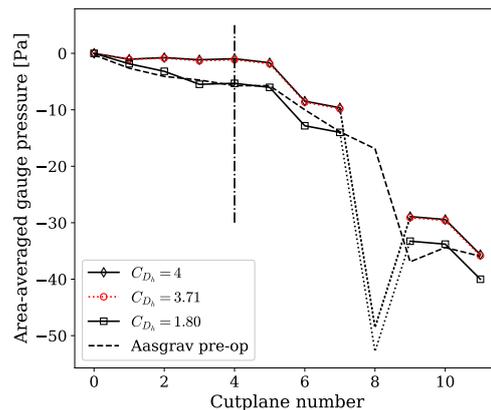


Figure 11: Pragmatic model simulations with cross-sections from Fig. 2 as input verified using CFD simulation performed through Aasgrav's specialization project (Aasgrav, 2017). The dashed and dotted line indicates where the nasal passages coincide.

5. Conclusions

As a part of creating a pragmatic flow simulator, the hydraulic diameter is used to transform the human upper airway geometry into a piping system. A test case has been created and presented in this paper to determine if this is a valid assumption to make in the nasal cavity. The test case used two variations of a simple duct-like geometry with a single duct to begin with, which goes over to a split up section with a separating wall between two passages and a coinciding geometry at the end of the wall. The simulations were performed assuming a laminar steady incompressible airflow that with a uniform velocity of 1 m/s at the inlet of the numerical domain. The results from these simulations showed the expected decline in the area-averaged gauge pressure over the length of the geometry. The resulting pressure curve was used as a basis for comparison to the pragmatic flow simulations. Pragmatic flow simulations were executed, with the only

loss contribution being the frictional losses. Comparing the results to the CFD simulations proved that the pressure slope generated by the pragmatic simulator was a little off. Analysing different values of the hydraulic diameter coefficient C_{D_h} for the two test cases presented gave two slightly different values. The two values were $C_{D_h} = 3.72$ and $C_{D_h} = 3.69$ for the case with and without a wedge at the leading and trailing ends of the wall respectively. Averaging the two values and implementing the new equation is given by $D_h = 3.71 \cdot \frac{A}{P_e}$ into the pragmatic model hardly improved its accuracy for flow in the upper airways of an OSA patient. However, $C_{D_h} = 1.80$ in the pragmatic model proved to give good agreement of the pressure with the CFD results. This investigation proved that the accuracy of the pragmatic simulations could be improved by altering the hydraulic diameter coefficient C_{D_h} . To conclude, the coefficients found in this study did not provide accurate results on the actual human upper airway, but imply that altering the hydraulic diameter can lead better accuracy. This implies that the redefinition of the hydraulic diameter proposed in this study can be used as an optimization parameter in the pragmatic model.

References

- Aasgrav, E. (2017), CFD Simulations of Turbulent Flow in the Human Upper Airways, master's thesis, NTNU.
- ANSYS (2021a), *Ansys® Fluent, 2021R2, User's Guide*.
- ANSYS (2021b), *Ansys® Meshing, 2021R2, ANSYS Meshing User's Guide*.
- ANSYS (2021c), *Ansys® Workbench, 2021R2, User's Guide*.
- Cengel, Y. A. and Cimbala, J. M. (2010), *Fluid Mechanics: Fundamentals and Applications, 2nd ed.*, McGraw Hill.
- Fettman, N., Sanford, T. and Sindwani, R. (2009), 'Surgical management of the deviated septum: techniques in septoplasty', *Otolaryngologic Clinics of North America* **42**(2), 241–252.
- Friedman, M., Vidyasagar, R., Bliznikas, D. and Joseph, N. (2005), 'Does severity of obstructive sleep apnea/hypopnea syndrome predict uvulopalatopharyngoplasty outcome?', *The Laryngoscope* **115**(12), 2109–2113.
- Gharibeh, T. and Mehra, R. (2010), 'Obstructive sleep apnea syndrome: natural history, diagnosis, and emerging treatment options', *Nature and Science of Sleep* **2**, 233.
- Jordal, M. R., Johnsen, S. G., Dahl, S. K. and Müller, B. (2017), 'Patient Specific Numerical Simulation of Flow in the Human Upper Airways for Assessing the Effect of Nasal Surgery', *Progress in Applied CFD - CFD - 2017* pp. 153–162.
- Kezirian, E. J., Weaver, E. M., Yueh, B., Deyo, R. A., Khuri, S. F., Daley, J. and Henderson, W. (2004), 'Incidence of serious complications after uvulopalatopharyngoplasty', *The Laryngoscope* **114**(3), 450–453.
- Kim, S. K., Na, Y., Kim, J.-I. and Chung, S.-K. (2013), 'Patient specific CFD models of nasal airflow: overview of methods and challenges', *Journal of Biomechanics* **46**(2), 299–306.
- Lee, T. and Mateescu, D. (1998), 'Experimental and numerical investigation of 2D backward-facing step flow', *Journal of Fluids and Structures* **12**(6), 703–716.
- Martonen, T. B., Quan, L., Zhang, Z. and Musante, C. (2002), 'Flow simulation in the human upper respiratory tract', *Cell Biochemistry and Biophysics* **37**(1), 27–36.
- Matthias, C. (2007), 'Surgery of the nasal septum and turbinates', *GMS Current Topics in Otorhinolaryngology, Head and Neck Surgery* **6**.
- Punjabi, N. M. (2008), 'The epidemiology of adult obstructive sleep apnea', *Proceedings of the American Thoracic Society* **5**(2), 136–143.
- Sawyer, A. M., Gooneratne, N. S., Marcus, C. L., Ofer, D., Richards, K. C. and Weaver, T. E. (2011), 'A systematic review of CPAP adherence across age groups: clinical and empiric insights for developing CPAP adherence interventions', *Sleep Medicine Reviews* **15**(6), 343–356.
- Sparrow, E., Abraham, J. and Minkowycz, W. (2009), 'Flow separation in a diverging conical duct: Effect of Reynolds number and divergence angle', *International Journal of Heat and Mass Transfer* **52**(13-14), 3079–3083.
- Weisz, T. G. (2021), Pragmatic modeling of flow in the human upper airways, Specialization project, NTNU.
- Weisz, T. G. (2022), Pragmatic modeling of flow in the human upper airways for sleep apnea treatment, Master's thesis, NTNU.
- White, F. M. (2008), *Fluid Mechanics, 6th ed.*, McGraw Hill.
- Wikipedia contributors (2022), 'Residual sum of squares — Wikipedia, the free encyclopedia'. [Online; accessed 29.05.2022].
- URL:** https://en.wikipedia.org/wiki/Residual_sum_of_squares

Steam reforming of methane over a nickel-based catalyst

Rakhi ^{a,*}, Vivien Günther ^b, Jana Richter ^a, Fabian Mauss ^a

^a Brandenburg University of Technology, 03046 Cottbus, Germany,

^b LOGE AB, 03044 Cottbus, Germany

* rakhi.rakhi@b-tu.de

Abstract

Steam reforming is a promising route to convert natural gas into syngas - a mixture of H₂ and CO, used as a feed stock e.g. for ammonia, methanol and Fischer-Tropsch synthesis processes. For the industrial application of steam reforming, a detailed understanding of the process is a prerequisite. Models that capture the detailed homogeneous and heterogeneous reaction kinetics and the comprehensive transport processes as well as their interaction have the potential to optimize the catalytic process without expensive experimental campaigns. In the present work, a one-dimensional model, LOGEcat is used to carry out a detailed investigation considering a multi-step reaction mechanism for modeling steam reforming of methane over nickel-based catalyst. The model is computationally cost effective due to the reduction in dimensionality, in contrast to experimental investigations which are not always feasible or 2D/3D simulations which are computationally expensive. The 1D tool is based on a series of perfectly stirred reactors (PSR) and is applicable to the simulation of all standard after-treatment catalytic processes of combustion exhaust gas as well as other chemical processes involving heterogeneous catalysis such as the Sabatier process. We have applied the model to perform the simulations for various reactor conditions in terms of parameters such as temperature, pressure, velocity and steam-to-carbon (S/C) ratio. Several chemical reaction terms have been analyzed and the results are compared with 2D simulation and experimental reference data. We note a very good agreement of the various profiles produced with the cost-effective reduced order model in comparison to the reference data.

Keywords: Nickel-based catalyst; One-dimensional modeling; Methane; Steam reforming;

1. Introduction

Steam reforming of hydrocarbons is a crucial chemical process [1, 2] providing synthesis gas (H₂ and CO) which plays a key role as a feedstock in many catalytic processes, for example, synthesis of methanol, oxo-synthesis, and Fischer-Tropsch synthesis [3]. The produced synthesis gas, hydrogen, is used in the manufacturing of ammonia [3]. The most prominent and widely used industrial steam reforming process is the methane or gas (natural) reforming. This is one of the most efficient technologies for hydrogen and synthesis gas production from fossil fuels in large scale facilities reaching yields close to the thermodynamic equilibrium [2, 3]. Thus, conventional steam reformers deliver high concentrations of hydrogen at high fuel conversion rates [4], however, this process is disadvantageous in small scale operation units because of the highly endothermic reactions and the requirement of efficient external energy supply.

Thermodynamics control the products of the reaction and favour the formation of methane at lower temperature along with hydrogen at higher ones. Steam reforming of methane accompanied by water-gas shift reactions on a Ni/MgAl₂O₄ catalyst by intrinsic rate equations derived from a Langmuir-Hinshelwood mechanism is described in [5]. Recently, a catalytic sequence for reactions of CH₄ with CO₂ and H₂O on Ni/MgO catalysts has been postulated in [6] and a microkinetic model for steam reforming reactions over a Ni/MgAl₂O₄ catalyst has been proposed in [7] by reactions for CO₂ reforming of

methane and deactivation by carbon formation.

Due to the potential to reduce the cost of synthesis gas production and environmental concerns, partial oxidation over noble metal catalysts [8–11] as well as CO₂ reforming [12–14] of natural gas to synthesis gas have attained much interest. The sequence and interaction of the reaction routes have been considered in several investigations in order to understand the reaction mechanism of synthesis gas formation from methane. Earlier, a direct catalytic partial oxidation route has been followed [11], however, in later studies the overall conversion is realized in a two-step process (indirect route) [8–10, 15]. In [11, 15–17], steps for steam reforming for the catalytic partial oxidation of methane over platinum and rhodium are published. The reaction kinetics of methane steam reforming over nickel catalyst has been extensively investigated experimentally and theoretically in [18, 19].

All the investigations discussed above are either performed experimentally (not always feasible) or using 2D/3D tools which are computationally expensive specially when the full reaction mechanism is included. The computational cost increases drastically with increasing number of species and alternatives need to be explored in order to capture the flow physics and chemistry at a reduced cost. A good strategy is to reduce the reaction mechanism or dimensionality from 2D/3D to 1D which are time efficient alternatives and are suitably applicable to catalyst simulations.

In this paper, selecting the latter approach (reduction in dimensions) a one-dimensional model which is discussed in next section, has been elaborated to test the steam reforming of methane over nickel/alumina monoliths in the temperature interval of 900-1350 K. The results are compared with the data available in the literature, for the 2D simulations as well as experiments. Note that this work forms the basis for further detailed investigations presented in our recent study [20].

2. Model Description

The one-dimensional model, LOGecat [21] is used for the simulations presented in this paper. The model is based on the single-channel 1D catalyst model and is applicable to the simulation of all standard after-treatment catalytic processes of combustion exhaust gas, for example, three-way catalyst (TWC), diesel oxidation catalysts (DOC), NO_x storage and reduction (NSR) catalysts and selective catalytic reduction (SCR) catalysts.

The model has been successfully applied and tested in previous studies [20, 22]. However, a detailed investigation of the steam reforming of methane over nickel has not been done in any of the previous studies. So, here we aim to analyze the case considered in a general and detailed way in order to check the predictability of model and to know how far the one-dimensional model can capture the flow physics and the chemistry involved with it.

Now, we discuss the modeling of the conservation and flow equations for the readability of this paper. A number of representative channels are used to model the three-way catalytic converter. The gas and surface properties are calculated as a function of axial distance in these channels and also, the heat conduction between the channels is modeled. In contrast, no radial heat conduction between the channels is considered in case of a single-channel (discussed below) set-up representing the whole catalyst using one channel.

2.1. Single Channel Model

Figure 1 shows the single channel which is divided into a finite number of cells with Δx as their length. Each cell is treated as a perfectly stirred reactor (PSR) and the pressure gradient along with inhomogeneity of the mixture can be neglected as the diameter of the catalytic channel is small. A thin layer represented by a separate pore gas zone close to the wall is used to model the external diffusion and this pore layer is depicted by the area between the bulk gas and the washcoat as shown in the figure.

The model used to carry out the simulations is the part of the LOGEsoft software suite [21] for chemical reaction calculation. The conservation equations (next section) for gas species mass fraction, gas enthalpy, surface temperature, pore layer gas species mass fraction, and surface site fractions are solved in each PSR. These equations are solved for each time step and additionally, the 1D Navier-Stokes equations for pressure as well as flow velocity are solved over all cells by an operator splitting method.

2.1.1. Conservation Equations

Assuming constant pressure during the time step Δt in the operator splitting loop, the bulk gas in each cell is modeled by a PSR. The mass transfer coefficient accounts for the species transport between bulk gas and pore volume layer and the conservation equation for bulk gas species mass fraction is given as,

$$\begin{aligned} \rho \frac{\partial Y_{i,g}}{\partial t} = & \frac{(\rho v)_{in} A}{V_g} (Y_{i,in} - Y_{i,g}) + W_i \omega_{i,g} \\ & - \frac{P \Delta x}{V_g} W_i K_m k_{m,i} (C_{i,g} - C_{i,p}) \\ & + Y_{i,g} \frac{P \Delta x}{V_g} \sum_{j=1}^{N_g} W_j K_m k_{m,j} (C_{j,g} - C_{j,p}) \end{aligned} \quad (1)$$

The subscript g denotes the bulk gas, in the inflow from the upstream cell and p gas in the pore layer. $Y_{i,g}$ is the mass fraction of species i , V_g is the gas volume in the current cell, $w_{i,g}$ is the species source term for gas phase reaction, K_m is the tunable parameter for the overall mass transfer, $k_{m,i}$ is the conservation mass transfer coefficient of species i , $C_{i,g}$ is the concentration of species i in the bulk gas, and $C_{i,p}$ is the concentration of species i in the pore layer and P is the geometric wetted perimeter of the channel. For more details we refer the reader to [21].

The pore volume layer considers gas phase as well as surface reactions which further includes diffusion into the pores and the conservation equation of the gas phase species is given as,

$$\begin{aligned} \rho_{p,l} \frac{\partial Y_{i,p,l}}{\partial t} = & \frac{P \Delta x}{V_{p,l}} W_i K_m k_{m,i} (C_{i,g} \\ & - C_{i,p,l})|_{l=1} + W_i \omega_{i,p,l} \\ & + W_i K_e \sum_{m=1}^{N_{surf}} \frac{A_m}{V_{p,l}} \omega_{i,m} \\ & + D_i \frac{C_{i,p,l} - C_{i,p,l+1}}{W_{l+1,l}} \frac{A_m}{V_{p,l}} \\ & - Y_{i,p,l} \sum_{j=1}^{N_g} \left[\frac{P \Delta x}{V_{p,l}} K_m k_{m,j} W_j (C_{j,g} \right. \\ & - C_{j,p,l})|_{l=1} + W_j K_e \sum_{m=1}^{N_{surf}} \frac{A_m}{V_{p,l}} \omega_{j,m} \\ & \left. + D_i \frac{C_{j,p,l} - C_{j,p,l+1}}{W_{l+1,l}} \frac{A_m}{V_{p,l}} \right] \end{aligned} \quad (2)$$

In the above equation, $V_{p,l}$ is the gas volume of the pore volume layer in washcoat layer l in the current cell, $w_{i,p,l}$ is the species source term for gas phase reactions in the pore layer in washcoat layer l , N_{surf} is the number of different surface materials present in the catalytic converter (usually 1), A_m is the catalytic surface area in the current cell, k_e is the tuning parameter for the overall reaction efficiency. The parameter D_i accounts for an additional term for diffusion through multiple washcoat layers and is appropriate diffusion coefficient for species i , the subscript l is the current washcoat layer, $w_{l+1,l}$ is the radial distance through the washcoat calculated as $(w_{l+1} - w_l)/2$ for diffusion between washcoat layer l and $l + 1$. Note that the source term for bulk gas species transport into the washcoat is only used for the first washcoat layer (denoted as $|_{l=1}$).

Next, the conservation equation for surface species site fraction is given as,

$$\frac{\partial \theta_{i,n}}{\partial t} = \sigma_{i,n} K_e \frac{\omega_{i,n}}{\tau_n} \quad (3)$$

$\Theta_{i,n}$ indicates the site fraction of species i at site n , τ_n is the site density, $w_{i,n}$ is the species source term from

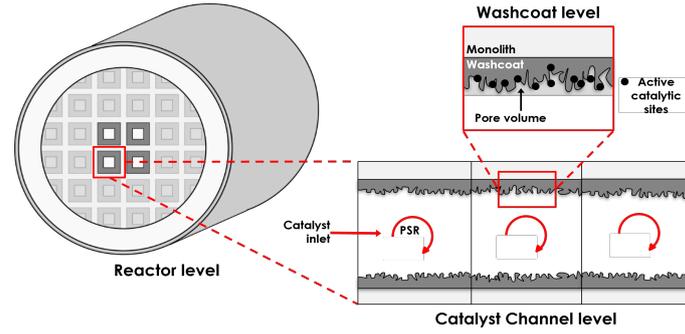


Figure 1: Schematic illustration of the modeling approach

reactions at site n and $\sigma_{i,n}$ is the site occupancy number of species i at site n .

The heat transport by convection and molecular transport is taken into account by the bulk energy (specific enthalpy) conservation equation given as,

$$\begin{aligned} \frac{\partial h_g}{\partial t} = & -K_h h_T \frac{P \Delta x}{m_g} (T_g - T_w) + \frac{(\rho v)_{in} A}{m_g} (h_{in} - h_g) \\ & + \frac{P \Delta x}{m_g} \sum_{j=1}^{N_g} W_j K_m k_{m,j} (C_{j,g} \\ & - C_{j,p}) (h_g - h_{j,g \leftrightarrow p}) \end{aligned} \quad (4)$$

In equation 4, h_g is the bulk gas specific enthalpy, h_T is the convective heat transfer coefficient between bulk gas and surface, T_g is the bulk gas temperature, T_w the pore layer temperature, h_{in} is the specific enthalpy of the gas from the upstream cell and $h_{j,g \rightarrow p}$ the specific enthalpy of species j transported between the bulk gas and the pore layer. The bulk gas enthalpy is used in case of the species being transported from the bulk gas and pore layer enthalpy is used if it is transported to the bulk gas.

The pore layer temperature is assumed to be homogeneous for the substrate as well as for the gas and the pressure is assumed constant in the pore layer. The kinetic energy due to gas movement is neglected. Considering these assumptions, the conservation equation for the surface temperature is given as,

$$\begin{aligned} & (\rho_s V_s c_{p,s} + \rho_p V_p c_{p,p}) \frac{\partial T_w}{\partial t} \\ = & \\ & V_s \frac{\partial}{\partial x} \left(k_s \frac{\partial T_w}{\partial x} \right) + K_h h_T P \Delta x (T_g - T_w) \\ & - \sum_{j=1}^{N_{g,p}} W_j P \Delta x K_m k_{m,j} (C_{j,g} \\ & - C_{j,p}) (h_{j,p} - h_{j,g \leftrightarrow p}) \\ & - \sum_{j=1}^{N_{g,p}} W_j \left[h_{j,p} V_p K_e \omega_{j,p} + h_{j,p} K_e \sum_{m=1}^{N_m} A_m \omega_{j,m} \right] \\ & - K_e \sum_{m=1}^{N_m} \sum_{j=1}^{N_{s,m}} h_{j,m} W_j A_m \omega_{j,m} \\ & + \frac{k_{s,l+1} \left(\frac{\partial T_{w,l+1}}{\partial w_{l+1,l-1}} \right) - k_{s,l-1} \left(\frac{\partial T_{w,l-1}}{\partial w_{l+1,l-1}} \right)}{w_l} \end{aligned} \quad (5)$$

where, V_s is the volume of the solid wall material (washcoat and substrate) in the current cell, $C_{p,s}$ is the specific heat capacity of the solid material at constant pressure, $C_{p,p}$ is the specific heat capacity in the pore volume layer at constant pressure and K_h is a tunable parameter for the heat transfer. Hence, the above equation accounts for heat conduction along the channel, heat convection/ diffusion to the bulk gas, molecular heat transport as well as heat released by reactions. $k_{s,l}$ is the thermal conductivity of washcoat layer l and for single washcoat pore diffusion is mimicked by the tunable parameter k_e . The washcoat diffusion for the surface temperature is also included in case if there are multiple washcoats. Additionally, heat flow term is used to account for heat losses through the material and the catalyst at the periphery of the substrate.

The heat and mass transfer coefficients, h_T and $k_{m,i}$, used in the conservation equations are calculated from the Nusselt and Sherwood numbers [23]. For simultaneously developing velocity, concentration and thermal boundary layer flow, the correlations for Sherwood and Nusselt numbers are used as [24],

$$Sh_i(x) = \begin{cases} \frac{0.35}{Sc_i^{1/6}} \sqrt{\left(\frac{d_h}{4}\right)^2 v}, & 0 < x < \frac{\left(\frac{d_h}{4}\right)^2 v}{D_i} \left(\frac{1}{Sc_i}\right)^{1/3} \left(\frac{1.4}{Sh_{T,\infty}}\right)^2 \\ Sh_{T,\infty}, & x \geq \frac{\left(\frac{d_h}{4}\right)^2 v}{D_i} \left(\frac{1}{Sc_i}\right)^{1/3} \left(\frac{1.4}{Sh_{T,\infty}}\right)^2 \end{cases} \quad (6)$$

$$Nu(x) = \begin{cases} \frac{0.35}{Pr^{1/6}} \sqrt{\left(\frac{d_h}{4}\right)^2 v}, & 0 < x < \frac{\left(\frac{d_h}{4}\right)^2 v}{D_T} \left(\frac{1}{Pr}\right)^{1/3} \left(\frac{1.4}{Nu_{T,\infty}}\right)^2 \\ Nu_{T,\infty}, & x \geq \frac{\left(\frac{d_h}{4}\right)^2 v}{D_T} \left(\frac{1}{Pr}\right)^{1/3} \left(\frac{1.4}{Nu_{T,\infty}}\right)^2 \end{cases} \quad (7)$$

Here, D_i is the species diffusion coefficient for species i , D_T is the thermal diffusion coefficient, v is the fluid velocity along the channel and $Sh_{T,\infty}$ and $Nu_{T,\infty}$ are the asymptotic Sherwood and Nusselt numbers, respectively, for constant flux boundary conditions (their values are taken from [25]). The Schmidt number for species i , Sc_i and the Prandtl number, Pr are given as,

$$Sc_i = \frac{\mu_i}{\rho D_i} \quad (8)$$

$$Pr = \frac{c_p \mu}{k_g} \quad (9)$$

where, μ_i is the dynamic viscosity, D_i is the diffusion coefficient of species i , C_p is the heat capacity at constant pressure, μ is the dynamic viscosity and k_g is the thermal conductivity of the gas.

2.1.2. Flow equations

With the assumption that the flow is in steady state, the conservation equations are given as,

$$\frac{\partial(\rho v)}{\partial x} = -\frac{A_{w,G}}{m_g} \sum_{i=1}^{N_g} W_i k_{m,i} (C_{i,g} - C_{i,p}) \quad (10)$$

$$\frac{\partial(\rho v^2)}{\partial x} + \frac{\partial p}{\partial x} = -\frac{f_F}{2} (\rho v) \left| v \right| \frac{p}{A} \quad (11)$$

as mass and momentum equations, respectively. In the above equations, A is the cross-sectional channel area. The friction factor, f_F for laminar and fully developed flow is given as,

$$f_F = \frac{16}{Re} = \frac{16\mu}{\rho v d_h} \quad (12)$$

For more details related to the model and the derivations for the equations, we refer the reader to [21].

3. Surface reaction mechanism

We have used the reaction mechanism from [3] which contains 6 gas-phase and 13 surface species in total. The reaction mechanism consists of 42 reactions. The reaction mechanism indicates that adsorbed carbon species (CH, CH₂, CH₃ etc) formed from activated methane reacts with adsorbed atomic oxygen O(s), formed from the adsorption of oxygen or from the decomposition of water and CO₂, and produce carbon oxide. The mechanism also comprises the reactions of partial oxidation and steam reforming of methane and is based on the key reaction intermediate - adsorbed atomic oxygen O(s). For the details of the reaction mechanism, we refer the reader to [3].

Note that the sticking coefficients are used as kinetic data for the adsorption of reactants and products (H₂, O₂, CH₄, H₂O, CO, CO₂) given in the reaction mechanism.

3.1. Thermodynamic Consistency

The equilibrium of a chemical reaction is given as,

$$\sum_i v'_{ik} A_i \rightleftharpoons \sum_i v''_{ik} A_i \quad (13)$$

and is defined by the thermodynamic properties of the participating species. In terms of equilibrium constant, K_{pk} , the equilibrium activities, a_i , obey the expression,

$$K_{pk} = \prod_i (a_i^{eq})^{v_{ik}} = \exp\left(-\frac{\Delta_k G^0}{RT}\right) \quad (14)$$

$v_{ik} = v_{ik}' - v_{ik}''$ is the stoichiometric coefficient, R gas constant and T temperature. The change of free enthalpy at normal pressure p^0 is given as,

$$\Delta_k G^0 = \sum_i v_{ik} G_i^0(T) \quad (15)$$

The activation can be approximated by their partial pressure in case of gases and by coverage in case of surface species. Considering the dependence of the heat capacity on temperature by a forth-order polynomial and standard

enthalpies and entropies of formation, in that case the standard free enthalpies can be expressed in terms of seven coefficients,

$$G_i^0(T) = a_{0,i} + a_{1,i}T + a_{2,i}T^2 + a_{3,i}T^3 + a_{4,i}T^4 + a_{5,i}T^5 + a_{6,i}T \ln T \quad (16)$$

The rate coefficients for the forward and the reverse reaction must obey the below equation to predict the equilibrium correctly,

$$\frac{k_{fk}}{k_{rk}} = K_{pk} \prod_i (c_i^0)^{v_{ik}} \quad (17)$$

where, c_i are reference concentrations at normal pressure. Nonetheless, the forward and the reverse reactions are defined separately with their own rate laws due to the problems encountered in setting up a reaction mechanism. It is difficult to define the thermodynamic data for intermediate surface species. The thermodynamic consistency is ensured in a sense that the thermodynamic equilibrium of the participating gas-phase species is matched for a range of temperatures, while writing all reversible reactions as pairs of independent forward and backward reactions. The thermodynamic data of the intermediate species is, therefore not needed for the evaluation of the reaction rates [3]. We do not aim to cover the more details in this section and hence leave this to the readers interest. Note that the reaction mechanism and the thermodynamic data for all the species used in the present study are taken from [26].

4. Simulation Set-up

The simulation set-up used for the present study follows from [3]. The investigation is extended for different reactor conditions in terms of parameters, such as, temperature, S/C ratio, flow rate and pressure. Analysis is done for four different temperatures, T=920, 1020, 1120, and 1220 K while keeping all other parameters (S/C, \dot{f} and P) constant. Similarly, the S/C ratio is varied as S/C=1.9, 2.77, and 3.67, flow rate as \dot{f} =296, 593 and 1186 mL/min and pressure as P=1, 10 and 100 atm with other parameters fixed.

We have considered a single channel being 3.0×10^{-2} m in length with a catalyst radius of 7.5×10^{-3} m (640 cpsi) and it is uniformly divided into 25 cells. One layer of washcoat is used for the simulations. The overall heat transfer efficiency factor, mass transfer efficiency factor and efficiency factors for surface chemistry are taken as unity. The surface site density, τ for Ni is 2.6×10^{-5} mol/m² [2]. The surface area per catalyst length is used as 6.9×10^{-2} m²/m. 75% Argon dilution is used.

5. Results

The model explained above is used to perform the simulations of steam reforming of methane over a nickel catalyst and all the kinetic parameters are taken from [3]. Some of the important terms encountered in chemical reaction engineering are conversion, selectivity and yield which are discussed in the following section. It is important to check these quantities to observe if the system is consistent. The variation of these quantities are shown with different temperature. The conversion describes the ratio of how much of a reactant has reacted and lies between zero and one. The yield shows the formation of a desired product and it also falls in between zero and one. The selectivity defines the ratio of the desired product to

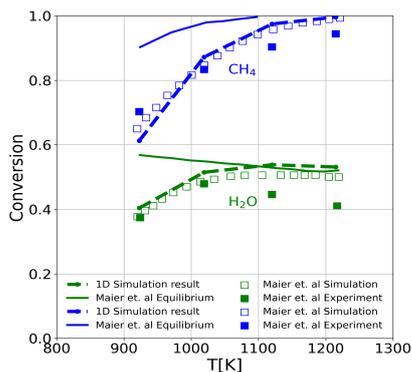


Figure 2: Methane and water conversion as a function of temperature for $S/C=2.77$ and 75% Ar along with the reference data.

the undesired products. The formulas for these quantities are given where they are first discussed in the section.

Figure 2 shows the conversion of methane and water as a function of temperature along with the reference data. The conversion is calculated as, $X_i = (y_{i,o} - y_{i,e})/y_{i,o}$. The simulations are performed for fixed S/C ratio as 2.77 and 75% Argon dilution. The 1D simulation results using LOGEcat model are in very good agreement with the experimental and simulation results from [3]. It is observed that the thermodynamic equilibrium is attained at the higher temperatures. However, at higher temperature the simulation results, both from LOGEcat model as well as 2D reference data, deviates from the experimental data.

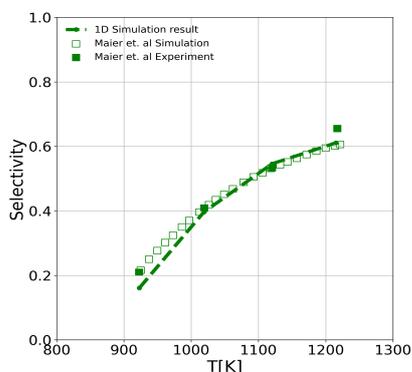


Figure 3: CO Selectivity variation with temperature in methane steam reforming for $S/C=2.77$ and 75% Ar along with the reference data.

The CO selectivity variation with temperature in methane steam reforming for fixed S/C ratio is shown in Figure 3 along with the reference simulation and experimental data. The selectivity for CO is calculated as, $S_{CO} = X_{CO}/(X_{CO} + X_{CO_2} + X_{CH_4})$ and we observe a very good agreement for the 1D LOGEcat model results with [3].

Figure 4 illustrates the H_2/CO ratio variation with the temperature in methane steam reforming for S/C ratio 2.77 and 75% Argon along with the reference results. As explained in [3], the over-prediction of the H_2/CO ratio in comparison to the experimental measurements at the given S/C ratio might be due to the underestimation of water-gas shift reaction at low temperature in the 2D as well as 1D model. The H_2/CO ratio for simulation, Maier's 2D simulation as well as results from 1D model, is higher

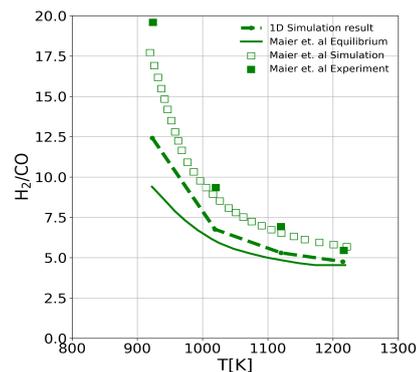


Figure 4: H_2/CO ratio variation with temperature in methane steam reforming for $S/C=2.77$ and 75% Ar along with the reference data.

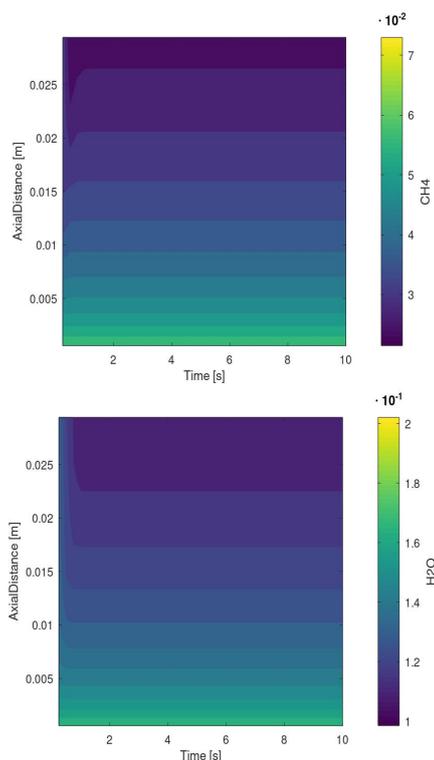


Figure 5: (a) Methane and (b) water concentration along the reactor for $T=920$ K.

compared to the equilibrium calculation at temperature ≤ 1000 K. The 1D results lie in between the experimental and simulation results from [3]. Nevertheless, it is worth noticing that the 1D model capture this profile, qualitatively as well as quantitatively very well.

Further, the variation of concentration for reactants and products along with the axial distance/ length of the reactor is shown in Figure 5 and 6. The simulation results are shown for $T=920$ K. These figures show that the reactants (Figure 5), methane and water, are being used in the first few seconds, i.e., within 2s. Then the thermal equilibrium is reached and no change in the concentration can be observed after 2s but the simulations do run a little longer in order to make sure that steady state is reached. Similarly, the formation of products (Figure 6), H_2 , CO , CO_2 , takes place within first few seconds and then it

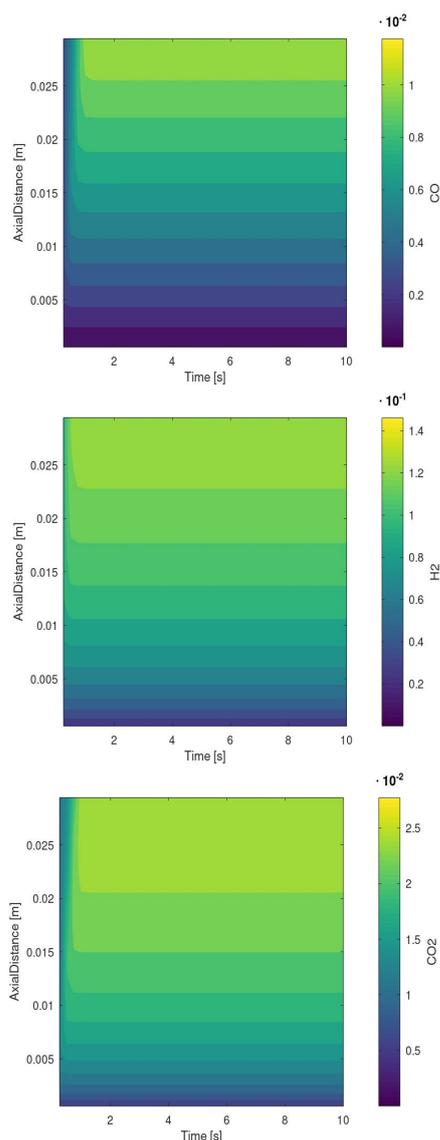


Figure 6: (a) CO, (b) H₂ and (c) CO₂ concentration along the reactor for T=920 K.

ceases. We expect this behaviour as the endothermic reactions are dominant on the catalytic surface in the initial phase which cause the major changes in concentration of different species only in the beginning and then the system attains thermal equilibrium. However, if we observe these concentration plots at higher temperature (not shown in this paper due to limited space), it takes slightly longer to reach steady state and hence, it depends on the temperature.

In Figure 7, the methane and water conversion is shown as a function of temperature for fixed S/C ratio (S/C=3) along with the reference data. The figure captures the conversion behaviour for a wide range of temperature, $T \in [600, 1300]$ in order to check the predictability of the 1D model for higher temperature. The 1D simulation results are in good agreement with the reference 2D simulations and experiments for the entire temperature range. The reference data is available only for methane conversion, however, we have also shown the water conversion for the given temperature range for our simulation set-up.

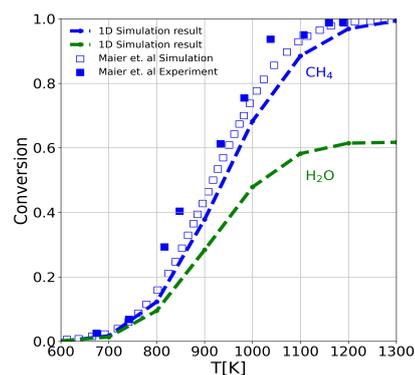


Figure 7: Methane and water conversion as a function of temperature for S/C=3 along with the reference data.

6. Conclusions

This paper presents the kinetics of the steam reforming of methane over nickel catalyst using a one-dimensional tool, LOGEcat. The results are compared with literature [3] and the investigation is carried out for different temperatures. Various quantities, such as, selectivity, yield and conversion has been discussed.

The results show that the conversion, selectivity, H₂/CO ratio for temperature $\in [920, 1020, 1120, 1220]$ K are in very good agreement with the reference data considered for the comparison purpose. This proves the capability of the model to capture the basic flow physics and the chemistry and hence, the model can be used in multiple directions for further investigations at a reduced computational cost.

Acknowledgment

The authors thank the Graduate Research School (GRS) of the BTU Cottbus-Senftenberg for the financial support.

References

- [1] B. T. Schädel, M. Duisberg, and O. Deutschmann, "Steam reforming of methane, ethane, propane, butane, and natural gas over a rhodium-based catalyst," *Catalysis Today*, vol. 142, pp. 42–51, 2009.
- [2] J. R. Rostrup-Nielsen, In: *Anderson JR, Boudart M (eds) Catalytic steam reforming in catalysis—science and technology*. Berlin: Springer-Verlag, 1984.
- [3] L. Maier, B. Schädel, K. H. Delgado, S. Tischer, and O. Deutschmann, "Steam Reforming of Methane Over Nickel: Development of a Multi-Step Surface Reaction Mechanism," *Topics in Catalysis*, vol. 54, pp. 845–858, 2011.
- [4] B. C. Michael, A. Donazzi, and L. D. Schmidt, "Effects of H₂O and CO₂ addition in catalytic partial oxidation of methane on Rh," *Journal of Catalysis*, vol. 265, pp. 117–129, 2009.
- [5] R. Quiceno, O. Deutschmann, J. Warnatz, and J. Pérez-Ramírez, "Modelling of the high-temperature catalytic partial oxidation of methane over platinum gauze. Detailed gas-phase and surface chemistries coupled with 3D flow field simulations," *Applied Catalysis A*, vol. 303, pp. 166–176, 2006.
- [6] J. Wei and E. Iglesia, "Isotopic and kinetic assessment of the mechanism of reactions of CH₄ with CO₂ or H₂O to form synthesis gas and carbon on nickel catalysts," *Journal of Catalysis*, vol. 224, pp. 370–383, 2004.
- [7] L. M. Aparicio, "Transient Isotopic Studies and Microkinetic Modeling of Methane Reforming over Nickel Catalysts," *Journal of Catalysis*, vol. 165, pp. 262–274, 1997.
- [8] D. Dissanayake, M. P. Rosynek, K. C. C. Kharas, and J. H. Lunsford, "Partial Oxidation of Methane to Carbon

- Monoxide and Hydrogen over a Ni/Al₂O₃ Catalyst,” *Catalysis Today.*, vol. 132, pp. 117–127, 1991.
- [9] W. J. M. Vermeiren, E. Blomsma, and P. A. Jacobs, “Catalytic and thermodynamic approach of the oxyreforming reaction of methane,” *Catalysis Today.*, vol. 13, pp. 427–436, 1992.
- [10] S. Hannemann, J. D. Grunwaldt, N. van Vegten, A. Baiker, P. Boye, and C. G. Schroer, “Distinct Spatial Changes of the Catalyst Structure inside a Fixed-Bed Microreactor during the Partial Oxidation of Methane over Rh/Al₂O₃,” *Catalysis Today.*, vol. 126, p. 54, 2007.
- [11] D. A. Hickman and L. D. Schmidt, “Steps in CH₄ oxidation on Pt and Rh surfaces: High-temperature reactor simulations,” *American Institute of Chemical Engineers AIChE.*, vol. 39, pp. 1164–1177, 1993.
- [12] A. M. Gadalla and M. E. Sommer, “Carbon dioxide reforming of methane on nickel catalysts,” *Chemical Engineering Science.*, vol. 44, pp. 2825–2829, 1989.
- [13] Z. W. Liu, H. S. Roh, and K. W. Jun, “Important factors on carbon dioxide reforming of methane over nickel-based catalysts,” *Journal of Industrial and Engineering Chemistry.*, vol. 9, pp. 753–761, 2003.
- [14] M. C. J. Bradford and M. A. Vannice, “Catalytic reforming of methane with carbon dioxide over nickel catalysts II. Reaction kinetics,” *Applied Catalysis A: General.*, vol. 142, pp. 97–122, 1996.
- [15] R. Schwiedernoch, S. Tischer, C. Correa, and O. Deutschmann, “Experimental and Numerical Study of the Transient Behavior of a Catalytic Partial Oxidation Monolith,” *Chemical Engineering Science.*, vol. 58(3), pp. 633–642, 2003.
- [16] O. Deutschmann and L. Schmidt, “Modeling the partial oxidation of methane in a short-contact-time reactor,” *American Institute of Chemical Engineers AIChE.*, vol. 44, pp. 2465–2477, 1998.
- [17] A. B. Mhadeshwar and D. G. J. Vlachos, “Hierarchical Multiscale Mechanism Development for Methane Partial Oxidation and Reforming and for Thermal Decomposition of Oxygenates on Rh,” *The Journal of Physical Chemistry: B.*, vol. 109(35), pp. 16819–16835, 2005.
- [18] J. Xu and G. F. Froment, “Methane steam reforming, methanation and water-gas shift: I. Intrinsic kinetics,” *American Institute of Chemical Engineers AIChE.*, vol. 35, pp. 88–96, 1989.
- [19] J. R. Rostrup-Nielsen and J. H. B. Hansen, “CO₂-Reforming of Methane over Transition Metals,” *Journal of Catalysis.*, vol. 144, pp. 38–49, 1993.
- [20] Rakhi., V. Günther, J. Richter, and F. Mauss, “Steam reforming of methane over nickel catalyst using a one-dimensional model,” *International Journal of Environmental Science.*, vol. 5(1), pp. 1–32, 2022.
- [21] “LOGEsoft v1.10.” www.logesoft.com, 2008.
- [22] K. Fröjd and F. Mauss, “A Three-Parameter Transient 1D Catalyst Model,” *SAE International Journal of Engines.*, vol. 4 (1), pp. 1747–1763, 2011.
- [23] F. Incopera, D. DeWitt, T. Bergman, and S. Lavine, “Fundamentals of Heat and Mass Transfer 6E,” 2006.
- [24] K. Ramanathan, V. Balakotaiah, and D. West, “Light-off criterion and transient analysis of catalytic monoliths,” *Chemical Engineering Science.*, vol. 58, pp. 1381–1405, 2003.
- [25] H. Santos and M. Costa, “Modelling transport phenomena and chemical reactions in automotive three-way catalytic converters,” *Chemical Engineering Journal.*, vol. 148, pp. 173–183, 2009.
- [26] D. Schmider, L. Maier, and O. Deutschmann, “Reaction Kinetics of CO and CO₂ Methanation over nickel,” *Engineering Chemistry Research.*, vol. 60, pp. 5792–5805, 2021.

CFD modeling of the transport of human respiratory droplets in an indoor environment

Ole Martinius Harket Norbeck ^{a,*}, Oda Martine Sundsdal ^a, Suresh Kumar Nambully ^b, Arnab Chaudhuri ^a

^a *Department of Civil Engineering and Energy Technology, OsloMet – Oslo Metropolitan University, Oslo, Norway*

^b *Convergent Science GmbH, Linz*

*olenorbeck@hotmail.com

Abstract

For the last couple of years, the world has faced the global pandemic COVID-19. The viral transmission could occur via different modes like large respiratory droplets, direct contact with contaminated surfaces and airborne microdroplets or aerosol. This work revisits and focuses on human cough, and breathing sequence together with cough in confined spaces. We consider the Eulerian dispersion medium as a multicomponent ideal gas mixture consisting of oxygen, nitrogen and water vapor and the Lagrangian dispersed phase of human cough/breath is modeled as pure liquid water. The unsteady complex flow is resolved with an advanced three-dimensional multiphase flow solver utilizing adaptive mesh refinement (AMR). A simplified rectangular block with a rectangular mouth area is considered to mimic human beings to inject exhaled gas and liquid droplets associated with cough and or breathing instances. The evaporation model is switched off for the particles of diameter less than $5 \mu\text{m}$ to resolve the dynamics of the airborne particles. The results clearly demonstrate the efficacy of the novel approach toward gaining more knowledge about viral transmission in indoor environments.

1. Introduction

We are all affected by the ongoing global pandemic, COVID 19 (alternatively called SARS-CoV-2). The serious repercussions following the pandemic have caused communities all over the world to shut down and create a new way of living. The pandemic has highlighted the importance of mitigation procedures concerning the viral spread from humans. Since the beginning of the pandemic, scientists all around the world have shifted their priorities, resulting in a significant surge in research within this field.

Viral transmission takes place in a variety of ways, including inhalation of very fine airborne droplets and aerosols, through droplets and particles depositing on exposed mucous membranes in the mouth, nose, and eyes, and by being in contact with surfaces containing infectious virus. During breathing, talking, singing, coughing, or sneezing, an infected person exhales particles that could be contagious to others. Apart from experimental approaches, Computational fluid dynamics (CFD) has proven to be an effective tool for detailed realistic analysis and eventually plausible design of mitigation measures for indoor environments. One needs to carefully consider several parameters like the local airflow, source proximity, droplet dispersion/evaporation, air change per hour and relative humidity to understand the physics and design such preventive measures. However, resolving the 3D complex flow dynamics of the involved physics is computationally challenging with a high CPU cost. Since the beginning of the pandemic many research articles reported CFD-based analysis of human sneezing [1–3], coughing, breathing, talking [2, 4–12] in various social setups to address the plausible mechanism of

viral transmission, social distancing, risk assessments and measures of mitigation. CFD studies in this context mostly used Eulerian-Lagrangian numerical approach to resolve the multiphase flow physics of the governing equations. In a very recent review article [13], an insightful overview of CFD approaches related to respiratory events in buildings towards modeling the airborne and aerosol pathogen transmission is presented highlighting the efficacy of the CFD tool together with the uncertainties and limitations to resolve such complex flow physics. In their study Dbouk and Drikakis [2] and Feng et al. [10] investigated the influence of wind and relative humidity related to human coughs. An increase in droplet sizes can occur due to hygroscopic growth effects which can cause increased deposition fraction on both humans and ground [10]. Gomez et al. [14] mentioned the importance of considering the effects of droplet interactions, especially for the high droplet concentration via scale resolved simulations like large eddy simulation (LES). The studies presented in [15–17] investigated cough instances with LES. Liu et al. [18] performed LES in a restaurant set up highlighting the spatial map of airborne infection risk which revealed the existence of a high aerosol exposure index connected to the reported infection pattern. Burgmann and Janoske [19] presented a study of the transmission of aerosols in a classroom and mentioned the effective application of the air purifier in addition to window ventilation. Both mathematical (Wells–Riley) and numerical (CFD) approaches have been used by Foster and Kinzel [20]. In this regard, Wang et al. [21] also used a coupled CFD and Wells–Riley model to predict the infection probability for passengers on long-distance trains.

The design of ventilation systems is critical to reduce the risk, and many studies [9, 22–24] dealt with the ventilation efficiency, ACH, air filters, and usage of UVC light in different indoor environments. A literature study reported in [25] focused on the existing ventilation strategies of school classrooms. They mentioned that neither natural nor mixing mechanical ventilation are fully capable of dealing with long-range and short-range airborne transmissions, and states that health-based design is required compared to comfort-based design. Their study reported the potential of personalized ventilation systems for the protection of the occupants.

Despite an outpouring of new research articles emerging during the present pandemic, detailed studies of a more realistic combination of human respiratory droplet-producing instances are not abundant. Keeping social distance is challenging in an elevator setup and addressed in studies presented in [5, 6, 9, 11] considering human coughs and breathing events. However, a realistic combination of breathing and cough instances are not considered in these studies. Another important issue is to resolve the smaller respiratory particles (post evaporation) which contribute to forming aerosol particles. For example, Arpino et al. [26] used the Eulerian-Lagrangian model to study non-isothermal aerosol airborne dispersion in a passenger car cabin. They have considered the post-evaporation number and volume distributions in their simulation. Here we attempt to revisit human cough and breathing events in an elevator, considering the Eulerian dispersion medium as a multicomponent ideal gas mixture consisting of O_2 , N_2 and H_2O and pure liquid water as the Lagrangian dispersed phase of human cough/breathe. A three-dimensional AMR-based multiphase flow solver (CONVERGE 3.0) [27] is used to resolve the complex unsteady flow dynamics. Some preliminary results using this solver are reported in [28] revealing the effect of relative humidity with fixed air change per hour in an elevator setup for a human cough instance. In this work, we considered a simplified human model and realistic respiratory events of a human cough and breathing sequence in an elevator setup. To resolve the microdroplets after evaporation we set a cutoff particle diameter of $5 \mu\text{m}$ to switch off evaporation models which effectively mimics the aerosol particles in reality. It is reported in various studies that droplets smaller than approximately $5 \mu\text{m}$ in diameter can remain airborne for a long time, potentially carrying infectious viruses [29, 30]. The article is organized as follows. In section 2, we present the Eulerian-Lagrangian numerical approach. The computational setup for the different cases is presented in section 3. This is followed by the results and discussion in section 4. Finally, the conclusions are drawn in section 5.

2. Method

The unsteady flow dynamics during instances like breathing, coughing, sneezing, talking or singing in an indoor environment essentially involves fundamental mass, momentum and energy transfer in a multi-component gaseous dispersion medium together with coupled interaction with the dispersed phase arising from aerosol-producing processes. The dispersion medium consisting of oxygen, nitrogen and water vapor, is governed by the 3D compressible Navier-Stokes system of equations together with mass conservation, species conservation, and energy conservation equations. We solve the unsteady Reynolds averaged Navier-Stokes (RANS) formulation and the effects of turbulence are

resolved by the Realizable $\kappa - \epsilon$ model. The multi-component dispersion medium is assumed to follow the ideal gas law and the fluid is considered a Newtonian fluid. A Lagrangian approach is adopted for the dispersed liquid phase. The equation of motion is solved for this, considering the drag force and gravity source terms as well as buoyancy and temperature effects. Furthermore, droplet evaporation, droplet collisions, droplet turbulent dispersion, droplet-wall interactions, and droplet breakups are also considered. The detail of the governing equations is not presented here for brevity. We briefly summarize the numerical procedure below.

A three-dimensional finite volume method (FVM) and AMR-based multiphase flow solver (CONVERGE 3.0) is used for this purpose. The predictor-corrector-based Pressure Implicit with the Splitting of Operator (PISO) algorithm is chosen as the solution procedure for the momentum equation and subsequently other transport equations. The solver is based on a collocated finite volume approach together with a Rhie-Chow interpolation scheme to eliminate checker-boarding issues. A 2nd order central scheme is chosen for convective terms which switch to a 1st order upwind scheme with a step limiter to address the non-monotonic behavior for the convective fluxes. On the other hand, the diffusive terms are handled with a 2nd order central scheme. Point-wise successive over-relaxation (SOR) algorithm-based linear solver is used for each governing equation and a first-order implicit Euler scheme is used for time advancement. For the discrete phase, virtual injectors are defined to mimic the droplet-producing instances like breathing and coughing in this work (see figure 2). In accordance with previous studies in the literature, suitable size distributions, injected liquid mass [31, 32] and injection velocity signals are utilized in those instances. For the cough, Rosin-Rammler size distribution, and for breathing, a user-defined log-normal size distribution is employed for the injection of the liquid phase.

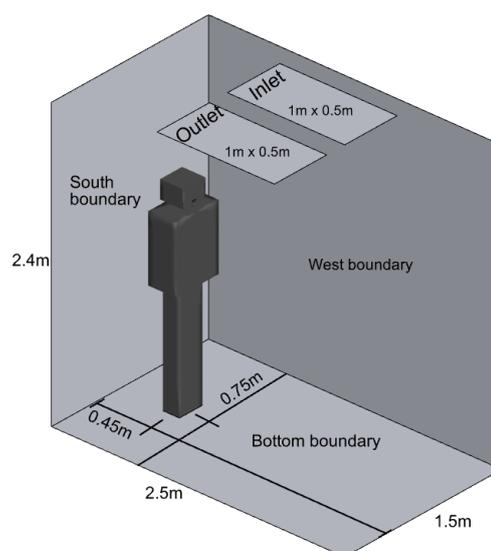


Figure 1: Computational domain of the elevator setup.

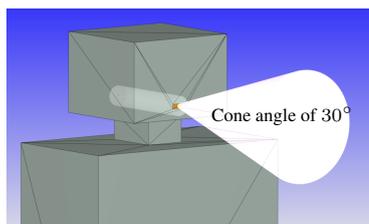


Figure 2: The mouth boundary of the standing human.

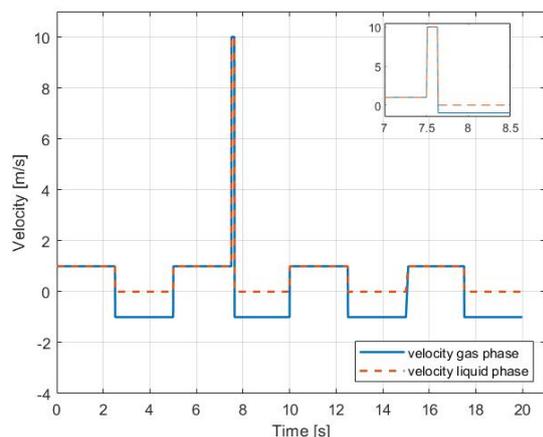


Figure 3: Velocity profiles for breathing signal and cough instance.

3. Problem setup

In this study, we focus on realistic combinations of breathing and cough instances in a confined space. The simulations are performed on a model of an elevator setup (see figure 1), with a base grid of 0.1 m resulting in a cell count of 8911 without embedding or AMR. The air supply and exhaust are placed in the center of the elevator ceiling, each with an area of 0.5 m². Air is supplied to the elevator with a supply temperature of 18 °C and relative humidity of 50 %, which is specified in terms of N_2 , O_2 and H_2O , with fractions of 0.760775, 0.231109, and 0.008116 respectively. The airflow rate of the supply air is 270 m³/h which is equivalent to an air velocity of 0.15 m/s. The simplified model of a standing human is considered with a surface area of 1.9 m² and a height of 1.75 m. The mouth is modeled as a rectangle with dimensions of 0.028 m × 0.0088 m. Figure 2 shows the mouth boundary and the virtual nozzle for the injection of the dispersed phase. The cough and breath are considered as a multiphase mixture consisting of liquid water and air ($O_2 + N_2$) ejected from the mouth through the virtual nozzle, with a specified cone angle of 30 ° for breathing and 40 ° for the cough. The turbulence intensity and length scale at the air supply inlet is set to 2 % and 0.003 m respectively, while for the mouth boundary these are assigned as 9.9 % and 0.0011 m respectively. The length scales are 7 % of the hydraulic diameter of corresponding inflow boundaries, as recommended by [33]. The Rosin-Rammler distribution for the cough is assigned with a sauter mean diameter of 73 μm and shape-parameter n=8. On the other hand, the size distribution for breathing particles is based on a log-normal distribution with a mean diameter of 5 μm. The injected droplets are represented by parcels, which is a collection of identical drops with the same properties

(same mass, radius etc.), table 1 summarizes the input data for the cough and breathing.

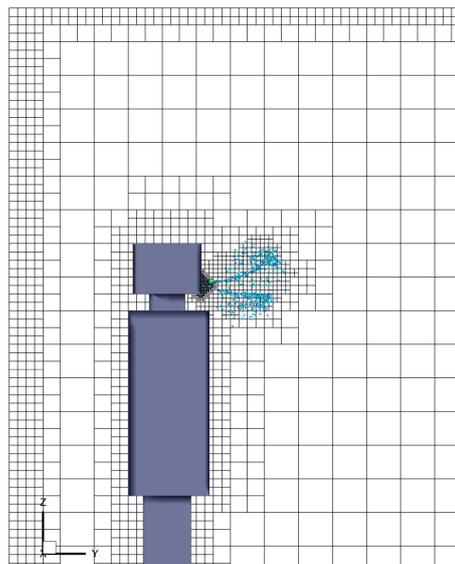


Figure 4: Mesh resolution near the wall boundaries and AMR-based resolution near the mouth at 2 s.

Table 1: Overview of input data

Cough	
Total injected mass	7.7 mg
Number of parcels	3500
Velocity	10 m/s
Spray cone angle	40 °
Duration of injection	0.12 s
Sauter mean diameter (SMD)	73 μm
Breathing	
Total injected mass	0.033 mg
Number of parcels	1600
Velocity	1 m/s
Spray cone angle	30 °
Duration of injection	2.5 s
Diameter range	0.3-15 μm
Spray model	
Evaporation	Frossling
Collision and coalescence	O'Rourke
Wall interaction	O'Rourke film splash
Droplet breakup	KH-RT
Simulation parameters	
Simulation time	20 s
Initial time step	1-e06 s
Convection CFL limit	1
Diffusion CFL limit	2
Mach CFL limit	500

KH ≡ Kelvin-Helmholtz and RT ≡ Rayleigh-Taylor

We first simulate a single cough instance (Case-1) followed by a realistic breathing (exhale-inhale) sequence embedded with a cough instance (Case-2). Figure 3 illustrates the velocity boundary conditions of the complete signal. Finally, in Case-3, a user-defined function is set to switch off evaporation models when the particle diameter is < 5 μm. Simulations are performed till the final physical time reaches 20 seconds for all cases. The flow solver is equipped with AMR functionality which efficiently reduces the computational cost by automatically applying higher mesh resolution in

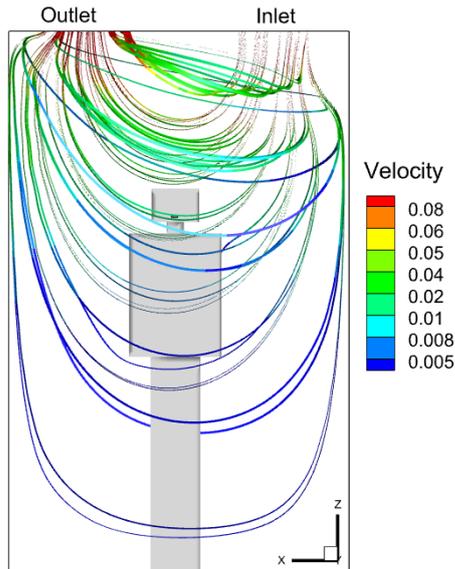


Figure 5: Steady state precursor simulations: streamlines flooded with velocity magnitude.

desired regions where flow-field variations are high while keeping the mesh resolution low elsewhere. The AMR is invoked using a velocity scalar and a passive scalar criteria in CONVERGE. Figure 4 depicts the near wall fixed mesh refinement and AMR-based mesh near the mouth region. For the mouth boundary, we have used an embedding of 10 with a scaling of 5, while for the rest of the boundaries an embedding and scaling of 2 is implemented.

4. Results and Discussions

The transient simulations of the test cases are performed with an initial state of established ventilated space inside the elevator. To achieve this, a precursor steady-state simulation is performed with the desired air supply conditions at the inlet boundary of the computational domain. The simulation is performed until a steady average velocity at the outlet boundary is reached starting from an initially stagnant space. For this steady state simulation, the mouth boundary is considered an impermeable wall. A flow visualisation is presented in figure 5 for this. The wall y^+ values lie in the range of log-layer (not shown) and can be considered sufficient for the two-equation turbulence model equipped with wall functions.

4.1. Single cough

In Case-1, we simulate a single cough instance starting at the beginning. The evolution of the dispersed phase has been illustrated in figure 6. Note that the droplets of the dispersed phase are colored with droplet radius. In this case, we allowed droplets to evaporate completely. It is clear from the colored droplets that the heavier droplets are falling at a faster rate. Noticeably, the dynamics of the droplets depend on the RH and existing airflow pattern in the elevator. A reduction in the mean diameter of the distribution of the evaporating and falling droplets is predicted by the solution (see figure 7). The velocity magnitude reduces to ≈ 0.5 m/s from the cough injection velocity of 10 m/s. From the dynamics of the droplet and wall film accumulation we noticed that it takes ≈ 11 s for all the droplets to reach the floor. This is consistent with the acceptable ranges of the evaporation-falling curves

mentioned by Xie et al. [34]. Experimental data are required for additional validation. However, the solver has been validated by other studies with similar applications [15–17]. The extent of the spread of the droplet particles under 50 % RH is found to be ≈ 0.5 m far from the mouth location in front of the human being.

4.2. Breathing and cough

Although the solution of Case-1 can be considered as a verification of the solution procedure, a realistic respiratory signal should consider background breathing instances. In Case-2, we introduce this with a signal as shown in figure 3. Note that a background breathing with exhale-inhale of 5 s is set for this. A single cough is set to occur at 7.51 s during the last phase of the second exhale of the breathing sequence. Two different virtual nozzles are assigned to realize this background breathing signal together with the single cough (figure 2). The time evolution of the droplet dynamics is shown in figure 8. Note that the exhaled breathing particles completely evaporate within a very short time. From previous literature, it can be realized that the micro-droplets of diameter less than $5 \mu\text{m}$ may remain airborne for a significant amount of time and travel a longer distance. To account for this, a user-defined function is implemented to switch off the evaporation for particle diameters below $5 \mu\text{m}$. In Case-3, we simulate the same respiratory signal as Case-2 together with this evaporation cutoff. The gas phase velocity distribution is shown in figure 9. Note that the cough jet velocity decreases rapidly from 10 m/s at the mouth to a low range ≈ 1 m/s very close to the mouth region. The dynamics of the dispersed phase clearly reveal the breathing and coughing particles. Note that the droplets are colored by their size. The violet-colored droplets are arising from the breathing sequence while larger particles are associated with the cough instance. At a later stage, (see figure 10) the dispersion of the droplets in the elevator is dramatically different compared to that predicted in Case-1. The liquid particles reach the east, south and bottom walls as well as the human body surface. The smaller droplets (diameter $< 5 \mu\text{m}$) now represent the combination of the original injection as well as arising due to evaporation of the larger droplets during this time. The distribution of the droplets (having a radius larger than $2.5 \mu\text{m}$) at different time instants is illustrated in figure 11. This corroborates the time variation of SMD of the droplet distribution. The present simulation strategy effectively resolves the aerosol behavior under the influence of the existing airflow pattern. The smaller particles remain airborne throughout the simulation time of 20 s for Case-3.

Time evolution of some of the important parameters is compared among the three cases in figure 12. It is evident that the injected respiratory particles are in accordance with the input data mentioned in Table 1 for all cases. The decrease in the particles for Case-1 and Case-2 implies the droplet coalescence and complete evaporation. On the other hand, the signature of the background breathing and cough signal in the evolution of droplets for Case-3 can be noticed. This is also reflected in the liquid mass time-evolution. The orange curve (Case-3) remains above the red curve (Case-2). The droplets reaching any impermeable boundary of the computational domain result in film accumulation accounted for by the wall-droplet interaction model. It is clear that, for a single cough, the film accumulation occurs only on the floor. The film accumulation at this bottom boundary is shown in figure

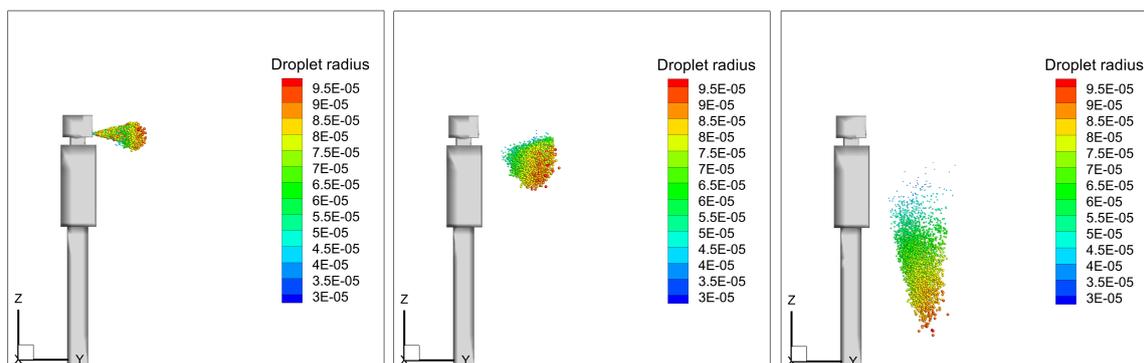


Figure 6: Evolution of dispersed phase for Case-1 at 0.1s, 0.5s and 1.9s.

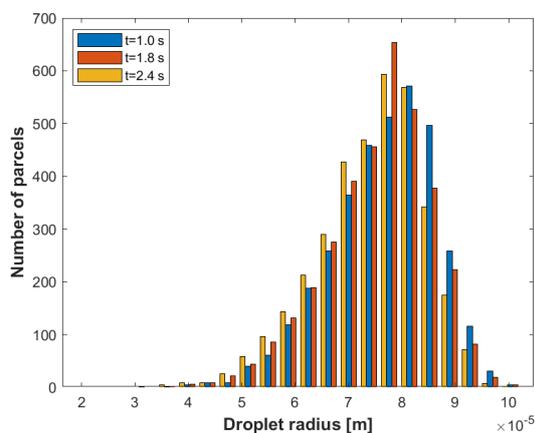


Figure 7: Size distribution for Case-1 at different time instants.

12 for all the cases. A higher film accumulation for Case-2 and Case-3 compared to Case-1 clearly depicts the incorporation of breathing in the respiratory action. Additionally, a marginal film accumulation (not shown) at the south boundary (see figure 1, the wall behind the standing man) is noticed for Case-2. However, for Case-3 due to evaporation cutoff, smaller micro-droplets remain in the domain and film accumulation is noticed in several boundaries (south, east and on the human body surfaces) in addition to the floor. The AMR-based approach to solving transient complex multiphase flow physics is advantageous to achieve lower overall computational costs. We compare the computational cells and total memory usage for Case-2 and Case-3 in figure 13, and it shows that the memory usage increases with time as more particles are injected into the domain. The simulations are performed using 96 processors, resulting in a maximum cost of 144 CPU hours for Case-2, 82 CPU hours for case-3, and less than 10 CPU hours for Case-1. The time-varying cell counts are in accordance with the breathing and cough events. The maximum cell number reaches ≈ 0.158 million when cough action triggers. Evidently, the evolution of the cell counts clearly corroborates the breath-cough signal. The memory usages are found to be in the comparable range among these cases.

We have performed a mesh sensitivity analysis with Case-3. Three different base sizes of the mesh are chosen for this, namely Mesh-1 with 0.2 m, Mesh-

2 with 0.1 m and Mesh-3 with 0.05 m. The mesh resolution at the wall boundaries and the mouth boundary are kept consistently similar by varying the embed scale and number of embed layers of CONVERGE setup, and the AMR-based embedding was kept constant as stated before. Figure 14 clearly shows that the results of Mesh-2 and Mesh-3 vary marginally.

5. Conclusions

In this study, we explored a more realistic approach to human cough instances in an elevator setup, where background breathing combined with a cough instance is analyzed. An FVM-based 3D multiphase flow solver equipped with an AMR functionality is used to solve the multi-component compressible dispersion media with the Eulerian approach while the dispersed phase is resolved via the Lagrangian approach. First, a single cough with an active evaporation model is presented to show the capability of the solver to resolve complex transient flow dynamics and evaporation-falling patterns are found to be in accordance with the literature. The respiratory signal consisting of breathing and cough is then studied with the active evaporation model. Since the particles of a diameter less than $5 \mu\text{m}$ are believed to remain airborne, a user-defined function is implemented to switch off the evaporation to account for this. This implementation makes it clear that the particle dispersion is significantly different from that expected in Case-1, with particles traveling great distances and accumulating on the walls, floor, and human body. Throughout the entire 20 s of the simulation, the particles are still airborne, demonstrating the seriousness of the risk of infection through airborne particles. Results of Case-3 clearly demonstrate the efficacy of the present method to capture the behavior of airborne particles. To the best of our knowledge, a combination of realistic breathing-coughing together with cutoff evaporation to resolve micro-droplets has not been reported in the literature before.

This work forms a basis to gain new knowledge towards understanding airborne viral transmission and designing guidelines for preventive measures in indoor establishments. The future work will involve sensitivity/uncertainty analysis and address the risk assessment in different types of indoor environments with different ventilation strategies utilizing the present method.

Acknowledgment

This work utilized the supercomputing facility of Fram

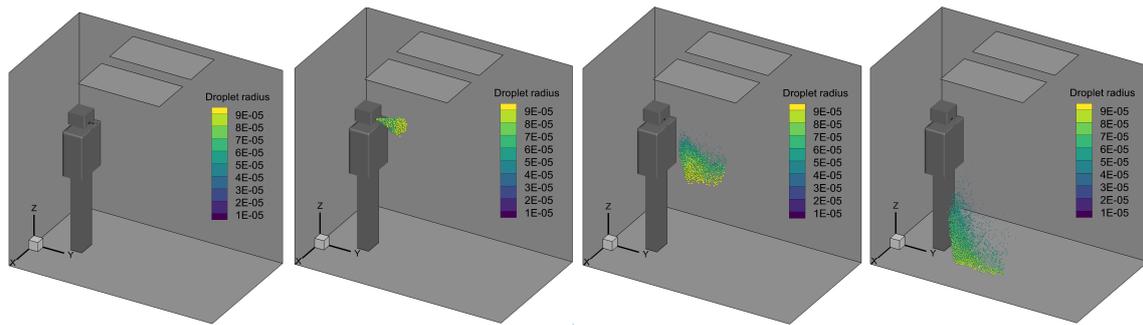


Figure 8: Evolution of dispersed phase for Case-2 at 7.3s, 7.6s, 8.3s and 10.3s.

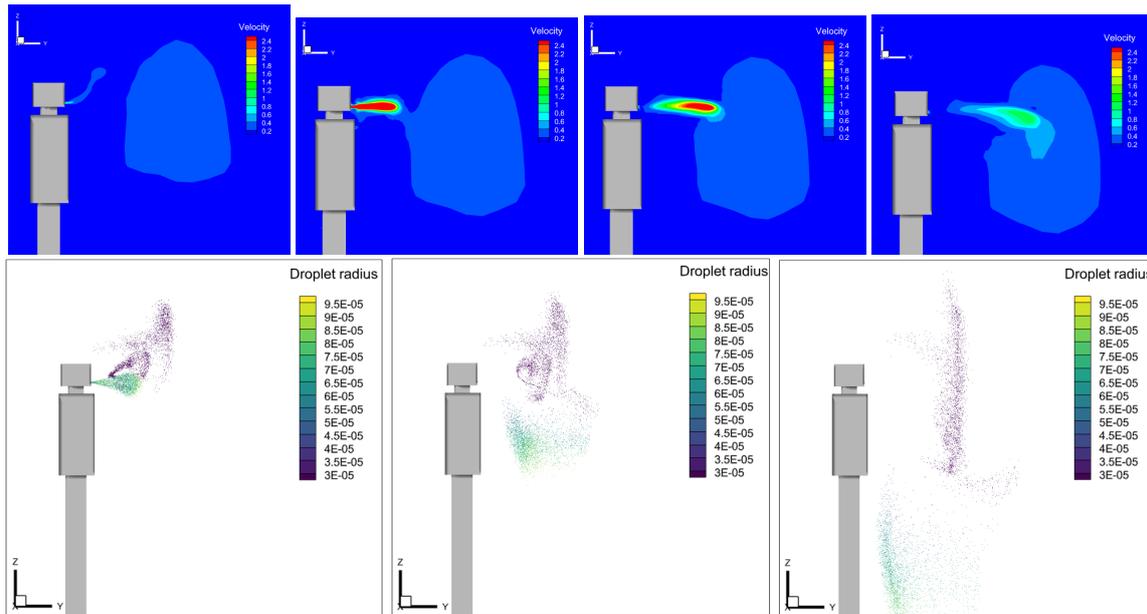


Figure 9: Top row: gas phase velocity contours for Case-3 at 7.0s, 7.6s, 7.7s and 7.9s. Bottom row: evolution of dispersed phase for Case-3 at 7.6s, 8.3s and 10s.

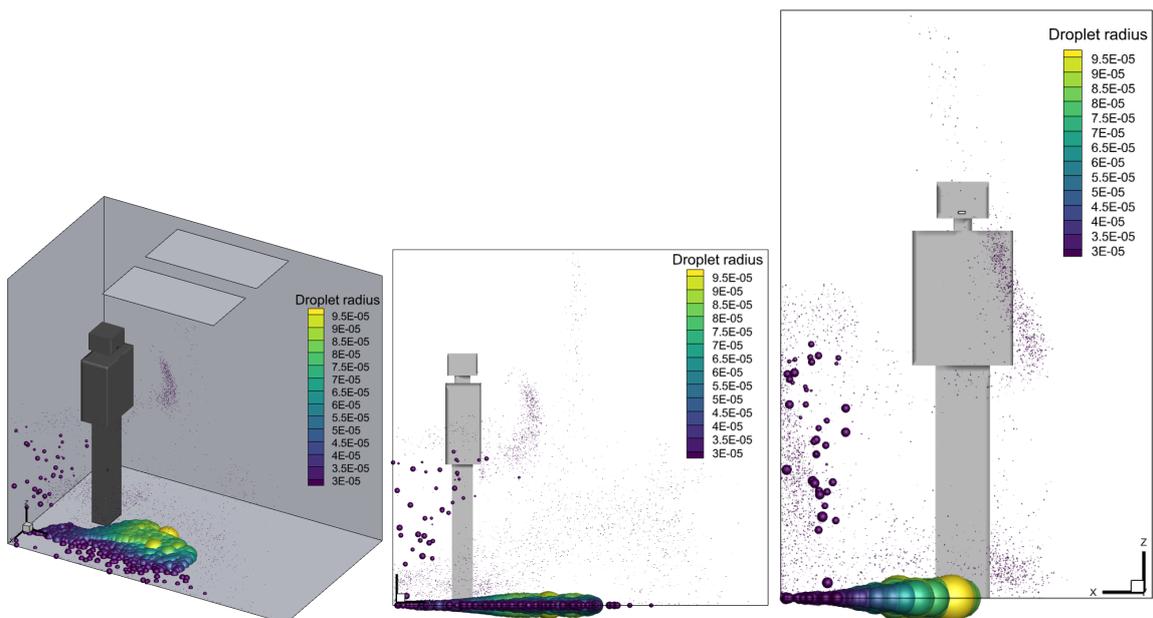


Figure 10: Dispersed phase visualisation for Case-3 at 20s.

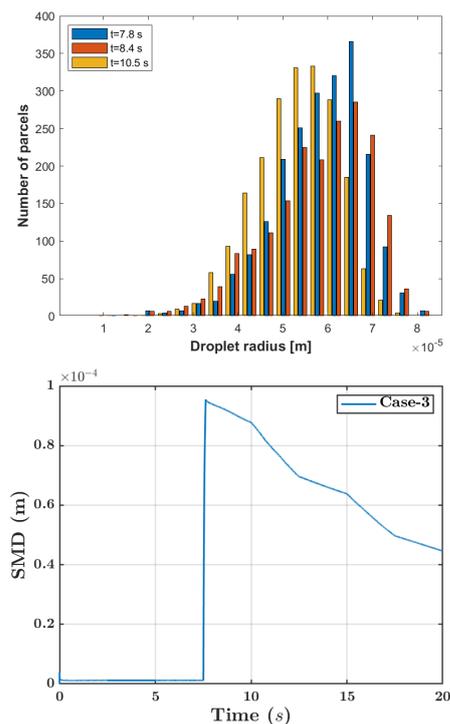


Figure 11: Liquid phase distribution for Case-3 above cutoff size.

computer cluster (UiT Arctic University of Norway) under the project NN8005K at Notur, UNINETT Sigma2 AS (National infrastructure for scientific computational science in Norway) as well as local stand alone Windows/Linux machines at OsloMet. The authors of OsloMet greatly acknowledge the collaboration with Dr. Suresh Kumar Nambully of Convergent Science GmbH.

References

- [1] G. Busco, S. R. Yang, J. Seo, and Y. A. Hassan, "Sneezing and asymptomatic virus transmission," *Physics of Fluids*, vol. 32, no. 7, p. 073309, 2020.
- [2] T. Dbouk and D. Drikakis, "On coughing and airborne droplet transmission to humans," *Physics of Fluids*, vol. 32, no. 5, p. 053310, 2020.
- [3] M.-R. Pendar and J. C. Páscoa, "Numerical modeling of the distribution of virus carrying saliva droplets during sneeze and cough," *Physics of Fluids*, vol. 32, no. 8, p. 083305, 2020.
- [4] V. D'Alessandro, M. Falone, L. Giammichele, and R. Ricci, "Eulerian-lagrangian modeling of cough droplets irradiated by ultraviolet-c light in relation to sars-cov-2 transmission," *Physics of Fluids*, vol. 33, no. 3, p. 031905, 2021.
- [5] N. Sen, "Transmission and evaporation of cough droplets in an elevator: Numerical simulations of some possible scenarios," *Physics of Fluids*, vol. 33, no. 3, p. 033311, 2021.
- [6] T. Dbouk and D. Drikakis, "On airborne virus transmission in elevators and confined spaces," *Physics of Fluids*, vol. 33, no. 1, p. 011905, 2021.
- [7] Y. Zhou and S. Ji, "Experimental and numerical study on the transport of droplet aerosols generated by occupants in a fever clinic," *Building and Environment*, vol. 187, p. 107402, 2021.
- [8] J. K. Gupta, C.-H. Lin, and Q. Chen, "Flow dynamics and characterization of a cough," *Indoor Air*, vol. 19, no. 6, pp. 517–525, 2009.
- [9] S. Shao, D. Zhou, R. He, J. Li, S. Zou, K. Mallery, S. Kumar, S. Yang, and J. Hong, "Risk assessment

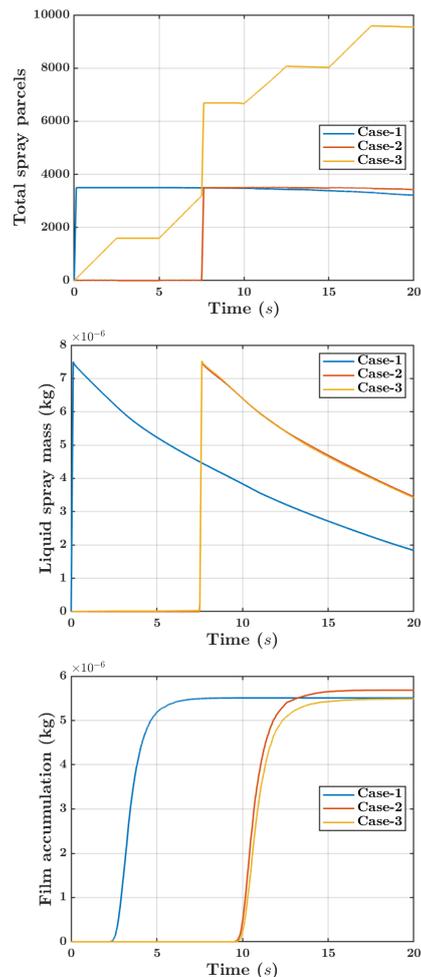


Figure 12: Time evolution of different parameters for three cases.

- of airborne transmission of covid-19 by asymptomatic individuals under different practical settings," *Journal of Aerosol Science*, vol. 151, p. 105661, 2021.
- [10] Y. Feng, T. Marchal, T. Sperry, and H. Yi, "Influence of wind and relative humidity on the social distancing effectiveness to prevent covid-19 airborne transmission: A numerical study," *Journal of Aerosol Science*, vol. 147, p. 105585, 2020.
- [11] S. Liu, X. Zhao, S. R. Nichols, M. W. Bonilha, T. Derwinski, J. T. Auxier, and Q. Chen, "Evaluation of airborne particle exposure for riding elevators," *Building and Environment*, vol. 207, p. 108543, 2022.
- [12] C. Crowley, B. Murphy, C. McCaul, R. Cahill, and K. P. Nolan, "Airborne particle dispersion by high flow nasal oxygen: An experimental and cfd analysis," *Plos One*, vol. 17, no. 1, p. e0262547, 2022.
- [13] Y. Sheikhnejad, R. Aghamolaei, M. Fallahpour, H. Motamedi, M. Moshfeghi, P. A. Mirzaei, and H. Bordbar, "Airborne and aerosol pathogen transmission modeling of respiratory events in buildings: An overview of computational fluid dynamics," *Sustainable Cities and Society*, p. 103704, 2022.
- [14] A. Gomez-Flores, G. Hwang, S. Ilyas, and H. Kim, "A cfd study of the transport and fate of airborne droplets in a ventilated office: The role of droplet-droplet interactions," *Front. Environ. Sci. Eng.*, vol. 16, no. 3, pp. 1–14, 2022.
- [15] S. Trivedi, S. Gkantonas, L. C. Mesquita, S. Iavarone, P. M. d. Oliveira, and E. Mastorakos, "Estimates of the

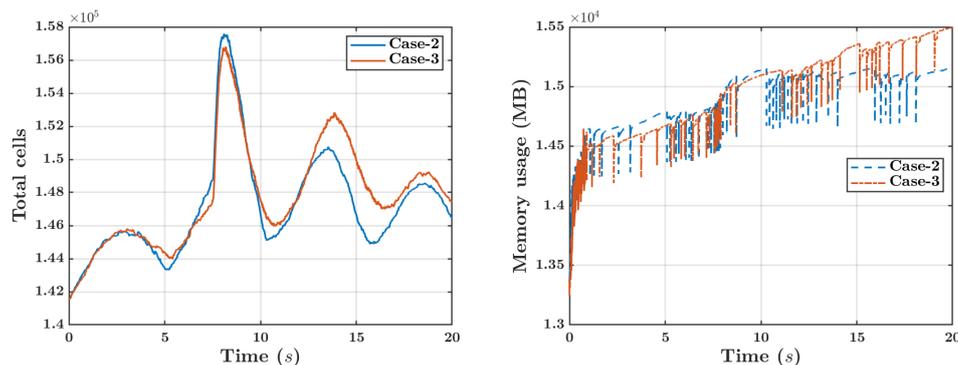


Figure 13: Computational cells and total memory usage for cough and breathing cases.

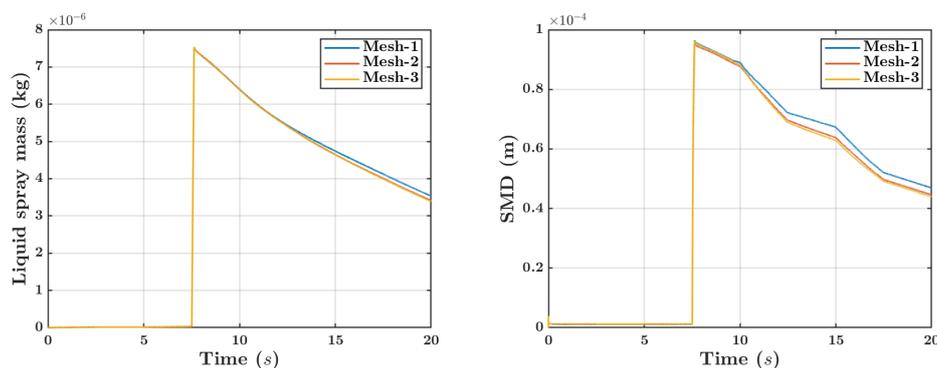


Figure 14: Mesh sensitivity analysis for Case-3.

- stochasticity of droplet dispersion by a cough,” *Physics of Fluids*, vol. 33, no. 11, p. 115130, 2021.
- [16] H. Ge, L. Chen, C. Xu, and X. Cui, “Large-eddy simulation of droplet-laden cough jets with a realistic manikin model,” *Indoor and Built Environment*, p. 1420326X211032247, 2021.
- [17] H. Ge, P. Zhao, S. Choi, T. Deng, Y. Feng, and X. Cui, “Effects of face shield on an emitter during a cough process: A large-eddy simulation study,” *Science of The Total Environment*, vol. 831, p. 154856, 2022.
- [18] H. Liu, S. He, L. Shen, and J. Hong, “Simulation-based study of covid-19 outbreak associated with air-conditioning in a restaurant,” *Physics of Fluids*, vol. 33, no. 2, p. 023301, 2021.
- [19] S. Burgmann and U. Janoske, “Transmission and reduction of aerosols in classrooms using air purifier systems,” *Physics of Fluids*, vol. 33, no. 3, p. 033321, 2021.
- [20] A. Foster and M. Kinzel, “Estimating covid-19 exposure in a classroom setting: A comparison between mathematical and numerical models,” *Physics of Fluids*, vol. 33, no. 2, p. 021904, 2021.
- [21] Z. Wang, E. R. Galea, A. Grandison, J. Ewer, and F. Jia, “A coupled computational fluid dynamics and wells-riley model to predict covid-19 infection probability for passengers on long-distance trains,” *Safety Science*, vol. 147, p. 105572, 2022.
- [22] F. Mohamadi and A. Fazeli, “A review on applications of cfd modeling in covid-19 pandemic,” *Arch Computat Methods Eng*, pp. 1–20, 2022.
- [23] H. Motamedi, M. Shirzadi, Y. Tominaga, and P. A. Mirzaei, “Cfd modeling of airborne pathogen transmission of covid-19 in confined spaces under different ventilation strategies,” *Sustainable Cities and Society*, vol. 76, p. 103397, 2022.
- [24] C. Anuraghava, K. Abhiram, V. N. S. Reddy, and H. Rajan, “Cfd modelling of airborne virus diffusion characteristics in a negative pressure room with mixed mode ventilation,” *International Journal for Simulation and Multidisciplinary Design Optimization*, vol. 12, p. 1, 2021.
- [25] E. Ding, D. Zhang, and P. M. Bluyssen, “Ventilation regimes of school classrooms against airborne transmission of infectious respiratory droplets: A review,” *Building and Environment*, vol. 207, p. 108484, 2022.
- [26] F. Arpino, G. Cortellessa, G. Grossi, and H. Nagano, “A eulerian-lagrangian approach for the non-isothermal and transient cfd analysis of the aerosol airborne dispersion in a car cabin,” *Building and Environment*, vol. 209, p. 108648, 2022.
- [27] K. Richards, P. Senecal, and E. Pomraning, “Convergent science, converge 3.0*.” <https://convergecf.com/>.
- [28] P. Patrick, “Analysis of human cough in confined spaces: a numerical study,” master’s thesis, Institutt for bygg- og energiteknikk, Oslo Metropolitan University, 2021.
- [29] S. Asadi, N. Bouvier, A. S. Wexler, and W. D. Ristenpart, “The coronavirus pandemic and aerosols: Does covid-19 transmit via expiratory particles?,” 2020.
- [30] K. A. Prather, C. C. Wang, and R. T. Schooley, “Reducing transmission of sars-cov-2,” *Science*, vol. 368, no. 6498, pp. 1422–1424, 2020.
- [31] T. Dbouk and D. Drikakis, “On respiratory droplets and face masks,” *Physics of Fluids*, vol. 32, no. 6, p. 063303, 2020b.
- [32] X. Xie, Y. Li, H. Sun, and L. Liu, “Exhaled droplets due to talking and coughing,” *Journal of the Royal Society Interface*, vol. 6, no. suppl_6, pp. S703–S714, 2009.
- [33] H. K. Versteeg and W. Malalasekera, *An introduction to computational fluid dynamics: the finite volume method*. Pearson Education Limited, 2. edition ed., 2007.
- [34] X. Xie, Y. Li, A. Chwang, P. Ho, and W. Seto, “How far droplets can move in indoor environments—revisiting the wells evaporation-falling curve,” *Indoor Air*, vol. 17, no. 3, pp. 211–225, 2007.

Uncertainty quantification and sensitivity analysis during the development and validation of numerical artery models

Friederike Schäfer ^{a,*}, Jacob Sturdy ^a, Mateusz Mesek ^b, Aleksander Sinek ^b, Ryszard Białecki ^b, Ziemowit Ostrowski ^b, Bartłomiej Melka ^b, Marcin Nowak ^b, Leif Rune Hellevik ^a

^a Division of Biomechanics, Department of Structural Engineering, NTNU, Trondheim, Norway,

^b Biomedical Engineering Lab, Department of Thermal Technology, SUT, Gliwice, Poland,
friederike.e.schaefer@ntnu.no

Abstract

Increasing age and cardiovascular diseases lead to stiffening of the vasculature. Knowledge about an individual's arterial stiffness gives insights into the current state of the cardiovascular system and it is considered to be a valuable diagnostic index. However, arterial stiffness cannot be measured directly. Numerical modelling based on measurements of flow and deformation in an individual's artery enable an indirect means. Our research aims to develop a method to estimate the local arterial stiffness of an artery from non-invasive measurements through inverse modelling. Experimental measurement limitations and the unmeasurable nature of model input parameters lead to uncertainties in the model prediction. Uncertainty quantification and sensitivity analysis (UQSA) inform about how the model prediction is influenced by these uncertainties. Due to the computational expenses of 3D fluid-structure interaction (FSI) models, we reduced the model's complexity to a 1D model. To verify the 3D-FSI implementation and validate the 1D implementation we performed simulated inflation tests and compared the results with analytical theory. 3D-FSI simulations were performed and compared to the 1D-model predictions for different simplification assumptions. To quantify the impact of uncertainties in input data, polynomial chaos expansion for UQSA was applied to the 1D-model. This analysis revealed the model input parameters which lead to the highest variability in model prediction. UQSA showed that variations in the Young's modulus and the lumen radius lead to the largest variability in the 1D-model prediction. Thus, we focused in the validation process on the comparison between the the arterial wall behaviour between the 1D and the 3D-FSI model.

1. Introduction

Arterial stiffening is an ubiquitous process associated with human ageing and can be observed in all humans over age 30 [1]. Stiffening itself contributes to cardiovascular disease, and some diseases, such as arteriosclerosis, directly cause arterial stiffening [2]. Clinicians recognize arterial stiffness as a valuable biomarker describing the overall state of an individual's cardiovascular system. However, arterial stiffness cannot be measured directly [3]. Using numerical models and an inverse problem formulation, we seek to estimate arterial stiffness from non-invasive flow and deformation measurements.

The accuracy of an inverse problem's solution depends on the accuracy of the forward model response, which is limited by experimental measurement uncertainties and unavailable model parameters. Through uncertainty quantification (UQ), the propagation of input uncertainties through a numerical model can be investigated. Subsequent sensitivity analysis (SA) attributes the uncertainty in the model response to individual model input parameters and their interactions [4]. These results can help to identify key sources of error as well as focus efforts on developing and validating influential portions of the forward model. However, UQSA require the evaluation of the entire uncertain input parameter space increasing the computational expenses with the number of uncertain model parameters and dimensions.

Three-dimensional (3D) models of vessel segments enable the simulation of patient-specific geometries giving spatially resolved insights into a patients' haemodynamics. Of all model types, fluid-structure interaction (FSI) models represent the anatomical and physiological state in the greatest detail. However, the computational expenses for such models are high, making it impractical to conduct several thousand model evaluations for UQSA.

Reducing the spatial model dimension to one (1D) decreases the information which can be retrieved from the model, but it also reduces the computational expenses significantly. In vessel segments where the vascular flow is independent of the geometry, 1D models represent hemodynamic quantities like pressure and flow rate accurately [5]. Low computational expenses and the ability to evaluate specific hemodynamic quantities accurately make 1D models suitable for UQSA.

The aim of this work is to compare the 1D model prediction of pressure, flow rate, and radial deformation of the common carotid artery (CCA) against a 3D-FSI model and to determine focus for model development. To gain insights into the uncertainty propagation from the input parameters to the 1D model response, we applied polynomial chaos (PC) expansion. PC is an efficient way of representing random model inputs as a polynomial function of another random variable enabling

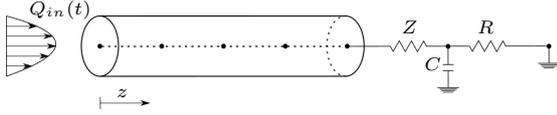


Figure 1: Representation of the 1D-model used in this validation work. As an outlet boundary condition, a Windkessel model with arterial impedance Z , compliance C , and resistance R is applied.

the evaluation of qualitative and quantitative uncertainty measures [6]. Information about the most sensitive 1D model input parameters can be used in the process of validating low-dimensional models against higher fidelity models and experimental measurements.

2. Methodology

In this work, steps for the verification and validation of a 1D-model of the CCA were conducted. The workflow included verification of the 1D-model, UQSA, as well as initial steps for the validation against a 3D-FSI model.

2.1. 1D-model formulation

The CCA in the 1D model is represented as a straight deformable tube, depicted in Fig. 1. The quantities of interest, pressure p , volumetric flow rate Q , and radius change ΔR , are evaluated at five equidistant nodes located on the tube's centreline. In the vessel, laminar and axisymmetric blood flow is assumed. Impermeability and homogeneity characterize the vessel wall. Deformation of the vessel wall is purely in the circumferential direction. Blood is modelled as an incompressible Newtonian fluid. Based on these assumptions, the conservation laws of mass and momentum can be formulated as

$$\frac{\partial A}{\partial t} + \frac{\partial(Au)}{\partial z} = 0 \quad (1a)$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial z} + \frac{1}{\rho_f} \frac{\partial p}{\partial z} = \frac{f}{\rho_f A}, \quad (1b)$$

where the quantities u and p are the velocity and pressure averaged over the vessel cross-section A , ρ_f is the fluid density, and f represents the frictional force per unit length [7]. The velocity profile of the fluid flow determines the magnitude of f which accounts for the wall shear stress and convective inertia terms. A commonly used symmetric polynomial model of the velocity is applied:

$$u_r(z, r, t) = u(z, t) \frac{\zeta + 2}{\zeta} \left[1 - \left(\frac{r}{R} \right)^\zeta \right], \quad (2)$$

where u_r is the velocity at a given radial distance r from the centerline. R denotes the vessel radius, and the polynomial order ζ determines the shape of the profile, where a higher ζ corresponds to a blunter profile. For this profile and a constant dynamic blood viscosity μ , the friction term becomes $f = -2(\zeta + 2)\mu\pi u$. In all simulations of this work, a parabolic velocity profile with $\zeta = 2$ was applied.

The arterial wall was modelled as a thin, incompressible, homogenous, isotropic, elastic material. The tube law describes the fluid-structure interaction by relating the pressure to the area with

$$p = p_{dia} + \frac{\beta}{A_{dia}} (\sqrt{A} - \sqrt{A_{dia}}) \quad (3)$$

with $\beta = \frac{\sqrt{\pi} E h}{(1 - \nu^2)}$,

where p_{dia} and A_{dia} are the diastolic pressure and cross-sectional area, respectively. Generally, A_{dia} is larger than the reference area A_0 because A_0 is the lumen area when $p = 0$. Material properties of the vessel wall are the Young's modulus E , wall thickness h , and Poisson's ratio ν .

Eqs. 1 resemble a hyperbolic system of equations such that at each boundary one boundary condition (BC) needs to be specified. A representative volumetric flow rate $Q_{in}(t)$ of the CCA was imposed at the inlet boundary [8]. At the outlet, the 1D-model was coupled with a three element Windkessel model, mimicking the behaviour of the downstream vasculature. The elements of the Windkessel model are a resistor Z modelling the arterial impedance, a second resistor representing the peripheral resistance R , and a capacitor C , which mimics the arterial compliance. Relating pressure and flow, the Windkessel model equation reads

$$\frac{\partial p}{\partial t} + \frac{p}{RC} = \left(\frac{1}{C} + \frac{Z}{RC} \right) Q + Z \frac{\partial Q}{\partial t}. \quad (4)$$

As an initial condition, pressure in the entire domain was set to $P_d = 74.5$ mmHg [9] and the initial flow rate was 7.7 mL/s. An explicit MacCormack scheme was used to solve the system of equations. This solution method is second order in space and time. To grant stability of an explicit scheme, the Courant-Friedrich-Lewy (CFL) condition needs to be satisfied. Information travels in the domain from one element to the next with a forward travelling wave, which is the sum of the fluid flow speed u and the pulse wave velocity c . Using the constitutive equation given in Eq. 3, the pulse wave velocity is

$$c = \sqrt{\frac{\beta}{2\rho A_{dia}}} A^{1/4}. \quad (5)$$

The CFL condition becomes with the grid spacing Δx and the time step Δt

$$CFL = (u + c) \frac{\Delta t}{\Delta x} \leq 1. \quad (6)$$

Since Eqs. 1 form a hyperbolic system of equations, physiological conditions have two Riemann invariants that travel in opposite directions, thus at each boundary the BCs can only be specified to determine the single Riemann invariant entering the domain [10]. We imposed a volumetric inflow at the inlet and a relationship between pressure and flow at the outlet.

2.2. Grid independence of the 1D model

Model verification is essential for every numerical study [11]. In this work, a periodic solution of pressure, flow and radius change was reached within twelve cardiac cycles, demonstrating iterative convergence. Consistency of the model was shown through mass conservation over five consecutive heart beats and a pressure drop from inlet to outlet. In total, three grids with increasing number of grid points were tested. The coarsest grid with a total of five nodes returned the same values for pressure, flow, and radius change as the finer grids with ten and 15 nodes. Therefore, all simulations used a grid of five nodes. The maximum allowed time step Δt_{max} during the simulations was determined with a maximal CFL number of 0.8 and the five grid points following Eq. 6. To check temporal convergence, the time step was set manually in two subsequent simulation runs to $0.5 \Delta t_{max}$ and $0.25 \Delta t_{max}$. Decreasing the time step did not alter the

Table 1: Baseline hemodynamic parameters and uncertainties of the 1D CCA-model of a healthy 45-year-old. Each value is given with their lower and upper bound.

	Lower	Upper	Unit	
R	2.525	3.495	mm	[12]
h	0.8875	1.3725	mm	[13]
E	236	1306	kPa	[9]
ν	0.4908	0.4912	-	[14]
ρ	1043	1054	kg/m ³	[15]
μ	3.784	5.516	mPa s	[16]
R_{tot}	1.9066	2.3302	10 ⁹ Pa s/m ³	[17]
C	1.2192	1.4901	10 ⁻¹⁰ m ³ /Pa	[17]

accuracy of pressure, flow rate, and radius change values. Therefore, the maximal allowed time step Δt_{max} was used for the simulations.

2.3. Validation of the 1D model

As a first step of the validation process, we conducted UQSA of the 1D-model based on population variations in parameter values. The results of SA motivated us to place special focus on the structural mechanics of the FSI-solver. Therefore, we compared a finite element (FE) inflation test with the analytical solution of thin- and thick-walled cylinder theory.

Uncertainty quantification and sensitivity analysis

Measurement errors and lack of knowledge lead to uncertainties about model inputs and thus propagate through numerical models to contribute to uncertainty about model outputs [18]. Quantifying this uncertainty is an integral part of the model verification and validation process [11], as well as a prerequisite for its implementation in clinical decision support [18]. However, computational costs for UQ increase with increasing number of uncertain model parameters and model dimension, limiting efficient analysis to low-fidelity models. This motivates to conduct UQSA on the 1D-model prior to its validation against the 3D-FSI model since UQSA will give insights into the model characteristics. Validation test cases can then be designed in such a way to capture these characteristics [11].

We considered a total of eight uncertain parameters based on the variations in a 45-year old healthy population given in literature. A uniform distribution was assumed for all parameters, summarized with their lower and upper bounds in Tab. 1. Due to lack of data and lack of knowledge, the uncertainty of the Windkessel model parameters, C and $R_{tot} = R + Z$, were assumed with a deviation of $\pm 10\%$ from their respective reference value. Compliance was adjusted from reference values in order to represent a 45-year old subject, whereas vascular resistance was considered to be age-independent [17, 19]. The vessel length was held constant at 126 mm for UQSA. The 1D-model's governing equations can be summarized with the deterministic inputs \mathbf{z} and deterministic output Y with a black box functional f as

$$y = f(\mathbf{z}). \quad (7)$$

If model input parameters are uncertain, then the model output becomes uncertain as well. This is emphasized by rewriting Eq. 7 in terms of a stochastic vector of input variables, \mathbf{Z} , that yields a stochastic output Y

$$Y = f(\mathbf{Z}). \quad (8)$$

PC expansion was used for estimating uncertainty and sensitivity measures of this stochastic model, which can

Table 2: Material parameters and dimensions of the 3D-FSI model. Material parameters, ρ_s , ρ_f , μ , and ν , are taken from [20].

Property	Value	Unit
Length L	115	mm
Unstressed radius R_d	2	mm
Wall thickness h	0.5	mm
Young's modulus E	700.0	kPa
Wall density ρ_s	1120	kg/m ³
Wall Poisson ratio ν	0.49	-
Fluid density ρ_f	998.2	kg/m ³
Fluid dynamic viscosity μ	1.003	mPa s

be written in the form of Eq. 8. A finite number of polynomials N approximate Y through a sum of expansion coefficients c_p and orthogonal polynomials Φ_p

$$Y \approx \sum_{p=0}^N c_p \Phi_p(\mathbf{Z}), \quad (9)$$

where the orthogonality of Φ_p is with respect to the distributions in \mathbf{Z} [6].

SA attributes model output variance to particular model inputs, as well as to the interactions between the uncertain input parameters [4]. From the PC expansion, total variance of the model output, $\text{Var}[Y]$, was computed as

$$\text{Var}[Y] \approx \text{Var}[Y_{PC}] = \sum_p \text{Var}[c_p \Phi_p(\mathbf{Z})]. \quad (10)$$

The main sensitivity index S_i quantifies the direct effect of particular input parameter z_i on $\text{Var}[Y]$. With the set A_i containing all basis functions depending only on z_i , the main sensitivity index S_i describing the fraction of output variance due to z_i , can be approximated as

$$S_i \approx \frac{1}{\text{Var}[Y_{PC}]} \sum_{p \in A_i} \text{Var}[c_p \Phi_p]. \quad (11)$$

The set $A_{T,i}$ contains all basis functions where the random input z_i and all its interactions with $z_{\sim i}$ are involved, such that the total model output variance with respect to z_i can be represented with the total sensitivity index ST_i as

$$ST_i \approx \frac{1}{\text{Var}[Y_{PC}]} \sum_{p \in A_{T,i}} \text{Var}[c_p \Phi_p]. \quad (12)$$

PC expansion of orders one to three were tested, showing that the sensitivity indices were already converged for the third order. At the mid-point of the vessel, we computed the sensitivity indices through time-averaging and weighting by the last cardiac cycle's variance, as well as normalizing with the model output variance [21].

3D-model formulation

The CCA is modelled as an idealized, hollow cylinder. Tab. 2 summarizes the dimensions and material properties of the model. The mathematical formulation of the 3D problem assumes an incompressible Newtonian fluid and a linear elastic vessel wall. Mass and momentum conservation of the fluid domain Ω_f are

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega_f \quad (13a)$$

$$\rho_f \frac{\partial \mathbf{u}}{\partial t} + \rho_f \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \nabla \cdot \boldsymbol{\tau}_f + \mathbf{b} \quad \text{in } \Omega_f \quad (13b)$$

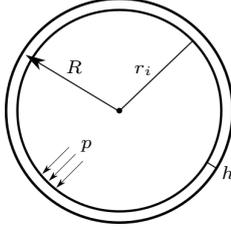


Figure 2: Representation of a thick-walled cylinder subjected to internal pressure p and external pressure $p_{ext} = 0$ Pa.

where ρ_f denotes the fluid density, \mathbf{u} , the fluid velocity, p the pressure inside the vessel, $\boldsymbol{\tau}_f = \mu_f(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$ the viscous stress tensor for a Newtonian fluid, and \mathbf{b} body forces which are assumed to be zero. At the fluid-structure interface, a no-slip boundary condition was applied. The governing equations for the solid domain Ω_s are

$$\rho_s \frac{\partial \mathbf{u}_s}{\partial t} - \nabla \cdot \boldsymbol{\sigma}_s = \mathbf{b}_s \quad \text{in } \Omega_s \quad (14a)$$

$$\boldsymbol{\sigma}_s \cdot \mathbf{n}_s = (-P\mathbf{I} + \boldsymbol{\tau}_f) \cdot \mathbf{n}_s \quad \text{on } \Gamma \quad (14b)$$

with the wall density ρ_s , Cauchy stress tensor $\boldsymbol{\sigma}_s$, body forces on the solid domain \mathbf{b}_s , which are assumed to be zero, and the outward normal vector \mathbf{n}_s on Γ .

Static and transient inflation test

In order to assure the correct settings for the comparison between the 3D-FSI and the 1D-model, we tested the mechanical solver individually with a static and transient inflation test. Longitudinal displacement at the ends was either prohibited or allowed, representing thick-walled cylinder theory with fixed surfaces at the end or free surfaces without stress, respectively. Radial displacement was allowed in both cases. The pressure inside the tube was increased linearly from the reference pressure $p_0 = 0$ Pa to 20000 Pa over ten seconds. Large deformation of the domain was suppressed.

The wall displacement of the inflation test cases was compared with the analytical solution of the thick-walled cylinder theory, represented in Fig. 2. For a thick-walled cylinder with fixed ends and no external pressure acting on the outer surface is radial displacement described as [22]

$$u(R) = \frac{1}{2G} \frac{r_i}{1 - \left(\frac{r_i}{r_i+h}\right)^2} \left(\frac{r_i}{R} + (1 - 2\nu) \left(\frac{r_i}{r_i+h} \right)^2 \frac{R}{r_i} \right) P, \quad (15)$$

where r_i is the inner radius and G the shear modulus as

$$G = \frac{E}{2(1 + \nu)}. \quad (16)$$

In order to represent the thick-walled cylinder with free ends, a correction term $\Delta u(R)$ needs to be added to Eq. 15

$$\Delta u(R) = -\frac{\nu^2}{G(1 + \nu)} \frac{-\left(\frac{r_i}{r_i+h}\right)^2 P}{1 - \left(\frac{r_i}{r_i+h}\right)^2} R. \quad (17)$$

To investigate inertia effects, we tested simulation times of five seconds while maintaining the maximum pressure. This set of test cases resulted in a total of two static and four transient simulations. Mesh

independence was demonstrated through three mesh refinements. All simulations were performed in Ansys Mechanical (Version 2021 R2) with the standard solver MAPDL. Time integration was performed with an implicit scheme.

In addition to the analytic thick-walled cylinder theory, the radial displacement was analysed under the assumption of a thin-walled cylinder with free ends. This analysis was conducted because the current 1D-model contains this assumption. The circumferential stress σ_c in a thin-walled cylinder [22] is

$$\sigma_c = \frac{P r}{h}, \quad (18)$$

with the middle radius $r = r_i + 0.5h$, where r_i is the inner vessel radius and h the vessel wall thickness. From the linear stress-strain relation it follows for the displacement

$$u_{thin} = \frac{P r^2}{E h}. \quad (19)$$

One-way FSI simulation without flow

To ensure a correct coupling between the fluid and the solid domain in the study, we conducted a one-way FSI simulation with the same pressure increase as in the inflation test, fixed ends of the artery, and a resting fluid domain. Ansys (Version 2020 R2) Fluent was used to generate and solve the fluid domain. Both solid and fluid domains consisted of hexahedral elements. A coupled pressure-velocity solver evaluated the fluid problem, and the standard MAPDL solver evaluated the solid problem. Since Ansys uses the segregated approach to solve a FSI problem, the equations are not solved simultaneously within one matrix [20]. Instead, each of the participant programs calculates its own problem. Fluent first iterates the fluid domain for a set number of iterations, then passes the evaluated force which the fluid exerts on the structure to Mechanical so that the displacement of the structural part can be obtained. This constitutes to one coupling iteration in one-way FSI simulations. One coupling step is equivalent to one time step and it contains a set number of coupling iterations. Ansys guidelines suggest between one and five. The number of iterations in Fluent was set to seven, Mechanical's iteration number was program controlled. The maximum amount of coupling iterations was five, with three usually being sufficient to meet the convergence criteria after a few time steps had passed since the start of the simulation [20].

3. Results

In Fig. 3, the time-averaged main and total sensitivity indices for the pressure, flow rate, and radius change predicted by the 1D-model, Eqs. 1-3, are presented for the mid-point of the artery. Further, the numerical values of the time-averaged main and total sensitivity indices for the quantities of interest are given in Tab. 3. For all quantities of interest, the uncertainty in the Young's modulus is responsible for the largest part of the model output variation. Lumen radius and wall thickness have a minor influence, whereas the fluid properties ρ_f and μ , the Poisson ratio, and the total resistance of the Windkessel model have no influence on model output variation. Additionally, all quantities of interest are slightly sensitive to the compliance value of the Windkessel model. For all uncertain input parameters the main and total sensitivity indices are approximately the same, indicating that parameter interaction does not play a significant role.

Over one cardiac cycle, Fig. 4 displays the 95 % prediction interval for the pressure and the relative radial

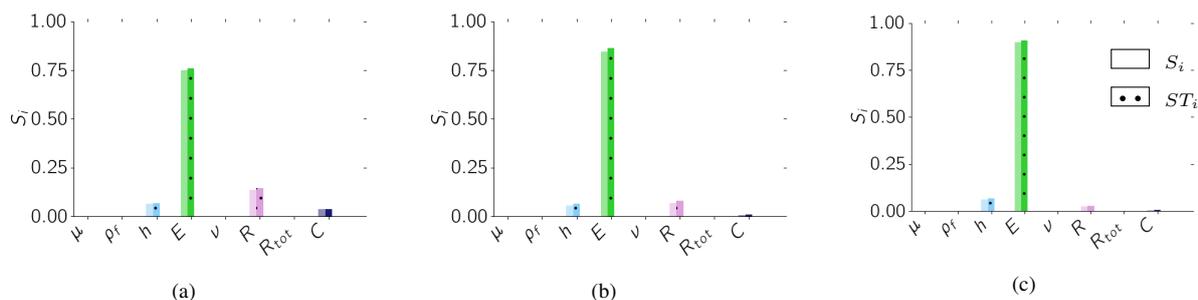


Figure 3: Main and total sensitivity indices, S_i and ST_i , respectively, evaluated at the mid-point of the 1D-model. All sensitivity indices are time-averaged over the last cardiac cycle. The quantities of interest are (a) pressure p , (b) flow rate Q , and (c) radius change ΔR with reference to the minimum radius of the respective model evaluation.

Table 3: Numerical values for the main and total sensitivity indices, S_i and ST_i , respectively, evaluated at the mid-point of the 1D-model for pressure p , flow rate Q , and radius change ΔR with reference to the minimum radius of the respective model evaluation. All sensitivity indices are time-averaged over the last cardiac cycle.

	μ		ρ_f		h		E		ν		R		R_{tot}		C	
	S	ST	S	ST	S	ST	S	ST	S	ST	S	ST	S	ST	S	ST
p	0.00	0.00	0.00	0.00	0.06	0.07	0.75	0.76	0.00	0.00	0.14	0.14	0.00	0.00	0.04	0.04
Q	0.00	0.00	0.00	0.00	0.06	0.07	0.85	0.86	0.00	0.00	0.07	0.08	0.00	0.00	0.01	0.01
ΔR	0.00	0.00	0.00	0.00	0.06	0.06	0.90	0.91	0.00	0.00	0.03	0.03	0.00	0.00	0.01	0.01

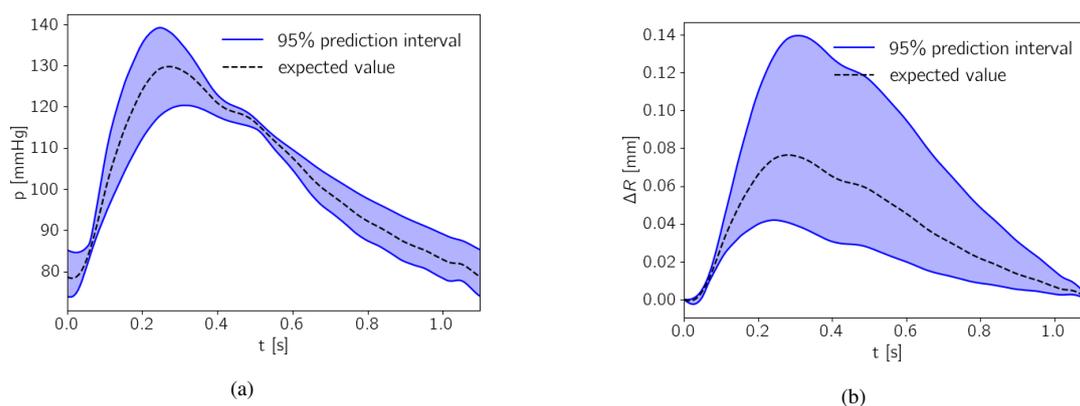


Figure 4: 95% prediction interval for the last cardiac cycle at the mid point of the artery for the (a) pressure p and the (b) relative radius change ΔR with reference to the minimum radius of the respective model evaluation.

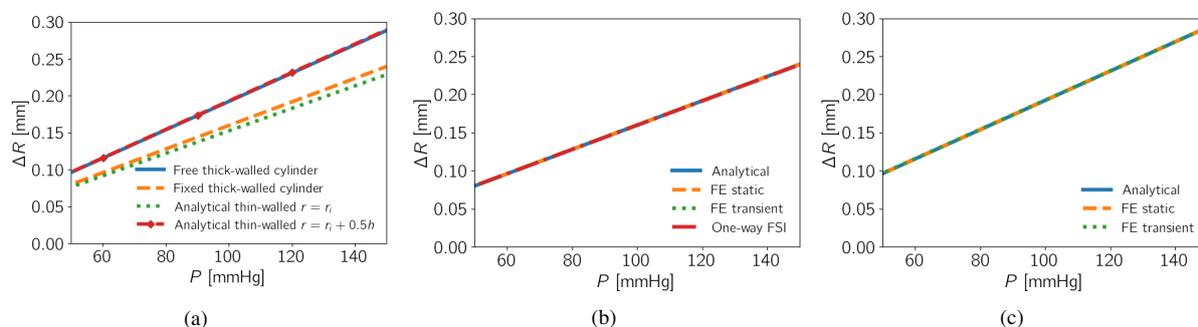


Figure 5: Comparison of analytical solutions of thin- and thick-walled cylinders with static and transient FE and one-way FSI inflation test. (a) analytical solutions with different underlying assumptions, (b) solution comparison with the assumption of fixed ends, and (c) solution comparison with the assumption of free ends.

displacement with reference to the minimum radius of the respective model evaluation predicted by the 1D-model at the mid-point of the artery. During early systole and around the diastolic notch is the prediction band for the pressure very close to the expected pressure values. The relative radial displacement shows large variations at peak systole.

Results from the comparison of the inflation tests with analytical solutions for thick- and thin-walled cylinders are presented in Fig. 5. In our analysis, static and transient inflation tests for the 3D-FE model of the artery coincide with the analytical solution for a cylinder with fixed and free ends. The radial displacement of the one-way FSI-simulation with fixed ends agrees with the analytical and FE solution. Fig. 5a depicts the analytical solutions visualizing the underlying assumption on the radial displacement ΔR . For a thick-walled cylinder with free ends, ΔR is higher than for a thick-walled cylinder with fixed ends. For the thin-walled simplification, Eq. 19, using the internal lumen radius results in the largest discrepancy between 1D and 3D; however, using the mid-wall radius results in a near agreement with the thick-walled theory and 3D results.

4. Summary and Discussions

In this work, we verified that the simulations of a 1D arterial model were grid independent. Subsequently, UQSA of the 1D arterial model was performed using PC expansion. This UQSA served a role in the validation process to focus on comparison of the 1D wall model, Eq. 3, with results from the 3D-FSI model. To this end, we verified the solution during inflation of the 3D-FSI model against thick-walled cylinder theory and evaluated the 1D arterial model predictions against the 3D-FSI model of the CCA.

The SA suggests that the most sensitive input parameters are related to the arterial wall mechanics, namely the wall thickness h , the Young's modulus E , and the lumen radius R . Small sensitivity indices were estimated for the compliance C which might be due to the relation of the downstream vasculature's compliance to arterial wall properties. Since variations in the fluid properties, the Poisson ratio, and the total resistance do not lead to variations in the quantities of interest, this result suggests that these parameters can be set to average population values in further simulations. This reduces the number of parameters which need to be explored and with this decrease computational expenses as well.

The SA findings motivated a detailed analysis of the structural part of the FSI-model. We showed agreement between the analytical thick-walled cylinder theory and the FE inflation tests indicating that the boundary conditions and material properties were set correctly in the simulation. Since the transient and static simulation results coincided, the inflation tests were not influenced by inertia effects. Furthermore, the consensus between the FSI simulation and the analytical solution assure correct coupling between the fluid and the solid domain.

The discrepancy between the analytical thin- and thick-walled cylinder with fixed ends radial displacement under internal pressure indicates that the wall model in the current 1D-model needs to be revised in order to lead to the same displacement as the 3D-FSI model.

5. Conclusions

With this work we showed that UQSA during the validation process can elucidate model characteristics where specific emphasis needs to be placed on during

the validation process. In our specific case, special focus needs to be placed on the wall mechanics because variations in the arterial wall input parameters lead to the largest variations in the model output.

In a next step, the FSI inflation test will be simulated with a two-way coupling and the 1D wall model will be advanced such that the average cylinder radius in the thin-walled theory is used for the computation instead of the inner radius. After successful completion, a physiological flow rate and pressure wave will be applied at the boundaries of the 3D-FSI model as well as at the 1D-model. The discrepancy between the model predictions will be quantified closing the validation process.

Acknowledgment

The authors would like to thank the Polish National Science Center GRIEG programme, (Grant no: UMO-2019/34/H/ST8/00624) for providing financial support to this project.

References

- [1] M. F. O'Rourke, C. O'Brien, and T. Weber, "Arterial stiffness, wave reflection, wave amplification: Basic concepts, principles of measurement and analysis in humans," in *Blood Pressure and Arterial Wall Mechanics in Cardiovascular Diseases* (M. E. Safar, M. F. O'Rourke, and E. D. Frohlich, eds.), pp. 3–13, Springer, 2014.
- [2] P. M. Nabeel, V. R. Kiran, J. Joseph, V. V. Abhidev, and M. Sivaprakasam, "Local pulse wave velocity: Theory, methods, advancements, and clinical applications," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 74–112, 2020. Conference Name: IEEE Reviews in Biomedical Engineering.
- [3] J. Alastruey, A. W. Khir, K. S. Matthys, P. Segers, S. J. Sherwin, P. R. Verdonck, K. H. Parker, and J. Peiró, "Pulse wave propagation in a model human arterial network: Assessment of 1-d visco-elastic simulations against in vitro measurements," *Journal of Biomechanics*, vol. 44, no. 12, pp. 2250–2258, 2011.
- [4] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis the primer*. John Wiley, 2008.
- [5] G. Ninos, V. Bartzis, N. Merlemis, and I. E. Sarris, "Uncertainty quantification implementations in human hemodynamic flows," *Computer Methods and Programs in Biomedicine*, vol. 203, p. 106021, 2021.
- [6] V. G. Eck, W. P. Donders, J. Sturdy, J. Feinberg, T. Delhaas, L. R. Hellevik, and W. Huberts, "A guide to uncertainty quantification and sensitivity analysis for cardiovascular applications," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 32, no. 8, p. e02755, 2016.
- [7] S. Sherwin, V. Franke, J. Peiró, and K. Parker, "One-dimensional modelling of a vascular network in space-time variables," *Journal of Engineering Mathematics*, vol. 47, no. 3, pp. 217–250, 2003.
- [8] C. A. Figueroa, I. E. Vignon-Clementel, K. E. Jansen, T. J. Hughes, and C. A. Taylor, "A coupled momentum method for modeling blood flow in three-dimensional deformable arteries," *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 41, pp. 5685–5706, 2006.
- [9] W. A. Riley, R. W. Barnes, G. W. Evans, and G. L. Burke, "Ultrasonic measurement of the elastic modulus of the common carotid artery. the atherosclerosis risk in communities (ARIC) study," *Stroke*, vol. 23, no. 7, pp. 952–956, 1992.
- [10] L. Formaggia, D. Lamponi, and A. Quarteroni, "One-dimensional models for blood flow in arteries," p. 26, 2003.
- [11] A. E. Anderson, B. J. Ellis, and J. A. Weiss, "Verification, validation and sensitivity studies in computational

- biomechanics,” *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 10, no. 3, pp. 171–184, 2007.
- [12] C. Sass, B. Herbeth, O. Chapet, G. Siest, S. Visvikis, and F. Zannad, “Intima–media thickness and diameter of carotid and femoral arteries in children, adolescents and adults from the stanislas cohort: effect of age, sex, anthropometry and blood pressure,” *J Hypertens*, vol. 16, no. 11, pp. 1593–1602, 1998. Place: Hagerstown, MD Publisher: Hagerstown, MD: Lippincott Williams & Wilkins, Inc.
- [13] M. R. Skilton, L. Bousset, F. Bonnet, S. Bernard, P. C. Douek, P. Moulin, and A. Serusclat, “Carotid intima–media and adventitial thickening: Comparison of new and established ultrasound and magnetic resonance imaging techniques,” *Atherosclerosis*, vol. 215, no. 2, pp. 405–410, 2011.
- [14] A. Karimi, T. Sera, S. Kudo, and M. Navidbakhsh, “Experimental verification of the healthy and atherosclerotic coronary arteries incompressibility via digital image correlation,” *Artery Research*, vol. 16, p. 1, 2016.
- [15] T. Kenner, “The measurement of blood density and its meaning,” *Basic Research in Cardiology*, vol. 84, no. 2, pp. 111–124, 1989.
- [16] C. Irace, C. Carallo, F. Scavelli, T. Esposito, M. S. De Franceschi, C. Tripolino, and A. Gnasso, “Influence of blood lipids on plasma and blood viscosity,” *Clinical Hemorheology & Microcirculation*, vol. 57, no. 3, pp. 283–290, 2014. Publisher: IOS Press.
- [17] N. Xiao, J. Alastruey, and C. A. Figueroa, “A systematic comparison between 1-d and 3-d hemodynamics in compliant arterial models,” *International Journal for Numerical Methods in Biomedical Engineering*, vol. 30, no. 2, pp. 204–231, 2014. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cnm.2598>.
- [18] W. Huberts, S. G. Heinen, N. Zonnebeld, D. A. van den Heuvel, J.-P. P. de Vries, J. H. Tordoir, D. R. Hose, T. Delhaas, and F. N. van de Vosse, “What is needed to make cardiovascular models suitable for clinical decision support? a viewpoint paper,” *Journal of Computational Science*, vol. 24, pp. 68–84, 2018.
- [19] P. H. Charlton, J. Mariscal Harana, S. Vennin, Y. Li, P. Chowienzyk, and J. Alastruey, “Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 317, no. 5, pp. H1062–H1085, 2019.
- [20] A. Inc., “Ansys FSI lecture trainee.” 2021.
- [21] V. Eck, J. Sturdy, and L. Hellevik, “Effects of arterial wall models and measurement uncertainties on cardiovascular model predictions,” *Journal of Biomechanics*, vol. 50, pp. 188–194, 2017.
- [22] F. Irgens, *Continuum mechanics*. Springer, 2008.

Numerical modelling of fin side heat transfer and pressure loss for compact heat recovery steam generators

Johan R. Espelund ^{a,b*}, Ole H. H. Meyer ^b, Geir Skaugen ^b

^a NTNU - Department of Energy and Process Engineering, ^b SINTEF Energy Research,
johan.espelund@sintef.no

Abstract

An optimization tool for offshore bottoming cycle and heat recovery steam generator (HRSG) design has previously been developed. The tool is based on empirical correlations to obtain hydraulic and thermal quantities for the HRSG. However, as these correlations are based on experiments with typical onshore designs, they may not be valid for the compact designs encountered in offshore HRSGs.

In order to extend the validity range of the optimization tool, this work presents a numerical model able to predict heat transfer and pressure loss in finned tube bundles by means of Computational Fluid Dynamics (CFD), utilizing a periodic domain to reduce computational costs. Both steady-state and transient models were applied, using the Spalart-Allmaras turbulence model, and their performance compared. To validate the model, results were compared with available experimental data, and then the model's performance was compared with a selected empirical correlation.

Three different fin-tube geometries were investigated (two serrated and solid) with varying tube layout angles. A parameterized grid generation tool was developed and used to generate grids for the selected geometries. The CFD results were found to be within 20 % of the experimental data, and were in most cases more accurate than the empirical correlation. The steady-state simulations did, however, not converge for the geometry with the largest layout angle. The steady-state framework should therefore be applied only to compact tube layouts. The transient simulations, though being computationally more intensive, are also able to model large layout angles.

1. Introduction

Oil and gas production contributes significantly to the global CO₂-emissions. In Norway, it is the industry that emits most greenhouse gases, accounting for 27% of CO₂-emissions from Norwegian territory in 2019. The largest contributor to these emissions are the gas turbines used for power generation offshore, which amounts to around 85% of these emissions. [1]

The large emissions of the gas turbines makes them an attractive candidate for emission reduction, and installing steam bottoming cycles has been proposed as a way to achieve this. Most of today's offshore power systems utilize the hot exhaust gases to some degree, e.g. for heating crude oil, but the heat lost to the atmosphere is still significant. With a steam bottoming cycle the heat is utilized for power generation using a steam generator, lowering the demand for power production from the gas turbines, and can reduce the turbine CO₂-emissions with as much as 25% [2]. One reason for why there is no widespread use of offshore steam bottoming cycles today, is weight and size limitations.

The heat recovery steam generator (HRSG) is a crucial and large component of the steam bottoming cycle. The limitations necessitate compact designs, and the once through steam generator (OTSG) has been found to be the most suitable HRSG type for offshore steam bottoming cycles [2]. One of the key factors when optimizing OTSGs for weight and size is small diameter heat exchanger tubes, compared with their onshore equivalents [3, 4].

The optimization procedures used in the design process rely on correlations for finned tube banks in order to

predict the heat transfer and pressure drop of the OTSG. Correlations are almost exclusively based on empirical data, and their region of validity is therefore limited to the range of experiments that they are based on. This has proved to be a challenge for the optimization of the compact offshore OTSG designs, which have fin and tube geometries outside the validity range of the correlations. The result is that different correlations tend to give significantly different predictions when compared to the same experimental data sets. Holfeld [5] reported up to 77% spread between the correlations for heat transfer, and up to 410% for pressure drop when comparing different correlations to the same experimental data.

Ideally, new correlations would be developed based on experiments that are performed under conditions close to those expected for the offshore OTSGs. However, performing experiments to produce enough data for new correlation development is both costly and time-consuming, and therefore Computational Fluid Dynamics (CFD) is proposed as a way to predict heat transfer and pressure drop in compact OTSGs, and to validate the designs produced by the optimization procedure.

Numerical simulations of finned tube banks, and particularly serrated fins, have only started to emerge in the course of the last two decades, as CFD has become more widespread in the engineering field and the computational power available to researchers and engineers has increased. The earliest works on the field were limited to solid annular fins, with few tube rows being modelled. Jang et al. [6] performed both experiments and numerical modelling of

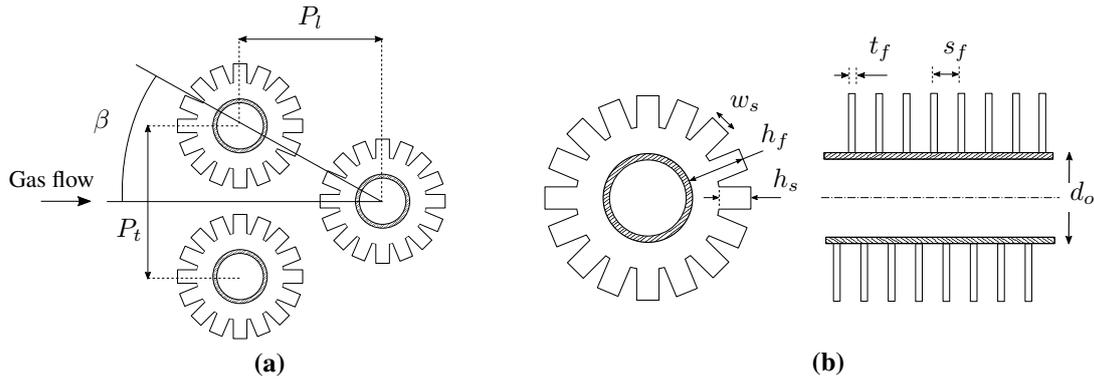


Figure 1: (a) Tube layout parameters. (b) Fin and tube geometry parameters

laminar flow through 4 tube rows in a staggered layout. The numerical model was able to accurately predict pressure drop, but over-estimated heat transfer by 20 - 30%.

Mon [7] performed turbulent simulations for 23 different tube layouts, both in-line and staggered, with the number of tube rows ranging from 2 to 6. The model was able to qualitatively describe the flow, but no comparison with experimental data was done. New correlations based on the numerical results and available experimental data were proposed. In a later paper, Mon & Gross [8] compared the results from selected layouts with existing correlations, where they were found to agree from $\pm 15\%$ to over 50%. Torresi et al. [9] were among the first to model flow in tube banks with serrated fins. They simulated only one tube row, without modelling heat transfer, and then used equivalent porous medium zones in order to model the full HRSG. The results were not compared with experimental data, but showed good agreement with a proprietary 1D code.

McIlwian [10] compared the performance of solid and serrated fins in a single tube row, and gave qualitative insight into how serrated fins improve heat transfer compared with solid fins. In a later study [11], McIlwian extended the model and looked at the effects of adding a 2nd, 3rd and 4th row. The results were compared with correlations, but neither of the studies were validated with experimental data.

As with Jang et al., laminar flow was also assumed by Lemouedda et al. [12], where fin tube bundles for Re between 600 and 2 600 were investigated. Fins with and without serration were compared, and the effect of twisting of the serrated fins was also investigated. No comparisons with experiments were made.

Hofmann & Walter [13] performed simulations and experiments for both solid and serrated fins, with both helical and angular fin attachment. Both local and overall heat transfer and pressure drop was investigated for turbulent flow with Re ranging from 3500 to 50 000. Results showed good agreement of the simulations with experimental data, being within $\pm 15\%$ uncertainty.

Ó Cléirigh and Smith [14] investigated the effects of degree of serration, modelling fully serrated, partially serrated and solid fins. They found that the Nusselt number increased with 23% from partially to fully serrated fins, a distinction that is not made in most correlations. However, no validation against experimental data was performed in this study either.

Where the previous studies all have used standard inlet-outlet boundary conditions in the stream wise direction, Martinez et al. [15] utilized periodic boundary conditions also in this direction, thus assuming fully periodic flow. Local flow features were compared with experimental

measurements, but global heat transfer and pressure drop were only compared with selected correlations, though with good agreement. In a consecutive paper [16], they went on to model six tube rows using standard inlet-outlet boundary condition, and showed that the velocity, temperature and turbulence fields indeed display periodic behaviour after the third tube row.

Lindqvist & Næss [17] also used a periodic domain model, and applied it to four different cases, both serrated and solid fins. For one of them, a full domain model with eight tube rows was also considered, which was shown to match very closely with the periodic domain model. All four cases were validated against experimental data, and were found to be within 15% for both heat transfer and pressure drop. In addition, three correlations and two fin efficiency corrections were compared with the numerical and experimental results and their performances assessed. In a consecutive study [18], the authors investigated vortex-shedding frequencies by performing transient simulations with the same model, not including heat transfer. Both studies were limited to geometries with compact tube layout angles of 30° .

The aim of the present work is to develop a numerical model that is able to readily and accurately predict heat transfer and pressure drop for a range of geometrical parameters that is representative of compact heat exchangers. Both serrated and solid fins will be considered, and will be restricted to helical fin attachment. Both steady-state and transient simulations will be carried out and their performance assessed. To model a wider range of fin-tube geometries, the present work will use geometries with different layout-angles. The model will be validated with available experimental data and compared with an empirical correlation.

2. Selected geometries

Solid and serrated fin tubes are characterized by a fixed set of parameters, illustrated in Fig. 1b, with the solid fin being a special case of the serrated fin ($h_s = 0$). The tube bundle layout is characterized by the longitudinal and transverse pitches P_l and P_t , respectively, or the tube layout angle $\beta = \arctan\left(\frac{P_t}{2P_l}\right)$, as shown in Fig. 1a.

The three tube and fin geometries considered in this study, listed in Tab. 1, are selected in order to span a sufficiently wide range of parameters to represent the possible designs encountered in compact heat exchangers. The first two geometries, N1 and N2, are geometries 1 and 2 from the experimental study by Næss [19], respectively. The two cases from Næss are using the same serrated fin tube geometry, but with different layout angles. Næss showed that correlations tended to perform poorly for serrated

Table 1: Layout, fin and tube geometries that are modelled.

Geometry Type	H8	N1	N2
P_t [mm]	38.7	46.1	65.2
P_l [mm]	33.5	39.9	32.6
d_o [mm]	13.5	20.89	20.89
h_f [mm]	10	8.61	8.61
h_s [mm]	-	8.61	8.61
s_f [mm]	2.81	5.08	5.08
t_f [mm]	0.50	0.91	0.91
w_f [mm]	-	3.97	3.97
β [°]	30	30	45

geometries with layout angles deviating from 30°, and it is therefore desirable to investigate the predictive accuracy of the numerical model in this case. The third case, H8, is based on geometry 8 from the experimental study by Holfeld [5], and is a solid finned tube with the same layout angle as N1 (30°).

3. Numerical method

The open-source CFD library OpenFOAM v2112 was used to solve the coupled equations and SALOME v6.7.0 was used for the grid generation.

3.1. Domain

The numerical domain consists of two regions: a fluid region (gas) and a solid region (fins), as conduction in the fins is also modelled. Note that only the outer surface of the tube is modelled, *not* the tube wall itself. The numerical

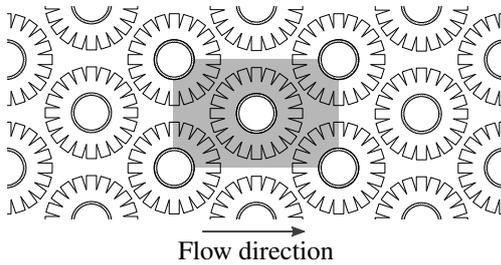


Figure 2: The numerical domain (shaded in grey) inside an infinite tube bank, with flow direction from left to right.

domain is based on the works of Maritz et al. [15] and Lindqvist & Næss [17, 18], and can be thought of as a "unit cell" within an infinite tube bank, as illustrated in Fig. 2. This implies periodicity in all directions, and reduces the computational requirements significantly, compared to domains that include several tube rows.

3.2. Governing equations

The incompressible continuity, momentum and energy equations are solved in the gas region. The continuity equation reads

$$\nabla \cdot \mathbf{u} = 0, \quad (3.1)$$

where \mathbf{u} is the velocity field. Cyclic boundary conditions in the stream wise direction necessitate the addition of a source term in the momentum equations to drive the flow, as demonstrated by Patankar & Liu [20], and takes the form,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = \mathbf{S} - \frac{1}{\rho} \nabla p + \nabla \cdot \left[\nu_{\text{eff}} \left(\nabla \mathbf{u} + (\nabla \mathbf{u})^T \right) \right], \quad (3.2)$$

where p is the pressure field, ρ is the density and ν_{eff} is the effective kinematic viscosity. The source term \mathbf{S} acts as an imposed pressure gradient to drive the flow through the domain. The energy equation is formulated using the specific enthalpy, viz.

$$\frac{\partial(\rho h + e_K)}{\partial t} + \nabla \cdot (\mathbf{u}(\rho h + e_K)) - \frac{Dp}{Dt} = -\nabla \cdot \mathbf{q} \quad (3.3)$$

where the heat flux is given by Fourier's law $\mathbf{q} = \rho \alpha_{\text{eff}} \nabla h$, where $\alpha_{\text{eff}} = \kappa_{\text{eff}} / (\rho c_p)$ is the effective thermal diffusivity and $e_K = \frac{1}{2} \rho |\mathbf{u}|^2$ is the specific kinetic energy. The energy equation is also solved in the solid region, being a special case of Eq. (3.3) by setting $\mathbf{u} = \mathbf{0}$, $Dp/Dt = 0$ and use ρ , c_p and κ for the solid.

3.3. Choice of turbulence model

The model of choice in the present study is the *Spalart-Allmaras* turbulence model [21], which has shown good performance for flow over finned tube bundles [17, 22]. Being a one-equation model, where a transport equation for the modified turbulent viscosity $\tilde{\nu}$ is solved, it is computationally advantageous to the more common two-equation models.

3.4. Thermal properties

The unit cell allows for the assumption of constant thermal properties, as the temperature differences will be moderate since only one tube row is considered. As a consequence, the incompressibility assumption is used in the governing equations.

Both the fluid and solid regions are modelled using constant thermal properties, given in Tab. 2. The fluid is modelled as dry air at atmospheric pressure and 300 K, which matches the experimental conditions closely. The fins are modelled as carbon steel and aluminium 6060 for the Næss (N1 and N2) and Holfeld (H8) cases, respectively.

3.5. Grid generation

One of the main tasks has been to develop a parameterized grid generation procedure able to produce quality grids from a given set of geometry parameters, e.g. the result from an HRSG optimization. Lindqvist [17] had a similar approach, but limited the tube layout angle to $\beta = 30^\circ$. The present grid generation procedure is not limited to one layout angle, but rather lets β be a free variable. Fig. 3 shows one of the grids used (N2), with $\beta = 45^\circ$.

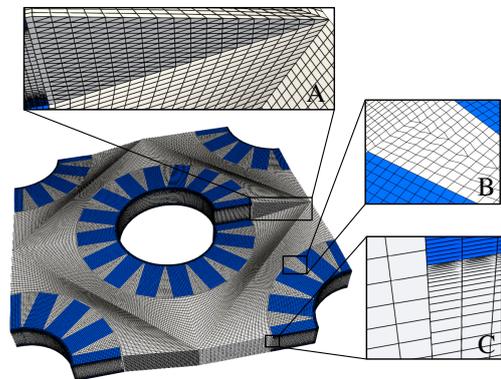


Figure 3: The numerical grid for geometry N2, showing the gas (■) and fin (■) regions.

The grid is dominated by hexagonal cells, though this is not achieved in the serrated regions, where prismatic wedge cells also are included, shown in box B in Fig.

Table 2: Thermal properties for dry air, carbon steel and aluminium A6060 at atmospheric pressure and 300 K [23, 24].

Property	Dry air	Carbon Steel	Aluminium	Unit
Thermal conductivity λ	0.0263	48.5	210	W/(m·K)
Specific heat capacity c_p	1007	434	898	J/(kg·K)
Density ρ	1.1614	7854	2700	kg/m ³
Viscosity ν	$15.89 \cdot 10^{-6}$	-	-	kg/(m·s)

3. Hexagonal cells are desired in order to maximize the accuracy of solution, and consideration was also given to skewness and cell growth to ensure a smooth grid.

The boundary layer is resolved in the inter-fin regions and on the tube surface, as shown in box C in Fig. 3. Ensuring a first cell height of $y^+ < 1$ yields a resolved turbulent boundary layer in these regions. The boundary layer cells are set to a growth rate of 20 % with a smooth transition to the inter-fin region. On the fins sides and ends however, the boundary layer was not resolved, and the turbulent boundary layer is modelled using Spalding's unified wall function [25]. At the interface between the bulk grid and the inter-fin grid, polyhedral cells are used to make the two grid regions conform. This can result in skewed faces, but is always kept within the grid criterion set in OpenFOAM. Polyhedral cells are also used in the cut-plane directly downstream of the tube, shown in box A in Fig. 3. This is a consequence of the helix angle of the fins, and results in wedge-shaped faces in order to make the periodic boundaries conform, shown in box A in Fig 3.

3.6. Boundary conditions

The fin and tube surfaces are prescribed the no-slip and no penetration boundary conditions for velocity, i.e. $\mathbf{u}_w = \mathbf{0}$ (where the subscript w denotes the quantity evaluated at the wall) and zero gradient for pressure, $\partial p_w / \partial n = 0$.

For the modified turbulent viscosity $\tilde{\nu}$, the wall boundary conditions should also be zero, i.e. $\tilde{\nu}_{t,w} = 0$. This is done on all walls except for the fin sides, where the boundary layer is not resolved. Here, wall functions are used for the turbulent viscosity directly to impose the theoretical turbulent boundary layer profiles onto the flow field. This does not yield as accurate results as the fully resolved boundary layers do, but as these constitute only a minor part of the total wall area, the use of wall functions here is deemed acceptable.

The temperature at the tube surface and fin bases is set to a uniformly fixed temperature $T_w = 300$ K at the tube walls and the base of the fins. At the interface between the gas and fins, the boundary conditions are set to a conserved heat flux through the interface as well as identical temperature for both regions, viz.

$$T_{w,\text{fluid}} = T_{w,\text{solid}}, \quad \mathbf{q}_{w,\text{fluid}} = -\mathbf{q}_{w,\text{solid}}. \quad (3.4)$$

All but the wall surfaces are periodic boundaries, and for velocity and pressure stream wise periodicity is implemented in Eq. (3.2), and the modified turbulent viscosity $\tilde{\nu}$ is also assumed to be fully periodic. The temperature is also cyclic between inlet and outlet, but an offset is prescribed to account for the temperature drop over the tubes. The offset is defined such that the left cyclic boundary (the inlet) $T_{\text{inlet}}(\mathbf{x})$ always is kept at a constant average temperature $T_{\text{in}} = 320$ K while at the same time maintaining the constant wall temperature $T_w = 300$ K,

$$T_{\text{inlet}}(\mathbf{x}) = T_w + \left(\frac{T_{\text{in}} - T_w}{T_{\text{out}} - T_w} \right) (T_{\text{outlet}}(\mathbf{x}) - T_w), \quad (3.5)$$

where $T_{\text{outlet}}(\mathbf{x})$ is the temperature field at the right cyclic boundary (the outlet) and T_{out} is the average temperature at the outlet.

For the steady-state simulations, the stream wise velocity boundary conditions are implemented by mapping the velocity and pressure fields between inlet and outlet, as were proposed by Lindqvist & Næss [17]. This improves the stability of the simulation and makes it easier to arrive at the steady-state solution. The inlet boundary is initially prescribed a uniform profile with a fixed mass flux and a zero gradient pressure, whereas the outlet is zero-gradient velocity and fixed pressure boundary condition. The velocity field is then mapped from outlet to inlet after every 1000 iterations, while the pressure field is mapped from inlet to outlet and scaled to ensure a constant pressure at the outlet boundary. This mapping is repeated until the simulation converges, which is assessed by total heat transfer and pressure drop over the central tube.

For the transient simulations, the inlet and outlet boundaries were cyclic for both the velocity and pressure fields, which is achieved by a momentum source term \mathbf{S} in Eq. (3.2), acting as an imposed pressure gradient. The pressure gradient was fixed for each simulation and corresponded to a Reynolds number. To estimate the pressure gradient to achieve a desired Reynolds number, the Weierman correlation [26] was used. The transient simulations are carried out for 20 domain flow-through cycles to ensure fully developed flow conditions.

3.7. Discretization and solution algorithm

All convective terms were discretized with the linear upwind scheme, and a linear blend between Euler (0.3) and the Crank-Nicolson (0.7) scheme was used for transient terms, giving 2nd order accuracy in space and 1st-2nd order in time.

The discretized equations were first solved with a steady state solver using the SIMPLE algorithm. The PISO algorithm was used to perform the transient simulations, with adaptive time stepping ensuring that $Co < 0.5$ for all cells in the grid. To accelerate the transient simulations, the initial conditions were obtained by carrying out steady simulations for 1000 iterations.

3.8. Data reduction

To allow for direct comparison with experimental results, non-dimensional parameters are calculated by integrating the raw numerical data and normalizing with appropriate characteristic scales.

3.8.1. Reynolds number

The Reynolds number that characterizes the flow field is defined as

$$Re = \frac{d_o u_{F_{\text{min}}}}{\nu}, \quad (3.6)$$

where $u_{F_{\text{min}}}$ is the average velocity in the minimum free flow area. In the steady-state simulations, $u_{F_{\text{min}}}$ is fixed by the inlet velocity, whereas for the transient simulations it is obtained by sampling the mean velocity at the inlet and time-averaging.

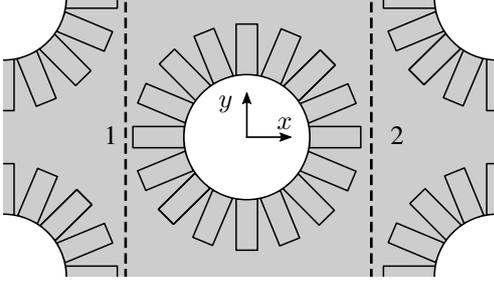


Figure 4: Sampling planes for temperature and pressure.

3.8.2. Euler number

The pressure drop is normalized by the dynamic pressure to form the Euler number,

$$Eu = \frac{\Delta p}{\frac{1}{2} \rho u_{F_{min}}^2} \quad (3.7)$$

where Δp is the pressure drop across one tube row. For the steady-state simulations, Δp is obtained by sampling the pressure in front of and behind the central tube (at $x = \pm P_l/2$). For the transient simulations, Δp is obtained by multiplying the momentum source from Eq. (3.2) with the longitudinal tube pitch P_l .

3.8.3. Nusselt number

The outside heat transfer coefficient α_o is defined as,

$$\alpha_o = \frac{\dot{Q}_{tot}}{[\eta_f A_f + A_t] \Delta T}, \quad (3.8)$$

where \dot{Q}_{tot} is the total heat transferred to both the fin and tube surface, η_f is the fin efficiency and A_f and A_t is the surface area of the fin and bare tube surfaces, respectively. ΔT is the average temperature difference that drives the heat transfer between the gas and the surface of the finned tube. For the cyclic domain with only one tube row, a local arithmetic average temperature is used to approximate the mean temperature difference,

$$\Delta T = \frac{1}{2} [(T_{b,1} - T_w) + (T_{b,2} - T_w)], \quad (3.9)$$

where $T_{b,1}$ and $T_{b,2}$ are the average bulk fluid temperatures on sampling planes located at $x = \pm P_l/2$, respectively (see Fig. 4).

In experimental studies, the fin efficiency is usually estimated using theoretical and corrected predictions, and therefore the calculated heat transfer coefficient will depend on the chosen fin efficiency calculation. On the other hand, CFD results provides a full description of the temperature field, which allows for direct computation of the actual fin efficiency. However, to compare with the experiments by Næss [19] and Holfeld [5] on a consistent basis, the corrected and theoretical approaches will be used here.

$$\eta_{th} = \frac{\tanh[m(l_e + t_f/2)]}{m(l_e + t_f/2)}, \quad (3.10)$$

where $m = \sqrt{\frac{2\alpha_o(t_f + w_f)}{k_f \cdot t_f \cdot w_f}}$.

Due to the non-uniform distribution of the heat transfer coefficient, Weierman [26] proposed the following correction

$$\eta_f = \eta_{th} \cdot (0.9 + 0.1 \cdot \eta_{th}), \quad (3.11)$$

for serrated fin tube bundles. This correction is used by Næss [19], i.e. for cases N1 and N2. For solid fins, the theoretical fin efficiency is given as,

$$\eta_{th} = C \frac{I_1(mr_f) K_1(mr_o) - I_1(mr_o) K_1(mr_f)}{I_0(mr_o) K_1(mr_f) + I_1(mr_f) K_0(mr_o)}, \quad (3.12)$$

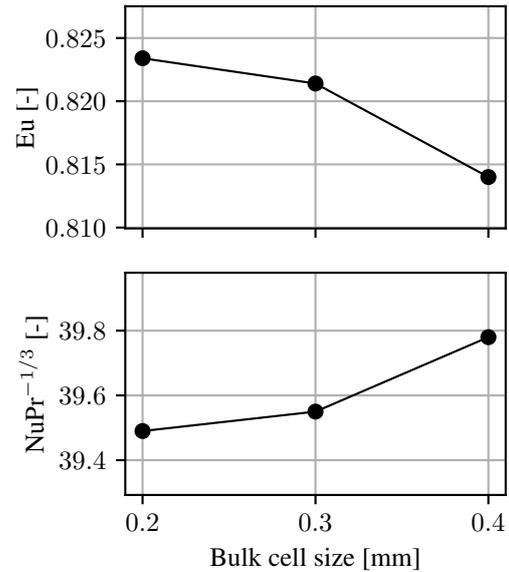
where $C = \frac{2r_o}{m(r_f^2 - r_o^2)}$ and I_n and K_n are the modified Bessel functions of first and second kind, respectively. $r_o = d_o/2$ is the tube outside radius and $r_f = r_o + h_f$ is the fin radius.

The heat transfer coefficient α_o is then normalized with d_o and κ to yield the Nusselt number, defined as

$$Nu = \frac{d_o \alpha_o}{\kappa}. \quad (3.13)$$

3.9. Grid Refinement Study

A grid refinement study was performed using the steady-state solution method on geometry H8 at $Re = 5000$. Three different grid resolutions were investigated, where the bulk mesh was refined, and the boundary layer mesh was kept at a constant y^+ and cell growth ratio. Both Eu and $NuPr^{-1/3}$ were used as integral parameters. The results from the grid refinement study can be seen in Fig. 5.

Figure 5: Results from the grid refinement study, where Eu and $NuPr^{-1/3}$ are used as integral parameters.

4. Results and Discussions

The numerical results from case H8 are shown in Fig. 6, and are compared with the experimental data from Holfeld [5] and the Weierman [26] correlation. The steady-state results are within 15% for both heat transfer and pressure drop, whereas the transient results are within 15% for pressure drop and 20% for heat transfer. All numerical results are predicting better than the Weierman correlation. Though both solution methods match the experimental results closely, it is worth noting that the transient results, which are computationally more intensive, are generally performing poorer than the steady-state simulations.

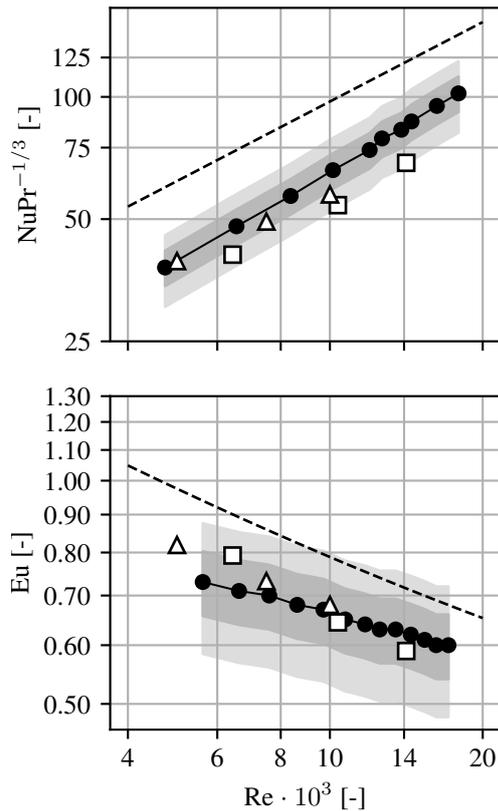


Figure 6: Results for geometry H8. Steady-state (Δ) and transient (\square) CFD results compared with the experimental results (\bullet) from Næss [19] and the Weierman [26] correlation (---). The dark and light shaded areas denote $\pm 10\%$ and $\pm 20\%$ deviation from the experimental results, respectively.

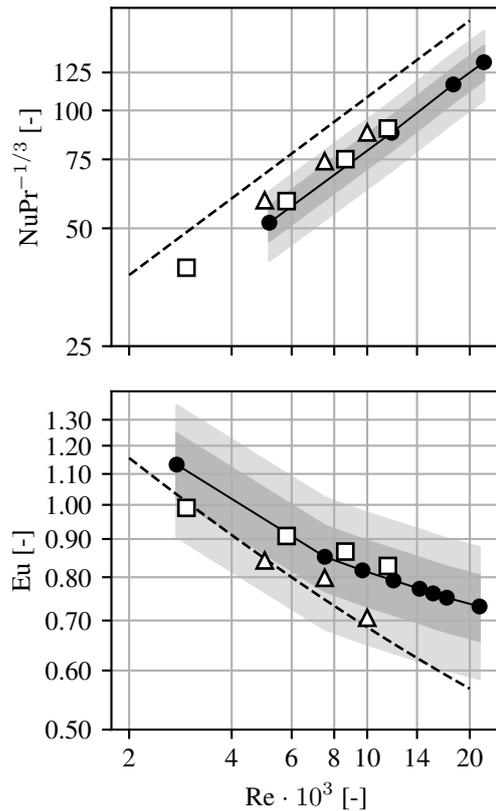


Figure 7: Results for geometry N1. Steady-state (Δ) and transient (\square) CFD results compared with the experimental results (\bullet) from Næss [19] and the Weierman [26] correlation (---). The dark and light shaded areas denote $\pm 10\%$ and $\pm 20\%$ deviation from the experimental results, respectively.

The numerical results from case N1 are shown in Fig. 7, and are compared with the experimental data from Næss [19] and the Weierman correlation. The steady-state results are within 20% for both heat transfer and pressure drop, whereas the transient results are within 10% for pressure drop and heat transfer. It is evident that the transient simulations are performing better in this case, particularly when considering the pressure drop. Eu from the steady-state simulations follows the Weierman correlation – which diverges from experiments at higher Re – closely, whereas the transient simulation matches the experimental results, but with larger deviation for the lowest Re .

The numerical results from case N2 ($\beta = 45^\circ$) are shown in Fig. 9, and are compared with the experimental data from Næss [19] and the Weierman correlation. The steady-state results did not converge, and thus only transient results are presented from this case. The convergence issues are due to the backflow at the outlet boundary. This occurs because the layout angle is so large that the wake extends beyond the outlet, resulting in unphysical and unstable behaviour. Fig. 8 shows how the wake is being limited by the zero-backflow boundary condition at the outlet. This problem is not encountered in the transient cases, where actual cyclic boundary conditions are used, and not mapping between inlet and outlet. Cyclic boundary conditions were also tested on the steady-state cases for all geometries, but all simulations displayed unstable behaviour and no converged results were obtained. The transient heat transfer results show good agreement with experimental data, being within 10% and performing significantly better than the Weierman

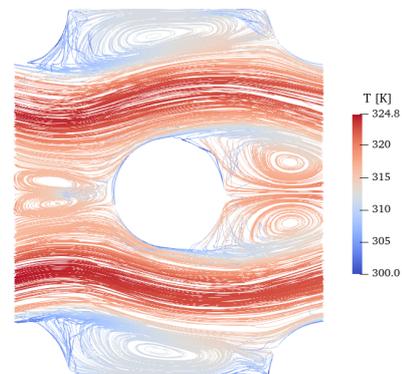


Figure 8: Streamlines coloured by temperature at $Re = 10\,000$ for geometry N2, from steady-state simulation with mapped inlet and outlet.

correlation. The pressure drop results are within 20% for all Re , but are more accurate for lower Re . The Weierman correlation displays similar accuracy. The inability of the steady-state solution method to model high layout angles is not ideal, as the transient method is computationally far more demanding and time-consuming than the steady-state approach. As the compact tube layout angle is the most used, especially for compact HRSGs, this will not be an issue in most practical cases. However, in the general case, where optimized designs may have a higher layout angle, the numerical model will not provide results as rapidly as for compact layouts.

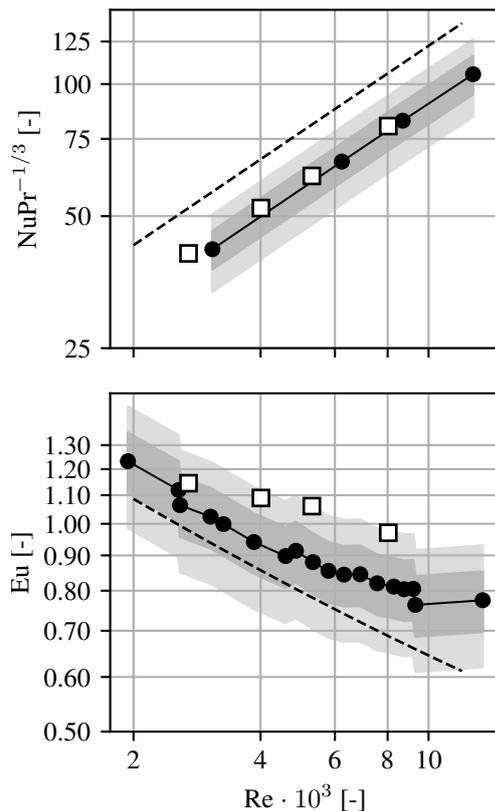


Figure 9: Results for geometry N2. Transient (\square) CFD results compared with the experimental results (\bullet) from Næss [19] and the Weierman [26] correlation ($- - -$). The dark and light shaded areas denote $\pm 10\%$ and $\pm 20\%$ deviation from the experimental results, respectively.

5. Conclusion and further work

In this study, a numerical model has been used to predict pressure drop and heat transfer in fin tube banks. A parameterized grid generation procedure was developed and used to generate grids for three different fin-tube geometries, representable for compact Heat Recovery Steam Generators. A combination of both solid and serrated fins were considered, where the serrated geometry were investigated at two different layout angles. Steady-state and transient CFD-simulations were performed on all geometries, and available experimental data was used to validate the numerical results in addition to comparison with an empirical correlation.

- All simulations were within $\pm 20\%$ of experimental data for both pressure drop and heat transfer, except for the non-converging steady-state simulations of the geometry with the largest tube layout angle.
- For compact layout angles ($\beta = 30^\circ$), steady-state simulations should be used, as they require less CPU-hours than the transient solution method. For larger layout angles, transient methods must be employed. Further work is needed to find the critical layout angles where
- To ensure sufficiently small time-steps for the transient simulations, a temporal convergence study should be performed in addition to spatial (grid) convergence studies for all geometries.
- Though being computationally more intensive, the transient solution method can also be used to perform

vibrational analyses of fin tube banks, in addition to predicting pressure drop and heat transfer.

Acknowledgment

This article is based on a Master's thesis from NTNU with computations performed at the IDUN [27] cluster. This work has been funded by HighEFF - Centre for an Energy Efficient and Competitive Industry for the Future, an 8 year Research Centre under the FME-scheme (Centre for Environment-friendly Energy Research, 257632/E20). The authors gratefully acknowledge the financial support from the Research Council of Norway and user partners of HighEFF.

References

- [1] "Klima- og miljørapport," tech. rep., Norsk Olje og Gass, 2021.
- [2] L. Nord and O. Bolland, "Steam bottoming cycles offshore – Challenges and possibilities," *Journal of Power Technologies*, vol. 92, pp. 201–207, Jan. 2012.
- [3] R. M. Montañés, G. Skaugen, B. Hagen, and D. Rohde, "Compact Steam Bottoming Cycles: Minimum Weight Design Optimization and Transient Response of Once-Through Steam Generators," *Frontiers in Energy Research*, vol. 9, p. 687248, June 2021.
- [4] M. J. Mazzetti, B. A. L. Hagen, G. Skaugen, K. Lindqvist, S. Lundberg, and O. A. Kristensen, "Achieving 50% weight reduction of offshore steam bottoming cycles," *Energy*, vol. 230, p. 120634, Sept. 2021.
- [5] A. Holfeld, *Experimental investigation of heat transfer and pressure drop in compact waste heat recovery units*. Doctoral theses, Norwegian University of Science and Technology, Trondheim, June 2016.
- [6] J.-Y. Jang, J.-T. Lai, and L.-C. Liu, "The thermal-hydraulic characteristics of staggered circular finned-tube heat exchangers under dry and dehumidifying conditions," *International Journal of Heat and Mass Transfer*, vol. 41, pp. 3321–3337, Nov. 1998.
- [7] M. S. Mon, *Numerical Investigation of Air-Side Heat Transfer and Pressure Drop in Circular Finned-Tube Heat Exchangers*. PhD thesis, Technischen Universität Bergakademie Freiberg, Freiberg, Germany, Feb. 2003.
- [8] M. S. Mon and U. Gross, "Numerical study of fin-spacing effects in annular-finned tube heat exchangers," *International Journal of Heat and Mass Transfer*, vol. 47, pp. 1953–1964, Apr. 2004.
- [9] M. Torresi, A. Saponaro, S. Camporeale, and B. Fortunato, "CFD Analysis of the Flow Through Tube Banks of HRSG," (Berlin, Germany), pp. 1–11, 2008.
- [10] S. R. McIlwain, "A Comparison of Heat Transfer Around a Single Serrated Finned Tube and a Plain Finned Tube," *International Journal of Research and Reviews in Applied Sciences*, vol. 2, pp. 88–94, Feb. 2010.
- [11] S. R. McIlwain, "A CFD Comparison of Heat Transfer and Pressure Drop Across Inline Arrangement Serrated Finned Tube Heat Exchangers With an Increasing Number of Rows," *International Journal of Research and Reviews in Applied Sciences*, vol. 4, pp. 162–169, Aug. 2010.
- [12] A. Lemouedda, A. Schmid, E. Franz, M. Breuer, and A. Delgado, "Numerical investigations for the optimization of serrated finned-tube heat exchangers," *Applied Thermal Engineering*, vol. 31, pp. 1393–1401, June 2011.
- [13] R. Hofmann and H. Walter, "Experimental and Numerical Investigation of the Gas Side Heat Transfer and Pressure Drop of Finned Tubes—Part II: Numerical Analysis," *Journal of Thermal Science and Engineering Applications*, vol. 4, p. 041008, Dec. 2012.
- [14] C. T. Ó Cléirigh and W. Smith, "Can CFD accurately predict the heat-transfer and pressure-drop performance of finned-tube bundles?," Dec. 2014. Publisher: Elsevier.

- [15] E. Martinez, W. Vicente, M. Salinas-Vazquez, I. Carvajal, and M. Alvarez, "Numerical simulation of turbulent air flow on a single isolated finned tube module with periodic boundary conditions," *International Journal of Thermal Sciences*, vol. 92, pp. 58–71, June 2015.
- [16] E. Martinez-Espinosa, W. Vicente, M. Salinas-Vazquez, and I. Carvajal-Mariscal, "Numerical Analysis of Turbulent Flow in a Small Helically Segmented Finned Tube Bank," *Heat Transfer Engineering*, vol. 38, pp. 47–62, Jan. 2017. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/01457632.2016.1156396>.
- [17] K. Lindqvist and E. Naess, "A validated CFD model of plain and serrated fin-tube bundles," *Applied Thermal Engineering*, vol. 143, pp. 72–79, Oct. 2018.
- [18] K. Lindqvist and E. Naess, "NUMERICAL MODELING OF VORTEX SHEDDING IN HELICALLY WOUND FINNED TUBE BUNDLES IN CROSS FLOW," in *International Heat Transfer Conference 16*, (Beijing, China), pp. 1843–1850, Begellhouse, 2018.
- [19] E. Naess, "Experimental investigation of heat transfer and pressure drop in serrated-fin tube bundles with staggered tube layouts," *Applied Thermal Engineering*, vol. 30, pp. 1531–1537, Sept. 2010.
- [20] S. V. Patankar, C. H. Liu, and E. M. Sparrow, "Fully Developed Flow and Heat Transfer in Ducts Having Streamwise-Periodic Variations of Cross-Sectional Area," *Journal of Heat Transfer*, vol. 99, pp. 180–186, May 1977.
- [21] P. Spalart and S. Allmaras, "A one-equation turbulence model for aerodynamic flows," in *30th Aerospace Sciences Meeting and Exhibit*, (Reno, NV, U.S.A.), American Institute of Aeronautics and Astronautics, Jan. 1992.
- [22] H. Nematy and M. Moghimi, "Numerical Study of Flow Over Annular-Finned Tube Heat Exchangers by Different Turbulent Models," vol. 6, p. 13, 2014.
- [23] F. Incropera, D. DeWitt, T. Bergman, and A. Lavine, *Principles of Heat and Mass Transfer*. John Wiley & Sons Singapore Pte. Limited, 2013.
- [24] "Information on Aluminium 6060 - Thyssenkrupp Materials (UK)."
- [25] D. B. Spalding, "A Single Formula for the "Law of the Wall"," *Journal of Applied Mechanics*, vol. 28, pp. 455–458, Sept. 1961.
- [26] C. Weierman, "Correlations Ease the Selection of Finned Tubes," *Oil and Gas Journal*, pp. 94–100, Sept. 1976.
- [27] M. Sjalander, M. Jahre, G. Tufte, and N. Reissmann, "EPIC: An Energy-Efficient, High-Performance GPGPU Computing Research Infrastructure," *arXiv:1912.05848 [cs]*, Sept. 2021. arXiv: 1912.05848.

Checking data informativity as the first step in data-driven modeling – case study

Amir Farzin ^{a,*}, Kateryna Rabchuk ^b, Bernt Lie ^a, Nils-Olav Skeie ^a

^a *Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway,*

^b *Process Modeling and Control Department, Yara International ASA.*

{amir.farzin, bernt.lie, nils-olav.skeie}@usn.com, kateryna.rabchuk@yara.com

Abstract

This paper reviews and introduces the strategies for testing a given dataset sampled from an unknown dynamic process to determine if it is sufficiently informative to model the system's behavior. The presented tests should be done as the first step in data-driven modeling to avoid an endless search for a proper model which may not exist based on the available data. It is unrealistic that available data holds complete information about the system at hand. The tests also allow us to estimate how good the established model can be. Finally, the presented methodologies are applied to an actual process as the case study: modeling the decarbonization section in an ammonia plant.

1. Introduction

By model, we will mean anything for which an experiment can be used to answer questions about the system [1]. Modeling is the act of developing a model. In this definition, the model can be a physical instance of the system or a mathematical representation of the system. The latter is what the model means in this paper.

In many engineering cases, modeling is the first step before other analysis techniques. Therefore, the quality of the model directly affects the solution of the final problem by putting an upper bound on its quality [2]. This fact makes modeling the bottleneck of many engineering problems and raises the need for putting more effort into finding high-quality models.

Data-driven modeling is a rapidly evolving field with great potential to transform engineering science [3]. The concept of data-driven modeling contrasts with physics-based modeling. In the former methodology, the data is the core element that illustrates the behavior and expresses the properties of the regarded phenomenon or object. In data-driven modeling, the scientist does not need to know the underlying physical interactions. These physical interactions form the basis for and needs to be known in physics-based modeling. In the case of data-driven modeling, the relationship between a given set of available measurements (i.e., model input variables or features) and desired behaviors or values (i.e., model output variables or targets) is called the model of the process/system. The

construction of a model for a process involves three basic entities:

1. Dataset
2. Model structure
3. Rules for assessing the model from data

The dataset comes first among the items mentioned above, indicating its importance and fundamental effects on the other two entities. The data can be recorded based on designing a proper experiment (e.g., [4]). However, it is not always possible to affect the experiment, and historical data from the plant's operation must be used. Therefore, although the informativity of the data during experiment design is guaranteed, while dealing with a historical dataset, the informativity of data should be evaluated. In addition, the historical data often includes many measurements where only a few are useful for building a model. Therefore, feature selection is a way of reducing the size of the dataset by removing non-informative variables.

This paper presents a method for determining (1) whether the available data is informative enough to develop models and (2) which features contain information about the target variable. The novelty of the paper comes from the informativity of the data. Without considering the fact that the data is informative, one may spend weeks/months to find a model which does not exist. Therefore, at a point, this struggle should be ended by a conclusion that "based on the current data, the process cannot be modeled" or "the data is not enough for the modeling". The current paper helps to make this decision quickly.

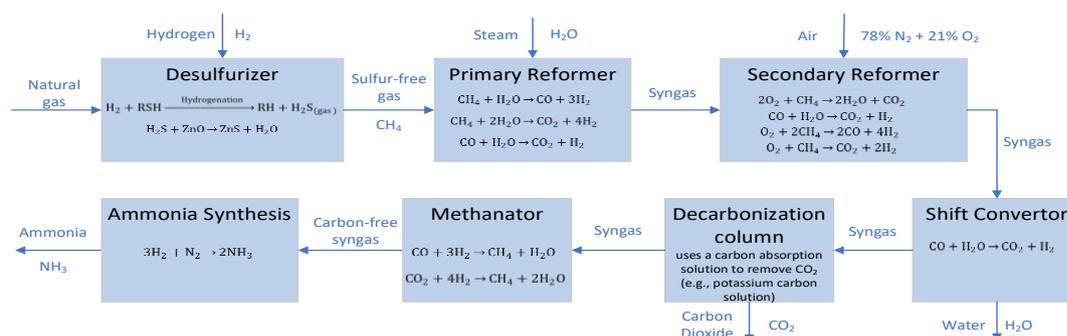


Figure 1: Overview of ammonia typical production steps

The viewpoint of this paper is more practical than theoretical, and represents a walkthrough on how to start modeling given a dataset without any further knowledge about the process. A case study from an ammonia plant is used to illustrate the methods. The case study is an example of non-informative dataset which serves to illustrate the topic of this paper.

In the sequel, Section 2 gives a brief presentation of the dataset. Section 3 reviews the tests for checking the informativity of the dataset and selecting the features. For each represented test, the result of applying it to the case study dataset is shown with further discussion on the results. The authors of the present paper suggest to apply the proposed methods to the dataset in the suggested order, and if applicable, the latest results from each step are used as the basis in the subsequent steps. Section 4 provides a discussion on the results and the conclusion is given in Section 5.

2. Case study dataset

Yara International ASA, one of the largest ammonia producers globally [5], has provided the historical data used in this work. The data is coming from a conventional steam reforming ammonia plant with natural gas used as both feed and fuel. The production capacity is approximately 400,000 tons/year.

As mentioned above, the assumption in the following approaches is that no technical information about the data is available. In other words, it does not matter whether the data comes from an ammonia plant, a refinery, or any other industrial process. However, as the essence of the features and target is known, it is helpful for evaluation purposes. Therefore, let us briefly explain the process and data for a more straightforward interpretation of the results.

2.1. Ammonia plant process

Ammonia plants normally use natural gas to produce ammonia. In the process, natural gas is converted into hydrogen, and then the hydrogen is

combined with nitrogen to produce ammonia using the Haber-Bosch process [6]. Figure 1 illustrates the whole process with the main chemical reactions in each block. The process is described briefly in the following.

Natural gas contains hydrogen sulfide components that can deactivate the catalysts used in the further steps of the process. Therefore, the first step is to remove sulfur components from the inlet gas; the so-called desulfurization. The sulfur-free natural gas is then sent to the primary reformer to react with super-heated steam, where H_2 , CO , and CO_2 are the products. Then, the gas, called synthesis gas or syngas, is mixed with air in the second reformer. The air's nitrogen is needed in the final synthesis, and the oxygen reacts with syngas to produce more hydrogen. To convert the CO contents of the syngas into CO_2 , it is sent to the shift converter section. At this point, the residual water is also removed from the syngas. Then, in the decarbonization section, CO_2 is absorbed from the syngas. Next, the outlet of the decarbonization section passes through another purification section called the methanator to remove small traces of residual CO/CO_2 from the syngas by converting them to methane. Finally, to produce ammonia, the purified syngas (which now contains almost only H_2 and N_2) enters the ammonia converter or synthesis section, where the ammonia is the final product.

2.2. Target variable: CO_2 slip

In the methanator, the reaction between CO/CO_2 and H_2 is extremely exothermic. Therefore, high amounts of CO/CO_2 supplied to the methanator can increase the reactor temperature leading to a series of complex reactions where the consequence is temperature runaway [7]. In addition, removing higher amounts of CO_2 in the methanator consumes more power and reduces the efficiency of the plant. Therefore, it is crucial to control the CO_2 residual in the syngas that enters the methanator (leave the decarbonization column) below a desired limit. This value is the so-called CO_2 slip, the target variable for modeling in the case study dataset.

2.3. Features

Based on experts' knowledge¹, 45 values from historical data are provided for the modeling. Although the selected variables are initially assumed to be promising, there is no guarantee whether they carry sufficient information about the target. In addition, there is a possibility that some variables are redundant, which means they have almost identical information. Using two redundant features does not help the modeling process and may make the model problematic by wrongly assuming the redundant information too important in the training phase. Interestingly, this is the case regarding the dataset used here.

The values of 39 features come from 13 controllers (C1-C13). For each controller, three measurements are available: (1) set-point (SP), (2) sub-process output measurement (PV), and (3) control signal (OUT). If the controller is doing its job well (the case in most running plants), the PV value tracks SP, which means they are almost identical. In some cases, if a simple PID controller is used, the OUT value is nearly proportional to the PV values. However, this is not always true because of non-linearity and saturation, or in the case of cascade controllers. All 13 controllers with a brief explanation are listed in Table 1.

Table 1: List of controllers in the dataset

Code	Type	SP/PV	OUT
C1	Flow	Steam flow (plant load)	Valve
C2	Ratio	Steam-to-carbon ratio	Ctrl (gas flow)
C3	Ratio	Gas-to-air ratio	Ctrl (air flow)
C4	Temp.	Prim. reformer temp. (last 1/3, near outlet)	Ctrl (pressure of last burner nozzle)
C5	Temp.	Prim. reformer temp. (first 1/3, near inlet)	Ctrl (pressure of first burner nozzle)
C6	Δpressure	between air and syngas inlets, sec. reformer	Valve
C7	Flow	Semi-lean solution	Valve
C8	Flow	Semi-lean solution	Valve
C9	Flow	Purge gas into unit	Valve
C10	Flow	Purge gas into prim. reformer	Valve
C11	Pressure	Syngas compressor suction	Ctrl (compressor)
C12	Flow	Lean solution	Valve
C13	Flow	Lean solution	Valve

The remaining six variables (S1 to S6) are measured signals from sensors, where five of them are temperature sensors, and one is a gas sensor. Table 2 summarizes the features, including the responsibilities of controllers and sensors.

Table 2: List of sensors in the dataset

Code	Type	Description
S1	Temp.	Cooling water temperature
S2	Temp.	Ambient temperature
S3	Temp.	Secondary reformer catalyst temperature
S4	Temp.	Secondary reformer outlet temperature
S5	Gas	Secondary reformer methane slip
S6	Temp.	Syngas temperature exits shift convertor

¹ Engineers at Yara International ASA

2.4. Samples

The process historical data is stored in one-minute intervals. The dataset consists of data from 30 consecutive days in August and September 2020. Therefore, there are 43200 samples, each having 46 measurement values (i.e., 45 features + target). Note that, for secrecy, the actual values of the variables are normalized into range [0,1].

3. Step-by-step methodology

Assume a dataset from an unknown industrial process is given, and the desired output (i.e., target) is pre-defined. Therefore, the rest of the variables are inputs to the model (i.e., features). In this section, a practical walkthrough for the starting phase of modeling is presented to check the informativity of data and select the useful features.

The walkthrough consists of three main steps: (1) data visualization, (2) data splitting and initial modeling, and (3) feature selection. The feature selection itself, is divided into two steps.

Note that, for the computations, data science packages in Python such as NumPy, Pandas, scikit-learn, Seaborn, etc., are used.

3.1. Data visualization

Data visualization is data representation by employing visual elements like charts, graphs, and maps, which helps to comprehend the trends, outliers, and patterns in data. Therefore, it is recommended to start data-driven modeling by data visualization prior to the other steps.

As we are dealing with time-series data, the first essential plot is a simple time plot of each variable. To save space, only the target variable and C9_SP are shown in Figure 2. In addition, the weekends are highlighted in the plot, which sometimes helps to realize if they have different trends.

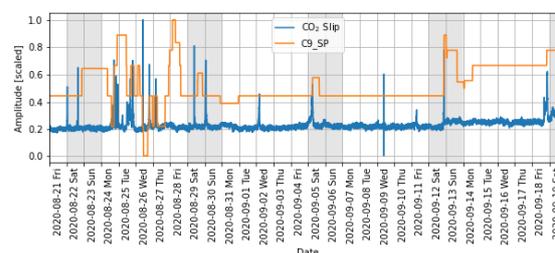


Figure 2: Time-series plot of the target variable (CO₂ Slip) and one feature (C9_SP)

In this example plot, one obvious outlier in target data can be seen on Sep. 9. Another visible issue is a rise in the mean values of both variables after Sep. 12, which suggests the usefulness of C9_SP for modeling CO₂ slip. Note that this relevancy is not the case for all features. Several sharp peaks in

CO₂ slip values indicate a skewed probability distribution of this variable.

Suggestion: It is beneficial to plot the time-series data using interactive plots. Python, for instance, has rich libraries for interactive plots where the programmer can plot all features and targets together. Then it is possible to switch on/off any variable or zoom in/out to have better comparison and inspection.

Another helpful visualization technique is the histogram plot which reveals the data distribution. For instance, Figure 3 depicts the histogram for CO₂ slip where the data bins before and after Sep. 12 are shown in different colors. In addition, the skewness of the data during the whole period is also shown, which is a relatively high number.

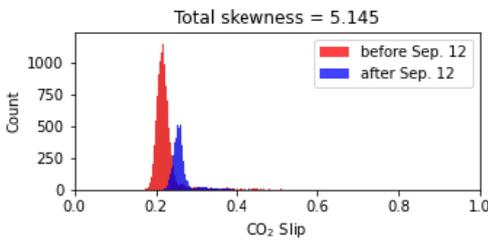


Figure 3: Histogram plot for the target variable

Suggestion: Most linear system identification methods (e.g., ARX, FIR, etc.) assume normally distributed data and Gaussian noises [8]. In cases similar to this example, it may be beneficial to try different models for each part. Later, using a clustering method can define which model should be used in each situation.

Several other charts and plots (e.g., bar chart, radar chart, etc.) can give an initial insight into the data. However, for saving space, this is skipped here.

3.2. Initial regression and dataset split

Before examining the features and selecting the promising ones, making an initial regression model is beneficial. This model reveals if the data has the potential to predict the target variable. In addition, the regression results can help for feature selection later.

In data-driven modeling, splitting data to train and test sets is a routine. However, instead of breaking the dataset into two simple parts, let us divide it into four parts with the following breaking dates: Aug. 28, Sep. 5, and Sep. 12. Then, four train-test pairs are made by taking each part as the test set and the rest as the train set (Table 3).

Table 3: Four train-test pairs

Index	Train	Test
1	Part 1 + Part 2 + Part 3	Part 4
2	Part 1 + Part 2 + Part 4	Part 3
3	Part 1 + Part 3 + Part 4	Part 2
4	Part 2 + Part 3 + Part 4	Part 1

* Part 1: Aug. 21~Aug. 28, part 2: Aug. 29 ~ Sep. 5, part 3: Sep. 6 ~ Sep. 12, part 4: Sep. 13 ~ Sep. 19

The idea comes from k-fold cross-validation [4]. Here, one may be concerned about the test sets' validity as the system is dynamic and the prediction needs the data to be continuous in time. However, as each train set has more than 30,000 samples, and typically not more than 100 past values are used in the models, this will not cause any problem.

The form of the ordinary linear regression (OLS) model for a dataset with n features and N samples is as follows:

$$y(k) = \theta_0 + \sum_{i=1}^n x_i(k)\theta_i \quad (1)$$

Using the extended feature vector $X(k) = [1, x_1(k), x_2(k), \dots, x_n(k)]$ allows writing Eq. (1) in the following compact form:

$$y(k) = X(k)\theta \quad (2)$$

where $\theta = [\theta_0, \theta_1, \dots, \theta_n]^T$ is the coefficient vector. The coefficient vector θ can be calculated by minimizing the mean square error (MSE) between model prediction and measured values:

$$\theta = (X^T X)^{-1} X^T Y \quad (3)$$

Here, $X_{N \times n}$ is the feature matrix where each row presents one measured sample, and $Y_{N \times 1}$ is a column vector containing values of the target. While dealing with dynamic systems, it is beneficial to use the finite impulse response (FIR) instead of a simple (static) OLS. From one point of view, FIR is the dynamic version of static OLS where shifted values of the features are added as new features to the model. Therefore, the FIR model has the following structure:

$$y(k) = \theta_0 + \sum_{i=1}^n \sum_{j=0}^m x_i(k-j)\theta_{ij} \quad (4)$$

where m is called the order of the model and represents the maximum backward shifts for x . To calculate the coefficients θ_{ij} , Eq. (3) can still be used. However, the coefficient vector θ , feature matrix X , and target vector Y should be constructed as follows:

$$\theta_{1 \times (nm+1)} = [\theta_{01}, \theta_{02}, \dots, \theta_{0m}, \dots, \theta_{n1}, \dots, \theta_{nm}]^T \quad (5)$$

$$X_{(N-m) \times (nm+1)} = [X_1^{(0)}, X_1^{(1)}, \dots, X_1^{(m)}, \dots, X_n^{(0)}, \dots, X_n^{(m)}] \quad (6)$$

$$Y_{1 \times (N-m)} = [y(m+1), y(m+2), \dots, y(N)]^T \quad (7)$$

where:

$$X_i^{(j)} = [x_i(m+1-j), x_i(m+2-j), \dots, x_i(N-j)]^T \quad (8)$$

for $i = 1, \dots, n$ and $j = 0, \dots, m$.

The advantage of FIR compared to static OLS is that the past values of features affect the current target prediction. Therefore, if some features affect the target with delay, FIR model includes those delays in the correspondence coefficients.

Now, let us build three models for each train-test pair. The first model is a static OLS, the second and the third are FIR with orders 10 and 30, respectively.

As we want to make a linear model for the process, the dataset is informative enough if the residual values have zero mean. Therefore, for each model, the distributions of residuals in the prediction of test sets are plotted in Figure 4.

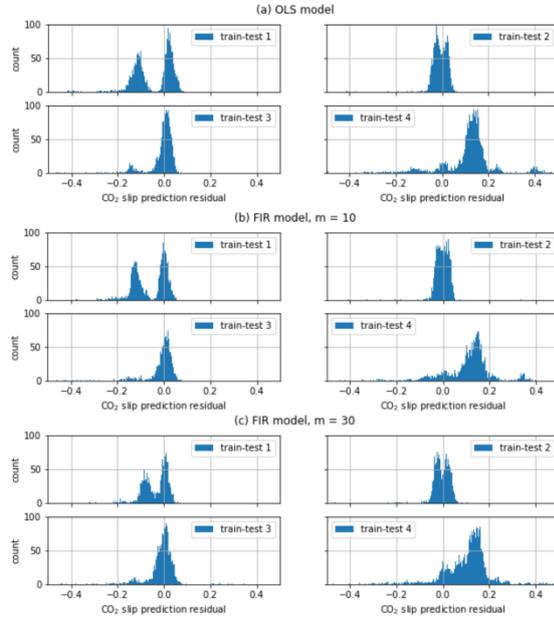


Figure 4: CO₂ slip prediction residual using in (a) OLS model, (b) FIR model $m=10$, FIR model $m=30$

The above plots show that the trends in the first part (i.e., train-test 4) and the last part (i.e., train-test 1) of the data are not wholly similar to the other parts. However, using a FIR model of higher order makes it possible to predict the mean value change discussed earlier, which occurred on Sep. 12. In contrast, for the first part of data, it is almost impossible to accurately predict the target values using either FIR or static OLS models.

As mentioned earlier, redundant or irrelevant features can also cause the above problem. Therefore, let us examine the features and select the promising ones in the following subsections.

3.3. Feature selection: filter methods

Up to this point, we have some idea about the dataset. However, we need some metrics to select valuable features. Generally speaking, feature selection methods are classified into three major groups [9]:

1. Filter methods
2. Wrapper methods
3. Embedded methods

Going through all feature selection methods is out of the scope of this part. Instead, a few filter methods are explained and applied to the data in this subsection. The next subsection briefly reviews the wrapper methods. The embedded methods are not discussed in this paper.

Filter methods are simple statistical tests independent of the model structure. They are computationally cheap and easy to apply to big datasets. Therefore, it is suggested to use them before other methods.

The first and most important filter method is the Pearson correlation coefficient test which measures the linear relationship between variable pairs. The correlation of two random variables is the division of the covariance by their standard deviations:

$$\text{corr}(A, B) = \frac{\text{cov}(A, B)}{\sigma_A \sigma_B} \quad (9)$$

The correlation values are in the range $[-1, 1]$ (if the variances of both variables are non-zero²), where the value of $1/-1$ means that the variables are positively/negatively proportional. Several modeling methods use linear regression as the assessment rule. Hence, a feature that has a near-zero correlation with the target variable almost contains no information about the target.

In addition, if two features are highly correlated (e.g., $|\text{corr}| > 0.95$), they almost contain identical information, making one of them redundant.

The correlation matrix is a symmetric matrix that presents the correlation values of each variable pair. For the case study, the matrix is 46 by 46 and is massive to show here. However, the matrix for the first 24 variables is shown in Figure 5 using a heatmap. The correlations between SP and PV values for all controllers are almost 1.0. This was expected, as discussed earlier. Therefore, from each pair, one of them is redundant. In addition, SP/PV values of controllers C4 and C5 are also correlated. C4 and C5 control the temperature of the last 1/3 and first 1/3 of the primary reformer, respectively. Therefore, it is logical that both temperatures are kept proportional. The same thing is true about controller pairs C7-C8 and C12-C13. C7 and C8 are two parallel controllers for the semi-lean solution, and C12 and C13 are two parallel

² A random variable with zero variance has no information and can be removed from the dataset.

controllers for the lean solution. Therefore, identical set-point settings and their correlations make sense.

Figure 5: Correlation matrix of the first 24 features

Also, as expected, some OUT values have high correlations with their corresponding SP/PV values. However, we keep all of them.

A correlation test can also be done between the target and each feature. Figure 6 depicts the correlations between the target and all features.

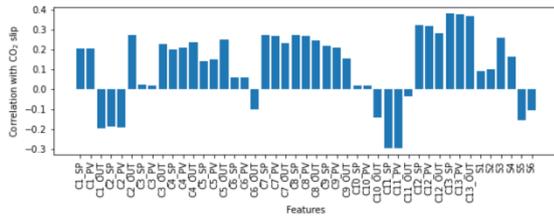


Figure 6: Correlations between the target and features

The correlations shown in Figure 6 help the selection between redundant features. SP is kept for each controller except C2 and C4, and PV is removed from the dataset based on correlations. Regarding the C2 and C4, PVs are kept instead of SPs as they have higher correlations with the target. Therefore, the number of features is reduced to 32.

3.4. Feature selection: wrapper methods

In contrast to filter methods, wrapper methods depend on the type of final model in evaluating the performance of features. Therefore, let us assume the model to be OLS or FIR. Although this assumption may lead to removing features with valuable non-linear information about the target, these models make a fair basis for feature selection for linear modeling in general.

Two common general ideas of wrapper methods are backward and forward feature selection. In addition to these standard methods, two innovative algorithms are also used in this paper. Let us explain the methods in brief:

Forward selection: several models, each including one of the features, are made. Then the feature

corresponding to the best-performed model is selected as the most promising one. In the next rounds, the same procedure is used where all models include the selected variable(s) from the previous round. Finally, the algorithm stops when adding new features does not increase the model performance.

Backward selection: the algorithm starts with a model that includes all features. Then in each step, one variable whose removal leads to the highest performance of the model is selected and removed from the dataset. This procedure continues until removing none of the variables leads to a better model.

Backward-forward ver. 1: in each round, one complete forward selection procedure follows by one complete backward selection. If the forward or backward selection in one of the rounds does not change the selected set, the algorithm stops.

Backward-forward ver. 2: after each round of adding a variable to the list, one backward round is run. Therefore, each backward or forward round removes or adds only one feature.

Note that both backward-forward algorithms (i.e., ver. 1 and 2) can start from an empty or complete set. Therefore, each algorithm has two variants (i.e., empty set start and full set start).

A metric is needed for the model's performance to decide on adding or removing features in all the above-mentioned algorithms. The most common metric is the mean square error (MSE) between predicted and actual values of the target in the test set.

$$\text{MSE} = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k))^2 \quad (10)$$

where $\hat{y}(k)$ is the model prediction for k th sample in the train/test set.

Another helpful metric is the R^2 -Score which is the proportion of the variation in the target variable that is predictable from the features:

$$R^2\text{-Score} = 1 - \frac{\text{MSE}}{\frac{1}{N} \sum_{k=1}^N (y(k) - \bar{y})^2} \quad (11)$$

where \bar{y} is the total mean of the target.

Using each algorithm (i.e., 6 in total) with the two mentioned metrics (i.e., MSE and R^2 -Score) on each train-test pair yields 12 different sets for the features. For instance, using the forward selection algorithm and MSE metric on train-test set 1 gives the following features as the promising ones: C13_SP, C9_OUT, C6_SP, C10_OUT, C12_SP, S3. To save space, the full results are not shown here. Instead, the appearance counts of the features in the final models separated by the train-test sets are shown in Figure 7.

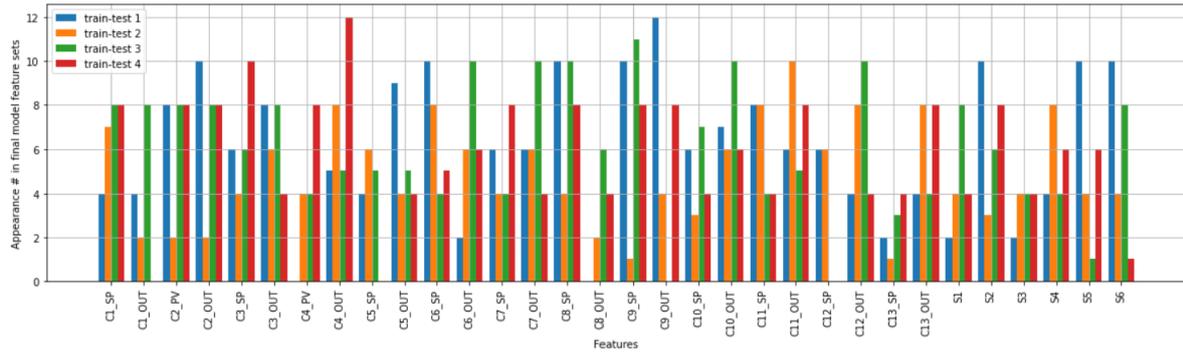


Figure 7: Appearance count of features

One first observation is that the occurrence of some features does not really show their importance. For example, feature S2 (i.e., ambient temperature) is in the final models 10, 3, 6, and 8 times. Obviously, the ambient temperature has almost no meaningful effect on the CO₂ slip values. Therefore, this suggests that most other variables do not contain more information about CO₂ slip than the ambient temperature.

Table 4 lists all variables present in more than half of the final models on average. Based on the information given in Table 4, the most important feature is C8_SP (i.e., set-point value for semi-lean solution controller), which is reasonable. The other important features are C4_OUT (i.e., primary reformer temp.), C9_SP (i.e., purge gas flow into the unit), C10_OUT (i.e., purge gas into the primary reformer), C11_OUT (i.e., syngas compressor suction pressure), and C2_OUT (i.e., steam-to-carbon ratio), where all relate to the plant load and affect the average of CO₂ slip rather than its trends. The rest of the features, as mentioned, are not more informative than ambient temperature. Another observation from Table 4 is the difference between train-test set #2 and the other sets. This suggests a different operation regime in the plant between Sep. 6 and Sep. 12.

Table 4: The most important features based on wrapper methods

Feature	Number of occurrences				Average
	Train-test 1	Train-test 2	Train-test 3	Train-test 4	
C1_SP	4	7	8	8	6.75
C2_PV	8	2	8	8	6.50
C2_OUT	10	2	8	8	7.00
C3_SP	6	4	6	10	6.50
C3_OUT	8	6	8	4	6.50
C4_OUT	5	8	5	12	7.50
C6_SP	10	8	4	5	6.75
C7_OUT	6	6	10	4	6.50
C8_SP	10	4	10	8	8.00
C9_SP	10	1	11	8	7.50
C10_OUT	7	6	10	6	7.25
C11_OUT	6	10	5	8	7.25
C12_OUT	4	8	10	4	6.50

S2	10	3	6	8	6.75
----	----	---	---	---	------

4. Discussion

Based on the features found in previous section, let us make our final models for each train-test set. The models are made based on the train sets, and prediction results on the test set are shown in Figure 8, together with the measured values. The initial OSL model predictions (section 3.2) are also plotted for better comparison.

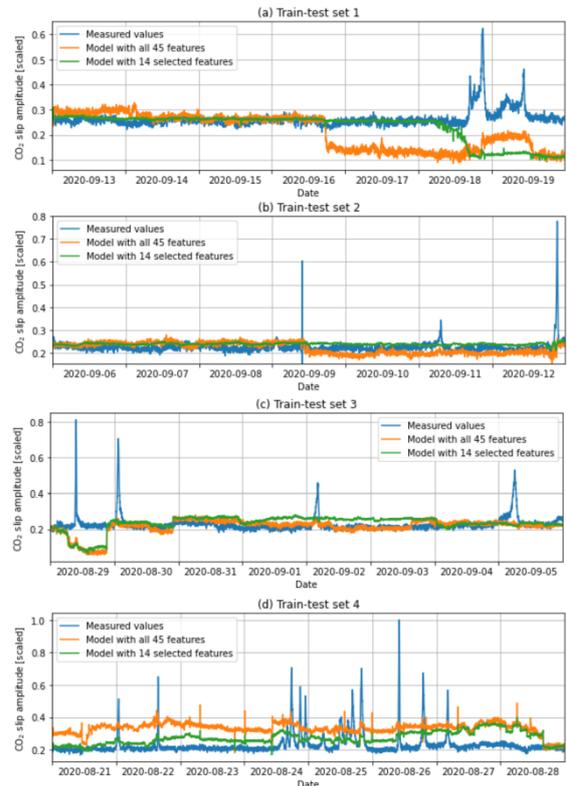


Figure 8: Performance of different models on test sets

The plot's first conclusion is that the feature selection procedure leads to models with better predictions. The second conclusion is that the models suffer from noise that is not white. In other words, some other crucial variables are not

available in the given dataset. This makes linear modeling almost impossible.

Additional attempts were made on this dataset to evaluate the correctness of above-mentioned conclusions. For example, some pre-processing methods used before applying the feature selection methods. Among them:

- the data was smoothed using moving average filter,
- Hampel filter used for detecting and removing outliers, and
- non-scaled and standardized data (i.e., scale to data to have mean of 1 and standard deviation of 0) were used instead of normalized data.

None of the pre-processing methods changed the results noticeably.

In addition, instead of Eq. (3), partial least square (PLS) method was used for finding coefficients in both static OSL and FIR models. PLS automatically reduces the number of features using SVD decomposition. However, those features would be transformed features that have no physical interpretation. Also, using PLS did not contribute to the performance of the models.

To find out how non-linear models perform on this dataset, deep learning methods such as long short-term memory (LSTM) and convolutional neural networks (CNNs) models were also tried. However, all of the models failed to follow the trends in the data. Therefore, “missing feature(s)” is the best description for the case study dataset.

5. Conclusions

The outset of this work was to develop a data-driven model for use in model predictive control (MPC). When considering such models, it is vital to know whether the data contain sufficient information for such a model. In other words: whether the data is informative. In this paper, a walkthrough for feature selection given a historical dataset measured from a process was reviewed. In addition, the informativity of the data is also checked during the feature selection process. The step-by-step method was applied to a case study dataset from an ammonia plant. However, the fact that irrelevant features such as ambient temperature are among the selected features suggests the dataset is not informative enough to predict the target variable. In addition, the prediction results show some trends that the models cannot follow. Further works (i.e., non-linear modeling approaches – not presented in the current paper) illustrate the correctness of the presented walkthrough and non-informativity of the data for modeling the process.

References

- [1] P. Fritzson, Introduction to modeling and simulation of technical and physical systems with Modelica: John Wiley & Sons, 2011.
- [2] O. Nelles, Nonlinear System Identification: From Classical Approaches to Neural Networks, Fuzzy Models, and Gaussian Processes: Springer Nature, 2020.
- [3] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, Dynamic mode decomposition: data-driven modeling of complex systems: SIAM, 2016.
- [4] K. H. Esbensen, D. Guyot, F. Westad, and L. P. Houmoller, Multivariate data analysis: in practice: an introduction to multivariate data analysis and experimental design: Multivariate Data Analysis, 2002.
- [5] Yara Fertilizer Industry Handbook, Yara International ASA, Oslo, 2018.
- [6] V. Pattabathula, and J. Richardson, “Introduction to ammonia production,” CEP magazine, vol. 2, pp. 69-75, 2016.
- [7] J. Gao, Y. Wang, Y. Ping, D. Hu, G. Xu, F. Gu, and F. J. R. a. Su, “A thermodynamic analysis of methanation reactions of carbon oxides for the production of synthetic natural gas,” vol. 2, no. 6, pp. 2358-2368, 2012.
- [8] V. Stojanovic, N. Nedic, D. Prsic, and L. Dubonjic, “Optimal experiment design for identification of ARX models with constrained output in non-Gaussian noise,” Applied Mathematical Modelling, vol. 40, no. 13-14, pp. 6676-6689, 2016.
- [9] I. Guyon, and A. Elisseeff, “An introduction to variable and feature selection,” Journal of machine learning research, vol. 3, no. Mar, pp. 1157-1182, 2003.

Predictive Maintenance of Pumps at ‘Den Magiske Fabrikken’, Using Machine Learning Techniques

Martin Holm^{1,2}, Ozgur Yalcin¹, Carlos Pfeiffer¹, Håkon Viumdal¹

¹ Faculty of Technology, Natural Sciences, and Maritime Studies, University of South-Eastern Norway, Norway, ozgurylc@gmail.com, {[carlos.pfeiffer](mailto:carlos.pfeiffer@usn.no), [hakon.viumdal](mailto:hakon.viumdal@usn.no)}@usn.no

²Lindum AS, Norway, martin.holm@lindum.no

Abstract

In this work, we investigate machine learning methods to predict the failures of progressive cavity pumps (PCP). The PCPs are located in a biogas plant, Den Magiske Fabrikken, in Norway, which is transforming food waste and animal manure to biogas and biofertilizer. Available measurements were pump on-signal, speed, current, torque and control signal, inlet flow, inlet pressure and outlet pressure, and several vibrations derived signals.

Five categories were defined to categorize the operation of the pumps as: *stopped*, *normal running*, *7 days from failure*, *1 day from failure* and *1 hour from failure*. The objective was to train a Machine Learning model to predict these categories. The data was pre-processed to clean gross outliers and scale the signals using different techniques.

This paper presents results from the same Long Short-Term Memory (LSTM) model using two different approaches for scaling the data. The results are evaluated using confusion matrices where one scaling method clearly improves the results when testing on new data points.

Keywords: Machine Learning, Predictive Maintenance, Long-Short Term Memory, Progressive Cavity Pump

1 Introduction

This project investigates and evaluates progressive cavity pump failures used in a waste processing plant by applying machine learning (ML) methodology. In the industry, maintenance costs account for significant losses in profit for companies. Predictive maintenance methods and their tools have changed the way in how to approach problems through advanced control analysis.

[Lindum](#) operates a waste management company that produces biogas and biofertilizer as a result of processing animal manure and food waste. During the processing phase, the highly corrosive and acidic liquid flows through the pipes and causes severe effects on the pumps. To prevent any production losses and increase the pump lifetime, pumps are maintained periodically.

Obviously, excessive manual supervision of the pumps may result in increased labor force demand and increase the spare part costs. However, increased periodic surveillance do not prevent unexpected failures completely. Consequently, the objective of this project is to develop methods for preventing pump failures by analyzing pump parameters with ML algorithms and proposing a model to detect faults.

1.1 Progressive Cavity Pumps and Predictive Maintenance

Positive displacement pumps can handle solids, high viscosity and low flow rates. Besides, progressive cavity pumps are one type of positive displacement pump. Centrifugal pumps, on the other hand, are suitable for low viscosity and high flow rates. The pump efficiency will decrease at both higher and lower pressures for centrifugal pumps, whereas the pump efficiency will increase with increasing pressure in positive displacement pumps.

In this project, the analyzed pump type is ‘Nemo’ brand progressive cavity pumps produced by Netzsch Pumpen & Systeme GmbH. These types of pumps provide a large capacity and pressure range. During the operation of the process in the factory, the pumps suffer from changing viscosity and corrosive materials in each batch. Figure 1 illustrates the progressive cavity pump that is used in the process. The pump has the following components: rotor (1), stator (2a, 2b), drive chain (3), shaft sealing (4), suction and discharge housing (5). Typical problems in progressive cavity pumps are elastomer expansion, rotor, and stator material corrosion which are caused by high temperature or fluid type (Lea et al., 2003).

There are some points to avoid pump failure specifically in progressive cavity pumps. These are:

- Choosing the right elastomer type by taking into account temperature and fluid physical properties
- Avoiding dry running conditions

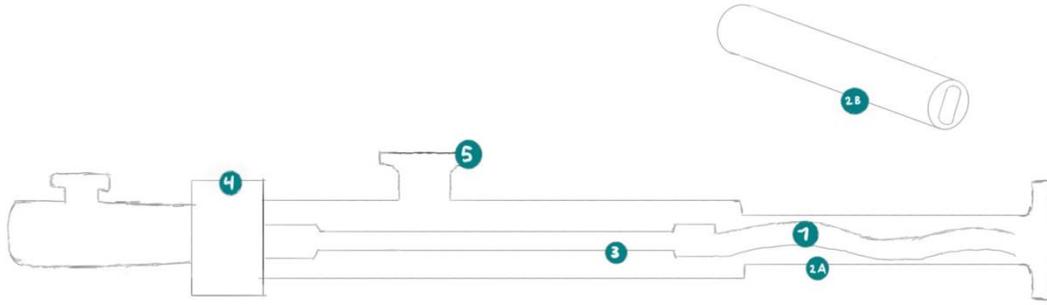


Figure 1. Illustration of the progressive cavity pump that is used in the process.

- Selecting suitable rotor material to stay away from abrasive wear on the rotor

Predictive maintenance aims to transform advanced analytical and process data into valued outcomes. Hence, equipment failure or breakdown can be prevented just before it occurs. Additionally, predictive maintenance may take advantage of ML algorithms to build a systematic approach. Besides, predictive maintenance minimizes the cost of maintenance and improves the equipment lifetime without causing unpredicted production losses. Thus, the process will run as long as possible without interruption.

1.2 Machine Learning Methods

Various type of data is gathered from the process equipment. ML algorithms are able to unveil unseen or hidden patterns and relationships within a data set. With the progressively increase of computational power, and development of new ML algorithms, there is an increasing trend in publications in the literature related to data analysis through ML algorithms (Carvalho et al., 2019). One method is the LSTM algorithm, which is considered especially successful in time series applications, where long-term dependencies in the data needs to be detected (Géron, 2019, pp.511-523). Simply, the function stores a value and determines how long it should be stored. This makes long short - term memory one of the most common models when working with time-dependent data (Rivas et al., 2019). Wisyaldin (2020) compared Autoregressive Moving Average (ARMA), Recurrent Neural Network (RNN), and LSTM models for analyzing vibration signals to predict the health condition of bearings of a water circulation pump and LSTM produced better accuracy. Even though LSTM is used to calculate remaining useful time and anomaly detection in various processes, there are few studies for progressive cavity pump failure analysis with LSTM found in the literature.

2 System Description

2.1 Features

The system under scrutiny in this paper consists of a progressive cavity pump with measurements control signal [%], current [A], torque [%] and speed [%] from a frequency converter. In addition, inlet pressure [Bar], outlet pressure [Bar] and inlet flow [m³/h] is measured. These will be used as the features for the machine learning model. The sampling rate for all the measurements is 30 seconds. Although the selected pump is part of a system of pumps and may be impacted by other pumps earlier in the process, this potential impact has been ignored in this work. The system cyclically pumps fluid for 45-60 minutes, it will always start the cycle again after 60 minutes whether it has just ended or ended 15minutes ago.

The analyzed feature data spans 17 months, with some missing data. During this period, the pump considered has been replaced 14 times due to pump failures.

2.2 Predictions

The goal is to predict one of the five operational categories: pump is (0) *stopped*, (1) *running normally*, (2) *less than one week from failure*, (3) *less than one day from failure* or (4) *less than one hour from failure*. Where running normally is assumed to be anything which is not covered by the other categories.

These categories have been assumed useful as there was little information concerning the breakdown of the pumps, only sparse information about when they had been replaced was available.

3 Methods and Methodology

3.1 Long Short-Term Memory Configuration

The LSTM model architecture was set up as a two layered LSTM block with a dense output layer as seen in Figure 2. The first layer has 7 feature inputs with a sequence length of 120 and 32 output neurons. The layer

has the parameter `return_sequence` set as `true` (Chollet, LSTM layer, 2015) which means a sequence will be returned, compared to only return the last estimate of the sequence, which is the case when set to `false`. The sequence length of 120 samples correspond to one hour which is the cycle time for the pump sequence. The pump sequence is determined by the process operation. The 32 neurons from the first layer serves as inputs to the second layer. However, it outputs only 16 neurons as the return sequence is set to `false`. Both the LSTM layers are using standard configurations for all other parameters. Lastly, a dense layer using the softmax activation function with the 16 neurons from the previous layer as inputs and outputting a probability for each of the categories. The output with the highest probability is assumed to be correct for a given sequence thus giving a positive for one of the five categories. The loss function, categorical crossentropy (Chollet, 2015), is minimized using the Adam optimizer (Kingma, 2017).

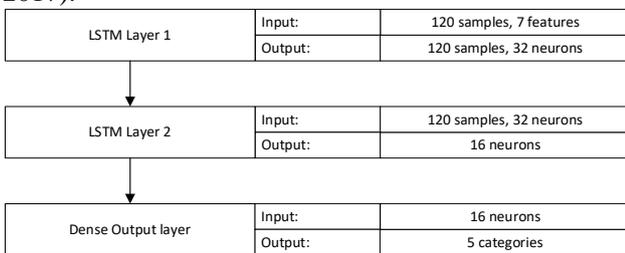


Figure 2. LSTM model architecture

3.2 Scaling

Standardization is used to scale the data, using Eq. (1) where z is the scaled sample, x is the sample that should be scaled, μ is the mean and σ is the standard deviation.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

The standardization is used in two ways, one where the data from all the pumps are scaled using the same scaler for the merged dataset, from hereon named Merged Dataset Scaling (MDS). The other approach is to collect data during one hour of operation for each pump and use this data to calculate individual means and standard deviations to scale the new data. This second method is called Separate Dataset Scaling (SDS). Both approaches are shown in Figure 3.

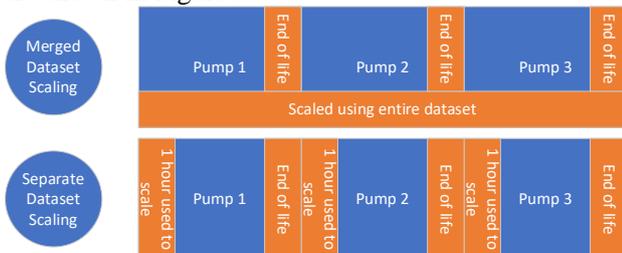


Figure 3. Two methods to scale training data.

The reasoning behind this approach can be seen by inspecting parts of the data shown in Figure 4. There are

several normal distribution-like structures in the current when including data from all the pumps in the same histogram. This may indicate that the level of current (and other variables) may vary between the pumps that have been replaced, and thus the level near the end of the lifetime may vary. This gives the ML method ambiguous signals as to what is considered a degraded pump.

Figure 5 shows one single pump scaled with the first hour of its own data. The distribution now seems closer to a single normal distribution, yet, it still has two distinct tops. Investigating other plots reveals that many look like Figure 5 and some are a lot closer to a narrow normal distribution.



Figure 4. Histogram including all pump failures with previous original scaling method for current.



Figure 5. Histogram for one pump failure for the feature current, scaled.

It thus appears that there are individual characteristics of the pumps, and hence they have various distributions. This may confuse the LSTM-model as there will be many levels of data points where the pump is ok for one pump, but not for another.

Continuing this approach and using data from only the first hour of the pump's active lifetime yields a distribution as seen in Figure 6. This distribution looks more coherent, yet there are more outliers and a higher

span.

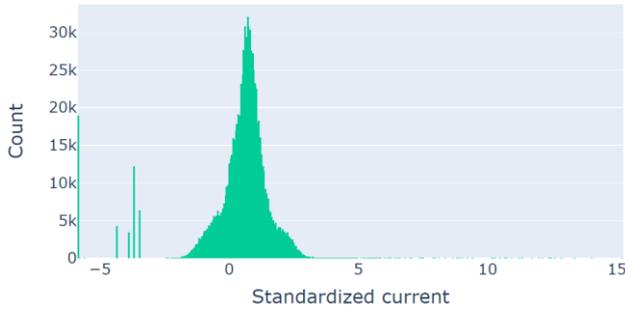


Figure 6. Histogram including all failures with scaling Method 2.

3.3 One hot encoding

The outputs are one hot encoded from integer encoded, meaning that the labels have been converted to numbers as seen in Table 1. These in turn has been transformed into a one hot encoded format, where each row indicates an example where only one label is true, and others are false. As per definition of one hot encoding (Géron, 2017).

Table 1. Integer encoded labels

Label	Description
0	Stopped
1	Normal running
2	Less than one week before failure
3	Less than 24 hours before failure
4	Less than 1 hour before failure

4 Results and Discussion

4.1 Scaling pump data using Merged Dataset Scaling

Before the data is used for training, the features are standardized. All the 14 failures are scaled using one scaler and the data is split into sequences. The outputs are one hot encoded. After this, the data is split into training, validating and testing datasets with 60% used for training, 20% for validation and 20% used for testing. The validation data is used during training to check if the model is improving or not, while the test set from this distribution is used in Figure 7.

Using the MDS method, the results on the confusion matrix based on the training set can be seen in Figure 7 and appears very good. Figure 8 however shows the results in a more realistic manner where the data tested on was not involved in training the model. The model was trained on data from March 2020 to September 2021 and was then tested on data from September 2021 to October 2021. The confusion matrix for the test data shows that all the three categories where it was less than one week before failure of the pump, was considered “normal operation”, or in some rare cases “stopped”. Some of the reason for this might be related to an

extensive number of “normal operation” in the data, compared to the other. That will affect the model. Comparing Figure 7 and Figure 8, there is a clear indication of a generalization problem with the model. As already mentioned, the MDS method has its shortcomings, which is improved in the SDS method.

		Predicted label				
		Stopped	Normal	1W from fail	24h from fail	1h from fail
True label	Stopped	92,4 %	6,7 %	0,7 %	0,2 %	0,0 %
	Normal	0,7 %	99,3 %	0,0 %	0,0 %	0,0 %
	1W from fail	0,8 %	0,3 %	98,6 %	0,2 %	0,0 %
	24h from fail	1,2 %	0,4 %	0,9 %	97,4 %	0,2 %
	1h from fail	2,6 %	0,5 %	0,0 %	8,3 %	88,6 %

Figure 7. Confusion matrix based on training data

		Predicted label				
		Stopped	Normal	1W from fail	24h from fail	1h from fail
True label	Stopped	96,8 %	3,1 %	0,0 %	0,0 %	0,0 %
	Normal	0,8 %	99,1 %	0,0 %	0,0 %	0,0 %
	1W from fail	1,5 %	98,4 %	0,0 %	0,0 %	0,1 %
	24h from fail	1,8 %	98,2 %	0,0 %	0,0 %	0,0 %
	1h from fail	2,5 %	97,5 %	0,0 %	0,0 %	0,0 %

Figure 8. Confusion matrix based on test data

4.2 Scaling pump data using Separate Dataset Scaling

Using the SDS method for scaling the data has reduced the overall accuracy of the model based on the training data, as seen in Figure 9. It can however be noted that most false positives in the failure categories for the most part end up in another failure category.

The test set shows that the model is greatly improved by using the SDS method in Figure 10. The total accuracy becomes 78.5% where the total accuracy is defined by how many samples are correct for each label divided by number of samples tested on.

		Predicted label				
		Stopped	Normal	1W from fail	24h from fail	1h from fail
True label	Stopped	93,6 %	5,4 %	0,8 %	0,2 %	0,0 %
	Normal	1,0 %	98,9 %	0,0 %	0,0 %	0,0 %
	1W from fail	0,7 %	0,0 %	98,8 %	0,5 %	0,0 %
	24h from fail	0,9 %	0,0 %	8,4 %	90,6 %	0,2 %
	1h from fail	0,5 %	0,0 %	14,7 %	37,8 %	47,0 %

Figure 9. Confusion matrix on test data with SUS method from training data

		Predicted label				
		Stopped	Normal	1W from fail	24h from fail	1h from fail
True label	Stopped	97,1 %	1,2 %	0,6 %	1,2 %	0,0 %
	Normal	0,8 %	99,2 %	0,0 %	0,0 %	0,0 %
	1W from fail	1,3 %	0,0 %	55,2 %	43,5 %	0,0 %
	24h from fail	1,1 %	0,0 %	45,0 %	53,9 %	0,0 %
	1h from fail	0,9 %	0,0 %	50,5 %	48,6 %	0,0 %

Figure 10. Confusion matrix on test data completely separate from training data using the SUS method

The expectation to be able to predict one hour before failure may have been high, however, the other failure categories appear to be reasonable, at least a combination of them. Assuming all failure modes are merged to one, the total accuracy of the model becomes 98.9%. One week from fail: 98.7%, One day from fail: 98.9%, and one hour from fail: 99.1% when adding together each row.

However, the model does never predict a normal operation as an upcoming failure for the given test set. The results indicate that there is a small risk of having a false alarm, and performing pump replacements without any good reason, with this approach. Simultaneously, there is a good chance of being able to detect an approaching error within a pump in advance. On the other side, it is a risk for not being able to detect precisely when the pump failures will occur.

The model was only tested on one pump failure, while trained on many. As such there may be a lucky draw that the model was able to predict as well as it did. It has already been seen that the data varies from pump to pump, and it may be that other pump failures are not that well picked up.

In the process of LSTM modelling, noise was not removed from the input data and only raw data was fed into the model. One might argue that noise filtering can increase accuracy. However, it was concluded that noise in data still can hold valuable information, and disregarding noise in data might reduce the model performance.

5 Conclusions and Further Work

This paper has aimed at evaluating and predicting of progressive cavity pump failures in the waste processing plant, maintained by Lindum AS. After gathering information from field instruments, measurement data classified 5 different pump working time cycles such as stopped or normal condition, 1h, 24h, and 1 week from failure. Analyzed data covered 17 months of operation that consist of 14 replacements, and with 30s sampling rate. The time series data was handled well by the LSTM algorithm and produced reasonable results. However, it became evident that scaling for the entire dataset led to information loss on pump failures, that is using the MDS method. Instead, improved results were obtained by scaling pump data with one hour of operational data for each pump replacement, the SDS method. Thus, each pump data was captured on the scaled dataset, separately. The total accuracy of the model with the proposed scaling method becomes 78.5%.

Further work is being done on trying to generalize the model such as to fit onto similar pumps in the process. This requires some features to be removed and use data from many more pumps (Holm, 2022).

As the stopped label is already known, it is not really needed to predict and will be removed in further studies. The one hour from failure label is never predicted outside the training data and is thus removed from further studies.

More work should be done on setting correct parameters of the LSTM structure.

The initial project (Holm et al., 2021) also explored other ML techniques such as Support Vector Machine, Naïve Bayes and Principal Component Analysis. These were not tested with the new scaling method and may be worthwhile to investigate further.

Acknowledgements

We acknowledge the collaboration with and support Lindum AS for providing data. Credits for Ronnie André Horne Moe (ronnie.moe@boliden.com), who

was in the project team, for building LSTM architecture and coding support throughout the project.

References

- François Chollet. LSTM layer, 2015. Retrieved from https://keras.io/api/layers/recurrent_layers/lstm/.
- François Chollet. A. Probabilistic losses, 2015. Retrieved from https://keras.io/api/losses/probabilistic_losses/.
- Aurélien Géron. Handling Text and Categorical Attributes. In *Hands-On Machine Learning with Scikit-Learn*. In pages 92-93. O'Reilly Media, Inc, 2021. ISBN: 9781492032649.
- Sepp Hochreiter and Jürgen Schmidhuber. Long-Short Term Memory, *Neural Computation*, 9(8):1735 - 1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- Martin Holm 2022. *Specifying a machine learning operational framework, refinement and scaling of machine learning models for progressive cavity pumps at Den Magiske Fabrikken* Master's thesis. Retrieved from <https://openarchive.usn.no/usn-xmlui/handle/11250/3000638>
- Martin Holm, Ozgur Yalzin, and Ronnie André Horne Moe, 2021. *Machine Learning for Predictive Maintenance of pumps at "Den Magiske Fabrikken"* Master's pre-project. Retrieved from: <https://github.com/martinsgugge/mast-predictive-maintenance-PCP/blob/master/FM4017%20-%20Project%20Report%20-%20MP-05-21.pdf>
- James Lae, Henry V. Nickens, and Mike. R. Wells. *Gas Well Deliquification*, Gulf Professional Publishing, 2003. doi: 10.1016/B978-0-7506-8280-0.X5001-X.
- Diederik P. Kingma and Jimmy Lei Ba. Adam: A Method for Stochastic Optimization. Retrieved from arXiv: <https://arxiv.org/abs/1412.6980>.
- Thyago P. Carvalho, Fabrizzio A. A. M. N. Soares, Roberto Vita, and João P. Basto. A systematic literature review of machine learning methods applied to predictive maintenance, *Computers & Industrial Engineering*, 137(106024), 2019, ISSN 0360-8352, doi: 10.1016/j.cie.2019.106024.
- Alberto Rivas, Jesús M. Fraile Pablo Chamoso, Alfonso González-Briones, Inés Sittón, and Juan M. Corchado. A. Predictive Maintenance Model Using Recurrent Neural Networks, In *14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019), Advances in Intelligent Systems and Computing*, Springer, pages 261-270, 2019. doi: 10.1007/978-3-030-20055-8_25.
- Muhammad Kamal Wisyaldin, Gita Maya Luciana, and Henry Pariaman. Using LSTM Network to Predict Circulating Water Pump Bearing Condition on Coal Fired Power Plant, *2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*, pages 54-59, 2020, doi: 10.1109/ICT-PEP50916.2020.9249905.

Modeling and analysis of secondary oil recovery with water flooding from a heterogeneous reservoir through advanced wells.

Onkar Bhujange, Ali Moradi, Amaranath Kumara , Britt Moldestad

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway.
onkarbhujange123@gmail.com, Ali.Moradi@usn.no, britt.moldestad@usn.no, Amaranath.S.Kumara@usn.no

Abstract

Oil and gas will remain the most important source of energy for the foreseeable future according to DNV's Energy Transition Outlook 2021, and there is an urgent need to improve oil and gas recovery with less carbon footprint to meet the future energy demands. For the extraction of oil from the reservoir, horizontal drilling is being applied due to its higher recovery rate. These horizontal wells are advanced wells equipped with downhole Flow Control Devices (FCDs), sand screens, zonal isolation as well as monitoring and control systems, etc. FCDs are the key elements of advanced wells. The main objective of this paper is to modelling and simulation of secondary oil recovery with water flooding from a heterogeneous reservoir through advanced wells completed by main types of FCDs. In this paper, a computational study of oil recovery from a black oil reservoir is performed through modelling and simulation using the Petrel[®] software. The modified 'Egg Model' is used which is a synthetic heterogeneous reservoir model consisting of an ensemble of relatively small three-dimensional realizations of a channelized oil reservoir produced under water flooding conditions. A comparative analysis of reservoir models with vertical open hole wells and horizontal open hole wells is presented where it is found that the horizontal well produces 23% more oil. Also, horizontal open-hole wells are further equipped with ICDs which have reduced the water production by 91.4% and have also delayed the water breakthrough in comparison to open-hole horizontal wells.

1. Introduction

The extraction of oil from a reservoir starts by drilling a well into the oil zone. Initially, due to the high pressure, the oil is pushed towards the surface. But as the pressure inside the reservoir drops a recovery method such as water injection is required to maintain a high pressure in the reservoir. This process is called secondary oil recovery. (Lie, n.d.) To achieve cost-effective and efficient oil recovery it is necessary to maximize the well-reservoir contact, and this is achieved using long horizontal wells. However, this method has its challenges. Due to the heel-toe effect and heterogeneity along the horizontal wells, early gas and/or water breakthrough is a big challenge. To tackle this problem, advanced wells are widely used today. Advanced wells are horizontal wells equipped with downhole Flow Control Devices (FCDs), sand screens, zonal isolation as well as monitoring and control systems. (Moradi et al., 2022) FCDs are the key elements of advanced wells. The most widely used flow control device today is passive Inflow Control Devices (ICDs). To achieve a successful design of advanced wells, a suitable dynamic model of oil field and advanced wells must be developed. Generally, it is difficult to observe and understand the dynamics of fluid flow in a porous medium and this is one of the main barriers to developing such dynamic models. Also, it is not possible to measure all the parameters that influence the multiphase flow behavior inside a

reservoir. Consequently, predicting how a reservoir will produce over time and respond to different drive and displacement mechanisms is important to understand. This paper aims to provide more insight into the operation of a passive inflow control device (ICD). This is achieved through the reservoir simulation of black oil production from a vertical well and then a horizontal well equipped with ICDs. The reservoir model used for simulation is an enlarged egg model developed in Petrel[®].

2. Passive inflow control devices

Since the 1990s, ICDs have been used to reduce the danger of early water and/or gas breakthrough in horizontal wells. ICDs are passive flow restrictor devices that are installed on the production tubing and do not have any moving parts. (Moradi and Moldestad, 2022)

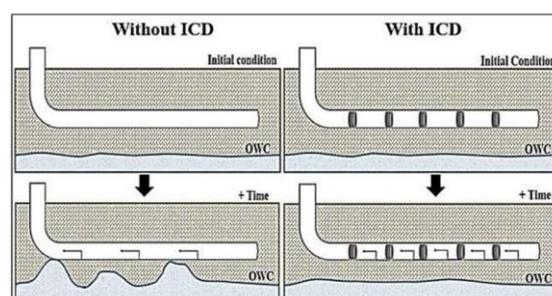


Figure 1: Mitigation of early water breakthrough through the application of ICDs. (Moradi and Moldestad, 2022)

ICDs are used to add extra pressure drop to compensate for non-uniform inflow along the length of the horizontal well. Fig. 1 depicts how such devices work to prevent early water breakout by balancing intake along the well. One type of ICD is the Orifice/nozzle-type ICD. Nozzle-type ICDs enable fluid to achieve the desired pressure drop by limiting the fluid flow. To induce flow resistance, fluid is pushed to pass through small openings (orifices) in a pipe. Fig. 2 shows the nozzle ICD where the fluid path is shown by red arrows. The reservoir fluid flows into the well through the annulus via sand screen and then through the nozzle ICD. As stated in Equation 1, when the fluid flows through the small nozzle, the pressure drop is generated as a function of fluid density, velocity squared, and the geometry of the ICD. Also, in the case of nozzle ICD, the pressure drop is not dependent on the fluid viscosity. The nozzle size and the pressure drop for a specific fluid are set for the nozzle ICD before the installation. (Elverhøy et al., 2018)

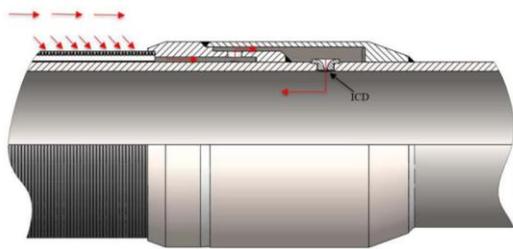


Figure 2: Flow through nozzle ICD. (Elverhøy et al., 2018)

$$\Delta P = C \frac{1}{2} \rho v^2 \quad (1)$$

$$v = \frac{q}{A} \quad (2)$$

Here, q is the volume flow rate of oil, gas, or water depending on the fluid being referred. ΔP is the pressure drop, v is the velocity, C is the geometrical constant, ρ is the density of fluid referred to, and A is the cross-sectional area of the nozzle opening.

Figure 3 shows the performance curve for the ICD which shows ΔP on the y-axis and q on the x-axis.

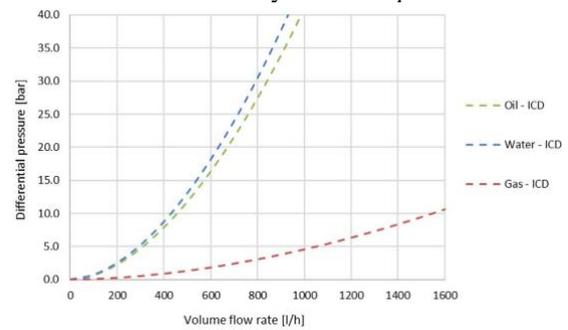


Figure 3: Performance curves for nozzle ICD for oil, water, and gas.

3. Development of Reservoir Model

The reservoir model is developed using the Petrel[®] 2021 software. The geological model used in this simulation case is the Egg Model. The Egg Model is a synthetic heterogeneous reservoir model consisting of small three-dimensional realizations of an oil reservoir produced under water flooding conditions with eight water injectors and four oil producers. This model has been used to demonstrate a variety of aspects related to water flooding simulations. The model consists of a reservoir with discrete permeability fields modeled with $60 \times 60 \times 7 = 25,200$ grid cells of which 18,553 cells are active. The non-active cells are all at the outside of the model, which leaves an egg-shaped model of active cells. (Jansen et al., 2014)

The model is modified in Petrel[®] and is referred to as Enlarged Egg Model which is used to demonstrate the horizontal well.

3.1. Characteristics of the Enlarged Egg Model.

For simulation of oil production through vertical and then horizontal wells a heterogeneous reservoir model namely the Enlarged Egg Model is created in Petrel[®]. The datum of the reservoir is at -4000 m. Each grid block has a width and length of 40 m and a height of 8 m. The model expands 2400 m in x and y directions, while the height is 56 m with 7 layers. Fig. 4 shows the Enlarged Egg Model created in Petrel[®].

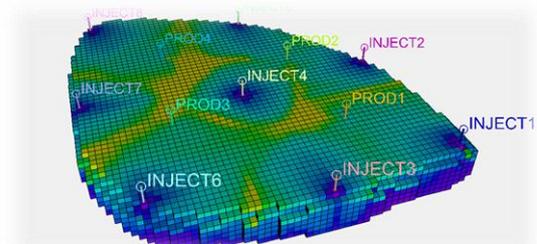


Figure 4: Enlarged Egg Model developed in Petrel[®].

The Permeability of the reservoir in the x-direction is equal to the permeability in the y-direction which is equal to 500mD. The permeability in the z-

direction is 10 % of the permeability in the x/y direction. Since the reservoir is heterogeneous it has high and low permeability zones which forms the channels as shown in Fig. 5.

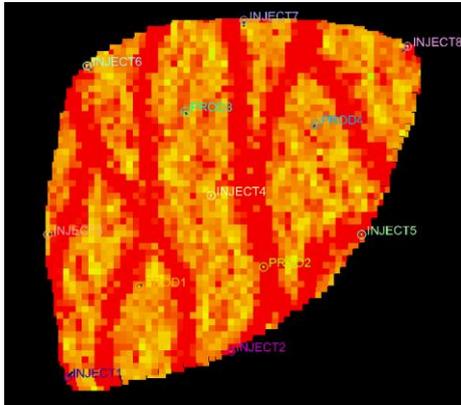


Figure 5: Highly permeable channels.

Tab. 1 shows the characteristics of the reservoir model used for simulation.

Table 1: Reservoir characteristics.

Variable	Value	Unit
Porosity	0.2	-
Oil compressibility	1.1×10^{-10}	Pa ⁻¹
Rock compressibility	0	Pa ⁻¹
Water compressibility	1.0×10^{-10}	Pa ⁻¹
Oil dynamic viscosity	5.0×10^{-3}	Pa s
Water dynamic viscosity	1.0×10^{-3}	Pa s
End-point rel. perm., oil	0.8	-
End-point rel. perm., oil	0.75	-
Corey exponent, oil	4.0	-
Corey exponent, water	3.0	-
Residual-oil satur.	0.1	-
Connate water-satur.	0.2	-
Capillary pressure	0.0	Pa
Ini. Reservoir pressure	40×10^6	Pa
Ini. Water satur	0.1	-
Water inj. rate/well	1650	m ³ /day
Production well BHP	39.5×10^6	Pa
Well-bore radius	0.1	m
Simulation time	7200	day

3.2. Vertical and horizontal well patterns.

Figure 6 shows the 8 injectors (INJECT1 to INJECT8) and 4 vertical producer wells (PROD1 to PROD4) set in a staggered line drive pattern. Each well is 80 m in length vertically with a well-head at -3976 m up to the base of the reservoir at -4056m. The open hole diameter of the well is 8 inches.

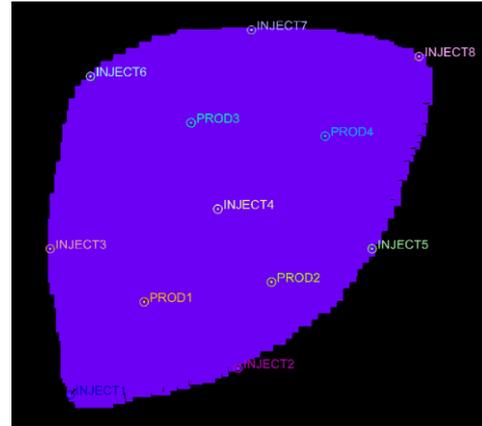


Figure 6: Pattern of producers and Injectors in the reservoir model.

The vertical producer wells are converted to horizontal wells. The horizontal wells rest in the top three layers of the reservoir. The horizontal wells are optimized by trying different positions and directions of the wells before coming to the final well pattern. The length of PROD1 is 533m, PROD2 is 696m, and PROD3 is 660m. PROD4 is 634m. The pattern of the horizontal wells is shown in Fig. 7.

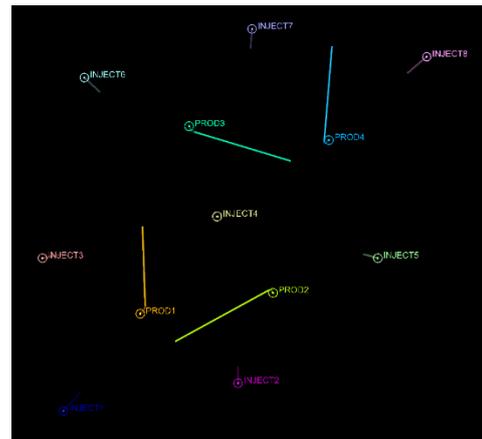


Figure 7: Horizontal well pattern.

3.3. Well Completion

For the open hole of 8 inches, the diameter of the casing is 7 inches, and the diameter of the tubing is 5.5 inches which rest inside the casing. The ICDs are implemented inside the compartment. The length of each compartment is 125 m which is the distance between two packers. Each compartment is equipped with 10 ICDs as shown in Fig. 8.

The ICD in this case has a cross-sectional area of 3.3653×10^{-6} m² as calculated from Equations (1) and Equation (2) where the value of ΔP and q are taken from Fig. 3.

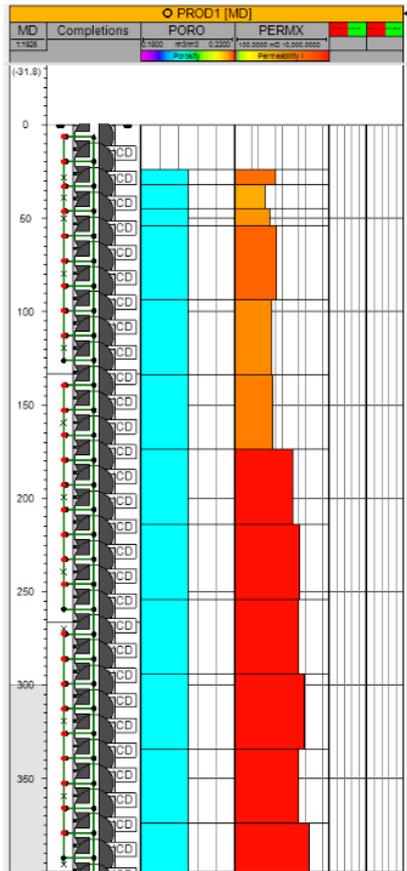


Figure 8: ICD completion developed in Petrel®.

3.4. Well segmentation.

The standard well model cannot be used to model the frictional pressure losses, acceleration pressure loss, and pressure drop across the flow control device. To overcome the shortcomings of the standard well model in the case of horizontal wells, a more rigorous well model is used which is a multi-segment well model. Multi-segment well analysis breaks the well into a series of continuous segments with 0, 1, or more connections to the reservoir grid blocks as shown in Fig. 9. Each segment will consist of four equations three material balance equations and one pressure drop equation. These equations contain the elements that define hydrostatic, acceleration, and friction effects. The equations are solving the pressure, flow rate, and fluid composition in each segment. (Youngs et al., 2010)

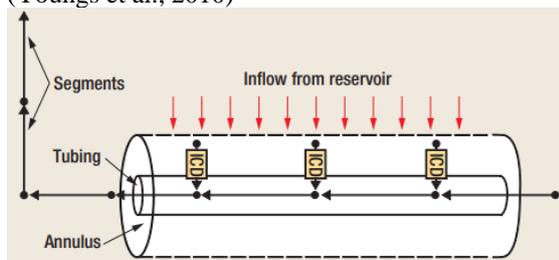


Figure 9: ICDs implemented in multisegmented model as individual segments. (Youngs et al., 2010)

The four producers are segmented in Petrel®. A separate segment is created for each grid block. Each segment is shown by a cross sign (×) in Fig 8. The components of pressure drop in the segmentation are phase slip, friction, and acceleration.

4. Result and discussion

In this chapter results obtained from different simulation cases run in Petrel® are presented and discussed. The reservoir production is controlled by the rate of water injection which is 1650 m³/day and bottom-hole pressure (BHP) which is 395 bar. The different simulation cases compared in this chapter are the production of vertical open holes vs the production of the horizontal open hole. The other case is where all the horizontal open hole production well are equipped with the well completions like casing, tubing, packers, and ICDs. Also, all the producers are provided with segmentation so that the effect of phase slip, friction, and acceleration are taken into consideration while the production. Then the horizontal open hole is compared with the multi-segmented horizontal well equipped with ICD.

3.1. Vertical open hole vs horizontal open hole.

The result for the horizontal well is given by the solid line while the result for the vertical open hole is given by the dashed line, the green color shows the result for oil and the blue is for water.

Field production rate

From the results plotted in Fig. 10, the rate of production of oil for the horizontal well is significantly higher than the vertical producer which at its peak gives the production rate of 6000 m³/day in the case of horizontal and 3600 m³/day in case of vertical producers. After day 2400 the production rate in horizontal cases drops just below vertical production. This is because in the initial 2000 days most of the oil is produced. Also, the production rate of water is very high in the case of a horizontal well. The reason for the higher production rate in the horizontal well is because the horizontal well provides a larger surface area compared to the vertical well given the difference in their lengths. Also, the simulation results of all the production wells show that there is a rapid increase in oil production before the water breakthrough. This is caused because the viscosity of water is 5 times lower than that of the oil and hence water is 5 times more mobile than oil. This highly mobile water tends to push oil rapidly towards the production well which leads to a sudden increment. After the water breakthrough, the rate of production of oil starts dropping.

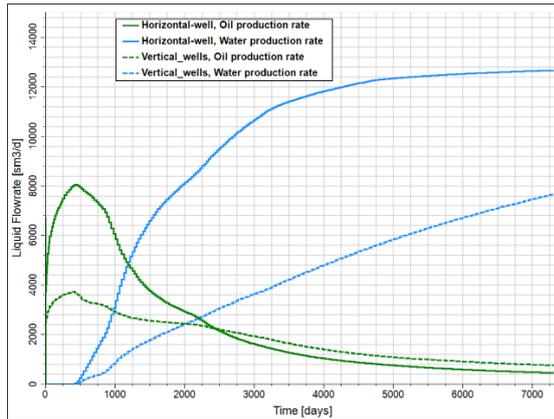


Figure 10: Production rate of oil and water for the horizontal wells vs vertical wells.

Cumulative field production.

The total production of oil and water over time of 7300 days (20 years) is shown in Fig. 11. The production of the horizontal well is higher than that of the vertical. Over time of 20 years, the horizontal well has produced a little over 7×10^7 m³ of water and 1.6×10^7 m³ of oil. While the vertical open hole wells have produced 3×10^7 m³ of water and 1.3×10^7 m³ of oil. So, the horizontal well has produced a higher quantity of fluid from the reservoir. Although the production of water is notably high in the horizontal case. This is because the horizontal wells, in this case, are not equipped with any kind of FCDs and hence these wells are prone to heterogeneity effect, and the water breakthrough is observed quite early.

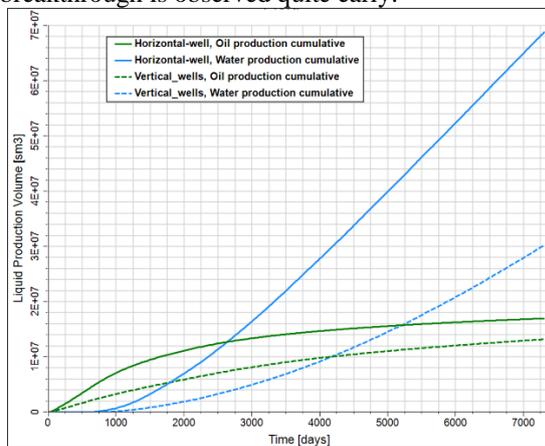


Figure 11: Cumulative production of oil and water for the horizontal well and vertical well vs time.

Tab. 2 shows the total oil and water production of the field in the case of vertical open holes and horizontal open holes.

Table 2: Cumulative production after 7300 days

Well	Oil m ³	Water m ³
Vertical	1.3×10^7	3×10^7 m ³
Horizontal	1.6×10^7	7×10^7 m ³

Dynamic results

The dynamic results in Fig. 12. show that in the case of vertical open hole lot of oil is trapped in the upper layers of the reservoir. Since the result of cumulative oil production shows no further increment. In the case of horizontal open hole more oil recovery is observed.

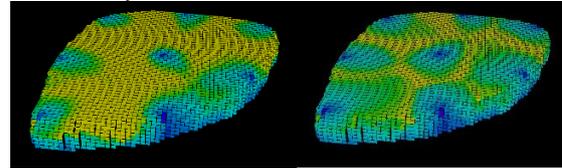


Figure 12: Dynamic simulation result of vertical (left) producer and horizontal producers (right)

3.2. Horizontal open hole well vs horizontal well with ICDs.

In this case, all the horizontal wells are equipped with ICDs, and the production and production rate of the whole reservoir field are analyzed. The result of the production rate of horizontal open hole wells is compared with the horizontal wells with ICD

Field production rate

The result of production rate over time shown in Fig. 13 shows that the rate of production after implementing the ICDs has dropped significantly in the first 2000 days. Also, the initial hype of oil production is flattened. This is because ICDs help to reduce the pressure in horizontal wells by restricting the flow of fluid through a small passage. This also helps to obtain a smoother flow profile over the length of a horizontal well. Implementation of ICD also helped to delay the water breakthrough significantly. Without ICDs the water breakthrough was seen on day 450 while after implementation of ICD the water breakthrough is obtained on day 1700. ICDs have also been helpful to reduce water production as the water production rate at its peak without ICDs is 12800 m³/day and with ICDs is just 2000 m³/day.

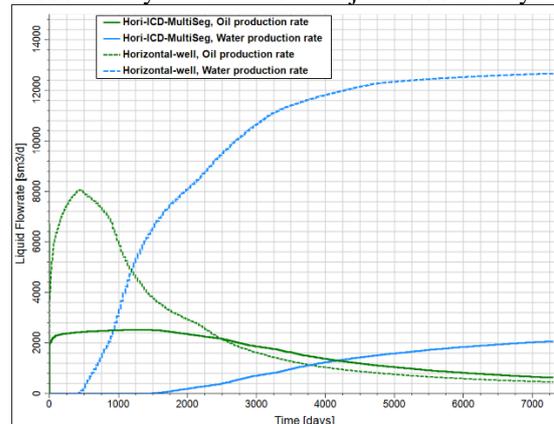


Figure 13: Production rate of oil and water for the horizontal well and horizontal well with ICDs vs time.

Cumulative field production

The simulation results for the total production of the field show that the field with open-hole horizontal wells has produced more oil and water over 20 years (7300 days). However, the lower production of oil is not a bad sign at all. Fig. 14 shows that after 7300 days the oil produced is almost 1.7 times that of oil which stands at 1×10^7 m³ of oil in the case of horizontal wells with ICD. Whereas, in the open hole the water produced is 4 times more than the oil being produced.

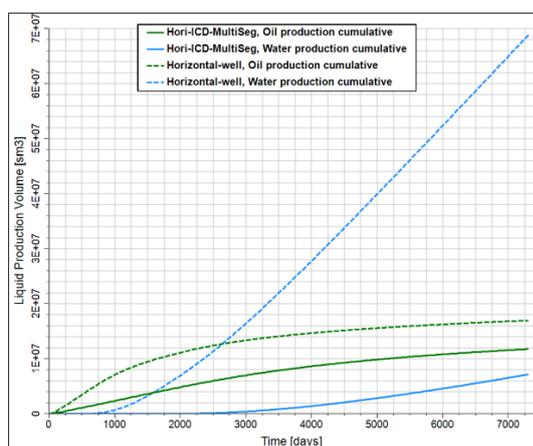


Figure 14: Cumulative production of oil and water for the horizontal well and horizontal well with ICDs.

Tab. 3 shows the total oil and water production of the field in the case of horizontal open hole and horizontal well with ICD.

Table 3: Cumulative production after 7300 days

Well	Oil m ³	Water m ³
Vertical	1.6×10^7	7×10^7 m ³
Horizontal	1×10^7	0.6×10^7 m ³

Dynamic results

The dynamic results in Fig. 15. show that in the case of horizontal open hole producers most of the oil is produced. While in case of the horizontal wells with ICD large amount of oil is still to be produced and the water is in most of the lower layers.

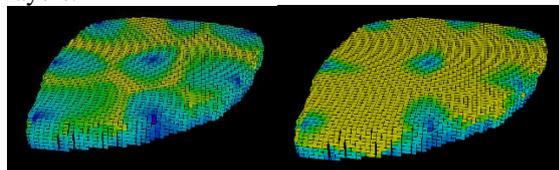


Figure 12: Dynamic simulation result of horizontal open h. (left) producer and horizontal with ICD (right)

5. Conclusion

The simulation results show that the horizontal wells help to improve the production from the reservoir by increasing the surface area in contact with the reservoir. However, they do not necessarily improve the quality of production. Replacing the open hole vertical wells with open-

hole horizontal wells increased the oil production by 23% but it will also increase the production of water by 133%. This is because horizontal wells are prone to heterogeneity effect, especially in the heterogeneous reservoir which may lead to the early water breakthrough which further leads to the drop in oil production.

This effect of early water breakthrough in the open hole horizontal well can be mitigated using the FCDs such as ICD. The comparative analysis of production from horizontal wells equipped with ICDs and the horizontal well with open hole shows that the use of ICDs helps to avoid the early water breakthrough. Results show that the production of water is reduced by 91.4% in wells with ICD. Although there is also a reduction in oil production which is about 37%, the results show that more oil and less water will be produced over time which makes it more economical and environmentally friendly since the amount of water produced is reduced remarkably. Hence, it can be concluded that applying the advanced wells with ICD can improve the quality of production.

Acknowledgment

We gratefully acknowledge the economic support from The Research Council of Norway and Equinor ASA through Research Council project “308817 - Digital wells for optimal production and drainage” (DigiWell).

References

- Elverhøy, A.B., Aakre, H., Mathiesen, V., 2018. Autonomous Inflow Control for Reduced Water Cut and/or Gas Oil Ratio, in: Day 2 Tue, May 01, 2018. Presented at the Offshore Technology Conference, OTC, Houston, Texas, USA, p. D021S024R007. <https://doi.org/10.4043/28860-MS>
- Jansen, J.D., Fonseca, R.M., Kahrobaei, S., Siraj, M.M., Van Essen, G.M., Van den Hof, P.M.J., 2014. The egg model - a geological ensemble for reservoir simulation. *Geosci. Data J.* 1, 192–195. <https://doi.org/10.1002/gdj3.21>
- Lie, K.-A., n.d. AN INTRODUCTION TO RESERVOIR SIMULATION USING MATLAB/GNU OCTAVE 678.
- Moradi, A., Moldestad, B.M.E., 2022. Simulation of Heavy Oil Production using Smart Wells. Presented at the The First SIMS EUROSIM Conference on Modelling and Simulation, SIMS EUROSIM 2021, and 62nd International Conference of Scandinavian Simulation Society, SIMS 2021, September 21–23, Virtual Conference, Finland, pp. 263–270. <https://doi.org/10.3384/ecp21185263>
- Moradi, A., Samani, N.A., Kumara, A.S., Moldestad, B.M.E., 2022. Evaluating the performance of advanced wells in heavy oil reservoirs under uncertainty in permeability parameters. *Energy Rep.* 8, 8605–8617. <https://doi.org/10.1016/j.egy.2022.06.077>
- Youngs, B., Neylon, K., Holmes, J., 2010. Multisegment Well Modeling Optimizes Inflow Control Devices. *World Oil* 4.

Influence on the fluidization pattern of a freely bubbling fluidized bed with different modes of air supply

Rajan Jaiswal *, Marianne S. Eikeland, Britt M. E. Moldestad, Rajan K. Thapa

University of South-Eastern Norway, Faculty of Technology, Kjølnes rig 56, 3918, Porsgrunn, Norway

rajan.jaiswal@usn.no

Abstract

Bubbling fluidized bed (BFB) reactors are extensively used in several process applications like gasification, pyrolysis, drying, and combustions due to their excellent mixing properties and good temperature control. The bubble dynamic and particle movement in the reactor is primarily responsible for uniform heat and mass transfer and mixing. The properties of bubbles in BFB are governed by the gas distribution inside the reactor or supply of the fluidizing gas. This work investigates the influence on the fluid dynamic behaviour of the BFB reactor at different fluidizing gas injection systems using the Computational Particle Fluid dynamic model. Three different modes of fluidizing gas injection include uniform injection, air injection via twenty-five nozzles, and air supply via side nozzles along the reactor height in a gasification reactor of 10.04 cm diameter. Air is used as the fluidizing gas and silica sand as the bed material. The CPFD model is developed in Barracuda Virtual reactor 20.01. The CPFD model is validated against the experimental data obtained from the Electrical Capacitance Tomography (ECT) sensors. The result depicts the better fluidization quality of the bed with uniform air supply as flow boundary and air injection via twenty-five nozzles located at the bottom of the reactor. With air injection via two side nozzles along the reactor height, the bed is fluidized with large bubbles and particle entrainment in the freeboard zone of the reactor. A method is proposed to improve the fluidization quality of the bed while using side nozzles as inlet flow Boundary Conditions (BC) for air injection. The proposed method includes addition of four nozzles along the reactor wall instead of two which improves the fluidization quality of the bed in terms of smaller size bubbles without particle entrainment in the freeboard region.

Keywords: Gas distribution, flow boundary conditions, bubbling-fluidized bed, bubble properties, CPFD

1. Introduction

Fluidized bed reactors are widely used for several industrial applications like waste to energy conversion, chemical synthesis, granulation, drying of pharmaceutical products and raw agricultural products, chemical looping, catalyst regeneration, biomass gasification, pyrolysis etc. (Jaiswal *et al.*, 2020; Bandara *et al.*, 2021; Singh and Gbordzoe, 2017; Chang *et al.*, 2013). The efficiency of the fluidized bed reactors largely depends on the gas distribution inside the reactor since the gas distribution influences the conversion process and the fluidization regime under which the reactor is operated. For instance, during the gasification of biomass or wastes using a bubbling fluidized bed reactor, the carbonaceous feedstock is converted into higher calorific value gases in the presence of limited amount of oxidizing agent (Jaiswal *et al.*, 2020). The amount of oxidizing medium present for the feedstock conversion depends on how well the fluidizing gas is distributed across the reactor cross-section. Similarly, the hot bed material which is in a continuous motion in such reactors acts as the thermal flywheel and provides the required heat for thermal degradation of the feedstocks. The particle

motion is governed by the bubbles rising in the bed and the properties of the bubbles for example bubble rise velocity, bubble diameter and bubble frequency is determined by the gas distribution inside the reactor. Additionally, the mixing phenomena of large biomass particles with bed material and the operating regime of the reactor is determined by the gas distribution (fluidizing gas) to the reactor. The fluidizing gas can be supplied to the particle bed through the distributor or nozzles (Basu *et al.*, 2006). The most common method has been the use of a distributor plate that allows to distribute the fluidizing gas uniformly, supports the bed material, provides good gas-solid mixing, prevents channeling, and minimizes dead zones in the reactor (Depypere and Dewettinck, 2004). While using the air distributor or distributor plate, the supply gas velocity has to overcome the distributor plate resistance and lift entire mass of the particles against the gravity in order to fluidize. Alternatively, the fluidizing gas can be supplied to the reactor via orifice or nozzles which can be on the side of the reactor wall or at the bottom of the reactor. It is essential to characterize the fluidized bed behaviour for the specified fluidizing gas flow boundary

conditions for a smooth operation of the reactor (Sasic and Johnsson, 2005). Each of the fluidizing gas supply methods, with or without distributor, nozzles, orifice has its own limitation and advantages. For instance, the use of a distributor plate increases the auxiliary power consumption required to pump the gas through the reactor. Additionally, the distributor plate has to be selected depending on the reactor types and process as the distributor plays a critical role in the reactor performance (Raza *et al.*, 2021). In addition, there are challenges in operating the reactor with the distributor plate, which requires cleaning and maintenance due to clogging of the pores of the distributor plate by the fine particles and sintering. The blockage of the pores in the distributor plate can lead to local de-fluidization and dead zones in the reactor. In this regard, operating the fluidized bed reactor without a distributor can be of great advantages as it can save the operational cost, and the construction and design cost of the distributor plate. At the same time, it helps to avoid the problem associated with the air distributor. Several studies have been conducted on design of the distributor plate and the influence of the distributor plate on the reactor performance (Geldart and Baeyens 1985; Saxena *et al.*, 1979). However, the number studies on operation of a bubbling fluidized bed reactor without an air distributor and comparing the fluid dynamics behaviour of the bed operated with and without distributor are scarce (Agu *et al.*, 2018). There are a few studies on operation of a fluidized bed reactor without an air distributor where large size particles are used at the lower region of the bed and the bed is fluidized by passing the gas through the bed of stationary large particles (Agu *et al.*, 2018).

In a single study, reported on the operation of reactor without an air distributor Agu *et al.* has studied the bed behaviour with different types of particles and gas velocity based on the pressure and solid circulation using CPFDF simulation. However, no information is provided about the grid size and number of computational cells in the article (Agu *et al.*, 2018). The grid number or computational domain in such simulations have significant impact on the fluid dynamics behaviour of the bed and the pressure and flow boundary conditions. Additionally, there are no studies in the literature that mention the influence on the bubble properties in a fluidized bed reactor operated without a distributor plate. Therefore, more work is required to characterize the fluid dynamics behaviour of the bed without an air distributor and compare the fluid dynamics behaviour of the bed with a distributor and different modes of gas supply to the reactor.

The objective of this work is to investigate the fluid dynamics behaviour of the bed without air

distributor and compare with uniform air supply methods. The uniform air supply methods include use of an air distributor and nozzles. A cold flow model of the fluidized bed reactor with different air injection methods are simulated using a CPFDF model developed in Barracuda VR 20.01. For the case without air distributor the bed is fluidized with air supply from two nozzles (holes) located at the opposite side of the reactor wall. The fluid dynamics behaviour of the bed in terms of solid fraction fluctuation and bubble properties are reported briefly. The results from the CPFDF simulations are compared with experimental data obtained from a cold fluidized bed equipped with ECT sensors and an air distributor. A method to improve the fluidization quality of the bed without an air distributor is proposed.

2. Material and methods

2.1. Experimental set up

A cold fluidized bed reactor equipped with ECT sensors, and a data acquisition system are used for the experiments. The reactor is 10.04 cm in internal diameter and 150 cm in height. The reactor column is fitted with an air distributor at the bottom and is open at the top. The air distributor is 3mm thick with a 10.04 cm internal diameter. It has a porosity of 40% with a flow area of 36.6 cm² which allows the fluidizing gas to pass through the bed uniformly. The reactor is equipped with twin-plane ECT sensors that are located at 15.7 cm and 28.7 cm from the air distributor. Each of the sensors consists of 12 electrodes mounted on the outer wall of the reactor that allows to capture raw data in the form of matrix or images. The online images are extracted from the capacitance measurements using the Linear Back Projection algorithm. The cross-section of each sensor is divided into 32*32 square pixels of which 812 are the effective pixels that lie within the bed. Each pixel holds a normalized relative permittivity value between 0 and 1, which represents the solid-gas fraction. The details of the reactor set up can be found elsewhere (Agu *et al.*, 2019). The reactor is filled with sand material and fluidized by using compressed air. The fluctuation of solid volume fraction is measured from the transient data for each gas velocity. The raw data from the experiments were processed in MATLAB to obtain the solid fraction fluctuation, bubble properties, and bed flow dynamics behaviour.

2.2. Simulation set up

The CPFDF simulations in this work have been carried out using a commercial software Barracuda VR which consists of numerical codes specially designed for applications in the multiphase systems like industrial fluidized bed reactors.

Initially, the reactor is filled with bed material with a static bed height of 30 cm. The properties of bed

material like density (2650 kg/m^3), and particle size distribution, sphericity (0.86), close pack volume fraction (0.63), are defined similarly to that of the experiment. The particle size distribution was measured from the sieve analysis with mean diameter $423 \mu\text{m}$.

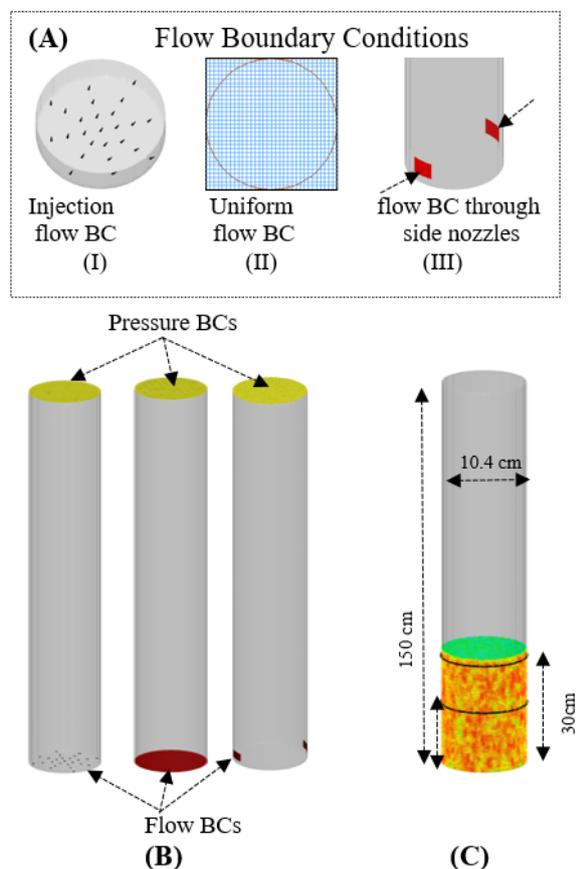


Figure 1: (A-B) flow and pressure boundary conditions, (C) flux planes, initial bed height and reactor dimension.

For the simulation of gas-particle flows, 3D multiphase particle-in-cell approach is used where solid particles are modeled as discrete Lagrangian methods and the fluid is modeled as Eulerian grid of cells. To create a virtual reactor, a CAD geometry equal to the experimental column is imported to Barracuda VR. A uniform grid of total 102400 cells were defined which provides the control volume for all fluid fields calculations. The reactor is operated at atmospheric pressure. Therefore, the pressure boundary condition is defined at the top of the reactor. Three different flow boundary conditions were set up for different modes of gas supply to the reactor as shown in Fig. 1(A). The fluidizing gas is supplied to the reactor using injection points, uniformly distributed along the reactor cross-section (grids or distributor) and through side nozzles (holes). 25 injection points were defined along the reactor cross-section shown in Fig. 1A(I). The mass flow rate through each nozzle (injection points) was equally distributed. Similarly, for the

second case, uniform flow boundary conditions are defined at the bottom of the reactor. The reactor cross-section is divided into 1024 cells. The 812 effective pixels or grids that cover the reactor cross-section as shown by the red circle in Fig. 1A(II) is used as the flow BC. For the third case, two nozzles or holes located on the opposite side of the reactor wall are defined as the flow boundary conditions (as shown in Fig. A(III)). The size of the nozzles is 0.5 cm in diameter. Also, two planes are set up at heights 15.7 cm and 28.7 cm along the height of the reactor column to measure the transient raw data for each of the flow boundary conditions. The drag model used in fluidized bed simulations is an important factor that determines the force acting on a particle by the flow of fluid around it. In this work, the Wen-Yu and Ergun blended drag model is used where the Wen-Yu drag model is suitable for the dilute phase and the Ergun model is used for the dense phase. The details of the drag model, and the governing momentum and force equation can be found elsewhere (O'Rourke and Snider, 2012; Sinder, 2001; Andrews and O'Rourke, 1996; Weber et al., 2013).

3. Results and Discussion

For maximum conversion efficiency, it is important to operate a BFB reactor above minimum fluidization velocity and within a bubbling regime. In this work employs methods to identify the fluidization quality of the bed based on solid volume fraction fluctuation and bubble properties measurement. In addition, influence of different air supply modes on the fluid dynamics behaviour of the bed are presented.

3.1. Model Validation

The CPFDF model is validated by comparing the solid volume fraction fluctuation measured from the experimental data and CPFDF simulations at different gas velocities as shown in Fig. 2. As the air was supplied through the static bed, initially the bed expanded as it reached the superficial gas velocity 0.065 m/s . The bed exhibited into the fluidization regime at superficial gas velocity 0.075 m/s . The results depict that the CPFDF model followed a similar trend to that of minimum fluidization velocity and bed expansion. However, the solid fraction fluctuation during bed expansion was higher as predicted by the CPFDF model compared to the experimental data. This may be due to the difference in the initialization (packing) of the bed material in the CPFDF model compared to the experimental conditions.

Further, the bubbles rising in the bed at the superficial gas velocity of 0.1 m/s at different time steps are compared to check the model robustness. The bubbles rising in the bed at a pre-defined plane in the simulation setup and ECT sensors in the experimental set up were captured for both the

CPFD simulation data and the experimental measurement as shown in Fig. 3. The size of the bubbles and the path along which the bubbles move upward (along the center of the bed) in the bed were similar for both experimental tests and the CPFD simulations. However, the number of bubbles predicted by the CPFD model is higher at a lower superficial gas velocity (0.1 m/s). Due to increase in area of the flow boundary conditions in case of CPFD simulation, smaller size bubbles appeared in the CPFD simulation compared to experimental data.

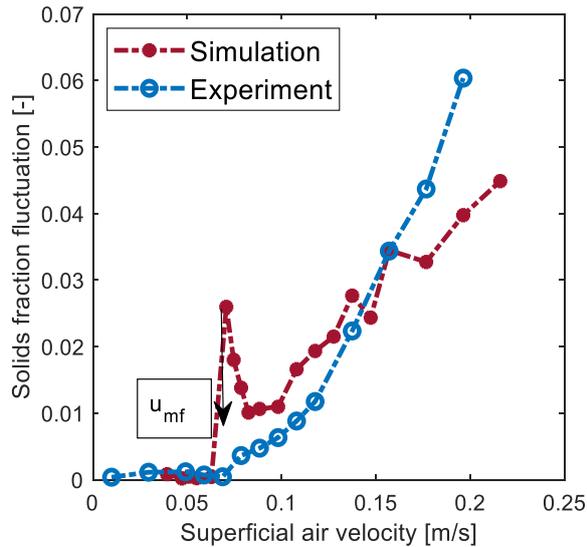


Figure 2: Solid fraction fluctuation at different superficial gas velocities obtained from CPFD simulation and experiment.

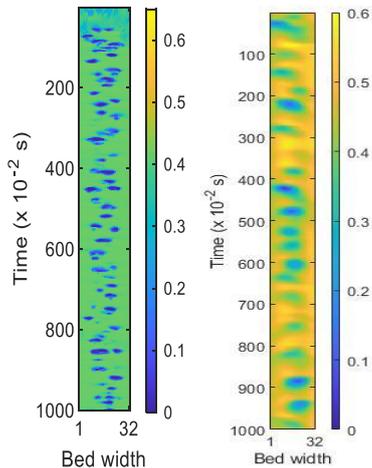


Figure 3: Comparing the rising bubbles in the bed obtained from the CPFD simulation (left side) and experiment (right side) as it reaches the plane at 15.7 cm at the superficial gas velocity 0.1 m/s.

3.2. Influence on solid fraction fluctuation

For a given superficial gas velocity, the variation of the gas-solid fraction in the bed can be used to characterize the fluidization pattern of the bed. The

solid fraction fluctuation across the cross section of the bed at height 27.5 cm and at the superficial gas velocity 0.15 m/s were captured over the measurement of 60s with time steps of 0.001. The average solid fraction fluctuation over the measurement period for different flow boundary conditions (injections BC, nozzle BC, uniform BC) from the CPFD simulation and experimental measurements with air distributor are compared in Fig. 4. The result illustrates that for the experimental measurements with an air distributor the solid fraction was lower towards the center of the bed, and it increases near to the wall of the reactor. A similar trend of solid fraction fluctuation was predicted by the CPFD model with the injection flow BC and the uniform flow BC. However, fluctuation in the solid fraction for both the cases were not smooth compared to the experimental measurements because of smaller bubbles in the bed with the uniform and injection flow BCs. Multiple smaller size bubbles were formed in case of the CPFD simulations due to an increase in flow boundary area as compared to that of the experiment. The decrease in the solid fraction at the lower region of the bed near to the center reveals a higher gas fraction in the region. The gas in a fluidized bed rises from the lower region of the bed to the upper region in the form of bubbles. Therefore, the depression of the solid fraction near to the center of the reactor and the increase near to the wall of the reactor illustrate that the bubbles rise upwards following the path near to center of the reactor. For the flow BCs with two nozzles near to the wall of the reactor, the solid fraction fluctuation was lower near to the wall of the reactor where the nozzles were defined and near to the center of the bed. The fluidizing gas passes in the form of bubbles where some of the bubbles pass near to the wall while some follow the path near to the center of the bed.

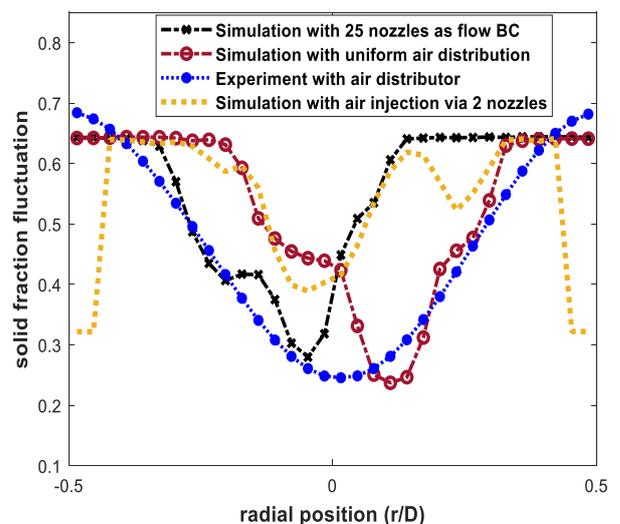


Figure 4: Radial distribution of the solid volume fraction for different flow boundary conditions at superficial gas velocity 0.15 m/s.

3.3. Influence on bubble properties and bed fluid dynamics behaviour

The bubble properties in a BFB reactor significantly influences the fluid dynamics behaviour of the reactor and its performance. The appearance and the movement of the bubbles in the bed is governed by the flow of fluidizing gas through the bed. For the same mass flow rate of the fluidizing gas with different flow boundary conditions, the bubble properties were measured. The influence on the bubble properties and fluidization quality of the bed with different flow boundary condition are compared. The bubbles in BFD can be distinguished from the dense phase by using a bubble solid threshold value. In this work the bubble solid threshold value of 0.2 is used to identify bubbles in the bed. For the CPFD simulations, the bubbles were identified as the zones (object) and the volume of the bubbles were measured by counting (with algorithm written in MATLAB) the number of the cells the bubbles occupied. The bubbles diameter were than calculated from equivalent spherical area.

Fig.5 compares the bubble frequency with respect to the bubble diameter for different flow boundary conditions over the measurement of 30 s. With the uniform flow boundary conditions, the result shows that different bubble sizes within the range of $2.5 \text{ cm} < d_b < 5 \text{ cm}$ appeared in the bed with dominant frequency of the bubbles with 3 cm. Here, d_b is the bubble diameter in cm. While for the flow with two nozzles, the bubble diameters are comparatively large for the same mass flow rate (1.5 kg/hr.) of the air as compared to that of uniform and injection flow BCs. With the injection flow (using 25 nozzles) the bubbles in the bed are smaller and uniform in size as shown in Fig.5(b). The smaller and uniform size bubbles in the bed means better fluidization quality of the bed which contributes to uniform heat transfer and better mixing. On contrary, the large size bubbles in case of flow BC via two nozzles (as shown in Fig.5(c)) can bypass the bed if such bubbles rise in the bed near to the wall of the reactor. Such large bubbles can grow into slugging bubbles as they rise in the bed and turn the bed into a slugging regime. Additionally, with large size bubbles in the bed, the bubble rise velocity in the bed increases significantly which may transform the bed into turbulent regime. When the bed is in the turbulent regime during gasification of biomass in BFB reactor, the fine particles can entrain in the freeboard region which may contaminate the product gases. The entrainment of fine particles in the freeboard region when the air is supplied to the reactor via two side nozzles is shown in Fig.7 F(a). Also, due to increase in gas velocity and bypassing of the fluidizing gas through the side of the bed (near to the wall of the reactor), less oxidizing medium is present for thermochemical conversion of feedstock in the bed where biomass is present. As a result,

more oxidizing medium is present in the free board region which can convert the carbon monoxide into carbon dioxide. Therefore, in order to supply the fluidizing gas to the reactor without a distributor (with side nozzles) it is essential to optimize the flow behaviour of the fluidizing gas to the reactor. This can be achieved by increasing the number of nozzles (or holes) along the reactor wall. In this work, the CPFD model was used to simulate a case with four different nozzles along the reactor wall. By increasing the number of side nozzles, entrainment of the bed particles were prevented as shown in the Fig.7(b). Similar, the bubble size is reduced (as shown in Fig.5(d)) and the bubble frequency is increased with large number of smaller size bubbles in the bed. With addition of two more side nozzles, better fluidization quality of the bed is achieved (as shown in Fig.6 and Fig.7).

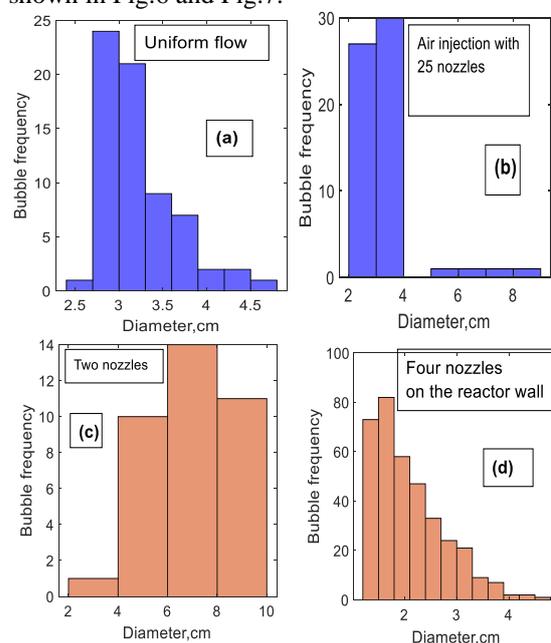


Figure 5: Bubble frequency vs bubble diameter for different air supply systems.

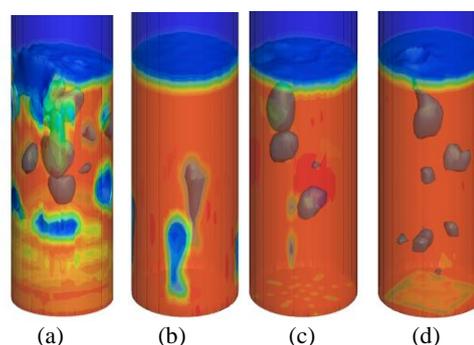


Figure 6: (a-d), Iso-surface of the bubbles in the bed obtained during, air injection via four nozzles, two side nozzles, 25 nozzles at the bottom of the reactor, and uniform flow BC respectively.

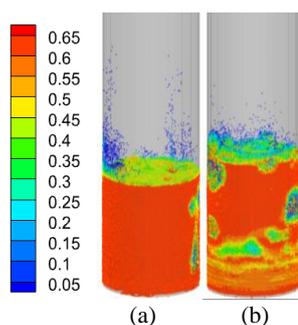


Figure 7: Particle volume fraction of the bed with (a) two side nozzles and (b) four side nozzles.

4. Conclusion

Bubbling fluidized bed reactors are extensively used for several industrial application due to uniform heat and mass transfer. The advantages of a such reactor can only be achieved with proper distribution of the fluidizing gas to the reactor. This work investigates three different methods to supply fluidizing gas to the reactor including uniform flow BC, air injection with twenty-five nozzles and air supply via side nozzles along the reactor wall. A CPFD model has been developed in Barracuda VR 20.01 and the model is used to investigate the fluidization quality of the bed in terms of solid fraction fluctuation and bubble properties. The CPFD model is validated against experimental data obtained from a cold fluidized bed equipped with ECT sensors and air distributor. The result depicts that a better fluidization quality of the bed is achieved with uniform air supply as flow boundary and air injection via twenty-five nozzles located at the bottom of the reactor. With air injection via two side nozzles along the reactor height, the bed fluidized with large bubbles and particle entrainment in the freeboard zone of the reactor. A method was proposed to improve the fluidization quality of the bed while using side nozzles as inlet flow BC for air injection. The proposed method includes addition of four nozzles along the reactor wall instead of two. With the addition of two nozzles, fluidization quality of the bed was improved in terms of smaller bubble sizes without particle entrainment in the freeboard region.

References

Agu, C.E., et al. (2019) 'Investigation of Bubbling Behavior in Deep Fluidized Beds at Different Gas Velocities using Electrical Capacitance Tomography', *Industrial & Engineering Chemistry Research*, 58(5): p. 2084-2098. <https://doi.org/10.1021/acs.iecr.8b05013>

Agu, C.E. and B.M.E. Moldestad. (2018) 'Distribution of Solids in a fluidized bed operated without a gas distributor', *Linköping Electronic Conference Proceedings*. (153), 248-254. <http://dx.doi.org/10.3384/ecp18153248>.

Andrews, M.J. and P.J. O'Rourke. (1996) 'The multiphase particle-in-cell (MP-PIC) method for dense particulate flows', *International Journal of Multiphase Flow*, 22(2): p. 379-402. [https://doi.org/10.1016/0301-9322\(95\)00072-0](https://doi.org/10.1016/0301-9322(95)00072-0).

Bandara, J.C., et al. (2021) 'Air gasification of wood chips, wood pellets and grass pellets in a bubbling fluidized bed reactor', *Energy*, 233: p. 121149. <https://doi.org/10.1016/j.energy.2021.121149>

Basu, P. et al. (2006) *Combustion and gasification in fluidized beds*, Taylor & Francis Group, LLC, chapter (11).

Chang, J., et al. (2013) 'Computational Investigation of a Turbulent Fluidized-bed FCC Regenerator', *Industrial & Engineering Chemistry Research*, 52(11): p. 4000-4010. <https://doi.org/10.1021/ie3013659>

Depypere, F., J.G. Pieters, and K. Dewettinck. (2004) 'CFD analysis of air distribution in fluidised bed equipment', *Powder Technology*, 145(3): p. 176-189. <https://doi.org/10.1016/j.powtec.2004.06.005>

Geldart, D. and J. Baeyens, 'The design of distributors for gas-fluidized beds', *Powder Technology*, 1985. 42(1): p. 67-78. [https://doi.org/10.1016/0032-5910\(85\)80039-5](https://doi.org/10.1016/0032-5910(85)80039-5)

Jaiswal, R., R.K. Thapa, and B.M.E. Moldestad. (2020) 'Study of agricultural waste gasification in an air-blown bubbling fluidized bed using a CPFD model', *In Proceedings of The 61st SIMS Conference on Simulation and Modelling SIMS 2020 - Linköping Electronic Conference Proceedings*, 176(56). <https://doi.org/10.3384/ecp20176393>

Jaiswal, R., et al. (2020) 'A CPFD model to investigate the influence of feeding positions in a gasification reactor', *International Journal of Energy Production and Management*, 5(3): p. 223-233. <https://doi.org/10.2495/EQ-V5-N3-223-233>

O'Rourke, P.J. and D.M. Snider. (2012) 'Inclusion of collisional return-to-isotropy in the MP-PIC method', *Chemical Engineering Science*, 80: p. 39-54. <https://doi.org/10.1016/j.ces.2012.05.047>.

Raza, N., et al. (2021) 'Computational Analysis of the Hydrodynamic Behavior for Different Air Distributor Designs of Fluidized Bed Gasifier', *Front. Energy Res*, 9: p. 692066. <https://doi.org/10.3389/fenrg.2021.692066>

Sasic, S., B. Leckner, and F. Johnsson. (2005) 'Fluctuations and waves in fluidized bed systems: The influence of the air-supply system', *Powder Technology*, 153(3): p. 176-195. <https://doi.org/10.1016/j.powtec.2005.03.012>

Singh, R. and E. Gbordzoe. (2017), 'Modeling FCC spent catalyst regeneration with computational fluid dynamics', *Powder Technology*, 316: p. 560-568. <https://doi.org/10.1016/j.powtec.2016.10.058>

Saxena, S.C., A. Chatterjee, and R.C. Patel, 'Effect of distributors on Gas—Solid Fluidization', *Powder Technology*, 1979. 22(2): p. 191-198. [https://doi.org/10.1016/0032-5910\(79\)80026-1](https://doi.org/10.1016/0032-5910(79)80026-1).

Snider, D.M. (2007) 'Three fundamental granular flow experiments and CPFD predictions', *Powder Technology*, 176(1): p. 36-46. <https://doi.org/10.1016/j.powtec.2007.01.032>.

Snider, D.M. (2001) 'An Incompressible Three-Dimensional Multiphase Particle-in-Cell Model for Dense Particle Flows', *Journal of Computational Physics*, 170(2): p. 523-549. <https://doi.org/10.1006/jcph.2001.6747>.

Weber, J.M., et al. (2013) 'Fluid bed characterization using Electrical Capacitance Volume Tomography (ECVT), compared to CPFD Software's Barracuda', *Powder Technology*, 250: p. 138-146. <https://doi.org/10.1016/j.powtec.2013.10.005>.

Study of gasification behavior for a biorefinery lignin waste in a fluidized bed gasification reactor

Saugata Ghosh^a, Ramesh Timsina^a *, Britt Moldestad^a

^a Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Porsgrunn

*Corresponding author: ramesh.timsina@usn.no

Abstract

Lignocellulosic biorefineries, paper and pulp industries across the globe can convert cellulose and hemicellulose parts of the biomass into higher valued products. However, lignin from biomass is an underutilized biorefinery waste. Value-added applications of lignin waste should be investigated to produce low-molecular-weight compounds as an alternative to petrochemical compounds. Valorization and lignin recovery play an important role in 'green shifts' for such industries. In this work, the authors performed gasification of lignin pellets obtained from one biorefinery located in Finland. A 20-kW pilot-scale bubbling fluidized bed gasifier was used for the experiments. A computational particle fluid dynamics model based on a multi-phase particle in cell approach was developed for the same process. The developed model was validated against the experimental results.

The experimental results showed good conversion of lignin pellets into permanent light gases such as carbon monoxide, hydrogen, methane, etc. The average production of product gas and the lower heating value were 5.74 Nm³/hr and 4.95 MJ/Nm³, respectively. The average molar gas compositions obtained from the experimental study were 0.04 for CH₄, 0.16 for CO, 0.15 for CO₂, 0.13 for H₂ and 0.51 for N₂.

Keywords: lignin waste, lignin gasification, fluidized bed, CPFD

1. Introduction

Biomass is the world's fourth largest primary energy source, after coal, petroleum, and natural gas. Biomass now accounts for around 15% of total global energy use (Ankolekar and Kulkarni, 2018). The carbon present in the biomass is non-fossil in origin and considered as a renewable energy source. Biomass is a biobased material with a wide variety of sources such as lignocellulosic biomass, food waste, fish ensilage, etc. Lignocellulosic biomass primarily consists of cellulose, hemicellulose, and lignin. Lignin forms a part of the secondary cell walls of plants that helps to maintain the structural integrity of the plants. Lignin is the most abundant natural aromatic polymer which is found in most of the terrestrial plants on earth. Lignin accounts for 15–40% of dry weight (Ragauskas *et al.*, 2014) in terrestrial plants. The exact composition and the content of lignin in plants depend upon the botanical species. Millions of tonnes of lignin are produced as a byproduct from the paper, pulp industries, biorefineries, etc every year. The majority of lignin produced from such processes is currently used as a low-cost fuel to balance energy needs for the same process. Annually around 5–36×10⁸ tonnes of lignin are produced across the globe (Gellerstedt and Henriksson, 2008). Lignin residue can be used directly as fuel in combustion furnaces, boilers, etc. or can be used as binders, emulsifiers, concrete

additives, dyestuff dispersants, components for composites and copolymers, etc after chemical transformation (Calvo-Flores and Dobado, 2010). The major challenge for lignin conversion is its heterogeneous properties, such as molecular weight, functionality, and thermal properties depending upon the different sources and processing methods (Saito *et al.*, 2014). There are no significant applications of lignin waste on a commercial scale other than burning it as low-grade boiler fuel for an energy source.

Different alternatives for lignin conversion include biochemical conversion such as hydrolysis, fermentation, etc, and thermochemical conversions such as pyrolysis and gasification. Gasification converts lignin waste into low molecular weight gases that can be used for various applications. Gasification is primarily a thermochemical conversion process where thermal heat breaks down the solid materials into product gases such as carbon monoxide (CO), hydrogen (H₂), methane (CH₄), carbon dioxide (CO₂), etc. The mixture of CO and H₂ is also known as syngas, a common valuable product for the production of alternative fuels. Gasification of lignin waste can handle a certain variation of lignin properties.

A bubbling fluidized bed (BFB) gasifier uses solid bed materials to heat up the solid fuels in the presence of a limited supply of oxidizing agents (air,

oxygen, or steam) to produce light hydrocarbon gases such as CO, H₂, CH₄, CO₂, etc. The BFB gasifiers operate in the temperature range of 700-1100°C (Franco *et al.*, 2003). Good mixing and high heat capacity of the bed material result in a higher heat transfer rate to the biomass particles, thus resulting in higher conversion. The operation of the BFB gasifiers involves multiphase flow, different chemical reactions, and heat transfer. The effect of different parameters and designs of gasifiers can be studied with the help of modelling and simulation in a relatively short time period. Computational fluid dynamics (CFD) is used to model the systems handling the fluid flow. Conventional CFD is a well-accepted technique for single-phase systems while multiphase CFD is used to investigate the systems handling both solids and fluids.

Multiphase CFD models are based on either Eulerian-Eulerian (EE) or Eulerian-Lagrangian (EL) approaches. Both the solid and gas phases are treated as a continuous phases in the EE approach where two phases are differentiated by their volume fraction. This method lacks the discrete nature of solid particles and the transient information of the two-phase interactions (Bin *et al.*, 2009). EL approach preserves the discrete nature of solid particles where the solid phase can exchange the mass, momentum, and energy with the fluid phase, i.e., strong coupling between the phases. The EL approach gives high loading to a computer central processing unit due to the need of tracking individual particles in the system and the requirement of the small-time steps for solving the particle collisions (Ku *et al.*, 2015). The EL approach is computationally expensive and is limited to only 2×10^5 particles (Gidaspow *et al.*, 2004).

Multi-Phase Particle-In-Cell (MP-PIC) modelling incorporates the EL approach that eliminates the need for tracking individual particles. Computational particles for MP-PIC modelling are a group of particles (called parcels) with similar properties such as size, density, residence time, velocity, etc. The parcels are modelled in a discrete frame and the particle interactions are modelled in the Eulerian frame. Barracuda Virtual Reactor (VR) is a commercial software based on the MP-PIC approach, also known as a computational particle fluid dynamics (CPFD) approach. The rapid development of the graphic process unit in computers has made the CPFD simulation capable of simulating the real process in a short time. The major advantage of CPFD is that it can downsize the billions of particles in a large commercial plant to millions of computational particles (Chen *et al.*, 2013). The governing equations for the CPFD approach can be found in the study of Snider *et al.* (Snider *et al.*, 2011).

1.1 Background

Lignin is a major byproduct of biomass-based biorefinery, paper and pulp industries. This led to different researchers focusing on lignin as a low-cost renewable raw material. Biochemical conversion of lignin waste is widely studied over time (Boerjan *et al.*, 2003; Bugg and Rahmanpour, 2015). However, heterogeneity and higher resistance to chemical and physical actions (Ralph *et al.*, 2019; Strassberger *et al.*, 2014) limit the lignin valorization via biochemical conversion pathways. The biochemical conversion technologies still depend upon improvements and innovation in terms of product separation and catalysts (Strassberger *et al.*, 2014). Therefore, in this study, the authors attempt to convert pelletized lignin waste into syngas via gasification, a thermochemical conversion technology. Different technological hurdles must be overcome to make lignin processing feasible, and the authors believe this study gives a step forward toward lignin waste valorization. The present study aims to gasify lignin pellets obtained from a biorefinery in a pilot scale bubbling fluidized bed gasifier located at the University of South-Eastern Norway (USN).

Literature study in the field of lignin gasification gave a limited number of recent studies, suggesting that lignin gasification is an emerging topic. Liakakou *et al.* (Liakakou *et al.*, 2019) have studied the gasification of lignin rich residues in an updraft, bubbling fluidized bed, and MILENA indirect gasifier. The study shows the feasibility of gasification for lignin extracted from a steam explosion (lignin A) and enzymatic hydrolysis (lignin B). Fluidized bed gasification gave the lowest tar content compared to the other two technologies. Different combinations of lignin and gasification technologies were able to give the H₂/CO ratio of 0.6 to 1.0 and the tar yield was between 18-108 g/Nm³. The gasification behaviour of the three technologies was quite different and was able to give the overall gasification behaviour of lignin residue (Liakakou *et al.*, 2019).

Liakakou *et al.* (Liakakou *et al.*, 2021). performed gasification of lignin rich residue obtained after a steam explosion and enzymatic hydrolysis of wheat straw in the MILENA gasifier. The authors compared the results with the gasification of beechwood chips from the same gasifier. The average volume percentage of the major product gas is presented Tab 1.

Table 1: Product gas composition from Liakakou *et al.*

	CO	H ₂	CO ₂	CH ₄	N ₂
Beech wood	28.8	32.1	25.2	8.8	1.2
Lignin	19.8	35.5	24.4	11.4	1.4

The study showed reasonable product gas composition from the gasification of lignin waste

(Liakakou *et al.*, 2021). The basic background on the discussed topic is presented in the beginning. The relevant literature, theory, research gap, and the nobility of this study are presented in Chapter 1. The detailed methods and the properties of the materials are presented in Chapter 2. The results and discussions of this study are described in detail in Chapter 3, and the conclusion is drawn at last.

2. Methodology

2.1. Experimental setup

Experiments were performed at the pilot scale BFB gasifier located at USN Porsgrunn. Fig. 1 shows the isometric view of the BFB gasifier, and the gasifier is designed to operate at atmospheric pressure.

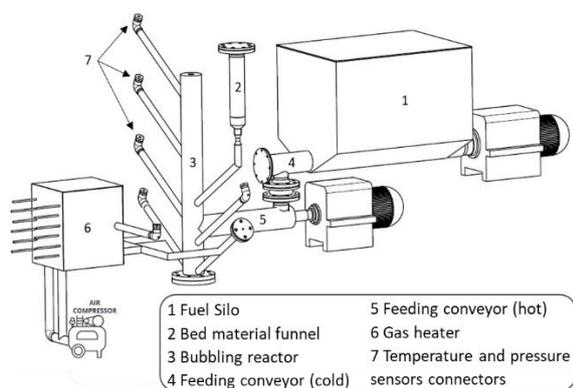


Figure 1: BFB gasifier with auxiliary connections at USN.

The setup consists of a reactor (3), biomass storage silo (1), biomass feeding screws (4 and 5), bed material funnel (2) and a chimney with flare. The setup is facilitated by three electrical heaters with 3kW each, one is the gas heater (6) used to heat the

gasifying agent, and the other two are used to heat up the reactor. The reactor has an inner diameter and height of 0.1 m and 1 m, respectively. An air preheater heats the gasifying agent (compressed air) before entering the reactor to a temperature of around 450°C. Fuel is stored in a sealed silo and is conveyed to the reactor using two screw conveyors. Biomass is fed at a height of 250 mm from the distributor. The gasifier has different pressure and temperature measurement sensors at different positions. The facility also has sensors for the detection of H₂, CO and N₂ to identify any gas leakage. Silica sand in the range of 850 – 1000 μm and density of 2650 kg/m³ was used as bed materials. The initial bed height was 0.22 m. Characterization of the lignin pellets was performed at the Eurofins testing facility to know the elemental composition. The proximate and ultimate analyses are presented in Tab. 2.

Table 2: Characterization of lignin pellets.

Proximate analysis (wt.%, wet basis)	
Fixed carbon	25.68
Volatiles	65.8
Moisture	8.2
Ash	0.32
Ultimate analysis (wt.%, dry basis)	
Ash	0.32
C	54.8
H	6.3
N	0.78
S	0.11
Cl	0.01
O (by difference)	37.68
LHV (MJ/kg, dry)	23.51

Table 3: Reaction kinetics for the CPFDF model.

Reactions/reactions name	Reaction rate kinetics (mol m ⁻³ S ⁻¹)	Enthalpy (kJ/mol)
2C + O ₂ ↔ 2CO		
Char partial oxidation (Snider <i>et al.</i> , 2011)	$4.34 \times 10^7 m_s \text{Texp}\left(\frac{-13590}{T}\right)[\text{O}_2]$	+221.0
C + O ₂ ↔ CO ₂		
Char oxidation (Sreejith <i>et al.</i> , 2015)	$5.7 \times 10^{12} m_s \text{exp}\left(\frac{-4595}{T}\right)[\text{O}_2]^{0.78}$	-394.0
C + CO ₂ ↔ 2CO		
CO ₂ gasification (Thapa <i>et al.</i> , 2014)	$1.12 \times 10^8 m_s P^{0.31} \theta_r \text{exp}\left(\frac{-29518}{T}\right)[\text{CO}_2]$	+172.0
H ₂ + 0.5O ₂ ↔ H ₂ O		
H ₂ oxidation (Bates <i>et al.</i> , 2017)	$5.69 \times 10^{11} \text{exp}\left(\frac{-17610}{T}\right)[\text{H}_2][\text{O}_2]^{0.5}$	-241.9
CO + 0.5O ₂ ↔ CO ₂		
CO oxidation (Xie <i>et al.</i> , 2013)	$5.62 \times 10^{12} \text{exp}\left(\frac{-16000}{T}\right)[\text{CO}][\text{O}_2]^{0.5}$	-283.0
CH ₄ + 1.5O ₂ ↔ CO + 2H ₂ O		
CH ₄ oxidation (Bates <i>et al.</i> , 2017)	$5.0118 \times 10^{11} \text{exp}\left(\frac{-24357}{T}\right)[\text{CH}_4]^{0.7}[\text{O}_2]^{0.8}$	-519.4
CO + H ₂ O ↔ CO ₂ + H ₂		
WGS (Xie <i>et al.</i> , 2013)	$7.68 \times 10^{10} \text{Texp}\left(\frac{-36640}{T}\right)[\text{CO}]^{0.5}[\text{H}_2\text{O}]$	-41.0
CH ₄ + H ₂ O ↔ CO + 3H ₂		
Methane reforming (Solli <i>et al.</i> , 2018)	$3 \times 10^5 \text{exp}\left(\frac{-15042}{T}\right)[\text{CH}_4][\text{H}_2\text{O}]$	+201.9

Lignin pellets were gasified in a bubbling fluidized bed reactor with air as a fluidizing agent. Experiments were performed at different airflow rates (thus giving a different equivalence ratio) and the product gas was analyzed in a gas chromatograph.

2.2 Computational model

A simulation model was developed using Barracuda VR 21.1.1. WenYu-Ergun drag model was used with 40% momentum retention after a particle-particle collision. The reactor was modelled as an open cylinder with a diameter of 0.1 m and a height of 1 m. The developed geometry was divided into 10000 cells with the help of a built-in mesh generator. Fig. 2 shows the boundary conditions, computational cells, and initial bed materials of the developed model.

The formation of the tar and the higher hydrocarbons were neglected during this study. Arrhenius reaction rate model was used to model the chemistry in the reactor. The reaction kinetics is presented in Tab. 3. Major reactions were considered for the model, and their reaction rate kinetics were taken from the literature.

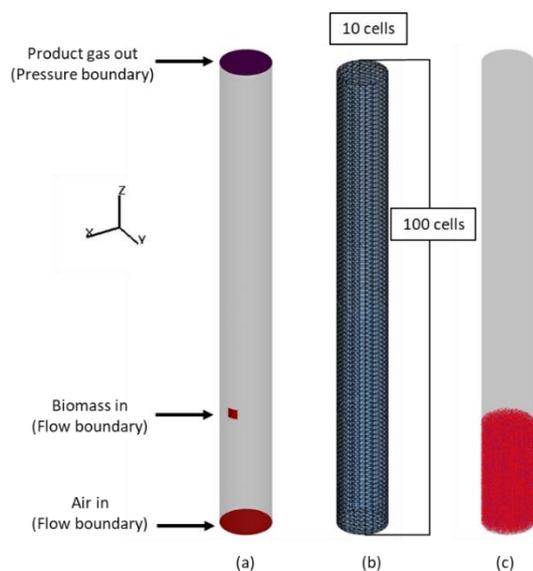


Figure 2. (a) boundary conditions, (b) computational cells and (c) initial bed materials.

3. Results and discussions

The reactor temperature and pressure for a window of 30 minutes during the experiments are presented in Fig. 3.

The results show a certain variation in the bed temperature during the experiments. Biomass was fed via a screw conveyor which gives some degree of variation from the mean feeding rate. Reactor temperature decreases with higher biomass feed rate and increases with lower biomass feed rate. This is because biomass gasification is an endothermic

process, which takes heat from the heated bed materials.

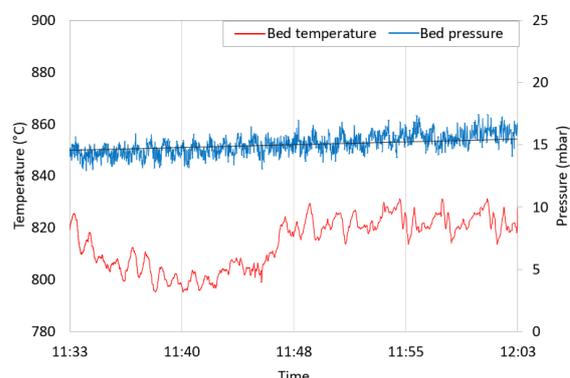


Figure 3: Temperature and pressure variation during the experiment.

A higher biomass feed rate consumes larger amounts of reactor heat, thus lowering the reactor temperature. Lower biomass feeding could lead to a higher reaction rate which in turn increases the reactor temperature. Therefore, variation in reactor temperature gave a certain level of variation in the product gas composition for different samples. Average gas composition was taken for further analysis which minimizes a certain level of measurement uncertainties during the experimental process. The reactor's bed pressure increased slightly with time due to the accumulation of char and ash during the gasification process. As the ash composition was relatively low for the lignin pellets, char gasification is an important aspect of these types of reactors. Char gasification is relatively slow compared to the devolatilization and homogeneous phase reactions (Bermudez and Fidalgo, 2016). Therefore, an alternative configuration such as a dual fluidized bed could be the better alternative to fully utilize the char generated during the gasification process. Char is combusted in a combustor reactor to heat the bed materials which are recirculated back into the bubbling fluidized bed reactor to optimize the biomass conversion.

A simulation model was developed using Barracuda V 21.1.1 (Software[©], 2022) and the simulation results were compared with the experimental results. The average gas compositions for the experimental study and simulation study are presented in Fig. 4. Average molar composition is taken from 40 to 80 seconds for the simulation. The model predicts the fractions of product gases comparatively well and there is a good agreement between the experimental and simulation results regarding the major gas species such as carbon monoxide (CO), carbon dioxide (CO₂), hydrogen (H₂) and nitrogen (N₂). The difference observed between the simulation results and the experimental results could be due to the limited number of reactions considered for the

simulation model, a significant number of reactions occurs during an experimental process. However, the difference is not significant and therefore the developed model was used to study the lignin pellets gasification in depth.

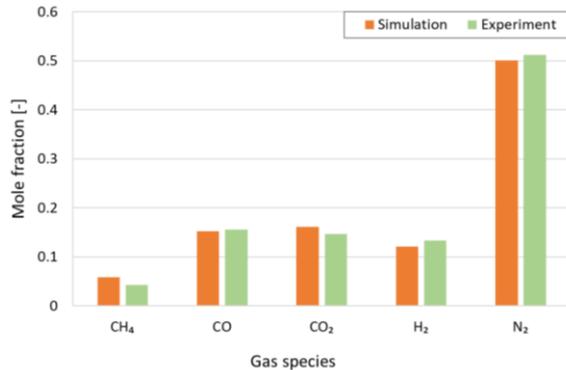


Figure 4. Comparison of average gas species at ER 0.165.

A number of chemical reactions occur during the biomass gasification process. However, only a certain number of major chemical reactions were included in the simulation model for reasonable accuracy and simulation time. The average oxygen concentration during the simulations was zero whereas around 1% of oxygen was present during the experimental analysis. This shows some degree of air contamination during the sampling process. The CFPD model shows a comprehensive result regarding the molar concentration and reactor hydrodynamics during a gasification process. The product gas compositions were monitored during the simulations with respect to time at the reactor outer boundary. Fig. 5 shows the composition of the product gases over simulation time. The reactor reaches steady state conditions after around 10 seconds of simulation time. The CO₂ production started abruptly after around 3 seconds, which counterbalances the O₂ inside the reactor. The high fraction of carbon dioxide at the start represents the combustion process due to the presence of an excess amount of oxygen.

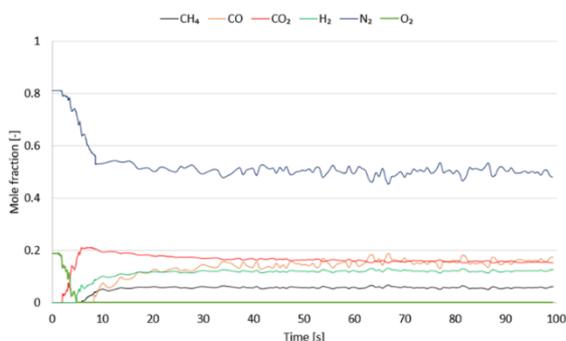


Figure 5. Mole fraction of product gases with respect to time.

The start of the production of CO, H₂ and CH₄ after around 10 seconds of simulation time represents the start of the gasification process. The process gradually shifted from combustion to gasification process. The average molar composition fluctuates around their mean value at steady state conditions. A certain level of variation in the composition illustrates different physical and chemical transformations inside the reactor.

The product gas composition during the simulations was monitored in the reactor. Fig. 6 shows the mole fraction for methane, carbon monoxide and hydrogen at the center plane of the reactor.

Non-distinguishable gas variation inside the bed represents that the heterogeneous reactions are less significant as compared to the homogeneous phase reactions. This suggests that the heterogeneous reactions in the gasification process are slow and are often considered as a rate-limiting process compared to the homogeneous phase reactions. Certain abrupt changes in gas composition can be seen in the biomass feeding regions. This is the region where biomass first enters the reactor, thus altering the flow behaviour. A number of phenomena occur simultaneously giving distinct chemical variations in this region.

Additionally, devolatilization and char mixing with bed materials give different dynamics and properties variation in this region. A certain level of variation in gas composition can be seen in the freeboard regions. The carbon monoxide concentration increased slightly along with the reactor height that indicating the dominance of methane reforming reaction and backward water gas shift reaction.

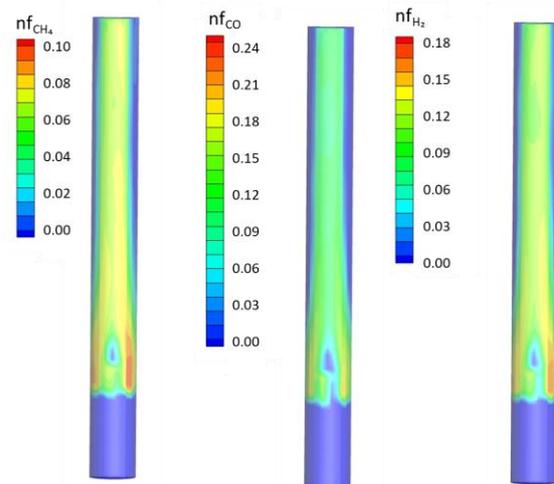


Figure 6. Product gas composition (mole fraction) along xz plane (at y=0.05m).

Gasification of lignin is an emerging aspect of lignin valorization. The literature study gave only certain studies on this topic. It is important to compare the quality of the product from this study to the results

published in the literature. The mole percentage of major gas species excluding N₂, C₂H₂, C₂H₄, C₂H₆ and other minor gas components is compared in Tab. 4.

Table 4: Comparison of molar percentage.

	CH ₄	CO	CO ₂	H ₂
Current study	9.05	32.51	30.59	27.85
Liakakou et al	12.51	21.73	26.78	38.97

As seen from the table, there is considerable variation between the two results. Liakakou et al have used the MILENA gasifier, a dual fluidized bed gasifier operated with steam as a fluidizing agent. However, the air was used as the gasifying agent in the current study which gives lower hydrogen production compared to the MILENA gasifier. Further, variation in operating conditions, bed materials, equivalence ratio, etc would normally give the variation in the product composition. Therefore, further experiments in different gasifiers at different conditions would give a broader perspective of the generated product gas applications.

4. Summary and conclusions

This study gave a comprehensive result for the air gasification of lignin pellets. Experiments were performed in a pilot scale BFB gasification reactor. A simulation model based on MP-PIC modelling approach was developed to study the reaction chemistry and the bed hydrodynamics inside the reactor. The average product gas obtained from the model was validated against the experimental results. The product gas varied with a variation of equivalence ratio in the gasifier. The simulation study showed a higher fraction of CO₂ at the start, suggesting combustion has taken place. The conversion shifted towards gasification at steady-state conditions. The average gas compositions from the experimental study were 0.04 for CH₄, 0.16 for CO, 0.15 for CO₂, 0.13 for H₂ and 0.51 for N₂. The lower heating value and the carbon conversion efficiency were 4.95 MJ/Nm³ and 40.3% respectively. Different sets of experiments are needed to generalize the product gas quality and quantity from the gasification of lignin waste. The developed CPFD model can be used to optimize the model in terms of gasifier performance, fuel flexibility, and reduced emissions. The model can be used to scale up the reactor to an industrial scale reactor to investigate the behaviour at a commercial scale.

Acknowledgment

The authors would like to thank the University of South-Eastern Norway (USN) for providing the necessary software for the simulation works.

References

- Ankolekar, V., and Kulkarni, S. (2018). Briquetting of agricultural biomass: an overview. *International Journal for Research in Applied Science and Engineering*, 6(2), 1681-1685.
- Bates, R. B., Ghoniem, A. F., Jablonski, W. S., Carpenter, D. L., Altantzis, C., Garg, A., . . . Field, R. P. (2017). Steam-air blown bubbling fluidized bed biomass gasification (BFBBG): Multi-scale models and experimental validation. *AIChE Journal*, 63(5), 1543-1565. doi:10.1002/aic.15666
- Bermudez, J. M., and Fidalgo, B. (2016). 15 - Production of bio-syngas and bio-hydrogen via gasification. In R. Luque, C. S. K. Lin, K. Wilson, & J. Clark (Eds.), *Handbook of Biofuels Production (Second Edition)* (pp. 431-494): Woodhead Publishing.
- Bin, C., Cong, W., Zhiwei, W., and Liejin, G. (2009). Investigation of gas-solid two-phase flow across circular cylinders with discrete vortex method. *Applied Thermal Engineering*, 29(8-9), 1457-1466. doi:10.1016/j.applthermaleng.2008.06.025
- Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annual review of plant biology*, 54(1), 519-546.
- Bugg, T. D., and Rahmanpour, R. (2015). Enzymatic conversion of lignin into renewable chemicals. *Current opinion in chemical biology*, 29, 10-17.
- Calvo-Flores, F. G., and Dobado, J. A. (2010). Lignin as renewable raw material. *ChemSusChem*, 3(11), 1227-1235.
- Chen, C., Werther, J., Heinrich, S., Qi, H.-Y., and Hartge, E.-U. (2013). CPFD simulation of circulating fluidized bed risers. *Powder Technology*, 235, 238-247. doi:10.1016/j.powtec.2012.10.014
- Franco, C., Pinto, F., Gulyurtlu, I., and Cabrita, I. (2003). The study of reactions influencing the biomass steam gasification process. *Fuel*, 82(7), 835-842.
- Gellerstedt, G., and Henriksson, G. (2008). Lignins: major sources, structure and properties. *Monomers, polymers and composites from renewable resources*, 201-224.
- Gidaspow, D., Jung, J., and Singh, R. K. (2004). Hydrodynamics of fluidization using kinetic theory: an emerging paradigm: 2002 Flour-Daniel lecture. *Powder Technology*, 148(2-3), 123-141. doi:10.1016/j.powtec.2004.09.025
- Ku, X., Li, T., and Løvås, T. (2015). CFD-DEM simulation of biomass gasification with steam in a fluidized bed reactor. *Chemical Engineering Science*, 122, 270-283. doi:10.1016/j.ces.2014.08.045
- Liakakou, E., Infantes, A., Neumann, A., and Vreugdenhil, B. (2021). Connecting gasification with syngas fermentation: Comparison of the performance of lignin and beech wood. *Fuel*, 290, 120054.
- Liakakou, E., Vreugdenhil, B., Cerone, N., Zimbardi, F., Pinto, F., André, R., . . . Girio, F. (2019). Gasification of lignin-rich residues for the production of biofuels via syngas fermentation: Comparison of gasification technologies. *Fuel*, 251, 580-592.
- Ragauskas, A. J., Beckham, G. T., Biddy, M. J., Chandra, R., Chen, F., Davis, M. F., . . . Keller, M. (2014). Lignin valorization: improving lignin processing in the biorefinery. *science*, 344(6185), 1246843.
- Ralph, J., Lapierre, C., and Boerjan, W. (2019). Lignin structure and its engineering. *Current opinion in biotechnology*, 56, 240-249.
- Saito, T., Perkins, J. H., Vautard, F., Meyer, H. M., Messman, J. M., Tolnai, B., and Naskar, A. K. (2014). Methanol fractionation of softwood kraft lignin: Impact on the lignin properties. *ChemSusChem*, 7(1), 221-228.

Snider, D. M., Clark, S. M., and O'Rourke, P. J. (2011). Eulerian–Lagrangian method for three-dimensional thermal reacting flow with application to coal gasifiers. *Chemical Engineering Science*, 66(6), 1285-1295.

Software[®]. C. (2022). CPFD Software. Retrieved from <https://cpfd-software.com/>

Solli, K.-A., Thapa, R. K., and Moldestad, B. M. E. (2018). *Screening of Kinetic Rate Equations for Gasification Simulation Models*. Paper presented at the Proceedings of The 9th EUROSIM Congress on Modelling and Simulation, EUROSIM 2016, The 57th SIMS Conference on Simulation and Modelling SIMS 2016.

Sreejith, C., Muraleedharan, C., and Arun, P. (2015). Air–steam gasification of biomass in fluidized bed with CO₂ absorption: A kinetic model for performance prediction. *Fuel Processing Technology*, 130, 197-207.

Strassberger, Z., Tanase, S., and Rothenberg, G. (2014). The pros and cons of lignin valorisation in an integrated biorefinery. *Rsc Advances*, 4(48), 25310-25318.

Thapa, R., Pfeifer, C., and Halvorsen, B. J. I. J. E. E. (2014). Modeling of reaction kinetics in bubbling fluidized bed biomass gasification reactor. 5, 35-44.

Xie, J., Zhong, W., Jin, B., Shao, Y., and Huang, Y. (2013). Eulerian–Lagrangian method for three-dimensional simulation of fluidized bed coal gasification. *Advanced Powder Technology*, 24(1), 382-392.

Oil Production Forecasting with Uncertainty Description Using Data Driven Proxy Model

Javad Tavakolifaradonbe*, Ali Moradi, Britt M. E. Moldestad

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway
h.javadt@yaho.com, ali.moradi@usn.no, britt.moldestad@usn.no

Abstract

The petroleum industry operates under great uncertainty. Achieving an efficient approach to quantify uncertainty in oil production models is of key importance in supporting decision-makers to find suitable strategies for mitigating risks and maximizing profit. Uncertainty quantification is commonly performed based on the Monte Carlo approach and this is a very time-consuming process by using the physics-based models developed by reservoir simulators. To solve this challenge, data-driven proxy models which are less complex and computationally efficient can be used as an alternative. This paper aims to investigate the functionality of the ANN method in developing proxy models for uncertainty quantification of oil production from advanced wells. The investigation is conducted through a case study for uncertainty assessment of cumulative oil and water productions from a long horizontal well with ICD completion and zonal isolation in a synthetic reservoir for 10 years. In this study, the Eclipse® reservoir simulator is used for developing the base case model and it is coupled with MATLAB® for generating the required data sets to train and test the ANN proxy model. According to the obtained results, the trained and developed ANN proxy model can predict the production of oil and water from advanced wells accurately with a mean error of less than 4%. Besides, the proxy model is 150 times faster than the Eclipse model and can solve the challenge of the time-consuming process of uncertainty quantification.

1. Introduction

Nowadays, to develop long-term oil production models, engineers need to predict the behavior of reservoirs by utilizing geological features [1]. However, the defining variables have a wide range and are many and various. Thereby integrating the parameters one by one cannot describe the reservoirs accurately. Hence some simplifications should be applied to solve the model's difficulties [1], [2]. First, the most impactful parameters should be determined. Sensitivity analysis is a good solution for figuring out the input parameters with the most likely influence [3]. By considering some assumptions on variables and their distribution, numerical models will be generated to simulate the reservoir. However, there is still an unsolved issue, namely the treatment of uncertainty.

Defeating the uncertainty in anticipating the production requires a huge number of simulations, which is unfeasible because of the long computation time. [4], [5]. Different techniques have been focused on, and among them, proxy models have received more attention. The first proxy model was innovated by utilizing a bilinear polynomial of inputs. These analytical models have been trained and developed to behave like simulators while consuming less time. In this way prediction, analysis, and finally optimization will be performed more efficiently [6], [7].

So far it can be mentioned that accuracy and acceleration are two important characteristics that should be considered during the reservoir

simulations. Artificial Neural Networks (ANNs) have been introduced as a practical solution. In a study, ANNs were applied as a proxy model to assess the uncertainty in production prediction [8]. Another study investigated different architectures of the neural networks in reducing the time consumption of reservoir simulation [9]. Artificial neural network separately or in combination with the genetic algorithm was utilized to grab nonlinearities of problems [2], [10]. Apart from optimization of the algorithm, there are several studies related to the application of ANN in engineering and production. Shaik et al. [2] predicted the life time of a pipeline by applying ANN. Otchere et al. [7] forecasted the features of a petroleum reservoir by using supervised machine learning paradigms. Moreover, the application of neural networks in production prediction was also proposed by Yuan et al. [11]. Through all the previous studies it is mentioned that the quality and accuracy of a proxy model highly depend on the training step.

This study focuses on the applicability of ANN as a proxy model for assessing the uncertainty in production prediction. Based on the Design of Experiments (DOE), a set of reliable data is produced by coupling MATLAB and Eclipse. Then the proxy model is trained, and the trained model is used to assess the uncertainty based on the Monte Carlo sampling principle. The main purpose of this paper is to propagate a methodology to allow for a more reliable decision about the productivity of a reservoir based on geological parameters.

1.1. Definition of Artificial Neural Network

A neural network is simply a set of neurons that are connected. Each neuron, as it is depicted in Fig. 1, takes one or more inputs and gives an output based on defined functions. These functions may be Relu, Sigmoid, Gaussian, or the sign functions. It should be mentioned that synaptic weight is considered for different neurons [4].

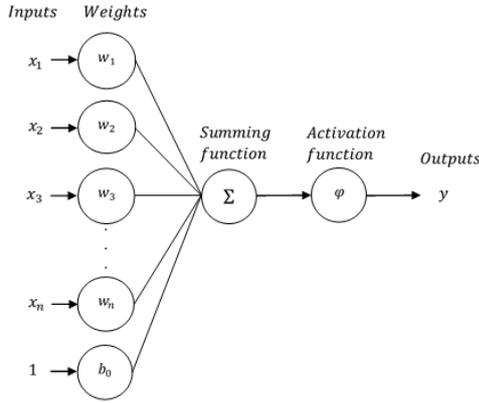


Figure 1: Schematic of a artificial neuron.

According to Fig. 1, x represents inputs, w is allocated to weight vector and b_0 is bias. Mathematically each neuron should be displayed by its function (f). So, if a neuron receives n input it would be represented as below:

$$f: R^{n+1} \times R^n \rightarrow R \quad (1)$$

Satisfying

$$\begin{aligned} &1: g: R \rightarrow R \\ &2: W \in R^{n+1}, W = (w_1, w_2, \dots, w_n, b_0) \\ &3: \forall x \in R^n, f(W, x) = g\left(\sum_{i=1}^n w_i z_i + b_0\right) \\ &x = (z_1, \dots, z_n) \end{aligned} \quad (2)$$

Where g is the transfer function. This basic function could model the higher-order functions by utilizing the collective behavior of a set of neurons which is called a layer. Indeed, a network is made of multilayer, consisting of an input layer, one or more intermediate or hidden layers, and an output layer, while each layer is composed of several neurons (Fig. 2).

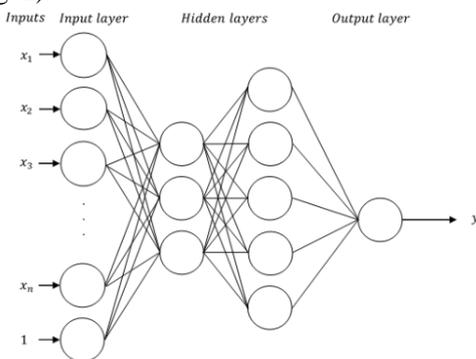


Figure 2: Schematic diagram of a neural network with two hidden layers.

Learning is the most important ability of a neural network, and a neural network will be able to generalize, classify and foresee [4], [12]. In other words, because of having experienced, neural networks will have recognition ability. But according to the learning, networks are divided into two classes, supervised and unsupervised networks. In a supervised learning class, input and output are fed into the network at the same time [7]. Then, the machine will learn how to reconfigure itself. On the other side, under unsupervised learning, the proxy is exposed to unlabeled input solely for clustering or comparing. It should be mentioned that to make the quadratic error of output at least, backpropagation is utilized in the supervised learning network, indeed it is a method of more accurate weight calculation [13].

Generally, based on the structure of the network and the operation of neurons, neural networks carry out a quite simple differentiable function. Indeed, after the learning phase and stabilizing the weight, the machine as a black box forecasts the phenomenon for new inputs [9]. Despite all of these, there are still some deterrents against utilizing ANNs. In other words, configuring the architecture of ANNs, namely the number of layers and number of neurons in each layer, should be found, while there is not any identification of a better architecture [5].

2. Methodology

2.1. Creating Proxy Model

Fig. 3 represents the steps of an algorithm to model a proxy. The Data sets and the previously proxy algorithms are considered as the most important factors for qualifying a proxy model. To make sure that all aspects of the model are dealt with, an infinite size dataset is required, which is practically impossible [3], [9], [12]. Some techniques of experimental design are enlisted to extract the utmost information with the least simulations.

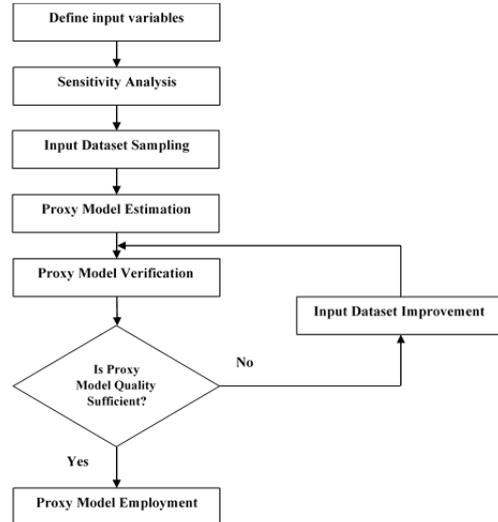


Figure 3: Schematic proxy model development [14].

Selecting the input variable highly depends on the type of problem and the level of knowledge of the project. It is recommended to consider all input variables at the beginning and then omit unimportant parameters through the sensitivity analysis step. Indeed, sensitivity analysis filters out less significant parameters on a simulation model. Consequently, an appropriate dataset will be prepared.

The accuracy of the proxy model highly originates from training. By utilizing a decent dataset sampling, the proxy model will be trained appropriately, and in the following, the estimation will be performed accurately. Thereby, verifying the model which is based on prediction accuracy will be satisfied [14].

2.1.1. Architecture of Artificial Neural Network

A neural network is considered a good proxy when it predicted a new case with acceptable error. Therefore, evaluation of the model should be performed sequentially to avoid overtraining. In this way, cross-validation, as one of the most popular methods, lets defined architecture stop learning when a validation error is raised [11], [12].

Overtraining also comes from a poorly structured network. Thereby identifying the appropriate number of hidden layers and their neurons is so requisite [1]. Moreover, the complexity of neural networks should also be limited. For this purpose, after finishing the learning phase, the pruning method will eliminate the connections with the smallest effect on the output error.

To obtain an optimal neural network for the defined method, both pruning and cross-validation were utilized.

2.2. Development of The Physic-based Model With Uncertainty Description

2.2.1. Defining Uncertain Input Domains

This study is conducted through modeling and forecasting oil and water production from an advanced horizontal well in a synthetic reservoir with uncertain properties for 10 years. The reservoir properties are the model inputs, and it is assumed that the value of some of the properties is uncertain. The uncertain reservoir parameters with their uncertainty range are reported in Tab. 1.

Table 1: Uncertain reservoir properties with their range.

Parameter	Min	Mean	Max
Porosity	0.15	0.23	0.27
Permeability in x-dir. [mD]	200	500	1000
Permeability in y-dir. [mD]	150	600	1200
Permeability in z-dir. [mD]	20	100	500
Irreducible water sat.	0.1	0.15	0.2
Residual oil saturation	0.05	0.1	0.15
Max. rel. perm. of water	0.2	0.4	0.5
Max. rel. perm. of oil	0.85	0.95	1

Initial water saturation	0.12	0.2	0.25
Capillary pressure [bar]	4	2.7	2
Aquifer prod. Index [m ³ /d/bar]	2000	10000	15000

2.2.2. Determining the Most Impactful Uncertain Parameters

Uncertainty quantification based on the Monte Carlo approach requires many simulations. For each simulation run, a random combination of model input values is chosen, and the corresponding model outputs are calculated by using the simulator. This is a very time-consuming process when the system has several inputs. By filtering the less important inputs out and focusing on the most impactful input variables on the accuracy of the models, a bit of prediction accuracy is sacrificed but the speed of uncertainty assessment highly increases. The sensitivity analysis assesses the contribution of the uncertainty of each model input to the accuracy of the model outcomes and identifies the most important parameters of the system. Cumulative oil and water production are the most important outputs of oil models and are the model outputs in this paper. By performing sensitivity analysis on the uncertain reservoir parameters given in Tab. 1, the sensitivity coefficient of each reservoir parameter for the cumulative oil and water production is calculated. The obtained results are depicted as a tornado diagram in Fig. 4. Based on the presented results, the five most important input variables for predicting oil and water production are determined and given in Tab. 2. These input variables are the model inputs for the proxy model development and uncertainty quantification.

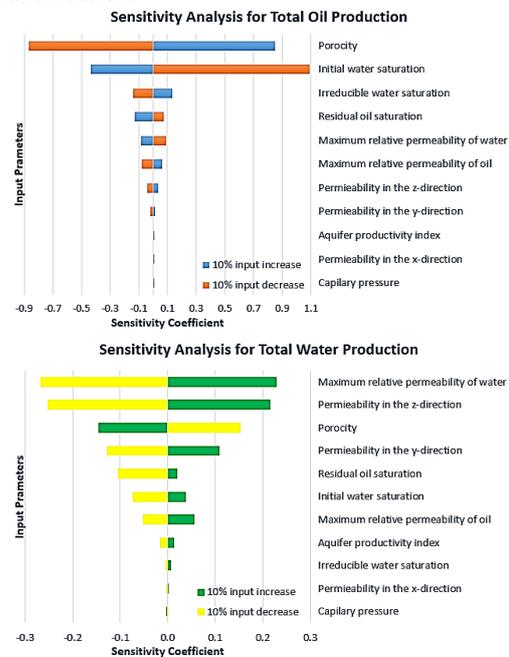


Figure 4: Sensitivity analysis of uncertain reservoir parameters.

Table 2: Uncertain input variables for uncertainty assessment.

Parameter	Min.	Mean	Max.
Porosity	0.15	0.23	0.27
Irreducible water saturation	0.1	0.15	0.2
Initial water saturation	0.12	0.2	0.25
Permeability in z-dir	20	100	500
Max. rel. perm. of water	0.2	0.4	0.5

2.2.3. Development of The Physics-based Model in Eclipse.

In this paper, the production prediction and uncertainty assessment are performed for primary oil production from a medium oil reservoir with a water drive. The reservoir fluid properties, as well as the temperature and pressure of the reservoir, are given in Tab. 3.

Table 3: Reservoir characteristics and fluid properties.

Parameter	Value
Oil density and Viscosity	900 kg/m ³ , 2.5 cP
Water density and Viscosity	1050 kg/m ³ , 0.45 cP
Gas-oil ratio (GOR)	50
Temperature and Pressure	60 °C, 200 bar

It is assumed that oil is produced from the reservoir near an advanced horizontal well with a length of 1000 m completed with Inflow Control Devices (ICDs) and zonal isolation. The diameters of the wellbore, the production tubing, and the ICDs are 8.5 inches, 5.5 inches, and 0.01 m respectively. The thickness and width of the reservoir are assumed to be 30 m and 70 m respectively. It is also assumed that the well is located 5.5 m below the top of the drainage area. The schematic of the near-well reservoir is shown in Fig. 5.

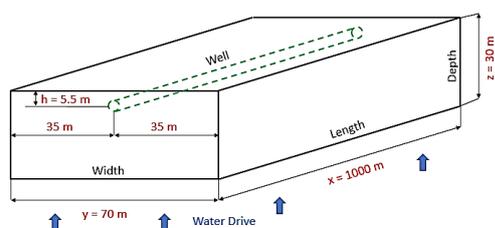


Figure 5: Schematic of the near-well reservoir.

To achieve a suitable grid setup, in the Y and Z directions finer meshes have been set near the wellbore and uniform meshes are considered in the X-direction. It is assumed that the horizontal well has 8 equivalent joints, each 125 m long. As a result, 8 uniform cells are considered for the reservoir in the X-direction. The grid resolution in Y and Z directions is illustrated in Fig. 6.

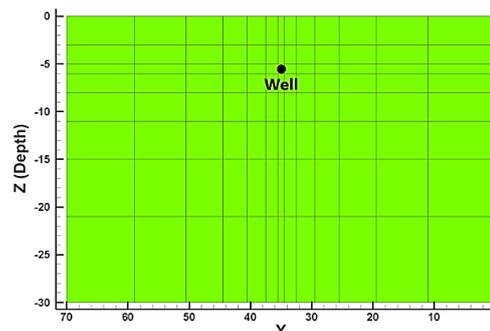


Figure 6: Grid resolution in the Y-Z plane.

In this study, Eclipse® which is a robust physics-based simulator is applied as a simulation tool. Due to the high pressure and low temperature of the reservoir, the reservoir condition is located well to the left-hand side of the critical point, and the black-oil model can be used for modeling fluid flow from the reservoir to the production tubing. Moreover, the multisegmented well model in the Eclipse simulator is used for developing the well model with ICD completion and zonal isolation. The well is considered to be controlled by the Bottom Hole Pressure (BHP), and the BHP is assumed to be 190 bar. Based on the mean value of the reservoir properties and the mentioned considerations and assumptions, a base model is developed in the Eclipse simulator. The cumulative oil and water production based on the base model for 10 years is shown in Fig. 7.

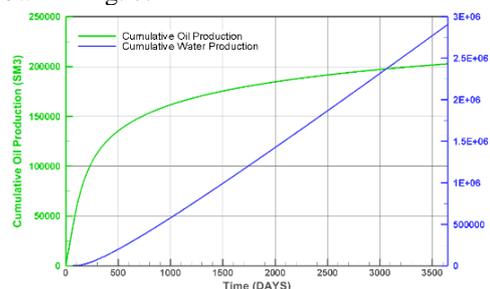


Figure 7: Base model water and oil production prediction for 10 years.

3. Results

3.1. Training and Test the ANN

The training data sets were generated by the Eclipse reservoir simulator. Inputs consisted of 5 variables and each variable accounted for 5 values, which gives the data sets with a size of 5⁵. For each variable, the min. value, max. value, mean value, a value between min. and mean, and a value between mean and max. were opted. In addition, the main dataset accounted for 67% training, 25% validation, and 8% test.

Before feeding inputs to the machine for training, all values were normalized between [0,1] based on min value and max value. Outputs in the dataset

accounted for accumulative oil and water production through 10 years. Because of that, for predicting fluid production of each year, fluid production of the previous year was considered as input too, Fig. 8.

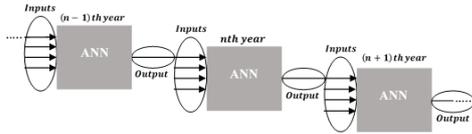


Figure 8: Schematic flow chart of predicting in sequential years.

For testing the proxy, the values of inputs opted ununiformly in a way that the machine had never experienced, although these values were between the minimum and the maximum values.

During testing the machine, prediction errors were calculated point by point within 10 years and for each state. State refers to the specific set of an input.

$$Error = \frac{simulation - prediction}{simulation} \quad (3)$$

The machine predicted the oil production better than the water production. The mean error values are presented as a horizontal line in Fig. 9, 0.96 and 3.79 for oil and water respectively.

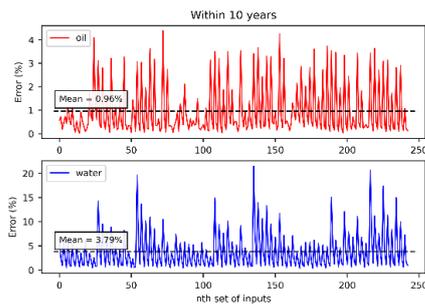


Figure 9: Prediction error for testing the proxy model.

3.2. Uncertainty Assessment in Production Prediction

Input datasets play a great role in studying uncertainty in production prediction. In addition, the uncertainty of input parameters makes it necessary to use a probabilistic approach [4], [6]. Although there are different methods for probabilistic data sampling, Latin Hypercubic Sampling (LHS) was chosen as an efficient sampling method [14]. Indeed, by considering the min and max of 5 effective parameters, the 100000 most probable sets of inputs were extracted according to the LHS approach. In this way, P (10), P (50), and P (90) are to be predicted.

After performing 100000 predictions for each time step. Cumulative probability distribution and probability density for accumulative oil and water production within 10th year is shown in Fig. 10, 11, 12 and 13.

Fig. 10 depicts a normal distribution for accumulative oil production. On the other side, Fig. 12 shows a lag normal distribution of water production.

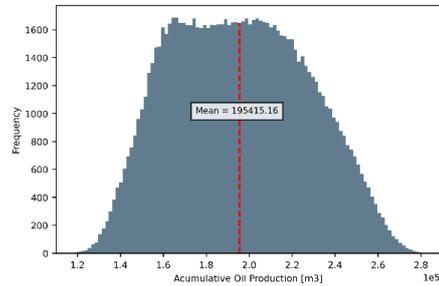


Figure 10: Probability distribution of cumulative oil production after 10 years.

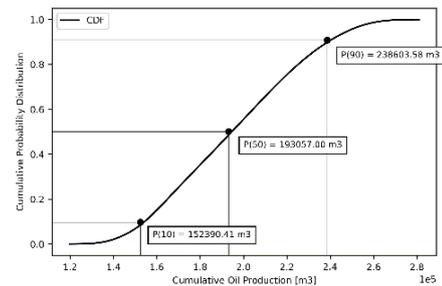


Figure 11: Cumulative probability distribution diagram for predicting oil production after 10 years.

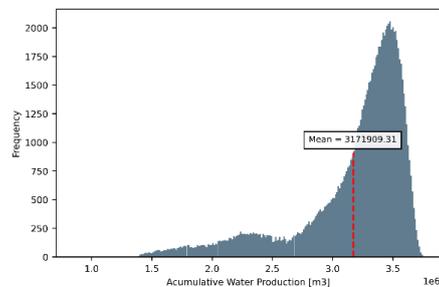


Figure 12: Probability distribution of cumulative water production after 10 years.

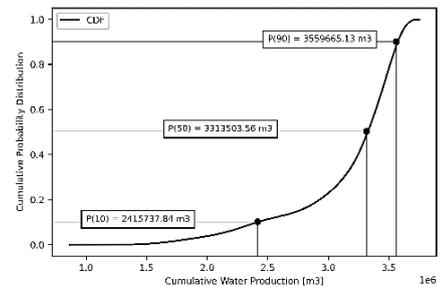


Figure 13: Cumulative probability distribution diagram for predicting water production after 10 years.

As it is presented in Tab. 4, there is a 90% chance to produce less than $2.4E+05$ m³ oil and less than $3.6E+06$ m³ water after 10 years. In addition, cumulative oil and water production will be less than $1.5E+05$ m³ and $2.4E+06$ m³ respectively with a chance of 10%. In this case, the best estimation (P50) for cumulative oil and water production is $1.9E+05$ m³ and $3.3E+06$ m³ respectively.

Table 4: Summary of uncertainty results.

	P (10)	P (50)	P (90)
Oil	1.5E+05	1.9E+05	2.4E+05
Water	2.4E+06	3.3E+06	3.6E+06

4. Conclusion

As the oil reservoirs are developing, uncertainty assessment has become a priority. An unreliable prediction of the producibility of a reservoir may cause investment bankruptcy. Although this type of reservoir engineering problem could be solved by applying a reservoir simulator, the time-consuming process is a significant deterrent. Therefore, an artificial neural network with a stochastic approach has been enlisted to analyze uncertainty.

This paper presents an appropriate methodology to deal with assessing the uncertainty during the fluid production prediction. The paper compiles MATLAB and Eclipse and built an amendable optimal Neural Network by utilizing mathematical procedures, to reduce the time consumption of data extraction and prediction of an oil reservoir.

The results show that an appropriate proxy model can predict the production with an acceptable error of less than 4%. In addition, when utilizing ANN, the time consumption was reduced by a ratio of 150 times. It is also concluded that by increasing the size of a dataset, the time-consumption effectiveness of ANN will raise.

References

- [1] J. Navrátil, A. King, J. Rios, G. Kollias, R. Torrado, and A. Cudas, "Accelerating Physics-Based Simulations Using End-to-End Neural Network Proxies: An Application in Oil Reservoir Modeling," *Frontiers in Big Data*, vol. 2, p. 33, 2019, doi: 10.3389/fdata.2019.00033.
- [2] T. Foroud, A. Seifi, and B. AminShahidi, "Assisted history matching using artificial neural network based global optimization method - Applications to Brugge field and a fractured Iranian reservoir," *Journal of Petroleum Science and Engineering*, vol. 123, 2014, doi: 10.1016/j.petrol.2014.07.034.
- [3] A. F. Teixeira and A. R. Secchi, "Machine learning models to support reservoir production optimization," *IFAC-PapersOnLine*, vol. 52, no. 1, pp. 498–501, 2019, doi: 10.1016/j.ifacol.2019.06.111.
- [4] D. Guérillot and J. Bruyelle, "Uncertainty assessment of hydrocarbon in place with geological models satisfying gravity-capillary equilibrium," *Society of Petroleum Engineers - SPE Kuwait Oil and Gas Show and Conference 2019, KOGS 2019*, 2019, doi: 10.2118/198099-ms.
- [5] D. D. Monteiro *et al.*, "Uncertainty analysis for production forecast in oil wells," *SPE Latin American and Caribbean Petroleum Engineering Conference Proceedings*, 2017, doi: 10.2118/185550-ms.
- [6] C. Wei *et al.*, "Uncertainty assessment in production forecasting and optimization for a giant multi-layered sandstone reservoir using optimized artificial neural network technology," *Society of Petroleum Engineers - SPE Reservoir Characterisation and Simulation Conference and Exhibition, RCSC 2017*, pp. 1179–1188, 2017, doi: 10.2118/186069-ms.
- [7] D. A. Otchere, T. O. Arbi Ganat, R. Gholami, and S. Ridha, "Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models," *Journal of Petroleum Science and Engineering*, vol. 200, 2021, doi: 10.1016/j.petrol.2020.108182.
- [8] D. R. Guérillot and J. Bruyelle, "History Matching Methodology Using an Optimal Neural Network Proxy and a Global Optimization Method," vol. 2016, no. 1, pp. 1–5, 2016, doi: <https://doi.org/10.3997/2214-4609.201602403>.
- [9] P. C. Silva, C. Maschio, and D. J. Schiozer, "Use of Neuro-Simulation techniques as proxies to reservoir simulator: Application in production history matching," *Journal of Petroleum Science and Engineering*, vol. 57, no. 3, pp. 273–280, 2007, doi: <https://doi.org/10.1016/j.petrol.2006.10.012>.
- [10] D. R. Guérillot and J. Bruyelle, "Uncertainty assessment in production forecast with an optimal artificial neural network," in *SPE Middle East Oil and Gas Show and Conference, MEOS, Proceedings*, 2017, vol. 2017-March, doi: 10.2118/183921-ms.
- [11] Z. Yuan, H. Huang, Y. Jiang, and J. Li, "Hybrid deep neural networks for reservoir production prediction," *Journal of Petroleum Science and Engineering*, vol. 197, 2021, doi: 10.1016/j.petrol.2020.108111.
- [12] N. B. Shaik, S. R. Pedapati, and F. A. B A Dzubir, "Remaining useful life prediction of a piping system using artificial neural networks: A case study," *Ain Shams Engineering Journal*, 2021, doi: <https://doi.org/10.1016/j.asej.2021.06.021>.
- [13] L. A. N. Costa, C. Maschio, and D. Schiozer, "Study of The Influence of Training Data set in Artificial Neural Network Applied to The History Matching Process," 2010.
- [14] D. I. Zubarev, "Pros and cons of applying proxy-models as a substitute for full reservoir simulations," *Proceedings - SPE Annual Technical Conference and Exhibition*, vol. 5, pp. 3234–3256, 2009, doi: 10.2118/124815-ms.

Simulation of adsorption and desorption of VOC on activated carbon

Sviatoslav Eroshkin^a, Eivind Johannessen^b, Even Solbraa^c

^{a, c} Department of Energy and Process Engineering, Norwegian University of Science and Technology

^{b, c} Equinor

eroshkin.sviatoslav@gmail.com

Abstract

The study's primary purpose is to simulate the adsorption and desorption of volatile organic components on activated carbon bed using python. The model was verified with experiments for the separation of methane (55 mol %) - carbon dioxide (45 mol %) mixture using vacuum pressure swing adsorption on an activated carbon molecular sieve. For a six-component mixture (methane, ethane, propane, butane, carbon dioxide and nitrogen), the model was verified using Aspen Adsorption flowsheet simulator and experimental results of vacuum pressure swing adsorption on activated carbon. Even though the model showed a relatively low error in comparison with provided experiments, some experimental cases need to be investigated more to get a better model prediction.

1. Introduction

Adsorption of VOC (Volatile Organic Components) on a fixed-bed AC (Activated Carbon) is commonly used to reduce VOC emissions. VRU (Vapor Recovery Units) gained popularity in the 1990s. Today 95 % of all new VRU are based on AC adsorption followed by vacuum regeneration.

Many papers have studied the PSA (Pressure Swing Adsorption) for other components than VOC. From the technical point of view, these works are interesting because the mechanism of VOC VPSA (Vacuum Pressure Swing Adsorption) is in many ways similar. The PSA process exploits the change of the adsorption equilibrium with a change in system pressure. The process efficiency depends on many factors such as the number, sequence, and time of PSA steps, flow rate of the gas, gas composition, adsorption pressure, and others. For example, Kim et al. (2015) investigated biogas mixture adsorption on AC molecular sieve with four beds and a seven steps PSA system. A significant improvement to the PSA process was the equalization step: reducing the energy consumption in the pressurization step by employing the purified product from the second adsorber. The experimental setup and numerical model showed a cyclic steady-state process after 13 cycles.

Furthermore, Kim et al. (2015) indicate that optimum conditions for the separation highly depend on such parameters as the adsorption pressure, desorption pressure, purge, and feed rates (the ratio of the purge gas flow rate to the flow rate

of desirable component feed rate). Ahn et al. (2013) also showed the importance of purge to feed ratio, adsorption pressure, feed flow rate, step times, and carbon ratio for layered two- and four-bed 6-step PSA processes for H₂ recovery from coal gas. The result of the study is that higher purity of the product can be achieved with layered beds; however, less recovery can also be observed.

One of the main difficulties in mathematical adsorption modeling is the correct definition of adsorption thermodynamics and mass transfer. The EL (extended Langmuir) (Duong, 1998) isotherm is widely used due to the model's simplicity. However, the accuracy of this isotherm type is questionable. Even though it has a theoretical basis behind it, the assumptions place significant restrictions on the applicability of this isotherm. The EL model is an explicit model that is preferred over its implicit counterparts, for example, IAST (Ideal Adsorption Solution Theory), due to the computation complexity of the last one. An essential shortcoming of the EL model is the neglect of the adsorbate size effect (Tom et al., 2021). However, the results of studies that use EL are usually coherent with experimental data. In this study, more components will be considered, which may have a negative effect when a simple EL model is used.

The adsorption mass transfer model defines the mass accumulation in the solid phase. The most common is to use the LDF (Linear Driving Force) for estimating the mass transfer. Kim et al. (2015) indicate that the optimal parameters of mass transfer constants should be obtained by comparing

the numerical results with experimental data. These parameters should be chosen such that the model solution is the closest to the whole range of available experimental data.

Even though PSA systems are well known, the multicomponent VOC VPSA systems are not properly investigated. The VOC VPSA modeling is becoming critical to the industry as increasingly stringent emission standards are being set. In addition, modeling in python is more and more advantageous since it provides practical optimization tools.

1.1 The VPSA process

The VPSA process can be done in different ways (with two or more adsorbers, with three or more steps, and with co-current and counter-current flow configurations). The simplified flow diagram of VOC adsorption process studied in this work is shown in Fig. 1. The process consists of two beds: one in adsorption mode and another in regeneration mode. The adsorber outlet is a pure product during the adsorption process and highly concentrated vapor during the regeneration process. This vapor is then recycled in the absorption system.

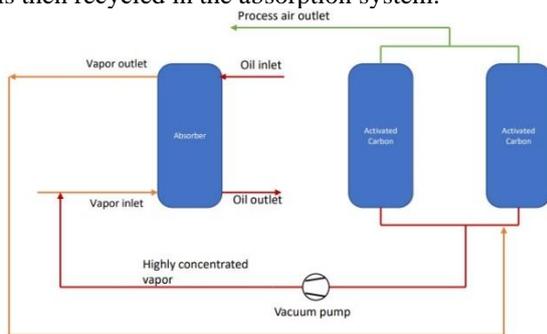


Figure 1: The simplified process flow diagram of VOC adsorption

One way of VPSA organization is co-current pressurization, adsorption, counter-current desorption, and counter-current purge. The steps are switched by employing valves. The top of the adsorber is closed during pressurization, and the system's pressure increases (see Fig. 2, upper left). The pressure increase occurs from a specific low pressure after regeneration to that required for adsorption. Next, the valve at the top of the adsorber opens, and the pure product is achieved at the outlet (see Fig. 2, upper right). After a certain period, the breakthrough of specific components occurs, and it is necessary to switch the adsorber to the regeneration mode (see Fig. 2, lower left). There is no inflow at the bottom during regeneration. Also, the valve to the vacuum pump opens, and the vacuum pump turns on. After some time, inert gas is supplied from the top while the pump operates. This step is called "purge" (see Fig.

2, lower right). As a result, the adsorbates' partial pressure decreases, which causes them to desorb better. After the purge, the whole cycle repeats. A certain amount of non-desorbed gas remains in the adsorber, which affects the process and changes the result of the next cycle. However, the system gradually converges with an increase in number of cycles to the so-called "cyclic-steady-state".

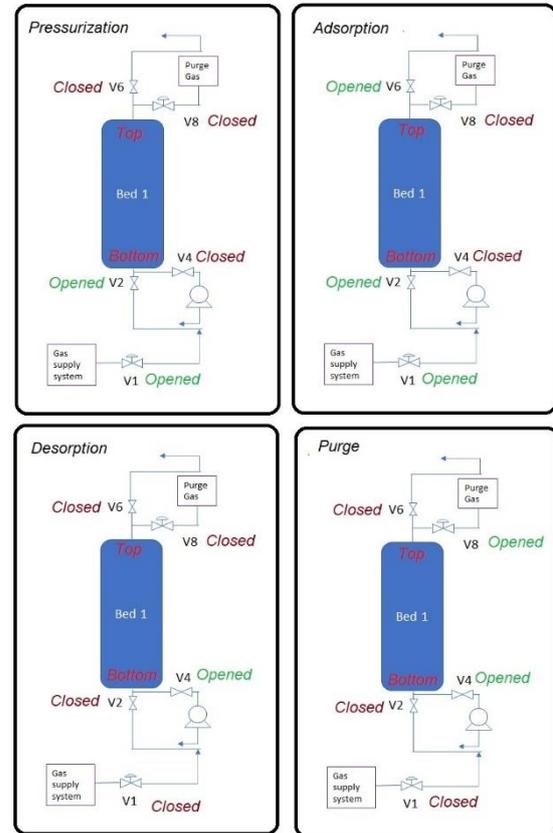


Figure 2: Cycle organization

1.2 Adsorption Equilibrium and Mass transfer

The extended Langmuir (EL) isotherm is widely used to describe the adsorption equilibrium due to the model's simplicity (see Eq. (1)).

$$\frac{w_{eq,i}}{w_{max,i}} = \frac{b_i P_i}{1 + \sum_{j=1}^N b_j P_j} \quad (1)$$

here, $w_{eq,i}$ is an equilibrium concentration of the adsorbate in the solid phase [kmol/kg], $w_{max,i}$ is the maximum adsorbate concentration in the solid phase [kmol/kg], b_i is an affinity coefficient of component i [1/bar], P_i is a partial pressure of component i [bar], N - number of components, b_j is an affinity coefficient of component j [1/bar], P_j is a partial pressure of component j [bar].

Coefficient b is the parameter that shows the ratio between adsorption and desorption rates. This parameter is called affinity coefficient and may be expressed in terms of adsorption heat, temperature, and affinity constant:

$$b_i = b_\infty \exp\left(\frac{Q}{RT}\right) \quad (2)$$

here, b_∞ is an affinity constant [l/bar], Q is an isosteric heat of adsorption [J/mol], R is a universal gas constant $R = 8.314$ [$J/molK$], T is temperature [K].

The IAST is widely used to describe multicomponent adsorption. One of the challenges in the theory is that it requires the evaluation of the spreading pressure, which is not available in the analytical form for many isotherms. However, any isotherm that can express spreading pressure analytically can be used in Fast IAST (Fast Ideal Adsorption Solution Theory), such as one-component Langmuir isotherm. The calculation procedure is described in more detail by Duong (1998).

The mass transfer resistance between gas and solid phases plays a great role in adsorption modeling. The mass transfer model defines the mass accumulation in the solid phase. The most common is to use the LDF model (see Eq. (3)) for estimating the mass transfer coefficient. The lumped mass transfer coefficient (overall mass transfer coefficient k_i in Eq. (3)) considers all the resistances to the mass transfer. They are the film around each particle, surface diffusion, macropore, and micropore diffusion in each particle.

The optimal mass transfer constants can be obtained from the comparison of the numerical results with experimental data.

$$\frac{\partial w_i}{\partial t} = \rho_{bulk} k_i (w_{eq,i} - w_i) \quad (3)$$

here, ρ_{bulk} is a solid bulk density [kg/m^3], k_i is an overall mass transfer coefficient of a component i [l/s].

1.4 Governing transport equations

The transport equations include mass, momentum, and energy balance equations. The following assumptions are used:

1. Bed porosity is homogeneous and constant along the bed.
2. Gas flows only in axial direction and there are no gradients in radial direction (one-dimensional model).
4. Diffusion in axial direction are negligible.

5. Ideal gas law is used for gas state calculations.
 6. Solid thermal conductivity is negligible and heat capacity is assumed to be constant and independent of temperature.
 7. Gas thermal conductivity is negligible and heat capacity is assumed to be constant (independent of temperature and gas composition).
 8. There is no thermal resistance between gas and solid phase.
 9. The heat of adsorption is assumed to be constant, independent of loading and temperature.
 10. Constant wall temperature assumption. Wall temperature is equal to the ambient temperature.
- The mass balance equation:

$$\frac{10^2 \epsilon}{RT} \frac{\partial (y_i P)}{\partial t} - \left(\frac{10^2}{RT}\right) \left(\frac{B}{\mu}\right) (y_i P \frac{\partial^2 P}{\partial x^2} + y_i \left(\frac{\partial P}{\partial x}\right)^2) + P \frac{\partial P}{\partial x} \frac{\partial y_i}{\partial x} + \frac{\partial w_i}{\partial t} = 0 \quad (4)$$

here, ϵ is a total porosity, R is a universal gas constant [$J/kmolK$], T is temperature [K], B is bed permeability [m^2], μ is a dynamic gas viscosity [$bar s$], P is pressure [bar], y_i is a molar fraction of component i , w_i is a solid loading [$kmol/m^3$].

The energy equation:

$$\begin{aligned} & (c_{p,solid} \rho_{bulk} + \epsilon c_{p,gas} \left(\frac{P \cdot 10^2}{RT}\right)) \frac{\partial T}{\partial t} \\ & - c_{p,gas} \left(\frac{P \cdot 10^2}{RT}\right) \left(\frac{B}{\mu}\right) \left(\frac{\partial P}{\partial x} \frac{\partial T}{\partial x}\right) \\ & = \frac{4h_W}{d} (T_{amb} - T) + \sum_{i=1}^N \frac{\partial w_i}{\partial t} \Delta H_i \end{aligned} \quad (6)$$

here, $c_{p,solid}$ is a solid specific heat capacity [J/kgK], $c_{p,gas}$ is a gas specific heat capacity [$J/kmolK$], h_W is a wall heat transfer coefficient [W/m^2K], d is a diameter of the adsorber [m], T_{amb} is an ambient temperature [K], ΔH_i is a heat of adsorption of component i [$J/kmol$].

1.5 Boundary and initial conditions

The prescribed equations are subject to boundary and initial conditions:

Table 1: Boundary and initial conditions for pressurization and adsorption steps.

Pressurization	Adsorption
$P(x, t = 0) = P_{initial}$	$P(x, t = 0) = P_{press}$
$P(x = 0, t) = P_1(t)$	$P(x = 0, t) = P_{inlet}$
$P(x = L, t) = P_2(t)$	$\frac{\partial^2 P}{\partial x^2}(x = L, t) = 0$
$y_i(x = 0, t) = y_{i,feed}$	$y_i(x = 0, t) = y_{i,feed}$
$T(x, t = 0) = T_{initial}$	$T(x, t = 0) = T_{press}$
$T(x = 0, t) = T_{inlet}$	$T(x = 0, t) = T_{inlet}$

Table 2: Boundary and initial conditions for desorption and purge steps.

Desorption	Purge
$P(x, t = 0) = P_{ads}$	
$P(x = 0, t) = P_1(t)$	$P(x, t = 0) = P_{des}$
$P(x = L, t) = P_2(t)$	$P(x = 0, t) = P_1(t)$
$\frac{\partial y_i}{\partial x}(x = L, t) = 0$	$P(x = L, t) = P_2(t)$
$T(x, t = 0) = T_{ads}$	$y_{N2}(x = L, t) = 1$
$\frac{\partial T}{\partial x}(x = L, t) = 0$	$T(x, t = 0) = T_{des}$
	$T(x = L, t) = T_{inlet}$

here, $P_{initial}$ is the initial pressure along the adsorber for the first cycle or pressure along the adsorber after purge step for the next cycles [bar], $P_1(t)$ is the pressure at the bottom (inlet) of the adsorber [bar], $P_2(t)$ is the pressure at the top (outlet) of the adsorber [bar], $y_{i,feed}$ is the molar fraction of component i in the feed stream, $T_{initial}$ is the initial temperature along the adsorber for the first cycle or temperature along the adsorber after purge step for the next cycles [K], T_{inlet} is the gas inlet temperature [K], subscripts *press*, *ads*, *des* mean the results of temperature or pressure from the previous step.

$P1(t)$ and $P2(t)$ are the boundary pressures at the bed bottom (inlet) and top (outlet) respectively. These pressures are calculated based on the ideal gas law using the following equations:

1. Pressurization

$$\frac{\partial P_1}{\partial t} = \frac{P_1}{V_1} \left(\frac{FRT \cdot 10^{-2}}{P_1} - u_{out}A \right) \quad (7)$$

$$\frac{\partial P_2}{\partial t} = \frac{P_2}{V_2} (u_{in}A) \quad (8)$$

2. Desorption

$$\frac{\partial P_1}{\partial t} = \frac{P_1}{V_1} (-q_{pump} - u_{in}A) \quad (9)$$

$$\frac{\partial P_2}{\partial t} = \frac{P_2}{V_2} (u_{out}A) \quad (10)$$

3. Purge

$$\frac{\partial P_1}{\partial t} = \frac{P_1}{V_1} (-q_{pump} - u_{in}A) \quad (11)$$

$$\frac{\partial P_2}{\partial t} = \frac{P_2}{V_2} (q_{purge} + u_{out}A) \quad (12)$$

here, F is the flow rate coming into the adsorber [kmol/s], V_1 is an inlet volume not filled with adsorbent [m^3], V_2 is an outlet volume not filled with adsorbent [m^3], u_{out} is a velocity of gas going to the adsorber [m/s], u_{in} is a velocity of gas going out of the adsorber [m/s], q_{pump} is a volumetric flow rate to the vacuum pump [m^3/s], q_{purge} is a purge volumetric flow rate to the adsorber [m^3/s]. It should be noted that velocity has a negative sign during desorption and purge steps.

2. Methodology

2.1 Python numerical solution

The set of PDEs (Partial Differential Equations) and ODEs (Ordinary Differential Equations) need to be solved with the initial and boundary conditions as shown in Tables 1 and 2.

The spatial and time derivative terms can be approximated. The variables are defined on a grid with finite difference methods.

The first-order upwind difference scheme may be beneficial in the case of sharp front propagation, which is vital in the simulation beginning and in the systems where breakthrough curves are steep. The first-order upwind method is also recommended because of the fast simulation. The scheme is first-order accurate and may give a significant numerical diffusion. However, the method does not produce oscillations (unconditionally stable). The upwind difference scheme for the positive velocity is shown below:

$$\frac{\partial f}{\partial x} = \frac{f_i - f_{i-1}}{\Delta x} \text{ if } u > 0,$$

$$\frac{\partial f}{\partial x} = \frac{f_{i+1} - f_i}{\Delta x} \text{ if } u < 0 \quad (13)$$

here, f is a function, l is a space grid point, u is velocity.

The second order derivative approximation is shown below:

$$\frac{\partial^2 f}{\partial x^2} = \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2} \quad (14)$$

The system of PDE and ODE will be solved in Python.

2.2 Aspen Adsorption numerical solution

Aspen Adsorption will be used to compare the results. Aspen Adsorption is a comprehensive flowsheet simulator for adsorption modeling. The developed flowsheet with one adsorber for VPSA process is shown in Fig. 3.

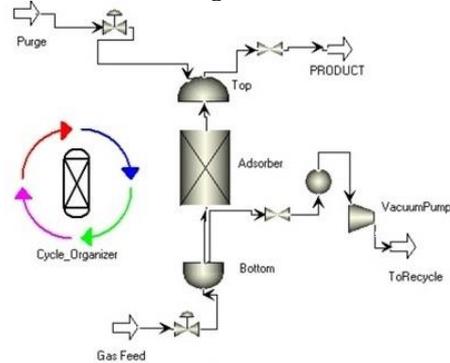


Figure 3: Aspen Adsorption VPSA flowsheet

3. Results

3.1 Model verification

3.1.1 Comparison with Cavenati et al. (2005) experiments

The simulation becomes more complicated with an increase in the number of components. Therefore, it is better first to compare the results with fewer components to investigate if the simulation works correctly. For example, Cavenati et al. (2005) considered only two components' adsorption on AC molecular sieve 3K: methane and carbon dioxide.

Cavenati et al. (2005) focused on vacuum pressure swing adsorption for the separation of methane (55 mol %) - carbon dioxide (45 mol %) mixture with a total flow between 1 and 1.5 SLPM (Standard Liter per Minute). A four-step cycle consisted of pressurization, adsorption, counter-current blowdown, and counter-current purge. The results for binary methane-carbon dioxide adsorption at a constant pressure of 320 kPa are presented. The authors did the experiments in a column of 0.83 m length and 0.021 m in diameter with a bulk density of 715 kg/m^3 . The ambient temperature during the experiments was 303 K.

Equilibrium and kinetic parameters for methane-carbon dioxide mixture adsorption on Carbon Molecular Sieve 3K can be found in Vilardi's work (Vilardi et al. 2020).

Fig. 4 shows the simulated breakthrough curves and experimental results for non-isothermal conditions. The model describes well the trend of breakthrough for methane and carbon dioxide.

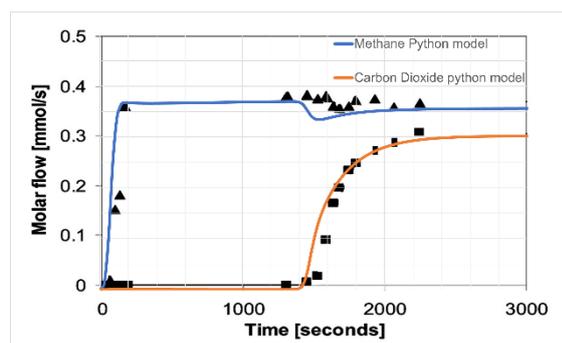


Figure 4: Comparison of breakthrough results with experiments (experimental results from Cavenati et al. (2005))

An essential part of the simulation is energy balance. Carbon dioxide adsorption produces a significant amount of heat. Cavenati et al. (2005) also reported a substantial amount of heat loss to the environment that needs to be considered. Therefore, wall heat transfer coefficient of $4\text{E-}5 \text{ MW/m}^2\text{K}$ was used as a fitting parameter, and the

result is shown in Fig. 5. The experimental data includes temperature profiles at three positions of the adsorption bed (0.17, 0.43, and 0.68 m). These temperature profiles were compared with simulated ones.

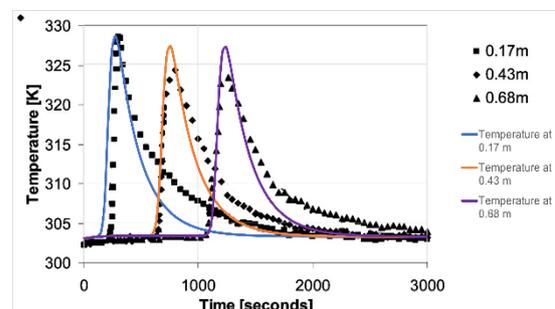


Figure 5: Comparison of temperature results with experiments (experimental results from Cavenati et al. (2005))

The results reflect the thermal wave well, but temperature profiles are more sharp than experimental. That may be due to the inaccuracy of the component properties (such as the heat capacity of gas and adsorbent) and neglect of axial thermal conductivities.

Next, Cavenati et al. (2005) presented the results after 46 cycles for VPSA system. The results of the experiments as well as the developed python simulation results are shown in Fig. 6. During the first 80 seconds of a cycle, the pressure rises from 10 kPa to 320 kPa with the constant feed flow rate of gas (see Fig. 2, upper left). The outlet flow rate of methane and carbon dioxide is 0 during this step because the top of the column is closed. During the next 100 seconds, the top of the column is opened, and the adsorption step with constant pressure and inlet flow rate begins (see Fig. 2, upper right). Only methane is obtained as the product at this step. From 180 to 300 seconds, the top of the column is closed, and counter-current blowdown occurs with the vacuum pump (see Fig. 2, lower left). The characteristic of the MZ 1C vacuum pump was used in the developed model. The final desorption pressure of 10 kPa coincides with what the authors presented in the article. In the next 50 seconds, there is a counter-current purge to decrease the partial pressure of carbon dioxide inside the adsorber (see Fig. 2, lower right).

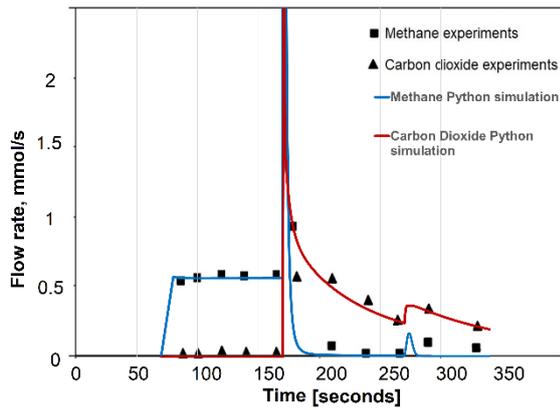


Figure 6: Comparison of temperature results with experiments (experiment results from (Cavenati et al., 2005))

3.1.2 VOC simulation and comparison with Aspen Adsorption results and Equinor experiments

The central part of the study deals with simulation of VOC adsorption on AC. The experiments were conducted at Equinor to optimize the adsorption of VOC on AC adsorption bed (Eroshkin, 2022). Adsorber design, cycle details and operating conditions are described in Tables 3-5. The adsorption cycle in the experiments was as described in Section 1.1 (see Figure 2):

Table 3: Bed parameters

Parameter	Value
Adsorber height	1 m
Adsorber height (filled with AC)	0.77 m
Adsorber diameter	0.032 m
Clean adsorbent bulk density	430 kg/m ³

Table 4: Cycle steps

Cycle step	Time
Pressurization	Until 1 bara is reached
Adsorption	900 sec/1500 sec
Desorption	600 sec / 1200 sec
Purge	300 sec

Table 5: Parameters of the main case study

Parameter	Value
Gas inlet flow rate	1248/1254 Nml/min
Purge gas flow rate	50 Nml/min
Composition	CH ₄ : 8.6 / 8 mol% C ₂ H ₆ : 16.1 / 16 mol% C ₃ H ₈ : 22.4 / 20.4 mol% C ₄ H ₁₀ : 10 / 11.7 mol% CO ₂ : 4.7 / 4.6 mol%, N ₂ : 38.2 / 39.3 mol%
Inlet temperature	303 K

Adsorbent initial temperature 303 K

Since the temperature variations do not exceed 10 degrees in the experimental results of VOC adsorption, neglecting temperature calculations should not radically change the solution. Thus, the isothermal simulation is presented below. The main parameters fitting to the model are the mass transfer coefficients. The mass transfer coefficients are initially assumed to be high, and there is almost no adsorption resistance.

The following parameters are used for equilibrium description. However, these parameters still need to be improved with proper experiments.

Table 6: Equilibrium parameters

Methane	Ethane
$w_{max}: 0.00032 \frac{kmol}{kg}$	$w_{max}: 0.0008 \frac{kmol}{kg}$
$b: 0.65 \text{ bar}^{-1}$	$b: 3.06 \text{ bar}^{-1}$
Propane	Butane
$w_{max}: 0.0018 \frac{kmol}{kg}$	$w_{max}: 0.0018 \frac{kmol}{kg}$
$b: 3.8 \text{ bar}^{-1}$	$b: 8 \text{ bar}^{-1}$
Carbon Dioxide	Nitrogen
$w_{max}: 0.00075 \frac{kmol}{kg}$	$w_{max}: 0 \frac{kmol}{kg}$
$b: 0.73 \text{ bar}^{-1}$	$b: 0 \text{ bar}^{-1}$

Fig. 7 shows the change in pressure for the first case (15 min cycle and 1248 Nml/min inlet flow rate). When the pressure reaches just above 1 bar, the adsorption process begins while the pressure remains constant. The adsorption cycle lasts 15 minutes, after which the vacuum pump is turned on, and the pressure reduces rapidly to 0.06 bar during 300 seconds of regeneration. Then pressure rises slightly after desorption due to the desorption of heavy hydrocarbons (HC) with the purge. The model developed in Aspen Adsorption (see Fig. 3) shows approximately the same pressure result as the model developed in Python.

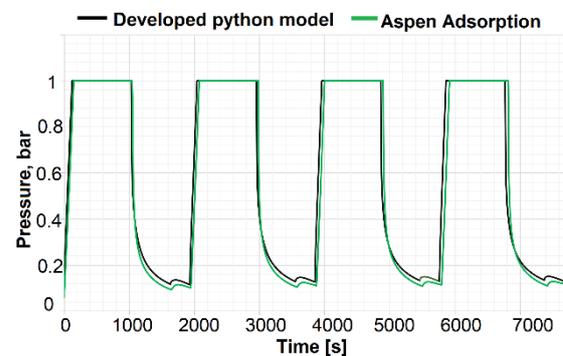


Figure 7: Pressure development during VPSA cycles

The graphs shown in Figs. 8 represent the molar flow rate of each gas exiting the adsorber after 10 simulated cycles. Cycle 1 means that all four stages were carried out only once: pressurization, adsorption, desorption, and purge. Thus, the number of cycles means how many times this 4-step cycle was simulated, using the final state from the previous cycle as the initial condition for the next cycle. Initially, the adsorber is assumed to be in equilibrium with nitrogen at 0.06 bar. However, some other adsorbed components remain adsorbed after the first cycle, affecting the next cycle. When the differences between cycles begin to be less than 1%, one can argue that the system has come to the so-called cyclic-steady-state.

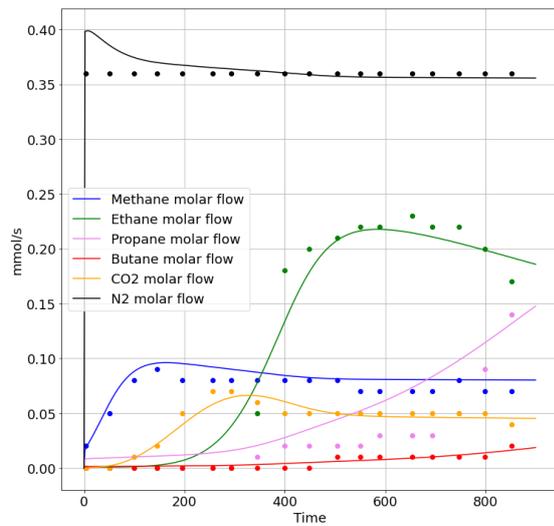


Figure 8: Comparison of the experimental and simulated results of VOC adsorption on AC with 15 min cycle and inlet flow rate 1248 Nml/min

Figures 9, and 10 show the capture efficiency when the cyclic-steady-state is reached for the cases with 15 min cycle and 25 min cycle. Capture efficiency of 100 % means no flow rate of a component exiting the column. Capture efficiency of 0 % means that the outlet molar flow rate of a component is equal to the inlet molar flow rate. The negative efficiency means that the outlet flow rate is higher than the inlet. As one can see from the model the capture of various VOC happens differently. The competitive behavior can be noticed. Once an equilibrium zone is formed in the AC, components with stronger affinity will push components with less affinity out of the adsorbent. Consequently, the outlet flow of “weak” components will increase above the inlet flow. It is seen from the graph that heavier components bind to the AC stronger than lighter components.

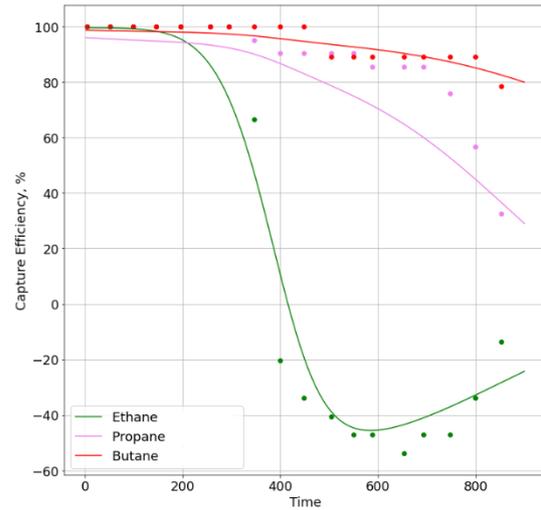


Figure 9: Capture efficiency. Molar flow 1248 Nml/min and cycle time 15 min

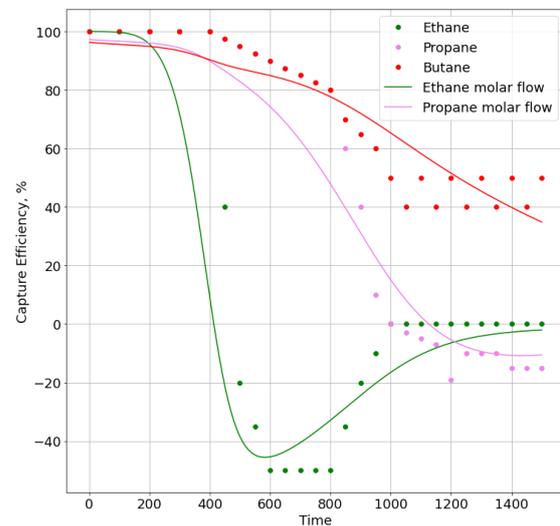


Figure 10: Capture efficiency. Molar flow 1254 Nml/min and cycle time 25 min

4. Summary and Discussion

The developed python model for VPSA simulation showed good agreement with Aspen Adsorption simulations, Cavenati et al. (2005) experiments and some multicomponent VOC experiments. It can be confirmed that the model calculates correctly; however, the equilibrium and kinetics parameters for the six-component mixture need to be improved. The available experimental data on equilibrium and kinetics is limited, and more accurate model predictions can be achieved by conducting more equilibrium and kinetic experiments.

5. Conclusion

The proper development of the adsorption process with high performance involves the design of a model that can describe the dynamic of adsorption, considering all relevant transport phenomena. The developed model in python can be used for prediction and theoretical description of competitive VOC adsorption. The multicomponent adsorption isotherm and LDF equation for mass transfer reasonably well predict the adsorption behavior of 2 component VOC mixture (methane and carbon dioxide). When it comes to the six-component mixture, some adjustment required to get a good comparison with experimental data.

References

- Ahn, S. et al. (2013) 'Layered two- and four bed PSA processes for H₂ recovery from coal gas', *Chemical Engineering Science*, pp. 413–423. doi:10.1016/j.ces.2011.09.053
- Cavenati, S. et al. (2005) 'Upgrade of methane from Landfill gas by pressure swing adsorption', *Energy and Fuels*, pp. 5108–5117.
- Duong, D. (1998) 'Adsorption Analysis: Equilibria and Kinetics'. *Imperial College Press*. Vol. 2, 11
- Eroshkin, S (2022) 'Volatile Organic Components Adsorption on Activated Carbon'. Master's Thesis. Norwegian University of Science and Technology.
- Esteves, I. et al. (2003). 'Adsorption of Natural Gas Components on Activated Carbon for Gas Storage Applications'. C-H. Lee (Ed.), *Proceedings of the Third Pacific Basin Conference on Adsorption Science and Technology* (pp. 479-483). World Scientific Publishing Co. Pte Ltd. doi:10.1142/9789812704320_0086
- Kim, Y.J. et al. (2015). 'Study on a numerical model and PSA (pressure swing adsorption) process experiment for CH₄/CO₂ separation from biogas', *Energy (2015)*, pp. 732–741. doi:10.1016/j.energy.2015.08.086.
- P.Cruz et al. (2003) 'Cyclic adsorption separation processes: analysis strategy and optimization procedure', *Chemical Engineering Science* 58, pp. 3143–315. doi:10.1016/S0009-2509(03)00189-1
- Shuangjun, L. et al. (2018). 'Mathematical modeling and numerical investigation of carbon capture by adsorption: Literature review and case study'. *Applied Energy* 221, pp. 437–449.
- Tefera, D.T. et al. (2014) 'Modeling competitive adsorption of mixtures of volatile organic compounds in a fixed-bed of beaded activated carbon', *Environ Sci Technol.* 48(9), pp. 5108-17. doi: 10.1021/es404667f. Epub 2014 Apr 8. PMID: 24670053.
- Tom, R. C. et al. (2021) 'Adsorption Size Effects for Langmuir Systems in Process Simulators: Case Study Comparing Explicit Langmuir-Based Models and FASTIAS', *Industrial & Engineering Chemistry Research* 60 (32), pp. 12092-12099. doi: 10.1021/acs.iecr.1c01657.
- Vacuum pump VacuuBrand MZ1C. url: <https://shop.vacuubrand.com/en/chemistry-diaphragm-pump-mz-1c-20724100.html#description>
- Vilardi, G. et al. (2020). 'Exergy and energy analysis of biogas upgrading by pressure swing adsorption: Dynamic analysis of the process', *Energy Conversion and Management* 226. doi: 10.1016/j.enconman.2020.113482.

Integrating energy in the conceptual design stage to optimize building form

Niloozar Zaker Vafae^{a,*}, Melika Sandani^a, Tahmineh Akbarinejad Khameneh^a,
Matteo Tagnocchetti^a, Bunji Izumi^a, Luca Finocchiaro^a

^a *Department of Architecture and Technology, Norwegian University of Science and Technology*
Niloozafaz@stud.ntnu, melikas@stud.ntnu.no, Tahmineh.akbarinejad@ntnu.no, matteo.tagnocchetti@ntnu.no,
bunji.izumi@ntnu.no, luca.finocchiaro@ntnu.no

Abstract

In architecture, Integrated Energy Design (IED) entails considering energy during each design phase, especially in the early design stage. The form of a building is an important factor in this stage due to its considerable impact on energy consumption. Finding the optimal form is a time-consuming process, and computational design techniques can help designers to facilitate this process and achieve a design solution with acceptable performance in terms of CO₂ emission. Moreover, the surrounding buildings, trees and urban elements can affect the energy and daylight of the project by casting shadows. Considering all these elements throughout the design process can be very demanding and take several working days. Today, digital tools make it possible to parametrically analyze morphological characteristics of buildings to identify the most efficient solution. The present study proposes an environmental-simulation based design workflow to be used in the early design stage to determine the building's form parameters (height, angle,...) in a given urban area based on the weather data and the surrounding context. This process is done by parametric design tools and environmental simulations in Rhino3D[®], Grasshopper[®], and ladybug Tools[®]. The typical Norwegian cabin's form parameters are applied in the visual coding program (Grasshopper[®]) to generate the initial geometry for optimization. Due to the great effect of the energy consumption on the CO₂ emission, minimizing energy, maximizing thermal comfort and the sky view percentage were the main objectives. To test the workflow the weather data of Tromsø (Norway) and 3d model of the surrounding context of a design location was applied as inputs. The output of this application was several building's form alternatives for that specific location. This study showed using the digital tools and parametric design thinking can help the designers to apply the climatic data in the design process to narrow down the design solutions.

Keywords: parametric-design, optimization, morphology, Integrated-Energy-Design

1 Introduction

The building industry is responsible for 40% of the carbon emission (Huang et al., 2018). The decisions made in the early design stage are having a great effect in terms of CO₂ emission. In this stage, many design alternatives are generated, and their performances are evaluated (Miles et al., 2001). Among the building factors the building's form can have an considerable impact on energy consumption. A study showed that by changing design parameters such as the roof slope and skylight length and width, daylighting performance can be increased to nearly 40 percent, and energy demand reduced by around 20 percent (Miles et al., 2001). Zou et al. (2021) showed that the average performance of the building could increase by up to 24% by optimizing the design variables, including wall length and glazing ratio with the objective of having minimum air conditioning and lighting

energy and maximizing the average Useful Daylight Illuminance (UDI). Zhang et al. (2017) investigated the effect of the building shape parameters such as window-to-wall ratio, room depth, orientation, and shading type on energy consumption and thermal discomfort. The optimized solution could perform better near 13% and 4% in energy and thermal comfort, respectively. Konis et al. (2016) introduced a new method to improve the performance of the passive strategies in conceptual design and investigated the building morphology optimization to achieve desired daylighting and to minimize energy consumption in four different climates. The result showed that the performance of the building improved, especially in warmer climates. It also revealed a huge impact of shape optimization on daylight improvement, between 24% to 65% depending on the local context and the climate. Harkouss et al. (2018) presented a

comprehensive study on optimal passive design. In this study the optimal solution showed the potential of saving up to 54% of cooling and 87% of heating demands compared to the initial values. The present study aims to introduce a workflow for an investigation of building morphology with geometry generation parameters as variables and optimization using an evolutionary multi-objective optimization, so-called Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002), regarding the energy, occupants' thermal indoor comfort, and sky view percentage.

2 Methodology

The present study is addressed to designers and engineers who need to evaluate the environmental performance of their design in the early design stage but they do not have computational design knowledge to write the analysis by digital tools. In this study Grasshopper® (Visual programming plug-in in Rhino3D®) is used as the parametric design tool. The defined workflow is applied to the design of small cabin with the sloped roof but the process can be applied more widely. The order of used applications and the process are shown in Fig. 2. Procedure is explained in detail in 2.1.



Figure 1: left) a typical Norwegian cabin outline (svenskoedegaard, 2020), and right) the extracted geometry showing different parameters.

2.1 Simulation Procedure

The first step in the workflow was generating an initial 3D geometry as the building's form in Grasshopper®. Inspired by a typical Norwegian cabin form with a polygon as the plan and the sloped roof (Fig. 1), the main framework of the geometry was defined, and the geometry parameters such as plan shape, the height of the wall, wall angle, and roof slope were set as the variables in Grasshopper® (Table. 2) for which numeric values are assigned. Changing each variable's numeric value could generate a new form with different height and angles. Next step is to evaluate the performance of the form in Energy consumption, occupants' thermal comfort, and the sky view by the Ladybug Tools® and EnergyPlus® plug-in in Grasshopper®. The Energy analysis is conducted by running the Energy component in EnergyPlus® inside Grasshopper®. After that the indoor thermal comfort will be calculated by the same plug in. Finally, the sky view

will be analyzed by Ladybug Tools® in Grasshopper®. The optimization is conducted by Wallacei X® plug-in in Grasshopper®. This workflow could generate thousands of form possibility and calculate their environmental performance and compare each form to the others regarding the result of the performances. Therefore, it can find the numeric values for geometry parameters that generate the optimal form, with having an average of good performance in all the objectives (Energy, comfort, sky view). The inputs of this workflow are weather data of the desired location and the context geometry as a 3D model in Rhino3D®, and the output is several form alternatives with minimized energy consumption, maximized occupants' indoor comfort, and sky view among all the form possibilities (design solutions).

2.2 Setting the geometry parameters

The initial geometry framework was inspired by the typical Norwegian cabin having a polygon as the plan and a sloped roof (Fig. 1). The building geometry parameters was decoded and translated into visual coding (Grasshopper®) to generate the initial geometry. Considering the initial construction form, a polygon was generated as the base surface. Since this study aimed to use a parametric design approach, the location of the corners of that polygon could be defined in a parametric way. Therefore, a circle was chosen as the base of the polygon creation. Then by connecting the subdivision points on the circle the plan polygon is created. After that, getting the center point of the polygon will produce the sloped roof geometry. By connecting each corner to the center point and moving the center point, the degree of the roof is changed (Fig. 3). In the morphology investigation process, the geometry parameters generating the shape of the building were considered as the variables (Fig. 3). To achieve the appropriate percentage of the glazing, the area of the walls was multiplied by a number between 0.1 to 0.8, resulting in the window-to-wall ratio. Then the location of the glazing is made by a random function. Random selection of the exterior wall surfaces produced the location of the glazing in the initial solid geometry. Although this is a random function, the possibilities of window locations can be modified.

2.3 Setting the Objectives

2.3.1 Total Energy Consumption

In this study, energy demand minimization was the first objective calculated using the EnergyPlus® engine (Delgarm et al., 2016) in Ladybug Tools®. The output of this analysis was heating and cooling energy load. The addition of these values is used in this study to calculate the total energy demand. The Energy and comfort simulations are done with the opaque material of the EnergyPlus® database, with a

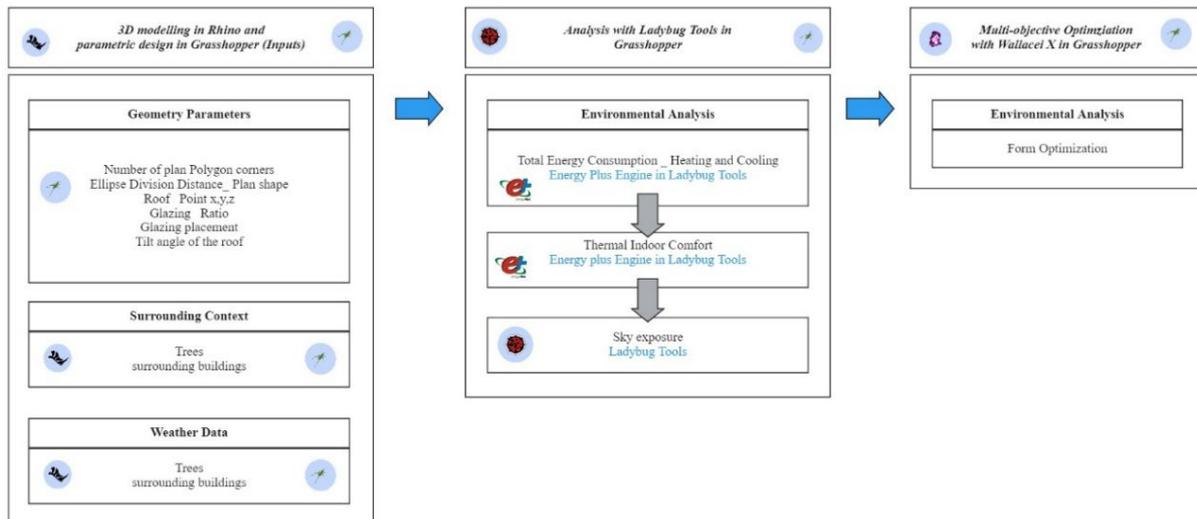


Figure 2: Diagram showing the workflow process and the order of used applications.

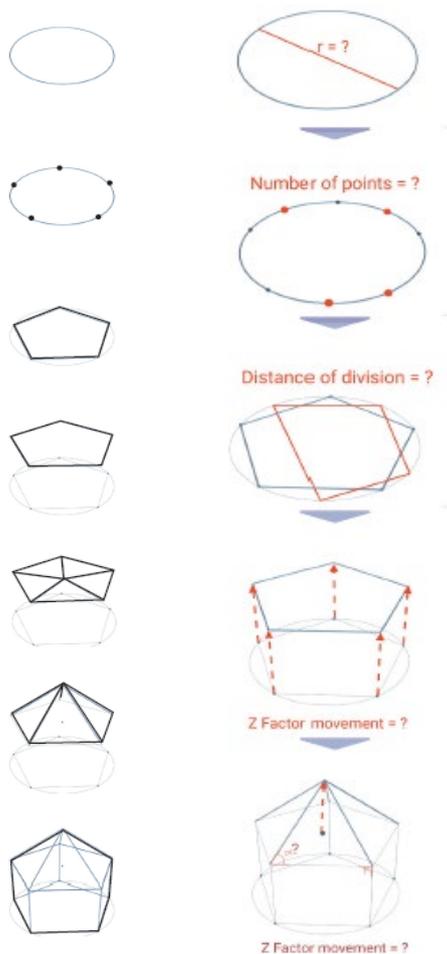


Figure 3: left) generating the simple geometry of typical Norwegian house with parametric design tools, and right) showing the geometry parameters as the variables (in red) to generate different forms.

U-value of $0.13 \text{ W/m}^2\text{K}$, which is the heat transfer rate through the opaque material. The glazing U value is set as $0.7 \text{ W/m}^2\text{K}$ for the initial analysis. In the current study material was not the in focus of the study, therefore the simulations are considered without insulation. The general comparison of the form's performances has been under investigation. The numeric result of the simulations were higher than standards expectedly.

2.3.2 Occupants' Indoor thermal comfort

Thermal comfort is a state of mind in which a person expresses full satisfaction with their thermal surroundings (Moser et al., 2001). Design variables under the architect's control can affect the indoor environment. These design variables are general layout, shape, location of windows, and insulation. Comfort is evaluated by predicted mean vote (PMV), a well-known example of a thermal comfort performance indicator which was developed by Povl Ole Fanger as an empirical fit to the human sensation of thermal comfort. It was later adopted as an ISO standard. PMV is a seven-point sensation scale from -3 to +3. According to the ASHRAE standards, this should be kept at 0 with a tolerance of ± 0.5 to ensure a comfortable indoor environment (Srebric et al., 2015). The PMV Comfort in Ladybug Tools® has three main inputs, including dry bulb temperature, mean radiant temperature of the surrounding surfaces in degrees Celcius, and metabolic rate of the human and the output will be predicted mean vote showing the degree of the occupants' comfort. Wallacei X® Plug-in can only minimize the objectives' values; Therefore, the result of the comfort calculation was changed into the formula below and then connected to the Wallacei X® to be minimized.

$$N = 3 - |X| \quad (1)$$

where, N is an objective to be minimized, and X is the result of the comfort simulation from Ladybug Tools®. The closer the result of the comfort simulation be to zero, the better comfort is provided. The result of formula (1) should be close to three. (Tab. 1) gives the assumed ranges for nine main variables.

2.3.3 Sky View

If the energy was the only objective of the workflow the result would be a form with no window to save the energy consumption due to less heat loss. Another objective is needed to balance the glazing ratio. Therefore, sky view percentage from inside of the form is chosen as an objective.

Since Wallacei X® can only minimize the objectives, to maximize the objective it was multiplied by minus one (-1) and then connected to input of Wallacei X®.

2.3.4 Area

Optimization results showed disregarding the floor area can lead to solutions with smallest floor area and high energy efficiency. Therefore, floor area was also added as an objective to drive the optimization in favor of the floor area as close as 35 m². The objective function appears as formula(2).

$$A = |35 - (\text{Floor area of the case})| \quad (2)$$

2.4 Optimization

In this project, Wallacei X® version 2.7 is used as the optimization tool using an evolutionary algorithm (NSGA-II) suitable for multi-objective optimization with four objectives and nine geometry variables. The evolutionary algorithm is the genetic algorithm that uses the natural selection principles to evolve a set of solutions towards an optimum solution (Machairas et al., 2014). Wallacei X® is the key built-in and integrated multi-objective optimization algorithm widely employed in many studies (Wang et al., 2021). This tool tests each numeric value for each variable, test the results with objective functions, compares the results and goes to another set of variables which generate another solution (form). For the first generation (iteration) it conducts random numeric values to evaluate the resulted forms. Then using the analyzed data it produces another set of geometries in the next generation (iteration) and compared them with the previous generation (iteration). The goal is to minimize the numeric value of the result. So, it is expected after multiple generation the last one be containing the optimized solutions (Deb et al., 2002). The overview of optimization process is shown in Fig. 4.

2.5 Case Study (Tromsø)

To test the developed workflow in Grasshopper®, a real location in Tromsø was chosen because of its severe cold climate and critical energy demand. The surrounding of a residential area in Tromsø in Petterburggate is modeled in Rhino3D® as the cabin location (Fig. 5(Right)).

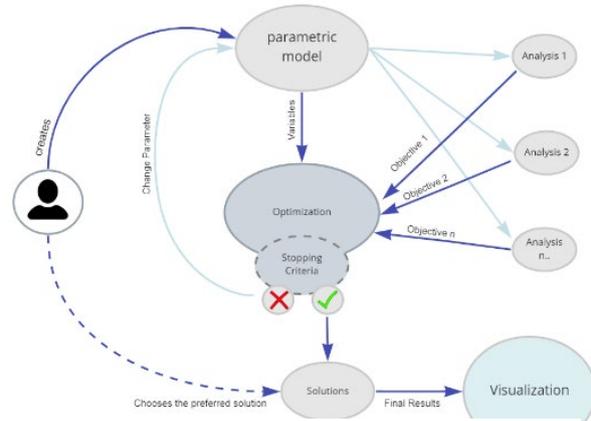


Figure 4: Diagram showing the workflow process.

2.5.1 Tromsø climate

Tromsø lies in Northern Norway (69.6492° N, 18.9553° E). The temperature typically varies from -6 to 15 °C and is rarely below -13 °C or above 21°C. The coldest month in Tromsø is January, with an average low of -6 °C and a high of -1 °C. Fig. 5(right) shows the total radiation rose in Tromsø, having the most radiation on the south side up to near 770 kWh/m².

2.5.2 Applying Tromsø Weather Data

The weather data¹ of Tromsø was the climate input for the environmental simulation. The surrounding area was modelled in Rhino3D® and connected to the context input of the Ladybug Tools® plug-in.

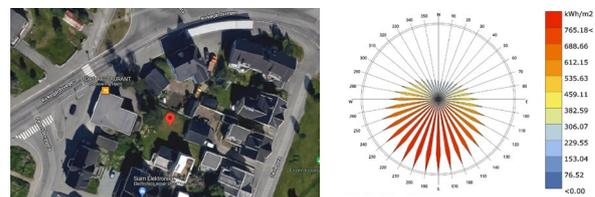


Figure 5: Left) Tromsø location in "Petterburggate," with the surrounding Right) Radiation Rose of Tromsø extracted from Ladybug Tools® weather data analysis.

The optimization is operated with the default setting of Wallacei X®, having 50 as the number of

¹ Extracted from <https://climate.onebuilding.org/>

generations to iterate and a maximum population of 30 forms for each iteration.

Table 1: Variables' assumed ranges

<i>Variables</i>	<i>Range</i>
Number of plan Polygon corners	3 < <10
Ellipse Division Distance_ Plan shape	0.0 < <1.0
Roof Point x Coordinate	-3.0 < <+3.0
Roof Point y Coordinate	-3.0 < <+3.0
Roof Point z Coordinate	-3.0 < <+3.0
Glazing Ratio	0.1 < < 0.7
Glazing Plane location (Randomness between choosing which surface to be glazing)	0.0 < <1.0
Scale factor of the roof surface (responsible for the wall angle)	0.0 < <1.0
Tilt angle of the roof	-25 < <+25

Table 2: Variables' assumed ranges and chosen values for sensitivity analysis.

<i>Variables</i>	<i>Range</i>
Number of plan polygon corners	3 < <10
Ellipse division distance plan shape	0.0 < <1.0
Roof point X coordinate	-3.0, -1.5, 0, +1.5, +3.0
Roof point Y coordinate	-3.0, -1.5, 0, +1.5, +3.0
Roof point Z coordinate	-3.0, -1.5, 0, +1.5, +3.0
Glazing Ratio (customized based on location)	0.1, 0.2, 0.3, 0.4
Glazing plane location (randomness between choosing which surface to be glazing)	0.2, 0.3, 0.4, 0.5
Scale factor of the roof surface (responsible for the wall angle)	1.0, 1.3, 1.5, 1.7, 2.0
Tilt angle of the roof	-25, -12.5, 0, +12.5, +25

Algorithm parameters for optimization were as follows: 0.9 for crossover probability, 20 for crossover distribution index, 20 as mutation distribution index, 1 as random seed, and the mutation probability is $1/n$. The result was 30 individual forms in Generation (iteration) 49, 12

forms of the 3d models are shown in Fig. 7 in the plan view.

Due to the time-consuming nature of the optimization and simulation in this study, energy and comfort simulations were done for the most critical month of the year in Tromsø which is January, with the highest demand for heating. Regarding the nine main variables (Tab. 1) and to narrow down the number of simulations, the limited number of numeric values as variables were chosen among all the possibilities so that the number of generated 3D models was reduced from millions of shapes to a few thousands of models to do the sensitivity analysis which narrows down the number of possibilities to see the most affective range of the numeric values of the variables (Tab. 2).

3 Results

The generated solutions were exported from Wallacei X[®] as shown 3D in Fig. 6 and 2D in Fig. 7. The optimization process showed that with a ready to use workflow developed by parametric design tools, achieving an efficient form based on the climatic data will take less than one working day.

The plan shape of the pareto front solutions Fig. 7 shows that according to the weather data in January and surrounding buildings in that context, the southeast has a wider side to receive more radiation, while in the north, the walls are more compact. The roof is tilted toward the sun in the south, and in most cases of the pareto front solutions, the skylight in the north side. As Fig. 6 shows, the wall angles are slightly leaning towards the outside, helping them to receive better solar radiation since the sun's altitude is low in that location. Observations showed that no single solution is doing the best in all the objective values. The solution that performs the best in Energy consumption is expected to be the worst in sky view since the glass ratio is low to avoid heat loss.

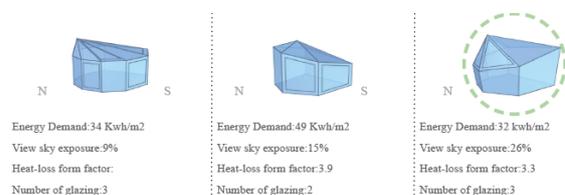


Figure 6: showing three different alternatives with their objective values as the outputs from Wallacei X[®]. The highlighted form is showing more efficiency.

5 Discussion

Due to the environmental simulations' complexity and the time-consuming optimization process, not every designer and engineer know the environmental simulation techniques to apply to the project's location.

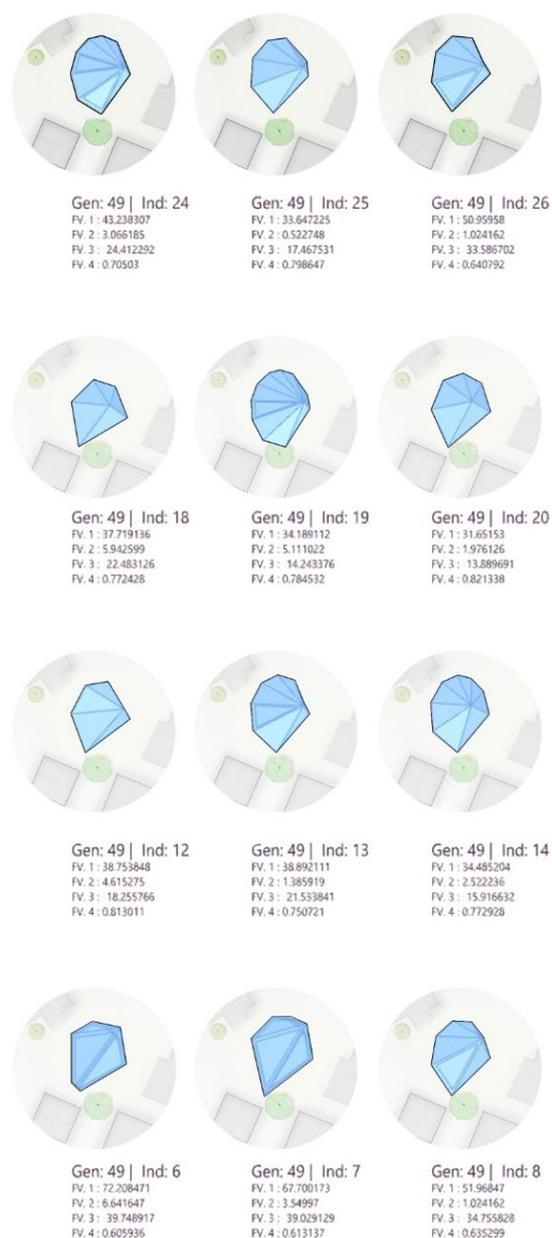


Figure 7: 20 Pareto front Solutions top view in generation (iteration) 49 exported from *Wallacei X*[®]
 Fv.1 : Energy Demand , Fv.2 : A, Fv.3 : percentage of View, Fv.3 : N gen= generation, Ind = Individual.

Therefore, an integrated workflow is needed to facilitate the decision-making in the design space regarding the performance of the building in environmental simulations to enhance the use of the climatic data and energy efficiency in the design process. Previous studies mentioned earlier have investigated the possibility of optimization for the constructed shape. In this study the generation of a new form is studied to prepare a workflow regarding the energy efficiency. This workflow shows the power of using the digital tools to narrow down design solutions.

6 Conclusion and further study

The introduced workflow used parametric design tools such as Rhino3D[®] and Grasshopper[®], and Ladybug Tools[®] for environmental simulation to investigate the morphology of buildings and introduce a form generation method inspired by typical Norwegian cabin geometry. The geometry parameters were considered as numeric variables, and the objectives for optimization were total energy demand, occupants' thermal indoor comfort, sky view percentage, and area. This workflow helps the designers to narrow down the design solutions and make better decisions based on the environmental performance of the. Still, the designer's role is to choose among the generated form solutions manually. This innovative approach introduced the potential of generating geometry and its modification based on the climatic data.

This study considered energy, occupants' comfort, the sky view percentage, and the area as the objectives for optimization, and it is suggested that further work be conducted to include daylight availability, wind analysis for natural ventilation, and cost of the elements. Further studies can be done to find another geometry generation process, such as a complex geometry for the form of the building. Interesting study as further work can be suggested by allocating material and analyzing the different material options and their effect on each objective value. This workflow has the potential to be used as a plug-in for Grasshopper[®], which generates a set of optimum design solutions for given weather data. Comparing the optimized shape for different locations also can be another interesting topic in future studies.

Acknowledgment

Thanks to Marcin Luczkowski for his help in the form-finding process, and Steinar Hillersøy Dyvik for helping with the investigation of digital tools.

References

- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197. <https://doi.org/10.1109/4235.996017>
- Delgarm, N., Sajadi, B., Kowsary, F., & Delgarm, S. (2016). Multi-objective optimization of the building energy performance: A simulation-based approach by means of particle swarm optimization (PSO). *Applied Energy*, 170, 293–303. <https://doi.org/10.1016/j.apenergy.2016.02.141>
- Harkouss, F., Fardoun, F., & Biwole, P. H. (2018). Passive design optimization of low energy buildings in different climates. *Energy*, 165, 591–613. <https://doi.org/10.1016/j.energy.2018.09.019>
- Huang, L., Krigsvoll, G., Johansen, F., Liu, Y., & Zhang, X. (2018). Carbon emission of global construction sector.

Renewable and Sustainable Energy Reviews, 81, 1906–1916.
<https://doi.org/10.1016/j.rser.2017.06.001>

Konis, K., Gamas, A., & Kensek, K. (2016). Passive performance and building form: An optimization framework for early-stage design support. *Solar Energy*, 125, 161–179.
<https://doi.org/10.1016/j.solener.2015.12.020>

Machairas, V., Tsangrassoulis, A., & Axarli, K. (2014). Algorithms for optimization of building design: A review. *Renewable and Sustainable Energy Reviews*, 31, 101–112.
<https://doi.org/10.1016/j.rser.2013.11.036>

Miles, J. C., Sisk, G. M., & Moore, C. J. (2001). The conceptual design of commercial buildings using a genetic algorithm. *Computers & Structures*, 79(17), 1583–1592.
[https://doi.org/10.1016/S0045-7949\(01\)00040-2](https://doi.org/10.1016/S0045-7949(01)00040-2)

Moser, A., Schäulin, A., Davidson, L., Corrado, V., Dorer, V., & Koschenz, M. (2001). 11—DESIGN WITH MODELING TECHNIQUES. In H. Goodfellow & E. Tähti (Eds.), *Industrial Ventilation Design Guidebook* (pp. 1025-p3). Academic Press.
<https://doi.org/10.1016/B978-012289676-7/50014-X>

Rizzo, G., Beccali, M., & Nucara, A. (2004). Thermal Comfort. In C. J. Cleveland (Ed.), *Encyclopedia of Energy* (pp. 55–64). Elsevier. <https://doi.org/10.1016/B0-12-176480-X/00551-9>

Santamouris, M. (2019). Chapter 3—Urban Heat Island and Local Climate Change. In M. Santamouris (Ed.), *Minimizing Energy Consumption, Energy Poverty and Global and Local Climate Change in the Built Environment: Innovating to Zero* (pp. 65–102). Elsevier. <https://doi.org/10.1016/B978-0-12-811417-9.00003-9>

Srebric, J., Heidarinejad, M., & Liu, J. (2015). Building neighborhood emerging properties and their impacts on multi-scale modeling of building energy and airflows. *Building and Environment*, 91, 246–262.
<https://doi.org/10.1016/j.buildenv.2015.02.031>

Wang, W., Liu, K., Zhang, M., Shen, Y., Jing, R., & Xu, X. (2021). From simulation to data-driven approach: A framework of integrating urban morphology to low-energy urban design. *Renewable Energy*, 179, 2016–2035.
<https://doi.org/10.1016/j.renene.2021.08.024>

Zhang, A., Bokel, R., Van den Dobbelaars, A., Sun, Y., Huang, Q., & Zhang, Q. (2017). The Effect of Geometry Parameters on Energy and Thermal Performance of School Buildings in Cold Climates of China. *Sustainability*, 9(10), 1708.
<https://doi.org/10.3390/su9101708>

Zou, Y., Zhan, Q., & Xiang, K. (2021). A comprehensive method for optimizing the design of a regular architectural space to improve building performance. *Energy Reports*, 7, 981–996.
<https://doi.org/10.1016/j.egy.2021.01.097>

A fast and effective method for modelling and optimizing district heating systems in the Modelica language

Haoran Li ^{a,*}, Juan Hou ^a, Natasa Nord ^a

^a *Department of Energy and Process Technology, Norwegian University of Science and Technology, Kolbjørn Hejes vei 1 B, Trondheim 7491, Norway*
corresponding. haoranli@ntnu.no

Abstract

A district heating system is a centralized energy system that supplies heat to end users such as buildings and industrial facilities. This centralized system may have multiple heat sources, a complex distribution network, and a large number of end users. Moreover, the heat distribution and utilization processes entail tumultuous thermal dynamics. Therefore, modelling and optimizing such a system generally demands arduous labour and necessitates powerful computing resources. To overcome these difficulties, this study introduced a fast and effective method for modelling and optimizing district heating systems using the Modelica language. Firstly, a simplified district heating system model was developed. This simplified model lumped all the end-users into a single thermal pinot with critical physical constraints. Meanwhile, the distribution network was simplified into two pipelines: supply and return. In addition, a one-dimensional discrete model was used to describe the behaviours of water tank thermal energy storage. Other essential components, like central and distributed heat sources, were modelled using basic mass and energy balance equations. Afterwards, two optimization frameworks were formulated, which incorporated the developed system model: a long-term optimal operation framework aimed at a yearly level open-loop optimization with a two-hour resolution, and a model predictive control framework aimed at a daily level close-loop optimization with a one-hour resolution. The proposed method was tested numerically on a university campus district heating system in Norway on a personal computer. Model validation showed that the proposed modelling approach could capture the key characteristics of the studied system. Optimization results demonstrated the effectiveness of the proposed optimization frameworks both for the long-term and short-term optimization.

1. Introduction

Buildings consume a considerable amount of energy and contribute significantly to global warming. In the European Union (EU), buildings account for roughly 40% of overall energy consumption and 36% of greenhouse gas emissions (*In focus: Energy efficiency in buildings*). As important parts of building energy systems, space heating (SH) and domestic hot water (DHW) systems play a crucial role in the energy usage of buildings. For example, in the EU's residential sector, SH and DHW account for over 80% of energy use (*Heating and cooling- European Commission*). District heating (DH) systems are a cost-effective and environmentally responsible approach to meeting buildings' heat demand (Li & Nord, 2018). Because of these advantages, DH systems are competitive with alternative heating methods, particularly in urban areas with high heat demand. In Europe, around 4,000 DH systems are currently operational (Sayegh et al., 2017), with DH systems accounting for up to 60% of the national heat market share in some countries

(Åberg et al., 2020; Connolly et al., 2014; Werner, 2017). Despite these advantages of DH systems, no viable open-source platform focusing on the optimal design of DH systems and their control systems exists.

Modelica is a promising open-source language for modelling energy systems with numerous libraries. Recently, the Modelica language is further promoted by several large-scale international projects. Among these projects, IBPSA Project 1 has built the basis of the next generation computing tools for district energy and control systems (*IBPSA Project 1*, Jan 2021). However, the current Modelica based platforms show inadequate performance on computing DH systems and their control systems, especially for the cases with large scale and complex systems. Furthermore, modelling such a system is typically time-consuming, which adds to the difficulty of modelling and optimizing DH systems. To overcome these challenges, this article proposes a fast and effective method for modelling and

optimizing DH systems using the Modelica language. This article reports the research outcomes from several recent publications (Hou et al., 2019; Hou, Li, & Nord, 2022; Hou, Li, Nord, et al., 2022; Li et al., 2021; Li, Hou, Hong, et al., 2022; Li, Hou, Tian, et al., 2022; Li & Nord, 2019).

2. Modelling DH systems in Modelica

As illustrated in Figure 1, the proposed DH system model included a building, distribution pipeline, water tank thermal energy storage (WTES), main substation, and distributed heat source (DHS) component. Due to the object-oriented nature of the Modelica language, it is possible to model a complete DH system by integrating these components. The detailed modelling work for these components is presented in Sections 2.1 to 2.4.

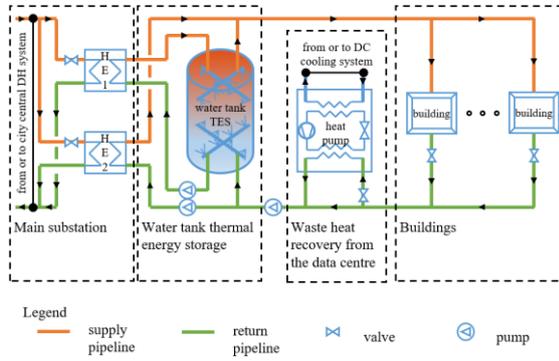


Figure 1: Configuration of the studied DH system.

2.1. Main substation and distributed heat source model

The energy balance equation was used to model the main substation and distributed heat source, as shown in Equations (1) to (4). The main substation has two heat exchangers, heat exchanger 1 and heat exchanger 2, working as charging and boosting heat sources, respectively. The distributed heat source may be solar thermal plants and waste heat recovery facilities.

$$\dot{Q}(t) = \dot{Q}_{HE1} + \dot{Q}_{HE2} \quad (1)$$

$$\dot{Q}_{HE1} = c \cdot \dot{m}_{HE1} \cdot (T_{HE1,sup} - T_{HE1,ret}) \quad (2)$$

$$\dot{Q}_{HE2} = c \cdot \dot{m}_{HE2} \cdot (T_{HE2,sup} - T_{HE2,ret}) \quad (3)$$

$$\dot{Q}_{DHS} = c \cdot \dot{m}_{DHS} \cdot (T_{DHS,sup} - T_{DHS,ret}) \quad (4)$$

where \dot{m}_{HE1} , \dot{m}_{HE2} , and \dot{m}_{DHS} are the water flow rate of the heat exchanger 1, heat exchanger 2, and distributed heat source, respectively. \dot{Q}_{HE1} , \dot{Q}_{HE2} , and \dot{Q}_{DHS} are the heat flow rate of the heat exchanger 1, heat exchanger 2, and distributed heat source, respectively. $T_{HE1,sup}$, $T_{HE2,sup}$, and

$T_{DHS,sup}$ are the supply water temperature of the heat exchanger 1, heat exchanger 2, and distributed heat source, respectively. $T_{HE1,ret}$, $T_{HE2,ret}$, and $T_{DHS,ret}$ are the return water temperature of the heat exchanger 1, heat exchanger 2, and distributed heat source, respectively. c is the specific heat capacity of water.

2.2. Buildings model

The overall performance of all the buildings in a DH system was represented by a single-equivalent building model to improve computing efficiency. Following this simplification, the thermal behaviour of all the buildings was described using Equation (5). Inequality constraints for the water temperature and flow rate variables were given by Equations (6), (7), and (8).

$$\dot{Q}_{Bui} = c \cdot \dot{m}_{Bui} \cdot (T_{sup} - T_{ret}) \quad (5)$$

$$\Delta T_{Bui,L} \leq \Delta T_{Bui} = T_{sup} - T_{ret} \leq \Delta T_{Bui,U} \quad (6)$$

$$T_{sup,L} \leq T_{sup} \leq T_{sup,U} \quad (7)$$

$$\dot{m}_{Bui,L} \leq \dot{m}_{Bui} \leq \dot{m}_{Bui,U} \quad (8)$$

where \dot{Q}_{Bui} is the total building heat demand, including the heat demand for SH and DHW systems as shown in Equation (9). \dot{Q}_{SH} can be further divided into the demand for the radiator heating system \dot{Q}_{rad} and the demand for the ventilation system \dot{Q}_{ven} , as described in Equation (10). \dot{m}_{Bui} and ΔT_{Bui} are the mass flow rate and temperature difference for water at the primary side of the building substation, respectively. T_{sup} and T_{ret} are the supply and return temperature of water at the primary side of the building substation, respectively. $\Delta T_{Bui,L}$, $T_{sup,L}$, and $\dot{m}_{Bui,L}$ are the lower bounds for ΔT_{Bui} , T_{sup} , and \dot{m}_{Bui} , respectively. $\Delta T_{Bui,U}$, $T_{sup,U}$, and $\dot{m}_{Bui,U}$ are the upper bounds for ΔT_{Bui} , T_{sup} , and \dot{m}_{Bui} , respectively.

$$\dot{Q}_{Bui} = \dot{Q}_{SH} + \dot{Q}_{DHW} \quad (9)$$

$$\dot{Q}_{SH} = \dot{Q}_{rad} + \dot{Q}_{ven} \quad (10)$$

Buildings' dynamics were described using a simplified-lumped-capacity model generated from resistance-capacitance networks analogous to electric circuits, as stated in Equations (11), (12), and (13).

$$C_{env} \cdot \frac{dT_{env}}{dt} = \frac{T_{ia} - T_{env}}{R_{i,e}} + \frac{T_{oa} - T_{env}}{R_{o,e}} \quad (11)$$

$$C_{ia} \cdot \frac{dT_{ia}}{dt} = \frac{T_{ma} - T_{ia}}{R_{i,m}} + \frac{T_{env} - T_{ia}}{R_{i,e}} + \frac{T_{oa} - T_{ia}}{R_{o,e}} + \frac{R_{win}}{T_{oa} - T_{ia}} + \frac{R_{ven}}{T_{oa} - T_{ia}} + \dot{Q}_{rad} + \dot{Q}_{ven} + \dot{Q}_{in} \quad (12)$$

$$C_{ma} \cdot \frac{dT_{ma}}{dt} = \frac{T_{ia} - T_{ma}}{R_{i,m}} \quad (13)$$

where C and R are the heat capacitance and resistance, and T refers to the temperature. Building envelopes (including exterior walls and roofs), indoor air, outdoor air, internal thermal mass, window, and ventilation (including infiltration and mechanical ventilation) are denoted by the subscripts *ven*, *ia*, *oa*, *ma*, *win*, and *ven*, respectively. $R_{i,e}$ represents the heat resistance between indoor air and building envelopes, $R_{o,e}$ represents the heat resistance between outdoor air and building envelopes, and $R_{i,m}$ represents the heat resistance between indoor air and interior thermal mass. \dot{Q}_{in} is the internal heat gain.

As described in Equation (14), the lower bound of the supply temperature should be high enough for the SH and DHW systems to maintain a comfortable indoor temperature while avoiding hygiene problems. Equation (15) defined the lower bound of the supply temperature for the SH system (He et al., 2009), whereas Equation (16) determined the lower bound for the DHW system, which is required by European standard CEN/TR16355 ("CEN/TR16355 Recommendations for prevention of Legionella growth in installations inside buildings conveying water for human consumption," 2012).

$$T_{sup,L} = \max(T_{sup,SH,L}, T_{sup,DHW,L}) \quad (14)$$

$$T_{sup,SH,L} = T_{ia} + 0.5 \cdot (T_{sup,SH,des} + T_{ret,SH,des} - 2 \cdot T_{ia,des}) \cdot \left(\frac{T_{ia,des} - T_{oa}}{T_{ia,des} - T_{oa,des}} \right)^{1/b} + 0.5 \quad (15)$$

$$\cdot (T_{sup,SH,des} - T_{ret,SH,des}) \cdot \left(\frac{T_{ia,des} - T_{oa}}{T_{ia,des} - T_{oa,des}} \right) T_{sup,DHW,L} = 60^\circ\text{C} \quad (16)$$

where $T_{sup,SH,L}$ and $T_{sup,DHW,L}$ are the lower bounds for the SH and DHW system's supply temperatures, respectively. T_{ia} and T_{oa} are the

indoor and outdoor air temperatures, respectively. $T_{sup,SH}$ and $T_{ret,SH}$ are the supply and return temperatures for the SH system, respectively. b is a parameter defining the radiator's characteristic. des is a subscript that refers to the design conditions.

The lower bound of the water mass flow rate $\dot{m}_{Bui,L}$ was zero, while the upper bound $\dot{m}_{Bui,U}$ was constrained by the distribution system's capacity. In addition, the lower bound of the water temperature difference $\Delta T_{Bui,L}$ was zero, and the upper bound of the water temperature difference $\Delta T_{Bui,U}$ was obtained from linear regression Equation (17).

$$\Delta T_{Bui,U} = a_0 + a_1 \cdot T_{sup} \quad (17)$$

where a_0 and a_1 are parameters.

2.3. Distribution pipeline model

Equations (18), (19), and (20) were used to describe the heat loss from pipelines.

$$\dot{Q}_{loss,pip} = \dot{Q}_{loss,pip,sup} + \dot{Q}_{loss,pip,ret} \quad (18)$$

$$\dot{Q}_{loss,pip,sup} = L \cdot \pi \cdot d \cdot \frac{(R_g + R_i) \cdot \Delta T_{pip,sup} - R_c \cdot \Delta T_{pip,ret}}{(R_g + R_i)^2 - R_c^2} \quad (19)$$

$$\dot{Q}_{loss,pip,ret} = L \cdot \pi \cdot d \cdot \frac{(R_g + R_i) \cdot \Delta T_{pip,ret} - R_c \cdot \Delta T_{pip,sup}}{(R_g + R_i)^2 - R_c^2} \quad (20)$$

where $\dot{Q}_{loss,pip}$, $\dot{Q}_{loss,pip,sup}$, and $\dot{Q}_{loss,pip,ret}$ represent the overall heat loss from pipes, supply pipe heat loss, and return pipe heat loss, respectively. L refers to the route length for the pair of pipes. d is the outer pipe diameter. R_i , R_g , and R_c are the resistances for the insulation, ground, and coinciding, respectively, and they can be obtained by Equations (21), (22) and (23), respectively. $\Delta T_{pip,sup}$ and $\Delta T_{pip,ret}$ are the temperature differences for the supply and return pipes, respectively, and can be obtained using Equations (24) and (25).

$$R_i = \frac{d}{2 \cdot \lambda_i} \cdot \ln \frac{D}{d} \quad (21)$$

$$R_g = \frac{d}{2 \cdot \lambda} \cdot \ln \frac{4 \cdot h}{D} \quad (22)$$

$$R_c = \frac{d}{2 \cdot \lambda} \cdot \ln \left(\left(\frac{2 \cdot h}{s} \right)^2 + 1 \right)^{0.5} \quad (23)$$

$$\Delta T_{pip,sup} = T_{pip,sup} - T_{grou} \quad (24)$$

$$\Delta T_{pip,ret} = T_{pip,ret} - T_{grou} \quad (25)$$

where D is the outer insulation diameter, h is the distance between the pipe centres and the ground

surface, s is the distance between pipe centres, and λ and λ_i are the heat conductivity for the ground and insulation, and T_{grou} is the ground temperature.

2.4. Water tank thermal energy storage model

The dynamics of the water tank were described using a one-dimensional WTES model (Powell & Edgar, 2013).

$$c \cdot \rho \cdot A_{XS} \cdot \frac{\partial T}{\partial t} = c \cdot (\dot{m}_{sou} - \dot{m}_{use}) \cdot \frac{\partial T}{\partial x} - U \cdot P \cdot (T(t, x) - T_{amb}) + \varepsilon \cdot A_{XS} \cdot \frac{\partial^2 T}{\partial x^2} \quad (26)$$

where T is the water temperature. x is the height of the tank. t is the time. ρ is the water density. A_{XS} and P are the cross-sectional area and perimeter of the tank, respectively. \dot{m}_{sou} and \dot{m}_{use} are the water mass flow rate of the heat source and user side, respectively. T_{amb} is the ambient temperature. U is the U-value of the tank wall. ε is a parameter representing the combined heat transfer effect of water through diffusion, conduction, and mixing due to turbulent flow.

By discretizing the tank into n nodes, spatial derivatives were approximated using numerical techniques. Equation (27) shows the ordinary differential equation for the i th node. Equations (28) and (29) were used to compute the heat loss and heat flow rate of the i th node, while Equations (30) and (31) were used to get the total heat loss and heat flow rate of the WTES.

$$c \cdot \rho \cdot A_{XS} \cdot \Delta x \cdot \frac{dT_i}{dt} = c \cdot \dot{m}_{use} \cdot (T_{i-1} - T_i) + c \cdot \dot{m}_{sou} \cdot (T_{i+1} - T_i) - U \cdot P \cdot \Delta x \cdot (T_i - T_{amb}) + \frac{\varepsilon \cdot A_{XS}}{\Delta x} \cdot (T_{i+1} - 2 \cdot T_i + T_{i-1}) \quad (27)$$

$$\dot{q}_{loss, TES, i} = U \cdot P \cdot \Delta x \cdot (T_i - T_{amb}) \quad (28)$$

$$\dot{q}_{TES, i} = c \cdot \dot{m}_{sou} \cdot (T_{i+1} - T_i) \quad (29)$$

$$\dot{Q}_{loss, TES} = \sum_{i=1}^{n-1} \dot{q}_{loss, TES, i} \quad (30)$$

$$\dot{Q}_{TES} = \sum_{i=1}^{n-1} \dot{q}_{TES, i} \quad (31)$$

where Δx is the node length, and T_i is the water temperature of the i th node. $\dot{q}_{loss, TES, i}$ and $\dot{q}_{TES, i}$ are

the heat loss and heat flow rate of the i th node, respectively.

3. Optimization frameworks in JModelica

Two optimization frameworks were formulated: a long-term optimal operation framework aimed at a yearly level open-loop optimization with a two-hour resolution, and a model predictive control (MPC) framework aimed at a daily level close-loop optimization with a one-hour resolution. These two optimization frameworks used the same objective function, minimizing heating costs while tracking the reference indoor temperature, as shown in Equation (32). In addition, these two optimization frameworks used the same system dynamic models and inequality constraints introduced in Section 2, as shown in Equations (33), (34), (35), and (36).

$$\int_{t_0}^{t_f} EP(t) \cdot \dot{Q}(t) dt + LP \cdot \dot{Q}_{pea} + W \cdot \int_{t_0}^{t_f} (T_{ia}(t) - T_{ia}^{ref}(t))^2 \cdot dt \quad (32)$$

subject to:

$$\dot{Q}(t) \leq \dot{Q}_{pea} \quad (33)$$

$$F(t, \mathbf{z}(t)) = 0 \quad (34)$$

$$F_0(t_0, \mathbf{z}(t_0)) = 0 \quad (35)$$

$$z_L \leq \mathbf{z}(t) \leq z_U \quad (36)$$

where $\dot{Q}(t)$ is the heat flow rate supplied from the central DH to the main substation. \dot{Q}_{pea} and LP is the peak load and the peak load related heating price, respectively. $EP(t)$ is the heating price for the heat use related heating cost. $T_{ia}(t)$ and $T_{ia}^{ref}(t)$ are the simulated indoor temperature and its reference value at time t . $\mathbf{z} \in \mathbb{R}^{n_z}$ represents the time-dependent variables, which includes the manipulated variable $\mathbf{u} \in \mathbb{R}^{n_u}$ to be optimized, the differential variable $\mathbf{x} \in \mathbb{R}^{n_x}$, and the algebraic variable $\mathbf{y} \in \mathbb{R}^{n_y}$. Equation (34) defines the system dynamics and Equation (35) is the initial conditions of the system. $z_L \in [-\infty, \infty]^{n_z}$ and $z_U \in [-\infty, \infty]^{n_z}$ are the lower and upper bounds, respectively.

The long-term open-loop optimization framework is illustrated in Figure 2. This long-term optimization framework computes the optimal operation trajectory for a whole operating year. This framework was used for optimal design and operation in the research (Li, Hou, Hong, et al., 2022; Li, Hou, Tian, et al., 2022).

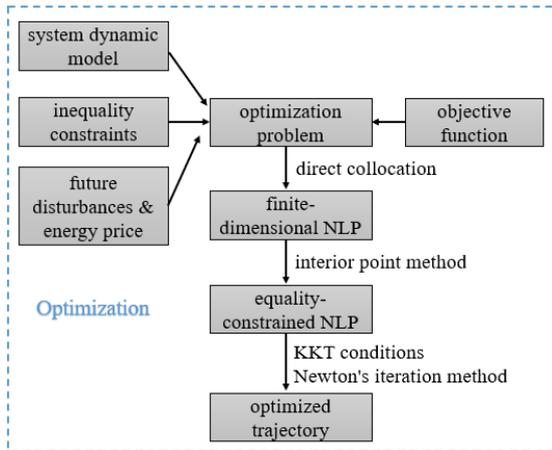


Figure 2: Long-term open-loop optimization framework for optimal design and operation.

The short-term close-loop optimization framework is illustrated in Figure 3. This short-term optimization framework computes the optimal operation trajectory within the prediction horizon (mostly from half to two days). This framework was used for MPC in a study presented in the results section 4.2.

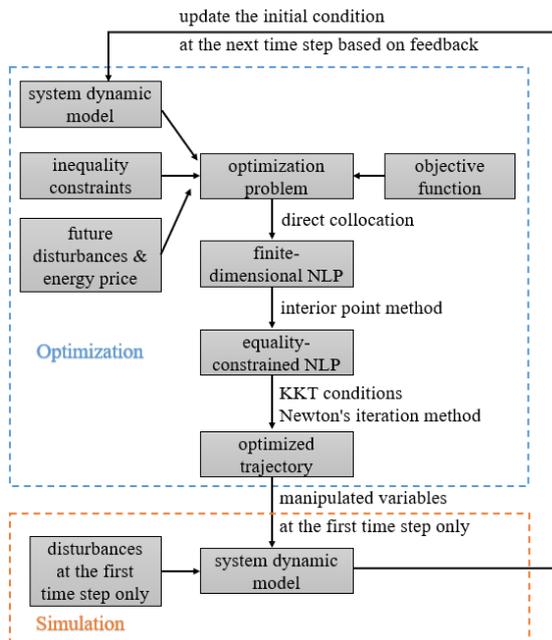


Figure 3: Short-term close-loop optimization framework for optimal control.

4. Case study

A campus DH system in Trondheim, Norway, was chosen as the case study. The campus DH system is a prosumer with a distributed heat source, as shown in Figure 4. The distributed heat source is the university data centre, which recovers the condensing waste heat from the data centre's cooling system. The campus DH system supplies heat for the university buildings with a total building area of 300,000 m². The main substation

is used to connect the campus DH system with the city's central DH system. According to the measurements from June 2017 to May 2018, the total heat supply for the campus DH system was 32.8 GWh. About 80% of the heat supply came from the central DH system through the main substation. The other 20% came from the waste heat recovery from the data centre.

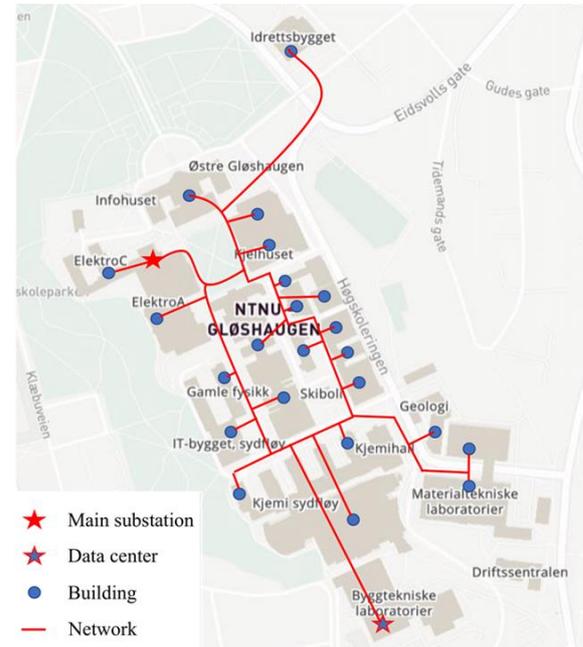


Figure 4: Campus district heating system.

5. Results

The developed DH system model presents high accuracy, and the results of model validation can be found in articles (Li, Hou, Hong, et al., 2022; Li, Hou, Tian, et al., 2022). This article only presents the key results on energy and economic performance of the long-term and short-term optimization frameworks.

5.1. Energy and economic performance of the long-term optimization framework

The long-term optimization framework tested the idea of introducing a WTES into the campus DH system. Figure 5 and Figure 6 present the annual heat use and the yearly peak load for the scenario before and after introducing the WTES, respectively. These two indicators quantified the heat supply from the central DH system to the heat prosumer through the main substation. It can be observed from Figure 5 that introducing the WTES reduced the annual heat use from 26.2 GWh to 25.9 GWh, meaning a heat use saving of 1%. Compared to this less significant heat use saving, a more obvious peak load shaving was obtained as shown in Figure 6, the yearly peak load was shaved from 12.4 MW to 9.5 MW, a shaving of 24%.

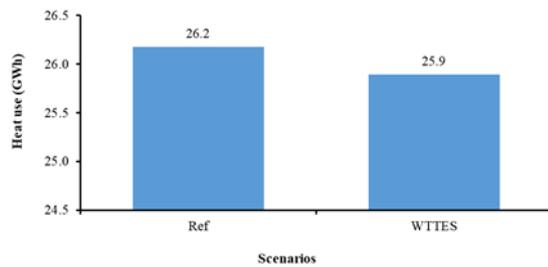


Figure 5: Annual heat use for the scenario before and after introducing WTTS.

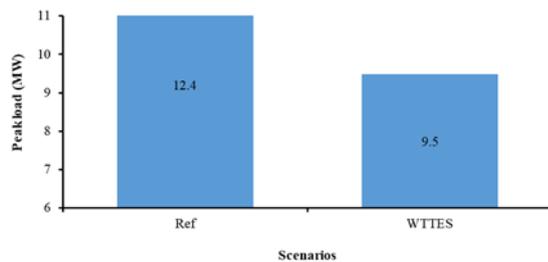


Figure 6: Yearly peak load for the scenario before and after introducing WTTS.

The resulting annual heating cost for the scenario before and after introducing WTTS is presented in Figure 7. Introducing WTTS cut the annual heating cost from 20.7 million NOK to 19.3 million NOK, which meant a cost saving of 7% was achieved.

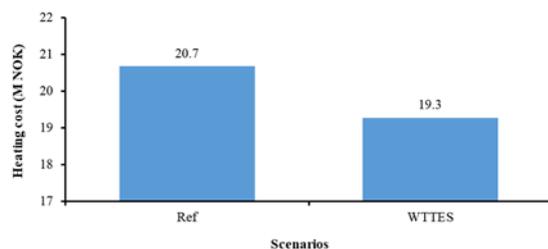


Figure 7: Annual heating cost for the scenario before and after introducing WTTS.

5.2. Energy and economic performance of the short-term optimization framework

The short-term optimization framework evaluated the potential of the MPC strategy. In this study, two scenarios, one MPC scenario and one rule-based control (RBC) scenario, were designed and compared. Figure 8 and Figure 9 present the heat use and peak load for the MPC and RBC scenario, respectively.

It can be observed from Figure 8 that the MPC scenario reduced the heat use from 5.12 GWh to 5.02 GWh in January, meaning a heat use saving of 2%. Meanwhile, in April the reduction was from 2.33 GWh to 2.25 GWh, a saving of 3%. Similar to the long-term optimization problem with less significant heat use saving, a more obvious peak load shaving was obtained as shown in Figure 9, the peak load was shaved from 11.6 MW to 10.9

MW in January, a shaving of 6%. In addition, the shaving was from 8.0 MW to 7.1 MW in April, a shaving of 11%.

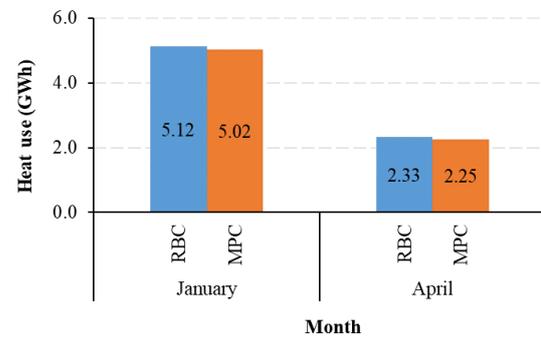


Figure 8: Heat use for the MPC and RBC scenario.

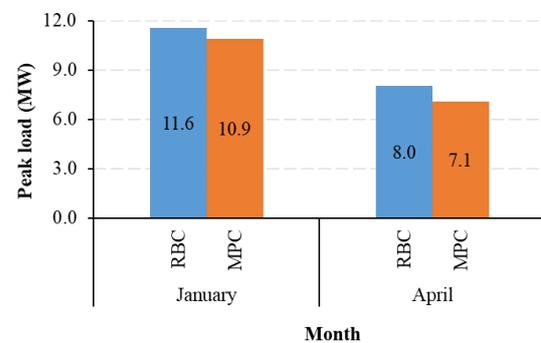


Figure 9: Peak load for the MPC and RBC scenario.

The resulting energy costs (including heating and electricity cost) for the MPC and RBC scenarios is presented in Figure 10. In Figure 10, the electricity cost included the spot price-related fee and surcharges, which had prices in NOK/kWh. The proposed MPC strategy cut the energy cost from 3.47 million NOK to 3.40 million NOK in January, which meant a cost saving of 2% was achieved. Meanwhile, the cutting was from 1.92 million NOK to 1.86 million NOK in April, a saving of 3%. These cost savings were subjected to the specified energy price models in this study. For other cases, the results may be different.

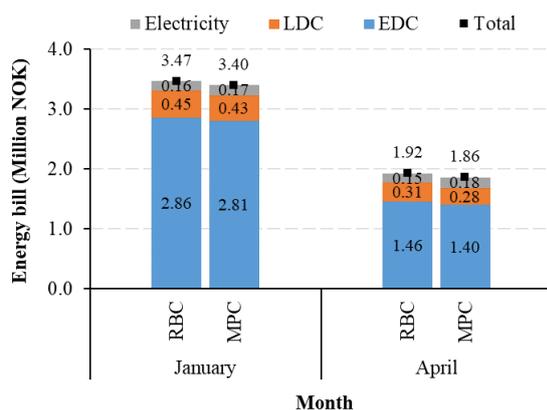


Figure 10: Heating cost for the MPC and RBC scenario.

6. Conclusion

This study presents a fast and effective method for modelling and optimizing DH systems using the Modelica language. A case study on a university campus DH system in Norway showed that the method was effective both for long-term optimal operation and short-term optimal control problems. For the studied case, the approach achieved energy cost saving by energy use reduction and peak demand shaving. It is worth noting that the achieved results may be subjected to specified energy price models, however, this study provided a generalized method to solve this type of research problem.

Acknowledgement

The authors gratefully acknowledge the support from the H2020 EU-funded project Climate Positive Circular Communities (grant agreement ID: 101036723), which aims at creating climate positive circular communities in Europe and increasing the building renovation rate in the continent.

References

Åberg, M., Fälting, L., Lingfors, D., Nilsson, A. M., & Forssell, A. (2020). Do ground source heat pumps challenge the dominant position of district heating in the Swedish heating market? *Journal of Cleaner Production*, 254, 120070. <https://doi.org/https://doi.org/10.1016/j.jclepro.2020.120070>

CEN/TR16355 Recommendations for prevention of Legionella growth in installations inside buildings conveying water for human consumption. (2012). In.

Connolly, D., Lund, H., Mathiesen, B. V., Werner, S., Möller, B., Persson, U., Boermans, T., Trier, D., Østergaard, P. A., & Nielsen, S. (2014). Heat Roadmap Europe: Combining district heating with heat savings to decarbonise the EU energy system. *Energy Policy*, 65, 475-489. <https://doi.org/10.1016/j.enpol.2013.10.035>

He, P., Sun, G., Wang, F., Wu, H., & Wu, X. (2009). *Heating engineering (in Chinese)*. China Architecture & Building Press.

Heating and cooling- European Commission. <https://ec.europa.eu/energy/en/topics/energy-efficiency/heating-and-cooling>

Hou, J., Li, H., & Nord, N. (2019). Optimal control of secondary side supply water temperature for substation in district heating systems. *E3S Web Conf.*, 111, 06015. <https://doi.org/10.1051/e3sconf/201911106015>

Hou, J., Li, H., & Nord, N. (2022). Nonlinear model predictive control for the space heating system of a university building in Norway. *Energy*, 253, 124157. <https://doi.org/https://doi.org/10.1016/j.energy.2022.124157>

Hou, J., Li, H., Nord, N., & Huang, G. (2022). Model predictive control under weather forecast uncertainty for HVAC systems in university buildings. *Energy and Buildings*, 257, 111793. <https://doi.org/https://doi.org/10.1016/j.enbuild.2021.111793>

IBPSA Project 1. (Jan 2021). <https://ibpsa.github.io/project1/index.html>

In focus: Energy efficiency in buildings. (Jan 2021). https://ec.europa.eu/info/news/focus-energy-efficiency-buildings-2020-feb-17_en#:~:Text=collectively%2c%20buildings%20in%20the%20eu.%2c%20usage%2c%20renovation%20and%20demolition

Li, H., Hou, J., Hong, T., Ding, Y., & Nord, N. (2021). Energy, economic, and environmental analysis of integration of thermal energy storage into district heating systems using waste heat from data centres. *Energy*, 219, 119582. <https://doi.org/https://doi.org/10.1016/j.energy.2020.119582>

Li, H., Hou, J., Hong, T., & Nord, N. (2022). Distinguish between the economic optimal and lowest distribution temperatures for heat-prosumer-based district heating systems with short-term thermal energy storage. *Energy*, 248, 123601. <https://doi.org/https://doi.org/10.1016/j.energy.2022.123601>

Li, H., Hou, J., Tian, Z., Hong, T., Nord, N., & Rohde, D. (2022). Optimize heat prosumers' economic performance under current heating price models by using water tank thermal energy storage. *Energy*, 239, 122103. <https://doi.org/https://doi.org/10.1016/j.energy.2021.122103>

Li, H., & Nord, N. (2018). Transition to the 4th generation district heating- possibilities, bottlenecks, and challenges. *Energy Procedia*, 149, 483-498. <https://doi.org/https://doi.org/10.1016/j.egypro.2018.08.213>

Li, H., & Nord, N. (2019). Operation strategies to achieve low supply and return temperature in district heating system. *E3S Web Conf.*, 111, 05022. <https://doi.org/10.1051/e3sconf/201911105022>

Powell, K. M., & Edgar, T. F. (2013). An adaptive-grid model for dynamic simulation of thermocline thermal energy storage

systems. *Energy Conversion and Management*, 76, 865-873.
<https://doi.org/https://doi.org/10.1016/j.enconman.2013.08.043>

Sayegh, M. A., Danielewicz, J., Nannou, T., Miniewicz, M., Jadwischak, P., Piekarska, K., & Jouhara, H. (2017). Trends of European research and development in district heating technologies. *Renewable and Sustainable Energy Reviews*, 68, 1183-1192. <https://doi.org/10.1016/j.rser.2016.02.023>

Werner, S. (2017). International review of district heating and cooling [Review]. *Energy*, 137, 617-631.
<https://doi.org/10.1016/j.energy.2017.04.045>

Temporal Fusion Transformer for thermal load prediction in district heating and cooling networks

Fabian Behrens ^a, Stefan Leiprecht ^a, Jonas Brantl ^a, Matthias Finkenrath ^{a,*}

^a *University of Applied Sciences Kempten*
matthias.finkenrath@hs-kempten.de

Abstract

Accurate forecasting of thermal loads is a critical factor for operating district heating and cooling networks economically, efficiently and with minimized emissions. If thermal loads are known with high accuracy in advance, use of renewable energies can be maximized, and fossil generation, in particular in peaking units, can be avoided. Machine learning has already proven to be an efficient tool for time series forecasting in this context. One recent advancement in machine learning is the "Temporal Fusion Transformer" (TFT), which shows especially good results in the area of time series forecasting. This paper examines the performance of TFT in the concrete context of thermal load forecasting for district heating and cooling networks. First, a brief summary of differences between TFT and other machine learning methods is given. Secondly, it is described how the method can be adopted to train a machine learning model for thermal load forecasting. The data to train and evaluate the neural network is based on 8 years of hourly operating data made available from the district heating network of the city of Ulm in Germany. The presented technique is used to produce 72 hours of heating load forecasts for three different district heating grids in the city of Ulm. The results are compared to forecasts of other machine learning methods that have been previously made as part of the publicly funded research project "deepDHC", in order to evaluate if TFT is an improvement to further reduce forecasting uncertainties.

1. Introduction

Precise forecasting of thermal loads is crucial for operating district heating networks efficiently, economically and environmentally friendly. If precise load forecasts are available to the operator, the use of fossil-fuelled peaking boilers can be significantly reduced. In addition, integration of fluctuating renewable into the grid can be maximized. A precise long-term load forecast several days ahead also simplifies fuel ordering, or planned maintenance. Hence this work focuses on thermal load forecasts throughout 72 hours in advance. The data used for the process is based on hourly data from the district heating network in Ulm, a medium-sized city in southern Germany with about 130,000 inhabitants.

2. Related Work

Accurate prediction of heat loads has become an interesting field of application for modern time series forecasting methods. Its importance even increases with a rising global energy demand, decreasing reserves of fossil fuels and the impact of using fossil fuels on climate change (Benalcazar and Kamiński, 2019). District heating and cooling can be a sufficient way to reduce carbon dioxide emissions by optimizing fuel consumption (Werner, 2017). Machine learning has proven to be an attractive option for generating accurate thermal load predictions also in the context of district heating and cooling (e.g. (Saloux and Candanedo, 2018; Leiprecht et al., 2021)).

Different algorithms have been evaluated in recent years for this purpose, such as Adaptive Boosting (AdaBoost) (Freund et al., 1996) and its derivative Extreme Gradient

Boosting (XGBoost) (Friedman, 2000), recurrent neural networks (RNNs) like Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997) and the Seasonal Autoregressive Integrated Moving Average Exogenous model (SARIMA) (Fang and Lahdelma, 2016).

While the traditional machine learning methods produce decent results, they all have the problem of not being significantly better than statistical methods on many time series problems. In many cases the larger overhead of implementing these methods make them economically less attractive than statistical approaches (Lim and Zohren, 2021). Current research tries to solve this issue by improving the abilities of models of learning from the past, which helps these methods to further improve their accuracy in time series forecasting by reducing their overfitting.

One approach to this issue is the Temporal Fusion Transformer (TFT) (Lim et al., 2021). This new method is an attention based network. TFT is already used in a number of areas for time series forecasting, like meteorology (Wu et al., 2022), medicine (Phetrittikun et al., 2021) and the stock market (Hu, 2021). While there will be some commonalities in input data with forecasting in meteorology, right now there is no research about the performance of TFT for energy demand forecasting. This paper therefore aims to give a first estimation of what results can be expected in this area.

Most of the work that was done on the topic of energy demand forecasting focused on a 24 hour time horizon (Benalcazar and Kamiński, 2019; Xue et al., 2019). This paper instead focuses on an extended forecasting period of up to 72 hours, in order to allow further optimised dispatch

planning of power plants and thermal energy storages.

3. Temporal Fusion Transformer

TFT (Lim et al., 2021) is a new approach explicitly developed for time series forecasting. Therefore it brings a number of qualities that are very helpful for training robust forecasting models. Usually machine learning methods use information of the past to learn the behaviour of a time series in order to create accurate forecasts. In order to learn patterns in historic data, neurons have to memorize data they have seen earlier during training. Most of the time this is currently achieved with RNNs (Hochreiter and Schmidhuber, 1997; Jaeger, 2001). However RNNs oftentimes face the problem of expecting that all input data is known, even for time steps in the future. However this is not always possible. For example the temperature in the future can not be known for sure. It can be estimated with the help of weather forecasts, but there exist parameters for which there is no way of knowing them upfront. TFT on the other side does not belong into the class of RNNs, it instead uses a transformer architecture (Vaswani et al., 2017). Transformers use a more advanced method to learn patterns in historic data. As a consequence they support a variety of different types of input data, which enables the usage of features whose values can not be known during prediction. Next some of the advantages of TFT for time series forecasting will be explained.

Firstly TFT supports multi step forecasting. This means that multiple forecasts can be done in one prediction call. For example in the case of this paper one prediction creates forecasts for the next 72 hours with an interval of one hour. Single step methods on the other side would only predict one step at a time. To predict further ahead than one hour this would require the user to do one prediction, add it to the input data and then run the next prediction until the size of the targeted prediction interval is reached. This is undesirable because prediction errors in early time steps can influence the prediction of later time steps. Many new machine learning methods support this type of forecasting, but it is still important to have and can't be done with every of the methods mentioned in this paper. As transformers are not RNNs, another way to learn relations between historic data is needed. TFT uses an attention based method for solving this issue.

TFT supports three different types of input data: temporal data which is known in the future, temporal data which is unknown in the future and static variables. The first group is the most common type of data as known from other forecasting problems. For example the hour of the day for which a prediction is made is such a feature. It is known for historic data during training, but it can also be determined for every future time step. The second type of data is only needed during training, but can be missing when the model is used for predictions. A good example is weather data. Usually the historic weather is known but can not be determined for the future. Other machine learning algorithms would require to guess the weather data or for example use a weather forecast instead of real weather data. However these approaches don't deliver the actual correct values. The model however is trained assuming the provided values are correct, which leads to the predictions being inherently wrong when a forecast is used as input instead. TFT on the other hand makes it possible to use any feature in training even when it is not possible to provide it during prediction phase. The last type allows to add static data that will not change over time, e.g. the holidays of the location of the prediction or the location itself.

Another feature is the support of predicting multiple time series at once. Usually every time series that should be predicted needs its own neural network that is fitted to the training data, in order to create the best possible forecast based on the provided data. TFT provides the possibility to add multiple sets of input data to a model. The model then learns which dataset is used for predicting which time series and fits its model in a way that can predict all time series at the same time. This process can be very helpful because this can save a lot of time. Usually in the process of fitting a model, the hyperparameters will be optimized to. If each time series would need its own neural model multiple of this hyperparameter optimizations would be required. In TFT only one for the model as a whole is needed.

Additionally TFT tries to make the process of working with it more interpretable. Usually neural networks are black boxes that can not be understood in their way of calculating a result. This makes the process of improving a model especially tough when the model just does not seem to get better. TFT solves this issue by a so called multi head attention mechanism. This process works as follows. TFT always calculates the importance of different input features as part of its attention system. These importances can be analysed and can be provided to the machine learning developer. They can then examine which features are important or which impact different features had during one training. Altering features and then evaluating the impact of the change to the performance of a parameter makes it much easier to optimize and understand a model. The prior explained advantage already includes a last advancement TFT provides. Since the TFT calculates the importance of all input features, it can also realize that a given feature has no importance to the prediction problem. In this case TFT can weigh the effect of the feature with a zero which leads to the feature having no effect in training and prediction. This can also save a lot of time, because the right features do not have to be selected up front by a data scientist.

4. Model Training

The training was done with a little bit more than six years of historic data beginning at 02.09.2014 until the 31.12.2020. This time frame was split into a training and validation dataset with the first 70% being the training dataset and the last 30% being used for validation.

As TFT can analyse features itself in terms of their importance for the problem, almost all features that were available to us were used to train the model. In total those were more than 37 features. Some of the most important ones can be found in Tab. 1

Table 1: Used features for TFT training.

Name	Description
Last Load	The thermal load of the prior hour in MW
avgLoad6/12/24	The average thermal load over the last 6/12/24 hours in MW
Temperature	Current air temperature in Celsius
avgTemp6/12/24	The average temperature over the last 6/12/24 hours in Celsius
Dewpoint	Current dewpoint in Celsius
Season Sin	One period of a sine wave mapped onto the period of a whole year

Key features are temperature and the current thermal load profile of the district heating network. The temperature is the main factor that changes heating behaviour, especially in residential areas. Therefore the required thermal load strongly correlates with the temperature. The dew point acts like an amplifier of the temperature. In our case both values are very similar most of the time so it also is a good indicator for how the heating demand evolves.

The last load is a good indicator, because most of the time the thermal load demand does not change drastic over a short period of time. Therefore, it usually acts as a good estimation of the next thermal load required. Both of these parameter can also be used as averages over the last few hours. These averages can indicate the overall trend of the current thermal load demand which can help to estimate if the load demand will rise or decline over the next few hours.

The Season Sin feature encodes which day of the year it is at a given prediction point. This can be helpful in improving the understanding of time in the neural networks. While the weather is not exactly the same at the same day over multiple years it can be similar, because it is usually around the same time of the year when the weather gets warmer or colder. Season Sin helps to learn to take this periodic behaviour into account.

The implementation of TFT was not done by ourselves. Pytorch(Paszke et al., 2019) already provides an implementation of TFT which was used in this paper.

5. Methodology

5.1. Metrics

Benchmarking the forecasts is not an easy task, since there is no standardized metric available. The Mean Absolute Percentage Error (MAPE) is probably the most commonly used metric for measuring forecast accuracy. It is widespread in finance or other forecasting applications, especially if enough data is available. The MAPE is dimensionless and independent of the magnitude of the values considered. At the same time, it can be clearly interpreted. A MAPE of zero corresponds to a perfect forecast (Clark, 2013; Armstrong and Collopy, 1992). Its equation can be seen in (1). It is the mean of the sum of the absolute error e_i divided by the real value d_i . n is the number of prediction-load pairs that are used to calculate the error.

$$MAPE = \frac{1}{n} \sum_{i=0}^n \frac{|e_i|}{d_i} \quad (1)$$

In addition the Mean Absolute Error (MAE)(Willmott and Matsuura, 2005) was used as a second metric. It is the mean of the sum of n errors. The errors are the absolute deviation of the prediction y_i from the real thermal load x_i . The MAE oftentimes has the disadvantage of being hard to interpret. In many cases the range of values the target of a prediction can have is not known. In this situations it is hard to argue if the measured absolute improvement is significant or not. However in our context this is not the case. For each of the district heating networks considered the thermal loads that can be expected are known. Moreover the unit of the MAE in this case is Megawatts, a unit that is very easily interpretable. The MAE should not be used to compare different district heating networks, because their load profile can differ significantly, however for each individual network the metric can be very helpful for the power plant. It knows which ways it has to provide the thermal load to a given

district heating network and how much energy each of these options can provide. In this context absolute values can be very helpful to optimize the energy production for a given district heating network.

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (2)$$

So while the MAPE is a good indicator for making general assumptions about the performance of a machine learning method for district heating networks, the MAE can be used for closer evaluation for concrete scenarios. Moreover the MAE can relativise a high MAPE if the absolute target values are quite low.

5.2. District Heating Networks

For the comparison three different district heating networks of the city of Ulm were used. The first network acts as a good general baseline for the performance of a machine learning method, as the network is very consistent and most algorithms evaluated so far perform best on it. It has a total length of 40 km and provides space heating for over 7500 households with an average annual heating demand of 75 GWh, and a heat load ranging from 2 to 22 MW. Water with a temperature between 70°C and 110°C is used as a heat transfer fluid.

Additionally two more networks were selected which have a more complex thermal load profile. The second district heating network uses a combination of steam and hot water for heat transfer. Steam transfer uses steam at a heat of 130°C while the temperature of the water varies between 70°C and 110°C. The network provides space heating to over 13.000 households and has a heat load ranging from 1 to 19 MW.

The third and last district heating network considered, mainly supplies industrial buildings instead of private households, which leads to one more different thermal load profile. It is run with 120°C hot water as transfer fluid. Additional to many factories the district heating network supplies 220 households. It has a heat load ranging from 2 to 25 MW.

While still providing promising forecasts, many of the machine learning methods evaluated prior produced far less optimal results on district heating network two and three. Evaluating the results of TFT on these networks too, can show if the strengths of its new approach help dealing with overall harder to predict scenarios. Also the addition of a more industrial focused district heating network provides more insights for a wider range of use cases.

5.3. Time Frames

As time frame for the comparison the whole year of 2021 was used. With this time frame the evaluation should hold meaningful results for the active usage of the model in a power plant by covering many different scenarios and load profiles of different seasons. Moreover the data is very new, thereby the results can be extrapolated into the future of the net more easily than an older time frame.

In addition to the comparison over the whole year, several shorter time frames are evaluated too. These are:

- Winter (01.01.2021-28.02.2021 and 01.12.2021-31.12.2021)
- Spring (01.03.2021-31.05.2021)
- Summer (01.06.2021-31.08.2021)
- Fall (01.09.2021-30.11.2021)

These four intervals resemble four parts of the year which have different load profiles. Evaluating them makes it possible to further investigate how TFT performs in different scenarios. For example the load profile is very consistent in winter and summer which resulted in pretty good predictions for the already evaluated methods. However the older approaches struggle far more in the spring and fall time frame. During these periods the load profile is much more ambivalent. This could be a problem for these older models because they tend to overfit. Comparison of these time frames will show if TFT can adapt better to learning more unpredictable time series.

5.4. Machine Learning Methods

TFT will be compared to the results of three machine learning methods. Namely LSTM, AdaBoost and XGBoost. Since these methods can not abstract which features are not important to them, these methods were not trained by providing all possible features available as input. To find the most important features for the given machine learning method a feature reduction was performed. The used method was the scikit learn (Pedregosa et al., 2011) implementation of Recursive Feature Elimination (Guyon et al., 2002) with cross validation. The features used for each method can be seen in Tab. 2

Table 2: Features used for training of different machine learning methods.

Method	Used Features
LSTM	loads of the last 6 hours, temperature, season sin, avgTemp24, hour, temperature forecast, dewpoint forecast
AdaBoost	loads of the last 6 hours, season sin, avgTemp12, avgLoad24, hour, temperature forecast for the next 5 hours
XGBoost	season sin, avgTemp12, avgLoad24, hour, loads of the last 6 hours, temperature forecast for the next 3 hours

The features correspond to the features explained in section 4. The amount of last loads and forecasts used in AdaBoost and XGBoost vary between the models for each of the three different nets in order to further improve the results of the nets. The range in which those parameters lie is three to six hours. The forecasts of all LSTM neural networks are the same as the LSTM predicts the load as a multi step while AdaBoost and XGBoost use a single step method.

6. Discussion of Results

The promising results of TFT can be seen in Fig. 1. TFT beats all other machine learning methods on every district heating network evaluated. Moreover this is special, because it is the first time in our investigation that one method is the best one for any district heating network it was tested on. For example without TFT, LSTM would be the best way to predict network one while XGBoost is the best method for network two and three. Moreover the reduction of the error is impressive. On the easiest network it beats LSTM by two percent points. This is a good result, however it is not that relevant for the

facility, because looking at the absolute error, it already lies beneath one MW. However the improvement is much stronger in the tougher to predict networks two and three. In both cases the MAPE of TFT is almost half as high as the MAPE of XGBoost. This indicates how the attention-based approach of TFT is way better in generalizing the problem than older machine learning methods and thereby avoids overfitting.

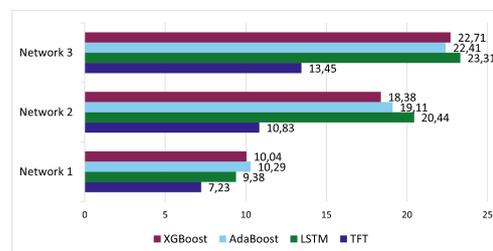


Figure 1: MAPE for the whole year of 2021

As a next step the four different time frames will be evaluated to gain a better understanding why TFT beats the other machine learning methods by such a huge margin. The results for the spring time frame can be seen in Fig. 2. The results of this time frame early in the year are quite similar to the overall results of the evaluation. TFT performs best on every net, but while the difference is only around two percent points better for network one the error is around 4 percent points better for network two and 8 percent points better for network three. The MAPE for all methods is higher in the spring than in the overall year. The reason is that spring and fall is more difficult to predict because the load profile does not behave as similar as it does in the summer and winter months.

Comparing the difference between the overall MAPE and the spring MAPE for each of the machine learning methods shows that the difference in percent points is quite similar across all methods. This indicates that TFT is not per definition better in predicting spring times. The big improvement is only so huge because the predictions of TFT are overall better.



Figure 2: MAPE of the spring time frame.

Fig. 3 shows the different results for the summer time frame. Errors in the summer time frame are generally lower and closer to the overall year MAPE than the results of the spring time frame. Again TFT is the best machine learning method evaluated for all three district heating networks. Comparing the summer results with the spring results, district heating network three stands out. While in network one and two the MAPE is much lower in summer than in spring, in district heating network three the reduction of the error is not that significant. TFT even performs worse in summer than in spring on network

three. All three district heating networks are located in the same city, so the weather during this period was the same for all three networks and should not have been more unpredictable for network three. This leaves two possibilities for why TFT is worse in summer than in spring. Either the network changed which leads to the historic data being less optimal for a prediction of the current form of the district heating network, or the reported load values during the prediction time frame were erroneous, which lead to bad input values for the prediction. The load data during the prediction time frame did not have any issues, so the first problem probably caused this result.

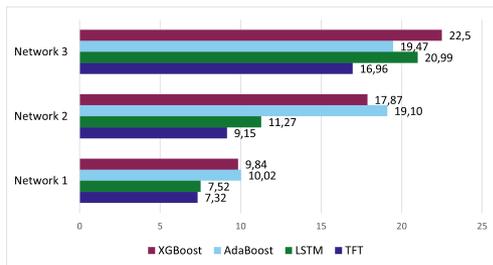


Figure 3: MAPE of the summer time frame.

Fig. 4 shows the results of all machine learning methods for the fall time frame. Similar to the spring time frame TFT really performs much better than its alternatives during this period. For network two and three the error is reduced by almost 50% over the second best method and even on the already well performing network one the error is reduced by around 30%. Together with the results of Fig. 2 the assumption of TFT being especially good for less predictable time frames can be proved. This is a very good trait for a machine learning method used for thermal load forecasting, because some of the most important features are weather data and weather forecasts. Even if a forecast is very accurate, the nature of a forecast is, that it is never a safely known value. This makes a very adaptable system like TFT preferable. If compared with the spring time frame also all MAPEs are a bit better. In both time frames the thermal load provided by the district heating networks is quite similar, which again would indicate that the spring time frame of 2021 was less predictable in its behaviour compared to the fall time frame. Moreover when compared to the summer time frame it seems like the predictions in fall would be better than in the summer. This is a misconception created by the MAPE having a percentage as error unit. In summer the thermal load is far lower in all of the evaluated district heating networks. This results in larger MAPEs even for small absolute errors. Considering the absolute errors, the predictions for the summer are actually better than for fall, e.g. with TFT for network one the MAE for summer is 0.22MW and for fall it is 0.49MW. So the summer predictions are actually better even if the MAPE is worse.

Lastly the winter time frame is considered. The results can be seen in Fig. 5. Again TFT is the strongest method for each of the three inspected networks. For network two and three TFT beats XGBoost by about 10 percent points. Furthermore, TFT also is a notable improvement in network one even though the error reduction is just 2.3 percent points. Moreover the MAPEs in the winter time frame is the best of the whole year. This has two reasons. First the prediction is overall very good. 5.52% MAPE resembles a MAE of 0.78MW which is fairly low

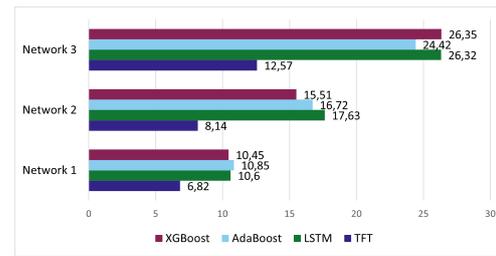


Figure 4: MAPE of the fall time frame.

in the context of the evaluated district heating network. The second fact is again one of the properties of the MAPE. In the winter months the thermal load is quite high, which leads to lower MAPEs even on similar large absolute errors.

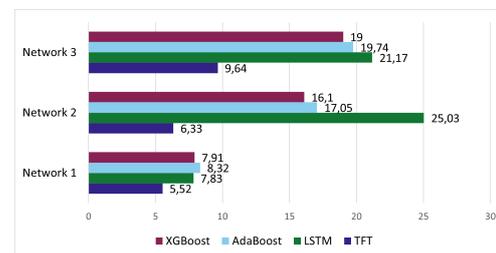


Figure 5: MAPE of the winter time frame.

For more insight on the actual absolute error, Tab. 3 shows all predicted MAEs for TFT. A good indicator for the performance of the TFT is that most errors are below 1MW. Only in the very unpredictable network three the absolute error lies on average above 1MW. The networks considered have thermal loads in the range of 10 to 20 MW for most of the year, so the calculated error is very little. The absolute errors of summer also show the issue of comparing the MAPEs of different time frames. The MAEs are the lowest for every network in the summer time frame but still the summer MAPEs were the worst.

Table 3: Absolute errors of TFT in MW.

Timeframe	Network 1	Network 2	Network 3
Overall	0.57	0.56	1.00
Spring	0.79	0.96	1.28
Summer	0.22	0.19	0.36
Fall	0.49	0.45	0.92
Winter	0.78	0.67	1.44

Evaluation of all time frames has shown that out of the list of evaluated methods, TFT is the best machine learning method for the use case of time series forecasting in any scenario. Further comparing the MAPEs in different time frames showed that TFT is not really better in predicting any of the time frames as it also struggled with predicting spring and fall more than summer and winter. However TFT predictions were overall always way better than their competitors. This results in overall lower errors in all time frames which leads to very competitive overall errors because the reduction of the error in spring and fall also reduces the overall error far more than improvements in the summer and winter forecasts. A low overall error translates to a more robust system.

7. Conclusion-and-Future-Work

Use of TFT could contribute to further improve thermal load forecasting. This paper presents first results of benchmarking TFT against different machine learning-based forecasting approaches for district heating and cooling networks. As a starting point, the predictions of these different methods were analysed on multiple time frames and over multiple district heating networks. All measured data, including statistically optimized point weather forecasts, were automatically pre-processed prior to the actual training and validation steps. The models predicted 72 hours in advance. The predictions were benchmarked against three other machine learning methods that were evaluated in previous works. TFT showed to have better MAEs and MAPEs over all experiments, making it a very strong candidate for thermal load forecasting in any scenario. Especially the improvements in spring and fall forecasts above other methods is a big improvement.

This paper used one specific model for each of the networks to predict the thermal loads. One of the advantages of TFT is to be able to train one model for the prediction of multiple time series. As a next step it should be investigated if a model trained to predict the thermal load of multiple networks still holds the same results. This could be a very important step in making machine learning for thermal load forecasting more viable, because it would reduce training effort and cost for the power plants immensely.

Acknowledgments

This work was funded by the Federal Ministry of Economics and Energy of Germany under the funding code 03EN3017, with additional financial and in-kind support by the project partners AGFW, Fernwärme Ulm GmbH and ZAK Energie GmbH. The responsibility for the content of this publication lies with the authors.

References

- Armstrong, J. and Collopy, F. (1992), 'Error measures for generalizing about forecasting methods: Empirical comparisons', *International Journal of Forecasting* **8**(1), 69–80. doi: [https://doi.org/10.1016/0169-2070\(92\)90008-W](https://doi.org/10.1016/0169-2070(92)90008-W).
- Benalcazar, P. and Kamiński, J. (2019), 'Short-term heat load forecasting in district heating systems using artificial neural networks', *IOP Conference Series: Earth and Environmental Science* **214**, 012023. doi: [10.1088/1755-1315/214/1/012023](https://doi.org/10.1088/1755-1315/214/1/012023).
- Clark, M. (2013), 'A comparison of correlation measures', *Center for Social Research, University of Notre Dame* **4**.
- Fang, T. and Lahdelma, R. (2016), 'Evaluation of a multiple linear regression model and sarima model in forecasting heat demand for district heating system', *Applied Energy* **179**, 544–552. doi: <https://doi.org/10.1016/j.apenergy.2016.06.133>.
- Freund, Y., Schapire, R. E. et al. (1996), Experiments with a new boosting algorithm, in 'icml', Vol. 96, Citeseer, pp. 148–156.
- Friedman, J. (2000), 'Greedy function approximation: A gradient boosting machine', *The Annals of Statistics* **29**. doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451).
- Guyon, I., Weston, J., Barnhill, S. and Vapnik, V. (2002), 'Gene selection for cancer classification using support vector machines', *Machine Learning* **46**, 389–422. doi: [10.1023/A:1012487302797](https://doi.org/10.1023/A:1012487302797).
- Hochreiter, S. and Schmidhuber, J. (1997), 'Long Short-Term Memory', *Neural Computation* **9**(8), 1735–1780. doi: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
URL: <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hu, X. (2021), Stock price prediction based on temporal fusion transformer, in '2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)', pp. 60–66. doi: [10.1109/MLBDBI54094.2021.00019](https://doi.org/10.1109/MLBDBI54094.2021.00019).
- Jaeger, H. (2001), 'The "echo state" approach to analysing and training recurrent neural networks—with an erratum note', *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report* **148**.
- Leiprecht, S., Behrens, F., Faber, T. and Finkenrath, M. (2021), 'A comprehensive thermal load forecasting analysis based on machine learning algorithms', *Energy Reports* **7**, 319–326. The 17th International Symposium on District Heating and Cooling. doi: <https://doi.org/10.1016/j.egypr.2021.08.140>.
- Lim, B., Arik, S. Ö., Loeff, N. and Pfister, T. (2021), 'Temporal fusion transformers for interpretable multi-horizon time series forecasting', *International Journal of Forecasting* **37**(4), 1748–1764. doi: <https://doi.org/10.1016/j.ijforecast.2021.03.012>.
- Lim, B. and Zohren, S. (2021), 'Time-series forecasting with deep learning: a survey', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **379**(2194), 20200209. doi: [10.1098/rsta.2020.0209](https://doi.org/10.1098/rsta.2020.0209).
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J. and Chintala, S. (2019), Pytorch: An imperative style, high-performance deep learning library, in 'Advances in Neural Information Processing Systems 32', Curran Associates, Inc., pp. 8024–8035.
URL: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al. (2011), 'Scikit-learn: Machine learning in python', *Journal of machine learning research* **12**(Oct), 2825–2830.
- Phetrittikun, R., Suvirat, K., Pattalung, T. N., Kongkamol, C., Ingviya, T. and Chaichulee, S. (2021), Temporal fusion transformer for forecasting vital sign trajectories in intensive care patients, in '2021 13th Biomedical Engineering International Conference (BMEiCON)', pp. 1–5. doi: [10.1109/BMEiCON53485.2021.9745215](https://doi.org/10.1109/BMEiCON53485.2021.9745215).
- Saloux, E. and Candanedo, J. A. (2018), 'Forecasting district heating demand using machine learning algorithms', *Energy Procedia* **149**, 59–68. 16th International Symposium on District Heating and Cooling, DHC2018, 9–12 September 2018, Hamburg, Germany. doi: <https://doi.org/10.1016/j.egypro.2018.08.169>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I. (2017), 'Attention is all you need', *CoRR* **abs/1706.03762**.
URL: <http://arxiv.org/abs/1706.03762>
- Werner, S. (2017), 'International review of district heating and cooling', *Energy* **137**, 617–631. doi: <https://doi.org/10.1016/j.energy.2017.04.045>.
- Willmott, C. and Matsuura, K. (2005), 'Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance', *Climate Research* **30**, 79. doi: [10.3354/cr030079](https://doi.org/10.3354/cr030079).
- Wu, B., Wang, L. and Zeng, Y.-R. (2022), 'Interpretable wind speed prediction with multivariate time series and temporal fusion transformers', *Energy* p. 123990. doi: <https://doi.org/10.1016/j.energy.2022.123990>.
- Xue, P., Jiang, Y., Zhou, Z., Chen, X., Fang, X. and Liu, J. (2019), 'Multi-step ahead forecasting of heat load in district heating systems using machine learning algorithms', *Energy* **188**, 116085. doi: <https://doi.org/10.1016/j.energy.2019.116085>.

Offshore CO₂ Capture from gas turbine with low investment optimized using Aspen HYSYS

Lars Erik Øi, Fatemeh Fazli, Rajan Thapa

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway

Lars.oi@usn.no

Abstract

CO₂ capture from gas turbine exhaust gas is a possibility for CO₂ emission reduction on oil and gas production platforms. A standard process is based on absorption in monoethanol amine (MEA). A challenge is that the cost of size and weight for the process equipment is higher than on a land-based process. A standard process based on CO₂ absorption into (MEA) is simulated in Aspen HYSYS™. The equipment cost was obtained from Aspen In-plant Cost Estimator™. The base case is based on assumptions which are in earlier works assumed to be close to optimum for a land-based process with a heat consumption of 3.5 MJ/kg removed CO₂. Different parameters as the number of stages in the absorption column and the minimum temperature approach are varied in the direction expected to be more optimum for an offshore application. It is expected that a lower absorption column and smaller heat exchangers are more optimum offshore even though the heat consumption will increase. Parametric studies were performed at 90 % capture efficiency. Suggested conditions for an offshore application with 87 % capture efficiency are 13 m absorber packing height and 15°C minimum approach temperature due to a decrease in equipment cost, size and weight. This is expected to balance the increase of heat consumption to approximately 5.5 MJ/kg CO₂ removed.

Keywords: *Carbon capture, Aspen HYSYS, gas turbine, offshore, simulation.*

1. Introduction

There are published numerous articles on CO₂ capture in general (Liang et. al, 2015; Li et al., 2016). There are also many articles about cost estimation of CO₂ capture (Rao and Rubin, 2002; Ali, 2019). Some include power production and CO₂ capture in their studies (de Ruick, 1992; Schach et al., 2010).

Research work on the combination of simulation, cost estimation and cost optimization have been performed by Kallevik (2010), Øi (2012) and Shirdel et al. (2022). In a PhD Thesis by Ali (2019) the EDF (Enhanced Detailed Factor) method was presented. Øi et al. (2021) evaluated automated calculation of cost optimum process parameters in the CO₂ capture process. Typical parameters to optimize in an amine-based process is the number of stages in the absorption column and minimum temperature approach in the heat exchangers.

In this study, there was a big challenge to find open literature on process simulation of CO₂ capture from a gas turbine offshore application. There are very few literature references that contain both the CO₂ capture process from offshore gas turbines in a combined cycle by process simulation.

In order to dispose and capture CO₂ De Ruyck et al. (1992) proposed a combined cycle with higher efficiency for a more economical CO₂ capture process. They concluded that the combined steam-

CO₂ gas turbine cycle was viable with no major issues in practical manner.

Bjerve and Bolland (1994) assessed six alternatives in power generation with the purpose of reducing CO₂ emission from exhaust gas released from a gas turbine in an offshore natural gas combustion. They concluded that the option with exhaust gas recycling in the combined cycle, was the best alternative in the CO₂ capture process, it resulted in the highest efficiency of CO₂ removal, and second lightest option. Falk-Pedersen et al. (1995) presented a concept of CO₂ removal from an off-shore gas turbine based on absorption into monoethanol amine (MEA). Falk-Pedersen and Dannström (1987) suggested an off-shore process based on absorption into an amine solution through a membrane.

Flatebø (2012) evaluated several offshore combined cycle configurations to meet the power need for the offshore installation and accordingly reduce the cost. The tools Aspen Plus, GTPRO and GTMASTER were used. One of the case studies focused on the design of the offshore installation, and another assigned to obtain high efficiency. Liu and Karimi (2019) used the tool GateCycle for similar analysis. Nord et al. (2017), studied gas turbines as a power production resource in an offshore oil and gas installation by increasing the level of CO₂ emission. To supply energy demand in the reboiler, a heat resource is required. To compensate the energy

demand entirely or partially, they suggested an additional system of a compact steam bottoming cycle that can supply power need included in the CO₂ capture system. Through his study, a weight evaluation for main equipment has been performed to investigate a simple relation between the outlet gas turbine mass flow and the size of the oil and gas installation.

There are very few studies on offshore CO₂ capture analysis and almost none about optimizing the suggested solutions with parameter values. In this work, similar procedure of designing, simulating, and cost estimation for the CO₂ capture process have been performed and more consideration regarding the offshore application have been considered. This work presents results from the Master Thesis work by Fatemeh Fazli (2022).

2. Process Description and Specifications

2.1 Process description

Figure 1 shows the process before the CO₂ capturing process. This combined process cycle is assigned to meet the platform's need for power energy.

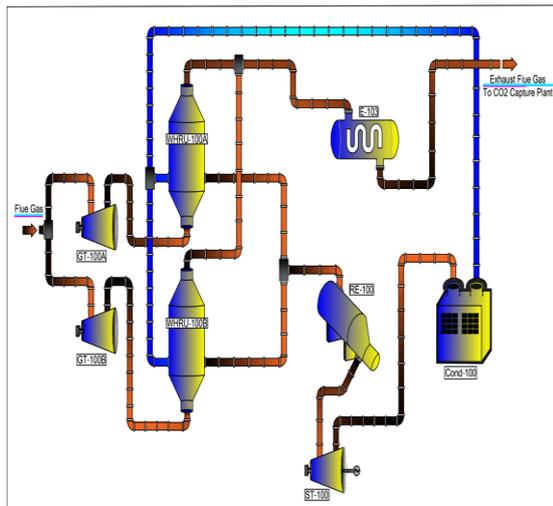


Figure 1: Upstream process before CO₂ capture.

This combination is a set of heat engines collaborating sequentially together with the same heat source and for a typical offshore project is included a set of gas turbine, a WHRU unit, and a generator followed by a condenser and a pump. The flue gas enters the two parallel gas turbines which are assigned to produce power. Then heat will be recovered in a waste heat recovery unit (WHRU) to supply additional heat and generate more energy. One of the exhaust gases from the WHRU unit will be sent to the reboiler and one to the steam turbine to generate electricity.

Figure 2 shows the CO₂ capture process for the defined project. The exhaust gas from the WHRU unit is entering the absorption column and the pressure and temperature of the flue gas need to be

adjusted. The exhaust gas will be led to an absorber where the CO₂ will be removed from the gas with the help of a mixture of MEA and water injected into the absorber from the top of the column. The dissolved CO₂ then exit the absorber column from the bottom and is sent to the rich amine pump to pressurize for further process in the desorber unit. Meanwhile, the clean gas will exit the column from the top. To regenerate the MEA and separate the CO₂ captured from the flue gas in the absorber unit, the gas flows to the desorber unit where this is done with the help of a reboiler and a condenser. The temperature should be increased to fulfill the reboiler requirement. The required energy to heat up the gas is supplied by the outlet of the reboiler. As the regeneration process in the desorber column is done, CO₂ removed from the gas is obtained from the top of the stripper for further storage processes. The regenerated amine will exit the stripper from the bottom and will be sent back to the cycle where the MEA is injected into the absorber column.

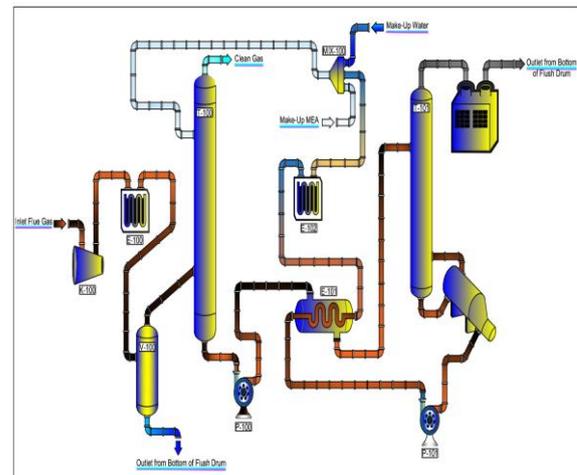


Figure 2: CO₂ capture process.

2.2 Process Specifications and simulation

The process specifications used for the base case simulation are presented in Table 1. The process simulation in this work applies the same strategy used in Øi (2007) and Aromada et al. (2015). The simulations were conducted using the equilibrium-based Aspen HYSYS Version 10.

The base case was simulated to capture 90 % CO₂ from exhaust gas from a natural gas combined cycle (NGCC) power plant. The process consists of an absorber with 16 packing stages (16 m), a desorber with 6 packing stages (6 m), and 10 °C temperature difference in the main heat exchanger.

The Aspen HYSYS simulation process flow diagram showing all the equipment included in the scope of the study is shown in Figure 3.

Table 1: Specifications for process simulation.

Parameter	Value	Unit
Inlet flue gas temperature	40	°C
Inlet flue gas pressure	101	kPa
Inlet flue gas flow rate	1.0×10^5	kgmol/h
CO ₂ content in inlet gas	3.73	mol %
Water content in inlet gas	6.71	mol %
Lean amine temperature before and after pump	120	°C
Lean amine pressure before pump	200	kPa
Lean amine pressure after pump	300	mol %
Lean amine pressure to absorber	101	kPa
Lean amine rate to absorber	25945	kgmol/h
CO ₂ content in lean amine	2.942	mol %
Number of stages in absorber (base case)	16	-
Rich amine pressure before pump	111	mol %
Rich amine pressure after pump	200	mol %
Number of stages of stripper	6 + Reboiler + Condenser	-
Reboiler temperature	120	°C

2.3 Equipment Sizing

To determine the packing height, a constant stage (Murphree) efficiency corresponding to 1 meter of packing was assumed. Murphree efficiencies of 0.15 and 0.5 were specified for the absorber and the

desorber respectively (Table 1). For the absorber and desorber internals, structured packing was assumed. The absorption column diameter was calculated based on a gas velocity of 2.0 m/s and the desorption column based on a gas velocity of 1 m/s as in Park and Øi (2017) and Øi et al. (2021). The total height of the absorption column and desorption column is specified to be packing height plus 34 m (Kallevik, 2021). The extra height is due to distributors, water wash packing, demister, gas inlet, outlet, and sump. Centrifugal pumps with 75 % adiabatic efficiency were used in the process simulations. Overall heat transfer coefficient values have been specified for the lean/rich heat exchanger to 500 W/(m²K). These values are close to the same as in Øi (2012) and Park and Øi (2017) and slightly less than the numbers in Øi et al. (2021) which are regarded as optimistic.

2.4 Capital and Operating Cost Estimation

The equipment costs were calculated in Aspen In-plant Cost Estimator (v.12), which gives the cost in Euro (€) for Year 2020. A generic location (e.g. Rotterdam) was assumed. Stainless steel (SS316) with a material factor of 1.3 was assumed for all equipment units.

In the detailed factor method and the EDF method, each equipment cost (in carbon steel) was multiplied with its individual installation factor to get equipment installed cost. The detailed installation factor is a function of the site, equipment type, materials, size of equipment and includes direct costs for erection, instruments, civil, piping, electrical, insulation, steel and concrete, engineering cost, administration cost, commissioning, and contingency. The updated installation factors for 2020 (Aromada, 2021) were used. Specifications for operating cost estimation are found in Table 2. More details can be found in Fazli (2022).

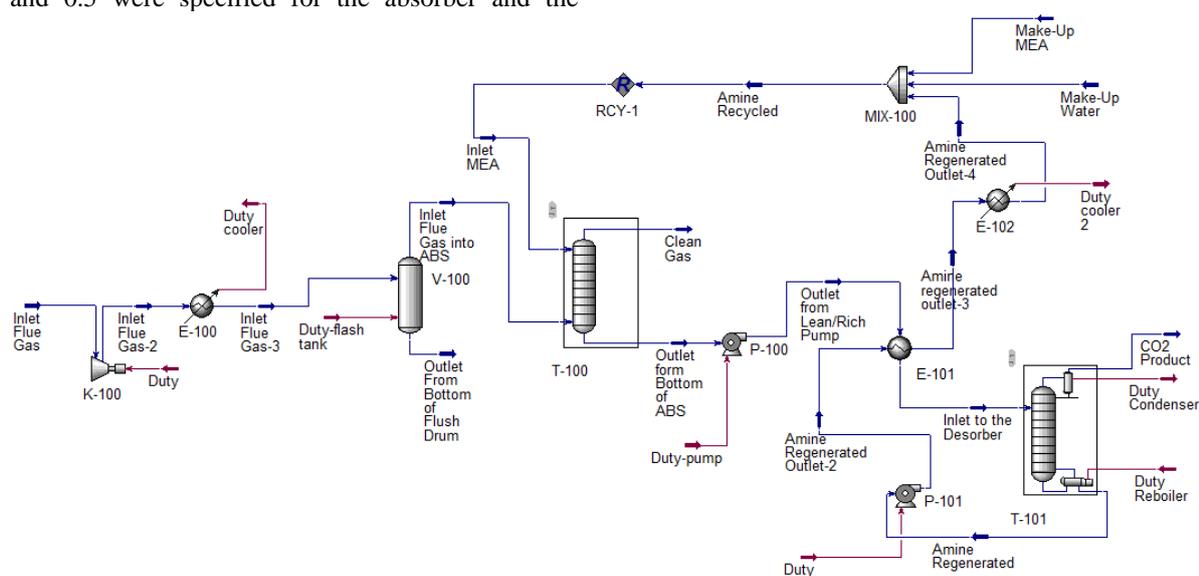
Figure 3: CO₂ capture process model in Aspen HYSYS.

Table 2: Cost calculation specifications.

Parameter	Value
Plant lifetime	25 years
Discount rate	8%
Maintenance cost	3 % of installed cost
Electricity price	0.078 Euro/kWh
Steam price	0.032 Euro/kWh
Annual operational time	8000 hours
Currency exchange rate 2021	9.78
Cost index 2020	301
Cost index September 2021	317

2.5 Annual Cost and Capture Cost

A cost optimization can be based on minimization of different cost measures. A common measure is the CO₂ capture cost defined by Equation 1 for a defined process plant and a defined time of operation.

$$CO_2 \text{ capture cost} = \frac{\text{Total annual cost}}{\text{Mass of } CO_2 \text{ Captured}} \quad (1)$$

In this work the annual capital cost is calculated and is added to the yearly operating cost to obtain the total annual cost. The annual capital cost is obtained as by Equation 2:

$$\text{Annual capital cost} = \frac{\text{capital cost}}{\text{Annualized factor}} \quad (2)$$

The annualized factor is calculated by Equation 3 which is based on a constant interest rate (Aromada et. al, 2021).

$$\text{Annualised factor} = \sum_{i=1}^n \left[\frac{1}{(1+r)^i} \right] \quad (3)$$

where n is the years of operation and r is the interest rate.

3 Results and Discussion

3.1 Simulation Results

Table 3 presents the process simulation results for the base case and parametric optimization.

Table 3: Main simulation results.

	Reboiler heat [MJ/kg CO ₂]	Optimum parameter
Base case	3.58	-
Base case packing height	-	16 meter
Base case minimum temperature difference	-	10°C

The reboiler specific heat consumption in this work is 3.58 MJ/kg CO₂. This is close to the 3.7 MJ/kg CO₂ calculated by (Øi, 2007) for a similar process with 85 % CO₂ capture.

3.2 Sensitivity Analysis

In Figure 4, the cross-section area in the absorber is simulated and calculated as a function of number of stages and minimum temperature approach (ΔT_{MIN}).

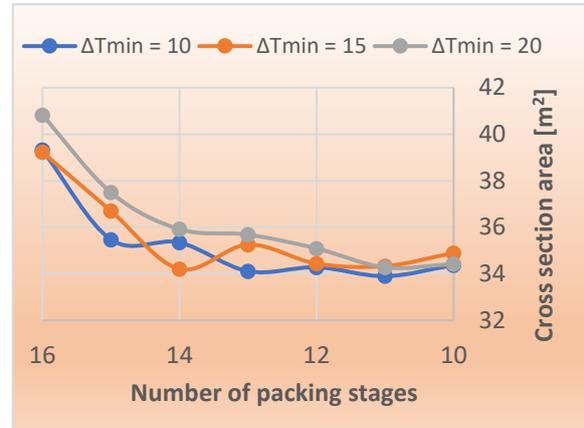


Figure 4: Absorber cross section area per no. of stages.

The figure shows that the absorber cross section increases slightly with the number of packing stages and increases slightly with increasing minimum temperature difference.

In Figure 5, the heat exchanger area in the main lean/rich amine heat exchanger absorber is simulated and calculated as a function of number of stages.

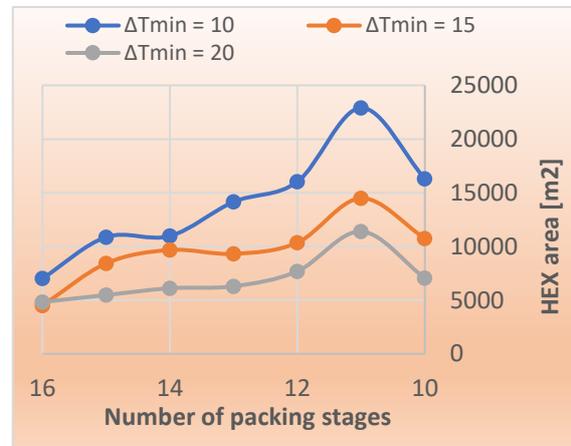


Figure 5: HEX area per number of stages area.

The figure shows that the heat exchanger area decreases with the number of packing stages and increases with the minimum heat exchanger area. This means that a low column and a high temperature difference (ΔT_{MIN}) give the lowest heat exchanger area.

In Figure 6, the reboiler duty in the desorber is simulated and calculated as a function of number of stages.

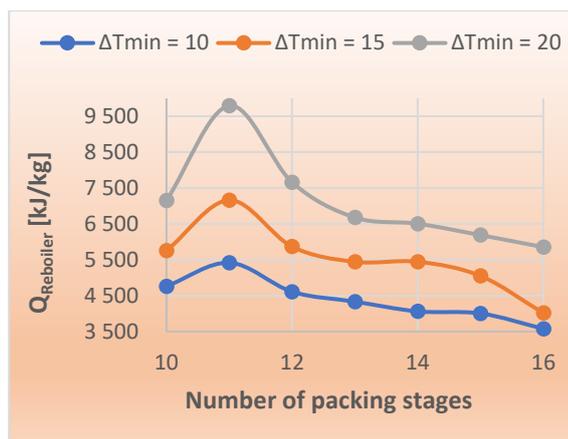


Figure 6: Q_{Reboiler} [kJ/kg CO_2 removed] dependency on number of stages.

The figure shows that the heat consumption decreases considerably with the number of packing stages and with the minimum temperature difference (ΔT_{MIN}). This means that some packing height and a reasonably low minimum temperature approach is necessary to avoid a too high heat consumption. Table 4 shows a summary of the sensitivity analysis results.

A lower column than 12 meter (12 packing stages) is probably not optimum due to high energy consumption. A higher column than 15 meter gives only a slight reduction in energy consumption. Due to this, a 13 meter packing height is suggested for an offshore

application. Compared to $\Delta T_{\text{MIN}} = 15$ °C, a reduction to 10 °C increases the heat exchanger area substantially while the heat consumption decreases only slightly. When the ΔT_{MIN} increases to 20 °C, the heat consumption increases substantially. Because of this, $\Delta T_{\text{MIN}} = 15$ °C is suggested as a close to optimum value. These values can not be calculated as exact optimums, because accurate cost data are not available.

3.2 General discussion

In a traditional land-based plant, the accuracy of a capital cost and operating cost estimate based on an early phase study is in order of magnitude 30-50 %. The uncertainty in a cost estimate for an offshore installation is much higher. One reason for this is that there are more unknown factors in such an estimate. The uncertainty of purchased equipment cost and energy consumption has about the same accuracy for land-based and offshore installations. But the additional factors have much higher uncertainty for an offshore plant. If there are no space and weight limitations, capital cost estimates based on land-based methods can be reasonable. But with space and weight limitations, equipment cost factors can be an order of magnitude higher for an offshore installation. Because of this, cost estimates for an offshore plant in this study can only be order of magnitude estimates.

Table 4: Summary of the results.

	cost optimum packing height	cost optimum temperature difference	90% route (1)	92% route (2)	Combined optimum parameters
Capital cost (million €)	167.7	171.7	192.2	190.6	174.9
Annualized capital cost (million €/yr)	16.5	16.8	18.9	18.7	17.2
Annual operating cost (million €/yr)	52.5	51.9	76.6	60.2	50.8
Total annual cost (million €/yr)	69.0	68.7	95.5	78.9	67.9
CO_2 capture cost (k€/t CO_2)	63.9	63.8	84.5	67.3	62.9
Specific reboiler heat (MJ/kg CO_2)	3.50	3.41	5.24	3.55	3.33
Annual cost savings (%)	-2	-2	29	3	-4
Energy savings (%)	-7	-10	39	-6	-12

4 Conclusion

A challenge for an offshore based CO_2 capture plant is that the cost of size and weight for the process equipment is higher and more uncertain than on a land-based process. Possibilities to reduce the size and cost are evaluated using simulations in an equilibrium-based model in Aspen HYSYSTM. The effects of reducing the size and weight of the absorber and the amine/amine heat exchanger and especially the effect on heat consumption are calculated and evaluated.

A standard process based on CO_2 absorption into mono ethanolamine (MEA) is simulated. The base case is based on assumptions which are in earlier works assumed to be close to optimum for a land-based process with a heat consumption of 3.5 MJ/kg removed CO_2 . In this work, different parameters as the number of stages in the absorption column and the minimum temperature approach are varied in the direction expected to be more optimum for an offshore application. It is expected that a lower absorption column and smaller heat exchangers are more optimum offshore even though the heat consumption will increase.

Suggested conditions for an offshore application with 87 % capture efficiency are 13 m absorber packing height and 15°C minimum approach temperature due to a decrease in equipment cost, size and weight. This is expected to balance the increase of heat consumption to approximately 5.5 MJ/kg CO₂ removed.

References

- Ali, H. (2019) 'Techno-economic analysis of CO₂ capture concepts', PhD Thesis, University of South-Eastern Norway. [Online]. Available: <http://hdl.handle.net/11250/2622802>
- Aromada, S. A. et al. (2021) 'Capital cost estimation of CO₂ capture plant using Enhanced Detailed Factor (EDF) method: Installation factors and plant construction characteristic factors', *Int. J. Greenh. Gas Control*, vol. 110, no. 1, p. 103394, doi: 10.1016/j.ijggc.2021.103394.
- Bjerve, Y. and Bolland, O. (1994) 'Assessment of Power Generation Concepts on Oil Platforms in Conjunction with CO₂ Removal', *Proc. ASME Int. Gas Turbine Aeroengine Congr. Expo.*, vol. 3, [Online]. Available: <https://doi.org/10.1115/94-GT-378>
- De Ruyck, J. (1992) 'Efficient CO₂ capture through a combined steam and CO₂ gas turbine cycle', *Energy Convers. Manag.*, vol. 33, no. 5–8, pp. 397–403, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/019689049290036V>
- Falk-Pedersen, O. et al. (1995) 'Separation of carbon dioxide from offshore gas turbine exhaust', *Energy Convers. Manag.*, vol. 36, no. 6–9, pp. 393–396, [Online]. Available: [https://doi.org/10.1016/0196-8904\(95\)00028-C](https://doi.org/10.1016/0196-8904(95)00028-C)
- Falk-Pedersen, O. and Dannström, H. (1997) 'Separation of carbon dioxide from offshore gas turbine exhaust', *Energy Convers. Manag.*, vol. 38, pp. S81–S86.
- Fazli, F. (2022) 'Process design of CO₂ capture from the gas turbine at an oil platform.' Master's thesis. Telemark University College, Norway.
- Flatebø, Ø. (2012) 'Off-design Simulations of Offshore Combined Cycles', Norwegian University of Science and Technology. [Online]. Available: <http://hdl.handle.net/11250/234879>
- Kallevik, O. B. (2010) 'Cost estimation of CO₂ removal in HYSYS.' Master's thesis. Telemark University College, Norway.
- Kvamsdal, H. M. et al. (2007) 'A quantitative comparison of gas turbine cycles with CO₂ capture', *Energy*, vol. 32, no. 1, pp. 10–24, doi: 10.1016/j.energy.2006.02.006.
- Li, K. et al. (2016) 'Systematic study of aqueous monoethanolamine (MEA)-based CO₂ capture process: Techno-economic assessment of the MEA process and its improvements. *Applied Energy*, 165, 648–659.
- Liang, Z. et al. (2015) Recent progress and new developments in post-combustion carbon-capture technology with amine-based solvents. *International Journal of Greenhouse Gas Control*, 40, 26–54. doi: 10.1016/j.ijggc.2015.06.017.
- Nord, L. O. R. et al. (2017) 'CCS on Offshore Oil and Gas Installation - Design of Post-Combustion Capture System and Steam Cycle', *Energy Procedia*, vol. 114, no. 7, pp. 6650–6659, Jul. 2017, doi: 10.1016/j.egypro.03.1819.
- Rao, A. B., and Rubin, E. S. (2002) 'A Technical, Economic, and Environmental Assessment of Amine-Based CO₂ Capture Technology for Power Plant Greenhouse Gas Control', *Environ. Sci. Technol.*, vol. 36, no. 20, pp. 4467–4475, doi: 10.1021/es0158861.

- Schach, M. O. et al. (2010) 'Techno-economic analysis of post combustion processes for the capture of carbon dioxide from power plant flue gas', *Industrial & Engineering Chemistry Research*, 49(5), 2363–2370,
- Shirdel S. et al., (2022) 'Sensitivity Analysis and Cost Estimation of a CO₂ Capture Plant in Aspen HYSYS', *ChemEngineering*, vol. 6, no. 2, p. 28, doi: 10.3390/chemengineering6020028.
- Øi, L. E. (2012) 'Removal of CO₂ from Exhaust Gas', Ph.D. Thesis, Department of Process, Energy and Environmental Technology, Telemark University College, Porsgrunn, Norway.

Development of central-upwind-WENO scheme for two-phase 1-D drift-flux model

Prasanna Welahettige¹, Christian Berg², Bernt Lie¹

¹*Department of Electrical Engineering, Information Technology and Cybernetics, University of South-Eastern Norway, Porsgrunn 3918, Norway.*

²*Kelda Drilling Controls, Porsgrunn, 3933, Norway.*
prasanna.welahettige@usn.no

Abstract

The two-phase drift flux model is extensively used in multiphase flow applications. In this study, we focus on possible numerical schemes for solving the drift flux model. Due to the complexity of the primitive equations and empirical parameters, it is challenging to achieve stability of the numerical scheme used for the drift flux model. The high resolution second order central scheme, the high resolution second order central-upwind scheme, and the high resolution third order and fifth order weighted essentially non-oscillatory schemes (WENO) were successfully implemented for the drift flux model. The schemes were tested with the shock tube discontinuity problem. The central-upwind-WENO scheme was developed and applied to the drift flux model. In the central-upwind-WENO scheme, the cell interface values were taken from the WENO reconstruction, and the monotone flux is calculated from the central-upwind flux. The central-upwind-WENO scheme can achieve higher order accuracy than the central-upwind scheme by using the same stencils which are used for the central-upwind scheme. The central-upwind-WENO scheme gives more accurate results than the central scheme, central-upwind scheme and the WENO scheme. Especially at the rarefaction and shock wave fronts, the central-upwind-WENO scheme gives sharper gradients compared to the other schemes. Instead of a limiter function, the central-upwind-WENO scheme uses a smoothness indicator. All the schemes used in the study are suitable for two-phase drift flux model simulation.

1. Introduction

Two-phase models are important in the oil and gas production process and in well drilling operation due to the multiphase properties of the fluid. The two-fluid model and the mixture model have been used for over two decades in the oil and gas industry. The two-fluid model treats each phase with a set of conservation equations [1]. The phase interaction appears as source terms in the conservation equations in the two-fluid model. The number of transport equations can be reduced by correlating the relative velocity between the phases, the slip velocity, with the flow variables [2]. The drift-flux model can be derived by adding together the momentum equations from both phases in the two-fluid model [3]. The drift-flux model consists of two mass conservation equations and a mixture momentum equation. The drift flux model uses a single momentum equation; therefore, it might give weak results for the phasic velocities compared to the two-fluid model. The drift flux model requires a number of empirical parameters. Those are the drawback of the drift flux model.

2. The two-phase 1-D drift flux model

Equations 1 and 2 give the mass balance for the liquid and gas phases, respectively. A single momentum equation for the liquid and gas mixture

in Equation 3 gives the momentum balance for the mixture,

$$\frac{\partial(\rho_l \alpha_l)}{\partial t} + \frac{\partial(\rho_l \alpha_l v_l)}{\partial x} = 0, \quad (1)$$

$$\frac{\partial(\rho_g \alpha_g)}{\partial t} + \frac{\partial(\rho_g \alpha_g v_g)}{\partial x} = 0, \quad (2)$$

$$\frac{\partial(\rho_l \alpha_l v_l + \rho_g \alpha_g v_g)}{\partial t} + \frac{\partial(\rho_l \alpha_l v_l^2 + \rho_g \alpha_g v_g^2 + p)}{\partial x} = 0. \quad (3)$$

Here, subscripts l and g denote the liquid and gas phases, ρ is the density, α is the phase volume fraction, v is the velocity. In this study, we have not considered source terms. The drift-flux model assumes that the phases are mechanically at equilibrium, such that gas and liquid phases have same pressure, $p(x, t) = p_l(x, t) = p_g(x, t)$. The speed of sound is comparatively larger than the fluid velocity, therefore the fluids are assumed to be weakly compressible [4]. However, the phases are compressible, in other words, the densities of the phases change with time and space. The system is assumed to be isothermal; hence, the energy equation is neglected. The phases are immiscible, and there is no phase transfer between them. The pipe cross sectional area is assumed to be constant. There is set of closure laws related to the two-phase drift-flux model. There is no empty space and the total volume fraction is unity,

$$\alpha_l + \alpha_g = 1. \quad (4)$$

The closure law for the liquid density is a function of pressure,

$$\rho_l = \rho_{l,0} + \frac{1}{c_l^2}(p - p_{l,0}), \quad (5)$$

where $\rho_{l,0} = 1000 \text{ kg/m}^3$ and $p_{l,0} = 1 \text{ bar}$ are constants related to a reference point, $c_l = 1000 \text{ m/s}$ is the sound velocity in the liquid phase. The closure law for the gas density as a function of pressure is

$$\rho_g = \frac{1}{c_g^2} p, \quad (6)$$

where $c_g = 316 \text{ m/s}$ is the sound velocity in the gas phase. The slip velocity is the relative velocity between the phases, and can be defined as a function of flow variables, $v_g - v_l = \Phi(\alpha_g, p, v_g)$ [2]. The gas velocity can be presented in term of mixture velocity, $v_m = \alpha_l v_l + \alpha_g v_g$, and gas drift velocity, v_d ,

$$v_g = k v_m + v_d. \quad (7)$$

Here, k is the profile parameter (or distribution parameter); it describes the velocity and concentration profile, and is varying from 1.0 to 1.5 in a cylindrical pipe flow [3]. For simplicity, we assume $k = 1.2$; this implies that the ratio of maximum to the average flow velocity is approximately equal to 1.2. The drift velocity is 0.216 [5]. Here, we denote Equations 4, 5, 6, and 7 by the *primitive equations*.

Equation 1, 2, and 3 can be presented in compact vector form with the conservative variables,

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = 0. \quad (8)$$

Here, \mathbf{U} is the conservative variable vector, and $\mathbf{F}(\mathbf{U})$ is the flux vector function.

$$\mathbf{U} = \begin{pmatrix} \alpha_l \rho_l \\ \alpha_g \rho_g \\ \alpha_l \rho_l v_l + \alpha_g \rho_g v_g \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad (9)$$

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} \alpha_l \rho_l v_l \\ \alpha_g \rho_g v_g \\ \alpha_l \rho_l v_l^2 + \alpha_g \rho_g v_g^2 + p \end{pmatrix} = \begin{pmatrix} u_1 v_l \\ u_2 v_g \\ u_1 v_l^2 + u_1 v_g^2 + p \end{pmatrix}, \quad (10)$$

while, u_1 , u_2 and u_3 are conservative variables. Due to the non-linearity of the drift flux model and the closure laws, it is difficult to express explicit conservative formulas for v_l , v_g and p in the flux vector function. It is possible to express the v_l , v_g and p in terms of conservative variables. By embedding the closure laws in Equations 5 and 6 in 4, a quadratic equation can be found expressing the pressure $p(u_1, u_2)$ with conservative variables, $p^2 + (-p_0 + c_l^2 \rho_{l,0} - c_l^2 u_1 - c_g^2 u_2)p + (c_g^2 p_0 - c_l^2 c_g^2 \rho_{l,0})u_2 = 0$. (11)

Equation 11 gives two real values for the pressure; however, Equation-12 is the physically realistic root,

$$p(u_1, u_2, u_3) = 0.5 \left(-(-p_0 + c_l^2 \rho_{l,0} - c_l^2 u_1 - c_g^2 u_2) + \sqrt{(-p_0 + c_l^2 \rho_{l,0} - c_l^2 u_1 - c_g^2 u_2)^2 - 4(c_g^2 p_0 - c_l^2 c_g^2 \rho_{l,0})u_2} \right) \quad (12)$$

$$\sqrt{(-p_0 + c_l^2 \rho_{l,0} - c_l^2 u_1 - c_g^2 u_2)^2 - 4(c_g^2 p_0 - c_l^2 c_g^2 \rho_{l,0})u_2} \quad (12)$$

The velocities can be derived by embedding conservative variables in Equation-7,

$$v_l(u_1, u_2, u_3) = \frac{(1-k\alpha_g)u_3 - v_d u_2}{(1-k\alpha_g)u_1 + k\alpha_l u_2}, \quad (13)$$

$$v_g(u_1, u_2, u_3) = \frac{k\alpha_l u_3 + v_d u_1}{(1-k\alpha_g)u_1 + k\alpha_l u_2}. \quad (14)$$

The fluids are assumed to be compressible. Because oil and gas reservoir pressure is comparatively large, the wells are deep, and pressure is varying along the well. Fluid densities are functions of pressure such as $\rho_l = \rho_l(u_1, u_2)$ and $\rho_g = \rho_g(u_1, u_2)$. The compact conservative form of the drift flux model, after discretization, can be presented without source terms,

$$\frac{d\mathbf{U}}{dt} = \mathbf{L}(\mathbf{U}), \quad (15)$$

Here, $\mathbf{L}(\mathbf{U})$ is the discretization of the spatial operator,

$$\mathbf{L}(\mathbf{U}) = -\frac{1}{\Delta x} \left(\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}} - \mathbf{F}(\mathbf{U})_{i-\frac{1}{2}} \right). \quad (16)$$

In most applications, the challenge is to discretize the cell interface flux function, $\mathbf{F}(\mathbf{U})_{i\pm\frac{1}{2}}$, accurately.

The flux function is the most important part of semi-discretization for partial differential equations. Monotonicity of the numerical flux, $\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}} =$

$\mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^+, \mathbf{U}_{i+\frac{1}{2}}^-)$, can be achieved by using monotonicity preserving schemes such as Godunov flux, Engquist-Osher flux, Lax-Friedrichs flux [6]. The positive and negative fluxes should be Lipschitz continuous functions, and should be consistent with the physical flux. Here, $\mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^+, \mathbf{U}_{i+\frac{1}{2}}^-)$ is non-decreasing in the first argument and nonincreasing in the second argument. For the lower order reconstructions, there can be a big difference between results obtained by different monotone fluxes. The order of the reconstruction polynomial higher than two helps to achieve less smearing effect at the discontinuities [7]. The problem is that most of the high order monotone schemes move to first order at a discontinuity.

3. Time integration for the semi-discrete schemes

Time integration is applied to the Equation-15. The TVD Runge-Kutta method is used to preserve the total variation diminishing (TVD) properties under the CFL condition [8]. A fourth order TVD Runge-Kutta scheme for time iterations is [9]

$$\mathbf{U}^{(1)} = \mathbf{U}^{(0)} + \frac{1}{2} \Delta t \mathbf{L}(\mathbf{U}), \quad (17)$$

$$\mathbf{U}^{(2)} = \frac{1}{2} \mathbf{U}^{(0)} + \frac{1}{2} \mathbf{U}^{(1)} - \frac{1}{4} \Delta t \mathbf{L}(\mathbf{U}^{(0)}) + \frac{1}{4} \Delta t \mathbf{L}(\mathbf{U}^{(1)}), \quad (18)$$

$$\mathbf{U}^{(3)} = \frac{1}{9}\mathbf{U}^{(0)} + \frac{2}{9}\mathbf{U}^{(1)} + \frac{2}{3}\mathbf{U}^{(2)} - \frac{1}{9}\Delta t\mathbf{L}(\mathbf{U}^{(0)}) - \frac{1}{3}\Delta t\mathbf{L}(\mathbf{U}^{(1)}) + \Delta t\mathbf{L}(\mathbf{U}^{(2)}), \quad (19)$$

$$\mathbf{U}^{(4)} = \frac{1}{3}\mathbf{U}^{(1)} + \frac{1}{3}\mathbf{U}^{(2)} + \frac{1}{3}\mathbf{U}^{(3)} + \frac{1}{6}\Delta t\mathbf{L}(\mathbf{U}^{(1)}) + \frac{1}{6}\Delta t\mathbf{L}(\mathbf{U}^{(3)}). \quad (20)$$

Here, $\mathbf{U}^{(0)}$ is related to time step n , $\mathbf{U}^{(4)}$ is related to the time step $n + 1$. High order TVD Runge-Kutta methods can be found in [9]. Under the CFL restriction, the time step can be defined as,

$$\Delta t \leq c \frac{\Delta x}{\max\{\lambda_i\} \forall i}, \quad (21)$$

Here c is the CFL number, λ is the eigenvalues of the Jacobian of the hyperbolic system.

4. High-resolution second order central scheme for the drift flux model

The piecewise polynomials are functions of cell averages. For the third order central schemes, they couple with a piecewise quadratic approximation, such as the essentially non-oscillatory (ENO) reconstruction [10]. The integral values are approximated by the midpoint rule. The mid-point values are predicted by Taylor expansion. For more details, see [11], [12]. The cell interface values are

$$\mathbf{U}_{i+\frac{1}{2}}^{n,+} = \mathbf{U}_{i+1}^n - \frac{\Delta x}{2}(\mathbf{U}_x)_{i+1}^n, \quad (22)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{n,-} = \mathbf{U}_i^n + \frac{\Delta x}{2}(\mathbf{U}_x)_i^n. \quad (23)$$

Here, the spatial derivatives, \mathbf{U}_x , are reconstructed from the cell averages. A scalar total variation diminishing (TVD) property is achieved from a limiter function. The family of MINMOD-like limiters are used to calculate the numerical derivative,

$$(\mathbf{U}_x)_i^n = \text{MINMOD}\left(\theta \frac{\mathbf{U}_i^n - \mathbf{U}_{i-1}^n}{\Delta x}, \theta \frac{\mathbf{U}_{i+1}^n - \mathbf{U}_i^n}{\Delta x}\right), \quad (24)$$

where, $1 \leq \theta \leq 2$ is the limiter parameter; for more details about the limiter functions, refer [13], [14]. It can be shown that for large θ values, a highly oscillatory solution is found, while for small θ values, a highly diffusive solution is found. In this study, we use the $\theta = 1$ for simplicity. The multivariable MINMOD function is defined as

$$\text{MINMOD}(z_1, z_2, \dots) = \begin{cases} \min_i\{z_i\}, & \text{if } z_i > 0, \quad \forall i, \\ \max_i\{z_i\}, & \text{if } z_i < 0, \quad \forall i, \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

The central scheme's numerical viscosity coefficient is $\frac{\Delta t}{\Delta x} a_{i+1/2}$ which is less than the numerical viscosity of the Lax-Friedrichs scheme, corresponding to the Courant-Friedrichs-Lewy (CFL) condition. There can be an accumulation of numerical dissipation, $\mathcal{O}\left(\frac{(\Delta x)^{2r}}{\Delta t}\right)$; this is one of the disadvantages of the central scheme [15]. The cell interface flux is calculated with the modified-Lax-Friedrichs scheme [12],

$$\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^n = \frac{1}{2} \left(\mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n,-}) + \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n,+}) \right) + \frac{1}{2} a_{i+\frac{1}{2}}^n \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} - \mathbf{U}_{i+\frac{1}{2}}^{n,+} \right). \quad (26)$$

Here, a is the one-sided local speed. In the original Lax-Friedrichs scheme, a is set as $\Delta t/\Delta x$. In this approach, a is calculated from the maximum eigenvalue of the Jacobian matrix,

$$a_{i+\frac{1}{2}}^n = \max \left\{ \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \right\}. \quad (27)$$

The eigenvalues of the drift flux model are

$$\lambda = \{v_l + c_m, v_g, v_l - c_m\}. \quad (28)$$

The eigenvalues are derived with the assumption of having an incompressible liquid phase [16]. Here, c_m is the sound velocity in the mixture of gas and liquid. Based on the assumption of incompressible liquid and no slip for each phase, assuming $\alpha_g \rho_g \ll \alpha_l \rho_l$ and $0 < \alpha_g < 1$, the mixture sound velocity can be presented as [1], [17],

$$c_m = \sqrt{\frac{p}{\alpha_g \rho_l (1 - k \alpha_g)}}. \quad (29)$$

Here we introduce the stencils for the central scheme in a form similar to that of ENO schemes. To calculate the cell interface function, the second order central scheme uses one stencil from one side, because the MINMOD limiter function selects a single stencil from a side.

5. Second order central-upwind scheme for the drift flux model

As an improvement of the high-resolution central scheme, the central-upwind scheme was developed by Kurganov et al. [15]. The scheme is a semi-discrete scheme, Godunov-type central scheme since it is based on integration over the Riemann fan. The upwind nature of the scheme is adapted by measuring the one-sided local speeds which present the directions of wave propagation. The one-sided propagation speeds are the largest and smallest eigenvalues of the Jacobian matrix at the cell interface. The central-upwind scheme is a Godunov-type central scheme because the evolution employs the integration over the Riemann solver, and it does not require a Riemann solver and a characteristic decomposition. Equation 22 and 23 calculate the cell interface values, $\mathbf{U}_{i+\frac{1}{2}}^{n,\pm}$, with the MINMOD limiter

function. Compared with the high-resolution central scheme, the high-resolution central-upwind scheme only differs is the cell interface flux calculation. Specifically, the central-upwind scheme consists of negative and positive direction one-sided local speeds at the cell interface. These are called the discontinuities propagation speeds. The discontinuities propagation speeds in right and left sides are

$$a_{i+\frac{1}{2}}^{n,+} = \max \left\{ \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \right\}, \quad (30)$$

$$a_{i+\frac{1}{2}}^{n,-} = \min \left\{ \min \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \min \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \right\}. \quad (31)$$

The central-upwind scheme bounds the Riemann fan by $(a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}) \Delta t$ [15]. The numerical flux of the high resolution central-upwind scheme is

$$\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^n = \frac{a_{i+\frac{1}{2}}^{n,+} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n,-}) - a_{i+\frac{1}{2}}^{n,-} \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}^{n,+})}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} + \frac{a_{i+\frac{1}{2}}^{n,+} a_{i+\frac{1}{2}}^{n,-}}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} - \mathbf{U}_{i+\frac{1}{2}}^{n,-} \right). \quad (32)$$

The numerical flux, $\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^n$, is independent of Δt , therefore the numerical viscosity is independent of $\mathcal{O}\left(\frac{1}{\Delta t}\right)$ in both the central and the central-upwind schemes. This is an advantage for using these two schemes for steady state calculations. The order of the scheme is second order.

6. WENO scheme for the drift flux model

Harten et al. [10] made the fundamental step for developing the ENO scheme. The ENO scheme uses a non-linear adaptive procedure to select the smoothest stencil where it can avoid the crossing discontinuities as much as possible [7]. ENO schemes map the cell averages in the stencil to the value of cell interface where there exists constant values. As the first step of the ENO scheme, a polynomial reconstruction is used to approximate the cell averages. If the degree of the interpolation polynomial function is r , then the order of the ENO scheme becomes r . The weighted ENO (WENO) scheme can be generated from the same stencil node in the ENO scheme with higher order accuracy. Liu et al. [6] developed the WENO scheme and it is further developed by Jiang and Shu [19]. The WENO scheme uses a convex combination of all the candidate stencils instead of just one used in the ENO scheme. Therefore, the WENO scheme can achieve higher order accuracy, $2r - 1$, by using the same ENO stencils.

6.1. Fifth order WENO scheme for the drift flux model

Parallel to the third order ENO scheme, the fifth order WENO scheme uses a convex combination of the all three stencils [20].

The third order ENO approach is used to calculate the cell interface variable related to each stencil. The negative direction cell interface with related to the stencils s_0 , s_1 , and s_2 are

$$\mathbf{U}_{i+\frac{1}{2}}^{-,0} = \frac{1}{3} \mathbf{U}_{i-2} - \frac{7}{6} \mathbf{U}_{i-1} + \frac{11}{6} \mathbf{U}_i, \quad (33)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{-,1} = -\frac{1}{6} \mathbf{U}_{i-1} + \frac{5}{6} \mathbf{U}_i + \frac{1}{3} \mathbf{U}_{i+1}, \quad (34)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{-,2} = \frac{1}{3} \mathbf{U}_i + \frac{5}{6} \mathbf{U}_{i+1} - \frac{1}{6} \mathbf{U}_{i+2}. \quad (35)$$

The positive direction cell interface with related to the stencils s_0 , s_1 , and s_2 are

$$\mathbf{U}_{i+\frac{1}{2}}^{+,0} = \frac{11}{6} \mathbf{U}_{i+1} - \frac{7}{6} \mathbf{U}_{i+2} + \frac{1}{3} \mathbf{U}_{i+3}, \quad (36)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{+,1} = \frac{1}{3} \mathbf{U}_i + \frac{5}{6} \mathbf{U}_{i+1} - \frac{1}{6} \mathbf{U}_{i+2}, \quad (37)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{+,2} = -\frac{1}{6} \mathbf{U}_{i-1} + \frac{5}{6} \mathbf{U}_i + \frac{1}{3} \mathbf{U}_{i+1}. \quad (38)$$

Here we have skipped the superscript n in time. The cell averages are used to calculate the cell interface value $\mathbf{U}_{i+\frac{1}{2}}^{\pm}$ in each stencil. Here we only present the

left side stencils approach. The convex combination of the left biased stencils of the ENO approach gives the WENO approach, at the cell interface $i + \frac{1}{2}$,

$$\mathbf{U}_{i+\frac{1}{2}}^{-} = \omega_0 \mathbf{U}_{i+\frac{1}{2}}^{-,0} + \omega_1 \mathbf{U}_{i+\frac{1}{2}}^{-,1} + \omega_2 \mathbf{U}_{i+\frac{1}{2}}^{-,2}. \quad (39)$$

The weight functions are

$$\omega_0 = \frac{\alpha_0}{\alpha_0 + \alpha_1 + \alpha_2}, \quad (40)$$

$$\omega_1 = \frac{\alpha_1}{\alpha_0 + \alpha_1 + \alpha_2}, \quad (41)$$

$$\omega_2 = \frac{\alpha_2}{\alpha_0 + \alpha_1 + \alpha_2}, \quad (42)$$

Here, $\omega_0 + \omega_1 + \omega_2 = 1$, and

$$\alpha_0 = \frac{1/10}{(\epsilon + \beta_0)^2}, \quad (43)$$

$$\alpha_1 = \frac{6/10}{(\epsilon + \beta_1)^2}, \quad (44)$$

$$\alpha_2 = \frac{3/10}{(\epsilon + \beta_2)^2}. \quad (45)$$

$\epsilon = 10^{-6}$ is used to avoid division by zero β is the smoothness indicator. The smoothness indicator is defined based on undivided differences [20]. The smoothness indicator consists of a $(r - 1)$ -th order accurate polynomial functions,

$$\beta_0 = \frac{13}{12} (\mathbf{U}_{i-2} - 2\mathbf{U}_{i-1} + \mathbf{U}_i)^2 + \frac{1}{4} (\mathbf{U}_{i-2} - 4\mathbf{U}_{i-1} + 3\mathbf{U}_i)^2, \quad (46)$$

$$\beta_1 = \frac{13}{12} (\mathbf{U}_{i-1} - 2\mathbf{U}_i + \mathbf{U}_{i+1})^2 + \frac{1}{4} (\mathbf{U}_{i-1} - \mathbf{U}_{i+1})^2, \quad (47)$$

$$\beta_2 = \frac{13}{12} (\mathbf{U}_i - 2\mathbf{U}_{i+1} + \mathbf{U}_{i+2})^2 + \frac{1}{4} (3\mathbf{U}_i - 4\mathbf{U}_{i+1} + \mathbf{U}_{i+2})^2. \quad (48)$$

By symmetry, the $\mathbf{U}_{i+\frac{1}{2}}^{+}$ values can be calculated.

Here we choose the local Lax-Friedrichs flux to calculate the monotone flux in the WENO scheme.

The local Lax-Friedrichs flux is

$$\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^n = \mathbf{F} \left(\frac{\mathbf{U}_{i+\frac{1}{2}}^{n,-} + \mathbf{U}_{i+\frac{1}{2}}^{n,+}}{2} \right) - \frac{1}{2} a_{i+\frac{1}{2}}^n \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} - \mathbf{U}_{i+\frac{1}{2}}^{n,+} \right), \quad (49)$$

where $a_{i+\frac{1}{2}}^n$ is

$$a_{i+\frac{1}{2}}^n = \max \left\{ \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\} \right\}. \quad (50)$$

6.2. Third order WENO scheme for the drift flux model

The third order WENO scheme is developed from the second-order ENO scheme. Parallel to the fifth order WENO scheme, the third order WENO scheme has two stencils in each side. Here we briefly explain the third order WENO scheme, one can compare the fifth order WENO scheme parallelly. The cell interface values for the stencil s_0 and s_1 respectively are

$$\mathbf{U}_{i+\frac{1}{2}}^{-,0} = -\frac{1}{2}\mathbf{U}_{i-1} + \frac{3}{2}\mathbf{U}_i, \quad (51)$$

$$\mathbf{U}_{i+\frac{1}{2}}^{-,1} = \frac{1}{2}\mathbf{U}_i + \frac{1}{2}\mathbf{U}_{i+1}. \quad (52)$$

The negative cell interface value is

$$\mathbf{U}_{i+\frac{1}{2}}^- = \omega_0\mathbf{U}_{i+\frac{1}{2}}^{-,0} + \omega_1\mathbf{U}_{i+\frac{1}{2}}^{-,1}. \quad (53)$$

The weight functions are

$$\omega_0 = \frac{\alpha_0}{\alpha_0 + \alpha_1}, \quad (54)$$

$$\omega_1 = \frac{\alpha_1}{\alpha_0 + \alpha_1}, \quad (55)$$

where,

$$\alpha_0 = \frac{1}{(\epsilon + \beta_0)^2}, \quad (56)$$

$$\alpha_1 = \frac{1}{(\epsilon + \beta_1)^2}, \quad (57)$$

and,

$$\beta_0 = (\mathbf{U}_{i-1} - \mathbf{U}_i)^2, \quad (58)$$

$$\beta_1 = (\mathbf{U}_i - \mathbf{U}_{i+1})^2. \quad (59)$$

The right side (positive direction) to the cell interface $i + \frac{1}{2}$.

7. Proposing a high resolution central-upwind-WENO scheme for the drift flux model

Here we use the above mentioned central scheme, the central-upwind scheme, and the WENO scheme for development of a central-upwind-WENO scheme. The main idea of the central-upwind-WENO scheme is to calculate the numerical flux from the central-upwind flux, and the cell interface values are calculated from the WENO reconstruction. The objective is to adapt the central-upwind nature in WENO scheme which can produce more accurate and stable results. The order of the scheme is decided by the polynomial interpolation used to calculate the cell interface flux.

7.1. Development of high resolution fifth order central-upwind-WENO scheme for the drift flux model

The cell interface flux is calculated with the central-upwind flux. Here we rewrite Equation 32 which gives the central-upwind flux,

$$\mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^n = \frac{a_{i+\frac{1}{2}}^{n,+}\mathbf{F}\left(\mathbf{U}_{i+\frac{1}{2}}^{n,-}\right) - a_{i+\frac{1}{2}}^{n,-}\mathbf{F}\left(\mathbf{U}_{i+\frac{1}{2}}^{n,+}\right)}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} + \frac{a_{i+\frac{1}{2}}^{n,+}a_{i+\frac{1}{2}}^{n,-}}{a_{i+\frac{1}{2}}^{n,+} - a_{i+\frac{1}{2}}^{n,-}} \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} - \mathbf{U}_{i+\frac{1}{2}}^{n,-} \right). \quad (60)$$

The one-sided local speeds in Equation 60 can be calculated from the largest and smallest eigenvalues of the system. Here we rewrite the Equations 25 and 26 form of the central-upwind scheme:

$$a_{i+\frac{1}{2}}^{n,+} = \max \left\{ \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \max \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \right\}, \quad (61)$$

$$a_{i+\frac{1}{2}}^{n,-} = \min \left\{ \min \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,-} \right) \right\}, \min \left\{ \lambda \left(\mathbf{U}_{i+\frac{1}{2}}^{n,+} \right) \right\}, 0 \right\}. \quad (62)$$

The cell interface values $\mathbf{U}_{i+\frac{1}{2}}^{n,\pm}$ can be calculated from

the fifth order WENO reconstruction. Here, we rewrite Equation 39, which gives the negative direction cell interface value,

$$\mathbf{U}_{i+\frac{1}{2}}^- = \omega_0\mathbf{U}_{i+\frac{1}{2}}^{-,0} + \omega_1\mathbf{U}_{i+\frac{1}{2}}^{-,1} + \omega_2\mathbf{U}_{i+\frac{1}{2}}^{-,2}. \quad (63)$$

The weight functions can be calculated from Equations 40 to 48. The fifth order central-upwind-WENO scheme is developed from the central-upwind flux and the fifth order WENO reconstruction for the cell interface values. Once we know the cell interface flux, it is possible to calculate flux gradients.

7.2. Development of high resolution third order central-upwind-WENO scheme for the drift flux model

Like the fifth order central-upwind-WENO scheme, the third order central-upwind-WENO scheme can be developed. The numerical flux function is calculated from the Equations 60 to 62. The cell interface values can be calculated from Equations 51 to 59.

8. Results

The shock tube problem is a benchmark case for testing the capability of numerical schemes in pipe flow. Evje and Flåtten [1] used the following initial conditions for a discontinuity in a pipe flow, see Figure 1. Here, subscript L is left side and subscript R is right side.

$p_L = 80450 \text{ Pa}$	$p_R = 24282 \text{ Pa}$
$\alpha_g = 0.55$	$\alpha_g = 0.55$
$v_L = 10.37 \text{ m/s}$	$v_L = 0.561 \text{ m/s}$

Figure 1. Initial condition of the shock tube problem-1 in the pipe flow

The initial gas velocity, v_g , can be calculated as, $v_g = (kv_l\alpha_l + v_d)/(1 - k\alpha_l)$. For the test case, we assume that the diameter of the pipe is 0.1 m, the length of the pipe is 100 m, the step length is $\Delta x = 0.5$ m. The initial discontinuity is at $x = 50$ m when $t = 0$ s. The friction force and gravity force terms are neglected for the numerical test, and $k = 1.07$ and $v_d = 0.216$ m/s. The test problem is purely a convection dominated flow; this initial condition generates discontinuities for all three conservative variables. The fifth order WENO scheme with fine mesh is used as the reference solution; we used 2.5

times more cells in the fluid domain compared to the other schemes. The fifth order WENO scheme gives accurate results compared with [16]. Figure 2 shows the shock tube problem result comparison between the central-upwind-WENO scheme and the central scheme. The results are shown after 1 s of simulation for pressure, gas velocity, liquid velocity, and liquid volume fraction. The central-upwind-WENO scheme gives a better solution than the central scheme, especially at the square wave in the solution. The second order central scheme is more diffusive compared to the other schemes used in this study. Figure 3 shows the shock tube result comparison between the third order central-upwind-WENO scheme vs. the second order central-upwind scheme. Compared to the central scheme, the central-upwind scheme gives more accurate results. However, the central-upwind-WENO scheme gives even higher accuracy than the central-upwind scheme. Figure 4 shows the shock tube problem result after 1s of simulation for the third order central-upwind-WENO scheme vs. the third order WENO scheme. The third order WENO scheme gives more accurate results than the central scheme and the central-upwind scheme. However, the third order central-upwind-WENO scheme gives more accurate results than the third order WENO scheme. The third order central-upwind-WENO scheme gives sharper gradient for rarefaction waves and shock wave fronts compared to the third order WENO scheme, the second order central scheme, and the second order central-upwind scheme. In other words, the third-order central-upwind-WENO gives the highest accuracy.

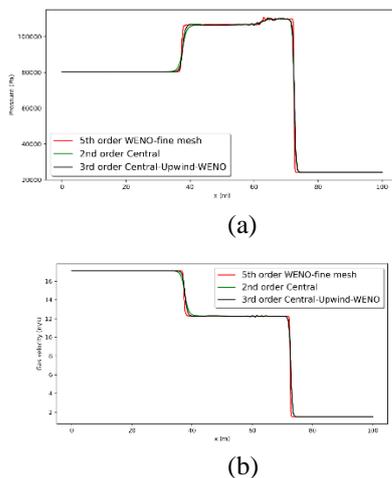


Figure 2. Shock tube problem results comparison after 1 s simulation between the third order high-resolution central-upwind-WENO scheme vs. the second order high-resolution central scheme. The fifth order WENO scheme with a fine mesh is used as the reference: (a)Pressure, (b)Gas velocity.

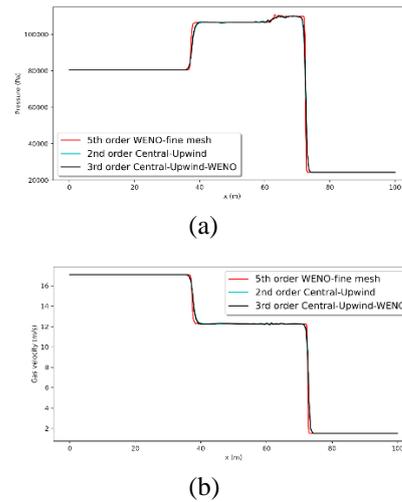


Figure 3. Shock tube problem results comparison after 1 s simulation between the third order high-resolution central-upwind-WENO scheme vs. the second order high-resolution central-upwind scheme. The fifth order WENO scheme with a fine mesh result is used as the reference: (a)Pressure, (b)Gas velocity.

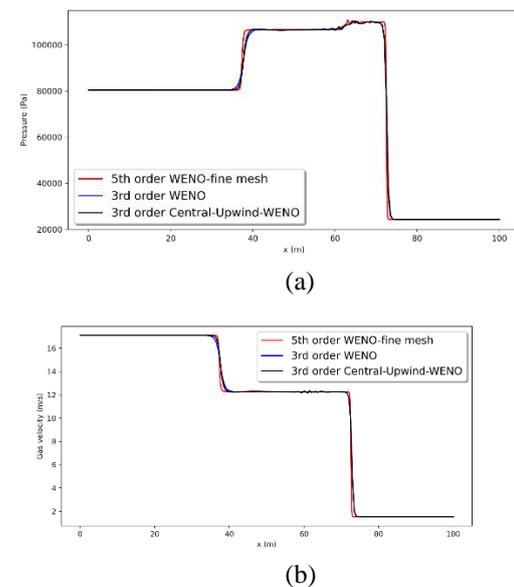


Figure 4. Shock tube problem results comparison after 1 s simulation between the third order high-resolution central-upwind-WENO scheme vs. the third order high-resolution WENO scheme. The fifth order WENO scheme with a fine mesh result is used as the reference: (a)Pressure, (b)Gas velocity.

Table 1 shows the computational speed comparison for the schemes. The simulations have conducted for the shock tube problem up to 1s, and all the schemes have used same coarse mesh which has 200 cells. The CPU times presented in the table are averages of five simulation runs for each case. The third order WENO scheme is the fastest scheme, it has higher

accuracy than the second order central scheme and second order central-upwind scheme. Compared to the third order WENO scheme, the central-upwind scheme is a slower scheme. The MINMOD limiter function has logical “if” closure structures, or equivalent mathematical formulae, which are not very efficient. The WENO scheme solves algebraic equations with weight functions. Therefore, the central and the central-upwind second order schemes are slower than the third order WENO scheme. However, all the schemes are in same speed range, this is no big differences. One can argue that fifth order WENO scheme has comparatively high speed for the coarse mesh. However, the fifth order WENO scheme gives higher oscillatory results for the coarse mesh than the third order WENO schemes, see Figure 5. The fifth order WENO scheme is very accurate and less oscillatory for finer mesh. The fifth order central-upwind-WENO scheme is more accurate than the fifth order WENO scheme. However, the fifth order WENO scheme and the fifth order central-upwind-WENO scheme produce more oscillatory results than their third order schemes.

Table 1: Computational time comparison: Computation time was calculated for 1s simulations. All the schemes were used same coarse mesh, 200 cells

Scheme	CPU Time (s)
Third order central-upwind-WENO	0.90
Fifth order WENO	0.95
Third order WENO	0.80
Second order central	0.88
Second order central-upwind	0.92

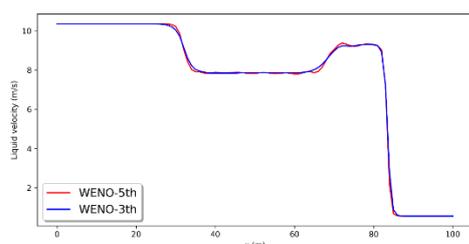


Figure 5. Coarse mesh simulation of the fifth order WENO scheme and the third order WENO scheme: The mesh has 200 cells. The results show liquid velocity for the shock tube problem

9. Concluding remarks

In this study, we considered possible numerical schemes for the two-phase drift flux model. The drift flux model is typically used in well-drilling and oil-gas production process in the petroleum industry. The high resolution second order central scheme, the high resolution second order central-upwind scheme and the third order and fifth order WENO schemes

were used for numerical simulations. The stencil selection procedure and numerical discretization procedure were explained in detailed for each scheme. We developed the high resolution third order and fifth order central-upwind-WENO scheme for the drift flux model. The main idea of the central-upwind-WENO scheme is to calculate the numerical flux from the central-upwind flux, and the cell interface values are calculated from the WENO reconstruction. This method helps to achieve higher order of the scheme with minimum number of stencils. The developed central-upwind-WENO scheme produces more accurate results than the central scheme, the central-upwind scheme, and the WENO scheme. All the schemes used in the study are suitable for the drift-flux model simulation. The source term effect, especially friction and gravity, will be discussed in future publications.

Acknowledgement

We gratefully acknowledge the economic support from The Research Council of Norway and Equinor ASA through project no. 255348/E30 “Sensors and models for improved kick/loss detection in drilling (Semi-kidd)”.

References

- [1] S. Evje and T. Flåtten, “Hybrid flux-splitting schemes for a common two-fluid model,” *Journal of Computational Physics*, vol. 192, no. 1, pp. 175–210, 2003, doi: 10.1016/j.jcp.2003.07.001.
- [2] S. T. Munkejord, *Analysis of the two-fluid model and the drift-flux model for numerical calculation of two-phase flow*, no. November. 2005.
- [3] H. Shi *et al.*, “Drift-flux modeling of two-phase flow in wellbores,” *SPE Journal*, vol. 10, no. 1, pp. 24–33, 2005, doi: 10.2118/84228-PA.
- [4] Z. Li, G. Oger, and D. Le Touzé, “A finite volume WENO scheme for immiscible inviscid two-phase flows,” *Journal of Computational Physics*, vol. 418, p. 109601, Oct. 2020, doi: 10.1016/j.jcp.2020.109601.
- [5] C. Berg, “Estimation of Influx and Loss During Drilling Operations Modeling for Automatic Control and Estimation of Influx and Loss During Drilling Operations A PhD dissertation in,” University of South-Eastern Norway, 2020.
- [6] X.-D. Liu, S. Osher, and T. Chan, “Weighted Essentially Non-oscillatory

- Schemes,” *Journal of Computational Physics*, vol. 115, no. 1, pp. 200–212, 1994, doi: <https://doi.org/10.1006/jcph.1994.1187>.
- [7] C. W. Shu, “Essentially Non-Oscillatory and Weighted Essentially Non-Oscillatory Schemes for Hyperbolic Conservation Laws,” *ICASE Report*, no. 97–65, 1997, doi: [10.1007/bfb0096355](https://doi.org/10.1007/bfb0096355).
- [8] S. Gottlieb and C. W. Shu, “Total variation diminishing Runge-Kutta schemes,” *Mathematics of Computation of the American Mathematical Society*, vol. 67, no. 221, 1998, doi: [10.1090/s0025-5718-98-00913-2](https://doi.org/10.1090/s0025-5718-98-00913-2).
- [9] C.-W. Shu and S. Osher, “Efficient implementation of essentially non-oscillatory shock-capturing schemes,” *Journal of Computational Physics*, vol. 77, no. 2, pp. 439–471, Aug. 1988, doi: [10.1016/0021-9991\(88\)90177-5](https://doi.org/10.1016/0021-9991(88)90177-5).
- [10] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, “Uniformly high order accurate essentially non-oscillatory schemes, III,” *Journal of Computational Physics*, vol. 71, no. 2, pp. 231–303, 1987, doi: [10.1016/0021-9991\(87\)90031-3](https://doi.org/10.1016/0021-9991(87)90031-3).
- [11] H. Nessyahu and E. Tadmor, “Non-oscillatory central differencing for hyperbolic conservation laws,” *Journal of Computational Physics*, vol. 87, no. 2, pp. 408–463, Apr. 1990, doi: [10.1016/0021-9991\(90\)90260-8](https://doi.org/10.1016/0021-9991(90)90260-8).
- [12] A. Kurganov and E. Tadmor, “New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection-Diffusion Equations,” *Journal of Computational Physics*, vol. 160, no. 1, pp. 241–282, May 2000, doi: [10.1006/jcph.2000.6459](https://doi.org/10.1006/jcph.2000.6459).
- [13] E. F. Toro, *Riemann solvers and numerical methods for fluid dynamics-A Practical Introduction*, 3rd ed. Heidelberg: Springer Science & Business Media, 2009. doi: [10.1007/b7976](https://doi.org/10.1007/b7976).
- [14] R. J. LeVeque, *Finite volume methods for hyperbolic problems*, 1st ed. Cambridge: Cambridge University Press, 2002.
- [15] A. Kurganov, S. Noelle, G. Petrova, and S. J. C. Sci, “Semidiscrete central - upwind schemes for hyperbolic conservation laws and Hamilton – Jacobi equations,” *SIAM Journal on Scientific Computing*, vol. 23, no. 3, pp. 707–740, 2001, doi: [10.1137/S1064827500373413](https://doi.org/10.1137/S1064827500373413).
- [16] K. K. Fjelde and K. H. Karlsen, “High-resolution hybrid primitive-conservative upwind schemes for the drift flux model,” *Computers and Fluids*, vol. 31, no. 3, pp. 335–367, Mar. 2002, doi: [10.1016/s0045-7930\(01\)00041-x](https://doi.org/10.1016/s0045-7930(01)00041-x).
- [17] M. H. Abbasi *et al.*, “A Godunov-type Scheme for the Drift Flux Model with Variable Cross Section,” *Journal of Petroleum Science and Engineering*, vol. 179, no. February 2019, pp. 796–813, 2019, doi: [10.1016/j.petrol.2019.04.089](https://doi.org/10.1016/j.petrol.2019.04.089).
- [18] C. W. Shu and S. Osher, “Efficient implementation of essentially non-oscillatory shock-capturing schemes, II,” *Journal of Computational Physics*, vol. 77, no. 2, pp. 439–471, 1988, doi: [10.1016/0021-9991\(88\)90177-5](https://doi.org/10.1016/0021-9991(88)90177-5).
- [19] G. S. Jiang and C. W. Shu, “Efficient implementation of weighted ENO schemes,” *Journal of Computational Physics*, vol. 126, no. 1, pp. 202–228, 1996, doi: [10.1006/jcph.1996.0130](https://doi.org/10.1006/jcph.1996.0130).
- [20] G. S. Jiang and D. Peng, “Weighted ENO Schemes for Hamilton-Jacobi Equations,” *Journal of Applied Mathematics and Stochastic Analysis*, vol. 2008, pp. 1–24, 2008, doi: [10.1155/2008/359142](https://doi.org/10.1155/2008/359142).

An improved delay and state observer for SISO LTI systems with known delay lower bound

Jonas Ingebretsen¹ Damiano Rotondo*¹

¹Department of Electrical and Computer Engineering (IDE), University of Stavanger, Stavanger, Norway
damiano.rotondo@uis.no

Abstract

It is known that the presence of delays hinders the performance achievable by a feedback control system, and it can even lead to closed-loop instability if not considered during the design. For this reason, predictors are often included in the loop, although they typically require the knowledge of the exact value of the delay, which in some applications is hard to obtain in practice. This paper presents a method to design an observer that simultaneously estimates the unknown state and the time-varying input delay of a plant based on an available model and the measurements coming from the sensors. In particular, the main contribution of this paper is to show that by accounting for a known lower bound of the input delay, it is possible to improve the observer's performance when compared to state-of-the-art approaches encountered in the literature. Simulations are used to illustrate the efficiency of the proposed design method.

Keywords: input delay, delay estimation, state observer

1 Introduction

An input delay system is a particular case of time-delay system in which the delay affects only the input signal. There are many possible causes for input delays, of both computational and physical nature. In many situations, these delays must be explicitly considered during the design, because failing to do so can lead to a non-acceptable degradation of the performance or to the loss of stability.

The most widely used approach to account for delays is the Smith predictor (Smith, 1959) which was extended to work with time-varying delays in (Normey-Rico et al., 2012) and (Franklin and Santos, 2020). An alternative approach that has drawn attention lately relies on the linear parameter varying (LPV) framework (Rotondo et al., 2019) and achieves the desired performance/stabilization using a delay-scheduled controller (Briat et al., 2009b,a). However, even when the time delay is explicitly taken into account, the presence of uncertainty on the delay value can produce fragility in the control system. A possible way to address this issue and decrease the amount of uncertainty in the delay value is to estimate it online. For this reason, many works in the literature have focused on the problem of time delay estimation.

The approaches used for time delay estimation can be classified into signal processing and control-oriented (Léchappé et al., 2018). In the former class of methods, the value of the delay that minimizes some criterion (cost function) based on available data collected from the process is sought (Knapp and Carter, 1976; Jacovitti and Scarano, 1993). However, this comes at the cost of a high computation time and they require the knowledge of the delayed signal, which is not always realistic. For this reason, control-oriented approaches consider the delay as a

parameter to be estimated online. (Agarwal and Canudas, 1987) approximated the delay term using the Padé form, then used least-squares to minimize an objective function. (Tuch et al., 1994) proposed a recursive least-squares algorithm in the frequency domain, which had the drawback of requiring perfectly known initial conditions.

Among the control-oriented approaches, observer-based methods have been under consideration lately. For instance, (Cacace et al., 2015) proposed an augmented observer able to estimate a constant state delay. A Kalman filter-based solution was proposed in (Léchappé et al., 2015) and (Léchappé et al., 2018). On the other hand, adaptive and sliding mode observers were applied to time delay estimation in (Wu et al., 2013) and (Drakunov et al., 2006).

In this paper, we consider the observer-based solution proposed by (Léchappé et al., 2015), and we improve it under the assumption that a known lower bound is available for the time-varying delay. We discuss how this knowledge can be used when applying the Taylor's theorem, so that a generally smaller remainder is obtained. Consequently, a better performing observer is implemented, as demonstrated by simulations under different realizations of the delay and input signals.

The paper is organized as follows. Section 2 presents the problem formulation and summarizes the state and delay observer proposed by (Léchappé et al., 2015). Section 3 shows that by considering a lower bound for the delay, it is possible to perform a different sequence of calculations that lead to an improved state and delay observer structure. The discussion about the observer stability is provided in Section 4. The performance of the proposed improved observer is demonstrated using simulation examples in Sec-

tion 5. Finally, the main conclusions are drawn in Section 6.

2 Problem and background

Consider the following SISO LTI system:

$$\begin{cases} \dot{x}(t) = Ax(t) + bu(t - d(t)) \\ y(t) = c^T x(t) \end{cases} \quad (1)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}$ is the known input, $y \in \mathbb{R}$ is the measured output, and $d \in \mathbb{R}$ is the unknown input delay, which should be estimated. The matrix A and the vectors b, c are assumed to be known and such that the pair (A, c^T) is observable. On the other hand, it is assumed that the time-varying delay $d(t)$ is continuous and differentiable, with $|\dot{d}(t)| \leq H$. Finally, it is assumed that the signal $u(t)$ is smooth, which means $u \in \mathcal{C}^2$ and there exists a bound $M > 0$ such that $|\ddot{u}(t)| \leq M$ for all $t \geq -\bar{d}$.

The problem under consideration is the design of an observer that can estimate the state $x(t)$ and the input delay $d(t)$ from the knowledge of $y(t)$, $u(t)$ and $\dot{u}(t)$.

The work by (Léchappé et al., 2015) assumed that the time-varying delay $d(t)$ satisfies $d(t) \in [0, \bar{d}]$, where $\bar{d} > 0$ denotes the upper bound of $d(t)$. Then, they proceed to apply the Taylor's theorem to expand $u(t - d(t))$ about t , obtaining:

$$u(t - d(t)) = u(t) - d(t)\dot{u}(t) + \gamma(t - d(t)) \quad (2)$$

where $\gamma(\cdot)$ is the so-called *remainder*, for which a uniform bound can be obtained given that $|\ddot{u}(t)| \leq M$, as follows:

$$|\gamma(t - d(t))| \leq M \frac{d(t)^2}{2} \quad (3)$$

This allowed (Léchappé et al., 2015) to rewrite (1) as:

$$\begin{cases} \dot{x}(t) = Ax(t) + bu(t) - b\dot{u}(t)d(t) + b\gamma(t - d(t)) \\ y(t) = c^T x(t) \end{cases} \quad (4)$$

and, by augmenting the state vector as $z(t) = [x(t)^T \quad d(t)^T]^T$:

$$\begin{cases} \dot{z}(t) = \begin{bmatrix} A & -b\dot{u}(t) \\ 0 & 0 \end{bmatrix} z(t) + \begin{bmatrix} b \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} b\gamma(t - d(t)) \\ \dot{d}(t) \end{bmatrix} \\ y(t) = [c^T \quad 0] z(t) \end{cases} \quad (5)$$

which can be brought to the compact form:

$$\begin{cases} \dot{z}(t) = \bar{A}(\dot{u}(t)) z(t) + \bar{B}u(t) + \Gamma(t, t - d(t)) \\ y(t) = \bar{C}z(t) \end{cases} \quad (6)$$

where the definition of \bar{A} , \bar{B} , \bar{C} and $\Gamma(\cdot)$ is straightforward.

Then, the following observer was proposed to reconstruct and estimate $\hat{z}(t)$ of $z(t)$ (and, in turn, to obtain estimates $\hat{x}(t)$ and $\hat{d}(t)$ of $x(t)$ and $d(t)$):

$$\dot{\hat{z}}(t) = \bar{A}(\dot{u}(t)) \hat{z}(t) + \bar{B}u(t) - L(t) (\bar{C} \hat{z}(t) - y(t)) \quad (7)$$

where the discussion about how to choose the observer gain L will be omitted, since it is not relevant for further developments (the interested reader is referred to (Léchappé et al., 2015)).

3 Improved state and delay observer

As already mentioned, the approach proposed by (Léchappé et al., 2015) and summarized in the previous section works under the assumption that $d(t) \in [0, \bar{d}]$. In many practical situations, it is realistic to assume that not only an upper bound \bar{d} is available, but a lower bound \underline{d} as well, which means that $d(t) \in [\underline{d}, \bar{d}]$. In this case, expanding $u(t - d(t))$ about $t - \underline{d}$ would lead to a reduced magnitude of the remainder term, which acts as an unknown exogenous disturbance in the model (5), thus being responsible for degrading the state/delay estimate obtained via the observer (7).

Then, the following is obtained:

$$u(t - d(t)) = u(t - \underline{d}) - d(t)\dot{u}(t - \underline{d}) + \underline{d}\dot{u}(t - \underline{d}) + \tilde{\gamma}(t - d(t)) \quad (8)$$

where $\tilde{\gamma}(\cdot)$ is a new remainder term, for which a uniform bound is calculated as follows:

$$|\tilde{\gamma}(t - d(t))| \leq M \frac{(d(t) - \underline{d})^2}{2} \quad (9)$$

By comparing (3) and (9), it is clear that given the same value of the delay $d(t)$, it can be generally expected that:

$$|\tilde{\gamma}(t - d(t))| \leq |\gamma(t - d(t))| \quad (10)$$

will likely hold.

Based on (8), Eq. (1) can be rewritten as:

$$\begin{cases} \dot{x}(t) = Ax(t) + bu(t - \underline{d}) - b\dot{u}(t - \underline{d})d(t) + b\underline{d}\dot{u}(t - \underline{d}) + b\tilde{\gamma}(t - d(t)) \\ y(t) = c^T x(t) \end{cases} \quad (11)$$

and, using the augmented state vector $z(t)$:

$$\begin{cases} \dot{z}(t) = \begin{bmatrix} A & -b\dot{u}(t - \underline{d}) \\ 0 & 0 \end{bmatrix} z(t) + \begin{bmatrix} b \\ 0 \end{bmatrix} u(t - \underline{d}) + \begin{bmatrix} b\tilde{\gamma}(t - d(t)) \\ \dot{d}(t) \end{bmatrix} \\ y(t) = [c^T \quad 0] z(t) \end{cases} \quad (12)$$

which can be brought to a form similar to (6):

$$\begin{cases} \dot{z}(t) = \bar{A}(\dot{u}(t - \underline{d})) z(t) + \bar{B} \begin{bmatrix} u(t - \underline{d}) \\ \dot{u}(t - \underline{d}) \end{bmatrix} + \tilde{\Gamma}(t, t - d(t)) \\ y(t) = \bar{C}z(t) \end{cases} \quad (13)$$

with an appropriate definition of the matrices \bar{A} , \bar{B} , \bar{C} and $\tilde{\Gamma}(\cdot)$.

For the system (13), let us use the following observer:

$$\dot{\hat{z}}(t) = \bar{A}(\dot{u}(t - \underline{d})) \hat{z}(t) + \bar{B} \begin{bmatrix} u(t - \underline{d}) \\ \dot{u}(t - \underline{d}) \end{bmatrix} - L(t) (\bar{C} \hat{z}(t) - y(t)) \quad (14)$$

where, following (Léchappé et al., 2015), the gain $L(t)$ is chosen according to a Kalman filter-like structure, i.e.:

$$L(t) = S(t)^{-1} \bar{C}^T R \quad (15)$$

where R is a positive diagonal matrix that acts as a filter, chosen as $R = I$ in the noise-free scenario, and the symmetric matrix S is obtained as the solution of the following matrix differential equation:

$$\begin{aligned} \dot{S}(t) = & -\rho S(t) - \bar{A}(\dot{u}(t-d))^T S(t) \\ & - S(t)\bar{A}(\dot{u}(t-d)) + \bar{C}^T R \bar{C} \end{aligned} \quad (16)$$

with a positive constant $\rho > 0$ that affects the convergence speed of S and an initial condition $S(0) \succ 0$.

4 Observer stability

This section provides the observer stability proof, which follows the theoretical steps discussed in (Léchappé et al., 2015). For the sake of proving the stability, let us recall the following lemma (Khalil, 2002).

Lemma 1. *Let $x = 0$ be an exponentially stable equilibrium point of the nominal system:*

$$\dot{x}(t) = f(t, x(t)) \quad (17)$$

and let $V(t, x(t))$ be a Lyapunov function for (17) that satisfies for all $t \geq 0$ and $\forall x \in \mathcal{D} = \{x \in \mathbb{R}^n : \|x\|_2 < r\}$:

$$c_1 \|x\|^2 \leq V(t, x(t)) \leq c_2 \|x\|^2 \quad (18)$$

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t, x(t)) \leq -c_3 \|x\|^2 \quad (19)$$

$$\left\| \frac{\partial V}{\partial x} \right\| \leq c_4 \|x\| \quad (20)$$

for some positive constants c_1, c_2, c_3, c_4 . Also, assume that a perturbation term $g(t, x)$ satisfies:

$$\|g(t, x)\| \leq \gamma(t) \|x\| + \delta(t) \quad \forall t \geq 0, \forall x \in \mathcal{D} \quad (21)$$

where $\gamma: \mathbb{R} \rightarrow \mathbb{R}$ is a nonnegative and continuous function such that:

$$\int_{t_0}^t \gamma(\tau) d\tau \leq \varepsilon(t - t_0) + \eta \quad (22)$$

for some nonnegative constants ε and η where:

$$\varepsilon < \frac{c_1 c_3}{c_2 c_4} \quad (23)$$

and $\delta: \mathbb{R} \rightarrow \mathbb{R}$ is nonnegative, continuous, and bounded for all $t \geq 0$. Provided that:

$$\|x(t_0)\| < \frac{r}{\rho} \sqrt{\frac{c_1}{c_2}} \quad (24)$$

$$\sup_{t \geq t_0} \delta(t) < \frac{2c_1 \theta r}{c_4 \rho} \quad (25)$$

with:

$$\theta = \frac{1}{2} \left[\frac{c_3}{c_2} - \varepsilon \frac{c_4}{c_1} \right] > 0 \quad (26)$$

$$\rho = \exp\left(\frac{c_4 \eta}{2c_1}\right) \geq 1 \quad (27)$$

then, the solution of the perturbed system:

$$\dot{x}(t) = f(t, x(t)) + g(t, x(t)) \quad (28)$$

satisfies:

$$\|x(t)\| < \sqrt{\frac{c_2}{c_1}} \rho \|x(t_0)\| e^{-\theta(t-t_0)} + \frac{c_4 \rho}{2c_1} \int_{t_0}^t e^{-\theta(t-\tau)} \delta(\tau) d\tau \quad (29)$$

In addition, let us recall the following lemma given by (Besançon et al., 1996), which characterizes the matrix $S(t)$ obtained as a solution of (16).

Lemma 2. *Assume that \dot{u} is regularly persistent, and consider the differential equation (16). Then $\exists \rho_0 > 0$ such that for any symmetric positive definite matrix $S(0)$, $\forall \rho \geq \rho_0$, there exist $\bar{\alpha}, \bar{\beta}, t_0 > 0$ such that $\forall t \geq t_0$:*

$$\bar{\alpha} I \preceq S(t) \preceq \bar{\beta} I \quad (30)$$

Let us define the Lyapunov candidate function:

$$V(e(t)) = e(t)^T S(t) e(t) \quad (31)$$

where $S(t)$ is the solution of (16) and $e(t) = \hat{z}(t) - z(t)$. Given the assumption of regularly persistent $\dot{u}(t)$, in virtue of Lemma 2, there exists a positive scalar ρ_0 so that $\forall \rho \geq \rho_0$ there exist $\bar{\alpha}, \bar{\beta}, t_0 > 0$ such that $\forall t \geq t_0$:

$$\bar{\alpha} \|e(t)\|^2 \leq V(e(t)) \leq \bar{\beta} \|e(t)\|^2 \quad (32)$$

so that (18) holds with $c_1 = \bar{\alpha}$ and $c_2 = \bar{\beta}$. The dynamics of the estimation error is described by:

$$\dot{e}(t) = \hat{z}(t) - z(t) \quad (33)$$

which, using (13)-(15), follows:

$$\begin{aligned} \dot{e}(t) = & \bar{A}(\dot{u}(t-d)) \hat{z}(t) + \bar{B} \begin{bmatrix} u(t-d) \\ \dot{u}(t-d) \end{bmatrix} - L(t) (\bar{C} \hat{z}(t) - y(t)) \\ & - \bar{A}(\dot{u}(t-d)) z(t) - \bar{B} \begin{bmatrix} u(t-d) \\ \dot{u}(t-d) \end{bmatrix} - \tilde{\Gamma}(t, t-d(t)) \\ = & (\bar{A}(\dot{u}(t-d)) - L(t) \bar{C}) e(t) - \tilde{\Gamma}(t, t-d(t)) \\ = & (\bar{A}(\dot{u}(t-d)) - S(t)^{-1} \bar{C}^T R \bar{C}) e(t) - \tilde{\Gamma}(t, t-d(t)) \end{aligned} \quad (34)$$

Let us neglect the perturbation term $\tilde{\Gamma}(t, t-d(t))$ so that:

$$\dot{V}(e(t)) = \dot{e}(t)^T S(t) e(t) + e(t)^T \dot{S}(t) e(t) + e(t)^T S(t) \dot{e}(t) \quad (35)$$

becomes, after taking into account (16) and (34):

$$\begin{aligned} \dot{V}(e(t)) = & e(t)^T [\bar{A}(u(t-d)) - S(t)^{-1} \bar{C}^T R \bar{C}]^T S(t) e(t) \\ & - \rho e(t)^T S(t) e(t) - e(t)^T \bar{A}(\dot{u}(t-d))^T S(t) e(t) \\ & - e(t)^T S(t) \bar{A}(\dot{u}(t-d)) e(t) + e(t)^T \bar{C}^T R \bar{C} e(t) \\ & + e(t)^T S(t) [\bar{A}(u(t-d)) - S(t)^{-1} \bar{C}^T R \bar{C}] e(t) \\ = & -\rho e(t)^T S(t) e(t) - e(t)^T \bar{C}^T R \bar{C} e(t) \end{aligned} \quad (36)$$

Due to the positive definiteness of the matrix $\bar{C}^T R \bar{C}$, (32) and (36) lead to:

$$\dot{V}(e(t)) \leq -\rho \bar{\alpha} \|e(t)\|^2 \quad (37)$$

so that (19) holds with $c_3 = \rho \bar{\alpha}$. Finally, from (32) it follows that:

$$\left\| \frac{\partial V(e(t))}{\partial e} \right\| \leq 2\bar{\beta} \|e(t)\| \quad (38)$$

which means that (20) holds with $c_4 = 2\bar{\beta}$.

Let us consider now the perturbation term $\tilde{\Gamma}(t, t-d(t))$. Notably, in this case the domain \mathcal{D} corresponds to the entire state-space ($r = \infty$), which means that by choosing $\gamma(t) = 0$ and $\delta(t) = \sup \|\tilde{\Gamma}(\cdot)\|$ in (21), then (24)-(25) always hold and we can choose $\varepsilon = \eta = 0$ in (22)-(23). Consequently, (26)-(27) lead to $\rho = 1$ and $\theta = \bar{\alpha}/2\bar{\beta}$. Then, according to Lemma 1, $e(t)$ will satisfy the following inequality:

$$\begin{aligned} \|e(t)\| &< \sqrt{\frac{\bar{\beta}}{\bar{\alpha}}} \|e(t_0)\| e^{-\frac{\bar{\alpha}}{2\bar{\beta}}(t-t_0)} \\ &+ \frac{\bar{\beta}}{\bar{\alpha}} \int_{t_0}^t e^{-\frac{\bar{\alpha}}{2\bar{\beta}}(t-\tau)} \sup \|\tilde{\Gamma}(\cdot)\| d\tau \\ &= \sqrt{\frac{\bar{\beta}}{\bar{\alpha}}} \|e(t_0)\| e^{-\frac{\bar{\alpha}}{2\bar{\beta}}(t-t_0)} \\ &+ \frac{\bar{\beta}}{\bar{\alpha}} \int_0^{t-t_0} e^{-\frac{\bar{\alpha}}{2\bar{\beta}}s} \sup \|\tilde{\Gamma}(\cdot)\| ds \\ &= \sqrt{\frac{\bar{\beta}}{\bar{\alpha}}} \|e(t_0)\| e^{-\frac{\bar{\alpha}}{2\bar{\beta}}(t-t_0)} \\ &+ \frac{2\bar{\beta}^2}{\bar{\alpha}^2} \left(1 - e^{-\frac{\bar{\alpha}}{2\bar{\beta}}(t-t_0)}\right) \sup \|\tilde{\Gamma}(\cdot)\| \end{aligned} \quad (39)$$

which shows the ultimate boundedness of the estimation error $e(t)$.

It is worth remarking that the regular persistency of $\dot{u}(t)$ is a key point for the above discussion to hold true, which is the reason why in practice the estimation error spikes at times where this assumption does not hold true, as shown in the results in the next section.

5 Simulation results

Let us consider the second-order example proposed by (Léchappé et al., 2015):

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t-d(t)) \quad (40)$$

$$y(t) = [1 \quad 0] x(t) \quad (41)$$

for which we assume that $x(0) = [1.5, 1]^T$. Then, the extended system is described by matrices:

$$\bar{A}(\dot{u}) = \begin{bmatrix} 0 & 1 & 0 \\ -2 & -3 & -\dot{u} \\ 0 & 0 & 0 \end{bmatrix} \quad \bar{B} = \begin{bmatrix} 0 & 0 \\ 1 & \underline{d} \\ 0 & 0 \end{bmatrix} \quad (42)$$

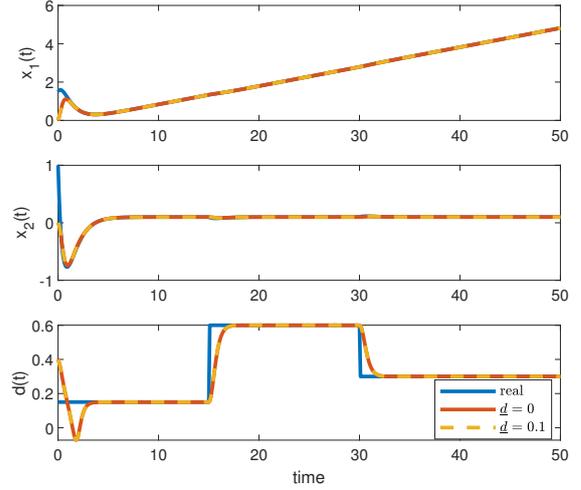


Figure 1. Simulation results (Scenario 1).

where \bar{d} is the known lower bound for $d(t)$. We will assume that the observer (14)-(16) is obtained with $\rho = 5$, $S(0) = I_3$, $\hat{x}(0) = [0, 0]^T$, $\hat{d}(0) = 0.4$, and we will compare the performance of the observer proposed in this paper with the results in (Léchappé et al., 2015).

5.1 Scenario 1

Let us first consider the delay sequence:

$$d(t) = \begin{cases} 0.15 & \text{for } 0 \leq t \leq 15 \\ 0.6 & \text{for } 15 < t \leq 30 \\ 0.3 & \text{otherwise} \end{cases} \quad (43)$$

with a ramp input signal $u(t) = 0.2t$ for which $\dot{u}(t) = 0.2$, and let us consider a known lower bound for the delay $\underline{d} = 0.1$. Under these conditions, we obtain the simulation results shown in Figure 1. In this scenario, the relationship (2) holds with a zero remainder term, so that (6) reduces to:

$$\dot{z}(t) = \bar{A}(\dot{u}(t)) z(t) + \bar{B}u(t) \quad (44)$$

and the estimation errors for both the state variables and the delay tend asymptotically to zero. Notably, no difference between the case with $\underline{d} = 0$ and $\underline{d} \neq 0$ is perceivable in this case.

5.2 Scenario 2

Let us now consider the delay $d(t) = 0.4 + 0.2 \sin(0.4t)$ while keeping the ramp input signal $u(t) = 0.2t$, with the known lower bound for the delay $\underline{d} = 0.2$. In this case, $\dot{d}(t) \neq 0$ acts as an exogenous disturbance that prevents the observer from estimating the delay correctly, as shown in Figure 2, where $d(t) - \hat{d}(t)$ exhibits a clear steady-state error. Notably, also in this case the knowledge of a precise lower bound \underline{d} for $d(t)$ does not play any role.

5.3 Scenario 3

Let us now consider the following modification to Scenario 1: $u(t) = \sin(0.1t)$, so that the terms $\gamma(t-d(t))$

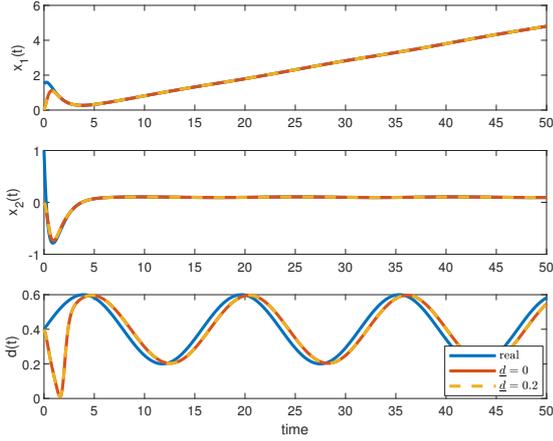


Figure 2. Simulation results (Scenario 2).

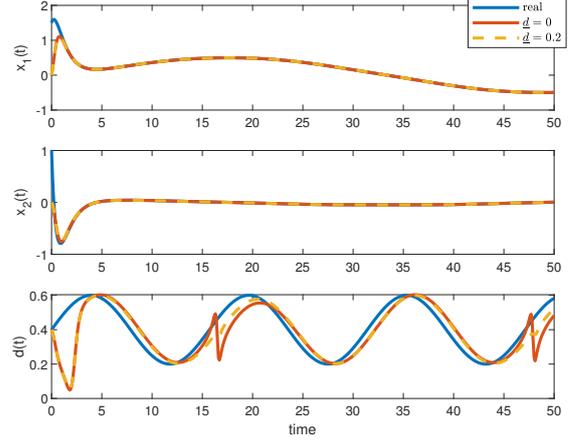


Figure 4. Simulation results (Scenario 4).

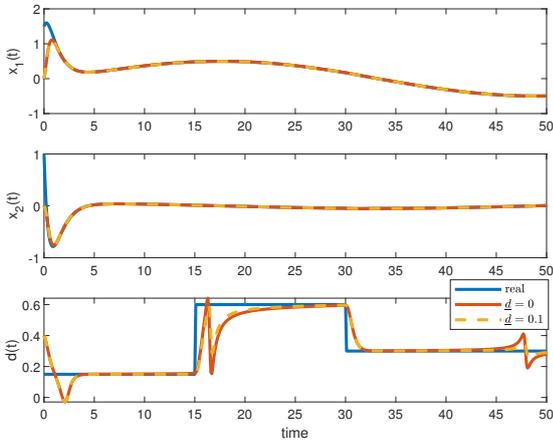


Figure 3. Simulation results (Scenario 3).

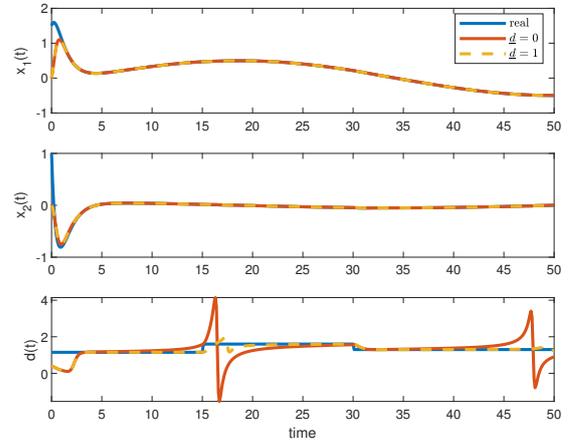


Figure 5. Simulation results (Scenario 5).

and $\tilde{\gamma}(t-d(t))$ appear in (2) and (8), respectively. Given that (10) is likely to happen, we would now expect to see a difference between the performance of the observer with $\underline{d} = 0$ and that of the observer with $\underline{d} \neq 0$. This is confirmed by the simulation response shown in Figure 3, where a slight improvement brought by the proposed observer in the estimation of $d(t)$ can be perceived.

5.4 Scenario 4

Let us now consider the delay from Scenario 2 $d(t) = 0.4 + 0.2\sin(0.4t)$ with known lower bound for the delay $\underline{d}(t) = 0.2$, and the input signal from Scenario 3 $u(t) = \sin(0.1t)$. The corresponding results are showed in Figure 4. It can be seen that the estimate obtained with $\underline{d} = 0$ suffers much more from the loss of observability of the system (5) when $\dot{u}(t) = 0$, which is the reason for the sudden changes in $\hat{d}(t)$ at approx. 16.5s and 47.5s.

5.5 Scenario 5

We will now analyze the behavior of the proposed observer under relatively big delay and known lower bound.

Under such situation, the difference between the right-hand terms in the inequalities (3) and (9) becomes more significant, which should lead to a bigger difference between the performance of the observer without known lower bound ($\underline{d} = 0$) and that of the proposed observer. More specifically, let us consider $d(t)$ as follows:

$$d(t) = \begin{cases} 1.15 & \text{for } 0 \leq t \leq 15 \\ 1.6 & \text{for } 15 < t \leq 30 \\ 1.3 & \text{otherwise} \end{cases} \quad (45)$$

with $u(t) = \sin(0.1t)$, with a bound $\underline{d} = 1$. The corresponding results are shown in Figure 5, where it can be seen that the proposed observer estimates $d(t)$ with a much reduced error and mitigates the impact of the loss of observability when $\dot{d}(t) = 0$.

5.6 Scenario 6

Finally, Figure 6 shows the results obtained under the same conditions as Scenario 5, except for the delay being $d(t) = 1.4 + 0.2\sin(0.4t)$. Also in this case, the benefit of

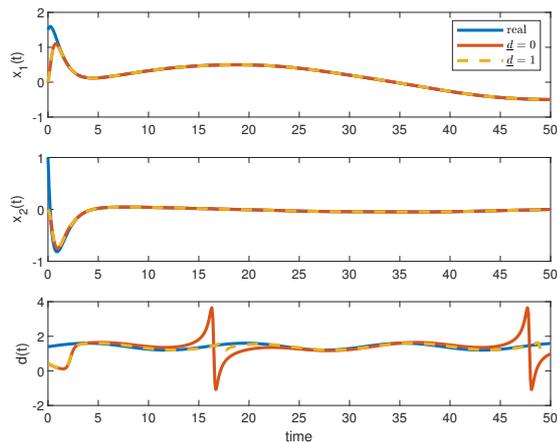


Figure 6. Simulation results (Scenario 6).

considering a lower bound for \underline{d} is evident by looking at the delay estimation error.

6 Conclusions

This paper has proposed an improved observer design for joint state and delay estimation. In particular, the improvement with respect to a similar design previously proposed in the literature comes from the knowledge of a lower bound for the time-varying delay, which can be taken into account during the application of Taylor's theorem, so that a generally smaller remainder is obtained. The simulation results have shown that the proposed design does not improve the state estimation. On the other hand, the delay estimation is improved sensibly in cases where approximation errors become non-negligible or when the slowly changing nature of the input signal ($\dot{u} \approx 0$) causes loss of observability issues. In particular, it was observed that for large delays, the performance improvement is outstanding. As a side note, it has been observed so far that this type of method is fragile when the system is affected by parametric uncertainties or nonlinearities, so that future research should be devoted to increase the robustness properties of the observer.

References

- M. Agarwal and C. Canudas. On-line estimation of time delay and continuous-time process parameters. *International Journal of Control*, 46(1):295–311, 1987.
- G. Besançon, G. Bornard, and H. Hammouri. Observer synthesis for a class of nonlinear control systems. *European Journal of control*, 2(3):176–192, 1996.
- C. Briat, O. Sename, and J.-F. Lafay. h_∞ delay-scheduled control of linear systems with time-varying delays. *IEEE Transactions on Automatic Control*, 54(9):2255–2260, 2009a.
- C. Briat, O. Sename, and J.-F. Lafay. Delay-scheduled state-feedback design for time-delay systems with time-varying delays—a lpv approach. *Systems & Control Letters*, 58(9):664–671, 2009b.
- F. Cacace, F. Conte, A. Germani, and G. Palombo. A new method for the simultaneous estimation of state and delay in time delay systems. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5666–5670. IEEE, 2015.
- S. V. Drakunov, W. Perruquetti, J.-P. Richard, and L. Belkoura. Delay identification in time-delay systems using variable structure observers. *Annual reviews in control*, 30(2):143–158, 2006.
- T. S. Franklin and T. L. M. Santos. Robust filtered smith predictor for processes with time-varying delay: A simplified stability approach. *European Journal of Control*, 56:38–50, 2020.
- G. Jacovitti and G. Scarano. Discrete time techniques for time delay estimation. *IEEE Transactions on signal processing*, 41(2):525–533, 1993.
- H. K. Khalil. Nonlinear systems third edition. *Prentice Hall*, 115, 2002.
- C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE transactions on acoustics, speech, and signal processing*, 24(4):320–327, 1976.
- V. Léchappé, J. De León, E. Moulay, F. Plestan, and A. Glumineau. Delay and state observer for siso lti systems. In *2015 American Control Conference (ACC)*, pages 4585–4590. IEEE, 2015.
- V. Léchappé, J. De Leon, E. Moulay, F. Plestan, and A. Glumineau. Delay and state observation for siso nonlinear systems with input delay. *International Journal of Robust and Nonlinear Control*, 28(6):2356–2368, 2018.
- J. E. Normey-Rico, P. Garcia, and A. Gonzalez. Robust stability analysis of filtered smith predictor for time-varying delay processes. *Journal of Process Control*, 22(10):1975–1984, 2012.
- D. Rotondo, H. S. Sanchez, F. Nejjari, and V. Puig. Analysis and design of linear parameter varying systems using lmis. *Rev. Iberoam. Autom. Inform. Ind*, 16:1–14, 2019.
- O. J. M. Smith. A controller to overcome dead time. *ISA J.*, 6:28–33, 1959.
- J. Tuch, A. Feuer, and Z. J. Palmor. Time delay estimation in continuous linear time-invariant systems. *IEEE Transactions on Automatic Control*, 39(4):823–827, 1994.
- X. Wu, Z. Sun, F. Liang, and C. Yu. Online estimation of unknown delays and parameters in uncertain time delayed dynamical complex networks via adaptive observer. *Nonlinear Dynamics*, 73(3):1753–1768, 2013.

Simulation and optimization of screw feeder in a bubbling fluidized bed gasification reactor

Rajan K. Thapa, Trym Fehn Vaa, Oddvin Vaa, Rajan Jaiswal, Mladen Jecmenica

Department of Process, Energy and Environmental Technology, University of South-Eastern Norway

Rajan.k.thapa@usn.no

Abstract

A fluidized bed biomass gasification reactor is used to produce syngas from biomass and municipal wastes. Gasification is a flexible technology where many different types of feedstocks can be used. The University of South-Eastern Norway has a 20kW gasification reactor which is used to investigate the quality and quantity of the syngas produced using different types of feedstocks. At the present, the reactor has the challenge to supply feedstock to the reactor via transport screws. The main challenge consists of achieving continuous feeding and reduction in the feed rate. Therefore, this work is focused on the optimization of the screw feeder in the gasification reactor to obtain a reduced feed rate while maintaining a continuous feeding rate. The aim is to reduce the feed rate from approximately 8 kg/h to 3-4 kg/h without missing continuity of the feeding.

A model of the screw feeder is developed using CAD software SolidWorks and the model is simulated using open-source simulation software LIGGGHTS to investigate feed rate and continuity using different combinations of transport screw parameters. The simulation results are processed using Excel and viewed graphically with the open-source visualization software ParaView. The simulation results are compared to the experimental measurements in the gasification reactor. The validated model is further used to investigate the feed rate with different combinations of transport screw parameters and the results are compared and discussed.

Keywords: Biomass, Screw Feeder, Open-source LIGGGHTS, CPFD, Simulation

1. Introduction

Feeding the biomass to the gasification reactor at the required rate is essential to achieve higher-quality syngas gas. Therefore, the feeding system for such gasification reactors plays a major role in the utilization of biomass. The most common method to transport feedstock to the reactor has been the use of a screw feeder (Cummer et al., 2002). Due to physical characteristics, moisture content, and heterogeneity of feedstocks, trouble-free transportation of feedstock to the reactor is difficult to obtain (Bandara et al., 2021; Elliott, 1989). The major challenge during the transportation of feedstock to the reactor includes blockage and bridging of the supply line which disrupts the uniform and continuous flow of the feedstock (Bandara et al., 2021; Castleman et al., 1994). During a gasification process, it is important to maintain the required air-to-fuel ratio for the smooth operation of such a reactor at higher efficiency (Jaiswal et al., 2020). The air-to-fuel ratio in a gasification reactor is determined by the amount of oxidizing medium available for supplied feedstock (fuel). When the feedstock supply is lowered to the required air-to-fuel ratio, the quality of syngas is reduced due to the increased carbon dioxide fraction

in the product gas due to partial combustion instead of gasification. On the other hand, if the feedstock is supplied at a higher rate, the supplied heat may not be sufficient for the endothermic reaction to gasify the feedstock. Then, the reactor temperature may drop significantly, reducing the conversion efficiency. Therefore, it is vital to maintain the required flow rate of the biomass.

The University of South-Eastern Norway has a gasification reactor with a capacity of 20 kW and is placed in SINTEF's building that is used for various research.

This work aims to optimize the feeding system for this reactor. Currently, the feeding system of the reactor delivers a high amount of biomass if the reactor is run in continuous mode. The plan for the optimization is to design and produce new feeding screws while keeping the rest of the reactor system in its present condition. These feeding screws need to be able to run continuously while keeping the feed rate not exceeding 3.0 kg/h to 4.0 kg/h.

The optimization process is started by measuring the rotation speed and feed rate of the current screw feeder in the gasification reactor. This measurement gave the baseline data for further optimization. In the next step, a mechanical model of the screw

feeding system is prepared using the CAD software SolidWorks and the model is simulated using an open-source simulation software LIGGGHTS (LAMMPS Improved for General Granular and Granular Heat Transfer Simulations) where LAMMS is Large-scale Atomic/Molecular Massively Parallel Simulator. A series of simulations were performed using different types of screw feeders with different parameters. From the simulation results, three of the screw feeders that gave the best results were taken further for investigation. The screw feeders were made in the mechanical workshop located at the University of South-Eastern Norway.

The screws are then calibrated, and the feed rates are registered. The goal is to optimize the feed rate without changing the reactor or the electrical motors. Other feeding systems than feeding screws are not looked at in this work.

Under the calibrating and registry, blockage and sealing failure and other actual challenges where different biomass will be used are not investigated. Here, the investigations will be done using only the wood pellets as feedstock.

Experiments are run to verify the best screw for the system. This will be done with help from people with a wider range of knowledge when it comes to running the reactor.

2. Material and Methods

First, an attempt has been made to simulate the feeding system using the commercial software SolidWorks. It is found that the program is not qualified for these types of simulations because SolidWorks is more focused on flow simulations and not particle simulations.

After this, another program that USN is licensed to use was tested. This program is called Barracuda Virtual Reactor and is more focused on a mixture of particle and fluid flow simulations. This was acknowledged as a possible candidate for the simulations.

Another attempt was made to use the open-source simulation software LIGGGHTS. The program is found to work a little bit differently than the previously mentioned programs but seemed ok to use for the simulations. This software uses a Discrete Element Method (DEM), and it seems like this gives sufficiently good results for the simulation of wood pellets transferred through the screw feeders. The benefit of using LIGGGHTS is that it uses less processor power than the other software needed for the same situations. It's also easier to use, both with the setup and the change of parameters and variables. Accessibility also has something to say, as CPFD only can be used at USN's campus, which is different from LIGGGHTS which can be utilized anywhere due to it being open-source. It was also checked if LIGGGHTS had been used in earlier

articles and research. It was found in Science Direct 552 articles related to LIGGGHTS simulations. The software doesn't have any completed run files accessible, only the source code. Therefore this will be downloaded and compiled from (LIGGGHTS, 2022). Some small changes were made to the source code to get it to be compiled.

Further, to visualize the results, it was necessary to use a program that could process the files that LIGGGHTS was producing. For this, ParaView was found as open-source software (Ayachit, 2015), and downloaded (ParaView, 2022).

2.1. Simulation of existing feeding system

To make a model that could be used further in the optimizations, it first must be validated. This means that the simulation results produced by LIGGGHTS, must be, if not exact, approximately accurate with the feeding rates from the existing screw feeding system. A series of measurements were carried out in the gasification reactor to measure the feed rate of the biomass pellets using the existing screw feeder. The feeding rate was after 1 hour, 7.67 kg/h at 16 % motor efficiency. This was done by defining the density and size of the particles, different material parameters, contact models for particles against particles, and particles against walls. Friction coefficients also had to be defined for the last-mentioned parameters.

First, a CAD model of the feeding system has been built. The model can be seen in Figure 1. To make the simulations run faster, the model was simplified. This means removing all geometry that doesn't affect the feeding rate. Everything except where the particles flow from the silo is removed. The hot screw also is removed because it was assumed to not affect the feeding rate. The hot screw rotates at a significantly higher speed and has no variation in the pitch.

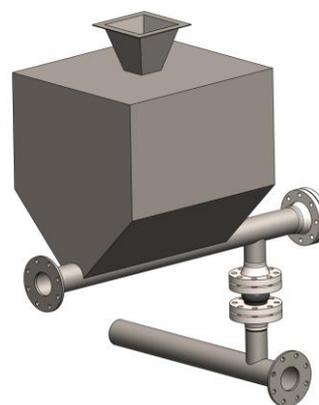


Figure 1: CAD model.

A container was also modeled and is intended to collect all the particles that fall through the descent between the two pipes. This results in a model that can be seen in Figure 2. The origin is set in the back

of the screw because this is the easiest way to make the rotation on the screw, as the screw needs an axis to be defined that it can rotate around.

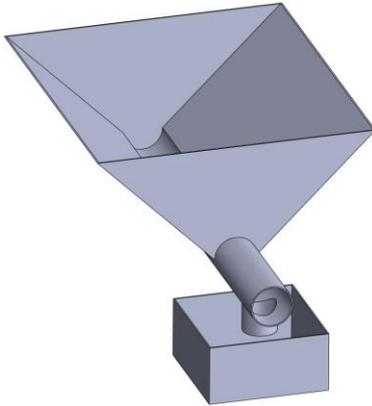


Figure 2: Simplified silo and container.

The model was then exported into STL files that LIGGGHTS could process, and it was only selected surfaces for each part because this is the only thing the software can read. This means 4 different files, one file for each of the components such as a silo, container, feeding screw, and the feed plane. These were exported with a deviation tolerance and an angular tolerance of 1.0 mm and 20.0 °, respectively.

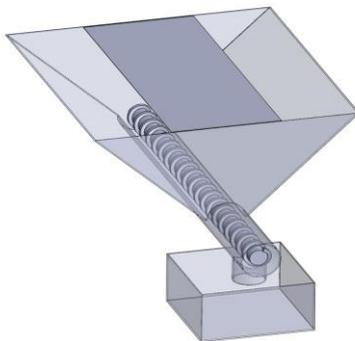


Figure 3: Simplified model.

For the silo, only the inner surfaces were selected, because these surfaces are the only ones that come in contact with the particles. This can be seen in Figure 4.

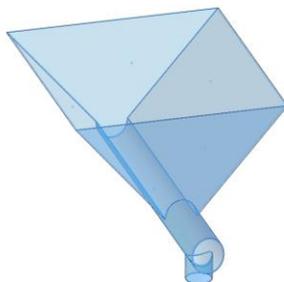


Figure 4: Selected surfaces for the silo.

For the screw, only the outer surfaces were selected, because these are the ones that meet the particles. This can be seen in Figure 5.



Figure 5: Selected surfaces for the screw.

For the container, only the outer surfaces were selected. This can be seen in Figure 6.

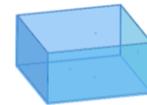


Figure 6: Selected surfaces for the container.

For the feeding plane, only the upper surface was selected from the modeled plate. This was put in the XY plane. This can be seen in Figure 7.



Figure 7: Selected surfaces for the feeding plane.

2.2 Explanation and setup for LIGGGHTS

DEM is an intuitive method with discrete particles colliding with each other and other surfaces through dynamic simulation. Usually, each DEM particle represents a type of granular material. In simulations where particles get exposed to complicated deformations, DEM is not a good method to use. (DEM,2022)

Generally speaking, the contact definitions are easily expandable to include contacts between DEM particles and FEM-based or analytic surfaces. Big relative displacements between particles are typical for applications where DEM is utilized. Particle-to-particle interaction can involve both similar and different particles. Each particle can be involved in many interactions at the same time. (DEM,2022)

The three situations above in Figure 8 show two particles that are just touching, two deformed spherical particles that push each other, and two rigid spherical particles that get pushed into each other with some penetration. (DEM,2022)

To be able to use LIGGGHTS, a program accessed in the form of a text file, that can be obtained from (DEM,2022) and examples found in (LIGGGHTS, 2022).

The decent in the model creates problems when the simulation is starting to run because it gets exposed to many points, and therefore the angles between the points are too small for LIGGGHTS to process.

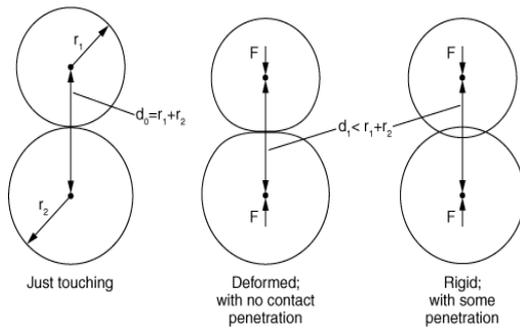


Figure 8: Interaction between spherical particles [5].

Therefore, this pipe will be drawn as a hexagon. The particles have an approximate cylindrical form but can be simplified to be simulated as spherical particles by converting them to spherical equivalent particles. These particles get a density of 1139.0 kg/m³ and 8.96 mm. (Agu et al., 2019) The contact models that are being used are hertz, tangential history, and rolling friction EPSD2 (LIGGGHTS, 2022).

3 Results and discussion

A series of simulations were performed using the simulation software LIGGGHTS. The main parameters used in the simulation are given in Table 1.

Table 1: Simulation Parameters.

Parameter	Value
Particle	-
Density	1139.0 kg/m ³
Diameter	8.96 mm
Modulus of elasticity	-
Particle	5.0 N/mm ²
Wall	2.1e5 N/mm ²
Poissons number	-
Particle	0.45
Wall	0.33
Coefficient of restitution	-
Particle to particle	0.10
Particle to wall	0.20
Frictions coefficient	-
Particle to particle	0.60
Particle to wall	0.52
Rolling coefficient of friction	-
Particle to particle	0.50
Particle to wall	0.45
Simulation time	> 900.0 s
Time increment	5e-5 s
Filling mass (pellets)	7.0 kg

Filling time	4.0 s
Rotational speed	32.43 s ⁻¹
Gravitation	9.81 m/s ²
Timestep	-
Visual data	0.2 s
Mass data	0.2 s

Simulations were run first, using the existing screw feeder. Each simulation was run for a minimum of 900 seconds of simulation time. The flow of particles from the silo through the screw feeder is shown in figure 9.

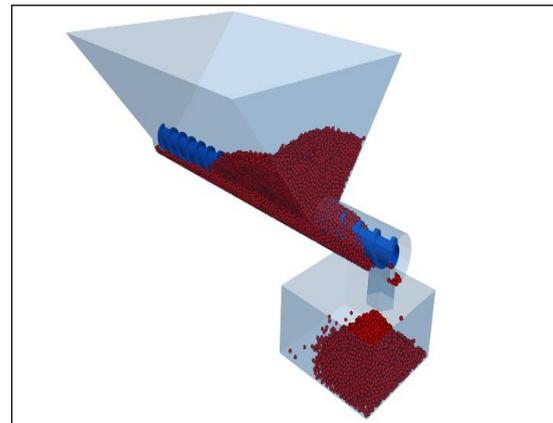


Figure 9: Visualization of the model after 912-second simulation.

Results from the simulation for the existing screw feeder are shown in Figure 10. The results for the first 300 seconds of simulation are not used in validation to avoid starting fluctuation effects.

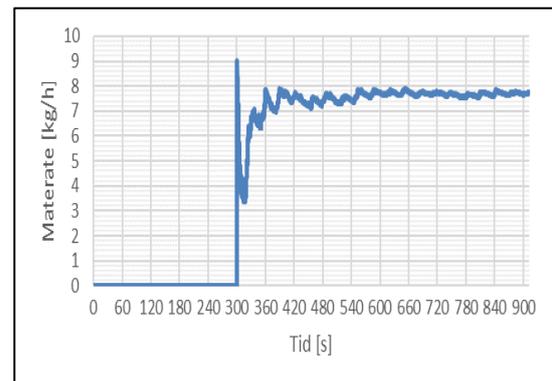


Figure 10: Results of simulation for the feed rate of the existing screw feeder.

The simulation and experimental results for feed rate are found to be 7.30 kg/h and 7.73 kg/h, respectively. The results show that the experimental and simulation results are in good agreement with each other thus the model can be used for further investigations.

Further, three candidate screw feeders with respective diameters of 48mm, 51mm, and 60mm are simulated using different parameters as shown in Figure 11.

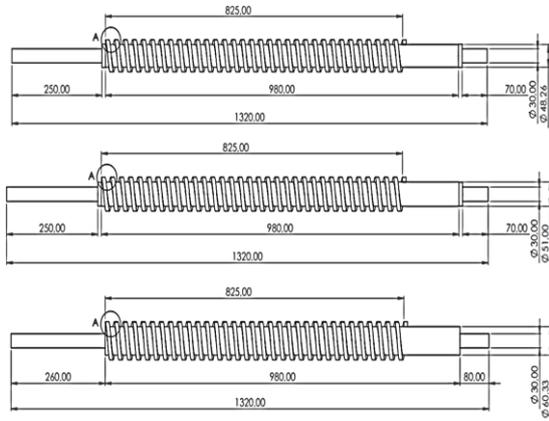


Figure 11: Screw candidates used in simulations.

Each of the screw-feeder was simulated with a varying number of pitches to check the effect of screw diameter as well as the pitch on the feed rate. Figure 12 shows the variation of feed rate with varying diameter and pitch.

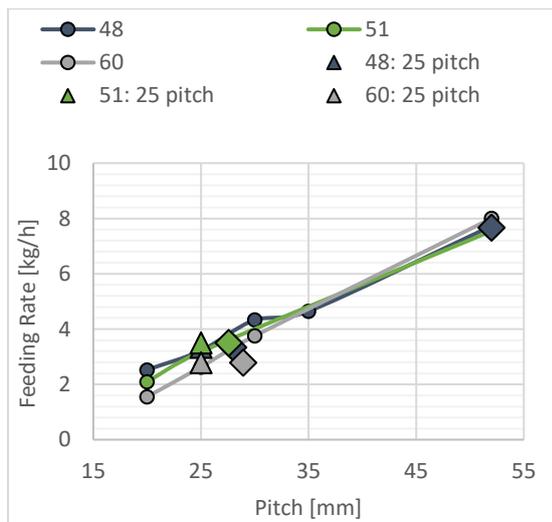


Figure 12: Variation of feed rate with diameter and pitch.

The results of the simulation show that increasing the pitch increases the feed rate. The results of the simulations for the three candidate screw feeders are shown in Figure 13. The results show that all screw feeders give the desired reduction of the feed rate. However, the screw with a diameter of 60 mm gives the minimum feed rate. Therefore, the screw is taken further for testing in the gasification reactor which gives the desired results.

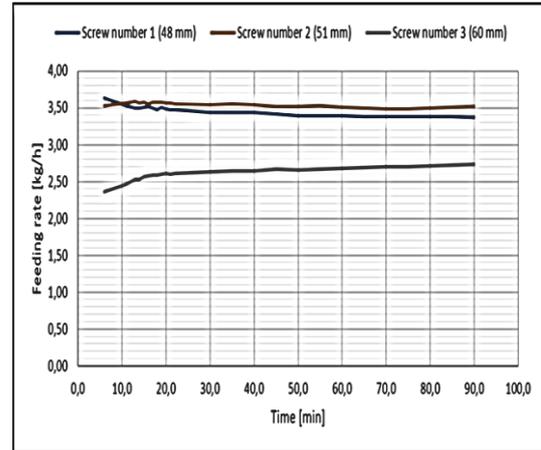


Figure 13: Results of simulation for a feed rate of the three candidate screws.

The simulation results for the candidate screw feeders are compared with the experimental measurements and presented in Table 2.

Table 2: Comparison of the results

Diameter [mm]	Measured [kg/h]	Simulated [kg/h]	Deviation [kg/h]
48	3.37	4.04	0.67
51	3.51	3.67	0.16
60	2.74	3.45	0.71

The results show that the experimental and simulated feed rates for those three candidate screw feeders have good agreements. However, the measured feed rate for the screw feeder with a diameter of 60 mm gave the minimum feed rate. Therefore the screw is installed in the feeding system of the gasification reactor. A gasification experiment was performed with the screw to investigate the gas composition which is compared with the gas composition obtained from the existing screw feeder. The results are shown in Table 3.

Table 3: Comparison of gas composition

Gass	With new screw [%]	With old screw [%]
Oxygen	0.69	0.89
Nitrogen	47.47	53.25
Methane	4.78	3.44
Carbon-monoxide	19.09	14.63
Carbon-dioxide	15.33	16.08
Hydrogen	12.62	11.68

The results show that the energy-containing gasses have increased while nitrogen and carbon dioxide has decreased with the use of the new screw.

4 Conclusions

Experiments were carried out to investigate a feed rate for biomass pellets to a biomass gasification reactor located at USN. A model of the screw feed system of the gasification reactor was made using the CAD software SolidWorks and the model is simulated using fluid-particle flow software LIGGGHTS. The simulation results were compared with experimental measurements to validate the model. A series of simulations were performed with different diameters and pitches of the screw. The screw with a diameter of 60mm gives the minimum feed rate which is desired for the reactor to get continuous feed. Experiments were performed using the new screw to verify again the simulation results. The experimental and simulation results are in good agreement.

Acknowledgments

We would like to thank Øyvind Johanson for providing the materials needed to produce the screws and the use of instruments in the mechanical workshop at USN.

References

- Agu, C. E. et al.(2019), 'Behaviour of biomass particles in a bubbling fluidized bed: A comparison between wood pellets and wood chips, Chemical Engineering Journal, bd. 363,s.84–98,doi:0.1016/j.cej.2019.01.120.
- Ayachit, U. (2015) 'The ParaView Guide: A Parallel Visualization Application.' Clifton Park, NY, USA: Kitware, Inc.
- Bandara, J.C., et al. (2021), ' Air gasification of wood chips, wood pellets and grass pellets in a bubbling fluidized bed reactor. Energy, **233**: p. 121149.
- Castleman, J., C. et al. (1994) 'Fluidized bed combustion and gasification: A guide for biomass waste generators.' Western Regional Biomass Energy Program, Reno, NV (United States).
- CFDEMproject/LIGGGHTS-PUBLIC. DCS Computing GmbH, (2022). Åpnet: 8. February. [Online]. Available on: <https://github.com/CFDEMproject/LIGGGHTS-PUBLIC>
- Cummer, K.R. and Brown, R.C.(2002) 'Ancillary equipment for biomass gasification, Biomass, and Bioenergy, **23**(2): p. 113-128. DEM.[https://abaqus\(2017\)docs.mit.edu/2017/English/SIMACAEANLRefMap/simaanl-c-demanalysis.htm](https://abaqus(2017)docs.mit.edu/2017/English/SIMACAEANLRefMap/simaanl-c-demanalysis.htm) (Accessed 10. May 2022).
- DEM Tutorial 2, 'Screw Conveyor Engineerdo', Engineerdo, 19.september2020. <https://www.engineerdo.com/2020/09/19/dem-tutorial-2-screw-conveyor/> (accessed 1. February 2022).
- Elliott, T.C. (1989), *Standard handbook of powerplant engineering*. 'Gran rolling_friction epsd2 model — LIGGGHTS v3.Documentation. https://www.cfdem.com/media/DEM/docu/gran_rolling_friction_epsd2.html (Åpnet 15. April 2022).
- Jaiswal, R., et al. (2020) 'Study of agricultural waste gasification in an air-blown bubbling fluidized bed using a CPFD model.' ParaView. (accessed 28. February 2022. [Online]. Available on: <https://www.paraview.org/>

Aspen Hysys simulation of the methanol synthesis based on gas from biomass gasification

Morten A. Fossen^a, Johann Halvorsrød^a, Thomas Narvestad^a, Sander Tjemsland^a, Ramesh Timsina^a, Marianne S. Eikeland^{a*}

^a *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway (USN), Norway*
corresponding author: marianne.eikeland@usn.no

Abstract

From the gasification of biomass, the produced gas mainly consists of CO, H₂, CO₂, and CH₄. After gas cleaning and conditioning, the syngas obtained can be used to produce methanol, dimethyl ether, polymers, biofuels, etc. Methanol is one of the important industrial chemicals that can be used directly as a fuel or can be blended into conventional fuels. Methanol produced from renewable biomass sources can limit greenhouse gas emissions. Based on data on gas composition from experiments and simulations of the fluidized bed gasification reactor at the University of South-Eastern Norway, the methanol process is simulated with Aspen Hysys. It is desirable to assess how different process parameters affect the yield of methanol production. A Gibbs reactor was modeled with three gas-phase exothermic reactions. The product from the reactor is depressurized to separate gas from the liquid. The liquid enters the distillation column to give CH₃OH in the distillate and the water as the bottom product. The yield of methanol is evaluated with regard to pressure, temperature, and H₂/CO ratio. The theoretical H₂/CO ratio for the methanol synthesis should be approximately between 1.5-2. However, different gas compositions were obtained from different types of fluidizing agents in the gasification reactor, and a low ratio of H₂/CO gives a low yield of methanol. Fluidized bed gasification with steam as gasifying agent is preferable to increase the H₂/CO ratio and give a higher methanol yield.

Keywords: Methanol, Aspen Hysys, Gasification, Biomass

1. Introduction

According to the report, Methanol Market - Growth, Trends, and Forecast (2021-2026), the market share for methanol in 2020 was 83.8 million tons, and it is predicted that from 2021 to 2026 the market would experience an average annual growth of around 5% (Research and Markets, 2018). All large-scale methanol plants currently use steam-reformed natural gas and pure oxygen as raw materials in the process. Research to improve current production methods to reduce the climate footprint is constantly evaluated. At the same time, new studies are constantly being evaluated on whether methanol can be produced from other sources, such as using synthesis gas from biomass gasification, to produce green methanol.

Gasification of biomass is a thermochemical conversion process where the chemical bonds are broken due to high temperature and partial oxidation with air or steam as the oxidizing agent. The product gas from the gasification reactor consists of CH₄, CO, CO₂, and H₂, as well as other light gases such as ethane (C₂H₆) and propane (C₃H₈) in addition to

various condensable gases. The gas will also contain some nitrogen (N₂). The proportion of N₂ will vary based on the type of fluidizing agent used. When using air, the nitrogen content will be relatively high, while when using pure oxygen or steam, the proportion will be significantly lower. The product gas will also have proportions of sulfuric compounds, chlorine, heavy metals, and other trace elements from the biomass (Bandara, 2021).

The composition of feed gas used in methanol synthesis has changed in recent years. Earlier, only H₂/CO was used, as the producers were convinced that the process was re-hydrogenation of CO and therefore removed all CO₂ from the gas with absorption. However, Waugh (Waugh, 2012) found that after running parallel experiments with H₂/CO₂/CO and H₂/CO, it was discovered that H₂/CO₂/CO produced methanol faster than H₂/CO. Tests were performed with a Cu/ZnO/Al₂O₃ catalyst, and it turns out that the highest production was achieved with a composition of CO/CO₂/H₂ in the ratio of 10:10:80 at a temperature of 290°C (Waugh, 2012).

In an exothermic reaction, energy is released, which results in an increase in the temperature. According to Le Chatelier's principle, the equilibrium will shift towards the reactants as energy is consumed from the reaction mixture. If the pressure in the mixture is increased, the equilibrium of the mixture will be shifted towards the product as there will be fewer molecules in the mixture (Moulijn J. A., 2013). Methanol synthesis from syngas is a low-temperature, high-pressure exothermic reaction (Yang L., 2016). The reactions and the reaction enthalpies are given in Tab. 1. Methanol is usually produced with high selectivity, although minor amounts of side products are found. The synthesis is usually conducted at 200–300 °C and 35–100 bar (Klerk, 2020). Common by-products of methanol formation are methane, methyl-formate, higher alcohols, and acetone.

Table 1: Reactions and reaction enthalpies for the methanol synthesis (Moulijn J. A., 2013)

Reactions	Reaction enthalpy (kJ/mol)
$\text{CO} + 2 \text{H}_2 \rightleftharpoons \text{CH}_3\text{OH}$	-90.64
$\text{CO}_2 + 3\text{H}_2 \rightleftharpoons \text{CH}_3\text{OH} + \text{H}_2\text{O}$	-49.67
$\text{CO} + \text{H}_2\text{O} \rightleftharpoons \text{CO}_2 + \text{H}_2$	-41

The University of South-Eastern Norway has a pilot plant bubbling fluidized bed (BFB) gasification reactor. Fig. 1 is a sketch of the reactor. The BFB reactor with optional equipment consists of a silo tank for biomass (1), a hopper for filling bed material (2), the gasification reactor (3), feed screws for biomass, cold and hot (4 and 5), gas heater for the fluidizing agent (6), and a compressor for the fluidizing agent (7). At the bottom of the reactor, there is an outlet for solid material (8). In several places of the reactor, it is mounted connection points for pressure and temperature readings (9). The gasification reactor is operated within the temperature range of 700-1100°C, with atmospheric pressure. The temperature is achieved by the heated fluidizing agent, heat generated in exothermic reactions, and electric heating cables mounted around the reactor (Timsina, 2022). Experiments were performed using wood chips of an irregular shape and with lengths ranging from 1 to 15 mm. with air as a fluidizing agent (Timsina R., 2020). The experiment was performed at 1100 K. Computational Particle Fluid Dynamics (CPFD) modeling is applied to simulate the bubbling fluidized bed gasifier reactor aiming at finding consistency between the experimental results and the simulation results. After validating the CPFD hydrodynamic model, related to the experimental results, the model has been used to investigate the effects of temperature on the steam gasification

process (Samani N.A., 2022). Based on the experimental data from the gasification pilot plant and the computational particle fluid dynamic simulations, the focus of this study is to determine the yield of methanol production based on the gas composition from the gasification of wood chips. The gas from the gasifier must be purified, cooled, and compressed before entering the methanol reactor for optimal conversion.

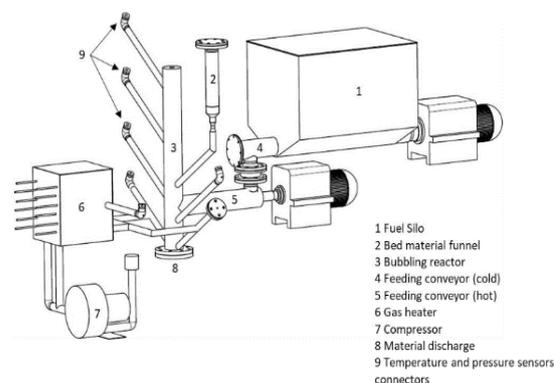


Figure 1: A schematic sketch of the bubbling fluidized bed gasifier at the University of South-Eastern Norway (Timsina, 2022)

2. Methodology

Simulations have been performed on the methanol synthesis using Aspen Hysys. The flowsheet is presented in Fig. 2. The feed consists of purified product gas from the gasification of wood chips with either air or steam as the fluidizing agent. It is also possible to supply the feed with a supplemental stream that goes inside the MIX-100 to give the possibility to adjust the feed gas ratio. Stream 1 is compressed in the K-100 and sent to the MIX-101 where stream 2 is mixed with 7_RCY which is a recycle stream. After the compression, the gas is heated or cooled before the gas enters the GBR-100, a Gibbs reactor. The gases are converted to methanol, and the reactions taking place are shown in Tab. 1. Stream 5 is cooled down in the cooler E-101, before entering V-100, a gas-liquid separator. The TEE-100 split the stream 7_gas in 7_Purge, while the stream to be recycled goes to 7_RCY and back into the MIX-101 together with the feed gas. The stream 7_Liquid is depressurized and enters the distillation column T-100 where methanol and water are separated. The remaining gases go out in the stream 8_Gas.

The Gibbs Reactor of Aspen HYSYS can work solely as a separator, a reactor that minimizes the Gibbs free energy without an attached reaction set or as a reactor using equilibrium reactions. When a reaction set is attached, the stoichiometry involved in the reactions is used in the Gibbs Reactor

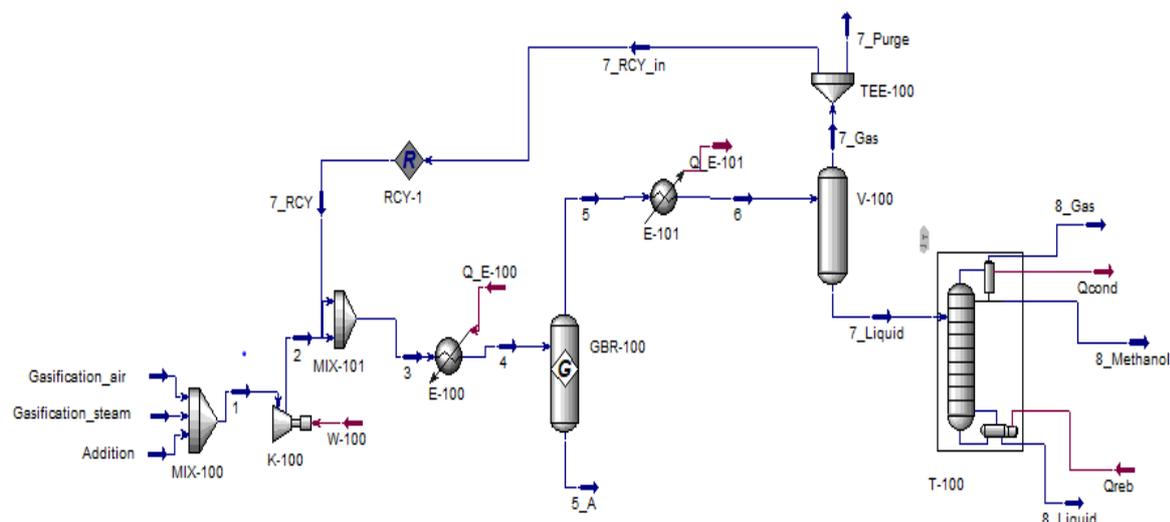


Figure 2: Process flowsheet for the Aspen Hysys Simulation (M Fossen, 2022)

calculations (Haydari, 2018). In this study, the reaction set is defined based on the stoichiometric reactions given in Tab. 1.

Most industrial methanol synthesis plants produce methanol with a purity of between 95-99.9% by weight (Equinor, 2022), (Timsina R., 2020). The requirements set for the distillation column in the simulations are 99% by weight.

Other requirements are set to obtain the lowest possible amount of methanol in the bottom product and the gas flow is the overhead in the distillate. The distillation column has been simulated with 10 stages, at a pressure of 1 bar.

Since the methanol synthesis is enhanced by low temperatures and high pressures, 200°C and 100 bar are chosen for this study. The feed is set to 200 kmole/h. The composition of the feed gas for methanol synthesis is taken from experiments and simulations done at the University of South-Eastern Norway, and the ideal ratio suggested by Waugh (Waugh, 2012) is used for comparison. The gas composition data are presented in Tab. 2. For the simulation, the gases N₂, H₂O, and CH₄ are not included. A recirculation rate of 1100 kmol/h is used in the simulation.

3. Results

The gas mixtures compositions defined in Tab. 2, are used in the simulation in Aspen Hysys. All cases have been simulated with and without recirculation. The single-pass conversion of the components is defined as (Felder R, 1986):

$$\text{Single - Pass Conversion} = \frac{\text{reactant in to reactor} - \text{reactant out from reactor}}{\text{reactant into reactor}} \quad (1)$$

Table 2: Gas composition in mol% from bubbling fluidized bed gasification of wood chips. a. composition of produced gas, b. recalculated without N₂, H₂O and CH₄

	Ideal ratio	Experimental data from gasification with air (Timsina, 2022)		CPFD Simulation data for gasification with steam (Samani N.A., 2022)	
		a.	b.	a.	b.
CO	10	22.59	42.55	49.47	54.29
CO ₂	10	20.46	38.54	9.36	10.27
H ₂	80	10.04	18.91	32.3	35.44
CH ₄		7.5	-	8.66	-
H ₂ O		0	-	2.01	-
N ₂		38.4	-	0	-

The result of single conversion calculations shows - that hydrogen is the limiting reactant in gas mixtures from the gasification system. In the case of the ideal ratio, CO is the limiting reactant.

The overall conversion of the process is given in Tab. 3 and defined as (Felder R, 1986):

$$\text{Overall Conversion} = \frac{\text{reactant in to process} - \text{reactant out from process}}{\text{reactant in to process}} \quad (2)$$

For all cases, the overall conversion is high, where losses are in the purge stream 7_Purge, and the stream 8_gas, overhead gas from the distillation column.

Table 3: Overall gas conversion

	Limiting component	without recycling	with recycling
Ideal ratio	CO	98.8 %	99.9%
Gasification with air	H ₂	88.3%	95.9%
Gasification with steam	H ₂	93.6%	98.9%

Fig. 3 shows the methanol fraction in stream 5, after the reactor, as a function of temperature. The

methanol formation is higher at low temperatures; however, a low temperature will also reduce the kinetic energy, and an elevated temperature is needed to obtain the activation energy that leads to reactions between molecules. Therefore, in this study 200°C is used in the further simulations. The formation of methanol for the gas mixture from ideal ratio and gasification with steam is approximately the same at 200°C and higher temperatures, however, at lower temperatures, a high ratio of H₂/CO gives a higher formation of methanol. Fig. 4 shows the methanol fraction in stream 5, as a function of pressure.

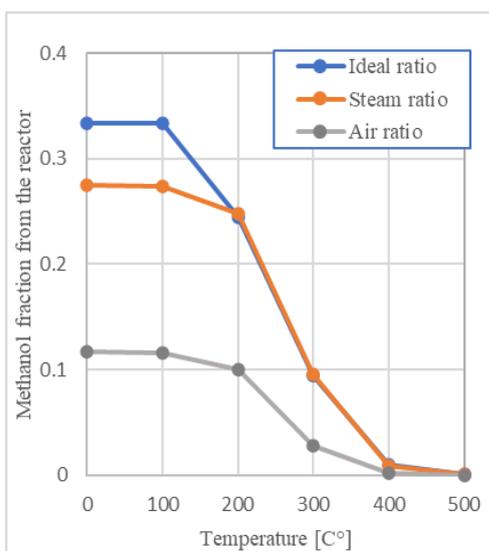


Figure 3: Mole fraction methanol from the reactor vs. temperature at 100 bar

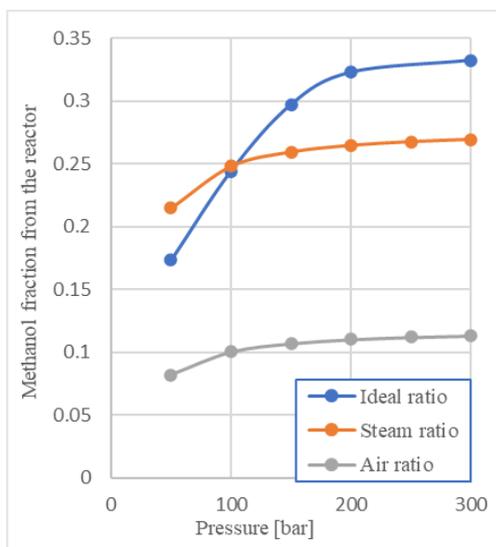


Figure 4: Mole fraction methanol from reactor vs. pressure at 200°C.

The methanol mole fraction increases with increasing pressure. However, an increase in pressure would also be considered related to the

increased cost of equipment, and energy in the compressor and cooling/heating systems. A pressure of 100 bar is used in this study. At this pressure the formation of methanol for the gas mixture from ideal ratio and gasification with steam is approximately the same, however, at higher pressures, a higher ratio of H₂/CO gives a higher formation of methanol. At lower pressure, the methanol formation is lower for the ideal ratio mixture, indicating a shift in the reaction equilibriums.

The methanol yield is defined as (Felder R, 1986):

$$\text{Yield} = \frac{\text{moles of desired products}}{\text{mols of reactant fed}} \quad (3)$$

The results are given in Tab. 4. A higher molar yield is obtained with a higher H₂/CO ratio in the gas mixture. However, the yield is approximately the same for the ideal ratio mixture and the mixture from the gasification with steam. Recirculation in the system will also enhance a higher methanol yield.

	without recycling	with recycling
Ideal ratio	15.9%	19.2%
Gasification with air	7.6 %	8.3%
Gasification with steam	16.1%	17.1%

Table 5 gives the mass flow rates of methanol produced in the different cases simulated.

	without recycling	with recycling
Ideal ratio	1016.8	1230.8
Gasification with air	488.5	532.4
Gasification with steam	1030.7	1097.5

15-35% of the reactor product, on a mole basis, is sent to the distillation column when recycling is not installed. The rest is unreacted gases that can be sent out of the system or recycled. Fig. 5 shows that the molar flow of methanol is increasing when the unreacted gas is sent to recirculation. However, the increase is limited to a recycling molar flow of around 300 kmol/h. The methanol yield will be approximately the same above this recycle flow.

Fig. 6 illustrates how the mole fraction of methanol changes when extra hydrogen is added to the system. A hydrogen feed of 150 kmole/h gives a methanol mole fraction of 0.82, for the gas mixture from the steam gasification. Also, the gas mixture from air gasification gives the highest mole fraction with an H₂ feed at 150 kmol/h. The methanol yields for these two feed mixtures are approximately 28% and

22% respectively. The H_2/CO ratio in the feed is after adding H_2 at 2,0 and 2,2 respectively.

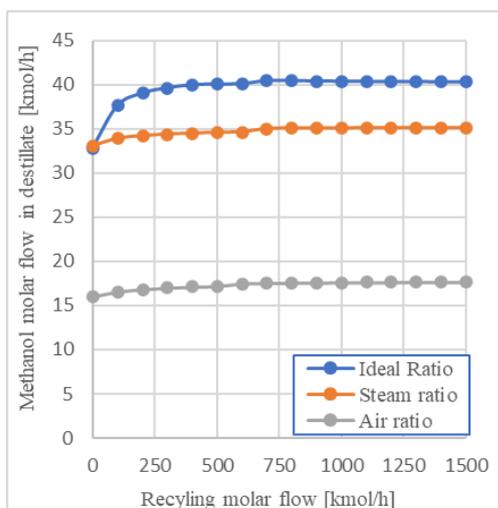


Figure 5: The molar flow of methanol in distillate vs. recycled flow.

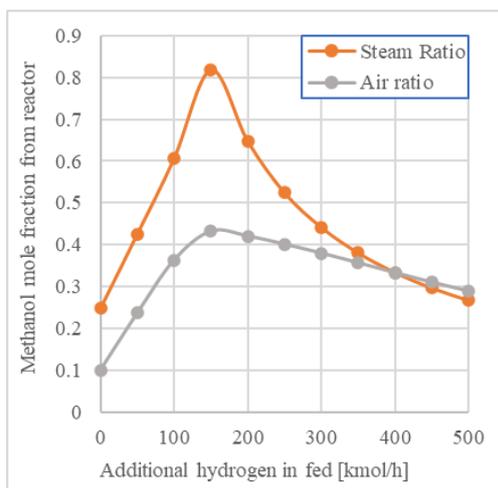


Figure 6: Mole fraction methanol from reactor vs. additional hydrogen fed to the system

4. Summary and Discussions

In this study, Aspen Hysys has been used to evaluate the conversion of gas mixtures to methanol. Gibbs reactors are based on the calculation of minimization of Gibbs free energy for the reactions and the results are based on reaction equilibrium. In a practical situation, the residence time for the reaction is short and equilibrium is not likely to occur. The results given in this study are thus an optimistic approach, and the methanol yield will be lower than estimated. Performing the simulation with a continuous stirred tank reactor (CSTR), which needs the reaction kinetics, would have given a better insight into both reaction speed, residence time, and how pressure and temperature affect the reaction. In CSTR catalytic reactions can also be studied. The results from the Gibbs reactor nevertheless give a good

indication of the process parameters, and the energy needs of the process.

The methanol synthesis is enhanced by low temperature and high pressure. In this study, different gas mixtures have been evaluated regarding temperature, pressure, and H_2/CO ratio.

From the simulation of the mole fraction of methanol, the gas mixture from gasification with air as a fluidizing agent has the lowest H_2/CO and H_2/CO_2 ratios of 0.44 and 0.49 respectively, giving the lowest methanol mole fraction from the reactor, and the lowest methanol yield. The H_2/CO ratio for the gas mixture from gasification with steam as a fluidizing agent has the H_2/CO and H_2/CO_2 ratios of 0.65 and 3.45 respectively. The H_2/CO and H_2/CO_2 ratios for the ideal mixture are both 8. The methanol mole fraction becomes the same for these two mixtures at 200°C and higher temperatures. The impact of the H_2/CO and H_2/CO_2 ratios is not significant when reaching these temperatures. For the pressure, low H_2/CO and H_2/CO_2 ratios give low methanol mole fraction out of the reactor, however increasing the gas ratios an increase in the methanol formation and yield is observed. An interesting observation is that the gas mixtures from steam gasification give a better mole fraction at pressures lower than 100 bar compared to the ideal ratio. Indicating that there is an optimal gas ratio between the two mixtures. At higher pressures high H_2/CO ratio is preferable. The methanol yields are given in Tab. 4 results from the simulation at 200°C and 100 bar, and the approximately same result for the gas mixtures from steam gasification and the ideal ratio is directly related to the analysis above. In this study a recycle of 1100 kmole/h is used in the simulation. The methanol yield reported in the results will be approximately the same with a recirculation stream above 300 kmole/h. The disadvantages of having too large recirculation streams are the increased energy requirements for cooling and heating the gas flows. Also, high dilution of stream 5 gives low methanol concentrations in the reactor product, which will be challenging to the gas-liquid separation in the two-phase separator and to the distillation columns, increasing the cost of the system.

The addition of extra hydrogen increases the methanol yield of the system. The simulation shows that an optimal addition of 150 kmole/h hydrogen gives the best result. Hydrogen cost versus the extra income for the methanol produced must be evaluated before hydrogen is chosen as feed to the system. The calculations with H_2 addition show that an H_2/CO ratio of approximately 2 is preferable.

Fluidized bed gasification with steam as gasifying agent is preferable compared to air gasification to increase the H_2/CO ratio to achieve a higher methanol yield in the methanol synthesis.

References

- Bandara, J., 2021. *Simulation and parameter optimization of fluidized-bed and biomass gasification*, Ph.D. Thesis. University of South-Eastern Norway.
- Equinor, 2022. *Tjeldbergodden Industriaregg equinor.com (in Norwegian)*. (Webpage assessed 30.05.2022)
<https://cdn.sanity.io/files/h61q9gi9/global/0cf648ed90b924c02d73e7b4e1d0fc877b5805a1.pdf?equinor-tjeldbergodden-2019.pdf>].
- Felder R, Rousseau R., 1986. *Elementary Principles of Chemical Processes*. Second Edition, John Wiley & Sons Inc.
- Haydary, J., 2018. *Chemical Process Design and Simulation: Aspen Plus and Aspen Hysys Applications.*, Wiley Online Library. <https://doi.org/10.1002/9781119311478.ch5>
- Klerk, A. d., 2020. Transport Fuel: Biomass-, Coal-, Gas- and Waste-to-Liquids Processes. *Future Energy (Third Edition)*, pp. 199-226.
- Fossen M., Halvorsrød J., Narvestad T., Tjemsland S., 2022. *Aspen HYSYS simulation of methanol synthesis based on gas from biomass gasification*, Bachelor thesis, University of South-Eastern Norway.
- Moulijn J. A., Makkee. M. and Van Diepen A. E., 2013. *Chemical Process Technology*. John Wiley & Sons, Inc.
- Timsina R., Thapa R. K., Moldestad B. M. E. and Eikeland M. S., 2022. *Methanol Synthesis from Syngas: a Process Simulation*. Proceedings of SIMS EUROSIM 2021, Virtual, Finland, 21-23 September DOI: 10.3384/ecp211854442021
- Research and Markets, 2018. Research and Markets. (Webpage assessed 30.05.2022)
<https://www.researchandmarkets.com/reports/4514928/methanol-market-growth-trends-and-forecast>.
- Samani N.A., Thapa. R.K., Moldestad. B. M. E. , Eikeland. M. S., 2022. Evaluating the impacts of temperature on a bubbling fluidized bed biomass gasification using CFPD simulation model. *MATHMOD 2022*. Vienna, Austria 27-29 July 2022
- Timsina R., Thapa R. K., Moldestad B. M. E., Eikeland M. S., 2020. Experiments and Computational Particle Fluid. *International Journal of Energy Production and Management*, Vol. 5 Iss 2, pp. 102-114.
- Timsina, R., 2022. *Modelling and simulations of bubbling fluidized bed and entrained flow biomass gasification reactors*, Ph.D Thesis, University of South-Eastern Norway.
- Waugh, K. C., 2012. *Methanol Synthesis*, Catalysis Letters, 142, pp. 1153-1166.
- Yang L., Ge. X., 2016. Biogas and Syngas Upgrading. *Advances in Bioenergy*, Volum 2016, pp. 125-188.

Mesh Sensitivity Analysis of an Entrained Flow Biomass Gasifier: A CPFD Study

Ramesh Timsina^{a, *}, Zahir Barahmand^a

^a *Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Porsgrunn*
ramesh.timsina@usn.no

Abstract

Biomass such as agricultural waste, forestry waste, municipal solid waste, and industrial waste, are renewable energy sources that may be used to produce biofuels. Biomass gasification is an effective and promising technology for converting any biomass into valuable products that can contribute considerably to renewable energy generation. In the manufacturing industry, computer-based simulations, improving production processes while incorporating sustainable industrial strategies, are rising. In the Computational Fluid Dynamics (CFD) scientific community, the reliability of computational prediction of findings is a rising problem. Mesh independence is crucial since it may determine if the solution obtained is independent of the mesh resolution. In CFD models, there are a variety of strategies for discovering a mesh independence test such as the grid resolution, general Richardson extrapolation, and Grid Convergence Index (GCI). In the grid resolution technique, the mesh size gradually increases until no meaningful performance improvement can be seen due to the larger mesh size. The present study aims to analyze the mesh independence test using the grid resolution method on an entrained flow biomass gasifier and investigate the model's sensitivity to parameters such as reactants' inlet temperature, product gas compositions and flow rate. To achieve this goal, four different scenarios were defined employing a series of Computational Particle Fluid Dynamics (CPFD) simulations using Barracuda® v21.0.1. The results confirmed that within the range of 25000 and 200000 cells, synthesis gas production decreased by almost 2 percent, which is not significant.

Keywords: mesh sensitivity analysis, mesh independence test, CPFD, gasification, entrained flow gasifier

1 Introduction

As the fourth-largest source of energy after coal, petroleum, and natural gas, biomass contributes a significant percentage of global primary energy consumption (Shah and Venkatramanan, 2019). Biomass now accounts for around 15% of total global energy use in all forms (Ankolekar and Kulkarni, 2018).

Biofuels, such as agricultural, forestry, municipal solid waste, and industrial waste, are renewable energy sources that may be used to produce solid or liquid fuels. Gasification, pyrolysis, and direct combustion are the main thermochemical conversion technologies (Pereira *et al.*, 2012), where gasification is the most efficient process (Purohit, 2009). Gasification is the partial oxidation of biomass (carbonaceous materials) at elevated temperatures to generate synthesis gas (commonly known as syngas), primarily carbon monoxide and hydrogen (Zamarripa *et al.*, 2013). The product gas from the gasification process consists of CH₄, CO, CO₂, and H₂, as well as other light gases such as ethane (C₂H₆) and propane (C₃H₈) in addition to various condensable gases. Moreover, this process produces some amounts of biochar, tars, and ashes (Pereira *et al.*, 2012). Biomass gasification is an effective and promising technology for converting

any biomass into valuable products by thermochemical conversion, which contributes considerably to renewable energy generation.

There are various developments and studies of numerical simulation tools, such as Computational Fluid Dynamics, with high prediction accuracy within a reasonable simulation time to predict such complex flows. Due to the presence of three-phase in the systems, entrained flow reactors are the most challenging and most complex systems in multiphase modelling. The CFD scientific community is becoming increasingly concerned with the accuracy of computational outcomes prediction. Therefore, several important concerns arise (Seeni *et al.*, 2021): Are computational results reliable? How can the accuracy or validity of CFD predictions be evaluated? These questions are posed because of the uncertainty associated with CFD-generated data. As a result, procedures for verification and validation have been created to address this developing issue. Code and solution verification are both components of verification. Analytical, very precise hybrid analytical-numerical, and manufactured solutions to mathematical models can be used for verification. Validation often entails determining the accuracy of a mathematical model's representation of the

physical processes of interest using carefully designed and conducted experimental data (Baliga and Likhmanets, 2016).

A mesh independence study (or so-called mesh sensitivity) determines whether or not simulation results are independent of the underlying mesh (McDavid, 2001). There are several methods for determining a mesh independence test in CFD issues. The grid resolution, the general Richardson extrapolation, and the Grid Convergence Index (GCI) are three often utilized methods (Seeni et al., 2021). Grid Convergence Index - GCI is a technique for estimating discretization error even when subsequent mesh refinements are not integer multiples (Castedo *et al.*, 2019). Richardson's extrapolation is a numerical analysis approach for predicting the error in the answer by solving the issue with two alternative grid sizes, assuming the solution's functional form is known (Rao, 2001). In the grid resolution technique, the mesh size gradually increases until the performance improvement cannot be noticed due to the increased mesh size.

The current study aims to conduct a mesh sensitivity analysis using the grid resolution method on an entrained flow biomass gasifier and investigate the model's sensitivity for different parameters such as molar concentrations, flow rate, and temperature.

The paper is organized as follows: Chapter 2 provides a brief background, and Chapter 3 provides a mathematical description of the CPFDF model. Chapter 4 details the developed computational model for the selected gasifier and simulation setup. The study's outcome is described in Chapter 5, and the conclusion is drawn at last.

2 Background

There are different CFD models available in the literature to study an entrained flow biomass gasification reactor at different levels of accuracy and depth (Fletcher et al., 1998, 2000; X. Gao et al., 2018; X. Y. Gao et al., 2014; Ku et al., 2014; Slezak et al., 2010). Most models account for the chemical and physical properties inside the gasifier, commonly known as the non-equilibrium/kinetic model. The models are generally validated using data from a certain gasifier. This can create a certain room for uncertainty as the operation of commercial gasifiers could vary significantly. Variations in various aspects include air-blown, oxygen-blown, non-pressurized, pressurized, single-stage, multiple-stage, dry feed, slurry feed, swirling flow, non-swirling flows, refractory insulations, etc. are the common example. The current study aims to develop a CPFDF model considering these different aspects to simulate an entrained flow biomass gasification reactor within a reasonable time frame with certain efforts and accuracy. Fig. 1 shows the

relationship between time, effort, and accuracy for any computational model.



Figure 1: Inter-relation between time, efforts, and accuracy for CFD models.

The higher the time and effort invested while developing a CFD model, the higher the model's accuracy and vice versa. Therefore, a compromise must be made between efforts to build the model, accuracy of calculations, and calculation time (Kaczor *et al.*, 2020). In addition to this, model validation is an important aspect of their applications. A validated CFD model becomes a credible engineering tool for further applications.

3 Mathematical modelling equations

The continuity gives the gas-phase mass and momentum conservation equations, and the Navier-Stokes equations are represented by Eq. (1) and (2), respectively.

$$\frac{\partial(\alpha_g \rho_g)}{\partial t} + \nabla \cdot (\alpha_g \rho_g \vec{u}_g) = \delta \dot{m}_p \quad (1)$$

$$\begin{aligned} \frac{\partial}{\partial t} (\alpha_g \rho_g \vec{u}_g) + \nabla \cdot (\alpha_g \rho_g \vec{u}_g \vec{u}_g) \\ = -\nabla p + F + \alpha_g \rho_g g \\ + \nabla \cdot (\alpha_g \tau_g) \end{aligned} \quad (2)$$

where α , ρ and \vec{u} represent the volume fraction, density, and velocity vector, respectively. $\delta \dot{m}_p$ is the gas mass production rate per volume formed from the particle-gas chemical reaction. In the case of the cold flow model with no chemical reaction, $\delta \dot{m}_p$ becomes zero. P is the mean flow gas pressure, g is the acceleration due to gravity, τ_g is the fluid phase stress tensor, and F is the inter-phase momentum transfer rate per unit volume (particle to fluid phase). For a Newtonian fluid, the gas phase stress tensor for each species, τ_g is given by:

$$\tau_{g,ij} = \mu \left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} \mu \delta_{ij} \frac{\partial u_k}{\partial x_k} \right] \quad (3)$$

where μ is the shear viscosity, which is the sum of the laminar shear viscosity and the turbulence viscosity defined in the Smagorinsky turbulence model (Smagorinsky, 1963), the model is given in equation 4 (Snider *et al.*, 2011).

$$\mu_t = C_s \rho_g \Delta^2 \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (4)$$

The Smagorinsky coefficient C_s has a default value of 0.01. Δ is the subgrid length and is given by:

$$\Delta = (\delta x \delta y \delta z)^{1/3} \quad (5)$$

The interphase momentum transfer (F) in Equation 2 is given by:

$$F = \iiint f \left[m_p \left\{ D_p (\vec{u}_g - \vec{u}_p) - \frac{\nabla P}{\rho_p} \right\} + \vec{u}_p \frac{dm_p}{dt} \right] dm_p d\vec{u}_p dT_p \quad (6)$$

In terms of f , the fluid mass source in Eq. (1) is:

$$\dot{m}_p = - \iiint f \frac{dm_p}{dt} dm_p d\vec{u}_p dT_p \quad (7)$$

where the time-rate-of-change of particle mass dm_p/dt is the rate of change of the particle mass produced by chemical processes. The particle acceleration can be calculated by:

$$\frac{d\vec{u}_p}{dt} = D_p (\vec{u}_g - \vec{u}_p) \frac{\nabla P}{\rho_p} - \frac{\nabla \tau_p}{\alpha_p \rho_p} + g + \frac{\vec{u}_p - \vec{u}_p}{\tau_D} \quad (8)$$

where ρ_p is the particle mass density, τ_p which can be derived by Eq. (9) is the solids contact stress, which depends on spatial location, D_p is the drag function, which depends on the particle size, velocity, position, and time. \vec{u}_p is the local mass-averaged particle velocity. τ_D is a particle collision damping time (O'Rourke and Snider, 2010).

$$\tau_p = \frac{P_s \alpha_p^\beta}{\max[(\alpha_{cp} - \alpha_p), \varepsilon(1 - \alpha_p)]} \quad (1)$$

Particle normal stress is exerted on a solid until the solid reaches the particle-mean velocity (Snider et al., 2011). P_s is a constant (Pa), α_{cp} is the particle volume fraction at close packing, β is a constant (between 2 – 5) and ε is a very small number in the order of 10^{-8} . The solids volume fraction is related to the PDF f by:

$$\alpha_p = \iiint f \frac{m_p}{\rho_p} dm_p d\vec{u}_p dT_p \quad (2)$$

4 Simulation setup

The reactor dimensions as well as the reaction kinetics were adopted from the author's previous studies (Timsina et al., 2021; Timsina et al., 2020). As shown in Fig. 2, the reactor has 0.52 m in diameter with a conical outlet and 1.67 m in height. The biomass and the fluidizing agent were modelled

as injection boundaries at the top, and a pressure boundary was defined at the bottom for product outflow from the reactor. The details of the injection boundary can be found in the author's previous study. Barracuda® v21.0.1 was the CFD software used for the simulations (Software®, 2022).

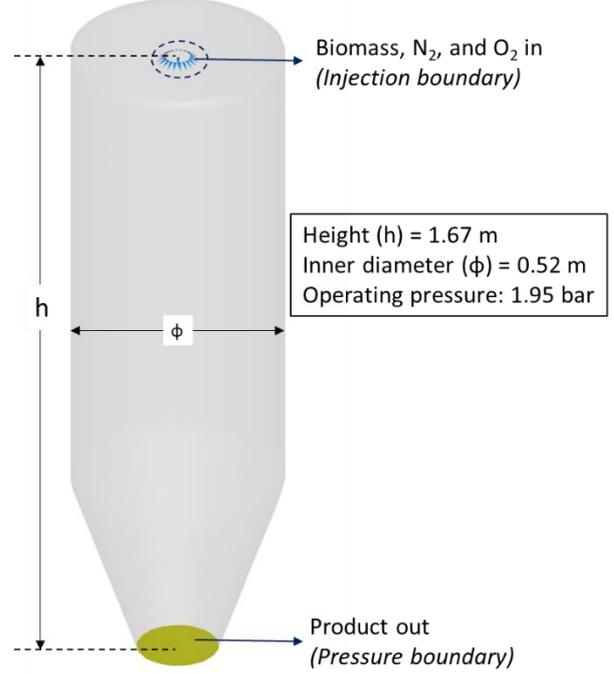


Figure 2: Boundary conditions

The Wen-Yu drag model was chosen due to the dilute solid phase where the gas volume fraction is higher than 0.8 (Cho et al., 2020; Jayarathna et al., 2019; Patel et al., 1993), and the chosen particle parameters are presented in Tab. 1.

Table 1: Particle phase model parameter

Drag model	Wen-Yu
Closed pack volume fraction	0.2
Maximum momentum redirection from collision	40%
Tangent to wall momentum retention	0.85
Normal to wall momentum retention	0.15

An injection boundary was chosen for the introduction of biomass particles into the reactor as it does not need an assistance of a fluid stream. The blue triangles with spheres at the top represent the injection points. A total of 20 injection points along the circle and one in the middle were defined in the model (Timsina et al., 2021).

The grid dimensions in x, y, and z directions for each mesh are given in Tab. 2. The normalized grid size was checked to ensure they lie below the warning line. The grid refinements at the wall were not performed for all the meshes to have uniformity in simulation conditions. The built-in grid generator was used in Barracuda to generate the grid, and the

cells having a volume fraction less than 0.04 and aspect ratio higher than 5:1 were neglected. As a result, the Cartesian mesh of 25000 (25k), 50000 (50k), 100000 (100k), and 200000 (200k) cells gave the corresponding number of cells in Tab. 2.

Table 2: cell dimensions

No. of cells	ΔX	ΔY	ΔZ	Case
22743	19	19	63	1
48080	24	24	80	2
97061	31	31	101	3
193167	39	39	127	4

Four different meshes with 22743, 48080, 97061, and 193167 cells were tested, and the cross-sectional views are illustrated in Fig. 3. From here onwards, the four cases will be mentioned as 25k cells, 50k cells, 100k cells, and 200k cells for case numbers 1, 2, 3, and 4, respectively.

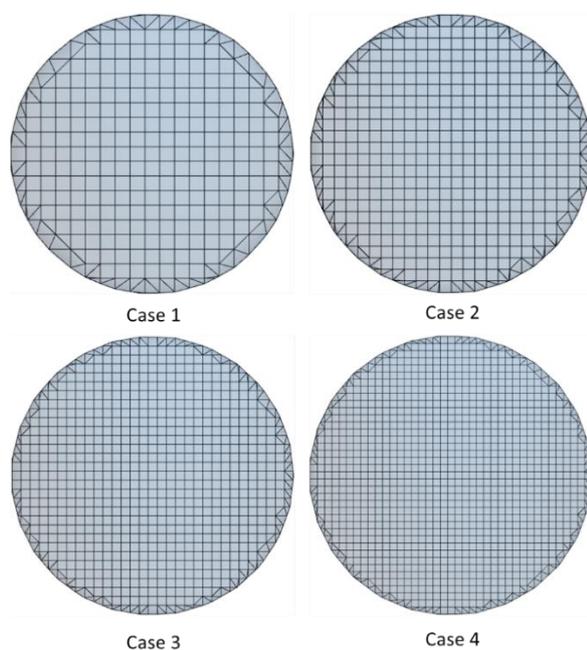


Figure 3: Cross-sectional view of different grids (top view)

5 Results and Discussions

Simulations were performed for four different grids with 25k, 50k, 100k, and 200k cells. In the grid resolution technique, the mesh size gradually increases until no meaningful performance improvement can be seen due to the larger mesh size. The computation is a three-dimensional non-isothermal with homogeneous and heterogeneous gasification chemistry. Simulations were carried out for 300 seconds of simulation time with a number density of 125000, and the gas composition, temperature, residence time, and flow rates were monitored. Average gas compositions were taken from the final 150 seconds of simulations. The

bottom plane of the reactor gives the product gas from the reactor.

It is important to monitor the different grids' fluid temperature along the reactor. Fig. 4 shows the average fluid temperature along the gasifier. The temperature value was radially averaged at $t=150s$. The figure shows the highest fluid temperature at the reactor injection burner, suggesting that some degree of combustion prevails around this region of the gasifier. This is beneficial as it supplies the generated heat to the devolatilization of the biomass. The average fluid temperature in this region for 25k and 50k cells seems to fluctuate more than the 100k and 200k cells. This gives some stability to the reactor temperature profile in the burner region of the gasifier. The average temperature is almost the same for all the cases as we move downwards along the gasifier.

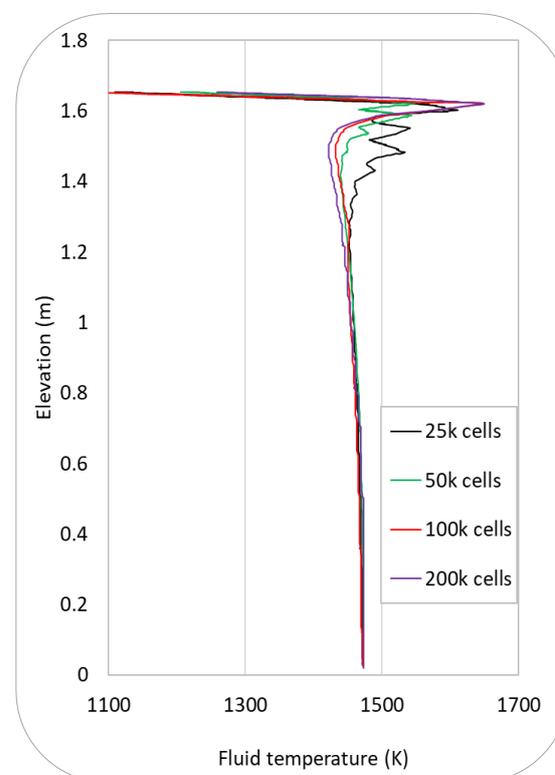


Figure 4: fluid temperature versus elevation. The fluid temperature is radially averaged.

As the thermal and chemical behavior are coupled together in such systems, a change in one parameter affects the changes in the other ones. Heat is supplied from reactor walls, and the gas is fed into a biomass gasifier. Chemical transformation, such as breaking chemical bonds, gives sensible heat, which changes the temperature. Therefore, product gas composition and flow rates were monitored for all the cases during the simulations. Fig. 5 shows the product gas flow rate from the outlet boundary of the reactor. After the reactor reached the steady state, which took around 50 seconds, the mass flow rate of

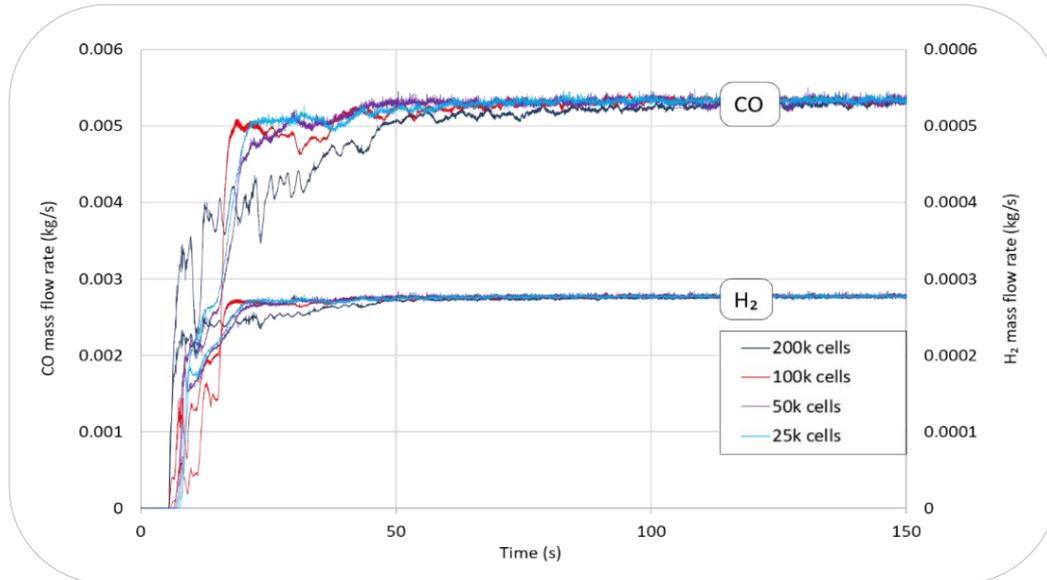


Figure 5: carbon monoxide and hydrogen mass flow rate at the outlet boundary

CO and H₂ stabilized around their median value. Production of gases starts after around 5 seconds of simulation. There was a high fraction of the CO₂ at the start representing the combustion process; however, after the steady-state, the average composition for all the gas species varied around their mean value. To look closely, only CO and H₂ are taken to compare the four different grids. A certain level of variation at steady in the gas production illustrates different physical and chemical transformations occurring inside the reactor. To compare the different cases, the average molar composition of major gas species (CO, H₂, and CH₄) is compared in Tab. 3.

Table 3: average molar concentrations

Case	Average molar concentration			% Variation wrt 25k cells		
	CO	H ₂	CH ₄	CO	H ₂	CH ₄
1	0.20	0.01	0.01	-	-	-
2	0.20	0.01	0.01	-0.32	0.16	-0.92
3	0.20	0.01	0.01	-0.52	0.08	-0.92
4	0.20	0.01	0.01	-1.72	-0.22	-2.09

As the data show that there is little variation between the four different cases. The absolute percentage variation of the molar composition of CO, H₂, and CH₄ varies from 0.32% to 2.1%. This shows a minor and not significant variation, and there should be a trade-off between the number of cells and the simulation timing.

Looking only at the reactor hydrodynamics can be tricky when doing resolution studies because the number of pixels drawn in Tecplot (an output result viewer for Barracuda) corresponds to the number of

clouds in the simulation. This makes the high-resolution cases look much denser even if they have the same volume fraction and mass of particles. This limitation of the Tecplot leads us to present the simulation timing for the considered cases. The simulation time increased exponentially with the increase in the number of cells in the system. The simulation time in hours (*t*) varied according to the $t = 24.185e^{0.00002n}$, where *n* represents the total number of cells (25000, 50000, 100000 & 200000). In numbers, the corresponding simulation time were 29, 78, 230 and 1100 hrs respectively for the case 1, 2, 3, and 4 respectively. In order to visualize the simulation timing in terms of simulation results, CFL is monitored for all the cases. The timestep for any transient CFD model is an important parameter. The time step must be small enough to represent any rapidly changing variables of interest. If the time step is too big, an accumulation of errors will occur (Zhang *et al.*, 2000). To resolve this problem, a varying time step can be utilized with the help of the Courant-Friedrichs-Lewy (CFL) number, as shown in Eq. (11):

$$CFL = \frac{v \Delta t}{\Delta x} \quad (3)$$

where *v* is velocity, Δt is time step and Δx is cell size. Fig. 6 shows the CFL for all the cases.

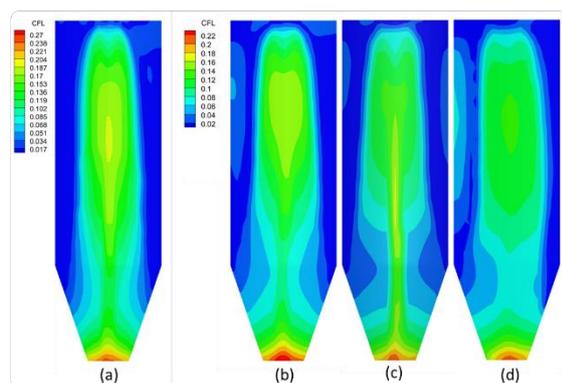


Figure 6: CFL for cases with (a) 200k, (b) 100k, (c) 50k and (d) 25k cells

With an increase in cell size, CFL seems to increase linearly. However, the velocity of the particle plays an important role. As shown in the figure, CFL decreased with an increase in the number of cells. CFL is comparatively lower for case number four compared to other cases. The higher the number of cells, the more continuous the CFL profile inside the reactor. Courant–Friedrichs–Lewy number gives the amount of information traveling across a computational cell in a unit of time. CFL number higher than one gives an inaccurate solution and could potentially lead to divergence of the results. The timestep is an important aspect during a CFD simulation and should be selected carefully such that all the physics of interested parameters are resolved in that time step. The additional cells are allowing the model to capture more details of the flow profile. However, the 100k-cell resolution looks similar to that of 200k-cell resolution, so that the 100k-cell case could be sufficient in terms of cell resolution.

6 Conclusions

A CPFD model was developed in Barracuda® v21.0.1 using the MP-PIC modelling approach. The model was used to simulate a biomass gasification process in an entrained flow gasifier. To conduct a mesh sensitivity analysis through a mesh resolution technique, the model was simulated with four different numbers of cells (from 25k to 200k) to see the effect of grid resolution in terms of molar concentrations, flow rate, and temperature along the gasifier. Although there was a considerable difference in hydrodynamics inside the reactor in different cases, the results show minor variations (maximum 2 percent) in main output variables such as reactor temperature, gaseous product flow rate, and composition. Therefore, it can be concluded that the system was not sensitive to the number of cells within the selected mesh sizes. For further studies, it is suggested to perform this sensitivity analysis with higher resolutions (over millions); however, it will be costly.

Acknowledgment

The authors would like to thank the University of South-Eastern Norway (USN) for providing the necessary software for the simulation works.

References

- Ankolekar, V., and Kulkarni, S. (2018). Briquetting of agricultural biomass: an overview. *International Journal for Research in Applied Science and Engineering*, 6(2), 1681-1685.
- Baliga, B. R. R., and Likhmanets, I. Y. (2016). Generalized Richardson extrapolation procedures for estimating grid-independent numerical solutions. *International Journal of Numerical Methods for Heat & Fluid Flow*.
- Castedo, R., Reifarth, C., Santos, A. P., Losada, J. J., López, L. M., Chiquito, M., and Mancilla, J. M. (2019). Application of grid convergence index to shock wave validated with LS-DYNA and ProsAir. *Ingeniería e Investigación*, 39(3), 20-26. doi:10.15446/ing.investig.v39n3.81380
- Cho, S., Park, C., Lee, J., Lyu, B., and Moon, I. (2020). Finding the best operating condition in a novel process for explosive waste incineration using fluidized bed reactors. *Computers & Chemical Engineering*, 107054. doi:10.1016/j.compchemeng.2020.107054
- Jayarathna, C. K., Balfe, M., Moldestad, B. M., and Tokheim, L.-A. (2019). Improved multi-stage cross-flow fluidized bed classifier. *Powder Technology*, 342, 621-629. doi:10.1016/j.powtec.2018.10.026
- Kaczor, Z., Buliński, Z., and Werle, S. (2020). Modelling approaches to waste biomass pyrolysis: a review. *Renewable Energy*, 159, 427-443.
- McDavid, R. M. (2001). Macroscopic Modeling. In K. H. J. Buschow, R. W. Cahn, M. C. Flemings, B. Ilshner, E. J. Kramer, S. Mahajan, & P. Veyssièrè (Eds.), *Encyclopedia of Materials: Science and Technology* (pp. 4720-4732). Oxford: Elsevier.
- Patel, M., Pericleous, K., and Cross, M. (1993). Numerical modelling of circulating fluidized beds. *International Journal of Computational Fluid Dynamics*, 1(2), 161-176. doi:10.1080/10618569308904470
- Pereira, E. G., Da Silva, J. N., de Oliveira, J. L., and Machado, C. S. (2012). Sustainable energy: a review of gasification technologies. *Renewable and sustainable energy reviews*, 16(7), 4753-4762.
- Purohit, P. (2009). Economic potential of biomass gasification projects under clean development mechanism in India. *Journal of Cleaner Production*, 17(2), 181-193.
- Rao, S. S. (2001). FINITE ELEMENT METHODS. In S. Braun (Ed.), *Encyclopedia of Vibration* (pp. 530-544). Oxford: Elsevier.
- Seeni, A., Rajendran, P., and Mamat, H. (2021). A CFD Mesh Independent Solution Technique for Low Reynolds Number Propeller. *CFD Letters*, 11(10), 15-30. Retrieved from <https://www.akademiabaru.com/submit/index.php/cfdl/article/view/3187>
- Shah, S., and Venkatramanan, V. (2019). Advances in microbial technology for upscaling sustainable biofuel production. In *New and future developments in microbial biotechnology and bioengineering* (pp. 69-76): Elsevier.
- Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Monthly weather review*, 91(3), 99-164. doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2
- Snider, D. M., Clark, S. M., and O'Rourke, P. J. (2011). Eulerian–Lagrangian method for three-dimensional thermal reacting flow with application to coal gasifiers. *Chemical engineering science*, 66(6), 1285-1295. doi:10.1016/j.ces.2010.12.042

Software[®]. C. (2022). CPFD Software. Retrieved from <https://cpfd-software.com/>

Timsina, R., Thapa, R. K., Moldestad, B. M., and Eikeland, M. S. (2021). Computational particle fluid dynamics simulation of biomass gasification in an entrained flow gasifier. *Chemical Engineering Science: X*, 12, 100112.

Timsina, R., Thapa, R. K., Moldestad, B. M. E., and Eikeland, M. S. (2020). *Simulation of entrained flow gasification reactor with Multi Phase Particle in Cell (MP-PIC) approach*. Paper presented at the 61st SIMS Conference on Simulation and Modelling SIMS 2020, Online.

Zamarripa, M., Hjaila, K., Silvente, J., and Espuña, A. (2013). Simplified model for integrated Supply Chains Planning. In *Computer Aided Chemical Engineering* (Vol. 32, pp. 547-552): Elsevier.

Zhang, G.-Q., Ernst, L. J., and de Saint Leger, O. (2000). *Benefiting from thermal and mechanical simulation in micro-electronics* (1st ed.). Netherlands: Springer.

Enhanced oil recovery using CO₂ injection and inflow control devices

Stian Juvet Sørensen^a, Joachim Rød Knarrum^a, Ole Christian Kristoffersen Sannes^a, Lars Lia^a, Ramesh Timsina^a, Haavard Aakre^b, Britt M. E. Moldestad^a

^a Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Norway

^b InflowControl AS, Norway

britt.moldestad@usn.no

Abstract

CO₂ capture is an important aid to achieve the goals of reduced emissions. Today the market for captured CO₂ is limited. Injection of CO₂ in oil reservoirs is one way of utilizing the captured CO₂ and provide permanent storage of the CO₂. By injecting CO₂, the mobility of oil is increased, and the amount of residual oil is reduced. This is called enhanced oil recovery. The goal of this study is to show the effects of CO₂ injection combined with autonomous inflow control valves on oil production. Simulations are carried out using the software OLGA in combination with ROCX. The input data to the simulations are based on information obtained from a literature study. To show the effect of the autonomous inflow control valves and CO₂ injection, the simulations were compared with simulations performed without autonomous inflow control valves and CO₂ injection. The results from the simulations show that CO₂ injection contributes to increase the mobility of both oil and water which leads to an increase in both oil and water/CO₂ production. Autonomous inflow control valves reduce the amount of water produced by choking the production in areas with water breakthrough. The combination of CO₂ injection and autonomous inflow control valves results in a higher oil-water ratio and a considerably lower water production.

1. Introduction

Due to global warming, countries around the world have signed a climate agreement which obliges the countries to reduce the emission of climate gases, among them CO₂. CO₂ can be captured from power plants, and the captured CO₂ can further be injected into oil reservoirs or aquifers to reduce the CO₂ emission to the atmosphere. Injection of CO₂ into oil reservoirs increases the oil recovery and at the same time, the CO₂ can be permanently stored. CO₂ capture in combination with CO₂ enhanced oil recovery (CO₂-EOR) and CO₂ storage is called Carbon Capture, Utilization and Storage (CCUS), and is a promising way to reduce the CO₂ emission to the atmosphere. CO₂-EOR is used in US and Canada, but mainly by using CO₂ from natural deposits. So far, CO₂-EOR is not utilized on the Norwegian shelf. However, a mapping including 46 oil fields has been carried out on the Norwegian shelf, and it is indicated that the potential for increased oil recovery using EOR methods, including CO₂-EOR, is 700 MSm³ [1]. The Norwegian Petroleum Directorate has also mapped the potential of CO₂ storage on the Norwegian shelf and has concluded that the storage potential is more than 80 billion tonnes of CO₂ [2]. This corresponds to today's Norwegian CO₂ emissions for 1600 years.

When CO₂ is injected into a reservoir, the CO₂ is physically mixed with the oil and the properties of oil are changed. CO₂ reduces the oil viscosity and in addition it has a swelling effect which increases the oil volume in the reservoir pores [3, 4, 5]. The effect of CO₂ injection is most significant for heavy oils. The solubility of CO₂ in oil increases with decreasing temperatures, and the effect of CO₂ injection is therefore highest in reservoirs with moderate temperatures. Injection of CO₂ can contribute to decrease the oil viscosity up to 25% [6]. A combination of the effect of reduced oil viscosity and the swelling contributes to increase the mobility of oil in the reservoir, and thereby decrease the residual oil significantly [7]. Laboratory tests with CO₂ injection in core samples from the Oseberg field, have shown that the residual oil saturation can be reduced to 0.1 [8]. Fig. 1 gives a schematic of the transition zone of CO₂ between the injection and the production well. Supercritical CO₂ is injected into the oil reservoir at high pressure. Hydrocarbons from the reservoir oil vaporize into the CO₂ and a part of the injected CO₂ dissolves into the oil. The two phases become completely miscible without any interface effects and contribute to develop a transition zone that is miscible with oil in the front and with CO₂ in the back [9].

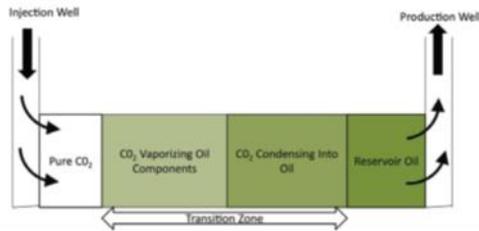


Figure 1: The schematic of the CO₂ transition zone between the injection and production well [9].

One of the main problems related to CO₂-EOR is a direct breakthrough of CO₂ or CO₂ solved in water into the production well. The breakthrough can be due to fractures and zones with high permeability in the reservoir, or due to frictional pressure drop in long horizontal wells. Fig. 2 shows a scenario where CO₂ is injected into a reservoir and flows together with water directly to the production well without being distributed in the reservoir and mixed with the oil. To avoid this type of breakthrough, different types of inflow control devices can be installed in the production well. Inflow control devices have the option to equalize the production rates along a horizontal well as shown in Fig. 3. This will solve or reduce the early breakthrough problem of CO₂ into the production well, and thereby make the CO₂-EOR process more energy effective.

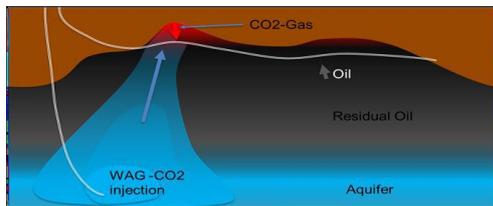


Figure 2: Short circuiting of CO₂ between the injector and the production well.

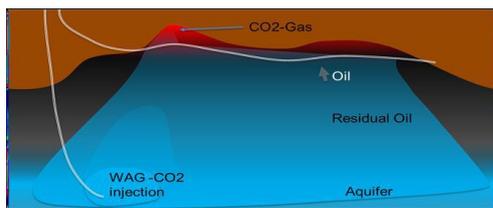


Figure 3: Improved distribution of CO₂ in the reservoir by using inflow control devices.

Different types of inflow control devices are developed to delay or avoid early breakthrough of unwanted fluids like water and gas/CO₂ into the production wells. The inflow control devices can be divided into two main categories, passive inflow control devices (ICDs) and autonomous inflow control devices (AICD). In this study, an orifice ICD and an autonomous inflow control valve (AICV) are used to delay or avoid direct breakthrough of CO₂ and water to the production well. The description and functionality of ICD and AICV are presented in the literature [10, 11, 12, 13]. ICDs are commonly

used in oil fields all over the world to delay water and gas breakthrough. AICVs have so far mainly been installed in Canada, USA, and the Middle East, and have also been tested in a CO₂ EOR well in Canada [14, 15]. There is a lack of production data available regarding CO₂-EOR and inflow control devices. Aakre *et al.* [15] simulated the effect of AICV in a vertical CO₂-EOR well in Canada and compared the results with real well data. Hansen and Moldestad [16] simulated oil production in a field with CO₂-EOR and horizontal wells completed with ICDs and AICVs. Water/CO₂ breakthrough usually occurs first in zones with high drawdown or with higher permeability than other zones along the production well. To avoid water/CO₂ to reach the other zones via annulus, zone isolation is used by installing packers between the different sections. Fig. 4 shows a schematic of the production line including packers, sand screen and ICDs/AICVs.

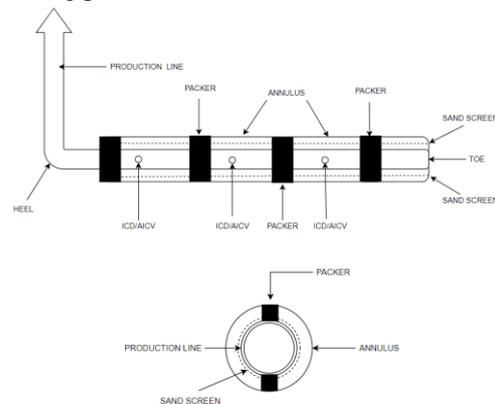


Figure 4: Production line with packers, sand screen and ICDs/AICVs.

2. Methodology

2.1. Simulation tools

The simulations have been performed using the simulation tool OLGA in combination with ROCX. OLGA is a software developed to simulate multiphase fluid flow in networks of wells, flowlines, pipelines and process equipment. ROCX is a near-well reservoir simulator and can be combined with OLGA. Due to the coupling between OLGA and ROCX the dynamic interactions between the wellbore and the reservoir are considered [17]. Fig. 5 shows an overview of inputs needed for the simulations of oil, gas, and water production from a reservoir.

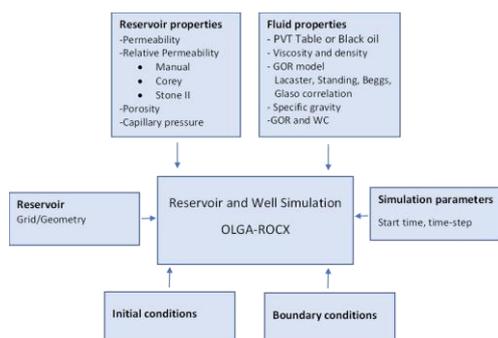


Figure 5: Overview of input to combined OLGA and ROCX simulations.

2.1.1. ROCX

ROCX calculates the fluid flow in a porous medium based on the conservation equations for water, oil and gas. Reservoir properties such as porosity, fluid saturation, permeability, relative permeability, capillary pressure, temperature, and pressure are implemented in ROCX. The required fluid properties are viscosity, density, bubble point and gas/oil ratio. Initial and boundary conditions are set for the reservoir and the sources coupled to OLGA. ROCX receives information from OLGA regarding pressure and pressure drop in the well and through the inflow control devices. Based on the reservoir and fluid information and the information given in OLGA, the production rates into the wellbore are calculated for each phase [17].

The simulations have been performed in two different types of reservoirs: a homogeneous reservoir and a heterogeneous reservoir. The size of the reservoir is 1000 m in x-direction (length), 100 m in y-direction (width) and 20 m in z-direction (height). The reservoir is divided into 2100 cells, 10 in the x-direction, 21 in the y-direction and 10 in the z-direction. The size of the cells is constant in the x- and z-direction, 100 m and 2 m respectively. In the y-direction the cell sizes are reduced from 10 to 1 m towards the well. The horizontal well is located along the x-axis 2 m from the top of the reservoir and in the middle (cell 11) in the y-direction. The porosity of the reservoirs was set to a constant value of 0.3. The viscosity of oil was 2 cP for the case without CO₂ injection and was reduced to 1.5 in the cases with CO₂ injection to simulate the effect of CO₂ on viscosity. The reservoir pressure and temperature were 130 bar and 100°C for all the simulations. Simulations were carried out in a homogeneous and a heterogeneous reservoir. The permeability in the homogeneous reservoir was 1000 mD in the x- and y-direction and 100 mD in the z-direction. In the heterogeneous reservoir, a zone with permeability 10000 mD in x- and y-direction and 1000 mD in z-direction was defined in the heel section of the reservoir. Fig. 6 shows the permeability distribution in the heterogeneous reservoir.

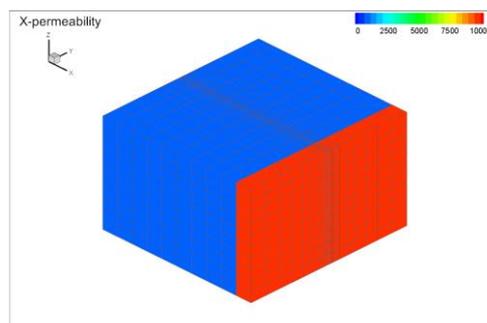


Figure 6: Permeability zones in heterogeneous reservoir.

Tab. 1 shows the input data to ROCX for the basecase and the case with CO₂ injection.

Table 1: Input data to Rocx.

Input data	Basecase	CO ₂ -EOR
Reservoir pressure	130 bar	130 bar
Reservoir temperature	100°C	100°C
Pressure, heel (boundary)	120 bar	120 bar
Porosity	0.3	0.3
Oil viscosity	2 cP	1.5 cP

The relative permeability curves for the basecase and the case with CO₂-EOR are presented in Fig. 7. The input data for calculation of the relative permeability curves are presented in Tab. 2. The relative permeability curves are calculated based on the Corey equation for the water and the Stone II correlation for the oil. The relative permeability curves when CO₂ is injected in the reservoir are modified based on information presented in the introduction.

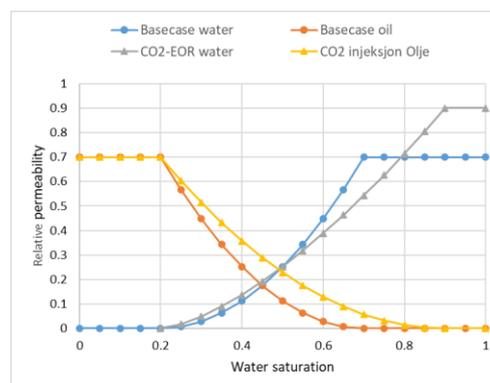


Figure 7: Relative permeability curves for basecase and CO₂-EOR case.

Table 2: Input data for relative permeability curves.

Input data	Basecase	CO ₂ -EOR
S_{wc}	0.2	0.2
S_{or}	0.3	0.1
K_{rwoc}	0.7	0.9
K_{rowc}	0.7	0.7
n_w	2	1.5
n_{ow}	2	2

S_{wc} is the irreducible water saturation, S_{or} is the residual oil saturation, K_{rwoC} is the end point relative permeability for water at maximum water saturation, K_{rowC} is the endpoint relative permeability for oil in water at irreducible water saturation, n_w is the Correy exponent, and n_{ow} is a fitting parameter for oil [18].

2.1.2. OLGA

OLGA is a transient dynamic multi-phase simulator used to predict flow in pipelines and connected equipment. The OLGA simulator is governed by conservation of mass, momentum, and energy [19, 20].

The simulations are carried out using ICD and AICV completion. Fig. 8 shows one section of the well set-up with ICD completion. To be able to include the effect of zone isolation, the set-up includes two pipelines. The upper one is the production pipe, and the lower one is illustrating the annulus and the transition of fluids from the annulus via the ICD to the production pipe. The source represents the fluid flow from the reservoir, and the leak represents the fluid flow to the well. The packer is represented by a closed valve. The total set-up includes 10 sections all including Packer, Source, ICD and Leak. Packers are used to isolate the production zones from each other, and thereby avoid fluids to flow from one zone to another through annulus.

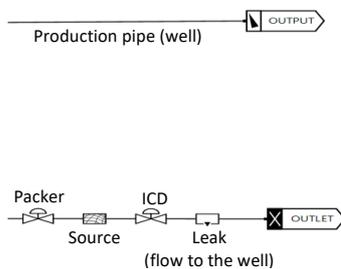


Figure 8: Set-up for ICD in OLGA.

Fig. 9 shows two sections of the OLGA-setup for a well with AICV completion. The installations Source, Packer and Leak are defined in the same way as for the ICD case. The AICV is presented by a control valve with transmitter and PID controller. Tab. 3 shows the input data to OLGA. The PID controller starts to close the valve when the water cut (WC) is 75%. The AICV closes gradually from 100% to 1% opening.

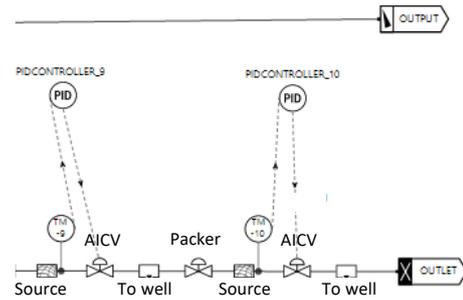


Figure 9: Set-up for AICV in OLGA.

Table 3: Input data to OLGA.

Parameters	
ICD/AICV diameter	0.02 m
Leak diameter	0.035 m
Annulus diameter	0.15 m
Set point AICV (WC)	0.75
Discharge coefficient ICD/AICV	0.84
Length production well	992 m
AICV opening (closed position)	1%
Number of ICDs/AICVs	10

3. Results

Simulations are performed for a homogeneous reservoir with and without CO₂ EOR. In addition, simulations of a heterogeneous reservoir with CO₂ EOR are carried out. These simulations are performed using wells with ICD and AICV completion. The CO₂ is assumed solved in water, and when water is mentioned it also includes CO₂.

3.1. Homogeneous reservoir

Two cases were simulated for the homogeneous reservoir, one basecase and one case with CO₂ EOR. In both cases, the production well is completed with ICDs. Fig. 10 shows the water cut versus time for the two cases.

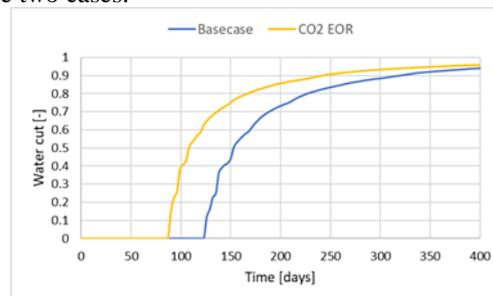


Figure 10: Water cut versus time for Base-case and CO₂-EOR case

The water breakthrough occurs after 90 days for the CO₂-EOR case and after 126 days for the basecase. The associated oil production versus time is presented in Fig. 11.

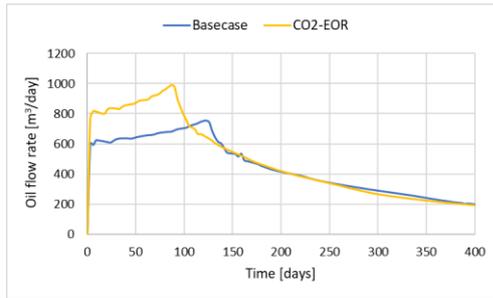


Figure 11: Comparison of oil flow rates for basecase and CO₂-EOR case

Initially, the oil production is significantly higher when using CO₂-EOR compared to the basecase. However, after water breakthrough, the oil production decreases, and after about 111 days the oil production for the CO₂-EOR case is lower than the oil production rate in the basecase. After 150 days, the oil production is about equal for the two cases. As can be seen from Fig. 10, the water cut increases when CO₂ is injected into the reservoir. Tab. 4 gives an overview of the production data for the basecase and the CO₂-EOR case.

Table 4: Production data

	Basecase	CO ₂ -EOR
Accumulated oil	190000 m ³	210000 m ³
Accumulated water	430000 m ³	702000 m ³
Water breakthrough	126 days	90 days
WC (400 days)	94%	96%
ΔP toe-heel (400 days)	4 bar	4.6 bar

The accumulated oil production increased by 10% and the water production increased by 63% when CO₂ was injected into the reservoir. The increase in oil production is because CO₂ reduces the oil viscosity, and the oil becomes more mobile as shown in Fig. 11. When the oil mobility increases, the water will follow the oil and move faster. The viscosity of water is lower than the oil viscosity and therefore the water production exceeds the oil production after water/CO₂ breakthrough. In addition, the changes in the relative permeability curves influence significantly on the mobility of the fluids. This is observed by the earlier water breakthrough for the CO₂ case. Due to frictional pressure loss in the production pipe, the pressure in the toe section is higher than in the heel section, and therefore, the first water breakthrough occurs in the heel of the production pipe. When the results from the basecase and CO₂-EOR case are compared, it is shown that the pressure difference (toe-heel) after 400 days of production is 0.6 bar higher when CO₂ - injection is used. This is due to the higher total flow rate, which creates higher friction in the production pipe.

3.2. Heterogeneous reservoir

Two different cases were simulated for the heterogeneous reservoir, one case with CO₂-EOR and ICD completion, and one case with CO₂-EOR and AICV completion. The heterogeneous reservoir has a high-permeability zone in the heel section. These simulation cases were performed to determine whether AICVs can reduce the water production compared to the ICDs.

3.2.1. Well with ICD completion

Fig. 12 shows the volume of accumulated oil and water versus time. The water breakthrough occurs after 12 days, and the first breakthrough is located in the heel section. Fig. 13 shows the water cut versus time for the heel, middle and toe sections of the well. After about 100 days, water breakthrough is observed in all the zones and the water production increases rapidly compared to the oil. After 400 days, the accumulated oil production is 212000 m³ and the overall water cut is 81%. The water cut in the heel section exceeds 90% after 150 days and is close to 100% at day 400.

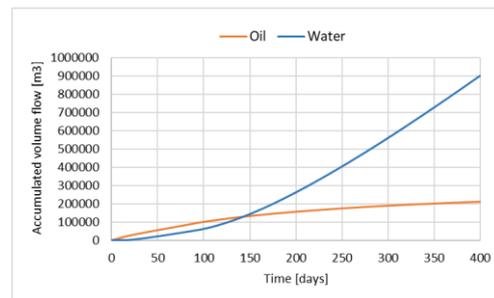


Figure 12: Accumulated oil and water for the ICD case.

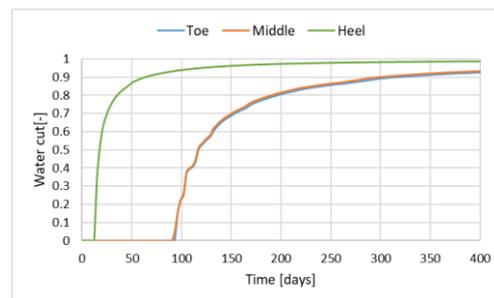


Figure 13: Water cut versus time for the ICD case.

Fig. 14 represents the oil and water flow rates as a function of time. After water breakthrough has occurred in all the zones, the total flow rate increases significantly. Fig. 15 shows the pressure along the well at different times. The pressure in the toe section (location 0) increases significantly with time due to increased total volume flow and thereby increased frictional pressure drop in the well.

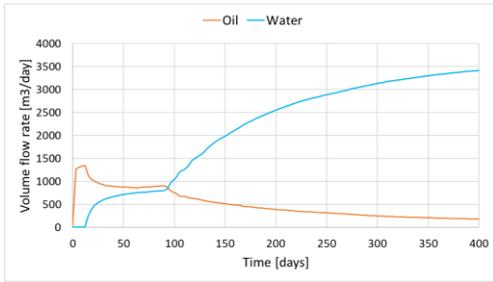


Figure 14: Volume flow rate for oil and water for the ICD case.

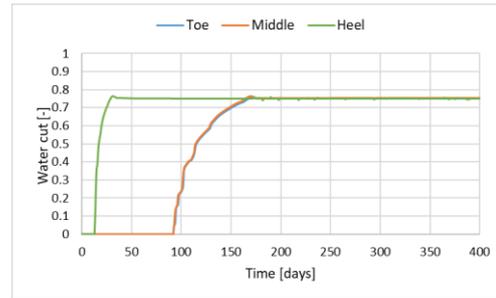


Figure 17: Water cut versus time for the AICV case.

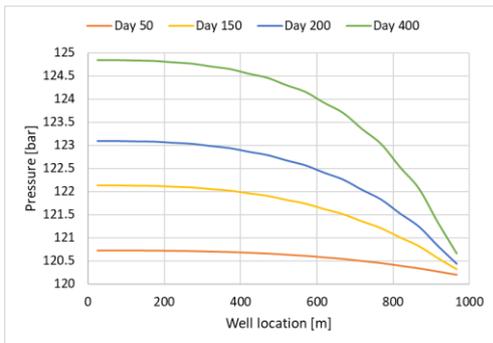


Figure 15: Pressure versus well position for the ICD case.

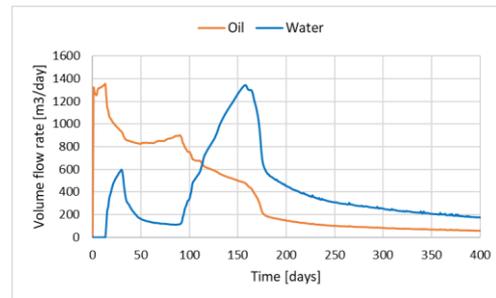


Figure 18: Volume flow rate for oil and water for the AICV case.

3.2.2. Well with AICV completion

Fig. 16 shows the accumulated oil and water production versus time for the AICV case. The water/CO₂ breakthrough is observed in the heel section after 12 days of production. This is the same as for the ICD case. After 400 days, the oil production is 164000 m³ and the water production is about the same. The water cut curves for the toe, middle and heel sections are presented in Fig. 17. The AICVs start to close when the water cut reaches 75% in the current zone. After about 170 days, all the AICVs are partly closed, and the production capacity is reduced to about 1%. The flow rates versus time for oil and water/CO₂ are presented in Fig. 18. The flow rates for both oil and water/CO₂ vary with time due to the breakthrough of water/CO₂ in the different zones followed by the choking of the AICVs.

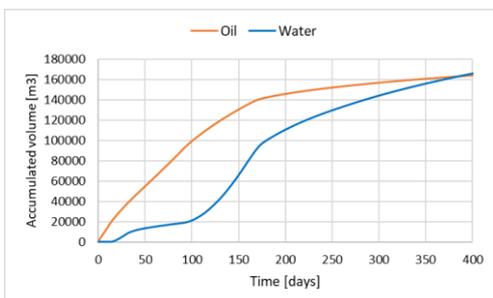


Figure 16: Accumulated oil and water for the AICV case.

Fig. 19 presents the pressure profile in the AICV completed well at different times. The well pressure in the heel is set constant to 120 bar, and the decreasing pressure from the toe to the heel is due to the frictional pressure drop in the production pipe. The pressure in the toe section decreases from 122 bar at day 150 to close to 120 bar at day 400. The reduction in the pressure is due to the decreasing production rates after the choking of the AICVs. The low pressure drop at day 50 is due to the water/CO₂ breakthrough in the high permeability zone and the following choking of this zone.

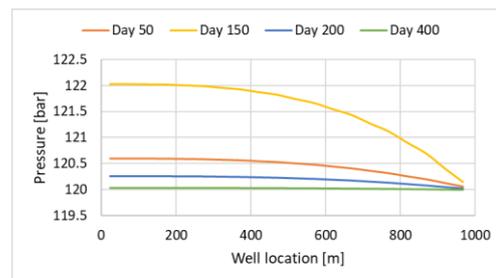


Figure 19: Pressure versus well position for the AICV case.

4. Summary and Discussions

Simulations were performed in a homogeneous reservoir to study the effect of CO₂ EOR on oil production. It was observed that CO₂ EOR increases the oil production rate until breakthrough of water/CO₂ occurs. The total oil production increased by 10% after 400 days of production, but since CO₂ influences the mobility of both oil and water, the water/CO₂ production increased by with 63%. The water cut at day 400 increased from 96% to 98% when CO₂ was injected. The average water cut based on 400 days of production is 69% and 77% for the

basecase and the CO₂ EOR case, respectively. This indicates that CO₂-EOR has the potential to decrease the residual oil in the reservoir and increase the total oil production. However, a significant amount of water will be produced together with the oil. The oil and water/CO₂ production depends very much on the relative permeability. In this study, it was not possible to find exact information about relative permeability for the simulated type of reservoir. The next step was to study the effect of using AICVs compared to ICDs in a heterogeneous reservoir with CO₂ EOR. Fig. 20 shows the comparison between water production from a well with AICV completion and a well with ICD completion. The AICV well is reducing the water production by 82% compared to the ICD well.

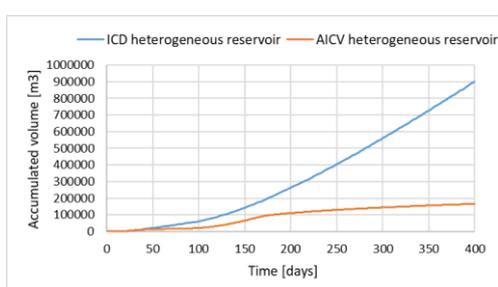


Figure 20: Comparison of accumulated water for AICV and ICD in the heterogeneous reservoir.

The simulated oil flow rates versus time for the two cases are presented in Fig. 21. The ICD well is producing more oil than the AICV well after breakthrough of water/CO₂ has occurred in all the zones. The reason why more oil is produced in the case of ICD is that no quantity restrictions have been set, which is unrealistic. In the case with ICD completion, the accumulated oil/water ratio is significantly lower than in the case with AICV completion. The normal scenario is to use a choke valve to control the total production rate and avoid overloading the top-side separation processes. A choking of the total flow will result in an even lower oil/water ratio for the ICD case because the main production will occur from the high permeability zone, and after breakthrough, the water/CO₂ production will increase significantly with time.

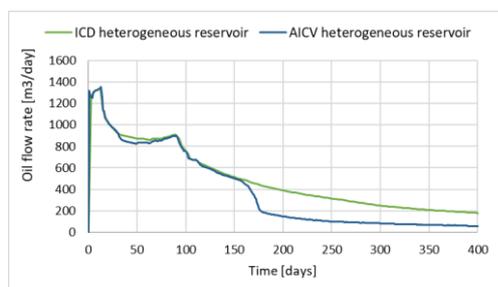


Figure 21: Comparison of oil flow rate for AICV and ICD in the heterogeneous reservoir.

Tab. 5 summarizes the production results from the heterogeneous reservoir.

Table 5: Production data for ICD and AICV completed wells in heterogeneous reservoir with CO₂-EOR

	ICD	AICV
Accumulated oil	212000 m ³	164000 m ³
Accumulated water	902000 m ³	166000 m ³
Water breakthrough	15 days	15 days
Water cut (day 400)	98%	75%
Total oil/water ratio	0.230	0.988

Future simulations will be performed with choking of the total flow for both the ICD case and the AICV case to obtain more realistic and comparable results for the two cases. The set point for a choke valve will be chosen based on the capacity of the top-side process equipment.

References

- [1] Oljedirektoratet. 'Avanserte utvinningsmetoder (EOR)'. <https://www.npd.no/fakta/produksjon/ior/avanserte-utvinningsmetoder-eor/> (accessed 25.04.2022).
- [2] Oljedirektoratet. 'CO₂-håndtering'. <https://www.npd.no/fakta/co2-handtering-ccs/> (accessed 05.05.2022).
- [3] OMICS International. 'Oil Swelling'. <https://www.omicsonline.org/oil-swelling.php> (accessed 13.03.2022).
- [4] Preston, C. *et al.* (2009), 'IEA GHG Weyburn-Midale CO₂ monitoring and storage project—moving forward with the final phase', *Energy Procedia*, 1(1), pp. 1743-1750. doi: 10.1016/j.egypro.2009.01.228.
- [5] Tadesse N. S., Torsæter O. and Slotte, P. A. (2018) 'Simulation of CO₂ Injection in a Reservoir with an Underlying Paleo Residual Oil Zone', NTNU, 2018.
- [6] Kandil, M. E. (2019) 'Measurements and Modeling of Phase Behavior, Density, and Viscosity of Oil + Carbon Dioxide: Squalane + CO₂ at Temperatures (313 to 363) K and Pressures up to 77 MPa', *J. Chem. Eng. Data*, 64(1), pp. 273-285, 2019, doi: 10.1021/acs.jced.8b00786.
- [7] The Energy lab. 'Carbon Dioxide Enhanced Oil Recovery.' https://www.netl.doe.gov/sites/default/files/netl-file/co2_eor_primer.pdf (accessed 13.03.2022).
- [8] Majlaton, N.I. (2012) 'A visual study of CO₂ injection at the pore scale using micromodels,' ed. Department of physics and technology, University of Bergen.
- [9] Verma, M. K. (2015) 'Fundamentals of carbon dioxide-enhanced oil recovery (CO₂-EOR): a supporting document of the assessment methodology for hydrocarbon recovery using CO₂-EOR associated with carbon sequestration,' Reston, VA, Report 2015-1071, 2015. [Online].
- [10] Aakre, H. Halvorsen, B., Werswick, B. and Mathiesen, V. (2013) 'Smart Well With Autonomous Inflow Control Valve Technology,' *SPE Middle East Oil and Gas Show and Conference*, 2013, vol. All Days, SPE-164348-MS, doi: 10.2118/164348-ms. <https://doi.org/10.2118/164348-MS>
- [11] Mathiesen, V., Werswick, B., Aakre, H. (2014) The Next Generation Inflow Control, the Next Step to Increase Oil Recovery on the Norwegian Continental Shelf. SPE-169233-MS (2014) *SPE Bergen One Day Seminar, 2 April, Bergen, Norway. Society of Petroleum Engineers.*
- [12] Krinis, D. *et al.* (2009) 'Optimizing Horizontal Well Performance In Non-Uniform Pressure Environments Using Passive Inflow Control Devices.' *Offshore Technology Conference (2009). Society of Petroleum Engineers.*
- [13] Al-Khelaiwi, F. T. and Davies, D. R. (2007), 'Inflow Control Devices: Application and Value Quantification of a Developing

Technology.' *SPE-108700-MS International Oil Conference and Exhibition in Mexico (2007)*,) *Society of Petroleum Engineers*

[14] Kais, R. *et al.* (2016)'First Autonomous Inflow Control Valve (AICV) Well Completion Deployed in a Field under an EOR (Water & CO₂ Injection) Scheme.' *SPE-181552-MS, SPE Annual Technical Conference and Exhibition held in Dubai, UAE, 2016.*

[15] Aakre, H., Mathiesen, V., Moldestad, B. (2018) 'Performance of CO₂ flooding in a heterogeneous oil reservoir using autonomous inflow control.' *Journal of Petroleum Science and Engineering* (2018).

[16] Hansen, O. C. and Moldestad, B. M. E. (2020) 'Computational study of CO₂ injection at Johan Sverdrup for enhanced oil recovery and storage.' *Linköping Electronic Conference Proceedings*, (176) pp. 311-317.

[17] Bin Hu, A., 'Transient Wellbore/Reservoir Model: Integrated wellbore/reservoir transient simulation' http://www.spebergen.no/publish_files/3.1_SPT_group_Bin_Hu.pdf

[18] Tiab, D. and Donaldson, E. C. (2012). 'Petrophysics, Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties', 3rd edn., USA : Gulf Professional Publishing. ISBN: 978-0-12-383848-3.

[19] Schlumberger, *Rocx Reservoir Simulator* (2007), 1.2.5.0 edn.

[20] Thu, E. S. (2013) 'Modeling of transient CO₂ flow in pipeline and well,' Master Thesis, Norwegian University of Science and Technology.

Eulerian-Lagrangian simulation of air-steam biomass gasification in a bubbling fluidized bed gasifier

Nastaran A. Samani^{a, *}, Marianne S. Eikeland^a

^a Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, nastaran.samani@usn.no

Abstract

To numerically study biomass gasification in a three-dimensional bubbling fluidized bed, a CFD-DEM (computational fluid dynamics – discrete element method) model with heat transfer and homogeneous and heterogeneous chemical reactions is implemented. An ideal reactor model is used for the air-steam bubbling fluidized bed (BFB) gasification reactor assuming perfectly mixed solids and plug flow. A validated computational particle fluid dynamics (CPFD) model has been applied to investigate the sensitivity analysis of mesh grids as well as to find the optimum number of grids. The result shows that 7452 grid cells are the optimal number of cells for the existence BFB gasifier. The effects of key process operating parameters such as steam to biomass ratio (SB), as well as temperature shows that by enhancing the SB ratio or reactor temperature, gas yields increase. H₂ and CO₂ concentrations promote by increasing the steam to biomass ratio while CO and CH₄ production drop. The optimal value of SB for the gasification process can be found in the range of 0.3 to 1.

Keywords: bubbling fluidized bed gasifier, CPFD method, Steam-to-biomass ratio, mesh sensitivity analysis, Eulerian-Lagrangian simulation

1. Introduction

Gasification is the partial oxidation of the carbon in the biomass in the presence of a gasifying carrier such as air, oxygen, steam, or carbon dioxide. The biomass gasification process can convert a solid or liquid organic material into a multifunctional gaseous compound and a solid phase. The gas phase, called syngas, has a high heating value and can be applied to produce biofuel and generate power. Char consists of organic residue and inert material and comprises the solid phase (A. Samani et al., 2022). The gasification of biomass consists of a sequence of successive, endothermic, and exothermic reactions that can be divided into the main subprocesses as can be seen in Fig. 1.

The gasification process consists of drying, pyrolysis, partial oxidation, reduction (char gasification), and cracking (Arena, 2012, Mahinpey and Gomez, 2016).

Heating and drying: The moisture content makes up 5-35% of the biomass feed. During the drying stage, the moisture is released from the feed, and it is decreased to 5%. The operating temperature of the drying stage depends on the pressure inside the reactor, for example, if the reactor is operating on a pressure of 1-60 bars, the boiling temperature of the water can be between 373 to 550 K. High moisture content can lead to some troubles in feeding or fluidization like agglomerate formation and

jamming problems. Furthermore, it can decrease the lower heating value (LHV) of the production as well as the energy efficiency of the gasification. Therefore, the tar content in the syngas is more likely to increase due to the reduction of the reaction temperature (Arena, 2012, Mishra and Upadhyay, 2021).

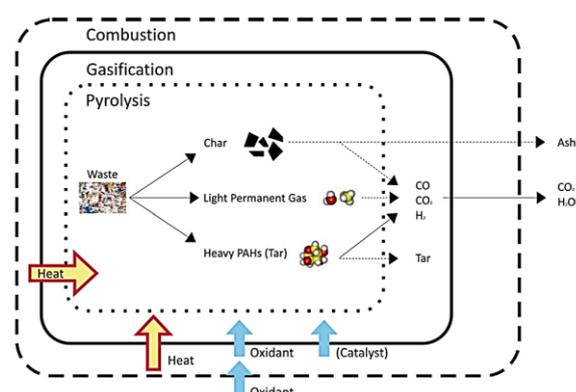


Figure 1: Schematic representation of pyrolysis, gasification, and combustion processes (Arena, 2012)

Devolatilization (pyrolysis): During the pyrolysis stage, the organic materials are decomposed into volatile and biochar (carbonaceous solid residue) in an oxygen-deprived atmosphere at higher temperatures (398-773 K)(Mishra and Upadhyay,

2021). This process involves the breaking of the transitory bonds of the aromatic clusters in the feedstock. Thus, light gases (volatile matter) in the cooler part of biomass condense and produce high molecular weight hydrocarbons (tar). When dried biomass is heated to temperatures ranging from 200 to 500 °C during the pyrolysis step, it decomposes into solid char and volatiles (tar and gases), as shown in Fig. 2.

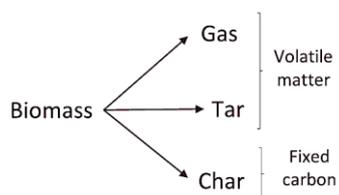


Figure 2: Pyrolysis output (Safarian et al., 2019)

Combustion(oxidation): Although the easiest and the most direct way of decomposition of biomass is the combustion, the overall heat released from the biomass in the combustion zone is lower than in the gasification process. Due to the exothermic chemical reactions, the temperature can be elevated to 1373 to 1773 K in the combustion zone. By controlling the gasifying agents, the temperature can be controlled in such a way that it does not approach the ash-slugging temperature and causes operational problems. The amount of pure oxygen or air is in the range of 25 to 40 percent. CO, CO₂, H₂, and H₂O are formed as the final products, and the heat released from the exothermic reactions can be applied to dry the particles in the pyrolysis process (Zhang et al., 2010).

Gasification/reduction: In this zone, some CO₂ and H₂O are reduced in the reaction of the remaining carbonized residue produced from pyrolysis. It can produce a gas with a high CO and H₂ content. The required energy for endothermic reduction reactions can be provided by the combustion of char and volatiles. Hydrogen, carbon monoxide, and methane can be generated through a sequence of reactions. Tar is one of the main problems in gasification of the biomass and can reduce the overall efficiency of the process and increase the cost of the plant (Mularski et al., 2020).

Consequently, gasifier reactors must be designed using either experimental data or numerical modeling and simulation. Modeling should be combined with experimental results as the most reliable option. Thus, at the R&D level, modeling, and simulation play an important role in the design and study of the gasification process. Modeling and simulations offer a low-cost method for optimizing existing gasifiers as well as scaling up and designing new gasifiers in terms of key operational parameters. Exploration of these operational parameters can provide insight into the relationship between the

influences of the gasification variables and trends in process, cost, and implementation risks. (Baruah and Baruah, 2014, Safarian et al., 2019).

Some of the important operating parameters influencing the gasification process are feedstock flow rate, gasifying agent flow rate, equivalence ratio, reactor pressure, and reactor temperature. Any parameter change has a significant impact on the end-gas composition and, as a result, the gasifier's performance. Mathematical and computational fluid dynamics (CFD) models are created to provide a good representation of the chemical and physical phenomena that occur inside the gasification reactors (Basu, 2010). An Eulerian-Lagrangian model has been applied to study the influence of increasing the temperature, steam-biomass ratio, and equivalence ratio on the solid phase's gas voidage, fluid force, collision force, dispersion coefficient, as well as the various homogeneous reaction rates. The model offers important information about the hydrodynamics of the bubbling fluidized bed for biomass gasification, and they are expected to be useful for the operation, scale-up, and optimization of such systems for sustainable energy production (Yang et al., 2019). To simulate the gasification of pine sawdust in the presence of both air and steam, a comprehensive model was developed. The proposed model improved on the premise of an existing biomass gasification model based on ASPEN PLUS. The accuracy of the model's predictions was compared to actual experimental results to confirm validity. Furthermore, the comprehensive model's relative accuracy was compared to the original base model to see if there was any improvement. The model predicts the composition of H₂, CO, CO₂, and CH₄ with reasonable accuracy in varying temperature, steam-to-biomass, and equivalence ratio conditions (Pauls et al., 2016).

Since gasification is a thermodynamic conversion containing many solid-gas reactions, the gasification parameters such as temperature, pressure, feed composition, type of the gasification agents, and kinetics of the reaction are more likely to affect the syngas production. For this reason, the purpose of this paper is to apply CPFD simulation to evaluate the impacts of the steam to biomass ratio on the gasification process.

2. The CPFD method

CPFD is a numerical method based on the Eulerian-Lagrangian approach for the simulation of a large-scale multiphase (particle-fluid) flow system in three dimensions by adopting the multiphase particle-in-cell (MP-PIC) method and the particle parceling algorithm (Andrews and O'Rourke, 1996). The basis of the Eulerian-Lagrangian method is that a continuum model is considered for the fluid phase and a Lagrangian method is applied for the particle

phase. This can provide an appropriate numerical solution for a wide range of particle sizes, shapes, and velocities. For the fluid phase, the Navier-Stokes equation with coupling between the discrete particles is utilized (Samani et al.). However, in order to solve the particle phase, the direct element method (DEM) fits into the Lagrangian method (Jiang et al., 2014). The method of solving is that the computational domain is divided into several computational parcels including particles with the same characteristics. Furthermore, conservation equations of momentum, mass, and energy are computed, including the coupling between the gas/solid phases. Therefore, the spatial distribution of key process parameters such as temperature, pressure, and velocities in the system can be calculated. Such information is necessary to find out the fluidization process, and that is why the CFPD scheme is widely used for the simulation of gas-particle fluidized reactors (Samani et al., 2020).

3. CFPD simulation set-up

Computational Particle Fluid Dynamics (CPFD) modeling is used to simulate the bubbling fluidized bed gasifier reactor at the University of South-Eastern Norway, with the goal of mesh sensitivity analysis to find the optimal number of grid cells to achieve accurate results. The CFPD hydrodynamic model has been validated against the results of Timsina's experiment for wood chips as biomass air gasification at 1000 K (Timsina et al., 2020) by A. Samani (A. Samani et al., 2022). In this paper, the model was used to investigate the effects of steam to biomass ratio as a key operating parameter on the steam gasification process.

The reactor was modelled as a cylinder with an inner diameter of 0.1 m and a height of 1 m. SolidWorks was used to create a computer-aided design (CAD) model of the reactor in stereolithography (STL) format, which was then imported into The CFPD Barracuda® VR software. Fig. 3 illustrates the initial bed conditions and boundary conditions.

The proximate and ultimate analysis of biomass species used in the experimental studies is shown in Tab.1. The rest of the physical and operational conditions used in the simulations are represented in Tab. 2 (Timsina et al., 2020).

In Barracuda, the Wen-Yu model takes into account the particle packing by including a dependence on the fluid volume fraction in addition to the single particle drag models upon which it is based. The Wen-Yu drag model is appropriate for more dilute systems, the Wen-Yu drag model has been applied with 60% momentum retention for the particle collision.

Table 1: Characterization of biomass (A. Samani et al., 2022)

Type of biomass	Wood chips
Proximate analysis (wt % basis)	
ash	1.16
volatiles	80
Fixed carbon	18.84
Moisture	11.1
Ultimate analysis (wt %)	
C	51.0
H	6.1
O	42.2
N	0.11
S	0.011
Cl	0.011
Ash	0.58

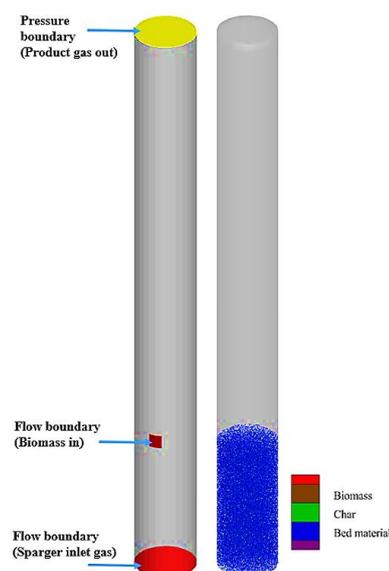


Figure 3: Boundary conditions and c) initial particle in bed (silica sand+char)

Table 2: Gasifier operating conditions (A. Samani et al., 2022)

Initial bed mass	Particle: Silica Particle size: 100-425µm 0.49 volume fraction Density: 2650 kg/m ³ Initial height: 250 mm
	Particle: Char Particle size: 500µm 0.05 volume fraction Density: 300 kg/m ³ Initial height: 250 mm
Solids close pack volume fraction	0.54
Initial gas in the reactor	Air (80% N ₂ , 20% O ₂)
Initial solid and gas temperature	1000 K
Initial gasifier pressure	101325 Pa
Sparger inlet gas composition by mass	80% H ₂ O, 20% air (80% N ₂ , 20% O ₂)
Sparger inlet gas temperature	1000 K
Sparger inlet gas superficial velocity	0.17 m/s
Biomass	Temperature: 500 K Mass flow rate: 2.08 kg/h

3.1 Main reactions of gasification:

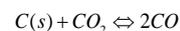
The main global gasification reactions are divided into two groups, heterogeneous and homogenous.

3.1.1 Heterogeneous (solid-gas phase) reactions

The heterogeneous reactions are (Hejazi et al., 2017, Snider et al., 2011):

For all reactions, the kinetics rates [$\text{molm}^{-3}\text{s}^{-1}$] are represented. r_{f1} and r_{r1} are kinetics rates for the forward and reverse reactions. m_s is the mass of the solid, the solid volume fraction, and is the temperature of the reaction.

Boudouard:

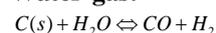


$$\Delta H_{rx}^{\circ} = +172 \text{ kJ / mol}$$

$$r_{f1} = 1.272 m_s T \exp\left(\frac{-22645}{T}\right) [CO_2]$$

$$r_{r1} = 1.044 \times 10^{-4} m_s T^2 \exp\left(\frac{-2363}{T} - 20.92\right) [CO]^2$$

Water-gas:

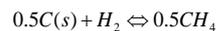


$$\Delta H_{rx}^{\circ} = +131 \text{ kJ / mol}$$

$$r_{f1} = 1.272 m_s T \exp\left(\frac{-22645}{T}\right) [H_2O]$$

$$r_{r1} = 1.044 \times 10^{-4} m_s T^2 \exp\left(\frac{-6319}{T} - 17.29\right) [H_2][CO]$$

Methanation:

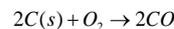


$$\Delta H_{rx}^{\circ} = -75 \text{ kJ / mol}$$

$$r_{f3} = 1.368 \times 10^{-3} m_s T \exp\left(\frac{-8078}{T} - 7.087\right) [H_2]$$

$$r_{r3} = 0.151 m_s T^{0.5} \exp\left(\frac{-13578}{T} - 0.372\right) [CH_4]^{0.5}$$

Combustion:



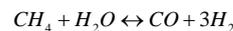
$$\Delta H_{rx}^{\circ} = -111 \text{ kJ / mol}$$

$$r_{f6} = 4.34 \times 10^7 \varphi_p T \exp\left(\frac{-13590}{T}\right) [O_2]$$

3.1.2 Homogeneous (gas-phase) reaction

Simple global homogeneous reactions are listed as (Snider et al., 2011):

Steam Methane reforming (SMR):

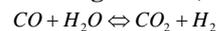


$$\Delta H_{rx}^{\circ} = +206 \text{ kJ / mol}$$

$$r_{f4} = 3 \times 10^{-5} \exp\left(\frac{-15042}{T}\right) [H_2O][CH_4]$$

$$r_{r4} = 0.0265 \exp\left(\frac{-32900}{T}\right) [CO][H_2]^2$$

Water-gas shift (WGS):



$$\Delta H_{rx}^{\circ} = -41 \text{ kJ / mol}$$

$$r_{f5} = 7.68 \times 10^{10} \exp\left(\frac{-36640}{T}\right) [H_2O][CO]^{0.5}$$

$$r_{r5} = 6.4 \times 10^9 \exp\left(\frac{-39260}{T}\right) [H_2]^{0.5} [CO_2]$$

4. Results and discussion

4.1. Sensitivity mesh analysis

The CPFDF method of solving is that the computational domain is divided into several computational cells including particles with the same velocity and properties. Furthermore, conservation equations of momentum, mass, and energy are computed, including the coupling between the gas/solid phases. The grid defines the spatial resolution for calculating the governing equations that result in gas-particle flow properties like pressure, velocity, and temperature. In order to achieve accurate results, it is necessary to define an optimum number of grid cells. As a result, the CPFDF model for the BFB gasifier has been applied at the temperature of 1000 K, and the pressure of 101325 Pa for wood chips as biomass. The steam to biomass ratio of 0.8 and other gasification operating conditions are identical to Tab.2. The reason for choosing the SB ratio is that based on the results of the paper by Sharma (Sharma and Sheth, 2016), at this SB ratio, the hydrogen production could be maximum. The simulations have been carried out for 5 different mesh cells, 1450, 3577, 5376, 7452, and 10000 as shown in Fig.4. With 1450 grid cells, it was not possible to achieve results. This number of mesh is too coarse to achieve the results.



Figure 4: Computational grids, a: 1450, b: 7452, and c: 10000 grid cells

Fig. 5 and 6 investigate the effects of the grid sizes on the production of gas compositions and the concentration of CO production at the outlet cross-sectional area of the reactor. According to the results, since the error for 7452 and 10000 are very negligible (almost identical results), 7452 grid meshes are selected as an appropriate number of meshes for further analysis. By increasing the number of cells to more than 7452 grid cells, the results of gas production have not changed.

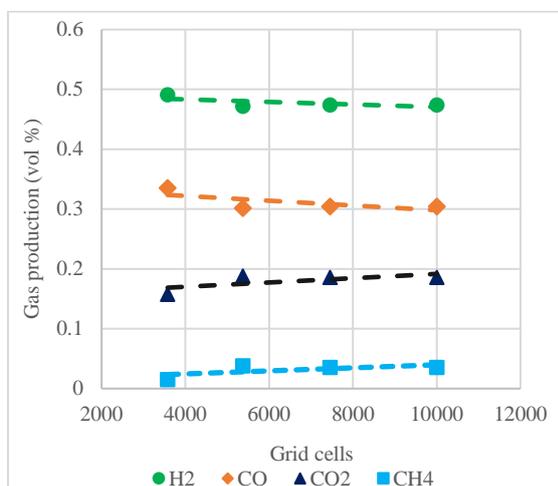


Figure 5: Effect of different grid cells on dry and N₂-free product gas composition for steam gasification of wood chips at a reactor temperature of 1000 K and S/B=0.8

4.2. The impacts of the gasification parameters

To investigate the impacts of SB on gas production as well as the temperature effects on the gas yield, the CPFD method has been applied to simulate the BFB gasifier for wood chips. The simulations have been conducted for 100 s to make sure that steady-state condition has been achieved.

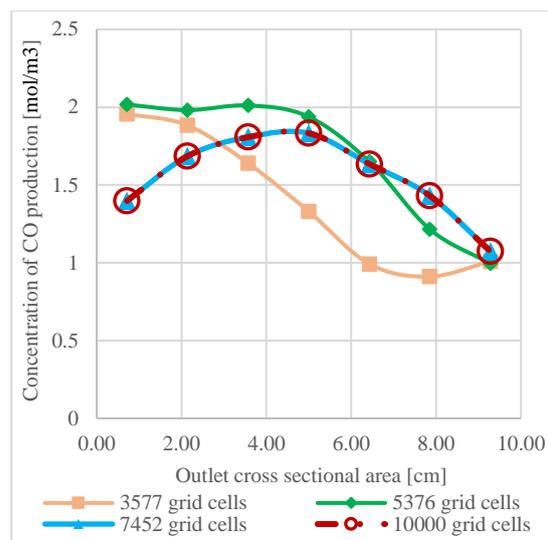


Figure 6: Concentration of CO production at the outlet cross-sectional of the cylinder for different mesh sizes

4.2.1. The effects of the steam to biomass ratio

This section discusses the variation of steady-state values of component production on the condition of dry and N₂ free, for air-steam gasification of wood chips at the temperature of 1000 K with steam to biomass ratio (SB) to identify the quantitative effects of steam addition on gasifier performance. Fig.7 presents particle volume fraction, particle temperature distribution, speed of the particle, and particle species distribution at the temperature at the simulation time of 100 s for SB equal to 0.8 and reactor temperature of 1000 K.

Owing to water gas shifts, which are aided by steam and are predominant at higher temperatures, increasing SB enhances H₂ and CO₂ concentrations and the heating value of syngas while decreasing CO and tar concentrations. Furthermore, because of the water gas and steam reforming reactions, the H₂ mole fraction in the product gas rises as the SB ratio is boosted. An excess of steam reduces temperature, favoring tar formation; additionally, the higher the SB, the more energy is required by the gasification process. The mole fractions of CO and CH₄ drop as the SB ratio promotes owing to increased WGS and SMR reactions as the results show in Fig.8. Higher H₂ and CO₂ concentrations were discovered for SB values ranging from 1.5 to 3. As can be seen in Fig.8, for the gasification process, the optimal value of SB can be found in the range of 0.3 to 1.

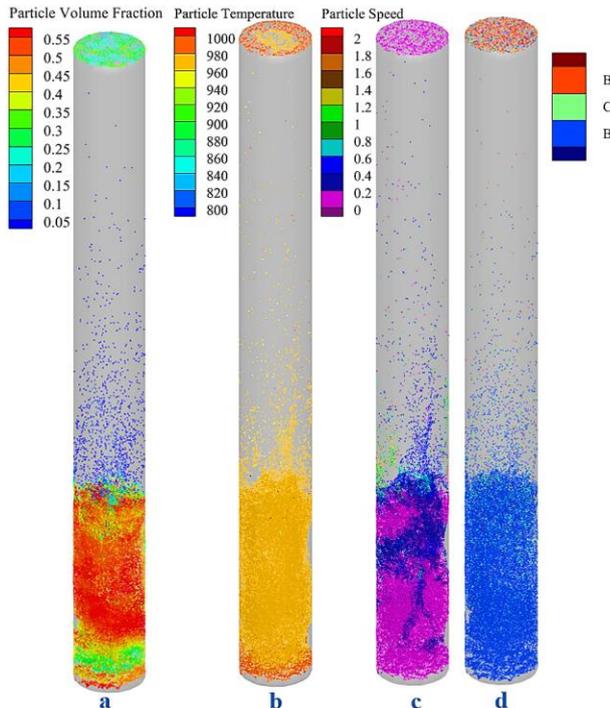


Figure 7: Bed hydrodynamics at 100 s simulation time, (a) Particle volume fractions, (b) Particle temperature distribution, (c) the speed of particle, and (d) Particle species distribution

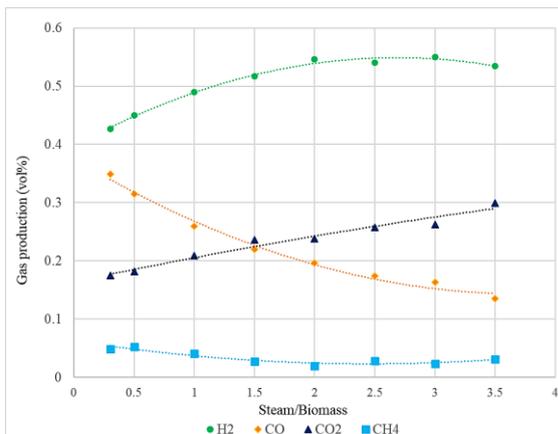


Figure 8: Effects of steam to biomass ratio on dry and N₂-free product gas composition for steam gasification of wood chips at a reactor temperature of 1000 K

4.2.2. The effects of the temperature on products yield

Gasification performance, syngas yield, and composition are all determined by the following operating parameters: partial pressure of Gasifying Agent (GA), heating rate and temperature, and gasification pressure. Fig.9 shows that by running the gasification process at high temperatures, it is possible to achieve high char conversion (the conversion of char into gases), low tar content, and high CO and H₂ contents owing to endothermic reactions of water-gas, Boudouard, and SMR. 750–

850°C, 800–900°C, and 850–950°C are typical temperature ranges for gasification of agricultural waste, refused derived fuel (RFD), and woody biomass, respectively (Molino et al., 2018). Temperatures above 1000 °C have two major drawbacks: ash melting and strict reactor specification requirements. Thus, selecting a temperature range for operating the reactor is one of the most important parameters.

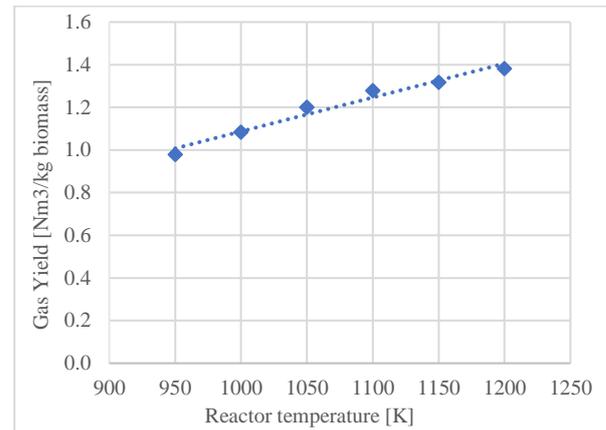


Figure 9: Predicted yield distribution of products from air-steam gasification of wood chips at different temperatures at SB=0.8

5. Conclusion

A validated three-dimensional numerical simulation was carried out to characterize the effects of the gasification parameters on the air-steam biomass gasification process in a BFB gasifier. In the first step, the mesh study has been carried out to find the optimum number of grid cells to have accurate results. It is illustrated that for the geometry of BFB in this paper, 7452 grid cells are chosen as the optimal number of meshes. By raising the steam-to-biomass (SB) ratio or reactor temperature, gas yields promote. Owing to the water gas, SMR, and WGS reactions, the H₂ mole fraction in the product gas rises by enhancing the SB ratio for steam gasification of wood chips in a BFB gasifier at a temperature of 1000 K. The CO₂ mole fraction rises as the SB ratio rises, most likely because the water-gas shift reaction becomes dominant at high gasifier temperatures. Due to increased WGS and SMR reactions, increasing the SB ratio brings the mole fractions of CO and CH₄ to drop. It is concluded that the optimal value of SB for the gasification process can be in the range of 0.3 and 1. By applying a more complex reaction for tar reforming and cracking, as well as using fuel-specific biomass gasification reaction kinetics to account for the catalytic influence of ash, the major factors affecting syngas production and composition such as gasification technologies (fixed bed reactors, fluidized bed reactors, entrained flow reactors), feedstock properties (biomass type, moisture content, particle

size, ash content), and operating gasification conditions (bed material, temperature, pressure, gas agent, equivalence ratio, SB) can be investigated.

References

- A. Samani, N., K. Thapa, R., Moldestad, B. M. E., & S. Eikeland, M. (2022) 'Evaluating the impacts of temperature on a bubbling fluidized bed biomass gasification using CPFD simulation model', *MATHMOD 2022*.
- Andrews, M. J., & O'Rourke, P. J. (1996) 'The multiphase particle-in-cell (MP-PIC) method for dense particulate flows', *International Journal of Multiphase Flow*, 22(2), 379-402.
- Arena, U. (2012). 'Process and technological aspects of municipal solid waste gasification. A review', *Waste management*, 32(4), 625-639.
- Baruah, D., & Baruah, D. (2014) 'Modeling of biomass gasification: A review', *Renewable and Sustainable Energy Reviews*, 39, 806-815.
- Basu, P. (2010) 'Biomass gasification and pyrolysis: practical design and theory', *Academic press*.
- Hejazi, B., Grace, J. R., Bi, X., & Mahecha-Botero, A. s. (2017) 'Kinetic model of steam gasification of biomass in a bubbling fluidized bed reactor', *Energy & Fuels*, 31(2), 1702-1711.
- Jiang, D., Liu, Q., Wang, K., Qian, J., Dong, X., Yang, Z., Du, X., & Qiu, B. (2014) 'Enhanced non-enzymatic glucose sensing based on copper nanoparticles decorated nitrogen-doped graphene', *Biosensors and Bioelectronics*, 54, 273-278.
- Mahinpey, N., & Gomez, A. (2016) 'Review of gasification fundamentals and new findings: Reactors, feedstock, and kinetic studies', *Chemical engineering science*, 148, 14-31.
- Mishra, S., & Upadhyay, R. K. (2021) 'Review on biomass gasification: gasifiers, gasifying mediums, and operational parameters', *Materials Science for Energy Technologies*, 4, 329-340.
- Molino, A., Larocca, V., Chianese, S., & Musmarra, D. (2018) 'Biofuels production by biomass gasification: A review', *Energies*, 11(4), 811.
- Mularski, J., Pawlak-Kruczek, H., & Modlinski, N. (2020) 'A review of recent studies of the CFD modelling of coal gasification in entrained flow gasifiers, covering devolatilization, gas-phase reactions, surface reactions', models and kinetics, *Fuel*, 271, 117620.
- Pauls, J. H., Mahinpey, N., & Mostafavi, E. (2016) 'Simulation of air-steam gasification of woody biomass in a bubbling fluidized bed using Aspen Plus: A comprehensive model including pyrolysis, hydrodynamics and tar production', *Biomass and bioenergy*, 95, 157-166.
- Safarian, S., Unnþórsson, R., & Richter, C. (2019) 'A review of biomass gasification modelling', *Renewable and Sustainable Energy Reviews*, 110, 378-391.
- SAMANI, N. A., JAYARATHNA, C. K. & TOKHEIM, L.-A. (2020) 'Fluidized bed calcination of cement raw meal: Laboratory experiments and CPFD simulations', *Linköping Electronic Conference Proceedings (Proceedings of the 61st SIMS, September 22nd-24th, virtual conference)*, pp. 399-406.
- Samani, N. A., Jayarathna, C. K., & Tokheim, L.-A. (2020) 'CPFD simulation of enhanced cement raw meal fluidization through mixing with coarse, inert particles'.
- Sharma, S., & Sheth, P. N. (2016) 'Air-steam biomass gasification: experiments, modeling and simulation', *Energy conversion and management*, 110, 307-318.
- Snider, D. M., Clark, S. M., & O'Rourke, P. J. (2011) 'Eulerian-Lagrangian method for three-dimensional thermal reacting flow with application to coal gasifiers', *Chemical engineering science*, 66(6), 1285-1295.
- Timsina, R., Thapa, R. K., Moldestad, B. M., & Eikeland, M. S. (2020) 'Experiments and computational particle fluid dynamics simulations of biomass gasification in an air-blown fluidized bed gasifier', *International Journal of Energy Production and Management*. Vol. 5. Iss. 2, 5(2), 102-114.
- Yang, S., Wang, H., Wei, Y., Hu, J., & Chew, J. W. (2019) 'Eulerian-Lagrangian simulation of air-steam biomass

gasification in a three-dimensional bubbling fluidized gasifier', *Energy*, 181, 1075-1093.

Zhang, L., Xu, C. C., & Champagne, P. (2010) 'Overview of recent advances in thermo-chemical conversion of biomass', *Energy conversion and management*, 51(5), 969-982.

CFD simulations of flow jetting impact and high erosion region in a production choke and its downstream spool

Agastian Perinpasivam ^{a,*}, Nikhil Bagalkot ^b, Arvind Keprate3 ^b

^a Westcon Yard AS Haugesund Norway,

^b Department of Mechanical Engineering, Oslo Metropolitan University, Norway
nikhilba@oslomet.no

Abstract

Erosion wear is a rather well-known problem in the petroleum and transport industry. Over the years there have been many different models suggested to estimate the erosion. Each model uses unique equations and is suited for different types of geometries, which gives different flow profiles and erosion patterns. It is critical to know where the erosion wear occurs and at what magnitude the system is located at to predict and economically create a choke valve design. The erosion study additionally, helps in accessing location and thickness of cladding required to prolong the life of components. In this study, computational fluid dynamics (CFD) is used to investigate erosion on a production choke valve and its downstream spool. There are three main steps to predicting erosion wear using CFD analysis: flow model, particle tracking and calculating the erosion wear from particle interaction. The results indicate that pressure drop affects the velocity jet shape and impact region. In the 50% opening case, the pressure drop creates a wide jet stream region that, in turn, will cause an increased wall impact region. When the opening decreases, to 35% and 25%, the jet stream gets more focused. The velocity jet impacts a smaller area of the pipe wall, which in turn creates possibilities for increased erosion rate. The high pressure drop in the 15% opening case creates a high-focused jet stream in the middle of the downstream pipe, leading to low wall interaction.

1. Introduction

For the production of oil and gas (surface and subsea), Production Choke Valves (PCVs) are used to control the flowrate and pressure further downstream. PCVs are valves designed to take the brunt of the pressure of the line components, as this helps to increase the life expectancy of the system. The primary task of the PCV is to lower the pressure to a manageable level, so that the oil and gas can be transported safely (McLaury, Shirazi et al. 2000). By restricting the flow to a very small orifice, it reduces the well pressure downstream. PCVs have the ability to change the cross-section area (throat) either manually or by using sensors. The change in the area would lead to varying of the pressure resulting in a pressure drop and increase in velocity, which makes the choke valve most subjected to erosion.

Generally, a mixture of oil, gas and water is flowing through the system, which will contain small impurities like sand particles. In most of the cases there will be sand filters installed, which are effective in removing the large sand particles. However, smaller size sand particles (<150 micrometer), end up in the production lines, which are the main reason for erosion. As the production choke reduces the pressure, the loss in pressure energy will result in increase in the kinetic energy. Therefore, the reduction in pressure will cause the

flow and the particles to move up to sonic velocities. Erosion rates at these high velocities will become substantially high, which will eventually lead to repair, replacement or leakage of the components.

The damage of choke valves and the downstream production flow lines is dangerous for workers in the area, and the replacement cost is extremely high (Haugen, Kvernfold et al. 1995, Raghavendra, Shivashankar et al. 2014). Thus, the need for good materials (Wheeler, Wood et al. 2006) and design optimization is very desirable. Increasing the longevity of a choke valve will have a significant cost reduction on top of the area being more secure.

Erosion is an occurrence that happens when mass is worn away from the material surface due to either chemical or physical interaction. In the industry, any process that transports solid particles in a fluid phase is vulnerable to erosion damage (Haugen, Kvernfold et al. 1995, Oka, Okamura et al. 2005, Desale, Gandhi et al. 2009). It often occurs in line components such as pipe bends, tubes and structures that alter the flow field. Pipelines that bend cause shifts in the flow, which makes the particles hit the wall of a surface with high pressure or velocity.

Over the years researchers have spent a significant amount of time trying to understand erosion models, what type of mechanisms and material needed when developing pipeline systems (DNV 2015). Both

fluid (velocity, density, pressure to name a few) and particle properties (size, shape, density, and concentration) play a critical role in erosion phenomena and has been well researched (Wallace, Dempster et al. 2004, Mathew 2017, Bishnoi, Kumar et al. 2021). The effect of particle size and shape have been a big topic of investigation with many researchers. Oka, Okamura et al. (2005) created an erosion model that defines that erosion rate caused by sand particles on stainless steel is dependent on the shape, velocity and impact angle of the particle. In their study they find that erosion rate increases as the particle size becomes bigger. Similarly, Feng and Ball (1999) came to the conclusion that particle size increases the erosion rate. The size affects both the particle impact velocity and the kinetic energy.

In addition to fluid and particle properties, the design of the production choke plays a critical role. Changes in valve geometry can impact the erosion wear significantly. Given time, erosion can drastically change the flow field, which can affect production and later result in equipment failure. McLaury, Shirazi et al. (2000) explains this in a computational study where the effects of erosion wear and geometry changes were compared between experimental erosion results and an erosion prediction model. A waterflow mixed with sand particles was directed through a choke geometry with a sharp entrance profile, and the rounding of the leading edge began immediately after starting the test. In addition, the turbulent kinetic energy near the entrance was large which, in turn, led to turbulent fluctuations in the flow. As a result of the high turbulent fluctuations, more sand particles struck the geometry wall, causing increased erosion. Results showed that the predicted erosion rates were larger than the experimental results, but when the rounded edge was accounted for, the predictions matched the experimental results very well. This shows the importance of accounting for changes in the geometry to attain accurate predictions.

Wallace, Dempster et al. (2004) investigated the capability of computational fluid dynamics techniques to estimate the erosion rate in two different valve geometries, a simple geometry with basic geometric features and a more complex choke geometry. Measurements from a parallel experiment were used as comparison regarding erosion rates and flow coefficient predictions. The test resulted in underestimated erosion rates; however, erosion location matched the experimental data, along with the flow characteristics. It is suggested that neglecting the changes in the model geometry due to erosion could be one contributing factor to the faulty erosion rate predictions.

The main objective for the current work is to simulate erosion in a PCV. In the industry, and in

previous research on this topic it is always assumed that the maximum erosion or the erosion hot spot is mainly at the first “U” bend downstream of the choke. Therefore, there is a thicker cladding on the “U” bend section than other regions. Some of the analytical models are also based on this assumption. The novelty of the current work is identifying the hot zones (high erosion zones) with respect to opening of the choke valve and mass flow rates. The findings of the study will help in accessing where the cladding is necessary. The analysis will cover the danger zones on three critical parts of the choke valve, the needle, the seat, and the downstream piping region. To achieve this, we will also need to simulate multiple high speed, compressible flow cases. We will discuss the relation between downstream jet impact regions and the pressure drop. Finally, we will compare erosion impact areas between the cases and see which cases produce the highest erosion rates.

2. Methodology

Figure 1 shows the methodology employed in the present study. The study starts with the gathering of a CAD model and simplifying it to suit the CFD simulations; this is followed by meshing and finally CFD simulations. The following section 2.1 to 2.3 will give a detailed description about each of the steps mentioned in Figure 1.

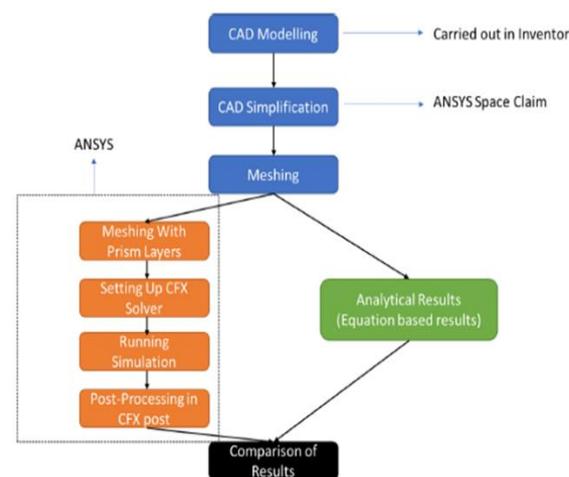


Figure 1: Flowchart of the methodology.

2.1. CAD Model Simplification

The model shown in the Figure 2A, was obtained from open-source platform called GrabCAD. A needle seat type production choke valve is used to carry out the present study. The needle and seat valve is adjustable, which means it can be used to change and control the flow and the pressure parameters. When the choke is completely closed, the needle sits tight into the seat, restricting all flow. The current study investigates a valve where the inlet and outlet dimensions are 2” and 1”,

respectively. When simulating a flow field, the important part of the valve geometry is the internal volume. The internal volume was isolated, and all other unnecessary parts of the valve were discarded. The inlet and outlet regions were also extended to ensure fully developed flow through the choke valve as seen in Figure 2B.

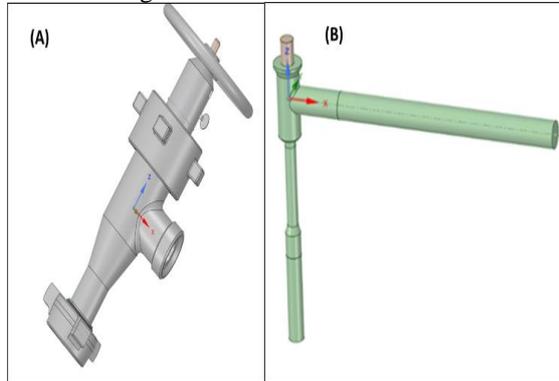


Figure 2: Original and Simplified CAD model.

2.2. Meshing

A finite volume unstructured meshing was used to mesh the CFD domain shown in the Figure 2B using the inbuilt ANSYS Meshing tool. To decide the number of the mesh elements and to get accurate results using CFD, a mesh sensitivity analysis is necessary. The idea is to run the same simulation with different mesh element sizes and compare the results. When the results do not change even though the number of elements is increasing, the desired mesh size to continue the analysis is found. CFX settings for mesh sensitivity analysis, were to set steady state, with air ideal gas at 25 degrees Celsius. A mass flow inlet rating 25 kg/s, and a static pressure outlet that is 200 bar.

The analysis consisted of five different mesh sizes with 1.5, 3.1, 4.6, 6 and 8.1 million elements, respectively. Maximum velocity at a fixed point in the downstream region and maximum velocity at inlet was compared between the different meshes and results are shown in the Figure 3 below. As the graph shows (Figure 3), the values do not change noteworthy after 4.6 million elements, so choosing a finer mesh would result in increased computational requirements without increased accuracy. Further, inflation layers were created, whose first layer size is driven by the size of the particle to be analyzed. The particle size used in the present work is $100 \cdot 10^{-6}$ m, hence the first layer inflation size is $105 \cdot 10^{-6}$ m, with 10 layers and growth factor of 1.2. Figure 4 shows the final mesh with inflation layer distribution.

2.3. CFD Simulation and Validation

A transient CFD simulations was carried out using ANSYS CFX. A SST turbulence model was

employed and the simulations were carried out for four opening of the valve (15%, 25%, 35% and 50%) and for two flow rates of gas (2.5 kg/s and 12.5 kg/s). A minimum of 30,000 particles were injected. The erosion model and its prediction is based on the DNV method (DNV 2015). First, Steady state simulation (only flow) was carried out, the results of the steady state simulations was then used as initialization for the transient state simulation (flow and sand injection), which was run for 1 second, with a max time step of 0.001s and minimum timestep of 0.00001s. During the transient simulations the sand particles were injected at the rate of 100000 per second. After one second the sand injection was stopped and simulation was run until at least 90% of the sand particles were traced back at the outlet.

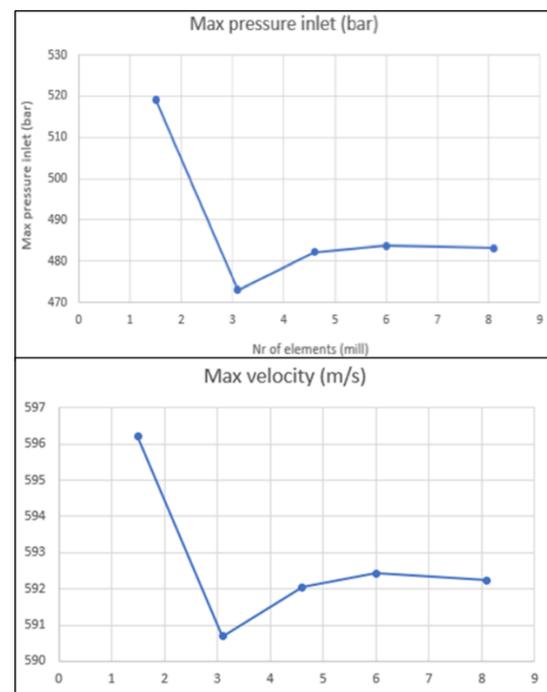


Figure 3: Mesh sensitivity analysis.

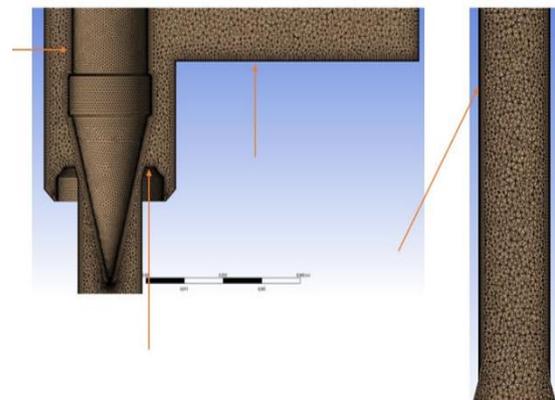


Figure 4: Mesh Distribution on the model.

3. Results

In this section results are discussed. Figure 5 shows the results for 50% opening at 2,5 kg/s mass flow rate showing the results of pressure, velocity, and particle tracks. The results show a pressure drop around the needle tip, causing the velocity to increase. At 50% opening, the velocity jet stream spreads the particles downstream, causing a wide impact region, which is different than the results are for 15% opening at 2,5 kg/s mass flow rate shown in Figure 6. At 15% opening there is greater pressure drop around the needle tip compared to 50% opening, causing the velocity to increase. The velocity jet stream carries the particles far downstream with a focused jet shape, causing almost none wall impact. Therefore, flow details the possible hot spots for the erosion to occur.

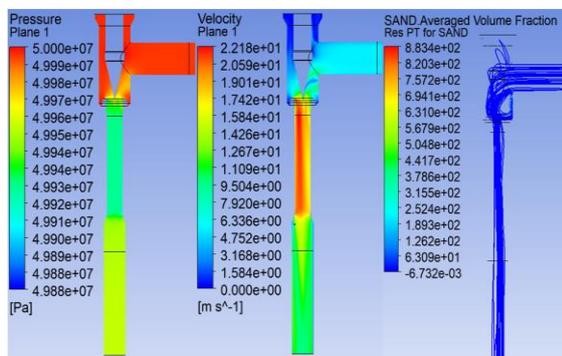


Figure 5: Pressure, velocity, and particle tracks for 50% opening and 2.5kg/s mass flow rate.

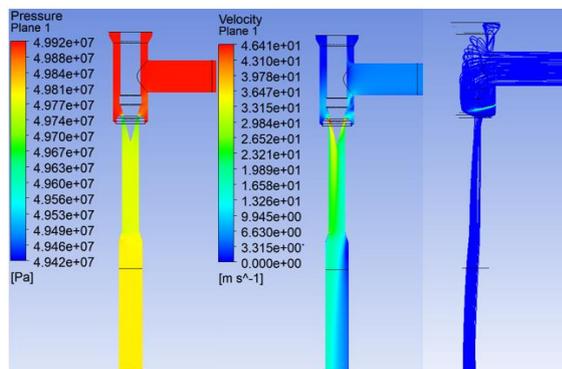


Figure 6: Pressure, velocity, and particle tracks for 15% opening and 2.5kg/s mass flow rate.

Figure 7 and 8 takes the analysis into erosion estimation. As mentioned before, the erosion estimation was carried out using DNV model, with the inputs like velocity and angle of impact calculated by the CFD analysis. Figure 7 shows the results are for the downstream region at 50% opening and 2,5 kg/s mass flow rate and shows erosion rate density at user defined values. It may be observed that most of the erosion happened immediately downstream of the choke, as predicted by the velocity results. Similar, Figure 8 shows the

erosion profile for 15% opening at 2.5 kg/s. It may be clearly observed that the jet does not significantly impact the downstream walls. Again inclining with the velocity observations.

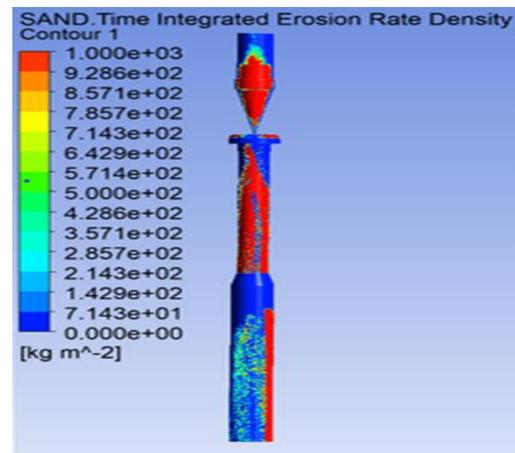


Figure 7: User defined erosion values at 50% opening and 2.5 kg/s mass flow for entire geometry.

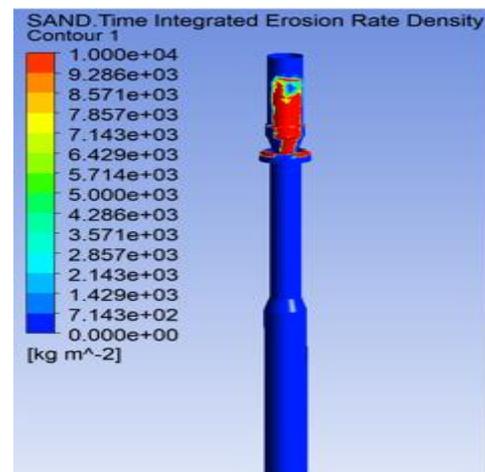


Figure 8: User defined erosion values at 15% opening and 2.5 kg/s mass flow for entire geometry.

The simulations were also carried out at different flows rate and when comparing the simulated maximum erosion rates, it is seen that the highest value appears at the needle with 15% opening and 17,5 kg/s mass flow rate. This is also the scenario where the highest pressure drop and velocity appears. Overall, the needle and seat are subject to high erosion rate, especially at higher flow rates. Erosion in the downstream (DS) region seems to depend on the velocity profile. At 50% opening the velocity jet profile is wide and reaches far DS. The erosion is spread on a large portion of the pipe walls. When the opening decreases, the velocity jet profile gets more focused, leading to higher erosion rate at a smaller area of the pipe wall. At 15% the jet is so focused that the particles are carried out of the model nearly without impacting the walls. An example of

the observed erosion rates at 15% open for different flow rates is given in Table 1.

Table 1: Comparing erosion rates at 15% opening.

Opening	Mass flow rate kg/s	Max erosion simulated (kg/m ²)		
		DS	Needle (10 ⁸)	Seat (10 ⁸)
15 %	2.5	4382	6.74	0.535
15 %	7.5	0	90.6	9.37
15 %	12.5	0	180	11.8
15 %	17.5	306161	835	35.8

4. Summary and Discussions

The carried-out simulations have in the study shown that pressure drop affects the velocity jet shape and impact region. In the 50% opening case, the pressure drop creates a wide jet stream region that, in turn, will cause an increased wall impact region. When the opening decreases, to 35% and 25%, the jet stream gets more focused. The velocity jet impacts a smaller area of the pipe wall, which in turn creates possibilities for increased erosion rate. The high pressure drop in the 15% opening case creates a high-focused jet stream in the middle of the DS pipe, leading to low wall interaction. Overall, the 50% valve opening shows the lowest pressure drops and velocities. This is the case with the largest erosion regions, but the lowest erosion rates. The erosion is evenly distributed between the three critical parts, relative to the other opening cases.

Acknowledgment

We would like to thank Eric Stephen, Kristian Ringstad, Jon Ramstad for their contribution.

References

- Bishnoi, P., et al. (2021). Effect of Velocity Variation on the Rate of Erosion Wear at the Elbow of the Slurry Pipe by Using CFD. *Advances in Fluid and Thermal Engineering*, Springer: 305-312.
- Desale, G. R., et al. (2009). "Particle size effects on the slurry erosion of aluminium alloy (AA 6063)." *Wear* **266**(11-12): 1066-1071.
- DNV, G. (2015). "Recommended Practice–DNVGL-RP-O501–Managing Sand Production and Erosion." *DNV GL: Akershus, Norway*: 3-60.
- Feng, Z. and A. Ball (1999). "The erosion of four materials using seven erodents—towards an understanding." *Wear* **233**: 674-684.
- Haugen, K., et al. (1995). "Sand erosion of wear-resistant materials: Erosion in choke valves." *Wear* **186**: 179-188.

Mathew, A. P. (2017). "Erosion rate prediction in single and multiphase flow using CFD." *Int. Res. J. Eng. Technol* **4**(7): 766-769.

McLaury, B. S., et al. (2000). "Effect of entrance shape on erosion in the throat of chokes." *J. Energy Resour. Technol.* **122**(4): 198-204.

Oka, Y. I., et al. (2005). "Practical estimation of erosion damage caused by solid particle impact: Part 1: Effects of impact parameters on a predictive equation." *Wear* **259**(1-6): 95-101.

Raghavendra, H., et al. (2014). "Simulation of erosion wear in choke valves using CFD." *International Journal of Engineering Research & Technology* **3**(7).

Wallace, M., et al. (2004). "Prediction of impact erosion in valve geometries." *Wear* **256**(9-10): 927-936.

Wheeler, D., et al. (2006). "Application of diamond to enhance choke valve life in erosive duties." *Wear* **261**(10): 1087-1094.

Simulation of different scenarios for resource utilization. Uncertainties, availability, and fair distribution of global resources

Erik Dahlquist ^{a,*}, Stefan Hellstrand ^a, Madeleine Martinsen ^a

^a *Malardalen University, Vasteras, Sweden*

*erik.dahlquist@mdu.se

Abstract

In this paper scenarios are made simulating how demand can be predicted for different materials and resources for the future when we develop new technologies to build the sustainable society for everyone. How does recycling efficiency impact total demand for virgin resources? This also means a fair distribution including all people, but with a bit different demand depending on where people live. In hot countries cooling is important, while heating in cold. But how will global warming affect this? And how can solutions like EV with batteries and H2 with fuel cells (FC), new materials for heat and cold storage etc change the demand for different materials and resources? Different scenarios are simulated, and the results discussed.

Key words: Resources, simulations, scenarios

1. Introduction

Normally we don't think too much about what resources we have available. Most people want to see that everyone should have access to what we see as a minimum standard, and as this minimum standard is continuously increasing, more resources will be needed. Renewable energy conversion techniques demand a lot of material, as well as all other appliances we demand, like vehicles and communication appliances. When a resource is scarce, we try to find other solutions instead. This makes it very difficult to predict the use of different elements by just extrapolating from the last few years utilization. On the other hand, these types of extrapolations can be very useful to give an insight in what will happen if we don't take actions.

2. Simulation model

The assumptions now can be to look for what it would mean if all persons globally would have the same usage pattern as EU-27 population today. How would that change the annual demand for materials? At the same time there are strong indications that the world population will increase significantly during the next 50 – 80 years. Especially in Africa south of Sahara the birth rate is very high, although hopefully it will decrease if the population get better living conditions. This has been the case in many parts of Asia during the last 50 years. As no one knows how many we will be year 2100 we calculate for an

increase by 1500 million respectively 3000 million, which will cover the predictions made by different organizations and researchers. A second aspect is how much virgin materials are needed respectively how much can be recycled when we have reached steady state at the living standard of the average EU-27 citizen. In table 1 below we see estimates for material use from Eurostat [1] respectively the authors own calculations [2]. As Sweden is a producer of a lot of virgin metals and paper products the energy use is higher than for the average EU citizen but gives a reasonable average if we include that there is a lot of import to EU from other parts of the world. Thus, energy figures for EU may be too low if we don't include energy for production of the imported materials.

EU-27 had 447 million inhabitants 2019, while the global population was 7 674 million. Of the global population 1 236 million are in high-income countries, 5 769 in middle-income and 668 in low-income countries according to World bank data [3]. The average GDP in US\$/capita was 44 618 for high-income, 5573 for middle-income and 810 for low-income 2019.

Table 1.

Per capita use of materials in EU-27 according to Eurostat and energy per capita in Sweden for different uses. Aggregated material groups 2019 in EU-27 with 447 million inhabitants.

	EU-27	Sweden
	kg/c	MWh/c
Agri+food	2235	3
Wood+paper	534	0,89
Metals	1436	2,3
Fossil oil,coke,.....	526	
Chemicals+pharma+rubber	232	1,2
Sand, cement (construction)	4690	
-Electricity		12,7
House hold el ca 1.5-3		
-Heat		13,9
-Transports		9,2
Total counted here	9653	43,19
Total including everything (2019)	14445	
Energy MWh/c (2019)	30,3	43,6

We use 2019 figures as 2020 expectional due to Covid-19

Sources: [1],[2] and [4].

EU-27 will represent the high income countries, but we should also be aware of the uneven distribution of wealth inside EU as well as in eg China and India. For the calculations we will use the equations below for material use:

A basic unit is $a_i = \text{ton/capita,year}$ for component/material i .

If we multiply this with the number of people in a group (N_k) like human population today (7 674 million people), forecasts for the end of this century according to different UN analysis, 9 000 resp 10 500 million people. We also can look at the number of people in low-income, middle-income, or high-income societies according to world bank definitions [3].

The amount consumed per year $X_{i,k}$ then will be for component or material i for the population-group k :

$$X_{i,k} = N_k * a_i \quad (\text{in ton/y}) \quad (1)$$

We identify the known or estimated sources of component or material i to be z_i million tons. The amount of years Y_i the known resources will last then will be calculated by

$$Y_i = z_i / X_{i,k} \quad (2)$$

For many important materials we will recycle a large portion of the material. The recycling rate is R %. The consumed amount $X_{i,k,V}$ of virgin material i then will be:

$$X_{i,k,V} = X_{i,k} (100-R) \quad (3)$$

and the number of remaining years $Y_{i,k,R}$ of nown resources will be

$$Y_{i,k,V} = z_i / X_{i,k,V} \quad (4)$$

If we want to study how much would be consumed of material i if all countries had the same use level as EU-27 we calculate this for the world (G

$i, \text{EUlevel}$) from per capita figures for EU-27 for eg the global population (N_k)

$$G_{i, \text{EUlevel}} = a_{i, \text{EU27level}} * N_k \quad (5)$$

Sometimes it is interesting also to look for usage of several different materials and we then can summarize G_i for M components or materials i :

$$G_{\text{tot}} = \sum_{i=1}^M G_i \quad (6)$$

Concerning energy, we have a slightly different procedure as the usage is split at different applications like transport, residential, industry and others.

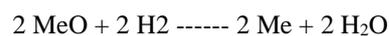
The total energy E_i used for material i will be the sum of mass in kg times energy per kg for virgin material $E_{\text{pkgV},i}$ respectively recycled material $E_{\text{pkgR},i}$:

$$E_i = (100-R)/100 * X_{i,k,V} * E_{\text{pkgV},i} + R/100 * X_{i,k,R} * E_{\text{pkgR},i} \quad (7)$$

When it comes to replace fossil fuels by renewables, we will have three main cases. The first will be just to replace fossil fuel with a biofuel. Then we will have the energy per kg for fossil fuel x , $E_{x,i}$ with efficiency for conversion $\eta_{x,i}$ for eg reduction of metal oxide MeO to metal Me and the corresponding energy per kg for biomass, $E_{\text{bio},i}$, and conversion efficiency $\eta_{\text{bio},i}$. The correlation between the two will be

$$E_{\text{bio},i} = E_{x,i} * (\eta_{\text{bio},i} / \eta_{x,i}) \quad (8)$$

For some specific reaction and processes we can also add several conversion steps with separate conversion efficiencies. If we take the case with using H_2 instead of C we will first have the actual chemical reactions to consider:



What we can see here is that we need twice as many moles to convert 1 MeO with H_2 compared to C , assuming the same conversion efficiency η_c for both cases. If we then look at the losses on the way, we will have small extra losses for the C case while the production of H_2 from water using an electrolyzer will have efficiency $\eta_{\text{electrolyser}}$ and compression of the H_2 gas will have efficiency $\eta_{\text{compression}}$. If the gas is to be used for EVs (Electric vehicles) using Fuel Cells (FC) we will need to add the fuel cell efficiency η_{FC} .

The total efficiency ($\eta_{\text{tot}, \text{H}_2}$) for the H_2 compared to C then would be:

$$\eta_{\text{tot}, \text{H}_2} = 1/2 (\text{Cmol}) * \eta_{\text{electrolyser}} * \eta_{\text{compression}} * \eta_{\text{FC}} \quad (9)$$

where $\eta_{\text{FC}} = 1.0$ if no Fuel Cell in the system. Normally we can assume approximately the following efficiencies today: $\eta_{\text{electrolyser}} = 0.5-0.7$; $\eta_{\text{compression}} = 0.9$ if 250 bar ; $\eta_{\text{FC}} = 0.5-0.7$. If we multiply this assuming reasonable figures of today

we get $\eta_{\text{tot,H}_2} = \eta_{\text{electrolyser}} * \eta_{\text{compression}} * \eta_{\text{FC}} = 0.6 * 0.9 * 0.6 = 0.32$ or 32 %. To this we have twice as many moles of H₂ compared to C.

For a conversion of a vehicle with ICE (internal combustion engine) using fossil fuel to an EV we have the conversion efficiency $\eta_{\text{c,ICE}}$ in the ICE engine at approximately 0.30-0.45 depending on driving in the city or at the highway. This means that we normally need some 0.5-0.7 liter/10 km or some 5-7 kWh/10 km with the internal combustion engine. If we drive an EV with battery the corresponding calculation η_{battery} including both charging and discharging is 0.9-0.95. The electric engine efficiency η_{eleng} will be very high; 0.95-0.97 is realistic. This gives a total efficiency $\eta_{\text{tot,el}}$ around 0.86-0.92.

$$\eta_{\text{tot,el}} = \eta_{\text{battery}} * \eta_{\text{eleng}} \quad (10)$$

For the case with H₂ we can use the equation from earlier and then get approximately $\eta_{\text{tot,FC}} = 0.32$. This shows that from an energy perspective it is much more efficient to use battery electric solutions compared to H₂/FC. On the other hand – H₂ can be stored easier at a large scale than electricity in batteries, which can still give advantages with H₂ from a storage perspective, at least for more long-term storages. Hydrogen also can be stored in chemicals like NH₃.

Concerning emission of CO₂ equivalents from power production we have principally the conversion relation that 1 C consumes 1 O₂ to produce 1 CO₂. If we know the fuel composition C_xH_yO_vN_zS_p we thus can easily calculate the emissions, if we know the amount of fuel being used. Unfortunately, the composition can vary a lot for coal, oil and biomass, which can give high uncertainties. In table 2 we have the total global figures for energy “consumption” respectively “production” in ton per capita in ton oil equivalent per year.

Table 2. ‘

Global use in ton per capita for different usage areas and different energy sources [5]

Globally ton/capita	NG	oil	coal	electricity	"Total"	% of con
Total Mtoe final consumption	0,20993	0,527886	0,129528	0,250065	1,295022	(excl bio)
Total Mtoe production	0,441723	0,578447	0,482799	0	1,861089	of all
Industry	0,077912	0,038168	0,100514	0,105027	0,321622	24,8356
Residential	0,062744	0,028506	0,009715	0,067268	0,168232	12,9909
Transport	0,015272	0,344182	0	0,004251	0,363705	28,0853
Commercial & public service	0,027	0	0	0,053764	0,080764	6,23662
Non-energy use	0,025192	0,088157	0,006735	0	0,120084	9,272
other	0,001889	0,029034	0,012564	0,019755	0,063243	4,88359
Total exclud bio+waste+other					1,11765	86,3050

Unfortunately collected data is primarily distributed on NG, oil, coal, and electricity as the values for biomass and waste are so uncertain. Thus, the sum of use here is only giving 86.3% of the estimated total energy use. In the last column thus % of what is measured is given as well. What we can see is that Industry and Transport sectors are dominating.

Another interesting variable is how much CO₂ equivalents that is emitted in the different applications as well. In table 3 we see the energy utilization annually as well as estimated emission of CO₂ equivalents for three of the major emitters.

Table 3.

Global use of energy and emission of CO₂ eq 2018.

Data for energy use [6]	TWh/y	Gton _{CO2eq} /y	Mtoe/y
Global use 2018			
TWh/y and Gton CO _{2eq} /y	TWh/y	Gton _{CO2eq} /y	Mtoe/y
Industry	43 610	8,5	3180-3677
Transport	30479	7,3	2783
Agriculture&forestry (non-ene)	5800	9,3	922
Buildings	34881	10	2981- 3478
Total use	116269	36	9938

The high and low figures for buildings and industry is if construction industry should be included in one or the other post. The total emission globally is approximately 36 GtonCO_{2eq}/y, where most of the rest is for energy use (electricity and heat/cooling). When we summarize these components, we reach 35.1 GtCO_{2eq}/y. Especially battery or fuel cell/Hydrogen vehicles or industrial processes as replacement of fossil fuels can be studied as described earlier. If we assume EVs with battery the global energy use can be reduced as batteries with electric motors are much more efficient than ICE internal combustion engines. Assuming that the electricity is made from renewables or nuclear the CO_{2,eq} emissions will be dramatically reduced, but also the total energy reduced to roughly 1/3 of the demand today [7]. For production of Hydrogen by electrolysis, storage and then combustion in a fuel cell gives an increased energy demand by approximately 2.5-2.7 times compared to what is used with coal or other fossil fuel [8]. On the other hand, also here the CO₂ emissions can be reduced in the same way as with batteries.

So, what would it mean in material use to produce the electricity with wind, PV, hydropower or biomass-CHP? In table 4 below approximate figures for how much material is needed per MW installed capacity. The electricity production still will depend on the capacity factor Cp. Eurostat gives the average figure Cp for EU 2019 to be 26%, with 24% on-shore and 38% off-shore. For PV cells it is lower but depend a lot on the geographic site. For CHP and hydropower with reservoir for storage the Cp can be varied intentionally, which is an advantage compared to wind and PV.

Table 4.

How much material is needed per MW installed capacity and how much of first year electricity production is needed to compensate for energy input in building plants. [9], [10], [11], [12] and [13]

	PVcells	Wind	CHP or CC
Concrete	0	73%	75.8%
Steel	0	25%	24%
Glass	76%		
Aluminium	8%	0	0.2%
Copper			0.16%
Plastic	10%	1.7%	
Oil/lubricant 20y		0.1%	
Total ton/MW	840	515	129
Energy input as % of one year production	39%	4-18%	8%

To operate the transport fleet with EVs we would need some 15 000 TWh/y of electricity. From Table 4 we can see that the amount of materials in ton per MW is 129-840 ton/MW. In table 5 we have how many GWh/GW we get according to statistics on installed capacity and production per year. If we assume a mix with 1/3 of each Wind, PV and Thermal power we get $5000\ 000\ \text{GWh}/2197 = 2275\ \text{GW}$ wind, $5\ 000\ 000\ \text{GWh}/1145 = 4366\ \text{GW}$ solar power and $5\ 000\ 000\ \text{GWh}/4366 = 1145\ \text{GW}$ thermal, assuming same efficiency for coal fired and biomass fired CHP plants.

Table 5.

Energy as TWh/y, GW installed capacity 2019 and GWh/GW

Type	TWh/y	GW install	GWh/GW
	2019	2019	
PV	720	629	1145
Wind	1430	651	2197
Coal	9168	2100	4366
NG	6250	1812	3449
Nuclear	2500	390	6410

From table 4 we have 840 ton/MW for PV cell system, meaning $840\ \text{ton/MW} * 4\ 366\ 000\ \text{MW} = 3,67\ \text{Gton}$. For wind power we get $515\ \text{ton/MW} * 2\ 275\ 000\ \text{MW} = 1.17\ \text{Gton}$ and for thermal power $129\ \text{ton/MW} * 1\ 145\ 000\ \text{MW} = 0.148\ \text{Gton}$ materials. For wind and thermal power, we will put in approximately 10 % of first year production into the manufacturing, while approximately 4 times as much for PV cells with today's technology [9]. From this we see that CHP demands least materials for construction but on the other hand demand fuels for operation.

3. Results – Scenarios

There are two major challenges for the human population. The first is the issue with fair distribution of resources. What would it mean if everyone would have the same living standard as the average EU-27 person, for world resource utilization? The second big challenge is what it will

mean if world population proceed to increase significantly. In table 6 we see how demand for resources would be impacted under these conditions.

Table 6.

Scenarios for demand for resources if everyone globally should use same amount of resources as an average EU-27 citizen respectively how much will be needed if population increase to 9000 respectively 10500 million inhabitants.

	Population			Assuming EU-27 level			Total today	Diff today - to EU-27	Diff today - to 9000	Diff today - to 10500
	EU-27	Sweden	Wrlid if Eulev	Mton/y	Mton/y	Mton/y				
Population in millions	447		7674	9000	10500					
Agri+food	2,235	3	17151	20115	23468	8000	9151	12115	15468	
Food including stalks etc			17151	20115	23468	16000	1151	4115	7468	
Wood+paper	0,534	0,89	4098	4806	5607	600	3498	4206	5007	
Forest products totally			4098	4806	5607	2999	1099	1807	2608	
Metals	1,436	2,3	11020	12924	15078	2000	9020	10924	13078	
Fossil oil,coal, NG	2,35	10,9	18034	21150	24675					
Plastics, Chemicals+pharma+ru	0,232	1,2	1780	2088	2436	300	1480	1788	2136	
Cement	0,317		2433	2853	3329	4100	-1667	-1247	-772	
Sand, cement (construction)	4,69		35991	42210	49245					

What we can see here is that food and agriculture is of highest importance. The total production and use today are approximately 8 000 million tons/y but would be more than double this if we extrapolate EU-27 figures to all people globally. Roughly twice as much biomass is produced including stalks and leaves giving approximately 16000 Mton/y. Approximately 30% of what we produce is wasted due to different reasons like poor storage, food getting to old and thus thrown away etc. In figure 1 we see annual consumption today compared to if all at EU-27 level, and if we become 9000 or 10500 million inhabitants at EU-27 level.

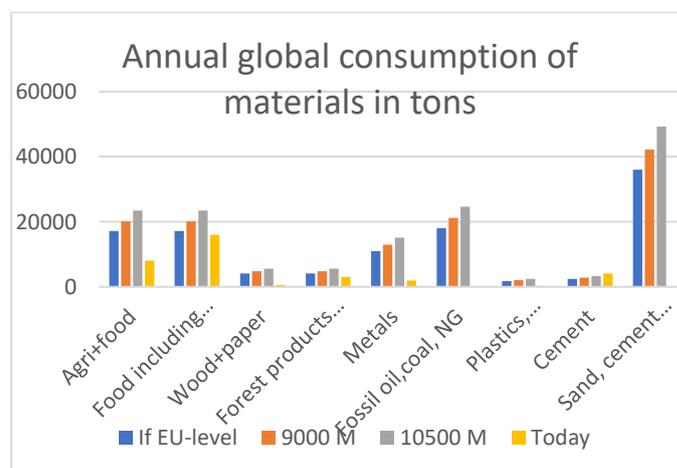


Figure 1. Annual consumption in million tons for different world populations – A. As reference we have the total global use today, 7 674 million 2021 (today). B. same assuming same level as EU-27 and scenarios with C. 9 000 respectively D. 10 500 million inhabitants.

For wood and paper products we have a similar discussion. If we use EU-27 average per capita we get 4098 Mton/y globally while official figures on

global level says 2999 Mton forest products/y. If we then take the production of pulp + paper this is 600 Mton/y. A significant amount of the rest is for direct use of wood for constructions but also as biomass fuel. Still, we have a difference between the average global use per capita and the EU-27 level. As internet trading is increasing, we see more packages and boxes, but less printing and journal paper demanded. Tissue like toilet paper is increasing where it has not been used so much before. But also, we see a new demand for replacement of oil and plastic by wood fibers.

For metals it is complex to take averages as in some countries we have a very high recycling rate of metals in relation to virgin materials compared to others. Even if you use the same amount as kg/capita it can still be very big differences with respect to amount of virgin material. This gives a very high amount if we multiply average use in kg/c.y in EU-27 with number of people in the world, as a lot of the EU-27 use is recycled metals. For Iron for instance, recycling rate is in the range of 88%. Thus, it may be more interesting to look at the total metal production which is 2000 Mton/y. A fact is that if everyone should have their own Electric vehicle the amount of metal will increase a lot, as many of the existing population still hasn't any car. If we then add 1500 million or 3000 million people, we can see that the demand increases dramatically (see figure 1). From weight perspective cars are very important, but also armament in concrete buildings or metal as construction material. Batteries and wind power plants as well as power transmission and distribution will affect the total use a lot next 20 years for transfer into non-fossil energy, but also even more if the world population increases even further. This will include also more demand for cement and other bulk materials like sand and gravel.

Energy is another aspect of resources. In table 2 we see how energy is used per capita in Sweden, which is one of the EU-27 countries with a high energy demand per capita. The total use is 43 MWh/c.y from which 10.9 MWh/c.y from fossil sources. Approximately 3 MWh/c.y is for food and food production while 0.89 MWh/c.y for wood and paper. 2.3 MWh/c.y is for metals and metal production while 1.2 MWh/c.y for plastics and chemicals. This is for a country with relatively high amount of virgin wood fibers and metals, so average EU-27 will be significantly less if the share of recycled materials is high, as the case with pulp and paper and metals generally. We still can see that the amount of energy for materials (7.4 MWh/c.y) is significantly less than the amount used for transport (9.2 MWh/c.y), heating (13.9 MWh/c.y) and electricity (12.7 MWh/c.y). The total use varies between countries and especially the national power mix varies a lot. In Sweden more than 90% of electric power is from

renewables or nuclear and most of the fossil CO₂ emissions comes from other countries through different products, fuel for transportation or for industry use. As at average 23 TWh out of 158 TWh produced is exported annually, the "net" electricity production is close to 100% non-fossil. In a country like Poland with large amount of coal the situation is the opposite, although also Poland is driving renewable energy a lot, and the same goes for Germany, Denmark and Finland, although with different mix of technologies.

Without going too deep into the relation between energy and materials, we can see that the average energy input per kg for several products was presented in table 1. When recycling e.g. iron the kWh/kg decreases by 75%, which shows that we both get a positive effect with respect to demand for new metal, but also reduces the energy demand per kg dramatically. This shows us that from a material use perspective increased material recycling is an important move to take.

The power demand for electric vehicles, replace fossil fuels in process industries and similar will be both higher and lower than today. For an electric vehicle consuming 5-10 kWh/10 km an electric vehicle may consume 1.5-2 kWh/10 km. This means only 1/3 of the energy demand compared to today. For energy calculations we use the set of equations presented earlier.

4 Discussion about uncertainties

Concerning energy, we have a lot of data on production and consumption of oil, coal and natural gas, but also electricity, biomass, waste and "others". Usually, you recalculate all to be TWh/y or Mtoe/y, million-ton oil equivalents, for global figures. Then we must be aware of that heating values differ a lot between different qualities. For NG we have 13.5+- 1.8 kWh/kg; for crude oil 12.4+- 0.7; for coal 5.8+- 1.1. Even for Hydrogen, H₂, we have a span 36.4+- 3.1 [14]. If you have detailed analysis of all materials and summarize these it should be very accurate, but we will have to expect that some values are more accurate than others. Even if everyone is trying to collect and report as correct as possible, we still will have uncertainties as this is done in many stages and with several recalculations from one sort to another. We see that the reported total production is 14 282 Mtoe 2018 for all products, while final consumption ends up at 9938 Mtoe 2018. The difference is losses in one way or another. If we for instance look at the deviation between produced and consumed electric power, we see losses of approximately 7% in many countries, but with extremes of more than 50% in e.g. Benin. Do people steal power or are the measurements just poor? If we extend the transmission of power over

long distances we can expect higher losses, but with new technologies they can be reduced as well.

Another uncertainty is how materials are defined. When we look at official figures of cement the value is 4100 Mton/y, but in some official sources we have 4100 Mton/y of concrete as well, while others report 30 000 Mton/y for concrete, which makes more sense. These are all official figures, but the nomenclature has obviously not been correct.

When it comes to grouping data into categories it is also difficult to compare data from one source to data from another. If we look at the EU-27 data these are valid from 2020, but EU-28 before UK left the union, up to 2019. It is also difficult to know what components are aggregated in different sources.

Concerning CO₂ equivalent emissions, we come to even more tricky considerations. For agriculture we see figures stating that 9.3 GtCO₂eq/y is emitted, from which 5.3 Gt from crops and livestock and 4 Gt from change in land use. These figures are built on some few measurements in a few countries and then extrapolated for the globe. The uncertainty is huge. If we look at N₂O we have seen in measurements that almost all is emitted during a few weeks when the land is covered by melting ice, and there has been a lot of Nitrogen left in the soil. If we avoid fertilizing before the crop has come up to some 10 cm, we can almost eliminate this. For methane the water level in wet land is very important. When water is almost up to the surface, we get larger emissions than if it is some 0.8 m below the surface, but it also depends on how the water is flowing through the wet land. We mostly guess how much is emitted by extrapolating from very uncertain measurements.

4. Summary and Conclusions

From the simulations we can determine effects of different developments with respect to population increase, fair distribution of materials, energy use from different sources and how long it will take until resources are depleted without recycling respectively with recycling. This is important for how politicians should make their decisions on rules to direct use of materials and how to reduce fossil CO₂ emissions. If all had EU-27 level of living standard the impact on metals, plastic and some chemicals will be very high while marginal on food, forest products and concrete. For increased population the increase will be in proportion to how many more people there will be.

Acknowledgment

We thank our cooperation partners LKAB (SUM Academy), Mälarenergi and Eskilstuna Strängnäs Energi och Miljö (ESEM) and EU ERA NET projects IFAISTOS and DISTRHEAT for financial support and technical input.

References

- [1] Material flow accounts in raw material equivalents by final uses of products - modelling estimates [ENV_AC_RMEFD\$DEFAULTVIEW]. 27/09/2021. <https://ec.europa.eu/eurostat/web/environment/material-flows-and-resource-productivity>
- [2] Dahlquist, E., & Hellstrand, S. (2017). Natural resources available today and in the future: How to perform change management for achieving a sustainable world (pp. 1-304). Springer International Publishing. <https://doi.org/10.1007/978-3-319-54263-8>
- [3] World bank indicators, (down loaded data 2022). <https://data.worldbank.org/indicator>
- [4] SCB, <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/environment/environmental-accounts-and-sustainable-development/system-of-environmental-and-economic-accounts/pong/tables-and-graphs/economy-wide-material-flow-accounts/domestic-material-consumption-per-category-of-material-sweden-1998-2020/>
- [5] Eurostat: Primary energy consumption [T2020_33]. By "Primary Energy Consumption" is meant the Gross Inland Consumption excluding all non-energy use of energy carriers. Last update 21 Dec 2021. Energy distribution on activities and fuels
- [6] Fatih Birol: Key World Energy Statistics 2020. CO₂ equivalents. IEA August 2020.
- [7] U.S. Energy Information Administration. International Energy Outlook 2016.
- [8] Hydrogen efficiency (IVA report 2022, not yet published May 2022)
- [9] Elena Tomas-Aparicio, Erik Dahlquist, Jinyue Yan, Konstantinos Kyprianidis, Bertil Moritz: Comparison between different renewable energy solutions from a materials and CO₂ perspective. Conference proceedings for International Conference on Applied Energy 2020, Dec. 1 - Dec. 10, 2020, Bangkok / Virtual, Paper ID: 518.
- [10] Haapala, Karl R. * and Preedanood Prempreeda: Comparative life cycle assessment of

2.0 MW wind turbines. Int. J. Sustainable Manufacturing, Vol. 3, No. 2, 2014

[11] Jan Wenske: Article Monitoring the Oil of Wind-Turbine Gearboxes: Main Degradation Indicators and Detection Methods Diego Coronado 1 and Jan Wenske 2,* 1 OELCHECK GmbH, Kerschelweg 29, 83098 Brannenburg, Germany; dcg@oelcheck.de 2 Fraunhofer Institute for Wind Energy Systems IWES, Am Luneort 100, 27572 Bremerhaven, Germany * Correspondence: jan.wenske@iwes.fraunhofer.de Received: 15 March 2018; Accepted: 5 June 2018; Published: 8 June 2018.

[12] Spath Pamela L. , Margaret K. Mann [2000] : Life Cycle Assessment of a Natural Gas Combined-Cycle Power Generation System. September 2000 i NREL/TP-570-2771

[13] Spath Pamela L., Margaret K. Mann, and Dawn R. Kerr: Life Cycle Assessment of Coal-fired Power Production.

<https://www.nrel.gov/docs/fy99osti/25119.pdf>,
June 1999 • NREL/TP-570-25119.

[14] Higher heating values: <https://world-nuclear.org/information-library/facts-and-figures/heat-values-of-various-fuels.aspx>

Modelling of liquid injection of ammonia in a direct injector using Reynolds-averaged Navier–Stokes simulation

Jessica Gaucherand^{a,*}, Corinna Netzer^a, Michał T. Lewandowski^a, Terese Løvås^a

^a Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Kolbjørn Hejes vei 1a, Trondheim, 7491, Trøndelag, Norway, jessica.gaucherand@ntnu.no

Abstract

Ammonia as a fuel has gotten attention in the past years to enable decarbonization for internal combustion engines. There is a need to understand the behavior of the liquid fuel in direct injection engines, a crucial step of the engine cycle. Injection impacts the mixture formation in the cylinder, equivalence ratio, combustion and pollutants formation. Liquid ammonia is expected to behave significantly different than traditional fuels during the injection phase and hence requires to be investigated. Indeed, recent experimental research has highlighted the appearance of flash boiling during injection of ammonia spray under engine-relevant conditions. The high volatility is also expected to influence the cavitation behavior. Cavitation is the partial vaporization of the liquid typically caused by locally increased velocity resulting in a pressure drop, when the fluid enters an orifice with sharp edges. One parameter controlling cavitation is therefore the geometry, but cavitation is also influenced by the fuel's property and the boundary conditions. This study presents 3-D RANS simulations performed with CONVERGE CFD of the internal flow of a Gasoline Direct Injector (GDI), operating with liquid ammonia. The transient simulations account for the injector needle movement. Simulations capture the presence of both vapor and liquid in the nozzle head. These results from the simulations will provide input data for separate spray simulations with the same engine geometry, as well as support the development of a 0-D model that will be important for design purposes. Preliminary results predict a liquid discharge coefficient of less than 0.1 at the outlet for injection in atmospheric conditions.

1. Introduction

Reduction of greenhouse gas emissions to mitigate climate change requires a change for the transportation sector both in the form of vehicles electrification or towards the usage of carbon-free fuels for internal combustion engines (ICE). Ammonia (NH_3) is one of the fuels of interest and has been investigated as a promising fuel for internal combustion engines (Lhuillier et al., 2020; Mounaïm-Rousselle et al., 2022): it can be stored in liquid form at an ambient temperature at 9 bar, has a high hydrogen content and no carbon. However, the fuel characteristics are very different from conventional fuels, including its behaviour in the injection system. The fuel injection in the engine is an important process deciding on mixture formation, equivalence ratio, combustion, and pollutant formation. So far the majority of ammonia applications in ICE have been performed with port injection, with the fuel in gaseous phase, leading to purely premixed combustion (Mounaïm-Rousselle et al., 2021; Niki et al., 2016). On the other hand, direct liquid ammonia injection allows for other combustion modes with more strategies to explore and can potentially reduce ammonia slip (unburned ammonia in the exhaust) (Reiter and Kong, 2008). In other words, studying the injection of liquid ammonia is of interest to optimize internal combustion engine performances and emissions with this fuel. There is little literature on ammonia sprays relevant for engine conditions, yet recent experimental works (Pelé et al., 2021; Cheng et al., 2022) shed some light on this topic. To

gain further insights, numerical investigations of fuel in-nozzle flow allows to compare the fuel's behavior during injection, capture phase change in the injector (Torelli et al., 2017), and study spray formation in the engine (Saha et al., 2017).

This study aims to investigate numerically the internal flow of liquid ammonia in a gasoline direct injector (GDI). GDIs, which are solenoid activated injectors for high-pressure direct injection, are used in ICEs as they allow rapid response, increased engine efficiency, and reduced emissions (Wang et al., 2014). The flow travels throughout the injector due to a strong pressure gradient between the fuel system, where the fuel is stored at high pressure, and the engine's combustion chamber. The injector's internal geometry, the nozzle material, the tank and engine pressure and temperature, and the fuel's property affect the fluid flow. It is therefore expected that the fuel arriving at the injector's outlet will not have the same properties as in the tank. In particular, the sharp edges of the injector's nozzle hole cause a large pressure drop inducing vapor formation in the injector if the pressure drops below the saturation pressure of the fuel. This phenomenon is called cavitation. Cavitation can be destructive as a vapor bubble can collapse under high pressure and create a shockwave that can damage the nozzle material. Flash boiling denotes a condition where liquid undergoes a rapid pressure reduction and partial vaporization occurs due to the fuel's property and local pressure and temperature. These conditions occur during

spray formation in the combustion chamber but may also happen inside the injector during the injection: the latter phenomenon is referred to as internal flash boiling. Therefore, two sources of energy can be the cause for fuel vapor formation inside an injector : kinetic energy, or heating of the fuel. Studying the fuel's evolution, and more specifically its interaction with the injector needle motion and the nozzle holes, will provide useful inputs on the thermodynamic state of the fuel when it is injected into the engine. Previous research has been carried out with various fuels such as diesel or biodiesel, showing that each fuel behaves differently and cavitates at varying degree in identical injectors due to their physical properties (Torelli et al., 2017). Experiments have shown that in-nozzle flow cavitation for traditional fuels impact spray formation (Serras-Pereira et al., 2010). However, it remains unknown how ammonia behaves during the injection and how this contributes to the untypical ammonia spray characteristics observed in a recent studies (Pelé et al., 2021). This forms the motivation for the present work to gain further insights by looking at the internal flow of ammonia inside of an injector numerically. Additionally, the obtained results will be needed in future work for two purposes: as valuable information and input data for Lagrangian spray setup in separate simulations in similar ICE configurations; and to support the development of a 0-D model that can be used during design process of new and optimised injectors.

First, the numerical setup will be outlined with the description of the model equations, the injector's geometry, and a mesh convergence study. Then in the second part of the paper, results will be discussed with a focus on the flow characteristic at the injector's outlet. A comparison with iso-octane will also be presented. Finally, parameters such as the output pressure, the wall temperature will be modified to see the impact on the injection.

2. CFD setup and mesh study

In this work, a six-hole counter bore GDI fuel injector with ammonia is investigated. The geometry of the injector was obtained using molds of the physical injector and measurements conducted with a scanning electron microscope. The 3-D CFD simulations are carried out using the software CONVERGE 3.0 (Richards et al., 2017) within the Reynolds-averaged Navier Stokes (RANS) framework. The Navier Stokes solver is set to Pressure Implicit with Splitting of Operator algorithm (PISO). The simulation is transient with automatically adapted variable time step. The adaptative mesh refinement as allowed with CONVERGE is used to capture the fine nozzle details and the flow around the needle as it moves. The needle in the injector moves to open and close the passage of the fluid at the nozzle head. Turbulence is modeled with a Re-Normalization group (RNG) $k-\epsilon$ model. For the sake of brevity, details of the models are not discussed here.

2.1. Multi-phase flow model

To capture the effect of cavitation in the injector, a multi-phase approach is needed to describe both liquid and vapor flow and the interactions between the two. In this work, this is done using the volume of fluid (VOF) method. The liquid phase is represented with the subscript l , and the gaseous phase by the subscript g . The global density equation is written as follow:

$$\rho = \alpha\rho_g + (1 - \alpha)\rho_l. \quad (1)$$

Where α , the void fraction, is equal to 0 when the cell contains only liquid, 1 when the cell contains only gas, and in between when the cell contains both gas and liquid. The void fraction is transported indirectly. The species transport equation solves the species mass fraction in the cell:

$$\frac{\partial\rho_m}{\partial t} + \frac{\partial\rho_m u_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\rho D \frac{\partial Y_m}{\partial x_j} \right), m = 1 \dots n \quad (2)$$

with $\rho_m = Y_m \rho$ being the density of the species m , Y_m its mass fraction and n is the total number of species. D is the diffusion coefficient. The total mass of gas (m_g) in the cell is calculated and the mass of liquid (m_l) can be deducted knowing $m_l = 1 - m_g$ and $m_g = \sum_{m=1}^{n_g} Y_m$ with n_g as the number of gas species. Finally, the void fraction is calculated with:

$$\alpha = \frac{m_g}{\rho_g} \left(\frac{m_g}{\rho_g} + \frac{m_l}{\rho_l} \right)^{-1}. \quad (3)$$

2.2. Cavitation model

Cavitation is here modeled using the flash-boiling hypothesis of Shields et al. (Shields et al., 2011). The mass exchange between vapor and liquid is modeled by describing the rate at which the mass fraction of vapor (Y_v) in a two-phase mixture approaches its equilibrium value. The rate follows the equation by Bilicki and Kestin (Bilicki and Kestin, 1990):

$$\frac{DY_v}{Dt} = \frac{\bar{Y}_v - Y_v}{\theta} \quad (4)$$

with Y_v being the instantaneous mass, \bar{Y}_v the equilibrium mass and θ the time scale over which Y_v relaxes to \bar{Y}_v (Richards et al., 2017). \bar{Y}_v is a function of the local enthalpy:

$$\bar{Y}_v = \frac{h - h_l}{h_v - h_l} \quad (5)$$

and the time scale for evaporation is defined as (Downar-Zapolski et al., 1996) :

$$\theta = 3.84 \cdot 10^{-7} \alpha^{-0.54} \phi^{-1.76} \quad (6)$$

with ϕ as the non dimensional pressure ratio:

$$\phi = \frac{P_{sat} - P}{P_{crit} - P_{sat}}. \quad (7)$$

Here P_{sat} and P_{crit} are the saturation pressure and the critical pressure respectively. The ideal gas equation of state is used to couple density, pressure and temperature of the vapor phase:

$$\frac{P}{\rho} = \frac{RT}{W} \quad (8)$$

with R as the universal gas constant and W the molecular weight of the gas. The pressure below which ammonia becomes gaseous at the range of temperature studied is low enough for the ideal gas law to be sufficiently accurate in this study.

2.3. Domain description

The geometry of the 6 holes counter bore injector is shown in Fig. 1. The nozzle holes have a diameter of 227 μm and a counter bore diameter of 453 μm as determined following the technique in (Macian et al., 2003).

At the inlet, liquid ammonia at a temperature of 300 K and a pressure of 200 bar enters the domain. At the outlet, a

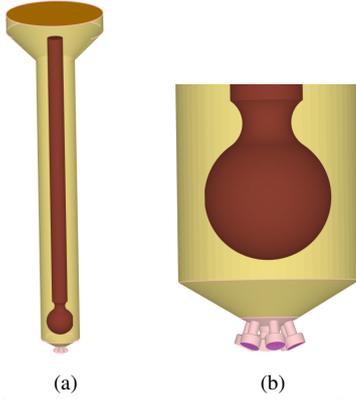


Figure 1: Geometry of the simulated injector: full injector and zoom on the nozzle head.

pressure of 1 bar and a temperature of 300 K is imposed. Liquid ammonia is specified in case of backflow. The rest of the domain is composed of walls following the law-of-the-wall boundary condition for velocity, and the heat transfer model is set to O'Rourke and Amsden (Amsden, 1997). The needle inside the injector is a moving wall with a translating motion up and down. The needle movement mimics open and closed conditions during which the fuel is either injected or not injected into the engine. As common in numerical simulation, the injector cannot be entirely shut by resting the needle directly against the nozzle seat. The minimum needle lift is here defined with $2e-4$ mm and a numerical separation between the two volumes. A secondary region at the bottom of the nozzle is defined in order to capture this (see Fig. 1b). When the needle is at its minimal lift, the injector is considered closed, and no flow travels between the region, and when the needle lifts again, the injection starts, and the flow travels between the region.

2.4. Grid convergence

A grid convergence study is carried out with a base grid of $110 \mu\text{m}$, $70 \mu\text{m}$, and $50 \mu\text{m}$. Details about the setup are presented in Table 1. The mesh for a slice in the middle of the domain is shown in Fig. 2.

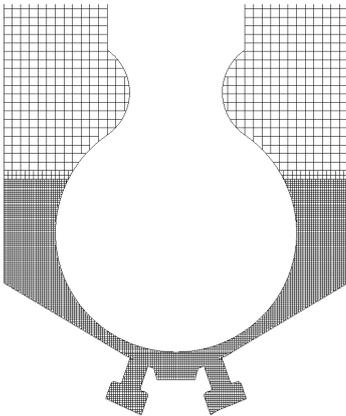


Figure 2: Mesh of the nozzle head.

Figure 3 shows the comparison of the three meshes' influence on the mass flow rate at the outlet. The normalized lift profile is also plotted. Mass flow rate is selected for the grid convergence study here. The internal

Table 1: Comparison of mesh convergence.

	Coarse	Medium	Fine
Base mesh size [μm]	110	70	50
Average time step [s]	$3.97e-8$	$3.25e-8$	$7.27e-8$
Peak cell count [cells]	$2e6$	$2.1e6$	$4.8e6$

flow distribution may require further grid refinement, which would be relevant to study for spray injection, but this is outside this work's scope.

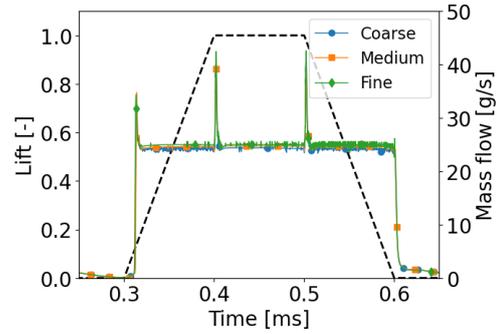


Figure 3: Comparison of the liquid ammonia flow rate during injection from the three meshes: coarse, medium and fine. The normalized lift profile is also plotted.

The results indicate accurate numerical prognosis for the three meshes. Oscillations seem to be diminished with the finer mesh but are reasonable for the other meshes. Going forward, the base grid of $70 \mu\text{m}$ is selected to study the results.

3. Results and discussion

3.1. Reference case : Ammonia at 300K

The reference case is as described in the domain description : liquid ammonia at 300 K is injected at 200 bar at the inlet, the outlet is set to 1 bar, the wall temperature is at 300 K. The mass flow rate, discharge coefficient and temperature evolution during the fuel's injection are investigated hereafter.

3.1.1. Mass flow rate during injection

Figure 3 shows the total mass flow rate through the 6 holes. When the needle reaches its final lifted position at 0.4 ms, a jump in mass flow rate injection is seen, then again at 0.5 ms when the needle starts to move down again. This is linked to a pressure surge that occurs as the needle stops in its trajectory, and then again when it starts to move from its idle position. This pressure jump, shown Fig. 4 is a numerical artefact, linked to the simulated movement of the needle in the injector. The fuel at the inflow is injected at a pressure of 200 bar, but as can be seen in Fig. 4, the average outflow pressure during the injection is around 40 bar due to the pressure losses in the injector and cavitation effect. The shown pressure is an average of the liquid and vapor pressure at the outlet. In future studies it would be interesting to evaluate the partial pressures of vapor and liquid, and also identify pressure losses due to friction.

3.1.2. Discharge coefficient

One of the objectives of this paper is to identify the discharge coefficient, a coefficient comprised between 0 and 1 that is an indication of how smoothly the flow

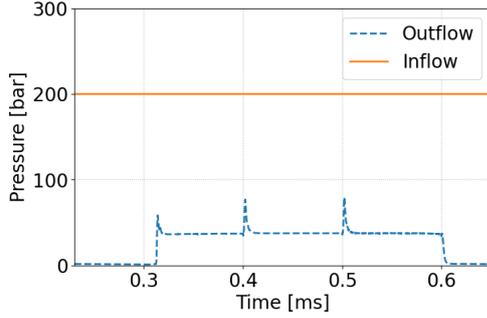


Figure 4: Time evolution of the pressure during injection at the inlet and the outlet.

passes through an orifice. This information can then be used as input for Lagrangian spray simulations or 0-D modeling. The discharge coefficient C_d is defined as the ratio between the theoretical ideal liquid mass flow rate through the outlet and the actual liquid mass flow rate. The theoretical liquid mass flow rate is determined using the following formula based on Bernoulli equation:

$$\dot{m}_{ideal} = A\sqrt{2\rho\Delta P} \quad (9)$$

with A being the nozzle outlet area, ρ the inlet density, and ΔP the difference between the upstream and downstream pressure. The discharge coefficient is defined as:

$$C_d = \frac{\dot{m}_{real}}{\dot{m}_{ideal}} \quad (10)$$

with \dot{m}_{real} as the computed mass flow rate of liquid ammonia through the outlet.

To obtain the liquid mass flow rate, the fraction of liquid at the outlet can be computed using the averaged density at the outlet, as well as the density of gaseous ammonia (10 kg/m^3) and liquid ammonia (600 kg/m^3):

$$\rho_{avg} = X_g\rho_g + X_l\rho_l \quad (11)$$

with

$$X_l + X_g = 1. \quad (12)$$

The resulting liquid fraction of ammonia is shown in Fig. 5, comparing the outflow and inflow liquid fraction. A significant reduction of the liquid fraction is predicted inside the injector, from a fully liquid fuel at the inlet to only 20% of liquid fuel remaining at the outlet. This indicates a significant degree of cavitation that has to be taken into account when studying ammonia sprays in practical systems.

Finally, a liquid discharge coefficient of less than 0.05 is computed, as seen Fig. 6. This is a very low value, and is likely due to the shape of the injector, especially the impact of the counter bore that is causing cavitation, vaporizing the fuel inside the nozzle.

Figure 7 shows the void fraction during the fuel injection when the injected flow has reached steady-state corresponding to Fig. 5. Here it is noticeable that vapor is starting to form when the flow enters the nozzle, caused, among other things, by the sharp edge of the orifices. As the flow interacts with the counter bore, more vapor is formed due to the area expansion. Indeed, the nozzle geometry causes the flow to accelerate (Fig. 8 and Fig. 9). The strong acceleration causes a pressure drop in the nozzle as seen in Fig. 10 due to the high kinetic energy of the fuel. When the pressure is below the

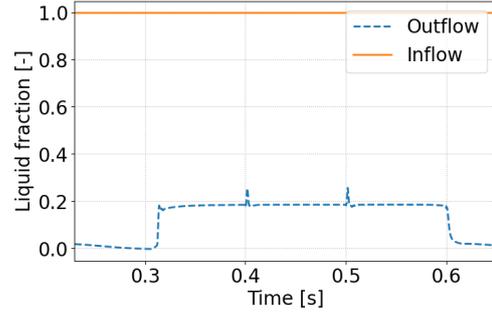


Figure 5: Time evolution of the fraction of liquid during injection at the inlet and the outlet.

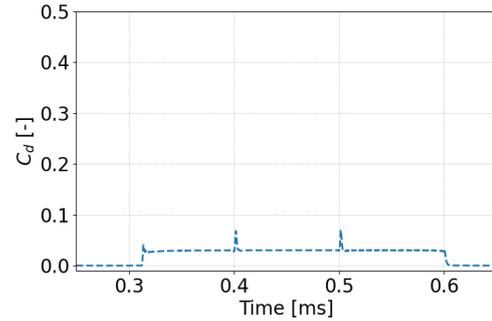


Figure 6: Time evolution of the discharge coefficient at the outlet during injection.

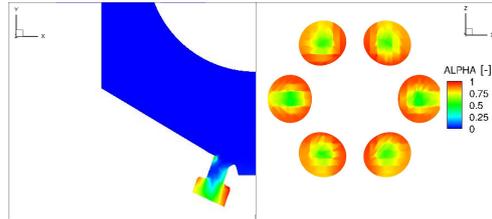


Figure 7: Contours of void fraction at steady state condition.

saturation pressure at the given temperature, the phase change occurs. This is observed here in the nozzles as the pressure drops below 11 bar at 300 K.

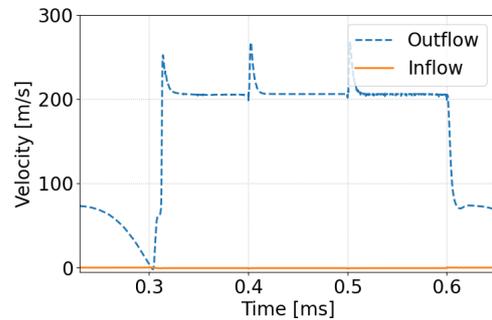


Figure 8: Average velocity at the outlet during the injection.

3.1.3. Temperature evolution during the injection

The average temperature at the inlet of the injector and its outlet during the injection is shown in Fig. 11. This

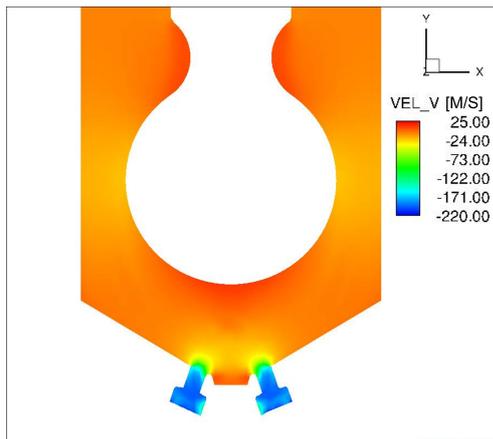


Figure 9: Contour of velocity component v at steady state condition (the values are negative due to the axis orientation).

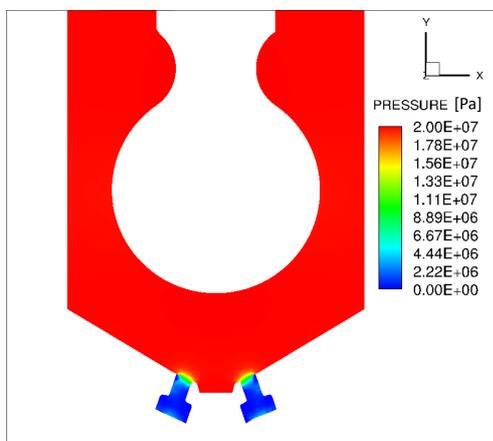


Figure 10: Contour of pressure at steady state condition.

plot shows the average contributions of both the vapor and the liquid phase. Before the start of injection ($t < 0.3$ ms), the fuel at the outlet is fully vaporized. The velocity (Fig. 8) indicates that the gas is slowing down due to its expansion in the nozzle counter bore. This is because the flow is subsonic as seen in Fig. 12. The gas undergoing expansion cools down, explaining the temperature drop to almost 270 K. This can be explained by the positive Joule-Thomson coefficient at 300 K that describes the cooling of gaseous ammonia during expansion (Beattie, 1930). During the injection ($0.3 > t > 0.6$ ms), the liquid fuel undergoes phase change with the liquid fuel in the injector vaporizing. Energy from the fuel is lost to the system during the phase change with latent heat exchanged with the boundaries such as the walls. The injected fuel's temperature therefore remains below the inlet temperature. Figure 13 compares the fuel's heat and kinetic energy. The negative heat illustrates that energy from the fuel is lost to the system. Despite this, vapor formation still occur due to the pressure drop caused by the increased kinetic energy of the fuel, as explained in the previous section.

3.2. Influence of the fuel : Comparison with iso-octane

An additional simulation is set up with iso-octane to compare ammonia's behavior to a traditional fuel in the same injector. Iso-octane is a surrogate for gasoline, a well-known fuel used in engines, notably in direct injection (Zhuang et al., 2017). Because iso-octane will

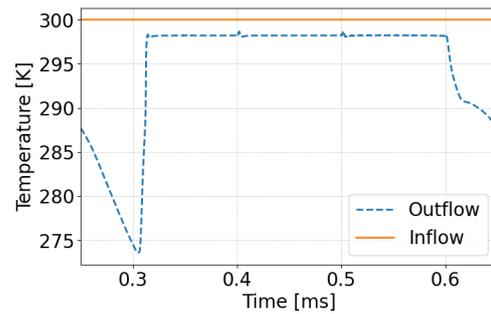


Figure 11: Time evolution of the temperature during injection at the inlet and the outlet.

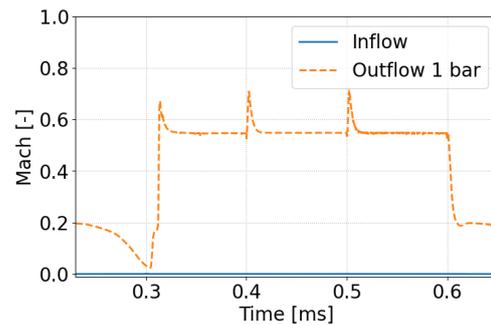


Figure 12: Time evolution of Mach number during injection at the inlet and the outlet.

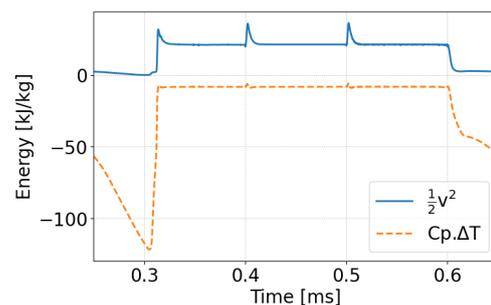


Figure 13: Energy evolution at the outlet during the injection.

not undergo phase change at the given pressure range at 300 K, the temperature is raised to 500 K in this simulations (see Fig. 14). Comparing the two fuels allows the evaluation of the fuel's property's influence on the injection process. Figure 15 shows the time evolution during the injection for both ammonia and iso-octane. A higher liquid fraction of fuel is found at the outlet for ammonia, resulting in a higher liquid discharge coefficient. Despite similar saturation pressure for the two fuels at the operating temperature, the saturation heat of vaporization of iso-octane is lower at the present study's temperature (Fig. 16). This can be the reason for higher vapor generation for iso-octane at the given temperature since less energy from the boundary is needed to vaporize the fuel. Furthermore, it seems based on Fig. 17 that more kinetic energy is provided to the fuel, likely due to iso-octane's lower viscosity that makes it flow faster than ammonia. This higher kinetic energy then cause higher vapor formation. This confirms that the injector's geometry and its counter bore are not the sole responsible for cavitation generation and that the fuel's property and boundaries have a strong influence.

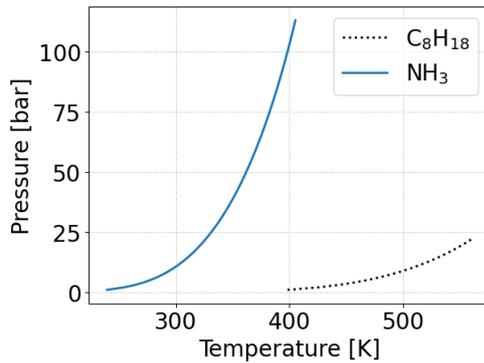


Figure 14: Liquid-vapor curve for ammonia and iso-octane: below the saturation curve, the fuel is gaseous, above it is liquid.

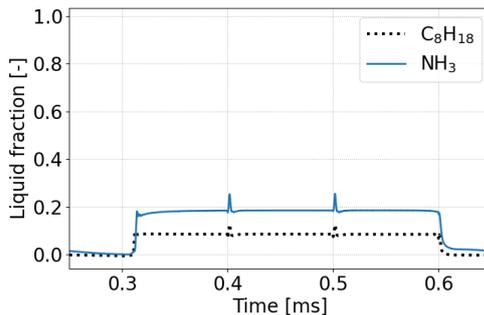


Figure 15: Time evolution of the fraction of liquid during injection at the outlet for ammonia and iso-octane.

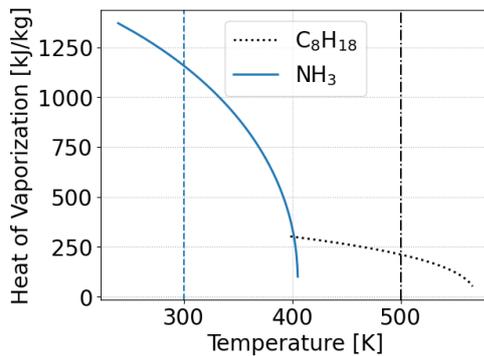


Figure 16: Heat of vaporization saturation curve for ammonia and iso-octane. The vertical lines indicate the temperature of the fuels in the present study (300 K for ammonia, 500 K for iso-octane).

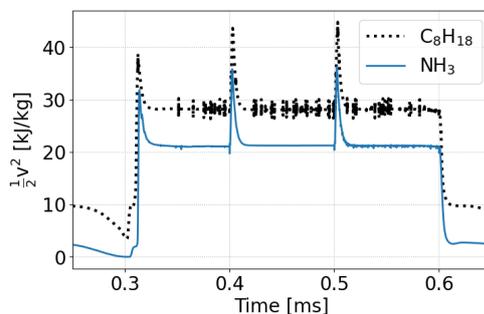


Figure 17: Time evolution of kinetic energy during injection at the outlet for ammonia and iso-octane.

3.3. Influence of the outlet condition: Pressure

Other parameters at the outlet need to be investigated

to get more realistic engine conditions. The study at 1 bar shows that there was no backflow of fuel in the injector, therefore, the specified temperature at the outlet will not be investigated further in this study. The wall temperature can have an impact as it would be heating the fuel and will be investigated in the next section. In this section, the pressure at the outlet is modified to imitate the pressure in an engine. Figure 18 shows the liquid discharge coefficient at different outlet pressures. As the outlet pressure increases, the discharge coefficient is increased. Indeed, since the pressure difference between the inlet and the outlet is smaller, the velocity decreases, the proportion of liquid at the outlet is higher and therefore the predicted cavitation is lower (Fig. 19 and Fig. 20). At an outlet pressure of 40 bar, almost none of the fuel has vaporized. The surrounding conditions and especially the pressure difference drives the velocity of the flow across the nozzle and dictates whether or not cavitation will occur. In (Saha et al., 2017), liquid fraction between 0.4 and 0.5 was found when injecting iso-octane pressurized at 200 bar in a chamber at 6 bar. In our case, at 5 bar the liquid fraction is between 0.3 and 0.4, a similar order of magnitude. The difference in value can be linked to the nature of the fuel used and the different geometries.

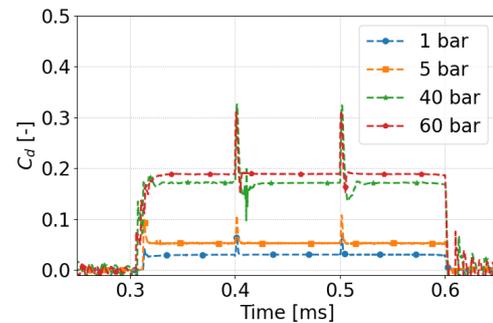


Figure 18: Time evolution of discharge coefficients during injection for different outlet pressures.

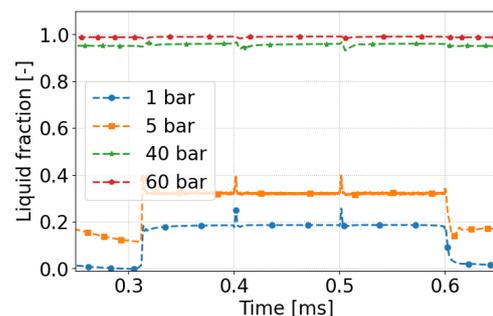


Figure 19: Time evolution of fraction of liquid at the outlet during injection for different outlet pressures.

3.4. Influence of the boundary condition: Wall temperature

To assess the impact of the heat transfer on the fluid, the wall temperature of the nozzle is varied. The outlet pressure is kept the same as the reference study at 1 bar to isolate the effect of the heat transfer. Figure 21 shows the liquid fraction of the fuel during the injection with different wall temperatures imposed: 300 K, 400

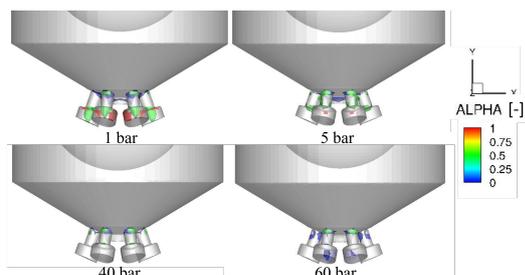


Figure 20: Contour of the void fraction during injection for different outlet pressures at 3.7 ms.

K and 600 K. As the wall temperature increases, energy is provided to the fuel as illustrated Fig. 22, the fuel temperature also increases as expected. The kinetic energy however remains the same for all cases. When the wall temperature is at 300 K, the fluid loses energy to the system, and the kinetic energy only causes vapor formation. However, for the wall temperature of 400 K and 600 K, positive values of heat shows that energy coming from the wall is provided to the fuel. This causes the fuel to further vaporize.

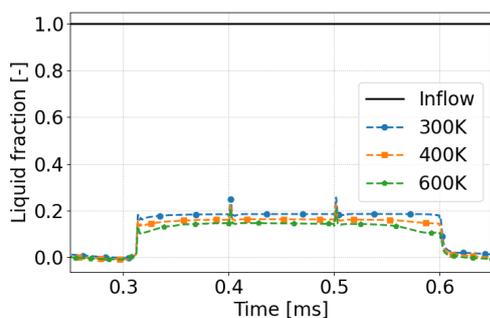


Figure 21: Time evolution of fluid liquid fraction during injection at the inlet and at the outlet for different wall temperature.

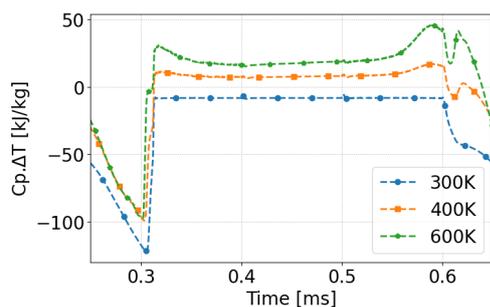


Figure 22: Time evolution of the fuel's heat during injection at the outlet for different wall temperature.

4. Conclusion

3-D RANS simulations were performed using CONVERGE CFD to study the internal flow of a 6-holes injector. Two energy sources seem to cause vapor formation: heat provided to the fuel from the boundary, and kinetic energy causing a pressure drop below the saturation pressure. This study first showed the impact of the nozzle geometry on the vapor formation

of ammonia: the sharp edges in this given geometry (based on a real commercialized injector) accelerate the flow causing a pressure drop. When the pressure drops below the saturation pressure, vaporization of the liquid fuel occurs. The presence of the counter bore inducing an expansion causes the gaseous fuel to cool down due to the Joule-Thomson effect. The amount of fuel vaporizing depends not only on the geometry, but also on the outlet conditions, and especially the pressure at the outlet. Higher pressure at the outlet will decrease the fuel vaporization and improve the discharge coefficient of the nozzle. At engine-relevant condition, the discharge coefficient is typically around 0.2. Furthermore, it is found that the wall's temperature affect the amount of fuel cavitating, which is increased with increased wall temperature.

The fuel's properties also impact the amount of cavitation. Fuels with high saturation pressure will cavitate more than fuels with lower saturation pressures. Furthermore, the fuel's viscosity impact the velocity of the flow and hence the cavitation amount.

This study shows that it is indeed possible to adapt to ammonia fuel in ICE with possibly minor changes to the injector design. More work will be needed in the future to understand the cavitation of ammonia in more detail and in different injector's geometry. Coupling of in-nozzle flow with Lagrangian spray simulations will provide further information on ammonia spray formation before computing reacting flows in full engines simulations.

Acknowledgment

The Authors gratefully acknowledge the support from the Norwegian Research Council and related partners in the Low Emission Centre. The computations were performed on resources provided by Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway as well as on the IDUN cluster at the Norwegian University of Science and Technology. Convergent Science provided CONVERGE licenses and technical support for this work.

References

- Amsden, A. (1997), 'Kiva-3v: A block-structured kiva program for engines with vertical or canted valves'. doi: 10.2172/505339. URL: <https://doi.org/10.2172/505339>
- Beattie, J. A. (1930), 'A simple equation for the joule-thomson effect in real gases', *Physical Review* **35**, 643. doi: 10.1103/PhysRev.35.643. URL: <https://journals.aps.org/pr/abstract/10.1103/PhysRev.35.643>
- Bilicki, Z. and Kestin, J. (1990), 'Physical aspects of the relaxation model in two-phase flow', *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* (1875), 379–397. doi: 10.1098/rspa.1990.0040. URL: <https://doi.org/10.1098/rspa.1990.0040>
- Cheng, Q., Ojanen, K., Diao, Y., Kaario, O. and Larmi, M. (2022), 'Dynamics of the ammonia spray using high-speed schlieren imaging', *SAE Technical Paper*. doi: 10.4271/2022-01-0053. URL: <https://doi.org/10.4271/2022-01-0053>
- Downar-Zapolski, P., Bilicki, Z., Bolle, L. and Franco, J. (1996), 'The non-equilibrium relaxation model for one-dimensional flashing liquid flow', *International Journal of Multiphase Flow* pp. 473–483. doi: 10.1016/0301-9322(95)00078-X. URL: [https://doi.org/10.1016/0301-9322\(95\)00078-X](https://doi.org/10.1016/0301-9322(95)00078-X)
- Lhuillier, C., Brequigny, P., Contino, F. and Mounaïm-Rousselle, C. (2020), 'Experimental study on ammonia/hydrogen/air combustion in spark ignition engine conditions', *Fuel*. doi:

- 10.1016/J.FUEL.2020.117448.
URL: <https://doi.org/10.1016/j.fuel.2020.117448>
- Macian, V., Bermúdez, V., Payri, R. and Gimeno, J. (2003), 'New technique for determination of internal geometry of a diesel nozzle with the use of silicone methodology', *Experimental Techniques* pp. 39–43. doi: 10.1111/j.1747-1567.2003.tb00107.x.
URL: <https://doi.org/10.1111/j.1747-1567.2003.tb00107.x>
- Mounaïm-Rousselle, C., Bréquigny, P., Dumand, C. and Houillé, S. (2021), 'Operating limits for ammonia fuel spark-ignition engine'. doi: 10.3390/en1414141.
URL: <https://doi.org/10.3390/en1414141>
- Mounaïm-Rousselle, C., Bréquigny, P., Medina, A. V., Boulet, E., Emberson, D. and Løvås, T. (2022), 'Ammonia as fuel for transportation to mitigate zero carbon impact', *Energy, Environment, and Sustainability* pp. 257–279. doi: 10.1007/978-981-16-8717-4_11.
URL: https://link.springer.com/chapter/10.1007/978-981-16-8717-4_11
- Niki, Y., Yoo, D.-H., Hirata, K. and Sekiguchi, H. (2016), 'Effects of ammonia gas mixed into intake air on combustion and emissions characteristics in diesel engine'. doi: 10.1115/ICEF2016-9364.
URL: <http://asmedigitalcollection.asme.org/ICEF/proceedings-pdf/ICEF2016/50503/V001T03A004/2443686/v001t03a004-icef2016-9364.pdf>
- Pelé, R., Mounaïm-Rousselle, C., Bréquigny, P., Hespel, C., Bellettre, J., Begg, S. and Emekwuru, N. (2021), 'First study on ammonia spray characteristics with a current gdi engine injector', *Fuels 2021, Vol. 2, Pages 253-271* pp. 253–271. doi: 10.3390/fuels2030015.
URL: <https://doi.org/10.3390/fuels2030015>
- Reiter, A. J. and Kong, S. C. (2008), 'Demonstration of compression-ignition engine combustion using ammonia in reducing greenhouse gas emissions', *Energy and Fuels* **22**, 2963–2971. doi: 10.1021/ef800140f.
URL: <https://doi.org/10.1021/ef800140f>
- Richards, K. J., Senecal, P. K. and Pomraning, E. (2017), *CONVERGE Manual v3.0*.
- Saha, K., Quan, S., Battistoni, M., Som, S., Senecal, P. K. and Pomraning, E. (2017), 'Coupled eulerian internal nozzle flow and lagrangian spray simulations for gdi systems', *SAE Technical Papers*. doi: 10.4271/2017-01-0834.
URL: <https://doi.org/10.4271/2017-01-0834>
- Serras-Pereira, J., Romunde, Z. V., Aleiferis, P. G., Richardson, D., Wallace, S. and Cracknell, R. F. (2010), 'Cavitation, primary break-up and flash boiling of gasoline, iso-octane and n-pentane with a real-size optical direct-injection nozzle', *Fuel* **89**, 2592–2607. doi: 10.1016/J.FUEL.2010.03.030.
URL: <https://doi.org/10.1016/j.fuel.2010.03.030>
- Shields, B., Neroorkar, K. and Schmidt, D. P. (2011), 'Cavitation as rapid flash boiling', *ILASS-Americas 23rd Annual Conference on Liquid Atomization and Spray Systems, Ventura, CA, May 2011*.
URL: [http://refhub.elsevier.com/S0017-9310\(21\)00036-3/sbref0038](http://refhub.elsevier.com/S0017-9310(21)00036-3/sbref0038)
- Torelli, R., Som, S., Pei, Y., Zhang, Y., Voice, A., Traver, M. and Cleary, D. (2017), 'Comparison of in-nozzle flow characteristics of naphtha and n-dodecane fuels', *SAE Technical Papers*. doi: 10.4271/2017-01-0853.
URL: <https://doi.org/10.4271/2017-01-0853>
- Wang, C., Xu, H., Herreros, J. M., Wang, J. and Cracknell, R. (2014), 'Impact of fuel and injection system on particle emissions from a gdi engine', *Applied Energy* pp. 178–191. doi: 10.1016/J.APENERGY.2014.06.012.
URL: <http://dx.doi.org/10.1016/j.apenergy.2014.06.012>
- Zhuang, Y., Qian, Y. and Hong, G. (2017), 'The effect of ethanol direct injection on knock mitigation in a gasoline port injection engine', *Fuel* pp. 187–197. doi: 10.1016/J.FUEL.2017.08.060.
URL: <https://doi.org/10.1016/j.fuel.2017.08.060>

Anaerobic Co-Digestion of Products from Biosolids Pyrolysis – Implementation in ADM1

Gudny Øyre Flatabo^{a,b,*}, Wenche Hennie Bergland^a

^a Department of Process, Energy and Environmental Technology, University of South-Eastern Norway, Kjølnes ring 56, 3918 Porsgrunn, Norway, ^b Scanship AS, Nedre Langgate 19, 3126 Tønsberg, Norway.
gudny.flatabo@scanship.no

Abstract

Pyrolyzing biosolids can decrease volume and increase value of solids while anaerobic digestion of gas and liquids from the process could increase overall methane production. Prediction of process behavior and biogas yield through simulation is valuable when considering new substrates for anaerobic digestion. In this study, gas and liquids from biosolids pyrolysis were implemented in Anaerobic Digestion Model No 1 (ADM1) together with a stream of thermally hydrolyzed sludge/food waste used in an industrial biogas plant. Average operational data from the industrial plant was used to calibrate the base scenario in ADM1, achieving a good fit. Simulation scenarios evaluating two hydrolysis constants for the pyrolysis liquid showed minor differences at the load simulated and simulated variations in composition of the liquid showed minor differences. Simulation of adding a relevant stream of pyrolysis liquid and gas together increased methane production by 7 % but decreased overall methane yield from 63 % to 61 % compared to the base scenario.

1. Introduction

Pyrolysis is a thermochemical process that converts dry biomass into gas, solids, and liquids at high temperatures without oxygen. The process can yield value added products such as biochar, pyrolysis gas, and condensable liquids. Biochar offers carbon and nutrient capture (Lehmann *et al.*, 2021) and pyrolysis gas and condensable liquids are energy carriers or chemicals precursors (Jahangiri *et al.*, 2021). Anaerobic digestion (AD) is a biochemical process where microorganisms decompose wet biomass in the absence of oxygen, to simple chemicals such as acetate, H₂, and CO₂. These simple chemicals are then converted to methane by methanogens. The resulting mix of methane and CO₂ is called biogas. The residual biomass, digestate, is normally dewatered to reduce volume and is then called biosolids. Biogas can be used to produce heat and electricity or be upgraded to biomethane for use as a fuel similar to fossil (“natural”) gas. Biosolids can normally be applied to land as a fertilizer, but if sewage sludge is the source, there are many limitations to its use making its disposal a large cost for AD plants.

Pyrolysis gas contains varying amounts of H₂, CO₂, CO, CH₄, and other C₂₊ gases, and has been successfully converted to biogas via anaerobic digestion (Luo and Angelidaki, 2012; Luo *et al.*, 2013; Li *et al.*, 2020; Torri *et al.*, 2020). AD of pyrolysis liquids from various feedstocks has also been studied (Seyedi *et al.*, 2020; Ghimire *et al.*, 2021). Successful co-digestion of liquid and gaseous pyrolysis products in an industrial process might

increase biogas production from initial feedstock, but it might also upset process behavior. Prediction through simulation is valuable when considering new AD substrates or new process conditions. The Anaerobic digestion model No. 1 (ADM1) was developed by the International Water Association to predict AD process behavior by many substrates and conditions (Batstone *et al.*, 2002). However, simulations using the standard ADM1 cannot predict process behavior for AD of unusual substrates such as pyrolysis gas and liquids and needs modification when implementing such substrates.

Currently, some works have focused on implementing mixtures of CO, H₂, and CO₂ (syngas) into ADM1, providing explanations of the model and necessary extensions (Shah *et al.*, 2017; Sun *et al.*, 2021). Simulation of combined pyrolysis liquid and gas has not been explored in the literature, and in the few experimental works on AD of real pyrolysis gas, it has been cleaned and condensed first (Giwa *et al.*, 2019; Torri *et al.*, 2020). The authors have argued that the liquid may be inhibitory and have either disposed of or digested it in a separate reactor from the gas. Possible inhibition from pyrolysis liquid from lignocellulosic materials implemented in ADM1 was presented at EUROSIMS 2021 (Raya *et al.*, 2021), while Seyedi (2020) experimented with several methods to decrease AD toxicity from biosolids pyrolysis liquid. Results from Seyedi (2020) were not implemented in any AD model, but the data showed most success in reducing toxicity from biosolids pyrolysis liquid when using acclimated inoculum

and a low loading rate (0.03 gCOD L⁻¹d⁻¹) with correspondingly long solids retention time.

Pyrolysis liquid from biosolids have different properties compared to liquid from lignocellulosic materials, such as much higher total ammonia nitrogen (TAN) content (Seyedi *et al.*, 2019). In ADM1, inhibition from NH₃ on acetate degrading organisms is modelled using a non-competitive reversible inhibition term (Eq. 1) (Batstone *et al.*, 2002).

$$I = \frac{1}{1 + S_1/K_1} \quad (1)$$

Where the total inhibition, I, is calculated using the inhibitor concentration S₁ (NH₃, M) and the inhibition parameter K₁, set to 0.0018 M as the inhibitory free ammonia concentration. The higher the inhibitor concentration S₁ goes, the lower the total inhibition value, I, gets, corresponding to more inhibition. A lower loading rate, such as suggested by Seyedi (2020) would reduce inhibitor concentration and reduce total inhibition (increase I). Simulations of co-digestion with high TAN pyrolysis liquid in an industrial process can give indications of tolerance limits for the microbial processes. Simulating at industrially relevant loading rates and co-digestion ratios gives rapid results on whether the process and loading are hypothetically viable in an industrial setting.

1.1. Goal and scope

Through this study, we aim to implement gas and liquid from biosolids pyrolysis in ADM1 and simulate anaerobic co-digestion of pyrolysis products using industrially relevant ratios of pyrolysis products to main substrate.

The study includes the following: a base scenario simulating and calibrating an industrial AD process using operational and laboratory data, an implementation of pyrolysis liquid using a combination of laboratory and literature based compositional data, an evaluation of hydrolysis constant for the implemented pyrolysis liquid, a comparison of effect of pyrolysis liquid composition on simulation results, and a combination scenario including pyrolysis liquid and gas.

We do not address additional inhibition from pyrolysis liquid products, as we hypothesize that our loading rate is sufficiently low and diluted by the main substrate, and that the high total ammonium nitrogen concentration will make inhibition by ammonia more important than other potential inhibitors.

2. Methodology

Samples of the main substrate, hydrolyzed sludge (HS), digestate, and pyrolysis products were collected and analyzed for relevant compositional data to implement in ADM1. ADM1 extended with syngas addition from Shah *et al.* (2017) was used as

a basis and extended with an extra product stream of pyrolysis liquid with separate hydrolysis constant.

2.1. Materials

Main substrate and digestate was sampled from a mesophilic sewage sludge/food waste continuously fed stirred tank (CSTR) industrial AD plant. At this plant, a thermal hydrolysis process (THP) at 160°C is used to pre-treat and sterilize the substrate. HS for analysis was sampled after the THP, prior to entering the AD bioreactors.

Pyrolysis liquid was provided by Scanship AS and produced from pyrolysis at 600°C of dried and pelletized biosolids from AD at the same plant, using the Biogreen® technology. Pyrolysis gas was condensed down to 5-8°C and the condensate, pyrolysis liquid (PL), was collected. The PL was a heterogeneous liquid: an emulsion of unevenly distributed organic phases in an aqueous phase. It was not possible to separate the liquid into an organic and an aqueous phase by gravity.

2.2. Analytical methods

Main substrate and digestate was analyzed for Chemical Oxygen Demand (total and soluble: tCOD and sCOD), total ammonium nitrogen (TAN), pH and volatile fatty acids (VFAs: acetic acid, propionic acid, butyric acid, iso-butyric acid, valeric acid, iso-valeric acid, caprioic acid, iso-caprioic acid and heptanoic acid), previously described (Bergland *et al.* 2015).

Dried samples of substrate and digestate were analysed for Total Kjeldahl Nitrogen (TKN) to estimate protein content, crude fat (EC, 2009) to estimate lipids content, and water soluble carbohydrates (Randby *et al.*, 2010) to estimate sugars.

Pyrolysis liquid was analysed for TAN, pH, VFAs, and total acid capacity (Supelco test kit no. 101758, for an indication of inorganic carbon) in addition to elemental analysis (ASTM D1552-16, 2021; ASTM D5291-21, 2021).

Pyrolysis gas was sampled in a gas bag during pyrolysis and composition was analysed using gas chromatography with two detectors: Pulsed Discharge Helium Ionisation Detector (PDHID, VIDI) and Thermal Conductivity Detector (TCD, Agilent). Helium was used as carrier gas and two different columns were used (Restek MXT Molsieve 5Å, 30 m, and Varian Poraplot Q, 25 m).

2.3. Modelling and Simulation

A full-scale AD bioreactor was simulated (base scenario) using ADM1 in AquaSim Version 2.1f. Relevant biogas production data was obtained from biogas plant operators and used as input data for the model. Known compositional data from the pyrolysis products PL and pyrolysis gas was then added, using relevant pyrolysis yields obtained from

pyrolysis operators. Simulation of PL evaluates hydrolysis rate and PL composition. Simulation of pyrolysis gas addition does not include the process of diffusion solubilizing the gas into the bioreactor volume and assumes the whole syngas volume flow is available for the microorganisms. Operational data to compare with simulation results was only available for the base scenario.

2.3.1. Base scenario

The base scenario was calibrated against average operational data (Tab. 1) from an ongoing industrial biogas plant running on thermally hydrolyzed sludge/food waste.

Table 1: Average production data from industrial AD plant, used for calibrating the base scenario. HS = hydrolyzed sludge, HRT = hydraulic retention time.

Production data	Unit	Average value
HS _{load}	m ³ d ⁻¹	183
Volume _{bioreactor}	m ³	3500
HRT	d	19.1
Production _{biogas}	Nm ³ d ⁻¹	8374
%vol CH ₄	%vol	63.0
tCOD _{digestate}	kg m ⁻³	54.0
sCOD _{digestate}	kg m ⁻³	5.30
pH _{digestate}		7.87
TAN _{digestate}	kmol N m ⁻³	0.104

Compositional results of the dry substrate for protein, lipids and sugar were adjusted to 9 % Total Solids (TS, typical value for the AD plant), while measured VFAs were input additionally, assuming these had been lost during drying while measuring TS. Complex VFAs (caproic acid, iso-caproic acid and heptanoic acid), not included in the model, were input as sugars. Amount of inerts were estimated from operational data. Lipids were estimated based on crude fat results, while proteins were estimated based on TKN results after subtracting TAN values. Carbohydrates were estimated as the difference needed to make up measured tCOD. Tab. 2 lists the inputs used for the simulation.

Table 2: Input compositional concentrations for base scenario. COD values add up to a total COD of 130 kg m⁻³. Soluble input values for amino acids, inerts, and long chain fatty acids are not listed.

Input data	Unit	Value
Lipids	kg COD m ⁻³	36.7
Carbohydrates	kg COD m ⁻³	53.4
Protein	kg COD m ⁻³	11.7
Sugar	kg COD m ⁻³	0.223
Acetic acid	kg COD m ⁻³	1.84
Propionic acid	kg COD m ⁻³	0.686
Butyric acid	kg COD m ⁻³	0.806
Valeric acid	kg COD m ⁻³	0.491
Total inert	kg COD m ⁻³	23.4
Inorganic nitrogen	kmol N m ⁻³	0.0976
Inorganic carbon	kmol HCO ₃ ⁻ m ⁻³	0.0800

Measured total COD at 9 % TS was 130 kg m⁻³, while soluble COD was 21.8 kg m⁻³. Soluble amino acids were set to 7.2 % of protein COD, as the average solubilization from two sludges treated with thermal hydrolysis in Chen *et al.* (2019). Soluble inerts were set to 8.9 % of inerts based on operational data. Soluble long chain fatty acids were used to make up measured sCOD difference. Inorganic carbon is not known but is assumed based on biogas composition operational data (63% methane, Tab. 1). Additional ions may be present, used in pH simulation, requiring adjustment of the ion balance. The base scenario has a relatively high pH (7.87, Tab. 1), so an adjustment term was added to the cations (0.35 M) to reach the target.

2.3.2. Pyrolysis liquid input

Due to the emulsified nature of the PL, only water-soluble COD could be tested with good accuracy. Because of this, the total COD of PL was represented by theoretical oxygen demand, calculated based on elemental composition (OECD, 1992).

PL was added to ADM1 as an additional substrate load, without changing volume load from HS or total volume of reactor, thus reducing the overall hydraulic retention time (HRT).

The PL load (Tab. 3) was based on the pyrolysis mass yield when pyrolyzing biosolids from the same AD process, assuming all PL goes back into the process. Fig. 1 outlines the mass load estimation for PL and pyrolysis gas. In our case, residual mass in digestate is dewatered and dried to 86 % TS, pyrolyzed, yielding 50 % mass pyrolysis gases and liquids, divided between 33 % mass liquids (4 t) and 17 % mass gases (2 t).

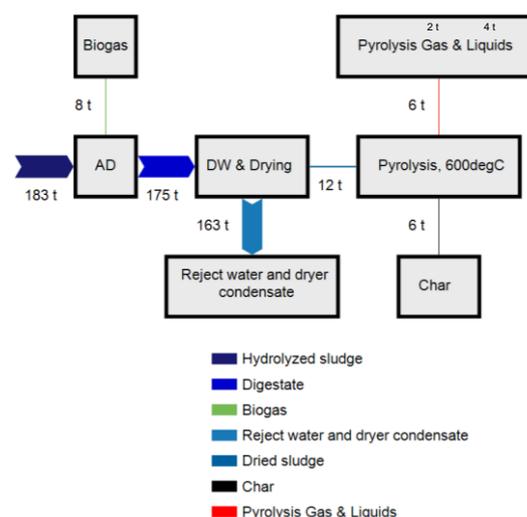


Figure 1: Simplified mass balance of industrial anaerobic digestion (AD) with pyrolysis of biosolids. DW = dewatering.

The biomethane potential (BMP) of the PL is unknown, and to simulate PL digestion the amount

of inert must be estimated. Seyedi *et al.* (2019) tested the BMP of pyrolysis liquid produced from biosolids, spiked with acetate. They found that methane production was highly affected by the organic load of PL. To estimate the concentration of inerts in our PL, a linear curve was fitted to methane % yield data estimated from the three lowest COD loadings from Seyedi *et al.* (2019) adjusted for the control (Fig. 2). The two lowest loadings gave relatively high methane yield (%), while the third lowest, 0.6 g COD L⁻¹, resulted in less methane production than the control and was adjusted to 0 when fitting a linear model. The linear model gave a methane yield of 40 % based on our calculated loading of 0.41 g COD L⁻¹ (Fig. 2, X), corresponding to 60 % inert.

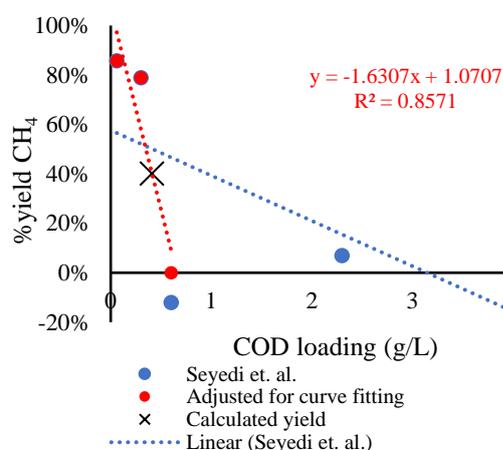


Figure 2: Methane yield of pyrolysis liquid from biosolids pyrolysis based on COD loading. Adjusted for control sample, based on data from (Seyedi *et al.*, 2019). Blue dots show all COD loadings tested by Seyedi *et al.* (2019), some have a negative yield because they produced less than the control sample. Red dots show the two loadings with a positive yield and the first negative yield adjusted to 0. The black X shows the calculated value for our data.

Compositional data was not available for PL, but analyses of the digestate gave an indication of protein and lipid content going into the pyrolysis. As a simplification and approximation of PL composition, the mass yield of compounds relevant for AD (Tab. 4) were used to adjust the protein and

lipid content in the digestate to input values for protein and lipids from PL (Wang *et al.*, 2017).

To account for the uncertainties concerning the composition of PL, an additional three compositional variations were simulated (Tab. 5). For all scenarios PL inorganic nitrogen and inorganic carbon was set to the measured TAN and total acid capacity values, 1.085 mol/L for IN and 0.700 mol/L for IC. Soluble inerts were used to add up the sCOD and are added together with particulate inerts in Tab. 5.

Table 3: Load data when adding pyrolysis liquid (PL) to the base scenario. HRT = hydraulic retention time.

Production data	Unit	Value
PL _{load}	m ³ d ⁻¹	4.08
Total load _(HS+PL)	m ³ d ⁻¹	187
Volume _{bioreactor}	m ³	3500
HRT	d	18.7
tCOD _{PL}	kg m ⁻³	352
sCOD _{PL}	kg m ⁻³	194
Load _{tCOD_PL}	kg d ⁻¹	1440
Load _{tCOD_(HS+PL)}	kg d ⁻¹	25200
HS _{loading rate}	kg COD m ⁻³ d ⁻¹	6.80
PL _{loading rate}	kg COD m ⁻³ d ⁻¹	0.410
(HS+PL) _{loading rate}	kg COD m ⁻³ d ⁻¹	7.21

Table 4: Mass yields (%) of relevant compounds for AD, from liquid from pyrolysis of proteins and lipids extracted from a microalga. From Wang *et al.* (2017).

Compounds type	Protein	Lipid
Amines and Amides (<C8)	1.94	0.43
Carboxylic acids	0.45	44.46
Ketones and Aldehydes (<C8)	4.82	0.53
Alcohols (<C8)	0	2.23
Alcohols (>C8)	0	0.36
Phenols	2.87	0
Esters	8.19	15.25
Ethers	1.92	0.12
Total	20.19	63.38

2.3.3. Hydrolysis constants

In the original ADM1, Batstone *et al.* (2002) used a complex particulates variable as a general variable for particulates in a substrate.

Table 5: Input compositional data for four scenarios with additional pyrolysis liquid. All units are concentration, kg COD m⁻³. COD values add up to a total COD of 352 kg m⁻³.

Input data	PL _{most relevant}	PL _{lipids=carbs}	PL _{low protein}	PL _{high protein}
Lipids	37.4	23.5	41.2	21.6
Carbohydrates	9.53	23.5	9.53	9.53
Protein	7.66	7.66	3.83	23.5
Sugar	18.1	18.1	18.1	18.1
Acetic acid	40.6	40.6	40.6	40.6
Propionic acid	9.45	9.45	9.45	9.45
Butyric acid	10.3	10.3	10.3	10.3
Valeric acid	8.23	8.23	8.23	8.23
Inert	111	111	111	111

A disintegration process was the rate limiting step for this variable. The variable was especially designed for complex sludge such as waste activated sludge and for microbial biomass and uses a disintegration constant of 0.5 d⁻¹ as standard. However, this input decomposes into 30 % each of lipids, carbohydrates, protein and 10 % inert. Our substrates have more accurate compositional data to input directly, making the hydrolysis process the rate limiting step. The standard hydrolysis constant, k_{hyd} , in the original ADM1 is 10 d⁻¹, which is a very high value meaning that the hydrolysis step goes very quickly and is not rate limiting.

Koch and Drewes (2014) compared two common methods of fitting biomethane potential (BMP)-curves to estimate the hydrolysis constant and derived an alternative approach where the only input necessary is the time, t , it takes for the daily methane production to drop to <1 % of accumulated value, and keep this low production for at least three days, according to the German Guideline VDI 4630 (VDI 4630, 2016). The alternative calculation of k_{hyd} based on BMP-time, t , is shown as Eq. 2 below:

$$k_{hyd} = \frac{t - 100}{t - t^2} \quad (2)$$

Raya (2021) tested a sample of the main substrate inoculated with its digestate and got a clear end of the experiment by day 10. Substituting t with days gives $k_{hyd} = 1$ d⁻¹ for the main substrate, double of the original disintegration constant. This seems like a probable value for thermally hydrolyzed sludge, which normally is digested faster than untreated sludge.

When evaluating hydrolysis constant for the pyrolysis liquid (PL), the data, however, is not clear. The BMP tests performed by Seyedi *et al.* (2019) drops to <1 % daily production after 17 days, but the production goes higher again some days later, the latest being 30 days when no daily production higher than 1% is seen. This corresponds to k_{hyd} values of 0.3 (17 days) and 0.08 (30 days). To evaluate the effect of the different k_{hyd} values, both were tested using the most relevant PL composition (Tab. 5).

2.3.4 Gas addition

Evaluating the gas effect on the processes occurring in AD is done by simplification where technical dissolution challenges are assumed solved by ignoring physical barriers and adding the gas as an extra source of COD without increasing the total liquid substrate volume stream (Shah *et al.*, 2017). This has been achieved in laboratory experiments (Corbellini *et al.*, 2021) where hydrogen gas has been injected at similar organic loading rate below surface level and with high mixing speed, achieving 90-99 % conversion efficiency, which might improve with further acclimation. Since hydrogen gas is the least soluble in water, we assume carbon monoxide can be transferred with similar efficiency

(solubility H₂ at 38°C: 0.0014 g gas/kg water, solubility CO at 38°C: 0.021 g gas/kg water, (Engineering ToolBox, 2008)).

Table 6: Load of non-condensable gases, based on composition of gas measured during pyrolysis.

Gas loaded	Unit	Value
CO	kg COD d ⁻¹	182
H ₂	kg COD d ⁻¹	434
CH ₄	kg COD d ⁻¹	600
COD _{gas loading rate}	kg COD m ⁻³ d ⁻¹	0.347
COD _{HS+PL+Gas loading rate}	kg COD m ⁻³ d ⁻¹	7.56
CO ₂	kmol d ⁻¹	32.8
CO ₂ input conc. in HS+PL vol	kmol m ⁻³ d ⁻¹	0.170

3. Results

3.1 Base scenario

The simulation results had a good fit with AD operational data and were within 1 % deviation (Tab. 7) except for total VFA.

Table 7: Measured and simulated values for the base scenario. Value units are kg COD m⁻³ for all unlabeled except for gas production (m³), Total Ammonium Nitrogen (TAN, M d⁻¹), and pH.

Target data	Value _{meas}	Value _{sim}	%Diff
Production _{biogas}	8374	8409	0.418
Production _{CH₄}	5275	5278	0.0569
% vol CH ₄	63.0	62.8	-0.317
tCOD _{digestate}	54.0	54.0	0.000
sCOD _{digestate}	5.30	5.33	0.566
pH _{digestate}	7.87	7.83	-0.508
TAN _{digestate}	0.104	0.103	-0.962
totVFA _{digestate}	0.172	3.22	1782

3.2 Variation of hydrolysis constant

The simulation results of two scenarios with PL added with two different k_{hyd} , 0.3 d⁻¹ and 0.08 d⁻¹, had very small differences, less than 0.6 % for all parameters. The main effect from the reduction of k_{hyd} was a higher tCOD concentration in the digestate (+0.4 %), a small decrease in biomass and biomass growth rates, and a minor decrease in biogas production (-0.3 %).

Due to the low difference between the simulation results using different hydrolysis constants, the higher and most likely value (0.3 d⁻¹) was used for further simulations on composition and gas addition.

3.3 PL composition

3.3.1 Lipids and carbohydrates

To see the impact on simulation from lipid and carbohydrate distribution in PL, lipids concentration was reduced, and carbohydrate concentration was increased to a scenario where these would contribute equal amounts of COD. This scenario was compared to the most relevant compositional distribution (Tab. 5). All the differences in simulation results at stable conditions were less than 1 %, except for simulated

total VFA in digestate, which was reduced by 1.15 % compared to the most relevant scenario.

3.3.2 Changing PL protein input

Higher protein content changed the simulation results to a much larger extent than the other compositional changes. The parameters most affected by the changes (Fig. 3) were: Acetic acid (9% increase), sCOD (4% increase), amino acid degrading organisms' biomass and growth rate (3% increase), TAN (2% increase) and inorganic carbon (1% decrease). The low protein content scenario decreased acetic acid and sCOD by 2 and 1 %, respectively, and otherwise showed small changes (Fig. 3).

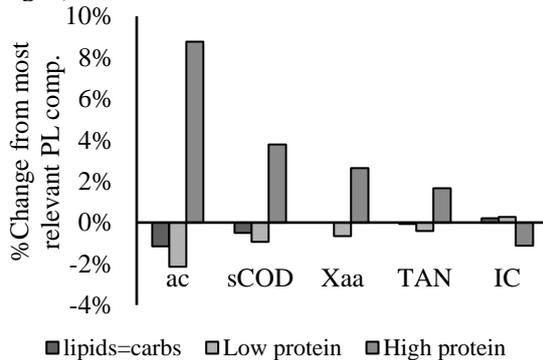


Figure 3: The parameters with the largest % change when comparing the three different composition scenarios to the most relevant composition (Tab. 5). ac = acetic acid left in digestate, sCOD = soluble chemical oxygen demand left in the digestate, Xaa = amino acid degrading organisms and their growth rate, TAN = total ammonium nitrogen left in digestate, IC = inorganic carbon in digestate.

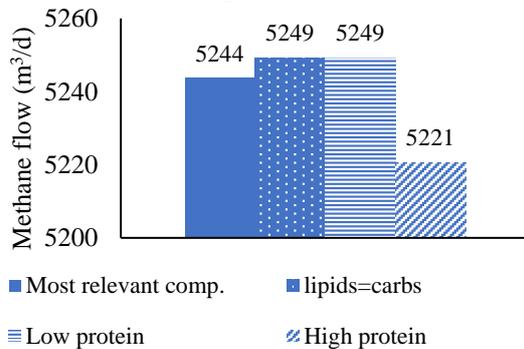


Figure 4: Daily simulated methane production when different PL input compositions are tested.

Simulated methane production only changed 0.1-0.5% by changing the protein load to the concentrations used in these scenarios (Tab. 5), but when comparing the results scaled (Fig. 4), it looks like a small trend where higher protein content affect methane production negatively. Like for the methane production, inhibition from NH_3 increases slightly (lower inhibition value) with higher protein input (Fig. 5).

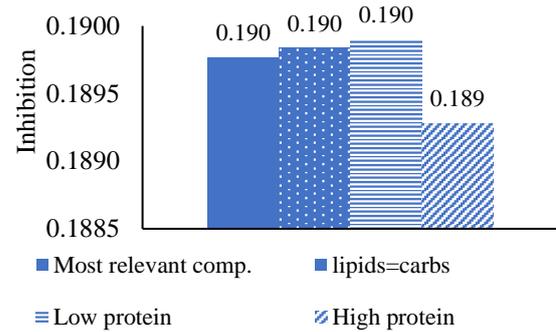


Figure 5: Inhibition values from ammonia (NH_3) on acetoclastic methanogens and the difference between simulated PL compositions. Lower inhibition value means more inhibition.

3.4 Simulation of gas addition

Gas addition increased the simulated methane production with 7 %, $375 \text{ Nm}^3 \text{ d}^{-1}$ more than the simulated base scenario (Fig. 6 and 9). The pyrolysis gas contained 210 Nm^3 of CH_4 , which accounts for about 56% of the increase in methane production. Inhibition from ammonia that increased (lower inhibition value) with PL addition, normalized somewhat when gas was added (from 0.19 with PL to 0.196 with PL and gas, Fig. 7).

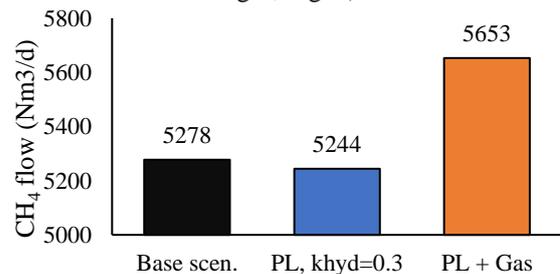


Figure 6: Daily simulated methane production for base scenario, added pyrolysis liquid and liquid and gas added together.

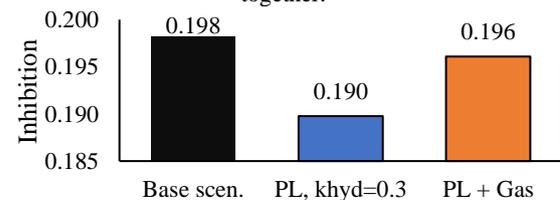


Figure 7: Inhibition values from NH_3 on acetoclastic methanogens for base scenario, added pyrolysis liquid and liquid and gas added together. Lower inhibition value means more inhibition.

The largest change from the base scenario was the acetic acid concentration in digestate that nearly tripled (from $1.6 \text{ kg COD m}^{-3}$ in base scenario to $4.5 \text{ kg COD m}^{-3}$, Fig. 8) when adding PL alone. The largest change for the gas scenario was the residual sCOD in the digestate, which increased by 55 % from the base scenario (Fig. 8). Biogas also had a significant increase of 14 % for the PL+Gas scenario (Fig. 9). TAN increased by about 20 % for both the scenarios (Fig. 8).

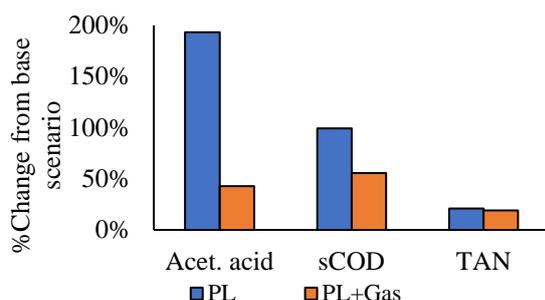


Figure 8: Percentage change from base scenario on acetic acid, soluble COD (sCOD) and TAN (total ammonium nitrogen) when adding pyrolysis liquid (PL) or pyrolysis liquid and gas (PL+Gas).

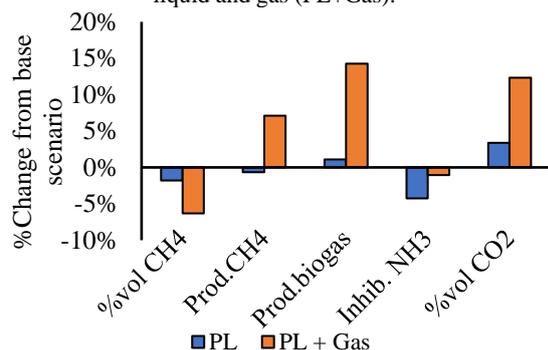


Figure 9: Percentage change from base scenario on methane concentration (% vol CH₄), produced methane, produced biogas, ammonia inhibition (Inhib. NH₃) and CO₂ concentration (% vol CO₂) when adding pyrolysis liquid (PL) or pyrolysis liquid and gas (PL+Gas).

4. Discussions and Summary

4.1 Base scenario

The simulated base scenario was well calibrated to match operational and measured data for the industrial process, except for VFAs (Tab. 8), these are previously reported with low fit for thermally hydrolyzed sludge in ADM1 (Donoso-Bravo *et al.*, 2020).

4.2 Variations of PL hydrolysis and composition

Despite reducing the hydrolysis constant for the added PL to a fourth of the most relevant value, from 0.3 d⁻¹ to 0.08 d⁻¹, the differences in the simulation results after stabilization were small at the simulated PL load. Reductions in biomass growth rates and increase in residual tCOD (section 3.3) are expected with slower solubilization of particulates caused by reducing the hydrolysis constant without changing the HRT. The % difference in biogas production, however, was less than the % difference between the simulated base scenario and measured operational data, so they would most likely not be noticed in an experimental or industrial setting.

Changing the simulated protein input had more effect than changing lipid and carbohydrate distribution. More protein leads to higher TAN (0.127 M for the high protein scenario, up from 0.125 M for the most relevant scenario), which

increases inhibition by ammonia of acetoclastic methanogenesis (Fig. 5) causing a higher concentration of acetic acid in the digestate (Fig. 3, 8 % increase in this scenario) and lower overall methane production (Fig. 4). In our scenario, the process is still stable as the pH has not decreased by much, but it gives a clear indication of what happens with too much loaded protein.

Most likely, however, the PL does not contain “traditional” lipids, carbohydrates and protein, as these are decomposed and changed into other compound groups through pyrolysis (e.g. Tab. 4 and Wang *et al.* (2017)). PL from biosolids does contain a lot of nitrogen, however. From our total nitrogen analysis there is still 1.4 M nitrogen bound to unknown compounds, after correcting for ammonium-N. Whether these compounds are anaerobically digested with an effect like proteins, whether they stay inert or whether they are inhibitory to the process is difficult to predict without conducting carefully planned, long-term continuous experiments.

4.3 Effects on industrial process from pyrolysis liquid and gas

The simulated methane production drops slightly when adding PL alone but increases 7 % when gas is added. From the overview figures (Fig. 8 and 9) we see that acetic acid concentration triples when PL is added, and TAN concentration increases, causing more inhibition from ammonia on acetoclastic methanogens. This might explain why the methane production drops slightly despite the PL contributing extra COD to the process (0.4 kg COD m⁻³ d⁻¹, Tab. 3).

The overall methane yield drops compared to the base scenario when adding these pyrolysis products (Fig. 10), most likely due to the high load of inorganic nitrogen in the liquid fraction.

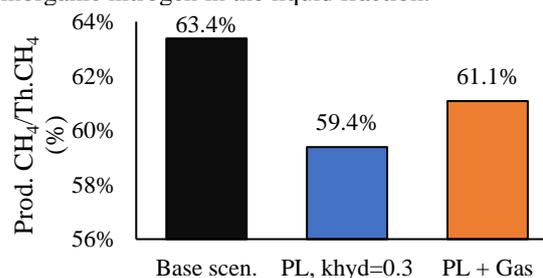


Figure 10: %Methane yields for the overall processes.

In combination, the slightly inhibitory effect from the PL is “diluted” by the gas, so the actual methane production still increases overall, however with less overall efficiency compared to the extra added COD. Despite the possible problems with PL in AD, the result from the combined addition indicates that the processes occurring in AD may tolerate an uncondensed pyrolysis gas co-substrate and increase methane production moderately.

Acknowledgment

The authors would like to thank and acknowledge master student Nasim Mohajeri Nav who has performed and participated extensively in the lab work for this study. The study was funded by the Norwegian Research Council and Scanship AS through the Industrial PhD-program.

References

- ASTM D1552-16 (2021) 'Standard Test Method for Sulfur in Petroleum Products by High Temperature Combustion and Infrared (IR) Detection or Thermal Conductivity Detection (TCD)', in Subcommittee 03 (ed.) *Book of Standards*. ASTM International, pp. 1–5. Available at: <https://www.astm.org/d1552-16r21.html> (Accessed: 23 May 2022).
- ASTM D5291-21 (2021) 'Standard Test Methods for Instrumental Determination of Carbon, Hydrogen, and Nitrogen in Petroleum Products and Lubricants', in Subcommittee 03 (ed.) *Book of Standards*. ASTM International, pp. 1–8. doi: 10.1520/D5291-21.
- Batstone, D. J. *et al.* (2002) 'The IWA Anaerobic Digestion Model No 1 (ADM1)', *Water Science and Technology*. IWA Publishing, 45(10), pp. 65–73. doi: 10.2166/WST.2002.0292.
- Bergland, W. H. *et al.* (2015) 'High rate manure supernatant digestion', *Water Research*. Elsevier Ltd, 76, pp. 1–9. doi: 10.1016/j.watres.2015.02.051.
- Chen, S. *et al.* (2019) 'Effects of thermal hydrolysis on the metabolism of amino acids in sewage sludge in anaerobic digestion', *Waste Management*. Pergamon, 88, pp. 309–318. doi: 10.1016/J.WASMAN.2019.03.060.
- Corbellini, V. *et al.* (2021) 'Performance Analysis and Microbial Community Evolution of in Situ Biological Biogas Upgrading with Increasing H₂/CO₂Ratio', *Archaea*. Hindawi Limited, 2021, pp. 1–15. doi: 10.1155/2021/8894455.
- Donoso-Bravo, A. *et al.* (2020) 'Exploitation of the ADM1 in a XXI century wastewater resource recovery facility (WRRF): The case of codigestion and thermal hydrolysis', *Water Research*. Pergamon, 175, p. 115654. doi: 10.1016/J.WATRES.2020.115654.
- EC, T. C. of the E. C. (2009) 'COMMISSION REGULATION (EC) No 152/2009 of 27 January 2009 laying down the methods of sampling and analysis for the official control of feed', *Official Journal of the European Union*, 54, p. 37.
- Engineering ToolBox (2008) *Solubility of Gases in Water vs. Temperature*. Available at: https://www.engineeringtoolbox.com/gases-solubility-water-d_1148.html (Accessed: 26 August 2022).
- Ghimire, N., Bakke, R. and Bergland, W. H. (2021) 'Liquefaction of lignocellulosic biomass for methane production: A review', *Bioresource Technology*. Elsevier, 332, p. 125068. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0960852421004077> (Accessed: 8 April 2021).
- Giwa, A. S. *et al.* (2019) 'Pyrolysis of difficult biodegradable fractions and the real syngas bio-methanation performance', *Journal of Cleaner Production*. Elsevier Ltd, 233, pp. 711–719. doi: 10.1016/j.jclepro.2019.06.145.
- Jahangiri, H. *et al.* (2021) 'Thermochemical Conversion of Biomass and Upgrading of Bio-Products to Produce Fuels and Chemicals', *Catalysis for Clean Energy and Environmental Sustainability*. Springer, Cham, pp. 1–47. doi: 10.1007/978-3-030-65017-9_1.
- Koch, K. and Drewes, J. E. (2014) 'Alternative approach to estimate the hydrolysis rate constant of particulate material from batch data', *Applied Energy*. Elsevier, 120, pp. 11–15. doi: 10.1016/J.APENERGY.2014.01.050.
- Lehmann, J. *et al.* (2021) 'Biochar in climate change mitigation', *Nature Geoscience*. Springer US, 14(12), pp. 883–892. doi: 10.1038/s41561-021-00852-8.
- Li, C., Zhu, X. and Angelidaki, I. (2020) 'Carbon monoxide conversion and syngas biomethanation mediated by different microbial consortia', *Bioresource Technology*. Elsevier, 314(May), p. 123739. doi: 10.1016/j.biortech.2020.123739.
- Luo, G. and Angelidaki, I. (2012) 'Integrated biogas upgrading and hydrogen utilization in an anaerobic reactor containing enriched hydrogenotrophic methanogenic culture', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 109(11), pp. 2729–2736. doi: 10.1002/bit.24557.
- Luo, G., Wang, W. and Angelidaki, I. (2013) 'Anaerobic digestion for simultaneous sewage sludge treatment and CO biomethanation: Process performance and microbial ecology', *Environmental Science and Technology*. American Chemical Society, 47(18), pp. 10685–10693. doi: 10.1021/es401018d.
- OECD (1992) *Annex IV - Calculation and Determination of Suitable Summary Parameters, OECD Guideline for Testing of Chemicals*. OECD. Available at: <https://www.oecd.org/chemicalsafety/risk-assessment/1948209.pdf> (Accessed: 3 January 2022).
- Randby, Å. T., Nørgaard, P. and Weisbjerg, M. R. (2010) 'Effect of increasing plant maturity in timothy-dominated grass silage on the performance of growing/finishing Norwegian Red bulls', *Grass and Forage Science*, 65, pp. 273–286. doi: 10.1111/j.1365-2494.2010.00745.x.
- Raya, D. *et al.* (2021) 'Anaerobic Digestion of Aqueous Pyrolysis Liquid in ADM1', *Linköping Electronic Conference Proceedings*, 185, pp. 458–464. doi: 10.3384/ecp21185458.
- Raya, D. (2021) *Analysing Aqueous Pyrolysis Liquid as feed for Anaerobic Digestion*. University of South-Eastern Norway. Available at: <https://openarchive.usn.no/usn-xmlui/handle/11250/2766578> (Accessed: 30 May 2022).
- Seyedi, S. (2020) *Overcoming Anaerobic Digestion Toxicity of Aqueous Liquid from Wastewater Solids Pyrolysis*. Marquette University. Available at: https://epublications.marquette.edu/dissertations_mu/1025 (Accessed: 3 February 2022).
- Seyedi, S., Venkiteswaran, K. and Zitomer, D. (2019) 'Toxicity of various pyrolysis liquids from biosolids on methane production yield', *Frontiers in Energy Research*. Frontiers Media S.A., 7(FEB), pp. 1–12. doi: 10.3389/fenrg.2019.00005.
- Seyedi, S., Venkiteswaran, K. and Zitomer, D. (2020) 'Current status of biomethane production using aqueous liquid from pyrolysis and hydrothermal liquefaction of sewage sludge and similar biomass', *Reviews in Environmental Science and Biotechnology*. Springer Science and Business Media B.V., pp. 237–255. doi: 10.1007/s11157-020-09560-y.
- Shah, S., Bergland, W. H. and Bakke, R. (2017) 'Methane from Syngas by Anaerobic Digestion', *Proceedings of the 58th Conference on Simulation and Modelling (SIMS 58) Reykjavik, Iceland, September 25th – 27th, 2017*. Linköping University Electronic Press, 138, pp. 114–120. doi: 10.3384/ECP17138114.
- Sun, H. *et al.* (2021) 'Modification and extension of an anaerobic digestion model No.1 (ADM1) for syngas biomethanation simulation: From lab-scale to pilot-scale', *Chemical Engineering Journal*. Elsevier B.V., 403, p. 126177. doi: 10.1016/j.cej.2020.126177.
- Torri, C. *et al.* (2020) 'Evaluation of the potential performance of hyphenated pyrolysis-anaerobic digestion (Py-AD) process for carbon negative fuels from woody biomass', *Renewable Energy*. Elsevier Ltd, 148, pp. 1190–1199. doi: 10.1016/j.renene.2019.10.025.
- VDI 4630 (2016) *VDI 4630 - Fermentation of organic materials - Characterization of the substrate, sampling, collection of material data, fermentation tests*, Verein Deutscher Ingenieure e.V. Available at: <https://www.vdi.de/richtlinien/details/vdi-4630-fermentation-of-organic-materials-characterization-of-the-substrate-sampling-collection-of-material-data-fermentation-tests> (Accessed: 30 May 2022).
- Wang, X., Sheng, L. and Yang, X. (2017) 'Pyrolysis characteristics and pathways of protein, lipid and carbohydrate isolated from microalgae *Nannochloropsis* sp.', *Bioresource Technology*. Elsevier, 229, pp. 119–125. doi: 10.1016/J.BIORTECH.2017.01.018.

Method for mean-line design and performance prediction of one-stage axial turbines

Lasse B. Anderson ^{a,*}, Roberto Agromayor ^a, Lars O. Nord ^a

^a NTNU - Norwegian University of Science and Technology, Department of Energy and Process Engineering, Trondheim, Norway

* lasse.b.anderson@ntnu.no

Abstract

Today, expanders for Organic Rankine Cycles (ORC) are either inefficient or expensive. One reason for this is that expanders of power conversion systems usually operate under different conditions over an annual perspective. Still, they are designed to perform best at a single operating point. In view of this limitation, it is suggested that the overall expander performance can be improved by taking into account off-design operation in the design process. As the first step in this direction, this paper presents a two-fold method for design optimization and performance analysis of one-stage axial turbines. The method utilizes the same mean-line model for performance analysis and design optimization, ensuring consistency between the two modes. In addition, the proposed method evaluates the turbine performance at three stations for each blade row: inlet, throat and exit, and employs a novel numerical treatment of flow choking that automatically determines which blade rows are choked as part of the solution. Furthermore, the method was validated against cold-air experimental data from three different one-stage axial turbines, at both on- and off-design conditions. The model predicts design point efficiencies between 1.1 and 4.5 percentage points off the experimental values. The model was also able to capture the trend of mass flow rate as a function of total-to-static pressure ratio and angular speed. However, an unphysical behavior was observed as the pressure ratio approaches the critical value, and further developments of the model are required. It is envisioned that the proposed method will serve as foundation for a robust design methodology that will enable higher expander performance over a range of operating conditions.

1. Introduction

The Rankine Cycle technology is used to convert heat from external sources into electrical power, making it suitable to exploit low-carbon energy sources such as geothermal reservoirs, concentrated solar radiation and waste heat from industrial processes (Colonna et al. 2015). The Rankine cycle in its simplest form consist of four elements: condenser, pump, evaporator and expander (Figure 1). Electrical power is generated as the working fluid flows through the expander, and produces a torque on a shaft coupled to an electrical generator. After the expander, the working fluid is condensed, pressurized through the pump and heated before it enters the expander once again. When the heat source temperature is limited, it becomes beneficial to utilize organic compound as working fluid, to avoid condensation within the expander, and the cycle is referred to as Organic Rankine Cycle (ORC).

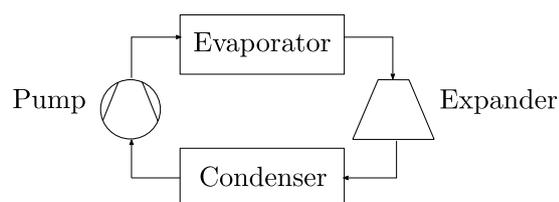


Figure 1: Simple ORC configuration.

Furthermore, adopting an organic fluid in the Rankine cycle paves the way towards compact and cost-efficient turbo-expanders. The higher molecular weight of these fluids leads to a lower enthalpy drop compared to conventional steam cycles (Macchi and Astolfi 2017). This enables turbo-expanders with a low number of stages, and operation at relatively low rotational speed. However, the expansion ratio for each stage is usually very high in compact ORC turbines, and combined with a low speed of sound this often leads to transonic or even supersonic flows. In addition, non-ideal gas effects could be prominent as the expansion process usually occurs in the dense vapor region. These effects combined result in additional losses and non-conventional turbine designs, where traditional design methods are insufficient (Persico and Pini 2017).

The first step of the design process for ORC turbines is to decide which type of architecture to use. This could be axial turbine, radial turbine or a hybrid solution. This choice is usually based on statistical data or overall machine correlations (Macchi and Astolfi 2017). The next step is the preliminary design phase. The purpose of this phase is to subdivide the the expansion process between turbine stages, thus specifying the thermodynamic and kinetic state at inlet and exit for each cascade. For this reason, the preliminary design is crucial for the whole design process. Furthermore, this phase makes use of mean-line models, which assume that the flow is uniform in the circumferential and spanwise directions (Denton 1993). Inlet and exit conditions for each cascade are determined

by solving the mass, momentum and energy conservation equations along with an equation of state. In addition, the mean line approach requires empirical correlations to estimate entropy generation throughout the turbine, referred to as loss models. The turbine performance is predicted with the mean-line model, given a geometry and a set of thermodynamic boundary conditions for the turbine. Thus, the preliminary design can be accomplished combining the mean-line model with an optimization routine that finds the best geometry that optimizes an objective function, commonly being the isentropic efficiency (Persico and Pini 2017).

In most cases, the turbine design process only takes into account the performance at one operation point (i.e., the design point). However, ORC power systems are frequently connected to heat sources and sinks with variable load. Heating and cooling duties from sources such as exhaust from gas turbines or cooling by ambient air are very likely to vary, and this will propagate to change expander conditions (Jiménez-Arreola et al. 2018; Macchi and Astolfi 2017). In order to improve efficiency of ORC expanders it is therefore suggested that the design process should account for variable expander conditions. In other words, rather than designing an expander for one single design point, the expander is designed considering a range of operating conditions. This multipoint optimization strategy has been adopted in several high-fidelity design methods, and shows great promise to rise the efficiency of the turbine at off-design conditions (Aïssa et al. 2019; Bonaiuti and Zangeneh 2009; Châtel, Verstraete, and Coussment 2020; M. Pini, G. Persico, and Dossena 2014; Sanchez Torreguitart, Verstraete, and Mueller 2018). However, to the authors' knowledge, the use of multipoint optimization strategies for preliminary turbine design has not been considered yet.

The objective in this paper is to formulate a two-fold method for the design optimization and performance analysis of one-stage axial turbines. The method uses the same mathematical model for performance analysis and design optimization, and both modes are formulated as nonlinear programming problems, which are solved using gradient-based optimization algorithms. The mathematical model follows a mean-line approach and, in contrast to most of the methods available in the open literature, evaluates the turbine performance at three stations for each blade row: inlet, throat and exit, making it suitable for choked flow conditions. The model is validated against numerical cases from the literature and experimental data, at both on- and off-design conditions. It is envisaged that the methods presented herein will serve as building blocks to develop a robust design methodology for axial turbines that will enable higher performance over a range of operating conditions.

The outline is as follows: the next section describes the mathematical model used for the two-fold method, followed by a section on the differences of the two modes. The model validation is then given before a design case study is presented. At last, the content is summarized before suggestions for further work are made.

2. Mathematical model

This section describes a mathematical model for assessment of one-stage axial turbines. This model is used for both performance analysis and preliminary design of one-stage axial turbines. The variables used to define the turbine geometry are presented first, followed by modelling

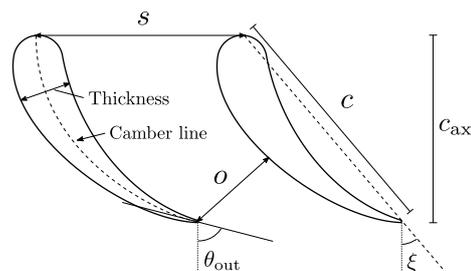


Figure 2: Geometry of the cross-sectional shape of the blades.

of velocity triangles, thermodynamic properties and loss prediction. The mathematical model presented here is based on the work of Agromayor and Nord, who proposed a method for the design optimization of axial turbines (Agromayor and Nord 2019). However, the objective here is to extend the method and include performance analysis in addition to design optimization, leading to certain model differences. Notably, the mass flow rate is computed, rather than being an input variable, because the performance analysis should predict the mass flow rate. Furthermore, thermophysical properties and velocity triangles are assessed at the throat, in addition to the inlet and exit of each cascade, to predict choked flow. Consequently, the set of input variables will deviate from the work of Agromayor and Nord.

The model is implemented in the Python programming language and is released under an open source license. This does not only increase the transparency and reproducibility of the results, but also enables other researchers and industry practitioners to use the code for their needs or contribute to its development.

2.1. Turbine geometry

The geometrical variables considered in the mean-line turbine model include parameters related to the cross-sectional shape of the blades, distance between adjacent blades and extent in radial direction. Figure 2 illustrates the geometrical variables related to the cross sectional shape of the blades. The chord c is the distance between leading and trailing edge, while the camberline is the line halfway between pressure and suction surfaces of the blade. Axial chord c_{ax} is the axial component of the chord. The blade thickness is the distance between pressure and suction surfaces, perpendicular to the camber. Furthermore, the metal angle θ is defined as the angle between axial direction and tangential component of the camber line. The pitch s is the distance between two adjacent blades, and the opening o is the distance between the trailing edge to the suction surface of the adjacent blade, measured normal to the outlet metal angle. The stagger ξ is the angle between axial direction and chord line. Maximum thickness t_{max} and trailing edge thickness t_{te} must also be considered for loss prediction (Kacker and Okapuu 1982).

In addition, the parameters describing the shape of the blade in radial direction are also included in the model, such as mean radius r_m , blade height H and tip-clearance gap t_{cl} (see Figure 3). The blade height may vary from inlet to outlet of the blade cascade, and the flaring angle δ_H describes this variation.

In order to reduce the number of input variables, some relations between geometrical variables are used. The first

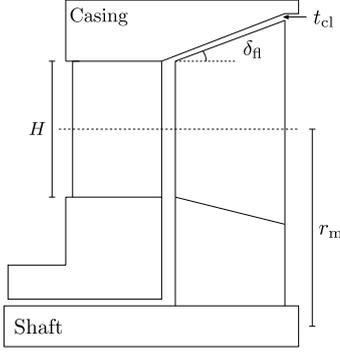


Figure 3: Geometry of turbine blades in radial direction.

relation is the cosine rule, which relates the blade opening with the exit metal angle (Saravanamutto et al. 2017):

$$o \approx s \cdot \cos(\theta_{out}) \quad (1)$$

Furthermore, Kacker and Okapuu proposed a formula to approximate the maximum thickness to chord ratio as a function of the blade camber angle ($\Delta\theta = |\theta_{in} - \theta_{out}|$) (Kacker and Okapuu 1982):

$$\frac{t_{max}}{c} = \begin{cases} 0.15, & \Delta\theta \leq 40^\circ \\ 0.15 + 1.25 \cdot 10^{-3} \cdot (\Delta\theta - 40) & \\ 0.25, & 120^\circ \leq \Delta\theta \end{cases} \quad (2)$$

If one assumes that the camber line is a circular arc, the stagger angle is given by (Dixon 2014):

$$\xi = \frac{1}{2}(\theta_{in} + \theta_{out}) \quad (3)$$

Using the definition of stagger angle, the axial chord can be determined from the chord:

$$c_{ax} = c \cdot \cos(\xi) \quad (4)$$

The mean radius is assumed to be constant, meaning the blade converge/diverge equally at the inner and outer wall of the flow passage, then the stator inlet and exit blade heights are determined by the mean blade height and stator flaring angle:

$$H_{in,s} = H_s - \tan(\delta_{fl,s}) \cdot c_{ax,s} \quad (5)$$

$$H_{out,s} = H_s + \tan(\delta_{fl,s}) \cdot c_{ax,s} \quad (6)$$

Furthermore, the blade height at stator exit is assumed to be equal to rotor inlet, meaning no flaring between blade rows. Consequently, for rotor row, the flaring angle is determined from inlet and mean blade height, which further gives the exit blade height:

$$\delta_{fl,r} = \arctan\left(\frac{H_r - H_{out,s}}{c_{ax,r}}\right) \quad (7)$$

$$H_{out,r} = H_r + \tan(\delta_{fl,r}) \cdot c_{ax,r} \quad (8)$$

Another parameter needed for loss calculations is the hub-to-tip ratio:

$$r_{ht} = \frac{r_h}{r_t} = \frac{r_m - H/2}{r_m + H/2} \quad (9)$$

At last, the area at inlet, exit and throat of each blade row is calculated by the following equations:

$$A_{in} = 2\pi \cdot r_m \cdot H_{in} \quad (10)$$

$$A_{out} = 2\pi \cdot r_m \cdot H_{out} \quad (11)$$

$$A_{throat} = 2\pi \cdot r_m \cdot H_{throat} \cdot \cos(\theta_{out}) \quad (12)$$

Table 1: Geometrical variables of the axial turbine.

Variable	Stator	Rotor
Mean radius	r_m	
Blade height	H_s	H_r
Aspect ratio	$(H/c)_s$	$(H/c)_r$
Pitch to chord ratio	$(s/c)_s$	$(s/c)_r$
TE thickness to opening ratio	$(t_{te}/o)_s$	$(t_{te}/o)_r$
Inlet metal angle	-	$\theta_{in,r}$
Outlet metal angle	$\theta_{out,s}$	$\theta_{out,r}$
Flaring angle	$\delta_{fl,s}$	-
Tip clearance	-	t_{cl}

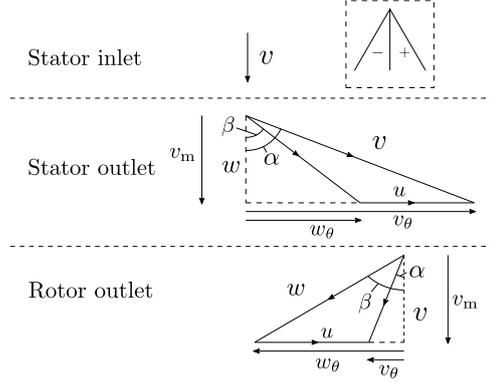


Figure 4: Velocity triangles. Sign convention for the angles is illustrated in the upper right corner.

In total, the geometry of the stage is defined using fourteen variables. The geometrical variables are listed in Table 1 and are inputs to the model.

2.2. Velocity triangles

The velocity triangles (Figure 4) are evaluated at each station of the turbine, and are calculated with the following equations:

$$\begin{aligned} v_m &= v \cdot \cos(\alpha) \\ v_\theta &= v \cdot \sin(\alpha) \\ w_m &= w \cdot \cos(\beta) \\ w_\theta &= w \cdot \sin(\beta) \\ w_\theta &= v_\theta - u \end{aligned} \quad (13)$$

Where the blade velocity u is calculated from the mean radius and the rotational speed Ω :

$$u = \Omega \cdot r_m \quad (14)$$

The stator inlet flow angle is fixed to zero, while the flow angles at throats are assumed to equal the exit metal angles, which imply zero deviation:

$$\beta_{throat} = \theta_{out} + \delta \approx \theta_{out} \quad (15)$$

Thus, the velocity at each station, exit flow angles and rotational speed are input variables. These are listed in Table 2, except for the rotational speed, which is treated similarly as the geometrical variables (see Section 3).

Table 2: Velocity, flow angles and static pressure variables that are inputs for the model.

Variable	Symbol
Stator inlet velocity	$v_{in,s}$
Stator throat velocity	$v_{throat,s}$
Stator exit velocity	$v_{out,s}$
Rotor throat relative velocity	$w_{throat,r}$
Rotor exit relative velocity	$w_{out,r}$
Stator exit absolute flow angle	$\alpha_{out,s}$
Rotor exit relative flow angle	$\beta_{out,r}$
Stator throat static pressure	$p_{throat,s}$
Stator exit static pressure	$p_{out,s}$
Rotor throat static pressure	$p_{throat,r}$

Table 3: Thermodynamic boundaries for the turbine.

Variable	Symbol
Working fluid	-
Inlet stagnation temperature	$T_{0,in}$
Inlet stagnation pressure	$p_{0,in}$
Outlet static pressure	p_{out}

2.3. Thermodynamics

The thermodynamic boundaries for the turbine model must also be specified. These variables are usually obtained from a system-level analysis of the power cycle and include states at inlet and outlet of the turbine, as well as the working fluid. These variables are listed in Table 3.

In addition to these input variables, other thermophysical variables need to be evaluated through the turbine stage. Similarly as for the velocity triangles, these are assessed at each station of the turbine. For simple compressible systems, the state is completely specified for two independent intensive properties (Moran et al. 2015). A library of thermophysical properties is required for these calculations, and in this method CoolProp is used (Bell et al. 2014). Since the stagnation state is prescribed at the turbine inlet, entropy and enthalpy are used to calculate thermophysical properties at this station:

$$[\rho, a, p, \mu]_{in} = f(s_{in}, h_{in}) \quad (16)$$

where ρ is the fluid density, a speed of sound, and μ is the dynamic viscosity. Through the rest of the turbine, static pressure and enthalpy are used to determine the dependent thermophysical properties. This means static pressure at stator throat, between blade rows and rotor throat must be provided as input (see Table 2). This was found to be more robust than using entropy and enthalpy as input variables.

$$[\rho, a, s, \mu] = f(p, h) \quad (17)$$

The enthalpy at each station of the turbine is calculated by the conservation of rothalpy:

$$h + \frac{1}{2}w^2 - \frac{1}{2}u^2 = \text{constant} \quad (18)$$

Through the stator row this is reduced to conservation of stagnation enthalpy:

$$h_{0,s} = \text{constant} \quad (19)$$

With the velocity triangles and thermophysical properties, the mass flow rate, Mach and Reynolds number may be evaluated:

$$\dot{m} = \rho v A \quad (20)$$

$$\text{Ma} = \frac{v}{a} \quad (21)$$

$$\text{Re} = \frac{\rho V c}{\mu} \quad (22)$$

2.4. Loss model

What is left is to predict the entropy generation in the turbine. For this purpose, a loss coefficient must be defined, and a loss model to predict this loss coefficient must be selected. In this work, the Kacker and Okapuu loss model is adopted for its maturity and accuracy (Kacker and Okapuu 1982), but the loss model may easily be switched to another if desired. The loss coefficient may be defined in several ways (Denton 1993), but is here formulated as the total pressure loss between inlet and exit of a blade row, divided by the exit dynamic pressure:

$$Y_{\text{definition}} = \frac{p_{0,\text{rel},in} - p_{0,\text{rel},out}}{p_{0,\text{rel},out} - p_{out}} \quad (23)$$

This is referred to as the total pressure loss coefficient, and it is adopted because the Kacker and Okapuu loss model is based on this definition. The deviation between the total pressure loss coefficient computed from its definition and the value determined by the loss model must be assessed:

$$Y_{\text{error}} = Y_{\text{definition}} - Y_{\text{lossmodel}} \quad (24)$$

This error should be zero, and the states through the turbine must be such that this is satisfied. This is discussed further in Section 2.5 and 3. Furthermore, the loss coefficient is here assessed at both throat and exit for each blade row. The loss model itself is formulated to predict losses between inlet and exit of a blade row and not between inlet and throat. However, the alternative is to assume isentropic flow between inlet and throat, but it was observed that this strategy predicted losses less accurately compared to the original formulation. A logical explanation behind this is that the state at the throat is more likely to be similar to the exit rather than the inlet.

2.5. Closing of equations

With the use of the equations above, the mass flow rate, loss coefficient and Mach number can be evaluated at each station using the variables in Tables 2 and 3. However, by only using the equations above, the mass flow rates at each station are not necessarily equal, the Mach number at the throat is not necessarily maximum one and the error between the loss model coefficient and loss coefficient obtained from its definition is not necessarily zero. This forms a set of equations that must be satisfied for each blade row in order for the flow to be physically feasible:

$$\begin{aligned} \dot{m}_{in} - \dot{m}_{throat} &= 0 \\ \dot{m}_{in} - \dot{m}_{out} &= 0 \\ \text{Ma}_{throat,rel} - \min(1, \text{Ma}_{out,rel}) &= 0 \\ (Y_{\text{def}} - Y_{\text{loss}})_{throat} &= 0 \\ (Y_{\text{def}} - Y_{\text{loss}})_{out} &= 0 \end{aligned} \quad (25)$$

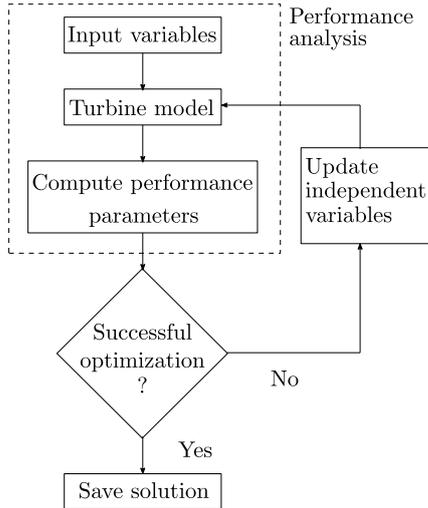


Figure 5: Workflow for the preliminary design objective and performance analysis.

In other words, this is a set of ten equations on the form $\vec{F}(\vec{x}) = 0$, where \vec{x} is a vector with the ten variables in Table 2. The thermophysical states and velocity triangles can be determined by ensuring that Eq. (25) is satisfied. The strategy to ensure this is different for the two modes of the two-fold method, and are presented in the next section.

3. Optimization vs. performance analysis

The model described above is used for two purposes: preliminary design and performance analysis of one-stage axial turbines. Figure 5 presents the workflow for the preliminary design mode, and where the performance analysis fits into this. The turbine model is used for both preliminary design and performance analysis to ensure consistency between the two modes, but there are certain differences in problem formulation, which will be described in this section.

3.1. Optimization

The goal of the optimization objective is to find the optimal geometry given the thermodynamic fixed variables. In this case, the thermodynamic boundaries (Table 3) are fixed parameters, while the geometry, rotational speed and input velocity/static pressure (Tables 1 and 2) are independent variables. This mode makes use of an optimization algorithm to tune the independent variables such that the optimal value of the objective function is reached while a set of bounds and constraints are respected. The optimization algorithm used in this method is the sequential quadratic programming (SQP) algorithm, available in the Scipy library (Virtanen et al. 2020). The objective function is defined as the total-to-static efficiency (Eq. (26)), but may easily be changed to other performance related parameters such as total-to-total efficiency or work output. The closing equations (Eq. (25)) are specified as a set of equality constraints to ensure a physical solution, while bounds may be defined for all independent variables. Other constraints may be added, both equality and inequality. For example, if the user desires to treat mass flow rate as an input variable, it can be imposed with an equality constraint.

$$\eta_{ts} = \frac{h_{0,in} - h_{0,out}}{h_{0,in} - h_{0,is}} \quad (26)$$

Table 4: Operating conditions and comparison of performance parameters between the numerical design reference (Agromayor and Nord 2019) and model presented in this paper.

Variable	Agromayor and Nord 2019	Present work
Fluid	R125	-
$T_{0,in}$	428.15 K	-
$p_{0,in}$	36.18 bar	-
p_{out}	15.34 bar	-
Ω	34 738.2	-
η_{ts}	88.72%	88.44%
\dot{m}	11.90 kg/s	11.85 kg/s
\dot{W}	227.76 kW	226.12 kW

3.2. Performance analysis

The goal of the performance analysis is to assess the thermodynamic and kinetic variables given a geometry, rotational speed and thermodynamic boundaries. This means that the thermodynamic, geometrical variables (Tables 3 and 1) and rotational speed are fixed parameters, while the velocity and static pressure inputs (Table 2) are unknown variables. The unknown variables are determined by solving the system of equations from section 2.5 (Eq. (25)), with a root finder method from the Scipy library (Virtanen et al. 2020).

4. Model validation

The turbine model has been used to simulate the performance of three experimentally investigated turbines, in addition to one numerical test case. The numerical case corresponds to the case considered in (Agromayor and Nord 2019), and present work is validated against this case due to the similarities of the models. The three experimental cases are cold-air tests of one-stage axial turbines, both at subsonic and transonic conditions (Haas and Kofskey 1975; Moffitt et al. 1980; Nusbaum and Kofskey 1972). The experimental values were attempted to be replicated at design pressure ratio and design rotational speed. The metal angles were assumed to equal the design flow angles (Eq. (15)). Note that the values in the experimental reports are given in terms of equivalent conditions, and are here converted to ordinary values, using the conditions at standard sea-level (temperature: 288.15 K; pressure: 1.01325 bar).

4.1. Numerical case

The operating conditions, and corresponding performance parameters are presented in Table 4. Note that the reference case includes a diffuser model after turbine exit, which is not considered here, and the efficiency is determined with the state at the exit of the turbine. As expected the results match well as the models used in the reference case and this work are similar.

4.2. Experimental case 1

The first experimental case to be simulated is the turbine documented in (Nusbaum and Kofskey 1972). Table 5 show the operating conditions and corresponding performance parameters at design pressure ratio and rotational speed. The results show a deviation in total-to-static efficiency of 3.52 percentage points, which is larger than

Table 5: Operating conditions and comparison of performance parameters between the first experimental reference (Nusbaum and Kofskey 1972) and model presented in this paper.

Variable	Nusbaum and Kofskey 1972	Present work
Fluid	Air	-
$T_{0,in}$	295.6 K	-
$p_{0,in}$	13.8 bar	-
$p_{0,in}/p_{out}$	2.298	-
Ω	15533	-
η_{ts}	80.00%	76.48%
\dot{m}	2.696 kg/s	2.740 kg/s
\dot{W}	134.14 kW	131.58 kW

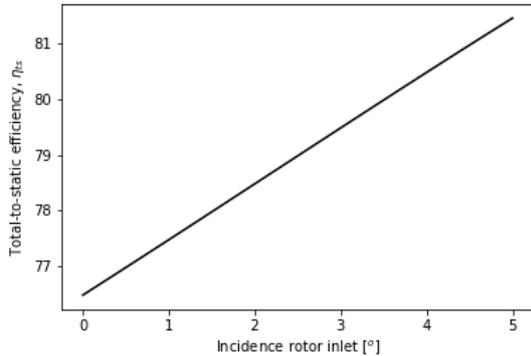


Figure 6: Total-to-static efficiency as a function of rotor inlet incidence. The relation looks linear, but it is observed that it becomes nonlinear outside this region.

the uncertainty of the loss model by Kacker and Okapuu (± 1.5 percentage points) (Kacker and Okapuu 1982). The predicted mass flow rate exceed the experimental value with 1.64%, while the predicted work out is 1.91% less than the actual value. The deviation in predicted design-point efficiency was unexpected as Agromayor and Nord used the same loss model and same reference case for validation, but predicted an efficiency within the loss model uncertainty (78.85%). Nevertheless, the report explains that the rotor inlet whirl is greater than design, and hence the assumption from Eq. (15) is not satisfied. Therefore, simulations with incidence were performed and Figure 6 shows the total-to-static efficiency as a function of rotor incidence. The figure illustrate that an efficiency of 80% was achieved at about 3.5° incidence.

In addition, the predictions of the turbine model are also evaluated at off-design conditions. Figure 7 show the mass flow rate as a function of total-to-static pressure ratio and rotational speed. The lines represent predicted mass flow rate, while the markers illustrate the experimental values. The figure show that the model predicts similar behaviour as the experimental case, but with certain differences. Most noticeable is the unphysical "bump" that occur as the pressure ratio approaches the critical value. It is observed that the bump is non-existent for isentropic flow, and becomes more significant as the efficiency decreases. Thus, it is likely that this behaviour is related to the way losses are predicted throughout the turbine. Furthermore, similar as for the design point comparison, the predicted choked mass flow rate exceed the experimental value slightly. The choked mass flow difference in percentage is approximately 1.12% and 2.61% at design ro-

Table 6: Operating conditions and comparison of performance parameters between the second experimental reference (Haas and Kofskey 1975) and model presented in this paper.

Variable	Haas and Kofskey 1975	Present work
Fluid	Air	-
$T_{0,in}$	300 K	-
$p_{0,in}$	8.27 bar	-
$p_{0,in}/p_{out}$	3.16	-
Ω	32 100 rpm	-
η_{tt}	83.2%	84.3%
\dot{m}	0.185 kg/s	0.184 kg/s
\dot{W}	11.73 kW	11.92 kW

tational speed and 30% of design rotational speed, respectively. As expected the deviation is larger for off-design rotational speed. Moreover, it can be observed that the critical pressure ratio is slightly higher then experimental value (approximately 2.7 vs. 2.6).

4.3. Experimental case 2

The second experimental reference is the case presented in (Haas and Kofskey 1975). The operating conditions and performance parameters are shown in Table 6. Note that the efficiency here is the total-to-total efficiency. The results show that the efficiency is predicted well, with a deviation of 1.1 percentage points, while the predicted mass flow rate and work output differs from the experimental value by 0.66% and 1.56% respectively. The original report indicates that the stator throat area was fabricated 5.5% smaller than the design specification, which is accounted for in the simulations. Otherwise, the deviation in both mass flow rate and work output would exceed 4.5%.

4.4. Experimental case 3

Information on the third experimental turbine is presented in (Moffitt et al. 1980). This study presents a turbine operating at transonic conditions with relatively high specific work output. The experimental conditions and results are shown in Table 7. Note that the pressure ratio here is the total-to-total pressure ratio. The results show a deviation of 4.5 percentage points in efficiency, 3.36% in mass flow rate and 6.11% in work output. The results suggest that the model predicts performance of transonic turbines less

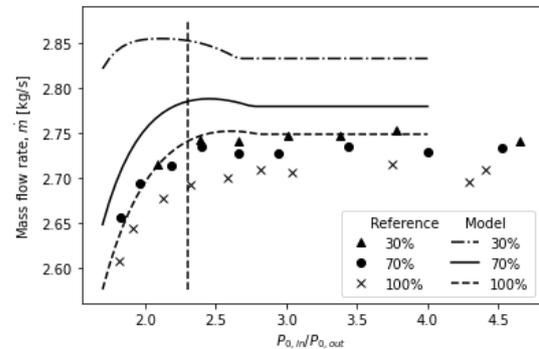


Figure 7: Mass flow rate as a function of total-to-static pressure ratio and percentage of design rotational speed. The lines represent predicted values, while the markers are the measured values for Experimental case 1. The vertical dashed line indicates the design pressure ratio.

Table 7: Operating conditions and comparison of performance parameters between the third experimental reference (Moffitt et al. 1980) and model presented in this paper.

Variable	Moffitt et al. 1980	Present work
Fluid	Air	-
$T_{0,in}$	378 K	-
$p_{0,in}$	24.13 bar	-
$p_{0,in}/p_{0,out}$	3.44	-
Ω	10 600.8 rpm	-
η_{tt}	88.6%	84.1%
\dot{m}	8.018 kg/s	7.748 kg/s
\dot{W}	799.33 kW	750.52 kW

Table 8: Case definition and results for the design case study. The results are compared to the reference design case described in (Agromayor and Nord 2019).

Variable	Agromayor and Nord 2019		Present work	
Fluid	R125		-	
$T_{0,in}$	428.15 K		-	
$p_{0,in}$	36.18 bar		-	
p_{out}	15.34 bar		-	
\dot{m}	11.90 kg/s		-	
η_{ts}	88.72%		89.56%	
Ω	34 738.2 rpm		28 647.9 rpm	
\dot{W}	227.76 kW		229.89 kW	
Row	Stator	Rotor	Stator	Rotor
r_m	4.04 cm		4.30 cm	
H	2.02 cm	2.02 cm	2.02 cm	2.02 cm
(H/c)	1.772	2.000	1.900	2.000
(s/c)	0.579	0.693	0.588	0.690
(t_{te}/o)	0.0479	0.0489	0.0500	0.0500
θ_{in}	-	-15.00°	-	-15.01°
θ_{out}	80.00°	-74.49°	80.00°	-74.44°
δ_{fl}	0°	0°	-8.83°	10.00°

accurately than subsonic turbines, which is consistent with the high uncertainty associated with the supersonic correction factor of the Kacker and Okapuu loss model.

5. Design case study

A design case study was conducted to showcase the design optimization method. The results obtained for this design exercise are compared to the design obtained with the method described in (Agromayor and Nord 2019). As mentioned, the mass flow rate is an input variable for reference model. Thus, the mass flow rate was here constrained to match the reference case and obtain comparable results. The rest of the bounds and constraints were set such that they also match for the two models. The fixed input variables, key performance parameters and geometry are presented in Table 8.

The results show that this model predicts a design with 0.84 percentage points higher efficiency, and a very similar geometry except for the flaring angle. Indeed, the proposed model predicts highest efficiency at -8.83° flaring of stator row, and 10° of rotor row, with $\pm 10^\circ$ being the upper and lower bounds. The exact same constraint is applied for the reference case, while that model predicts highest efficiency at 0° flaring.

6. Summary and Discussions

In this paper a method for preliminary design and performance analysis of one-stage axial turbines have been presented. This method is intended to be the basis for a robust design optimization method that takes into account off-design operating conditions. Both the performance analysis and design optimization modes have been tested and compared against relevant data.

The validation section showed that the model performed variably for the different reference cases. As expected, the model predicted similar performance for the numerical case study, but deviated significantly in some of the experimental cases. It is likely that the assumption of zero incidence/deviation is violated, which could cause the error in predicted efficiency. Differences in fabricated turbine vs. design could also cause deviations in performance parameters. Moreover, the model error was also observed to be greater for transonic conditions compared to subsonic, which is explained by the high uncertainty of the supersonic correction factor of the Kacker and Okapuu loss model. In addition, unphysical behavior of the mass flow vs. pressure ratio curve was observed when the pressure ratio approached the critical value, which could cause the model to be less accurate. It is likely that this behaviour is a consequence of how the losses are calculated at the blade throats. As discussed previously, the loss coefficient is evaluated at the throat in the same way as the exit, while a strategy to split the losses could solve this problem.

The design case study resulted in a geometry similar to the reference case, with converging-diverging flaring angles being the most noticeable difference. This affects the flow area and hence the flow velocities, making the velocity higher at stator exit and lower at rotor exit compared to the reference case. The difference of flaring angle could arise as the reference case included a diffuser model at turbine exhaust, which recovers a fraction of the exit kinetic energy. Nevertheless, the preliminary design optimization method is confirmed to give similar results as comparable preliminary design tools.

7. Further work

To improve the accuracy of the model, it is suggested to do a comparative analysis of several loss models for axial turbines, and to account for incidence in the loss calculation. All loss models considered should be compared against experimental data, and a data collection should be conducted, either from available literature or experimental investigation. The data collection should also contain data from transonic/supersonic turbines to validate models in this flow regime. Only when this is conducted one can assess which loss model to use in this method.

In addition, the design method will be extended to account for off-design conditions. The objective function for this purpose will be extended to be a weighted average of efficiencies at several design points. The efficiencies could be weighted according to the occurrence of the corresponding conditions over an annual perspective.

Acknowledgment

The research leading to these results has received funding from the EEA Grants and the Technology Agency of the Czech Republic within the KAPPA Programme. Project code TO01000160.

References

- Agromayor, Roberto and Lars O. Nord (2019). “Preliminary design and optimization of axial turbines accounting for diffuser performance”. eng. In: *International journal of turbomachinery, propulsion and power* 4.3. Place: Basel Publisher: MDPI AG, p. 32. ISSN: 2504-186X.
- Aissa, Mohamed Hassanine et al. (2019). “Optimisation of a turbine inlet guide vane by gradient-based and metamodel-assisted methods”. eng. In: *International journal of computational fluid dynamics* 33.6-7. Place: Abingdon Publisher: Taylor & Francis, pp. 302–316. ISSN: 1061-8562.
- Bell, Ian H et al. (2014). “Pure and Pseudo-pure Fluid Thermophysical Property Evaluation and the Open-Source Thermophysical Property Library CoolProp”. eng. In: *Industrial & engineering chemistry research* 53.6. Place: WASHINGTON Publisher: American Chemical Society, pp. 2498–2508. ISSN: 0888-5885.
- Bonaiuti, Duccio and Mehrdad Zangeneh (2009). “On the Coupling of Inverse Design and Optimization Techniques for the Multiobjective, Multipoint Design of Turbomachinery Blades”. eng. In: *Journal of turbomachinery* 131.2. Place: NEW YORK Publisher: ASME, 21014–021014 (16). ISSN: 0889-504X.
- Châtel, Arnaud, Tom Verstraete, and Grégory Coussement (2020). “Multipoint optimization of an axial turbine cascade using a hybrid algorithm”. eng. In: *Journal of turbomachinery*. Publisher: ASME, pp. 1–15. ISSN: 0889-504X.
- Colonna, Piero et al. (2015). “Organic Rankine Cycle Power Systems: From the Concept to Current Technology, Applications, and an Outlook to the Future”. eng. In: *Journal of engineering for gas turbines and power* 137.10. Place: NEW YORK Publisher: ASME. ISSN: 0742-4795.
- Denton, J.D (1993). “Loss mechanisms in turbomachines”. eng. In: *Journal of turbomachinery* 115.4. Place: New York, NY Publisher: American Society of Mechanical Engineers, pp. 621–656. ISSN: 0889-504X.
- Dixon, S.L. (2014). *Fluid mechanics and thermodynamics of turbomachinery*. eng. Edition: 7th ed. ISBN: 9780124159549 Place: Amsterdam.
- Haas, J. E and M. G Kofskey (1975). *Cold-air performance of a 12.766-centimeter-tip-diameter axial-flow cooled turbine. 1: Design and performance of a solid blade configuration*. eng.
- Jiménez-Arreola, Manuel et al. (2018). “Thermal power fluctuations in waste heat to power systems: An overview on the challenges and current solutions”. eng. In: *Applied thermal engineering* 134. Place: OXFORD Publisher: Elsevier Ltd, pp. 576–584. ISSN: 1359-4311.
- Kacker, S. C and U Okapuu (1982). “A Mean Line Prediction Method for Axial Flow Turbine Efficiency”. eng. In: *Journal of engineering for gas turbines and power* 104. Publisher: ASME, pp. 111–119. ISSN: 0742-4795.
- Macchi, Ennio and Marco Astolfi (2017). *Organic rankine cycle (ORC) power systems : technologies and applications*. eng. ISBN: 0-08-100511-3 Place: Amsterdam, Netherlands Series: Woodhead Publishing in energy ; Volume: Number 107.
- Moffitt, T. P et al. (1980). *Design and cold-air test of single-stage uncooled turbine with high work output*. eng.
- Moran, M.J. et al. (2015). *Principles of Engineering Thermodynamics*. Wiley. ISBN: 978-1-118-96088-2. URL: <https://books.google.no/books?id=zFQwBgAAQBAJ>.
- Nusbaum, W. J and M. G Kofskey (1972). *Design and cold-air investigation of a turbine for a small low-cost turbofan engine*. eng.
- Persico, G and M Pini (2017). “Fluid dynamic design of Organic Rankine Cycle turbines”. eng. In: *Organic Rankine Cycle (ORC) Power Systems: Technologies and Applications*, pp. 253–297. ISBN: 978-0-08-100510-1.
- Pini, M., G. Persico, and V. Dossena (June 2014). *Robust Adjoint-Based Shape Optimization of Supersonic Turbomachinery Cascades*. Vol. Volume 2B: Turbomachinery. Turbo Expo: Power for Land, Sea, and Air. DOI: 10.1115/GT2014-27064. URL: <https://doi.org/10.1115/GT2014-27064>.
- Sanchez Torreguitart, Ismael, Tom Verstraete, and Lasse Mueller (2018). “CAD and Adjoint Based Multipoint Optimization of an Axial Turbine Profile”. eng. In: *Evolutionary and Deterministic Methods for Design Optimization and Control With Applications to Industrial and Societal Problems*. Computational Methods in Applied Sciences. ISSN: 1871-3033. Cham: Springer International Publishing, pp. 35–46. ISBN: 978-3-319-89889-6.
- Saravanamutto, H.I.H. et al. (2017). *Gas Turbine Theory*. eng. 7th. Harlow: Pearson Education Limited. ISBN: 978-1-292-09309-3.
- Virtanen, Pauli et al. (2020). “Author Correction: SciPy 1.0: fundamental algorithms for scientific computing in Python”. eng. In: *Nature methods* 17.3. Backup Publisher: SciPy 1.0 Contributors Place: United States Publisher: Nature Publishing Group, pp. 352–352. ISSN: 1548-7091.

