Doctoral theses at NTNU, 2022:193

Andres Carranza Abaid

Integration of Artificial Neural Networks and First Principles Modelling

Application to Thermodynamics and CO_2 Capture Processes

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Natural Sciences
Department of Chemical Engineering

NTNU
Norwegian University of
Science and Technology

Andres Carranza Abaid

Integration of Artificial Neural Networks and First Principles Modelling

Application to Thermodynamics and CO_2 Capture Processes

Thesis for the Degree of Philosophiae Doctor

Trondheim, June 2022

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Chemical Engineering

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Natural Sciences
Department of Chemical Engineering

ISBN 978-82-326-6235-7 (printed ver.)
ISBN 978-82-326-5215-0 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:193

Printed by NTNU Grafisk senter

Abstract

Hybrid modelling methods consolidate first-principles models and machine learning algorithms such as Artificial Neural Networks (ANNs). This combination is claimed to have superior modelling capabilities while preserving the physics rigorousness of the first-principles model. The open literature available hybrid modelling methods usually set up a structure where the mechanistic model and the ANN exchange information but still are independent models. Therefore, the available hybrid modelling methods do not exploit important features of the ANN training algorithms.

This thesis introduces a new hybrid modelling paradigm called Neural Network Programming (NNP), in order to maximize the potential of the ANN training algorithms. NNP utilizes first-principle-based equations and structural analysis to formulate a physically coherent Algorithmically Structured Neural Network (ASNN). The key difference between a typical fully connected ANN and an ASNN is that the former utilizes generic architectures that do not represent the unique modelled phenomena while an ASNN has a priori knowledge that is transcribed in its architecture. NNP enables chemical engineers to comprehend and shape the ANN architecture according to their needs, instead of borrowing generic neural network architectures.

The first section of this thesis presents applications of the NNP method to a diverse chemical engineering subdisciplines with emphasis on equilibrium thermodynamics. Four main points are made regarding the integration of thermodynamics and ANNs: (1) it is demonstrated that basic principles such as the Gibbs' phase rule is paramount to formulate an appropriate Vapor-Liquid Equilibrium (VLE) model, (2) the typical fully connected neural networks cannot be thermodynamically consistent for VLE modelling. Hence, the importance of analyzing the neural network structure for

achieving thermodynamic consistency is stressed, (3) it is shown that NNP can be applied to formulate thermodynamically consistent models by transcribing the Wilson and NRTL to the architecture of an ASNN, and (4) the formulation of thermodynamic functions that exploit the universal approximator theorem is proposed. In particular, an activity coefficient model that overrides the need of performing chemical equilibrium calculations is suggested.

The second section presents a simulation framework and algorithmic improvements for the techno-economic evaluation of solvent-based acid gas treating technologies. Specifically, processes involving aqueous amines and organic solvents for biogas upgrading and CO_2 capture. Lastly, ideas and concepts for the implementation of solvent screening simulation frameworks based on ANN-driven algorithms is discussed.

Acknowledgements

I wish to begin thanking my supervisor Prof. Jana P. Jakobsen for giving me the opportunity for pursuing a PhD degree at the Department of Chemical Engineering in NTNU. I must say that it was a pleasure working with you during these 4 years. I really appreciate the trust and support that you gave me during the development of this research project. Your fundamentals-oriented perspective and encouragement for out-of-the box thinking has always been an inspiration to me. Aside from the professional side, I am also grateful for our nice and interesting conversations that we had.

I would like to acknowledge Prof. Hallvard Svendsen for the valuable advice, discussions, and motivating feedback that he always gave to me. Furthermore, I want to thank Prof. Hanna Knuutila for her advice and for being understanding and supportive through these years. Additionally, I am grateful for all the interesting and philosophical discussions that I shared with Assoc. Prof. Tore Haug-Warberg. Although quite fruitful, these conversations always reminded me that thermodynamics are hard, and I barely understand them.

I want to express my gratitude to the people who made my work at NTNU amusing, interesting, and, above all, pleasant. Eirini, the nicest friend whom I shared wonderful conversations and countless late sittings. All the members of the B. Ny. Eggplant crew. Tobias and Ricardo who were nice enough to teach me the most important skill I acquired during the PhD: cooking. Canberk for being such a cool guy. The nice gals: Martina, Stine, and Ainara who always inspired me to improve my guacamole recipe. Bahareh K., who is the coolest Bahareh of this global warmed world. Vanja, a great office mate (5 out of 5 stars, 100 % recommended). I am also thankful for the nice

conversations that I have had throughout the years with Andressa, Visha, Karen, Lucas, Maxime, Lodin, Seniz, Umesh, Ida, Ina, Jasper, Blanca, Putta, Oscar, Mohammed, Saravanan, Pedro, Katherina, Trieu, Rouzbeh, and Ljubiša “the stranger”. I want to thank Mikael for his technical support and to everyone who, at some point, opened the K4 doors every time I forgot my card and saved me from the evil weather.

Above everything, I want to thank and dedicate this thesis to my family Maria Isabel, Gabriel, Gabriel (T), Eleni, Ulises, Marko, Dory “Gorda”, and Rocky. Your constant support through different stages of my life has always meant so much to me that I cannot describe it with some rushed sentences that I came up with at 5 AM before delivering this thesis.

Finally, I want to thank everyone who stoically listened to my rants about the most insignificant topics. Every. Single. Day.

And thanks to you, who picked up this thesis from a dark corner of the internet or a dusty and long-forgotten shelf.

TLDR: GG EZ and thanks y’all. :v

Trondheim, 2022

Andres Carranza Abaid

Contents

Chapter 1. Introduction 1

Chapter 2. VLE modelling with relaxed constraints 23

Chapter 3. Neural network programming 87

Chapter 4. Constrained VLE modelling 161

Chapter 5. Universal approximators as generating functions 245

Chapter 6. A simulation framework for solvent-based biogas upgrading processes 295

Chapter 7. A computationally efficient algorithm for designing unit operations 395

Chapter 8. Concluding remarks 441

List of symbols

Acronyms related to artificial intelligence

AI	Artificial Intelligence
ANN	Artificial Neural Network
ASNN	Algorithmically Structured Neural Network
BR	Bayesian Regularization training algorithm
CNN	Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
FFNN	Feed Forward Neural Network
LM	Levenberg-Marquardt training algorithm
ML	Machine Learning
NNP	Neural Network Programming
RNN	Recurrent Neural Network
SNN	Shallow Neural Network

Acronyms related to CO_2 capture

AM	Aqueous Amine chemical absorption without CO_2 recovery
AMCR	Aqueous Amine chemical absorption with CO_2 recovery
AMP	2-amino-2-methyil-1-propanol

BECCS	Bioenergy with Carbon Capture and Storage
CCS	Carbon Capture and Storage
MDEA	N-methyl diethanolamine
MEA	2-aminoethano-1-ol / Monoethanolamine
PW	Physical/Water absorption without CO_2 recovery
PWCR	Physical/Water absorption with CO_2 recovery
PZ	1,4-Diazacyclohexane / Piperazine

Variables related to thermodynamics

G	Gibbs' energy
H	Enthalpy
P	Pressure
R	Gas constant
T	Temperature
x	Liquid molar fraction
y	Vapor molar fraction
z	Feed molar fraction

Important mathematical operators

$\odot$	Schur / Hadamard product
$\oslash$	Schur / Hadamard division

Chapter 1.

Introduction

A summary of the history and evolution of AI and ANNs in the past 80 years is first presented in this chapter. The modelling potential of ANNs as well as their challenges of being physically coherent are discussed. After the background, this chapter also states the motivation, objective, hypothesis, and an outline of the work that comprises this thesis.

1.1. Background

1.1.1. Artificial intelligence and chemical engineering

The importance of Artificial Intelligence (AI) in today's world is unquantifiable and unquestionable. AI is critical in high-tech projects such as in the development of self-driving planetary rovers or for trivial tasks such as the recommendation system in the Steam store. The development and evolution AI since its conception is shown in Figure 1.1. Additionally, milestones towards a successful integration of AI and chemical engineering are highlighted in green in Figure 1.1.

The history of AI started with the notion that "machines think", idea envisioned by Alan Turing in 1950 [1]. He proposed that a machine can be deemed as intelligent if its behavior is comparable to that of a human. However, eight decades later, the AI definition has morphed, and it is dependent on the context and environment in which the AI is implemented.

An interesting and particularly useful subset of AI is Machine Learning (ML). ML is a set of algorithms that utilize advanced computational methods to make inferences about datasets. Artificial Neural Networks (ANNs) are a prominent subset of ML algorithms that have the flexibility to be adapted to a wide variety of problems. ANNs can be applied for regressing data, analyze uncertainty, diagnose faults, classify data, image recognition, or even perform complex decisions based on a multiagent scheme and beat the best chess and videogame players in the world. Because of this and extensive media coverage, ANNs have surfaced as a promising tool for chemical engineering modelling.

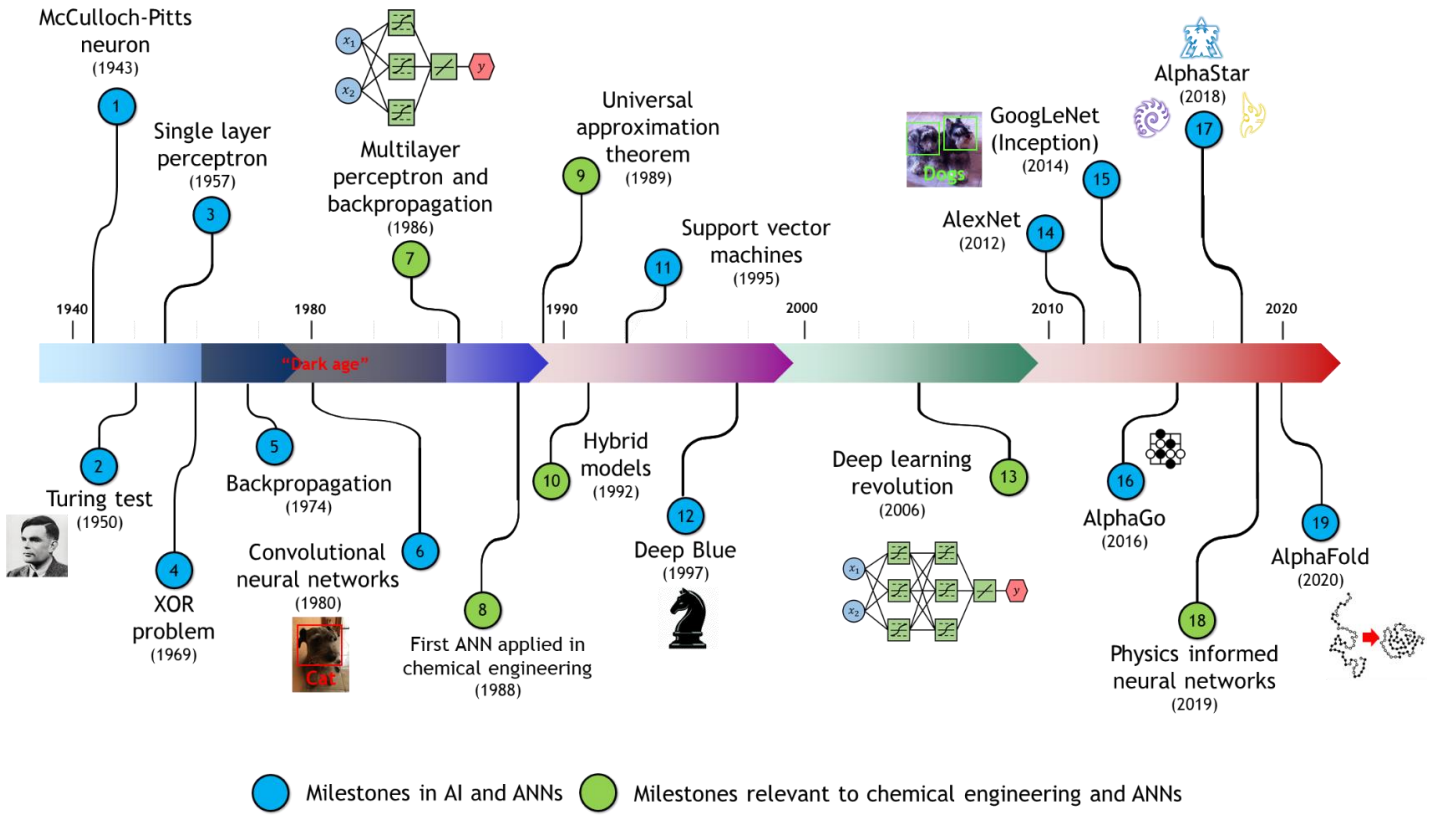


Figure 1.1. Milestones in AI, ANNs and chemical engineering.

Despite the fact that ANNs are a subset of AI, the history of ANNs starts prior to Alan Turing's work. This is because McCulloch and Pitts [2] formulated the first artificial neuron model for applications in neuroscience rather than for computer science. The goal of their model was to replicate the process that occurs in human brains. The McCulloch and Pitts neuron (MCN) model takes an input signal, processes the information with a linear combination and a decision function, and finally transmits the output to the user. MCN is a binary model that utilizes thresholding logic. Hence, it only accepts binary inputs, provides binary outputs, and uses a tunable thresholding parameter instead of individually weighting the inputs.

Regardless of its limitations, the MCN model served as basis for the formulation of better alternatives based on a similar philosophy. A remarkable improvement of the MC model is the single layer perceptron (SLP) proposed by Rosenblatt in 1957 [3,4] (an illustration of an SLP is shown in Figure 1.2 a)). In this model, the SLP receives an input signal, performs a non-linear transformation of the linear combination of the inputs, and finally provides the output to the user. As opposed to the MCN, the SLP accepts any real input, can perform both Boolean and non-Boolean predictions, and weights the inputs.

The future of neuron-based models seemed promising until 1969 when Minsky and Papert [5] showed that a SLP was incapable of learning the XOR operation. This is due to the fact that SLP can only converge on linearly separated data. This discovery began what is known as the dark age in the history of neuron-based AI technology since other options could handle this problem. Thus, the interest in neural networks and their corresponding funding waned [6].

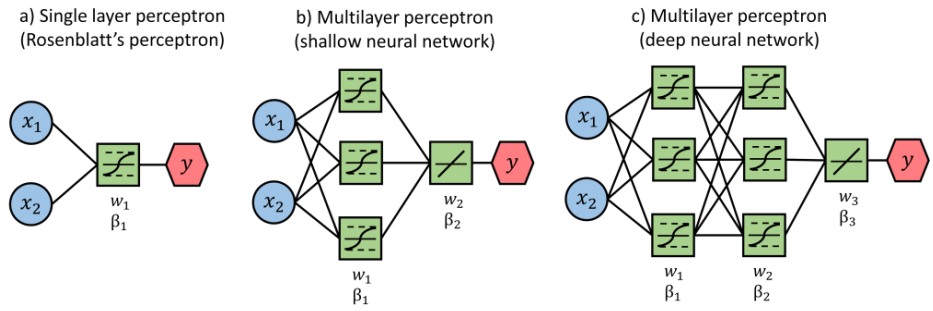


Figure 1.2. Examples of the artificial neural network architectures.

The renaissance of the perceptron-based technology began with the proposal of multilayer perceptrons (MLP) by Rumelhart, Hinton and Williams in 1986 [7]. Nowadays the models that utilize MLP as the basic building block are known as ANNs (e.g., ANNs are shown in Figure 1.2 b) and Figure 1.2 c)). They suggested using sequentially connected perceptron layers whose weights were trained with a backpropagation algorithm in order to approximate different mathematical functions. Each layer in the MLP processes the information in the same fashion as the SLP proposed by Rosenblatt. With this new idea, ANNs finally managed to solve the XOR problem that haunted the research field for decades.

One of the most important features of the MLP approach is that it introduced the backpropagation algorithm for optimizing the MLP weights (fitting parameters). Even though the backpropagation algorithm was proposed in several instances in the 1960s [8], it was Werbos in 1974 [9,10] who first suggested its application to neural networks. Independently of who proposed this algorithm first, the integration of backpropagation and MLP changed ANNs forever.

Although the work by Hinton suggested that utilizing ANNs could solve a broad set of problems, it was until 1989 when it was demonstrated that ANNs can serve as

universal approximators [11–13]. These papers demonstrated that by superpositioning sigmoid functions, it is possible to approximate any function. Later, in 1991, Hornik showed that the universal approximation feature was due to the architecture of the neural network and not the choice of nonlinear function [14].

Regardless of the mathematical implications, the fact that an ANN is capable of approximating any mathematical function independently of its complexity created countless areas of opportunity. Several research fields, like chemical engineering, found the perceptron technology as an enhanced modelling alternative. In fact, the first application of ANNs to a chemical engineering problem was done just a couple of years later after the conception of the MLP by Hoskins and Himmelblau [15]. They utilized an ANN approach for fault detection and diagnosis on a series of continuously stirred tank reactors. The impact of ANNs has been so significant in chemical engineering that there are more than 15,000 peer-reviewed publications combining these two topics (SCOPUS database as of 02/2022).

By definition, utilizing an ANN to find numerical relationships in a dataset is a merely empirical modelling approach. This implies that there is no certainty that basic concepts and constraints are fulfilled within the ANN. In order to address this, the concept of hybrid modelling was proposed by Psichogios et al. in 1992 [16]. Hybrid models are composed of a mechanistic model based on first principles equations and an ANN. Therefore, they leverage ANNs as to substitute the empirical parameters< of mechanistic models. For example, instead of utilizing the Antoine's equation parametrization, the model uses an ANN. Hybrid models allow the user to exploit the multidimensional and universal approximator properties of the ANN. This allows finding better correlations between the model inputs and outputs and while still complying with physics laws and constraints. The application of hybrid modelling has

been quite successful and has inspired more than 1000 research documents in chemical engineering alone. Examples of their application have been observed in process monitoring, process control, process optimization, feasibility analysis, model-plant mismatch analysis, environmental science, predictive maintenance, virtual flow metering, among others [17,18]. Almost two decades later, the Physics-Informed Neural Networks (PINNs) were proposed by Raissi et al. in 2018 [19]. PINNs are a deep learning framework for solving problems that involve partial differential equations. PINNs are trained by summing two errors namely, the neural network error and the error of the mechanistic model function. PINNs have been quite successful in several research fields for instance in fluid dynamics, plasma physics or biophysics [20].

By looking at Figure 1.1, the reader might ask why it took so long from envisioning of MLP (1986), to the deep learning revolution (2006)? The simplest answer is related to the computer processing capabilities and the availability of “big data”. However, the advent of deep learning came with the greedy layer-wise training algorithm developed for training Deep Neural Networks [21,22]. Improved algorithms, together with more computational power created the best scenario for the exploration and improvement of DNN architectures aimed at solving different problems.

Image processing is one of the research fields that has been benefitted the most with Deep Learning. Some of the most remarkable projects in image recognition are ImageNet (AlexNet) [23] and GoogLeNet (Inception) [24], which provided enhanced accuracy than previous systems. Another field that has fascinated computer scientists for decades is the deployment of AIs that are smarter than humans (according to Turing’s definition of machine intelligence). Although not based on ANN technology, one of the most memorable AIs is “Deep Blue”, that managed to beat the current

chess world champion in 1997, Gary Kasparov. Eventually, AI enthusiasts (DeepMind research group) have developed outstanding AIs for classical games such as Go (2016) or chess (2017) [25,26] and even multi-agent schemes for playing real-time strategy videogames such as StarCraft II (2018) [27]. More recently, in 2018 the DeepMind team deployed the AlphaFold model which is capable of accurately solving protein structure problems [28].

Considering these fantastic achievements in AI and computer science, there is no denying that sooner than later chemical engineering will be transformed by machine and deep learning.

1.1.2. Applicability of ANNs to chemical engineering

As usual with new promising technologies, there are skeptics and naysayers (like me :v). This is unsurprising, especially given that it is the first time since the scientific revolution that a technology surpasses the understanding of its makers. To my knowledge, Machine/Deep Learning is the only computer-based technology that has a specialized research field focused on understanding how and why they work [29]. One of the most incomprehensible characteristics about ANNs is the unpredictability and highly entropic nature of the algorithm. In other words, if the same data and ANN model are used, different fitting parameters will be obtained. Despite this, in most cases, all the models have competitive and comparable prediction capabilities.

From a more fundamental and conservative viewpoint, perceptron-based models lack the rigorousness and finesse of first-principles models. In this context, the phrase commonly attributed to J.S. Denker, “neural networks are the second-best way to do almost anything” [30] suggests that mechanistic modelling is a better alternative.

It is true that due to these factors, ANNs do not seem too promising for understanding physics phenomena in detail. However, ANN-based models are exceptional at pointing out correlations between the inputs and outputs due to their universal approximator feature. This characteristic together with the entropic nature of ANN modelling makes them ideal for temporarily replacing highly entropic sections of mechanistic models. The highly entropic sections are usually empirical parameters or correlations. They are utilized to compensate for knowledge gaps (or ignorance according to Jaynes' definition of entropy [31]).

The word temporarily was highlighted because, from a rigorous perspective, an accurate model without fitting parameters would be the ultimate goal. Yet, the truth is that developing such models either would take an unfeasible amount of time or is straightforward impossible.

Considering the above discussion, ANNs have high potential to be used for temporarily substituting unknown functions. Thus, speeding up the implementation of simulation and optimization frameworks without the need of describing every phenomenological detail.

1.2. Motivation

The universal approximation feature is one of the most important features of ANNs in order to be used for modelling physics phenomena. Unfortunately, it is also its main drawback. The reason for this is that, as the name of the feature implies, it is an approximator that is being used for exact sciences. Based alone on the feature name, discrepancies between the physics concepts and the ANN model are expected. Thus, the models cannot be considered to be "correct". But how can a model be defined as correct? A famous answer to this question was given by George E.P. Box who

declared “all models are wrong, some are useful”. Although from a philosophical point of view, this statement is plausible, using this definition of correct does not move forward this discussion.

The typical method for assessing the correctness of ANN-based models is by using a “consequential” approach. This indicates that it only evaluates its validity based on statistical methods (e.g., a performance function). However, I find this approach to be shallow because it disregards first principles concepts.

It must not be ignored that most physics laws are mathematized human interpretations of nature. In many cases physics laws are used in the data acquisition, e.g., the Beer-Lambert law is used to relate the energy absorbed by a solution and its concentration. Therefore, one cannot just ignore physics laws in the development of the models since the training data comes from the application of physics laws in the first place. I consider that in order for an ANN model to be correct, it should comply with the physics framework it is approximating. The priority of this “deontological” approach aims at preserving the physics coherence. Nonetheless, it is evident that the usefulness of a model depends on whether it can make accurate predictions or not. Therefore, the model performance is equally important and should be considered as equally important.

It should be commented that the universal approximation feature of ANNs is quite impressive. In fact, the discrepancy that can exist between the physics laws and constraints can be “reasonably small”. While this is not a problem for several chemical engineering subfields like process control, process monitoring or feasibility analysis, the same cannot be said for more fundamental disciplines like kinetics, transport phenomena or thermodynamics. Of particular interest is equilibrium thermodynamics since it is the foundation of most process models. The intrinsic thermodynamic

relationships that exist between the different variables provide valuable information (e.g., predicting the heat of vaporization from the saturation pressure through the Clausius-Clapeyron equation). Thus, if the model is not thermodynamically consistent, then it is not appropriate to utilize these relationships

Due to the recent developments, it is reasonable to assume that in the coming decades machine learning algorithms will become a dominant modelling approach. Therefore, the motivation of this work is to propose methods and algorithms that allow the integration of artificial neural networks and scientific knowledge. This, with the purpose of making better use of available data, improve hybrid modelling, and lifting up the skepticism towards ANNs so that they can be finally accepted as a reliable modelling tool.

1.3. Hypothesis and objective

The structural characteristics of the ANNs and the universal approximation theorem suggest that a large set of mathematical functions can be represented with an ANN. Considering this, the **hypothesis** of this work is formulated as: "Every physics law can be transcribed into an Artificial Neural Network with an appropriate architecture and parametrization". Since testing this hypothesis for every physics problem is time-wise unfeasible, the scope of this thesis is narrowed to deterministic physics laws expressed with algebraic equations. This work emphasizes on equilibrium thermodynamics and process modelling. Equilibrium thermodynamics is particularly interesting since it is one of the most constrained subdisciplines. Thus, it is expected that if the methods proposed can formulate physically coherent models, there are reasonable chances that they will be useful in several subdisciplines. The physics-

coherence concept describes whether a model follows fundamental physics relationships or not. This term is akin to the thermodynamic consistency concept.

Considering this, the thesis **objective** is to: "Formulate algorithms based on ANNs for automating and speeding-up the development of physically coherent chemical engineering models".

1.4. Thesis outline

This thesis is divided in three sections, namely introduction, results, and concluding remarks. The first section (Chapter 1) discusses the possibilities and challenges of utilizing ANNs for chemical engineering problems. Chapters 2 – 7 include the results obtained during the PhD timeframe which in turn are subdivided in two topics. The first set of topics (chapters 2 – 5) involves the implementation of ANN-based hybrid modelling algorithms for chemical engineering and equilibrium thermodynamics. The second set of topics (chapters 6 - 7) is about process analysis and unit operation design algorithms. Although the methods presented in this thesis are applicable to a broad set of chemical engineering problems, this work focuses focus on CO_2 capture processes. This is because of environmental concerns and scientific curiosity. Moreover, the high complexity of the absorption phenomenon makes it an excellent candidate to highlight the advantages of the proposed algorithms. Chapter 8 discusses the applicability of the methods to help in addressing some of the challenges in modelling gas-treating processes. Additionally, it presents the conclusions of this thesis.

1.5. Research dissemination

1.5.1. List of peer-reviewed publications

ID	Publication	Status
1	Carranza-Abaid, Andres; Svendsen, Hallvard F.; Jakobsen, Jana P.; *Surrogate modelling of VLE: Integrating machine learning with thermodynamic constraints*. Chem. Eng. Sci.: X, Vol. 8, 2021. https://doi.org/10.1016/j.cesx.2020.100080	Accepted (Chapter 2)
2	Carranza-Abaid, Andres; Jakobsen, Jana P.; *Neural network programming: Integration of first principles into machine learning models*. Comput. Chem. Eng., Vol. 154, 2021.	Accepted (Chapter 3)
3	Carranza-Abaid, Andres; Svendsen, Hallvard F.; Jakobsen, Jana P.; *Thermodynamically Consistent Vapor-Liquid Equilibrium Modelling with Artificial Neural Networks.* Submitted to Fluid Phase Equilib.	Submitted (24.05.2022) (Chapter 4)
4	Carranza-Abaid, Andres; Svendsen, Hallvard F.; Jakobsen, Jana P.; *Thermodynamic Modelling using Artificial Neural Networks as Generating Functions: Application to Vapor-Liquid Equilibrium.* To be submitted.	Waiting for item #3 to be accepted (Chapter 5)
5	Carranza-Abaid, Andres; Wanderley Ricardo R.; Knuutila, Hanna K., Jakobsen, Jana P.; *Analysis and selection of optimal solvent-based technologies for biogas upgrading.* Fuel, Vol. 303, 2021. https://doi.org/10.1016/j.fuel.2021.121327	Accepted (Chapter 6)

6	Carranza-Abaid, Andres; Jakobsen, Jana P.; *A non-autonomous frame of reference for unit operation*. Computer Aided Chemical Engineering. pp. 151-156, 2020. https://doi.org/10.1016/B978-0-12-823377-1.50026-4	Accepted (Chapter 7)
7	Carranza-Abaid, Andres; Jakobsen, Jana P.; *A computationally efficient formulation of the governing equations for unit operation design*. Comput. Chem. Eng., Vol. 154, 2021. https://doi.org/10.1016/j.compchemeng.2021.107500 0098-1354	Accepted (Chapter 7)
8	Carranza-Abaid, Gonzalez-Garcia, Raul; *A Petlyuk distillation column dynamic analysis: Hysteresis and bifurcations*. Chem. Eng. Process.: - Process Intensif., Vol. 149, 2020. https://doi.org/10.1016/j.cep.2020.107843 (Not included in this thesis but written during the PhD timeframe)	Accepted.

My contribution to the publications with ID 1-4 and 6-8 were: research conceptualization, method formulation, software/model development, model validation, and writing/edited the manuscript.

On a larger picture, this work will help computational algorithm for a simulation framework. This tool is expected to perform techno-economic assessments of CO_2 removal technologies. Based on this, the simulation framework presented in research article #5 was formulated. My role on this work consisted in implementing the simulation framework, formulate/analyze the results, and wrote an important part of the manuscript.

1.5.2. List of international conferences

ID	Conference
1	Carranza Abaid, Andres; Jakobsen, Jana Poplsteinova. (2019) “Assessing CO2 absorption techno-economics”, The 10th Trondheim Conference on Carbon Capture, Transport and Storage (TCCS-10). SINTEF - NTNU; Trondheim.
2	Wanderley, Ricardo Ramos; Carranza-Abaid, Andres; Jakobsen, Jana Poplsteinova; Knuutila, Hanna K. (2019) “The energy cost of CO2 absorption: When can physical solvents outperform chemical solvents?”, The 10th Trondheim Conference on Carbon Capture, Transport and Storage (TCCS-10). SINTEF - NTNU; Trondheim.
3	Carranza-Abaid, Andres; Svendsen, Hallvard Fjøsne; Jakobsen, Jana Poplsteinova. (2020) “Developing Machine Learning Surrogate Vapor – Liquid Equilibrium Models”. University of Texas Conference on Carbon Capture and Storage. University of Texas; 2020-01-28 - 2020-01-29.
4	Carranza-Abaid, Andres; Jakobsen, Jana Poplsteinova. (2020) A Non-Autonomous Relative Frame of Reference for Unit Operation Design. European Symposium of Computer Aided Chemical Engineering (ESCAPE).
5	Carranza-Abaid, Andres; Jakobsen, Jana Poplsteinova. (2019) “Tailor-Making a Robust Model for Gas – Liquid Contactor Design”. 5th Post Combustion Capture Conference. Research Institute of Innovative Technology for the Earth; Kyoto. 2019-09-17 - 2019-09-19.

ID	Conference
6	Carranza-Abaid, Andres; Svendsen, Hallvard Fjøsne; Jakobsen, Jana Poplsteinova. (2021) “A New Apparent Approach to Create Machine Learning Based Thermodynamic Models for CO2 Capture”, 15th International Conference on Greenhouse Gas Control Technologies GHGT-15, Abu Dhabi.
7	Carranza-Abaid, Andres; Svendsen, Hallvard F.; Jakobsen, Jana Poplsteinova. (2021) “Deep Learning Programming for Modelling CO2 Capture Processes”. 6th Post Combustion Capture Conference. UK CCS Research Centre; UK. 2021-10-19 - 2021-10-21.

1.6. References

[1] A.M. Turing, I.—COMPUTING MACHINERY AND INTELLIGENCE, Mind. LIX (1950) 433–460. https://doi.org/10.1093/mind/LIX.236.433.

[2] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Phys. (1943) 113–133.

[3] F. Rosenblatt, Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, 1962.

[4] F. Rosenblatt, The Perceptron—a perceiving and recognizing automaton. Report 85-460-1., 1957.

[5] M. Minsky, S. Papert, Perceptrons: An Introduction to Computational Geometry, 1969.

[6] R.C. Eberhart, R.W. Dobbins, Early neural network development history: the age of Camelot, IEEE Eng. Med. Biol. Mag. 9 (1990) 15–18. https://doi.org/10.1109/51.59207.

[7] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature. 323 (1986) 533–536. https://doi.org/10.1038/323533a0.

[8] J. Schmidhuber, Deep Learning in neural networks: An overview, Neural Networks. 61 (2015) 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

[9] P. Werbos, Beyond regression: New tools for prediction and analysis in the behavioral sciences, Harvard University, 1974.

[10] P.J. Werbos, Backpropagation Through Time: What It Does and How to Do It, Proc. IEEE. 78 (1990) 1550–1560. https://doi.org/10.1109/5.58337.

[11] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks. 2 (1989) 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

[12] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control. Signals, Syst. 2 (1989) 303–314. https://doi.org/10.1007/BF02551274.

[13] K.I. Funahashi, On the approximate realization of continuous mappings by neural networks, Neural Networks. 2 (1989) 183–192. https://doi.org/10.1016/0893-6080(89)90003-8.

[14] K. Hornik, Approximation Capabilities of Multilayer Neural Network, Neural Networks. 4 (1991) 251–257.

[15] J.C. Hoskins, D.M. Himmelblau, Artificial neural network models of knowledge representation in chemical engineering, Comput. Chem. Eng. 12 (1988) 881–890. https://doi.org/10.1016/0098-1354(88)87015-7.

[16] D.C. Psichogios, L.H. Ungar, A hybrid neural network-first principles approach to process modeling, AIChE J. 38 (1992) 1499–1511. https://doi.org/10.1002/aic.690381003.

[17] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L.H. Chiang, M.S. Reis, Recent trends on hybrid modeling for Industry 4.0, Comput. Chem. Eng. 151 (2021) 107365. https://doi.org/10.1016/j.compchemeng.2021.107365.

[18] T. Bikmukhametov, J. Jäschke, Combining machine learning and process engineering physics towards enhanced accuracy and explainability of data-driven models, Comput. Chem. Eng. 138 (2020). https://doi.org/10.1016/j.compchemeng.2020.106834.

[19] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.

[20] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (2021) 422–440. https://doi.org/10.1038/s42254-021-00314-5.

[21] G.E. Hinton, R.R. Salakhutdinov, Reducing the Dimensionality of Data with Neural Networks, Science (80-.). 313 (2006) 504–507. https://doi.org/10.1126/science.1127647.

[22] G.E. Hinton, S. Osindero, Y.-W. Teh, A Fast Learning Algorithm for Deep Belief Nets, Neural Comput. 18 (2006) 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527.

[23] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM. 60 (2017) 84–90. https://doi.org/10.1145/3065386.

[24] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: 2015 IEEE Conf. Comput. Vis. Pattern Recognit., IEEE, 2015: pp. 1–9. https://doi.org/10.1109/CVPR.2015.7298594.

[25] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature. 529 (2016) 484–489. https://doi.org/10.1038/nature16961.

[26] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, D. Hassabis, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, Science (80-.). 362 (2018) 1140–1144. https://doi.org/10.1126/science.aar6404.

[27] O. Vinyals, I. Babuschkin, W.M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D.H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J.P. Agapiou, M. Jaderberg, A.S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T.L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, D. Silver, Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature. 575 (2019) 350–354. https://doi.org/10.1038/s41586-019-1724-z.

[28] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu,

P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature. 596 (2021) 583–589. https://doi.org/10.1038/s41586-021-03819-2.

[29] L. Zdeborová, Understanding deep learning is also a job for physicists, Nat. Phys. 16 (2020) 602–604. https://doi.org/10.1038/s41567-020-0929-2.

[30] G. National, H. Pillars, Intelligent Scene Modelling Information Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. https://doi.org/10.1007/978-3-540-92902-4.

[31] E.T. Jaynes, Information Theory and Statistical Mechanics, Phys. Rev. 106 (1957) 620–630. https://doi.org/10.1103/PhysRev.106.620.

Chapter 2.

VLE modelling with relaxed constraints

This is the first work related to the integration of thermodynamic concepts into ANN modelling. The importance of considering the degrees of freedom (Gibbs' phase rule) in thermodynamic systems is emphasized.

Surrogate Modelling of VLE: Integrating Machine Learning with Thermodynamic Constraints

Andres Carranza-Abaid*, Hallvard F. Svendsen, Jana P. Jakobsen

[a]Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), NO-7491, Trondheim, Norway

*Corresponding author: andres.c.abaid@ntnu.no

Keywords: Thermodynamics, Machine Learning, Surrogate Modelling, CO_2, MEA

2.1. Abstract

An easy-to-implement methodology to develop accurate, fast and thermodynamically consistent surrogate machine learning (ML) models for multicomponent phase equilibria is proposed. The methodology is successfully applied to predict the VLE behavior of a mixture containing CO_2, MEA, and H_2O. The accuracy of the surrogate model predictions of VLE for this system is found to be satisfactory as the results provide an average absolute relative difference of 0.50 % compared to the estimates obtained with a rigorous thermodynamic model (eNRTL + Peng-Robinson). It is demonstrated that the integration of Gibbs phase rule and physical constraints into the development of the ML models is necessary, as it ensures that the models comply with fundamental thermodynamic relationships. Finally, it is shown that the speed of ML based surrogate models can be ~10 times faster than interpolation methods and ~1000 times faster than rigorous VLE calculations.

2.2. Introduction

2.2.1. Motivation and Literature Review

The emission of anthropogenic greenhouse gases has been one of the main subjects of environmental concerns over the past decades. Development of new, clean and enhanced industrial processes has become a must in order to reach international and national sustainability goals. One of the most promising approaches to operate cleaner industrial processes is the implementation of CO_2 capture and storage. It has been labeled as one of the key technologies that will assist in achieving a global temperature increment of no more than 1.5 °C by the end of 2030 (IPCC, 2018). Albeit there are several other technologies for CO_2 capture, chemical absorption of CO_2 with aqueous amines has been, and seems to be in the foreseeable future, the most commercially ready and competitive technology [2,3]. The high energy demand associated with the CO_2 capture process is its main challenge [4].

Extensive experimental and modelling research has been conducted aiming to find a system with favorable vapor-liquid equilibrium (VLE) behavior that will lead to a more energetically efficient process. Modelling of VLE is of great importance for these efforts. The knowledge of phase behavior is necessary in order to assess the solvent performance in CO_2 capture processes. Traditionally, the VLE models are divided into two categories: semi-empirical models (also called rigorous models) and empirical models, each with their respective advantages and disadvantages.

In this work, an alternative approach to VLE modeling is proposed which attempts to merge all the main features of both categories.
The main objectives of this work are:

- To present an easy-to-implement method based on machine learning technology that combines the robustness of rigorous models with the computational efficiency of an empirical model.
- Demonstrate that machine learning models can properly predict the dependencies between the input and output variables if a proper analysis of the thermodynamic variables is performed.
- Give an insight into how to select properly the input and the output variables of a machine learning model so that it is thermodynamically consistent.

Rigorous models may use different mathematical representations like an equation of state for both phases (phi-phi formulation) or excess Gibbs energy model for liquid phase and equation of state for the gas phase (gamma-phi formulation). These models make use of governing thermodynamic equations (e.g. Henry's law, Raoult's law), semi-theoretical models (e.g. excess Gibbs energy models for aqueous electrolytic solutions [5–7]) together with empirically fitted parameters. Some successful applications of gamma-phi models for the prediction of the CO_2 solubility in several aqueous amines are: 2-aminoethan-1-ol (MEA) [8], piperazine (PZ) and 2-amino-2-methylpropan-1-ol (AMP) [9], 2-piperidineethanol (PE) [10], AMP [11] to name a few.

Although semi-empirical models have proven to give accurate predictions of VLE, the difficulty of developing a thermodynamic model for each CO_2 aqueous amine solvent and the high computational overhead of these computationally complex models have led to the development of empirical models.

A simple empirical VLE model that describes the CO_2 solubility in primary, secondary and tertiary amines was proposed by [12]. This model was developed by summing up all the reactions into a single overall reaction and lumping the equilibrium constant with the CO_2 physical solubility coefficient to obtain the CO_2 partial pressure. Another empirical approach was taken by [13–16] where empirical correlations that relate the molar compositions and the temperature with the CO_2 partial pressure were formulated. All these empirical correlations are computationally inexpensive due to their simplicity when compared to gamma-phi models. The main drawback of these empirical models is their low dimensionality, which translates into their limited validity range. Hence, the mathematical functions are different among different systems.

Machine learning has been used as an alternative to create empirical models. Several studies in the literature have used this approach to estimate the VLE of CO_2 capture related thermodynamic systems. These studies have mostly focused on the estimation of CO_2 solubility in different liquid solvents: triisopropanolamine – MEA aqueous solutions [17], various alkanol systems [18], several aqueous amine and diamine systems containing MEA, MDEA, PZ, 2-amino-2-methyl-1-propanol (AMP) [19–21], pure water [22], mixtures of ethanol and ionic liquids [23], aqueous PZ solutions [24], piperazine (PZ) and ionic liquids [25], aqueous sodium salt of L-phenylalanine [26], aqueous potassium lysinate mixed with MEA [27], among others. In general, the authors of these models claim that their machine learning models have better prediction capabilities than the semi-empirical models. This is expected, since these multidimensional models were fitted to a limited range of operating conditions. However, that is unfortunately also the reason why they cannot be extrapolated. In addition, due to their empirical nature and to the fact that they are constrained to an univariable output, the reported models are not capable of predicting other important

thermodynamic quantities such as the CO_2 heat of absorption or speciation in the liquid phase.

Because of these major limitations, a direct application of these models into a process simulation framework is not practical. Besides, in the context of CO_2 absorption into aqueous amine systems, the molar compositions of all the ionic species and the physical solubility of CO_2 are needed for the evaluation of kinetic and mass transfer rate expressions. Even though the CO_2-MEA-H_2O mixture is the most studied amine system, the available data alone is not enough to create an accurate machine learning model that will estimate the VLE behavior of all thermodynamic (e.g., liquid or vapor molar fractions) variables over a broad range of operating conditions.

2.2.2. What is Machine Learning?

Machine learning has become popular over the past years in both industrial and research environments. The sudden increase in popularity is due to the fast improvement of the technical capabilities of current-day computers. Machine learning is a computational modelling tool for data management that enables the classification, pattern recognition, clustering and data prediction [28]. Machine learning is closely related to artificial intelligence, hence, it aims at implementing intelligent agents that are capable of mimicking the cognitive functions of a biological brain in order to learn or solve complex problems [29].

One of the most prominent and notable methods of machine learning is artificial neural networks (ANN), which were first conceptualized by McCulloch and Pitts several decades ago [30]. An ANN can be defined as a nonlinear vector of functions (Ω) that needs a vector of input variables ($\underline{X}$) and a set of weight parameters ($\underline{\omega}$) in order to estimate a vector of output variables $\underline{\hat{Y}}$ [31].

In general, an ANN has the form:

$$\underline{\hat{Y}} = \Omega\left(\underline{X}; \underline{\omega}\right) \tag{2.1}$$

Due to the high dimensional nature of ANN, any set of experimental or observed values ($\underline{Y}^e$) corresponding to a certain set of experimental input values (X^e) can be approximated using Eq. (2.1) as long as an appropriate $\underline{\omega}$ is used.

The process of obtaining $\underline{\omega}$ that, together with $\underline{X}^e$, allows an estimation of the predicted output $\underline{\hat{Y}}$ which numerically resembles the experimental values $\underline{Y}^e$ is called ANN training. This process is based on the biological analogy on how a person is trained to perform a task correctly through a feedback procedure. Following this scheme, a person performs a task and receives feedback on the task performance in order to improve the result. This process goes on repeatedly until the person performs the task well enough to meet a standard. Following this analogy, the ANN is trained by utilizing a loss function ($\mathcal{L}$) that evaluates its performance by comparing the experimental values $\underline{Y}^e$ and the predicted output $\underline{\hat{Y}}$:

$$Loss\ function = \mathcal{L}\left(\underline{\hat{Y}}, \underline{Y}^e\right) \tag{2.2}$$

The loss function indicates the total error of the predicted values with respect to the observed values.

Figure 2.1 illustrates the training process of a feedforward neural network (FFNN) that has 3 input values, 4 neurons in the hidden layer and that estimates 2 values in the output layer. The FFNN inside the grey square in Figure 2.1 and is said to be a feedforward propagation model because the information flows only in one direction. First, the input vector $\underline{X}^e$ is transformed to calculate the hidden layer vector $\underline{Z}$ using

an activation function (e.g. an hyperbolic function or sigmoid function) and 2 subsets of the weight parameter vector ($\underline{\omega}^{(0)}$ and $\underline{\omega}^{(1)}$, also note that: $\underline{\omega}^{(0)} \cup \underline{\omega}^{(1)} \cup \underline{\omega}^{(2)} = \underline{\omega}$). Afterwards, vector $\underline{Z}$ uses a linear transformation together with another subset of weight parameters $\underline{\omega}^{(2)}$ to estimate $\underline{\hat{Y}}$. The output vector ($\underline{\hat{Y}}$) is then forwarded to the loss function (marked with a red square in Figure 2.1) outside the FFNN. The loss function assesses the ANN performance by comparing $\underline{\hat{Y}}$ and $\underline{Y}^e$ and, depending on this evaluation, a backpropagation signal is sent to modify $\underline{\omega}$ so that the loss function is minimized. This process keeps going until the variation of the parameters stops improving the predictions. The loss function can be optimized through different backpropagation training methods such as Levenberg-Marquardt, resilient backpropagation, Bayesian regularization propagation, the scale conjugate gradient, among others [32].

In summary, training an ANN is quite similar to the other approaches that have been used for decades or even centuries to develop mathematical models; one of such methods is polynomial fitting. For example, let's consider a case where it is desired to develop a linear model of a process. One must have a set of experimental data that contains the values of the independent variable $\underline{X}^e$ and the dependent variable $\underline{Y}^e$. A linear model contains only the slope and the intersect as the model weight parameters $\underline{\omega}$, so the «training» of a linear model should find the $\underline{\omega}$ that minimize the difference between $\underline{Y}^e$ and the values calculated with the linear model $\underline{\hat{Y}}$. Machine learning models use the same concept as polynomial fitting. The main difference is that the machine learning model equations are more complicated, hence the optimization algorithms are more sophisticated. Despite the similarities between ANN and other mathematical modelling methods, ANN has the advantage of its inherent high dimensionality that allows it to capture non-linear behaviors.

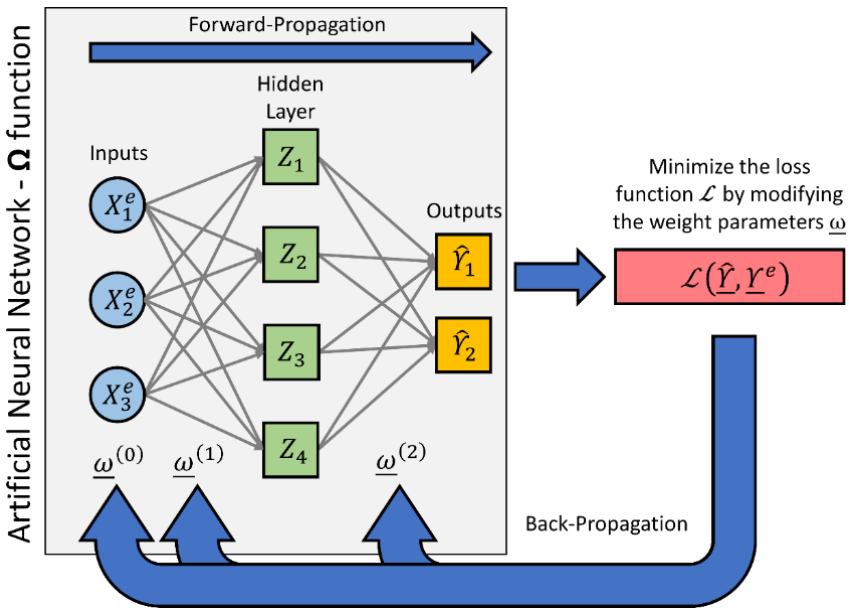


Figure 2.1. Diagram of a forward propagation artificial neural network (FFNN) trained by a backpropagation algorithm.

2.3. Methodology

2.3.1. The Thermodynamic System

The vapor-liquid equilibrium of the reactive mixture of CO_2-MEA-H_2O is a highly non-ideal thermodynamic system due to the distinct chemical nature of the components that constitute the liquid phase. There is a gas (CO_2) dissolved in an already non-ideal mixture of MEA and H_2O. Furthermore, some species are electrolytes, and the presence of ions induces highly non-ideal long-distance interactions that contribute to a highly non-linear behavior. The reactions taking place in the liquid phase have been reported to be [33]:

$$2H_2O \leftrightarrow H_3O^+ + OH^- \tag{2.3}$$

$$2H_2O + CO_2 \leftrightarrow H_3O^+ + HCO_3^- \tag{2.4}$$

$$H_2O + HCO_3^- \leftrightarrow H_3O^+ + CO_3^{2-} \tag{2.5}$$

$$MEAH^+ + H_2O \leftrightarrow H_3O^+ + MEA \tag{2.6}$$

$$MEACOO^- + H_2O \leftrightarrow MEA + HCO_3^- \tag{2.7}$$

A sketch of the component distribution in the system of CO_2 chemically absorbed in MEA and H_2O is shown in Figure 2.2. It illustrates the presence of 3 volatile components that are distributed between the two phases and the non-volatile components that correspond to the cations and anions formed in the liquid phase due to the chemical reactions. Note that the water is considered as the diluent of the ions and the volatile components in Figure 2.2.

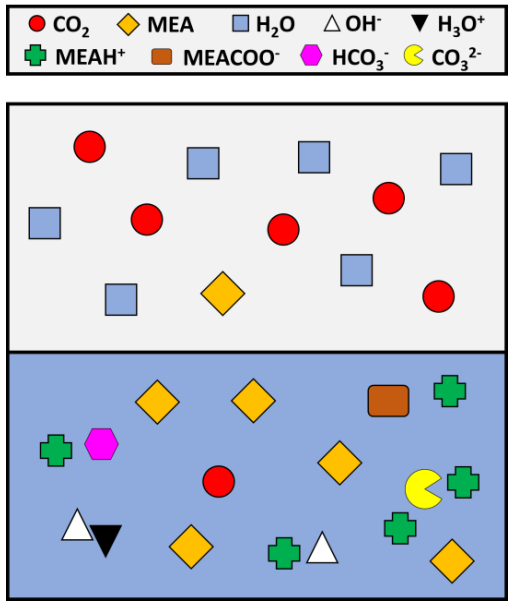


Figure 2.2. Sketch of the modelled thermodynamic system with water as the diluting component of the liquid phase.

2.3.2. Essential Thermodynamic Relations

2.3.2.1. Physical Constraints and Auxiliary Equations

The semi-empirical gamma-phi VLE model that is replicated in this work (also referred to as “base model”) considers the phase equilibria of the volatile components and the speciation reactions in the liquid phase. This gamma-phi model is composed of equations for phase equilibria (Henry’s law and Raoult’s law), chemical equilibria, thermodynamic constraints and complementary auxiliary equations. The complementary auxiliary equations can be either of empirical or semi-empirical nature and have fitted parameters (e.g., an activity coefficient model).

The base model estimates the activity coefficients with the eNRTL model [6] along with the Peng-Robinson equation of state for the vapor phase [34]. The binary

parameters of the eNRTL activity coefficient model, the equilibrium constant correlations and the implementation of the semi-empirical model were taken from a previously published paper [35].

Since machine learning models use alternate pathways to compute the numerical values of the VLE, there is no certainty that these estimations comply with thermodynamic rules. Therefore, it is mandatory to ensure that the calculated thermodynamic system fulfills physical constraints. The summation constraints for the vapor and liquid phases must be considered in the implementation of machine learning models:

$$\sum_i y_i = 1 \tag{2.8}$$

$$\sum_i x_i = 1 \tag{2.9}$$

The electro-neutrality constraint is a restriction that only exists in electrolytic systems. This equation arises from the principle that the overall sum of local charges must be 0 in a system at thermodynamic equilibrium [36]. This relation is:

$$\sum_i n_i \Gamma_i = 0 \tag{2.10}$$

Here n is the total number of moles of the component and Γ is the ion charge relative to a hydrogen ion. For a more detailed description of the additional fundamental equations and the parameters of the semi-empirical model please refer to the supporting information.

It is customary in the CO_2 capture research field to represent the component concentrations in the liquid phase with non-natural thermodynamic variables: CO_2 loading (α_{CO_2}) and amine weight percent in the liquid solvent on a CO_2 free basis (w_{MEA}). The CO_2 loading is estimated with the following relation:

$$\alpha_{CO_2} = \frac{x_{CO_2,App}}{x_{MEA,App}} \tag{2.11}$$

Where the subscript App refers to apparent and indicates that the molar compositions of CO_2 and MEA are calculated as if they had not reacted in the liquid phase. Therefore, the apparent molar compositions only consider the CO_2, MEA and H_2O molar fractions.

The following equation relates w_{MEA} with the MEA apparent molar fraction on a CO_2 free basis $x^*_{MEA,App}$:

$$x^*_{MEA,App} = \frac{\left(\frac{w_{MEA}}{\bar{M}_{MEA}}\right)}{\left(\frac{w_{MEA}}{\bar{M}_{MEA}} + \frac{100 - w_{MEA}}{\bar{M}_{H_2O}}\right)} \tag{2.12}$$

Where $\bar{M}_{MEA}$ and $\bar{M}_{H_2O}$ are the molecular weights of MEA and H_2O respectively. It is possible to calculate the apparent molar compositions of the liquid phase using Eqs. (2.11) and (2.12).

2.3.2.2. Enthalpy of Phase Change

The CO_2 enthalpy of phase change (usually referred to as CO_2 enthalpy of absorption) is of fundamental importance in the field of acid—gas treating because it is one of the main characteristics that determines the techno-economic potential of the technology

[37,38]. It is this thermodynamic quantity that is needed to estimate the required heat duty of the solvent regeneration system in the process.

As opposed to non-supercritical fluids (MEA-H_2O), the enthalpy of vaporization of permanent gases (CO_2) cannot be measured in pure state. Therefore, it is a function of the solvent into which the CO_2 is absorbed. Moreover, experimentally it is not possible to measure the enthalpy of vaporization and the enthalpies of the individual reactions separately. Therefore, the overall effect is commonly measured and reported as a single CO_2 enthalpy of absorption [39,40].

An alternative way to obtain the heat of phase change from VLE data is by using an expression derived from the Gibbs-Helmholtz relation (the vant' Hoff equation):

$$\left(\frac{\partial\, ln\left(P_{CO_2}\right)}{\partial\left(\frac{1}{T}\right)}\right)_{P,n} = -\frac{\Delta H_{CO_2}}{R} \tag{2.13}$$

Here P_{CO_2} is the partial pressure of CO_2 in the vapor phase and ΔH_{CO_2} is the CO_2 enthalpy of absorption. For an in-depth discussion of the derivation and the inherent assumptions to obtain Eq. (2.13) see [4].

2.3.2.3. Gibbs Phase Rule Analysis

The Gibbs phase rule for reactive systems provides an information about how many degrees of freedom exist in a closed reactive thermodynamic system. The degrees of freedom refer to the number of independent variables that can be simultaneously set in order to fully specify the state of the thermodynamic system.

The equation for the Gibbs phase rule for reactive systems is:

$$F = 2 + N - \pi - r - s \tag{2.14}$$

Where *F* is the number of degrees of freedom, *N* is the number of components, π is the number of phases, *r* is the number of reactions and *s* is the number of non-summation constraints. It is imperative to follow the Gibbs phase rule in the development of machine learning models. If this rule is broken, thermodynamic quantities cannot be estimated with thermodynamic consistency.

2.3.3. Developing Surrogate Thermodynamic Models

The principles of the proposed methodology to formulate surrogate machine learning VLE models is summarized in Figure 2.3. Figure 2.4 presents detailed the algorithms of the steps 4,5 and 7 shown in Figure 2.3.

Step 1: The semi-empirical model to be replicated is chosen. As previously mentioned, an eNRTL thermodynamic model for CO_2-MEA-H_2O is utilized in this work as the case study to exemplify the use of the proposed methodology.

Step 2: The number of independent variables that can be selected in the input layer of the ANN is calculated. In the current example of the CO_2-MEA-H_2O mixture, the evaluation of Eq. (2.14) yields ***F = 3***, which designates that only 3 variables can be selected.

Step 3: The parameters needed to perform the next steps of the algorithm are set. The ANN training parameters are independent variables, dependent variables, limits of the independent variables, architecture of the ANN, simulations to parameters ratio, sampling method, training algorithm, preprocessing function, loss function, accuracy target.

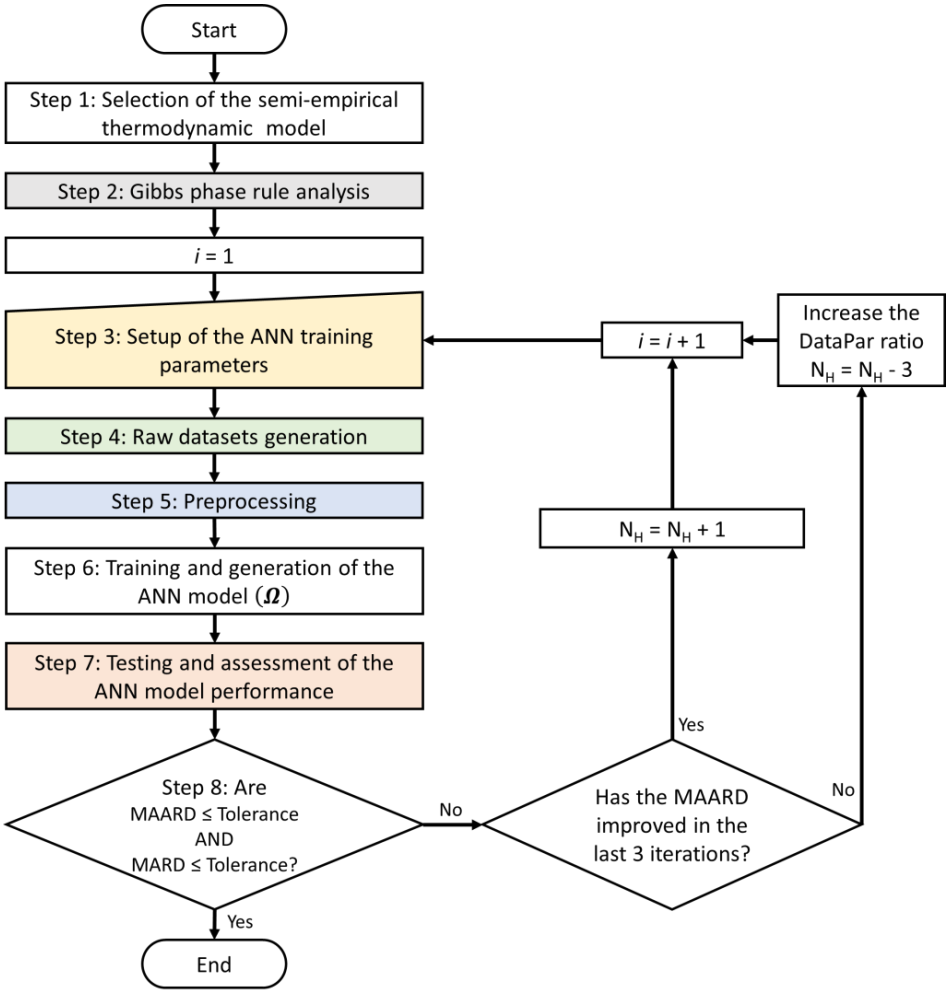


Figure 2.3. Algorithm for the development of a surrogate thermodynamic machine learning based model.

Independent variables: any set of independent variables can be chosen as long as the number of independent variables is equal to the degrees of freedom and the variables are independent of each other (i.e. all the apparent molar fractions of CO_2, MEA and H_2O cannot be chosen simultaneously due to the summation constraint). The fact that ANN can estimate the relationships between different numerical values provide the possibility to select any non-natural thermodynamic variable as independent variable (i.e. CO_2 loading can be selected as independent variable). The independent variables selected in the presented case study are α_{CO_2}, x_{MEA} and temperature (T).

Dependent variables: any number of dependent variables may be chosen. In addition, any variable can be chosen as a dependent variable as long as it is a function of at least one independent variable. In the present case study, there are 13 dependent variables: 9 liquid molar fractions (x_i), 3 vapor molar fractions (y_i) and the total pressure (P).

Limits of the independent variables: the upper and lower boundaries in which the surrogate model is valid are set for each one of the independent variables. The limits set for the surrogate ANN model in this work are presented in Table 2.1. They were specified so that the machine learning model can be used for applications under common industrial operation conditions in a CO_2 amine-scrubbing plant.

Architecture of the ANN: the number of hidden layers and the number of neurons in the hidden layer must be chosen. In this work a feedforward neural network (FFNN) was chosen as it is the simplest ANN architecture without internal cycles or loops. Therefore, it is expected that the lack of recursive operations will create the most computationally efficient ANN-based surrogate models.

Table 2.1. Limits of the independent variables for the case study.

Variable	Min	Max
$\alpha_{CO_2-MEA}/(mol\ CO_2/\ mol\ MEA)$	0.001	0.60
$MEA\ wt/\%$	0.1	0.60
T/K	293.15	393.15

Additionally, the number of hidden layers for all ANN was fixed to 1 because a higher number of hidden layers is used for more complex purposes where deep learning is specifically needed, for example for image pattern recognition [41] or in the development of advanced artificial intelligence for intricate decision-making models for the video-game industry [42]. Furthermore, there is no need to choose more complex architectures according to the universal approximation theorem that states that any continuous function can be approximated using a single hidden layer feed-forward ANN with a finite number of hidden neurons [43]. The effect of the number of neurons in the hidden layer on the ANN model performance is analyzed and discussed in section 3.1.

Simulations to parameters ratio: the number of simulations present in the training dataset is selected. Since the overall idea of our proposed methodology is to use data generated by a semi-empirical model to train an ANN, the number of data points generated by the base model is an important training variable.

The simulations to parameters ratio is given by:

$$DataPar = \frac{N_S}{N_P} \tag{2.15}$$

Here $DataPar$ is the simulations to parameters ratio, N_S is the number of datapoints in the datasets and N_P is the number of parameters in the ANN. The number of parameters in a single layer FFNN is calculated by:

$$N_P = (N_H N_I + N_H N_O) + (N_H + N_I) \tag{2.16}$$

Where N_H is the number of neurons in the hidden layer, N_I is the number of inputs and N_O is the number of outputs in the FFNN. It is important to remark that $DataPar$ must always be greater than 1 so that the ANN training can be an optimization problem. The effect of $DataPar$ on the accuracy of the ANN models is studied in section 3.1.

Sampling method: the values of the independent variables in the raw matrix must be generated with a sampling method. Three different sampling methods are analyzed in this manuscript: random, structured and combined. The random method generates the input raw matrix with a Monte Carlo sampling scheme. The structured method generates permuted vectors that form an evenly distributed grid. The combined scheme divides the number of simulations in two, the first half is generated with the random method while the second half with the structured method. All the values generated with the sampling methods must be within the limits shown in Table 7.1.

Training algorithm: selection of the optimization method used to fit ANN models. The present work focuses on the most prominent methods to train FFNN: the Levenberg-Marquardt backpropagation method and the Bayesian regularization method.

The Levenberg-Marquardt (LM) backpropagation algorithm is considered because, according to the Matlab documentation [44], it is the recommended method due to its higher computational speed during the training of feedforward neural networks. The LM method uses a maximum neighborhood search method that is an hybrid between the Taylor series and gradient optimization methods [45]. This method was later adapted and applied to perform the backpropagation training of ANN [46]. The Bayesian Regularization (BR) backpropagation algorithm combines the Bayesian interpolation method developed by [47] with the Levenberg-Marquadt (LM) optimization method [48]. It was chosen because it has been shown in several applications like in biological studies with mice [49] or in the cement industry [50] that the Bayesian regularization may give better generalization properties than the LM, and thereby, better prediction capabilities. The effect of the training algorithm on the ANN performance is discussed in section 3.1.

Preprocessing function: any mathematical function can be used to transform the raw data into preprocessed data. Data normalization is a common preprocessing function that scales the data within a certain range usually from 0 to 1. The normalization function is not shown in Figure 2.3 or Figure 2.4 because Matlab 2019b always includes this step inside the ANN model and therefore it is not necessary to explicitly program it. Caution is advised when using a different programming platform.

Loss function: the loss function assesses the performance of the ANN based on the preprocessed values.

The Matlab framework for performing the ANN training uses the Mean Square Error (MSE) as the optimization function. Nonetheless, in the case of models that are valid over a broad range of molar compositions, using MSE will focus the optimization on the higher values rather than distribute the error evenly throughout the dataset. To

avoid this, the optimization function of the model was changed to Mean Relative Square Error (MRSE) instead of the default mean square error (MSE). The optimization function for MRSE has the form:

$$\min \mathcal{L} = \frac{1}{N_S} \sum \left(\frac{\underline{B}' - \underline{\hat{B}}'}{\underline{B}'} \right)^2 \tag{2.17}$$

Where $\underline{B}'$ is the output matrix generated with the base model and $\underline{\hat{B}}'$ is the output matrix from the ANN model.

Accuracy target: the degree of desired exactness of the surrogate model with respect to the base model must be specified (see step 8).

Step 4: Here, the datapoints are generated with the semi-empirical model. Figure 2.4 a) shows that in the first iteration of the algorithm, the input matrix $\underline{X}^e$ and the output $\underline{Y}^e$ testing matrices are generated. This testing dataset is used in step 7 to determine if the ANN model can properly predict values for which the ANN model was not trained for. If the predictions of the testing dataset are significantly worse than the ones done with the training dataset, it means that the ANN is overfitted and cannot generalize (this usually happens if the DataPar value is too small). Note that a similar algorithm is used to generate the raw input matrix $\underline{A}$ and the raw output $\underline{B}$ matrices used for the ANN training of each one of the ANN models.

Step 5: The raw input matrix $\underline{A}$ and the raw output matrix $\underline{B}$ are preprocessed using transformation functions (see Figure 2.4 b)). In the present work the matrices are transformed as follows: $\underline{A}' = \underline{A}$ and $\underline{B}' = \ln(B)$.

Using the natural logarithm of the output variables helps in addressing one of the main challenges in the modelling of VLE with ANN: the big variance between the orders of

magnitude in the molar fractions of the individual electrolytes. A large variation in the orders of magnitude of the output variables results in an ill-conditioned or badly scaled training dataset which reduces the accuracy of the model predictions.

An example of badly scaled data points can be obtained by comparing the orders of magnitude of the unreacted molar fraction of CO_2 $\left(x_{CO_2}\right)$ of two different thermodynamic systems. The first case considers a thermodynamic system at $\alpha_{CO_2} = 10^{-5}$, $T = 293.15\,K$ and $w_{MEA} = 30\,\%$ which yields molar fraction $x_{CO_2} = 3\,x\,10^{-16}$. This small value is caused by the low CO_2 loading and the exothermic nature of the chemical reactions which promote the product formation. Hence, most of the solubilized CO_2 is chemically bound to the amine. On the other hand, the second scenario considers the system to be at $\alpha_{CO_2} = 0.60$, $T = 393.15$ and $w_{MEA} = 30\,\%$ which leads to an absolutely contrasting value of molar fraction $x_{CO_2} = 3\,x\,10^{-3}$. This is because the high CO_2 loading together with the high temperature significantly decreases the molar fraction of CO_2 that can be chemically bound to MEA and, consequently, increases the amount of CO_2 solubilized by van der Waals forces.

In order to overcome this problem, a scaling through a logarithmic transformation can be performed. An additional advantage of this transformation is that it avoids the calculation of negative molar fractions when any of the molar compositions is close to 0.

Step 6: The ANN weight parameters $\underline{\omega}$ are calculated and the ANN function Ω is generated. This step needs the following ANN training parameters: training algorithm, ANN architecture, number of neurons in the hidden layer and the loss function.

In Matlab 2019b, the training datasets are usually divided into 3 parts: the training fraction, the validation fraction and the test ratio. We selected 0.90 for the training

fraction, 0.1 for the validation fraction and 0 for the testing ratio. The testing ratio is used to assess the ANN generalization capability of the ANN model to predict values that were not used in the ANN training. Since the ANN model testing in this methodology is performed with an independent dataset, the value was set to be a very low value.

Step 7: The model is tested and assessed by evaluating the ANN model Ω using the testing matrix $\underline{X^e}$ as input and then comparing the output predictions with the testing output matrix $\underline{Y^e}$ (see Figure 2.4 c)).

First, the preprocessing transformation function p is applied on $\underline{X^e}$ to obtain $\underline{X'^e}$. Then, the transformed matrix $\underline{X'^e}$ and $\underline{\omega}$ are used to evaluate Ω in order to get $\underline{\hat{Y}''}$. The values of $\underline{\hat{Y}''}$ do not have physical meaning yet because they are in a different mathematical space due to the preprocessing. Hence, a postprocessing procedure must be performed with an anti-transformation function of the form q^{-1} to get physically meaningful values $\underline{\hat{Y}'}$. In this work, the molar compositions of the liquid and vapor phase in $\underline{\hat{Y}'}$ are evaluated afterwards with a compositional normalization function (η) in order to comply with the restrictions of Eqs. (2.8) and (2.9). This process finally calculates $\underline{\hat{Y}}$.

After postprocessing, the accuracy of the model can be evaluated by comparing $\underline{\hat{Y}}$ and the testing output matrix $\underline{Y^e}$. The ANN model capabilities to replicate the base model can be assessed by calculating the "model average absolute relative difference" (MAARD) and "model average relative difference" (MARD) of each one of the variables between the base model and the ANN model.

The equations for the MAARD and the MARD are:

$$AARD = \frac{1}{N_C}\sum_{N}\frac{\left|\hat{Y} - Y^*\right|}{\sqrt{\hat{Y}\,Y^*}} * 100\ \% \tag{2.18}$$

$$ARD = \frac{1}{N_C}\sum_{N_C}\frac{\hat{Y} - Y^*}{\sqrt{\hat{Y}\,Y^*}} * 100\ \%. \tag{2.19}$$

Where N_C is the total number of datapoints in the testing dataset ($N_C = 250{,}000$ in the case study).

Step 8: This decision-based step will stop the algorithm if the accuracy target is reached. In the case that the accuracy target is not reached, the algorithm estimates if the number of neurons in the hidden layer or $DataPar$ should be increased. It should be remarked that although ANN can replicate the base model, one should specify reasonable values (i.e., do not specify very small MAARD).

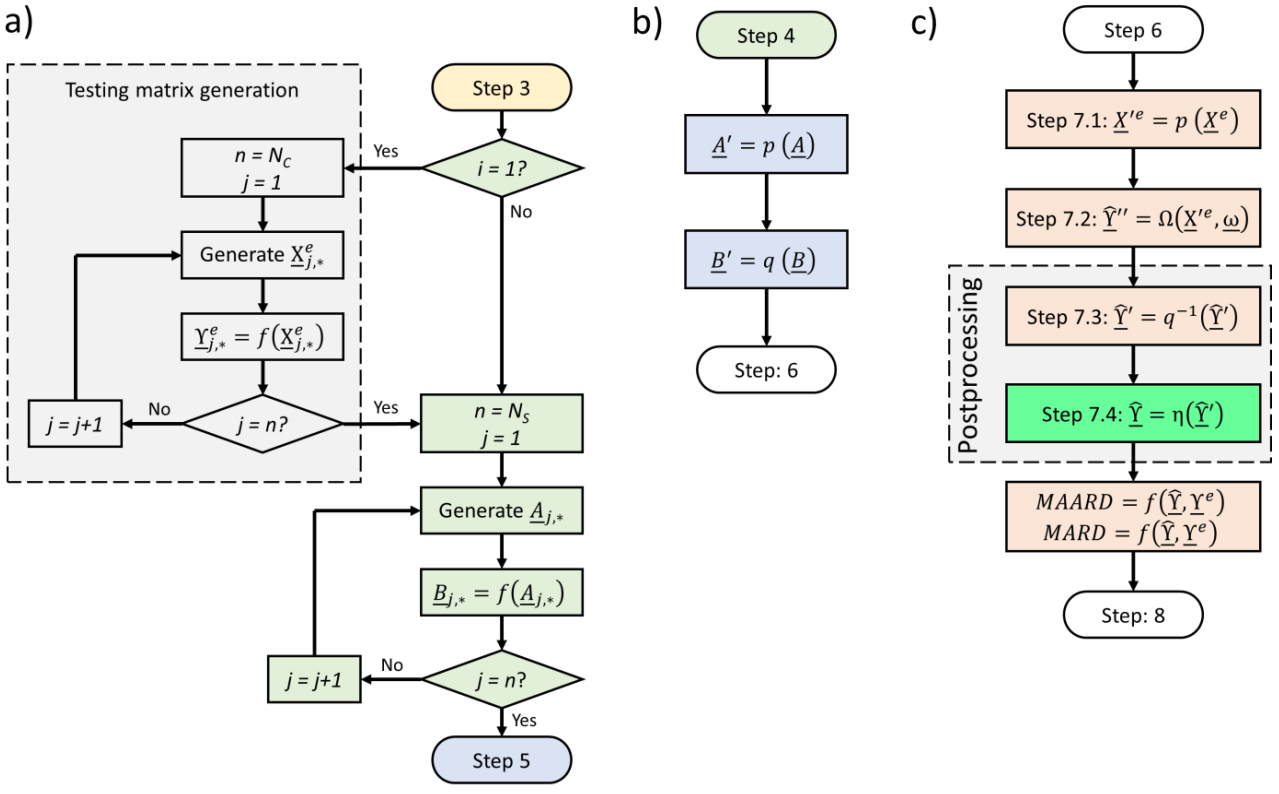


Figure 2.4. Expanded algorithms of: a) step 4: raw datasets generation, b) step 5: preprocessing and c) step 7: testing and assessment of the ANN model performance.

2.4. Results and Discussion

2.4.1. Surrogate Model Parametrization

Since the number of ANN training parameters is somewhat extensive, a factorial-based study was done to determine the best parameter selection. The analyzed ANN training parameters as well as their factors are presented in Table 2.2. A total of 54 ANN models were developed considering all permutations. These models were generated by only performing the steps from 1 to 7 of the algorithms shown in Figure 2.3 and Figure 2.4. Step 8 was not done because this parametric study aims at understanding the effect of the training parameters on the ANN model performance.

Training method: It can be seen from Table 2.3 that the best models were trained using the BR back propagation method. Moreover, the average MAARD of the models trained with the BR method was 23 % better than the models trained with LM.

Sampling method: The results in Table 2.3 demonstrate that both the combined and random sampling methods provide quite similar results and there is not a clear trend on which one is better. The MAARD of the random sampling ANN models have, on average, a MAARD of 3.78 % while for the combined method it is 4.72 %.

Architecture of the ANN: The effect of the number of neurons in the hidden layer is clear, as the best models have 50 neurons. This is expected as the number of neurons and number of fitting parameters increase together. The average MAARD of all the models with 10, 30 and 50 neurons are 13.2 %, 5.6 % and 4.2 % respectively, which confirms that a higher number of neurons increase the model accuracy.

Simulations to parameters ratio: As seen in Table 2.3, a higher DataPar has a higher chance of having higher accuracy. For instance, models #35, #33 and #31 have the same training parameters except for the DataPar and it is seen that the model with

the highest DataPar has the best MAARD. It is important to note that the DataPar value does not have a significant impact in the computational performance of the surrogate model as it is only used in the ANN training. However, choosing a large DataPar (e.g. 100) may significantly increase the training time and, as seen in Table 2.3, the MAARD improvement is so minimal that it may not justify the considerable additional training time.

According to this study, we recommend using the Bayesian Regularization training method and the random sampling method to develop thermodynamic ANN surrogate models for new systems. However, the optimal number of neurons in the hidden layer and DataPar may vary from system to system. Therefore, we suggest starting the procedure shown in Figure 2.3 with $N_H = 10$ neurons and $DataPar = 2$.

Table 2.2. ANN training parameter values.

Variable	Values
Number of neurons in the hidden layer (N_H)	10 / 30 / 50
Simulations to parameters ratio ($DataPar$)	2 / 3 / 5
Training method	BR / LM
Sampling method	R / S / C

* In sampling method: R, S and C stand for random, structured, and combined respectively

Table 2.3. ANN model performances using different training parameters.

ANN Model	# Neurons in the hidden layer	DataPar	Training method	Sampling method	Mean MAARD (%)	Mean MARD (%)
35	50	5	BR	Random	0.50	0.03
33	50	3	BR	Random	0.52	0.04
13	50	2	BR	Combined	0.52	-0.01
15	50	3	BR	Combined	0.54	0.00
17	50	5	BR	Combined	0.57	-0.01
31	50	2	BR	Random	0.57	0.09
36	50	5	LM	Random	0.63	0.11
16	50	3	LM	Combined	0.66	-0.02
18	50	5	LM	Combined	0.67	0.00
34	50	3	LM	Random	0.80	0.17

2.4.2. Surrogate Model Assessment

2.4.2.1. Statistical and Graphical Analysis

In this section, the application and validation of the model #35 is performed (the parameters can be obtained from the authors upon request). This model was chosen because it is ranked as the best model and because it has a MAARD of 0.50 %. The MAARD and the experimental AARD (EAARD) of model #35 and the eNRTL model are presented in Table 2.4. All fitted variables in model #35 show good agreement with the base model, as none surpasses a MAARD value of 1.00 %.

The EAARD values were calculated by comparing the model predictions against experimental data: 131 CO_2 partial pressure data points [8], 80 total pressure data points [8], vapor molar fractions data points [51] and 16 liquid phase speciation data points [52]. The operating conditions in the cited references that were outside the validity range of the ANN models were omitted. The MAARD seems to be insignificant as the difference between the EAARD of the base model and model #35 is negligible.

Parity plots between the base model and the ANN model #35 are presented in Figure 2.5. The 1,000 datapoints used in Figure 2.5 were randomly chosen from the testing dataset. It can be seen in Figure 2.5 a) – b) that the highest residual errors are not the same as the largest relative deviations. The residual errors are higher in the carbamate x_{MEACOO^-} at high loadings because this is when the carbamate concentrations are larger, so even if the relative error is small at these conditions, the residual error will be high.

Table 2.4. Relative deviations between model #35 and the base model and experimental data.

Variable	Model #35 MAARD (%)	Base model EAARD (%)	Model #35 EAARD (%)
x_{H_2O}	0.05	-	-
x_{CO_2}	0.63	-	-
x_{MEA}	0.41	14.6	14.8
$x_{H_3O^+}$	0.86	-	-
x_{MEAH^+}	0.46	28.4	28.6
x_{OH^-}	0.57	-	-
$x_{HCO_3^-}$	0.41	34.6	34.5
$x_{CO_3^{2-}}$	0.67	27.0	26.9
x_{MEACOO^-}	0.35	12.5	12.6
y_{H_2O}	0.39	1.7	1.7
y_{CO_2}	0.79	23.8	23.9
y_{MEA}	0.50	31.5	31.5
P_T	0.41	13.9	14.0
P_{CO_2}	-	20.7	20.8
Overall (%)	0.50	24.4	24.5

Figure 2.5 c) – d) on the other hand presents the vapor molar fraction of CO_2. The residual plot shows that at low loadings, the residuals are small. This is because at these loadings, the CO_2 is mostly absorbed in the liquid phase by chemical reactions, therefore the free CO_2 in the liquid solution does not exert high pressures of CO_2 in the vapor phase, resulting in a small CO_2 molar fraction. This gives low vapor molar fractions and, therefore, the residuals must be small. In contrast, at higher loadings the CO_2 is absorbed in the liquid phase by van der Waals forces as well, thus the pressure exerted by CO_2 on the vapor phase increases drastically, and the CO_2 vapor molar fraction is higher.

When the relative deviations are compared, the error is well distributed along the entire range of CO_2 loadings and temperatures. There is no trend that suggests that the CO_2 loading and temperature have any effect on the differences in CO_2 molar fractions between the ANN model #35 and the base model.

In order to show the VLE prediction accuracy, the total pressure of the CO_2-MEA-H_2O system was calculated at different conditions and compared against experimental data. The results are presented in Figure 2.6 to Figure 2.9, where the continuous lines represent the predictions of model #35. The data reported for total pressure in literature [8] are mainly from medium to high pressures, hence there are few data points at low CO_2 loadings or temperatures. Figure 2.6 shows that the cited experimental data and the model predictions are in good agreement over a broad range of temperatures and CO_2 loadings.

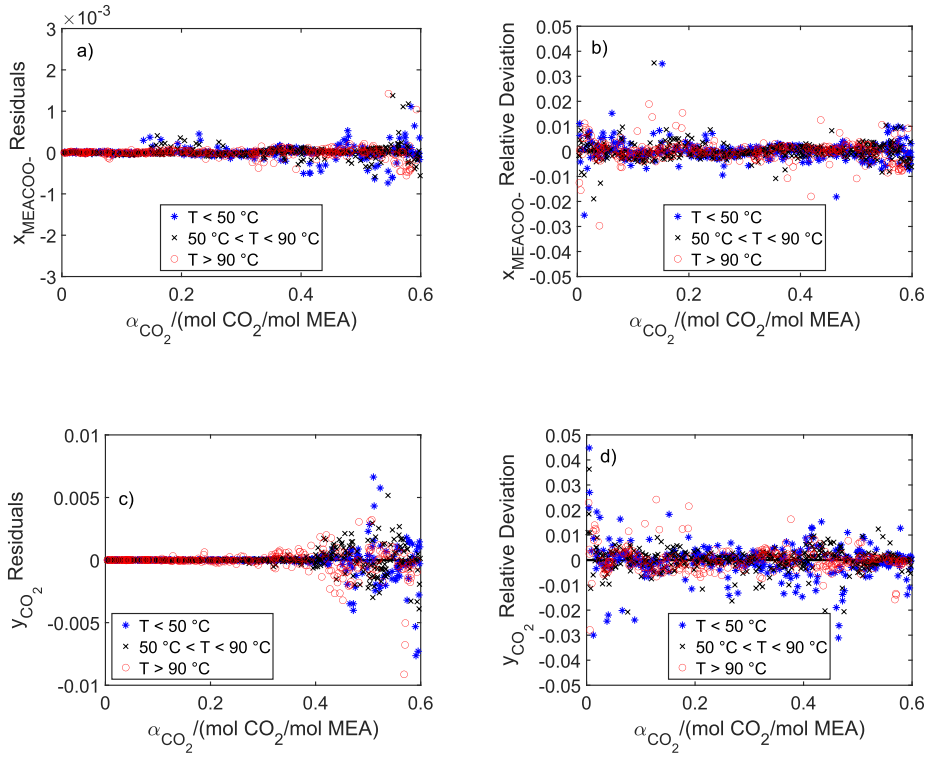


Figure 2.5. Parity plots between the base model and the ANN model #35: a) x_{MEACOO^-} residual error parity plot, b) x_{MEACOO^-} relative deviation plot, c) y_{CO_2} residual parity plot and d) y_{CO_2} relative deviation plot.

The speciation predictions of model #35 were calculated and are presented in Figure 2.7 and compared against experimental speciation data points [52]. Even at low loadings, the smooth behavior of the speciation curves was accurately reproduced by the ANN model. It is important to underline that the estimation of the molar compositions need to be reasonably accurate as the liquid molar compositions are often used in kinetic and mass transfer models.

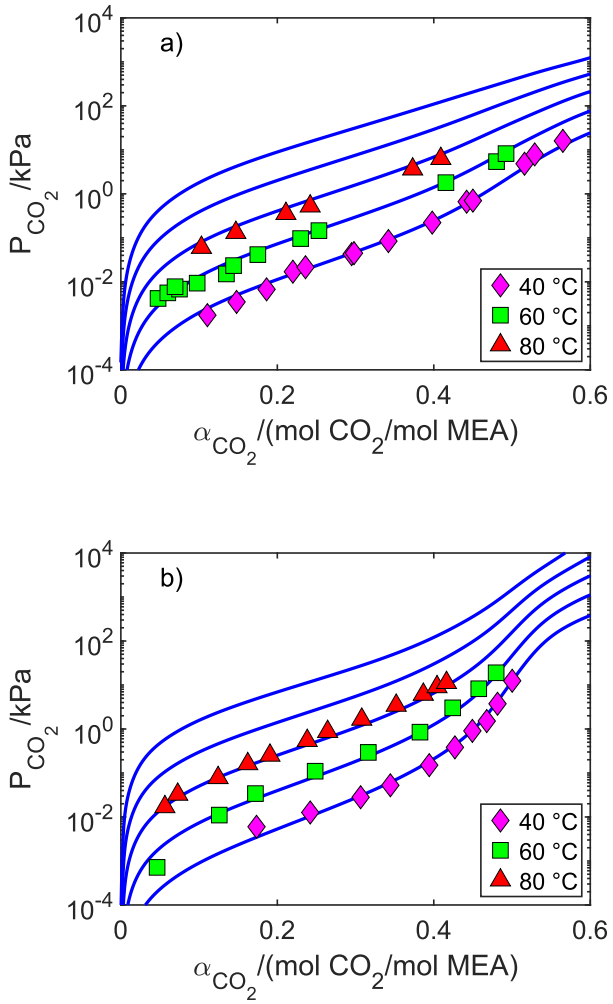


Figure 2.6. Total pressure plots at different MEA weight fractions: a) $w_{MEA} = 15$ % and b) $w_{MEA} = 60$ %. Continuous lines: model prediction with ANN model #35. The top solid lines without experimental datapoints are computed at $T = 100°$ C and $T = 120°$ C (the topmost correspond to the highest T). Experimental data: [8].

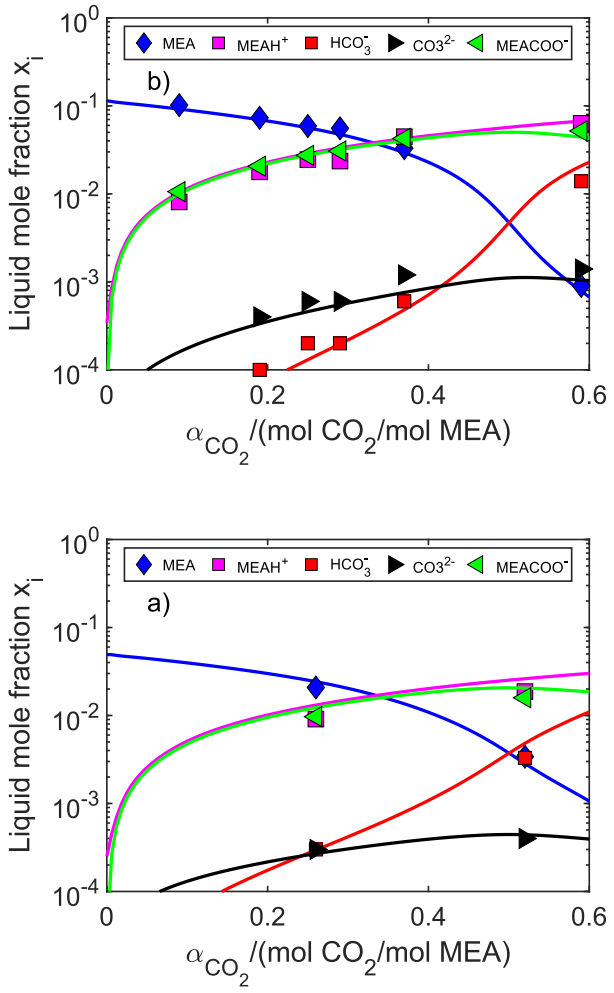


Figure 2.7. Speciation plots at different conditions. a) $w_{MEA} = 30$ % and 40 °C, b) $w_{MEA} =$ 45 % and 20 °C. Continuous lines: model prediction with ANN model #35. Experimental data: [52].

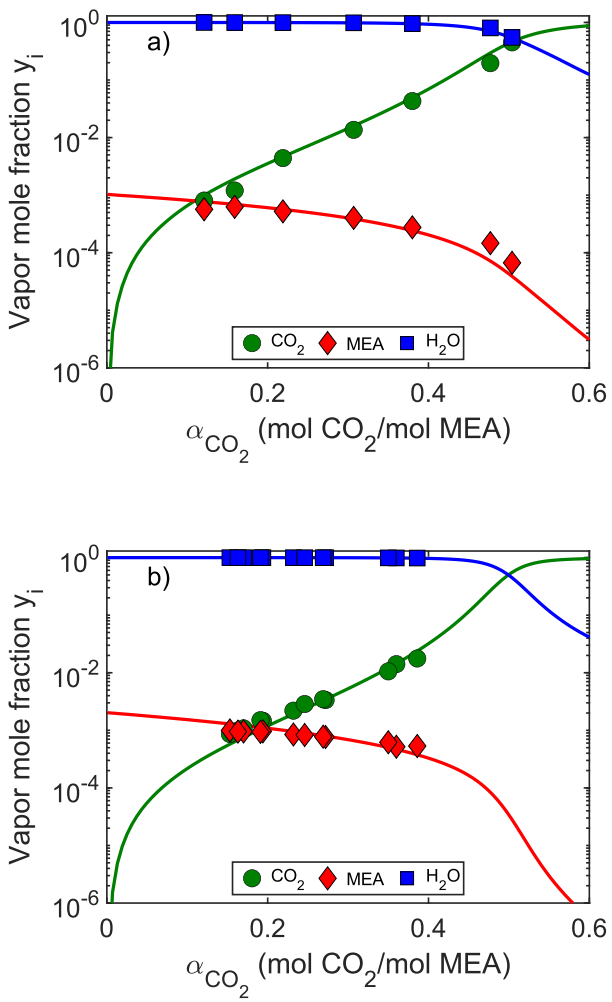
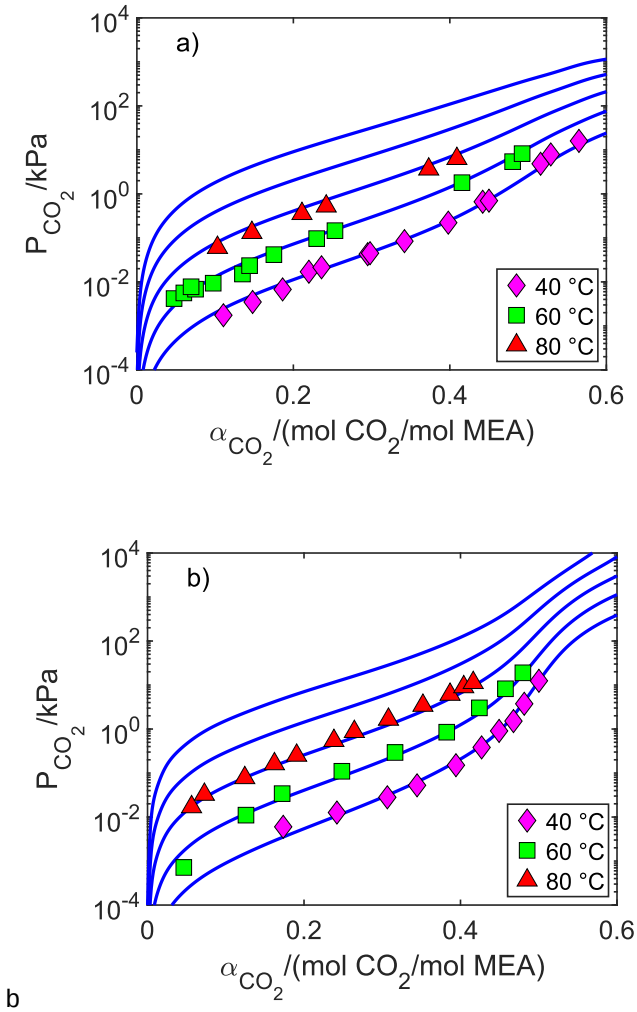


Figure 2.8. Molar fraction plots at a) 3.5 MEA molarity and 60 °C and b) 7.0 MEA molarity and 40 °C. Continuous lines: model prediction with ANN model #35. Experimental data: [51].



b

Figure 2.9. CO_2 partial pressure plots at different w_{MEA}: a) 15 % b) 60 %. Continuous lines: model prediction with ANN model #35. The top solid lines without experimental datapoints are computed at $T = 100°$ C and $T = 120°$ C (the topmost correspond to the highest T). Experimental data: [8].

Although the developed ANN models in this work do not explicitly predict the CO_2 partial pressure P_{CO_2}, it can be calculated by multiplying the total pressure and the

CO_2 vapor phase molar fraction. Since the ANN model #35 accurately estimates both variables independently, the calculated P_{CO_2} is in good agreement with the experimental data as shown in Figure 2.9. This indicates that the surrogate model is not only accurate, but also represents the thermodynamic system in the same fashion as calculations done with traditional VLE models.

Figure 2.10. presents a relative deviation plot for predicted P_{CO_2} and P_T as a function of α_{CO_2} compared with to experimental data. Figure 2.10. shows four different sets of relative deviations between the predictions of the base model and model #35 and the experimental values of P_{CO_2} and P_T. It is seen that model #35 and the base model agree very well with each other, as almost all datapoints in Figure 2.10 calculated by the base model are covered by the estimations of model #35. Additionally, the relative deviation of model #35 with respect to the experimental data is well distributed and there is no sign of bias.

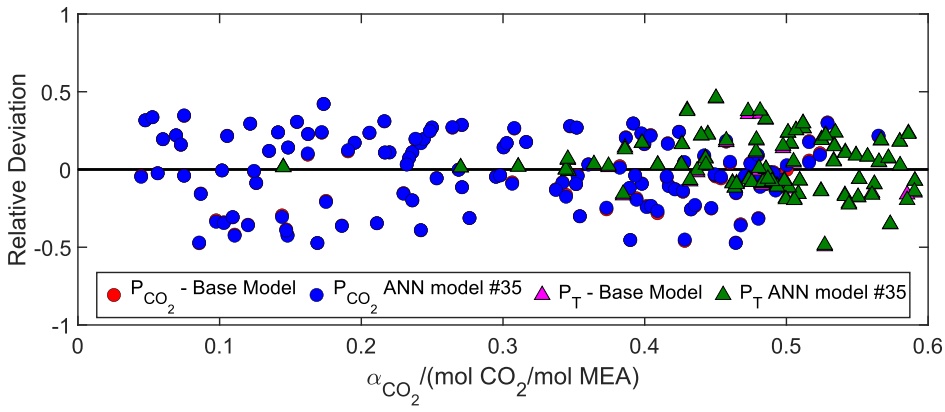


Figure 2.10. Relative deviation plot between the experimental values and the predicted values from the ANN model #35 and the base model. Experimental data: [8].

The predictions of the base model and model #35 were plotted together so their differences could be illustrated. For instance, if only the datapoint of the ANN model #35 is seen, it is an indication that both models predicted the same values. Few datapoints show both model datapoints, whereby it can be concluded that the base model was accurately replicated by model #35.

2.4.2.2. Thermodynamic Constraints

The black-box nature of ANN models may draw skepticism on the physical validity of the model, which in turn, may inhibit their implementation into process engineering frameworks. The ANN models are said to be of black-box nature because, to this day, it is not possible to deduce a physical meaning from the ANN parameters alone. Additionally, the difference between the semi-empirical model and the ANN models is that the first model considers the electronegativity constraint and the summation constraints in the solution algorithm while the ANN models do not (the summation equation restrictions are used in the postprocessing calculations but not in the neural network itself).

Table 2.5 presents the mean difference between the values calculated by the ANN model #35 with respect to Eqs. (2.8) - (2.10) without performing a compositional normalization. The second column presents the absolute mean error when the molar fractions are compositionally normalized. Note that the absolute error of Eqs. (2.8) and (2.9) is 0 on the second column because the normalization redistributes the molar fractions so that their sum is 0.

Table 2.5 shows that although there is a small error in the summation constraints, the fact that the absolute error in Eq. (2.10) is the same between the second and third column. This indicates that the error in the molar fractions is well distributed between

all the species. The base model fulfills the electronegativity constraint within error less than 10^{-15} while the error in the summation constraints is 0.

Table 2.5. Thermodynamic constraints check of ANN model #35.

Eq.	Absolute error without normalization - 10^5	Absolute error with normalization - 10^5
(2.8)	2.34	0
(2.9)	8.19	0
(2.10)	12.6	12.6

2.4.2.3. Why Does the Gibbs Phase Rule Matter?

Complying with the Gibbs phase rule is crucial, regardless of whether if the model is either semi-empirical or machine learning based. For example, let us consider an over specified system where there are more fixed variables than degrees of freedom. In this case, the thermodynamic system will not have any physical meaning as it cannot exist at the specified conditions. The mathematical effect of overspecification in a thermodynamic system depends on the model type. If one tries to estimate the VLE behavior of a thermodynamic system with a semi-empirical model, a solution cannot be achieved as the mathematical model will be mathematically inconsistent. Unfortunately, an ANN does not have a similar “safe-lock” and it can predict VLE values that may seem reasonable at first glance, but do not have a physical meaning.

Consistent thermodynamic models allow evaluating thermodynamic quantities that are related to the VLE behavior through fundamental equations. In the context of CO_2

capture, an important solvent quantity is the CO_2 enthalpy of absorption which can be estimated using Eq. (2.13) and the VLE model.

In order to highlight the importance of the Gibbs phase rule and thermodynamic consistency in the ANN model development, an inconsistent model of the VLE system was formulated following the methodology previously presented but omitting the Gibbs phase rule analysis. The inconsistency was imposed on the model by specifying one extra variable and breaking the Gibbs phase rule. The "inconsistent model" was trained with 4 inputs: α_{CO_2}, MEA wt%, T and P_T. The model was developed and compared with the testing dataset, and a MAARD of 0.37 % was obtained. The P_T values of the training dataset were given as inputs in the "validation" of the inconsistent model. Hence, the MAARD seems to be low and in agreement with the base model.

Figure 2.11 and Figure 2.12 show the CO_2 enthalpy of absorption on the left axis and the P_{CO_2} on the right axis, both as functions of α_{CO_2}. This was done demonstrate that a low MAARD is not necessarily a satisfactory indication that the model was "properly understood" by the ANN.

Figure 2.11 presents the thermodynamic system at 40 °C, it shows that model #35 accurately reproduced the P_{CO_2} behavior of the base model as the curves of both models are overlapping completely. The heat of absorption predictions between the two models are in excellent agreement (MAARD = 0.30 %). The difference between the experimental and the estimated CO_2 heat of absorption is caused by the inherent simplifications to the use of Eq. (2.13) [4,40,53].

Figure 2.12 on the other hand, presents the predictions of the inconsistent model and the base model at: $w_{MEA} = 30$ %, 120 °C and different loadings and pressures. The

heat of absorption predictions with the inconsistent model are inaccurate (relative deviations up to 50 %). This is a clear indication of a model inconsistency. An ANN model may look “good” in a statistical analysis even if it was with a wrong conception of the system physics. The inconsistent ANN model created a fallacious dependency of P_{CO_2} on the total pressure even though it is thermodynamically proven that, for this case, it should only be a function of α_{CO_2}, w_{MEA}, and T. This demonstrates that the misuse of machine learning or artificial neural networks can lead to correlation-causation fallacies.

If the P_{CO_2} predictions of the inconsistent model are observed, one would expect that the CO_2 heat of absorption estimations would be accurate as well, but they are not. This might raise the question: why are the heat of absorption values so inaccurate? This can be explained by considering the similitudes between the ANN and biological organisms. Much of the knowledge of animals or humans is based on experience or correlations. They learn to react accordingly to environmental stimuli by creating a cause-effect correlation that does not necessarily indicate that the living being has understood the root cause of the phenomenon. A similar situation occurs when training an ANN. During the process, the ANN learns and finds the correlations between the variables, but that does not mean that it found right dependencies of the system, it just found a mathematical correlation.

Another situation in which thermodynamic inconsistencies in ML models may occur is if there is an interdependence among the input variables. Let us consider a situation where the ANN model was developed using the CO_2 molarity concentration, T and w_{MEA} as inputs. While the use of three input variables complies with the degrees of freedom of the system, the interdependency between the molar concentration and the temperature will produce thermodynamic inconsistencies if a physically unfeasible

combination of input variable values is chosen. Once again, this may not be detectable by the ANN as the VLE predictions do not directly solve the thermodynamic equations. For this reason, concentrations on a molarity basis are not recommended and it is preferable to use concentrations on a molal basis or a molar fraction basis.

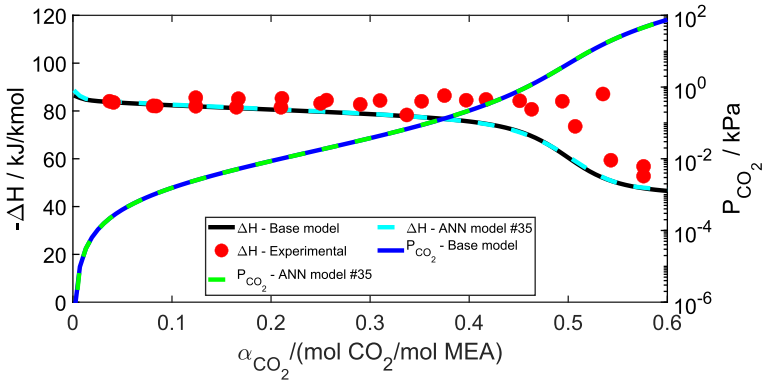


Figure 2.11. Enthalpy of absorption calculated as a function of CO_2 loading using 30 MEA wt% at 40 °C. Experimental data: [54].

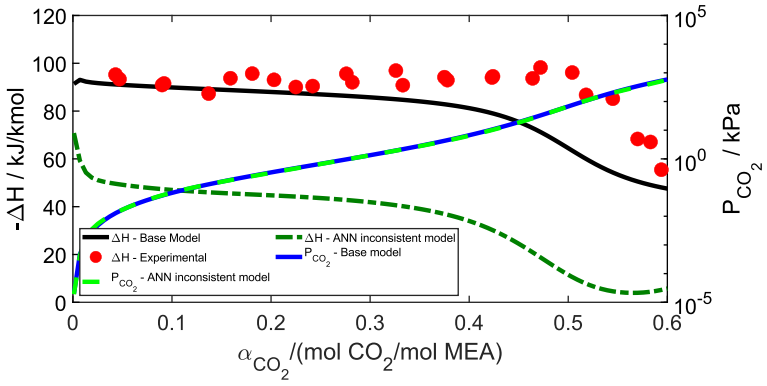


Figure 2.12. Enthalpy of absorption calculated as a function of CO_2 loading using 30 MEA wt% at 80 °C. Experimental data: [54].

2.4.3. Surrogate Model Application

Since the goal of developing the surrogate model is to improve the computational time, the computational speed of the different calculation methods was compared. Table 2.6 presents the comparison of the computational speed of the gamma-phi calculations, different interpolation methods and the ANN surrogate model method proposed in this work. The molar fraction of CO_2 in the vapor phase was chosen as the output variable. A total of 100 batches with 50 simulations each were run in order to estimate the relative computational speed with respect to the gamma-phi calculations shown in Table 2.6. The values of each batch are different between each other and were generated randomly.

The interpolations were done using the Matlab built-in function interp3. Two interpolating algorithms were chosen: the linear interpolation and the cubic spline interpolation algorithm [55]. The best models with 10, 30 and 50 neurons in the hidden layer are also presented in Table 2.6. This table shows that the linear interpolation can accelerate the computational speed. However, its intrinsic relative error is high and the savings in computational time may not be high enough to compensate and justify the error.

Cubic spline interpolation has better accuracy when compared to the linear interpolation, but at the cost of lower computational speed. Making a balance between the relative computational speed and the MAARD, cubic spline interpolation appears to outperform the linear interpolation. It is important to underline, that due to the algorithm of the cubic spline interpolation, it may happen that the algorithm estimates values of y_{CO_2} less than 0 whenever there is a steep change in y_{CO_2} when the value is close to infinite dilution.

Table 2.6. Computational speed comparison between different VLE calculation schemes.

	Elements in the interpolation matrix	Relative computational speed	y_{CO_2} MAARD (%)
Gamma-phi calculations (base model)	-	1	-
Linear interpolation	1000	10	28.17
	8000	7	8.78
	27000	6	4.67
Cubic spline interpolation	1000	8	9.20
	8000	6	1.74
	27000	5	0.93
ANN with 10 neurons	-	1202	8.03
ANN with 30 neurons	-	987	1.12
ANN with 50 neurons	-	966	0.49

The machine learning models clearly outperforms the interpolation in both the computational speed and the capabilities for estimation accuracy. The main reason of the computational speed superiority is that the ANN calculations are simple and non-recursive calculations. Another reason for the machine learning outstanding computational characteristics is that the ANN structure allows calculating the entire input matrix in parallel (in this example there were 50 simulations in parallel) as

opposed to the interpolation algorithms that must process each one of the inputs individually.

The last three rows of Table 2.6 suggest that there is a tradeoff between the number of hidden neurons and the computational speed. Figure 2.13 presents how the speed of the ANN models is affected by the number of parallel simulations and the number of hidden neurons. The tradeoff between the number of neurons and the computational speed is evident as the model with 10 neurons has the fastest relative speed, however its MAARD is somewhat high, thus the models with 30 and 50 hidden neurons seem to be better. A higher number of neurons implies more operations inside an ANN. Therefore, it is important to set a reasonable target of the MAARD in order to avoid a large number of neurons that would reduce the computational speed.

From Figure 2.13 a) it is evident that the model with 50 neurons is around 3 orders of magnitude faster than the gamma-phi calculations when performed with more than 100 simulations done in parallel. The computational gain at smaller number of parallel simulations is unclear, hence, an amplification was done and shown in Figure 2.13 b). The minimum relative speed is reached for all models when there is only one simulation. However, even at this minimum point the relative speed of all 3 models is 40 times the speed of the base model. At 30 parallel simulations the computational gain is around 3 orders of magnitude, while for 400 simultaneous calculations, the speed is around 2,500 times faster if the model with 50 neurons is used.

So, one may wonder: is it really important to speed-up the VLE computation time? If the surrogate model is to be used to model the behavior of a single vapor-liquid separator, the calculations would take 8.0×10^{-5} s instead of 3.2×10^{-3} s. Looking at this scenario, from the end user perspective there is not a noticeable difference that justifies the use of a surrogate model. But what happens if there is a complex problem

that requires a non-steady state 2-D formulation? Let us assume that the geometry uses a 400 x 400 grid that requires a VLE evaluation in each discretized point. Furthermore, the time is discretized in 400 timesteps and it is assumed that each timestep requires 10 iterations to solve. Considering these conditions, the VLE simulations would take 400 x 400 x 100 x 10 x 3.2 x 10^{-3} s = 510,000 seconds (142 hours) if the base model is used. In contrast, the simulation would take 0.15 hours to solve if the surrogate model with 50 neurons is used (this scenario assumes 400 simultaneous VLE simulations). The use of ANN surrogate models instead of performing the traditional semi-empirical models may thus give extraordinary computational advantages in complex problems that require VLE calculations in multidimensional problems. One of such examples may be the solution of the model needed for the characterization of aerosol emissions from CO_2 capture plants [56].

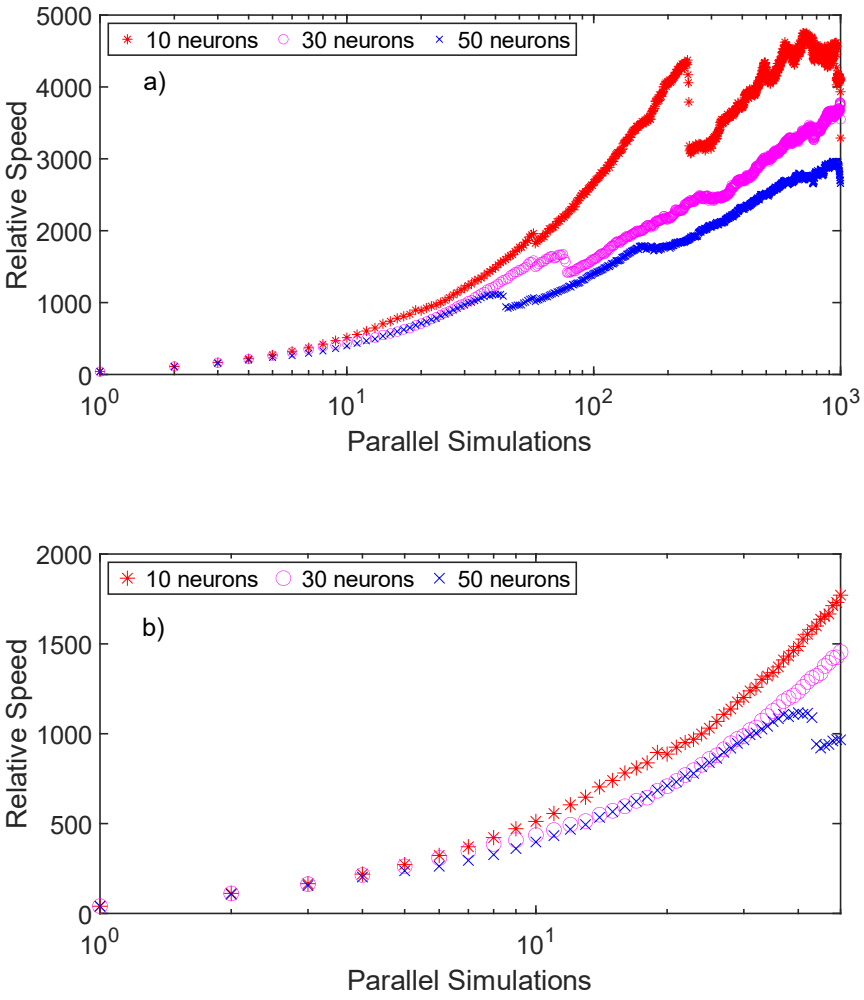


Figure 2.13. a) Computational efficiency tradeoff between the number of neurons and the number of parallel simulations. b) Amplification of a).

2.5. Conclusions

An easy-to-implement method based on machine learning Artificial Neural Networks (ANN) was developed and proved to be a feasible alternative for the development of accurate, consistent and computationally fast surrogate models. The proposed method was employed to develop a surrogate machine learning thermodynamic model of a ternary system CO_2-MEA-H_2O. The surrogate models were based on a semi-empirical gamma-phi model framework (eNRTL for the liquid phase and Peng Robison for the vapor phase). The advantage of the proposed method is that it can be easily extrapolated to other thermodynamic quantities (e.g. activity coefficients, enthalpy calculations or heat capacities) and systems (e.g. more components or more phases).

A quantitative assessment of the effect of the training parameters on the prediction capabilities of the ANN models was performed. The first conclusion from this study is that a single hidden-layer FFNN architecture is enough to represent the behavior of reactive multiphase systems. Therefore, the need for more complex architectures (e.g., multiple hidden-layers or feedback) may be unnecessary as they may introduce extra training parameters or iteration loops that may make the ML model implementation difficult without significantly improving the prediction capabilities.

Through a statistical analysis, it was found that the accuracy of the surrogate models was improved using Bayesian regularization back-propagation algorithm and a random sampling method to generate the training datasets. The DataPar value is suggested to be as low as possible ($DataPar = 2$ for similar applications). The analysis involved the development of 54 ANN models wherefrom the one with the best prediction capabilities with respect to the semi-empirical model has a $MAARD = 0.50\ \%$.

It was shown that the integration of the Gibbs phase rule and physical constraints in the ANN framework is of utmost importance as their inclusion allows avoiding thermodynamic inconsistencies that may lead to inaccurate predictions.

It was also demonstrated that good prediction capabilities in ANN models are not necessarily a satisfactory indication of thermodynamic consistency, proper dependency between the variables or that the model complies with the physical constraints. Therefore, the integration of the Gibbs phase rule and physical constraints in the ANN model is a viable method to ensure that the thermodynamic model has the same behavior and dependencies as the semi-empirical models.

The developed surrogate ANN models can be around ~1000 times faster than rigorous calculation methods because the ANN models do not have recursive operations that jeopardize the computational efficiency of the calculations. Additionally, the computational speed of ANN surrogate models outperforms the ones of linear and non-linear interpolation methods.

2.6. Supplementary information

2.6.1. Fundamental equations of the base model

The Henry's relation for the physical solubility of CO_2 is

$$y_i \hat{\varphi}_i P = x_i \gamma_i H_i^{\infty} \exp\left(\frac{\bar{v}_i^{\infty}(P - P_i^{sat})}{RT}\right). \quad (2.20)$$

Where the subscript i indicates that the subscripted variable is related to component i, the superscript sat indicates that the variable is at its saturation point, y is the vapor molar fraction, $\hat{\varphi}$ is the vapor fugacity coefficient, P is the total pressure of the system, x is the liquid molar fraction, γ is the liquid activity coefficient, T is the temperature, R is the gas constant and finally, H_i^{∞} and $\bar{v}_i^{\infty}$ are the Henry's coefficient and partial molar volume of component i at infinite dilution in water [57]. For the case of the other volatile components (MEA and H_2O), the modified Raoult's law is used:

$$y_i \hat{\varphi}_i P = x_i \gamma_i P_i^{sat} \exp\left(\frac{V_i(P - P_i^{sat})}{RT}\right). \quad (2.21)$$

Where V is the pure component liquid molar volume. The general equation governing the chemical equilibria in terms of molar fractions and activity coefficients is given by:

$$K_j = \prod_i (x_i \gamma_i)^{\nu_i}. \quad (2.22)$$

Where K is the chemical equilibrium constant, the subscript j indicates the reaction number and ν is the stoichiometric coefficient according to the reaction set given in section 2.1 of the manuscript. The right side of Eq. (2.22) is the product of the volatile and electrolyte components activities in each one of the reactions. The activity coefficients must be calculated using a usually computationally intensive set of

equations that consider the contribution of the molecule-molecule, ion-molecule and ion-ion binary interactions [58].

An extra equation must be added to electrolyte systems: the electro-neutrality constraint, it arises from the principle that the overall sum of the local charges must be 0 in a system at thermodynamic equilibrium [36]. This relation is:

$$\sum_i n_i \Gamma_i = 0. \tag{2.23}$$

Where n is the total number of moles of the component and Γ is the ion charge relative to a hydrogen ion. Note that only ionic species can be evaluated with Eq. (2.10). Since the equations are written in terms of molar compositions, the summation constraint for the vapor and liquid phases respectively must be set

$$\sum_i y_i = 1, \tag{2.24}$$

$$\sum_i x_i = 1. \tag{2.25}$$

The apparent composition approach has been widely used to represent the VLE data of gas-liquid systems due to its convenience in data analysis and representation. Apparent compositions represent the system as if only unreacted species existed in it, therefore only CO_2, MEA and H_2O have apparent compositions. In order to relate the apparent mole compositions with the actual compositions of the reactive system, the following relations must be used:

$$x_{CO_2,App} = \frac{x_{CO_2} + x_{HCO_3} + x_{CO_3^{2-}} + x_{MEACOO^-}}{x_{Total}} \tag{2.26}$$

$$x_{MEA,App} = \frac{x_{MEA} + x_{MEAH^+} + x_{MEACOO^-}}{x_{Total}} \tag{2.27}$$

$$x_{H_2O,App} = \frac{x_{H_2O} + x_{H_30^+} + x_{OH^-} + x_{HCO_3^-} + x_{CO_3^{2-}}}{x_{Total}} \tag{2.28}$$

$$x_{Total} = x_{CO_2} + x_{MEA} + x_{H_2O} + x_{H_30^+} + x_{OH^-} + 2x_{HCO_3^-} + 2x_{CO_3^{2-}} + x_{MEAH^+} + 2x_{MEACOO^-}. \tag{2.29}$$

Where the subscript App, indicates that it is an apparent composition. The denominator (Eq. (2.29)) in Eqs. (2.26) - (2.28) is given by summing the numerators of all equations. Note that one of the Eqs. (2.26) - (2.28) can be substituted if the summation constraint in apparent basis is used.

It is customary in the CO_2 capture field of research to represent the apparent molar compositions in terms of the relative composition of CO_2 with respect to the amine:

$$\alpha_{CO_2} = \frac{x_{CO_2,App}}{x_{MEA,App}}. \tag{2.30}$$

Where α_{CO_2-MEA} is the well-known CO_2 loading. The solvent composition is usually specified in terms of amine weight percent in a CO_2 free basis (w_{MEA}). free basis molar fraction percent for MEA is related to the amine weight percent with:

$$x_{MEA,App} = \frac{\left(\frac{w_{MEA}}{\bar{M}_{MEA}}\right)}{\left(\frac{w_{MEA}}{\bar{M}_{MEA}} + \frac{100 - w_{MEA}}{\bar{M}_{H_2O}}\right)}. \tag{2.31}$$

Where $\bar{M}_{MEA}$ and $\bar{M}_{H_2O}$ are the molecular weights of MEA and H_2O respectively.

2.6.2. Semi-empirical parameters

The Henry's relation for the solubility of CO_2 (Eq. (2.20)) in pure water at infinite dilution is given by:

$$\ln H_{CO_2}^{\infty} = A + \frac{B}{T} + \frac{C}{T^2} + \frac{D}{T^3}. \tag{2.32}$$

While the equilibrium constants for the reactions in the liquid phase have the following form:

$$\ln K_j = A + \frac{B}{T} + C \ln T + DT. \tag{2.33}$$

The values of the Henry's coefficient and equilibrium constants used in the base model are provided in Table 2.7.

Table 2.7. Parameters for the Henry's coefficient and the chemical equilibrium constants.

Reaction:	A	B	C	10^4D	Eq.	Reference
$H_{CO_2}^{\infty}$	-6.8346	12817	-3.76E6	29970	(2.32)	[59]
H_2O diss	132.899	-13455.9	-22.4773	0	(2.33)	[57]
CO_2 diss	231.465	-12092.12	-36.7816	0	(2.33)	[57]
CO_3	216.049	-12431.70	-35.4819	0	(2.33)	[57]
$MEAH^+$	-4.9074	-6166.12	0	-9.84816	(2.33)	[60]
$MEACOO^-$	-4.47	-2354.14	0.168549	77.7289	(2.33)	[35]

2.7. References

[1] V. Masson-Delmotte, P. Zhai, H.O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, W. A. Pirani, Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, T. Waterfield, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate c, 2018.

[2] G.T. Rochelle, Amine Scrubbing for CO2 Capture, Science (80-.). 325 (2009) 1652–1654.

[3] X. Wu, M. Wang, P. Liao, J. Shen, Y. Li, Solvent-based post-combustion CO2 capture for power plants: A critical review and perspective on dynamic modelling, system identification, process control and flexible operation, Appl. Energy. 257 (2020) 113941. https://doi.org/10.1016/j.apenergy.2019.113941.

[4] H.F. Svendsen, E.T. Hessen, T. Mejdell, Carbon dioxide capture by absorption, challenges and possibilities, Chem. Eng. J. 171 (2011) 718–724. https://doi.org/10.1016/j.cej.2011.01.014.

[5] K.S. Pitzer, Thermodynamics of electrolytes. I. Theoretical basis and general equations, J. Phys. Chem. 77 (1973) 268–277. https://doi.org/10.1021/j100621a026.

[6] C. -C Chen, L.B. Evans, A local composition model for the excess Gibbs energy of aqueous electrolyte systems, AIChE J. 32 (1986) 444–454. https://doi.org/10.1002/aic.690320311.

[7] J.O. Lloret, L.F. Vega, F. Llovell, A consistent and transferable thermodynamic model to accurately describe CO2 capture with monoethanolamine, J. CO2 Util. 21 (2017) 521–533. https://doi.org/10.1016/j.jcou.2017.08.018.

[8] U.E. Aronu, S. Gondal, E.T. Hessen, T. Haug-Warberg, A. Hartono, K.A. Hoff, H.F. Svendsen, Solubility of CO2 in 15, 30, 45 and 60 mass% MEA from 40 to 120°C and model representation using the extended UNIQUAC framework, Chem. Eng. Sci. 66 (2011) 6393–6406. https://doi.org/10.1016/j.ces.2011.08.042.

[9] S.K. Dash, A.N. Samanta, S.S. Bandyopadhyay, Solubility of carbon dioxide in aqueous solution of 2-amino-2-methyl-1-propanol and piperazine, Fluid Phase Equilib. 307 (2011) 166–174. https://doi.org/10.1016/j.fluid.2011.05.009.

[10] B.J. Sherman, A.F. Ciftja, G.T. Rochelle, Thermodynamic and mass transfer modeling of carbon dioxide absorption into aqueous 2-piperidineethanol, Chem. Eng. Sci. 153 (2016) 295–307. https://doi.org/10.1016/j.ces.2016.07.019.

[11] A. Hartono, R. Ahmad, M. Usman, N. Asif, H.F. Svendsen, Solubility of CO2 in 0.1M, 1M and 3M of 2-amino-2-methyl-1-propanol (AMP) from 313 to 393K and model representation using the eNRTL framework, Fluid Phase Equilib. 511 (2020). https://doi.org/10.1016/j.fluid.2020.112485.

[12] J. Gabrielsen, M.L. Michelsen, E.H. Stenby, G.M. Kontogeorgis, A model for estimating CO2 solubility in aqueous alkanolamines, Ind. Eng. Chem. Res. 44 (2005) 3348–3354. https://doi.org/10.1021/ie048857i.

[13] P. Brúder, A. Grimstvedt, T. Mejdell, H.F. Svendsen, CO2 capture into aqueous solutions of piperazine activated 2-amino-2-methyl-1-propanol, Chem. Eng. Sci. 66 (2011) 6193–6198. https://doi.org/10.1016/j.ces.2011.08.051.

[14] P. Brúder, K.G. Lauritsen, T. Mejdell, H.F. Svendsen, CO 2 capture into aqueous solutions of 3-methylaminopropylamine activated dimethyl-monoethanolamine, Chem. Eng. Sci. 75 (2012) 28–37. https://doi.org/10.1016/j.ces.2012.03.005.

[15] X. Luo, A. Hartono, S. Hussain, H.F. Svendsen, Mass transfer and kinetics of carbon dioxide absorption into loaded aqueous monoethanolamine solutions, Chem. Eng. Sci. 123 (2015) 57–69. https://doi.org/10.1016/j.ces.2014.10.013.

[16] V. Plesu, J. Bonet, A.E. Bonet-Ruiz, A. Chavarria, P. Iancu, J. Llorens, Surrogate model for carbon dioxide equilibrium absorption using aqueous monoethanolamine, Chem. Eng. Trans. 70 (2018) 919–924. https://doi.org/10.3303/CET1870154.

[17] N. Daneshvar, M.T. Zaafarani Moattar, M. Abedinzadegan Abdi, S. Aber, Carbon dioxide equilibrium absorption in the multi-component systems of CO2 + TIPA + MEA + H2O, CO2 + TIPA + Pz + H2O and CO2 + TIPA + H2O at low CO2 partial pressure: Experimental solubility data, corrosion study and modeling with artificial neural netwo, Sep. Purif. Technol. 37 (2004) 135–147. https://doi.org/10.1016/j.seppur.2003.09.004.

[18] B. Zarenezhad, A. Aminian, Predicting the vapor-liquid equilibrium of carbon dioxide+alkanol systems by using an artificial neural network, Korean J. Chem. Eng. 28 (2011) 1286–1292. https://doi.org/10.1007/s11814-010-0492-0.

[19] D. Bastani, M.E. Hamzehie, F. Davardoost, S. Mazinani, A. Poorbashiri, Prediction of CO2 loading capacity of chemical absorbents using a multi-layer perceptron neural network, Fluid Phase Equilib. 354 (2013) 6–11. https://doi.org/10.1016/j.fluid.2013.05.017.

[20] M.E. Hamzehie, S. Mazinani, F. Davardoost, A. Mokhtare, H. Najibi, B. Van der Bruggen, S. Darvishmanesh, Developing a feed forward multilayer neural network model for prediction of CO2 solubility in blended aqueous amine solutions, J. Nat. Gas Sci. Eng. 21 (2014) 19–25. https://doi.org/10.1016/j.jngse.2014.07.022.

[21] M.E. Hamzehie, M. Fattahi, H. Najibi, B. Van der Bruggen, S. Mazinani, Application of artificial neural networks for estimation of solubility of acid gases (H2S and CO2) in 32 commonly ionic liquid and amine solutions, J. Nat. Gas Sci. Eng. 24 (2015) 106–114. https://doi.org/10.1016/j.jngse.2015.03.014.

[22] N. Ghasemian, M. Kalbasi, G. Pazuki, Experimental Study and Mathematical Modeling of Solubility of CO2 in Water: Application of Artificial Neural Network and Genetic Algorithm, J. Dispers. Sci. Technol. 34 (2013) 347–355. https://doi.org/10.1080/01932691.2012.667293.

[23] M. Mirarab, M. Sharifi, M.A. Ghayyem, F. Mirarab, Prediction of solubility of CO2 in ethanol-[EMIM][Tf2N] ionic liquid mixtures using artificial neural networks based on genetic algorithm, Fluid Phase Equilib. 371 (2014) 6–14. https://doi.org/10.1016/j.fluid.2014.02.030.

[24] S. Norouzbahari, S. Shahhosseini, A. Ghaemi, Modeling of CO2 loading in aqueous solutions of piperazine: Application of an enhanced artificial neural network algorithm, J. Nat. Gas Sci. Eng. 24 (2015) 18–25. https://doi.org/10.1016/j.jngse.2015.03.011.

[25] K. Golzar, H. Modarress, S. Amjad-Iranagh, Evaluation of density, viscosity, surface tension and CO2 solubility for single, binary and ternary aqueous solutions of MDEA, PZ and 12 common ILs by using artificial neural network (ANN) technique, Int.

J. Greenh. Gas Control. 53 (2016) 187–197. https://doi.org/10.1016/j.ijggc.2016.08.008.

[26] S. Garg, A.M. Shariff, M.S. Shaikh, B. Lal, H. Suleman, N. Faiqa, Experimental data, thermodynamic and neural network modeling of CO2 solubility in aqueous sodium salt of l-phenylalanine, J. CO2 Util. 19 (2017) 146–156. https://doi.org/10.1016/j.jcou.2017.03.011.

[27] Z. Zhang, H. Li, H. Chang, Z. Pan, X. Luo, Machine learning predictive framework for CO2 thermodynamic properties in solution, J. CO2 Util. 26 (2018) 152–159. https://doi.org/10.1016/j.jcou.2018.04.025.

[28] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A.E. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey, Heliyon. 4 (2018). https://doi.org/10.1016/j.heliyon.2018.e00938.

[29] S.J. Russel, P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed., Prentice-Hall, Upper Saddle River, New Jersey, 2009.

[30] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Phys. (1943) 113–133.

[31] C.M. Bishop, Pattern Recognition and Machine Learning, Springer Science+Business Media, LLC, 2006.

[32] C. Gonzalez Viejo, D.D. Torrico, F.R. Dunshea, S. Fuentes, Development of Artificial Neural Network Models to Assess Beer Acceptability Based on Sensory Properties Using a Robotic Pourer: A Comparative Model Approach to Achieve an Artificial Intelligence System, Beverages. 5 (2019) 33. https://doi.org/10.3390/beverages5020033.

[33] I. Kim, K.A. Hoff, E.T. Hessen, T. Haug-Warberg, H.F. Svendsen, Enthalpy of absorption of CO2 with alkanolamine solutions predicted from reaction equilibrium constants, Chem. Eng. Sci. 64 (2009) 2027–2038. https://doi.org/10.1016/j.ces.2008.12.037.

[34] R. Peng, P-R.a New Equation of State., Proc. Natl. Acad. Sci. U. S. A. 15 (1976) 11–18. http://dns2.asia.edu.tw/~ysho/YSHO-English/1000 CE/PDF/Ind Eng Che Fun15, 59.pdf.

[35] K.R. Putta, D.D.D. Pinto, H.F. Svendsen, H.K. Knuutila, CO2 absorption into loaded aqueous MEA solutions: Kinetics assessment using penetration theory, Int. J. Greenh. Gas Control. 53 (2016) 338–353. https://doi.org/10.1016/j.ijggc.2016.08.009.

[36] J.M. Prausnitz, R.N. Lichtenthaler, E. Gomes de Azevedo, Molecular Thermodynamics of Fluid-Phase Equilibria, 3rd ed., Prentice-Hall, 1999.

[37] J. Oexmann, A. Kather, Post-combustion CO2capture in coal-fired power plants: Comparison of integrated chemical absorption processes with piperazine promoted potassium carbonate and MEA, Energy Procedia. 1 (2009) 799–806. https://doi.org/10.1016/j.egypro.2009.01.106.

[38] A. Raksajati, M.T. Ho, D.E. Wiley, Reducing the cost of CO2 capture from flue gases using aqueous chemical absorption, Ind. Eng. Chem. Res. 52 (2013) 16887–16901. https://doi.org/10.1021/ie402185h.

[39] F.Y. Jou, F.D. Otto, A.E. Mather, Vapor-Liquid Equilibrium of Carbon Dioxide in Aqueous Mixtures of Monoethanolamine and Methyldiethanolamine, Ind. Eng. Chem. Res. 33 (1994) 2002–2005. https://doi.org/10.1021/ie00032a016.

[40] I. Kim, K.A. Hoff, T. Mejdell, Heat of absorption of CO2 with aqueous solutions of MEA: New experimental data, Energy Procedia. 63 (2014) 1446–1455. https://doi.org/10.1016/j.egypro.2014.11.154.

[41] A.A.M. Al-Saffar, H. Tao, M.A. Talab, Review of deep convolution neural network in image classification, Proceeding - 2017 Int. Conf. Radar, Antenna, Microwave, Electron. Telecommun. ICRAMET 2017. 2018-Janua (2017) 26–31. https://doi.org/10.1109/ICRAMET.2017.8253139.

[42] O. Vinyals, I. Babuschkin, W.M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D.H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J.P. Agapiou, M. Jaderberg, A.S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T.L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, D. Silver, Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature. 575 (2019) 350–354. https://doi.org/10.1038/s41586-019-1724-z.

[43] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control. Signals, Syst. 2 (1989) 303–314. https://doi.org/10.1007/BF02551274.

[44] Matlab, trainlm: Levenberg-Marquardt backpropagation, Mathworks Doc. (2019). https://se.mathworks.com/help/deeplearning/ref/trainlm.html (accessed December 20, 2019).

[45] D.W. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, J. Soc. Ind. Appl. Math. 11 (1963) 431–441. https://doi.org/https://doi.org/10.1137/0111030.

[46] M.T. Hagan, M.B. Menhaj, Training Feedforward Networks with the Marquardt Algorithm, IEEE Trans. Neural Networks. 5 (1994) 989–993. https://doi.org/10.1006/brcg.1996.0066.

[47] D.J.C. MacKay, Bayesian Interpolation, Neural Comput. 4 (1992) 415–447. https://doi.org/10.1162/neco.1992.4.3.415.

[48] F. Dan Foresee, M.T. Hagan, Gauss-Newton approximation to bayesian learning, IEEE Int. Conf. Neural Networks - Conf. Proc. 3 (1997) 1930–1935. https://doi.org/10.1109/ICNN.1997.614194.

[49] H. Okut, D. Gianola, G.J.M. Rosa, K.A. Weigel, Prediction of body mass index in mice using dense molecular markers and a regularized neural network, Genet. Res. (Camb). 93 (2011) 189–201. https://doi.org/10.1017/S0016672310000662.

[50] H. Garoosiha, J. Ahmadi, H. Bayat, The assessment of Levenberg–Marquardt and Bayesian Framework training algorithm for prediction of concrete shrinkage by the artificial neural network, Cogent Eng. 6 (2019) 1–14. https://doi.org/10.1080/23311916.2019.1609179.

[51] M.D. Hilliard, A Predictive Thermodynamic Model for an Aqueous Blend of Potassium Carbonate, Piperazine, and Monoethanolamine for Carbon Dioxide Capture from Flue Gas, Dr. Thesis Tech. Univ. Texas Austin. (2008) 1083.

[52] J.P. Jakobsen, J. Krane, H.F. Svendsen, Liquid-phase composition determination in CO2-H2O-alkanolamine systems: An NMR study, Ind. Eng. Chem. Res. 44 (2005) 9894–9903. https://doi.org/10.1021/ie048813+.

[53] A.E. Sherwood, J.M. Prausnitz, The heat of solution of gases at high pressure, AIChE J. 8 (1962) 519–521. https://doi.org/10.1002/aic.690080419.

[54] I. Kim, H.F. Svendsen, Heat of Absorption of Carbon Dioxide (CO 2) in Monoethanolamine (MEA) and 2-(Aminoethyl)ethanolamine (AEEA) Solutions, Ind. Eng. Chem. Res. 46 (2007) 5803–5809. https://doi.org/10.1021/ie0616489.

[55] C. de Boor, A Practical Guide to Splines, Springer-Verlag, New York, 1978.

[56] H. Majeed, H.F. Svendsen, Characterization of aerosol emissions from CO2 capture plants treating various power plant and industrial flue gases, Int. J. Greenh. Gas Control. 74 (2018) 282–295. https://doi.org/10.1016/j.ijggc.2018.04.016.

[57] D.M. Austgen, G.T. Rochelle, X. Peng, C.C. Chen, Model of Vapor—Liquid Equilibria for Aqueous Acid Gas—Alkanolamine Systems Using the Electrolyte—NRTL Equation, Ind. Eng. Chem. Res. 28 (1989) 1060–1073. https://doi.org/10.1021/ie00091a028.

[58] E.T. Hessen, T. Haug-Warberg, H.F. Svendsen, The refined e-NRTL model applied to CO2–H2O–alkanolamine systems, Chem. Eng. Sci. 65 (2010) 3638–3648. https://doi.org/10.1016/j.ces.2010.03.010.

[59] J.J. Carroll, J.D. Slupsky, A.E. Mather, The Solubility of Carbon Dioxide in Water at Low Pressure, J. Phys. Chem. Ref. Data. 20 (1991) 1201–1209. https://doi.org/10.1063/1.555900.

[60] R.H. Weiland, T. Chakravarty, A.E. Mather, Solubility of Carbon Dioxide and Hydrogen Sulfide in Aqueous Alkanolamines, (1993) 1419–1430. https://doi.org/10.1021/ie00019a016.

Chapter 3.

Neural network programming

The novel NNP hybrid modelling method and the ASNNs are presented in this chapter. NNP is contrasted against the current hybrid modelling methods in order to highlight its modelling advantages. The NNP is used for the development of the ASNNs described in Chapter 4 - Chapter 5. The NNP method is applied for modelling a diverse set of chemical engineering processes such as flash tanks, distillation columns or biogas upgrading processes.

Neural Network Programming: Integrating First Principles into Machine Learning Models

Andres Carranza-Abaid*, Jana P. Jakobsen

Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), NO-7491, Trondheim, Norway

*Corresponding author: andres.c.abaid@ntnu.no

Keywords: Hybrid Modelling, Physics-informed machine learning, Surrogate Modelling, Numerical Analysis, Thermodynamics

3.1. Abstract

This work introduces Neural Network Programming (NNP) as an integrated hybrid modelling approach. NNP consists in formulating a set of first principles equations that is later decomposed and transcribed to an Algorithmically Structured artificial Neural Network (ASNN). NNP leverages the advantages of the universal approximation theorem and neural network optimization algorithms in order to generate physically coherent machine learning models. Since ASNNs are not mere approximations of physics equations, it is not necessary to modify either the gradient or performance function in order to account for errors with respect to the first principles equations. ASNNs are automatically differentiable, hence, they are trained faster and more accurately than typical hybrid models. It is shown that the same ASNN architecture is transferable between processes with similar characteristics.

3.2. Introduction

The complex nature of the physics involved in engineering processes is usually reflected in highly intertwined and sophisticated mechanistic models. This complexity has always been both a challenge and a motivation for the formulation of new modelling strategies and techniques. Although quite rigorous, mechanistic models often exhibit considerable deviations from experimental measurements, perhaps due to unreasonable assumptions about the physics or maybe because some effects were not properly characterized. Recently, the modelling community has shifted its attention to data-driven modelling tools. From quite a few modelling alternatives, machine learning algorithms seem to have the potential to become the dominant tool in the Industry 4.0 era [1,2]. From all the ML methods, Artificial Neural Networks (ANNs) are particularly interesting since they have been used to predict the unfolding mechanism of macromolecules [3], redesign proteins [4] or even perform activities that require complex and high-level decision making [5,6].

The outstanding performance of ML algorithms together with important media coverage has brought the awareness of their capabilities, not only to academia, but also to the general public. However, ML has been around for decades or even centuries if one considers that linear regression was pivotal in the development of these algorithms. Figure 3.1 shows that the overall relative amount of ML research publications indexed in SCOPUS with respect to the end of the previous decade has been rapidly increasing. We expect that, in the same fashion as with other novel technologies, the annual research output will reach a maximum and then a decline will come afterwards. Nevertheless, due to the advances in computer science we expect that this technology will be around for several decades.

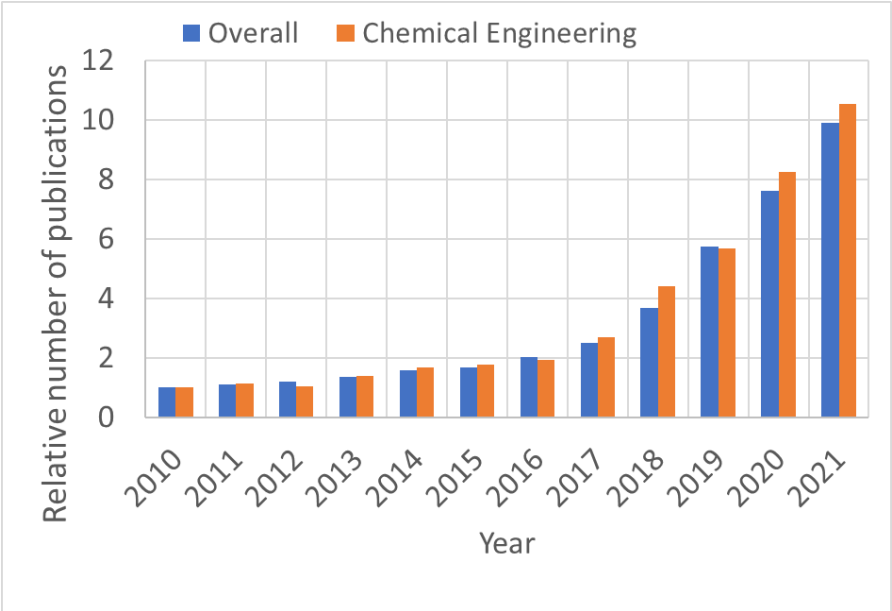


Figure 3.1. Relative number of research documents indexed in SCOPUS (The relative number of publications was calculated by dividing the yearly number of publications over the number of publications available in 2010). Keywords: “machine learning” OR “artificial neural networks”. “Limited to” Chemical engineering whenever applicable.

There might be some well-founded skepticism about the reliability and applicability of ML and ANNs in many chemical engineering subdisciplines. This might be due to the lack of transparency and the fact that purely data-driven models override first principles relationships. Some of these concerns have been partially addressed in the past with the introduction of hybrid modelling algorithms. However, from our perspective, in most cases the traditional hybrid modelling methodologies address the physics problems more from a computer science perspective rather than from a chemical engineering or physics perspective. Although in some chemical engineering subdisciplines utilizing relaxed physics models is convenient, this is not valid in several subdisciplines in which it is mandatory that certain mathematical relationships

are exactly conserved and not only approximated. One example of this is found in thermodynamics, where several restrictions must be accounted for in order to be able to call a set of equations a “thermodynamic model”. Some other subfields that require a more rigorous description of the physics are transport phenomena, kinetics modelling or thermophysical property modelling, which we consider, will find this work especially interesting.

This work proposes a hybrid modelling method named Neural Network Programming (NNP) that utilizes both theoretical knowledge and the universal approximator capabilities of ANNs. In essence, NNP consists in integrating a first-principles model within a customized neural network type called Algorithmically Structured Neural Network (ASNN). NNP not only provides a new set of hybrid models but also gives a new perspective on how to utilize and approach neural networks. In this way, chemical engineers can formulate customized architectures instead of only utilizing borrowed predefined generic architectures (e.g., typical fully connected neural networks). The NNP method has some advantages over purely data-driven algorithms and typical hybrid modelling configurations including their rigorousness, extrapolation capabilities, interpretability, auto differentiability (optimized faster and more accurately), ability of representing limit cases (e.g., an appropriate ASNN of a flash separator will not compute vapor molar fractions of a component whose molar composition is 0), and parallel computations (in order to simulate several systems a single call to the ASNN is required instead of using “for” or “while” cycles). Due to the parallel computation advantages given by quantum computing [7], this last feature might become critical in the coming decades.

3.3. Modelling methodologies

3.3.1. First principles models

First principles models are also known as mechanistic, semi-empirical, phenomenological, or white-box models. They are catalogued as rigorous although they are founded on an idealized human interpretation of the physics phenomena. Due to their inherent simplifications, these models usually possess empirical parameters that account for unmodelled dynamics (i.e., unmodelled phenomena). Mechanistic models can be regarded as interpretable and physically coherent because the mathematical equations are compelled to be consistent with physics laws (e.g., conservation, kinetics, or thermodynamics laws) despite the use of empirical parameters.

There are different mathematical frameworks that can be utilized for mechanistic modelling in chemical engineering. The first type of frameworks is known as equation-oriented modelling. These frameworks utilize their own user-friendly programming language, auxiliary chemical-engineering-oriented routines, and requires defining all involved variables and equations. Some notable examples of equation-oriented modelling frameworks are gPROMS, Aspen Custom Modeler, Dymola [8], ASCEND II [9], Omola/Omsim [10,11] , EMSO [12], ICAS-MoT [13], JModelig.org, Mosaic/Optimica [14,15], DAE Tools [16], Daedalus modelling framework [17], among others. The second category corresponds to phenomenological modelling frameworks. These frameworks explicitly utilize first principles to semiautomatically generate equation-oriented models, where the user specifies information about the problem such as assumptions or the topology of the system. Some notable phenomenological modelling frameworks are Model.la [18], Techtool [19], ModKit

[20], OntoCAPE [21], Mobatec [22], computer-aided modelling template [23] and an ontology builder for physics-chemical processes [24].

In some cases, these frameworks are either unknown by the modelers or are not suitable for specific studies. For instance, there are not any open packages for bifurcation analysis that can be utilized for thermally coupled distillation columns, thus, a combination of different libraries (MatCont and Aspen) was needed in order to perform the hysteresis study [25,26]. Another issue regarding the development of mechanistic models is their high requirement for multiple model subroutines that complement the first-principles model. For example, in order to model an absorption column, models for vapor-liquid equilibrium, kinetics, viscosity, surface tension, diffusivity and packing correlations are needed [27]. The number of subroutines needed in many unit operations can be inconvenient. Thus, in many cases, models based on machine learning seem to be a more practical option.

3.3.2. Data-driven models (Artificial neural networks)

Machine learning models are data-driven models that consist of an arrangement of equations that do not explicitly describe the causal interactions between the data, thus, they entirely rely on the data and their quality instead of the shape of the equations. Because of their lack of interpretability, these models are usually known as black boxes. Some of the most relevant data-driven modelling methods for chemical engineering are linear regression, artificial neural networks (ANNs), support vector machines (SVM), multivariate adaptative regression splines (MARS), latent variable methods, dimensionality reduction methods, to name a few.

ANNs —also known as multilayer perceptrons — are the most prominent data-driven models because of their ability to model highly non-linear systems. They were

developed as representations of how biological brains process information [28–30]. In other words, ANNs are mathematical models that we, as humans, developed to "understand" how we "understand" things. However, due to the extreme complexity of biological brains, ANNs seem to have found broader use in data and computer science rather than in neurological sciences. Their modelling robustness is due to the universal approximation feature of the multilayer perceptrons demonstrated in the late 1980s [31–34].

ANNs are constituted by input layers, output layers and hidden layers. The input layers transfer the information specified by the user to the neural network and the output layers deliver the results back to the user. Hidden layers are constituted by smaller building blocks called artificial neurons that transform the information in order to provide a numerical prediction of the outputs as a function of the inputs. Each artificial neuron (hence all hidden layers) must follow the artificial neuron signal transformation process, which consists in combining the input vectors and applying a transfer function afterwards. The combination of the input vectors is usually done through a linear combination. However, the input vectors can also be combined using a Schur / Hadamard product (element-by-element product). The Hadamard product as a means of combining input functions has been mostly used for image processing and classification (e.g., [35–38]). On the other hand, transfer functions are usually nonlinear or logic functions like the linear hyperbolic tangent function, log-sigmoid function, or rectified linear unit (ReLU). In some cases, it is not necessary to perform a non-linear transformation and the transfer function is simply a linear function.

In order to illustrate the artificial neuron signal transformation process, a hidden layer formed of p artificial neurons that linearly combines the inputs and uses a hyperbolic

tangent transfer function (Figure 3.2 a)). This is hidden layer is mathematically described as

$$L_1 = \tanh\left(\mu_1 X_1 + \mu_2 X_2 + \beta\right) \quad [p], \tag{3.1}$$

where L_1 is the output vector with p x 1 dimensions (p is the number of artificial neurons), X_1 is an input vector with q_1 x 1 dimensions while X_2 has q_2 x 1 dimensions. The square brackets to the right of the equation are utilized to encompass the number of artificial neurons in a particular hidden layer. The fitting parameter matrices μ_1 and μ_2 have p x q_1 and p x q_2 dimensions, respectively, while the bias β has p x 1 dimensions. In every hidden layer, the fitting parameter matrices reshape the size of the input vectors so that the new column vector has as many elements as number of artificial neurons.

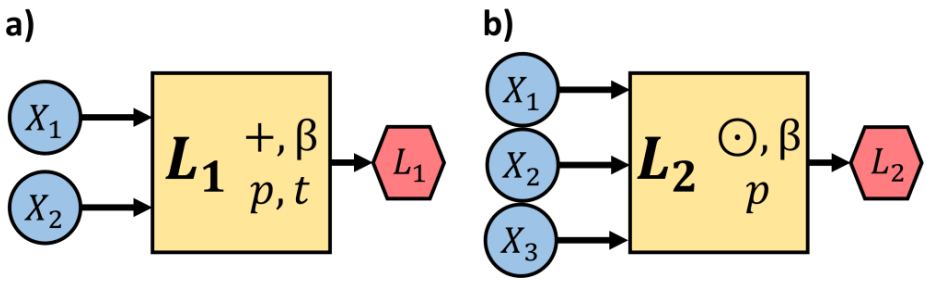


Figure 3.2. Examples of hidden layers:

a) Linear combination with hyperbolic tangent transfer function and b) Hadamard product.

A second example containing the Schur product (Figure 3.2 b)) is described with the following equation

$$L_2 = \left(\mu_1 X_1\right) \odot \left(\mu_2 X_2\right) \odot \left(\mu_3 X_3\right) \odot (\beta) \quad [p], \tag{3.2}$$

where the $\odot$ symbol stands for the Schur product. Once again, the fitting parameters matrices μ_1, μ_2 and μ_3 reshape the input vectors (X_1, X_2 and X_3) to the size of the output vector (p x 1 dimensions) to perform the operations.

As seen in Figure 3.2, this work introduces a new way of expressing the architecture of an ANN. This was proposed because we consider that the typical fully connected representation does not provide valuable information aside from the overall input-output relationship. The symbols in bold represent the name of the hidden layer while the subindexes and superscripts are related to the characteristics of the hidden layer. Every layer in this representation may have up to 2 subscripts and 2 superscripts. The first subscript indicates represents the number of artificial neurons p in the layer. The second subscript symbolizes the use transfer function: t for the hyperbolic tangent transfer function, e for the exponential function, l for the natural logarithm transfer function, $\oslash$ for the Hadamard division function, s for the saturating linear transfer function, r for the rectified linear transfer function, and a blank space for the linear transfer function. The first superscript indicates whether the linear combination ($+$) or the Hadamard product ($\odot$) is used. The second superscript indicates whether there is a bias (marked with β) or not (no second superscript).

This work focuses on the application of the NNP method to Feedforward Neural Networks (FFNNs). FFNNs are a static time-invariant neural network and are divided in two subclasses: shallow artificial neural networks (e.g., Figure 3.3 a)) and deep

neural networks (e.g., Figure 3.3 b)). The difference between a shallow neural network and a deep neural network is that the former must have 2 hidden layers while latter should have 3 or more hidden layers. In general, shallow neural networks are related to “machine learning” while deep neural networks are related to “deep learning”. Despite their differences, according to the universal approximation theorem, both shallow and deep neural networks are universal approximators.

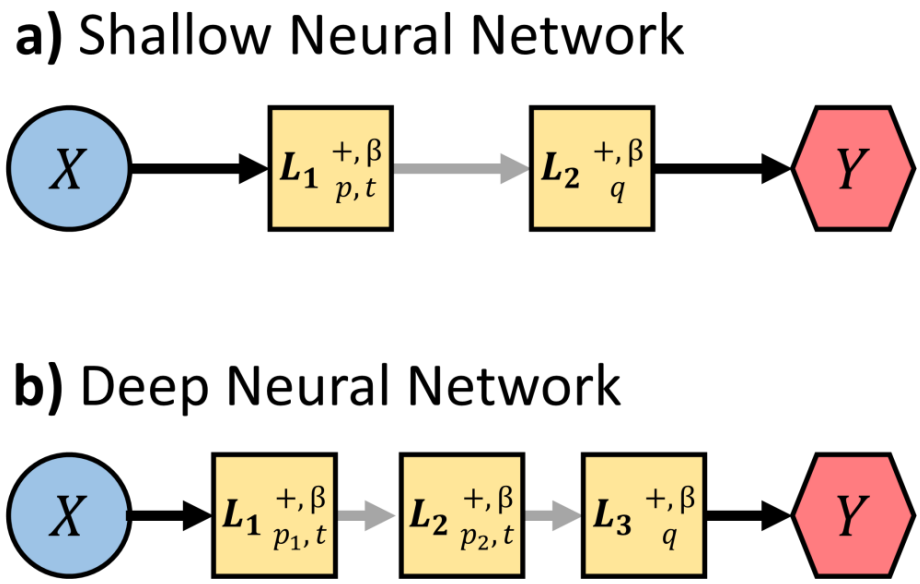


Figure 3.3. FFNN sketches:

a) shallow neural network and b) deep neural network.

3.3.3. Hybrid models

Data-driven models that are, to some extent, physics-driven are known as hybrid models or grey-box models. Hybrid modelling combines the flexibility of ML-based models with the rigorousness of mechanistic models. Specifically, they utilize the physics framework inherent to mechanistic models and a data-driven model to reduce the discrepancy between the hybrid model and the experimental data. The typically reported benefits of using hybrid models over the mechanistic and data-driven modelling methodologies are: lower data requirement, more interpretability, more accuracy, and more compliance with physics than purely data-driven models [39–41]. Several data-driven models have been utilized in hybrid models such as ANNs, support vector machines [42], Padé expansion [43], extended Kalman filter [44], multivariate adaptive regression splines (MARS) [45], multivariate discrete-time models [46], principal component analysis [47], fuzzy systems [48], to name a few.

The development of hybrid models can be done by using serial or parallel hybrid modelling structures [49]. Although in both modelling structures the core idea is alike, the mechanistic and the data-driven model interact differently as illustrated in Figure 3.4.

In the serial hybrid modelling paradigm, the data-driven and mechanistic models are set up in sequence (Figure 3.4 a)). The input variables are fed to the black box model thereafter fed to the first principles model (the sequence between the black and white boxes can also be reversed [50]). The mechanistic model has the first principles equations (e.g., conservation laws) while the black box is utilized as the empirical section of the model (e.g., mass transfer coefficients or kinetic constants). The pioneering work by Psichogios and Ungar (1992) for sequential hybrid models was

firstly applied to fed batch bioreactors in order to substitute the traditional empirical parts of a kinetic model with an ANN. Some other hybrid modelling examples have been used to predict the pulp quality [52], kinetics of polymerization processes [53], fermentation processes [54], scaling of a pilot plant catalytic cracking [55], crystallization rates [56–58], process identification of an ethylene glycol process [59], analysis of fixed-bed reactor processes [60], cheese fermentation [61].

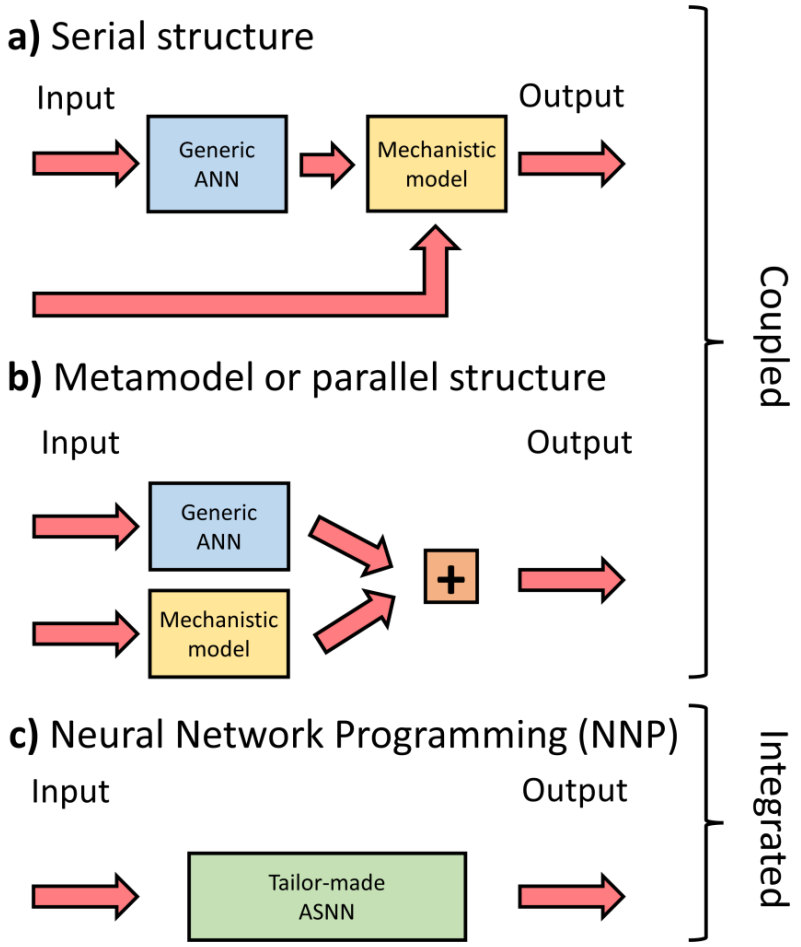


Figure 3.4. Hybrid modelling structures.

The parallel hybrid modelling paradigm consists in formulating a mechanistic model independently of the black box model (Figure 3.4 b)). The data-driven model is used to predict the difference between the mechanistic model (with its own empirical parameters) and the experimental measurements. The first reported parallel model was done by Su et al. (1992) for the modelling of a reactor system where the output signals of both individual models were summed. A similar approach was proposed by Thompson and Kramer (1992) for the predictions of the kinetics in a fermentation study. Other examples are the knowledge based modular networks for yeast production processes [64], process control with knowledge-enhanced algorithms [65] or monitoring and control of bioreactors [66]. More recently, improved algorithms and structures that provide enhanced accuracy in parallel hybrid models have been proposed for flowmeters [41].

The decision of which modelling structure is better is a function of the amount of available data and the accuracy of the mechanistic model. In general, it has been proposed by Sansana et al. (2021) as generally rule to choose parallel models if the deviation between the mechanistic model and the measurements is considerable; otherwise, a serial model is preferred.

On the other hand, there are other important types of hybrid modelling paradigms that are not based on these structures. They add information through the gradient of the performance function rather than with a mechanistic model. Their main feature is the utilization of a loss function comprised of different error terms, which are specific to each method. Some of the most remarkable examples is the solution of Partial Differential Equations (PDE) based on Physics-Informed Neural Networks (PINNs) [67] or the theory/physics guided neural networks [68,69]. For a more in-depth review

of these hybrid modelling approaches the reader can refer to the review by Karniadakis et al., 2021.

The neural network programming (NNP) approach proposed in this work is an integrated method that transcribes first principles equations on the architecture of an ASNN (Figure 3.4 c)). Applying NNP to a problem yields a new set of neural networks: Algorithmically Structured Neural Network (ASNNs). NNP simplifies the ASNN training since the performance function does not need to be modified. Utilizing NNP allows the formulation of models that are completely coherent with the physics framework and are not mere approximations (as gradient-based or parallel structure methods). These features make the NNP an ideal tool for fields like thermodynamics or kinetic modelling where the consistency of the model is as important as the accuracy.

ASNNs can be equivalent to serial hybrid models if the same model structure, variables, and information flow is utilized. Nevertheless, as opposed to the typical hybrid model paradigms, ASNNs are automatically differentiable. Automatic differentiation is a technique in which the derivative of an operation is computed by evaluating elementary arithmetic operations, thus, the numerical error is only caused by computational precision limits. Since the typical hybrid model structures are usually not automatically differentiable, they perform numerical differentiations with an inherent numerical error. The numerical error associated to numerical differentiation can be highly inaccurate due to round-off and truncation errors [71]. Even though for small processes with low amount of experimental data the computational efficiency might not be significant, for large processes training an ANN might become economically prohibitive. Another important factor is the coding efficiency imposed by

the ASNNs framework, which compels the model to be written in matricial operations, hence, exploiting the parallel computing technology.

Although ASNNs and PINNs have similar conceptual ideas, the main differences between the two is that PINNs account for the physics and constraints through the performance function while the ASNNs account for them through the architecture. This suggests that ASNNs require fewer fitting parameters than PINNs (which usually require several layers with dozens or hundreds of fitting parameters). However, it is not fair to compare these two approaches in this fashion since PINNs and ASNNs were designed to solve different sets of problems.

3.3.4. Interpretability

The lack of interpretability in machine learning models might be one of its main drawbacks. In order to address this issue, two method categories for interpreting data-driven models have been reported model-specific methods and model-agnostic methods [72]. The model specific methods are those that can be understood or "read" by analyzing the model structure like the gradient booster algorithm XG-boost [73], the use of rule lists [74,75], or additive models [76]. The model agnostic methods evaluate the feature importance and, therefore, are applicable to any machine learning method. For a more in-depth review we recommend the following references Ribeiro et al. (2016) and Bikmukhametov and Jäschke (2020).

The ANN training procedure heavily relies on random processes that induce entropy of modeling into the optimized parameters and, hence, reducing the chances to interpret the model. Randomness can be observed in the selection of the datapoints used for training / validation, the initial values of the optimizable parameters, or the optimization process (e.g., the stochastic random descent method). Furthermore, the

fact that the ANNs utilize generic equations of the same form (e.g., Eq. (3.1)) does not facilitate their interpretation. Because of this, it is unlikely to produce the same optimized parameter values in different runs. The highly entropic behavior of the ANN training can be compared to the molecular movement in an ideal mixture in an isolated system (like in the Figure 3.5 a) sketch). Due to the high entropy associated to the molecular movement, it is unlikely that the same molecule arrangement will be obtained even if both runs start at the same conditions (chances are 1 in 24).

Mechanistic models are expected to have a lower entropy since the equation parameters are constrained by the first principles equations. This behavior can be compared to the molecular movement in a nonideal mixture in an isolated system (as illustrated in Figure 3.5 b)). Although in principle, these molecules can “freely” move, the molecule interactions can compel that some molecule arrangements are always formed (the fact that molecules associate has been used for developing local composition activity coefficient models [77,78]).

Considering the above, we hypothesize that an ASNN with an appropriate architecture will have a more uniform distribution of the optimized parameters (if trained several times) even if it is trained with the highly entropic algorithms commonly used in ANNs. Therefore, the proper construction of the ASNN might be equally or more important than the training procedure. The interpretability of ASNNs is further discussed in section 3.5.1.2.

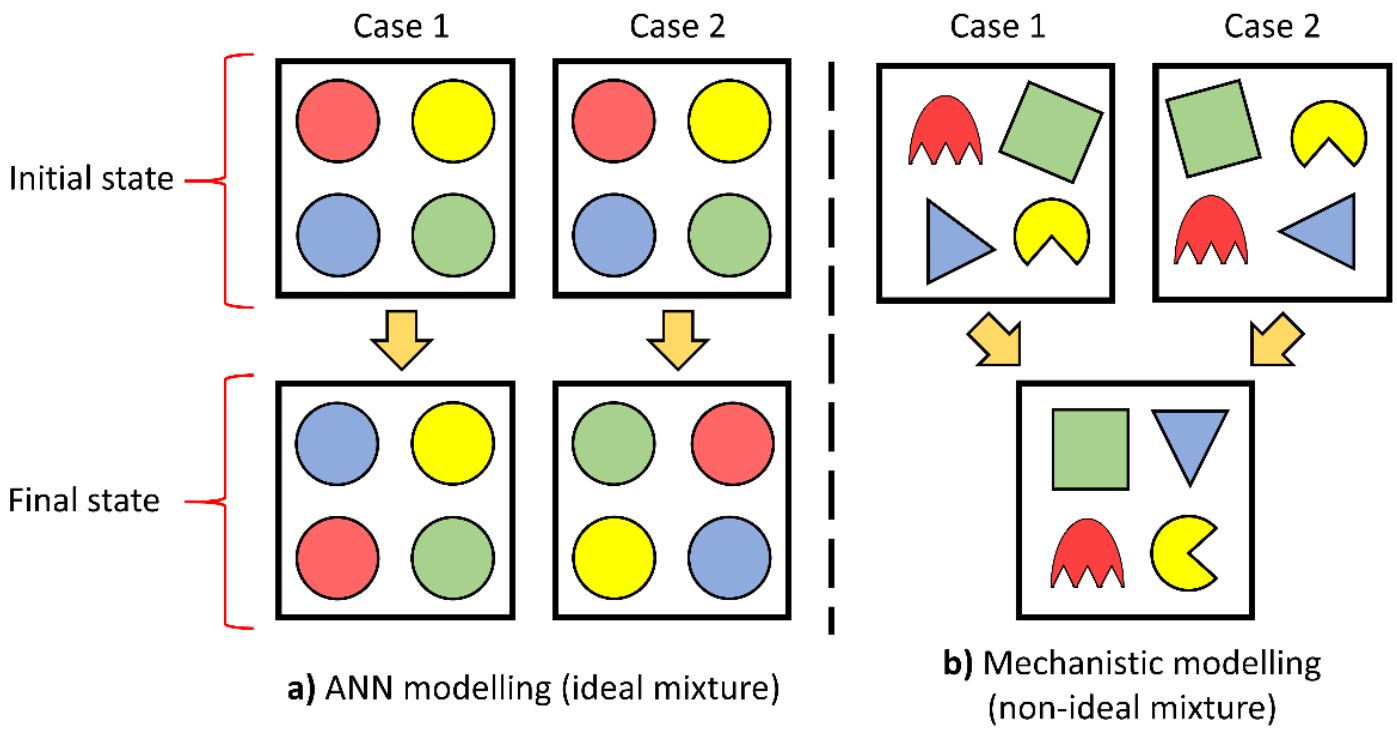


Figure 3.5. Analogy between the thermodynamic entropy of an isolated system and the parameter entropy in a) ANNs and b) mechanistic models.

3.4. Methodology

The NNP algorithm is divided in six steps, which are briefly described below and exemplified in section 3.5.

Step 1. Setting the modelling objective and data collection. Definition of the phenomenon or process to be modelled, delimitation of the system boundaries, gathering of experimental data or auxiliary data that can aid the ANN training (e.g., vapor liquid equilibrium data may be useful if a flash separator is to be modelled).

Step 2. Definition of the assumptions, physics laws and constraints. Formulation of the first-principles equations and auxiliary equations to be transcribed. Determination of the model inputs and outputs in accordance with the Degrees of Freedom (DoF) of the first principles model. Defining a first principles system of equations inherently assumes a model structure, therefore, the first principles system of equations must be completely determined. Acknowledging the assumptions and the limitations of the physics equations is paramount in order to modify the ASNN in case that the proposed structure does not meet the performance expectations.

Step 3. Identification of the uncertain/surrogate sections of the model. The selection of the empirical parameters/sections that are to be substituted with a universal approximator substructure. A universal approximator substructure is any ANN with a shallow neural network or a deep neural network architecture. It should be remarked that the user must ensure that the uncertain/surrogate sections of the model completely satisfy the system of equations proposed in step 2.

The user has the flexibility to either use traditional semi-empirical parameters (e.g., saturation pressure) or new model parametrizations. The universal approximators can

also be utilized as surrogates in order to simplify the modelling process by lumping several parameters into a single substructure.

Step 4. Equation decomposition. Every first principles equation must be decomposed into simpler equations so that they can be transcribed and subsequently arranged. It must be remarked that the decomposition must follow the rules imposed by the artificial neuron signal transformation process (e.g., Eqs. (3.1) - (3.2)). The decomposition algorithm for an arbitrary equation $\boldsymbol{f}$ is presented in Figure 3.6, where X_j is the input j fed to the equation Y_k and m is the total number of inputs fed to Y_k (includes the bias/constant, if any). The function $\boldsymbol{f}$ is defined by placing on the right side the part of the equation to be decomposed. As seen in Figure 3.6, the decomposition is done according to the hierarchy of operations (HO), in other words, in the same order as if the equation was being numerically evaluated in a calculator. Therefore, the decomposition should be done from left to right, from the innermost to the outermost parentheses, and multiplication operations have priority over the addition / subtraction operations. Note that if function Y_k^* is not encompassed in a nonlinear function, then $Y_k = g(Y_k^*) = Y_k^*$. The algorithm ends when the right side of the first principles equation cannot be decomposed anymore (i.e., when $Y_k = f$).

It must be remarked that divisions cannot be expressed as a/b in an ANN, but rather as ab^{-1}. Therefore, divisions should be treated as nonlinear functions in NNP (this is valid for the Deep Learning Toolbox from Matlab).

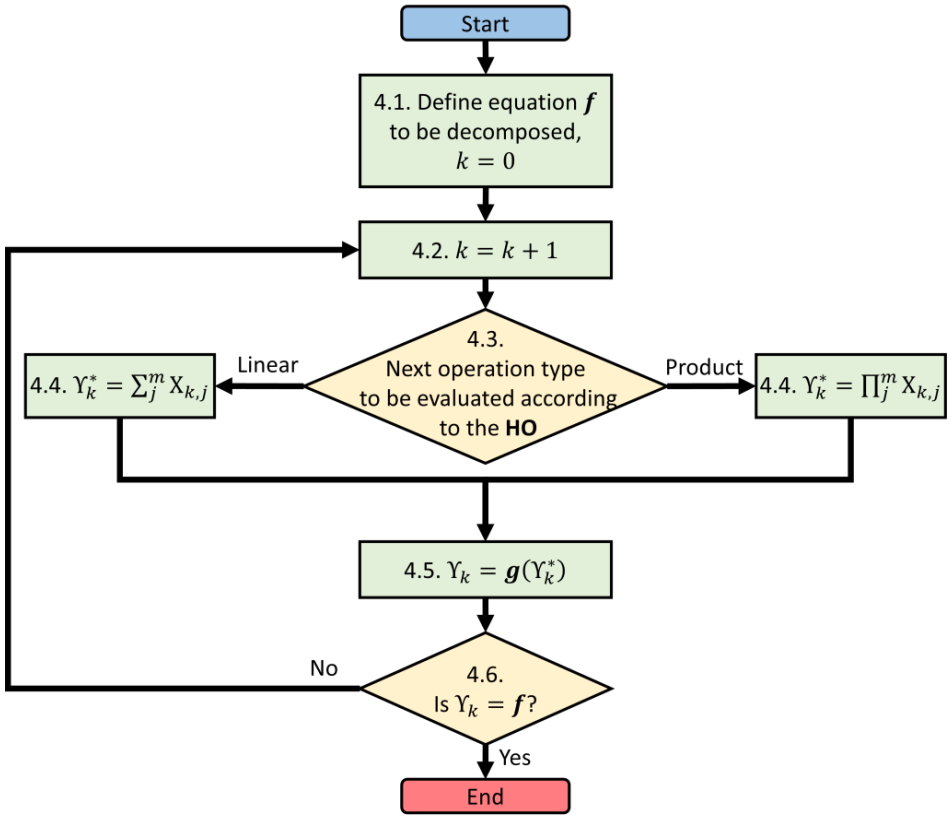


Figure 3.6. Equation decomposition algorithm. **HO:** hierarchy of operations.

Step 5. Construction of the ASNN. A solution algorithm should be formulated in order to solve the decomposed equations. In this work, we propose two alternatives for implementing the solution procedure: sequential solution algorithm and matrix inversion algorithm (further explained in section 3.5). After the solution algorithm is proposed, the ASNN is constructed following the order of computation given by the solution algorithm.

The input and output layers correspond to the inputs and outputs selected in steps 2 and 3. The inputs can be fed to the ASNN in different mathematical forms as long as

they are not dependent on the computations done by the neural network (this is further discussed in section 3.5.2.1).

The algorithm shown in Figure 3.6 automatically generates the solution sequence (inverted) of the decomposed equations. However, if more than one function $\boldsymbol{f}$ is to be decomposed, then the user should decompose each function and subsequently arrange all the decomposed equations from every function.

It must be remarked that an ASNN must have an associated weight matrix (W) to every vector that is fed to a layer (e.g., if the decomposition of an arbitrary equation yields $Y_k = X_1 + X_2$, the neural network will represent it as $Y_k = W_1 X_1 + W_2 X_2$).

Constructing the ASNN architecture is the only mandatory requirement in this step. Nonetheless, minor structural modifications can be performed in order to simplify the ASNN architecture or to constraint the predictions. The first type of minor modification merges the output of a universal approximator with the equation that is feeding. However, this can only be done if the output of the universal approximator feeds exactly one equation. A second type of minor modification consists in adding process constraints in order to help the ASNN to compute unrealistic results if the model is extrapolated too far away from the training conditions (e.g., avoid negative volumes). The minor modifications are no limited to the ones listed above. These modifications are done as a function of each modelled system and, in general, are not a requirement for successfully implementing NNP.

Step 6. Training the ASNN. The parameters with low entropy are those that help to transcript the physics laws within the ASNN architecture (identified in step 2) while the high entropy parameters are the ones used for modelling semi-empirical parameters

(identified in step 3). Considering this, the low entropy parameters must be fixed so that the physics equations are properly transcribed.

The fixed parameters can either be used to only transfer the signal (e.g., if the identity matrix is used), to perform summations of the input vector (e.g., if an all-ones row vector), or to arrange an output vector according to the solution algorithm (e.g., to set up the constant vector C for the matrix inversion algorithm). Utilizing the weight parameter matrices for quite different purposes is one of the main strengths of the NNP algorithm. The high entropy parameters are set free so that the ASNN can adjust to the data.

Since the unfixed parameters have a high entropic behavior, it is recommended to train the model several times and pick the set of parameters that provides the best performance.

3.5. Results and discussion

This section presents the application of the NNP method to two study cases. The first case uses NNP to model a thermodynamic vapor-liquid equilibrium (VLE) problem and the second study case applies the NNP to two-product separators.

3.5.1. Case study 1: ideal VLE system

3.5.1.1. Application of the NNP method

Step 1. Setting the modelling objective and data collection. This example develops a VLE model for a benzene (1) and toluene (2) system. A dataset containing the results of 200 simulations was utilized to train the ASNN. This dataset was generated by using the ideal thermodynamic package available in Aspen Plus v8.8. The molar fraction of component 1 (x_1) and the temperature (T) are the 2 independent variables and were randomly generated according to $0 < x_1 < 1$ and $333\ K \leq T \leq 473\ K$. The output variables taken from the Aspen Plus model were the partial pressure p_i and the total pressure P.

Step 2. Definition of the assumptions, physics laws and constraints. The physics laws and constraints to be considered are Raoult's law (Eq. (3.3)), and Dalton's law (Eq.(3.5))

$$\ln x_i + \ln p_i^{sat}(T) = \ln p_i \tag{3.3}$$

$$p_i = \exp\left(\ln p_i\right) \tag{3.4}$$

$$\sum_i p_i = P. \tag{3.5}$$

Considering that the partial pressure vector in Eq. (3.3) is in logarithmic and the partial pressure vector in Eq. (3.5) is not, an auxiliary equation (Eq. (3.4)) is needed to solve the system of equations. Whenever there is a subscript i, the equation refers to component i, otherwise, it refers to a vector (e.g., x is the vector of liquid molar compositions and x_i is only for component i)

Imprinting physics laws into the ASNN inherently assumes that the theoretical framework in which they were developed is correct. Hence, the available DoF must be the same as in the mechanistic model. It was previously shown that ignoring the DoF for modelling a VLE system (Gibbs' phase rule) makes the ANN to find correlations foreign to equilibrium thermodynamics [79]. The Gibbs' phase rule determines the number of intensive variables that can be independently selected. For vapor-liquid non-reactive systems, the Gibbs' phase rule is

$$DoF = 2 + n - \pi\,. \tag{3.6}$$

Where n is the number of components and π is the number of phases. By evaluating Eq. (3.6), the DoF is 2.

Step 3 Identification of the uncertain/surrogate sections of the model. In this example, the source of entropy (or ignorance according to the definition of entropy [80]) comes from the uncertainty about the form of the pure component saturation pressure equation. At low pressures, p_i^{sat} can be safely assumed to be only a function of T. Therefore, a shallow neural network function (η) with the general form can be utilized

$$\ln p_i^{sat} = \eta\,(T). \tag{3.7}$$

Note that η is formed by two sequential equations.

Step 4. Equation decomposition. Since the equations utilized in this example are relatively simple, the decomposition procedure does not change the form of Eqs. (3.3), (3.4), (3.5), and (3.7). Therefore, they can be directly transcribed into an ASNN.

Step 5. Construction of the ASNN. This example applies the functionality matrix equation-ordering algorithm from Book and Ramirez [81,82] to arrange Eqs. (3.3), (3.5), (3.7). The functionality matrix of this example is

$Eq.\backslash Var.$	T	$\ln p_i^{sat}$	$\ln x_i$	$\ln p_i$	p_i	P
(3.7)	**x**	**x**				
(3.3)		**x**	**x**	**x**		.
(3.3)				**x**	**x**	
(3.5)					**x**	**x**

(3.8)

Note that Var. stands for variable. The **x** marks indicate that the variables are present in the equation that matches the given row. For instance, $\ln p_i^{sat}$, $\ln x_i$ and $\ln p_i$ are related by Eq. (3.3).

The functionality matrix indicates that the solution procedure should evaluate, in order, Eqs. (3.7), (3.3), (3.4) and (3.5). The resulting ASNN is illustrated in Figure 3.7 and given by the following set of equations

$$L_1 = \tanh\left(W_1 T^* + \beta_1\right) \quad [3] \tag{3.9}$$

$$L_2 = W_2 L_1 + \beta_2 \quad [n] \tag{3.10}$$

$$L_R = W_{R,1} \ln(x) + W_{R,2} L_2 = \ln(p) \quad [n] \tag{3.11}$$

$$L_A = \exp(W_A L_R) \quad [n] \tag{3.12}$$

$$L_D = \ln(|W_D L_A|) = \ln(P) \quad [1]. \tag{3.13}$$

Neither the input nor the output variables are normalized, however, the absolute value of T is reduced according to $T^* = T/1000$.

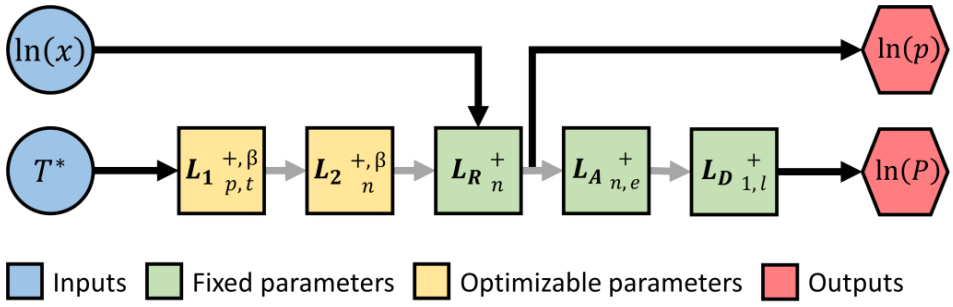


Figure 3.7. Visual representation of the ASNN that represents the VLE.

Layers L_1 and L_2 are the representation of the shallow neural network substructure that implicitly calculates the logarithm of the saturation pressure as a function of T^*. According to Eq. (3.3), $\ln(x)$ and $\ln(p^{sat})$ must be linearly combined in order to estimate $\ln p$. The auxiliary equation (Eq. (3.4)) is represented in layer L_A, and Dalton's law corresponds to layer L_D. Note that Layer L_2 could have been merged with the Raoult's law layer.

The ASNN could have utilized 4 layers if physics laws were not represented in their logarithmic form. This emphasizes that an adequate selection of the input and output variables is essential since it can needlessly complicate the ASNN architecture.

The total number of sets of parameters (W) in Eqs (3.9) - (3.13) must be equal to the number of connections that feed a non-output layer (six in this example: W_1, W_2, $W_{R,1}$, $W_{R,2}$, $W_{A,}$ and W_D. The number of neurons in layers L_2, L_R, L_A, and L_D is equal to the

number of components (n) since Raoult's law is used n number of times ($n = 2$ in this example). The number of neurons in layer L_1 was determined by a trial-and-error procedure and it was found that using 3 neurons provided excellent prediction capabilities of the model without incurring into overfitting.

Step 6. Training the ASNN. Since layers L_R, L_A and L_D represent two physics laws and an exact mathematical relationship, the transformation functions and their parameters are fixed as follows:

- Layer L_R: the set of parameters available in this layer are $W_{R,1}$ and $W_{R,2}$. Both weight matrices should be equal to an identity matrix with $n \times n$ dimensions.
- Layer L_A: the set of parameters that correspond to this layer is W_A which should be an identity matrix of $n \times n$ dimensions.
- Layer L_D: W_D is an all-ones vector of $1 \times n$ dimensions so that its dot product with L_A is equivalent to Dalton's law (Eq. (3.5)).

The nonfixed parameters (W_1 and W_2) were trained afterwards using the Bayesian regularization algorithm [83–85]. The deep learning toolbox from Matlab 2020b was used for implementing and optimizing the ASNN.

The data was divided using the default parameters from Matlab (70 % for training, 15 % for validation and 15 % for testing). The learning rate was set to 0.001 and the weights of each observation were divided by the square of the experimental value in order to emulate an AARD performance function. No special hyperparameter tuning was required.

The NNP approach showed adequate representation of the ideal VLE system. The absolute average relative deviation (AARD) of the partial pressures p_1 and p_2 were

estimated to be 0.09 % and 0.08 % respectively while for total pressure P is 0.10 %. One of the main features of NNP is that it is possible to retrieve and analyze the parameters calculated inside the ASNN. In this case, if the output values of L_2 are analyzed, one can notice that they correspond to those of $\ln(p^{sat})$, hence, the values of p^{sat} can be compared to those calculated with Aspen Plus. In this example, the AARD of p_1^{sat} and p_2^{sat} are 0.12 % and 0.15 %, respectively. These values are satisfactory considering that the ASNN was not explicitly trained for modelling these values.

3.5.1.2. Entropy of the modelling process

It was discussed in subsection 3.3.4 that providing a structure to an ANN can reduce the entropy of the optimization process. Thus, sharpening the distribution of the fitting parameters. We tested this hypothesis by comparing the parameter distribution values of the ASNN (Figure 3.7) and a shallow neural network (SNN) with 4 artificial neurons (see Figure 3.8). In this study, the parameters fixed in step 6 (e.g., W_A or W_D) were fitted by the ANN training algorithm instead of being fixed. Moreover, biases were added to Eqs. (3.11) - (3.13) and obtain, respectively,

$$L_R = W_{R,1}\ln(x) + W_{R,2}L_2 + \beta_R \quad [n] \tag{3.14}$$

$$L_A = \exp\left(W_A L_R + \beta_A\right) \quad [n] \tag{3.15}$$

$$L_D = \ln\left(\left|W_D L_A + +\beta_D\right|\right) \quad [n]. \tag{3.16}$$

The biases were added for the purpose of this study. The ASNN trained in this study is formed by Eqs. (3.9), (3.10), (3.14), (3.15), and (3.16). The optimized parameter

distribution was analyzed by training both neural networks 1000 times with the same dataset used in subsection 3.5.1.1.

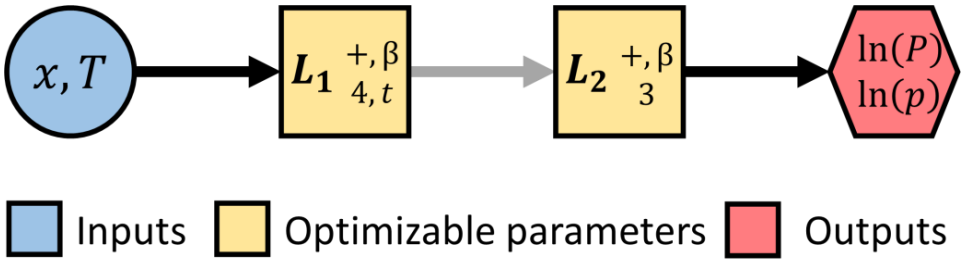


Figure 3.8. Shallow neural network (SNN) architecture of the benzene (1) – toluene (2) VLE system.

Figure 3.9 a) shows the empirical probability distribution of the weight parameters of the second layer of the shallow neural network (W_2^*). This distribution is broader and flatter than the optimized parameter distributions of the ASNN (Figure 3.9 b) and 9 d)). These characteristics imply that there is a low likelihood that the same fitting parameter values will be obtained during different training procedures. Therefore, the W_2^* parameters have high entropy and consequently low interpretability. It is not possible to obtain definite conclusions from their numerical values despite having good agreement with the dataset (the AARD of all shallow neural networks is between 0.05 and 0.25 %). If more artificial neurons are utilized in the SNN, more variability of the fitting parameters are expected, hence, the chances of interpreting overfitted ANNs are even lower than non-overfitted ANNs.

The empirical probability distribution of the optimized parameters of the ASNN are presented in Figure 3.9 b) – d). Figure 3.9 b) shows two highly sharp distributions that correspond to $W_{R,1}$. These distributions indicate that there is an 80 % chance that the

optimized $W_{R,1}$ parameter values are either 0 or 1. It is worth mentioning that the models whose $W_{R,1}$ parameters are 0 and 1 have an AARD less than 1 %. This suggests that the optimized parameters which have the highest empirical probability are the ones with better predictions.

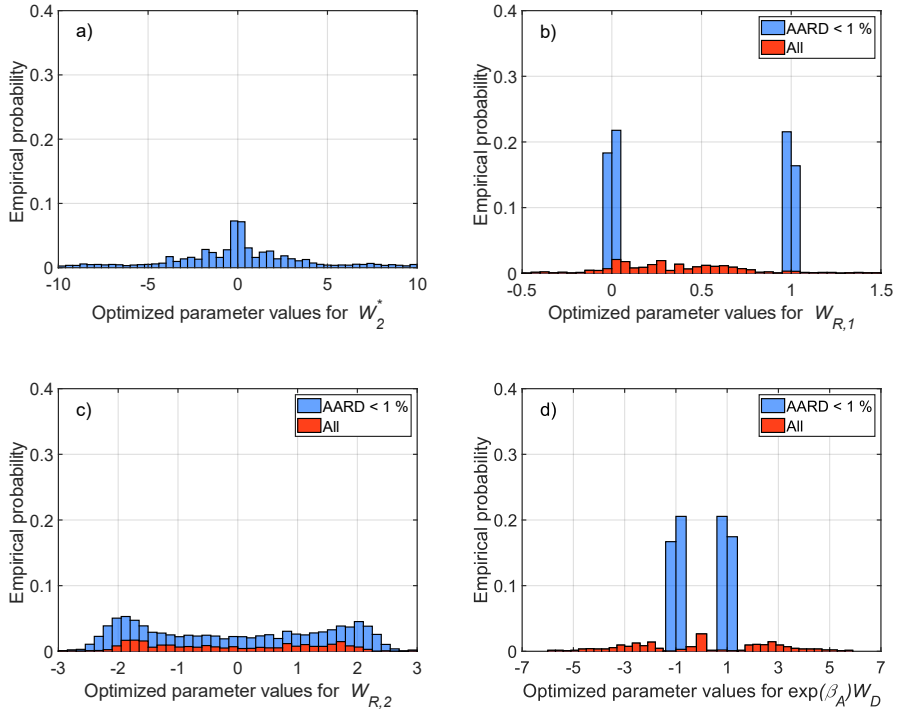


Figure 3.9. Histograms of the empirical probability of the optimized parameter values of: a) W_2^* (shallow neural network), b) $W_{R,1}$, c) $W_{R,2}$ and d) combined effect of β_A and W_D.

From now on, we refer to the model developed in section 4.1.1 as the fixed parameter neural network (FNN) while the models developed in this section are called unfixed parameter neural networks (UNN).

If one analyzes the parameters found in W_2, $W_{R,2}$, W_D, β_R, and β_A, it is not possible to obtain definite conclusions about the parameter behavior due to their blunt distribution. For example, Figure 3.9 c) shows a broad and flat empirical probability distribution of the $W_{R,2}$ parameters. The characteristics of the distribution shown in Figure 3.9 c) resemble those of a random probability distribution, hence, neglecting the possibility to draw objective conclusions from it. The results of Figure 3.9 c) indicate that $W_{R,2}$ do not interfere with the prediction capabilities. The accurate predictions of the UNN together with the fact that the $W_{R,2}$ and β_D values agree with those of the FNN indicate that these models, despite being different, perform equivalent calculations. For example, due to lack of information in the ASNN, the parameters $W_{R,2}$, and β_R become highly dependent on the parameter values of W_2 and $W_{R,2}$. In order to interpret this, let us combine Eqs. (3.10) and (3.14)

$$L_R = W_{R,1}\ln(x) + W_{R,2}(W_2 L_1 + \beta_2) + \beta_R. \tag{3.17}$$

Eq. (3.17) indicates that the product $W_{R,2}\left(W_2 L_1 + \beta_2\right)$ has the condition of having infinite number of solutions. This condition suggests that the parameter distribution of these two sets of parameters should be relatively flat (as shown in Figure 3.9 c)). This signifies that the parameter $W_{R,2}$ has high entropy, therefore, the terms $W_{R,2}\left(W_2 L_1 + \beta_2\right) + \beta_3$ in Eq. (3.17) become equivalent to $\ln(p^{sat})$. The difference between the $\ln(p^{sat})$ values and the values calculated from $W_{R,2}\left(W_2 L_1 + \beta_2\right) + \beta_3$ is quite low (0.03 - 0.06 %). This suggests that the ASNN "knows" that $\ln(p^{sat})$ must be calculated, however, it is difficult to interpret it because we provided a sequence of operations with poor interpretability.

There is a similar effect in the Dalton's law section of the ASNN. The bias β_A (constant) in Eq. (3.15) is transformed to its exponential form and multiplies the parameters in W_D. If the overall interaction of these parameters is analyzed, two narrow and definite distributions can be seen (Figure 3.9 d)). By only considering the UNN with an AARD < 1 %, there is a 50 % probability that the optimized parameter is -1 and 50 % that it is 1. This apparent disagreement with the FNN (where all parameters should be equal to +1) is caused by the fact that the argument of the logarithmic function uses absolute values, thus making -1 and 1 equivalent.

3.5.2. Case study 2: two-product separator

3.5.2.1. Sequential solution algorithm

Step 1. Setting the modelling objective and data collection. The objective of this example is to develop a model of a vapor-liquid separator of a multicomponent non-ideal mixture (see Figure 3.10). This model is expected to predict the molar flows of the components in each product stream. The components present in the mixture are methanol (1), ethanol (2) and water (3).

A dataset containing the results of 50 simulations was utilized to train the ASNN. This dataset was generated with the equilibrium two-phase separator model and the NRTL thermodynamic package available in Aspen Plus v8.6. The input variables used were the molar fractions in the feed stream z_i ($0 \leq z_i \leq 1$), the separator temperature T ($343\ K < T < 413\ K$) and the vaporization fraction Ψ ($0 \leq \Psi \leq 1$). The values of the independent variables z_i and T were randomly generated.

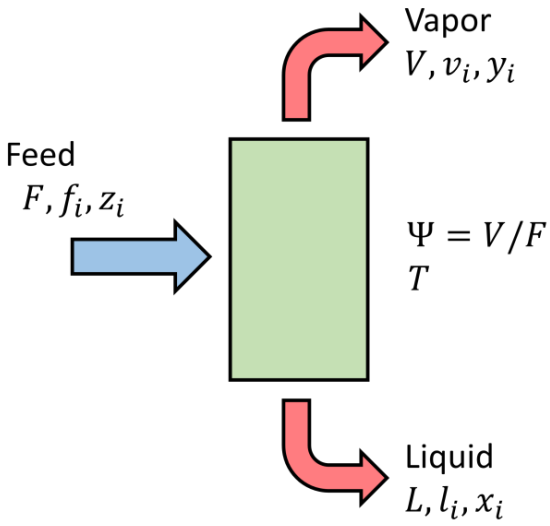


Figure 3.10. Diagram of the modelled two-phase separator.

Step 2. Definition of the assumptions, physics laws and constraints. The mass conservation equation for each component in the vapor-liquid separator can be written in terms of molar fractions and the separation or vaporization fraction $(\Psi = V/F)$:

$$z_i F = y_i \Psi F + x_i (1 - \Psi) F \tag{3.18}$$

where, z_i, y_i and x_i are the component i molar fractions in the feed, vapor, and liquid streams, respectively. The total molar flows of the feed, vapor and liquid streams are respectively F, V and L. Assuming thermodynamic equilibrium, the molar fraction of the vapor and the liquid phases should hold the following relationship (independently of the magnitude of F)

$$\lambda_i = \frac{x_i}{y_i}, \tag{3.19}$$

where λ_i is the inverse of the distribution coefficients (k-values) commonly used in vapor liquid equilibrium calculations. By combining Eqs. (3.18) - (3.19) and multiplying the result by Ψ/F, the following equation is obtained

$$v_i^* = y_i \Psi = z_i \Psi \left[\frac{1}{\Psi + \lambda_i(1-\Psi)}\right]. \tag{3.20}$$

The operation Ψ/F was performed in order to use scalable variables instead of total molar flows. Therefore, the "reduced" vapor flow v_i^* is independent of the magnitude of F (valid for systems in equilibrium). The reduced mass balance for each component in the liquid phase is

$$l_i^* = z_i^* - v_i^* \tag{3.21}$$

According to the first principles, the DoF are 4 (T, Ψ and 2 independent molar fractions).

Step 3. Identification of the uncertain/surrogate sections of the model. Although the λ_i parameters are not explicitly semi-empirical, they are calculated in mechanistic models with the product of semi-empirical models (e.g., activity coefficient models or pure component saturation pressure), hence, this is a source of entropy. Therefore, λ_i can be calculated with a universal approximator substructure. Consequently, a shallow neural network can be used to predict the λ vector

$$\lambda = \eta\,(z, T, \Psi). \tag{3.22}$$

The Gibbs' phase rule establishes the inputs to η in Eq. (3.22), therefore, it should not be biased.

Steps 4 - 5. Equation decomposition and construction of the ASNN architecture. Since Eqs. (3.20) - (3.22) are somewhat complex, it is necessary to break down the equations and organize the solution algorithm. The detailed procedure of the equation decomposition, structural analysis, and equation arrangement of this example can be found in the supplementary information.

The ASNN that models the two-phase separator is illustrated in Figure 3.13 and is described by the following set of equations

$$L_1 = \tanh(W_1 I_1) \quad [p] \tag{3.23}$$

$$L_2 = \left(W_{2,1}\Psi^*\right) \odot \left(W_{2,2}L_1\right) \quad [n] \tag{3.24}$$

$$L_3 = 1 \oslash \left(W_{3,1}\Psi + W_{3,2}L_2\right) \quad [n] \tag{3.25}$$

$$L_4 = \left(W_{4,1}\Psi\right) \odot \left(W_{4,2}z\right) \odot \left(W_{4,3}L_3\right) \quad [n]$$
$$L_4 = 0 \quad if \quad L_4 < 0 \quad [n] \tag{3.26}$$

$$L_5 = W_{5,1}z + W_{5,2}L_4 \quad [n]$$
$$L_5 = 0 \quad if \quad L_5 < 0 \quad [n] \tag{3.27}$$

$$L_6 = W_{6,1}z + W_{6,2}L_5 \quad [n], \tag{3.28}$$

where I_1 is a vector containing z, T and Ψ, and $\Psi^* = 1 - \Psi$. Layer L_1 and the operation $\left(W_{2,2}L_1\right)$ in L_2 is the shallow neural network substructure (representing Eq. (3.22)). Eq. (3.20) was decomposed into smaller operations (Eqs. (3.24) - (3.26)). For instance, $\left(W_{2,1}\Psi^*\right) \odot \left(W_{2,2}L_1\right)$ in Eq. (3.24) is equivalent to the $\lambda(1-\Psi)$ term in Eq.

(3.20). Layers L_2 to L_6 have 3 neurons each because there are $n = 3$ components and L_1 has p =8 neurons (adjusted by a trial and error procedure).

The homogeneous characteristic of the ASNN (Figure 3.11) comes from the fact that z is used as input in I_1 in order to estimate λ (because of the equilibrium assumption). If the molar flows f were utilized instead of z, it would mean that the ASNN is assuming a non-homogeneous behavior, therefore, the system could not be in thermodynamic equilibrium. Conversely, if the input z is substituted with the molar flows f, the predicted values molar flows would not be reduced molar flows (i.e., v and l instead of v^* and l^*).

Step 6. Training the ASNN. The architecture of the ASNN is composed of 4 input layers, 6 hidden layers and 2 output layers. The hidden layers L_1 and L_2 contain adjustable parameters while layers L_3, L_4, L_5 and L_6 containg fixed parameters layers (some of the parameters in L_2 are fixed and some are adjustable). In order to represent the first principles equations, the following considerations must be done:

- The parameter weight matrices $W_{3,2}$, $W_{4,2}$, $W_{4,3}$, $W_{5,2}$, and $W_{6,2}$ should be the unit matrix with 3 x 3 dimensions for this ternary mixture.
- The parameter weight matrices $W_{5,1}$ and $W_{6,1}$ are the negative identity matrix with 3 x 3 dimensions.
- The parameter weight vectors $W_{2,1}$, $W_{3,1}$ and $W_{4,1}$ should be a unit vector with 3 x 1 dimensions.
- Only the inputs in layer I_1 are normalized using the “mapminmax” function from Matlab (transforms all variables to values between -1 and 1). Conversely, the remaining layers should not be normalized.

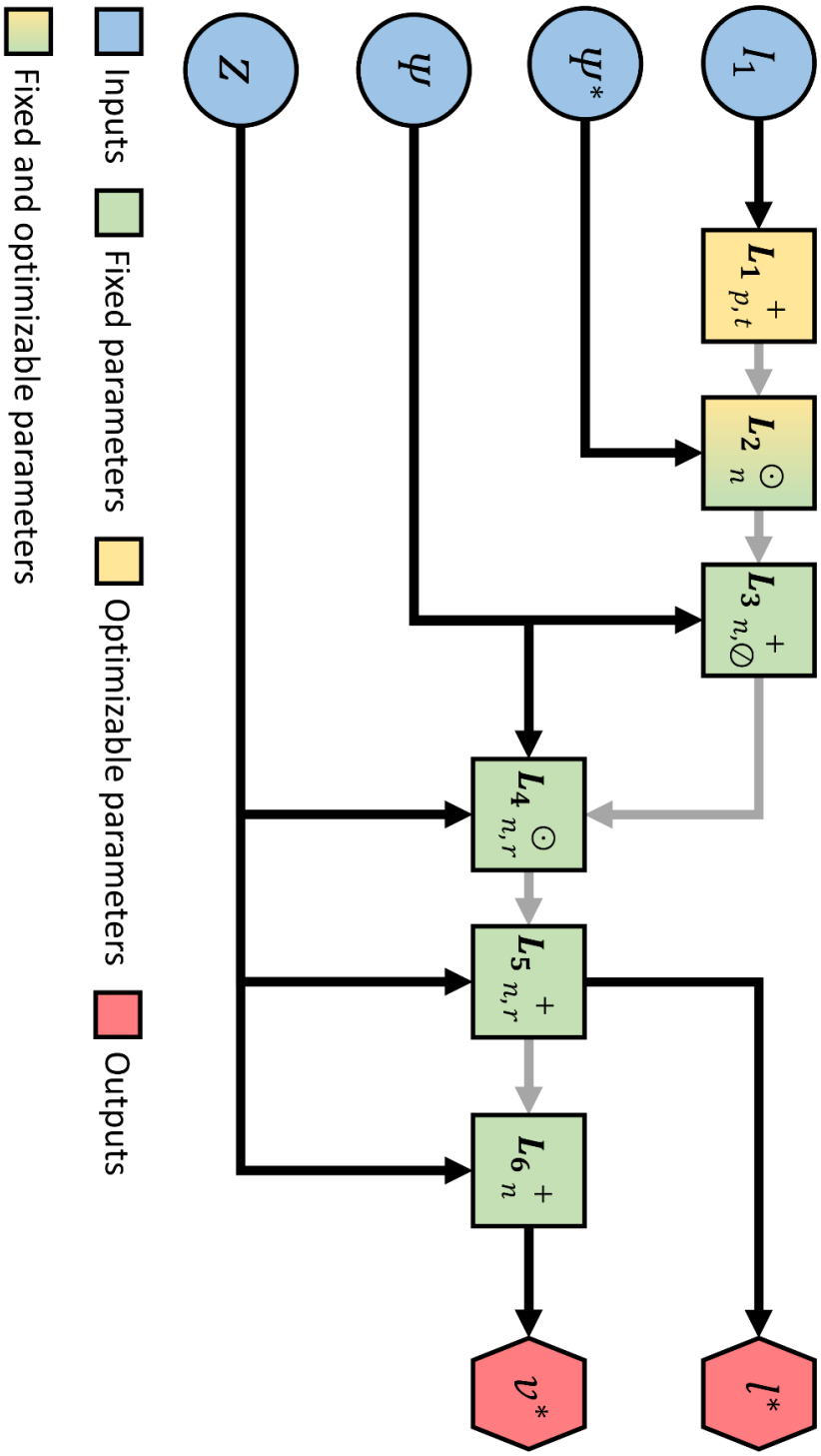


Figure 3.11. ASNN architecture of the two-product separator using the exact first principles representation.

The data used to train the model was divided in the training, validation, and test sets. The training set has 70 % of the data points while the validation and test sets are 15 % each. In order to distribute the error evenly across all the datapoints, it was necessary to weight the observations. The datapoint weights for l^* and v^* were calculated with $(1/l^*)^2$ and $(1/v^*)^2$ respectively. Due to the stochastic nature of the neural network training algorithms, the ASNN was trained 100 times and subsequently the model with the lowest AARD was selected. A learning rate of 0.0001 was utilized.

An extra validation dataset consisting of 100,000 datapoints was generated in the same fashion as the training dataset. This was done to test the generalization capabilities of the ASNN on the entire parameter space. The AARDs of important process variables are presented in Table 3.1 (under the column ASNN, ± 0 %). The molar fractions were calculated by normalizing the corresponding l^* and v^*. The AARDs between the ASNN and Aspen's calculations are small considering that few datapoints (50 runs) were used for the training. It is possible to have better predictions if more training data is used. For example, if using 300 datapoints, the average error can be reduced to 0.2 %. In view of the small AARD, it is reasonable to consider the proposed ASNN as a suitable surrogate model for flash separation of multicomponent mixtures. Utilizing ASNN for overriding the need for iterative calculations may speed-up complex equilibrium calculations with several components.

It is remarkable that the ASNN performance is so high despite the fact that a small dataset (only 50 simulations) made with a random sampling method was used. In the context of surrogate modelling, this suggests that using NNP together with more advanced sampling methods (e.g., [86,87]) can provide more extrapolable and robust models.

It is important to note that strictly applying the NNP method would yield an ASNN with 5 hidden layers instead of 6. The purpose of adding layer L_6 was to verify the robustness of the NNP method. Layer L_6 computes v^* a second time (v^* is computed in L_4 as well) in order to ensure that vector v^* does not provide negative values. It was observed that the results do not substantially change if the last layer is removed, in fact, if the mass balances are not solved twice in the ASNN, only 0.008 % of the solutions have an error in the mass balance larger than 10^{-15} while in the case with 6 layers is 0 %. Therefore, it can be determined that the method is robust.

Table 3.1. AARD (%) between the neural networks and the extra validation dataset when using different noise levels in the training dataset.

Variable	NNP – based hybrid models			Serial – based hybrid models	
	ASNN, ± 0 %	ASNN, ± 5 %	ASNN, ± 20 %	SHM, ± 5 %	SHM, ± 20 %
l^*	1.4	3.1	5.0	4.4	8.6
v^*	0.5	1.7	4.0	2.0	6.7
x	1.1	2.9	4.6	4.1	7.5
y	0.4	1.6	3.8	1.9	6.3
Average	0.9	2.8	4.3	3.1	7.3

As discussed earlier, ASNNs can be applied for fitting of experimental or process data where noise is expected. In order to test the proposed model, randomly distributed noise was added to the training dataset (the validation dataset remains without noise). The results of training a model with a noisy data set are shown in Table 3.1 ((ASNN, ± 5 % and ASNN, ± 20 %). The noise was added to the reduced molar flows l^* and v^*, then the values of x, y, Ψ and λ were recalculated. It can be observed that, as expected, the average error increases when the noise in the datapoints is increased. Despite the aggregated noise, the difference between the trained ASNN and the validation dataset remains within a reasonable range.

A hybrid model based on the serial paradigm was developed in order to compare it with the ASNN. The columns labeled as "SHM, ± 5 % and SHM, ± 20 % in Table 3.1 report the deviation of models generated by a serial hybrid model framework (Figure 3.4 a)). The λ_i parameters were fitted through a shallow neural network (SHM) with 8 artificial neurons (same as in the ASNN). The reduced product flows l^* and v^* were calculated with the rigorous mass balances while the molar compositions (x and y) were calculated by normalizing l^* and v^*. The NNP-based models outperform the prediction capabilities of the coupled hybrid models. In cases where the noise is high (± 20 %), the AARD of the ASNNs can be 40 to 50 % lower than when compared to the serial hybrid models. We consider that there are two causes of the superior performance of the ASNNs over the serial hybrid models. The first cause is due to the ASNN architecture which helps the optimization process. Secondly, the numerical differentiation error negatively affects the optimization step of the serial hybrid model.

It is worth mentioning that the proposed ASNN structure can also work for a $P\Psi$ flash problem as well. The only difference would be the input layer I_1, where T should be

substituted by P. On the other hand, the solution of a PT flash would require a different ASNN architecture. Due to the characteristics of the PT flash problems, it has been suggested in the literature [88], to use regression and classifier neural networks together to solve a PT flash problem. The authors reported that over 25 million datapoints were used to train the model for a binary mixture. This points out that using the NNP approach for a TP flash will require designing a possibly more complex architecture.

The generality and easy adaptability of ASNN structures to other processes is a practical advantage. In particular, the hidden layers with fixed parameters will be the same as in Figure 3.11. This means that as long as the modelled system has a feed stream and two product streams, the ASNN structure can be utilized independently of the modelled process. For example, the same ASNN architecture was utilized for modelling the processes shown in Figure 3.12. The ASNN architecture for modelling the flash separator and the stripping process (Figure 3.12 a)) are quite similar, since the only differences are the involved components (the rich amine stream has CO_2, MEA and H_2O) and the input variables to layer I_1 (molar fractions (z), temperature in the reboiler (T_R), feed temperature (T_F), and the bottoms to feed ratio (Ψ)).

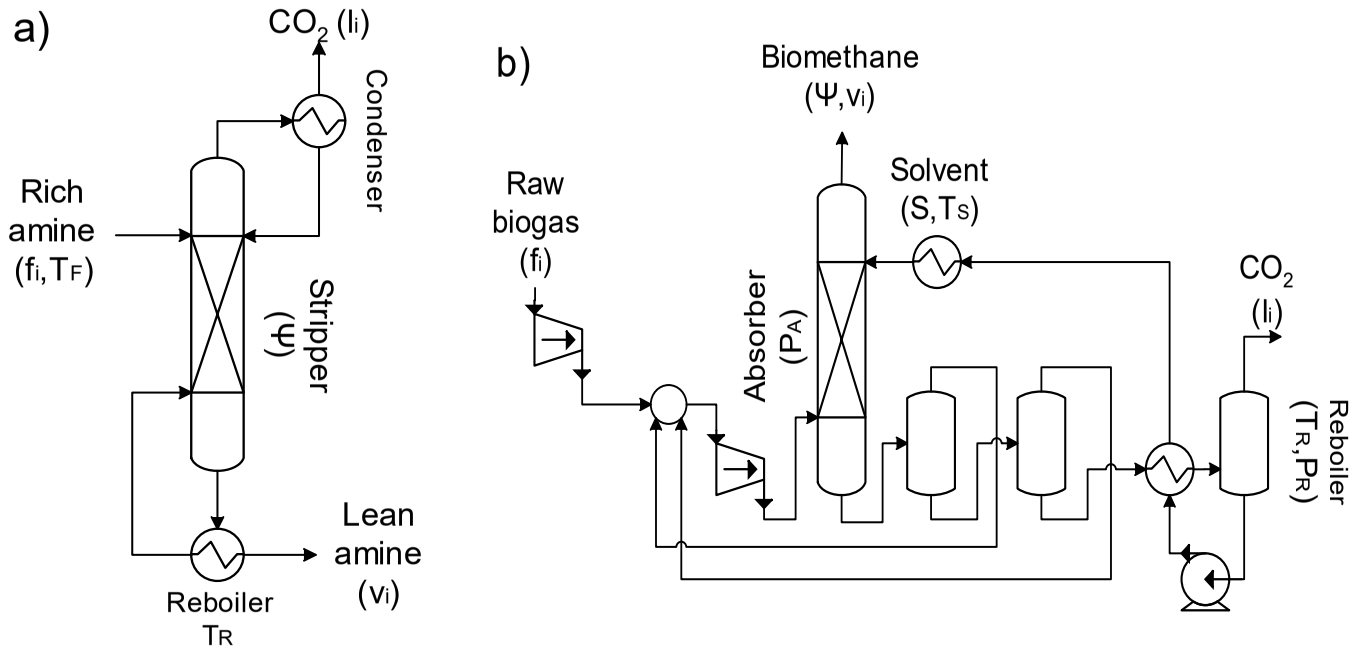


Figure 3.12. Additional processes that can be modelled by utilizing the ASNN architecture shown in Figure 3.11. a) Stripper column, and b) Biogas upgrading process.

The process shown in Figure 3.12 b) has two inlet components in the feed (the raw biogas has CH_4 and CO_2) and 6 process parameters which are the reboiler temperature (T_R), reboiler pressure (P_R), solvent flow (S), solvent temperature (T_S), absorber pressure (P_A) and biomethane to raw biogas ratio (Ψ). Therefore, the layer I_1 has 8 input parameters and layers L_2 to L_7 have 2 neurons. The ASNN were trained by utilizing data generated from models validated in previous works [89–91]. The same ASNN architecture was used for both processes and showed an AARD of 1-2 % in both cases (8 artificial neurons were used in L_1, however, the AARD can be lower if more parameters are included).

There are important points that must be considered when applying the two-product separator ASNN architecture to other processes:

- In cases where the process is quite complex or there is more available data, additional layers can be inserted in between or before layer L_1 in Figure 3.11. Adding more layers should not modify layers L_2 to L_7 (except for the number of neurons which must be equal to the number of components).
- As opposed to the flash process, the λ_i parameters might not have a formal definition in other processes (e.g., distillation or gas stripping). Despite this, the λ_i parameter indicates that, independently of the process, there is always a relationship between the component compositions of the two products that is independent of the extensive properties of the system.
- If the process is modelled with the ASNN structure but the mass balances do not behave homogeneously, then the proposed ASNN architecture might not be optimal. However, knowing that the system does not behave

homogeneously is also valuable modelling information that can be utilized in order to modify the ASNN to account for extensive variables.

3.5.2.2. Matrix inversion algorithm

The application of NNP using a sequential algorithm is feasible for relatively simple models. However, for moderate to large processes, solving conservation equations sequentially is not ideal. To overcome this conundrum, this section utilizes a matrix inversion solution algorithm to solve the linear equations like mass balances. Steps 1 and 2 of the NNP method are the same as in subsection 3.5.2.1.

Step 3. Identification of the uncertain/surrogate sections of the model. The high entropy section (η) of the ASNN is set to substitute the square bracketed term in Eq. (3.20). Therefore, the reduced vapor molar flows are given by

$$v^* = z \bigodot \Psi \bigodot \eta\,(z, T, \Psi). \tag{3.29}$$

Steps 4 - 5. Equation decomposition and construction of the ASNN architecture. Applying the matrix-inversion solution algorithm to the construction of an ASNN must be done in two parts. The first part requires to organize the non-linear equations according to a sequential solution algorithm. The second part uses the values computed by the nonlinear sections of the model to solve the linear system of equations given by the mass balances. To solve the linear system of equations, the user should represent the system of equations such that

$$A = M^{-1}C, \tag{3.30}$$

where A is the solution vector containing the mass/mole flows, M is the characteristic mass balance matrix, and C is the constant vector.

In this example, the constant C must be given by the output of the nonlinear section of the model (v^*) and the known input mass flow variable (z). Considering this, the mass balance matrix M and its corresponding constant vector C can be expressed as

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, \qquad C = \begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \\ z_1 \\ z_2 \\ z_3 \end{bmatrix}. \tag{3.31}$$

The first three columns of M correspond to v_i^* while the last three columns to l_i^*. Therefore, the first three rows of M ensure that v_i^* is equal to the output of the nonlinear section of the model and the last three rows guarantee that $f_i^* = v_i^* + l_i^*$ (separated version of Eq. (3.21)). In order to apply this approach to other processes it is necessary to update the mass balance matrix M, the constant vector C and the variable vector A in accordance with the process flowsheet and the components involved.

Once the M and C are identified, the construction of the ASNN consists in first placing all the hidden layers corresponding to the nonlinear section of the model (already organized) and then adding 2 layers afterwards. The first extra layer must utilize the weight parameters to construct the constant vector C and the second extra layer must transcript Eq. (3.30). Therefore, the parameter matrix of the second layer must always be equal to M^{-1}.

The ASNN architecture is shown in Figure 3.13 and the set of equations is

$$\mathrm{L}_1 = \tanh\left(W_{1,1} I_1 + \beta_1\right) \quad [p] \tag{3.32}$$

$$L_2 = \left(W_{2,1}\Psi\right) \odot \left(W_{2,2} z\right) \odot \left(W_{2,3} L_1\right) \;\; if \;\; 0 < L_2 < 1 \quad [n]$$

$$L_2 = 0 \;\; if \;\; L_2 < 0 \quad [n] \tag{3.33}$$

$$L_2 = 1 \;\; if \;\; L_2 > 1 \quad [n]$$

$$L_3 = \left(W_{3,2} z\right) + \left(W_{3,1} L_2\right) = C \quad [2n] \tag{3.34}$$

$$L_4 = \left(W_4 L_3\right) = M^{-1} C \;\; if \;\; 0 < o_1 < 1 \quad [2n]$$

$$L_4 = 0 \;\; if \;\; L_4 < 0 \quad [2n] \tag{3.35}$$

$$L_4 = 1 \;\; if \;\; L_4 > 1 \quad [2n].$$

Layer L_2 is composed of $n = 3$ neurons because it predicts v_1^*, v_2^* and v_3^*. L_3 and L_4 have 6 neurons because they carry the information of both l_i^* and v_i^*. Figure 3.13 and Eq. (3.33) show a product between the universal approximator predictions (composed of L_1 and part of L_2) with Ψ and z. The advantage of using Eq. (3.29) over Eq. (3.20) is that it is no longer assuming equilibrium between both phase. Eq. (3.29) guarantees that: in the absence of a vapor phase no vapor molar flows different than 0 will be predicted, and a positive component molar flow cannot be predicted if it is not present in the feed.

Optional modification to the ASNN: In order to aid that the predicted v^* and l^* values are between 0 and 1, the linear saturation nonlinear function (satlin in Matlab) was added to L_2 and L_4 (Eqs. (3.33) and (3.35)).

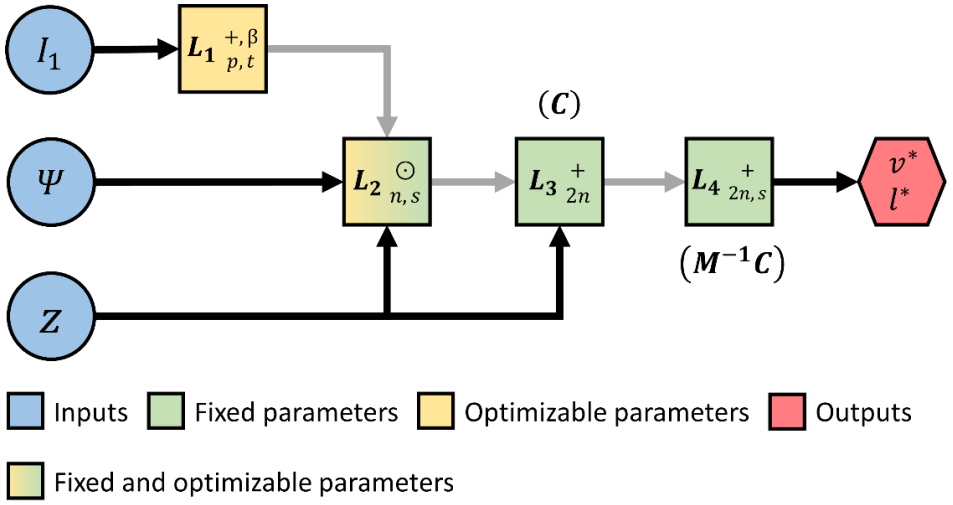


Figure 3.13. ASNN architecture of the two-product separator using the matrix inversion solution algorithm.

Step 6. Training the ASNN. In order to transcript Eq. (3.29), the weight parameter matrix $W_{2,1}$ was fixed to be an all-ones vertical vector with n elements and $W_{2,2}$ was fixed to an identity matrix with n x n elements. In order to construct the C vector, the following weight parameter matrices were defined

$$W_{3,1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad W_{3,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.36}$$

The purpose of $W_{3,1}$ is to bypass the values of z to the lower part of vector C while $W_{3,2}$ transfers the values estimated by Layer L_3 in order to formulate the upper part of vector C. The parameter matrices in layer L_3 are used as mathematical artifices to reshape the C vector to solve the linear system of equations.

In order to understand the behavior of the ASNNs as a function of the number of training parameters, several ASNNs models were formulated. The results are presented in Table 3.2. The universal approximator substructure indicates how many hidden layers with a sigmoid function were utilized (Eq. (3.32)). For example, model #1 has one hidden layer with 8 neurons while model #7 has 2 layers with sigmoid functions (one with 8 neurons and one with 4 neurons which means that there is an additional layer before L_2). The universal approximator substructures are connected to layer L_2 (Eq. (3.33)). The models with 50 datapoints used the same training database as in section 3.5.2.1 while the ones with 200 and 500 datapoints utilized new databases.

Each model was trained 20 times and the model with the lowest AARD was selected. The weights of the datapoints were calculated with $(1/l^*)^2$ and $(1/v^*)^2$ in order to optimize for the AARD instead of the mean square error. Note that the input layer I_1 that feeds the universal approximator substructure uses a normalized input while input layers I_2 and I_3 do not. The models were trained using the Bayesian stochastic optimization algorithm available in the Deep Learning Toolbox of Matlab 2020b.

The AARD presented in Table 3.2 was calculated with the same extra validation dataset used in section 3.5.2.1. Comparing the AARD (models 1 and 2 in Table 3.2) against the results of Table 3.1 shows that utilizing the ASNN1 (Figure 3.11) has, in general, better prediction capabilities than the ASNN2 (Figure 3.13). This is

understandable since the ASNN1 architecture has more information regarding the relationship between the input and output variables than the ASNN2. This reveals a clear tradeoff between the ASNN complexity and its prediction capabilities. The more physics information is provided to the ASNN architecture, the less data and fitting parameters will be needed; thus, the model will have a lower entropy. In fact, adding more parameters without including more datapoints is detrimental to the model performance since it overfits the model to the training data (compare models #1 and #2 in Table 3.2). One can notice that in order to have a similar AARD between the ASNN1 and ASNN2, the number of datapoints is increased roughly tenfold (from 50 to 500) and the number of neurons in the hidden layer is increased two times (from 8 to 16).

The type of universal approximator substructure can be different from a shallow neural network. In fact, Table 3.2 shows that the model AARD can be reduced 50 % if a deep neural network substructure that has 2 hidden layers is utilized. Moreover, better prediction capabilities can be achieved than in the ASNN1 model if more datapoints and a deep neural network substructure is utilized.

It should be remarked that applying NNP with exact first principles representations for every equipment in a large process can be a time-consuming task, therefore, in many cases a compromise between accuracy and practicality will emerge. For example, in a process with multiple unit operations and subprocesses and large amount of data, it is more practical to apply the inverse matrix approach since the mass balances will be solved in a single step instead of multiple layers. On the other hand, for systems with limited amount of measured data or where keeping certain mathematical relationships is paramount to maintain the cohesion between the data, one should explicitly represent all the first-principles relationships.

Table 3.2. Characteristics and results of every ASNN2 architecture formulated.

ID	Universal approximator architecture	No. of datapoints	AARD / %
1	8	50	12
2	16	50	23
3	8	500	2.0
4	16	500	1.2
5	8x4	50	10
6	16x4	50	6.4
7	16x8	50	6.7
8	32x8	50	8.2
9	16x4	500	1.0
10	16x8	500	0.9
11	32x8	500	0.8

3.6. Conclusions

This work presents the Neural Network Programming (NNP) hybrid modelling paradigm. It integrates a first principles modelling approach with a data-driven algorithm. The models developed with NNP are Algorithmically Structured Neural Networks (ASNNs). The idea is to generate a first principles model, decompose it and transcript the mathematical equations to an ASNN. This allows a physically coherent over the entire solution space, including limit cases (e.g., ASNNs do not predict a positive molar flow of a component if it is not present in the mixture). Due to the features of ASNNs, there is no need to utilize sophisticated performance functions in order to account for physics constraints and the hyperparameter tuning is not critical as in data-driven models. Since the first principles equations are in-built in the ASNN, the data is more efficiently utilized for fitting process parameters rather than trying to rediscover physics concepts that must hold. This causes the ASNNs to exhibit superior performance than black-box and conventional serial hybrid model configurations.

Three examples on how to model chemical engineering problems with the NNP method are presented in this work. The first example consisted in the rigorous representation of the physics equations and relationships in the same way as how it is done in a mechanistic model (section 3.5.1). Being able to transcript physics laws within an ASNN is of utmost importance in subfields where the models must comply with a large set of requirements in order to be deemed as “correct” (e.g., transport property modelling or thermodynamics). The second example (section 3.5.2.1) shows a more relaxed representation of the first principles equations. It is highlighted that through a structural analysis of the first principles model it is possible to lump several parameters onto a single parameter without disrupting the reliability of the model. The

third example (section 3.5.2.2) presents an alternative approach for performing mass balances where the parameter matrix of a hidden layer is used to solve the mass balances. This approach is of fundamental importance in the application of NNP to large processes since it can be used to model processes with several process streams.

There is a compromise between the complexity of the ASNN architectures and the amount of physics knowledge embedded into them. Therefore, a careful assessment of the assumptions and the information that wants to be predicted with the model must be done before constructing the ASNN. Performing a structural analysis allows the user to select the parts of the model that can be substituted with a universal approximator substructure and how to connect them with the rest of the model. A structural analysis allows ASNNs to be effectively utilized for the formulation of surrogate models by removing the iterative loops that are commonly seen in mechanistic models.

An interesting feature of ASNNs is their transferability between processes with akin characteristics. For example, this work presented the ASNN architecture needed to model a flash separator which was later applied to model biogas upgrading processes without substantial modifications to the original ASNN architecture.

In contrast to the typical hybrid configurations, NNP only relies on using a single auto-differentiable model that complies with the physics laws rather than independently utilizing a first principles model and an artificial neural network model. This allows a more accurate and faster training because no numerical differentiation techniques are utilized. As opposed to mechanistic models, the NNP framework directs the user to develop models in matricial form, which means that multiple simulations can be utilized with a single call to the model function (i.e., instead of utilizing for or while

cycles). We expect that these features will be exploited even further in the coming decades due to the current advances in quantum computing.

If the architecture of an ANN has cohesion with the physics phenomenon description, the parameter entropy will be lower and therefore, there is high likelihood that the model will provide repeatable parameters. NNP can be utilized to discard incorrect assumptions about the physics phenomena.

Interpreting the parameters of a black-box ANN might is an overcomplicated task due to the high variability of the optimized parameters. In other words, it is unlikely to find a physical meaning of parameters whose numerical values are a consequence of the entropy associated to the ANN training. Hence, it is more effective to assume an ANN architecture based on *a priori* knowledge and find a pattern in the optimized parameters rather than using generic ANN architectures. When performing numerical analyses on ANNs, it is not enough to analyze the numerical values of the optimizable parameters in order to interpret them.

Further work of NNP includes the application of this method for the development of consistent thermodynamic and transport property models. Additionally, we hypothesize that NNP can be utilized to develop NNP-based PINNs in order to guarantee the exact execution of first principles equations.

3.7. Supplementary information

This section presents a detailed description of the steps taken in the decomposition and construction of the ASNN used to model the two-product separator presented in section 4.2.1. The equations that are to be decomposed are Eqs. (3.20) - (3.22) whose vectorial representation is

$$v^* = z \odot \Psi \odot \left[\frac{1}{\Psi + \lambda \odot (1 - \Psi)}\right] \tag{3.37}$$

$$l^* = z - v^* \tag{3.38}$$

$$\lambda = \eta\,(z, T, \Psi), \tag{3.39}$$

respectively. First, Eq. (3.37) is decomposed according to the algorithm presented in Figure 3.6. Therefore, the function $\boldsymbol{f}$ to be decomposed is Eq. (3.37) and the result for the iteration steps are shown in Table 3.3.

Only one decomposition iteration can be done for Eq. (3.38) because it is already a linear equation. Therefore, Eq. (3.38) can be “decomposed” as

$$\Upsilon_l = l^* = z^* - v^*. \tag{3.40}$$

Table 3.3. Results of every iteration of the application of the decomposition procedure to Eq.(3.37).

Step 4.2.	Step 4.3.		Step 4.4.		Step 4.5.	Step 4.6
k	Operation	Type	$X_{k,j}$	Y_k^*	Y_k	
1	$(1-\Psi)$	Linear	$X_{1,1}=1$ $X_{1,2}=\Psi$	$Y_1^*=X_1+X_2$	$Y_1=Y_1^*$	No
2	$Y_1 \odot \lambda$	Product	$X_{2,1}=Y_1$ $X_{2,2}=\lambda_i$	$Y_2^*=X_1 \odot X_2$	$Y_2=Y_2^*$	No
3	$\Psi+Y_2$	Linear	$X_{3,1}=\Psi$ $X_{3,2}=Y_2$	$Y_3^*= X_1+X_2$	Y_3 $=1 \oslash Y_3^*$	No
4	$z_i \odot \Psi$ $\odot Y_3$	Product	$X_{4,1}=z$ $X_{4,2}=\Psi$ $X_{4,3}=Y_3$	Y_4^* $=X_1 \odot X_2 \odot X_3$	$Y_4=Y_4^*$	Yes

The set of equations that derives from the decomposition of the first principles equations and the shallow neural network substructure is

$$\Upsilon_1 = 1 + W_0\Psi \quad [1] \tag{3.41}$$

$$\Upsilon_2 = \left(W_{2,1}\Upsilon_1\right) \odot \left(W_{2,2}\lambda\right) \quad [n] \tag{3.42}$$

$$\Upsilon_3 = 1 \oslash \left(W_{3,1}\Psi + W_{3,2}\Upsilon_2\right) \quad [n] \tag{3.43}$$

$$\Upsilon_4 = \left(W_{4,1}\Psi\right) \odot \left(W_{4,2}z\right) \odot \left(W_{4,3}\Upsilon_3\right) = v^* \quad [n] \tag{3.44}$$

$$\Upsilon_l = W_{5,1}z + W_{5,2}v^* \quad [n] \tag{3.45}$$

$$\lambda_0 = \tanh(W_1 I_1) \quad [p] \tag{3.46}$$

$$\lambda = W_{2,2}\lambda_0 \quad [n], \tag{3.47}$$

Where I_1 is a vector formed by z, T and Ψ. Eqs. (3.41) - (3.44) are the decomposed equations that correspond to Eq. (3.37). Eq.(3.45) is the neural network representation of the mass balances (Eq. (3.38)). Eqs. (3.46) - (3.47) are the decomposition equations of Eq. (3.39). The signs of the equations (e.g., Eq. (3.41)) are given by the selection of weight parameters in step 6.

Considering that the selected input variables are z, T, and Ψ, the equations should be solved in the following order: (3.46) → (3.47) → (3.41) → (3.42) → (3.43) → (3.44) → (3.45). The ASNN constructed by following this procedure is illustrated in Figure 3.7 a).

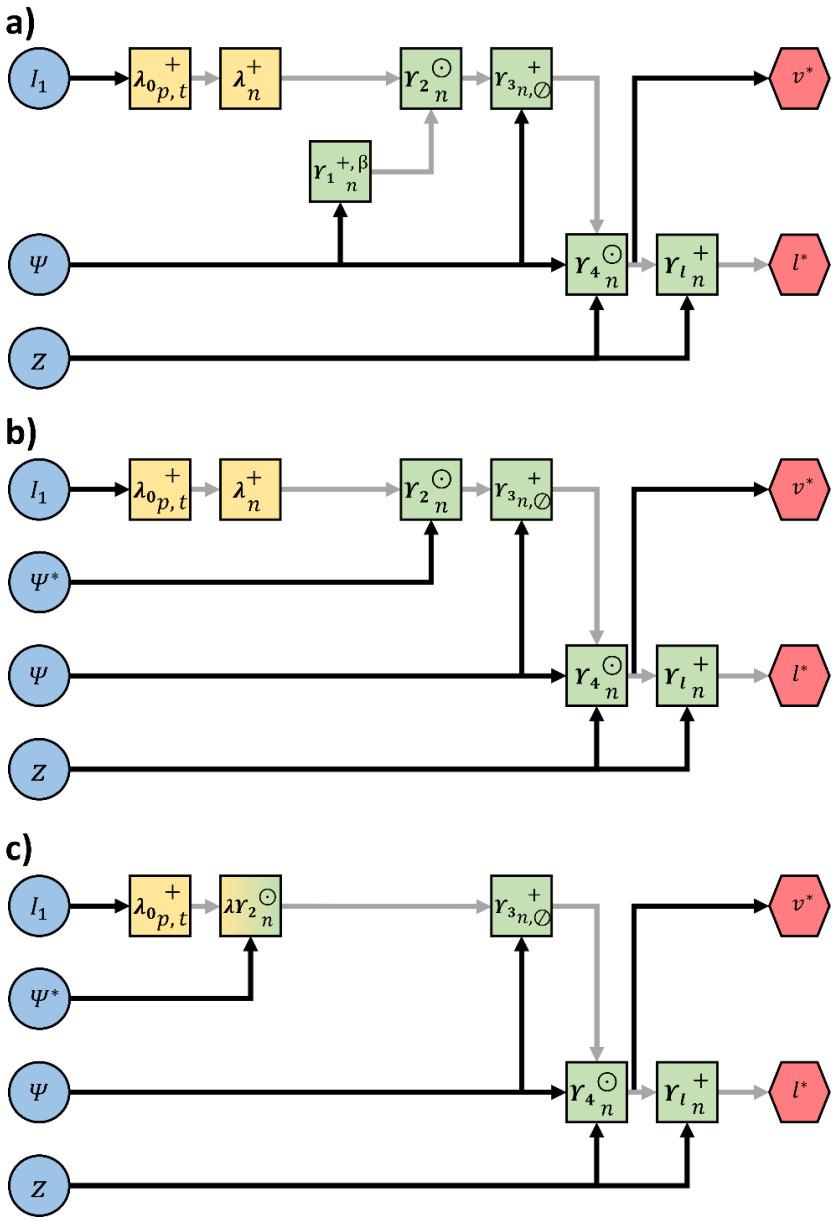


Figure 3.14. Representations of the ASNN architecture: a) unsimplified and unconstrained, b) substituting hidden layer Υ_1 with the input Ψ^*, and c) merging the output of the universal approximator and layer Υ_2.

By comparing Figure 3.11 and Figure 3.14 a), it can be noticed that the ASNNs are different. The reason for this is the fact that the ASNN shown in Figure 3.14 a) does not have any of the optional modifications proposed in step 5, hence it is unsimplified and unconstrained. Note: the notation of the layer names was changed with respect to the manuscript (e.g., from L_3 to Υ_2) in order to be consistent with the notation given in Figure 3.6.

The first modification that can be done is performing Eq. (3.41) outside the ASNN. This can be done because it can be calculated only by using the input variables (z, T and Ψ). Doing so, removes Eq. (3.41) and adds $\Psi^* = 1 - \Psi$ as an input layer (as shown in Figure 3.14 b)). If this ASNN is further simplified by combining the output of the universal approximator and layer Υ_2 into $\lambda\Upsilon_2$, the architecture of Figure 3.14 c) can be obtained. As mentioned in section 4.2.1, the ASNN in the manuscript added an extra layer (L_6), in order to verify the robustness of the NNP method. Therefore, strictly applying the NNP method without the proposed modifications would yield Figure 3.14 c). After the architecture is done, the user can add the ReLU functions to Figure 3.14 c) in order to get Figure 3.11.

3.8. References

[1] V. Venkatasubramanian, The promise of artificial intelligence in chemical engineering: Is it here, finally?, AIChE J. 65 (2019) 466–478. https://doi.org/10.1002/aic.16489.

[2] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L.H. Chiang, M.S. Reis, Recent trends on hybrid modeling for Industry 4.0, Comput. Chem. Eng. 151 (2021) 107365. https://doi.org/10.1016/j.compchemeng.2021.107365.

[3] M. Torrisi, G. Pollastri, Q. Le, Deep learning methods in protein structure prediction, Comput. Struct. Biotechnol. J. 18 (2020) 1301–1310. https://doi.org/10.1016/j.csbj.2019.12.011.

[4] Y. Xu, D. Verma, R.P. Sheridan, A. Liaw, J. Ma, N.M. Marshall, J. McIntosh, E.C. Sherer, V. Svetnik, J.M. Johnston, Deep Dive into Machine Learning Models for Protein Engineering, J. Chem. Inf. Model. 60 (2020) 2773–2790. https://doi.org/10.1021/acs.jcim.0c00073.

[5] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature. 529 (2016) 484–489. https://doi.org/10.1038/nature16961.

[6] O. Vinyals, I. Babuschkin, W.M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D.H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J.P. Agapiou, M. Jaderberg, A.S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J.

Molloy, T.L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, D. Silver, Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature. 575 (2019) 350–354. https://doi.org/10.1038/s41586-019-1724-z.

[7] H.L. Huang, D. Wu, D. Fan, X. Zhu, Superconducting quantum computing: a review, Sci. China Inf. Sci. 63 (2020) 1–32. https://doi.org/10.1007/s11432-020-2881-9.

[8] H. Elmqvist, A Structured Model Language for Large Continuous Systems, Lund Institute of Technology (LTH), 1978. https://lup.lub.lu.se/search/ws/files/4602422/8570492.pdf.

[9] P.C. Piela, T.G. Epperly, K.M. Westerberg, A.W. Westerberg, ASCEND: an object-oriented computer environment for modeling and analysis: The modeling language, Comput. Chem. Eng. 15 (1991) 53–72. https://doi.org/10.1016/0098-1354(91)87006-U.

[10] F.E. Cellier, H. Elmqvist, Automated Formula Manipulation Supports Object-Oriented Continuous-System Modeling, IEEE Control Syst. 13 (1993) 28–38. https://doi.org/10.1109/37.206983.

[11] H.E.S.M. Karl Johan Åström, Evolution of Continuous-Time Modeling and Simulation., Esm. (1998) 1–10. http://www.modelica.org/publications/index_html/papers/esm98his.pdf.

[12] R. de P. Soares, A.R. Secchi, EMSO: A new environment for modelling, simulation and optimisation, in: Comput. Aided Chem. Eng., 2003: pp. 947–952. https://doi.org/10.1016/S1570-7946(03)80239-0.

[13] M. Heitzig, A.A. Linninger, G. Sin, R. Gani, A computer-aided framework for development, identification and management of physiologically-based pharmacokinetic models, Comput. Chem. Eng. 71 (2014) 677–698. https://doi.org/10.1016/j.compchemeng.2014.07.016.

[14] S. Kuntsche, T. Barz, R. Kraus, H. Arellano-Garcia, G. Wozny, MOSAIC a web-based modeling environment for code generation, Comput. Chem. Eng. 35 (2011) 2257–2273. https://doi.org/10.1016/j.compchemeng.2011.03.022.

[15] J. Åkesson, K.E. Årzén, M. Gäfvert, T. Bergdahl, H. Tummescheit, Modeling and optimization with Optimica and JModelica.org-Languages and tools for solving large-scale dynamic optimization problems, Comput. Chem. Eng. 34 (2010) 1737–1749. https://doi.org/10.1016/j.compchemeng.2009.11.011.

[16] D.D. Nikolić, DAE Tools: Equation-based object-oriented modelling, simulation and optimisation software, PeerJ Comput. Sci. 2016 (2016). https://doi.org/10.7717/peerj-cs.54.

[17] J.R. Leal, A. Romanenko, L.O. Santos, Daedalus Modeling Framework: Building First-Principle Dynamic Models, Ind. Eng. Chem. Res. 56 (2017) 3332–3346. https://doi.org/10.1021/acs.iecr.6b03110.

[18] G. Stephanopoulos, G. Henning, H. Leone, MODEL.LA. A modeling language for process engineering-I. The formal framework, Comput. Chem. Eng. 14 (1990) 813–846. https://doi.org/10.1016/0098-1354(90)87040-V.

[19] A.A. Linninger, H. Krendl, TechTool - computer-aided generation of process models (part 1-a generic mathematical language), Comput. Chem. Eng. 23 (1999) S703–S706. https://doi.org/10.1016/S0098-1354(99)80172-0.

[20] R. Bogusch, B. Lohmann, W. Marquardt, Computer-aided process modeling

with ModKit, Comput. Chem. Eng. 25 (2001) 963–995. https://doi.org/10.1016/S0098-1354(01)00626-3.

[21] J. Morbach, A. Yang, W. Marquardt, OntoCAPE-A large-scale ontology for chemical process engineering, Eng. Appl. Artif. Intell. 20 (2007) 147–161. https://doi.org/10.1016/j.engappai.2006.06.010.

[22] M.R. Westerweele, J. Laurens, Mobatec Modeller - A flexible and transparent tool for building dynamic process models, Comput. Aided Chem. Eng. 25 (2008) 1045–1050. https://doi.org/10.1016/S1570-7946(08)80180-0.

[23] M. Fedorova, G. Sin, R. Gani, Computer-aided modelling template: Concept and application, Comput. Chem. Eng. 83 (2015) 232–247. https://doi.org/10.1016/j.compchemeng.2015.02.010.

[24] H.A. Preisig, Constructing an ontology for physical-chemical processes, Comput. Aided Chem. Eng. 37 (2015) 1001–1006. https://doi.org/10.1016/B978-0-444-63577-8.50012-7.

[25] A. Carranza-Abaíd, R. González-García, A Petlyuk distillation column dynamic analysis: Hysteresis and bifurcations, Chem. Eng. Process. - Process Intensif. 149 (2020). https://doi.org/10.1016/j.cep.2020.107843.

[26] A. Dhooge, W. Govaerts, Y.A. Kuznetsov, MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs, ACM Trans. Math. Softw. 29 (2003) 141–164. https://doi.org/10.1145/779359.779362.

[27] L. Faramarzi, G.M. Kontogeorgis, M.L. Michelsen, K. Thomsen, E.H. Stenby, Absorber Model for CO2 Capture by Monoethanolamine, (2010) 3751–3759. https://doi.org/10.1021/ie901671f.

[28] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Phys. (1943) 113–133.

[29] F. Rosenblatt, Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms, 1962.

[30] C.M. Bishop, Pattern Recognition and Machine Learning, Springer Science+Business Media, LLC, 2006.

[31] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control. Signals, Syst. 2 (1989) 303–314. https://doi.org/10.1007/BF02551274.

[32] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks. 2 (1989) 359–366. https://doi.org/10.1016/0893-6080(89)90020-8.

[33] K. Hornik, Approximation Capabilities of Multilayer Neural Network, Neural Networks. 4 (1991) 251–257.

[34] K.I. Funahashi, On the approximate realization of continuous mappings by neural networks, Neural Networks. 2 (1989) 183–192. https://doi.org/10.1016/0893-6080(89)90003-8.

[35] L. Manevitz, M. Yousef, One-class document classification via Neural Networks, Neurocomputing. 70 (2007) 1466–1481. https://doi.org/10.1016/j.neucom.2006.05.013.

[36] F. Gao, T. Huang, J. Wang, J. Sun, E. Yang, A. Hussain, Combining deep convolutional neural network and SVM to SAR image target recognition, Proc. - 2017 IEEE Int. Conf. Internet Things, IEEE Green Comput. Commun. IEEE

Cyber, Phys. Soc. Comput. IEEE Smart Data, IThings-GreenCom-CPSCom-SmartData 2017. 2018-Janua (2018) 1082–1085. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.165.

[37] E. Perez, F. Strub, H. De Vries, V. Dumoulin, A. Courville, FiLM: Visual reasoning with a general conditioning layer, 32nd AAAI Conf. Artif. Intell. AAAI 2018. (2018) 3942–3951.

[38] B.J. Jain, P. Geibel, F. Wysotzki, SVM learning with the Schur-Hadamard inner product for graphs, Neurocomputing. 64 (2005) 93–105. https://doi.org/10.1016/j.neucom.2004.11.011.

[39] M.L. Thompson, M.A. Kramer, Modeling chemical processes using prior knowledge and neural networks, AIChE J. 40 (1994) 1328–1340. https://doi.org/10.1002/aic.690400806.

[40] L.F.M. Zorzetto, R. Maciel Filho, M.R. Wolf-Maciel, Process modelling development through artificial neural networks and hybrid models, Comput. Chem. Eng. 24 (2000) 1355–1360. https://doi.org/10.1016/S0098-1354(00)00419-1.

[41] T. Bikmukhametov, J. Jäschke, Combining machine learning and process engineering physics towards enhanced accuracy and explainability of data-driven models, Comput. Chem. Eng. 138 (2020). https://doi.org/10.1016/j.compchemeng.2020.106834.

[42] X. Wang, J. Chen, C. Liu, F. Pan, Hybrid modeling of penicillin fermentation process based on least square support vector machine, Chem. Eng. Res. Des. 88 (2010) 415–420. https://doi.org/10.1016/j.cherd.2009.08.010.

[43] K.C. Tan, Y. Li, Grey-box model identification via evolutionary computing,

Control Eng. Pract. 10 (2002) 673–684. https://doi.org/10.1016/S0967-0661(02)00031-X.

[44] B. Sohlberg, E.W. Jacobsen, Grey Box Modelling – Branches and Experiences, IFAC, 2008. https://doi.org/10.3182/20080706-5-kr-1001.01934.

[45] B.P.M. Duarte, P.M. Saraiva, Hybrid models combining mechanistic models with adaptive regression splines and local stepwise regression, Ind. Eng. Chem. Res. 42 (2003) 99–107. https://doi.org/10.1021/ie0107744.

[46] H.J.A.F. Tulleken, Grey-box modelling and identification using physical knowledge and bayesian techniques, Automatica. 29 (1993) 285–308. https://doi.org/10.1016/0005-1098(93)90124-C.

[47] F. Destro, P. Facco, S. García Muñoz, F. Bezzo, M. Barolo, A hybrid framework for process monitoring: Enhancing data-driven methodologies with state and parameter estimation, J. Process Control. 92 (2020) 333–351. https://doi.org/10.1016/j.jprocont.2020.06.002.

[48] P.F. Van Lith, B.H.L. Betlem, B. Roffel, Combining prior knowledge with data driven modeling of a batch distillation column including start-up, Comput. Chem. Eng. 27 (2003) 1021–1030. https://doi.org/10.1016/S0098-1354(03)00067-X.

[49] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L.H. Chiang, M.S. Reis, Recent trends on hybrid modeling for Industry 4.0, Comput. Chem. Eng. 151 (2021) 107365. https://doi.org/10.1016/j.compchemeng.2021.107365.

[50] A.Y. Di Tsen, S.S. Jang, D.S.H. Wong, B. Joseph, Predictive control of quality in batch polymerization using hybrid ANN models, AIChE J. 42 (1996) 455–465. https://doi.org/10.1002/aic.690420215.

[51] D.C. Psichogios, L.H. Ungar, A hybrid neural network-first principles approach to process modeling, AIChE J. 38 (1992) 1499–1511. https://doi.org/10.1002/aic.690381003.

[52] H.C. Aguiar, R.M. Filho, Neural network and hybrid model: A discussion about different modeling techniques to predict pulping degree with industrial data, Chem. Eng. Sci. 56 (2001) 565–570. https://doi.org/10.1016/S0009-2509(00)00261-X.

[53] Y. Tian, J. Zhang, J. Morris, Modeling and optimal control of a batch polymerization reactor using a hybrid stacked recurrent neural network model, Ind. Eng. Chem. Res. 40 (2001) 4525–4535. https://doi.org/10.1021/ie0010565.

[54] A. Bazaei, V.J. Majd, Feedback linearization of discrete-time nonlinear uncertain plants via first-principles-based serial neuro-gray-box models, J. Process Control. 13 (2003) 819–830. https://doi.org/10.1016/S0959-1524(03)00027-1.

[55] G.M. Bollas, S. Papadokonstadakis, J. Michalopoulos, G. Arampatzis, A.A. Lappas, I.A. Vasalos, A. Lygeros, Using hybrid neural networks in scaling up an FCC model from a pilot plant to an industrial unit, Chem. Eng. Process. Process Intensif. 42 (2003) 697–713. https://doi.org/10.1016/S0255-2701(02)00206-4.

[56] Y. Meng, S. Yu, J. Zhang, J. Qin, Z. Dong, G. Lu, H. Pang, Hybrid modeling based on mechanistic and data-driven approaches for cane sugar crystallization, J. Food Eng. 257 (2019) 44–55. https://doi.org/10.1016/j.jfoodeng.2019.03.026.

[57] Z.K. Nagy, B. Szilagyi, K. Pal, I.B. Tabar, A novel robust digital design of a network of industrial continuous cooling crystallizers of dextrose monohydrate: From laboratory experiments to industrial application, Ind. Eng. Chem. Res. 59 (2020) 22231–22246. https://doi.org/10.1021/acs.iecr.0c04870.

[58] P. Georgieva, M.J. Meireles, S. Feyo de Azevedo, Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena, Chem. Eng. Sci. 58 (2003) 3699–3713. https://doi.org/10.1016/S0009-2509(03)00260-4.

[59] O. Kahrs, W. Marquardt, Incremental identification of hybrid process models, Comput. Chem. Eng. 32 (2008) 694–705. https://doi.org/10.1016/j.compchemeng.2007.02.014.

[60] A. Azarpour, T. N.G. Borhani, S. R. Wan Alwi, Z. A. Manan, M. I. Abdul Mutalib, A generic hybrid model development for process analysis of industrial fixed-bed catalytic reactors, Chem. Eng. Res. Des. 117 (2017) 149–167. https://doi.org/10.1016/j.cherd.2016.10.024.

[61] M. Ebrahimpour, W. Yu, B. Young, Artificial neural network modelling for cream cheese fermentation pH prediction at lab and industrial scales, Food Bioprod. Process. 126 (2021) 81–89. https://doi.org/10.1016/j.fbp.2020.12.006.

[62] H.-T. Su, N. Bhat, P.A. Minderman, T.J. McAvoy, Integrating Neural Networks with First Principles Models for Dynamic Modeling, IFAC Proc. Vol. 25 (1992) 327–332. https://doi.org/10.1016/s1474-6670(17)51013-7.

[63] M.A. Kramer, M.L. Thompson, P. Bhagat, Embedding theoretical models in neural networks, Proc. Am. Control Conf. 1 (1992) 475–479.

https://doi.org/10.23919/acc.1992.4792111.

[64] J. Peres, R. Oliveira, S. Feyo de Azevedo, Knowledge based modular networks for process modelling and control, Comput. Aided Chem. Eng. 8 (2000) 247–252. https://doi.org/10.1016/S1570-7946(00)80043-7.

[65] Q. Xiong, A. Jutan, Grey-box modelling and control of chemical processes, Chem. Eng. Sci. 57 (2002) 1027–1039. https://doi.org/10.1016/S0009-2509(01)00439-0.

[66] R. Oliveira, Combining first principles modelling and artificial neural networks: A general framework, Comput. Chem. Eng. 28 (2004) 755–766. https://doi.org/10.1016/j.compchemeng.2004.02.014.

[67] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.

[68] A. Daw, A. Karpatne, W. Watkins, J. Read, V. Kumar, Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling, (2017). http://arxiv.org/abs/1710.11431.

[69] A. Karpatne, G. Atluri, J.H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, V. Kumar, Theory-guided data science: A new paradigm for scientific discovery from data, IEEE Trans. Knowl. Data Eng. 29 (2017) 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168.

[70] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, Nat. Rev. Phys. 3 (2021) 422–440. https://doi.org/10.1038/s42254-021-00314-5.

[71] A. Güneş Baydin, B.A. Pearlmutter, A. Andreyevich Radul, J. Mark Siskind, Automatic differentiation in machine learning: A survey, J. Mach. Learn. Res. 18 (2018) 1–43.

[72] M.T. Ribeiro, S. Singh, C. Guestrin, Model-Agnostic Interpretability of Machine Learning, (2016). http://arxiv.org/abs/1606.05386.

[73] T. Chen, C. Guestrin, XGBoost, in: Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., ACM, New York, NY, USA, 2016: pp. 785–794. https://doi.org/10.1145/2939672.2939785.

[74] B. Letham, C. Rudin, T.H. McCormick, D. Madigan, Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model, Ann. Appl. Stat. 9 (2015) 1350–1371. https://doi.org/10.1214/15-AOAS848.

[75] F. Wang, C. Rudin, Falling rule lists, J. Mach. Learn. Res. 38 (2015) 1013–1022.

[76] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, N. Elhadad, Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission, Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 2015-Augus (2015) 1721–1730. https://doi.org/10.1145/2783258.2788613.

[77] H. Renon, J.M. Prausnitz, Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures, AICHE J. 14 (1968) 135–144.

[78] G.M. Wilson, Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing, J. Am. Chem. Soc. 86 (1964) 127–130. https://doi.org/10.1021/ja01056a002.

[79] A. Carranza-Abaid, H.F. Svendsen, J.P. Jakobsen, Surrogate modelling of

VLE: Integrating machine learning with thermodynamic constraints, Chem. Eng. Sci. X. 8 (2020) 100080. https://doi.org/10.1016/j.cesx.2020.100080.

[80] E.T. Jaynes, Information Theory and Statistical Mechanics, Phys. Rev. 106 (1957) 620–630. https://doi.org/10.1103/PhysRev.106.620.

[81] N.L. Book, W.F. Ramirez, Structural analysis and solution of systems of algebraic design equations, AIChE J. 30 (1984) 609–622. https://doi.org/10.1002/aic.690300412.

[82] W.F. Ramirez, Computational Methods in Process Simulation, 2nd ed., Elsevier, 1997. https://doi.org/https://doi.org/10.1016/B978-0-7506-3541-7.X5000-8.

[83] D.W. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, J. Soc. Ind. Appl. Math. 11 (1963) 431–441. https://doi.org/https://doi.org/10.1137/0111030.

[84] M.T. Hagan, M.B. Menhaj, Training Feedforward Networks with the Marquardt Algorithm, IEEE Trans. Neural Networks. 5 (1994) 989–993. https://doi.org/10.1006/brcg.1996.0066.

[85] D.J.C. MacKay, Bayesian Interpolation, Neural Comput. 4 (1992) 415–447. https://doi.org/10.1162/neco.1992.4.3.415.

[86] J. Eason, S. Cremaschi, Adaptive sequential sampling for surrogate model generation with artificial neural networks, Comput. Chem. Eng. 68 (2014) 220–232. https://doi.org/10.1016/j.compchemeng.2014.05.021.

[87] A. Nuchitprasittichai, S. Cremaschi, Optimization of CO2 capture process with aqueous amines - A comparison of two simulation-optimization approaches,

Ind. Eng. Chem. Res. 52 (2013) 10236–10243. https://doi.org/10.1021/ie3029366.

[88] J.P. Poort, M. Ramdin, J. van Kranendonk, T.J.H. Vlugt, Solving vapor-liquid flash problems using artificial neural networks, Fluid Phase Equilib. 490 (2019) 39–47. https://doi.org/10.1016/j.fluid.2019.02.023.

[89] A. Carranza-Abaid, R.R. Wanderley, H.K. Knuutila, J.P. Jakobsen, Analysis and selection of optimal solvent-based technologies for biogas upgrading, Fuel. 303 (2021) 121327. https://doi.org/10.1016/j.fuel.2021.121327.

[90] A. Carranza-Abaid, J.P. Jakobsen, A computationally efficient formulation of the governing equations for unit operation design, Comput. Chem. Eng. 154 (2021) 107500. https://doi.org/10.1016/j.compchemeng.2021.107500.

[91] A. Carranza-Abaid, J.P. Jakobsen, A Non-Autonomous Relativistic Frame of Reference for Unit Operation Design, in: Comput. Aided Chem. Eng., 2020: pp. 151–156. https://doi.org/10.1016/B978-0-12-823377-1.50026-4.

Chapter 4.

Constrained VLE modelling

The application of the NNP method to the modelling of non-reactive VLE systems is presented. The traditional Wilson and NRTL activity coefficient models are used as a basis to formulate the ASNNs. These models were utilized in order to show that an ASNN can be thermodynamically consistent if its architecture is appropriately built. Additionally, using well-known models makes the NNP method more accessible to thermodynamic researchers.

This paper is awaiting publication and is not included in NTNU Open

Chapter 5.

Universal approximators as generating functions

This chapter shows an instance on how to exploit the universal approximator feature of ANNs for the formulation of generating function for thermodynamics. This approach allows modelling reactive mixtures in equilibrium as if they were a non-reactive system. The goal of this activity coefficient model is to facilitate the modelling of systems containing acid gases and electrolyte solutions.

This paper is awaiting publication and is not included in NTNU Open

Chapter 6.

A simulation framework for solvent-based biogas upgrading processes

The techno-economic assessment of different processes containing reactive and unreactive solvents is presented in this chapter. The goal of the study is to point out in which operating conditions physical solvents outperform amine solvents. Additionally, it is stressed that the cost of emitting greenhouse gases to the atmosphere should be considered when selecting separation technologies.

Analysis and Selection of Optimal Solvent-Based Technologies for Biogas Upgrading

Andrés Carranza-Abaid*, Ricardo R. Wanderley, Hanna K. Knuutila, Jana Poplsteinova Jakobsen

Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), NO-7491, Trondheim, Norway

Corresponding Author: andres.c.abaid@ntnu.no

Keywords: Biogas upgrading, CO_2 capture, chemical absorption, physical absorption

6.1. Abstract

Biogas upgrading is an important industrial process for producing biomethane, a sustainable energy source with low carbon footprint. There are three main solvent-based alternatives for biogas upgrading: water scrubbing, physical scrubbing and chemical scrubbing with amines. Though assessments have been published regarding which technologies are more cost-effective and energetically efficient, these often either neglect inspecting the impact of raw biogas concentrations and biomethane delivery pressures on the overall performance of the plant, or they do not consider that the separated CO_2 has to be conditioned for transportation for properly fulfilling the requirements of BECCS (bioenergy with carbon capture and storage). In fact, many assessments of physical scrubbing processes forfeit CO_2 recovery

altogether. This work intends to rehabilitate physical scrubbing as an alternative for BECCS by showing that, depending on the conditions of the raw biogas, the requirements for biomethane delivery, and the cost of carbon taxes, biogas upgrading can be feasible by scrubbing either with amines or with organic solvents. We introduce a review on organic physical solvents for CO_2 recovery, a simulation framework for the evaluation of energetical operational costs of biogas upgrading, and a simplified economic analysis. High biomethane delivery pressures and high CO_2 concentrations in raw biogas benefit the use of physical solvents such as N-formyl-morpholine, N-methyl-2-pyrrolidone and poly(ethylene glycol)dimethyl ether, whereas the opposite conditions are advantageous to aqueous monoethanolamine. Finally, the implementation of carbon taxes of around 10 USD/ton CO_2 emitted are sufficient to increase the attractiveness of CO_2 recovery as opposed to CO_2 wasting.

6.2. Introduction

In past decades, the development of sustainable energy generation technologies has become an important asset in mitigating climate change and environmental degradation. These include technologies such as improved solar panels, efficient wind-powered equipment, hydrogen fuel cells and biogas production facilities [1]. Of these, biogas production can be inserted in the wider context of negative emissions [2] and, therefore, deserves a careful assessment.

Biogas is a mixture containing mostly methane (CH_4, 40–75 %v/v) and carbon dioxide (CO_2) [3], with typical secondary impurities being hydrogen sulfide (H_2S), ammonia (NH_3), siloxanes, halogenates and volatile organic carbon (VOC) compounds such as ketones, alkanes and terpenes [3–6]. *Biogas upgrading* is the name given to the process of removing CO_2 from raw biogas. Since biogas is produced biologically

through anaerobic digestion of organic matter, both the nature of the digestate and the conditions of the biological fermentation will affect the composition of raw biogas [6–8]. This is of interest to the industry, since this raw biogas must often be treated before it is delivered as high purity biomethane, and the degree of complexity required for this treatment naturally impacts the sort of technologies needed to perform the task [3,4,6]. There are numerous biogas producing plants that include a biogas upgrading process [9].

Utilization of biogas through combustion inevitably generates CO_2 and cannot be strictly considered a green energy alternative. However, if biogas is produced from a biomass source that participates in a stable carbon cycle (i.e., in which the production of biomass by plants or algae consuming atmospheric CO_2 happens at rate comparable to that of combustion of the resulting biogas), then the net amount of CO_2 emitted at the end of such cycle can be said to approach zero [10]. Surely enough, assuring that a process is “net-zero” is a delicate matter which requires careful evaluation of all of its intermediary steps. One practical way of reducing the odds of having a disbalanced biogas production cycle is by capturing part of the CO_2 that is generated together with biomethane before delivering the latter to its final user. Then, the technology can be said to have achieved *negative emissions*: part of the CO_2 utilized for biomass production in the beginning of the process ends up removed from the carbon cycle through CCS (*carbon capture and storage*) alternatives [2,11].

This is an important facet of the so-called BECCS (*bioenergy with CCS*): from an environmental perspective, the process of biogas upgrading together with CO_2 recovery is not only practical for the consumer (e.g., to generate a stream with higher heating value), but it is also indispensable towards achieving negative emissions in the industry. In this aspect, biogas upgrading *without* CO_2 recovery – i.e., removing

CO_2 from raw biogas and then emitting it to the atmosphere – is a practice that should be ultimately discouraged.

One of the several policies suggested in the Paris agreement in order to achieve the global zero greenhouse gas emissions target by 2050 is the implementation of CO_2 taxes. These taxes force companies or production plants to pay a fee for each ton of CO_2 emitted into the atmosphere. Since biogas upgrading plants *may* or *may not* emit the removed CO_2 to the atmosphere, and since carbon taxes *may* or *may not* apply to bioenergy-based plants, the inclusion of these taxes in economic assessments is an uncertainty expected to hold a crucial role in process feasibility, thus reflecting on the optimal CO_2 capture technology for each scenario.

In order to assess this uncertainty, the present study considers the feasibility of biomethane production plants both in the scenario in which they must pay for their CO_2 emissions and in the scenario where this is not a requirement. This may vary depending on the CO_2 emissions policies of the country where the plant is located (see section 6.4 for further discussion). We suggest that the performance of different biogas upgrading technologies and the selection of the optimal alternative will change as a function of the following external conditions: the raw biogas properties, the treated biomethane specifications and the value of the CO_2 tax. Our intention is, therefore, to evaluate how these external conditions will impact the choice of a proper upgrading technology.

Ultimately, we repeat, biogas upgrading with CO_2 recovery is environmentally beneficial even when not economically so. It is our goal to identify which upgrading technologies are able to better align financial and environmental considerations by minimizing the cost of BECCS.

This work focuses on solvent-based solutions (i.e., absorption by physical and chemical means) due to their prominence in the biogas upgrading market. We present:

- **BACKGROUND (section 6.3).** A summary of biogas upgrading alternatives (section 6.3.1), followed by a review of physical (section 6.3.2) and chemical (section 6.3.3) solvents for CO_2 separation. Though sections 6.3.1 and 6.3.3 are rather cursory, we went into great lengths to make a proper comprehensive review of physical solvents for CO_2 absorption since we have identified a lack of such a source in the available published literature.
- **PROCESS MODELLING (section 6.4).** A methodical explanation of our approach to solvent performance evaluation. Though we employ this approach in the assessment of BECCS in the present study, there is nothing deterring anyone of using such a methodology in the evaluation of other solvent-based processes in different contexts. Hence, we consider that section 6.4 in itself can be valuable for future researchers.
- **RESULTS AND DISCUSSION (section 6.5).** An energy analysis and an operative-cost-oriented assessment of the economic performance of 4 different technologies: **P**hysical/**W**ater absorption without CO_2 recovery (PW), **P**hysical/**W**ater absorption with **CO_2 R**ecovery (PWCR) (both on section 4.1), **A**queous a**M**ine chemical absorption without CO_2 recovery (AM) and **A**queous a**M**ine chemical absorption with **CO_2 R**ecovery (AMCR) (both on section 4.2). This includes an evaluation of different physical solvents, in particular with regards to PWCR (section 4.1). Additionally, we report an easy-to-use guide on optimal process selection

as a function of raw biogas conditions and biomethane delivery pressure. As we present results *with* and *without* the inclusion of carbon taxes, section 0 is insightful even in the context where there are no penalties for the emission of CO_2 by biogas upgrading plants.

Therefore, the middle bulk of this article contains three very distinct albeit interlinked sections, two laying down important groundwork and one delivering our main results. These sections can be read by themselves, hence we recommend that readers who are interested merely in the outcome of our energetic analysis skip directly to section 0 and then to section 6.6 (the Conclusions). However, we hope to have showed with the table of contents above that each section has its individual value, and that they all contribute to a proper understanding of our results.

6.3. Background

6.3.1. Biogas upgrading technologies

There are currently six main biogas upgrading technologies. These are:

- Water scrubbing – using pressurized water to physically absorb CO_2 followed by decompression and/or stripping with an inert gas for regeneration [3,6,8,12] (formerly, the resulting water was just directly wasted as an effluent, but this is generally not acceptable anymore [6,8]).
- Physical scrubbing – similar to water scrubbing, but using an organic solvent instead of water [6,8,13,14]. Solvent regeneration can be of three types: flash desorption (by solvent decompression), stripping with an inert gas, and hot regeneration [15]. More on these organic solvents will be discussed in section 0.

- Chemical scrubbing – a solvent containing a species that chemically reacts with CO_2 is employed for upgrading [4,8,13,14]. This chemical reaction increases the solvent capacity for CO_2 absorption, but also makes regeneration more difficult. Desorption is performed by supplying heat to the system [13,14,16]. More on these chemical solvents will be discussed in section 6.3.3.
- Pressure swing adsorption (PSA) – CO_2 is adsorbed over a porous material at high pressures and desorbed at low pressures in the so-called *Skarstrom cycle* [13,17]. The criterium for separation here is the higher CO_2 diffusivity when compared to that of methane (due to its lower molecular size) [6,8,12,14]. Typical materials for CO_2 adsorption are zeolites, silicates, silica gel and activated carbon [8,12], though current research in the field of metal-organic frameworks render these materials particularly attractive.
- Membrane technology – CO_2 is separated from methane due to its lower molecular size in a porous interface using differential partial pressure as the driving force [3,12–14]. Development of such membranes is rapidly advancing [18,19]. For industrial applications, the materials employed are typically polymeric in nature [14], though fast developments might render this information anachronistic.
- Cryogenic separation – CO_2 is condensed at high pressures and low temperatures while methane remains in the gas phase due to the difference between the boiling points [3,6,8]. Due to the high energy-intensity of this process, designing an optimal system configuration is essential, and therein lie most recent advances in this field [20].

Comparisons between these six different techniques can be performed across a variety of criteria, with none being necessarily more or less important than the other. In terms of modularity and flexibility, for example, membranes are commonly argued to be superior to other technologies [21,22] – and yet, they often require an integration of parallel and/or sequential modules to be fully capable of separating CO_2 [18,19], bringing down their energy efficiency [12,14]. Energy consumption is often also high when operating PSA, and its footprint (i.e., the amount of physical space a processing plant occupies) can be problematic [17]. Cryogenic methodologies are expensive both in terms of equipment as well as operational costs, but are able to deliver highly purified biomethane ready for condensation into LBG (*liquid biogas*) together with pressurized CO_2 proper for transportation and storage [8,20]. In terms of robustness to stand impurities, water scrubbing is able to handle most of them easily [12], whereas cryogenic separation can be designed so that each contaminant is removed sequentially according to their relative volatilities [20]. All of the other techniques are vulnerable to one impurity or another, which can devolve in operational issues such as membrane degradation [19] or solvent decomposition [23].

Perhaps the best way to assess the competitiveness of these different technologies is by looking at how they are actually implemented in the biogas upgrading market. Almost every year, IEA Bioenergy publishes a list of the new developed biogas upgrading plants implemented by their member countries (which critically does not include neither China nor the U.S.A.) categorized by chosen upgrading technology. Figure 6.1 shows the cumulative number of biogas upgrading plants deployed from 2000 up to 2019 in those member countries. We unfortunately had to ignore the entries in the list for which no installation year was given, but those were very few overall.

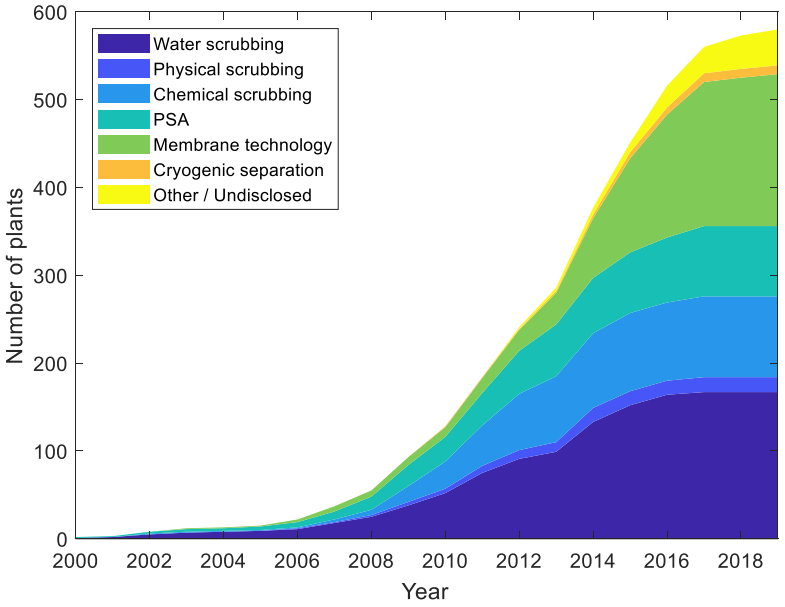


Figure 6.1. Cumulative number of biogas upgrading plants installed by IEA Bioenergy member countries divided by their chosen biogas upgrading technologies. List obtained from the IEA Bioenergy website [9].

Figure 6.1 shows a fragmented market between several technologies, with physical scrubbing and cryogenic separation techniques arguably lagging the others. From the perspective of someone coming from either the pre-combustion or post-combustion carbon capture background, this is a somewhat surprising image. In those fields, one typically thinks of chemical absorption as the state-of-the-art CO_2 capture process [24,25]. In the biogas upgrading field, however, chemical scrubbing with amines not only fares poorly compared to membrane technologies, but it is also overshadowed by water scrubbing. Conversely, one finds no instances of water scrubbing ever being used for natural gas treatment nor CO_2 removal from flue gases.

A 2012 study carried by the Vienna University of Technology has come up with a comparison between five of the six technologies mentioned previously, which is summarized on Table 6.1. Being from 2012, Table 6.1 is perhaps anachronistic, but its overall message is representative of what has been reported until quite recently regarding biogas upgrading (see [26] for more contemporaneous data). The electrical energy consumption of all different technologies is very similar, with the remarkable exception of chemical scrubbing. This, of course, comes with two caveats. First, chemical scrubbing requires a high thermal energy consumption for its solvent regeneration step, so that looking at electricity consumption alone might be misleading. Second, chemical scrubbing is the technology in which biomethane is delivered at its lowest pressure. As biomethane is most often pressurized for injection into either a medium-pressure (e.g., 16 bars) or high-pressure (e.g., 55 bars) natural gas grid [1], the gas produced by chemical scrubbing is still in need of compression unless it is directly used for energy production. Even with this limitation, however, one can see at the bottom of Table 6.1 that the operational costs of chemical scrubbing end up above those of the various other technologies.

Finally, there is something to be said about biomethane purity and methane slip. The demanded biomethane purity is defined by the constraints imposed for its utilization. For biomethane injection into the natural gas grid, the 2016 Deliverable D5.2 from the ISAAC project to the European Commission [27] compiles some of the current national standards: biomethane should contain less than 2.5 %v/v CO_2 in Germany and France, less than 3 %v/v CO_2 in Sweden, Denmark, Austria and the U.K., and less than 6 %v/v in the Netherlands. Though national standards may vary, basic standards for the European Union are given in EN 16723-1 for biomethane injection in the gas grid and in EN 16723-1 for biomethane use in road transport [28] – notice

that EU standards do not specify a maximum CO_2 content but a minimum methane number: 65 according to EN 16723-1, 65 or 80 according to EN 16723-2 (depending on fuel grade) [29]. Clearly then, all technologies addressed on Table 6.1 are capable of producing high quality biomethane. The main difference among processes is methane slip. Other than the obvious loss of profit from reduced productivity, methane slip is an environmental hazard that is typically addressed through regenerative catalytic oxidation (RCO) or regenerative thermal oxidation (RTO) (see the following references [30,31] for a comparison between RCO and RTO), i.e., burning, which involves additional energetic and economic penalties. Chemical scrubbing has a high selectivity for CO_2 separation and consequentially the lowest methane slip values amongst all technologies. The reason for this is the very low solubility of CH_4 into the aqueous amine mixture and the low pressures commonly used in the amine scrubbing processes.

One could then wonder what the most environmentally benign technology for biogas upgrading is. As it turns out, it depends heavily on the origin of the electricity employed in the biogas upgrading plant [33,34]. This highlights the importance of choosing an energy efficient pathway for biogas upgrading. Therefore, both for deployment considerations and environmental considerations, energy saving is key.

Table 6.1. Comparison between biogas upgrading technologies carried by the Vienna University of Technology [32] in 2012. Values typical of plants with a capacity to produce around 500 m^3/h biomethane (average sized biogas upgrading plants).

	Water scrubbing	Physical scrubbing	Chemical scrubbing	PSA	Membrane technology
Biomethane purity (%)	95.0 – 99.0	95.0 – 99.0	> 99.0	95.0 – 99.0	95.0 – 99.0
Methane slip (%)	2.0	4.0	0.04	2.0	20.0 – 0.5
Delivery pressure (bar)	5 – 9	5 – 9	1	5 – 8	5 – 8
Electric energy demand (kWh/m^3 biomethane)	0.46	0.49 – 0.67	0.27	0.46	0.25 – 0.43
Heating demand	None	Medium	High	None	None
Typical investment costs (€/m^3/h biomethane)	3500	3500	3500	3700	3500 – 3700
Typical operational costs (€/m^3/h biomethane)	0.091	0.090	0.112	0.092	0.065 – 0.101

6.3.2. Physical solvents for CO_2 absorption

In physical scrubbing, the CO_2 binds to the physical solvents by relatively loose intermolecular forces. It has been demonstrated that CO_2 physical solubility relies on Lewis acid-base interactions between the acid gas and the absorbent [35–37], being thus more pronounced in solvents with strong electronegative groups. Because of this, the calorific energy requirement to regenerate the solvents in the physical absorption processes is lower than the one from chemical absorption processes.

Table 6.2. Physical solvents for CO_2 absorption and their processes.

Abbreviation	Name of chemical	Process it appears in
Methanol	Methanol	Rectisol™ / Ifpexol™
NFM	N-formyl-morpholine	Morphysorb™
NMP	N-methyl-2-pyrrolidone	Purisol™
PC	Propylene carbonate	Fluor™
PEGDME	Poly(ethylene glycol) dimethyl ether	Selexol™ / Genosorb™
TBP	Tributyl phosphate	Estasolvan™
TMS	Tetramethylene sulfone	Sulfinol™

There have been many physical scrubbing processes developed commercially for acid gas absorption [38], and new candidates for physical solvents are developed each year [39,40]. Many ionic liquids operate essentially as physical solvents [41], and so do the silicone-based hydrophobic physical solvents [42,43] recently presented in literature. In this study, however, we will focus only on commercial

solvents. The physical solvents included in our analysis are listed in Table 6.2. A good review on many of the physical absorption technologies is given by Vega et al. [44], therefore we will only summarize some of the operational peculiarities of these solvents.

- Methanol: Both the Rectisol™ and the Ifpexol™ processes operate with methanol at moderate to high pressures (p_{CO2} ≥ 1 MPa [45]) and fairly low temperatures, −70 to −10 °C though the details vary from author to author [38,44–47]. These low temperatures both enhance CO_2 solubility and help avoiding methanol losses due to volatilization.
- N-formyl morpholine: The Morphysorb™ process operates with mixtures of N-formyl-morpholine and N-acetyl-morpholine. Absorption is also carried at moderate pressures, such as P_{CO_2} ≈ 0.8 MPa in the Kwoen power plant [48] and temperatures between −20 to 40 °C [44]. These low temperatures both enhance CO_2 solubility and help avoiding methanol losses due to volatilization. It is important to remark that NFM has a relatively high freezing point of approximately 23 °C, and thus the addition of N-acetyl-morpholine to the Morphysorb™ solvent comes as a solution for enabling operation at lower temperatures [49].
- N-methyl-2-pyrrolidone: The Purisol™ process uses chilled N-methyl-2-pyrrolidone at temperatures as low as −15 °C [44,47,49], with solvent volatility becoming possibly an issue in case refrigeration is not employed [47,49–51]. Once again, higher CO_2 partial pressures are preferred, for example P_{CO_2} = 2.4 MPa [15].
- Propylene carbonate: The Fluor™ process operates with propylene carbonate at high pressures (between 3 and 8 MPa total pressure) and

ambient temperatures [44], though chilling can also be employed [47]. There seems to be evidence that the propylene carbonate selectivity for methane instead of CO_2 increases with pressure [38], indicating some sort of competitive absorption. This, coupled with the low tolerance of propylene carbonate to H_2S, makes the Fluor™ process more popular for syngas treating and not for natural gas applications [38,47].

- Poly(ethylene glycol) dimethyl ether: The Selexol™ and Genosorb™ processes are popular alternatives for CO_2 separation [45,52]. The solvent is a mixture of polyethylene glycol dimethyl ethers with chain lengths of between 3–9 monomers [45,47,52,53] (the details of this mixture potentially make the difference between the Selexol™ and Genosorb™ solvents [53]). The Selexol™ operates between 0 and 5 °C [44,47], with lower temperatures being avoided due to large solvent viscosity issues [47]. Processes are typically operated in a pressure range from 2 to 14 MPa and treat gases of 5 to 60 %v/v of CO_2 content [50,54].
- Tributyl phosphate: Not much is spoken about the Estasolvan™ nowadays, perhaps because of the low solubility of CO_2 in tributyl phosphate [38,50,51,55]. To our knowledge, this process has never been implemented commercially [38].
- Tetramethyl sulfone: Contrarily to the other solvents mentioned in this section, there is no physical scrubbing process employing tetramethyl sulfone (commonly called *sulfolane*) as a solvent in itself. Instead, sulfolane is mixed with an amine such as diisopropanolamine or N-methyldiethanolamine to form the Sulfinol-M™ solvent. The Sulfinol™ is

often called a hybrid process, and has arguably more similarities to chemical scrubbing than to physical scrubbing [38,56–58].

In terms of solvent stability, most organic solvents are apparently resistant to degradation. Evaluation of the Morphysorb™ process in the Kwoen plant shows that N-formyl-morpholine suffers little degradation, being safe to operate at temperatures below 80 °C [48]. The Selexol™ solvent is stable, nontoxic and biodegradable [38,45]. Propylene carbonate is apparently less stable, so that operations should be kept below 65 °C [47] and mixing with water should be avoided [49].

The physical solubility of CO_2 and methane in these organic solvents can be estimated with Henry's law, Eq. (6.1), where P_i is the partial pressure of the gas and x_i is the equilibrium molar fraction of the gas in the solvent. Eq. (6.2) shows a temperature-dependent expression for calculating the Henry's coefficient H_i, wherein H_i has the unit of MPa and T has the unit of K. The Henry's relation and the correlation used to estimate the solubility of the gas i is given by:

$$H_i = \frac{P_i}{x_i} \tag{6.1}$$

$$\ln(H_i) = A + \frac{B}{T} \tag{6.2}$$

Table 6.3 shows the A and B parameters for different gases in a series of solvents. These parameters have been regressed from different sources in literature, all of which are referred to in Table 6.3. For regressing the parameters of Eq. (6.2), we have refrained from using any set of data above 100 °C or in which the molar fraction of gas in the solvent was above x_i = 0.30. These measures were taken to reinforce that the resulting parameters are fitted precisely for our region of applicability (between 20

°C and 80 °C, with lower temperatures if possible), and also to make sure that the solubility data does not extend beyond the span of validity of Henry's law approach. It is also true that high temperature solubility data tends to be less reproducible than data at lower temperatures, as can be observed in Figure 6.2, this being another reason for avoiding using high temperature data for parameter fitting.

Table 6.3. Henry's law parameters for gas absorption in physical solvents.

Gas	Solvent	A	B	T range	Sources
CO_2	Methanol	8.328	−1709.8	[−60, 90 °C]	[59], [60], [61]
	NFM	7.710	−1716.2	[25, 100 °C]	[62], [63]
	NMP	7.567	−1682.8	[−20, 100 °C	[64], [65], [66], [67]
	PC	7.858	−1730.5	[25, 100 °C]	[63], [64]
	PEGDME	6.032	−1411.3	[25, 70 °C]	[63], [53]
	TBP	4.493	−1024.5	[0, 40 °C]	[68],[69]
	TMS	7.123	−1471.5	[30, 100 °C]	[64], [70]
CH_4	Methanol	5.086	123.26	[−60, 40 °C]	[71], [72], [73]
	NFM	4.786	101.00	[25, 100 °C]	[62], [74]
	NMP	4.926	−166.02	[−20, 25 °C]	[67], [74]
	PC	4.808	76.269	[−29, 100 °C]	[75]
	PEGDME	3.137	111.68	[25, 60 °C]	[76]
	TMS	5.086	123.26	[25, 100 °C]	[70]

We have not included in our fitting sets of data that differed too much from their counterparts, or that expressed an abnormally high CO_2 solubility. This applies to Rajasingam et al. [77] for NMP and to Chen et al. [78] for TPB. It is difficult to find (and thus evaluate) experimental data for CO_2 solubilities in tributyl phosphate, but the results obtained by Li et al. [68] and by Thompson et al. [69], and molecular simulation estimates give credibility to the observations of the latter two research groups [79].

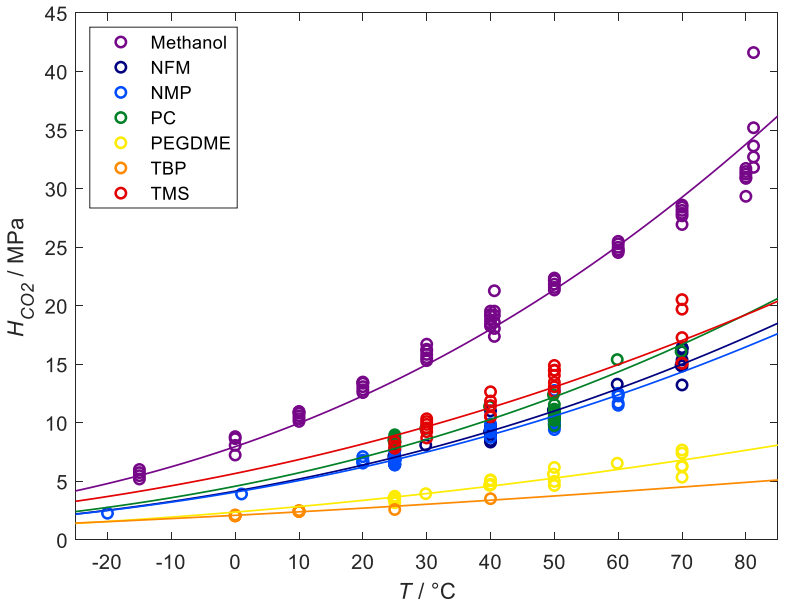


Figure 6.2. Experimental Henry's coefficients at different temperatures obtained in literature for a series of organic solvents. The bold lines have been obtained with Eq. (6.1 and the regressed parameters shown on Table 6.3.

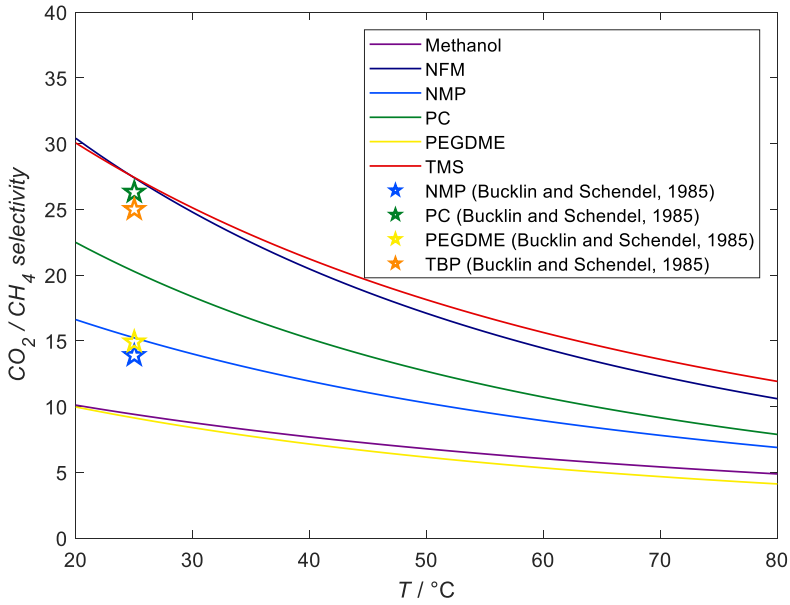


Figure 6.3. Selectivity for CO_2 absorption with respect to methane in a series of physical solvents at different temperatures. The lines have been obtained with Eq. (6.1 and the regressed parameters shown on Table 6.3. The stars are published data reported by Bucklin and Schendel [50] for 25 °C only.

The experimental Henry's coefficients of all physical solvents are shown in Figure 6.2, together with the curves obtained with Eq. (2) and the parameters presented on Table 6.3. In molar basis, methanol has the lowest CO_2 solubility and TBP has the highest, and the remainder solvents follow roughly an order of TMS < PC < NFM < NMP < PEGDME.

In terms of selectivity for CO_2 in detriment of methane, the solvents perform as shown on Figure 6.3. Unfortunately, we were unable to find published data regarding methane solubility in TBP. Nevertheless, Bucklin and Schendel [50] report the selectivity of some physical solvents for CO_2, and among them TBP. As shown on

Figure 6.3, if the Henry's coefficient data from Table 6.3 for CO_2 and methane is taken into account, the selectivity for CO_2 follows the order PEGDME < methanol < NMP < PC < TBP < NFM < TMS, though there are discrepancies between the data reported by Bucklin and Schendel [50] and the curves obtained through modelling (also notice that, while Kohl and Nielsen [38] mention the low selectivity of propylene carbonate, this is not observed by Bucklin and Schendel [50]). It is important to remark that the CO_2 selectivity for all solvents increase with a decrease in absorption temperature.

Physical solvents do not interact strongly with CO_2 during the absorption process, and evidence of this is the low exothermicity of the chemical phenomenon. The CO_2 heat of absorption may be calculated by application of the van 't Hoff equation (Eq. (6.3) to Eqs. (6.1) and (6.2), resulting in Eqs. (6.4) and (6.5) respectively:

$$\left(\frac{\partial\, ln(P_i)}{\partial T}\right)_{x_i} = -\frac{\Delta H}{R \cdot T^2} \tag{6.3}$$

$$\left(\frac{\partial ln P_i}{\partial T}\right)_{x_i} = -\frac{B}{T^2} = -\frac{\Delta H}{R \cdot T^2} \tag{6.4}$$

$$\Delta H = B \cdot R. \tag{6.5}$$

The van 't Hoff equation can be used to calculate the heat of absorption ΔH by using the differential of the Napierian logarithm of P_{CO_2} with respect to temperature. Remarkably, one convenience of employing Eq. (6.2) for the modelling of CO_2 solubilities instead of a more complex expression is that each B-parameter fitted in Table 6.3 translates linearly into the CO_2 heat of absorption in its respective solvent, as seen on Eq. (6.5). Note that Eqs. (6.3) - (6.5) apply for both CO_2 and CH_4.

The heat of absorption values calculated with Eqs. (6.4) - (6.5) and shown on Table 6.4 are very similar to the ones obtained empirically through calorimetric experiments for physical solvents [80,81]. These calculations were performed with the parameters shown on Table 6.3 at T = 20 °C and P_{CO_2} = 100 kPa. Tributyl phosphate has the lowest heat of absorption of all solvents. A low heat of absorption is also observed for PEGDME and TMS. It is an interesting fact that the solvents which show less CO_2 solubility also show the lowest exothermicities for CO_2 absorption in Table 6.4. As for methanol, NFM, NMP and PC, their heat of absorption is essentially the same.

The solubility of CO_2 in water has been obtained by several authors, and the works by Dodds et al. [82] and Diamond and Akinfiev [83] offer good summaries of the published data. For our parametrization of Eq. (6.5), we have employed four representative datasets covering the temperature span between 0 and 100 °C with special emphasis on low temperature data. Our list of references, as well as the parameters obtained through the regression, can be seen on Table 6.5. The heat of absorption of CO_2 and CH_4 into water can be calculated with Eq. (6.5), from which their values are calculated as −17.55 kJ/mol and −11.51 kJ/mol respectively.

Table 6.4. Heat of absorption for a series of physical solvents. Values obtained through Eq. (6.7) with the parameters presented on Table 6.3.

Solvent	ΔH / kJ·mol CO_2^{-1}
Methanol	−14.2
NFM	−14.3
NMP	−14.0
PC	−14.4
PEGDME	−11.7
TBP	−8.5
TMS	−12.2

Table 6.5. Henry's law parameters for gas absorption in water.

Gas	A	B	T range / °C	Sources
CO_2	12.2616	−2110.7	[1, 100]	[84], [85], [86], [87]
CH_4	13.0312	−1384.4	[1, 71]	[88], [89], [90], [91]

An important solvent characteristic that should be considered in addition to its absorption capacity and selectivity is the solvent volatility. This solvent property has an important effect on the selection of the operating temperatures. With the sole exception of methanol, all organic solvents evaluated in this study have lower volatility than water (see Figure 6.4). Note that some liquid vapor pressures were extrapolated

for the sake of completion. In fact, these solvents often have vapor pressures below those of aqueous amines like MEA. Moreover, all physical absorption processes proposed in this investigation are carried out at lower temperatures, and the little amount of solvent that is carried over with the CO_2 product (i.e., due to desorption at higher temperatures) is recovered after pressurization of the CO_2. Hence, the loss of physical absorbent via volatilization is deemed to be negligible in most applications covered in this research. This will be further discussed in section 4.1.1.

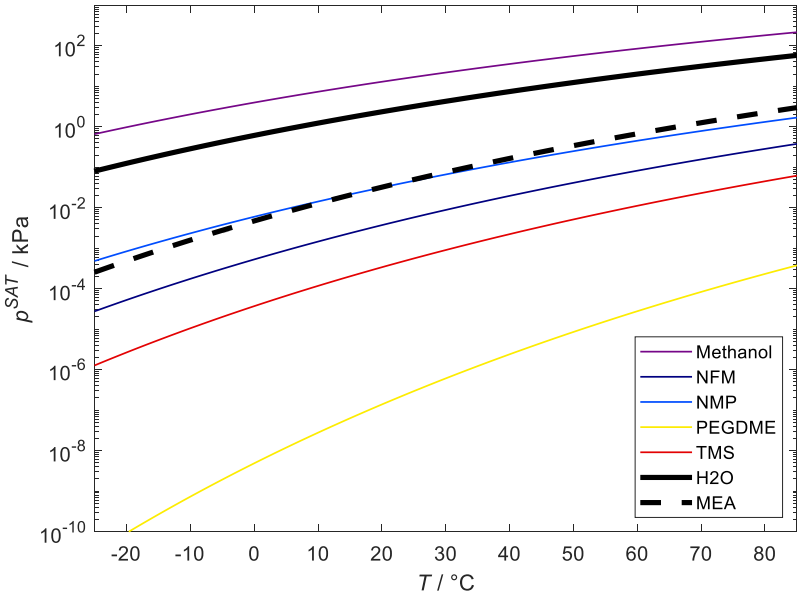


Figure 6.4. Saturation pressure (kPa) of different components vs temperature (°C).

Finally, it should be pointed out that our analyses do not consider the impurities in the biogas to be a problem with regards to solvent degradation in the case of physical solvents. As previously mentioned, the main impurities in biogas besides CO_2 are H_2S, NH_3, and then trace amounts of siloxanes, halogenates and VOC compounds

[3–6]. Most reports do not mention issues regarding organic physical solvent degradation due to H_2S, with the exception of propylene carbonate [38,47] (in fact, the tolerance of solvents such as PEGDME with regards to H_2S is often mentioned as one of their strengths [47,50]). Siloxanes are harmless to physical absorbents – their main cause of concern is the risk of microcrystalline silica formation, which might be abrasive to equipment downstream of the upgrading plant [92]. Similarly, NH_3 and other contaminants have little to no effect on solvent stability. In summary, degradation in the case of organic physical solvents should not be above what is observed for most chemical solvents. All other issues posed by these secondary contaminants c<an be easily addressed with a dedicated cleaning station, which is a common feature of most biogas production factories [5].

6.3.3. Chemical solvents for CO_2 absorption

The field of CO_2 absorption with amines is vast [25,93] and this section does not intend to carry a full review. Instead, we will focus solely in discussing the nature of the chemical solvent itself.

Solvents for chemical scrubbing of CO_2 are targeted to address a series of practical issues regarding the CO_2 capture process. These issues encompass:

- Capacity and cyclic capacity – Meaning how much CO_2 a solvent in thermodynamical equilibrium can pack between its absorption and desorption cycles. High cyclic capacities mean that less solvent is needed to perform the same amount of CO_2 removal, which implies smaller equipment and perhaps less heating, cooling, and pumping duties.

- Rates – Meaning how fast the CO_2 is absorbed into the solvent. Without fast absorption rates, whatever high cyclic capacities are in equilibrium cannot be achieved in practice.
- Volatility – If one employs volatile solvents, an extra care must be taken to avoid solvent emissive losses. This can result in costly equipment and complex treatment processes [94].
- Viscosity – High viscosities mean all transport phenomena are carried out with more morosity, which affects equipment sizing and energy performances [95].
- Degradation – Thermal and oxidative degradation result in constant reclamation and costly solvent make-up issues. Degradation has an additional impact on corrosion and emissions, being an environmental as well as a financial issue [96].
- Corrosion – Corrosion has been observed to happen in more than half of the typical CO_2 capture plant equipment [97], thus drastically reducing the useful life of the installation.
- Emissions – Emissions might stem from the volatilization of the amine or of its degradation products, presenting a threat to health and environment [98,99]. It must be noted that, in the case of biomethane production, wherein the treated gas will ultimately be combusted, the issue of emissions can be considered perhaps less of a problem than in cases where the treated gas is released to the atmosphere.
- Toxicity – Solvent toxicity to plant and animal life is clearly an issue of health and environmental concern [100].
- Price – Amines for CO_2 capture should not be too expensive [93].

Previously, the benchmark chemical solvent for CO_2 absorption was aqueous monoethanolamine (MEA). This in itself came out of a technological evolutionary process. In fact, the first amine employed in the chemical scrubbing industry was triethanolamine (TEA) [94]. This amine presented a series of issues (e.g., high viscosity, low absorption rates) and quickly lost ground to MEA, diethanolamine (DEA), N-methyldiethanolamine (MDEA) and diglycolamine (DGA). Of these, MEA presented the largest number of advantages: it has fast mass transfer rates, low viscosity, high capacity, and whatever issues it has that contribute to its losses (e.g., its degradation rates [101], its relatively high volatility) are swiftly compensated by its very low price [93].

The development of amine solvents for CO_2 capture did not end with aqueous MEA. Investigations veered into new aqueous amine solvents containing one single amine [102] and then amine blends [103,104] with great success. Investigations attempted to remove the water off of these mixtures to create water-lean solvents [105], or to shift from alkanolamines to naturally occurring amino acids [106,107]. From single-phase solvents, the solvent has been allowed to form two phases either with liquid-liquid demixing [108] or solid precipitation [109]. This is an extremely long and intricate history of developments that go beyond the scope of this section. What can be said is that aqueous MEA is in a process of losing its relevance as the benchmark amine solvent for CO_2 capture, being substituted either by aqueous piperazine (PZ) [110] or by blends of PZ and 2-amino-2-methyl-1-propanol (AMP) [111]. Meanwhile, it seems that industrial biogas upgrading plants have been using aqueous mixtures of MDEA and PZ more often than aqueous MEA [112–114]. This is supported by Bauer et al. [8,12]. This blend of MDEA with PZ is often called aMDEA (*activated MDEA*) [8,115].

Regardless, most reviews on chemical scrubbing for biogas upgrading still mention aqueous MEA as the benchmark [113,114,116]. For our purposes, we shall join our peers in focusing on aqueous MEA as a representative chemical solvent. This is obviously convenient, as aqueous MEA has been extensively studied and its properties are widely catalogued. Moreover, since we have decided to carry our analysis on the basis of thermodynamic performance while ignoring issues of kinetics and degradation/corrosion (in which one would find vast discrepancies between aMDEA and MEA performances), for us, the main differences between aMDEA and MEA are their distinct vapor-liquid equilibrium behaviors (aMDEA has a slightly higher cyclic capacity [117]) and enthalpies of absorption. With this in mind, results for MEA should not differ excessively from those for aMDEA, as the true advantages and disadvantages of the aMDEA solvent over aqueous MEA are obfuscated from a purely thermodynamic perspective.

In a practical sense, there are two essential thermodynamic quantities that must be well understood for assessing the energy consumption of CO_2 absorption-desorption into a process with aqueous amines. The first one is the solvent capacity and how it shifts with temperature. This is often referred to as the vapor-liquid equilibrium (VLE) of CO_2 in the solvent. The second thermodynamic quantity is the CO_2 heat of absorption, i.e., how much heat is released in the exothermic absorption of CO_2 into the solvent. Even though roughly the same amount of heat must be given to the solvent for its endothermal desorption of CO_2, the energy penalty is still considerably high. Together, these two thermodynamic quantities define both how much CO_2 is released by raising the temperature of the solvent and how much energy has to be spent for releasing CO_2 at high temperatures.

As discussed in the literature [118–120], the use of the van 't Hoff equation in amine systems inherently implies that, among other assumptions, only one reaction is occurring in the system, some species are disregarded and the relation between the molar fractions and the activity coefficients is constant. Even though for physical solvents this approach is reasonable, for amine systems considerable errors have been reported, even if the thermodynamic model is consistent and accurate with respect to the VLE [121,122]. Hence, the empirical method proposed by Kim and Svendsen [118] using differential calorimetry seems to produce more precise results, as it is directly targeted at measuring enthalpy variations in the solvent upon absorption of CO_2.

Some simplified thermodynamic models have been developed and report a constant heat of absorption for the CO_2-MEA-H_2O system [123] of 88.0 kJ/mol. However, for typical reboiler operating temperatures (120 °C), using this averaged value underpredicts the energetic requirements of the reboiler, hence the experimental values reported by Kim and Svendsen are used in this work [118]. By comparing this value to the ones reported by physical solvents (Table 6.4), it is possible to acknowledge that the chemical solvents will require more calorific energy to carry out the separation, as the CO_2 heat of absorption in those is around 4 to 5 times larger than in physical solvent processes.

A review on the chemical scrubbing process would be superfluous in this stage of our study, as a detailed description of how this process is performed is carried out throughout section 6.4. For a different approach on process modelling, however, we invite the reader to consider the works of Moioli et al. [107,124] and Øi et al. [125] employing Aspen Plus and Aspen HYSYS respectively.

6.4. Process modelling

This section describes the in-house models utilized in the techno-economic assessment of this work. The models were implemented in Matlab 2019b and the thermodynamic properties not referenced in section 2 were taken from the Aspen Plus v8.6 databank. These parameters are provided in the supplementary information.

Our thermodynamic framework assumes that the vapor phase behaves as an ideal gas, the liquid phase behaves as an ideal mixture and the condensers, reboilers and flash tanks are in thermodynamic equilibrium. Although real gases deviate from the ideal behavior at moderate to high pressures, using the same assumption when comparing the different technologies should not jeopardize the findings of this work.

The main goal of any biogas upgrading plant is to process the feed of a raw biogas stream (F_F) at temperature (T_F), at pressure (P_F) and with a methane composition $\left(z_{CH_4}\right)$ in order to produce a biomethane stream (F_B) with a delivery temperature (T_B), pressure (P_B) and methane composition $\left(y_{CH_4}\right)$. Furthermore, the processes also deliver a carbon dioxide stream (F_C) at a certain temperature (T_C), pressure (P_C) that mainly contains CO_2 $\left(x_{CO_2}\right)$ and usually small quantities of CH_4 $\left(x_{CH_4}\right)$.

6.4.1. Physical and water absorption process

6.4.1.1. With CO_2 recovery

The flowsheet of the **P**hysical/**W**ater absorption process with **C**O_2 **R**ecovery (PWCR) is shown in Figure 6.5. The role of the absorber in the process is to remove the necessary amount of CO_2 from the absorber vapor inlet (F_G) in order to comply with the biomethane composition specification $\left(y_{CH_4}\right)$. The CO_2 removal requires the absorber to be operated at a certain pressure (P_A) with a lean liquid solvent stream

(F_L) at a given temperature (T_L). The PWCR absorber uses the material balances, equilibrium relations, summation equations and enthalpy balances (MESH equations). Thermodynamic equilibrium is assumed at the bottom of the column (the compressor outlet temperature is adjusted to match the outlet liquid temperature calculated with the energy balance in order to simplify the solution algorithm). The lean liquid solvent stream flow is specified using the liquid to gas ratio (L/G):

$$L/G = \frac{m_L}{m_G} \tag{6.6}$$

where L/G can be in mass or mole basis. Since both CH_4 and CO_2 are solubilized in the solvent, and the carbon dioxide stream has a specified maximum limit for the amount of methane, part of the methane absorbed must be recovered. This is done by depressurizing the rich solvent stream and then recirculated the vapor that is vaporized from the recycle adiabatic flash tanks. Therefore, only the output pressure of each flash tank can be tuned in order to meet the x_{CH_4} specification. The pressure of the flash tanks must be between the absorber pressure P_A and the reboiler pressure P_R, thus the following relations are proposed:

$$P_1 = (P_A - P_R) \cdot r_P + P_R \tag{6.7}$$

$$P_2 = (P_1 - P_R) \cdot r_P + P_R \tag{6.8}$$

where P_1 and P_2 are the outlet pressures of the flash tank 1 and 2 respectively, while r_P is the depressurization ratio (must be between 0 and 1) which represents the fraction of the pressure difference between the inlet pressure and the reboiler. This arrangement is convenient for numerical stability as it ensures that $P_A > P_1 > P_2 > P_R$

while eliminating one degree of freedom. Considering this, r_P can be directly iterated upon in order to comply with the specification in x_{CH_4}.

The heat exchanger shown in Figure 6.5 is the result of the heat integration of the biogas upgrading plant and uses a fixed temperature difference in the cold side equal to 10 °C. The computational framework has an implemented algorithm that does not perform the heat integration in cases where heat recovery is not thermodynamically feasible.

The reboiler operation finishes the depressurization cycle by decreasing the pressure from P_2 to P_R and heating the solvent up to a specified reboiler temperature T_R. A temperature increase is usually needed to increase the partial pressure of CH_4 and CO_2 so that the lean solvent has the least amount of solubilized gases. The make-up stream reintroduces the solvent lost due to the evaporation in the reboiler.

Since the carbon dioxide stream is assumed to be fed to a CO_2 transport and storage system, the PWCR requires to comply with a maximum permissible amount of CH_4 in the carbon dioxide stream. Considering this, using a reboiler as the last separation stage instead of a third adiabatic flash tank provides an extra degree of freedom (T_R) that enables the regulation of the carbon dioxide stream compositions. Tuning the compositions without thermal regeneration constraints the operating ranges and, in most cases, it is not possible to find an optimal solution that complies with the specifications of the biomethane, and the carbon dioxide streams.

There are four different compressor sections in the PWCR process, each of which has different delivery pressures. The delivery pressure of the feed compressor is the same as the pressure calculated from mixing the vaporized streams from the flash tanks. The absorber compressor section outlet pressure value is identical to P_A. The biogas

compressor elevates the pressure from P_A and delivers it at P_B. The CO_2 compressor increases the pressure from the reboiler up to P_C.One of the main concerns in biogas upgrading technologies is the loss of methane in the process. This is quantified with the methane slip parameter (θ):

$$\theta = 1 - \frac{F_{B,CH_4}}{F_{F,CH_4}} \tag{6.9}$$

where F_{B,CH_4} is the CH_4 material flow in the biomethane stream and F_{F,CH_4} is the CH_4 material flow in the raw biogas stream.

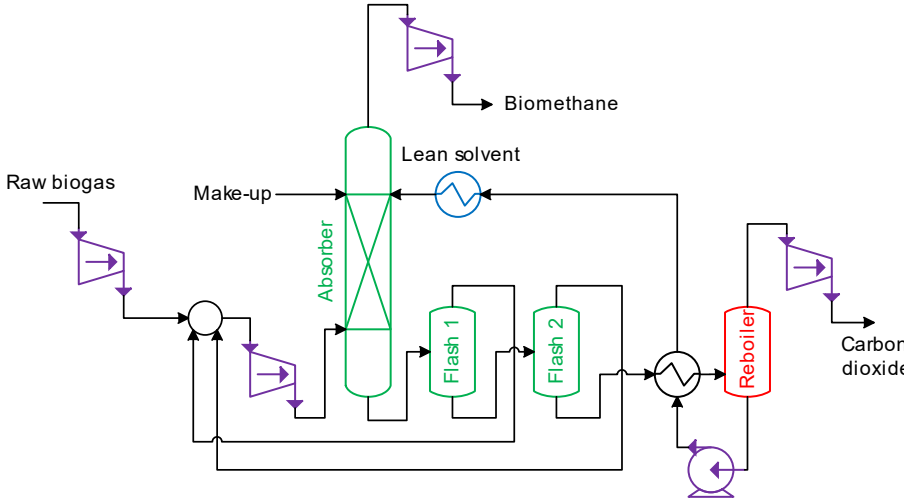


Figure 6.5. Proposed process configuration of the physical and water scrubbing technologies with CO_2 recovery (PWCR).

6.4.1.2. Without CO_2 recovery

The traditional process configuration of physical scrubbing processes regenerates the solvents using air stripping, i.e. stripping the solvent with an air stream that reduces the partial pressures of the CO_2 and CH_4 in the vapor phase and allows the

vaporization of the gases by creating a difference in the chemical potential. The **P**hysical/**W**ater process configuration (PW) analyzed in this work has been reported in different sources [126,127] and is presented in Figure 6.6. One can notice that, as opposed to the PWCR, the solvent regeneration system of the PW has a heat exchanger and a reboiler instead of a desorber column. The L/G ratio between the rich solvent and the air supply stream in the desorber column is specified and it is used to calculate the composition of the lean solvent exiting the bottom of the absorber.

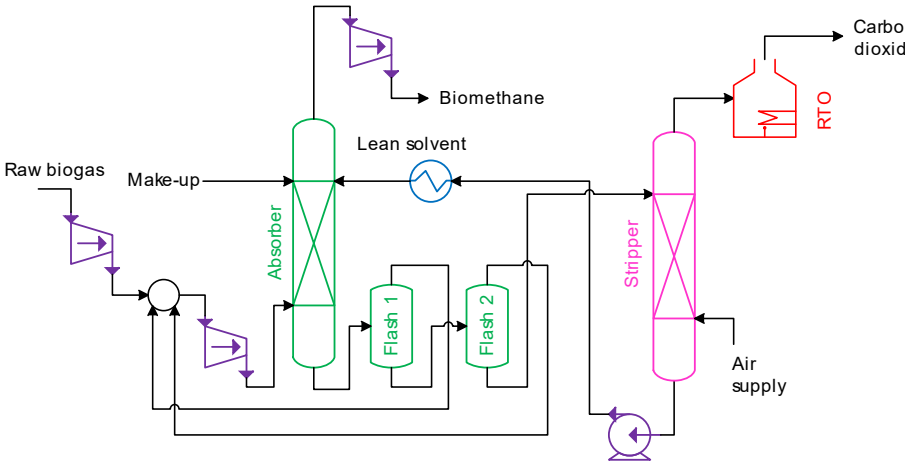


Figure 6.6. Proposed process configuration of the physical and water scrubbing technologies without CO_2 recovery (PW).

The vapor stream released from the desorber contains both CH_4 and CO_2. While the amount of CH_4 is expected to be lower than that of CO_2, the greenhouse effect and environmental impact of CH_4 methane is significantly higher when compared to CO_2 [128]. Therefore, the costs associated to the release of CH_4 to the atmosphere are more than 20 times higher than those from releasing CO_2 [129]. In order to eliminate CH_4 emissions, the gas released from the stripper is treated with a regenerative

thermal oxidation (RTO) unit that transforms the CH_4 into CO_2. The destruction efficiency of thermal oxidizers can be close 100% [130], hence, in this work, we assume that the CH_4 conversion is 99% and we consider that the RTO operates at 750 °C in order to ensure a complete conversion of methane (the autoignition temperature of CH_4 is 550 °C).

6.4.2. Chemical absorption process

The chemical process flowsheet of the **A**queous a**M**ine absorption process without CO_2 Recovery (AM) and the **A**queous a**M**ine absorption process with **C**O_2 **R**ecovery (AMCR) are shown in Figure 6.7. As opposed to the physical solvent processes, the process configuration does not change significantly whether CO_2 recovery is considered or not. In fact, the only difference lies on the inclusion of the CO_2 compressor in the carbon dioxide stream.

Considering that in this process most of the CO_2 is solubilized by chemically binding itself to the amine, solvent regeneration is usually carried at low pressures. The vapor-liquid equilibrium behavior is calculated using the machine-learning based surrogate model validated in a previous work [122]. On the other hand, H_2O must be included in the phase equilibria calculations because its partial pressure can be equal or greater than the CO_2 partial pressure. MEA is assumed to be non-volatile because its boiling point is considerably smaller than that of the other components.

The role of the absorber in this process is the same as in the physical absorption processes, but the difference is that the absorber pressure P_A is assumed to be constant and equal to 100 kPa. This process also uses a lean liquid solvent stream (F_L) at a given temperature (T_L), CO_2 loading (α_{Lean}) and a 30 %wt. MEA solvent.

Analogously to the physical and water scrubbing process, the (L/G) ratio is used to specify F_L.

The model considers the absorber to be isobaric, adiabatic, and that both phases are in thermal equilibrium at the top. This simplification is justified when one looks at the pilot plant data from Tobiesen et al. [131]. It is reported that, in average, the liquid temperature is 5 °C higher than the vapor phase temperature at the top.

The assumptions done in the absorber allow the estimation of the compositions and the temperature of the outlet streams without using complex differential models like the ones presented elsewhere [131–133]. An inconvenience of this simplification is that it does not estimate the temperature profile, hence the location of a temperature bulge is unknown. A large temperature bulge can lead to an undesired pinch, and the CO_2 mass transfer flux can decrease significantly [134]. Not considering the temperature bulge inevitably underestimates the minimum L/G ratio and the energetic requirement for the separation. In order to consider the temperature bulge effect, a maximum theoretical temperature ($T_{L,Max}$) is calculated by performing an overall mass and energy balance and neglecting water evaporation. Note that the value of $T_{L,Max}$ is always higher than the actual liquid temperature flowing from the bottom of the absorber. In this work we assumed that if $T_{L,Max}$ is estimated to be greater than 90 °C, then the absorber is deemed unfeasible.

The heat exchanger recovers the heat from the reboiler outlet and fixes a temperature difference in the hot side constant and equal to 10 °C. The reboiler temperature (T_R) is specified, and along with α_{Lean} and the aqueous solvent composition, the reboiler pressure can be estimated (P_R). The desorber pressure is assumed to be equal to the reboiler pressure. The temperature difference between the reboiler and the top of the

desorber ΔT_D is specified to be equal to 20 °C, while the condenser temperature of the partial condenser of the desorber is 20 °C in accordance to reported plant data [135]. Although a small ΔT_D would require less energy consumption, the size of the desorber would become unfeasibly large.

The carbon dioxide stream (F_C) is produced by condensing most of the water content that exits the top of the desorber. It is important to underline that since the solubility of methane is disregarded in the chemical absorption plant model, x_{CH_4} is equal to 0. The make-up W compensates the amount of water lost in the absorber or desorber.

As opposed to the physical scrubbing processes, the amine-based processes require less compression sections. The AM process only requires the biomethane compressor while the AMCR process requires the compressors for the biomethane and the carbon dioxide streams.

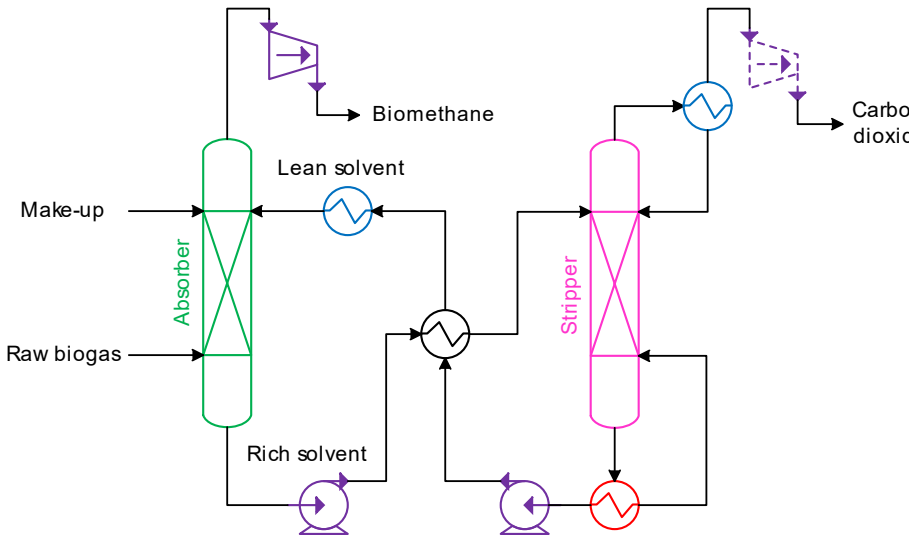


Figure 6.7. Process configuration of the chemical scrubbing technology with CO_2 recovery (AMCR) and without CO_2 recovery (AM). The dashed compressor indicates that it is only present in the AMCR process.

6.4.3. Auxiliary systems

6.4.3.1. Compressors

The compressors are modelled as isentropic; hence the following equation can be applied to determine their power consumption [136]:

$$W_K = \frac{m \cdot (h_2^* - h_1)}{\eta_{K,S} \cdot \eta_{K,M}} \tag{6.10}$$

where W_K is the compressor power consumption, m is the material flow of the unit operation, h_2^* is the specific enthalpy of the outlet for the isentropic process, h_1 is the specific enthalpy of the inlet, $\eta_{C,S}$ is the isentropic efficiency and $\eta_{C,M}$ is the mechanical efficiency. The efficiency values for the compressors and other auxiliary systems are shown in Table 6.6.

Since the biogas upgrading processes requires considerable pressure increments, the compression trains must be divided in compression stages. The number of compression stages in each compression section is calculated using the compressor ratio equation [137]:

$$\log(CR) = \frac{\log\left(\frac{P_a}{P_b}\right)}{n} \tag{6.11}$$

where P_a is the pressure feed to the compressor sequence, P_b is the outlet pressure of the compressor train, *n* is the number of compressors and CR is the compression ratio. It is advised that compressors do not work at high temperatures (this work avoids operating compressor temperatures above 150 °C), therefore the maximum value allowed in this work for CR is 4.0 in order to comply with this restriction.

Compression sequences usually include intercooling stages that lower the temperature from the outlet of the compressors to a low intercooling temperature (T_I). This is done to improve the energetic performance of the compressor system and to remove any undesired volatile components from the gas mixture. T_I in this work is fixed and equal to 25 °C for all solvents except for methanol which uses 0 °C because of its high volatility. The intercooling stages in the biomethane and the carbon dioxide streams are used for cooling the compressed gas, removing any organic solvent/water from the gas stream, and recirculating the condensed phase to the make-up stream of the process.

6.4.3.2. Pumps

Hydraulic pumps are required to compress and transport the liquid solvent. The electrical energy required by a pump (W_P) is given by:

$$W_P = \frac{m \cdot \Delta P}{\rho \cdot \eta_P} \tag{6.12}$$

where m is the material flow of the fluid, ΔP is the pressure difference between the outlet and the inlet of the pump, ρ is the fluid density and η_P is the mechanical efficiency of the pump. The ΔP is calculated as the pressure difference between the pressure of the connected equipment plus a constant value of 300 kPa in order to account for the pressure drop caused by hydrostatic and frictional forces.

6.4.3.3. Cooling system

The cooling system is used to cool the service fluid (water is used for this purpose) to lower the temperature of the main unit operations in the biogas upgrading plant. This system is composed of a cooling tower, a blower and a hydraulic pump.

The equation for estimating the blower power (W_B) is given by:

$$W_B = \frac{m \cdot \Delta P}{\rho \cdot \eta_B} \tag{6.13}$$

where η_B is the blower efficiency. It is considered that the minimum temperature which can be reached by cooling water with air is $T_U = 20\,°C$. The value of T_U is dependent on the geographical location of the biogas upgrading plant. A temperature increase of 10 °C is allowed on the cooling water before it is sent to the cooling tower again. Moreover, a minimum temperature gradient of 5 °C is considered as needed in order to cool a stream, hence, any stream that requires to be cooled less than $T_U + 5\,°C$ must use a fluid service cooled by the chilling system.

6.4.3.4. Chilling system

The energy needed by the chilling system is calculated with the coefficient of performance (COP), which is the ratio between the electrical energy consumption of the chilling system (W_F) and the amount of energy that must be removed using the refrigerant (W_{Chill}). The energetic consumption of the chilling system can be calculated by:

$$W_F = COP \cdot W_{Chill} \tag{6.14}$$

There has been extensive research in order to improve the energetic efficiency of different chilling systems [138]. The COP of the system is a function of the chilling temperature (T_{Chill}), hence the COP can be calculated with the following empirical equation

$$COP = 0.001 \cdot exp(0.301 \cdot T_{Chill}), \tag{6.15}$$

which is valid for chilling systems operating with CO_2 as refrigerant and the M-SC configuration as reported by Bellos and Tzivanidis [138].

6.4.3.5. Reboiler

This system provides the necessary calorific energy to heat up the solvent in the reboiler in order to regenerate it. It is considered that the produced biomethane is used as the fuel source. Henceforward, the reboiler duty (Q_R) can be calculated in terms of combusted biomethane (F_{Burnt}):

$$F_{Burnt} = \frac{Q_R}{y_{CH_4} \cdot \Delta H_{CH_4}} \tag{6.16}$$

where ΔH_{CH_4} is the combustion enthalpy of methane and is equal to 890.5 kJ/mol. In the context of the PW process, the term Q_R is the energy associated to the RTO unit.

6.4.4. Economic evaluation

The economic assessment of the biogas upgrading technologies of this work captures the effect of the biomethane sold to the user, the operational costs, and the taxes/penalties associated to the release of greenhouse gases to the atmosphere. The concept of integrating the environmental impacts with monetary values has been standardized within the ISO 14008:2019 framework [139]. By considering this, the following equation is proposed:

$$\begin{aligned} \Pi = \big(0.844 \cdot C_Q \cdot y_{CH_4} \cdot (F_B - F_{Burnt}) - 0.278 \cdot C_E \cdot E - 0.044 \cdot C_T \\ \cdot \big(x_{CO_2} \cdot F_C + 27.8 \cdot x_{CH_4} \cdot F_C + F_{Burnt}\big) - x_S \cdot F_C \cdot M_S \cdot C_S\big) \cdot \frac{1}{Q_F}, \end{aligned} \tag{6.17}$$

where Π is the unit profit in USD/Nm^3 of processed raw biogas, Q_F is the normal cubic meter of raw biogas, C_Q is the value of the biogas per energy unit in USD/MMBTU, C_E

is the electricity price in USD/kWh, E is the total electricity consumption of all the auxiliary equipment in MJ/s, C_T is the CO_2 tax paid in USD/ton CO_2, C_S is the cost of the solvent in USD/ton, the material flows F_B, F_{Burnt} and F_C are in kmol/s and M_S is the molecular weight of the solvent in kg/kmol. The values of C_Q, C_E and C_T are presented in Table 6.6. The values of C_S were set as follows: 0 USD/ton for water, 600 USD/ton for methanol [140], 2,000 USD/ton for MEA [93], 1000 USD/ton for PEGDME, NMP and NFM (assumed value).

The first term of Eq. (6.17)) is the combined effect of the revenue generated by producing the biomethane minus the cost associated to the reboiler, the second term is the cost associated to the electricity consumption of the plant, the third term is the tax or penalty that must be paid for releasing greenhouse gases (CO_2 or CH_4) to the atmosphere, and the last term represents the solvent loses due to solvent vaporization (PW/PWCR) or amine degradation (AM/AMCR). The numeric factor associated to the CH_4 flow implies that the CH_4 emissions are 27.8 times more expensive than those of CO_2. This ratio was estimated by dividing the average emission values reported by Lombardi and Francini [129]. In this equation we assume that the emission taxes ratio remains constant.

In the case that the policies do not require the biogas upgrading plants to pay CO_2 taxes, the last term C_T can be the value of a CO_2 credit. These credits can be sold between different commercial entities. If one wants to use Eq. (6.17) for the process evaluation with CO_2 credits, then the last constant should be changed from −0.044 to +0.044.

6.4.5. Key variables summary

The most important variables for each one of the assessed technologies in this work are presented in Table 6.6. Each variable with an independent tag is an available degree of freedom. The values for the compressor efficiencies were taken from recommended values [141]. The price of biomethane is an average value taken from the energy market [142] and the C_E value is the average retail electricity price for industrial consumers in the U.S. in 2019 [143].

It is important to remark that the integration of CO_2 recovery affects the parameters in which the process can be feasible. If Table 6.6 is looked at, one can realize that there are 4 degrees of freedom in the 4 processes (x_{CH_4} can be considered as a degree of freedom because it is a restriction).

The reboiler/desorber pressures P_R for both the PW and PWCR processes were fixed to be equal to the atmospheric pressure. The reason for this is that operating at lower P_R implies that P_A will be low as well. Operating at low absorber pressure decreases the electricity output of the compressors, thus making the processes economically more competitive.

Table 6.6. Summary of the key variables used in the biogas upgrading processes.

Variable	Process			
	PWCR	PW	AMCR	AM
Process variables				
L/G (absorber)	Independent			
L/G (desorber)	N/A	Independent	Dependent	
T_L / °C	Independent			
T_R / °C	Independent	N/A	Independent	
P_R / kPa	100	100	Dependent	
r_P	Dependent	Independent	N/A	
α_{Lean}	N/A		Independent	
P_A / kPa	Dependent		100	
Specifications				
z_{CH_4}	0.5 – 0.9			
y_{CH_4}	0.98			
x_{CH_4}	≤ 0.04	Dependent	≤ 0.04	Dependent
T_B / K	298.15			
P_B / kPa	100 – 5,500			

Table 6.6. Summary of the key variables used in the biogas upgrading processes.

Variable	Process			
	PWCR	PW	AMCR	AM
T_C / K	283.15			
P_C / kPa	150,000			
Equipment efficiency				
$\eta_{C,S}$ / %	80			
$\eta_{C,M}$ / %	97			
η_P / %	80			
η_B / %	70			
Economic variables				
C_Q / $USD \cdot MMBTU^{-1}$	2.50			
C_E / $USD \cdot kWh^{-1}$	0.0683			
C_T / $USD \cdot ton^{-1}$	≥ 0			

* P_R is the desorber pressure in the PW process

6.4.6. Optimization algorithm

Comparing different technologies is a difficult task, especially if the processes greatly differ one from another. Although the PWCR and the PW processes may seem alike at first glance, the fact that the former alternative includes a reboiler entirely changes the process energetic and economic performance. For this reason, in order to have a fair comparison and judgement of each technology, it is suggested to compare the different technologies only when their operating parameters are optimized.

The optimization algorithm in this work aims to identify operating areas in which the different technologies may be economically competitive. Hence, a relaxed optimization method such as the response surface analysis is ideal. This method has been successfully applied to other complex separation methods [144].

The first step involves selecting the specifications of the process and setting the limits of the independent variables. The limits of important variables for the physical and chemical solvents are presented in Table 6.7.

The minimum and maximum temperature values for the physical solvents were taken from reports of already-existing processes described in section 2. In the case of NFM and NMP, the maximum temperatures were set so that their partial pressure does not exceed 1.0 kPa. The minimum solvent molar composition of water must be kept high in order to avoid hydrate formation [87]. Note that the NFM case assumes that the solvent is a mixture of NFM and N-acetyl-morpholine in order to decrease the freezing point, so that operations at low temperatures are feasible.

Table 6.7. Limits of the operating parameters of the solvents in the optimization.

	PEGDME	NFM	NMP	Methanol	H_2O	MEA
Minimum temperature / °C	0	−20	−15	−60	10	40
Maximum temperature / °C	175	100	60	−10	60	120
Maximum CO_2 molar fraction	0.3	0.3	0.3	0.3	0.03	0.5*

* CO_2 loading: in mol CO_2/mol MEA.

The response surface method relies on creating a design of experiments (DOE) and then performing the simulations. The process performance is particularly sensitive to the L/G ratio, therefore it was found that it was best to perform the optimization in two steps. First, the simulations were carried at certain conditions (e.g., in the PWCR process, at fixed values of T_L, T_R and x_{CH_4}) and then an univariable optimization method was used to find the best L/G ratio. The full optimization scheme is presented in Figure 6.8**.**

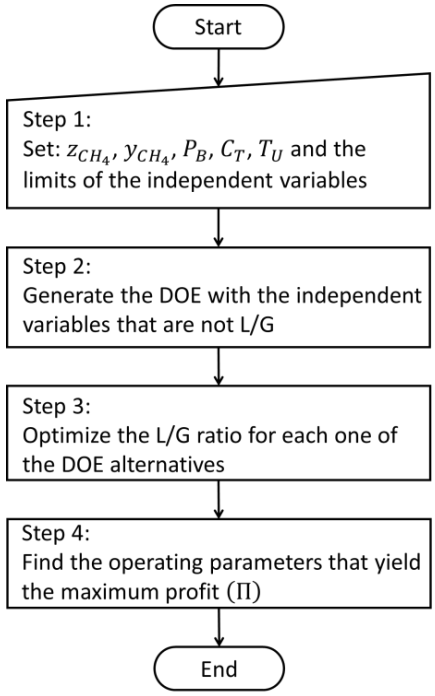


Figure 6.8. Optimization algorithm.

6.5. Results and discussion

6.5.1. Physical absorption processes

An overview of the operating conditions and their effect on the energetic and economic performance of both physical scrubbing technologies PWCR and PW is presented in this subsection. The physical solvent PEGDME is used as the case study to exemplify how the techno-economic analysis of the PWCR and PW processes were made. We perform the economic evaluation of four different organic solvents and water.

6.5.1.1. With CO_2 Recovery

An overview of the operating conditions and their effect on the energetic and economic performance of both physical scrubbing technologies PWCR and PW is presented in this section. The physical solvent PEGDME is used as the case study to exemplify how the techno-economic analysis of the PWCR and PW processes was made.

Usually, the physical scrubbing technologies are denominated as pressure swing processes because they rely on a pressure difference to perform the separations. This is true for the PW process. However, for the case of PWCR, the inclusion of a reboiler enables a mixed temperature-pressure swing separation. Considering this, the optimal process operating conditions are expected to differ to those from the PW process.

An analysis of the PWCR process is presented in Figure 6.9 in the form of contour plots. The optimization was done using a 7 level DOE for the T_R and the T_L while $z_{CH4} = 0.7$, $P_B = 1.2\ MPa$ and $T_U = 20\ °C$. The variables that are not explicitly defined use the values presented in Table 6.6.

Figure 6.9 (a) shows that T_L has a direct relationship with the optimized L/G ratio (kg/kg). This is expected as lower temperatures allow more gas solubilization into the

liquid phase, hence less solvent will be required. On the other hand, the reboiler temperature does not seem to have a noticeable effect on the optimized L/G ratio.

The behavior of P_A with respect to the operating temperatures is presented in Figure 6.9 (b). It shows that at larger T_L, the absorber operating pressure will be lower. The fact that P_A decreases when T_L is higher may seem counterintuitive, but this behavior is a combined effect caused by the tradeoff that exists between the optimized L/G ratio and T_L: a larger T_L means a larger L/G, which conversely means that the absorber pressure P_A can be lowered.

The effect of T_L on the depressurization ratio is presented in Figure 6.9 (c). The recirculation rate of the flash recycle tanks and their corresponding pressure drops are inversely affected by r_P. With a higher T_L, the selectivity CO_2/CH_4 decreases for all physical solvents, and thus the depressurization ratio must be lower to avoid high methane losses.

Figure 6.9 (d) presents the calorific energy requirements, where one can observe that the calorific requirement for the PWCR process increases at larger T_L. Although one may expect that larger temperatures would imply lower Q_R, this is not true, because the optimized L/G ratios increase with larger T_L values. This means that the system will spend more calorific energy to heat up the solvent and to desorb CO_2 (note that the calorific energy and all other variables are normalized with respect to the amount of processed raw biogas, hence, Nm^{-3} refers to the normal cubic meters fed to the process).

The electrical energy requirement as a function of the operating temperatures of the process is presented in Figure 6.9 (e). The electrical energy consumption is a lumped effect of several energy requirements needed in the process: compression, pumping,

cooling system and chilling system duties. Larger T_L values allow the operation of the process at lower pressures (higher L/G), which in turn will reduce the pumping and compression duties. Moreover, the cooling duty is drastically reduced when $T_L \geq 25\ °C$ because the cooling is done with water instead of employing the chilling system. The combination of these effects causes larger T_L to induce less electrical energy requirements.

The combined effect of the calorific and electrical energy can be assessed by observing the profit (Eq. (6.17)) surface plot in Figure 6.9 (f). The effect of T_R on the profit equation indicates that heating the solvent up to 60 to 80 °C has a beneficial effect in the process performance. This is because a larger temperature enables the desorption of more gas from the liquid phase, hence the lean solvent will have less dissolved gases and more capacity to absorb. It is interesting to observe that the surface topology of Figure 6.9 (f) is almost symmetrical to that of the electrical energy. This is not a surprise if we consider that, according to Table 6.6, the cost ratio between the electrical and calorific energy is 8:1. In fact, if the price of electricity is increased significantly, the topology of the electrical requirement and profit response surfaces become symmetrical.

A similar response surface analysis was performed on the PWCR process at different concentrations of CO_2 in the feed and the results are presented in Table 6.8. It is shown that the optimal L/G ratio has a direct relationship with the value of z_{CH_4}. This behavior is a result of the interaction between the L/G ratio and the molecular weight of the gases. Although the total amount of gas in mass basis is decreasing with larger CO_2 concentrations, the total amount of gas in mole bases is increasing, which means that a larger higher quantity of solvent will be needed to keep a lower pressure.

The analysis on these results also suggests that the dependency of the optimized T_L and T_R with respect to z_{CH_4} is not significative. However, this is not the case for all process variables as r_P, θ and E have a smooth inverse relationship when compared to z_{CH_4}. The results show that at larger z_{CH_4} the depressurization ratio must be lower, hence the recirculation of the recycle tanks must increase as well. The cause of this is that larger z_{CH_4} involves more CH_4 dissolved in the mixture, which must be removed from the solvent so that the carbon dioxide stream can meet the specification of x_{CH_4}. The reduction in the methane loss and the operating costs eventually cause that the profit increase proportionately with respect to z_{CH_4}. This is reasonable as it is easier to perform a separation from $z_{CH_4} = 0.9$ to $z_{CH_4} = 0.98$ than if the separation starts at $z_{CH_4} = 0.5$.

The share of the costs associated to the conditioning of the carbon dioxide stream (K_C) is shown in Table 6.8. These costs include the electricity consumed by the compressors and the cooling power needed in their interstage cooling sections. The inverse relationship between K_C and z_{CH_4} is caused by the reduction of the material flow in the carbon dioxide stream with a larger CH_4 feed concentration.

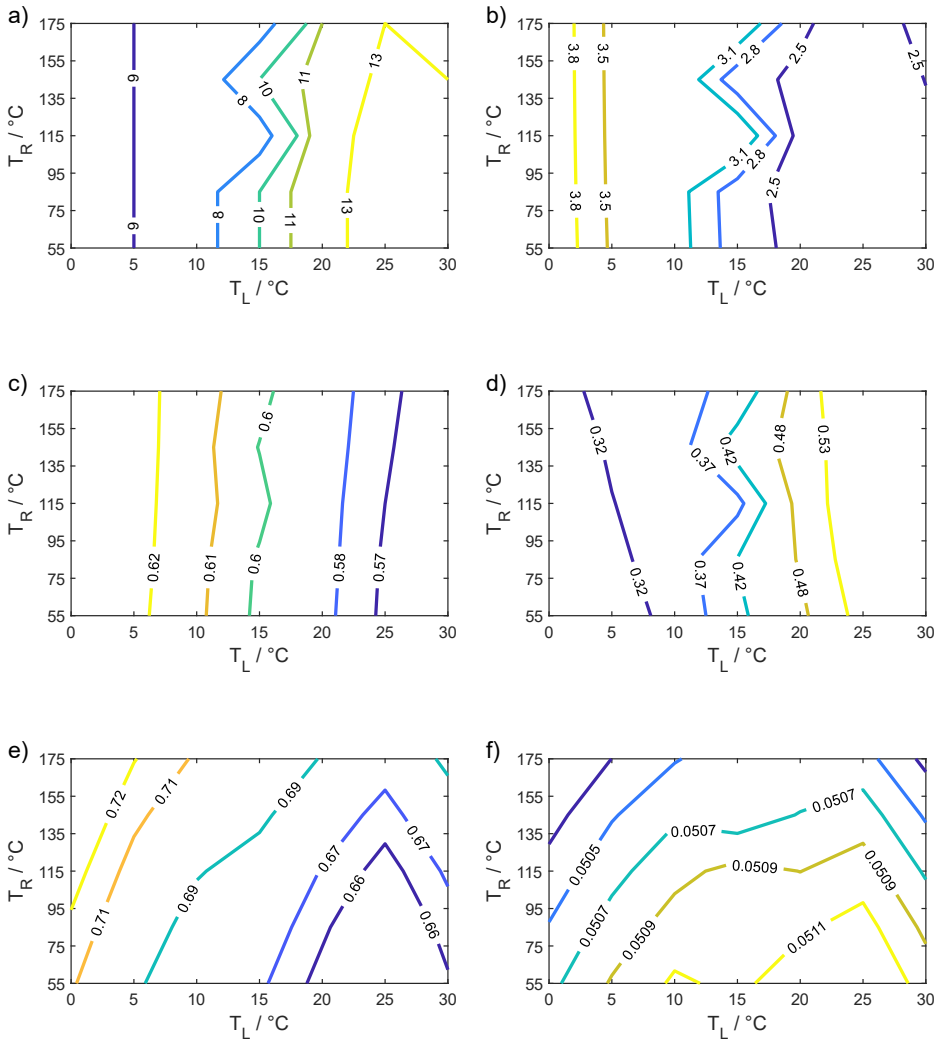


Figure 6.9. PWCR process contour plots of: (a) L/G in kg/kg, (b) absorber pressure P_{Abs} in MPa, (c) depressurization ratio r_P, (d) Q_R in MJ/Nm^3, (e) E in MJ/Nm^3 and (f) Profit Π in USD/ Nm^3. $z_{CH4} = 0.7$, $P_{Bio} = 1.2\ MPa$ and $T_U = 20\ °C$. Solvent: PEGDME.

Table 6.8. Optimized parameters of the PWCR process using PEGDME as the solvent.

z_{CH_4}	0.5	0.6	0.7	0.8	0.9
L/G (kg/kg)	10.5	12.5	14.5	17	20
T_L / °C	25	25	25	25	25
T_R / °C	70	70	70	70	70
P_A / MPa	2.4	2.3	2.2	2.1	2.1
r_P	0.76	0.67	0.57	0.44	0.27
P_R / MPa	0.1	0.1	0.1	0.1	0.1
x_{CH_4}	0.04	0.04	0.04	0.04	0.04
θ	0.041	0.027	0.017	0.010	0.004
Lean solvent molar fraction	0.99	0.99	0.99	0.99	0.99
Rich solvent molar fraction	0.67	0.72	0.76	0.80	0.84
Q_R / MJ/Nm^3	0.59	0.58	0.56	0.54	0.51
E / MJ/Nm^3	0.77	0.70	0.64	0.60	0.57
P_B / MPa	1.2	1.2	1.2	1.2	1.2
K_C / %	59.4	52.5	43.6	31.8	18.2
Profit Π / USD/Nm^3	0.029	0.040	0.051	0.062	0.072

A comparison of five different physical solvents (4 organic solvents and water) is presented in Figure 6.10. The organic solvents PEGDME, NFM and NMP seem to be equally competitive over almost all compositions of z_{CH_4}. Due to the high methanol volatility, the process must be operated at cryogenic temperatures. This causes considerable electricity demands due to the additional dependence on the chilling system. However, this is not the only cause of economic loses, the methanol that cannot be recovered within the compression system should also be replaced. While more than 99.99 % of the evaporated PEGDME, NFM and NMP can be recovered in the interstage cooling systems, only 98-99 % of the methanol can be recovered. Although the methanol loss may seem low one should consider that adding an extra component into the CO_2 transport and storage systems may raise unforeseen operating issues. Despite the fact that the water absorption capabilities are worse than those of methanol, its low volatility and negligible prices make it more economically competitive at lower z_{CH_4}.Nonetheless, neither methanol nor water seem to be economically competitive.

It is interesting that, despite the fact that PEGDME, NFM and NMP have different thermodynamic properties, they seem to be equally competitive. For example, PEGDME has the worst selectivity of all the compared solvents, however, this is compensated with having the lowest Henry's coefficient (3.7 MPa at 25 °C) and a high boiling point (> 250 °C). This last property implies that the solvent can be regenerated at higher temperatures without incurring into considerable solvent losses. On the other hand, NFM and NMP have higher Henry's coefficients (e.g. 7.1 and 6.8 MPa at 25 °C respectively) than PEGDME, but their higher selectivity and lower molecular weight outweigh the lower CO_2 solubility. The lower Henry's coefficient of PEGDME may be

advantageous from a capital costs perspective as it means that the shell thickness of the absorber will be smaller and, therefore, more economical.

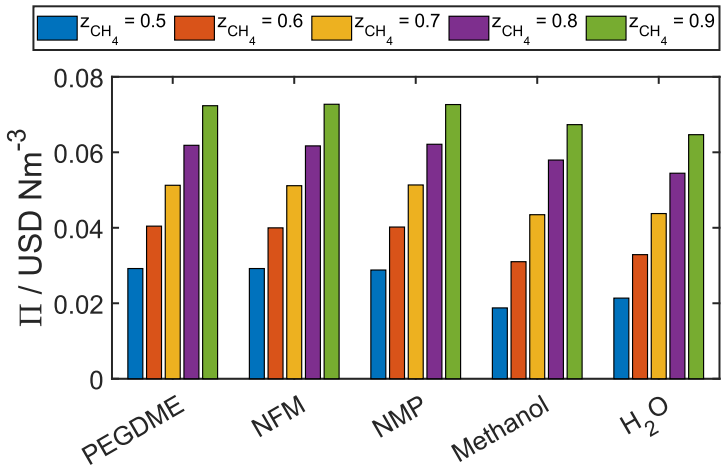


Figure 6.10. Profit Π of each one of the solvents using the PWCR process technology at different methane concentrations in the feed.

Water seems to not be energetically or economically competitive in the PWCR because its Henry's coefficient is around 10 to 20 times greater than in the studied organic solvents. This indicates that the process will be required to operate at larger pressures, and hence the electricity consumption will be larger. Moreover, larger L/G ratios are needed in order to counter the large Henry's coefficient values and to avoid hydrate formation. This causes that its economic performance is around 10 to 25% worse than the competing technologies. In addition to this, since the solubilization of CO_2 into water produces an acidic mixture, it is necessary to use stainless steel to avoid corrosion and the capital costs of the equipment will increase by a factor of 2.4–3.4 (depending on whether stainless steel 304L or 316L is used) [130].

Due to the similar competitiveness of the PEGDME and NMP, it is important to consider the transport properties. In particular, viscosity is an important property as it may increase significantly the equipment size if it is too large. In Table 6.9 we include the minimum and maximum viscosity values that can be found in the PWCR process. From this perspective, the solvent competitiveness could be rearranged as H_2O > NMP > Methanol > PEGDME > NFM (the actual viscosity values the NFM-based solvents also accounts for N-acetyl-morpholine). Although NFM, methanol and H_2O have competitive viscosities when compared to the other organic solvents, their thermodynamic and corrosion-related disadvantages are too substantial to be countered by low viscosities alone.

Table 6.9. Solvent viscosity as a function of the operating temperatures.

Solvent	Viscosity / $\mathrm{mPa \cdot s}$ at lowest T_L	Viscosity / $\mathrm{mPa \cdot s}$ at highest T_R	Source
PEGDME	6.5 at 25 °C	2.4 at 70 °C	[146]
NFM*	17.1 at −20 °C	8.3 at 25 °C	Aspen Plus v8.6
NMP	1.5 at 15 °C	1.1 at 54 °C	[147]
Methanol	3.25 at −60 °C	0.8 at 0 °C	[148]
H_2O	1.31 at 10 °C	0.5 at 47 °C	[149]

* These values account for the mixture between N-formyl-morpholine and N-acetyl-morpholine.

Providing a definite conclusion about whether PEGDME or NMP is better might not be possible without performing a detailed engineering design. However, an analysis on the optimal T_L for each solvent, might suggest that PEGDME has a slight advantage because the T_L in all cases is at 25 °C while the minimum temperature of

NMP is at 15 °C. This implies that PEGDME, might not need to require a chilling system. Nonetheless, we remark, the decision of the optimal solvent requires a thorough analysis that considers the plant location, solvent availability, solvent price and capital investment requirements.

6.5.1.2. Without CO_2 Recovery

It is not possible to fully understand the difference between the processes with CO_2 recovery and without CO_2 recovery by only analyzing the effect the of compression system. For this reason, the same optimization procedure was performed on the PW process at different z_{CH_4} conditions when using PEGDME as the solvent. The results of these calculations are shown in Table 6.10.

The L/G ratios in the PW process are highly influenced by z_{CH_4}. In contrast to the PWCR process, the L/G ratios in the PW process are significantly larger than in the PWCR process. This suggests that the PWCR process relies more on the pressure increase to perform the separation while the PW process depends more on the solvent flow. Since the PWCR process requires higher absorber pressure, the pressure drop between the absorber and the recycle tanks is enough to be able to vaporize the methane and reduce its concentration in the desorber vapor outlet. Nonetheless, it is more expensive to operate the process at larger pressures, therefore, the optimal L/G values in the PW process tend to be higher than in the PWCR process.

Another significant difference is the optimal T_L. While for the PWCR process larger temperatures are preferred, for the PW it seems to be the opposite. This can be explained by considering the fact that a low T_L reduces the absorber pressure and that the PWCR has to operate at larger pressures in order to comply with the x_{CH_4} restriction.

Although the PWCR process relies on a chemical potential shift caused by a combined temperature and pressure swing in contrast to the PW process, the PWCR requires less input calorific energy. This is caused by the implementation of the RTO. We can observe in Table 6.10 that the amount of calorific energy used to abate the methane in the PW process is significantly higher than in the reboiler of the PWCR process. This is because the PW process increases the temperature from low temperatures to around 750 °C, while the PWCR process increases the temperature of a liquid around 50–70 °C. Moreover, the share of the costs associated to the carbon dioxide stream conditioning – K (i.e., the operating costs of the RTO system) is significantly higher at low z_{CH_4}. This can be explained by analyzing the absorber L/G, the smaller the L/G rate is, the less amount of liquid will be fed to the desorber and, hence, less air supply flow rate will be needed to be heated. Note that the carbon dioxide stream in the PW process is a mixture of CH_4, CO_2, air and evaporated solvent. This additional complexity of the PWCR process gives an economic disadvantage with respect to the PW process. More than 95 % of the calorific energy penalty can be removed if a cross heat exchanger is used to recover the energy of the carbon dioxide stream. If a heat exchanger is added, the profit increases between 1 % for larger z_{CH_4} and 5 % for smaller z_{CH_4}. Comparing both processes when there is no penalty for the CO_2 emissions is intrinsically unfair, as the PWCR process was developed to comply with the specifications of both biomethane and carbon dioxide streams.

Table 6.10. Optimized parameters of the PW process using PEGDME as the solvent when $C_T = 0$.

z_{CH_4}	0.5	0.6	0.7	0.8	0.9
Absorber mass basis L/G	143	123	107	82	42
Desorber mass basis L/G	300	300	300	300	300
T_L / °C	0	0	0	0	0
r_P	0.05	0.05	0.05	0.05	0.05
P_R / kPa	100	100	100	100	100
$x^*_{CH_4}$/ 10^6	161	183	239	299	284
θ	0.02	0.02	0.01	0.01	0.00
Lean solvent molar fraction	0.98	0.98	0.98	0.98	0.98
Rich solvent molar fraction	0.95	0.95	0.94	0.94	0.91
Q_R / MJ/Nm^3	1.38	1.08	0.76	0.47	0.21
E / MJ/Nm^3	0.23	0.27	0.29	0.33	0.38
P_B / MPa	1.2	1.2	1.2	1.2	1.2
K / %	40.2	31.5	22.6	13.7	5.7
Profit Π / USD/Nm^3	0.039	0.048	0.058	0.067	0.077
Profit with integrated heat Π / USD/Nm^3	0.042	0.051	0.060	0.069	0.077

* In an air-free basis.

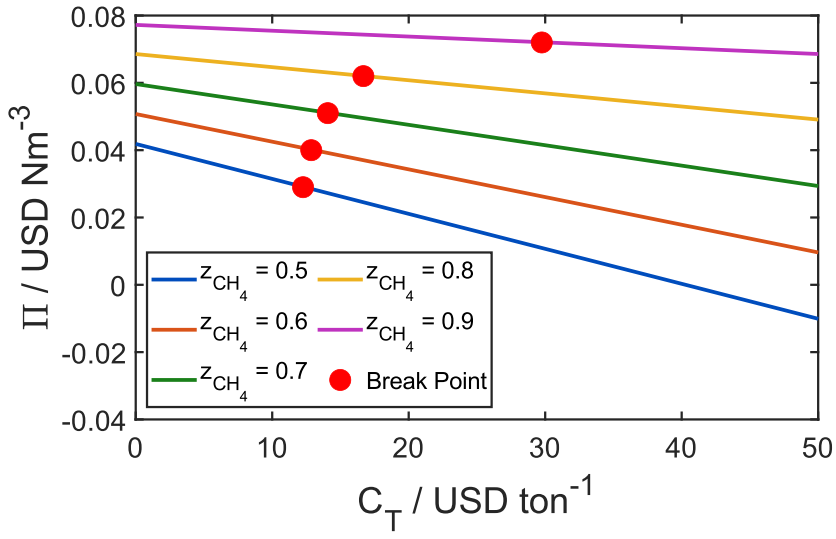


Figure 6.11. Dependence of the PW process profit Π as a function of the CO_2 tax value. The break point is the CO_2 tax value in which the PW and PWCR have the same economic performance. Solvent: PEGDME.

The economic performance of the PW process as a function of C_T is presented in Figure 6.11 using PEGDME as a study case. Figure 6.11 highlights that processes with less z_{CH_4} will be more affected if the biogas upgrading process has not implemented CO_2 recovery. Furthermore, Figure 6.11 also shows the break points in each one of the methane concentrations isolines. A break point is the value of the CO_2 tax in which the PWCR process and the PW processes have the same economic competitiveness. From the break point onwards, the PWCR process will always be economically more competitive. It is important to underline that since the processes with CO_2 recovery do not have CO_2 emissions, their optimal operating conditions are independent of the C_T value. In most cases, the value of the CO_2 tax in the break point is around $12\ USD/ton$.

The values of the CO_2 tax depend on the geographical area in which the biogas upgrading plant is to be built. Therefore, analyzing the breakpoint may be important to forecast the upgrading plant revenues, and most importantly, to decide whether the implementation of CO_2 recovery into the process is feasible or not.

These results assume that the biomethane plant must pay CO_2 taxes. If the biomethane plant sells CO_2 credits, then the results would be linearly translated to higher profit values (this can be seen in the fact that Eq. (6.17) has a linear dependence on C_T).

6.5.2. Chemical absorption processes

In the case of chemical scrubbing, the difference between the AMCR and AM processes is minimal because the only discrepancy is the carbon dioxide stream compressor. For this reason, the AMCR and AM processes share the same optimal operating conditions.

In the same fashion as the physical scrubbing processes, the operating parameters were optimized following the algorithm of Figure 6.8 and considering the constraints given in Table 6.7. The results of the optimization algorithm are presented in Table 6.11.

The main advantage of the chemical absorption processes over the physical absorption processes is the high selectivity of CO_2 and negligible loses of methane, hence, the production of biomethane is maximized with chemical scrubbing. The results show that larger z_{CH_4} require lower solvent flow rates because a smaller amount of CO_2 needs to be captured. Furthermore, a low concentration of CO_2 in the lean solvent requires less calorific and electric energy because smaller liquid solvent flow rates imply lower calorific and electric energy requirements. However, low CO_2

loading values (less than 0.2) are not desirable, because at lower loadings the energy required to perform the regeneration increases considerably.

Assuming that the effect of methane on the AM process is similar to nitrogen, the last case in Table 6.11 can be compared to the removal of CO_2 from the flue gas of a coal plant using MEA as a solvent. The value of $Q_R = 0.69\ MJ/Nm^3$ when $z_{CH_4} = 0.9$ is equivalent to 3.64 GJ/ton of CO_2 captured, which is similar to what is usually reported for the conventional setup of CO_2 capture with aqueous MEA at 30 %wt. [135]. Contrarywise, for $z_{CH_4} = 0.5$, estimated reboiler duties are reduced to 1.62 GJ/ton of CO_2 captured. Although the amount of heat needed is less than in the case with high methane concentration, the economic performance is inferior. This is caused by two main factors: higher CH_4 concentration implies more biomethane sold and lower CO_2 tax/credits.

One of the main operating concerns in amine-based chemical scrubbing plants is the thermal and oxidative degradation of the solvent. Chemical absorption plants using MEA to remove the CO_2 from a coal-fired plant have reported amine degradation values between 0.3 kg and 3.6 kg/ton CO_2 [150]. Due to the lack of oxygen in the biogas upgrading feed, we expect a lower degradation rate, hence, the degradation boundaries are set to be between 0 and 3 kg/ton CO_2. Table 6.11 shows that the economic impact of amine degradation decreases with larger z_{CH_4} due to a lower requirement of amine solvent. This effect is quite pronounced at low z_{CH_4}, where degradation can reduce the profit up to 35 % at maximum degradation conditions.

The difference between the AMCR and AM processes in the electrical energy consumption is the carbon dioxide compressor. In the case of the AM process, the electricity demand increases slightly when z_{CH_4} increases because the amount of gas

that needs to be compressed increases as well. Conversely, the electrical demand decreases when the amount of methane increases because the delivery pressure specification of the biogas $P_B = 1.2\ MPa$ is significantly lower than the delivery pressure of the carbon dioxide stream $P_C = 15\ MPa$.

Table 6.11. Optimized parameters of the AM and AMCR processes using 30 %wt. MEA as the solvent.

z_{CH_4}	0.5	0.6	0.7	0.8	0.9
Absorber mass basis L/G	13.5	12	10	7	4.5
α_{Lean}	0.24	0.24	0.24	0.24	0.30
T_L / °C	40	40	40	40	40
T_R / °C	120	120	120	120	120
α_{Rich}	0.50	0.50	0.50	0.50	0.50
Q_R / MJ/Nm^3	4.13	3.28	2.42	1.55	0.69
E – AM / MJ/Nm^3	0.28	0.31	0.35	0.38	0.42
E – AMCR / MJ/Nm^3	0.97	0.86	0.75	0.64	0.54
Profit Π (no degradation) / USD/Nm^3	0.019	0.032	0.046	0.059	0.073
Profit Π (maximum degradation) / USD/Nm^3	0.012	0.027	0.041	0.056	0.071

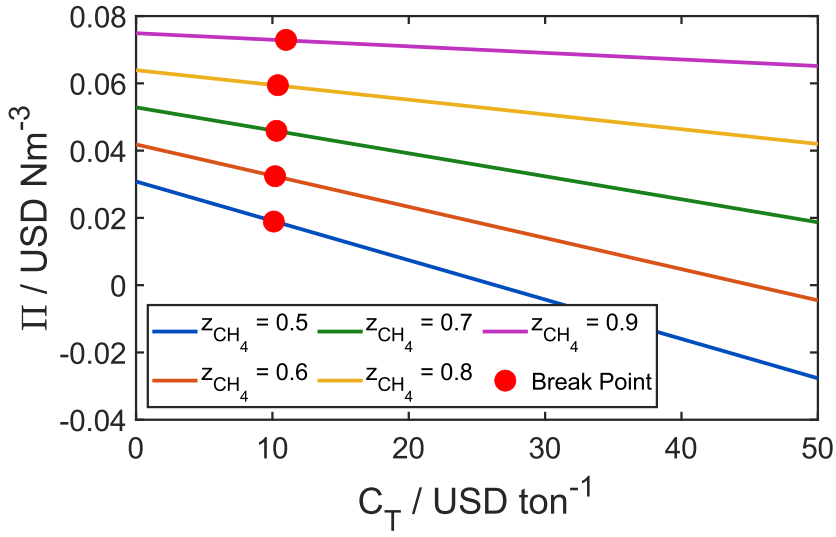


Figure 6.12. Dependence of the AMCR process profit Π as a function of the CO_2 tax value. The break point is the CO_2 tax value in which the AM and AMCR have the same economic performance. Solvent: 30 %wt. MEA with no degradation.

The breakpoint plot of the AMCR and the AM process is presented in Figure 6.12. One can realize that z_{CH_4} does not have a significant influence in the location of the breakpoints, which means that most of them are located when the CO_2 tax is around $10\ USD/ton$. Furthermore, the breakpoints are located at lower C_T values, which indicates that the extra costs related to the compression of the carbon dioxide stream are low compared to the tax that must be paid for CO_2 emissions.

The capital costs of the AM process will increase with larger amounts of methane in the biogas, because the biomethane compressor must become larger, and thus, the investment costs become larger as well. However, this is not the case for the AMCR process, where capital costs associated to the compressors will be reduced when z_{CH_4} is increased. This is caused by the fact that $P_B < P_C$, hence, the less amount of gas

that is needed to be pressurized to 150 bar, the smaller the compressors will be (the same effect is found in the PWCR process). We should remark that the AMCR process will always be more expensive than the AM process because of the extra compressor system needed.

The size of the absorber will not change significantly at different z_{CH_4} because its diameter is mainly a function of the gas surface velocity [151] and the total amount of biogas entering the absorber is always the same. Conversely, the larger the amount of CO_2 in the feed (low z_{CH_4}), the more gas the desorber must process and, hence, it will be larger and more expensive.

Since the amine-based processes are operated at mild temperatures and pressures, the equipment thickness is expected to be minimum when compared to the PWCR process. Unfortunately, due to the corrosive nature of aqueous amine mixtures, it is required to use stainless steel 304L or 316L (the latter is recommended in order to avoid failures [152]). Stainless steel 316L can be 3.4 times more expensive [130] and will increase significantly the capital costs of the plant.

6.5.3. Selection of the optimal technology

In the previous sections, it was shown that, at moderate CO_2 tax values, the PWCR and the AMCR processes seem to be economically more feasible that their counterparts without CO_2 recovery. For this reason, and because of environmental interests, this subsection will focus only on the processes that involve the CO_2 recovery.

If the profit values reported for the PWCR and the AMCR are compared, one can notice that the PWCR process is more profitable than the AMCR process when z_{CH_4} is between 0.5 and 0.8. Conversely, the chemical scrubbing process has better

performance at higher methane compositions. One of the main reasons on why the PWCR process has better performance than the AMCR process is because it uses a combined pressure and temperature swing to perform the separation. Contrarywise, the AMCR process relies only on a temperature swing, hence, the energy used to regenerate the solvent can be said to be lost because it does not give any added value to the final product (the biomethane stream at P_B). In contrast, the electrical energy spent in the PWCR process to increase the pressure not only serves to separate the raw biogas mixture but also gives an added value to the final product.

The profit of each one of the processes should change as a function of P_B (as shown in Figure 6.13). One can observe that the AMCR process profit increases at higher z_{CH_4} while it decreases at higher P_B. The reason for this effect is because $P_B < P_C$, which indicates that the larger the CO_2 concentration, the higher the electric consumption. The profit decreases with larger pressures because more electrical power is required to compress the biomethane stream. On the other hand, for the PWCR process, the profit (Π) seems to not be affected by P_B. This is because the absorber pressure P_A is larger than the delivery pressure of biomethane P_B, therefore, the last compressor is deemed unnecessary and, therefore, less capital cost investments are required.

If the PWCR and AMCR processes are compared, it is possible to notice that when P_B is equal to 0.1 MPa, there is a crossover in the profit lines, which means that when P_B is low, then it is preferable to use an AMCR process at larger feed methane compositions ($z_{CH_4} > 0.7$) values. Conversely, when the delivery pressure P_B is $1.2\ MPa$, the AMCR process loses competitivity due to the extra costs caused by the higher compression duties.

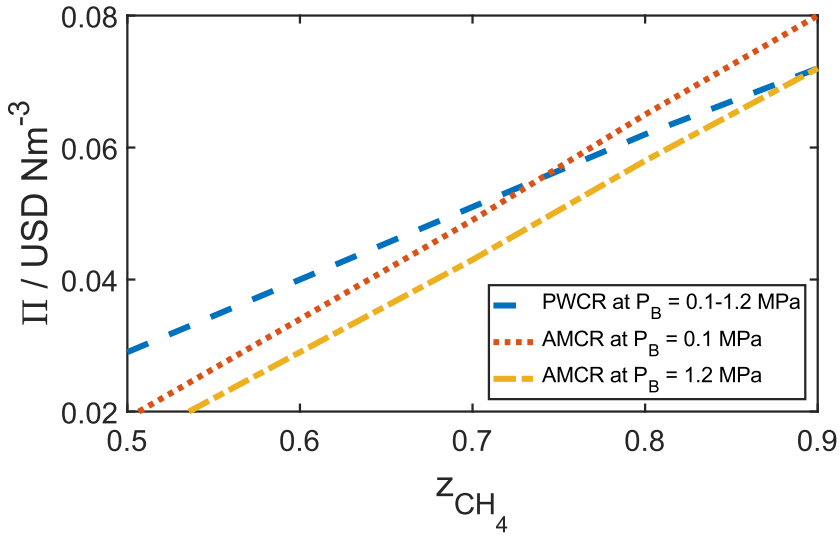


Figure 6.13. Effect of P_B and z_{CH_4} on the profit Π of the PWCR and AMCR processes using PEGDME and MEA as solvents.

In order to provide a guide on which technology to choose based on the methane concentration z_{CH_4} and the delivery pressure P_B, Figure 6.14 was generated. A technology is reported as being superior if the difference between the optimized profit between the AMCR and the PWCR process is larger than 5%. Since the degradation values in the AMCR process are somewhat uncertain, an average value of 1.5 kg of degraded MEA/ton CO_2 was used for the calculations in Figure 6.14. In the case where the AMCR and PWCR processes have very similar competitiveness (marked with green in Figure 6.14), further evaluations are required to determine the best option. These evaluations should consider that both processes have inherent disadvantages caused by their high operation pressures (PWCR) or corrosive solvents (AMCR) that might reflect on higher capital costs.

Figure 6.14 shows that the competitiveness of the PWCR processes increases at higher delivery pressures. This is because the PWCR process uses the pressure increase for two purposes: to perform the separation and to reach the delivery pressure P_{CH_4}. The AMCR process spends the same amount of calorific energy just to regenerate the solvent, while the electric energy always increases when P_B is augmented.

We want to remark that when the study case fixes $z_{CH_4} = 0.9$ and $P_B = 0.1\ MPa$, the biogas upgrading plant behaves similarly to CO_2 capture processes from flue gases (usually the CO_2 composition is 10–12% mol/mol in coal-fired power plants). At these conditions, the AMCR process seems to have better economic performance. This finding agrees with what has been widely discussed about CO_2 capture processes from flue gases: the amine-based separation processes are expected to outperform processes that require pressurization because of the compression work [24,153]. This is because the CO_2 capture systems have a “clean” stream that flows out from the absorber with a low concentration of CO_2 and a low pressure equal to 1 bar. For this reason, any kind of energy spent to modify the pressure or temperature of the gases that are not CO_2 can be considered to be unnecessary.

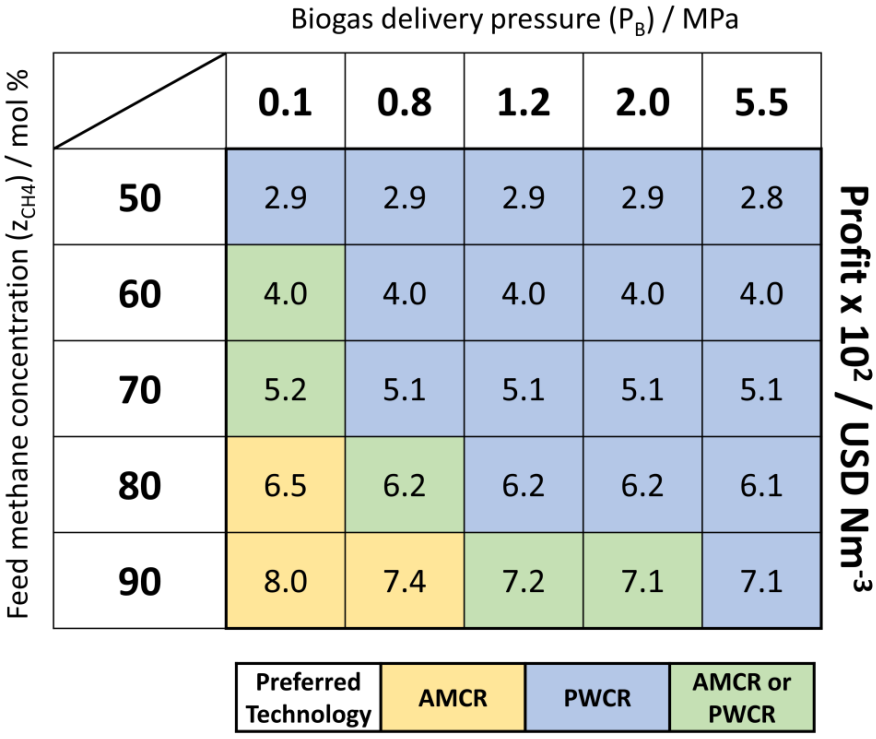


Figure 6.14. Technology selection for processes with CO_2 recovery.

6.6. Conclusions

The energetic and economic assessment of 5 physical solvents was performed. These solvents are PEGDME, NFM, NMP, methanol and water. It was concluded that all of them are capable of upgrading biogas with a CO_2 recovery system, delivering both biomethane and CO_2 within the required specifications. The performance of the solvents can thus be arranged in terms of economic, energetic and expected capital costs in the following order: PEGDME ≈ NMP > NFM > methanol > water.

The feasibility of implementing CO_2 recovery into liquid solvent-based processes was evaluated. It was found that it is necessary to consider the CO_2 tax into the assessment of the biogas upgrading technologies because, at relatively low CO_2 tax values (between 10 and 15 USD/ton CO_2), the processes with CO_2 recovery are economically more competitive than their counterparts without CO_2 recovery.

The optimal technology for biogas upgrading plants is a function of the methane concentration in the feed (z_{CH_4}) and the delivery pressure of the biogas (P_B). Processes using amine solvents are favored at lower z_{CH_4} and lower P_B values, while the physical solvents are favored at larger P_B. The physical absorption processes outperform the modelled chemical solvent process based on 30 %wt. MEA at moderate to high pressures because the physical absorption uses the compression duty to separate the raw biogas mixture, partly regenerate the solvent, and achieve the delivery pressure specification P_B. In this way, the energetic usage is optimized. On the other hand, the amine processes use the calorific energy to regenerate the solvent, hence all the energy spent on the CO_2 desorption is not used on any other part of the process. It should be remarked that increasing the absorber pressure in case of chemical absorption would lead to different results.

6.7. Supplementary information

The complementary component properties used for the pumping systems and the CO_2 compression systems in the simulation framework are shown in Table 6.12. Pure component properties used in this work were taken from Aspen's Plus databank v8.6. The gas heat capacity (CP_i^{IG}) equation used in this work is:

$$CP_i^{IG} = C_1 + C_2\left(\frac{C_3/T}{sinh(C_3/T)}\right) + C_4\left(\frac{C_5/T}{sinh(C_5/T)}\right) \quad (6.18)$$

This equation is one proposed by the Design Institute for Physical Properties (DIPPR) and the constants used are given in Table 6.13 and were taken from Aspen's Plus v8.6 databank.

The pure liquid vapor pressure of the components (P_i^{sat}) used in this work is given by the extended Antoine equation is:

$$\ln P_i^{sat} = C_1 + \frac{C_2}{T + C_3} + C_4 T + C_5 \ln T + C_6 T^{C_7} \quad (6.19)$$

The parameters can be found in Table 6.14 and were taken from Aspen's Plus v8.6 databank.

Table 6.12. Pure component properties used in this work.

Component	MW / g/mol	Critical temperature / K	Critical pressure / MPa	Critical volume / mol/cm^3	Critical compressibility factor	Acentric factor
CH_4	16.04	190.56	4.599	98.6	0.286	0.011
CO_2	44.01	304.21	7.383	94	0.274	0.224
H_2O	18.02	647.1	22.06	55.94	0.229	0.344
MEA	61.08	678.2	7.12	225	0.284	0.447
NMP	99.13	721.6	4.52	310	0.234	0.373
NFM	115.13	779.3	5.08	328	0.257	0.403
Sulfolane	120.17	853	5.030	300	0.213	0.382
PEGDME	266.33	769.8	1.67	821	0.214	1.139
Methanol	32.04	512.5	8.08	117	0.222	0.566

Table 6.13. Parameters for the CP^{IG} equation (kJ/kmol).

Component	C_1	C_2	C_3	C_4	C_5
CH_4	33.30	79.93	2086.90	41.60	991.96
CO_2	29.37	34.54	1428.00	26.40	588.00
H_2O	33.36	26.79	2610.50	8.90	1169.00
MEA	72.14	181.50	2030.00	131.40	860.00
NMP	73.44	346.70	1874.50	249.10	821.50
NFM	108.67	227.28	783.07	140.37	3174.20
Sulfolane	115.56	202.29	968.69	-1130.20	28.10
PEGDME	231.42	786.85	1628.10	565.85	702.04
Methanol	39.25	87.90	1916.50	53.65	896.70

Table 6.14. Parameters for the P_i^{sat} equation (kPa).

Component	C_1	C_2	C_3	C_4	C_5	C_6	C_7
H_2O	66.74	-7258.2	0	0	-7.3037	4.17E-06	2
MEA	85.72	-10367	0	0	-9.4699	1.90E-18	6
NMP	61.57	-8467.9	0	0	-6.3622	3.22E-18	6
NFM	78.27	-10241	0	0	-8.6143	3.39E-18	6
Sulfolane	143.75	-13283	0	0	-19.429	0.013441	1
PEGDME	133.22	-17794	0	0	-15.547	3.09E-18	6
Methanol	75.81	-6904.5	0	0	-8.8622	7.47E-06	2

Table of symbols

Symbol	Unit	Meaning
A, B		Henry's coefficient parameters
b	mol/kg solvent	Molality
COP		Coefficient of performance
CR		Compression ratio
C_T	USD/Nm^3	Carbon tax
E	MJ/Nm^3	Electric energy
F	kmol/s	Molar flow rate
H	MPa	Henry's coefficient
h	kJ/kg	Enthalpy
K	%	Share of the costs associated to the carbon dioxide stream conditioning
L/G		Liquid-to-gas ratio
M	kg/kmol	Molecular weight
m	kg/s	Mass flow rate
n		Number of compression stages
P	kPa	Pressure
Q	MJ/Nm^3	Calorific energy
r_P		Depressurization ratio
T	K	Temperature
x		Mol fraction in CO_2 stream

Symbol	Unit	Meaning
W	kW	Compressor work
y		Mol fraction in biomethane stream
z		Mol fraction in raw biogas stream
Greek symbols	**Units**	**Meaning**
α	mol CO_2/mol amine	Amine loading
ΔH	kJ/mol CO_2	Heat of absorption
$\eta_{K,S}$		Isentropic efficiency of compressor
$\eta_{K,M}$		Mechanical efficiency of compressor
η_P		Mechanical efficiency of pump
θ		Methane slip
Π	USD/Nm^3	Profit
ρ	kg/m^3	Density
Subscripts	**Meaning**	
A	Referent to absorber	
B	Referent to biomethane	
$Burnt$	Referent to the biogas that has to be burnt	
C	Referent to the carbon dioxide stream	
CH_4	Referent to methane	
$Chill$	Referent to chilling system	
D	Referent to desorber	
G	Referent to gas entering absorber	

Subscripts (Cont.)	Meaning
i	Referent to component *i*
L	Referent to liquid entering absorber
$Lean$	Referent to lean loading
R	Referent to reboiler
ref	Reference value
$Rich$	Referent to rich loading
S	Referent to solvent

6.8. References

[1] Urban W. Biomethane injection into natural gas networks. In: Wellinger A, Murphy J, Baxter D, editors. Biogas Handb., Woodhead Publishing Limited; 2013, p. 378–403.

[2] Yang F, Meerman JC, Faaij APC. Carbon capture and biomass in industry: A techno-economic analysis and comparison of negative emission options. Renew Sustain Energy Rev 2021;144:111028. https://doi.org/10.1016/j.rser.2021.111028.

[3] Ryckebosch E, Drouillon M, Vervaeren H. Techniques for transformation of biogas to biomethane. Biomass and Bioenergy 2011;35:1633–45. https://doi.org/10.1016/J.BIOMBIOE.2011.02.033.

[4] Petersson A, Wellinger A. Biogas upgrading technologies - developments and innovations. 2009.

[5] Petersson A. Biogas cleaning. In: Wellinger A, Murphy J, Baxter D, editors. Biogas Handb., Woodhead Publishing Limited; 2013, p. 329–41.

[6] Persson M. Evaluation of upgrading techniques for biogas - Rapport SGC 142. 2003.

[7] Weiland P. Biogas production: Current state and perspectives. Appl Microbiol Biotechnol 2010;85:849–60. https://doi.org/10.1007/s00253-009-2246-7.

[8] Bauer F, Hulteberg C, Persson T, Tamm D. Biogas upgrading - Review of commercial technologies - SGC Rapport 270. 2013.

[9] IEA Bioenergy - Task 37 - Plant Lists 2019. http://task37.ieabioenergy.com/plant-list.html (accessed May 20, 2020).

[10] Nogueira GP, McManus MC, Leak DJ, Franco TT, Oliveira de Souza Dias M, Nakao Cavaliero CK. Are eucalyptus harvest residues a truly burden-free biomass source for bioenergy? A deeper look into biorefinery process design and Life Cycle Assessment. J Clean Prod 2021;299:126956. https://doi.org/10.1016/j.jclepro.2021.126956.

[11] Halliday C, Hatton TA. Net-Negative Emissions through Molten Sorbents and Bioenergy with Carbon Capture and Storage. Ind Eng Chem Res 2020;59:22582–96. https://doi.org/10.1021/acs.iecr.0c04512.

[12] Bauer F, Persson T, Hulteberg C, Tamm D. Biogas upgrading - technology overview, comparison and perspectives for the future. Biofuels, Bioprod Biorefining 2013;7:499–511. https://doi.org/10.1002/bbb.1423.

[13] Beil M, Beyrich W. Biogas upgrading to biomethane. In: Wellinger A, Murphy J, Baxter D, editors. Biogas Handb., Woodhead Publishing Limited; 2013, p. 342–77.

[14] Niesner J, Jecha D, Stehlík P. Biogas Upgrading Technologies: State of Art Review in European Region. Chem Eng Transac 2013;35:517–22. https://doi.org/10.3303/CET1335086.

[15] Hochgesand G. Rectisol and Purisol. Ind Eng Chem 1970;62:37–43. https://doi.org/10.1021/ie50727a007.

[16] Gamba S, Pellegrini LA. Biogas Upgrading: Analysis and Comparison between Water and Chemical Scrubbings. Chem Eng Trans 2013;32:1273–8.

[17] Riboldi L, Bolland O. Evaluating Pressure Swing Adsorption as a CO2 separation technique in coal-fired power plants. Int J Greenh Gas Control 2015;39:1–16. https://doi.org/10.1016/j.ijggc.2015.02.001.

[18] Zhang Y, Sunarso J, Liu S, Wang R. Current status and development of membranes for CO2/CH4 separation: A review. Int J Greenh Gas Control 2013;12:84–107. https://doi.org/10.1016/j.ijggc.2012.10.009.

[19] Vrbová V, Ciahotný K. Upgrading Biogas to Biomethane Using Membrane Separation. Energy and Fuels 2017;31:9393–401. https://doi.org/10.1021/acs.energyfuels.7b00120.

[20] Baena-Moreno FM, Rodríguez-Galán M, Vega F, Vilches LF, Navarrete B, Zhang Z. Biogas upgrading by cryogenic techniques. Environ Chem Lett 2019;17:1251–61. https://doi.org/10.1007/s10311-019-00872-2.

[21] Criscuoli A, Drioli E. New metrics for evaluating the performance of membrane operations in the logic of process intensification. Ind. Eng. Chem. Res., vol. 46, American Chemical Society ; 2007, p. 2268–71. https://doi.org/10.1021/ie0610952.

[22] Brunetti A, Scura F, Barbieri G, Drioli E. Membrane technologies for CO2 separation. J Memb Sci 2010;359:115–25. https://doi.org/10.1016/j.memsci.2009.11.040.

[23] Maurya R, Tirkey SR, Rajapitamahuni S, Ghosh A, Mishra S. Recent advances and future prospective of biogas production. Adv. Feed. Convers. Technol. Altern. Fuels Bioprod. New Technol. Challenges Oppor., Elsevier; 2019, p. 159–78. https://doi.org/10.1016/B978-0-12-817937-6.00009-6.

[24] Rochelle GT. Amine scrubbing for CO2 capture. Science (80-) 2009;325:1652–4. https://doi.org/10.1126/science.1176731.

[25] Feron PHM. Absorption-Based Post-combustion Capture of Carbon Dioxide. Elsevier Inc.; 2016. https://doi.org/10.1016/c2014-0-03382-5.

[26] Kapoor R, Ghosh P, Kumar M, Vijay VK. Evaluation of biogas upgrading technologies and future perspectives: a review. Environ Sci Pollut Res 2019;26:11631–61. https://doi.org/10.1007/s11356-019-04767-1.

[27] Maggioni L, Pieroni C. Deliverable D5.2: Report on the biomethane injection into national gas grid. 2016.

[28] Prussi M, Julea A, Lonza L, Thiel C. Biomethane as alternative fuel for the EU road sector: analysis of existing and planned infrastructure. Energy Strateg Rev 2021;33:100612. https://doi.org/10.1016/j.esr.2020.100612.

[29] GIIGNL. Position paper on the impact of including methane number in natural gas regulation n.d. https://giignl.org/publications/position-paper-impact-including-methane-number-natural-gas-regulation (accessed April 27, 2021).

[30] Gosiewski K, Pawlaczyk A. Catalytic or thermal reversed flow combustion of coal mine ventilation air methane: What is better choice and when? Chem Eng J 2014;238:78–85. https://doi.org/10.1016/j.cej.2013.07.039.

[31] Moscato I, Munoz DC, Italiano P. Sustainable biomethane: methane slip removal applying regenerative catalytic oxidation (RCO) post combustion technology. Environ Eng Manag J 2020;19:1831–3. https://doi.org/10.30638/eemj.2020.175.

[32] TUV. Biogas to Biomethane Technology Review. 2012.

[33] Lorenzi G, Gorgoroni M, Silva C, Santarelli M. Life cycle assessment of biogas upgrading routes. Energy Procedia, vol. 158, Elsevier Ltd; 2019, p. 2012–8. https://doi.org/10.1016/j.egypro.2019.01.466.

[34] Starr K, Gabarrell X, Villalba G, Talens L, Lombardi L. Life cycle assessment of biogas upgrading technologies. Waste Manag 2012;32:991–9. https://doi.org/10.1016/j.wasman.2011.12.016.

[35] Hyatt JA. Liquid and Supercritical Carbon Dioxide as Organic Solvents. J Org Chem 1984;49:5097–101. https://doi.org/10.1021/jo00200a016.

[36] Raveendran P, Wallen SL. Cooperative C-H0 ⋯ O hydrogen bonding in CO2-Lewis base complexes: Implications for solvation in supercritical CO2. J Am Chem Soc 2002;124:12590–9. https://doi.org/10.1021/ja0174635.

[37] Tande B, Seames W, Benson S. Efficient Regeneration of Physical and Chemical Solvents for CO2 Capture. 2013.

[38] Kohl AL, Nielsen RB, Kohl AL, Nielsen RB. Physical Solvents for Acid Gas Removal. Gas Purif 1997:1187–237. https://doi.org/10.1016/B978-088415220-0/50014-8.

[39] Gui X, Tang Z, Fei W. Solubility of CO 2 in Alcohols, Glycols, Ethers, and Ketones at High Pressures from (288.15 to 318.15) K. J Chem Eng Data 2011;56:2420–9. https://doi.org/10.1021/je101344v.

[40] Li X, Jiang Y, Han G, Deng D. Investigation of the Solubilities of Carbon Dioxide in Some Low Volatile Solvents and Their Thermodynamic Properties. J Chem Eng Data 2016;61:1254–61. https://doi.org/10.1021/acs.jced.5b00893.

[41] Aghaie M, Rezaei N, Zendehboudi S. A systematic review on CO2 capture with ionic liquids: Current status and future prospects. Renew Sustain Energy Rev 2018;96:502–25. https://doi.org/10.1016/j.rser.2018.07.004.

[42] Enick RM, Koronaios P, Stevenson C, Warman S, Morsi B, Nulwala H, et al. Hydrophobic polymeric solvents for the selective absorption of CO 2 from warm gas streams that also contain H2 and H 2O. Energy and Fuels 2013;27:6913–20. https://doi.org/10.1021/ef401740w.

[43] Siefert NS, Agarwal S, Shi F, Shi W, Roth EA, Hopkinson D, et al. Hydrophobic physical solvents for pre-combustion CO2 capture: Experiments, computational simulations, and techno-economic analysis. Int J Greenh Gas Control 2016;49:364–71. https://doi.org/10.1016/j.ijggc.2016.03.014.

[44] Vega F, Cano M, Camino S, Fernández LMG, Portillo E, Navarrete B. Solvents for Carbon Dioxide Capture. Carbon Dioxide Chem. Capture Oil Recover., InTech; 2018. https://doi.org/10.5772/intechopen.71443.

[45] Rackley SA. Absorption capture systems. Carbon Capture and Storage, Elsevier; 2017, p. 115–49. https://doi.org/10.1016/b978-0-12-812041-5.00006-4.

[46] Higman C. Gasification processes and synthesis gas treatment technologies for carbon dioxide (CO2) capture. Dev. Innov. Carbon Dioxide Capture Storage Technol., Elsevier; 2010, p. 243–79. https://doi.org/10.1533/9781845699574.2.243.

[47] Burr B, Lyddon L. A comparison of physical solvents for acid gas removal. Digit Refin 2008. https://www.digitalrefining.com/article/1000560#.XqE2B8gzZPY (accessed April 23, 2020).

[48] Nagaraju Palla, Dennis Leppin. Technical and Operating Support for Pilot Demonstration of Morphysorb Acid Gas Removal Process. Pittsburgh, PA, and Morgantown, WV: 2003. https://doi.org/10.2172/822134.

[49] Boll W, Hochgesand G, Higman C, Supp E, Kalteier P, Müller W-D, et al. Gas Production, 3. Gas Treating. Ullmann's Encycl. Ind. Chem., Weinheim, Germany:

Wiley-VCH Verlag GmbH & Co. KGaA; 2011. https://doi.org/10.1002/14356007.o12_o02.

[50] Bucklin RW, Schendel RL. Comparison of Physical Solvents Used for Gas Processing. In: Newman SA, editor. Acid Sour Gas Treat. Process., 1985, p. 42–79.

[51] Bhattacharyya D, Turton R, Zitney SE. Acid gas removal from syngas in IGCC plants. Integr. Gasif. Comb. Cycle Technol., Elsevier Inc.; 2017, p. 385–418. https://doi.org/10.1016/B978-0-08-100167-7.00011-1.

[52] Bell DA, Towler BF, Fan M. Sulfur Recovery. Coal Gasif. Its Appl., William Andrew Publishing; 2010, p. 113–36. https://doi.org/10.1016/B978-0-8155-2049-8.10006-3.

[53] Rayer A V., Henni A, Tontiwachwuthikul P. High pressure physical solubility of carbon dioxide (CO2) in mixed polyethylene glycol dimethyl ethers (Genosorb 1753). Can J Chem Eng 2012;90:576–83. https://doi.org/10.1002/cjce.20615.

[54] Ordorica-Garcia JG, Elkamel A, Douglas PL, Croiset E. Clean-Coal Technology: Gasification Pathway. Environ. Conscious Foss. Energy Prod., Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2009, p. 243–76. https://doi.org/10.1002/9780470432747.ch7.

[55] Bucklin RW, Schendel RL. Comparison of Fluor Solvent and Selexol Processes. Energy Prog 1984;4:137–42.

[56] Isaacs EE, Otto FD, Mather AE. Solubility of hydrogen sulfide and carbon dioxide in a Sulfinol solution. J Chem Eng Data 1977;22:317–9. https://doi.org/10.1021/je60074a024.

[57] Ghanbarabadi H, Khoshandam B, Wood DA. Simulation of CO2 removal from ethane with Sulfinol-M+AMP solvent instead of DEA solvent in the South Pars phases 9 and 10 gas processing facility. Petroleum 2019;5:90–101. https://doi.org/10.1016/j.petlm.2018.06.004.

[58] Yih S-M, Lai H-C. Simultaneous absorption of CO2 and H2S in sulfinol solutions in a set of packed absorber-stripper. Chem Eng Commun 1990;88:119–25. https://doi.org/10.1080/00986449008940551.

[59] Décultot M, Ledoux A, Fournier-Salaün MC, Estel L. Solubility of CO2 in methanol, ethanol, 1,2-propanediol and glycerol from 283.15 K to 373.15 K and up to 6.0 MPa. J Chem Thermodyn 2019;138:67–77. https://doi.org/10.1016/j.jct.2019.05.003.

[60] Lyu Z, Ma H, Zhang H, Ying W. Solubility of carbon dioxide in methanol from 213.15 K to 273.15 K: Measurement and modeling. Fluid Phase Equilib 2018;471:40–54. https://doi.org/10.1016/j.fluid.2018.04.014.

[61] Xia J, Jödecke M, Kamps APS, Maurer G. Solubility of CO2 in (CH3OH + H2O). J Chem Eng Data 2004;49:1756–9. https://doi.org/10.1021/je049803i.

[62] Jou FY, Deshmukh RD, Otto FD, Mather AE. Solubility of H2S, CO2 and CH4 in N-formyl morpholine. J Chem Soc Faraday Trans 1 Phys Chem Condens Phases 1989;85:2675–82. https://doi.org/10.1039/F19898502675.

[63] Xu Y, Schutte RP, Hepler LG. Solubilities of carbon dioxide, hydrogen sulfide and sulfur dioxide in physical solvents. Can J Chem Eng 1992;70:569–73. https://doi.org/10.1002/cjce.5450700321.

[64] Murrieta-Guevara F, Romero-Martinez A, Trejo A. Solubilities of carbon dioxide and hydrogen sulfide in propylene carbonate, N-methylpyrrolidone and

sulfolane. Fluid Phase Equilib 1988;44:105–15. https://doi.org/10.1016/0378-3812(88)80106-7.

[65] Murrieta-Guevara F, Trejo Rodriguez A. Solubility of carbon dioxide, hydrogen sulfide, and methane in pure and mixed solvents. J Chem Eng Data 1984;29:456–60. https://doi.org/10.1021/je00038a027.

[66] Bohloul MR, Vatani A, Peyghambarzadeh SM. Experimental and theoretical study of CO2 solubility in N-methyl-2-pyrrolidone (NMP). Fluid Phase Equilib 2014;365:106–11. https://doi.org/10.1016/j.fluid.2013.12.019.

[67] Melzer WM, Schrödter F, Knapp H. Solubilities of methane, propane and carbon dioxide in solvent mixtures consisting of water, N,N-dimethylformamide, and N-methyl-2-pyrrolidone. Fluid Phase Equilib 1989;49:167–86. https://doi.org/10.1016/0378-3812(89)80014-7.

[68] Li Y, Liu Q, Huang W, Yang J. Below the room temperature measurements of solubilities in ester absorbents for CO2 capture. J Chem Thermodyn 2018;127:71–9. https://doi.org/10.1016/j.jct.2018.07.021.

[69] Thompson R, Culp JT, Tiwari SP, Shi W, Siefert N, Hopkinson DP. CO2 Solubility in Organophosphate Physical Solvents Wherein Alkyl Groups Are Replaced with Poly(Ethylene Glycol) Groups 2019. https://doi.org/10.26434/CHEMRXIV.11348051.V1.

[70] Jou FY, Deshmukh RD, Otto FD, Mather AE. Solubility of H2S, CO2, CH4 and C2H6 in sulfolane at elevated pressures. Fluid Phase Equilib 1990;56:313–24. https://doi.org/10.1016/0378-3812(90)85111-M.

[71] Schnabel T, Vrabec J, Hasse H. Molecular simulation study of hydrogen bonding mixtures and new molecular models for mono- and dimethylamine. Fluid Phase Equilib 2008;263:144–59. https://doi.org/10.1016/j.fluid.2007.10.003.

[72] Kapateh MH, Chapoy A, Burgass R, Tohidi B. Experimental Measurement and Modeling of the Solubility of Methane in Methanol and Ethanol. J Chem Eng Data 2016;61:666–73. https://doi.org/10.1021/acs.jced.5b00793.

[73] Frost M, Karakatsani E, Von Solms N, Richon D, Kontogeorgis GM. Vapor-liquid equilibrium of methane with water and methanol. Measurements and modeling. J Chem Eng Data 2014;59:961–7. https://doi.org/10.1021/je400684k.

[74] Henni A, Tontiwachwuthikul P, Chakma A. Solubility Study of Methane and Ethane in Promising Physical Solvents for Natural Gas Sweetening Operations 2006. https://doi.org/10.1021/je050172h.

[75] Jou F-Y, Mather AE, Schmidt KAG. Solubility of Methane in Propylene Carbonate. J Chem Eng Data 2015;60:1010–3. https://doi.org/10.1021/je500849m.

[76] Rayer A V., Henni A, Tontiwachwuthikul P. High-pressure solubility of methane (CH 4) and ethane (C 2H 6) in mixed polyethylene glycol dimethyl ethers (Genosorb 1753) and its selectivity in natural gas sweetening operations. J Chem Eng Data 2012;57:764–75. https://doi.org/10.1021/je200905z.

[77] Rajasingam R, Lioe L, Pham QT, Lucien FP. Solubility of carbon dioxide in dimethylsulfoxide and N-methyl-2-pyrrolidone at elevated pressure. J Supercrit Fluids 2004;31:227–34. https://doi.org/10.1016/j.supflu.2003.12.003.

[78] Chen H, Li HQ, Tu WX, Bao WJ, Li SP. Measurement of CO2 solubility in tributyl phosphate. Guocheng Gongcheng Xuebao/The Chinese J Process Eng 2012;12:206–11.

[79] Shi W, Thompson RL, Macala MK, Resnik K, Steckel JA, Siefert NS, et al. Molecular Simulations of CO2 and H2 Solubility, CO2 Diffusivity, and Solvent Viscosity at 298 K for 27 Commercially Available Physical Solvents. J Chem Eng Data 2019;64:3682–92. https://doi.org/10.1021/acs.jced.8b01228.

[80] Svensson H, Zejnullahu Velasco V, Hulteberg C, Karlsson HT. Heat of absorption of carbon dioxide in mixtures of 2-amino-2-methyl-1-propanol and organic solvents. Int J Greenh Gas Control 2014;30:1–8. https://doi.org/10.1016/J.IJGGC.2014.08.022.

[81] Wanderley RR, Evjen S, Pinto DDD, Knuutila HK. The Salting-out Effect in Some Physical Absorbents for CO2 Capture. Chem. Eng. Trans., vol. 69, 2018, p. 97–102. https://doi.org/10.3303/CET1869017.

[82] Dodds WS, Stutzman LF, Sollami BJ. Carbon Dioxide Solubility in Water. Ind Eng Chem - Chem Eng Data Ser 1956;1:92–5. https://doi.org/10.1021/i460001a018.

[83] Diamond LW, Akinfiev NN. Solubility of CO2 in water from -1.5 to 100°C and from 0.1 to 100 MPa: Evaluation of literature data and thermodynamic modelling. Fluid Phase Equilib 2003;208:265–90. https://doi.org/10.1016/S0378-3812(03)00041-4.

[84] Lucile F, Cézac P, Contamine F, Serin JP, Houssin D, Arpentinier P. Solubility of carbon dioxide in water and aqueous solution containing sodium hydroxide at temperatures from (293.15 to 393.15) K and pressure up to 5 MPa: Experimental measurements. J Chem Eng Data 2012;57:784–9. https://doi.org/10.1021/je200991x.

[85] Chapoy A, Mohammadi AH, Chareton A, Tohidi B, Richon D. Measurement and Modeling of Gas Solubility and Literature Review of the Properties for the Carbon Dioxide-Water System. Ind Eng Chem Res 2004;43:1794–802. https://doi.org/10.1021/ie034232t.

[86] Dalmolin I, Skovroinski E, Biasi A, Corazza ML, Dariva C, Oliveira JV. Solubility of carbon dioxide in binary and ternary mixtures with ethanol and water. Fluid Phase Equilib 2006;245:193–200. https://doi.org/10.1016/j.fluid.2006.04.017.

[87] Valtz A, Chapoy A, Coquelet C, Paricaud P, Richon D. Vapour-liquid equilibria in the carbon dioxide-water system, measurement and modelling from 278.2 to 318.2 K. Fluid Phase Equilib 2004;226:333–44. https://doi.org/10.1016/j.fluid.2004.10.013.

[88] Servio P, Englezos P. Measurement of dissolved methane in water in equilibrium with its hydrate. J Chem Eng Data 2002;47:87–90. https://doi.org/10.1021/je0102255.

[89] Culberson OL, McKetta JJ. Phase Equilibria in Hydrocarbon-Water Systems III - The Solubility of Methane in Water at Pressures to 10,000 PSIA. J Pet Technol 1951;3:223–6. https://doi.org/10.2118/951223-g.

[90] Clever HL, Young CL, Battino R, Hayduk W, Wiesenburg DA. Methane. Pergamon; 1987.

[91] Lekvam K, Raj Bishnoi P. Dissolution of methane in water at low temperatures and intermediate pressures. Fluid Phase Equilib 1997;131:297–309. https://doi.org/10.1016/s0378-3812(96)03229-3.

[92] Abatzoglou N, Boivin S. A review of biogas purification processes. Biofuels, Bioprod Biorefining 2009;3:42–71. https://doi.org/10.1002/bbb.117.

[93] Rochelle GT. Conventional amine scrubbing for CO2 capture. Absorption-Based Post-Combustion Capture of Carbon Dioxide, Elsevier Inc.; 2016, p. 35–67. https://doi.org/10.1016/B978-0-08-100514-9.00003-2.

[94] Kohl AL, Nielsen RB, Kohl AL, Nielsen RB. Alkanolamines for Hydrogen Sulfide and Carbon Dioxide Removal. Gas Purif 1997:40–186. https://doi.org/10.1016/B978-088415220-0/50002-1.

[95] Yuan Y, Rochelle GT. Lost work: A comparison of water-lean solvent to a second generation aqueous amine process for CO 2 capture. Int J Greenh Gas Control 2019:82–90. https://doi.org/10.1016/j.ijggc.2019.03.013.

[96] Vega F, Sanna A, Navarrete B, Maroto-Valer MM, Cortés VJ. Degradation of amine-based solvents in CO2 capture process by chemical absorption. Greenh Gases Sci Technol 2014;4:707–33. https://doi.org/10.1002/ghg.1446.

[97] Krzemień A, Wieckol-Ryk A, Smoliński A, Koteras A, Wiecław-Solny L. Assessing the risk of corrosion in amine-based CO2 capture process. J Loss Prev Process Ind 2016;43:189–97. https://doi.org/10.1016/j.jlp.2016.05.020.

[98] Karl M, Castell N, Simpson D, Solberg S, Starrfelt J, Svendby T, et al. Uncertainties in assessing the environmental impact of amine emissions from a CO2 capture plant. Atmos Chem Phys 2014;14:8533–57. https://doi.org/10.5194/acp-14-8533-2014.

[99] Azzi M, White S. Emissions from amine-based post-combustion CO2 capture plants. Absorption-Based Post-Combustion Capture of Carbon Dioxide, Elsevier Inc.; 2016, p. 488–504. https://doi.org/10.1016/B978-0-08-100514-9.00020-2.

[100] Gjernes E, Helgesen LI, Maree Y. Health and environmental impact of amine based post combustion CO 2 capture. Energy Procedia, vol. 37, Elsevier Ltd; 2013, p. 735–42. https://doi.org/10.1016/j.egypro.2013.05.162.

[101] Lepaumier H, Da Silva EF, Einbu A, Grimstvedt A, Knudsen JN, Zahlsen K, et al. Comparison of MEA degradation in pilot-scale with lab-scale experiments.

Energy Procedia, vol. 4, Elsevier Ltd; 2011, p. 1652–9. https://doi.org/10.1016/j.egypro.2011.02.037.

[102] Bernhardsen IM, Knuutila HK. A review of potential amine solvents for CO2 absorption process: Absorption capacity, cyclic capacity and pKa. Int J Greenh Gas Control 2017;61:27–48. https://doi.org/10.1016/J.IJGGC.2017.03.021.

[103] Conway W, Bruggink S, Beyad Y, Luo W, Melián-Cabrera I, Puxty G, et al. CO2 absorption into aqueous amine blended solutions containing monoethanolamine (MEA), N,N-dimethylethanolamine (DMEA), N,N-diethylethanolamine (DEEA) and 2-amino-2-methyl-1-propanol (AMP) for post-combustion capture processes. Chem Eng Sci 2015;126:446–54. https://doi.org/10.1016/j.ces.2014.12.053.

[104] Conway W, Beyad Y, Feron P, Richner G, Puxty G. CO2 absorption into aqueous amine blends containing benzylamine (BZA), monoethanolamine (MEA), and sterically hindered / tertiary amines. Energy Procedia, vol. 63, Elsevier Ltd; 2014, p. 1835–41. https://doi.org/10.1016/j.egypro.2014.11.191.

[105] Wanderley RR, Pinto DDD, Knuutila HK. From hybrid solvents to water-lean solvents – A critical and historical review. Sep Purif Technol 2021;260:118193. https://doi.org/10.1016/j.seppur.2020.118193.

[106] Sanchez-Fernandez E, Heffernan K, Van Der Ham L, Linders MJG, Goetheer ELV, Vlugt TJH. Precipitating amino acid solvents for CO2 capture. Opportunities to reduce costs in Post combustion capture. Energy Procedia, vol. 63, Elsevier Ltd; 2014, p. 727–38. https://doi.org/10.1016/j.egypro.2014.11.080.

[107] Moioli S, Pellegrini LA, Ho MT, Wiley DE. A comparison between amino acid based solvent and traditional amine solvent processes for CO2 removal. Chem Eng Res Des 2019;146:509–17. https://doi.org/10.1016/j.cherd.2019.04.035.

[108] Pinto DDD, Knuutila H, Fytianos G, Haugen G, Mejdell T, Svendsen HF. CO2 post combustion capture with a phase change solvent. Pilot plant campaign. Int J Greenh Gas Control 2014;31:153–64. https://doi.org/10.1016/j.ijggc.2014.10.007.

[109] Sanchez Fernandez E, Heffernan K, Van Der Ham L V., Linders MJG, Eggink E, Schrama FNH, et al. Conceptual design of a novel CO2 capture process based on precipitating amino acid solvents. Ind Eng Chem Res 2013;52:12223–35. https://doi.org/10.1021/ie401228r.

[110] Rochelle G, Chen E, Freeman S, Van Wagener D, Xu Q, Voice A. Aqueous piperazine as the new standard for CO2 capture technology. Chem Eng J 2011;171:725–33.

[111] Feron PHM, Cousins A, Jiang K, Zhai R, Garcia M. An update of the benchmark post-combustion CO2-capture technology. Fuel 2020;273:117776. https://doi.org/10.1016/j.fuel.2020.117776.

[112] Lee WY, Park SY, Lee KB, Nam SC. Simultaneous Removal of CO2 and H2S from Biogas by Blending Amine Absorbents: A Performance Comparison Study. Energy and Fuels 2020;34:1992–2000. https://doi.org/10.1021/acs.energyfuels.9b03342.

[113] Capra F, Fettarappa F, Magli F, Gatti M, Martelli E. Biogas upgrading by amine scrubbing: Solvent comparison between MDEA and MDEA/MEA blend. Energy Procedia, vol. 148, Elsevier Ltd; 2018, p. 970–7. https://doi.org/10.1016/j.egypro.2018.08.065.

[114] Privalova EI, Mäki-Arvela P, Eränen K, Avetisov AK, Mikkola J-P, Murzin DY. Amine Solutions for Biogas Upgrading: Ideal versus Non-Ideal Absorption Isotherms. Chem Eng Technol 2013;36:740–8. https://doi.org/10.1002/ceat.201200161.

[115] Costa C, Demartini M, Di felice R, Oliva M, Pagliai P. Piperazine and methyldiethanolamine interrelationships in CO2 absorption by aqueous amine mixtures. Part I: Saturation rates of single-reagent solutions. Can J Chem Eng 2019;97:1160–71. https://doi.org/10.1002/cjce.23320.

[116] Abdeen FRH, Mel M, Jami MS, Ihsan SI, Ismail AF. A review of chemical absorption of carbon dioxide for biogas upgrading. Chinese J Chem Eng 2016;24:693–702. https://doi.org/10.1016/j.cjche.2016.05.006.

[117] Closmann F, Nguyen T, Rochelle GT. MDEA/Piperazine as a solvent for CO2 capture. Energy Procedia, vol. 1, Elsevier; 2009, p. 1351–7. https://doi.org/10.1016/j.egypro.2009.01.177.

[118] Kim I, Svendsen HF. Heat of absorption of carbon dioxide (CO2) in monoethanolamine (MEA) and 2-(aminoethyl)ethanolamine (AEEA) solutions. Ind Eng Chem Res 2007;46:5803–9. https://doi.org/10.1021/ie0616489.

[119] Svendsen HF, Hessen ET, Mejdell T. Carbon dioxide capture by absorption, challenges and possibilities. Chem Eng J 2011;171:718–24. https://doi.org/10.1016/j.cej.2011.01.014.

[120] Kim I, Hoff KA, Mejdell T. Heat of absorption of CO2 with aqueous solutions of MEA: New experimental data. Energy Procedia 2014;63:1446–55. https://doi.org/10.1016/j.egypro.2014.11.154.

[121] Aronu UE, Gondal S, Hessen ET, Haug-Warberg T, Hartono A, Hoff KA, et al. Solubility of CO2 in 15, 30, 45 and 60 mass% MEA from 40 to 120 °C and model representation using the extended UNIQUAC framework. Chem Eng Sci 2011;66:6393–406.

[122] Carranza-Abaid A, Svendsen HF, Jakobsen JP. Surrogate modelling of VLE: Integrating machine learning with thermodynamic constraints. Chem Eng Sci X 2020;8:100080. https://doi.org/10.1016/j.cesx.2020.100080.

[123] Gabrielsen J, Michelsen ML, Stenby EH, Kontogeorgis GM. A model for estimating CO2 solubility in aqueous alkanolamines. Ind Eng Chem Res 2005;44:3348–54. https://doi.org/10.1021/ie048857i.

[124] Moioli S, Pellegrini LA. Modeling the methyldiethanolamine-piperazine scrubbing system for CO2 removal: Thermodynamic analysis. Front Chem Sci Eng 2016;10:162–75. https://doi.org/10.1007/s11705-016-1555-5.

[125] Øi LE, Brathen T, Berg C, Brekne SK, Flatin M, Johnsen R, et al. Optimization of configurations for amine based CO2 absorption using Aspen HYSYS. Energy Procedia, vol. 51, Elsevier Ltd; 2014, p. 224–33. https://doi.org/10.1016/j.egypro.2014.07.026.

[126] Beil M, Beyrich W. Biogas upgrading to biomethane. 2013. https://doi.org/10.1533/9780857097415.3.342.

[127] Budzianowski WM, Wylock CE, Marciniak PA. Power requirements of biogas upgrading by water scrubbing and biomethane compression: Comparative analysis of various plant configurations. Energy Convers Manag 2017;141:2–19. https://doi.org/10.1016/j.enconman.2016.03.018.

[128] Masson-Delmotte V, Pörtner H-O, Skea J. IPCC report Global Warming of 1.5 □C. 2018.

[129] Lombardi L, Francini G. Techno-economic and environmental assessment of the main biogas upgrading technologies. Renew Energy 2020;156:440–58. https://doi.org/10.1016/j.renene.2020.04.083.

[130] Smith R. Chemical Process Design and Integration. 2nd ed. McGraw-Hill; 2005. https://doi.org/10.1529/biophysj.107.124164.

[131] Tobiesen FA, Svendsen HF, Juliussen O. Experimental Validation of a Rigorous Absorber Model for CO2 Postcombustion Capture 2007;53. https://doi.org/10.1002/aic.

[132] Faramarzi L, Kontogeorgis GM, Michelsen ML, Thomsen K, Stenby EH. Absorber Model for CO2 Capture by Monoethanolamine 2010:3751–9. https://doi.org/10.1021/ie901671f.

[133] Carranza-Abaid A, Jakobsen JP. A Non-Autonomous Relativistic Frame of Reference for Unit Operation Design. vol. 48. 2020. https://doi.org/10.1016/B978-0-12-823377-1.50026-4.

[134] Kvamsdal HM, Rochelle GT. Effects of the Temperature Bulge in CO 2 Absorption from Flue Gas by Aqueous Monoethanolamine 2008. https://doi.org/10.1021/ie061651s.

[135] Cousins A, Cottrell A, Lawson A, Huang S, Feron PHM. Model verification and evaluation of the rich-split process modification at an Australian-based post combustion CO 2 capture pilot plant. Greenh Gases Sci Technol 2012;2:329–45. https://doi.org/10.1002/ghg.1295.

[136] J.M. Smith, H.C. Van Ness MMA. Introduction to Chemical Engineering Thermodynamics. 3rd ed. McGraw-Hill, Inc.; 2004.

[137] Stewart M, Mcguire R, White R. Surface Production Operations: Design of Oil-Handling Systems and Facilities. Elsevier; 1999. https://doi.org/10.1016/B978-0-88415-821-9.X5000-3.

[138] Bellos E, Tzivanidis C. A comparative study of CO2 refrigeration systems. Energy Convers Manag.2019;1:100002. https://doi.org/10.1016/j.ecmx.2018.100002.

[139] International Organization for Standarization. ISO 14008: 2019—Monetary Valuation of Environmental Impacts and Related Environmental Aspects. 2019.

[140] Methanol Price and Supply/Demand 2021.

[141] Campbell JM. Gas Conditioning and Processing, Vol. 2: The Equipment Modules. 9th ed. Norman, Oklahoma: Campbell Petroleum Series; 2014.

[142] Henry hub natural gas spot price 2019.

[143] Statista. Average retail electricity price for industrial consumers in the U.S. from 1970 to 2019 2019.

[144] Zapiain-Salinas JG, Barajas-Fernández J, González-García R. Modified method to improve the design of Petlyuk distillation columns. Chem Cent J 2014;8:2–11. https://doi.org/10.1186/1752-153X-8-41.

[145] Mokhatab S, Poe WA, Mak JY. Natural Gas Treating. 2019. https://doi.org/10.1016/b978-0-12-815817-3.00007-1.

[146] Conesa A, Shen S, Coronas A. Liquid densities, kinematic viscosities, and heat capacities of some ethylene glycol dimethyl ethers at temperatures from 283.15 to 423.15 K. Int J Thermophys 1998;19:1343–58. https://doi.org/10.1023/A:1021927417610.

[147] Langan JR, Salmon GA. Physical Properties of N-Methylpyrrolidinone as Functions of Temperature. J Chem Eng Data 1987;32:420–2. https://doi.org/10.1021/je00050a009.

[148] Xiang HW, Laesecke A, Huber ML. A new reference correlation for the viscosity of methanol. J Phys Chem Ref Data 2006;35:1597–620. https://doi.org/10.1063/1.2360605.

[149] Korson L, Drost-Hansen W, Millero FJ. Viscosity of water at various temperatures. J Phys Chem 1969;73:34–9. https://doi.org/10.1021/j100721a006.

[150] Moser P, Wiechers G, Schmidt S, Goetheer E, Garcia J, Monteiro M, et al. ALIGN-CCUS: Results of the 18-month test with MEA at the pilot plant at Niederaussem. IEAGHG 5th Post Combust Capture Conf 2019:1–7.

[151] Seader JD, Henley EJ. Separation Process Principles. 2nd ed. Wiley; 2004.

[152] Panahi H, Eslami A, Golozar MA, Ashrafi Laleh A, Aryanpur M, Mazarei M. Failure analysis of type 304 stainless steel amine exchanger sheets in a gas sweetening plant. Case Stud Eng Fail Anal 2017;9:87–98. https://doi.org/10.1016/j.csefa.2017.08.002.

[153] Wu X, Wang M, Liao P, Shen J, Li Y. Solvent-based post-combustion CO2 capture for power plants: A critical review and perspective on dynamic modelling, system identification, process control and flexible operation. Appl Energy 2020;257:113941. https://doi.org/10.1016/j.apenergy.2019.113941.

Chapter 7.

A computationally efficient algorithm for designing unit operations

The description of a new unit operation design algorithm is presented in this chapter. This algorithm can be one order of magnitude faster than the typical iterative algorithm. Moreover, a rigorous modelling framework for gas-liquid contactors is presented here. The resulting model can substitute the equilibrium-based model in Chapter 6. This model serves to highlight which properties are needed for implementing a rigorous absorber model.

A Computationally Efficient Formulation of the Governing Equations for Unit Operation Design

Andres Carranza-Abaid*[a], Jana P. Jakobsen[a]

[a]Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), NO-7491, Trondheim, Norway

*Corresponding author: andres.c.abaid@ntnu.no

Keywords: Modelling, Process Design, Numerical Analysis, Gas-Liquid Contactor, CO_2 Capture

7.1. Abstract

A computationally efficient numerical method that uses a Pseudo-Eulerian formulation (PEF) for the design calculation of unit operations is presented and validated. This method is applicable to any unit operation that can be modelled using a system of ODEs. Performing the design of a unit operation in the PEF is tenfold faster than with the conventional Eulerian formulation (EF). The mathematical equivalence between the PEF and the EF is demonstrated by proving that the solution of different unit operation design problems provides the same numerical result independently of the formulation. It is shown that reducing the computation of the unit operation design problems also speeds the computation time of a process design or an optimization flowsheet. PEF allows the accurate estimation of the concentration or temperature profiles of complex unit operations such as multiphase reactor systems.

7.2. Introduction

The simulation of chemical engineering processes is a necessary task in the assessment of the techno-economic performance of chemical engineering projects. The flourishment of process simulations can partly be attributed to the steady and constant development of enhanced computer technologies in the recent decades. Although the computational capabilities of modern-day computers have been significantly enhanced over the past years, the computational resources may appear limited with respect to conceptual design and optimization superstructure frameworks. The complexity of these frameworks ultimately may lead to long overall computation times. Undesirable long times are greatly accentuated if the process has unit operations that are designed with ordinary differential equations (ODE). Due to the relevance of ODE-based models for process design, this work will focus on reducing the computation time of ODE systems by proposing a method that derives an alternative formulation of the governing equations.

The unit operation design methods can be divided in two main branches: short-cut methods and rigorous methods. The short-cut methods utilize simplified physics and, hence, do not require complex solution procedures (e.g., the Rayleigh equation in batch distillation, performance equation in reactor design or the height equivalent theoretical plate method for packed columns [HETP]). These methods are usually presented in process design textbooks (e.g. [1–3]) to highlight the fundamental concepts behind the unit operation rather than providing a rigorous description of the physics. On the other hand, rigorous models include different physics phenomena (e.g., thermodynamics and transport phenomena) into the mass and energy conservation equations to account for important effects. It is a common practice to design ODE-based unit operations using the models available in commercial software

such as Aspen Plus, CO2SIM or Mathcad. In these cases, a parameter is set (design specification) and an iterative procedure is utilized to find the operating parameter or equipment size that will yield the design specification (e.g., [4,5]). An alternative procedure, applicable for simplified physics, consists in transforming the mathematical model into a set of nonlinear algebraic equations and then using shifted-Legendre polynomials together with the orthogonal collocation method to design the unit operation [6].

Despite being computationally efficient, the applicability of the short-cut approaches is limited because they neglect important effects that affect the unit operation design and performance. This causes the short-cut models to have worse prediction capabilities than their rigorous counterparts. For example, by comparing the predictions done by a short-cut model (HETP) [7] and a rigorous model [8] of an amine-based CO_2 scrubber, it can be observed the later model has superior prediction capabilities of the concentration and temperature profiles. The proper estimation of these profiles is paramount in several unit operations (e.g., heat exchangers or gas-liquid contactors) because their operation and performance can be affected by mass or heat transfer pinch conditions.

Considering the need to reduce the computational costs of the simulations without losing the physical meaningfulness of the model. PEF provides a method for developing the governing equations in the Pseudo-Eulerian Formulation (PEF) (This work extends on the method previously presented [9]). This work presents the application of the PEF to the design of a plug-flow reactor and a gas-liquid contactor for the removal of CO_2. The results highlight the computational advantages of using the PEF to reduce the computational costs without oversimplifying the physics.

7.3. Methodology

7.3.1. Eulerian Formulation

The differential equations describing the property flows (mass and energy flows) can be derived using the Eulerian formulation (EF) or the Lagrangian formulation (LF). The difference between both formulations is the frame of reference used to derive the equations and the chosen control volume. The EF uses a control volume that is fixed in space; thus, it quantifies the property flow field from a stationary location. On the other hand, the control volume in the LF moves with the flow field, hence it quantifies the property flow field using a moving location. Of the two, the EF is more used, perhaps because of its computational advantages [10] or its simpler way of formulating and solving the governing equations.

The main characteristic of the EF is that the property flows are the state variables, and the spatial dimension of the system is the independent variable. The general form of a 1-D model can be expressed as a vector of state variables ($\underline{\beta}$) that are a function of an independent variable (θ) and a set of parameters ($\underline{\pi}$):

$$\frac{d}{d\theta}\left(\underline{\beta}(\theta)\right) = \underline{f}\left(\underline{\beta}(\theta), \underline{\pi}\right). \tag{7.1}$$

Note that the independent variable can be either a spatial or temporal dimension. The simulation of a unit operation described by Eq. (7.1) usually means solving a boundary value problem (BVP) where the boundary conditions of each one of the dependent variables must be specified in addition to the boundary where those conditions apply (θ_0 or θ_f).

Unit operation design calculation implies that the value of a state variable at a given location in the unit operation is specified (usually it is when $\theta = \theta_f$). The unit operation design specification fixes the state variable $\left(\beta^*_{spec}\right)$ at a certain θ_f. Since the value of θ_f is an unknown variable of the design problem and, at the same time, it is needed to solve Eq. (7.1), an iterative procedure is needed (see Figure 7.1 a)). A discrepancy function is then applied to evaluate how far the calculated specified state variable (β^*) is from the specified value (β^*_{spec}). For a constant set of input parameters $\underline{\pi}$, the evaluation of the discrepancy function (f) must be equal or below a specified tolerance (ε)

$$f = \left|\beta^*(\theta) - \beta^*_{spec}(\theta)\right| \leq \varepsilon. \tag{7.2}$$

Figure 7.1 a) illustrates that it is computationally expensive to solve a BVP problem in order to comply with Eq. (7.2) because of the iterative nature of the algorithm which is causing a computational bottle neck in the unit operation design calculation.

It is important to remark that a unit operation can be designed if and only if the solution of the state variables satisfies $\underline{\beta} \in \mathbb{R}$. Unit operation design calculation implicitly assumes that the solution is within physical boundaries, hence if there is no solution, the algorithm will not converge (the program may even crash in some cases if complex numbers appear during the calculations).

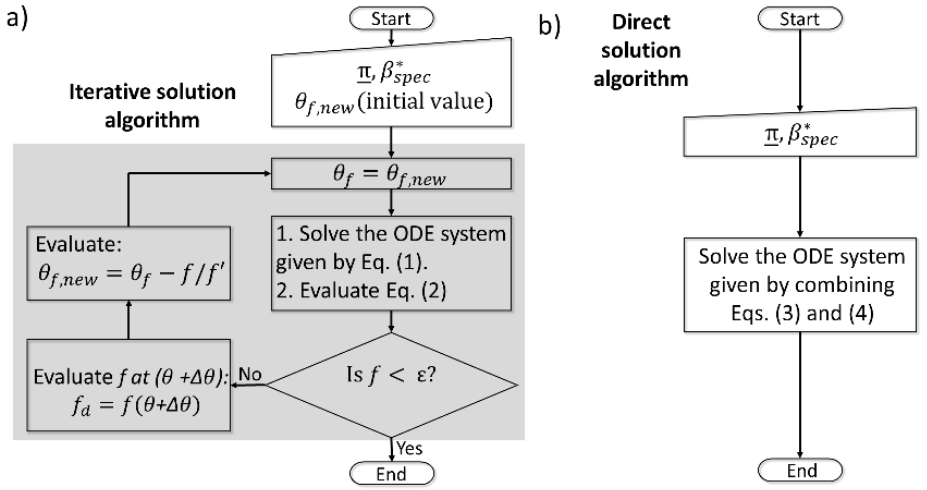


Figure 7.1. Algorithm to design a unit operation with a) Eulerian formulation (EF) assuming a Newton-Raphson iterative method and b) Pseudo-Eulerian Formulation (PEF).

7.3.2. Pseudo-Eulerian Formulation

This formulation is based on deriving the governing equations as a function of β^* instead of θ. The way to do this is to apply the chain rule between the vector of derivatives and the inverse of the ODE that describes the specified state variable. In this way, the new ODE system expresses the state variables (the ones that are not the specified state variable $\underline{\beta}^r$) as a function of the specified state variable.

This procedure gives the following governing equations with the PEF:

$$\frac{d}{d\beta^*}\left(\underline{\beta}^r(\beta^*)\right) = \underline{f}^r\left(\underline{\beta}^r(\beta^*), \underline{\mu}\right). \tag{7.3}$$

Note that $\underline{\beta}^r$ is a reduced vector with *c-1* variables. The PEF ODE system has a reduced geometry when compared to the ODE of the EF. This is because the procedure to derive the PEF equations makes the equation corresponding to the

specified variable a trivial solution. The form of Eq. (7.3) implies that when the system can be conceived within the physical boundaries ($\underline{\beta} \in \mathbb{R}$), all the mass and energy balances are relative to each other, hence, one can calculate how the state variables behave as a function of the specified state variable.

As seen in Figure 7.1 b), solving the ODE set given by Eq. (7.3) provides the solution of the mass and energy balances, but it does not calculate the final design of the unit operation. This is because the ODE set is relative to $\beta^*$and the solution is then in a dimensionless mathematical space that is not dependent on the spatial or temporal dimensions. Because of its mathematical properties and the physical implications of the PEF of the governing equations, it can be considered that the mass and energy balances are formulated in a different frame of reference than in the EF.

It should be noted that if the problem does not require to design the unit operation (i.e., the equipment volume is not needed), then it is enough to solve the equations with the form of Eq. (7.3). This can save further computation time since the extra computational resources needed to give the system some spatial dimensions are avoided.

On the other hand, if the unit operation design is needed, one must bring the solution from the dimensionless mathematical space to the spatial or temporal dimensions. This is done by adding a “dimensioning” function to the ODE system. This equation has the general form:

$$\frac{d\theta}{d\beta^*} = f^*\left(\underline{\beta}^r(\beta^*), \underline{\mu}\right) \tag{7.4}$$

Note that the ODE of Eq. (7.4) is the reciprocal of the ODE that describes the state variable behavior in the EF.

7.4. Results

This section presents the application of the PEF for two different unit operation design problems. The first example highlights the numerical and conceptual equivalence between the EF and the PEF models and shows the computational gains of using the PEF. The second example validates PEF-based model of a multiphase multicomponent reactor and discusses the effect of the computational advantages of using the proposed method on a superstructure framework.

7.4.1. Multiple Reactions in an Ideal Plug-Flow Reactor

This subsection presents the implementation of the PEF for the design of an ideal plug-flow reactor (Figure 7.2). The reaction mechanism is illustrated in Figure 7.2 and each one of the reactions has a first order kinetic behavior. The model considers a reactive liquid phase system with constant density that operates under an isothermal plug-flow hydrodynamic regime where the diffusivity of the components is infinitely slow compared to the reaction rate.

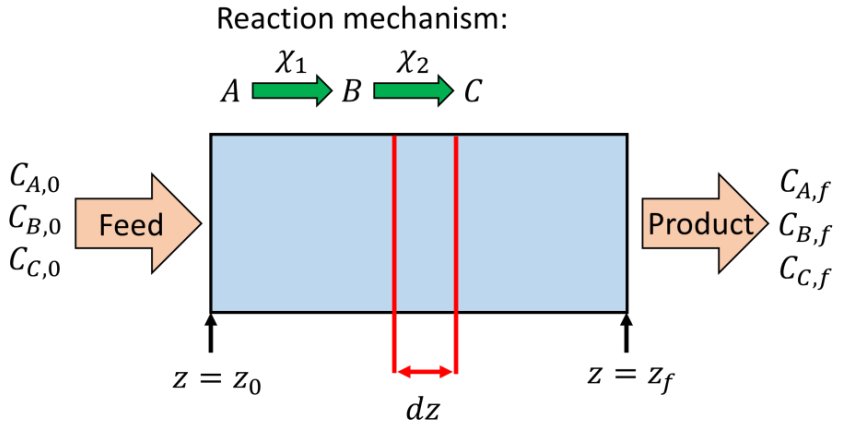


Figure 7.2. Sketch of the modelled tubular reactor with two reactions in series.

Eulerian Formulation: The equations and the boundary conditions that model the plug-flow reactor are:

$$\frac{dC_A}{dz} = -\frac{\chi_1 C_A}{v} \qquad B.C.: \; C_A = C_{A,0} \;\; for\; z = z_0 \tag{7.5}$$

$$\frac{dC_B}{dz} = \frac{\chi_1 C_A}{v} - \frac{\chi_2 C_B}{v} \qquad B.C.: \; C_B = C_{B,0} \;\; for\; z = z_0 \tag{7.6}$$

$$\frac{dC_C}{dz} = \frac{\chi_2 C_B}{v} \qquad B.C.: \; C_C = C_{C,0} \;\; for\; z = z_0 \tag{7.7}$$

The design problem consists in finding the reactor volume that yields a certain amount of component A, therefore, the discrepancy function is:

$$f = \left| C_{A,f} - C_{A,f|spec} \right| \leq \varepsilon \tag{7.8}$$

Pseudo-Eulerian formulation: considering the same assumptions and using the concentration of component A as the independent variable, the equations are:

$$\frac{dC_B}{dC_A} = -\frac{\chi_1 C_A - \chi_2 C_B}{\chi_1 C_A} \qquad B.C.: \; C_B = C_{B,0} \quad for\; C_A = C_{A,0} \tag{7.9}$$

$$\frac{dC_C}{dC_A} = -\frac{\chi_2 C_B}{\chi_1 C_A} \qquad B.C.: \; C_C = C_{C,0} \;\; for\; C_A = C_{A,0} \tag{7.10}$$

$$\frac{dz}{dC_A} = -\frac{v}{\chi_1 C_A} \qquad B.C.: \; z = z_0 \qquad for\; C_A = C_{A,0} \tag{7.11}$$

The key difference between the EF and the PEF is that while the PEF express the concentration values of components B and C when the concentration A has certain value ($C_{A,0}$), the B.C. in the EF express the concentration values of A, B, C where the inlet is being fed. From a practical point of view, both formulations “answer” different

questions. While the EF answers: “what happens to the component concentrations in a reactor with certain spatial dimensions?”, the PEF answers: “what should happen in the reactor to achieve this concentration of this component?”.

The PEF can also be utilized in the design of processes where time is the independent variable (batch or semi batch processes like reactors or drying systems). If we apply the reaction scheme shown in Figure 7.2 to a batch reactor, Eqs. (7.9) - (7.11) will have a similar form with minor differences: the independent variable would be time (t) instead of length (z) and Eq. (7.11) will not have the surface velocity term (v). The resulting set of ODEs would find the time t needed to reach certain concentration. It should be noted that in this case, the design of the unit operation does not include the estimation of the physical dimensions of the reactor, which can be easily estimated with an algebraic equation that involves time and the initial amount of reactants as it has traditionally been done in reactor design [3].

Transforming an already-implemented EF model into a PEF is a straightforward task that requires few modifications in the EF programming code. It is enough to evaluate the equations in their EF (form of Eq. (7.1)) and then multiply the numerical result by the inverse of the ODE containing the design specification. Note that the ODE containing the design specification must be multiplied twice in order to obtain the inverse of the design specification derivative (Eq. (7.11) was estimated by multiplying Eq. (7.5) by its inverse two times). Moreover, the boundary conditions should also be modified according to the change of independent variables, however, as seen in Eqs. (7.9) - (7.11), the numerical values are the same as in the EF.

In order to illustrate the equivalence between both formulations, a design calculation was performed of a plug-flow reactor that converts 98 % of an inlet stream that only contains 1 $mol\ m^{-3}$ of component A (see Figure 7.2). Figure 7.3 presents the obtained

concentration profiles of the 3 components as a function of the independent variable. Figure 7.3 a) is done by solving Eqs. (7.5) - (7.8) with the algorithm presented in Figure 7.1 a) while Figure 7.3 b) is the solution of Eqs. (7.9) - (7.11). The “spaceless” concentration profiles of Figure 7.3 b) can be transformed into the spatial-dependent concentrations profiles shown in Figure 7.3 a) by evaluating Eq. (7.11).

Both ODE systems are solved with the orthogonal collocation fifth-order method available in Matlab 2019b (bvp5c function). For the case of the EF, the Newton-Raphson numerical method was used to find the solution to the design problem by solving the discrepancy function.

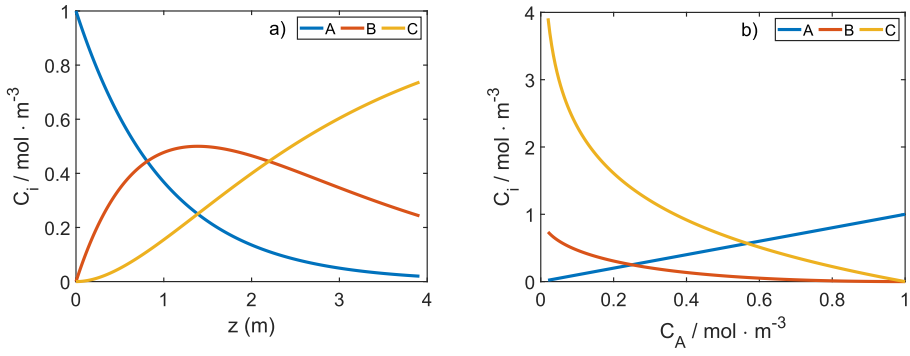


Figure 7.3. Concentration profiles obtained with solution of the ODE set given by the a) Eulerian formulation (EF) and b) Pseudo-Eulerian formulation (PEF). Parameters $(\underline{\pi})$: $C_{A,0} = 1\ mol\ m^{-3}$, $C_{B,0} = C_{C,0} = 0\ mol \cdot m^{-3}$, $\chi_1 = 1\ s^{-1}$, $\chi_2 = 0.5\ s^{-1}$ and $v = 1.0\ m \cdot s^{-1}$. Design spec: $C_{A,f} = 0.02\ mol \cdot m^{-3}$.

The equivalence between the EF and the PEF can be assessed by calculating the average absolute relative deviation (AARD) between the numerical results of both formulations.

The AARD was calculated using the following equation

$$AARD = \frac{1}{N}\sum_{1}^{N}\left|\frac{(Y_{EF} - Y_{PEF})}{\sqrt{(Y_{EF})(Y_{PEF})}}\right|, \tag{7.12}$$

where N is the total number of compared simulations, Y_{EF} and Y_{PEF} are the output variables calculated using the EF and PEF respectively.

An analysis of the computational speed and the numerical equivalence between both formulations was done by performing the simulations for the presented problem at different operating conditions. The Monte Carlo method was used to randomly generate the input parameters of 1,000 design problems from which the computational speed and the numerical difference between selected output variables ($C_{A,f}, C_{B,f}, C_{C,f}$ and z_f) was assessed. The varied input parameters and their respective ranges are shown in Table 7.1 and the remaining parameters have the same values as used in the calculations needed in Figure 7.3. The initial value to start the iterations in the EF calculations to obtain the required z_f was set to be equal to 1 m.

Table 7.1. Limits of the varied parameters in the Monte Carlo simulations.

Parameter	Min	Max
$\chi_1 \,/\, mol \cdot m^{-3} \cdot s^{-1}$	0.5	3.0
$\chi_2 \,/\, mol \cdot m^{-3} \cdot s^{-1}$	0.01	0.5
$v \,/\, m \cdot s^{-1}$	0.5	1.5
$C_{A,F} \,/\, mol \cdot m^{-3}$	0.01	0.99

Although both models use the same BVP solver, it is important to remark that the PEF does not require a discrepancy function to solve the design problem specifications, therefore, the entire computation time is spent on solving a single BVP problem. Contrarily, the EF requires to specify the tolerance of the discrepancy function (ε), which impacts the computational cost of the design problem as seen in Figure 7.4. The computational cost is defined as the number of times the ODE system is evaluated to solve a single design problem (i.e., the number of times the Matlab function containing the ODE system is evaluated). The relative computational cost (Ω) is defined as

$$\Omega = \frac{EF\ computational\ cost}{PEF\ computational\ cost} \tag{7.13}$$

Figure 7.4 presents two probability histograms as a function of Ω. The histograms use the same sampling sets but different discrepancy function tolerances. The tolerance in Figure 7.4 a) is set to be $\varepsilon = 10^{-5}$ while for Figure 7.4 b), $\varepsilon = 10^{-9}$. In both cases the probability of decreasing the computational costs by more than 5 times is > 95 % while for an entire order of magnitude is over 60 % for Figure 7.4 a) and 75 % for Figure 7.4 b). If one compares Figure 7.4 a) and Figure 7.4 b), it is possible to notice that the lower the tolerance, the larger the computational speed-up. This situation occurs because lower tolerances require more iterations in the Newton Raphson while the computational cost for the PEF remains constant because the design specification is fed as the upper boundary of the independent variable ($C_{A,f}$). This proves that the removal of the iteration loop in the algorithm (Figure 7.1) makes the operation unit design calculation more efficient.

One may wonder, how does the computational cost relate with the computational speed and does it speed-up the calculations? To answer this question, a parity plot between Ω and the inverse of the relative computational speed ($1/\kappa$) is shown in Figure 7.5. The relative computational speed is calculated with the following equation:

$$\kappa = \frac{EF\ computational\ speed}{PEF\ computational\ speed} \tag{7.14}$$

Figure 7.5 shows that the relative computational cost is inversely proportional to the relative computational speed. Therefore, reporting Ω or $1/\kappa$ is approximately equivalent. It is important to remark that, Although reporting the relative computational time may give a more meaningful insight to the end-user, computation times are subject to the available computing resources and other variables that are outside the scope of this discussion. For this reason, it is more convenient to compare the computational costs of algorithms instead of the computation time.

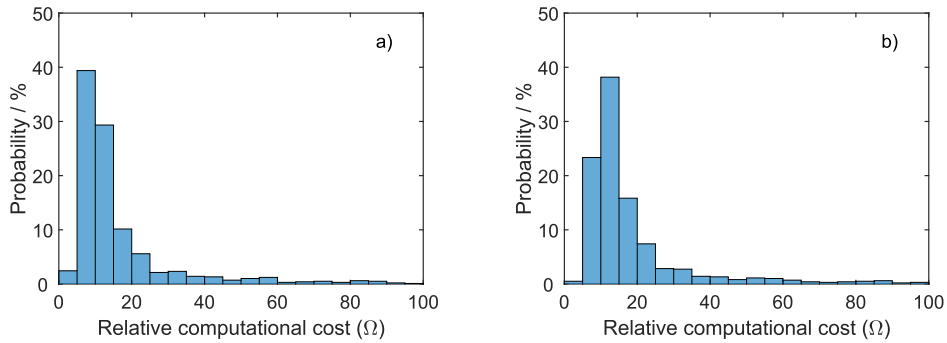


Figure 7.4. Probability distribution of the relative computational cost (ε) at different discrepancy function tolerance: a) $\varepsilon = 10^{-5}$ and b) $\varepsilon = 10^{-9}$.

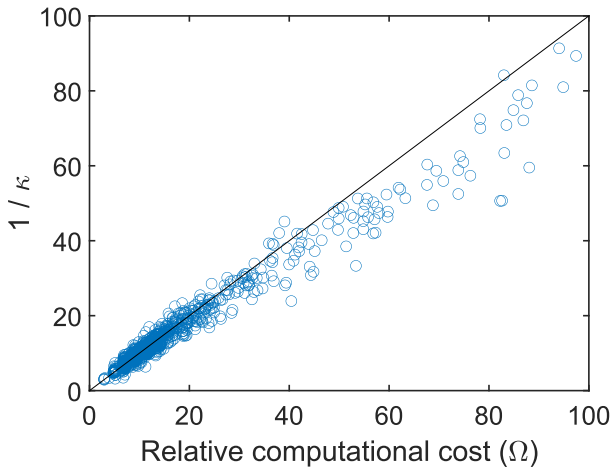


Figure 7.5. The reciprocal of the relative computational speed versus the relative computational costs between the Eulerian and the pseudo-Eulerian formulations.

Both formulations are equivalent from a mathematical standpoint. However, the comparison of obtained solutions may exhibit discrepancies caused by the non-exact nature of the numerical method used to solve the ODEs and the fact that the EF uses a discrepancy function, hence the calculated $C_{A,f}$ will be different from $C_{A,f|spec}$. In contrast, the PEF uses $C_{A,0}$ and $C_{A,f}$ as the limits of the independent variable, therefore, $C_{A,f} = C_{A,f|spec}$. The difference between solutions obtained from both formulations was assessed and it is presented in Figure 7.6. It shows that the AARD substantially decreases when the discrepancy function of the EF is $\varepsilon \geq 10^{-7}$ for all the output variables.

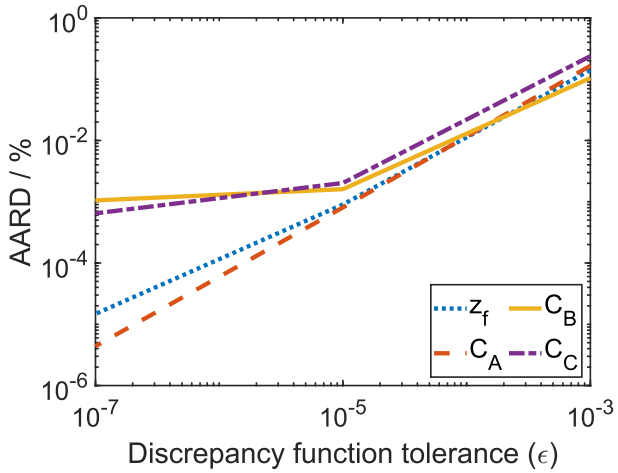


Figure 7.6. The AARD between the numerical solutions of the EF and the PEF equations for the ideal plug-flow reactor study.

7.4.2. Multicomponent Multiphase Reactor

7.4.2.1. PEF Governing Equations

This subsection presents the validation of a multicomponent multiphase reactor PEF-based model. The studied case is the CO_2 capture from a flue gas using an amine-based absorption relevant to both the chemical and the environmental engineering field. Absorbers for CO_2 capture are usually designed to remove a specified amount of CO_2 from a given flue gas, hence, developing the governing equations in the PEF can be particularly advantageous.

A sketch of the modelled system is presented in Figure 7.7. The following considerations were done in the model development:

1. The process is adiabatic, isobaric and is operated in steady state.
2. The vapor and liquid phases have a plug-flow hydrodynamic regime.

3. There are four components in the system: CO_2, monoethanolamine (MEA), H_2O and N_2.
4. N_2 is not soluble in the liquid phase.
5. The direction of the mass and energy transfer is assumed to be from the vapor phase to the liquid phase.
6. All thermal effects related to phase-shifting take place in the liquid phase.
7. The energy transport in the liquid phase is infinitely fast compared to the vapor phase.

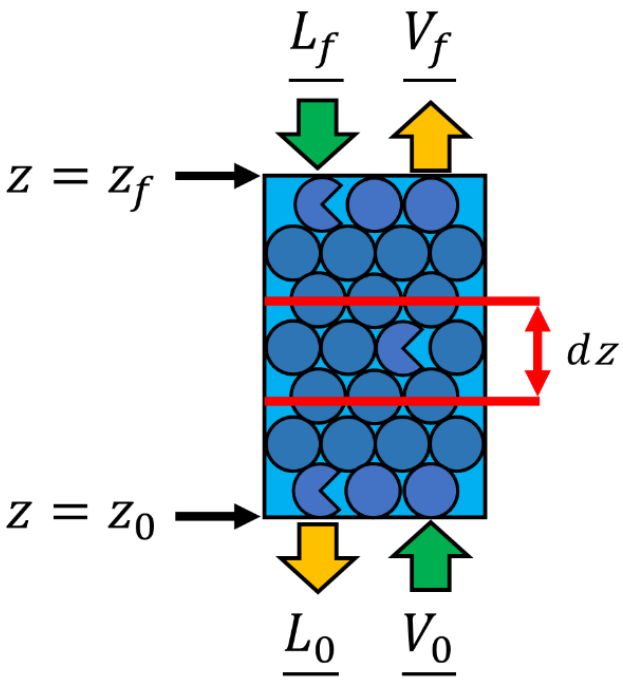


Figure 7.7. Sketch of the modelled multicomponent multiphase reactor. $\underline{V}$ and $\underline{L}$ represent the property flows of the vapor and liquid phases respectively.

According to assumptions #1 to #3, the model is independent of time and the state variables only vary along the axial dimension. Following the nomenclature of Figure 7.7, the resulting equations in the PEF are

$$\frac{dV_i^r}{dV_{CO_2}} = \frac{r_i^r}{r_{V,CO_2}} \quad (i \neq CO_2) \qquad B.C.: \; V_i^r = V_{i,0}^r \;\; for\, V_{CO_2} = V_{CO_2,0} \tag{7.15}$$

$$\frac{dL_i}{dV_{CO_2}} = \frac{r_i}{r_{V,CO_2}} \qquad B.C.: \; L_i = L_{i,f} \;\; for\, V_{CO_2} = V_{CO_2,f} \tag{7.16}$$

$$\frac{dE_V}{dV_{CO_2}} = \frac{r_{E,V}}{r_{V,CO_2}} \qquad B.C.: \; E_V = E_{V,0} \;\; for\, V_{CO_2} = V_{CO_2,0} \tag{7.17}$$

$$\frac{dE_L}{dV_{CO_2}} = \frac{r_{E,L}}{r_{V,CO_2}} \qquad B.C.: \; E_L = E_{L,1} \;\; for\, V_{CO_2} = V_{CO_2,f} \tag{7.18}$$

$$\frac{dz}{dV_{CO_2}} = -\frac{1}{A_C r_{V,CO_2}}. \qquad B.C.: \; z = z_0 \quad for\, V_{CO_2} = V_{CO_2,0}. \tag{7.19}$$

The form of the mass energy balance equations for gas-liquid contactors in the EF can be consulted in the literature [8,11–14].

The system of ODEs contains 8 equations from which 2 ODEs with the form of Eq. (7.15), 3 ODEs with the form of Eq. (7.16) and one ODE for each one of the remaining equations. Because of assumption #4, the evaluation of the ODEs corresponding to N_2 are always 0 and therefore can be eliminated from the ODE set. Furthermore, it is important to remark that although there are electrolyte compounds in the liquid mixture (caused by the solubilization of CO_2), their mass balances are lumped into the apparent mass balances of CO_2, MEA and H_2O.

The source terms in the above equations are estimated using an overall mass or energy transfer coefficient. Considering assumption #5, the rate term for component i can be defined by

$$r_i = \left[\frac{1}{\frac{ZRT_V}{a_e k_{V,i}} + \frac{H_i \varphi_i P^x}{\Xi a_e k_{L,i} C_L}}\right](P_i - P_i^x). \tag{7.20}$$

Note that the overall mass transfer coefficient is the term inside the square brackets. The first term in the denominator is the resistance to the mass transfer in the vapor phase while the second term corresponds to the liquid phase resistance. The enhancement factor term (Ξ) only applies to the CO_2 rate of mass transfer.

Since assumption #6 implies that the liquid phase gives or removes the necessary amount of energy from the molecule that is going to be transferred between both phases the energy source terms are:

$$r_{E,V} = \left[\frac{1}{\frac{1}{a_e h_V}}\right](T_V - T_L) + \sum_i^n r_i H_{V,i} \tag{7.21}$$

$$r_{E,L} = \left[\frac{1}{\frac{1}{a_e h_V}}\right](T_V - T_L) + \sum_i^n r_i \left(H_{V,i} - \lambda_i\right). \tag{7.22}$$

Assumption #7 implies that the overall energy transfer coefficient is only a function of the vapor phase resistance to energy transfer. The enthalpy of phase change (λ_i) is estimated using a rigorous method that sums the enthalpies of vaporization and reaction [15].

7.4.2.2. Complementary Equations and Parameters

The model validation was performed by comparing the simulated results with the measured values from different pilot plant systems reported in the open literature. A summary of the equations used to describe the absorber transport phenomena and thermodynamics is presented in Table 7.2. Two different systems are used as the benchmark to validate the absorber model. System 1 refers to the data measured in a random packed column [11] and system 2 refers to the experimental data obtained in a structured packed column [8]. Since the packing material in the experimental setups is different, the correlations to describe the mass transfer coefficients will differ as well.

Table 7.2. Description of the parameters used in the model validation of the gas-liquid contactor.

System	Parameter	Comment	Reference
1,2	Z	Peng - Robinson	[16]
1,2	φ_i	Peng - Robinson	[16]
1,2	H_i, P^*, P_i^*	Machine learning based surrogate thermodynamic model	[17]
1	$a_e, k_{V,i}, k_{L,i}$	Random packing (Berl saddles)	[18,19]
2	$a_e, k_{V,i}, k_{L,i}$	Structured packing (Mellapak 250Y)	[20,21]
1,2	E	Irreversible enhancement factor	[22]
1,2	h_V	Chilton-Colburn analogy	[23]

The correlations used in the gas-liquid contactor model require certain properties of the gas and liquid phases to be estimated. The methods used to estimate the physical properties are shown in Table 7.3. The properties labeled as machine learning models in Table 7.3 use a shallow neural network architecture. The models, their parameters and their statistical analysis are presented in the supporting information.

Table 7.3. Methods to estimate the properties of the gas-liquid contactor.

Property	Comment	Reference
ρ_V	Peng-Robinson	[16]
K_V	Non-linear mixing rule and correction for higher pressures	[24–26]
μ_V	Non-linear mixing rule and correction for higher pressures	[25–27]
$D_{i,V}$	Predictive method of the binary diffusion coefficients	[28]
CP_V	DIPPR equation	Parameters taken from Aspen Plus v8.6 Databank
ρ_L	In-house machine learning model	[29–31]
σ_L	In-house machine learning model	[29–32]
μ_L	In-house machine learning model	[33–36]
$D_{i,L}$	Empirical correlation for alkanolamine solutions*	[37]

*$D_{H_2O,L}$ is held constant and equal to 10^{-9} m^2/s

7.4.2.3. Model Application

This subsection shows that, as opposed to shortcut methods, the PEF allows the proper estimation of the profiles of the process parameters. As discussed in section 1, it is important for the unit operation models to properly estimate the location of the mass and energy transfer pinch points because they affect the absorber performance [38]. As it can be seen in the profiles of α_{CO_2}, y_{CO_2}, T_V and T_L in Figure 7.8, the calculations not only have good agreement with the stream outlet values, but also on the mass and energy balances profiles. Figure 7.8 shows a proper prediction of the temperature bulges that are commonly observed in CO_2-amine systems. In the cases of Figure 7.8 a) – c), the bulge is located at the bottom of the absorber whereas in Figure 7.8 d) is located close to the top of the absorber. This is because the relation between the liquid and vapor flows (L/V ratio) in system 1 is larger than in system 2, which means that lower L/V ratios move the bulge location higher in the column. The validation of other important process variables is presented in the supporting information.

In the same fashion as in the previous example, a Monte Carlo method was used to perform 1,000 simulations in order to assess the difference between using the EF or the PEF. The simulations were done using the packing specifications of system 2 and varying the inputs inside the ranges shown in Table 7.4 considering a tolerance for the discrepancy function of $\varepsilon = 10^{-7}$. The results of the Monte Carlo simulations are presented in Table 7.5. The difference between the solutions of both models is negligible for practical purposes as the AARD between all the important variables is less than 10^{-6} which means that the relative deviation is caused by the inherent error of the BVP solver used and the iterative loop needed to obtain the dimensions of the absorber in EF.

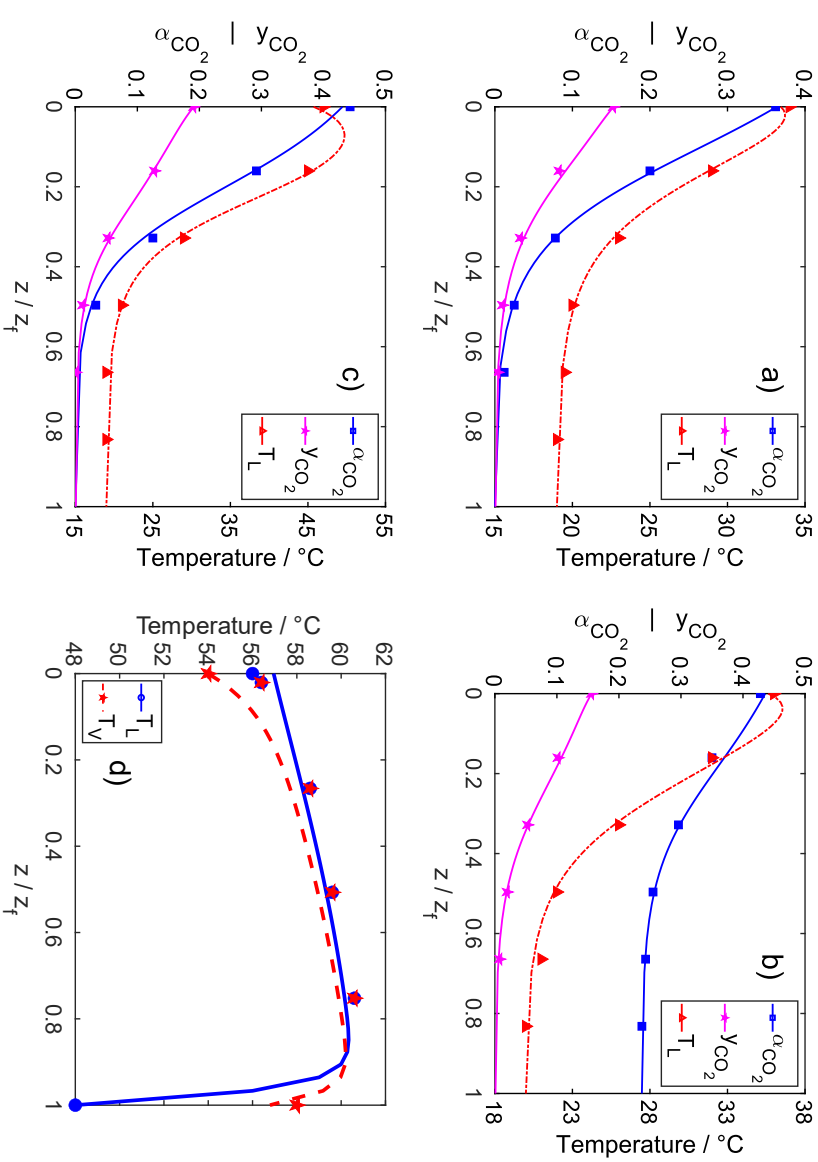


Figure 7.8. Comparison between the model predictions and the corresponding experimental values at different operating conditions: a) run 13 (system 1), b) run 17 (system 1), c) run 22 (system 1) and d) run 10 (system 2).

Table 7.4. Limits for the Monte Carlo simulations in the absorber case study.

Variable	Min.	Max.
$y_{CO_2,0}$	0.003	0.10
$y_{CO_2,0}$	0.01	0.2
$\alpha_{CO_2,0}$	0.1	0.25
$L\ /\ V$	6	12
$T_L, T_V\ /\ °C$	40	
$V_{0,T}\ /\ mol/s$	1	
$w_{MEA}\ /\ \%$	30	

The comparison of the computational costs between both formulations is given in Figure 7.9. This histogram has a similar shape as in the first plug-flow example (Figure 7.4). It shows that for a gas-liquid contactor, a speed-up of around one order of magnitude is expected. Although the computational costs may differ from system to system, the relative computational costs of the PEF seem to be, in average, around one order of magnitude.

Another advantage of the PEF over the EF is when an unfeasible design is proposed as a specification (i.e., when the solution of the state variables $\underline{\beta} \notin \mathbb{R}$). It was noticed that when the EF is used to solve the problem, the procedures run for a long time because the iterative algorithm diverges until either the program "crashes" or the maximum number of iterations is reached. Whereas the PEF quickly finds which

solution is not feasible and hence stops the calculations earlier without spending additional computational resources.

Table 7.5. Difference between the EF and PEF of selected output variables of the absorber model.

Variable	AARD / %	Max. AARD / %
z_f	8 x 10^{-7}	1 x 10^{-5}
$y_{CO_2,1}$	5 x 10^{-5}	2 x 10^{-4}
$\alpha_{CO_2,0}$	5 x 10^{-9}	1 x 10^{-8}
$T_{V,1}$	2 x 10^{-6}	8 x 10^{-5}
$T_{L,0}$	9 x 10^{-7}	2 x 10^{-5}

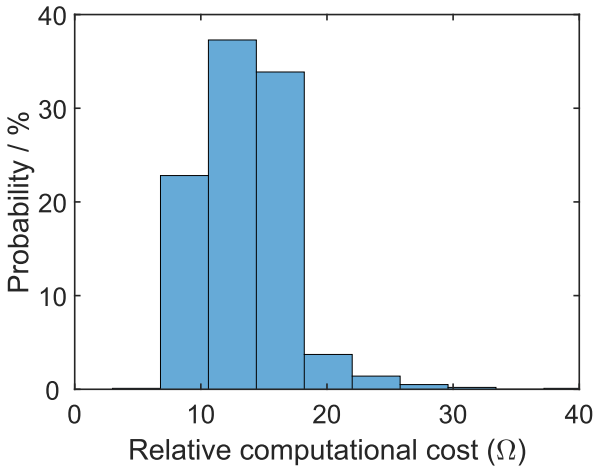


Figure 7.9. Results of the relative computational cost between the EF over the PEF.

From the end-user point of view, spending 1.5 seconds instead of 0.1 seconds to design a single unit operation is more a minor inconvenience rather than a fatal issue. However, unit operations are usually part of larger processes and optimization superstructure frameworks. Using the PEF in the evaluation of these frameworks is beneficial because if the unit operation speed can be reduced more than ten times, their overall computation speed should also be reduced accordingly.

To illustrate this, let us take the examples of the biogas upgrading processes shown in Figure 7.10 that remove the CO_2 from the raw biogas to produce biomethane [39]. Figure 7.10 a) shows the amine-based biogas upgrading process, which contains two unit operations that are designed with ODE-based models (marked in red). In order to test the PEF computational advantages, this process was designed and optimized by manipulating the solvent flowrate, amine concentration and desorber pressure. It was found that utilizing the PEF for optimization frameworks can be 9-20 times faster than with the EF. These values agree with those shown in Figure 7.9 because more than 99 % of the computational time is spent on the design of the absorber and the stripper. Conversely, designing the physical solvent process (Figure 7.10 b)) with the PEF, a computational speed is around 7 to 16 times faster. Although still quite high, the computational advantages of the PEF are not as high as in the amine process because in the second process. In this case only 80 % of the computational time is spent on the absorber because the calculations involved in the flash tanks, recirculation compressors and reboiler are computationally intensive as well. The same relative computational speed up is observed when any of the two processes is designed at fixed conditions or when the processes are optimized.

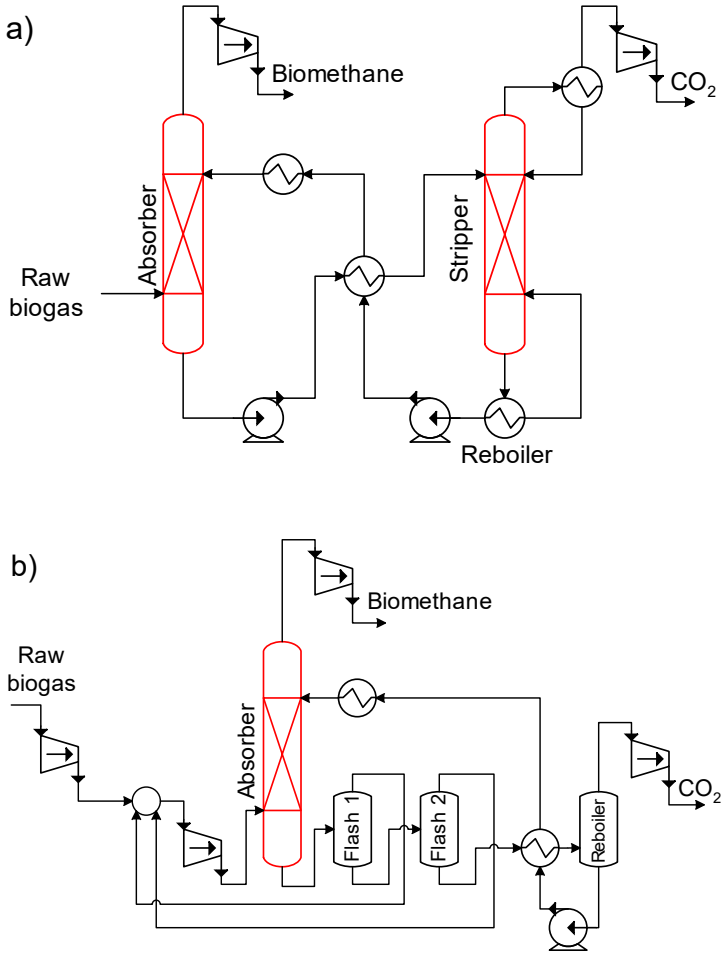


Figure 7.10. Biogas upgrading technologies

a) Aqueous amine process b) Physical-solvent process.

Reducing the computation time of any process superstructure framework has great practical potential since it is quite convenient for the analysis and selection of materials such as chemical solvents for CO_2 removal. If one wants to assess the techno-economic potential of a novel solvent for CO_2 capture, it necessary to have a mixed-integer optimization framework that considers multiple process parameters,

several process configurations, and other relevant variables (e.g., plant location or solvent supplier). Instead of unnecessarily spending days or even weeks in evaluating all the possible scenarios for an optimization, one could perform the same comprehensive analysis in few hours by reformulating the models with the PEF. This approach is not only limited for solvent-based technologies since the PEF can be utilized in other superstructure frameworks that involve other processes such as simple distillation (packed column), reactive distillation, membrane separation, adsorption or batch processes.

7.5. Conclusions

A Pseudo-Eulerian formulation (PEF) was proposed to develop the governing equations for the design calculation of unit operations. It was shown that this alternative formulation (PEF) speeds-up the unit operation design algorithm by removing the inherent iterative loop that arises from developing and solving the governing equations in the Eulerian framework. It was demonstrated that utilizing the PEF algorithm for unit operation design is more than tenfold faster than the EF algorithm. The solution of the PEF-based model gives similar numerical results as the EF-based models with minimal numerical differences that can be attributed to the numerical methods.

The PEF approach was used, as an example, to model a multiphase multicomponent reactive system (CO_2 absorber). It was shown that, as opposed, to the short-cut methods, it can properly represent the mass and temperature profiles. This shows that there is no need to oversimplify the physics of the unit operations to implement computationally fast unit operation design algorithms.

Using the PEF to develop unit operation design models can be utilized for greatly improving the computational speed of rigorous superstructure frameworks that involve conceptual process design such as the ones utilized in sensitivity, optimization, or uncertainty quantification studies. The computational advantages of the PEF are more significant when there is a larger number of unit operations that are modelled with an ODE.

7.6. Supplementary Information

7.6.1. Density, surface tension and viscosity models

The machine learning models. for density (ρ_L), surface tension (σ_L) and viscosity (μ_L) in the liquid phase are a function of α_{CO_2}, w_{MEA} and T_L. All the machine learning models were developed by using a shallow artificial neural network architecture and the hyperbolic tangent sigmoid function. Hence, they have the machine learning models have the following general form:

$$\underline{x_n} = \left(\underline{x} - \underline{x_A}\right) .* \underline{x_B} + x_C \tag{7.23}$$

$$\underline{n} = \underline{b_1} + \underline{I} * \underline{x_n} \tag{7.24}$$

$$\underline{a_1} = 2./\left(1 + exp\left(-2\underline{n}\right)\right) - 1 \tag{7.25}$$

$$\underline{a_2} = \underline{b_2} + \underline{O^T} * \underline{a_1} \tag{7.26}$$

$$y = Q * exp\left((a_2 - yC)/y_B + y_A\right) \tag{7.27}$$

Where the vectors are assumed to be column vectors, $\underline{x}$ is the vector of input variables (they must be input in the following order: α_{CO_2}, w_{MEA} and T_L) and y is the predicted property. The operator "$.*$" indicates that a Schur product (or element by element product). The remaining unknown variables are the parameters of each model and are presented in Tables Table 7.6 - Table 7.10. Note that ";" separates rows and "," separates columns. Parity plots of the experimental properties and the predicted properties with the machine learning models are presented in

The average relative deviation (ARD) and the AARD of with respect to the experimental data are presented in Figure 7.11. The AARD shows an acceptable deviation between the model predictions and the experimental data while the small ARD values imply that there is no sign of model bias, and that the experimental data is well distributed around the model.

Table 7.6.Parameter values for $\underline{x_A}$, $\underline{x_B}$ and x_C.

Property	x_A	x_B	x_C
ρ_L	[0;0.062;293.15]	[4;2.132;0.0333]	-1
σ_L	[0;0;298.15]	[4;2;0.05]	-1
μ_L	[0;0;0]	[4;2;0.00566]	-1

Table 7.7. Parameter values for $\underline{b_1}$, and b_2.

Property	b_1	b_2
ρ_L	[0.8048;0.1141;2.6733]	-1.435
σ_L	[1.623;-0.7740; 3.381]	-0.03411
μ_L	[-0.8828;-1.054;-0.09768;1.165]	0.4406

Table 7.8. Parameter values for $\underline{I}$.

Property	$\underline{I}$
ρ_L	[0.2863, -0.7952, -0.1194; 0.4250, 1.137, -0.09251; -0.4639, 2.283, -0.3126]
σ_L	[-0.8323, 0.4304, 1.164; -0.2820, 0.2855, 0.09529; 0.3012, 2.192, 0.08269]
μ_L	[0.2324, 0.3806,1.900; 0.1911, 1.090, 2.075; 0.1965, 0.3745, -0.8437; 0.8272, 0.4359, 2.088]

Table 7.9. Parameter values for $\underline{O}$.

Property	$\underline{O}$
ρ_L	[1.654, 1.157, 0.3694]
σ_L	[-0.07337; -2.485; -1.1549]
μ_L	[-2.116; 1.084; 2.645; 2.139]

Table 7.10. Parameter values for Q.

Property	Q
ρ_L	1.0 (kg/m^3)
σ_L	1.0 (N/m)
μ_L	10^{-6} (Pa s)

Table 7.11. Deviation of the machine learning property models with respect to the experimental values. Application range for

Model	α_{CO_2}	w_{MEA}	T	No. Datapoints	AARD / %	ARD / %
ρ_L	$0-0.5$	$0.062-1$	$293-353$	271	0.311	0.003
σ_L	$0-0.5$	$0-1$	$298-358$	278	1.67	0.162
μ_L	$0-0.5$	$0.062-1$	$293-353$	307	4.93	0.130

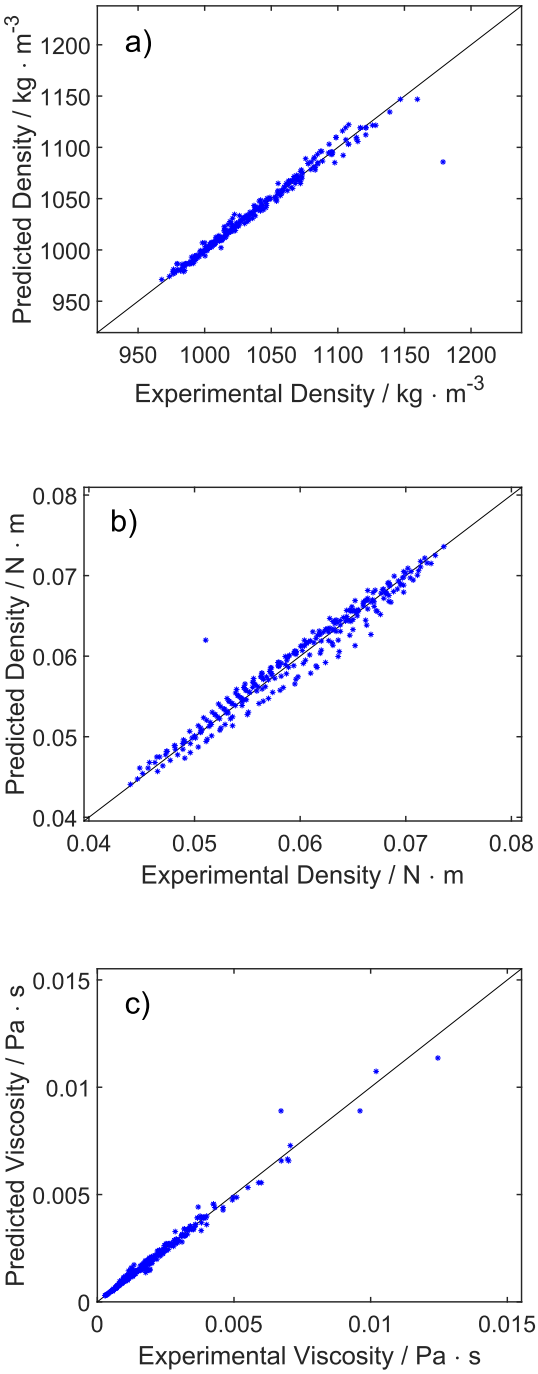


Figure 7.11. Parity plots of the experimental property values and the predictions with the machine learning models: a) density, b) surface tension, and c) viscosity. The parity plots present all the experimental datapoints within the ranges reported in Table 7.11.

7.6.2. Validation of the CO_2 scrubbing unit operation

The model validation was performed using the input values in the ranges shown in Table 7.12. Since the PEF is aimed to design absorber columns and not to simulate already existing columns, the amount of CO_2 at the outlet ($V_{CO_2,1}$ that is expressed as $y_{CO_2,0}$ in Table 7.12) was iterated until the total height of the absorber (z_f) was obtained. If the EF is used for the validation, then it is not necessary to perform this iterative procedure. In summary, if the model is to be used for a simulation of an already existing process it is better to develop the model in the EF, but if it is going to be used for process design and optimization it is better to use the PEF. The error of the mass balances in the PEF is quantified by comparing the total amount of CO_2 removed in the outlet ($y_{CO_2,1}$) and the rich CO_2 loading error ($\alpha_{CO_2,1}$). The CO_2 loading is defined as the apparent number of CO_2 moles over the apparent number of MEA moles. The mass balances show reasonable deviation percentages in $y_{CO_2,1}$ and $\alpha_{CO_2,1}$. Although both variables quantify the amount of CO_2 in the vapor and liquid phases respectively, the deviation of $\alpha_{CO_2,1}$ is considerably lower. This may be attributed to a wrong estimation of water vapor in the inlet gas stream (Tobiesen, Svendsen, & Juliussen, 2007). The error in the energy balances is quantified as the difference in the outlet temperatures of the vapor phase ($T_{V,1}$) and the liquid phase ($T_{L,0}$). The deviations are small and have good agreement with the measured values. It is seen that the ARD and AARD of $T_{L,0}$ are equal. This means that the temperature outlet is slightly underpredicted (a maximum difference of -3 °C).

Table 7.12. Results of the validation of the PEF model against experimental data.

		System 1	System 2
Dimensions (Input)	A_C / m^2	0.0177	0.00785
	z_f / m	4.36	6.55
Bottom Vapor (Input)	V_0 / mol/s	0.11 – 0.14	1.46 – 1.68
	$y_{CO_2,0}$	0.15 – 0.19	0.015 – 0.11
	$T_{V,0}$ / $°C$	34	39 – 69
Top Liquid (Input)	L_1 / mol/s	1.00 – 1.49	2.27 – 6.95
	$\alpha_{CO_2,0}$	0 - 237	0.18 – 0.41
	$w_{MEA,0}$ / %	12 - 18	30
	$T_{L,0}$ / $°C$	19 – 20	40 - 66
Top Vapor (Output)	AARD/ARD: $y_{CO_2,1}$	0.88 / 0.79	5.22 / -0.60
	AARD/ARD: $T_{V,1}$ / $°C$	N/A	0.49 / -0.09
Bottom Liquid (Output)	AARD/ARD: $\alpha_{CO_2,0}$	1.90 / 0.42	1.01 / -0.23
	AARD/ARD: $T_{L,0}$ / $°C$	0.26 / 0.24	0.46 / -0.46

Nomenclature

A_C: cross-sectional area [m^2]

a_e: effective specific interfacial area [m^2/m^3]

$\underline{f}$: vector of functions

C: concentration [mol/m^3]

CP: heat capacity [kJ/mol K]

D: diffusivity [m^2/s]

E: energy flow [kJ/s]

H: thermodynamic factor ($H = y/x$) [=]

h_V: heat transfer coefficient [kJ/m^2 s K]

k: mass transfer coefficient [m/s]

K: conductivity [kJ/m K]

L: liquid phase mole flow [mol/s]

P: pressure [kPa]

R: ideal gas constant [kPa m^3/mol K]

T: temperature [K]

V: vapor phase mole flow [mol/s]

v: superficial velocity [m/s]

w: MEA weight fraction in a CO_2-free basis

x: liquid molar fraction

y: vapor molar fraction

Z: compressibility factor [=]

z: length of the unit operation [m]

A_C: cross-sectional area [m^2]

Superscript

*: related to the specified variable

r: reduced vector

x: thermodynamic variable *at equilibrium

Subscript

0: bottom of the unit operation

E: related to energy

f: top of the unit operation

i: related to component i

L: related to the liquid phase

T: total

Greek Letters

α_{CO_2}: CO_2 loading [mol CO_2/mol MEA]

$\underline{\beta}$: state variable vector

ε: absolute error tolerance

θ: independent variable

κ: relative computational speed

λ: heat of phase change [kJ/mol]

μ: viscosity [Pa s]

Ξ: enhancement factor [=]

ρ: density [kg/m^3]

σ: surface tension [N/m]

φ: fugacity coefficient

χ: reaction constant [1/s]

Ω: relative computational cost

7.7. References

[1] W. McCabe, J. Smith, P. Harriott, Unit Operations of Chemical Engineering, 7th ed., McGraw-Hill, 2004.

[2] J.D. Seader, E.J. Henley, Separation Process Principles, 2nd ed., Wiley Series, 2004.

[3] O. Levenspiel, Chemical Reaction Engineering, 3rd ed., 1998. https://doi.org/10.1016/0009-2509(80)80132-1.

[4] U. Lee, J. Burre, A. Caspari, J. Kleinekorte, A.M. Schweidtmann, A. Mitsos, Techno-economic Optimization of a Green-Field Post-Combustion CO2Capture Process Using Superstructure and Rate-Based Models, Ind. Eng. Chem. Res. 55 (2016) 12014–12026. https://doi.org/10.1021/acs.iecr.6b01668.

[5] A.A. Cuadri, J.E. Martín-Alfonso, J. Urbano, Using Mathcad to facilitate the design of chemical reactors involving multiple reactions, Comput. Appl. Eng. Educ. 28 (2020) 293–303. https://doi.org/10.1002/cae.22192.

[6] R. Garma, H. Binous, H. Dhaouadi, A. Bellagi, Design of continuous contacting countercurrent unit operations: An approach based on the usage of orthogonal collocation and <scp>Matlab</scp>, Comput. Appl. Eng. Educ. 27 (2019) 1308–1332. https://doi.org/10.1002/cae.22153.

[7] A. Alhajaj, N. Mac Dowell, N. Shah, A techno-economic analysis of post-combustion CO2 capture and compression applied to a combined cycle gas turbine: Part I. A parametric study of the key technical performance indicators, Int. J. Greenh. Gas Control. 44 (2016) 26–41. https://doi.org/10.1016/j.ijggc.2015.10.022.

[8] F.A. Tobiesen, H.F. Svendsen, O. Juliussen, Experimental Validation of a Rigorous Absorber Model for CO2 Postcombustion Capture, 53 (2007). https://doi.org/10.1002/aic.

[9] A. Carranza-Abaid, J.P. Jakobsen, A Non-Autonomous Relativistic Frame of Reference for Unit Operation Design, in: Comput. Aided Chem. Eng., 2020: pp. 151–156. https://doi.org/10.1016/B978-0-12-823377-1.50026-4.

[10] H.A. Jakobsen, Chemical reactor modeling: Multiphase reactive flows: Second edition, 2014. https://doi.org/10.1007/978-3-319-05092-8.

[11] P. Tontiwachwuthikul, A. Meisen, C.J. Lim, CO2 absorption by NaOH, monoethanolamine and 2-amino-2-methyl-1-propanol solutions in a packed column, Chem. Eng. Sci. 47 (1992) 381–390. https://doi.org/10.1016/0009-2509(92)80028-B.

[12] R. Taylor, R. Krishna, Multicomponent Mass Transfer, Wiley Series, 1993.

[13] J. Gabrielsen, H.F. Svendsen, M.L. Michelsen, E.H. Stenby, G.M. Kontogeorgis, Experimental validation of a rate-based model for CO2 capture using an AMP solution, Chem. Eng. Sci. 62 (2007) 2397–2413. https://doi.org/10.1016/j.ces.2007.01.034.

[14] L. Faramarzi, G.M. Kontogeorgis, M.L. Michelsen, K. Thomsen, E.H. Stenby, Absorber Model for CO2 Capture by Monoethanolamine, (2010) 3751–3759. https://doi.org/10.1021/ie901671f.

[15] I. Kim, K.A. Hoff, E.T. Hessen, T. Haug-Warberg, H.F. Svendsen, Enthalpy of absorption of CO2 with alkanolamine solutions predicted from reaction equilibrium constants, Chem. Eng. Sci. 64 (2009) 2027–2038. https://doi.org/10.1016/j.ces.2008.12.037.

[16] R. Peng, P-R.a New Equation of State., Proc. Natl. Acad. Sci. U. S. A. 15 (1976) 11–18. http://dns2.asia.edu.tw/~ysho/YSHO-English/1000 CE/PDF/Ind Eng Che Fun15, 59.pdf.

[17] A. Carranza-Abaid, H.F. Svendsen, J.P. Jakobsen, Surrogate modelling of VLE: Integrating machine learning with thermodynamic constraints, Chem. Eng. Sci. X. 8 (2020) 100080. https://doi.org/10.1016/j.cesx.2020.100080.

[18] K. Onda, H. Takeuchi, Y. Okumoto, Mass Transfer Coefficients Between Gas and Liquid Phases in Packed Columns, J. Chem. Eng. Japan. 1 (1968) 56–62. https://doi.org/10.1252/jcej.1.56.

[19] J.L. Bravo, J.R. Fair, Generalized Correlation for Mass Transfer in Packed Distillation Columns, Ind. Eng. Chem. Process Des. Dev. 21 (1982) 162–170. https://doi.org/10.1021/i200016a028.

[20] J.A. Rocha, J.L. Bravo, J.R. Fair, Distillation Columns Containing Structured Packings: A Comprehensive Model for Their Performance. 1. Hydraulic Models, Ind. Eng. Chem. Res. 32 (1993) 641–651. https://doi.org/10.1021/ie00016a010.

[21] P. Suess, L. Spiegel, Hold-up of mellapak structured packings, Chem. Eng. Process. 31 (1992) 119–124. https://doi.org/10.1016/0255-2701(92)85005-M.

[22] X. Luo, A. Hartono, S. Hussain, H.F. Svendsen, Mass transfer and kinetics of carbon dioxide absorption into loaded aqueous monoethanolamine solutions, Chem. Eng. Sci. 123 (2015) 57–69. https://doi.org/10.1016/j.ces.2014.10.013.

[23] T.H. Chilton, A.P. Colburn, Mass transfer (Absorption) Coefficients, Ind. Eng. Chem. Res. 26 (1934) 1183–1187. http://infoscience.epfl.ch/record/60296.

[24] A. Wassiljewa, Warmeleitung in gasgemischen, Phys. Z. (1904).

[25] D. Misic, G. Thodos, The thermal conductivity of hydrocarbon gases at normal pressures, AIChE J. 7 (1961) 264–267. https://doi.org/10.1002/aic.690070219.

[26] L.I. Stiel, G. Thodos, The thermal conductivity of nonpolar substances in the dense gaseous and liquid regions, AIChE J. 10 (1964) 26–30. https://doi.org/10.1002/aic.690100114.

[27] L.A. Bromley, C.R. Wilke, Viscosity Behavior of Gases, Ind. Eng. Chem. 43 (1951) 1641–1648. https://doi.org/10.1021/ie50499a046.

[28] E.N. Fuller, P.D. Schettler, J.C. Giddings, New Method for Prediction of Binary Gas-Phase Diffusion Coefficients, Ind. Eng. Chem. 58 (1966) 18–27. https://doi.org/10.1021/ie50677a007.

[29] T.G. Amundsen, L.E. Oi, D.A. Eimer, Density and Viscosity of Monoethanolamine plus Water plus Carbon Dioxide from (25 to 80) degrees C. Journal of Chemical and Engineering Data, 54(11), 3096–3100. https://doi.org/10.1021, J. Chem. Eng. Data. 54 (2009) 3096-3100Amundsen, T. G., Oi, L. E., Eimer, D. A. https://doi.org/10.1021/je900188m.

[30] A. Hartono, E.O. Mba, H.F. Svendsen, Physical Properties of Partially CO2 Loaded Aqueous Monoethanolamine (MEA), J. Chem. Eng. Data. 59 (2014) 1808–1816. https://doi.org/10.1021/je401081e.

[31] R.H. Weiland, J.C. Dingman, D.B. Cronin, G.J. Browning, Density and viscosity of some partially carbonated aqueous alkanolamine solutions and their blends, J. Chem. Eng. Data. 43 (1998) 378–382. https://doi.org/10.1021/je9702044.

[32] L. Korson, W. Drost-Hansen, F.J. Millero, Viscosity of water at various temperatures, J. Phys. Chem. 73 (1969) 34–39. https://doi.org/10.1021/j100721a006.

[33] S.A. Jayarathna, C.K. Jayarathna, D.A. Kottage, S. Dayarathna, D.A. Eimer, M.C. Melaaen, Density and surface tension measurements of partially carbonated aqueous monoethanolamine solutions, J. Chem. Eng. Data. 58 (2013) 343–348. https://doi.org/10.1021/je300920t.

[34] Z. Idris, J. Han, S. Jayarathna, D.A. Eimer, Surface Tension of Alkanolamine Solutions: An Experimentally Based Review, Energy Procedia. 114 (2017) 1828–1833. https://doi.org/10.1016/j.egypro.2017.03.1310.

[35] S.A. Jayarathna, A. Weerasooriya, S. Dayarathna, D.A. Eimer, M.C. Melaaen, Densities and surface tensions of CO2 loaded aqueous monoethanolamine solutions with r = (0.2 to 0.7) at T = (303.15 to 333.15) K, J. Chem. Eng. Data. 58 (2013) 986–992. https://doi.org/10.1021/je301279x.

[36] G. Vázquez, E. Alvarez, J.M. Navaza, R. Rendo, E. Romero, Surface tension of binary mixtures of water + monoethanolamine and water + 2-amino-2-methyl-1-propanol and tertiary mixtures of these amines with water from 25 °C to 50 °C, J. Chem. Eng. Data. 42 (1997) 57–59. https://doi.org/10.1021/je960238w.

[37] E.D. Snijder, M.J.M. Riele, G.F. Versteeg, Diffusion Coefficients of Several Aqueous Alkanolamine Solutions, (1993) 475–480. https://doi.org/10.1021/je00011a037.

[38] H.M. Kvamsdal, G.T. Rochelle, Effects of the temperature bulge in CO2 absorption from flue gas by aqueous monoethanolamine, Ind. Eng. Chem. Res. 47 (2008) 867–875. https://doi.org/10.1021/ie061651s.

[39] A. Carranza-Abaid, R.R. Wanderley, H.K. Knuutila, J.P. Jakobsen, Analysis and selection of optimal solvent-based technologies for biogas upgrading, Fuel. 303 (2021) 121327. https://doi.org/10.1016/j.fuel.2021.121327.

Chapter 8.

Concluding remarks

This chapter presents a discussion on how NNP can be utilized for improving the development of modelling frameworks. It also provides conclusions on what can be done to effectively utilize neural networks in thermodynamic modelling. Finally, a brief personal commentary what would be an ideal thermodynamic modelling AI.

8.1. Improving chemical engineering modelling with NNP

This subsection discusses some instances in which NNP and ASNNs can be utilized to tackle some modelling challenges. The discussion is centered on processes where CO_2 is absorbed by liquid solvents. The phrase “improving chemical engineering modelling” does not exclusively refer to enhance the model accuracy. But, also to the simplification and speeding up the overall modelling process.

Two alternatives for modelling processes with CO_2 capture were discussed in this thesis. The first alternative is based on equilibrium thermodynamics and heuristics (the biogas upgrading model in Chapter 6). The second alternative utilizes a model based on thermodynamics, transport phenomena, kinetics, thermophysical properties, and heuristics (model in Chapter 7). From now on, the first approach is referred to as “equilibrium-based” while “rate-based” is used for the second alternative. The rate-based allows the estimation of estimating the operational costs (OPEX) and the capital costs (CAPEX). Conversely, the equilibrium-based models can only estimate the OPEX.

Despite the limitations of the equilibrium-based approach, it could be used to map the operating conditions in which physical solvents outperform chemical solvents. From a broader perspective, the tools presented in this thesis can be utilized to determine the economic feasibility of a new solvent. For example, it is possible to establish whether the model can outperform a benchmark solvent or not. In this framework, only the activity coefficient ASNN (Chapter 4), VLE data, and the biogas upgrading simulation framework (Chapter 6) are needed. The key advantage of ASNN-based VLE models is that their development requires a short time. This is in contrast to the typical VLE

models that require speciation data and complex solution algorithms that may take a long time to develop (e.g., chemical equilibrium solvers [1–3]).

Rate-based models, usually require submodels for several thermodynamic and transport properties (e.g., viscosity, surface tension, or volume). In many cases, empirical equations are used to model these properties (e.g., [4–6]). However, they are usually unsystematically made and might not be applicable to other systems. Although the models generated by NNP are empirical models as well, a generalized structure based on an excess function can be used instead (e.g., the excess Gibbs' function proposed in Chapter 4). This would allow having a single model for every system where the only difference is the numerical value of the parameters. More accurate models than the polynomial-based models are expected due to the universal approximation theorem and the automatized optimization framework. More importantly, this task would be simplified while still maintaining a degree of coherence in the physics.

The reaction kinetics are one of the most significant modelling challenges in rate-based models. Setting aside the experimental difficulties, one of the main issues is the computational overhead. As discussed in Chapter 7, it is common to utilize the enhancement factor which, in simple terms, lumps the effect of mass transfer and kinetics into a single parameter. This workaround has been widely used since it was proposed by Hatta in 1928 [7]. Although newer and more sophisticated enhancement factor models have been proposed (e.g., the work by Gaspar and Fosbøl in 2015 [8]), it was shown by Putta et al. that they are as competitive as the simple pseudofirst order enhancement factor. Nonetheless, different enhancement factor models should be used if a primary or tertiary amine is being modelled since the speciation is different. For example, primary amines (e.g., MEA) form a carbamate while tertiary

amines (MDEA) do not [9]. Moreover, depending on the diluent, the electrolytes available in the liquid phase might be different whether it is an aqueous or a water-lean solvent [10]. Considering this, formulating a generalized ASNN architecture for predicting an enhancement factor would be highly valuable for solvent assessment.

Another advantage of using NNP for modelling kinetics is the possibility of overriding the need for modelling the diffusivity of CO_2. Since it is not feasible to measure the diffusivity of CO_2 on aqueous amines mixtures, it is customary to utilize the N_2O analogy [11] (in my opinion this analogy is not correct but it is necessary). Customizing a neural network with NNP to handle the kinetics and diffusivity modelling together may help to simplify and reduce the modelling uncertainty.

Another important area of interest is the one concerning systems in unsteady state. Physics informed neural networks (PINNs) proposed by Raissi et al. [12] can be used for that purpose. However, ASNNs can perform exact mass balances and hold thermodynamic relationships. Hence, keeping the physics coherence within the model at all times. Further research is required to apply NNP to solve differential equations since FFNNs are time invariant, therefore, Recurrent Neural Networks, would be more appropriate instead.

8.2. Conclusions

It was demonstrated that complex mathematical relationships can be explicitly represented by applying the NNP method (Chapter 3). Therefore, an ASNN with a tailor-made architecture and parameter sets can represent a broad set of physics laws. This allows the formulated ASNNs to be coherent with the physics and, hence, reliable, and trustworthy. As opposed to most machine learning modelling approaches, the NNP is a “deontological” approach. This means that it prioritizes the preservation of the physics relationships between the inputs and outputs rather than focusing only on the performance.

Analysis of the degrees of freedom (DoF) is fundamental in the development of models based on ANNs. As shown in Chapter 2, The Gibbs’ phase rule (DoF in thermodynamic systems in equilibrium) should play a key role in the selection of input variables. Ignoring the DoF can result in an accurate model that is not necessarily coherent with the physics. It was shown that ignoring the DoF in VLE modelling may result in accurate but inconsistent thermodynamic models. The consistency of a thermodynamic model is not only dependent on the Gibbs’ phase rule but also how the variables interact within the ANN. In fact, in Chapter 4 -Chapter 5, it was demonstrated that an ANN cannot be thermodynamically consistent if the inputs are not interacting properly within the ANN. Therefore, the architecture of the neural network must be built according to the definition and constraints of the thermodynamic function. This implies that both the Gibbs’ phase rule and proper understanding of the information flow is needed when developing models based on ANNs.

As discussed in Chapter 4, the role of DoF in process modelling is different than in equilibrium thermodynamics. Ignoring the DoF in process modelling does not automatically make the model incoherent with the physics. However, it sets

assumptions on what is being calculated by the model. For example, consider a thermodynamically overspecified flash tank ASNN model. This model implies that there is no thermodynamic equilibrium between the output phases. Therefore, it is assuming that other phenomena are occurring in the flash tank (e.g., mass transfer limitations or an error in the temperature reader).

The universal approximation theorem can be exploited to serve as a generator of thermodynamic functions. Utilizing the approach presented in Chapter 5 definitely simplifies and speeds-up the development time of acid gas VLE models. Moreover, it showcases a completely new perspective on how to create new thermodynamic models. Not only from the embedment thermodynamic models into neural networks, but also the possibility of exploiting the universal approximation theorem in different manners. ANNs and machine learning algorithms should improve the modelling process by removing uncertain parameters, not to entirely substitute physics laws.

By applying the VLE modelling approach proposed in Chapter 5, the techno-economic potential of solvent candidates can be simplified. The simulation framework presented in Chapter 6 can be utilized to discard solvents that are clearly uncompetitive by only utilizing VLE measurements. However, if the solvent is reasonably competitive, a rate-based gas-liquid contact model like the one presented in Chapter 7 would be required.

The NNP method has great potential for implementing models that make better use of the available data while still maintain physics coherence. Because of these two characteristics, NNP has potential to enhance models in thermodynamics, kinetics, thermophysical/transport property, and process simulations. The most remarkable advantages of NNP allows using some of the best computational technologies developed in the last century. These features are backpropagation (optimization algorithm), multilayer perceptron (the building blocks of ANNs), and automatic

differentiation (the numerical errors are essentially removed). This is the first hybrid modelling algorithm that allows using these three features while maintaining exact physics coherence in a neural network.

Two important questions might raise after finishing reading this thesis: (1) is NNP an actual machine learning method? and (2) are ASNNs artificial intelligence? Short answers: yes and no.

The first question might be raised from a computer science perspective. It might be argued that the NNP method compels and restricts the ASNN to perform the physics laws. Although this statement is true, the NNP also has potential to find the appropriate set of parameters if a suitable architecture is proposed (see Chapter 3). This implies, that the power of the NNP model relies on providing a structured and human frame of mind to the ASNN.

The above can be compared on how people learn basic mathematics at school. First, the general rules about mathematics are taught (the ASNN structure). Afterwards, the students are trained with examples and homework (learning data). This allows the teacher to provide feedback to the student (optimization algorithm). In this example, the NNP method has the same role as the teacher and ruleset. Contrarily, in the typical black box modelling with ANNs, the structure is only given by the data. Therefore, the model will exclusively be as good as the data.

Regarding the second question, typical ANNs and ASNNs are excellent at making input / output correlations but they do not possess a degree of understanding about physics (i.e., AI is absent). My main argument against regarding typical ANNs as AI is founded on their inability of properly predicting limit cases. The fact that a typical ANN cannot predict a limit case implies that it did not understand the data, it just

approximated it. Utilizing ANNs with several layers and thousands of parameters might decrease the error between the predictions and datapoints in limit cases. However, the error reduction is not because the ANN “understood” the physics. This is because whenever there are more neurons in a layer, each neuron will have a lower contribution to the prediction. Moreover, with more neurons there are higher chances to cancel the effect of each other. Furthermore, as discussed in Chapter 4 - Chapter 5, independently of how close is to the limit point there is an infinite set of operating conditions in which the model is not correct. However, it is unfair to blame the ANN when its structure is not coherent with the physics in the first place. Perhaps it is us, the users, the ones that do not understand neural networks.

Lastly, I believe that a thermodynamic AI would be considered as such if it is capable of formulating thermodynamically consistent models on its own. This implies that it would require to utilize a universal approximator function as a starting function and automatically build the ASNN architecture. In the context of the work presented in this thesis, the algorithm would read the G^E/RT function and generate the corresponding ASNN along with its parametrization (including the activity coefficient and excess enthalpies sections). Making an AI with these features seems like an impossible challenge, but who knows? Years ago, nobody would have guessed that there was going to be an AI capable of solving protein structures.

8.3. References

[1] C. Tsanas, E.H. Stenby, W. Yan, Calculation of multiphase chemical equilibrium in electrolyte solutions with non-stoichiometric methods, Fluid Phase Equilib. 482 (2019) 81–98. https://doi.org/10.1016/j.fluid.2018.10.008.

[2] C. Tsanas, E.H. Stenby, W. Yan, Calculation of simultaneous chemical and phase equilibrium by the method of Lagrange multipliers, Chem. Eng. Sci. 174 (2017) 112–126. https://doi.org/10.1016/j.ces.2017.08.033.

[3] M.L. Michelsen, J.M. Mollerup, Thermodynamic Models: Fundamentals & Computational Aspects, 2nd ed., Tie-Line Publications, 2007.

[4] S. Evjen, R. Wanderley, A. Fiksdahl, H.K. Knuutila, Viscosity, Density, and Volatility of Binary Mixtures of Imidazole, 2-Methylimidazole, 2,4,5-Trimethylimidazole, and 1,2,4,5-Tetramethylimidazole with Water, J. Chem. Eng. Data. 64 (2019) 507–516. https://doi.org/10.1021/acs.jced.8b00674.

[5] E. Skylogianni, R.R. Wanderley, S.S. Austad, H.K. Knuutila, Density and Viscosity of the Nonaqueous and Aqueous Mixtures of Methyldiethanolamine and Monoethylene Glycol at Temperatures from 283.15 to 353.15 K, J. Chem. Eng. Data. 64 (2019) 5415–5431. https://doi.org/10.1021/acs.jced.9b00607.

[6] X. Luo, A. Hartono, S. Hussain, H.F. Svendsen, Mass transfer and kinetics of carbon dioxide absorption into loaded aqueous monoethanolamine solutions, Chem. Eng. Sci. 123 (2015) 57–69. https://doi.org/10.1016/j.ces.2014.10.013.

[7] S. Hatta, Tohoku Imperial University Technical Reports, 8, 1932.

[8] J. Gaspar, P.L. Fosbøl, A general enhancement factor model for absorption and desorption systems: A CO2 capture case-study, Chem. Eng. Sci. 138 (2015) 203–215. https://doi.org/10.1016/j.ces.2015.08.023.

[9] J. Gabrielsen, M.L. Michelsen, E.H. Stenby, G.M. Kontogeorgis, A model for estimating CO2 solubility in aqueous alkanolamines, Ind. Eng. Chem. Res. 44 (2005) 3348–3354. https://doi.org/10.1021/ie048857i.

[10] R.R. Wanderley, K.K. Høisæter, H.K. Knuutila, Signs of alkylcarbonate formation in water-lean solvents: VLE-based understanding of pKa and pKs effects, Int. J. Greenh. Gas Control. 109 (2021). https://doi.org/10.1016/j.ijggc.2021.103398.

[11] G.F. Versteeg, W.P.M. van Swaal, Solubility and Diffusivity of Acid Gases (CO2, N2O) in Aqueous Alkanolamine Solutions, J. Chem. Eng. Data. 33 (1988) 29–34. https://doi.org/10.1021/je00051a011.

[12] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707. https://doi.org/10.1016/j.jcp.2018.10.045.

[13] J. Sansana, M.N. Joswiak, I. Castillo, Z. Wang, R. Rendall, L.H. Chiang, M.S. Reis, Recent trends on hybrid modeling for Industry 4.0, Comput. Chem. Eng. 151 (2021) 107365. https://doi.org/10.1016/j.compchemeng.2021.107365.

ISBN 978-82-326-6235-7 (printed ver.)
ISBN 978-82-326-5125-0 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology