NTNU

Norwegian University of
Science and Technology

Advanced Process Control at Hammerfest LNG

Developing model predictive control based on subspace and classical system identification principles for the mixed fluid cascade liquefaction process

Andreas Burheim Volden

Master of Science in Cybernetics and Robotics
Submission date: June 2016
Supervisor: Morten Hovd, ITK
Co-supervisor: Sigurd Skogestad, IKP
Eirik Heggelund, Statoil

Norwegian University of Science and Technology
Department of Engineering Cybernetics

Norwegian University of
Science and Technology

Advanced process control at Hammerfest LNG

Developing model predictive control based on subspace and classical system identification principles for the mixed fluid cascade liquefaction process

Andreas Burheim Volden

Master of Science in Cybernetics and Robotics

Submission date:	June 2016
Supervisor:	Sigurd Skogestad, IKP
Co-supervisors:	Morten Hovd, ITK
	Eirik Heggelund, Statoil ASA

Faculty of Information Technology, Mathematics and Electrical Engineering
Department of Engineering Cybernetics
Norwegian University of Science and Technology

Preface

This master thesis is written as a compulsory part of my MSc degree at the department of Engineering Cybernetics at NTNU, spring of 2016. The thesis work has been carried out under the supervision of Professor Sigurd Skogestad at the department of Chemical Engineering. Professor Morten Hovd at the department of Engineering Cybernetics has been co-supervisor. The external supervisor from Statoil ASA, has been Mr. Eirik Heggelund.

The initial motivation behind this thesis was to investigate use of model predictive control on the Statoil/Linde developed mixed fluid cascade natural gas liquefaction process by utilizing the internal Statoil MPC software - SEPTIC. Data from the process has been retrieved from a simulator developed in D-Spice. In addition, the thesis work considers MATLAB computations and OPC communication.

This experience have truly been a Bayesian stairway: the planned progress and aspects have changed somewhat during the course of this thesis as new information has surfaced and required focus. This feature, I believe, is a necessity when considering problems which comprises unexplored territories - which essentially applies for all problems. Originally, the work was planned to consider model predictive control to a larger extent, but due to challenges in obtaining a satisfactory model from system identification, the focus shifted towards the latter. Additionally, the author have the impression from research literature that the process of system identification is considered trivial. Certainly, as this thesis emphasize, this is not the case, and for complex control structures the degree of difficulty considering system identification substantially increases. Furthermore, an interesting investigation has been to compare system identification routines provided in MATLAB and SEPTIC.

To complete this thesis has been a long journey, and even though the thesis has its complete form, the reader will only be able to read what the author have *chosen* to materialize. Not mentioned further is the well of frustrations, long hours, and, certainly, new experiences and joys of learning. The fact that the thesis eventually is a result of post-rationalization which don't display dead-ends or digressions is a part of the purpose.

The thesis is primarily written for people with knowledge on control engineering - at least to some degree. Additionally, knowledge on system identification is an advantage, but not required.

Declaration of compliance

I hereby declare that this project has been an individual work in agreement with the rules and regulations of the Norwegian University of Science and Technology.

Andreas Burheim Volden	Date

Norwegian University of Science and Technology, Trondheim

Acknowledgment

The last years of my life have been somewhat unilateral. Being a student have been a truly life-changing experience in which I've devoted a lot of time in - most certainly some would say I've sacrificed a little too much on the hunt for this degree. Although that might be true, the decision to do so have been well-considered - drinking from the well of knowledge have truly been motivational and inspiring. Leaving a well paid and interesting job to the benefit of taking the red pill and exploring the rabbit hole[1] further is a decision I've never regretted. Its been a long road - not only the last 5 months which comprises the thesis work - likewise it has been a mixture of fun, hard work, frustration, but all in all it's been a fruitful journey.

A number of people deserves recognition for me being able to write this finishing last words. Above all, Sigurd Skogestad deserves a huge appreciation. His profound knowledge combined with a humble and including approach have been a constant source of inspiration and motivation. His singular abilities as supervisor have guided me steady through this thesis. Likewise, Morten Hovd deserves recognition for valuable feedback on the report. So for both of you, I truly hope that we can keep in touch after I graduate, and on some level continue with fruitful discussions - hopefully intriguing for you as well.

Hammerfest LNG represented by Eirik Heggelund have provided assistance on the LNG plant simulator, and his accommodating attitude is acknowledged. Although, no formal attachment to this thesis, Morten Fredriksen at Statoil Rotvoll have provided assistance on the SEPTIC application and related questions and discussions. He deserves considerable appreciation for being so cooperative. I certainly don't fear the future for Statoil if all employees are so forthcoming as you two are.

In addition, I want to thank my girlfriend, Mathilde, for generally being patient and still standing ground with a stubborn and quarrelsome person. Although, the stubborn and quarrelsome features are not likely to change (at least to a degree worth mentioning), but I envisage a future being less hung up with work.

- A.B.V.

[1] Yes, since this thesis deal with mathematical matrices considerably, the author take the liberty to refer to `THE MATRIX`.

"And following our will and wind we may just go where no one's been.
We'll ride the spiral to the end and may just go where no one's been."

- Maynard James Keenan's lyrics on the unprecedented masterpiece "Lateralus" by the band TOOL

Abstract

Although it principally resembles a refrigerator, production of LNG is a complex and energy demanding process. At Hammerfest LNG the Statoil/Linde developed mixed fluid cascade process liquefy natural gas from gaseous to liquid at temperatures at 110 K. From an operational perspective, the process is sensitive to disturbances and demonstrate multivariable phenomena such as interactions. Further, sea water temperature ultimately renders operation of cooling cycle compressors at their constraints. In addition, some control loops suffer from poor closed-loop control. These features outline the motivation behind this study towards more advanced process control.

The MFC process is vast and comprises a complex control structure although primarily based on single loop controllers. Thus, a necessity to render this project feasible, an initial task was to define the MPC extent in terms of MVs and CVs. By considering findings from previous work on optimization and analysis of the MFC process, the MPC extent was defined to comprise pressure control of the subcooling cycle.

Towards the development of an MPC, a large amount of labour have been invested in the model identification work. Some initial results were unsatisfactory, and required a thorough investigation as to why identified models consistently scored too low on validation data. The reason behind was found in a feature of the control structure. A split range and selector controller rendered the initial system infeasible for proper identification. Thus, the suggestion MPC extent was altered. New identification experiments returned satisfactory results. Even though, a range of analyses and comparisons were performed to ensure the proposed model's suitability.

Two MPC configurations were developed utilizing two different control systems. Simulations were run for setpoint manipulations of different CVs and disturbances. In terms of constraint handling, both MPC configurations demonstrated superior performance compared to the isolated regulatory layer. Although, there were some aspects regarding time spent to move CVs that need to be investigated further. The isolated regulatory layer outperforms the MPC applications in terms if this aspect for most of the simulations. The direct MPC application is almost consistently slower compared to the supervisory MPC. In terms of setpoint changes for CVs, the supervisory MPC shows promising tight control for accurate priorities.

Since the results are based on simulator data, one would indeed expect some differences between simulator data and plant operational data. However, the D-Spice simulator utilized in this work is not expected to possess the accuracy required to demonstrate such results. Optimization is nevertheless a difficult task for this system, and it certainly requires more rigorous modelling. Additionally, the complexity of the control system is not straightforward to handle. This may cause additional issues. In total, these factors rule out the feasibility to provide an unambiguous conclusion. Although, the work carried out in this thesis would hopefully be useful as a starting basis for further development.

Sammendrag

Selv om prosessen i prinsippet ligner et kjøleskap, er produksjon av flytende naturgass en kompleks og energikrevende prosess. På Hammerfest LNG benyttes den Statoil/Linde-utviklede blandet-væske-kaskadeprosessen(MFC) for å omgjøre naturgass i gassform til flytende ved temperaturer ved 110 K. Fra et driftsperspektiv er prosessen følsom for forstyrrelser og viser multivariable fenomener i form av interaksjoner. I tillegg er sjøvannstemperatur en variabel som ved høye temperaturer gjør at kjølesykelkompressorene drives mot sine begrensninger. I tillegg eksisterer det flere reguleringsløyfer som ikke fungerer tilfredsstillende med tilbakekobling. Disse elementene utgjør motivasjonen bak denne studien mot mer avansert prosesskontroll.

MFC-prosessen er kompleks og stor i utbredelse, og styres av en tilsvarende kompleks reguleringsstruktur først og fremst basert på enkeltsløyferegulering. For å gjøre dette prosjektet gjennomførbart var en initiell oppgave å definere omfanget av den modellprediktive regulatoren i form av kontrollerte og regulerte variable. Ved hjelp av tidligere dokumentert forskningsarbeid på optimalisering og analyse av MFC-prosessen ble MPC-regulatoren i utstrekning definert til å omfatte trykkontroll av underkjølingsyklusen.

Under utviklingen av MPC-regulatoren, har en større mengde arbeid blitt lagt ned i å identifisere modeller av systemet. De initielle resultatene var utilfredsstillende, og krevde en grundig undersøkelse på hvorfor de identifiserte modellene konsekvent returnerte lav skår på validering av data. Bakgrunnen til dette ble funnet i en varierende funksjon i reguleringsstrukturen. En "split range"-regulator i samspill med en velger gjorde identifikasjon av gode modeller umulig. For å hanskes med dette ble omfanget av MPC-regulatoren utvidet. Nye identifikasjonseksperimenter gav deretter tilfredsstillende resultater. Likevel ble en rekke analyser og sammenligninger utført for å sikre den foreslåtte modellens egnethet.

To MPC-konfigurasjoner ble utviklet for to forskjellige reguleringsstrukturer. For verifikasjon ble simuleringer hvor manipulasjoner av ulike CV-er og forstyrrelser gjennomført. Når det gjelder håndtering av beskrankninger, viste begge MPC konfigurasjoner overlegen ytelse i forhold til enkeltsløyfereguleringen. Aspekter vedrørende tidsbruk for å flytte regulerte variable var noe utilfredsstillende, ogde tte bør utredes videre. Basert på dette, isolert sett gir enkeltsløyferegulering bedre ytelse sammenlignet med begge MPC-applikasjonene. Den direkte MPC-applikasjonen er nesten konsekvent tregere i forhold til den indirekte MPC-applikasjonen. I form av settpunktsendringer for regulerte variable, viser den indirekte MPC-applikasjonen lovende regulering og hensyntar prioriteringer.

Siden resultatene er basert på en numerisk simulator, må man forvente noen forskjeller mellom simulatordata og driftsdata. D-Spice-simulatoren benyttet i dette arbeidet forventes ikke å ha den nøyaktighet som kreves for å demonstrere slike resultater. Optimalisering er uansett en vanskelig oppgave for dette systemet, og det krever mer grundig modellering. I tillegg er kompleksiteten i reguleringsstrukturen en faktor som kan føre til ytterligere utfordringer. Totalt sett utelukker

disse faktorene muligheten for å nå en entydig konklusjon. Likevel bør noe av arbeidet i denne avhandlingen forhåpentligvis være nyttig som et utgangspunkt for videre utvikling.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

PMR: Precooling Mixed Refrigerant
SMR: Subcooling Mixed Refrigerant
LMR: Liquefaction Mixed Refrigerant
PRC: Precooling Refrigeration Circuit
LRC: Liquefaction Refrigeration Circuit
SRC: Subcooling Refrigeration Circuit
LNG: Liquid Natural Gas
LPG: Liquid Petroleum Gas
MPC: Model Predictive Control
MTPA: Million Tonnes Per Annum
NGL: Natural Gas Liquids
CWHE: Coil Wound Heat Exchanger
MEG: Monoethylene Glycol
SEPTIC®: Statoil Estimation and Prediction Tool for Identification and Control
MFC®: Mixed Fluid Cascade
SISO: Single Input Single Output
MIMO: Multiple Input Multiple Output
RTO: Real Time Optimization
MMBTU: Million British Thermal Unit
bcm/a: Billion Cubic Meters Annually
PRBS: Pseudo-Random-Binary-Sequence
AMPRBS: Amplitude-Modulated-Pseudo-Random-Binary-Sequence
MOESP: Multivariable Output Error State-Space model
CVA: Canonical Variate Analysis
PEM: Prediction Error Method
SNR: Signal-to-Noise Ratio
FIR: Finite Impulse Response
ARX: AutoRegressive with eXogenous term

RGA: Relative Gain Array
PRGA: Performance Relative Gain Array
NRGA: Non-square Relative Gain Array
HIIA: Hankel Interaction Index Array

Controller Variables and Parameters

DV, d: Disturbance variable
MV, u: Manipulated variable
CV, y: Controlled variable
N: Prediction horizon
M: Control horizon

Physical Variables

Sm^3: Standard cubic metres
Nm^3: Normal cubic metres
Barg: Bar gauge pressure
Bara: Bar absolute pressure
K: degrees kelvin

Chapter 1

Introduction

In this section, a brief description of the Snøhvit natural gas field and Hammerfest LNG[2]is presented. Additionally, the background and motivation for this project is presented. Further, challenges and potential areas of improvement are briefly mentioned. Furthermore, scope of the thesis and work previously carried out is presented. In closing, an outline of the report is given and some relevant research literature is listed.

1.1 Hammerfest LNG and some background material

Natural gas(NG) is a mixture of fluids containing primarily hydro carbon gases. It is the cleanest fossil fuel with the lowest carbon dioxide emissions, and it is colorless and odorless in its pure form. Natural gas is an important fuel source as well as a major feedstock for fertilizers and petrochemicals.

Liquefaction of natural gas is an established mean to lower transport costs over great distances. By liquefying natural gas the gas is compressed to 1/600th of its normal volume in gaseous form, which increase transport efficiency significantly. The liquefaction process is a cryogenic process, which indicate operating temperatures below 123 K(Venkatarathnam [86]), and the LNG product is about 113 K in its final form at atmospheric pressure. From figure 1.1, which depicts transport costs of natural gas in pipelines and LNG, it is demonstrated that transport above 5000 kilometres favours LNG with respect to cost(Schwimmbeck [63]).

[2]The author use the designation Hammerfest LNG throughout this report. Though other interchangeable designations have been used in various literature; Melkøya LNG and Snøhvit LNG, they all indicate the same. LNG is an abbreviation for Liqufied Natural Gas.

For readers interested in a less technical perspective on Hammerfest LNG and Snøhvit, the author recommends reading the Norwegian book "Snøhvit: Historien om olje og gass i Barentshavet" by the writer Alf R. Jacobsen.

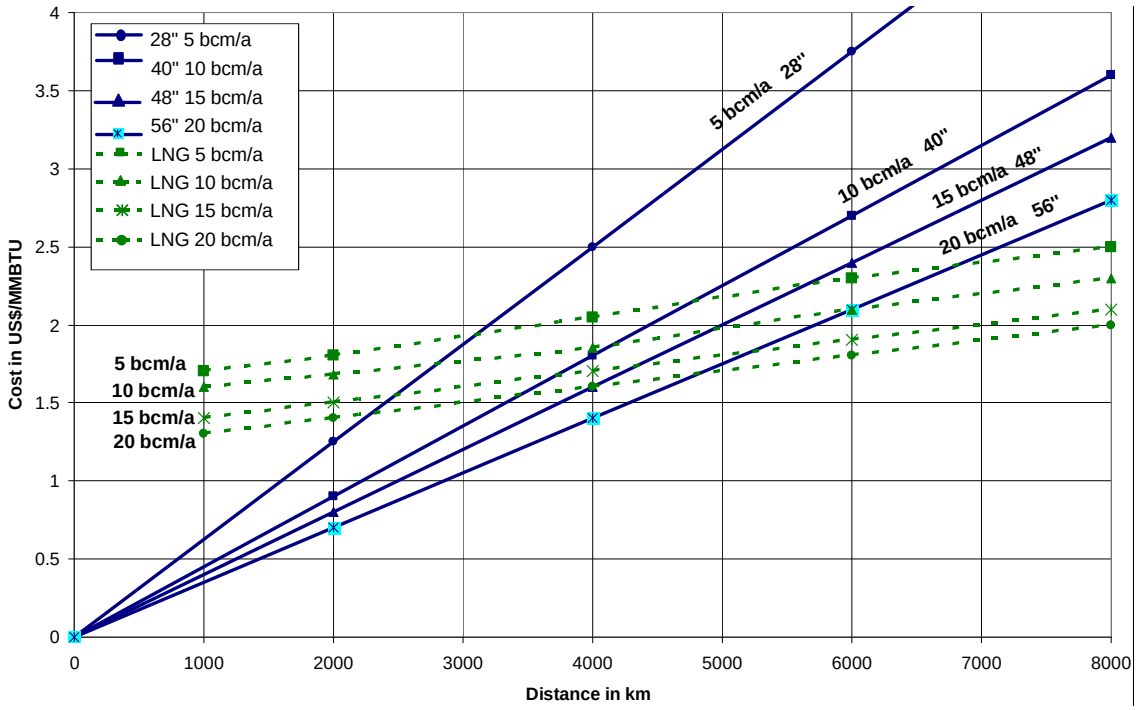


Figure 1.1 Comparison of transport cost of LNG and NG in terms of volume and pipeline dimensions. From Schwimmbeck [63].

The Snøhvit natural gas field was discovered by Statoil exploration in late 1984, and by mid 2002 the project development plan for Hammerfest LNG was approved in the Norwegian parliament. Official start up date of the plant was 21th of August 2007(Statoil [78]). Hammerfest LNG is a Statoil operated LNG processing plant owned by Statoil(37%), Petoro(30%), Total(18%), GDF SUEZ(12%), and RWE Dea Norge(2.8%)(Norwegian Petroleum Directorate [54]).

(a) Geographical location.

(b) Hammerfest LNG.

Figure 1.2 Location of the Snøhvit field(a) and overview of the LNG plant(b). From Statoil [77] and Teknisk Ukeblad [83], respectively.

Location of the plant is outside the city of Hammerfest in the northernmost part of Norway. It is

the first development in the Barents sea area and the first installation on the Norwegian continental shelf without surface installations offshore. Because of this, Hammerfest LNG is considered a milestone in the Norwegian oil history. By comparison, the plant is one of the worlds most efficient LNG producing plants, primarily due to design and location(Schmidt et al. [62]). The arctic climate plays a major role and to a high degree, it affects several operating constraints which influence both efficiency and plant throughput.

Some 140 kilometers north-west for the plant, the subsea templates are located on top of wells which provides natural gas from three fields, Snøhvit, Askeladd and Albatross. Snøhvit and Albatross started production in 2007, while Askeladd was set for production in 2015(Statoil [78]). The sea depth varies from 250 to 350 meters at the production templates, while the reservoir depth is about 2000 meters. In addition to producing wells, there exists a CO_2 reinjection well for carbon storage(Statoil [75]).

The LNG plant is dimensioned for production with basis in a feed gas rate of $2.08 \cdot 10^7 Sm_3$ every day(Statoil [74]. Today the plant has an annual export of 4.3 MTPA LNG, 3.1-5.7 million barrels of condensate and 150 000 - 250 000 tonnes of liquid petroleum gas(LPG). At ordinary operation the plant power consumption is 230 MW and the energy demand to run the entire plant is covered by a fuel consumption of less than 6% of the feed flow(Skjerven and Vist [66]).

The well output is a mixture of natural gas, carbon dioxide, condensate and natural gas liquids(NGL), and is led through a multi phase flow pipe back to the plant where each substance is separated and individually processed. During the processing, CO_2 is removed and sent to the injection well. Additionally, mercury, nitrogen and alkanes are separated from the feed gas prior to liquefaction. After separation, the ambient temperature natural gas is led through two fractionation columns, an expansion turbine and a series of cascade heat exchangers. These operations essentially make up the liquefaction system. The end LNG product is a purified, liquefied natural gas at temperatures about 113 K and pressures at 0.2-4 Bar gage pressure(Barg). Some key LNG product specifications are found in table 1.1. After processing, three main products; LNG, LPG and condensate is stored and later shipped to the market(Norwegian Petroleum Directorate [54]).

Table 1.1 Selection of some vital LNG product specifications. Several additional elements are expected to comprise the LNG product, but methane is obviously the crucial one. Thus, methane is the only component subject for a minimal constrained LNG composition value. From Statoil [74].

LNG composition specifications		
Component	**Min. value**	**Max. value**
Methane	84.55 mol%	100 mol%
Nitrogen	N/A	1.1 mol%
Oxygen	N/A	0.008 mol%
Carbon dioxide	N/A	0.008 mol%
Mercury	N/A	10 ng/Sm^3

1.2 Motivation

This thesis is a continuation the work carried out in Volden [88], where a preliminary study of the liquefaction system at Hamemrfest LNG was conudcted. The overall goal has been to develop an MPC controller to gain more robust and less energy consuming operation of the subcooling system. Today the process is SISO controlled with extensive use of regular feedback, cascade, selective and split range controlled loops without any advanced control implemented on top.

The LNG plant has been operational for almost 10 years, and even though studies have been conducted, optimization based control is utilized only to a small degree at the plant. Upon initialization of production, a study of advanced process control for the liquefaction system was conducted. However, this study rendered the system immature at that time. This project will focus on development of advanced process control at the plant, and to narrow the scope of the project, it has been decided to further investigate the subcooling cycle in the liquefaction system. Instead of utilizing real plant data for the MPC development, a simulator plant model developed in D-Spice[3] have been used.

To develop and utilize an MPC application is not unknown in Statoil. Since the end of the 1990's, the company have developed an in-house software for model identification and MPC applications(SEPTIC)Skofteland et al. [67]. This software have been used extensively in conventional downstream operations in Statoil. According to Strand [79], by 2011 there were approximately 80 SEPTIC applications running within Statoil and 80% av these were either installed on Kårstø, Mongstad or Kalundborg, which all are established refineries and gas processing plants. This emphasize the unrealized potential at other plant locations, e.g. Hammerfest LNG.

The general understanding of MPC compared to simpler control strategies is unambiguous, and the superior performance for MIMO processes is proven in numerous research and industry applications. The process performance, controller disturbance rejection and process energy consumption are variables that in general are improved by using an MPC application. See for example Morari and Lee [52] and Qin and Badgwell [58]. These properties have helped the MPC gain popularity over the last decades. The advantages of MPC have expanded its area of usefulnesses. From initially being used in chemical processes, the range of MPC applications is increasing over a vast range of industries.

For several available software systems there have been reported a payback time of reasonable short time, i.e., 3 months as stated in Bassett and van Wijck [5] or 2-4 months as stated in Strand and Sagli [80]. The global competition is rapidly increasing in terms of pricing and quality along with higher cost of energy. Hence, the industry is forced to search for optimal solutions and thus reduce cost to stay competitive(Glandt et al. [24]). Optimization in the economical sense is now increasingly being conducted solely in steady-state processes by a separate overlying controller. This is referred to as the real time optimization(RTO). This enables optimizing energy consumption towards stable operation(Downs and Skogestad [14]).

[3]Information on D-Spice are found in Fantoft [16] and Fantoft [15].

In addition, by utilizing the in-house MPC software and on site simulator models, one can expect to minimize support costs. The software enables to build a network of MPC applications which are interdependent. This expands the control hierarchy as we know it today, where we in general place the MPC in the supervisory layer.

To build the D-Spice simulator which is primarily used for operator training and system testing, Statoil have divided the simulator model into 10 sections, where one computer runs one section. By using ten computers and thus dividing workload, the simulator manages to run at desired speeds above real time. Communication is done by using OPC and TCP/IP protocols(Skjerven and Vist [66]).

1.3 Objectives

At Hammerfest LNG large volumes of natural gas are processed continuously. Liquefying natural gas from ambient temperatures about 286 K to about 113 K requires significant amounts of energy. In addition, the system is sensitive to disturbances from of varying feedstock impurity and rates, external and internal temperatures, and system interactions. To better handle disturbances and simultaneously operate the subcooling system more energy efficient, development and testing of an MPC application has been suggested. An MPC would additionally require less interaction from operators, thus enabling operators to devote time to other tasks. It is this author's impression that operations and maintenance have been granted significant attention during the plants lifetime up until now, while research, development and modifications have gained little attention. This project aims to concentrate on the latter while focusing on the subcooling cycle.

The objective of this work is to develop an application such that performance and product specifications are kept within their respective limits while minimizing energy consumption. To obtain a feasible MPC configuration, several aspects must undergo investigation. These conditions make out the ruling research objectives for this project work:

- Define scope and extent of MPC application
- Identify model and disturbances
- Identify operational conditions and constraints
- Identify and investigation on optimal selection of available measurements, controlled variables and manipulated variables
- Optimization problem formulation for the MPC application
- Simulations of MPC connected to dynamical simulator
- (Suggestion of MPC applications in other parts of liquefaction system)

During the work on this thesis, some of the above mentioned aspects have been thoroughly investigated. Based on the author's academical interest and by supervisor suggestion, the facet of obtaining a process model have been a major part of this work. This is reflected by the thesis content.

1.3.1 Limitations

Limitations pertain to this work, and by limitations we do not consider aspects beyond the scope. Limitations refer to e.g. uncertainties in results and so forth. This section aims to inform the reader in this respect.

The obtained results in terms of computations and simulations is based solely on numerical solvers and thus choice of solver plays a significant role. Additionally, the simulator model is developed by Statoil and runs in D-Spice which relies on numerical solvers and thermodynamic tables as well. In reality the results may deviate due to obvious differences between the numerical domain and the true operation.

In addition, simulator behaviour compared to true plant behaviour is affected by differing time delays between the two domains. The time delay in a running simulator is affected by certain factors, while the factors affecting plant time delays are incomparable. Hence, there will be differences in running parameters which ultimately will lead to somewhat deviating behaviour between the two domains, due to the different effective time delays.

Towards the end of this work, a discrepancy in the D-Spice model was detected. This involve the part of the MFC process which interacts with the fractionation system. Sections 2.1 and 7.1 provide a description of said interaction. As mentioned in above section 1.2, the D-Spice model is divided in sub-models for the intention to run sections of the modelled LNG plant individually and distribute computational load. However, when running the MFC section individually, interaction with the fractionation system becomes faulty as perturbing the MFC sub-model leads to operation that slowly drifts from equilibrium. This ultimately forces a model restart, but takes considerable time. To avoid said feature the solution is to run the entire plant model, but this increase computational load significantly and the simulator is not able to run above speeds twice real time. However, this solution was not considered further and all results in the theses are obtained while running the MFC section individually.

Further, the simulator model may, on a detailed level, deviate from the plant configuration. Modifications to the control system, altering of measurements, and parameter adjustments may occur without the author's knowledge. Thus, any results from this study may not be directly applicable to the control configuration of the plant.

1.4 Scope and outline of the report

As specified more in detail in section 1.3, the aim for this report is to investigate and formulate an MPC problem, and test the application. Intention of the work carried out in the thesis is to provide a controller which render the operation of subcooling cycle more energy efficient and accurate while simultaneously keeping within operational limits and ensuring robust performance. By increasing

energy efficiency, the plant throughput increases for the same amount of fuel, and thus energy costs are minimized.

Chapter 2 provides a superficial description of Hammerfest LNG plant including its major systems. More in detail, the mixed fluid cascade process and the subcooling cycle is presented. This includes presenting main components and process objectives. The basic thermodynamics and MPC preliminaries are covered in Volden [88], and will not be considered in this thesis. Further, the control system and its structure is briefly described including degrees of freedom and constraints.

Based the work of obtaining a decent process model, some general theory on system identification is provided in chapter 3. This includes a particular focus on subspace identification. Since the identification experiments are executed in open-loop configurations, this chapter discusses features of closed-loop identification as well.

In chapter 4, the author describe a potential pitfall when considering system identification and verification of experiment data. This chapter is based on a first-hand experience. Further, the continuation work of obtaining a satisfactory model is presented. A number of different approaches are tested and results are presented.

Chapter 5 presents several measures for model validation and analysis. Based on validation results, several candidate models are further analysed based on different aspects such as model order, signal-to-noise ratio and statistical measures.

In chapter 6, a description of the MPC software, and some important features are introduced. Additionally, this includes classical system identification results obtained using the internal MPC software for system identification.

Chapter 7, provides an outline of the MPC development. Further, some arguments behind the defined MPC extent is specified. Additionally, results from simulations including the MPC are depicted.

In chapter 8 the author summarizes the results from the work and discusses different aspects around the obtained results. Chapter 9 list main conclusions and provides suggestions for further work.

1.5 Earlier work

This section provides an overview of the literature utilized in this work. Additionally, chapters and sections based on the preliminary project(Volden [88]) is listed.

Table 1.2 Sections in this thesis based on previous project

Section in thesis	Based on section in Volden [88]	To which degree in %
Chapter 1	Chapter 1	60
Chapter 2	Chapter 3, 4.2, 4.3, 4.4	75
Appendix D	Chapter 5.2	95

The project leading up to this thesis was carried out in Volden [88]. The project considered theoretical preliminaries regarding thermodynamics of heat transfer, compression and expansion, and model predictive control. In addition, an introduction to the D-Spice simulator and OPC communication was provided. Furthermore, a complete description of the MFC process was given, and an analysis of present control structure and comprising control loops was carried out. Additionally, some initial work on system identification was executed. The conclusion rendered the system identification results not satisfactory, thus a large part of this work considers further system identification in order to analyse and obtain satisfactory models for MPC development.

In terms of research literature utilized in the thesis work, a challenge has been to retrieve literature including MPC development for LNG production. In Jensen [39] the author provides information on refrigeration cycles and operational analyses. In addition, the author has investigated the MFC process in terms of optimal operation and degrees of freedom for operation. However, little research on control in that sense is provided. Heldt [28] provides a somewhat similar basis for his work. Thorough analyses on several LNG processes are given including the MFC process. Still, the perspective is from an operational point of view, and additionally the MFC process considered is somewhat unsimilar to the one studied in this work. In Sturm et al. [81], the authors present an MPC development by Shell for the C3-MR LNG process. However, this paper contain very limited information and does not provide any guidelines which may have been useful for this thesis. Regarding analyses for optimal operation and selection of self-optimizing variables for an LNG process in addition to Jensen [39], the papers Michelsen et al. [51] and Michelsen et al. [50] have been useful, but to a limited degree since they consider a different LNG process.

Chapter 2

Process description and current control system

This chapter aims to describe the natural gas liquefaction process at Hammerfest LNG. Initially, an overall plant description including major systems at Hammerfest LNG are briefly described. Further, the subcooling cycle and its control system is described more in detail. In addition, a degrees of freedom analysis of the subcooling cycle is included. For other sections of the MFC process, the reader is advised to consult chapter 3 in Volden [88]. Unless otherwise specified, this chapter is based on Volden [88] and Statoil [72, 73, 74, 76].

2.1 Process description of Hammerfest LNG

To state the obvious and what the reader will gradually comprehend reading this thesis and this chapter, particularly: liquefaction of natural gas is a complex and energy demanding process. At the Hammerfest plant production of LNG is realized through a wide scope of processes, including utility and support systems. Figure 2.1 depicts how the LNG plant and interfacing systems are related, and includes both upstream and downstream facilities to the battery limit at the loading docks.

Between the subsea equipment and the onshore plant, a pipeline is installed for landing the producing stream from the wells. The pipeline contain a multiphase flow composed of an aqueous phase consisting of water and monoethylene glycol(MEG), a condensate phase(oil) and a gaseous phase. The pipeline stream is cooled along the pipeline, and due to decrease in temperature, gas will condense along the pipeline prior to landing. After landing onshore, the pipeline flow is led to the slug catcher(see figure 2.1).

The slug catcher, which principally is a buffer volume between the pipeline and plant, smooths variations in flow and pressure, and provides a simple separation between water, gas and condensate.

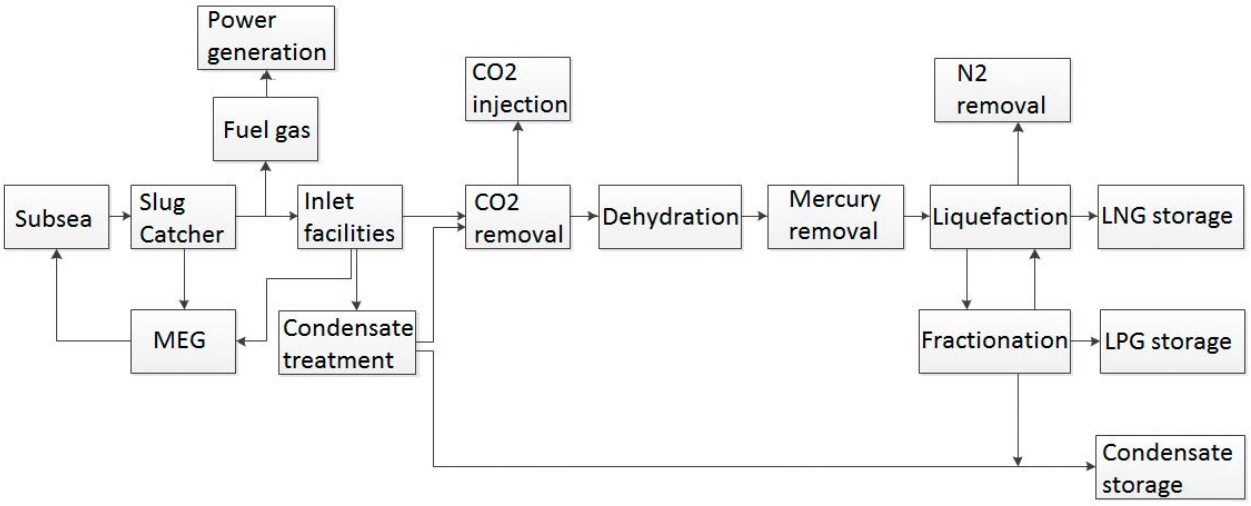


Figure 2.1 Plant system overview at Hammerfest LNG

Additionally, condensate slugs are separated. Upon leaving the slug catcher, the gas is sent to the inlet facilities where it is pressurized, and a second separation to further remove condensate takes place. Additionally, sand and similar particles are removed, and subsequently, the gas is then sent to gas conditioning. Gas condition comprises three separate processes; CO_2 removal, dehydration and mercury removal. These operations prevent corrosion, freezing, hydration and blocking of the producing stream downstream. An activated amine wash unit is used to remove CO_2, while the water is separated in an absorber station. The CO_2 from the wash unit is then re-compressed and re-injected subsea to maintain the reservoir producing pressure. To prevent hydration and slugging upstream of the slug catcher, MEG is added to the producing fluid at the subsea templates prior to entering the landing pipeline.

The condensate treatment system receives streams from several parts of the plant. Downstream of the slug catcher and inlet facilities, the condensate enters a separator, where gas and the water-MEG mixtures are removed. Subsequently, light components are removed from the condensate in a distillation column and transferred to gas processing downstream. The aqueous phase from the slug catcher and condensate separator is sent to MEG recycling, where particles, salt and most of the water is removed. Recycled MEG is stored and eventually pumped offshore to be re-injected into the well stream subsea.

From gas conditioning, the gas is sent to the first sub-process of liquefaction; precooling and heavy hydrocarbon fractionation. This section comprises the first part of the Statoil/Linde developed mixed fluid cascade liquefaction process(MFC). Initially, the ambient temperature gas enters the precooler, which comprises the first cooling cycle in the cascade. Here the gas is cooled before being led to fractionation, where heavier natural gas liquids(NGL), such as propane and butane (C3 and C4) are separated and condensed to liquefied petroleum gas(LPG). Additionally, these components are utilized for refrigerant make-up for the three cooling cycles in the MFC process; precooling, liquefaction and subcooling. Refrigerant make-up is a necessity to compensate for the

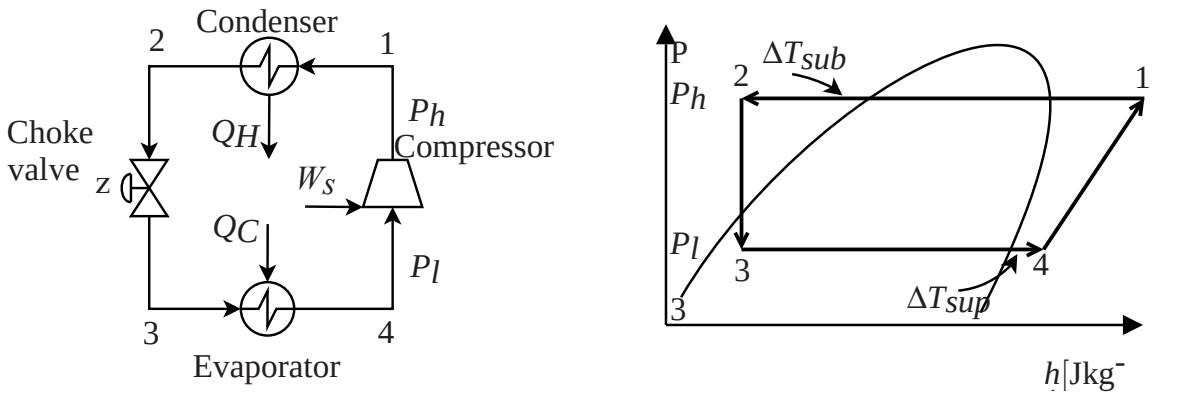
losses through in the cooling cycles, and to render cooling refrigerant composition adjustments feasible. After fractionation and LPG production, the LPG is temporarily stored in a cryogenic tank from where it is loaded onto tankers. In the same way, condensate is stored in a tank at atmospheric pressure from where it is loaded onto shuttle tankers.

Upon leaving the heavy hydrocarbon fractionation column(HHC) the natural gas is led to the liquefaction and subcoolong cycles. Since the LNG contains above specification of nitrogen subsequent of liquefaction and subcooling, the feed gas is led to a splitting column for nitrogen removal. After liquefaction and nitrogen removal, the product LNG has a proper composition and temperature of about 110 K at slightly above atmospheric pressure, and the purified LNG is then pumped to storage tanks. The gas formed during loading onto shuttle tankers which comprises displacement, flash and boil-off gas, is returned to the onshore tanks, from where it is re-compressed jointly with the boil-off gas from the land LNG tanks.

As stated in section 1.2, the liquefaction system is the core process of the plant and accounts for a major influence on the plants efficiency and energy consumption. Volden [88] lists several design aspects in terms of efficiency and bounds on capacity. Additionally, a more complete description of the system in entirety, and additional information on relevant thermodynamics are found in Volden [88] and references therein. Thus, only an introduction to the MFC process is given here.

2.2 The mixed fluid cascade process for liquefaction of natural gas

The main tasks of the MFC process is to liquefy and subcool the feed gas, in addition to remove heavy hydrocarbon and nitrogen fractions prior to storage. Composition adjustments of the LNG are performed in the HHC and nitrogen removal columns. The refrigeration duty for precooling, liquefaction and subcooling of the feed gas is provided by three cascaded mixed refrigerant cycles. An electrical driven refrigerant compressor is dedicated to each cycle. Refrigeration duty is also supplied to other consumer systems in the plant. Principally, one might say that each refrigerant cycle resembles a refrigerator; the cycle comprises compression, condensing, expansion and evaporation. A principal refrigeration cycle with corresponding phase diagram is depicted in figure 2.2.



(a) Principal refrigeration cycle

(b) Phase diagram for refrigeration cycle. Subcooling and superheating are indicated.

Figure 2.2 Principal refrigeration cycle and phase diagram. From Jensen [39].

As depicted in above figure, the refrigeration cycle has four states(1-4). The energy input is the compressor work(Ws) that increases the pressure and temperature of the vapour entering the compressor($4 \rightarrow 1$). After compression the high pressure vapour is saturated, condensed and further subcooled($1 \rightarrow 2$) in a condenser. The gas is then liquid. To decrease pressure, the liquid is expanded through a choke valve to a resulting two-phased low temperature vapour($2 \rightarrow 3$). Further, heat is added in the evaporator($3 \rightarrow 4$), thus making the liquid gaseous and additionally, super heating takes place before leaving the evaporator and entering the compressor suction side.

The gas is led through three cascaded cooling cycles in the MFC process. The three cooling cycles are labeled precooling refrigerant cycle(PRC), liquefaction refrigerant cycle(LRC) and subcooling refrigerant cycle(SRC). This includes their respective mixed refrigerants; precooling mixed refrigerant(PMR), liquefaction mixed refrigerant(LMR) and subcooling mixed refrigerant(SMR). Upon leaving the gas conditioning sections, the feed gas enters the MFC process at high pressure and ambient temperature.

Figure 2.3 provides a principal outline of the mixed fluid cascade process.

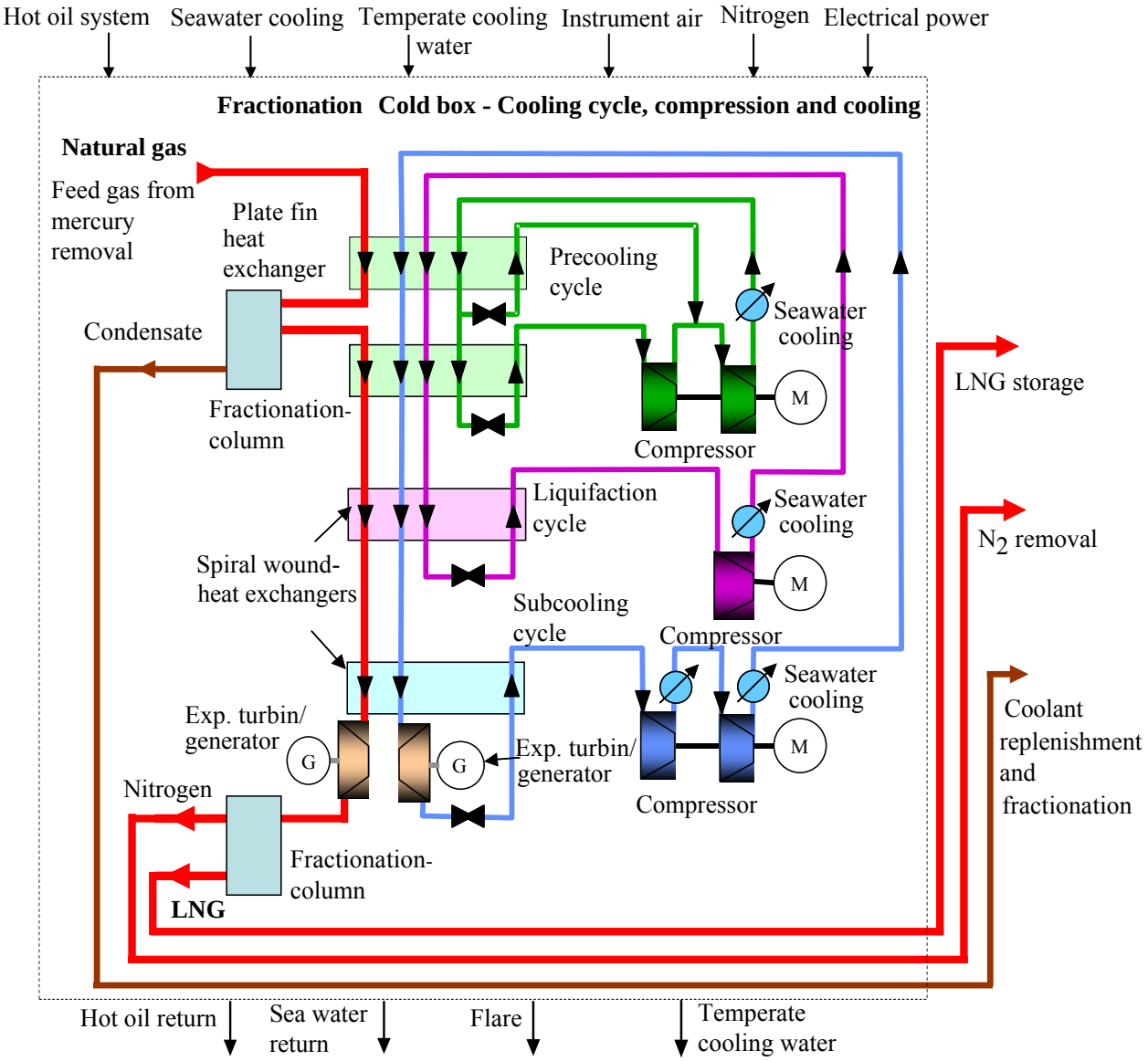


Figure 2.3 Principal overview of system 25 and utility systems at Hammerfest LNG. Modified from Statoil [74].

2.2.1 Subcooling refrigeration cycle

The subcooling refrigeration circuit(SRC) utilizes a subcooling mixed refrigerant(SMR) for cooling duty. In reading this section, figure 2.5 depicts the subcooling cycle in a somewhat detailed fashion. Additionally, to understand the process principle figures 2.4 and 2.6 may be useful.

Upon leaving the liquefier the feed gas has a temperature of about 197 K, and subsequently enters the subcooler, which is a coil-wound heat exchanger(CWHE). When leaving the subcooler, the gas temperature is lowered to the neighborhood of 113 K, while the pressure is kept relatively constant

from leaving the gas conditioning sections and until the LNG is pumped to storage.

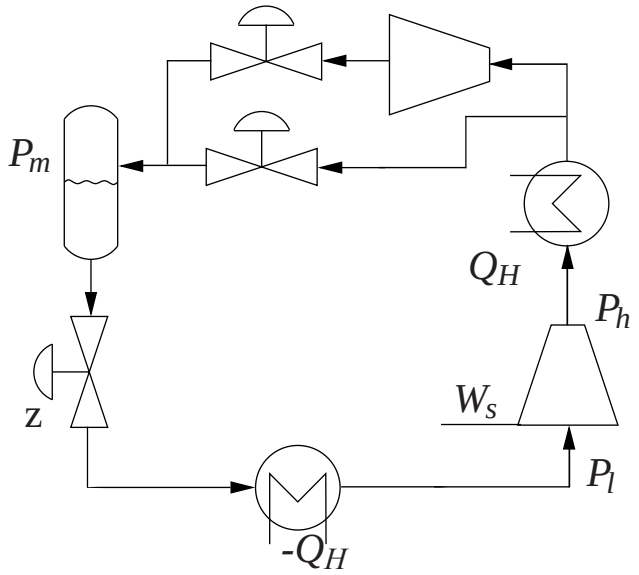


Figure 2.4 Principal sketch of the subcooling cycle. Modified from Jensen [39].

Liquid SMR is stored in the subcooling cycle tank at low pressure and temperature. The tank is pressure controlled via a bypass opening to the shellside of the subcooler. In the case of low pressure it is controlled by directing hot gas downstream of the compressor aftercooler into the tank. In the case of high pressure, gas from the tank can be directed to flaring.

The liquid SMR from the tank is led to the shell-side top of the subcooler. As the liquid SMR flows down the subcooler shell-side, it evaporates due to heat exchange. At the shell-side bottom gaseous SMR exits at a higher temperature, and then enters the subcooling cycle compressor. The low pressure SMR gas from the subcooler enters the subcooling cycle compressor 1st stage. From 1st stage compressed gas is led to an intercooler before entering the 2nd compressor stage. Leaving the compressor, the gas temperature and pressure is substantially increased. Subsequently leaving the compressor, the compressed gas is cooled in a aftercooler against seawater.

The SMR is then distributed to the precooling cycle condenser. As the gas flows through the condenser it is partially condensed. It leaves the bottom of the condenser as two-phase at a low temperature, and is led to the liquefaction cycle. As the SMR passes through liquefier it is fully condensed. The liquid SMR then leaves the liquefier at a lower temperature.

The liquid SMR enters the bottom of the subcooler. As the SMR moves up through the subcooler it is subcooled against the liquid SMR falling and vapoursing over the tubes on the shellside. The subcooled liquid SMR leaves at the top of the exchanger with a lowered temperature and high pressure.

After leaving the subcooler, the SMR stream is de-pressurized. This is achieved in two ways; if the plant load is low the liquid stream is flashed over an expansion valve. At higher load rates, the

process stream is directed through an expansion turbine. The discharge of the turbine is pressure controlled in order to avoid flashing and thus prevent damage to the turbine by vapour generation. The final pressure let down after the turbine is done via a expansion valve. By converting the kinetic energy of the turbine, the subcooling stream pressure reduction runs an electrical generator. Subsequently, the SMR stream is then led back to the SRC tank.

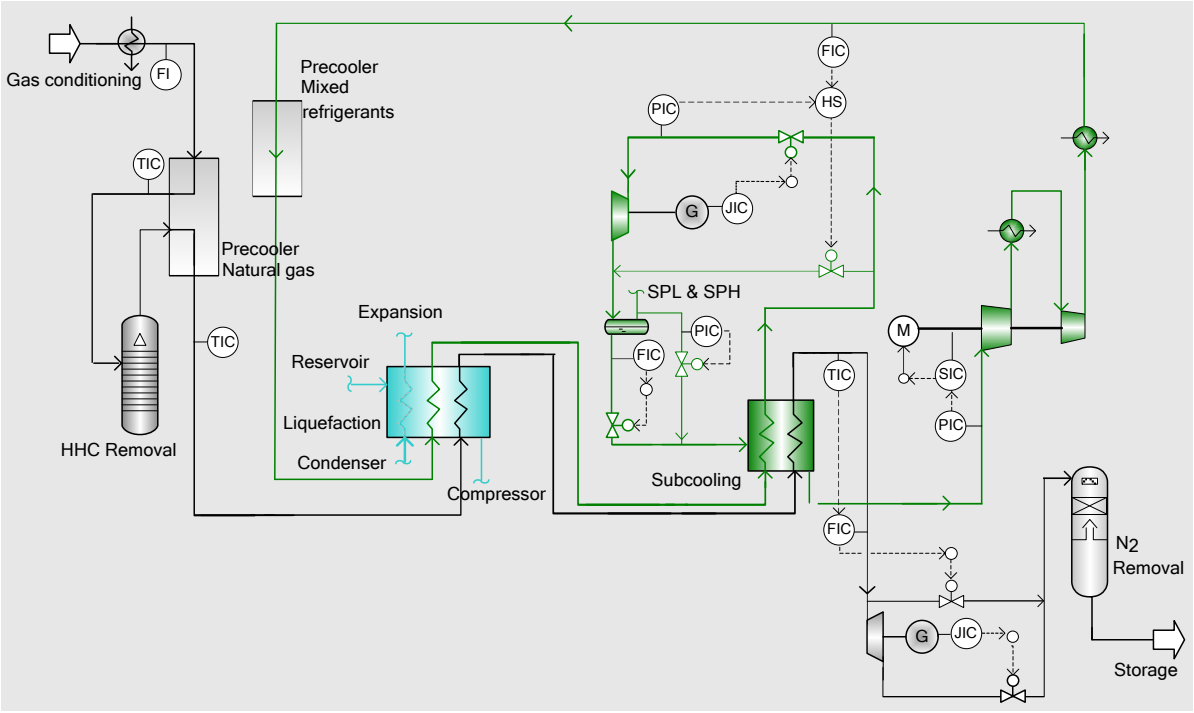
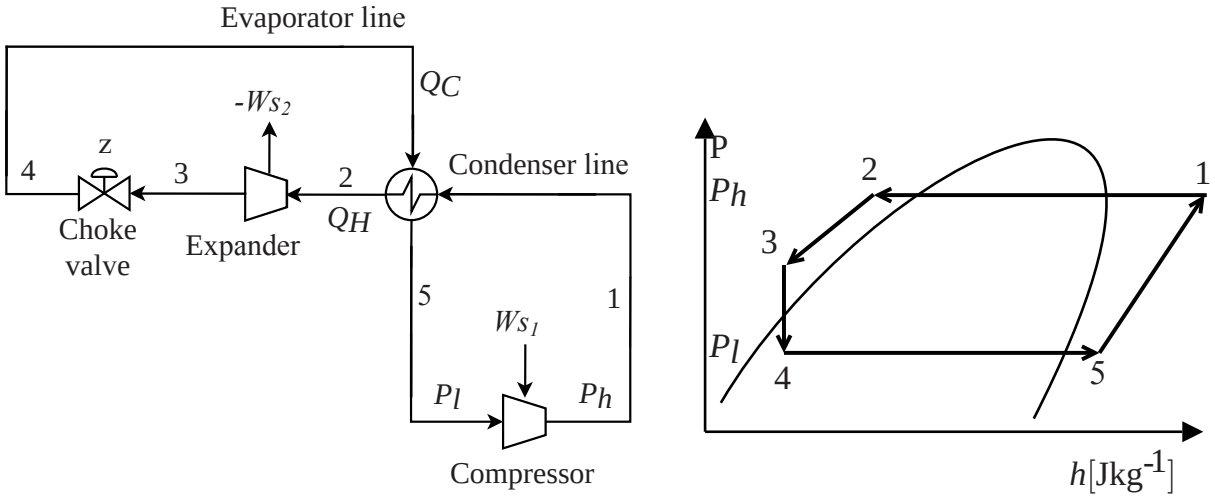


Figure 2.5 Subcooling section overview in system 25

In light of this information, the subcooling cycle does not properly resemble the refrigerator as depicted in figure 2.2. Thus, a more precise conceptual depiction is the figure 2.4. To further illustrate the use of subcooling refrigerant for cooling and heating purposes, in addition to a corresponding phase diagram, the figure 2.6 provides a more accurate description.



(a) Principal sketch of the subcooling cycle

(b) Phase diagram for the subcooling cycle.

Figure 2.6 Principal refrigeration cycle and phase diagram. Modified from Jensen [39].

2.3 Operation and degrees of freedom for the subcooling cycle

In Volden [88], the author performed an analysis on the current control structure of the MFC system. Additionally, an investigation on potential and actual degrees of freedom was executed.

2.3.1 Constraints

Below lists provide some information on relevant constraints for the MFC process. The lists are essentially taken from Volden [88].

Potential active constraints

- When considering optimal operation, in general, two objectives for optimality are considered; maximize profit or minimize cost. For the LNG system this translates to maximizing throughput or to minimize compressor work. Either objective implies that the product conditions, i.e., the LNG product pressure and temperature must be controlled fast and tight, and that seawater cooling is maximized, reducing degrees of freedom by 7. The production rate can be controlled much slower(to a certain degree) as the LNG is not directly delivered but stored in tanks.
- For the objective of maximizing profit, the subcooling compressor load become an active constraint and we loose a degree of freedom. This could potentially be the case for the

precooling and liquefaction compressor if the load is distributed evenly and plant throughput is further increased. To minimize the cost, we consider the feed as given, but in reality this is a manipulated variable.
- To allow for optimal operation, maintaining the inlet pressure to the compressors at the lower bound is necessary. This implies keeping the inlet pressure at its bound, and the constraints thus become active. This potentially reduces degrees of freedom by 3 for the objectives of both minimizing work and maximizing throughput.

Operational constraints

- For some LNG processes one need to ensure that no liquid enters the compressors. This is usually solved by applying a some degree of superheating of the vapour exiting heat exchangers. For the MFC process studied here, this is not an operational constraint as any liquid exiting the heat exchangers are vaporized using a knockout drum. However, this feature is not regarded for optimal operation since the knockout drum employs the compressor discharge for vaporizing liquid.
- The temperature of LNG is vital and cannot under any circumstances increase above the upper temperature limit.
- The operating pressure must stay within bounds. In particular for the compressors.
- Compressor surge and stonewall.
- Fraction of LNG after subcooling must be 100% and within product specification.
- Levels in separators and tanks, and valves are hard constraints.
- Compressor capacities and heat exchanger duties.
- Feed rates and production limits.

As mentioned in section 1.5, there exist a lot of research on model predictive control in various applications. Additionally, cooling or heating applications e.g. as in heat exchanger networks are also known research topics. However, finding literature on model predictive control for LNG production returned few hits, and the only similar application is superficially described in Sturm et al. [81]. Figure 2.7 depicts the constraints applicable for the C3-MR LNG process, and thus define a feasible operating window. Although there are some dissimilarities between the MFC and C3-MR processes, they do share a number of constraints.

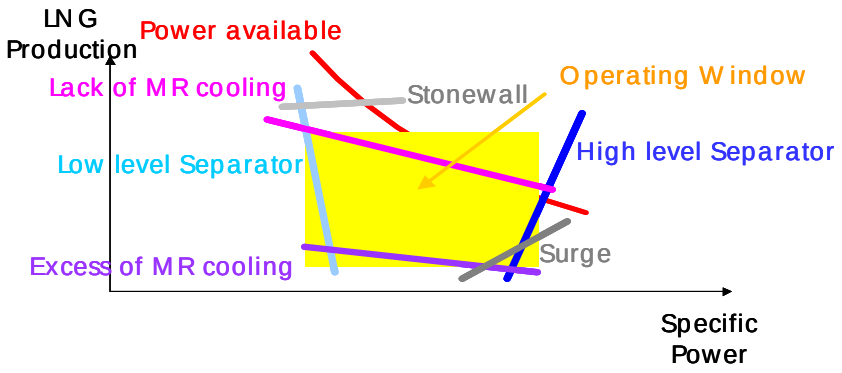


Figure 2.7 Operational window for C3-MR LNG process. From Sturm et al. [81].

2.3.2 Degrees of freedom

In chapter 4 of Volden [88] functional descriptions and degrees of freedom analyses are given. Figure 2.8 depicts actual degrees of freedom for the MFC system. Since only the subcooling cycle is investigated in this work, the degrees of freedom related to other parts of the system are not considered here.

Figure 2.8 depicts the found degrees of freedom on the control system. This provides a more detailed perspective in terms of control and how the degrees of freedom are related. The table 2.1 provide additional information on the degrees of freedom for control.

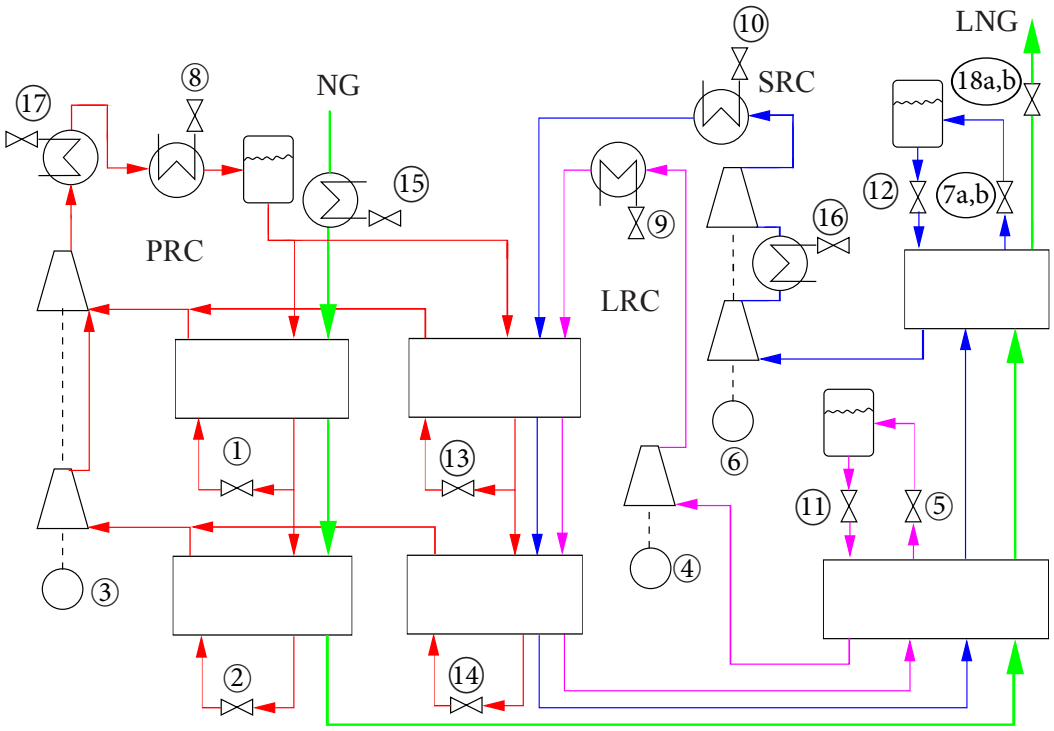


Figure 2.8 Recap of actual degrees of freedom for the MFC process. From Volden [88].

Related to figure 2.8, the table 2.1 which is derived from table 4.3 in Volden [88], lists relevant degrees of freedom for the subcooling cycle. The numbering correspond to figure 2.8. The table list primary and secondary variables. This definition is related to when the variables become constraints, and a further explanation on this matter is provided in section 4.4 in Volden [88].

The primary plant throughput is set by adjusting the subcooling refrigerant flow through the subcooler. This variable set the plant load as it indirectly set the LNG production rate by manipulating the LNG temperature. Large volumes of LNG flow through the system and the production temperature specification have small tolerances. Thus, this is an obviously vital variable to control within tight bounds.

Additionally, accurate pressure in the subcooling refrigerant tank is important to maintain as fluctuation in the tank pressure may affect the flow of refrigerant downstream the tank. Low tank pressure requires non-optimal compressor operation since hot gas is diverted downstream the compressor and into the tank to increase pressure. If tank pressure becomes too high, the pressure is reduced by flaring which obviously is not optimal.

Table 2.1 Degrees of freedom and their control purposes in current control structure for the subcooling cycle. The table is based on table 4.3 in Volden [88], but is altered to some extent.

Degrees of freedom and their primary and secondary control variables for the subcooling cycle	
7a: Bypass valve for SMR upstream tank	**Primary variable:** SMR pressure upstream expansion turbine. **Secondary variable:** N/A.
7b: Choke valve for SMR upstream tank	**Primary variable:** SMR pressure upstream expansion turbine. **Secondary variables:** Low pressure downstream expansion turbine.
10: Aftercooling of SMR	**Primary variable:** N/A. Active constraint.
12: Choke valve for SMR cycle	**Primary variable:** SMR flow entering subcooler. This is identified as the primary throughput manipulator. **Secondary variable:** Low temperature on SMR upstream the compressor.
16: Intermediate cooling of SMR in compressor	**Primary variable:** N/A. Active constraint.
18a: Bypass valve for LNG after subcooling	**Primary variable:** LNG flow upstream the expansion turbine. **Secondary variable:** Bottom level in nitrogen removal column.
18b: Choke valve for LNG after subcooling	**Primary variable:** Flow of LNG set by temperature controller. **Secondary variables:** Low pressure downstream expansion turbine.

It is assumed that for any operational condition, all sea water coolers are at full operation, i.e., they become active constraints as indicated in above table. This is elaborated in chapter 4.3 in Volden [88].

2.3.3 Control structure

When discussing control structures, large plants separate control application in a hierarchical fashion based primarily on different operating time domains. This depicted in figure 2.9 and elaborated in more detail in Skogestad [68] and Skogestad and Postlethwaite [69]. Due to the thesis scope, only the regulatory layer and the supervisory layer will be further considered.

The general purpose of the regulatory control layer is to locally stabilize the process, i.e., stabilize unstable modes and prevent drifting of operating states. This is maintained using single-input-

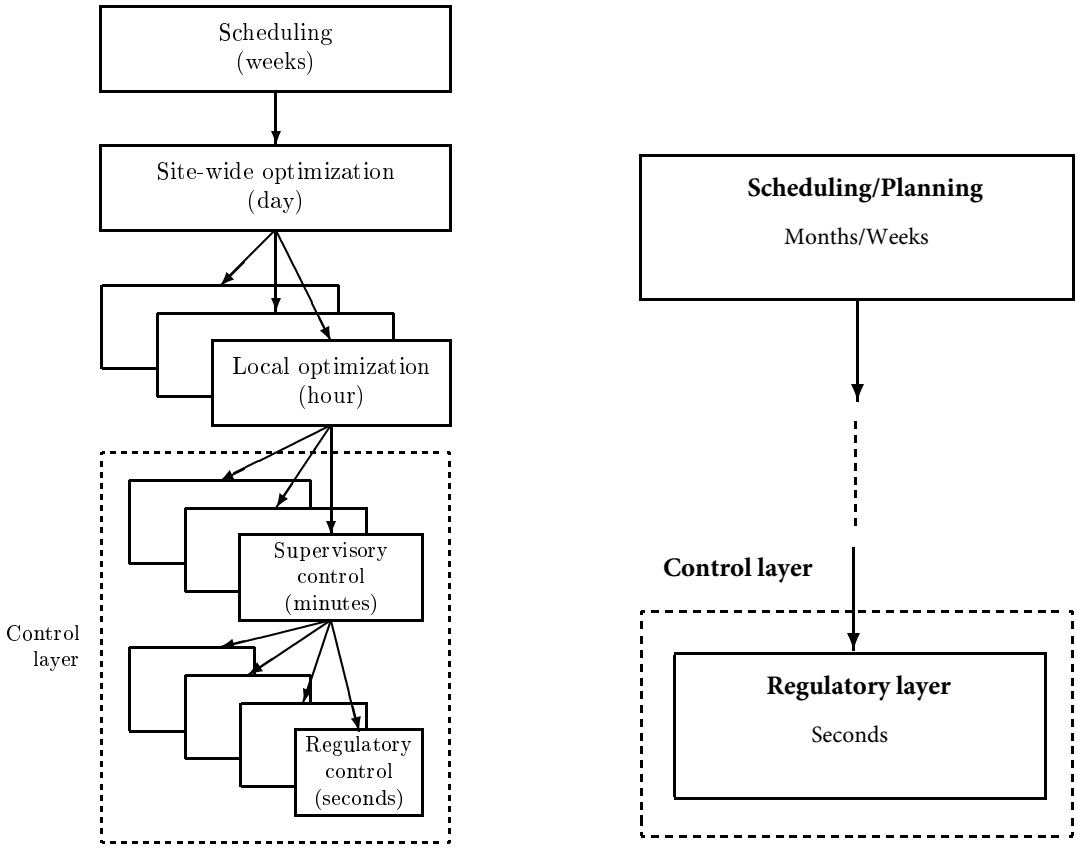
single-output(SISO) control loops. In the theory provided in Skogestad and Postlethwaite [69], it is outlined how to choose appropriate control variables for stabilization and disturbance rejection, to render a supervisory control layer that can handle disturbances by determining the regulatory layer setpoint. This requires somewhat tight control in the regulatory layer, and for the subcooling cycle especially, tight compressor control is a vital task.

The principle of the supervisory control layer is to maintain the controlled variables at optimal setpoints using the setpoints for the regulatory control layer as the manipulated variables. In order to achieve this, Skogestad and Postlethwaite [69] proposes two control strategies; decentralized and multivariable control. The decentralized approach is preferred for non-interacting processes and in cases where the active constraints remain constant, while the multivariable approach is preferred for interacting processes and in cases where the active constraints changes.

In the MFC system, the control structure is different compared to illustrations in Skogestad and Postlethwaite [69], as the control layer primarily consist of a regulatory layer and scheduling layer(see figure 2.9b). This emphasize some of the motivation behind this thesis which is to close the gap between scheduling and regulatory control. The work of defining the MPC application for the subcooling cycle is further considered in sections 7.1 and 7.2.

Some additional motivations for this thesis scope are mentioned in Volden [88]. These are among others; disturbances influencing plant throughput and several open-loop control configurations. These are features which were emphasized in discussions with two operators. An additional motivation behind the choice of multivariable control opposed to decentralized control is the degree of process interactions, which is depicted in section 5.5. In addition, advanced process control and optimisation at the plant has not been a priority since start-up. Recently, the focus has shifted from ordinary operations towards more optimal operation and increasing throughput, which stem from gained experience and the conviction that today's operation at best is sub-optimal. Additionally, as mentioned in section 1.2, SEPTIC have since launch proved its advantages for downstream operations.

For additional information on the control structure features, consult Volden [88].



(a) Typical plant control structure. From of Skogestad and Postlethwaite [69].

(b) Hammerfest LNG liquefaction system control structure today.

Figure 2.9 A typical control structure(a) and the control structure in the Hammerfest LNG MFC process(b).

Chapter 3

System identification theory

In this chapter, some background material on system identification is covered prior to the continuation of model identification initialized in Volden [88]. Predominantly, due to thesis scope, subspace identification theory is treated, while little attention is devoted to other methods. Definitions and preliminaries regarding subspace identification are found in appendices C.4 through C.7. Furthermore, some general principles and models are considered. An introduction and further reading on other identification schemes are mentioned in Volden [88] and references therein.

Although this chapter focus on subspace identification, no attempt will be made to cover this eventually vast topic. For more complete literature on the subject, numerous sources exist. A selection of papers may be Favoreel et al. [18], Overschee and Moor [55], and Ruscio [60]. More in-depth knowledge are found in the books Overschee and Moor [56] and Katayama [41]. For a more surveyable perspective, the papers Ljung [46] and Qin [57], and to a certain degree, Viberg [87], are prominent.

The term *identification for control* introduced in the papers of Gevers [21] and Hjalmarson [30] considers identification of experimental models for control synthesis, tuning and updating. A model which represent the true process in the vicinity of the closed-loop bandwidth is vital for feedback control. Hof and Schrama [33] emphasizes that robust control in the closed-loop bandwidth neighbourhood relies on an accurate model. Outside the bandwidth vicinity, the closed-loop sensitivity function decreases, and thus emphasizes a forgiving effect which make the feedback loop more insensitive to modelling errors far from the closed-loop bandwidth.

In Jacobsen et al. [37], the authors stress the importance of high frequency behaviour accuracy in an identified model, and that steady-state characteristics are of less importance. In addition, the potential pitfalls of fitting individual transfer functions are emphasized. Some of the problems faced in both Jacobsen et al. [37] and Jacobsen and Skogestad [38] are avoided by acquiring a multiple-input multiple-output(MIMO) model directly from subspace identification, and this is one of the arguments of choosing subspace identification. In closing of this chapter, some additional advantages and disadvantages regarding subspace identification are included.

3.1 Subspace identification

From the identification we consider a general continuous state-space model depending on the parameter vector θ:

$$x(t+1) = A(\theta)x(t) + B(\theta)u(t) + v(t) \tag{3.1.1}$$

$$y(t) = C(\theta)x(t) + \bar{e}(t) \tag{3.1.2}$$

where $v(t)$ and $\bar{e}(t)$ are white noise sequences. Choosing a proper model structure, i.e., the dimension of the state vector(x) and the way in which the parametrization vector(θ) enterers the model, is essential. A general and obvious rule of thumb is to tailor the model to its purpose. There are indeed a range of different elements which influence this choice. Söderström and Stoica [70] lists four:

- Flexibility
- Parsimony
- Algorithm complexity
- Properties of the criterion function

Algorithm complexity and properties of the criterion function are more in-depth subjects which are beyond the scope of this work. Hence, only the first two topics are further explained in this thesis. Consult e.g. Söderström and Stoica [70] for a complete explanation on all four topics.

Flexibility describes the model's ability to describe the system dynamics which can be expected in an operational context. Flexibility relates to the number of free parameters in the model, and the way the parameters enter the model itself. *Parsimony* relates to the model representing only the deterministic part of the system. When considering the a parsimony, it pertains to choosing the lowest number of free parameters to obtain an adequate model(Box et al. [6]). An essential consideration on this subject is avoid modelling the stochastic properties i.e., noise properties, which would unnecessary increase the model order and lead to undesired model behaviour. This is also referred to as *overfitting*.

When considering the model flexibility, there are two main approaches; examining plots e.g. evaluating fit of model against true data, and statistical tests and residual analyses. This is considered more in detail in section 4.4 and chapter 5. Certainly, one would never expect a perfect fit, and since the process investigated in this work is nonlinear, developing a linear model *will* lead to deviations.

Subspace identification algorithms are built on four important concepts, as briefly described here:

- **Observability:** The extended observability matrix[4] plays a critical role in subspace identification. Essentially, it advocate determination of the system order n since it has the rank of an observable minimal realizable state-space system. Additionally, its definition allows

[4]The extended observability matrix is defined in appendix C.5

estimation of A and C matrices. Further, it determines the capability to generate the initial state vector x_0 by output and input data(Tangirala [82]).

- **State estimator:** Generally, all system identification problems are principally defined by the fact that states are unknown variables. Therefore, any least squares formulation would result in an nonlinear least-squares problem. By knowing the state-space matrices a priori allows use of the Kalman filter for optimal estimation of states, but this is not the case in any real application. Subspace identification methods deploy a numerical Kalman filter to obtain states without the knowledge of system matrices. This is followed by a least-squares algorithm on the state-space model to estimate the matrices.

- **Realization theory:** This is concerned with building state-space systems which combine deterministic and stochastic partial systems. In general, this is either achieved from empirical or mechanistic descriptions.

  Empirical realizations from data have been subject for major research starting decades ago. The work of deterministic realizations is initialized in the paper of Ho and Kalman [31]. In terms of stochastic data processes, the works of Faurre [17] and Akaike [2] are mentioned. Ho and Kalman [31] formulated deterministic state-space models from Hankel matrices of impulse responses. On the other hand, stochastic realization construct state-space models for random process data. More recently, Katayama [41] developed a realization that constructs state-space models of stochastic data in presence of exogenous effects. Subspace identification essentially extend these concepts to developing models from input-output data which combine deterministic and stochastic data.

  Numerical implementation of these methods is made efficiently by using QR and SVD factorizations[5]. Although subspace realization is based on linear algebra, the mathematical methods utilized are complex and will not be covered in this work. For further details on the `N4SID` subspace algorithm utilized in this work, see appendix C and references therein.

- **Projections:** Subspace identification essentially employ orthogonal and oblique projections of output data onto appropriate spaces[6]. Hence, knowledge of projections is necessary for understanding these methods. However, estimates obtained from subspace identification can be shown as solutions to the minimization of multi-step prediction-error criteria(Tangirala [82]).

3.2 Experimental input design

In general, input design is formulated as an optimization problem with the objective of minimizing a cost function subject to constraints. Additionally, a large class of input design problems try to

[5] QR, LU and SCD factorizations are defined in appendices C.2 and C.4

[6] Orthogonal and oblique projections are defined in appendix C.6.

maximize the Fisher's information[7] to obtain efficient estimates, some of which are mentioned in Tangirala [82].

In general, from a mathematical perspective persistance of excitation(PE)[8] provides desirable inputs for identification, although there are no general solution which meet the needs for different systems. Certainly, an input signal tailored for an linear time invariant(LTI) system is not the best for linear time-varying or nonlinear systems. The papers of Braun et al. [7] and [13] illustrates this well. Nevertheless, persistent excitation is mathematically necessary for input design but not necessarily sufficient. It does not take into account the identification signal cost on the system. The cost could be defined in terms of input amplitudes which cause wear and tear on the actuator, rate of change of inputs, bias and variance considerations(Tangirala [82]). An input that meets the practical requirements of a process operation is said to be plant friendly(Rivera et al. [59]). Some primary considerations apart from persistent excitation are:

- The asymptotic properties of the estimate only depend on the input spectrum, and not on the actual waveform.
- The input must have limited amplitude: $u_{min} < u_t < u_{max}$ to consider physical constraints(e.g., actuators), and to ensure operation of the system within the linear domain. This is more critical for processes under severe noise influence, which is not the case within the MFC simulator regime.
- The parameter estimate covariance matrix is typically inversely proportional to the input power(variance). This emphasizes the advantage of perturbing with as much power as possible into the input. However, in light of the above aspect this is a conflicting factor, which in this work is neglected. Since main dynamics and gains are within tolerable limits, the investigated models are thus fit for purpose. Additionally, the models display varying, but satisfactory uncertainties. This is later depicted in chapter 5 and appendix E.1.
- For multivariable systems the different input signals must be uncorrelated with each other to allow the identification method to separate how the input signals affects the system response.

A property that impose a balance between the amplitude and the variance requirements is the crest factor, C_r:

$$C_r^2 = \frac{max_t\, u^2(t)}{\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u^2(t)} \tag{3.2.1}$$

The smaller the crest factor is, the larger the total energy into the system is. As a small crest factor means a high amount of energy delivered into the system it also means an enhanced signal-to-noise ratio. This is obviously preferable in general. A good signal waveform has a small crest factor, which translates to a minimum amplitude in the numerator for a maximum variance contained in the denominator. The theoretical lower bound of C_r is 1, which is achieved for binary symmetric signals. To achieve this, the input signal should deliver as much input power into the system as

[7]

[8] Persistance of excitation is defined in C.1.

possible(Ljung [44]). As earlier stated, attention was devoted to designing an input signal within the physical constraints while simultaneously keeping the process behaviour in the linear regime. Thus, the power of the excitation signal was not increased for the benefit of a lowered crest factor.

3.2.1 Inputs generally suited for identification

Below, the author summarizes four input signals suited for identification. Since PRBS excitation is chosen for this project, the author place particular emphasis said input signal. Unless otherwise noted, the list is based on Söderström and Stoica [70] and Ljung [44].

- **White noise:** It contains all frequencies uniformly. Theoretically a preferable input signal. Decouples the impulse response parameter estimation problem. Provides uniform fit at all frequencies. However, possesses a high crest factor.
- **Random binary:** Signal generated by starting with a Gaussian sequence and then passing it through a filter depending on the input spectrum requirements. The sign of the filtered signal is the RBS. No proper control over the spectrum. The "sign" operation distorts the spectrum of the input sequence. The RBS has a low crest factor.
- **Pseudo-random binary:** Not strictly a random signal. It is a deterministic signal with the properties of a random signal. Generated using a Linear Feedback Shift Register(LFSR) of n bits, a *maximum length* PRBS is $M = 2n - 1$ sequences long. A PRBS possesses white noise-like properties, although this requires a full length sequence. For an n coefficient PRBS the white noise spectrum is derived using the Wiener-Khinchin theorem relations for periodic signals:

$$P_{xx} = (f_n = \frac{n}{M}) = \frac{1}{M}\sum_{l=0}^{M-1}\sigma_{uu}[l]e^{-j2\pi f_n l} = \begin{cases} \dfrac{U^2}{M^2}, & n = 0 \\ \dfrac{U^2}{M}\left(1 + \dfrac{1}{M}\right), & n = \dfrac{1}{M}, \ldots, \dfrac{M-1}{M} \end{cases} \tag{3.2.2}$$

  As M becomes large the spectrum is uniform at all frequencies, which essentially realize white noise spectral characteristics.

  The frequency content is easily altered, and pertain to generating a band-limited content PRBS. This is a simple operation based on a full-length PRBS, followed by an extension of the constant fractions of the original sequence, while keeping the overall length fixed. Accordingly, the full-length PRBS is re-sampled P times faster than the frequency at which it is generated. This is feasible since the PRBS generation depends on two frequencies, one for the shift register and one for the internal clock. The clock frequency defines a minimum number of sampling intervals after which the sequence is allowed to shift. The resulting signal has the same properties as passing the PRBS through a simple moving average filter of order P(Tangirala [82]).

Mathematically, a disadvantage is that only maximum length PRBS possess the desired properties. For a given amplitude range, PRBS packs the maximum variance or energy, hence it has the lowest a crest factor. It is suited for linear systems only; since it switches between two states, it cannot detect nonlinearities.

- **Sum of sines or chirp signals:** A combination of sinusoids of different frequencies, which are generally known beforehand.

$$u_k = \sum_{i=1}^{M} a_j \sin(\omega_i k + \phi_i), \quad 0 \leq \omega_1 < \omega_2 < \cdots < \omega_M \leq \pi \tag{3.2.3}$$

These signals provide good estimates of the transfer function at the respective frequencies. Both amplitudes and phases are design parameters, but the phase additionally have a influence on the amplitude, and hence the crest factor. A useful guideline is to keep the sines as much as "out of phase" as possible to keep the crest factor low. Since the spectrum is not continuous, the estimates at other frequencies are not available.

3.2.2 Preliminaries for experiment design

As Volden [88] states, some preliminary tests have to be conducted in order to arrive at the optimal input design, unless this information is known a priori. A short summary list from Volden [88] including aspects from Tangirala [82] follows:

- Perform a preliminary step test on the system, to render information on gain, time constant, delay, and inverse response. These tests are depicted in appendix B.2.
- To check for nonlinearities and range of linearisation, steps in both directions are required.
- Identify effective time constant of the process, t.
- Compute effective bandwidht: $\Omega_{BW} = 1/t$.
- Set upper bound on maximum frequency to $10 - 20$ times Ω_{BW}.
- Design the input sequence accordingly.

Ljung [44] states some additional general guidelines:

- Choose excitation frequencies and input energy in neighbourhood where a model is intended, and where the disturbance is insignificant.
- Under open-loop conditions and for linear systems, apply binary, periodic inputs, and tailor the input energy.
- The error in a transfer function estimate is inversely proportional to the sample size and SNR.

Additionally, rate of sampling is also mentioned. This aspect will not be granted any further attention given the environment the experiments are executed within; the simulator and OPC interface, which allows for easy manipulation and adjustment of sampling. Although, keeping this aspect in mind for the MPC development is advantageous. This is further considered in chapter 7.

After obtaining the experiment data, some checks are considered, and if needed, performed. This could be some pre-filtering, detection and elimination of outliers, in addition to identifying drifts, offsets and trend in the investigated data.

Linear models satisfy superposition and homogeneity principles, and for nonlinear processes, linear models relate deviation variables, i.e., deviations from a nominal operating point which typically is chosen as the steady-state condition. Hence, it is important to construct deviation variables from the data as an initial step for identification. However, steady-state values or the nominal point may not be available, and in these situations it is a common practice to replace it with a sample mean of the available data. An important aspect when obtaining multiple data sets is the necessity to run the experiments with a consistent estimated nominal point. This is especially emphasized for nonlinear systems where varying nominal values may cause discrepancies between the data sets. However, when considering models which include gain scheduling this may be a method to identify appropriate gains for the model. Although, not suited for identification as merging different data sets may result in inconsistent deviation variables.

3.2.3 Model structure comparison

Although the model structure is defined for this work, the author summarizes some features of several general model structures, and advantages and disadvantages for each structure are included. Note that this comparison should be seen in the context of `MATLAB` and its system identification toolbox([47]).

Table 3.1 General comparison of model structures

Model structure	**Advantages**	**Disadvantages**
ARX models	Simple input-output relation. Linear regression. Preferable when the model order is high.	Disturbances are part of the system dynamics. Limited freedom to describe the disturbance.
ARMAX models	Flexible disturbance model. Useful when there are dominating disturbances entering early in the process.	Disturbances are part of the system dynamics. Not a linear regression.
Output Error models	System dynamics and the disturbance are described separately.	Not a linear regression. No disturbance model.

Box-Jenkins models	System dynamics and the disturbance are described separately. Useful to identify processes where the disturbance enters late in the process.	Not a linear regression.
Low order transfer function models	Ideal for identifying physical systems. Provides estimation of process delays. The identified coefficients has a physical interpretation, i.e., preferable for simple mechanistic and nonparametric modelling.	No disturbance model. Maximum model order is 3. Limited freedom in choice of zeros.
State-space models	Represents the relationship between input, output and noise as a system of first order differential equations. Easy to incorporate insight into physical features of the system if based on white or grey box approaches.	Prone to overfitting. Some statistical tools provide guidance, but may not be consistent. See section 5.4.1.1 and appendix E.2. Example in appendix C.2 emphasizes this affirmatively.

3.3 Why not a rigorous modelling approach?

For complex chemical processes rigorous modelling is in general a complex and time consuming process. In addition, the cooling cycle studied in this project include phase transitions and flashing which further increase the complexity significantly. By using equations of state, one is principally able to accurately predict properties of hydrocarbon fluids over a wide range of conditions. Two equations of state which are proven successful in applied thermodynamics are Peng-Robinson(PR) and Soave-Redlich-Kwong(SRK)(Ghosh [22]). These equations are utilized in the works of e.g., Heldt [28] and Jensen [39], but not considered in this work.

In Volden [88], arguments for both mechanistic and empirical modelling were presented, and a summary is here included: In Foss et al. [20], the authors compare model development costs of mechanistic and empirical models, and find that empirical models are about 10% of the cost of an empirical model. This is supported by Hauge et al. [27], where 200 days are spent on developing a mechanistic model for a paper machine while a substantially shorter time is spent on initially obtaining an empirical model. Due to the most present constraint - time - a mechanistic modelling approach is not considered further, and focus is rather shifted towards system identification and the MPC development. For the sake of completing the argument, empirical and mechanistic modelling are considered briefly.

A mechanistic model holds unique extrapolating properties and comprises a reasonable number of parameters which is comprehensive to model behaviour. This makes model adjustment and

re-tuning manageable. In addition, non-manipulatable disturbances makes empirical modelling of their output influence difficult. A long-term advantage is that the mechanistic modelling increases process knowledge which will always be useful in later relations.

For empirical models, an additional advantage besides the time and cost spent on development are no requirement of considerable process knowledge. Some disadvantages for empirical modelling is the increased model complexity and increased risk of overfitting the model. Overfitting occur when the developed model comprises excess parameters from what is identifiable, and the identification focus too much on local features rather than global features. This may arise if stochastic data is treated as deterministic data, e.g., noise is treated is deterministic features and order of model is increased to capture noise properties. Although, no general rules apply for prevention of overfitting, but some heuristics are usually beneficial to utilise(Söderström and Stoica [70] and Ljung [44]).

3.4 Some notes on closed-loop subspace identification

Since this thesis essentially consider open-loop identification experiments, some notes on closed-loop identification is included in this section. The purpose is to enlighten the reader of the advantages and limitations that pertain to closed-loop experiments.

Some dynamical systems operating in closed-loop may be unstable, poorly damped or prone to unknown disturbances making identification in closed-loop the only option(Hof [32] and Söderström and Stoica [70]). This may come of plant instability, requirements of production control at all time, economic, or safety reasons. It could additionally contain inherent feedback mechanisms. Still, one must be able to identify the open-loop system from the closed-loop data. When discussing closed-loop identification, Söderström and Stoica [70] states two factors which consider the suitability for an identified model: identifiability and accuracy. Identifiability describes a system whose corresponding parameter estimates are consistent, while accuracy pertain to the ability of obtaining an exact model from closed-loop experiments.

The majority of subspace algorithms are developed for open-loop identification, i.e., the algorithm expects uncorrelated inputs and noise in the control loop investigated. Ordinary subspace algorithms does not handle closed-loop data well due to the use of an extended future horizon that introduces correlation between inputs and past noise[9]. Later developments account for this by utilising e.g. some pre-estimation for separation of the two terms, or using the parity space rather than the observability subspace(Qin [57]). For the ordinary case, closing the loop would certainly lead to correlation between inputs and noise, thus providing biased estimates. This is due to the use of instrumental variables which assume no correlation between said data to filter noise in the investigated data.

The basic closed-loop data problem is the loss of information caused by feedback which make

[9]This is the case for the `N4SID` subspace algorithm as demonstrated in appendices C.5 and C.7.

the closed-loop system less sensitive to changes in the open-loop system. Performing closed-loop identification, depending on method, may complicate the search for the real process model as the controller and model are identified as a combined model. Limited or no knowledge of the controller would pose additional modelling uncertainties. Furthermore, varying controller parameters play a significant role in closed-loop identification as demonstrated in Bakke et al. [3].

Since the process input is determined by the control law, the user has no direct control over input excitation. Consequently, the user defined perturbations must be set indirectly utilizing setpoint changes. Given no external excitation to the closed-loop system provides data which renders separation of the process model and the inverse of the controller infeasible. Obviously, without setpoint changes or external perturbations, the system cannot be identified from regulatory data. This is because the input excitation is derived entirely from that of the output, and in this context the input is said to be endogenous to the output. This emphasize the necessity of some additional excitation in the input uncorrelated with the output.

Ljung [44] state some general notes on closed-loop identification and its limitations:

- A too simple controller may lead to inconclusive closed-loop experiments, even if the input in itself is persistently exciting.
- For open-loop data, output error models will give consistent estimates of the transfer function, even if the additive noise is not white. This is not true for closed-loop data.
- The subspace method will typically not give consistent estimates when applied to closed-loop data. As pointed out in Qin [57], this issue is resolved in more recent subspace algorithms.

Methods which may provide consistent estimates for open-loop data may fail when applied for closed-loop identification. This includes e.g. impulse responses, spectral and especially correlation analyses since feedback conditions produce biased estimates, but also subspace methods may yield erroneous estimates.

Additionally, when closing the loop less comprehensible responses due to controller action may be introduced. Just consider a simple example of level control in a tank by manipulation of flow. If we assume a tank of uniform diameter, one would expect a pure integrator with gains dependent on the valve (inlet or outlet) position. Assuming a stable controller, and performing a closed-loop experiment with a step in controller reference would reveal a stable model when considering the reference as input and the tank level as output. An inexperienced eye may be tempted to model this system as the tank model, which obviously lead to an erroneous model due to inclusion of the controller. On the other hand, if e.g., bounds on the input and outputs are critical, conducting the experiment in closed-loop is favourable due to the inherent control over the variables in question.

There exist several methods of closed-loop identification methods:

- Direct identification
- Indirect identification
- Joint input-output identification

Direct identification is the most simple and straightforward method. This method simply ignore the existence of feedback, and the data is treated similar as for an open-loop experiment. Although the direct method is easy to apply, some limitations persist in terms of identifiability(Söderström and Stoica [70] and Ljung [44]):

- Identifiability is not guaranteed if input is determined through a noise-free linear low-order feedback.
- Using a high-order noise-free linear feedback may enable identifiability. This is dependent on order of the unknown system.
- Some easy-to-implement schemes to obtain identifiability is; inclusion of external varying setpoint or to utilise a controller with internal setting shifts during the experiment.

Indirect identification describes a two-step approach where it is assumed the external setpoint is measurable and the feedback law is known. The first step is to identify the entire closed-loop system. Then the known feedback term is excluded from the model to obtain the open-loop model. This method is limited to include linear and known controllers for obvious reasons stated earlier in this section(Söderström and Stoica [70]). A known controller model render the indirect method useful, and its advantage, even for unstable open-loop processes, are accurate estimates without estimating any noise model. However, since a majority of controllers incorporate nonlinearities, potential pitfalls arise. An inaccurate controller model will propagate and cause erroneous estimates of the process model(Ljung [44]).

Joint input-output identification considers the known input and output signals as outputs from a multivariable system driven by white noise. It treats the data as mutlivariable series with dimensions according to the input and output signal dimensions, and the system is identified using the original parameters as unknowns. The method relies on computing the inverse of a possibly close-to-singular transfer function, which limit its scope(Bakke et al. [3]).

Apart from influencing the identifiability, feedback may additionally influence the accuracy. From an open-loop perspective, closing the loop may considerably lower the variations in process measurement signals. Inherently, the controller tries to keep the process within bounds, and by doing so opposes variability and excitation in vital signals. This leads to less information being present which ultimately results in a model of lower accuracy. However, this is not the case when in the majority of applications, some constraint on the input or output variations are indeed present. An example which favour closed-loop identification for a minimum variance controller under output variance constraints is depicted in Söderström and Stoica [70].

3.5 Advantages of subspace identification methods

Originally, system identification considered computation and development of polynomial models, but due to risk of facing ill-conditioned problems, especially for MIMO systems, subspace identi-

fication has become the preferred method during the last decades. This is mainly due to numerical robustness and the algorithms non-iterative nature. The challenges of facing ill-conditioned models in MIMO identification prior to subspace methods are discussed in Jacobsen and Skogestad [38] and Jacobsen et al. [37]. For higher order MIMO systems, finding the appropriate parametrization increases in difficulty and is by far trivial, which further favours subspace methods. The numerical robustness in subspace methods is mainly due to the use of singular value decomposition and QR factorisation[10]. The above advantages applies when the user has no preference on structure for the particular system or a specific basis for the states. Tangirala [82] name these descriptions unstructured state-space models contrary to structured definitions.

The mathematical features ensures convergence to a *global minima* of the criterion function while remaining insensitive to initial estimates. This emphasize that there is no practical difference between zero and nonzero initial states. Prediction error methods are iterative, may perform unsatisfactory for problems with slow or lack of convergence and may even become numerically unstable. Additionally, the risk of honing onto a local minima is present. In classical approaches, an *a priori* parametrization, and knowledge of the order and observability indices are required(Overschee and Moor [55]). Parametrisation for prediction error methods rely on canonical forms which may be sensitive to even small perturbations, leading to ill-conditioned problems.

A major drawback when considering use of classical input-output approaches for multivariable systems is that the delays and orders that have to be determined grow substantially with the dimensionality of the system. Subspace methods, on the other hand, are equipped with an automated method for order determination in principle. Although, it is emphasized that automatic order determination may be inconsistent for non-ideal cases as demonstrated in appendix C.2.

3.6 Some disadvantages on subspace identification

As with practically anything, there are some disadvantages by utilizing subspace identification compared to other methods. Some important negative features are listed here:

- **The need of a large set of input/output data:** The statistical properties of the geometrical methods used in subspace identification require a large amount of input/output data samples. This could limit application in some areas, e.g. where data are sparse, such as economic modelling, and generally for dynamical systems which suffers from pathological sampling[11]. One example is the Hall-Héroult process for primary aluminium production where manual measurements have a low and pathological sampling frequency[12].

[10] See appendices C.2 and C.4 for some basic definitions and further reading.

[11] A term defined in Chen and Francis [12] which indicates that insufficient sampling ruin controllability and observability properties of a system.

[12] The works of Hestetun [29] and Kolås [42] provide more information on the Hall-Héroult process and challenges related to low measurement frequency.

- **Theoretically complexity:** The algorithms utilize geometrical projections in the high dimensional row or column spaces of matrices. This increases the threshold for a mathematical understanding, however, they can be interpreted in other well known frameworks.
- **Applications for recursive on-line identification:** The basic algorithms were developed to identify the system parameters from off-line data, i.e., identification from a given complete sequence of input/output data. However, for practical reasons, some industrial applications favour a recursive algorithm to identify in the real-time from on-line data. This is a more recent field of development which the initial subspace algorithms not were suited for. Though, recent works have shown promising results, see e.g., Bako et al. [4] and Mercere et al. [49].
- **Utilization of *a priori* knowledge:** Prior knowledge can not be easily incorporated into subspace methods. These methods have a pure black-box approach to the identification problem, however there exists often some a priori information which could be exploited to increase the quality and the robustness of the identification. This could for instance be an educated guess on system order, which for some cases may be estimated without much effort by some superficial modelling. Although, this approach has obvious limitations.
- **Identifiability issues:** Since a dynamic system may be described by an infinite number of state-space models, subspace methods does not provide an unique estimation. Thus, we run into identifiability issues, i.e., the problem of estimating a unique model, which is impossible for any state-space realization, including subspace methods. Subspace identification methods also suffer from the identifiability issue since they do not explicitly impose identifiable structures on the state-space models. We can only identify state-space models in some basis, over which we may have little or no control. Formally, we say that subspace algorithms identify state-space models uniquely only up to a nonsingular transformation, or that they construct a realization of the system.
- **Utilizing algorithm parameters:** Besides *a priori* knowledge, one may obtain better results by utilizing the subspace algorithm parameters. As stated earlier in this work, and in Ljung [45], parameter adjustments essentially relies on heuristics as no research have provided unambiguous results on how to exploit these parameters. Sima et al. [65] offer some general, non-rigid guidance, but no global solution[13]. Additionally, statistical tools are used to provide advices, but this does not exempt the user to employ practical knowledge in the identification process. This is demonstrated in section 4.3.

[13]This is demonstrated in section 4.2.1

Chapter 4

Practical system identification

This chapter is a continuation of chapter 6 in Volden [88], and the main part of this work has been to improve and attain a satisfactory model for the system. Some initial results from Volden [88] are included as the basis for the progressing work on model identification. Further, an important experience regarding the system identification process is included. In closing, some results from a new iteration on system identification are included.

4.1 Previous work from autumn project

Based on the works of Braun et al. [7], Deflorian and Zaglauer [13], Hauge et al. [27], Ljung [44], and the information in Foss et al. [20], Volden [88] performed several approaches to obtain an empirical model of the chosen 2x3 MIMO system. Based on the conclusions in Favoreel et al. [18], Overschee and Moor [56] and Ruscio [60] regarding the comparison of prediction error methods[14]and subspace algorithms, an initial model based on two subspace algorithms were obtained as the basis. The two subspace algorithms utilized in this work, `N4SID` and `DSR`, are further described in Overschee and Moor [55] and Ruscio [60, 61], respectively.

[14] A review of closed-loop identification using prediction error methods are conducted in Forssell and Ljung [19].

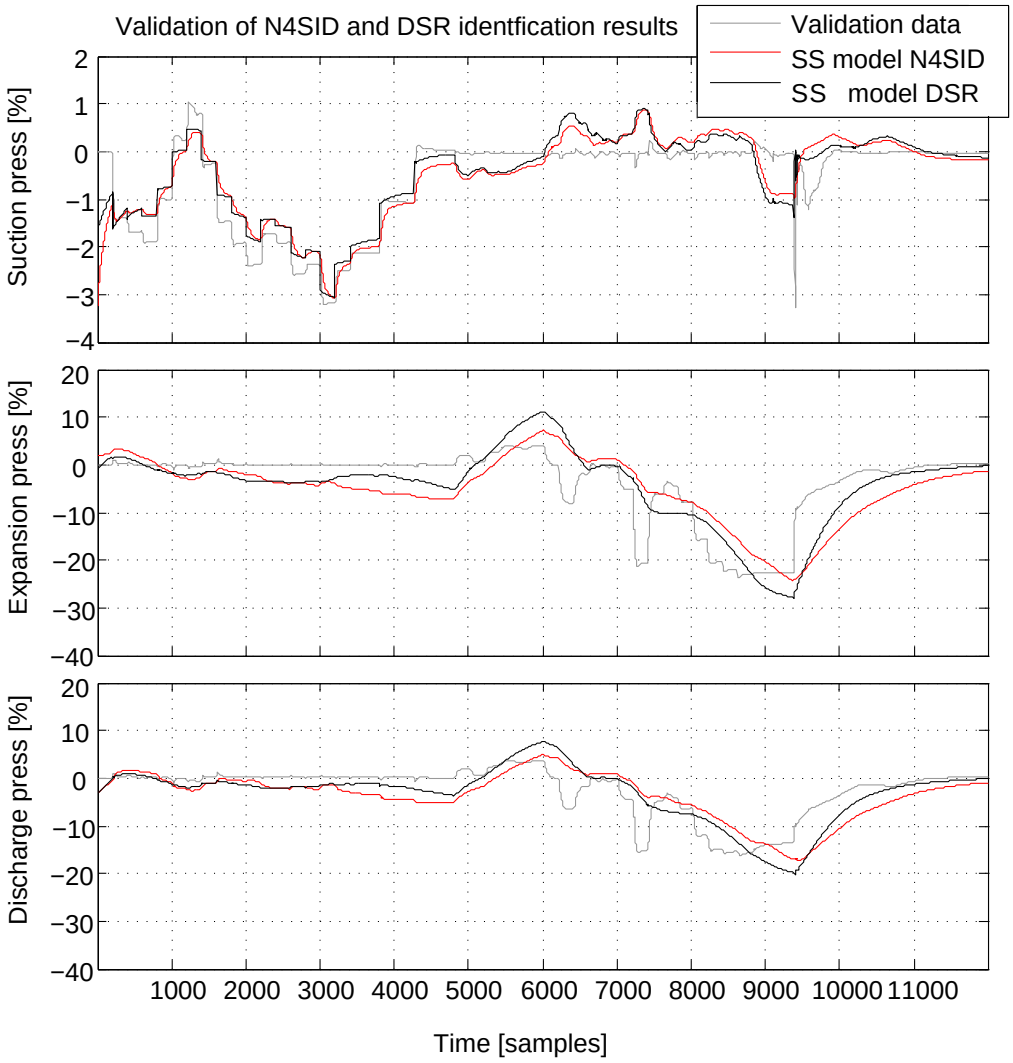


Figure 4.1 Comparison plot of the two state-space models against the validation data. From Volden [88].

The identification results achieved for the initial model were not satisfactory. An indication on this were obtained in figures 6.6 and 6.7 in Volden [88] together with the results in figure 4.1. Figures 6.6 and 6.7 in Volden [88] illustrates the data sets used for model identification and validation. By observing the identification and validation data sets in Volden [88], it is inevitable to observe the resemblance between the two data sets. Thus, we would expect a high fit for the validation set essentially, compared to a more random validation set. Still, the goodness of fit value was not acceptable.

From the comparison figure 4.1, we observe both derived models failing to capture faster dynamics. Only the main dynamics are somewhat similar. One reason behind this could be the low order of both derived models. A higher order may render a model which holds faster dynamics and thus render a better model fit against validation data. Another prominent feature is the method utilized for identification and validation; the AMPRBS procedure, which is, as shown in Deflorian and Za-

glauer [13], primarily tailored for nonlinear identification. Below table list the initial identification results based in the NRMSE value.

Table 4.1 Validation results of returned state-space models from subspace identification. From Volden [88].

Validation of state-space model			
Model order		NRMSE fit	
N4SID:3	DSR:2	N4SID	DSR
Y1: Compressor suction pressure		47	51
Y2: Expansion pressure		21	34
Y3: Compressor discharge pressure		21	32

To develop a linear model from an experiment suited for nonlinear models is in general not advantageous. Thus, to obtain a linear model, an identification procedure tailored for linear systems is preferred. In the continuation of the model development work, only system identification procedures suited for linear system were considered, in accordance with the aspects covered in sections 3.1 and 3.2.

In Volden [88], several identification experiments were carried out, both for open-loop and closed-loop configurations. Three types of excitation signals were utilized, PRBS, AMPRBS and ramping step tests. For open-loop tests, the ramping step test and AMPRBS experiments were conducted, and for closed-loop an initial step test was carried out in addition to several PRBS experiments. The additional work performed in relation to system identification in Volden [88] is not covered in this report, and the reader is advised to consult said reference. Although, a table listing the best identification results based on closed-loop identification data from Volden [88] is included below.

Table 4.2 Identification and validation set results no. 7. Note that the fit criteria is higher for the validation set for almost all variables. This is not what one would normally expect, although it does not render the results invalid. The reason may be close to identical identification and validation data sets.

Identification set 1	**Validation set 1**	**Model order**	**Output**	**Fit against Id set**	**Fit against validation set**
PRBS test no.3 PC1669	PRBS test no.7 PC1669	N4:16	Y1	**N4:74.2%**	N4:67.8%
				DSR:69.4%	**DSR:76.2%**
Id set 2	**Val set 2**		Y2	**N4:66.3%**	**N4:68.1%**
		DSR:24		DSR:52.8%	DSR:53.2%
PRBS test no.1 PC1282	PRBS test no.4 PC1282		Y3	**N4:67.2%**	**N4:67.4%**
				DSR:54.2%	DSR:53%

As the results indicate, achieving an sufficiently accurate model using the old identification data were

not easy, even with a particularly high model order. In addition, observations from the validation data revealed somewhat different results achieved by the subspace algorithms. The N4SID algorithm returned a higher fit almost consistently compared to the DSR algorithm(see appendix D in Volden [88]). Based on the results in appendix D in [88] and table 4.2, which favours the N4SID algorithm, only the N4SID algorithm is further used for subspace identification in this thesis. Although, some contrary results are found in Ruscio [61] and Hauge et al. [27], but to investigate the reason behind this is beyond the thesis scope. One reason may be the user-friendly differences; N4SID is well documented, at least to some extent, and there exist numerous applications in research and industry, while the DSR algorithm is less documented and there exist little information on parameter adjustment for the DSR case.

4.2 Options to improve model accuracy

As the results in table 4.2, and in appendices C and D in Volden [88] indicate, attaining a satisfactory model has not been straightforward. To improve the model, attempts to increase the identification accuracy were performed. As stated in Ljung [45]; although subspace identification have been researched thoroughly for the last decades, the different algorithms and adjustments of algorithm parameters are essentially based on heuristics. There exist, generally speaking, sparse information on how to utilise features within subspace algorithms in order to improve model accuracy. Some of these features were investigated in this work.

4.2.1 Altering subspace algorithm and prediction horizon

One option to improve identification results is to alter the subspace algorithm scheme within the N4SID environment, which includes various weighting schemes used for the singular value decomposition(SVD). In this work, two options were investigated:

- The multivariable output error state-space(MOESP) scheme described in Veerhagen [85].
- The canonical variate analysis(CVA) scheme described in Larimore [43].

Different trials revealed that the results obtained by using the CVA scheme were almost consistently surpassing the results obtained using the MOESP scheme. Thus, CVA was the preferred scheme for the contiuation of model development. This is depicted in figure 4.2.

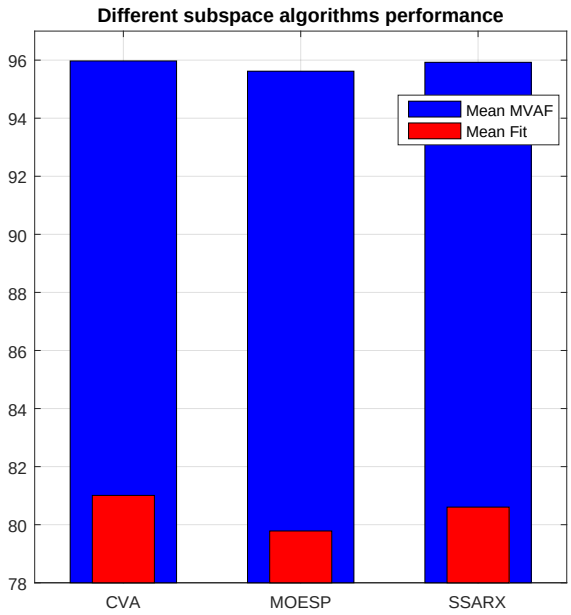


Figure 4.2 Comparison of three subspace algorithms on a simulator obtained data set. `CVA` yields the best result. `MOESP` provides the lowest result. Although, the author emphsizes that this result is not valid in general and valid only in this context.

In addition, altering of the forward and backward-prediction horizons were performed due to the fact that the experimental data from the simulator is considered almost free of noise[15]. By not considering the noise, and thus neglecting modelling of noise properties, we consider the model as an output error model. This is elaborated below, and translates to choice of horizon for the subspace algorithm.

We have the discrete state-space representation when considering noise as

$$x(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + w(t) \tag{4.2.1}$$
$$y(t) = C\boldsymbol{x}(t) + D\boldsymbol{u}(t) + v(t) \tag{4.2.2}$$

When considering the input-output dynamics, we alter the state-space representation to

$$x(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \tag{4.2.3}$$
$$y(t) = C\boldsymbol{x}(t) + D\boldsymbol{u}(t) + v(t) \tag{4.2.4}$$

Since we neglect modelling of noise and thus system orders corresponding to the noise model, we expect a lower order model. Because the input and the output sequences, u and y, are given, the challenge is to find the state vector x. From the general system given by the impulse response, we

[15]The simulator is considered almost free of any process noise or measurement noise. On the other hand, due to its domain i.e., numerical solvers and thermodynamic tables, some numerical noise is indeed present, though, this is considered of minor significance in this work.

have

$$y(t) = \sum_{j=0}^{\infty} \left[h_u(j)u(t-1) + h_e(j)e(t-1) \right] \tag{4.2.5}$$

where u is the input and e is the innovations. Defining the true k-step ahead predictors by deleting the input to $y(t)$ from $e(j)$, $u(j)$: $j = t..., t-k+1$:

$$\hat{y}(t|t-k) = \sum_{j=k}^{\infty} \left[h_u(j)u(t-j) + h_e(j)e(t-j) \right] \tag{4.2.6}$$

Remembering that the noise model is not included in the state-space model, the terms including $e(j-1)$ are neglected(Ljung [44]).

The true predictor relies on infinite data, and is thus not practical. An approximation of the true predictor includes finite data for both s_2 past inputs and s_1 past outputs, and takes the form

$$\begin{aligned} \hat{y}(t+k-1|t-1) = \alpha_1 y(t-1) + ... + \alpha s_1 y(t-s_1) \\ +\beta u(t-1) + ... + \beta s_2 u(t-s_2) \end{aligned} \tag{4.2.7}$$

When considering identification of an output-error model the past innovation terms is neglected as mentioned, which implies that the above predictor only considers past inputs, i.e., $s_1 = 0$. Even though, utilizing this method did not increase the `NRMSE` fit significantly as seen in figure 4.3, although the `MVAF` fit was improved. Additionally, the horizon of past inputs and outputs were chosen in accordance with Akiake's criterion which provided almost identical results. More information on this criterion is found in e.g. Ljung [44] and Akaike [1]. Though, it is emphasized that the criterion is purely statistically based, and does not consider practical aspects in the identification routine.

In Sima et al. [65], the authors state that a large number of step-ahead predictors, s is favourable e.g., $s = 2n$, where n is estimated system order. A large s usually produces a more accurate result. A typical lower bound is $s \geqq n$. In terms of using past inputs and outputs, a general approach is $s = p = q$, where p denotes past inputs and q denotes past outputs. Although the output-error model approach did not increase fit, the method yield practically identical results as the best performing algorithm horizon, which in this case was the Akiake's criterion. This is depicted in figure 4.3.

Another option to obtain better quality models may be to combine the subspace method with the prediction error method(PEM)(Ljung [44]). This is further supported in Viberg [87]. One procedure is to use the subspace identified model as an initial estimate for the prediction error method. This procedure requires a suitable parametrization for the initial model.

We have the general PEM estimate of θ_N as

$$\hat{\theta}_N = \arg\min V_N(\theta, Z^N) \tag{4.2.8}$$

where for the multivariable case

$$V_N(\theta, Z^N) = h(Q_N(\theta, Z^N)) \text{ , and } Q_N(\theta, Z^N) = \frac{1}{N}\sum_{t=1}^{N} \epsilon(t,\theta)\epsilon^T(t,\theta) \tag{4.2.9}$$

and the prediction error, ϵ, for a specific model is $\epsilon(t,\theta) = y(t) - \hat{y}(t|\theta)$. The method of merging subspace identification and PEM to obtain superior results were carried out briefly, but the improvements were insignificant. In addition, the procedure is not particularly user-friendly and time consuming. This is mainly due to the resulting problem which now becomes a nonlinear optimization problem. Thus, the conclusion was to keep the results from the subspace identification as these were considered above satisfactory.

Figure 4.3 depicts 11 horizons tested for the subspace algorithm and how well the different horizons score in terms of two fitness values. These results stem from a test where pre-defined matrix contained different horizons. This matrix was defined as:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad A \in \mathbb{Z}^{mxn} \tag{4.2.10}$$

where $m = 150$ and $n = 3$. To test a wide range of horizons, the step ahead predictors, past inputs and past outputs were altered such that

$$a_{ij} = \{0 \leq a \leq 150\}, \quad i \in \mathbb{Z}^m \ \wedge \ j \in \mathbb{Z}^n \tag{4.2.11}$$

While running the algorithm with various horizons, remaining parameters were kept at a fixed value. The horizon based on Akiake's criterion gave the best result. Additionally, as mentioned above, considering the system as an output-error model gave almost identical results, i.e., setting past inputs to 0. As a matter of fact, the `MVAF` value for the output-error case and for the case of horizon = $[75, 141, 100]$ yield better results compared to the Akiake case. Still, the `NRMSE` value were identical for these cases, and this criterion is regarded of higher importance in this work. Too large step-ahead predictors yield low score, as the biggest one depicted is 125. Past inputs and past outputs vary more, thus these results are somewhat inconclusive and does not give an unambiguous indication on preferred neighbourhood of values to use.

Another interesting finding is comparing the recommendation from Sima et al. [65] which states to set all three values identical. From the plot, this approach yield low results, and altering the identical numerical values render insignificant improvements. Still, the best result from the Akiake criterion is close to similar for all three values: $[27, 24, 24]$. This emphasizes the non-triviality of choosing the best horizon.

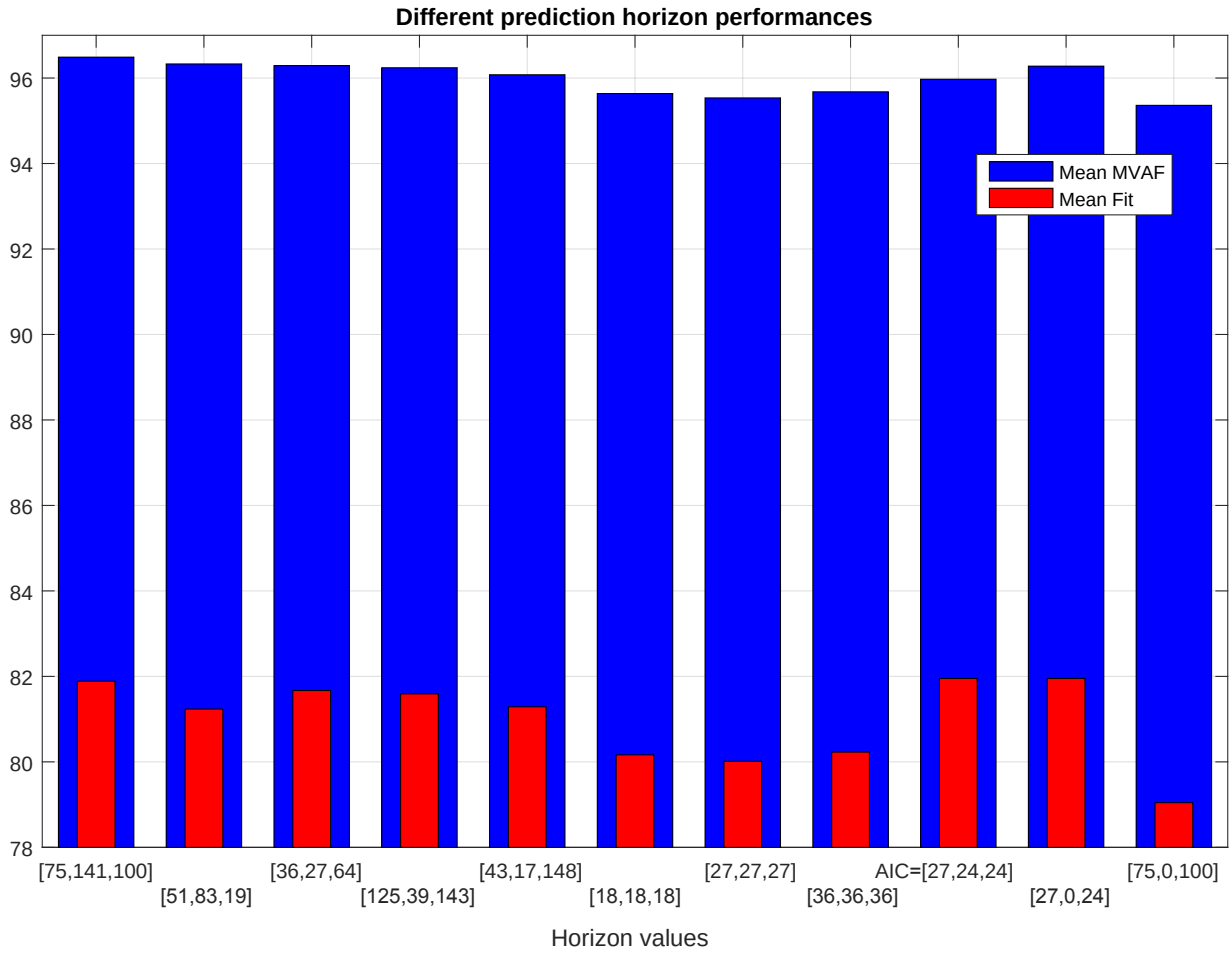


Figure 4.3 Performance of 11 best subspace horizons from a selection of 150 candidates. Selection is based on a combination of the best `NRMSE` and `MVAF` values.

To further emphasize the non-trivial part of choosing the algorithm horizon considered in this section, additional plots are included in appendix E.5. These figures depict results based on a slightly altered algorithm horizon.

An additional option to investigate diverse subspace parameters was briefly considered. Apart from comparing different algorithms and algorithm horizons, focusing on a particular frequency region and discarding dynamics outside that region was regarded in order to obtain a satisfactory model. This approach may be advantageous if e.g., if control bandwidth is limited by actuator dynamics which emphasize to discard higher-frequency dynamics. The motivation behind this approach are demonstrated in Jacobsen et al. [37] and Jacobsen and Skogestad [38]. The resulting comparison plot is depicted in figure 4.4.

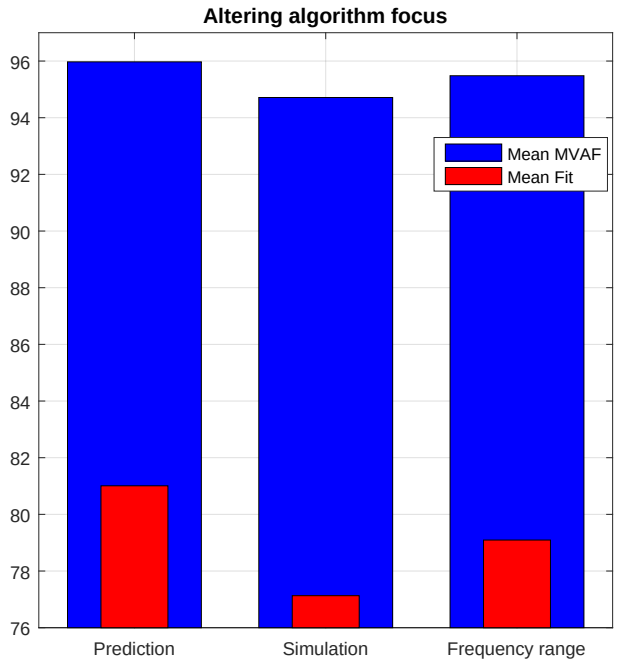


Figure 4.4 Altering subspace algorithm focus

As the figure indicate, focusing on a smaller range of frequency for the model identification did not improve the fitness criteria. This may be explained by the ideal process data utilized in the identification procedure. For example, including actuator dynamics is not carried out in the simulator development. No further investigation behind these results was carried out. The frequency interval tested were defined by the preliminary tests outlined in section 3.2.2.

4.2.2 SISO modelling approach

Since MIMO models are inherently more complex to identify, an alternative is to utilize the superposition principle which holds for linear systems(Seborg et al. [64]). This approach allows for development utilizing SISO principles, i.e., perturbing input-output wise, while holding the other channel inputs constant. For the 2x3 system investigated in this thesis, the resulting system based on SISO models comprises thus 6 SISO models. The idea of working with subsets of inputs and outputs is also described as a potentially easier way to handle difficult modelling procedures in Ljung [44].

Some potential advantages for this particular approach is:

- Testing input channels individually is a simple measure to reveal if all measured inputs have a significant influence on the outputs. This helps reveal if system simplifications may be realized without facing non-modelled discrepancies.
- Generally speaking, the fit increases when adding more inputs, and contrary, the fit decreases when more outputs are included. This is due to the fact that any identification which include

several outputs is more difficult than modelling systems where one simply must account for a single output.

- If MIMO modelling fall short, SISO modelling allows for identification of troublesome outputs in the process. This could contribute to achieving a satisfying model, since one is able to take additional measures for the particular output regarding modelling effort. However, models for control purposes will generally produce better results if constructed for all outputs together. In addition, knowledge of all previous output channels provides a better basis compared to just knowing the past outputs in a channel.

An additional motivation to this approach is the practice of system identification in Statoil. System identification by a SISO approach is included in SEPTIC, and it is the preferred method to obtain empirical models for suitable applications[16]. It is the author's impression that empirical modelling is an extensive used approach in SEPTIC applications, and this is further supported in Strand and Sagli [80].

When considering the system based on number of inputs and outputs we have the models:

$$\begin{aligned} \frac{Y_1}{U_1} &= G_{p11}(s) \quad \frac{Y_2}{U_1} = G_{p21}(s) \quad \frac{Y_3}{U_1} = G_{p31}(s) \\ \frac{Y_1}{U_2} &= G_{p12}(s) \quad \frac{Y_2}{U_2} = G_{p22}(s) \quad \frac{Y_3}{U_2} = G_{p32}(s) \end{aligned} \tag{4.2.12}$$

The above transfer functions describes the change in Y_j from U_i, where $j \in [1, 2, 3]$ and $i \in [1, 2]$. In addition, from the superposition principle, we have that a change in manipulated variables render an additive effect on each controlled variable:

$$Y_j(s) = G_{pj1}(s)U_i(s) + G_{pj2}(s)U_{i+1}(s) \tag{4.2.13}$$

$$\tag{4.2.14}$$

and for the total system we get

$$\underbrace{\boldsymbol{Y}(s) = \boldsymbol{G}_p(s)\boldsymbol{U}(s)}_{\Downarrow}$$

$$\boldsymbol{Y}(s) = \begin{bmatrix} Y_1(s) \\ Y_2(s) \\ Y_3(s) \end{bmatrix} \quad \boldsymbol{U}(s) = \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix} \quad \boldsymbol{G}_p(s) = \begin{bmatrix} G_{p11}(s) & G_{p12}(s) \\ G_{p21}(s) & G_{p22}(s) \\ G_{p31}(s) & G_{p32}(s) \end{bmatrix} \tag{4.2.15}$$

The six individual SISO models were estimated using simple step tests. To determine the order of each model, a Hankel singular value inspection was performed on each input-output relation initially. The resulting SISO models were either second or third order models. To achieve best fit, several model features were adjusted; model order, inclusion of zeros and imaginary poles.

[16] An impression based on discussions with control engineering researchers in Statoil.

Figures 4.5 and 4.6 illustrates the best model fits for each of the six models. As demonstrated in equation 4.2.15, the six models were then combined to a deterministic MIMO state-space model wich resulted in a 15^{th} order system. However, the MIMO model built on SISO models did not increase fit significantly compared to the MIMO identified models. This does, however, not support a conclusion towards the SISO approach being inadequate, but rather raise awareness on faulty identification and/or validation data due to the poor validation results. This issue is considered further in the next section.

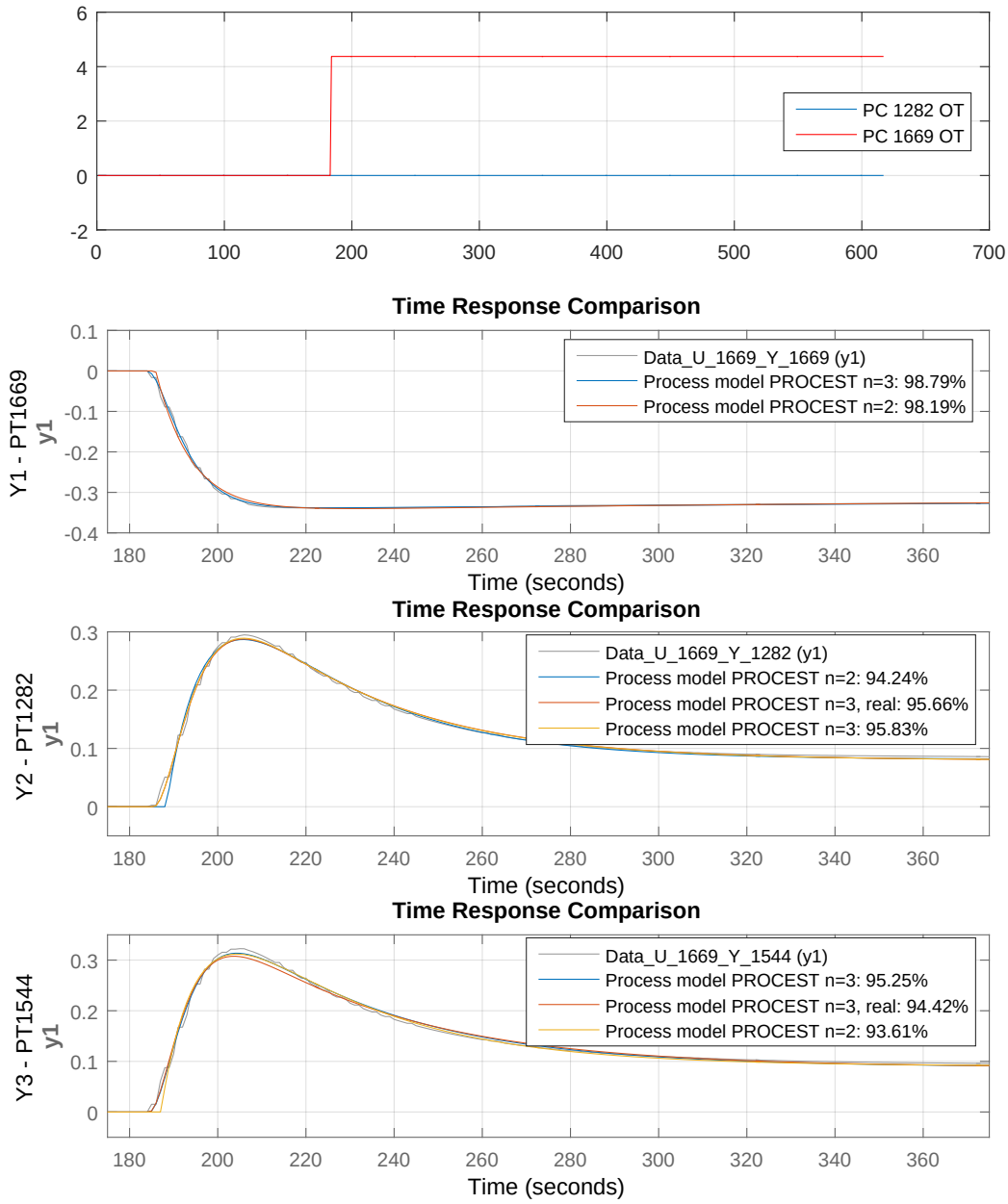


Figure 4.5 Comparison of several candidate models compared to identification data from input U1. A SISO model represent all channel pairings possible, resulting in 3 models for one input to three outputs. The figure legends provide estimation algorithm, model order, n, and if poles only comprises real values, in addition to the goodness of fit value. Same annotations are used for figure 4.6.

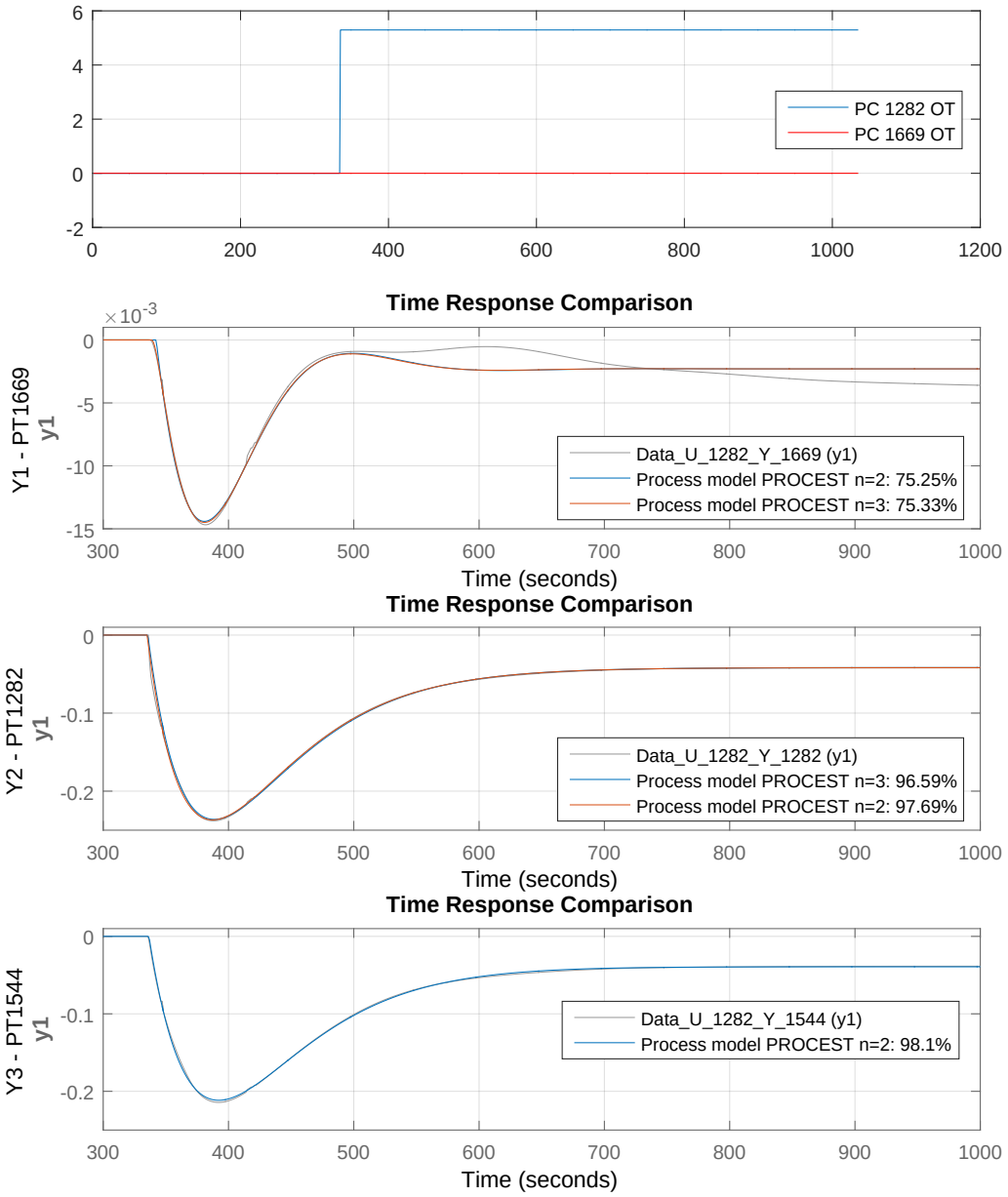


Figure 4.6 Comparison of several candidate models compared to identification data from input U2.

4.3 Practical issues on system identification: some author-based experiences

Given the data obtained in Volden [88] and in initially in this work, an exhaustive number of both open-loop and closed-loop system identification procedures were performed to develop an accurate linear model. Results from these experiments are found in appendices C and D in Volden [88].

Since the system is indeed nonlinear, model discrepancies were expected, but only to a certain degree since perturbations were kept reasonable small. Still, to obtain satisfactory results were difficult, and as it turned out, there was a reason behind. This chapter describes a valuable experience for the author when considering system identification experiments and results; the importance of visual checks of experiment data, and the understatement that even the best algorithm and experiment design cannot replace manual observations.

4.3.1 System identification data discrepancies from Volden [88]

For a full overview over the numerous datasets and success of model validation, the reader should consult appendices C and D in the latest revision of Volden [88]. A short summary which includes some initial work on the continuation of Volden [88] will, however, follow:

- Out of 24 candidate data sets, 14 were selected for the work of achieving a model fit for purpose. Selection were based on suitability and perturbations.
- `N4SID` gave almost consistently better results than `DSR`, and thus, `N4SID` was the only algorithm considered in the further work. Selection were primarily based on fitness results.
- Even though, based primarily on observed behaviour against identification and validation data sets in addition to `NRMSE` values, the results achieved were poor. Algorithm adjustments mentioned in Juricek et al. [40], Söderström and Stoica [70], Tangirala [82], Viberg [87] and Ljung [44, 45] were tested without significant and unambiguous improvements.
- As further described in 4.2.2, a SISO modelling approach were performed.
- Based on the derived models from the data sets in Volden [88], 12 candidate models were compared.

The comparison of the derived models is found in figure 4.7. Although, 12 models were originally derived, in this comparison, the two best state-space models based on closed-loop data and three state-space models based on open-loop data are compared in figures 4.7 and 4.8. The selection of best models were based on plot information and two numerical criteria, namely NRMSE and MVAF. The criteria are further described in section 5.3.

For the sake of clarity, the SRC compressor suction pressure - Y_1, and SRC upstream expander pressure - Y_2, and SRC compressor discharge pressure - Y_3, correspond to PT1669, PT1282 and PT1544, respectively. The two manipulated variables PIC1669 and PIC1282 correspond to the SRC compressor speed and energy consumption, and SRC expander bypass and discharge, respectively. Further details are included in table 4.3 in Volden [88].

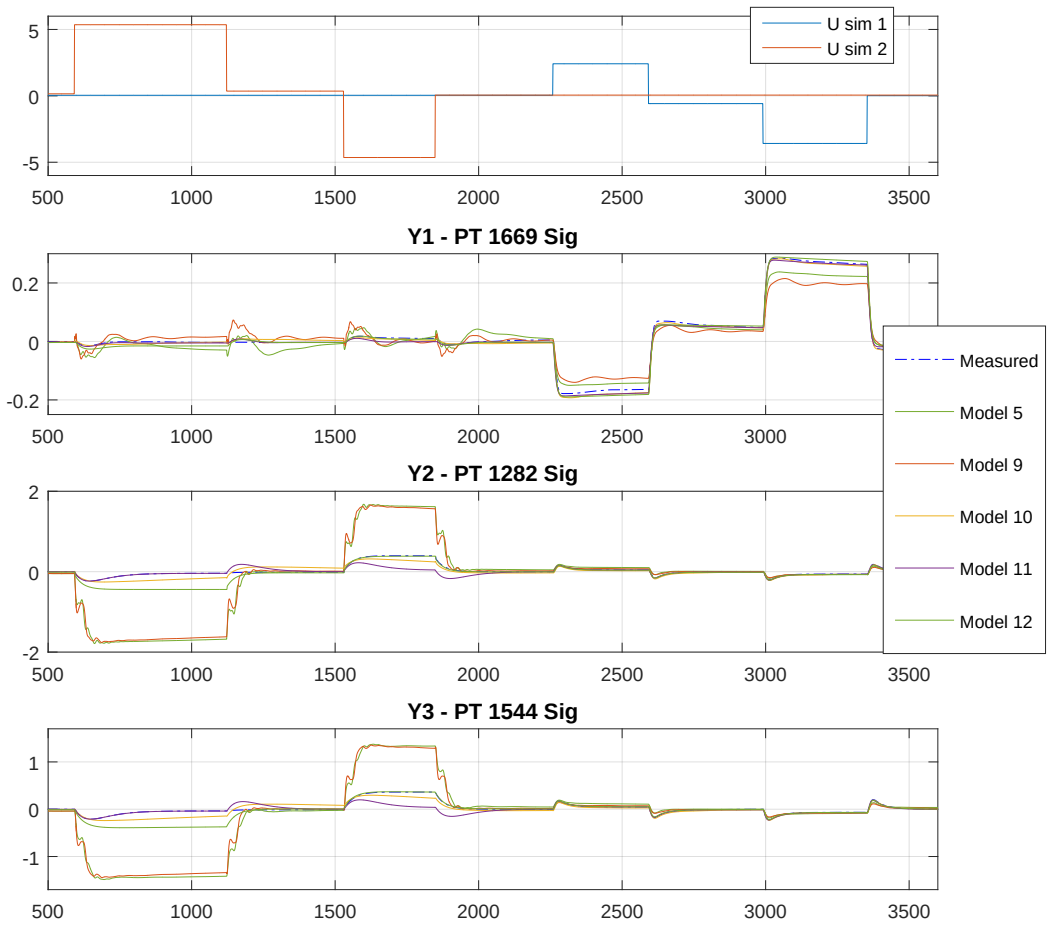


Figure 4.7 Open-loop data comparison of process data and identified models. Model 5 and 9 are closed-loop data identified models, whereas models 10-12 are open-loop data identified models. All models considered here stem from Volden [88].

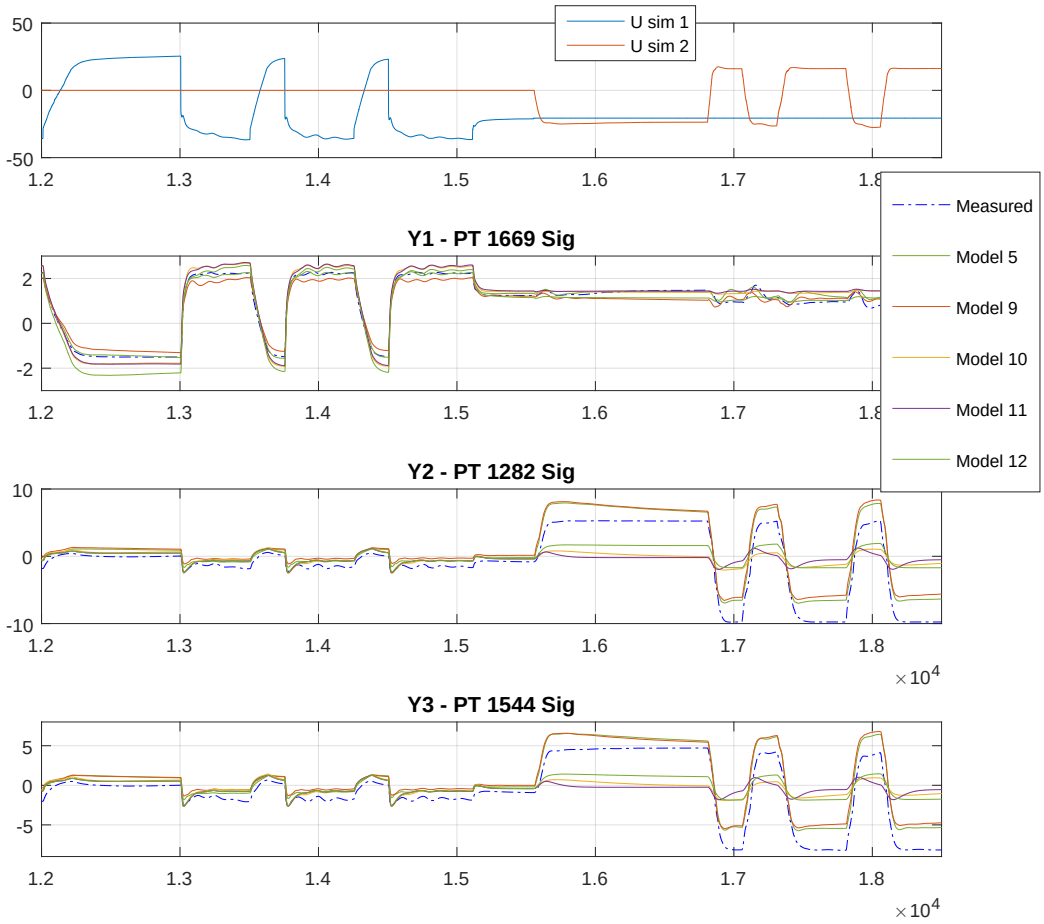


Figure 4.8 Closed-loop data comparison of process data and identified models. All models from Volden [88].

The plots depict a segment of the data to increase detail visibility. By observation, it is easy to see where the models are deviating from the process. Additionally, closed-loop data models perform superior of open-loop data models against closed-loop verification data, whereas the opposite for open-loop verification data. This is expected due to the nature of identification experiments for the different open-loop data identified and closed-loop data identified models. For both inputs excited, Y_1 for all models are somewhat accurate. This is *partly* the case for Y_2 and Y_3. For the open-loop data set, Y_2 and Y_3 are inaccurate for one of the inputs. Looking at the closed-loop data set reveals the same pattern for Y_2 and Y_3. Further examination reveals that the model inaccuracies are related to a specific input, the expander pressure controller PIC1282.

Figure 4.9 depicts the open-loop data set. Notice that Y_1 are almost unaffected by PIC1282, and respond reasonably for PIC1669 manipulations. This indicate a weakly coupled system for PIC1282 $\rightarrow$ $Y1$. For Y_2 and Y_3, the responses to PIC1669 manipulation are reasonable, although the responses does not resemble an ordinary open-loop setup. This is due to an external compensation of the pressure, which causes the state to "move back" to its former operating point, and is not considered a modelling issue. For the sake of brevity, this section will not investigate this further, but see section

7.1 for an explanation. Additionally, the closed-loop data models performed superior compared to the open-loop data models against closed-loop validation data. For open-loop validation, the open-loop data models were most accurate. This was as expected due to the differing framework which the models were identified within.

On manipulating PIC1282 in figure 4.9, some discrepancies for Y_2 and Y_3 are clearly observed. On the first step up, the response is reasonable, but on the same step down, the response is quite different and almost non-existing. The next steps up and down additionally reveals different gain in different directions for the process, which clearly indicate nonlinearities and different gains.

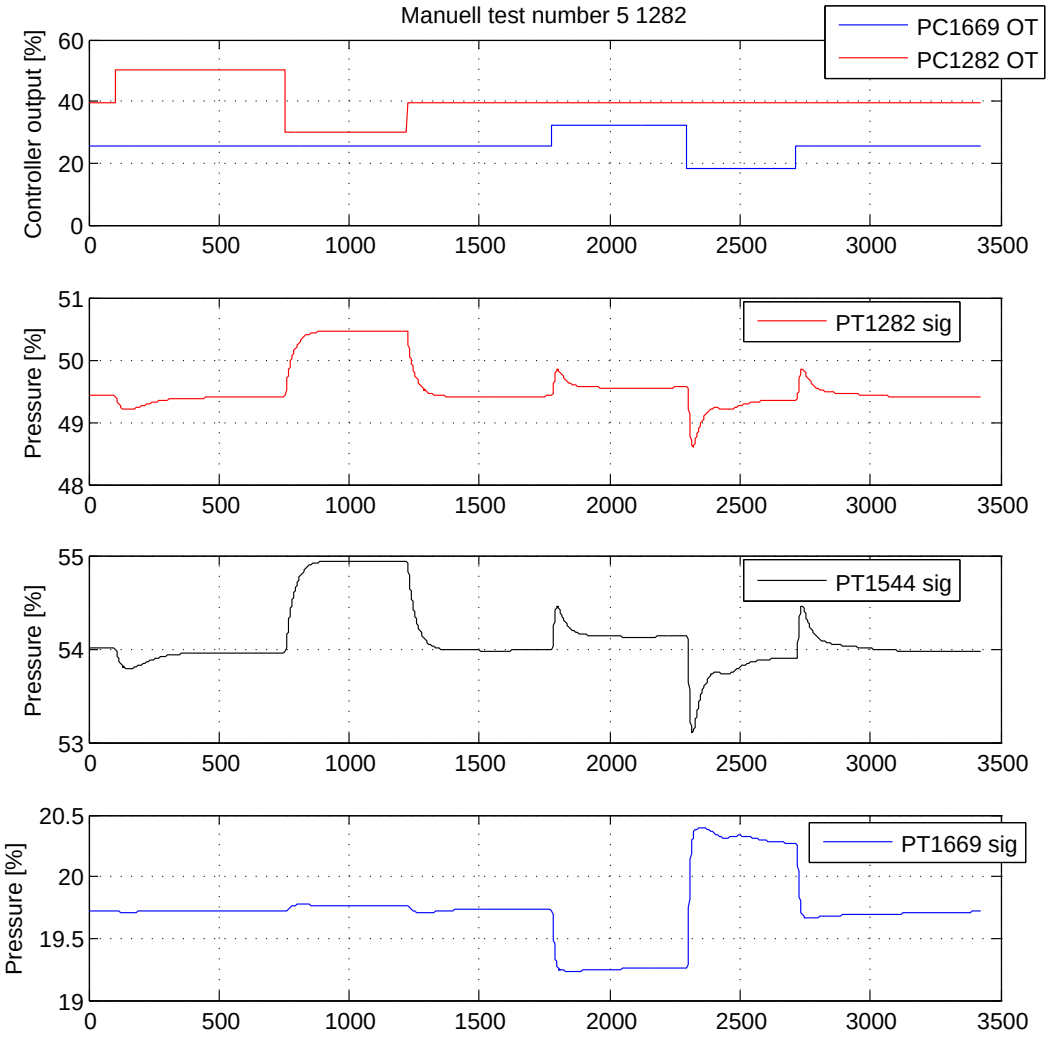


Figure 4.9 Open-loop data set for the 2x3 system

4.3.2 Altering the identification procedure to improve modelling results

The nonlinear phenomena mentioned in previous section was investigated, and as it turns out, the feature is due to a selector switching controller signals within the control structure. Because of this, all system identification experiments were rendered faulty, and accordingly, all identified models were considered invalid.

The figures 4.10 and 4.11 are intended to illustrate this altering of the control structure. A closed-loop experiment is executed where the setpoint of PIC1282 is manipulated. In these figures, a new manipulated variable - PIC1281 - and a new controlled variable - PT 1281 - are introduced. These variables constitute the extended system to the initially investigated 2x3 MIMO system, which now becomes a 3x4 system.

The aforementioned selector switches between either PIC1282 or PIC1281 outputs, and consistently chose the smallest input. In addition, PIC1282 is a split range controller[17], where $0 \rightarrow 50\%$ output controls a primary valve $0 \rightarrow 100\%$, and $50 \rightarrow 100\%$ output controls a secondary valve $0 \rightarrow 100\%$. To render this feasible, the controller output is multiplied by 2. This feature is illustrated in figure 4.10 and thoroughly described in figure 4.11, which depict a closed-loop experiment where PIC1282 setpoint is manipulated. Additional plots and information on these features are mentioned in appendix E.3.

Note that by observing the top left subplot in figure 4.10, the closed-loop response does not reveal the mentioned features. One might believe, just by observing the subplot, that the closed-loop experiment is valid because of the seemingly sensible response, which is *not* the case. This was a misleading and contributing factor to the faulty initial identification experiments.

[17] See e.g. Volden [88] and references therein for a brief introduction to split range control.

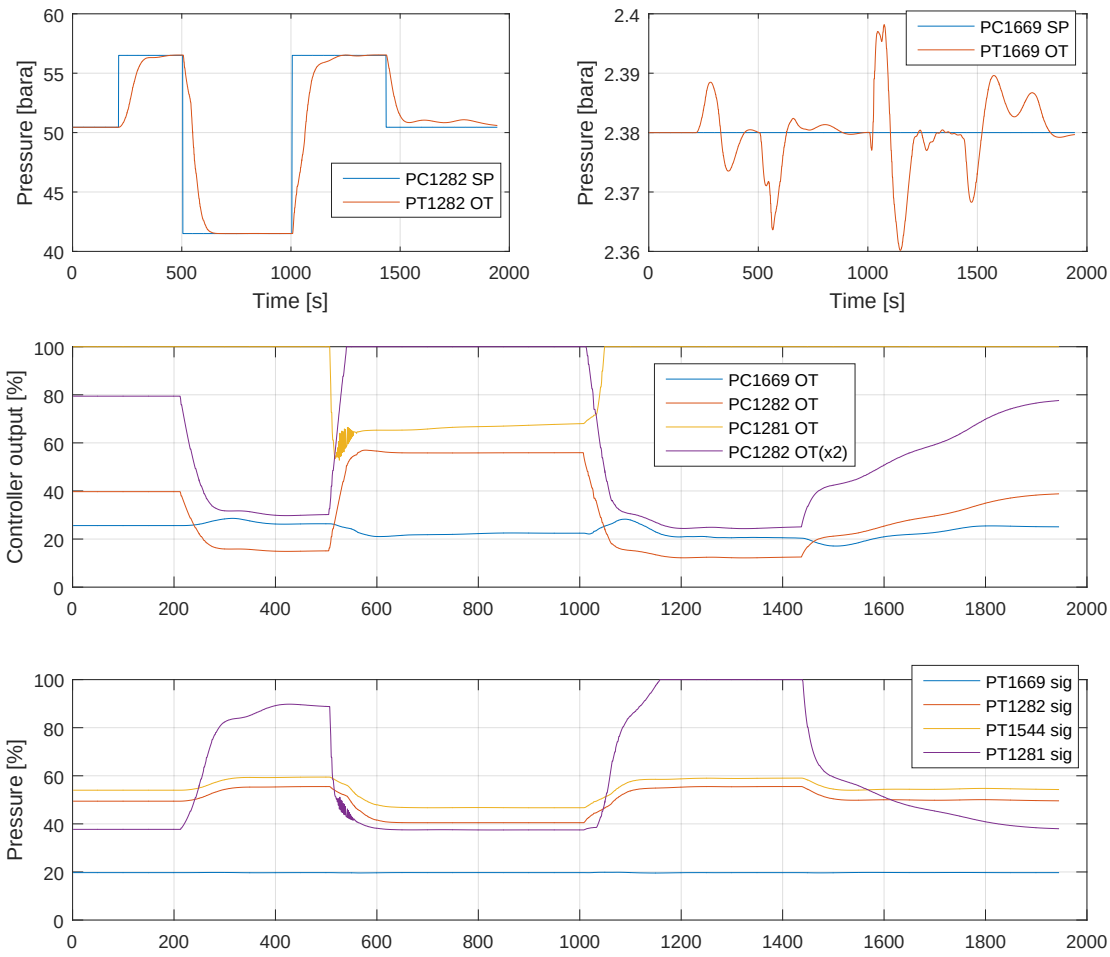


Figure 4.10 Closed-loop experiment for the new 3x4 MIMO system which depicts selector control.

An additional feature is the predictive response of the PC1281 controller. As observed in bottom subplot of figure 4.10, the pressure PT1281 does not violate its operational setpoint below 39% prior to controller action. This is not the case for a conventional PI controller, where one indeed would expect some violation of setpoint prior to controller action. To achieve this, there are several controller configurations available. In this case, the controller is operating in velocity form, which imply that the controller respond to *changes* in the value of the controlled variable, rather than its *actual* value at a given point in time. For a discrete PI controller on position form the control law is:

$$u_k = \bar{u_0} + K_p \left[e_k + \frac{\Delta t}{\tau_I} \sum_{j=1}^{k} e_j \right], \tag{4.3.1}$$

where $\bar{u_0}$ is the controller bias(steady-state) output and Δt is the sampling interval. For a PI controller

on velocity form, the control law is:

$$\Delta u_k = u_k - u_{k-1} = K_p \left[(e_k - e_{k-1}) + \frac{\Delta t}{\tau_I} e_j \right], \tag{4.3.2}$$

From the integral term in equation 4.3.2, it is clear that the velocity form incorporates anti-windup inherently[18]. This feature is observed in both figures 4.10 and 4.11.

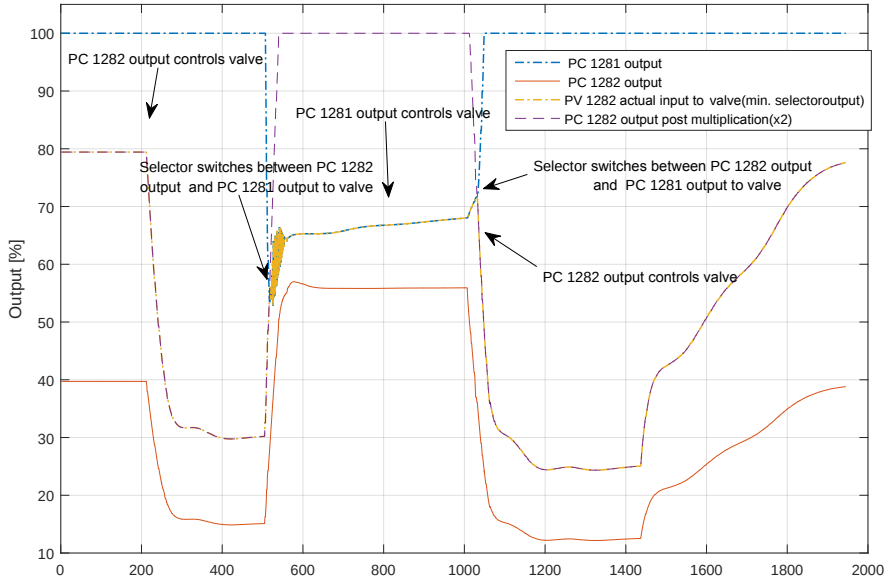


Figure 4.11 Controller outputs: Explanation of the selector switching control.

As mentioned, these findings result in a new perspective on the control system and renders all system identification experiments invalid. To familiarise the reader, a simplified sketch of the control system including the relevant control loops are depicted figure 4.12.

[18] Additional information on velocity form control is found in e.g. Seborg et al. [64].

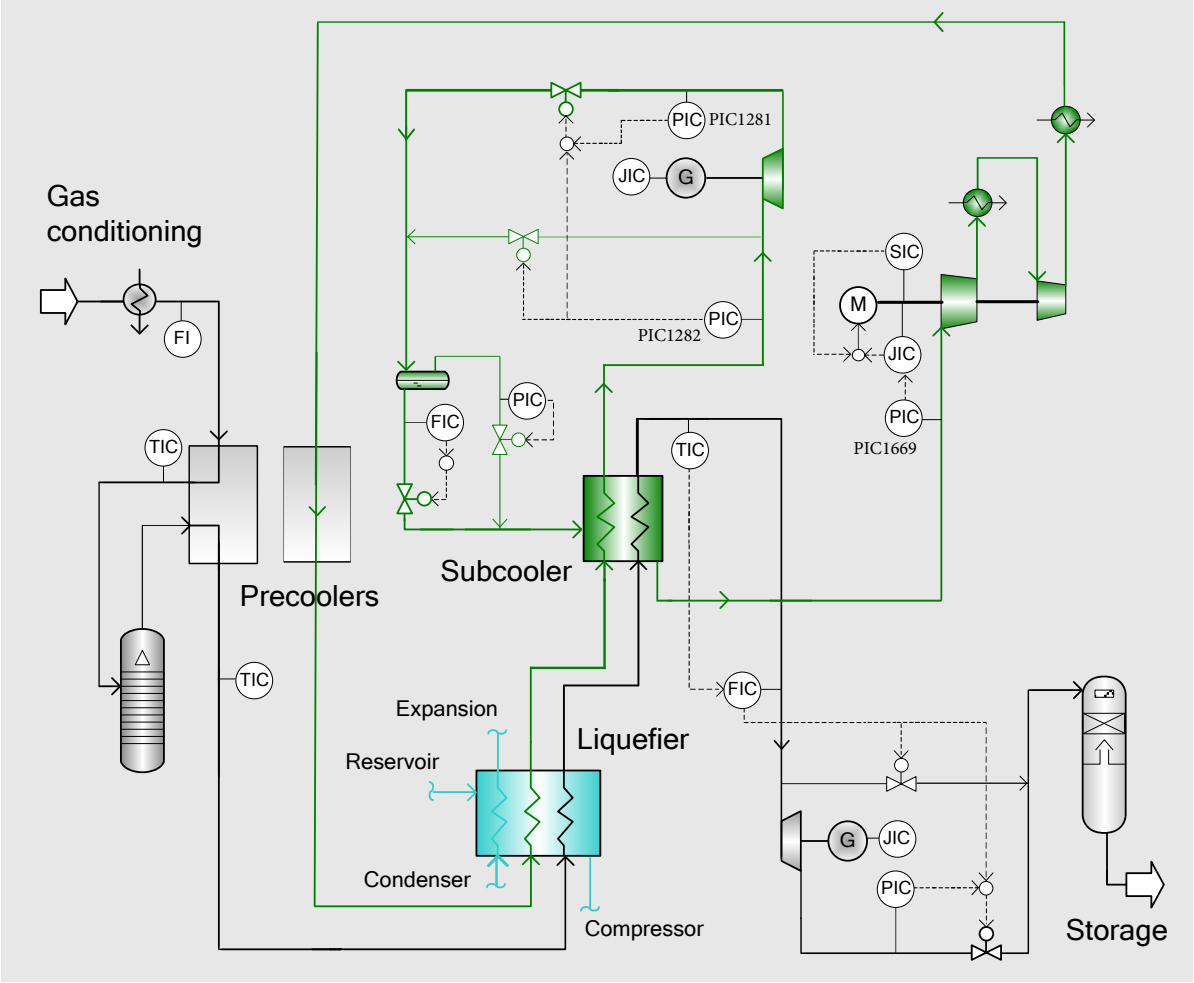


Figure 4.12 3x4 MIMO control structure of the subcooling cycle

Dealing with a selector configuration complicate the modelling process. A selector is obviously a nonlinear feature, and it additionally alters the model when switching between candidate variables. One option for handling a selector is to utilise multiple models which correspond to selector settings, i.e., one model for each selector candidate output and alter between these models depending on state of the selector. A second option which coincides with aforementioned scheme is to use mixed integer linear programming(MILP) in the work of defining a suitable optimization problem formulation for the MPC. A third option is to alter the control system and thus circumvent the use of selector.

In this project, the option of altering the control system is chosen. The reason behind this is to keep within the project scope as including MILP in the problem formulation is cumbersome and challenging. However, it is not unusual to alter the control system when implementing MPC, since certain control features are advantageous to include in the MPC. To omit the selector, the PIC1281 controller is not considered, i.e., set to manual and set to 100% output. Hence, the selector will consistently choose the smaller controller input from PIC1282. This renders a 2x4 MIMO system,

where the last controlled variable is the pressure downstream of the expansion turbine. This pressure provides a lower constraint on pressure downstream the turbine to avoid flashing of the liquid. Liquid entering a two-phase state may cause cavitation and thus harm the turbine. To account for the lost degree of freedom by setting PIC1281 to manual operation, the downstream turbine pressure is incorporated in the model with its constraints included in the optimization problem.

4.4 Performing new system identification experiments

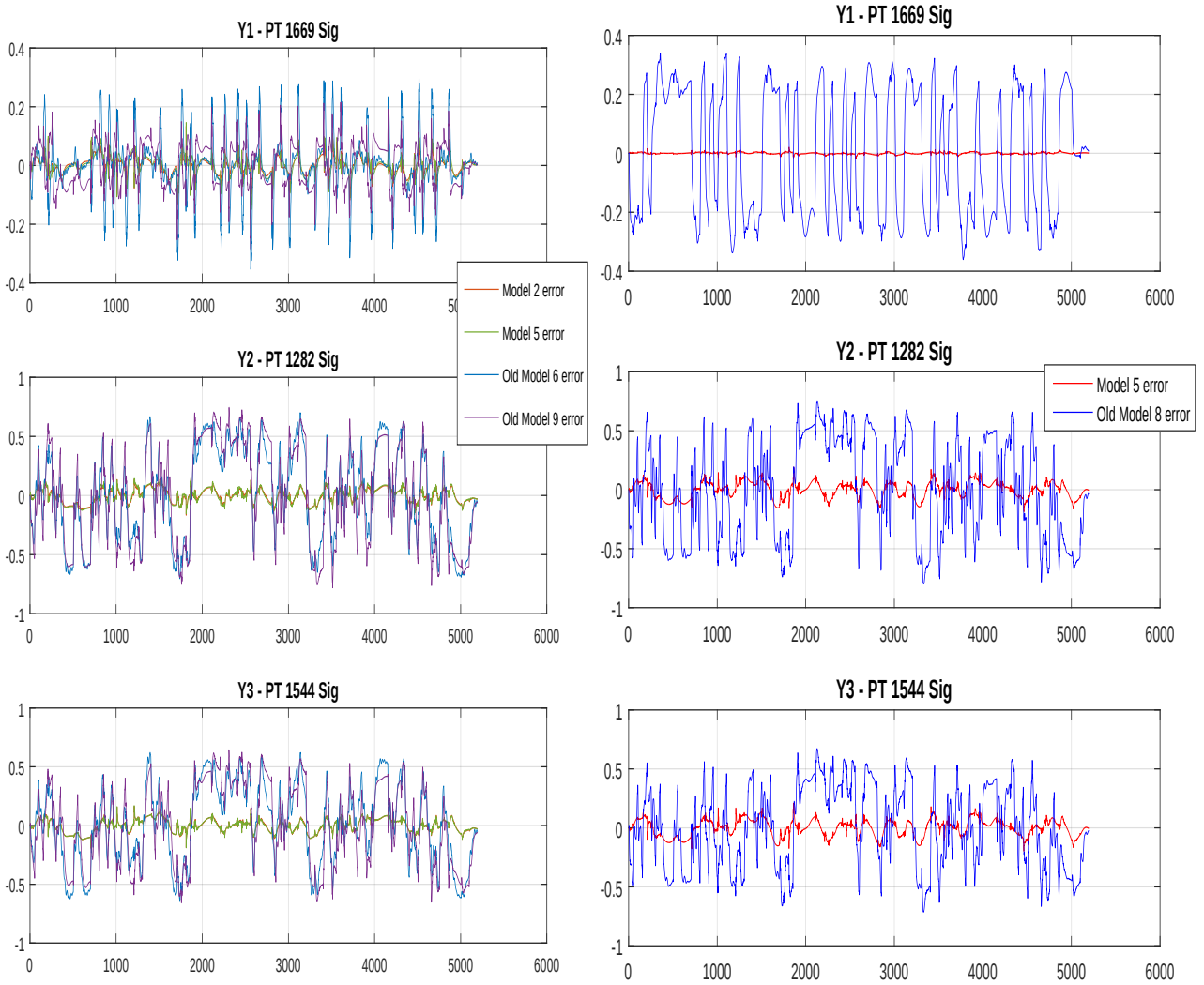
Based on the new information, an additional iteration of subspace identification experiments were conducted. The experiments were conducted in open-loop configuration in accordance with Statoil standards, but also due to its user-friendliness. Both simultaneously excitation and channel-wise excitation experiments were executed. As mentioned in section 4.3.2, the selector control were omitted by setting controllers to manual and increase output signal to 100%, keeping the selector state steady. In accordance with section 3.2.2, some preliminary step tests were performed to ensure integrity of the system in focus for the identification procedures. A selection of the various system identification experiments are included in appendix B.

Several 2x4 candidate models were derived and validated based on the new system identification data, and the results were unambiguously improved compared to the old models. In addition to MIMO models, a similar principle described in section 4.2.2 were utilised to provide a MIMO model based on SISO identification data. The resulting SISO data developed MIMO model has a relatively high order of 20 compared to the MIMO models which vary in order from 7 to 19.

A comparison between the old and new models is depicted in figure 4.13. Since the old models were 2x3, the 4^{th} output from the new 2x4 system is neglected for the sake of comparison in this case. Two old models and two new models were compared, and an error plot for the three particular outputs is included along with the numerical fitness values:

Table 4.3 Numerical fitness values for two old and two new models.

	Fit Y1	Fit Y2	Fit Y3	Mean fit	MVAF Y1	MVAF Y2	MVAF Y3	Mean MVAF
Model 1	91.9	82.5	84	86.1	99.4	96.9	97.5	97.9
Model 2	90.3	80.5	84.3	85	99.1	96.2	97.5	97.6
Old model 1	71.6	-16.9	5	19.9	92.3	-40.8	7.5	19.7
Old model 2	78.7	-27.2	3.3	18.3	95.5	-64.3	4.9	12.1



(a) Error plots for 3 CV's in old and new models

(b) Comparison of new candidate model and an old model developed on faulty data

Figure 4.13 Error plots for comparison of old data models and new data models

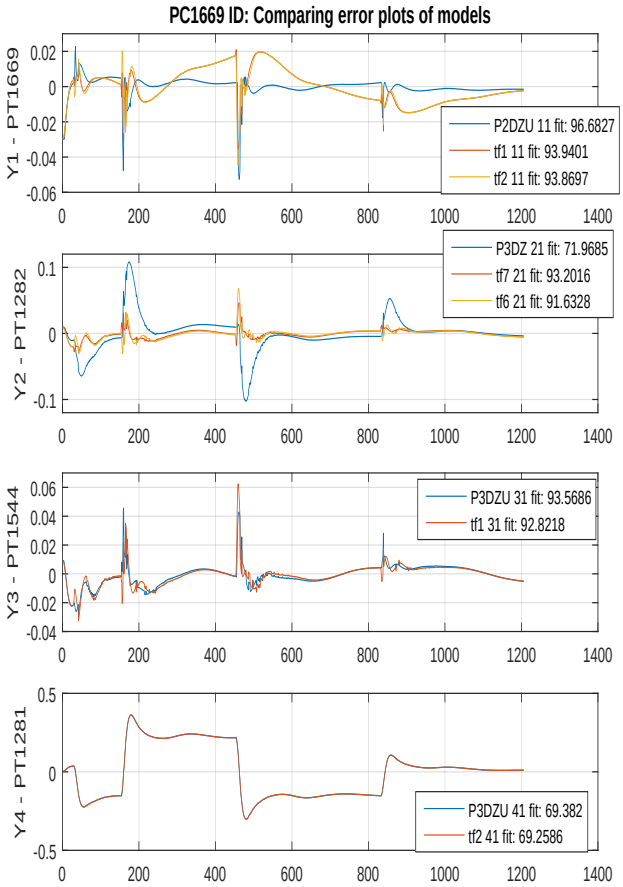
Both graphically and numerically, the new models were far better compared with the old models. This obviously support the findings in section 4.3.2, which were expected. Based on these results, the work of assesing and analysing suitable models for the MPC application were initialized.

A comparison of the SISO developed MIMO model compared to the directly identified subspace MIMO models was additionally performed. Since the system investigated is 2x4, a fourth CV is added. Since the system was expanded with an additional CV, two additional SISO models for the fourth ouput were derived. Thus, we have the MIMO model acquired by the SISO data models as:

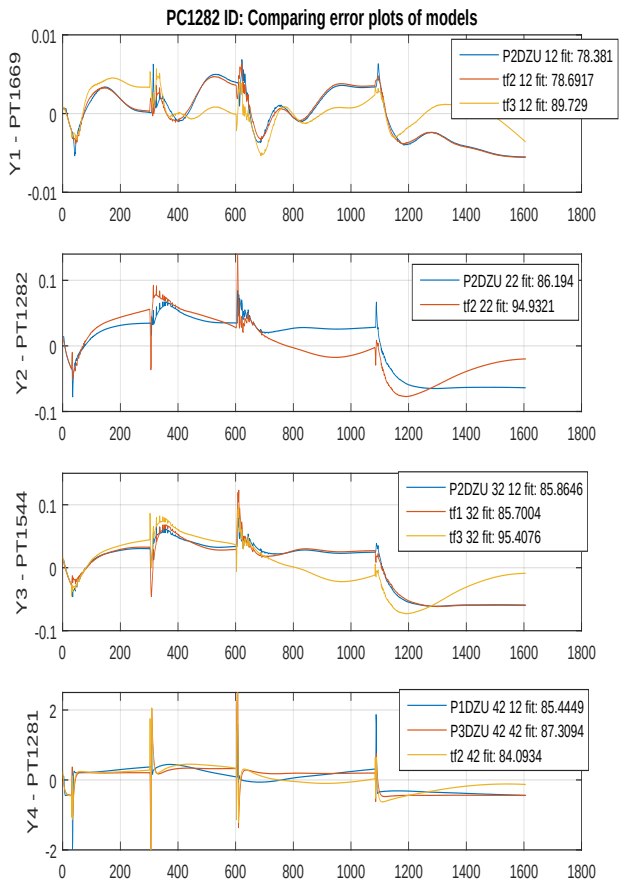
$$\underbrace{\boldsymbol{Y}(s) = \boldsymbol{G}_p(s)\boldsymbol{U}(s)}$$
$$\Downarrow$$
$$\boldsymbol{Y}(s) = \begin{bmatrix} Y_1(s) \\ Y_2(s) \\ Y_3(s) \\ Y_4(s) \end{bmatrix} \quad \boldsymbol{U}(s) = \begin{bmatrix} U_1(s) \\ U_2(s) \end{bmatrix} \quad \boldsymbol{G}_p(s) = \begin{bmatrix} G_{p11}(s) & G_{p12}(s) \\ G_{p21}(s) & G_{p22}(s) \\ G_{p31}(s) & G_{p32}(s) \\ G_{p41}(s) & G_{p42}(s) \end{bmatrix} \tag{4.4.1}$$

For the sake of clarity, Y_1, Y_2, and Y_3, are defined the same way as in section 4.3.1. The new CV, Y_4, is the downstream expander pressure - PT1281.

All channels were individually investigated, and several SISO modelling procedures were carried out. The best SISO models for each channel were chosen based on visual comparison and a numerical fitness value. The resulting MIMO system based on SISO models performs satisfactory on validation data. The figures 4.14a and 4.14b depicts the SISO models simulated against a validation data set.



(a) Error plot for SISO models from input PIC1669 onto four CVs



(b) Error plot for SISO models from input PIC1282 onto four CVs

Figure 4.14 Error plots for the different SISO data individual candidate models

From several candidate MIMO data models, only the most accurate model were chosen for comparison against the SISO data model. The fitness values are listed below. Figure 4.15 depict an error plot. As observed, both numerically and graphically, the validation differences were small, and both models perform satisfactory.

Table 4.4 Fitness values for validation of MIMO data model and SISO data model comparsion

	Fit Y1	Fit Y2	Fit Y3	Fit Y4	Mean fit
MIMO data model	91.5	82.2	83.6	75.8	83.3
SISO data model	91.7	84.5	84.8	74.3	83.8
	MVAF Y1	MVAF Y2	MVAF Y3	MVAF Y4	Mean MVAF
MIMO data model	99.3	96.9	97.3	94.1	96.9
SISO data model	99.3	97.6	97.7	93.4	97

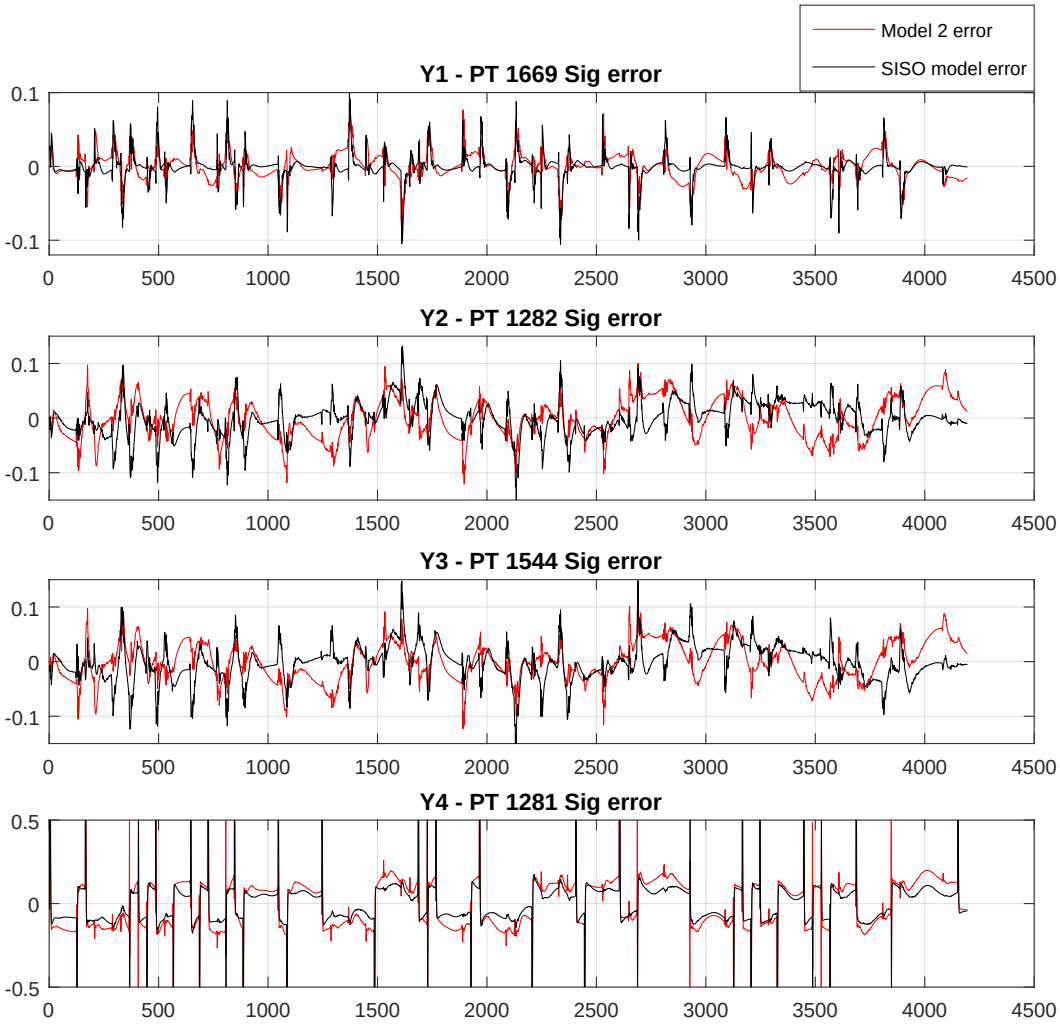


Figure 4.15 Error plot comparison of SISO data and MIMO data candidate models.

Chapter 5

Results from system identification experiments and model analysis

To evaluate and conclude on the best models, several measures were investigated. Based on the aspects mentioned in chapters 4.2.1, comparisons of different candidate models were performed. This chapter aims to provide a thorough graphical analysis by utilizing three measures; the `NRMSE` and `MVAF` fitness criteria, and error plots comparing model responses against process simulator data. In addition, model uncertainties and residual analyses provide a wide base for model assesment. The simulator data utilized are retrieved from PRBS experiments similar of those in appendix B.1. Thus, excitation is carried out using input channel excitation individually and combined. Based on the approach mentioned in section 4.2.1, systems of an increasing order were compared. In closing, the degree of interaction is investigated.

The first part of this chapter aims to provide a thorough graphical analysis by utilizing three measures; the `NRMSE` and `MVAF` fitness criteria, and error plots comparing against simulator data. Although, the model discrepancies can be due to noise or model errors, and only the combined effect is observable. further, it is worth noting that the numerical results based on error plots are tailored for each individual comparison. Hence, no general features of the various results are applicable.

5.1 Numerical fitness measures

To investigate the feasibility of the models, numerous tests against the selected data sets from the system identification procedures were performed. Numerous MIMO models from the MIMO identification and the MIMO model from SISO identification were simulated against the controller outputs from the selected data sets and compared to the simulator output for the actual data set. In addition to a graphical comparison, the normalized root mean squared error(NRMSE) from Ljung [44] and the mean variance-accounted-for(MVAF) from Sotomayor et al. [71] were investigated to

indicate the best models:

$$NRMSE = 100 \times \left(1 - \frac{||y - \hat{y}||}{||y - mean(y)||}\right) \tag{5.1.1}$$

$$MVAF = 100 \times \left(1 - \frac{variance(y_i - \hat{y}_i)}{variance(y_i)}\right) \tag{5.1.2}$$

where y is validation data and $\hat{y}$ is model output. The NRMSE criterion provides numerical information on the differences between measured and predicted output from the identified model, both at steady-state and dynamically. The MVAF criterion emphasizes the dynamical differences essentially, and is a more important measure than steady-state deviations, which is demonstrated in Jacobsen et al. [37] and Jacobsen and Skogestad [38].

5.2 Combined data sets

In this section, the simulator data sets are retrieved from combined input channel excitation sequences. Figure 5.1 depict an error plot for comparison. Models of incrementing order, $n = [5, \ldots 30]$, were simulated with the simulator data input signals.

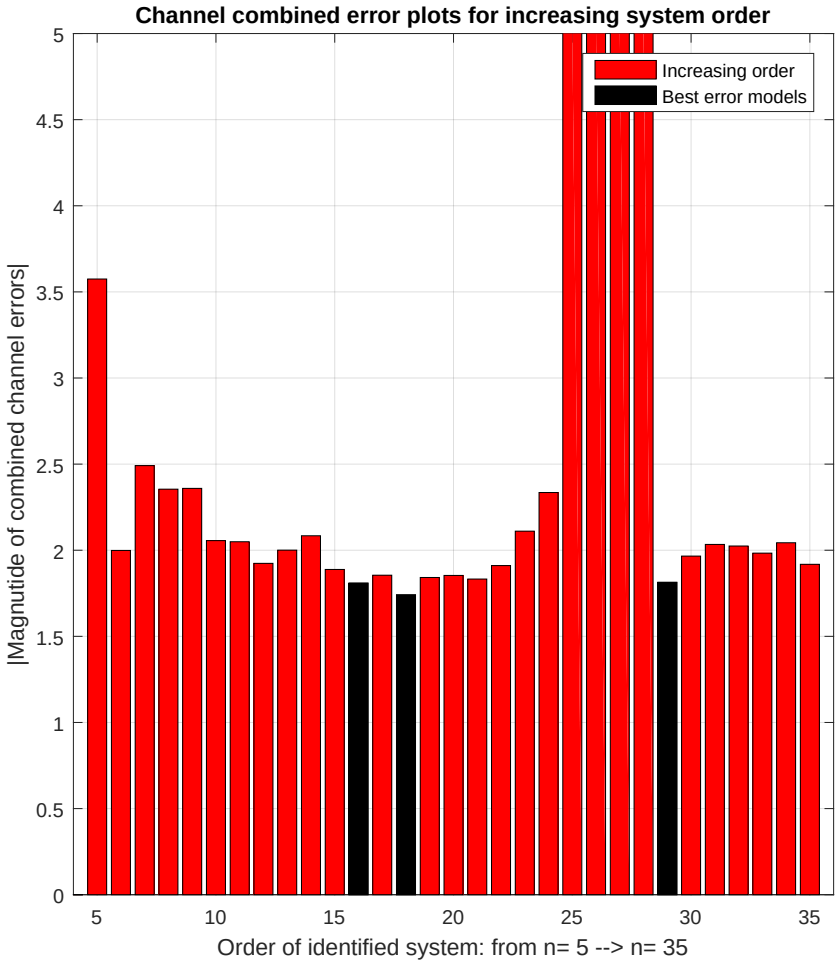


Figure 5.1 Combined input channel excitation data set. Uses AIC chosen horizon. Black bars mark the best results. Combined error values of 5 indicate open-loop unstable models.

For the same models as in above figure, numerical fitness values were calculated. This is depicted in figure 5.2. The results coincide somewhat with the error plot. Thus, models around order $\approx 15 \rightarrow 18$ return high scores on both fitness values and combined errors. Additionally, the error plot indicate a high score for the model order 29. This does not coincide with the fitness values. Hence, this particular model order is not considered further.

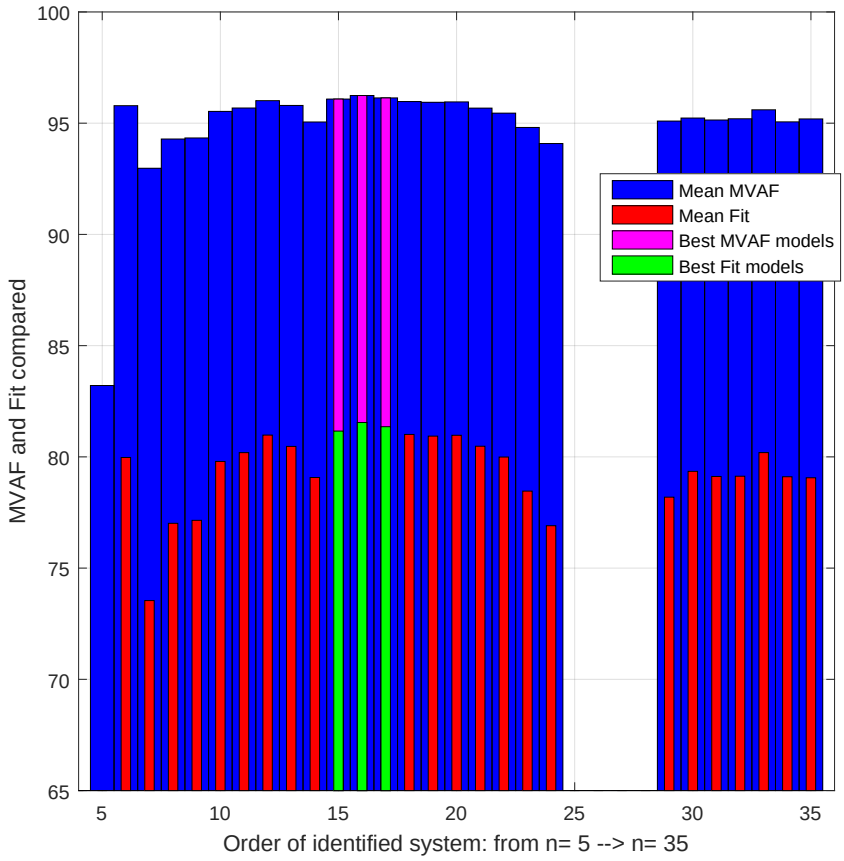


Figure 5.2 Combined input channel excitation data set. Uses AIC chosen horizon. Best results are green and magenta colour marked. Fitness values below 65 indicate open-loop unstable systems as correspondingly illustrated in figure 5.1.

5.3 Individual channel data sets

In addition to data sets which combine input channel excitation, data sets for individual input channel excitation were utilized in the work of obtaining a good model. Figure 5.3 depict the error plot of the comparison in a similar fashion as figure 5.1.

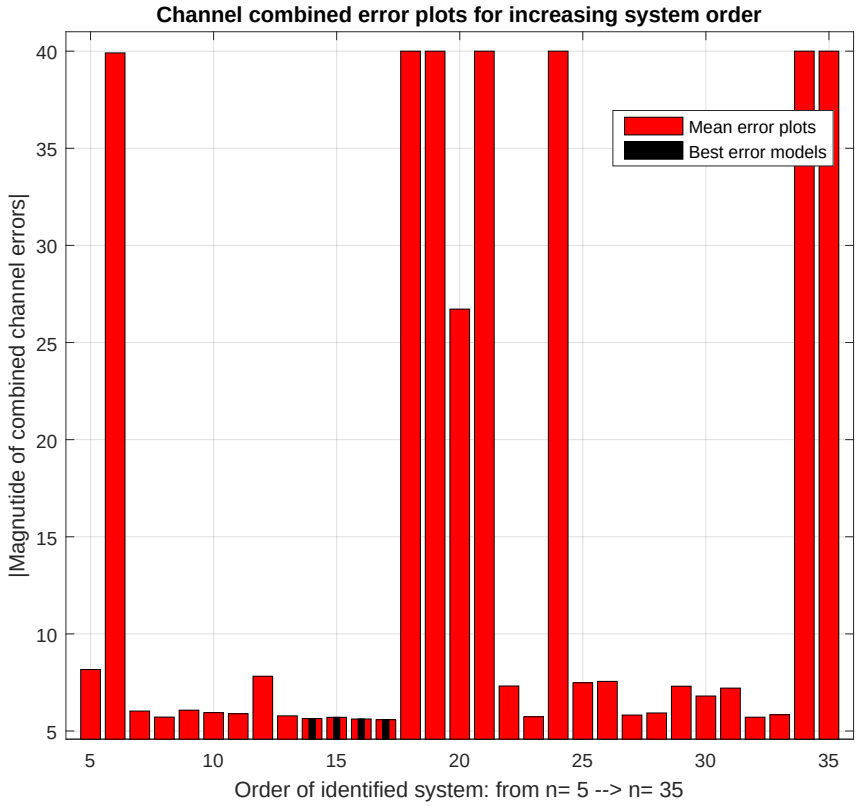


Figure 5.3 Individual channel excitation data sets. Uses AIC chosen horizon. Values of 40 indicate open-loop unstable models.

Accordingly, the numerical fitness values were calculated, and the result is depicted in figure 5.4. For closer comparison, the three best models based on the fitness values are illustrated in figure 5.5.

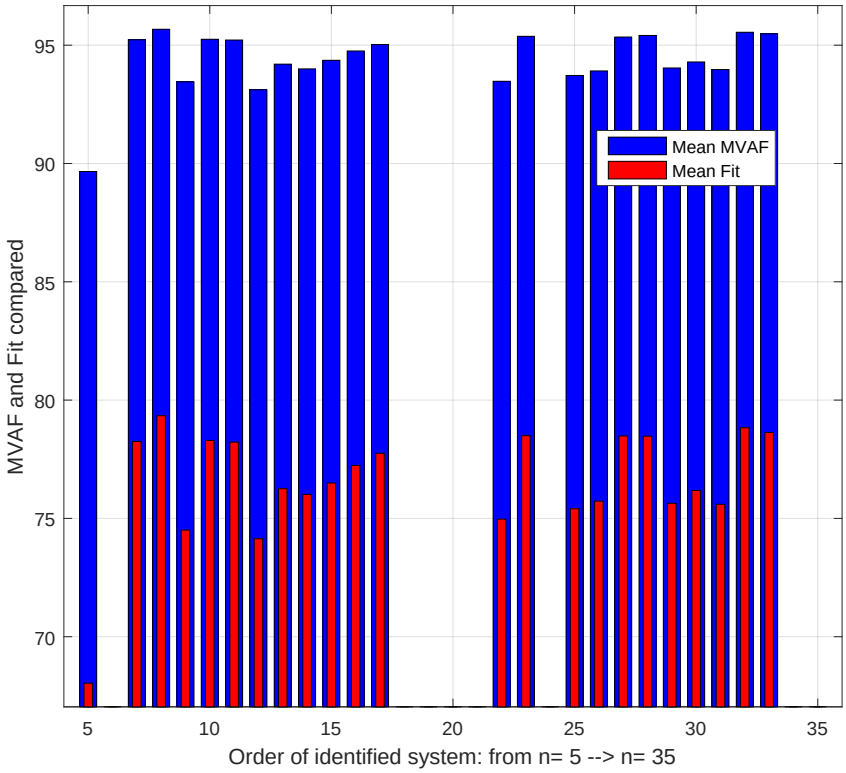


Figure 5.4 Individual channel excitation data sets. Uses AIC chosen horizon. Open-loop unstable systems are removed from the plot.

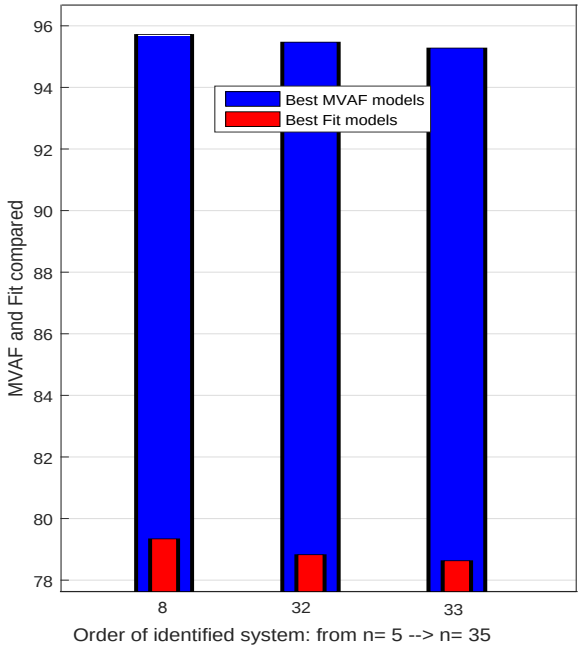


Figure 5.5 Best results from figure 5.4. Uses AIC chosen horizon.

Compared to the fitness results in section 5.2 where the best models were located in the vicinity of

each other, this is not the case here. As depicted in figure 5.5, the best fitness results favour a low order model and two higher order models. However, for the error plot, the results are similar to the results in section 5.2. Thus, models in the order of $15 \rightarrow 18$ are accentuated for the continuing work of model validation. Certainly, these results signify that it may be advantageous to further asses the candidate models using different measures to ensure that the best possible model is chosen.

5.4 Further analysis of model development and suitability

To evaluate and conclude on the best model, additional measures apart from the criteria considered in sections 5.2 and 5.3 were investigated. Based on various approaches, comparisons of different candidate models were performed. These measures utilizes statistically anchored approaches for validation. All comparisons are based on the three candidate models chosen from sections 5.2 and 5.3.

5.4.1 Model uncertainty

In the work of developing the model, an important issue is to evaluate the modelling uncertainties. Model variability is estimated from random disturbances in the output, and describes the alteration of model parameters when estimation is repeated using a different data set and the same model structure. Large uncertainties in model parameters may originate from high model orders, inadequate excitation, or poor signal-to-noise ratio in the data.

Uncertainty in the model is labelled model covariance, and this information is stored in the covariance matrix which originate from the estimated parameters stored with the model. The covariance matrix is used to provide all uncertainties in the model output. Computing the covariance matrix is based on the assumption that the model structure provides an accurate description of the process dynamics, and one method to determine whether the estimated model uncertainty values are reliable is to perform a residual analysis. This is further discussed in section 5.4.2.1.

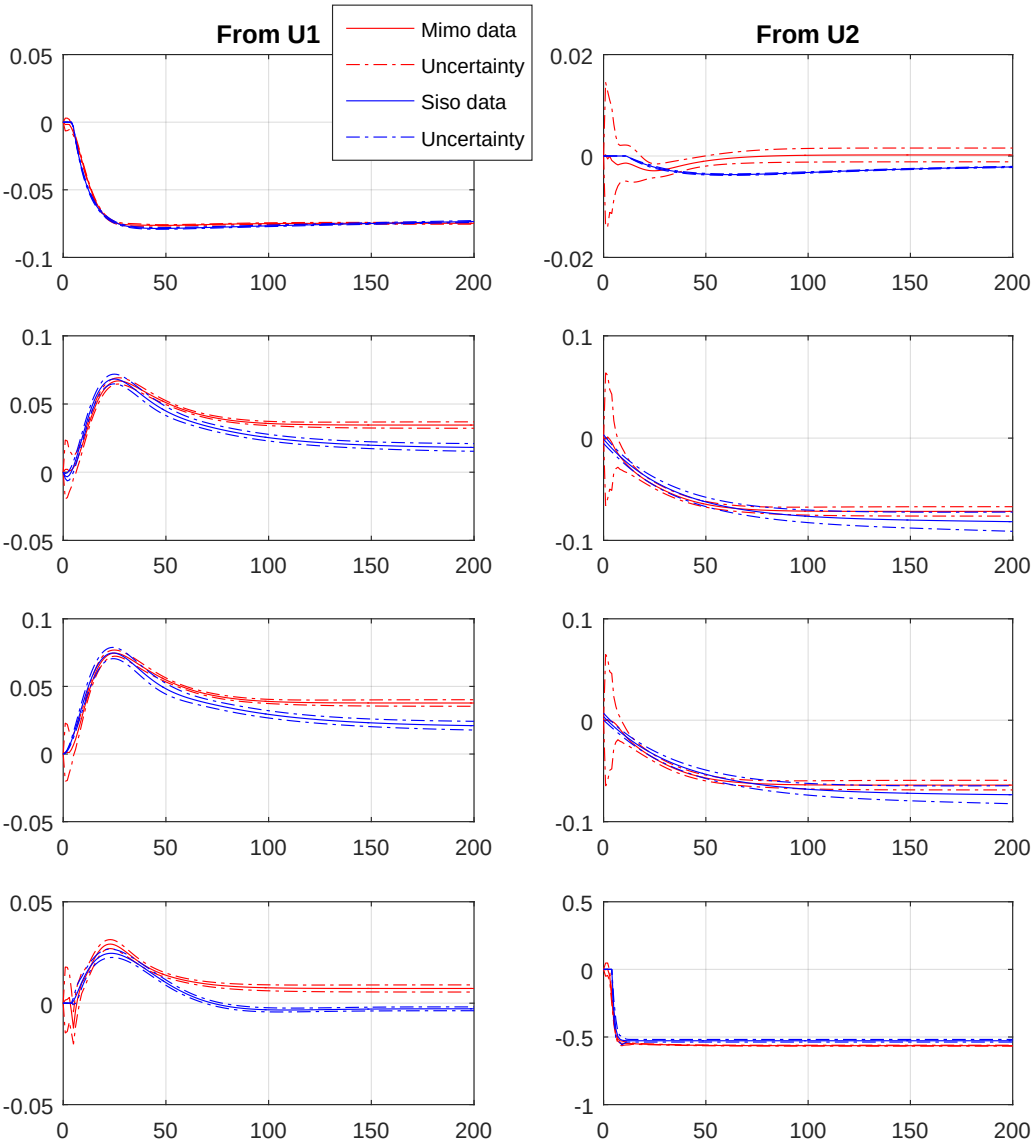


Figure 5.6 Uncertainty plot for MIMO data model and SISO data model. Confidence interval is 3σ.

Besides performing a residual analysis for model uncertainty, a useful method may be to utilize the standard deviations present. In this regard, a useful procedure is to check whether the current model contains too many parameters by comparing the estimates with the corresponding estimated standard deviations. Although, this check is less difficult to interpret if the corresponding parameter reflects a physical structure e.g. such as time delays. If the estimated standard deviations are all large, the information matrix is close to singular. This is an indication of too large model orders.

The model uncertainties depicted in figure 5.6 converges fast and its steady-state uncertainty is well within acceptable limits. Although, for the MPC application it is emphasized that the dynamical part of the model is most important. However, to obtain a small confidence interval is not trivial and varies dependent on a number of factors. Some additional, and more uncertain results are depicted

in appendix E.1.

5.4.1.1 Choosing appropriate order of the dynamical system

Prior to subspace model development, there is one vital parameter to consider: the system order. For the algorithms considered in this work, this parameter is the only input required from the user apart from data sets. Additionally, both the `N4SID` and `DSR` algorithm provide some general guidelines in terms of choosing a proper system order. This can be observed in figure 5.7. To make these guidlines general, the criteria behind the proposed order is purely statistical. As shown in the example in appendix C.2, this may yield erroneous or inconclusive system order estimates. Thus, the author have chosen a more time consuming manual approach to estimate the system order.

When considering subspace identification methods, the input-output matrix equation below is the basis, and all subspace methods initially starts here:

$$Y_f = \Gamma_i X_i + H_i^d U_f + H_i^s M_f + N_f \tag{5.4.1}$$

where we recognise the extended observability and state sequence matrices as Γ_i and X_i, respectively. Further, H_i^d is the deterministic lower block triangular Toeplitz matrix, and H_i^s is the stochastic lower block triangular Toeplitz matrix. U_f and Y_f are the future input and output block Hankel matrices, respectively. M_f and N_f are the future block Hankel matrices formed with process noise and measurement noise, respectively. The input-output matrix equation states that the block Hankel matrix containing the future outputs Y_f is linearly related to the future input block Hankel matrix U_f and the future state sequence X_i.

The first step of any subspace method is to perform a weighted projection of the row space of the Hankel matrices. In the second step, the system matrices are computed. There are basically two classes of subspace algorithms in this respect; algorithms that use the extended observability matrix Γ_i to obtain the state-space matrices, and those using the estimated state sequence $\hat{X}_i$.

To estimate the system order, the matrix input-output equation is used to obtain the extended observability matrix and state sequence, Γ_i and X_i. An estimate of the term $\Gamma_i X_i$ is found utilizing projection, and there are basically two projections utilized in the reviewed algorithms; orthogonal and oblique projections. Further, the term $\Gamma_i X_i$ is a rank deficient term of the system order. This mean that knowledge of the term $\Gamma_i X_i$ enables Γ_i, X_i and system order n to be extracted by simply utilising SVD. We have that $\Gamma_i \tilde{X}_i = O_i$ where $\tilde{X}_i$ is denoted the forward Kalman filter state sequence, and O_i is defined:

$$O_i \triangleq Y_f /_{U_f} W_p \tag{5.4.2}$$

where W_p is the past outputs and inputs, and Y_f are the future outputs. Finding the order is, however, dependent on chosen algorithm, and it certainly differs. One may believe that what seemingly is a

straight forward algebraic computation renders identical results, this is not the case as the author will elaborate further.

The computed Hankel singular values reflect the joint controllability and observability of the states of a realization(Skogestad and Postlethwaite [69]). For visualisation, note that logarithmic plots are preferred in order to distinguish magnitude differences between the Hankel singular values. This is depicted in figure 5.7, where two subspace algorithms; `DSR` and `MOESP`, have computed the Hankel singular values for a data set from the MFC process simulator. The observed differences are partly due to altering weighting schemes and, as mentioned, different projections utilised in the algorithms. However, when the matrix O_i is obtained, the method to estimate the system order is identical for all reviewed algorithms:

Definition 5.1. *The order of the identified state-space system is equal to the number of nonzero singular values in the equation*

$$W_1 O_i W_2 = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{5.4.3}$$

where $W_1 \in \mathbb{R}^{li \times li}$ and $W_2 \in \mathbb{R}^{j \times j}$ are the pre-defined weights which are dependent on algorithm.

In this section, additional details on subspace methods are not considered. However, some definitions related to the first step in subspace identification are included in appendix C.5.

Based on mentioned factors for choosing a proper model structure, the manual approach outlined in section 4.2.1 was employed. To ascertain the best model, the different order models were compared using an independent data set.

Essentially, the choice of model order is based on singular value decomposition of the Hankel matrix, but there are other methods which could provide additional information. This involves tests for pole-zero cancellations(see section 5.4.1.2) and singularities in the Fisher information matrix.

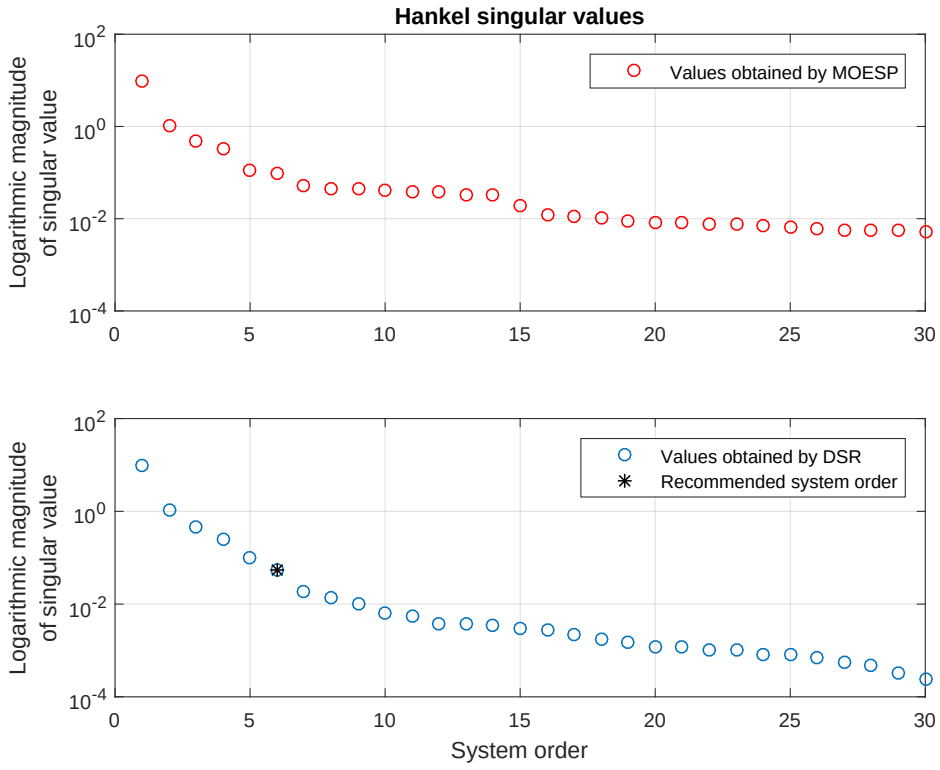


Figure 5.7 Computation of Hankel singular values for based on simulator data. Two algorithms are utilised for the sake of comparison; `MOESP` and `DSR`. Additionally, recommended system order is depicted using the criterion for estimating model order utilized in the `DSR` algorithm.

A simple, yet proven effective practice when considering the Hankel singular values is stated in Juricek et al. [40]: *The order is selected ad hoc by the user, usually looking for a "knee" in the plot of singular values, or by noting where the singular values fall below a specified critical value.*

The choice of parameters depends heavily on the chosen model structure. For an empirical model where the number of inputs m and outputs l are known, and the system order n is set, the number of free parameters for an canonical parametrisation requires $n(2l + m) + ml$ parameters(Viberg [87]). For example: considering the 2x4 system investigated in this work and assuming a model order $n = 10$ yield 108 free parameters.

5.4.1.2 Potential of reduced model

Reduced order models can simplify analysis and control design compared to higher-order models. Additionally, simpler models are easier to understand, physically interpret and manipulate. Higher-order models obtained by data from complex systems may include states that does not contribute to the dynamics of interest for the control application. Hence, it may be useful to reduce model order while preserving model characteristics that are important for the control application. Thus,

verification to ensure that the reduced model preserves important features is necessary, and for control design, it is useful to verify that the reduced closed-loop system is stable. Additionally, it is useful to verify that the reduced open-loop transfer function adequately matches the original models where the open-loop gain is close to 1 in the gain crossover region(Ljung [44]).

The pole-zero cancellation test is simply to observe at which order the cancellations appear, which indicate a too-large model structure. However, there is a risk of choosing a too high model order, even for a very large number of data. This is mainly due to inconsistency between the various tests.

Figure 5.8 depicts an input-output defined pole-zero map of a candidate MIMO data model obtained using MATLAB.

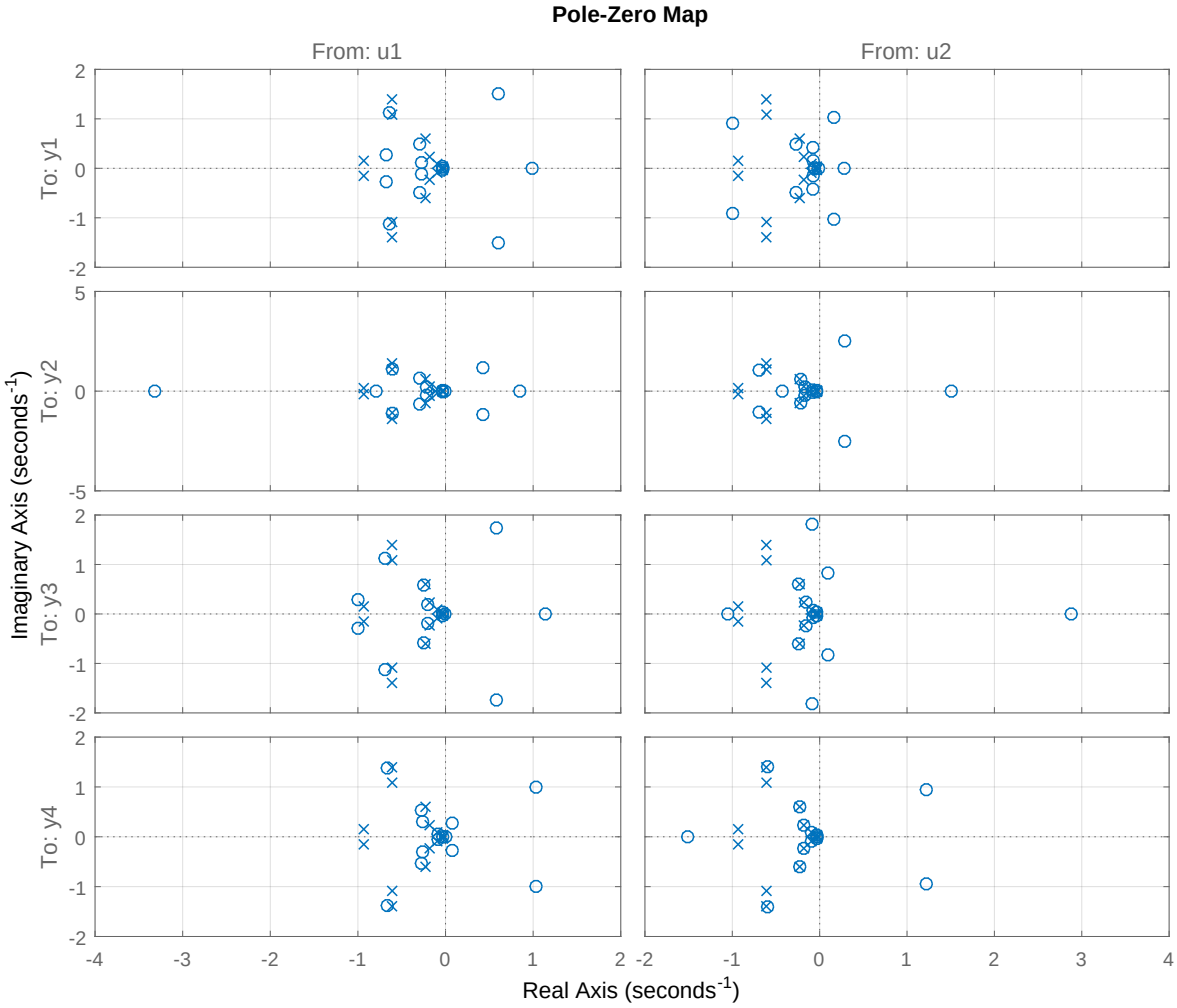
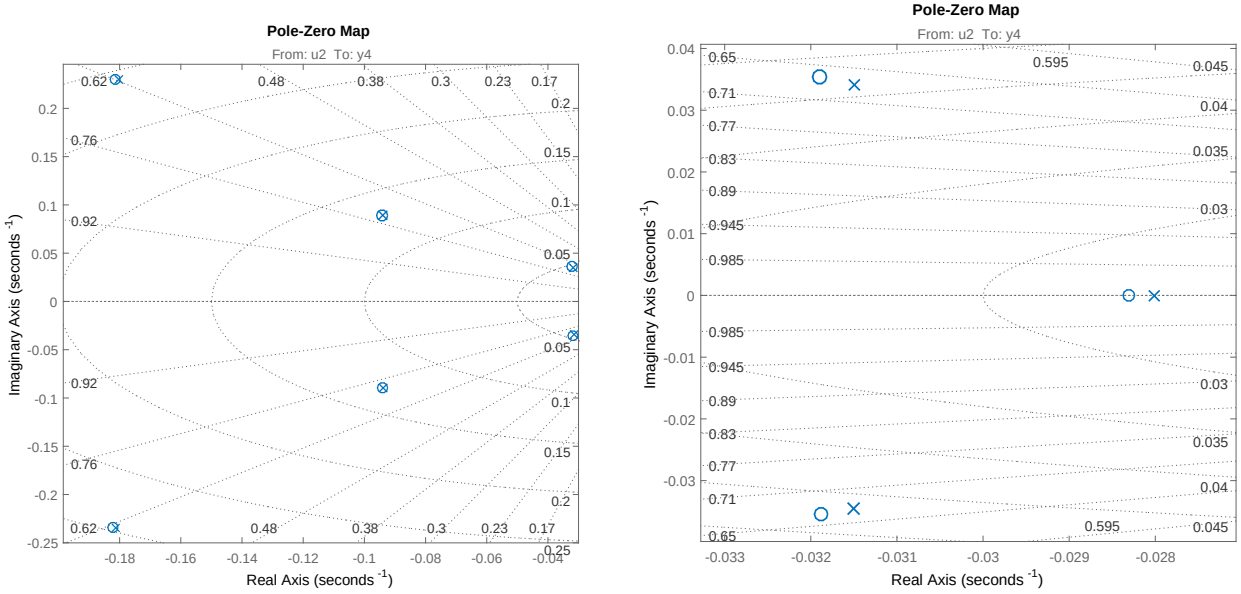


Figure 5.8 Pole zero map of candidate 15 order state-space model



(a) Enlarged pole-zero location for $U_2 \rightarrow Y_4$ channel. 3 potential pole-zero pairs which may be suitable for model reduction.

(b) Further enlarged pole-zero location for $U_2 \rightarrow Y_4$ channel. 3 pole-zero locations in close vicinity, but no pole-zero cancellation feasible.

Figure 5.9 3 candidate pairs of poles and zeros for model reduction. Plot b provides a closer look which render model reduction based on pole-zero location infeasible.

If the pole-zero plot indicates pole-zero cancellations in the dynamics, this suggests that lower order models can be used. In particular, for certain model structures it may be straightforward to indicate an increase in model order to obtain a good fit. However, if pole-zero cancellations are indicated, the extra poles are just introduced to describe the noise. The solution is then to investigate different model structures of similar order as the number of non-cancelled poles. However, when considering pole-zero cancellations precaution must be taken: this approach is purely theoretical and prone to numerical errors.

Observing above results reveal that model reduction is not possible. This was further confirmed using the `minreal` MATLAB function which returned a reduced order model of identical order as the original system. In addition, some practical issues limit the use of pole-zero cancellations. This is demonstrated in appendix C.8. Further information on a number of approaches to model reduction are covered in Skogestad and Postlethwaite [69].

5.4.1.3 Signal-to-noise ratio

In general, the signal-to-noise ratio plays a significant role in estimation. Even with persistent excitation, the stochastic effects in measurements may effect the model quality. The measured signal-to-noise ratio(SNR) is the quantification of relative contributions of deterministic excitation

and random variations, given by the ratio between variance of signal and variance of noise:

$$SNR = \frac{\sigma_y^2}{\sigma_N^2} \tag{5.4.4}$$

where the signal and noise variance is denoted by σ_y^2 and σ_N^2, respectively. If a significant portion of the variations in the measurement is due to noise, then the contribution of the input weakens and hence the ability to precisely estimate the model as well. In Ljung [44], the estimate of the open-loop transfer function covariance is defined as

$$\text{Cov } \hat{G}_0 = \frac{n}{N} \frac{\Phi_v(\omega)}{|S_0|^2 \Phi_u(\omega)} \tag{5.4.5}$$

where $\Phi_v(\omega)$ and $\Phi_u(\omega)$ is the power spectrums of the noise and excitation signals, respectively. $|S_0|^2$ is the squared absolute sensitivity function value. The estimated covariance is inversely proportional to the signal-to-noise ratio. That is, from above equation we have that the estimated covariance is proportional to the input power and inversely proportional to the noise variance. This feature directly indicates how well the open-loop transfer function may be estimated in the presence of noise.

An alternative viewpoint is that the SNR represents the ratio of effects due to known variables versus the uncertainties. Thus, a lower SNR, renders a more ambiguous estimate of the input-output model. An illustrating example is included in appendix C.3. However, in the case of utilizing simulator data, the noise is insignificant and we get a theoretically very large SNR. This allows for small perturbations, which is favorable when considering a linear model, when conducting the identification procedure. For further reading on SNR, see e.g., Ljung [44] and Tangirala [82].

5.4.2 Selection of model

A models' ability to reproduce input-output data in terms of simulations has been key in the previous sections for evaluation. In this section, the goal is to employ statistical tools to further investigate aspects which may assist in the search for the best possible model. In total, three MIMO models undergo various analyses in the subsequent sections.

5.4.2.1 Residual analyses

The prediction errors evaluated for the parameter estimates $\hat{\theta}$ are residuals, which describe the differences between the one-step-predicted output from the model and the measured output from the validation data set. Additionally, residuals represent the validation data not explained by the model. Thus, the residuals contain important information about the quality of the estimated model. In addition, the autocorrelation of the residuals indicates whether or not the error model is accurate.

The residual analysis consists of two tests: the whiteness test and the independence test. According to the whiteness test criteria, a good model places the residual autocorrelation function inside the confidence interval of the corresponding estimates, indicating that the residuals are uncorrelated. Furthermore, according to Ljung [44], the cross correlation between residuals and input of a suitable model does not go significantly outside its confidence region. Simply put, if the residuals don't move significantly outside the confidence region, the basic properties of the system has been captured by the estimated model.

For the independence test criteria, a good model has residuals uncorrelated with past inputs. Evidence of correlation indicates that the model does not describe how part of the output relates to the corresponding input. For example, a peak outside the confidence interval would indicate improper modelling for the part investigated.

The prediction error, technically termed as residual, serves as the key quantity of interest in assessing the goodness of the model.

$$\varepsilon(t) = y_t - \hat{y}_t \tag{5.4.6}$$

where y_t and $\hat{y}_t$ represents the simulator data and empirical model response, respectively. The part of the data which the developed model can not reproduce are the residuals. It is useful to study the correlation between residuals and past inputs to determine if the residuals depend on the input used. As defined on previous page, this is the independence test, and it is performed to ensure that the model behaves satisfactory for a range of different inputs. We have the correlation between residuals and past inputs:

$$\hat{R}^N_{\varepsilon u}(\tau) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t)u(t-\tau) \tag{5.4.7}$$

Returning a low covariance is, in general, favourable as it may indicate a consistent model behaviour independent of input. If there are traces of past inputs in the residuals, then there is a part of $y(t)$ that originates from the past input and that has not been properly picked up by the model.

A step further is to evaluate correlation between the residuals, which is defined on the previous page as the whiteness test:

$$\hat{R}^N_{\varepsilon}(\tau) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t)\varepsilon(t-\tau) \tag{5.4.8}$$

A relatively large residual correlation indicate that $\varepsilon(t)$ could have been predicted from past data. This means that $y(t)$ could have been better predicted, which is a sign of model deficiency. Further, the residual correlation $\hat{R}^N_{\varepsilon}(\tau)$ reveal if the residuals can be regarded as white. A large number of additional test which will not be covered in this work are listed in Ljung [44] and Tangirala [82].

5.4.2.2 Model integrity based on residuals

An accurate model should not produce residuals that offer further scope for predictions, while avoiding overfitting. Accordingly, Tangirala [82] lists some important test to determine the integrity of the model. Further, Ljung [44] states a useful rule of thumb: *A slowly varying cross correlation function outside the confidence region is an indication of too few poles, while sharper peaks indicate too few zeros or wrong delays.*

Verifying correlation between deterministic residuals and model inputs:

The autocorrelation function of ε and the cross-correlation between ε and the inputs are computed. The 99% confidence intervals for these values are additionally computed based on the assumption that ε is white and independent of the inputs. This assumption render computation of the autocorrelation function and the cross-correlation between ε and the inputs, feasible only for open-loop configurations. The correlation functions are computed and displayed up to lag 25. The 99% confidence region indicate that the insignificant correlations are within the shaded region around the X-axis.

Correlation between the residuals and lagged(time shifted) inputs are computed for the models. This is also known as the cross-correlation function, and cross correlation indicate if the investigated models leave behind any unidentified input effects. Three candidate models were investigated, and the cross correlations are depicted in figure 5.10. A considerable correlation between $\varepsilon(t)$ and $u(t)$ at positive lags imply that input effects on the process response is unexplained. Hence, model 3 does not perform well in this respect as seen from the figure. Based on the cross correlation plots in this subsection, the observed correlation past lag 0 indicate which model that describes the deterministic part best. In this case, model 3 performs clearly worst. It is, however, somewhat difficult to distinct models 1 and 2. Hence, to better distinguish models 1 and 2 appendix E.6 includes individual residual correlation plots for all three models.

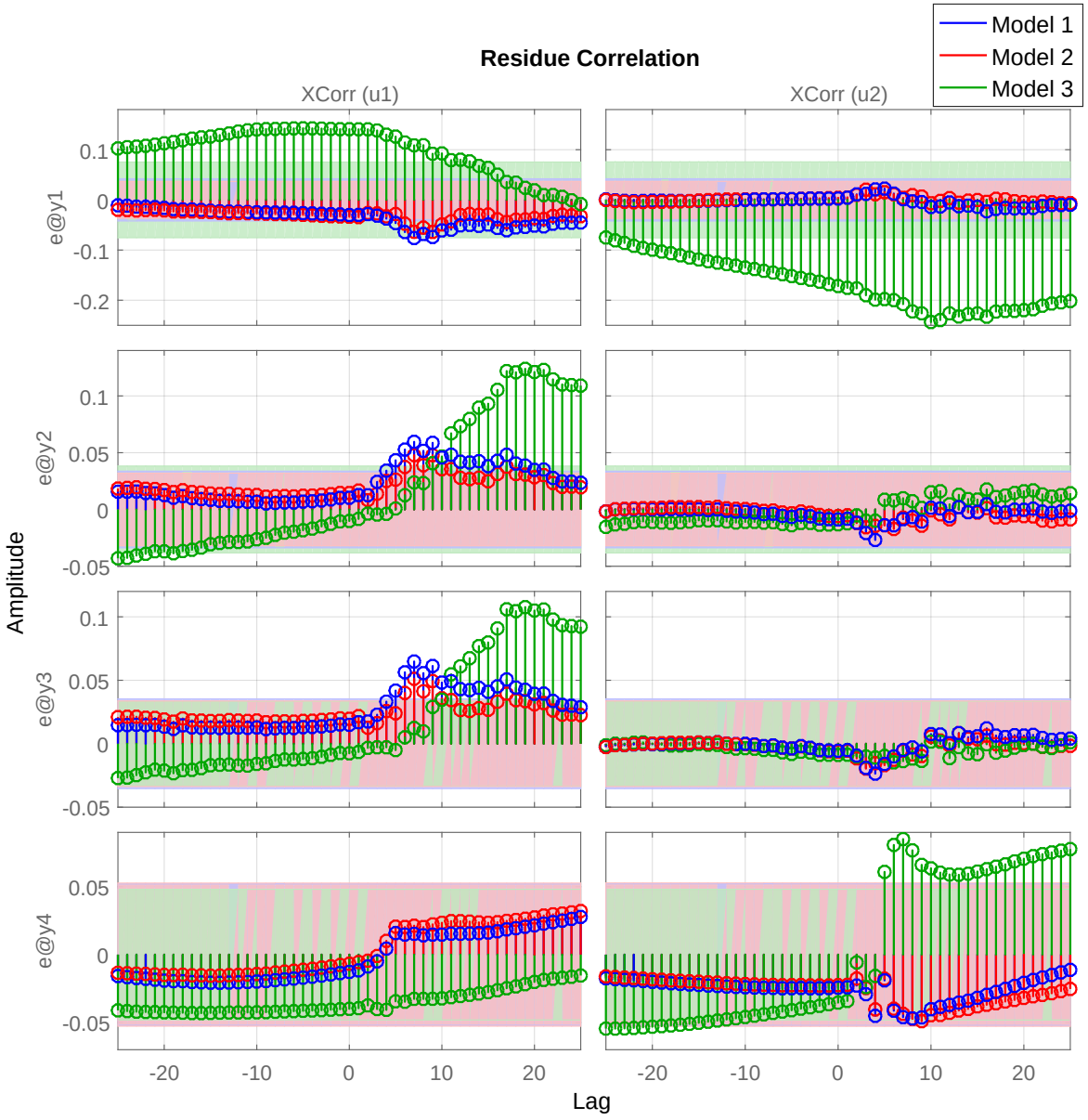


Figure 5.10 Correlation between residuals and lagged inputs for 3 candidate models.

Verify predictability in stochastic resdiduals:

This test assesses the models in terms of stochastic properties, and how well they demonstrate random effects, and for this purpose a plot of the auto-correlation function of the residuals is used. Essentially, it is the correlation between two arbitrary samples separated by a lag l. Predictability is indicated as non-zero correlation at a non-zero lag l. Missing considerable correlation at any non-zero lag indicates no space for predictability within the residuals. By definition, the auto-correlation function is unity at lag zero, i.e., any sample is best correlated with itself. This is observed in figure

5.11. Additionally, considering correlation at any non-zero lags yields no clear conclusion of best performing model when considering model 1 and model 2. Although, comparing all outputs may indicate that model 1 has least scope for predictability within the residuals, and thus is the preferred model. Furthermore, model 3 proceed being the worst model evaluated.

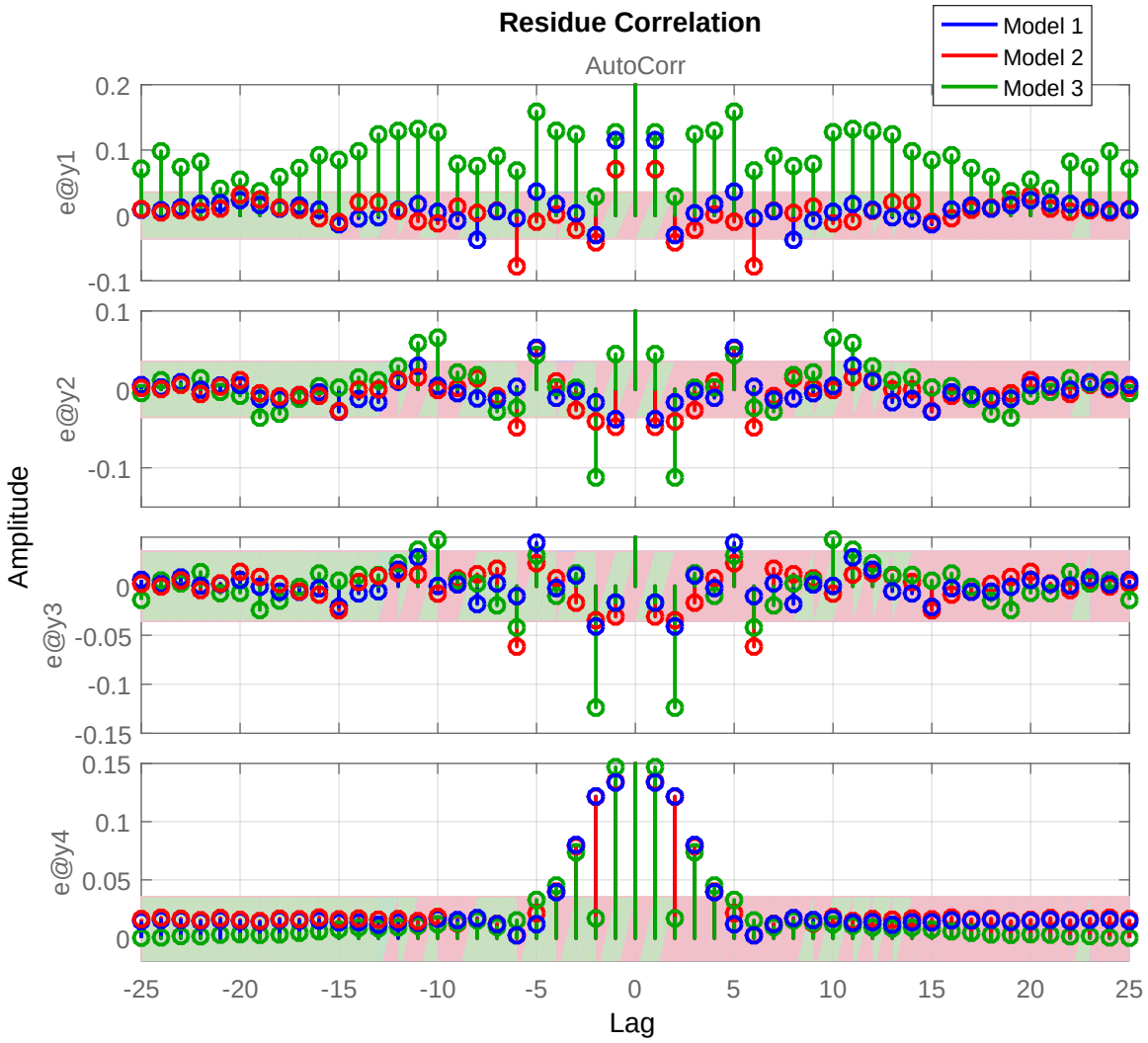


Figure 5.11 Auto-correlation of residuals for 3 candidate models.

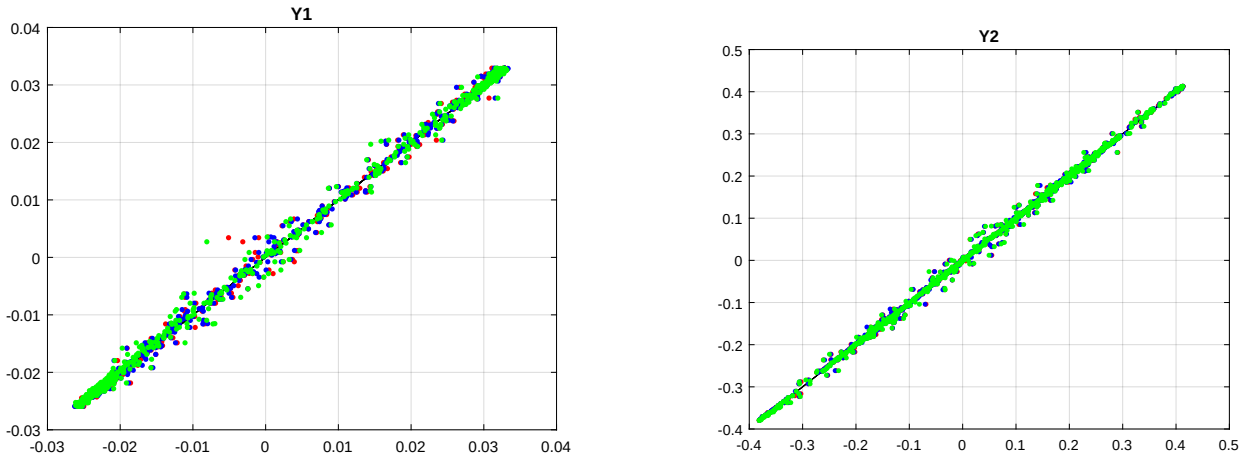
Further, individual plots autocorrelation of residuals for the three models considered here are additionally provided in appendix E.6.

Test for cross validation:

A good model yields adequate predictions on any independent data. In this case, linearity is though necessary due to the models being linear themselves. In this context, a model's ability to provide satisfactory predictions are investigated. The one-step ahead predictions of the output at a given

instant, k are based on knowledge of the observations until previous instant$(k-1)$. Figures 5.12 through 5.15 depicts scatter plots and zoomed plots for the three candidate models.

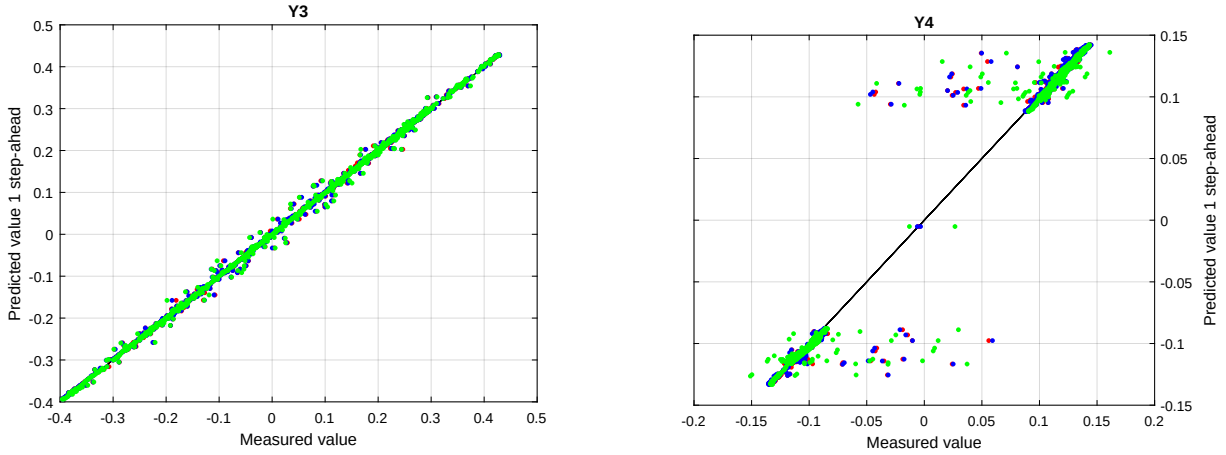
Further investigation is performed using infinite-step ahead predictions. These predictions are based on observations provided at infinite time in the past. This means that the infinite-step ahead predictions does not rely on information about the observations, but only on the input. For scatter plots and zoomed plots in a similar fashion as in figures 5.12 through 5.15 for the infinite step-ahead predictions, the reader is referred to appendix E.7.



(a) Y1 output scatter plot

(b) Y2 output scatter plot

Figure 5.12 Scatter plot comparing one step-ahead predictions against process data for outputs Y_1 and Y_2 on three models. Scatter plots comprises roughly 1050 samples. Red scatter represents model 1, blue scatter represents model 2 and green scatter represents model 3. Black lines indicate simulator data values.

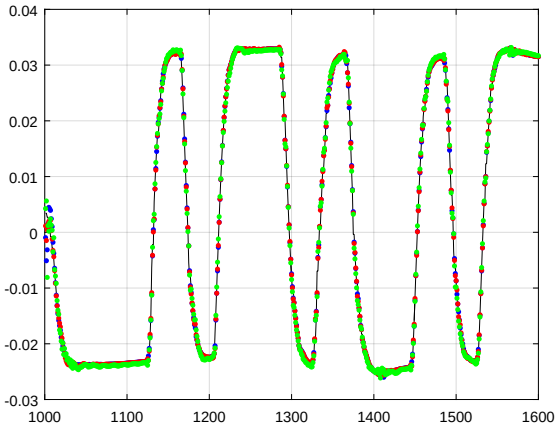


(a) Y3 output scatter plot

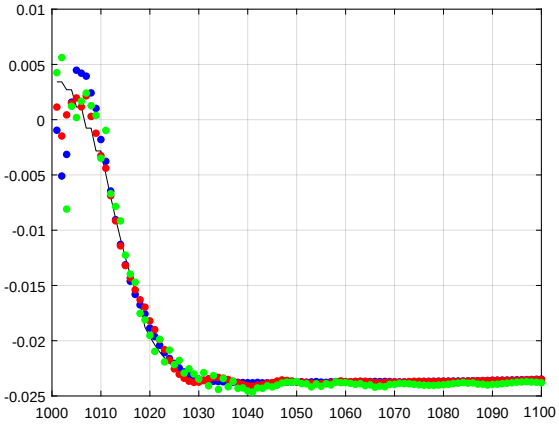
(b) Y4 output scatter plot

Figure 5.13 Scatter plot comparing one step-ahead predictions against process data for outputs Y_3 and Y_4 on three models. Scatter plots comprises roughly 1050 samples. Red scatter represents model 1, blue scatter represents model 2 and green scatter represents model 3. Black lines indicate simulator data values.

It is somewhat difficult to distinct the model performances purely based on the one step-ahead scatter plot. Hence, a more detail-focused plot of the one step-ahead prediction outputs are considered. These plots include two ranges of samples for each output: $1001 \rightarrow 1600$ and $1001 \rightarrow 1100$.

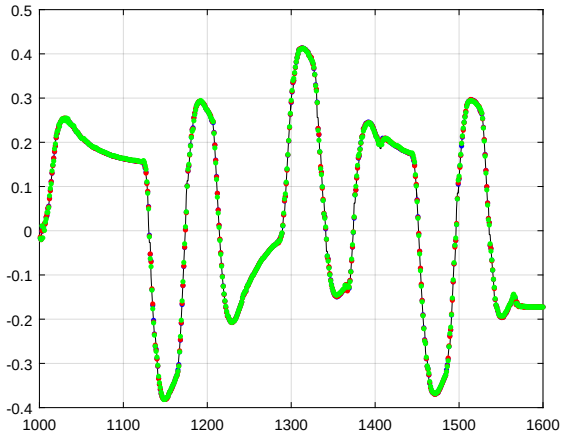


(a) Y1 output plot

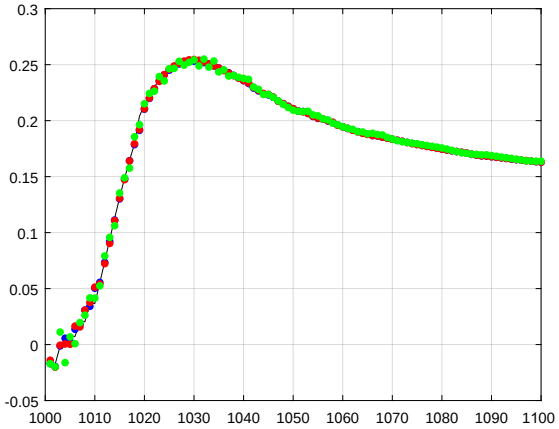


(b) Y1 zommed output plot

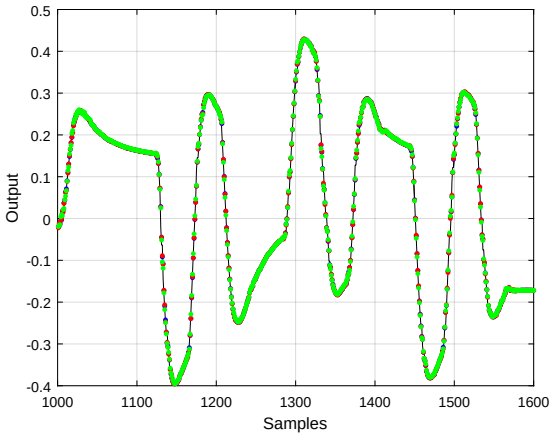
Figure 5.14 Comparing one step-ahead predictions against process data for three models. Samples 1001 to 1600 plotted to the left and 1501 to 1600 zoomed plots to the right. Red represents model 1, blue represents model 2 and green represents model 3. This figure depicts output Y_1. Remaining outputs are depicted on next page. Black lines indicate simulator data values.



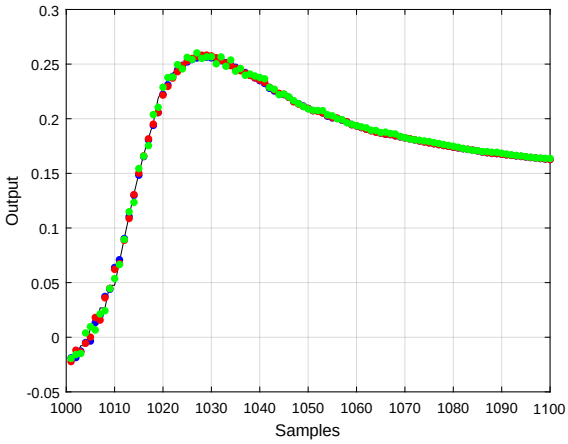
(a) Y2 output plot



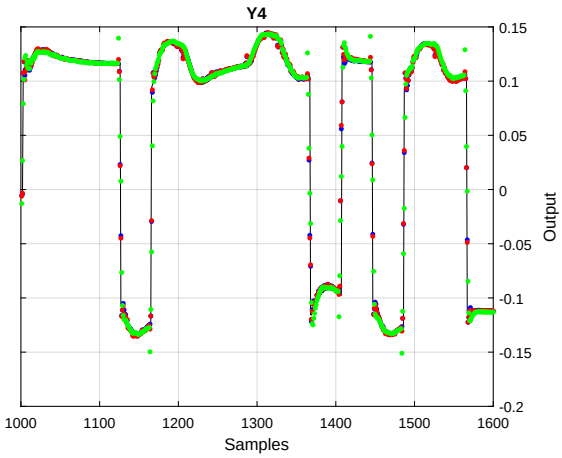
(b) Y2 zommed output plot



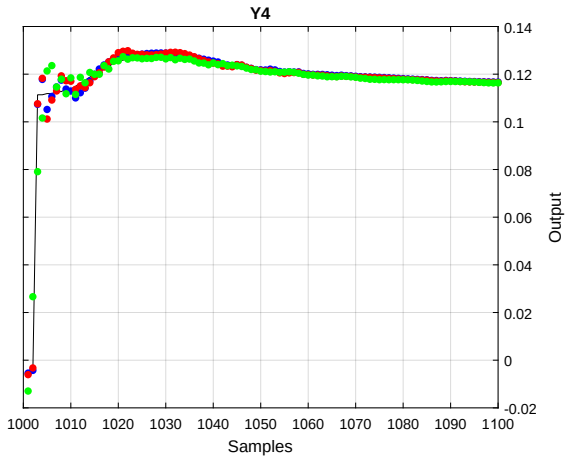
(c) Y3 output plot



(d) Y3 zommed output plot



(e) Y4 output plot



(f) Y4 zommed output plot

Figure 5.15 Comparing one step-ahead predictions against process data for three models. Samples 1001 to 1600 plotted to the left and 1501 to 1600 zoomed plots to the right. Red represents model 1, blue represents model 2 and green represents model 3. This figure depicts outputs Y_2, Y_3 and Y_4. Black lines indicate simulator data values.

As observed in the scatter and output plots, the one step-ahead predictions provide generally satisfactory results for all 3 models.

5.4.3 Final thoughts on methods of model validation

Based on the results derived in this section, some indications on model validity and suitability are provided. Still, as these are purely theoretical and statistical anchored approaches, the most important aspect is to consider application of the model. For example, the model may be required for controller design or state estimation. A typical validation is then to investigate if the problem that motivated the modelling can be solved using the obtained model. Simply put; if model-based control performs satisfactory within a given reference frame, the model is valid. However, to validate all candidate models may be impossible, costly or dangerous regarding the intended use. Thus, alternative measures are taken to verify model behaviour. An important feature for mechanistic models is the feasibility check for physical parameters(Tangirala [82]). Since this deals with empirical models, no focus on this aspect is overlooked.

Further, the model application often determines bounds on deviations, uncertainty and complexity. Although some deviations will always be expected independent of application, for certain systems bounds on deviations are tight. This could for instance be safety critical systems or systems based on state estimation with few measurements available. To analyse this, two key questions are useful:

- Similarity between model and validation data, and is it sufficient?
- Does model serve intended use properly?

Another question is how well the model describes the true system. To answer this question may be non-trivial. For multivariable systems, it is notably more difficult as for monovariable systems. For large multivariable systems, difficulty increases substantially as this may require a rigorous approach. Thus, this point is neglected in this work, although process knowledge to any extent will always be preferable.

Additionally for empirical models, input-output properties are of interest. To validate the consistency of empirical model input-output behaviour, and for linear models, Bode diagrams are useful. Figures 5.16 and 5.17 depicts the Bode plot including peaks and confidence intervals for models 1 and 3, respectively. However, a thorough and in-depth study of the Bode plots will not be carried out here. The complexity of the models simply render a full interpretation difficult.

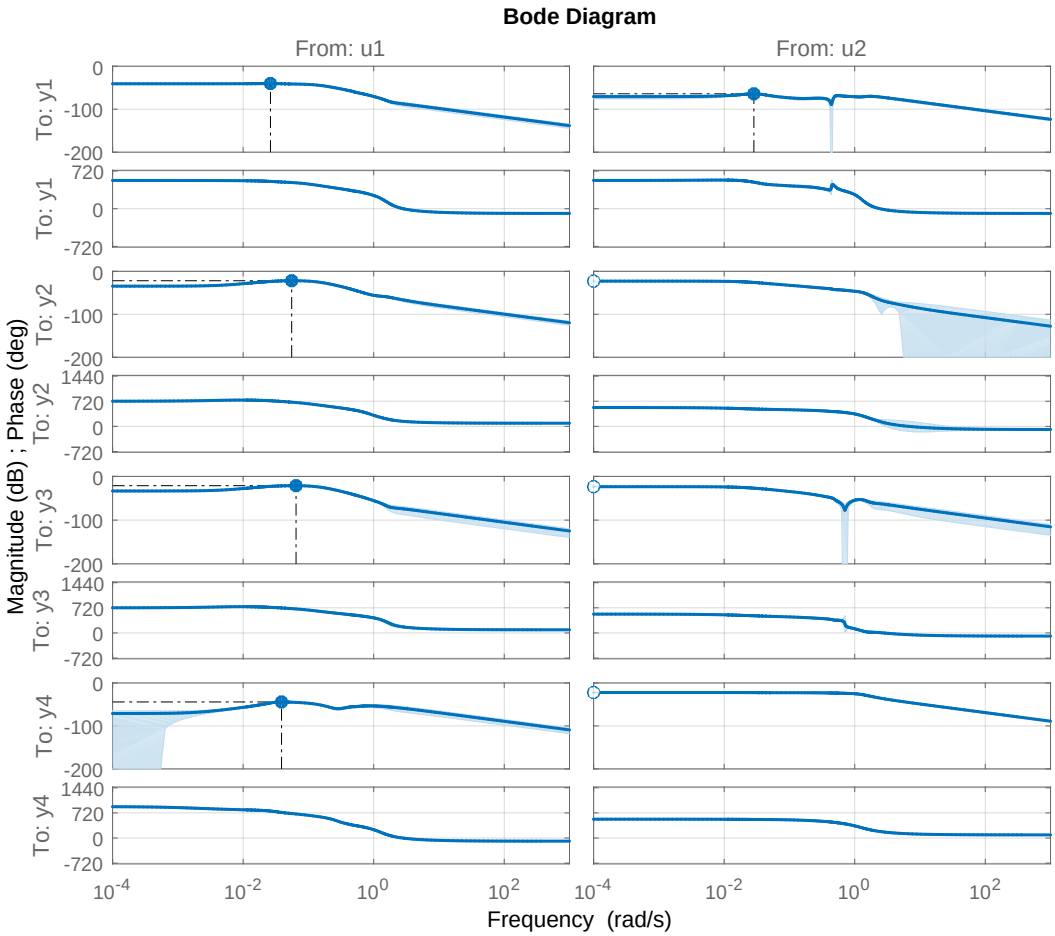


Figure 5.16 Bode plot of model 1 including peaks and confidence region.

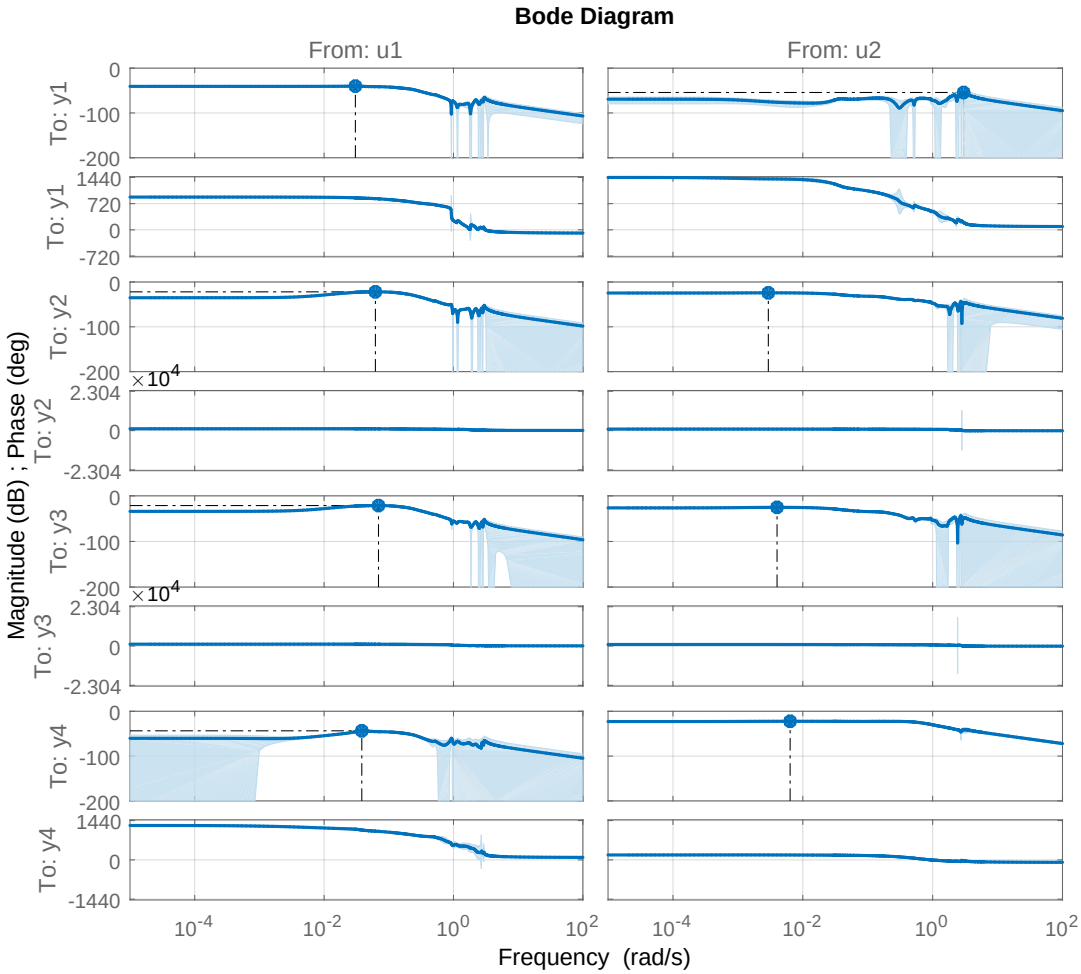


Figure 5.17 Bode plot of model 3 including peaks and confidence region.

Observed from the Bode plots, the models compared have some similar features while other features remain individual. Comparing gains for the U_1 input reveal somewhat similar frequency behaviour for both models, and certainly, for both models peaks are located in close vicinity. Considering gains for the U_2 reveal bigger differences; peaks on distinctive frequencies and less similar frequency behaviour.

Comparison of phase behaviour for the U_1 input is somewhat futile. Due to different scaling, phase comparison of outputs Y_2 and Y_3 become practically impossible. This applies for both inputs considering model 3. Additionally, comparing outputs Y_1 and Y_4 reveals significant differences, and output phases for model 3 are significantly bigger compared to model 1. Moreover, the confidence intervals are more prominent for model 3 which emphasize the uncertainties, and favours model 1 which has less prominent confidence intervals.

5.5 Assesing degree of interaction

It is earlier in this thesis emphasized that one challenge to achieve satisfactory control is the inherent multivariable features in the MFC process. To deal with multivariable systems essentially comprises two strategies which are outlined in Skogestad and Postlethwaite [69]: decentralized and multivariable control. Decentralized control is preferred for non-interacting processes and in cases where the active constraints remain constant. Multivariable control is thus preferred for interacting processes and in cases where the active constraints changes(Skogestad [68]).

To further emphasize the motivation behind the choice of utilizing an MPC, which indeed is a multivariable controller, this part assesses the interactions in the 2x4 system. The analyses are based on model 1 from the previous sections.

The relative gain array(RGA) measure for interactions was introduced in Bristol [8], and will be utilized in this work. However, in said reference the RGA was defined for square systems, i.e., for plants with a equal number of inputs and outputs. Thus for non-square plants, some additional research led to a general case of RGA, namely the NRGA, in which RGA is a special case of NRGA. This feature was demonstrated in Chang and Yu [11]. From a more recent perspective the RGA is given a thorough presentation in Skogestad and Postlethwaite [69]. Additional interaction measures are briefly mentioned in this thesis, but no attempt will be made to cover these topics entirely in this work. Interested readers may consult the thorough review of various methods for input-output selection and pairing given in van de Wal and de Jager [84].

The reader should note that the main motivation behind this section is to demonstrate interactions which emphasize the use of a multivariable controller. Input-output pairings and other features avaliable through RGA analyses are only superficialy considered. Thus, no conclusions apart from the degree of interaction will be derived here. Although, additional ideas which may be of further interest are introduced.

5.5.1 RGA of non-square plants

When considering control of a non-square system, there are several approaches available for control; non-square controller, conventional SISO control with decoupling, square down of system based on NRGA analysis or to utilize an advanced controller e.g., MPC. Considering a non-square system with more outputs than inputs($m > n$), it is obvious that to control all CVs to their setpoints perfectly is an impossibility. In general, with n inputs, for the ideal case n outputs are possible to keep under perfect control.

When considering the system as a transfer matrix($\boldsymbol{G}$), the NRGA is defined as

$$\boldsymbol{\Lambda^N} = \boldsymbol{G} \otimes (\boldsymbol{G}^{\dagger})^T \tag{5.5.1}$$

where $\otimes$ is the Hadamard or Schur product, and $\dagger$ is the Moore-Penrose pseudoinverse. For the case of frequency dependent NRGA, we have(Halvarsson [26]):

$$\boldsymbol{\Lambda^N} = \boldsymbol{G}(j\omega) \otimes (\boldsymbol{G}(\boldsymbol{j\omega})^{\dagger})^T \tag{5.5.2}$$

Within the RGA matrix each relative gain is defined as:

$$\lambda_{ij}^N = \frac{\left[\frac{\partial y_i}{\partial u_j}\right]_{ol}}{\left[\frac{\partial y_i}{\partial u_j}\right]_{cl}} = g_{ij} g_{ji}^{\dagger} \tag{5.5.3}$$

where subscripts $_{ol}$ and $_{cl}$ represent open-loop and closed-loop gains, respectively.

Combining above definitions, the steady-state NRGA matrix is derived for current system below(Chang and Yu [11]):

$$\boldsymbol{\Lambda^N} = \begin{bmatrix} \lambda_{11}^N & \lambda_{12} \dots & \lambda_{1n}^N \\ \vdots & & \vdots \\ \lambda_{m1}^N & \dots & \lambda_{mn}^N \end{bmatrix} = \begin{bmatrix} 0.2036 & 0.0012 \\ 0.2778 & 0.1456 \\ 0.509 & 0.0038 \\ 0.0096 & 0.8494 \end{bmatrix} \tag{5.5.4}$$

where column sum vector is

$$\begin{aligned} \boldsymbol{CS} &= \left[\sum_{n}^{i=1} \lambda_{i1}^N, \sum_{n}^{i=1} \lambda_{i2}^N, \dots, \sum_{n}^{i=1} \lambda_{in}^N \right]^T \\ &= [1.0 \;\; 1.0]^T \end{aligned} \tag{5.5.5}$$

and the row sum vector is

$$\begin{aligned} \boldsymbol{RS} &= \left[\sum_{n}^{j=1} \lambda_{1j}^N, \sum_{n}^{j=1} \lambda_{2j}^N, \dots, \sum_{n}^{j=1} \lambda_{mj}^N \right]^T \\ &= [0.2048 \;\; 0.4234 \;\; 0.5129 \;\; 0.8590]^T \end{aligned} \tag{5.5.6}$$

From Chang and Yu [11] some properties of the NRGA matrix are

- Sum of each column is is always 1. This property is consistent with RGA properties.
- Sum of each row, $0 \leq rs(i) \leq 1, \; \forall \, i$. A row sum less than unity indicate imperfect control.
- NRGA is invariant under input scaling, but variant under output scaling. Input scaling pertain to post-multiplying with a diagonal scaling matrix, and output scaling represents

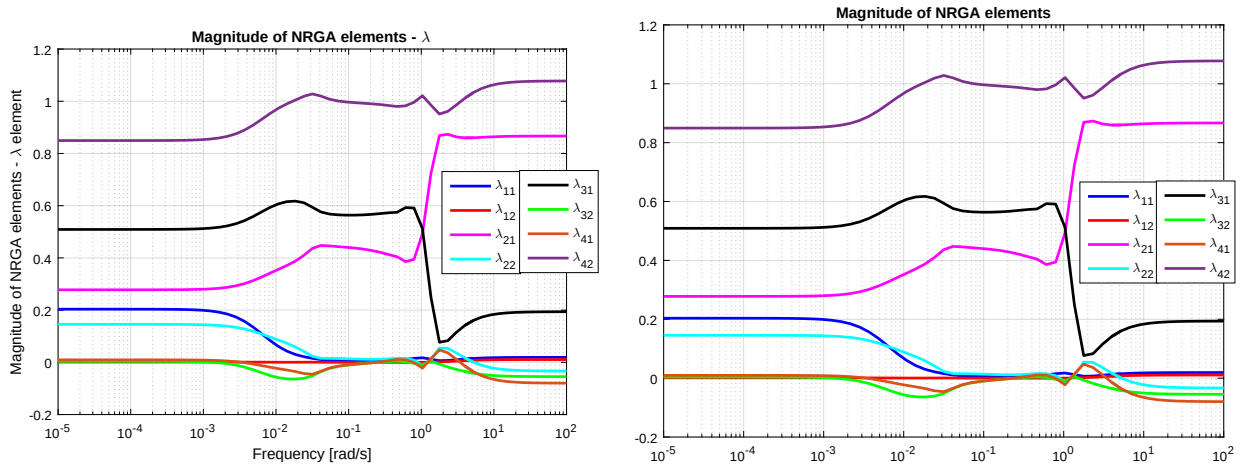
pre-multiplying with a diagonal scaling matrix. This property hold only for cases $m > n$. For the opposite case, the opposite is true.

Based on steady-state values, the NRGA analysis indicate that Y_1 is difficult to control using a conventional feedback application, e.g., PI control. Considering that two degrees of freedom are available for control, outputs Y_1 and Y_2 should be removed based on the NRGA steady-state analysis. However, this approach is not recommended for operation. Knowing that Y_1 is perhaps the most vital CV to control, a solution based purely on NRGA indications may not be feasible. This important consideration is not included in the NRGA analysis, and emphasizes that process knowledge is superior to any tools such as NRGA analyses which provide a pure numerical approach.

In addition, the dynamical NRGA(equation 5.5.2) is computed for frequencies in the span

$$\omega = [1 \times 10^{-5} \text{ rad/s}, \ldots, 1 \times 10^{2} \text{ rad/s}] \tag{5.5.7}$$

This is depicted in figures 5.18a and 5.18b. Unlike the steady-state NRGA, the dynamical NRGA indicate altering interactions which clearly are frequency dependent. In general, the main goal is to derive input-output pairings close to identity in the bandwidth region, although this feature is not considered in this work. However, by observing the dynamical NRGA figures, we observe an altering interaction around $\omega = 2$ rad/s. Additional information on the dynamical (N)RGA is found in Skogestad and Postlethwaite [69].



(a) Dynamical NRGA for 2x4 system. Vectorial absolute λ values.

(b) Dynamical NRGA for 2x4 system. Real part λ values.

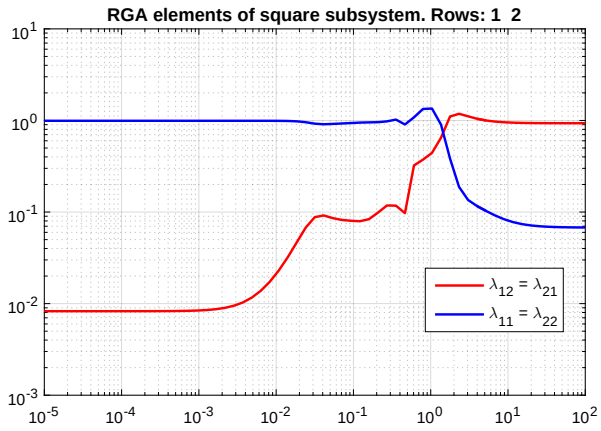
Figure 5.18 Dynamical NRGA for 2x4 system.

5.5.2 RGA for a square plant subsection

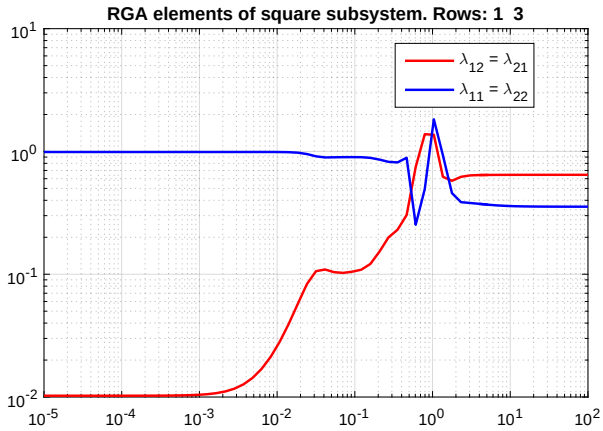
As earlier mentioned, an alternative approach is to concentrate on a square part of the system, i.e., select an equal number of outputs corresponding to available inputs. This is not uncommon

in chemical plants(Chang and Yu [11]). To choose an appropriate square part of the plant, the NRGA analysis was performed for the square case, thus making it an RGA analysis for all possible configurations of the 2x4 system.

In addition to the NRGA approach for both non-square and square plants, the performance relative gain array(PRGA) has been investigated for the square plant cases. As earlier mentioned, the RGA indicate two-way interactions which may lead to destabilization. In the case of a triangular plant, the RGA does not indicate the presence of couplings through off-diagonal elements. Thus, one-way interactions are not indicated by the RGA, and what may seem straightforward using RGA may be difficult or even erroneous based on the PRGA. The diagonal elements of the PRGA is equal to the RGA diagonal elements. The off-diagonal PRGA elements are different from off-diagonal RGA elements, and additionally output scaling dependent. Large PRGA elements indicate slow process interactions and may imply that satisfactory control is difficult to obtain subjected to individual loop control. However, small PRGA elements indicate that interactions actually may improve performance. This is observed from equation 10.94 in Skogestad and Postlethwaite [69]. Figures depicting the dynamical PRGA are included in appendix E.8, and the findings are unambiguously indicating that satisfactory decentralized control is difficult to obtain. This, of course, favours the use of MPC. Additionally, the dynamical singular values computed are found in appendix E.8.

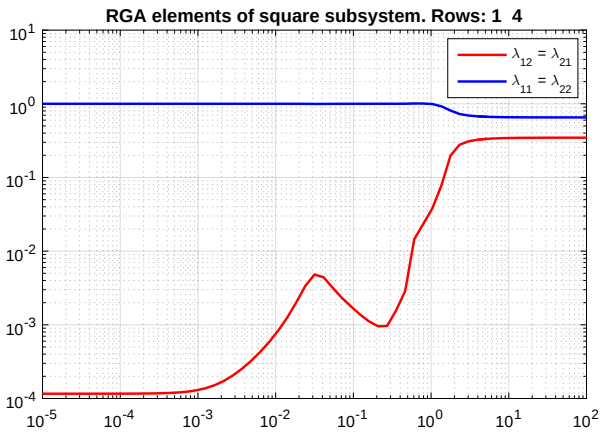


(a) RGA rows 1 & 2 from NRGA matrix

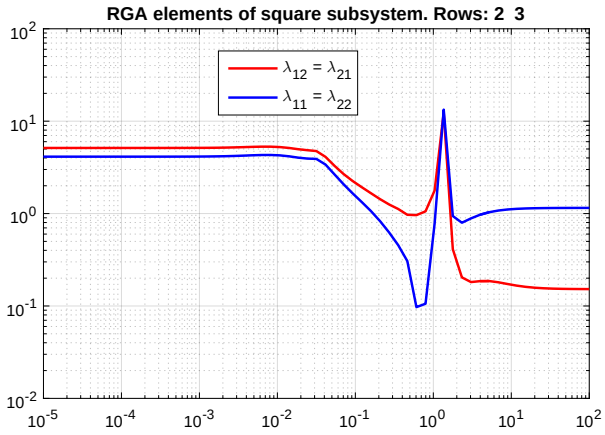


(b) RGA rows 1 & 3 from NRGA matrix

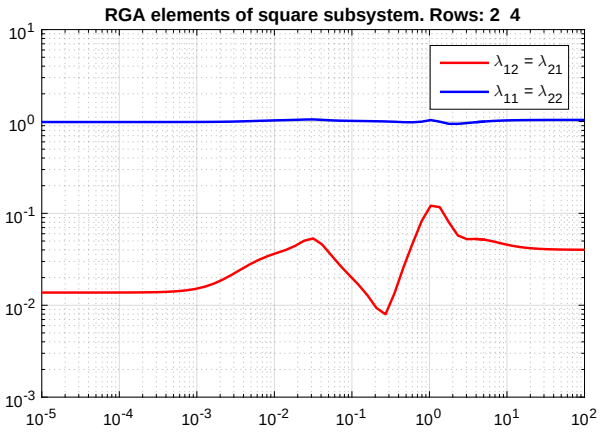
Figure 5.19 Comparing possible input-outputs pairings based on down-squaring of the non-square plant.



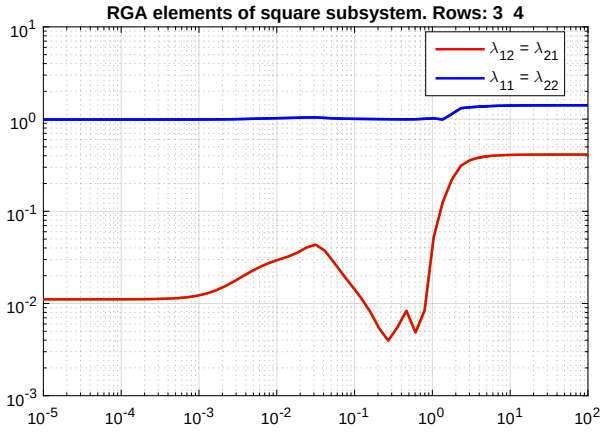
(a) RGA rows 1 & 4 from NRGA matrix



(b) RGA rows 2 & 3 from NRGA matrix



(c) RGA rows 2 & 4 from NRGA matrix



(d) RGA rows 4 & 3 from NRGA matrix

Figure 5.20 Comparing possible input-outputs pairings based on down-squaring of the non-square plant.

5.5.3 Additional measures for interactions

SVD may be utilized to handle non-square systems. However, this method relies on input scaling, and thus, uncontrolled variables may change dependent on scaling. This is demonstrated in Chang and Yu [11]. Gramian based interaction measures, such as the Hankel Interaction Index Array(HIIA) are considered in the works of Halvarsson [26] and Wittenmark and Salgado [89]. This interaction measure is however, scaling dependent as well. Wittenmark and Salgado [89] demonstrates that the HIIA outperforms RGA when dealing with systems that have interactions with non-monotonic frequency behaviour. However, the HIIA may not give reliable pairing suggestions when the state space realization has a nonzero direct matrix D. The reason is that the D matrix is not used when computing the Gramians(Halvarsson [26]). Additionally, another measure for interactions is Gershgorin bands. Although, no analysis by utilizing said measure is carried out here, an

introduction is found in Skogestad and Postlethwaite [69]. Due to the fact that deriving the condition number relies on SVD analysis, and thus is scaling dependent, no effort has been made to make use of the condition number. Scaling is essentially based on heuristics which makes it a potential pitfall that, if performed incorrect, may provide erroneous results(Halvarsson [26]). An example which demonstrate that one cannot generally conclude that a plant with a large condition number is sensitive to uncertainty is provided in Skogestad and Postlethwaite [69].

However, an introduction to condition numbers are presented in appendix C.2. Based on the definition, it is possible to estimate condition numbers for the investigated system by observing the dynamical singular values. For the similar square systems depicted in figures 5.20 and 5.19, the dynamical singular values were computed for the span in 5.5.7. By observing the largest and smallest singular values, one is able to make an educated guess on the condition number size. Figures E.20 and E.21 in appendix E depict the dynamical singular values for the chosen model.

Additional measures which may be of interest for control analyses are:

- Ability to control outputs independently. The controllability matrix:

$$C_o = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \tag{5.5.8}$$

  has full rank if the system is controllable. For example, for the 16 order MIMO model $rank(C_o) = 16$. Which indicate a controllable system.

- Further investigation of poles and zeros, and related directions.
- Compute a disturbance model $G_d(j\omega)$ to obtain the lower bandwidth(ω_d) for control.

Pole and zero directions and obtaining a disturbance model have not been investigated further in this work. Consult e.g. Skogestad and Postlethwaite [69] for additional information.

Chapter 6

Comparison of modelling results

This section provides some graphical results on comparison of the various candidate models. At first, two MATLAB derived MIMO models are compared with the identified SISO models from SEPTIC. Additionally, SISO models derived in MATLAB and SEPTIC are compared. In closing, 3 MATLAB developed models and the SEPTIC developed models are compared against simulator data from various step perturbations.

6.1 Obtaining models in SEPTIC

Strand and Sagli [80] authors emphasize the use of linear step response models in a majority of the existing MPC applications in 2003. The software has been developed since initialization and new measures for obtaining models were included. This includes a scheme for automatic closed-loop identification based on the work of Zhu and Butoyi [91]. This resulted in a commercial software package(Zhu [90]) which was included in the SEPTIC application. Although the software in Zhu [90] was earlier employed, as far as the author knows, step response models has been, and still is the preferred practice, and identification is no longer based on external software.

In addition to step response models, identification of ARX and FIR models are included. This is a necessity to identify models where no means of manipulation allow for step responses. This could for instance be variables manipulated by a controller, where the only option for excitation is to manipulate either the setpoint or controller output directly. For example, the feed gas temperature is a potential disturbance investigated in this work and may render an inaccurate model. Manipulating this temperature in steps does not resemble true variations. The more realistic case could for instance be a periodic signal which simulate the feed gas temperature as a function of plant surrounding air temperatures. Including this aspect in the identification exclude the use of step response models. Thus, FIR and ARX models are more appropriate.

By a mathematical point of view, simple step responses does not perturb the system sufficient to

capture essential dynamics, and for nonlinear and multivariable systems this is to a greater extent emphasized. In general, a system of a high order requires a higher degree of excitation signals to highlight all states. From a practical point of view the discussion is many-facetted, and the arguments found in Strand and Sagli [80] more or less defines today's practice within the company. These observations indicate a gap between industry and academia in the field of system identification, which in this work, the author will make no attempt close. Having said that, a comparison of models identified from different mathematical and practical principles have been conducted.

One idea behind the step response approach is to keep things simple. By assuming a unit step input, i at time instant t, we have the step response for output,j

$$y_j(t) = \sum_{k=0}^{t} g_{ji}(t-k) \tag{6.1.1}$$

By using the simplification

$$\sum_{k=0}^{t} g_{ji}(t-k) = \sum_{k=0}^{N} g_{ji}(k) \tag{6.1.2}$$

we get

$$y_{ji} = \sum_{k=0}^{t} g_{ji}(k) \tag{6.1.3}$$

By sampling the response over a sufficiently large horizon until steady-state($t \rightarrow N$), we have

$$g_{ji}(N) \approx g_{ji}(N+t) \tag{6.1.4}$$

Based on the control structure of 2 manipulated variables and 4 controlled variables, the resulting step response matrix for the total system become

$$G(t) = \begin{bmatrix} g_{11}(t) & g_{12}(t) \\ g_{21}(t) & g_{22}(t) \\ g_{31}(t) & g_{32}(t) \\ g_{41}(t) & g_{42}(t) \end{bmatrix} \tag{6.1.5}$$

A response to an arbitrary input for an arbitrary output is thus given by

$$y(t) = \sum_{k=0}^{t} G(t-k)u(k) \tag{6.1.6}$$

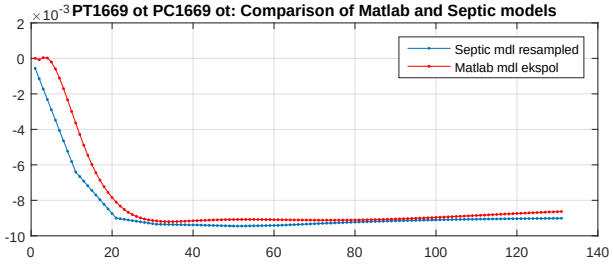
Since this approach coincides with the approach described in section 4.2.2, the possibility of MIMO or MISO identification is not possible. Thus, SEPTIC allows only for pure SISO identification.

6.1.1 Comparing SEPTIC and MATLAB obtained models

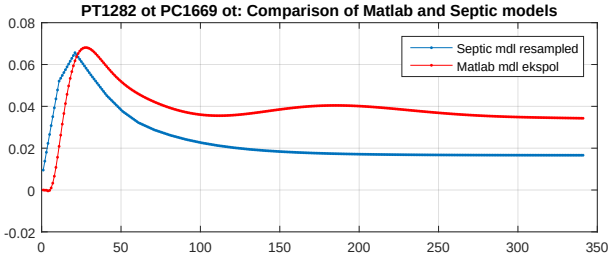
An additional aspect consecutive of obtaining the models was to investigate the differences between the models. This is not a straightforward matter as SEPTIC and MATLAB have different schemes to store the models. Independent of model obtained; ARX, FIR or step response model, SEPTIC stores the model information in a text file format. This information is universal and is translated to a unit step response. Thus, all models develop in SEPTIC constitute a step response, and as mentioned, all models are SISO. MATLAB stores models inherently in a defined format. The model dimension is user-defined, and there exist a range of models available. To coincide with the SEPTIC models the MATLAB model was unit step perturbed for each channel, and the basis for comparison were the individual step responses.

As for user-defined parameters, SEPTIC models are influenced by sample time. This parameter determines the number of samples to use for model description. The model length, i.e., time to reach steady-state following the step response is automatically computed. Based on discussions on the SEPTIC setup with a control engineering researcher at Statoil, the sample interval was defined to 10 seconds. Thus, a model which contains 10 samples takes 100 seconds to reach steady-state consecutive a unit step perturbation. As observed in figures 6.1 and 6.2, the time constants for the MFC variables investigated are not slow nor fast, but somewhat in-between. This renders a relatively fast application to handle the fast responses well. To ensure feasible solutions within the given sample instant, parameters regarding computational efficiency were considered. This is further described in section 7.3.

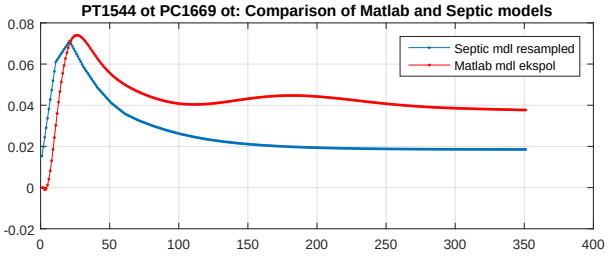
To realize this model comparison, interpolation of the SEPTIC models and in some cases extrapolation of the MATLAB models were necessary. All SEPTIC models were linearly interpolated to match the sample time of 1 seconds for the MATLAB model. For each output channel compared, the slowest model set the range of time for comparison. Thus, if a MATLAB model reach steady-state prior to the SEPTIC model, the MATLAB model response is extrapolated. For the other way around, the length of the SEPTIC model sets the boundary length of comparison and the MATLAB response is shortened accordingly. A comparison of unit step responses for the SEPTIC models was performed against a MATLAB MIMO model, and a MATLAB SISO model are found in figures 6.1 and 6.2.



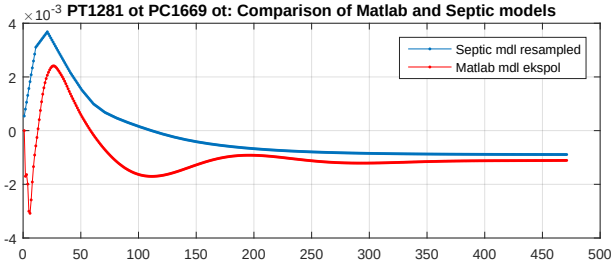
(a) PIC1669 → Y1



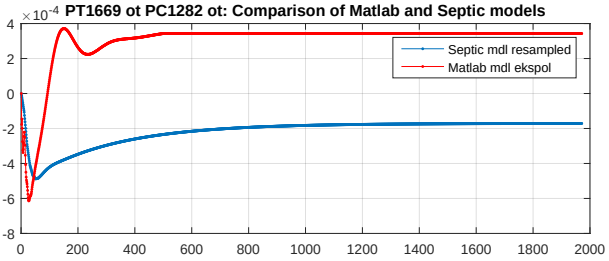
(b) PIC1669 → Y2



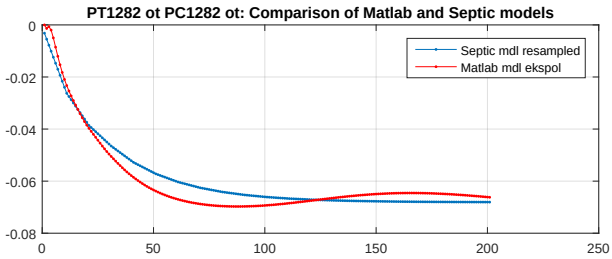
(c) PIC1669 → Y3



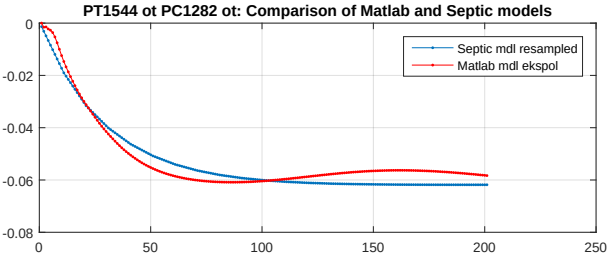
(d) PIC1669 → Y4



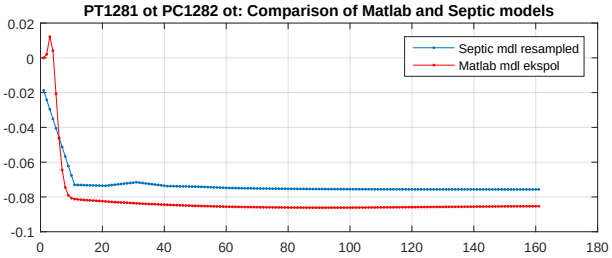
(e) PIC1282 → Y1



(f) PIC1282 → Y2

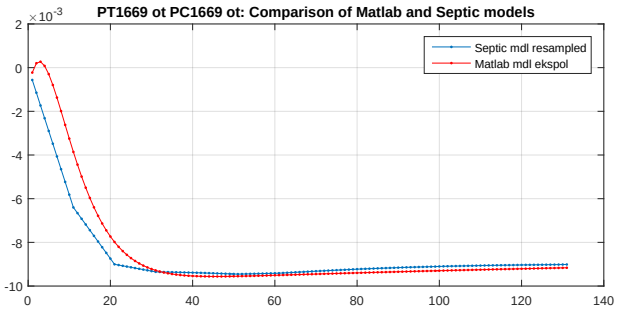


(g) PIC1282 → Y3

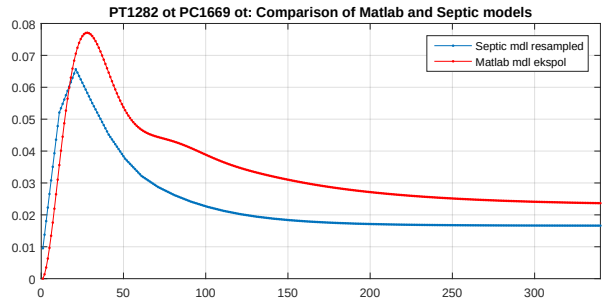


(h) PIC1282 → Y4

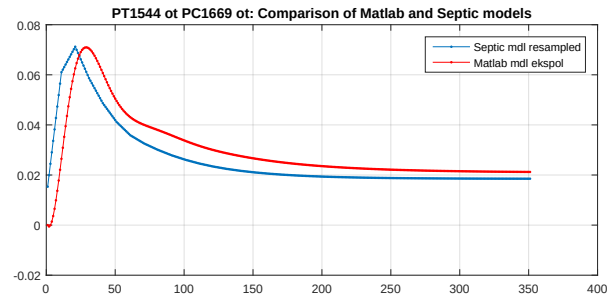
Figure 6.1 Step response comparison of SEPTIC and MATLAB MIMO data model



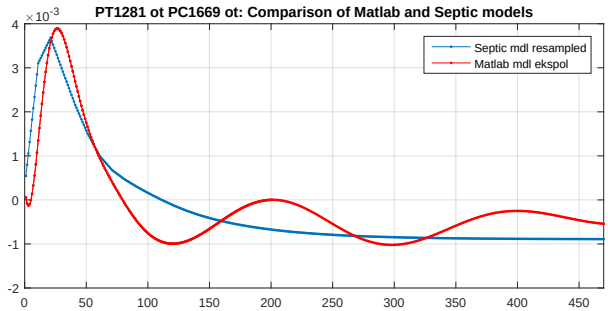
(a) PIC1669 → Y1



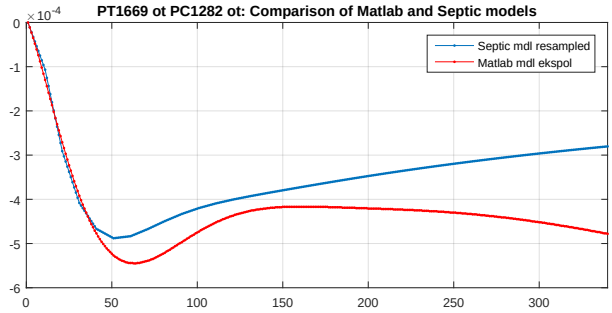
(b) PIC1669 → Y2



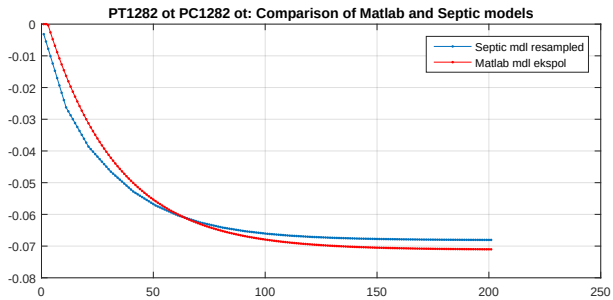
(c) PIC1669 → Y3



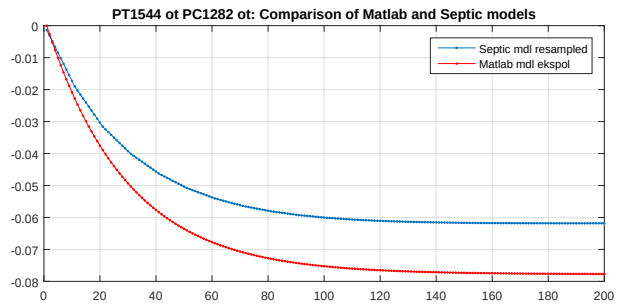
(d) PIC1669 → Y4



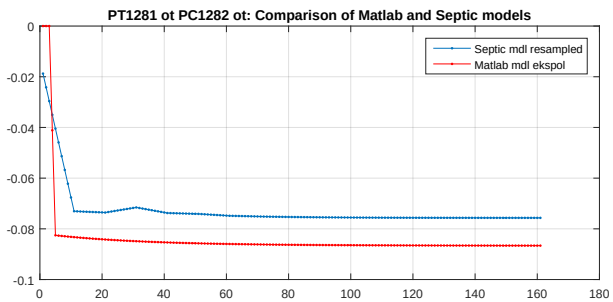
(e) PIC1282 → Y1



(f) PIC1282 → Y2



(g) PIC1282 → Y3



(h) PIC1282 → Y4

Figure 6.2 Step response comparison of SEPTIC and MATLAB SISO data model

From the figures 6.1 and 6.2, the shortest open-loop response time is ≈ 10 seconds, and the longest open-loop response time is ≈ 350 seconds. Additionally, there are differences both between the

SEPTIC model and MATLAB models, but also between the individual MATLAB models. This applies for both steady-state and dynamical differences to varying degrees, but graphically most distinctive is the PIC1282 → Y1 channel. Though, as the numerical differences are very small, i.e., $5 \cdot 10^{-4}$ for the biggest steady-state error, this may indicate numerical noise or properties that are of no model interest. Thus, no additional attention is devoted to decrease this particular error. Both the SISO data developed model and the MIMO data model developed in MATLAB demonstrated satisfactory behaviour as observed in the comparison figures.

It is emphasized that the initial dynamics are of most interest. This part of the response represents the faster dynamics, and as demonstrated in Jacobsen et al. [37] and Jacobsen and Skogestad [38], capturing the faster dynamics is essential in modelling. Additionally, the MPC and its iterative scheme render responses predicted beyond the next iteration somewhat insignificant.

Even though the modelling in SEPTIC is less laborious compared to the MATLAB approach, it does not exclude modelling on faulty input data. Thus, an example of faulty SEPTIC modelling is included in appendix E.4.

6.2 Comparison of various identified models and simulator data

As a final check for the candidate models, a comparison was conducted simultaneously while running the simulator and perturbing by steps manually. The result is observed in figures 6.4 through 6.12. Perturbations in the form of steps in various amplitudes were utilized, and the largest steps were of amplitude 10. Both directions were perturbed. In the comparison, four models are compared to simulator data; a MIMO model derived from a combined excitation data set, a MIMO model derived from two individual channel excitation data sets, a MIMO model derived from SISO identification and a SEPTIC derived step response model.

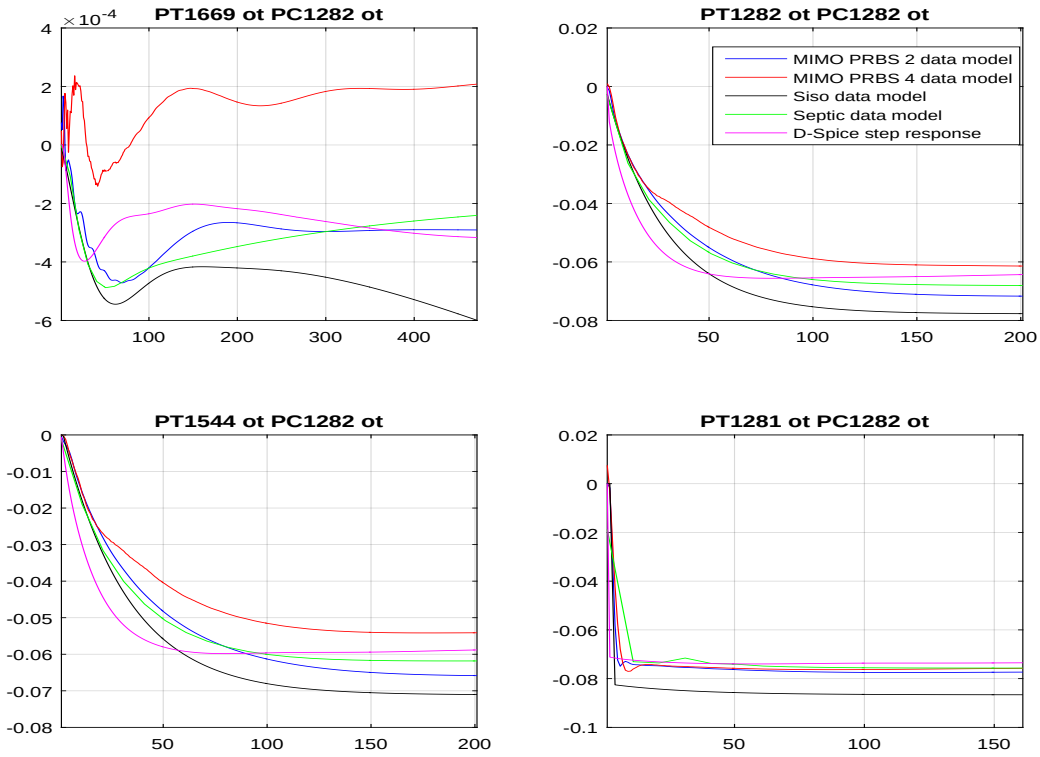


Figure 6.3 PIC1282 unit step up - simulator and model CV responses

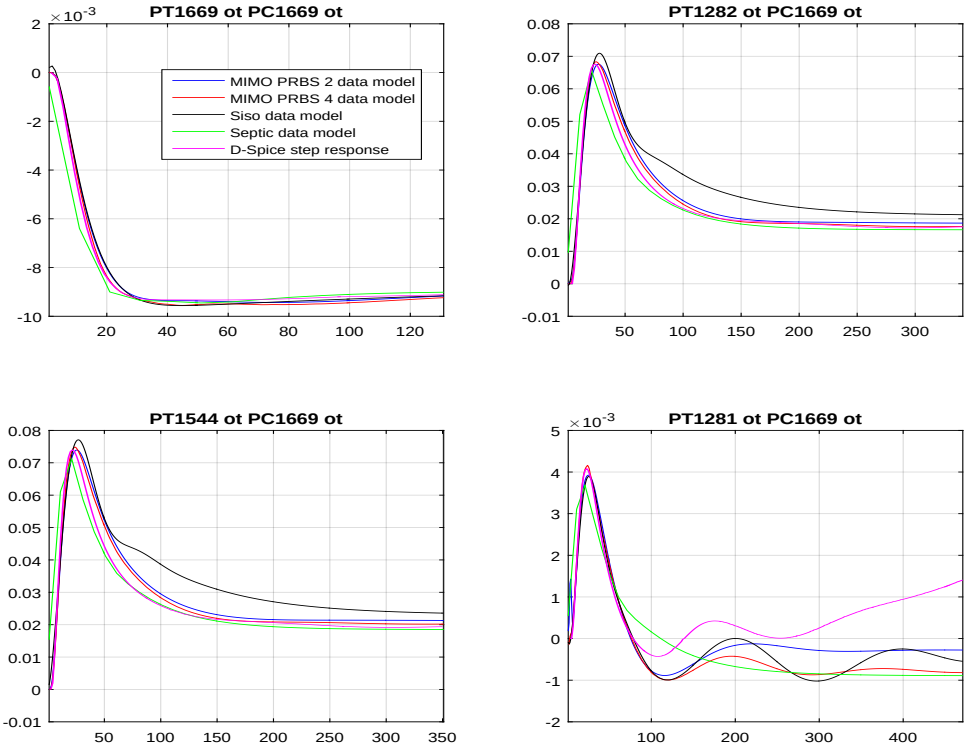


Figure 6.4 PIC1669 unit step up responses - simulator and model CV responses

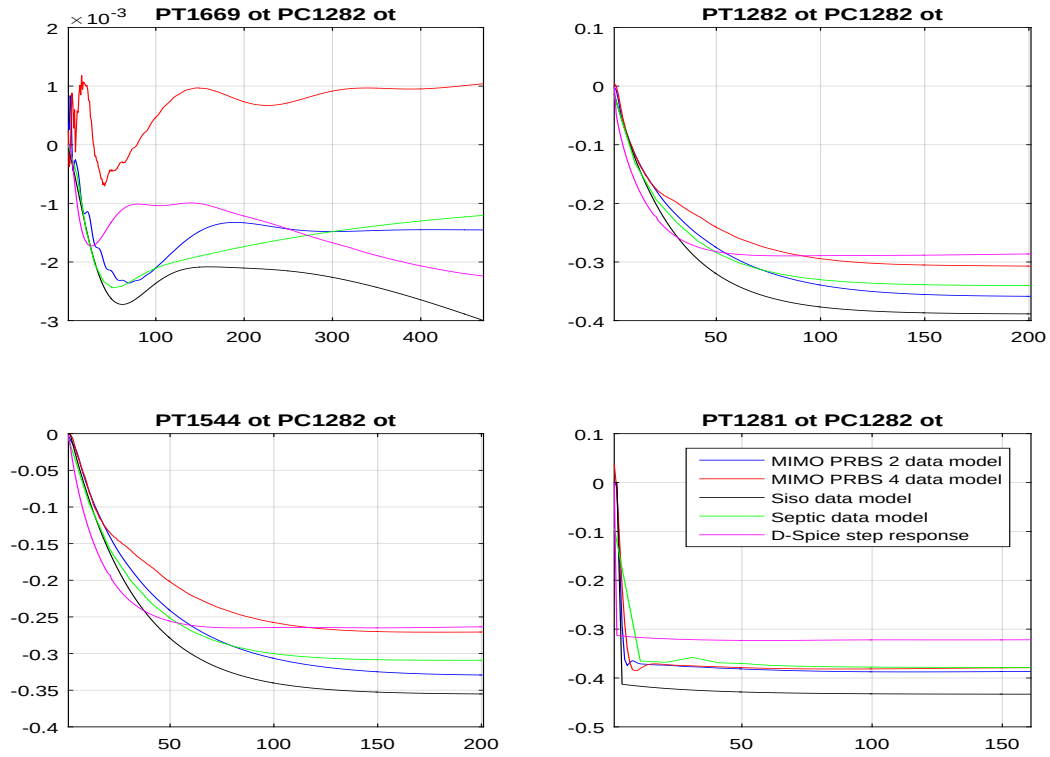


Figure 6.5 PIC1282 step of 5 up - simulator and model CV responses

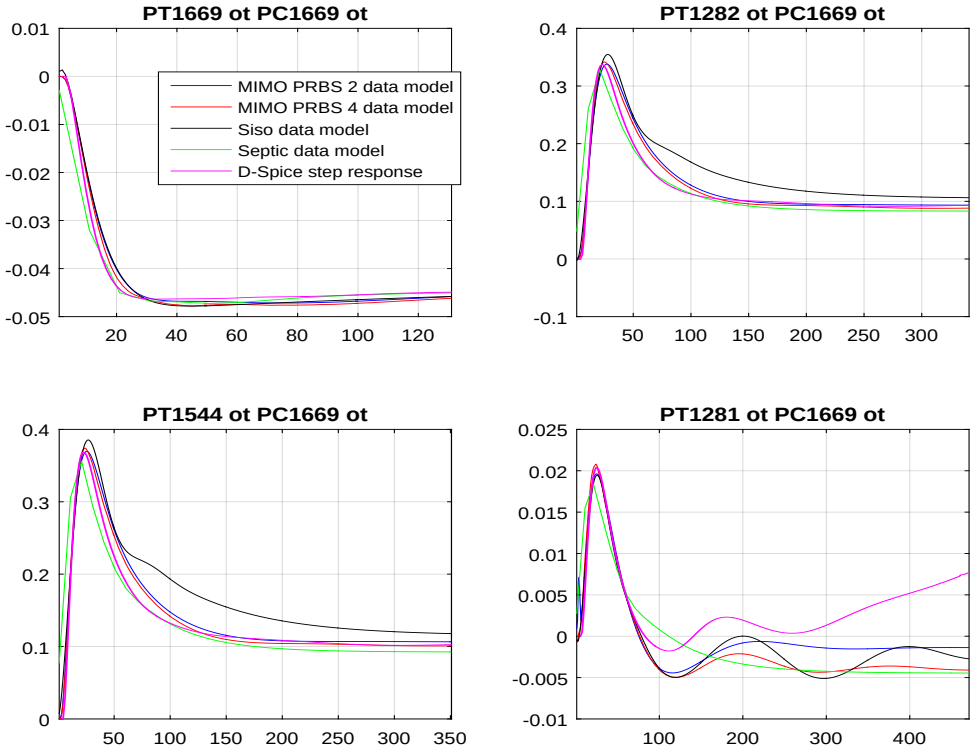


Figure 6.6 PIC1669 step of 5 up - simulator and model CV responses

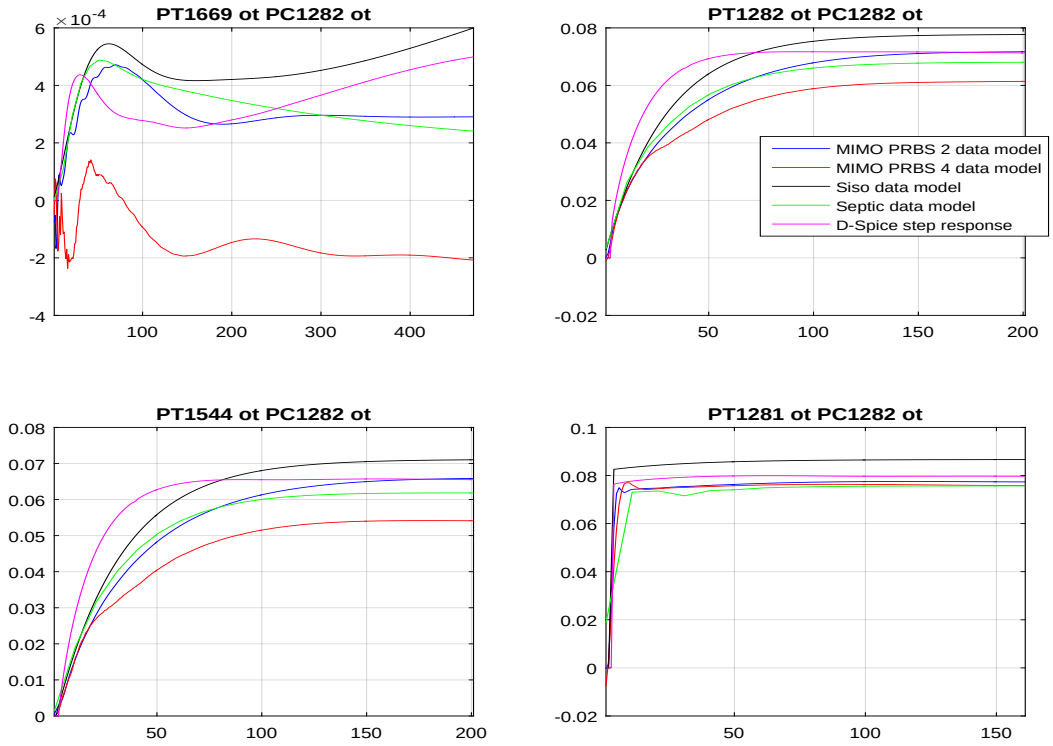


Figure 6.7 PIC1282 unit step down - simulator and model CV responses

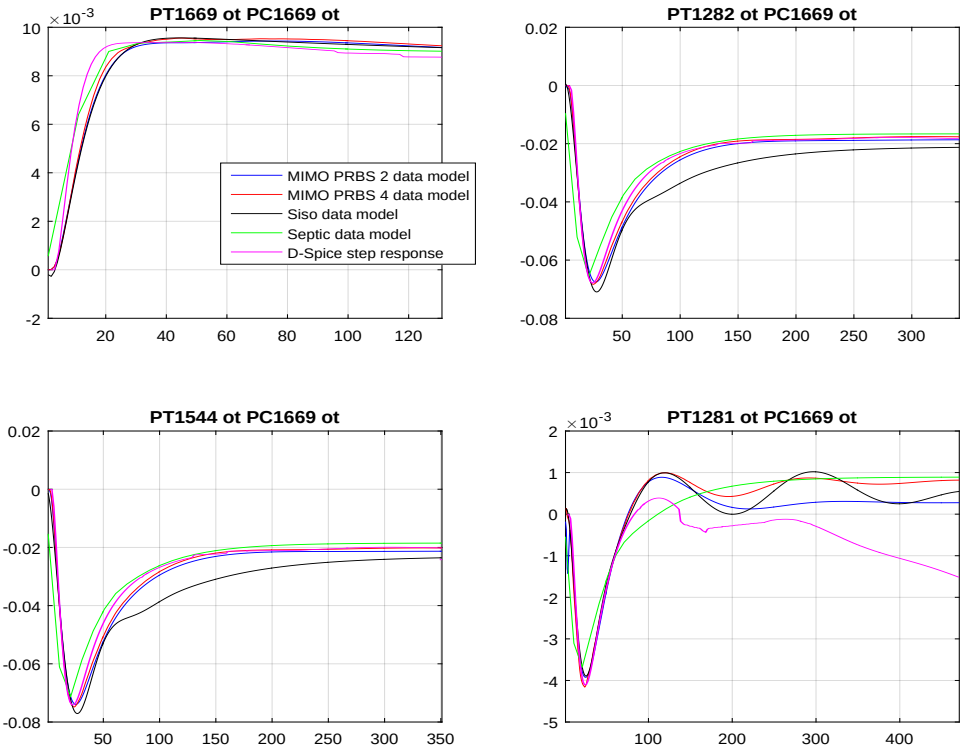


Figure 6.8 PIC1669 step of 1 down responses - simulator and model CV responses

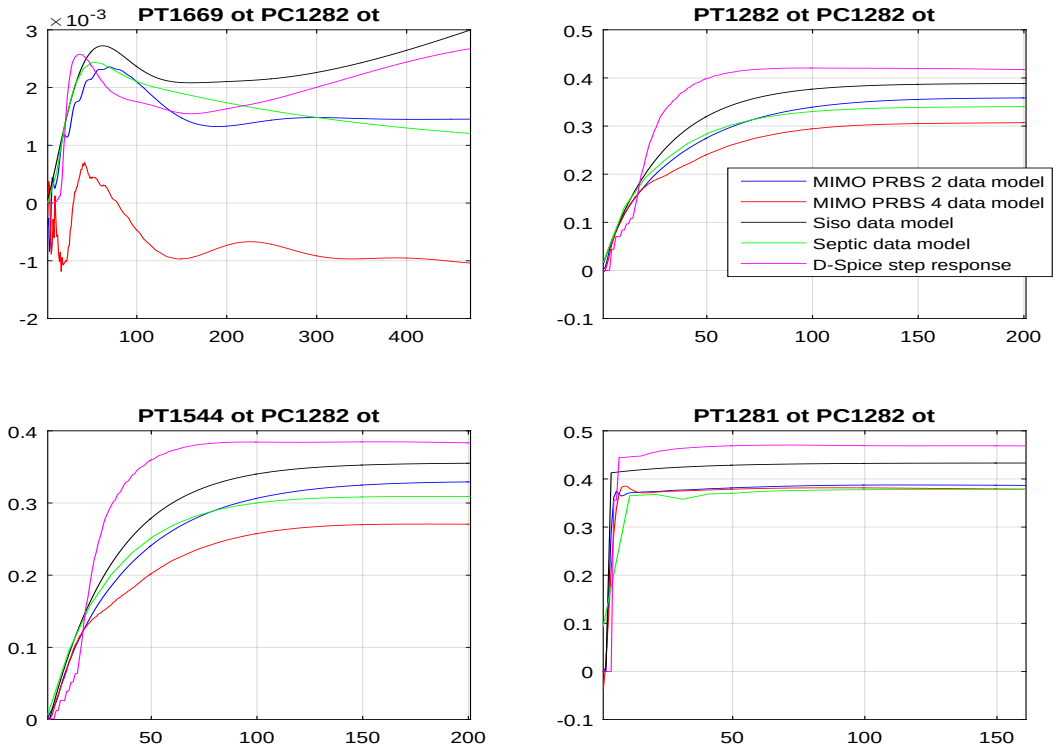


Figure 6.9 PIC1282 step of 5 down - simulator and model CV responses

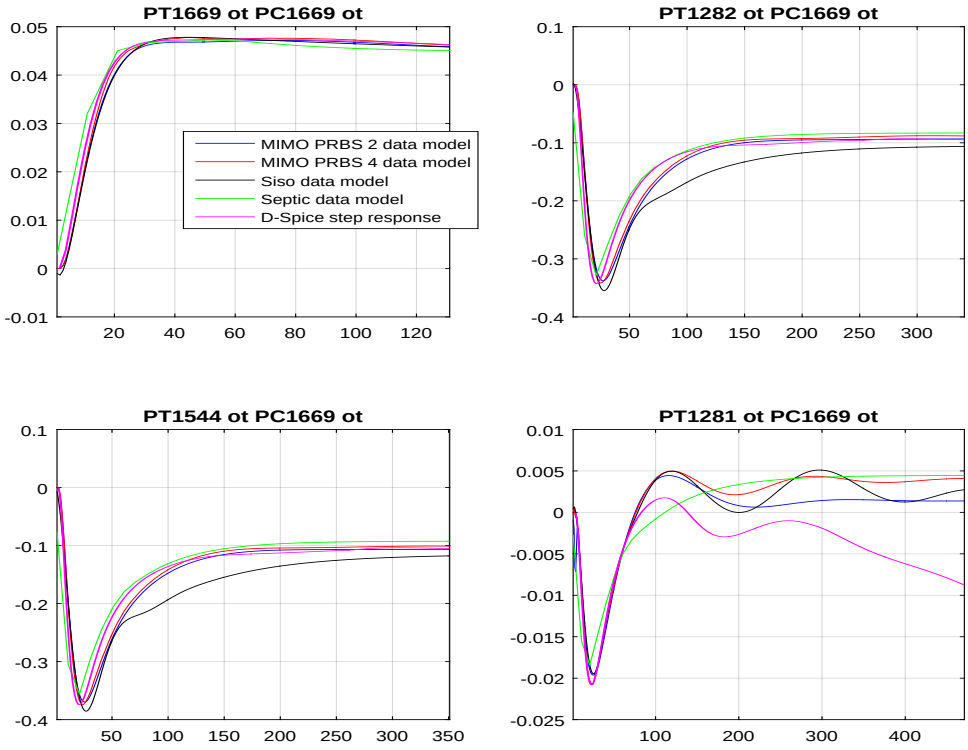


Figure 6.10 PIC1669 step of 5 down - simulator and model CV responses

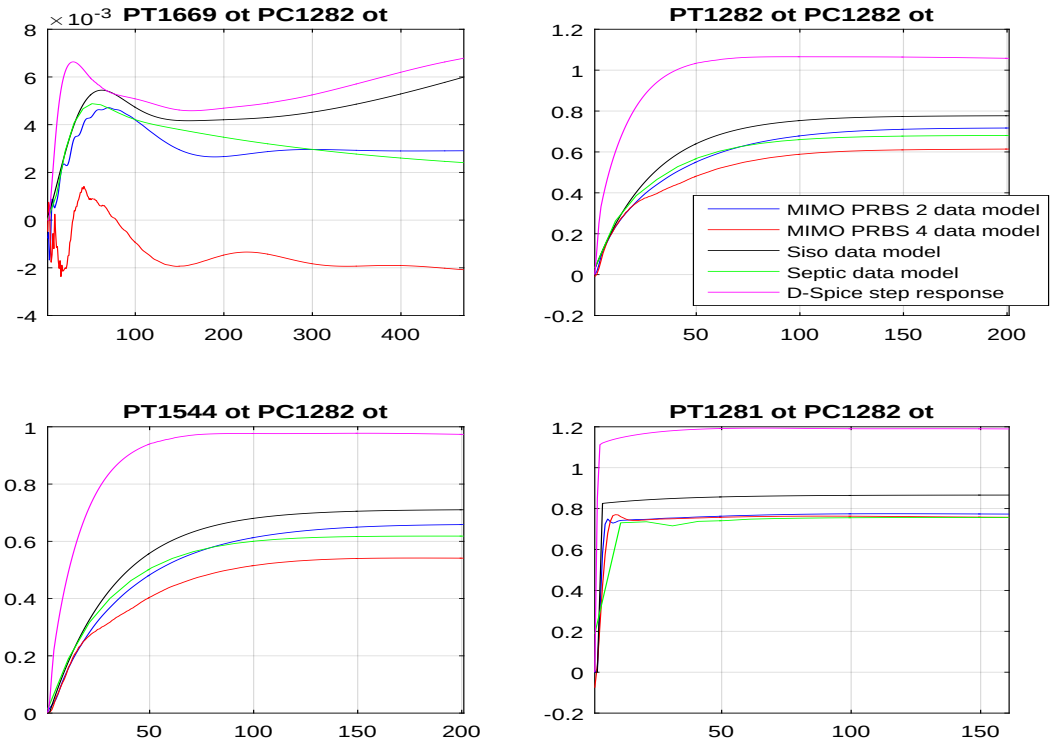


Figure 6.11 PIC1282 step of 10 down - simulator and model CV responses

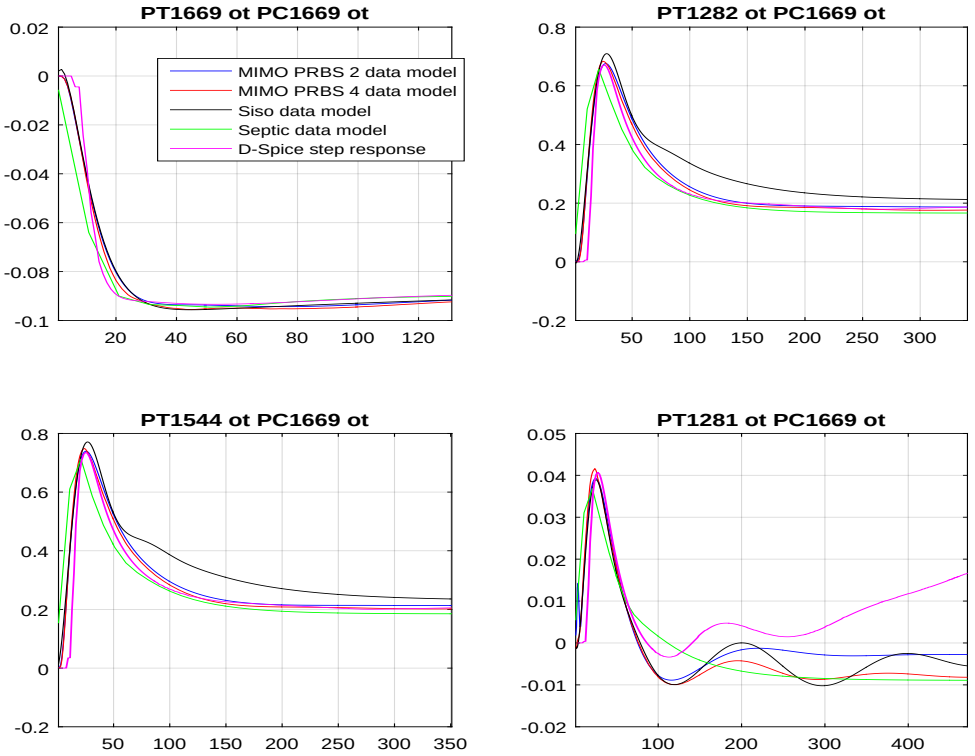


Figure 6.12 PIC1669 step of 10 up - simulator and model CV responses

As observed, there are some features which the identified models deviate from the simulator. As earlier emphasized, the most important aspect for comparison is the initial dynamics. Thus, steady-state deviations are of little interest due to the re-iterative scheme the MPC offers. However, in total the different models demonstrate satisfactory behaviour, and there are small differences between the models tested. Step response from controller PIC1282 to CV PT1669 depicts the biggest dynamical differences graphically. However, due to the small numerical values, these model deviations are not considered important, and considering the results in section 6.1.1 this was expected.

Chapter 7

Testing and simulations of two SEPTIC MPC applications

This chapter considers development of two MPC applications using Statoil's internal software SEPTIC. Since the author has not managed to find research on this particular problem, this work has been somewhat groundbreaking. The applications consider an identical process, and the difference is how the MPC is implemented in the control structure. Furthermore, an outline of the MPC extent and control problem is provided. Additionally, some notes on parameters regarding computational efficiency and controller tuning are provided. In closing, some simulation results are included.

7.1 Defining MPC extent

The initial extent of the MPC application was defined in Volden [88]. According to findings in chapter 4, the MPC applications were designed to comprise a 2x4 MIMO system including the initial MVs and adding an extra CV in the downstream expander pressure.

By analysing the regulatory layer and its corresponding variables for control, the control problem considered in this work was focused on pressure control through the subcooling cycle. The choice is based on some essential findings regarding selection of controlled variables for refrigeration cycles in Jensen [39]. Another prominent feature supporting this choice of MPC extent is the assumption of opposing CVs. This is further explained below figure 7.1.

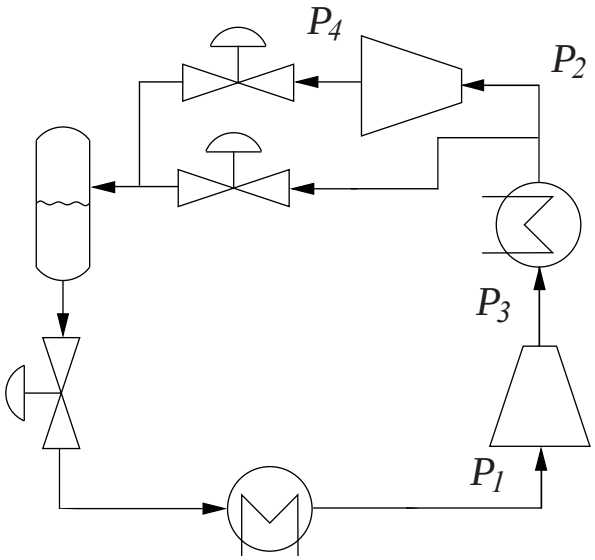


Figure 7.1 Principal sketch of the subcooling cycle with accompanying 4 CVs. Numbering of CVs correspond to Y1, Y2, Y3 and Y4 used throughout this work.

The current control structure maintain accurate pressure control in the subcooling refrigerant cycle. All 4 CVs, i.e., $P_1 \rightarrow P_4$, depicted in figure 7.1 are controlled by their own SISO controller. This essentially translates to all CVs having identical priorities. Thus, each CV's respective controller does not consider that its actions certainly influence the other CVs to a greater or lesser degree. Considering the subcooling cycle as a closed system, which in reality is not entirely true, it is however obvious that this pressure control scheme renders the system stiff and difficult to control in terms of pressure.

The subcooling cycle principally is considered as a closed system, but as shown in system responses throughout this work and mentioned in section 4.3.1, this is not the case. There are two system interactions outside the MFC systems which render the subcooling system open. One of these systems is the refrigerant make up system briefly mentioned in chapter 2 and further discussed in Volden [88]. Although this system provide additional flow in the subcooling cycle, it is not considered in this work. As the main task of the make-up system is to ensure correct refrigerant mixture, one may assume little interaction from this system in nominal operations. During a leakage this assumption would obviously fail, but this is beyond the scope of this work.

The major interaction stem from the fractionation system, where SMR is extracted upstream and added downstream the compressor. From the principal overview of the plant in figure 2.1, this interaction is identified between the MFC and the fractionation system. In nominal operations, this interaction is continuous and influence the subcooling cycle. Although, in terms of flow, the quantities supplied to the fractionation system is well below 10% of the total flow in the subcooling cycle, which makes the assumption of a closed cycle somewhat feasible. Additionally, as observed in the simulator data, the changes are dynamical and the cycle attain stationary behaviour within

reasonable time. As stated in section 1.3.1, interaction with the fractionation system was not completely correct since it failed to attain steady-state consecutive of perturbations. This feature is valid for running the individual MFC section of the simulator. However, the interaction was considered of minor significance and since it was discovered late in the development, time to investigate and resolve the error was not sufficient. Nevertheless, the author in cooperation with a Statoil employee verified a solution which involve to run the entire plant model. The consequence is significantly more computational load, and a simulator model incapable of running at speeds appreciable above real time. This certainly demonstrates the advantage of sharing simulator load over a network of computers as illustrated in Skjerven and Vist [66].

The initially launched MPC application considered 2 MVs and 4 CVs. After some initial runs, the application was expanded to include potential disturbances. Volden [88] lists initially considered disturbances. A new iteration on identifying potential disturbances considered the ability for measurement of the investigated disturbance variables. Thus, a revised summary of disturbances is presented in table 7.1.

Table 7.1 Considered variables for MPC applications

Considered MPC variables for the subcooling cycle applications		
MV's	**CV's**	**DV's**
Compressor speed	Compressor suction pressure	Sea water temperature
Pressure/flow control of expansion	Expansion turbine upstream pressure	Feed gas inlet flow from gas conditioning
	Compressor discharge pressure	Feed gas inlet temperature from gas conditioning
	Expansion turbine downstream pressure	LMR flow in liquefier cycle
		SMR inlet pressure on subcooler(downstream liquefier)
		SMR tank gas pressure
		SMR flow in subcooling cycle
		LNG temperature downstream subcooler
		Feed gas pressure upstream liquefier

Considering that some of the potential disturbances originate far upstream of the subcooling cycle, it was investigated how much impact each disturbance made on the CVs. A thorough analysis of 36 potential disturbance models was carried out. The conclusion was to discard four disturbances that simply had insignificant impacts for the subcooling cycle variables. These disturbances were essentially rejected upstream by regulatory layers of the precooling and liquefaction cycles. Hence, five disturbances were considered for further development.

Table 7.2 Actual variables for MPC applications

MPC variables for subcooling cycle applications		
MV's	**CV's**	**DV's**
Compressor speed	Compressor suction pressure	Sea water temperature
Pressure/flow control of expansion	Expansion turbine upstream pressure	SMR tank gas pressure
	Compressor discharge pressure	Feed gas pressure upstream liquefier
	Expansion turbine downstream pressure	LNG temperature downstream subcooler
		SMR inlet pressure on subcooler(downstream liquefier)

Appendix E.10 depicts a unit step response for a disturbance selected from table 7.2 on the Cvs considered in this work.

7.2 MPC configurations

Figure 7.2 depicts two control structures including MPC for a process plant. These control structures are outlined more in detail in Skogestad [68], but a brief presentation is given here. As observed in the figure, one control structure implements MPC and regulatory control in the same control layer as depicted in figure 2.9a. This configuration is well known, and the MPC manipulates setpoints to the regulatory layer controllers based on regulatory layer feedback. This is considered a common way of implementing MPC. Alternatively the MPC may be used to manipulate inputs directly, like indicated to the right in figure 7.2. To render a pure MPC approach as depicted in figure 7.2 feasible, an obvious requirement is the time scale. For fast systems, the sample time of the MPC application is particularly vital. However, for any application considered, the sampling interval defines upper bounds on the computational time. An advantage with this approach is the simplicity in which control is provided by a single application. Considering the supervisory MPC structure renders state of the regulatory layer vital. For example, altering parameters in the regulatory layer controllers affect closed-loop responses. This may degrade performance and overall operation if not regarded in the MPC application. Thus, this is a potential pitfall for the said control structure.

In the work of developing the MPC applications, the two control structures outlined above were the basis for development. Hence, one application was developed to provide setpoints to the regulatory layer, while the second application directly manipulate the process. Thus, the regulatory layer is removed.

For development of the supervisory MPC, the open-loop models identified were unsuitable due to the exclusion of controller models. Thus, a closed-loop identification was carried out to include

the regulatory layer for the supervisory MPC application. This identification was performed solely in SEPTIC, since the `N4SID` algorithm is not primarily tailored for closed-loop identification. The identified closed-loop models and some additional information on the identification for the 4 CVs are included in appendix E.9.

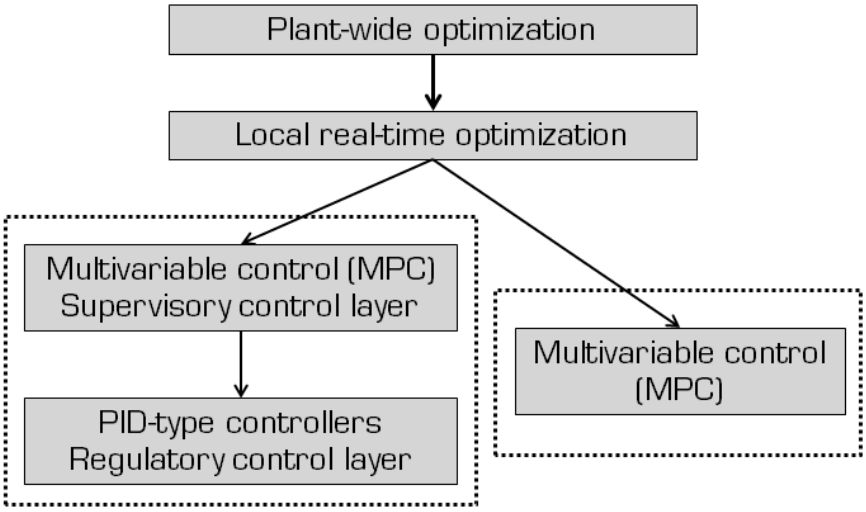


Figure 7.2 Two control structures including MPC. From Gjertsen [23].

7.3 Tuning and computational efficiency of MPC applications

To gain satisfactory process performance in closed-loop, several options pertain to the MPC application. There are different opinions on what parameters which primarily serves as tuning parameters. The author distinguishes between variables which primary purpose serve to increase efficiency/reduce computational load and tuning parameters. It should be noted that this distinction not necessarily is the common comprehension across industry and academia, but rather a personal perception.

In the general MPC formulation there exist several parameters for adjusting the closed-loop performance, and some general information on variables for increasing efficiency and tuning of MPC are found in e.g., Imsland [35], Camacho and Bordons [10] and Maciejowski [48]. This work will primarily consider the SEPTIC application specifically and its variables. However, in terms of parameters to reduce computational load/increase efficiency, SEPTIC resembles the general application, i.e., the parameters are well known and discussed in numerous papers and books. Although, some aspects of the variables for increasing efficiency are included as they are vital for optimizing the application.

7.3.1 Increasing computational efficiency

Reducing the degrees of freedom for the optimization problem reduces computational time. Cagienard et al. [9] present the number of degrees of freedom as the product of number of inputs times samples in the prediction horizon. Hovd [34] states that the solution time may grow cubic with the number of degrees of freedom for a QP problem. This emphasizes the importance of investigating the aspects of reducing computational load. There exist several approaches, and for the SEPTIC application four parameters are available; sample time, prediction horizon, control horizon and input blocking.

Sample time of the application is important as the controlled variables have different open-loop dynamics. Accordingly the desired closed-loop responses differ depending on the problem investigated, and consequently the upper bound on controller sampling is given by the shortest closed-loop response time. The sample time is essential, both regarding control and computational efficiency. Qualitatively, a decrease of the sampling time improves rejection of unknown disturbances especially, but for control purposes this is generally preferable. However, a sample time reduction increases the computational effort substantially. Thus, the optimal choice is a balance of performance and computational effort(Camacho and Bordons [10]).

For slow chemical processes, which initiated the use of MPC(Maciejowski [48]), the sampling time is slow compared to systems such as automotive, electrical and aerospace systems. However, in certain chemical processes, the sampling time vary substantially dependent on open-loop dynamics. The MFC process variables investigated in this work have a varying response time; the most time critical being the compressor pressure control, and the less time critical upstream expander pressure. A rule of thumb is to set the sample time in the range 10% to 30% of the smallest desired closed-loop response time(Maciejowski [48]). In applications where problem solving time exceeds the sampling interval, the explicit MPC option is recommended(Imsland [35]). However, for this application, the explicit MPC scheme is unnecessary since the MPC computational time not exceeds the bounds.

The prediction horizon is the number of future sampling intervals the controller must evaluate by prediction when optimizing its MVs at the sampling interval instants. The prediction horizon should ensure internal controller stability and that constraint violations are anticipated early enough to allow corrective action. A general rule of thumb is to set the prediction horizon approximately equal to the ratio between desired closed-loop response time and sample time. If the open-loop dynamics are unstable, the rule of thumb is to set the prediction horizon equal to the maximum sample intervals for a step response to become approximately infinite(Seborg et al. [64]). As increasing the prediction horizon allocates more degrees of freedom for the MPC to reach a solution, more computer power is required. An alternative to increasing the prediction horizon, is to enlarge the terminal region in order to more easily obtain feasible solutions. More on the terminal region is found in e.g. Hovd [34].

The control horizon is the number of MV moves to be optimized at sampling instant k. The control horizon is generally between 1 and the prediction horizon. Seborg et al. [64] provides two general

rules of thumb:

$$5 \leq M \leq 20 \quad \text{or} \tag{7.3.1}$$

$$N/3 < M < N/2 \tag{7.3.2}$$

where M and N is the control horizon and prediction horizon, respectively. Grüne [25] found an optimal control horizon: $M = N/2 + 1$, although this case considers a nonlinear MPC. At each sample instant, the optimized MV move at the beginning of the horizon is used and the others are discarded. This describes the receding horizon principle utilized in predictive control and estimation. A small control horizon means fewer variables to compute in the optimization problem solved at each control interval, which promotes faster computations. Though, a short control horizon may provide too few degrees of freedom to obtain satisfactory control.

Input blocking is a unique tuning parameter for each of the MVs. This technique reduces the number of optimization points and divides the optimal input sequence into time intervals with constant value. Furthermore, the input blocking is defined as a row vector covering the control horizon. Commonly, blocking intervals are of increasing length. Furthermore, the input blocking is defined as a row vector covering the control horizon. The first blocking element specifies the length of the first constant input interval. The last constant input interval, beginning directly after the last specified constant input interval, has a length equal to the remaining control horizon. An important aspect when considering input blocking and control horizon is to allow enough time to arrive at steady-state between the last control input block and the end of current control horizon.

In this thesis, the blocking is chosen to consist of seven and eight elements and is derived from the CV with the longest settling time, that is, compressor downstream and upstream expander pressure responses. This is in accordance with the rule of thumb from Seborg et al. [64]. Moreover, the blocking elements should be selected such that the initial prediction of optimal input sequence resembles the applied input. Hence, the input blocking is selected to be

$$\text{MV 1 input blocking} = [1\ 2\ 3\ 5\ 8\ 12\ 17\ 22] \tag{7.3.3}$$

$$\text{MV 2 input blocking} = [2\ 4\ 6\ 9\ 14\ 20] \tag{7.3.4}$$

The optimization problem is further simplified by evaluating specified points. The application automatically evaluate the optimization at the end of each input interval. In addition, a number of equally distributed evaluation points is specified by the developer, and are commonly within the range of 5-20. Furthermore, during dead-time or inverse responses, it is desirable to avoid evaluation of the MPC optimization problem. The MPC controller action to counteract these events could result in instability. The variable $EvalDT$ defines the number of samples to ignore initially on the prediction horizon for each of the CVs(Strand and Sagli [80]).

On the other hand, utilizing input blocking may result in stability issues and loss of constraint satisfaction guarantees. This may additionally have unfortunate effects on the closed-loop performance.

The reason behind is the impossibility of shifting input when blocked, and the issue of obtaining an input that satisfies the constraints for all future time. There are alternative blocking schemes treated in Cagienard et al. [9], but these are not included in SEPTIC. Although, one point from Cagienard et al. [9] is worth mentioning: input blocking tends to provide better results for stable systems compared to alternative blocking schemes. Since the models utilized in this work are stable, input blocking is accordingly the preferred scheme.

7.3.2 Tuning of MPC application

In appendix D, the optimization problem formulation in SEPTIC is given. To customise an MPC application for the problem considered, tuning is an intuitive way of doing so. Tuning of the SEPTIC application resembles to a certain degree LQR control, and the author will utilize analogies from LQR control to describe SEPTIC tuning. Thus, it is expected that the reader has knowledge on LQR control or is able to consult literature on said subject.

Tuning of the MPC application The SEPTIC scalar problem formulation in equation D.0.2, three scalar variables pertain to controller tuning. In the vector formulation D.0.1a these variables are referred to as weight matrices(equations D.0.1f, D.0.1g and D.0.1h). For simplicity, we consider the tuning parameters for the scalar formulation, and they are defined

$$q_{y,i} = \left(\frac{Fulf}{Span}\right)^2, \qquad q_{u,i} = \left(\frac{Fulf}{Span}\right)^2, \qquad p_i = \left(\frac{MovePnlty}{Span}\right)^2 \tag{7.3.5}$$

where $Span$ is a scaling parameter unique for all variables considered for control in the MPC application. The parameter defines the range of variation for the respective variable. Hence, for CVs where small moves are desired and expected, the $Span$ parameter is a low value. $Fulf$ is the parameter which states cost of deviation from setpoint and ideal value for the CVs and MVs, respectively. This resembles the Q matrix in LQR control, and increasing this cost lead to a tighter control for the variable in question. The cost of MV action is defined by the $MovePnlty$ parameter which resembles the R matrix from LQR control. This parameter alter the degree of MV freedom, and is for instance useful to reduce MV wear and tear on valves or pumps. Essentially, *Span* and to a small degree *MovePnlty* were the parameters utilized in MPC tuning in this work. As seen from above definitions(7.3.5), the $Span$ parameter alone influences all tuning variables. This certainly limits the confusion opposed to when tuning is performed using several variables.

To render focus on defined variables, SEPTIC includes a hierarchy of control priorities. In this work, the hierarchy is utilized to differentiate desired operating states for the variables in question. For example, the compressor suction pressure priority is considered vital and thus have a higher priority compared to upstream expander pressure setpoint. More on the general priority hierarchy feature is found in appendix D. However, for this application a ranking of CV priorities are found in table 7.3. In addition, the table list actual priority values implemented.

Table 7.3 Suggestion for SEPTIC priorities

CV	**Suggested CV priorities**	**Implemented SEPTIC priorities**
Y_1 low limit	1	1
Y_4 low limit	2	1
Y_1 high limit	3	1
Y_1 setpoint	4	2
Y_3 high limit	5	2
Y_3 low limit	6	3
Y_4 high limit	7	5
Y_2 high limit	8	8
Y_2 low limit	9	8
Y_2 setpoint	10	10
Y_4 setpoint	11	20
Y_3 setpoint	12	20

The difference in suggested priorities and implemented SEPTIC priorities are basically to better differentiate, but also equalize the considered aspects. One may believe that identical priorities given to several variables may cause a conflict, but this is not the case here. If the application encounter variables of identical priority, it solves the potential conflict by considering the applicable weight matrices and slack variables for the given variable.

For the Y_3 and Y_4 setpoints, the low priority emphasizes that no consideration shall be given these variables. This allows more freedom for the controller to focus on the more important priorities, and is a suggestion to make the pressure control less stiff as discussed in section 7.1.

7.4 MPC Simulations

This section includes the simulation results from testing of both MPC applications. For simplicity, the configurations are labelled supervisory and direct MPC. The supervisory application translates to the supervisory layer MPC, while the direct application translates to the isolated MPC without a regulatory layer. Both MPC applications were implemented with disturbance models, i.e., feed-forward configurations. For comparison reasons, the identical tests were carried out for the original control structure which involves the regulatory layer. Two main tests were executed: altering setpoints for CVs and rejection of the 5 disturbances defined in section 7.1. In addition, test on two variables to demonstrate the priorities were performed.

7.4.1 Changing setpoint of CVs

The figures obtained from simulations depict 6 subplots. For the case of supervisory MPC the top subplots depict setpoints for the regulatory controllers. However, for the cases of direct MPC or regulatory layer configurations, the top subplots illustrate controller outputs. Thus, setpoints are naturally excluded in said cases. The subsequent figures depict three tests; one test to display the MPC priority functionality, one test for manipulating the Y_1 setpoint, and a test for manipulating the Y_2 setpoint. In addition, appendices E.11 and E.12 provides more information on the priority functionality and disturbance rejection, respectively.

Interpreting the MPC priority functionality

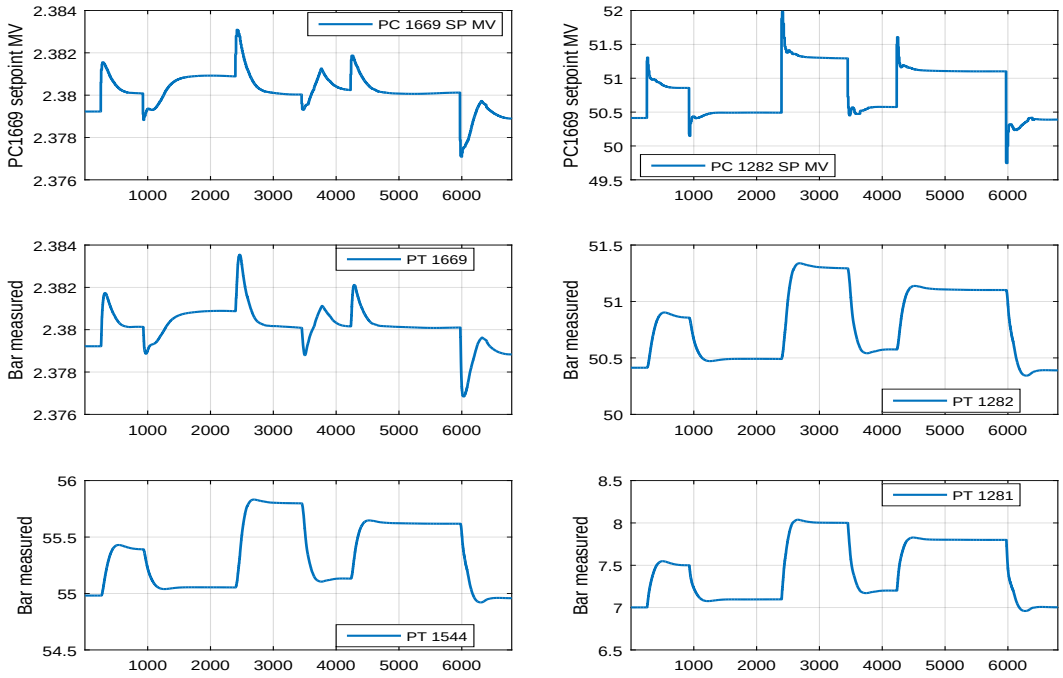


Figure 7.3 Y_4 setpoint manipulation for SEPTIC supervisory application. The priority is set to 1.

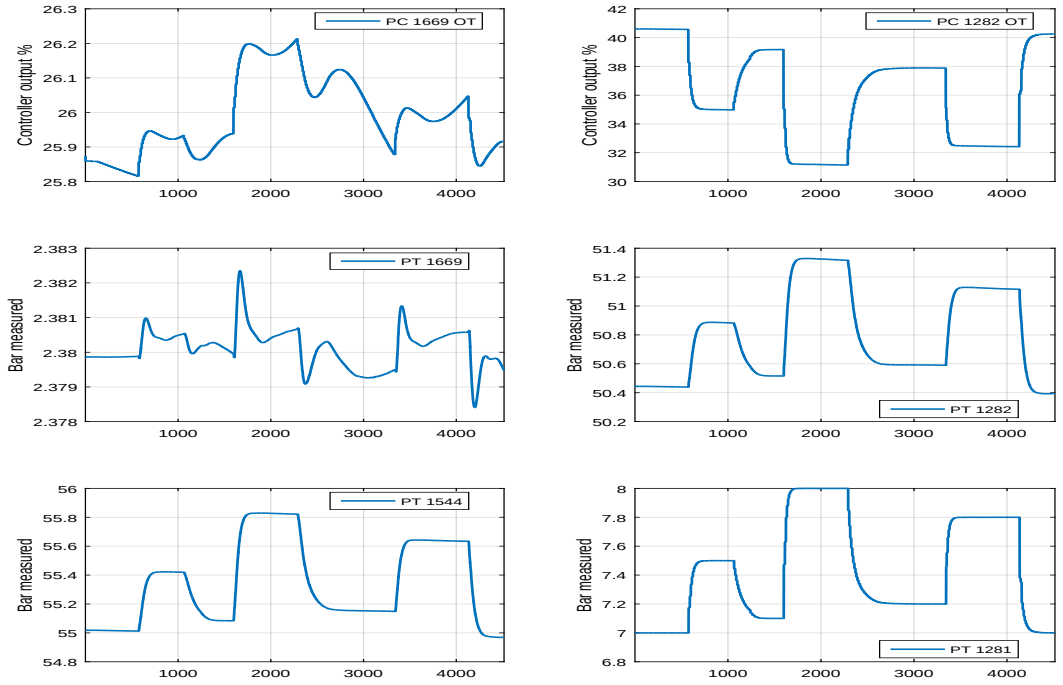


Figure 7.4 Y_4 setpoint manipulation for SEPTIC direct application. The priority is set to 1.

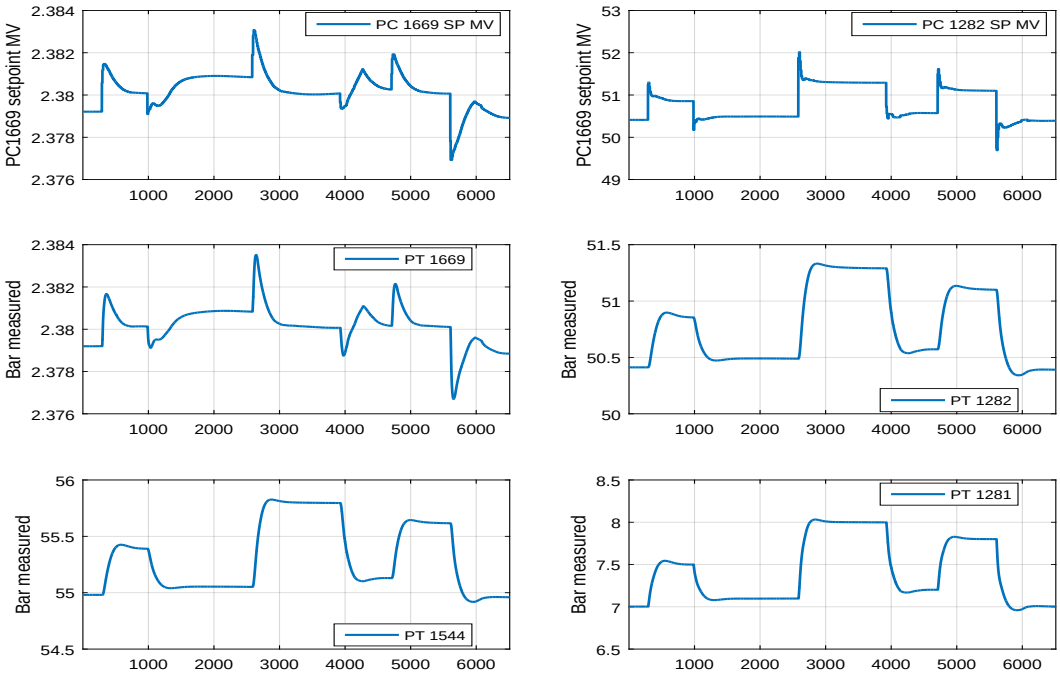


Figure 7.5 Y_4 setpoint manipulation for SEPTIC supervisory application. The priority for Y_4 setpoint is set to 10.

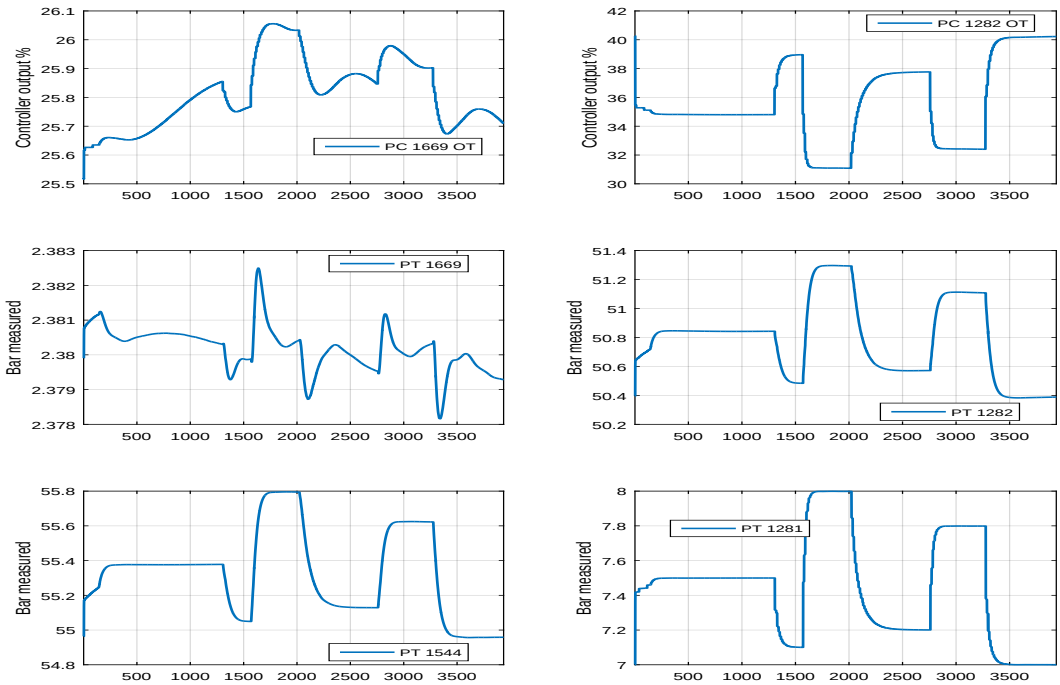


Figure 7.6 Y_4 setpoint manipulation for SEPTIC direct application. The priority for Y_4 setpoint is set to 10.

Independent on priority setting, control of the CVs are accomplished. Y_4 reaches the desired setpoints for both priorities. Some expected differences is depicted, such as slightly less control of

Y_1 when the priority for Y_4 is set to 1. Appendix E.11 illustrates this functionality further.

CV1 - PT1669 setpoint manipulation

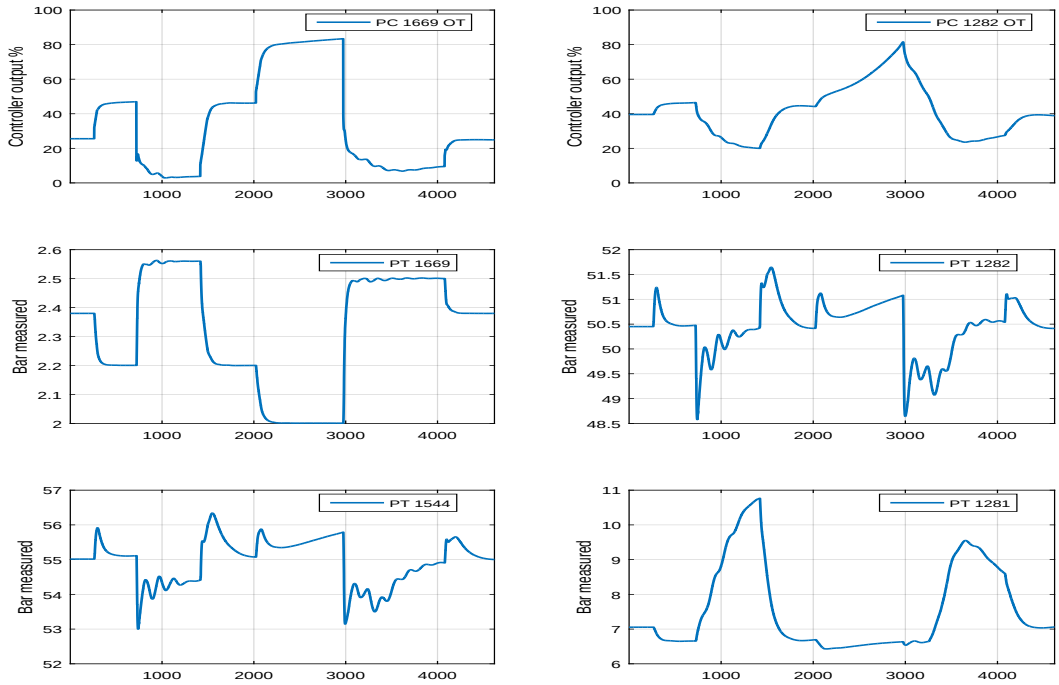


Figure 7.7 Y_1 setpoint manipulation from regulatory layer

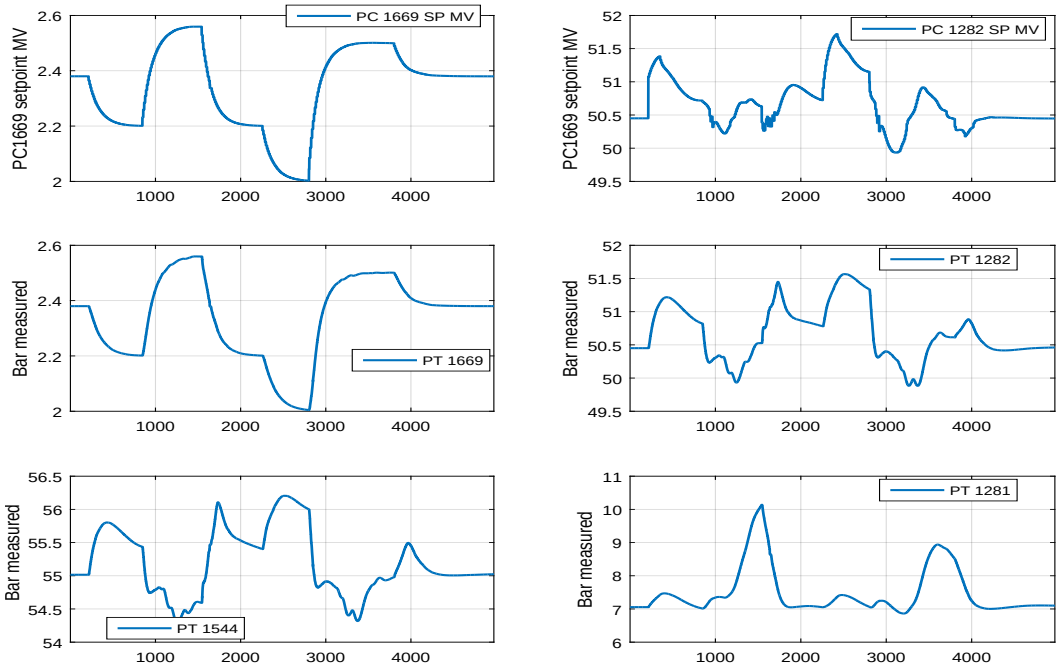


Figure 7.8 Y_1 setpoint manipulation for SEPTIC supervisory application. Priority is set to 2.

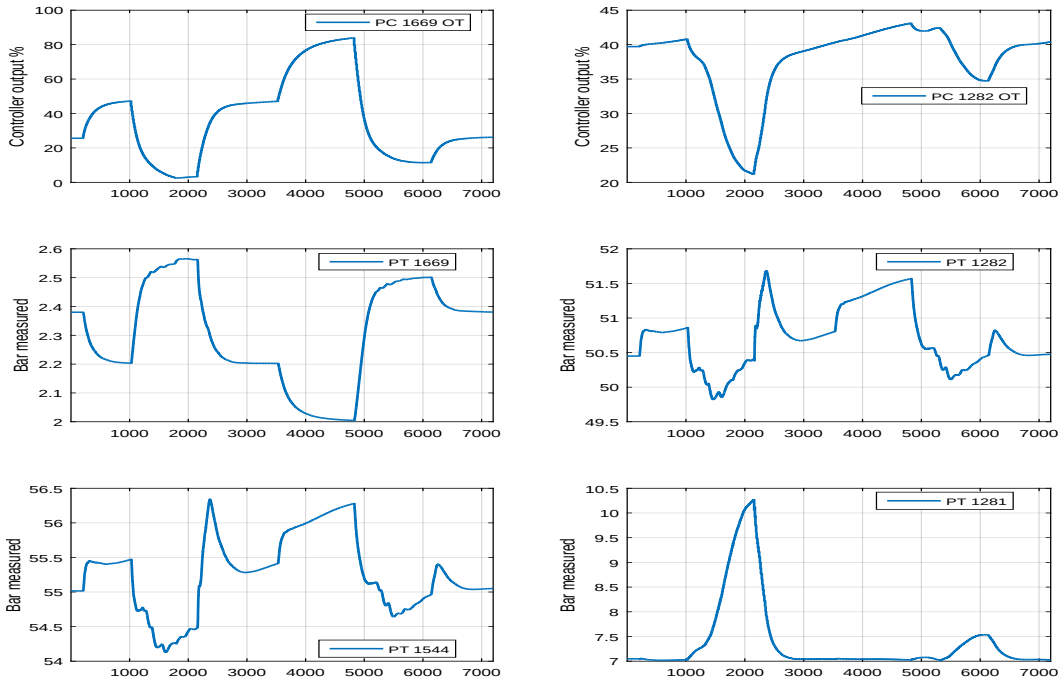


Figure 7.9 Y_1 setpoint manipulation for SEPTIC direct application. Priority is set to 2.

Altering the the setpoints for Y_1 reveals a more robust control by the MPC applications. The direct application is, however, substantially slower compared to the supervisory application. In addition, one main advantage of MPC control is depicted in the lower left subplots for Y_4: handling of CV constraints. For the regulatory layer we clearly observe the Y_4 pressure moving below 7 bar. For the MPC cases, this is not the case, where the Y_4 pressure does not move below 7 bar. This is due to a CV constraint which does not allow the pressure to move below 7 bar.

CV2 - PT1282 setpoint manipulation

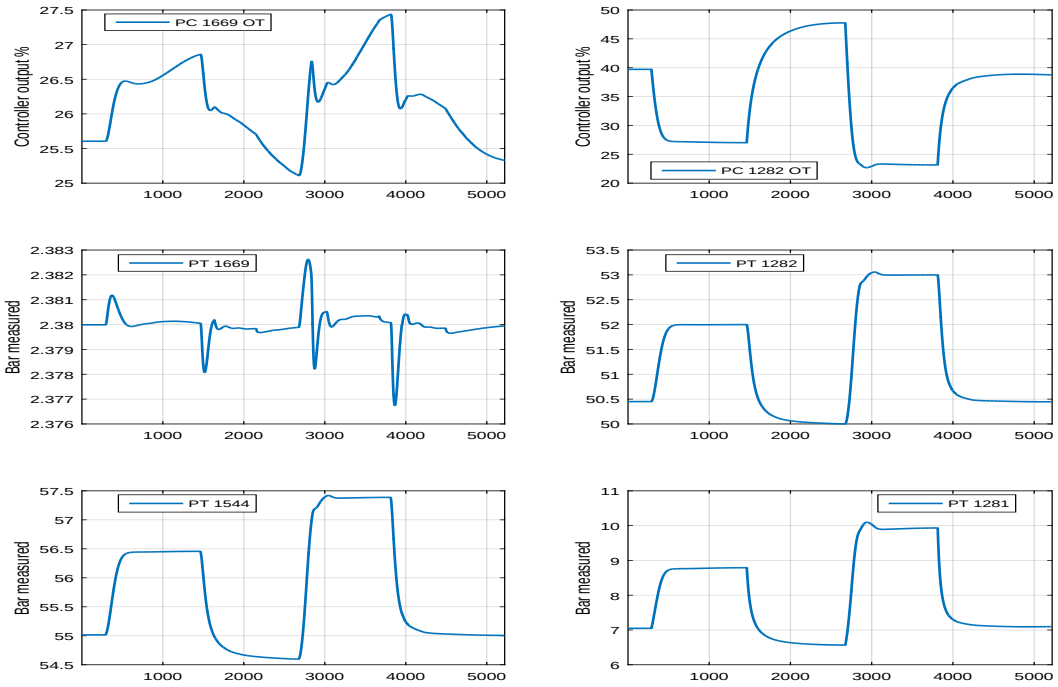


Figure 7.10 Y_2 setpoint manipulation from regulatory layer

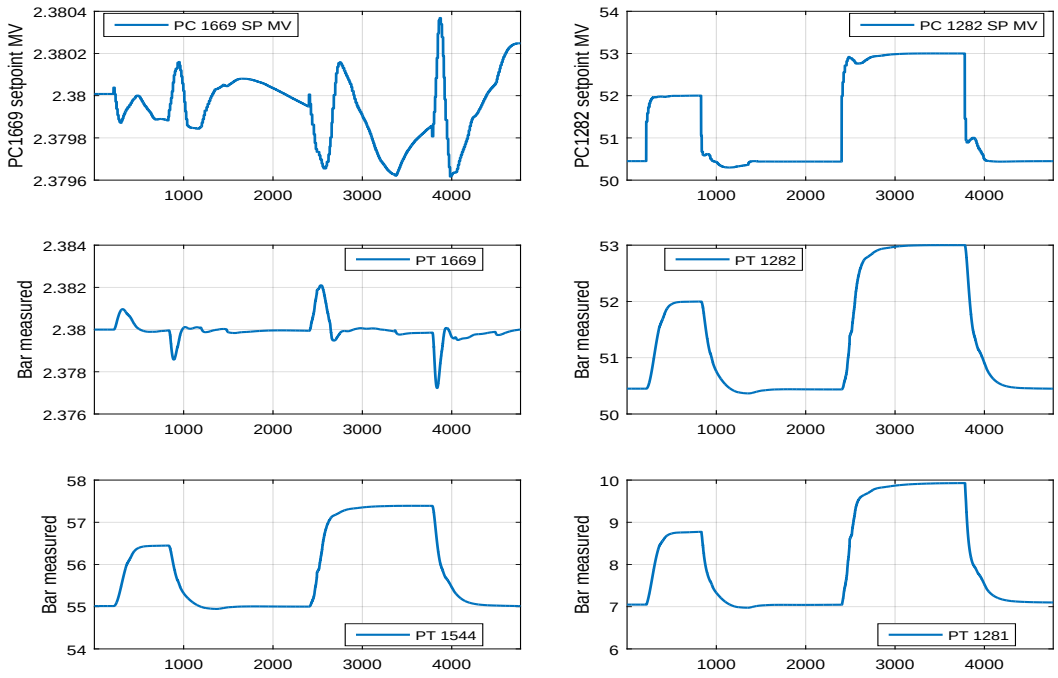


Figure 7.11 Y_2 setpoint manipulation for SEPTIC supervisory application. Priority is set to 10.

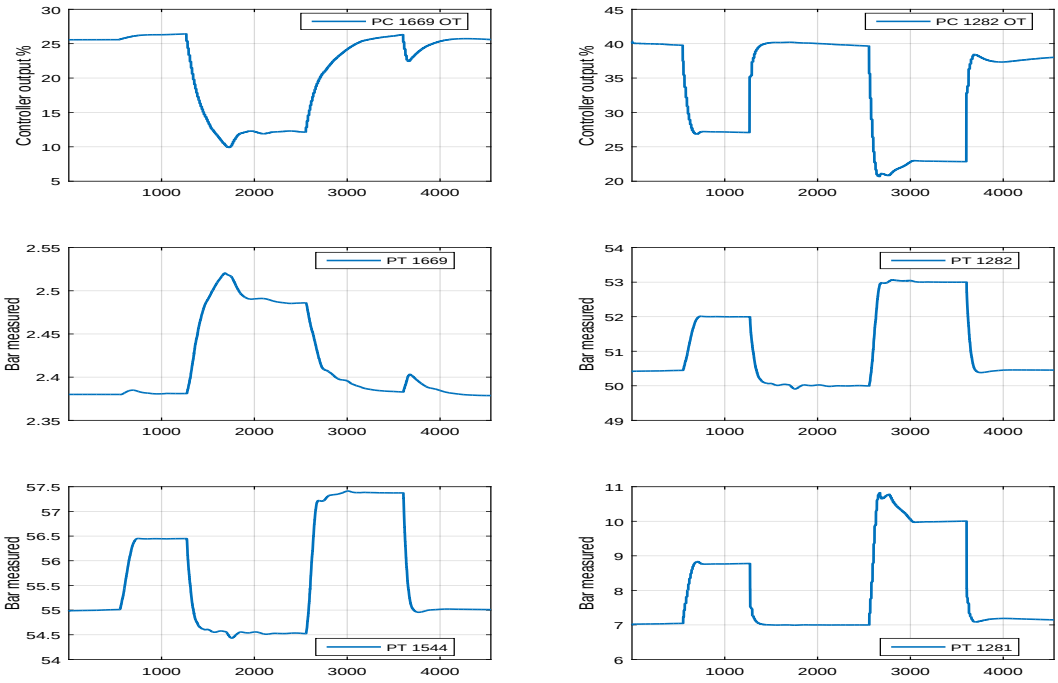


Figure 7.12 Y_2 setpoint manipulation for SEPTIC direct application. Priority is set to 1.

Similar features are recognized from setpoint simulations for Y_2. From the regulatory control we observe constraint violations for Y_4. In addition, the pressure Y_1 is slightly tighter controlled at setpoint for the supervisory MPC. For the direct MPC, the priority of Y_2 is set to 1, and thus, setpoint violations for Y_1 are allowed. However, no constraint violation occur for the MPC applications.

7.4.2 Rejection of disturbances

SMR vessel pressure

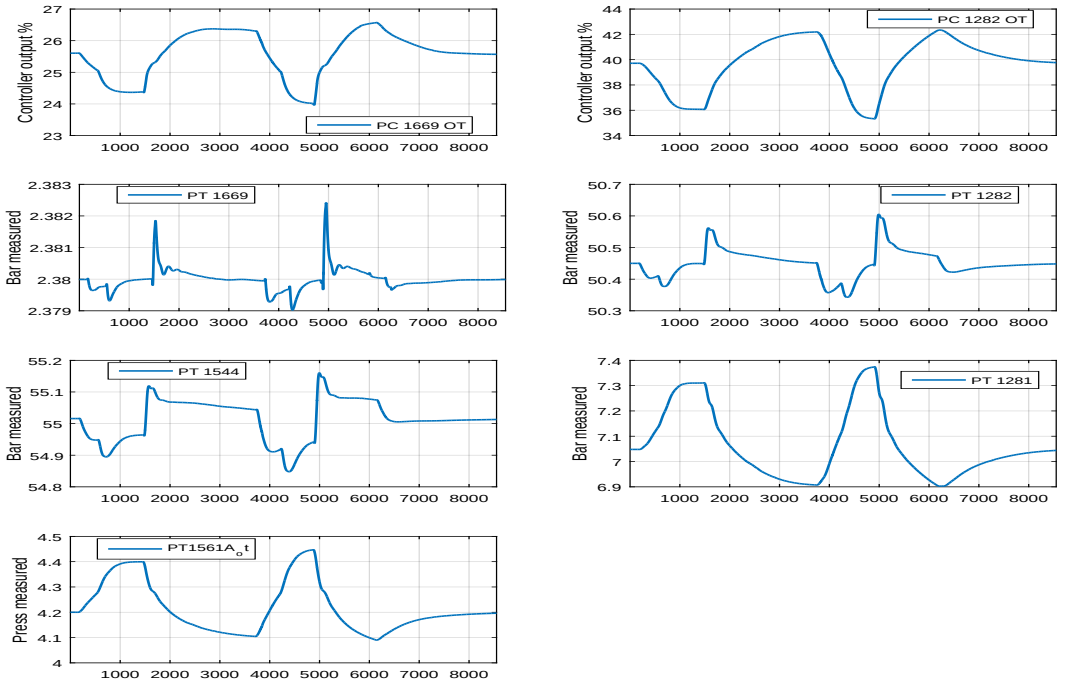


Figure 7.13 Rejection of SMR vessel pressure disturbance from regulatory layer. SMR vessel pressure is manipulated as depicted in bottom subplot.

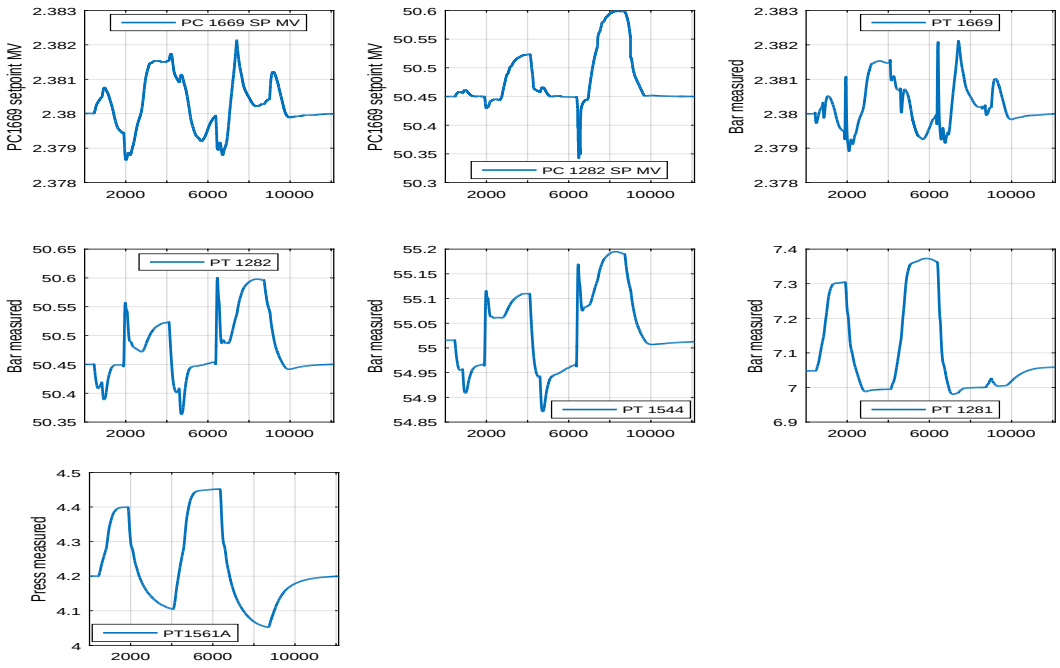


Figure 7.14 Rejection of SMR vessel pressure disturbance from SEPTIC supervisory application. SMR vessel pressure is manipulated as depicted in bottom subplot.

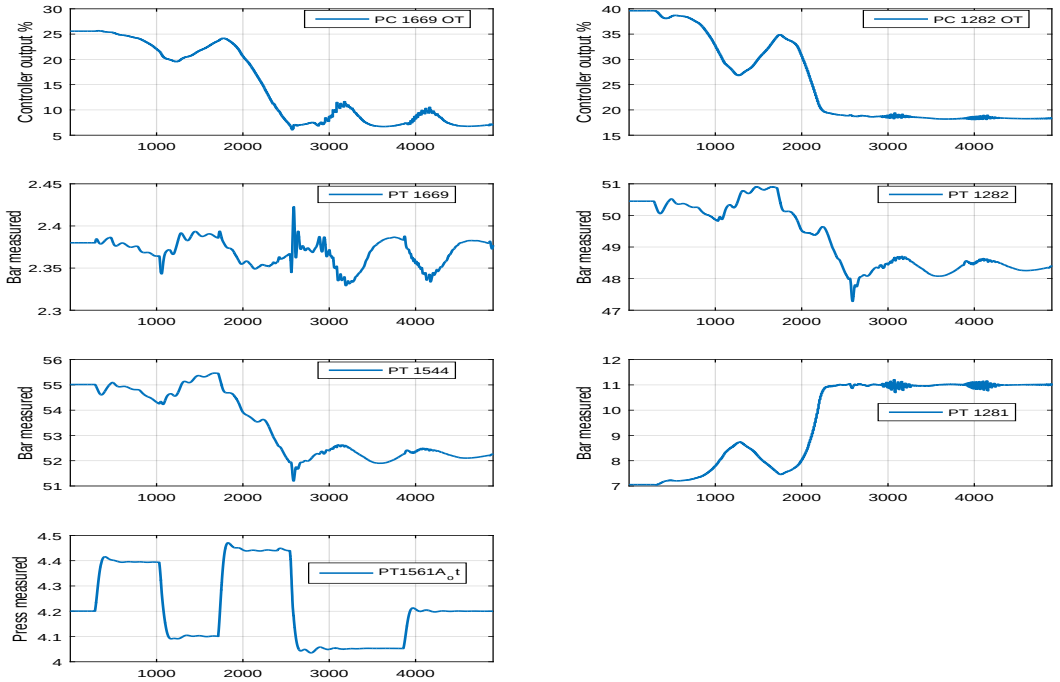


Figure 7.15 Rejection of SMR vessel pressure disturbance from SEPTIC direct application. SMR vessel pressure is manipulated as depicted in bottom subplot.

Considering pressure disturbance, we observe constraint violation for the regulatory case, while this is not the case for the MPC applications. The direct MPC application is a bit aggressive at the

consequence of less strict control for Y_1. To investigate this is not considered here, but the aspects discussed in 7.3 is certainly relevant for resolving this.

LNG temperature upstream subcooling cycle

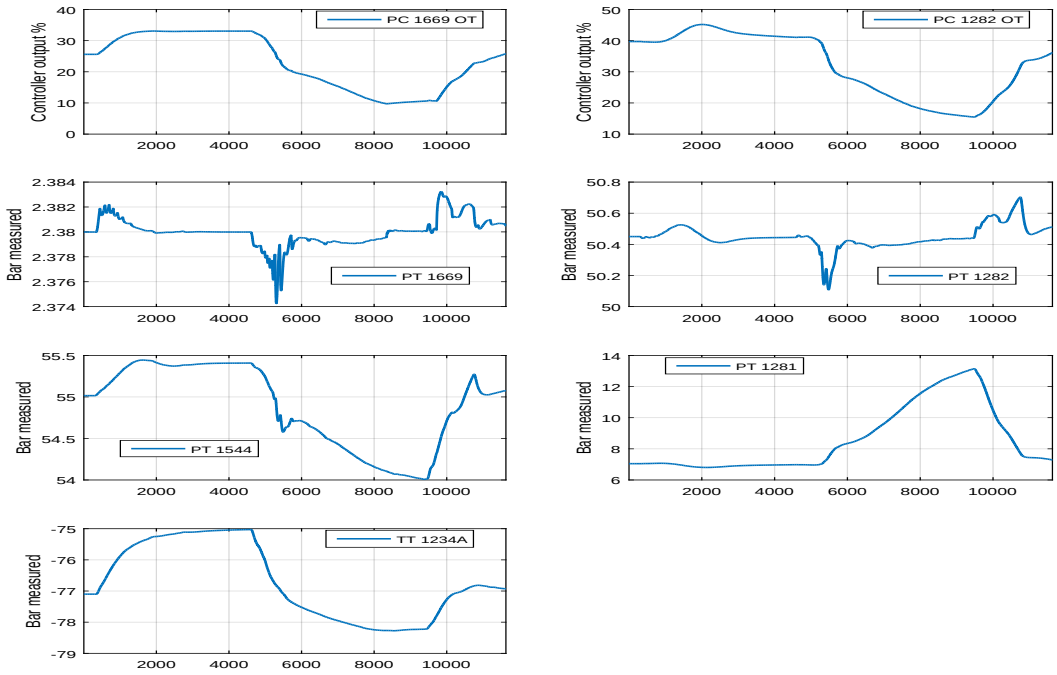


Figure 7.16 Rejection of LNG temperature disturbance entering subcooling cycle from regulatory layer. LNG temperature is manipulated as depicted in bottom subplot.

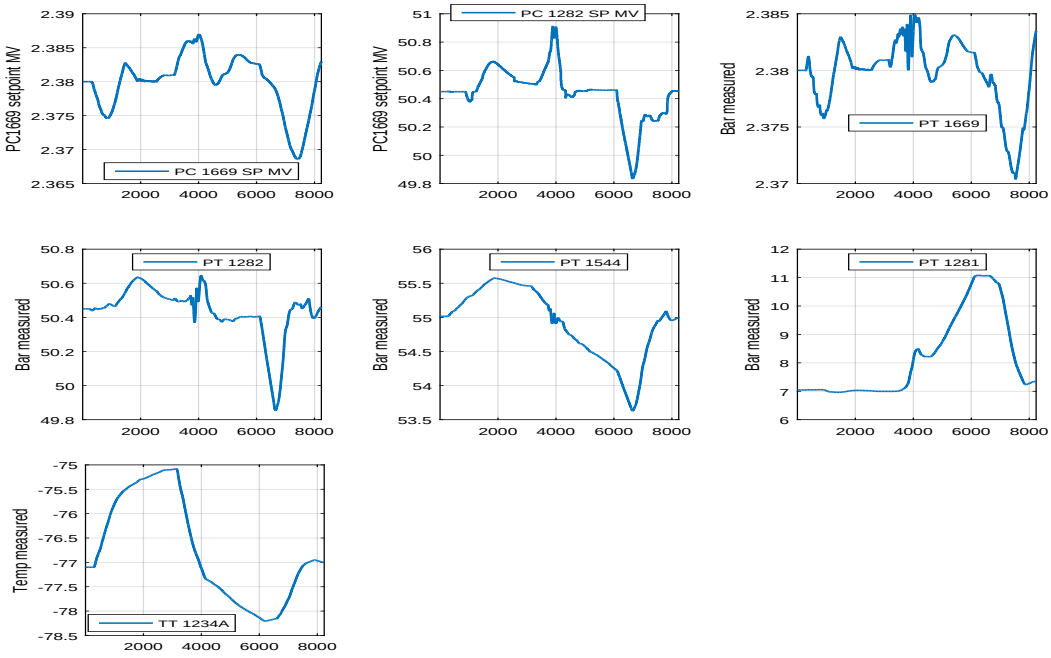


Figure 7.17 Rejection of LNG temperature disturbance entering subcooling cycle from SEPTIC supervisory application. LNG temperature is manipulated as depicted in bottom subplot.

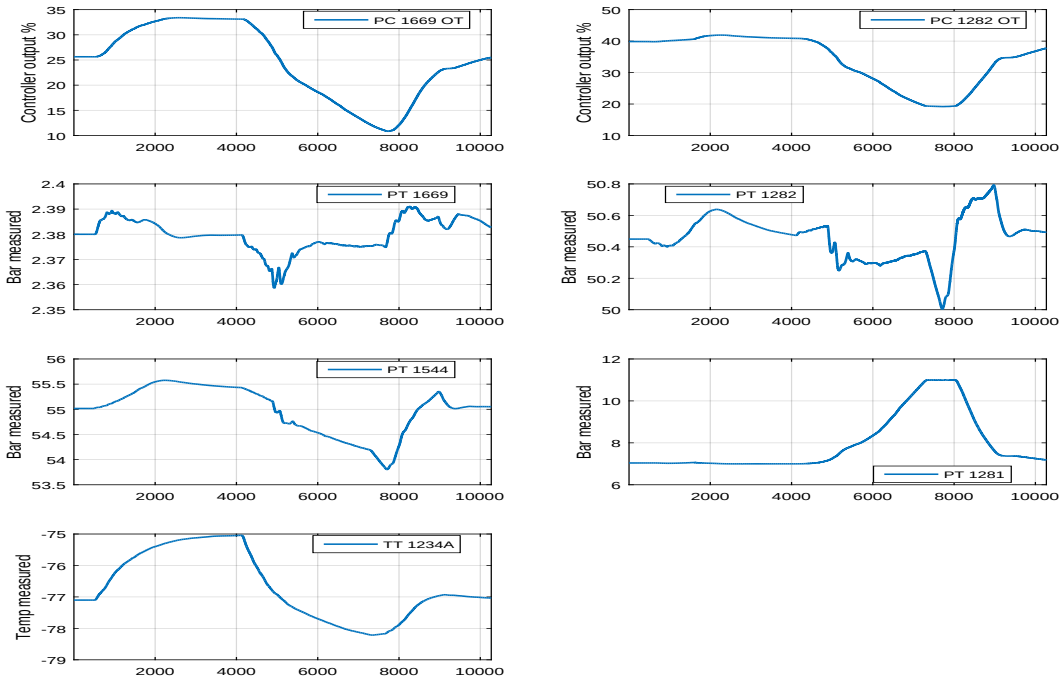


Figure 7.18 Rejection of LNG temperature disturbance entering subcooling cycle from SEPTIC direct application. LNG temperature is manipulated as depicted in bottom subplot.

The upstream LNG temperature disturbance depict somewhat inconclusive results. For the Y_1, regulatory control performs best. However, this includes a slight constraint violation for which the

MPC applications, certainly is not the case.

LNG temperature downstream subcooling cycle

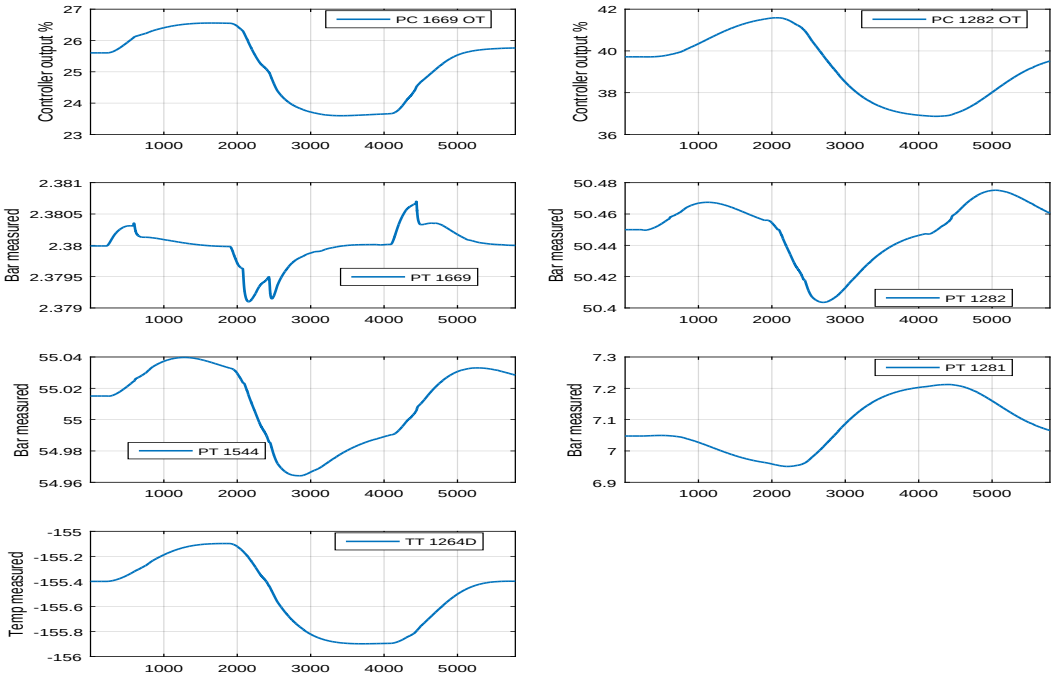


Figure 7.19 Rejection of LNG temperature disturbance downstream subcooling cycle from regulatory layer. LNG temperature is manipulated as depicted in bottom subplot.

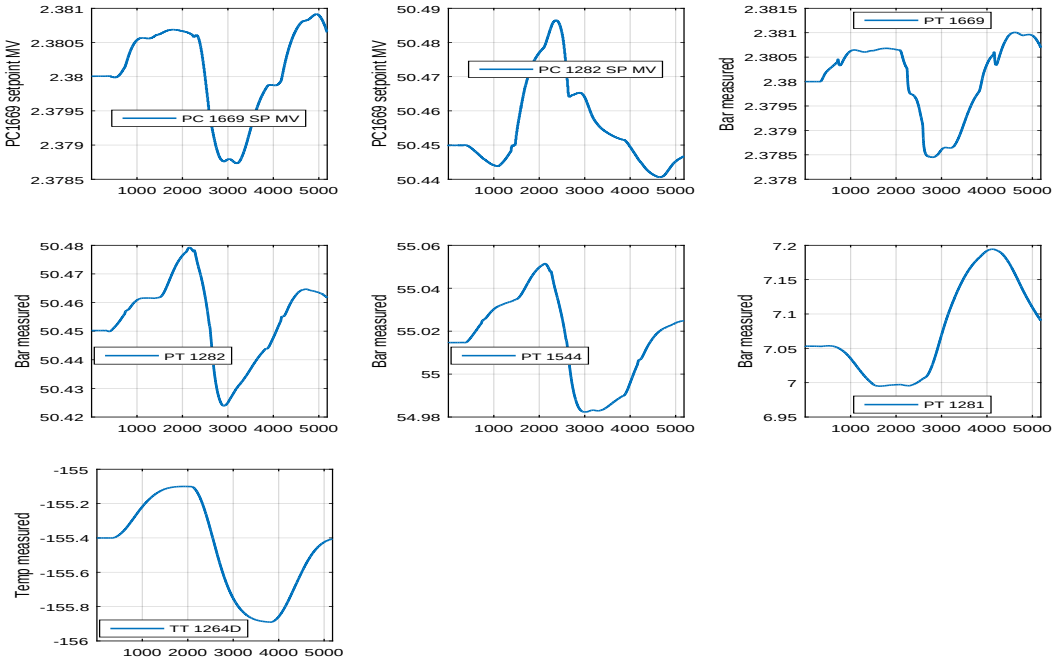


Figure 7.20 Rejection of LNG temperature disturbance downstream subcooling cycle from SEPTIC supervisory application. LNG temperature is manipulated as depicted in bottom subplot.

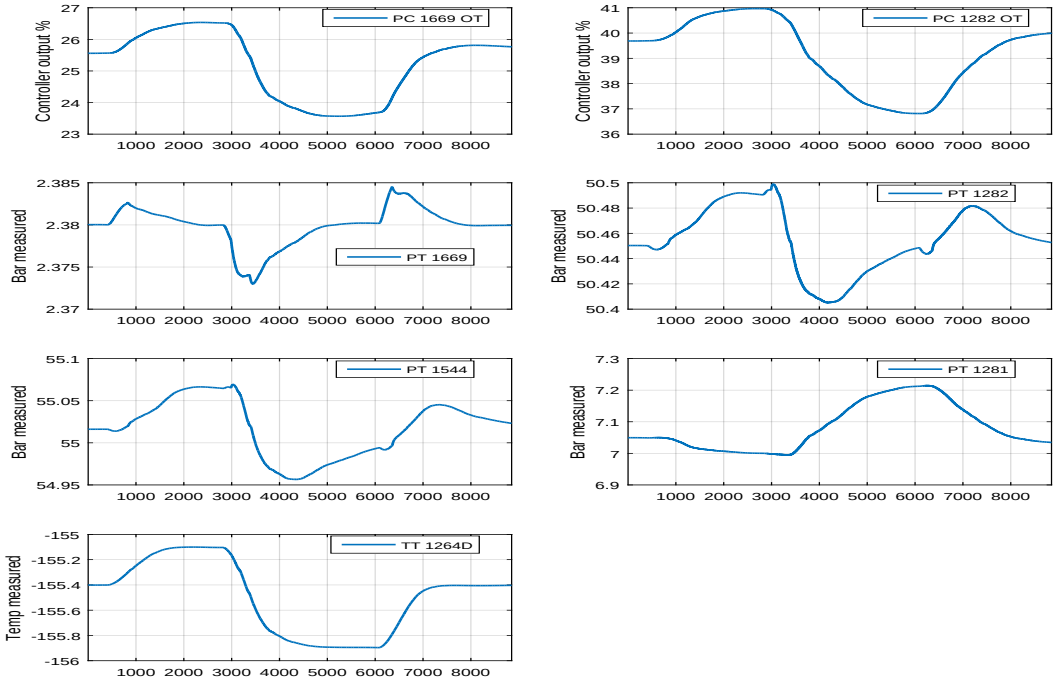


Figure 7.21 Rejection of LNG temperature disturbance downstream subcooling cycle from SEPTIC direct application. LNG temperature is manipulated as depicted in bottom subplot.

Considering varying LNG temperature downstream the subcooling cycle provides somewhat similar results as for the previous disturbance case. The control if Y_1 is equally proper for the regulatory

control and the supervisory MPC case. The direct MPC case provides slightly less tight control of said variable. Considering constraint violations, regulatory does not perform satisfactory as expected. Constraints are well handled for both MPC applications.

Chapter 8

Discussions, conclusions and suggestions for further work

8.1 Discussions and conclusions

The MFC process is a result of the Statoil/Linde alliance, and the control system was primarily designed by Linde. Upon production initialization in 2007, a study to illuminate potential sections suitable for more advanced process control was carried out. The MFC process section was investigated during this study. However, the conclusion rendered the MFC process immature for advanced process control. This was the foundation for the work carried out in this thesis.

The overall goal was to do another iteration on investigating possibilities for use of advanced process control of the MFC process at Hammerfest LNG. The control system for the MFC process is based on pure SISO control, and lacks supervisory and optimization layers. Further, the system is prone to disturbances and some loops are difficult to handle in closed-loop configuration. Additionally, during times when the sea water temperature peaks, the subcooling compressor reaches its operational limit and become a bottleneck for plant throughput.

From the work carried out in Volden [88], the basis was to obtain a satisfactory linear black box model for a chosen extent of the system, namely the subcooling cycle. However, to obtain said model was not a trivial task, and a laborious approach to diagnose why was executed. The answer was found in the control structure, where the extensive use of nonlinear control features render identification difficult. In particular, the use of selector and split range controllers complicated the identification process. Considering these findings imposed an altering of the original modelling extent, and the resulting system was defined to a 2x4 MIMO system.

Two methods for identification was essentially utilized. Subspace identification using the `N4SID` algorithm constitute the major part of this work. Additionally, classical system identification procedures such as step response, FIR and ARX modelling were utilized. To ensure validity of

the experiment data different identification experiments were performed . This includes both SISO and MIMO modelling. Considering the findings on control system, and making the necessary adjustments for identification provided significantly improved models. Based on the best initial model validation results, a selection of candidate models was considered further. A part of the work regarding subspace identification discusses the declared misconception that subspace identification is trivial.

A great deal of labour were used to analyse and verify the candidate models acquired. This part represents a major contribution to this thesis. A range of different measures to investigate the candidate models; numerical fitness criteria, graphical comparisons against simulator data, assessment of model uncertainties and residual analyses. In addition, the potential model reduction by using pole-zero cancellations was briefly investigated. To outline one of the main motivations behind this project and, in general advanced control, an assessment to emphasize the degree of interactions was carried out by using RGA and related analyses. These findings state that satisfactory decentralized control principally is difficult to obtain due to interactions. Based on these analyses a model for MPC implementation was chosen. Lastly, a verification based on simulator tests was carried out, and the results indicated satisfactory model behaviour.

The last part of this work considered MPC development and configuration. Two MPC applications for different control structures were developed. This is the supervisory and the direct MPC configurations in accordance with figure 7.2. Performance of both applications was tested through simulations of setpoint changes and manipulated disturbances. In terms of constraint handling, both MPC configurations demonstrated superior performance compared to the isolated regulatory layer. Although, there were some aspects regarding time spent to move CVs that need to be investigated further. The isolated regulatory layer outperforms the MPC applications in terms if this aspect for most of the simulations. The direct MPC application is almost consistently slower compared to the supervisory MPC. In terms of setpoint changes for CVs, the supervisory MPC shows promising tight control for accurate priorities. Although, for some cases the tightest control pertain to the isolated regulatory layer. A potential solution to this may be investigation and re-tuning of the regulatory layer for the supervisory MPC application. Due to the time constraint on this project, this is not considered further in this thesis. It is, however, a task suggested for further work.

In general, it is a laborious task to demonstrate if this system including MPC is or has the potential to reduce energy costs with the suggested applications and considerations on the control structure. Since the results are based on simulator data, one would indeed expect some differences between simulator data and plant operational data. However, the D-Spice simulator utilized in this work is not expected to possess the accuracy required to demonstrate such results. Optimization is nevertheless a difficult task for this system, and it certainly requires more rigorous modelling. Additionally, the complexity of the control system is not straightforward to handle. This may cause additional issues. In total, these factors rule out the feasibility to provide an unambiguous conclusion. Although, the work carried out in this thesis would hopefully be useful as a starting basis for further development.

8.2 Suggestions for further work

Undoubtedly, there are still issues to be solved before considering advanced process control for the MFC process feasible. By no means does the author acknowledge this development finished, and the following topics state suggestions for further work to continue on this thesis:

- There is a potential for improved performance for the supervisory MPC application by investigating and implementing a more aggressive tuning of the regulatory layer controllers. However, this is not a trivial task which require a substantially amount of labour.
- An MPC application based on this work may be utilized for the liquefaction system as the subcooling and liquefaction sections are similar. This ideally does not require much labour as the foundation for the application is laid out in this work.
- When considering the modelling aspect, gain scheduling may be useful as the figures 6.4 through 6.12 indicate. Simulator data reveal that responses from PC1282 perturbations especially are suited for gain scheduling. Ultimately, this may initialize a hierarchical control structure comprising several layers of MPC and RTO on top. This control structure is not uncommon for the downstream applications mentioned in Strand and Sagli [80], and thus well known to Statoil engineers.
- It should be noted that the depicted MPC performance is based on an initial application. No time has been devoted to consider tuning beyond the performance achieved in section 7.4. Thus, by no mean is the application considered complete, but hopefully findings in this thesis will be useful for further utilization.

For the specific applications considered in this work, there are two additional aspects which may be of interest for further development:

- Include disturbance model for PV1282B bypass valve.
- Extend MPC application to include plan throughput manipulator and LNG temp as an additional CV.

Bibliography

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd International Symposium on Information Technology*, pages 267–281.

[2] Akaike, H. (1974). Stochastic theory of minimal realization. *IEEE Transactions on Automatic Control*, 19:667–674.

[3] Bakke, M., Johansen, T. A., and Skogestad, S. (2010). Effect of varying control parameter in closed-loop subspace identification. In *9th International Symposium on Dynamics and Control of Process Systems*.

[4] Bako, L., Mercere, G., Lecoeuche, S., and Lovera, M. (2009). Recursive subspace identification of hammerstein models based on least squares support vector machines. *IET Control Theory and Applications*, 3(9).

[5] Bassett, S. and van Wijck, M. (1999). Application of Predictive Control Technology at BP's Crude Oil Terminal at Grangemouth. *IEE Colloquium Digest*, 95.

[6] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2013). *Time Series Analysis*. Wiley, 4 edition.

[7] Braun, M. W., Rivera, D. E., Stenman, A., Foslien, W., and Hrenya, C. (1999). Multi-level Pseudo-Random Signal Design and "Model-on-Demand Estimation Applied to Nonlinear Identification of a RTP Wafer Reactor. Technical report, Department of Electrical Engineering at Linköping University.

[8] Bristol, E. H. (1966). On a new measure of interaction for multivariable process control. *IEEE Transactions on Automatic Control*, 11:133–134.

[9] Cagienard, R., Grieder, P., Kerrigan, E. C., and Morari, M. (2004). Move blocking strategies in receding horizon control. In *43rd IEEE Conference on Decision and Control*.

[10] Camacho, E. F. and Bordons, C. (2007). *Model Predictive Control*. Springer, 2 edition.

[11] Chang, J.-W. and Yu, C.-C. (1990). The relative gain for non-square maultivariable systems. *Chemical Engineering Science*, 45(5):1309–1323.

[12] Chen, T. and Francis, B. A. (1995). *Optimal Sampled-Data Control Systems*. Springer.

[13] Deflorian, M. and Zaglauer, S. (2011). Design of Experiments for Nonlinear Dynamic System Identification. In *The 18th IFAC World Congress*.

[14] Downs, J. and Skogestad, S. (2011). An industrial and academic perspective on plantwide control. *Annual Reviews in Control*, 1:99–110.

[15] Fantoft (2004). *D-Spice User Guide*. Fantoft Process Technologies.

[16] Fantoft (2006). *Introduction to D-Spice*. Fantoft Process Technologies.

[17] Faurre, P. L. (1976). Stochastic realization algorithms. *Mathematics in Science and Engineering*, 126:1–25.

[18] Favoreel, W., Moor, B. D., and Overschee, P. V. (2000). Subspace state space system identification for industrial processes. *Journal of Process Control*, 10:149–155.

[19] Forssell, U. and Ljung, L. (1999). Closed-loop identification revisited. *Automatica*, 35:1215–1241.

[20] Foss, B. A., Lohmann, B., and Marquardt, W. (1998). A field study of the industrial modeling process. *Journal of Process Control*, 8:325–338.

[21] Gevers, M. (2004). Identification for control: achievements and open problems. Technical report, Center for Systems Engineering and Applied Mechanics (CESAME) Université Catholique de Louvain.

[22] Ghosh, P. (1999). Prediction of vapor-liquid equilibria using peng-robinson and soave-redlich-kwong equations of state. *Chemical Engineering and Technology*, 22(5):379–399.

[23] Gjertsen, F. (2014). A modeling framework for control of smart-scale tubular polymerization reactors. Master's thesis, NTNU.

[24] Glandt, E. D., Klein, M. T., and Edgar, T. E. (2001). *Optimization of Chemical Processes*. McGraw-Hill Chemical Engineering Series.

[25] Grüne, L. (2009). Analysis and design of unconstrained nonlinear mpc schemes for finite and infinite dimensional systems. *SIAM Journal on Control and Optimization*.

[26] Halvarsson, B. (2010). *Interaction Analysis in Multivariable Control Systems - Applications to Bioreactors for Nitrogen Removal*. PhD thesis, Uppsala University.

[27] Hauge, T. A., Slora, R., and Lie, B. (2005). Application and roll-out of infinite horizon mpc employing a nonlinear mechanistic model to paper machines. *Journal of Process Control*, 15:201 – 213.

[28] Heldt, S. (2011). *Near-optimal operation of LNG liquefaction processes by means of regulation*. PhD thesis, Technical University of Berlin.

[29] Hestetun, K. (2009). *Use of data from anode current distribution for state and parameter estimation and fault detection in an aluminium prebake electrolysis cell*. PhD thesis, NTNU.

[30] Hjalmarson, H. (2004). From experiment design to closed loop control. *Automatica*, 41:3939–438.

[31] Ho, B. and Kalman, R. (1966). Effective construction of state variables from input-output data. In *3rd Annual Allerton Conference on Circuit and System theory*, pages 449–459.

[32] Hof, P. M. J. V. D. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, 22:173–186.

[33] Hof, P. M. J. V. D. and Schrama, R. J. P. (1995). Identification and control - closed loop issues. *Automatica*, 31(12):1751–1770.

[34] Hovd, M. (2015). Lecture notes for the course advanced control of industrial processes.

[35] Imsland, L. (2007). Introduction to model predictive control.

[36] Ioannou, P. A. and Sun, J. (2012). *Robust Adaptive Control*. Wiley.

[37] Jacobsen, E. W., Lundström, P., and Skogestad, S. (1991). Modelling and identification for robust control of ill-conditioned pplant - a distillation case study. In *Proceedings of the American Control conference, Boston*.

[38] Jacobsen, E. W. and Skogestad, S. (1994). *The Modeling of Uncertainty in Control Systems: Proceedings of the 1992 Santa Barbara Workshop*, chapter Identification of Ill-conditioned plants — A benchmark problem, pages 367–376. Springer Berlin Heidelberg, Berlin, Heidelberg.

[39] Jensen, J. B. (2008). *Optimal Operation of Refrigeration Cycles*. PhD thesis, NTNU.

[40] Juricek, B. C., Seborg, D. E., and Larimore, W. E. (2001). Identification of the tennessee eastman challenge process with subspace methods. *Control Engineering Practice*, 9:1337–1351.

[41] Katayama, T. (2005). *Subspace methods for system identification*. Springer.

[42] Kolås, S. (2008). *Estimation in nonlinear constrained systems with severe disturbances*. PhD thesis, NTNU.

[43] Larimore, W. E. (1983). System identification, reduced-order filtering and modeling via canonical variate analysis. In *American Control conference*, pages 445–451.

[44] Ljung, L. (1999). *System Identification Theory for the User, second edition*. Prentice Hall.

[45] Ljung, L. (2003). Aspects and experiences of user choices in subspace identification. In *13th IFAC Symposium on System Identification*.

[46] Ljung, L. (2010). Perspectives on System Identification. *Annual Reviews in Control*, 34:1–12.

[47] Ljung, L. (2015). *System Identification Toolbox User's Guide*, 2015b edition.

[48] Maciejowski, J. (2002). *Predictive Control with Constraints*. Prentice Hall.

[49] Mercere, G., Lecoeuche, S., and Lovera, M. (2004). Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *International Journal of Adaptive Control and Signal Processing*, 18(9-10):771–797.

[50] Michelsen, F. A., Halvorsen, I. J., Lund, B. F., and Wahl, P. E. (2010a). Model and Simulation for Control of the TEALARC Liquefied Natural Gas Process. *Industrial & Engineering Chemistry Research*, 49:7389–7397.

[51] Michelsen, F. A., Lund, B. F., and Halvorsen, I. J. (2010b). Selection of Optimal, Controlled Variables for the TEALARC LNG Process. *Industrial & Engineering chemistry Research*, 49:8624–8632.

[52] Morari, M. and Lee, J. H. (1999). Model predictive control: past, present and future. *Computers and Chemical Engineering*, 23:667–682.

[53] Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer.

[54] Norwegian Petroleum Directorate (2015). Snøhvit facts (norwegian). http://factpages.npd.no/factpages/Default.aspx?culture=nb-no&nav1=field&nav2=PageView|All&nav3=2053062. [Accessed 2015-09-02].

[55] Overschee, P. V. and Moor, B. D. (1994). N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1):75–93.

[56] Overschee, P. V. and Moor, B. D. (1996). *Subspace Identification for Linear Systems. Theory - Implementation - Applications*. KLUWER ACADEMIC PUBLISHERS.

[57] Qin, S. J. (2006). An overview of subspace identification. *Computers & Chemical engineering*, 30:1502–1513.

[58] Qin, S. J. and Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764.

[59] Rivera, D., Lee, H., Braun, M., and Mittelman, H. (2003). Plant-fiendly system identification: a challenge for the process industries. In *13th IFAC Symposium on System Identification*.

[60] Ruscio, D. D. (1995). A method for the identification of state space models from input and output measurements. *Modeling, Identification and Control*, 16:129–143.

[61] Ruscio, D. D. (1997). A method for identification of combined deterministic stochastic systems. *Applications of Computer Aided Time Series Modeling*, 119:181–235.

[62] Schmidt, W. P., Ott, C. M., Liu, Y. N., and Wehrman, J. G. (2012). Arctic LNG Plant Design: Taking Advantage of the Cold Climate. [Internet: http://www.airproducts.com/~/media/Files/PDF/industries/lng/arctic-lng-plant-design.pdf][Accessed 2015-09-30].

[63] Schwimmbeck, R. G. (2008). Pipeline vs. LNG. 3rd Pipeline Technology Conference.

[64] Seborg, D. E., Edgar, T. F., Mellichamp, D. A., and Doyle, F. J. (2011). *Process Dynamics and Control*. Wiley.

[65] Sima, V., Sima, D. M., and Huffel, S. V. (2004). High-performance numerical algorithms and software for subspace-based linear multivariable system identification. *Journal of Computational and Applied Mathematics*, 170:371–397.

[66] Skjerven, Ø. and Vist, S. (2007). Snøhvit lifecycle simulator from wellhead through pipeline and LNG liquefaction to offloading. In *PS7-3, 15th International Conference on Liquefied Natural Gas*.

[67] Skofteland, G., Strand, S., and Lohne, K. (1998). Successful use of a model predictive control system at the statfjord a platform. In *IIR Conference on Offshore Separation Processes*.

[68] Skogestad, S. (2004). Control structure design of complete chemical plants. *Computers and Chemical Engineering*, 28:219–234.

[69] Skogestad, S. and Postlethwaite, I. (2005). *Multivariable Feedback Control.Analysis and Design*. Wiley.

[70] Söderström, T. and Stoica, P. (2001). *System Identification*. Prentice Hall.

[71] Sotomayor, O. A. Z., Park, S. W., and Garcia, C. (2003). Model reduction and identification of wastewater treatment plants - a subspace approach. *Latin American Applied Research*, 33:135–140.

[72] Statoil (2005). *Detailed Process Description with Experience Transfer*. Statoil.

[73] Statoil (2006). *Overall Operation Manual - Overall Plant Operation*. Statoil.

[74] Statoil (2012). *SO Documentation system 25 - LNG Liquid Natural Gas*. Statoil.

[75] Statoil (2013). Carbon capture and storage. http://www.statoil.com/en/TechnologyInnovation/NewEnergy/Co2CaptureStorage/Pages/Snohvit.aspx. [Accessed 2015-08-30].

[76] Statoil (2014). *System 25 - LNG Liquefaction Operation Manual*. Statoil.

[77] Statoil (2015a). Snøhvit facts. http://www.statoil.com/en/ouroperations/explorationprod/ncs/snoehvit/pages/default.aspx. [Accessed 2015-09-28].

[78] Statoil (2015b). Snøhvit facts (norwegian). http://www.statoil.com/no/ouroperations/explorationprod/ncs/snoehvit/pages/default.aspx. [Accessed 2015-09-30].

[79] Strand, S. (2011). Mpc in statoil. Presentation. [Accessed 02.09.2015].

[80] Strand, S. and Sagli, J. R. (2003). Mpc in statoil - advantages with in-house technology. pages 97–103.

[81] Sturm, W., Parra-Calvache, M., Chantant, F., and van Opstal, J. (2009). Unlocking the Potential of Modern Control and Optimization Strategies in LNG Production. In *Proceedings of the first Annual Gas Processing Symposium*.

[82] Tangirala, A. K. (2015). *Principles of System Identification: Theory and Practice*. CRC Press.

[83] Teknisk Ukeblad (2008). Snøhvit article (norwegian). http://www.tu.no/nettarkiv/2008/01/29/snohvit-er-pa-bedringens-vei. [Accessed 29.09.15].

[84] van de Wal, M. and de Jager, B. (2001). A review of methods for input/output selection. *Automatica*, 37:487–510.

[85] Veerhagen, M. (1994). Identification of the deterministic part of mimo state space models given on innovations form from input output data. *Automatica*, 30(1):61–74.

[86] Venkatarathnam, G. (2008). *Cryogenic Mixed Refrigerant Processes*. Springer.

[87] Viberg, M. (1995). Subspace-based methods for the identification of linear time-invariant systems. *Automatica*, 31(12):1835–1851.

[88] Volden, A. B. (2015). On The Investigation of a Model Predictive Controller Applied to a Natural Gas Liquefaction Process.

[89] Wittenmark, B. and Salgado, M. E. (2002). Hankel-norm based interaction measure for input-output pairing. In *15th Triennial World Congress*.

[90] Zhu, Y. (2000). Automatic closed-loop identification package for model based process control. In *IFAC symposium on System Identification*.

[91] Zhu, Y. and Butoyi, F. (2002). Case studies on closed-loop identification for mpc. *Control Engineering Practice*, 10(4):403–417.

Appendix A

A generalized system identification procedure

There exist numerous examples of a system identification procedures, some more superficial than others. The author list the following example, taken from Hauge et al. [27], which provides thorough information on the various sequences in this somewhat iterative scheme of research. Some aspects of this particular procedure may be excess for certain cases, while other cases may require additional routines to achieve satisfactory results. However, the illustration captures all essential aspects.

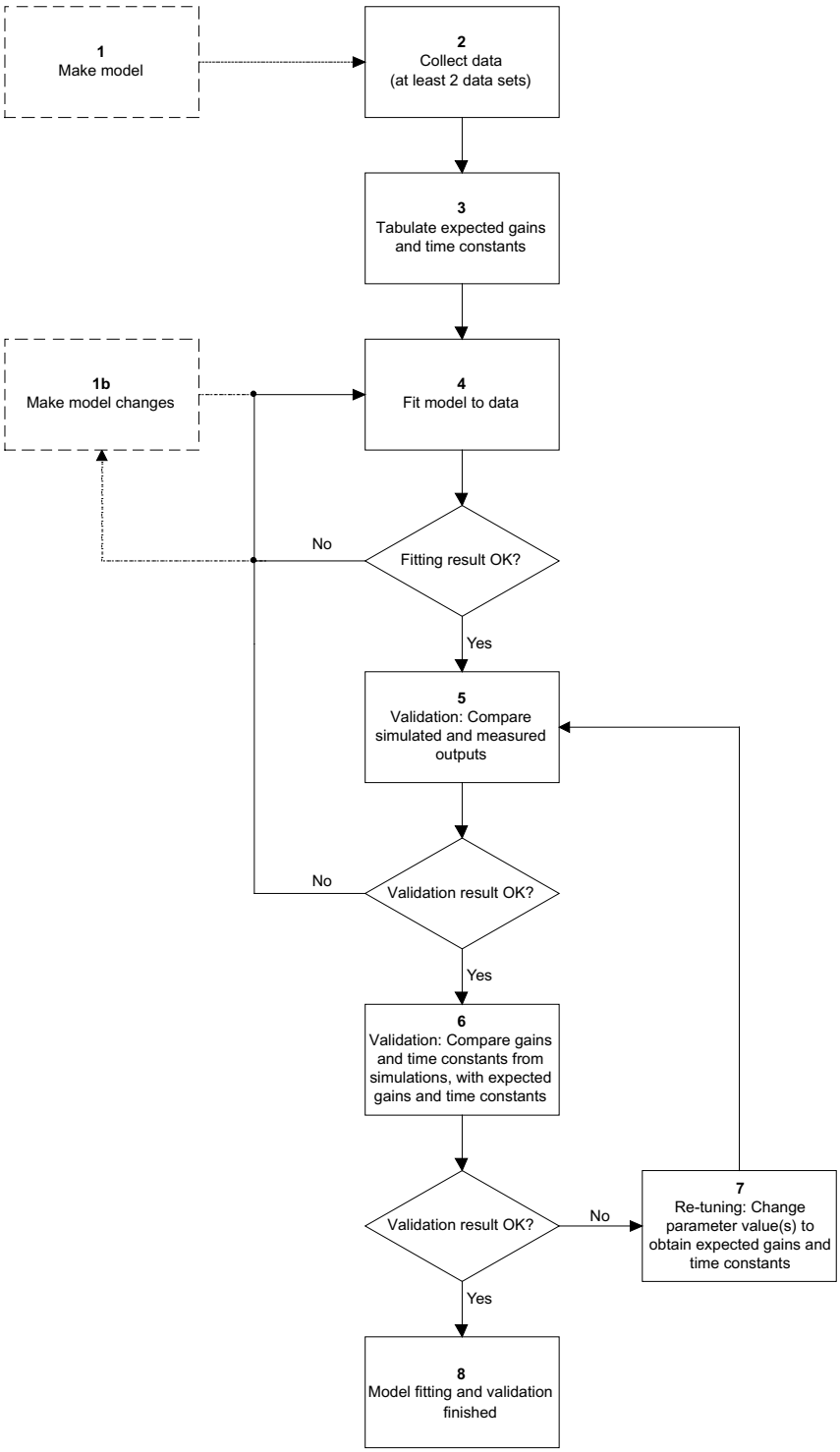


Figure A.1 A comprehensive system identification procedure. From Hauge et al. [27].

Appendix B

System identification data from open-loop experiments

In this appendix, a selection of data sets from the system identification experiments is depicted. All experiments is performed in open-loop configuration. Three approaches for identification are utilized; single channel perturbation and multiple channels perturbation using PRBS excitation, and step test perturbations channel by channel for a simple feasibility check of process behaviour, as in accordance with section 3.2.2.

B.1 Perturbations using PRBS

The subsequent data sets depicted are a selection of data sets used in the model identification process.

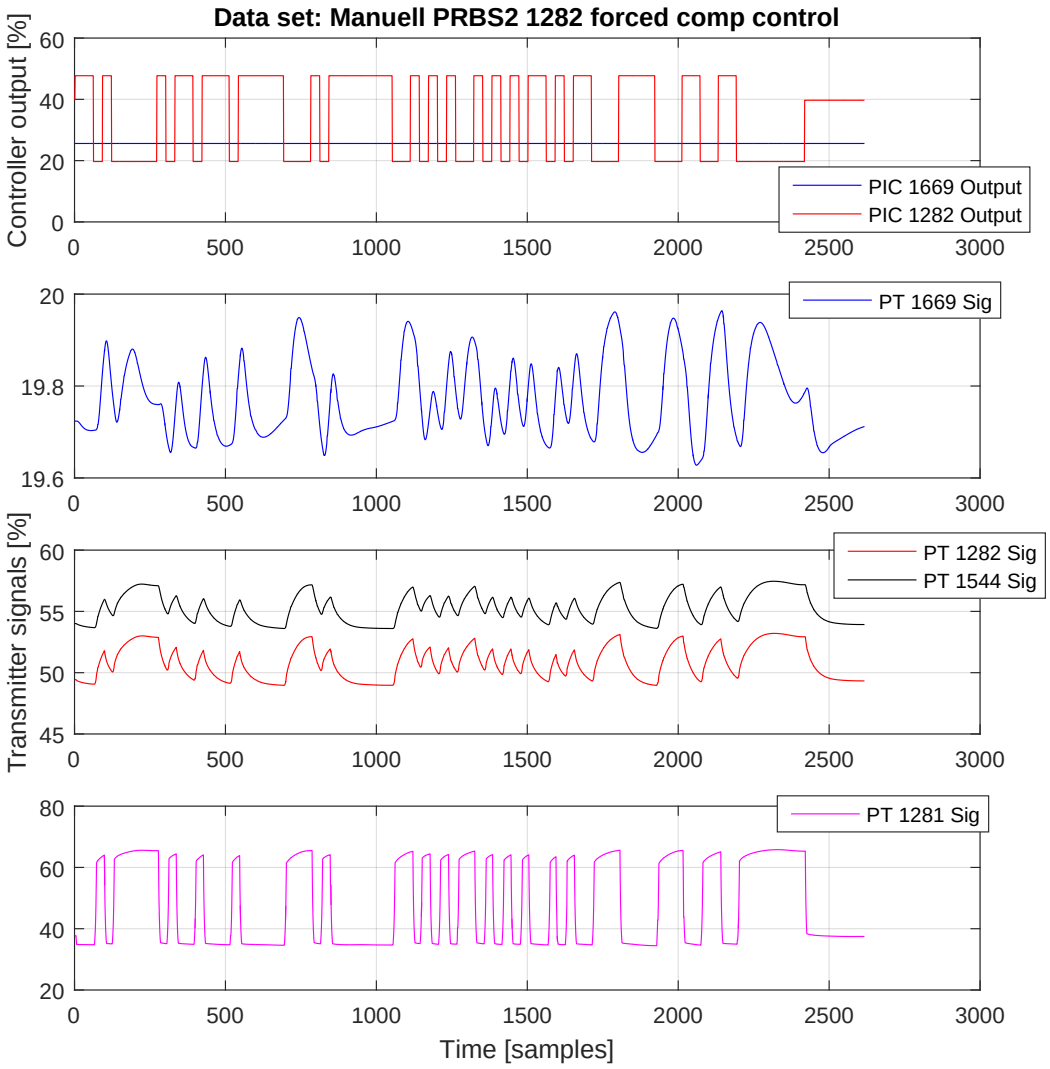


Figure B.1 PC1282 channel PRBS excitation

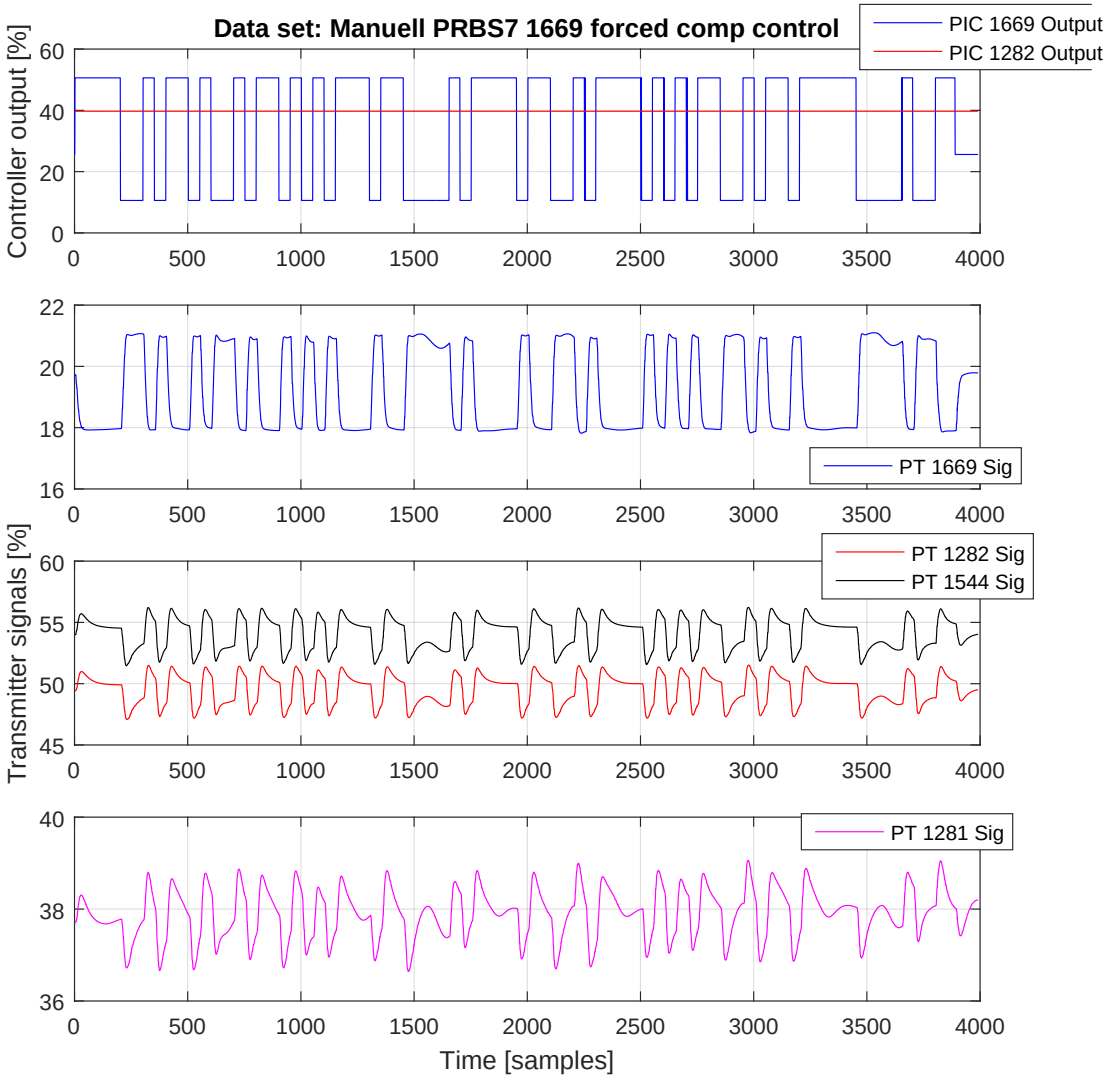


Figure B.2 PC1669 channel PRBS excitation

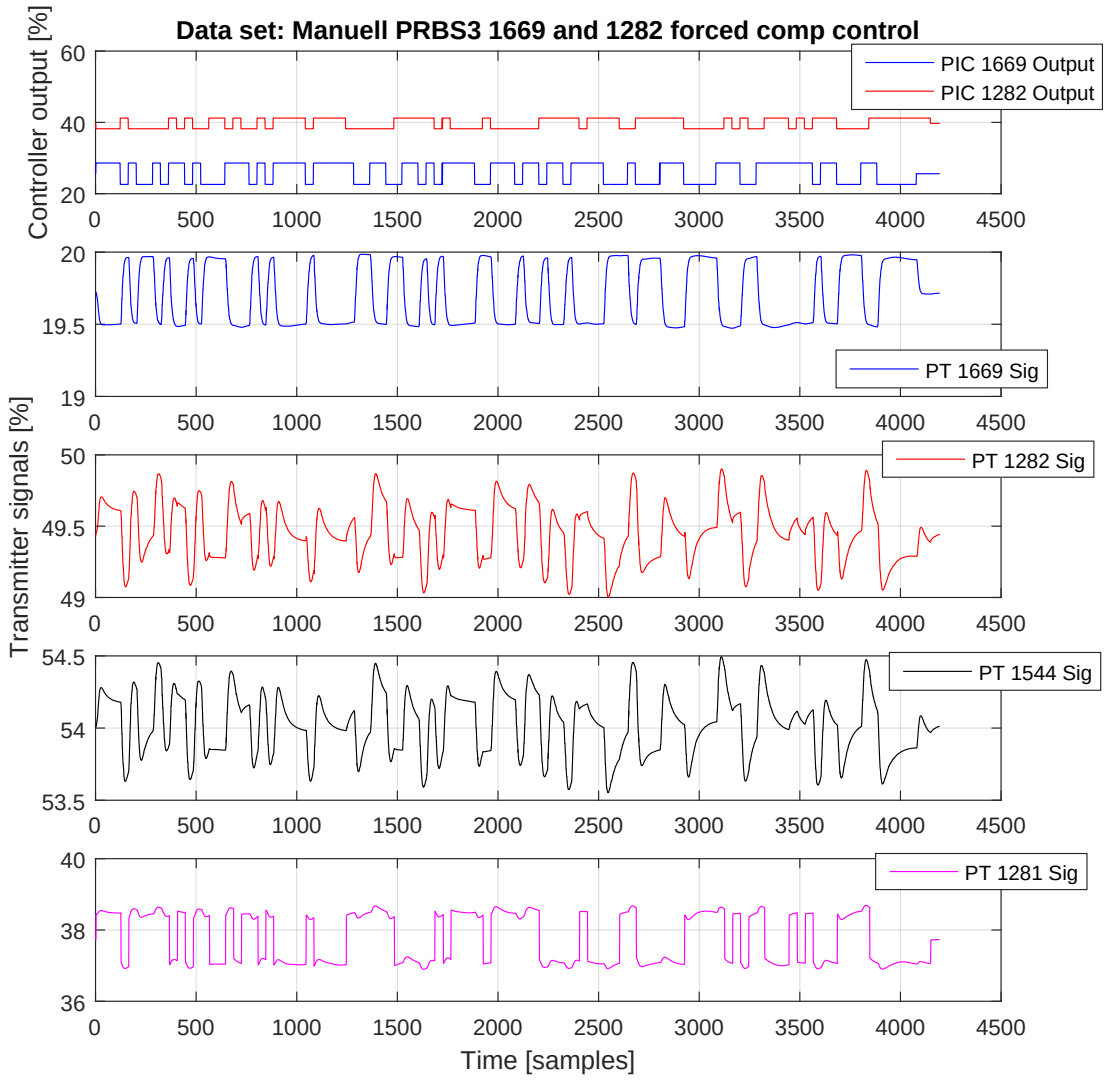


Figure B.3 PC1282 and PC1669 channels simultaneously PRBS excitation

B.2 Step test perturbations

In accordance with Tangirala [82], a feasibility check of the experiment data was performed. As demonstrated in section 4.3.1, the value of this test cannot be overstated. Steps of two different amplitudes and for both directions were imposed on the process as recommended in Tangirala [82].

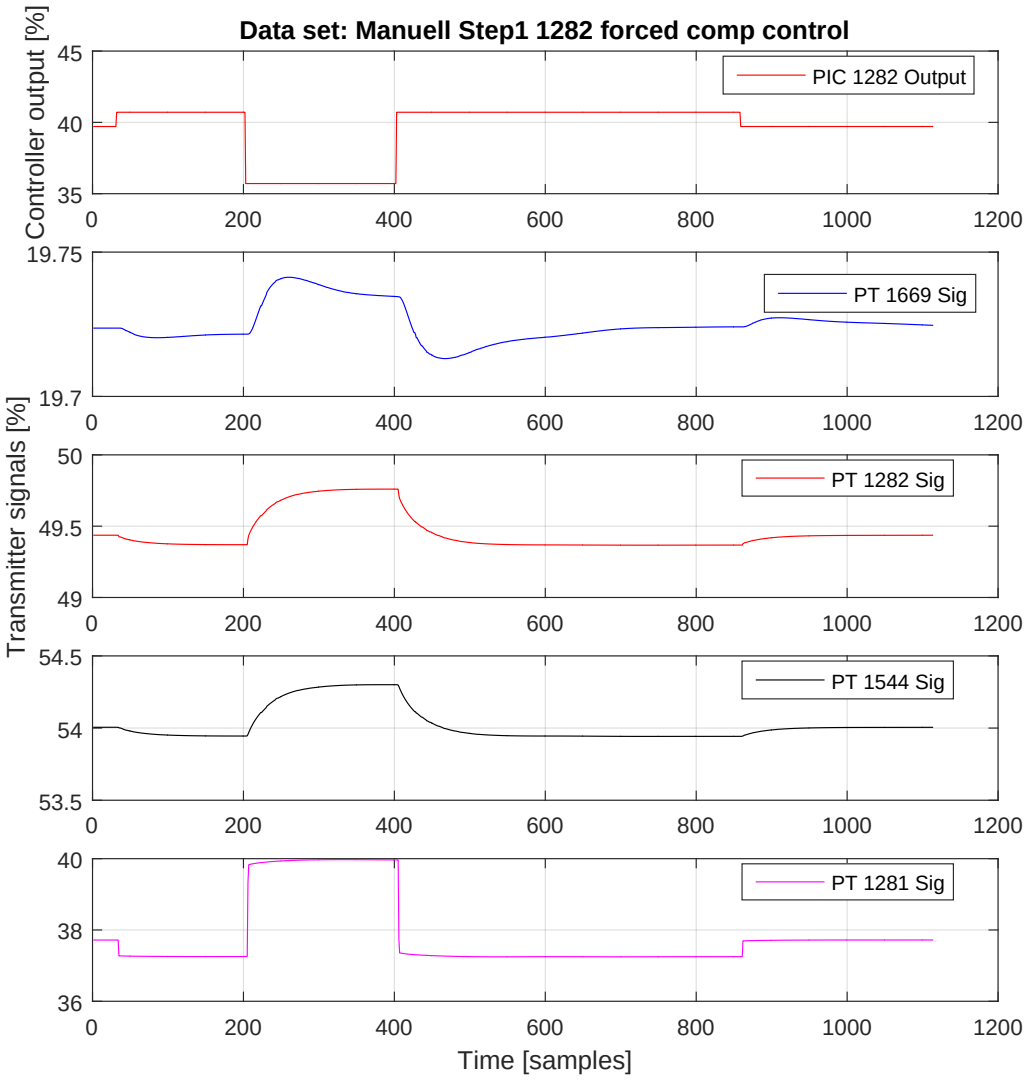


Figure B.4 Initial step test PC1282

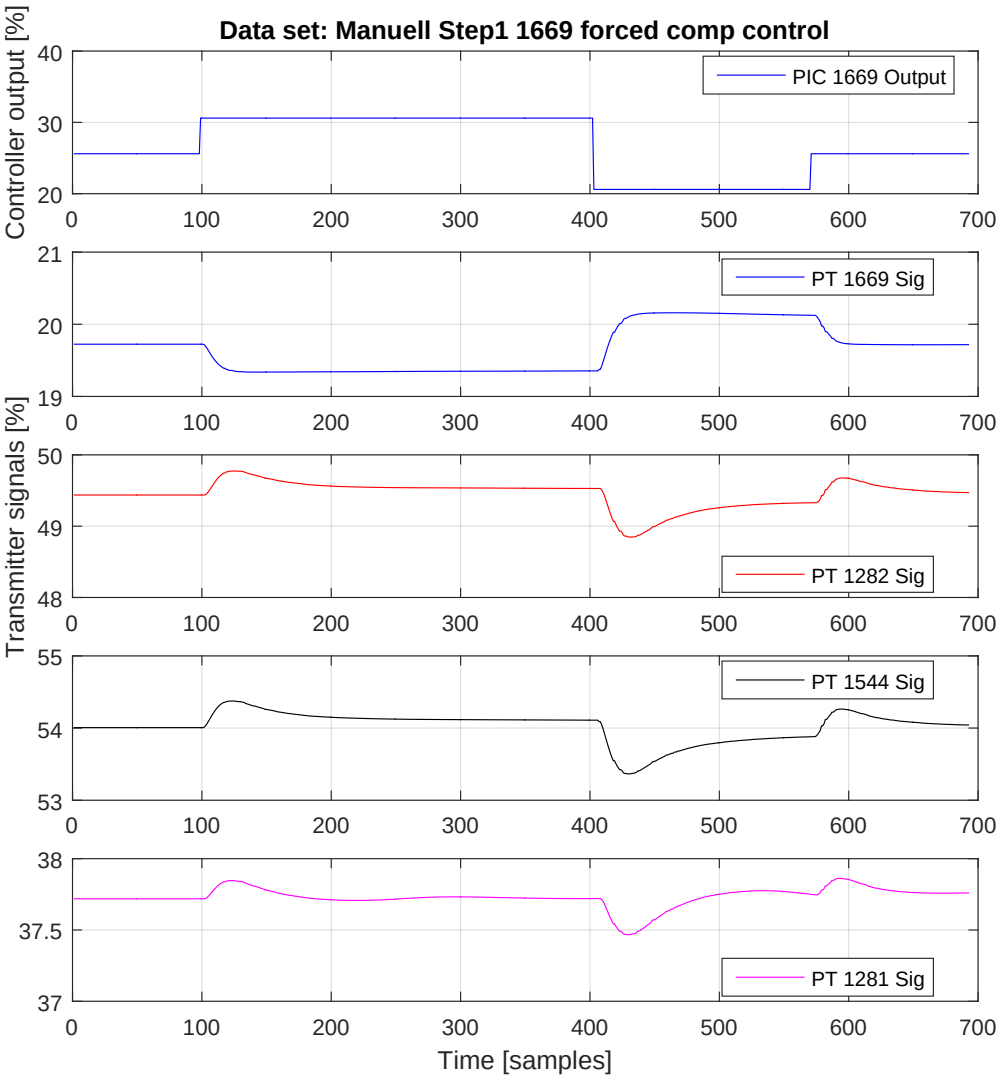


Figure B.5 Initial step test PC1669

Appendix C

Selection of mathematical preliminaries and useful examples

C.1 Persistence of Excitation

In Ioannou and Sun [36], the authors define persistence of excitation:
Definition C.1. *A piecewise continuous signal vector u: $\mathcal{R}^+ \rightarrow \mathcal{R}^n$ is persistent exciting in $\mathcal{R}^n$ with a level of excitation $\alpha_0 > 0$ if there exists constants $\alpha_1, T_0 > 0$ such that:*

$$\alpha_1 \boldsymbol{I} \geq \frac{1}{T_0} \int_t^{t+T_0} u(\tau) u^T(\tau) d\tau \geq \alpha_0 \boldsymbol{I}, \forall t \geq 0 \tag{C.1.1}$$

This definition gives a lower bound on richness of the input signal u, and in general, examples of a rich signal may be a PRBS signal, sum of various periodic signals or white noise. In addition, the matrix $u(\tau)u^T$ is singular for each τ, and the above definition requires that u makes the integral of the matrix $u(\tau)u^T$ uniformly positive definite over any time interval $[t, t + T_0]$. For further reading on the subject, the reader may consult Ioannou and Sun [36], Ljung [44], Söderström and Stoica [70], and Tangirala [82] and references therein.

Various identification procedures may require a specific input for excitation. Nonparametric procedures such as frequency analyses require input signals which must be a sinusoid, step or impulse. For transient analyses and correlation analyses, different input signals pertain. For different procedures the input requirement is in general persistently exciting. Identification of an nth-order model generally requires that the input signal is PE of order $2n$, but this will vary depending chosen procedure(Söderström and Stoica [70]).

C.2 Singular Value Decomposition(SVD) and condition number(cN)

In order to determine the order of the state-space system, inspection of the Hankel matrix's singular values are the standard approach. An approach for attaining the singular values is to utilise singular value decomposition(SVD). A definition on SVD from Nocedal and Wright [53] states:
Definition C.2. *Any complex $l \times m$ matrix A may be factorized using SVD*

$$A = U\Sigma V^H \tag{C.2.1}$$

where U and V are orthogonal matrices of dimensions $l \times l$ and $m \times m$, respectively, and may be unitary[19] *, and the $l \times m$ matrix Σ contains a diagonal matrix σ_1 of real, non-negative singular values, σ_i, $i = 1, 2, ..., n$, arranged in a descending order.*

Unitary matrices U and V form orthonormal bases for the column(output) space and the row(input) space of matrix A. The above decomposition is not unique, but for any case, the singular values σ_i are indeed unique.

The *condition number* which is defined as:

$$cN(A) = ||A||\,||A^{-1}||$$

can be defined based on SVD, where we get

$$cN(A) = \frac{\sigma_1}{\sigma_n}$$

and σ_n denotes the smallest singular value. It is then easy to se that a rank deficient matrix will have at least one singular value equal zero, thus producing a condition number $cN \to \infty$. This is an indication of an ill-conditioned plant, and one should expect some difficulties in regards of achieving an unambiguous control scheme(Skogestad and Postlethwaite [69]). Singular values are discussed in section 5.5.3, and the dynamical singular values for the best MIMO model developed in this work are depicted in figures E.20 and E.21.

Figure C.1 illustrates how severe noise influences the inspection of singular values in the Hankel matrices. A 4^{th} order stable system is perturbed using a PRBS approach in open-loop. The system is multivariable, and consists of two inputs and two outputs. The simulation is run in an approximately noise free configuration, and subsequent under severe noise influence. As observed, noise render the choice of system order inconsistent. Although this is an extreme case, it illustrates the uncertainty related to the choice of model order which seldom is unambiguous when inspecting the Hankel singular values, at least for non-trivial systems. Appendix C.3 discusses the significance of SNR in relation to model fitting further.

[19] A complex matrix M is unitary if $M^H = M^{-1}$ or for real matrices $MM^T = I$. Nevertheless, for both the absolute eigenvalues and its singular values are equal to 1.

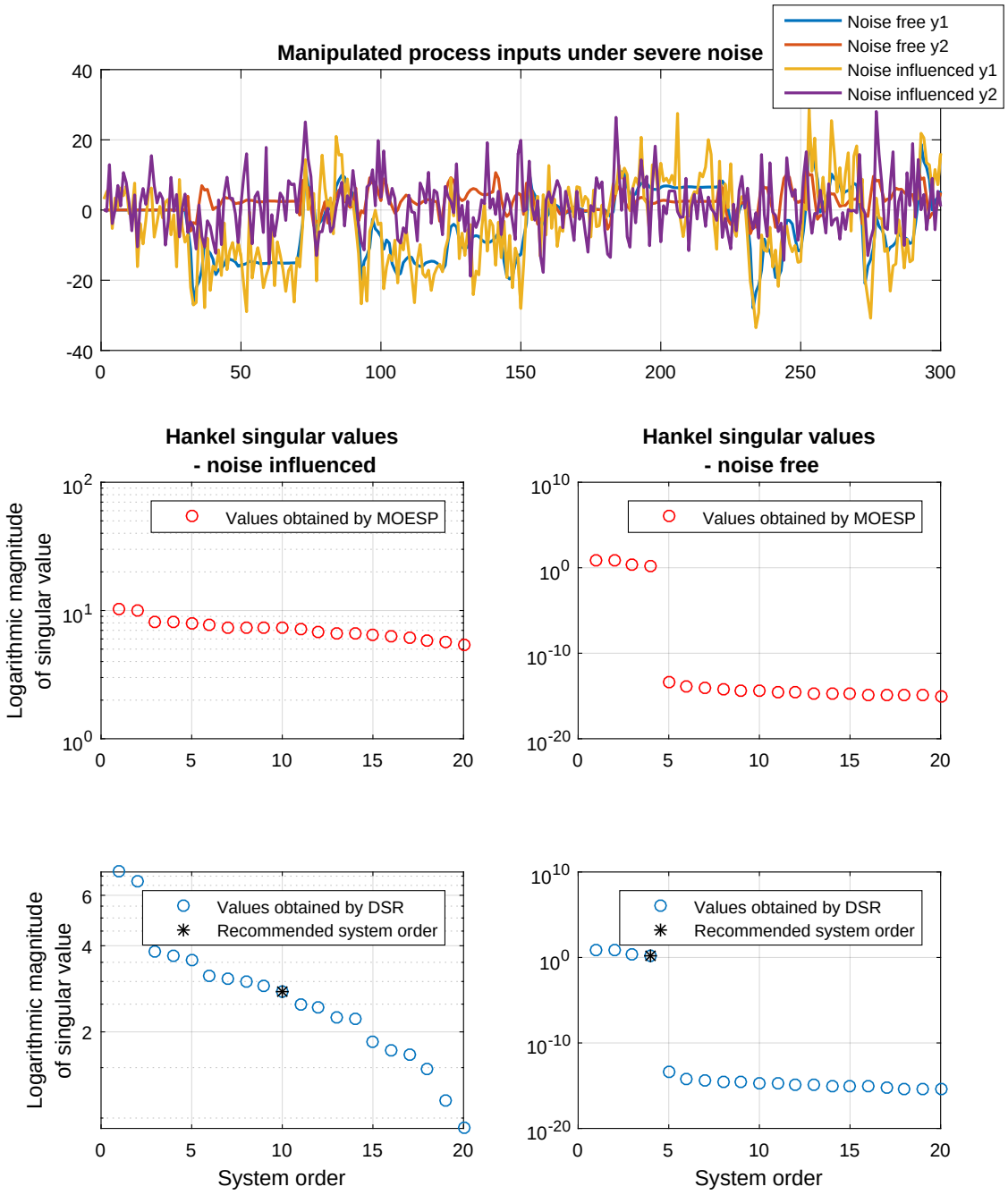


Figure C.1 4thorder system identification under severe noise, and how it influences the Hankel matrix singular value computation. As observed, noise impact the choice of system order which, compared to the unambiguous noise-free case, includes the noise. This inclusion will lead to overfitting as the noise is treated as a deterministic part.

C.3 Example to illustrate the SNR significance related to curve fitting

This example demonstrates the importance of sufficient signal-to-noise ratio. We investigate a simple nonlinear model on the form

$$y_i = b_1 u_i^2 + b_0$$

where the true values of b_1 and b_0 are 4. The model may resemble some physical relation, e.g., friction or turbulent flow of an incompressible fluid. A measurement of the output, y_m, is available

$$y_{meas} = y_i + v_i$$

where a measurement uncertainty, v_i, present. Two different measurement uncertainties are included, where the signal-to-noise ratios are (1): 10, and (2): 3, respectively. To estimate the model parameters, a least squares computation is utilized. An algorithm computes the model estimates, $\hat{b_{i1}}$ and $\hat{b_{i0}}$, and the estimation error. Estimation error is defined as $\tilde{b_{ij}} = b_{ij} - \hat{b_{ij}}$, where $i = \{1, 2\}$ and $j = \{0, 1\}$. The two returned estimates for both cases then become

$$\hat{b_{11}} = 3.99 \;\; \hat{b_{10}} = 3.76$$
$$\hat{b_{21}} = 4.06 \;\; \hat{b_{20}} = 3.7$$

As observed, the estimates remain somewhat similar despite of the SNR difference, though, the case 1 estimates are more accurate. This is not the case for the estimation error, where the error returned increases with lower SNR. This is expected, and the estimation is observed in figure C.2.

Table C.1 Estimation results and errors

	i = 1	i = 2	Error
$\tilde{b_{i1}}$	0.123	0.186	1.51
$\tilde{b_{i0}}$	0.078	0.117	1.5

Theoretically, the error value is given by the square of the fraction of the two SNR: $\sqrt{\frac{10}{3}} = 1.826$. This indicates that a lower SNR ultimately will lead to a lower reliability of the resulting estimate. Figure C.2 illustrates both cases investigated in this example.

C.4 QR and LU factorisation

As emphasized in chapter 3, matrix factorization is essential in subspace algorithms. Different algorithms use different factorizations, and this feature is vital to achieve the mathematical robustness

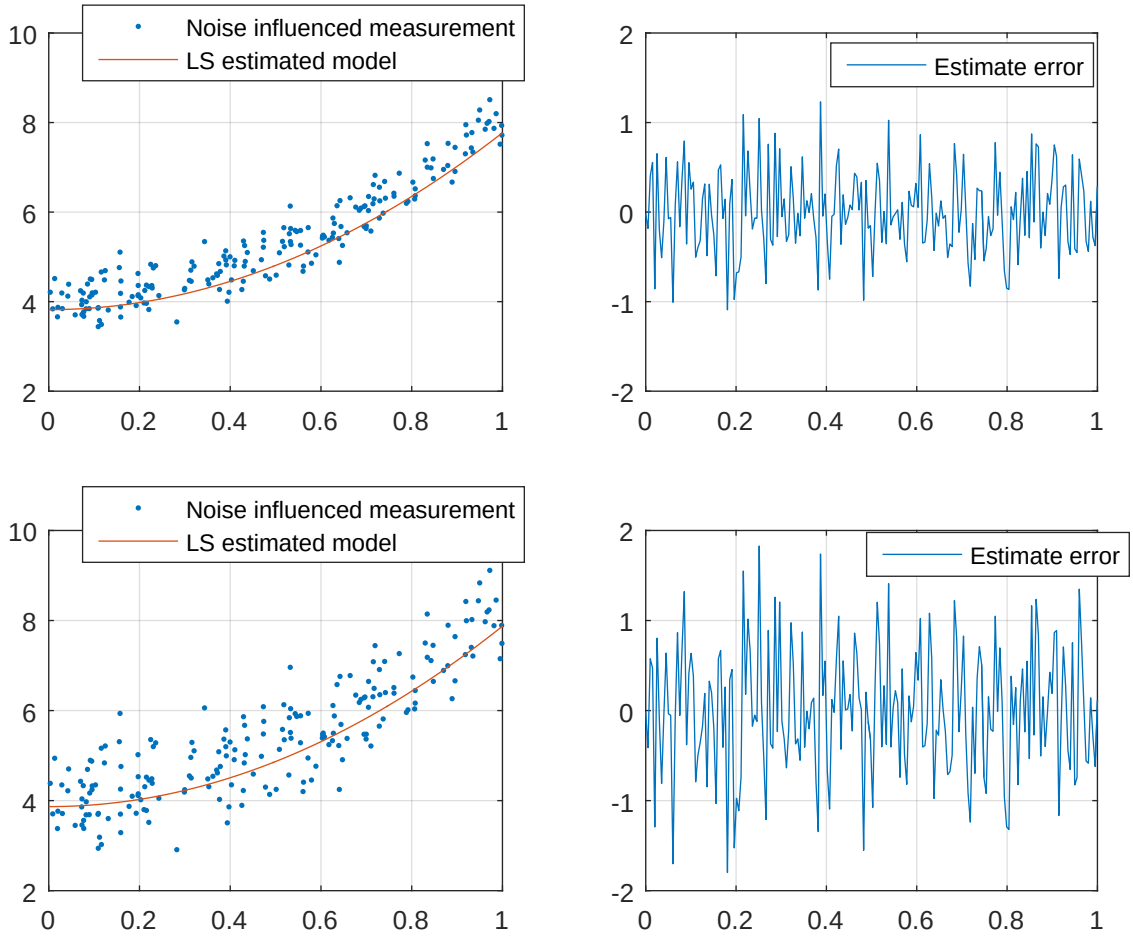


Figure C.2 Estimation plots and error plots of the curve fitting example. Case of SNR=10 are the top plots, while the case of SNR=3 are the bottom plots.

and fast computations which makes subspace identification superior to prediction error methods in many ways. The following subsections defines two factorizations utilized in various subspace algorithms.

C.4.1 QR

For a rectangular matrix $A \in \mathbb{R}^{m \times n}$, a useful factorisation is the QR factorisation(Nocedal and Wright [53])

$$AP = QR = Q \begin{bmatrix} R \\ 0 \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R \tag{C.4.1}$$

where

P is an $n \times n$ permutation matrix and thus orthogonal;

Q is $m \times m$ orthogonal;

Q_1 is the first n columns of Q while Q_2 contains the last $m - n$ columns;

R is $m \times n$ upper triangular.

C.4.2 LU

For a square matrix, the LU factorisation of a matrix $A \in \mathbb{R}^{n \times n}$ is defined as

$$PA = LU \tag{C.4.2}$$

where

P is an $n \times n$ permutation matrix(obtained by rearranging the rows of the $n \times n$ identity matrix);

L is unit lower triangular(lower triangular with diagonal elements equal to 1) and;

U is upper triangular.

C.5 Some notational details regarding subspace identification

In the following, a few definitions related to subspace identification are given. Some common variables pertain for all definitions: m is number of process inputs, l is the number of process outputs, i denotes number of block rows, and j denotes number of columns.

We have the past inputs and outputs Hankel matrix:

$$\begin{bmatrix} U_p \\ U_f \end{bmatrix} = \left[\begin{array}{ccccc} u_0 & \dots & u_1 & \dots & u_{j-1} \\ u_1 & \dots & u_2 & \dots & u_j \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{i-1} & \dots & u_i & \dots & u_{i+j-2} \\ \hline u_i & \dots & u_{i+1} & \dots & u_{i+j-1} \\ u_{i+1} & \dots & u_{i+2} & \dots & u_{i+j} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{i+h-1} & \dots & u_{i+h} & \dots & u_{i+h+j-2} \end{array}\right] \tag{C.5.1}$$

and the modified past inputs and outputs Hankel matrix:

$$\begin{bmatrix} U_p^+ \\ U_f^- \end{bmatrix} = \left[\begin{array}{ccccc} u_0 & \dots & u_1 & \dots & u_{j-1} \\ u_1 & \dots & u_2 & \dots & u_j \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{i-1} & \dots & u_i & \dots & u_{i+j-2} \\ u_i & \dots & u_{i+1} & \dots & u_{i+j-1} \\ \hline u_{i+1} & \dots & u_{i+2} & \dots & u_{i+j} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{i+h-1} & \dots & u_{i+h} & \dots & u_{i+h+j-2} \end{array}\right] \tag{C.5.2}$$

Notice that the entries of Hankel matrices can be the vectors $u_k \in \mathbb{R}^m$, hence they are called block Hankel matrices with the dimensions $U_p \in \mathbb{R}^{im\times j}, U_f \in \mathbb{R}^{hm\times j}, U_p^+ \in \mathbb{R}^{(i+1)m\times j}, U_f^- \in \mathbb{R}^{(h-1)m\times j}$. The parameters i and h allow for the different number of block rows for past U_p and U_f future. This is different to some sources, where both parameters are assumed equal. The values of the coefficients i and h are usually selected slightly larger then the upper bound of expected system order and the coefficient j is approximately equal to the number of measured data at disposal ($j \gg i, j \gg h$)(Tangirala [82]). Matrices U_p^+ and U_f^- are created from U_p and U_f by moving the first block row from U_f to the end of U_p. This variation is later used to retrieve the system matrices. For the outputs y_k and the noise e_k similar Hankel matrices Y_p, Yf and E_p, E_f can be constructed.

We have the combined past input and output data matrix, W_p:

$$W_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix} = \left[\begin{array}{ccccc} u_0 & \dots & u_q & \dots & u_{j-1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{i-1} & \dots & u_{i+q-1} & \dots & u_{i+j-2} \\ \hline y_0 & \dots & y_q & \dots & y_{j-1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{i-1} & \dots & y_{i+q-1} & \dots & y_{i+j-2} \end{array}\right] \tag{C.5.3}$$

The extended observability matrix, , Γ_i:

$$\Gamma_i \triangleq \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{bmatrix} \quad \in \mathbb{R}^{li\times n} \tag{C.5.4}$$

The lower block triangular deterministic Teoplitz matrix H_i^d is defined as:

$$H_i^d \triangleq \begin{bmatrix} D & 0 & 0 & \dots & 0 \\ CB & D & 0 & \dots & 0 \\ CAB & CB & D & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \dots & D \end{bmatrix} \in \mathbb{R}^{li \times mi} \tag{C.5.5}$$

The lower block triangular stochastic Teoplitz matrix H_i^s is defined as:

$$H_i^s \triangleq \begin{bmatrix} F & 0 & \dots & 0 \\ CE & F & \dots & 0 \\ \dots & \dots & \dots & \dots \\ CA^{i-2}E & CA^{i-3}E & \dots & F \end{bmatrix} \in \mathbb{R}^{li \times li} \tag{C.5.6}$$

C.6 Orthogonal and oblique projection

In Tangirala [82], the author defines orthogonal and oblique projection.

Definition C.3. *The orthogonal projection of the row space of a matrix onto the row space of B is given by*

$$A/B \triangleq AB^T \left(BB^T\right)^\dagger B \tag{C.6.1}$$

Definition C.4. *The oblique projection of the row space of $A \in \mathbb{R}^{r \times l}$ along the row space of $B \in \mathbb{R}^{s \times l}$ onto the row space of $C \in \mathbb{R}^{p \times l}$ is given by*

$$A/_B C \triangleq A \begin{bmatrix} C^T & B^T \end{bmatrix} \left(\begin{bmatrix} CC^T & CB^T \\ BC^T & BB^T \end{bmatrix}^\dagger \right) \tag{C.6.2}$$

where † denotes the Moore-Penrose pseudoinverse of a matrix.

C.7 Outline of the `N4SID` subspace algorithm

The subspace identification algorithm mainly used in this work is the `N4SID` algorithm developed by Overschee and Moor [55]. It utilises SVD of an oblique projection of the future output onto past data and future input spaces to estimate the states before it estimates the state-space matrices, which identify the state-space model. The algorithm is a combined deterministic and stochastic

identification procedure, which means that all data treated in the algorithm is divided into a stochastic and deterministic part, i.e., for the state sequence we have

$$X_i^d \triangleq \begin{bmatrix} x_i^d & x_{i+1}^d & \dots & x_{i+j-2}^d & x_{i+j-1}^d \end{bmatrix} \in \mathbb{R}^{n \times j}$$
$$X_i^s \triangleq \begin{bmatrix} x_i^s & x_{i+1}^s & \dots & x_{i+j-2}^s & x_{i+j-1}^s \end{bmatrix} \in \mathbb{R}^{n \times j}$$

where X_i^d and X_i^s denotes the deterministic and stochastic parts, respectively.

Assumption C.1. *Without further details, which can be found in Overschee and Moor [56], we state the initial assumptions under which the subspace identification is feasible:*

- *The deterministic input u_k is uncorrelated with the process noise w_k and the measurement noise v_k. This essentially imply that the algorithm expects open-loop data.*
- *The input is persistently exciting of order $2i$, where i equals the number of block rows used in the block Hankel matrices.*
- *The number of measurements goes to infinity $j \to \infty$.*
- *The process noise w_k and the measurement noise v_k are not identically zero.*

The number of block rows used in the block Hankel matrices, i, define an upper bound on estimation of system order and influences the computational time significantly. Overschee and Moor [56] states that the computational time is proportional with the quadrate of i. The maximum order that can be estimated is equal to $i \times l$, where l is number of outputs. Typically the parameter is set to:

$$i = 2 \times \frac{\text{Maximum order}}{\text{Number of outputs}}$$

Algorithm 1 Combined deterministic and stochastic subspace identification

1: Arrange I/O data into Hankel matrices U_p, U_f, Y_p, Y_f with their respective modifications;

$$U_p^+, U_f^-, Y_p^+, Y_f^-$$

2: Compute oblique projections:

$$\mathbb{O}_h = Y_f /_{U_f} W_p$$
$$\mathbb{O}_{h+1} = Y_f^- /_{U_f^-} W_p^+$$

3: Compute SVD of oblique projection with weights W_1 and W_2:

$$W_1 \mathbb{O}_h W_2 = USV^T$$

4: Determine the system order by inspecting the singular values of S and partition the SVD accordingly to obtain U_1 and S_1:

$$\mathbb{O}_h = USV^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

5: Determine Γ_h and $\underline{\Gamma_h}$, where:

$$\Gamma_h = W_1^{-1} U_1 S_1^{1/2}$$
$$\underline{\Gamma_h} = \begin{bmatrix} \gamma_{1,1} & \cdots & \gamma_{1,h-1} \\ \vdots & \ddots & \vdots \\ \gamma_{i,1} & \cdots & \gamma_{i,h-1} \end{bmatrix}$$

6: Determine the state sequences:

$$\hat{X}_i = \Gamma_h^{\dagger} \mathbb{O}_h$$
$$\hat{X}_{i+1} = \underline{\Gamma_h}^{\dagger} \mathbb{O}_{h+1}$$

7: Estimate the matrices $A,\ B,\ C$ and D by using linear equations:

$$\begin{bmatrix} \hat{X}_{i+1} \\ Y_i \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \hat{X}_i \\ U_i \end{bmatrix} + \epsilon$$

8: Estimate the stochastic parameters R_e and K from the covariance estimate of the residuals:

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \frac{1}{j - (n+m)} \epsilon \epsilon^T$$

where

$$R_e = \Sigma_{22}$$
$$K = \Sigma_{12} \Sigma_{22}^{-1}$$

As earlier mentioned(see sections 4.3.1 and 5.4.1.1), the weights W_1 and W_2 are different depending

on the chosen subspace algorithm. As this project considers the `N4SID` algorithm, the weights are set to:

$$W_1 = I_{li} \quad \text{and} \quad W_2 = I_j$$

C.8 Faulty pole-zero cancellation

This example demonstrates a pitfall when considering pole-zero cancellation for model reduction. We start with a simple transfer function:

$$G(s) = \frac{1}{s^2 + 2s - 3}$$

The system $G(s)$ has a RHP pole at 1, and consequently, it is open-loop unstable. Utilizing the idea behind pole-zero cancellation for model reduction, we try to cancel the RHP pole by adding a RHP zero at the same position. Thus, we get the modified system:

$$\tilde{G}(s) = (s-1)G(s) \Rightarrow \tilde{G}(s) = \frac{s-1}{s^2 + 2s - 3}$$

By closing the feedback loop and perturbing the closed-loop system by a unit step at $t = 0$ we get the following response:

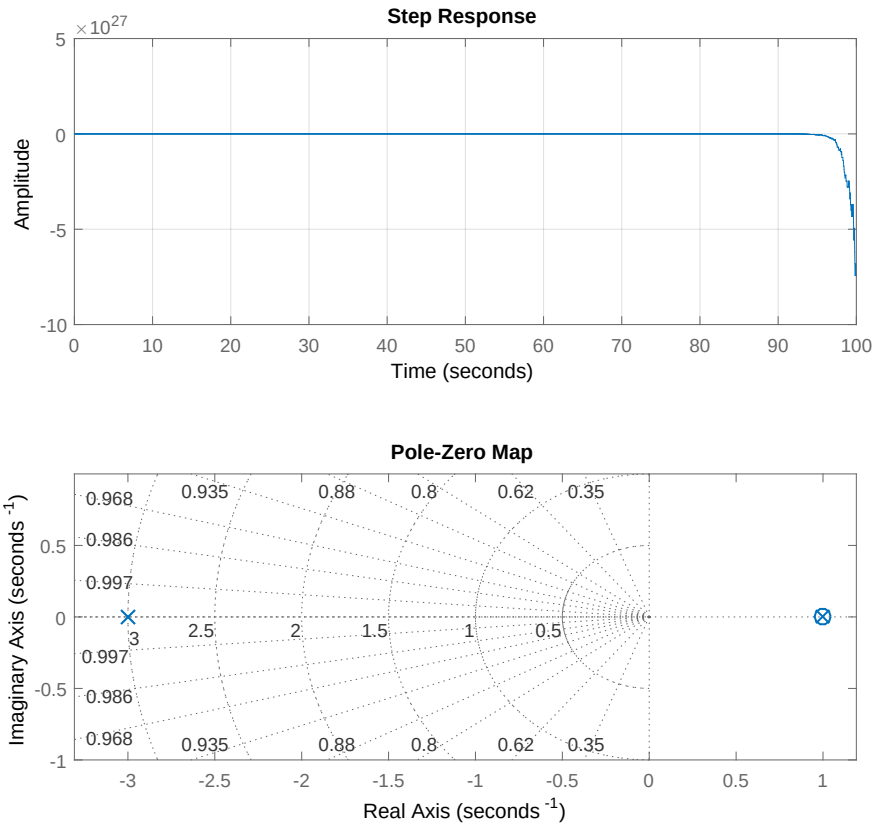


Figure C.3 Poles(X) and added RHP zero(O) over the unstable pole for cancellation.

Observing the figure emphasizes the unreliability of this method. Adding a zero will never, at least for practical cases, cancel the unstable pole entirely. Thus, a part of the root locus will be trapped in the right-half plane. This renders an unstable closed-loop response.

The conclusion is that pole-zero cancellation is theoretically feasible, but impractical. Even a perfect pole-zero cancellation will not be practically implementable because of internal stability issues and numerical round-off inaccuracies.

Appendix D

Introduction to SEPTIC model predictive control

This appendix provides an outline to the optimization problem formualtion for the MPC application developed in SEPTIC. Thus, the reader in search for general information on MPC is advised to consult Volden [88] and references therein.

For development of MPC applications, STATOIL's in-house MPC tool, SEPTIC (Statoil Estimation and Prediction Tool for Identification and Control) is used. SEPTIC was first introduced in 1997(Skofteland et al. [67]), and has since then been implemented on numerous installations(Strand and Sagli [80]). Most of these applications have been implemented using linear and experimental SISO models, as these have proved sufficiently accurate for the respective processes. Additional incentives to utilize linear models is; simpler understanding and troubleshooting, easy to build and maintain, and reduced process sensitivity to perturbations. There is however, possible to use first principles nonlinear models, either as models programmed in SEPTIC, or through an interface against an external simulator. Either way, SEPTIC treats the model as a black-box model. In addition, features from more conventional control such as anti wind-up, gain scheduling and logic operations are included in SEPTIC.

The quadratic optimization problem is formulated on the form

$$\min_{\Delta u} \quad \boldsymbol{y}_{dev}^T \boldsymbol{Q}_y \boldsymbol{y}_{dev} + \boldsymbol{u}_{dev}^T \boldsymbol{Q}_u \boldsymbol{u}_{dev} + \Delta \boldsymbol{u}^T \boldsymbol{P} \Delta \boldsymbol{u} \tag{D.0.1a}$$

subjected to

$$\underline{\boldsymbol{u}} < \boldsymbol{u} < \bar{\boldsymbol{u}}, \tag{D.0.1b}$$
$$\underline{\Delta \boldsymbol{u}} < \Delta \boldsymbol{u} < \bar{\Delta \boldsymbol{u}}, \tag{D.0.1c}$$
$$\underline{\boldsymbol{y}} < \boldsymbol{y} < \bar{\boldsymbol{y}} \tag{D.0.1d}$$
$$\boldsymbol{y} = \boldsymbol{M}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{d}, \boldsymbol{v}) \tag{D.0.1e}$$

where

$\underline{\boldsymbol{u}}$ and $\bar{\boldsymbol{u}}$ is the low and high bounds, respectively.

and the diagonal weight matrices which penalizes setpoint deviations, ideal-value deviations and input usage become

$$\boldsymbol{Q}_y \geq 0, \tag{D.0.1f}$$
$$\boldsymbol{Q}_u \geq 0, \tag{D.0.1g}$$
$$\boldsymbol{P} \geq 0, \tag{D.0.1h}$$

The bounds are imposed on both the manipulated variables, the rate of change for the manipulated variables and the controlled variables(equations D.0.1b, D.0.1c and D.0.1d, respectively). Equation D.0.1e correspond to the equality constraint, i.e., the dynamical model. The model realizes prediction of CV responses from past and future CV and MV values, in addition to past DV values. The prediction horizon and the control horizon are implicitly stated as the dimension of the vectors $\boldsymbol{y}_{dev}$ and $\boldsymbol{u}_{dev}$. Thus, in element formulation, the objective function becomes

$$\sum_{t=1}^{N} q_i(y(t+i) - y^d(t+i))^2 + r_i(u(t+i) - u^d(t+i))^2 + p_i \Delta u(t+i)^2 \tag{D.0.2}$$

where $\Delta u(t+i)$ is defined as $\Delta u_{t+1} = u_{t+1} - u_t$, and N is the prediction horizon.

SEPTIC introduces a priority level hierarchy for optimizing targets, which leads to solving several steady-state quadratic optimization problems. The reason behind this is the fact that a typical application does not have the required number degrees of freedom to satisfy all control specifications. The variables in the priority level hierarchy are given their priority depending on the severity of a constraint violation. The priority level hierarchy in decreasing order is given in table D.1.

By the table, the highest priority fall to the MV's rate of change. This priority is always respected. At each of the following stages, a quadratic steady-state optimization problem is solved with respect to the remaining priorities. The solution at each stage respects the results from the previous stages and yields the smallest possible deviation from the original specifications. Should there be more CV's to control than there are MV's, the number of degrees of freedom will become negative and the MPC controller will successively drop the least important priority levels until a solution can be

Table D.1 SEPTIC priority rating

SEPTIC priority rating	
Priority	**Constraint**
1	MV rate of change
2	MV upper and lower bounds
3	CV hard constraints
4	CV setpoint, CV upper and lower constraints, and MV ideal values with priority level 1
5	CV setpoint, CV upper and lower constraints, and MV ideal values with priority level n
6	CV setpoint, CV upper and lower constraints, and MV ideal values with priority level99

obtained. Omitting such a hierarchy when considering hard CV constraints may lead to infeasible problems.

The ability for optimal constraint handling is an important feature of the MPC controller. In SEPTIC, the constraints can be defined as either soft or hard. Hard constraints pertain to MV's and physical relations, e.g valve openings or pump capacities. Hard constraints are not allowed to be broken, while soft constraints are allowed to be broken occasionally, but only if necessary. By softening the constraints, one may use slack variables which may look like this:

$$\underline{y} - \epsilon < y < \bar{y} + \epsilon, \quad \epsilon \in \mathbb{R}^{n_y} \geq 0$$

To minimize the constraint violation, the slack variables are included with a tuning parameter in the objective function which returns a higher cost upon constraint violation. In SEPTIC, hard constraints on CV's are principally omitted. Soft CV constraints may render a difficult problem in terms of realisation of all requirements for the different combinations.

Appendix E

Additional plots

This appendix depicts some additional plots. Although not included in the report, they may be of some interest for the reader.

E.1 Model uncertainty plots

In addition to figure 5.6, additional plots of three candidate MIMO data models compared to the SISO data model were compared in terms of model uncertainty. In these plots, the confidence interval is 5σ. Perturbation is a unit step.

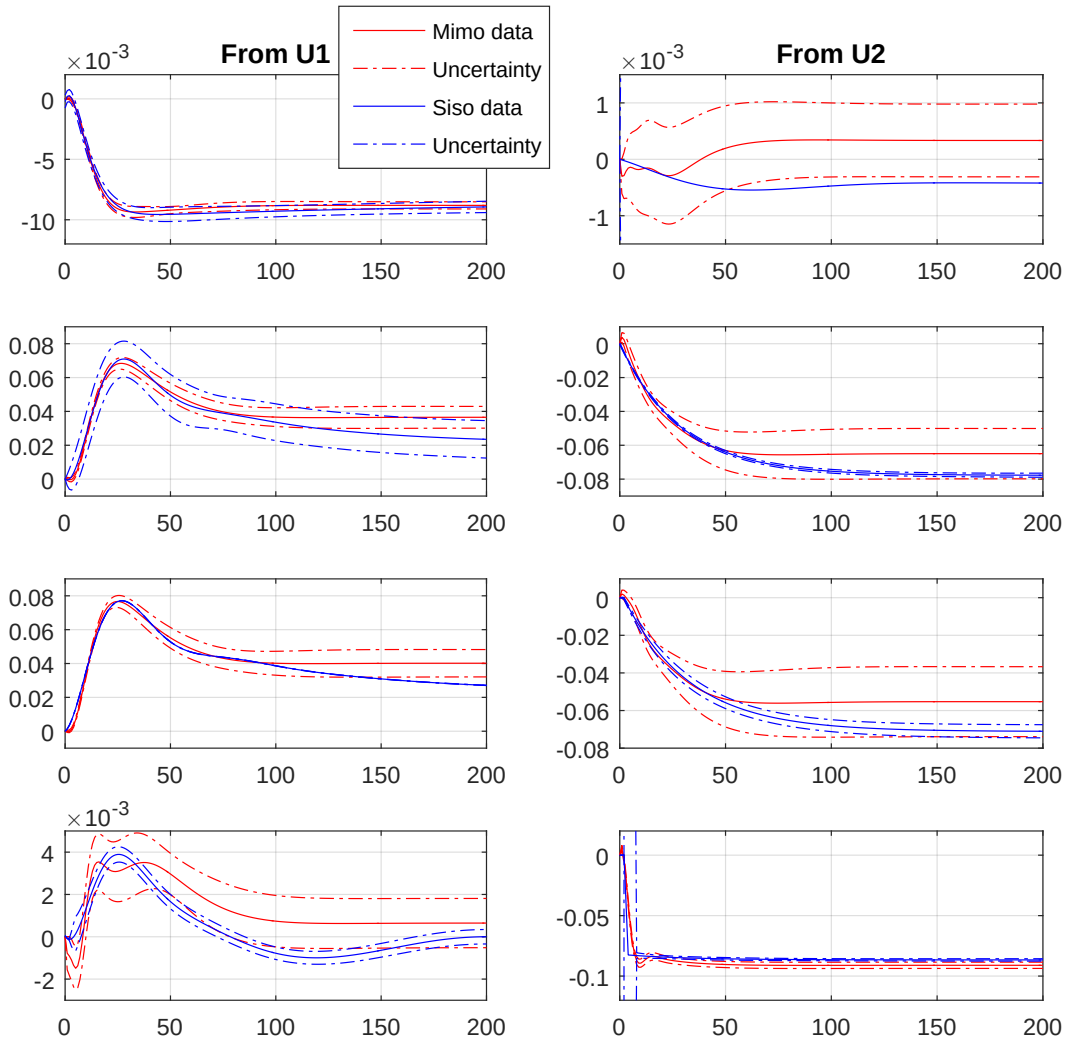


Figure E.1 MIMO data model 1 and SISO data model uncertainties

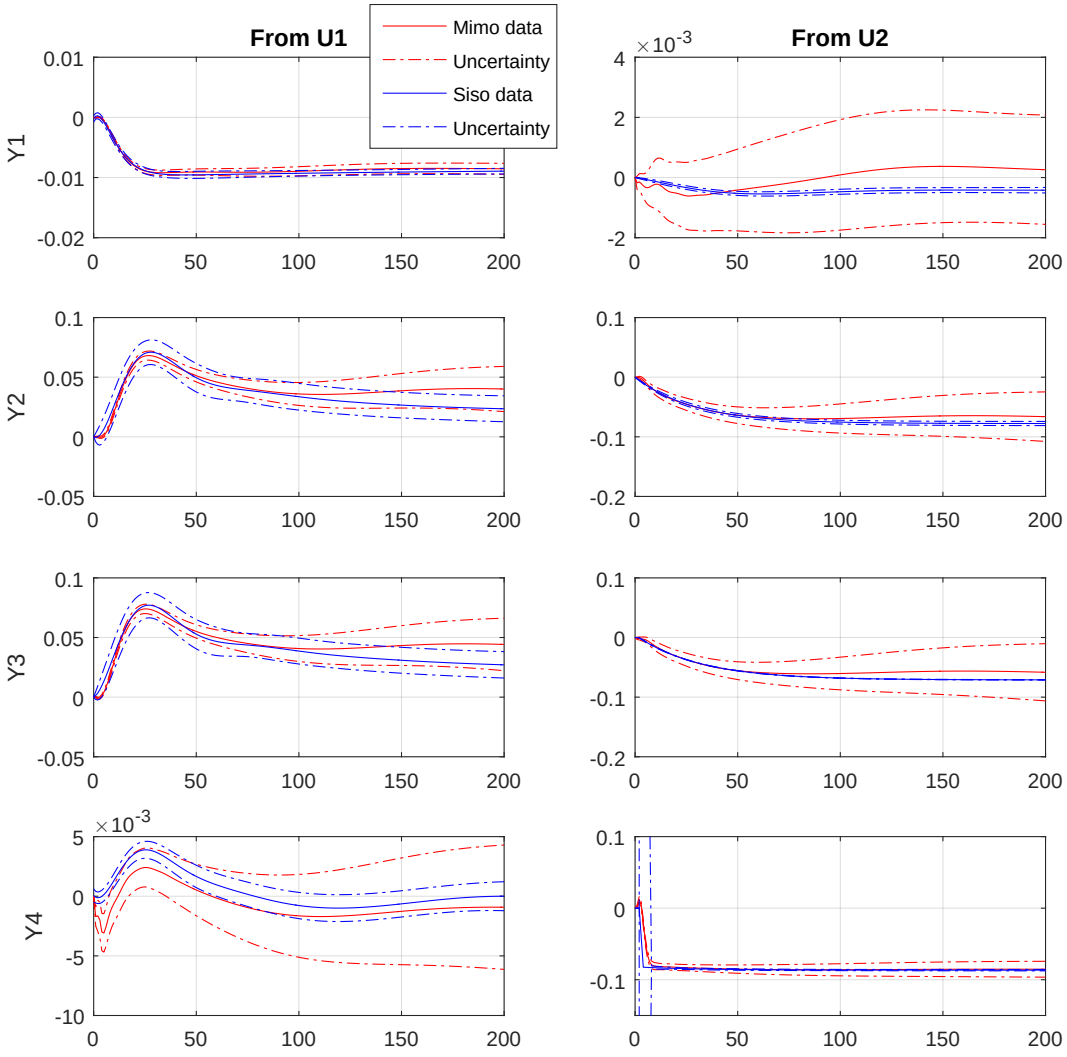


Figure E.2 MIMO data model 2 and SISO data model uncertainties

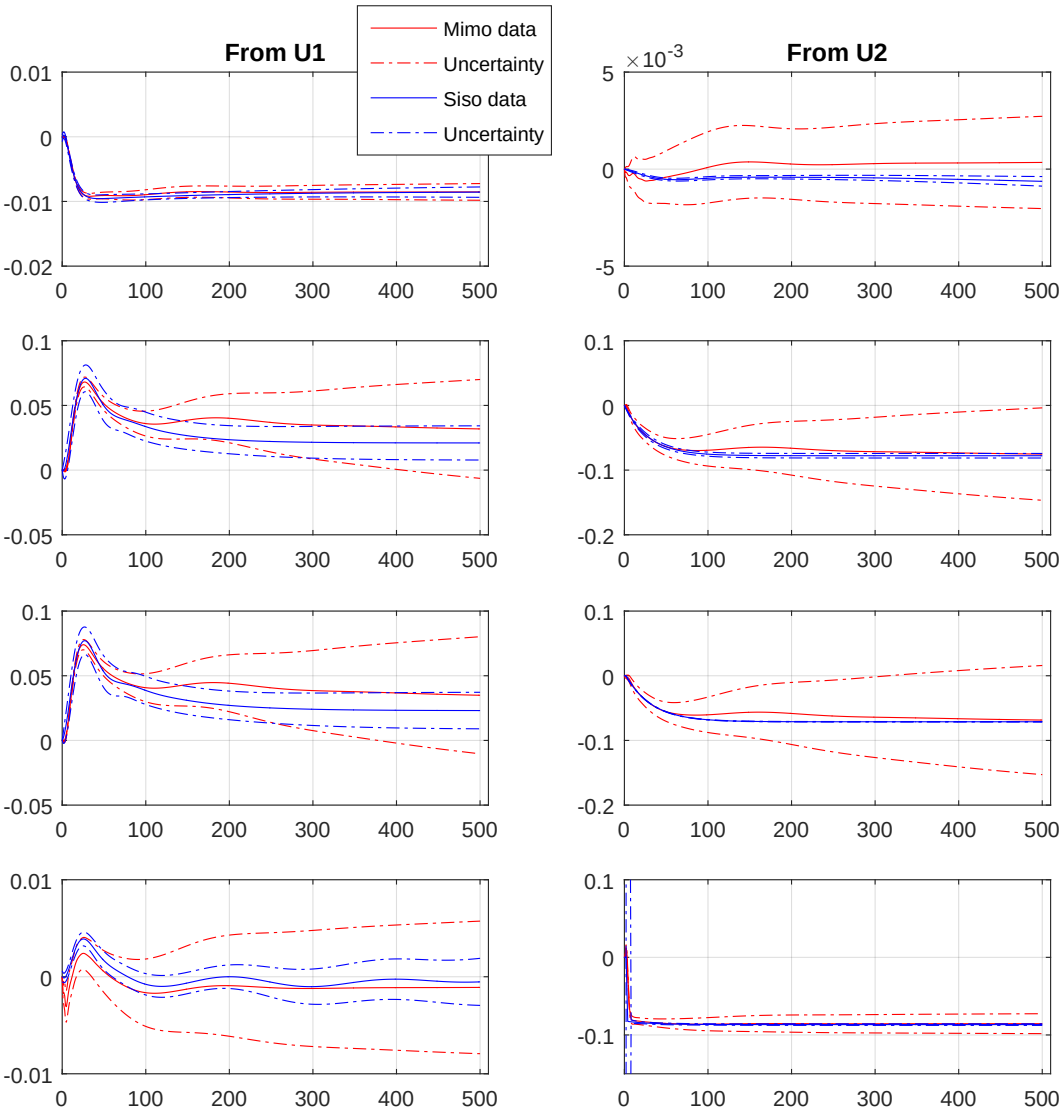


Figure E.3 MIMO data model 3 and SISO data model uncertainties

E.2 Inspection of Hankel singular values

Figure E.4 depicts an additional inspection of the Hankel singular values for an identification data set. The data set is similar to the one used in figure 5.7., but not identical.

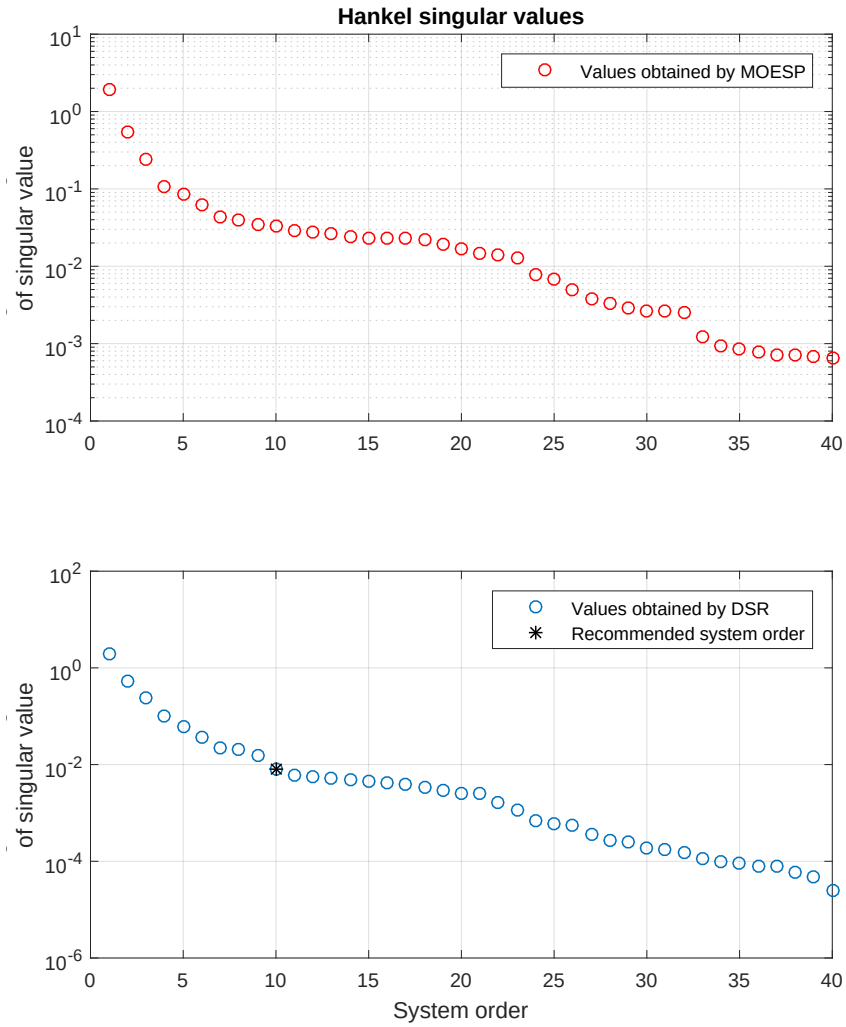


Figure E.4 Hankel singular value inspection of data and recommendation of system order

Comparing the figure E.4 and figure 5.7 reveals a weakness of the statistical criterion used for system order estimation. In section 5.4.1.1 the estimated system order is 6, while observed from figure E.4, the value has increased to 10. This emphasizes that additional measures to obtain model order may be advantageous. An iterative approach used in this work is described in section 4.3.1.

E.3 Additional plots on the selector functionality and velocity form controller

To further describe the selector behaviour and the velocity form controller PIC1281, some additional plots are included. All figures depict a closed-loop sequence of setpoint changes and corresponding responses. Further, the top subplot for both figures E.5 and E.6 depicts the setpoint changing

sequence. In addition, the split range configuration of PIC1282 mentioned in section 4.3.2 is demonstrated.

In the middle subplot of the first figure we observe that the selector compares the PIC1282 and PIC1281 controller outputs, and consistently selects the lowest value. The MV in this case is the valve PV1282A. The bottom subplot depicts the PT1281 CV and its setpoint.

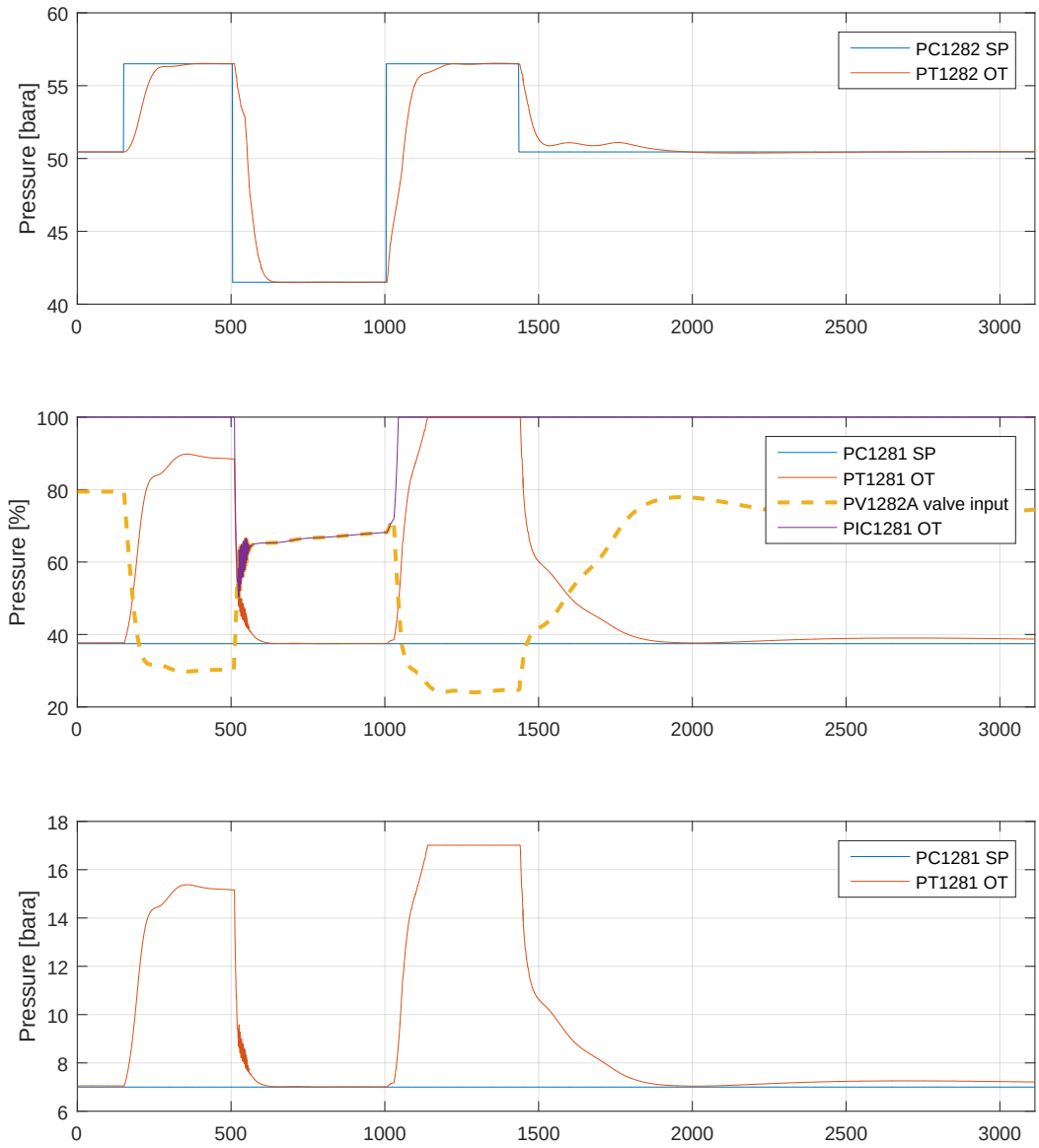


Figure E.5 Closed-loop response to setpoint change depicting input shift to PV1282A from a controller output selector

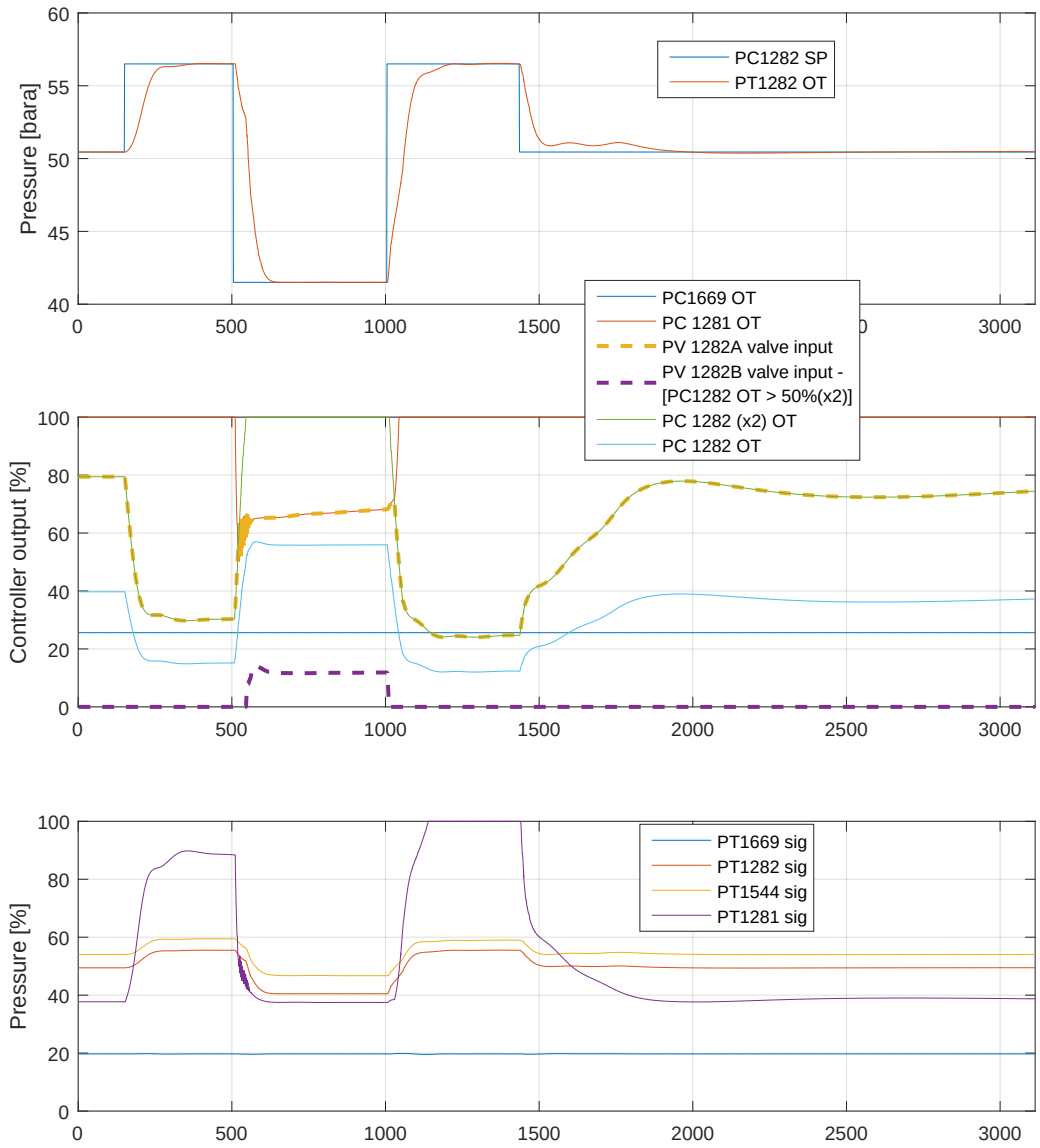
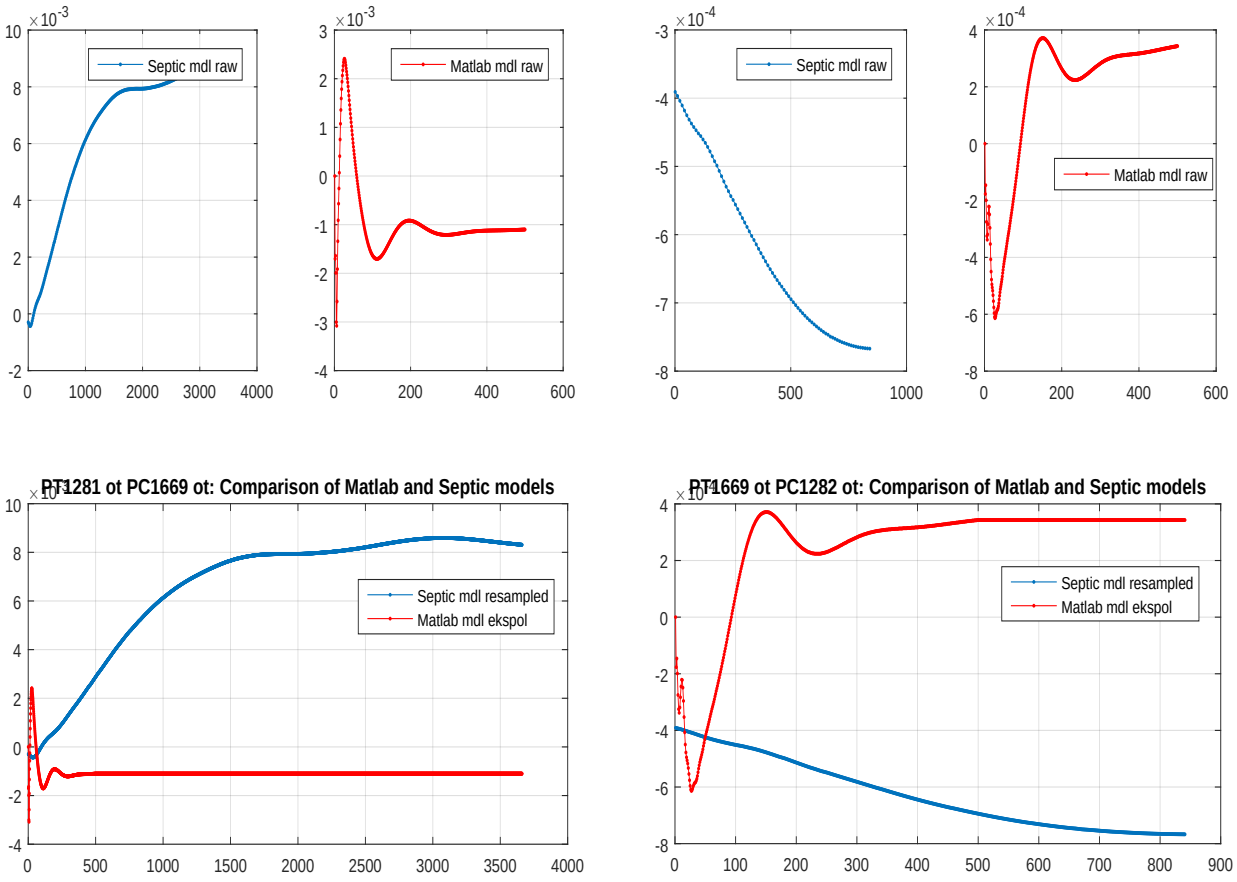


Figure E.6 Closed-loop response to setpoint change. Input shift to PV1282A from selector and input to PV1282B bypass from split-range controller PIC1282.

For the middle subplot in figure E.6, some additional variables are added to include a second MV, PV1282B. This is a bypass valve for PV1282A. PIC1282 is a split range controller where the control output $u : 0 \rightarrow 50\%$ applies $0 \rightarrow 100\%$ for PV1282A, whereas $u : 50 \rightarrow 100\%$ applies for $0 \rightarrow 100\%$ PV1282B. Thus, PV1282B serves as a bypass only when PV1282A is fully open at its limit. To render this configuration feasible, the PIC1282 output is doubled prior to utilization. PV1282B is solely controlled by PIC1282. The bottom plot depicts the four normalized CV's response for the setpoint sequence.

E.4 Faulty SEPTIC modelling

Figure E.7 depict a faulty case of modelling in SEPTIC. Two SEPTIC developed SISO models are compared with a multivariable MATLAB model. The reason behind the faulty modelling is faulty data for the SEPTC case. For modelling purposes the SEPTIC application directly relies on data retrieved from the simulator. This implies the faulty data were provided by the simulator which later was confirmed(see section 6.1.1).



(a) SEPTIC CV 1 SISO modelling error compared

(b) SEPTIC CV 2 SISO modelling error compared

Figure E.7 Two faulty SEPTIC SISO models comparisons.

E.5 Additional error and fitness plots

Figures E.8 and E.9 depicts models developed from increasing order and fixed subspace algorithm horizon. The algorithm horizon utilized in this case is not based on Akiake's criterion directly, but chosen in the neighbourhood for the sake of comparison. Moreover, the only part of the horizon altered is the number of step-ahead predictors. Thus, the horizon is set to [15 24 24] as opposed to

the AIC chosen horizon which was [27 24 24].

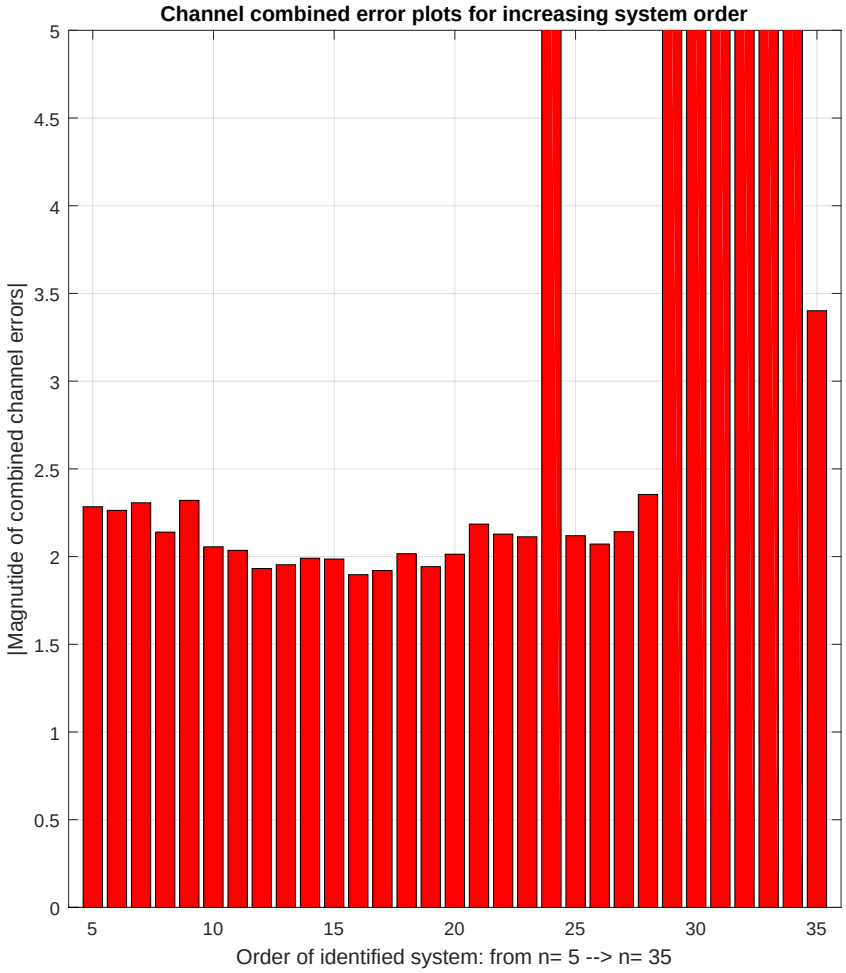


Figure E.8 Combined excitation and fixed horizon comparison for specific algorithm horizon. A error value of 5 indicates a diverging open-loop system response. Thus, the system is open-loop unstable and produces errors $\rightarrow \infty$.

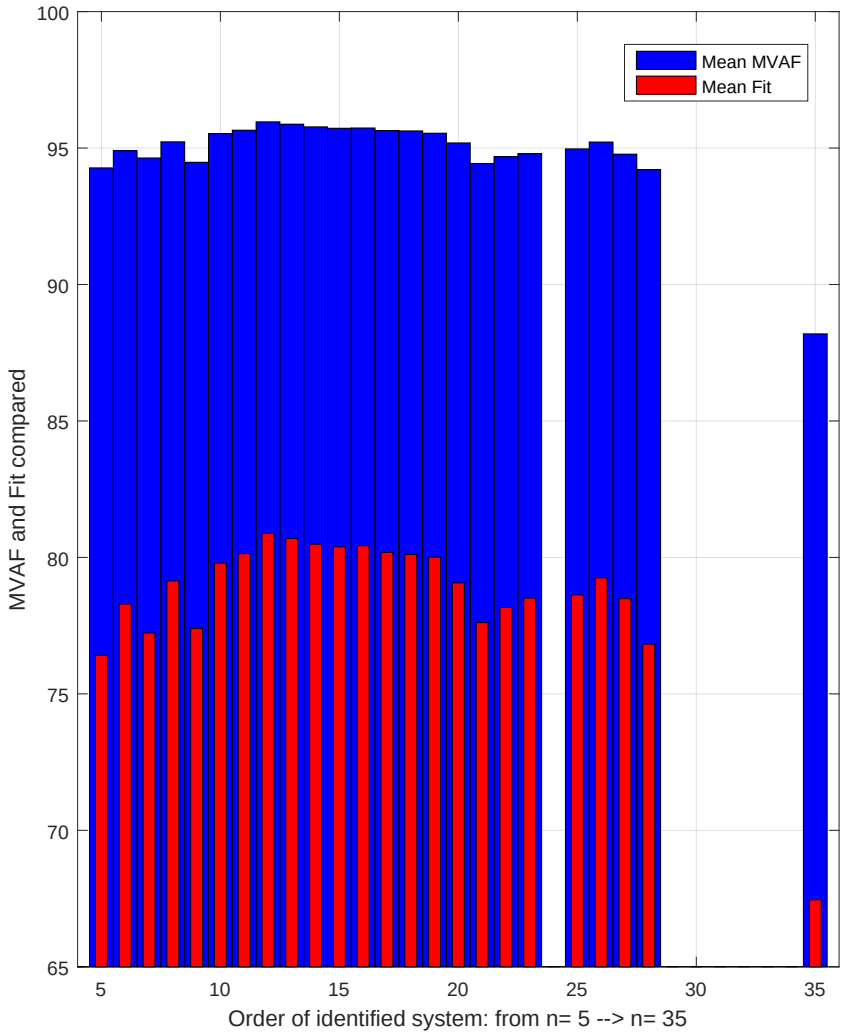
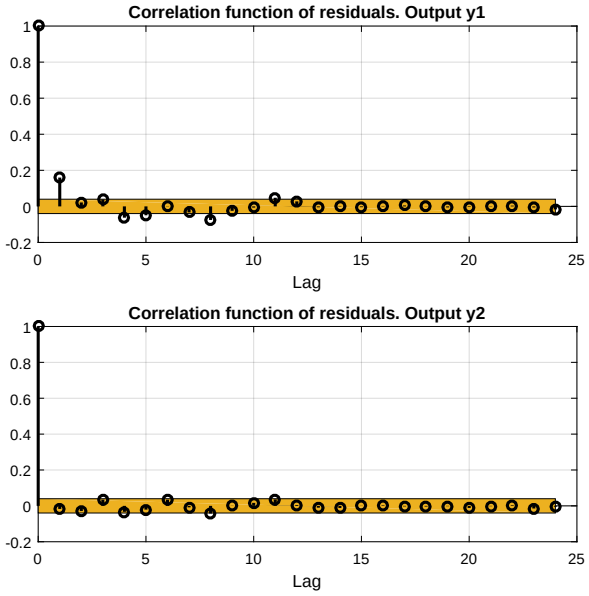


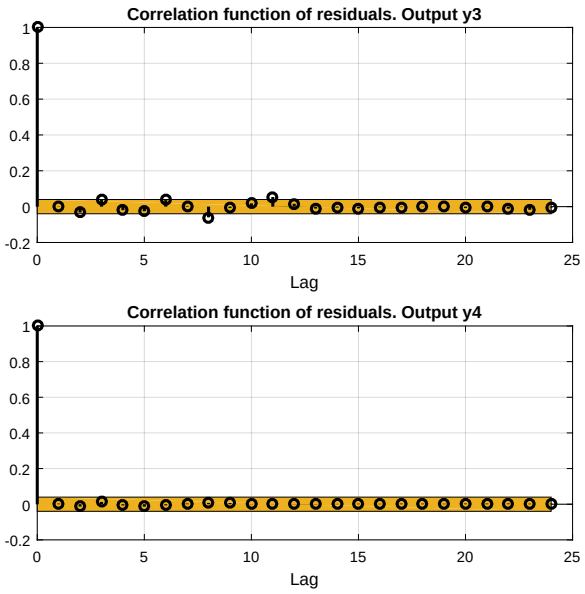
Figure E.9 Combined excitation and fixed horizon comparison for specific algorithm horizon. As in figure E.8, the open-loop unstable systems have low or negative NRMSE and MVAF values. Hence, they are removed from the plot.

E.6 Comparing 3 models based on residuals

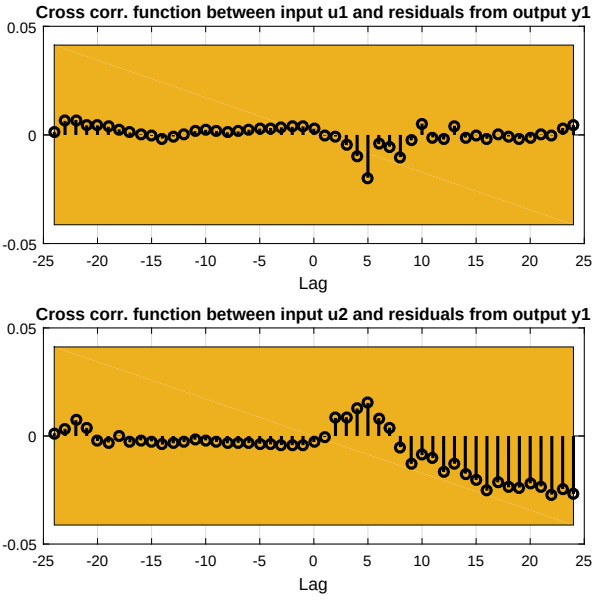
This section depicts three individual residual analyses for the three candidate models depicted in section 5.4.2.2. To be consistent, model 1 in 5.4.2.2 correspond to model 1 in this section, and so on. Although, a small difference pertain: in the top plots, the autocorrelation function is depicted for lags $[0 \rightarrow 25]$, and not $[-25 \rightarrow 25]$ as in section 5.4.2.2. The cross correlation plots is depicted for lags $[-25 \rightarrow 25]$. These figures may provide a simpler understanding of the differences depicted in figures 5.10 and 5.11.



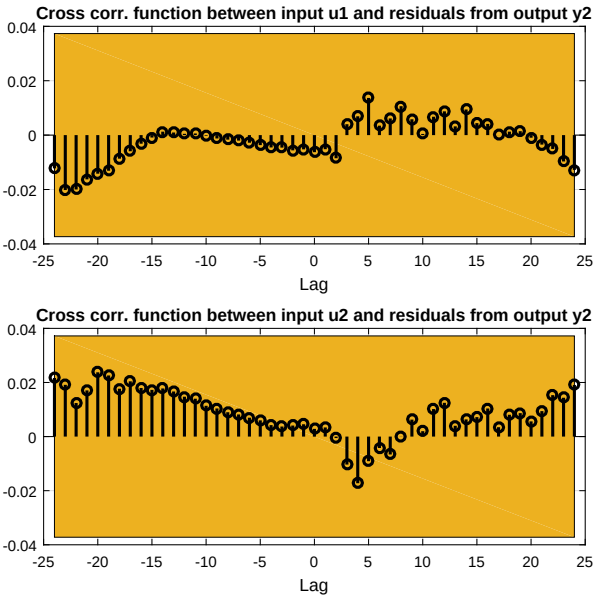
(a) Correlation between residuals and lagged inputs for Y_1 and Y_2 for model 1.



(b) Correlation between residuals and lagged inputs for Y_3 and Y_4 for model 1.

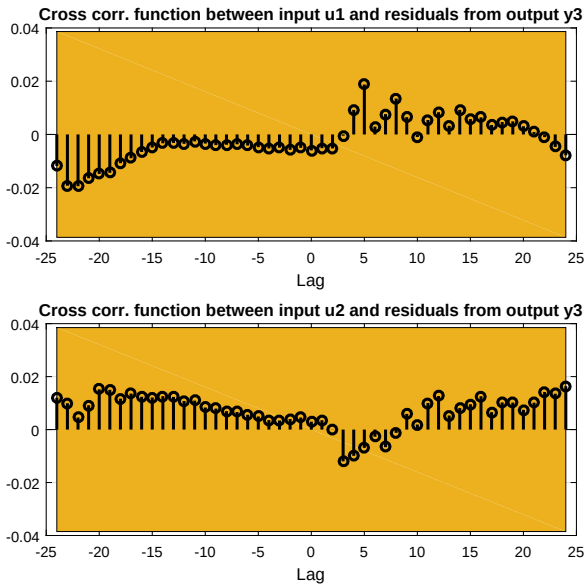


(c) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_1 for model 1.

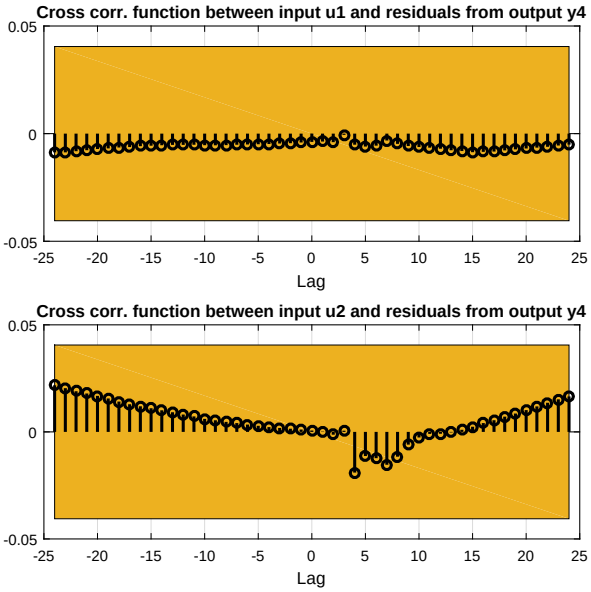


(d) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_2 for model 1.

Figure E.10 Correlation of residuals and cross correlation between lagged inputs and residuals of outputs Y_1 and Y_2 for model 1.

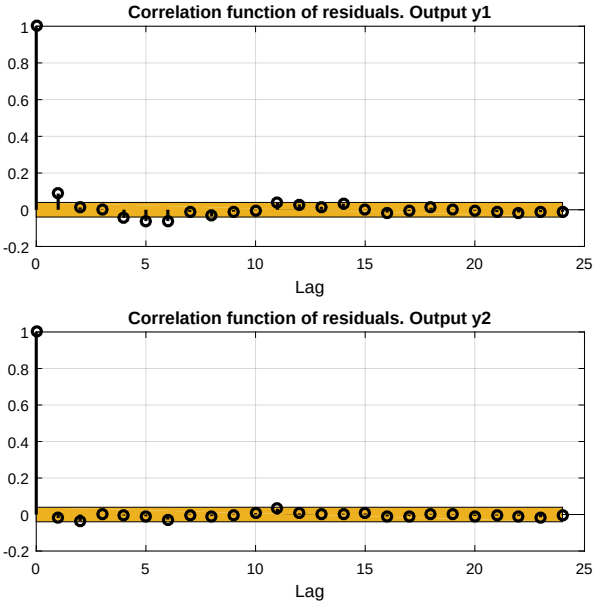


(a) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_3 for model 1.

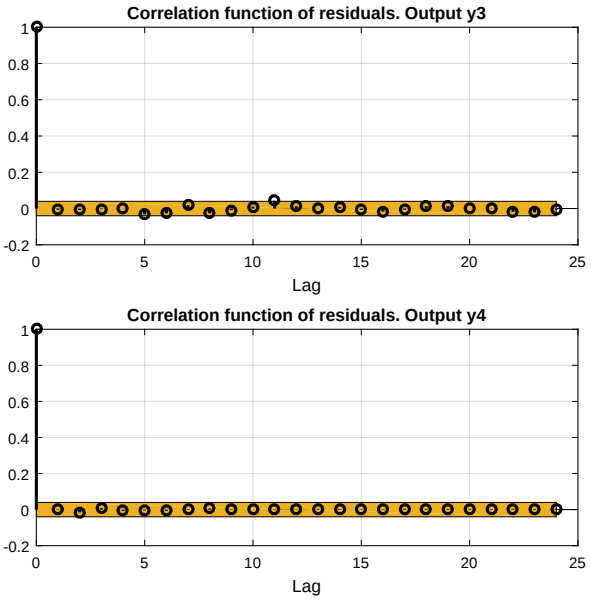


(b) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_4 for model 1.

Figure E.11 Cross correlation between inputs and residuals of outputs Y_3 and Y_4 for model 1.

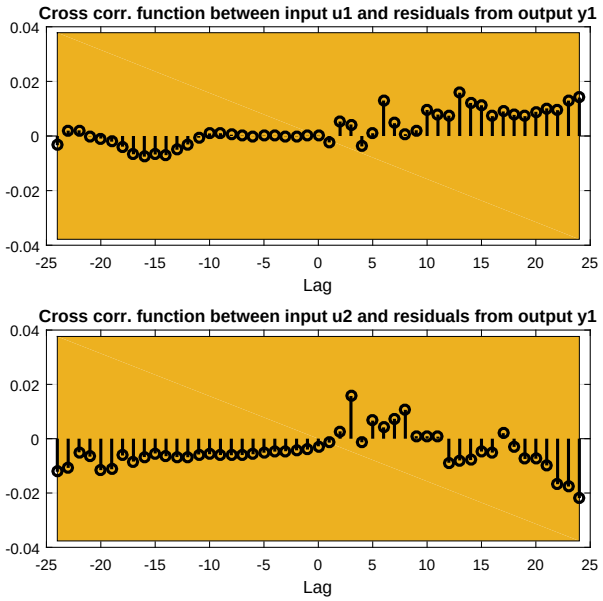


(a) Correlation between residuals and lagged inputs for Y_1 and Y_2 for model 2.

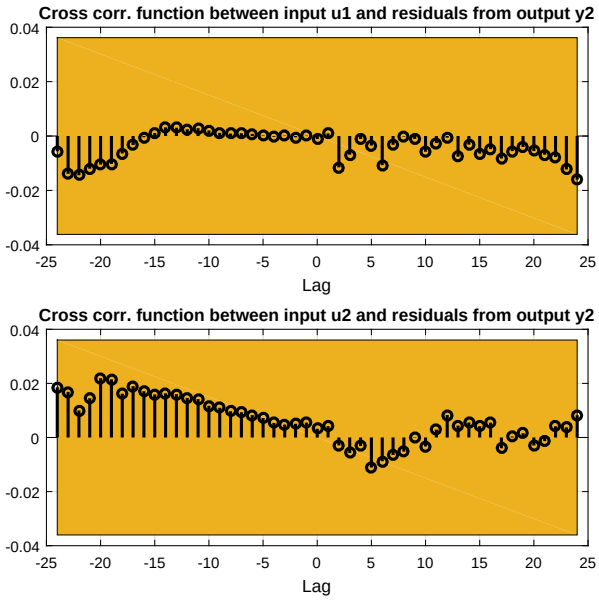


(b) Correlation between residuals and lagged inputs for Y_3 and Y_4 for model 2.

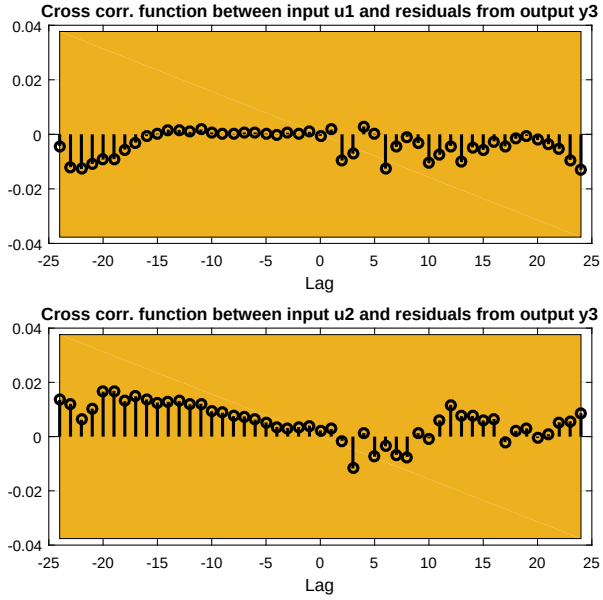
Figure E.12 Correlation between lagged inputs and residuals of outputs Y_1 and Y_2 for model 2.



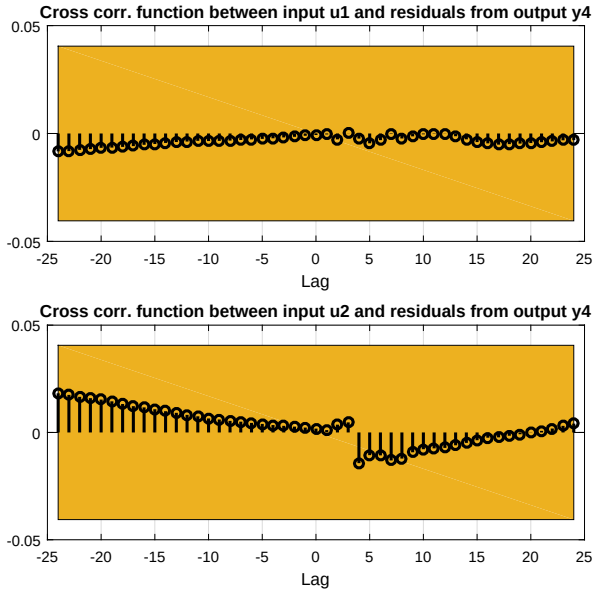
(a) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_1 for model 2.



(b) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_2 for model 2.

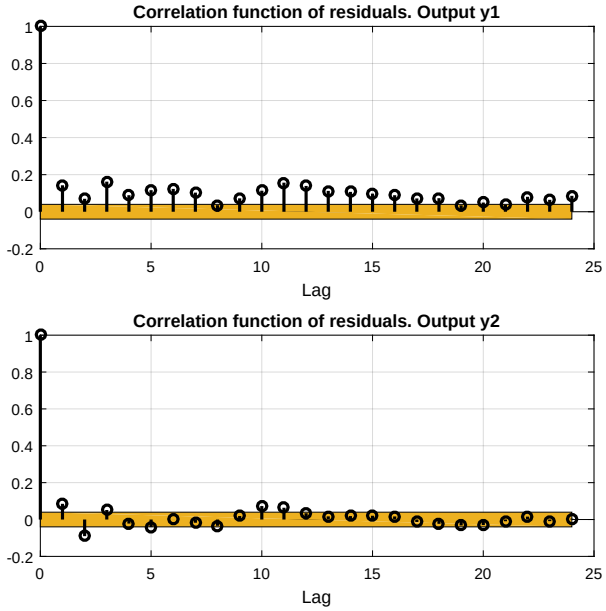


(c) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_3 for model 2.

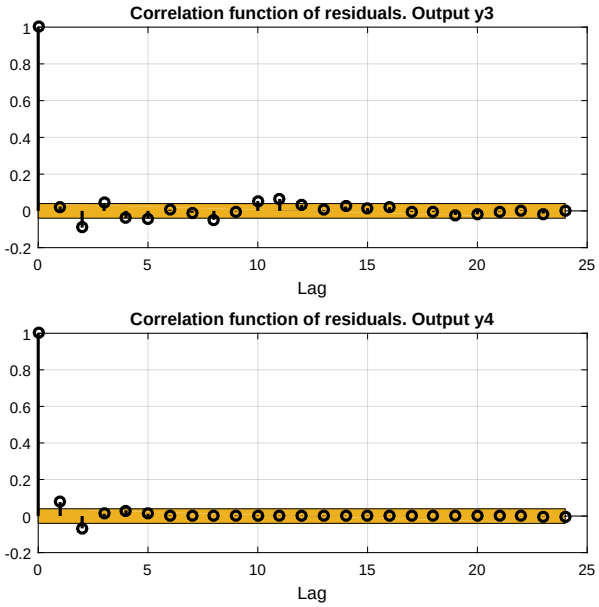


(d) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_4 for model 2.

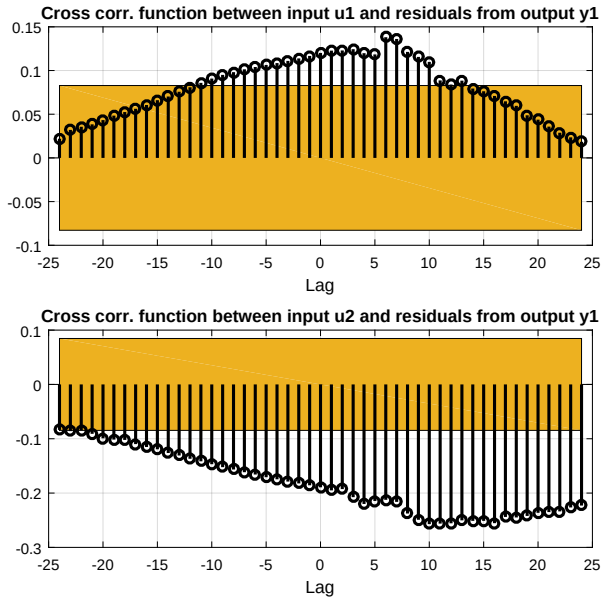
Figure E.13 Cross correlation between lagged inputs and residuals of outputs Y_1, Y_2, Y_3 and Y_4 for model 2.



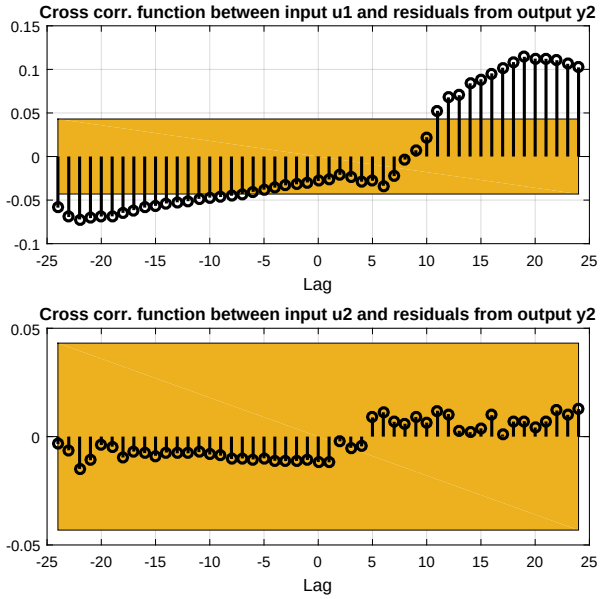
(a) Correlation between residuals and lagged inputs for Y_1 and Y_2 for model 3.



(b) Correlation between residuals and lagged inputs for Y_3 and Y_4 for model 3.

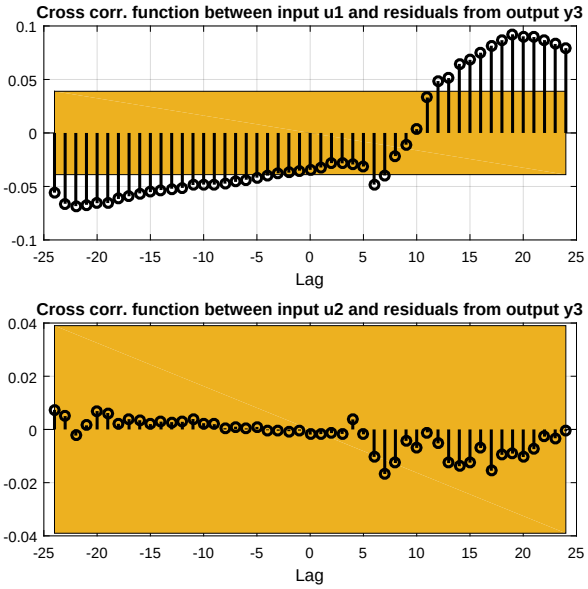


(c) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_1 for model 3.

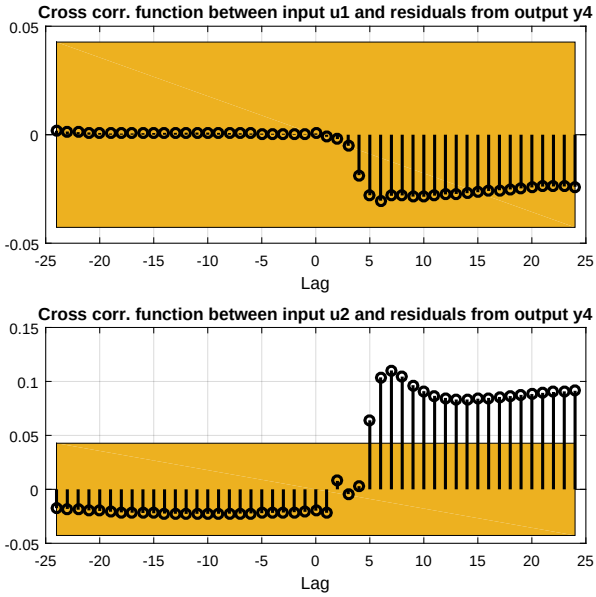


(d) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_2 for model 3.

Figure E.14 Correlation between residuals and lagged inputs for outputs Y_1, Y_2, Y_3 and Y_4. Cross correlation between lagged inputs and output residuals for outputs Y_1 and Y_2 of model 3.



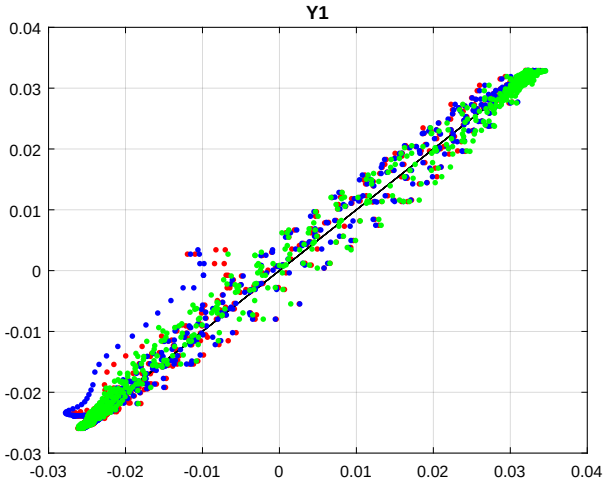
(a) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_3 for model 3.



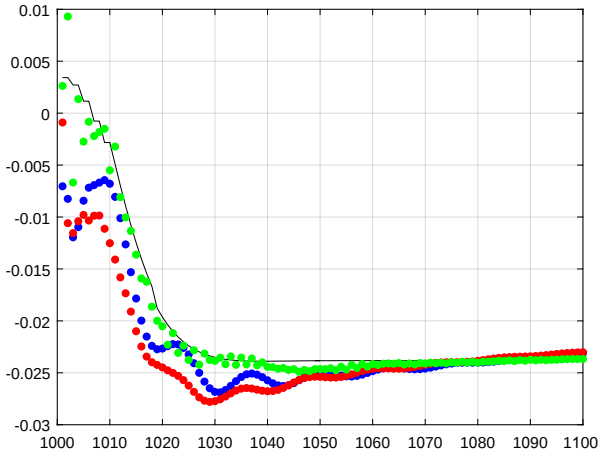
(b) Cross correlation between lagged inputs U_1 and U_2, and residuals for Y_4 for model 3.

Figure E.15 Cross correlation between lagged inputs and output residuals of model 3. Outputs Y_3 and Y_4.

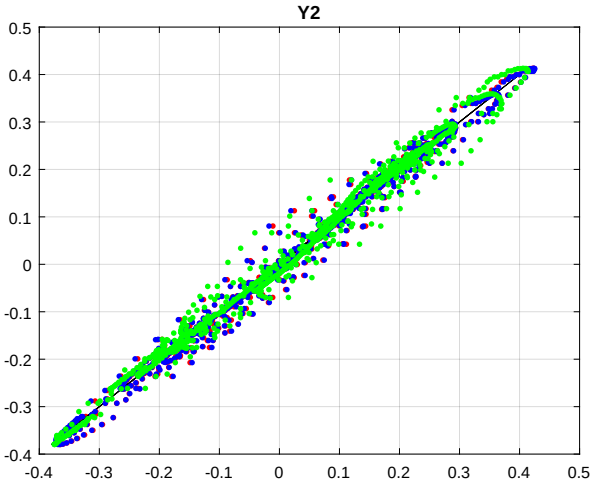
E.7 Cross validation of infinite step-ahead predictions



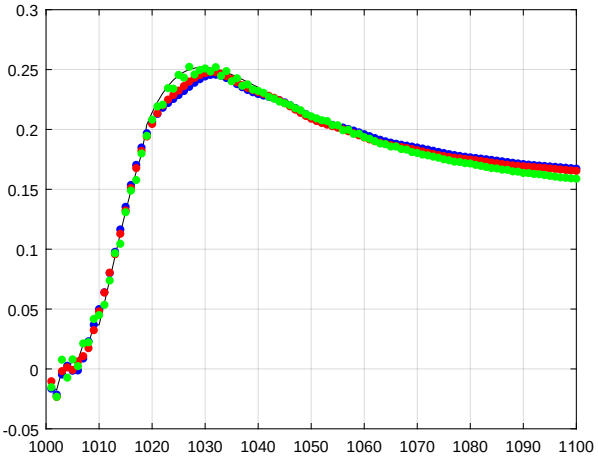
(a) Y1 output scatter plot



(b) Y1 zommed output plot

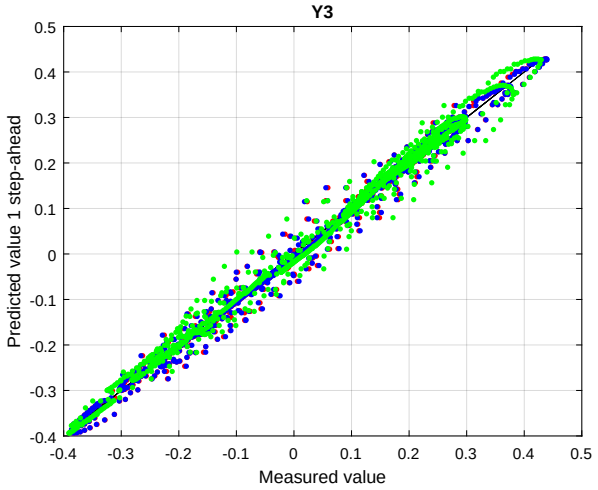


(c) Y2 output scatter plot

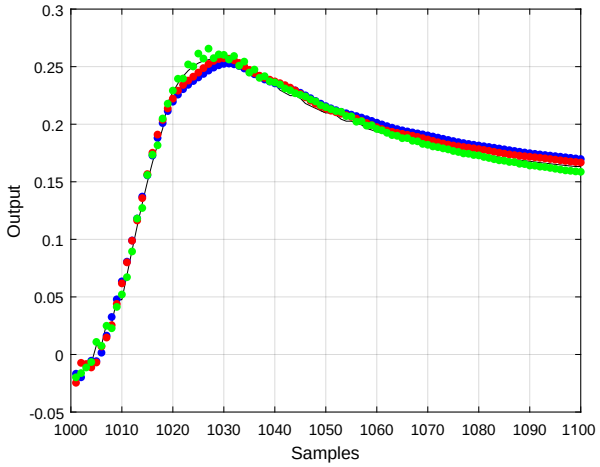


(d) Y2 zommed output plot

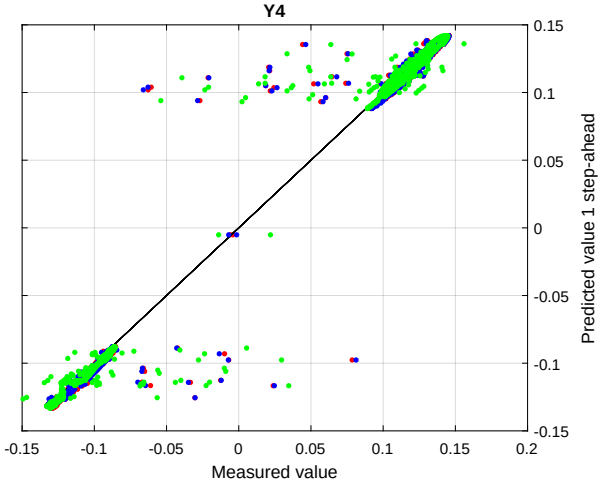
Figure E.16 Comparing infinite step-ahead predictions against process data for three models. Scatter plots on the left and zommed plots to the right. Red scatter represents model 1, blue scatter represents model 2 and green scatter represents model 3. Outputs Y_1 and Y_2 depicted. Black lines indicate simulator data values.



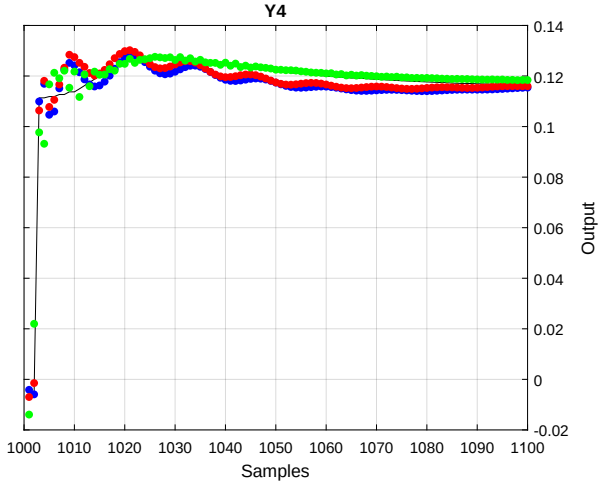
(a) Y3 output scatter plot



(b) Y3 zommed output plot



(c) Y4 output scatter plot

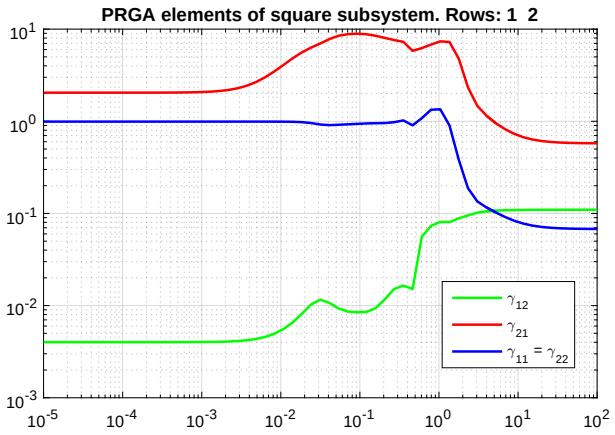


(d) Y4 zommed output plot

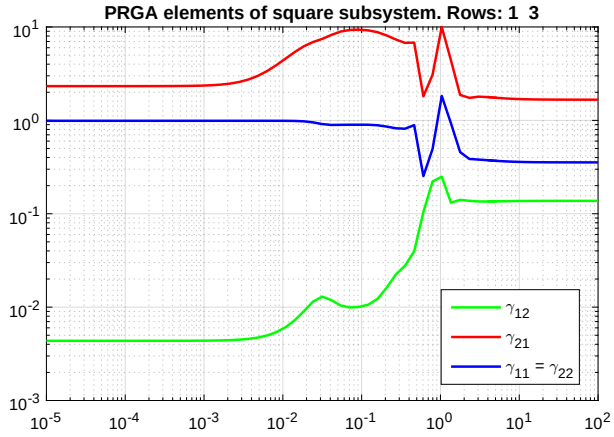
Figure E.17 Comparing infinite step-ahead predictions against process data for three models. Scatter plots on the left and zommed plots to the right. Red scatter represents model 1, blue scatter represents model 2 and green scatter represents model 3. Outputs Y_3 and Y_4 depicted. Black lines indicate simulator data values.

E.8 Dynamical PRGA and singular value plots

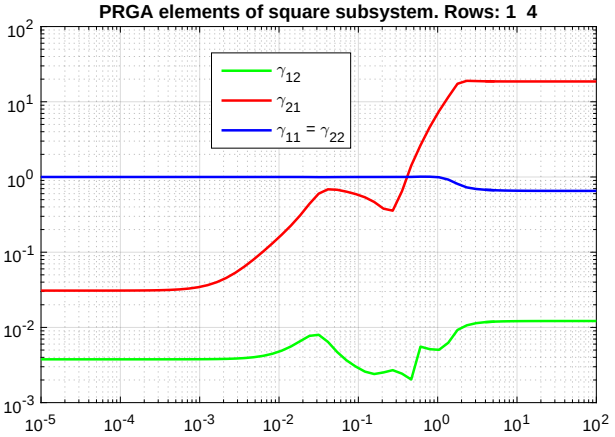
Dynamical PRGA of the different square plant subsections are computed in this section. The main motivation is to illuminate one-way interactions not which are not indicated by the (N)RGA computations. For considering decentralized control, this is an important aspect.



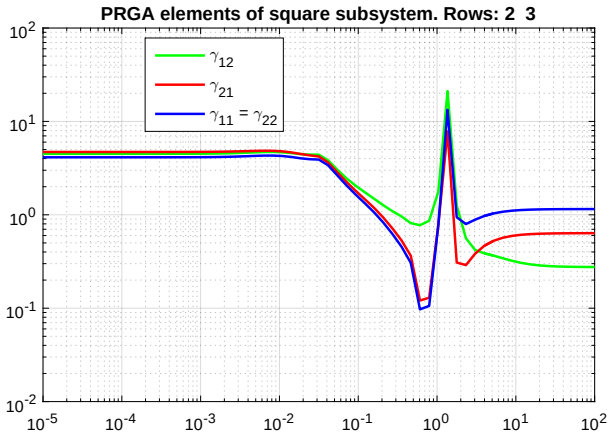
(a) PRGA rows 1 & 2 from NRGA matrix



(b) PRGA rows 1 & 3 from NRGA matrix

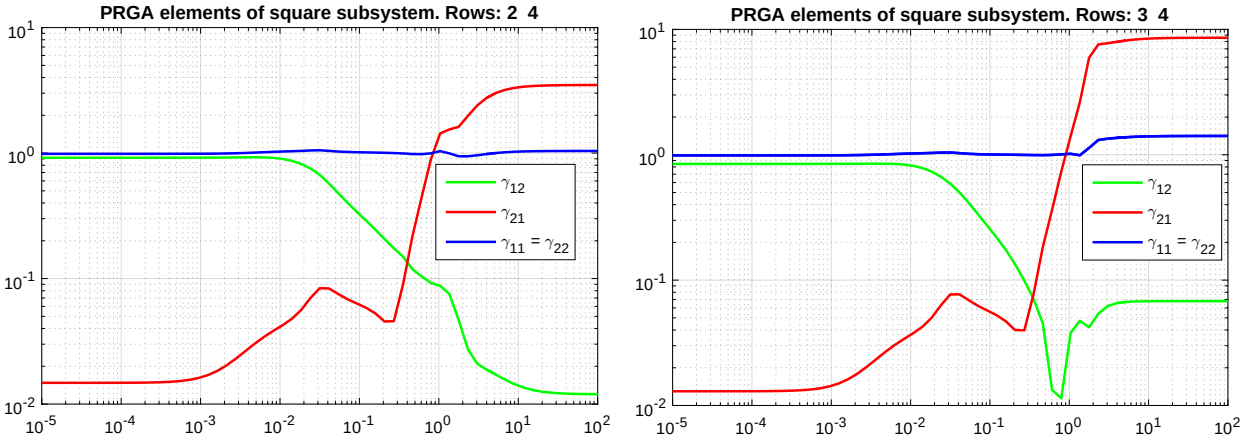


(c) PRGA rows 1 & 4 from NRGA matrix



(d) PRGA rows 2 & 3 from NRGA matrix

Figure E.18 Comparing possible 2x2 pairings based on PRGA input-output pairings based on down-squaring of the non-square plant.

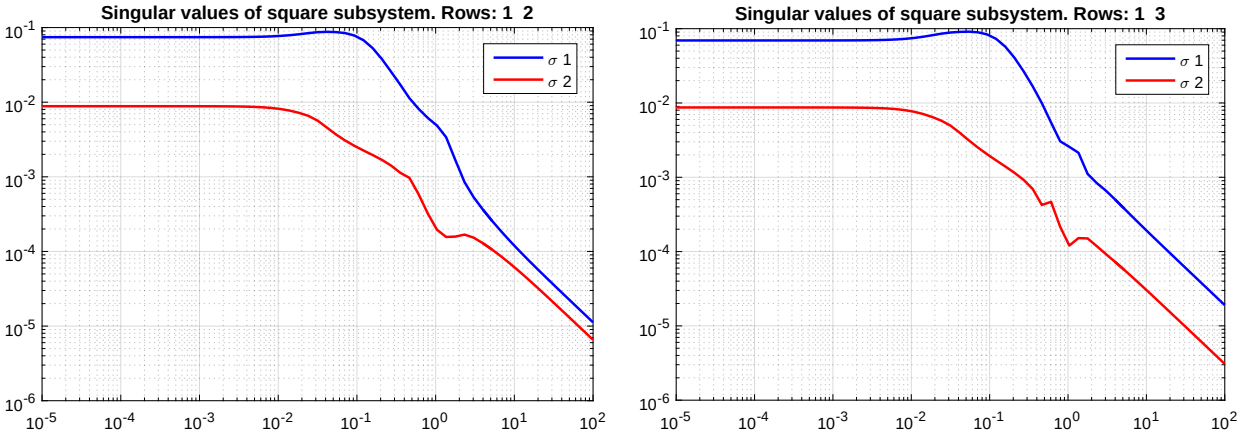


(a) PRGA rows 2 & 4 from NRGA matrix

(b) PRGA rows 4 & 3 from NRGA matrix

Figure E.19 Comparing possible 2x2 pairings based on PRGA input-output pairings based on down-squaring of the non-square plant.

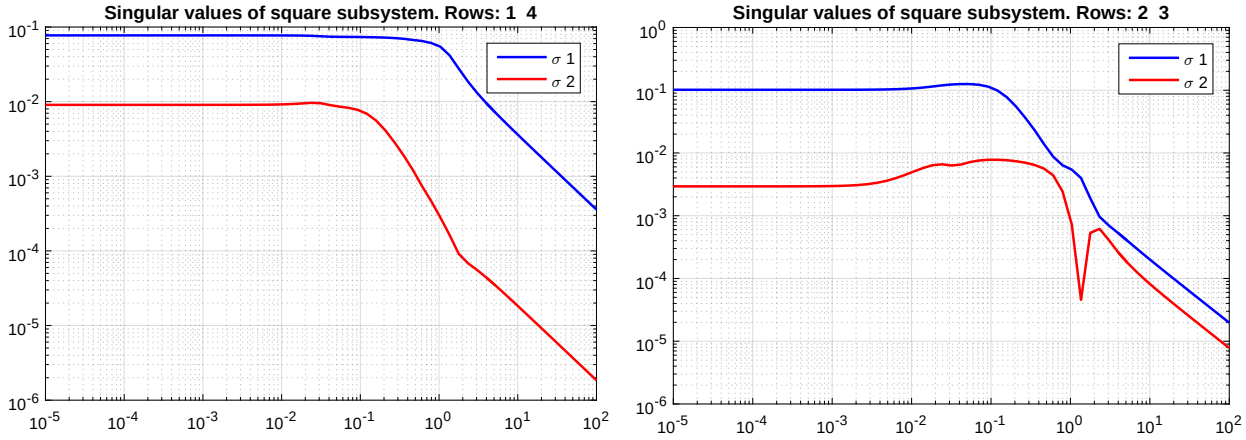
Dynamical singular values. Skogestad and Postlethwaite [69] state that it is impossible with independent control of outputs when $\underline{\sigma}(G(j\omega)) < 1$. As observed from the singular value figures, they indicate small values which satisfy above statement. However, considering the condition number indicate the opposite.



(a) Singular values for rows 1 & 2

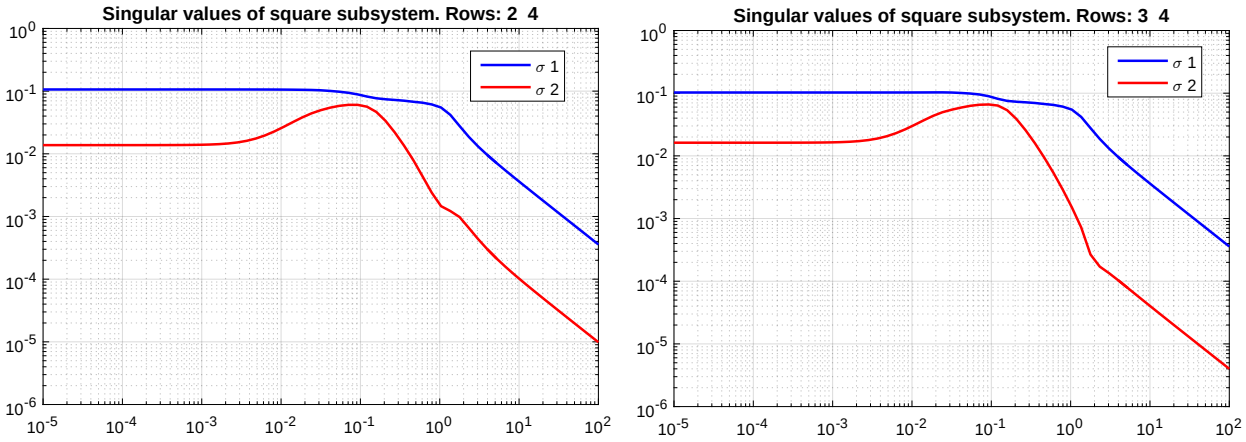
(b) Singular values for rows 1 & 3

Figure E.20 Dynamical singular values for input-output pairings based on down-squaring of the non-square plant.



(a) Singular values for rows 1 & 4

(b) Singular values for rows 2 & 3



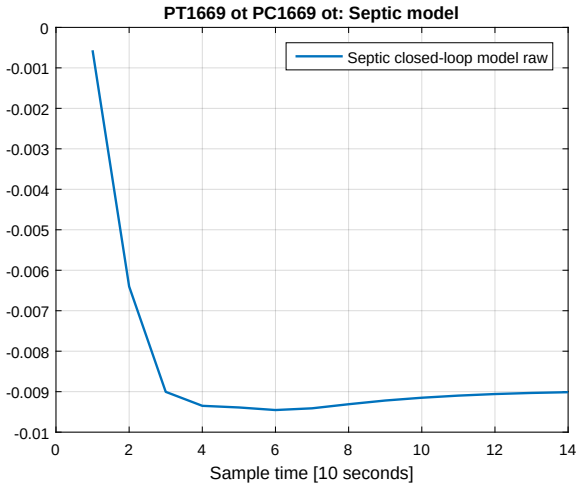
(c) Singular values for rows 2 & 4

(d) Singular values for rows 3 & 4

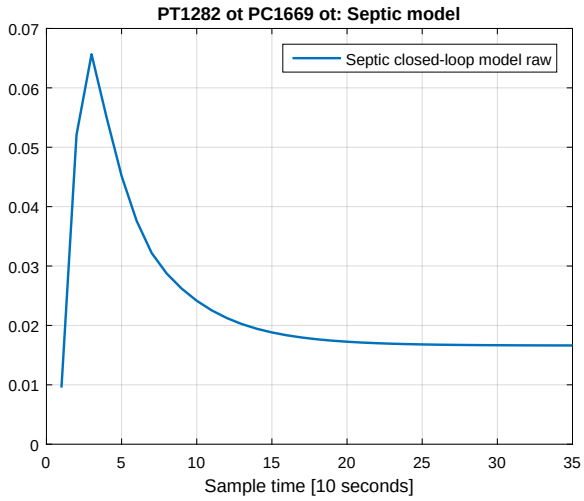
Figure E.21 Dynamical singular values for input-output pairings based on down-squaring of the non-square plant.

E.9 SEPTIC identified closed-loop models

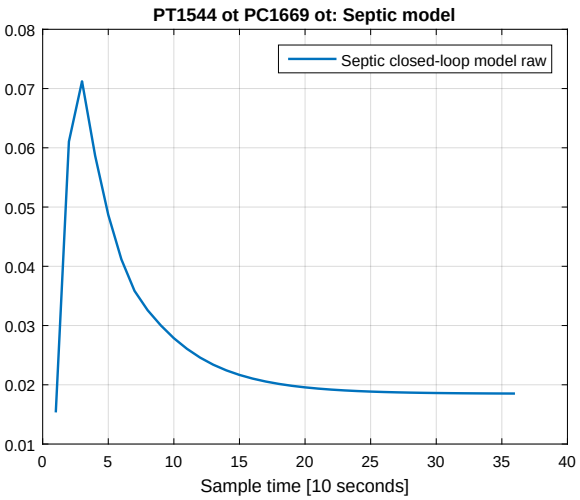
As mentioned in section 7.2, closed-loop models were identified for the supervisory MPC apllication. Identification was carried out in SEPTIC and figures E.22 and E.23 depict the closed-loop responses. For simplicity, the identification experiments were executed using step perturbations in the controller references. An alternative to this approach of obtaining closed-loop models is to manually extend the open-loop model by adding the controller term and feedback. This resembles a twist of the better known *indirect* closed-loop identification discussed in section 3.4. Obviously, this method requires a known controller term. The alternative method was not utilized in this work, but it some general information



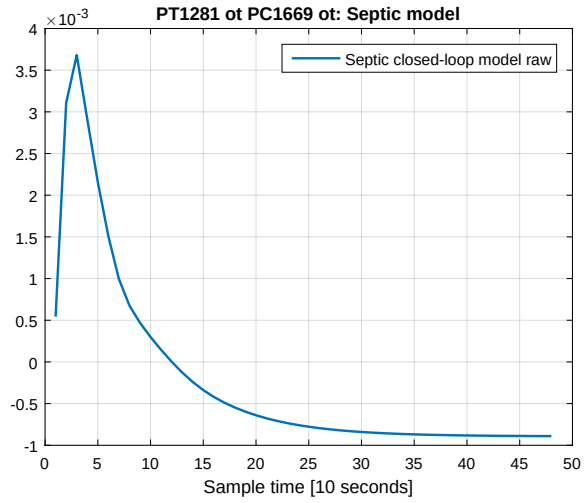
(a) Closed-loop model of $PC1669 \rightarrow Y_1$



(b) Closed-loop model of $PC1669 \rightarrow Y_2$

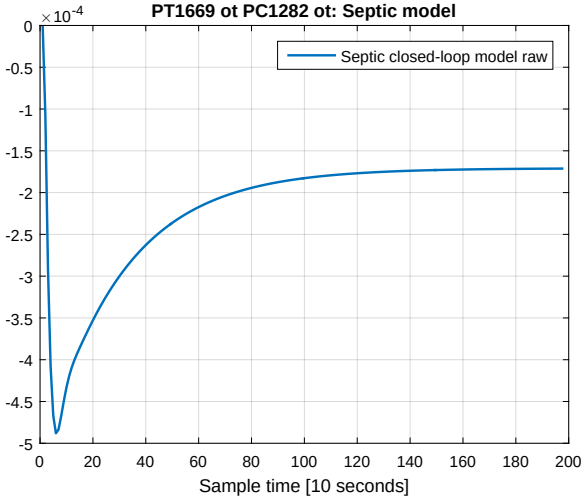


(c) Closed-loop model of $PC1669 \rightarrow Y_3$

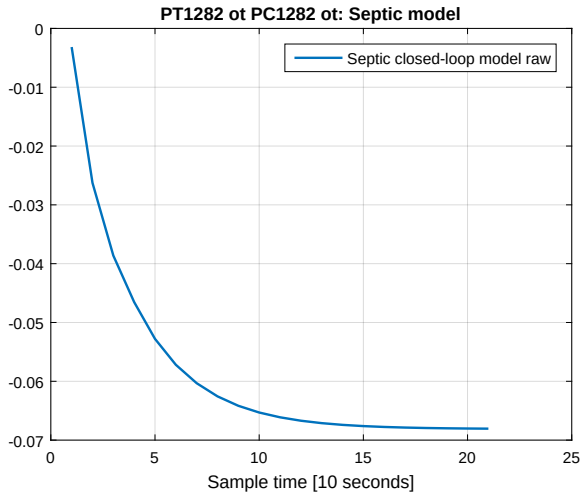


(d) Closed-loop model of $PC1669 \rightarrow Y_4$

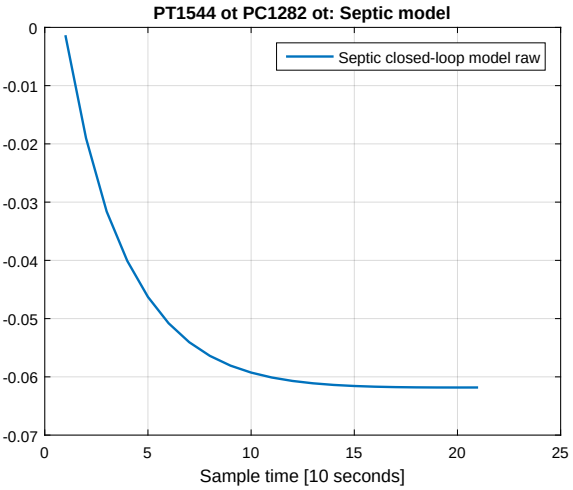
Figure E.22 SEPTIC identified closed-loop models for controller PC1669.



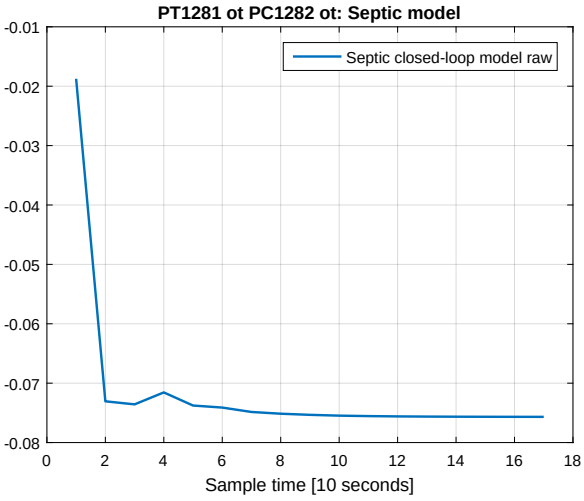
(a) Closed-loop model of $PC1282 \rightarrow Y_1$



(b) Closed-loop model of $PC1282 \rightarrow Y_2$



(c) Closed-loop model of $PC1282 \rightarrow Y_3$



(d) Closed-loop model of $PC1282 \rightarrow Y_4$

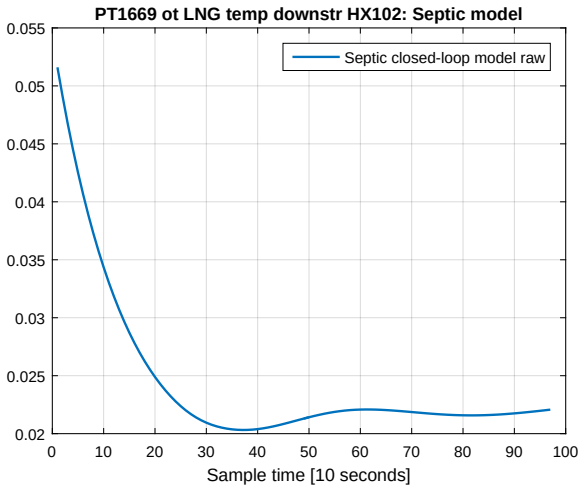
Figure E.23 SEPTIC identified closed-loop models for controller PC1282.

Although these are closed-loop models from the system investigated, they are depicted as step responses. As earlier mentioned, independent of which model structure one chose to develop, SEPTIC translates all models to unit step response models.

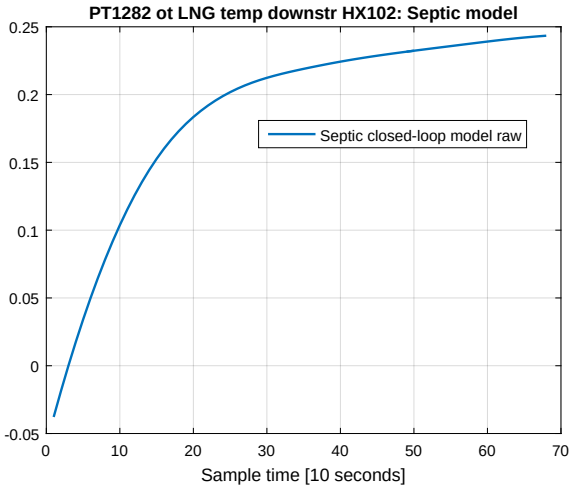
E.10 Step responses for a selected disturbance

The figures included in this appendix display how a change in the LNG temperature prior to entering the subcooling cycle affects the four CVS considered in this work. The figures depict a unit step response although the disturbance models are identified either by ARX or FIR models.

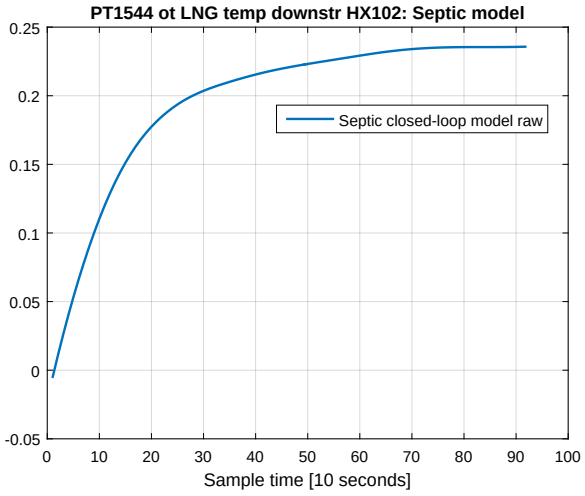
This disturbance correspond to the entry labelled *LNG temperature downstream subcooler* in table 7.2. For the reader it may seem peculiar to consider a disturbance downstream of the subcooler cycle. However, this disturbance affects the flow of LNG which directly affects flow of SMR which certainly influences the four CVs considered. The disturbance may originate from adjustments in the nitrogen removal column or issues with the LNG expander turbine. For further reading on this part of the MFC process, the reader may consult shapter 3 in Volden [88]. Although not depicted in this thesis, there exist 16 additional models for the total 5 disturbances on 4 Cvs considered in this work.



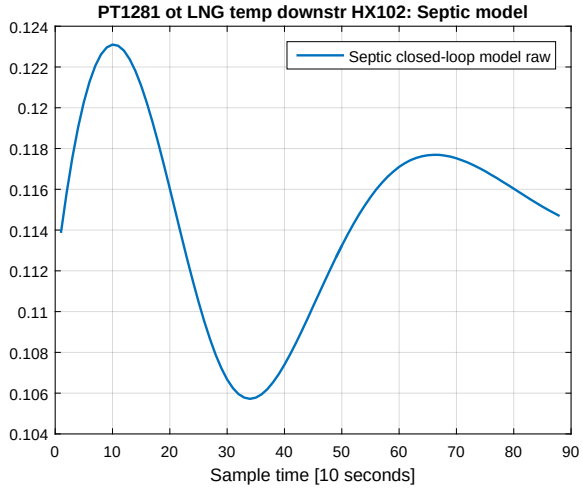
(a) LNG temperature disturbance on Y_1



(b) LNG temperature disturbance on Y_2



(c) LNG temperature disturbance on Y_3



(d) LNG temperature disturbance on Y_4

Figure E.24 Selected disturbance step response models. Depicted disturbance is LNG temperature subsequent of liquefaction prior to entering subcooler.

E.11 Additional figures depicting SEPTIC priority

This appendix includes two supervisory MPC simulations. The intention is to provide additional plots to emphasize the significance of priorities. The simulations include a sequence of changing setpoints for a choosen CV. The difference is the setpoint priority for said variable, which is altered from 10 → 1. This is indicated in the subsequent figures.

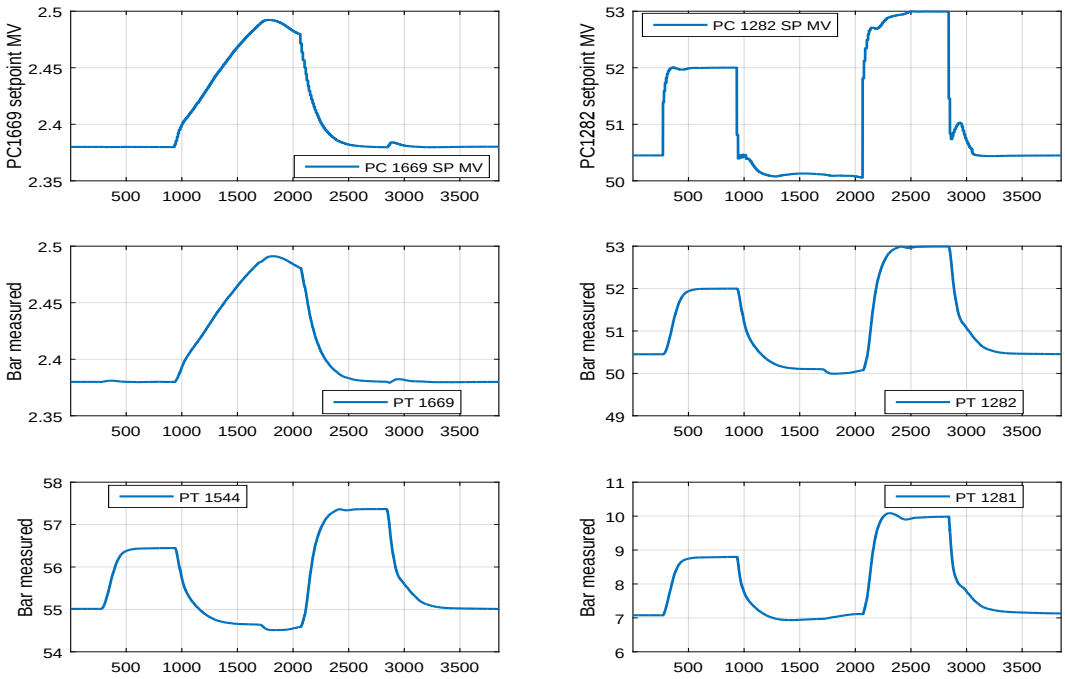


Figure E.25 Setpoint altering for Y_2 with priority 1.

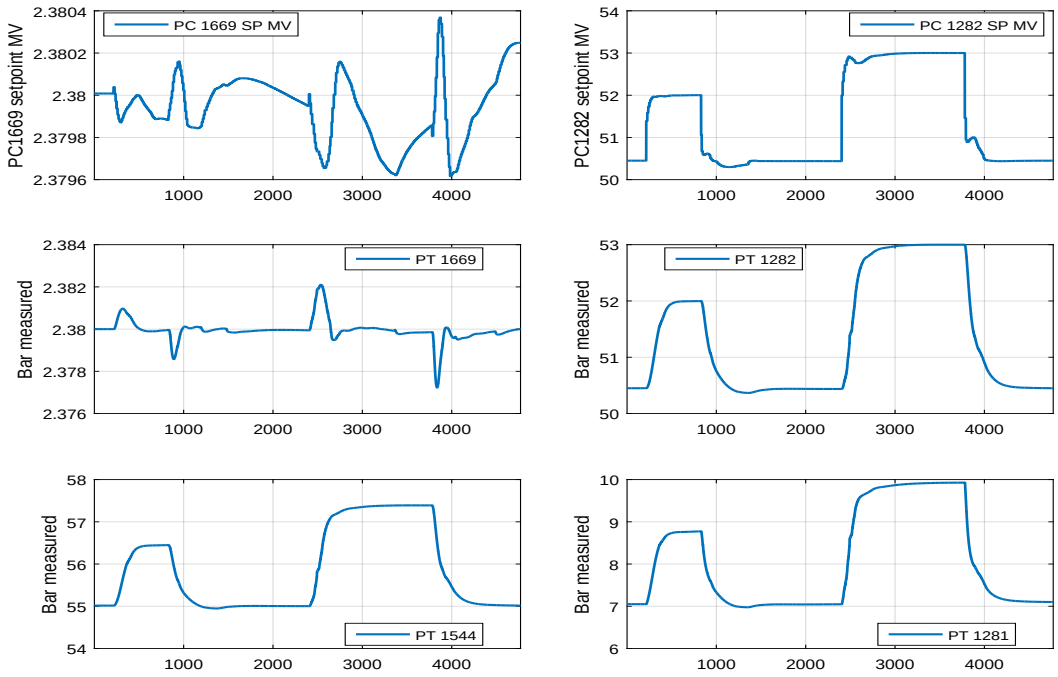


Figure E.26 Setpoint altering for Y_2 with priority 10.

Giving the variable a higher priority allows more freedom to alter the regulatory layer setpoints, which depicted in the figures. The consequence is less consideration for variables, and this is clearly observed in the PT1669 subplot. In figure E.26 the PT1669 values are kept close to the setpoint which is 2.38 bar. However, altering the priority to 1 for PT1281 as figure E.25 depicts, PT1669 is not controlled tight. Thus, the MPC controller focus to a greater extent on manipulating PT1281 to its setpoint.

E.12 Additional disturbance simulation figures

The two remaining disturbances which are not depicted in section 7.4 are included in this appendix. The disturbances are simulated for regulatory layer control, SEPTIC supervisory application and SEPTIC direct application, respectively.

Sea water temperature disturbance

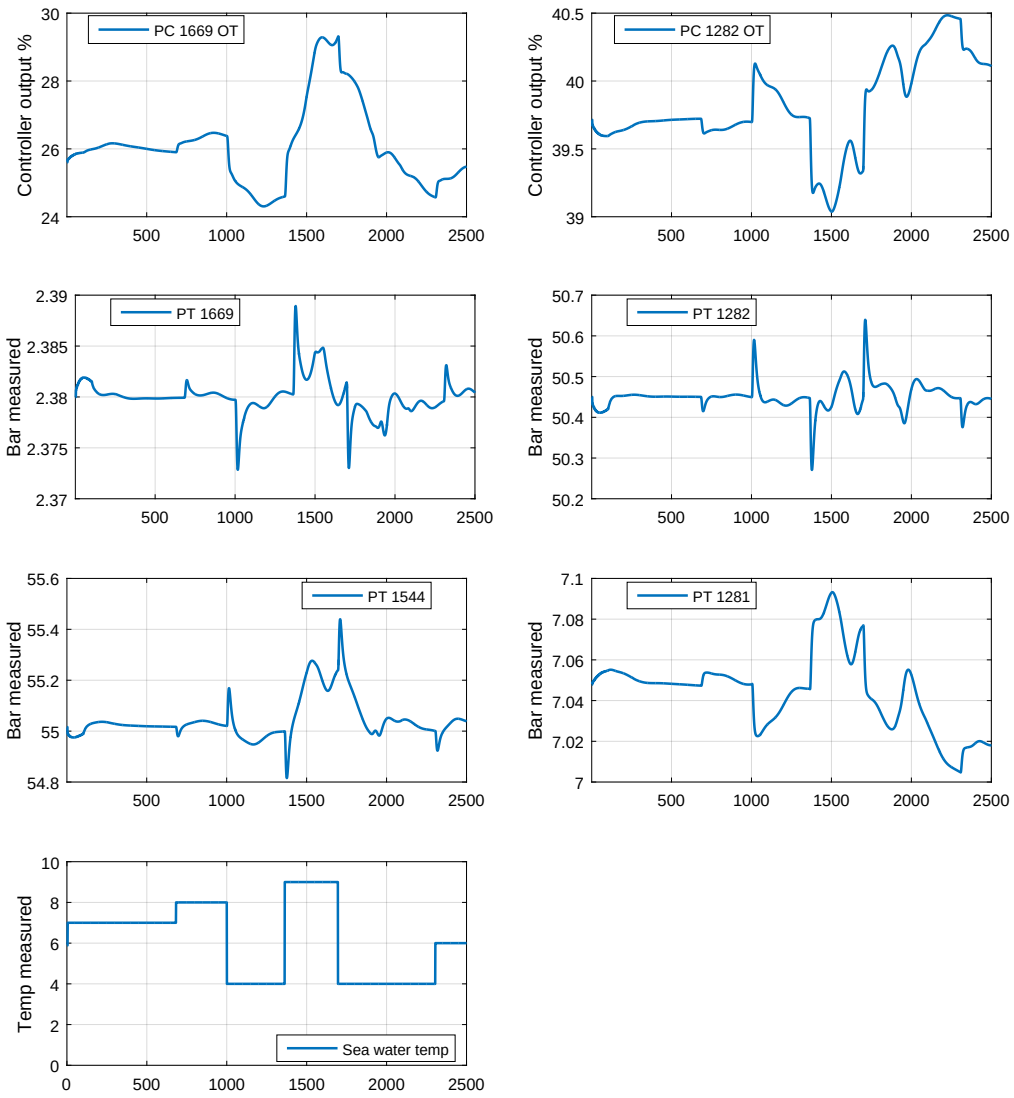


Figure E.27 Rejection of varying sea water temperature from regulatory layer. Sea water temperature is manipulated in steps as depicted in bottom subplot.

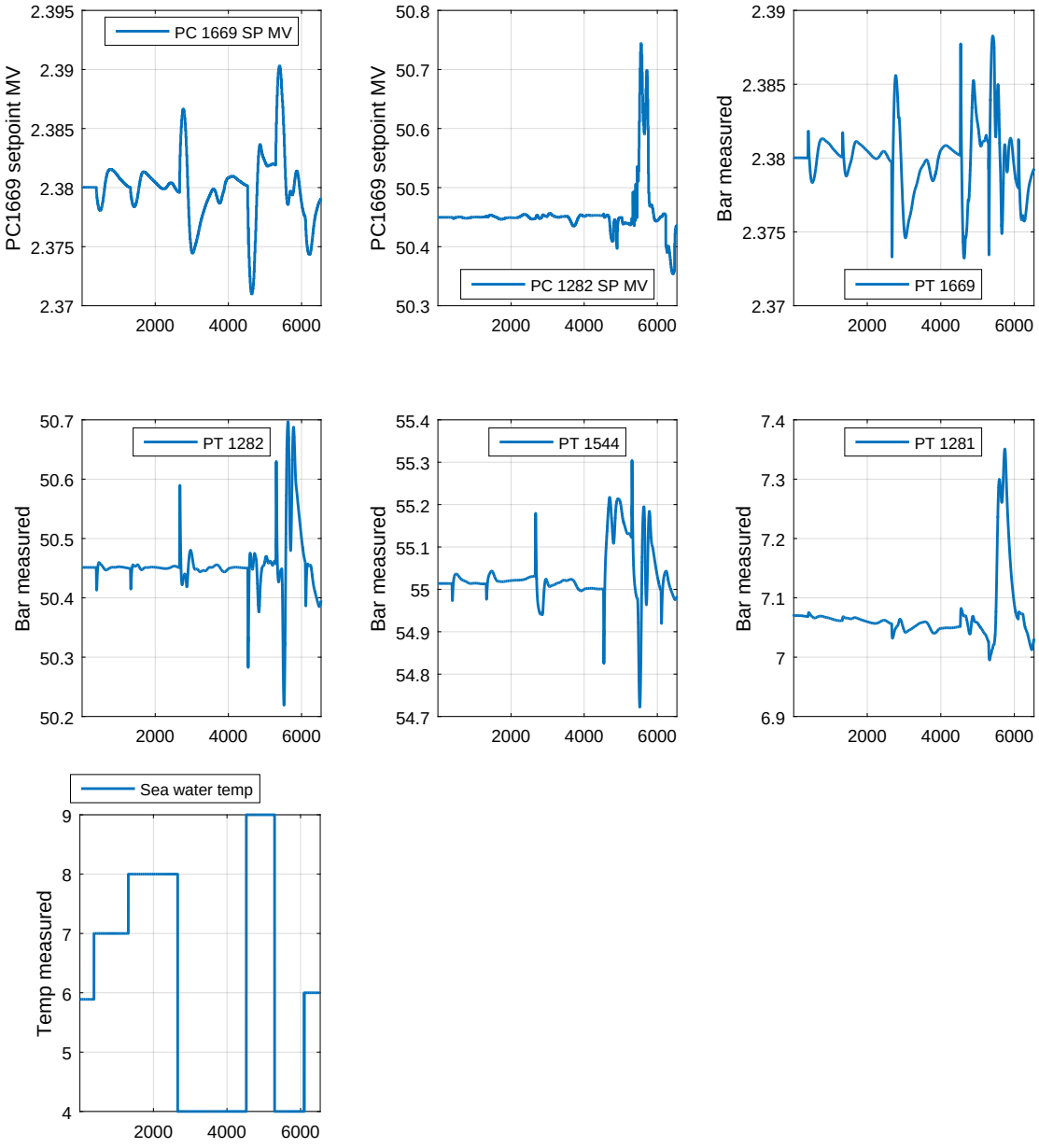


Figure E.28 Rejection of varying sea water temperature from SEPTIC supervisory application. Sea water temperature is manipulated in steps as depicted in bottom subplot.

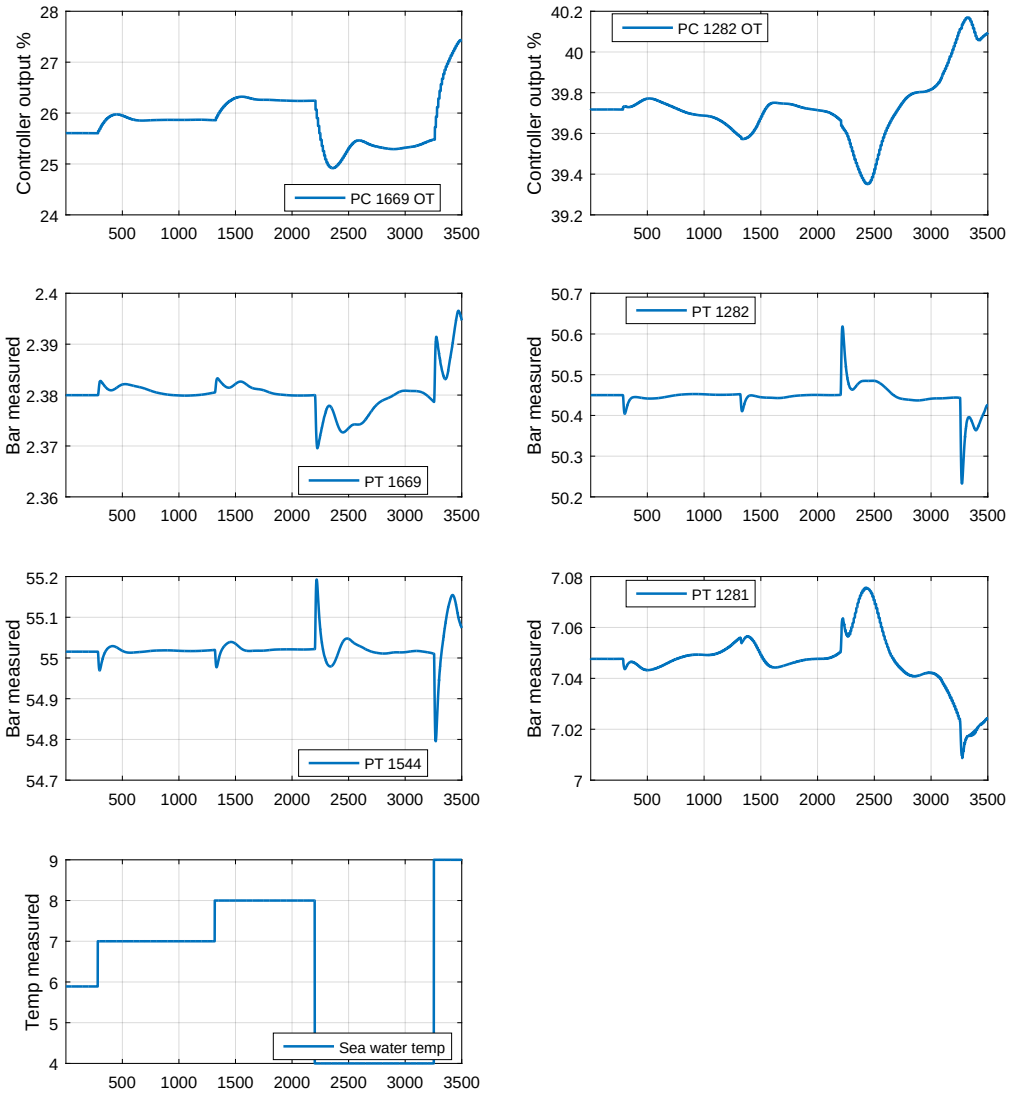


Figure E.29 Rejection of varying sea water temperature from SEPTIC direct application. Sea water temperature is manipulated in steps as depicted in bottom subplot.

Upstream liquefaction cycle feed gas pressure disturbance

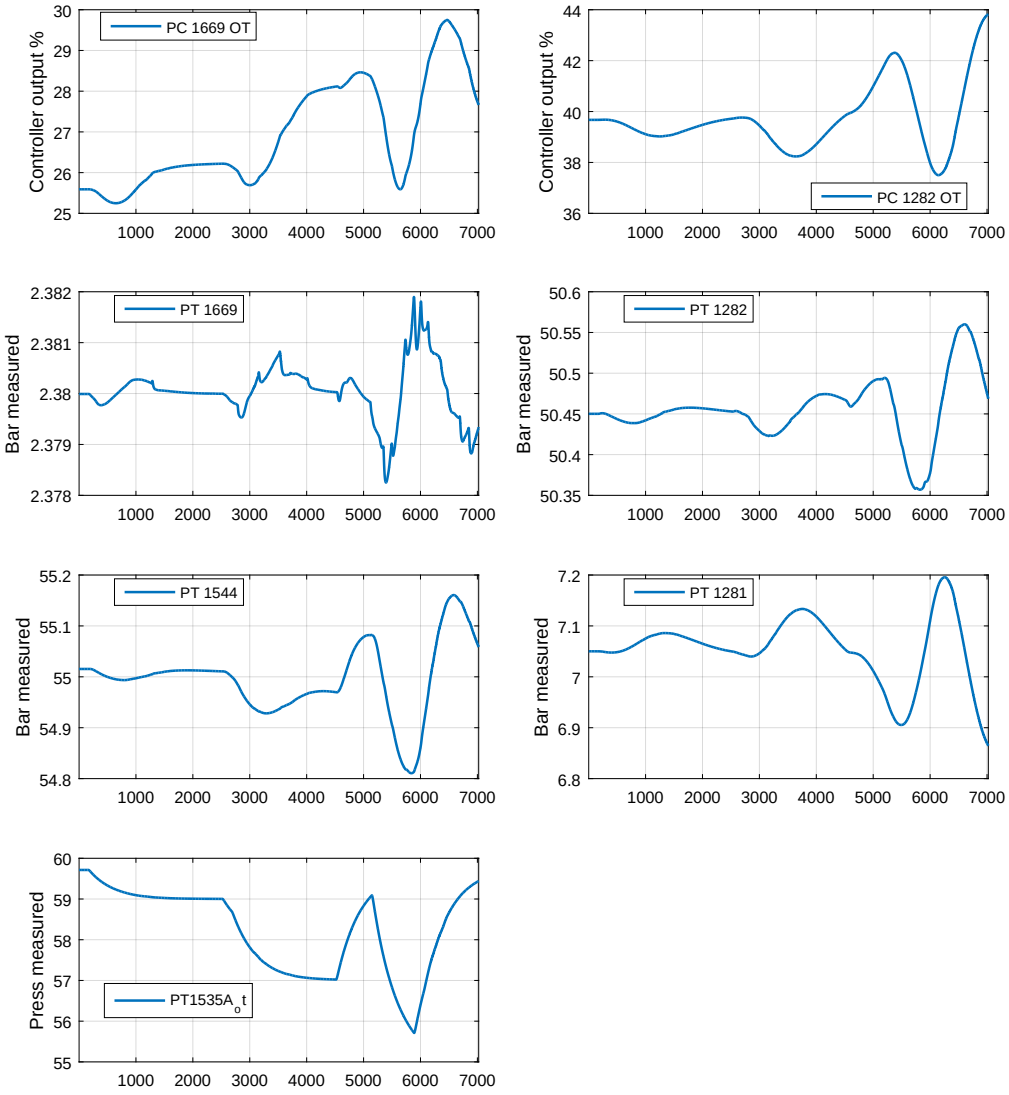


Figure E.30 Rejection of varying feed gas pressure upstream liquefaction from regulatory layer. Upstream feed gas pressure is manipulated as depicted in bottom subplot.

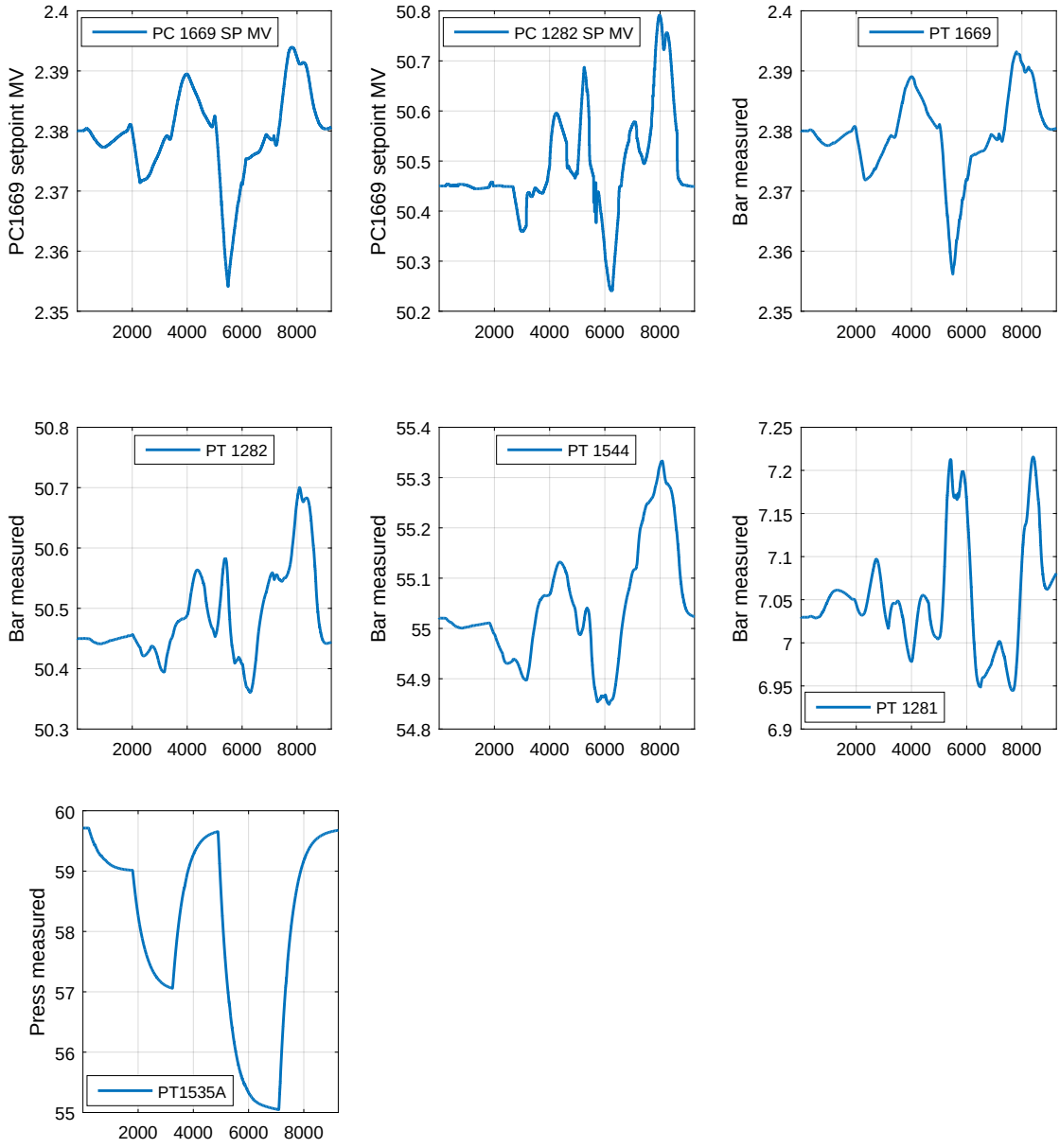


Figure E.31 Rejection of varying feed gas pressure upstream liquefaction from SEPTIC supervisory application. Upstream feed gas pressure is manipulated as depicted in bottom subplot.

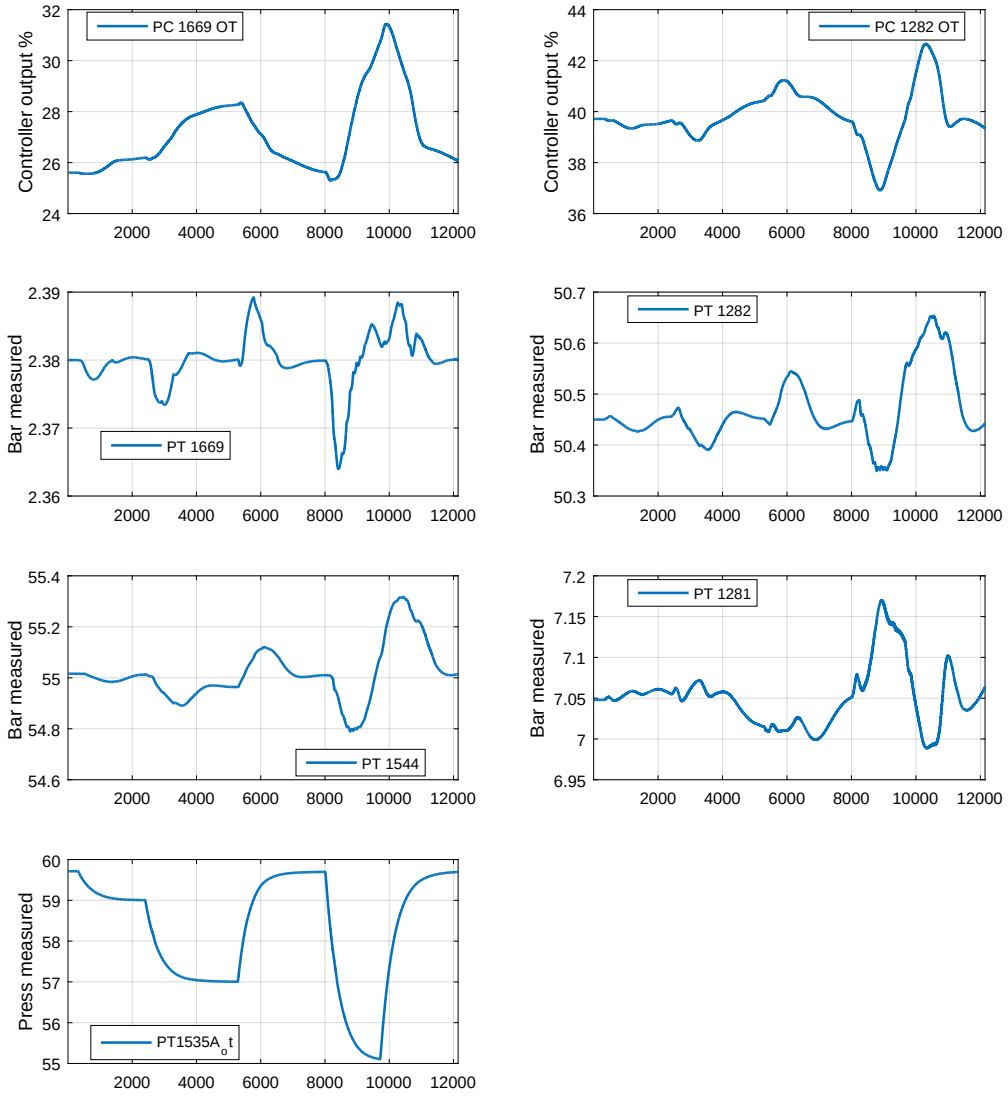


Figure E.32 Rejection of varying feed gas pressure upstream liquefaction from SEPTIC direct application. Upstream feed gas pressure is manipulated as depicted in bottom subplot.