Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Industrial Economics and Technology Management

Rakel Sjaastad Andreassen
Eline Myhrum Westby

Integrating Machine Learning Techniques in Real-Time Production Optimization

Master's thesis in Industrial Economics and Technology Management
Supervisor: Henrik Andersson
June 2019

NTNU
Kunnskap for en bedre verden

MASTER'S THESIS

Integrating Machine Learning Techniques in Real-Time Production Optimization

Authors:
RAKEL SJAASTAD ANDREASSEN
ELINE MYHRUM WESTBY

Supervisors:
HENRIK ANDERSSON
BJARNE GRIMSTAD

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF ECONOMICS AND MANAGEMENT

DEPARTMENT OF INDUSTRIAL ECONOMICS AND TECHNOLOGY MANAGEMENT

June 6, 2019

Preface

This Master's Thesis constitutes the completion of our Master of Science degrees in Industrial Economics and Technology Management, with specialization in Managerial Economics and Operations Research, at the Norwegian University of Science and Technology (NTNU). The thesis is a continuation of our project report, written as part of the subject TIØ4500 - Managerial Economics and Operations Research in the Fall of 2018.

The motivation for this work is our interest in how data and machine learning techniques can be used for the modelling of real-life systems, and the opportunity to work with the highly relevant research area of real-time petroleum production optimization. It has been truly interesting and challenging to work with this problem.

We would like to express our gratitude to industry partner Solution Seeker AS, and especially main representative Bjarne Grimstad, for valuable insights and interesting discussions. We would also like to thank our supervisor, Professor Henrik Andersson, for helpful guidance and valuable advice throughout the year.

Trondheim, June 6th, 2019.

Rakel Sjaastad Andreassen and Eline Myhrum Westby

Summary

The topic of this thesis is the integration of machine learning techniques in real-time optimization (RTO) in the upstream sector of the petroleum production industry. The main target of the real-time optimization problem is to maximize oil production over a short time horizon by performing small adjustments to the operation settings. The output from the optimization are suggestions for how the choke openings of the wells in the system should be adjusted from a given operation point.

Historical measurement data from a field is applied to model relationships between choke opening, wellhead pressure and well flow rates for each well, in addition to pressure drops in pipelines. Multivariate adaptive regression splines (MARS), a regression-based machine learning technique, is applied for this purpose. A novel method of integrating MARS models directly in optimization is proposed. Including both pressure and choke opening in the optimization, and applying data-driven modelling to find these relationships, has not been seen in previous literature. These are important contributions towards attaining more realistic models of the production system.

A stochastic mixed integer linear programming (MILP) problem, taking the uncertainty in the MARS models into account, is introduced. The models are integrated as scenarios in the optimization by building them on varying subsets of the data source. The models therefore differ from one scenario to another. An extension to this problem includes the additional objective of minimizing the conditional value-at-risk (CVaR) with regards to gas production. The problem provides a method of studying uncertainty related to the risk of breaching the gas capacity constraint, and enables a varying level of conservativeness of the solution.

The problem of cyclic pressure and flow instability in risers, known as slugging, is included in another extension of the problem. Using the classification tree (CT) technique, a model that predicts the risk of slugging for different operation areas is built based on measurement data. The CT is integrated in the stochastic MILP problem with the additional objective of minimizing the cost of slugging, providing decision support by identifying the optimal solution for a given level of risk aversion.

The main challenge when applying machine learning in linear stochastic optimization is found to be the trade-off between accurate machine learning models and including a sufficient number of scenarios in the optimization representing the stochastic variables. While the processes are more correctly modelled with advanced models, the uncertainty can be better captured with more scenarios. This thesis explores how to allocate the processing power available in order to make the best possible decisions within a reasonable time frame. Results show that the scalability as the number of scenarios in the problem increases is highly dependent on the complexity of the machine learning models. Optimal machine learning models might not be optimal when integrated in stochastic optimization, and methods for finding suitable complexity levels with respect to both the bias-variance trade-off and solution time should be further developed.

Sammendrag

Denne avhandlingen undersøker hvordan å integrere maskinlæringsteknikker i sanntidsoptimering (RTO) i oppstrømssektoren i petroleumsproduksjonsindustrien. Hovedmålet med optimeringsproblemet er å maksimere oljeproduksjonen over en kort tidshorisont ved å justere innstillingene i produksjonssystemet. Mer spesifikt er resultatet fra optimeringen anbefalinger til hvordan chokeåpningene til brønnene i systemet skal justeres.

I oppgaven benyttes brønnmålingsdata fra et petroleumsfelt til å modellere relasjonene mellom chokeåpning, brønnhodetrykk og produksjonsrater for hver brønn, i tillegg til trykkfall over rørledninger. Multivariate adaptive regression splines (MARS), en regresjonsbasert maskinlæringsteknikk, benyttes i modelleringen. En ny formulering som muliggjør integrering av MARS-modeller direkte i lineær optimering foreslås. Data-drevet modellering av trykk og chokeåpninger i optimeringen er ikke beskrevet i tidligere litteratur, og er viktige bidrag til å oppnå mer realistiske modeller av produksjonssystemet.

Avhandlingen introduserer et stokastisk blandet heltallsproblem (MILP), som tar hensyn til usikkerheten i MARS-modellene. Modellene er bygget på ulike subsett av datagrunnlaget og er integrert som scenarier i optimeringen. Modellene varierer derfor fra ett scenario til et annet. For å ta hensyn til usikkerheten knyttet til risikoen for å bryte gasskapasiteten, er problemet utvidet med objektivet om å minimere conditional value-at-risk (CVaR) for gassproduksjon. Utvidelsen gjør det mulig å styre hvor konservativ løsningen skal være.

Slugging er et fenomen som kan oppstå i rørledninger i et produksjonssystem og som kan føre til oscillerende trykk og ujevn strømning. Slugging tas hensyn til i enda en utvidelse av optimeringsproblemet. Maskinlæringsteknikken klassifiseringstrær brukes til å predikere risikoen for slugging for forskjellige operasjonsområder basert på måledata fra et felt. Klassifiseringstreet integreres direkte i det stokastiske MILP-problemet, hvor risikoen for slugging minimeres samtidig som oljeproduksjonen maksimeres. Modellen gir beslutningsstøtte ved å identifisere den optimale løsningen for ulik grad av risikoaversjon.

Hovedutfordringen ved bruk av maskinlæring i lineær stokastisk optimering er avveiningen mellom nøyaktige maskinlæringsmodeller og et tilstrekkelig antall scenarier i optimeringen, der scenariene representerer de stokastiske variablene. Selv om prosessene kan modelleres mer nøyaktig med avanserte modeller, kan usikkerheten bedre fanges opp med flere scenarier. Denne oppgaven undersøker hvordan tilgjengelig prosessorkraft kan fordeles for å kunne ta de beste beslutningene innenfor en rimelig tidsramme. Resultatene viser at skalerbarheten er svært avhengig av kompleksiteten til maskinlæringsmodellene. Optimale maskinlæringsmodeller kan vise seg å være mindre optimale når de integreres i stokastisk optimering, og rammeverket for å finne passende kompleksitetsnivåer med hensyn til bias og varians samt løsningstid bør videreutvikles.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Crude oil, or petroleum, is the unrefined composition of hydrocarbon deposits and other organic material. It is formed under high pressure and temperature and exists as a liquid mixture in reservoirs underneath the Earth's surface. Petroleum has been central to all aspects of the modern industrial society. It is used as an energy source and as raw material in the production of different chemicals and synthetic materials. Petroleum has been particularly important as a source of transportation fuels, and it is difficult to find equally performing, cost-effective substitutes (Vassiliou, 2018). Using petroleum as an energy source causes emissions that are damaging to the environment, and the use of alternative energy sources is therefore increasing. Even though there is progress in the development of renewable energy sources, the world's energy consumption is still highly dependent on petroleum production. The global oil consumption is today around 100 million barrels per day, which is more than twice the consumption 50 years ago (Cooper and Johnson, 2018).

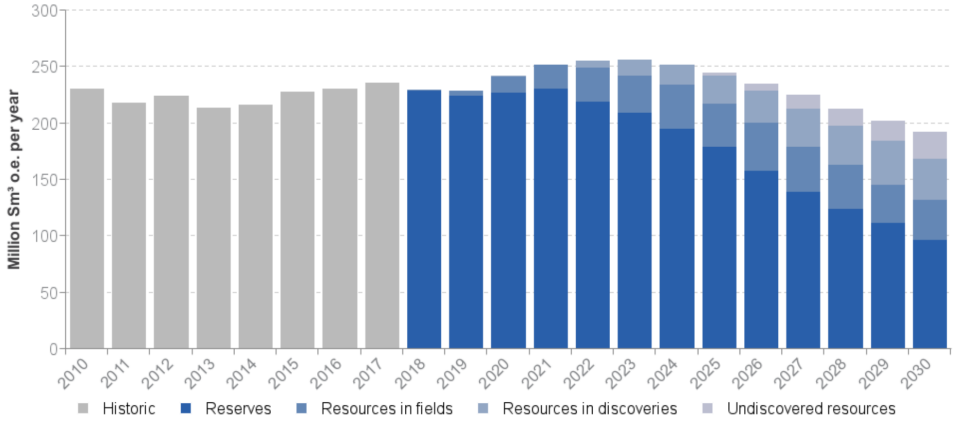

Figure 1.1: Production history and forecast per resource category (Norwegian Petroleum Directorate, 2018)

The level of activity on the Norwegian continental shelf is expected to be high over the next decades. The Norwegian Petroleum Directorate (2018) estimates that there are about 55 per cent left of the total recoverable resources on the continental shelf. Figure 1.1 shows the production history and forecast per resource category from 2010 to 2030.

Even though the yearly average basket price has increased by about 50 per cent since 2016 (OPEC, 2019), the industry is still facing lower margins than previous years. This has resulted in an increased focus on production efficiency when operating a field, and thereby increasing the importance of short-term operational planning. Operational production planning is often done on a weekly basis, involving the preparation of well production plans aiming to minimize the short-term costs or maximize short-term profit. However, unexpected shifts in conditions during production can happen, and thus the operators may want to update the planning problem several times a day. This is the focus of real-time optimization (RTO). RTO aims to obtain higher throughput by improving the utilization of the capacity of the production system, for example by small adjustments to the choke openings.

Increasing computational power and measurements from sensors have been crucial for the development of RTO of petroleum production. A common petroleum production system has thousands of sensors to monitor how the system is performing. However, even with the vast amount of data generated by these sensors, it is a challenging task to optimize the day-to-day throughput of the system. The main issues are the complexity of modelling the physics, lack of key measurements such as well rates and uncertainty in the available data.

Flow assurance is the process of ensuring a safe flow of fluids from the well through the system, and is a critical task during oil production. The flow assurance issue of slugging, which is the phenomenon of gas traveling in large plugs through a pipeline, may cause overflow or damage to the production system. This will further decrease production and thus profitability. With the low margins the industry is facing, it is important to find methods to prevent and manage issues such as slugging.

This thesis focuses on RTO in the upstream sector of the petroleum production industry. The choice of scope is motivated by the increased focus in the industry of improving production efficiency, supported by the increasing number of articles and literature on the topic. Another important motivation for this research is the potentially growing amount of data available from sensors, which can be used in machine learning techniques. Machine learning techniques are already thoroughly studied in the literature. However, there is a limited amount of literature on combining machine learning with RTO, and literature on how the techniques can be directly integrated in RTO is close to non-existent. Slugging is also extensively studied, but there is very limited literature on slugging in production optimization.

In this thesis, multivariate adaptive regression splines (MARS) are used to model the relationships between operational settings, flow rates and pressures throughout the production system, based on historical measurement data. The possibilities and challenges related to integrating these models directly in linear optimization are explored.

A stochastic mixed integer linear programming (MILP) problem, taking the uncertainty in the MARS models into account, is introduced. The purpose is to provide decision support to a production engineer with the aim of maximizing oil production. The model is extended with the additional objective of minimizing the conditional value-at-risk (CVaR) with regards to gas production. The problem provides a method of studying uncertainty related to the risk of breaching the gas capacity constraint, and enables a varying level of conservativeness of the solution. An additional extension is the objective of minimizing

a given cost of slugging, providing decision support by identifying the optimal solution for a given level of risk aversion. The cost of slugging is modelled with decision tree learning, where a predictive model is used to find factors that contribute to slugging.

This thesis presents novel methods of directly integrating MARS models and classification trees in linear optimization. The formulations of a regression-based and a classification-based machine learning technique are important contributions to the new area of integrating machine learning techniques in optimization problems.

Solution Seeker, a tech start-up and spin-off developing the world's first artificial intelligence for real-time production optimization, provides the data source and guidance for the thesis. The models presented are applied to a simplified representation of an oil field on the Norwegian Continental Shelf.

In the next chapter, Chapter 2, relevant background on petroleum production is presented. A literature review with research related to this problem follows in Chapter 3, while Chapter 4 presents relevant theory for methods used in the modelling. Chapter 5 contains a problem description of the problem considered in this thesis. Further, Chapter 6 presents a stochastic optimization model and bi-objective optimization models. Chapter 7 describes how the data-driven models can be integrated in the optimization models, while Chapter 8 explains the implementation of the models and other methods. The case study is presented in Chapter 9. Furthermore, an analysis of the data-driven models is presented in Chapter 10, followed by a computational study of the optimization models in Chapter 11. Chapter 12 presents concluding remarks, and finally, future research topics are suggested in Chapter 13.

Chapter 2

Background

In this chapter, background information and concepts from petroleum production are presented. Section 2.1 explains the petroleum supply chain and production planning. The next sections include background necessary to model different parts of the production system. In Section 2.2, the components of the petroleum production system are described, while Section 2.3 explains the pressure and flow dynamics in a petroleum production system. The chapter is based on the similar chapter in the project report by Andreassen and Westby (2018).

2.1 Supply Chain and Production Planning

A supply chain in the petroleum industry consists of complex problems and decision-making, where the goal is to maximize production in a cost-effective manner (Economides et al., 2013). The petroleum production supply chain is illustrated in Figure 2.1. The chain can be divided into three different sectors: upstream, midstream and downstream. The upstream sector finds and produces crude oil and gas, and is often called the exploration and production sector. The midstream sector consists of the transportation of the petroleum products from the production sites to refineries, while the downstream sector is the marketing and distribution of the refined products to end users (Haug, 2013).

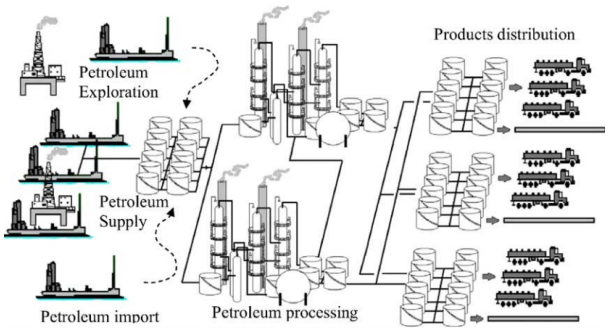

Figure 2.1: The petroleum production supply chain (Neiro and Pinto, 2004)

Operating within the different sectors requires extensive planning, and the complexity of the supply chain leads to planning problems that vary in difficulty, scope and time horizon. The planning problems can thus be divided into four different levels determined by their time horizon, as illustrated in Figure 2.2.

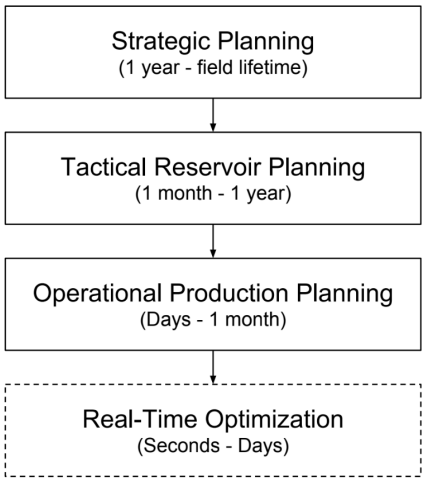

Figure 2.2: Planning levels in the petroleum industry (Morken and Sandberg, 2016)

Strategic planning ensures long-term profitability by considering aspects such as market conditions and field properties, in addition to supply chain design (Shah et al., 2010). The decisions from strategic planning will influence the other planning phases, for instance decisions regarding the level of capacity of the field (Gunnerud and Langvik, 2007). Tactical reservoir planning has a shorter time horizon. It includes decisions related to allocation of resources and maintenance planning (Ulstein et al., 2007). Operational production planning has a planning horizon of a month to days, and the planning is often done on a weekly basis. It involves preparation of production plans for each well where the goal is to minimize the short-term cost or maximize the short-term profit (Wang, 2003).

Real-time optimization (RTO) aims to obtain higher throughput by improving capacity utilization of the production system (Bieker et al., 2007a). Through a process of validating data, updating models and solving optimization problems based on the models, the production system may be operated to a near desired optimum at all times. The goal is to maximize daily production rates. RTO uses real-time measurements together with data about physical properties to operate the control settings of the well (Langvik and Dzubur, 2012). The control settings of the well are mainly the choke opening levels and gas-lift, which is injection of gas into the well. These operational adjustments are explained further in Section 2.2.

Figure 2.3 illustrates the traditional workflow setup in a production system. Historical and real-time production data, such as flow rates, pressure and temperature, is received through sensors at the production facilities. Traditional software is used for simulation and visualization of data. The data is analyzed by a production team, using both their own experience and traditional software tools, potentially including optimization software. The result of the different analyses is a set of adjustments of several input parameters that may be implemented by the well operator. The number of adjustments proposed should be constrained to a reasonable number, as the adjustments have to be manually carried out by operators on-site. In addition, it will be difficult to allocate the changes in production rates to certain wells if a lot of adjustments are done at the same time. Changes implemented at the oil field will generate new, informative data that can be analyzed.

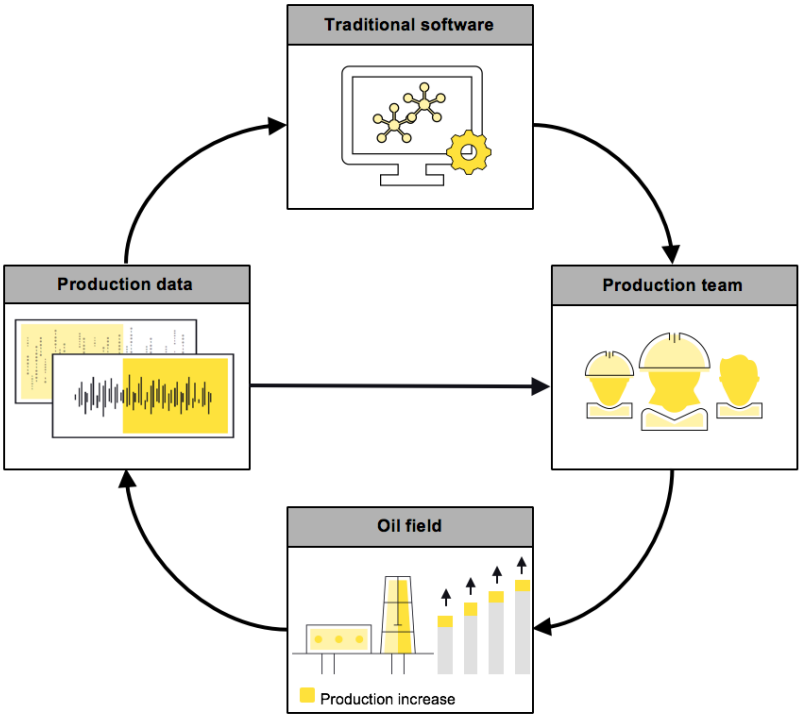

Figure 2.3: Traditional production optimization workflow setup

2.2 Components of a Petroleum Production System

An offshore petroleum production system consists of several interconnected modules. Figure 2.4 shows some of the important parts of a typical integrated petroleum production system. Starting from the bottom of the production system is the reservoir, which is a porous medium with dynamic properties that consists of an accumulation of oil, gas and water.

The production wells are structures located at the seabed with direct contact with the reservoir. The region of the well inside the reservoir is called the well bore, as illustrated in the figure. Each well consists of two pipe segments, the well tubing and the well flow line. Oil, gas and water flow in these

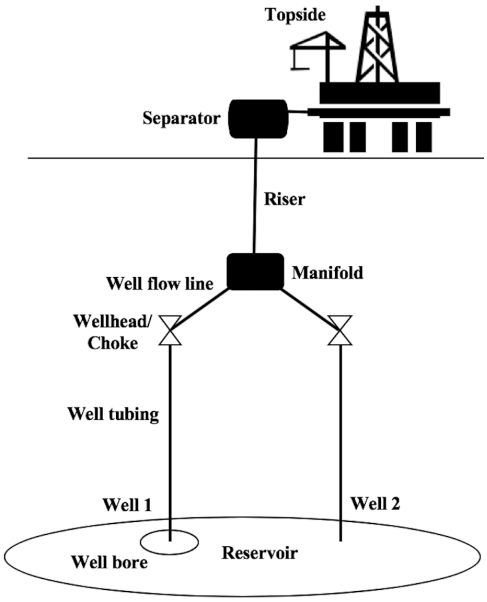

Figure 2.4: An integrated oil and gas production system

pipe segments. Between the two segments is the wellhead, illustrated with a choke. Schlumberger (2019) defines the wellhead as the system of spools, valves and assorted adapters that provide pressure control of a production well. The choke is an important part of the wellhead used to control fluid flow rate or downstream system pressure of the well. The choke level describes the percentage opening size of the wellhead valve. There is no flow through the well pipeline when the choke is closed, while increasing the choke level increases flow from the well.

A well is categorized as either a subsea well or a topside well. A subsea well has the wellhead located on the seabed, while a topside well has the wellhead located above the surface. Figure 2.4 illustrates two subsea wells. Both subsea wells and topside wells may be satellite wells, which are wells directly linked to a topside manifold.

The fluids from the reservoir are transferred by the production wells to the manifold, where the streams are mixed. A manifold is an arrangement of piping or valves designed to control, distribute and monitor the flow (Schlumberger, 2019). From the manifolds, the fluids flow through the riser, which connects the manifold to a separator. The separator separates oil, gas and water from the total fluid stream from the well. The separator is a cylindrical or spherical vessel that can be either horizontal or vertical. A two-phase separator handles only oil and gas, while a three-phase type handles water as well. From the separator, the fluids flow topside, as seen in the upper part of Figure 2.4. Topside is the upper part of the structure, which includes processing facilities and accommodation space for the crew that works on the platform.

In order to optimize the total production, it is important to monitor the flow rates of oil, gas and water. There are two main types of measurements on the flow rates: total export and multiphase flow meter (MPFM). The export measurement records the total production on the platform, and the MPFM registers the individual flow rates of the fluids flowing through a pipeline. Most modern offshore production

systems are equipped with MPFMs, however, these measurements are uncertain due to the complexity of the fluids. Furthermore, the MPFM devices are complex and expensive, and they need to be calibrated with well tests. An alternative to MPFMs is periodic single-well rate testing. This is done by routing a well to a test separator and estimating the flow rates of the individual wells.

2.3 Pressure and Flow Dynamics

Principles of pressure and flow dynamics relevant for modelling the physical relations in a petroleum production system is explained in this section. Production principles are explained in Section 2.3.1, while the issue of severe slugging is presented in Section 2.3.2.

2.3.1 Production Principles

Pressure is an important physical factor for the flow of fluids in a production system. In order for the flow to move upwards from the reservoirs via wells to the platform, there must be pressure differences in the system. The pressure difference of two points in the system is defined as a pressure drop. The pressure drop over a given part of the pipeline is mainly the result of frictional forces acting on the fluid flowing through the tube, in addition to the volume of the fluid and fluid characteristics such as density and viscosity. Other relevant factors contributing to pressure drop are the pressures at the inlet and outlet of the pipeline.

The pressure in the reservoir decreases over time as oil is extracted. When the pressure decreases to a certain level, it might not be high enough to ensure flow of the fluids. To maintain a sufficient pressure difference, gas-lift may be applied. If gas-lift is used at the platform, one can describe the current operating point as a combination of the choke levels and gas-lift injection rates. The offshore well operator can thus control the well by adjusting the choke level and gas-lift settings.

When operating a well, there are limits on the production of the well, such as the amount of gas and water that can be produced. The limit of water is due to the fact that the water has to be rinsed before being sent back into the ocean. Limits on the production can be strict, meaning that the production can never exceed the limit. However, the limit often only expresses an average that the production should lie below. With such softer limits, the operator can allow for temporary violations.

2.3.2 Flow Patterns and Severe Slugging

Flow assurance is the process of ensuring a safe and profitable flow of fluids from the well to the point of sale (Gao, 2008). This is a critical task during oil production, and becomes increasingly difficult as hydrocarbon production moves into deeper water. The increasing interest in deep water production in the upstream industry, with increasing number of deepwater wells, has thus made flow assurance a critical

part of the production system (Joshi et al., 2003).

The multiphase flow in a pipeline can take many possible shapes depending on the specific operating conditions, one being slug flow. Slug flow occurs in pipelines connecting the wells to production platforms. It is characterized by irregular flows and surges from the accumulation of gas and liquid in any cross-section of a pipeline (Sausen et al., 2012). Severe slugging is riser-based slugging where the buildup of liquid slugs is equal to or longer than one riser height (Jansen et al., 1996). Severe slugging is referred to as slugging for the rest of the thesis.

The cyclic behavior in a riser pipeline when slugging occurs is illustrated in Figure 2.5. The slugging is illustrated in four states, all taking place in the vertical riser pipeline. In the first state, liquid accumulates in the bottom of the riser, depicted in blue. When more gas and liquid enters the system, the pressure will increase and the riser will be filled with liquid, as seen in the second state. The third state illustrates the blow-out that happens as a result of the blocked gas building up in combination with high pressure. In the final and fourth state, after the blow-out, the liquid falls back and starts to build up in the bottom of the riser. After the fourth state, the cycle repeats.

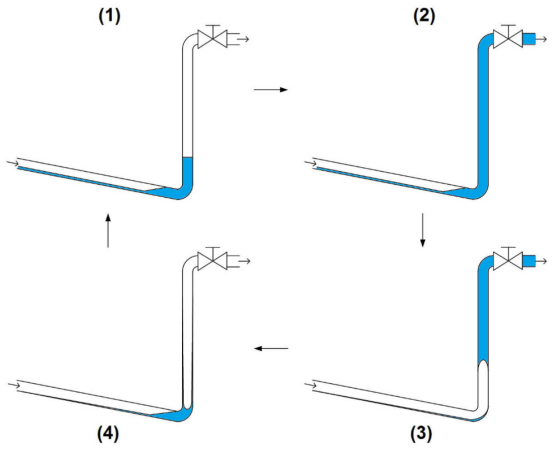

Figure 2.5: The cyclic behavior in a riser pipeline when slug occurs (Biltoft et al., 2013)

The pressure fluctuations from slug flow lead to instability in the production flow, which causes undesired consequences. The main negative impacts are liquid overflows and high pressures in the separators. Large slug blowouts result in varying flow inputs to the separator, and the liquid volume might be larger than what the separator can handle. Another issue is overload on gas compressors, which occurs if the slugs produce a much larger pressure and flow rate than the equipment is designed for. Fatigue and corrosion in the pipelines are also unwanted consequences. In addition, slug flow leads to production stop when gas surges result in increased pressure over the safety level. The consequence of this is significant financial losses, mainly due to reduction in oil production capacity (Sausen et al., 2012). To avoid the identified problems, it is important to predict the risk of slugging before production starts.

When using methods to predict and avoid slugging, oil and gas production may be reduced simultaneously. This is a big concern, and one of the reasons for the difficulty and complexity of the slugging issue. The production engineers aim to eliminate the slug flow while optimizing the production

rate (Pedersen et al., 2017). It is very difficult to balance these two objectives. Producing at a high rate is often correlated with a higher risk of slug flow occurring.

It is today common practice that the production engineers use simulators, intuition and experience to handle slugging issues. Analysis of scenarios from simulation can initiate active approaches such as choke valve control or gas-lift control, or passive approaches like avoiding certain operation areas, as is discussed in Section 3.4.

Chapter 3

Literature Review

This chapter presents a literature study of relevant topics related to optimization in the petroleum industry. The review is a continuation of the study found in Andreassen and Westby (2018). The chapter starts with an overview of optimization in the petroleum industry in Section 3.1. Further, Section 3.2 presents research where uncertain parameters are handled. Section 3.3 presents different methods used to model well parameters, while Section 3.4 presents approaches for modelling slug and methods to limit slug from occurring. Lastly, Section 3.5 presents selected references and places this thesis in a literary context.

3.1 Optimization in the Petroleum Industry

The problem of how to maximize profit with available petroleum field production capacities has been researched since the start of crude oil production from petroleum fields. Attra et al. (1961) were one of the first to address the problem with a new approach of linear programming. They conclude that this technique is effective when applied to field problems and may significantly improve the means of controlling field operations. Bodington and Baker (1990) state that such mathematical programming techniques have been developed and applied already from the 1940's, with a dominant use of linear models.

Other articles addressing linear formulations are Nygreen et al. (1998) and Wang et al. (2002a). Nygreen et al. (1998) state a mixed integer linear programming (MILP) problem, where the objective is to maximize the net present value. As the model has been in practical use for over fifteen years and is under continuous development, it has shifted from being discrete to including continuous decision variables. Wang et al. (2002a) maximize oil production while ignoring flow interactions, causing potentially inaccurate flow rates produced from each well. The issue is partially solved by a genetic algorithm.

Compared to linear formulations, non-linear optimization models are more common in both earlier and

newer literature. An example is Lang et al. (1983) optimizing production schedules, solving the problem by using an approximation of the non-linear processes by stepwise linear programs and by dynamic programming. The authors find dynamic programming to be most efficient. Ursin-Holm and Shamlou (2013) address upstream and operational planning of petroleum fields, using a special purpose branch & bound algorithm to solve the non-convex problem. As with Ursin-Holm and Shamlou (2013), Grimstad et al. (2016) formulate a mixed integer non-linear programming (MINLP) problem including spline constraints, using B-splines to approximate pressure drop correlations. The problem is solved with a Convex Envelopes for Spline Optimization (CENSO) framework. Approximation of pressure drop in well pipelines is also seen in Bieker (2007), solving the problem with a branch & bound MILP solver.

Piecewise linearization is a method designed to solve non-linear problems, seen in early literature in Fang and Lo (1996) solving a MINLP gas-lift problem. The problem is solved by using the simplex method with application of separable programming principles. Gunnerud et al. (2012) use piecewise linearization to optimize large scale production networks. By using SOS2 sets in the linearization of the well and pipe behaviour, the problem is formulated as a MILP approximation, removing the complex non-linear equations of the pipes and wells. The problem is decomposed into subproblems that are solved with column generation and a branch & price framework. A similar piecewise approximation can be found in Gunnerud and Foss (2010), where gas coning and routing of wells are included in the model. The use of piecewise linearization in a MINLP gas-allocation and routing problem is also addressed in Codas and Camponogara (2012). They propose two different formulations: the first linearizes the well production functions, while the other combines routing variables with piecewise-linear binary variables. The results show that the latter formulation is more efficient when combined with cutting-plane generation.

Gas-lift allocation is a highly studied field in production optimization. Gas-lift is together with choke adjustment input controls that production engineers adjust, and that directly affect the production. The relationship between gas injection rates and oil production was first analyzed by Mayhill et al. (1974), resulting in the gas-lift performance curve (GLPC). Alarcón et al. (2002) present a new fit to the GLPC, using the GLPC for several wells in a non-linear problem and finding the optimal gas-lift rates by sequential quadratic programming. Nishikiori et al. (1995) find the optimal gas injection rates for a group of wells by using physical restrictions and a quasi-Newton non-linear optimization technique. Another approach is to balance buoyancy improvements from additional gas in the production stream with the increased flowline back-pressure (Dutta-Roy and Kattapuram, 1997). Wang et al. (2002b) solve a similar problem and handle flow interactions among wells by simultaneously optimizing the allocation of well rates and gas-lift rates. The article builds on the MILP method from Wang et al. (2002a), and uses a sequential quadratic programming algorithm.

Several authors look more directly at the financial aspects of oil production, maximizing the profits of a system. Simmons (1972) combines a mathematical model of an oil well with an economical model for optimizing gas-lift. Input data include unit costs for injected gas and daily fixed costs. Redden et al. (1974) find each well's contribution to the total profit of the system based on distribution of available gas, calculating the distribution using performance curves for gas-input rates compared to production rates. Mora Buitrago et al. (2005) also look at production rates and gas-lift, and find that the optimal operating conditions when maximizing net present value do not always occur at maximum production or maximum

gas-lift efficiency. This is tested at a mature offshore field, using simulation to predict parameters.

With the low margins in the petroleum industry, real-time optimization (RTO) is increasingly important. Jansen et al. (2009), building on Jansen et al. (2008) where the topic is introduced, find that using real-time reservoir management strongly increases the net present value compared to a more conventional strategy. Ofonmbuk et al. (2015) further argue that introducing RTO tools to achieve real-time updates of well models will give huge monthly savings in engineering hours.

There is a limited number of examples of RTO being implemented in oil fields. One example is by Linden and Busking (2013), who use a physical model including dynamic changes in the system. This requires continuous measuring and processing of the state of the system. They conclude that modelling in RTO requires dynamic optimization, intelligent data processing and real-time data. de Boer et al. (2016) build on the article by Linden and Busking, suggesting that RTO technology is ready to be implemented for gas asset-performance evaluation.

Applying RTO also gives rise to different challenges (Bieker et al., 2007a). The optimization has to include a model and a model-updating component at a minimum to give relevant results in real-time. The models are typically steady-state models, and this has to be fitted to transient data. Handling model uncertainty is also an important challenge in RTO, as this will affect both the model and the production. The updating of constraints in the model is a main component of handling this uncertainty.

3.2 Uncertainty in Petroleum Production Optimization

Historically, research within production optimization has not taken uncertain parameters into account, resulting in most articles having a deterministic approach. However, incorporating uncertainty is important as to find optimal solutions more applicable to real-life systems, and has therefore been given increased focus in later times.

There are several articles concerning strategic and tactical decisions within the production supply chain and field development while incorporating uncertainty. Jørnsten (1992) formulates a stochastic model for petroleum field scheduling with uncertain future oil and gas prices. The model is solved using a set of heuristics. A broader approach can be found in Al-Othman et al. (2008). They develop a two-stage stochastic linear planning model with fixed recourse where the uncertainties lie in market demands and prices. The first-stage decisions are the production profiles of each type of crude oil, while the second-stage decisions are quantities produced of the different products. Dempster et al. (2000) address similar uncertainties within market demand and product prices, and develop a strategic and tactical planning problem for an oil consortium formulated as a dynamic recourse problem. A multiperiod supply, transformation and distribution scheduling problem is used as a basis for the stochastic formulation. Another example of a stochastic formulation with recourse can be found in Ribas et al. (2010). They look at an integrated oil supply chain, and formulate a two-stage model with fixed recourse for strategic planning of the supply chain. They assess uncertainty in crude oil supply, demand for final products and in products and prices in the market.

Furthermore, Tarhan et al. (2009) propose a multistage stochastic programming model for planning of offshore oil field infrastructure under uncertainty, where the uncertainties within oil flow rates and volume are gradually reduced with increased design and operating decisions. The model presents investment strategies for reducing uncertainties before making the facility investments. Goel and Grossmann (2004) also consider investment and operational planning of gas field developments, where the uncertainties lie in the gas reserves. The shape of the scenario tree depends on the investment decisions, and this decision-dependence of the scenario tree is incorporated in the stochastic model.

Within articles on operational planning in the petroleum industry, Elgsæter et al. (2010) handle measurement and model uncertainty in multiphase flow. In the article, they combine an iterative two-step approach to optimization and post-optimization feasibility assurance. Setpoints are iteratively changed until capacity constraints are reached. However, the uncertainty is only used to evaluate the solution, and is not taken into account in the optimization problem. An example where the uncertainty is formulated in the optimization problem is Bieker et al. (2007b), researching uncertain gas and water oil ratios. The uncertainties are handled by a stochastic formulation where the values are sampled by Monte Carlo simulation. The solution is the order of opening and closing wells, presented as a prioritization list. In another article regarding the operational level, Huseby and Haavardsson (2009) aim to find optimal choke settings for wells in reservoirs within the same processing facility. As with Bieker et al. (2007b), the result of the stochastic models is an order priority strategy.

Hanssen and Foss (2015) find another strategy for operating wells, using a two-stage stochastic programming formulation. They argue that, because of model uncertainty, the problem has to be modelled so that the solution is feasible with a high probability, and that a deterministic formulation may yield an unreachable solution. Chance constraints or robust formulation may solve this challenge. In another article with stochastic programming by Hanssen et al. (2015), conditional value-at-risk is used to handle uncertain capacity constraints. The solution is more conservative than the common expected value solution, but less than a robust formulation.

As articles on both RTO and articles on production optimization with uncertainty are limited, there are very few articles on uncertainty in daily production optimization (Foss et al., 2018). One contribution is made by Zhang et al. (2002) who propose a RTO formulation assessing parametric uncertainty, solved using a stochastic approach. The formulation incorporates both individual and joint chance constraints. Morken and Sandberg (2016) also handle uncertainty in input parameters. They apply stochastic programming techniques such as two-stage recourse models, and reduce uncertainty in well models by including accumulated export measurements to multiphase meter measurements.

Further, Glæserud and Syrdalen (2009) use both chance constraints and recourse models to handle uncertain well flows. Their simple recourse model penalizes capacity constraint breaches in the objective function. Krishnamoorthy et al. (2016) take a scenario-based optimization approach when evaluating daily production optimization in upstream oil and gas, seeking to find optimal decision variables, including the uncertainty in the gas-oil ratio (GOR) parameters. They argue that the scenario-based approach is favorable to worst-case optimization as it reduces the conservativeness of the solution, while still being feasible.

3.3 Modelling Well Parameters

When modelling a complex petroleum structure, it is necessary to assess the physical factors in the production system. The flow rates and pressure in the system, here referred to as well parameters, can be modelled to capture the underlying physics of the system. In addition, the number of sensors on platforms are increasing, and new methods can be developed to utilize the new data streams. A combination of physical principles and utilizing sensor data may enable modelling the system as close to real life as possible.

Models based on physical relationships can be seen in Trangenstein and Bell (1989). They use a black-oil model, which is the modelling of oil rates in a reservoir with partial differential equations. The method assumes that there is no exchange of mass between water and other phases. Fang and Lo (1996) also use a black-oil model, however transferred to simulation, when maximizing oil production. The method integrates reservoir performance with other factors such as wellbore hydraulics and gas-lift allocation.

Production rates from a reservoir are proportional to the pressure differential between the reservoir and the wellbore. The relationship between the production rate and the pressure of a well is described as an inflow performance relationship (IPR). This curve is concave in its origin, implying that an increase in production has an associated decrease in pressure (Gilbert et al., 1954). An IPR curve is a common way of modelling the production from wells, as presented in Vogel et al. (1968). They argue that the IPR curve can be applied for most solution-gas drive reservoirs, which are oil reservoirs that develop free gas on pressure depletion. Jahanbani et al. (2009) perform a method using well testing that makes the curve applicable to a wider variety of cases. Similarly, Camponogara et al. (2010) develop well performance curves (WPCs) based on a limited number of downhole pressure measurements. They argue that by conducting production tests on a test separator, the IPR curves that the WPCs are based on will be more accurate.

How pressure changes throughout the lifetime of the production system depending on flow rates, gas-lift injections and other factors is very complex, and thus difficult to model. When deriving a formulae for the reasonable injection rate under fraction and no-fraction conditions, Ling et al. (2008) find influencing factors of pressure loss on the horizontal pipelines being diameter of casing pipe, production rate and length of horizontal section. Baker et al. (1953) argue that energy required to move liquid through a pipeline is the greatest factor of high pressure drops. Decreased pipe diameters when liquid accumulates in the pipelines and increased pipeline roughness are also contributing factors. Griffith and Wallis (1961) calculate pressure drops in two-phase slug flows in vertical pipelines, using physical measures of the flows as input to a mathematical model. Building on the work by Griffith and Wallis, Orkiszewski et al. (1967) develop a method which accurately predicts pressure drops in production wells with two-phase flow based on different well conditions. The method from Griffith and Wallis is extended by including high-velocity flow range and better approximations of friction losses and flow density.

Using simulation techniques to predict flow parameters in optimization problems is a historically common method. April et al. (2003) present common approaches for optimizing simulations, such as stochastic approximation, sequential response surface methodology, random search and sample path

optimization. Simulation is often seen when modelling reservoirs. A numerical simulator for two-phase flow for reservoirs is developed by Kazemi et al. (1976). The simulator equations are solved by a finite-difference method. Ewing (1983) presents several fluid displacement techniques using different physical laws, and derives partial differential equations that act as basis for the mathematical models describing flow rates in reservoirs. Bratvedt et al. (1992) also develop a reservoir simulator, but use approximations of fractional-flow functions by piecewise linearization, together with equations for flow in a porous media, pressure equations and saturation equations. Pauchon et al. (1994) take simulation into the design of multiphase flow in pipelines. Flow characteristics are predicted using a hydrodynamic model, and the simulation models PEPITE and WELLSITE are uses for horizontal and vertical wells, respectively.

Simulation approaches are also widely used when modelling pressure, as seen in papers such as Ursin-Holm and Shamlou (2013). Ursin-Holm and Shamlou simulate the pressure drop over a pipeline using the oil, gas and water rates as input, and call the simulator directly in the optimization model. Similarly, Langvik and Dzubur (2012) simulate wellhead pressure and pressure drops using the PIPESIM Simulator from Schlumberger for multiphase flow simulation. Jiang et al. (2012) use a mathematical model when simulating pressure transients within a pipeline under non-cavitating conditions. The model has been thoroughly studied before, but Jiang et al. include a genetic algorithm to handle the problem of lack of deterministic data.

An alternative to simulation based on physical principles is to apply data-driven modelling for finding well parameters. A problem benefits from data-driven modelling if a sufficient amount of data is available, if there are no larger changes to the system during the time period of the model and if building similar models based on knowledge of underlying processes is difficult (Solomatine and Ostfeld, 2008).

Ziegel et al. (2014) use a data-driven approach where several data-driven models that represent a well's performance are integrated. Available real-time data, including bottomhole pressure, wellhead pressure, choke changes and gas-lift rate, are used in the models. This enables the asset investigated to utilize the sensor data to manage and optimize the wells. Li and Chan (2010) use decision tree learning to predict petroleum production. They develop several different models, each involving different combinations of parameters affecting oil production. Furthermore, a data-driven approach for pressure modelling is applied by Teixeira et al. (2014). They extend the traditional method of finding soft sensors for estimating downhole pressure to building neural models on historic data collected from gas-lift oil wells.

Using neural networks has often been applied to build parameter models, as seen in Shokir et al. (2017). In the article, an artificial neural network model is developed, yielding accurate predictions of bottomhole pressure and well production rates. The accuracy stems from including gas injection rates to predict well parameters. Another application is seen in Stoisits et al. (1999), based on the earlier work found in Stoisits et al. (1992). They use neural networks with backpropagation as a replacement for simulation methods applied in production optimization, modelling wells and surface lines. The neural nets are found to reduce the computational time of the output of the model when responding to changes in input variables. Neural network models using backpropagation are also found in Gharbi et al. (1997), where PVT properties of crude oil are predicted. A review of the usage of neural networks in the petroleum

production industry can be found in Saputelli et al. (2002). They find that artificial neural networks have been used to predict pressure and production, PVT properties, bottomhole pressure and flow regimes, among others. An alternative to neural networks, a functional network, is used in Oloso et al. (2017) to predict the bubblepoint pressure of black oil. The functional network is combined with clustering of the input data, producing a hybrid functional network.

Another data-driven approach is using ensemble machine learning methods, which is the combination of different machine learning techniques. Anifowose et al. (2017) present a thorough study of applications of such methods in the petroleum industry, with several examples of the method being used to model well parameters. One example is Chen and Lin (2006), predicting reservoir characteristics by using a committee neural network. Using different models is also found in Panja et al. (2018), providing two models based on least square support vector machine and neural networks to predict hydrocarbon production from wells.

Both for physical and data-driven methods, there is very limited literature on using choke parameters as input, and combining this with production data. Some examples seen are Goh et al. (2007), Bieker (2007) and Morken and Sandberg (2016). In Bieker (2007), the choke settings limit the maximum production rates and pressures at each well. Morken and Sandberg (2016) use regression techniques to model the relationships between production system settings, such as the choke valve, and production rates. They use certain operation modes to handle the non-linear relationships between the production settings and rates.

3.4 Slug Modelling and Management

Petroleum production systems have big challenges related to slugging. As described in Section 2.3.2, slug flow leads to severe consequences for production, and the most severe slugs occur in long vertical risers. As presented in the following section, the issue of slugging has been thoroughly researched, but there is limited literature on slugging in production optimization, nor in a data-driven context.

Mathematical modelling of slug has been used to investigate how slugging is induced and to support anti-slug control. Slug modelling is a challenging task that has been investigated for many years, as reviewed by Pedersen et al. (2015). The early studies in the 1980s focused on simulations of steady-state movement of the flow (Taitel et al., 1980). Viggiani et al. (1988) develop a mathematical model using pressures, length and heights of the pipeline sections and liquid levels. Taitel (1986) also develop a mathematical model, based on the work by Schmidt et al. (1980), where the model predicts conditions where slugging is not possible.

In the 1990s, simple transient models emerged as the main focus area. The transient model for pipeline-riser systems presented by Sarica and Shoham (1991) is based on one-dimensional gravity-dominant flow in both the pipeline and riser. The model predicts slugging better than other models at that time. The limitation of these earlier models is that they were developed to simulate the flow, and not developed for anti-slug control. Software simulators like Schlumbergers OLGA and

Kongsbergs LedaFlow have been developed in recent years. They have been used to assess different low-dimensional control-oriented models developed to describe slug (Pedersen et al., 2017). See Bendiksen et al. (1991), Tang et al. (2006) and Danielson et al. (2011) for more detailed descriptions of the simulators.

Slug elimination or control is critical due to the negative consequences of slugging. There are many ways of preventing slug, and it is an area that has been investigated for many years. Pedersen et al. (2015) classify methods of slug elimination into active and passive approaches. With active approaches, the slug elimination is realized by an automatic feedback control strategy based on a given process, and manipulated actuators are involved. Slug elimination with passive approaches is conducted by a dedicated system or process design, and no actuators are involved.

Active slug control approaches manipulate the actuators based on feedback mechanisms, and are subject to signals from pressure, temperature and flow transmitters (Pedersen et al., 2017). The most researched approach is using choke valve control, which is the choking of a controllable valve in the riser as to reduce flow. This was assessed in the early 1970s by Yocum et al. (1973). They consider the efficiency of choking as a slug mitigation method, and find that the method has to be used together with a sufficient riser system design. Building on this work, Schmidt et al. (1979) state that severe flow can be eliminated with careful choke control, while at the same time giving steady production rates. Another known active approach is gas-lift control. Asheim et al. (1988) argue that this is an effective approach, but it may require very large amounts of gas to see an effect on the slugging. Choke valve control and gas-lift control may be combined into a multi-input multi-output (MIMO) control system. Pagano et al. (2008) present such a system where the gas-lift control stabilizes the gas-lift injections, and the choke adjusts the topside pressure.

Yocum et al. (1973) present the fundamental basis for passive approaches, which includes reducing the incoming line diameter near the riser to get a more stable flow regime, creating multiple risers and using a liquid remix device, which mixes the content in the riser to avoid accumulation. An example of the latter is the usage of flow conditioners, as the Wavy pipe developed and presented by Xing et al. (2013). A flow conditioner is a device installed in the pipeline to change the current flow pattern. The Wavy pipe artificially creates small slugs to reduce the possibility of severe slugging. Other passive approaches include using slug catchers, and the compression and subsequent separation of gas upstream in the riser base, explored by Tengesdal et al. (2002).

Available literature on slugging in production optimization is limited, and examples of slugging used directly in a production optimization problem are not found. In Willersrud et al. (2013), slugging is used as an input disturbance variable in a non-linear model with a predictive control formulation, yielding pressure control in the system. Short-time production optimization is then performed in combination with the predictive control. The authors note that the pressure control may be used without any optimization, and still increase oil production.

3.5 Summary of Literature Review

The literature review conducted starts by introducing early optimization in the petroleum industry, and presents the development of both models and solution methods in the field. The literature shows that the area of RTO is increasingly important, but has potential for additional modelling components and increased complexity. The review continues by introducing the topic of handling uncertainty in production optimization, showing that this is a field, as with RTO, that has been given more focus in later years. The review further discusses how well parameters are modelled. The traditional methods of simulating parameters and using physical relationships may give inaccurate results, implying that data-driven modelling is a potentially better alternative for some cases. Lastly, literature on slugging is presented, introducing methods of handling the problem.

When modelling a complex petroleum structure, it is necessary to assess the physical factors in the production system. The flow rates and pressure in the system, here referred to as well parameters, can be modelled to capture the underlying physics of the system. In addition, the number of sensors on platforms are increasing, and new methods can be developed to utilize the new data streams. A combination of physical principles and utilizing sensor data may enable modelling the system as close to real life as possible.

A selection of the articles from Section 3.1 and Section 3.2 is presented in Table 3.1. This thesis, Andreassen and Westby (2019), is included as the first article for comparison. The columns describe different characteristics of the models in the selected articles. In the first two columns from the left, the authors, article year and objective function being maximized for each article are presented. Next, the problem type of the article, whether or not the optimization problem is real-time and the solution method used to solve the problem are shown. Lastly, the decision variables used in the articles are presented.

The selected references in Table 3.1 illustrate the common objectives, problems and solution methods in production optimization. They illustrate the fact that production optimization problems are mainly non-linear in nature, requiring advanced solution methods. Compared to the references, this thesis uses a combination of decision variables, choke and pressure, that is rarely seen in previous literature. The study of including slugging in optimization has also rarely been conducted, even though the issue of slugging is common and assessed before. Though not included in the Table 3.1, the articles assessed in this literature review use several different techniques to model the physics of the system, as seen in Section 3.3. While several of the articles use various machine learning techniques, none directly integrate machine learning techniques into the optimization with mathematical formulations, which is the topic of this thesis. Additionally, the usage of multivariate adaptive regression splines in production optimization has not been seen before.

Table 3.1: Selected references on production optimization

					Decision variable		
Article	**Objective function**	**Problem type**	**RTO**	**Solution method**	**Choke**	**Gas-lift**	**Pressure**
Andreassen and Westby (2019)	Oil production	MILP	Yes	Branch & bound	Yes	No	Yes
Krishnamoorthy et al. (2016)	Profits	Non-linear	Yes	Scenario-based approach	No	Yes	Yes
Morken and Sandberg (2016)	Oil production	MILP	Yes	Branch & bound	Yes	No	No
Grimstad et al. (2016)	Oil production	MINLP	No	CENSO	No	Yes	Yes
Linden and Busking (2013)	Profits	Non-linear	Yes	Iterative partial derivation	Yes	No	Yes
Ursin-Holm and Shamlou (2013)	Oil production	MINLP	No	Branch & bound	No	Yes	Yes
Gunnerud et al. (2012)	Oil production	Piecewise linear MILP	No	Branch & price	No	No	Yes
Codas and Camponogora (2010)	Oil production	Piecewise linear MILP	No	Cut & price	No	Yes	No
Bieker (2007)	Oil production	Piecewise linear MILP	No	Branch & bound	Yes	Yes	Yes
Mora Buitrago et al. (2005)	Net present value	Non-linear	No	Simulation	No	Yes	Yes
Wang et al. (2002b)	Oil production	MILP	No	Sequential quadratic programming	No	Yes	Yes
Alarcón et al. (2002)	Oil production	Non-linear	No	Sequential quadratic programming	No	Yes	No
Nygreen et al. (1998)	Net present value	MILP	No	Branch & bound	No	No	No
Dutta-Roy and Kattapuram (1997)	Oil production	Non-linear	No	Sequential quadratic programming	No	Yes	No
Fang and Lo (1996)	Oil production	Piecewise linear MILP	No	Separable programming	No	Yes	No
Nishikiori et al. (1995)	Oil production	Non-linear	No	Quasi-Newton	No	Yes	No

Chapter 4

Theory

Theory relevant for understanding the models in this thesis is presented in this chapter. In Section 4.1, relevant machine learning concepts and methods are explained. An introduction to selected optimization concepts is given in Section 4.2.

4.1 Supervised Machine Learning

Supervised learning is the machine learning task of mapping a function from the input to the output. Supervised learning problems can be categorized into regression and classification problems. The difference between the two methods is the output variable, which is numerical for regression problems and categorical for classification problems.

Relevant concepts related to supervised machine learning is presented in this section. Model selection and evaluation are discussed in Section 4.1.1. The concept of multivariate adaptive regression splines (MARS) is explained in Section 4.1.2, and classification trees (CT) are explained in Section 4.1.3.

4.1.1 Model Selection and Evaluation

There are many different methods and approaches for supervised learning. The choice of method should be done based on the given problem characteristics. Criteria for selecting a suitable model for the problem at hand are discussed in this section.

Bias-Variance Trade-off

Prediction errors in a prediction model can be decomposed into two subcomponents, namely the error due to bias and the error due to variance. The ability of a prediction model to minimize bias and variance represents a trade-off. Over- and underfitting a prediction model can be avoided by understanding these

two main types of error (Fortmann-Roe, 2012).

Denoting the variable to be predicted Y and the predictor variables X, it can be assumed that there is a relationship $Y = f(x) + \epsilon$, where ϵ is an error term that is normally distributed with a mean of zero, $\epsilon \sim (0, \sigma_\epsilon)$. Let $\hat{f}(x)$ denote the model that approximates $f(X)$. The squared-error loss of $\hat{f}(x)$ can be expressed as:

$$Err(x) = E[(Y - \hat{f}(x))^2] \tag{4.1}$$

where the error can be further decomposed into the bias and variance components:

$$Err(x) = (E[\hat{f}(x)] - f(x))^2 + E[\hat{f}(x) - E[\hat{f}(x)]^2] + {\sigma_\epsilon}^2 \tag{4.2}$$

$$Err(x) = Bias^2 + Variance + Irreducible\ Error \tag{4.3}$$

The terms in Equation (4.2) represent the squared of the bias, the variance and the irreducible error, respectively. The irreducible error is the noise term that cannot be reduced by any model. The bias is the difference between the expected prediction of the model and the correct value, while the variance is the expected squared deviation of the model around its mean. The bias and variance terms can in theory be reduced to zero, given the true model and an infinite data source to calibrate the model. However, in real-world applications, the assumption of an infinite data source is unreasonable, and there is a trade-off between minimizing the bias and variance of the model (Hastie et al., 2009).

In general, if model complexity is increased, the bias is often reduced and the variance is increased. Oppositely, the bias is often increased and the variance is reduced when the model complexity is reduced. The goal of a prediction model should be to decrease the overall error. Thus, a level of complexity where an increase in bias corresponds to an equivalent reduction in variance is optimal. This level is illustrated in Figure 4.1, and can be expressed as:

$$\frac{dBias}{dComplexity} = -\frac{dVariance}{dComplexity} \tag{4.4}$$

Underfitted models, associated with high bias, do not capture trends in the dataset well and model the training data poorly. On the other hand, overfitted models perform well on the training data, but do not generalize well to new data. The models follow the noise in the data too closely, and pick up patterns in the training data that are caused by random chance rather than by true properties. With a higher level of complexity, the model is overfitted, and oppositely underfitted with a lower level of complexity. In practice, it is difficult to find the optimal model complexity (Hastie et al., 2009). To search for the right level of complexity, accurate measures of prediction error should be used. Different levels of complexity can be explored, and the level that minimizes the error measures should chosen.

Uncertainty Quantification

Many sources of uncertainty exist, and they are in general categorized as either aleatory and epistemic. The uncertainties are categorized as aleatory if the uncertainty cannot be reduced by the modeler, and

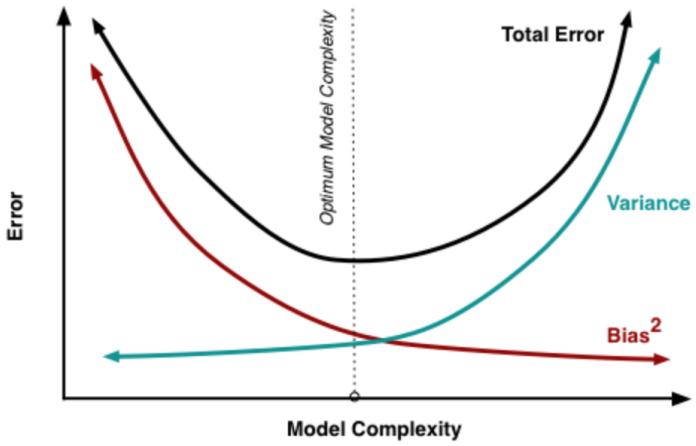

Figure 4.1: Bias-variance trade-off (Fortmann-Roe, 2012)

oppositely epistemic if there is a possibility of reducing uncertainty by gathering more data or refining the models (Der Kiureghian and Ditlevsen, 2009). Quantification of uncertainty works toward reducing epistemic uncertainty to aleatoric uncertainty.

The prediction uncertainty is based on epistemic and aleatoric uncertainty, and is expressed as

$$Var_{pred} = Var_{epistemic} + Var_{aleatoric} \tag{4.5}$$

Epistemic uncertainty, also known as systematic or model uncertainty, occurs due to effects that in principle could be known, thus because of lack of knowledge. An example would be inaccurate measurements because certain effects are neglected by the model. In order to evaluate epistemic uncertainties, better knowledge of the process or system should be attained.

Aleatoric uncertainty is also referred to as statistical uncertainty. The uncertainty represents the effect of unknown parameters that change every time the system or process executed. That is, it captures inherent noise in the observations.

Training, Validation and Testing

A supervised machine learning model should be able to generalize to input samples it has never seen before. In order to get a model with sufficient accuracy, the model must be exposed to a variety of input samples.

The data is divided into a training, validation and test dataset. The training dataset is used to train the machine learning model. Further, the validation set is used to provide an evaluation of the trained model that is not biased, in addition to tune the hyperparameters of the model. Lastly, the test dataset is used to provide an evaluation of the final model fit on the training dataset that is not biased. It is important that the test dataset is only used after a model has been fully trained.

The choice of dataset split ratio mainly depends on the number of samples in the data and the type of model that is trained. A visualization of the three splits is presented in Figure 4.2. It is common practice to split the dataset into training and test datasets, and using cross-validation on subsets of the training dataset. The concept of cross-validation is explained in the next section.

Figure 4.2: Visualization of train, validation and test dataset

Cross-Validation

The explanation of cross-validation is based on the work by James et al. (2013). Cross-validation is a method of finding the test error of a model, i.e. the average error when predicting a measurement that was not used in the training set. The validation set approach randomly divides the dataset into a training set and a validation set. Fitting the model on the training set, the model predicts responses for all observations in the validation set. The validation set error rate, often measured by MSE, estimates the test error rate.

The validation set approach has certain drawbacks. The estimate of the test error rate may vary, depending on how the data is split into training and validation. In addition, as only a subset of the data is used for fitting the model, the approach tends to overestimate the test error rate for the same model fit on the whole data.

Based on the two identified problems of the validation set approach, cross-validation tries to address these. Leave-one-out cross-validation (LOOCV) is the most similar method to the validation set approach. This method splits the dataset into two parts. However, the validation set consists of only one observation, and the remaining observations make up the training set. The resulting prediction $\hat{y}_1$ for the observation x_1 in the validation set yields an estimate of the test error rate, $MSE_1 = (y_1 - \hat{y}_1)^2$, where y_1 is the actual observed output. Denote the number of observations in the dataset n. Repeating the process n - 1 times for the other observations in the dataset, the method produces an error rate from the average of all error estimates:

$$CV_{(n)} = \frac{1}{n} \sum_{i=1}^{n} MSE_i \tag{4.6}$$

An alternative to using LOOCV is k-fold cross-validation. In this method, the dataset is divided into k different groups of similar sizes. The first group, or fold, is used as a validation set, and the other k - 1 folds are used to fit the model. MSE_1 is computed for the first fold, and the process is repeated k times

for each fold. The resulting error rate is an average of the k MSE-values:

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} MSE_i \tag{4.7}$$

Typical values for k-fold is setting $k = 5$ or $k = 10$. Setting $k = n$ gives LOOCV.

Evaluation Metrics

There exist many different approaches to measuring the error of a fitted function to the real function. Different metrics provide different insight, and there is not a single best metric. Evaluation of regression and classification models require different metrics, since the output variable of the former is numerical, while the output variable of the latter is categorical.

Evaluation of regression models is commonly performed using measures such as mean squared error (MSE), as used in cross-validation, and the coefficient of determination, R^2. MSE is used when quantifying how close the predicted response value for an observation is to the true value (James et al., 2013). MSE is given by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.8}$$

The MSE value is small when the predicted and the true values are close, and large when some of the predicted values differ substantially from the true values.

R^2 describes the proportion of variance in a response variable that is explained by using a predictor variable, taking a value between zero and one. A score close to one indicates that a large portion of the variance in the response variable is explained, while a score of zero means that no variance is explained. R^2 is given as

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{4.9}$$

where $\bar{y}_i$ is the mean of the target variables.

Another possible evaluation metric is Root Mean Squared Error (RMSE). RMSE is the squared of MSE,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4.10}$$

The standard deviation (SD) is a measure of the spread of the values. A low SD implies that the numbers are close to the mean of the values, while a high SD implies that the deviation from the mean is large. SD is calculated by the following formula:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}} \tag{4.11}$$

Classification models must be evaluated by other metrics than regression models since the output is categorical. There are many ways to evaluate the performance of a classification model. A common method is to calculate the classification accuracy, which is the percentage of correctly classified examples out of all predictions made:

$$Accuracy = \frac{Correct\ predictions}{Total\ predictions} * 100\% \tag{4.12}$$

4.1.2 Multivariate Adaptive Regression Splines

Multivariate adaptive regression splines (MARS) is an approach to multivariate non-parametric regression. The approach was introduced in 1991 by Jerome H. Friedman (Friedman et al., 1991).

The MARS model is a weighted sum of basis functions $B_i(x)$, where c_i is a constant coefficient:

$$\hat{f}(x) = \sum_{i=1}^{k} c_i B_i(x) \tag{4.13}$$

A basis function $B_i(x)$ can be either a constant, a hinge function or a product of two or more hinge functions. A hinge function has the form $max(0, x - k)$ or $max(0, k - x)$, where x is an input variable and k is a constant, or knot. The MARS algorithm selects variables and values for the knots of the hinge functions, and decides how the hinge functions can be combined. When the basis function is a product of two or more hinge functions, the interactions between different variables can be modelled.

There are two phases in the MARS building process: the forward and the backward pass.

The forward pass: Basis functions in pairs are repeatedly added to the MARS model. A greedy algorithm that chooses the pair of basis functions that will contribute to the largest reduction of the sum-of-squares residual error, RSS, is applied in each step. The process of adding terms will continue until the maximum number of terms specified by a user is reached or the change in residual error is too small. The forward pass will usually build an overfitted model which will not generalize well to new data.

The backward pass: The backward pass prunes the model in order to improve the generalization ability of the model. In each step, it removes the term that is least effective. The generalized cross-validation (GCV) score is evaluated to choose the best model subset in the backward pass. The GCV trades off goodness-of-fit against model complexity. The formula for GCV is given by

$$GCV = \frac{RSS}{(1 - \frac{C}{N})^2} \tag{4.14}$$

with

$$RSS = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{4.15}$$

and

$$C = 1 + cd \tag{4.16}$$

Here, N is the number of observations in the dataset and C is the effective number of parameters. For the parameter C, c is the penalty for adding a new basis function and d is the number of independent basis functions.

4.1.3 Classification Trees

Classification and Regression Trees (CART) are used to make predictions by stratifying and segmenting the predictor space into simple regions. The splitting rules for a given prediction task can be summarized in a tree structure, which is either a classification tree or a regression tree. This section focuses on classification trees (CT), and is based on the work by James et al. (2013). Tree-based methods can be simpler to interpret than other machine learning approaches, and are easily explained graphically. They may however be non-robust, so that they are sensitive to change in data.

Building Classification Trees

The process of building a classification tree can be divided into two steps:

- Step 1: The predictor space is divided into J distinct and non-overlapping regions, R_1, R_2, ..., R_J. The predictor space is the set of possible values for X_1, X_2, ..., X_p, i.e. the set of possible values for the predictors. The predictors can be both quantitative or qualitative.
- Step 2: The same prediction is made for every observation that falls into the same region R_J. The prediction is the most commonly occurring class of training observations in the region. A region R_J is a leaf node in the tree.

The regions may in theory have any shape suited for the data. However, a more structured way to divide the predictor space is by using boxes. Boxes R_1, , ..., R_J are found by minimizing the classification error rate, which is the fraction of the training observations in that region that do not belong to the most common class:

$$E = 1 - \max_k(\hat{p}_{mk}) \tag{4.17}$$

Here, $\hat{p}_{mk}$ equals the proportion of training observations in the mth region that are from the kth class.

However, the classification error rate may not be sufficiently sensitive, and two other measures are therefore preferred: the Gini index and entropy.

The Gini index is defined as

$$G = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \hat{p}_{mk}) \tag{4.18}$$

The index measures the total variance across the K different classes. As a small value is an indication of a node containing predominantly observations from a single class, the index is also referred to as a measure of node purity.

The entropy measure is given by

$$D = -\sum_{k=1}^{K} \hat{p}_{mk} log(\hat{p}_{mk}) \tag{4.19}$$

The entropy will be close to zero if the $\hat{p}_{mk}$ values are all near zero or near one. It will therefore take on a small value if the mth node is pure. Numerically, the entropy is similar to the Gini index.

As considering every possible partition space is computationally infeasible, a top-down, greedy approach called recursive binary splitting is used. The approach starts at the top of the tree, in the root node, where all observations belong to a single region. Moving down the tree, the predictor space is split such that each split is the best in regards to minimizing either the Gini index or the entropy at that particular step. When building the tree, these two measures are often chosen as they are more sensitive to node purity than the classification error rate. A split is indicated by adding two new branches further down the tree.

From the top node, the first split is found by splitting the predictor space into two regions, $\{X|X_j < s\}$ and $\{X|X_j \geq s\}$, leading to the greatest possible reduction in the Gini index or entropy. X_j is the predictor and s is a cutpoint. The process is then repeated, but this time splitting one of the regions into two new regions, again minimizing the chosen measure. The process stops by a stopping criteria, for example that the amount of observations in each region should be below a given number.

Pruning Classification Trees

Using recursive binary splitting to find the classification tree may lead to good predictions on the training set. However, it is likely to overfit the data, especially if the resulting tree is too complex. To avoid overfitting, a very large tree T_0 is built, and a subtree of T_0 is obtained by pruning. The subtree should lead to a low test error rate, measured by cross-validation or the validation set approach. As these are very complex with a large number of possible subtrees, the method of cost complexity pruning is used instead.

Cost complexity pruning considers a sequence of trees, each sequence indexed by the non-negative tuning parameter α. For each value of α, there is a corresponding subtree $T \subset T_0$, such that the subtree minimizes

$$\sum_{m=1}^{|T|} \sum_{i:x_i \in \mathcal{R}_m} (y_i - \hat{y}_{R_m})^2 + \alpha|T| \tag{4.20}$$

Here, $|T|$ equals the number of terminal nodes, R_m is the subset of the predictor space for the mth terminal node and $\hat{y}_{R_m}$ is the mean of the training observations in R_m.

As α increases from zero, branches are pruned from the tree, and there is a trade-off between the subtree's fit to the training data and the complexity. An α is chosen by using k-fold cross-validation or a validation

set, and the corresponding subtree is obtained from the α value.

4.2 Optimization

A selection of relevant concepts related to stochastic optimization and multi-objective optimization are presented in this section. First, Section 4.2.1 explains stochastic optimization, and Section 4.2.2 presents how stochastic solutions can be evaluated. Further, Section 4.2.3 explains how conditional value-at-risk can be used to account for uncertainty. Lastly, the concept of multi-objective optimization is presented in Section 4.2.4.

4.2.1 Stochastic Optimization

Stochastic optimization refers to optimization when randomness is present. That is, random variables are generated and used. Stochastic programming models decision problems in order to obtain solutions that perform well under uncertainty (King and Wallace, 2012).

A recourse model allows decisions to be made at different stages in time, and thus handles uncertainty in the model parameters. At a later stage, there is an opportunity to adjust decisions that are made in a previous stage. The adjustments are done when new information becomes available, as a response to the outcome of one or more random variables. Recourse models consist of at least two stages. The two-stage recourse problem is an example of stochastic optimization where the decisions are made in two distinct stages.

Define two decision variables x and y, where decision x is made at the first stage, and y is the recourse action. In addition, c is the cost vector for the first-stage decisions. Denoting a stochastic variable $\widetilde{\xi}$, and ξ a realization of this variable, the mathematical formulation of the two-stage recourse problem is

$$min\ \phi(x) = c^T x + Q(x) \tag{4.21}$$

subject to

$$Ax = b \tag{4.22}$$

$$x \geq 0 \tag{4.23}$$

where the recourse function $Q(x)$ is the expected value of the second-stage problem:

$$Q(x) = E[Q(x, \widetilde{\xi})] \tag{4.24}$$

The distribution of $\widetilde{\xi}$ can be continuous, which implies that one must integrate over $\widetilde{\xi}$ when calculating $Q(x)$. That can be a difficult task, and it is thus often assumed that $\widetilde{\xi}$ is a discrete random variable with a finite set of realizations.

Letting q_0 be a cost vector in the second-stage decision, the second-stage problem $Q(x, \xi)$ is defined as

$$min\ q_0^T(\xi)\ y \tag{4.25}$$

subject to

$$W(\xi)y = \xi - T(\xi)x \tag{4.26}$$

$$y \geq 0 \tag{4.27}$$

In the formulation presented, no assumptions are made about the structure of the parameters of the second-stage problem. All of the parameters can potentially be uncertain, and this situation is referred to as general recourse. In the case of simple recourse, the recourse action is uniquely defined by the value of the first-stage variables and the realization of the stochastic variables. Fixed recourse is the case where there are no uncertain parameters in the constraints. For complete recourse, there exist second-stage variables that satisfy the constraints in the second-stage problem for all x. When the recourse problem has certain properties, there are computational advantages that can be exploited (Higle, 2005).

4.2.2 Evaluation of Stochastic Solution

It is important to evaluate the stability of the stochastic solution to ensure that the solution is representative. In-sample and out-of-sample stability, and relevant measures of the value of incorporating uncertainty, are explained in the following.

In-Sample and Out-of-Sample Stability

Stochastic programming models cannot be solved when the random variables are expressed with a continuous distribution, except for some trivial cases. A discretization of the probability distribution of the random variables is therefore necessary. The number of outcomes from the discrete distributions must be constrained due to limited available computing power and to limit the complexity of the decision model. The discretization is often called a scenario tree, depicted in Figure 4.3 for a two-stage problem (Kaut and Wallace, 2003).

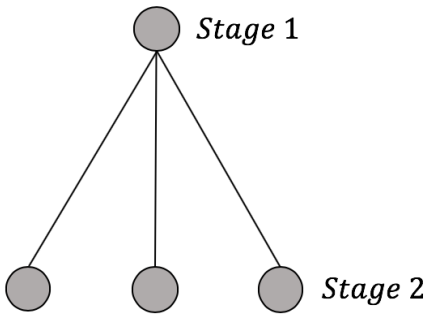

Figure 4.3: Scenario tree for a two-stage problem

The discretization should be such that it gives results that are as similar as possible to the results from using the true, continuous distribution. It is desirable that the discrete scenarios capture the important aspects of the true distribution. It is however very difficult to verify whether or not this is the case.

Scenario generation techniques are often based on drawing samples from an underlying distribution. Thus, the scenario tree is different every time a scenario is generated, implying that the solution found may also vary for different scenario trees. This is not a problem if the objective values corresponding to the solutions are similar. If the objective values however vary with the given scenario tree, the solution found cannot be trusted to be optimal. The quality of a solution is then dependent on the scenario generated. In-sample and out-of-sample stability are often used to measure the quality of the generated scenarios. The theory related to stability in the rest of the section is obtained from King and Wallace (2012).

In-sample stability ensures that roughly the same objective value is achieved for each generated scenario tree. Assume that a scenario generation method is run several times on the same data. Denote two resulting scenario trees $\mathcal{T}_i$ and $\mathcal{T}_j$, respectively, where the optimal solutions are denoted $\hat{x}_i$ and $\hat{x}_j$. In-sample stability is obtained if the optimal objective function values are approximately the same. That is, for the two scenario trees $\mathcal{T}_i$ and $\mathcal{T}_j$, they have in-sample stability if

$$f(\hat{x}_i, \mathcal{T}_i) \approx f(\hat{x}_j, \mathcal{T}_j) \tag{4.28}$$

Out-of-sample stability means that the solutions from different scenario trees have approximately the same objective function value as if the true distributions were used on the same solutions. Let $\boldsymbol{\xi}$ denote the true distribution of the random variables in the given problem. For the two scenario trees $\mathcal{T}_i$ and $\mathcal{T}_j$, they have out-of-sample stability if

$$f(\hat{x}_i, \boldsymbol{\xi}) \approx f(\hat{x}_j, \boldsymbol{\xi}) \tag{4.29}$$

It should be noted that evaluating a solution using the true random distribution is often difficult. If the true distribution is discrete but with too many scenarios, out-of-sample stability may be found by calculating $f(\hat{x}_i, \boldsymbol{\xi})$. However, if the true distribution is continuous, the distribution has to be sampled. Other approaches are using a simulation model or performing the weaker out-of-sample test

$$f(\hat{x}_i, \mathcal{T}_j) \approx f(\hat{x}_j, \mathcal{T}_i) \tag{4.30}$$

The weaker out-of-sample test states that when evaluating the solutions of a model from one scenario tree in another scenario tree, the objective values should be approximately the same.

Expected Value of Perfect Information and Value of the Stochastic Solution

Solving a stochastic optimization problem is in general more difficult than solving its deterministic counterpart. However, stochastic formulations can capture aspects of a problem that deterministic models cannot. Measures of the value of incorporating uncertainty can be defined in order to compare the

different solutions.

The expected value of perfect information (EVPI) is the difference between the objective function values from solving the stochastic problem, compared to solving the same problem given perfect information. Thus, EVPI represents the amount a decision maker would be willing to pay to get complete and accurate information. Mathematically, EVPI is the difference in objective value between the recourse problem (RP), expressed as a minimization problem in Section 4.2.1, and the wait-and-see solution (WS) (Madansky, 1960). In order to calculate the WS, the model is first solved for each scenario with perfect information about the distribution of the random variables. The expectation of the objective value of these solutions is referred to as the WS. For a minimization problem, the EVPI is given by

$$EVPI = RP - WS \tag{4.31}$$

The value of the stochastic solution (VSS) is a measure of the difference in objective value when planning with uncertainty incorporated, compared to using expected values. The VSS therefore measures the expected loss when using the deterministic solution. Thus, the VSS represents how much a decision maker is willing to pay to incorporate uncertainty in the model. The VSS is the difference between the expectation of the expected value model (EEV) and the RP, given by

$$VSS = EEV - RP \tag{4.32}$$

The EEV is found by replacing the uncertain parameters with their expected value, and finding the expectation in objective function value of the solution in all scenarios.

For a conventional stochastic minimization problem with recourse, the EVPI and the VSS are always non-negative. This gives that for a minimization problem, the WS is smaller than the RP, and the RP is smaller than the EEV.

4.2.3 Value-at-Risk and Conditional Value-at-Risk

Value-at-risk (VaR) and conditional value-at-risk (CVaR) are widely used performance measures of the maximum loss with a specified confidence level.

Denoting a random variable X describing loss and its cumulative distribution function $F_X(z) = P\{X \leq z\}$, VaR can be expressed as

$$VaR_\alpha(X) = min\{z|F_X(z) \geq \alpha\} \tag{4.33}$$

for a confidence level $\alpha \in [0, 1]$. The expression defines VaR as the highest loss that can occur with a given probability α.

Methods of calculating VaR use linear approximations of risks and assume the joint normal distribution of underlying parameters. The measure can be controlled in a mathematical programming problem with

VaR constraints. However, since VaR is non-convex, it is very difficult to solve with formal optimization techniques (Uryasev, 2000).

CVaR is an alternative measure of losses that can be more easily applied in optimization. CVaR is also known as mean excess loss, mean shortfall, or tail VaR. Rockafellar et al. (2000) introduced the approach of optimizing a portfolio to reduce the risk of high losses. With respect to a specified probability level α, the α-VaR of a portfolio is the lowest amount β such that the loss will not exceed β with probability α, whereas the α-CVaR is the conditional expectation of losses above that amount. This ensures that the α-VaR never exceeds α-CVaR. Thus, portfolios with low CVaR must also have low VaR. For $\alpha \in [0, 1)$, CVaR is expressed as

$$CVaR_\alpha(X) = \int_{-\infty}^{+\infty} z dF_x^\alpha(z) \tag{4.34}$$

where

$$F_X^\alpha(z) = \begin{cases} 0 & z < VaR_\alpha(X) \\ \frac{F_X(z)-\alpha}{1-\alpha} & z \geq VaR_\alpha(X) \end{cases} \tag{4.35}$$

VaR and CVaR are represented in Figure 4.4 where the frequency of different loss values are represented in a histogram.

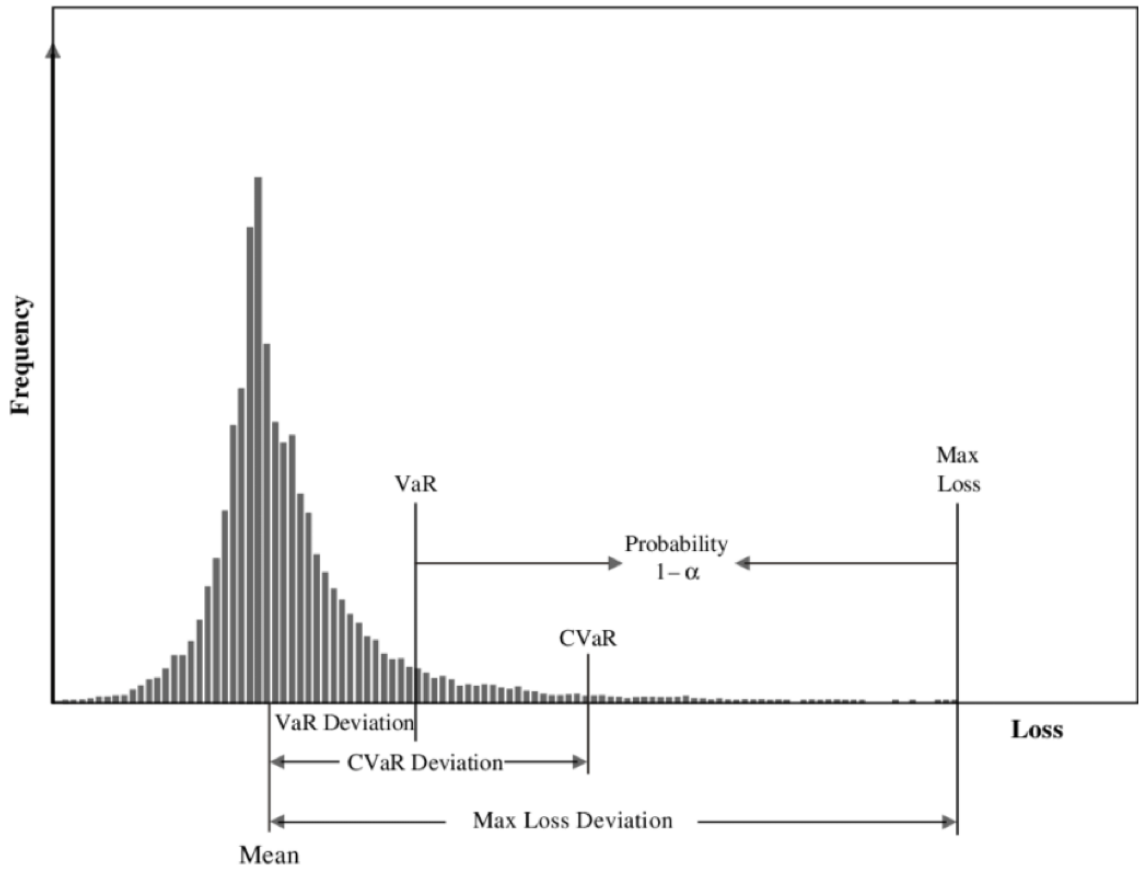

Figure 4.4: Illustration of VaR and CVaR (Uryasev et al., 2010)

Which approach to choose for handling uncertainty is based on the given application and decision

maker. However, VaR has certain shortcomings that would make CVaR more suitable, as described by Rockafellar and Uryasev (2002). It is unstable and difficult to work with numerically without normal distribution on the losses. In addition, it does not control scenarios with losses exceeding the value. This means that it cannot distinguish between scenarios with losses that are only a little bit worse, and losses that go way above the value. The measure can only provide a lowest bound for losses in the tail of the distribution. In addition, VaR is not a coherent measure.

Oppositely to VaR, CVaR has an ability of quantifying the losses in the tail of the loss distribution. In addition, CVaR has a linear implementation, which makes it easier to use in linear optimization. It also yields the same results as VaR for normal distributions. Even though CVaR has certain advantages, both measures may be used. If good models for the tails are not available, VaR may be better. Oppositely, if good models for the tails are available, CVaR should be used.

4.2.4 Multi-Objective Optimization

Multi-objective optimization problems (MOPs) are problems that involve more than one objective function that are to be minimized or maximized. Problems that include only two objective functions, are often referred to as bi-objective optimization problems. The solution from a MOP is a set of solutions that define the best trade-off between the given competing objectives. Multi-objective optimization has an important practical importance, since real-world optimization problems often involve multiple conflicting objectives that must be taken into account (Deb, 2014).

The general form of a MOP with k objectives is defined as

$$min\ f(x) = [f_1(x), f_2(x), ..., f_k(x)]^T \tag{4.36}$$

subject to

$$x \in \mathcal{X} \tag{4.37}$$

where $\mathcal{X}$ is the set of feasible solutions.

The Pareto dominance relation says that a solution x dominate another solution y if x is no worse than y in all the relevant objectives and better than y in at least one of the objectives. x is strictly dominated by y if y is strictly better for all objectives in the problem.

A solution is said to be Pareto optimal if there does not exist another solution that dominates it. A solution is weakly Pareto optimal if it is not strictly dominated by another solution. The Pareto optimal solutions can be represented in a Pareto optimal front as depicted in Figure 4.5. Including weakly Pareto optimal solutions yields a weak Pareto front.

ϵ-Constrained Method

The ϵ-constrained method is a well-known scalarization technique used to generate a Pareto front for MOPs, as described in Jaimes et al. (2009). An advantage of the method is its ability to generate

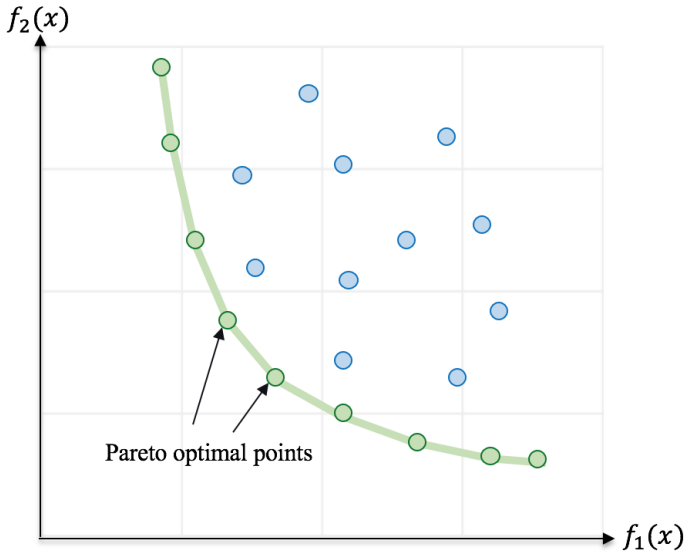

Figure 4.5: Pareto front for two objective functions $f_1(x)$ and $f_2(x)$

non-convex and non-concave parts of the Pareto front. Compared to other methods, such as the lexicographic method, it does not avoid weakly Pareto optimal solution.

The method uses one of the objective functions as the objective of the optimization, while the others are restricted by varying ϵ-values. The model is solved in a given number of iterations to generate a front. Given k objectives, $k \in \mathcal{K}$, and a chosen objective used in the objective function, $f_i(x)$, a MOP is given in the following single objective problem:

$$min\ f_i(x) \tag{4.38}$$

subject to

$$f_k(x) \leq \epsilon_k, \quad k \in \mathcal{K}, k \neq i \tag{4.39}$$

$$x \in X \tag{4.40}$$

where ϵ_k is the target value for objective k, and X is the feasible domain.

To acquire reasonable values for ϵ_k in each iteration, the optimization problem is first solved without any restrictions on the objectives in the constraints. Denoting the maximum value of $f_k(x)$ as f^*, the ϵ_j-values for the iterations for each $f_k(x)$ are given as

$$\epsilon_j = \frac{j}{N} f^*, \quad j = 1, .., N \tag{4.41}$$

where N is the number of iterations performed to generate the Pareto front.

Chapter 5

Problem Description

In the following chapter, the real-time production optimization problem for petroleum production is described. The modelling and optimization is based on historic data measurements from a petroleum field. The main goal is to maximize oil production rates in a short time horizon, i.e. from seconds to days, while considering different operational constraints. The output is explicit recommendations to the production engineers on operational adjustments.

The petroleum field structure varies for different fields. However, all fields consist of a number of wells producing a multiphase flow of oil, gas and water. The multiphase flow is conserved throughout the production system, which is the connected system from the wells to the platform. From a subsea well in the system, the flow is routed to a pipeline via a manifold. The flow in a pipeline may pass through a chain of manifolds, where the flow from other wells are routed to the given pipeline. Multiple wells can be connected to the same manifold. The last manifold in a chain is connected to a separator. The flow in each pipeline is therefore routed to one separator, and there may be several separators on a platform. In addition, there can be multiple chains of manifolds connected to the same separator. The separators split the phases and route the different streams to an export line, where the export line leads the phases to off-site locations.

Each well in the production system is associated with a wellhead pressure, and there are pressure drops over the choke openings of the wells. The manifolds are also associated with a pressure that depends on the pressure drop over the pipelines. In order for the fluids to be able to flow downstream, the pressure must increase upstream in the system.

Different components in the production system have certain production capacity limits. The platform has a maximum capacity on the amount of water it is able to process at a given time, as the water has to be rinsed before being sent back into the ocean. In addition, the platform has an upper gas limit, limiting the total gas in the production system. This comes from limited compressor capacity. Lastly, separators and wells have a capacity limit on the amount they are able to process per time unit for the respective phases of the flow. These capacities are individual for each separator, and the same for all wells.

It is important to take the work process and preferences of the operators into account in order to get an implementable and appropriate optimization model. The output from the optimization represents the physical adjustments that the operators are recommended to perform. Operators have individual preferences on how many operational adjustments to perform, and they typically want to avoid a large number of adjustments. One reason is the difficulty of finding the effect of each change when several changes are made simultaneously. The number of recommended changes should therefore be limited to the maximum number of changes an operator is willing to perform. The operators also prefer to make small, iterative changes to the operation settings, and the value of the adjustment should therefore also be restricted.

The petroleum production system is complex, and the relationships between physical parameters can be modelled to give a realistic picture of the system. The well inflow rate, the pressure drop over the choke openings and the pressure drop in pipelines are such physical factors. The relationships can be modelled by machine learning techniques using historic data measurements. The well inflow and the pressure drop over the choke openings in the production system can be modelled for each well. However, these physical properties cannot be directly controlled by the operator. In order to perform the real-time optimization, relationships that couple well inflow rates, pressure and operational settings per well need to be found. The pressure drops in the pipelines vary with the amount and composition of the flow in the line. The pressures in the manifolds can thus be modelled by finding a mathematical relationship between the pressures and the flow rates in the pipelines.

It is important to take the uncertainty of the models created by historic data measurements into account. The uncertainty of the response of the system increases when bigger changes are made to the operational settings, and the value of the change in choke opening should therefore be restricted by the uncertainty in the well inflow. Another possibility of controlling the uncertainty is with regards to breaching a capacity constraint.

Severe slugging can be investigated to ensure a profitable and safe flow of fluids from the well to the point of sale. The risk of slugging can be predicted by modelling the explanatory variables that lead to slugging. Including a model that predicts the cost of slugging can enhance the recommendations given to the production engineers by taking into account their accepted risk of slugging.

Chapter 6

Optimization Models

In this chapter, mixed integer linear programming (MILP) optimization models are presented. Section 6.1 presents assumptions and simplifications of the models, while Section 6.2 presents a stochastic optimization model for petroleum production with the objective of maximizing real-time oil production. A stochastic bi-objective optimization model that maximizes oil production and minimizes the conditional value-at-risk (CVaR) of gas production is presented in Section 6.3. Further, a stochastic bi-objective model maximizing oil production and minimizing the cost of slugging is presented in Section 6.4. If all stochastic variables in the stochastic models are set to a given scenario, the models are deterministic.

The three stochastic optimization models presented are two-stage recourse models. The first-stage decisions for the models are well adjustment recommendations. In the second stage, the flows and pressures in each scenario are calculated based on the well settings and data-driven models associated with the scenario. The values of the variables in the second stage are thereby direct results from the first-stage action. In the models, the first-stage decisions are variables that are not dependent on a given scenario, whereas the second-stage decisions are scenario-dependent.

6.1 Assumptions and Simplifications

Section 6.1.1 presents assumptions and simplifications for the production system, while Section 6.1.2 presents simplified physical relations in the production system, represented by well models and pipeline pressure models.

6.1.1 Production System

The objective is to maximize real-time oil production. The optimization may be in conflict with other long-term goals, such as maximizing production throughout the whole supply chain. Gas-lift is not applied in the model, thus changing the choke opening is the only decision variable the operators can affect directly.

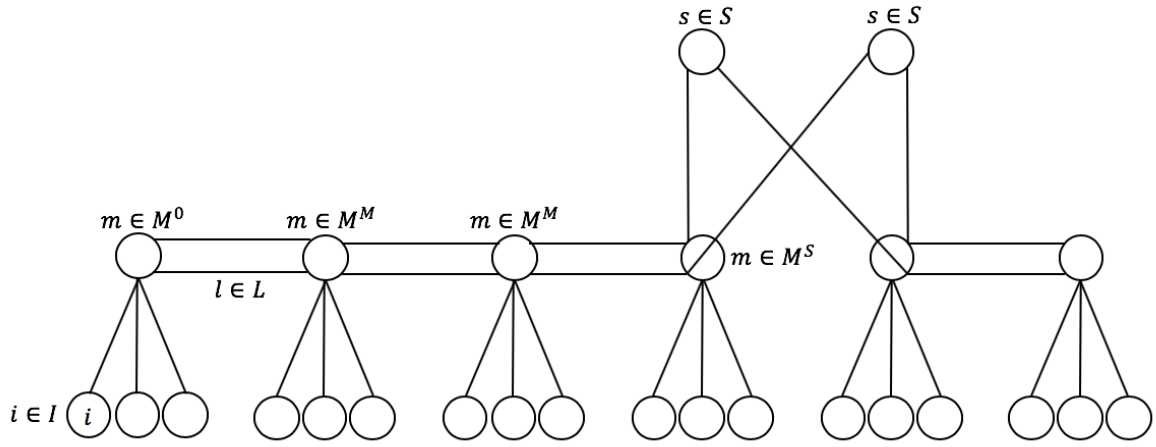

Figure 6.1: Structure of petroleum production system used in the optimization model

Figure 6.1 illustrates the general structure of the petroleum production system used in the optimization problem formulation. The system consists of chained manifolds, indexed m, connected to separators, indexed s. The manifolds first in each chain are in the $\mathcal{M}^O$ set, while the next manifolds are in the $\mathcal{M}^M$ set. The last manifolds in the chain, directly connected to the separators, are defined by the $\mathcal{M}^S$ set. Each well, indexed i, is connected to one manifold. All components are connected via pipelines, indexed l. A well can be routed to different pipelines, but they cannot be rerouted after one pipeline is chosen. One pipeline is connected to one separator, so implicitly, flow from one well is routed to one separator. The number of each component may vary between different production systems.

For the optimization, it is assumed that the system is described with an incidence matrix defining the feasible arcs in the system. The matrix consists of elements A_{ij}, which describe the connection between node i and edge j. As the edges have a specified direction associated with them, the incidence matrix describes a directed graph. If A_{ij} = -1, the edge j leaves node i, if A_{ij} = 1, the edge j enters node i and A_{ij} = 0 otherwise. The nodes are represented by all wells, manifolds and separators, while an edge is the part of a pipeline between two nodes.

6.1.2 Physical Relations in the Production System

The pressures and pressure drops throughout the production system are illustrated in Figure 6.2. There are pressure drops over the chokes and in pipelines. The pressure drop over choke is illustrated between the wellhead pressure p_i and downstream choke pressure p_i^D. Assuming that there is no pressure drop between downstream choke pressure and the pressure of the connected manifold, the downstream choke pressure equals the manifold pressure, p_{ml}. There is a pressure drop in pipelines between the manifold pressures, and for the last manifold there is a pressure drop from the manifold pressure to the separator inlet pressure, p_s^S. The separator pressure is assumed to be constant.

The flow rates from the wells in a production system are impacted by physical factors like pressure, friction and temperature. Pressure drops in a production system are in reality connected by complicated physical relations that are difficult to include in a linear optimization problem. They depend on physical factors such as flow rates, temperature, friction and pressures in other parts of the system. The pressure drop over choke also depends on the choke opening. In the following, simplifications made for these relations are presented.

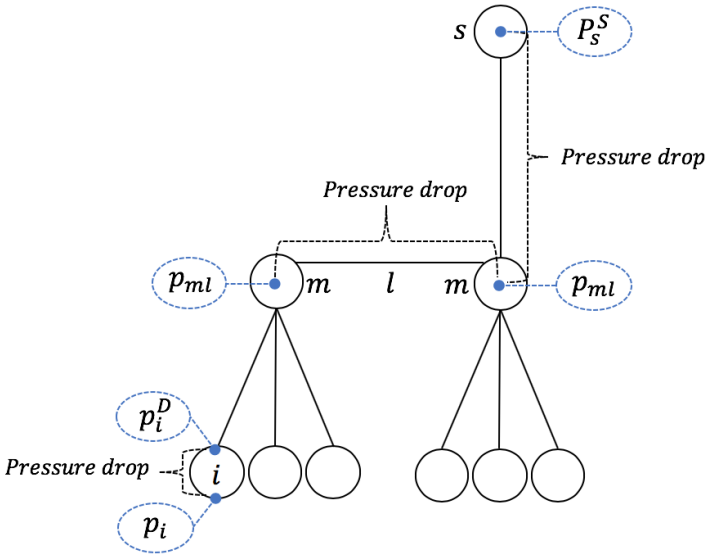

Figure 6.2: Pressure relations in a production system

Well Models

There are two models per well that describe the oil flow rate and the pressure drop over the choke, respectively. These two models, the oil flow rate model and the wellhead pressure model, are collectively referred to as well models. A simplifying assumption that the oil flow rate models are functions of wellhead pressure and choke opening, is made. It is further assumed that the gas-oil ratio (GOR) and water cut (WCT) are constant, so that the gas and water flow rates can be deducted from the oil flow rate. The water cut is the ratio of water to total liquid. The pressure drop over a given choke, represented by the wellhead pressure models, is simplified to a function of the choke opening, the well inflow rates and the pressure downstream choke pressure.

The choke openings are not allowed to change more than a certain limit, which is determined by the uncertainty of the models and an operational limit. The limits for which the models are valid depend on the given data source. The models are created around the current choke openings in order to use the relevant data. The relevant data is in this case the data within the allowed limits of the given choke opening. The well models are delta models, which implies that they describe the change of the output variable given the changes in the input variables. Thus, the operation points after adjustments are described by the initial operation point and the changes suggested.

In reality there are non-linear relationships between choke opening, pressure and production rates for a well. The non-linear relationships can be simplified and described by different machine learning techniques, where the choice of technique is discussed in Section 7.1. In the model formulation in this section, the relationships are described by unknown functions.

Pipeline Pressure Models

There is a model for each manifold and pipeline in the system that describes the pressure drop from a given manifold to the next, in the given pipeline. These models are referred to as pipeline pressure models. The model can be simplified to a function that is dependent on an aggregation of flows from wells flowing from the given manifold to the next, and pressure at a certain point. For all manifolds except the last in a chain, this pressure is the pressure at the subsequent manifold. For the last manifold in a chain, it equals the inlet pressure of the separator.

Similarly to the well models, the pipeline pressure models are expressed as delta models. The relationships are described by unknown functions in the model formulation in this section, and the modelling technique is discussed in Section 7.1.

6.2 Stochastic Optimization Model

A stochastic model for maximizing oil production from a petroleum field is formulated. Each scenario describes a given outcome for the well models and pipeline pressure models. The full formulation of the optimization model can be found in Appendix A.

Sets

Ω	Set of scenarios, indexed ω
$\mathcal{I}$	Set of wells, indexed i
$\mathcal{L}$	Set of pipelines, indexed l
$\mathcal{M}$	Set of manifolds, indexed m
$\mathcal{P}$	Set of phases (o for oil, g for gas, w for water), indexed p
$\mathcal{S}$	Set of separators, indexed s
$\mathcal{N}$	Set of nodes, $\mathcal{N} \in \{I \cup \mathcal{M} \cup \mathcal{S}\}$, indexed j, k
$\mathcal{I}_m$	Set of wells connected to manifold m, $\mathcal{I}_m \in \mathcal{I}$, indexed i
$\mathcal{L}_m$	Set of pipelines connected to manifold m, $\mathcal{L}_m \in \mathcal{L}$, indexed l
$\mathcal{L}_s$	Set of pipelines connected to separator s, $\mathcal{L}_s \in \mathcal{L}$, indexed l
$\mathcal{M}^O$	Set of manifolds first in the chain, $\mathcal{M}^O \in \mathcal{M}$, indexed m
$\mathcal{M}^M$	Set of manifolds connected to two other manifolds, $\mathcal{M}^{\mathcal{M}} \in \mathcal{M}$, indexed m
$\mathcal{M}^S$	Set of manifolds directly linked to a separator, $\mathcal{M}^{\mathcal{S}} \in \mathcal{M}$, indexed m

Parameters

U_i	Initial choke opening of well i
P_i	Initial wellhead pressure for well i
P_i^D	Initial downstream choke pressure for well i
P_s^S	Inlet pressure for separator s
Q_{pi}	Initial flow of phase p in well i
$\overline{U}_i$	Upper limit of choke opening for well i
$\underline{U}_i$	Lower limit of choke opening for well i
$\Delta\overline{U}_i$	Maximum change in choke opening for well i
$\overline{P}_i$	Maximum wellhead pressure for well i
$\overline{P}_{ml}$	Maximum pressure at manifold m in pipeline l
GOR_i	Initial gas-oil ratio for well i
WCT_i	Initial water cut for well i
C_p	Maximum capacity of phase p from wells to manifolds
C_{ps}	Maximum capacity of phase p in separator s
C^G	Total gas capacity of platform
C^W	Total water capacity of platform
D	Maximum number of chokes that can be adjusted

Variables

Δu_i	Delta choke opening of well i
$\Delta p_{i\omega}$	Delta wellhead pressure for well i in scenario ω
$\Delta p^D_{i\omega}$	Delta downstream choke pressure for well i in scenario ω
$\Delta p_{ml\omega}$	Delta pressure at manifold m in pipeline l in scenario ω
$\Delta q_{pi\omega}$	Delta flow of phase p from well i in scenario ω
$\Delta q_{pjkl\omega}$	Delta flow of phase p from node j to node k in pipeline l in scenario ω
$\Delta q^S_{ps\omega}$	Delta flow of phase p to separator s in scenario ω
u_i	Choke opening of well i
$p_{i\omega}$	Wellhead pressure for well i in scenario ω
$p^D_{i\omega}$	Downstream choke pressure for well i in scenario ω
$p_{ml\omega}$	Pressure at manifold m in pipeline l in scenario ω
$q_{pi\omega}$	Flow of phase p from well i in scenario ω
$q_{pjkl\omega}$	Flow of phase p from node j to node k in pipeline l in scenario ω
$q^S_{ps\omega}$	Total flow of phase p to separator s in scenario ω
x_i	1 if well i is open, 0 otherwise
y_{iml}	1 if well i connected to manifold m is open and connected to pipeline l, 0 otherwise
h_i	1 if well i is adjusted, 0 otherwise

Functions

$f^{OIL}_{i\omega}(X)$	Function mapping from the inputs X to the output for the oil flow rate model for well i in scenario ω
$f^{PRS}_{i\omega}(X)$	Function mapping from the inputs X to the output for the wellhead pressure model for well i in scenario ω
$f^{PRS}_{ml\omega}(X)$	Function mapping from the inputs X to the output for the pipeline pressure model for manifold m in pipeline l in scenario ω

Objective Function

The objective is to maximize the mean of the total oil production from all separators to the platform for the scenarios.

$$max\ w = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{s \in \mathcal{S}} q^S_{os\omega} \tag{6.1}$$

Gas Capacity

The total amount of gas flowing via the separators to the platform must be less than or equal to the maximum gas capacity of the platform for all scenarios.

$$\sum_{s \in \mathcal{S}} q^S_{gs\omega} \leq C^G, \quad \omega \in \Omega \tag{6.2}$$

Water Capacity

The total amount of water flowing via the separators to the platform must be less than or equal to the maximum water capacity of the platform for all scenarios.

$$\sum_{s \in \mathcal{S}} q^S_{ws\omega} \leq C^W, \quad \omega \in \Omega \tag{6.3}$$

Choke Opening

When a choke opening of a well is adjusted, the new choke opening is given by the initial choke opening and the value of the change.

$$u_i = U_i + \Delta u_i \quad i \in \mathcal{I} \tag{6.4}$$

Number of Well Adjustments

As the well operator has to manually adjust the wells, it is preferred to keep the number of adjustments below a certain level. The variable h_i, indicating if a well is adjusted or not, is forced to one when a choke is adjusted in constraints (6.5) and (6.6). The maximum positive delta choke opening for a well is $\overline{U}_i - U_i$ and the maximum negative delta choke opening is $U_i - \underline{U}_i$. The limits for a well, $\underline{U}_i$ and $\overline{U}_i$, are determined by the uncertainty of the data source for a given well, and can give tighter limits than shut off and fully open. Constraint (6.7) ensures that a maximum number of D wells are adjusted.

$$\Delta u_i - h_i(\overline{U}_i - U_i) \leq 0, \quad i \in \mathcal{I} \tag{6.5}$$

$$\Delta u_i + h_i(U_i - \underline{U}_i) \geq 0, \quad i \in \mathcal{I} \tag{6.6}$$

$$\sum_{i \in \mathcal{I}} h_i \leq D \tag{6.7}$$

Well Routing

Constraints (6.8) ensure that a well is closed if the choke opening is zero. Similarly, constraints (6.9) ensure that a well must be open if the choke opening is greater than zero. Constraints (6.10) state that there is no flow from a well routed to a pipeline if the well is closed. They also assure that the flow from each well can be routed to maximum one pipeline via the manifold.

$$u_i \geq \underline{U}_i x_i, \quad i \in \mathcal{I} \tag{6.8}$$

$$u_i \leq \overline{U}_i x_i, \quad i \in \mathcal{I} \tag{6.9}$$

$$\sum_{l \in \mathcal{L}_m} y_{iml} - x_i = 0, \quad m \in \mathcal{M}, i \in \mathcal{I}_m \tag{6.10}$$

Mass Balances in Manifolds

Constraints (6.11) - (6.15) ensure mass balances in the manifolds for all scenarios. Mass balances for delta rate variables are defined similarly, and can be found in Appendix A. Constraints (6.11) state that the flow produced in well i equals the sum of the flow sent to connected pipelines via a manifold m. Constraints (6.12) ensure that the flow into a manifold from its connected wells equals the flow out for the first manifold in a chain. Constraints (6.13) and (6.14) are similar for the other manifolds in the sets, $\mathcal{M}^M$ and $\mathcal{M}^S$, where the flow into the manifold is the sum of flow from its connected wells and the flow from the previous manifold in the chain. Figure 6.3 illustrates the flow balance for a manifold $m \in \mathcal{M}^M$ for a given scenario. There should be no flow from a well into a manifold if no well is routed to the manifold, as stated in constraints (6.15).

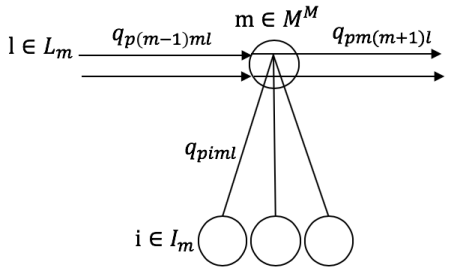

Figure 6.3: Manifold mass balance for $m \in \mathcal{M}^M$

$$\sum_{l \in \mathcal{L}_m} q_{piml\omega} = q_{pi\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, \omega \in \Omega \tag{6.11}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} = q_{pm(m+1)l\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^O, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.12}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} + q_{p(m-1)ml\omega} = q_{pm(m+1)l}, \quad p \in \mathcal{P}, m \in \mathcal{M}^M, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.13}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} + q_{p(m-1)ml\omega} = q_{pmsl\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{6.14}$$

$$q_{pimlω} \leq C_p y_{iml}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.15}$$

Mass Balances in Separators

Constraints (6.16) ensure mass balance in the separators for all scenarios, where the flow into a separator equals the sum of flows in each pipeline connected to the separator. Constraints (6.17) restrict the flow of each phase into the separator below the separator capacity.

$$\sum_{m \in \mathcal{M}^S} \sum_{l \in \mathcal{L}_s} q_{pmslω} = q^S_{psω}, \quad p \in \mathcal{P}, s \in \mathcal{S}, \omega \in \Omega \tag{6.16}$$

$$\sum_{m \in \mathcal{M}^S} q_{pmslω} \leq C_{ps}, \quad p \in \mathcal{P}, s \in \mathcal{S}, l \in \mathcal{L}_s, \omega \in \Omega \tag{6.17}$$

Well Inflow

The oil flow rate models for each well in a given scenario are represented by equations (6.18), as functions of the change in the wellhead pressure and choke opening. The flow of oil from a well is given by the initial value of the rate and the value of the change, as stated in equations (6.19). Constraints (6.20) force the flow of oil to zero if the well is shut off, and limit the oil flow to the maximum capacity from a well to a pipeline if open. The flow of the gas and water phase from a well are represented by constraints (6.21) and (6.22). The gas and water flow are deducted from the oil flow by their GOR and WCT values. The delta flows of gas and water are similarly described, and can be found in the full formulation of the optimization model in Appendix A.

$$\Delta q_{oiω} = f^{OIL}_{iω}(\Delta p_{iω}, \Delta u_i), \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.18}$$

$$q_{oiω} = Q_{oi} + \Delta q_{oiω} \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.19}$$

$$q_{oiω} \leq C_o x_i, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.20}$$

$$q_{giω} = GOR_i q_{oiω}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.21}$$

$$q_{wiω} = \left(\frac{WCT_i}{1 - WCT_i} \right) q_{oiω}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.22}$$

Pressure Drop over Choke

The wellhead pressure model for each well in a scenario is modelled as a function of different variables, including the change in choke opening, as stated in equations (6.23). The wellhead pressure for a well is given by the initial value of the pressure and the change in pressure, described by constraints (6.24). Similarly, constraints (6.25) state that the downstream choke pressure for each well is given by the initial value of the pressure and the change in pressure.

$$\Delta p_{iω} = f^{PRS}_{iω}(\Delta p^D_{iω}, \Delta u_i, \Delta q_{oiω}, \Delta q_{giω}, \Delta q_{wiω}), \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.23}$$

$$p_{i\omega} = P_i + \Delta p_{i\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.24}$$

$$p^D_{i\omega} = P^D_i + \Delta p^D_{i\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.25}$$

Pressure Drop in Pipelines

The pipeline pressure models describes the pressure drop in pipelines per scenario. The change in manifold pressure for all manifolds excluding the last in a chain is expressed by equations (6.26), as a function of pressure in the next manifold in the chain and the given flow rates. For the last manifold in a chain, represented by equations (6.27), the pressure drop is a function of the flow rates. It should be noted that the function does not include the inlet pressure for the separator, as this is assumed to be constant. Constraints (6.28) enforce the manifold pressure to be equal to the downstream pressure of the wells connected to the given manifold, as the pressure drop is assumed to be zero between the two points. The manifold pressure is given by the initial value of the pressure and the change in pressure, as ensured in constraints (6.29).

$$\begin{aligned}\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta p_{(m+1)l\omega}, \Delta q_{om(m+1)l\omega}, \Delta q_{gm(m+1)l\omega}, \Delta q_{wm(m+1)l\omega}),\\ m \in \mathcal{M}\backslash\{\mathcal{M}^S\}, l \in \mathcal{L}_m, \omega \in \Omega\end{aligned} \tag{6.26}$$

$$\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta q_{omsl\omega}, \Delta q_{gmsl\omega}, \Delta q_{wmsl\omega}), \quad m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{6.27}$$

$$p_{ml\omega} = p^D_{i\omega}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.28}$$

$$p_{ml\omega} = P_{ml} + \Delta p_{ml\omega}, \quad m \in \mathcal{M}, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.29}$$

Pressure Relations

The pressure must increase upstream in the system in order for the fluids to be able to flow downstream for all scenarios. This is represented by constraints (6.30) - (6.32). For constraints (6.30), when y_{iml} is set to zero there is no flow from any of the wells connected to a manifold, and the pressures may take an arbitrarily value. When y_{iml} is set to one there is flow, and a pressure drop is enforced.

Constraints (6.31) and (6.32) state that there must be a flow in the connected pipeline of a manifold in order for the pressure drops to be enforced. There is a flow in the pipeline if either of the wells connected to the pipeline earlier in the chain produce oil, thus y_{iml} for at least one well is greater than zero. For constraints (6.31), a manifold m^* is connected to the same pipelines as a manifold m. In constraints (6.32), a manifold m^* equals the manifold m when $m^* \in \mathcal{M}^S$.

Constraints (6.30) and (6.31) are non-linear, and are therefore relaxed in constraints (6.44) and (6.45).

$$p_{ml\omega} y_{iml} \leq p_{i\omega}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.30}$$

$$p_{ml\omega} y_{im^*l} \leq p_{(m-1)l\omega}, \quad m^*, m \in \mathcal{M}\backslash\{\mathcal{M}^O\}, m^* < m, i \in \mathcal{I}_{m^*}, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.31}$$

$$P^S_s y_{im^*l} \leq p_{ml\omega}, \quad m^* \in \mathcal{M}, m \in \mathcal{M}^S, i \in \mathcal{I}_{m^*}, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{6.32}$$

Variable Bounds

The variable bounds for Δu_i are defined in constraints (6.33). The limits on the choke opening $\underline{U}_i$ and $\overline{U}_i$ are decided by uncertainty in the data source. Parameters $\Delta\overline{U}_i$ are defined by operational considerations. The other delta variables are implicitly bounded by the value of Δu_i, and are therefore created without bounds.

It should be noted that variables are defined and generated based on the network of the production system, further defined by an incidence matrix. As an example using constraints (6.39), the variable $q_{pisl\omega}$ will not be created, as there is no edge found in the incidence matrix directly connecting a well i and a separator s.

$$\Delta u_i \in (max\{\underline{U}_i - U_i, -\Delta\overline{U}_i\}, min\{\overline{U}_i - U_i, \Delta\overline{U}_i\}), \quad i \in \mathcal{I} \tag{6.33}$$

$$u_i \in (\underline{U}_i, \overline{U}_i), \quad i \in \mathcal{I} \tag{6.34}$$

$$p_{i\omega} \geq 0, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.35}$$

$$p^D_{i\omega} \geq 0, \quad i \in \mathcal{I}, \omega \in \Omega \tag{6.36}$$

$$p_{ml\omega} \geq 0, \quad m \in \mathcal{M}, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.37}$$

$$q_{pi\omega} \geq 0, \quad p \in \mathcal{P}, i \in \mathcal{I}, \omega \in \Omega \tag{6.38}$$

$$q_{pjkl\omega} \geq 0, \quad p \in \mathcal{P}, j \in \mathcal{N}, k \in \mathcal{N}, l \in \mathcal{L}, \omega \in \Omega \tag{6.39}$$

$$q^S_{ps\omega} \geq 0, \quad p \in \mathcal{P}, s \in \mathcal{S}, \omega \in \Omega \tag{6.40}$$

$$x_i \in \{0, 1\}, \quad i \in \mathcal{I} \tag{6.41}$$

$$y_{iml} \in \{0, 1\}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m \tag{6.42}$$

$$h_i \in \{0, 1\}, \quad i \in \mathcal{I} \tag{6.43}$$

Relaxation of Constraints (6.30) **and** (6.31)

Constraints (6.30) and (6.31) are relaxed by introducing new sets of constants, $\overline{P}_i$ and $\overline{P}_{ml}$. They represent the maximum wellhead pressure for a well i connected to manifold m and the maximum pressure in pipeline l at manifold m, respectively. Constraints (6.44) and (6.45) describe the new pressure relation constraints. When y_{iml} is set to zero there is no flow in the pipeline, and the pressures may take any value below the maximum pressures. When y_{iml} is set to one there is flow, and a pressure drop is forced.

$$p_{ml\omega} \leq p_{i\omega} + \overline{P}_i(1 - y_{iml}), \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.44}$$

$$p_{ml\omega} \leq p_{(m-1)l\omega} + \overline{P}_{(m-1)l}(1 - y_{im^*l}), \quad m^*, m \in \mathcal{M}\backslash\{\mathcal{M}^O\}, m^* < m, i \in \mathcal{I}_{m^*}, l \in \mathcal{L}_m, \omega \in \Omega \tag{6.45}$$

6.3 Bi-Objective Stochastic Model with Conditional Value-at-Risk

In this model, uncertainty in gas production is studied. This provides a method of limiting the uncertainty aggregated over the wells, and thus avoiding individual well uncertainty limits. The gas capacity constraint is used for illustrative purposes, but the method can also be used for the water capacity constraint. The concept of the conditional value-at-risk (CVaR) is applied to handle this uncertainty. It is referred to Section 4.2.3 for a theoretical background on CVaR and VaR.

CVaR is an appropriate method as it is a coherent risk measure which quantifies only the downside risk, compared to measures such as standard deviation. For the application in this thesis, the downside risk is the issue of too high gas production. In addition, as the method is linearizable, it does not significantly increase the complexity of the model (Weskamp et al., 2019).

In the following section, Section 6.3.1 presents additional notation in the model regards to the stochastic optimization model. Further, Section 6.3.2 formulates the capacity constraint as a CVaR constraint. Lastly, Section 6.3.3 presents a formulation where the CVaR is included in the objective function.

6.3.1 Additional Notation in the Model

The same constraints as in Section 6.2 still apply, but there are two exceptions. Firstly, the gas capacity constraints, constraints (6.2), do not apply and are replaced by a CVaR constraint. Secondly, the uncertainty is no longer studied per well, such that the limits $\underline{U}_i$ and $\overline{U}_i$ are not defined by the uncertainty in the data source, but are set to its minimum and maximum limits, shut in and fully open. Additional notation is presented in the following tables.

Parameters

α	Confidence level

Variables

z_ω^G	Excess value over value-at-risk for gas in scenario ω
t^G	Value-at-risk for gas
$CVaR_\alpha$	The conditional value-at-risk for confidence level α
θ	The negative of the expected value of the total oil production
ψ^G	The expected excess value over value-at-risk for gas

6.3.2 Formulation of CVaR Constraint

The gas capacity constraints from the stochastic optimization model in Section 6.2 are restated in constraints (6.46). This is a robust formulation that requires the gas handling capacity to hold for all scenarios. It can, however, be argued that this is an overly conservative approach, especially as the gas

capacity often expresses an average that the production should lie below as described in Section 2.3.1.

$$\sum_{s \in \mathcal{S}} q^S_{gs\omega} \leq C^G, \quad \omega \in \Omega \tag{6.46}$$

Another approach is to require the expected value over all scenarios to hold, which may be inadequate. This approach ignores the uncertainty, and can lead to unwanted solutions, in addition to big variations in gas production where many scenarios may breach the constraint.

CVaR is an approach that allows the decision maker to control the conservativeness of the solution, and can be seen as an approach between the strict robust approach and the non-conservative expected value approach. Using CVaR to handle the gas constraints, scenarios producing very high gas rates are limited. As the gas capacity limit may be temporary violated, but should not be breached by very high values, CVaR is a suitable measure.

Similarly to Hanssen et al. (2015), constraints (6.47) and (6.48) express the CVaR constraint for the gas handling capacity, where the first constraints define the variables as described in the notation, and the second constraint requires the CVaR to be less than the gas capacity.

$$\sum_{s \in \mathcal{S}} q^S_{gs\omega} - t^G \leq z^G_\omega, \quad \omega \in \Omega \tag{6.47}$$

$$t^G + \frac{1}{1-\alpha} \psi^G \leq C^G \tag{6.48}$$

Equation (6.49) expresses the expected value of gas production as an average over the scenarios.

$$\psi^G = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} z^G_\omega \tag{6.49}$$

Variable Bounds

$$z^G_\omega \geq 0, \quad \omega \in \Omega \tag{6.50}$$

$$t^G \geq 0 \tag{6.51}$$

6.3.3 Formulation of the Model

Objective Function

In this formulation, CVaR is included in the objective function. Moving the CVaR-constraint to the objective enables the decision maker to study the trade-off between oil production and uncertainty in gas production expressed by CVaR. By moving the CVaR to the objective function, the value of the gas capacity is removed and the decision maker can study the capacity as a soft limit. Since temporary violations of the capacity constraints sometimes can be allowed, this might create value for the decision maker. For example, if there is a relatively high risk of breaching the gas capacity constraint, it may be allowed if the associated oil production rate is high.

As explained in Section 4.2.3, CVaR minimizes the uncertainty, or risk, such that the expected losses in the worst $(1-\alpha)$-percentile of the scenarios are minimized. However, in this case, the CVaR represents the average of the $(1-\alpha)$-percentile highest gas rates. The gas rates are chosen for the CVaR value as they are bounded by gas capacities that are difficult to know for certain. In addition, the gas rates themselves are uncertain. Instead of restricting the gas to a capacity limit, values where a small increase in total gas production gives a larger increase in total oil production can be found. This would be valuable to know for decision makers when operating with uncertain gas rates and capacity limits on the platforms.

The objective is to maximize the expected oil production and minimize the expected value of the $(1-\alpha)$-percentile highest gas production rates. The objective is rewritten to obtain a minimization objective.

$$min\ w = [\theta, CVaR_\alpha] \tag{6.52}$$

$$CVaR_\alpha = (t^G + \frac{1}{1-\alpha}\psi^G) \tag{6.53}$$

Equation (6.54) presents the negative expected value of the total oil production, which is an average over the scenarios.

$$\theta = -\frac{1}{|\Omega|}\sum_{\omega\in\Omega}\sum_{s\in\mathcal{S}} q^S_{os\omega} \tag{6.54}$$

CVaR Constraints

While constraint (6.48) is moved to the objective, constraints (6.47) are kept as restrictions in the model.

$$\sum_{s\in\mathcal{S}} q^S_{gs\omega} - t^G \leq z^G_\omega, \quad \omega \in \Omega \tag{6.55}$$

Variable Bounds

$$z^G_\omega \geq 0, \quad \omega \in \Omega \tag{6.56}$$

$$t^G \geq 0 \tag{6.57}$$

6.4 Bi-Objective Stochastic Model with Slug Severity

In this section, a bi-objective stochastic optimization model where the objectives are to maximize oil production and minimize the cost of slugging from the scenarios is presented. The slug severity models integrated in the optimization model are explained in Section 6.4.1, additional notation for this model is presented in Section 6.4.2, and the formulation is presented in Section 6.4.3.

6.4.1 Assumptions for Slug Severity Models

The flow assurance issue of slugging is described in Section 2.3.2, and approaches for modelling slug are reviewed in Section 3.4.

Slugging in a riser can be described by complicated physical relations that would be difficult to include in a linear optimization problem. Predicting slugging is therefore complicated. A data-driven technique that through explanatory variables predicts the risk of slugging for the current operation area, is proposed. The output from the predictive slug severity models are categories representing the cost of slugging for the current operation area. The input variables for the models are simplified to the decision variables that are available from the optimization model and that makes physical sense as a driver of slugging. The variables used as input are oil, gas and water rates in the riser, in addition to the manifold pressure in the manifold connected to the separator.

Different machine learning methods may be used to develop such predictive slug severity models, and the choice of a suitable technique is discussed in Section 7.2. In the model formulation in this section, the models are described by unknown functions.

6.4.2 Additional Notation in the Model

The same constraints as in Section 6.2 still apply. Additional notation is presented in the following tables.

Sets

$\mathcal{C}^{\mathcal{SLUG}}$	Set of categories for cost of slugging, indexed c

Variables

$c_{l\omega}$	Slugging category of riser belonging to pipeline l in scenario ω
η	The expected cost of slugging
θ	The negative of the expected value of the total oil production

Function

$f_l^{SLUG}(X)$	Function mapping from the inputs X to the output of the model for riser belonging to pipeline l

6.4.3 Formulation of the Model

Objective Function

The expected total oil production over the different scenarios is maximized, while the expected cost of slugging is minimized. The objective is rewritten to obtain a minimization objective.

$$min\ w = [\theta, \eta] \tag{6.58}$$

Expected Value Functions

Constraint (6.59) expresses the negative expected value of the total oil production, which is an average over the scenarios. Constraint (6.60) expresses the expected cost of slugging as an average over the scenarios, summarized for all pipelines.

$$\theta = -\frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{s \in \mathcal{S}} q^S_{os\omega} \tag{6.59}$$

$$\eta = \frac{1}{|\Omega|} \sum_{l \in \mathcal{L}} \sum_{\omega \in \Omega} c_{l\omega} \tag{6.60}$$

Slug Severity Models

The slug severity values, or the risk of slug for a given operation point, can be described with cost categories representing the risk. The output from the slug severity models are the slug cost categories of the current operation point for each of the scenarios.

$$c_{l\omega} = f^{SLUG}_l(q_{omsl\omega}, q_{wmsl\omega}, q_{gmsl\omega}, p_{ml\omega}), \quad m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \mathcal{L}_S, \omega \in \Omega \tag{6.61}$$

Variable bounds

$$c_{l\omega} \in \mathcal{C}^{\mathcal{SLUG}}, \quad l \in \mathcal{L}, \omega \in \Omega \tag{6.62}$$

Chapter 7

Application of Machine Learning Methods

In this chapter, the models required in the optimization problem are described. Section 7.1 presents the well models and pipeline pressure models, while Section 7.2 presents the slug severity models. Both sections present the choice of machine learning technique, give a thorough explanation of the models and state a general formulation of the technique chosen. The formulations can be directly integrated in a mixed integer linear programming (MILP) problem. Multivariate adaptive regression splines (MARS) is used for the well models and the pipeline pressure models, while a classification tree (CT) approach is used for the slug severity models. MARS is a regression-based approach and CT is a classification-based approach, thus the integration of two different machine learning approaches in optimization is studied.

7.1 Well Models and Pipeline Pressure Models

MARS is the chosen technique for the well models and the pipeline pressure models. Section 7.1.1 presents the choice of machine learning technique, Section 7.1.2 gives a further explanation of the models, and Section 7.1.3 provides a general formulation of a MARS model that can be integrated with a MILP problem.

7.1.1 Choice of Machine Learning Technique

The well models presented in the stochastic optimization formulation in Section 6.2 are restated in equations (7.1) and (7.2). The first equations describe the oil flow rate models and the second equations describe the wellhead pressure models.

$$\Delta q_{oi\omega} = f_{i\omega}^{OIL}(\Delta p_{i\omega}, \Delta u_i), \quad i \in \mathcal{I}, \omega \in \Omega \tag{7.1}$$

$$\Delta p_{i\omega} = f_{i\omega}^{PRS}(\Delta p_{i\omega}^{D}, \Delta u_i, \Delta q_{oi\omega}, \Delta q_{gi\omega}, \Delta q_{wi\omega}), \quad i \in \mathcal{I}, \omega \in \Omega \tag{7.2}$$

The pipeline pressure models from Section 6.2, describing the pressure drop in pipelines, are also restated in equations (7.3) and (7.4). The first equations describe the pressure drop in the pipelines connecting manifolds, and the second equations describe the pressure drop in the riser.

$$\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta p_{(m+1)l\omega}, \Delta q_{om(m+1)l\omega}, \Delta q_{gm(m+1)l\omega}, \Delta q_{wm(m+1)l\omega}), \quad m \in \mathcal{M}\backslash\{\mathcal{M}^S\}, l \in \mathcal{L}, \omega \in \Omega \tag{7.3}$$

$$\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta q_{omsl\omega}, \Delta q_{gmsl\omega}, \Delta q_{wmsl\omega}), \quad m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \mathcal{L}_S, \omega \in \Omega, \tag{7.4}$$

Different machine learning techniques could have been used to model the well and pipeline pressure models. MARS, presented in Section 4.1.2, was chosen as the technique for these models, seeing that it has some very desirable characteristics.

MARS is an adaptive method that does automatic variable selection, which means that it includes the important variables in the model and excludes the others. The method identifies important independent variables through the basis functions when many potential variables are considered. This is an advantage that makes the models very flexible and easy to update when new data becomes available.

Furthermore, MARS can model non-linearities and find complex relationships in a dataset. Additionally, the method has a relatively short training process, and can thus save modelling time when the dataset is large. MARS models are also easy to understand and interpret, compared to for example neural nets, which makes them easier to explain to the users of the optimization model.

The most important factor for choosing MARS models is that the models can be integrated with a MILP problem. The MARS models can be described in the optimization using binary variables for the hinge functions. Since binary variables are used, an increasing number of hinge functions will have a big effect on the solution time. The interaction terms in the MARS models cannot be directly integrated in the optimization since they are non-linear. An approximation to interaction terms that is linear is proposed in the following. The default interaction terms of MARS models will be referred to as multiplicative interaction, and the approximation presented in this thesis will be referred to as linear interaction.

7.1.2 Explanation of the Models

It is referred to Section 4.1.2 for a theoretical background on MARS. As explained in that section, multiplicative interaction terms between variables can be created in the MARS algorithm. Denote two explanatory variables x_1 and x_2, and the target variable $y = B(x_1, x_2)$. The interaction terms are expressed as a product of hinge functions of the form

$$max(0, x_1 - A) \cdot max(0, x_2 - B) \tag{7.5}$$

Here, A and B are constants.

Figure 7.1 illustrates a MARS model with a multiplicative interaction term. The product of the hinge functions are non-zero over the space of predictors where both hinge functions are positive.

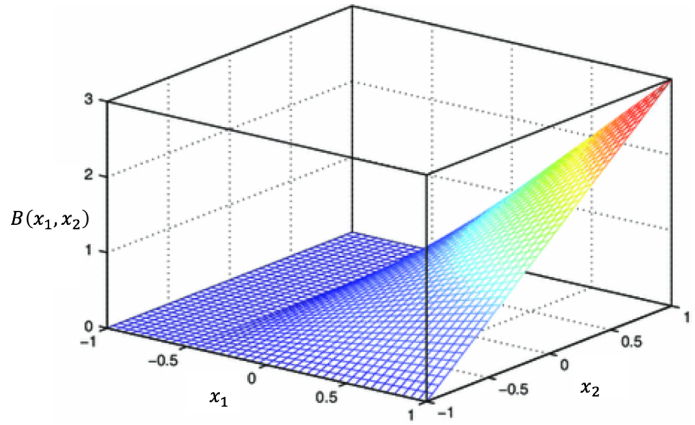

Figure 7.1: Possible MARS model for target variable $y = B(x_1, x_2)$ with interaction terms between variables x_1 and x_2 (Koc and Bozdogan, 2015)

The function $B(x_1, x_2)$ can be expressed

$$max(0, x_1 + 0.5) \cdot max(0, x_2 + 1) \tag{7.6}$$

Since the models must be linear in order to be integrated in a MILP problem, products of hinge functions cannot be directly applied. A method of linearizing multiplicative interaction terms from the MARS algorithm can be found in Martinez et al. (2017). In this thesis, another approximation to interaction terms is proposed. Instead of using a multiplicative interaction term, a linear interaction term created by weighting the relevant variables using α-values, $\alpha \in [0, 1]$, is proposed. Denoting a constant C and a weighted variable w, the linear interaction terms can be written

$$max(w - C, 0) \tag{7.7}$$

where

$$w = \alpha_1 \cdot x_1 + \alpha_2 \cdot x_2 \tag{7.8}$$

The new expression is compliant with a MILP problem as it is linear. The α-parameters, α_1 and α_2, sum up to one. Only linearly independent combinations are generated. While the method proposed by Martinez et al. (2017) performs linearization after the MARS algorithm is executed, the approach in this thesis provide linear interaction terms as input to the MARS algorithm.

Note that this is only an approximation, and not a replacement for non-linear terms. However, it provides an opportunity to model linear interactions between variables. The weighted variable could also be a combination of more than two variables, and the general formulation provided in the next section allows for this. The variable w will be referred to as an α-weighted variable.

7.1.3 Formulation of MARS Model

A formulation of a MARS model that describes a target variable based on a set of explanatory variables is presented in this section. The MARS model is a general formulation, and can be used in a variety of different applications. As explained in Section 7.1.2, products of hinge functions are not applied in this formulation as it is linear. The basis functions are therefore referred to as hinge functions, and the basis functions that are constants are aggregated to the intercept of the output. This formulation will therefore refer to hinge functions.

Sets

$\mathcal{A}$	Set of explanatory variables, indexed a
$\mathcal{H}$	Set of hinge functions, indexed h

Parameters

α_{ah}	Weighting parameter for explanatory variable a in hinge function h
β_h	Coefficient of the hinge function h
κ_h	Knot of the hinge function h
β_0	Intercept of the output value
$\overline{\Psi}_h$	Upper value of the hinge function h
$\overline{M}_h$	Upper value of the negative component of hinge function h
$\underline{X}_a$	Lower value of the explanatory variable a
$\overline{X}_a$	Upper value of the explanatory variable a
$\underline{\Gamma}_h$	Lower value of the variable associated with hinge function h
$\overline{\Gamma}_h$	Upper value of the variable associated with hinge function h
$\underline{Z}$	Lower value of the output variable z
$\overline{Z}$	Upper value of the output variable z

Variables

x_a	Explanatory variable a
l_h	Variable associated with hinge function h
ψ_h	Hinge function h
μ_h	Negative component of the output from hinge function h
z	Output variable
y_h	1 if μ_h is forced to zero, 0 if ψ_h is forced to zero, in hinge function h

Output

The output is described in constraint (7.9), and each hinge function is described by equations (7.10). Equations (7.10) are non-linear, and are therefore reformulated in constraints (7.18)-(7.20). The variable in a hinge function might be a single variable or a linear combination of multiple variables, represented in equations (7.11). If only one of the α_{ah}-variables equals one for a hinge function h, the variable in the hinge function is a single variable.

$$z = \beta_0 + \sum_{h \in \mathcal{H}} \beta_h \psi_h \tag{7.9}$$

$$\psi_h = max(l_h - \kappa_h, 0), \quad h \in \mathcal{H} \tag{7.10}$$

$$l_h = \sum_{a \in \mathcal{A}} \alpha_{ah} x_a \quad h \in \mathcal{H} \tag{7.11}$$

Variable bounds

The limits in constraints (7.13) and constraints (7.16) are found from the limits on x_a in constraints (7.12).

$$x_a \in [\underline{X}_a, \overline{X}_a], \quad a \in \mathcal{A} \tag{7.12}$$

$$l_h \in [\underline{\Gamma}_h, \overline{\Gamma}_h], \quad h \in \mathcal{H} \tag{7.13}$$

$$\psi_h \geq 0, \quad h \in \mathcal{H} \tag{7.14}$$

$$\mu_h \geq 0, \quad h \in \mathcal{H} \tag{7.15}$$

$$z \in [\underline{Z}, \overline{Z}] \tag{7.16}$$

$$y_h \in (0, 1), \quad h \in \mathcal{H} \tag{7.17}$$

Reformulation of constraints (7.10)

Constraints (7.10) are reformulated by decoupling the output of the hinge function into a positive and a negative component, as described in (7.18). At least one of the components ψ_h and μ_h must be zero to ensure uniqueness of the solution as described in constraints (7.19) and (7.20). When y_h is set to zero, the value of the hinge function, the positive component, is forced to zero. When y_h is set to one for a given h, the negative component μ_h is zero, and the hinge function takes the value $l_h - \kappa_h$. The upper limit $\overline{\Psi}_h$ of the hinge function and the upper limit $\overline{M}_h$ of the negative component are directly given by the limits on l_h and the knot κ_h.

$$l_h - \kappa_h = \psi_h - \mu_h, \quad h \in \mathcal{H} \tag{7.18}$$

$$\psi_h \leq \overline{\Psi}_h y_h, \quad h \in \mathcal{H} \tag{7.19}$$

$$\mu_h \leq \overline{M}_h (1 - y_h), \quad h \in \mathcal{H} \tag{7.20}$$

7.2 Slug Severity Models

Decision tree learning in the form of a classification tree (CT) is the technique chosen for the slug severity models. Section 7.2.1 presents the rationale behind the choice of the machine learning technique, Section 7.2.2 explains the models, and Section 7.2.3 presents a formulation of a CT that can be integrated in a MILP problem.

7.2.1 Choice of Machine Learning Technique

$$c_{l\omega} = f_l^{SLUG}(q_{omsl\omega}, q_{wmsl\omega}, q_{gmsl\omega}, p_{ml\omega}), \quad m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \mathcal{L}_S, \omega \in \Omega \tag{7.21}$$

The functions representing the slug severity models given in Section 6.4.3 are restated in equations (7.21). There is one slug severity model representing each of the risers in a production system. The riser is the part of a pipeline between the last manifold in a chain and a separator.

Similarly to the well and pipeline models, different techniques can be used to model the slug severity. The choice of using CTs, presented in Section 4.1.3, is mainly based on the simplicity of the technique, and how it is easily integrated with a linear optimization problem. In addition, the goal of the slug severity models is to be able to predict a slugging risk or value based on the current operation point for each riser, and CTs are suited for this purpose.

The CT divides the data into different categories, and each new data sample will fall into one of these categories. This is a simple, but efficient way to model the operation area, that fits the purpose of indicating a slug cost. Further, the resulting classification model can be easily interpreted and visualized, and it is easy to follow the reasoning behind a category by following the branches of a tree.

Given a big dataset, a CT is able to analyze the data in a reasonable amount of time. Given additionally new data, the method may however be non-robust, meaning that there may be large changes in the tree. As our problem is real-time, the tree will not be updated with new data, and thus this problem is avoided.

A decisive element is that the resulting decision tree can be integrated with the optimization model, by transforming the constraints associated to the tree to model operation areas. Binary variables are used to model the operation areas. Since binary variables are used, an increasing number of operation areas will have an effect on the solution time.

7.2.2 Explanation of the Models

The CT seeks to explain how the input variables influence the cost of slugging. One CT corresponds to one riser in a petroleum production system. Figure 7.2 illustrates the structure of a CT. The root node represents the first branching of the tree, and each leaf node is connected to the root node via a path. A connection between two different nodes in a path is a branch, so the number of branches in a given

path equals at most the depth of the tree. A branch is associated with its corresponding constraint, which can be seen in the tree to the right in the figure. The CT builds branches on different input variables (features), where each leaf node contains a value indicating the expected cost category associated with the slugging for that path. The input variables in the slug severity models are primarily based on factors that are known to affect the risk of slugging in a riser.

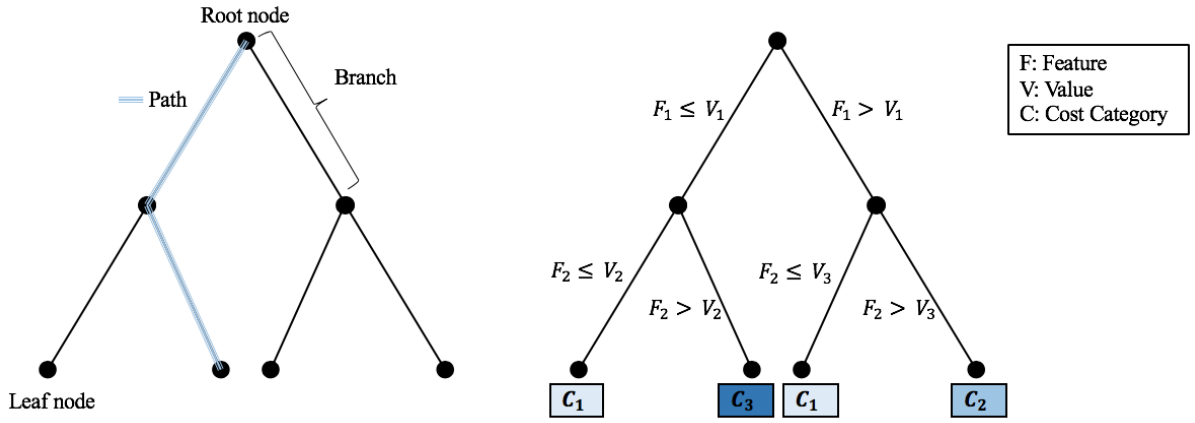

Figure 7.2: Classification tree of depth two with two features

The depth and the amount of branches in a tree may vary. A more complex tree will give more detailed classification of the slug severity value, but may also increase the possibility of the tree being overfitted. The depth and chosen explanatory variables should therefore be evaluated and chosen for each specific case. It is important to emphasize that the method is a simplification of the real values. Building a tree where all observations fall into one of a number of different categories will only yield an expected cost or risk of slugging for the given observation, which will deviate from the actual risk.

A simple example of how certain operation areas can be avoided based on the structure of the CT, is given in the following. Looking at Figure 7.2, the second leaf node from the left has a cost of slugging of C_3. In this example, this cost is too high for the decision maker, and the associated operation point should be avoided. That is, the combination of Feature 1 (F_1) being smaller than Value 1 (V_1) and Feature 2 (F_2) being greater than Value 2 (V_2) is not allowed. This implies that the path leading to this leaf node must be avoided. Figure 7.3 shows another visualization of the operation areas based on the CT presented in Figure 7.2.

Turning the inequality signs in the constraints related to the relevant path defines feasible areas for the given variables. At least one of the corresponding branch inequalities in the optimization model has to be satisfied to avoid the area with cost of slugging C_3. These branch inequalities derived from the relevant path in Figure 7.2 are formulated in constraints (7.22) and (7.23). When one or both of the constraints are satisfied, described by constraint (7.24), the field is operating in the areas with cost of slugging C_1 or C_2 in Figure 7.3.

$$F_1 > V_1 x_1 \tag{7.22}$$

$$F_2 \leq V_2 + M(1 - x_2) \tag{7.23}$$

$$x_1 + x_2 \geq 1 \tag{7.24}$$

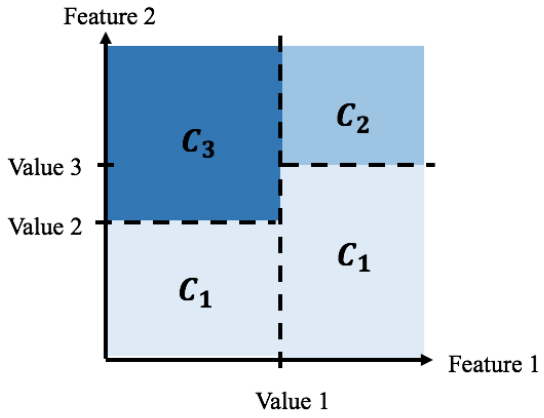

Figure 7.3: Operation areas for a classification tree with two features

7.2.3 Formulation of a Classification Tree Model

A formulation of a CT that models operation areas based on its input variables is presented in this section. The model returns a target variable representing the cost of the slugging category based on the chosen operation area. It should be noted that other target variables in a wide variety of applications also can be modelled with this methodology. In addition, only small adjustments would be required in order to formulate a regression tree rather than a classification tree.

Sets

$\mathcal{F}$	Set of paths in the CT, indexed f
$\mathcal{B}_f$	Set of branches in the CT in path f, indexed b
$\mathcal{B}_f^G$	Set of branches in the CT in path f described with a $>$ constraint, $\mathcal{B}_f^G \in \mathcal{B}_f$, indexed b
$\mathcal{B}_f^L$	Set of branches in the CT in path f described with a $\leq$ constraint, $\mathcal{B}_f^L \in \mathcal{B}_f$, indexed b

Parameters

V_b	Value of variable representing the decision variable of branch b in the CT
M_b	Maximum value of the decision variable of branch b in the CT
C_f	The cost of slugging in path f
ϵ	A small positive value used to represent strict inequalities in the formulation

Variables

c	The cost of slugging in the solution
δ_f	1 if the path f is the path of the current operation area, 0 otherwise
x_b	1 if the constraint defined by branch b in the CT is broken, 0 otherwise
w_b	Variable representing the decision variable of branch b in the CT

Output

The output from the CT is the cost associated with the current operation area. The cost is calculated from predetermined cost categories of slugging associated with a given path.

$$c = \sum_{f \in \mathcal{F}} C_f \delta_f \tag{7.25}$$

Path constraints

Constraints (7.26) force δ_f to one if none of the branches in the given path are broken. Constraints (7.27) force δ_f to zero if at least one of the branches in the given path are broken. The operation area only corresponds to one path, since no areas are identical, as expressed in equation (7.28).

$$\delta_f \geq 1 - \sum_{b \in \mathcal{B}_f} x_b, \quad f \in \mathcal{F} \tag{7.26}$$

$$\delta_f \leq 1 - \frac{\sum_{b \in \mathcal{B}_f} x_b}{M_b}, \quad f \in \mathcal{F} \tag{7.27}$$

$$\sum_{f \in \mathcal{F}} \delta_f = 1 \tag{7.28}$$

Branch constraints

Constraints (7.29) describe *greater than or equal to* constraints in the CT. The constraint has to be satisfied for a branch b to be broken. Constraints (7.30) describe *less than or equal to* constraints, where similarly a constraint has to be satisfied for the branch b to be broken.

$$w_b \leq V_b + M_b(1 - x_b), \quad f \in \mathcal{F}, b \in \mathcal{B}_f^G \tag{7.29}$$

$$w_b \geq (V_b + \epsilon)x_b, \quad f \in \mathcal{F}, b \in \mathcal{B}_f^L \tag{7.30}$$

Variable bounds

$$c \geq 0 \tag{7.31}$$

$$\delta_f \geq 0, \quad f \in \mathcal{F} \tag{7.32}$$

$$x_b \in \{0, 1\}, \quad f \in \mathcal{F}, b \in \mathcal{B}_f \tag{7.33}$$

$$w_b \geq 0, \quad f \in \mathcal{F}, b \in \mathcal{B}_f \tag{7.34}$$

Chapter 8

Implementation

Implementation of the models and methods used in this thesis are presented in this chapter. An overview of the components of the framework used is given in Section 8.1. Section 8.2 presents the method of hyperparameter tuning for multivariate adaptive regression splines (MARS) models and classification trees (CTs), respectively. Further, Section 8.3 presents the method used for scenario generation in the stochastic optimization models, and lastly, Section 8.4 assesses well model uncertainty.

8.1 Framework

Historical data from petroleum production fields is used to develop the data-driven models. The given case in this thesis is presented in Chapter 9, but the framework can be applied to different fields. The pipeline from data to output for decision support is presented in Figure 8.1. The different components are summarized in the list below.

- A: Measurements from the petroleum production field recorded by sensors and other instruments. The measurements from the data source used in this thesis are stored in comma separated values (.csv) files.
- B: The well models use data measurements as input to MARS models to model relationships between choke opening, pressure and flow rates. The well models consist of the oil flow rate models and the wellhead pressure models.
- C: The pipeline pressure models use data measurements as input to MARS models to model the pressure drop in pipelines.
- D: The slug severity models use a slug indicator developed by Solution Seeker and other data measurements to model CTs. The trees describe operation areas with regards to risk of slugging.

While the indicator from Solution Seeker is based on historic data and indicates the risk of slugging for a given system at a given time, the slug severity models are predictive.

- E: Scenarios used in the stochastic optimization model are generated from both of the data-driven MARS models, including the well models and the pipeline pressure models. These models describe the relationships between physical parameters in the system. The stochastic optimization model maximizes average oil production over all scenarios.

- F: Scenarios used in the bi-objective stochastic optimization model with conditional value-at-risk (CVaR) are generated similarly as for the stochastic optimization model. The model maximizes average oil production from the scenarios and minimizes the CVaR with regards to the gas capacity constraint.

- G: Scenarios used in the bi-objective stochastic optimization model with risk of slugging are also generated similarly as for the two other optimization models. The slug severity models are used to model the risk of slugging for the different scenarios. The model maximizes average oil production from the scenarios while minimizing the average slugging cost.

- H: Constraints limiting the feasible area and the initial operating point are problem specific inputs that are needed in the optimization models.

- I: The output from the optimization models is explicit recommendations to the production engineer regarding choke adjustments. The recommendations serve as decision support.

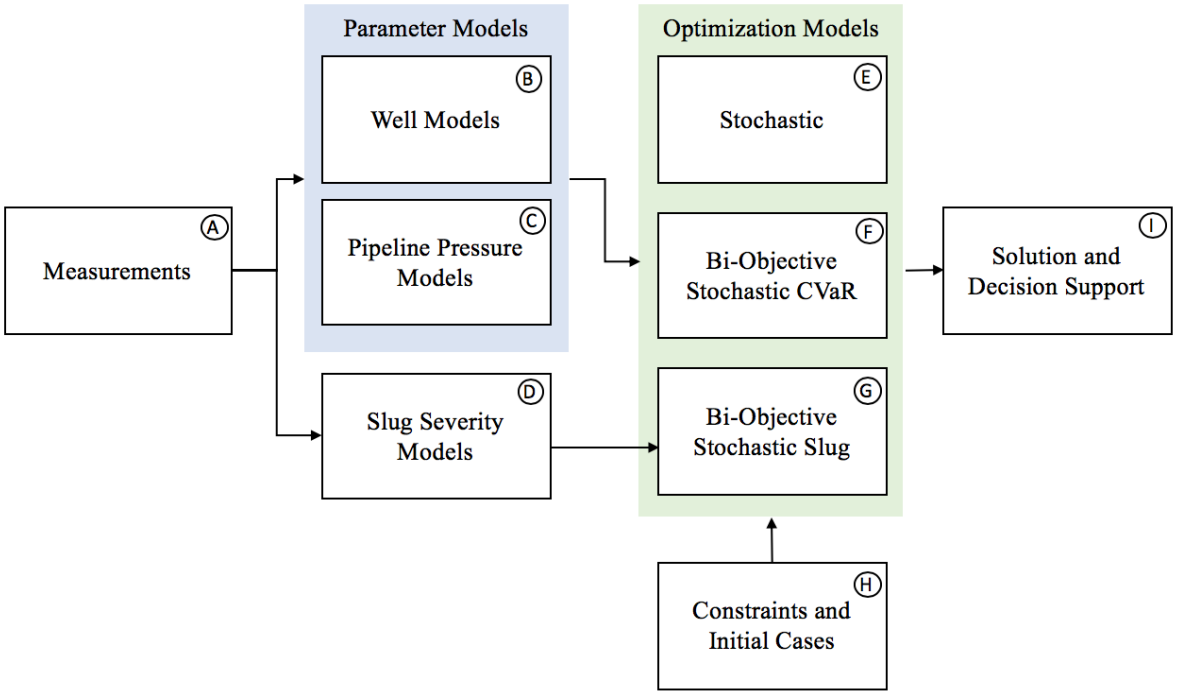

Figure 8.1: Pipeline of the solution framework

Pycharm version 2018.1.4 is used for the implementation of the optimization models. The dataset is read into a Python script from an excel.csv-file, and all scripts are written in Python. The data is preprocessed before it is used in the optimization model, as described in Section 9.2. The data-driven models are imported to the optimization file, where parameters from and structures of the models are used in the optimization. Gurobi Optimizer is used to execute the mixed integer linear programming (MILP) problems by using the library Gurobipy in Python.

MARS Models Implementation

The well models and pipeline pressure models are built using the Pyearth library for MARS models in Python. The hinge functions described by maximum functions cannot be directly treated by the solver, as the maximum functions provided by Gurobi only enables the comparison of two single variables with no constants. In this thesis, the linearization of max expressions formulated in Section 7.1.3 is used in the implementation.

CT Models Implementation

The CT, used to model the slug severity, is implemented in Python using the Scikit-learn library, which is the machine learning library of Python. A tree object is imported from the library, and the DecisionTreeClassifier function is used to create the CT. The slug severity models are used directly in the optimization by transforming paths into linear constraints, as formulated in Section 7.2.3.

8.2 Hyperparameter Tuning

A hyperparameter is an external parameter that is set before the learning process begins. Machine learning models have different hyperparameters as input, and these parameters should be set at reasonable values. Setting the hyperparameters can be done with cross-validation on the training data.

8.2.1 MARS

Selected important hyperparameters for MARS models are described in Table 8.1.

Table 8.1: Description of selected important hyperparameters for MARS models

Hyperparameter	Description
Max terms	The maximum number of terms generated by the forward pass.
Max degree	The maximum degree of terms generated by the forward pass.
Penalty	A smoothing parameter used to calculate GCV. The parameter is used in the pruning pass and to determine whether to add a hinge function or linear basis function during the forward pass.
Min span	The minimal number of data points between knots.

The max degree parameter will in this case be set to one in order to get a linear formulation, as described in Section 7.1.3. The parameter tuning is therefore based on max terms, penalty and min span.

GridSearchCV is a grid search strategy used as decision support to find the hyperparameters. GridSearchCV performs an exhaustive search over specified parameter values for an estimator, where the parameters are optimized by a cross-validated grid search over a parameter grid. Before the grid search is performed, the range for the parameter grid search is set as described in Table 8.2. The same range is applied when performing grid search on the different MARS models.

Table 8.2: Hyperparameter search space for MARS models

Hyperparameter	Symbol	Search space
Max terms	m	[4,5,...,100]
Penalty	p	[1,2,...,100]
Min span	s	[1,2,...,30]

A detailed explanation of k-fold cross-validation is found in Section 4.1.1. The grid search uses 5-fold cross-validation, which implies that the data is divided into 5 sets, where one of the sets are used as testing sets in each iteration. The datasets that are not used for testing in an iteration are used for training.

8.2.2 Classification Trees

Similarly to MARS models, the decision tree classifier has different hyperparameters as input, and these parameters should be set at reasonable values. Selected important hyperparameters are described in Table 8.3

Table 8.3: Description of selected important hyperparameters for classification trees

Hyperparameter	Description
Criterion	The function to measure the quality of a split. The mean squared error and mean absolute error are possible criteria.
Splitter	The strategy used to choose the split at each node. One possible choice are ”best”, which for each node considers all the features and chooses the best split. The other choice is ”random”, which considers a random subset of features to choose the best split.
Max depth	The maximum depth of the tree.
Max leaf nodes	The maximum number of leaf nodes in the tree.
Min samples leaf	The minimum samples required in a leaf node. A split point at any depth is only considered if it leaves at least this number of training samples in each of the left and right branches.
Min samples split	The minimum number of samples required to split an internal node.

The criterion is used with the default value of mean squared error, and the splitter with the default of choosing the best split at each node. Maximum leaf nodes has the same purpose as maximum depth of the tree. A maximum depth of n produces a maximum of 2^n leaves. It is therefore decided to focus on the maximum depth of the tree, and not study the maximum leaf nodes. The minimum samples split is

similar to minimum samples leaf. The main difference between the two is that minimum samples leaf guarantees a minimum number of samples in a leaf, while minimum samples split can create arbitrary small leaves. As it is preferable to control the number of samples in a leaf rather than the minimum samples split, the latter is set to its default value of 2. The parameter tuning will therefore be focused on the maximum depth of the tree and the minimum samples leaf, hereby referred to as depth and min samples.

Similarly to the MARS models, the optimal hyperparameters are found by refining a grid over the hyperparameter space, and using GridSearchCV to find these parameters. 5-fold cross-validation is performed on the hyperparameter search presented in Table 8.4.

Table 8.4: Hyperparameter search space for classification trees

Hyperparameter	**Search space**
Maximum depth	[3,4,...,10]
Minimum samples leaf	[1,2,...,5,10,...,300]

8.3 Scenario Generation

The MARS models, including the well and pipeline pressure models, are used as scenarios in the optimization model. The scenario generation technique is inspired by ensemble learning, which is a method that combines the predictions from multiple machine learning models. Ensemble learning is used to improve prediction performance of a model and reduce the likelihood of selecting a poor model. Different learning algorithms can be used to create an ensemble of models, but the ensemble can also be created by using the same learning algorithm on different training datasets. The ensemble of models is often used to obtain one robust machine learning model, often based on the average prediction of the ensemble.

In this thesis, a type of ensemble models are used as scenarios in the optimization. The procedure is illustrated in Figure 8.2. Two MARS models per well and one MARS model per manifold make up a scenario. That is, a scenario consists of an oil flow rate and a wellhead pressure model per well, and a pipeline pressure model per manifold. The scenarios are created by resampling the training dataset used to train the models. Specifically, a split of the data, where a subset of the data is used to fit the models, is performed for each model associated with a scenario. Figure 8.2 depicts the grey area as the data, consisting of random samples, that is not used to fit a model. It is assumed that there is no correlation between models in the same scenario nor between scenarios.

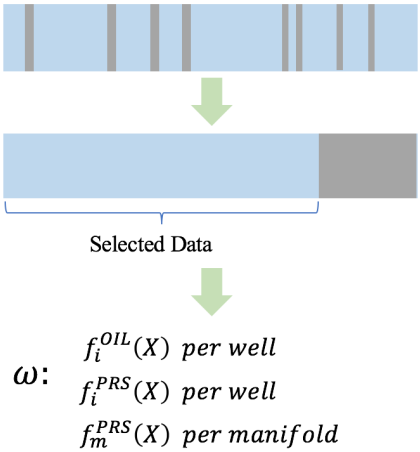

Figure 8.2: Scenario generation procedure

The procedure results in sets of different MARS models for each scenario, thus each scenario yields different parameters for oil rates and pressure relations. The different parameters generated cover the model uncertainty. By resampling datasets, the models use different parts of the training data, and will make different errors. It is assumed that the scenarios will make better predictions and be more robust when they are combined. An important goal of the technique is to avoid overfitting to specific samples of the training data by using a sufficient number of scenarios.

There are different ways to resample the training dataset, and a simple method using a random split of the data is used in this thesis. It is emphasized that there are several other ways to generate an ensemble of MARS models. An example is a Bayesian approach to MARS fitting, which applies a probability distribution over a space of possible MARS models (Denison et al., 1998).

8.4 Well Model Uncertainty from Variance

As described for the well models in the stochastic formulation in Section 6.2, uncertainty in the models is taken into account by restricting the choke openings of the wells. It is assumed that the restrictions on choke openings implicitly restrict other variables, such as flow rates and pressures. Only data samples where the choke opening lies within its limits are used for training of the oil flow rate model and the wellhead pressure model.

Equation (4.3) from Section 4.1.1 is repeated in equation (8.1) to point out the scope of the uncertainty limits presented in this section. The equation states that the error of a model stems from its bias, variance and an irreducible error. The uncertainty limits developed only take the variance of the machine learning models into account, and thus the level of bias in a model will not affect the limits.

$$Err(x) = Bias^2 + Variance + Irreducible\ Error \tag{8.1}$$

A general method of deciding the allowed interval of change for the current choke opening is proposed. The oil flow rate models are repeated in equations (8.2). The uncertainty in the oil flow rates determines the choke limits of the wells. In order to study the well model uncertainty, each well is studied separately and decoupled from the rest of the production system. When isolating a well, it is assumed that $\Delta p_{i\omega}^{D}$ is zero, and the pipeline pressure models are excluded since they connect the wells. The wellhead pressure models can thus be expressed as in equations (8.3).

$$\Delta q_{oi\omega} = f_{i\omega}^{OIL}(\Delta p_{i\omega}, \Delta u_i), \quad i \in \mathcal{I}, \omega \in \Omega \tag{8.2}$$

$$\Delta p_{i\omega} = f_{i\omega}^{PRS}(\Delta u_i, \Delta q_{oi\omega}, \Delta q_{gi\omega}, \Delta q_{wi\omega}), \quad i \in \mathcal{I}, \omega \in \Omega \tag{8.3}$$

A grid of choke openings is defined in the range [0,100]. The second-stage optimization problem for an isolated well is solved for each fixed choke opening in 100 iterations, and the average and standard deviation of the oil flow rate are calculated. The output is thus a sorted set of choke openings with associated mean values and standard deviations of oil flow rates.

The algorithm in Appendix B shows the method of finding the upper and lower limit for a choke, or uncertainty bands, of a well. The coefficient of variation (CV) is used to set the limit for the uncertainty bands. CV is a standardized measure of dispersion, defined as the ratio of the standard deviation to the mean. The measure shows the extent of variability in relation to the mean for the given choke opening. Thus, using the same value of the measure, higher choke openings are allowed a higher standard deviation.

Models with high variance yield strict limits, while models with low variance are able to make bigger adjustments from the current choke values. The uncertainty of the data source will vary from well to well, and from case to case. The sources of uncertainty are discussed in Section 9.4, and the uncertainty study is performed on a case in Section 10.3.3.

Chapter 9

Case Study

To evaluate the optimization model, input data is needed. For this thesis, steady-state production data from an oil field on the Norwegian Continental Shelf is used. The dataset was generated by Solution Seekers steady-state classification algorithm, and is from the time period of January 2014 to January 2018. The dataset is split into a training and test dataset. The training data contains 1998 samples from January 2014 to January 2017. The test data contains 639 samples from January 2017 to January 2018. Each sample in the dataset represents measurement data taken over a period of steady-state production. The measurements include uncertainty, so for each measurement in the dataset, the mean and the standard error of the mean (SEM) is given. In the following chapter, the specific field structure for this case is presented in Section 9.1. The preprocessing of the dataset is explained in Section 9.2, while Section 9.3 presents some observed data characteristics for the specific field. The sources of uncertainty are discussed in Section 9.4. Finally, Section 9.5 presents the initial cases used for analysis in Chapter 11.

9.1 Field Structure

The production system consists of eleven wells producing oil, gas and water through two risers. Figure 9.1 shows the production system topology of the specific field. Wells with a high gas-oil ratio (GOR) produce via the gas riser to the 1st stage separator. The other wells produce via the oil riser to 2nd and 3rd stage separators. Only the oil riser and the connected oil wells are studied in this case. This is mainly because slugging is a more common phenomena in the oil riser, and it is therefore more interesting to study this pipeline. Well 8 (W08) and Well 9 (W09) are connected to the same manifold, but only W08 is routed to the oil riser. Well 4 (W04), Well 5 (W05), Well 6 (W06) and Well 7 (W07) are connected to the same manifold and routed to the oil riser. Well 1 (W01), Well 2 (W02) and Well 3 (W03) are connected to the same manifold, with Well 1 and Well 3 routed to the oil riser. The manifold structure for the oil riser is illustrated in Figure 9.2. It is further assumed that a given well is routed to exactly one pipeline, and that rerouting is not allowed. This is also a valid assumption based on the dataset given, as there has been very few cases of routing.

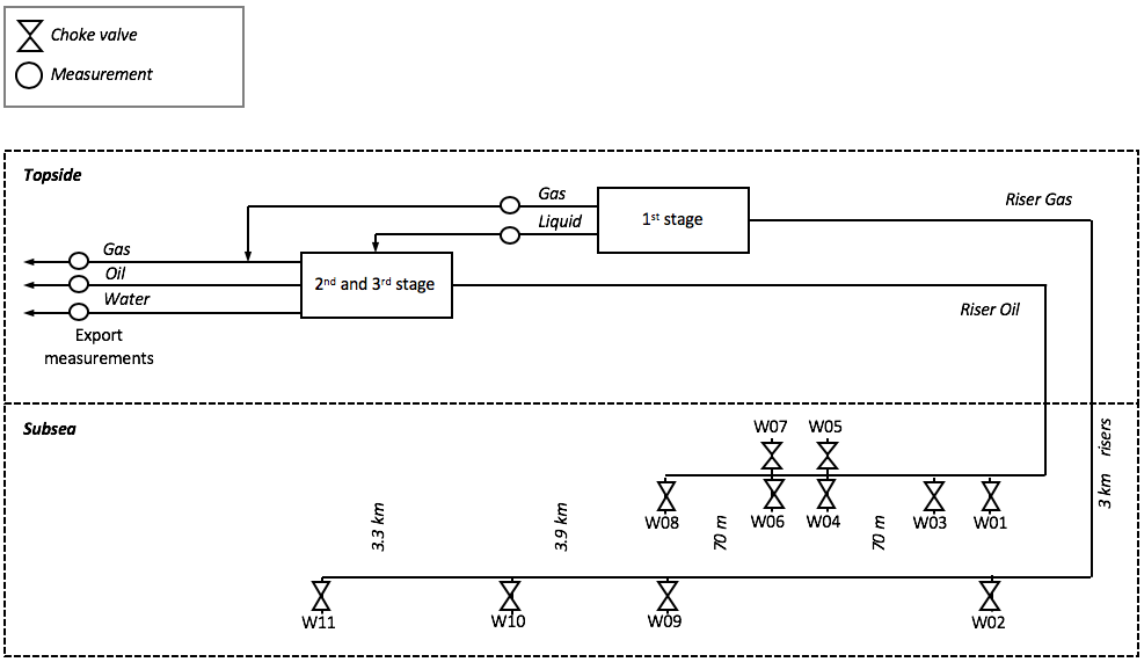

Figure 9.1: Production system topology

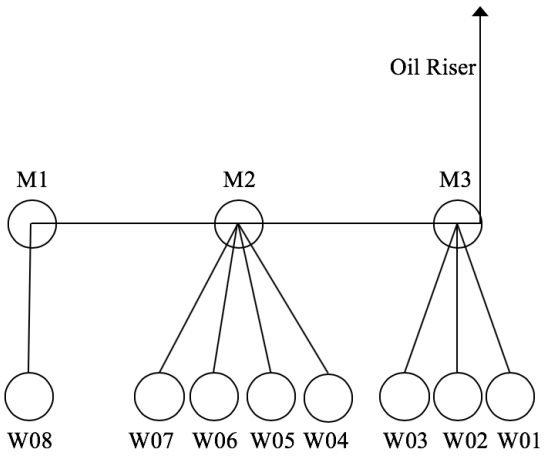

Figure 9.2: Structure of manifolds and wells for oil riser

The dataset is comprised of a set of measurement from the wells. The sensor placement for the production system is shown in Figure 9.3. The subsurface measurements are not used, and therefore not illustrated. The relevant parameters used are the wellhead pressure (PWH) and the downstream choke pressure (PDC) as formulated in the optimization models in Chapter 6. As the upstream choke pressure (PUC) is assumed to be approximately equal to the wellhead pressure, it is not used. The MPFM is placed on the wellhead, and gives the oil (QOIL), gas (QGAS) and water (QWAT) flow rates. The flow rates are measured in the unit Sm^3/h. Each of the measurements has a time stamp, which is useful for weighting in the machine learning models. The choke openings are also included in the measurements.

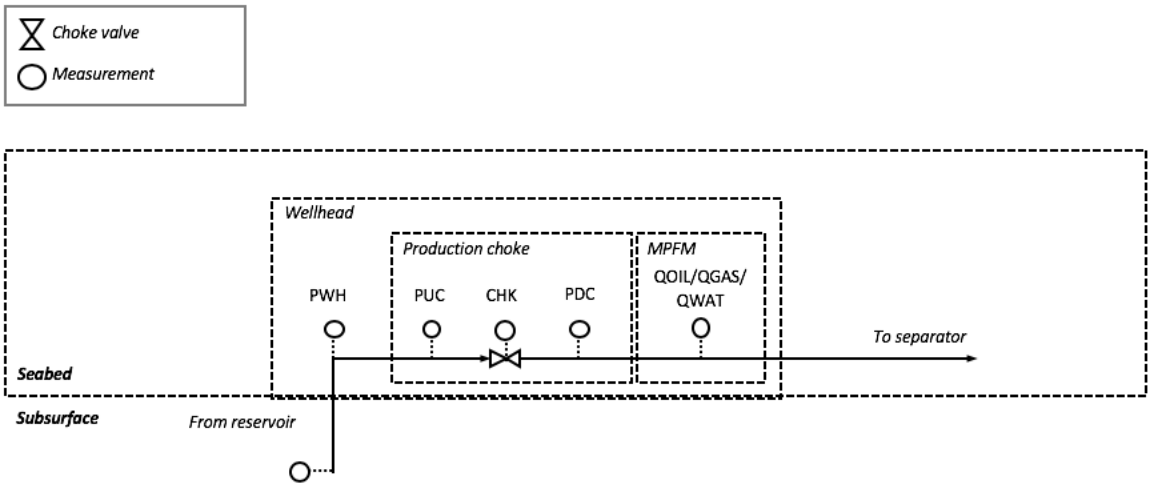

Figure 9.3: Sensor placement in production system

9.2 Data Preprocessing

The dataset is preprocessed before use. Data cleansing steps are described in Section 9.2.1, and data transformation steps are described in Section 9.2.2. Data cleansing involves detecting and correcting corrupt or inaccurate measurements, while data transformation involves transforming data measurements into appropriate forms.

9.2.1 Data Cleansing

The dataset contains uncertainty, and some of the sensors are failing. The dataset is therefore cleansed before used for modelling. Some of the issues assessed include the handling of missing data, removing outliers and reducing noise. In general, when the data values seem highly unlikely or impossible, they should be removed or replaced so they do not affect the data models and analysis. In order to keep as much original data as possible, data is only removed or replaced per well, and not for all wells during a given time interval.

To start with, data is sorted after the start of the time interval, making it possible to see the change over time. Values for choke openings are then preprocessed, setting openings with a negative value to zero and values greater than 100 to 100, as this is the interval of a choke opening. Furthermore, choke openings less than five are assumed to be closed, as this would be an unlikely low value. That is, these choke openings are also set to zero.

For the flow rates, the oil, gas and water rates are set to zero if they have a negative value, since a negative value is not physically possible. Inaccurate oil, gas and water measurements are also removed. These are cases such as when the choke opening is closed and the production output is greater than a certain value, and when the choke is open and there are no values for the output production data. In these cases,

the values are set to NaN for the production output and choke opening, respectively.

Continuing, outliers from the dataset are removed or replaced. Spikes and outliers are identified by plotting the measurements with correlated factors. Spikes in pressure are removed by studying pressure nearby in the dataset, and identifying unrealistic measurements. A specific case related to pressure outliers is for Well 3, which was routed to the gas riser in a given period. This caused a jump in the measurements. Measurements of production rates are also removed by using the standard error of the mean (SEM), and removing points with a SEM value above a certain percentage of the mean of all SEM values for the measurement. Furthermore, the slug severity indicator in the dataset is supposed to have a value between zero and one, thus all slug severity indicators above one are set equal to one, and negative values are set to zero.

9.2.2 Data Transformation

In addition to the data cleansing, some more preparation of the data must be done. Some variables introduced in the optimization model in Chapter 6 do not have corresponding measurements in the data. Therefore, some measurements were derived from other measurements and constructed as new parameters in the dataset. Examples of such measurements are GOR and water cut (WCT) for each well, in addition to oil, water, gas and downstream choke pressure mean for each manifold.

The MARS models are modelled as delta models, and it is therefore necessary to calculate the difference between measurements before training of the models. In addition, the well models are local, and are only trained on the data that lies inside certain limits as described in Section 6.2. Data measurements outside of these limits are therefore excluded before training of the given well model.

A linear weighting of the data points with respect to time is applied to take into account the development of the reservoir. As oil is extracted from the reservoir, this has implications for the dynamics of the reservoir. It is therefore argued that the more recent measurements are of higher relevance to the current operation point. The development of the reservoir is only relevant for the well models. The pipeline pressure models therefore have no linear weighting.

As discussed in Section 7.1.3, multiplicative interaction terms in the MARS models are avoided in order to obtain a linear formulation, and α-weighted variables that can model linear interactions are applied. A pre-study on the given data source can be done in order to find which variables that have an interaction and the number of α-weights to include for the given interaction.

Slug severity indicators within a certain interval are assigned a corresponding cost. These cost categories are set based on conversations with Solution Seeker. The cost of slugging can be estimated to be 5 % of the production. As an approximation, the average production rate from the dataset is used. The average production rate for the oil riser is 112 Sm^3/h. 5 % of this rate is estimated to be the cost of slugging for the category high slug, and 4% for the low slug category. The intervals and the corresponding categories and costs are presented in Table 9.1.

Table 9.1: The cost of slugging for different operation areas

Operation area	Indicator Interval	Cost
Safe	0.0 - 0.6	0
Low Slug	0.6 - 0.8	4.5
High Slug	0.8 - 1.0	5.6

9.3 Data Characteristics

Plotting of parameters as a time series gives interesting observations. By plotting the choke opening for each well per time, it is observed that the wells are operated within an interval that does not cover the whole operation space. The choke opening is then often either within this given interval or closed. Using Well 6 as an example, one can see in Figure 9.4 that there are many data points where the well is operated around choke openings from 40 to 60 or closed, but not for other levels.

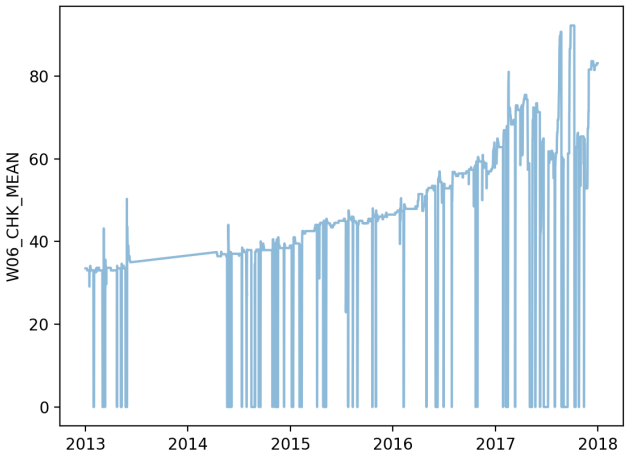

Figure 9.4: Choke opening over time for Well 6

Figure 9.5 shows a scatter plot of choke openings and oil flow rates for Well 6, and similar plots can be generated for the other wells. The vertical lines consisting of scatter points in the figure arise from the operational adjustments of the choke. It is not continuously varied, as observed in Figure 9.4, and it often stays in the same setting for some period.

Figure 9.6 shows the slug severity indicator value over time for an aggregated train and test dataset. It is observed that most of the values are less than approximately 0.4, and there is a significant rise in the values in 2017. This corresponds to the high slug indicator values in the test dataset.

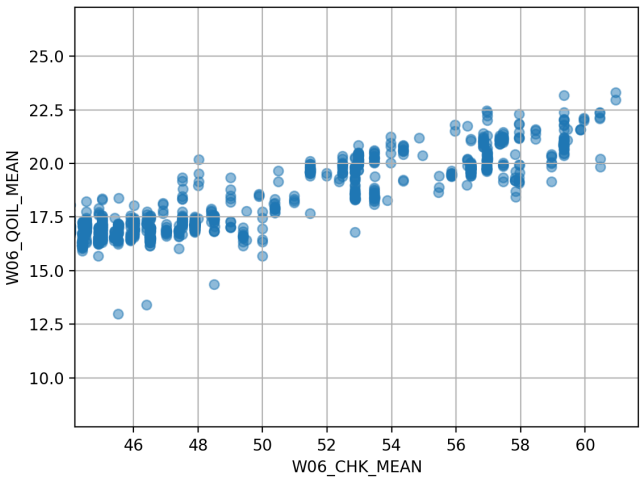

Figure 9.5: Oil flow rates for an interval of choke opening for Well 6

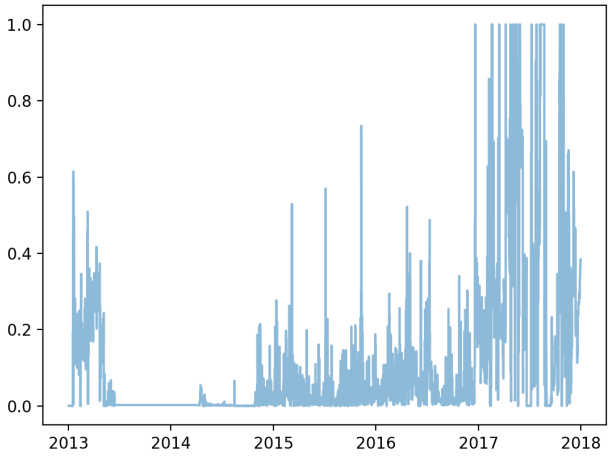

Figure 9.6: Slug severity indicator over time for oil riser

9.4 Sources of Uncertainty

The terms epistemic and aleatoric uncertainty are introduced in Section 4.1.1. Both types of uncertainty are present in the petroleum production optimization problem. Aleatoric uncertainty, or data uncertainty, captures the inherent noise in observations, which in this case is given by the variance in the measurements. Epistemic uncertainty in this application captures the uncertainty in model parameters, and is therefore referred to as model uncertainty.

Data Uncertainty

Data uncertainty stems from the fact that the measurements themselves are uncertain. The multiphase flow meters (MPFMs) can be biased, and the readings always have a connected error. The MPFM

is placed at the wellhead, and the final measurement is based on a combination of measurements from different instruments and sensors. All the instruments have a measurement uncertainty, and it is therefore a total uncertainty associated with the measurement from the MPFMs. Even though the MPFMs improve and become more reliable, measurement uncertainty and failure cannot be avoided. Obtaining more data will not reduce this uncertainty, because the noise is inherent in the data.

Model Uncertainty

Data-driven models built on historical production data are integrated in the optimization problem to predict flow rates and pressures from choke openings. The models may fail to accurately predict due to their model uncertainty. Given infinite data, model uncertainty can be removed. However, as for most real applications, there is limited data available. When machine learning models are built from production data, they will therefore be associated with varying levels of uncertainty. They are approximations of the underlying physics in the production system, and will therefore not be able to model the flow rates and pressures in a way that precisely corresponds to the real world. This would not be possible by using any modelling technique, regardless of complexity and precision.

The uncertainty of the models are affected by the scarcity of the data. The wells have often only been operated at certain settings, implying that the variance is larger in areas that the wells have not been operated. Models will therefore often be accurate in the regions where the data source is rich, and poor when evaluated further away from these operation points. The data often represents a trajectory of an operation setting, and the settings outside might never been measured. These challenges come from the fact that well operators seldom change the choke settings, which does not generate new information to further build the data source. The frequency and range of changes will however vary significantly from field to field.

Experimenting with different choke openings to increase the data source can contribute to attaining more realistic models and reducing uncertainty. There is however a trade-off associated with this, as experimenting with the choke openings in order to gain more information will increase cost when performing the changes.

9.5 Initial Cases

The two initial cases used in the computational study are presented in Table 9.2 and Table 9.3. The general parameters used for both cases can be found in Appendix C. Initial Case 1 represent a real operation setting from the dataset, with some adjustments where measurements are failing. It represents a feasible operation setting that has a low risk of slugging. Initial Case 2 is not a real operation setting from the dataset. The case has the same initial choke openings as Initial Case 1, but the WCT of the wells are assumed to be significantly higher, such that the case starts in an operation area with a relatively high risk of slugging.

Table 9.2: Operation settings with low risk of slugging

Initial Case 1						
Well	**Choke Opening**	**PWH**	**PDC**	**Oil Rate**	**Water Rate**	**Gas Rate**
W1	50.00	61.56	59.10	20.93	8.97	10173.86
W3	37.61	97.94	59.10	11.15	3.72	42295.96
W4	46.94	82.26	59.86	19.05	8.56	47231.81
W5	57.79	81.66	59.86	13.76	11.72	49153.06
W6	50.47	77.81	59.86	19.67	18.15	46760.70
W7	56.60	82.40	59.86	14.51	16.36	41416.32
W8	50.82	88.38	59.95	9.91	1.35	43971.26
Sum	-	-	-	108.98	68.84	281002.99
Total gas capacity						300000.00

Table 9.3: Operation settings with high risk of slugging

Initial Case 2						
Well	**Choke Opening**	**PWH**	**PDC**	**Oil Rate**	**Water Rate**	**Gas Rate**
W1	50.00	61.56	59.10	20.93	71.28	10173.86
W3	37.61	97.94	59.10	11.15	33.45	42295.96
W4	46.94	82.26	59.86	19.05	29.80	47231.81
W5	57.79	81.66	59.86	13.76	11.72	49153.06
W6	50.47	77.81	59.86	19.67	18.16	46760.70
W7	56.60	82.40	59.86	14.51	16.36	41416.32
W8	50.82	88.38	59.95	9.91	9.19	43971.26
Sum	-	-	-	108.98	189.96	281002.99
Total gas capacity						300000.00

Chapter 10

Analysis of Data-Driven Models

The results from the analysis of the data-driven models are presented in this chapter. An analysis and evaluation of the multivariate adaptive regression splines (MARS) and classification trees (CT) are presented in Section 10.1 and Section 10.2, respectively. Section 10.3 discusses considerations that should be taken into account when integrating these data-driven models in optimization, and concludes with the choice of model complexities that are further used in the computational study. The analysis is performed on the case described in Chapter 9. The results in this study are therefore specific for the case, but a similar study could be performed for other petroleum fields or synthetic cases. Initial Case 1 presented in Section 9.5 is used in the generation of the models in this chapter.

10.1 Analysis of the MARS models

Results from the MARS models are presented in the following section. The models are implemented as described in Section 8.1. Section 10.1.1 presents an analysis of α-weighted variables. Further, the results from the hyperparameter tuning are found in Section 10.1.2. Lastly, Section 10.1.3 studies the feature importance of the models.

10.1.1 α-Weighted Variables

The α-weighted variables, as stated in equation (7.8), are studied in this section. The equation for an α-weighted variable, where the sum of α_1 and α_2 equals one, is restated here for convenience:

$$w = \alpha_1 \cdot x_1 + \alpha_2 \cdot x_2 \tag{10.1}$$

Which variables that should be expressed as linear combinations, thus which variables should be set to x_1 and x_2, are decided. The x_1 and x_2 variables are chosen from the input variables of each model. The

suitable number of α-weighted variables to include as input for the different models is also decided.

A representation of the study can be found in the figures in Appendix D. The study is conducted by training the models on a subset of the data in 20 runs, yielding 20 different versions per oil flow rate model, wellhead pressure model and pipeline pressure model. The α-weighted variables are determined by varying both the x_1 and x_2 variables, and the number of α-weighted variables used in each model. The α-values are found by equally partitioning the range of $[0,1]$ at n values, n equal to the number of α-weighted variable for each iteration. An example is seen in the equations (10.2) - (10.4) below, where n equals three.

For simplicity, the figures only show the GCV scores for the x_1 and x_2 variables that on average give the overall greatest generalized cross-validation (GCV) reduction for each model. The results show that the effect of the α-weighted variables vary between the models, and within models between wells and manifolds.

When the α-weighted variables are included in a model, it follows that the number of variables in the model increase, and thus also the solution time. It is therefore important to include these variables only when they give a significant increase in the model accuracy. Based on the analysis of error reduction, it is observed that for only a few number of wells in the oil flow rate models, seen in Figure D.1, and wellhead pressure models, Figure D.2, do the GCV notably decrease. For simplicity of the optimization model, the α-weighted variables are either included or not included for all wells or manifolds associated with one model. They are therefore not included for the well models. On the other hand, the pipeline pressure models, figures D.3-D.5, show a decrease in GCV for all three manifolds. α-weighted variables for these models are therefore included.

The number of α-weighted variables to include in the pipeline pressure models is set to three. These variables will hereby be included as input variables when these models are referenced, if not stated otherwise. Adding the three variables will not yield too complicated models, but still reduce the GCV values significantly. The α-weighted variables, represented by w, used in the pipeline pressure models for Manifold 1 and Manifold 2 are linear combinations of the manifold pressure Δp_{m+1} and the water flow rate Δq_w:

$$w_1 = 0.25\Delta p_{m+1} + 0.75\Delta q_w \tag{10.2}$$

$$w_2 = 0.50\Delta p_{m+1} + 0.50\Delta q_w \tag{10.3}$$

$$w_3 = 0.75\Delta p_{m+1} + 0.25\Delta q_w \tag{10.4}$$

For Manifold 3, which does not have manifold pressure as an input variable, the α-weighted variables are linear combinations of the oil flow rate Δq_o and the water flow rate Δq_w:

$$w_1 = 0.25\Delta q_o + 0.75\Delta q_w \tag{10.5}$$

$$w_2 = 0.50\Delta q_o + 0.50\Delta q_w \tag{10.6}$$

$$w_3 = 0.75\Delta q_o + 0.25\Delta q_w \tag{10.7}$$

10.1.2 Hyperparameter Tuning

The hyperparameter tuning for the MARS models are performed as described in Section 8.2.1. The optimal hyperparameters from the grid search are presented in Table 10.1 and Table 10.2. They describe the MARS models from the grid search that yield the greatest reduction of overall error, corresponding to the optimal complexity level. R^2 and MSE are used as evaluation metrics in the grid search, and the values of the metrics can be found in the tables. As described in Section 4.1.1, a lower MSE and a higher R^2 indicate more accurate models. The MSE scores vary between the different models, while all the models have high R^2 scores. In addition, it can be found from the tables that the values of the parameters vary greatly between wells, manifolds and models. There may be several reasons for this, one being the varying data source between wells and between manifolds, leading to different parameters that best fit the data.

Table 10.1: Hyperparameters and test scores for the well models found from grid search

	Oil flow rate					Wellhead pressure				
Well	**MSE**	R^2	**Penalty**	**Max terms**	**Min span**	**MSE**	R^2	**Penalty**	**Max terms**	**Min span**
W1	1.52	0.94	35	12	1	2.98	0.88	7	30	1
W3	0.32	0.95	20	10	3	1.19	0.86	20	30	10
W4	0.94	0.94	6	6	5	3.37	0.68	45	12	10
W5	0.36	0.98	7	14	2	0.61	0.96	3	40	1
W6	0.65	0.97	2	12	2	1.79	0.90	50	30	20
W7	0.84	0.95	1	12	1	1.35	0.92	15	20	1
W8	1.59	0.91	25	8	20	1.16	0.90	35	40	1

Table 10.2: Hyperparameters and test scores for the pipeline pressure models found from grid search

			Hyperparameters		
Manifolds	**MSE**	R^2	**Penalty**	**Max terms**	**Minspan**
M1	1.51	0.98	1	14	2
M2	1.18	0.98	1	18	2
M3	2.30	0.46	9	18	25

10.1.3 Feature Importance

The input variables of the data-driven models are presented in Chapter 6, and the equations are restated in equations (7.1) - (7.4) in Section 7.1.1. The choice of input variables is based on knowledge of how physical relations interact in petroleum production systems, in addition to discussions with Solution Seeker. As the data source varies from well to well and manifold to manifold, the models for each well and each manifold may differ in what inputs are found to give models with the highest accuracy. If new data becomes available, the feature importance of the variables might change. Furthermore, as previously discussed in Section 7.1, the MARS technique performs automatic variable selection, including important variables and excluding unimportant ones. Based on this, all relevant input variables are included.

Figures 10.1 - 10.4 show feature importance for the well models and pipeline pressure models. The feature importance per well and manifold are stacked, and the sum of feature importance per well or manifold sum up to one. The measures are based on the *number of subsets* criterion, which counts the number of model subsets that include the variable. Variables that are included in more subsets are considered more important. Subsets are defined as the subsets of terms that are generated by the backward pass in the MARS algorithm.

For both the oil flow rate model in Figure 10.1 and the wellhead pressure model in Figure 10.2, it is seen that choke is the most important input, yet the other features also play an important role. It is observed that the feature importance vary drastically from well to well, which can be explained by the large variation in quality of the data source. As described in Section 9.2.1, the data set contains several failing measurements, especially from the multiphase flow meters (MPFMs), which are individual for each well.

Figures 10.3 and 10.4 illustrate the feature importance for the pipeline pressure models for Manifold 1 and 2, and Manifold 3, respectively. For Manifold 1 and 2, it is observed that the weighted variable w_3, a linear combination of the manifold pressure and water flow, is the most important explanatory variable.

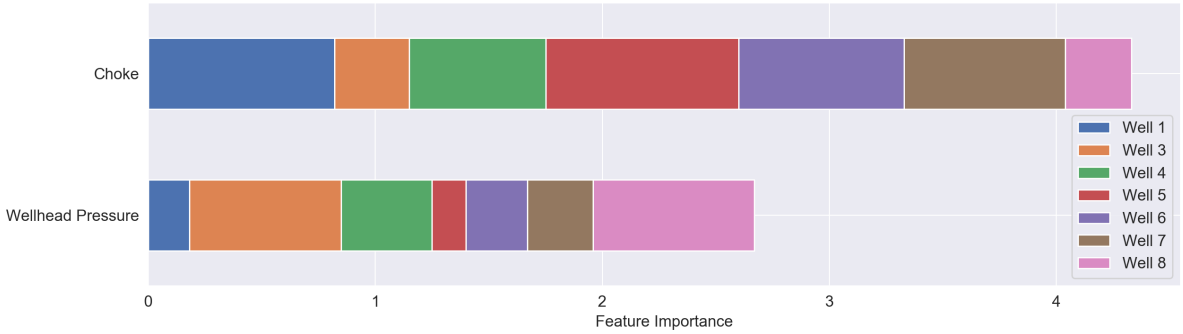

Figure 10.1: Feature importance for oil flow rate model

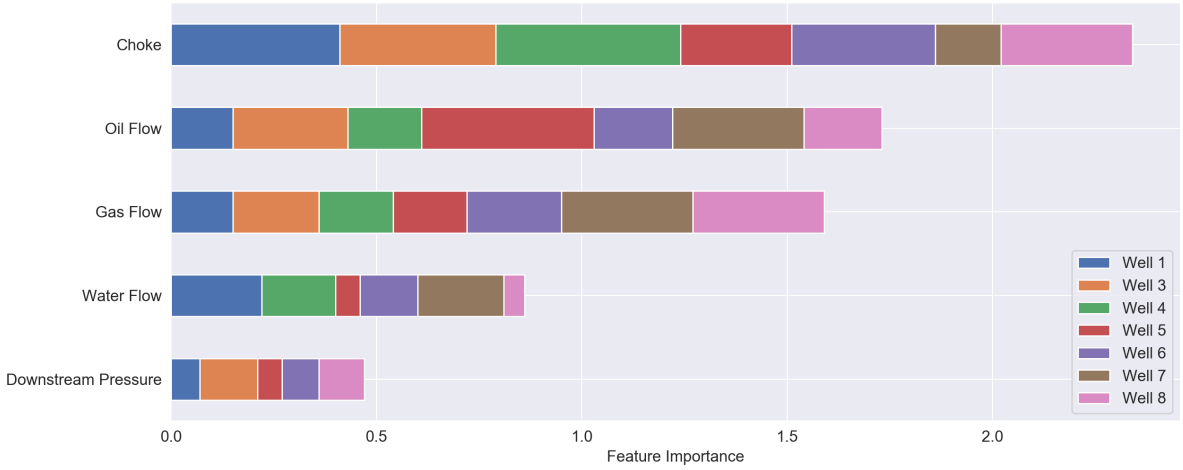

Figure 10.2: Feature importance for wellhead pressure model

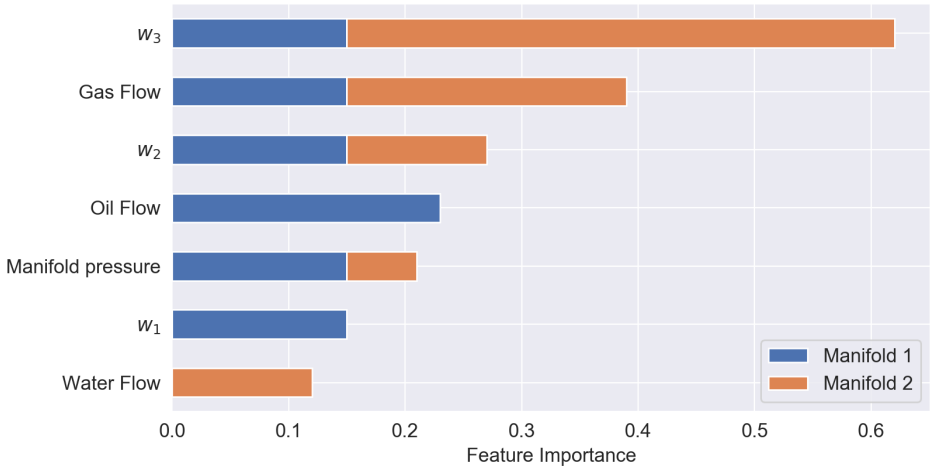

Figure 10.3: Feature importance for pipeline pressure model for Manifold 1 and 2 with α-weighted variables

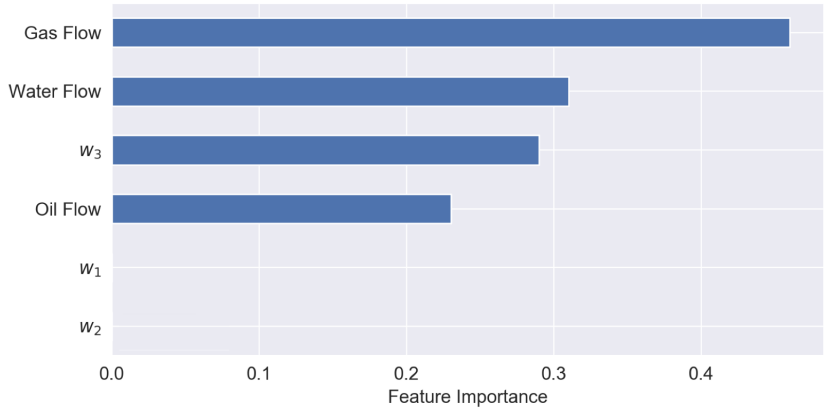

Figure 10.4: Feature importance for pipeline pressure model for Manifold 3 with α-weighted variables

10.2 Analysis of the Classification Tree

Results from the CT are presented in this section. As described in Section 7.2, the purpose of the slug severity model is to apply an indicator developed by Solution Seeker to make a predictive slug severity model. The indicator from Solution Seeker is based on historic data, and indicates the level of slugging for a given system at a given time, while the slug severity model is a predictive model.

In this section, the hyperparameters and input parameters for the CT are studied. The hyperparameter tuning in Section 10.2.1 finds the optimal hyperparameters for the classification tree based on the grid search described in Section 8.2.2, while the input parameters, or features, are studied in Section 10.2.2. As only the oil riser is studied, there is only one slug severity model.

Before the hyperparameter tuning, a new split of the train and test dataset is done, as the test dataset contains most of the data points with high slug severity indicators. Both datasets are merged to an aggregated dataset, and a split where 80% is used for training and the remaining 20% is used for testing is performed.

10.2.1 Hyperparameter Tuning

The hyperparameter tuning is performed with a 5-fold cross-validation, and the best hyperparameters are based on the accuracy score explained in Section 4.1.1. The best hyperparameters from the grid search are a maximum depth of 4 and minimum samples 9. The cross-validation score for this decision tree is 0.967. The performance should also be evaluated based on the test dataset, as this consists of the remaining 20% of the data. The accuracy for the test dataset is 0.965, which means that 96.5% of the labels are classified correctly in the test dataset. The accuracy for the training dataset is slightly higher than for the test dataset, which could be expected as the grid search is fitted to the training data. However, the difference is negligible.

Classification accuracy alone can be misleading for this case, since there is an unequal number of observations in each class. 94% of the data samples have a slug severity indicator classified as safe with regards to slugging. Therefore, one would get a very high accuracy score of 94% if all samples were classified as safe before using the classification tree for the prediction. A confusion matrix is therefore presented in Figure 10.5 for the test dataset, and summarizes the performance of the classification tree. The diagonal elements represent the number of samples for which the predicted label is equal to the true label, while off-diagonal elements are the samples that are mislabeled. It is observed that the diagonal values of the confusion matrix are high, which indicates many correct predictions. The error the classification tree makes mislabeling the actual high slug samples are the most critical errors.

The classification tree with the optimal hyperparameters is presented in Figure 10.6, and an enlarged version can be found in Appendix F. The green leaf nodes represent safe areas, the yellow nodes represent areas with low risk of slugging, and the red nodes represent areas with high risk of slugging. As expected, most of the operation areas are classified as safe.

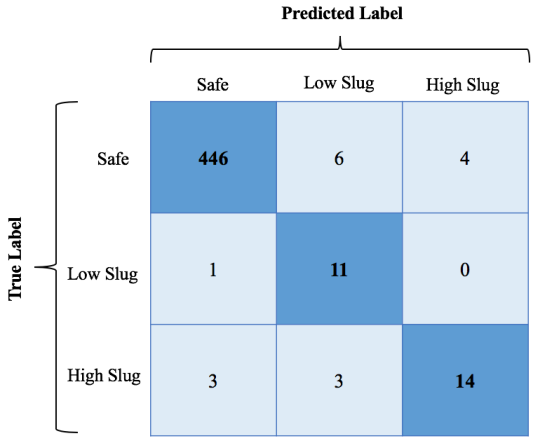

Figure 10.5: Confusion matrix for classification tree on the test dataset

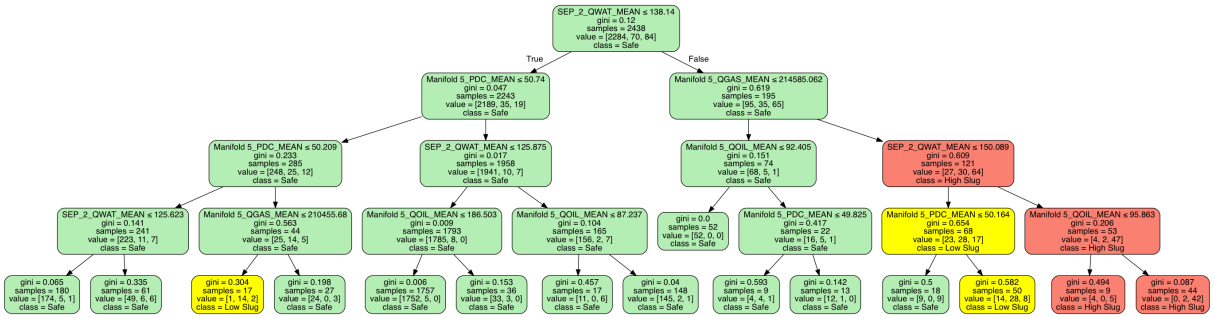

Figure 10.6: Final classification tree

10.2.2 Feature Importance

When setting the input parameters, using measurements per well and measurements for the oil riser are both tested. Using measurements associated with each well in the CT gives poor scores. This might be caused by the difficulty of relating slugging to for example the water rate in a certain well, as slugging is known to be caused in the riser. The CT that uses measurements for the riser therefore expectantly performed better.

Generally, it is common to use many variables in decision tree learning. In this case, however, because of the research conducted on slugging and the limited relevant measurements, the choice of input parameters is based on the factors that are known to contribute to slugging. The variables used in the decision tree are therefore oil, gas and water rates in the riser, in addition to the manifold pressure in the manifold connected to the separator. The variables are presented in equations (7.21).

The importance of the input variables in the final model is presented in Figure 10.7. The importance of a feature is computed as the normalized total reduction of the accuracy criterion brought by that feature.

The figure shows that the water flow rate is the most important feature. This is also what is found in the production systems, as the slug flow described in Section 2.3.2 consists of liquid and gas accumulating. However, all the factors contribute to slugging, and this is reflected in the figure where every input

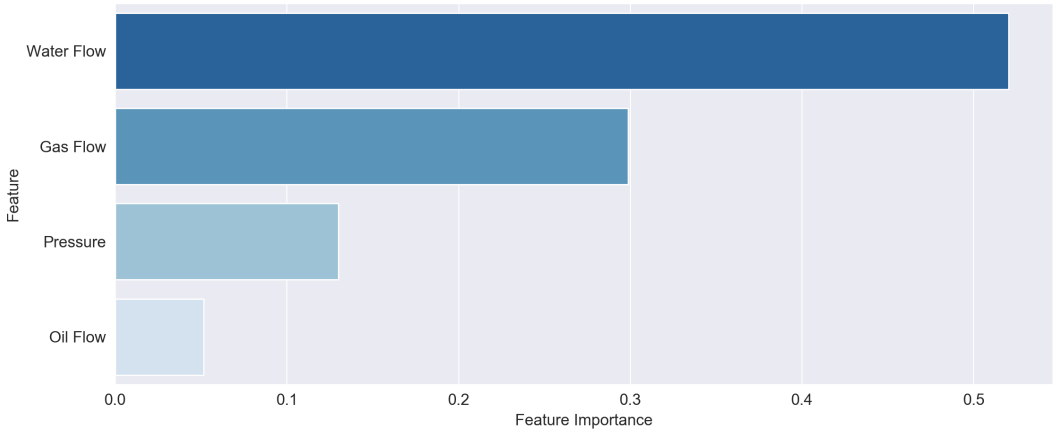

Figure 10.7: Feature importance for the classification tree

variable has an effect on the tree.

10.3 Data-Driven Models in Optimization

The optimal hyperparameters of the models were found in Section 10.1 and Section 10.2. When integrating the optimal models with these hyperparameters in the linear optimization, there are some considerations that must be taken into account. As found in the MARS formulation in Section 7.1.3 and the CT formulation in Section 7.2.3, binary variables are needed to formulate the models. The number of binary variables, which is directly related to the complexity of the models, has a large impact on the solution time of the problem.

The impact complexity has on solution time can be studied for both the MARS models and the CT. However, only the MARS models are studied in this thesis, as they have the greatest impact. While there are two MARS models per well and one per manifold for each scenario, there is only one classification tree per relevant riser. The classification tree can therefore be trained with the optimal hyperparameters, knowing that this will only have a small impact on the increase of binary variables in the optimization model.

In the following, Section 10.3.1 discusses implications of MARS model complexity on solution time. Further, Section 10.3.2 presents three levels of complexity that are used in the computational study, while Section 10.3.3 presents the uncertainty limits related to the three levels.

10.3.1 Complexity and Solution Time

The complexity of a MARS model is here defined as the number of hinge functions in the model. Thus, increasing the number of hinge functions will increase the complexity. Introducing a set $\mathcal{M}_\omega$, describing

the set of MARS models m in a scenario ω, and a set $\mathcal{H}_m$, describing the set of hinge functions h in a given model m, the number of binary variables, denoted N, that are added to the problem can be expressed as

$$N = \sum_{\omega \in \Omega} \sum_{m \in \mathcal{M}_\omega} \sum_{h \in \mathcal{H}_m} y_{hm\omega} \tag{10.8}$$

This can be derived from the formulation in Section 7.1.3, where one binary variable is needed to represent one hinge function. The variable $y_{hm\omega}$ represents one hinge function h in a model m in scenario ω. Adding hinge functions to an optimization model will not only lead to the addition of binary variables, but also continuous variables that represent the functions. However, binary variables have the highest impact on solution time, and are therefore highlighted here.

The complexity of the models should be determined based on the trade-off between bias and variance, presented in Section 4.1.1. The bias-variance trade-off states that the prediction errors in a prediction model can be decomposed into the error due to bias and the error due to variance, and there is a trade-off between minimizing these errors. In general, if model complexity is increased, the bias is often reduced while the variance is increased. Oppositely, the bias is often increased and the variance is reduced when the model complexity is reduced. This can be seen in the illustration in Figure 10.8, where the models to the left have a high complexity and variance, whereas the models to the right are simple but biased. The parameters of the machine learning models from the hyperparameter tuning are determined by minimizing these errors, and thus finding an optimal complexity level from a machine learning perspective.

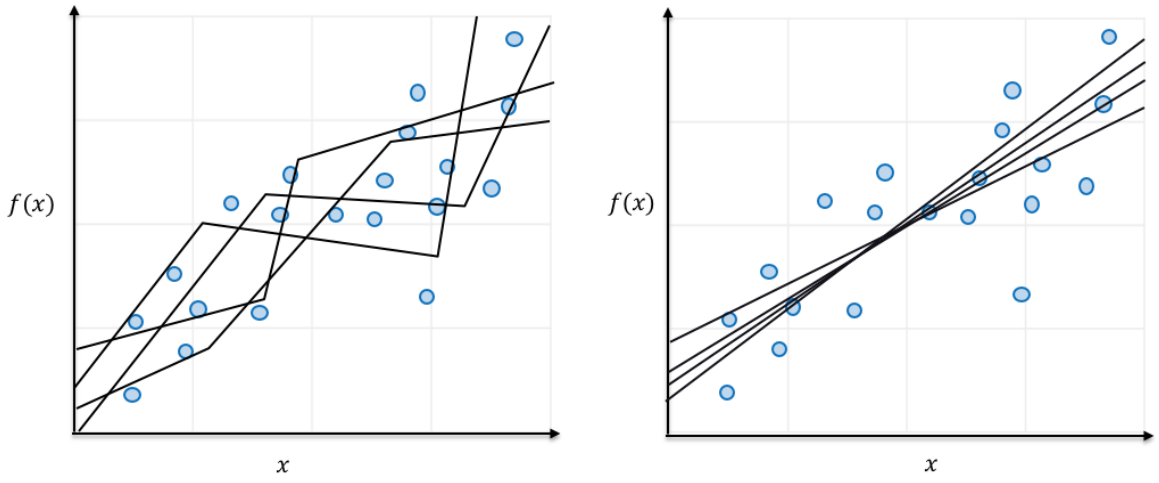

Figure 10.8: High and low model complexity representing the bias-variance trade-off for MARS models

The bias-variance trade-off implies that many scenarios are needed to cover the variance of complex models. Oppositely, fewer scenarios are needed when the model variance is low. It should however be noted that the error stemming from bias might be large in low complexity models. In real-time optimization, the solution time is decisive, and it should be restricted to a reasonable maximum limit.

Complex models, which also require many scenarios to cover their uncertainty, might therefore lead to a problem that is too large to solve within the time limit. This is based on the fact that when increasing the number of scenarios in the optimization model, the number of binary variables will increase linearly, since the MARS models are used directly as scenarios in the optimization.

Based on the discussion above, there are two ways the complexity of MARS models will impact the solution time, illustrated in Figure 10.9. Firstly, the complexity directly impacts the solution time because of the increase in binary variables. Secondly, an increase in model complexity leads to an increase in variance. As an increase in variance is an argument to include more scenarios in order to capture the uncertainty sufficiently, the solution time is further increased. It should be noted that when complex models require more scenarios in the optimization, simple models have a bias error that cannot be reduced by including more scenarios.

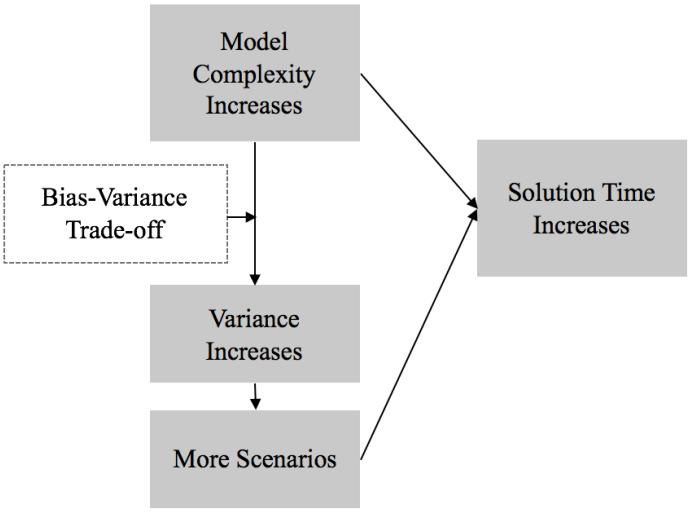

Figure 10.9: Relationship between machine learning model complexity and solution time in optimization

10.3.2 Three Levels of Complexity

To be able to test the effect of complexity on solution time, three levels of complexity for MARS models are introduced, where the highest complexity level corresponds to the optimal models found in Section 10.1.2. The three different complexities of the models are applied in Chapter 11, and relevant parameters are presented in Table 10.3. For the low and medium complexity models, the min span parameters from the optimal models are constant, while the penalty and max terms parameters are varied. The penalty values, represented by a p, and the max terms, represented by an m, are the optimal parameters found in tables 10.1 and 10.2. The combinations of parameters for the low and medium complexity levels are chosen based on yielding the best possible result from the evaluation metrics while at the same time giving the desired complexities. The hyperparameters used for low and medium complexity, in addition to the MSE and R^2 values from the grid search, can be found in tables E.1 - E.3 in Appendix E.

Table 10.3: Three levels of complexity for well models and pipeline pressure models

	Complexity level	MSE	R^2	Penalty	Max terms	Hinge functions
Oil flow rate	Low	2.60	0.62	4*p	m/4	2.30
	Medium	4.19	0.35	2*p	m/2	4.13
	High	4.85	0.26	p	m	6.21
Wellhead pressure	Low	17.71	0.44	8*p	m/8	3.16
	Medium	21.66	0.29	4*p	m/4	5.82
	High	19.74	0.43	p	m	15.94
Pipeline pressure	Low	2.70	-1.82	4*p	m/4	3.54
	Medium	5.51	-4.30	2*p	m/2	6.69
	High	6.28	-4.88	p	m	10.29

The test scores given for the models are averages over all wells for the well models and all manifolds for the pipeline pressure models. For all complexities, the number of hinge functions and test scores are based on an average of 50 runs of each model, training on a subset of the training data, and testing on the test data. Test scores for all the models with the three complexities can be found in tables E.4 and E.5 in Appendix E. It should be noted that the test scores and number of hinge functions may vary between wells and between manifolds, and are stated here as averages for simplification. In addition, these scores are not comparable with the scores found from the grid search, as the grid search scores are found from cross-validation within the grid search algorithm.

As seen from the table, the high complexity models yield worse test scores than the low and medium complexity, on average, which is the opposite of what is expected. Additionally, the R^2 scores for the pipeline pressure models are very poor for all levels. One important reason for this is that the train and test datasets are from two different distributions. That is, the test dataset is from a later time period. The process generating the data is not stationary because the reservoir develops as oil is extracted. The underlying distribution will therefore change over time, and the test dataset will be different than the train dataset. In addition, the GOR and WCT of the reservoir changes over time, and this is not represented by the models generated. This is a potential problem for the models, and it is therefore difficult to interpret the test scores given.

The low complexity models have the fewest hinge functions, as defined. They give simple representations of the data source, and are not able to model the relationships as accurately as desired. Since the complexity in these models are restricted compared to the optimal complexity, the models are underfitted, and might miss relevant relations between the input and output variables. The bias and variance decrease and increase, respectively, for the next two complexity levels. They also give the possibility of more overfitted models, as random noise in the training data may be modelled. However, it is assumed that the high complexity models represent a suitable trade-off between underfitting and overfitting, since they are based on the optimal hyperparameters from the grid search.

10.3.3 Well Model Uncertainty from Variance

In this section, the uncertainties stemming from variance in the three complexities of the well models are quantified based on the method described in Section 8.4. The standard deviations when moving away from a current operation point are used to set suitable limits for the choke openings, and the limits found are applied in Chapter 11. There are unique limits for each level of complexity, as the variance of the models are varying.

Example of uncertainty bands for Well 1 for each of the complexities are depicted in figures 10.10, 10.11 and 10.12. The darker blue area in the figures represents one standard deviation away from the mean, while the lighter blue represents two standard deviations. It should be noted that the oil flow model is dependent on two variables, and a plot of the flow as a function of choke is a simplification for illustration purposes. It is observed that the uncertainty band for the low complexity model is tighter than for the medium and high complexity model. That is because the low complexity model is not as flexible as the higher complexities, and will therefore not be as sensitive to small changes in the data source. This gives a model with lower variance, but as discussed, it is likely to be biased. As model complexity increases, the variance similarly increases.

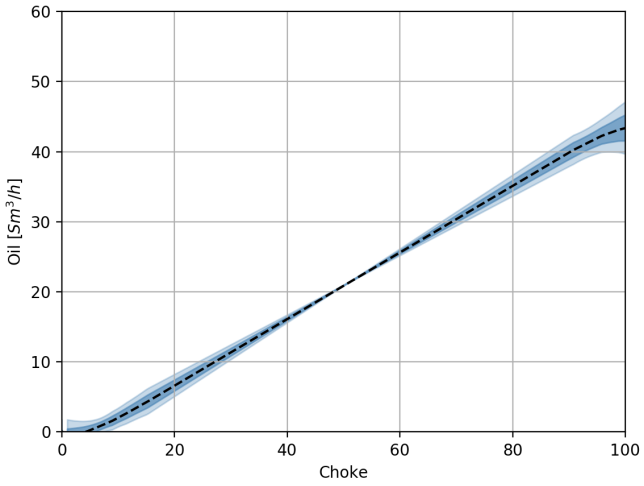

Figure 10.10: Well 1 model uncertainty for low complexity

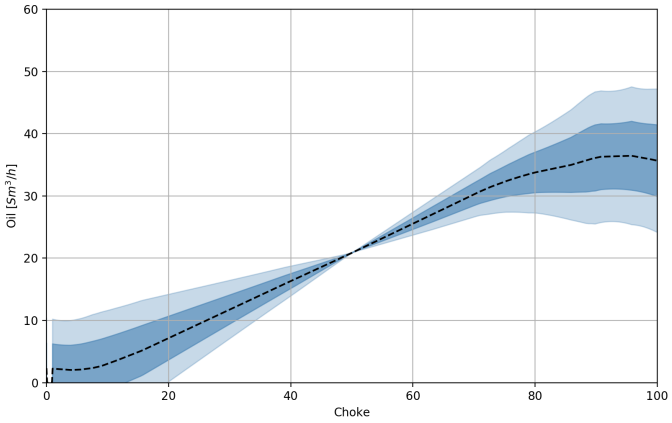

Figure 10.11: Well 1 model uncertainty for medium complexity

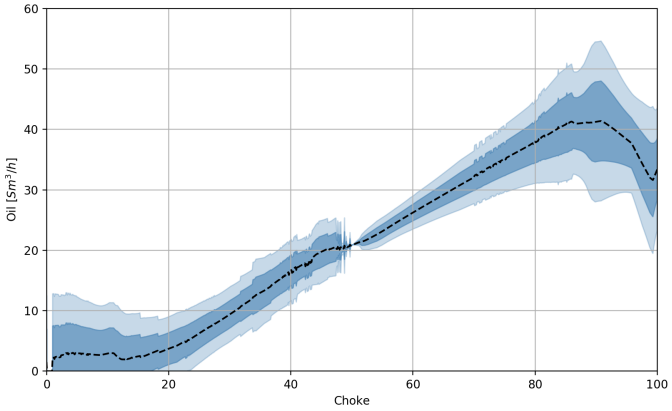

Figure 10.12: Well 1 model uncertainty for high complexity

Table 10.4 presents the uncertainty limits for all complexities and wells for Initial Case 1, with the coefficient of variation set to 0.2. The table shows that the limits vary depending on the initial choke opening for each well, where some of the choke openings lie in more uncertain areas. It is also found that when the model complexity increases, the choke limits generally decrease.

Table 10.4: Delta choke limits for the operation point in Initial Case 1

	W01		**W03**		**W04**		**W05**		**W06**		**W07**		**W08**	
	-Δu	+Δu	-Δu	+Δu	-Δu	+Δu	-Δu	+Δu	-Δu	+Δu	-Δu	+Δu	-Δu	+Δu
Low	32.7	50.0	34.7	40.3	40.2	53.1	24.6	42.2	8.3	49.5	14.0	43.4	23.3	49.2
Medium	21.7	50.0	25.5	34.8	40.6	41.2	40.7	21.8	20.2	49.5	0.9	43.4	25.8	23.5
High	31.3	20.5	11.5	11.4	36.9	26.2	32.6	20.6	42.5	34.5	0.8	20.6	23.5	22.5

As described in Section 9.4, the uncertainties related to the different models stem from several sources, one being the lack of data. When operating in areas with large uncertainty, the possibility of breaching operational constraints or getting a lower production increase. To account for some of this uncertainty, limiting the chokes will restrict the solution so that it cannot move to new operation points where the

uncertainty is high.

It should be noted that the uncertainty limits only takes the variance of the models into account, and not the bias. For the models of low and medium complexity, the bias error term might be considerable. The solutions from the optimization might therefore not be valid if the models are highly underfitted, which is further discussed in Chapter 11. It is outside the scope of this thesis to study how a bias can be included in the optimization, but it is discussed as a topic of future research in Chapter 13.

Chapter 11

Computational Study

The results from the computational study are presented in this chapter. The computational study is performed on the case described in Chapter 9. The results in this study are therefore specific for the case, but a similar study could be performed for other petroleum fields or synthetic cases. The two initial cases used in this chapter are as stated in Section 9.5. Note that all objective value results stated in this chapter are based on the oil rate of the original operation point and the change in oil rate that comes from the delta oil flow rate models. The chapter begins with the stochastic optimization study presented in Section 11.1. Further, the results from the bi-objective optimization problem where the conditional value-at-risk (CVaR) is minimized are presented in Section 11.2. Lastly, Section 11.3 presents the results from the bi-objective optimization problem where cost of slugging is minimized.

11.1 Stochastic Optimization Study

As discussed in Section 10.3, integration of machine learning techniques in stochastic optimization involves a trade-off between accurate machine learning models and the number of scenarios. While the processes are modelled more correctly with advanced models, the uncertainty can be better captured with more scenarios. In addition, it is desirable to get an optimal solution within a reasonable time frame. The focus of this study is thus how to allocate computing power available in order to make best possible decisions, within reasonable time.

The study is performed on the stochastic optimization model with the objective function as presented in equation (6.1), and the constraints (6.2) - (6.45). In this study, the scenarios are generated as described in Section 8.3, where a random selection of 90% of the total data source is used to generate each of the models that together make up a scenario. Initial Case 1 is applied for the analysis in this section, and the maximum number of choke adjustments is set to three.

The chapter begins with a technical study in Section 11.1.1. Section 11.1.2 investigates the distribution

of objective values for the different complexity levels for fixed first-stage solutions. Further, Section 11.1.3 presents the expected value of perfect information (EVPI) and the value of the stochastic solution (VSS) results. Lastly, the choice of suitable MARS complexity level for the case studied is discussed in Section 11.1.4.

11.1.1 Technical Study

A technical study of the stochastic optimization model formulated in Section 6.2 is presented in this section. The technical study consists of a presentation of the problem size, followed by stability testing.

Problem Size

Real-time optimization (RTO) implies that solution time is of essence. Solution times of a RTO problem should ideally not exceed 30 minutes. Relevant factors that affect the solution time include the number of scenarios, the complexity of the MARS models that are applied and the number of well adjustments allowed, in addition to the magnitude of the adjustments.

The number of scenarios in the problem versus solution time is presented for the three complexities of MARS models in Figure 11.1, with a linear scale and a logarithmic scale for the solution time to the left and right, respectively. The solution time is an average of 20 runs for each number of scenarios.

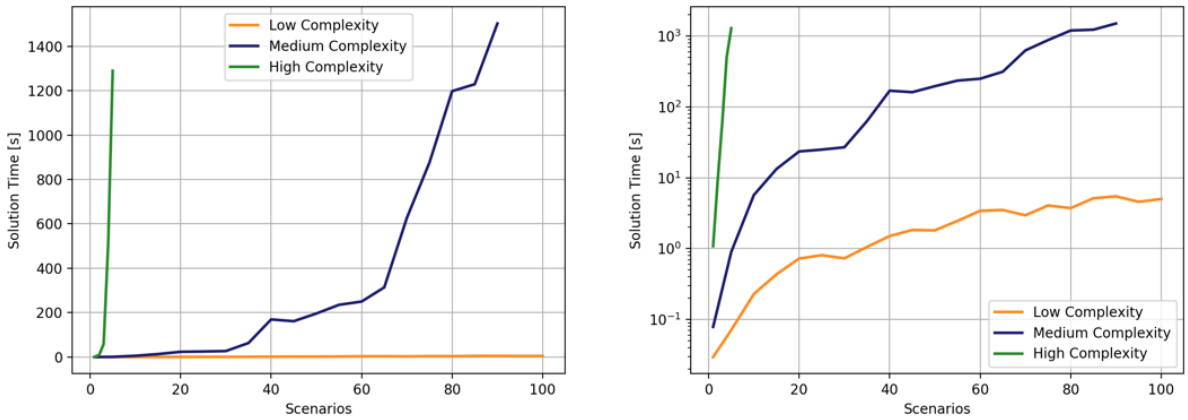

Figure 11.1: Solution time for varying number of scenarios for the three complexities on a linear and a logarithmic y-scale

The results show that the solution time is higher for MARS models with higher complexity, which is expected based on the discussion in Section 10.3.2. The solution time with low complexity models is around five seconds for 100 scenarios. In comparison, the solution time with only four scenarios for the high complexity models is over 1200 seconds. The medium complexity lies between the low and high complexities, with a solution time around 1200 seconds for 80 scenarios. If the high complexity level is used in the optimization problem, the problem is too computationally expensive to solve. In addition, this complexity is highly restrictive in terms of number of scenarios the optimization problem can handle.

Table 11.1 summarizes the maximum number of scenarios possible when the problem is solved to optimality and the limit of the solution time is set to 30 minutes. The MILP solver terminates and returns an optimal solution when the optimality gap is less than 0.0001%. The table presents the maximum number of scenarios the MILP solver has been able to solve in the first column, and the maximum number of scenarios that are applicable to use in further analysis in the second column. The latter represents the maximum number of scenarios where 80% of runs have given an optimal solution within the time limit.

Table 11.1: Maximum scenarios for varying level of complexity of MARS models

Complexity	Max Scenarios	Max Applicable Scenarios
Low	550	400
Medium	100	70
High	5	4

Based on the results from the problem size, it is concluded that the high complexity MARS models are too computationally heavy, and will not give results within reasonable time. As seen in Section 10.3, there are two ways that the MARS complexity impacts the solution time. In addition to directly affecting the solution time through binary variables, the variance of the model should be accounted for by scenarios, which involves more binary variables. However, as seen above, the high complexity model can only handle a maximum of five scenarios. It is concluded that such a small amount of scenarios is not enough to account for the variance of the models. The MARS models with high complexity are thus not studied further in the computational study, but are used as a basis of comparison for low and medium complexity models.

Stability Testing

It is referred to Section 4.2.2 for a theoretical background on stability measures. In-sample stability and out-of-sample stability are presented for the low and medium complexity levels in the following section. The low complexity models are run up to 150 scenarios, and the medium complexity models up to 70 scenarios. The approximation to out-of-sample stability, the test presented in equation (4.30), is applied.

Testing in-sample stability involves finding if approximately the same objective value is achieved for each generated scenario tree. In-sample stability is important to ensure that the optimization problem is able to find the solutions that perform well. The in-sample stability of the optimization model using low and medium complexity scenarios is measured. The optimization model is solved 50 times for each number of scenarios. New MARS models are generated each time the model is solved, according to the scenario generation procedure described in 8.3. The objective values from the different runs are recorded, and the averages are plotted in Figure 11.2 for low and medium complexity scenarios to the left and right, respectively. To quantify the in-sample stability, the standard deviations of the objective values are calculated for each number of scenarios. This is plotted in Figure 11.3 for low and medium complexity scenarios to the left and right, respectively.

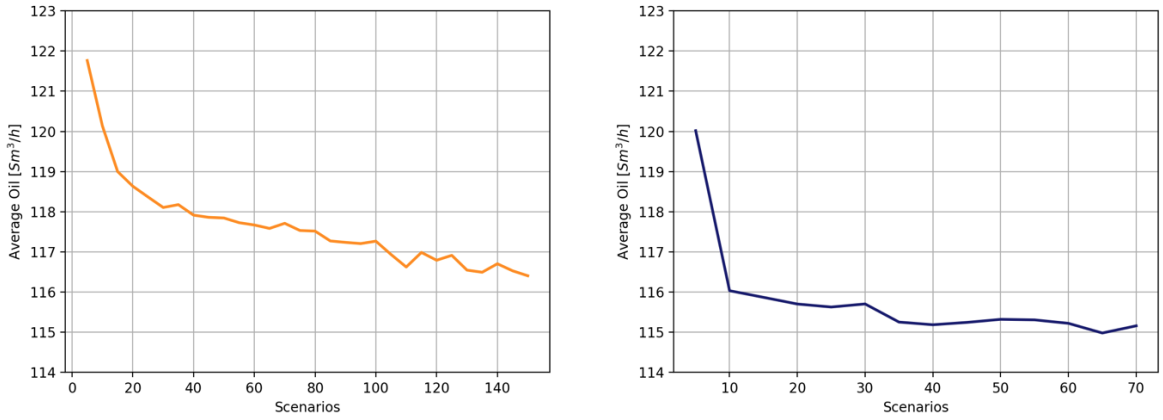

Figure 11.2: Average objective function value for in-sample stability for low and medium complexity

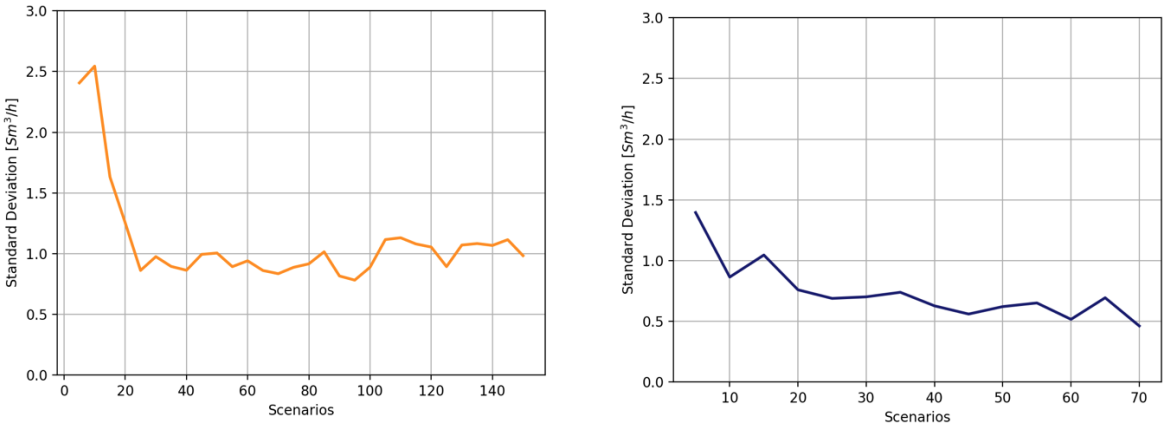

Figure 11.3: Standard deviation for in-sample stability for low and medium complexity

Figure 11.3 shows that the standard deviation is high, even for a high number of scenarios. In addition, the standard deviation of the low complexity models is higher than medium complexity, which is opposite of what may be expected. However, variance within a model has to be studied by fixating choke openings, which is not the case for in-sample stability where the choke openings vary. It is therefore not possible to make a conclusion on the actual level of variance in the models based on the in-sample stability testing. Model variance with fixed solutions is studied in Section 11.1.2. Assessing the stability of the complexities, it can be found that the complexities are relatively stable. The low complexity models stabilize from around 30 scenarios, and the medium complexity models around 50.

Measuring out-of-sample stability involves checking if the solutions from different scenario trees have approximately the same objective function value as if the true distributions were used on the same solutions. The out-of-sample stability can be used to measure the quality of the scenario generation method. Out-of-sample stability is important to ensure that the solution will perform well when applied to a real situation. For this thesis, a solution is represented by a set of adjustments for choke values. The recommended set of adjustments for a scenario tree is applied for other scenario trees. That is, the first-stage solutions are fixed to a given solution. For each number of scenarios, solutions from 20 scenario trees are applied. The standard deviations of the objective values are plotted in Figure 11.5.

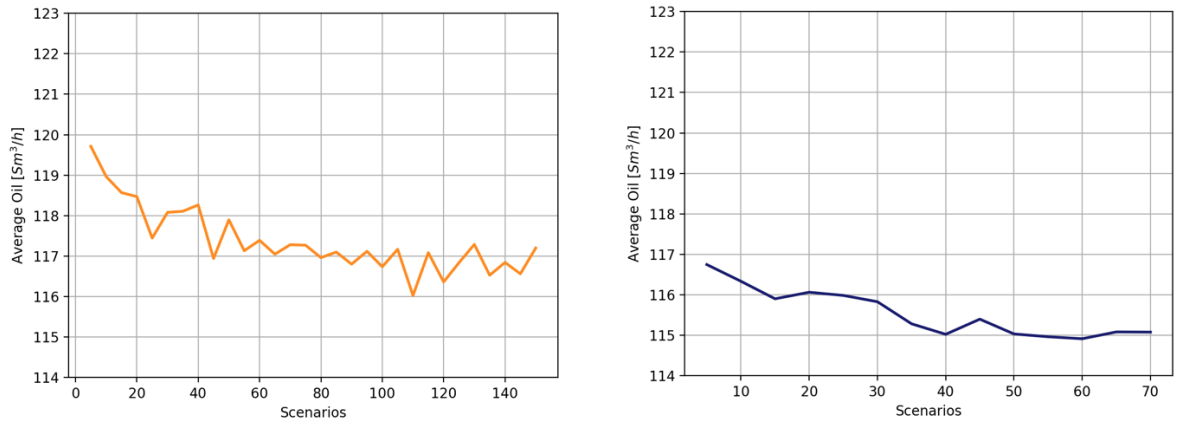

Figure 11.4: Average objective function value for out-of-sample stability for low and medium complexity

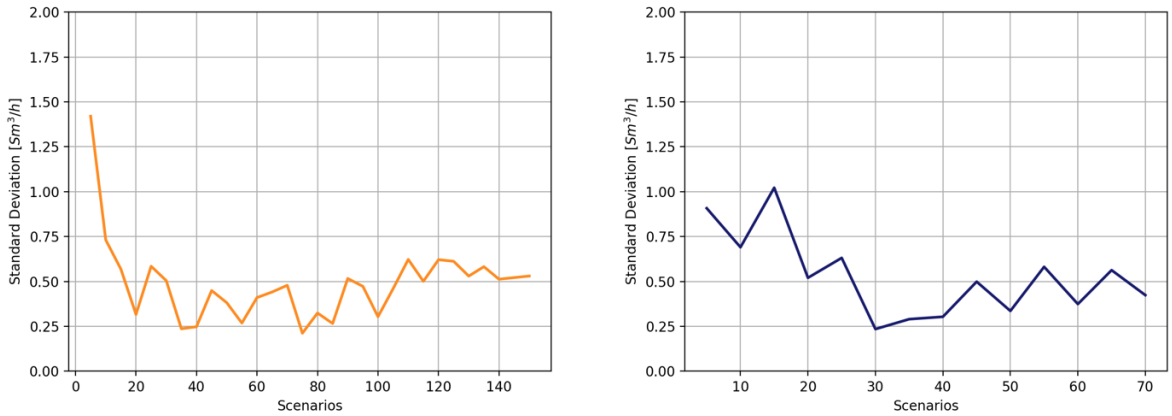

Figure 11.5: Standard deviation for out-of-sample stability for low and medium complexity

Similarly to the in-sample stability results, Figure 11.4 shows that the average objective function decreases for more scenarios. It is observed from Figure 11.5 that the standard deviations for the out-of-sample stability are varying for both complexity levels, and not as stable compared to in-sample. It would be preferred to study the stability with a higher number of scenarios, especially for the medium complexity, but this is beyond the models computational limitations. It therefore cannot be concluded whether or not the model would have out-of-sample stability for more scenarios. However, some stability seems to be achieved for about 30 scenarios for the low complexity and around 45-50 scenarios for the medium complexity, even though the standard deviation still varies.

From figures 11.2 and 11.4, it is observed that the average objective function decreases for an increasing number of scenarios. The reason for this might be that the solution must satisfy restrictions for more scenarios. For example, more scenarios must adhere to the gas capacity constraint, which leads to a lower expected oil production. The average objective function seems to be lower for medium complexity, both for the in-sample and out-of-sample results. This might be due to the higher flexibility the medium complexity models have to the data source combined with the strict gas capacity limit. The flexibility implies that the scenarios vary more, and cases with higher gas production can occur. Another reason might be the uncertainty limits that are only based on the variance of the models. Since the medium

complexity models have a larger variance than the low complexity models, they are more restricted by the limits. This is a disadvantage, as the low complexity models in reality might be more uncertain due to their bias.

With scenarios built from an uncertain data source, the scenarios will vary considerably. Additionally, the scenario generation method does not generate a fixed set of stochastic variables, but rather new variables for each scenario. This could possibly yield an infinite amount of stochastic variables, thus needing an infinite amount to cover the variance. This is another reason for why it is difficult to set an exact limit on the number of scenarios.

Even though it is difficult to determine the exact limit of scenarios, it is assumed that 30 scenarios are sufficient for the low complexity level, and 50 for the medium complexity level, based on the stability tests. Since the low complexity models only takes a few seconds to solve in the optimization problem, 50 scenarios are applied for both complexity levels.

11.1.2 Evaluation of Predictive Performance

This section studies how well the low and medium complexity models perform in estimating the objective value.

Representative choke recommendations for selected number of scenarios are recorded in Table 11.2 for low and medium complexity models. Δu_i for well i represents the change in choke opening recommended from the optimization. The recommendations presented are not averages, but the most common solution combination seen from the in-sample stability testing.

Table 11.2: Representative choke recommendations for selected number of scenarios for low and medium complexity

	Scenarios						
	10	**30**	**50**	**70**	**100**	**200**	**400**
Low	$\Delta u_6 = 20.0$ $\Delta u_7 = 19.4$ $\Delta u_8 = -13.3$	$\Delta u_6 = 17.2$ $\Delta u_7 = 20.0$ $\Delta u_8 = -13.8$	$\Delta u_6 = 16.3$ $\Delta u_7 = 20.0$ $\Delta u_8 = -13.0$	$\Delta u_6 = 17.1$ $\Delta u_7 = 20.0$ $\Delta u_8 = -15.9$	$\Delta u_1 = 4.1$ $\Delta u_4 = -6.4$ $\Delta u_6 = 20.0$	$\Delta u_1 = 4.0$ $\Delta u_6 = 20.0$ $\Delta u_8 = -14.0$	$\Delta u_1 = 3.7$ $\Delta u_6 = 17.3$ $\Delta u_8 = -3.2$
Medium	$\Delta u_4 = 16.8$ $\Delta u_5 = -20.0$ $\Delta u_6 = 20.0$	$\Delta u_4 = 11.5$ $\Delta u_5 = -16.5$ $\Delta u_6 = 20.0$	$\Delta u_4 = 15.2$ $\Delta u_5 = -20.0$ $\Delta u_6 = 18.8$	$\Delta u_4 = 16.1$ $\Delta u_5 = -20.0$ $\Delta u_6 = 17.7$	-	-	-

As mentioned in the stability testing, the testing does not take into account that the scenario generation methods might introduce a bias in the solution. As discussed in Section 10.3, the model complexity should be studied with regards to the bias-variance trade-off. In order to get an indication of the predictive performance of the low and medium complexity levels, the distributions of the objective function values are studied. The solutions from applying the high complexity level models for fixed first-stage solutions are used for evaluation purposes, since it is expected to be the most correct and unbiased distribution. In the following, the solutions recorded in Table 11.2 are evaluated for the three complexity levels. Tables with the corresponding objective values are found in Appendix G.

By fixating the choke openings when solving the optimization problem for the different complexities, the objective values show how the models give varying results based on the realization of the uncertain parameters. This is difficult to study when using several different choke openings. Using the low complexity models as an example, the models should give similar parameters from one scenario to another based on their relatively low variance. However, small differences in certain combinations of scenarios may in total enable a higher objective value than another set of scenarios. Several of such scenario combinations may therefore yield a higher variance than one model would indicate.

Distributions for Solutions from Low Complexity

Figure 11.6 presents estimates of the objective values for the choke recommendations obtained from using the low complexity level when solving the optimization model. Estimates are found by solving 1000 runs of the second-stage problem for the given complexity level with fixed choke openings.

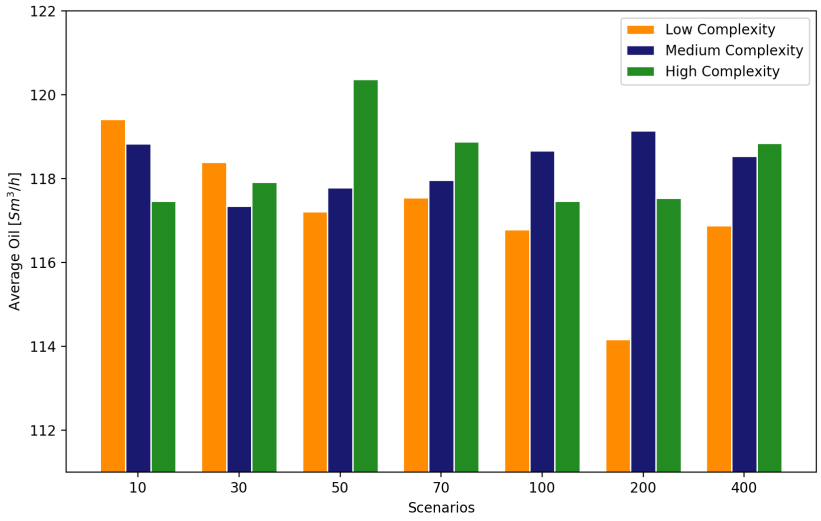

Figure 11.6: Average objective values for fixed first-stage solutions from low complexity scenarios

In general, it is observed from Figure 11.6 that the estimates from the medium complexity level are closer to the high complexity level estimates for the number of scenarios studied. However, for the solution from 30 scenarios, the low complexity level is closer.

The solution from 50 scenarios and the low level of complexity is used as an example to illustrate the distributions of objective values for the three complexities. Figure 11.7 depicts the distributions in a histogram. It is observed that the solutions from the low complexity level have a much smaller variance. The estimated value from the low complexity models differ by 3.16 from the estimate from the high complexity models. This is a substantial difference, and it indicates that the models have a bias error. The same applies for the medium complexity models, even though they are closer to the estimate from the high complexity models. It is also found that these models have higher variance, as the objective values have a much higher spread.

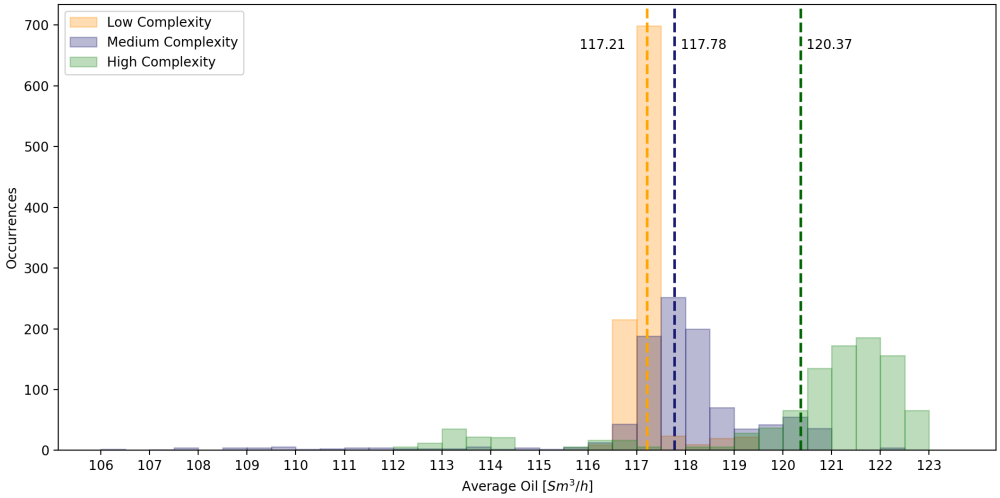

Figure 11.7: Distributions for the solution from low complexity and 50 scenarios

Distributions for Solutions from Medium Complexity

Similarly to the previous section, first-stage solutions found from solving the optimization problem with medium complexity models are studied. The estimates of the objective values are presented in Figure 11.8.

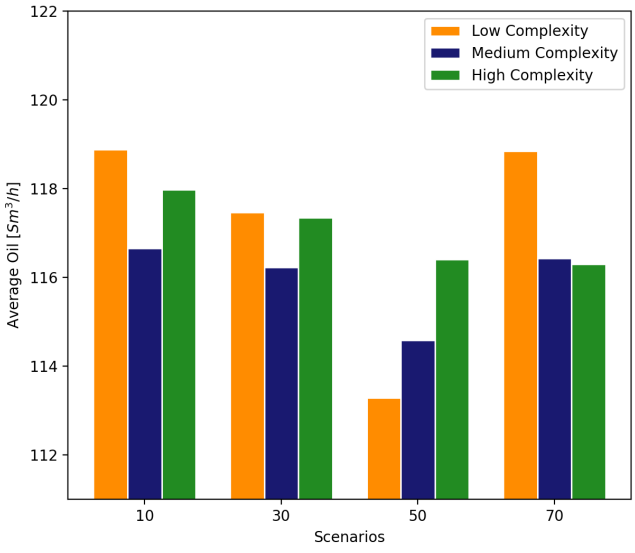

Figure 11.8: Average objective values for fixed first-stage solutions from medium complexity scenarios

The figure shows varying results for low and medium complexity models. The low complexity scenarios give average objectives that are closer to the estimate from the high complexity for 10 and 30 scenarios. Since each scenario consists of multiple MARS models, two per well and one per manifold, there are many factors that influence the predictions. There may therefore be certain cases where low complexity

estimates have a lower gap to the high complexity estimates than medium.

The solution from 50 scenarios and a medium level of complexity, presented in Table 11.2 and Figure 11.8, is further studied in Figure 11.9. It is observed from the histogram that the medium complexity models have a greater spread of objective values than the low complexity, but at the same time, the mean oil rate of 114.58 is closer to the high complexity mean of 116.41. This illustrates the higher variance, but lower bias, of the medium complexity models for this fixed first-stage solution, similarly to what was found in the histogram of the low complexity solution.

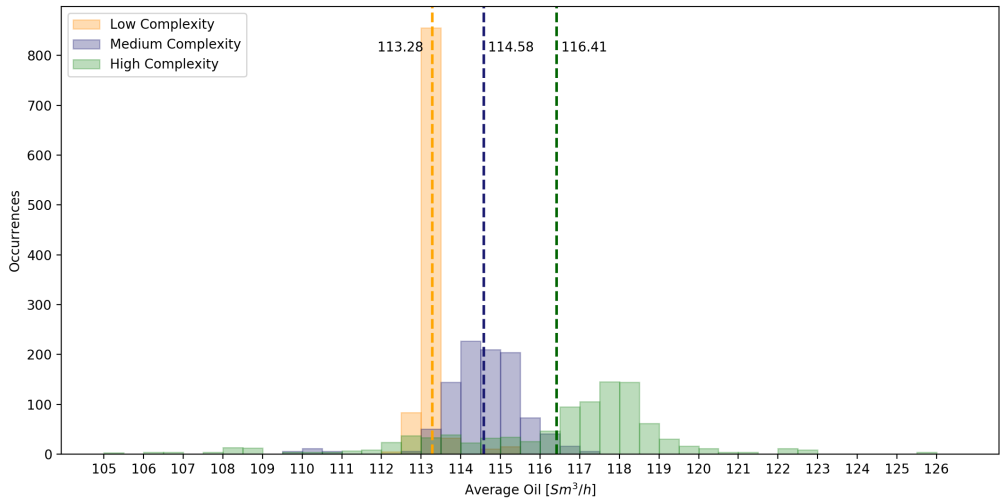

Figure 11.9: Distributions for the solution from medium complexity and 50 scenarios

The error from bias vary between each of the different well models and pipeline pressure models, and the objective value from the optimization is a result of this combination. The objective function can be viewed as an aggregation of the bias and variance effect from the individual well and pipeline pressure models. It is thus difficult to make a conclusion on the predictive performance only based on the gap to the high complexity objective function value. However, the gaps give an indication of how far off the models are from the high complexity models. The estimates show that objective values from both the low and medium complexity levels vary compared to the high complexity, while the medium complexity on average lie closer to the high complexity mean.

11.1.3 Expected Value of Perfect Information and the Value of the Stochastic Solution

The concepts of EVPI and VSS are explained in Section 4.2.2, and are calculated for the low and medium complexity levels in this section. EVPI is the difference between the objective function values from solving the stochastic problem, compared to solving the same problem given perfect information. VSS is a measure of the difference in objective value when planning with uncertainty incorporated, compared to using expected values. Note that if either complexity does not yield a feasible solution when finding either EVPI and VSS, the chokes are adjusted down until they reach feasibility.

The EVPI and VSS are illustrated for low and medium complexity in Figure 11.10 and Figure 11.11, respectively. RP is found by solving the two-stage recourse problem. WS is found by solving the optimization problem with perfect information, and calculating the expected objective value from evaluating the solution for all scenarios. The interpretation of the EEV is not straightforward in this application. The scenarios are MARS models, and it is not clear how to define the expected value of the uncertain parameters given by the models. In this case, a model that is trained on the complete data source is interpreted as an expected value model. The solution from the expected value model is evaluated for the relevant scenarios.

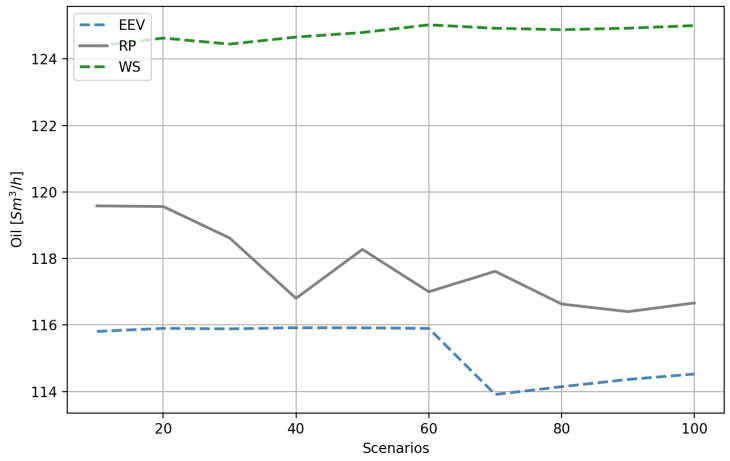

Figure 11.10: EVPI and VSS for a varying number of low complexity scenarios

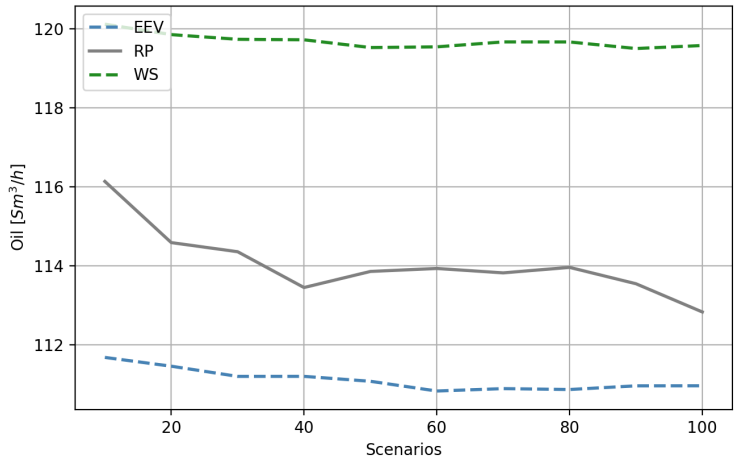

Figure 11.11: EVPI and VSS for a varying number of medium complexity scenarios

For this study, EVPI can be interpreted as the value of having models that perfectly represent the data, and thus each scenario represents a true realization of the well parameters. As seen in the figures for the two complexity levels, the EVPI increases as the number of scenarios increases. This might come from the conservative gas capacity constraint, that requires the gas production from all scenarios to lie below the gas handling capacity. With more scenarios, this gives a lower average oil production, and thus the

value of perfect information increases. Both low and medium complexity have high EVPI values for 100 scenarios, 8.34 and 6.75, respectively. This represents a 7.14% increase for the low complexity and a 6.01 % increase for the medium complexity from the recourse problem to having perfect models. The medium complexity represents a higher variance, and one would therefore expect a higher EVPI as there would be more to gain in knowing the actual value. However, the low complexity has a higher bias, which is not covered by the scenarios. This could also contribute to a higher EVPI, as the uncertain parameters would in general be further away from the actual realizations.

The EVPI shows that the correctness of the models are vital as to optimize the oil production. Such perfect models represented by WS cannot be obtained with the current data source, as discussed throughout this thesis, but it gives a good indication of the value of choosing the appropriate models.

The VSS for the two complexity levels is the difference from training the models on the whole dataset and using scenarios with models trained on subsets of the data. The EEV is interpreted as one set of models being representative for all the data, and giving the expected parameters in all cases. As seen in the Figure 11.10, The VSS for low complexity and 100 scenarios is 2.14, representing an increase of 1.87 % for the recourse problem from using the expected values. For medium complexity, as seen in Figure 11.11, the VSS is 2.51 for 100 scenarios, which represents an increase of 2.21 %. However, as mentioned above, these expected values are very difficult to quantify, and it may therefore be difficult to state the added value of the results. Yet, the VSS may indicate that there is in fact a value of varying the models, as they cover more of the potential outcomes of the uncertain parameters.

11.1.4 Choice of Complexity Level

Section 11.1.1 concludes that the high complexity of MARS models are not suitable for the case studied in this thesis, as the optimization model becomes too computationally expensive. The choice of complexity level is therefore based on comparing the low and medium complexity levels.

From machine learning, it is known that less advanced lower complexity MARS models are biased, and therefore unable to capture important relationships in the well and pipeline pressure models. This effect is illustrated in the histograms in Section 11.1.2. Using fewer hinge functions to model the complex relationships gives underfitted models, which is discussed in Section 10.3.

Since the low complexity models have a small variance, increasing the number of scenarios will not give the same improvement in predictions as increasing the number of scenarios for medium complexity models, which have a larger variance. The processing power available should be utilized to its full extent, and it could be argued that the trade-off between using complex machine learning models and including enough scenarios is not optimal when the low complexity level is applied.

Another reason for why medium complexity might be more applicable than low, is how the scenario-generation method is based directly on the models. The chosen scenario-generation method should preferably not introduce any bias into the solution. However, when processing power available is restricted, and complexity of machine learning models thus must be restricted, it is difficult to eliminate

the bias completely. The bias introduced should however be as small as possible, indicating an advantage of medium complexity.

It is emphasized that this thesis provides a method of covering the variance of the MARS models in the optimization, but not a way of including the bias in the optimization. Further research within this area would be useful for choosing suitable complexity levels.

Based on the analysis in this section, it is concluded that the medium level of complexity is most suitable for optimization purposes. It is emphasized that the results might be different for a different case. If for example only one MARS model is used in an optimization model, it is very likely that choosing the optimal model from hyperparameter tuning would be preferable. The requirements for the solution time are thereby decisive for the conclusion in this thesis.

11.2 Results from Bi-Objective Optimization with CVaR

The bi-objective optimization problem presented in Section 6.3 is examined in this section. The problem consists of the objective function in equation (6.53), constraints (6.2) - (6.45) and (6.54) - (6.57), with some modifications presented in Section 6.3. The model is solved with Initial Case 1 presented in Section 9.5, but without a limit on gas production. The maximum number of choke adjustments is set to three, and models of medium complexity level and 20 scenarios are applied.

As described in Hanssen and Foss (2015), using expected values of uncertain parameters for an unconstrained problem is a viable approach, but there is no guarantee of optimality. Using expected values when a problem is constrained, in this case by a gas capacity, has certain drawbacks. If the solution of the optimization problem is at the constraint, and this solution is implemented by the production engineers, there may be a chance that the capacity is violated. Oppositely, there may be a chance that there is actually spare capacity left. Using CVaR will assess some of these drawbacks, as there is uncertainty connected with both the gas rates and the capacity limits. Other uncertain parameters could also be studied, such as the GOR and WCT of each well, but this is outside the scope of this analysis.

The problem is solved with the ϵ-constrained method described in Section 4.2.4. The ϵ-value can be interpreted as the degree of risk aversion of the decision maker, and is in the range [206000, 516000] with a step length of 5000. When the ϵ-value is low, the CVaR has to be below a strict capacity limit, representing that the decision maker would want to limit the possibility of high gas production. Oppositely, a high ϵ-value will enable a high CVaR value. This represents a risk-neutral decision maker. Varying the α-value is another way of controlling the conservativeness, where a higher confidence level is connected with a more risk-averse decision maker, and a lower with risk-neutral. The α-value should therefore be set in regards to the preferred level of risk. α-values of 0.80 and 0.90 are used in this study, with α = 0.99, the closest α-value possible to a more robust formulation, used for comparison. An α-value of 0.80 means that the CVaR is the expected gas production of the 20th percentile highest gas rates.

The Pareto front can be viewed as a set of options for production engineers with varying degrees of risk aversion. The trade-off between oil production maximization and risk minimization can be seen in Figure 11.12, representing the α-values 0.80, 0.90 and 0.99.

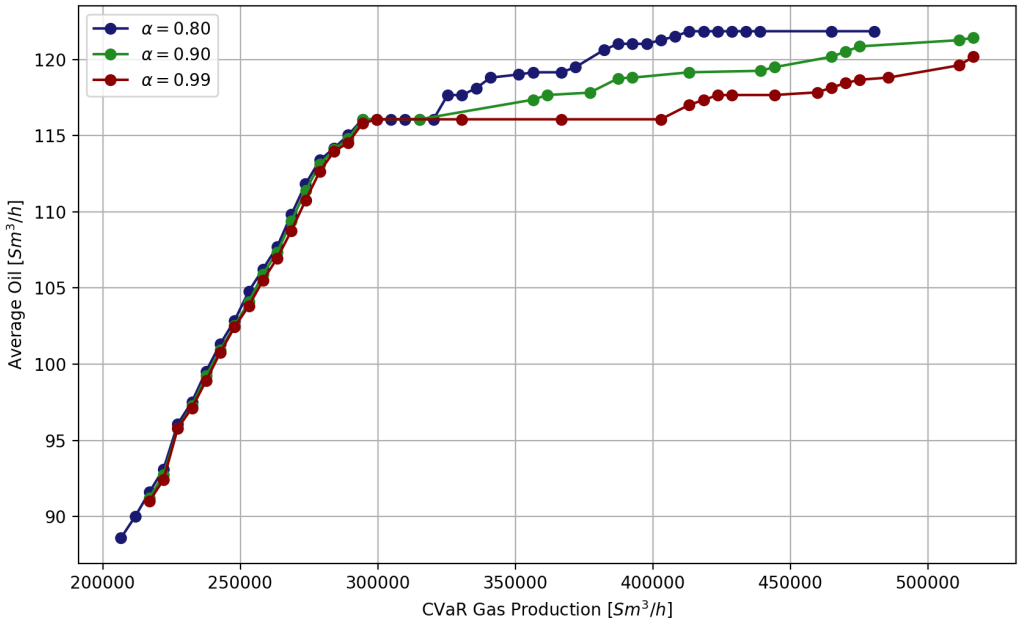

Figure 11.12: Pareto front for bi-objective optimization with CVaR

The figure shows that as the ϵ-values increase, the CVaR is less restricted, and both the value of the CVaR and the oil production increase. Compared to the robust formulation of the gas constraint in the single-objective stochastic optimization problem, the objective value increases as some of the scenarios are allowed to exceed the capacity. For the Pareto fronts, the oil productions are at first restricted by the low CVaR-values. As this value increases, the oil rate follows.

At CVaR equal to 300000, the increase in oil compared to CVaR is reduced. This happens when the changes in choke openings stabilize. Before this point, choke openings belonging to different wells are adjusted for each increase in ϵ-value. After this point, only Well 4, Well 5 and Well 6 adjust the openings for nearly all ϵ-values, with Well 4 and Well 6 adjusting upwards and Well 5 downwards. A possible explanation may be that both Well 4 and Well 6 have lower GOR levels, while the GOR of Well 5 is high.

At each ϵ-value, the CVaR values for the three confidence levels are the same. However, these values are interpreted differently for the different levels. For $\alpha = 0.80$, the expected value of the 20th percentile highest gas production scenarios is minimized. Using the CVaR value equal to 356000 depicted in the three fronts as an example, this results in an average z-value of 13124.84, average gas production of 301535.72 and average oil rate of 119.27. The z-value for each scenario equals the excess value over VaR for gas. For this case, four scenarios produce over VaR, and thus have $z > 0$. For $\alpha = 0.90$, there are only two scenarios that have gas production exceeding VaR. For this case, the average z-value is 6237.39,

while the average gas production over all scenarios is 295876.93. Seeing that the 10th percentile highest gas production scenarios are minimized, fewer scenarios can have a z-value above zero compared to α = 0.80. The oil rate for α = 0.90 is 117.36. Compared to α = 0.99, which has an oil production of 116.08 and gas production below 300000 for all scenarios, allowing a risk of exceeding the gas capacity also enables on average a higher oil production.

The Pareto fronts show that a lower confidence level α allows more scenarios to have gas productions above VaR, while also yielding a higher oil production. A small average increase in gas production can thus give a higher expected oil output. It is also seen that after a certain point, the lower α confidence level reaches higher oil rates quicker, as the higher confidence levels are more restricted in gas production. The confidence level reflects the risk level of the decision maker, as lower levels yield higher risks of the actual gas production being high. It should be noted that a drawback of the method is that when minimizing a given percentile of the highest gas rates, one cannot limit the value these gas rates take. Thus, it is possible that the there is a big difference from the highest gas rates to the others. If this is something that should be avoided, constraining the average gas production below a limit could be a possible alternative.

11.3 Results from Bi-Objective Optimization with Slug Severity

The results from the bi-objective stochastic optimization problem presented in Section 6.4 are examined in this section. The computational study is performed on the objective function in equation (6.58), and the constraints (6.2) - (6.45) and (6.59) - (6.62). The problem is solved for Initial Case 2, presented in Section 9.5, and the medium level of complexity for MARS models.

In Section 11.3.1, the relevant paths from the classification tree (CT) based on the optimal parameters are repeated for convenience. The solutions and objective function values from the bi-objective model are investigated in a Pareto front in Section 11.3.2.

11.3.1 Classification Tree

The CT used in the bi-objective optimization is the optimal tree presented in Section 10.2, and the four paths with an associated cost of slugging are repeated in Figure 11.13. There are two paths with low risk of slugging, given a cost of 4.5, and two paths with a high risk of slugging, given a cost of 5.6, as described in Section 9.2.2. The rest of the paths are classified as safe operation areas, and are therefore not associated with any cost. Each scenario in the optimization model will have a cost corresponding to its path, which might vary based on the parameters in the models that make up the scenario.

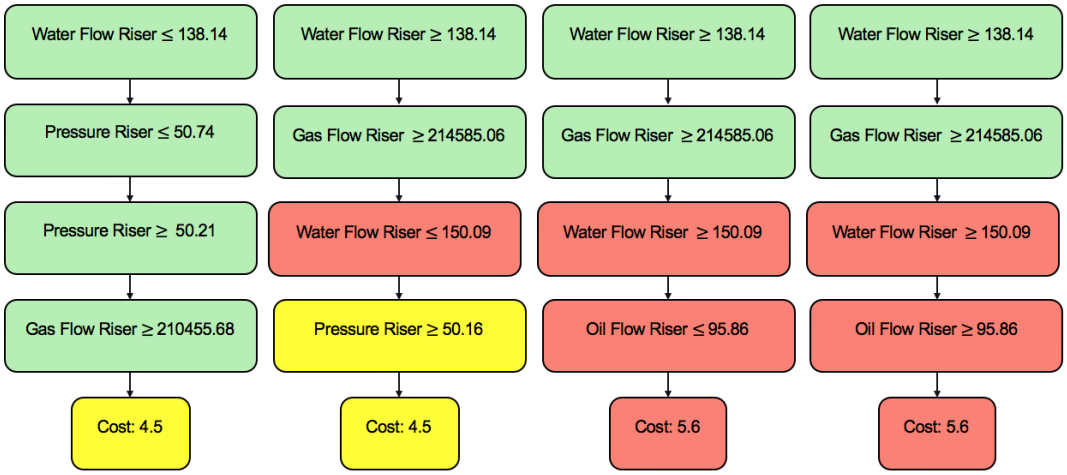

Figure 11.13: Paths in the optimal classification tree with an associated cost of slugging

11.3.2 Pareto Front

The Pareto front for medium complexity MARS models and 20 scenarios is illustrated in Figure 11.14. It is based on Initial Case 2, and is generated with maximum number of choke adjustments of three.

The problem is solved with the ϵ-constrained method. The ϵ-value is in the range $[0.2, 5.6]$, with a step length of 0.2. The front can be interpreted as a set of options for the decision maker, in this case the production engineer, on whether to minimize the risk of slug or to maximize oil production. The risk aversion of the production engineer determines which solution is preferable. By studying the risk of slugging in a bi-objective problem and in a Pareto front, the possible trade-off of a small increase in slug cost yielding a larger increase in oil production is taken into account. As expected, operating in an area associated with a high risk of slugging increases the oil production rate.

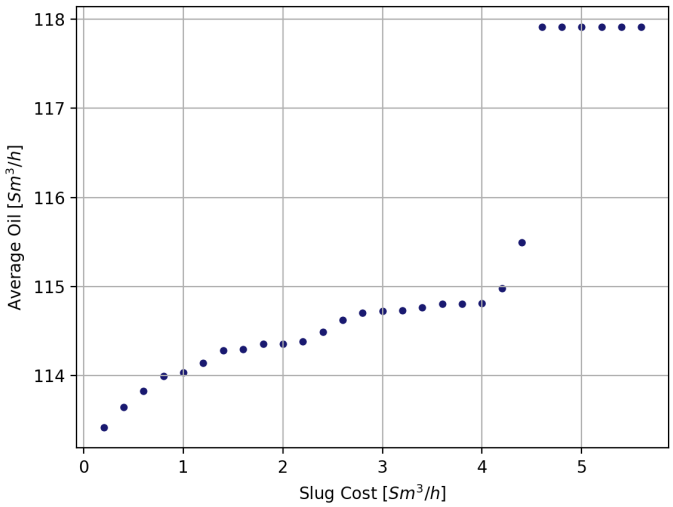

Figure 11.14: Pareto front for bi-objective optimization with slug severity

Associated with the Pareto front, it can be found how the solutions, the choke recommendations, change as the cost of slugging is restricted by ϵ. Figure 11.15 shows the choke openings when the only objective considered is maximizing oil production. This solution corresponds to an average slugging cost of 4.5. That is, the operation area gives that all scenarios end up in a path with a low slug cost. The corresponding objective value of these choke settings is 117.91, which can be found as the points in the top right of the Pareto front.

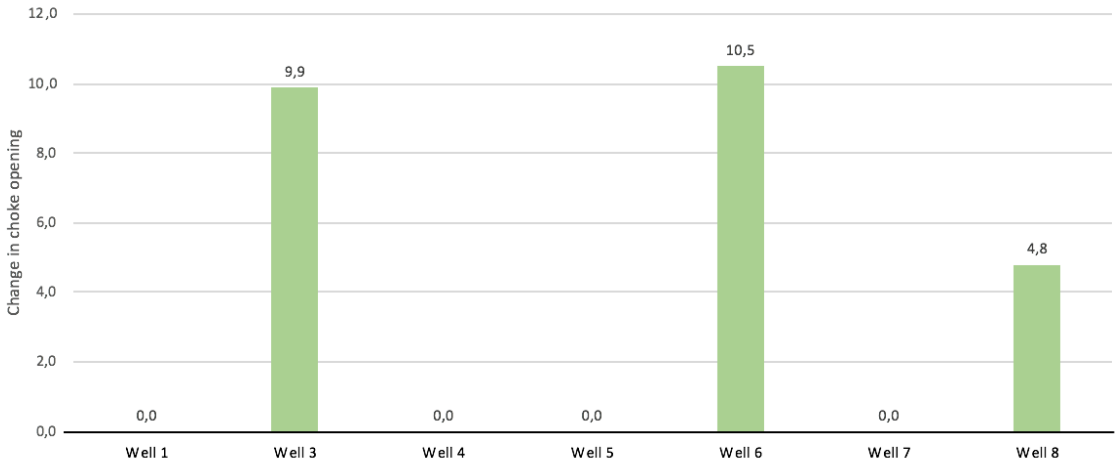

Figure 11.15: Choke recommendations with no cost of slugging considered

The solution does not change until ϵ is set to a value less than 4.5. At this level, there is a large change in the average objective value, and thus also changes in the current choke openings. The new solution is illustrated in Figure 11.16. The recommendations for ϵ-values less than 4.5 consist of the same three wells, but the absolute value of the change vary with the cost of slugging allowed.

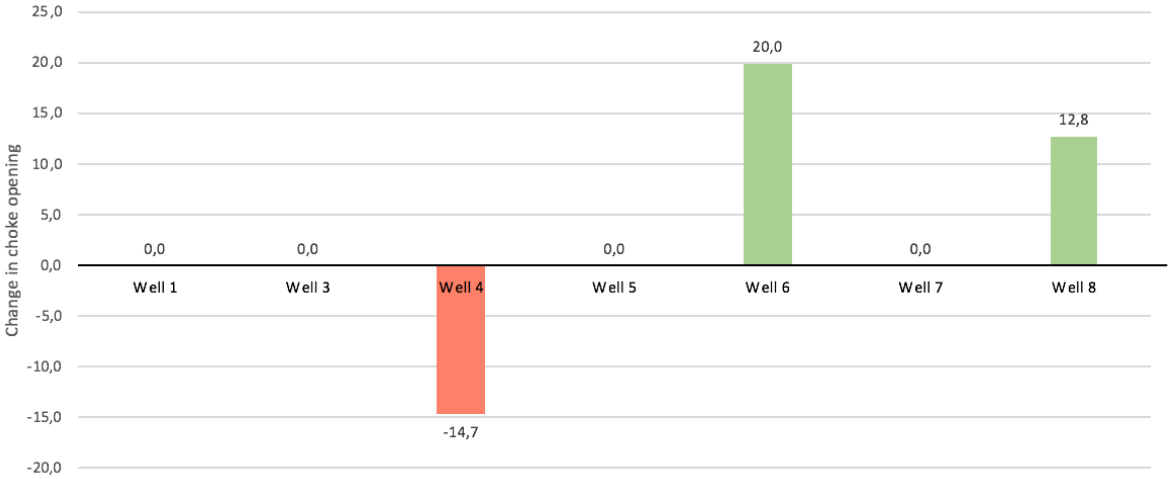

Figure 11.16: Choke recommendation with an average cost of slugging of 4.4

When the cost of slugging is further restricted by ϵ, the choke opening of Well 4 is gradually decreased to the limit of -20, and Well 6 and Well 8 stay at approximately the same values. It is also noted that Well 3 is not in the solution, in contrast to the solution in Figure 11.15 where the slugging objective is not

considered. The reason for this might be that Well 3 and Well 4 are the wells with the highest water cut (WCT) values. The water flow rate in the riser is the most important explanatory variable for slugging, as seen in Section 10.2.2. It is therefore expected that the wells with a higher WCT is adjusted to lower values when slugging is considered. Well 6, on the other hand, has a lower WCT, which can explain why the recommended adjustment for the choke value is at the maximum upper limit.

Results from the bi-objective problem show that the recommendations from the optimization will generally move to a safer operation area with regards to slugging when the risk aversion of the decision maker is higher. Operating in a safer area will often decrease the total oil production. Operating in high slug risk areas increases the oil production, but will ultimately lead to decreased profits if slugging occurs.

Chapter 12

Concluding Remarks

The real-time optimization problem for petroleum production is considered in this thesis. The problem is concerned with increasing oil production from a field by performing minor adjustments on the operation settings. Historical measurement data is used to model well performance and relationships between different pressure measurements in the system, referred to as well models and pipeline pressure models, respectively. Historical data is also used to develop a slug severity model that predicts the flow assurance issue of slugging.

A stochastic mixed integer linear program (MILP) problem that integrates the well and pipeline pressure models and uses these data-driven models as scenarios, is formulated. The scenarios are generated by training models on varying subsets of the data. An extension to the stochastic model includes the objective of minimizing the conditional value-at-risk (CVaR) with regards to the gas capacity constraint. An additional extension adds the objective of minimizing the cost of slugging, taking the risk profile of the decision maker into account.

Important contributions of this thesis are novel approaches of integrating a regression-based and a classification-based machine learning technique in linear optimization. Pressures, flow rates and slugging are modelled with the two techniques, and formulated to be compliant with a MILP problem. Results show that the inclusion of the data-driven models highly impacts the solution time. The multivariate adaptive regression splines (MARS) models used to model pressures and flow rates are tested with three different complexities. The more advanced models with the highest complexity, yielding the optimal bias-variance trade-off, are incompatible with real-time optimization, as the solution time for the optimization problem is too high for the case studied.

The main challenge identified when integrating machine learning models in optimization is therefore how to allocate the processing power available in order to make best possible decisions within a reasonable time frame. This can be described as a trade-off between accurate machine learning models and the number of scenarios in the optimization. For the case in this thesis, the medium complexity models are found more appropriate than the low complexity models. The bias from the medium complexity models

are lower, and the variance can be covered by sufficiently many scenarios.

The bi-objective model that minimizes CVaR provides a method to take uncertainty into account with regards to an operational constraint. That is, in contrast to the stochastic optimization, there are no uncertainty limits on choke openings. The model allows the decision maker to control the conservativeness of the solution, and the optimal solutions are studied in a Pareto front. Results show that enabling the possibility of some of the scenarios to produce gas over the hard capacity, but at the same time minimizing these gas productions, the oil rates increase. The potential increase is connected with the risk aversion of the production engineer. Reducing the confidence level for CVaR enables the production engineers to extract more oil, with an increased risk of the true gas output to exceed a given capacity limit. This is seen in the Pareto front for the three confidence levels 0.80, 0.90 and 0.99, as the lowest confidence level on average has a 2.33% higher oil production than the highest confidence level, but thereby also higher gas production.

The bi-objective model that minimizes the cost of slugging presents a new data-driven approach of including risk of slugging directly in the optimization. Results from the bi-objective optimization verify the trade-off between risk of slugging and oil production, decreasing the objective function value when avoiding high-risk areas of slug. It is found that the water rate in the separator has the highest impact on slugging, as the wells that have high water cut (WCT) values are adjusted downwards to reach safe areas.

The fields of machine learning and mathematical programming are increasingly intertwined. Optimization is essential in machine learning algorithms, but this thesis substantiates that applying machine learning techniques in optimization also has great potential. Machine learning algorithms learn from experience, which in principle resembles the way operators at a platform learn to control the processes at the platform. However, the algorithms can accumulate unlimited experience, and will have good performance if the dataset provided is of a certain quality. The possibility of directly integrating these methods in optimization can automate many costly processes. The integration of machine learning techniques in optimization is relevant for many industries, and the formulations provided in this thesis might create value for other applications.

In conclusion, the formulations presented provide a comprehensive framework for modelling a petroleum production system with physical relations, and utilizing data measurements from fields. The novel methods of incorporating a regression-based and a classification-based machine learning technique offer a way of applying machine learning techniques directly in linear optimization. Decision support is provided for a decision maker who wants to consider the additional objectives of minimizing uncertainty with regards to gas production or minimizing risk of slugging. These extensions of the stochastic optimization model enable the decision maker to identify optimal solutions for varying levels of risk aversion.

Chapter 13

Future Research

Suggestions for further research topics are given in this chapter. Main challenges in this thesis have been identified in the computational study, with emphasis on the topics related to machine learning model complexity and solution time of the optimization model. With an increasing amount of data and available processing power, the research area of machine learning models in optimization has big potential. As found in the literature review, there is a wide range of research related to machine learning applied with optimization. However, the direct integration of the models, thus not only in a pre-optimization stage, has not been researched so far.

This thesis uses two different machine learning techniques, multivariate adaptive regression splines (MARS) and classification trees (CT). However, there are other machine learning techniques that could be studied that might be capable of achieving a higher level of accuracy. A possible technique for the well and pipeline pressure models are neural networks, given that the rectified linear unit activation function is similar to a hinge function. However, due to the increased number of binary variables from linear formulations of neural nets, the scalability would potentially be poor. Future research can seek to overcome scalability issues, or explore other techniques. An extension from simple decision trees that may decrease variance and bias is random forest, which combines several decision trees to get better predictive performance.

The scope of this thesis is how to cover the variance of the machine learning models in the optimization. However, in order to develop a framework for integration of machine learning models, a method of handling the total prediction uncertainty of the models, thus both variance and bias, should be further researched. Covering bias is a difficult task, as the quantification of this error is based on knowing the true values of the output variables in a model. A method of approximating the bias would have to be used, and could potentially be applied during the optimization by adjusting the output according to the bias of the model.

There are various options for performing scenario generation that might better capture the uncertainty related to the models. The scenario generation method used in this thesis was found difficult to interpret

with some of the results, such as the variance found in the low and medium complexity models for stability testing. In addition, the method does not generate a distribution of possible values of the stochastic variables, but rather a potentially infinite number of values. One possible scenario generation method that would similarly be based on the data-driven models is applying the Bayesian approach to MARS fitting, proposed by Denison et al. (1998) as mentioned in Section 8.3.

As described in Section 9.4, a challenge with data measurements from petroleum fields is that most wells have only been operated at a few settings. When the lack of data is significant, which is the situation for several of the wells in the case studied, using physical relationships might be more reliable. This gives rise to the topic of combining simulation and data-driven modelling. Both areas are highly researched, however not commonly in combination. Furthermore, the uncertainty in the data itself is not included in this thesis. It should be investigated further how such aleatoric uncertainty can be included into the optimization problem.

With regards to extending the model, routing is a relevant factor to include in a new model formulation. This will give a more realistic representation of the production systems. The current model only allows the routing from a well to one of the connected lines, but no further routing in the pipelines. Routing will give a new solution far away from the current operation point, which has to be handled in the data-driven models.

Bibliography

Al-Othman, W. B., Lababidi, H. M., Alatiqi, I. M., Al-Shayji, K., 2008. Supply chain optimization of petroleum organization under uncertainty in market demands and prices. European Journal of Operational Research 189 (3), 822–840.

Alarcón, G. A., Torres, C. F., Gómez, L. E., 2002. Global optimization of gas allocation to a group of wells in artificial lift using nonlinear constrained programming. Journal of energy resources technology 124 (4), 262–268.

Andreassen, R. S., Westby, E. M., 2018. Modelling and optimization of real-time petroleum production with slug management.

Anifowose, F. A., Labadin, J., Abdulraheem, A., 2017. Ensemble machine learning: An untapped modeling paradigm for petroleum reservoir characterization. Journal of Petroleum Science and Engineering 151, 480–487.

April, J., Glover, F., Kelly, J. P., Laguna, M., 2003. Simulation-based optimization: practical introduction to simulation optimization. In: Proceedings of the 35th conference on Winter simulation: driving innovation. Winter Simulation Conference, pp. 71–78.

Asheim, H., et al., 1988. Criteria for gas-lift stability. Journal of Petroleum Technology 40 (11), 1–452.

Attra, H., Wise, W., Black, W., et al., 1961. Application of optimizing techniques for studying field producing operations. Journal of Petroleum Technology 13 (01), 82–86.

Baker, O., et al., 1953. Design of pipelines for the simultaneous flow of oil and gas. In: Fall Meeting of the Petroleum Branch of AIME. Society of Petroleum Engineers.

Bendiksen, K. H., Maines, D., Moe, R., Nuland, S., et al., 1991. The dynamic two-fluid model olga: Theory and application. SPE production engineering 6 (02), 171–180.

Bieker, H. P., 2007. Topics in offshore oil production optimization using real-time data. PhD. The Thesis, Norwegian University of Science and Technology, Trondheim.

Bieker, H. P., Slupphaug, O., Johansen, T. A., et al., 2007a. Real-time production optimization of oil and gas production systems: A technology survey. SPE Production & Operations 22 (04), 382–391.

Bieker, H. P., Slupphaug, O., Johansen, T. A., et al., 2007b. Well management under uncertain gas or water oil ratios. In: Digital Energy Conference and Exhibition. Society of Petroleum Engineers.

Biltoft, J., Hansen, L., Pedersen, S., Yang, Z., 2013. Recreating riser slugging flow based on an economic labsized setup. In: IFAC International Workshop on Periodic Control, 5th. pp. 47–52.

Bodington, C. E., Baker, T. E., 1990. A history of mathematical programming in the petroleum industry. Interfaces 20 (4), 117–127.

Bratvedt, F., Bratvedt, K., Buchholz, C., Holden, L., Holden, H., Risebro, N. H., et al., 1992. A new front-tracking method for reservoir simulation. SPE Reservoir Engineering 7 (01), 107–116.

Camponogara, E., Plucenio, A., Teixeira, A. F., Campos, S. R., 2010. An automation system for gas-lifted oil wells: Model identification, control, and optimization. Journal of Petroleum Science and Engineering 70 (3-4), 157–167.

Chen, C.-H., Lin, Z.-S., 2006. A committee machine with empirical formulas for permeability prediction. Computers & Geosciences 32 (4), 485–496.

Codas, A., Camponogara, E., 2012. Mixed-integer linear optimization for optimal lift-gas allocation with well-separator routing. European Journal of Operational Research 217 (1), 222–231.

Cooper, A., Johnson, C., 2018. Now near 100 million bpd, when will oil demand peak?
URL https://www.reuters.com/article/us-oil-demand-peak/now-near-100-million-bpd-when-will-oil-demand-peak-idUSKCN1M01TC

Danielson, T. J., Bansal, K. M., Djoric, B., Duret, E.-D., Johansen, S. T., Hellan, O., et al., 2011. Testing and qualification of a new multiphase flow simulator. In: Offshore Technology Conference. Offshore Technology Conference.

de Boer, J., van der Linden, R., Renes, W., et al., 2016. Results of real-time production optimization of a maturing north sea gas asset with production constraints. In: SPE Intelligent Energy International Conference and Exhibition. Society of Petroleum Engineers.

Deb, K., 2014. Multi-objective optimization. In: Search methodologies. Springer, pp. 403–449.

Dempster, M., Pedron, N. H., Medova, E., Scott, J., Sembos, A., 2000. Planning logistics operations in the oil industry. Journal of the Operational Research Society 51 (11), 1271–1288.

Denison, D. G., Mallick, B. K., Smith, A. F., 1998. Bayesian mars. Statistics and Computing 8 (4), 337–346.

Der Kiureghian, A., Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter? Structural Safety 31 (2), 105–112.

Dutta-Roy, K., Kattapuram, J., 1997. A new approach to gas-lift allocation optimization. In: SPE western regional meeting. Society of Petroleum Engineers.

Economides, M. J., Hill, A. D., Ehlig-Economides, C., Zhu, D., 2013. Petroleum production systems. Pearson Education.

Elgsæter, S. M., Slupphaug, O., Johansen, T. A., 2010. A structured approach to optimizing offshore oil and gas production with uncertain models. Computers & chemical engineering 34 (2), 163–176.

Ewing, R. E., 1983. The mathematics of reservoir simulation. SIAM.

Fang, W., Lo, K., 1996. A generalized well management scheme for reservoir simulation. SPE Reservoir Engineering 11 (02), 116–120.

Fortmann-Roe, S., 2012. Understanding the bias-variance tradeoff.

Foss, B., Knudsen, B. R., Grimstad, B., 2018. Petroleum production optimization–a static or dynamic problem? Computers & Chemical Engineering 114, 245–253.

Friedman, J. H., et al., 1991. Multivariate adaptive regression splines. The annals of statistics 19 (1), 1–67.

Gao, S., 2008. Investigation of interactions between gas hydrates and several other flow assurance elements. Energy & Fuels 22 (5), 3150–3153.

Gharbi, R., Elsharkawy, A. M., et al., 1997. Neural network model for estimating the pvt properties of middle east crude oils. In: Middle East Oil Show and Conference. Society of Petroleum Engineers.

Gilbert, W., et al., 1954. Flowing and gas-lift well performance. In: Drilling and production practice. American Petroleum Institute.

Glæserud, P., Syrdalen, J. A., 2009. Optimization of petroleum production under uncertainty: Applied to the troll c field. Master's thesis, Norges teknisk-naturvitenskapelige universitet, Fakultet for samfunnsvitenskap og teknologiledelse, Institutt for industriell økonomi og teknologiledelse.

Goel, V., Grossmann, I. E., 2004. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers & chemical engineering 28 (8), 1409–1429.

Goh, K.-C., Moncur, C. E., Van Overschee, P., Briers, J., et al., 2007. Production surveillance and optimization with data driven models. In: International Petroleum Technology Conference. International Petroleum Technology Conference.

Griffith, P., Wallis, G. B., 1961. Two-phase slug flow. Journal of Heat Transfer 83 (3).

Grimstad, B., Foss, B., Heddle, R., Woodman, M., 2016. Global optimization of multiphase flow networks using spline surrogate models. Computers & Chemical Engineering 84, 237–254.

Gunnerud, V., Foss, B., 2010. Oil production optimization – a piecewise linear model, solved with two decomposition strategies. Computers & Chemical Engineering 34 (11), 1803–1812.

Gunnerud, V., Foss, B. A., McKinnon, K., Nygreen, B., 2012. Oil production optimization solved by piecewise linearization in a branch & price framework. Computers & Operations Research 39 (11), 2469–2477.

Gunnerud, V., Langvik, E., 2007. Production planning optimization for the troll c field. Master's thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Hanssen, K. G., Foss, B., 2015. Production optimization under uncertainty-applied to petroleum production. IFAC-PapersOnLine 48 (8), 217–222.

Hanssen, K. G., Foss, B., Teixeira, A., 2015. Production optimization under uncertainty with constraint handling. IFAC-PapersOnLine 48 (6), 62–67.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Model assessment and selection. In: The elements of statistical learning. Springer, pp. 219–259.

Haug, E. R., 2013. Subsea investments and well placements in the oil production industry. Master's thesis, Institutt for industriell økonomi og teknologiledelse.

Higle, J. L., 2005. Stochastic programming: optimization when uncertainty matters. In: Emerging Theory, Methods, and Applications. Informs, pp. 30–53.

Huseby, A. B., Haavardsson, N. F., 2009. Multi-reservoir production optimization under uncertainty. Reliability, Risk and Safety. Theory and Applications, 407–413.

Jahanbani, A., Shadizadeh, S., et al., 2009. Determination of inflow performance relationship (ipr) by well testing. In: Canadian International Petroleum Conference. Petroleum Society of Canada.

Jaimes, A. L., Martınez, S. Z., Coello, C. A. C., 2009. An introduction to multiobjective optimization techniques. Optimization in Polymer Processing, 29–57.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning. Vol. 112. Springer.

Jansen, F., Shoham, O., Taitel, Y., 1996. The elimination of severe sluggingexperiments and modeling. International journal of multiphase flow 22 (6), 1055–1072.

Jansen, J.-D., Bosgra, O. H., Van den Hof, P. M., 2008. Model-based control of multiphase flow in subsurface oil reservoirs. Journal of Process Control 18 (9), 846–855.

Jansen, J.-D., Brouwer, R., Douma, S. G., et al., 2009. Closed loop reservoir management. In: SPE reservoir simulation symposium. Society of Petroleum Engineers.

Jiang, D., Li, S., Edge, K. A., Zeng, W., 2012. Modeling and simulation of low pressure oil-hydraulic pipeline transients. Computers & Fluids 67, 79–86.

Jørnsten, K. O., 1992. Sequencing offshore oil and gas fields under uncertainty. European Journal of Operational Research 58 (2), 191–201.

Joshi, N. B., Muhammad, M., Creek, J. L., McFadden, J., et al., 2003. Flow assurance: A challenging path to well completions and productivity. In: Offshore Technology Conference. Offshore Technology Conference.

Kaut, M., Wallace, S. W., 2003. Evaluation of scenario-generation methods for stochastic programming.

Kazemi, H., Merrill Jr, L., Porterfield, K., Zeman, P., et al., 1976. Numerical simulation of water-oil flow in naturally fractured reservoirs. Society of Petroleum Engineers Journal 16 (06), 317–326.

King, A. J., Wallace, S. W., 2012. Modeling with stochastic programming. Springer Science & Business Media.

Koc, E. K., Bozdogan, H., 2015. Model selection in multivariate adaptive regression splines (mars) using information complexity as the fitness function. Machine Learning 101 (1-3), 35–58.

Krishnamoorthy, D., Foss, B., Skogestad, S., 2016. Real-time optimization under uncertainty applied to a gas lifted well network. Processes 4 (4), 52.

Lang, Z., Horne, R., et al., 1983. Optimum production scheduling using reservoir simulators: a comparison of linear programming and dynamic programming techniques. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Langvik, A. S., Dzubur, L., 2012. Optimization of oil production-applied to the marlim field. Master's thesis, Institutt for industriell økonomi og teknologiledelse.

Li, X., Chan, C. W., 2010. Application of an enhanced decision tree learning approach for prediction of petroleum production. Engineering Applications of Artificial Intelligence 23 (1), 102–109.

Linden, R. V. D., Busking, T., 2013. Real-time optimization of a maturing north sea gas asset with production constraints. In: SPE Middle East Intelligent Energy Conference and Exhibition. Society of Petroleum Engineers.

Ling, Z.-f., Wang, L.-j., Hu, Y.-l., Li, B.-z., 2008. Flood pattern optimization of horizontal well injection. Petroleum Exploration and Development 35 (1), 85–91.

Madansky, A., 1960. Inequalities for stochastic linear programming problems. Management science 6 (2), 197–204.

Martinez, N., Anahideh, H., Rosenberger, J. M., Martinez, D., Chen, V. C., Wang, B. P., 2017. Global optimization of non-convex piecewise linear regression splines. Journal of Global Optimization 68 (3), 563–586.

Mayhill, T., et al., 1974. Simplified method for gas-lift well problem identification and diagnosis. In: Fall Meeting of the Society of Petroleum Engineers of AIME. Society of Petroleum Engineers.

Mora Buitrago, O. S., Starzman, R. A., Saputelli, L. A., et al., 2005. Maximizing net present value in mature gas lift fields. In: SPE Hydrocarbon Economics and Evaluation Symposium. Society of Petroleum Engineers.

Morken, M. L., Sandberg, P. T., 2016. Modelling and optimization of real-time petroleum production-using robust regression, bootstrapping, moment matching, and two-stage stochastic optimization. Master's thesis, NTNU.

Neiro, S. M., Pinto, J. M., 2004. A general modeling framework for the operational planning of petroleum supply chains. Computers & Chemical Engineering 28 (6-7), 871–896.

Nishikiori, N., Redner, R., Doty, D., Schmidt, Z., 1995. An improved method for gas lift allocation optimization. Journal of energy resources technology 117 (2), 87–92.

Norwegian Petroleum Directorate, 2018. Oil and gas production.
URL https://www.norskpetroleum.no/en/production-and-exports/oil-and-gas-production/

Nygreen, B., Christiansen, M., Haugen, K., Bjørkvoll, T., Kristiansen, Ø., 1998. Modeling norwegian petroleum production and transportation. Annals of Operations Research 82, 251–268.

Ofonmbuk, U., Shnaib, F., Nadar, M., Davies, D., Ifechukwu, M., Oladepo, D. A., et al., 2015. Real-time production optimization application in offshore gas lifted fields. In: SPE Nigeria Annual International Conference and Exhibition. Society of Petroleum Engineers.

Oloso, M. A., Hassan, M. G., Bader-El-Den, M. B., Buick, J. M., 2017. Hybrid functional networks for oil reservoir pvt characterisation. Expert Systems with Applications 87, 363–369.

OPEC, 2019. Opec basket price.
URL https://www.opec.org/opec_web/en/data_graphs/40.htm

Orkiszewski, J., et al., 1967. Predicting two-phase pressure drops in vertical pipe. Journal of Petroleum Technology 19 (06), 829–838.

Pagano, D., Plucenio, A., Traple, A., Gonzaga, C., 2008. Controlling oscillations and re-starting operations in gas-lift wells. In: Proc. XVII Congresso Brasileiro de Automática, Juiz de Fora, Brasil.

Panja, P., Velasco, R., Pathak, M., Deo, M., 2018. Application of artificial intelligence to forecast hydrocarbon production from shales. Petroleum 4 (1), 75–89.

Pauchon, C., Dhulesia, H., et al., 1994. Tacite: A transient tool for multiphase pipeline and well simulation. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Pedersen, S., Durdevic, P., Yang, Z., 2015. Review of slug detection, modeling and control techniques for offshore oil & gas production processes. Ifac-papersonline 48 (6), 89–96.

Pedersen, S., Durdevic, P., Yang, Z., 2017. Challenges in slug modeling and control for offshore oil and gas productions: A review study. International Journal of Multiphase Flow 88, 270–284.

Redden, J. D., Sherman, T., Blann, J. R., et al., 1974. Optimizing gas-lift systems. In: Fall Meeting of the Society of Petroleum Engineers of AIME. Society of Petroleum Engineers.

Ribas, G. P., Hamacher, S., Street, A., 2010. Optimization under uncertainty of the integrated oil supply chain using stochastic and robust programming. International Transactions in Operational Research 17 (6), 777–796.

Rockafellar, R. T., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. Journal of banking & finance 26 (7), 1443–1471.

Rockafellar, R. T., Uryasev, S., et al., 2000. Optimization of conditional value-at-risk. Journal of risk 2, 21–42.

Saputelli, L., Malki, H., Canelon, J., Nikolaou, M., et al., 2002. A critical overview of artificial neural network applications in the context of continuous oil field optimization. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers.

Sarica, C. t., Shoham, O., 1991. A simplified transient model for pipeline-riser systems. Chemical engineering science 46 (9), 2167–2179.

Sausen, A., Sausen, P., De Campos, M., 2012. The slug flow problem in oil industry and pi level control. In: New Technologies in the Oil and Gas Industry. InTech.

Schlumberger, 2019. Schlumberger glossary.

Schmidt, Z., Brill, J., Beggs, H., 1979. Choking can eliminate severe pipeline slugging. Oil & gas journal 77 (46), 230.

Schmidt, Z., Brill, J. P., Beggs, H. D., et al., 1980. Experimental study of severe slugging in a two-phase-flow pipeline-riser pipe system. Society of Petroleum Engineers Journal 20 (05), 407–414.

Shah, N. K., Li, Z., Ierapetritou, M. G., 2010. Petroleum refining operations: key issues, advances, and opportunities. Industrial & Engineering Chemistry Research 50 (3), 1161–1170.

Shokir, E. M. E.-M., Hamed, M. M., Ibrahim, A. E.-S., Mahgoub, I., 2017. Gas lift optimization using artificial neural network and integrated production modeling. Energy & Fuels 31 (9), 9302–9307.

Simmons, W., 1972. Optimizing continuous flow gas lift wells. pt. 2. Pet. Eng.;(United States) 44 (10).

Solomatine, D. P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. Journal of hydroinformatics 10 (1), 3.

Stoisits, R., Batesole, E., Champion, J., Park, D., et al., 1992. Application of nonlinear adaptive modeling for rigorous representation of production facilities in reservoir simulation. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Stoisits, R., Crawford, K., MacAllister, D., McCormack, M., Lawal, A., Ogbe, D., et al., 1999. Production optimization at the kuparuk river field utilizing neural networks and genetic algorithms. In: SPE mid-continent operations symposium. Society of Petroleum Engineers.

Taitel, Y., 1986. Stability of severe slugging. International journal of multiphase flow 12 (2), 203–217.

Taitel, Y., Bornea, D., Dukler, A., 1980. Modelling flow pattern transitions for steady upward gas-liquid flow in vertical tubes. AIChE Journal 26 (3), 345–354.

Tang, Y., Danielson, T. J., et al., 2006. Pipelines slugging and mitigation: case study for stability and production optimization. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Tarhan, B., Grossmann, I. E., Goel, V., 2009. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. Industrial & Engineering Chemistry Research 48 (6), 3078–3097.

Teixeira, B. O., Castro, W. S., Teixeira, A. F., Aguirre, L. A., 2014. Data-driven soft sensor of downhole pressure for a gas-lift oil well. Control Engineering Practice 22, 34–43.

Tengesdal, J., Sarica, C., Thompson, L., et al., 2002. Severe slugging attenuation for deepwater multiphase pipeline and riser systems. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Trangenstein, J. A., Bell, J. B., 1989. Mathematical structure of the black-oil model for petroleum reservoir simulation. SIAM Journal on Applied Mathematics 49 (3), 749–783.

Ulstein, N. L., Nygreen, B., Sagli, J. R., 2007. Tactical planning of offshore petroleum production. European Journal of Operational Research 176 (1), 550–564.

Ursin-Holm, S., Shamlou, S. S., 2013. Simulation based optimization of petroleum production problems: Development of a special purpose b&b for a non-convex minlp. Master's thesis, Institutt for industriell økonomi og teknologiledelse.

Uryasev, S., 2000. Conditional value-at-risk: Optimization algorithms and applications. In: Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr)(Cat. No. 00TH8520). IEEE, pp. 49–57.

Uryasev, S., Sarykalin, S., Serraino, G., Kalinchenko, K., 2010. Var vs cvar in risk management and optimization. In: CARISMA conference.

Vassiliou, M. S., 2018. Historical dictionary of the petroleum industry. Rowman & Littlefield.

Viggiani, M., Mariani, O., Battarra, V., Annunziato, A., Bollettini, U., et al., 1988. A model to verify the onset of severe slugging. In: PSIG Annual Meeting. Pipeline Simulation Interest Group.

Vogel, J., et al., 1968. Inflow performance relationships for solution-gas drive wells. Journal of petroleum technology 20 (01), 83–92.

Wang, P., 2003. Development and applications of production optimization techniques for petroleum fields. Ph.D. thesis, Stanford University Stanford.

Wang, P., Litvak, M., Aziz, K., 2002a. Optimization of production from mature fields. paper no. 32152. In: 17th World Petroleum Congress, Rio de Janeiro, Brazil, September. pp. 1–5.

Wang, P., Litvak, M., Aziz, K., et al., 2002b. Optimization of production operations in petroleum fields. In: SPE annual technical conference and exhibition. Society of Petroleum Engineers.

Weskamp, C., Koberstein, A., Schwartz, F., Suhl, L., Voß, S., 2019. A two-stage stochastic programming approach for identifying optimal postponement strategies in supply chains with uncertain demand. Omega 83, 123–138.

Willersrud, A., Imsland, L., Hauger, S. O., Kittilsen, P., 2013. Short-term production optimization of offshore oil and gas production using nonlinear model predictive control. Journal of Process Control 23 (2), 215–223.

Xing, L., Yeung, H., Shen, J., Cao, Y., 2013. Experimental study on severe slugging mitigation by applying wavy pipes. Chemical Engineering Research and Design 91 (1), 18–28.

Yocum, B., et al., 1973. Offshore riser slug flow avoidance: mathematical models for design and optimization. In: SPE European Meeting. Society of Petroleum Engineers.

Zhang, Y., Monder, D., Forbes, J. F., 2002. Real-time optimization under parametric uncertainty: a probability constrained approach. Journal of Process control 12 (3), 373–389.

Ziegel, E., Shirzadi, S., Wang, S., Bailey, R., Griffiths, P., Ghuwalela, K., Ogedengbe, A., Johnson, D., et al., 2014. A data-driven approach to modelling and optimization for a north sea asset using real-time data. In: SPE Intelligent Energy Conference & Exhibition. Society of Petroleum Engineers.

Appendix A

Full Formulation of Optimization Model

This appendix presents the full formulation of the stochastic optimization model presented in Section 6.2.

Sets

Ω	Set of scenarios, indexed ω
$\mathcal{I}$	Set of wells, indexed i
$\mathcal{L}$	Set of pipelines, indexed l
$\mathcal{M}$	Set of manifolds, indexed m
$\mathcal{P}$	Set of phases (o for oil, g for gas, w for water), indexed p
$\mathcal{S}$	Set of separators, indexed s
$\mathcal{N}$	Set of nodes, $\mathcal{N} \in \{I \cup \mathcal{M} \cup \mathcal{S}\}$, indexed j, k
$\mathcal{I}_m$	Set of wells connected to manifold m, $\mathcal{I}_m \in \mathcal{I}$, indexed i
$\mathcal{L}_m$	Set of pipelines connected to manifold m, $\mathcal{L}_m \in \mathcal{L}$, indexed l
$\mathcal{L}_s$	Set of pipelines connected to separator s, $\mathcal{L}_s \in \mathcal{L}$, indexed l
$\mathcal{M}^O$	Set of manifolds first in the chain, $\mathcal{M}^O \in \mathcal{M}$, indexed m
$\mathcal{M}^M$	Set of manifolds connected to two other manifolds, $\mathcal{M}^{\mathcal{M}} \in \mathcal{M}$, indexed m
$\mathcal{M}^S$	Set of manifolds directly linked to a separator, $\mathcal{M}^{\mathcal{S}} \in \mathcal{M}$, indexed m

Parameters

U_i	Initial choke opening of well i
P_i	Initial wellhead pressure for well i
P_i^D	Initial downstream choke pressure for well i
P_s^S	Inlet pressure for separator s
Q_{pi}	Initial flow of phase p in well i
$\overline{U}_i$	Upper limit of choke opening for well i
$\underline{U}_i$	Lower limit of choke opening for well i
$\Delta\overline{U}_i$	Maximum change in choke opening for well i
$\overline{P}_i$	Maximum wellhead pressure for well i
$\overline{P}_{ml}$	Maximum pressure at manifold m in pipeline l
GOR_i	Initial gas-oil ratio for well i
WCT_i	Initial water cut for well i
C_p	Maximum capacity of phase p from wells to manifolds
C_{ps}	Maximum capacity of phase p in separator s
C^G	Total gas capacity of platform
C^W	Total water capacity of platform
D	Maximum number of chokes that can be adjusted

Variables

Δu_i	Delta choke opening of well i
$\Delta p_{i\omega}$	Delta wellhead pressure for well i in scenario ω
$\Delta p_{i\omega}^D$	Delta downstream choke pressure for well i in scenario ω
$\Delta p_{ml\omega}$	Delta pressure at manifold m in pipeline l in scenario ω
$\Delta q_{pi\omega}$	Delta flow of phase p from well i in scenario ω
$\Delta q_{pjkl\omega}$	Delta flow of phase p from node j to node k in pipeline l in scenario ω
$\Delta q_{ps\omega}^S$	Delta flow of phase p to separator s in scenario ω
u_i	Choke opening of well i
$p_{i\omega}$	Wellhead pressure for well i in scenario ω
$p_{i\omega}^D$	Downstream choke pressure for well i in scenario ω
$p_{ml\omega}$	Pressure at manifold m in pipeline l in scenario ω
$q_{pi\omega}$	Flow of phase p from well i in scenario ω
$q_{pjkl\omega}$	Flow of phase p from node j to node k in pipeline l in scenario ω
$q_{ps\omega}^S$	Total flow of phase p to separator s in scenario ω
x_i	1 if well i is open, 0 otherwise
y_{iml}	1 if well i connected to manifold m is open and connected to pipeline l, 0 otherwise
h_i	1 if well i is adjusted, 0 otherwise

Functions

$f_{i\omega}^{OIL}(X)$	Function mapping from the inputs X to the output for the oil flow rate model for well i in scenario ω
$f_{i\omega}^{PRS}(X)$	Function mapping from the inputs X to the output for the wellhead pressure model for well i in scenario ω
$f_{ml\omega}^{PRS}(X)$	Function mapping from the inputs X to the output for the pipeline pressure model for manifold m in pipeline l in scenario ω

A.1 Objective Function

$$max\ w = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \sum_{s \in \mathcal{S}} q^S_{os\omega} \tag{A.1}$$

A.2 Constraints

$$\sum_{s \in \mathcal{S}} q^S_{gs\omega} \leq C^G, \quad \omega \in \Omega \tag{A.2}$$

$$\sum_{s \in \mathcal{S}} q^S_{ws\omega} \leq C^W, \quad \omega \in \Omega \tag{A.3}$$

$$u_i = U_i + \Delta u_i \quad i \in \mathcal{I} \tag{A.4}$$

$$\Delta u_i - h_i(\overline{U}_i - U_i) \leq 0, \quad i \in \mathcal{I} \tag{A.5}$$

$$\Delta u_i + h_i(U_i - \underline{U}_i) \geq 0, \quad i \in \mathcal{I} \tag{A.6}$$

$$\sum_{i \in \mathcal{I}} h_i \leq D \tag{A.7}$$

$$u_i \geq \underline{U}_i x_i, \quad i \in \mathcal{I} \tag{A.8}$$

$$u_i \leq \overline{U}_i x_i, \quad i \in \mathcal{I} \tag{A.9}$$

$$\sum_{l \in \mathcal{L}_m} y_{iml} - x_i = 0, \quad m \in \mathcal{M}, i \in \mathcal{I}_m \tag{A.10}$$

$$\sum_{l \in \mathcal{L}_m} q_{piml\omega} = q_{pi\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, \omega \in \Omega \tag{A.11}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} = q_{pm(m+1)l\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^O, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.12}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} + q_{p(m-1)ml\omega} = q_{pm(m+1)l}, \quad p \in \mathcal{P}, m \in \mathcal{M}^M, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.13}$$

$$\sum_{i \in \mathcal{I}_m} q_{piml\omega} + q_{p(m-1)ml\omega} = q_{pmsl\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{A.14}$$

$$q_{piml\omega} \leq C_p y_{iml}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.15}$$

$$\sum_{m \in \mathcal{M}^S} \sum_{l \in \mathcal{L}_s} q_{pmsl\omega} = q^S_{ps\omega}, \quad p \in \mathcal{P}, s \in \mathcal{S}, \omega \in \Omega \tag{A.16}$$

$$\sum_{m \in \mathcal{M}^S} q_{pmsl\omega} \leq C_{ps}, \quad p \in \mathcal{P}, s \in \mathcal{S}, l \in \mathcal{L}_s, \omega \in \Omega \tag{A.17}$$

$$\sum_{l \in \mathcal{L}_m} \Delta q_{piml\omega} = \Delta q_{pi\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, \omega \in \Omega \tag{A.18}$$

$$\sum_{i \in \mathcal{I}_m} \Delta q_{piml\omega} = \Delta q_{pm(m+1)l\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^O, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.19}$$

$$\sum_{i\in\mathcal{I}_m} \Delta q_{piml\omega} + \Delta q_{p(m-1)ml\omega} = \Delta q_{pm(m+1)l}, \quad p \in \mathcal{P}, m \in \mathcal{M}^M, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.20}$$

$$\sum_{i\in\mathcal{I}_m} \Delta q_{piml\omega} + \Delta q_{p(m-1)ml\omega} = \Delta q_{pmsl\omega}, \quad p \in \mathcal{P}, m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{A.21}$$

$$\Delta q_{piml\omega} \leq C_p y_{iml}, \quad p \in \mathcal{P}, m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.22}$$

$$\sum_{m\in\mathcal{M}^S} \sum_{l\in\mathcal{L}_s} \Delta q_{pmsl\omega} = \Delta q^S_{ps\omega}, \quad p \in \mathcal{P}, s \in \mathcal{S}, \omega \in \Omega \tag{A.23}$$

$$\sum_{m\in\mathcal{M}^S} \Delta q_{pmsl\omega} \leq C_{ps}, \quad p \in \mathcal{P}, s \in \mathcal{S}, l \in \mathcal{L}_s, \omega \in \Omega \tag{A.24}$$

$$\Delta q_{oi\omega} = f^{OIL}_{i\omega}(\Delta p_{i\omega}, \Delta u_i), \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.25}$$

$$q_{oi\omega} = Q_{oi} + \Delta q_{oi\omega} \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.26}$$

$$q_{oi\omega} \leq C_o x_i, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.27}$$

$$q_{gi\omega} = GOR_i q_{oi\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.28}$$

$$q_{wi\omega} = \left(\frac{WCT_i}{1 - WCT_i}\right) q_{oi\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.29}$$

$$\Delta q_{gi\omega} = GOR_i \Delta q_{oi\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.30}$$

$$\Delta q_{wi\omega} = \left(\frac{WCT_i}{1 - WCT_i}\right) \Delta q_{oi\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.31}$$

$$\Delta p_{i\omega} = f^{PRS}_{i\omega}(\Delta p^D_{i\omega}, \Delta u_i, \Delta q_{oi\omega}, \Delta q_{gi\omega}, \Delta q_{wi\omega}), \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.32}$$

$$p_{i\omega} = P_i + \Delta p_{i\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.33}$$

$$p^D_{i\omega} = P^D_i + \Delta p^D_{i\omega}, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.34}$$

$$\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta p_{(m+1)l\omega}, \Delta q_{om(m+1)l\omega}, \Delta q_{gm(m+1)l\omega}, \Delta q_{wm(m+1)l\omega}), \quad m \in \mathcal{M}\backslash\{\mathcal{M}^S\}, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.35}$$

$$\Delta p_{ml\omega} = f^{PRS}_{ml\omega}(\Delta q_{omsl\omega}, \Delta q_{gmsl\omega}, \Delta q_{wmsl\omega}), \quad m \in \mathcal{M}^S, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{A.36}$$

$$p_{ml\omega} = p^D_{i\omega}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.37}$$

$$p_{ml\omega} = P_{ml} + \Delta p_{ml\omega}, \quad m \in \mathcal{M}, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.38}$$

$$p_{ml\omega} y_{iml} \leq p_{i\omega}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.39}$$

$$p_{ml\omega} y_{im^*l} \leq p_{(m-1)l\omega}, \quad m^*, m \in \mathcal{M}\backslash\{\mathcal{M}^O\}, m^* < m, i \in \mathcal{I}_{m^*}, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.40}$$

$$P^S_s y_{im^*l} \leq p_{ml\omega}, \quad m^* \in \mathcal{M}, m \in \mathcal{M}^S, i \in \mathcal{I}_{m^*}, s \in \mathcal{S}, l \in \{\mathcal{L}_m \cap \mathcal{L}_s\}, \omega \in \Omega \tag{A.41}$$

A.3 Variable Bounds

$$\Delta u_i \in (max\{\underline{U}_i - U_i, -\Delta\overline{U}_i\}, min\{\overline{U}_i - U_i, \Delta\overline{U}_i\}), \quad i \in \mathcal{I} \tag{A.42}$$

$$u_i \in (\underline{U}_i, \overline{U}_i), \quad i \in \mathcal{I} \tag{A.43}$$

$$p_{i\omega} \geq 0, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.44}$$

$$p^D_{i\omega} \geq 0, \quad i \in \mathcal{I}, \omega \in \Omega \tag{A.45}$$

$$p_{ml\omega} \geq 0, \quad m \in \mathcal{M}, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.46}$$

$$q_{pi\omega} \geq 0, \quad p \in \mathcal{P}, i \in \mathcal{I}, \omega \in \Omega \tag{A.47}$$

$$q_{pjkl\omega} \geq 0, \quad p \in \mathcal{P}, j \in \mathcal{N}, k \in \mathcal{N}, l \in \mathcal{L}, \omega \in \Omega \tag{A.48}$$

$$q^S_{ps\omega} \geq 0, \quad p \in \mathcal{P}, s \in \mathcal{S}, \omega \in \Omega \tag{A.49}$$

$$x_i \in \{0,1\}, \quad i \in \mathcal{I} \tag{A.50}$$

$$y_{iml} \in \{0,1\}, \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m \tag{A.51}$$

$$h_i \in \{0,1\}, \quad i \in \mathcal{I} \tag{A.52}$$

A.4 Relaxation of Constraints (6.30) and (6.31)

$$p_{ml\omega} \leq p_{i\omega} + \overline{P}_i(1 - y_{iml}), \quad m \in \mathcal{M}, i \in \mathcal{I}_m, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.53}$$

$$p_{ml\omega} \leq p_{(m-1)l\omega} + \overline{P}_{(m-1)l}(1 - y_{im^*l}), \quad m^*, m \in \mathcal{M}\backslash\{\mathcal{M}^O\}, m^* < m, i \in \mathcal{I}_{m^*}, l \in \mathcal{L}_m, \omega \in \Omega \tag{A.54}$$

Appendix B

Well Model Uncertainty Algorithm

The well model uncertainty algorithm is presented in this appendix.

Algorithm 1: Finding upper and lower delta limits of a choke

Data: Sorted set of $\mathcal{N}$ choke openings for a well, named $chokes$, sorted set of mean values related to each of the $\mathcal{N}$ choke openings for a well, named $averages$, sorted set of standard deviation values related to each of the $\mathcal{N}$ means of a well, named SD, and initial choke opening for a well, named U

Result: Upper and lower limit of change for a particular choke opening

index ← index of *U* in chokes;
CV ← the coefficient of variation;
original SD← sd[*index*];
original mean← averages[*index*];
if *original SD/original mean* > *CV* **then**
    return *infeasible operation area*;
end
pos ← *index*;
current SD ← *original SD*;
current mean ← *original mean*;
while *current SD/current mean* < *CV and pos* < *length of* $\mathcal{N}$ **do**
    increase *pos* by 1;
    current SD ← sd [*pos*];
    current mean ← averages [*pos*];
end
upper delta ← chokes [*pos*-1] - U ;
pos ← *index*;
current SD ← *original SD*;
current mean ← *original mean*;
while *current SD/current mean* < *CV and pos* > *0* **do**
    decrease *pos* by 1;
    current SD ← sd [*pos*];
    current mean ← averages [*pos*];
end
lower delta ← chokes [*pos*+1] - U;
return *upper delta, lower delta*

Appendix C

Computational Study Parameters

This appendix presents the additional parameters used in both initial cases in Section 9.5.

Table C.1: General parameter values for the initial cases

Parameter symbol	Parameter	Value
P_s^S	Inlet pressure for separator s	21.56
$\Delta\overline{U}_i$	Maximum change in choke opening for well i	20.00
$\overline{P}_i$	Maximum wellhead pressure for well i	10000.00
$\overline{P}_{ml}$	Maximum pressure in line l at manifold m	10000.00
C_p	Maximum capacity of phase p of a pipeline from a well to a manifold	10000.00
C_{ps}	Maximum capacity of phase p in separator s	10000.00
C^G	Total gas capacity of platform	300000.00
C^W	Total water capacity of platform	10000.00

Note that the capacities equal to 10 000 in Table C.1 are set so they cannot restrict solutions.

Appendix D

α-Weighted Variables Analysis

This appendix presents the results from the analysis of α-weighted variables in the MARS models as described in Section 10.1.1. The average GCV-score for different number of α-weighted variables are plotted for the MARS models.

In the following tables, PDC refers to the downstream choke pressure which equals the manifold pressure, QOIL refers to the oil flow rate, QGAS refers to the gas flow rate, and QWAT refers to the water flow rate.

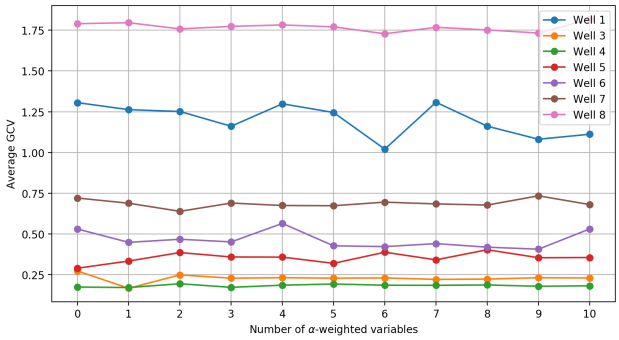

Figure D.1: Alpha analysis of oil flow rate models with interaction between choke and wellhead pressure

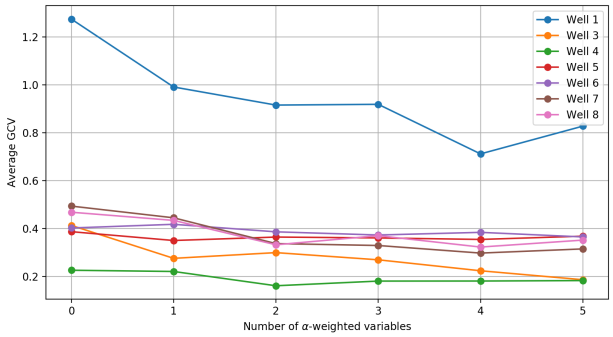

Figure D.2: Alpha analysis of wellhead pressure models with interaction between choke and downstream pressure

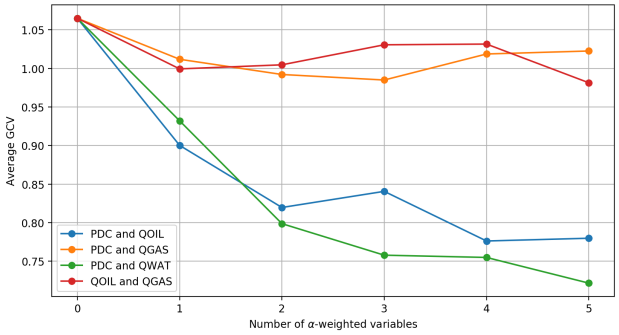

Figure D.3: Alpha analysis of pipeline pressure model for Manifold 1

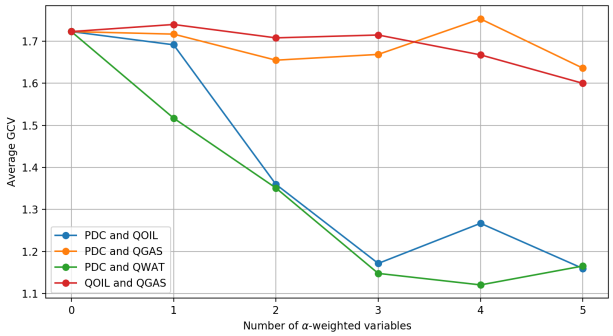

Figure D.4: Alpha analysis of pipeline pressure model for Manifold 2

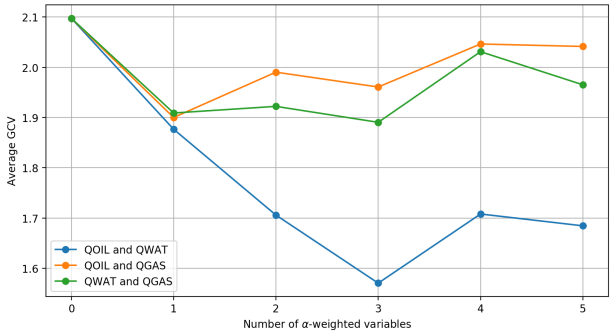

Figure D.5: Alpha analysis of pipeline pressure model for Manifold 3

Appendix E

Grid Search Results MARS Models

This appendix presents the grid search results from the hyperparameter tuning of the MARS models. The presentation of the grid search is found in Section 10.1.2.

Table E.1, Table E.2 and Table E.3 present the hyperparameters and test scores for the low and medium complexities for the oil flow rate models, wellhead pressure models and pipeline pressure models, respectively. The test scores, MSE and R^2, are found from the grid search.

Table E.1: Hyperparameters for the low and medium complexity oil flow rate models

	Low					Medium				
Well	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2
W1	140	3	1	1.94	0.92	70	6	1	1.88	0.92
W3	80	2	3	1.16	0.84	40	5	3	0.56	0.91
W4	24	2	5	1.70	0.89	12	3	5	1.81	0.85
W5	28	3	2	0.49	0.97	14	7	2	0.60	0.96
W6	8	3	2	0.89	0.96	4	6	2	0.78	0.96
W7	4	3	1	1.14	0.94	2	6	1	0.98	0.95
W8	100	2	20	2.96	0.84	50	4	20	1.72	0.90

Table E.2: Hyperparameters for the low and medium complexity wellhead pressure models

	Low					Medium				
Well	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2
W1	480	3	3	21.1	0.13	240	7	3	7.64	0.69
W3	160	3	10	2.38	0.69	80	7	10	1.57	0.81
W4	360	2	10	7.44	0.34	180	3	10	7.77	0.29
W5	24	5	1	2.48	0.80	12	10	1	1.69	0.87
W6	40	6	3	2.42	0.87	20	12	3	2.08	0.89
W7	120	3	1	21.1	0.1	60	5	1	3.49	0.83
W8	280	5	1	4.59	0.50	140	10	1	3.58	0.65

Table E.3: Hyperparameters for the low and medium complexity pipeline pressure models

	Low					Medium				
Manifold	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2	**Penalty**	**Max terms**	**Minspan**	**MSE**	R^2
M1	4	3	2	2.57	0.96	2	7	2	1.90	0.97
M2	4	4	2	2.45	0.96	2	9	2	1.48	0.98
M3	36	4	25	2.95	0.28	18	9	25	2.46	0.42

Table E.4 and Table E.5 present the test scores for all three complexities of well models and pipeline pressure models. The averages of the scores are found in Table 10.3. The test scores for each well and complexity are found from the average of 50 different train subsets when tested on the test dataset.

Table E.4: Test scores on test dataset for low, medium and high complexity well models

	Oil flow rate model						Wellhead pressure model					
	Low		**Medium**		**High**		**Low**		**Medium**		**High**	
Well	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2
W1	4.39	0.77	4.89	0.74	5.22	0.72	11.95	0.90	16.01	0.87	19.53	0.84
W3	1.05	0.21	2.59	-0.94	2.91	-1.18	11.45	-0.33	9.09	-0.06	4.97	0.42
W4	1.85	0.90	2.72	0.85	6.52	0.64	1.93	0.98	3.64	0.96	13.58	0.86
W5	1.32	0.39	1.49	0.31	1.45	0.33	32.65	0.24	34.52	0.20	28.30	0.35
W6	3.79	0.77	10.62	0.36	8.43	0.49	32.23	0.41	41.19	0.25	34.42	0.37
W7	2.68	0.77	3.29	0.71	4.04	0.65	30.23	0.10	28.31	0.15	21.58	0.36
W8	3.09	0.52	3.76	0.42	5.36	0.17	3.55	0.74	18.84	-0.37	15.77	-0.15

Table E.5: Test scores on test dataset for low, medium and high complexity pipeline pressure models

	Low		**Medium**		**High**	
Manifold	**MSE**	R^2	**MSE**	R^2	**MSE**	R^2
M1	4.56	-2.00	12.07	-6.94	14.63	-8.62
M2	0.78	0.19	1.13	-0.18	1.29	-0.35
M3	2.73	-2.72	4.41	-5.04	4.60	-5.28

Appendix F

Classification Tree

This appendix presents an enlarged version of the classification tree presented in Section 10.2.

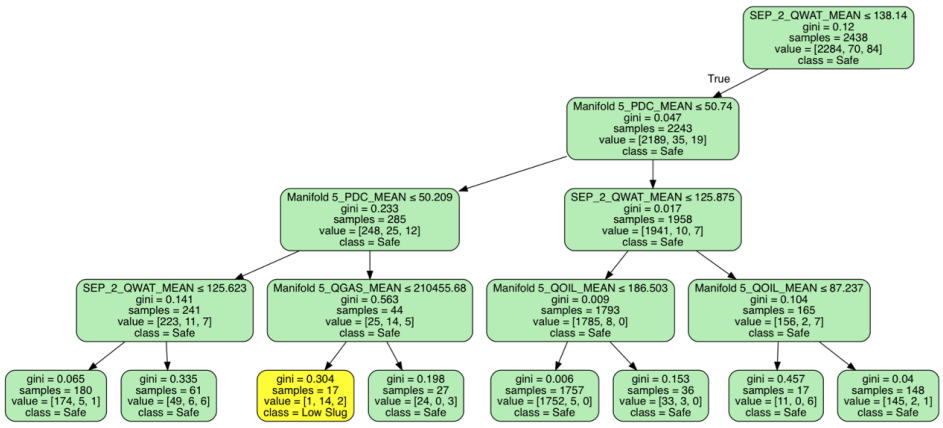

Figure F.1: Enlarged version of left branch of final classification tree

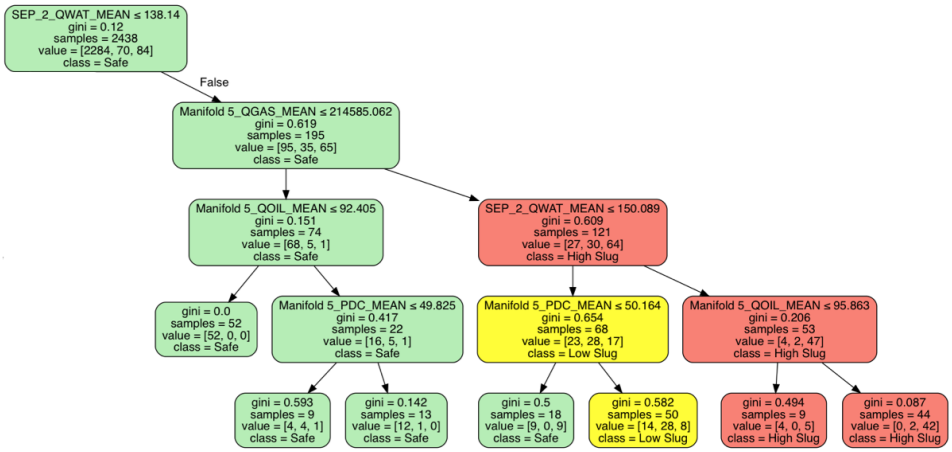

Figure F.2: Enlarged version of right branch of final classification tree

Appendix G

Predictive Performance Tables

This appendix presents the objective values from the predictive performance evaluation in Section 11.1.2.

Table G.1: Average objective values for fixed first-stage solutions from the low complexity level

	Scenarios						
	10	**30**	**50**	**70**	**100**	**200**	**400**
Low	119.42	118.39	117.21	117.55	116.79	114.17	116.88
Medium	118.83	117.34	117.78	116.30	119.88	119.14	118.53
High	119.88	119.14	120.37	118.88	117.46	117.53	118.84

Table G.2: Average objective values for fixed first-stage solutions from medium complexity

	Scenarios			
	10	**30**	**50**	**70**
Low	118.88	117.46	113.28	118.84
Medium	116.65	115.83	114.58	115.26
High	117.98	117.34	116.41	116.30

Integrating Machine Learning Techniques in Real-Time Production Optimization

NTNU

Kunnskap for en bedre verden