NTNU
Norwegian University of
Science and Technology

Modelling and Optimization of Real-Time Petroleum Production

Using robust regression, bootstrapping, moment matching, and two-stage stochastic optimization

Marius Lie Morken
Peter Tomassen Sandberg

Industrial Economics and Technology Management
Submission date: June 2016
Supervisor: Henrik Andersson, IØT
Co-supervisor: Vidar Gunnerud, Solution Seeker

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

Problem Description

The purpose of this work is to develop a way to use real-time petroleum production data to optimize petroleum production. This will include data modelling and the development of stochastic optimization models.

The optimization models will be tested both on real field data, as well as on synthetic data to illustrate general model applicability.

Main content:

1. Description of the problem
2. Literature study of relevant papers
3. Data analysis and model building
4. Development of different solution approaches
5. Implementation of the suggested approaches
6. Perform a computational study
7. Compare and discuss the results

The picture on the cover page is taken from Norwegian Petroleum Directorate (2016).

Abstract

This work is concerned with the upstream sector of the petroleum industry and seeks to develop methods for real-time optimization of petroleum production under uncertainty. The real-time production optimization problem deals with planning over a short time horizon, where the objective is to maximize oil production. The output from the optimization is a well defined plan of how the choke valves of the wells at the petroleum field should be adjusted. In order to formulate an optimization model, models describing the relationship between decision variables and production must be determined. We call these models *well models* and evaluate different ways of constructing them. The well models are then used to generate input for several *optimization models*. In order to capture different aspects in the production optimization process and handle the uncertainty in the input parameters, we apply stochastic programming techniques.

Three well models for modelling the relationship between production system settings and production from the wells are presented. The first model, the *export well model*, is built on the measurements of total exported production, while the second model, the *multiphase meter well model*, applies the measurements from the individual wells. The third model, the *combined well model*, uses all of these measurements, aiming to reduce the uncertainty related to the individual well measurements by including the accumulated export production. Several different linear regression techniques are tested for each model, and bootstrapping is used on the best model to find the probability distributions for the well model parameters. Moment matching is thereafter applied to generate a representative set of discrete scenarios. We handle the non-linear relationship between the settings of the production system and production from the wells by introducing a *mode formulation*. Instead of modelling the relationship for all possible settings of the system, we only allow for settings within certain intervals, or *operating modes*. The allowed modes are determined by the available historical data as a combination of previously seen modes and modes that can be estimated based on the available data.

In a case study of the petroleum field Gjøa in the North Sea, we show that the combined model is preferred for modelling both oil and gas production. The robust linear regression method Huber's T is found to be the best regression technique. Through bootstrapping, we find that all the parameters for all wells have probability distributions close to the normal distribution.

In order to include uncertainty in a proper way, two stochastic models which we name the *penalty model* and the *strategy model* are developed. The assumptions behind each of these models rely on insights from the company Solution Seeker AS on the production optimization process. In the penalty model, a constraint can be violated at a cost of lower oil production. The decision variables suggest how much each choke valve should be adjusted to optimize production. In the strategy model, it is assumed that changes to the production system are implemented slowly, so that it is observed when the constraint is met. The decision variables suggest both how much each choke should be adjusted and a sequence for the adjustments.

For the case study at Gjøa, we show that the stochastic models yield a small extra improvement over the deterministic models. However, a potential for significant improvement of oil production is found by applying either model. The solutions found have very good in-sample and out-of-sample stability, especially when moment matching is applied to create the scenario sets. From four synthetic cases, we find that the size of the benefit of incorporating uncertainty in the models depend on the input parameters. Uncertain oil prediction parameters give no additional benefit, while uncertain gas prediction parameters give a small extra benefit using the stochastic models. The penalty model and the strategy model usually return the same objective value, but both models should be tested as variations can occur.

Sammendrag

Dette arbeidet omhandler oppstrømssektoren av petroleumindustrien og søker å utvikle metoder for sanntidsoptimering av petroleumproduksjon under usikkerhet. Sanntidsoptimeringsproblemet av produksjonen dreier seg om planlegging med en kort tidshorisont, hvor objektivet er å maksimere oljeproduksjon. Resultatet fra optimeringen er en veldefinert plan for hvordan ventiler for ulike brønner på et petroleumsfelt bør justeres. For å kunne formulere en optimeringsmodell må man finne modeller som beskriver forholdet mellom beslutningsvariabler og produksjon. Vi kaller disse modellene *brønnmodeller* og evaluerer forskjellige måter å konstruere dem. Brønnmodellene blir deretter brukt for å generere inndata for flere *optimeringsmodeller*. For å kunne fange opp forskjellige aspekter i prosessen med produksjonsoptimering og for å håndtere usikkerheten i inndataparametrene, anvender vi teknikker fra stokastisk programmering.

Vi presenterer tre brønnmodeller for å modellere forholdet mellom innstillinger av produksjonssystemet og påfølgende produksjon fra brønnene. Den første modellen, *eksport-modellen*, bygger på målingene av total eksportert produksjon, mens den andre modellen, *multifasemeter-modellen*, anvender målinger fra hver enkelt brønn. Den tredje modellen, den *kombinerte brønnmodellen*, bruker alle disse målingene, med mål om å redusere usikkerheten forbundet med hver enkelte brønns målinger ved å inkludere akkumulert eksportproduksjon. Flere ulike regresjonsteknikker testes for hver modell, og bootstrapping blir brukt på den beste modellen for å finne sannsynlighetsfordelinger for brønnmodelparametrene. Moment matching blir deretter anvendt for å generere et representativt sett av diskrete scenarier. Vi håndterer det ulineære forholdet mellom innstillingene i produksjonssystemet og produksjon fra brønnene ved å introdusere en *modus-fomulering*. I stedet for å modellere forholdet for alle mulige systeminnstillinger, så tillates bare innstillinger innenfor visse intervaller, eller *driftsmoduser*. De tillatte modusene blir bestemt av de tilgjengelige historiske dataene som en kombinasjon av tidligere sette moduser og moduser som kan bli estimert basert på disse.

I et casestudie av petroleumsfeltet Gjøa i Nordsjøen, viser vi at den kombinerte brønnmodellen er foretrukket for å modellere både olje- og gassproduksjon. Den robuste lineære regresjonsmetoden Hubers T blir funnet å være den beste regresjonsteknikken. Gjennom bootstrapping finner vi at alle brønnparametre for alle brønner har sannsynlighetsfordelinger som ligger nært en normalfordeling.

For å inkludere usikkerhet på en passende måte utvikles to stokastiskte modeller kalt

straffmodellen og *strategimodellen*. Antagelsene bak hver modell bygger på innsikt om produksjonsoptimeringsprosessen gitt fra selskapet Solution Seeker AS. I straffmodellen kan en restriksjon bli brutt med en kostnad gitt på bakgrunn av lavere oljeproduksjon. Beslutningsvariablene foreslår hvor mye hver brønnventil bør bli justert for å optimere produksjonen. I strategimodellen antas det at forandringer i produksjonssystemer implementeres sakte, slik at det kan observeres når restriksjonen blir nådd. Beslutningsvariablene foreslår både hvor mye hver brønnventil bør justeres og en sekvens for justeringene.

For casestudiet på Gjøa viser vi at de stokastiske modellene gir en liten forbedring i produksjon utover det den deterministike modellen finner. Imidlertid finner vi potensiale for betydelig forbedring av oljeproduksjon med alle modeller. Løsningene som finnes har svært god in-sample og out-of-sample stabilitet, spesielt når moment matching blir anvendt for å generere scenariesett. Fra fire syntetiske tilfeller finner vi at størrelsen på fordelen med å inkorporere usikkerhet i modellene avhenger av inndataparametrene. Usikre oljeproduksjonsparametre gir ingen utbytte, mens usikker gassproduksjon gir en liten fordel av å bruke stokastiske modeller. Straffmodellen og strategimodellen resulterer som oftest i samme objektivverdi, men begge modeller bør testes siden variasjoner kan forekomme.

Preface

This Master's Thesis is a part of the specialization in Managerial Economics and Operations Research at the Department of Industrial Economics and Technology Management at the Norwegian University of Science and Technology (NTNU).

The motivation behind this work is the authors' previous experience with real-time petroleum production optimization from a specialization project performed during the fall of 2015 (Morken et al., 2015). During the specialization project we gained a lot of useful knowledge and we found interesting results that we were eager to improve using new ideas and more advanced methods. We do note that parts of this thesis is based on or taken from this specialization project.

We have truly enjoyed working on this thesis. We have acquired in-depth knowledge of mathematical optimization and gained an understanding of general concepts within the area that are clearly useful in later work. We are motivated to continue expanding our knowledge within this research area and applying what we have learned on related real-world problems.

We would like to thank our supervisor, Professor Henrik Andersson, for helpful advice and guidance through the work of this thesis. We are also grateful for informative and inspiring discussions with our co-supervisor Dr. Vidar Gunnerud from Solution Seeker. The insights and explanations on petroleum optimization and parameter modelling given by PhD candidate Kristian Hanssen, Dr. Øyvind Nistad Stamnes, and Dr. Stein Krogstad from SINTEF are also highly appreciated. Finally, we would like to thank Sarah Ustvedt for the help and cooperation during our previous specialization project concerning the same issues as discussed in this thesis.

Trondheim, 6th of June, 2016.

Marius Lie Morken and Peter Thomassen Sandberg

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

Crude oil, or petroleum, is the unprocessed mix of hydrocarbons formed by dead organisms under high pressure and temperature. Many important products are produced from crude oil, with the largest shares being fuel oil and gasoline. Asphalt and chemicals used to make plastics are other examples. Using petroleum as an energy source leads to emissions damaging the environment, and continuous efforts are therefore made to develop renewable energy alternatives. Although the technology for producing from sources such as solar energy and wave energy have made large progress in recent years, the world's energy consumption is still highly dependent on petroleum production.

Since crude oil is found in reservoirs underneath the Earth's surface, often below the seabed, a complex industry is concerned with the exploration, extraction, refinement, and transportation of petroleum products based on crude oil. The industry is often divided into three segments: upstream, midstream and downstream. While upstream companies retrieve the petroleum from its reservoirs, the midstream companies are responsible for transportation, storage, and marketing. Often this involves delivering crude oil from platforms and production facilities to processing and refinement. Downstream companies refine the petroleum to produce the finished products and distribute these to the markets.

All players in the petroleum industry need to make decisions regarding how to operate while seeking to maximize their profits. This leads to planning problems spanning all horizons, from hours to several years, involving decisions such as investments in new fields and how to operate to increase oil production. Based on the time-span relevant for a planning problem, it is common to categorize it as strategic, tactical, or operational.

A strategic problem is concerned with long-term decisions with a horizon spanning from a year to several years, while tactical planning decisions involve medium-term decisions with a horizon of a few months to a year. Operational planning has the objective of maximising daily profits with a short-term perspective. In the petroleum industry, continuous optimization with an even shorter time horizon is sometimes referred to as Real-Time Production Optimization (RTPO) (Bieker et al., 2006). This is planning on the operational level where the goal is to increase oil production by making small adjustments to the operating system. Often, decisions have to be made under uncertainty in parameters, such as measurements on the platform and market conditions. To plan and produce optimally, these uncertainties need to be accounted for in the planning problem.

In this thesis, the upstream operational planning problem is considered. The choice of topic is motivated by the increased focus in the petroleum industry of producing more oil, while making minimal new investments. This is a consequence of the low margins faced by the industry after the oil price has dropped drastically in recent years. The problem of how a petroleum field should be operated on a day-to-day basis, namely the RTPO problem, has therefore become more important. This optimization problem seeks to increase daily oil production by performing small adjustments to the production system. This is an iterative process, as information from new measurements can continuously be included and the optimization repeated. There is a time constraint on solution time, which preferably should be less than 10 minutes, as the solutions are to be used as decision support and continuously be implemented on the field.

Although the research on optimizing petroleum production is extensive, little emphasis is laid on how the solutions should be implemented on an actual field. For instance, the decision variables in many models involve the pressure in different parts of the production system, which is not a variable that can be directly controlled on a field. In addition, the parameters included in the models are commonly based on highly inaccurate simulators, reducing the models' usefulness. Finally, the uncertainty in the parameters is seldom accounted for when solving the optimization problem, leading to suboptimal solutions. Although there are many uncertainties in the data available for petroleum fields, the production optimization problem has historically been treated in a deterministic manner. There have been attempts to include uncertainty when modelling, often resulting in a model where the actual problem at hand has been adapted to fit a certain method. Ideally, it should be the other way around. Uncertainty should be incorporated in a way that best fits the nature of the problem.

To handle the challenges mentioned above, this thesis tests alternative methods for estimating the production system parameters relying mainly on historical data. Two stochastic optimization models are developed, with different assumptions on how the production system is operated. Although the methods in this thesis are relevant for a wide range of fields and industries, the main driver for the project is the planning problem at the Gjøa field in the North Sea. Therefore, the background material provided is focused on the offshore petroleum sector. The authors have collaborated with a start-up company named Solution Seeker to gain access to production data from ENGIE, which is the field operator. This data is used to conduct a case study on the methods proposed. Four synthetic cases are also suggested to illustrate different aspects of the proposed models.

On the data input side, three different data models are presented and compared using least squares and robust regression techniques. These are calibrated on a training set of the production data available and evaluated on a test set using several metrics. The best model is applied to generate constant and slope parameters for the optimization models. A bootstrap procedure is applied to get an estimate of the underlying probability distribution and moments for the chosen constants and slopes. A scenario generation technique based on moment matching is used to find a small set of discrete scenarios that represent the uncertainty well. Two recourse formulations are used to handle the uncertain parameters and to generate solutions that can easily be implemented at an actual field. A thorough technical study is given where stability and stress tests are performed to test model robustness. The recourse models are additionally compared with deterministic models where uncertain parameters are replaced by expected values, to give an estimate of the value of incorporating uncertainty in the decisions taken.

In the next chapter of this thesis, background material for the petroleum industry and planning process is presented. Further, Chapter 3 contains a literature study of research related to the problem studied, while Chapter 4 presents relevant theory used to solve the problem discussed. Chapter 5 gives a detailed problem description. The well parameter models and optimization models are described in Chapters 6 and 7, while the case study is presented in Chapter 8. Chapter 9 presents the implementation methods and well model analysis, and also the problem instances used to perform the computational study in Chapter 10. The results from the computational study are discussed in Chapter 11 and a conclusion is given in Chapter 12. Finally, ideas for further work are included in Chapter 13, before a list of references and appendices are provided.

Chapter 2

Background

In Norway, the first licenses to produce petroleum offshore were given in April 1965. The first major production started at the Ekofisk field in 1971. From the 0.35 million standard cubic meters of oil equivalents (Sm^3 o.e.) produced off the Norwegian coast in 1971, the production in 2015 ended at 228 million Sm^3 o.e.. A modest start up with only a few operating companies has developed to more than 50 Norwegian and international companies being active on the Norwegian continental shelf. This ensures hard competition and a need for sensible resource management. (Norwegian Petroleum Directorate, 2015a,b)

When considering revenue, the petroleum industry had created a total value of over NOK 12000 billions on the Norwegian Continental Shelf by the end of 2012. Even though production has been going on for over 40 years, it is expected that only 47% of the total available resources has been extracted (The Ministry of Petroleum and Energy, 2015). Conserning resources available, these estimates should provide a bright future for the industry in general. However, the development of shale oil fields in the US and increased production in the middle east has led to an oversupply of oil in the market. As a result, prices have dropped significantly leading to an economic downturn for the industry. In just over a year, the price per barrel of oil has more than halved, while the cost of producing a barrel is almost unchanged. As the industry historically has enjoyed high profits, many companies have focused on exploring new fields in order to increase their production. However, the reduced profitability has reduced budgets and plans for further investments in exploration and development activities. Instead, companies now focus on measures that can increase production efficiency at existing fields, as improved planning methods need a fraction of the investments necessary for exploring and drilling a new field or well.

In this chapter, an overview of relevant principles in the petroleum industry is presented. Chapter 2.1 provides an introduction to planning in the petroleum sector supply chain, while Chapter 2.2 explains the offshore production system. Finally, the production planning process is described in Chapter 2.3.

2.1 Planning in The Petroleum Sector Supply Chain

A general supply chain in the petroleum industry consist of crude oil production, transportation to a processing plant, crude oil refinement, transportation to sellers, and end users. A simplified illustration is given in Figure 2.1.1. The crude oil production is done in many different ways. At the Norwegian continental shelf three common ways of producing offshore oil are using a platform, a floating production system or a sub-sea system. Further details on production are given in the next section of this chapter. Transportation to the oil refineries is either done by oil tankers or through pipelines. This varies from location to location, depending on distance the from offshore facilities to the on-shore refineries and investment costs for building pipelines. At the refineries the crude oil is used to make products such as gasoline, diesel fuel, kerosene, butane, and many more. From the refineries transportation is done either to end customers or to an intermediary selling to end users.

To be able to deliver according to market and customer demand, extensive planning is needed. As mentioned in the introductory chapter, the planning problems can be divided into strategic, tactical and operational problems. A fourth phase, real-time planning can also be added. Figure 2.1.2 summarizes these planning phases and their time horizons. Decisions made at a longer time horizon give boundaries for the decisions possible to make in all shorter time horizons. A strategic decision would in other words impose limits on decision-making in tactical, operational and real-time problems.

Decisions made on a strategic planning horizon seek to maximize the total profits earned from a field development and can involve project portfolio selection, annual delivery planning, and vessel purchasing (Shakhsi-Niaei et al., 2013). Ensuring that the total drainage from the reservoir is at its optimum is also important in this time frame. Tactical planning often consider how oil and gas should be routed in different flows,

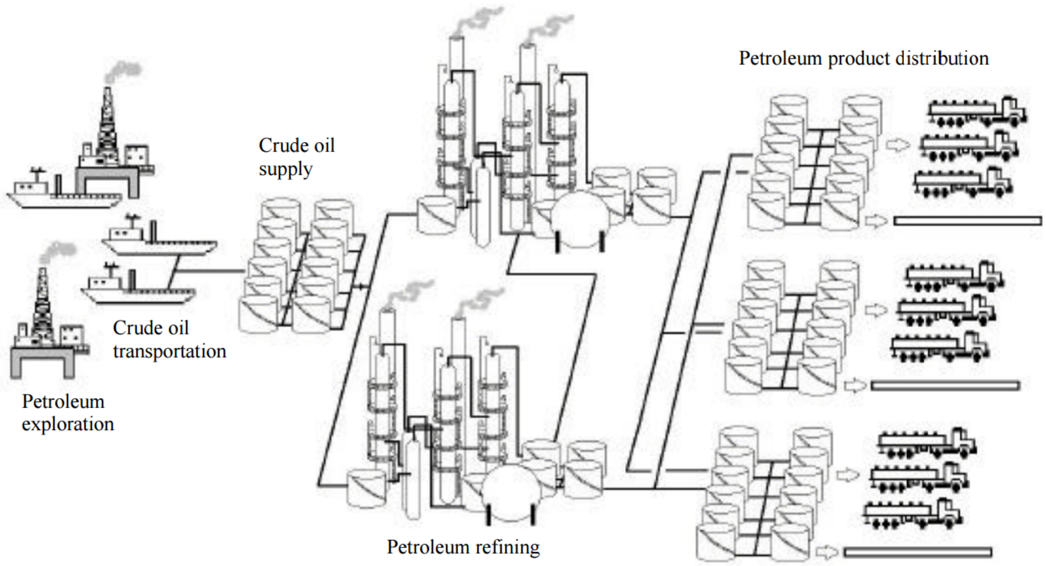


Figure 2.1.1: General petroleum supply chain (Neiro and Pinto, 2003)

or transportation of the recovered products (Ulstein et al., 2007). Capacity requirements for transportation and refinement are elements that need to be considered for this horizon. Operational planning narrows the focus even further, focusing on issues such as minimizing short-term costs or maximizing short-term profit by adjusting how the oil fields are operated (Wang, 2003). Real-time optimization uses production measurements applied real-time to continuously improve operations (Bieker et al., 2006).

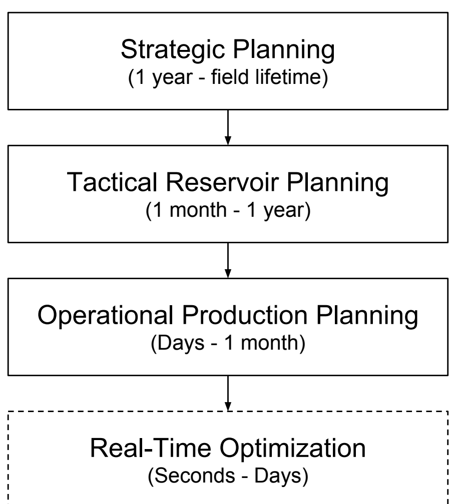


Figure 2.1.2: Strategic, tactical, operational and real-time planning

Common to all planning decisions are the uncertainties associated to the input parameters of the planning problem. When addressing the uncertainties related to planning, it is natural to separate between *internal uncertainties*, which to some extent can be reduced by the modeller, and *external uncertainties* which originate outside the model scope (de Weck et al., 2007). In the petroleum industry, internal uncertainties are typically concerned with the measurements of geological conditions and the production process. By investing in better equipment, these uncertainties can often be reduced, but likely not eliminated. Important external uncertainties faced by petroleum companies are the development and demand in relevant markets, which the companies have little control over.

2.2 Production System

The production system consists of a whole range of different modules. Figure 2.2.1 shows a simplified illustration of some important parts of an offshore installation. Topside, this illustration has a platform, a possible main production facility. This is where the production engineers live and work. Additionally, some oil processing is often performed here. A manifold, either topside or sub-sea, diverts the oil and gas from several wells either to the production platform, to a tank for measurements or directly to a production line, depending on the system configuration (Schlumberger Limited, 2015). The separator in turn, has the job of separating the fluids being extracted from the well. The most common separators are either two-phase separators separating gas and liquid, or three-phase separators separating gas, oil, and water.

The production wells are located at the sea-bed. A production well is a drilled borehole going from the sea-bed and into a reservoir located somewhere beneath. A well is often either a satellite well or a sub-sea well. A satellite well is a single well often separate from a permanent drilling structure, having a direct line to a topside manifold. A sub-sea well on the other hand has the wellhead and production control equipment located on the sea-bed. The wellhead consists of a system of valves, spools and different adapters to control the pressure of the fluids being extracted. An important part of the wellhead is the choke, which is the actual device a production engineer can adjust to control the fluid flow-rate and pressure out of the well. When a choke is closed, there is no flow through the pipeline, while opening a choke will in general lead to increased production from the well. The openings of the chokes at a certain point in time is often referred

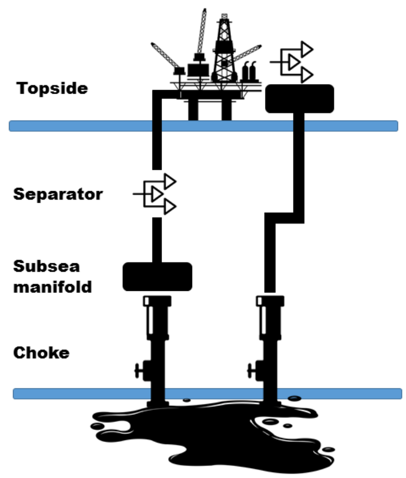


Figure 2.2.1: Petroleum production system

to as the *operating point* at that time.

There are two main types of measurements recording the amount of oil, gas, and water produced on the platform. The *export measurement* records the total production on the platform, while a *multiphase meter (MPM) measurement*, is recorded for each individual well. As multiphase meters are expensive to install, not all platforms or fields have invested in them, and therefore only rely on the export measurement. Both these measurements are uncertain. The multiphase meter is a device measuring the flow of oil, gas, and water through the well at any point in time. The complex flow of oil, gas, water, and other substances are one of the main reasons why the MPMs often have uncertain measurements (The Norwegian Society for Oil and Gas Measurement, 2005). Modern commercial three-phase MPMs can achieve relative uncertainty measures of about ±5% for liquid and gas flow-rates (Thorn et al., 2013). One of the ways used to find measurement errors, is to compare the sum of measurements of the individual subsea MPMs to the more accurate total flow of wells measured topside. This could reveal deviations, but will not tell the operators which MPMs that are not accurate (Folgerø et al., 2013). More extensive testing, for example using a test separator, is needed to reveal how the MPMs should be calibrated. There are also some uncertainty connected to the export measurement. Typically these measurements are more accurate, as the oil and gas are measured after the separators as a separate single-phase. A typical measure of inaccuracy is ±0.25% for oil and ±1% for gas (Thorn et al., 2013).

A production system can contain one or more reservoirs. A reservoir is a subsurface body of rock having sufficient porosity and permeability to store and transmit fluids

(Schlumberger Limited, 2015). A reservoir has dynamic properties meaning that pressure, gas-to-oil ratio (GOR) or other measures can change over time. Changes are minimal in an operational planning horizon, while for tactical and strategical planning, the reservoir dynamics can be substantial. Pressure in a reservoir decreases over time as fluids are extracted. At some point the pressure in the reservoir might not be high enough to send fluids through the wells toward the surface. Then, injection of gas into the well might be necessary to decrease the hydrostatic pressure in the well, so that the reservoir pressure again is high enough for the fluids to flow through the wellhead. This technique is known as gas-lift, and can be set directly by the production engineer. Often, the engineer chooses the gas-lift rate and then this rate is translated to a gas-lift choke opening by the system.

When operating a production facility, the goal is normally to maximize oil production. Increasing the opening of the chokes will generally lead to higher production. The reservoirs consist of a mix of oil, different gases, water, and other articles such as sand, meaning that an increase in choke opening will give an increased production of all these substances. A production facility often has a limit on how much of these substances (especially gas and water) that can be produced. In some cases these are hard limits which the production cannot exceed, but most often these are soft limits which the production on *average* should lie below. This can for instance be if a limit is imposed by the reservoir engineer, who is concerned with optimizing production and reservoir depletion in a long term perspective. The operator will then ensure that average production lies below this limit, but can allow for temporary violations without damaging the system.

2.3 Production Planning Process

In order to operate a platform at optimal conditions, many decisions must be taken. Figure 2.3.1 provides an illustration of a general information flow for some of these decisions. At the offshore production facility, continuous measurements are made in order to control and monitor production. These production measurements are often sent onshore, possibly through an optimization software, so that a team of skilled production engineers can analyze the data and decide on which actions to take. This operational planning must adhere to the restrictions set by the reservoir engineer, who might have conflicting goals when planning for total output of the reservoir in a tactical or strategical time horizon.

When onshore engineers have decided which changes that should be implemented to improve production, these changes are discussed with offshore well operators. This could be done in many ways, but a scheduled video conference is often used. The offshore well operators have the responsibility for the actual physical actions that have to be made to change the system. This includes changes in choke and gas-lift settings.

The proposed changes to the system, determined by the production engineers, are generally a set of wells with new choke and/or gas-lift settings. In many cases, the list of wells is unordered, so the well operator can implement the changes in any order. Since it is preferable to not violate any limits of the production system, reductions in choke openings are generally performed before any increases. Although less common, the list can be ordered in a sequence, which the well operator must follow when implementing the changes. These two ways for receiving instructions from the production engineers lead to different patterns for how the well operator works. In this thesis we divide the work process of an operator into two different categories. In the first approach, which we name the *penalty approach*, the operator receives an unordered list and implements all these changes, and later reverts some changes if any limits are observed to be violated. In the second approach, which we name the *strategy approach*, the operator implements a specific sequence of changes, not completing the list if any limit is met along the way. The work process is different for all fields, and our categorization does by no means cover all variations.

The evaluations done by both engineers and well operators are often based on a combination of analysis and previous experience. The experience each decision maker has in each situation may vary, possibly giving inaccurate or suboptimal solutions for production. Additionally, engineers often have to perform repetitive work, for example having to manually add new measurement data into analysis software or manually remove obvious invalid or bad data. This could lead to decisions being made at lower frequency than with optimal automatized software and equipment.

The optimization software stage shown in Figure 2.3.1 can range from simple excel-sheets to advanced programs. If the software is concerned with what we refer to as real-time optimization, live data is used to estimate the parameters needed for optimization, while the optimization models suggest changes submitted to the control system. Based on the changed settings, new data is recorded, which can be included to make more accurate parameter models. Very few fields apply software based on the

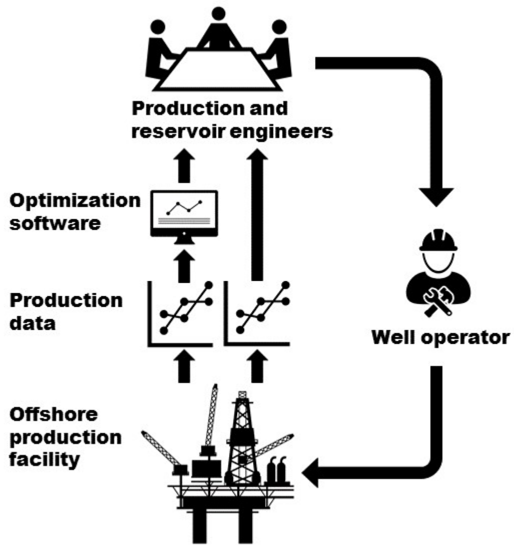


Figure 2.3.1: Information flow for petroleum field operations

methods described above, but rather often rely on simple manual methods. When there is no limit on for example the water processing, an heuristic method often applied by the operator is to decrease the choke opening of the wells with high GOR and increase the chokes of the wells with low GOR. This should lead to higher oil production for each unit of gas produced, but does not consider uncertainty and might therefore not be the optimal solution.

Chapter 3

Literature Study

This chapter gives an overview of literature relevant to optimization in the petroleum industry. The study has focus on how uncertainty is handled and how input parameters to the optimization models are generated. Although all planning horizons are considered, emphasis is given to the operational and real-time levels.

A top down approach is chosen for the discussion of literature. First, Chapter 3.1 provides a brief overview of optimization in the petroleum industry with examples from all parts of the value chain. Chapter 3.2 narrows the focus by presenting research where uncertain parameters are treated. Further, Chapter 3.3 discusses how the parameters relevant for planning in the petroleum industry can be estimated. Finally, Chapter 3.4 draws parallels between the span of research presented and comments how this thesis fits into the literature in general.

3.1 Optimization in the Petroleum Industry

In the petroleum industry, mathematical programming techniques have been applied since the 1940s (Bodington and Baker, 1990). Efforts have been made to optimize all levels of the industry, including exploration, development, production, and transportation. Problems subject to operations research have spanned from strategic planning to process control. A literature review of optimization techniques for petroleum fields by Wang (2003), found that almost all areas of the petroleum industry apply optimization techniques in some way or another. More specific examples are given within gas-lift and production rate allocation, production system and design, and reservoir development and management.

Nygreen et al. (1998) describe a Mixed Integer Linear Program (MILP), dealing with long term planning of Norwegian petroleum and production. Emphasis is given to scheduling of projects. The model has been in practical use for over 15 years, showing a willingness from the industry to invest in such decision support tools. Other authors investigating the project selection problem include Walls (2004), Orman and Duggan (1999) and Brashear et al. (1999).

An early contribution to the literature concerning production optimization on the operational level was made by Attra et al. (1961). They conclude that linear programming techniques effectively can be used to increase daily production from a multi-reservoir field. Improving daily production is also the objective of Dutta-Roy and Kattapuram (1997), who handle the problem of finding the optimal gas-lift rates when back pressure from additional gas in the flow line is accounted for. Wang et al. (2002) solve a similar gas-lift problem and also accounts for flow interactions among wells.

Although the production optimization problem in general is non-linear in nature, linear models are often used by applying some form of approximation. Gunnerud and Foss (2010) develop a MILP model using mostly physical field parameters as input. The model is solved through both Lagrangian and Dantzig-Wolfe decomposition. Non-linearities in the problem are modeled by a piecewise linearization. Gunnerud et al. (2012) uses column generation to solve a somewhat similar problem.

Kosmidis et al. (2005) and Ursin-Holm and Shamlou (2013) are authors that have formulated Mixed Integer Non-Linear Programming problems (MINLP). Kosmidis et al. present a physical model for daily well scheduling, while Shamlou and Ursin-Holm treat upstream and operational planning. The latter is modeled as a simulation based non-convex problem and uses different heuristics to solve the model. Nishikiori et al. (1995) also formulate a non-linear model, and uses a quasi-Newton non-linear optimization technique to find the optimal gas-lift rates.

As described in the introduction, real-time production optimization has been implemented at many oil and gas production sites. Some use physical models while others apply data-driven or black box models. A combination of these have also been attempted, an example being Linden and Busking (2013). The article describes an RTPO system based on a physical model where dynamical changes in the system are included. This requires the current state of the system to be continuously measured and processed. Process variables such as the relationship between pressure, volume, and temperature are included in the modelling. Information of more complex long term behaviour, such as the decline of production due to salt precipitation, is described using data-driven models. In a follow-up article by van der Linden et al. (2015) it is concluded that the real-time optimization technology is ready to be implemented for asset performance evaluation under real-life conditions. The RTPO system can also assist decision makers as immediate decision support. Tjønnås and Schjølberg (2011) also describe a combined physical and data-driven model, but in contrast to van der Linden and Busking, the main focus is on the data-driven model before including some additional physical parameter information. They suggest using simulation data to be able to optimize outside previously seen system settings.

To be successful with implementation of RTPO systems, some important aspects must be taken into account. Bakshi et al. (2015) describe their experiences from implementing such systems in Chevron. Implementation can take as much as 12-18 months and it is therefore critical to keep personnel motivated. The solution must also be easily and highly available and down-time due to for instance bad data quality must be avoided. An easy implementable and understandable system will therefore be of high importance. Some of these same issues are highlighted by Dutra et al. (2010) for Petrobras and Ramdial et al. (2009) for BP Trinidad & Tobago. Some of the challenges found were the need for gradual adaptation for engineers, as well as the need to keep the RTPO process available despite the lack of internet or other technological issues. Campos et al. (2010), describing implementation at Petrobras' Urucu field, Gunnerud and Foss (2010) and Gunnerud et al. (2012) at the Troll field and Dzubur and Langvik (2012) modeling Petrobras' Marlim field are other authors that address real field challenges. If these challenges are overcome, Ofonmbuk et al. (2015) conclude that an implementation of an RTPO system will give huge savings in engineering time and increased frequency of well model updates and re-calibrations. Additionally one can achieve quicker response to failure, increase efficiency in field operation, improve communication and improve well surveillance.

Several different types of commercial software are in use in the petroleum industry today to handle the RTPO problem petroleum companies face. Some examples are ReO and WellFlo from Weatherford (2015a,b), GAP from Petroleum Experts Ltd. (2015), FlowManager from FMC Technologies (2015) and FieldWare Production Universe from Shell Global Solutions (Cramer et al., 2009). FieldWare Production Universe is a good example of a software basing its estimates on a data-driven model. The data is obtained partly from well tests and partly from real-time production data. The software allows for the prediction of overall and single well production through changes in gas-lift rates and choke openings. A more extensive explanation of existing RTPO commercial software can be found in Grimstad (2015).

3.2 Uncertainty in Production Optimization

The majority of research performed on optimization in the petroleum industry has had a deterministic approach. According to Nygreen and Haugen (2010), the increased complexity of including uncertainty in the formulation has historically led to few applications of stochastic models. However, the uncertainties present on all levels of the petroleum value chain need to be accounted for in order to find optimal or close to optimal solutions. During the last decades, increased attention has therefore been given to different ways of incorporating uncertainty in production optimization.

In optimization problems concerning the strategic and tactical decisions of a field development, parameters such as reservoir characteristics and market demand are uncertain. Jørnsten (1992) propose a stochastic model for the scheduling of petroleum fields where the uncertainties lie in the future oil and gas prices. Haugen (1996) also attempts to solve the project scheduling problem, but applies dynamic programming. A broader approach is taken by Al-Othman et al. (2008). In their formulation, the first stage variables represent the production profile of each type of crude oil while the second stage recourse actions include the quantities produced of different products. A broad petroleum supply chain approach is also taken by Fernandes et al. (2015). They formulate a stochastic MILP problem for supply chain design and planning that maximizes the expected net present value of a multi-entity multi-product network. A potential for capturing market variations and obtain better solutions for their decision variables are found. Stochastic formulations with recourse are generally common in the literature on long-term planning in the petroleum industry, further illustrated by for example Goel and Grossmann (2004) and Dempster et al. (2000).

Literature accounting for uncertainties on the operational level is limited. One contribution is made by Huseby and Haavardsson (2009) who try to find the optimal choke settings for wells in reservoirs sharing the same processing facility. They define an optimal production strategy and consider uncertainties in the potential production rates.

However, their optimization problem is analysed deterministically and the strategy found is a vector of set-points for all future time periods.

Further, in a master thesis by Glæserud and Syrdalen (2009), the well flows and capacity limits are uncertain and handled using both chance constraints and recourse models. They present a simple recourse model, where violations of the gas constraint is penalized in the objective function. Authors having a more control theoretic perspective, are Elgsæter et al. (2010) who present a structured approach iteratively updating the set points of the production system in order to increase output. However, they do not incorporate uncertainty in the optimization problem, but instead use it to evaluate the quality of the solution.

Similar to Elgsæther et al., Bieker et al. (2007) propose a model where the decision variables include structured information on how the operator should adjust the wells. They handle uncertainties in the liquid oil ratio by a stochastic formulation where a priority list of wells is the solution. These ideas are applied by Hanssen and Foss (2015) to propose a recourse formulation where the first stage decision is a strategy on how to operate the wells and the recourse action is the executed strategy adapted to the realized well parameters. This model shows promising results when applied to synthetic cases. In addition, it could easily be incorporated in a real-time optimization framework. However, it needs to be adapted to more realistic conditions to be properly evaluated.

3.3 Modeling Well Parameters

In order to optimize petroleum production, models mapping decision variables to production rates are necessary. These models can be based on physical relationships, or rely on experimental data or historical measurements. Trangenstein and Bell (1989) describe the black-oil model, a set of partial differential equations modelling the flow of fluids in a reservoir where water is assumed not to exchange mass with the other phases. Coats (1980) present a compositional model, similar to the black-oil model, but handling each hydrocarbon component separately.

Although reservoir management is not mainly concerned with optimizing day-to-day production, reservoir models have often been used directly or indirectly in literature on production optimization. An example is Wang (2003) who investigated simulation methods where the reservoir and network models are coupled together, creating a full field simulator. He later applies this simulator to find the optimal production and gas-lift rates. Further, Fang and Lo (1996) use a black-oil simulator to formulate a production optimization scheme to maximize oil production under multiple facility constraints.

A moving-horizon based parametric model based on the decomposition of a physical reservoir model is applied by Awasthi et al. (2007). It is assumed that the model proposed is robust enough to be used for extrapolation outside the previously recorded historical data. This, in contrast to using modeling techniques based on regression or interpolation, gave the opportunity to predict phenomena unseen in the past. The assumption that extrapolation outside previously seen historic data is possible could however give inaccurate results, as it is hard to confirm if the obtained results are possible in a real world setting. Other authors suggesting a physical model are Gunnerud et al. (2013) and Shamlou et al. (2012) using a combination of modeling, simulation and optimization to try to avoid suboptimal solutions.

Instead of relying on the physical equations describing a system, experimental data can be used to create parameter models. The science of mapping input controls to output measurements by mathematical modelling is called system identification and is applied across a broad range of industries. A structured and extensive analysis of the field is given by Walter and Pronzato (1997). They thoroughly present the most important methods used in system identification and discuss both recursive and non-recursive approaches to developing linear and non-linear models. Other books covering a broad range of relevant theory are Keesman (2011) and Ljung (1998).

A common way to model production from wells is using the Inflow Performance Relationship (IPR), or the relationship between the flowrate and the flow pressure of the well. Camponogara et al. (2010) develop a method for identifying well performance curves from downhole pressure measurements. They use several equations to fit experimental data and estimate the gas, water, and oil rates. Examples of other authors who have applied IPR models are Fetkovich (1973) and Richardson and Shaw (1982). The disadvantage of these models is that they do not provide a mapping from choke opening directly to flowrate. When applying them for optimization purposes, the downhole pressure will therefore be a decision variable, which is not the case for an actual field.

Limited literature exists on the topic of fitting choke parameters to petroleum production data, and the majority of research performed involves models based on physical relationships. One contribution is Elgsæter et al. (2008) who uses ideas from system identification to develop local linear and non-linear models. In order to quantify the uncertainty in these models, bootstrapping is applied. This is a technique that can be used to estimate statistical properties of estimated parameters (Efron and Tibshirani, 1994).

System identification also encompasses the design of experiments that will generate the measurements needed to construct useful models. In terms of modeling production parameters, this entails adjusting the decision variables to gain more information and reduce model uncertainty, referred to as *planned excitation* by Elgsæter et al. (2008). In a case study, they analyse how low information content in production data affects uncertainty measures when fitting a model.

Another method from system identification, artificial neural networks, have also been applied to build parameter models. Neural networks uses observed data to build a non-linear statistical learning model. Saputelli et al. (2002) presents a review of the use of artificial neural networks in the petroleum industry. Another example is Stoisits et al. (1999), who experiment with replacing simulation methods with these networks and further apply it in production optimization. They implement this algorithm for a field in Alaska and are able to predict production rates with good accuracy.

3.4 Comments on Literature Study

This literature study was introduced with a broad overview of optimization in the petroleum industry. Research involving strategic, tactical, and operational decision making was presented, giving additional focus to RTPO implementations. The methods presented in this thesis are intended to fit into an RTPO framework and previous results in this area are therefore important. Common for the implementations considered, are the short solution time needed, as they seek to continuously improve production. Non-linear models require advanced solution techniques, often leading to long solution times. Since we seek to optimize production on a day-to-day basis, it is therefore desirable to keep the model linear.

The study continues with a thorough discussion of production optimization under uncertainty, including some examples with longer time horizons. It is noted that uncertainty is seldom included in literature on the operational level, especially in a framework that could be implemented at an actual field. Finally, a review of how well parameters are modelled is provided. The majority of work discussed use the physical relationships to estimate these parameters, often leading to complex and inaccurate models. Instead, this thesis seeks to build a simple mathematical model based on historical production data. A challenge noted in the literature is that limited excitation may lead to highly inaccurate models. When developing methods for modelling well parameters in this thesis, this is an important point to consider.

Figure 3.4.1 illustrates the placement of this work with respect to planning horizon and the use of data driven parameter models, among some of the articles presented. Compared to the majority of the literature reviewed, the problem handled in this thesis has a very short time horizon and relies heavily on data driven parameters.

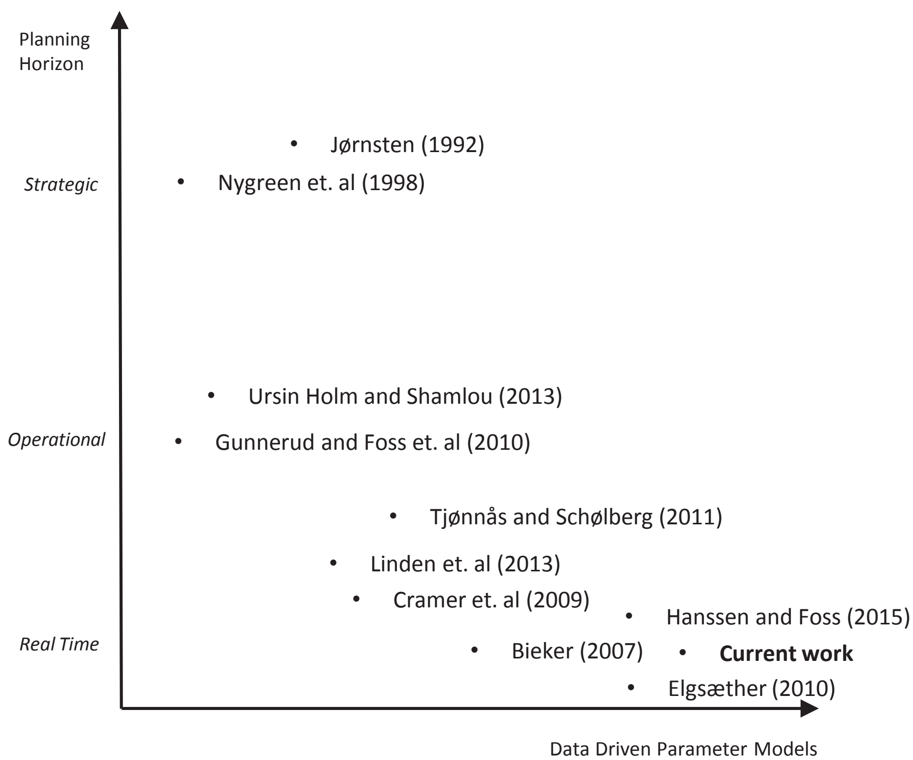


Figure 3.4.1: Placement of current work in selected literature

Chapter 4

Theory

In this chapter, theory relevant for understanding the models and calculations in the following chapters of this thesis is given. In Chapter 4.1 an overview of linear regression methods are provided. Basic familiarity to statistics and regression is assumed. The concept of bootstrapping is described in Chapter 4.2, before scenario generation and moment matching is explained in Chapter 4.3. It is further assumed that the reader has limited knowledge of optimization under uncertainty, and an introduction to the two-stage recourse model is therefore given in Chapter 4.4. Methods for evaluating the solution found from a stochastic model are presented in Chapter 4.5.

4.1 Linear Regression

The concept of linear regression is a mathematical technique that attempts to describe the relationship between a dependent variable and one or more explanatory variables. The technique could be used to make predictions about the future based on historical data or to draw inferences about a larger population or data set (Marill, 2004). Most of the subsequent statistical theory is explained more extensively in Walpole et al. (2012).

Multiple linear regression attempts to model the relationship between one response variable and several explanatory variables by fitting a linear equation to observed data. If Y is the vector of observations of the response variable, X is the matrix with only ones in the first column and the observed values of the explanatory variables in the remaining columns, β, is a vector of regression parameters, and ε is the vector of error terms, we can write the multiple linear regression equation:

$$Y = X\beta + \varepsilon$$

where

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix} \qquad X = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1p} \\ 1 & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_p \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

The value of the error terms in the ε vector are assumed to be normally distributed with a mean of zero and a certain variance independent of the explanatory variable. The assumptions about the underlying data necessary for applying linear regression, can be summarized as follows (Marill, 2004):

1. There is some linear relationship between the predictor and outcome variable.
2. The variation around the regression line is constant.
3. The variation of the data around the regression line follows a normal distribution at all values of the predictor variable.
4. The deviation of each data point from the regression line is independent of the deviation of the other data points.

4.1.1 Ordinary Least Squares

There are several ways to fit an estimated regression line to the data at hand. One of the more common ways are through Ordinary Least Squares (OLS). To find the line the vector β, OLS minimizes the sum of the squares of the residuals (SSE) of every data point, where a residual is given as

$$e_i = y_i - \hat{y}_i, \qquad i = 1, 2, ..., n,$$

that is the difference between the measured data value y_i and a fitted model $\hat{y}_i$ for that ith data point.

The method of least squares is therefore to minimize

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2.$$

OLS assign equal weight to all data points in a data set. This equal weight function is illustrated in Figure 4.1.1.

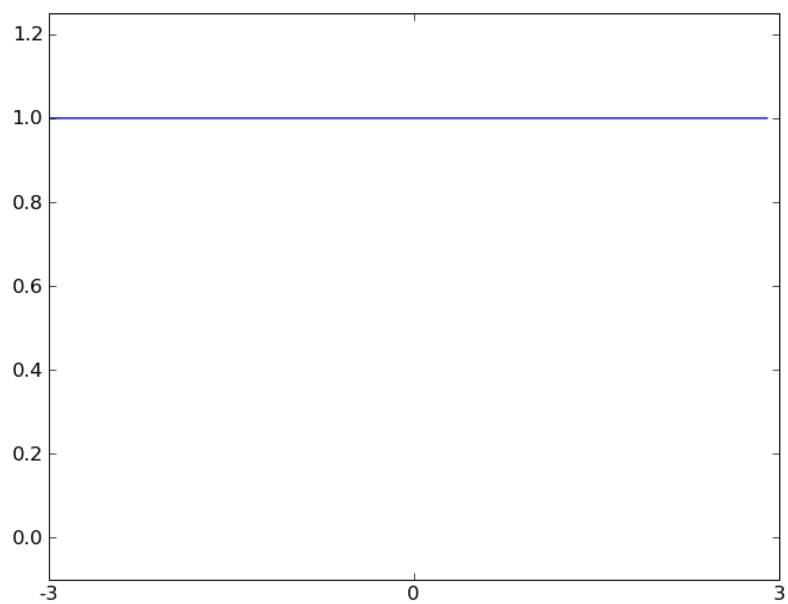


Figure 4.1.1: OLS weighting function (Perktold et al., 2013)

4.1.2 Robust Regression

Robust regression is a collection of regression methods designed to be less sensitive to outliers and give a better estimation of the regression line if the distribution of errors is assymetric or not following a normal distribution as assumed in OLS above. M-estimation is the most common method of robust regression (Alma, 2011). The method was first introduced by Huber (1964). Instead of minimizing the sum of squared errors as described above, a function ρ of the errors is minimized. Huber proposed a minimization of

$$\sum_{i=1}^{n} \rho(x_i, \theta),$$

where ρ is each residual's contribution to the objective function and the solutions

$$\hat{\theta} = arg \min_{\theta} \bigg(\sum_{i=1}^{n} \rho(x_i, \theta) \bigg)$$

are the M-estimators. This problem could be solved directly, but is often simpler if differentiating with respect to θ and solving for the root of the derivative. The function ρ or its derivative ψ can be given properties to make the estimator more desirable and accurate for each estimating problem. The ρ or ψ can for instance be chosen based on the preference of how much weight to assign to outliers. Using a monotone ψ function, large outliers would be weighted less than in OLS. Using a redescending ψ function, weight assigned to an outlier is increased until a specific distance from the estimated regression line, and from there decreased towards zero as the outlying distance gets larger. (Alma, 2011)

Andrews' Wave (Andrews, 1974) and Tukey's biweight (Mosteller and Tukey, 1977) are two ψ-functions using a wave as estimator. In this way a decreasing amount of weight is given to data points with a steeper slope the further away from the estimated line the points are located. Hampel's 17A (Andrews et al., 1972) and Huber's T (Huber, 1981) give equal weight to data points within a specified distance from the estimated regression line, before a decreasing weight is assigned to points beyond that distance. Outliers being far away are given a weight of zero. Minor uncertainty inherent in the data will in this way be given equal weight and data points which with a larger probability are outliers, are given less weight. Ramsay's E (Ramsay, 1977) use a slightly decreasing weight, treating small residuals in approximately a least square fashion, before gradually decreasing weight is given for larger residuals.

The ψ-functions for Hampel's 17A and Huber's T are given as examples to give the reader an insight into some of the properties of these functions. Figures describing the weight assigned to data points by these functions are presented in Figure 4.1.2 and

Figure 4.1.3. In the figures, the default tuning values given in the references are used (e.g. $a = 2$, $b = 4$, $c = 8$ for Hampel's 17A and $t = 1.345$ for Huber's T). ψ- and weighting functions for the other mentioned estimators can be found in Perktold et al. (2013).

$$\text{Hampel's 17A:} \quad \psi(z) = \begin{cases} z & |z| \leq a \\ a \cdot sin(z) & a < |z| \leq b \\ \frac{a \cdot sign(z)(c-|z|)}{c-b} & b < |z| \leq c \\ 0 & |z| > c \end{cases}$$

$$\text{Huber's T:} \quad \psi(z) = \begin{cases} z & |z| \leq t \\ t \cdot sign(z) & |z| > t \end{cases}$$

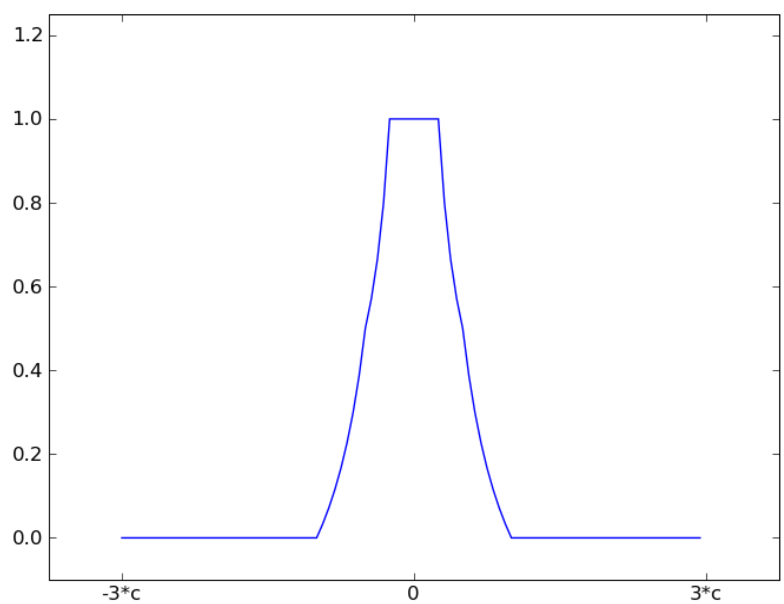


Figure 4.1.2: Hampel's 17A weighting function (Perktold et al., 2013)

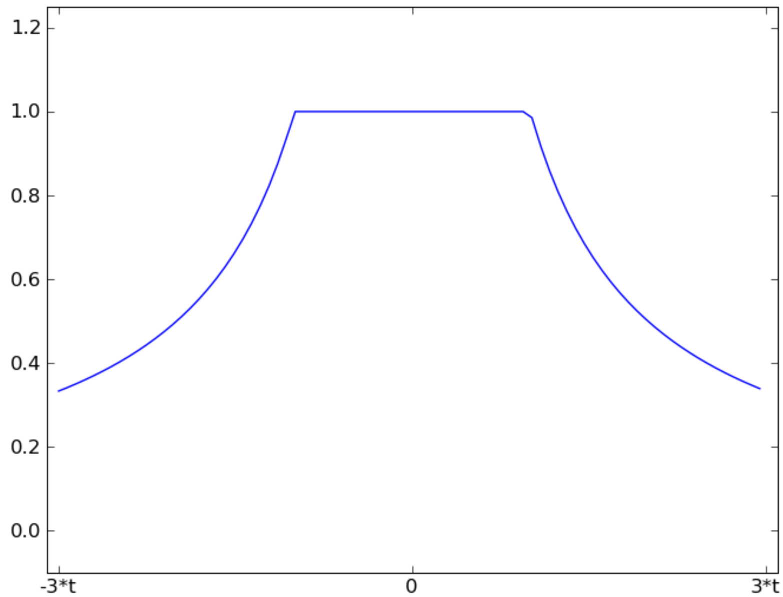


Figure 4.1.3: Huber's T weighting function (Perktold et al., 2013)

4.1.3 Evaluation of Regression Methods

Evaluation of regression models are often performed using *root mean squared error* (RMSE) or *mean absolute deviation* (MAD). MAD expresses the error in the same units as the data itself, making interpretation easier. Outliers are more heavily weighted and punished in RMSE than in MAD, giving two different views on the model error. Which measure to use depends on the assumptions taken and a combined evaluation is often beneficial. RMSE is given as

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

while MAD is given by

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|.$$

Additionally, the coefficient of determination denoted R^2, is often used as a measure for how much of the variance in the dependent variable that is predictable from the independent variable. In other words, it sets a value to how well the observed values are explained by the calculated model. Using the same notation as above, the mathematical equation is given as

$$R^2 = 1 - \frac{\sum_i e_i^2}{\sum_i (y_i - \overline{y})^2}.$$

A R^2 value of 1 would indicate that the model perfectly explains the underlying data.

4.2 Bootstrapping

Within statistics, bootstrapping in a general sense refers to some test or metric that relies on random sampling with replacement. The bootstrap was first published by Efron (1979), and has since been updated and developed by many authors. Bootstrapping is used to estimate the properties of some statistic estimator (e.g. its variance) by replacing the unknown population distribution by the known empirical distribution (Chernick, 2011). To achieve this, the bootstrapping needs to appropriately mimic the underlying data generating process that produced the "true" data set (Thai et al., 2013).

For example, bootstrapping can be used to estimate the distributions of the coefficients in a regression model. This can be achieved by resampling the original data several times to create new data sets, and then create a regression model for each of these new data sets. By doing this, a distribution of values for each regression coefficient is generated.

There exist many different types of bootstrap schemes. Which bootstrap that is appropriate to use depends on your assumptions and trust in your data and model. *Case resampling* and *Residual resampling* are examples of two fundamentally different bootstrap schemes. Most of the following theory is inspired by a bootstrap tutorial by Wehrens et al. (2000), where an even thorougher explanation is given.

In Algorithm 1 and 2 presented in the following sections, bootstrapping is used to estimate the distributions of regression coefficients. The `fit`-function in these algorithms can therefore be any chosen regression method to find a model describing the fit between $\mathbf{X}$ (independent variable) and Y (dependent variable). Further, the `model.predict`-

function uses the fitted model to predict the value of the dependent variable Y, for each row in the matrix of the independent variable $\mathbf{X}$.

Case resampling

In case resampling, whole rows from the original data set are resampled with replacement. Each data point is assumed to be an independent observation (Shalizi, 2013). Bootstrap samples of the same size as the original data set are created. For each bootstrap sample, the statistic of interest is computed. In order to obtain a more precise estimate of the bootstrap distribution of the statistic, the routine of resampling whole data rows to a new data set and computing the statistic is repeated many times.

Formally, a set of independent samples are drawn from $(x_1, y_1), ..., (x_n, y_n)$, yielding $(x_1^*, y_1^*), ..., (x_n^*, y_n^*)$. The pairs $(x_1^*, y_1^*), ..., (x_n^*, y_n^*)$ are then used to compute the simulated values $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ in the same way as $\hat{\beta}_0$ and $\hat{\beta}_1$ were computed from the original data set. Algorithm 1 explains the implementation to achieve this.

Algorithm 1: Case resampling

Data: N rows of data for the dependent variable Y and the independent variable $\mathbf{X}$.
Result: M possible values (i.e. a distribution) for each of the regression coefficients.

```
result ← array of length M;
for i ← 1 to M do
    Y* ← array of length N;
    X* ← array of length N;
    for j ← 1 to N do
        k ← sample a random integer between 1 and N;
        Y*[j] ← Y[k];
        X*[j] ← X[k];
    sample_model ← fit(Y*, X*);
    result[i] ← regression coefficients of sample_model;
return result
```

Residual resampling

In residual resampling a model is estimated from the original data, and residuals are computed for each row of this data set. Then a simulation is performed by resamling these residuals to make new bootstrap data sets. In a more precise manner, suppose $\hat{\beta}_0$

and $\hat{\beta}_1$ are estimates of β_0 and β_1 and that

$$e_1 = y_1 - \hat{\beta}_0 - \hat{\beta}_1 x_1, ..., e_n = y_n - \hat{\beta}_0 - \hat{\beta}_1 x_n$$

are the residuals. A new data set is then made by samling independently from $e_1, ..., e_n$, yielding, $e_1^*, ..., e_n^*$, and by setting $y_j^* = \hat{\beta}_0 - \hat{\beta}_1 x_j + e_j^*$. Then again, least squares (or another method) is used to compute the simulated values $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ from the bootstrap data $(x_1, y_1^*), ..., (x_n, y_n^*)$. Algorithm 2 explains the implementation to achieve this.

Algorithm 2: Residual resampling

Data: N rows of data for the dependent variable Y and the independent variable $\mathbf{X}$.

Result: M possible values (i.e. a distribution) for each of the regression coefficients.

```
result ← array of length M;
model ← fit(Y, X);
Ŷ ← model.predict(X);
ε̂ ← Y − Ŷ;
for i ← 1 to M do
    Y* ← array of length N;
    for j ← 1 to N do
        ε̂_j ← draw one random sample from ε̂;
        Y*[j] ← Ŷ[j] + ε̂_j;
    sample_model ← fit(Y*, X);
    result[i] ← regression coefficients of sample_model;
return result
```

Which Bootstrap to use

If the independent variables x_j are controlled as input by the system, Wehrens et al. (2000) suggest using the bootstrapping residuals method. Different versions of this method exist, depending on the assumptions made. *Wild bootstrap*, as described in Davison and Hinkley (1997), can for instance be applied if the variance of the errors depend on x_j. If on the other hand, x_js are random as well, case resampling should be chosen, as this method mimics the bivariate distribution $(x_1, y_1), ..., (x_n, y_n)$. Resampling cases are in general safer to use as it will also work under the assumptions made for the residual resampling method. Thai et al. (2013) support this, finding that bootstrapping only residuals often greatly underestimated the standard error of parameters and that case resampling on the contrary performed well and provided non-biased parameter estimates and standard errors. The confidence intervals on each estimated statistic is however wider for case resampling, giving a reason to choose residual reampling if model assumptions are trusted (Shalizi, 2013). An implementation of both methods and an evaluation of the probability of the output results could give an indication of which method to choose.

4.3 Scenario Generation

The problem of how to represent the random variables is always present in stochastic programming models. In order to find a solution to a stochastic program, a discretization of the probability distribution of these random variables is necessary. To keep computational time low, it is important that the number of scenarios is as small as possible, while still representing the underlying uncertainty in such a way that the solution found is representative, stable, and accurate.

If scenario generation is performed though a random draw from some known or assumed probability distribution, it is probable that the scenarios will end up representing *all* properties of the distribution, while you might only need some of the properties for the optimization to give good results. The disadvantage of this is that you will need to have an unnecessarily large amount of scenarios to represent the properties correctly, leading to longer computational time. A method known to overcome this is *moment matching*, as described in Høyland et al. (2003), with some updates published in Kaut (2003). In moment matching, an iterative procedure is used to produce a set of scenarios that give a good match with some specified first four marginal moments from the probability distribution. The corresponding correlations between the random variables are also taken into account. The four moments being matched are the mean, variance, skewness, and kurtosis. An even more flexible moment matching method can be found in Mehrotra and Papp (2013), where a column generation approach is used to generate scenarios matching any prescribed set of moments. The following explanation of the procedure matching the first four moments is based on King and Wallace (2012), as well as the two articles by Høyland et al. (2003) and Kaut (2003).

Description of procedure

In the following, the moment matching procedure is described. Generally, the target moments are first found through statistical analysis or by specifying them directly. At the start of the algorithm, the moments of each random variable are normalized. That is, each random variable gets a mean and a variance of 0 and 1 respectively. Skewness and kurtosis are normalized accordingly. The moments of each random variable are then corrected through a transformation at the end of the algorithm. After the normalization of the moments, a discretization of the random variables is performed, usually through random sampling from the standard normal distribution. The result of this random sampling is stored as $\mathbb{X}$. The algorithm then proceeds by iteratively correcting the moments and the correlations of the discrete random variables in $\mathbb{X}$.

To correct the correlations, the variables are transformed through some simple matrix operations. More specifically, at the beginning of the algorithm a Cholesky decomposition is performed on the target correlation matrix R to achieve the matrix L (i.e. $R = LL^T$). In a particular iteration i, the transformation of $\mathbb{X}$ is called $\mathbb{Y}_i$. In each iteration, the correlations of $\mathbb{Y}_i$ are computed and the result is stored as R_i. Then, a Cholesky decomposition is performed on R_i yielding the matrix L_i. To correct the correlations the following formula is then applied:

$$\mathbb{Y}_i^* = (L \times L_i^{-1}) \cdot \mathbb{Y}_i$$

In the general case, both the marginal distributions and the moments are distorted through this transformation. This means that $\mathbb{Y}_i^*$ has the right correlations, but incorrect moments. To correct the moments, the cubic transformation

$$\mathbb{Y}_i^{**} = a + b\mathbb{Y}_i^* + c\mathbb{Y}_i^{*2} + d\mathbb{Y}_i^{*3}$$

is performed. We need to find the parameters a, b, c and d, so that the target moments are matched. That is, we need to find the values of these parameters such that the moments of $\mathbb{Y}_i^{**}$ become as close to the target moments as possible. Four implicit equations in four unknowns are solved to find these parameters. For details on these equations, see Høyland et al. (2003). The result from the cubic transformation ($\mathbb{Y}_i^{**}$) is stored as $\mathbb{Y}_{i+1}$ and a new iteration is started. Since the cubic transformation is non-linear, correlations will be changed. The correlation-correcting transformation described above fixes this in the subsequent iteration. The iterative procedure runs until the correlations of the discrete random variable is sufficiently close to the target correlations.

The moment matching algorithm

Algorithm 3: Moment matching

Data: A set of target moments for each random variable (target_moments), their correlations (target_correlations) and the number of scenarios to be generated for each random variable (N).

Result: A discrete set of scenarios for each random variable with (almost) correct moments and correlations.

$L \leftarrow$ Cholesky decomposition of target_correlations;
normalized_moments $\leftarrow$ `normalize`(target_moments);
do
  $\mathbb{X} \leftarrow N$ scenarios drawn independently from a standard normal distribution for each of the random variables;
  $i \leftarrow 0$;
  $\mathbb{Y}_0 \leftarrow L \cdot \mathbb{X}$;
  $R_0 \leftarrow$ correlations of the random variables in $\mathbb{Y}_0$;
  do
    $L_i \leftarrow$ Cholesky decomposition of R_i;
    $\mathbb{Y}_i^* \leftarrow (L \times L_i^{-1}) \cdot \mathbb{Y}_i$;
    $\mathbb{Y}_i \leftarrow$ cubic transformation of $\mathbb{Y}_i^*$ using normalized_moments as target;
    $R_i \leftarrow$ correlations of the random variables in $\mathbb{Y}_i$;
    increase i by 1;
  while `distance`(R_i, target_correlations) $> \varepsilon$ *and* $i <$ max_iterations;
while $i ==$ max_iterations;
$\mathbb{Z} \leftarrow$ back transform the data of $\mathbb{Y}_i$ such that the moments are no longer normalized, but rather match target_moments;
return $\mathbb{Z}$

Some explanations of the moment matching algorithm are needed. First, normalizations of the moments are done by setting the mean μ and variance σ^2 of each random variable to 0 and 1 respectively. Then the third and forth moments are computed using the formulas in Equation 4.1.

$$\begin{aligned} normalized\ skew &= \frac{skew}{\sigma^3} \\ normalized\ kurtosis &= \frac{kurtosis}{\sigma^4} \end{aligned} \tag{4.1}$$

Second, back transformation of the data to achieve the target moments is done using the formula in Equation 4.2.

$$\mathbb{Z} = \sigma \mathbb{Y} + \mu \tag{4.2}$$

Third, the `distance` between two correlation matrices is computed as the root mean squared distance between the upper triangular parts (R^U) of the correlation matrices, as shown in Equation 4.3. Only the upper triangular parts are used to compute the distance, as the correlation matrices are symmetric across the diagonal. In Equation 4.3, `count` is a function that counts the number of entries in a vector.

$$distance = \sqrt{\frac{(R_1^U - R_2^U)^2}{\texttt{count}(R_1^U - R_2^U)}} \tag{4.3}$$

Fourth, the outer do-while loop is needed, as it may be that the initially drawn set of scenarios X will not converge to a set that matches the given moments and correlations. Thus, if the inner do-while loop is ended by the $i < max\_iterations$ condition, the moment matching procedure will restart from line 4.

The moment matching method described is intricate, and we therefore recommend the interested reader to explore the referenced literature for a more extensive and thorough explanation.

4.4 Two-Stage Recourse Model

A recourse model handles uncertainty in the model parameters, by allowing for decisions to be made at different stages in time. First presented by Dantzig (1955), this model has been widely used, as many real life planning problems involve stages in time where new information becomes available. Incorporating these stages into the optimization model yield a solution that is more robust to the possible outcomes of the stochastic parameters.

Recourse models consist of at least two stages where some decisions must be made in each stage. The term *recourse* comes from the opportunity to adjust a decision made in a previous stage when new information is obtained. A recourse action is therefore a decision made as a response to the outcomes of some of the random variables. The decisions made in the first stage are the same, independent of the outcome of the stochastic parameters. In the second stage, new information is revealed and this is used to make the next decisions.

Based on the nature of the stochastic parameters, a scenario tree can be constructed (Higle, 2005). Each path in the scenario tree corresponds to a certain realization of the random parameters. When dealing with continuous distributions, this means that there exist infinitely many scenarios. A discretization is therefore necessary. An example of a scenario tree for a problem involving parameters following a discrete distribution is given in Figure 4.4.1. Each node in the tree represents a subset of decisions to be made. In the root node, only the distribution of the parameters involved in the problem is known. When moving down the tree, new information is revealed until the leaf nodes are reached and the final decisions must be made.

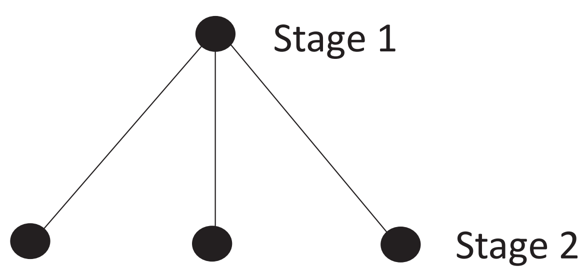


Figure 4.4.1: Scenario tree of two-stage problem

Letting $\tilde{\omega}$ represent a stochastic variable and ω a realization of this variable, the outcome of a parameter for this realization can be denoted with an ω-subscript. The mathematical formulation of a two-stage recourse problem then becomes

$$\min z = c^T x + E[h(x, \tilde{\omega})]$$

$$\begin{aligned} s.t. \quad Ax &\geq b \\ x &\geq 0 \end{aligned}$$

where

$$
\begin{aligned}
h(x,\omega) = \min g_\omega y & & \\
s.t. \quad W_\omega y & \leq & r_\omega - T_\omega x \\
y & \geq & 0
\end{aligned}
$$

Here, the decision x is made in the first stage, while y is the recourse action. $E[h(x,\tilde{\omega})]$ is the expected value function representing the effect of choosing x in the first stage on the possible realizations of the parameters. For each realization ω, the recourse action can compensate for a suboptimal choice of x at a cost g_ω. The optimal x is then robust in terms of the possible outcomes of the random parameters.

In the formulation above, we have assumed no specific structure on the parameters of the second stage problem and all of them can potentially be uncertain. Higle (2005) refers to this situation as having *general recourse*. In some situations, computational advantages can be achieved when the recourse problem has certain properties.

Simple Recourse occurs when the W-matrix in the second stage problem is the identity matrix, leading to the following formulation

$$
\begin{aligned}
h(x,\omega) = \min g_\omega^+ y^+ + g_\omega^- y^- & & \\
s.t. \quad Iy^+ - Iy^- & = & r_\omega - T_\omega x \\
y^+, y^- & \geq & 0
\end{aligned}
$$

For this formulation, the recourse action is uniquely defined by the value of the first stage variables and the outcome of the stochastic parameters. Hence, they represent the compensation necessary when the decision made in the first stage leads to constraint violations in a scenario (Kall and Mayer, 1976). This greatly simplifies the problem, as it is possible to separate the constraints by row. Each second stage variable appear in exactly one constraint and the optimal value is found by solving a simple optimization problem for each constraint.

Fixed Recourse is the case where the constraint matrix W does not depend on the stochastic variable ω. Simple recourse problems always have this property. The recourse subproblem is then given by:

$$\begin{aligned} h(x,\omega) = \min g_\omega y & \\ s.t. \quad Wy &\geq r_w - T_w x \\ y &\geq 0. \end{aligned}$$

When the objective function coefficients also are certain, the dual of this problem is:

$$\begin{aligned} h(x,\omega) = \max \pi^T(r_\omega - T_\omega x) & \\ s.t. \quad \pi^T W &\leq g^T \\ \pi &\geq 0 \end{aligned}$$

This problem has some interesting properties. Since there are no uncertain parameters in the constraints, the feasible region is the same for all outcomes of the random variable. This can be exploited when designing solution methods.

Complete Recourse is a property describing the feasibility of the second stage problem when the first stage variables are set to some value. A problem is said to have complete recourse if for all x, $h(x,\omega)$ is finite, meaning that there exist second stage variables that obey the constraints in the second stage problem. When the first stage variables are constrained so that no such infeasibility can happen, the problem is said to have *relatively complete recourse* (Shapiro and Philpott, 2007).

4.5 Evaluation of Stochastic Solution

It is important to evaluate the stability of the stochastic solution to be sure that the solution found is representative. Two measures, in-sample and out-of-sample stability, are therefore explained in Chapter 4.5.1. Further, since solving a stochastic problem is more advanced both mathematically and computationally, some measures of the value of incorporating uncertainty are explained in Chapter 4.5.2.

4.5.1 In-Sample and Out-of-Sample Stability

To solve a stochastic problem with continuous probability functions, one must discretize the functions. In order to evaluate whether a discretization method is satisfactory, stability tests can be performed. Since scenario generation techniques often rely on samples drawn from the underlying distribution, the resulting scenario tree is different every time the generation method is run. The solution found may then vary with the specific scenario tree. As long as the objective values of these solutions are similar, this is not a problem. However, if the objective values vary, the quality of a solution depends on the scenario tree, and we can no longer trust that the solution found is really optimal.

If a model has *in-sample stability*, the objective value found does not depend heavily on the generated scenario tree. Although the solutions may differ, they yield similar objective values when evaluated on their respective scenario trees. A model with in-sample stability is internally consistent, but can be arbitrarily bad (King and Wallace, 2012). Therefore, it is also necessary to consider how the solutions found perform when evaluated over the actual continuous distributions. By sampling a very large number of scenarios, the samples tend to accurately represent the underlying distributions. Since little computational power is needed to evaluate a solution over many scenarios, we can approximate the true objective value of a found solution. If a model has *out-of-sample stability*, the optimal solutions from applying different scenario trees will yield similar objective values when evaluated on the underlying distribution.

Out-of-sample stability is important to ensure that the solution will actually perform well when applied to a real situation. In-sample stability is important to ensure that the optimization problem is able to spot the good solutions. If there is poor in-sample stability, the optimization problem may interpret good solutions as bad, and the other way around. This could lead to suboptimal solutions, or solutions that do not perform well at all.

4.5.2 EVPI and VSS

Stochastic formulations are generally both harder to formulate and harder to solve than their deterministic counterparts. To be able to compare the different solutions, some measures of the value of incorporating uncertainty can be defined.

The *expected value of perfect information* (EVPI), described in for instance Schlaifer and Raiffa (1961), is defined as the difference between the objective function values of solving the stochastic problem and solving the same problem as if having perfect information. Another way of seeing EVPI is it being the maximum amount a decision maker would be ready to pay in return for complete and accurate information about the future (Birge and Louveaux, 1997). Mathematically, EVPI is the difference between the optimal value of solving the recourse problem (RP) and the wait-and-see (WS) solution (Madansky, 1960). For a minimization problem, the formula for EVPI is therefore

$$EVPI = RP - WS.$$

The RP is formulated as described in Chapter 4.1, with different decisions made at different points in time. An important condition is that a decision made in a stage only can be based on the information known in that stage. This means that a decision variable in a stage must take the same value in all scenarios with equal history. For the WS-solution, this constraint is relaxed, so the optimal decision is made in all possible scenarios.

Another important measure is the *value of the stochastic solution* (VSS). The VSS describes the difference in objective value when planning while incorporating uncertainty, compared to using the decision based on expected values. In other words, VSS is the amount a decision maker would be willing to pay to be able to incorporate uncertainty in the decision support tool used. Mathematically, VSS is the difference between the expected value of using the expected value model (EEV) and the recourse problem (RP). For a minimization problem, the formula for VSS is therefore

$$VSS = EEV - RP.$$

The EEV is found by replacing the uncertain parameters with their expected values in the optimization problem. The expected value of this solution is then found by evaluating it for different outcomes of the parameters. In Figure 4.5.1, the relationship between the different measures is illustrated. If the scenario tree properly represents the underlying distribution, the EVPI and VSS are always non-negative. This is because the solution to the expected value model is a feasible solution for the stochastic model. Similarly, the solution to the stochastic model is a feasible wait-and-see solution.

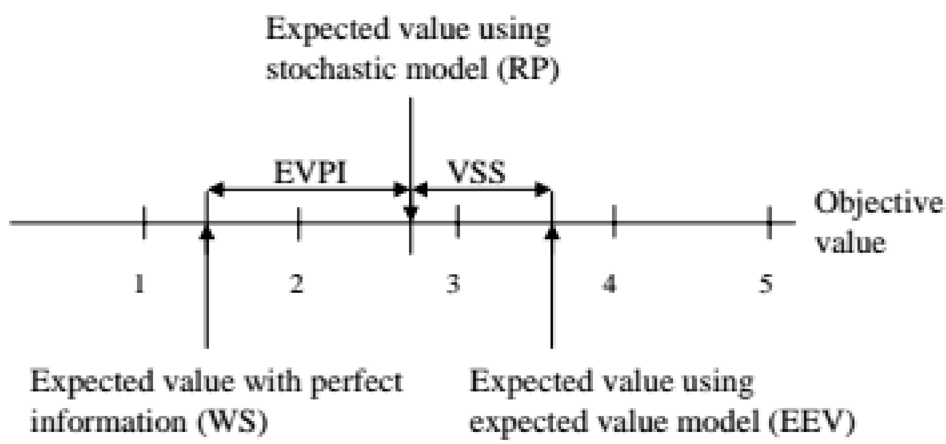


Figure 4.5.1: Relationship between RP, EEV and WS for a minimization problem (Birge, 1997)

Chapter 5

Problem Description

In the petroleum industry, optimization techniques are widely applied across multiple horizons. While the majority of the early applications were concerned with long-term planning, focus has shifted towards short-term continuous production optimization in recent years. This can be attributed to the increased access to data, improved computational power, and development of new optimization methods. The problem considered in this thesis is how to develop real-time production optimization models based on historical production data. Uncertainty in model parameters is to be incorporated and a suggestion for a directly implementable strategy for a petroleum field operator is to be made.

Well Parameters
In order to optimize the export of oil from a petroleum field, input parameters describing the production are required. An example is the mathematical relationship between input controls like choke opening and gas-lift, and the production of the different phases oil, gas, and water. Simulators are commonly used to model these parameters. However, these simulators are inaccurate as they attempt to model the system globally. In addition, since these relations are often non-linear and with discrete variables in nature, the resulting optimization problem is complex. An alternative approach is to use historical data to construct these models. Instead of finding the global relationship between input and output variables, one can restrict the optimization to small changes in the input controls. Then it suffices to model the local relationship around the allowed changes. If the adjustments of the operator are required to be significantly small, the assumption of a linear model might be realistic. This significantly reduces the complexity of the

optimization problem, which is desirable for real-time optimization.

The problem of mapping changes in choke opening to production rates is itself an optimization problem. The raw production data often contains large oscillations and bad measurements leading to inaccurate parameter models. This data therefore needs to be processed before a regression technique is applied. The choice of regression method can significantly affect the quality of the model and the best method might not be equal for all fields. In addition, different fields have different measurement equipment. While some fields have installed multiphase meters, others only rely on total production measurements. The data available therefore sets restrictions on the possible parameter models that can be built.

The Real-Time Production Optimization Problem

The real-time production optimization problem is concerned with maximizing oil production in a short time horizon. Maximization is done either with respect to the present operating conditions, or based on knowledge of increased or reduced production capacity on the field in the future. Increased capacity can for instance be the result of new and improved equipment or the completion of maintenance on the existing production facility. Reduced capacity might be the case when existing equipment is shut down for maintenance or be a strategic decision from the operators based on changed market conditions.

In order to formulate an appropriate optimization model, the work process of the operator is important. The solution of the model should be directly implementable at the petroleum field, with decision variables representing the physical adjustments that must be done to the system. Optimization models considering the physics of the underlying system often include several implicit decision variables, like pressure loss in the pipeline and production rates. These are not variables that can be directly controlled by the operator, but hopefully will take the desired value as a result of the explicit variables. The explicit decision variables are limited to the routing of flows from wells, the choke opening and the gas-lift rate on each well. The values of these variables therefore represent an actual action from the operator at a field.

When the input parameters to the optimization problem provide a direct mapping between the explicit decision variables and production rates, constraints regarding the physics of the system can in many cases be ignored. The main constraints to be fulfilled are then the capacity constraints for water and gas handling. These are often soft constraints that can be breached with a penalty of lower production when adjusting the system from the breach and back to normal operating conditions. Also the operator does not know the exact effect of applying a change to the system. This uncertainty needs to be included to find the optimal way to operate a field. For the real-time optimization problem, the number of wells that can be adjusted is often limited, as the operator does not want to impose too many changes to the system in one iteration. This is another constraint that must be accounted for.

Since the real-time optimization problem involves modelling of well parameters to generate input to the optimization, the objective of this thesis is two-fold. In addition to developing optimization models fit to the work process of the operator, we need to determine appropriate ways to model the well parameters needed as input for these these optimization models. This is illustrated in Figure 5.0.1. The importance of building good well parameter models must not be underestimated. If the built well parameter models do not reflect the real relationships between choke opening setups and production rates, the optimization problem will probably not provide optimal or profitable solutions. In addition, computational time is of great importance, as solutions are to be used in real-time both as decision support and as direct implementations on a field. Effective algorithms both for well parameter modelling, scenario generation and for solving the suggested stochastic problems are therefore needed.

To summarize, the purpose of this work is to find a way to use available real-time historical data, build models based on these, and use those in optimization models to find and suggest a strategy an operator at a petroleum field can follow.

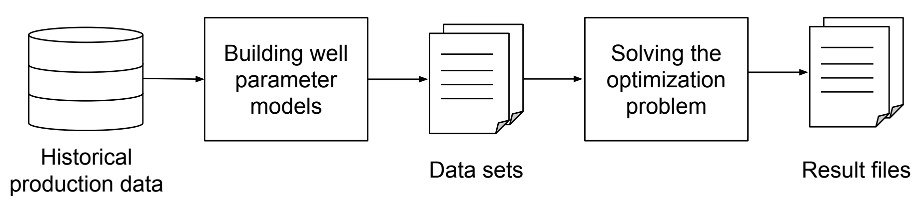


Figure 5.0.1: Real-time optimization process

Chapter 6

Well Models

In this chapter, an approach for modelling well parameters required as input for the optimization models is proposed. Chapter 6.1 explains the reasoning for assuming a linear relationship between the choke opening and the production from a well. Further, Chapter 6.2 proposes a method for modelling well chokes that are turned on and off. Finally, Chapter 6.3 presents three different well models, all applying historical production data.

6.1 Local Linear Relationship

Modelling the relationship between choke openings and total production for all possible settings of the chokes is generally extremely difficult due to uncertainty and lack of data. Therefore, simulators including the physical relationships of the system are commonly used for this purpose. However, the output from these simulators is often very inaccurate due to the dynamic characteristics of the reservoir amongst other things. In addition, since the relationship between choke and production rate is non-linear, the simulators lead to complex models. Since there is a hard solution-time restriction in real-time optimization, non-linear models may be too slow to be applicable. Instead of relying on the global parameter models, we assume that the relationship between chokes and production is linear when only considered locally, and use historical data to estimate the parameters.

To illustrate that a local linear relationship might be realistic, Figure 6.1.1 show the multiphase meter (MPM) measurement for gas production of a well at the Gjøa field. Based on this plot, the relationship between choke opening and gas production appears to be linear within a certain interval for the choke opening. For this linear relationship to hold globally, a choke opening of 10% would need to imply negative gas contribution. As this is physically impossible, we instead require the wells to have a choke opening larger than some constant $\underline{L}$ and smaller than some constant $\overline{L}$, in order for a linear relationship to hold. This interval must be determined for each individual well by analyzing its historical production. When a field has multiphase meter measurements, this is easily done by considering plots as in Figure 6.1.1, however it is more difficult when only the export measurement is available.

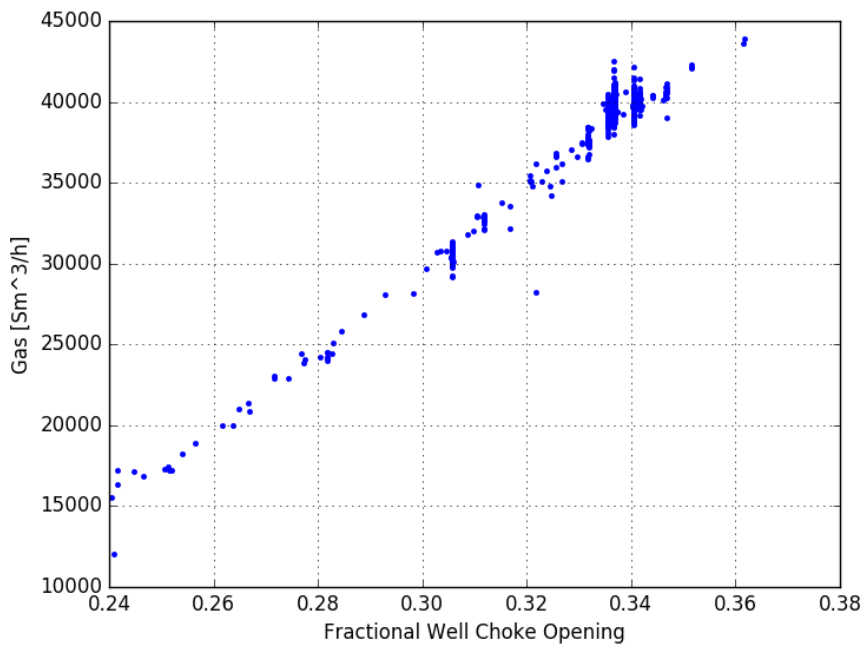


Figure 6.1.1: Gas MPM measurements from Gjøa well

A satellite well can have a separate flowline-system carrying oil and gas from the reservoir directly to a separator. In this case, its flow is not affected by the production from other wells. It is then reasonable to assume that the well has its own independent well model, with a linear relationship between its choke opening and production. Many fields do however have wells connected to a common riser to the separator, and the assumption that they are independent is then only an approximation. All the well models we present assume independent wells, with a linear contribution to production within some interval. All wells then have their own *constant term* and *slope*. This is illustrated in Figure 6.1.2. Although the linear model does not hold outside the defined area, a well can be turned off by subtracting its constant term from the production. Whether or not it is possible to determine each individual well model depends on the data available. While the multiphase meters provide production measurements for each separate well, the export measurement only gives the total production for the field. In the next section of this chapter, we introduce the concept of *operating modes* to handle adjustments where one or several wells are shut off.

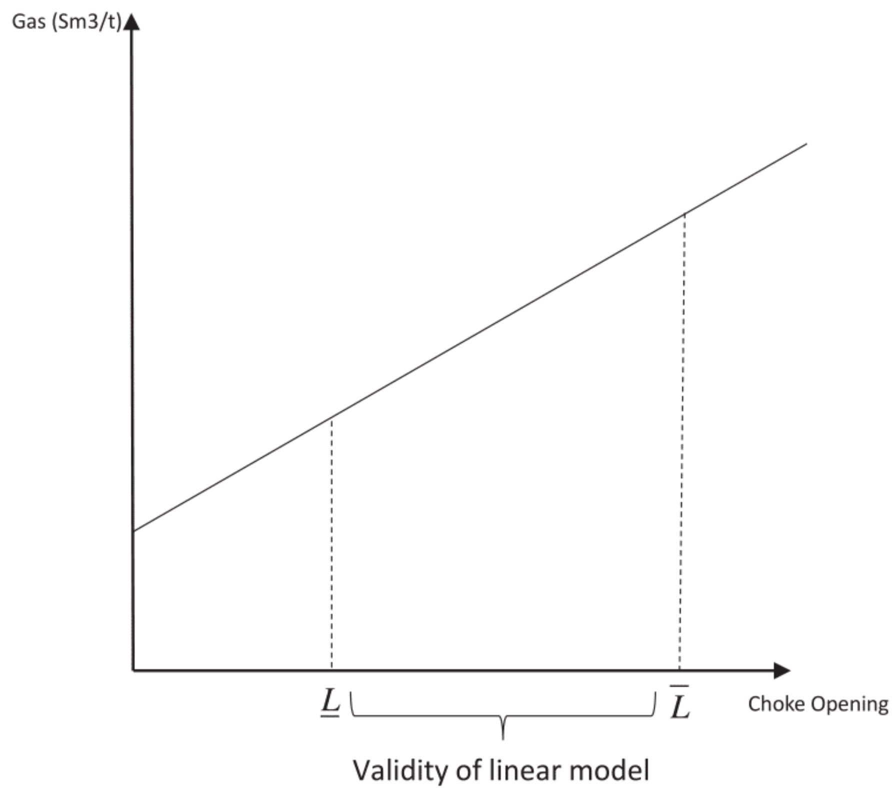


Figure 6.1.2: Local linear relationship for independent well

6.2 Operating Modes

When calculating the total production from a set of wells, the sum of the constants terms is taken in addition to each well's choke opening times its slope. When the choke of a well is closed, its constant term contribution must therefore be removed from the production function. If multiphase meter measurements are available, this can be achieved by introducing a binary variable in the summation, indicating whether a choke is on or off. However, many fields do not have this equipment on all wells. To generalize the optimization models to hold for all well models, we introduce *operating modes*, representing different on/off combinations of wells. One possible mode could for example be choke 1, 2 and 3 open, and choke 4 closed.

To estimate the constant reduction in production when a choke is closed, a unique *mode constant* is estimated for each mode. These constants are then used instead of a single fixed constant across all modes. When the operating mode is changed, the constant in the production function is replaced. Estimating the mode constant is easy when MPMs are available, as it is then the sum of the individual constant terms of all wells with open chokes. When only export data is available, finding the mode constant is more cumbersome, but still possible.

Applied to a case with only two wells, we would get the possible operating modes shown in Figure 6.2.1. For two wells we get four possible operating modes. In mode 1, both

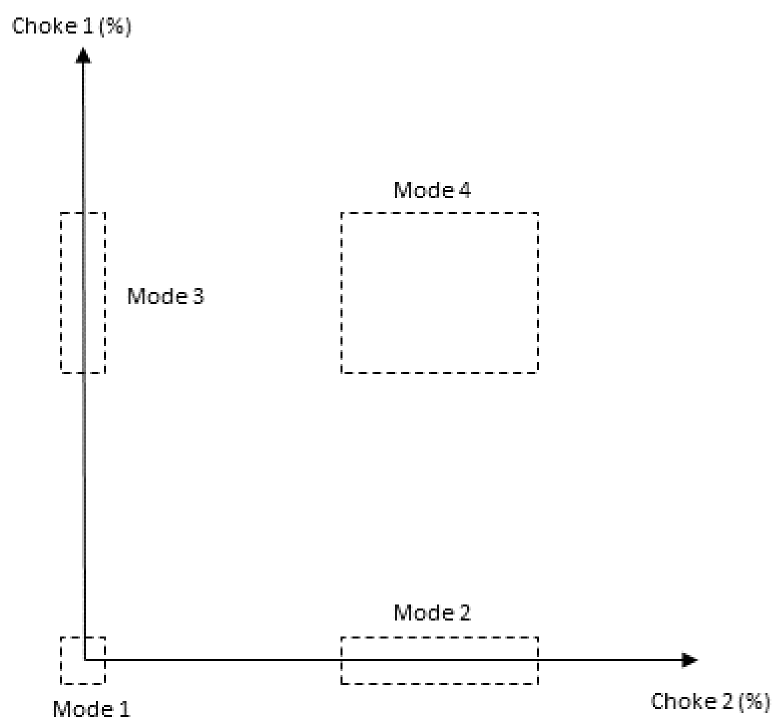


Figure 6.2.1: Operating modes for system with two wells

wells are turned off and the linear model is trivial as there is no production. In modes 2 and 3, one well is off while the other is allowed to move in some small interval, while both wells are on in mode 4. Increasing the number of wells increases the number of operating modes exponentially. Since each well can be either on or off, this gives 2^n possible modes for a system with n wells. However, it is unlikely that the historical data available support all these operating modes. Instead, only a subset of the modes have been visited in the past, or can be calculated from the available historical data. Only these modes can therefore be included in the model building and later the optimization problem.

6.3 Three Different Well Models

Depending on the measurements available at a field, different well models can be constructed. In the following, three well models that apply different measurements are presented.

6.3.1 Export Model

The total export measurements for oil and gas has less uncertainty than the individual multiphase meter measurements, but does not provide information about the contribution to production from each individual well. However, by employing multiple linear regression, the relationship between the choke openings of each well and total production can be modelled. Letting Q^o and Q^g be the vectors of export oil and gas measurements, and X be the matrix of ones in the first column (constant term) and choke opening measurements in remaining columns, the multiple linear regression equation for the total export measurement becomes

$$Q^o = X\beta^o + \varepsilon^o$$
$$Q^g = X\beta^g + \varepsilon^g$$

where β^o and β^g are best fitting lines to be estimated by the regression and ε^o and ε^g are the vectors of error terms. Since it is assumed that all wells have their own independent well model, the constant term contribution from a well must be removed from the total when a well is shut off.

In the export model, the constant terms in the models for oil and gas only hold for the case where all wells are turned on. When the system moves to a new operating mode by turning off one or more wells, it is not possible to remove the constant term contribution from these wells. Instead, a new linear model based on the historical data points where the system is in this new operating mode must be built. However, if the system is mostly operated in a mode with all wells on, there is not sufficient data to build a new model for each alternative mode. For optimization purposes, this means that the possible solutions is restricted to lie in one of the operating modes that have been seen historically and modeled for a field. This is illustrated in Figure 6.3.1, where the choke opening of a well at the Gjøa field is plotted. In the data set available, the choke has been turned off three times, and otherwise been operated between 30% and 50%. Assuming the other wells were on when this choke was turned off, only a limited amount of data points are available for modelling this operating mode.

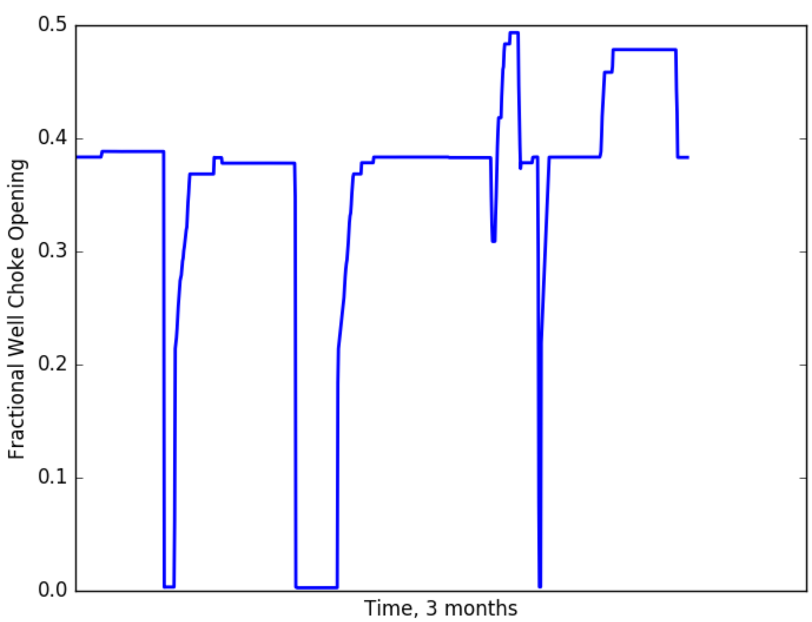


Figure 6.3.1: Choke opening of Gjøa well

6.3.2 Multiphase Meter (MPM) Model

For fields with multiphase meter measurements, it is possible to apply regression techniques to historical data, considering each well separately. Each well then has a set of measurements of its choke opening and related production of oil and gas for different points in time. For each well, a separate regression can be performed to estimate the individual slopes and constant terms. As a model is wanted for the total production of the field, the output of the regression must be combined. Letting $\hat{\beta}^o_{1i}$ and $\hat{\beta}^o_{0i}$ be the estimators for the slope and constant term describing the relationship between the choke opening of well i and oil production, and similarly for the gas production, we can combine the best fitting lines for each well, to find the model for the whole system:

$$q^o(x_1, x_2, ..., x_n) = x_1\hat{\beta}^o_{11} + \delta_1\hat{\beta}^o_{01} + x_2\hat{\beta}^o_{12} + \delta_2\hat{\beta}^o_{02} + ... + x_n\hat{\beta}^o_{1n} + \delta_n\hat{\beta}^o_{0n}$$
$$q^g(x_1, x_2, ..., x_n) = x_1\hat{\beta}^g_{11} + \delta_1\hat{\beta}^g_{01} + x_2\hat{\beta}^g_{12} + \delta_2\hat{\beta}^g_{02} + ... + x_n\hat{\beta}^g_{1n} + \delta_n\hat{\beta}^g_{0n}$$

where

$$\delta_i = \begin{cases} 0, & \text{if } x_i = 0 \\ 1, & \text{otherwise} \end{cases}$$

Note that x_i must lie in the semi-continuous interval of zero and $[\underline{L}_i, \overline{L}_i]$ in order for the model to be valid. Since we have independent models for each well, it is easy to remove the contribution from the wells that are turned off. Note that this makes it possible to make all 2^n different modes, which might lead to a complex optimization problem. However, many combinations can be excluded as it is unlikely that an operating mode where more than one or two wells are shut off is optimal.

6.3.3 Combined Model

The last approach is to use both the multiphase measurements and the total export measurements. Since the regression on the multiphase measurements is performed separately on each well, there is no coupling between the wells in the estimations. As it is assumed that the wells are independent, this should not be a problem. However, the multiphase meters can be unstable and return inaccurate measurements. In addition, when a multiphase meter is damaged for a well on a mature field, it is often not replaced as they are very costly to buy and install. Therefore, a model including both multiphase meter measurements and export measurements should provide more accurate predictions.

In this model, the matrix of measurements of the explanatory variables consists of several rows for each point in time. An example of the rows for the measurements at time t, for a field with three wells is then

$$X^t = \begin{pmatrix} 1 & 0 & 0 & x_{1t} & 0 & 0 \\ 0 & 1 & 0 & 0 & x_{2t} & 0 \\ 0 & 0 & 1 & 0 & 0 & x_{3t} \\ \underbrace{1 \quad 1 \quad 1}_{Constant} & & & \underbrace{x_{1t} \quad x_{2t} \quad x_{3t}}_{Chokes} & & \end{pmatrix}$$

where x_{it} is the choke opening of well i at time t. The first three lines in the matrix represent the multiphase measurements for the wells, while the last line is the export measurement during the same time period. When x_{it} is zero, well i is shut off at time t and the constant term corresponding to that well should be zero in the line for the total export measurement. The vector of measurements of oil production corresponding to the X_t matrix becomes:

$$Q_t^o = \begin{pmatrix} q_{1t}^o \\ q_{2t}^o \\ q_{3t}^o \\ q_t^o \end{pmatrix}$$

Here, q_{1t}^o is the multiphase meter oil production at time t from well 1, while q_t^o is the total oil production at the same point in time. If there was no uncertainty in the measurements, we would have:

$$q_t^o = q_{1t}^o + q_{2t}^o + q_{3t}^o$$

6.4 Assumptions behind use of Linear Regression

All well models presented apply linear regression when estimating appropriate well parameters. In order to apply linear regression, the underlying data must satisfy certain assumptions. However, the petroleum production data used as input for the well models does not necessarily satisfy all these assumptions. Firstly, the measurements applied in the regression are uncertain, meaning that we do not know whether the recorded production at some time was exactly what is given by the measurement. This may lead to bias in the estimated parameters (Yanez et al., 1998). In addition, the standard deviation of the parameters is in reality higher than predicted from the regression. These aspects are not treated in this thesis and are recorded as suggestions for further work.

When using the combined model, another assumption is violated. As the multiphase meter measurements have larger variance than the export measurements, the variation around the regression line is no longer constant. This can be solved by applying a weighting to the different measurements, so that more certain measurements are given more weight. In the regression, this weighting is multiplied with the error terms so that deviation from the export measurement is punished harder than deviation from the multiphase meter measurements.

Chapter 7

Optimization Models

In this chapter we first present the assumptions and simplifications for the developed optimization models in Chapter 7.1. In Chapter 7.2 we present a deterministic model, where the parameters take their expected values, and no uncertainty is incorporated. Then two different recourse models incorporating uncertainty are presented in Chapter 7.3. We introduce a model called the *penalty model*, where violations of the gas constraint leads to a penalty in the objective function. Another approach is also presented where the solution involves a sequencing of the wells, or an operational strategy, hence it is named the *strategy model*. An alternative formulation of the strategy model is also presented. In Chapter 7.4 an explanation of how the strategies given from the models are evaluated is given, before the chapter is ended by an algorithm explaining how the objective value is computed when perfect information is available in Chapter 7.5.

7.1 Assumptions and Simplifications

Application of Well Models

The real-time production optimization problem is complex due to the non-linear relationship between parameters such as flow rate and variables like choke opening. All three well models developed in Chapter 6 can be applied to generate the input parameters needed for the models presented in this chapter, resulting in linear optimization models. These well models estimate *slopes* for the oil and gas contribution of a well to the total production. Since the well models only have local validity, the chokes are required to lie between an upper and lower limit when they are not shut off. When a well is shut off, the system moves to a new operating mode, so a constant term must be subtracted from the total production of oil and gas. All optimization models presented are formulated with *delta* choke openings as decision variables. Therefore, the constant term for moving to a new operating mode is also a delta constant. When no chokes are turned off or on in the solution, the system is in the same operating mode as before the optimization. The delta constant is therefore zero for the operating mode representing the initial choke openings.

Objective and Constraints

This thesis is driven by the real-time optimization problem at Gjøa. The production at Gjøa is never close to its water handling capacity constraint and does not apply gas-lift. This is reflected in the optimization models presented. We have assumed that the only capacity constraint is on gas processing. This is not the case for all fields as there often are constraints on for example the amount of water handled. Although many fields use gas-lift to increase production, the only explicit decision variables used in these models are the choke openings. No options for routing of the well streams are included, and the capacity constraint is assumed to occur when the streams from all wells are gathered.

No considerations are made to other objectives than maximizing short-term oil production. Therefore, the solutions found may be in conflict with long-term objectives for example concerning the reservoir depletion plan.

7.2 Deterministic Model

When the well parameters are assumed to be certain, the real-time production optimization problem is deterministic. If there were no operating modes in the parameter models, and any number of wells could be adjusted, this would be a simple continuous knapsack problem. However, the model is complicated by the fact that it is possible to close a well completely and move to a new operating mode.

Indices

i	Well
k	Operating mode

Sets

$\mathcal{I}$	Set of wells
$\mathcal{K}$	Set of operating modes

Parameters

C_i	Slope of linear well model for oil production from well i
A_i	Slope of linear well model for gas production from well i
X_i	Initial opening of choke i
ΔQ	Remaining gas capacity at initial opening of chokes
$\overline{L}_i$	Upper limit on choke opening of well i
$\underline{L}_i$	Lower limit on choke opening of well i
N	Maximum number of chokes that can be adjusted
D_{ik}	Binary parameter indicating if choke of well i is on in mode k
B_k^o	Delta oil constant for moving to mode k
B_k^g	Delta gas constant for moving to mode k
ϵ	An infinitesimal value

Variables

x_i	Delta choke opening of well i
y_i	1 if the choke of well i is adjusted, else 0
w_k	1 if the system ends in operating mode k, else 0

Objective Function

The objective is to maximize total oil production. Since the x_i variables represent the delta choke opening for well i, the function to maximize is the change in oil production resulting from adjusting some wells. The second term in the objective function is the oil constant term contribution in the operating mode associated with the change in choke openings. The last term in the objective function is added to prevent the solution from including unnecessary steps. That is, if two solutions have the same expected production improvement, then the solution with the least number of steps used should be preferred.

$$\max w = \sum_{i \in \mathcal{I}} C_i x_i + \sum_{k \in \mathcal{K}} B_k^o w_k - \epsilon \sum_{i \in \mathcal{I}} y_i \tag{7.1}$$

Constraints

Number of Wells Adjusted A solution where many wells are adjusted is not optimal as the production engineer prefers not to make too many changes in one iteration. Therefore, a constraint on the maximum number of wells that can be changed is imposed. To achieve this, the y_i variables indicating whether a well is adjusted or not, must be forced to 1 when a choke is adjusted. Since the maximum positive delta choke opening for a well is $\overline{L}_i - X_i$ (when a well is initially shut and then opened to its upper limit) and the maximum negative delta choke opening is X_i (when a the choke of a well is turned off), this is achieved in the following equations:

$$x_i - y_i(\overline{L}_i - X_i) \leq 0 \qquad i \in I \tag{7.2}$$

$$x_i + y_i X_i \geq 0 \qquad i \in I \tag{7.3}$$

$$\sum_{i \in \mathcal{I}} y_i \leq N \tag{7.4}$$

Gas Constraint The gas produced for the delta choke openings must obey the remaining gas capacity (delta gas). Often, the production system is operated at the gas capacity, so that delta gas is zero.

$$\sum_{i\in\mathcal{I}} A_i x_i + \sum_{k\in\mathcal{K}} B_k^g w_k \quad \leq \quad \Delta Q \tag{7.5}$$

Operating Modes The final choke opening of each well must coincide with the final operating mode. Since the D_{ik} parameter is 1 when well i is open in mode k and w_k is a binary variable that is positive when k is the final operating mode, the final choke opening is forced to lie within the bounds of the final mode.

$$X_i + x_i - D_{ik} w_k \overline{L}_i - (1 - w_k)\overline{L}_i \quad \leq \quad 0 \qquad i \in \mathcal{I}, k \in \mathcal{K} \tag{7.6}$$

$$X_i + x_i - D_{ik} w_k \underline{L}_i \quad \geq \quad 0 \qquad i \in \mathcal{I}, k \in \mathcal{K} \tag{7.7}$$

$$\sum_{k\in\mathcal{K}} w_k \quad = \quad 1 \tag{7.8}$$

Variable Bounds The variables for delta choke opening are free as they can be both positive and negative. The values they can take however, is restricted by the final operating mode enforced in constraints (7.6)-(7.8). The remaining variables are binary.

$$x_i \quad free \qquad i \in \mathcal{I} \tag{7.9}$$

$$y_i \in \{0, 1\} \qquad i \in \mathcal{I} \tag{7.10}$$

$$w_k \in \{0, 1\} \qquad k \in \mathcal{K} \tag{7.11}$$

7.3 Recourse Models

In the following sections, two recourse models are presented. In addition to the assumptions described in Chapter 7.1, these models assume different characteristics regarding the work-process of the well operator.

7.3.1 Penalty Model

Assuming the *penalty approach*, described in Chapter 2.3, being employed by the operator, we develop an optimization model adapted to this operating pattern. According to this approach, the production engineer gives a list of instructions to the well operator, who implements the requested changes to the production system. However, as the exact relationship between the new settings and production is uncertain, the requested changes may lead to a violation of one or several constraints of the production system. A violation is not desirable as oil production then will be reduced, and the operator will in that case decrease the choke opening of some wells aggressively to be certain that the resulting gas production is again below the limit. The resulting production loss of this action can be included as a penalty in the objective function when modelling. There are many ways to model this penalty, and the most appropriate way may vary between fields. In this formulation, the relationship between the size of the necessary recourse action to obtain feasibility and associated reduction in oil and gas production is assumed to be linear.

The penalty model is a two-stage recourse model, where the first stage decision is a set of well adjustments to be implemented. In the second stage, these changes have been performed and new information is revealed regarding the total production of oil and gas resulting from the first stage actions. In addition, it is assumed that the production engineer will have gained more information on the individual oil and gas slopes of each well. Based on this information, the second stage decision, or the recourse action, is the wells that should be adjusted to achieve feasibility if the gas constraint is breached.

An example of a second stage solution for a scenario s, assuming a problem involving two wells is given in Figure 7.3.1. Each axis represents the choke opening of the wells, and X_1 and X_2 are the initial choke openings. The gas constraint included in the illustration is scenario-specific and depends on the outcome of the well parameters in each scenario. In a scenario where the slope of the gas contributions from the wells is low, the constraint indicated will shift outwards, as larger choke openings then are feasible. The first stage decision is not indexed by scenario as it must be the same for all scenarios, and is shown as the path indicated by the variables x_1 and x_2. These variables represent the adjustments made to each choke. As the first-stage decision violates the gas constraint in this scenario, the opening of a choke must be reduced. This recourse action is represented by the r_{2s} variable. r_{2s} is used to penalize constraint violations, by associating a penalty cost to these variables in the objective function.

In the following, the mathematical formulation of the penalty model is given. First, the necessary declarations are presented, followed by the objective function and constraints.

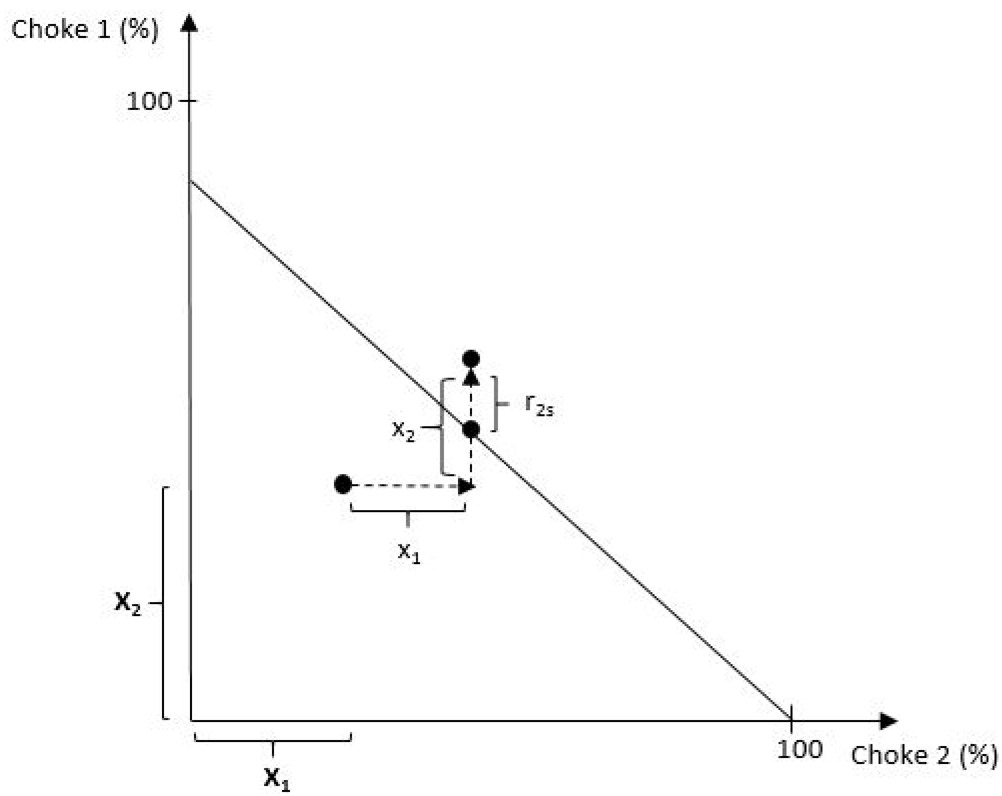


Figure 7.3.1: Example of second stage solution to the penalty model

Declarations

Indices

i	Well
k	Operating mode
s	Scenario

Sets

$\mathcal{I}$	Set of wells
$\mathcal{K}$	Set of operating modes
$\mathcal{S}$	Set of scenarios

Parameters

C_{is}	Oil gradient of well i in scenario s
A_{is}	Gas gradient of well i in scenario s
X_i	Initial opening of choke i
ΔQ	Remaining gas capacity at initial opening of chokes
$\overline{L}_i$	Upper limit on choke opening of well i
$\underline{L}_i$	Lower limit on choke opening of well i
N	Maximum number of chokes that can be adjusted
D_{ik}	Binary parameter indicating if choke of well i is on in mode k
B^o_{ks}	Delta oil constant for moving to mode k in scenario s
B^g_{ks}	Delta gas constant for moving to mode k in scenario s
C^p	Unit penalty cost of violating the gas constraint
ϵ	An infinitesimal value

Variables

x_i	Delta choke opening of well i
$\overline{y}_i$	1 if the choke of well i is adjusted up, else 0
$\underline{y}_i$	1 if the choke of well i is adjusted down, else 0
r_{is}	Amount to decrease choke of well i to abide gas limit in scenario s
w_{ks}	1 if the system ends in operating mode k in scenario s

Objective Function

The objective is to maximize total oil production. The first term in the function represents the delta choke openings adjusted for violations of the gas constraint. When the gas constraint is violated, one or more r_{is} is positive, leading to a punishment in the objective function. This punishment should be representative for the lost production resulting from having to adjust additional chokes after the constraint is violated. The lost production is here assumed to be linear in the violation of the gas constraint, as it is reasonable to assume that the production engineer is more aggressive when adjusting the system if the violation of the constraint is high. The last term in the objective function is added to prevent the solution from including unnecessary steps. That is, if two solutions have the same expected production improvement, then the solution with the least number of steps used is preferred.

$$\max w = \frac{1}{|S|}\left(\sum_{i\in\mathcal{I}}\sum_{s\in\mathcal{S}} C_{is}(x_i - r_{is}) - \sum_{i\in\mathcal{I}}\sum_{s\in\mathcal{S}} C^p r_{is} + \sum_{k\in\mathcal{K}}\sum_{s\in\mathcal{S}} B^o_{ks} w_{ks}\right) - \epsilon \sum_{i\in\mathcal{I}}(\overline{y_i} + \underline{y_i}) \tag{7.12}$$

Constraints

Number of Wells Adjusted These constraints have the same effect as in the deterministic model, with the new variables indicating whether a well has been adjusted.

$$x_i - \overline{y}_i(\overline{L} - X_i) \leq 0 \qquad i \in \mathcal{I} \tag{7.13}$$

$$x_i + \underline{y}_i X_i \geq 0 \qquad i \in \mathcal{I} \tag{7.14}$$

$$\sum_{i\in\mathcal{I}}(\overline{y}_i + \underline{y}_i) \leq N \tag{7.15}$$

Gas Constraint The gas constraint must hold for the final choke openings in all scenarios.

$$\sum_{i\in\mathcal{I}} A_{is}(x_i - r_{is}) + \sum_{k\in\mathcal{K}} B^g_{ks} w_{ks} \leq \Delta Q \qquad s \in \mathcal{S} \tag{7.16}$$

Operating Modes As in the deterministic formulation, the final choke openings must lie within the bounds of the final operating mode. In each scenario, exactly one operating mode is reached.

$$
\begin{aligned}
X_i + x_i - r_{is} - D_{ik} w_{ks} \overline{L}_i - (1 - w_{ks}) \overline{L}_i &\leq 0 \qquad i \in \mathcal{I}, k \in \mathcal{K}, s \in \mathcal{S} && (7.17)\\
X_i + x_i - r_{is} - D_{ik} w_{ks} \underline{L}_i &\geq 0 \qquad i \in \mathcal{I}, k \in \mathcal{K}, s \in \mathcal{S} && (7.18)\\
\sum_{k \in \mathcal{K}} w_{ks} &= 1 \qquad s \in \mathcal{S} && (7.19)
\end{aligned}
$$

Allowed Recourse Actions The opening of a well is only allowed to be reduced in the second stage if the opening was increased in the first stage.

$$
\begin{aligned}
x_i - r_{is} + (1 - \overline{y}_i) &\geq 0 \qquad i \in \mathcal{I}, s \in \mathcal{S} && (7.20)\\
r_{is} - \overline{y} &\leq 0 \qquad i \in \mathcal{I}, s \in \mathcal{S} && (7.21)
\end{aligned}
$$

Variable Bounds

$$
\begin{aligned}
x_i \quad & free && i \in \mathcal{I} && (7.22)\\
r_{is} \quad & \geq 0 && i \in \mathcal{I}, s \in \mathcal{S} && (7.23)\\
\overline{y}_i, \underline{y}_i & \in \{0, 1\} && i \in \mathcal{I} && (7.24)\\
w_{ks} & \in \{0, 1\} && k \in \mathcal{K}, s \in \mathcal{S} && (7.25)
\end{aligned}
$$

7.3.2 Strategy Model

If it is assumed that the changes imposed on the choke openings are implemented sufficiently slowly, the production engineer has the opportunity to observe when the gas constraint is met. In this case the gas constraint should never be violated, as a violation can result in reduced production. If the decision variables now include a sequencing of how the wells should be adjusted, the recourse action is the ability to stop following this sequence when a capacity constraint is met. The solution to the optimization problem is then a strategy that can be followed by the operator.

This is a recourse model with two stages. In the first stage, an operational strategy is set, defining a number of wells to be adjusted and a sequencing of these adjustments. It is assumed that the operator is informed when the gas constraint is met, and can then terminate the strategy planned in the first stage. The opportunity to adapt to the realization of the stochastic parameters in the second stage is the recourse of the formulation. Each outcome of the parameters will then lead to different final set-points for the system.

This model is based on the work by Hanssen and Foss (2015) with some important adaptions. Instead of assuming that all wells are turned off initially, this model optimizes from any operating point of the system. This is more realistic as it is uncommon for a production system to be completely shut down. The motivation for Hanssen and Foss to start with zero choke openings for all wells, was that it was difficult to assure that all scenarios were feasible in the starting point. We handle this by introducing a delta-model where the change in each choke opening is considered as the decision variable instead of the absolute opening. In addition, Hanssen and Foss assume global validity of the input parameter model, which is not realistic. By introducing operating modes and only allowing for small adjustments, we present a model better fit to actual real-life data.

In the mathematical model, a strategy is defined over a set of states, indexed by j. Each state represents a certain choke opening of the wells along a strategy. We use the term *step* to refer to what happens between two consecutive states. In Figure 7.3.2, an example of a second stage solution for a scenario s, in a problem with two wells, indexed by i, is illustrated. Both the first stage x_{ij}-variables and second stage z_{ij}-variables are indexed by both well and state.

Assuming that the first stage strategy is to first increase the choke of well 1 by 10%, followed by an increase in the choke of well 2 by 10%, the value of the different variables for this scenario are shown in Figure 7.3.3. In this example, the gas constraint is met between the second and third state, and the realized choke opening for well 2 is therefore only 5%. The v_{is} variable is binary and only 1 for the last state that is visited in a scenario. In this scenario, the constraint is met between the second and third states, so $v_{2s} = 1$. As only well 2 is adjusted between the last states, the final delta choke opening of well 1 is the same as in state two. The strategy needs to terminate between exactly two states in each scenario. A part of the optimization therefore becomes to find the optimal sequencing of the wells in the strategy.

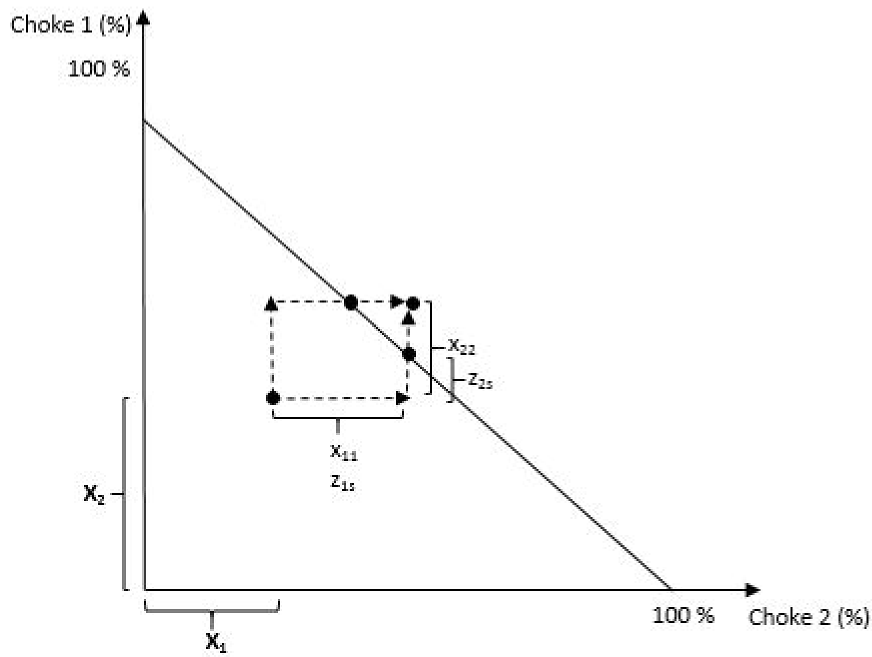


Figure 7.3.2: Example of second stage solution to the strategy model

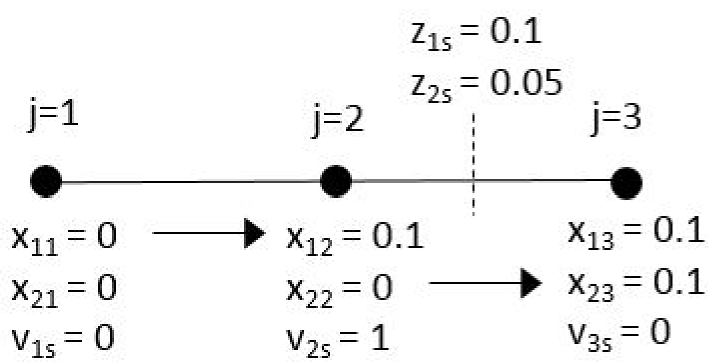


Figure 7.3.3: Illustration of steps in a strategy

In the following, the mathematical formulation of the strategy model is given. First, the necessary declarations are presented, followed by the objective function and constraints.

Declarations

Indices

i	Well
j	State
k	Operating mode
s	Scenario

Sets

$\mathcal{I}$	Set of wells
$\mathcal{J}$	Set of states
$\mathcal{K}$	Set of operating modes
$\mathcal{S}$	Set of scenarios

Parameters

C_{is}	Oil gradient of well i in scenario s
A_{is}	Gas gradient of well i in scenario s
X_i	Initial opening of choke i
ΔQ	Remaining gas capacity at initial opening of chokes
$\overline{L}_i$	Upper limit on choke opening of well i
$\underline{L}_i$	Lower limit on choke opening of well i
N	Maximum number of chokes that can be adjusted
D_{ik}	Binary parameter indicating if choke of well i is on in mode k
B^o_{ks}	Delta oil constant for moving to mode k in scenario s
B^g_{ks}	Delta gas constant for moving to mode k in scenario s
ϵ	An infinitesimal value

Variables

x_{ij}	Delta choke opening of well i in state j
y_{ij}	1 if the choke of well i is adjusted in state j, else 0
z_{is}	Final delta choke opening of well i in scenario s
v_{js}	1 if the constraint is met between states j and $j+1$, else 0
w_{jk}	1 if the system is in operating mode k in state j, else 0
δ_{ks}	1 if the realized operating mode is k in scenario s, else 0

Objective Function

The goal of the optimization is to maximize total oil production. Therefore, the objective in the recourse model with operating modes is the expected increase in oil production over all scenarios generated. A scenario is a realization of the slopes of the oil and gas contribution from each well. If it is optimal to change operating mode in a scenario, a constant term must be added to the objective. The last term in the objective function is added to prevent the solution from including unnecessary steps. That is, if two solutions have the same expected production improvement, then the solution with the least number of steps used is preferred.

$$\max w = \frac{1}{|\mathcal{S}|}\left(\sum_{s\in\mathcal{S}}\left(\sum_{i\in\mathcal{I}} C_{is}z_{is} + \sum_{k\in\mathcal{K}} B^{o}_{ks}\delta_{ks}\right)\right) - \epsilon\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}\setminus\{|\mathcal{J}|\}} y_{ij} \tag{7.26}$$

Constraints

Wells Adjusted in Each Step In each step of a strategy, only one well can be adjusted. The following constraints ensure that if the choke opening for some well, i, changes in step j, no other well can change in this step.

$$x_{ij} - x_{i(j+1)} - y_{ij}\overline{L_i} \leq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}\setminus\{|\mathcal{J}|\} \tag{7.27}$$

$$x_{ij} - x_{i(j+1)} + y_{ij}\overline{L_i} \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}\setminus\{|\mathcal{J}|\} \tag{7.28}$$

$$\sum_{i\in\mathcal{I}} y_{ij} \leq 1 \qquad j \in \mathcal{J}\setminus\{|\mathcal{J}|\} \tag{7.29}$$

Realized Choke Openings The realized delta choke opening for each well must obey the strategy defined by the first stage variables. The gas constraint is met when adjusting some well between exactly two states. Only if the constraint is met between states j and $j+1$, is the realized delta choke opening allowed to lie between the planned change in choke opening of two states. Since the slopes of the oil and gas contributions are assumed to be positive in all scenarios, the gas constraint will only be met between two states when the opening of a choke is increased.

$$
\begin{aligned}
x_{ij} - z_{is} - (1 - v_{js})\overline{L_i} &\leq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \qquad (7.30)\\
x_{i(j+1)} - z_{is} + (1 - v_{js})\overline{L_i} &\geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \qquad (7.31)\\
\sum_{j \in \mathcal{J} \setminus \{N^s+1\}} v_{js} &= 1 \qquad s \in \mathcal{S} \qquad (7.32)
\end{aligned}
$$

Gas Constraint Delta gas production resulting from the changes made to the set-points of the operating system must lie within the available gas capacity. If the system is in a new operating mode, a constant must be added to the constraint.

$$
\sum_{i \in \mathcal{I}} A_{is} z_{is} + \sum_{k \in \mathcal{K}} B_k^g \delta_{ks} \leq \Delta Q \qquad s \in \mathcal{S} \qquad (7.33)
$$

All States Must Be Feasible All states visited when following a strategy must be feasible with respect to the gas constraint.

$$
\sum_{i \in \mathcal{I}} A_{is} x_{ij} + \sum_{k \in \mathcal{K}} B_{ks}^g w_{jk} - M_s(1 - \sum_{t=j}^{t=|\mathcal{J}|} v_{ts}) \leq \Delta Q \qquad j \in \mathcal{J}, s \in \mathcal{S} \qquad (7.34)
$$

Where the parameter M_s is given by:

$$
M_s = \sum_{i \in \mathcal{I}} A_{is}(\overline{L}_i - X_i) + \max_{k \in \mathcal{K}}(B_{ks}) - \Delta Q \qquad s \in \mathcal{S} \qquad (7.35)
$$

Operating Modes In each state the system must be in exactly one operating mode.

$$
\begin{aligned}
\sum_{k \in \mathcal{K}} w_{kj} &= 1 \qquad j \in \mathcal{J} \qquad (7.36)\\
\sum_{k \in \mathcal{K}} \delta_{ks} &= 1 \qquad s \in \mathcal{S} \qquad (7.37)
\end{aligned}
$$

Legal States The operating mode in each state is defined by the value of the first and second stage variables.

$$X_i + x_{ij} - D_{ki} w_{kj} \overline{L}_i - (1 - w_{kj}) \overline{L}_i \leq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \tag{7.38}$$

$$X_i + x_{ij} - D_{ki} w_{kj} \underline{L}_i \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \tag{7.39}$$

$$X_i + z_{ij} - D_{ki} \delta_{ks} \underline{L}_i - (1 - \delta_{ks}) \overline{L}_i \leq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \tag{7.40}$$

$$X_i + z_{ij} - D_{ki} \delta_{ks} \underline{L}_i \geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K} \tag{7.41}$$

Variable Requirements

$$x_{ij} \quad free \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{7.42}$$

$$z_{is} \quad free \qquad i \in \mathcal{I}, s \in \mathcal{S} \tag{7.43}$$

$$y_{ij} \in \{0, 1\} \qquad i \in \mathcal{I}, j \in \mathcal{J} \tag{7.44}$$

$$v_{js} \in \{0, 1\} \qquad j \in \mathcal{J}, s \in \mathcal{S} \tag{7.45}$$

$$w_{jk} \in \{0, 1\} \qquad j \in \mathcal{J}, k \in \mathcal{K} \tag{7.46}$$

$$\delta_{ks} \in \{0, 1\} \qquad k \in \mathcal{K}, s \in \mathcal{S} \tag{7.47}$$

7.3.3 Alternative Formulation of the Strategy Model

The formulation of the strategy model presented in Chapter 7.3.2 only works when ΔQ is greater than or equal to zero, due to constraint 7.34 ("all visited states must be feasible"). As mentioned in Chapter 5, the operator may want to reduce gas production due to for example a planned maintenance activity. In our model, a reduction in gas production is represented by a negative ΔQ value. Since this makes the previous formulation of the strategy model integer infeasible, an alternative formulation of the strategy model is suggested. This alternative formulation works for both positive and negative values of ΔQ.

An underlying assumption of this formulation is that all down-adjustments of chokes will happen prior to any up-adjustments in an optimal strategy. We believe that moving all down-adjustments to the start of the strategy will rarely or never affect the quality of the solution. Intuitively this makes sense, as performing down-adjustments involves moving away from the gas capacity restriction. Moving away from the capacity restriction should increase the probability that all or most of the planned adjustments can be started. On the other hand, if an up-adjustment is planned before a down-adjustment, there is an increased probability of reaching the gas capacity constraint early. Thus, a strategy were all down-adjustments are performed prior to any up-adjustments should never end earlier than a strategy where the down-adjustments are placed in between the up-adjustments. On average, being able to perform as many steps of the strategy as possible must be a goal. The reason for this is that the model would not suggest adjustments that on average should not be performed. Therefore, we believe that the assumption that all down-adjustments should happen first is valid. In the rest of this section we present the additions and modifications that must be made to the previous strategy model in order to create this alternative formulation.

Note that an alternative formulation is only created for the *strategy* model. Modifying the *penalty* model in order to handle negative ΔQ values does not make sense. Recall that the recourse action in the penalty model is down-adjusting one of the chokes that are adjusted up in the first stage. This recourse action is penalized in the objective. If only down-adjustments are possible for a particular negative ΔQ, an alternative penalty model could be created where *up-adjustments* in the second stage were penalized. However, when ΔQ is negative, only allowing penalized up-adjustments in the second stage is not sufficient, as it may be that the first stage solution becomes infeasible when it is fixed and evaluated on a different scenario tree. When both up-adjustments and down-adjustments are possible in the recourse stage, the model has a lot of freedom to tailor the solution for each specific scenario. This should not be possible, and thus a penalty model that works for negative ΔQ is not realistic.

Declarations

New Variables

u_j	First up-adjustment of a choke happens in state j. Prior to this state, only down-adjustments may have been performed.

New Constraints

Down-Adjustments before Up-Adjustments If the strategy contains any down-adjustments of chokes, all the down-adjustments must happen prior to any up-adjustments. The constraints below ensure that prior to the first up-adjustment, adjustments can only be performed downwards, and after the first up-adjustment, adjustments can only be performed upwards. Note that there can only be a single switch between down- and up-adjustments, and that this switch may happen in the last state. A switch in the last state means that the strategy will not contain any up-adjustments at all.

With these new constraints, constraints (7.34) can be safely removed. Since the realized state must be feasible, and all down-adjustments must happen prior to any up-adjustments, there is no way to reach an infeasible state and then go back to a feasible state when ΔQ is greater than or equal to zero. Further, when ΔQ is negative, the first states will be infeasible, resulting in an integer infeasible solution in the original formulation of the model. With these new constraints however, infeasibility in the first states is allowed, and the strategy must move towards feasibility before any up-adjustments are performed, to ensure that the final state is feasible.

$$
\begin{aligned}
x_{ij} - x_{i(j+1)} + \sum_{k=0}^{j} u_k &\geq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\} && (7.48)\\
x_{ij} - x_{i(j+1)} - \sum_{k=j+1}^{|\mathcal{J}|} u_k &\leq 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\} && (7.49)\\
\sum_{j \in \mathcal{J}} u_j &= 1 && (7.50)
\end{aligned}
$$

Preventing Abortion of Strategy If the strategy includes swapping of chokes, that is adjusting down one or more chokes in order to adjust up one or more other chokes, the strategy may result in a worsened oil production in bad scenarios. Then, it is beneficial to abort the strategy and keep production settings as they are. However, in reality it is not possible to foresee whether the oil production will become better or worse, prior to having executed the strategy. In this alternative formulation of the strategy model, a legal solution may be to end the strategy in the initial state (when $\Delta Q \geq 0$), that is aborting the planned strategy. This should not be a legal solution. The following constraint is therefore added to ensure that if any up-adjustments are planned, the strategy cannot be ended before these adjustments are started. In other words, this constraint makes it impossible to *not* follow the strategy as far as possible.

$$\sum_{j \in \mathcal{J} \setminus \{|\mathcal{J}|\}} j \cdot v_{js} - \sum_{j \in \mathcal{J} \setminus \{|\mathcal{J}|\}} j \cdot u_j \quad \geq \quad 0 \qquad s \in \mathcal{S} \tag{7.51}$$

Realized Choke Openings If ΔQ is sufficiently negative, all planned adjustments will be downwards in order to reach a feasible state. Then, the strategy will end during a down-adjustment, i.e. as soon as the gas capacity constraint holds. In the original formulation this would not work, due to constraints (7.32). This constraint should only hold when the last step involves an up-adjustment. When the last step j involves a down-adjustment, some of the signs must be flipped such that $x_{ij} \geq z_{is} \geq x_{i(j+1)}$, rather than $x_{ij} \leq z_{is} \leq x_{i(j+1)}$ for the last adjusted choke i.

$$x_{ij} - z_{is} - \overline{L}_i(1 - v_{js}) - \overline{L}_i \sum_{k=j+1}^{|\mathcal{J}|} u_k \quad \leq \quad 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \tag{7.52}$$

$$x_{i(j+1)} - z_{is} + \overline{L}_i(1 - v_{js}) + \overline{L}_i \sum_{k=j+1}^{|\mathcal{J}|} u_k \quad \geq \quad 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \tag{7.53}$$

$$x_{ij} - z_{is} + \overline{L}_i(1 - v_{js}) + \overline{L}_i \sum_{k=0}^{j} u_k \quad \geq \quad 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \tag{7.54}$$

$$x_{i(j+1)} - z_{is} - \overline{L}_i(1 - v_{js}) - \overline{L}_i \sum_{k=0}^{j} u_k \quad \leq \quad 0 \qquad i \in \mathcal{I}, j \in \mathcal{J} \setminus \{|\mathcal{J}|\}, s \in \mathcal{S} \tag{7.55}$$

7.4 Evaluating a Strategy

To make sure that the solutions given by the optimization models are generalizable, i.e. not tailored to the random errors in the scenario set, the solutions are evaluated on an independent, much larger scenario tree. Similar to the scenario trees used in the optimization models, this evaluation tree is also generated using moment matching. Evaluating the solution given by the penalty model is easy, as this corresponds to fixing the first stage variables, and solving the optimization problem for the second stage variables. However, this method does not always work when evaluating a solution given by the strategy model, because ΔQ can be negative. When ΔQ is negative, a solution given by the stochastic strategy model or the deterministic model, may result in an infeasible problem when their first stage variables are fixed and a new scenario tree is used for evaluation.

To overcome this problem we choose to create an *evaluation algorithm* which is used for evaluating all strategies given by the stochastic strategy model or its deterministic counterpart. This is the algorithm presented in Algorithm 4 and Algorithm 5. We present the algorithm in two parts to make it more comprehensive, and because the two parts have different objectives.

The objective of Algorithm 4 is to evaluate the strategy *as given by the optimization model*. The strategy is followed as far as possible, until the gas capacity is possibly reached, or there are no more steps in the strategy. Reaching the gas capacity constraint is tested by the `if`-statement in line 8 and line 17, respectively. In line 8, we check if the gas capacity is reached during a down-adjustment, and in line 17 we test if it is reached by an up-adjustment. If the gas capacity is reached by a down-adjustment and the strategy contains down-adjustments only, we wish to stop as soon as possible. This stop point may be below the lower limit of the choke ($\underline{L}_i$), but above 0. Then, the choke must be turned completely off. If the plan was to turn the choke off, but it remains above its lower limit, the oil and gas constants must be given by the mode in state j rather then the mode in state $j+1$. Reaching the gas capacity by an up-adjustment does not involve such complexities. When the gas capacity is reached, or all adjustments have been implemented, the resulting oil and gas values are given as the return value by the algorithm. Note that these values may correspond to an illegal state, i.e. $gas + gas\_constant \geq \Delta Q$.

The objective of Algorithm 5 is to fix any infeasibility issues, or to ensure that the last step of the strategy brings you as close as possible to the gas capacity. Infeasibility is fixed by first trying to further adjust the choke that was used in the last step. If the resulting state is still not feasible, chokes that have not yet been adjusted down are adjusted down in order of expected GOR (high to low). This is continued until the gas capacity constraint is reached. Alternatively, it may be that the state given by evaluating the base strategy is feasible, but not *at* the gas capacity. In reality, the operator would then continue to adjust this choke until the capacity is reached. Therefore, this is tested for by the `if`-statement in line 4, and possibly executed. Note that this up-adjustment after the base strategy has been implemented results in a (in our opinion) more fair comparison of the stochastic and deterministic solutions. The reason is that the deterministic model always proposes strategies that in many scenarios will stop before the gas capacity is reached. This more rarely happens for the strategies given by the stochastic model, because this model 'knows' that there is no downside to adjusting the last choke to its limit (adjusting too far has no penalty in the strategy model). Therefore, we believe that the stochastic model would have a too large advantage if this final up-adjustment is not modified, and that this would give an unrealistic VSS.

Note also that this algorithm has an important underlying assumption, namely that the operator *cannot abort strategies*. We believe the best way to evaluate the strategies is to force them to be completed as far as possible. If strategies could be aborted, this information should be given to optimization model. The reason is that, when abortions are allowed, risky strategies are preferred. If strategies could simply be aborted in bad scenarios, only the *positive* part of the uncertainty is taken into account. Whether our assumption is good is up for discussion, as it may actually be easy to revert the system back to its old state. Further, in reality operators are often risk averse, and choose to cancel the strategy as soon as it does not behave as expected. We however believe that on average, a higher production improvement is achieved by following all steps of a strategy, and rather re-optimize if the strategy does not have the desired effect.

Algorithm 4: Evaluation of Base Strategy

Data: A choke opening solution (x^*), the corresponding operating mode solution (w^*) and the scenario which the solution is to be evaluated for (s). The other parameters used in this algorithm are those defined in Chapter 7.3.2.

Result: The oil and gas values after the solution (x^*, w^*) has been implemented as it was given by the optimization model. Note that this may be an infeasible state.

down_only $\leftarrow$ true if only down-adjustments are made, false otherwise;
end_strategy $\leftarrow$ false;
for $j \in \mathcal{J} \setminus \{|\mathcal{J}|\}$ **do**
    oil_constant $\leftarrow \sum_{k\in\mathcal{K}} B^o_{sk} w^*_{(j+1)k}$;
    gas_constant $\leftarrow \sum_{k\in\mathcal{K}} B^g_{sk} w^*_{(j+1)k}$;
    for $i \in \mathcal{I}$ **do**
        $\Delta x \leftarrow x^*_{i(j+1)} - x^*_{ij}$;
        if $\Delta x < 0$ *and* down_only *and* ($A_{is}\Delta x +$ gas_constant) $\leq \Delta Q$ **then**
            $\Delta x \leftarrow (\Delta Q - \sum_{k\in\mathcal{K}} B^o_{sk} w^*_{jk} -$ gas$)/A_{is}$;
            if $x^*_{i(j+1)} = 0$ **then**
                if $x^*_{ij} + \Delta x < \underline{L}_i$ **then**
                    $\Delta x = -x^*_{ij}$;
                else
                    oil_constant $\leftarrow \sum_{k\in\mathcal{K}} B^o_{sk} w^*_{jk}$;
                    gas_constant $\leftarrow \sum_{k\in\mathcal{K}} B^g_{sk} w^*_{jk}$;
            end_strategy $\leftarrow$ true;
        else if $\Delta x > 0$ *and* (gas $+ A_{is}\Delta x +$ gas_constant) $\geq \Delta Q$ **then**
            $\Delta x \leftarrow (\Delta Q -$ gas_constant $-$ gas$)/A_{is}$;
            end_strategy $\leftarrow$ true;
        if $\Delta Q \neq 0$ **then**
            oil $\leftarrow$ oil $+ C_{is}\Delta x$;
            gas $\leftarrow$ gas $+ A_{is}\Delta x$;
            break loop;
    if end_strategy **then**
        break loop;
return oil, oil_constant, gas, gas_constant

Algorithm 5: Evaluation of Strategy

Data: A choke opening solution (x^*), the corresponding operating mode solution (w^*) and the scenario which the solution is to be evaluated for (s). The other parameters used in this algorithm are those defined in Chapter 7.3.2.

Result: The evaluated objective value of the choke opening solution (x^*, w^*) in scenario s. This is the objective value after the *base* strategy has been evaluated, and infeasibility and optimality has been fixed.

oil, oil_constant, gas, gas_constant $\leftarrow$ `evaluate_base_strategy(`x^*, w^*`)`;
n $\leftarrow$ index of last adjusted choke;
m $\leftarrow$ index of last visited state;
if $(\mathsf{gas} + \mathsf{gas\_constant}) < \Delta Q$ **then**
    $\Delta x \leftarrow$ `min`$((\Delta Q - \mathsf{gas\_constant} - \mathsf{gas})/A_{ns},\ (\overline{L}_n - x^*_{nm}))$;
    $\mathsf{oil} \leftarrow \mathsf{oil} + C_{ns}\Delta x$;
    $\mathsf{gas} \leftarrow \mathsf{gas} + A_{ns}\Delta x$;
else
    while $(\mathsf{gas} + \mathsf{gas\_constant}) < \Delta Q$ **do**
        $\Delta x \leftarrow (\Delta Q - \mathsf{gas\_constant} - \mathsf{gas})/A_{ns}$;
        if $(x^*_{n(m+1)} + \Delta x) < \underline{L}_n$ **then**
            $\Delta x \leftarrow -x_{n(m+1)}$;
            $w_k^{New} \leftarrow$ find mode for new choke setup;
            $\mathsf{oil\_constant} \leftarrow \sum_{k \in \mathcal{K}} B^o_{sk} w_k^{New}$;
            $\mathsf{gas\_constant} \leftarrow \sum_{k \in \mathcal{K}} B^g_{sk} w_k^{New}$;
        $\mathsf{oil} \leftarrow \mathsf{oil} + C_{ns}\Delta x$;
        $\mathsf{gas} \leftarrow \mathsf{gas} + A_{ns}\Delta x$;
        n $\leftarrow$ select choke with worst (highest) GOR that has not yet been adjusted;
return $\mathsf{oil} + \mathsf{oil\_constant}$

7.5 Evaluating Perfect Information

To find the objective value that can be achieved with perfect information about oil production and gas production, we create an algorithm, Algorithm 6. This value could also be found by modifying the first stage variables of the stochastic optimization problems, such that each scenario can have a unique solution, and then solve the resulting problem. However, with many scenarios, testing shows that the resulting optimization model takes a long time to solve. Therefore, this algorithm is created. The algorithm is much more efficient than the *perfect information* model, and finds the objective value in a few seconds, compared to a few hours for the optimization model.

The algorithm is recursive, and for a particular call to the algorithm one step is evaluated. In a particular step you need to know:

- s: the scenario to be evaluated
- $\mathcal{I}^*$: the indices of the chokes that have not yet been considered
- j: the number of the step
- off: the number of chokes that are currently off
- d_gas: the current change in gas
- obj: the current change in oil (objective value)

In the evaluation of a step, the value of each possible option is added to an array, and then the algorithm returns the maximum value in this array. The possible options for each choke are to:

1. Stop at the gas capacity, if adjusting a choke towards its upper limit results in crossing the gas capacity constraint (line 10)
2. Stop at the gas capacity, if adjusting a choke towards its lower limit results in crossing the gas capacity constraint (line 13)
3. Adjust the choke to its upper limit (line 17)
4. Adjust the choke to its lower limit (line 19)
5. Turn the choke off, if the number of chokes turned off this far is less than 2 (line 21)

Note that only the extreme values of choke openings (upper limit, lower limit, and 0) are considered when evaluating, as all chokes *except the last choke adjusted* will take such a value. Only a strategy like this will be optimal when you have perfect information about oil production and gas production.

Algorithm 6: Objective Value with Perfect Information

Data: The parameters as defined in Chapter 7.3.2.

Result: The objective value that will be achievable with perfect information about future oil production and gas production

func `eval_scenario`(s, $\mathcal{I}^*$, j, off, d_gas, obj)
    if $j ==$ num_steps **then**
        if d_gas $\leq \Delta Q$ **then**
            return obj
        return $-\infty$;
    obj_options $\leftarrow$ empty array;
    if d_gas $\leq \Delta Q$ **then**
        add obj to obj_options;
    for $i \in \mathcal{I}^*$ **do**
        if d_gas $+ A_{is}(\overline{L}_i - X_i) > \Delta Q >$ d_gas **then**
            obj_new $\leftarrow$ obj $+ C_{is}(\Delta Q -$ d_gas$)/A_{is}$;
            add obj_new to obj_options;
        if d_gas $+ A_{is}(\underline{L}_i - X_i) < \Delta Q <$ d_gas **then**
            obj_new $\leftarrow$ obj $+ C_{is}(\Delta Q -$ d_gas$)/A_{is}$;
            add obj_new to obj_options;
        remove i from $\mathcal{I}^*$;
        obj_new $\leftarrow$ `eval_scenario`(s, $\mathcal{I}^*$, j+1, off, d_gas $+ A_{is}(\overline{L}_i - X_i)$, obj $+ C_{is}(\overline{L}_i - X_i)$);
        add obj_new to obj_options;
        obj_new $\leftarrow$ `eval_scenario`(s, $\mathcal{I}^*$, j+1, off, d_gas $+ A_{is}(\underline{L}_i - X_i)$, obj $+ C_{is}(\underline{L}_i - X_i)$);
        add obj_new to obj_options;
        if off < 2 **then**
            obj_new $\leftarrow$ `eval_scenario`(s, $\mathcal{I}^*$, j+1, off $+ 1$, d_gas $- A_{is}X_i - B^g_{is}$, obj $- C_{is}X_i - B^o_{is}$);
            add obj_new to obj_options;
    return `max`(obj_options);
off_strart $\leftarrow$ number of chokes that are turned off in the initial choke setup;
return $\sum_{s \in \mathcal{S}}$ `eval_scenario`(s, $\mathcal{I}$, 0, off_strart, 0, 0);

Chapter 8

Case Study

To test and evaluate the models suggested, input data is necessary. Production data from the Gjøa field has been made available. Chapter 8.1 gives a thorough introduction to the Gjøa field and our understanding of how production optimization and uncertainty is treated there today. Further, Chapter 8.2 provides a general insight to the properties of the given production data.

8.1 The Gjøa Field

The Gjøa field is located in the Norwegian North Sea, 60 kilometres west of the mainland city Florø. The field was discovered in 1989 by Norsk Hydro and development started in 2007 (GDF SUEZ E&P Norge, 2015). Statoil was responsible for the development phase, while ENGIE took over operations when production began in 2010. Several companies also hold licences, those with the highest stakes being Petoro and Wintershall (Wintershall Holding GmbH, 2015). The Gjøa field is estimated to have remaining recoverable reserves of around 22.5 billion cubic meters of gas and 3.3 million cubic metres of oil (Norwegian Petroleum Directorate, 2015c).

The field is divided into 7 sub-structures which lie in one of three areas: Gjøa North, Gjøa East, and Gjøa South. With a water depth of around 360 metres, the main reservoir of the field is the Viking group sandstone (Jørgensen, 2008). As shown in Figure 8.1.1, Gjøa is linked with the Vega and Vega South developments, with gas routed via Gjøa to land. Also illustrated, Gjøa receives part of its electricity from land, a measure to reduce emissions on the platform.

The field consists of a large gas cap above a thin oil rim (10-15 meters). Extraction is achieved by natural pressure reduction, and gas-lift is mainly used when bringing wells on-line (Statoil ASA, 2006). Therefore, the main control inputs managed by the production engineers are the choke valve openings.

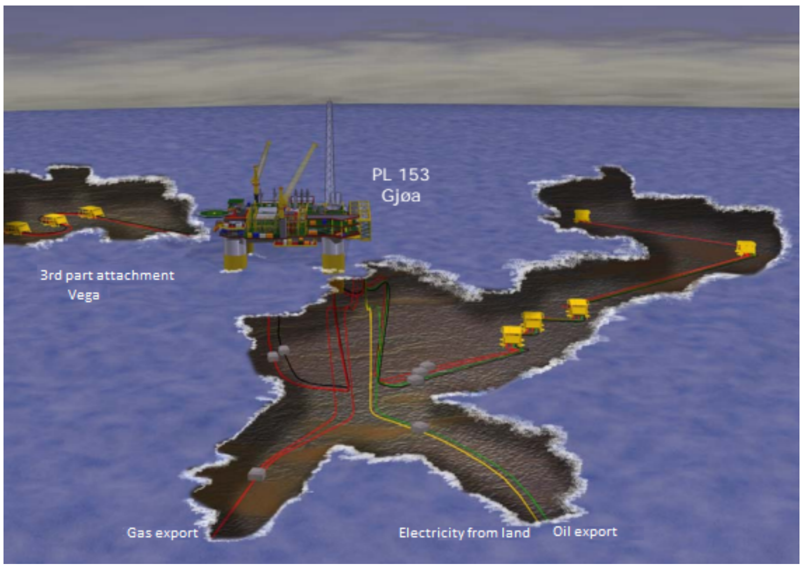


Figure 8.1.1: The Gjøa field (Statoil ASA, 2006)

8.1.1 Production System and Topology

The production system at the Gjøa field consists of a semi-submersible production and processing facility. Subsea there are two risers, one for oil producing wells and one for wells producing mostly gas. The gas riser has four wells connected and the oil riser has seven wells connected. Topside, different separators are used for the oil and gas riser. An illustration is presented in Figure 8.1.2.

Since the wells share either of two pipelines to the surface and the separators, the production from one well affects the others on the same pipeline. The production from the Gjøa field comes from pressure depletion, and more flow from one well decreases the pressure drop through the wellhead for the other wells. The assumption behind the well models developed for independent wells does then not appropriately represent the system dynamics and is therefore an approximation.

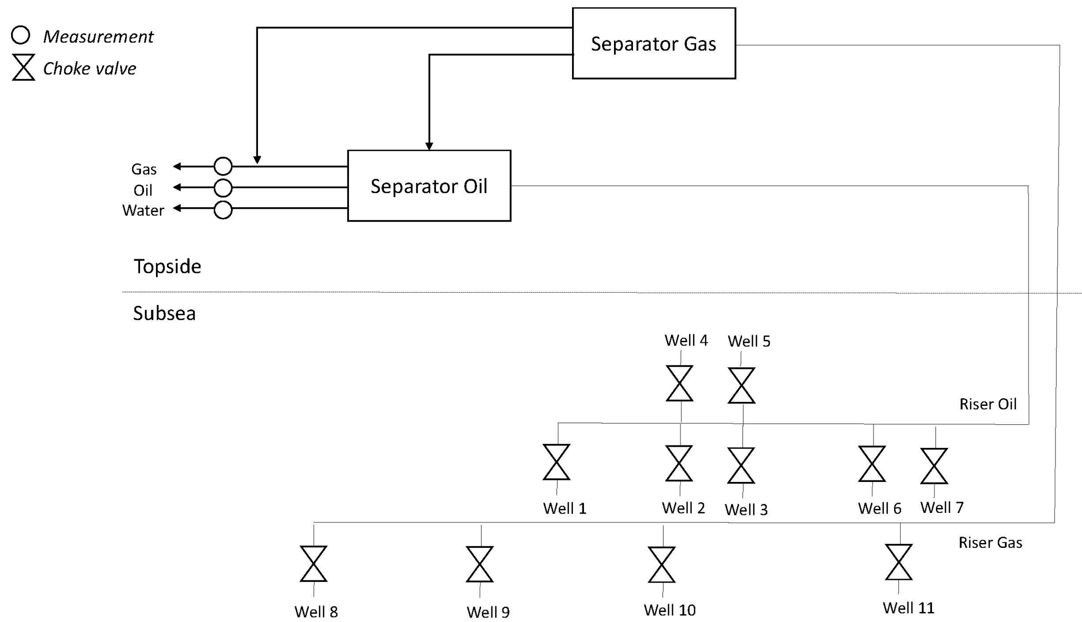


Figure 8.1.2: Gjøa topology (IO Center, n.d.; Statoil ASA, 2006)

Regarding measuring equipment, the Gjøa field has multiphase meters (MPM) installed on each individual well to give production operators and management an indication of the performance of each well. This equipment produces somewhat uncertain measurements as opposed to the export measurements for the total production, which are much more accurate.

8.1.2 Production Optimization at Gjøa

The production planning process at Gjøa today mainly follows the processes described in Chapter 2.3. Advanced optimization software is however not in use. The production engineers perform tests for each well by adjusting the choke openings for the wells and use this data to estimate which adjustments that should be performed. Excel and other basic optimization tools are used to support the decisions made. Historically, the production engineers have performed more extensive analysis and recognize increased oil production when doing so. However, the time and effort used to conduct such analysis have far exceeded the return (Solution Seeker AS with Vidar Gunnerud, 2015).

The real-time optimization problem at Gjøa is concerned with maximizing total oil production from the 11 wells in the system. There are no other capacity constraints than for gas handling. The seven oil wells (well 1-7 in Figure 8.1.2) have their own gas capacity constraint and can therefore be handled as a separate optimization problem. The four gas wells (well 8-11 in Figure 8.1.2) are part of a second optimization problem. Two of these wells lie in a separate gas reservoir, which the production and reservoir engineers wish to empty as fast as possible. The chokes of these wells are therefore fully open. Based on the remaining capacity, the second optimization problem is to maximize oil production from the last two gas wells.

In this thesis we are only concerned with the optimization of the seven wells using the oil riser, as this is the optimization problem with the highest potential for improvement. As explained earlier, the gas limit is not a hard constraint, but rather a limit the average gas production should lie below to satisfy other objectives. The gas limit at a certain time could then be set by the reservoir engineer, to ensure long-term optimal production, or for instance be reduced due to low gas prices.

8.2 Data Analysis

The historical production data used in this thesis consists of a three month extraction of raw data samples averaged for every 5 minutes. This very large raw data set contains information on choke opening for each well together with an estimate of each well's production of oil, gas, and water. These estimates for each well are provided using MPM measurements and are associated with a greater uncertainty. Total export production for the wells measured topside is also available. These measurements are more certain than the MPM measurements.

8.2.1 Pre-processing of Data

Some modifications are done to the raw data set to be able to efficiently analyse the data. Chokes measured to have a negative value are set to zero. The same procedure is applied to all negative numbers, as neither oil, gas, or water production physically can have a negative value. Furthermore, in some time intervals one or more of the wells have bad non-numeric data. To be sure that these periods are not affecting further analysis and regression, all data points for all wells are removed for those periods. Considering time intervals, this give a small area of discontinuity in the data set. This disadvantage is however considered less important than analysing an incomplete and faulty data set. Additionally, chokes with an opening less than 1% are assumed to be closed because of irregular production data in that region. In some periods the MPMs have not worked properly, resulting in an open choke, but no values for output production data. In these cases, choke value is set to zero. Furthermore, additional manual evaluations are performed to remove transient periods after system changes and other irrelevant or incorrect data.

8.2.2 Observed Data Characteristics

To find patterns and irregularities in the data sets, several plots are made where different parameters are combined. One of the main assumptions for applying the developed well models, is that a local linear relationship is present for parts of the data set. Solely based on the plots, this assumption seems to hold. However, there are large variations from well to well. An illustration of this is shown in Figure 8.2.1, where the well on the left seem to have a linear relationship between oil production and choke opening within a small interval, while the well on the right has a less clear relationship. Similar observations are made for gas production versus choke opening.

Correlation between oil and gas output for each choke opening seems very hard to find for most wells, while others again have a clearer relationship. An illustration for this is shown in Figure 8.2.2. For the right well, a large excitation in the data is observed, indicating that it will be hard to estimate the relationship between oil and gas production for a specific choke opening. The left well again indicates a somewhat linear relationship.

Another important observation is that most chokes are operated within a small interval or are shut off. This is shown in Figure 8.2.3. This gives limitations concerning which areas it is reasonable to assume that parameters found in the regression are valid. The left well has had some adjustments the last three months, while the right well has only been adjusted a few times and within a very small choke opening interval. With few adjustments, it is harder to find a correlation between choke opening and oil and gas production, as random variations can be assigned to the few changes in the chokes. To build a representative data set for a larger interval of choke openings, each well has to be adjusted more to test production output of oil and gas for different choke opening combinations.

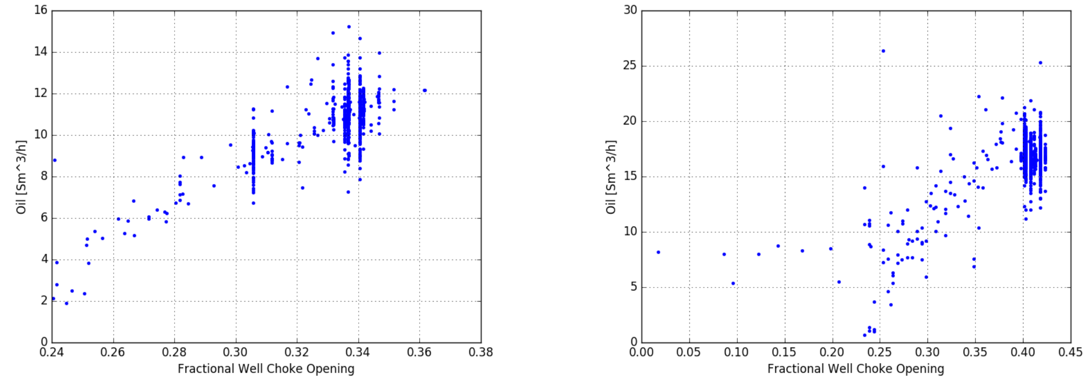


Figure 8.2.1: Oil production for different choke openings

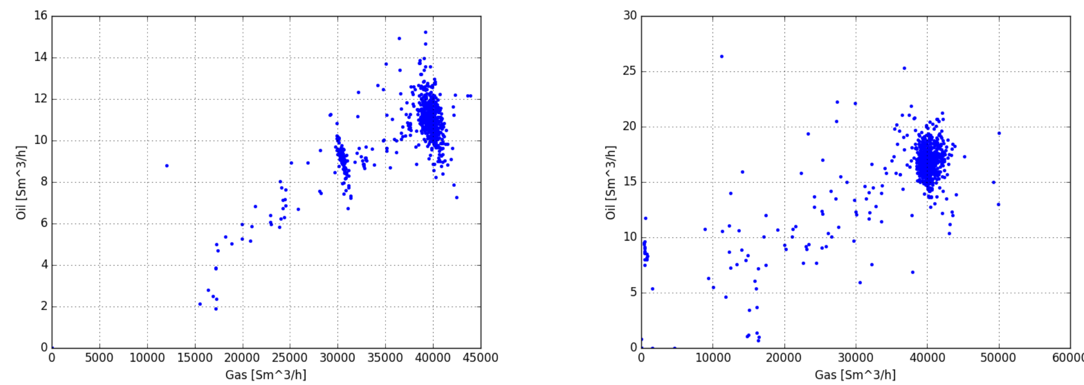


Figure 8.2.2: Relationship between oil and gas production

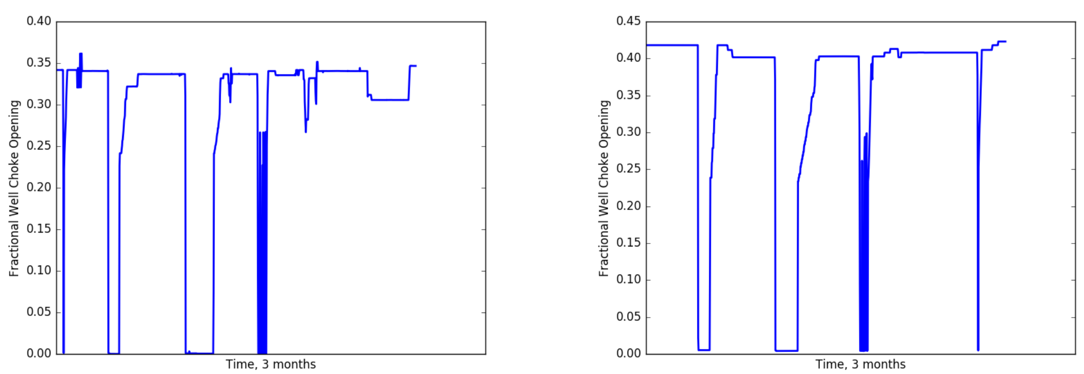


Figure 8.2.3: Choke variations over time

Chapter 9

Implementation and Well Model Analysis

This chapter starts with an explanation of the computer implementation procedure used for both well and optimization models in Chapter 9.1. Thereafter, the well model analysis is presented in Chapter 9.2, where evaluations and the results of the different regression methods are given. Chapter 9.3 then presents the problem instances that are used in the following optimization models and computational study.

An important note is that even though we are only interested in optimizing the seven oil wells, all eleven wells are included in the regression analysis. This is due to the fact that the available total export measurements is the sum of the oil and gas from both risers/separators. We however only show and evaluate the results for the relevant seven wells.

9.1 Implementation

The data sets used are firstly pre-processed in Microsoft Excel. Additionally, the data is set up in a format easily readable for most programming languages. Most of this procedure could easily be automatized if needed.

From Excel, the data sets are read into a script written in Python. As shown in Figure 9.1.1 this script is used for (1) data pre-processing and compression, (2) building regression models of the wells, (3) bootstrapping, (4) scenario generation, and (5) final analysis and writing of an input-file for the optimization model. Additionally, the script may be used to auto-invoke the solving of the optimization problem. However, this can also be done manually. The Python-language was chosen for these four tasks mainly due to its availability of good scientific and data-processing packages. The script is made fairly generic, such that new data may be easily generated if the data set is updated or exchanged with a completely different data set (e.g. data from a different petroleum field).

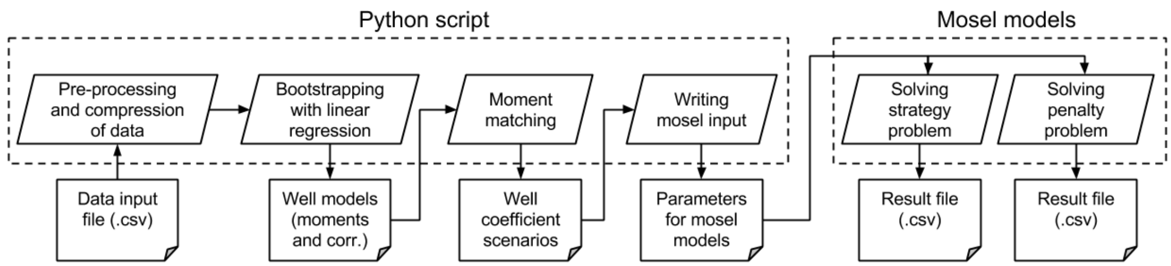


Figure 9.1.1: Processes in the implemented system

(1) Data pre-processing and compression

The script starts by converting the data set to a format beneficial for matrix regression. Thereafter, the data is run through a data compression algorithm.

The raw data set contains lots of data points that are very similar due to few changes to the production system in some time periods. There are two reasons why this data should be compressed in our case. Firstly, since computational time is important, compressing the data set to a fraction of its original size will help regression, cross-validation, bootstrapping and other data computations to run faster. Secondly, since we are interested in finding a regression line that best describes oil and gas output for the largest possible range of choke openings, lots of data points for a single choke opening will distort the regression result. This is because the values of the dependent variable usually varies much for some particular choke openings. With much variation and many data points, the regression line is forced close to the mean of these values, even though this may not be the overall trend across the complete range of choke openings. Therefore, compressing heavily populated intervals to a single or a few points, gives better results. The two plots in Figure 9.1.2 and Figure 9.1.3 show the difference between the two regressions lines found with and without data compression for one gas well, clearly indicating that data compression is beneficial. Similar results are found for all other wells for both oil and gas. For the interested reader, these are shown in Figure B.0.1, Figure B.0.2, Figure B.0.3 and Figure B.0.4 in Appendix B.

The data compression algorithm simply averages two rows in the data set if the average difference in choke opening is less than some threshold. Mathematically, if x is the choke opening, i is the well number and j is the row number in the data set, for each j two rows of data will be averaged and merged if

$$\sqrt{\frac{\sum_{i=1}^{n}(x_{ij} - x_{i(j+1)})^2}{n}} < \theta,$$

where θ is the chosen threshold. In our implementation $\theta = 0.01$.

An alternative way to compress the data set could be to split the range of possible choke openings into many small intervals, and for each interval pick the *newest* value for the compressed data set. This method could be better if the well models change significantly over time, making old data invalid. However, we have not tested this method in our thesis.

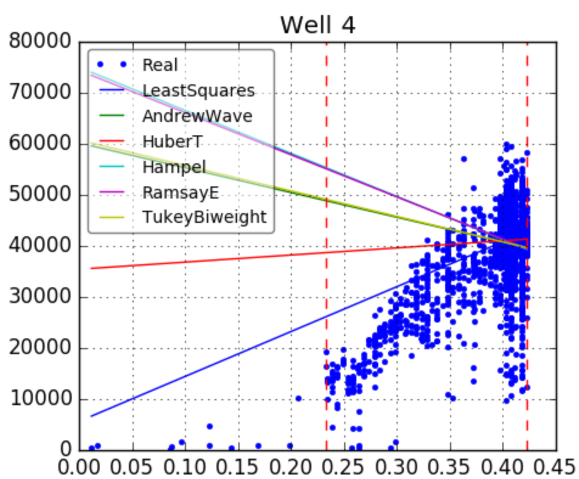


Figure 9.1.2: Regression, no compression

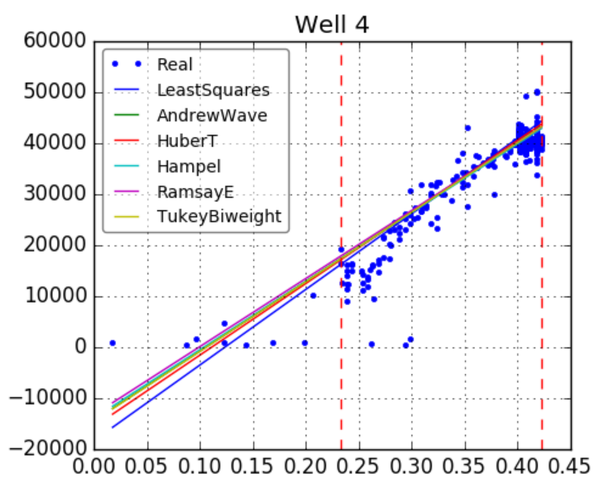


Figure 9.1.3: Regression with compression

(2) Building regression models of the wells

To build the well models, the Python package `statsmodels` is used. This package include implementations of all of the most common regression techniques. We use robust linear regression methods from this package to build models of oil and gas changes with respect to choke opening setups.

Ten-fold cross-validation is used for validating the regression models. This model validation technique is used to measure how well the models will generalize to independent data sets (Kohavi, 1995). It starts by partitioning the data into 10 random subsets.

Then, in each iteration, 9 of the sets are used for training the model, and 1 set is used for evaluating the trained model. This procedure is repeated 10 times, until all subsets have been used as test sets. In this way the training sets and the testing sets are independent, and all data is used for both training and testing. By using ten-fold cross validation, we reduce the chance of the chosen model being overfitted, that is a model that describes random errors instead of the relationship between the variables. If evaluation is done on the same set as the model was trained on, the error in the prediction will generally be low. However, this does not mean that the error is low when the same model is evaluated on an independent data set. As a result, these error metrics are biased, and not suitable for choosing a model that will generalize well. Ten-fold cross-validation is therefore chosen to avoid this problem and to help find the model that gives the best predictions on new data.

(3) Bootstrapping

The best robust regression method found in step (2), for building the oil model and gas model respectively, are used in the bootstrapping procedure. Residual resampling as described in Chapter 4.4 is used to estimate statistics for each of the regression coefficients (both the constants and the slopes). The statistics that are estimated are the first four moments of the distribution for each coefficient, as well as the correlations between all coefficients. Residual resampling is chosen with an assumption that the error variance is not dependent on x. Additonally, compared to case resampling it gives narrower and stronger bounds if the model assumptions are correct. 10000 bootstrap samples are taken in the simulation to estimate these statistics.

(4) Scenario generation

The implemented scenario generation algorithm is *moment matching*, as described in Chapter 4.5. The statistics found through the bootstrapping in step (3) are used as input data for the moment matching procedure. Further, the number of scenarios to be generated must be specified. If this number is too low, the moment matching procedure will fail as the specified statistics cannot be matched with very few scenarios. The amount of scenarios needed depends on the problem characteristics.

(5) Final analysis and writing an input-file for the optimization model

Further, the script is used to analyze the imported data set to find which combination of modes that has been seen in the past. This is done, as we assume the regression models are only valid within the modes observed in the data set and the modes that can be calculated from these (i.e. predicting the value of a setup far from anything observed will probably not work).

Additionally, the minimal (greater than 0) and maximal legal opening is computed for each choke. This is the range in which the choke may be adjusted if it is "on" in the operating mode considered. Algorithm 7 is used to find the lower limit of a choke. In short, this algorithm finds the first point in the data set that has many other points in close proximity. The data set will be densely populated from this point until the upper limit. The upper limit is found using the same algorithm by iterating from the other direction, i.e from N to 2. Also, when finding the upper limit the return value will be chokes[i+min_density-1], rather than chokes[i-min_density]. The limits found using this algorithm are illustrated in Figure 9.1.2 and Figure 9.1.3 as red dashed lines. For up-adjustments we additionally allow the choke to be adjusted up to 10% above the upper red dashed limit, as we trust that the linear model found will be valid in this area.

Algorithm 7: Finding the lower limit of a choke.

Data: A sorted set of N choke opening values for a particular well, named chokes.

Result: The lower limit for a particular choke.

```
density ← 0;
for i ← 2 to N do
    if chokes[i] == 0 then
        continue to next iteration;
    if chokes[i] - chokes[i-1] < ε then
        increase density by 1;
    else
        density ← 0;
    if density == minimum_density then
        break the loop;
return chokes[i-minimum_density].
```

The generated scenarios, list of allowable modes, choke limits, as well as a set of user defined parameters are then written to a text file.

(6) Optimization

The optimization models are solved in Mosel Xpress-MP. The text file generated from the Python script is taken as input. The results from Mosel Xpress-MP are written to

.csv-files and final analysis is done manually or with specialized Python scripts. Mosel Xpress-MP is run on a Hewlett Packard 64-bit Windows 7 Enterprise PC with Intel(R) Core(TM) i7-3770 3.40 GHz processor and 16,0 GB (15.9 GB usable) RAM.

9.2 Well Model Analysis

In this section the three well models presented in Chapter 6 are evaluated using different regression methods. Both ordinary least squares (OLS) and robust regression methods are applied to the oil and gas production data. The results of this are presented in Chapter 9.2.1. Using the results from this analysis, we present the final well parameters to use in the optimization in Chapter 9.2.2.

9.2.1 Well Model Regression Results

In the following, the results from the regression methods for the export, multiphase and combined well model are presented. The section is ended with a comparison of the well models and their results.

Evaluation Method

To evaluate and compare the quality of the well models proposed, defined metrics are necessary. Two metrics are used for this purpose. *Metric 1* is the average error between the prediction model and the data points for each method (MAD), while *Metric 2* is the root mean squared error (RMSE) for the regressions. Since the model preferred should have a low average error and minimize large outlier errors, a combined evaluation of both metrics decides which regression method to use.

The predictions from all well models are compared using the real export measurement for evaluation, as this measurement has low uncertainty. An alternative is to use the sum of the actual MPM measurements. However, the export measurement is assumed to be closest to the exact amount of oil and gas produced from the field, and is therefore preferred for evaluation.

Export Well Model

In Table 9.2.1, the results from the regressions based on the export measurements are given. An overall evaluation of the results using metric 1 and metric 2 show that Huber's T in total is the best regression method both for oil and gas. This is because it has the lowest combined error considering the two metrics. The difference between the metrics is however very small for all regression methods, indicating that more than one method can be used to obtain almost the same results.

The export oil measurements tend to oscillate more than the gas measurements and appears to have more outliers lying above or below the mean. The outliers of the oil production measurements are weighted very low by Huber's T, since its weighting function decreases very steeply outside a defined area around the mean. Points that are not considered outliers are weighted equally much. This sharp exclusion of outliers while still including an appropriate range of data around the mean might provide an explanation for why this method comes out best.

For the more certain gas measurements, it is still reasonable to assign equal value to points close to the mean as done by Huber's T. This argument is supported when considering OLS for metric 2, as this equal weighting method performs even better than Huber's T for those measurements.

Table 9.2.1: Regression results for the export well model

Export Well Model				
	Metric 1 (MAD)		Metric 2 (RMSE)	
	Oil	Gas [10^3]	Oil	Gas [10^3]
OLS	11.44	14.78	20.39	22.28
Andrews' Wave	10.84	14.11	21.90	23.15
Huber's T	10.77	14.01	20.85	22.54
Hampel's 17A	10.80	14.05	21.60	22.51
Ramsay's E	10.74	13.97	21.21	22.58
Tukey's Biweight	10.84	14.17	21.90	23.25

Multiphase Meter (MPM) Well Model

In Table 9.2.2, the results from the different regression methods for the MPM well model are presented. For this model, Ramsay's E is preferred for oil measurements, while OLS is preferred for gas prediction according to both metrics.

For oil we see that except from OLS, all robust regression methods provide about the same results, with Ramsay's E being slightly better in total concerning both metric 1 and 2. OLS provide poor results because it assigns a large weight to the large amount of outliers that is present in the oil multiphase measurements. In this case, all the robust regression methods successfully eliminate the impact of these outliers.

Gas measurements have fewer outliers and are more certain than oil measurements when using multiphase meters as well. It is therefore reasonable to assume that most of the data points should be included with full weight in the regression. OLS therefore performs well, and so does Huber's T which has the same properties as OLS in a range around the mean. Even though Hampel's 17A has a similar weighting as Huber's T it performs worse, since the area around the mean assigning equal weight to data measurements is smaller than for Huber's T. By this, the method excludes points that should be included in the regression. This also support the claim that there are few outliers, as we see that all data points should be included with equal weight to obtain the best regression results.

Table 9.2.2: Regression results for the multiphase meter well model

Multiphase Well Model				
	Metric 1 (MAD)		Metric 2 (RMSE)	
	Oil	Gas [10^3]	Oil	Gas [10^3]
OLS	14.84	14.09	24.13	23.19
Andrews' Wave	11.75	17.21	21.94	28.35
Huber's T	11.29	14.12	21.52	23.63
Hampel's 17A	11.40	15.51	21.75	26.28
Ramsay's E	11.23	15.35	21.53	25.89
Tukey's Biweight	11.75	17.25	21.94	28.45

Combined Well Model

The results from the regression methods for the combined well model are shown in Table 9.2.3. This well model includes both export measurements and multiphase measurements. A combined evaluation finds that Huber's T again is the preferred regression method. As described above, it in the same way as OLS assign full weight to all data points within a certain range of the mean and a decreasing weight for all points further away. In this way it may be able to capture the information in the export measurements with low variance while limiting the effect of the outliers of the MPM measurements.

For the gas measurements, the regression methods Andrews's Wave, Hampel's 17A, and Tukey's Biweight, perform considerably worse than the other methods. That is due to the decreasing weight the function gives to data very close to the mean. This again indicate that gas measurements contain fewer outliers and that all points should be evaluated as equal in the regression for best prediction.

Table 9.2.3: Regression results for the combined well model

Combined Well Model				
	Metric 1 (MAD)		Metric 2 (RMSE)	
	Oil	Gas [10^3]	Oil	Gas [10^3]
OLS	11.58	12.80	21.41	22.44
Andrews' Wave	11.15	14.68	21.69	25.70
Huber's T	10.88	12.50	21.54	22.60
Hampel's 17A	11.00	14.27	21.73	25.40
Ramsay's E	10.93	12.97	21.72	23.60
Tukey's Biweight	11.15	14.76	21.69	25.70

Figure 9.2.1 and Figure 9.2.2, are excerpts from plots showing the combined model prediction of oil and gas production using different regression methods. As can be seen, most regression methods follow the general variations of the real output well. For oil prediction, the outlier extremes are smoothly eliminated from the predicted values as expected. For gas measurements, an even better prediction is given due to the lower variance in the data. Full plots of predicted production of oil and gas for all available data are presented in Figure C.0.1 and Figure C.0.2 in Appendix C.

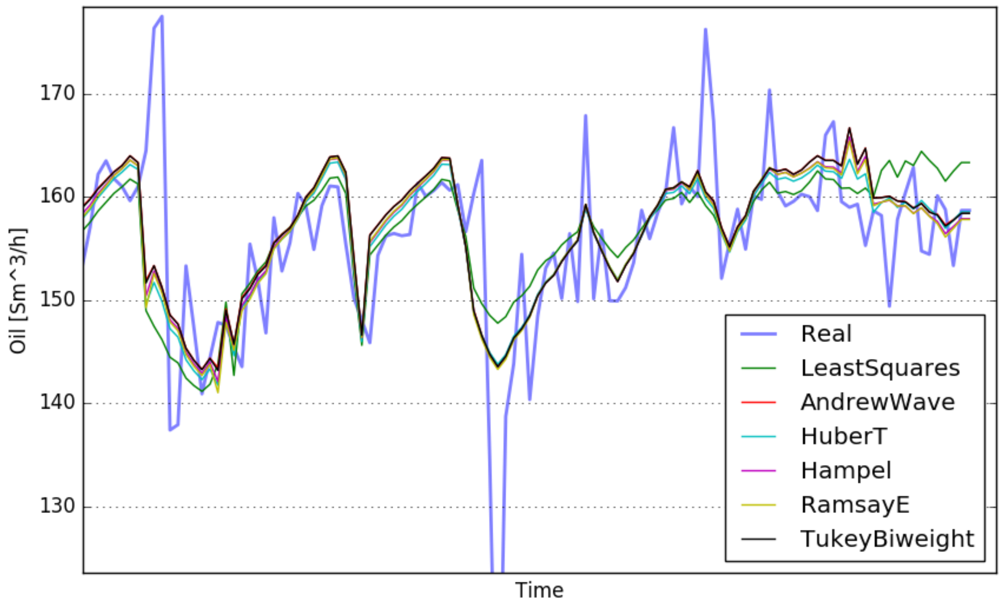


Figure 9.2.1: Plot of regression methods for oil production, combined well model

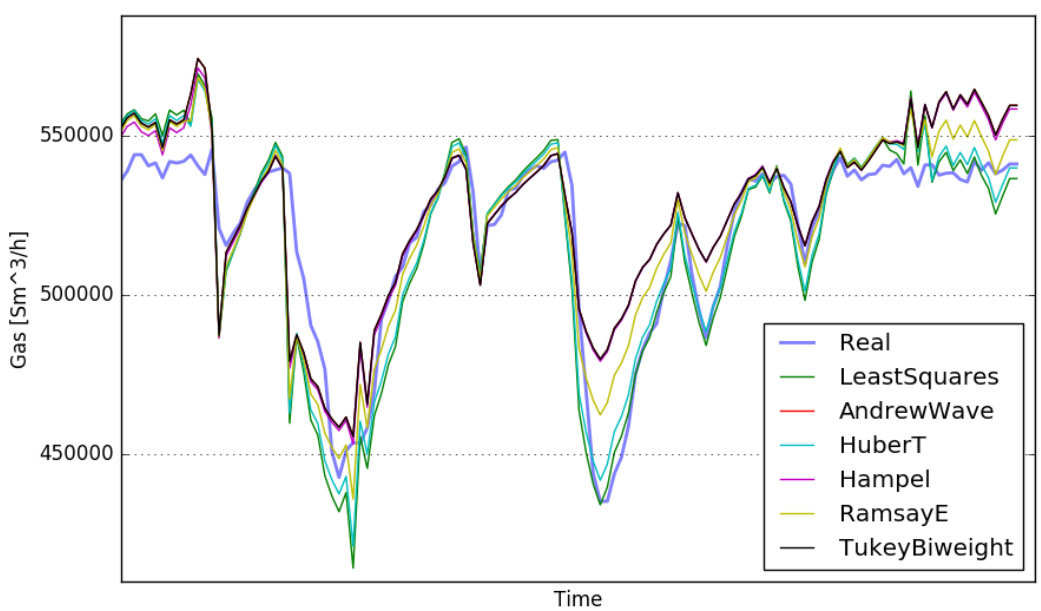


Figure 9.2.2: Plot of regression methods for gas production, combined well model

Comparison of Well Models

In deciding which well model and regression method to use to generate parameters for the optimization, we use a combined evaluation of metric 1 and 2. For the metric evaluations, we believe it is preferable to have a small average error in prediction as measured by metric 1. In addition, a low value for metric 2 is important as some large errors in prediction could have a strong impact on the result of the optimization.

Since the evaluations of all well models and regression methods are performed on the same data sets, we can directly compare the values of the different metrics across the tables presented. In general, we observe that the multiphase meter well model give the worst results. This model is only based on the measurements from the MPMs, which we know are very uncertain. An important point is also that some MPMs are more uncertain than others. The MPM model may therefore generate fair parameters for some wells, and fail completely for others. As the model does not apply the export measurements, it does not include any measurements where the wells are coupled together. Therefore, it cannot correct the parameters for the wells where the MPMs provide bad estimates.

The export model on the other hand is based on nearly certain data, and therefore provide good results when predicting total oil and gas output. It however attempts to estimate many parameters on fewer data points. It therefore fails to provide good predictions for each individual well.

An interesting finding is the good performance of the combined model. Compared to the export model, it performs better for predicting gas and about equal for oil using metric 1, and slightly worse using metric 2. An illustration of this is provided for well 1 in Figure 9.2.3 and Figure 9.2.4. We see that the predictions using combined data has a significantly better model fit than using export data.

Although the MPM measurements are uncertain and result in bad predictions when used alone, the combined model shows that it is possible to extract valuable information from them. We have shown here that they can improve some predictions, but that it is dangerous to apply their measurements without accounting for the export measurements.

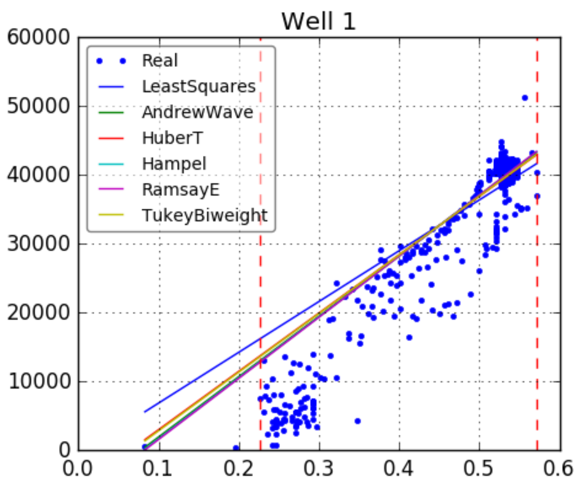


Figure 9.2.3: Regression on export data

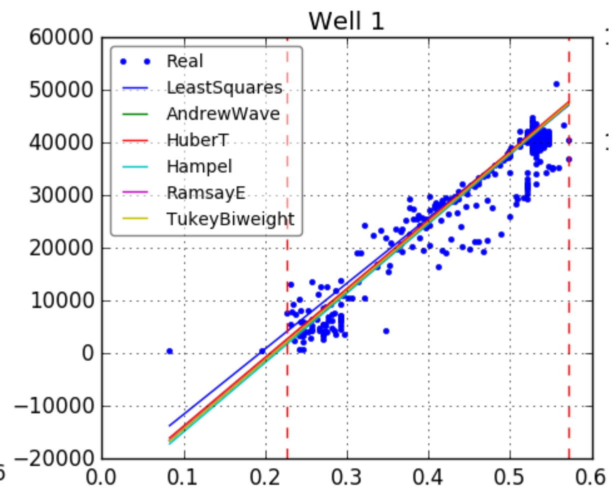


Figure 9.2.4: Regression on combined data

9.2.2 Final Well Parameters

Based on the good results of the combined model when predicting total production well, and the fact that the export model does not provide as good predictions for each individual well, we conclude that the combined model using Huber's T is applied to find well parameters. As described in Chapter 6, the combined well model allows for all possible operating modes by simply subtracting the value of the constant from total production when a well is closed. The distributions of the constants and slopes is obtained through the bootstrapping procedure residual resampling as described in Chapter 4.2.

Table 9.2.4 presents the oil parameter slopes and constants, and their moments. We observe that all wells have positive slopes indicating that a larger choke opening implies a larger oil output for each well. There are however large differences in the slope means. In short this implies that some wells with high slope coefficients will give considerably more (less) oil for the same choke increase (decrease). We also observe that the uncertainty of the wells vary when considering variance. Skewness and kurtosis are for some wells significant, but for most wells very small. Normalized skewness and kurtosis is given for clarity. '0' in normalized skewness and '3' for normalized kurtosis corresponds to the standard normal distribution. In total, the wells hold different properties, providing a good foundation for further analysis concerning which wells to adjust to improve production.

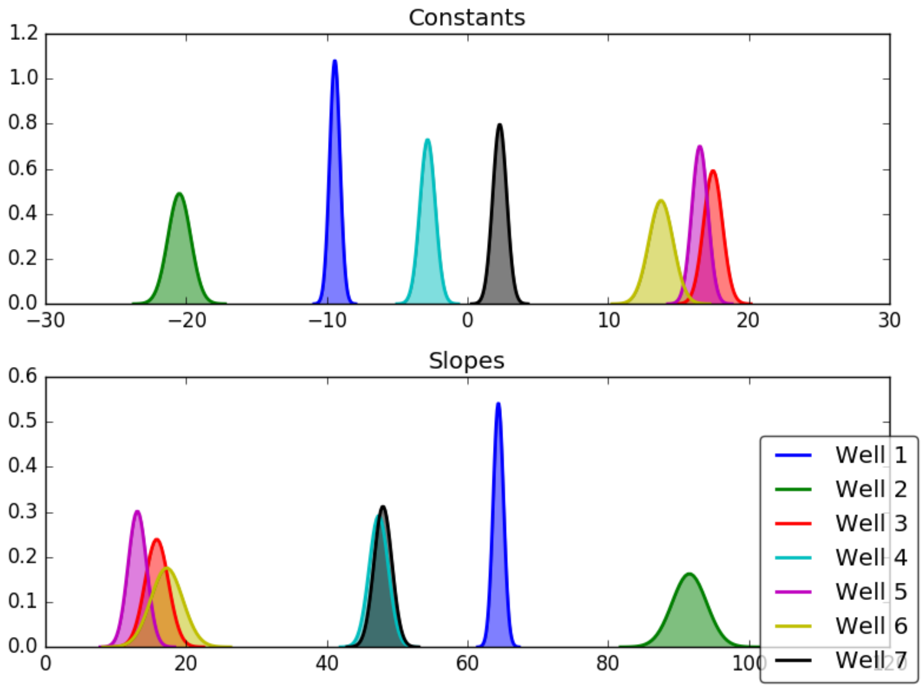
Table 9.2.5 presents the gas parameter slopes and constants, and their moments. Again all slopes are positive, suggesting more gas for a larger choke opening. Here, in contrary to the oil slope means, the relative difference between the gas slope means are much smaller. The variance for gas is also less considerable than for oil, which is expected due to its more certain measurements. Skewness and kurtosis is again very small for most wells.

Table 9.2.4: Oil parameters for Gjøa wells

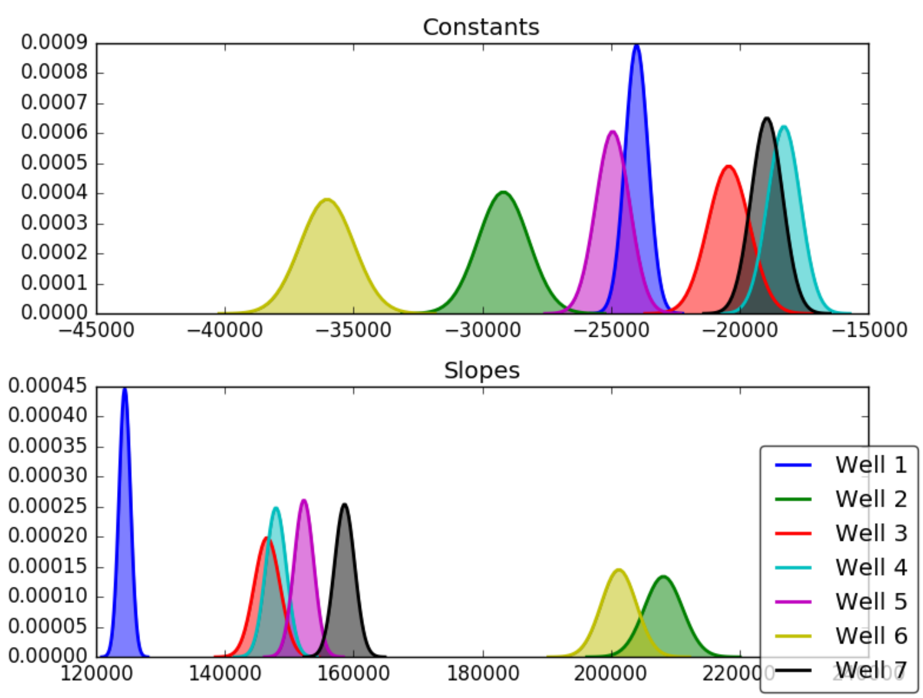
	Wells	**Mean**	**Variance**	**Skewness** Actual	**Skewness** Normalized	**Kurtosis** Actual	**Kurtosis** Normalized
Constants	1	-9.43	0.14	0.00	0.02	0.06	2.99
	2	-20.49	0.66	-0.02	-0.05	1.4	3.22
	3	17.45	0.45	-0.01	-0.03	0.62	3.03
	4	-2.84	0.30	0.00	-0.02	0.27	3.04
	5	16.51	0.32	0.00	0.01	0.31	3.00
	6	13.75	0.75	-0.01	-0.02	1.76	3.10
	7	2.29	0.25	0.00	-0.02	0.19	3.06
Slopes	1	64.39	0.54	-0.01	-0.02	0.88	3.00
	2	91.55	6.00	0.71	0.05	116.24	3.22
	3	15.81	2.78	0.15	0.03	23.28	3.02
	4	47.37	1.87	0.03	0.01	10.49	3.01
	5	13.05	1.75	-0.04	-0.02	9.14	3.00
	6	17.23	5.16	0.21	0.02	82.65	3.10
	7	47.97	1.63	0.02	0.01	8.14	3.07

Table 9.2.5: Gas parameters for Gjøa wells

	Wells	**Mean** [10^4]	**Variance** [10^5]	**Skewness** Actual [10^8]	**Skewness** Normalized	**Kurtosis** Actual [10^{13}]	**Kurtosis** Normalized
Constants	1	-2.40	1.99	-0.04	-0.05	0.01	2.90
	2	-2.92	9.69	-1.07	-0.11	0.30	3.18
	3	-2.04	6.59	-0.36	-0.07	0.13	3.01
	4	-1.83	4.09	-0.22	-0.09	0.05	3.05
	5	-2.49	4.32	-0.33	-0.11	0.06	2.98
	6	-3.60	11.00	-0.47	-0.04	0.37	3.07
	7	-1.89	3.75	-0.03	-0.02	0.04	3.05
Slopes	1	12.50	7.98	0.19	0.03	0.19	2.91
	2	20.80	88.50	30.40	0.12	24.90	3.19
	3	14.70	40.30	5.72	0.07	4.88	3.01
	4	14.80	25.80	3.27	0.08	2.03	3.05
	5	15.20	23.30	3.74	0.10	1.62	2.98
	6	20.10	75.10	8.20	0.04	17.30	3.07
	7	15.90	24.60	0.40	0.01	1.85	3.06



(a) Oil



(b) Gas

Figure 9.2.5: Distributions for constants and slopes

As a general observation for both oil and gas parameters, we see that variance around the mean for some wells might affect the calculations when optimizing. Parameters for skewness and kurtosis indicate that the probability distributions around the mean are very close to the properties of a normal distribution. For the interested reader, plots showing the regression line for gas and oil for each well is illustrated in Figure B.0.2 and Figure B.0.4 in Appendix B. An important note is that the regression lines are calculated not only based on the plotted data, but also using total export measurements. The regression lines therefore might seem shifted from what could be expected by visual inspection of the plotted data points. It should nonetheless provide a good indication of how the lines relate to the data. The distributions found for the constants and slopes of each well are illustrated in Figure 9.2.5 (a) for oil and (b) for gas. Plots showing the distributions for each well in individual plots are illustrated in Figure D.0.1 for gas, and Figure D.0.2 for oil respectively in Appendix D. Distributions very close to the normal distribution are observed for all wells.

In Table 9.2.6 we list the gas-to-oil ratios for each Gjøa well. The lowest GOR is observed for well 1 with 1941, and the highest for well 6 with 11667. This wide range of GORs provides an interesting insight into each well's properties, and is interesting to follow in the further analysis.

Table 9.2.6: Gas-to-oil ratio for each Gjøa well

Well	1	2	3	4	5	6	7
GOR	1941	2272	9298	3124	11648	11667	3315

Well Model Linearity

The coefficient of determination (R^2) is calculated for the linear equation of each well and for the total export prediction to test well model linearity. Note that R^2 is not an optimal indicator, since it is developed for ordinary least squares. We are nevertheless confident that it will provide a good indication even though the more robust Huber's T is used for regression. The R^2-values are presented in Table 9.2.7. As a general trend, it is observed that gas values are much better predicted than oil values. Considering previous analysis, this is not surprising.

For oil, very low R^2-values are found for well 3, 5, and 6, indicating that hardly no linear trend can be expected. For the other wells a small, but still significant relationship is observed. For the export prediction, 66% of the variance is explained using the found well model parameters. This provides an indication that using a linear model is a decent assumption.

For gas, high R^2-values are found for all wells. A linear model is shown to give a very good indication of what can be expected from a future prediction. For the export prediction, 95% of the variance is explained. This is a significant finding, showing that a linear assumption is a good assumption for these measurements.

Table 9.2.7: Coefficient of determination [R^2] for each Gjøa well

Well	1	2	3	4	5	6	7	Export
Oil	0.56	0.65	0.00	0.52	0.16	0.00	0.41	0.66
Gas	0.94	0.97	0.83	0.89	0.87	0.83	0.86	0.95

Well Model Correlation

From the bootstrap procedure, a correlation matrix between all slopes (A) and all other slopes, and between all slopes and all constants (B) for the different wells is given. Table 9.2.8 show the correlation matrix for the gas coefficients, while Table 9.2.9 show the correlation matrix for the oil coefficients. Obviously, the diagonal contains ones, as each parameter is perfectly correlated with itself. It is also observed that the diagonals between A_i and B_i, where i is the well number, has correlations close to or equal to minus one. This is explained by the fact that an increase in the slope of a well must lead to a decrease of the constant (and vice-versa), if the line still is to reasonably explain the data points it is modeled on. A slightly negative correlation trend is found in quadrant two and four; that is correlation between slopes or correlation between constants. Since the well models are built on total export measurements as well, this makes sense because an increase (decrease) in either the slope or constant of a well must lead to a decrease (increase) of one or more of the other wells' coefficients in order to not over- or underestimate the total production. A slightly positive correlation trend is observed in quadrant one and three. The slight positive correlation in the quadrants is explained by the fact that an increase in a constant for one well will reduce the slope coefficient of the same well (since it is perfectly correlated with itself), and that another well based on that must increase its slope to make up for the decrease of the first well.

All correlations between parameters for different wells are very small. The reason for this is that a change in one well's parameters cannot be directly described by one specific other well, but rather has to be explained by one or more of the other wells in combination. The correlation therefore might have an impact in the optimization even though we cannot conclude that a clear relationship exists.

Table 9.2.8: Correlation of gas coefficients, Gjøa case

Well	B1	B2	B3	B4	B5	B6	B7	A1	A2	A3	A4	A5	A6	A7
B1	1	−0.06	−0.09	−0.07	0.02	−0.07	−0.04	−0.98	0.06	0.09	0.07	−0.03	0.06	0.04
B2	−0.06	1	0.04	−0.13	0.01	0.02	0.06	0.06	−0.99	−0.04	0.12	−0.01	−0.02	−0.07
B3	−0.09	0.04	1	−0.05	−0.08	−0.09	0.03	0.08	−0.04	−0.99	0.05	0.07	0.09	−0.03
B4	−0.07	−0.13	−0.05	1	−0.07	0	−0.11	0.07	0.12	0.05	−0.99	0.07	0	0.11
B5	0.02	0.01	−0.08	−0.07	1	−0.08	−0.03	−0.03	−0.01	0.07	0.07	−0.99	0.07	0.03
B6	−0.07	0.02	−0.09	0	−0.08	1	−0.04	0.06	−0.01	0.09	0	0.08	−1	0.05
B7	−0.04	0.06	0.03	−0.11	−0.03	−0.04	1	0.03	−0.07	−0.04	0.1	0.02	0.04	−0.99
A1	−0.98	0.06	0.08	0.07	−0.03	0.06	0.03	1	−0.06	−0.08	−0.06	0.03	−0.06	−0.03
A2	0.06	−0.99	−0.04	0.12	−0.01	−0.01	−0.07	−0.06	1	0.04	−0.12	0.01	0.02	0.07
A3	0.09	−0.04	−0.99	0.05	0.07	0.09	−0.04	−0.08	0.04	1	−0.05	−0.07	−0.09	0.03
A4	0.07	0.12	0.05	−0.99	0.07	0	0.1	−0.06	−0.12	−0.05	1	−0.07	−0.01	−0.1
A5	−0.03	−0.01	0.07	0.07	−0.99	0.08	0.02	0.03	0.01	−0.07	−0.07	1	−0.07	−0.02
A6	0.06	−0.02	0.09	0	0.07	−1	0.04	−0.06	0.02	−0.09	−0.01	−0.07	1	−0.04
A7	0.04	−0.07	−0.03	0.11	0.03	0.05	−0.99	−0.03	0.07	0.03	−0.1	−0.02	−0.04	1

Table 9.2.9: Correlation of oil coefficients, Gjøa case

Well	B1	B2	B3	B4	B5	B6	B7	A1	A2	A3	A4	A5	A6	A7
B1	1	−0.05	−0.08	−0.07	0.03	−0.06	−0.06	−0.98	0.05	0.07	0.06	−0.03	0.05	0.06
B2	−0.05	1	0.05	−0.14	0.01	0	0.06	0.06	−0.99	−0.06	0.14	−0.01	0	−0.06
B3	−0.08	0.05	1	−0.06	−0.09	−0.1	0.04	0.07	−0.06	−0.99	0.06	0.09	0.09	−0.04
B4	−0.07	−0.14	−0.06	1	−0.07	0.04	−0.12	0.06	0.14	0.06	−0.99	0.06	−0.04	0.12
B5	0.03	0.01	−0.09	−0.07	1	−0.09	−0.04	−0.04	−0.01	0.09	0.06	−0.99	0.08	0.04
B6	−0.06	0	−0.1	0.04	−0.09	1	−0.03	0.06	0	0.1	−0.03	0.09	−1	0.04
B7	−0.06	0.06	0.04	−0.12	−0.04	−0.03	1	0.05	−0.06	−0.05	0.12	0.03	0.03	−0.99
A1	−0.98	0.06	0.07	0.06	−0.04	0.06	0.05	1	−0.06	−0.07	−0.06	0.04	−0.05	−0.04
A2	0.05	−0.99	−0.06	0.14	−0.01	0	−0.06	−0.06	1	0.06	−0.14	0.01	0	0.06
A3	0.07	−0.06	−0.99	0.06	0.09	0.1	−0.05	−0.07	0.06	1	−0.06	−0.08	−0.09	0.05
A4	0.06	0.14	0.06	−0.99	0.06	−0.03	0.12	−0.06	−0.14	−0.06	1	−0.06	0.03	−0.12
A5	−0.03	−0.01	0.09	0.06	−0.99	0.09	0.03	0.04	0.01	−0.08	−0.06	1	−0.08	−0.03
A6	0.05	0	0.09	−0.04	0.08	−1	0.03	−0.05	0	−0.09	0.03	−0.08	1	−0.03
A7	0.06	−0.06	−0.04	0.12	0.04	0.04	−0.99	−0.04	0.06	0.05	−0.12	−0.03	−0.03	1

9.3 Problem Instances

To evaluate the optimization models presented, they are tested on the Gjøa case which is presented in Chapter 9.3.1. To further study model robustness and determine the value of applying stochastic programming in different cases, we also test on four synthetic cases. These are presented in Chapter 9.3.2. The strategy model and the penalty model need the same input parameters for each individual well and we have therefore created one set of instances that is tested on both models.

9.3.1 The Gjøa Case

The parameters for oil and gas moments, correlations, and GOR-values for the Gjøa case are given in Chapter 9.2.2 above. The deterministic version of the Gjøa case has the same mean and contribution from each well, but no uncertainty attached to the parameters.

The Gjøa cases (both deterministic and stochastic) are tested with no spare gas capacity. This is the most relevant case as the gas production is generally balancing the gas capacity limit. However, we test these instances with spare capacity (positive delta gas) and excess production (negative delta gas) as well. Negative delta gas is tested to find which sequence an operator should down-adjust well chokes if the gas production limit is decreased due to for instance maintenance or changed market conditions as mentioned before. Negative delta gas is only tested for the strategy model, because the penalty model is per now not formulated in a way making this possible. Positive delta gas is similarly tested to find a suggestion for an up-adjustment sequence if new capacity is added. That could be due to new equipment or finished maintenance. Based on analysis of historical production data, we find -50000 Sm^3/h and $+10000$ Sm^3/h to be appropriate capacities to test with. A capacity change of -50000 Sm^3/h could for instance replicate maintenance overhaul of a separator, while $+10000$ Sm^3/h can replicate adding an improved valve or other equipment.

Each choke has a current opening and a limit for how much it is allowed to be adjusted. As described before, this is to insure that the assumption of a local linear relationship is still valid. The current choke openings, and the lower and upper limit for each choke are displayed in Table 9.3.1. The limits are found using the algorithm and description given under point 5 in Chapter 9.1. For the Gjøa case, both two and three allowed well choke changes are tested.

Table 9.3.1: Current fractional choke openings, and limits for well choke changes

	Well						
	1	2	3	4	5	6	7
Lower limit [$\underline{L}$]	0.2261	0.2214	0.2192	0.2336	0.2214	0.2367	0.2140
Current choke opening	0.5463	0.3466	0.4449	0.4229	0.4492	0.3963	0.3831
Upper limit [$\overline{L}$]	0.6714	0.4627	0.5548	0.5230	0.5592	0.5815	0.5935

Penalty Cost

The penalty model needs an additional parameter for the cost of violating the gas limit. This cost should reflect the lost production associated with the need to aggressively reduce a choke opening after a limit violation. In our model we have assumed that this cost increases linearly with the amount needed to reduce a choke after a breach of the limit. It is difficult to estimate the correct value of this cost, as it will vary between wells and be dependent on the topology between the wells at a the field. A cost of 10 are proposed and tested for the Gjøa case. The value of the cost is based on the slope of the oil contribution for the Gjøa wells. With a cost of 10, the penalty is in general lower than the slopes for the wells, possibly giving slightly optimistic objective values.

9.3.2 Synthetic Cases

To test different properties of the proposed optimization models, four synthetic cases are suggested. Each case is designed to provide a deeper understanding of how different combinations of parameters and levels of uncertainty affect the solution. Primarily, the cases are designed to illustrate interesting findings regarding the value of applying stochastic models and how the solutions differ between stochastic and deterministic models. Cases where wells for example have different GOR or where a good or bad well starts as being off is therefore not included, as results from these cases could easily be inferred from the results of the other tests. As with the real Gjøa case, the strategy cases are tested with delta gas being 0, +10000 Sm^3/h, and -50000 Sm^3/h. The penalty cases are tested with delta gas being 0 and +50000 Sm^3/h. All wells in the test cases have their initial choke opening set to 50%. The local linear range for each well is assumed to be within the range 50% $\pm$ 20%. As in the real case, up to two wells can be turned off. The constant term in the linear model is set to 0 for all wells. No correlation between the well parameters is assumed. In the penalty model, all the synthetic cases are again tested with a penalty cost of 10.

Synthetic case 1

The first synthetic case has uncertain oil production and certain gas production. All wells have the same mean oil production and mean gas production, and consequently the same GOR. The variance in oil production is however different for each of the wells. The distribution for oil production is not skewed, and has a normal kurtosis. Table 9.3.2 gives the parameters, while Figure 9.3.1 illustrates the distributions for the slopes coming from these parameters. Note that the x-axis shows the mean and three standard deviations, and that the axes have different scales. The red plotted line has the shape of a normal distribution.

Table 9.3.2: Synthetic case 1

Well	**Oil**				**Gas**			
	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)
1	50	16	0	3	50000	0	0	0
2	50	36	0	3	50000	0	0	0
3	50	64	0	3	50000	0	0	0
4	50	100	0	3	50000	0	0	0
5	50	144	0	3	50000	0	0	0

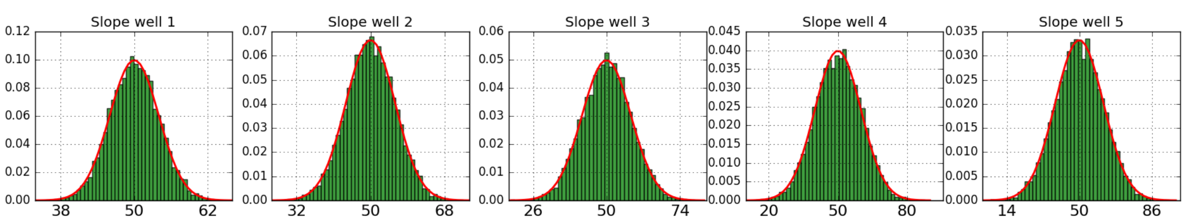


Figure 9.3.1: Case 1: Distribution for oil slope coefficient

Synthetic case 2

Similar to the first case, the second synthetic case also has uncertain oil production and certain gas production. All wells have the same mean oil production and mean gas production, and consequently the same GOR. In contrast to the first synthetic case, all wells have the same variance in oil production. However, the skewness of the oil production distribution is different for each of the wells. Table 9.3.3 gives the parameters, while Figure 9.3.2 illustrates the distributions for the slopes coming from these parameters.

Table 9.3.3: Synthetic case 2

Well	**Oil**				**Gas**			
	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)
1	50	25	-0.5	3	50000	0	0	0
2	50	25	-0.3	3	50000	0	0	0
3	50	25	0	3	50000	0	0	0
4	50	25	0.3	3	50000	0	0	0
5	50	25	0.5	3	50000	0	0	0

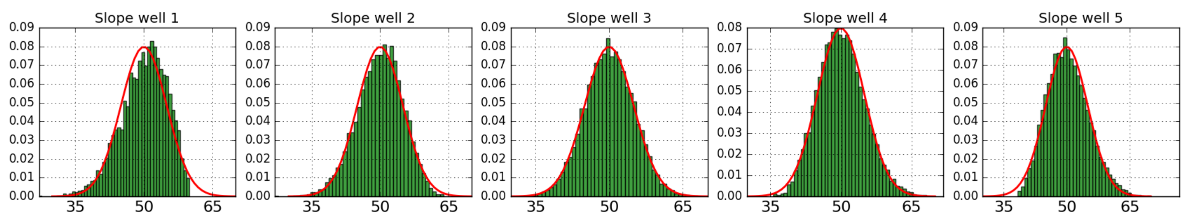


Figure 9.3.2: Case 2: Distribution for oil slope coefficient

Synthetic case 3

This case is similar to case 1, except that gas production is uncertain rather than oil production. This case has the same relative uncertainty of gas production, as the relative uncertainty of oil production in synthetic case 1 (i.e. the normalizations of the distributions in these two cases are the same). Table 9.3.4 gives the parameters, while Figure 9.3.3 illustrates the distributions for the slopes coming from these parameters.

Table 9.3.4: Synthetic case 3

Well	**Oil**				**Gas**			
	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)	Mean	Var. [10^8]	Skew. (Norm.)	Kurt. (Norm.)
1	50	0	0	0	50000	0.16	0	3
2	50	0	0	0	50000	0.36	0	3
3	50	0	0	0	50000	0.64	0	3
4	50	0	0	0	50000	1.00	0	3
5	50	0	0	0	50000	1.44	0	3

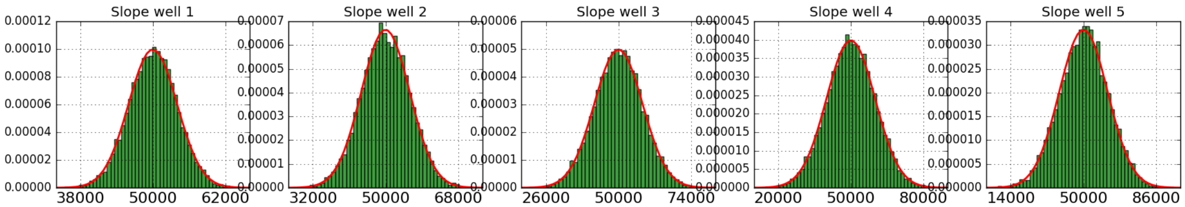


Figure 9.3.3: Case 3: Distribution for gas slope coefficient

Synthetic case 4

This case is similar to case 2, except that gas production is uncertain rather than oil production. This case has the same relative uncertainty of gas production, as the relative uncertainty of oil production in synthetic case 1 (i.e. the normalizations of the distributions in these two cases are the same). Table 9.3.5 gives the parameters, while Figure 9.3.4 illustrates the distributions for the slopes coming from these parameters.

Table 9.3.5: Synthetic case 4

Well	**Oil**				**Gas**			
	Mean	Var.	Skew. (Norm.)	Kurt. (Norm.)	Mean	Var. $[10^6]$	Skew. (Norm.)	Kurt. (Norm.)
1	50	0	0	0	50000	25	-0.5	3
2	50	0	0	0	50000	25	-0.3	3
3	50	0	0	0	50000	25	0	3
4	50	0	0	0	50000	25	0.3	3
5	50	0	0	0	50000	25	0.5	3

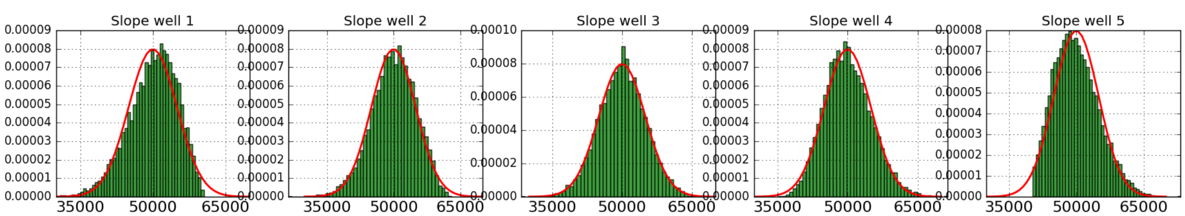


Figure 9.3.4: Case 4: Distribution for gas slope coefficient

Chapter 10

Computational Study

In this chapter we present the results from the computational study. In Chapter 10.1 we test the technical aspects of the optimization models and determine which parameters the models are suited for and which problem sizes that are possible to solve. Using the parameter settings found in this section, we solve the instances presented earlier for both the penalty and the strategy models. These results are given in Chapter 10.2.

10.1 Technical Testing

When testing the technical aspects of our optimization models, we have focused on three areas. In Chapter 10.1.1 we show how the problem size affects the solution time. In Chapter 10.1.2, we investigate the number of scenarios needed for in-sample and out-of-sample stability for the penalty and the strategy model. In Chapter 10.1.3 we explain how an improved branching strategy decreases computational time. Finally, in Chapter 10.1.4 we evaluate the technical results and conclude which model size and parameters that are appropriate for further testing.

In the technical testing, and later the case testing, the alternative formulation of the strategy model is used. We choose to use this model for all tests because (1) informal testing show that this formulation always gives the same values as the original formulation for both the decision variables and the objective value, and (2) informal testing shows that this formulation is more efficient than the original formulation, and more often results in a guaranteed optimal solution (zero optimality gap).

10.1.1 Problem Size

The two stochastic optimization models that have been developed are mixed integer linear programming (MILP) models, making them computationally hard to solve to optimality. The binary variables are handled in the solver using the branch-and-bound algorithm, which has an exponential worst case solution time. The number of variables in the optimization models depends on the problem instance being solved, involving parameters like number of wells and the number of scenarios used to represent uncertainty in parameters. To determine which types of problems our optimization models work best for, we test their performance over a range of different input combinations. Each instance is terminated after 10 minutes and we record the optimality gap. We believe 10 minutes are suitable in order to test whether the models are applicable for real-time optimization.

The results from the penalty model are shown in Table 10.1.1 and the results for the strategy model are shown in Table 10.1.2. The problem size tests are run with a realistic number of wells (3 to 20) and a realistic number of changes (2 to 4). In reality, if more changes are wanted, several iterations of the models can be run. The well model parameters are synthetically generated with realistic moments. It is observed that both the number of scenarios, number of wells and number of allowed choke changes affect the remaining dual gap. For both models, the problem is easily solved when only a few number of wells and scenarios are considered. For larger instances the problem rapidly becomes more computationally time consuming.

Except for the increased complexity due to more scenarios and more wells, there are some non-obvious reasons for the observed results. Firstly, using the notation from the mathematical formulation of the optimization models where $|\mathcal{I}|$ is the number of wells and $N = |\mathcal{J}| - 1$ is the maximum number of well changes allowed, there are $\binom{|\mathcal{I}|}{N}$ different combinations of wells that can be selected for adjustment. When N is either very small or very large compared to $|\mathcal{I}|$, the binomial coefficient is also relatively small and there are few possible ways to select the wells. This suggests a fast solvable problem, as can be seen for two allowed changes in for instance the strategy model with 15 wells, 100 and 200 scenarios or 20 wells, 50 and 100 scenarios. More allowed changes will complicate the problem up to a certain point. If the number of well changes were allowed to be zero or one, or close to or equal to the number of wells, the reduction in problem complexity would be obvious. The second non-obvious reason for increased complexity is the increase in possible modes when more wells are considered. Adding one extra well will add $N+1$ extra modes to consider (given that each mode contains no more than two shut down wells), where N is the total number of wells. That is, if increasing from four to five wells, the number of modes increase with $4+1=5$ modes. The total number of modes $|\mathcal{K}|$ with the possibility of shutting down two wells is $|\mathcal{K}| = 2 + \sum_{i=2}^{N} N$.

Table 10.1.1: Test of problem size capacity, penalty model

		Scenarios			
		50	100	150	200
Wells	Changes	% Gap	% Gap	% Gap	% Gap
3	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
7	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
15	2	0	0	0	36.3
	3	0	0	9.18	13.1
	4	0	0	0	1.56
20	2	0	89.4	205	-
	3	0	38.4	96.3	-
	4	0	23.2	68.6	-

Table 10.1.2: Test of problem size capacity, strategy model

		Scenarios			
		20	50	100	200
Wells	Changes	% Gap	% Gap	% Gap	% Gap
3	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
7	2	0	0	57.5	0
	3	0.01	0	0	3.06
	4	0.01	2.50	10.4	4.05
15	2	-	231	0	1.53
	3	-	276	431	364
	4	-	301	306	661
20	2	-	0	0	705
	3	-	4400	365	465
	4	-	292	309	613

In the two tables it can also be observed that for some instances, more allowed changes leads to a faster solved problem. A reason for this is that when allowing more changes to happen, the problem will have more possible solutions. The solution algorithm therefore has a greater probability for finding one of these solutions relatively fast. As soon as a solution or more are found, the gap is reduced and a part of the branch-and-bound tree can be excluded in the rest of the search. This is especially evident for the penalty model.

An important note to mention is the randomness involved when artificial wells and their parameters are generated. The somewhat varying results in the tables are partly affected by this. The extent of this impact is not investigated further. The general trend in the data is however clear.

For clarity, no gap is found for 20 wells, 200 scenarios for the penalty model, as no solution is found within the allocated computational time. Additionally, the moment matching algorithm does not converge for 20 scenarios with 15 and 20 wells, and no results are therefore presented for the strategy model for these instances.

Second stage relaxation to decrease computational time: A way to decrease the computational time of the problem could be through a relaxation of the binary second stage variables. Further, if the second stage relaxed solution gives the right signals to the first stage, Bender's decomposition (Benders, 1962) could be applied.

However, when testing the relaxation, we find that the solution of the relaxed problem varies considerably. The evaluated objective value also vary a lot. This indicates that the relaxed solution cannot be trusted, and we conclude that this method is to uncertain to be used.

10.1.2 Stability Tests

Stability is tested with seven wells and three allowed changes, as this is the most realistic setting concerning the Gjøa case. Each model is solved ten times for each number of scenarios and the standard deviation in the objective value is calculated. Both the in-sample and out-of-sample stability is tested for the penalty and strategy model in order to determine the suitable amount of scenarios for achieving acceptable stability. A comparison between the stability with and without moment matched scenarios is also performed. We assume that the results found for this problem instance is valid for all other variations of input data.

Figure 10.1.1 and Figure 10.1.2 show graphs over in-sample stability and out-of-sample stability for the strategy model. The standard deviation in objective value for random sampling and moment matched scenarios is plotted against the number of scenarios. Even though both have a low standard deviation in the objective value, moment matching significantly improves this measure, showing the importance of this implementation. For the moment matched scenarios, in-sample stability is very low, and out-of-sample stability is negligible. From these tests we conclude that anything above 50 scenarios will give very stable results.

For the penalty model, the in-sample and out-of-sample objective value standard deviation using moment matched scenarios is negligible with values around $1 \cdot 10^{-17}$ and $2 \cdot 10^{-8}$ respectively.

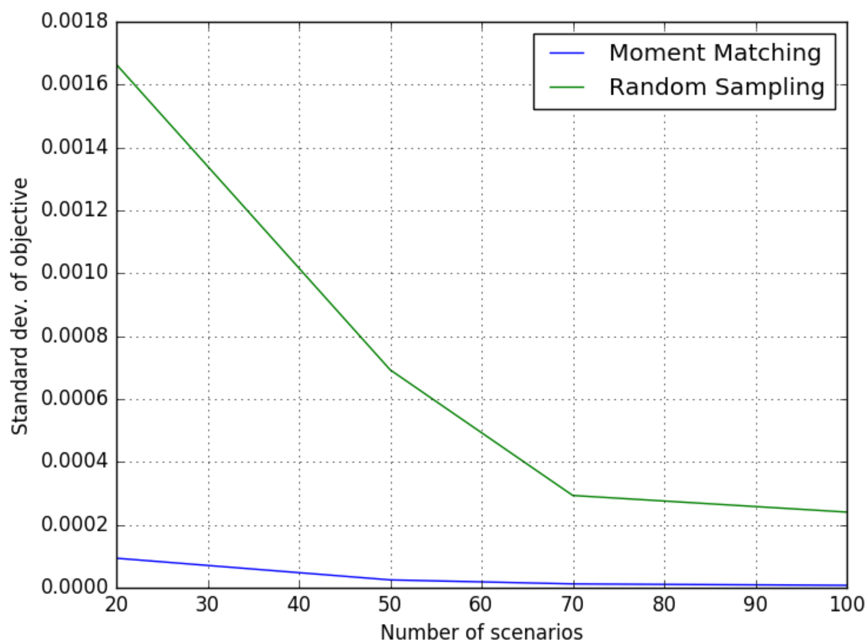


Figure 10.1.1: In-sample stability, strategy model

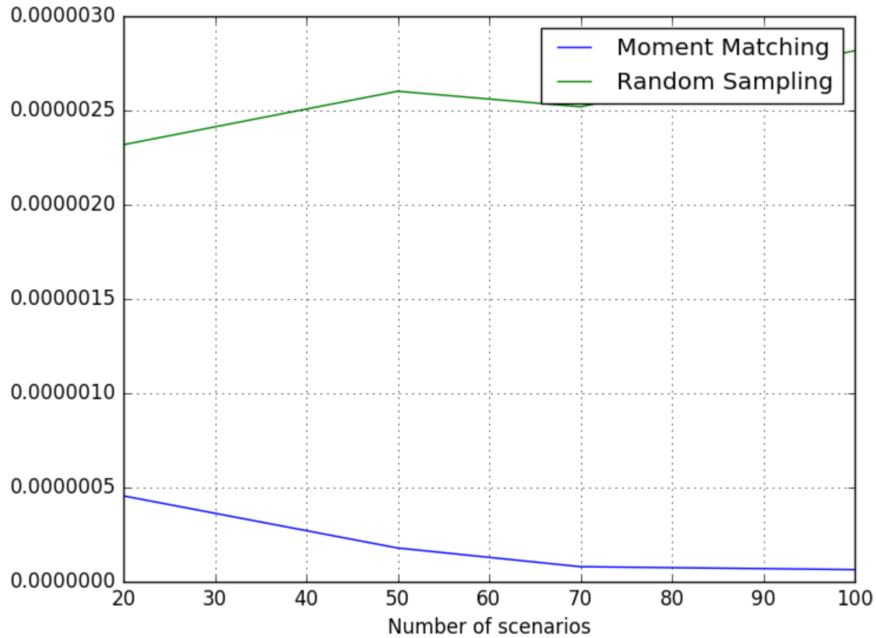


Figure 10.1.2: Out-of-sample stability, strategy model

10.1.3 Branching Strategy

The choice of branching strategy for the branch-and-bound algorithm can have a significant effect on the solution time. In our case, the default strategy set by Xpress, works well for the penalty model. However, the solution time of the strategy model can be significantly improved if choosing a customized branching strategy. This recourse model has binary variables in the second stage, so the number of binary variables increases linearly with the number of scenarios. Since the value of the first stage variables set limitations for the values of the second stage variables, a natural strategy is to begin with branching on the first stage binary variables.

For the strategy model, the first stage binary variables determine which wells are adjusted and the sequence defining the operational strategy, namely the y_{ij} variables. In order to prioritize these variables in the branch-and-bound tree produced by Xpress, we use the mosel-function `setmipdir(x:mpvar, t:integer, r:real)` which accepts a variable, directive type and a priority. The directive type is set to XPRESS-PR and priority equal to 1 is given to all y_{ij} variables.

The effect of this modified branching strategy compared to the default strategy set by

Xpress is shown in Figure 10.1.3. The default branching strategy provides the best result initially. However, the customized branching strategy quickly closes the gap and becomes the preferred strategy. The difference between the strategies is that the default strategy struggles to find new first stage solutions. The custom strategy however finds these much faster, and therefore outperforms the default strategy in the long run.

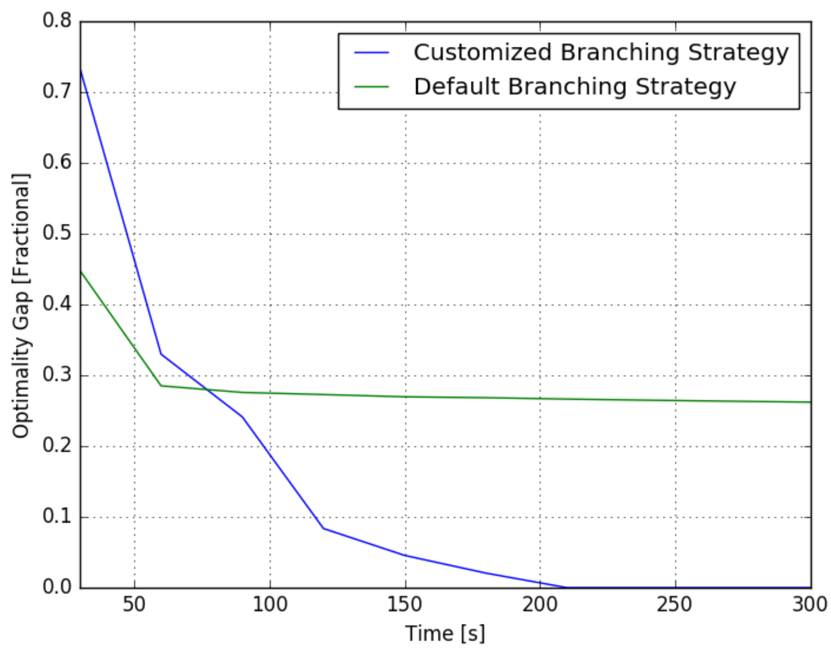


Figure 10.1.3: Branching strategy improvement

10.1.4 Evaluation of Technical Results

Comparing the results from the penalty and strategy problem size tests, we find that the penalty model can solve slightly larger problem instances and that the problem instances running out on time leaves a smaller remaining gap. For the strategy model the remaining gap is often very large for large instances. For 7 wells and 3 allowed changes, which is the case at the Gjøa field, the penalty model solve the problem to optimality for at least 200 scenarios, while the strategy model leaves a slight gap for the largest tested scenario instances.

The results in general indicate that the model user should be aware of the randomness in computational time and that there is a trade-off between the number of changes you allow and the chance of getting an optimal result. Increasing the number of well changes allowed, which is not preferable from an operator's point of view, could increase the chance of finding a solution in reasonable time. From the stability test, a significant improvement of using moment matching is found. Using moment matching there is negligible standard deviation for in-sample and out-of-sample stability for both the penalty and the strategy model, independent of the amount of scenarios. A significant improvement is also found for prioritizing the first stage variables in the branching strategy.

We therefore conclude that the optimization problem should manage the Gjøa case well, that the first stage variables should be prioritized under branching and that 50 scenarios are sufficient for accurate and representative results.

10.2 Case Testing

In this section, the instances given in Chapter 9 are solved. The results from the two different optimization models are presented and their solutions considered in an economic perspective. We especially focus on how uncertainty in the parameters affect the solutions and what can be gained by considering this uncertainty in the optimization. The Gjøa Case, with parameters generated by the preferred well model found in Chapter 9, is analysed closely to determine how the field should be operated.

In the results presented, we have recorded the objective value based on the scenario tree used in the optimization (Obj. Val) and the objective value when evaluating the solution on a large scenario tree of 10000 samples (Obj. Val*). This evaluation is performed to estimate how the solution would perform on the actual probability distributions of the uncertain parameters. Note that if the evaluated strategy does not provide a feasible solution, one or more well chokes are adjusted down according to GOR until feasibility is reached as mentioned before. Since we are interested in testing how the models perform in a real-time perspective, we set the maximum computing time to 10 minutes as mentioned, and record the optimality gap for the instances not completed within this time. The optimality gap is calculated as the current best found objective value minus the best bound, divided by the objective value. For all instances, both the deterministic and stochastic solutions are recorded and compared. VSS and EVPI are estimated, using the Obj. Val* of the stochastic, deterministic and perfect information solutions. In this way, the value of applying the stochastic models and the value of information about the outcomes of the uncertain parameters, can be estimated.

10.2.1 Results for the Penalty and the Strategy Model

In the following, the results for the penalty and strategy model are presented. Each case is presented separately, and then a comparison is performed at the end.

The main assumption behind the penalty model is that the gas constraint is soft and can be violated at a cost. When there is no uncertainty in the well parameters, the optimal solution is a set of choke changes moving the production system to the feasible operating point with highest oil production. However, when the parameters are uncertain, this point may no longer be feasible. Unlike the deterministic solution, the stochastic solution found from the penalty model is tailored to be more robust to this uncertainty. The reason that it is more robust, is that the penalty of doing recourse actions was taken into account when the optimal operating point was computed.

An underlying assumption for the strategy model is that the operator implements a set of well changes in a predefined sequence. Since it is assumed that the operator can observe when the gas constraint is met, the constraint is never violated. However, the solution from the deterministic model does not specify any sequence for the choke changes. To compare the stochastic and deterministic models, we therefore need a *deterministic strategy*, meaning a sequence of choke changes to be implemented. This is achieved by solving the deterministic model and subsequently ordering the wells. A natural ordering is to first perform all choke reductions before increasing the chokes in order of increasing GOR.

Gjøa Case

The results of the penalty model for the Gjøa case is shown in Table 10.2.1, and the results of the strategy model for the Gjøa case is shown in Table 10.2.2.

Some overall trends can be discovered if we study the results. First, by comparing the solutions given by the stochastic penalty model, and the solutions given by the stochastic strategy model, we can see that except for the test with $\Delta Gas = 0$ and 2 adjustments, they are almost identical. Even though the values of the decision variables differ for this test, we can see that the evaluated objective values are the same for both models. Therefore, these solutions may actually be equally good, and the solution given by the penalty model may also be optimal in the strategy model. Second, we observe that the order of adjustments given by the strategy model corresponds with the GOR order of the wells (cf. Table 9.2.6). For down-adjustments, the chokes of the wells with the highest GOR are adjusted first, and for up-adjustments, the chokes of the wells with the lowest GOR are adjusted first. Together, these two results have a particularly interesting implication, namely that the penalty model may actually be used to find a *strategy*. First, the penalty model can be used to find which chokes to adjust, and how much to adjust them. Then, the order of adjustments can be assigned heuristically, using the expected GOR values. This would, at least for these tests, result in the same optimal evaluated objective values as the strategies given by the strategy model. For most cases this is obvious. For the test with $\Delta Gas = 0$ and 3 adjustments however, this might be harder to see. Observe that the penalty model suggests adjusting the last choke up by only 0.095, while the strategy model suggest adjusting it by 0.116. Thus, one may think that a heuristically created strategy from the penalty model would not be as good as the strategy given by the strategy model. Recall however that, in the evaluation of *strategies*, the magnitude of the last adjustment has nothing to say on the evaluated result (as long as it is non-zero). In the evaluation of a strategy, if the last adjustment started does not bring you to the gas capacity, it will be continued until the capacity is reached. Therefore, even for this test, a strategy created heuristically out of the penalty model result would be optimal.

Further, comparing the stochastic solutions with the deterministic solutions, we can see that for most tests, they only differ slightly. The main difference is that for several tests, the stochastic model suggests adjusting down the choke of the well with the second worst GOR (well 5), while the deterministic model suggests the choke of the well with the worst GOR (well 6). Although well 5 has a GOR with an expected value that is higher than that of well 6, the GOR of well 5 is much more certain (compare the variance of the slopes for these chokes for both oil and gas, see Table 9.2.4 and Table 9.2.5). This may explain why well 5 is preferred over well 6 in the stochastic solutions. These stochastic solutions are interesting, as they make use of information not available to the deterministic model. In that way, solutions are found with higher expected production improvements than what is achievable with the deterministic approach (i.e. positive VSS). We can also observe that the VSS using the suggested models is in the range $0 < VSS < 0.05$. That is a quite small VSS, but it is non-negative. Even though it might look negligible, even a small improvement can be very valuable when large oil volumes are produced. Note that we could have gotten a negative VSS because this

Table 10.2.1: Results penalty model, Gjøa case

ΔGas	Changes	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Std. Dev.	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	Δx_6	Δx_7	*EVPI*	*VSS*
0	2	Stoch.	8.50	0	8.50	0.35	0	0.116	0	0	0	-0.122	0	0.21	0.03
		Det.	8.56	0	8.47	0.39	0	0.116	0	0	0	-0.120	0		
	3	Stoch.	13.45	0	13.45	0.39	0.125	0.095	0	0	-0.228	0	0	0.05	0.05
		Det.	13.49	0	13.40	0.38	0.125	0.092	0	0	-0.228	0	0		
10000	2	Stoch.	9.36	0	9.36	0.31	0	0.116	0	0	-0.095	0	0	0.13	0.03
		Det.	9.41	0	9.34	0.33	0	0.116	0	0	0	-0.070	0		
	3	Stoch.	16.07	0	16.07	0.39	0.125	0.116	0	0	-0.198	0	0	0.24	0.03
		Det.	16.13	0	16.04	0.39	0.125	0.116	0	0	-0.195	0	0		

Table 10.2.2: Results strategy model, Gjøa case

ΔGas	Changes	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Std. Dev.	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	Δx_6	Δx_7	Seq.	*EVPI*	*VSS*
-50000	2	Stoch.	-4.28	0	-4.28	0.38	0	0	0	0	-0.124	-0.160	0	6-5	0.12	0
		Det.	-4.28	0	-4.28	0.38	0	0	0	0	-0.117	-0.160	0	6-5		
	3	Stoch.	2.37	0	2.37	0.46	0.125	0	0	0	-0.228	-0.156	0	6-5-1	0.11	0.04
		Det.	2.44	0	2.32	0.48	0.125	0	0	0	-0.220	-0.160	0	6-5-1		
0	2	Stoch.	8.51	0	8.50	0.36	0	0.116	0	0	-0.161	0	0	5-2	0.20	0.03
		Det.	8.56	0	8.47	0.41	0	0.116	0	0	0	-0.120	0	6-2		
	3	Stoch.	13.49	0	13.49	0.43	0.125	0.116	0	0	-0.228	0	0	5-1-2	0.01	0
		Det.	13.49	0	13.49	0.43	0.125	0.092	0	0	-0.228	0	0	5-1-2		
10000	2	Stoch.	9.37	0	9.36	0.31	0	0.116	0	0	-0.095	0	0	5-2	0.13	0.02
		Det.	9.41	0	9.35	0.34	0	0.116	0	0	0	-0.070	0	6-2		
	3	Stoch.	16.08	0	16.07	0.39	0.125	0.116	0	0	-0.198	0	0	5-1-2	0.24	0.04
		Det.	16.13	0	16.04	0.47	0.125	0.116	0	0	0	-0.148	0	6-1-2		

value was based on evaluations on a different (larger) scenario tree than the tree that was used when solving of the stochastic model.

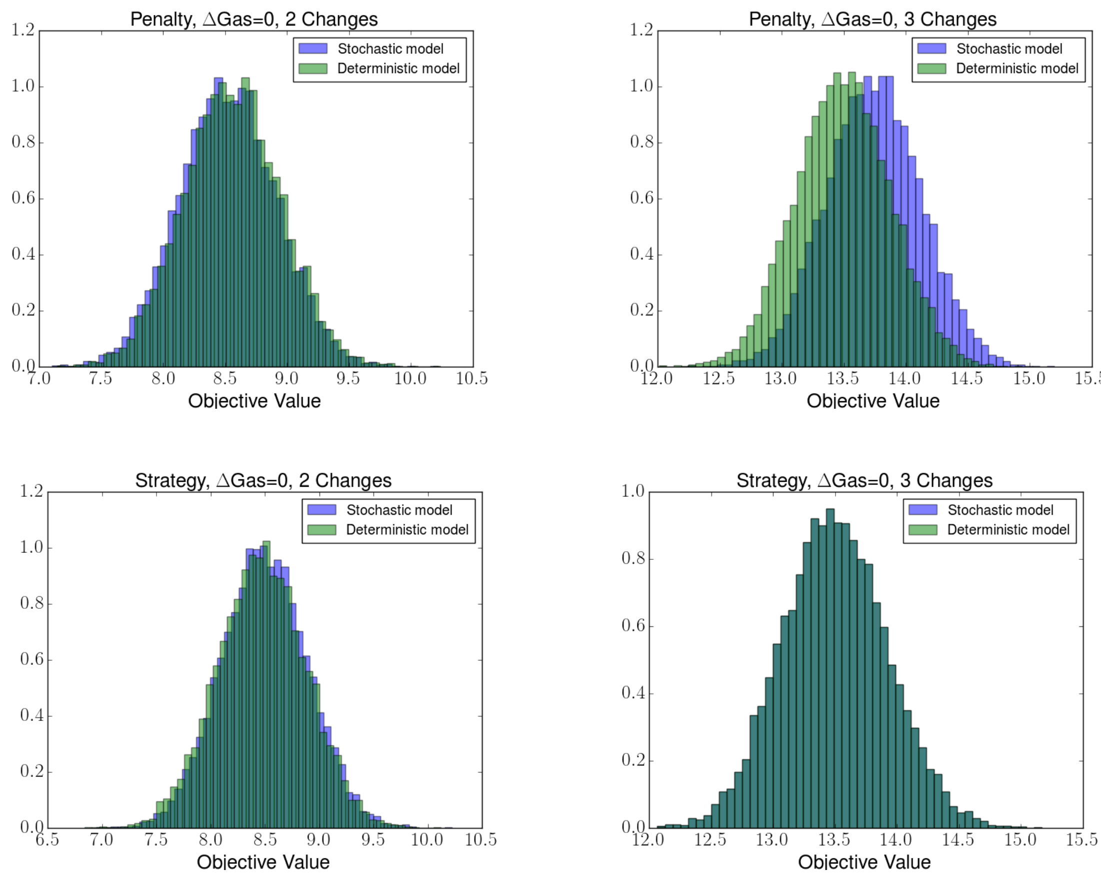


Figure 10.2.1: Histograms over possible objective values, Gjøa case

Finally, the objective values from the tests with $\Delta Gas = 0$ show that it is possible to make a profit by decreasing production from some wells, and increasing the production from other wells. That is, these results suggest that a significant improvement in oil production can be achieved without increasing gas production. All improvement in oil production that can be achieved without changing the gas production, can be considered pure profit. The reason is that making these adjustments costs almost nothing relative to the value of producing more oil. The historic data show that current oil production is around 160 Sm^3/h (see Figure C.0.2). With two available adjustments, an improvement of around 8.5 Sm^3/h is expected, and with three adjustments available, an improvement of around 13.5 Sm^3/h is expected. This equals approximately 5% and 8% respectively, which can be considered a significant improvement. Note however that the estimated improvements are based on the data models that were built on historic data. Therefore, this is the expected improvement under the assumption that the data models are good representations of reality. Based on the plots of predicted historic production versus real historic production (Figure C.0.2, Figure C.0.1, Figure B.0.4 and Figure B.0.2) and computed R^2 values (Table 9.2.7), this assumption seems to hold. Therefore, an improvement of 5% and 8% for two and three adjustments, may be what can actually be expected by implementing the suggested strategies. Further, if the estimated uncertainty of production matches reality, then we can say with 99.7% certainty that the production improvement with two adjustments available will be in the range [7.45 , 9.55] Sm^3/h and with three adjustments available will be in the range [12.3, 14.7] Sm^3/h. This can be seen in the plots in Figure 10.2.1. Based on these histograms, it may seem that the solution given by the penalty model is better than the solution given by the strategy model. However, this is not necessarily true, as the evaluated objective value is computed differently for each of the models, as discussed in Chapter 7.4. The strategies from these two models are almost identical, so the difference in objective value is most likely caused by the difference in the computation of this value.

Synthetic Case 1

The results of the penalty model for synthetic case 1 is shown in Table 10.2.3, and the results of the strategy model for synthetic case 1 is shown in table 10.2.4. All results, both for the strategy model and the penalty model, have an equal objective value (Obj. Val.), and evaluated objective value (Obj. Val.*). Further, for each similar test setup, the objective value is the same for both the stochastic version and deterministic version of these models. It can also be seen that the objective value is exactly equal to the sum of the realized choke adjustments times the mean oil contribution (which is 50 for all wells). That is, for the strategy model, with ΔGas equal to 10000, the expected oil production improvement is $(0.2 + 0 + 0 + 0 + 0) \cdot 50 = 10$. Note that for $\Delta Gas = -50000$, the result of the solution given by the strategy model includes two adjustments. However, as the gas contribution is certain, only the first will be realized in each scenario.

As the stochastic and deterministic versions give the same result, the VSS is zero for all the results in case 1. This should not be surprising, as only oil is uncertain in this case. Since gas production is certain and all wells have equal mean oil and gas contribution, there is no way to for example adjust one choke down and another fractionally more up, with the result being higher oil production. The optimal solution is therefore to do nothing when ΔGas is zero. When ΔGas is non-zero, which chokes that are adjusted is random.

Note that if the moments had been slightly off in the scenario set, the stochastic model would not have given the no-adjustment solution as an optimal solution for $\Delta Gas = 0$. If the moments in the scenario set were not perfectly matched, the wells would have unequal means, making some wells better than others in the stochastic model. Then, only solutions involving switching of chokes would have been optimal in the $\Delta Gas = 0$ test and probably also in the $\Delta Gas = 10000$ test.

Table 10.2.3: Results penalty model, synthetic case 1

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*EVPI*	*VSS*
0	Stoch.	0	0	0	0.20	0	0	0	-0.20	4.26	0
	Det.	0	0	0	0	0	0	0	0		
10000	Stoch.	10.00	0	10.00	0	0	0.20	0	0	4.69	0
	Det.	10.00	0	10.00	0	0	0.20	0	0		

Table 10.2.4: Results strategy model, synthetic case 1

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*Seq.*	*EVPI*	*VSS*
-50000	Stoch.	-50.00	5.28	-50.00	-0.50	-0.50	0	0	0	1-2	6.81	0
	Det.	-50.00	0	-50.00	-0.50	-0.50	0	0	0	2-1		
0	Stoch.	0	Inf.	0	-	-	-	-	-	-	4.26	0
	Det.	0	0	0	0	0	0	0	0	-		
10000	Stoch.	10.00	19.84	10.00	0	0.20	0	0.20	0	2-4	4.69	0
	Det.	10.00	0	10.00	0	0.20	0	0	0	2		

Synthetic Case 2

The results of the penalty model for synthetic case 2 is shown in Table 10.2.5, and the results of the strategy model for synthetic case 2 is shown in Table 10.2.6. The results for case 2 are almost the same as for case 1. All objective values are equal to those in case 1, and the values of the decision variables are mostly the same. As in case 1, only oil production is uncertain, and consequently there is no expected value of knowing about this uncertainty, that is solving the problem stochastically. This means that the VSS is zero for all the tests in this case as well.

Note that for the test with ΔGas equal to zero, the optimality gap for the solution given by the stochastic model is infinite. This is a consequence of how the optimality gap is computed, namely that the objective value is in the denominator of the optimality gap formula. Therefore, if the objective value is very close or equal to zero, this value will become very large or infinite. Such gaps are also found for some of the subsequent tests.

Table 10.2.5: Results penalty model, synthetic case 2

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*EVPI*	*VSS*
0	Stoch.	0	0	0	0	0	0	0	0	2.33	0
	Det.	0	0	0	0	0	0	0	0		
10000	Stoch.	10.00	0	10.00	0.20	0	0	0.20	-0.20	2.82	0
	Det.	10.00	0	10.00	0	0	0.20	0	0		

Table 10.2.6: Results strategy model, synthetic case 2

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*Seq.*	*EVPI*	*VSS*
-50000	Stoch.	-50.00	0	-50.00	-0.50	-0.50	0	0	0	1-2	4.16	0
	Det.	-50.00	0	-50.00	-0.50	-0.50	0	0	0	2-1		
0	Stoch.	0	Inf.	0	-	-	-	-	-	-	2.33	0
	Det.	0	0	0	0	0	0	0	0	-		
10000	Stoch.	10.00	6.83	10.00	0.20	0	0	0	0	1	2.82	0
	Det.	10.00	0	10.00	0	0.20	0	0	0	2		

Synthetic Case 3

The results of the penalty model for synthetic case 3 is shown in Table 10.2.7, and the results of the strategy model for synthetic case 3 is shown in Table 10.2.8. All strategies from the deterministic model are similar as in case 1 and 2, because the mean oil contributions and mean gas contributions are the same. The only difference compared to these cases, is which chokes that are adjusted to their limit. As mentioned earlier, this may vary as all wells have the same expected GOR.

When ΔGas is -50000, the strategy given by the stochastic strategy model and the strategy given by the deterministic model both start by closing the chokes of well 1 and 2. These are the wells with the most certain GOR. The strategy given by the stochastic model also includes a third step, which is adjusting down the choke of well 5 - the well with the least certain GOR. The solution given by the deterministic model does not include a third step. The reason is that, given deterministic contributions, a sum of adjustments equal to -1 would be sufficient to reduce the gas production by 50000. However, when the gas contributions are uncertain, a sum of adjustments equal to -1 might result in a reduction that is less than 50000. This is not a legal state according to the gas capacity constraint. As mentioned, when the strategy results in an illegal state, the last adjusted choke is adjusted further down (if it is not turned off), and if this is not sufficient, chokes are adjusted down in order of GOR (high to low). Since the GOR is the same for all wells, it is random which choke that is adjusted down in the realized strategy of the deterministic model. However, as VSS is zero, we can be very certain that the third realized adjustment must be the same for both models. Alternatively, it may be that another choke is adjusted down in the realization of the deterministic strategy, and that this results in the exact same objective values as for the realization of the stochastic strategy. However, this is more unlikely.

When ΔGas is 0, the stochastic penalty model and the stochastic strategy model give different solutions. The penalty model suggests no adjustments, while the strategy model suggests adjusting down choke 4 and 3 respectively, and then adjusting up choke 2. Interestingly, the strategy given by the strategy model results in a positive VSS. This means that the strategy model has been able to find a way to favourably exploit the uncertainty in gas contribution. By adjusting down choke 3 and 4 by a total of 0.14 percentage points, choke 2 can on average be adjusted up by more than 0.14. This is what causes the non-zero expected production improvement. This result is particularly interesting as it means that, even when all wells have the same expected contributions for both oil and gas and $\Delta Gas = 0$, there may be strategies that are better than doing nothing. A solution like this would not be found if only the expected values were taken into account. Further, we can see that neither the most certain, nor the least certain choke, is adjusted (1 and 5 respectively). Intuitively one may think that these chokes should be the ones to adjust, either upwards or downwards, to achieve the highest expected production improvement. These results however suggest otherwise. Note that the optimality gap was still very high when this test was ended. Therefore, it may be that the actual optimal solution includes choke 1 and/or 5.

When ΔGas is 10000, both the stochastic penalty model and the stochastic strategy model suggest solutions that include more steps than simply adjusting one choke to its upper limit. Adjusting one choke to its upper limit is the solution given by the deterministic model. The strategy given by the penalty model includes three up-adjustments, while the strategy given by the strategy model includes one down-adjustment and two up-adjustments. Both these solutions result in a higher expected production improvement than the solution given by the deterministic model.

Table 10.2.7: Results penalty model, synthetic case 3

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*EVPI*	*VSS*
0	Stoch.	0	0	0	0	0	0	0	0	4.33	0
	Det.	0	0	0	0	0	0	0	0		
10000	Stoch.	10.03	0	10.01	0	0	0.06	0.09	0.11	3.88	0.65
	Det.	10.00	0	9.36	0	0	0.20	0	0		

Table 10.2.8: Results strategy model, synthetic case 3

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*Seq.*	*EVPI*	*VSS*
-50000	Stoch.	-51.40	0	-50.29	-0.50	-0.50	0	0	-0.20	2-1-5	7.39	0
	Det.	-50.00	0	-50.29	-0.50	-0.50	0	0	0	2-1		
0	Stoch.	0.11	6106	0.08	0	0.20	-0.12	-0.02	0	4-3-2	4.25	0.08
	Det.	0	0	0	0	0	0	0	0	-		
10000	Stoch.	10.36	21.85	10.30	0.16	0	-0.07	0	0.20	3-1-5	3.60	0.71
	Det.	10.00	0	9.58	0	0.20	0	0	0	2		

Synthetic Case 4

The results of the penalty model for synthetic case 4 is shown in Table 10.2.9, and the results of the strategy model for synthetic case 4 is shown in table 10.2.10. Again, objective value given by the deterministic model are the same as for the three previous cases.

When ΔGas is -50000, the stochastic strategy model gives a solution that is different from the deterministic model in terms of which chokes that are adjusted. Both models suggest solutions where two of the chokes are turned off. However, the strategy given by the stochastic model also includes a third step, which is adjusting down choke 1. Probably, the strategy given by the deterministic model does not lead to a feasible solution in all scenarios, and therefore a third step is most likely executed when this strategy is evaluated. Interestingly, the difference in which chokes that are adjusted results in a positive VSS, even though the gas distributions for all wells have the same first two moments. This implies that, if building well models for a real production system, and these moments show to be non-normal, this non-normality should be represented in the generated scenario set for the optimization models.

When ΔGas is 0, the stochastic penalty model suggests no adjustments, while the stochastic strategy model suggests adjusting down choke 1 and choke 3, and adjusting up choke 2. For the strategy model, the VSS is again positive, which means that there is a way to exploit the uncertainty of production for the wells, to achieve a higher production improvement than if uncertainty is not taken into account. Further, this supports the finding for the test with $\Delta Gas = -50000$, namely that wells having the same first two moments for their production distributions, are not similar when higher moments are taken into account. Note however that the penalty model does not suggest the same solution as the strategy model for this test, and rather suggests a solution involving no adjustments. That is, the cost of doing penalized down-adjustments for bad scenarios most likely outweighs the added profit for good scenarios. Therefore, the optimal strategy in the penalty model is to keep the current setting of the production system. That is, in the penalty model the wells are not seen as *sufficiently different* for it to be valuable to take the risk involved with adjusting the system.

When ΔGas is 10000, both the stochastic penalty model and the stochastic strategy model again suggest solutions that include more steps than simply adjusting one choke to its upper limit. Similar to case 3, the penalty model suggests only up-adjustments, while the strategy model suggests adjusting down one choke, and adjusting up two others. Both solutions result in an improvement that is higher than the solution given by the deterministic model. Again, this supports the findings revealed in the two previous paragraphs.

Table 10.2.9: Results penalty model, synthetic case 4

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*EVPI*	*VSS*
0	Stoch.	0	0	0	0	0	0	0	0	2.47	0
	Det.	0	0	0	0	0	0	0	0		
10000	Stoch.	9.94	0	9.91	0	0.05	0.06	0	0.10	2.61	0.34
	Det.	10.00	0	9.57	0	0	0.20	0	0		

Table 10.2.10: Results strategy model, synthetic case 4

ΔGas	Model	*Obj. Val.*	*% Gap*	*Obj. Val.**	Δx_1	Δx_2	Δx_3	Δx_4	Δx_5	*Seq.*	*EVPI*	*VSS*
-50000	Stoch.	-51.38	0	-50.15	-0.14	0	-0.50	0	-0.50	5-3-1	4.77	0.04
	Det.	-50.00	0	-50.19	-0.50	-0.50	0	0	0	2-1		
0	Stoch.	0.08	2220	0.06	-0.07	0.20	-0.08	0	0	1-3-2	2.41	0.06
	Det.	0	0	0	0	0	0	0	0	-		
10000	Stoch.	10.00	5.08	10.07	0.20	0	-0.02	0.07	0	3-4-1	2.45	0.43
	Det.	10.00	0	9.64	0	0.20	0	0	0	2		

Chapter 11

Discussion

In this chapter, we further analyse and discuss findings from previous chapters. Chapter 11.1 give remarks on the tested regression methods and well models. Further, we compare the suggested optimization models and solutions in Chapter 11.2. In Chapter 11.3, we discuss practical considerations regarding how the well and optimization models could be applied in real life applications.

11.1 Remarks on Building of Well Models

Three different well models are suggested in this thesis, the *export model*, the *multiphase meter model* and the *combined model*. Ordinary least squares and different robust regression techniques are used to find parameters for each well. After the data compression algorithm has been used we see that the regression methods mostly give very similar results. This indicate that the compression algorithm smooths many of the inaccurate outliers. Using the coefficient of determination, we find that the assumption of local linearity is very good for gas predictions, but not that accurate for oil predictions. Based on regression plots and the knowledge gained from the analyses performed, we despite that believe that the results found are realistic. From the optimization results, we see that as long as the oil parameters give an indication of how oil production will vary with different choke values reasonable results can be achieved, as it is uncertainty in gas prediction that affects the value of incorporating uncertainty. We should also note that since real-time optimization is often concerned with increasing daily oil production by small increments, even minor errors in the well models could have a large impact. The results should therefore be treated with caution.

From the three months of historical production data, we observe that very few excitations in the chokes have been performed. To ensure an optimal solution, it is vital that valid parameter models can be built for all wells. The very few excitations for some wells means that it is hard to confirm model validity. To reduce this uncertainty, data sets containing more variations in the choke opening of each well is needed. That is, in order to fully capture the relationships between choke setups and production rates, the operators must physically excite each choke more than what has been done in the past. Operators tend not to be willing to do this as periods of decreased production might occur, loosing profits for the oil field. Since the well models' upper limit is set to be 10% above historically seen data points, more excited data can also be obtained by implementing suggested solutions. This will however take a lot more time than if exciting the chokes with the only goal being a richer data set. Adjusting the chokes more could also help find more certain correlation measures between the wells, which could further improve and help explain the found results. In this thesis, a bootstrapping procedure is applied to find the distributions and correlation of the parameters. Using bootstrapping we try to utilize the information inherent in the data to provide insight into the overall properties of each well. An even larger data set with more variation would also help the bootstrapping to provide better results. We note that other statistical methods and software than bootstrapping could be applied to find the same information, but that a larger variation in the data regardless of method is desirable. More certain and excited data should therefore be requested from the petroleum operators. This request can be explained and justified based on the possible economic benefit and the improved production that this thesis show can be achieved.

11.2 Comparison and Evaluation of the Results of the Optimization Models

The two optimization models developed rely on different assumptions regarding how the operator of a field performs adjustments to the production system. The penalty model is based on the assumption that the operator implements all changes in an arbitrary order and later makes down-adjustments if the gas constraint is violated. These down-adjustments are associated with a recourse cost. The strategy model is instead based on the assumption that the operator follows an operational strategy, meaning a defined sequence of adjustments to the production system. In addition, it is assumed that this sequence of well changes are implemented in a manner so that the the operator observes when the gas constraint is met and therefore never violates this constraint. For the operator to know exactly when the constraint is met, means that the changes must be implemented slowly so the system has time to stabilize.

For the Gjøa case, an interesting finding is that the objective value of the solution given by the penalty model is almost identical to the solution given by the strategy model, for most of the tested variants of this case. The decision variables however take slightly different values. The same trend is observed for most variants of the synthetic cases. This indicate that several different choke settings can give equally good oil production. In most cases we find that using the penalty model solution and ordering the adjustments heuristically using expected GOR can give the same objective value as the strategy found by the strategy model. In some cases we however find that the strategy model does find a better solution than the heuristically created strategy from the penalty model. Based on this we conclude that the models in general can be used interchangeably with little adaption, but that both models should be applied to obtain the most robust results.

From the Gjøa case we also find that the objective value significantly increases when more well changes are allowed. It is however not realistic to implement more than three changes at a time. Even though a better objective value could be achieved through more allowed changes, an operator would not consider this as large changes from the current operating point are considered very risky. In reality, an operator would implement the three changes suggested, observe the new system state, and thereafter optimize again to find three possible new changes. Over time, this will move the system towards an optimal setting.

Additionally for the Gjøa case, we find that the stochastic solution often prefer adjusting chokes with lower variance even though its GOR is worse, suggesting a different operational pattern than the deterministic solution. In this way, the stochastic solution hedge against uncertainty with the downside being that a well with less favorable average production is chosen. A similar pattern is found for the strategy model in synthetic case 1 and 3, but since all those solutions have a remaining optimality gap, we cannot conclude decisively.

For both the Gjøa case and synthetic case 3 and 4 we find a positive VSS. Incorporating uncertainty in the models does in these cases provide more profitable solutions than solving the problem deterministically. In synthetic case 1 and 2, no improvement is found using the stochastic models. The reason is that gas production must be uncertain if the stochastic models are to have a possibility for finding better solutions than the deterministic models. With only oil production being uncertain, the stochastic and deterministic models will always suggest the same changes, except for when the GOR is equal for several wells. When the GOR is equal for several wells, and only oil is uncertain, which chokes that are adjusted has no impact on the production improvement, and will therefore be arbitrary. We can regardless conclude that including uncertainty in the models does provide at least an equally profitable solution as if not considering uncertainty at all. We also find that the VSS in general is higher when ΔGas is 10000, rather than 0. This could indicate that using stochastic programming is even more useful when production is not close to the gas constraint. Note however that most of the time you will not operate far from the gas constraint. This only happens when for instance new and improved equipment is installed that would change the gas capacity. Normally, you would operate on the gas constraint and iteratively improve oil production for this level of gas production.

We can also note that testing with a higher penalty cost in the penalty model only gives slightly more conservative solutions, as the model then ensures with even greater certainty that the gas constraint is not breached. The difference in objective value is not significant.

11.3 Practical Considerations

Real-time production planning has a very short time horizon as it is concerned with day-to-day or even hour-to-hour adjustments of the system. As noted in the literature study, an implementation could amongst other things give huge savings in engineering time, improve well surveillance, get more frequent well model updates and increase efficiency in field operations. Preferably, an optimization tool should be able to suggest new solutions within minutes so that it can be applied as decision support during production planning meetings and that changes can be implemented almost continuously. The two optimization models presented in this thesis have been developed with this in mind. However, as shown in the technical study, a large number of wells and scenarios can lead to difficulties of finding an optimal solution within reasonable time, especially for the strategy model. Despite this, initial tests show that the solution found regarding which chokes to change seldom changes after five minutes of computing time. In most cases this solution can therefore be used for field decisions. An important note is that the problem in this thesis is run on a simple desktop computer, and that the models in a real life application would be run on a more powerful server with higher computing capacity. Optimal solutions can therefore be expected at a fraction of the time recorded in our results.

If the suggested models are to be implemented at the Gjøa field or another real life field, many considerations will have to be taken. Firstly, the properties of a proper information technology structure would have to be decided. Measurements from total export and MPMs will need to be stored and pre-processed so that it in an easy way can be read into and analysed by regression method software. An interface between this software and the chosen optimization software would also need to be implemented. The solutions suggested by the optimization software thereafter needs to be presented in an intuitive and understandable way to production engineers and other decision makers. This is important as the main reason that a solution is not implemented is that decision makers do not believe and trust the suggestions given by the models. Additionally, background information such as well models and plots of these and other relevant information should be made easily available to corroborate the results being presented.

Many trade-offs are also associated with such a system. A production engineer needs to consider the transient periods and other costs which is connected to changing choke openings frequently. These costs are not properly reflected in the suggested solutions from the proposed models, meaning that the engineer will have to assess the possible gain in production against cost of changes before deciding to go for a suggested solution or not. However, if the production engineer is able to provide an estimate on the cost of performing adjustments, it could easily be incorporated into the optimization models by changing the value of the ϵ parameter (cf. Chapter 7.2, Chapter 7.3.1 and Chapter 7.3.2).

In the data set used for input in the Gjøa models, very limited changes have been performed for each well. Another trade-off for a production engineer is therefore between the costs of changing wells to gain more information as input to the optimization model and the costs of performing these changes. Giving the optimization problem a larger and more diverse input data set will increase the possibility of finding a close to optimal solution for the choke openings. As mentioned this can also be achieved through trusting the suggested solutions and implementing these.

In the suggested penalty model, the recourse cost could be adjusted to fit an operators preferred risk profile. If the operator is in charge of a production facility where a breach of the gas production limit is very costly, a larger penalty could be assigned when solving the problem. Similarly, if a breach of the gas limit is not that costly, a smaller cost could be assigned to increase the chance of getting close to the gas production limit. A breach of the limit would only mean that the operator would have to adjust one or more chokes slightly down to again get within allowable production limits.

Chapter 12

Conclusion

In this thesis, an approach to solve the real-time petroleum production optimization problem is presented. This problem is concerned with increasing day-to-day oil production from a field by applying real-time production data to find the optimal way to operate. Hence, both methods for generating input parameters and optimization models are needed. Analyzing the previous literature on the topic, few authors apply historical petroleum production data to generate input to the optimization models, but instead rely on inaccurate simulators. In addition, the optimization models in the literature are seldom tailored to the work process of the operator and do not in general incorporate uncertainty in input parameters. In the work presented, these aspects are addressed and a complete approach to handle the real-time optimization problem is given.

We develop three well models based on different available oil and gas measurements at a field. When comparing them, we find that the *combined well model*, combining both export and multiphase meter measurements, give the most accurate prediction of oil and gas production. If a petroleum field only has export measurements available, investments in MPMs could therefore be considered if more accurate prediction models are wanted. Additionally we find that the best regression method depends on the input data available, and that several methods therefore should be tested. Also, the estimated parameters are sensitive to the underlying assumptions, for instance how the different measurements are weighted in the regression.

For the data available at the Gjøa field, we find that Huber's T is the best regression method. Bootstrapping is used to find distributions for the parameters for each well model. All parameters are shown to be close to normally distributed. Based on the linear model found, and the moments and correlation estimated through the bootstrapping procedure, moment matching is used to find the best possible set of scenarios representing the properties of each well. The assumption of local well model linearity is by the coefficients of determination shown to be very good for total export gas production, and somewhat more questionable for total export oil predictions.

Two optimization models for different work process patterns of the operator are presented. Both the *strategy model* and the *penalty model* are recourse models incorporating uncertainty in the well parameters. The technical study shows that the penalty model can handle larger problem instances than the strategy model and based on that it might be more suitable for larger fields with many wells. The binary second stage variables in the strategy model lead to very large branch-and-bound trees and long solution times. We show that branching on the first stage variables first, tends to give tighter bounds in the branch-and-bound tree and we therefore employ this branching strategy when solving the case studies. When testing the models' in-sample and out-of-sample stability, we find that moment matched scenarios give significantly lower objective value variance than randomly drawn scenarios, and that the issue of solution stability is negligible using those scenario sets. From the results of the case study at Gjøa, we find that the penalty model and the strategy model with a few exceptions give the same objective values even though the decision variables take slightly different values. If using the penalty model and sorting the suggested well changes heuristically after GOR, a strategy giving the same objective value as the strategy model is obtained. In some cases however, the strategy model can suggest a better solution than the penalty model, and both models should therefore be tested. An expected production improvement of about 8.5% and 13.5% is found for two and three allowed changes respectively for both stochastic models and the deterministic model. If assumptions and well models are correct, this is a significant improvement. We find a small VSS for most variants of the Gjøa case, suggesting that solving the problem stochastically is sligthly beneficial.

From the synthetic cases, we find that if only oil predictions are uncertain, the stochastic models are no better than the deterministic ones. If gas predictions are uncertain on the other hand, slightly better objective value is found using the stochastic models. Both the level of variance and the level of skewness for gas production is shown to affect the stochastic objective value. Significant EVPI is found for all cases, indicating that finding ways to improve the well models even further to get a better understanding of what might happen in the future is valuable.

To implement the suggested well and optimization models in an optimization software for use at an actual field, further development is necessary. The well models need to apply iterative regression techniques so that new measurements can be included in the parameter estimation continuously. Additionally, a user friendly interface must be developed, so that results is easily available and interpretable for end users. This thesis however show that there is potential for increased oil production by employing real-time optimization methods. Other positive effects such as more frequent well model updates and increased efficiency in field operations can also be achieved with an implementation. Further research on the topic is therefore encouraged.

Chapter 13

Further Work

The models in this thesis are subject to some assumptions and simplifications. In the following, suggestions on how to improve on these simplifications are presented. The suggested further work is recommended in order to increase confidence and credibility in the results, before using the models and solutions in practical applications.

Well models

The parameters for oil and gas could be estimated more accurately by partitioning the data set for each well in piecewise subsets and performing linear regression on each individual subset. In addition to assuming local linearity, this would include a possibility to move to a different linear well model modelled on historical data. In a mode formulation, this would give a significant increase in the number of modes and therefore input parameters. The increase in computational difficulty must be weighted against the possible increase in result and solution accuracy.

Additionally, the measurements in the data set applied in the regression are in fact uncertain themselves, meaning that we do not know whether the recorded production at some time was exactly what is given by the measurement. This can lead to biased parameters. If this is the case and how that can be prevented should be investigated further.

Model extensions and changes

Some model extension and changes can also be considered in future work. There are lots of historical data available at the Gjøa field and other relevant petroleum fields. Examples are the availability of pressure data over each wellhead valve or the amount of hydrogen sulfide gas (H_2S) in a well flow.

Instead of having the choke opening as the decision variable, wellhead pressure could be modelled and included in the model instead. Some operators prefer to base their decisions on pressure data and others on choke valve opening. The solution from the optimization problem would then be to which pressure each choke valve should be opened or closed instead of the actual choke opening as per now. Using pressure data could additionally provide greater control over the flow from each well, as this is not dependent on each valve's pressure profile.

Additional constraints could also be considered. This will be especially relevant if the optimization models are to be used on different petroleum fields. Two relevant suggestions are to include a constraint on water handling, and a constraint on the allowed amount of H_2S in each well flow. H_2S is unwanted because of its corroding properties, giving higher maintenance costs. If using the models in an even larger setting, constraints on for example routing and reservoir depletion plans can be considered included.

Testing of well and optimization models

The model building and optimization models are both formulated and implemented in a general manner. This implies that the models with minor changes can be applied to other similar problems. We suggest that the models are tested on other petroleum fields to get a better understanding of their applicability and results. More synthetic cases should also be considered to get a better picture of in which situations beneficial solutions can be expected. More specifically we suggest further testing on how different kurtoses and different combinations of variance, skewness and kurtosis in combination affect the value of a stochastic solution. Testing should also be performed on how the results of the well and optimization models change when the suggested implementation procedure is run iteratively.

At last, we suggest to test the models on a completely different application areas. An example could be within other production and processing industries (for example hydro-power), where you want to maximize some sort of production with some control variables, and where historical data can be used to build locally linear models for prediction.

Presentation of solution

In general, the presentation of the results should be improved so that they are easily understood by petroleum field operators and production engineers. A more compact and illustrative application could be developed. We also suggest that a few different solutions should be presented and compared. That is having the software presenting e.g. the three best solutions including their expected production improvement and its variance, and other properties. In that way an operator can make his or her own decision on which solution that is to be implemented.

Bibliography

Al-Othman, W. B., Lababidi, H. M., Alatiqi, I. M. and 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.

Alma, Ö. G. (2011). Comparison of robust regression methods in linear regression, *International Journal of Contemporary Mathematical Sciences* **6**(9): 409–421.

Andrews, D., Bickel, P., Hampel, F., Huber, P., Rogers, W. and J.W., T. (1972). *Robust estimates of location: survey and advances*, Princeton University Press.

Andrews, D. F. (1974). A robust method for multiple linear regression, *Technometrics* **16**(4): 523–531.

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

Awasthi, A., Sankaran, S., Nikolaou, M., Saputelli, L. and Mijares, G. (2007). Closing the gap between reservoir modeling and production optimization, *Digital Energy Conference and Exhibition*, Society of Petroleum Engineers.

Bakshi, A., Uniacke, E. and Dias, E. (2015). Lessons learned from world-wide deployments of model-based real time production optimization solutions, *SPE Digital Energy Conference and Exhibition*, Society of Petroleum Engineers.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems, *Numerische mathematik* **4**(1): 238–252.

Bieker, H. P., Slupphaug, O. and Johansen, T. A. (2006). Real time production optimization of offshore oil and gas production systems: A technology survey, *Intelligent Energy Conference and Exhibition*, Society of Petroleum Engineers.

Bieker, H. P., Slupphaug, O. and Johansen, T. A. (2007). Well management under uncertain gas or water oil ratios, *Digital Energy Conference and Exhibition*, Society of Petroleum Engineers.

Birge, J. R. (1997). State-of-the-art-survey-stochastic programming: Computation and applications, *INFORMS journal on computing* **9**(2): 111–133.

Birge, J. R. and Louveaux, F. (1997). *Introduction to Stochastic Programming*, Springer New York.

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

Brashear, J. P., Becker, A. B. and Gabriel, S. A. (1999). Interdependencies among e&p projects and portfolio risk management, *SPE annual technical conference*.

Camponogara, E., Plucenio, A., Teixeira, A. F. and 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): 157–167.

Campos, S. R., Teixeira, A. F., Vieira, L. F. and Sunjerga, S. (2010). Urucu field integrated production model, *SPE Intelligent Energy Conference and Exhibition*, Society of Petroleum Engineers.

Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*, Vol. 619, John Wiley & Sons.

Coats, K. H. (1980). An equation of state compositional model, *Society of Petroleum Engineers Journal* **20**(05): 363–376.

Cramer, R., Gob, K.-C. and Moncur, C. (2009). Real-time well optimization stabilizes, adds production, *Oil & Gas Journal* **107**(10): 41–41.

Dantzig, G. B. (1955). Linear programming under uncertainty, *Management science* **1**(3-4): 197–206.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*, Vol. 1, Cambridge university press.

de Weck, O., Eckert, C. M. and John, C. P. (2007). A classification of uncertainty for early product and system design, *Guidelines for a Decision Support Method Adapted to NPD Processes* .

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

Dutra, T. A., Machado, L., Rodriguez, M. A., Almeida, I., Montanha, B. N., Manzali, M., Dinis, M. G. M., Carbone, L., de Souza, M. F. and Nogueira Herdeiro, M. A. (2010). Monitoring and optimizing oil fields by a real-time production operation (rtpo) system, *SPE Latin American and Caribbean Petroleum Engineering Conference*, Society of Petroleum Engineers.

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

Dzubur, L. and Langvik, A. (2012). Optimization of oil production-applied to the marlim field, *Technical report*, Working Paper. Trondheim: Norwegian University of Science and Technology.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *The Annals of Statistics* **7**(1): 1–26.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*, CRC press.

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

Elgsæter, S., Slupphaug, O. and Johansen, T. A. (2008). Production optimization; system identification and uncertainty estimation, *Intelligent Energy Conference and Exhibition*, Society of Petroleum Engineers.

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

Fernandes, L. J., Relvas, S. and Barbosa-Póvoa, A. P. (2015). Petroleum supply chain network design and tactical planning with demand uncertainty, *Operations Research and Big Data*, Springer, pp. 59–66.

Fetkovich, M. (1973). The isochronal testing of oil wells, *Fall Meeting of the Society of Petroleum Engineers of AIME*, Society of Petroleum Engineers.

FMC Technologies (2015). Flowmanager, `http://www.fmctechnologies.`

com/en/SubseaSystems/Technologies/AdvancingTechnologies/ProductionMonitoringOpt/Flowmanagement/FlowManager.aspx.

Folgerø, K., Soldal, E. L., Kocbach, J., Frøysa, K.-E., Kleppe, K. and Åbro, E. (2013). Uncertainty analysis of multiphase flow meters used for allocation measurements: Field experiences and future challenges.

GDF SUEZ E&P Norge (2015). Gjøa, http://www.gdfsuezep.no/our-operations/gjoa.aspx?sc_lang=en.

Glæserud, P. and Syrdalen, J. A. (2009). Optimization of petroleum production under uncertainty: Applied to the troll c field.

Goel, V. and 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.

Grimstad, B. (2015). Daily production optimization for subsea production systems: Methods based on mathematical programming and surrogate modelling.

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

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

Gunnerud, V., Ursin-Holm, S. and Teixeira, A. (2013). Advanced optimization techniques applied to a petrobras fpso, *SPE Middle East Intelligent Energy Conference and Exhibition*, Society of Petroleum Engineers.

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

Haugen, K. K. (1996). A stochastic dynamic programming model for scheduling of offshore petroleum fields with resource uncertainty, *European Journal of Operational Research* **88**(1): 88–100.

Higle, J. L. (2005). Stochastic programming: optimization when uncertainty matters, *TutORials in operations research. INFORMS, Baltimore* pp. 30–53.

Høyland, K., Kaut, M. and Wallace, S. W. (2003). A heuristic for moment-matching scenario generation, *Computational optimization and applications* **24**(2-3): 169–185.

Huber, P. J. (1964). Robust estimation of a location parameter, *The Annals of Mathematical Statistics* **35**(1): 73–101.

Huber, P. J. (1981). *Robust Statistics*, Wiley, New York.

Huseby, A. B. and Haavardsson, N. F. (2009). Multi-reservoir production optimization, *European Journal of Operational Research* **199**(1): 236–251.

IO Center (n.d.). Solutions for efficient decision making and teamwork within planning in gdf suez e&p norge, `http://www.iocenter.no/system/files/sites/default/files/IO2/Presentations/GDF%20SUEZ.pdf`.

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

Jørgensen, K. O. (2008). Gjøa development and operations, `https://www.posccaesar.org/svn/pub/ISO15926AndSemanticWebTechnologiesNetwork/GDF-Suez.pdf`.

Kall, P. and Mayer, J. (1976). *Stochastic linear programming*, Vol. 21, Springer.

Kaut, M. (2003). Updates to the published version of a heuristic for moment-matching scenario generation by k. høyland, m. kaut, and sw wallace, *Computational Optimization and Applications* **24**: 169–185.

Keesman, K. J. (2011). *System identification: an introduction*, Springer Science & Business Media.

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

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection, *Ijcai*, Vol. 14, pp. 1137–1145.

Kosmidis, V. D., Perkins, J. D. and Pistikopoulos, E. N. (2005). A mixed integer optimization formulation for the well scheduling problem on petroleum fields, *Computers & Chemical Engineering* **29**(7): 1523–1541.

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

Ljung, L. (1998). *System identification*, Springer.

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

Marill, K. A. (2004). Advanced statistics: linear regression, part i: simple linear regression, *Academic emergency medicine* **11**(1): 87–93.

Mehrotra, S. and Papp, D. (2013). Generating moment matching scenarios using optimization techniques, *SIAM Journal on Optimization* **23**(2): 963–999.

Morken, M. L., Sandberg, P. T. and Ustvedt, S. (2015). Real-time optimization of petroleum production under uncertainty.

Mosteller, F. and Tukey, J. W. (1977). Data analysis and regression: a second course in statistics.

Neiro, S. M. and Pinto, J. M. (2003). Supply chain optimization of petroleum refinery complexes, *Proceedings of the fourth international conference on foundations of computer-aided process operations*, pp. 59–72.

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

Norwegian Petroleum Directorate (2015a). Norway's petroleum history, `http://www.norskpetroleum.no/en/framework/norways-petroleum-history/`.

Norwegian Petroleum Directorate (2015b). Production, `http://www.norskpetroleum.no/en/production/production-of-oil-and-gas/`.

Norwegian Petroleum Directorate (2015c). Reserves by field, `http://factpages.npd.no/factpages/Default.aspx?culture=en`.

Norwegian Petroleum Directorate (2016). Frontpage illustration, `http://www.npd.no/engelsk/cwi/pbl/field_jpgs/4467595_Vega.jpg`.

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

Nygreen, B. and Haugen, K. (2010). Applied mathematical programming in norwegian petroleum field and pipeline development: Some highlights from the last 30 years, *Energy, Natural Resources and Environmental Economics*, Springer, pp. 59–69.

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

Orman, M. and Duggan, T. (1999). Applying modern portfolio theory to upstream investment decision making, *Journal of petroleum technology* **51**(03): 50–53.

Perktold, J., Seabold, S. and Taylor, J. (2013). M-estimators for robust linear modeling, `http://statsmodels.sourceforge.net/devel/examples/notebooks/generated/robust_models.html`.

Petroleum Experts Ltd. (2015). Optimisation, `http://www.petex.com/products/?id=68`.

Ramdial, A., Hudson, N., Rodriguez, J. C. and Strathman, M. (2009). Real-time data, models optimize complex production off trinidad, *Oil & gas journal* **107**(17).

Ramsay, J. (1977). A comparative study of several robust estimates of slope, intercept, and scale in linear regression, *Journal of the American Statistical Association* pp. 608–615.

Richardson, J. and Shaw, A. H. (1982). Two-rate ipr testinga practical production tool, *Journal of Canadian Petroleum Technology* **21**(02).

Saputelli, L., Malki, H., Canelon, J. and Nikolaou, M. (2002). A critical overview of artificial neural network applications in the context of continuous oil field optimization, *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers.

Schlaifer, R. and Raiffa, H. (1961). Applied statistical decision theory.

Schlumberger Limited (2015). The oilfield glossary: Where the oil field meets the dictionary, `http://www.glossary.oilfield.slb.com/`.

Shakhsi-Niaei, M., Iranmanesh, S. H. and Torabi, S. A. (2013). A review of mathematical optimization applications in oil-and-gas upstream & midstream management, *International Journal of Energy and Statistics* **1**(02): 143–154.

Shalizi, C. (2013). Which bootstrap when?, `http://http://www.stat.cmu.edu/~cshalizi/uADA/13/lectures/which-bootstrap-when.pdf`.

Shamlou, S., Gunnerud, V. and Conn, A. R. (2012). Optimization of petroleum production networks–through proxy models and structural constraints', *8th IFAC International Symposium on Advanced Control of Chemical Processes, Advanced Control of Chemical Processes, Singapore*, pp. 543–548.

Shapiro, A. and Philpott, A. (2007). A tutorial on stochastic programming.

Solution Seeker AS with Vidar Gunnerud (2015). Interview held: 2015-10-22, personal communication.

Statoil ASA (2006). Plan for utbygging, anlegg og drift av gjøafeltet, `http://www.statoil.com/no/EnvironmentSociety/Environment/impactassessments/fielddevelopments/Downloads/Plan%20for%20utbygging,%20anlegg%20og%20driftav%20Gj%C3%B8afeltet%20-%20Utvinningstillatelse%20PL%20153%20-%20Del%202%20Konsekvensutredning%20-%20Oktober%202006.pdf`.

Stoisits, R., Crawford, K., MacAllister, D., McCormack, M., Lawal, A. and Ogbe, D. (1999). Production optimization at the kuparuk river field utilizing neural networks and genetic algorithms, *SPE mid-continent operations symposium*, pp. 391–399.

Thai, H.-T., Mentré, F., Holford, N. H., Veyrat-Follet, C. and Comets, E. (2013). A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models, *Pharmaceutical statistics* **12**(3): 129–140.

Thorn, R., Johansen, G. and Hjertaker, B. (2013). Three-phase flow measurement in the petroleum industry, *Measurement Science and Technology* **24**(1): 012003.

Tjønnås, J. and Schjølberg, I. (2011). Simple near well models for fast generation of dynamic gor.

Trangenstein, J. A. and 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. and Sagli, J. R. (2007). Tactical planning of offshore petroleum production, *European Journal of Operational Research* **176**(1): 550–564.

Ursin-Holm, S. and Shamlou, S. S. (2013). Simulation based optimization of petroleum production problems: Development of a special purpose b&b for a non-convex minlp.

van der Linden, R., Marck, J. and de Boer, J. (2015). Results of real-time production optimization of a maturing north sea gas asset with production constraints, *IFAC-PapersOnLine* **48**(6): 45–49.

Walls, M. R. (2004). Combining decision analysis and portfolio management to improve project selection in the exploration and production firm, *Journal of Petroleum Science and Engineering* **44**(1): 55–65.

Walpole, R., Myers, R., Myers, S. and Ye, K. (2012). *Probabilty & Statistics for Engineers & Scientists*, Pearson Education.

Walter, E. and Pronzato, L. (1997). Identification of parametric models, *Communications and Control Engineering* **8**.

Wang, P. (2003). *Development and applications of production optimization techniques for petroleum fields*, PhD thesis, Stanford University.

Wang, P., Litvak, M. and Aziz, K. (2002). Optimization of production operations in petroleum fields, *SPE Annual Technical Conference and Exhibition*, Society of Petroleum Engineers.

Weatherford (2015a). Reo production simulation and optimization software, `http://www.weatherford.com/doc/wft095013`.

Weatherford (2015b). Wellflo well-analysis software, `http://www.weatherford.com/doc/wft095007`.

Wehrens, R., Putter, H. and Buydens, L. M. (2000). The bootstrap: a tutorial, *Chemometrics and intelligent laboratory systems* **54**(1): 35–52.

Wintershall Holding GmbH (2015). Production: Ambitious plans, `https://www.wintershall.no/production.html`.

Yanez, N. D., Kronmal, R. A. and Shemanski, L. R. (1998). The effects of measurement error in response variables and tests of association of explanatory variables in change models, *Statistics in medicine* **17**(22): 2597–2606.

Appendix A

Electronical Attachments

The electronical folder attached contains code and data related to:

1. Data Pre-Processing, Data Modelling, Reading and Writing of Data, Scenario Generation etc.

    - `main.py`: the main script, used for starting the data model building (bootstrapping), scenario generation and to execute tests.

    - `tests.py`: a collection of all the tests that are run for the technical study and case study.

    - `data_modelling`: a package containing the files used for modelling of the data.

        - `bootstrap.py`: contains methods for *residual resampling* and for *case resampling*.

    - `my_collections`: a package of custom collections used in these scripts.

        - `petroleum_data.py`: storing processed Gjøa data.

    - `scenario_generation`: a package containing scripts used for scenario generation.

        - `moment_matching.py`: contains the main loop of the moment matching algorithm.

        - `cubic_transformation.py`: the cubic transformation part of the moment matching algorithm.

  - `random_sampling.py`: scenario generation by random sampling from the normal distribution.

- `input_output`: a package containing all scripts used for reading data from files, or writing data to files.
  - `data_io.py`: used to read the data sets from Gjøa.
  - `model_io.py`: used to read and write built data models.
  - `mosel_io.py`: used to write mosel-input files.

2. Mosel MILP-models
   - `strategy_v2.mos`: the original strategy model
   - `strategy_v3.mos`: the alternative formulation of the strategy model
   - `penalty.mos`: the penalty model

Appendix B

Regression plots

In this appendix, figures illustrating regression plots for each well are presented. Figure B.0.1 and Figure B.0.2 show the regression lines for gas parameters found with and without data compression respectively. Figure B.0.3 and Figure B.0.4 show the regression lines for oil parameters found with and without data compression respectively.

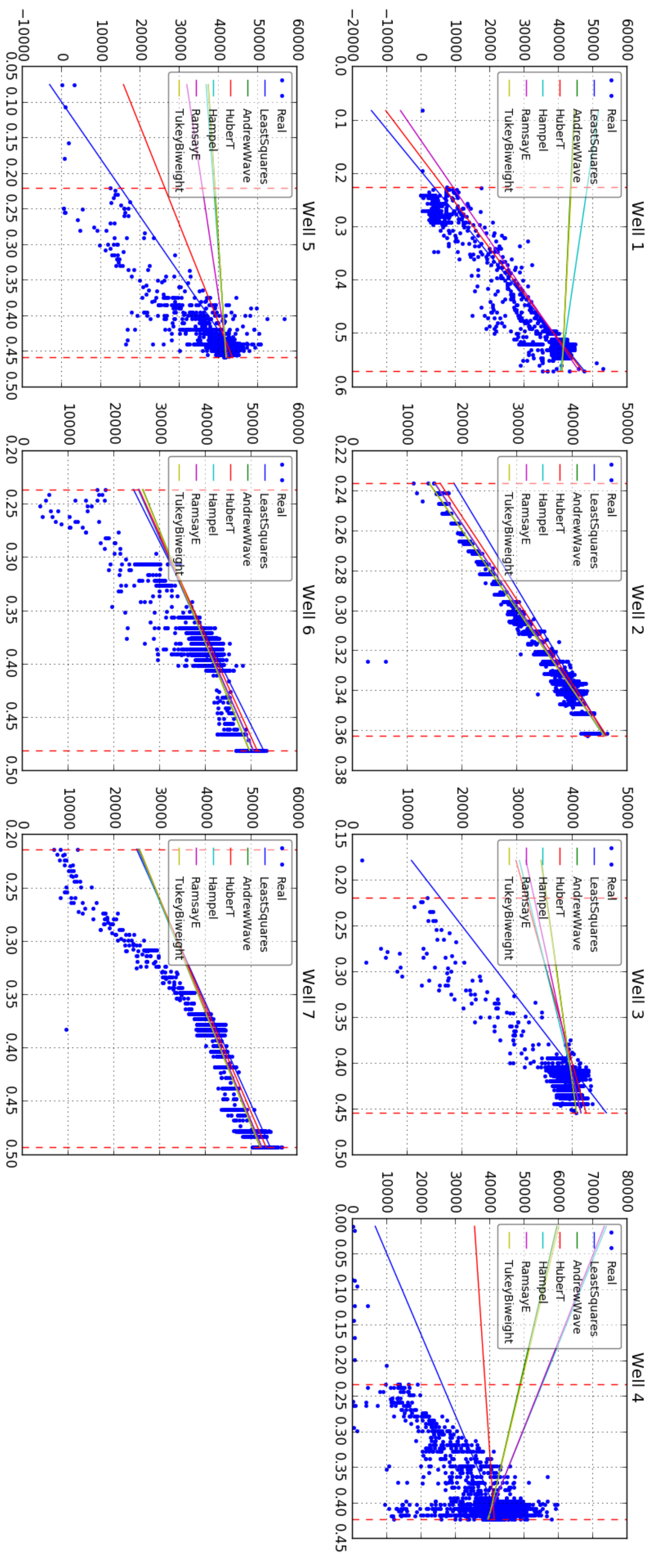


Figure B.0.1: Regression, no compression, gas plots

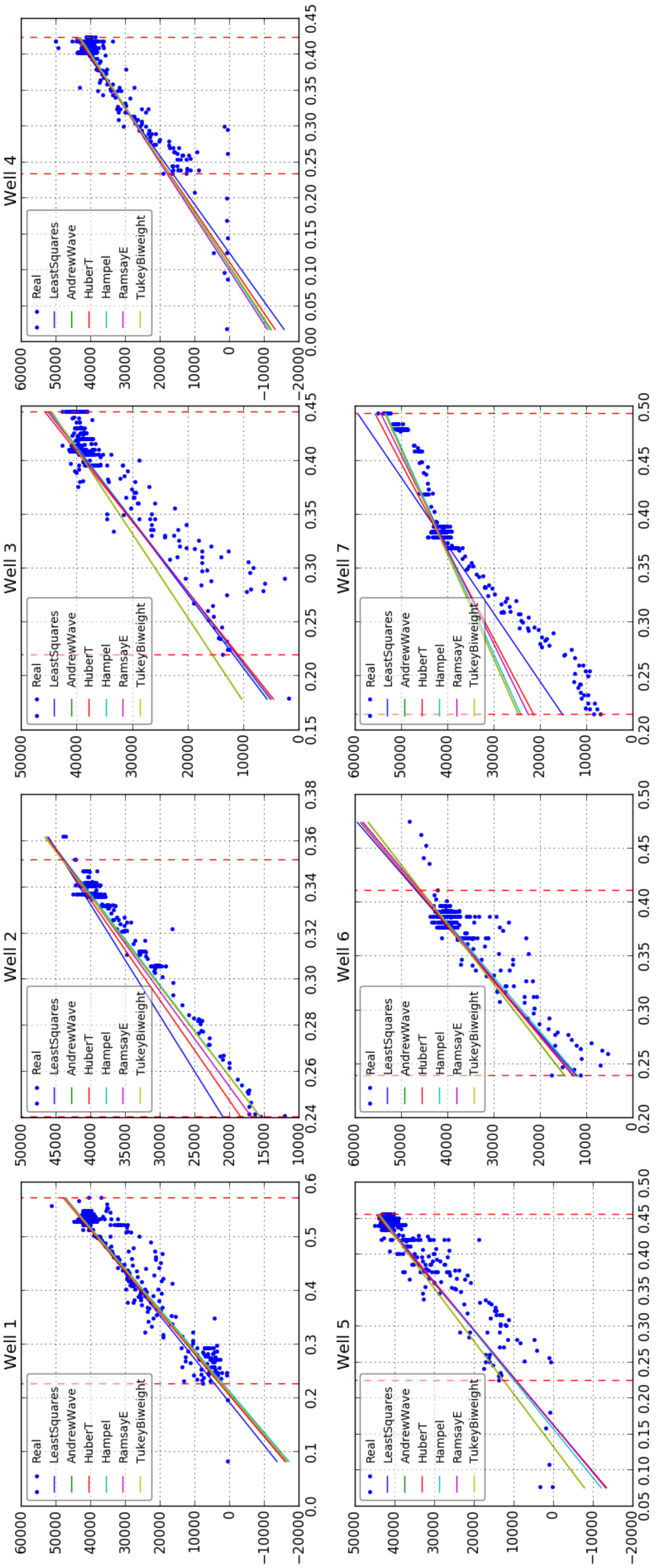


Figure B.0.2: Regression, with compression, gas plots

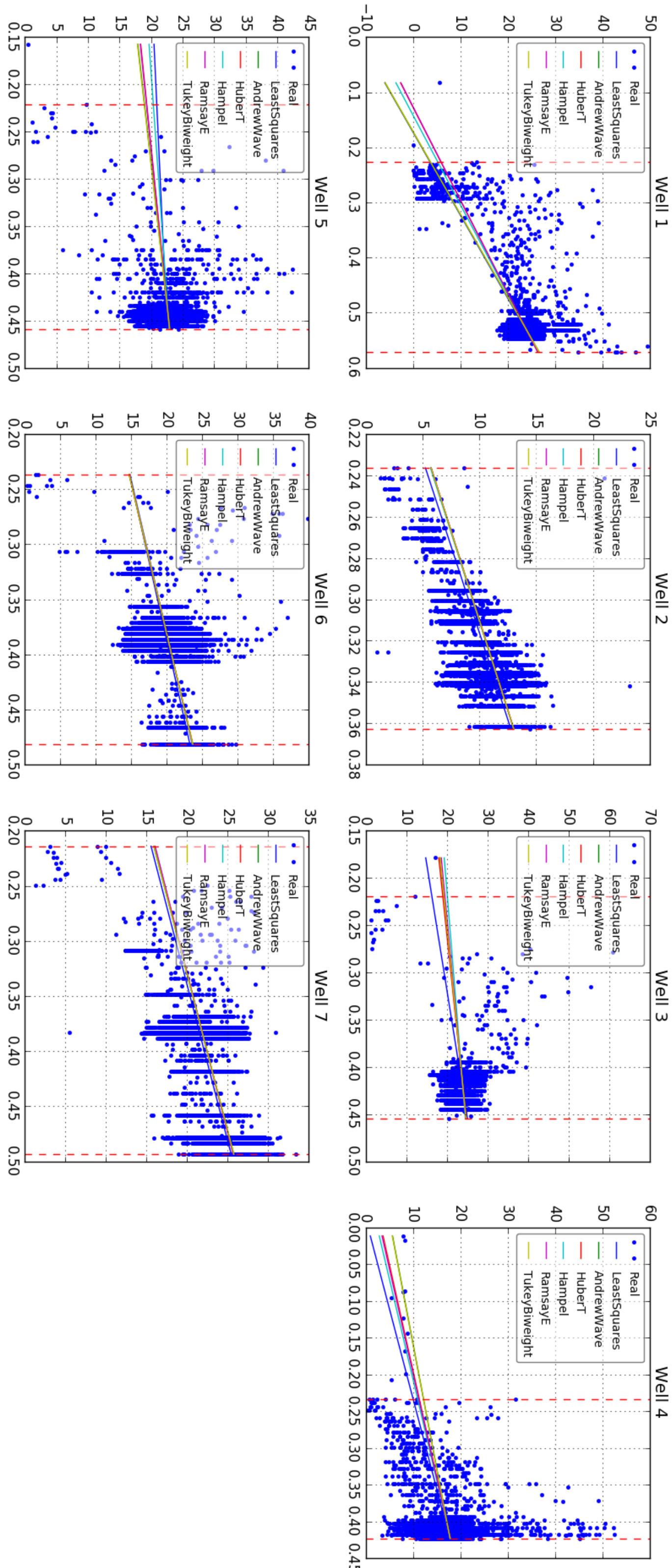


Figure B.0.3: Regression, no compression, oil plots

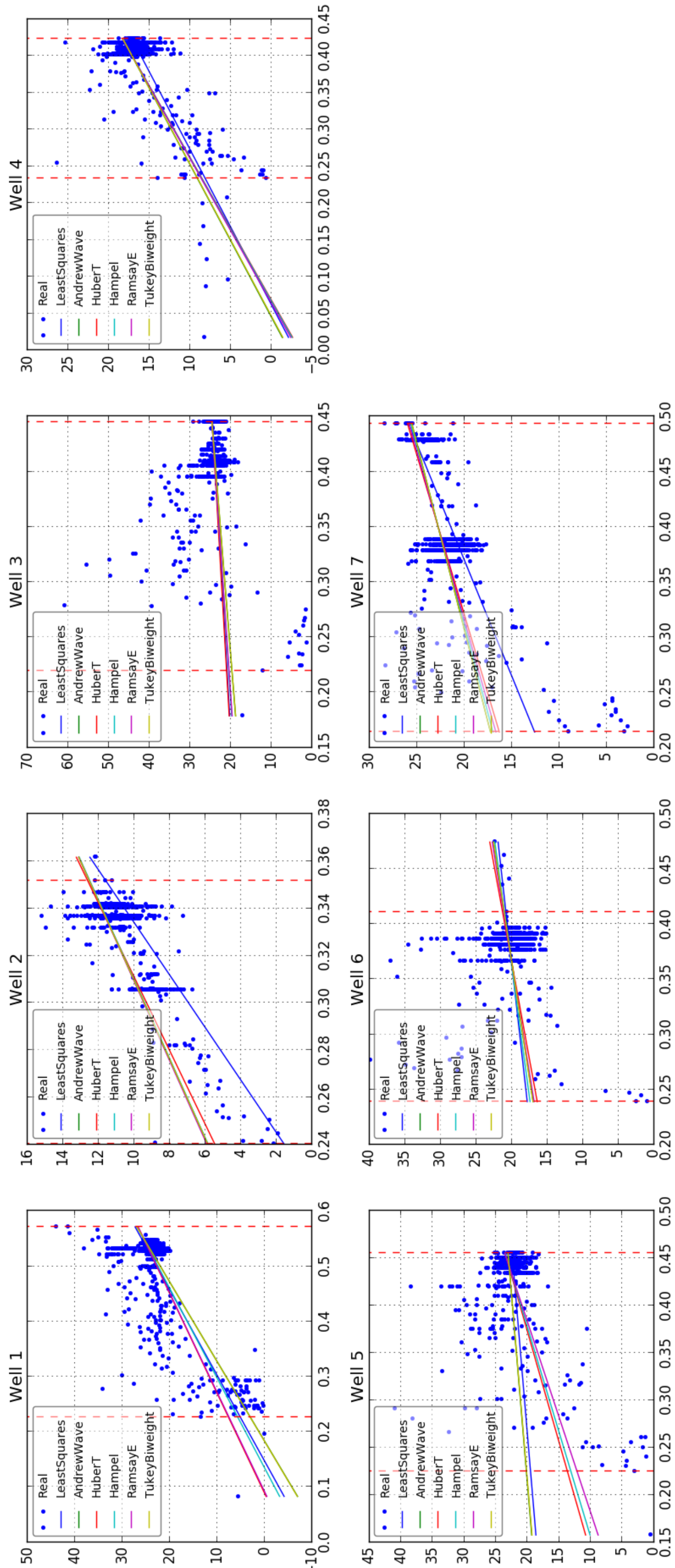


Figure B.0.4: Regression, with compression, oil plots

Appendix C

Export plots, all data

In this appendix, figures illustrating the gas and oil export prediction are presented. The predictions found using each of the linear regression methods are plotted against actual real production over a three month time period.

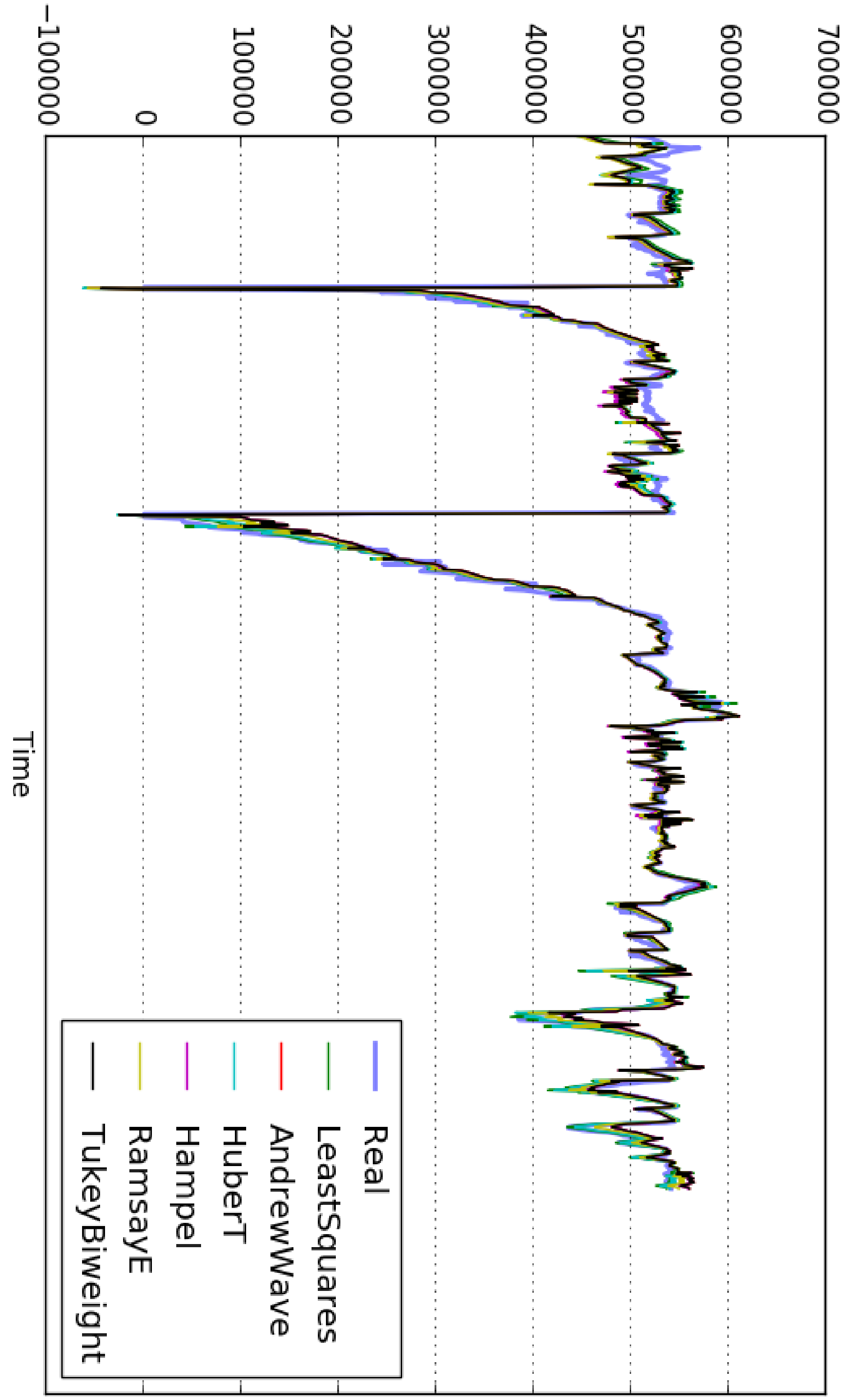


Figure C.0.1: Full gas export prediction

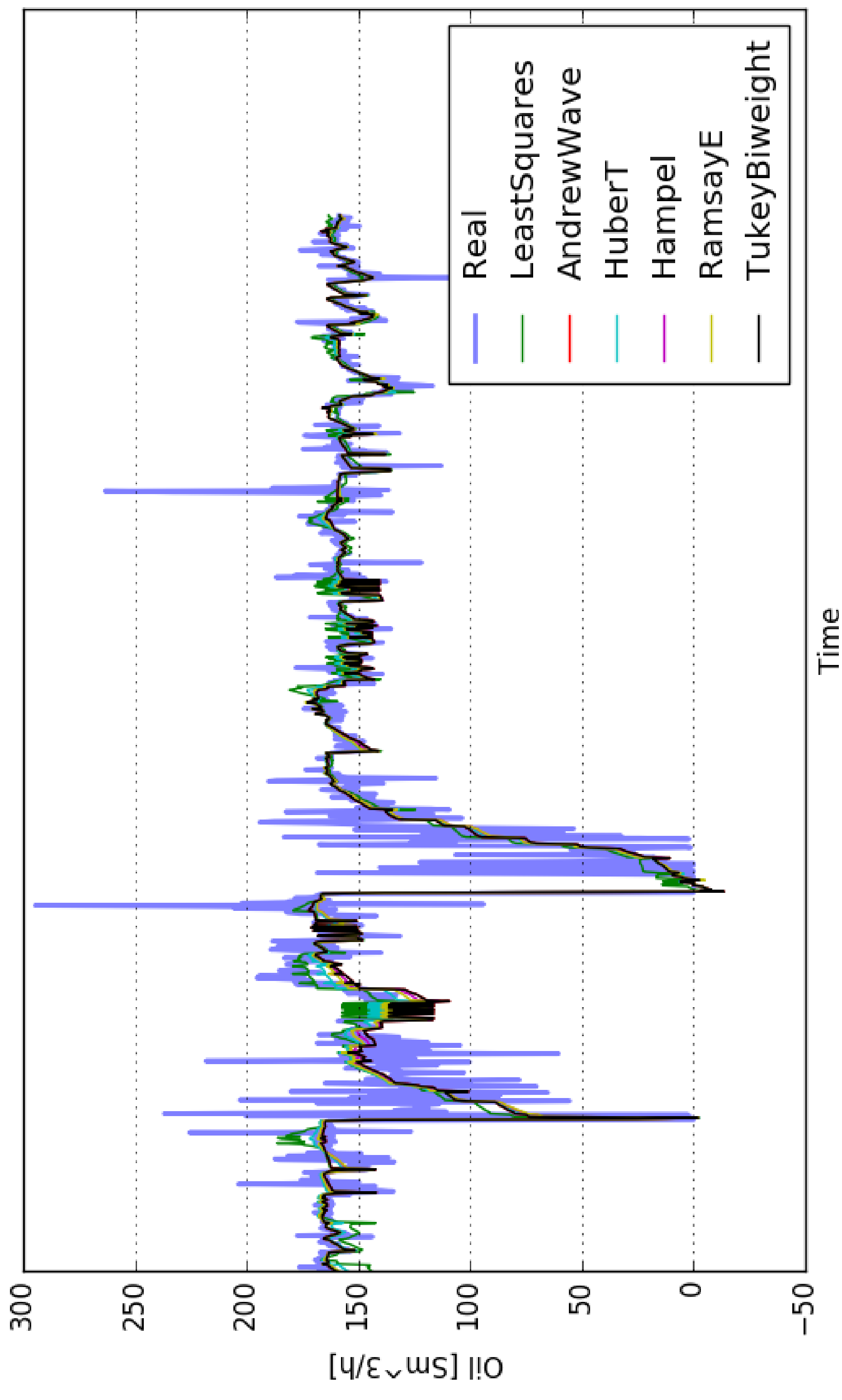


Figure C.0.2: Full oil export prediction

Appendix D

Bootstrap distributions

In this appendix, figures illustrating the distributions for the constants and slopes of oil and gas prediction are presented. The distributions are obtained through the residual resampling bootstrap procedure described in Chapter 5. The red lines show a normal distribution.

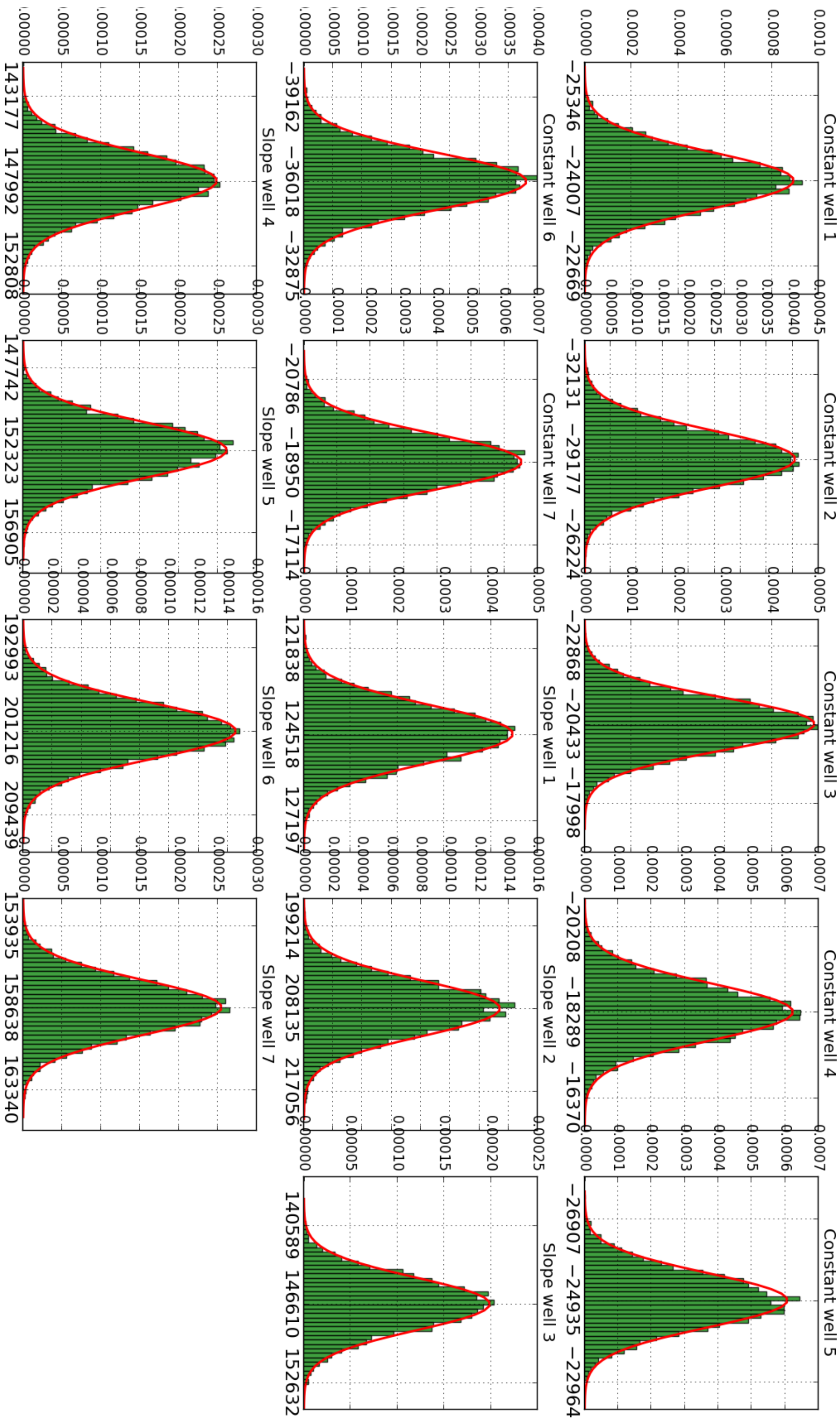


Figure D.0.1: Bootstrap distributions for gas constants and slopes [x-axis: μ and 3σ]

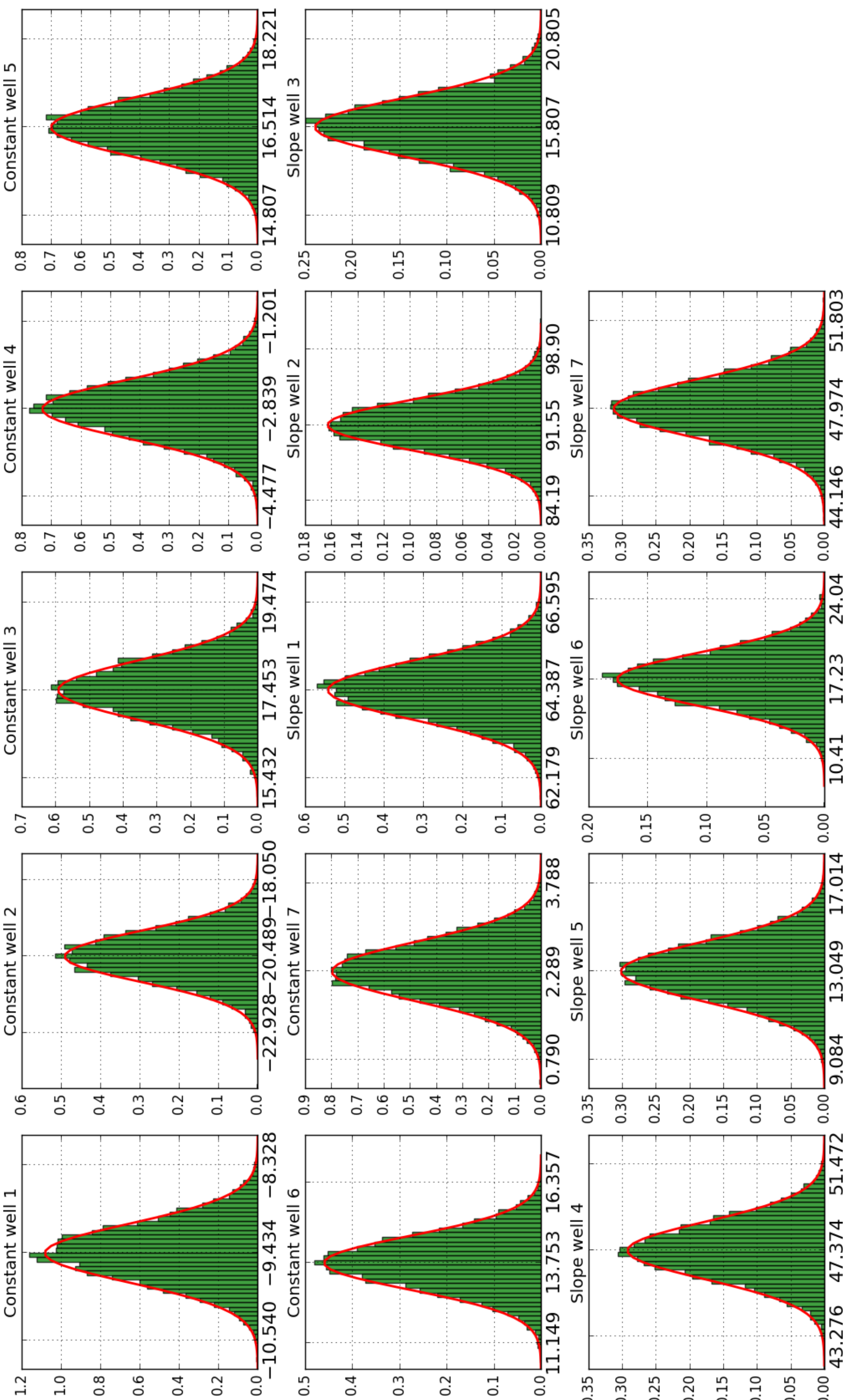


Figure D.0.2: Bootstrap distributions for oil constants and slopes [x-axis: μ and 3σ]