MASTER'S THESIS

Robust pipeline flow scheduling at an oil company

Author:
Sven B.J. Oosterhuis

Master Programme:
Industrial Engineering and Management

Specialization Track:
Production and Logistics Management

Supervisors:
Dr. ir. J.M.J. Schutten (University of Twente)
Prof. dr. J.L. Hurink (University of Twente)
Drs. J.S. de Wit (ORTEC)
P.H.W. Bom, MSc. (ORTEC)

Graduation date:
10-04-2015

In cooperation with:

UNIVERSITY OF TWENTE.

ORTEC
ADDING CREDIBILITY

Preface

I proudly present you my thesis, which is the result of seven months of hard work and marks the end of my student life. As an Industrial Engineering and Management student, I developed an increasing interest in mathematical methods for optimizing business processes. To increase my knowledge I attended some courses from the master's program Applied Mathematics, resulting in a master's thesis focussed on optimizing oil and gas flows. By doing my internship at the ORTEC Consulting Group, I hoped to find out if I wanted to start my career in the consulting business field, which I definitely did. I really enjoyed my time at ORTEC and certainly learned a lot, as my thesis was part of a new project in which I got a lot of freedom to include my own ideas. I learned that the consulting business is an exiting, dynamic, but also an uncertain business, as deadlines for model demonstrations showed up out of nothing, but really had to be met.

Unfortunately, business does not always go as planned, which resulted in a lack of information that definitely did not make the project easier. However, I am certain that the result is something to be proud of, as my developed model is a well performing and relevant model for the business. Hopefully a customer will be found in the near future, so the work can be applied in reality.

This thesis would not have been there without the help of my supervisors, ORTEC colleagues and fellow students. I especially want to thank my ORTEC supervisors Jan de Wit and Philip Bom, who both made me feel welcome at ORTEC from the first day. I really enjoyed our meetings in which we discussed, or philosophised as Jan would sometimes call it, which assumptions had to be made for the model and which aspects had to be included. I also want to thank my ORTEC colleague Wim Kuijsten for helping me out a lot in my everlasting struggle with AIMMS. Furthermore, I want to thank my supervisors from the university of Twente Marco Schutten and Johann Hurink who gave me (a lot of) valuable feedback that brought this thesis to a much higher level than I would have accomplished on my own.

Sven

Zoetermeer, March 2015

Summary

The goal of this thesis is to develop a mathematical model that optimizes the oil and gas flows of the International Company for Onshore Oil Operations (OilCO) and that supports operational decision making when disturbances occur. This mathematical model serves as a basis for the Oil-OperationsOptimizer (O3) that OilCO wants to implement for the management of its day to day oil and gas movements. The main question of our research is:

> *What is the best mathematical model for the optimization of OilCO's oil and gas flows including dealing with disturbances?*

The main goal of the O3 is to generate a schedule with a 30-day horizon which specifies the production rates of the reservoirs and the throughput rates of the reservoirs against minimal costs. The assets of OilCO that we model are part of a pipeline network where oil and gas are extracted from the earth at reservoirs, after which both flows are separated at separation facilities. Subsequently the oil is transported by pipelines to the customers, which are refineries and terminals. Both terminals and separation facilities have storage tanks, in which oil can be stored. Each type of asset has it own operational constraints that we consider in our model for the O3. The costs involved in the O3 are pumping costs, penalties for non-satisfied demand, penalties for deviating from the production targets, changeover costs at the reservoir, inventory costs, and penalties for having less on stock than the safety stock. The disturbances the O3 needs to deal with are: opportunities on the spot market, the scheduling of maintenance, asset breakdowns, and uncertainty in tanker arrivals.

The designed model for the O3 is a Mixed-Integer Linear Program based on an event based time representation. In an event based time representation the planning horizon is divided into time periods of unequal length based on pre-defined events that change the system, which in our case are the arrival and departure of tankers at the terminals and the start and end of a day. We model OilCO's network as generic as possible by modelling each type of asset using the same constraints, which allows us to make changes in the network if needed. In the objective function we assume that the energy costs have a non-linear relation with the throughput rate. We approximated this relation by creating an example of a pump, which is based on two existing models. Moreover, we propose a simplification for our model primarily based on the reduction of the number of integer variables, which turns out to decrease the average running time by 69%.

To be able to deal with disturbances, we extend the developed deterministic model. We indicate how the model can be used in a what-if analysis to deal with asset breakdowns, the scheduling of maintenance, and spot market opportunities. Furthermore, we apply RO to deal with the uncertainty in tanker arrivals by generating different scenarios of tanker arrival realizations as input for the model. Since the number of tanker arrival realizations is excessively large, we use the Sample Average Approximation method in which we minimize the average of a representative sample. Our model variables are divided into design and control variables. The design variables are variables of which the values cannot be changed when the tanker arrival moments get known, while the control variables can be changed and therefore depend on the realization. The robust model minimizes the sum of the cost resulting from the design variables and the sample average of the costs resulting from the control variables, together with a term for model robustness and a term for solution robustness. Model robustness is the extend to which a model is "almost" feasible for every realization of the uncertain parameters, where solution robustness is the extend to which the solution is "close" to optimal for every realization.

In the experiments we test a tanker arrival case in which the demand is stable and a case in which demand is less stable and has a peak. Results show that the deterministic model is capable of

generating a schedule very fast. However, the deterministic model only considers one tanker arrival scenario, which makes the solution inaccurate in some cases. The robust model considers 80 tanker arrival scenarios and therefore yields a solution which is on average cheaper for all possible scenarios. In the robust model it is possible to increase both model robustness and solution robustness at the prize of increasing total costs. Robustness is achieved by deviating more in the throughput rates of the pipelines to the customers, resulting in having less infeasibilities in the inventory levels of the terminal storage tanks. When robustness is further increased, shortages at the terminals occur when the model decides to transport less oil to the terminals to avoid infeasibilities.

All in all, our model is capable of generating an optimal or robust schedule with a 30-day horizon in a couple of minutes. Howerever, during this project we did not have direct contact with OilCO, so many assumptions are made about cost relations and parameter values. We formulate several alternatives for the model in case our assumptions about cost relations turn out to be incorrect, so the model can be adjusted quickly. Moreover, we perform a sensitivity analysis on the most important parameters of which we are uncertain about the value: the storage tank sizes, the peak efficiency throughput and the cost factors. The sensitivity analysis shows that the values of certain parameters have a large influence on the solution, but is also shows that our model is capable of dealing with different input values. In the future the model can be extended with the scheduling of the tankers at the terminals and the scheduling of maintenance.

Contents

1 Introduction

This thesis is written as part of an internship at ORTEC Consulting Group (OCG) in Zoetermeer. The graduation assignment involves a current project of OCG done in cooperation with ITCo, which is a large International IT company that received a request for proposal from the International Company for Onshore Oil Operations (OilCO) to design a company-wide system that manages its oil and gas flows on a day to day basis. ITCo consulted OCG to develop the optimization module of the company-wide system, which is called the OilOperationsOptimizer or O3. This graduation project involves the development of a mathematical model for the optimization module. Below, we first introduce the key players in this project: OCG (Section 1.1) and OilCO (Section 1.2). As ITCo is only an intermediary company between the key players, we not further introduce ITCo. Section 1.3 introduces the focus of our project, the O3 and Section 1.4 provides a research plan for the development of the mathematical model for the O3 optimization module.

1.1 ORTEC Consulting Group

OCG is a business unit within the ORTEC Group, where it combines business knowledge, operations research and IT in their products and services. It currently has around 150 consultants employed and operates in various industries, such as the airline industry, the oil and gas industry, and postal services. The company offers numerous optimization solutions in the field of process industry optimization, revenue management, supply chain optimization, and site logistics optimization. This research takes place at the Oil, Gas and Chemicals (OGC) department within OCG, which is currently involved in the tender for the development of a company-wide system for scheduling oil and gas movements at OilCO.

1.2 International Company for Onshore Oil Operations

OilCO is an oil and gas company that operates onshore and in shallow coastal water in several countries around the world. OilCO operates in the upstream and midstream oil business. The upstream business involves the extraction of oil and gas from the earth, while the midstream business covers the processing and transportation of crude oil products. OilCO produces its oil and gas from ten big oil reservoirs, which all contain around 300 oil wells. When the oil is pumped out of a well, it contains a fraction of gas, which is called associated gas[1]. The associated gas is separated from the oil at one of the four separation facilities near the reservoirs, after which the oil is transported to the customers via a large pipeline network. OilCO supplies the produced oil to two different sorts of customers: terminals and refineries. Two major terminals, called Terminal1 and Terminal2, export oil with oil tankers. Furthermore, there are two refineries called Refinery1 and Refinery2 that are also customers of OilCO's oil. The refineries are owned by the International Oil Refining Company. The length of OilCO's oil pipelines varies from a few kilometres up to around 400 kilometres (the pipeline to Terminal2), resulting in a total network of more than 1,500 km of pipelines. The size of the network means that the pumping costs of pumping oil through the pipelines are high.

OilCO wants to implement a company-wide system that is capable of optimizing the oil and gas production at the oil reservoirs and its transportation to meet the contractual agreements with their three customer groups (terminals, International Oil Refining Company, and BuyGas), which

[1] Next to associated gas, OilCO also extracts non-associated gas (NAG) from gas fields. Both types of gas are sold to BuyGas, which is an international gas company.

is currently done based on operational experience. The new company-wide system, the OilOperationsOptimizer, or O3, will be used for the real time scheduling of oil and gas production and distribution.

1.3 The O3

The O3 will be a company-wide system that will have an advisory role in the day to day management of the oil and gas movements of OilCO. As such, it should be able to recommend on the optimization of the production and pipeline movements collectively. The time frame in which the O3 optimizes OilCO's oil and gas flows is 30 days. The input of the O3 will be the current production plan, which states the production outputs of oil and gas per day. The main objective for the O3 is to recommend on production rates, separation rates, and throughput rates such that all customer demand is met against minimal costs. Large penalties for non-satisfied demand are included in the model as these are part of the contracts between OilCO and its customers. Furthermore, the changeover costs for changing the production rate of an oil reservoir are important, meaning that the production quantities of the oil reservoirs need to be as steady as possible. The changeover costs have to be minimized, together with the pumping costs, inventory costs at the storage tanks, and penalties for non-satisfied demand. Finally, the O3 has to take deviations from the original production plan into account.

The O3 has to be capable of answering what-if questions in case of disturbances concerning the daily operations. The model has to optimize oil and gas flows when a single disturbance or combinations of disturbances occur, which can occur both on the demand side (tanker is delayed or extra profit can be gained on the spot market) as well as on the supply side (breakdown of a plant, maintenance of a pipeline). The system has to recommend based on comparing different actions that could be taken. The recommendation could be a change in schedule such as a reallocation of oil or a temporary increase in production. Besides such a reactive scheduling approach, the O3 also needs to take tanker arrival uncertainties into account, which leads to a predictive scheduling approach.

The scope of research is the development of a mathematical model for the optimization module of the O3, which includes all functional requirements mentioned above and has to be implemented in the ORTEC Supply Chain Scheduling Tool, when finished. The running time of the model has to be reasonable, i.e. several minutes, since OilCO uses the O3 on a daily basis, which is a challenge given the size of the network, the interdependence of oil and gas flows, and the different what-if questions that have to be examined.

1.4 Research plan

The goal of this research is:

> *to develop a mathematical model that optimizes OilCO's oil and gas flows including operational decision making support when disturbances occur*

In order to achieve this goal, we have formulated research questions of which the main question is:

Main question: What is the best mathematical model for the optimization of OilCO's oil and gas flows including dealing with disturbances?

Since it is impossible to answer this question at once, we formulate several sub questions of which we treat one per chapter. First we need to describe the processes of OilCO, which is done by examining OilCO's oil and gas movements. Here, we want to know what we exactly have to consider for the mathematical model. Furthermore, we consider the functional requirements of OilCO for the O3 to find out what the exact requirements for our model are. Since OCG has no direct contact with OilCO, information is gathered from the functional design specification of the O3, the website of OilCO, and experience of OCG in similar projects.

Chapter 2: What are OilCO's relevant processes and what are the associated requirements for the O3?

- **Section 2.1:** Which processes are associated with OilCO's oil and gas movements?
- **Section 2.2:** What are OilCO's functional requirements for the O3?
- **Section 2.3:** Where do we lack information?

When we know the requirements for the O3, we review the existing mathematical models that cover pipeline flow scheduling. In order to develop a suitable model for OilCO's processes, more knowledge is required about how models for similar situations are formulated. Furthermore we look at the concept of Robust Optimization, since our model needs to take uncertainties into account when it is solved.

Chapter 3: Which pipeline flow scheduling models and Robust Optimization models exist that can be used for modelling OilCO's processes?

- **Section 3.1:** What optimization models for pipeline flow scheduling exist in literature?
- **Section 3.2:** Which of these models are useful for OilCO's optimization problem?
- **Section 3.3:** What is Robust Optimization and how is it applied in the upstream and midstream oil business?

With the information about existing models from Chapter 3 and the requirements resulting from Chapter 2, we formulate step by step an optimization model as the basis of the O3. Moreover, we use the experience of OCG on modelling similar processes.

Chapter 4: What is a suitable model formulation for the O3?

- **Section 4.1:** How to represent time in our model?
- **Section 4.2:** What assumptions have to be made?
- **Section 4.3:** How to model OilCO's oil and gas network?
- **Section 4.4:** How to formulate OilCO's processes in constraints for an optimization model?
- **Section 4.5:** What is a suitable objective function for the optimization model?
- **Section 4.6:** What are the consequences of the lack of information and how can we deal with them?
- **Section 4.7:** How to simplify the model to boost performance?

After formulating the model, we implement it in the ORTEC Supply Chain Scheduling tool. Our model needs to be able to handle disturbances by generating and evaluating what-if scenarios. To achieve this the model may have to be solved multiple times with different inputs and additional restrictions or some heuristic may have to be designed. We extend the model with a what-if analysis, such that it can be used to deal with disturbances. Moreover, we look at how to deal with the uncertainty in demand in our model, where we use the Robust Optimization knowledge we gathered in Section 3.3.

Chapter 5: How to use the proposed model formulation in order to deal with disturbances?

Section 5.1: How to use the model for what-if analysis?

Section 5.2: How to take the uncertainty in demand into account by using Robust Optimization?

In order to measure the performance of our model for the O3, we test it. First we test the deterministic model from Chapter 4 and then compare it with the robust model from Section 5.2. Furthermore, we test the simplification from Section 4.7, to see if it really boosts performance.

Chapter 6: What is the performance of the developed model?

Section 6.1: What are useful parameter values for the experiments?

Section 6.2: What is the performance of the deterministic model?

Section 6.3: What are the differences between the deterministic model and the robust model?

Section 6.4: What is the performance of the simplified model?

As indicated in Section 2.3, there is a lack of information. We examine the consequences of this lack of information on the model's solution and performance by doing a sensitivity analysis on uncertain parameters.

Chapter 7: How sensitive is our model to changes in uncertain parameters?

Finally, after all sub questions have been answered, Chapter 8 discusses the main question, a conclusion and we also provide recommendations for further research.

2 Situation description

This chapter describes the requirements for the O3 in terms of processes and assets it has to consider and functional requirements. We take a look at the current processes of OilCO in Section 2.1. A challenge in this respect is that ORTEC does not have direct contact with OilCO, but has contact with OilCO via ITCo. Section 2.2 considers the functional requirements OilCO that has formulated for our model. Section 2.3 examines the uncertainties in information that are caused by having no contact with OilCO, whereas Section 2.4 summarizes this chapter.

2.1 OilCO's oil and gas movements

To develop a model, information is required about the processes involving oil and gas movements. In this section we follow oil and gas from the reservoir to the customer, where we only discuss parts of the process that have to be modelled in the O3. The O3 considers the following units and customers: the reservoirs (Section 2.1.1), separation facilities (Section 2.1.2), pipelines (Section 2.1.3), refineries (Section 2.1.4), terminals (Section 2.1.5), storage tanks (Section 2.1.6), and BuyGas (Section 2.1.7).

2.1.1 Oil/Gas reservoirs

Oil is extracted from several oil reservoirs, which also contain (associated) gas. The proportion of this associated gas depends on the characteristics of the reservoir and is presented in the Gas Oil Ratio (GOR). The unit of the GOR is cubic feet per barrel (cf/bbl), as oil volumes are measured in barrels (bbl) and gas volumes in cubic feet (cf).

In total, 10 oil reservoirs are considered for the O3, which are the biggest 10 that OilCO exploits. These reservoirs provide almost all of the 2.5 million bbl per day that OilCO produces. The reservoirs are clustered into four so called production assets: Asset1, Asset2, Asset3, and Asset4. Table 1 shows an overview of the division of the production over the ten oil reservoirs and production assets. Each reservoir has a minimum and maximum capacity of oil that it can produce per day.

Production planning is currently done by the corporate planning department using production plans with a rolling horizon. The corporate plan is a long term plan that specifies of the amount of oil and gas to be extracted from the reservoirs and that must be in line with the shareholder agreements. The current planning model generates plans with different rolling horizons of which a rolling five year plan is the highest level plan. The automated planning model disaggregates this plan into a rolling one year and a three months plan, which is then further disaggregated into a short term operational rolling plan of one month which specifies the daily targets. The daily targets are the amounts of oil and associated gas that a certain reservoir should produce on a certain day.

The strategy of OilCO on production volumes is created in cooperation with its shareholders, who want OilCO to follow the strategy and therefore minimize deviations from the production targets. The O3 therefore has to take deviations from the production plan into account when generating a schedule, along with the changeover costs. Changeover costs occur when the production rate is changed, as this implies that individual wells, of which each reservoir has approximately 300, have to be started up or shut down.

2.1.2 Separation facilities

The mixture of oil and gas is transported from the reservoirs to a separation facility where the oil and gas flows are separated, because they have to be transported to different customers. The separation

Production Asset	Name of reservoir	Regular production per day (bbl)	Asset Total (bbl)
Asset1	Reservoir0	800,000	
	Reservoir1	152,000	
	Reservoir2	108,000	
	Reservoir3	60,000	
			1,120,000
Asset2	Reservoir4	713,000	
	Reservoir5	38,000	
	Reservoir6	66,000	
			817,000
Asset3	Reservoir7	84,000	
	Reservoir8	55,000	
			139,000
Asset4	Reservoir9	465,000	
			465,000
Grand Total			**2,541,000**

Table 1: Daily oil production (bbl) per reservoir and asset

process is optimized at the asset level by the CurrentSystem (CS), which is a local system that OilCO currently uses. The O3 has to operate at a company wide level rather than at the asset level, so the O3 complements the CS. Each production asset has its own separation facility, such that OilCO has four separation facilities in total. The separation facilities are called SepFac1, SepFac2, SepFac3, and SepFac4, where the number corresponds to the asset at which the facility is located. All separation facilities have storage tanks in which oil can be stored after separation. Gas is directly transported to the different BuyGas facilities, as these are close to the separation facilities. The separation facilities have a maximum daily separation capacity. Facilities can be in maintenance for a certain period in which they are unable to process at this maximum capacity. The availability percentage of a separation facility indicates which percentage of the daily maximum separation capacity can used. During the separation process 3% of the oil is lost, which is a factor that the O3 has to take into account.

2.1.3 Pipelines

All oil and gas transport takes place via unidirectional pipelines, meaning that oil and gas can only flow in one direction. Each pipeline has its own capacity, which is the maximum throughput rate it can handle. As oil in a pipeline cannot be compressed, the oil flow that enters a pipeline is equal to the flow that leaves the pipeline. Some pipelines in the network are connected to a pipeline node, which is a junction of pipelines where flows split up and/or merge. Here, the pipeline node near Asset4 is called Node 1 and the pipeline node near Refinery1 is called Node 2. Table 8 in Appendix A gives an overview of all pipelines in OilCO's network. Some locations are connected by multiple parallel pipelines, as can be seen in the overview.

Oil is transported through the pipelines using pumps. Pipelines of significant length require more than one pump to keep the oil flowing. The pumping costs depend on the throughput rate of a pipeline and are one of the cost components that have to be minimized. However, there are some units on the same location (e.g. the Reservoir9 and SepFac4) that are connected by pipelines that are so short that pumping costs are irrelevant and therefore have a length of 0 in the overview. Unfortunately, the information in the table is incomplete, so we have to make assumptions. If we do not know the capacity of a pipeline, we assume that the capacity is infinite. For some pipelines the capacity is insufficient with respect to the production. For example, the production of reservoir Reservoir6 is 66,000 bbl per day, while the capacity of the pipeline from Reservoir6 to SepFac2 is 36,000 bbl per day. We adjust the capacities of these pipelines, of which the result can be found in the column “Adjusted capacity” in the pipeline overview of Appendix A.

2.1.4 Refineries

OilCO delivers oil to two refineries (Refinery1 and Refinery2) that are owned by the International Oil Refinining Company. Both refineries have a pre-specified fixed oil demand per day and a penalty cost occurs for each barrel of oil that is undelivered. Furthermore, it is not possible to deliver more than given demand to the refinery, because the refinery can not handle the extra oil.

2.1.5 Terminals

The two terminals, Terminal1 and Terminal2, are also customers of OilCO's oil. Both terminals have loading platforms, where Terminal1 has three and Terminal2 has two. At these platforms oil is loaded into tankers with a rate of 50,000 bbl/hour. OilCO has contracts with oil tankers, which arrive in a three day window and have to pay a big fine when they are too late. The three day period is pre-arranged to provide tankers with flexibility in their travel times, since tankers have to cover long distances. At least three days before the beginning of three day window, tankers have to confirm an exact arrival date. The loading time of a tanker, which is usually around a day, consists of a constant component for berthing and unberthing and a variable component. The variable loading time, which depends on the size of the tanker, is calculated by dividing the total tanker volume by 50,000 bbl/hour. In the contracts with the tankers a deficit cost is included, which OilCO has to pay when demand is unfulfilled or fulfilled too late. Oil tankers can switch between the preassigned terminals due to the large distance (±500km) between the terminals. Both terminals also have storage tanks where oil can be stored for later use (more details, see Section 2.1.6). Next to that, oil is sold on the spot market, which is an opportunity for gaining extra profits. The spot market demand is also delivered via tankers.

2.1.6 Storage Tanks

As described above, some of OilCO's assets have storage tanks for oil. Storage tanks have a minimum and a maximum inventory limit that must be taken into account. A minimum inventory limit exists, since the storage tanks have floating roofs which are damaged if the tank gets empty. Next to the minimum inventory limit, storage tanks have a safety stock that must be taken into account, where it is possible to have less in stock than the safety stock against certain costs (so called safety stock penalties). Storage tanks can be used to cope with short term issues, e.g. the delay of a tanker. However, storing oil bears inventory costs, which is one of the cost components that has to be minimized.

2.1.7 BuyGas

BuyGas, which is OilCO's gas customer, operates multiple gas processing facilities located close to OilCO's separation facilities. In addition to associated gas from oil reservoirs, OilCO exploits non-associated gas to fulfill the gas demand, which is extracted from gas fields and therefore does not have to be separated from oil. This non-associated gas is transported to the BuyGas facilities in the same way as the associated gas. The O3 has to determine the volumes of associated and non-associated gas that are used to fulfill BuyGas' demand. The BuyGas facilities processes gas by extracting all its usable components. From the BuyGas facilities lean gas, which is gas without any usable components, is transported back to the oil reservoirs, since each reservoir requires an amount of gas injected per day in order to maintain pressure in the reservoir. In case there is a deficit in lean gas returned from BuyGas, nitrogen or other products are injected into the reservoir.

2.2 Functional requirements

Currently, OilCO uses a plan based on operational experience to match the demand from their three customer groups (terminals, the International Oil Refining Company, and BuyGas) to their supply capabilities. The O3 aims to replace this plan in the future by optimizing the scheduling of oil and gas movements in a time horizon of 30 days, which is complicated by uncertainty in the arrival times of the tankers at the terminals. Moreover, the schedule must minimize total costs of which pumping costs, penalties for non-satisfied demand, penalties for deviating from the production targets, changeover costs at the reservoir, inventory costs, and safety stock penalties are important cost components. Extra gains can be generated by selling oil on the spot market. As stated earlier, the O3 does not take decisions automatically, but recommends to the user. Specifically, the O3 advices on the following decisions:

- the production rate of the different oil reservoirs and associated separation facilities;
- the throughput rate of the different pipelines in the network

The inventory levels of the different storage tanks are a direct result from these decisions, since inventory levels depend on the ingoing and outgoing throughput rates of a tank. The inventory levels are therefore auxiliary variables.

The O3 must be able to handle disturbances as well, as it must consider delayed tankers, opportunities on spot market, maintenance of assets, and asset breakdowns. These disturbances are dealt with using both predictive scheduling as well as reactive scheduling. Predictive scheduling means that the disturbance is taken into account when the schedule is created and reactive scheduling means that the disturbance is taken into account after a schedule is realized. The O3 has to perform reactive scheduling on opportunities on spot market, maintenance of assets and asset breakdowns, as these disturbances do not occur frequently and are difficult to predict, meaning that preventive scheduling is not the best way of dealing with these disturbances. One could argue that for the scheduling of maintenance of assets predictive maintenance is possible, but this is out of scope for the O3. On these disturbances what-if analysis is applied, meaning that the O3 has to calculate the consequences of different reactions on disturbances, compare the outcomes, and finally present the best possible action to be taken. The recommendations need to be such that demand is still satisfied (by re-allocating the oil) when possible.

For the uncertainty in tanker arrivals we look at predictive scheduling techniques, since it is hard to predict the exact arrival time of a tanker due to the fact that tankers arrive in a three day window. The O3 has to yield solutions that are, by design, more capable of dealing with different arrival moments. We apply the method called Robust Optimization (RO) to our model to achieve that. Therefore, we look at RO techniques that are relevant for the O3 in Chapter 3.

Finally the O3 has to come up with an schedule within a couple of minutes, as the system has to be used in daily operations.

2.3 Information uncertainty

Unfortunately we do not have contact with OilCO during this project, as stated in the introduction of this chapter. In this section we indicate what the consequences are and how to deal with them. We split the uncertainty into process uncertainty (Section 2.3.1) and data uncertainty (Section 2.3.2). The process uncertainty is uncertainty about the processes that is modelled and data uncertainty is uncertainty about the parameter values of the model.

2.3.1 Process uncertainty

For some of OilCO's processes it is unclear what they exactly look like, meaning that some assumptions have to be made about how things exactly work at OilCO. The assumptions on how the energy and changeover costs should be modelled have a large impact on the model, since these are important cost components which our model has to minimize, but we do not know exactly what causes these costs and which factors have impact on them. For example, we do not know if the size of a changeover (which is the change in production rate) influences the changeover cost, or if the changeover costs are constant for each changeover. We examine how these cost components are modelled in similar models by performing a literature review in Chapter 3. Based on this literature review and the experience within the Oil, Gas and Chemicals department of OCG we make assumptions about how we model the cost components. Furthermore, in Section 4.6 we indicate what the consequences are if our assumptions are incorrect, which makes adjusting the model easier when more is known about the processes.

2.3.2 Data uncertainty

The data we have available for this thesis is incomplete and uncertain, as the pipeline overview in Appendix A indicates. The main goal of this thesis is to develop a mathematical model, which is possible without knowing the exact parameter values. We, however, require data in Chapter 6 to perform numerical experiments on our model to see how our model performs and make possible adjustments to the model if needed. We make educated guesses for the values of the parameters in the model to be able to do the experiments. There may be, however, some parameters (e.g. storage tank size) of which the value is uncertain, but that have a strong influence on the optimal schedule and model performance. In Chapter 7 we identify these parameters and perform a sensitivity analysis on them to examine the influence of the value of the parameter on the model.

There is also uncertainty regarding the outlook of OilCO's network. It could be that there are more storage tanks in the model, or that some pipeline connections are different from the pipeline overview in Appendix A. We want to deal with this by formulating a model that is as generic as possible, so the network can be changed easily. Every type of unit and customer is modelled using the same variables and constraints, such that a network layout that is different from OilCO's (but includes the same sort of units) can also be modelled using our model formulation.

2.4 Summary

In this chapter we examined OilCO's processes and the functional requirements for the O3, such that our mathematical model considers all relevant factors. It became clear that our model should consider a pipeline network of production facilities, separation facilities, pipelines, refineries, terminals, and storage tanks. Each type of asset has it own operational constraints that have to be considered. The O3 has to generate a schedule with a 30 day horizon with minimal total costs, which consists of pumping costs, penalties for non-satisfied demand, penalties for deviating from the production targets, changeover costs at the reservoir, inventory costs, and safety stock penalties. The schedule involves all assets mentioned and has to be generated in reasonable computation time. The most important decisions that the schedule has to specify are the processing rates of the reservoirs and throughput rates of pipelines in the network.

The O3 has to be able to handle disturbances by using both predictive as reactive scheduling techniques. The opportunities on the spot market, the scheduling of maintenance and asset breakdowns are dealt with by using what-if analysis and for the uncertainty in tanker arrivals we look at Robust Optimization (RO). In Chapter 3, a literature review is performed on models for situations similar to OilCO's.

3 Literature review

In this chapter we look into literature in order to get an overview of existing optimization models on pipeline flow scheduling and RO techniques. The sub question that this chapter answers is: "Which pipeline flow scheduling models and Robust Optimization models exist that can be used for modelling OilCO's processes?" The goal of the first part of this literature review is to get insight in how OilCO's pipeline network can be modelled, how the different assets can be modelled, and how the different cost components can be modelled. The research focuses on scheduling problems, since OilCO's optimization problem is on the scheduling of the production rates and throughput rates. Although the O3 has to schedule crude oil and gas flows, some models about other products (for example petroleum or water) may also be interesting as long as pipeline transportation is involved. The main focus of this literature review is on the scheduling of liquid flows, since most of the flows that have to be modelled are oil flows. We search, however, for literature about liquid pipeline scheduling problems that also consider gas flows.

OilCO operates in the upstream and midstream oil business. The upstream business involves the extraction of oil and gas from the earth and the midstream business involves the processing of crude oil and transportation of crude oil or oil products. Articles on optimization of downstream oil activities, which involves the distribution of refined products, are therefore omitted. We limit ourselves to upstream and midstream oil business in which pipelines are the main means of transportation, which means that articles that only cover shipment of crude oil by tankers are omitted. Section 3.1 gives an overview of existing optimization models and Section 3.2 indicates which of these models are relevant for OilCO.

Next to that, we look at RO techniques to deal with the uncertainty in tanker arrivals. Section 3.3 covers an overview of RO techniques and applications in the upstream and midstream oil business. Section 3.4 presents the conclusions of our literature review.

3.1 Optimization models for pipeline flow scheduling

Almost all literature on optimization models for scheduling flows of products through pipelines originates from the oil industry. These models not only consider crude oil, but also various refined oil products. Water is the only other product present in literature (Abbasi & Garousi, 2010). Mathematical models are used to support decision making that minimizes cost (Lee et al., 1996; Rejowski & Pinto, 2003; Moro & Pinto, 2004) or maximizes profit (Mas & Pinto, 2003; Reddy et al., 2004a; Pan et al., 2009). Most scheduling problems concerning a network of oil pipelines optimize the oil flow schedules, which contain information on the volume, the timing, and the destination of the oil that is pumped into the pipelines, which are decisions that are also relevant for OilCO's optimization problem.

Typical types of cost components that are minimized in the models are: pumping costs (Rejowski & Pinto, 2008; Herrán et al., 2010), inventory costs (in case storage tanks are considered) (Lee et al., 1996; Cafaro & Cerdá, 2012), and changeover costs (in case different products are considered, note that this is a different definition than changing the production rate) (Reddy et al., 2004b). All these are cost components are also relevant for OilCO's optimization problem. The cost components that are considered as well as the method by which they are determined differs per model.

When a capturing a business process as OilCO's into a mathematical model, often some assumptions have to be made. An important assumption that is made in every oil pipeline optimization model: the pipeline remains completely full with incompressible oil products at any time (Cafaro &

Cerdá, 2008b). The only way to get a volume of oil out of a certain pipeline is to inject an equal volume of oil at the origin, which means that the oil flow into a pipeline has to be equal to the flow out of the pipeline, and that there always is an initial volume of oil in the pipeline. This also holds for points in the network where pipelines split up or merge (MirHassani et al., 2011).

Literature on pipeline flow scheduling concerns both optimization of single pipelines (Rejowski & Pinto, 2003; Relvas et al., 2013) and networks of pipelines (Stebel et al., 2012; Cafaro & Cerdá, 2012). We are mainly interested in networks, since this best reflects OilCO's situation. In the literature the network is represented by a set of nodes that are connected by pipelines, and sometimes by using a graph formulation (Abbasi & Garousi, 2010). Nodes in the network can be supply nodes (or sources), demand nodes (or sinks), and intermediate nodes through which flow is transported. Flow is produced at the supply nodes, consumed at the demand nodes, and conserverd at the intermediate nodes, meaning that the incoming flow must be equal to the outgoing flow. In some models it is also possible to change the characteristics of the flow (Neiro & Pinto, 2004) or store flow (Rejowski & Pinto, 2008) at a node.

When an optimization model provides a schedule for oil flows, the way the time component is modelled is an important choice, because it influences the model basics and the performance of the model. Time can be represented in a discrete or a continuous way, which both can be found in literature. In a discrete time representation, the time horizon of the schedule is divided into time slots, which are of equal length in the original discrete time representation. (Lee et al., 1996; Rejowski & Pinto, 2003). The schedule determines what activities take place in which time periods and calculates inventories for the beginning and end of each period. A discrete formulation uses binary variables to indicate if an activity occurs in a time slot. To obtain a high accuracy for a discrete time formulation, a large number of time slots has to be used (Reddy et al., 2004a), which directly affects the number of binary variables in the model and therefore increases computation time. A continuous time representation means that the start and end times of activities are continuous variables (Mas & Pinto, 2003) and, in many cases, that the volumes of product batches are also modelled in as continuous variables (Relvas et al., 2006; Cafaro & Cerdá, 2008b). Some recent models also contain a hybrid version of both representations called an event based time representation, which is a continuous time representation by using discrete time intervals with unequal length (Saharidis et al., 2009). In a event based time representation the model variables are assumed constant when no external changes occur. The time horizon is therefore divided into events instead of hours, meaning that an interval is defined as a period between the moments an event starts and finishes. These events are predefined external events which change important model parameters, e.g. the arrival of a customer tanker, which causes that the model variables are only evaluated on moments that are relevant for the schedule and reduces the number of discrete time intervals.

Most work around the optimization of oil pipeline networks is done around two scheduling problems: the Crude Oil Operations Scheduling Problem and the Multi-product Pipeline Scheduling Problem. We look into these problems and associated models in more detail first. Section 3.1.1 considers the Crude Oil Operations Scheduling Problem and Section 3.1.2 considers the Multi-product Pipeline Scheduling Problem. Section 3.1.3 presents an overview of other existing models.

3.1.1 The Crude Oil Operations Scheduling Problem

Problem description

The Crude Oil Operations Scheduling Problem first received attention in literature in 1996 (Lee et al., 1996; Shah, 1996). The problem considers the unloading of crude oil tankers to deliver crude oil to an oil refinery. The tankers are mostly Very Large Crude Carriers (VLCCs), which are unloaded in ports at unloading berths. The moment they unload depends on the availability of the different berths, such that waiting time may occur. The unloaded oil is transported via pipelines to storage tanks in which the crude oil is stored. Since the refinery demands specific crude oil blends, the oil from different storage tanks is combined into a charging tank via pipelines. These charging tanks feed the Crude Distillation Units (CDUs) of the oil refinery. Figure 1 shows a graphical overview of the problem.

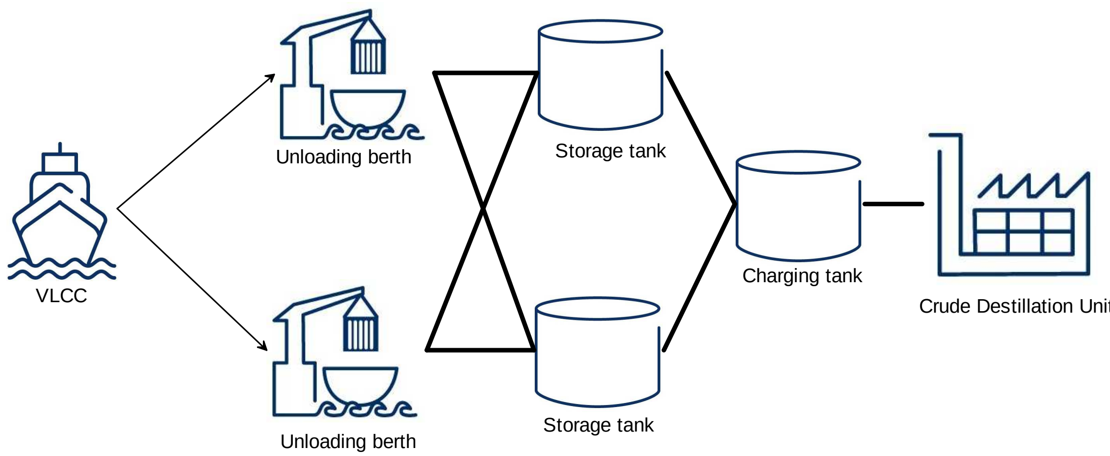

Figure 1: Graphical overview of the Crude Oil Scheduling Problem

The problem is to determine a schedule that specifies:

- The waiting time for each tanker in the sea
- Unloading duration time for each tanker
- The crude unloading rates from the tankers to the storage tanks
- The crude oil transfer rate and mixing rate from storage tank to charging tanks
- Inventory levels of all tanks involved
- CDU charging rates
- Sequence of mixed crude to be charged into each CDU

The different types of equipment have capacity limitations: pipelines have a maximum throughput that they can handle, storage and charging tanks have a maximum storage level, and unloading berths have a limited availability. The goal is to find a schedule that minimizes operating costs within the operational restrictions.

Solution methods for the problem

For the Crude Oil Operations Scheduling Problem, models using both discrete and continuous time representations can be found in literature. Lee et al. (1996) and Shah (1996) propose a discrete time formulation with time periods of equal length. Lee et al. (1996) assume that there is only one unloading berth at the port and ignores changeover times for the transition between two different types of crude mixes at the CDU. Lee et al. (1996) assume that perfect mixing occurs when crudes from different storage tanks are combined in one charging tank. Only specific key components in crude or mixed oil are considered, because calculating the properties takes a lot of computational time due to non-linear equations. Shah (1996) only considers one type of tanks, which receive crude oil from the tankers as well as charge it to the CDU. Shah (1996) decomposes the model in two sub models, an upstream and a downstream model and defines certain types of crudes on beforehand. Both papers propose a Mixed-Integer Linear Program (MILP) to solve the problem with mass balance equations at the end of each interval. The objective function of the model of Lee et al. (1996) minimizes the sum of the unloading costs, the waiting costs of tankers, the inventory costs of the tanks, and the changeover costs of the CDU, whereas the model of Shah (1996) minimizes the total value of the crude oil in storage when it is no longer consumed by a CDU.

Reddy et al. (2004b) extend the model of Lee et al. (1996) by allowing multiple parcels (separate compartments) of crude oil per tanker, which are unloaded via a single-buoy mooring pipeline and/or single parcel tankers. The properties of crude mixes are considered, which adds non-linear equations to the model. The model uses a combination between continuous and discrete time representation, as multiple activities can occur within one discrete time slot. The non-linear properties result in a Mixed-Integer Non-linear Program (MINLP) which is relaxed into a MILP. The relaxation results in crudes that are not mixed in the right mixture. The model maximizes profit as the margins per crude are included and also a penalty for having less in stock than the safety stock is included.

Mas & Pinto (2003) present the first continuous time representation of the problem, meaning that the start and end time of unloading and charging events are continuous variables. The model is decomposed into several sub models: a port model and a substation model (where the storage tanks are). First an assignment of tankers to unloading platforms is generated in the port model and subsequently a schedule for the loading and unloading operations of tanks and pipelines is generated using the substation model. Both sub models are modelled using a MILP formulation. The port model maximizes profit that consists of the revenue minus the crude oil cost minus the cost of utilizing the unloading platform and minus the interface costs (which separates two different types of crudes) and overstay costs (which are costs for delay) of the tankers. The substation model minimizes operating costs that consists of loading/unloading costs of the tanks and again the interface costs.

Reddy et al. (2004a) also present a continuous time formulation, which solves the same problem variant as solved by Lee et al. (1996) with as only difference that storage and charging tanks are combined and CDUs are loaded directly from the storage tanks, meaning that multiple tanks can feed a single CDU at the same time. The authors compare their model with the model of Reddy et al. (2004b) with discrete time representation, but encounter difficulties because of the differences in formulations. Reddy et al. (2004a) conclude that the discrete time representation outperforms the continuous time representation for smaller and more complex problems. Reddy et al. (2004a), however, expect that continuous time representations has the best potential of being the best choice in the future, as their model is the first one that uses a continuous time representation and it has to be developed further.

Jia & Ierapetritou (2004) use a decomposition approach to solve a similar problem as Mas & Pinto (2003) do. They extend the model with a product blending and delivery part, which occurs after the processing of the CDU. All sub models are solved using MILP formulations with a continuous time representation, which is compared with the discrete time model of Lee et al. (1996) and conclude that their continuous time model solves significantly faster.

Moro & Pinto (2004) also consider the crude oil operations from a continuous time perspective, but focus on the properties of the crude oil as different crudes are mixed. They consider the non-linear aspects of a model that takes the different oil properties into account, as Reddy et al. (2004b) also do. The model of Moro & Pinto (2004) maximizes the CDU feed flow rate while minimizing the costs associated with tank operations, where tanker unloading costs and pipeline transfer costs are omitted. Moro & Pinto (2004) propose both a MINLP formulation and a derived MILP formulation with linearized constraints. The authors conclude that the MINLP model is able to generate a more efficient schedule for a short term horizon.

Pan et al. (2009) base their work on the work of Reddy et al. (2004a) and try to come up with a more efficient model. A new continuous time model is proposed, which has non-linear equations due to the modelling of blending of crudes with different properties. The authors propose a heuristic to determine an unloading procedure and indirectly calculate the crude composition and properties in each tank, which results in a problem that can be modelled using linear equations. Their model outperforms the model of Reddy et al. (2004a), as it requires fewer binary variables and requires less computational time.

Saharidis et al. (2009) are the first to introduce the event based time representation (which is explained in Section 3.1). The authors show that an event based formulation outperforms discrete time formulation regarding computational time needed. The paper considers several modes of blending and various recipe preparations to cope with the non-linear mixing constraints. Their model, which minimizes the number of set-ups needed, assumes that a single oil tank can contain only one type of crude oil at a time.

Yadav & Shaik (2012) also propose an event-based time representation for the problem, based on a simplified state-task-network (STN) formulation which corresponds to a MINLP formulation. The non-linear constraints are relaxed to reduce computational effort, although it may give composition discrepancies of crude blends. The authors provide three different formulations of oil flows in a tank based on whether mixing is allowed or not, whether simultaneous input and output is allowed or not and whether bypassing (which happens when a volume of oil passes a storage tank without entering it) is allowed or not.

Not only Mixed-Integer Programs are proposed in literature, as Adhitya et al. (2007) propose a heuristic that can be used to reschedule operations when disruptions occur. Generating a new optimal schedule typically requires significantly large amounts of time, which is undesirable when disruptions require a fast response. Furthermore, a disruption causes the input of an optimization model to change, which can result in severe changes in a schedule. The authors propose a heuristic that overcomes both these shortcomings. The heuristic decomposes the schedule in certain “operation blocks”, which are rescheduled when a disruption occurs. Another advantage of the proposed heuristic is that it generates multiple feasible schedules instead of only one optimal schedule.

Wang & Rong (2010) propose a two-stage robust model to solve the crude oil scheduling problem under uncertain conditions. The first stage of the model is developed using chance-constrained programming and fuzzy programming that can be transformed into the deterministic counterpart problem, whereas the second-stage is scenario-based. Through the combination of approaches, the model can deal with uncertain parameters with both continuous and discrete probability distributions. The article primarily focuses on tanker arrival uncertainty and CDU charging demands uncertainty.

3.1.2 The Multi-product Pipeline Scheduling Problem

Problem description

The Multi-product Pipeline Scheduling Problem received first attention in 2003 (Rejowski & Pinto, 2003). Where the Crude Oil Scheduling Problem considers the processes of transporting crude oil to an oil refinery, the Multi-product Pipeline Scheduling Problem considers the process from the refinery to the customers. A refinery produces petroleum products that have to be transported to customers via a single pipeline, which is divided into segments with at the end of each segment a tank depot that is connected to a customer market. In the pipeline, each segment either transfers products to the depots or to the next customers. At the refinery, products have to be stored in dedicated tanks, which prevents mixing of crudes. Furthermore, an interface material is pumped into the pipeline between batches of different products to prevent mixing of crudes. Typical operational costs that have to be minimized are inventory costs at the different storage tanks, pumping costs and interface costs for the interface material between different products in the pipeline. Figure 2 shows a graphical overview of the problem.

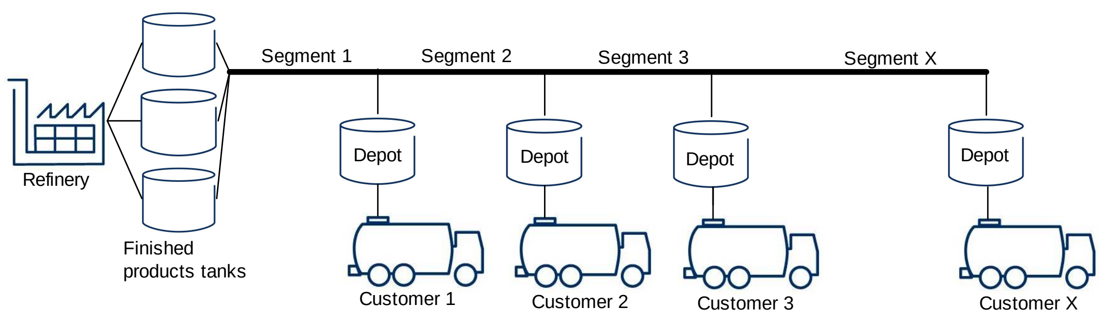

Figure 2: Graphical overview of the Multiproduct Pipeline Scheduling Problem

The problem consists of finding a schedule that specifies:

- Timing and volume of all batches that are pumped into the pipeline
- Distribution of batches among the different depots
- Locations and sequence of the different batches in the pipeline
- Inventory levels of all tanks involved

The operational constraints that have to be considered are the maximum inventory levels of the different tanks, the capacity of the pipeline, the maximum production rate of the refinery and the fact that only one product can be pumped into the pipeline at a time. The main assumption is the same as in the crude oil scheduling problem, namely that the pipeline should remain completely full at all times. It is also assumed that all finished products have constant densities.

Solution methods for the problem

Rejowski & Pinto (2003) are the first to treat the Multi-product Pipeline Scheduling Problem in literature. They propose a MILP formulation to solve the problem, which uses a discrete time formulation that also divides the pipeline into packs of equal volume. The volume of a certain product that is discharged from the finished products tanks into the pipeline is a multiplication of the volume of a pack (which is a constant) and the number of time slots that a discharge takes. The volumes at the different tanks are calculated using mass balance equations at the end of each time slot. This model still has drawbacks as it only allows one unloading operation in the entire system at a time. The model minimizes the sum of inventory costs of the tanks, the interface costs between different batches and the pumping costs. One year later, the same authors improved their model by adding some practical constraints (Rejowski & Pinto, 2004), which imply that the pumping of oil in a segment of the pipeline can only be stopped if the segment stores exactly one type of product. Also integer cuts were added to determine lower bounds on the number of times that a segment of the pipeline must operate within the time horizon. Both improvements increase the performance of the model in terms of calculation time needed significantly.

Rejowski & Pinto (2008) further develop the model of Rejowski & Pinto (2004) by changing the time representation from discrete to continuous. This is done by giving the time intervals a variable duration, using the same time representation as Moro & Pinto (2004). The model also considers pumping flow and yield rate variations and minimizes inventory costs, interface costs and pumping costs. The variable duration of time intervals results in non-linear terms in the objective function, which makes the proposed model a MINLP. The pumping costs are also non-linear due to the division of the unit pumping cost and the amount of product by the pumping yield rate (the flow rate that results from a certain pumping rate), which is considered since the power consumption of the booster stations changes significantly with the flow rate in the pipeline. The authors show that the proposed MINLP solves faster than the MILP of Rejowski & Pinto (2004), which considers a fixed flow and yield rate.

Herrán et al. (2010) further improve the discrete time model, using the same assumptions and objective function as proposed by Rejowski & Pinto (2003). The big difference is that their model involves multiple refineries, intermediate nodes, and destination nodes, which results in a larger network and increases complexity, meaning that solving a real life case to optimality takes more than 20,000 seconds.

Relvas et al. (2013) consider a simple network of a single refinery tank farm and one customer at the end of a multi-product pipeline. The authors also propose a MILP with a discrete time representation, however the pipelines are not discretized into packs. The receiving time of a batch can therefore assume intermediate values between points of the discrete time scale, which results in a model that solves problems with a medium-term time horizon in a relatively short time. The model minimizes the average flow rate, which indirectly minimizes pumping costs.

The first model that uses a continuous time representation for the Multi-product Pipeline Scheduling Problem is proposed by Cafaro & Cerdá (2004). The problem is the same as considered by Rejowski & Pinto (2003), but with the batch volumes and start and completion times of the different product batches modelled as continuous variables, which reduces the number of binary variables, constraints and computation time.

Relvas et al. (2006) also propose a continuous time model in which they consider a single multi-product pipeline network as done by Relvas et al. (2013). The proposed MILP model maximizes the amount of products transported plus the total inventory at the end of the time horizon. An extension of the model which takes a settling period (the period in which the oil has to stay in a tank) into account is also considered, which results in a more complex model with a lower performance.

Relvas et al. (2007) improve the model of Relvas et al. (2006) by adding a variable flow rate, pipeline stoppages, and variable settling periods. Next to that, a new rescheduling methodology is proposed, which tells how to re-optimize a given schedule when a disturbance occurs. The considered disturbances are: variation on client demands, imposition on product sequence, unpredicted pipeline stoppages, batch volume modifications, flow rate adjustments, and variation on maximum capacity storage. The model indirectly minimizes pumping costs by minimizing the pumping flow rate.

Relvas et al. (2009) develop a pre-processing heuristic to improve the model of Relvas et al. (2007). The heuristic tries to find the most desirable product sequences to be pumped into the pipelines, which results in a lower solving time for schedules with short-term and medium-term horizons.

Cafaro & Cerdá (2008a) consider the dynamic scheduling of a single multi-product pipeline over a multi-period moving horizon to extend the model of Cafaro & Cerdá (2004). At the end of the current period, the planning horizon moves forward and the re-scheduling process based on updated problem data is triggered again over the new horizon. The authors propose an efficient MILP formulation based on a continuous time representation, which minimizes the sum of pumping, transition, down-time, back-order, and inventory carrying costs.

Cafaro & Cerdá (2008b) consider the same real-world case study as Relvas et al. (2006) and adjust the proposed MILP formulation of Cafaro & Cerdá (2004) to this model. The authors consider a product-dependent settling time in the tanks, resulting in a much simpler model than Relvas et al. (2006). Batch-size tracing is considered as unimportant in the model, which makes the computation time drop by a factor of nearly 100.

Cafaro & Cerdá (2012) consider a multi-product pipeline network including mesh structures, which means that alternative paths between nodes can exist. The model allows simultaneous batch injections at multiple input stations. The proposed MILP formulation based on a continuous time representation is able to generate short-term operational schedules in reasonable computation time. The objective of the model is to timely meet all product demands at distribution terminals at minimum total cost including pumping, interface, pipeline utilization, and inventory costs.

Boschetto et al. (2010) consider a large network of refineries, harbors, customers, and depots that are connected by pipelines. This is the first article that considers bidirectional pipelines, which all have their own capacity. The problem is decomposed hierarchically into an assignment block (where resources are allocated), a heuristics block (where the sequence of products is determined), a pre-analysis block (where volumetric and flow rate limits are determined using simulation), and a timing block (where the timing of activities is determined using a MILP model). These steps result in a schedule that minimizes the makespan. In addition the costs for violating the time windows of pumping and receiving the product batches are minimized.

Stebel et al. (2012) extend the work of Boschetto et al. (2010) and add a MILP model for the tactical level. The model enables to test scenarios at the strategic and tactical levels to see the impact on the operational level.

MirHassani et al. (2011) present an algorithm for the long-term scheduling of a single multi-product pipeline. The algorithm uses a MILP with a continuous time representation for calculating a short-term schedule iteratively to come to a long term schedule that aims to minimize penalty costs for under utilizing pipeline capacity, interface costs, and costs resulting from delays in supplying the demand.

MirHassani & Fani Jahromi (2011) discuss the short-term scheduling of the distribution of multiple products from a single refinery to multiple depots through a tree-structure pipeline. The authors propose a continuous time MILP formulation to generate a schedule. The objective is again to minimize the sum of inventory costs of the tanks, the interface costs between different batches, and the pumping costs, as in the initial problem of Rejowski & Pinto (2003).

Fabro et al. (2014) consider a real-world case of four refineries, which are connected via intermediate nodes to a single port using unidirectional pipelines. An important consideration of the model is that all pipelines are heated and therefore heating constraints exist. The other new aspect of the model is that it allows to use the same tank for different products during the time horizon. The authors extend the hierarchical decomposition approach as introduced by Boschetto et al. (2010), which results in a method that generates a schedule for the coming 30 days in less than 300 seconds.

3.1.3 Other optimization models

We have clarified that a lot of the research done on pipeline flow scheduling is centered around the two previously mentioned problems, however also other models can be found in literature. Ortiz-Gomez et al. (2002) provide three multi-period optimization models for oil production planning in the wells of an oil reservoir, which generate a schedule involving the production of oil from the wells in each time period. The first model assumes that a well either is shut down or open to flow during a fixed time period of which the resulting model consists of a discrete time MILP formulation. The second model assumes that each well operates at full production capacity. In this model the time periods are disaggregated in a number of sub periods. This model has some non-linear constraints because of the so called well bore pressure behavior, which makes the model a MINLP model. The third model is also a MINLP model, which assumes a cyclic operation mode for all of the wells of the reservoir in each period of time. All three models minimize the sum of the variable costs associated to the production rate and the costs associated to each well, which change depending on whether the well is open to flow or shut down.

Neiro & Pinto (2004) present a general modelling framework for the operational planning of petroleum supply chains. The framework contains models for processing units, storage tanks, and pipelines, which are all combined in a large MINLP model. The objective is to maximize revenue, which is obtained by the product sales minus costs related to raw material, operation, inventory, and transportation. It is assumed that the transportation costs are linear with the flow rate. The non-linear equations result from the operational costs, which depend on both the variable feed flow rate of a unit and variables concerning operating modes. The authors state that applying a decomposition method to deal with the non-linear constraints is part of future work.

Finally Abbasi & Garousi (2010) consider the optimal scheduling of pump operations in fluid (such as oil or water) distribution networks. Since pumping costs are severe in these kind of networks, even slight improvements in the operations of these systems could lead to considerable savings. The proposed MILP model determines a optimal pump operation schedule while considering multi-tariff electricity supply. The model applies linearization techniques on the non-linear aspects of the model, for example some network hydraulics constraints and the calculation of power consumed.

3.2 OilCO's optimization problem

Now that we have an overview of all recent pipeline scheduling literature, we decide which of these models are relevant for OilCO. To compare the models, we give a schematic overview of OilCO's optimization problem in the same way as we did for the Crude Oil Operations Scheduling Problem in Figure 1 and the Multi-product Pipeline Scheduling Problem in Figure 2. Figure 3 illustrates a simplified graphical overview of OilCO's optimization problem, according to the different processes as explained in Section 2.1.

It becomes clear that the optimization problem of OilCO differs from the Crude Oil Scheduling Problem and the Multi-product Pipeline Scheduling Problem. However, both problems contain

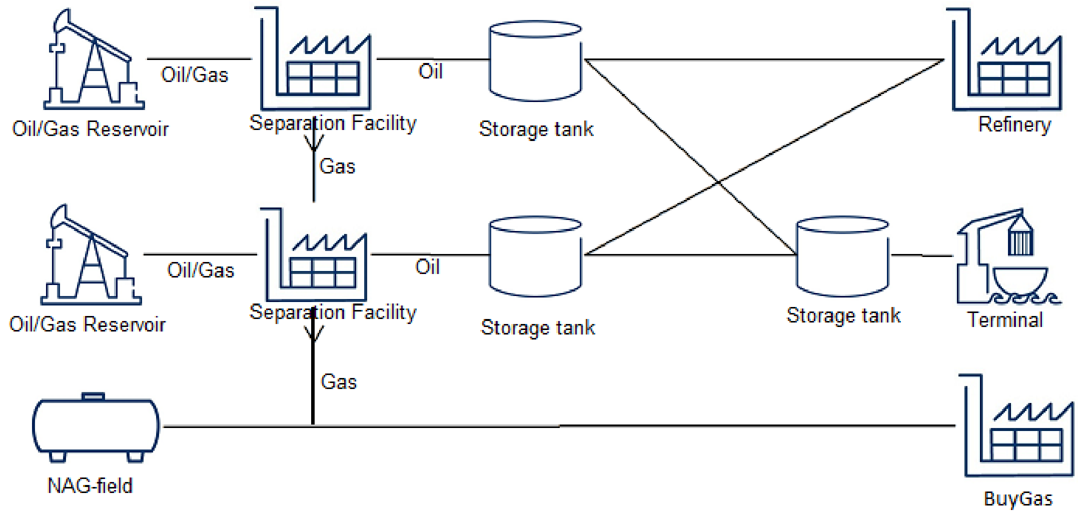

Figure 3: Graphical overview of OilCO's optimization problem

aspects that could be used for solving OilCO's optimization problem. Recall that the goal of this literature review is to get insight in how OilCO's pipeline network can be modelled, how the different assets can be modelled, and how the different cost components can be modelled. Section 3.2.1 explains how OilCO's pipeline network can be modelled and Section 3.2.2 treats how to model the assets in OilCO's optimization problem. We treat OilCO's optimization problem in a systematic manner by using the frequently used decomposition approach (Shah, 1996; Mas & Pinto, 2003; Jia & Ierapetritou, 2004; Neiro & Pinto, 2004), which divides the total optimization problem in several sub models based on the functionality of the different nodes in the model. Section 3.2.3 looks at how to model the different cost components in the objective function of OilCO's optimization problem and Section 3.2.4 looks at the functional requirements of the O3. All sections indicate the relevant literature for that part of OilCO's optimization problem.

3.2.1 Network

Pipeline networks can be modelled as a flow network of nodes that are connected by pipelines (Stebel et al., 2012; Cafaro & Cerdá, 2012), we can model OilCO's network in a similar way. In this network model all OilCO's assets are represented by nodes with specific characteristics. In OilCO's network the reservoirs and the NAG-field are the supply nodes which produce flow and the customers of OilCO are the consumers of the flow. Other assets of OilCO can be modelled as nodes that either conserve flow (pipeline nodes), store flow (storage tanks) or change the characteristics of the flow (separation facilities). How to model pipeline flows in the network strongly depends on the time representation that is chosen. Discrete time representations lead to intuitive model formulations, but lack in performance. A continuous time representation could be more suited than a discrete time formulation, as it is able to model times and volumes more precisely and it requires fewer binary variables in the model (Reddy et al., 2004a). More recently, a hybrid form called event-based time representation (Saharidis et al., 2009; Yadav & Shaik, 2012) also gives promising results.

3.2.2 Assets

Since we did not find any literature about fluid pipeline scheduling problems that also involve gas flows, we do not have any literature of this part of the model. For modelling the oil reservoirs, the model of Ortiz-Gomez et al. (2002) could be used as it considers multi-period optimization for oil production planning in the wells of an oil reservoir. Although this local optimization is done by the CS of OilCO, this information about well and reservoir behavior could be useful.

Separation facilities are not modelled explicitly in the literature, but the framework of Neiro & Pinto (2004) contains a sub model of a processing unit. A processing unit is defined here as "a piece of equipment that is able to physically or chemically modify the material fed into it". According to this definition, a separation facility could be modelled as a processing unit that changes the gas/oil mixture into gas and oil. This sub model considers different feed streams that are mixed before they are processed, which causes non-linear constraints when the properties of the total feed stream are calculated. It may be that this also causes non-linear constraints for OilCO, as the GORs from the reservoirs can vary.

Storage tanks are present in most of the found literature and most models also consider the inventory costs of the storage tanks in their objective function. OilCO's storage tanks can be modelled in the same manner as in existing literature. Some authors (Reddy et al., 2004b; Saharidis et al., 2009; Yadav & Shaik, 2012) give special attention to the mixing of different types of crudes in a storage tank. Non-linear constraints occur because the properties of the mixture are calculated using non-linear equations. Pan et al. (2009) have proposed a heuristic to deal with this by calculating the properties of the mixture in an indirect way. In OilCO's optimization problem only one type of crude is considered, so non-linear equations are avoided.

The different oil customers of OilCO can be modelled as the CDU in the Crude Oil Operations Problem and the customers in the Multi-product Pipeline Scheduling Problem, these are all customers with a certain demand for a certain product in a certain period. The modelling of the customers could even be simplified as all OilCO's oil customers demand the same product.

3.2.3 Objective function

The objective of OilCO's optimization problem is to minimize total costs which consist of: pumping costs, penalties for non-satisfied demand, penalties for deviating from the production targets, changeover costs at the reservoirs, safety stock penalties, and inventory costs. In certain articles (Rejowski & Pinto, 2003; Herrán et al., 2010; MirHassani & Fani Jahromi, 2011) the pumping costs are considered as the product of the pumped volume and the unit pumping cost and in some articles it is the product of the flow rate and a cost factor (Neiro & Pinto, 2004). Other authors (Rejowski & Pinto, 2008; Abbasi & Garousi, 2010) consider the pumping costs as non-linear, because the amount of power that pumps consume changes non-linearly when the flow rate in the pipeline changes. This consideration makes the model non-linear, but more accurate. Since OilCO's pumping costs are a significant part of the total costs, we want to model them as accurately as possible, so we use the non-linear approach.

The fines for non-satisfied demand are not modelled as a cost component in many articles as most authors model the fulfillment of demand as a constraint. MirHassani et al. (2011) include fines for delays in delivering the demand. Production target deviation costs are not considered in any of the articles, but Relvas et al. (2007) penalize deviations from the original schedule in their re-optimization methodology, which is done by multiplying the absolute deviation from the original schedule with a cost factor. Changeover costs at the reservoirs are also not considered in any of the articles, but some (e.g. Lee et al. (1996); Reddy et al. (2004b)) Crude Oil Operation Scheduling models do consider changeover costs of changing the tank that feeds the CDU. These models can be useful to see how changeover costs are modelled. Inventory costs are considered in almost every article that considers storage tanks, the inventory costs always are a simple multiplication of the average inventory level and a cost factor (Lee et al., 1996; Cafaro & Cerdá, 2012). Reddy et al. (2004b) model safety stock penalties in their objective function, this formulation is useful for OilCO.

3.2.4 Functional requirements

The most important functional requirement is that a schedule is generated in reasonable time, which is proved to be hard in large networks (Herrán et al., 2010) and/or on large time horizons (MirHassani et al., 2011). Some techniques were proposed in literature that could increase the performance of an optimization model by decomposing the problem hierarchically into several levels that add more detail to the schedule (Relvas et al., 2009; Boschetto et al., 2010; Fabro et al., 2014).

Another functional requirement is that the O3 has to be able to handle disturbances. Rescheduling techniques have been proposed by multiple authors (Relvas et al., 2007; Adhitya et al., 2007) who claim that solving a model all over again is inefficient in case of a disturbance and instead propose heuristics to reschedule only the relevant parts of a schedule. These heuristics can be used for the handling of disturbances by the O3.

3.3 Robust optimization

The O3 needs to deal with different kinds of disturbances (see Section 2.2). Disturbances such as assets breakdowns are events which are best modelled after realization by means of reactive scheduling (Li et al., 2012). The what-if analysis proposed by OilCO can be seen as a form of reactive scheduling as the "base case" is generated by solving the deterministic model. For disturbances in the form of delayed tankers a more robust approach would be better, since possible delay should be taken into account already when solving the model rather than after solving. Note, that the optimal solution of a deterministic model based on estimated arrival times does not have to be the solution that is able to withstand tanker delay. There may be solutions that are only close to optimality and are more robust with respect to tanker delay. Uncertainty related to demand and tanker arrival time can explicitly be taken into account through preventive approaches (Li et al., 2012). Robust Optimization (RO) is a preventive approach in which models are formulated that, by design, yield solutions that are less sensitive to uncertainty in model parameters (Mulvey et al., 1995). As the O3 needs to be able to deal with disturbances, solutions are needed that are robust to these disturbances and thus RO a is useful method to be applied to the model for the O3. Section 3.3.1 gives a brief introduction into RO where the concept of RO is explained. Section 3.3.2 concerns RO applications in the oil and gas industry.

3.3.1 Introduction in Robust Optimization

RO had its first applications in the early 1990s (Mulvey et al., 1995). In RO there are two forms of robustness: solution robustness and model robustness. Solution robustness means that the solution of the optimization model remains "close" to optimal for every realization of the uncertain parameters.

Model robustness means that the solution of the optimization model remains "almost" feasible for every realization of the uncertain parameters. The definitions of "close" and "almost" are subjective, but quantified by norms in the model.

A RO model is a model that has two sets of decision variables: design variables and control variables. Design variables are decision variables whose optimal values do not depend on the realization of the uncertain parameters. Design variables cannot be adjusted, once a realization of the uncertain parameters is observed. Control variables are decision variables that can be adjusted once a realization of the uncertain parameters is observed. Let the set x (we denote the set with a small letter, because the set denotes a vector of variables, while capital letters denote matrices of parameters) denote the set of design variables and the set y denote the set of control variables. The deterministic linear optimization model then has the following form (Mulvey et al., 1995):

$$\begin{aligned} \min\ & c^T x + d^T y \\ s.t.\ & Ax = b \\ & Bx + Cy = e \\ & x, y \geq 0 \end{aligned} \tag{1}$$

In (1) the second line denotes the constraints in which only design variables are included and the third line denotes the constraints in which the control variables are included together with the other design variables. In this model there is uncertainty in the coefficients d, B, C and e. Mulvey et al. (1995) introduce the set Ω, which denotes the set of scenarios of realizations. Every scenario s (with $s \in \Omega$) has an associated set $\{d_s, B_s, C_s, e_s\}$ of realizations of the uncertain parameters and a vector y_s of associated control variables. The authors also define a vector of penalty variables z_s and the probability p_s that scenario s occurs. The vector of penalty variables measures infeasibilities in the scenario, such that these can be penalized in the objective function. The resulting RO model is then:

$$\begin{aligned} \min\ & \sigma(x, y_1, ..., y_s) + \omega\rho(z_1, ..., z_s) \\ s.t.\ & Ax = b \\ & B_s x + C_s y_s + z_s = e_s \ \ \forall s \in \Omega \\ & x \geq 0, y_s \geq 0 \ \ \forall s \in \Omega \end{aligned} \tag{2}$$

The objective function minimizes the sum of the functions $\sigma(.)$, which is a function of the original variables, and $\rho(.)$, which is a function of the penalty variables. The objective function $\xi = c^T x + d^T y$ of (1) is now a random variable which has a probability p_s of taking value $\xi_s = c^T x + d_s^T y_s$. This implies that there is no single objective function, but there are multiple objective functions that can be used, denoted by the function $\sigma(.)$. If $\sigma(.) = \sum_{s \in \Omega} p_s \xi_s$ holds then the expected value of random variable ξ_s is minimized. If $\sigma(.) = \max_{s \in \Omega} p_s \xi_s$ holds then the maximum value of $p_s \xi_s$ is minimized, also known as a worst-case analysis. The advantage of RO over stochastic linear programming is that in stochastic linear programming only the expected value of ξ_s is minimized, where in RO every function of the random variable ξ_s can be minimized, depending on the choice of $\sigma(.)$. In high risk situations Mulvey et al. (1995) propose a function that minimizes the expected value of ξ_s plus a constant (λ) multiplied with the variance, so both the risk that a decision maker is willing to take and the distribution of ξ_s can be taken into account. The $\sigma(x, y_1, ..., y_s)$ part of the objective is then denoted as:

$$\sigma(x, y_1, ..., y_s) = \sum_{s \in \Omega} p_s \xi_s + \lambda \sum_{s \in \Omega} p_s (\xi_s - \sum_{s' \in \Omega} p_{s'} \xi_{s'})^2 \tag{3}$$

The parameter λ in (3) is used to make a trade-off between minimizing the expected value and the variance. The function $\rho(.)$ in (2) is called the feasibility penalty function, which penalizes violations by the control variables in the second line of constraints in (2). The specific choice of the penalty function is problem dependent. A commonly used penalty function when only positive violations of the constraints are relevant is: $\rho(z_1, ..., z_s) = \sum_{s \in \Omega} p_s max\{0, z_s\}$. The parameter ω in (2) is a weight factor that is used to make a trade-off between solution robustness (function $\sigma(.)$) and model robustness (function $\rho(.)$).

3.3.2 Robust Optimization applications in the oil and gas industry

In literature some applications of RO in upstream and/or midstream oil operations exist. Wang & Rong (2010) propose a RO model for the Crude Oil Operations Scheduling Problem (see Section 3.1.1) to deal with uncertainty in tanker arrival times and fluctuating oil demand. Uncertain demands are represented by chance-constrained programming and tanker arrival delay is represented by a scenario approach as in (2). They propose an $\sigma(.)$ function similar to (3), except that the variance term is replaced by $\sum_{s \in \Omega} p_s |\xi_s - \sum_{s' \in \Omega} p_{s'} \xi_{s'}|$ in order to have a linear model. Infeasibilities in violating the minimum and maximum storage capacity of a storage tank are penalized in the objective function. Li et al. (2005) also propose a RO model for dealing with uncertainty in demand and tanker arrivals in the Crude Oil Operations Scheduling Problem in which they do not penalize infeasibilities in the objective function, but penalize schedule changes. The $\sigma(.)$ function maximizes the expected profit over all scenarios. Li et al. (2012) propose a RO framework for dealing with demand uncertainty in the Crude Oil Operations Scheduling Problem, which contains solution methods for different kinds of uncertainty in demand. The resulting MINLP maximizes profit, which is the gross profit per scenario minus the safety stock penalties per scenario. In Section 5.2 we apply RO in order to come to a model for the O3 that yields solutions that are robust to disturbances.

3.4 Conclusion

This chapter presented a literature review to see which optimization models for pipeline flow scheduling are useful for OilCO. We found two major topics on this subject: the Crude Oil Operations Scheduling Problem and the Multi-product Pipeline Scheduling Problem. We gave an extensive overview about existing models for these problems and the differences in approaches of solving them. In modelling pipeline flow scheduling models, the manner in which time is represented is essential as it determines the model formulation and performance. For both the Crude Oil Operations Scheduling Problem (Jia & Ierapetritou, 2004) and the Multi-product Pipeline Scheduling Problem (Cafaro & Cerdá, 2004), continuous time representations outperform discrete time presentations. An event based time representation also outperforms the discrete time representation (Saharidis et al., 2009). No research is done so far in comparing continuous time representations with event based time representations. We therefore prefer a continuous time or an event based time representation for OilCO's optimization model.

Most of the proposed models are Mixed Integer Programs, where some are linear and some are non-linear. Non-linearity can occur when the mixing of different types of crudes are considered and when the efficiency of pumps is taken into account. We will model OilCO's network as a network of nodes which are connected by pipelines and where every node corresponds to one of OilCO's assets. We found in literature examples of how the different assets of OilCO can be modelled. We also found examples of how to model the different cost components we need to minimize. Aspects of other models can be used for a mathematical model for the O3, even though the problems considered in literature differ a lot from OilCO's optimization problem. In Chapter 4 we use these aspects in our model formulation.

This chapter also gave an overview of RO techniques and applications in upstream and/or midstream oil operations. Here we saw how a optimization model can be extended such that it yields solutions that are robust to disturbances. Furthermore, we found applications of RO in upstream and/or midstream oil operations (Wang & Rong, 2010; Li et al., 2005, 2012) that can be useful for the O3. Section 5.2 uses this knowledge to extend our optimization model such that it is able to deal with disturbances.

4 Model formulation

In this chapter we formulate a model for the optimization module of the O3 based on the good insight in OilCO's processes and the overview of relevant models from literature. The model will be basis for the optimization module of the O3. First, in Section 4.1 we choose the time representation we use in the model, since this decision influences the basics of the model. In Section 4.2, we make choices on where to allow simplifications by making assumptions, while Section 4.3 describes the basics of how we model OilCO's network. In Section 4.4, we formulate constraints for our model, Section 4.5 clarifies the objective function of our model, Section 4.6 explains the uncertainties in the model and which alternatives are possible, and Section 4.7 clarifies a possible simplification of the model. Finally, Section 4.8 summarizes this chapter.

4.1 Time representation

In Chapter 3 we found that the manner in which time is represented in an optimization model for pipeline flow scheduling determines the basics of the model and has impact on its performance. In this Chapter we discussed discrete time representations, continuous time representations, and event based time representations. Recall that a discrete time formulation divides the scheduling horizon in discrete time periods of equal length, a continuous time formulation models the start and end times of the events that have to be scheduled as continuous variables, and an event based representation divides the scheduling horizon into time periods of unequal length based on the start and end times of external events. These events are predefined external events that change important model parameters, e.g. tanker arrivals. The model variables are assumed to be constant during periods in which no such event occurs.

In case of OilCO, it is impossible to define events that have to be scheduled, which is necessary for a continuous time representation, since its reservoirs, separation facilities, and pipelines operate constantly. Furthermore, the linear models with continuous time representations assume the throughput rate to be constant, which is not the case for OilCO's problem, since the throughput rate is one of the decision variables of the model. The throughput rate can be variable by making the model non-linear, since both the throughput rate during an event and the duration of the event would be variable. However, since it takes more time to solve non-linear models, a continuous time representation is not an appropriate choice for our model, which leaves us with an event based time representation or a discrete time representation.

Both of these time representations divide the scheduling horizon in discrete time periods and model variables are evaluated at the end of each time period. An event based time representation can have the same accuracy as a discrete time representation, while needing less time periods (Saharidis et al., 2009). The accuracy of a discrete time representation depends on the chosen number of time periods, where the accuracy of an event based time representation is always the same as it evaluates model variables only at times (the start and end times of events) on which important model parameters change. For OilCO's situation these times would mostly be the start and end of a day, since the production (daily targets), refinery demand, and BuyGas demand are on a daily basis. The terminal demand, however, is not on a daily basis, because this demand depends on the start and end times of the loading of the tankers. Since we have no information about how many tankers OilCO serves, we make an assumption about this. If we assume that the maximum number of tankers that is loaded per day is five, which is reasonable given the number of loading platforms (also five), an event based time representation would result in at most twelve (start and end of day and the start and end of loading of five tankers) time periods per day. Furthermore, we may round every start and end moment to the nearest hour, since the level of detail required in our schedule is not minutes, but hours.

An event based time representation with twelve time periods per day corresponds with a discrete time representation with time periods of two ($\frac{24\,hours}{12\,periods}$) hours. Thus when twelve time periods per day are used, a discrete time representation has a base time period of two hours, while an event based time representation has a base time period of one hour, which is more accurate. We therefore choose to implement a event based time representation.

Figure 4 shows an example of the event based time representation in a tanker schedule of three days at the two terminals, in which the loading times of the tankers differ due to the differences in tanker volume. In this figure, the vertical blue lines indicate the start and end times of a time period. The set T[2] denotes the set of time periods we use in this model and the set D denotes the set of days. The parameter dur_t denotes the duration (in hours) of time period t and set T_d (with $T_d \subset T$) is the set of time periods that are part of day d. Note, that a time period can be only part of one day, since the start and end of a day are also moments that define new time periods. In the example of Figure 4 $dur_1 = 4$, $dur_2 = 6$, $dur_3 = 14$, $dur_4 = 2$, $dur_5 = 2$, $T_1 = \{1, 2, 3\}$, and $\{4, 5\} \in T_2$.

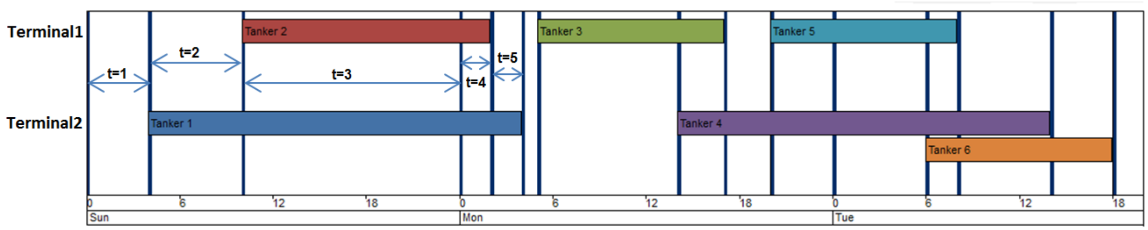

Figure 4: Example of an event based time representation

In our model, volumes are expressed in barrels (bbl), as this is the volumetric unit that OilCO uses for oil. Gas volumes are expressed in cubic feet (cf) to avoid the conversion of units in our model. All rates (i.e. throughput rate) are expressed in bbl/h. In our model the most important decision variables are the production rates of the reservoirs and the NAG-field, the separation rates at the separation facilities, and the throughput rates of the pipelines. The production rate variable $PR_{r,t}$ denotes the production rate of reservoir r in time period t in bbl/h, variable GAS_d denotes the production rate of the NAG-field on day d in cf/h, the separation rate variable $SR_{sf,t}$ denotes the separation rate of separation facility sf in time period t in bbl/h, and the throughput rate variable $TH_{p,t}$ specifies the throughput rate of pipeline p in time period t in bbl/h. In our model we use rates instead of volumes, since we use time periods of different duration and changes in production, separation or throughput are easier to notice if rates are compared. For some constraints (e.g. storage tank levels), however, we require volumes. These volumes simply can be calculated by multiplying the rate in a certain time period with the duration of that period.

[2]In our model capital letters denote sets and variables and small letters denote indices and parameters. The small letter that denotes the index for elements of a set is the same letter as the capital letter that denotes the set. If we require a second index for the same set, an apostrophe is added to the letter.

4.2 Assumptions

The following assumptions have to be made for our optimization model for the O3:

- A1 An oil pipeline remains completely full with incompressible oil products at any time, so the only way to get a volume of oil out of a certain pipeline is to inject an equal volume of oil at the origin. For convenience we also assume this for the pipelines in which the oil/gas mixture is transported.
- A2 Only one type of crude oil is considered, which means that different qualities of crudes are not taken into account.
- A3 All storage tanks are treated as aggregated capacities. In reality there are multiple smaller storage tanks at the tank locations, but we model this as a single storage tank with an overall total capacity.
- A4 Changeover times for changing the production rate at a reservoir are neglected, while we do take changeover costs into account.
- A5 Only one type of pump is used to realize throughput at each pipeline. OilCO transports oil though pipelines that require more than one pump due to their length. We assume that all these pumps are of the same type.
- A6 Transportation times of gas to BuyGas are neglected.

Assumption 1 is made for most oil pipeline scheduling problems, which means that an increase in throughput at the origin of a pipeline directly results in an increase in throughput at the destination of the pipeline. In reality there is some delay, but this assumption can be made, since we look at time periods of multiple hours. OilCO suggests Assumption 2, because they have other systems for monitoring crude qualities and this is therefore out of the scope of the O3. Assumption 3 simplifies the model significantly, while the same storage volume is available. We make Assumption 4, since taking changeover times into account would make the model more difficult, while changeover times are relatively short compared to the planning horizon and occur only rarely. Moreover, we do not know what the production rate during a changeover is, where it is possible that the production rate increases/decreases gradually or instantly or that the production rate is 0 during changeovers. Assumption 5 makes it easier to calculate the energy costs for a pipeline, since every type of pump at a pipeline requires a separate calculation of the energy costs. This is avoided by assuming all pumps are of the same type. We make Assumption 6, because BuyGas operates different facilities which are all close to separation facilities. We therefore assume that the gas is delivered to BuyGas at the moment it is separated from the oil or extracted from the oil extracted from the NAG-field. These assumptions make that we can translate OilCO's processes into an easier model.

4.3 Network modelling

Figure 5 shows a functional diagram of the total network we want to model, where all gas is directly transported from the separation facilities to BuyGas. It is uncertain if the network exactly looks like Figure 5. Therefore, we model the network as generic as possible (see also 2.3.2). Oil pipeline networks can be modelled as a set of nodes which are connected by pipelines (Stebel et al., 2012; Cafaro & Cerdá, 2012). The basics of OilCO's network can be modelled in a similar manner, since we assume that all oil is of the same quality and the flow into a pipeline is equal to the flow out of the pipeline (so transportation times can be neglected). In every node the flow has to be conserved (flow in = flow out), unless the node is a source (which produces flow) or or a sink (which consumes flow).

In our network, nodes are represented by OilCO's assets (set A) and arcs are represented by pipelines (set P). The set A is a collection of all assets of OilCO, which are all units mentioned in Section 2 excluding pipelines. Not all assets are actually owned by OilCO, since the refineries and BuyGas are customers. The set of assets consists of the set of reservoirs R, the set of separation facilities SF, the set of storage tanks ST, the set of refineries RF, the set of terminals TE, and the set of pipeline nodes N. In this network every pipeline connects two assets to each other, which is modelled using a set C which denotes all pipeline connections. Pipeline connections are represented by a combination of an asset (a), a pipeline (p), and another asset (a'), such that pipeline p transports flow from asset a to asset a'. We require pipeline p in this set to distinguish parallel pipelines between asset a and asset a'. The sources of this network that produce flow, are the reservoirs and the NAG-field. In the network the only place where the produced flow is consumed is at the customers, these are the sinks. In a network all intermediate nodes have to conserve flow. In OilCO's network this principle holds for the pipeline nodes N (Node 1 and Node 2 in Figure 5), because these are junctions of pipelines where flows split up and/or merge, which means that the incoming flow is equal to the outgoing flow.

The other assets are modelled differently as flow is either stored or converted at those assets. At the separation facilities SF an extra flow is generated, as a gas flow is extracted from the oil flow. The oil flow is conserved in a separation facility, which means that the incoming oil flow rate must also be equal to the the outgoing oil flow rate. The separation rate is also equal to the incoming oil flow rate into the separation facility, as that is the rate at which the oil is processed. For flow conservation we do not consider the gas flow rate of the incoming mixture, as gas is compressed in the pipeline. The incoming oil flow rate also has to be equal to the the outgoing oil flow rate. The outgoing gas flow rate is calculated by multiplying the oil flow rate with the GOR. At the storage tank the incoming flow can also be greater (stock level increases) or lower (stock level decreases) than the outgoing flow. Section 4.4 explains how this network model is translated into constraints for the optimization model.

4.4 Constraints

We structure our explanation of the optimization model according to the different assets of OilCO and formulate constraints for each asset. We model all assets of a certain type in the same way, since we want a model that is as generic as possible to be able to model network layouts that are different from OilCO's, as long as the same types of assets are used. Section 4.4.1 covers the reservoirs, Section 4.4.2 covers the processing facilities, Section 4.4.3 covers the pipelines, Section 4.4.4 covers the storage tanks, and Section 4.4.5 covers the demand locations.

4.4.1 Reservoir constraints

Recall that the set R denotes the reservoirs, which are the sources of the network flow model. Since the reservoirs have no storage tanks, all oil that is produced has to be transported. Equation (4) assures this by setting the sum of the throughput rates ($TH_{p,t}$) of the outgoing pipelines equal to the production rate ($PR_{r,t}$). The summation domain $(p, a)|(r, p, a) \in C$ ensures that the production of reservoir r is pumped into the pipelines p that connect asset a to reservoir r.

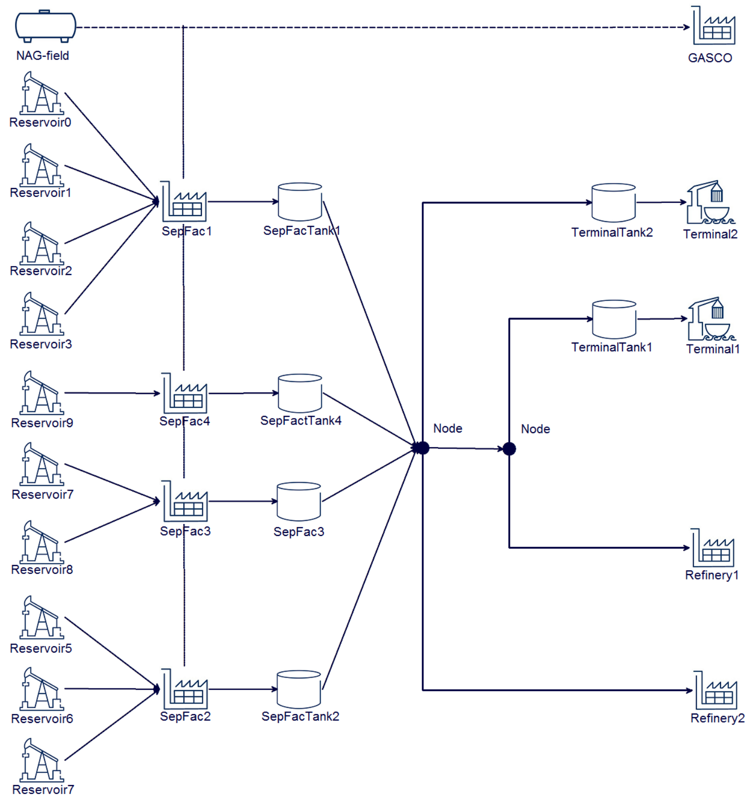

Figure 5: Functional diagram of OilCO's network

$$PR_{r,t} = \sum_{(p,a)|(r,p,a)\in C} TH_{p,t} \qquad \forall r \in R;\ t \in T \tag{4}$$

$$rmin_r\, avail_{r,t} \leq PR_{r,t} \leq rmax_r\, avail_{r,t} \qquad \forall r \in R;\ t \in T \tag{5}$$

The minimum and maximum production rate are taken into account in (5). Here, $rmin_r$ is the minimum production rate and $rmax_r$ is the maximum production rate of reservoir r, which are both input parameters. Parameter $avail_{r,t}$ is the availability (in %) of reservoir r during time period t. The maximum possible production rate in time period t is the maximum production rate of reservoir r multiplied with the availability of the reservoir in that time period. To ensure that the model remains feasible when the availability is low, the minimum production of reservoir r ($rmin_r$) is also multiplied with the availability.

4.4.2 Separation facility constraints

Recall that the set SF denotes the separation facilities. In the separation facilities gas is extracted from the oil. The separation rate ($SR_{sf,t}$) is the rate at which gas is separated from the oil, and has to be equal to the sum of the incoming oil flow rates, as that is the rate of the oil flow into the facility. Equation (6) sets the $SR_{sf,t}$ variable equal to the incoming oil flow rate. The ratio between the oil and the gas flow depends on the parameter gor_r, which is the GOR of reservoir r in cf/bbl. The gas flow rate ($GFR_{sf,t}$) of separation facility sf in time period t is determined by multiplying the oil flow rate with the GOR, which we model in (7). The inequality (8) bounds the separation rate in the same way as is done for the reservoirs in (5). Here the parameter $av_{sf,t}$ denotes the availability (in %) of separation facility sf in time period t and $srmax_{sf}$ denotes the maximum separation rate of separation facility sf.

$$\sum_{(r,p)|(r,p,sf)\in C} TH_{p,t} = SR_{sf,t} \qquad \forall sf \in SF;\ t \in T \tag{6}$$

$$\sum_{(r,p)|(r,p,sf)\in C} TH_{p,t}\, gor_r = GFR_{sf,t} \qquad \forall sf \in SF;\ t \in T \tag{7}$$

$$SR_{sf,t} \leq srmax_{sf}\, av_{sf,t} \qquad \forall sf \in SF;\ t \in T \tag{8}$$

$$(1 - oilloss)\, SR_{sf,t} = \sum_{(p,a)|(sf,p,a)\in C} TH_{p,t} \qquad \forall sf \in SF;\ t \in T \tag{9}$$

For the conservation of flow the separated oil has to leave the separation facility via pipelines, which we model in (9). We also account for the oil loss ($oilloss$) that occurs in the separation facilities in this constraint, which is the fraction of oil that is lost in the process. The transportation of gas out of the separation facility is not explicitly modelled, because the BuyGas facilities are close to the separation facilities.

4.4.3 Pipeline constraints

Recall that the set P denotes all pipelines and the set N denotes all pipeline nodes in the network. Our model considers a maximum and a minimum flow per pipeline, which are the maximum ($pmax_p$) and minimum ($pmin_p$) throughput of pipeline p, respectively. Both the minimum and maximum throughput are considered in (10). Here, the availability of pipeline p during time period t in % ($ava_{p,t}$) is also considered using the same formulation as for the reservoirs and separation facilities.

$$pmin_p\, ava_{p,t} \leq TH_{p,t} \leq pmax_p\, ava_{p,t} \qquad \forall p \in P;\ t \in T \tag{10}$$

$$\sum_{(a,p)|(a,p,n)\in C} TH_{p,t} = \sum_{(p,a)|(n,p,a)\in C} TH_{p,t} \qquad \forall n \in N;\ t \in T \tag{11}$$

In a pipeline node the flow is not converted or stored, but the flows are only merged and/or split up. Equation (11) is a flow conservation constraint, meaning that the incoming flow rate of oil has to be equal to the outgoing flow rate of oil.

4.4.4 Storage tank constraints

Recall that the set ST denotes all storage tanks. For storage tanks st the end inventory of time period t is modelled as variable $I_{st,t}$. In Equation (12) the end inventory of a certain time period is calculated by adding all incoming oil volumes to the end inventory of the previous period ($I_{st,t-1}$)

and subtracting all outgoing oil volumes from it. Here $I_{st,0}$ is an input parameter that denotes the inventory of storage tank st at the beginning of the planning horizon. The parameter dur_t is included in the equation to calculate the volume pumped into the tank by multiplying the throughput during time period t by the duration (dur_t) of the same period in hours, so the units of the variables $TH_{p,t}$ (bbl/h) and $I_{s,t}$ (bbl) are aligned.

$$I_{st,t} = I_{st,t-1} + \sum_{(a,p)|(a,p,st)\in C} TH_{p,t}\, dur_t - \sum_{(p,a)|(st,p,a)\in C} TH_{p,t}\, dur_t \qquad \forall st \in ST;\, t \in T \qquad (12)$$

$$mininv_{st} \leq I_{st,t} \leq maxinv_{st} \qquad \forall st \in ST;\, t \in T \qquad (13)$$

Inequality (13) ensures that the inventory of a storage tank stays between its minimum ($mininv_{st}$) and maximum ($maxinv_{st}$) volume.

4.4.5 Demand constraints

OilCO delivers to three sorts of customers: terminals, refineries, and BuyGas. All these customers have a demand that has to be fulfilled and OilCO receives a fine per bbl of demand that is not met. Recall that the set RF denotes the refineries and set TE denotes the terminals. Parameter $tdem_{te,t}$ specifies the terminal demand, which is the demanded volume at terminal te in time period t. Parameter $rdem_{rf,t}$ specifies the refinery demand, which is the demanded volume of refinery rf in time period t. The gas demand of BuyGas is a daily demanded volume and is denoted by $gdem_d$. In (14) the volume of oil delivered to a refinery in a time period has to be equal to the demanded volume in that time period. Variable $SHR_{rf,t}$ is added to measure the total shortage volume at refinery rf at the end of time period t. It is impossible to deliver more than the demand, because the refinery can not handle the extra oil. The volume of oil delivered in a time period is again calculated by multiplying the throughput of that time period with the duration of that time period. The shortage is added to the demand of the next time period, since variable $SHT_{rt,t}$ measures the total shortage instead of the shortage in time period t. We need to penalize this total shortage, because all demand has to be fulfilled as fast as possible, since the customer is waiting for the oil.

$$\sum_{(a,p,tp)|(a,p,rf)\in C} TH_{p,t}\, dur_t = rdem_{rf,t} - SHR_{rf,t} + SHR_{rf,t-1} \qquad \forall rf \in RF;\, t \in T \qquad (14)$$

$$\sum_{(a,p)|(a,p,te)\in C} TH_{p,t}\, dur_t = tdem_{te,t} - SHT_{te,t} + SHT_{te,t-1} \qquad \forall te \in TE;\, t \in T \qquad (15)$$

$$\sum_{(sf,t)|t\in T_d} GFR_{sf,t}\, dur_t + GAS_d \geq gdem_d \qquad \forall d \in D \qquad (16)$$

Equation (15) follows the same reasoning for the terminals. Here, the $SHT_{te,t}$ variable denotes the total shortage volume at terminal te at the end time period t. It it is also impossible to deliver more than the demand at the terminals, since the oil is directly pumped into tankers without spare capacity. Both $SHR_{rf,0}$ and $SHT_{te,0}$ denote the total shortages at the refineries and terminals at the beginning of the planning horizon. The dur_t parameter is again included to align the units bbl and bbl/h of the different variables. The production volume at the NAG-field on day d in cf is denoted by the unrestricted variable GAS_d. The NAG production closes the gap between the associated gas production ($GFR_{sf,t}$) and the BuyGas demand, which is ensured by (16). Recall that $t \in T_d$ when time period t is part of day d, so the summation in (16) sums over all time periods t in day d. It is possible to deliver more gas than demanded, since the surplus gas is flared.

4.5 Objective function

Our model needs to minimize the sum of pumping costs, changeover costs, inventory costs, safety stock penalties, penalties for non-satisfied demand (also called shortage penalties), and production target deviation penalties. In Chapter 3 we discussed examples of how these cost components can be modelled, which we use to determine the objective function. We treat the different cost factors one by one.

At OilCO the pumping costs are determined by the energy costs for operating the pipeline. The variable $EN_{p,t}$ denotes the energy costs for pipeline p during time period t in US Dollars (USD) per hour, which is multiplied with the duration of time period t (dur_t) in order to get the total amount of USD. Section 4.5.1 explains how the energy costs are calculated using the models of Rejowski & Pinto (2008) and Abbasi & Garousi (2010).

Changeover costs occur when the production rate of a reservoir in time period differs from the production rate in the previous time period, because this means that individual wells have to be started up or shut down. We multiply the number of changeovers with a cost factor per changeover, as is also done by Lee et al. (1996) and Reddy et al. (2004b). For this we introduce a binary variable $CH_{r,t}$ which has to take value 1 if a changeover occurs between time period t and time period $t-1$ at reservoir r and 0 otherwise. Furthermore, the parameter cc_r denotes the cost in USD of a changeover. We assume that the size of the changeover is not important for the changeover costs, because we do not know for sure if a bigger changeover results in higher costs. Section 4.5.2 explains the calculation of binary variable $CH_{r,t}$.

In literature the inventory costs are calculated by multiplying the average inventory in a time period by a cost factor (Lee et al., 1996; Cafaro & Cerdá, 2012). We apply this method by multiplying the average inventory ($AI_{st,t}$) of storage tank st in time period t by a cost parameter $cinv_{st}$, which denotes the cost per bbl per hour in storage at storage tank st, and the duration of the time period dur_t. $AI_{st,t}$ is calculated by (17).

$$AI_{st,t} = \frac{I_{st,t} + I_{st,t-1}}{2} \qquad \forall st \in ST;\ t \in T \tag{17}$$

The safety stock penalties are the fourth term of the objective function. The penalties are calculated by multiplying the safety stock penalty for having one bbl less than the safety stock in storage (ssp_{st}) at storage tank st by the volume that the inventory level is below safety stock ($BSS_{st,t}$) at storage tank st in time period t. Section 4.5.4 clarifies how we calculate the $BSS_{st,t}$ variable using the formulation of Reddy et al. (2004b).

The fifth and sixth cost components cover the penalties of non-fulfilled demand at the terminals and refineries. Variable $SHR_{rf,t}$ denotes the total shortage volume at refinery rf at the end of time period t and variable $SHT_{te,t}$ denotes the total shortage at terminal te at the end of time period t. Both variables are multiplied with a cost factor. Cost factor $csht_{te}$ is for shortages at terminal te and $cshr_{rf}$ is for shortages at refinery rf, which denote a penalty per barrel of non-fulfilled demand.

The last cost component covers the production target deviation costs, which are penalties for OilCO when it does not produce accordingly to the production targets. The production targets are expressed in a volume per reservoir per day. Relvas et al. (2007) also penalize changes in the original plan in the objective function, by multiplying the absolute deviation from the plan with a penalty cost factor. We apply the same method to OilCO's problem. Here, the variable $DEV_{r,d}$ denotes the absolute deviation from this target at reservoir r on day d. The amount is multiplied with cost parameter $cdev_r$ which denotes the penalty cost per bbl of deviation at reservoir r. Section 4.5.3 explains how we calculate the $DEV_{r,d}$ variable. The overall objective function is as follows:

$$\min \sum_{p,t} dur_t\, EN_{p,t} + \sum_{r,t} cc_r\, CH_{r,t} + \sum_{st,t} cinv_{st}\, dur_t\, AI_{st,t} + \sum_{st,t} ssp_{st}\, BSS_{st,t} + \sum_{te,t} csht_{te}\, SHT_{te,t} + \sum_{rf,t} cshr_{rf}\, SHR_{rf,t} + \sum_{r,d} cdev_r\, DEV_{r,d} \quad (18)$$

4.5.1 Energy costs for operating pipelines

In Chapter 3 we have seen that the energy costs for operating pipelines can be calculated in different ways. Rejowski & Pinto (2003), Herrán et al. (2010), and MirHassani & Fani Jahromi (2011) consider the energy costs to be linear with the pumped volume, where Neiro & Pinto (2004) consider the energy costs to be linear with the throughput rate. Both approaches result in the same energy costs formulation in our model, since we use predefined time intervals, so the volume transported in time period t is equal to $TH_{p,t} * dur_t$. Rejowski & Pinto (2008) and Abbasi & Garousi (2010) consider the energy costs to be non-linear with the throughput rate. We also choose this non-linear approach, because it is the most realistic approach for OilCO's situation. We want to model the pumping costs as realistic as possible, since the pumping costs at OilCO are significant as OilCO's network consists of more than 1,500 kilometres of pipelines. There are, however, pipelines with a length of (almost) 0, since these connect assets that are on the same location. For these pipelines the energy costs are not relevant. The set $EP \subset P$ denotes the pipelines for which the energy costs are relevant. We combine the approaches of Rejowski & Pinto (2008) and Abbasi & Garousi (2010) to get a formulation of the energy costs costs that is most in line with OilCO's situation. We, however, have no information on the characteristics of the pumps that OilCO uses, so modelling the energy costs exactly is impossible. Therefore, we model the behavior of the energy costs such that the model finds the solution with (almost) the lowest costs, although we do not know the exact value of the costs. We create an example of a pump based on both cited articles to illustrate the relation between throughput rate and energy costs. The efficiency of a pump is an important factor in this realtion. The efficiency of a pump indicates how efficient a pump operates at a certain production rate, which is indicated by a function called the efficiency curve. In the analysis we define TH_p^* as the peak efficiency throughput, which is the throughput rate at which maximum efficiency is achieved, and set it equal to 100 kbbl/hour in our example. Using the efficiency curve (which is already a function of the throughput) we define an equation in which the energy costs are a function of the throughput, which we need to include energy costs in our model. Appendix C covers the detailed analysis.

Figure 6 show the energy costs $EN_{p,t}$ as a function of the throughput rate $TH_{p,t}$, which is the result of the analysis. The figure shows that the energy costs behave differently at both sides of the peak efficiency throughput (the vertical dashed line). Therefore, we treat both parts separately in order to find a formulation for the behavior of the energy costs.

Throughput > Peak efficiency throughput
Figure 6 shows that the energy costs function is convex when the throughput is higher than the

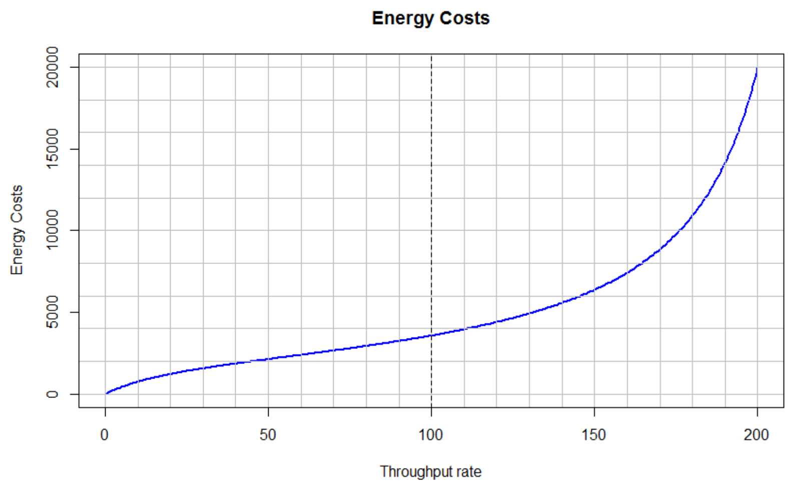

Figure 6: Energy costs of the example pump at different throughput rates

peak efficiency throughput, meaning that it is cheaper to spread increases above the peak efficiency point over multiple periods. For example, it is cheaper to have a throughput of 110 for five time periods than have a throughput of 150 for one period and a throughput 100 in the other four, while both cases result in the same volume of oil (550) that is transported.

To model the costs in case the throughput rate is larger than the peak efficiency throughput we introduce variable $PTD_{p,t}$, which is bounded by (19) and (20). Variable $PTD_{p,t}$ denotes the positive deviation from the peak efficiency throughput at pipeline p in time period t in bbl/hour.

$$PTD_{p,t} \geq TH_{p,t} - TH_p^* \qquad \forall p \in EP;\ t \in T \qquad (19)$$

$$PTD_{p,t} \geq 0 \qquad \forall p \in EP;\ t \in T \qquad (20)$$

As an example for expressing the costs of choosing a throughput rate above the peak efficiency throughput, we use the quadratic cost function $SPTD_{p,t} = (PTD_{p,t})^2$, which is a simple convex function. Adding this equation to our model, would change the model from linear to non-linear. However, a non-linear model requires more time and different methods to solve, so we want to keep our model linear. Using an approximation here is acceptable, because we only want to model the principle behavior of the energy costs instead of the exact cost, since the latter is impossible with the current information. A commonly used method to approximate a non-linear objective function is a piecewise linear approximation (Manthey, 2013), which we use, because it is easy to implement and does not require extra integer variables. For this piecewise linear approximation a set L of line segments is required next to the parameters $brp_{p,t,l}$ and $sbrp_{p,t,l}$ (with $sbrp_{p,t,l} = (brp_{p,t,l})^2$), which are the values of $PTD_{p,t}$ and $SPTD_{t,p}$ at the breakpoints (which are located at the start and end of the line segments). Using the piecewise linear function the value of $SPTD_{p,t}$ is approximated by a convex combination of two values at the breakpoints. (21) shows this approximation, in which the variable $\lambda_{p,tp,l}$ specifies the weight of the breakpoint at the end of line segment l.

$$\sum_{l} \lambda_{p,t,l}\, sbrp_{p,t,l} = SPTD_{p,t} \qquad \forall p \in EP;\, t \in T \tag{21}$$

$$\sum_{l} \lambda_{p,t,l}\, brp_{p,t,l} = PTD_{p,t} \qquad \forall p \in EP;\, t \in T \tag{22}$$

$$\sum_{l} \lambda_{p,t,l} = 1 \qquad \forall p \in EP;\, t \in T \tag{23}$$

The constraints (22) and (23) bound the variable $\lambda_{p,t,l}$. It can be shown that only two $\lambda_{p,t,l}$ variables are non-zero per pipeline and time period, since the function we approximate is convex. If this would not be the case, we would require extra integer variables to force only two $\lambda_{p,tp,l}$ to be non-zero (Manthey, 2013). The energy costs for $TH_{p,t} > TH_p^*$ are calculated by multiplying $SPTD_{p,t}$ with a cost factor (csq_p).

Throughput $\leq$ Peak efficiency throughput

The behavior of the energy costs is different when the throughput is lower than the peak efficiency throughput. Based on Figure 6 assume that the function is linear when $\frac{1}{5}TH_p^* \leq TH_{p,t} \leq TH_p^*$ (Figure 7 shows the linearization). The case in which the throughput is lower than $\frac{1}{5}TH_p^*$ is not relevant, since every pipeline has a minimum throughput that is higher than $\frac{1}{5}TH_p^*$. The linear part of the cost function is of the form $ca_p\,TH_{p,t} + cb_p$, where ca_p and cb_p are the cost factors. ca_p is the slope of the line and cb_p is equal to the value of the linearization energy costs when $TH_{p,t} = 0$. In this example cb_pwould lie around 700.

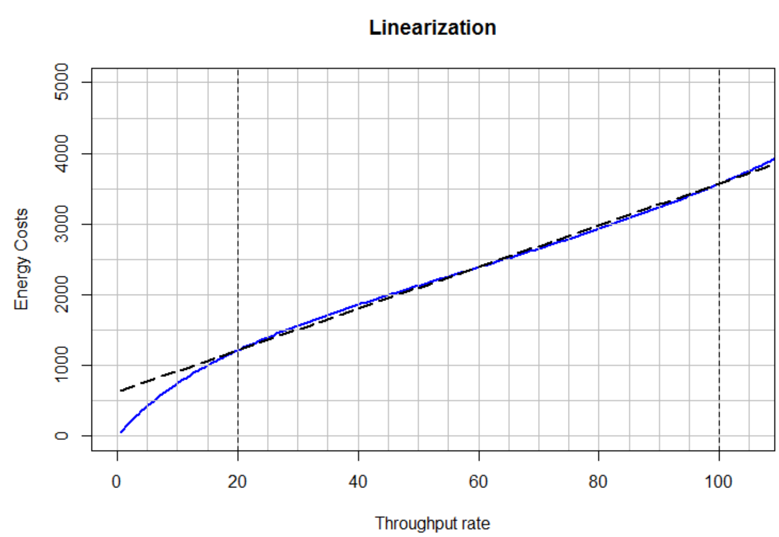

Figure 7: Linearization of the energy costs

The total energy costs thus consist of a linear element and a non-linear element. The advantage of this split is that less breakpoints are required to approximate the non-linear function and that the non-linear function is convex. (24) calculates the energy costs of pipeline p in time period t.

$$EN_{p,t} = ca_p\, TH_{p,t} + cb_p + csq_p\, SPTD_{p,t} \qquad \forall p \in EP;\ t \in T \tag{24}$$

Determining the values for the cost factors ca_p, cb_p, and csq_p is difficult. One of the factors that influences these cost factors is the length of pipeline p, as more pumps are required when the length of the pipeline increases. It is, however, not necessary to know the values of the cost factors exactly, as our main goal was to model the behavior of the energy costs.

4.5.2 Changeover costs at the reservoirs

Recall that the changeover costs are calculated by multiplying the number of changeovers by the fixed costs per changeover, as is also done by Lee et al. (1996) and Reddy et al. (2004b). A changeover occurs when the production rate at a reservoir is changed. The binary variable $CH_{r,t}$ has to be 1 when the production rate differs from the production rate in the previous time period, which can be both a positive and a negative difference. Inequalities (25)-(26) model this relation. Here, $PR_{r,0}$ is an input parameter that denotes the current production rate of reservoir r. The parameter M is an auxiliary parameter, which has the value of the maximum possible production deviation. This maximum possible production deviation is the production capacity minus the minimum production of reservoir r ($rmax_r - rmin_r$).

$$(PR_{r,t} - PR_{r,t-1}) \leq CH_{r,t}\, M \qquad \forall r \in R;\ t \in T \tag{25}$$

$$(PR_{r,t-1} - PR_{r,t}) \leq CH_{r,t}\, M \qquad \forall r \in R;\ t \in T \tag{26}$$

4.5.3 Production target deviation penalties

In the objective function we calculated the production target deviation penalties by multiplying the deviation from the production target ($DEV_{r,d}$) by a penalty factor (cc_r), as Relvas et al. (2007) propose. The deviation is the absolute deviation from the daily production target. Parameter $plan_{r,d}$ denotes the production target at reservoir r on day d. The summation $\sum_{t \in T_d} PR_{r,t}\, dur_t$ calculates the volume produced on day d by multiplying the production rate with the time period duration for all time periods in day d. Equations (27) and (28) bound variable $DEV_{r,d}$ at the absolute deviation of the daily production from the production target.

$$DEV_{r,d} \geq plan_{r,d} - \sum_{t \in T_d} PR_{r,t}\, dur_t \qquad \forall r \in R;\ d \in D \tag{27}$$

$$DEV_{r,d} \geq \sum_{t \in T_d} PR_{r,t}\, dur_t - plan_{r,d} \qquad \forall r \in R;\ d \in D \tag{28}$$

4.5.4 Safety stock penalties

To calculate the variable $BSS_{st,t}$, which denotes the volume that the inventory level of storage tank st is below safety stock at the end of time period t, we use the formulation of Reddy et al. (2004b). The safety stock of storage tank st is denoted by parameter $safinv_{st}$. Equations (29) and (30) bound variable $BSS_{st,t}$.

$$BSS_{st,t} \geq safinv_{st} - I_{st,t} \qquad \forall st \in ST;\ t \in T \tag{29}$$
$$BSS_{st,t} \geq 0 \qquad \forall st \in ST;\ t \in T \tag{30}$$

4.6 Cost modelling alternatives

We made assumptions about the cost components in the model from the previous sections, because we do not have contact with OilCO during the project. Moreover, we made assumptions about how we calculate the changeover costs and energy costs, as we already addressed in Section 2.3.1. In this section we want to point out these uncertainties in modelling the costs and discuss model changes that must be done when the cost components turn out differently than assumed. Section 4.6.1 covers the energy costs and Section 4.6.2 covers the changeover costs.

4.6.1 Uncertainty in modelling energy costs

Section 4.5.1 explains our energy costs approach. We created an example of a pump to examine the relation between the throughput rate of a pipeline and the associated energy costs. According to this relation we formulated the energy costs using both a linearization and an approximation of a convex function by a piecewise linear approximation. We modelled the non-linear behavior of the energy costs, such that the optimal solution is yielded, although we do not know the exact value of the costs. This results from the lack of information of OilCO's operational costs. If we get data from OilCO and our assumption about the behavior is right, the cost parameters ca_p, cb_p, and csq_p have to be estimated using the data.

If our assumption is not right, more work has to be done. Then first the relation between throughput rate and energy costs has to be derived from the data, of which the only thing we know for sure is is that it is non-linear. The definition of parameter $sbrp_{p,t}$ has to be changed according to this non-linear relation. Then the non-linear relation can again be approximated by constraints (21), (22), and (23) in the same manner as is done now. If the non-linear function, however, is not convex some variables and constraints have to be added to the model (Manthey, 2013). These variables and constraints have to make sure that at most two $\lambda_{p,t,l}$ variables are non-zero per pipeline and time period, so the approximated function can be expressed as a convex combination of two breakpoints. This is done by adding auxiliary binary variable $Y_{p,t,l}$ and constraints (31) and (32) (with $Y_{p,t,0} = 0$). In case of a convex function automatically exactly two $\lambda_{p,t,l}$ variables are non-zero.

$$\lambda_{p,t,l} \leq Y_{p,t,l-1} + Y_{p,t,l} \qquad \forall p \in P;\ t \in T;\ l \in L \tag{31}$$

$$\sum_{l} Y_{p,t,l} = 1 \qquad \forall p \in P;\ t \in T \tag{32}$$

This formulation, however, has a large impact on the model performance, since it adds a large number of integer variables (one for every combination of pipeline p, time period t, and line segment l) to the model. In Section 7 we perform a sensitivity analysis on the current behavior of the energy costs to examine how the model reacts on different values for the peak efficiency throughput.

In both cases it is possible to relax Assumption 6 from Section 4.2, if this assumption turns out to be unrealistic. The pipelines in the model should then be separated into segments that belong to a single pump. All variables and constraints would then hold for each segment instead of each pipeline. This would make the model larger with every segment that is added, because every extra segment means extra variables and constraints (including the piecewise linear approximation).

4.6.2 Uncertainty in modelling changeover costs

Section 4.5.2 explains the changeover costs. We assumed that the changeover costs are constant for every changeover at a reservoir meaning that the costs of a changeover are independent of the size of the changeover (which is the change in production rate). If the changeover costs, however, do depend on the size of the changeover, the model needs to be changed. It then could also be the case that the costs for increasing production are different from the costs for decreasing production. These relations can be modelled using linear equations if there exists a linear relation between the change in production rate and the changeover costs. We would require extra variables $PCH_{r,t}$ and $NCH_{r,t}$ which are respectively the positive and negative production changeover rates of reservoir r in time period t. These variables would denote the positive and negative difference in production rate between time period t and time period $t-1$, which means that $PCH_{r,t} = \max\{PR_{r,t} - PR_{r,t-1}, 0\}$ and $NCH_{r,t} = \max\{PR_{r,t-1} - PR_{r,t}, 0\}$. This definition can be modelled in constraints similar to (29) and (30), since $BSS_{st,t} = \max\{safinv_{st} - I_{st,t}, 0\}$. Both $PCH_{r,t}$ and $NCH_{r,t}$ are then multiplied with a cost factor in the objective function. With this definition variable $CH_{r,t}$ and constraints (25) and (26) would become obsolete en can be omitted, meaning that our model would change from a Mixed Integer Linear Program (MILP) to a Linear Program (LP), since $CH_{r,t}$ is the only integer variable in it.

If there is a non-linear relation between the change in production rate and changeover costs, then the relation can be modelled using variables $PCH_{r,t}$ and $NCH_{r,t}$ in the same way as is done for the relation between energy costs and variable $PTD_{p,t}$ in constraints (21), (22), and (23). There is a non-linear relation when, for example, increases in production have to be spread over all reservoirs in order to obtain the lowest cost. Parameter $sbrp_{p,t}$ then needs to be changed accordingly to the changeover costs as a function of change in production rate. This would increase the model size significantly, since a piecewise linear function is added for each combination of reservoir r and time period t. If the changeover costs are a convex function of the change in production rate then no extra variables are needed, otherwise extra integer variables are needed as in (31) and (32).

4.7 Model simplifications

In Section 2.2 it became clear that the model needs to be solved in a few minutes. We therefore propose a few simplifications for the model, which can be solved if the original model takes too much time to solve. In a MILP, which our model is, performance can be improved by reducing the number of integer variables in the model. Our model only has integer variables to model changeovers. Variable $CH_{r,t}$ takes value 1 if a changeover occurs between time period t and time period $t-1$ at reservoir r and 0 otherwise, meaning that there exists an integer variable for every combination of r and t in the model. Our simplification is based on reducing the number of integer variables by excluding certain of those combinations from the model. We use the assumption that the costs are independent of the size of the changeover, so the model will prefer one large changeover over multiple smaller ones. The simplifications of this section, therefore, do not hold if our assumption is wrong and one of the alternatives of Section 4.6.2 is implemented.

First, we look at the reservoirs r for which all integer variables are defined. From the instance that needs to be solved it can be deduced at which reservoirs takeovers will take place, if changeovers take place at all. For example, if the costs of a changeover are equal at all reservoirs and an increase in production rate is needed, the changeover will be realized at the reservoir that is the closest to Node 1 (all oil passes Node 1 on its way to the customer), so the energy costs for transporting the extra oil are minimized. On the other hand a decrease in production rate will be realized at the reservoir that is the farthest away from Node 1 to save the most energy costs. In most cases changeovers will only occur at one reservoir, since the costs are per changeover. Therefore, a changeover will mostly take place at a large reservoir so the increase or decrease can take place at one single reservoir. Based on this reasoning, the number of reservoirs where changeovers are allowed could be limited to only one or two, which reduces the number of changeovers by 80% or 90% and boosts model performance.

Second, we look at the time periods t. The start and end times of time periods are defined by the arrival and departure times of tankers and the start and end of a day, which results in a maximum of 12 time periods a day, so every few hours it is again evaluated if a changeover is needed. By only evaluating the decision of making a changeover at the start and end of a day the number of integer variables can also be reduced significantly. The timing of a changeover is then less precise in the solution, which could result in slightly higher costs, but the model can be solved again by allowing changeovers in the time periods that start and end in the days before and after the changeover(s) at the same reservoir(s) in the less precise solution. Solving two much smaller models after each other can still be faster than solving one larger model, while the solution is the same.

Finally we can also improve performance by looking at the size of parameter M in constraints (25) and (26), which is multiplied with the integer variables. We have put the value of M on the difference between maximum and minimum production rate, as that is the biggest possible changeover size. Based on the considered case M could be made smaller by increasing minimum production and/or decreasing maximum production. For example, if we have an instance for which an increase in production rate is expected we can set the minimum production equal to the current production, meaning that the solution space becomes smaller, since M is smaller and therefore the model is solved faster.

We test these simplifications and compare it with the model from Section 5.2.2 in Section 6.4 to see if the simplified model yields a solution faster.

4.8 Summary

In this chapter we have formulated our model for the optimization module of the O3, which can also be found in Appendix D. First, we explained how we deal with time by introducing the event based time representation. Next, we stated the assumptions we require to capture OilCO's processes into a model and clarified how we model OilCO's network. Moreover, we gave an overview of all operational constraints our model takes into account and the objective function and its cost components. Our proposed model is a MILP, since we defined binary variables to model the changeover costs. We split the energy costs into a linear part and a non-linear part of which the non-linear part is approximated using a piecewise linear function. Our model minimizes the sum of energy costs, changeover costs, inventory costs, shortage penalties, production target deviation costs, and safety stock penalties. Furthermore, we indicated the uncertainty in the modelling of the energy costs and changeover costs with possible alterations. Finally, we proposed some simplification for our model, which reduces the number of integer variables of our model and decreases the size of parameter M and therefore improves performance. In Chapter 5 we extend our model such that it can deal with disturbances.

5 Dealing with disturbances

In Chapter 4 we formulated a model for the O3. This is a deterministic model as all input parameters are considered as fixed values. In reality, however, uncertainty plays a role in OilCO's daily operations as disturbances can occur during daily operations. In Section 2.2 the relevant disturbances were mentioned: delayed tankers, opportunities on the spot market, maintenance of assets, and asset breakdowns. In Section 2.2 we also mentioned that we use both reactive as predictive scheduling to deal with these disturbances. The reactive scheduling is the what-if analysis as proposed by OilCO, which we treat in Section 5.1. The predictive scheduling is the RO as discussed in Section 3.3. Section 5.2 explains how the deterministic model of Section 4 has to be adjusted in order to yield a solution for OilCO that is more robust to disturbances. Section 5.3 summarizes this chapter.

5.1 What-if analysis

The planner at OilCO can use the what-if analysis to simulate different scenarios and compare the consequences of alternative actions when a disturbance occurs. The O3 does not have to take the decision on how to react on a disturbance, but it has to give the planner a clear overview of the different alternatives such that the planner can take the decision. In this section we treat the different disturbances for the what-if analysis separately. For each disturbance we indicate how the model is solved and what decisions can be taken using the analysis. Section 5.1.1 covers the breakdown of assets, Section 5.1.2 covers the asset maintenance, and Section 5.1.3 covers the spot market opportunities. In all these sections the base case is the current optimized schedule, which was implemented before the disturbance occurred.

5.1.1 Asset breakdown

At OilCO asset breakdowns occur, meaning that an asset (partly) stops working and cannot be used entirely for a certain period, while that was unplanned. A breakdown can happen at every reservoir, separation facility, and pipeline. When an asset breakdown occurs the planner has to decide if a reaction is required and if yes, which reaction is required. The main concern when an asset breakdown occurs is if the demand can still be fulfilled in the coming periods. First an estimate is required of how long it will take to solve the breakdown, which has to be expressed in a percentage of the maximum rate in a time period. For example a pipeline breakdown can result in an availability of 0% for the first three days after the breakdown, an availability of 25% on the fourth day and 60% on the fifth day. The pipeline is then fully available again on the sixth day. This estimate is then entered into the model using the availability parameters, for the example of pipeline p' this would be: $ava_{p',t} = 0 \; t \in \{T_1 \cup T_2 \cup T_3\}$, $ava_{p',t} = 25 \; t \in T_4$, and $ava_{p',t} = 60 \; t \in T_5$. If we solve the model with those parameter values we get optimal reaction on the breakdown. However, in case the broken down asset is a reservoir, this approach yields an unrealistic solution, because every change in production rate during maintenance results in a changeover. The production rate would be lower than is possible in the periods with 25% and 60% availability to avoid extra changeovers, which has to be avoided by fixing the production rates to the maximum possible production rates and neglecting the changeover costs for the reservoir in the first periods.

Furthermore, it is possible to solve the model with some fixed variables. If all variables that concern parts of the network that are not influenced by the broken down asset (all variables concerning assets upstream or downstream of the broken down asset are influenced) are fixed, then we can see the consequences of attaining the current schedule when solving the model. The what-if analysis can then show the difference between doing nothing when the breakdown occurs (the base case with a lower availability), or changing the schedule (re-optimizing with lower availability). The analysis can therefore help the planner in deciding what to do when an asset breaks down.

5.1.2 Asset maintenance

Sometimes an asset requires maintenance, which means that an asset cannot be (entirely) used for a certain time period. Reservoirs, separation facilities, and pipelines require maintenance periodically. The O3 has to advice on what the consequences are of scheduling maintenance in a certain period, using what-if analysis. First, an estimate is required of what the impact the maintenance has on the availability of the asset in the same way as for the asset breakdowns. Subsequently, the model can be solved for different scenarios of the moment the maintenance is scheduled. The planner can compare the different scenarios to determine when the maintenance takes place. The planner can also compare the scenarios with the base case to determine if it is a good idea at all to schedule maintenance in the coming 30 days.

5.1.3 Spot market opportunities

OilCO sometimes gets an opportunity to sell oil on the spot market. The spot market allows OilCO to sell its oil on the short term for higher prices than on the regular market, but this is only known a few days or weeks in advance. The planner has to decide if OilCO takes the opportunity or not. In the what-if analysis a tanker is added to the schedule for the moment and volume of the spot market opportunity. The planner can solve the model with this adjusted terminal demand (parameter $tdem_{te,t}$) to see what the consequences (compared with the base case) of taking the opportunity are. These consequences have to be compared with the revenue of the opportunity, since our model only concerns cost. If there is flexibility in the timing of the spot market opportunity is delivered, the model can be solved with different loading moments to determine the best moment to schedule the extra tanker (or to decide to decline the opportunity).

5.2 Robust Optimization applied to OilCO

In this section we apply Robust Optimization (RO) to OilCO's situation. In Section 3.3 we gave an overview of RO techniques that are useful for OilCO. We extend the model of Chapter 4 to come to a model that yields solutions that are less sensitive to tanker delay. In Section 5.2.1 we explain how we model tanker delay and in Section 5.2.2 we propose a RO model for OilCO. The notation for the RO is introduced in 5.2.2, but can also be found in Appendix B. The complete RO model can also be found in Appendix E.

5.2.1 Tanker arrival uncertainty

For OilCO we want to generate a schedule that is robust to uncertainty in tanker arrival times. All terminal demand at OilCO is fulfilled eventually, so there is no uncertainty in the total demand, but there is uncertainty in when that demand exactly occurs. In the model this means that parameters $tdem_{te,t}$, which are the demand volume at terminal te during time period t, can change if there is a tanker delay, but $\sum_t tdem_{te,t}$ is the same for every realization of the tanker schedule. We therefore cannot model demand as independent random variables as is done in the model of Li et al. (2012). We apply the scenario approach of Wang & Rong (2010) to model tanker arrival uncertainty, since tanker schedule scenarios at OilCO are easy to define and total demand is not affected. A tanker schedule scenario consists of realizations of all the tanker arrivals during the scheduling horizon. OilCO agrees on a three day window with the customers' tankers, meaning that the tankers have to arrive at the agreed terminal within that window. At least three days before the beginning of the three day window, tankers have to confirm an exact arrival day out of the three days from the window. As we have no data from OilCO on tanker arrivals, we have to make some assumptions here. We assume that tankers that have not confirmed a day yet always arrive within their three day window. To avoid being late, tankers always aim to arrive on the first day of the window. Note that, in the deterministic model we also assume that all tankers arrive on the first day. Tankers that have already confirmed a day can have a delay of one day (although only with a small probability), even if that means that they arrive outside their three day window.

For our event based time representation we want to keep the number of time periods to a minimum. Therefore, we use a discrete probability distribution to limit the number of possible realizations per tanker. Every tanker that has not confirmed a date yet has a probability to either arrive on the first, second, or on the third day of the three day window. A tanker can arrive on every whole hour in a day. We assume that a tanker always arrives on the same hour in a day, so we require a maximum of two extra events per tanker in our event based time representation. Every tanker that has not confirmed a date yet therefore has three possible arrival moments. Every tanker that has confirmed a date arrives either on the confirmed date or one day later.

Based on our assumption of a maximum of five tankers per day in our 30 day horizon, we have a maximum of $(27 * 5)^3 + (3 * 5)^2 = 2,460,600$ possible tanker arrival scenarios. In contrast to Wang & Rong (2010), who consider only three scenarios in their model and evaluate all three, it is impossible for us to evaluate all scenarios in our model. We therefore apply Sample Average Approximation (Verweij et al., 2003), where a number of scenarios is generated by random sampling and the sample average is minimized. The set Ω denotes all scenarios that can occur at OilCO during the scheduling horizon and set S (with $S \subset \Omega$) denotes the set of samples of tanker arrival scenarios that we consider for our model.

5.2.2 Robust Model

To define a RO model we have to identify which part of the model is affected by the uncertainty (control variables) and which part is not (design variables). The uncertain parameter $tdem_{te,t}$ is calculated from the tanker schedule by multiplying the number of tankers that is loaded at terminal te during time period t by the loading rate (50,000 bbl/hour) and the duration of time period t in hours. A variable that is directly influenced by parameter $tdem_{te,t}$ (as can be seen in (15)) is $SHT_{te,t}$, which is the total shortage volume at terminal te at the end of time period t. Based on this, also the throughput rates of the pipelines from the terminals to the terminal storage tanks have to be adapted. We call this subset the control pipelines CP (with $CP \subset P$). Variable $TH_{p,t}$, which is the throughput rate of control pipeline p (with $p \in CP$) in time period t, is a control variable, because the amount of oil pumped to the terminal depends on when the tanker exactly arrives. The inventory levels of the storage tanks at the terminals are indirectly influenced, because these inventory levels depend on when the oil is pumped from the tanks to the terminals. Again, we call this subset the control storage tanks CST (with $CST \subset ST$). The inventory levels of these tanks are therefore a control variable. The throughput rates of all pipelines for which energy costs are relevant (subset EP) have to be design variables, as these are long pipelines for which it is impossible to adjust the throughput rates on a short notice. Therefore, the variables concerning the terminal locations (terminal storage tanks, the pipelines at the terminals, and the terminals themselves) are control variables and variables concerning all other parts of the network are design variables.

All control variables are scenario dependent. Parameter $tdem^s_{te,t}$ denotes the demand in bbl at terminal te during time period t in scenario s. Variables $SHT^s_{te,t}$, $TH^s_{p,t}$ (with $p \in CP$), and $I^s_{st,t}$ (with $st \in CST$) are defined in the same manner, which also holds for the objective function variables that directly depend on the control variables: $AI^s_{st,t}$ and $BSS^s_{st,t}$ (both with $st \in CST$). The pipelines in CP are not in set EP ($CP \cap EP = \varnothing$), because the pipelines connect assets on the same location and energy costs are therefore not relevant. The random variable ξ_s, which is the objective function (18) of scenario s, is written in the form $\xi_s = c^T x + d_s^T y_s$:

$$\begin{aligned}\xi_s = &\sum_{p,t} dur_t\, EN_{p,t} + \sum_{r,t} cc_r\, CH_{r,t} + \sum_{rf,t} cshr_{rf}\, SHR_{rf,t} \\ &+ \sum_{r,d} cdev_r\, DEV_{r,d} + \sum_{st \notin CST,t} ssp_{st}\, BSS_{st,t} + \sum_{st \notin CST,t} cinv_{st}\, dur_t\, AI_{st,t} \\ &+ \sum_{st \in CST,t} ssp_{st}\, BSS^s_{st,t} + \sum_{st \in CST,t} cinv_{st}\, dur_t\, AI^s_{st,t} + \sum_{te,t} csht_{te}\, SHT^s_{te,t} \end{aligned} \tag{33}$$

For the RO model of OilCO we need to specify a $\sigma(.)$ function, which depends on variable ξ_s, and a penalty function $\rho(.)$. We do not know how much risk OilCO's planners are willing to take, but we take the $\sigma(.)$ function that includes risk (see Equation (3)) as a basis, just in case OilCO's planners are risk-averse. In this function parameter λ can be used to make a trade-off between minimizing the risk term or the expected value of the variables, where an increase in λ results in an increased focus on risk term. In case OilCO's planners do not want to take risk into account, λ can be set to 0. The expected value of random variable ξ_s ($E[\xi_s] = \sum_{s \in \Omega} p_s \xi_s$) can be written as $E[\xi_s] = c^T x + \sum_{s \in \Omega} p_s d_s^T y_s$. In the Sample Average Approximation the expected value is replaced by the sample average (Verweij et al., 2003), meaning that p_s is replaced by $\frac{1}{|S|}$ and thus every scenario gets the same weight in the objective. We also replace the variance by the sample variance term, this is also done by replacing p_s by $\frac{1}{|S|}$. Since there is no uncertainty in the design variables, taking all terms that do not depend on s out of the variance term does not influence the variance of ξ_s. As the quadratic term in the variance leads to a non-linear model, we apply the same method as (Wang & Rong, 2010) by replacing the variance term in (3) by the absolute deviation from the expected value. This leads to the following $\sigma(.)$ function:

$$
\begin{aligned}
\sigma(.) &= \sum_{s\in\Omega}\frac{1}{|S|}\xi_s + \lambda\sum_{s\in\Omega}\frac{1}{|S|}|\xi_s - \sum_{s'\in\Omega}\frac{1}{|S|}\xi_{s'}| \\
&= c^T x + \sum_{s\in\Omega}\frac{1}{|S|}d_s^T y_s + \lambda\sum_{s\in\Omega}\frac{1}{|S|}|d_s^T y_s - \sum_{s'\in\Omega}\frac{1}{|S|}d_{s'}^T y_{s'}| \\
&= \sum_{p,t} dur_t\, EN_{p,t} + \sum_{r,t} cc_r\, CH_{r,t} + \sum_{rf,t} cshr_{rf}\, SHR_{rf,t} \\
&+ \sum_{r,d} cdev_r\, DEV_{r,d} + \sum_{st\notin CST,t} ssp_{st}\, BSS_{st,t} + \sum_{st\notin CST,t} cinv_{st}\, dur_t\, AI_{st,t} \\
&+ \frac{1}{|S|}\sum_{s}(\sum_{st,t} ssp_{st}\, BSS^s_{st,t} + \sum_{st,t} cinv_{st}\, dur_t\, AI^s_{st,t} + \sum_{te,t} csht_{te}\, SHT^s_{te,t}) \\
&+ \lambda\frac{1}{|S|}\sum_{s}|\sum_{st\in CST,t} ssp_{st}\, BSS^s_{st,t} + \sum_{st\in CST,t} cinv_{st}\, dur_t\, AI^s_{st,t} + \sum_{te,t} csht_{te}\, SHT^s_{te,t} \\
&- \sum_{s'}\frac{1}{|S|}(\sum_{st\in CST,t} ssp_{st}\, BSS^{s'}_{st,t} + \sum_{st\in CST,t} cinv_{st}\, dur_t\, AI^{s'}_{st,t} + \sum_{te,t} csht_{te}\, SHT^{s'}_{te,t})| \qquad (34)
\end{aligned}
$$

In (34) the third, fourth and fifth line denote the sample average of ξ_s, of which the third and fourth line denote the value of the design variables and the fifth line denotes the sample average of the control variables. If we want to do a worst case analysis then the $\sum_s$ on the fifth line only has to be replaced by $\max_s$. The last two lines denote the risk term.

We now take a look at the penalty function $\rho(.)$. In the RO model of Wang & Rong (2010) infeasibilities in violating the minimum and maximum storage capacity are penalized in the objective function. For OilCO only infeasibilities in violating the maximum storage capacity are relevant, as no realization of tanker delay can result in a violated minimum storage capacity. A delayed tanker can cause an infeasibility in maximum storage capacity, because a tanker delay causes a delay in the discharge of oil from the terminal tanks, while this tank keeps receiving oil. In (2) penalties are measured by variable z_s,where we define variable $Z^s_{st,t}$ as the violation in bbl at storage tank st at the end of time period t in scenario s, with $Z^s_{st,t} \geq 0$. In order to measure the violation, we extend (13) from Section 4 as follows:

$$
mininv_{st} \leq I^s_{st,t} \leq maxinv_{st} + Z^s_{st,t} \qquad \forall st \in CST;\, t \in T\,;\, s \in S \qquad (35)
$$

For all $st \notin CST$, (13) still holds. Since $Z^s_{st,t}$ can only take positive values, we only have to consider positive violations of the constraints. A commonly used penalty function for that situation is $\sum_{s\in\Omega} p_s\, Z^s_{st,t}$ (Mulvey et al., 1995), which we also use. This leads to the following penalty function for OilCO:

$$
\rho(.) = \frac{1}{|S|}\sum_{s}(\sum_{st\in CST,t} pen_{st}\, Z^s_{st,t}) \qquad (36)
$$

Where pen_{st} is the penalty in USD per bbl for violating the maximum inventory of storage tank st, which is similar to the formulation of Wang & Rong (2010).

The objective function of the RO model consists of the sum of (34) and (36), so the overall objective function is:

$$\begin{aligned}
\min \sum_{p,t} dur_t\, EN_{p,t} + \sum_{r,t} cc_r\, CH_{r,t} + \sum_{rf,t} cshr_{rf}\, SHR_{rf,t} \\
+ \sum_{r,d} cdev_r\, DEV_{r,d} + \sum_{st \notin CST,t} ssp_{st}\, BSS_{st,t} + \sum_{st \notin CST,t} cinv_{st}\, dur_t\, AI_{st,t} \\
+ CSA + \lambda \frac{1}{|S|} \sum_{s} | \sum_{st \in CST,t} ssp_{st}\, BSS^{s}_{st,t} + \sum_{st \in CST,t} cinv_{st}\, dur_t\, AI^{s}_{st,t} \\
+ \sum_{te,t} csht_{te}\, SHT^{s}_{te,t} - CSA| + \omega \frac{1}{|S|} \sum_{s} (\sum_{st \in CST,t} pen_{st}\, Z^{s}_{st,t}) \qquad (37)
\end{aligned}$$

where $CSA = \frac{1}{|S|} \sum_s (\sum\limits_{st \in CST,t} ssp_{st}\, BSS^{s}_{st,t} + \sum\limits_{st \in CST,t} cinv_{st}\, dur_t\, AI^{s}_{st,t} + \sum\limits_{te,t} csht_{te}\, SHT^{s}_{te,t})$, which denotes sample average of the control variables. The parameter ω can be used by the planner to make trade-offs, just as the parameter λ. The parameter ω can be used to make a trade-off between model robustness and minimizing costs. A higher value for ω results in an increased focus on model robustness, as violations of the maximum storage capacity are penalized more. In Section 6 we test this model for different values of λ and ω.

5.3 Summary

In this chapter we clarified how to use the deterministic model of Chapter 4 to deal with disturbances as maintenance, asset breakdowns, spot market opportunities, and tanker delay. First, we explained how we can use our deterministic model in a what-if analysis (Section 5.1) for dealing with maintenance, asset breakdowns, and spot market opportunities. Every disturbance has its own approach here. For dealing with tanker delay we applied RO to our model (Section 5.2). We applied RO by dividing the variables in the deterministic model into design and control variables and using a scenario approach to model tanker delay. A scenario consists of a realization of delay for every tanker within the scheduling horizon. As the number of possible scenarios is large, it is impossible to evaluate all possible scenarios. Therefore, we use a Sample Average Approximation. We chose to incorporate a risk term in the objective function and to penalize infeasibilities in the storage tank capacity, which leads to a model where we had to redefine one set of constraints (35). The new objective function is shown in (37).

In Chapter 6 we compare our RO model with the deterministic model of Chapter 4 for different values of λ and ω. Finally we check if the simplifications from Section **??** yields solutions in less computational time.

6 Numerical results

In this chapter we perform numerical experiments to examine the performance of our models proposed in Chapters 4 and 5. In this chapter we perform three different experiments to test these models. First we describe the values of the different parameters we use in the model for the experiments (Section 6.1). In the first experiment we solve the deterministic model for different tanker arrival cases and compare the solutions with each other (Section 6.2). Next, we test the robust model in terms of solution value, required CPU time and robustness. We test the robust model for different values of λ and ω (Section 6.3) and compare it with the deterministic model. Finally, we test the simplifications from Section 4.7 to see if it indeed yields the same solutions and solves faster (Section 6.4). We perform all experiments on a computer with 2.60 GHz processor and 8.00 GB RAM using AIMMS 3.13.6.213 with solver CPLEX 12.6. Section 6.5 concludes this chapter.

6.1 Parameter input

In this section we determine the parameter data that we use in the experiments. In Section 2.3.2 we indicated that there is uncertainty about parameter values and the values of the parameters given in this section are therefore not certain. We make educated guesses for these values based on the limited information supplied by OilCO, the experience within OCG, and internet research. As most of the values in this chapter are guessed, we perform a sensitivity analysis in Chapter 7 on the parameters that have a large influence on the model. In this section we explain the parameter input by the different parts of the model: the network input (Section 6.1.1), demand input (Section 6.1.2), and cost factor input (Section 6.1.3).

6.1.1 Network input

The production volumes in the production plan are constant for every day during the planning horizon. The production volumes are oil volumes, which are equal to the volumes of Table 1. The gas volume is calculated using the Gas Oil Ratio (GOR) of the specific reservoir. The GOR differs per reservoir and ranges from 200 to 900, as these are normal values for oil that is extracted from reservoirs. The maximum production per reservoir is set on 1.2 times the production plan value and the minimum production per reservoir on 0.4 times the production plan value. The current production rate is equal to the rate that matches the production target of the first day. All reservoirs, separation facilities, and pipelines have a availability of 100%. Table 2 shows the reservoir input data.

The maximum separation rate of a separation facility is calculated by adding the maximum production rates of the reservoirs that are connected to that separation facility. The oil loss is 0.03 in all separation facilities. Most data on pipelines including the information required for set C (the set that indicates the pipeline connections between assets) can be found in Appendix A. Here, the pipelines of which the capacity is unknown are unconstrained and the minimum throughput is 20% of the pipeline capacity. The terminal storage tanks have a maximum inventory of 3,000,000 bbl, while the other storage tanks have a maximum inventory level of 1,500,000 bbl. The maximum inventory level parameter is examined in the sensitivity analysis of Chapter 7. The minimum inventory level of all storage tanks is 0 bbl, although it is higher in practice, to have round numbers for both the maximum inventory and the difference between maximum and minimum inventory. The safety stock levels and current inventory levels are $\frac{1}{6}$ and $\frac{1}{3}$ of the maximum inventory levels.

Reservoir	Regular production (kbbl/day)	Max production (kbbl/day)	Min production (kbbl/day)	GOR (cf/bbl)
Reservoir0	800	960	320	250
Reservoir1	152	182.4	60.8	350
Reservoir2	108	129.6	43.2	600
Reservoir3	60	72	24	400
Reservoir4	713	855.6	285.2	500
Reservoir5	38	45.6	15.2	700
Reservoir6	66	79.2	26.4	300
Reservoir7	84	100.8	33.6	650
Reservoir8	55	66	22	800
Reservoir9	465	558	186	450

Table 2: Reservoir input data

6.1.2 Demand input

The oil demand of both refineries is a constant 107,385 bbl per day, which leaves 2,250,000 bbl of the regular oil production for the terminals. The demand is spread over both terminals based on the number of loading platforms. Terminal1 has three loading platforms, whereas Terminal2 has two, which results in an average demand per day of 1,350,000 bbl for Terminal1 and 900,000 bbl for Terminal2. For the terminals this demand is not constant per day, because demand depends on when the tankers arrive. We consider two different demand cases in our model, which are different tanker scenarios for a 31 day period, while the average demand is the same. The first case is based on a tanker arrival schedule what we think is realistic for OilCO when taking tanker sizes[3], number of arrival platforms at the terminals, and production volumes into account. The second case is a demand case in which demand is less constant over time as all tankers have a maximum tanker size of 2,000 kbbl and demand is not spread in the same proportion over Terminal1 and Terminal2. The second also has a peak in demand in the second week, with this peak we want to force a changeover in the solution, so that we can examine how our model deals with that. Figure 9 shows the Gantt chart for both the stable and the unstable case. The Gannt chart shows per terminal the arrival and departure moments of the tankers. Recall that the loading duration of a tanker is linear with the size of the tanker, which results in an equal loading duration for all tankers in the unstable case. The arrival hour of a tanker on a day is randomly chosen, as a random integer between 0 and 23 is generated.

The Gantt chart of the unstable case shows that there is a peak in demand from day 11 to day 15, where the demand decreases significantly after day 15 as less tankers arrive. We consider this case to see how our model deals with less constant demand pattern and what the differences are with the case with stable demand pattern. We treat these two cases separately in Section 6.3 to be able to compare the results. At the terminals and the refineries the current total shortage volume is 0 and the BuyGas demand is 1,800,000,000 cf per day for the entire planning horizon.

Recall that we assume that tankers that have not confirmed a day yet always arrive within their

[3]Tanker sizes normally vary between 200 and 2,000 kbbl.

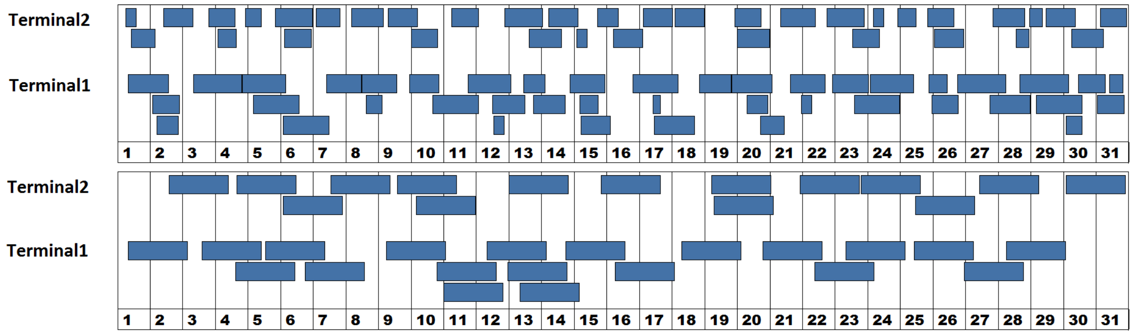

Figure 8: Gantt charts of the stable (above) and unstable (below) demand case

three day window. Furthermore, we assume that tankers always aim to arrive on the first day of the window, since tankers want to avoid being late due to the corresponding penalties. More precisely, we assume that tankers that have not confirmed a day have a probability of 0.6 to arrive on the first day of the window, a probability of 0.3 to arrive on the second day of the window and 0.1 to arrive on the third day of the window. Tankers that confirmed an arrival day have a probability of 0.2 to have a delay of one day.

6.1.3 Cost factor input

Our model has seven cost components and nine cost parameters, as all cost components have one cost parameter except the energy costs with three parameters. We do not know the exact values of these parameters. Therefore, we cannot retrieve the real costs that OilCO makes in the experiment. We, however, want the model to yield the best solution whatever the objective value may be. We therefore focus on the values of the cost components relative to each other. In the sensitivity analysis of Chapter 7 we change the cost parameters one by one and see what the influence is on the solution. For the energy costs we have three cost factors: ca_p, cb_p, and csq_p (see (24)). The value of cb_p has no influence on the solution value, since it is a constant. The ratio between ca_p and csq_q determines the shape of the cost function, where we want a function that looks like the function of our example in Figure 6, since this figure gives an example of what the energy costs as a function of the throughput are. In our example cb_p lies around 700. In Figure 9 we see the energy costs from Figure 6 including an approximation with $ca_p = 30$, $cb_p = 700$, and $csq_p = 1$. In this example the peak efficiency throughput is again half the maximum throughput ($TH_p^* = \frac{1}{2}TH_p^{max}$). The part of the approximation where throughput is higher than the peak efficiency throughput does not exactly match the function (indicated by the red plane), but both functions are convex so they have a similar behavior. Therefore, we take $ca_p = 30\,csq_p$ for every pipeline. To determine the ratio of these factors between pipelines we take the ratio of the lengths of the pipelines, since it takes more energy to realize a certain throughput rate at a longer pipeline.

For now we put the peak efficiency throughput on 95% of the "regular throughput". This regular throughput is determined by solving the deterministic model with constant average terminal demand. The regular throughput is calculated afterwards by taking the average throughput per pipeline. This approach is based on the assumption that OilCO designed its network such that it operates at the maximum efficiency throughput. We put the maximum efficiency throughput TH_p^* not on the regular

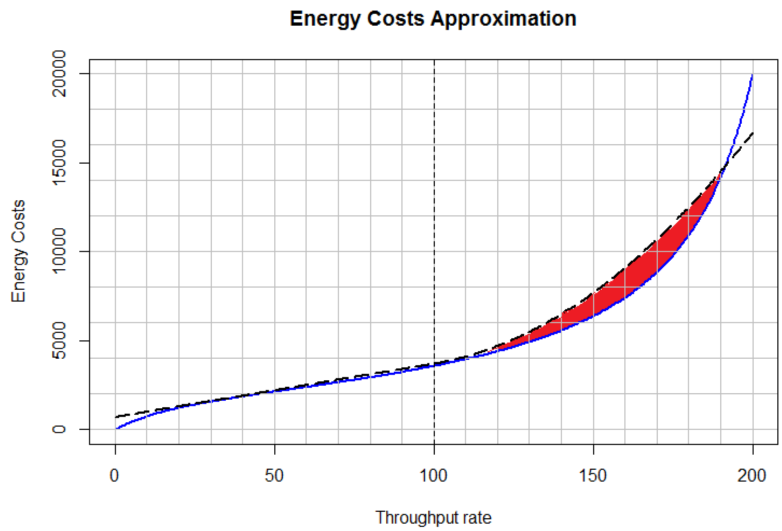

Figure 9: Energy cost approximation

throughput, but on 95% of the regular throughput, since we assume that the production has increased since the the network was designed. In Section 7 we perform a sensitivity analysis on the value of TH_p^*, since we are not sure about these assumptions, while the value of TH_p^* has a large influence on energy costs, since it determines for which values of the throughput rate the energy costs function is non-linear. We use 11 breakpoints for the piecewise linear approximation of the non-linear part of the function. This number is justified in Appendix F.

We put the costs of a changeover on 10,000 USD, since the changeover cost factor has to be the highest of all. Furthermore, the penalty of a shortage at the refinery ($cshr_{rf}$) is set on 40 USD per bbl and the penalty of a shortage at the terminals ($csht_{te}$) is set on 20 USD per bbl, since this implies that a terminal shortage is cheaper than a shortage at the refinery. A shortage at a terminal is preferred, since there is more flexibility in planning at the terminals which makes it possible to overcome the shortage in the realization of the schedule. The penalties for deviating from the production plan ($cdev_r$) are set to 1 USD per kbbl deviation from the plan on a certain day. The inventory costs have the lowest cost factor, these are 0.0001 USD per kbbl per day. The safety stock penalties ssp_{st} should be higher than the inventory cost to have effect on the inventory level, so they are set on 0.01 USD per kbbl per day.

In the robust model the maximum inventory level violations penalties (pen_{st}) are 0.1 USD per kbbl. In the robust model we use a sample size of 80 scenarios, of which the number is justified in Appendix G.

6.2 Deterministic model results

We solved the deterministic model for both cases. Table 3 shows the results of this experiment, in which all cost factors and the CPU time are shown. For the stable demand case only the energy costs, safety stock penalties, and inventory costs are larger than zero, since the cost factors of these cost components are high and these costs can be avoided. In Chapter 7 we perform experiments to determine for which values of the cost factors these cost components get larger than zero in the stable demand case. For the unstable demand case, there are also additional changeover costs, production plan deviation penalties, and terminal shortage penalties. Both cases are solved within 12 seconds, which is excellent considering our goal from Section 2.2 that the model needs to be solved in a couple a minutes.

Demand case	Stable demand	Unstable demand
Total costs (USD)	2,401	26,348
Energy costs (USD)	2,376	13,406
Changeover costs (USD)	0	10,000
Inventory costs (USD)	23	18
Safety stock penalties (USD)	2	90
Terminal shortage penalties (USD)	0	142
Refinery shortage penalties (USD)	0	0
Production target deviation penalties (USD)	0	2,883
CPU time (s)	4.5	11.9

Table 3: Deterministic model results for both cases

In the solution of the stable demand case the throughput rates of the pipelines are more or less constant, because that is the cheapest way to fulfill the demand. This constant rate is only possible since the demand is stable. For a more unstable demand case, changes in throughput rate are needed, which is more expensive due to the fact that the energy costs are a convex function of the throughput. When the demand is perfectly constant the energy costs are 2,376 USD, which is the same as for the stable demand case. In the solutions there is a difference between the behavior of the inventory levels of the terminal storage tanks and the storage tanks at the separation facilities. Figure 10 gives as an example of the inventory levels of both types of storage tanks for the stable demand case.

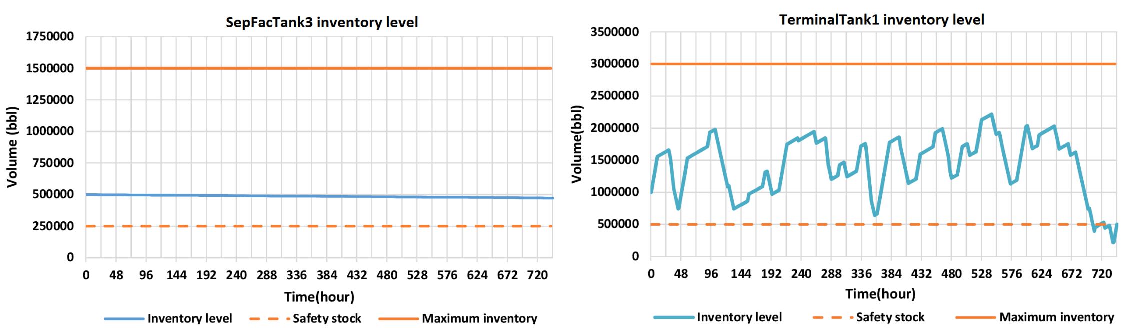

Figure 10: Inventory levels of a separation facility and a terminal storage tank for the stable demand case

The terminal storage tank inventory levels fluctuate, since the incoming flow is constant, but the outgoing flow fluctuates. The start and end times of the events of the event based time representation can be seen in Figure 10, since these are the points on which the throughput rate from the terminal storage tank to the terminal changes. The separation facility storage tank inventory levels are more or less a straight line, since both the incoming and outgoing flow are constant.

For the unstable demand case the solution is different, as expected, since the demand pattern is also different. As the throughput rate fluctuations are very large in certain periods to deal with the peak in demand, the energy costs are higher. In the solution a changeover occurs, as the production rate increases at the Reservoir9 to its maximum at the beginning of the planning horizon in order to have enough oil available for the peak demand. After the peak demand the production rate could be lowered back to the original production plan (or even lower), but this extra changeover is more expensive than the additional production plan deviation penalties and energy costs, so there is only one changeover. When the changeover cost factor is lowered to 1,505 USD or lower 2 changeovers or more occur in the solution: one at the beginning of the planning horizon and one after the peak demand. The extra production volumes caused by the changeover do not result in higher inventory costs as in the stable demand case, since the extra produced volume is pumped into tankers between day 10 and 14. As the safety stock is needed to fulfill the peak in demand, the safety stock penalties are higher for the unstable demand case.

Figure 11 shows the inventory levels of a separation facility storage tank and a terminal storage tank for the unstable demand case, these are different from the inventory levels of Figure 10, as the separation facility storage tank inventory levels fluctuate more and the terminal storage tank inventory level has larger fluctuations because the tanker loads are larger. As the extra production is not enough to cover the peak in demand, both tanks are empty at the end of the peak, which also causes shortages at Terminal1 However, this is only a shortage of 7 kbbl, which is small compared to tanker loads of 2,000 kbbl, resulting in terminal shortage penalties of 141.6 USD. A small shortage of 7 kbbl can probably be eliminated when the schedule is realized, as it takes around 8 minutes to pump 7 kbbl into a tanker.

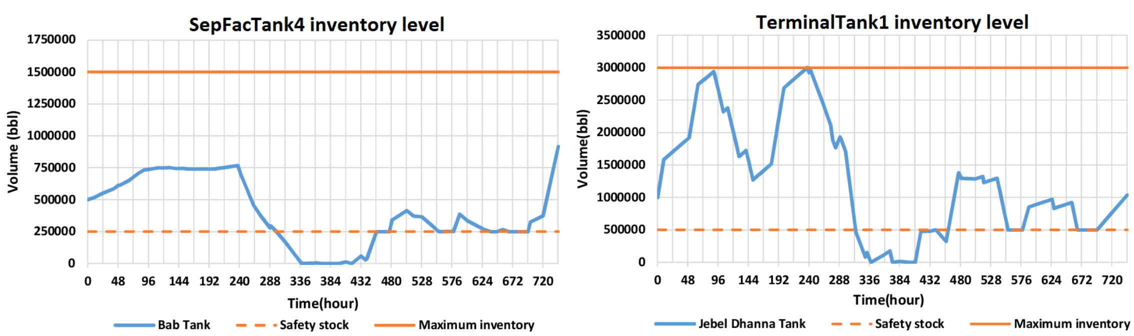

Figure 11: Inventory levels of a separation facility and a terminal storage tank for the unstable demand case

6.3 Robust model results

In this section we test the robust model from Chapter 5 and compare it with the deterministic model. Section 6.3.1 considers the experimental set-up, Section 6.3.2 considers the results of the stable demand case, and Section 6.3.3 considers unstable demand case.

6.3.1 Experimental set-up

In Section 3.3 we identified two sort of robustness: model robustness and solution robustness. We included both types of robustness in our model of which the weight is indicated by parameters λ and ω. Recall that λ determines the the solution robustness as it is multiplied with the risk term in the objective function. The value λ determines the trade-off between minimizing the expected value of the control variables or the "variance" (which we replaced by the absolute deviation from the mean) of the control variables. Recall that ω determines the model robustness as it is multiplied with a term that penalizes infeasibilities. Therefore, parameter ω quantifies the trade-off between minimizing the total costs or minimizing the infeasibility penalties. In this section we want to find out how different values of λ and ω influence our solution. We test the values 0, 1, 10, and, 100 for λ and the values 0, 1, 10, 100 and 1,000 for ω.

Increasing the solution and/or model robustness by increasing λ and/or ω in the model yields, by definition, a solution with higher cost. OilCO has to decide if the increase in costs is worth the increase in model and/or solution robustness. Therefore, we examine what the additional costs are when the robustness is increased and by how much the robustness is increased. We evaluate the solution robustness and model robustness for the different values of λ and ω. We measure solution robustness by the average absolute deviation of the objective value for different scenarios from the sample average. Model robustness is measured by the average violation of the storage tank capacity at the terminal storage tanks (in kbbl) and the percentage of scenarios (of the 80 scenarios we evaluate) that have infeasibilities in them. Also the deterministic schedule is evaluated for both solution and model robustness by re-solving the deterministic model for every scenarios, while fixing all variables concerning the parts of the network that are not influenced by tanker arrival uncertainty (which are the design variables of the robust model). This allows us to evaluate how the terminal tank inventory levels react to a scenario of tanker arrivals and check if the scenario becomes infeasible.

In Section 6.2 we saw that the case with stable demand and the case with unstable demand yield different solutions, which is why we treat them separately in this experiment. First, we consider stable demand case (Section 6.3.2) and subsequently the unstable demand case (Section 6.3.3).

We start the analysis with the results of solving the model with $\lambda = \omega = 0$, which means that both solution robustness and model robustness are not considered as both terms are multiplied with zero. The model then yields a solution by minimizing the average costs over the 80 scenarios, which we want to compare with the deterministic schedule where only one scenario is considered. Next we examine what the different solutions are as λ and/or ω increases, here we first look at how both parameters affect total costs, as increasing one or both parameters by definition increases total costs. Then, we look at how increasing ω influences model robustness and how increasing λ influences solution robustness. Finally, we examine how both parameters affect the CPU time to solve the model.

6.3.2 Experimental results for the stable demand case

Table 4 shows the computational results for this experiment with the stable demand case. The "Infeasibilities" column shows the average volume (in kbbl) of violation of the maximum inventory. The "% infeasible" column shows the percentage of scenarios with infeasibilities in them. The "Abs Dev" column show the average absolute deviation of the costs that result from the control variables of a realization from the sample average, which replaces the variance of the total costs (see Section 5.2.2).

Model	λ	ω	Objective value (USD)	Total costs (USD)	CPU Time (s)	Infeas-ibilities (kbbl)	% infeas-ble	Abs Dev (USD)
Deterministic	N/A	N/A	2,401	2,401	4.5	22,318	96	0.7
Robust	0	0	2,405	2,405	255	12,097	96	0.7
Robust	1	0	2,405	2,405	835	12,231	96	0.7
Robust	10	0	2,406	2,406	902	12,078	96	0
Robust	100	0	2,407	2,407	1,190	12,431	98	0
Robust	0	1	2,809	2,603	427	2,052	71	1.0
Robust	1	1	2,810	2,603	1,399	2,052	71	1.0
Robust	10	1	2,811	2,606	1,061	2,053	71	0
Robust	100	1	2,812	2,607	833	2,052	73	0
Robust	0	10	3,678	3,075	338	603	26	2.1
Robust	1	10	3,680	3,074	1,055	605	26	2.0
Robust	10	10	3,684	3,075	1,012	608	26	0.1
Robust	100	10	3,685	3,077	849	609	26	0
Robust	0	100	6,052	4,222	304	183	16	3.1
Robust	1	100	6,055	4,222	1,124	183	16	3.1
Robust	10	100	6,061	4,230	876	183	16	0.1
Robust	100	100	6,063	4,233	737	183	16	0
Robust	0	1,000	17,720	6,030	254	117	8	611.9
Robust	1	1,000	17,863	5,533	1,126	123	9	4.5
Robust	10	1,000	17,870	5,540	856	123	9	0.2
Robust	100	1,000	17,873	5,546	687	123	9	0

Table 4: Computational results for the stable demand case

Deterministic model vs robust model with $\lambda = \omega = 0$

Table 4 shows that the objective value of the robust model with $\lambda = \omega = 0$ (second row) is almost equal to the objective value of the deterministic model, while the robust model takes almost 50 times more CPU time to solve. This difference in CPU time is caused by the fact that the deterministic model takes one scenario into account when solved and the robust model 80 scenarios.

Table 9 in Appendix H shows the division of total costs over the different cost components of the solutions. Here only the energy costs, inventory costs, terminal shortage penalties, and the penalties for having less in stock than the safety stock are shown, since the other cost components are zero. The table shows that the robust model finds a pipeline schedule with equal energy costs as the deterministic model, but inventory costs are higher. The inventory costs are higher, since the robust model considers scenarios with delay and delay results in the fact that oil has to be stored longer and thus leads to higher inventory costs.

The average absolute deviation is low, meaning that for every scenario the deterministic model would yield a solution close to 2,405 USD and therefore solution robustness is high. For both models the percentage of infeasible scenarios is 96%, meaning that model robustness is low, since infeasibilities are not penalized in both models.

Influence of ω and λ on total costs

Increasing the value of λ decreases the absolute deviation and increases total costs for the stable demand case, which is expected by the definition of λ. However, in the experiments with $\lambda = 1$ and $\omega \geq 10$ a solution is found with lower total costs as in the same experiments with $\lambda = 0$. The influence of the value of λ on the total costs for most experiments, as can be seen in Table 4, is only small, since λ is multiplied with the average absolute deviation of the control variable costs of a scenario from the sample average. These control variable costs include the inventory costs, and safety stock penalties, which are low compared to the energy costs. The average absolute deviation is only significant for $\omega = 1,000$ and $\lambda = 0$, since this solution includes terminal shortage costs. Increasing λ decreases the absolute deviation again, where the total costs are remarkably also decreased. The total costs decrease, since increasing λ decreases the effect of ω.

Increasing ω has more influence on the total costs. Figure 12 shows the total costs for different values of ω, here we took the averages of all values of λ.

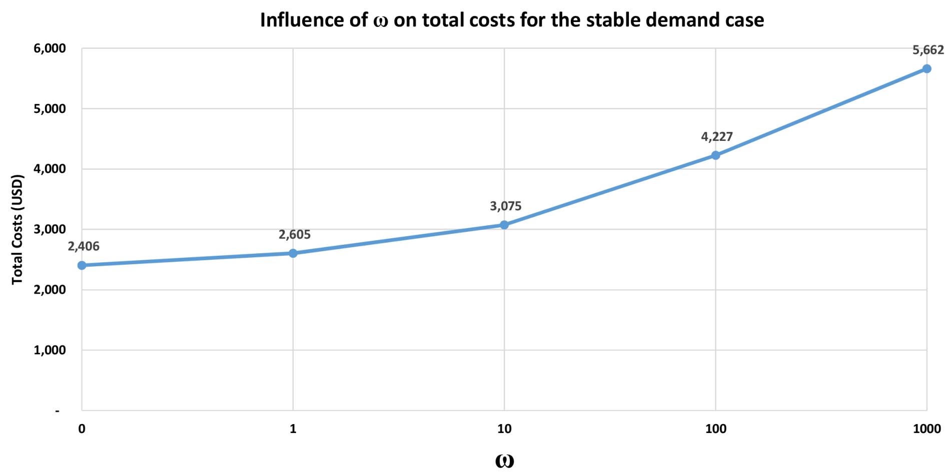

Figure 12: Influence of ω on total costs for the stable demand case

The figure shows that increasing ω increases total cost, which is primarily caused by an increase in energy costs. In the robust solution the terminal storage tank levels are kept below the maximum inventory levels by deviating more in the throughput of the pipelines, resulting in extra costs. Figure 13 shows an example of the difference in throughput between the deterministic solution and a robust solution. The example is from the experiment with $\lambda = 0$ and $\omega = 100$. The figure shows the throughput of one of the three pipelines between Node 1 and Node 2 (Node 2 is connected to TerminalTank1 (see Appendix A), which causes the fluctuations in the robust solution). The red line denotes the maximum capacity of the pipeline and the black dashed line the peak efficiency throughput (recall that the peak efficiency throughput is set at 95% of the "regular throughput"). In the optimal solution the throughput of this pipeline is constant over time, which is the cheapest way to transport a certain volume through the pipeline and exactly how we wanted to energy costs to behave when the throughput is more than the maximum efficiency throughput (see Section 4.5.1). The robust solution has higher costs because the throughput is much higher than the maximum efficiency throughput in certain time periods, which also causes the separation facility storage tanks inventory levels to fluctuate more than in the deterministic solution of Figure 10. We now examine if these extra costs result in improvements in model robustness.

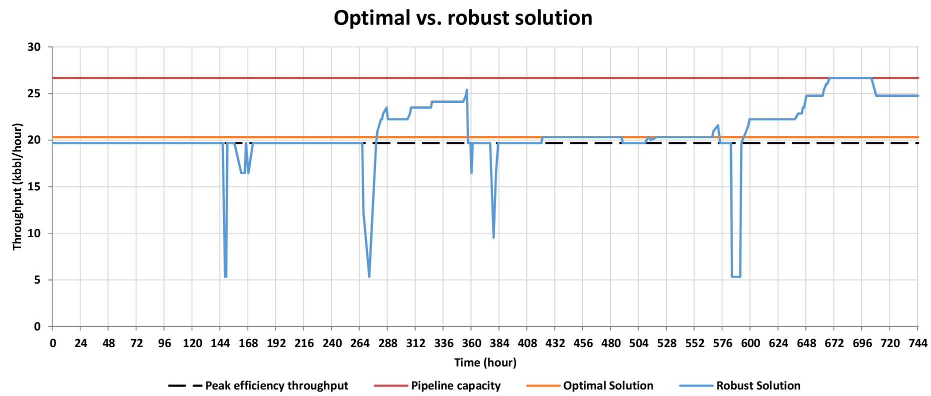

Figure 13: Difference between the optimal and a robust solution

Influence of ω on model robustness

The deterministic model and the robust model with $\omega = 0$ both had an infeasibility percentage of 96%. Figure 14 shows the infeasibility percentage and the average violation in kbbl for different values of ω. The average violation in kbbl is the average infeasibility volume per infeasible scenario (so not the average of all scenarios), which is calculated by correcting the "Infeasibilities" column of Table 4 for the number of infeasible scenarios.

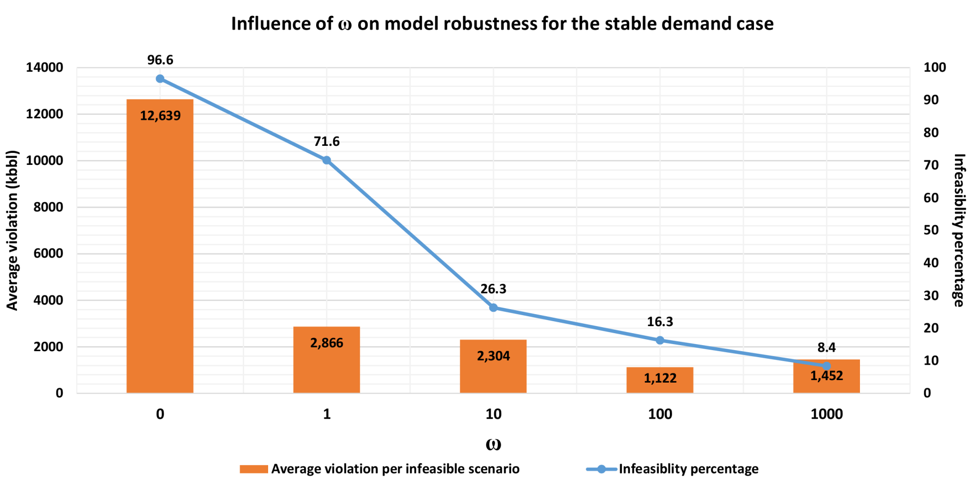

Figure 14: Influence of ω on model robustness for the stable demand case

For the deterministic model the infeasibility percentage is equal to the robust model with $\omega = 0$, but the average violation is more than three times as high (43,533 kbbl) as for $\omega = 0$. The figure shows that increasing ω increases model robustness as both the number of infeasible scenarios as the violation per infeasible scenario are reduced. However, it is remarkable that the average violation per infeasible scenario increases as ω is increased from 100 to 1,000. The number of infeasible scenarios, however, is almost halved as ω is increased from 100 to 1,000. The figure proves that our robust model indeed yields solutions that are, by design, less sensitive to uncertainty in tanker arrivals. It is up to OilCO to decide if this increase in robustness is worth the extra costs.

Influence of λ on solution robustness

We already concluded that increasing λ has a low influence of total costs, since the costs represented by the control variables are low compared to the energy cost. The absolute deviation is low for most experiments, meaning that solution robustness is high. However, the average absolute deviation is significant for $\omega = 1,000$ and $\lambda = 0$, since this solution includes terminal shortage costs. Increasing λ increases the solution robustness, as the absolute deviation is lowered and therefore the solutions for different realizations are closer to optimal. We can therefore conclude that improving the solution robustness only has a very small effect since it is already high. Furthermore, solution robustness only is low when terminal shortage costs are included in the solution, in that case solution robustness can be improved by increasing λ.

Influence of λ and ω on CPU time

The CPU time for the different models, which differs per combination of λ and ω as Figure 15 shows. We conclude that the CPU time increases when ω is increased from 0 to 1 for most values of λ. The CPU time slightly decreases when ω is further increased to 1,000 for most values of λ. The CPU time decreases as λ is increased, however it is on average 67% less when $\lambda = 0$, compared to the other values. The best value for λ in for the stable demand case is zero considering this performance difference and the fact solution robustness is already high. All running times are more than 50 times the running time of the deterministic model (4.5 seconds) and are on average 13.5 minutes, which is high compared to our goal of a couple of minutes. We try to improve the CPU time for this case by solving it with the simplified model in Section 6.4.

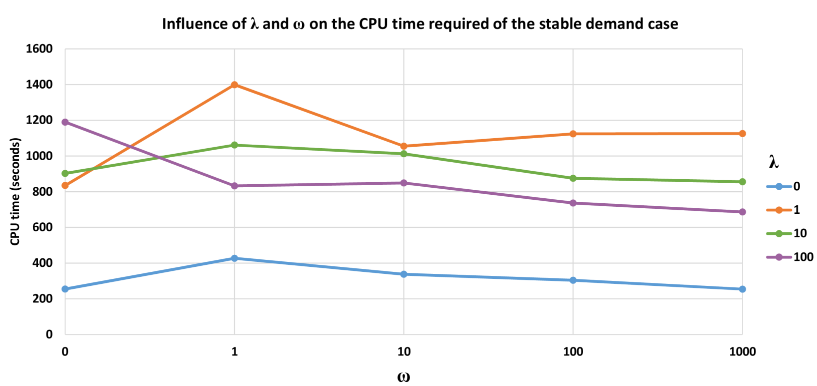

Figure 15: Influence of λ and ω on the CPU time for the stable demand case

6.3.3 Experimental results for the unstable demand case

In this section we examine the experimental results for the unstable demand case in the same manner as the stable demand case. We start with the results of solving the model with $\lambda = \omega = 0$, which means that both solution robustness and model robustness are not considered as both terms are multiplied with zero. After that we examine what the different solutions are as λ and/or ω increases, here we first look at total costs, then at model robustness, third at solution robustness, and finally at CPU time. Table 5 shows the results of the experiment. The division of the total costs over the different cost components can be found in Table 10 of Appendix I.

Model	λ	ω	Objective value (USD)	Total costs (USD)	CPU Time (s)	Infeasibilities (kbbl)	% infeasible	Abs Dev (USD)
Determin.	N/A	N/A	26,538	26,538	12	73,393	96	52.1
Robust	0	0	25,216	25,216	742	47,777	100	225.4
Robust	1	0	25,262	25,257	2201	48,387	100	5.2
Robust	10	0	25,269	25,268	1207	48,871	100	0.1
Robust	100	0	25,269	25,269	2519	48,910	100	0
Robust	0	1	28,170	26,037	900	21,323	100	403.8
Robust	1	1	28,267	26,097	3007	21,645	100	6.0
Robust	10	1	28,274	26,107	1683	21,652	100	0.2
Robust	100	1	28,275	26,109	1755	21,660	100	0
Robust	0	10	42,866	28,262	903	14,604	95	1395.6
Robust	1	10	43,703	27,721	1577	15,556	95	425.9
Robust	10	10	44,091	28,123	1353	15,965	95	0.3
Robust	100	10	44,093	28,128	1544	15,965	95	0
Robust	0	100	116,111	56,557	907	5,955	89	27390.6
Robust	1	100	137,231	46,467	1899	7,457	93	16197.9
Robust	10	100	155,624	49,220	1463	10,640	93	0.4
Robust	100	100	155,627	49,226	1716	10,640	93	0
Robust	0	1,000	333,416	202,862	287	1,306	44	116492.4
Robust	1	1,000	440,114	175,257	1428	1,657	44	99206.7
Robust	10	1,000	589,011	301,894	2981	2,869	61	22.0
Robust	100	1,000	589,210	301,634	3204	2,876	61	0

Table 5: Computational results for the unstable demand case

Deterministic model vs robust model with $\lambda = \omega = 0$

Table 5 shows that the solution value for the robust model with $\lambda = \omega = 0$ is lower than the solution value of the deterministic model, which is different from the results of the case with stable demand. The deterministic model only considers 1 scenario in which all tankers arrive on the first day of their 3-day window, where the robust model considers 80 scenarios with different realizations of delay. As delay causes the peak in demand to be flattened and to be spread over more days, the robust model yields a lower solution value, because the throughput rates are less high in the period with the peak in demand and some of the terminal shortage penalties are avoided. For both models the percentage of infeasible scenarios is close to 100%, since infeasibilities are not penalized in both models.

Influence of ω and λ on total costs Although the effect is small in the experiment with the stable demand case, the value of λ has a larger effect on the total costs in the unstable demand case. Increasing the value of λ increases the objective value significantly when $\omega = 100$ and $\omega = 1,000$, however in the experiments with $\lambda = 1$ and $\omega \geq 10$ a solution is found with lower total costs as in the experiments with $\lambda = 0$, just as in the stable demand case.

Increasing ω also has an effect on the total costs. The total costs are especially high when $\omega = 100$ or higher, since the model chooses to send less oil to the terminals to prevent infeasibilities, although that causes shortage penalties. In case $\omega \geq 100$, less penalties for deviating from the production plan occur, since the production rate at Reservoir9 is increased by less than in the other solutions, so less oil is pumped into the network and therefore infeasibilities are avoided. The solution does not include a changeover for the case that $\omega = 1,000$ and $\lambda = 0$ or $\lambda = 1$, since not increasing production prevents infeasibilities against high terminal shortage penalties. The increase in total cost is further explained by the deviations in throughput rate, which we already saw in Figure 13.

Influence of ω on model robustness

Figure 16 shows the infeasibility percentage and the average violation per infeasible scenario for different values of ω. Increasing the value of ω increases model robustness, since both the infeasibility percentage and the average violation per infeasible scenario decrease. Both are, however, higher as in the stable case (see Figure 14), because less infeasibilities are avoided as there is less flexibility in the period with the peak in demand in changing throughput rates without causing shortages at the terminals. The average violation per infeasible scenario decreases for every increase of ω, while the infeasibility percentage only is below 100% for $\omega > 1$.

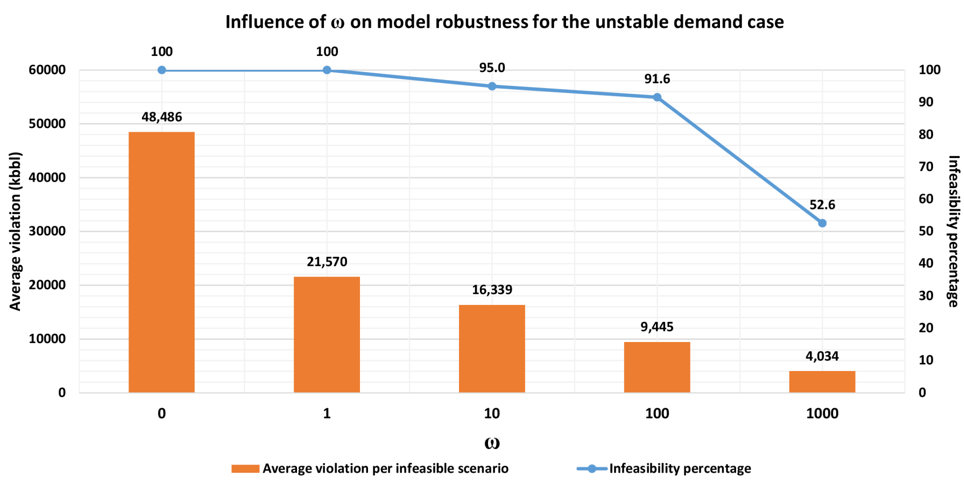

Figure 16: Influence of ω on model robustness for the unstable demand case

Influence of λ on solution robustness
The solution robustness is lower for experiments with $\lambda = 0$ than in the stable demand case, because the terminal shortage penalties increase, which are calculated per scenario and create differences in total costs per scenario. Increasing the value of λ does again decrease solution robustness as the absolute deviation decreases. The absolute deviation can be decreased to 0 if λ is set to 1,000, just as in the stable demand case.

Influence of λ and ω on CPU time
Figure 17 shows the CPU times for the different values of ω and λ. The CPU times to solve the unstable demand case are larger than for the stable demand case (see Figure 15), since a changeover is required in the solution and the model has to decide when and at which reservoir that changeover occurs. The figure shows that the CPU time is lower for $\lambda = 0$, because the risk term is excluded in those experiments. The relation between ω and the CPU time is the same as in the stable demand case when ω is between 0 and 100, but different for $\omega = 1000$ as the CPU times significantly increase for $\lambda \geq 10$. As the average CPU time is 28 minutes, which is high compared to our goal of a couple of minutes, we try to improve the CPU time of this case by solving it with the simplified model in Section 6.4.

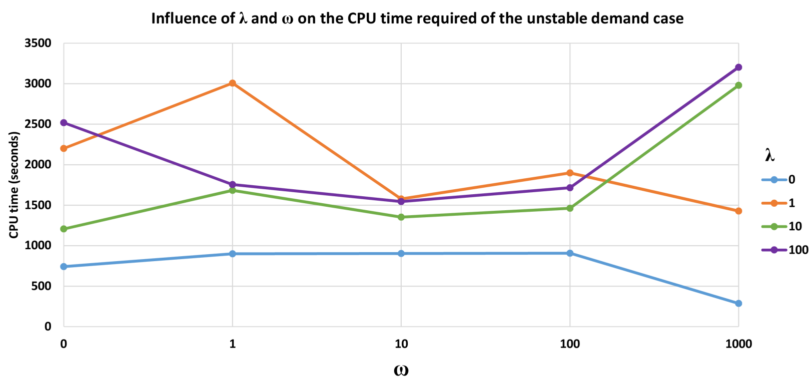

Figure 17: Influence of λ and ω on the CPU time for the stable demand case

6.4 Simplification results

In Section 4.7 we propose a few simplifications for the model, which reduce the number of integer variables in the model to boost performance. In this section we test if these simplifications improve the CPU time to solve the robust model. We examine if the simplifications yield the same solution as the robust model and where possible differences are. For this test we take the results of the experiments of the robust model from the experiment of Section 6.3 and compare these with the same experiments solved by the simplified model. For the simplified model we use the following simplifications, based on the analysis of the two cases:

1. Changeovers can only occur at the beginning of the day.
2. Changeovers can only occur at Reservoir9 (closest to Node 1) and Reservoir3 (farthest away from Node 1).

Demand case	Stable	Unstable	Average
CPU time without simplifications(s)	806	1,664	1,235
CPU time with only simplification 1(s)	791	736	764
CPU time with only simplification 2(s)	365	400	383
CPU time with simplification 3&4(s)	713	729	721
CPU time with all simplifications(s)	330	351	341
All simplifications - no simplifications(s)	-476	-1,313	-895

Table 6: CPU time difference between robust and simplified model

3. The minimum production rate is increased to 0.9 (was 0.4) times the production plan rate.
4. For the stable demand case the maximum production rate is decreased to 1.1 times the production plan rate, while for the unstable demand case this value is kept at 1.2.

Simplification 1 decreases the number of integer variables significantly, however it can cause differences in the solution. Simplification 2 further decreases the number of integer variables, but will not change the solution, as we know that changeovers occur at Reservoir9 (when an increase of production rate is required) or Reservoir3 (when a decrease of production rate is required). Since demand is stable in the first case, we only expect small changeovers (if any occur) so the difference between minimum and maximum production rate may be decreased as in simplifications 3 and 4 to decrease the size of the M parameter that is used in the constraints (Constraints (25) and (26)) that bound the changeover variable. In the unstable demand case we only increase the size of the minimum production rate, since the production rate of Reservoir9 is increased to its maximum in the previous experiments. Next to experiments with all simplifications, we also perform experiments with one of the simplifications to see which simplification has the largest impact. We perform an experiment with simplifications 3 and 4 together, since both simplifications decrease the size of the M parameter.

Table 6 shows the results of the experiments per case. The simplifications improves the CPU time by 69%. The largest improvement is yielded at unstable demand case. Simplification 2 has the largest impact on the CPU time and Simplification 1 has the smallest impact. Since the changeovers at the unstable demand case occur at the beginning of the day, Simplification 1 has no effect on the solution and, therefore, the simplified model yields the same solutions as the robust model for all experiments.

6.5 Conclusion

In this chapter we compared the different models we proposed in this thesis. In Section 6.2 we compared two different tanker arrival cases for the deterministic model. The conclusion is that the tanker arrival scenario has almost no impact on the energy costs of the stable demand case, as it only slightly increases inventory costs and the safety stock penalties. In the solution of the stable demand case the throughput of the pipelines is constant, which is in line with the energy costs behavior we chose in Section 4.5.1. In the solution of the unstable demand case a changeover occurs in order to deal with the peak in demand. Moreover, the throughput is increased to transport the extra produced oil to the terminals and terminal shortages occur.

In Section 6.3 we compared the deterministic model with the robust model. When comparing the deterministic model with the robust model with λ =0 and ω = 0 for the stable demand case it seems that including more tanker arrival scenarios is of little impact, as the deterministic model with one scenario yields almost the same solution as the robust model with 80 scenarios. However, when making the same comparison for the unstable demand case there is a much larger difference, since the delay of the tankers causes the peak in demand to flatten and therefore the energy costs to decrease. Solving the robust model with λ =0 and $\omega = 0$ is therefore preferred over solving the deterministic model.

Increasing ω increases the total costs but decreases the number of infeasible scenarios and the average infeasibility (in kbbl) per infeasible scenario. For the stable demand case the highest value of ω we tested yielded an infeasibility percentage of 8.4% (which is 96.6% when model robustness is not taken into account), while the total costs where increased by 135%. For the unstable demand case the infeasibility percentage could only be decreased to 52.6%, since the peak in demand gives less flexibility in avoiding infeasibilities. Deciding which value of ω is the best for OilCO is difficult, since we do not know if we modelled the energy costs correctly (see Section 4.6.1) and we do not know if the robust solution is totally realizable. In the robust solution the throughput rates are changed every few hours in certain periods (see Figure 13), but the question is if this is possible in practice. We do not have enough information on this subject to judge if the solution is realizable. However, we can conclude that the robust model yields solutions that are more capable of dealing with the uncertainty in tanker arrivals, meaning that taking model robustness into account by setting $\omega > 0$ is preferable.

The average absolute deviation of the objective value for different scenarios from the sample average is low for most experiments. This low absolute deviation causes that the value of λ has little influence on the solution. In solutions with terminal shortage costs the solution robustness is low, as terminal shortage costs are part of the control variable costs which vary per realization of the tanker arrival moments. Here, increasing λ decreases the absolute deviation and even decreases total costs when ω is high, but decreases model robustness. We propose to set $\lambda = 0$ in case no terminal shortage costs are expected, because this has a positive effect on model performance and almost no effect on model solution. When there are terminal shortage costs involved increasing λ increases solution robustness and even decreases total costs when ω is high. Increasing λ for high values for ω decreases the effect of ω as total cost and model robustness both decrease, but the question is if decreasing the value of ω instead of increasing the value for λ in those cases yields a "better" solution.

For the considered cases we decreased the number of integer variables and the size of parameter M in order to boost model performance, resulting in an average CPU time improvement of 69%. Therefore, we suggest to use this simplification. The average CPU time to solve the model then is 330 seconds for the stable demand case and 351 seconds for the non stable demand case, which are both in line with our goal of a couple of minutes.

In Chapter 7 we perform sensitivity analysis on input parameters that are uncertain to see how they influence the model in terms of solution, objective value, total costs, and CPU time.

7 Sensitivity analysis

In this chapter we perform a sensitivity analysis on the uncertain parameters in our model to find out how sensitive our model is to changes in these parameters. Moreover, we want to find out if our model is able to deal with all these different input values. In Section 2.3.2 we addressed the uncertainty in parameters for the first time. In this section we identified the storage tank size (also called the maximum inventory level) as one of the parameters on which we perform sensitivity analysis. In Section 6.1.3 we identified the peak efficiency throughput of the pipelines and the different cost factors of the objective function as uncertain, implying that we analyze these as well. First, Section 7.1 describes the analysis set-up. Subsequently, Section 7.2 considers the storage tank size analysis, Section 7.3 considers the peak efficiency throughput analysis, and Section 7.4 considers the cost factors analysis.

7.1 Analysis set-up

In this chapter we change one factor at a time in our analysis, since there are a large number of uncertain parameters and we need to keep the number of experiments in an acceptable range. We therefore choose a number of values for the uncertain parameters for which the model is solved and keep all other parameters constant. In the previous chapter we concluded that the value of λ has low influence on the solution in most tests, while setting λ equal to 0 boosts performance. Therefore, we set it equal to 0 in the analysis. We analyze both the effect of the changes in parameter values on the optimal solution ($\omega = 0$) and a robust solution ($\omega = 100$) for both cases to see if there is a difference in sensitivity. We treat the stable demand case and the unstable demand case separately in the analysis. We use the simplified model for most of this analysis, since we concluded that the simplified model outperforms the robust model (see Section 6.5). However, we use the model without simplification for the changeover cost analysis, as we want to include all possible changeover moments. After solving the model we analyze the difference in solution values, where we look at the difference in total costs and the different cost components. Furthermore, we look if the obtained schedule changes and what the robustness of the obtained schedule is. Finally, we look at the effect of the parameters values on the CPU time to solve the model.

7.2 Storage tank size analysis

We perform a sensitivity analysis on the storage tank size, since it determines the amount of variability in demand and operations that the network can handle and since we have no information on the size of the tanks. In Section 6.1 we set the storage tank size of the terminal tanks on 3,000,000 bbl and the size of storage tanks at the separation facilities on 1,500,000 bbl. We based this value on what we think is a logical tank size given OilCO's production and demand volumes, but we do not know the exact value. Figure 10 shows that the inventory levels at a terminal storage tank vary more than the inventory levels at a separation facility storage tank. We therefore perform most experiments on the terminal storage tanks. For the terminal storage tanks we perform experiments for the tank sizes: 1,000,000 to 5,000,000 bbl, with steps of 500,000 bbl. For the separation facility storage tanks we examine the values 750,000 to 2,000,000 bbl with steps of 250,000 bbl. The safety stock levels and the initial inventory are kept constant on the same absolute level as in Section 6.1. In Section 7.2.1 we analyze the separation facility storage tanks and in Section 7.2.2 we analyze the terminal storage tanks.

The assumption that every tank of a certain type (terminal storage tank or separation facility storage tank) has the same storage capacity may not be correct. It may also be possible that OilCO based its tank sizes on the volume of oil that passes through the tank on an average day. Table 7 shows these volumes.

Storage tank	Average daily volume in kbbl	Location
SepFacTank1	1,344	Separation facility
SepFacTank2	980.4	Separation facility
SepFacTank3	166.8	Separation facility
SepFacTank4	558	Separation facility
TerminalTank1	1,350	Terminal
TerminalTank2	900	Terminal

Table 7: Average daily volumes of the different storage tanks

In the table the largest daily volume of a separation facility storage tank (SepFacTank1) is more than eight times larger than the smallest one (SepFacTank3). We perform experiments for setting the tank sizes equal to 1 till 4 times the volumes of Table 7, with steps of 0.5. In contrary to the other storage tank experiments, we vary both the separation tank volumes and the terminal volumes at the same time. The safety stock levels are $\frac{1}{6}$ times the storage tank size and the initial inventories are $\frac{1}{3}$ times the storage tank size, just as in Chapter 6. In Appendix J we analyze the storage tank size based on the average daily volume, since this analysis is similar to the analysis of the storage tank size based on tank type.

7.2.1 Separation facility storage tank analysis

We solve the model for $\omega = 0$ and $\omega = 100$ for different separation facility storage tank sizes: 750, 1,000, 1,250, 1,500, 1,750, and 2,000 kbbl. In this section we treat some remarkable results of the experiment. Figure 18 shows an overview of the total costs, infeasibility percentages, average violation per infeasible scenario, and CPU time for the different separation facility storage tank sizes. The total costs are more or less stable for all tests as the separation facility storage tank size increases, however there is a small increase in energy costs for the experiments with $\omega = 100$ and a tank size of 750 kbbl. The throughput rates have to fluctuate more in those experiments, as a size of 750 kbbl is too small for the South East Tank. The difference in costs between ω =0 and $\omega = 100$ are caused by an increase in energy costs and the difference in costs between the stable and unstable demand case are caused by the peak in demand of the unstable demand case, which we both already saw in Section 6.3.2.

The infeasibility percentages are almost constant on 98% for the unstable demand case and on 100% for the stable demand case with $\omega = 0$, but for both cases with $\omega = 100$ there is a difference for a tank size of 750 kbbl. Increasing ω lowers the average infeasibility percentage to 16% for the stable case and to 89% for the unstable case. The average violation per infeasible scenario are more or less constant for both $\omega = 0$ and $\omega = 100$ for both cases, with a small increase at 750 kbbl for the stable case.

Furthermore, the CPU time is also more or less constant for all experiments. We conclude that the separation facility storage tank size only has an influence on the model if it is smaller than 1,000 kbbl.

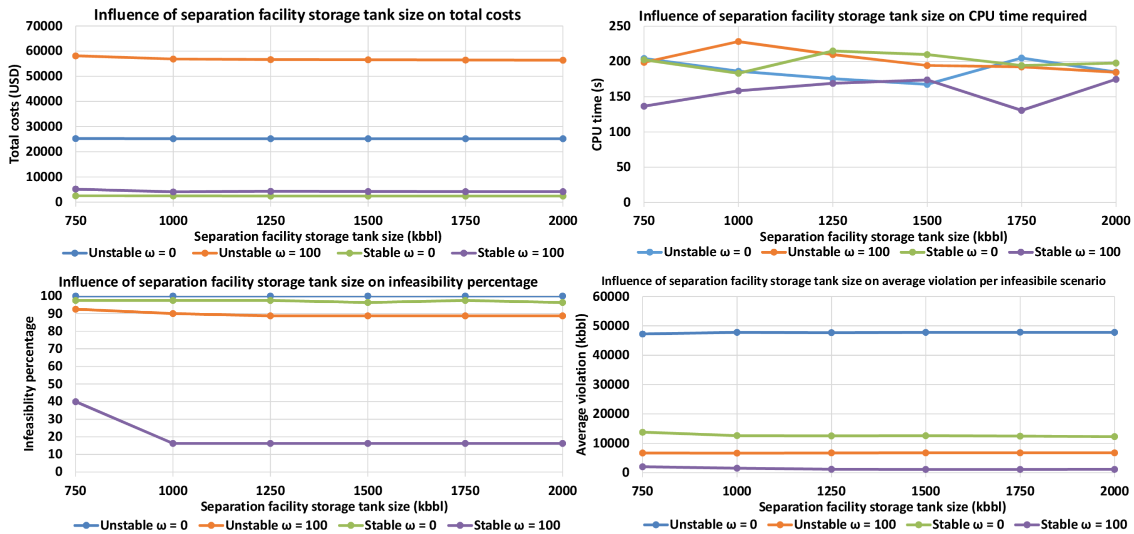

Figure 18: Separation facility storage tank experiment results

7.2.2 Terminal storage tank analysis

We solve the model for both cases with $\omega = 0$ and $\omega = 100$ and with different terminal storage tank sizes: 1,000, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 4,500 and 5,000 kbbl. In this section we look at the same indicators as in Section 7.2.1. Figure 19 shows the total costs, infeasibility percentages, average violation per infeasible scenario and CPU time for the different terminal storage tank sizes. The figure of the CPU time looks similar to the CPU time in Figure 18, where for most experiments the CPU time is constant. However, for the stable case with $\omega = 100$ the CPU time is higher for low values of the terminal storage tank size, since it is harder to find a robust solution.

The figures of the total costs, average violation per infeasible scenario, and infeasibility percentages differ a lot from Figure 18. The total costs are constant for $\omega = 0$, just as in Figure 18, but for $\omega = 100$ the total costs are very high when the terminal storage tank size is small and decrease as the terminal storage tank size increases. The difference in total costs are primarily caused by a difference in energy costs and terminal shortage costs. When the terminal storage tank size is low, the throughput rates of the pipelines that supply the terminal storage tanks have to fluctuate more in order to yield a robust solution. When the terminal storage tank size increases the total costs converge to the total costs for $\omega = 0$ (the optimal solution). It is remarkable that the experiments with the unstable case and $\omega = 100$ do not include the changeover when the terminal tank size is at most 2,000 kbbl, since the terminal tanks cannot handle the extra oil. Since there is no extra oil production in these experiments, the terminal shortage costs are high. For the same reasons the changeover is smaller than in the original solution (where the tank size was 3,000 kbbl) for tank sizes lower than 3,000 kbbl.

The infeasibility percentage is 100% for tank sizes smaller than or equal to 2,500 kbbl for the stable case with $\omega = 0$ and for tank sizes smaller than 3,500 kbbl for the unstable case with $\omega = 0$. When the tank size increases the infeasibility percentage decreases to 0 as the tank becomes large enough to deal with all delay scenarios. For $\omega = 100$ the infeasibility percentage decreases when the

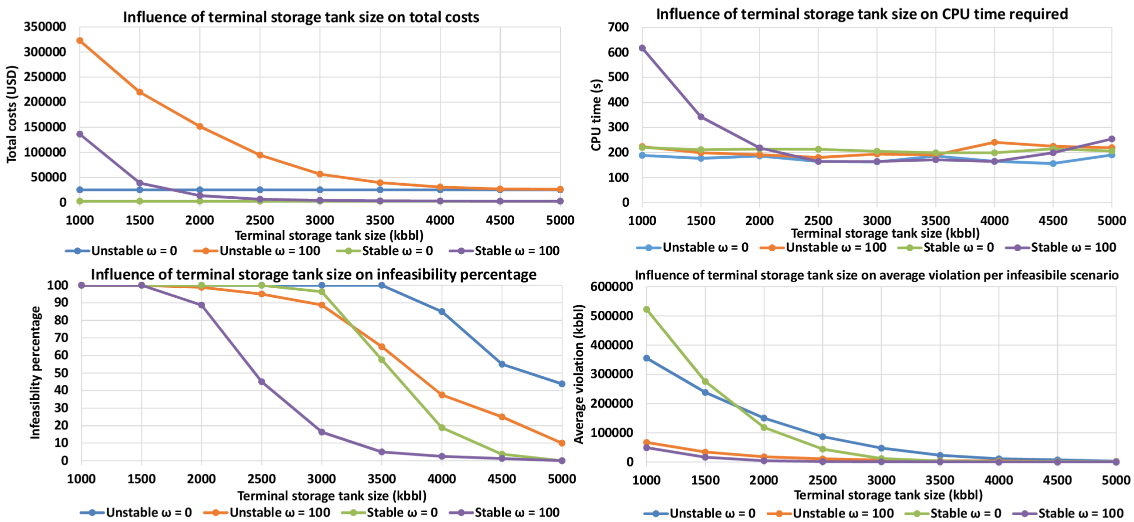

Figure 19: Terminal storage tank experiment results

tank size is at least 1,500 kbbl for both cases. There are no infeasible scenarios when the tank size is 5,000 kbbl for the stable case for both $\omega = 0$ and $\omega = 100$, while there are always infeasible scenarios in the experiments of the unstable case.

For all experiments the average violation per infeasible scenario decreases when the terminal tank size increases, since the violations are the volumes that the inventory level is above the maximum inventory level of the terminal storage tanks.

If we compare the results of this analysis with the results of 7.2.1, we can conclude that the influence of the terminal storage tank size on total costs and model robustness is larger than the separation storage tank size.

7.3 Peak efficiency throughput analysis

We analyze the impact of the peak efficiency throughput, since it is an important concept in the models, since we based our energy costs definition on it, and since we do not have any information about the exact value of it. Recall that the peak efficiency throughput is the throughput rate at which the pumps of the pipelines operate at maximum efficiency. In Section 6.1 we set the peak efficiency throughput of a pipelines equal to 95% of the regular throughput. This regular throughput is based on the average throughput in a network with perfectly constant demand.

It may be that this assumption is not correct and OilCO's regular throughput is equal to the peak efficiency throughput. Furthermore, it may be that OilCO designed its network to be prepared for an increase in demand, so that the peak efficiency throughput lies above the regular throughput. In this analysis we investigate values for a peak efficiency throughput of 90%, 95%, 100%, 105%, and 110% of the regular throughput.

However, it can also be possible that the peak efficiency throughput does not depend on the regular throughput at all, but that it is a fixed percentage of maximum pipeline throughput. We try values for a peak efficiency throughput of 40%, 50%, 60%, 70%, and 80% of the pipeline capacity,

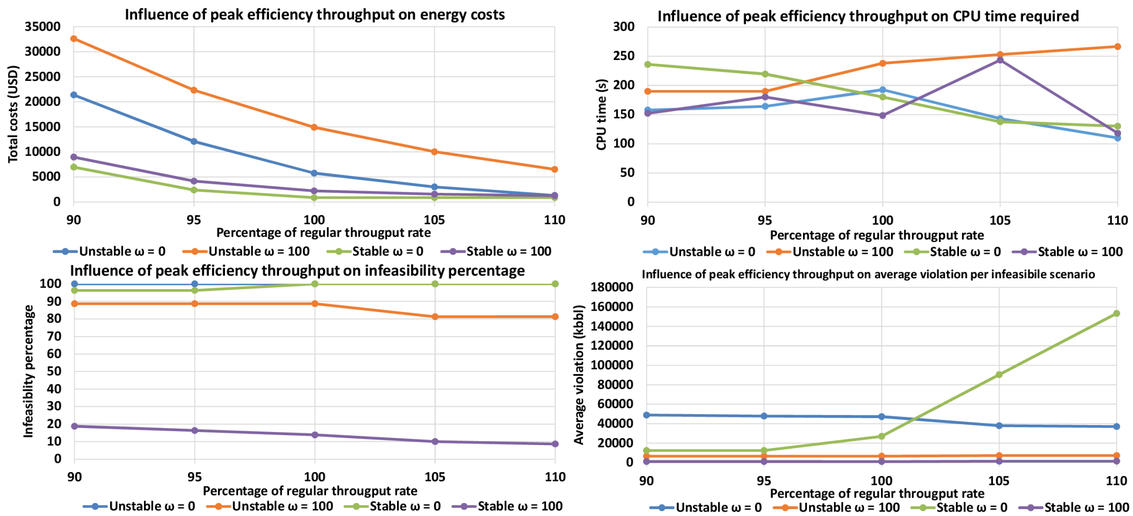

Figure 20: Peak efficient flow as a percentage of the regular throughput results

where the average pipeline utilization ($\frac{average\,throughput}{pipeline\,capacity}$) was 70.2% in the experiments of Section 6.2.

Per experiment we examine what the effect of the peak efficiency throughput is on energy costs, model robustness, and CPU time. Furthermore, we examine if changing peak efficiency throughput yields a different schedule, or that it yields the same schedule with a different objective value.

Section 7.3.1 gives the results of the experiment where the peak efficiency throughput is a percentage of the regular throughput. Appendix K gives the results of the experiment where the peak efficiency throughput is a percentage of the pipeline capacity, since the results are similar as in the experiment where the peak efficiency throughput is a percentage of the regular throughput.

7.3.1 Peak efficiency throughput as a percentage of the regular throughput

We solve the model for both cases, for $\omega = 0$ and $\omega = 100$, and for the mentioned values of the peak efficiency throughput. Figure 20 shows the energy costs, CPU time, infeasibility percentage, and average violation per infeasible scenario.

For all case the energy costs decrease as the peak efficiency throughput increases to 100% of the regular throughput, as this means that on average the throughput is above the peak efficiency throughput. As the peak efficiency throughput increases above 100% of the regular throughput the costs are stable for $\omega = 0$ for the stable demand case, but not for $\omega = 100$ and the unstable demand case as the throughput sometimes has to be above the peak efficiency throughput in order to deal with tanker arrival uncertainty. Moreover, in case the peak efficiency throughput increases above 100% of the regular throughput the CPU time to solve the model decreases for $\omega = 0$, which might be caused by the fact that the non-linear part of the energy costs function is approximated in less time periods. The infeasibility percentage slightly decreases for $\omega = 100$ for both cases if the percentage of the regular throughput increases, as it is cheaper to vary in throughput rate below the peak efficiency throughput than above the peak efficiency throughput. This allows the model to

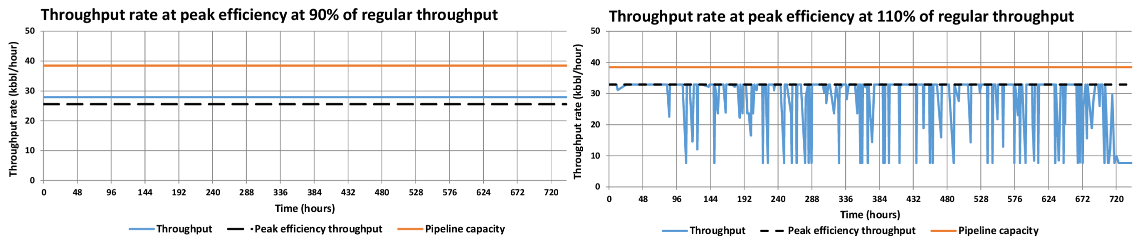

Figure 21: Difference in throughput in the optimal solution for a peak efficiency throughput below and above the regular throughput

avoid more infeasibilities without exceeding the peak efficiency throughput. The average violation per infeasible scenario remains constant if $\omega = 100$ for both cases. The average violation per infeasible scenario when ω =0 increases when the percentage of regular throughput increases for the stable demand case, while the opposite happens for the unstable demand case. To explain this difference in behavior, we look at the corresponding solutions. Figure 21 shows the throughput rate of one of the pipelines between Node 1 and Node 2 in the optimal solution of the stable demand case.

The throughput rate at peak efficiency of 90% of the regular throughput is constant for the same reasons as in Section 6.3.2, where the peak efficiency throughput was 95% of the regular throughput. The throughput rate fluctuates when the peak efficiency throughput is at 110% of the regular throughput. These fluctuations have a negative effect on model robustness (see Figure 20), where in the unstable demand case fluctuations in throughput have a positive effect on model robustness, so the effect on model robustness depends on the timing and size of the fluctuations. Since fluctuations throughput rate below the peak efficiency throughput have a low influence on the energy costs, solutions exist that are close to optimal and are more robust. For example, when $\omega = 0.01$ for the stable demand case in the 110% experiment the average violation per infeasible scenario decreases from 153,338 kbbl to 2,797 kbbl, while the total costs increase with 0.45 USD. In this solution the throughput rate fluctuates just as in Figure 21, but on different moments with different sizes. The question is, again, to what extend a fluctuating throughput rate is realizable in practice. The peak efficiency throughput has a high influence on the model as it both influences total costs as the solution in terms of throughput rates. The effect on model robustness in case $\omega = 100$ is however small.

7.4 Cost factor analysis

We perform sensitivity analysis on the cost factors, since they determine how heavy the cost components are weighted in the objective function and their values are hard to determine. We analyze the cost factors one by one, since the ratio between the factors has a big influence the resulting schedule. Our objective function has seven cost components: changeover costs, energy costs, inventory costs, and penalties for shortages at the refineries, shortages at the terminals, deviating from the production plan, and having less in stock than the safety stock. In most experiments of the stable case in Chapter 6 only the energy costs, inventory costs and penalties for having less in stock then the safety stock were larger than zero. For the cost components that were equal to zero, we want to find out by how much their cost factor has to decrease in order to make that cost component larger than zero. We divide the cost factors of these cost components by 2, 5, 10, 25, and 50 and analyze how the solution changes. For comparison, we perform the same analysis to the unstable demand case.

For the cost components that are larger than zero, we saw that the energy costs are much larger than the inventory cost and the safety stock penalties. We check if the solution changes if the safety stock penalty factor and inventory cost factor increases. We multiply both factors with 2, 5, 10, 25, and 50 such that the safety stock penalty factor stays 100 times larger than the inventory cost factor. For all experiments we again compare a robust solution with the optimal solution. We examine the cost factors one by one (except for the inventory cost and penalties for having less in stock then the safety stock, these are combined). Section 7.4.1 examines the terminal shortage penalty factor, Section 7.4.2 examines the refinery shortage penalties factor, Section 7.4.3 covers the changeover cost factor, Section 7.4.4 considers the penalty factor for deviating from the production plan, and Section 7.4.5 examines the inventory cost factor.

7.4.1 Terminal shortage penalty factor analysis

For the terminal shortage cost factor we evaluate the values 20, 10, 4, 2, 0.8, and 0.4 USD/kbbl. The results can be found in Appendix L. For the stable demand case decreasing the shortage factor indeed causes terminal shortage costs. The solution includes terminal shortage costs when the cost factor is at most 0.4 USD/kbbl for $\omega = 0$, where this is at most 4 USD/kbbl for $\omega = 100$. The shortage volumes also increase progressively as the cost factor decreases, as more and more energy costs can be saved.

The unstable case already had shortage costs in the solution for a cost factor of 20 USD/kbbl. The shortage volumes also increase as the cost factor decrease, although the terminal shortage costs decrease when the cost factor is at most 2 USD/kbbl for both $\omega = 0$ and $\omega = 100$, since from that point the cost factor decreases faster than the shortage volumes increase. The solution has no changeover any more when the cost factor is equal to 2 USD/kbbl or less for $\omega = 0$ or 4 USD/kbbl or less for $\omega = 100$, since then preventing shortages by increasing production is more expensive than not increasing production and paying the shortage penalties. For both cases the model robustness increases for $\omega = 100$ as the cost factor decreases, since less oil is sent to the terminals to avoid infeasibilities.

7.4.2 Refinery shortage penalty factor analysis

For the refinery shortage cost factor we evaluate the values 40, 20, 8, 4, 1.6, and 0.8 USD/kbbl. The results can be found in Appendix M. For the stable case with $\omega = 0$ the solutions are the same for all experiments, while for $\omega = 100$ there are refinery shortage costs if the cost factor is 1.6 USD/kbbl or less. However, these shortages are very small, as they are at most 50 kbbl in total for the entire planning horizon.

The unstable case yields refinery shortage costs for $\omega = 0$ when the shortage cost factor is at most 1.6 USD/kbbl, however these shortages are just as small as in the stable demand case experiments with shortages. The solutions for the experiments of the unstable demand case with $\omega = 100$ include refinery shortage costs if the cost factor is 4 USD/kbbl or less, however these shortages are much larger than in the other experiments as the solutions for cost factors of 1.6 and 0.8 USD/kbbl do have changeovers in them.The changeovers are excluded to prevent infeasibilities at the terminals, resulting in high shortages at the refineries.

7.4.3 Changeover cost factor analysis

For the changeover cost factor we evaluate the values 10,000, 5,000, 2,000, 1,000, 400, and 200 USD. The results can be found in Appendix N. In the stable demand case the solutions are the same, which is caused by the stable production plan in combination with the high penalties for deviating from the production plan.

In the unstable demand case, the number of changeovers increases from one to two (the production rate changes back to the production plan rate) in case the changeover cost factor is at most 1,000 USD to decrease the penalties for deviating from the production plan. In case $\omega = 0$ there are even four changeover in the solution when the changeover cost factor is 200 USD, caused by a short production increase at a second reservoir to save energy costs during the peak in demand.

7.4.4 Production plan deviation penalty factor analysis

For the production plan deviation penalty factor we evaluate the values 1, 0.5, 0.2. 0.1, 0.04, and 0.02 USD/kbbl. The results can be found in Appendix O. For the stable demand case all solutions are the same, since the changeover costs are to high to change the production rate.

For the unstable demand case the production plan deviation penalties decrease as the cost factor decrease, but for $\omega = 0$ this decrease is proportional to the decrease in cost factor, so the solution is the same. For $\omega = 100$ the solution changes, as the production rate at Reservoir9 increases to its maximum for a cost factor of 0.2 USD/kbbl or lower.

7.4.5 Inventory cost factor analysis

The inventory costs (combined with the penalties for having less in stock than the safety stock) are the only cost factor in this analysis that were already non-zero in the experiments of Section 6. For the inventory cost factor we evaluate the values 0.0001, 0.0002, 0.0005, 0.001, 0.0025, and 0.005 USD/kbbl per day. The results can be found in Appendix P. For the stable case the increase in cost factor results in an increase in total costs. The inventory costs increase proportionally to the increase in cost factor, which means the total inventory over the 31 day period stays the same, which is explained by the fact that the oil enters and leaves the network on exactly the same moments. The penalties for having less in stock than the safety stock increase also, but less than proportional to the cost factor increase. The total volume that is less in stock than the safety stock during the planning horizon therefore decreases as the inventory cost factor increases. The energy costs also increase, which means that the throughput rate has to deviate more above the peak efficiency throughput in order to prevent that there is less in stock than the safety stock on certain moments. These deviations also cause the model robustness to slightly decrease.

For the unstable case we see exactly the same effect. So for both cases we conclude that not only the total costs change when the inventory cost factor is changed, but also the solution in terms of throughput rates changes.

7.5 Conclusion

In this Chapter we performed a sensitivity analysis on the uncertain parameters of our model to examine what the effects of changes in these parameters are on our model. We performed sensitivity analysis on the storage tank sizes, the peak efficiency throughput, and the cost factors of the objective function.

In Section 7.2.1 we saw that the total costs and model robustness is more or less constant as the separation facility storage tank changes in size. The total costs increase and robustness decreases only when the tank size is smaller than or equal to 1,000 kbbl. The CPU time is insensitive for different values of the storage tank size. In Section 7.2.2 we saw that terminal storage tank size has a large influence on both model robustness and total costs when $\omega = 100$. Total costs decrease and model robustness increases as the terminal tank size increases. The infeasibility percentage and average violation per infeasible scenario eventually get zero if the terminal storage tank is large enough. From the analysis we conclude that the model is more sensitive to changes in the terminal storage tank size than the separation facility tank size.

The peak efficiency throughput value has a large effect on the model, as it both has a large influence on total costs and the throughput rates of the pipelines. The most important factor that determines this influence is if the peak efficiency throughput is below or above the average throughput needed to fulfill demand. In Section 7.3.1 we performed experiments where the peak efficiency throughput was a percentage of this average throughput. The experiments show that the optimal solution is to have constant throughput when the peak efficiency throughput is below the average throughput and fluctuates when the peak efficiency throughput is above average throughput. In the latter case there are many solutions that are close to optimal, which is caused by the fact that the energy costs function is linear when the throughput rate is below the peak efficiency throughput.

The cost factor analysis showed that the model the least sensitive for changing the production plan deviation penalty factor, as the solution only changes if demand is unstable, $\omega = 100$, and the penalty factor is divided by ten or more. Changing the refinery shortage penalty factor only yields refinery shortages if the factor is decreased enough, $\omega = 100$, and/or demand is unstable, although shortages are small in most experiments. Terminal shortage penalties increase as the terminal shortage penalty factor is increased, as less oil is pumped to the terminal to decrease energy costs and/or infeasibilities. Increasing both the inventory costs factor as the penalty factor for having less in stock then the safety stock yields solutions that have higher energy costs and lower robustness as throughput rates have to deviate more in order to keep the inventory level above the safety stock. For the stable demand case changing the changeover cost factor has no effect on the solution, while it increases the number of changeovers for the unstable demand case.

The model is therefore very sensitive to changes in storage tank size and peak efficiency throughput, while the sensitivity differs per cost factor. If we would have chosen an other value for the sensitive parameters in Section 6.1 then most experiments in Chapter 6 would yields different solutions. However, the sensitivity analysis gives no proof that the chosen method for our problem is insufficient, as our model comes up with in an optimal solutions for all experiments.

8 Conclusions and recommendations

This chapter gives the conclusions and recommendations of our research. Section 8.1 answers the main question of our research in a conclusion, Section 8.2 gives the limitations of this research, Section 8.3 provides recommendations to OCG to make the model commercially more interesting and Section 8.4 gives the recommendations for future theoretical research on this topic.

8.1 Conclusion

The main research question is:

> What is the best mathematical model for the optimization of OilCO's oil and gas flows including dealing with disturbances?

Our model is capable of generating a 30-day schedule that optimizes OilCO's oil and gas production and transportation. It considers OilCO's reservoirs, separation facilities, pipelines, refineries, terminals, storage tanks, and the gas customer while minimizing energy costs, changeover costs at the reservoirs, inventory costs, penalties for non-satisfied demand, penalties for deviating from the production targets, and penalties for having less in stock than the safety stock.

The proposed model is based on an event-based time representation, which proved to have a positive impact on model performance in literature, as it decreases the number of integer variables needed to achieve a solution, while having the same accuracy compared to a discrete time representation. The proposed simplification, based on decreasing the number of integer variables, improves model performance by 58%. As this is substantial it shows that event-based time representation, which requires less integer variables is indeed the best choice.

Our model deals with disturbances by solving it with different parameters and comparing the results in a what-if analysis. Moreover, by applying Robust Optimization solutions are yielded that are, by design, more capable of dealing with disturbances. By setting parameters λ and/or ω greater than zero gives a solution that has respectively a higher solution and/or model robustness. Increasing solution robustness means that the solution is "closer" to optimal for the realizations of the tanker arrivals, while increasing model robustness means that the solution is "closer to feasible" for the realizations of the tanker arrivals. Both types of robustness are realized by varying the throughput rates of the pipelines and pumping less oil to the terminals to prevent infeasibilities, of which the latter causes shortages. The question is to what extend a schedule in which the throughput rates vary heavily is realizable in practice, which is a question we cannot answer with our knowledge of the process. Another question we cannot answer is if solution robustness is really something OilCO wishes to achieve, as total costs have to be increased in order to decrease risk. If the answer is no, than the risk term can be removed from the objective function, i.e., $\lambda = 0$.

The average running time of the robust model is 330 seconds if the simplification is applied, which is in line with our goal that the model should yield the optimal solution within a couple of minutes. The sensitivity analysis shows that the solution is sensitive to changes in terminal storage tank size, the peak efficiency throughput of the pipelines and the cost factors, so changing these input parameters of the model results in different schedules with different costs.

All in all, our research results in a model that is capable of doing what it is supposed to do with a good performance. However, the lack of knowledge of the processes and the lack of data for the input parameters make it impossible to determine if this model really is the best model for OilCO. The sensitivity analysis shows that the values of certain parameters have a large influence on the solution, but is also shows that our model is capable of dealing with different input values.

8.2 Limitations

Our research has the following limitations:

- The lack of direct contact with OilCO made it impossible to confirm certain findings and assumptions during the process, resulting in some uncertainties in the model. The uncertainty in parameter values can have a impact on the model we designed in case the input parameters are incorrect, as the sensitivity analysis showed that the solution of the model is sensitive to changes in the parameters.

- The lack of direct contact with OilCO made it impossible to verify the solutions that are provided by the model. Therefore, we could not adjust the model in case it provides solutions that are impossible to realize.

- By taking the tanker arrival scenarios as fixed input for our robust model, we assumed that OilCO has no flexibility at all in scheduling tankers, meaning that every tanker has to be served on the moment it arrives. OilCO therefore has to be very flexible in deviating with the throughput rates of the pipelines in order to get a robust schedule in which all demand is fulfilled, while in reality there may be some flexibility in scheduling the tankers.

- The model was only applied on two demand test cases in this research, which is limited compared with the numerous demand cases that are possible for OilCO. However, both a stable and a unstable demand case are tested in the experiments. For both cases the model yields an optimal solution in reasonable time, so we expect that the model can be applied on most demand cases without additional problems

8.3 Recommendations

We recommend OCG the following with respect to the further development of an optimization model for pipeline flow scheduling that can be sold to different companies:

- Try to get direct contact with an oil company to be able to verify all assumptions made during the process on oil and gas extraction, separation, and transportation processes.

- In the what-if analysis we explained how our model can be used to schedule maintenance, however this method is a reactive scheduling approach. It could be possible to extend to model in order to include the scheduling of maintenance in the predictive scheduling approach of robust optimization, as was done with the tanker arrival uncertainty. To realize this, more research is needed on maintenance scheduling models and the manner oil companies deal with maintenance.

- The model could be extended with the scheduling of tankers at the loading platforms, as our research showed that the manner in which the tankers are scheduled can have a large impact on the solution. The model then has to determine the three day windows that are given to the tankers and the result of a certain schedule can be calculated with the current model.

- Investigate the possibilities to use the model in other industries. In literature we found an application of pipeline flow scheduling in water transportation, so water companies can be potential customers for OCG.

8.4 Future research

We suggest the following to further develop the model theoretically:

- As the energy costs are a large part of the total costs, more information is needed about the relation between the energy costs and the throughput rate of a pipeline. In our literature review we only considered mathematical models for pipeline flow scheduling optimization, while there might be more information about the energy costs of a pump to realize a certain throughput rate in (mechanical) engineering literature. Modelling the energy costs in a more accurate manner makes the model as a whole more accurate.
- More information is needed about how oil and gas extraction, separation, and transportation processes work at different oil companies. When we know more about these processes and the differences between oil companies, the generalizability of the model can be assessed.

References

Abbasi, E., & Garousi, V. (2010). An MILP-based formulation for minimizing pumping energy costs of oil pipelines: beneficial to both the environment and pipeline companies. *Energy Systems*, *1*(4), 393–416.

Adhitya, A., Srinivasan, R., & Karimi, I. (2007). Heuristic Rescheduling of Crude Oil Operations to Manage Abnormal Supply Chain Events. *American Institute of Chemical Engineers Journal*, *53*(2), 397–422.

Boschetto, S. N., Magata, L., Brondani, W. M., & Relvas, S. (2010). An Operational Scheduling Model to Product Distribution through a Pipeline Network. *Industrial and Engineering Chemistry Research*, *49*(12), 5661–5682.

Cafaro, D. C., & Cerdá, J. (2004). Optimal scheduling of multiproduct pipeline systems using a non-discrete MILP formulation. *Computers & Chemical Engineering*, *28*(10), 2053–2068.

Cafaro, D. C., & Cerdá, J. (2008a). Dynamic scheduling of multiproduct pipelines with multiple delivery due dates. *Computers & Chemical Engineering*, *32*(4-5), 728–753.

Cafaro, D. C., & Cerdá, J. (2008b). Efficient Tool for the Scheduling of Multiproduct Pipelines and Terminal Operations. *Industrial and Engineering Chemistry Research*, *47*(24), 9941–9956.

Cafaro, D. C., & Cerdá, J. (2012). Rigorous scheduling of mesh-structure refined petroleum pipeline networks. *Computers & Chemical Engineering*, *38*, 185–203.

Fabro, J. a. A., Stebel, S. L., Rossato, D., Polli, H. L., Arruda, L. V. R., Neves, F., ... Relvas, S. (2014). A MILP (Mixed Integer Linear Programming) decomposition solution to the scheduling of heavy oil derivatives in a real-world pipeline. *Computers & Chemical Engineering*, *66*, 124–138.

Herrán, a., de la Cruz, J., & de Andrés, B. (2010). A mathematical model for planning transportation of multiple petroleum products in a multi-pipeline system. *Computers & Chemical Engineering*, *34*(3), 401–413.

Jia, Z., & Ierapetritou, M. G. (2004). Efficient short-term scheduling of refinery operations based on a continuous time formulation. *Computers & Chemical Engineering*, *28*(6-7), 1001–1019.

Lee, H., Pinto, J. M., Grossmann, I. E., & Park, S. (1996). Mixed-Integer Linear Programming Model for Refinery Short-Term Scheduling of Crude Oil Unloading with Inventory Management. *Industrial and Engineering Chemistry Research*, *35*(5), 1630–1641.

Li, J., Karimi, I., & Srinivasan, R. (2005). Robust Scheduling of Crude Oil Operations under Demand and Ship Arrival Uncertainty. In *Aiche annual meeting.*

Li, J., Misener, R., & Floudas, C. a. (2012). Scheduling of crude oil operations under demand uncertainty: A robust optimization framework coupled with global optimization. *AIChE Journal*, *58*(8), 2373–2396. Retrieved from `http://doi.wiley.com/10.1002/aic.12772` doi: 10.1002/aic.12772

Manthey, B. (2013). Lecture Notes Optimization Modeling.

Mas, R., & Pinto, J. M. (2003). A Mixed-Integer Optimization Strategy for Oil Supply in Distribution Complexes. *Optimizaiton and Engineering*(4), 23–64.

MirHassani, S. A., & Fani Jahromi, H. (2011). Scheduling multi-product tree-structure pipelines. *Computers & Chemical Engineering*, *35*(1), 165–176.

MirHassani, S. A., Moradi, S., & Taghinezhad, N. (2011). Algorithm for Long-Term Scheduling of Multiproduct Pipelines. *Industrial and Engineering Chemistry Research*, *50*(24), 13899–13910.

Moro, L. F. L., & Pinto. (2004). Mixed-Integer Programming Approach for Short-Term Crude Oil Scheduling. *Industrial and Engineering Chemistry Research*, *43*, 85–94.

Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust Optimization of Large-Scale Systems.pdf. *Operations Research*, *43*(2), 264–281.

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

Ortiz-Gomez, A., Rico-Ramirez, V., & S., H.-C. (2002). Mixed-integer multiperiod model for the planning of oilfield production. *Computers & Chemical Engineering*, *26*, 703–714.

Pan, M., Li, X., & Qian, Y. (2009). New approach for scheduling crude oil operations. *Chemical Engineering Science*, *64*, 965–983.

Reddy, P. C. P., Karimi, I., & Srinivasan, R. (2004a). A new continuous-time formulation for scheduling crude oil operations. *Chemical Engineering Science*, *59*(6), 1325–1341.

Reddy, P. C. P., Karimi, I., & Srinivasan, R. (2004b). Novel Solution Approach for Optimizing Crude Oil Operations. *American Institute of Chemical Engineers Journal*, *50*(6), 1177–1197.

Rejowski, R., & Pinto, J. (2004). Efficient MILP formulations and valid cuts for multiproduct pipeline scheduling. *Computers & Chemical Engineering*, *28*(8), 1511–1528.

Rejowski, R., & Pinto, J. M. (2003). Scheduling of a multiproduct pipeline system. *Computers & Chemical Engineering*, *27*(8-9), 1229–1246.

Rejowski, R., & Pinto, J. M. (2008). A novel continuous time representation for the scheduling of pipeline systems with pumping yield rate constraints. *Computers & Chemical Engineering*, *32*(4-5), 1042–1066.

Relvas, S., Barbosa-Póvoa, A. P. F., & Matos, H. A. (2009). Heuristic batch sequencing on a multiproduct oil distribution system. *Computers & Chemical Engineering*, *33*(3), 712–730.

Relvas, S., Boschetto Magatão, S. N., Barbosa-Póvoa, A. P. F., & Neves, F. (2013). Integrated scheduling and inventory management of an oil products distribution system. *Omega*, *41*(6), 955–968.

Relvas, S., Matos, H. A., & Barbosa-Póvoa, A. P. F. (2006). Pipeline Scheduling and Inventory Management of a Multiproduct Distribution Oil System. *Industrial and Engineering Chemistry Research*, *45*, 7841–7855.

Relvas, S., Matos, H. A., & Barbosa-Póvoa, A. P. F. (2007). Reactive Scheduling Framework for a Multiproduct Pipeline with Inventory Management. *Industrial and Engineering Chemistry Research*, *46*(17), 5659–5672.

Saharidis, G. K. D., Minoux, M., & Dallery, Y. (2009). Scheduling of loading and unloading of crude oil in a refinery using event-based discrete time formulation. *Computers & Chemical Engineering*, *33*(8), 1413–1426.

Shah, N. (1996). Mathematical Programming Techniques for Crude Oil Scheduling. , *20*(96), 1227–1232.

Stebel, S. L., Magata, S. N. B., Povoa, A. P. F. D., Arruda, L. V. R., & Relvas, S. (2012). Mixed Integer Linear Programming Formulation for Aiding Planning Activities in a Complex Pipeline Network. *Industrial and Engineering Chemistry Research*, *51*(35), 11417–11433.

Verweij, B., Ahmed, S., Kleywegt, A. J., & Nemhauser, G. (2003). The Sample Average Approximation Method Applied to Stochastic Routing Problems :. *Computational Optimization and Applications*, *24*(1), 289–333.

Wang, J., & Rong, G. (2010). Robust Optimization Model for Crude Oil Scheduling under Uncertainty. *Industrial and Engineering Chemistry Research*, *49*, 1737–1748.

Yadav, S., & Shaik, M. A. (2012). Short-Term Scheduling of Refinery Crude Oil Operations. *Industrial and Engineering Chemistry Research*, *51*, 9287–9299.

Appendices

Appendix A Pipeline overview

#	From	To	Length (km)	Diameter (inch)	Capacity (kbbl/day)	Adjusted Capacity (kbbl/day)
1	Reservoir6	SepFac2	19.7	16	36	100
2	Reservoir5	SepFac2	unknown	unknown	25	100
3	Reservoir4	SepFac2	0	unknown	unknown	N/A
4	SepFac2	SepFacTank2	0	unknown	unknown	N/A
5	SepFacTank2	Node 1	35	24	360	460
6	SepFacTank2	Node 1	35	24	360	460
7	Reservoir0	SepFac1	0	unknown	unknown	N/A
8	Shah	SepFac1	68	12	60	60
9	Shah	SepFac1	68	16	100	100
10	Reservoir3	SepFac1	80	14	100	100
11	Reservoir1	SepFac1	45	12	120	120
12	Reservoir1	SepFac1	45	16	75	75
13	SepFac1	SepFacTank1	0	unknown	unknown	N/A
14	SepFacTank1	Node 1	86	32	720	820
15	SepFacTank1	Node 1	86	36	420	520
16	Reservoir9	SepFac4	0	unknown	unknown	N/A
17	SepFac4	SepFacTank4	0	unknown	unknown	N/A
18	SepFacTank4	Node 1	0	unknown	unknown	N/A
19	Reservoir8	SepFac3	29.6	12	55	100
20	Reservoir7	SepFac3	31	20	84	150
21	SepFac3	SepFacTank3	0	unknown	unknown	N/A
22	SepFacTank3	Node 1	70	20	170	170
23	Node 1	Refinery2	151	18	unknown	N/A
24	Node 1	Refinery2	395.5	48	1500	1500
25	TerminalTank2	Terminal2	0	unknown	unknown	N/A
26	Node 1	Node 2	112	32	640	640
27	Node 1	Node 2	112	36	924	924
28	Node 1	Node 2	112	24	312	312
29	Node 2	Refinery1	unknown	unknown	unknown	N/A
30	Node 2	TerminalTank1	unknown	unknown	unknown	N/A
31	TerminalTank1	Terminal1	0	unknown	unknown	N/A

Table 8: Overview of all oil pipelines in OilCO's network

Appendix B Model notation

Sets/indices

A - Assets, with $a, a' \in A$

R - Reservoirs (1,2,...,10), with $r \in R$ and $R \subset A$

SF - Separation Facilities (1,2,3,4), with $sf \in SF$ and $SF \subset A$

P - Pipelines for oil (1,2,...,34), with $p \in P$

EP - Pipelines of which the energy costs are relevant, with $ep \in EP$ and $EP \subset P$

T - Terminals (1,2), with $te \in TE$ and $TE \subset A$

RF - Refinery (1,2) with $rf \in RF$ and $RF \subset A$

ST - Storage tanks (1,2,...,6) with $st \in S$T and $ST \subset A$

T - Time period (1,2,...,nt) with $t \in T$

D - Days (1,2,...,30) with $d \in D$

N - Pipeline nodes $n \in N$ and $N \subset A$

L - Line segments of the piecewise linear function, with $l \in L$

C - Pipeline connections, which are pairs such that pipeline p connects asset a to asset a'

T_d - Set of time periods that are in day d, with $T_d \subset T$

Parameters

nt = Number of time periods

dur_t = Duration of time period t in hours

$rmax_r$ = Max gas/oil production rate of reservoir r in bbl/h

$rmin_r$ = Min gas/oil production rate of reservoir r in bbl/h

gor_r = Gas-oil ratio of reservoir r $(\frac{volume\,of\,gas}{volume\,of\,oil})$

$oilp_r$ = Oil fraction of the reservoir r $(\frac{volume\,of\,oil}{total\,volume})$

$plan_{r,d}$ = Production target of reservoir r on day d in bbl

$PR_{r,0}$ = Current oil production rate of reservoir r in bbl/h

$srmax_{sf}$ = Maximum separation capacity of separation facility sf in bbl/h

$av_{sf,t}$ = Availability of separation facility sf during time period t in %

$oilloss$ = Percentage of oil that is lost in the separation process

$pcap_p$ = Maximum throughput of pipeline p in bbl/h

$ava_{p,t}$ = Availability of pipeline p during time period t in %

$rdem_{rf,t}$ = Demanded volume in bbl of refinery rf in time period t

$SHR_{rf,0}$ = Current total shortage of refinery rf in bbl

$gdem_d$ = Demanded volume in bbl of BuyGas on day d

$tdem_{te,t}$ = Demanded volume in bbl of terminal te in time period t

$SHT_{te,0}$ = Current total shortage of terminal te in bbl

$mininv_{st}$ = Minimum inventory of storage tank st in bbl

$maxinv_{st}$ = Maximum inventory of storage tank st in bbl

$safinv_{st}$ = Safety stock of storage tank st in bbl

$I_{st,0}$ = Current inventory of storage tank st in bbl

ea_p = Constant that is used for calculating the efficiency of pipeline p

eb_p = Constant that is used for calculating the efficiency of pipeline p

ec_p = Constant that is used for calculating the efficiency of pipeline p

ed_p = Constant that is used for calculating the efficiency of pipeline p

$csht_{te}$ = Costs of one bbl of shortage at terminal t in USD

$cshr_{rf}$ = Costs of one bbl of shortage at refinery rf in USD

cc_r = Changeover costs of a changeover at reservoir r in USD

$cdev_r$ = Costs for deviating from the production plan at reservoir r in USD

$cinv_{st}$ = Costs for holding one bbl of oil in stock for one hour at storage tank st in USD

ssp_{st} = Safety stock penalty per bbl under safety stock at storage tank st in USD

ca_p = Variable pumping costs at pipeline p in USD

cb_p = Constant pumping costs at pipeline p in USD

csq_p = Cost factor of the squared positive deviation from the peak efficiency throughput at pipeline p in USD

TH_p^* = Peak efficiency throughput of pipeline p

TH_p^{max} = Zero efficiency throughput of pipeline p

EFF_p^* = Maximum efficiency of pipeline p

$brp_{p,t,l}$ = Value of $PTD_{p,t}$ at the breakpoint at the end of line l

$sbrp_{p,t,l}$ = Value of $SPTD_{p,t}$ at the breakpoint at the end of line l

M = Big M, an auxiliary parameter

Decision Variables

$PR_{r,t}$ = Production rate of oil of reservoir r in bbl/h in time period t in bbl/h (Continuous)

$SR_{sf,t}$ = Separation rate of separation facility sf during time period t in bbl/h (Continuous)

$TH_{p,t}$ = Throughput of pipeline p in time period t in bbl/h (Continuous)

Auxiliary Variables

$DEV_{r,d}$ = Deviation from the production target at reservoir r on day d in bbl (Continuous)

GAS_d = Gas extracted from the non-associated gas field on day d in cf (Continuous)

$GFR_{sf,t}$ = Gas flow rate after separation in separation facility sf during time period t in cf/h (Continuous)

$EN_{p,t}$ = Energy costs in USD at pipeline p during time period t (Continuous)

$EFF_{p,t}$ = Efficiency of the pumps at pipeline p during time period t (Continuous)

$PTD_{p,t}$ = Positive throughput deviation from TH_p^* in time period t in bbl/h (Continuous)

$SPTD_{p,t}$ = Squared positive throughput deviation from TH_p^* in $(\text{bbl/h})^2$ (Continuous)

$\lambda_{p,t,l}$ = Weight of the break point at the end of line l at pipeline p during time period t (Continuous)

$SHT_{te,t}$ = Total shortage at terminal te at the end of time period t in bbl (Continuous)

$SHR_{rf,t}$ = Total shortage at refinery rf at the end of time period t in bbl (Continuous)

$I_{st,t}$ = Inventory of storage tank st at the end of time period t in bbl (Continuous)

$AI_{st,t}$ = Average inventory of storage tank st during time period t in bbl (Continuous)

$BSS_{st,t}$ = Volume in bbl that the stock level at storage tank st is below the safety stock at the end of time period t (Continuous)

$CH_{r,t}$ = 1 if there is a changeover in production rate of reservoir r between time period t and time period t-1 (Binary)

Additional declarations for robust model

Sets/indices

S - Scenarios, sample of a tanker delay scenario that is used in the model, with $s \in S$ and $S \subset \Omega$

CST - Control storage tanks, which are the storage tanks at the terminals, with $cst \in CST$ and $CST \subset ST$

CP - Control pipelines, which are the pipelines that connect terminal tanks to the terminals, with $cp \in CP$ and $CP \subset P$

Ω - Set of all possible tanker arrival scenarios.

Parameters

$tdem^s_{te,t}$ = Demanded volume in bbl of terminal te in time period t of scenario s

pen_{st} = Penalty in USD per bbl for violating the maximum inventory of storage tank st (with $st \in CST$)

ω = Parameter that is used to make a trade-off between model robustness and solution robustness in the objective function

λ = Parameter that is used to make a trade-off between the sample average and sample variance in the objective function

Variables

$Z^s_{st,t}$ = Auxiliary variable to measure tank capacity violation of storage tank st (with $st \in CST$) in time period t of scenario s in bbl (Continuous)

$TH^s_{p,t}$ = Throughput of pipeline p (with $p \in CP$) in time period t of scenario s in bbl/h (Continuous)

$SHT^s_{te,t}$ = Total shortage at terminal te at the end of time period t of scenario s in bbl (Continuous)

$I^s_{st,t}$ = Inventory of storage tank st (with $st \in CST$) at the end of time period t of scenario s in bbl (Continuous)

$AI^s_{st,t}$ = Average inventory of storage tank st (with $st \in CST$) during time period t of scenario s in bbl (Continuous)

$BSS^s_{st,t}$ = Volume in bbl that the stock level at storage tank st (with $st \in CST$) is below the safety stock at the end of time period t of scenario s (Continuous)

ξ_s = Objective function of scenario s (Continuous)

Appendix C Energy costs analysis

In this analysis we want to find out the relation between the throughput rate of a pipeline and the associated energy costs by creating an example of a pump based on the work of Rejowski & Pinto (2008) and Abbasi & Garousi (2010). According to these authors an important variable in calculating the energy costs is the efficiency of the pumps at a pipeline. Here, variable $EFF_{p,t}$ denotes the efficiency of the pumps at pipeline p during time period t. This efficiency is a function of the throughput ($TH_{p,t}$) at pipeline p during time period t. Rejowski & Pinto (2008) and Abbasi & Garousi (2010) propose a third degree polynomial to describe this function, which (38) shows.

$$EFF_{p,t} = ea_p\, TH_{p,t}^3 + eb_p\, TH_{p,t}^2 + ec_p\, TH_{p,t} + ed_p \qquad \forall p \in EP;\ t \in T \tag{38}$$

The parameters ea_p, eb_p, ec_p, and ed_p are constants that depend on the characteristics of the pumps at pipeline p, where we assume that for each pipeline only one type of pump is used (see Section 4.2). Rejowski & Pinto (2008) state that the parameters ea_p, eb_p, ec_p, and ed_p can be requested at the pump manufacturer and give two examples for values of the parameters. We cannot use this approach as we have no contact with OilCO's pump manufacturers. Abbasi & Garousi (2010) estimate the parameters ea_p, eb_p, ec_p, and ed_p using extra parameters, for which they define EFF_p^* as the peak efficiency of pipeline p, which is the maximum efficiency that the pumps at pipeline$p$ can realize. TH_p^* is defined as the peak efficiency throughput of pipeline p, which is the throughput rate at which the peak efficiency is achieved. Since $ed_o = 0$ they approximate the efficiency with a cubic approximation as in (38) that goes through the points $(0, 0)$, (TH_p^*, EFF_p^*) and $(TH_p^{max}, 0)$, where TH_p^{max} is the throughput rate associated with the point where the efficiency is 0 again at pipeline p (zero efficiency point). When EFF_p^*, TH_p^*, and TH_p^{max} are known, the constants in (38) can be calculated by Equations (39)-(41).

$$ea_p = \frac{EFF_p^*(2\,TH_p^* - TH_p^{max})}{-(TH_p^*)^2(TH_p^{max} - TH_p^*)^2} \qquad \forall p \in EP \tag{39}$$

$$eb_p = \frac{-EFF_p^*(3\,(TH_p^*)^2 - (TH_p^{max})^2)}{-(TH_p^*)^2(TH_p^{max} - TH_p^*)^2} \qquad \forall p \in EP \tag{40}$$

$$ec_p = \frac{EFF_p^*\, TH_p^*\, TH_p^{max}(3\,TH_p^* - 2\,TH_p^{max})}{-(TH_{p,}^*)^2(TH_p^{max} - TH_p^*)^2} \qquad \forall p \in EP \tag{41}$$

Abbasi & Garousi (2010) state that TH_p^{max} is close to $2\,TH_p^*$. We have no information on what values for EFF_p^*, TH_p^*, and TH_p^{max} the pumps at OilCO have, so we illustrate an example to see what the relation between the throughput rate and the energy consumed is. According to the example pumps modelled by Rejowski & Pinto (2008), EFF_p^* lies around 0.70. Unlike Abbasi & Garousi (2010), Rejowski & Pinto (2008) do not assume that $ed_p = 0$, since assuming that $ed_p = 0$ causes unrealistic behavior of energy costs when the efficiency is close to 0 (efficiency is in the denominator of the function). To eliminate these large differences we need ed_p to be larger than 0, therefore we use $ed_p = 0.25$ in our example for illustration. To get the top of our efficiency curve at 0.70, $EFF_p^* = 0.70 - 0.25 = 0.45$. We show an example where $TH_p^* = 100$, $TH_p^{max} = 200$ (so $TH_p^{max} = 2\,TH_p^*$), and $EFF_p^* = 0.45$. We choose this values, because these result in a good illustration of the relation between the throughput and energy costs, while they are plausible according to Rejowski & Pinto

(2008) and Abbasi & Garousi (2010). These values result in $ea_p = 0$ (because $TH_p^{max} = 2TH_p^*$), $eb_p = -0.000045$, and $ec_p = 0.009$.

Rejowski & Pinto (2008) propose (42) to calculate the energy costs, so the energy costs depend on the throughput rate of a pipeline, a cost factor (ca_p) for pipeline p and the efficiency of the pipeline. This equation is non-linear, because a variable is divided by another variable. We can replace $EFF_{p,t}$ in (42) by its definition given in (38) of which (43) shows the result in which the energy costs just depend on one variable: the throughput rate.

$$EN_{p,t} = \frac{ca_p\, TH_{p,t}}{EFF_{p,t}} \qquad \forall p \in EP;\, t \in T \qquad (42)$$

$$EN_{p,t} = \frac{ca_p\, TH_{p,t}}{ea_p\, TH_{p,t}^3 + eb_p\, TH_{p,t}^2 + ec_p\, TH_{p,t} + ed_p} \qquad \forall p \in EP;\, t \in T \qquad (43)$$

Figure 22 shows the efficiency of this pump of different values of $TH_{p,t}$, as calculated by (38). Parameter TH_p^* is denoted by the vertical dashed line.

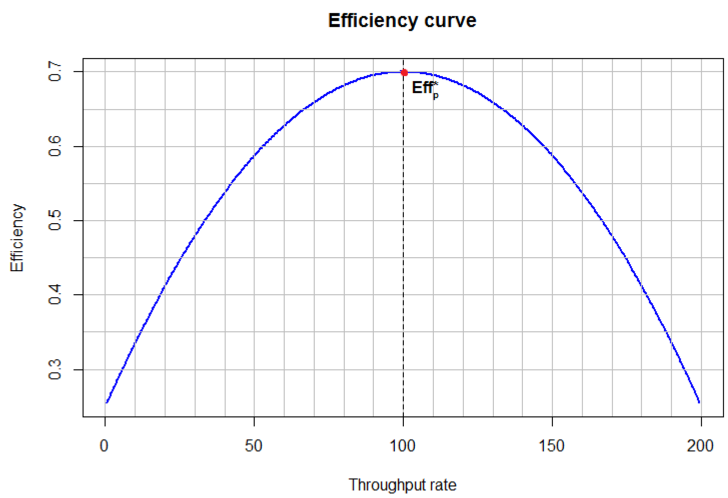

Figure 22: The efficiency curve of the example pump

We calculate the energy costs for different values of $TH_{p,t}$ using (43). In this example $ca_p = 1.8$, because this leads to a clear example of the relation between throughput and energy costs. Figure 6 in Section 4.5.1 shows the energy costs for different values of $TH_{p,t}$. We now have an idea of the relation between the throughput rate and the associated energy costs.

Appendix D Deterministic model

$$\min \sum_{p,t} dur_t \, EN_{p,t} + \sum_{r,t} cc_r \, CH_{r,t} + \sum_{st,t} cinv_{st} \, dur_t \, AI_{st,t} + \sum_{st,t} ssp_{st} \, BSS_{st,t}$$
$$+ \sum_{te,t} csht_{te} \, SHT_{te,t} + \sum_{rf,t} cshr_{rf} \, SHR_{rf,t} + \sum_{r,d} cdev_r \, DEV_{r,d}$$

s.t.
Operational constraints

$$PR_{r,t} = \sum_{(p,a)|(r,p,a)\in C} TH_{p,t} \qquad \forall r \in R;\ t \in T$$

$$rmin_r \, avail_{r,t} \leq PR_{r,t} \leq rmax_r \, avail_{r,t} \qquad \forall r \in R;\ t \in T$$

$$\sum_{(r,p)|(r,p,sf)\in C} TH_{p,t} = SR_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$\sum_{(r,p)|(r,p,sf)\in C} TH_{p,t} \, gor_r = GFR_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$SR_{sf,t} \leq srmax_{sf} \, av_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$(1 - oilloss) \, SR_{sf,t} = \sum_{(p,a)|(sf,p,a)\in C} TH_{p,t} \qquad \forall sf \in SF;\ t \in T$$

$$pmin_p \, ava_{p,t} \leq TH_{p,t} \leq pmax_p \, ava_{p,t} \qquad \forall p \in P;\ t \in T$$

$$\sum_{(a,p)|(a,p,n)\in C} TH_{p,t} = \sum_{(p,a)|(n,p,a)\in C} TH_{p,t} \qquad \forall n \in N;\ t \in T$$

$$I_{st,t} = I_{st,t-1} + \sum_{(a,p)|(a,p,st)\in C} TH_{p,t} \, dur_t - \sum_{(p,a)|(st,p,a)\in C} TH_{p,t} \, dur_t \qquad \forall st \in ST;\ t \in T$$

$$mininv_{st} \leq I_{st,t} \leq maxinv_{st} \qquad \forall st \in ST;\ t \in T$$

$$\sum_{(a,p,tp)|(a,p,rf)\in C} TH_{p,t} \, dur_t = rdem_{rf,t} - SHR_{rf,t} + SHR_{rf,t-1} \qquad \forall rf \in RF;\ t \in T$$

$$\sum_{(a,p)|(a,p,te)\in C} TH_{p,t} \, dur_t = tdem_{te,t} - SHT_{te,t} + SHT_{te,t-1} \qquad \forall te \in TE;\ t \in T$$

$$\sum_{(sf,t)|t\in T_d=1} GFR_{sf,t} \, dur_t + GAS_d \geq gdem_d \qquad \forall d \in D$$

Auxiliary constraints

$$
\begin{aligned}
AI_{st,t} &= \frac{I_{st,t} + I_{st,t-1}}{2} && \forall st \in ST;\ t \in T \\
PTD_{p,t} &\geq TH_{p,t} - TH_p^* && \forall p \in EP;\ t \in T \\
PTD_{p,t} &\geq 0 && \forall p \in EP;\ t \in T \\
\sum_l \lambda_{p,t,l}\, sbrp_{p,t,l} &= SPTD_{p,t} && \forall p \in EP;\ t \in T \\
\sum_l \lambda_{p,t,l}\, brp_{p,t,l} &= PTD_{p,t} && \forall p \in EP;\ t \in T \\
\sum_l \lambda_{p,t,l} &= 1 && \forall p \in EP;\ t \in T \\
EN_{p,t} &= ca_p\, TH_{p,t} + cb_p + csq_p\, SPTD_{p,t} && \forall p \in EP;\ t \in T \\
(PR_{r,t} - PR_{r,t-1}) &\leq CH_{r,t}\, M && \forall r \in R;\ t \in T;\ t > 1 \\
(PR_{r,t-1} - PR_{r,t}) &\leq CH_{r,t}\, M && \forall r \in R;\ t \in T;\ t > 1 \\
(PR_{r,t} - curp_r) &\leq CH_{r,t}\, M && \forall r \in R;\ t \in T;\ t = 1 \\
(curp_r - PR_{r,t}) &\leq CH_{r,t}\, M && \forall r \in R;\ t \in T;\ t = 1 \\
DEV_{r,d} &\geq plan_{r,d} - \sum_{t \in T_d} PR_{r,t}\, dur_t && \forall r \in R;\ d \in D \\
DEV_{r,d} &\geq \sum_{t \in T_d} PR_{r,t}\, dur_t - plan_{r,d} && \forall r \in R;\ d \in D \\
BSS_{st,t} &\geq safinv_{st} - I_{st,t} && \forall st \in ST;\ t \in T \\
BSS_{st,t} &\geq 0 && \forall st \in ST;\ t \in T
\end{aligned}
$$

Appendix E Robust model

$$
\begin{aligned}
\min & \sum_{p,t} dur_t\, EN_{p,t} + \sum_{r,t} cc_r\, CH_{r,t} + \sum_{rf,t} cshr_{rf}\, SHR_{rf,t} \\
& + \sum_{r,d} cdev_r\, DEV_{r,d} + \sum_{st \notin CST,t} ssp_{st}\, BSS_{st,t} + \sum_{st \notin CST,t} cinv_{st}\, dur_t\, AI_{st,t} \\
& + \frac{1}{|S|} \sum_{s} \Big(\sum_{st \in CST,t} ssp_{st}\, BSS^{s}_{st,t} + \sum_{st \in CST,t} cinv_{st}\, dur_t\, AI^{s}_{st,t} + \sum_{te,t} csht_{te}\, SHT^{s}_{te,t} \Big) \\
& + \lambda \frac{1}{|S|} \sum_{s} \Big| \sum_{st \in CST,t} ssp_{st}\, BSS^{s}_{st,t} \sum_{st \in CST,t} cinv_{st}\, dur_t\, AI^{s}_{st,t} + \sum_{te,t} csht_{te}\, SHT^{s}_{te,t} \\
& - \frac{1}{|S|} \sum_{s'} \Big(\sum_{st \in CST,t} ssp_{st}\, BSS^{s'}_{st,t} + \sum_{st \in CST,t} cinv_{st}\, dur_t\, AI^{s'}_{st,t} + \sum_{te,t} csht_{te}\, SHT^{s'}_{te,t} \Big) \Big| \\
& + \omega \frac{1}{|S|} \sum_{s} \Big(\sum_{st \in CST,t} pen_{st}\, Z^{s}_{st,t} \Big)
\end{aligned}
$$

s.t.
Operational constraints

$$PR_{r,t} = \sum_{(p,a)|(r,p,a) \in C} TH_{p,t} \qquad \forall r \in R;\ t \in T$$

$$rmin_r\, avail_{r,t} \leq PR_{r,t} \leq rmax_r\, avail_{r,t} \qquad \forall r \in R;\ t \in T$$

$$\sum_{(r,p)|(r,p,sf) \in C} TH_{p,t} = SR_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$\sum_{(r,p)|(r,p,sf) \in C} TH_{p,t}\, gor_r = GFR_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$SR_{sf,t} \leq srmax_{sf}\, av_{sf,t} \qquad \forall sf \in SF;\ t \in T$$

$$(1 - oilloss)\, SR_{sf,t} = \sum_{(p,a)|(sf,p,a) \in C} TH_{p,t} \qquad \forall sf \in SF;\ t \in T$$

$$pmin_p\, ava_{p,t} \leq TH_{p,t} \leq pmax_p\, ava_{p,t} \qquad \forall p \in P;\ t \in T$$

$$\sum_{(a,p)|(a,p,n) \in C} TH_{p,t} = \sum_{(p,a)|(n,p,a) \in C} TH_{p,t} \qquad \forall n \in N;\ t \in T$$

$$I_{st,t} = I_{st,t-1} + \sum_{(a,p)|(a,p,st) \in C} TH_{p,t}\, dur_t - \sum_{(p,a)|(st,p,a) \in C} TH_{p,t}\, dur_t \qquad \forall st \in ST;\ t \in T$$

$$I^{s}_{st,t} = I^{s}_{st,t-1} + \sum_{(a,p)|(a,p,st) \in C} TH_{p,t}\, dur_t - \sum_{(p,a)|(st,p,a) \in C \wedge p \in CP} TH^{s}_{p,t}\, dur_t \qquad \forall st \in ST;\ t \in T;\ s \in S$$

$$mininv_{st} \leq I_{st,t} \leq maxinv_{st} \qquad \forall st \in ST;\ t \in T$$
$$mininv_{st} \leq I^s_{st,t} \leq maxinv_{st} + Z^s_{st,t} \qquad \forall st \in CST;\ t \in T\,;s \in S$$
$$\sum_{(a,p,tp)|(a,p,rf)\in C} TH_{p,t}\, dur_t = rdem_{rf,t} - SHR_{rf,t} + SHR_{rf,t-1} \qquad \forall rf \in RF;\ t \in T$$
$$\sum_{(a,p)|(a,p,te)\in C} TH^s_{p,t}\, dur_t = tdem^s{}_{te,t} - SHT^s_{te,t} + SHT^s_{te,t-1} \qquad \forall te \in TE;\ t \in T;\ s \in S$$
$$\sum_{(sf,t)|t\in T_d} GFR_{sf,t}\, dur_t + GAS_d \geq gdem_d \qquad \forall d \in D$$

Auxiliary constraints

$$AI_{st,t} = \frac{I_{st,t} + I_{st,t-1}}{2} \qquad \forall st \notin CST;\ t \in T$$
$$AI^s_{st,t} = \frac{I^s_{st,t} + I^s_{st,t-1}}{2} \qquad \forall st \in CST;\ t \in T;\ s \in S$$
$$PTD_{p,t} \geq TH_{p,t} - TH^*_p \qquad \forall p \in EP;\ t \in T$$
$$PTD_{p,t} \geq 0 \qquad \forall p \in EP;\ t \in T$$
$$\sum_l \lambda_{p,t,l}\, sbrp_{p,t,l} = SPTD_{p,t} \qquad \forall p \in EP;\ t \in T$$
$$\sum_l \lambda_{p,t,l}\, brp_{p,t,l} = PTD_{p,t} \qquad \forall p \in EP;\ t \in T$$
$$\sum_l \lambda_{p,t,l} = 1 \qquad \forall p \in EP;\ t \in T$$
$$EN_{p,t} = ca_p\, TH_{p,t} + cb_p + csq_p\, SPTD_{p,t} \qquad \forall p \in EP;\ t \in T$$

$$(PR_{r,t} - PR_{r,t-1}) \leq CH_{r,t}\, M \qquad \forall r \in R;\ t \in T;\ t > 1$$
$$(PR_{r,t-1} - PR_{r,t}) \leq CH_{r,t}\, M \qquad \forall r \in R;\ t \in T;\ t > 1$$
$$DEV_{r,d} \geq plan_{r,d} - \sum_{t\in T_d} PR_{r,t}\, dur_t \qquad \forall r \in R;\ d \in D$$
$$DEV_{r,d} \geq \sum_{t\in T_d} PR_{r,t}\, dur_t - plan_{r,d} \qquad \forall r \in R;\ d \in D$$
$$BSS^s_{st,t} \geq safinv_{st} - I^s_{st,t} \qquad \forall st \notin CST;\ t \in T;\ \forall s \in S$$
$$BSS^s_{st,t} \geq 0 \qquad \forall st \notin CST;\ t \in T;\ \forall s \in S$$
$$BSS_{st,t} \geq safinv_{st} - I_{st,t} \qquad \forall st \in ST;\ t \in T$$
$$BSS_{st,t} \geq 0 \qquad \forall st \in ST;\ t \in T$$

Appendix F Number of breakpoints required for the piecewise linear approximation

In this appendix we determine the number of breakpoints required for the piecewise linear approximation of the energy costs. The more breakpoints we use for the approximation, the more accurate our approximation is, but the longer it takes to solve the model. We perform an analysis for 1 to 25 breakpoints in the model to see where the "optimum" lies. In the analysis we keep an equal distance between the breakpoints in terms of throughput. We solve the model for the stable demand case. The approximation yields energy costs that are higher than in reality, since we approximate a convex function. Per solution we calculate the "real" energy costs for the yielded solution to examine if the model finds a better solution. Figure 23 shows the real energy costs and the CPU time required to solve the model for a different number of breakpoints.

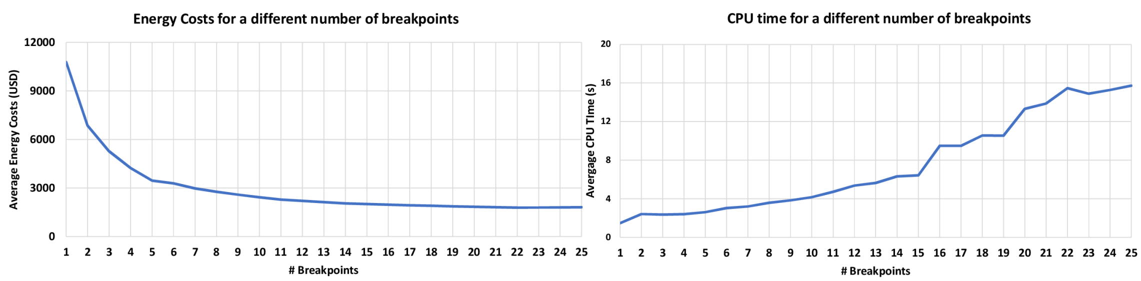

Figure 23: Energy costs and CPU time for a different number of breakpoints

The figures show that the real energy costs decrease as the number of breakpoints increases. The real energy costs, however, slightly increase if the number of breakpoints is 23 or more, since the model yields a different solution. The CPU time needed globally increases with the number of breakpoints, although sometimes adding a single breakpoint can decrease CPU time. We add a breakpoint if adding the breakpoint improves the approximation by 5% or more. When the number of breakpoints is 11 an additional breakpoint does not improve the approximation by more than 5% (only 3,6%). Therefore, we choose to take 11 breakpoints in our approximation. If CPU time was less an issue, we could take more breakpoints. For example, if our critical value is 1% a number of 22 breakpoints is yielded.

Appendix G Number of scenarios required in the robust model

In this appendix we determine the number of scenarios that we require in our sample set S in order for the solution of robust model to be significant. We do this by solving the robust model for a different number of scenarios and see to what extend the costs associated with the design variables change. Adding extra scenarios increases the solving time of the model, so we want to include as less scenarios as possible. If adding an extra scenario does not change the design variable costs significantly any more, then we have the number of scenarios needed. To be certain the number of scenarios applies is enough for all instances of the model, we take an instance for which we expect the design variable costs to vary the most. Therefore, we perform the experiment with the unstable demand case with $\omega = 100$. Figure 24 shows the design variable costs for a number of scenarios that ranges from 1 to 100.

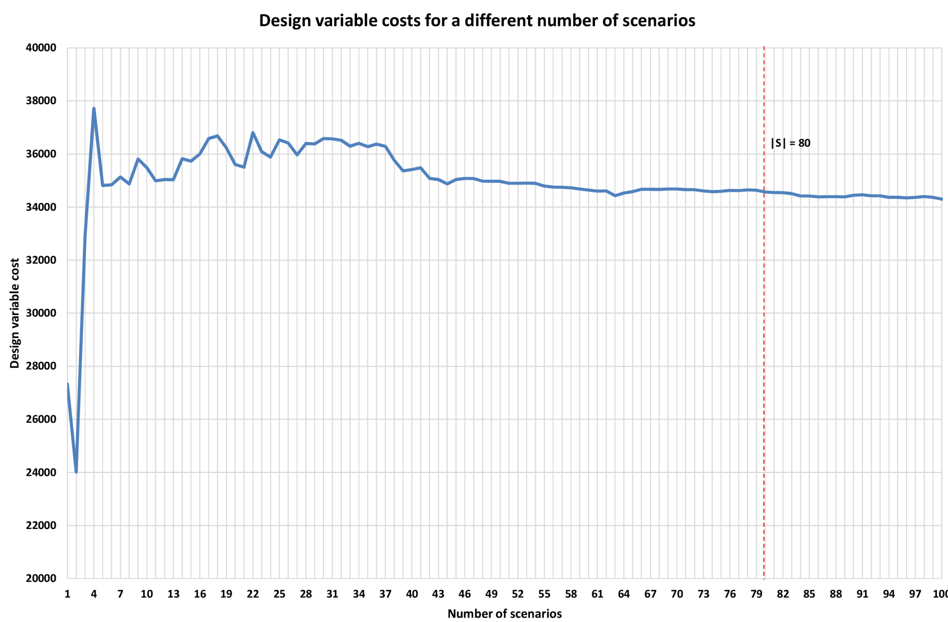

Figure 24: Total costs for a different number of scenarios

The figure shows that the costs fluctuate when the number of scenarios is low. When the number of scenarios is more than 80 (the vertical red dashed line), the total costs are more or less stable. Between 80 and 100 scenarios the decrease in total costs is around 265 USD, which is less than 1% of total costs. Therefore, we use 80 scenarios in our experiments.

Appendix H Division of costs in the results of the stable demand case

Model	λ	ω	Energy costs	Inventory costs	Safety stock penalties	Terminal shortage penalties
Det.	N/A	N/A	2,375.91	24.63	4.53	-
Robust	0	0	2,375.91	27.03	1.68	-
Robust	0	1	2,375.91	27.03	1.68	-
Robust	0	10	2,375.91	27.03	1.68	1.48
Robust	0	100	2,375.91	27.03	1.82	1.93
Robust	0	1,000	2,572.14	27.03	4.23	-
Robust	1	0	2,572.14	27.03	4.23	-
Robust	1	1	2,572.14	27.03	4.21	2.25
Robust	1	10	2,571.99	27.03	4.42	3.21
Robust	1	100	3,037.18	27.03	11.01	-
Robust	1	1,000	3,036.42	27.03	10.48	-
Robust	10	0	3,033.43	27.03	10.42	4.42
Robust	10	1	3,032.37	27.03	10.38	6.81
Robust	10	10	4,179.78	27.03	15.51	-
Robust	10	100	4,179.14	27.03	15.51	-
Robust	10	1,000	4,179.14	27.03	15.51	8.05
Robust	100	0	4,179.78	27.03	16.16	10.35
Robust	100	1	5,504.68	27.03	18.83	479.21
Robust	100	10	5,486.41	27.03	18.66	1.17
Robust	100	100	5,486.65	27.03	18.68	7.89
Robust	100	1,000	5,486.71	27.03	19.30	12.84

Table 9: Division of total costs over the different cost components for the stable demand case

Appendix I Division of costs in the results of the unstable demand case

Model	λ	ω	Energy costs	Change-over costs	Prod. plan deviation penalties	Safety stock penalties	Inven-tory costs	Terminal shortage penalties
Det.	N/A	N/A	13,405.75	10,000	2,883.00	89.96	18.17	141.57
Robust	0	0	12,083.10	10,000	2,883.00	90.27	22.56	137.44
Robust	0	1	12,261.48	10,000	2,883.00	89.94	22.56	0.19
Robust	0	10	12,261.22	10,000	2,883.00	90.66	22.56	10.91
Robust	0	100	12,261.12	10,000	2,883.00	90.90	22.56	11.88
Robust	0	1,000	12,805.94	10,000	2,883.00	78.64	22.56	247.32
Robust	1	0	13,112.63	10,000	2,883.00	77.00	22.56	1.59
Robust	1	1	13,114.02	10,000	2,883.00	76.87	22.56	10.94
Robust	1	10	13,113.30	10,000	2,883.00	76.45	22.56	14.13
Robust	1	100	14,449.60	10,000	2,766.46	78.35	22.30	945.29
Robust	1	1,000	14,430.14	10,000	2,870.68	78.24	22.53	319.44
Robust	10	0	14,605.53	10,000	2,883.00	76.56	22.56	535.35
Robust	10	1	14,604.14	10,000	2,883.00	77.32	22.56	540.98
Robust	10	10	22,318.03	10,000	2,174.32	72.72	20.99	21,971
Robust	10	100	22,129.69	10,000	2,174.32	75.80	20.98	12,066
Robust	10	1,000	22,483.13	10,000	2,216.37	102.76	21.08	14,396
Robust	100	0	22,477.47	10,000	2,216.37	694.21	21.08	13,816
Robust	100	1	37,916.09	-	-	157.53	16.24	164,771
Robust	100	10	37,691.47	-	-	160.66	16.21	137,388
Robust	100	100	33,756.97	10,000	1,552.40	67.59	19.87	256,497
Robust	100	1,000	34,130.91	10,000	1,552.40	6,231	19.86	249,699

Table 10: Division of total costs over the different cost components for the unstable demand case

Appendix J Sensitivity analysis: storage tank size based on average daily volume

We multiply the volumes of Table 7 by 1, 1.5, 2, 2.5, 3, 3.5, and, 4 and solve the model for both cases with both $\omega = 0$ and $\omega = 100$. Figure 25 shows the results in the total costs, infeasibility percentage, CPU time, and average violations per infeasible scenario. All the graphs look similar to the graphs of Figure 19, which is another proof that the influence of the terminal storage tank size on the model is larger than the influence of the separation tank size. In case the tank sizes are equal to the daily average volumes in the stable demand case, there are again changeovers at Reservoir9 and Reservoir0 in the solution. In Section 7.2.1 the changeovers were caused by the differences in daily average volumes between the separation facility storage tanks. In this experiment the storage tank sizes are proportional to the daily average volumes, so the changeovers have other causes. Since the storage tanks in this case are relatively small, throughput rates have to fluctuate more in order to get a feasible solution. Fluctuating in throughput rate is expensive, as we saw in Section 6.3.2, meaning that it is cheaper to produce more at the reservoir that is closest to Node 1 (Reservoir9) and produce less at reservoirs that are farther away to save energy costs, even though two changeovers are needed. In the unstable case for $\omega = 100$ when the storage tanks sizes are equal to the daily average volumes, even three changeovers occur in the solution to deal with the peak in demand. No changeovers occur in the unstable case if the storage tank sizes are at least 3 times the daily average volumes, while no changeovers occur if the storage tank sizes are at least 1.5 times the daily average volumes for the stable case.

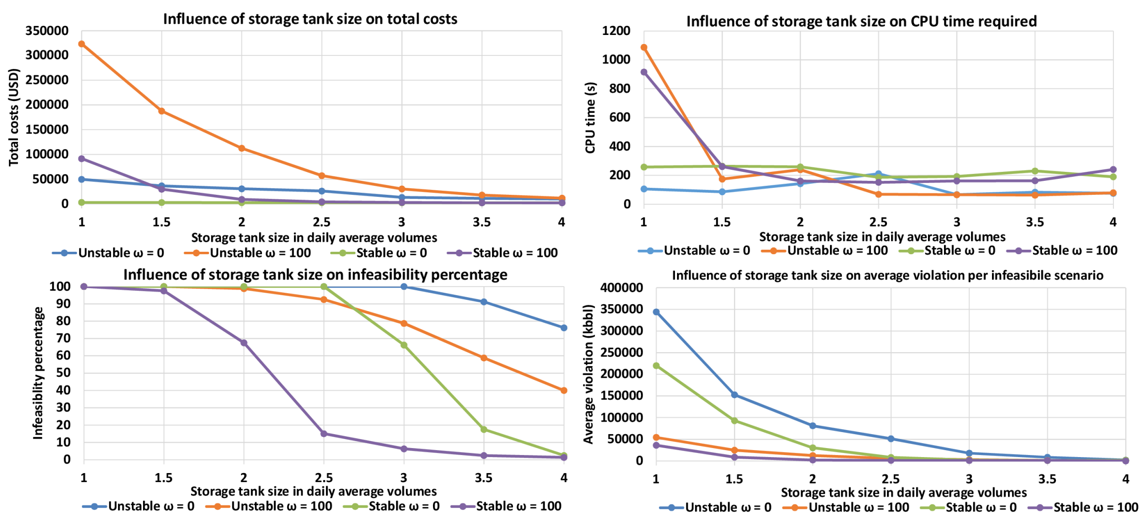

Figure 25: Storage tank volumes based on average daily volume experiment results

Appendix K Sensitivity analysis: peak efficiency throughput as a percentage of the pipeline capacity

We solve the model for both cases, for $\omega = 0$ and $\omega = 100$, and for the mentioned values of the peak efficiency throughput. Figure 26 shows the energy costs, CPU time, infeasibility percentage, and average violation per infeasible scenario. The energy costs figure looks the same as in Figure 20 with the only difference that the energy costs for 40% and 50% are much higher, which is because the average utilization of a pipeline is more than 30% and 20% above the peak efficiency throughput. It is important to note that the average utilization of the different pipelines in experiments of Section 7.3.1 varies from 36% to 81%, so in case the peak efficiency throughput as a percentage of pipeline capacity is increased from 40% to 80% more pipelines have an average throughput rate that is below the peak efficiency throughput. The figures for infeasibility percentage and the average violation per infeasible scenario for $\omega = 100$ are similar to Figure 20, because fluctuating in throughput rate becomes cheaper as the peak efficiency throughput increases, which allows that more infeasibilities are avoided for $\omega = 100$. In Section 7.3.1 we concluded that fluctuations in throughput can have both a positive and negative effect on model robustness when $\omega = 0$. Figure 26 shows that the effect in this case is positive for the unstable case as both the infeasibility percentage as the average violation per infeasible scenario decreases, while the effect is neutral for the stable case. The CPU time decreases as the peak efficiency throughput is increased for the stable case with ω=100, which might be caused by the fact that the non-linear part of the energy cost function is approximated in less time periods. For the other experiments the CPU time is more or less stable. Just as Section 7.3.1 we can conclude that also in this case the peak efficiency throughput has a high influence on the model as it influences both total costs and the solution in terms of throughput rates. The effect on model robustness in the experiments with $\omega = 100$ is again small.

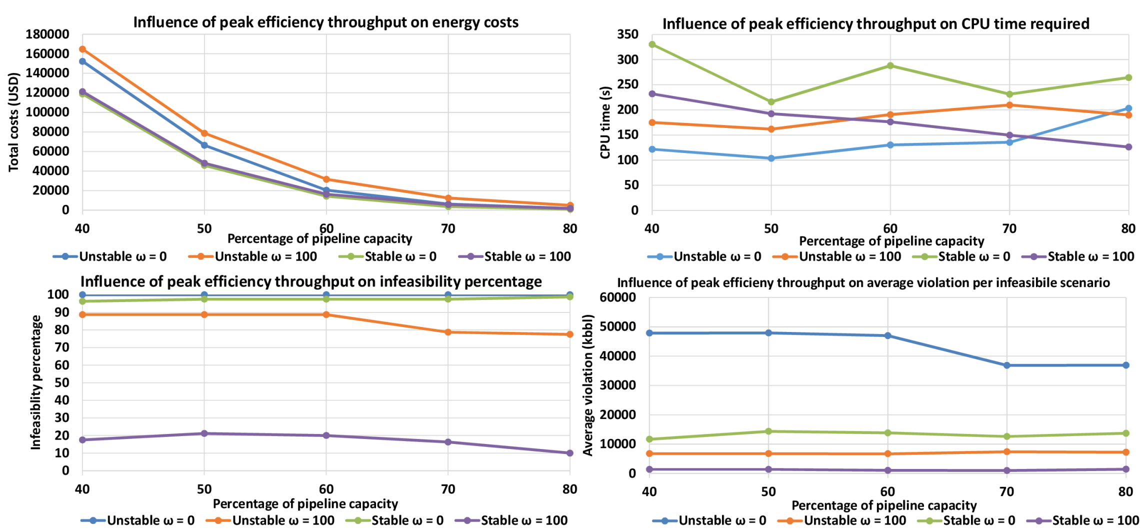

Figure 26: Peak efficient flow as a percentage of the pipeline capacity results

Appendix L Terminal shortage cost factor sensitivity analysis results

Case	Omega	Terminal shortage cost factor	Total costs	Terminal shortage costs	Prod. plan deviation penalties	Changeover costs	Energy costs
Stable	0	20	2,405	-	-	-	2,376
Stable	0	10	2,405	-	-	-	2,376
Stable	0	4	2,405	-	-	-	2,376
Stable	0	2	2,405	-	-	-	2,376
Stable	0	0.8	2,405	-	-	-	2,376
Stable	0	0.4	2,334	140	-	-	2,165
Stable	100	20	4,222	-	-	-	4,180
Stable	100	10	4,222	-	-	-	4,180
Stable	100	4	4,221	30	-	-	4,149
Stable	100	2	4,217	33	-	-	4,141
Stable	100	0.8	4,301	238	-	-	4,020
Stable	100	0.4	4,465	899	-	-	3,518
Unstable	0	20	25,216	90	2,883	10,000	12,083
Unstable	0	10	25,102	203	2,883	10,000	11,903
Unstable	0	4	24,817	542	2,883	10,000	11,281
Unstable	0	2	24,260	7,670	-	-	16,371
Unstable	0	0.8	17,884	6,293	-	-	11,365
Unstable	0	0.4	13,700	5,263	-	-	8,222
Unstable	100	20	56,557	21,971	2,174	10,000	22,318
Unstable	100	10	56,633	23,280	2,037	10,000	21,223
Unstable	100	4	52,340	29,975	-	-	22,188
Unstable	100	2	41,960	23,682	-	-	18,099
Unstable	100	0.8	29,959	15,417	-	-	14,363
Unstable	100	0.4	22,825	10,493	-	-	12,164

Table 11: Terminal shortage cost factor sensitivity analysis results

Appendix M Refinery shortage cost factor sensitivity analysis results

Case	Omega	Refinery shortage cost factor	Total costs	Refinery shortage costs	Prod. plan deviation penalties	Changeover costs	Energy costs
Stable	0	40	2,405	-	-	-	2,376
Stable	0	20	2,405	-	-	-	2,376
Stable	0	8	2,405	-	-	-	2,376
Stable	0	4	2,405	-	-	-	2,376
Stable	0	1.6	2,405	-	-	-	2,376
Stable	0	0.8	2,405	-	-	-	2,376
Stable	100	40	4,222	-	-	-	4,180
Stable	100	20	4,222	-	-	-	4,180
Stable	100	8	4,222	-	-	-	4,180
Stable	100	4	4,222	-	-	-	4,180
Stable	100	1.6	4,222	29	-	-	4,151
Stable	100	0.8	4,243	43	-	-	4,157
Unstable	0	40	25,216	-	2,883	10,000	12,083
Unstable	0	20	25,216	-	2,883	10,000	12,083
Unstable	0	8	25,216	-	2,883	10,000	12,083
Unstable	0	4	25,216	-	2,883	10,000	12,083
Unstable	0	1.6	25,193	43	2,883	10,000	12,048
Unstable	0	0.8	25,129	136	2,883	10,000	11,997
Unstable	100	40	56,557	-	2,174	10,000	22,318
Unstable	100	20	56,557	-	2,174	10,000	22,318
Unstable	100	8	56,557	-	2,174	10,000	22,318
Unstable	100	4	59,979	8,807	2,104	10,000	22,286
Unstable	100	1.6	59,167	12,895	-	-	29,321
Unstable	100	0.8	54,814	11,997	-	-	26,012

Table 12: Refinery shortage cost factor sensitivity analysis results

Appendix N Changeover cost factor sensitivity analysis results

Case	Omega	Changeover cost factor	Total costs	Changeover costs	Prod. plan deviation penalties	Refinery shortage costs	Energy costs
Stable	0	10000	2,405	-	-	-	2,376
Stable	0	5000	2,405	-	-	-	2,376
Stable	0	2000	2,405	-	-	-	2,376
Stable	0	1000	2,405	-	-	-	2,376
Stable	0	400	2,405	-	-	-	2,376
Stable	0	200	2,405	-	-	-	2,376
Stable	100	10000	4,222	-	-	-	4,180
Stable	100	5000	4,222	-	-	-	4,180
Stable	100	2000	4,222	-	-	-	4,180
Stable	100	1000	4,222	-	-	-	4,180
Stable	100	400	4,222	-	-	-	4,180
Stable	100	200	4,222	-	-	-	4,180
Unstable	0	10000	25,216	10,000	2,883	-	12,083
Unstable	0	5000	20,216	5,000	2,883	90	12,083
Unstable	0	2000	17,216	2,000	2,883	90	12,083
Unstable	0	1000	15,715	2,000	1,341	108	12,140
Unstable	0	400	14,515	800	1,341	108	12,140
Unstable	0	200	14,057	800	1,559	95	11,581
Unstable	100	10000	56,557	10,000	2,174	-	22,318
Unstable	100	5000	51,557	5,000	2,174	73	22,318
Unstable	100	2000	48,557	2,000	2,174	73	22,318
Unstable	100	1000	47,313	2,000	1,194	80	22,049
Unstable	100	400	46,113	800	1,194	80	22,049
Unstable	100	200	45,713	400	1,194	80	22,049

Table 13: Changeover cost factor sensitivity analysis results

Appendix O Production plan deviation penalty factor sensitivity analysis results

Case	Omega	Prod. plan deviation penalty factor	Total costs	Prod. plan deviation penalties	Energy costs	Changeover costs
Stable	0	1	2,405	-	2,376	-
Stable	0	0.5	2,405	-	2,376	-
Stable	0	0.2	2,405	-	2,376	-
Stable	0	0.1	2,405	-	2,376	-
Stable	0	0.04	2,405	-	2,376	-
Stable	0	0.02	2,405	-	2,376	-
Stable	100	1	4,222	-	4,180	-
Stable	100	0.5	4,222	-	4,180	-
Stable	100	0.2	4,222	-	4,180	-
Stable	100	0.1	4,222	-	4,180	-
Stable	100	0.04	4,222	-	4,180	-
Stable	100	0.02	4,222	-	4,180	-
Unstable	0	1	25,216	2,883	12,083	10,000
Unstable	0	0.5	23,775	1,442	12,083	10,000
Unstable	0	0.2	22,910	577	12,083	10,000
Unstable	0	0.1	22,622	288	12,083	10,000
Unstable	0	0.04	22,449	115	12,083	10,000
Unstable	0	0.02	22,391	58	12,083	10,000
Unstable	100	1	56,557	2,174	22,318	10,000
Unstable	100	0.5	55,350	1,209	22,079	10,000
Unstable	100	0.2	54,611	577	21,986	10,000
Unstable	100	0.1	54,323	288	21,986	10,000
Unstable	100	0.04	54,150	115	21,986	10,000
Unstable	100	0.02	54,093	58	21,986	10,000

Table 14: Production plan deviation penalty factor sensitivity analysis results

Appendix P Inventory cost factor sensitivity analysis results

Case	Omega	Inventory cost factor	Total costs	Inventory costs	Safety stock penalties	Energy costs
Stable	0	0.0001	2,405	27	2	2,376
Stable	0	0.0002	2,433	54	3	2,376
Stable	0	0.0005	2,519	135	8	2,376
Stable	0	0.001	2,663	270	17	2,376
Stable	0	0.0025	3,091	676	36	2,379
Stable	0	0.005	3,802	1,352	67	2,383
Stable	100	0.0001	4,222	27	16	4,180
Stable	100	0.0002	4,264	54	31	4,179
Stable	100	0.0005	4,392	135	78	4,179
Stable	100	0.001	4,597	270	151	4,176
Stable	100	0.0025	5,205	676	358	4,171
Stable	100	0.005	6,177	1,352	666	4,160
Unstable	0	0.0001	25,216	23	90	12,083
Unstable	0	0.0002	25,328	45	179	12,084
Unstable	0	0.0005	25,661	113	433	12,095
Unstable	0	0.001	26,193	226	819	12,128
Unstable	0	0.0025	27,690	564	1,843	12,277
Unstable	0	0.005	29,997	1,128	3,249	12,634
Unstable	100	0.0001	56,557	21	73	22,318
Unstable	100	0.0002	56,649	42	144	22,318
Unstable	100	0.0005	56,917	105	353	22,314
Unstable	100	0.001	57,372	210	702	22,316
Unstable	100	0.0025	58,681	535	1,526	22,295
Unstable	100	0.005	60,456	1,129	2,268	22,205

Table 15: Inventory cost factor sensitivity analysis results