



Norwegian University of
Science and Technology

Use of Clusters in a Route Generation Heuristic for Distribution of Fish Feed

Marius Gyberg Haugland
Sondre Thygesen

Marine Technology

Submission date: June 2017

Supervisor: Bjørn Egil Asbjørnslett, IMT

Co-supervisor: Inge Norstad, Marintek

Norwegian University of Science and Technology
Department of Marine Technology

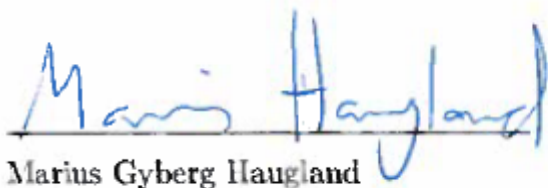
Preface

This master thesis is the finalization of our Master of Science degree at the Norwegian University of Science and Technology, Trondheim. The paper was written during the spring 2017, and it reflects aspects we have been focusing on during our specialization program in Marine Systems Design and Logistics.

We would like to thank all of those involved in the work of making this thesis. First of all, our supervisor Professor Bjørn Egil Asbjørnslett, and academic co-supervisor Inge Norstad from SINTEF Ocean. Your help and guidance have been valuable and greatly appreciated, especially for the expertise needed for making the route generation algorithm. Also the expertise of Professor Kjetil Fagerholt were valuable in the process of creating the essential distance matrix. We would also like to thank Ivar Ulvan at Egil Ulvan Rederi for providing us with information about relevant ship characteristics, and Claes Jonermark at Marine Harvest for information regarding the fish feed distribution industry. The meetings and email correspondence have been of great help to us for understanding the complexity of the system, and how it should be modeled.

A final thank you goes to our good classmates in office C1.084, Department of Marine Technology, Tyholt, for providing motivation, inspiration and good spirit throughout this final semester.

Trondheim 07.06.2017


Marius Gyberg Haugland


Sondre Thygesen

This page is intentionally left blank

Summary

Fish farming is an industry in growth which is also expanding its area of operation. Among the biggest cost-contributors are costs associated with the distribution of fish feed. The demand is heavily varied and the system is hard to model because of uncertainties and the complex problem structure. With the help of simplifications and problem reduction techniques, we have created a tool that can be used for minimizing operational costs in a desired system, and also for evaluating fleet-compositions in different demand scenarios.

An optimization model was created for the purpose of distributing a single cargo to a large number of customers from a given production location, using a pre-defined fleet. The objective function is based on fuel price, sailing distance, and specific fuel consumption for each ship. The model is based on a two-phased heuristic, that first creates all possible routing alternatives based on information in an input file, and then chooses the best combination of routes for a final solution which satisfy all requirements of the system. To be able to solve the model, a planning horizon must first be established. We categorize our routing problem as an operational decision problem, hence a planning period of around one week is suitable. For this short time period, variables such as demand can be treated as constant, hence the use of a time-continuous model can be used. To lower the calculation time to a practical level for the tool to be used as an operational planning tool, the number of customer nodes is reduced with clustering techniques and compatibility constraints. Three different methods are developed and compared for various dimensions and constraints; a regional clustering method, natural selection of clusters based on location, and a theoretical approach called k-means which are clustering the nodes based on distance between locations.

A computational study was performed to determine the best trade-off between cluster size and calculation time. One system configuration was chosen for each clustering method, and the techniques were compared in three different demand scenarios; low, medium, and high. The regional method included 11 zones, the natural method had 23 clusters with full compatibility between customers and ships, while the k-means approach included 30 clusters, but with restrictions of which customer each ship can visit. The regional method ended up being too coarse for consideration, meaning that the clusters became too large to have any effect of being optimized. In most cases, the natural method turned up to be less costly than the k-means approach, showing us that a fleet with full freedom to be optimized may be more beneficial than a fleet with constrained routing-options as are used today.

Finally the best clustering approach and system dimensions were used for evaluating Marine Harvest's current fleet composition in the same demand scenarios as before. Four alternative fleet configurations were compared to the original fleet, and one stood out as less costly in two out of the scenarios. This new composition is in accordance with the company's current plan of replacing one of the vessels in the fleet with a larger and faster ship. We see this as a validation that our model can be used for realistic and valuable evaluations, in addition to the main purpose of minimizing the operational costs.

This page is intentionally left blank

Sammendrag

Oppdrettsnæringen er i rask vekst og utvider også sitt operasjonsområde. Blant de største kostnadene er utgifter relatert til fordeling av fiskefôr. Det er stor variasjon i etterspørsel, og systemet er vanskelig å modellere på grunn av usikkerheter og en kompleks problemstruktur. Ved bruk av forenklinger og metoder for problemreduksjon har vi laget et verktøy som kan brukes til å minimere driftskostnadene i et ønsket system, og for å vurdere flåtesammensetninger i ulike etterspørsel-scenarier.

En modell for å optimere fordelingen av et enkelt produkt fra fabrikk til et stort antall kunder ved bruk av en gitt flåtestørrelse er utarbeidet. Målfunksjonen er basert på drivstoffpris, seildistanse og spesifikt drivstofforbruk for hvert enkelt skip. Modellen er basert på en to-steps heuristikk, som først lager en liste med alle mulige ruter basert på informasjon i en input-fil, og deretter setter sammen den beste kombinasjon av disse til en løsning som tilfredsstillende alle krav for systemet. For å kunne løse modellen må det etableres en planleggingshorisont. Siden vi har valgt å undersøke et operasjonelt problem er en planleggingsperiode på rundt en uke egnet. For en kort periode som dette kan variabler som etterspørsel bli sett på som tilnærmet konstant, og vi benytter derfor en tids-kontinuerlig modell. For å redusere beregningstiden til et nivå som er praktisk for et driftsplanleggingsverktøy, reduseres antall noder i problemet ved bruk av klustere og kompabilitet. Tre forskjellige metoder er utviklet og sammenlignet; en regional klusteringsmetode, naturlig klustering basert på lokasjon, og en teoretisk tilnærming kalt k-means som samler noder basert på avstand mellom oppdrettsanlegg.

En studie ble utført for å kunne bestemme det beste kompromisset mellom problemstørrelse og beregningstid. En systemkonfigurasjon ble valgt for hver av de tre metodene, og deretter sammenlignet i tre ulike etterspørsel-scenarier; lav, middels og høy. Metoden basert på regioner inneholdt 11 soner, den naturlige klusteringmetoden ble delt inn i 23 klustre, med full kompabilitet mellom kundene og skipene, mens k-means metoden ble oppdelt i 30 klustre med begrensninger på hvilke kunder hvert skip kan besøke. Den regionale metoden viste seg å være for grovt klustret, noe som innebar at klusterne var for store til å få noen effekt av å optimalisere rutingen. I de fleste tilfeller viste den naturlige klusteringsmetoden seg å ha lavere kostnader enn k-means metoden, noe som viser at en flåte uten restriksjoner kan være mer gunstig enn en flåte med begrensede rutingsalternativer som brukes i dag.

Tilslutt ble den beste klusteringsmetoden og dens tilhørende egenskaper brukt for å evaluere Marine Harvests nåværende flåtesammensetning i de samme etterspørsel-scenariene som tidligere. Her ble fire ulike flåtekonfigurasjoner sammenlignet med den opprinnelige flåten, og en skilte seg ut som billigere i to av scenariene. Denne endringen av flåtesammensetning er i samsvar med selskapets nåværende strategi om å erstatte et av de eldre fartøyene i flåten med et større og raskere skip. Dette ser vi på som en bekreftelse på at vår modell kan brukes for realistiske evalueringer, i tillegg til å minimere driftskostnadene i distribusjonen som var hovedformålet med oppgaven.

This page is intentionally left blank

Contents

Preface	i
Summary	iii
Sammendrag	v
1 Introduction	1
2 Problem Description	4
2.1 Problem Background	4
2.2 Problem Objective	6
2.3 Problem Scope and Limitations	6
3 The Fish Feed Industry	8
3.1 Production of Fish Feed	8
3.2 External Supply vs In-house Production	10
3.3 System Description	11
3.3.1 Structure	11
3.3.2 Feed Consumption and Variations	13
3.3.3 Planning the Distribution	15
3.4 Challenges in the Feed Distribution Today	16
3.5 Use of Multiple Depots	17
4 Related Literature	19
4.1 Comparison of Land-based and Maritime Transportation	19
4.2 Types of Transportation Problems	20
4.2.1 Vehicle Routing Problems (VRP)	20
4.2.2 Inventory Routing Problems (IRP)	21
4.3 Solution Methods	22
4.3.1 Exact methods	23
4.3.2 Heuristic methods	24
5 Optimization Model	28
5.1 Model Characteristics	28
5.2 Simplifications and Applications	31

5.3	Route Generation Algorithm	33
5.4	Optimization Algorithm	35
5.4.1	Mathematical Model Formulation	35
5.4.2	Xpress Modeling	38
6	Cluster Calculations	39
6.1	Cluster Generation	39
6.1.1	Theoretical Clustering Methods	39
6.1.2	Manual Clustering Methods	40
6.2	Selection of Clustering Methods	42
6.2.1	Contamination Zones	42
6.2.2	Natural Clustering Method	43
6.2.3	The K-means Algorithm	44
6.3	Parameter Estimations in the Clusters	45
6.3.1	Distance Estimations	45
6.3.2	Estimation of Service Time	47
7	Input Data	50
7.1	System Dimensions	50
7.2	Cluster Information	51
7.3	Distance Matrices	55
7.4	Cost Calculations	55
8	Computational Study	57
8.1	Planning Levels	57
8.2	Test Cases	58
8.3	Adjustments and Possible Improvements of the Model	62
8.4	Determination of Cluster Sizes	64
8.5	Manipulation of Compatibility Matrix	65
8.6	Consumption Scenarios	67
8.7	Results	75
8.7.1	Selection of Final Clustering Method and Characteristics	75
8.7.2	Evaluation of Fleet Configuration	76
9	Discussion	81

10 Conclusion	87
11 Further Work	89
References	91
A Explanation of Attached Files	95
B List of Marine Harvest’s Farm Locations	96
C Final Mathematical Model Formulation	99
D Optimization Outputs	100
D.1 Medium Demand Results	100
D.2 High Demand Results	103
D.3 Low Demand Results	105
E Problem Description Form	108

List of Figures

1	Amount and value of sold salmon in Norway (SSB, 2016).	1
2	Cost components in the fish farming industry.	2
3	Overview of Marine Harvest’s fish farm locations.	5
4	Production facilities at Marine Harvest’s factory in Bjugn (Marine Harvest, n.d.).	9
5	Typical fish feed content.	10
6	Annual temperatures of seawater.	14
7	Annual consumption of fish feed relative to fish biomass (Marine Harvest, 2016).	14
8	Illustration of a fish feed distribution system including multiple depots.	17
9	Model setup. Dotted boxes represent data files, and fixed boxes models.	28
10	Illustration of the route generation method.	34
11	Dendrogram displaying the hierarchical clustering method.	40
12	Example of large clusters based on geographical regions.	41
13	Proposed farming zones based on contamination-risk (Ådlandsvik, 2015).	41
14	Example of the natural clustering technique with 18 clusters.	43
15	Example of clustering using k-means algorithm with 10 clusters.	44
16	Example of route generated by Searoutes.com (2017).	46
17	Comparison of route generated by Searoutes.com (2017) (blue) and MATLAB’s mapping toolbox (red).	47
18	Illustration of the internal sailing time estimation method.	48
19	Illustration of the LNG ship ”With Harvest” (<i>ilaks.no, 2014</i>).	51
20	Definition of consumption zones with all Marine Harvest’s farms.	53
21	Test case 1 solution output.	59
22	Test case 2 solution output.	60
23	Overview over selected clustering methods.	67
24	Routing illustration for k-means method with 30 clusters and medium demand. .	69
25	Example of error with routes generated by Searoutes.com (2017).	84

List of Tables

1	Definitions for the model formulation.	36
2	Attributes of the ships used in the system.	50
3	Test-system dimensions and generation times.	58
4	Calculation times for natural clustering.	65
5	Calculation times for k-means clustering.	65
6	Calculation times for natural clustering with compatibility restrictions.	66
7	Calculation times for k-means clustering with compatibility restrictions.	66
8	Consumption scenarios based on consumption in zones.	68
9	Distribution costs for the medium demand scenario.	68
10	Distribution costs for the high demand scenario.	71
11	Distribution costs for the low demand scenario.	73
12	Fleet alternative 1.	77
13	Fleet alternative 2.	77
14	Fleet alternative 3.	78
15	Fleet alternative 4.	78
16	Comparison of the cost between the current fleet and new fleet configurations.	78

1 Introduction

The production and value of fish farming in Norway has never been larger than it is today. Salmon farming accounts for 95% of the total industry, and as of 2015 the value of farmed salmon in Norway was 45 billion NOK. As viewed in Figure 1, the industry is in rapid growth and it is expected to keep growing. One of the reasons why this industry is so attractive and is expected to keep growing can be seen from the feed conversion ratio. This is consumed feed divided by increase in biomass and hence measures how much feed is needed to increase the biomass. For a salmon, this ratio is close to one, which is superior to any other types of meat.

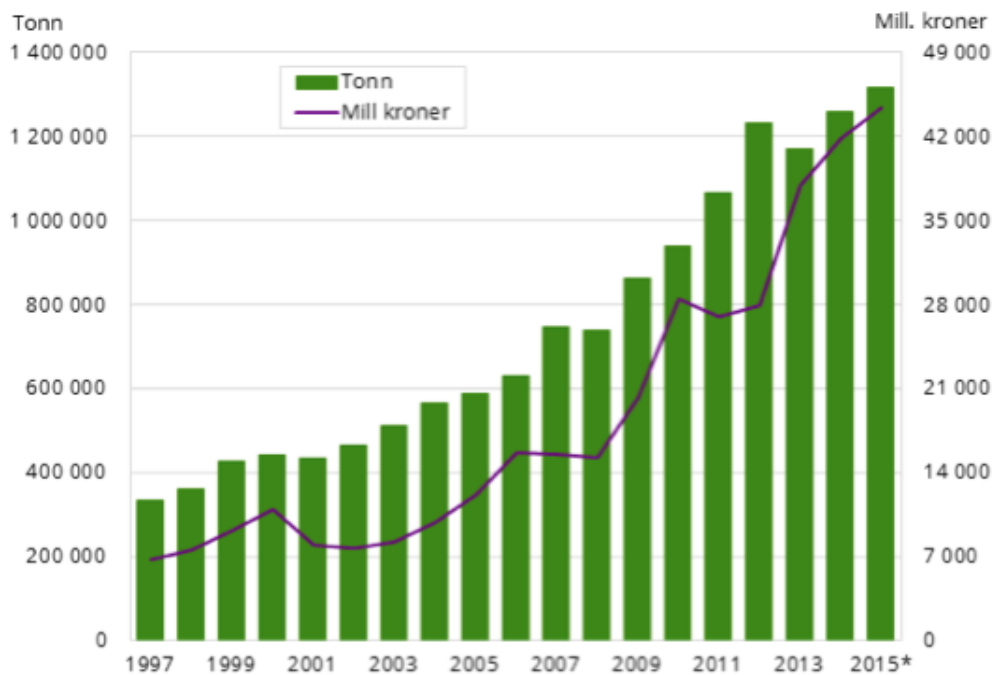


Figure 1: Amount and value of sold salmon in Norway (SSB, 2016).

Among the biggest expenses in the fish farming industry are costs associated with fish feed, which accounts for approximately half of the salmon production cost (Marine Harvest, 2016). To be able to keep the growing industry as profitable as possible, the importance of efficient scheduling and planning for the distribution of this large cost-contributor cannot be underestimated.

The figure below illustrates the main cost-components and their relative importance in the fish farming industry today. We see that feed-related costs are by far the biggest contributor and we hence see a big market opportunity here. If we are able to lower this cost-component, the total salmon production cost can also be lowered. Costs from importation of raw materials for the feed, and from the production process itself, are somewhat fixed, but a major part of this cost-component comes from the distribution of the feed. Sometimes the supply is not sufficient to meet all the different demands in the system, or unexpected delays appear, forcing the supplier to prioritize among customers. In worst case this could lead to a halt in the growth process for the fish, as it does not get the feed it needs to be able to follow the optimal growth process.

This will cause a drop in profit as the fish does not reach fully grown state at the optimal time, and the harvest has to be postponed. This shows the importance of well planned and robust delivery-schedules. By robust planning and flexibility in the schedules for the fleet, costs can be reduced by minimizing delays in deliveries, along with cost-savings related to fleet utilization and time utilization. Another aspect to take into account, is that the industry is moving offshore. This magnifies the problem of efficient routing of the fleet due to the longer sailing distance and higher operational costs.

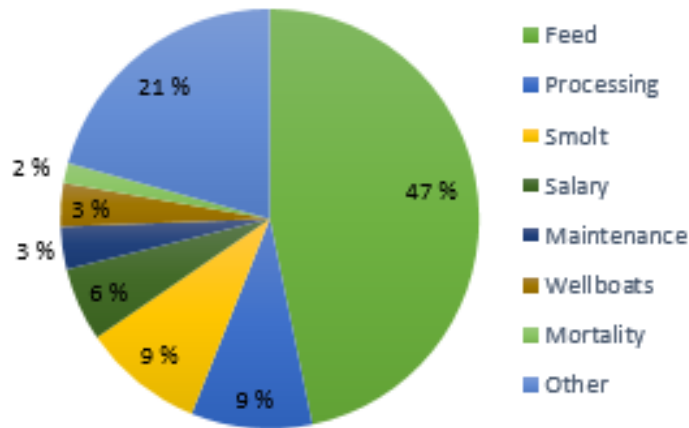


Figure 2: Cost components in the fish farming industry.

In order to create cost-effective and robust transportation plans there exists a variety of tools and models that can be used in the route and scheduling management. In the fish feed industry today, the scheduling is normally done by hand with daily changes implemented manually, but theoretically the routing can be done in a more optimal way to minimize desired parameters. In general, transportation problems are often different variants of so-called vehicle routing problems (VRP). According to Christiansen, Fagerholt, Nygreen, and Ronen (2013), the quantity of published research on problems related to ship routing and scheduling has been almost doubling every decade. They have about every 10 years published a survey of research on the field (see Christiansen et al., 2004, 2012; Ronen, 1993, 1983) for the purpose of providing a comprehensive source of published research to be easy accessed for researchers and students of this domain. These surveys show that an increasing focus on routing optimization in the maritime industry is necessary to remain competitive. Among the largest challenges of optimizing the distribution in the fish feed industry is the size and complexity of the system. The more complex and the more degrees of freedom the problem has, the computationally harder the problem is to solve. A large set of costumers increases this effects due to the great number of possible routes for the ship to sail. To deal with this in a general optimization problem, different algorithms has been created to more easily obtain a possible solution, so-called heuristics.

One widely used heuristic in maritime transportation problems is the route generation method. This heuristic is especially applicable for solving these kinds of problems, since the number of

feasible routes is usually not that enormous. Maritime problems are also often more tightly constrained, which reduces the number of feasible routes drastically compared to land-based problems. This approach is used by Wang (2014), where the use of this method for an integrated operational transportation planning problem for a heterogeneous fleet is investigated, much like the distribution problem faced in this paper. The objective for the problem is to minimize the total costs, including the fixed costs of the own fleet, the total routing costs, and charges for outsourcing requests. The result is a set of candidate routes of different fulfillment modes, and requests to be outsourced, in such a way that the total costs are minimized.

The overall objective of this paper is to develop a tool that can be used for minimizing the operational costs for the fish feed distribution process, while at the same time meeting all requirements and demands of the system at all times. Only operational costs are considered, since fixed costs related to ships are not affected by the decisions to be made. The final model should help the distribution planner to reduce operational costs in the feed distribution, and also reduce the time used for planning the routing of the fleet. In addition, the tool should be created in a way that it can be used for comparing different fleet configurations for evaluating the best way to distribute cargo in different systems.

The remainder of this paper is organized as follows: in Chapter 2 we start by presenting the background for this paper, and we give a detailed description of the objective and the scope of the problem we are looking at. Chapter 3 describes the fish feed industry in detail, how the fodder is made, the variation in consumption of fodder at the fish farms, how it is distributed and planned today, and also challenges in how the logistics are planned in today's industry. Then in Chapter 4 we perform a thorough review of similar work that has been performed over the years, looking at general transportation problem types, and known solution methods. After this, we start the work of developing our optimization model in Chapter 5, where we describe a two-phased model approach, along with the mathematical formula used for the optimization. Chapter 6 includes a discussion of how we can reduce the problem size by generating clusters of customer nodes, before Chapter 7 presents how we are modelling the real world system into an input data file. An extensive computational study is performed in Chapter 8, starting with an exploration of test cases to familiarize with the process, before we evaluate different modeling techniques and system configurations. Finally, the best dimensions are determined and used in a real world system evaluation. In the end, we discuss the discovered results and sources of error in Chapter 9, before making a final conclusion in Chapter 10, and suggestions for further work in Chapter 11.

2 Problem Description

In general, it is necessary to create cost-effective and robust transportation plans that ensure in-time delivery of cargo to customers. This paper will describe the challenges of producing these plans for distribution of fish feed from factory to a large number of fish farms, and why this is a complicated process to optimize. We will explore possible ways to reduce the problem size and complexity, and investigate possible methods of how to most optimal route a fixed fleet of ships to meet all requirements and demands in the system. Finally, we will test if Marine Harvest's current fleet is the optimal fleet for serving their farms along the Norwegian coast internally in different demand scenarios, and provide recommendation to improve the fleet composition. In Section 2.1 we present the background for the problem based on Marine Harvest's current distribution network along with general challenges in the industry today. In Section 2.2 we state the objective of this paper and what we want to accomplish, before giving a more detailed description of the scope of the thesis in Section 2.3.

2.1 Problem Background

The problem we are facing in this thesis is motivated by a real logistics problem that Marine Harvest is currently dealing with. The system they are serving are both large and complex, and hence the distribution planning is both an uncertain and time-consuming process. They have seen problems in how the planning process is performed today, as everything is done manually in large spreadsheets and is extensively based on experience. To stay competitive, and to more efficiently plan the routing for minimizing operational costs, a more automated process is desired.

The company points to several reasons why this system is a complex problem to optimize, mainly due to variations and uncertainties in demand. In Marine Harvest's case, planning for their around 150 fish farms located along the Norwegian coast, shown in Figure 3, amplifies the challenge of planning. Today they are using four vessels to perform the deliveries, where each ship supplies a given set of farms based on different restrictions. These restrictions include the size and location of the farms, the delivery possibilities at the farms, and the speed and capacities of each vessel. There are also various types of products and limited storage capacity at the factory, which means that the yearly production capacity cannot be fully utilized. Some farms contain smolt and smaller fish and hence need a certain type of feed, which complicates the problem. The demand of feed at each fish farm is determined by several factors such as fish size and water temperature, which is constantly changing, and hence the production schedule and delivery plans have to be changed accordingly. This means that both storage capacities and production/consumption rates constrain the scheduling of deliveries. Also, the decision of which fish farms to serve from the factory, and for which to order externally, must be taken several weeks in advance. Further complications of the planning problem comes from the possibility of delivering each customer just a fraction of their demand.

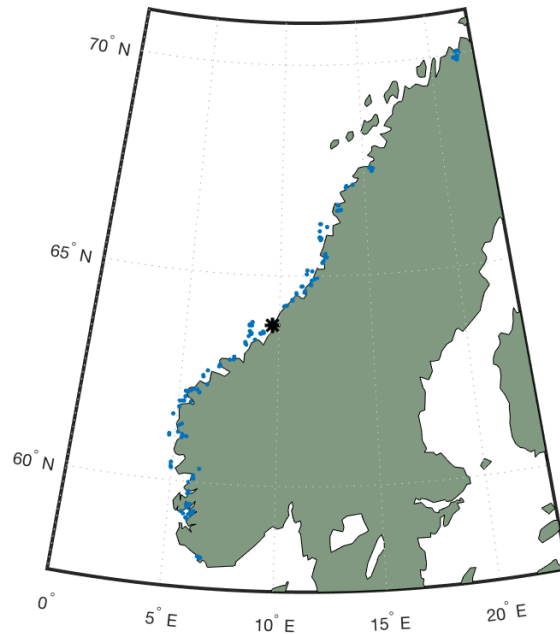


Figure 3: Overview of Marine Harvest's fish farm locations.

As we will describe in Section 3.3.3, the way the distribution is planned today is based on regular communication between the fish farm managers and the logistics planner. This is necessary because the final routing schedule is set only a few days prior to the vessel's departure from the factory. A feed delivery plan must consist of both the routing and scheduling of all the ships in the fleet, and the loading and unloading quantities of each feed type at the factory and fish farms. There is also a problem that fish farms do not keep their numbers up to date and mistakes are made due to human errors. The lack of planning opportunities and constant order changes, makes it difficult to optimize the routing, and find an optimal delivery schedule. Because of this, tools for better data-reporting from the farms and demand forecast are desired. The main problem comes from unexpected delays in the feed supply, or unexpected changes in demand that causes changes in the production schedule. This again leads to increased expenses due to the additional time needed for the fish to reach the ideal harvesting weight. The planning of routing and scheduling for the ships needs therefore to be robust, but also flexible. To create and run optimization models reflecting this large and complex system is almost impossible, and simplifications must be applied. Based on this, an operation support tool that is easy to alternate and that gives quick and usable results, is desired to help with the planning.

In addition, the company operates a fixed fleet of vessels, and this should be utilized the best way possible to reduce costs. Because of seasonal variation, there will be a high season and a low season with regard to demand. This means that the need for cargo capacity is drastically changing throughout the year, and the fleet size may not always be optimal. The tool should therefore also be able to evaluate different fleet configurations for different demand scenarios.

2.2 Problem Objective

The objective is to create a method for minimizing the operational costs for the fish feed distribution process in a more efficient and practical way than it is done today. This should be done while meeting all requirements and demands of the system at all times. Only operational costs are considered, since fixed costs related to ships are not affected by the decisions to be made. However, the possibility of external supply at a fixed rate for serving a customer is included to increase the feasibility area of the problem. The final model should help the distribution planner to reduce operation costs in the feed distribution, and also reduce the time used for planning the ship scheduling.

In addition, we want to use the model for investigation of the current fleet composition used by Marine Harvest. Today, their fleet consists of four vessels, where the two newest ships are specialized for the fish feed distribution. The company are considering replacing one of the older and slower ships in their fleet because of the higher operational cost associated with the vessel. We want to use the optimization model to test how their current fleet is performing in different demand situations compared to other fleet composition alternatives. Based on the results, we will provide recommendations to how the fleet possibly could be adjusted to improve the feed distribution, along with reducing the operational costs.

2.3 Problem Scope and Limitations

The overall task in this thesis is to provide a method of planning feed deliveries to a large number of costumers from a production facility in order to improve the efficiency of today's planning process. The decisions to be made, are which fish farms each ship should visit, and how much cargo to deliver to each fish farm for the lowest operational cost possible. The distribution problem must be considered on several planning levels when it comes to the determine the optimal delivery plan. Since decisions must be made rapidly, we categorize our routing problem as an operational decision problem, but the modeling of the system should be taken on a higher planning level.

Many parameters in the system are heavily variant, and the system is hard to model because of uncertainties and the complex problem structure. With the help of simplifications and problem reduction techniques, we want to reduce the complexity to a reasonable level so that the calculation time is lowered to a practical level for the tool to be used as an operational planning tool. To be able to do this, variables in the system must be treated as constants within the planning horizon, meaning that a relatively short planning period must be selected.

To be able to optimize Marine Harvest's real world distribution problem, we need information about their current distribution system. This includes information about ships used in the system with associated characteristics, information regarding all of Marine Harvest's fish farms, and the interaction between them. In addition, an understanding of the industry with its challenges and uncertainties are needed, along with general theories about routing optimization. We will not look into the procurement of raw materials, production scheduling or factory storage aspects, as the focus is on operational performance. When all the relevant information is gathered, we will need to reduce the size and complexity of the problem to be able to optimize it in a practical amount of time. This will be done by including realistic constraints to the model, and by collecting customers into clusters. This is a problem reduction technique that can be performed in several ways, and we will compare different approaches.

Since most optimization problems are hard to solve exact, a heuristic approach will be applied. The route generation method is a widely used heuristic for solving marine transportation problems and it has good applications for our problem as well. This is because of its nice structure, where the number of feasible routes are drastically reduced due to the tightly constrained restrictions in the problem. The nice structure of the heuristic also makes it able to give a high-quality solution (not necessary the optimal), consisting of a combination of routes that meets all requirements. To be able to create the most user-friendly tool for the system, a computational study will be performed to determine the best trade-off between cluster size and calculation time. We will also compare the different clustering methods for different demand scenarios to determine a final model that are able to represent the real world as accurate as possible.

Finally, the best clustering approach will be used for evaluating Marine Harvest's current fleet composition in the same demand scenarios as before. We will compare the original fleet to multiple alternatives to investigate whether other compositions could be more profitable. If we discover any significant cost-saving results, it will be interesting to see if this new composition is in accordance with the company's current plan of replacing one of the vessels in the fleet with a larger and faster ship.

3 The Fish Feed Industry

This chapter will provide necessary information needed to understand the overall fish feed industry today, along with challenges related to the distribution process.

3.1 Production of Fish Feed

To get a better understanding of the finished product distributed, we will in this section look at fish feed itself, focusing on the production process and the content and nutrition in the fodder. The information in this section is fully based on Marine Harvest's current production facilities in Bjugn, but the same principles are used at other feed factories as well. The factory produces over 320 000 tonnes of fish feed per year and it is operational all day and night throughout all year. It has a storage capacity of 16 000 tonnes of dry raw materials, 10 000 tonnes of wet raw materials, and over 10 000 tonnes of finished product. Because fish in different growth stages requires different nutrition and different sizes of feed, different feed types are produced. The finished product has the shape of pellets, which varies in size from small, with a diameter of 0.6 millimeters, to larger ones of about 13 millimeters.

The production process starts with the import of raw materials. The needed raw materials are shipped in from all over the world and are divided into three categories; macro materials, micro materials, and oils. The macro materials consist of fish meal, barley grain, wheat grain, or other plant and animal ingredients, while the micro materials are vitamins and minerals, which are the most important and costly raw materials. After the raw materials arrives the factory, they are stored in giant silos. The feed production starts by measuring and weighting the raw materials to get the wanted composition of nutrient, before they are grinded and mixed in big blenders. Water is added and the mixture is extruded into pellets of the desired size. The pellets are then dried, before oil is added and then placed under specific temperature and pressure to achieve the wanted characteristics. Finally, the finished products are stored in silos or big bags, waiting to be loaded onto ships. An overview over the production process can be viewed in Figure 4 on the next page.

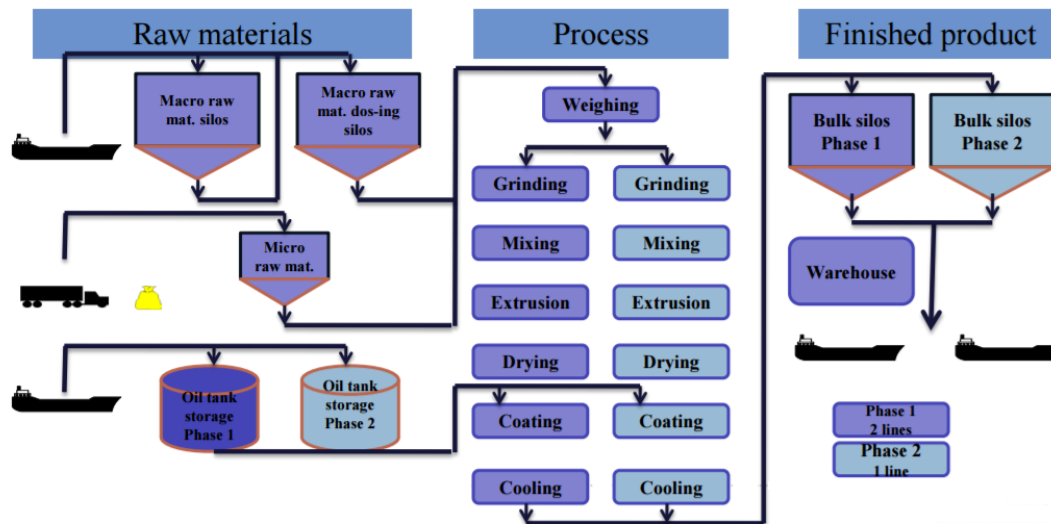


Figure 4: Production facilities at Marine Harvest's factory in Bjugn (Marine Harvest, n.d.).

There are several requirements for the finished product. First, the composition of the materials needs to be in such a way that the pellets can be transported without being destroyed. It is also important that the nutrients stay in the pellets and are not leaked out when the pellets enter the water. Lastly, the feed needs to sink at the right speed so that the fish are able to eat it before it hits the bottom of the cage. This also means that there is a requirement of taste so that the fish actually wants to eat it.

In general, there has been a major development in the fish feed content since fish farming industry started in the early 70's. In the beginning the farmers made their own fish feed consisting of fish waste, fish meal and shrimp-shell. Today, the feed is a complex and carefully planned mix of proteins, energy and essential nutrition to produce healthy and high-quality fish. To ensure a rapid and optimal growth, there are different kinds of recipes and compositions of feed depending on where in the growth stage the fish are. The feed consists mainly of fish meal, fish oil and vegetable oils. The marine materials are produced from small pelagic fish which are unattractive for humans, and trimmings from the fish processing industry. Because one kilo of industrial fish only gives around 3-5 % fish oil and 20-25 % fish meal, the exploitation of marine materials is not very effective. This has caused an increased lack of marine raw materials and hence increased the use of other vegetarian raw materials such as soy, sunflower, wheat, corn and rapeseed oil as Figure 5 shows (Marine Harvest, 2016).

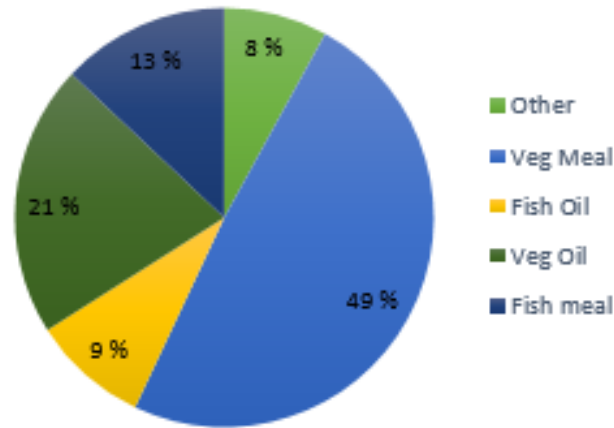


Figure 5: Typical fish feed content.

The fact that the nutrition in the fodder no longer exclusively comes from marine raw materials has raised the question of the nutrients in the fish that is consumed by humans. The fish require different amino acids, fatty acids, vitamins, and minerals to be as rich in nutrients as it normally is. These can in principle be provided by the different raw materials listed above, but this is not the natural source. Fish meal and other raw materials from animal origin have a more complete amino acid profile and generally a higher protein concentration compared to feed from vegetable origin. The fish normally gets their omega-3 fatty acids from the fish oil in the pellets they eat, but it is also possible that they can get the same omega-3s from plants like algae, so the fish product should contain the same amount of nutrients as wild fish (Craig and Helfrich, 2009).

3.2 External Supply vs In-house Production

It is a common discussion regarding the benefits of buying feed from an external supplier, or producing it in-house. As stated, costs related to feed are by far the biggest contributor to the total cost in the aquaculture industry. The fish feed market in Norway has in the later years mainly been controlled by three external global producers; Skretting, EWOS and BioMar. In 2014, Marine Harvest decided to take control over this major cost-contributor and supply their own farms with in-house produced feed.

There exist several arguments for a fish farmer to buy feed externally, as well as for producing the feed in-house. As in most production cases, the production of fish feed is most profitable when produced in large volumes. Companies specializing in fish feed are able to keep the unit production cost low, due to the large quantity produced. The low unit production cost also makes specialized feed companies capable of producing a large variety of feed types necessary in the aquaculture industry. These companies also have the the knowledge and competence to create feed pellets best suited for the different growth stages and for special situations, for example specialized feed needed during a disease breakout. Because there exist several specialized fish feed companies, the competition of providing the best feed is extensive. This competition benefits

the customer, as the investments in research and development increases, providing better feed types at lower costs.

For larger fish farming companies, in-house production of fish feed may be a better solution. In-house production only needs to produce feed for its own demand, and the variety in product types can be reduced, leading to a more efficient production. It can also reduce costs in the way that in-house production requires no sales or marketing costs, and also the logistics costs can be reduced. When producing and distributing own feed, the logistics planner only need to consider their own farms, making it easier to make a cheap and effective distribution plan. According to Marine Harvest, lost feed days was reduced when they changed from external supply to own-produced feed. The reduction in lost feed days was mainly due to more effectively communication between the farmers and the logistics planner, in addition to better feed demand forecasts.

3.3 System Description

This section will provide information about how the fish feed distribution system works today. We will look at which factors are causing decisions to be complex, and how they are dealt with in practice. This will give us important information for how to make simplifications to be able to create an optimization model for the real world system in the most realistic way possible.

3.3.1 Structure

It is important to have knowledge about the structure of the distribution network in order to fully understand the fish feed distribution system. It is possible to look at the network as simply consisting of one or several fish feed factories, a set of fish farms, and vessels transporting the feed from the factory to the fish farms. Even though the principle of the network seems simple, the system contains a lot of variables and restrictions that makes it a lot more complex. This complexity of the full system comes from uncertainties and changes in the different components in the system. For starters, the feed factory produces a verity of different products, and the production rate also varies throughout the year, depending of the total demand. This means that there are several products that needs to be shipped out from the factory to supply the different fish farms, at different times a year.

The different fish farms also have several variables to consider. As mentioned, the demand for feed varies throughout the year, as the feed consumption is depending on the fish size and seasonal variation. In addition, disease breakouts and unforeseen events may occur, hence changing the feed delivery schedule. The fish farms may also require multiple feed types based on the size and health of the fish. Different farms have different storage capacities for different feed types. Running out of any feed types at any time will lead to loss in income for the farmers due to lost feed days. There are also restrictions of how fish farms can receive the fodder. Some can receive it as bulk, while others require big bags.

Next, the vessels are restricted by their storage capacity and sailing speed. The storage capacity sets constraints to how much feed each vessel can deliver during one trip, and the sailing speed determines the distance it can travel during a trip. Marine Harvest have four vessels distributing fish feed in their system, with different capabilities, such as speed, capacity and transportation method. Which farms each vessel visits are therefore limited by both individual feed reception methods, and the speed and capacity of each vessel. The farms are also spread out all along the Norwegian coastline and within protected waters, setting restriction to which farms each vessel can visit determined by the size of the vessel. The combination of several variables in fish feed production and a variational consumption rate, together with the uncertainty of unforeseen events, makes this seemingly simple system a lot more complex. In cases where demand exceeds available distribution capacity, partial deliveries are conducted. This will ensure that no farm is left without feed for an entire day. This may however result in delays in the original schedule, as the rest of the partial demand are given urgent delivery priority and must be delivered before other scheduled visits. Due to the constant changes in routes and orders before and during a trip, the planning and scheduling of the feed distribution network is difficult to conduct and optimize. The routing plan must be flexible, as the routes can be rescheduled during a trip, creating changes to the original delivery plan.

Marine Harvest's fish farming system is divided into four regions; south, west, mid, and north. The feed distribution does not play a part in deciding how the regions are split up, this is based on the fact that they have a slaughter factory in each of these regions. The distribution of feed is then planned based on the size and placement of the farms to get practical delivery routes. The company is aiming for a total supply for all their farms within a radius of 30 hours sailing time from the factory, including stops along the route. The transit between farms are normally short and can vary from 10 minutes up to a couple of hours. Overall, the most important factor during transit is to maintain the schedule and not to cause any delays.

Today, four ships are included in their fleet to perform the distribution, and these mainly follow fixed routes. One large that covers the farms north of the factory in Bjugn, given that the farms are large enough and that they are equipped with modern silos capable of handling fish feed as bulk. Another large vessel covers the larger farms that are capable to accommodate the vessel, from Bjugn and down to Bergen. A smaller ship serve the farms in the mid region, as it has the highest cost relative to the distance traveled. The farms furthest south are smaller and often placed deeper inside the fjords, meaning that the larger ships cannot be used here because of their size. Some of these farms are also older and less modern, thus needing fish feed supplied in big bags. Pirholm, which is a smaller vessel that is transporting big bags, is therefore used to service these farms. Since this route travels furthest away from the factory in Bjugn, Pirholm is not capable to deliver fodder here within the desired 30 hours. This makes the planning of delivery harder, and the vessel is therefore planned to be replaced next year with a ship what is able to meet the requirement. In general, a vessel visits the same farms several times a month, giving the crew and captain knowledge and experience of sailing the specific routes.

Another important factor to consider during transportation is pellet erosion. That can occur when the pellets are shaken or rubbed together during transportation, and fine particles loosen. Pulverized fish feed is transferred to the fish farms and it can result in overgrowth of algae and have a negative impact of the water-quality if these particles sink and settles at the bottom of the cage. Marine Harvest states that among 0.5% of the fish feed pellets gets pulverized during bulk transport, whereas the number for pulverized feed for big bags are a bit higher.

3.3.2 Feed Consumption and Variations

The biggest advantage of fish farming is the feed to growth ratio, which is far more effective for fish than other animals. This is so effective that about 1.1 kilos of fish feed may result in one kilo growth for the fish. This is due to the fact that the fish is a cold-blooded animal, which means that it has the same temperature as the water around it, and hence it does not waste energy to keep warm (Marine Harvest, 2016).

There are mainly two feeding strategies that are used; meal feeding and continuous feeding. Continuous feeding supplies the fish in the farm with feed throughout the day in small doses, while meal feeding gives bigger doses a given number of times during a day. The amount of feed the fish are consuming varies for different reasons, mainly the size of the fish. Hence, which strategy to use depends on the fish size and sea temperature. As mentioned in Section 3.1, there are various sizes and types of feed so that the producer can customize the feeding process to fit the local environment, seasonal varieties, and nutrition-need for the fish depending on the size.

Because the salmon is cold-blooded, the growth is strongly dependent on the temperature of the sea. This means that the production of feed around the world varies accordingly as there are large deviations in sea temperature globally. The optimal temperature for growth of salmon is between eight and fourteen degrees Celsius. We can see from Figure 6 that Norway and Ireland have the largest seasonal variation in water-temperature, while Canada and Chile have a lower variation. Figure 7 shows the seasonal profile of relative feeding in Norway, Chile, the UK and North America. We also see that Norway has the largest seasonality in fish feed consumption relative to the fish biomass, and that the variation in Chile is the lowest. By comparing the two graphs, one can notice that there is a strong correlation between the water-temperature and feed consumption (Marine Harvest, 2013).

The low season for consumption and production of feed in Norway is from February to April, while the high season is from July to September. In the low season the consumption can be as low as 30% of the high season, which makes the distribution schedule of feed more flexible during this period as the fleet is design for a higher demand (Marine Harvest, 2016). The production rate at the factory is normally held fixed during the high season, but is reduced during winter when the demand is lower. Marine Harvest balance these variations by including or excluding some of the farms in their distribution plan, depending on the time of year.

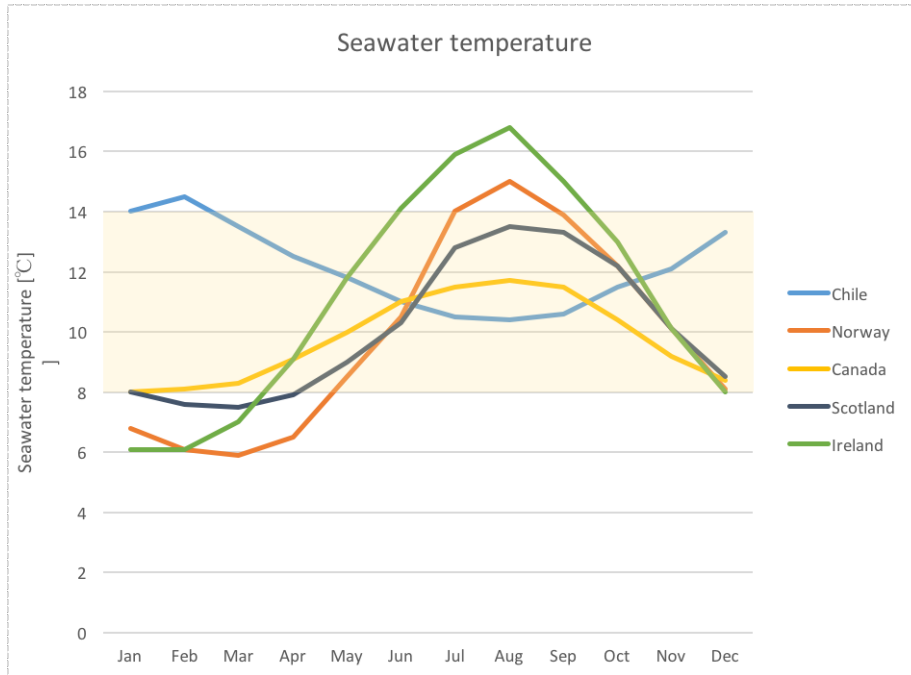


Figure 6: Annual temperatures of seawater.

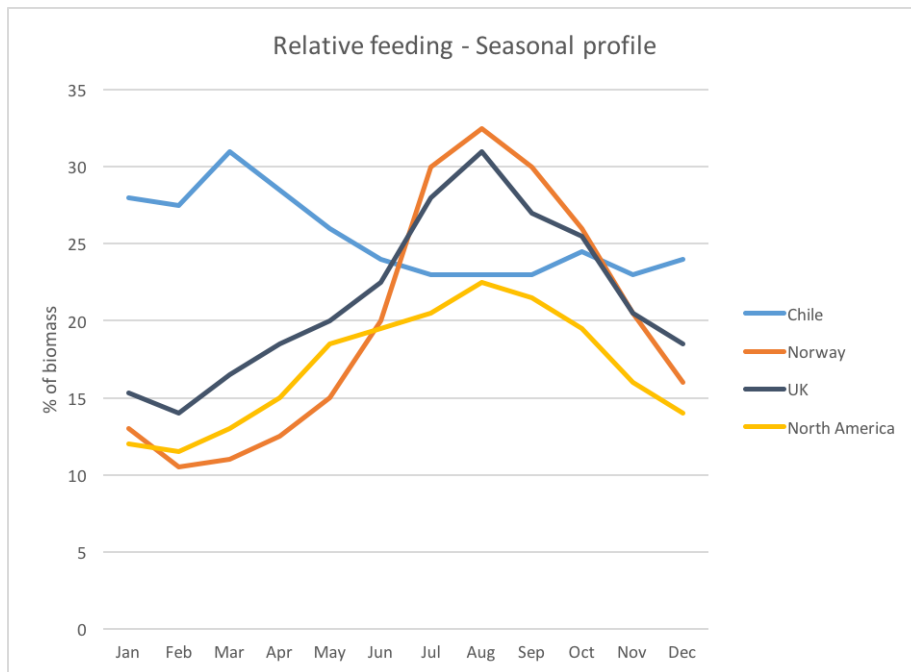


Figure 7: Annual consumption of fish feed relative to fish biomass (Marine Harvest, 2016).

In addition to the seasonal variation, the consumption rate naturally varies over the salmon growth process as their appetite also varies. In the beginning of the production cycle, the salmon naturally eats less feed than at the end of the cycle. During different stages in the life cycle, the fish eat different types of feed, giving a variation in which feed types each farm require. It is important that the factory has information about where in the production cycle each farm is, in order to plan the production and prevent running out of specific types of feed

at the storage. After the harvest period, the farms are fallowed for a period, thus there are no demand for feed during this period.

Another challenge in the fish feed industry comes from variations in the total demand due to unforeseen events. Unforeseen events can be disease breakouts at certain farms, requiring the salmon to be harvested or starved for a period, and hence changing the need for feed. This change may also come from mass-escapes of fish, also suddenly reducing the total need for feed for a customer. Both these scenarios lead to a decrease or halt in the need for feed deliveries, forcing the distribution planner to reschedule the delivery plan for one or more vessels. These uncertainties are further described in Section 3.4.

In defining the sequence of the customers visited by each ship, the planner must also consider the bio-security issues faced by the salmon farming industry. A ship that has visited a farm with a given risk level, cannot then service a farm with a lower risk classification without first being subjected to a disinfection process. The disinfection process can only be carried out at a major port, for the feed producer this means returning to its base port (Jonemark, 2016).

3.3.3 Planning the Distribution

Because of the complexity of the distribution network described in the previous sections, planning is essential to ensure cost-minimization. Since there are restrictions about which farms each vessel can serve, the routes are normally fixed. Each ship visits a calculated number of fish farms along the route for each trip, with small variations. How often a farm needs refill of fish feed depends on the storage capacity, and the feed consumption rate at the respective farm. Normally a farm needs feed 1-2 times per week. A vessel usually visits 15-20 farms per trip, and the planning of feed delivery is therefore very important to minimize the operational costs.

Since Marine Harvest delivers feed to its own fish farms, the logistics-planner at the factory uses forecasting estimations to predict the demand of feed at each farm. The forecast is based on the number of fish and their size at the different farms. This makes the planning of feed production easier and reduces the probability of a farm running out of certain types of feed. There is continuous communication between the location manager of each fish farm and the feed supplier, mainly over telephone or through email, to keep the orders correct. These orders are detailed with information about the farm, feed type, delivery date requested, and have to be placed at least two weeks before the delivery date in order to have enough time to produce and distribute the feed. A spreadsheet is used to set up a schedule for all four vessels. The spreadsheet contains information on which farm each vessel is visiting, when it expects to arrive at each farm, and the amount of each fish feed type the respective farm has ordered. The arrival time is based on a distance matrix between all farms and the factory, and the sailing speed of the vessel. The schedule is set up several weeks ahead, but it is constantly changed. The final order must be set four days before the vessel is leaving the factory. The reason for setting the final order date so close to vessel departure is due to the uncertainty in fish feed consumption, which varies based on fish size and appetite as described earlier.

The schedule for a vessel can also change during a trip. If weather conditions are not within the acceptable limits for unloading feed at the farms, the vessel may skip the farm and continue on the route. The farms normally have 1-2 days of feed available as a safety margin and if a farm is skipped, the vessel will have to come back on the return trip to unload the ordered amount of fish feed. In addition to bad weather, diseases at the fish farms can also affect the schedule, as certain diseases require different actions (further discussed in Section 3.4). Marine Harvest calculate a function of increased cost related to lost feed days, where one lost feed day is one day without sufficient feed at one cage at the farm. This means that if a farm has ten cages without any fodder for one day, the increased cost is then equal to ten lost feed days.

Today, 80-85% of Marine Harvest's farms are supplied from the factory in Bjugn. The rest are supplied by external fish feed factories, but the plan is to supply all of their farms in the near future. The farms supplied from external suppliers are on old outgoing contracts from the period before they build their own feed factory. Marine Harvest will from 2018 serve all of their farms with own produced fish feed, making them independent of other suppliers.

3.4 Challenges in the Feed Distribution Today

The feed-transport operations are determined solely based on work of the company's transport planners. Their job is to decide which feed orders to assign to which available ship, and in what sequence the deliveries are to be made. This opens for the possibility of human errors to be made. Since the work is mainly based on experience, it can be hard for other than the planner to have insight and understanding of the schedule and hence it is hard to have both control and redundancy of the work that is being done. Each order is manually entered into the system and a typing mistake can lead to a lower delivery than expected and farms can run out of fodder, leading to lost feeding days and loss in profit for the company. The transport planners are also responsible for determining the route for each ship, and have to avoid late or incomplete deliveries. During summer season, when the demand of fish feed is highest, determining the optimal route is difficult. The vessel has to follow a strict schedule in order to serve all farms. In this period, the schedules can be vulnerable to unforeseen events, which can lead to late or incomplete deliveries and loss in profit for the fish farms.

Due to the variations and uncertainties described in Section 3.3.2, it is hard to develop computer-based tools for the planning process, and hence most of the work is done manually in large spreadsheets. These can grow to be very large, and therefore not practical to use due to the sheer size. According to Jonermark (2016), the spreadsheets used by Marine Harvest was originally developed for distribution of four feed types, but today it has expanded to include a lot more types. This has made the spreadsheet more complex and less practical. It is not used for optimization of operation cost, but only as a planning tool. The planners add farms to the schedule in the order they think is optimal, but a more optimal solution may exist. If a farm has a critically low amount of feed at storage, the farm will be given an urgent priority. Then a nearby vessel will have to reschedule in order to visit the farm immediately, or the farm

has to be scheduled to the start of the route for another vessel. The same goes if a farm is exposed to diseases or bad weather, the farm must be rescheduled to be visited at the end of the route. These situations add further complications to the scheduling process, hence reducing the possibility to use tools to create routing-optimality.

The data to be implemented in the scheduling sheet is also delivered in a primitive and unpractical way. The planners are dependent on continuous communication with all the farms that require feed, and this is today done with individual phone calls or e-mails. This is currently leading to extra time-usage for the planners, but also introduces the possibility of communication-mistakes or misunderstandings. Marine Harvest admits that a web-based ordering system would have made the process easier, but this is not currently implemented.

3.5 Use of Multiple Depots

To try to compensate for the high uncertainty described above, we have thought about introducing external depots as part of the supply chain from factory to farm. The idea here is to include multiple depots closer to where the fish farms are located, for example in each region, for temporary storage. The fodder can then be shipped from the factory to the different depots in each region for storage, and from there be transported to the fish farms as illustrated in Figure 8 below. Marine Harvest already has¹ some external depots, but they are not implemented in the distribution chain yet. By 2018 they expect to take some of them in use as part of the delivery process, especially for the routes most north and south, which are furthest from the factory in Bjugn (Jonemark, 2016).



Figure 8: Illustration of a fish feed distribution system including multiple depots.

As we see it, there are various upsides and downsides with this alternative system. The inclusion of these depots could introduce the possibility to use larger vessels between the factory and depots, much like liner shipping with specific ports. A larger vessel usually has the advantage of a higher cargo to cost ratio compared to a smaller vessel. By only using these larger vessels for transporting the fodder from the factory, the port can be customized to suite exactly these ships, and by strategically placement of the depots, there will be no size restrictions for the ships transferring the fodder the first stage. This means that smaller, faster, and specialized vessels for specific regions can be used for the final stage of the distribution process. By increasing the speed, a higher number of farms can receive feed in a shorter period of time, hence increasing the efficiency and reducing the uncertainty. The use of regional depots also provides additional storage capacity and add flexibility to the varying demand of feed. Since the depots are closer to the farms, the time between orders and deliveries of feed will be reduced, decreasing the risk of lost feed days.

By using this system configuration, it can make the planning process easier. As mentioned, it is currently hard to use planning-tools as there are too many variables to take into account, but with external depots the big and complex system can be split up into smaller and simpler sub-systems. This means that simulations and optimization tools can be applied to the different individual parts of the system, and it will be easier to find optimal routes as there are fewer possibilities.

However, the use of larger vessels transporting the fodder from factory to the depots, means that more processing and transferring of the product. When loading the fodder as bulk, there will always be some losses. The introduction of an extra unloading and an extra loading process will make these losses more significant. This may cause the fodder to be transported in big bags. In order for the multiple depots to be beneficial, large changes in the distribution process is needed. Mainly, the system is dependent of major changes in the fleet composition to be optimal, which is a costly investment.

To conclude, we see this system as a great possibility to further optimize the distribution system in the future. Because the distribution system needs major changes for the multiple depots to be beneficial, we do not see this as applicable in the nearest future. The focus of this thesis is to optimize the current distribution system, but a study of the inclusion of a configuration that includes several storage facilities should be considered for further work.

4 Related Literature

One of the first studies of transportation problems was performed by Tolstói (1930). He published an article called *Methods of finding the minimal total kilometers in cargo-transportation planning i space*, where he studied the transportation problem and described a number of associated solution approaches. This was the beginning of a new research-field, where it today exists a large number of studies. Most of the research has been focused on land-based transportation problems, but over the years maritime transportation problems have gained substantially more interest. Christiansen et al. (2013) looked at the development of number of published papers and their research showed that the quantity of published research within the topic almost doubles every decade. This, combined with the possibilities of significant cost-savings which can be achieved by proper fleet scheduling and routing, is the motivation behind further research.

In order to create cost-effective and robust transportation plans in general, there exists a variety of models that can be used in the routing management process. In the fish feed industry today, the scheduling is normally done by hand with daily changes implemented manually, but theoretically the routing can be done in a more optimal way to minimize desired parameters. We will in this section point to typical characteristics for the transportation problem we are trying to solve, as well as present models for similar types of problems, and how these typically have been solved.

4.1 Comparison of Land-based and Maritime Transportation

Even though maritime transportation problems have gained more interest, there is still a great potential and need for research in the area. According to a report created by UNCTAD (2015), approximately 80% of the volume and 70% of the value of all goods transported worldwide is carried by sea. It is thus easy to imagine that the impact of research in planning of ship routing and scheduling can be huge, especially given the large concentration of players in this billion-dollar industry. Ronen (1983) discuss several explanations for the relatively low attention to ship scheduling problems. Among the explanations is the uncertainty in ship operations, where ships may be delayed due to weather conditions or mechanical problems. Ship scheduling problems are also less structured, as the problem structure and operation environment varies more than for standard vehicle routing problems. In addition, the major part of research on the topic has been performed in the US, where most cargo is transported by truck or railway. This, combined with the conservative tradition in the shipping industry, are some of the reasons to the low attention.

Maritime transportation differs from land-based transportation in many ways, mainly because it includes unique features that complicates the problem. Characteristics for cargo routing on land typically involve one supplier with several customers, capacities of vehicles much larger than the delivery to one customer, and a homogeneous fleet of vehicles where different products can be stowed together (Christiansen and Fagerholt, 2009). In maritime transportation, heterogeneous

fleet are much more common, and ships are different in their operation characteristics, such as speed, capacity, and costs. Even two identical ships can have different costs due to frequent fluctuations in the ship market (Ronen, 1983).

The nature of maritime transportation includes little slack in travelling schedules and capital intensive operations, as the transportation often include large volumes to be transported over long distances, overnight trips, and several loading and unloading ports. There is also an increased uncertainty related to breakdowns and weather conditions in the maritime industry compared to land based transportation. This varies somewhat between the three different modes of ship-transportation; liner, industrial and tramp. Liner vessels follow a fixed route according to a published schedule trying to maximize profit. Industrial shipping includes an operator that owns the cargo and controls the ships, and are trying to minimize the cost of delivering the cargoes. In a tramp operation, the vessels follow available cargoes trying to maximize profit (Christiansen et al., 2013).

Ronen (1983) also discusses other differences between maritime and land-based transportation. One important point is that ships not necessarily need to return to their origin. This is a situation which often occurs in tramp shipping, where the vessel may return to a new location where it is more likely to pick up cargo. Another point discussed is that the destination of a ship may be changed during a voyage. This is a scenario which is very applicable to the fish feed distribution industry, where fish farms may experience unexpected events, such as a disease breakout, causing the distribution schedule to be changed mid-sea.

4.2 Types of Transportation Problems

In this section, we will look into existing models which are relevant to our fish feed distribution problem and how the characteristics of the models relate to our problem. We are focusing on vehicle routing problems (VRP) and inventory routing problems (IRP), which are the two most common models used in the field of maritime transportation. Both models consider the transportation of cargo from one or several depots to a set of nodes, but the main difference is that the IRP is based on customer-usage rather than customer orders because of the inclusion of inventory management.

4.2.1 Vehicle Routing Problems (VRP)

In general, transportation problems are often different variants of so-called vehicle routing problems (VRP). It was first formulated and described by Dantzig and Ramser (1959), and has since then been a popular research-topic. The model includes a quantity of cargo to be transported from a depot to a given number of customers, subject to various constraints, such as the length of the route, vehicle capacity and time windows. The problem is then to determine how many vehicles that are optimal to use, and to decide which route each vehicle should take to satisfy all the demands. VRPs have a significant importance and it is faced on a daily basis by thousands

of distributors worldwide (Laporte, 2007). However, determining the optimal solution is a NP-hard problem, meaning that the size of the problem that can be solved optimally is limited. This is due to the fact that the solution-space increases exponentially with problem size. Commercial solvers therefore often use heuristics to find possible solutions (further discussed in Section 4.3).

One article that looks into optimization of fish feed distribution is written by Romeroa, Durán, Marencoe, and Weintraub (2013), where a VRP model to determine the feed delivery schedule for a feed supplier in Chile is investigated. They use characteristics such as ship capacities, ship/farm compatibility constraints, soft time windows, and priority definitions to model the problem. The priority constraint is due to bio-security, where a farm in a healthy condition cannot be visited after a farm in a risk condition. A heterogeneous fleet of ships is used, where each ship has a given capacity, average speed, fixed cost per day, and a variable cost for each traveled nautical mile. The authors also want to avoid partial deliveries, as they negatively affect the service level provided for the customer. A rolling horizon with a planning horizon of 3 days is used. To prevent lost feed days, any demand not delivered on its due date, is given the highest priority to be delivered the next planning period.

A paper with a wider application is written by Fagerholt (1999). He addresses the transportation of cargo from various production ports along the Norwegian Coast to markets in Europe and the United States. The focus of the paper is to decide an optimal fleet for the transportation of cargo to all customers, where weekly routes for the selected ships are the decision variables. The problem involves multiple trips as long as the total duration of the routes for each ship are within the planning period. It can thus be formulated as a multi-trip vehicle routing problem (VRPMT). The problem is solved using the route generation method (further explained in chapter 4.3.2), but solely for the important assumption of a homogeneous fleet. The paper also introduces depots in order to achieve an efficient shipping system. It consists of two independent transportation problems; transportation from production ports to the depots, and transportation from the depots to the markets. This approach has many similarities to our problem definition if multiple depots, as discussed in chapter 3.5, are introduced to the fish feed distribution system.

4.2.2 Inventory Routing Problems (IRP)

Inventory routing problems can be defined as a planning problem where an actor has the responsibility for both the inventory management at one or both ends of the delivery schedule, and also for the ships' routing and scheduling. It typically includes a planning horizon, inventory policy, number of products, product demand, no time windows, time considerations, and a fixed fleet of vehicles. The planning problem is to find routes and schedules for the fleet that minimize the transportation costs without interrupting production or failing to meet the demand. In contrast to the vehicle routing problem, the inventory problem has no predetermined number of visits in a given port during the planning horizon, neither is the quantity to be (un)loaded at each port. Predetermined pickup and delivery pairs are also often not considered, as opposed to cargo routing where each cargo has a specified pickup and delivery port.

The products transported in marine inventory routing problems (MIRPs) are usually bulk products, such as fish feed, where large quantities are transported. There are inventories at both the loading and the unloading ports, where the storage levels are constantly changing. The routing and scheduling of the fleet hence have to be synchronized with the inventory management at both production and consumption sites. This is often complicated, since one of the main issues in maritime supply-chains are often inventory control at processing facilities, and consumption at the customer sites.

One example considering such a MIRP is Agra, Christiansen, and Delgado (2013), where they look at the short sea fuel-distribution for an oil company. The problem is comparable to our distribution problem, consisting of routing and scheduling of ships between islands, so that the demands for various fuel products are met during the planning horizon. The paper also considers inventory management at the demand sites, and multiple time windows, as the ports in the problem have restricted opening hours. A small heterogeneous fleet is used to transport the fuel products between the islands, where each ship has a given capacity, speed, and cost-structure. Multiple products can be transported at the same time, but they cannot be mixed, and are thus separated into dedicated tanks. In this article, two types of formulations are provided and compared; a discrete time model where the production and consumption rates are varying, and a continuous time model used for constant consumption rates. Based on several test cases, they concluded that the discrete time model provided better bounds, but the running time was worse than for the continuous time model with constant consumption rates.

4.3 Solution Methods

Optimization problems are generally solved by maximizing or minimizing a function using variables that takes values within a pre-defined allowed set. If the problems have large numbers of variables, and a lot of input data, the solution calculation will be a complex process. The more complex and the more degrees of freedom the problem has, the more important it is to develop a tight formulation. The basic VRP is computationally hard to solve if the problem includes many customers with a large set of possible routes between them. The maritime inventory routing problem described earlier is even more challenging due to the additional degrees of freedom included.

In this section we will describe different types of solution methods that are often used for solving optimization problems. Normally we divide the solution methods into two categories; exact methods, that guarantee finding an optimal solution, and heuristic methods where there is no guarantee that an optimal solution is found. The effort of a solution method can be measured based on the usage of computer memory and computation time. Normally an exact method can be used if it is able to solve an optimization problem with an effort that grows polynomial with the problem size, otherwise the computation time and computer memory will be too extensive. For larger problems, a heuristic method, which is a trade-off between solution quality and effort, is a better choice (Rothlauf, 2011).

4.3.1 Exact methods

Exact solution methods refer to optimization methods that may be terminated before the process has investigated all possible solutions in the search-space. There exist methods that can solve optimization problems exact, but as mentioned this gets very time-consuming when the data-set increases in size. The following methods are some of the most used exact methods in transportation problems.

Branch-and-bound

The branch-and-bound method was first introduced by Lang and Doig (1960). The idea behind the algorithm is to break down the full problem into sub-problems which are easier to solve. The name branch-and-bound comes from the decomposing method, where the process of subsequently adding additional constraints is called branching. Bounding refers to removing sub-problems which are not considered any more, as better solutions exists (Rothlauf, 2011).

The method is the framework for almost all commercial software for solving mixed integer linear programs. One paper using the branch-and-bound method is Grønhaug and Christiansen (2009), which study the LNG supply-chain in cooperation with a leading actor in the LNG business. They test and solve both an arc-flow and a path-flow model by using Xpress Optimizer v17.1. As many other commercial software, Xpress uses the branch-and-bound method for finding the optimal solution.

Cutting-plane algorithms

Cutting-plane algorithms were first proposed by Gomory (1963), and are now the chosen solution method for a wide variety of problems. These algorithms are methods for solving integer linear problems (ILPs) and can either be used alone, or in combination with other methods, for example branch-and-bound. The concept of the algorithms is to add additional constraints to the problem so that infeasible fractional solutions are removed, but all integer solutions still remains feasible (Rothlauf, 2011). One example where a cutting-plane algorithm is used is in an article by Tang, Pan, Fung, and Lau (2009), where they propose and solve a vehicle routing problem with fuzzy time windows.

Branch-and-cut

The branch-and-cut method is a popular and successful algorithm for solving a variety of integer programming problems. The method is a combination of a cutting-plane method and a branch-and-bound algorithm. By combining these two methods, the reduction in the size of the solution-tree is considerable, leading to a more efficient solution method. The branch-and-cut method has in general a wide area of use which includes scheduling problems, network design problems, packing problems, and medical applications. Despite having a broad area of use, the method is best known for solving the traveling salesman problem (TSP). Grötschel and Holland (1991) and Padberg and Rinaldi (1991) are two papers discussing the research of the branch-and-cut method in large scale TSPs (Mitchell, 1999). More relevant for our problem is the use of branch-and-cut

for solving single vehicles IRPs, introduced by Archetti, Bertazzi, Laporte, and Speranza (2007).

4.3.2 Heuristic methods

To deal with the great number of possible solutions in an optimization problem, different algorithms have been created to be able to find a possible solution more easily, so-called heuristics. According to Rothlauf (2011), a heuristic algorithm is designed to solve a problem faster and more efficiently by sacrificing optimality, accuracy, precision, or completeness. Even though the methods perform a relatively limited exploration of the search space, they typically produce good quality solutions within reasonable amount of computation time. Heuristics can be used to provide a good baseline for the problem when approximate solutions are sufficient, or as supplement with optimization algorithms, when exact solutions are computationally extensive.

We will in this section present some general categories of heuristics, and briefly look at some relevant papers that successfully have been using heuristic approaches to find good approximations, to determine which ones that are most suitable for our case.

Constructive Methods

Constructive heuristics gradually build a feasible solution from scratch, where in each step parts of the solution are fixed. The heuristics do not attempt to improve the solution once it is constructed, but stop after obtaining a complete solution. This is because constructive heuristics are problem-specific, and hence different approaches are used for different kinds of problems.

One of the most common heuristic used for solving VRPs is the savings-algorithm developed by Clarke and Wright (1964). It applies to problems for which the number of vehicles is not fixed, meaning that it is a decision variable itself. It works by computing all possible alternatives for merging routes together, and list how much can be saved for each possibility. It then merges the routes with highest possible associated savings, until all merges are performed, and we can then decide how many vehicles that are needed in total to meet the requirements.

Some constructive algorithms are based on so-called greedy search. This is an iterative search approach that uses a heuristic function to guide the search process by estimating the minimal distance to an optimal solution. In each search step, a greedy search chooses the solution where the heuristic function becomes minimal, i.e. the improvement of the solution is maximal (Rothlauf, 2011).

Another constructive and improvement heuristic for the use of multiple trips for a mixed vehicle fleet, involving time windows, was proposed by Brandao and Mercer (1998). Their approach is based on the nearest neighbor rule and the insertion criterion to assign customers to routes served by a set of vehicles. This process is then repeated until all customers are included. By using real test-data, the results showed that their heuristic produced savings of 20% when compared to manual scheduling.

Metaheuristics

A metaheuristic is an algorithmic method that includes a set of guidelines to develop heuristic optimization solutions. The algorithm tries to find the best feasible solution in a solution space of an optimization problem by evaluating potential solutions, and then perform a series of operations in order to find new and better solutions. According to Sörensen and Glover (2013), metaheuristics have been demonstrated to be viable, and often a better alternative to more traditional methods of mixed-integer optimization. There are various types of metaheuristics that have been given a lot of attention, and we will in the following present some of them along with some relevant studies.

Tabu search is a metaheuristic proposed and developed by Glover (1989). The method is inspired by the principles of artificial intelligence, and is based on exploring the solution space by moving each iteration from one solution to the best solution in a subset of its neighborhood. This means that the solution may be updated from one iteration to the next. To avoid cycling, the method states that solutions containing some certain attributes of recently explored solutions are temporarily forbidden or declared tabu, hence the name of the method. Because it is using memory of previous solutions to develop new and better solutions, it is well suited for difficult and complex problems.

The tabu search heuristic is used by Brandao and Mercer (1997) to investigate a multi-trip vehicle routing and scheduling problem (MTVRSP). The problem is similar to the VRP, but has a larger number of constraints included, such as vehicles taking multiple trips during the planning period, time windows, and different vehicle capacities. In their problem, distribution plans must be produced daily, and the results must therefore be high-quality solutions produced in short computational times. They achieved this by developing an algorithm that takes advantage of both the classical tabu search theory, and an existing algorithm for the travelling salesman problem (TSP) called GENI. The algorithm created is capable of solving real distribution problems with practical constraints and actual costs. Applying the model to a real world system, they got an result which averaged about 20 % better than the manual solution created by an experienced planner.

Genetic algorithms are another kind of metaheuristics. They use computer procedures that are based on mechanics from natural selection and natural genetics to derive solutions. The method uses a population of candidate solutions instead of just a single solution, and then evolves this population by creating new generations. The new generations are created through three different mechanics; selection from two parents in the current population (crossover), reproduction of the best solutions from the current generation, or mutation that randomly modifies some of the genes of some of the solutions. Through an iterative process, the best possible solution is finally found based on convergence criterias.

Vaira and Kurasova (2014) have written a paper where they propose a genetic algorithm based on insertion heuristics for the vehicle routing problem with constraints. They first use a random insertion heuristic to construct an initial solution, or to reconstruct existing solutions. Then they

apply the crossover and mutation mechanics to some individuals, and finally select which parts should be reinserted. This increases the probability of the second solution to survive in the first population and thus increase the probability of finding the global optimum. The results show that the solutions found by the proposed algorithm are similar to the optimal solutions obtained by other genetic algorithms, but in most cases the proposed algorithm finds the solution in a shorter time, hence make it a competitive model.

Another metaheuristic is the multi-start metaheuristic. The method consists of two phases, where the first phase constructs several start solutions using a tour construction method, while the second phase uses a tour improvement method to improve the solution from the first phase. These two phases are alternated for a certain number of global iterations, where each global iteration produces a solution, usually a local optimum. The best overall solution found is then the algorithm's output (Martí, Resende, and Ribeiro, 2013).

Romeroa et al. (2013) uses a GRASP (Greedy randomized adaptive search procedures) algorithm, which is an example of a multi-start metaheuristic, to solve the unique features of the fish feed distribution industry. The paper focuses on several constraint types; capacity, compatibility, soft time windows, priorities, minimum load penalties for the vessels, and partial delivery penalty. The last two constraints originated from the local feed supplier in southern Chile. Their model was tested over a ten days period, where a rolling horizon of three days was considered in each run of the model. Compared to the actual scheduling plan, the model managed to reduce the total sailing distance by 6.39% and late deliveries by 65%.

The Two-Phase Heuristic

The two-phase heuristic (also called route generation method) includes two independent phases, a generation phase, and an integration phase. One advantage of the method is that it has a nice structure, which makes it easier to solve than a direct formulation. The disadvantage of the method is that it is necessary to perform both steps in order to find possible solutions, and the number of possible solutions grow exponentially with problem size.

In the generation phase, a construction heuristic is used to create a list of all possible routes for the problem. For this phase, any algorithm able to produce a large number of different routes can be used, but the approach is problem specific. This means that the number of routes generated is created by a set of constraints that vary from problem to problem. This can be time-windows, compatibility, maximum duration, or other restrictions. However, there is no guarantee that the construction heuristics will produce any feasible solutions for the next phase as there are multiple constraints included.

After all the possible routes are generated in the first step, a set-partitioning model is applied to select a subset of routes and assemble these into a final solution. The route combination is chosen according to an objective function, which typically involves minimizing a parameter, such as time or costs. The final combination must be in a way so that the solution fulfills all constraints in the set-partitioning model. The constraints are included so that the total system

is served in a desired way, for example that the total demand is met without exceeding any capacities.

The route generation method is especially applicable for solving maritime transportation problems, as the number of feasible routes is usually not that enormous. This is due to the fact that the problem sizes are normally not as big as in land transport because maritime transport serves fewer customers with larger deliveries. Maritime problems are also often more tightly constrained which reduces the number of feasible routes more drastically than for land based problems.

Wang (2014) investigates the use of this approach for an integrated operational transportation planning problem performed with a heterogeneous fleet. He defines sets of feasible routes that can be executed either by vehicles in an own fleet, or vehicles hired on a tour or daily basis. Next, different costs are introduced for different type of vehicles, only variable costs for own vehicles, and a tariff-rate for hired vehicles. The objective is to minimize the total request fulfillment costs including the fixed costs of the own fleet, the total operational costs, and charges for outsourcing requests to common carriers. As mentioned, the number of possible vehicle routes in these types of problems often increases exponentially with the problem size, making it impossible to enumerate all of them. In this case, Wang (2014) uses an iterative process in which only promising routes that can improve the objective function value are searched and added into the set of possible routes. In the end, he ends up choosing a set of candidate routes of different fulfillment modes and requests to be outsourced in such a way that the total cost is minimized.

Brønmo and Løkketangen (2007) also use a set partitioning type formulation for a shipping scheduling problem where ship routes are generated a priori. The problem is to find the most profitable assignment of cargoes to a heterogeneous fleet of ships, given a set of transportation requests within certain time windows. They state that finding a feasible schedule for serving all customers might be a difficult task itself, since the transportation mode usually is highly constrained. To solve the problem, they first generate routes by enumerating all feasible cargo sets for a given ship, and then for a given cargo set use a greedy insertion algorithm to find the route and schedule for the set. The model continues with this process until all cargoes are assigned to the given ship. To solve the second phase, a constructive heuristic is developed to solve the set partitioning formulation of the problem. The profit for each of the generated routes are listed, and the heuristic selects the best option. The selected route excludes all other routes for the given ship, and all routes containing at least one of the cargoes transported in the selected route. These routes are removed, and the candidate-set is reduced before the next iteration. The construction is finished when the candidate-set is empty. The model is then restarted a given number of times with a modified evaluation function. Finally, the method is tested on a set of test cases based on real world data, which are highly constrained, so that feasibility is difficult. The results showed that the heuristic is able to give several high-quality solutions for the planner to choose from.

5 Optimization Model

Based on the problem description in Chapter 2, we will in this chapter describe the development of the model used to solve the problem in hand. Since the size of the problem is so large that it would give unacceptable solution times, a heuristic approach is chosen, along with other simplifications. This method is based on earlier work performed by SINTEF Ocean, and in cooperation with Research Scientist, Inge Norstad, we have made alterations to the original work to make it suitable for our specific case.

In Section 5.1 we will first discuss the desired model characteristics, before we explain necessary simplifications and which applications our model is suited for in Section 5.2. Finally we will describe the route generation algorithm and the optimization algorithm in Section 5.3 and Section 5.4 respectively.

5.1 Model Characteristics

The model we explore is based on a route generation heuristic. As described in Section 4.3.2, the method has good properties suited for marine transportation problems. Possible solutions are easier obtained, which we see sufficient as we are only creating a decision support tool. Our approach consists of three steps; an input data file with all relevant information about the system, a route generation algorithm that creates possible routes for each ship based on a set of constraints, and finally a routing optimization model that selects the best set of routes so that all criterias are fulfilled, based on a desired objective function. The input file will be presented and discussed in Chapter 7.

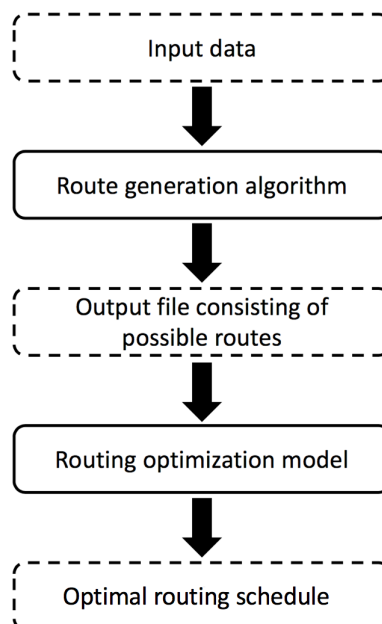


Figure 9: Model setup. Dotted boxes represent data files, and fixed boxes models.

The overall model is created in such a way that it works for a pre-defined system. As stated, we are using Marine Harvest's fish farming system as our basis. This means that we are looking at a system served by a fixed, heterogeneous fleet of ships, and only one production node where products are produced and stored. Maritime transportation problems often involve shipping of several different products, but since we are shipping the same type of product, and to reduce the complexity of the problem, we have chosen to simplify this to only one product type. There are positive and negative sides of using a heterogeneous fleet of ships instead of a homogeneous. According to Moutaoukil, Neubert, and Derrouiche (2014), the traditional VRP is usually based on a homogeneous fleet since the problem becomes more complex when the vehicle fleet is heterogeneous. On the other side, a fleet of vehicles with different carrying capacities provide capacity according to the customers' varying demand in a more cost-effective way, and this configuration has multiple advantages in real life. It is also a more realistic approach, as almost no vessels are equal.

The model setup includes the possibility for the different ships in the fleet to run on different types of fuel. In Marine Harvest's fleet, they have two ships running on LNG, and two ships running on a IFO-type of fuel. This implies different costs for different ships, as consumption and price varies for different kinds of fuel. Therefore, the opportunity to include input for what type of fuel each ship is using should be included in the model, and we can hence calculate the operational cost for a specific ship on a specific route based on this.

We have also included the possibility of having external suppliers delivering cargo to individual customers. This is to ensure that the model will generate solutions, even though the internal fleet is not able to meet the total demand of the system in any way. This also gives us the opportunity to manipulate how the system should be served by fixing some of the customers to receive cargo externally. The external service is included as a pre-defined fixed cost for serving each specific customer in the time period in question, on average a much larger cost than if the customer were served internally.

As described earlier, the system consists of several uncertainties and time-varying factors. When creating the model, we therefore have to set a planning horizon to avoid getting too many possible solutions. The length of the planning horizon depends on which decisions that are to be taken. The feed distribution problem must be considered on several planning levels when it comes to determining the optimal distribution plan and fleet configuration. In general, it is common to split the problem into a strategical, a tactical, and an operational planning level. Normally, longer planning periods are used for the tactical decisions than for the operational decisions, typically decisions applied for months, up to a year for tactical, and mostly on a daily or weekly basis for the operational (Schmidt and Wilhelm, 1999). A longer planning horizon often makes the problem more difficult to solve, since more options can be obtained. The purpose of the model is to function in the operational stage as a decision support tool, and hence the solver time should be as low as possible without having major impact on the optimal solution. Also, since our problem has a lot of short distances with much uncertainty, we decide to categorize our routing problem as an operational decision problem. We also need to choose whether to use a

time-continuous model or a time-discretized model. The latter is capable of handle time-varying variables, but is far more complex. Since the model is to function in the operational stage, with a short planning period, we can assume that the production and consumption rates are held fixed over this short period of time, hence the use of a time-continuous model is sufficient for our case.

Next, the customers are to be delivered a desired amount of cargo in the specific time period, at a calculated amount of time. The calculated unloading time depends on the amount of cargo delivered, and the offloading rate for the ship. In Marine Harvest's system today, most farms are highly automated and do not required any crew present. This means that all farms are able to receive deliveries at all times, and that all ships are allowed to unload at any time. This also goes for the loading process at the factory. An important aspect of the scheduling is the numeration of the nodes in the system. The factory is located in the middle of the country, and represent node 0. The 151 farms operating in the system are listed from north to south, meaning that the farm furthest north is node 1, and the farm furthest south is farm 151.

In the real world it is also possible to perform partial deliveries, meaning that the ordered amount of cargo can be delivered in several separate shipments. This is hard to model, but we introduce the possibility of delivering a part of a demand if the ship does not have the sufficient amount of cargo on-board while there is still time for more visits in the current planning period. This means that if a ship still has cargo left after supplying a demand, it is allowed to make another visits, even though it cannot fulfill the total demand of that customer. The rest of the demand, or more, will then have to be delivered on another route in the same time period so that the minimum demand is met for all customers. This model characteristic adds flexibility to the solution process, and possibly improves the utilization of the fleet, but also increases the complexity as it widens the solution space.

We have also included a constraint stating the minimum number of visits during the planning period. This is to give the distribution planner the opportunity of pre-specifying the minimum number of visits to one or several clusters during the planning period. This is mainly to make sure that all farms can be served if desired, even though it may not be in urgent need of a delivery. It also provides a fast and simple method of excluding clusters from the model, for example in situations where unforeseen events occur, or if a cluster is affected by disease outbreaks. This can be obtained by setting the minimum number of visits to zero.

The route generation model should in the end be able to create a list of possible routes for all ships within a set of constraints. The objective function is then to combine a set of these routes so that all requirements of the system are satisfied at a minimum operational cost. The operational cost consists of costs related to fuel consumption of the internal fleet, and costs of using external deliveries.

5.2 Simplifications and Applications

The real world system contains a large number of customer nodes, and serving these with a fleet consisting of several ships gives an enormous amount of possible routes to choose from. This makes it hard to model, and simplifications are thus needed. By making practical and realistic simplifications, we can create a model that can find actual solutions in a reasonable amount of time, but due to the inclusion of these simplifications, it also has its limitations.

To reduce the solution space for creating possible routes, the easiest way is to reduce the number of nodes to visit. The way we have chosen to investigate this, is to create clusters of nodes that are treated as one customer. There are several possibilities and problems related to this approach, and this will be more thoroughly discussed in the next chapter. The numeration of the clusters is done in the same way as for the farms, with the factory being node 0, the cluster furthest north being node 1, and the cluster furthest south being the same number as the total number of cluster included. To further reduce the number of possible routes we also introduce a compatibility matrix between clusters and farms. This means that we have the opportunity to manipulate the system so that not all ships are allowed to visit all customers. This can also be used to manually include time windows for short time periods, by fixing a specific customer to not be able to be visited by any ships in this exact period. By doing this, we get an approach that has similarities to how the routing is done in the real world, as some ships have close to preassigned fixed routes in the distribution process.

When clustering farms together it is hard to calculate the total demand of each cluster node since the collected farms does not necessarily have the same size and characteristics. As described in Section 3.3.2, the fish may be at different stages in the growth process, meaning that they consume different amounts of feed in the same time period. These reasons cause the estimation of the total demand to be both complex and time-varying. The need for real-time monitoring of consumption at each farm is therefore necessary to accurately calculate the total demand of each cluster at all times. The estimation made by us will only be a randomly generated number, based on scaling of the minimum and maximum consumption of feed by one farm. We also take advantage of the fact that the orders placed by customers almost never exceed the total amount of feed at storage, and we therefore assume that we are always able to meet the total demand of the system internally.

Another simplification is mentioned in the previous section where we choose to use a time-continuous model. This means that the variables such as consumption and production rates are held constant during the planning period, meaning that the model should be rerun once new information is available, and could hence only support weekly or even daily decisions for the delivery planning.

As previously presented in Section 3.1, there are several types and sizes of fish feed used depending on the fish size and health. The vessels used for the feed distribution are also equipped with multiple compartments for transporting different product at the same time. As discussed by Christiansen et al. (2013), the inclusion of multiple cargoes drastically increases the complexity

of the model, which also increases the calculation time and difficulty of solving the problem. Because our model is meant to be used as a decision support tool, and since the calculation time should be held at a minimum, we choose to simplify the system to only consist of transportation of a single cargo. To deal with this, the demand of the most critical type of feed could be used as a basis, and by scaling down capacities, the remaining available storage spaces could be assigned manually to other types.

Based on how our route generation algorithm is built, all ships must start and end their route at the factory in the planning period. This means that a ship cannot start on a new route unless it is possible to complete the whole route within the available time left in the period. When a ship completes a route, it can either wait at the factory until the end of the planning period, or if possible, start sailing a new route. This is a simplification that limits the application of the model, and can eliminate better and more optimal solutions, but as mentioned, this model is meant as a decision support tool during the planning phase of the distribution. Because all ships must start at the factory at the beginning, quay capacity at the factory is neglected.

Today, two of Marine Harvest's vessels can only deliver cargo during the farms working hours due to the unloading method on-board, while the other two vessels have the opportunity to unload at any time independent of the working hours at the farms. In the future it is desired that the feed unloading process can be performed around the clock for all ships, and thus increase the flexibility of the system. In our model we therefore assume that all vessels can deliver at any time during day and night. In the real world the ships have different unloading rates depending on the method they are transporting the cargo. Big bags have a lower unloading rate than bulk cargo, but because we allow delivery around the clock, we assume the same unloading method for all ships, and hence the same unloading rate. For simplicity when calculating this value, we assume that each ship always fully utilizes its cargo capacity when loading cargo at the factory, and that all ships are delivering exactly the loaded amount on a route before returning to the factory.

In order to estimate the time usage on a route, each ship is sailing with a given constant speed from the factory and between customers. This is pre-defined for each ship in the input data file, along with a specific fuel consumption for that given operational speed. Since the total operational cost depends on the specific fuel consumption, changing speed between model runs can help finding an optimal speed, reducing the total operational costs. The weather condition in the planning period is not considered, meaning that we assume that every vessel maintains the constant speed regardless of the weather and wave conditions, but this can be manually changed when using the tool for planning a specific period.

Because of all these simplifications, our model is only applicable to a limited number of problem types, and should only be used as an operational planning tool to support decisions on a short term delivery schedule. The model can help planning the distribution in an efficient way by avoiding certain expensive trips, for example the ships with the highest operational cost serving the farms furthest away. The model can also explore whether each ship should be fixed to serving a particular set of fish farms by fixing the compatibility matrix. We can then explore

the possibility of a certain ship sailing north and another south, while the two remaining are serving the ones in the middle, or other combinations. By changing the demand, we are also able to investigate the limits of the fleet in hand. With a longer planning horizon, the model can also be used as support for tactical planning, where decisions on a higher level can be made. An example of such a decision can be to determine if all four vessels are needed during low season. This will be further discussed in a computational study later on.

It is also worth mentioning that it is not a very extensive process to make alterations to the model to make it suitable for other systems than the one we are looking at. By changing the matrix dimensions and input data, other distribution problems can be evaluated with the same planning tool.

5.3 Route Generation Algorithm

The first part of the two-phase model that we are using to solve the distribution problem is the route generation algorithm. The main objective for the model is to transform the input data spreadsheet into a data-set containing all feasible routes for the fleet, as presented in Section 4.3.2. The output data-set must be on the right format and contain all necessary information needed by the optimization model. According to Research Scientist at SINTEF Ocean AS, Inge Norstad, the algorithm is most easily modelled in an object-oriented programming language such as Java. Since we have little or none experience with programming in Java, he suggested it would be best if he helped us with the code-writing part of the route generation algorithm. We will hence not go into detail on how the Java script is written, but rather how the feasible routes are generated using constraints.

In order to generate all feasible routes, the model has to satisfy a set of rules. The constraints are based on information from the input data-set and includes a planning period, cargo capacities of the vessels, and a compatibility matrix between the vessels and the clusters. The algorithm works by producing all feasible routes satisfying the constraints. One route can include several customers, as long as all constraints are satisfied, but each customer can only be visited once during a single route. The model checks all possibilities for expanding each feasible solution found, and add these to the current list of routes, and then this process is repeated for all solutions as illustrated in Figure 10. For each new route that satisfy all the mentioned constraints, a binary row is created containing information about which nodes that are included in the route. Figure 10 is an illustration of a system consisting of three farms, where eleven possible routes exists because of the inclusion of constraints.

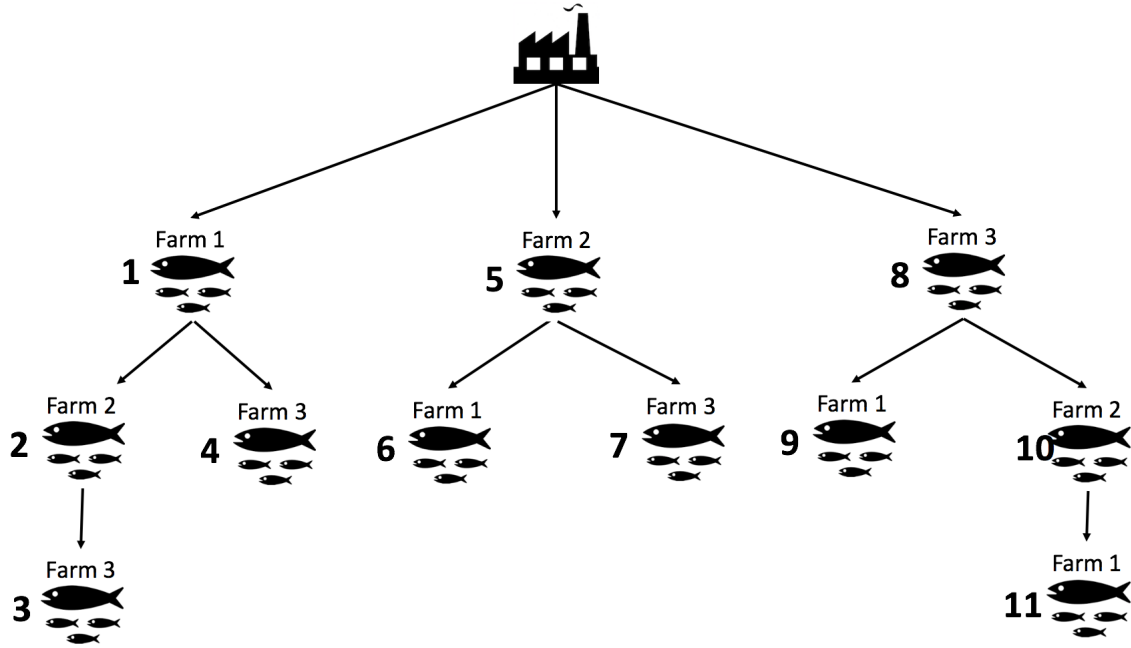


Figure 10: Illustration of the route generation method.

The algorithm starts by looking at how much time that is available in the current period. If no customers are added to the route, the available time equals the length of the planning period. It then connects all customer nodes to the factory, and calculates the time needed for visiting each of them alone. The time a full trip takes is based on the loading time at the factory, sailing time between the nodes, and time used within the clusters. If the total time this takes is shorter than the available time, the time-constraint is met for this specific route. Next the algorithm checks whether the inclusion of another visit after visiting the first customer node, without sailing back to the factory, is possible without exceeded the available time. If there is enough time, and all other constraints are met, the routes including two visits are added to the list of feasible routes. This goes on until all possible combinations are investigated.

All ships also have a capacity limit of how much feed they can transport on each trip, which works as a constraint on how many customers each ship can visit on a route before sailing back to the factory. Full freedom of split deliveries is not included in the model, because it will increase the solution space and thus increasing the calculation time. This means that it might be unused capacities at the vessels on some routes. As it is desired to fully utilize the cargo capacity of the vessels, the possibility of a partial split delivery solution is included. A vessel that has delivered cargo to one or several customers, and still has more cargo left, has the opportunity to deliver the remaining cargo before sailing back to the factory to end the route.

The final constraint defining the routes is the ship/cluster compatibility constraint. This is a matrix included in the input file, containing information of which clusters that can be visited by which vessel. If a vessel is not compatible with a cluster, the cluster cannot be added to a route sailed by the specific vessel, even if there is time and the vessel has sufficient cargo on-board.

The matrix is further explained in the input file discussion in Chapter 7.

The output file from the generation model does not contain information regarding the sequence of the visits on a route. This is because it is difficult to extract such information from the algorithm. In addition, we do not find it necessary to include, since the model will only be used as a decision support tool. The major advantage of excluding the routing sequence is the possibility to reduce the number of routes provided by the algorithm. When the routes are generated, the same customer nodes will be combined in many different orders, but the total cost of sailing the route will be the same. This means that all identical routes only need to be listed once, and the planning manager can choose in what sequence the customers should be visited. This will reduce the solution time for the optimization process, since there are less data to go through for the algorithm, which is desired. The objective of the optimization model is to reduce the costs, and therefore only the cheapest route of those including the same clusters is needed.

We do not see it as relevant to include all output files from the algorithm in this paper, as it only contains the same information as the input data-set along with a long list of routes. All files are included as attachments in the delivery of the thesis.

5.4 Optimization Algorithm

The second step in the two-phased method, is a model that selects the best combinations of possible routes into a complete solution. Based on the output file provided by the route generation algorithm, an Xpress model calculates the optimal composition of routes in order to find the best solution for the objective function. The model is based on the mathematical formulation presented in Section 5.4.1, and this is implemented in the commercial solver using the programming language Mosel as explained in Section 5.4.2.

5.4.1 Mathematical Model Formulation

We will in this section present the mathematical model for the final step of the optimization method, with the stated assumptions and simplifications from above. This model will use the information created by the route generation algorithm, and find the best possible solution based on a desired criterion. In Table 1 all sets, indices, parameters, and variables included in the model are presented.

Table 1: Definitions for the model formulation.

Sets	Definition	Unit
N	Set of nodes	[-]
V	Set of available vessels	[-]
E	Set of available charter vessels	[-]
R_v	Set of possible routes for ship v	[-]
Indices	Definition	Unit
v	Vessel v	[-]
r	Route r	[-]
i	Customer i	[-]
e	Charter vessel e	[-]
Parameters	Definition	Unit
C_{vr}	Transportation cost for vessel v to sail route r	[\$]
C_{ie}^e	Cost of serving customer i with charter vessel e	[\$]
T	Planning period	[hours]
T_{vr}	Time vessel v uses for completing route r	[hours]
A_{ivr}	=1 if customer i is visited by vessel v on route r , 0 otherwise	[bin]
S_i	Minimum number of visits during planning period for customer i	[int]
D_i	Demand of customer i	[tonnes]
Q_e	Capacity of charter vessel e	[tonnes]
F_{ivr}	Feed delivered to customer i by vessel v on route r	[tonnes]
Variables	Definition	Unit
x_{vr}	=1 if vessel v uses route r , 0 otherwise	[bin]
y_{ie}	=1 if customer i is served by charter vessel e , 0 otherwise	[bin]

We start by stating the objective function;

$$\min \sum_{v \in V} \sum_{r \in R_v} C_{vr} x_{vr} + \sum_{e \in E} \sum_{i \in N} C_{ie}^e y_{ie} \quad (5.1)$$

The objective function minimizes the sum of the costs for selecting route r for vessel v , plus the sum of the costs of using charter-vessels.

Next we will present and explain all of the constraints that are restricting the behavior of the model.

$$\sum_{v \in V} \sum_{r \in R_v} A_{ivr} x_{vr} + \sum_{e \in E} y_{ie} \geq S_i, \quad i \in N \quad (5.2)$$

Restriction 5.2 assures that the minimum number of visits required for each cluster is satisfied.

$$\sum_{r \in R} T_{vr} x_{vr} \leq T, \quad v \in V \quad (5.3)$$

Restriction 5.3 ensures that the total sailing time for each vessel is within the total length of

the planning period.

$$\sum_{i \in N} D_i y_{ie} \leq Q_e, \quad e \in E \quad (5.4)$$

Restriction 5.4 secures that cargo amount transported is within the cargo capacity of each charter vessel.

$$\sum_{i \in N} D_i y_{ie} \leq 1, \quad e \in E \quad (5.5)$$

Restriction 5.5 makes sure that only one charter ship visits one cluster.

$$\sum_{v \in V} \sum_{r \in R_v} F_{ivr} x_{vr} + \sum_{e \in E} D_i y_{ie} \leq Q_i \quad i \in N \quad (5.6)$$

$$\sum_{v \in V} \sum_{r \in R_v} F_{ivr} x_{vr} + \sum_{e \in E} D_i y_{ie} \geq D_i \quad i \in N \quad (5.7)$$

Restrictions 5.6 and 5.7 ensure that the amount of fish feed delivered to each cluster is below the storage capacity, and above the ordered amount of feed.

Finally, the binary variables are stated in constraint 5.8 and 5.9.

$$x_{vr} \in \{0, 1\} \quad v \in V, r \in R_v \quad (5.8)$$

$$y_{ie} \in \{0, 1\} \quad i \in N, e \in E \quad (5.9)$$

We will later in the paper explore the behavior of the model based on test cases in Section 8.2, and then discuss possible alterations and improvements of the model.

5.4.2 Xpress Modeling

In order to find the optimal routing solution, FICO Xpress-IVE Version 1.24.08, 64 bits, is used. This is a commercial optimization solver, used for solving linear, mixed-integer and quadratic programming problems. All parts of the problem in this paper are formulated using the Xpress-Mosel language. Mosel provides access to data in different formats (including spreadsheets and databases) and gives access to a variety of solvers, which can find optimal or near-optimal solutions to a problem (Guéret, Prins, and Sevaux, 2000).

The structure of the model usually contains declarations, data initialization and processing, a model section, solve-statement, and solution processing and presentation. In the Mosel language, all object must be declared in order to be implemented in the model, including sets, parameters and variables. In the data initialization and processing part, the input data files presented in Chapter 7 can be directly imported and processed. The modeling section consists of the mathematical model formulation described in the previous section, where the objective function is defined together with all constraints. The solve-statement, whether it is a maximum or a minimum function, is given as a separate command after all constraints have been declared. In the solution process, Xpress-IVE reduces the problem by using a variation of numerical methods, before finding the LP-relaxation of the problem. The last step of the solution process is to apply the branch and bound method. It is performed on the LP-relaxed solution by searching for an improved solution until the best feasible solution is found. Finally, the value of the optimal solution and the solution values for all variables is processed and presented in the output section (FICO, 2013).

6 Cluster Calculations

As briefly mentioned in Section 5.2, the system is way too large and complex to perform optimization on the full-size system. To be able to reduce the solution space of possible routes we have chosen to collect the farms into so-called clusters. This is not a very unrealistic approach, as a ship often visits farms close to each other when sailing on a pre-specified route. By doing this, the ships operating in the system will have fewer options when choosing which customer to sail to. The problem then becomes how the farms should be put together, how large the clusters should be, how we can calculate the distance between them, and how to calculate the internal service time when visiting one node representing several customers.

6.1 Cluster Generation

Clustering of nodes is a well know problem in the transportation problem area, and it has been developed several methods of doing this. Because our model is based on a route generation heuristic, we first need to cluster the farms together before the routes are created. We will therefore in this section focus on methods which are based on the cluster-first, route-second principle.

6.1.1 Theoretical Clustering Methods

Over the years, many clustering techniques have been proposed, and today there exists several methods of clustering customer nodes together in order to reduce the problem size. Some of the most common methods of clustering will be presented briefly in this section.

One commonly used method of clustering nodes is the k-means clustering approach introduced by MacQueen (1967). The procedure begins by selecting k initial cluster centers and then assign each customer node to its nearest cluster center. After a node is added to a cluster, the cluster center is adjusted in order to account for the new node. This is repeated until all nodes are assigned to clusters. The center of the cluster represents the mean of all the nodes included, hence the name k-means. It is a simple clustering algorithm that works well with large data sets, as in our case.

Closely related to k-means clustering technique is the k-medians clustering approach. The similarity lies in that both methods involve finding k cluster centers, but differs in how the center is defined. While k-means clusters use the mean location of the involved nodes, the k-medians clustering algorithm uses the median. The advantage of k-median is that it is very robust to outlier nodes in each cluster. This is in contrast to k-means, where one outlier can shift the cluster center away from the majority of the nodes (Whelan, Harrell, and Wang, 2015).

Another clustering method that can be considered as a cluster-first, route-second method is the sweep algorithm. It was first introduced by Gillett and Miller (1974), and it is clustering a group

of nodes into a route based on the polar angle between the nodes and the depot. This clustering method are normally used when several depots (or factories) are used, and determines which nodes each depot should serve. To transfer this clustering method to our problem, a selection of nodes can serve as depots, and the sweeping technique can be used from here to assign nearby farms to the basis-node and hence create clusters. If the idea of multiple depots (as discussed in Section 3.5) is introduced, this could be good method for determining which farms each depot should serve.

To most accurate represent the real world, all farms should be included in the optimization, but as stated this is extremely time-consuming to solve. Hierarchical clustering is a different way of cluster the farms. This approach starts with each node being a cluster and then repeatedly merging the two closest clusters to a single cluster. This can be performed until only one cluster remain, or until a satisfactory computational time is reached. A dendrogram, shown in Figure 11, is often used to display this technique. With this method, you can obtain the desired number of clusters by cutting the dendrogram at the wanted level (Tan, Steinbach, and Kumar, 2016). This method is also referred to as pair clustering, because of the repeatedly pairing of clusters.

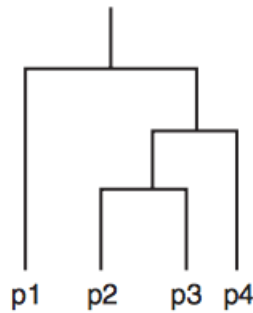


Figure 11: Dendrogram displaying the hierarchical clustering method.

6.1.2 Manual Clustering Methods

From the literature review we can see that there exist several theoretical ways of clustering nodes together, but most of these may not be relevant for our case as we are looking at a distribution along an uneven coastline. Instead, there may be other more practical methods that can be performed manually when collecting the farms in clusters. By simply looking at the map of the operational area, we should be able to create the clusters in a more reasonable way.

Maybe the easiest way to perform the clustering is to base it on practical sectioning of the coastline. By first dividing the coastline in a desired number of sections, and then define a geographical center based on the location of farms inside the area, we should be able to find the desired parameters. This is a simple method for Norwegian farms, as they are located almost on a straight line along the Norwegian coast. It is a practical way of splitting the system into sub-systems that also makes it easy to calculate the center and distances since each section can have approximately the same size and shape as Figure 12 illustrates.

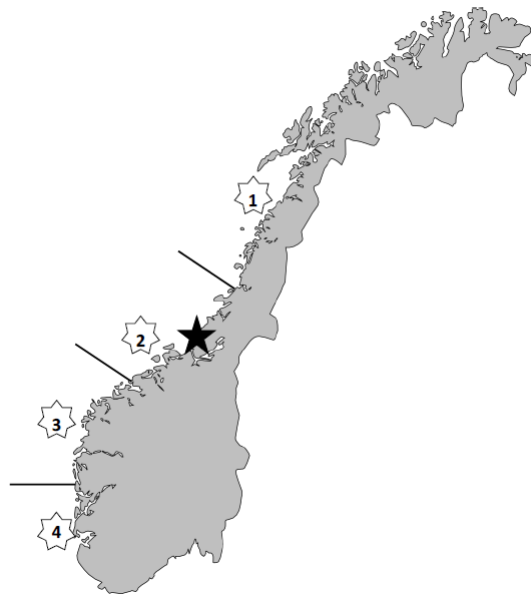


Figure 12: Example of large clusters based on geographical regions.

There is already a discussion going on about dividing the Norwegian coastline into production zones because of the fear for lice-deceases. The government has asked *The Institution of Ocean Research* and *The Norwegian Directorate of Fisheries* to come up with a proposal of how this should be done to minimize the risk of contamination between zones. In the report from Ådlandsvik (2015), on behalf of the research group, they propose a solution where the coastline is divided into 11 sections, showed in Figure 13.

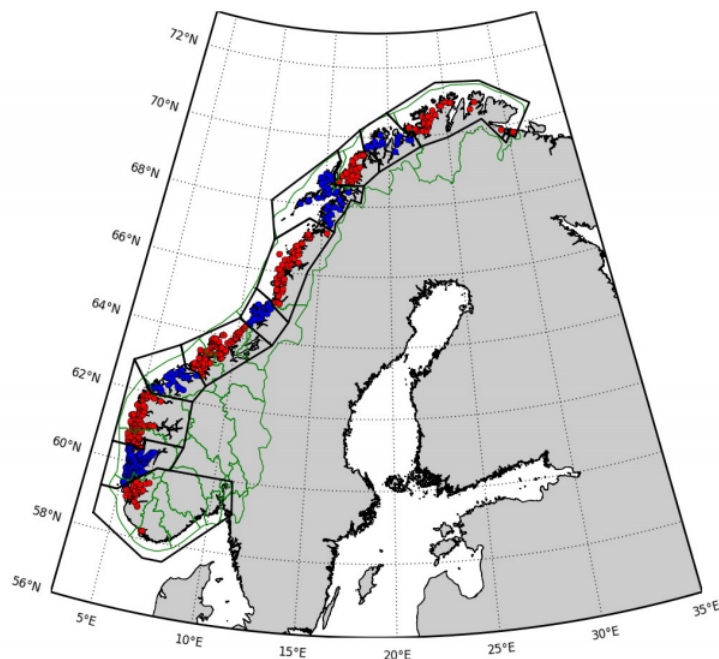


Figure 13: Proposed farming zones based on contamination-risk (Ådlandsvik, 2015).

This particular sectioning is based on proliferation modeling, where a influence matrix is created that is used to quantify the risk of contamination between farms. Even though we do not have any knowledge about this field of work, this is an interesting approach that our tool can applied to, in order to see how this will affect Marine Harvest's current distribution system.

Another manual way of clustering is to look at the map and study the geographical location of all the individual farms. In many cases they are placed in natural clusters, as there are limited number of optimal location for fish farming. By clustering the farms based on these natural locations, we get a very realistic system, but the clusters will have different sizes. This is a more time-consuming process, as it is not possible to generate these clusters automatically. It also gives us the possibility to manually exclude certain farms if desired. This may be farms that are located so far from the production facility that they should be excluded from the optimization process, and instead be considered as outliers and hence served externally.

6.2 Selection of Clustering Methods

With so many available methods for clustering, selecting the one best suited for our specific case is key. Our focus is on simplicity, calculation time and availability. Based on this we have chosen to look further into the k-means clustering approach, and also compare this method with the natural clustering approach mentioned. In addition we will look into the sectioning based on contamination zones, because it is currently being considered.

6.2.1 Contamination Zones

As a starting point, we want to use the sectioning proposed by Ådlandsvik (2015) as a basis to see how this would have affected Marine Harvest's system today if it is implemented. The coastline is already divided into sections by them, and we just have to place all the farms into their respective areas, and find the geographical center for serving as a measurement-basis. The reason for further exploring clustering based on the contamination zones, are the restrictions proposed in the report which are suggested to reduce the risk of contamination. If these restrictions are introduced, it can affect the current distribution scheduling, as ships cannot sail between farms in different zones on the same trip. However, if this system configuration is introduced, it also gives our model another possibility of use. When using the zones as clusters, we can use our model to optimize the sailing patterns inside each of the regions. In addition, depots can be placed out in each zone, and the distribution process could be broken into several smaller schedules which are easier to plan for. Because of this, we choose to include this clustering method further investigation.

6.2.2 Natural Clustering Method

A main study we want to look further into, is a comparison of an automatic approach and a logical manual approach. We therefore choose to include the natural clustering method, even though it is a more time consuming process. The advantage of using a manual clustering method like this, is how the clustering is performed. In an automatic approach, the clustering is based on straight lines between the farms. Even though farms are located closely in a straight line, they should not necessarily be clustered together. This is because of the many islands and fjords along the Norwegian coast, affecting the traveling distance between them. Tactical considerations like this can be taken when the farms are manually clustered, and we see this as a valuable characteristic for the method, and hence it is included in the further process. To implement this technique in our system model, we will look at the actual map with the position of the farms in the system, and create clusters based on what we believe are reasonable ways of collecting farms into groups. We will explore different alternatives when it comes to number of clusters, but one example can be seen in Figure 14 below.



Figure 14: Example of the natural clustering technique with 18 clusters.

6.2.3 The K-means Algorithm

Next we want to focus on the clustering method based on the k-means algorithm. The main reason for choosing this method is because of its fast and simple algorithm that works really well with large data sets. Because of the simplicity of the algorithm, it can easily be programmed in MATLAB. This gives us the advantage of generating several cluster sizes in a relatively short amount of time, using the same input data for the overall system. Another advantage when implementing the algorithm in MATLAB, is that the algorithm provides the coordinates for the cluster centers. We can then use this information directly to estimate the challenging distance-calculation, further discussed in Section 6.3.1.

The MATLAB script uses a built-in function called *kmeans* to partition all customer nodes into the k number of clusters. The objective of the function is to partition the farms in clusters so that the sum of the distances from farms to cluster centers are minimized. The function also provides a lot of other useful information, such as which cluster number each farm is placed in, all individual distances from each farm to its cluster center, and the sum of the distances inside each cluster. Figure 15 shows an example of a created clustered system where the k-means algorithm has gathered all customers nodes into ten clusters.

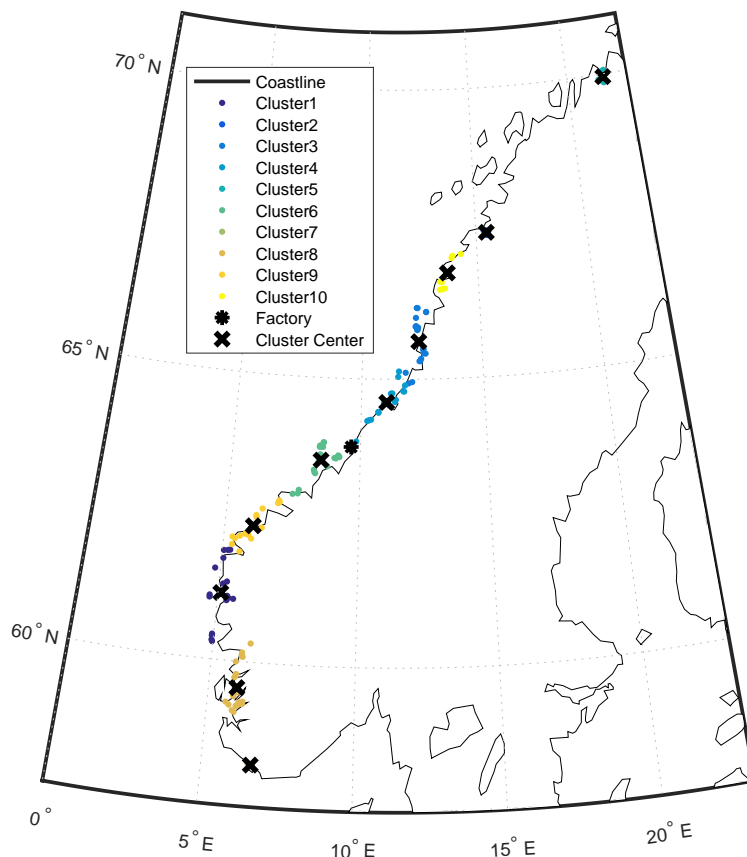


Figure 15: Example of clustering using k-means algorithm with 10 clusters.

We will use the three methods discussed above to further perform different computational studies to compare the number of farms in each cluster versus practical solution time. We will also look at how the theoretical clustering method behaves compared to a simple manual approach that is more similar to how the planning is done in practice today.

6.3 Parameter Estimations in the Clusters

After the clusters have been generated (independent of method), the question is then how to calculate parameters for the different clustered systems. As each cluster contains several farms with different geographical location and characteristics, it is not straight forward to calculate the overall characteristics for the full, complex system. There are mainly two parameters that we are interested in, and that will be discussed in this section; distance between the clusters, and the service time for each visit.

6.3.1 Distance Estimations

Since we are estimating distances at sea, and also along a coastline and in fjords, there are many infeasible sailing areas with shallows and other obstacles included. When measuring the distances, it will therefore in most cases not be correct to use a straight line. Also, since there are several customers to visit inside each cluster, it is hard to determine where to measure from, and where to measure to, when calculating the distance between generated clusters.

After discussing the problem with *Professor Kjetil Fagerholt, NTNU*, we agreed that there are several options for how to do this, and that this has been done in many different ways earlier. After the clusters have been generated, the question becomes if you should use the geographical center, or a center depending on density of farms, as the measurement-basis. Another less common possibility is to use the first farm the ship is arriving to as a super-node. It then becomes more critical how to generate the clusters, since the first farm a ship will be arriving at will be at the edge of the cluster-area. No matter how this is performed, we still have the main problem of finding the length of the sailing routes with obstacles and infeasible areas.

One way of dealing with this problem was performed by Fagerholt, Heimdal, and Loktu (2000), where they estimate the distances between ports as shortest paths in a network that models the land contours by nodes and arcs that represented obstacles, or infeasible sailing areas. By doing this they get good approximations, but the method involves extensive programming as the Norwegian coastline has so many contours, so this approach will not be considered.

One simple method of calculating the distances is using MATLAB's built-in mapping toolbox. It provides a tool and utilities for analyzing various geographic data. The distances are calculated using The Great Circle Mapper, which estimates the shortest path between two nodes on the surface of a sphere (The MathWorks, 2004). The drawback of using this built-in distance function is that it does not consider shore areas. This means that each distance is calculated to be as

short as possible and independent of land areas. Since Norwegian fish farms are mostly located sheltered between islands and in fjords, shore areas must be considered. By only calculating the shortest distance, the difference in sailing distance compared to the reality will be major and hence provide unreliable results.

Another more accurate method is to look at the sea traffic along the Norwegian coast and use real historic ship voyage-data recorded by AIS-signals to calculate the sea distances. This method provides good values of the sea distance from point-to-point, as the data material implemented comes from real ship routes. Searoutes.com (2017) is an existing commercial distance calculation service that makes use of the AIS-data to calculate length of sailing routes. Their algorithm can calculate an optimal ship route in both calm and rough waters. It also provides the possibility of calculating distances from point-to-point based on coordinates, in addition to port-to-port calculations. An example of a short route generated by the algorithm can be viewed in Figure 16.

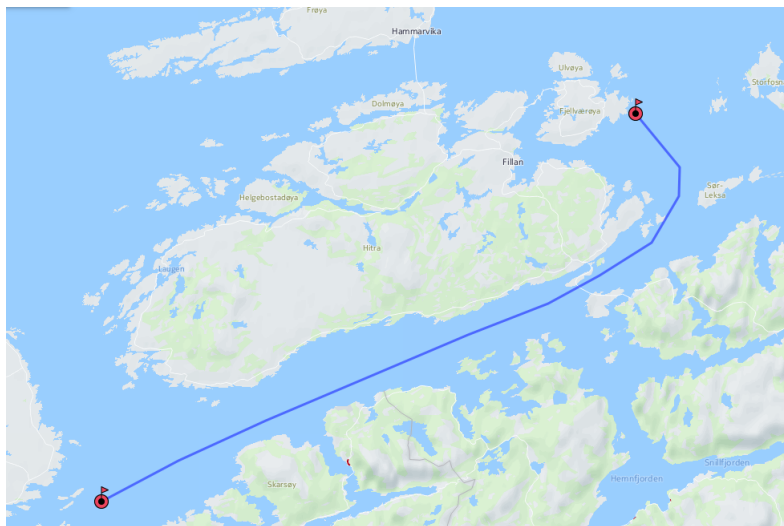


Figure 16: Example of route generated by Searoutes.com (2017).

As seen from Figure 17 below, the difference between using Searoutes.com (2017) and MATLAB's built-in mapping toolbox for estimating distances in our system is major. In this exact example, the measured distance is 20 % shorter with the MATLAB mapping tool compared to Searoutes.com (2017).

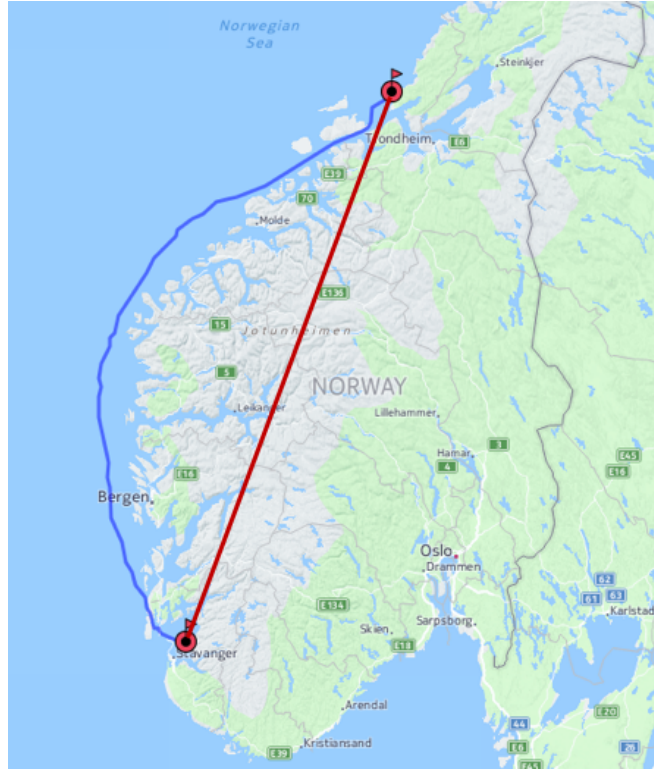


Figure 17: Comparison of route generated by Searoutes.com (2017) (blue) and MATLAB's mapping toolbox (red).

Even though several methods can provide good estimates, they will never manage to get 100 % exact values. Since it is easy to make manual alterations in the distance matrix in the input file, we do not see this calculation as the most important factor to be able to get good results. The best way of getting accurate data would be to consult with captains on ships that are sailing the actual routes on a regularly basis, and implement this data in the input file, but this will not be considered in this paper. We will therefore make an estimation using the web-page Searoutes.com (2017), as we see this as the most user-friendly and accurate method. This approach has the advantage that we can use the coordinates for a manually chosen center for all clustering methods that are being evaluated as a measurement point when calculating distances. If another approach is preferred, or if the data is available, this calculation can be done separately and the new values for distances can be inserted in the input file.

6.3.2 Estimation of Service Time

Another problem occurring when we are clustering farms together is the time usage of each customer visit. When using clusters as customers, the time of a visit comes from two different contributors; service time and internal sailing time.

The first contributor is normally easily calculated as how fast the cargo is transferred from the ship to the customer. However, in our case, visits to all farms in a cluster are represented by one single visit. The different offloading times must hence be scaled up to one single process.

As this time is dependent of how much cargo the vessel is delivering to each farm, and that each farm inside the cluster may have different demands, it is hard to calculate this directly. The total demand is a direct input that has to be changed every time the model is being run (see next section). In addition, we do not know how much feed that are being delivered to each cluster before the optimal routing schedule is generated. The service time for offloading must therefore be calculated from the total demand of the cluster divided by the offloading rate for each ship, since different ships have different offloading equipment and different cargo capacity. The loading process at the factory before starting on a route must also be included.

The second contributor is not part of the service time in a real scheduling problem, but since we are clustering several farms together and only using one visits to serve all, we also need to approximate the sailing time inside each cluster. The farms in a cluster are not necessarily close to each other geographically, and since the internal distance may vary from cluster to cluster, we need a way to approximate the total time spent on each cluster-visit. This is of course dependent of how many farms there are in the specific cluster, and also the total extent of the cluster. The most optimal solution would be to optimize the routing within each cluster based on consultations with ship captains. In our case, and since we are testing several cluster sizes and configurations, we need a simpler method of calculating the internal sailing distance. Because we have the coordinates for the center of each cluster, and also for all farm locations, we can calculate the distance from each farm in the cluster to the center, as illustrated in Figure 18. All the distances in each cluster can then be summed to find the total sailing distance within the cluster. This is of course not the optimal way of estimating the distances, and after comparing the calculated distances with more optimal and realistic internal routes for different clusters, we found that our method gave an overestimation of the total distance. We have therefore decided to divide the total distance found with this method by a factor of two. The total internal sailing time can then be found by dividing the total sailing distance inside each cluster with the speed of the ship visiting the specific cluster.

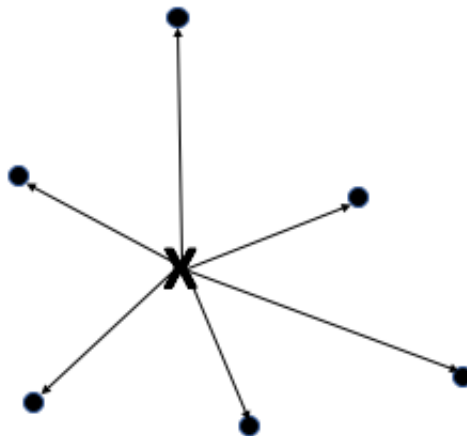


Figure 18: Illustration of the internal sailing time estimation method.

The total time-usage within each cluster finally becomes the sum of the offloading time and internal sailing time of each cluster. As mentioned this is just an estimation, normally a ship will follow a route based on the position of the farms. How many farms we choose to include in each cluster also affects the uncertainty of the final solution. If we have larger clusters (fewer clusters with more customers included) there will be more internal sailing, and a less accurate estimation of the total sailing time. If we instead choose many smaller clusters, this will give more accurate solutions, but the solution-time will increase. The internal sailing time is also depending on the ship characteristics, and will therefore be different for different ships. The ships are set with a given operational speed which is assumed constant during the voyages. When sailing internally between the farms in each cluster, the vessel in the real world may have to sail at a lower speed due to the archipelago along the coastline. To improve the validity of the estimations, the sailing- and service time should be acquired by consulting with ship captains and by performing tests of internal routing possibilities.

7 Input Data

To be able to use the route generation algorithm described in Section 5.3, we first need a data-set with all necessary information about the system we want to optimize. As mentioned earlier, we are basing our problem on Marine Harvest’s distribution system, and we hence need actual information about capacities, demands, distances and ship parameters to represent the system as accurate as possible. All the information is gathered in a simple excel-sheet which can be implemented by the Java script used for the route generation process.

7.1 System Dimensions

First of all, we need to state the overall dimensions of the system. This includes the number of ships operating, number of available charter ships, number of clusters in the system, and the length of the planning period.

To perform the distribution in the system, Marine Harvest are operating a fleet of four ships, plus the possibility of using external deliveries for desired locations. All ships are operated by Egil Ulvan Rederi. Two of the ships are identical with a cargo capacity of 3 000 tonnes, and running on LNG with an operational speed of about 13 knots. The two other ships are not completely similar, but for simplicity we use the same dimensions for these two as well. They operate at about 10 knots, are running on a IFO-type of fuel, and have a cargo capacity of about 1 500 tonnes. According to technical manager at Egil Ulvan Rederi, *Ivar Ulvan*, the two LNG-powered ships have a fuel consumption of about 27 kg per nautical mile at the operational speed, which gives a specific fuel consumption of about 0.35 tonnes per hour. The two other vessels use about 0.17 tonnes per hour.

Table 2: Attributes of the ships used in the system.

Attributes	Ship 1	Ship 2	Ship 3	Ship 4
Capacity [tonnes]	3 000	3 000	1 500	1 500
Operational speed [knots]	13	13	10	10
Fuel type	LNG	LNG	IFO	IFO
Fuel consumption [tonnes/hour]	0.35	0.35	0.17	0.17



Figure 19: Illustration of the LNG ship "With Harvest" (*ilaks.no, 2014*).

Because the number of farms in Marine Harvest's system is too large to efficiently find an optimal routing between all of them, we have chosen to collect the farms into clusters with the methods described in the previous chapter. The total number of clusters that all farms are assigned to is specified in the start of the input file as a main dimension, and will be used for calculating other relevant information about the system, such as demand, cluster storage capacity, and service time.

The planning horizon for this distribution system is set to one week. This is in coherence with the purpose of the model, which is to function as a decision support tool in the operational stage. In addition, a planning horizon of one week fits well with how Marine Harvest plans their feed distribution today, where fish farms normally are visited once per week.

If a charter ship is chosen to serve a customer, we assume that this is at a fixed pre-defined cost for this specific customer, depending on how many farms that are located in the cluster. The price implemented in the input file should be so high that for most cases it is not profitable to include any charter ships, but for some cases it will reduce the total cost. This cost represents all necessary service for the total demand to be met inside the time period (see Section 7.4 for further discussion of cost calculations). The number of charter ships must also be stated in the input file, and we restrict the maximum number of available charter ships in this case to be four, which is the same number of vessels included in the internal fleet. In addition, the capacities of each charter ship must be specified, but this is just included to be able to match the maximum demand of each customer, so that all customers have the possibility of being served externally.

7.2 Cluster Information

We also need different relevant information about the customers we are visiting. To begin with, we introduce a compatibility matrix between ships and clusters to reduce the size of the solution space as mentioned in Section 5.2. This is a binary matrix, where 1 means that a ship is capable of visiting the cluster, and 0 means that it is not. One possible approach using this application

could be to fix the two smaller and slower ships to not be allowed to sail to the farms furthest north and south, even though they are capable of doing so. Another is to restrict some ships to not sail further away from the factory than a pre-defined limit, or opposite, not visit farms closer to the factory than a certain distance. Also, if we have some outlier farms that are not included in Marine Harvest's distribution plan today, we can use this matrix to fix them to be served externally. In a computational study later in this paper, we will experiment with different compatibility scenarios to test the effects of changing this matrix.

The demand in the current time period of each cluster is determined from the sum of the consumption rates of the farms inside the respective cluster. This will vary throughout the year, but the dimensions of the feed orders are correlated with the planning period. A longer planning period naturally gives larger feed orders. This demand input must be changed for every planning period, but if put to real use, this information should be available for the distribution planner. To make an estimation, we calculate an upper and lower bound for weekly feed consumption for a single fish farm based on the feed conversion ratio for Atlantic salmon (Skretting, 2017). The daily growth for a salmon depends on seawater temperature, and the current size of the fish. Austreng, Storebakken, and Asgard (1987) provides a table for the daily growth of a salmon in percentage for these factors. This varies from 0.1% of its own weight, to 1.1% of its own weight. By combining this with the feed conversion ratio of 1.15, and assumptions of the fish farm's dimensions, we can calculate the demand for the desired period from Equation 7.1.

$$FC = \frac{FW * GR * FCR}{100} * n * nf * d \quad (7.1)$$

where FC is feed consumption, FW is the average weight of a fish in the farm, GR is the growth ratio in % of the current weight, FCR is the feed conversion ratio, n is the total number of fish in one cage, nf is the number of cages at the farm, and d is days in the planning period.

By assuming that each cage contains 100 000 fish, with fish weights ranging from 0.2 kg to 4.5kg (Marine Harvest, 2016) between farms, that each farm has eight cages, and that the planning period is one week, we can based on the numbers from Austreng et al. (1987) calculate a minimum demand of 25 tonnes per week, and a maximum demand of 300 tonnes per week. Since we do not have any information regarding the fish in each farm at the moment, we will create pre-defined consumption areas and explore different consumption scenarios for these areas. We start by dividing the coastline into four sections, where each section has the opportunity to have low, medium or high consumption rate for a specific time period.

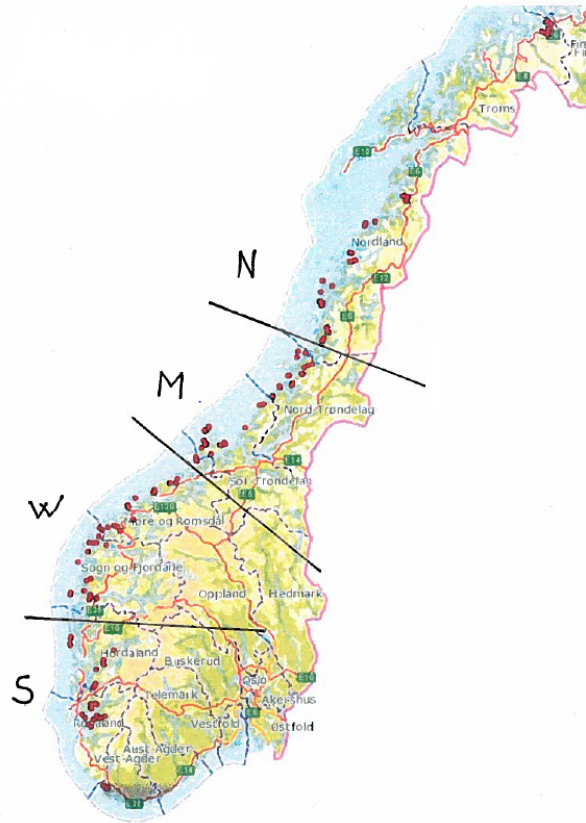


Figure 20: Definition of consumption zones with all Marine Harvest’s farms.

By using this sectioning, we can fix all farms inside a zone to have the same consumption rate, independent of the clustering method. To reduce the number of possibilities, only three possible rates are considered. A low consumption rate means that the fish are early in the growth phase, or/and that the seasonality are causing the fish to eat less. We have chosen this to be near the lower bound of consumption for a farm, about 75 tonnes per week. A medium rate corresponds to a farm consuming 150 tonnes per week, and a high rate is set to 250 tonnes per week. These number are not placed at any of the exact limits of consumption rates, but in the real world the scenario where all farms have equal demand will not occur.

We also include information about the storage capacity of each customer. According to *Claes Jonermark*, Operation Director at Marine Harvest Fish Feed Factory, their farms are able to store anything from 200 to 800 tonnes of feed at site. Since we for simplicity assumed that all farms contain 100 000 fish, we also choose to set the storage capacities for feed equal for all farms. This value is set to 300 tonnes per farm, which is equal to the calculated maximum demand per week. This number is then directly scaled up according to cluster size (number of farms inside cluster). Since this is included in the input file for each period, the inserted values can be treated as available storage space of the cluster, rather than total space. This is to make sure that the amount of delivered cargo does not exceed the available storage capacity at any point. Since the numbers are both approximated and scaled up, it will not be 100 % correct, but this can easily be changed by somebody with detailed knowledge about the system.

Since the sailing times between the clusters are not the only contributor to the time used on a route, we also need information about how long each ship is staying at each cluster to calculate the total time usage for each possible route. As stated in Section 6.3.2, the additional time use comes from service time when unloading feed at each farm in the cluster, and from internal sailing time inside each cluster. The loading process at the factory before starting on a route must also be included.

The service time related to the unloading process depends on the customer's demand, and the offloading rate for the specific ship delivering the cargo. Since we do not know which vessel that are visiting which cluster, and neither the amount of cargo to be delivered when creating the input file, we must assume that each vessel is delivering the exact demand for each visit when estimating the offloading time. This is calculated directly as the demand divided by the offloading rate, and then implemented in the input file. The same goes for when a ship is visiting the factory for loading new cargo. Here the service time will be calculated as the total cargo capacity of the ship, divided by the loading rate at the factory. According to Jonermark (2016), Operations Director at Marine Harvest, the newest ships have an offloading capacity of about 250 tonnes per hour, and a loading capacity at the factory of 300 tonnes per hour. We will use this as our basis, but as for the rest of the input parameters, this can easily be changed to suite a desired system. The internal sailing time for a specific cluster is found by using the estimation procedure presented in 6.3.

In order to calculate all the needed information about the clusters, a MATLAB script is created. The script is based on input data about all customers that are to be clustered. This information, regarding the individual fish farms, can be accessed from Fiskeridirektoratet (2017). From here we obtained a list of coordinates for all fish farms owned by Marine Harvest, which can be found in Appendix B. Because of different clustering methods, two different scripts are created. For the k-means method, the clusters are automatically generated and the coordinates are given for each cluster center. On the other hand, for the manual clustering methods, we are manually creating the clusters, and can thus manually decide the center of the clusters. In order to find a realistic center-location, adjusting for the dispersion and density of the cluster, the mean value of the farms latitude and longitude values are calculated in the script to represent the center of each cluster for the two manual methods. Based on the cluster center and the number of farms in each cluster, we can calculate the wanted information. The scripts are able to produce values for the total time spent in each cluster, as well as the demand of each cluster based on the consumption scenarios describes above, and also cluster capacities. The price of serving each customer with a charter ship, depending on the cluster size, is also calculated in the same scripts.

Both the script for the k-means clustering method, and the script for the manual clustering methods are attached with the delivery of this thesis (See Appendix A for list of attached files).

7.3 Distance Matrices

Since we are modeling for a fixed time period, we need to know the length of each route in order to not exceed the total available amount of time in the planning period. In addition to the internal time usage of each cluster, we need to calculate the sailing times between them. A distance matrix for each clustering method and system dimensions must therefore be created. The dimensions of the distance matrix can easily be changed depending on number of clusters in the system. As discussed in Section 6.3, the distance estimations are done dependent of how the cluster centers are defined. When the decision of how many clusters to collect the farm into is made, we calculate the distances between them based on the coordinates for all cluster centers for each of the selected clustering techniques, generated by the MATLAB scripts presented in the previous section.

The work of manually generating a distance matrix for all wanted cluster configurations is time-consuming. A MATLAB script that generates a distance matrix between all desired locations, based on their coordinates, is hence created. Because of the discovered drawbacks of using MATLAB's built-in mapping toolbox, we created a script that connects to Searoutes.com (2017), plots in coordinates between two locations, and finally store the distances between them in a matrix. This process is repeated until all the distances (in nautical miles) between locations are collected. Because the script is only able to calculate one distance at the same time, and needs to connect to the web-page for every calculation, the calculation time is high and dependent of the internet speed. In order to reduce the calculation time, we take advantage of the characteristics of a distance matrix. Since a distance matrix is symmetrical along the diagonal, we only need to calculate one side of the diagonal. This reduced the calculation time by 50%. The script is attached with the delivery of the thesis.

The total time usage for each route is finally calculated directly in the route generator algorithm as distance between clusters divided by operational speed for each ship, plus the total service time for each visit made on the route. This amount of time is then the basis for the cost calculations presented in the next section.

7.4 Cost Calculations

Finally, we need to calculate the cost of each of the possible routes in order to find the cheapest possible routing schedule. We are in this paper only focusing on the operational costs, ignoring product, crew, survey, maintenance, insurance, and other services as they are fixed and independent of which route we are choosing. This approach is usually sufficient for routing problems in contrast to fleet size optimization, where both the fixed costs and operating costs should be taken into account (Fagerholt, 1999). It should also be noted that all costs are rough estimations made by us for comparing solutions, and should not be interpreted as real cost values.

As explained previously, the ships in the fleet are running on different types of fuel, which have different fuel prices. The two fuel options Marine Harvest's fleet uses today are a type of IFO and

LNG, which per 23. April 2017 costs around 310 \$ and 260 \$ per tonne, respectively (Singapore Bunker Prices, 2017). The ships also have different specific fuel consumption at their operational speed, as they have different characteristics and specifications. Since most of the time spent on a route is sailing between clusters and internal sailing inside the clusters, the costs are calculated based on the total time of a route. This means that the reduction in fuel consumption during loading and unloading processes is neglected. The operational costs are then calculated from Equation 7.2.

$$SC = t * SFC * FP \tag{7.2}$$

SC is sailing costs [\$], t is total route time [hours], SFC is the specific fuel consumption at the operational speed for a vessel [tonnes/hour], and FP is the fuel price of the selected fuel type [\$/tonne].

The price of having feed delivered externally is, as mentioned earlier, a pre-defined fixed cost which can vary between different clusters. Normally the price includes the cost of the cargo itself as well, but we assume that this is calculated separately and that we are only looking at external operational costs, and the profit-margins of the external company. The price depends on several factors, such as distance from shore, external factory location, available service hours, or general availability. Because external deliveries are undesired, the cost is set to a high value compared to average deliveries included in the internal routing solution. This will then prevent the optimization model to include the option of external delivery if possible. We also assume that external feed suppliers have factories located in different areas, covering the whole Norwegian coastline, meaning that the external transportation costs are set equal for all farms. In order to adjust for the varying cluster sizes, the cost of external feed delivery to each cluster is correlated to the number of farms in each cluster. Based on the average operational cost for serving one farm internally, and including a profit margin, the total cost of serving one farm externally is set to 1 000 \$. This gives a relative high value, making internal deliveries the preferred delivery option. This can of course easily be changed by the distribution planner to any number or function desired to evaluate options more accurately.

8 Computational Study

We will in this chapter start the computational work by using the two-phase model described in Chapter 5 along with the input data-file described in Chapter 7. We will begin with making an input data-set for an artificial system and generate routes for this system to explore the behavior and applications of the optimization model. The mathematical formulation presented in Section 5.4.1 are written in Mosel programming language, and implemented in Xpress-IVE Version 1.24.08, 64 bits, to get solution outputs. When we are familiar with the procedure, we will look at the different clustering-alternatives from Chapter 6 and perform a computational study to determine the best suited dimensions for real world evaluation based on different demand scenarios. Finally, we will use all the discovered information to select a final system configuration, and implement this in our optimization model to evaluate the current fleet used for the distribution.

All models and scenarios for both algorithms in the two-phase model are run on the same computer, an Acer Aspire S 13 with Windows 10, two 2.3 GHz Intel Core i5 6200U processors, and 8 GB memory (RAM).

8.1 Planning Levels

As stated in the model characteristics in Section 5.1, the feed distribution problem must be considered on several planning levels when it comes to the determine the optimal distribution plan and fleet configuration. When evaluating different system dimensions to determine whether they are suited for further analysis, acceptable solution times are dependent of which level of planning we are looking at. Starting at the strategical level, we have decisions of how to allocate the key resources to pursue the firm's strategy. We see this as the determination of how the composition of the fleet of vessels should be, and since this is at the top level in the planning process, no time-frames should restrict this evaluation. We will look further into this in later sections. The two other planning levels we are looking at are tactical decisions and operational decisions. The first one includes how we decide to model the full and complex system. This mainly means the number of clusters we choose to represent all farms in the system. The latter involves how we want the fleet to operate in the chosen system, which is what we are mainly focusing on by looking at the routing scheduling for the system in detail.

Because the core purpose of the final model is to function in the operational stage as a decision support tool, the solver time should be as low as possible without having major impact on the optimal solution. However, this is also dependent of the tactical planning level as the routing optimization is affected by the number of customers to serve. It is important to be able to make rapid decisions in the fish feed distribution planning, due to the large uncertainty in the industry. It is also important to prevent ships from waiting for decisions of where to sail, causing possible necessary time for sailing on an optimal route to be spilled. We because of this explore the most practical dimensions for our system configuration to get output solutions within practical time

limits. By practical limits we mean a short enough solution time so that a ship is not extensively withheld when decision must be made. Therefore, after the system configuration is determined, operational decisions should be produced within a computational time frame of ten minutes. We will explore larger unconstrained configurations with longer solving time as well, to see how the different restrictions affect the solution time, but we set an absolute solving limit of one hour.

It should also be mentioned that the solution time is heavily dependent of the computer hardware that the models are run on. Even though the experiments could have been run on more powerful computers, we have run all models on the same private machine for simplicity.

8.2 Test Cases

We will first apply our optimization model to a smaller, artificial system to make sure that we get the desired outputs, and to familiarize with the process and how the model works. The main characteristics that are restricting our problem, in a fixed period of one week, are the number of customer nodes and the number of ships. The test cases we explore are systems with 10 and 20 clusters, served by two ships. The clusters are not generated based on any of the methods discussed in this paper, but there is coherence between the two artificial systems.

Since these are not real systems, we will make coarse approximations for values when it comes to distance between the clusters, service time when visiting a customer, and for the total demand of each customer. Even though the data is not real, the values are representing a system with similarities to the real world system in question. This means serving customers both north and south of the production factory, with some being much further away than others. The clusters are numerated from north to south, where cluster 1 is furthest north. The ships are homogeneous and able to serve all clusters. The planning period is set to one week, and the possibility of using external delivery is included. The demand is determined based on randomly generated values between the minimum and maximum possible feed consumption of a farm (25-300 tonnes), and the total demand of the system is similar in the two cases. In Table 3, the main dimensions of the artificial systems are presented, along with the total number of possible routes for each ship, generated by the route-generation algorithm. The calculation time includes both the time for generating the routes, and the time for finding the optimal combination of routes for a final solution. It should be noted that the distance matrices are not proportional between the two systems. This causes the outputs to not be directly comparable, but as mentioned above, this done to familiarize with the procedure and look at how it behaves for different sized systems. The full input data-sets for the test cases are attached with the delivery of the thesis (see Appendix A for list of attached files).

Table 3: Test-system dimensions and generation times.

Test	Ships	Ship Capacity	Clusters	Max Routes per Ship	Calculation Time
1	2	1 500	10	156	≈ 0 s
2	2	1 500	20	24 804	171 s

We see that the number of possible routes per ship increases drastically for the same time period when splitting the system into 20 clusters instead of ten. This is expected, as there are naturally more possibilities within the same time period. The calculation time also increases, going from almost immediately solving the problem, to a total calculation time of 171 seconds, where the largest portion of the computational time comes from the route generation part. This is an expected behavior of our clustered system as well, and it is the reason for further study. We know that including more customer nodes increases the solution space exponentially, but also the selection of routes for an optimal solution takes a significant amount of time, even for this small system.

Next we look at the optimization part, where the best combination of routes to meet the total demand within the constraints are determined. The test case results are presented below, and are in the same format as the final solution will be. The output shows which ships are serving which customer, how many routes each ship sails, the total demand and cargo delivered, along with customer capacity, time spent on each route, and the cost for each route. It also shows if any charter ships are being used, and at what cost. Finally, the total cost of the distribution plan is presented. The optimal routing solution for the two artificial systems can be found in Figure 21 and Figure 22.

```

Feed Distribution Plan
=====

```

	Ship 1	Ship 1	Ship 2	Ship 2	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	0	1041	2700
Cluster 2	---	---	---	826	826	826	1800
Cluster 3	---	---	---	528	528	528	1200
Cluster 4	---	---	510	---	510	510	1500
Cluster 5	---	---	964	---	964	964	2700
Cluster 6	726	323	---	---	1049	876	2700
Cluster 7	330	---	---	---	330	330	900
Cluster 8	---	1177	---	---	1177	1177	3900
Cluster 9	444	---	---	---	444	444	1200
Cluster 10	---	---	---	---	0	1267	4000
TOTAL	1500	1500	1474	1354	5828		
Ship Capacity	1500	1500	1500	1500			
Sail time	93	74	53	67			
Cost (€)	8511	6813	4910	6171			

```

-----
Cluster 1 supplied by Charter Ship 1 for 10000 €

Cluster 10 supplied by Charter Ship 3 for 10000 €

Total cost of distribution = 46405 €.

```

Figure 21: Test case 1 solution output.

Feed Distribution Plan

	Ship 1	Ship 1	Ship 1	Ship 1	Ship 2	Ship 2	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	0	252	1200
Cluster 2	---	---	402	---	---	---	402	402	1200
Cluster 3	---	---	288	---	---	---	288	288	1200
Cluster 4	---	---	310	---	---	---	310	310	1200
Cluster 5	---	---	---	---	---	873	873	873	1200
Cluster 6	---	---	---	---	---	567	567	567	1200
Cluster 7	---	---	219	---	---	---	219	219	1200
Cluster 8	---	---	---	498	---	---	498	498	1200
Cluster 9	---	---	112	---	---	---	112	112	1200
Cluster 10	---	752	---	---	---	---	752	752	1200
Cluster 11	---	354	---	---	---	---	354	354	1200
Cluster 12	168	---	---	---	---	---	168	168	1200
Cluster 13	547	---	---	---	---	---	547	547	1200
Cluster 14	711	---	---	---	---	---	711	711	1200
Cluster 15	---	---	---	---	236	---	236	236	1200
Cluster 16	---	---	---	---	416	---	416	416	1200
Cluster 17	---	---	---	---	---	---	0	632	1200
Cluster 18	---	---	---	---	264	---	264	264	1200
Cluster 19	---	---	---	---	150	---	150	150	1200
Cluster 20	---	---	---	---	280	---	280	280	1200
TOTAL	1426	1106	1331	498	1346	1440	7147		
Ship Capacity	1500	1500	1500	1500	1500	1500			
Sail time	53	26	73	15	114	36			
Cost (\$)	4912	2414	6659	1443	10435	3315			

Cluster 17 supplied by Charter Ship 1 for 10000 \$

Cluster 1 supplied by Charter Ship 2 for 10000 \$

Total cost of distribution = 49178 \$.

Figure 22: Test case 2 solution output.

From the outputs we see that when splitting the system into ten clusters, both ships sails two routes each, but for the 20 clusters system ship 1 sails four routes. By comparing the distribution plans, we also see that both systems include the use of two charter ships, even though these are more costly than any of the routes served internally. This is an expected behavior, as the charter ship should preferably only be used for the most expensive customers, which for this configuration is the clusters with the lowest and highest numbers (the ones furthest away).

With a planning period of one week (168 hours), both scenarios are dependent on charter ships to serve all clusters. In both scenarios, Ship 1 almost uses the full planning period, with 167 hours used for sailing, while ship 2 only uses 120 hours in the first scenario and 150 hours in the second. These are good indications that the ships are sailing optimal routes, exploiting the time period. In case 1, ship 2 has 48 hours left, but because the shortest route possible takes 53 hours, it is not possible to assign the ship to another route inside the current period. In case 2, the time left is reduced to 18 hours, which is due to the change in routes and the difference in internal sailing time between the two cases. With larger cluster sizes, each cluster contains more farms, increasing the total sailing time on each route for the same number of visits. With smaller cluster sizes, more and shorter routes exists, and a more optimal time usage for each

ship can be obtained.

The difference in costs between the systems comes from the difference in service time, and in this case, differences in sailing distances because of non-consistency when creating the matrices. In general, for smaller cluster sizes the cost is more accurate due to a more realistic description of the system. This is mainly because of the internal sailing time estimation, since we have calculated this with a more uncertain method. In general, with larger clusters, the center of the cluster is more inaccurately placed, and hence a distance estimation between these are also more inaccurate than for smaller clusters. Both these factors are affecting the results when it comes to the cost calculations, since the costs are proportional with distance. Because of the non-consistency in the distance matrices in the test cases, the cases are not directly comparable in this specific case, but in general there will be a difference in costs between different cluster configurations.

The total demand is equal for the two systems, but we can see that the total amount of delivered feed internally is 5 828 tonnes for the case with 10 clusters, and 7 147 tonnes for 20-clusters system. In the first case, the charter ships are transporting more feed than in the second case, leading to a larger total delivery of feed in the ten-cluster case. The reason for the larger amount of feed delivered by charter ships in case 1, is due to the larger cluster sizes. When a charter ship is used for the delivery, the whole demand of the cluster is met, and hence when the clusters are bigger, the more feed must be delivered at each visit. Because the change in cluster sizes affects the time spent internally in each cluster, the optimal solution is different for the cases. In case 1, some customers also receive more than their demand, while this is not the scenario in the second test case. An explanation for this is that with larger clusters, it is more important that the ships are fully utilized. This is because the demand of each cluster is much higher, and each ship can visit less clusters on each route. Therefore, selecting routes where the remaining cargo can be delivered to another cluster (discussed in Section 5.3) are more optimal. In addition, the unloading time in a cluster is calculated as the demand of each cluster divided by the unloading rate, meaning that delivering more feed than ordered does not affect the time usage in our model. For the second case with smaller cluster sizes, none of the selected routes delivers more feed than ordered. Here, the internal sailing time is smaller in each cluster, and the demand of each cluster is lower. Clusters therefore do not need to be visited more than once during the planning period, as ships can deliver the ordered amount of feed to each cluster on one route. Visiting the same clusters several times during the same planning period is time-consuming and thus avoided if possible.

Even though the model seems to work and is making reasonable choices when selecting routes, there also seem to be some issues. The ships are not fully utilized, with some routes only transporting a small amount of cargo. The problem may come from the fact that each ship is only delivering exactly the amount that are demanded during the period. If we could deliver more on each route so that each ship is fully utilized on every route, it would lower the demand in the next period. However, this could be solved by looking at the ships with remaining cargo capacity, and have the feed distribution planner manually assign cargo to fully utilize the ships

without exceeding any storage capacities. By doing this we will also lower the demand for the next period as more cargo will be at storage at the customer sites.

One of our main concerns is the computational time. Because the models are created to be used in the operational stage, where rapid decision must be taken, a low computational time is critical. With 20 clusters, the computational time is 171 seconds, which is reasonable. However, for real system configurations we want to include both more ships and clusters. Because we know that the computational time increases exponentially with the problem size, we will in the next section include some additional constraints to be able to keep the calculation time down when we are expanding the system.

8.3 Adjustments and Possible Improvements of the Model

Based on the test cases performed in the previous section, we see that both solution space and solution time seems to increase rather drastically when splitting the system in more clusters. Because of this, we want to make some adjustments and possible improvements. In general, it is desired to make the model formulation as tight as possible. That means adding more constraints to the model, which can lead to fewer possible routes, and thus reducing the solution time. Also, even though the first model generates valid solutions, they may not be 100% realistic. By restricting the behavior of the model further, we can get a model that better represent the real world. We are only adding restrictions to the optimization model, as the route generation algorithm is harder to modify because it is created based on previous work performed by SINTEF Ocean. This means that it is only the solution time of the optimization model that is being reduced.

The first thing we want to include is a pre-defined maximum number of routes for a ship to be able sail during each period. Since we are only looking at a short period of time, it is not natural for a ship to visit the factory a large number of times during this short period. In the test case, one of the ships is sailing four different routes, including four loading processes at the factory. In the real world, a ship will normally perform one or two trips each week. By including a constraint for maximum number of trips, the distribution planner can fix the optimization model to behave in a desired way. By adding this constraint, the model will also not consider solutions with many short routes and thus routes with high utilization will be selected. We thus add Equation 8.1 to the mathematical model in Section 5.4.1.

$$\sum_{r \in R} x_{vr} \leq n, \quad v \in V \quad (8.1)$$

Constraint 8.1 sets a maximum limit of n routes that each ship can sail during the planning period.

To add more flexibility during seasonal consumption variations, we also want to include a constraint that state the maximum number of ships that can be used for performing the distribution

in the current time period. By including this, we have the opportunity to reduce the fleet size in lower demand periods, and hence reducing the total cost drastically. In order to set a minimum number of ships that must be used in the model, a new variable w_v is introduced. This is a binary variable, which is 1 if ship v is used, and 0 otherwise. Since w_v gives information about whether ship v is used or not, it has to be connected with x_{vr} . If one or more routes are used by vessel v , w_v must be set to 1, and if not it has to be set to 0. In order to connect the two variables, a relative big value M has to be included, as seen in Restriction 8.2. By including the big value M , w_v must be one if the sum of all routes for x_{vr} is one. Restriction 8.3 specifies that the maximum number of ships to be used during the planning period must be less or equal to m number of ships. Restriction 8.4 is the binary constraint, saying that w_v must be 1 or 0.

$$\sum_{r \in R} x_{vr} \leq Mw_v, \quad v \in V \quad (8.2)$$

$$\sum_{v \in V} w_v \leq m \quad (8.3)$$

$$w_v \in \{0, 1\}, \quad v \in V \quad (8.4)$$

From the test cases we also see that not all ships are fully utilized at all times. This is because of the way the route generator is created, meaning that a customer only needs to receive its exact demand, not necessarily fully exploiting its capacity. To better utilize each ship's cargo capacity, we could introduce split deliveries or awarding utilization, as discussed in Section 5.3, but this would increase the calculation time, and the number of possible routes would drastically increase. Since this is only meant as a decision planning tool, we leave it for the planning manager to manually assign cargo to available capacities to better utilize the ships, after the optimal distribution plan is created. This simplifies the model, but also adds flexibility as the planning manager can assign cargo to where it is required rapidly.

When the model with the new constraints is run on the same test cases with maximum number of routes set to four, it is a significant drop in calculation time. The optimal solution for test case 2 was found in 13 seconds, compared to the unmodified version which found it in 18 seconds. This equals an improvement of 28 %, which we see as more than sufficient for including the alterations presented in this section in the final model. This drop in calculation time is a desirable effect, in order to be able to run the model for as many alternatives of clustering techniques as possible in a reasonable amount of time. We now see the model as finished, and will from here use the described altered model to perform further studies of the real distribution system. The final mathematical formulation can be viewed in Appendix C.

8.4 Determination of Cluster Sizes

We will in this section perform different studies to determine the best suited dimensions and system configuration for further evaluation. We will look at the three clustering methods discussed in Section 6.2, and vary parameters to see how it affects the solution space and solution time to be able to create the most user-friendly tool for the system possible.

As the clustering techniques now are determined, and most parameters are fixed, we next need to determine the number of clusters that the farms should be collected into. The total system has 151 farms along the coast. This gives us a lot of cluster opportunities, both with identical clusters, and non-identical. The regional clustering method already has a fixed number of clusters (11), but the two other methods are yet to be determined. In the start of the investigation, we keep the demand for the whole system at a fixed level of medium demand, explained in Section 7.2. We also give the system full compatibility, meaning that no constraints of which ships can visit which customers are included at this point. This is because we only want to look at the effect of the number of clusters included in each system configuration.

First, we test the model for all 151 farms for a time period of one week, with no constraints for ship/cluster compatibility. When running the route generation algorithm for this case, the calculations keeps running until the computer runs out of memory. This is due to the large data set that gives an enormous solution space when no compatibility constraints are included. We have now verified that it is not possible to calculate optimal routes between all farms with no constraints in the problem, hence clustering methods are needed.

When the regional sectioning is used as clustering method, the 151 farms are divided into the 11 contamination regional clusters based on their location. The division is completed by listing all the farms from north to south (shown in Appendix B) and manually giving each farm their selected cluster affiliation. This is a fixed system where we do not need to perform any computational study when it comes to number of clusters to include. The configuration gives us 177 possible routes, calculated within the first second. Also the optimal routing solution is found in an equally short amount of time. This shows us that this system configuration is easily solved due to the low number of customer nodes. In fact, there are so few customers that the demand of each cluster exceeds the largest ship's capacity, meaning that the clusters must be visited more than once during the planning period, and hence limit the possibility of optimizing the system. We will look further into this configuration with other input parameters in the next sections.

Next, we will test a variety of cluster sizes for the k-means clustering technique and for the natural clustering technique, to investigate how it affects the solution space and solution time. All cases use the same fixed fleet of vessels, have an equal total demand, and all distances and service times are determined with the approach described in Section 7.3. Table 4 and Table 5 show the cases we have studied, along with the time for generating all possible routes, and the time used for finding the optimal routing solution.

Table 4: Calculation times for natural clustering.

Clusters	Max Possible Routes	Generation Time	Solution Time
18	3 461	1 s	4 s
28	-	-	-
23	22 205	147 s	27 s

Table 5: Calculation times for k-means clustering.

Clusters	Max Possible Routes	Generation Time	Solution Time
50	-	-	-
30	-	-	-
23	27 357	232 s	29 s
25	-	-	-

A "-" in the tables means that the model was not able to find a solution within the computational time limit of one hour. We clearly see that this is the case for many of the system configurations, and hence it is not relevant to proceed with these cases. However, we will in the next section investigate how the solution time is affected by changing the compatibility matrix, and hence some of the unpractical cases can be used for the new alterations. We therefore choose to proceed the investigation of all the unsolved cases, and we also keep the two systems with 23 clusters created with different methods.

8.5 Manipulation of Compatibility Matrix

As the previous section shows, the number of possible solutions increases drastically when the number of customer nodes is increased, and many alternatives were not able to produce an output within the time-limit. To be able to include a larger number of clusters, and at the same time reduce the number of possible routes, we will try different manipulations of the compatibility matrix. In the current fleet, two of the vessels are smaller and have a lower operational sailing speed. Because of this, they will preferably serve the clusters closest to the factory. The two faster ships run on LNG, which are both more environment friendly, and have lower operational costs because of a lower fuel price. They are thus more optimal for visiting the clusters furthest away from the factory. This can be achieved by manipulating the compatibility constraints. In this way, we can create a system which reflects Marine Harvest's current routing plan more accurately, as they have assigned specific regions for their vessels. To investigate different approaches, we choose to define three categories for the compatibility matrix; strict, medium and full. By doing this we can explore the effect of reducing the computational time with compatibility versus the effect of reducing it with reduction in number of nodes.

In the strict category, the two smaller ships are set to only be allowed to sail to clusters less than 100 nautical miles north of the factory, or less than 200 nautical miles south of the factory. The two larger vessels are fixed to a pattern where one is only serving clusters north off 100 miles from the factory, and the other only clusters south of 200 miles from the factory. The reason for

this specific selection of southern and northern sailing borders of 200 and 100 nautical miles is to adjust for the density of farms along the coastline. By selecting these distances, the number of farms both south and north of the factory is about the same, around 40 farms. In the medium compatibility case, the smaller ships can sail up to 250 nautical miles away from the loading port, while the larger vessels can visit all customers except clusters closer than 100 nautical miles from the factory. The full compatibility configuration is a case without restrictions, hence giving all ships the possibility of visiting all customers in the system (same as in the previous section).

Tables 6 and 7 show different cases of cluster sizes and use of compatibility matrices for both the natural clustering technique and the k-means approach. The regional contamination system is still fixed with 11 clusters, and we do not see the need of including any compatibility restrictions for this configuration, as the solution time is already low.

Table 6: Calculation times for natural clustering with compatibility restrictions.

Clusters	Compatibility	Max Possible Routes	Generation Time	Solution Time
28	Mid	6 305	3 s	21 s
28	Strict	418	≈ 0 s	1 s
23	Mid	1 650	1 s	4 s

Table 7: Calculation times for k-means clustering with compatibility restrictions.

Clusters	Compatibility	Max Possible Routes	Generation Time	Solution Time
30	Mid	14 111	43 s	24 s
30	Strict	1 625	1 s	2 s
25	Mid	6 985	5 s	20 s
50	Strict	-	-	-

Compared to before introducing compatibility constraints, we now see that we are able to find optimal solutions for most cases within an acceptable time. From the tables we can also see that the number of maximum possible routes for a ship, generation time, and solution time, all have dropped drastically compared to before.

By studying the 23-natural cluster case with medium compatibility, we can see that the maximum number of routes is reduced from 22 205 in the full compatibility case, to 1 650 for the altered configuration. This affects the calculation time, reducing it with 99%, to only a total calculation time of five seconds. Two of the cases, the natural clustering method with 28 clusters, and the k-means clustering method with 30 clusters, provides a solution within the computational time limit when changing the compatibility constrains from full to medium. The 50-cluster system is still too large to produce any output inside the desired time, even with strict compatibility restrictions. We also see that when the compatibility matrix is strictly defined, there are fewer possible solutions and the calculation time drops. This shows us that this configuration may be too strict, removing routes that could have been included in an optimal routing schedule.

As expected, the introduction of compatibility constraints generally reduces the calculation time by reducing the number of possible routes. Even though it provides the possibility of including an increased number of clusters, it might also remove optimal routes. Because of this, it is desired to include configurations with both the use of compatibility constraints, and without it, in further evaluations. In addition, it is also desired to further explore alternatives with the highest possible number of routes that can generate optimal solutions within the acceptable solving time. The natural clustering method with 23 cluster and full compatibility, along with the k-means method with 30 cluster and medium compatibility constraints, are therefore chosen for further investigation in the next section. In addition, the 11 contamination regions are included, as it can be a relevant scenario in the future. By including all three clustering methods, we can compare the results and evaluate which method that represent the real world best. In Figure 23, an overview of the selected configurations is presented.

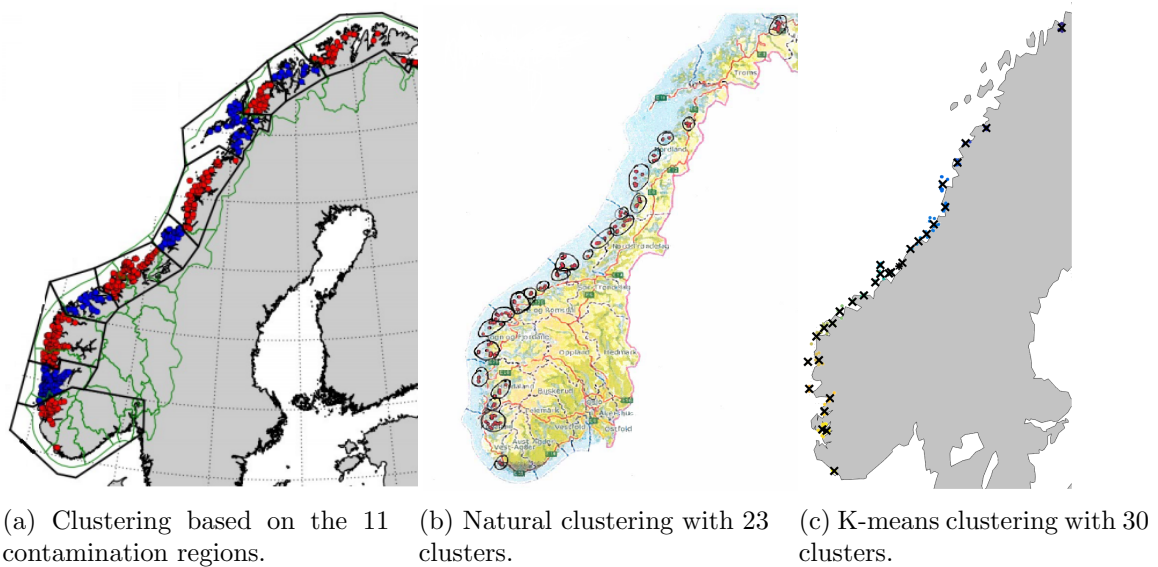


Figure 23: Overview over selected clustering methods.

8.6 Consumption Scenarios

To evaluate and compare the different methods and fleet configurations, we start by defining three consumption scenarios involving the four consumption areas described in Section 7.2. By fixing each of the four zones to a certain consumption rate, meaning that all farms inside this zone will have the same consumption rate, we can compare the different clustering approaches for variable rates. The different scenarios can represent different stages in the farming process, as the feed consumption ratio increases with the size of the fish, and they can be used for representing seasonal variations. By comparing results for the different scenarios, we can investigate which model that best suits the real feed-distribution system overall. Table 8 shows the consumption rates for each of the four defined zones in the three different scenarios.

Table 8: Consumption scenarios based on consumption in zones.

Scenario\Region	N	M	W	S
Low scenario	Low	Mid	Mid	Low
Mid scenario	Mid	Mid	Mid	Mid
High scenario	High	Mid	Mid	High

These three scenarios are chosen to test the limits of the distribution system. Especially the high scenario where the two regions furthest from the factory are given the highest demand. This means that the vessels supplying these farms needs multiple trips to serve all farms within the planning period and thus stressing the system. The low scenario will test the calculation time of the route generation algorithm, since lower demand will cause a higher number of possible routes per ship. The low demand scenario also provides the opportunity to test if all vessels are needed during low season, or if it is beneficial to reduce the sailing speed. The mid scenario is mostly used as a reference for comparing the high and low scenarios to a basis. It should be noted that the scenarios are extreme cases that would most likely not occur in the real world as the farmers want fish at different sizes all year to constantly be able to harvest. The demand in low season is only 30% of the high season, which is the limit of what occurs in the real world, but we will use these values to evaluate and compare the different configurations.

Medium Demand

In the previous section, three methods with different cluster sizes and compatibility constrains were chosen for further evaluation. All three methods were performed with medium consumption rates, with the results shown in Table 9. The full distribution plan for each method can also be found in Appendix D.1.

Table 9: Distribution costs for the medium demand scenario.

Method	Regions	Natural	K-means
Number of clusters	11	23	30
Compatibility	Full	Full	Mid
Ships used	4	4	4
Charter ships	1	0	0
Calculation time	1 s	174 s	67 s
Total cargo transported	23 700 t	22 650 t	22 650 t
Total cost	69 013 \$	41 857 \$	41 485 \$

From Table 9, we can see that there is a large variation in cost between the three methods with the same consumption rates. It varies between 69 013 \$ and 41 485 \$ for the different methods. The largest deviation is between the contamination regions method and the two others. The reason for the high cost of the contamination regions method, is because the distribution includes the use of one charter ship at a cost of 23 000 \$. With few regions, as in this method, the demand of each region is higher. This causes the demand of some clusters to be higher than the ships capacity, and hence it is more optimal to include charter ships for delivering the full demand than using multiple visits. When excluding the price of the charter ship, we can see that the

total cost is more similar to the other cases. Because of the difference in the number of clusters in these methods, the small difference in costs is a good indication that the internal sailing distance calculations are applicable.

One important factor to compare between the different methods is the utilization of the available time. This should be as close to 100% as possible, since it is not desired to keep any vessels stationary. By studying the feed distribution plans for the different methods in Appendix D.1, we can see that all ships, except ship 3 in the k-means method, are operating in over 75% of the planning period. One of the reasons for the low time usage of ship 3 in the k-means method is the inclusion of compatibility constraints, which constrain ship 3 and 4 to only supply the clusters closest to the factory. Because of the limited number of clusters to supply, the vessels will be restrained from supplying other clusters. This means that the ships are forced to wait the remaining time of the planning period after finishing the distribution to their possible farms. Even with the low utilization of ship 3, the k-means method manages to supply all clusters with medium demand without the use of charter ships. Looking at the natural clustering method without compatibility constraints, we can see that the total sailing time is better spread out between the vessels, compared to the k-means method with compatibility constraints. Here all ships are completing their routes within a time difference of 16 hours. This means that the distribution planner can start the next planning period earlier than by using the other methods.

An illustration of the routing solution of the k-means clustering method with 30 clusters and medium demand can be seen in Figure 24.

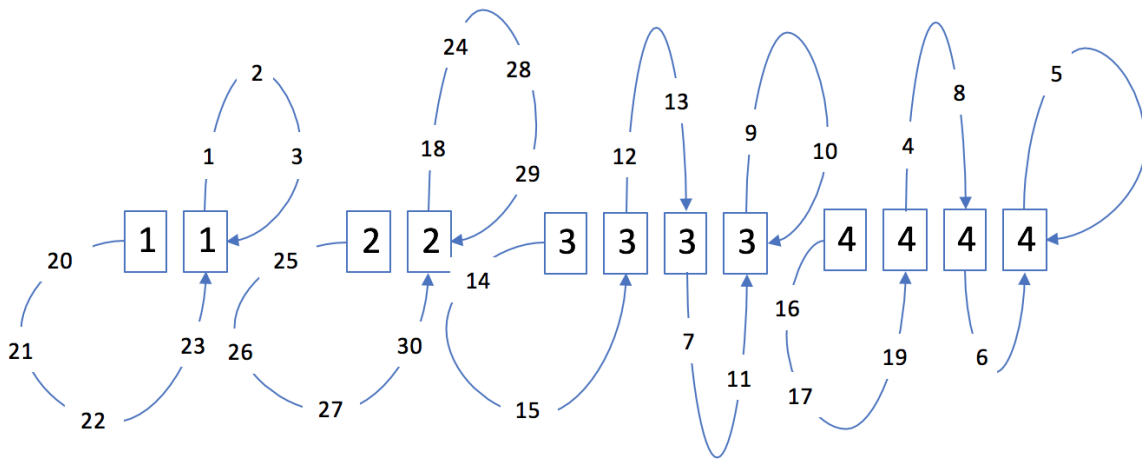


Figure 24: Routing illustration for k-means method with 30 clusters and medium demand.

We also notice that the total sailing time for all ships combined decreases with an increasing number of clusters. It does not necessarily only come from increased internal sail distance for larger clusters, but may also come from differences in the compatibility constraints. This causes different routes to be selected for the ships, and thus different sailing patterns and sailing distances. This comes clear by comparing the natural clustering method with the k-means clustering method. In the full compatibility case, the smaller ship 4 is used to serve the cluster

furthest north, while ship 1 and ship 3 are used for the clusters closer to the factory. Because ship 4 is sailing a long distance, and has lower operation speed and cargo capacity, it only manages to visit two clusters. This gives a larger work load for ship 3, leading to a higher utilization of the planning period than for the k-means method. In this method, different routes are selected, with both ship 3 and 4 serving the same area, while ship 1 and 2 are forced to visit the clusters furthest from the factory. The cost is about the same, but the time-utilization is different. For the method where the contamination zones are used for creating the clusters, a charter ship is used to serve the regions furthest south, while ship 1 is serving regions in the north. With this clustering method, the clusters include so many farms that only one cluster can be served on each route, and sometimes one visit is not even enough. This affects the routing, as some clusters may need multiple deliveries, which increases the total cost. Even though the cluster sizes are different, the total number of farms visited on each route is about the same in all three cases. Each route includes serving between 10 and 20 farms, depending on the ship type, which according to Marine Harvest is similar to the real world.

In order to reduce costs by lowering the demand in the next planning period, the utilization of the cargo capacity of the ships should also be as high as possible. Because the objective of the optimization model is to reduce operation cost, which depends on the sailing distance, the model will utilize the ships as best as possible. We can see in all three methods, that the cargo capacity utilization is relatively high, and in most of the routes the cargo capacity is fully utilized. Table 9 also shows that for both the natural clustering method with 23 clusters, and the k-means method with 30 clusters, the total transported amount of feed is 22 650 tonnes, which is exactly the quantity ordered in total. For the clustering based on the 11 contamination zones, the transported amount of feed is higher. This is because some of the clusters needs multiple visits in order to supply the demand. On the second visit, the remaining demand is lower than the total capacity of cargo on-board, and hence more feed is delivered than necessary. Even though this is more costly in the current time period, it will lower the total demand in the next period as the cluster will have more feed in storage. For the natural and k-means methods, we can see that the cost to transported feed-volume ratio is almost the same, while it is higher for the contamination regions. This is because of the external supply of feed for cluster 11, which is more expensive than being internally supplied. Since the model is meant as a planning tool, the distribution planner can allocate the remaining cargo capacity to supply extra feed to clusters which have capacity to receive more than ordered. This will help reducing the transported volume and cost ratio even more, and also lower the total demand in the next planning period for all methods.

High Demand

Next we want to investigate how the fleet is able to handle a scenario with an extremely high demand, and if there are any differences between the clustering methods. Table 10 shows the main parameters for each of the three selected system configurations when the total demand of each system is set to the high scenario. The routing schedules, along with the individual route costs and sailing times, for the natural and k-means methods can be found in Appendix D.2.

Table 10: Distribution costs for the high demand scenario.

Method	Regions	Natural	K-means
Number of clusters	11	23	30
Compatibility	Full	Full	Mid
Ships used	-	4	4
Charter ships	-	2	2
Calculation time	-	97 s	44 s
Total cargo transported	-	30 250 t	30 250 t
Total cost	-	58 546 \$	59 234 \$

There will naturally be fewer possible routes for the high demand scenario and hence a decrease in calculation time. The reduction in calculation time is around 50% compared the medium demand scenario. For the regional method, the amount of cargo required becomes so huge that the model is not able to find a possible solution with the fleet in hand and the maximum number of charter ships. As before, this is due to the large number of farms inside each cluster. For the two other methods, the total operational cost for the high demand scenario increases compared to the medium demand scenario, as expected. This is mainly because of the need of charter ships to supply all clusters in both the natural and k-means method. Both methods need two charter ships to serve two different clusters. In the natural method, the charter ships serve cluster 1 and 10 for a total cost of 10 000 \$, while the k-means method uses the charter ships to serve cluster 1 and 30 for a total cost of 17 000 \$. The reason for the large difference in charter prices, is that cluster 10 in the 23-cluster natural method only contains one fish farm, while cluster 30 in the k-means method contains 8 farms. However, even though there are a large difference in charter costs, the total costs are about the same for both methods, with the natural method only being 688 \$ cheaper. The increase in costs is also a result of transporting more feed, leading to longer unloading times at each cluster, and hence an increase in total sailing time for each route.

Compared to the medium demand scenario, we can see that the ships are sailing a higher number of routes in total to supply the higher demand. This is expected as more feed must be delivered. With the increase in demand, the number of clusters visited on each route becomes lower and more trips back to the factory for loading are needed. This can be seen by comparing the total number of routes needed to supply the southern clusters in the k-means method (cluster 20 to cluster 30), where the demand increases with 41% from 8 550 to 12 050 tonnes of feed. In the medium demand scenario, three routes were needed to meet the demand. Ship 1 with one trip serving four clusters, and ship 2 with two trips serving four and three clusters respectively. In the high demand scenario, the number have increased to four trips, in addition to one cluster

being served by a charter ship. Here ship 1 must sail two routes, visiting two and four clusters, while ship 2 and ship 3 are sailing one trip each, serving two clusters each.

By looking at the natural method with full compatibility, we can also see a change in the routing patterns with the increase in demand. The largest change is for the two smallest ships, which are now visiting clusters further from the factory than in the medium demand scenario, causing a better utilization of the planning period. Ship 4 is also sailing two more routes in the high demand scenario, increasing the total sailing time with 26 hours. In the medium demand scenario, the total length of the planning period is not needed to be able to supply all clusters. This causes the cheaper and faster vessels to be used to a greater extent than the other ships. In this high demand scenario, all vessels and all the available time is needed to supply all clusters, affecting the routing schedule to be more costly. We also see a change in the patterns for one of the largest vessels. In the medium scenario, ship 2 is used to serve customers south and north, while with higher demand it visits clusters in the south and middle regions instead. This shows the flexibility in the non-restricted model, as it changes distribution patterns with demand scenarios.

Because of the increase in demand, the total time used for completing the feed distribution increases. For the natural method with high demand, three of the ships uses 168 hours of the 169 hours planning period, while the last ship uses 167 hours. In the k-means method, the sailing time is almost similar for both the medium and high demand scenarios. This is due to the compatibility constrains. The constrains prevent the two slowest ships, ship 3 and 4 from supplying the clusters furthest north and south. It gives them a restricted area of clusters to supply, which they manage to visit within a shorter time than available. Since they are not able to supply any more clusters, they must wait at the factory while the two other ships continue to deliver feed. The compatibility constrains are therefore preventing more optimal solutions where the available time is better utilized. Because of the constrains, clusters far from the factory must be supplied by charter ships at a high price compared to the natural clustering method without compatibility constrains, where cheaper routes for the charter ships are used. By comparing the two methods, we can see that the total sailing time for the internal fleet is lower for the k-means method. This is because the charter ships included in this method serves more farms, which also are further from the factory. This again leads to a shorter sailing distance in total, less farms to supply for the fleet, and thus a lower utilization of the planning period. Since this is an extreme case with unnatural high demand, the inclusion of the two charter ships is realistic and would most probably also be necessary in the real world.

Low Demand

For the extreme low demand scenario, the number of possible routes becomes so large for the fleet in hand that the calculation times exceed the one hour mark. To be able to investigate how the routing should be done in this scenario, an extensive study should be performed to determine the optimal fleet size for this particular distribution. To compare the selected methods, we choose to shorten the planning period to four days in the natural and k-means method (marked with a * in the table below). This is done since the demand should be met in a shorter amount of

time by using the same fleet, now that the total demand is drastically lowered. By reducing the planning period, the number of maximum possible routes is reduced, and a solution can be obtained within the time limit. Table 11 shows the distribution characteristics for the low scenario for the three methods. The full routing schedules for all clustering methods can be found in Appendix D.3.

Table 11: Distribution costs for the low demand scenario.

Method	Regions	Natural*	K-means*
Number of clusters	11	23	30
Compatibility	Full	Full	Mid
Ships used	4	4	4
Charter ships	0	2	3
Calculation time	1 s	219 s	82 s
Total cargo transported	17 021 t	17 028 t	16 950 t
Total cost	44 347 \$	38 854 \$	44 415 \$

In general, we want a solution where the total cost is lower than in the other scenarios because of the decrease in transported quantity, even for the two methods where a shorter planning period is used. We see that this is not the case for the k-means method, where the cost is higher than in the medium demand scenario. This is because the optimal solution uses three charter ships to be able to deliver the required cargo at a cost of 17 000 \$. From the routing schedule in Appendix D.3, we see that the two clusters furthest north are served by charter ships. This means that the two largest ships do not have enough available time to supply the clusters furthest from the factory within the reduced planning period. The last charter ship is used in the middle region to help the two smaller vessels. Here both vessels are in operation throughout the planning period and in order to be able to serve all clusters a charter ship is needed.

For the natural method without compatibility constraints, two charter ships are also needed, but the total cost is significantly lower than for the medium demand scenario. This is because the ships are not restricted to certain areas and hence have more flexibility to meet the altered demand, and one less charter ship is thus needed. By closer studying the routes of the vessels for this method, we can see that one of the largest ships are used to supply the southern clusters, while the other large ship serves clusters closer to the factory. Ship 4 is now used to serve northern clusters, except the clusters furthest north, which is being served by a charter ship. This shows us that when constraints for sailing patterns are included, the limitation of freedom causes the solution to be more expensive, and that the fixed compatibility is not optimal for all cases.

Looking at the method based on the contamination regions, where the planning period is set to one week, the total demand is fully delivered internally. The cost however, is higher than with the natural method, but lower than in the medium demand scenario, which is a more realistic result. The reason for the higher cost compared to the other methods, is again the larger cluster sizes which gives an unnatural high internal sailing time, and because of the large collection of farms inside each cluster, the method still does not have enough possibilities to be optimized in

a good way.

The use of charter ships would most probably be unnecessary for both methods if we had not reduced the planning period. This can be seen by studying the total sailing time for all four ships. In both the natural and k-means method, the time usage for each ship is within 90% of the available time. The reduction of the planning period was introduced in order to reduce the computational time, but to deal with this problem the distribution planner can manually add clusters not visited to new routes outside the planning period if charter ships are not desired. If this schedule proves to be cheaper than the inclusion of charter ships, it will help reducing the total cost additionally.

Another thing to notice, is that the total number of routes is reduced compared to the other scenarios because of the reduced demand in the low demand scenario. With lower demand, each ship can visit a larger number of clusters on each route. When studying the method based on the contamination zones, which is the only method with the same planning period in all scenarios, we can see that the total number of routes is reduced by two, compared to the medium scenario. This is because the number of routes that includes more than one cluster-visit in the low scenario, is higher than in the medium demand scenario. This also introduces the possibility in this scenario to reduce the size of the fleet temporary. It is possible that only three of the vessels in the fleet would be able to meet the lowered demand alone, and hence the last ship could be used elsewhere. It could either be hired out to another company, or do maintenance or upgrades. However, for the k-means method with compatibility, this would create a whole new compatibility configuration which would lead to a non-consistency in the sailing patterns and cause direct comparisons to be inaccurate. Instead, alterations to the existing fleet size will be considered in Section 8.7.2. In general, the average cost of each route is higher in all methods compared to scenarios where less feed is delivered. This is because of the increase in the length of the routes, since each ship can serve more clusters on each route, and hence the total sailing distance naturally increases. Looking at the schedule for the natural method, one of the largest ships visits six clusters on one route, meaning that it does not have any remaining time for sailing additional routes.

Overall, the behavior of the models is as expected. In order to minimize the costs in a lower demand scenario, each ship should visit as many customers as possible without sailing back to the factory. If the demands on a route are so low that a ship is able to use the full time-period on the one route, this would most probably be included in the final solution. In general, the total demand is met in a shorter amount of time, which implies that the planning of the next period can begin earlier, and hence the optimization can be even more accurate.

8.7 Results

Based on the information discovered in the previous section, we will in the following two sections use this to investigate which system configuration that is best suited for modeling Marine Harvest's current system. When this is determined, we will use this final model to evaluate possible alterations to the current fleet used for performing the distribution.

8.7.1 Selection of Final Clustering Method and Characteristics

After exploring different demand scenarios for the different clustering methods and cluster dimensions, we are left with a lot of information. In general, the costs of distributing the feed decreases as the demand decreases. If the operational costs were a function of the quantity of cargo delivered as well, this would have been more clear. We are only looking at the transportation costs, and since we are visiting the same total number of farms in each scenario, the minimum total number of visits is always the same. Because of this, it is easier to directly compare the different approaches, as it shows us exactly which method that are able to create the cheapest routing of the same system with the same total demand. Based on this we will now select one of the methods for further evaluation of the current fleet configuration.

We are left with three distinctively different methods, where each one is using a characteristic of its own to reduce the problem size and complexity to about the same level, making them comparable. The contamination regions method stands out as it has a lot shorter calculation time than the two others, and hence may be too simplistic. This is amplified by the fact that Marine Harvest only have farms inside nine of the eleven regions, causing the clustering to be even more coarse. In some cases, the cluster demands become so large that a solution is not possible to produce. The method is also the most expensive in the medium demand case, and close to the most expensive in the low demand scenario. Overall, the approach does not seem suitable for our investigation, and further processing should be applied to the clustering technique before it is applicable to Marine Harvest's specific system. One way of doing this could be to divide each of the regions in two parts, and specify a compatibility between these pairs. This may be an important study if the contamination zones are introduced in the future.

The k-means method and the natural clustering method have a lot of similarities in all three demand scenarios, where the largest difference is in the low demand scenario. The similarity in the total cost is related to the small difference in cluster sizes. In the k-means method, compatibility constrains are used in order to reduce the number of possible routes instead of reducing the number of nodes directly. This creates less freedom in the solution space, and possible optimal solutions may be removed. This can especially be seen in the low demand scenario, where three charter ships must be used to serve all clusters, compared to the natural clustering method without compatibility constraints, where only two charter ships are used. Even though the compatibility constrains are set to reflect the distribution patterns used by Marine Harvest today, it is possible that solutions which are more optimal for their current fleet

exist. By restricting the sailing patterns for the ships, we remove the possibility to explore other distribution routes, which may be more effective and cost saving. It is thus preferred to reduce the number of clusters directly, and keep the flexibility in the sailing patterns to a maximum.

Throughout the determination of the cluster sizes, manipulation of the compatibility matrix, and demand scenarios, the natural clustering method with 23 clusters and full compatibility proved to be the best solution for optimizing the feed distribution in the given system. The method produces the cheapest distribution schedule in both the high and the low demand scenario, and was the only method with a lower cost in the low demand scenario than in the medium demand scenario. In the medium demand scenario, the cost is almost similar to the k-means method, but the k-means method has more clusters affecting the calculation time. The advantage of a method with no compatibility constraints is the high number of possible routes generated for all four ships within the desired time limit. Since there are no reduction in the solution space, all feasible routes for all ships are considered, not removing any optimal solutions. We therefore see the advantage of having no compatibility constraints as more valuable than increasing the number of clusters. The natural clustering method is also based on manually determined clusters. This gives the advantage of evaluating the cluster composition, for example clustering based on natural sailing routes, which is not considered in the mathematical k-means method. With a problem size of 23 clusters, the clusters are naturally located, and at the same time the model is able to generate solutions within a practical amount of time. By changing the configuration to 24 clusters, the increase in calculation time caused the solution to not be generated within the hour mark. If it is desired to reduce the calculation time, decreasing the cluster sizes, or applying compatibility constraints, can be considered. We will because of this continue with the natural clustering method with 23 clusters included for the final evaluation of the distribution process.

8.7.2 Evaluation of Fleet Configuration

When the best method of modeling the systems is determined, and the feed scenarios are stated, we can start to run the final experiments. Today, Marine Harvest use their fleet of four ships all through the year, but is this the best possible configuration of ships? In this section we want to investigate how different fleet configurations perform in the different scenarios, compared to the current fleet. This includes both periods with extremely high demand, to test if the internal fleet can handle the increased stress or if external supply must be more heavily used, and also during low season where the fleet utilization is at its lowest. Because of the large seasonal variation of feed consumption in Norway, a fleet that operates efficiently, and keeps the operational costs at a minimum in multiple demand scenarios, is desired. If the number of required vessels can be reduced, it is also possible that there are significant fixed costs savings associated.

In order to investigate whether the fleet Marine Harvest operates today is the most optimal, we will use the clustering method selected in the previous section, along with the three demand scenarios as used before. The current fleet configuration will be tested against different alterations

to evaluate the cost-savings and provide possible improvements for their distribution system. To create applicable alternative configurations, we first make some assumptions. Today, the fleet consists of two newer LNG-powered vessels which were built in 2014, and two smaller and older IFO-fueled vessels. Because the two newer vessels are only three years old, we will not consider fleet compositions where they are replaced. The two vessels are also the largest fish feed carriers in the market today, hence the inclusion of ships with larger capacity is assumed not possible. The total cargo capacity of the fleet is held close to the same amount since we are serving a fixed system. We finally also assume that there has been a development in IFO fuel-efficiency since the two oldest ships were built. This means that new ships operating with IFO as fuel option, are able to sail faster with the same consumption as the older IFO ships. We use the same planning horizon as before for the demand scenarios, one week for the high and medium demand scenarios, and four days for the low demand scenario. Four fleet alternatives, which we see as possible modifications, will be evaluated. These are alternatives that are in compliance with Marine Harvest’s strategy of substituting the oldest ships in near future. The alternative fleet compositions will differ in fleet size, total cargo capacity, fuel type, and operating speed.

The first fleet alternative we want to evaluate is a fleet where the two older ships have been replaced by one large LNG vessel, similar to their existing LNG vessels. This gives a homogeneous fleet of three ships as Table 12 shows.

Table 12: Fleet alternative 1.

Ship #	1	2	3
Capacity	3 000	3 000	3 000
Speed	13	13	13
Fuel	LNG	LNG	LNG

Next, we include Marine Harvest’s current strategy of replacing one of the older vessels with a similar, but newer and more efficient vessel. In our fleet, we have included a LNG vessel with 2 000 tonnes cargo capacity, and the same fuel specifications as the two larger LNG vessels. This is done based on predictions of stricter environmental regulations in the near future, and we hence see the inclusion of another LNG vessel as a good investment.

Table 13: Fleet alternative 2.

Ship #	1	2	3	4
Capacity	3 000	3 000	2 000	1 500
Speed	13	13	13	10
Fuel	LNG	LNG	LNG	IFO

We also want to look at the next possible step in the improvement process of the current fleet, where both of the older and smaller vessels are replaced with the same LNG vessel as in the previous fleet configuration. This means that the whole fleet is LNG powered as showed in Table 14, and even though this may not be very realistic in short term, we believe it is a case worth investigating for the future.

Table 14: Fleet alternative 3.

Ship #	1	2	3	4
Capacity	3 000	3 000	2 000	2 000
Speed	13	13	13	13
Fuel	LNG	LNG	LNG	LNG

The final configuration we want to investigate is the replacement of both of the smaller vessels, this time with three smaller and more updated IFO-powered vessels. As mentioned, we assume that these ships are state of the art, and hence can operate at a higher speed with the same specific fuel consumption as the two older ones. This configuration is mainly included because we believe it will add flexibility to the distribution process.

Table 15: Fleet alternative 4.

Ship #	1	2	3	4	5
Capacity	3 000	3 000	1 000	1 000	1 000
Speed	13	13	13	13	13
Fuel	LNG	LNG	IFO	IFO	IFO

After the four fleet configurations are established, we proceed by investigating the distribution performances in the different demand scenarios. The parameter we are most interested in is, as before, the total cost of the operation. We will run all consumption cases for each configuration, and compare them with the operational costs for the current fleet. All optimization processes are done with the selected natural model with the farms collected in 23 clusters. All fleets (1-4) have full compatibility to serve all of the 23 clusters generated by the natural clustering technique. A summary of the cost-differences compared to the original distribution schedule can be viewed in Table 16.

Table 16: Comparison of the cost between the current fleet and new fleet configurations.

Fleet Alternative	1	2	3	4
Difference cost mid	+ 881 \$	+ 454 \$	+ 2 372 \$	+ 2 532 \$
Difference cost high	- 155 \$	- 867 \$	+ 1 976 \$	+ 2 737 \$
Difference cost low	+ 4 118 \$	- 1 749 \$	- 401 \$	- 2 420 \$

In general, we can see that the difference in costs are both positive and negative in different scenarios for all alternative fleets, but none of the costs differs more than about 7% from the total cost of operating the current fleet configuration in any of the cases. It is hard to directly compare the different fleets against each other, but the factors that should be taken into account are total cost, flexibility, total capacity, and future applicability.

Fleet 1 has the same total cargo capacity of 9 000 tonnes as the current fleet, but one less ship. The reduced number of ships reduces the flexibility of the system. This is especially visible in the low demand scenario with a shorter planning period. The system has equally many farms to serve as in the longer planning period, but a shorter time to visit them all. Since the demand is

lowered, the number of possible routes for each ship increases, but with one less ship in the fleet, the total number of possible routes for the whole system is reduced. This gives fewer possible ways to construct the distribution schedule, thus reducing the flexibility of the system. With reduced flexibility, the distribution can be more costly in scenarios with low demand as there are less space for optimizing. The difference in operational costs in the low demand scenario for the two different fleet configurations is over 4 000 \$. The advantage of substituting the two smaller vessels with one larger, is the increased cargo capacity on each route. Instead of having routes visiting few clusters, and then returning to the factory more often, the larger ship can sail fewer routes and at the same time serve the same amount of clusters as the two smaller ships. Because the need of refilling cargo at the factory is lowered, the total sailing time without cargo can be reduced. This effect is visible in the high demand scenario, where this fleet configuration is cheaper to operate than the current fleet Marine Harvest are using.

Fleet 2 has 500 tonnes extra cargo capacity compared to the current fleet, and three ships running on LNG instead of two. The 500 tonnes of extra cargo capacity seems to be advantageous in both a low and high demand scenario compared to the original schedule. This is because it has the same flexibility as the current fleet with four vessels, and at the same time increased cargo capacity, making it able to serve more clusters on a single route. The reason why it has a higher cost in the medium demand scenario comes from the fact that a LNG powered ship has a higher operational cost. Even though the fuel price is lower than for the IFO-fuel, the specific fuel consumption is higher, giving a higher total operational cost. It is expected that the difference in fuel cost will be even greater in the future, which will give lower operational cost for LNG vessels compared to IFO-fueled vessels.

Fleet 3 consists of four vessels running on LNG, with a total cargo capacity of 10 000 tonnes. This configuration seems to be more expensive when the demand is medium or high. As mentioned above, the specific fuel consumption is higher for LNG-ships than for IFO-ships, which may cause the total cost of the fleet to increase. This seems to be affecting the total distribution cost in a larger scale than the increased cargo capacity. However, this may only be the case in a short term, as environmental restrictions are given an increased focus. Hence in a near future, it could be more profitable to only operating more environmental friendly LNG-vessels than only focusing on the current transportation costs. There is also a lower cost in the low demand scenario due to the increase in flexibility. Since the number of ships used in the system is the same, and the new ships have increased cargo capacity and operational speed, the number of possible ways to construct the routing schedule increases. Cheaper compositions of routes may then be created and hence reducing the total operational cost.

Fleet 4 operates with one more ship than the original fleet configuration, but has the same total cargo capacity. Instead of replacing the smaller ships with larger ships, we here reduce the cargo capacity of each ship, but instead increasing the flexibility by expanding the fleet size. This is especially profitable in the low demand scenario when it is desired to be able to visit many smaller customers in a shorter period of time. Equally negative is the effect in the high demand scenario, where the new ships are too small to effectively distribute the cargo, meaning that the

number of trips back to the factory is increased, and hence the total sailing time without cargo also increases. This causes the total costs to be higher than for the other configurations.

When comparing the possible new fleet configurations, it is clear that they all have advantages and disadvantages. There are multiple factors to take into consideration and analyze, but our main focus in this paper has been on cost-reduction. Based on the fleet configurations evaluated here, only one fleet configuration comes out as less costly in two of the scenarios than the current fleet composition. Even though it is more costly in the medium demand scenario, not by much, the overall cost-change in all scenarios combined is negative. By replacing one small, old, and slow ship with a newer, faster, and larger ship, significant cost savings can be achieved. Interestingly, this is also the same alteration as Marine Harvest are currently discussing. This shows that our tool may be applicable for actual evaluations, as their strategy matches our result.

We also believe that fleet 3 should be considered as an alteration of the fleet. At some point in the future, the other old vessel needs to be replaced with a new and more modern vessel. With the increasing focus on emission regulations, a fleet only consisting of LNG fueled vessels would be a good option for both the future, and also for certain scenarios today compared with the current fleet. In addition, the industry is growing, and Marine Harvest will probably also expand, meaning that they will need increased cargo distribution capacity in the future. On a short-term basis however, we consider fleet configuration 4 as a better solution. This is due to the high CAPEX related to investing in a new vessel, the current fuel prices, and current emission regulations, but fleet 3 should definitely be considered as an option on a long-term basis.

9 Discussion

Even though the structure of the fish feed distribution network appears simple, the size of the problem, combined with variations in feed consumption, and uncertainties in the delivery schedules, makes the system complex. In order to be able to model and optimize the problem, a lot of simplifications and assumptions are made throughout this paper, and including too many may cause the model to not represent the real world as accurate as desired. In addition, there are also direct sources of errors included in the calculations, affecting our results. However, in some cases these actions are necessary to be able to model the real world, and to get a problem structure that is solvable.

A major part of the inaccuracy in the modeling comes from simplifying the reality into simpler, solvable parameters and variables. One of these simplifications that is drastically different from the real world, is that the inclusion of several types of feed is neglected. In reality, multiple types are distributed to the different farms, and in some cases different types are demanded by the same customer. Because the inclusion of multiple products increases the problem size by adding several variables, one feed type is assumed. If multiple cargoes were to be distributed, both the model and the results would have been distinctively different. When a single product is transported, the full storage capacity of a ship can be used for the product, instead of multiple compartments. The quantities delivered would also have differed in the calculations, if multiple products were introduced, causing the distribution schedule to be changed. For a tool like ours to work, we see this specific simplification as absolutely necessary to be able to get results as quickly as desired.

We have also simplified some aspects of the reality when generating the possible routes. The biggest alteration to how the distribution is done in practice, is the inclusion of only partial split deliveries. Instead of full freedom to deliver any amount of cargo during a visit, we restrict our model to deliver exactly the demand of the customer, except when only having small quantities left on-board. This is mainly done to reduce to generation time, as this tend to grow exponentially with problem size. There exist models that are using split deliveries, but for a route generation algorithm, with a large number of customer nodes as in this system, the number of possible routes simply becomes too large to include full freedom. We could also have included time-windows for when the cargo is allowed to be delivered to reduce the solution space, but since most farms can receive feed at any time, we assumed this feature for all farms.

Another important aspect that is not in coherence with the reality, is the way the routes are starting and ending. We assume that all ships are located at the factory at the start of the planning period, which would never be the case in the real world. This problem is amplified if a vessel finishes its planned work in a shorter time than the length of the planning period, causing the ship to be stationary at the factory until the end of the period. In reality, Marine Harvest's factory only has the capacity to hold and load two ships at the same time. In order to simplify the model, we have assumed that the factory is able to load an infinite number of ships at the same time. This is mainly a problem in the beginning of the period, as all ships

are starting from the factory. If all ships would have been at the factory at the same time, as is possible in the model, the ships would have to wait in order to be loaded, which would reduce the distribution efficiency. In a future improvement of the model, a start location of each vessel should be included to deal with this problem. Since the tool is only meant for decision support, the utilization of the planning period could be assigned manually to remove the effects of stationary vessels. The same goes for the utilization of each ship's capacity. If there are available cargo capacity on any routes, additional feed can manually be assigned. We see this as a sufficient approach since the focus has been on optimal routing for the fleet.

Also, we have not considered the capacity of the production facilities in this paper. Since Marine Harvest is a fully vertically integrated company, inventory management planning is important as well. Some of the results discovered in this paper showed that sometimes the inclusion of external deliveries is advantageous. If we had considered production levels combined with the distribution problem, we could have experienced results where external orders would have been necessary in a greater extend in certain periods. This applies especially for the high demand scenarios discussed in this thesis, where it is possible that the current production facilities would not have the capacity to produce or store the large amount of feed needed. By evaluating the inventory management planning in our model, the structure would be very different and far more complex. It would have to include several additional factors, making the model much harder to solve. Because of a more complex structure than for the VRP, it would not have been possible to use the route generation heuristic used in this thesis. Another solution method would therefore have had to be used, which would require more calculation time. This would make it unusable for the operational stage, as we desire. Because of the motive of having a low calculation time, a VRP model formulation was necessary, even though a IRP formulation would represent the system in a better way.

When it comes to the cost calculation for each route, we have chosen to simplify it to only depend on distance, fuel prices and fuel consumption. Even though we are using real values for these parameters, they are actually variables that are constantly changing. The estimations are also not 100 % correct since the operational cost in reality depends on more factors, but the calculations are done in the same way for all methods, making them directly comparable for what we wanted to investigate. The same goes for calculating the cost of external deliveries. This is done solely depending on the number of customers to be served, and not the quantity of cargo to be delivered. In the real world, the external fodder price would have affected which farms that should be excluded from the internal distribution, and the amount of feed delivered externally should thus be a part of the optimization. As for the operational costs, the calculation is done in the same way for all methods, causing the optimization to only focus on the routing, which is what we wanted.

To be able to actually run a model, a planning horizon had to be chosen. This was determined based on a compromise between the real length of a planning period, and the desire to look at short enough periods to be able to treat variables as constants. With a planning horizon of one week, the assumption of constant feed consumption throughout the planning period can be

seen as a credible assumption. This is because in the real system, feed orders must be placed at least two weeks before the delivery date. The day-to-day variation in feed consumption is normally leveled out with the use of a safety stock at the farms. By including this, it can reduce the impact of unforeseen events which are affecting the distribution system. In addition, using a time-continuous model contributes to reduce the calculation time as demonstrated by Agra et al. (2013). This is why we have focused on a short calculation time, so that the model can be re-run once parameters are changing and then be able to re-optimize the current delivery plan.

As mentioned repetitively throughout the paper, reduction regarding the share size of the problem is also needed. By clustering farms together, the number of nodes is drastically reduced, which lower the calculation time to an acceptable value. To be able to evaluate the use of clusters in general, we should have compared the results with actual routing schedules used today. Sadly, this information is confidential and we hence do not have the necessary data material to discuss the exactness of our approach. However, we have tried to cluster the farms in the most realistic way possible. The location of all farms is publicly available, and in most cases they are placed out in a way that makes it reasonable to assume that if one is visited, the inclusion of the nearby farms on the same route is logical. The main problem with the use of clusters is that the routing inside each cluster is not optimized, but rather estimated in an inaccurate way. The larger the clusters are created, the worse the model represents the real world system. Larger clusters also amplify the problem of distance estimations between clusters. When the clusters include a larger number of farms, the center is representing the actual farm locations less accurate since the farms are more spread out. It is clear to us that when using clusters as problem reduction technique, it all comes down to the compromise between problem size and desired calculation time when modelling the system. In our case, the limit was found to be 23 customers.

In addition, all distances in the model are estimations based on commercial programs, and the method of calculating the distance matrix includes several sources of error. Since the farms are located among several islands and in narrow fjords, the distance calculations are difficult and often incorrect. One source of error is the routing performed by Searoutes.com (2017). As Figure 25 shows, some routes may not be feasible as it crosses land, and the calculated distance is thus not correct, most often too short. This can also cause the route generator to choose a route that does not exist in the real world, which gives an optimal routing schedule that are based on shorter distances than reality. In worst case, the time usage of the infeasible routing exceeds the available time, causing the whole routing solution to be a non-existing solution.

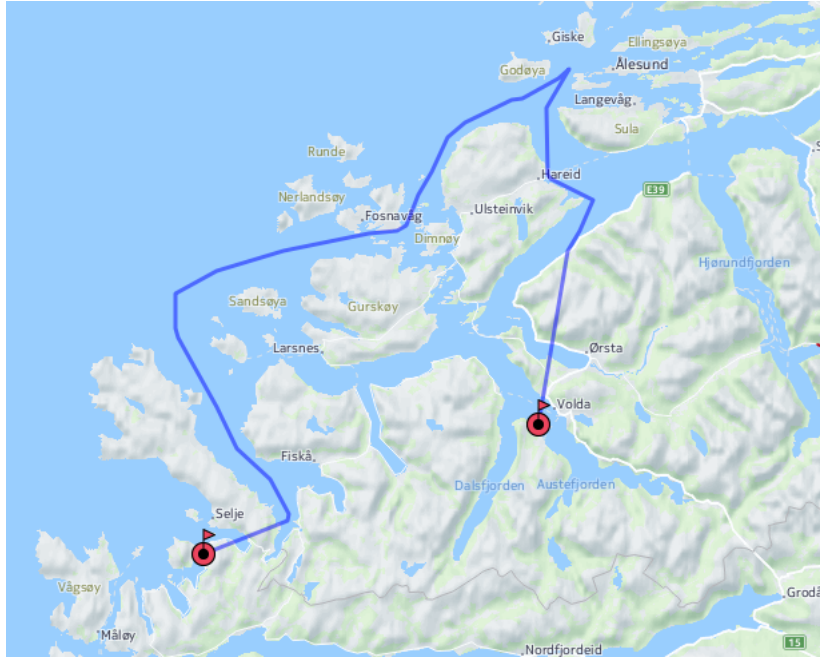


Figure 25: Example of error with routes generated by Searoutes.com (2017).

Even though we are aware of this problem, we do not see it as a critical simplification to include. The estimations are in most cases good approximations, and if the tool were to be implemented, the planning manager would have access to exact values for all distances in the system. This could then easily be implemented for the desired system configuration, and a fully feasible input file could be created.

Also, when we are calculating the internal sailing time for visiting a cluster, we are forced to simplify the calculations. In reality, the time spent in each cluster is based on the sailing distance along the distribution route within the cluster, various speed of the vessel, and time spent unloading feed at each location. The process of accurately calculate the time spent in a cluster is both difficult and time-consuming, and many simplifications are thus made. Preferably, individual routing optimization should be performed for each cluster to determine the optimal routing pattern within the cluster. Because this is affected by the demand of each farm, the routing may not be the same in each scenario. This internal pattern must then had to be determined whenever new information is provided, which will lead to a drastic increase in the total calculation time. Since it is desired to keep the calculation time at a minimum, we have assumed a sailing distance based on the total distance from each farm in the cluster, to the cluster center. This does not by any means minimize the sailing distance, and our time estimate is hence on the conservative side. This means that the contribution to the total cost of each route is also an over-estimation. An example of this effect can be seen by comparing the clustering method based on the contamination zones and natural clustering method, with 11 and 23 clusters. Even though the total number of farms and total demand is equal, there is a major difference in total costs, as the total internal distance for each cluster is different for the clustering techniques and cluster sizes.

The contribution to the internal service time from the offloading process in each cluster is on the other hand an under-estimation. The time should be dependent of the amount of feed actually delivered to the cluster, but the route generation method needs the internal sailing time before generating all possible routes. Because of this, the contribution from unloading time is not a function of delivered feed, but rather a function of the amount of cargo demanded. Since the ordered amount of feed is always the same or less, as the amount delivered, the unloading time calculated is an under-estimation of the reality. The pre-defined unloading rate will also in the real world not be constant as we treat it, but rather heavily dependent on weather conditions and other uncontrollable factors, causing the process to be even more under-estimated. Regardless of this, the implication-effect of these simplifications is small because of the short time spent unloading compared to the total time spent on the rest of the route.

Overall, there has been lot of focus on keeping the computation time low during the process of making this paper. It is possible that other system configurations and dimensions should have been considered if the calculations were run on a more powerful computer. This could have lowered the generation time compared to the experiments we have been running, and hence could have increased the size of the solvable solution space within practical time limits. If the improvements were significant, other time periods could have been considered and compared for finding the best way of modelling the system. This could also have been used for the low demand scenario where we have shortened the planning horizon, and we could possibly have found a solution where the need of external supply are not be needed. However, the reduction in planning period for two of the methods in the low scenario provides valuable information to how the system reacts to a shortened period. Also, the solution time for finding optimal routing schedules could decrease, meaning that we possibly could have optimized larger systems. This is especially applicable for higher planning levels, as these decisions normally are determined based on longer, and more copious calculation times than the operational planning level we are considering.

Today, some of Marine Harvests fish farms are supplied externally, but a part of their strategy is to be fully independent of external suppliers. From the results for the different scenarios this strategy may not be economically beneficial, but these extreme scenarios will seldom occur, and during this extraordinary period the inclusion of charter ships is therefore considered smart. This especially applies to the farms located furthest north and south. These particular farms are so far from the factory that it is seldom wise to include them in the internal schedule. Even though our calculations, for both operational costs and external costs, are done in a simplistic way and may not represent the actual costs in a sufficient way, deliveries to these farms are such a large cost contributor that special planning should be considered. Farms as remote as these farms should regardless of this be given extra attention since they affect the optimization process in a large extent. By comparing the inclusion in the internal schedules with fixed contracted deliveries, more certainty can be achieved to what should be done. Another possibility is mentioned in Section 3.5, where we discuss the inclusion of multiple depots along the coast for temporary storage closer to farm locations. This would be especially applicable to the remote farms furthest north and south, where a single optimized vessel could serve the

small number of farms from the depot alone. This would increase the possibility to optimize the system, as the inclusion of remote farms uses a lot of the available time, and hence removing them would increase the freedom of the system. This can easily be included in the current model by treating the depot as a customer node, and include the deliveries to this location to the routing schedule and reduce the time usage at the node.

When evaluating possible alterations to Marine Harvest's current distribution fleet, the same fuel prices are assumed in all fleet configuration. The optimization model only considers the best possible fleet configuration with constant fuel costs and consumption. Since the cost-function of using LNG fuel are higher than for IFO, IFO-driven vessels are preferred in the model. Even though IFO vessels are preferred in the model, it may not be the best solution for the future. It is an international effort to reduce emissions worldwide, and because LNG is a more environmental friendly alternative, a fleet consisting of LNG driven vessels may be a better solution in the near future. These are considerations which should be taken into account, in addition to the cost-reduction when evaluating different fleet alternatives.

Also when comparing the fleet configurations in the different scenarios, the different lengths of the planning periods causes the results to not be directly comparable. Some compositions may be able to meet the demand in the short time, while some are dependent of charter ships, and hence the total cost of the distribution is higher. However, there may exist solutions that would have a lower total cost if the planning horizon were extended, meaning that some of the compositions that are cheaper when we compare them, actually are more expensive in a longer planning period and hence not necessarily the best.

10 Conclusion

To help Marine Harvest remain a competitive and efficient business in an industry in rapid growth, an automatic and practical approach for minimizing operational costs for their fish feed distribution is proposed. Today, no practical tools exist for helping with this planning. In addition to minimize the operational costs, the focus has been on creating an optimization model reflecting the real world within a short calculation time to be able to replace long and unpractical planning methods applied in the business today. Because of the large problem size, many simplifications are applied to the model created, and the model should therefore only be used as an operational planning tool. It should ideally be rerun once new information is available, and could thereby support weekly or even daily decisions on delivery planning.

When we have modeled the reality, there has been a constant focus on the trade-off between the solution quality and calculation time. We first focused on the calculated time, and when this was at the desired level, we then evaluated the exactness of the solutions. When using the determined system dimensions, the model was able to recreate the reality in the way that the cost naturally decreased with lowered demand, even though the operational cost is not a direct function of the demand quantity. Even though values for variables such as demand, sailing-distances, and production capacity are estimated, we were still able to recreate the main characteristics of the distribution process as all farms were visited once or twice during a one week period. Ship and time-utilization were in general high, and the slack can easily be improved manually. With exact values for the estimated parameters, the model will be even closer to the reality.

The model is also able to generate results in the desired short amount of time, meaning that the desire of reducing the time used for planning today is met. Since the planners are using several hours each day to manually assign ships to routes, the total time of creating the input file and for the model to create a solution should be significantly lower. With a calculation time of less than 10 minutes, there are a lot of slack left for creating the system input-file for different configurations, and still be more efficient than the current planning process.

The computational study showed us that there were significant differences between the three clustering methods tested. The operational costs calculated with the k-means approach and the natural clustering approach first seemed to similar, but for more extreme demand cases the full freedom of the unconstrained natural method proved to produce more optimal routing schedules. Since the k-means approach with compatibility constraints represents how the distribution is performed today, the associated cost-savings with the full freedom modeling show signs of possible improvements of the routing done today. The difference in costs were not substantial in a medium or a low demand scenario, but over 5 000 \$, or 12.5 %, less costly in a high demand scenario performed with the same fleet. Because the distribution system evaluated in this thesis is affected by large seasonal variations, full freedom for the ships should be introduced in order to reduce the operational costs.

When finally using the determined model for evaluating the current fleet, we discovered that by replacing one of the older and smaller vessels with a new and larger, costs can be reduced in both high and low demand scenarios, which is also in compliance with Marine Harvest's strategy. The increased capacity with a new and larger vessel will also be beneficial for the company if they are to expand at the same rate as the rapidly growing aquaculture industry in Norway.

11 Further Work

Tools for optimizing the distribution of fish feed along a complex coastline does not exist at the moment, and the work presented in this paper are, because of this, heavily dependent of simplifications of the reality. In future work, a study of the inclusion of more realistic versions of the simplified aspects should be done. By investigating the individual effects of adding new constraints with regard to calculation time, a compromise reflecting the reality more accurately than the presented model can be found. This especially applies to the principles of multiple cargo types, split deliveries, and production limitations. For the tool to work in the current market, the option of time-windows should also be included in a future model because some of the farms are still not able to receive cargo at all times.

Further alterations should also be investigated to prevent the ships of returning to the factory at the end of the planning period. This reduces the optimality of the routing schedule, as ships may be stationary at the factory until the end of the planning period. An improvement of the model should thus include the possibility for a ship to start and end their route at any given node. This means that a new planning period can start whenever all vessels have completed their current route. The challenge of including this aspect is both concerned the programming and the increase in the solution space, leading to a higher calculation time.

Disregarding the possible improvements of the model formulation, the current model could also be used for better optimizing the system. Instead of only looking at the routing between clusters, multiple models could be created for each of the clusters. By doing this, an optimal schedule could be produced for when a vessel visited a specific cluster to determine the sailing pattern inside the cluster. If the model were first used on a more powerful computer in the tactical planning level to determine the best number of clusters to include, this would have been even more applicable. A more accurate estimation of the time used on each route could then be obtained.

An extensive study should also be performed regarding other ways to model the system configuration. We have determined that in most cases some of the farms can be considered as outliers, and hence should be served externally. If the optimization process were run without these farms included in the system, another routing schedule could have been generated. Another possibility that our method is able to investigate, is the inclusion of depots instead of the outlier farms. By replacing the location of the farms with the location of a possible depot, the delivery to the cluster could be including in the routing plan instead. By then creating a separate model for the distribution from these depots to the farms it is representing in the other model, the two different results combined could more accurately represent the real cost of serving all farms internally.

In the evaluation of different clustering methods, the regional method based on the contamination zones proved to be too coarse when including it unmodified. Since there are an increased focus on contamination of lice-diseases, and the possibility of a system configuration like this to be introduced is rather high, a future study how to implement this clustering method in the tool should be conducted. This could be accomplished by splitting each of the zones into smaller clusters, and introduce compatibility between the customers inside the zone, but not for connecting these farms with farms in other zones. This approach could also benefit of the use of multiple depots for temporary storage of feed, where each region-center serves as a depot for the distribution process inside the zone.

When it comes to the evaluation of Marine Harvest's current fleet size, a more extensive evaluation can be done. We have in this thesis only considered fleet configurations for their current distribution system, but the model could also be used as an evaluation-tool for a predicted expansion. Since the industry is growing, it is natural to believe that Marine Harvest also will increase the production volume, and hence the need for a larger fleet of vessels for distributing feed is probable. Also, the industry is expanding its area of operation to include offshore farming. This could imply that larger ships than the ones currently being used also should be considered in a future study.

References

- Ådlandsvik, B. (2015). *Forslag til produksjonsområder i norsk lakse- og ørretoppdrett*. Havforskningsinstituttet.
- Agra, A., Christiansen, M., & Delgado, A. (2013). Mixed integer formulations for a short sea fuel oil distribution problem. *Transportation Science*, Vol. 47 (1), 108–124.
- Archetti, C., Bertazzi, L., Laporte, G., & Speranza, M. G. (2007). A branch-and-cut algorithm for a vendor- managed inventory-routing problem. *Transportation Science*, Vol. 41 (3), 382–391.
- Austreng, E., Storebakken, T., & Asgard, T. (1987). Growth rate estimates for cultured atlantic salmon and rainbow trout. *Elsevier Science Publishers B.V.*
- Brandao, J. & Mercer, A. (1997). A tabu search algorithm for the multi-trip vehicle routing and scheduling problem. *European Journal of Operational Research*, Vol.100, 180–191.
- Brandao, J. & Mercer, A. (1998). The multi-trip vehicle routing problem. *The Journal of the Operational Research Society*, Vol.49, 799–805.
- Brønmo, G. & Løkketangen, A. (2007). *An adaptive constructive solution generator for ship scheduling dss*. Presented at the NIK 2007 conference.
- Christiansen, M. & Fagerholt, K. (2009). *Maritime Inventory Routing Problems*. In Floudas, C. and Pardalos, P., editors, *Encyclopedia of optimization*, pages 1947–1955. Springer Berlin, 2 edition.
- Christiansen, M., Fagerholt, K., Nygreen, B., & Ronen, D. (2013). Ship routing and scheduling in the new millennium. *European Journal of Operational Research*, Vol. 228, 467–483.
- Clarke, G. & Wright, J. (1964). Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, Vol.4, 568–581.
- Craig, S. & Helfrich, L. A. (2009). *Understanding Fish Nutrition, Feeds, and Feeding*. Retrieved October 8, 2016, from <https://pubs.ext.vt.edu/420/420-256/420-256.html>
- Dantzig, G. B. & Ramser, J. H. (1959). The truck dispatching problem. *Management Science*, Vol. 6 (1), 80–91.
- Fagerholt, K. (1999). Optimal fleet design in a ship routing problem. *International Transactions in Operational Research*, Vol. 6, 453–464.

- Fagerholt, K., Heimdal, S., & Loktu, A. (2000). Shortest path in the presence of obstacles: an application to ocean shipping, 683–688.
- FICO. (2013). *Xpress-mosel reference manual, release 3.4*. FICO.
- Fiskeridirektoratet. (2017). *Akvakultur*. Retrieved April 27, 2017, from <http://www.fiskeridir.no/Akvakultur>
- Gillett, B. & Miller, L. (1974). A heuristic algorithm for the vehicle-dispatch problem. *Operations Research, Vol. 22*, 340–349.
- Glover, F. (1989). Tabu search - part i. *ORSA Journal on Computing, Vol.1*, 190–206.
- Gomory, R. (1963). An algorithm for integer solutions to linear programs. In R. Graves & P. Wolfe (Eds.), *Recent advances in mathematical programming* (pp. 269–302). McGraw-Hill.
- Grønhaug, R. & Christiansen, M. (2009). Supply chain optimization for the liquefied natural gas business. In J. Nunen, M. Speranza, & L. Bertazzi (Eds.), *Innovations in distribution logistics, lecture notes in economics and mathematical systems* (Vol. 619, pp. 195–218). Springer.
- Grötschel, M. & Holland, O. (1991). Solution of large-scale travelling salesman problems. *Mathematical Programming, Vol. 51 (2)*, 141–202.
- Guéret, C., Prins, C., & Sevaux, M. (2000). Applications of optimization with xpress-mp. *Editions Eyrolles, Paris, France*.
- Jonermark, C. (2016). Information provided through email. Claes.Jonermark@marineharvest.com.
- Lang, A. & Doig, A. (1960). An automatic method of solving discrete programming problems. *Econometrika, Vol. 28 (1)*, 497–520.
- Laporte, G. (2007). What you should know about the vehicle routing problem. *Naval Research Logistics, Vol. 54*, 811–819.
- MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth symposium on math, statistics, and probability* (pp. 281–297). Berkley, CA: University of California Press.
- Marine Harvest. (n.d.). *Welcome to Valsneset - Marine Harvest Fish Feed*. Handed out material during visit at factory in Bjugn 27.10.2016.
- Marine Harvest. (2013). *Salmon Farming Industry Handbook 2013*. Retrieved October 17, 2016, from <http://hugin.info/209/R/1698446/559980.pdf>

- Marine Harvest. (2016). *Salmon Farming Industry Handbook 2016*. Retrieved October 17, 2016, from <http://www.marineharvest.no/globalassets/investors/handbook/2016-salmon-industry-handbook-final.pdf>
- Martí, R., Resende, M. G., & Ribeiro, C. C. (2013). Multi-start methods for combinatorial optimization. *European Journal of Operational Research*, Vol. 226 (1), 1–8.
- Mitchell, J. E. (1999). Branch-and-cut algorithms for combinatorial optimization problems. *Mathematical Sciences*.
- Moutaoukil, A., Neubert, G., & Derrouiche, R. (2014). A comparison of homogeneous and heterogeneous vehicle fleet size in green vehicle routing problem. *IFIP International Conference on Advances in Production Management Systems (APMS)*, 450–457.
- Padberg, M. & Rinaldi, G. (1991). A branch-and-cut algorithm for the resolution of large-scale symmetric travelling salesman problems. *SIAM Review*, Vol. 33 (1), 60–100.
- Romeroa, G., Durán, G., Marencoe, J., & Weintraub, A. (2013). An approach for efficient ship routing. *International Transactions in Operational Research*, Vol. 20, 767–794.
- Ronen, D. (1983). Cargo ships routing and scheduling: survey of models and problems. *European Journal of Operational Research*, Vol. 12 (2), 119–126.
- Rothlauf, F. (2011). *Design of modern heuristics: principles and application*. Natural Computing Series. Springer.
- Schmidt, G. & Wilhelm, W. (1999). Strategic, tactical and operational decisions in multi-national logistics networks. *International Journal of Production Research*.
- Searoutes.com. (2017). *Searoutes*. Retrieved April 27, 2017, from <http://www.searoutes.com/>
- Singapore Bunker Prices. (2017). *SHIPS AND BUNKER*. Retrieved April 23, 2017, from <https://shipandbunker.com/prices/apac/sea/sg-sin-singapore>
- Skretting. (2017). *How much feed is required to grow a farmed fish*. Retrieved April 17, 2017, from <http://www.skretting.com/en-AU/faqs/how-much-feed-is-required-to-grow-a-farmed-fish>
- Sörensen, K. & Glover, F. (2013). Encyclopedia of operations research and management science. In S. Gass & M. Fu (Eds.), (3rd ed., Chap. Metaheuristics, pp. 960–970). New York: Springer.
- SSB. (2016). *Aquaculture, 2015, final figures*. Retrieved September 12, 2016, from <https://www.ssb.no/en/jord-skog-jakt-og-fiskeri/statistikker/fiskeoppdrett/aar-forelopige>

- Tan, P.-N., Steinbach, M., & Kumar, V. (2016). Introduction to data mining. (Chap. 8. Cluster Analysis: Basic Concepts and Algorithms, pp. 487–568). Pearson Addison Wesley.
- Tang, J., Pan, Z., Fung, R. Y., & Lau, H. (2009). Vehicle routing problem with fuzzy time windows. *Fuzzy Sets and Systems, Vol. 160*, 683–695.
- The MathWorks. (2004). *Mapping toolbox for use with matlab user's guide version 2*. The MathWorks.
- Tolstoi, A. (1930). Metody nakhozhdeniya naimen'shego summovogo kilometrazha pri planirovanii perevozok v prostranstve [russian; methods of finding the minimal total kilometrage in cargo-transportation planning in space]. *Planirovanie Perevozok, Sbornik pervui [Russian; Transportation Planning, Volume I]*, 23–55.
- UNCTAD. (2015). *Review of maritime transport*. United Nation Conference on Trade and Development.
- Vaira, G. & Kurasova, O. (2014). Genetic algorithm for vrp with constraints based on feasible insertion. *INFORMATICA, Vol.25*, 155–184.
- Wang, X. (2014). *Operational transportation planning of modern freight forwarding companies*. Springer Gabler 2014.
- Whelan, C., Harrell, G., & Wang, J. (2015). Understanding the k-medians problem. (pp. 219–222). Athens: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). Retrieved from <https://search.proquest.com/docview/1705667495?accountid=12870>

A Explanation of Attached Files

Below is an explanation of the different folders and files attached to the delivery of this thesis.

1. Input files folder

- (a) Input file for test system 1.
- (b) Input file for test system 2.
- (c) Folder with input files for all systems discussed in Section 8.4, Determination of Cluster Sizes.
- (d) Folder with input files for all systems discussed in Section 8.5, Manipulation of Compatibility Matrix.
- (e) Folder with input files for all systems discussed in Section 8.6, Consumption Scenarios.
- (f) Folder with input files for all systems discussed in Section 8.7.2, Evaluation of Fleet Size.

2. Route generation folder

- (a) Generated routes for test system 1.
- (b) Generated routes for test system 2.
- (c) Folder with generated routes for all systems discussed in Section 8.4, Determination of Cluster Sizes.
- (d) Folder with generated routes for all systems discussed in Section 8.5, Manipulation of Compatibility Matrix.
- (e) Folder with generated routes for all systems discussed in Section 8.6, Consumption Scenarios.
- (f) Folder with generated routes for all systems discussed in Section 8.7.2, Evaluation of Fleet Size.

3. MATLAB files folder

- (a) Script used for calculating parameters and generating clustering using the k-means method.
- (b) Script used for calculating parameters and generating clustering using the natural selection method.
- (c) Script used for calculating sailing distances between cluster centers.

4. Route generator model

Java file delivered by Inge Norstad, SINTEF Ocean.

5. 5.Optimization model

Xpress optimization model written in Mosel language.

B List of Marine Harvest's Farm Locations

nr	navn	fylke	kommune	lat	lon	Feed region	Cluster	Demand
0	FORFABRIKK	Sør-Trøndelag	Bjugn	63,8171	9,63429			
1	HJELNES	Troms	KVÆNANGEN	70,0638	21,9035	N		
2	YTRE HAMNEBU	Troms	KVÆNANGEN	70,0492	21,8117	N		
3	SVARTBERGET	Troms	KVÆNANGEN	70,0186	22,0183	N		
4	KVITEBERG	Troms	KVÆNANGEN	69,9539	21,8804	N		
5	HJELLBERGET	Troms	KVÆNANGEN	69,9116	21,5972	N		
6	NØKLAN	Troms	KVÆNANGEN	69,9077	21,8331	N		
7	FJELLBUKT	Troms	KVÆNANGEN	69,9046	21,8768	N		
8	RAKKENES	Troms	KVÆNANGEN	69,8851	21,73	N		
9	KARVIKA	Troms	KVÆNANGEN	69,852	21,829	N		
10	JEKTVIKA I	Nordland	SØRFOLD	67,5313	15,4436	N		
11	MATVIKA	Nordland	SØRFOLD	67,528	15,3597	N		
12	KALVIK	Nordland	SØRFOLD	67,5132	15,5704	N		
13	NEDRE KVARV	Nordland	SØRFOLD	67,4585	15,5116	N		
14	KINES	Nordland	SØRFOLD	67,4488	15,488	N		
15	MULNESODDEN	Nordland	GILDESKÅL	67,1454	14,2812	N		
16	LANGSKJÆRAM	Nordland	GILDESKÅL	67,1184	13,8984	N		
17	FLEINA	Nordland	GILDESKÅL	67,0951	13,8752	N		
18	NORDFUGLØY	Nordland	GILDESKÅL	67,0839	13,8569	N		
19	DIGERMULEN	Nordland	RØDØY	66,6765	13,3169	N		
20	KVALVIKA	Nordland	RØDØY	66,6691	13,3916	N		
21	LIAN	Nordland	RØDØY	66,5579	13,5262	N		
22	STORSTROMPA	Nordland	RØDØY	66,5483	13,4	N		
23	LINESVIKA	Nordland	RØDØY	66,54	13,3	N		
24	RAMNØYA S	Nordland	DØNNA	66,2321	12,2803	N		
25	LILLE ÅSVÆR	Nordland	DØNNA	66,2299	12,3175	N		
26	BREIVIKA	Nordland	DØNNA	66,1609	12,6815	N		
27	BREIVIKA S	Nordland	DØNNA	66,1581	12,682	N		
28	LANGØYA NNV	Nordland	HERØY i Nordland	66,0608	12,2423	N		
29	BUKKHOLMEN	Nordland	ALSTAHAUG	65,9184	12,2447	N		
30	BRATTHOLMEN	Nordland	HERØY i Nordland	65,9074	12,221	N		
31	SKARVØY NV	Nordland	ALSTAHAUG	65,893	12,33	N		
32	BLOMSØRÅSA	Nordland	ALSTAHAUG	65,8564	12,3118	N		
33	MEFALDSKJÆR	Nordland	ALSTAHAUG	65,8515	12,2872	N		
34	SLÅTTVIKA	Nordland	BRØNNØY	65,5243	12,5466	N		
35	GAVELN	Nordland	BRØNNØY	65,511	12,524	N		
36	MORTENØY	Nordland	BRØNNØY	65,4572	12,5497	N		
37	ISTERVIKA	Nordland	BRØNNØY	65,4469	12,5618	N		
38	RÅKNESET	Nordland	BRØNNØY	65,4409	12,6269	N		
39	KLUBBEN	Nordland	SØMNA	65,3523	12,438	N		
40	JEKTVIKA II	Nordland	SØMNA	65,3127	12,4022	N		
41	OLDERBAKKEN	Nordland	SØMNA	65,3112	12,3686	N		
42	STEINFLESA	Nord-Trøndelag	LEKA	65,1345	11,5157	M		
43	JAKOBSTEIN SV	Nord-Trøndelag	LEKA	65,1127	11,7929	M		
44	SLOKKHOLMEN	Nord-Trøndelag	LEKA	65,0409	11,4848	M		
45	KALVHAGAN II	Nord-Trøndelag	NÆRØY	64,9508	12,0567	M		
46	STORVIKBUKTA	Nord-Trøndelag	NÆRØY	64,9289	11,9129	M		
47	TORGERHAUGE	Nord-Trøndelag	NÆRØY	64,893	11,763	M		
48	KIPHOLMEN	Nord-Trøndelag	NÆRØY	64,7939	11,6895	M		
49	DIGERNESET	Nord-Trøndelag	NÆRØY	64,7866	11,7279	M		
50	RAMSTADHOLM	Nord-Trøndelag	NÆRØY	64,7548	11,246	M		

51	GEITRYGGEN	Nord-Trøndelag	NÆRØY	64,7509	11,189	M		
52	KJELNESET	Nord-Trøndelag	FOSNES	64,6439	11,3885	M		
53	VEDØYA	Nord-Trøndelag	FOSNES	64,6074	11,3548	M		
54	BJØRGAN	Nord-Trøndelag	FLATANGER	64,5927	10,9083	M		
55	FEØYVIKA	Nord-Trøndelag	FLATANGER	64,5607	10,919	M		
56	ALMURDEN	Nord-Trøndelag	FLATANGER	64,4332	10,681	M		
57	AUSTVIKA	Nord-Trøndelag	FLATANGER	64,4185	10,696	M		
58	BREIDVIKA	Sør-Trøndelag	OSEN	64,2954	10,3975	M		
59	INDRE SKJERVØ	Sør-Trøndelag	OSEN	64,2858	10,3091	M		
60	FLATØYFJORDEN	Sør-Trøndelag	ROAN	64,273	10,2462	M		
61	PERSFLUA	Sør-Trøndelag	BJUGN	63,9135	9,82643	M		
62	SLETTHOLMENE	Sør-Trøndelag	FRØYA	63,88	8,55918	M		
63	GRØTTINGSØY	Sør-Trøndelag	FRØYA	63,8797	8,54377	M		
64	MANNBRUHOLM	Sør-Trøndelag	FRØYA	63,8	8,53	M		
65	KVEITSKJERET	Sør-Trøndelag	FRØYA	63,8	8,4	M		
66	MUNKSKJÆRA	Sør-Trøndelag	FRØYA	63,8197	8,40882	M		
67	VALØYAN	Sør-Trøndelag	FRØYA	63,8174	8,46248	M		
68	TENNØYA	Sør-Trøndelag	FRØYA	63,805	8,4444	M		
69	GÅSHOLMEN	Sør-Trøndelag	FRØYA	63,666	8,4414	M		
70	ILSØYA	Sør-Trøndelag	FRØYA	63,6642	8,43168	M		
71	OSHOLMEN	Sør-Trøndelag	HITRA	63,662	9,13558	M		
72	SVELLUNGEN	Sør-Trøndelag	HITRA	63,6553	9,1568	M		
73	VARDEN	Sør-Trøndelag	HITRA	63,6357	9,18647	M		
74	SENGSHOLMEN	Sør-Trøndelag	HITRA	63,6184	8,5277	M		
75	KÅHOLMEN	Sør-Trøndelag	HITRA	63,6172	8,99413	M		
76	HEGGVIKA	Sør-Trøndelag	HITRA	63,6106	9,05073	M		
77	LILLE TORSØY II	Sør-Trøndelag	HITRA	63,608	8,49463	M		
78	LILLE TORSØY	Sør-Trøndelag	HITRA	63,6056	8,48855	M		
79	KORSHOLMAN	Sør-Trøndelag	HITRA	63,5972	8,51575	M		
80	HELSØYA	Sør-Trøndelag	HITRA	63,5453	8,38309	M		
81	HÅØYA	Sør-Trøndelag	HITRA	63,4656	8,72723	M		
82	BREMNESSVAE	Møre og Romsda	SMØLA	63,398	8,21432	M		
83	BRETTINGEN	Møre og Romsda	SMØLA	63,3582	8,23727	M		
84	NØRHOLMEN	Møre og Romsda	SMØLA	63,3389	8,2718	M		
85	LEITE	Møre og Romsda	AVERØY	63,036	7,67812	W		
86	ROKSET	Møre og Romsda	AVERØY	62,9738	7,64618	W		
87	KORNSTAD	Møre og Romsda	AVERØY	62,9604	7,45015	W		
88	KRÅKNES	Møre og Romsda	AUKRA	62,8315	6,96273	W		
89	STORVIKA	Møre og Romsda	FRÆNA	62,8022	6,98127	W		
90	AUKRASANDEN	Møre og Romsda	AUKRA	62,7834	6,92605	W		
91	ROGNE	Møre og Romsda	HARAM	62,6695	6,35683	W		
92	RØYSA	Møre og Romsda	GISKE	62,5487	6,1417	W		
93	OKSEBÅSEN V	Møre og Romsda	GISKE	62,5174	6,15805	W		
94	SAGELVA	Møre og Romsda	ØRSTA	62,338	6,38165	W		
95	VOLDNES	Møre og Romsda	HERØY i Møre og Rom	62,2097	5,7849	W		
96	SANDVIKA	Møre og Romsda	VOLDA	62,1957	5,90442	W		
97	BJØRLYKKESTRA	Møre og Romsda	VANYLVEN	62,1666	5,5954	W		
98	BEITVEIT	Sogn og Fjordan	SELJE	62,1412	5,32675	W		
99	RØNSTAD	Møre og Romsda	VOLDA	62,1351	6,02143	W		
100	BRUDEVIKA	Møre og Romsda	VANYLVEN	62,1166	5,42072	W		

101	RUNDREIMSTRA	Sogn og Fjordan	SELJE	62,01	5,33192	W		
102	BRUNSVIK	Sogn og Fjordan	VÅGSØY	61,9102	5,29085	W		
103	VERPEIDE	Sogn og Fjordan	VÅGSØY	61,9022	5,20923	W		
104	JUVIKA B	Sogn og Fjordan	BREMANGER	61,8957	5,6331	W		
105	KRABBESTIG	Sogn og Fjordan	VÅGSØY	61,8946	5,0616	W		
106	GULESTØ	Sogn og Fjordan	BREMANGER	61,758	5,06988	W		
107	FLÅTEGRUNNEN	Sogn og Fjordan	FLORA	61,5762	4,80733	W		
108	HEGNES	Sogn og Fjordan	FJALER	61,3563	5,28467	W		
109	KYRAVIKA	Sogn og Fjordan	FJALER	61,3164	5,1563	W		
110	SKREDSTIVIK	Sogn og Fjordan	HYLLESTAD	61,1567	5,13398	W		
111	TVEIT	Sogn og Fjordan	HYLLESTAD	61,1101	5,34048	W		
112	OLDERØY	Sogn og Fjordan	SOLUND	61,1033	4,71223	W		
113	BEKKSNESET	Sogn og Fjordan	HYLLESTAD	61,0766	5,38595	W		
114	BRATTHOLMEN	Sogn og Fjordan	SOLUND	61,0684	4,71382	W		
115	INDRE OPPEDAL	Sogn og Fjordan	GULEN	61,0679	5,56383	W		
116	HJARTHOLM	Sogn og Fjordan	GULEN	61,0351	5,33105	W		
117	OKSEN	Hordaland	FJELL	60,4336	4,94062	S		
118	HAVERØY	Hordaland	FJELL	60,4048	4,93995	S		
119	SKORPEOSEN	Hordaland	FJELL	60,3325	4,94278	S		
120	SVÅSANDNESET	Hordaland	JONDAL	60,3242	6,31898	S		
121	SYLTØY	Hordaland	SUND	60,3046	4,96645	S		
122	HØYSTEINEN	Hordaland	KVINNHERRAD	60,1574	6,05012	S		
123	ÅKRE	Hordaland	KVINNHERRAD	60,1124	6,04248	S		
124	TROMMO	Hordaland	KVINNHERRAD	60,0824	6,08007	S		
125	ULVENESET	Hordaland	KVINNHERRAD	59,9928	5,85842	S		
126	UTÅKER	Hordaland	KVINNHERRAD	59,7783	5,9024	S		
127	TENDALSVIK	Hordaland	ETNE	59,7361	5,884	S		
128	MOLNES	Hordaland	ETNE	59,7199	5,85798	S		
129	LINDVIK	Rogaland	VINDAFJORD	59,4741	5,97643	S		
130	HERØY	Rogaland	VINDAFJORD	59,4664	6,0262	S		
131	HETTANESET	Rogaland	VINDAFJORD	59,4586	5,8491	S		
132	RINGJA	Rogaland	TYSVÆR	59,4055	5,8631	S		
133	HALSAVIKA	Rogaland	TYSVÆR	59,3923	5,88333	S		
134	VINDSVIK	Rogaland	HJELMELAND	59,31	6,2	S		
135	KJEAHOLA	Rogaland	FINNØY	59,2933	6,0729	S		
136	MUNKHOLMEN	Rogaland	TYSVÆR	59,2855	5,63132	S		
137	FOSSÅ	Rogaland	HJELMELAND	59,2699	6,14295	S		
138	BASTLI	Rogaland	HJELMELAND	59,2694	6,19308	S		
139	ANDREGGRUNN	Rogaland	FINNØY	59,2389	5,74062	S		
140	LANGAVIKA	Rogaland	HJELMELAND	59,2321	6,00832	S		
141	ATLATVEIT	Rogaland	FINNØY	59,2277	5,98423	S		
142	DJUPEVIK	Rogaland	FINNØY	59,1474	5,86378	S		
143	KOBBAVIKA	Rogaland	FINNØY	59,1223	5,93443	S		
144	NAPP	Vest-Agder	FLEKKEFJORD	58,2487	6,53675	S		
145	SKIPNINGSDALB	Vest-Agder	FLEKKEFJORD	58,2415	6,5958	S		
146	STØYTLAND	Vest-Agder	FLEKKEFJORD	58,2316	6,62853	S		
147	PINNEN	Vest-Agder	FLEKKEFJORD	58,2255	6,68272	S		
148	SALVÅGVIKA	Vest-Agder	FLEKKEFJORD	58,2234	6,63615	S		
149	TARMEVIKODDE	Vest-Agder	FLEKKEFJORD	58,2198	6,68563	S		
150	DJUPEVIKA	Vest-Agder	FLEKKEFJORD	58,2197	6,65217	S		
151	BUKSEVIKA	Vest-Agder	FLEKKEFJORD	58,2074	6,63505	S		
							Sum	0

C Final Mathematical Model Formulation

$$\min \sum_{v \in V} \sum_{r \in R_v} C_{vr} x_{vr} + \sum_{e \in E} \sum_{i \in N} C_{ie}^e y_{ie} \quad (\text{C.1})$$

$$\sum_{v \in V} \sum_{r \in R_v} A_{ivr} x_{vr} + \sum_{e \in E} y_{ie} \geq S_i, \quad i \in N \quad (\text{C.2})$$

$$\sum_{r \in R} T_{vr} x_{vr} \leq T, \quad v \in V \quad (\text{C.3})$$

$$\sum_{i \in N} D_i y_{ie} \leq Q_e, \quad e \in E \quad (\text{C.4})$$

$$\sum_{i \in N} D_i y_{ie} \leq 1, \quad e \in E \quad (\text{C.5})$$

$$\sum_{v \in V} \sum_{r \in R_v} F_{ivr} x_{vr} + \sum_{e \in E} D_i y_{ie} \leq Q_i \quad i \in N \quad (\text{C.6})$$

$$\sum_{v \in V} \sum_{r \in R_v} F_{ivr} x_{vr} + \sum_{e \in E} D_i y_{ie} \geq D_i \quad i \in N \quad (\text{C.7})$$

$$\sum_{r \in R} x_{vr} \leq n, \quad v \in V \quad (\text{C.8})$$

$$\sum_{r \in R} x_{vr} \leq M w_v, \quad v \in V \quad (\text{C.9})$$

$$\sum_{v \in V} w_v \leq m \quad (\text{C.10})$$

$$x_{vr} \in \{0, 1\} \quad v \in V, r \in R_v \quad (\text{C.11})$$

$$y_{ie} \in \{0, 1\} \quad i \in N, e \in E \quad (\text{C.12})$$

$$w_v \in \{0, 1\}, \quad v \in V \quad (\text{C.13})$$

D Optimization Outputs

D.1 Medium Demand Results

Regional Method

```

Feed Distribution Plan
=====

```

	Ship 1	Ship 1	Ship 2	Ship 2	Ship 2	Ship 3	Ship 3	Ship 4	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	---	---	0	0	0
Cluster 2	---	1350	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 3	---	---	---	---	---	---	---	---	---	---	0	0	0
Cluster 4	---	750	---	---	---	---	---	---	---	---	750	750	1500
Cluster 5	---	---	---	---	3000	---	---	---	---	---	4500	4050	8100
Cluster 6	---	---	---	---	---	1500	---	---	---	---	3000	2400	4800
Cluster 7	---	---	---	3000	---	---	---	---	---	---	4500	4500	9000
Cluster 8	---	---	2100	---	---	---	---	---	---	---	2100	2100	4200
Cluster 9	---	2850	---	---	---	---	---	---	---	---	2850	2850	5700
Cluster 10	---	---	---	---	---	---	---	1200	---	---	1200	1200	2400
Cluster 11	---	---	---	---	---	---	---	---	---	---	0	3450	6900
TOTAL	2850	2100	2100	3000	3000	3000	1500	1500	1500	1500	20250		
Ship Capacity	3000	3000	3000	3000	3000	1500	1500	1500	1500	1500			
Sail time	68	97	40	46	77	37	95	76	55	37			
Cost (\$)	6229	8865	3705	4255	7015	1987	5025	4026	2919	1987			

Cluster 11 supplied by Charter Ship 3 for 23000 \$

Total cost of distribution = 69013 \$.

Natural Method

Feed Distribution Plan

	Ship 1	Ship 1	Ship 1	Ship 2	Ship 2	Ship 3	Ship 3	Ship 3	Ship 3	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	---	---	1350	1350	1350	2700
Cluster 2	---	---	---	---	750	---	---	---	---	---	---	750	750	1500
Cluster 3	---	---	---	---	600	---	---	---	---	---	---	600	600	1200
Cluster 4	---	---	---	---	750	---	---	---	---	---	---	750	750	1500
Cluster 5	---	---	---	---	---	---	---	---	---	---	---	1500	1500	3000
Cluster 6	---	---	---	---	---	---	---	---	1200	---	---	1200	1200	2400
Cluster 7	---	---	---	---	450	---	---	---	---	---	---	450	450	900
Cluster 8	---	---	---	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 9	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 10	---	---	---	---	---	---	---	---	1050	---	---	150	150	300
Cluster 11	---	---	---	---	---	---	---	---	---	---	---	1950	1950	3900
Cluster 12	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 13	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 14	---	---	---	---	450	---	---	---	---	---	---	450	450	900
Cluster 15	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 16	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 17	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 18	---	---	---	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 19	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 20	---	---	---	---	---	---	---	---	---	---	---	900	900	1800
Cluster 21	---	---	---	---	---	---	---	---	---	---	---	750	750	1500
Cluster 22	---	---	---	---	---	---	---	---	---	---	---	1800	1800	3600
Cluster 23	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
TOTAL	3000	3000	3000	3000	2550	1500	1200	1200	1500	1350	1350	22650		
Ship Capacity	3000	3000	3000	3000	3000	1500	1500	1500	1500	1500	1500			
Sail time	75	57	25	85	57	40	17	32	42	27	115			
Cost (\$)	6892	5253	2278	7749	5233	2144	939	1699	2235	1424	6071			

Total cost of distribution = 41857 \$.

K-means Method

Feed Distribution Plan

	Ship 1	Ship 1	Ship 2	Ship 2	Ship 3	Ship 3	Ship 3	Ship 3	Ship 4	Ship 4	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	1350	---	---	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 2	---	750	---	---	---	---	---	---	---	---	---	---	750	750	1500
Cluster 3	---	600	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 4	---	---	---	---	---	---	---	---	---	---	---	---	750	750	1500
Cluster 5	---	---	---	---	---	---	---	---	---	---	---	---	1500	1500	3000
Cluster 6	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 7	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 8	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 9	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 10	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 11	---	---	---	---	---	---	---	---	---	---	---	---	150	150	300
Cluster 12	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 13	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 14	---	---	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 15	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 16	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 17	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 18	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 19	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 20	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 21	---	---	---	---	---	---	---	---	---	---	---	---	900	900	1800
Cluster 22	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 23	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 24	---	---	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 25	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 26	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 27	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 28	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 29	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 30	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
TOTAL	3000	2700	3000	3000	3000	1500	1500	1350	1050	1500	1350	1500	22650		
Ship Capacity	3000	3000	3000	3000	3000	1500	1500	1500	1500	1500	1500	1500			
Sail time	56	100	87	75	15	16	28	16	35	48	32	42			
Cost (\$)	5119	9136	7981	6874	798	866	1479	872	1868	2558	1699	2235			

Total cost of distribution = 41485 \$.

D.2 High Demand Results

* High demand results for regional method does not exist

Natural Method

Feed Distribution Plan

	Ship 1	Ship 1	Ship 1	Ship 2	Ship 2	Ship 2	Ship 3	Ship 3	Ship 3	Ship 3	Ship 4	Ship 4	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	0	2250	2700
Cluster 2	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1250	1250	1500
Cluster 3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 4	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1250	1250	1500
Cluster 5	---	---	2500	---	---	---	---	---	---	---	---	---	---	---	2500	2500	3000
Cluster 6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	2000	2000	2400
Cluster 7	---	---	450	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 8	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 9	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 10	---	---	---	---	---	---	---	---	---	---	---	---	---	---	0	150	300
Cluster 11	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1950	1950	3900
Cluster 12	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 13	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 14	---	---	450	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 15	---	---	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 16	---	---	1200	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 17	---	---	1200	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 18	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 19	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 20	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1500	1500	1800
Cluster 21	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1250	1250	1500
Cluster 22	---	---	---	---	---	---	---	---	---	---	---	---	---	---	3000	3000	3600
Cluster 23	---	---	---	---	---	---	---	---	---	---	---	---	---	---	2000	2000	2400
TOTAL	3000	2850	2950	3000	3000	2750	3000	1350	1050	1500	1250	1050	1350	1500	27850		
Ship Capacity	3000	3000	3000	3000	3000	3000	1500	1500	1500	1500	1500	1500	1500	1500			
Sail time	80	48	40	73	69	25	59	28	35	46	16	27	62	63			
Cost (\$)	7329	4434	3672	6650	6352	2278	3109	1475	1868	2452	861	1424	3317	3325			

Cluster 1 supplied by Charter Ship 3 for 9000 \$

Cluster 10 supplied by Charter Ship 4 for 1000 \$

Total cost of distribution = 58546 \$.

K-means Method

	Ship 1	Ship 1	Ship 1	Ship 2	Ship 2	Ship 2	Ship 3	Ship 3	Ship 3	Ship 3	Ship 3	Ship 4	Ship 4	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	0	2250	2700
Cluster 2	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1250	1250	1500
Cluster 3	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 4	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1250	1250	1500
Cluster 5	---	---	---	---	---	2500	---	---	---	---	---	---	---	---	---	2500	2500	3000
Cluster 6	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	2000	2000	2400
Cluster 7	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 8	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 9	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 10	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 11	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	150	150	300
Cluster 12	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 13	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 14	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 15	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 16	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 17	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 18	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 19	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 20	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 21	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	900	900	1800
Cluster 22	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 23	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 24	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 25	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 26	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 27	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1000	1000	1200
Cluster 28	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	2000	2000	2400
Cluster 29	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	1750	1750	2100
Cluster 30	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	0	2000	2400
TOTAL	3000	2950	2000	2750	2500	2250	1350	1500	1500	1500	1250	1500	1050	1200	1200	26000		
Ship Capacity	3000	3000	3000	3000	3000	3000	1500	1500	1500	1500	1500	1500	1500	1500	1500			
Sail time	76	62	29	74	38	54	42	35	15	46	20	14	21	26				
Cost (\$)	6998	5650	2649	6736	3459	4938	2255	1896	798	2452	1085	782	1136	1400				

Cluster 1 supplied by Charter Ship 1 for 9000 \$

Cluster 30 supplied by Charter Ship 4 for 8000 \$

Total cost of distribution = 59234 \$.

D.3 Low Demand Results

Regional Method

```

Feed Distribution Plan
=====

```

	Ship 1	Ship 1	Ship 2	Ship 2	Ship 2	Ship 3	Ship 3	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	0	0	0
Cluster 2	---	---	---	---	---	---	---	675	675	675	2700
Cluster 3	---	---	---	---	---	---	---	---	0	0	0
Cluster 4	---	---	---	---	---	---	---	375	375	375	1500
Cluster 5	---	---	---	---	2024	---	---	2024	2024	2025	8100
Cluster 6	---	---	---	2400	---	---	---	---	2400	2400	4800
Cluster 7	---	3000	---	---	---	---	---	---	4500	4500	9000
Cluster 8	---	---	---	---	---	900	---	---	2174	2100	4200
Cluster 9	---	---	2549	---	---	---	---	---	2549	2550	5700
Cluster 10	---	---	---	---	---	600	---	---	600	600	2400
Cluster 11	1724	---	---	---	---	---	---	---	1724	1725	6900
TOTAL	2998	3000	2549	2400	2024	1500	1500	1050	17021		
Ship Capacity	3000	3000	3000	3000	3000	1500	1500	1500			
Sail time	114	46	67	31	68	97	55	119			
Cost (\$)	10462	4255	6120	2841	6278	5152	2919	6320			

Total cost of distribution = 44347 \$.

Natural Method

Feed Distribution Plan

	Ship 1	Ship 2	Ship 2	Ship 2	Ship 3	Ship 3	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	0	675	2700
Cluster 2	---	---	---	---	---	---	---	---	375	375	1500
Cluster 3	---	---	---	---	---	---	---	---	300	300	1200
Cluster 4	---	---	---	---	---	---	---	---	375	375	1500
Cluster 5	---	---	---	---	---	---	---	---	750	750	3000
Cluster 6	---	---	---	---	---	---	---	---	600	600	2400
Cluster 7	---	---	---	---	---	---	---	---	450	450	900
Cluster 8	---	---	---	---	---	---	---	---	1350	1350	2700
Cluster 9	---	---	---	1050	---	---	---	---	1050	1050	2100
Cluster 10	---	---	---	---	---	---	---	---	150	150	300
Cluster 11	---	---	1950	---	---	---	---	---	1950	1950	3900
Cluster 12	---	---	---	1050	---	---	---	---	1050	1050	2100
Cluster 13	---	---	450	---	---	---	---	---	450	450	900
Cluster 14	---	---	---	---	---	---	---	---	0	450	900
Cluster 15	---	---	---	---	---	---	---	---	674	600	1200
Cluster 16	---	---	---	---	1199	---	---	---	1199	1200	2400
Cluster 17	---	1200	---	---	---	---	---	---	1200	1200	2400
Cluster 18	---	---	1350	---	---	---	---	---	1350	1350	2700
Cluster 19	---	---	---	---	---	---	---	---	300	300	1200
Cluster 20	---	---	---	---	---	---	---	---	450	450	1800
Cluster 21	---	---	---	---	---	---	---	---	375	375	1500
Cluster 22	---	---	---	---	---	---	---	---	900	900	3600
Cluster 23	---	---	---	---	---	---	---	---	600	600	2400
TOTAL	3000	3000	3000	1050	1498	1500	1350	1500	15898		
Ship Capacity	3000	3000	3000	3000	1500	1500	1500	1500			
Sail time	93	56	25	13	43	46	27	67			
Cost (\$)	8522	5129	2278	1231	2274	2439	1424	3557			

Cluster 1 supplied by Charter Ship 1 for 9000 \$

Cluster 14 supplied by Charter Ship 3 for 3000 \$

Total cost of distribution = 38854 \$.

K-means Method

Cluster 20 supplied by Charter Ship 4 for 3000 \$
 Total cost of distribution = 44415 \$.

Feed Distribution Plan													
	Ship 1	Ship 2	Ship 2	Ship 3	Ship 3	Ship 3	Ship 4	Ship 4	Ship 4	Ship 4	TOTAL	Demand	Capacity
Cluster 1	---	---	---	---	---	---	---	---	---	---	0	675	2700
Cluster 2	---	---	---	---	---	---	---	---	---	---	0	375	1500
Cluster 3	---	---	---	---	---	---	---	---	---	---	300	300	1200
Cluster 4	---	---	---	---	---	---	---	---	---	---	375	1500	1500
Cluster 5	---	---	---	---	---	---	---	---	---	---	750	750	3000
Cluster 6	---	---	---	---	---	---	---	---	---	---	600	600	2400
Cluster 7	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 8	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 9	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 10	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 11	---	---	---	---	---	---	---	---	---	---	150	150	300
Cluster 12	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 13	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 14	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 15	---	---	---	---	---	---	---	---	---	---	1200	1200	2400
Cluster 16	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 17	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 18	---	---	---	---	---	---	---	---	---	---	450	450	900
Cluster 19	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 20	---	---	---	---	---	---	---	---	---	---	0	450	900
Cluster 21	---	---	---	---	---	---	---	---	---	---	900	900	1800
Cluster 22	---	---	---	---	---	---	---	---	---	---	600	600	1200
Cluster 23	---	---	---	---	---	---	---	---	---	---	1050	1050	2100
Cluster 24	---	---	---	---	---	---	---	---	---	---	300	300	600
Cluster 25	---	---	---	---	---	---	---	---	---	---	300	300	1200
Cluster 26	---	---	---	---	---	---	---	---	---	---	300	300	1200
Cluster 27	---	---	---	---	---	---	---	---	---	---	300	300	1200
Cluster 28	---	---	---	---	---	---	---	---	---	---	600	600	2400
Cluster 29	---	---	---	---	---	---	---	---	---	---	525	525	2100
Cluster 30	---	---	---	---	---	---	---	---	---	---	600	600	2400
TOTAL	2625	2850	1050	1500	1200	1425	900	1500	1200	1200	15450		
Ship Capacity	3000	3000	3000	1500	1500	1500	1500	1500	1500	1500			
Sail time	95	66	26	16	18	62	20	15	34	27			
Cost (\$)	8711	6011	2396	879	968	3292	1078	816	1821	1443			

Cluster 1 supplied by Charter Ship 1 for 9000 \$
 Cluster 2 supplied by Charter Ship 3 for 5000 \$

E Problem Description Form



NTNU Trondheim
Norwegian University of Science and Technology
Department of Marine Technology

MASTER THESIS IN MARINE TECHNOLOGY

SPRING 2017

For stud.techn.

Sondre Thygesen & Marius Gyberg Haugland

“Use of Clusters in a Route Generation Heuristic for Distribution of Fish Feed”

Background

The production and value of fish farming in Norway has never been larger than it is today. Salmon farming accounts for 95% of the total industry, and the industry is also in rapid growth and it is expected to keep growing. Among the biggest expenses in fish farming are costs associated with fish feed, which accounts for approximately half of the salmon production cost. To be able to keep the growing industry as profitable as possible, the importance of efficient scheduling and planning for the distribution of this large cost-contributor cannot be underestimated.

The problem we are facing in this thesis is motivated by a real logistics problem that Marine Harvest is currently dealing with. The system they are serving are both large and complex, and hence the distribution planning is both an uncertain and time-consuming process. They have seen problems in how the planning process is performed today, as everything is done manually in large spreadsheets and is extensively based on experience. To stay competitive, and to more efficiently plan the routing for minimizing operational costs, a more automated process is desired.

Objective

The overall objective of this paper is to develop a tool that can be used for minimizing the operational costs for the fish feed distribution process, while at the same time meeting all requirements and demands of the system at all times. Only operational costs are considered, since fixed costs related to ships are not affected by the decisions to be made. The final model should help the distribution planner to reduce operation costs in the feed distribution, and also reduce the time used for planning the routing of the fleet. The tool should also be created in a way that it can be used for comparing different fleet configurations for evaluating the best way to distribute cargo in different systems.

Tasks

We have covered the following main tasks in the thesis:

- a. Review state of art within transportation problems.
- b. Describe the challenges of producing transportation plans for distribution of fish feed from factory to a large number of fish farms, and why this is a complicated process to optimize.
- c. Explore possible ways to reduce the problem size and complexity.
- d. Investigate possible models in different scenarios to determine the best way to create most optimal routes for a fixed fleet of ships to meet all requirements and demands in the system.
- e. Test if the current fleet used in the system is the optimal fleet for performing the distribution



General

In the thesis the candidate shall present his personal contribution to the resolution of a problem within the scope of the thesis work.

Theories and conclusions should be based on a relevant methodological foundation that through mathematical derivations and/or logical reasoning identify the various steps in the deduction.

The candidate should utilize the existing possibilities for obtaining relevant literature.

The thesis should be organized in a rational manner to give a clear statement of assumptions, data, results, assessments, and conclusions. The text should be brief and to the point, with a clear language. Telegraphic language should be avoided.

The thesis shall contain the following elements: A text defining the scope, preface, list of contents, summary, main body of thesis, conclusions with recommendations for further work, list of symbols and acronyms, reference and (optional) appendices. All figures, tables and equations shall be numerated.

The supervisor may require that the candidate, in an early stage of the work, present a written plan for the completion of the work.

The original contribution of the candidate and material taken from other sources shall be clearly defined. Work from other sources shall be properly referenced using an acknowledged referencing system.

Deliverable

- The thesis shall be submitted in two (2) copies:
- Signed by the candidate
- The text defining the scope included
- In bound volume(s)
- Drawings and/or computer prints that cannot be bound should be organized in a separate folder.
- The bound volume shall be accompanied by a CD or DVD containing the written thesis in Word or PDF format. In case computer programs have been made as part of the thesis work, the source code shall be included. In case of experimental work, the experimental results shall be included in a suitable electronic format.

Supervision:

Main supervisor: Bjørn Egil Asbjørnslett

Sub-supervisor: Inge Norstad

Company contact:

Deadline: 11.06.2017