NTNU – Trondheim
Norwegian University of
Science and Technology

Optimal Salmon Production

Producing Smolt Under Uncertainty

Kristian Frøystein
Endre Hegland Kure

Industrial Economics and Technology Management
Submission date: July 2013
Supervisor: Asgeir Tomasgard, IØT

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

NTNU
Det skapende universitet

SAMARBEIDSKONTRAKT

1. Studenter i samarbeidsgruppen

Etternavn, fornavn	Fødselsdato
Frøystein, Kristian	**06. mai 1988**
Etternavn, fornavn	Fødselsdato
Kure, Endre Hegland	**24. feb 1988**

2. Hovedveileder

Etternavn, fornavn	Institutt
Tomasgard, Asgeir	**Institutt for industriell økonomi og teknologiledelse**

3. Masteroppgave

Oppgavens (foreløpige) tittel
Optimal Salmon Production **Producing Smolt Under Uncertainty**

4. Bedømmelse

Kandidatene skal ha *individuell* bedømmelse ☐
Kandidatene skal ha *felles* bedømmelse ☒

Trondheim, 13. juni 2013
Sted og dato

Kristian Frøystein
Kristian Frøystein

Hovedveileder

Endre Hegland Kure

Originalen oppbevares på instituttet.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Frøystein, Kristian**	Fødselsdato **06. mai 1988**
E-post **kristian.froystein@gmail.com**	Telefon **47668661**

2. Studieopplysninger

Fakultet **Fakultet for samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **22. feb 2013**	Innleveringsfrist **19. jul 2013**
Oppgavens (foreløpige) tittel **Optimal Salmon Production** **Producing Smolt Under Uncertainty**	
Oppgavetekst/Problembeskrivelse This thesis will contain a model for parts of the salmon production value chain. The model should consider the most important restrictions and regulations to the production. The model is developed for a a salmon producer with many facilities. The thesis will include the following elements - Presentation and discussion of the problem - Discussion of the relevant uncertainty parameters - Formulation and implementation of a model - Discussion of the results	
Hovedveileder ved institutt **Professor Asgeir Tomasgard**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

Trondheim, 13. juni 2013

Sted og dato

Student **Hovedveileder**

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

NTNU
Det skapende universitet

MASTERKONTRAKT
- uttak av masteroppgave

1. Studentens personalia

Etternavn, fornavn **Kure, Endre Hegland**	Fødselsdato **24. feb 1988**
E-post **endre.kure@gmail.com**	Telefon **48078507**

2. Studieopplysninger

Fakultet **Fakultet for samfunnsvitenskap og teknologiledelse**	
Institutt **Institutt for industriell økonomi og teknologiledelse**	
Studieprogram **Industriell økonomi og teknologiledelse**	Hovedprofil **Investering, finans og økonomistyring**

3. Masteroppgave

Oppstartsdato **22. feb 2013**	Innleveringsfrist **19. jul 2013**
Oppgavens (foreløpige) tittel **Optimal Salmon Production** **Producing Smolt Under Uncertainty**	
Oppgavetekst/Problembeskrivelse This thesis will contain a model for parts of the salmon production value chain. The model should consider the most important restrictions and regulations to the production. The model is developed for a a salmon producer with many facilities. The thesis will include the following elements - Presentation and discussion of the problem - Discussion of the relevant uncertainty parameters - Formulation and implementation of a model - Discussion of the results	
Hovedveileder ved institutt **Professor Asgeir Tomasgard**	Medveileder(e) ved institutt
Merknader **1 uke ekstra p.g.a påske.**	

4. Underskrift

Student: Jeg erklærer herved at jeg har satt meg inn i gjeldende bestemmelser for mastergradsstudiet og at jeg oppfyller kravene for adgang til å påbegynne oppgaven, herunder eventuelle praksiskrav.

Partene er gjort kjent med avtalens vilkår, samt kapitlene i studiehåndboken om generelle regler og aktuell studieplan for masterstudiet.

13/6 -2013
....
Sted og dato

....
Student

....
Hovedveileder

Originalen lagres i NTNUs elektroniske arkiv. Kopi av avtalen sendes til instituttet og studenten.

Sammendrag

Laksenæringen har endret karakter fra å være preget av entreprenører til industrialisering og kostnadseffektivisering. Store aktører kjøper opp de små aktørene og næringen konsolider. Optimering er derfor blitt mer relevant, men for ferskvannsdelen av verdikjeden er det publisert lite akademisk arbeid.

Denne avhandlingen presenterer en lineær stokastisk programmeringsmodell som minimerer de totale forventede kostnader tilknyttet smoltproduksjon. En av dens nyvinninger er en lineær formulering ved hjelp SOS2, slik at vanntemperaturen kan være en beslutningsvariabel. Den største usikkerheten, ferskvanntemperaturen, er modellert gjennom bruk av scenarioer, som er generert ved hjelp av en sesongavhenging AR(1)-modell. Testing og tilbakemeldingsøkter med biologer ble gjennomført for å sikre modellkvaliteten.

Modellen ble anvendt på to case-studier av Marine Harvest sitt ferskvann anlegget på Slørdal i Sør-Trøndelag. Det første studiet undersøker en typisk to årlig produksjonsplan, mens det andre studiet utforsker produksjon av et nytt produkt, smolt på 500 gram. Basert på modellkjøringene kan vi trekke tre konklusjoner. For det først, optimering kan ha en betydelig innvirkning på de totale kostnadene. En typisk to års produksjonsplan for anlegget på Slørdal fikk en total kostnadsreduksjon på 11 %. For det andre, smoltbestillinger kan bli oppfylt med redusert oppvarming av vann i forhold til dagens praksis hvis yngelen blir tatt tidligere inn i anlegget. For det tredje, 500 gram smolt bør leveres om vinteren for å utnytte de naturlige temperatursvingningene. Smolt levert om vinteren har betydelig lavere kostnader og kortere produksjonstid enn ved andre årstider.

Abstract

The salmon farming industry is shifting from an entrepreneurial spirit towards industrialization and cost efficiency. Large players are purchasing the small players, and the industry is consolidating. Optimization is therefore becoming more relevant, but little academic work has been done on the freshwater part of the value chain.

This thesis presents a linear stochastic programming model that minimizes the total expected costs related to smolt production. One of its innovations is a linear formulation using SOS2, enabling water temperature to be a decision variable. The main uncertainty, freshwater intake temperature, is modeled through the use of scenarios, which were generated using a seasonal AR(1)-model. Testing and feedback sessions with biologists were conducted to ensure model quality.

The model was applied to two case studies involving the Marine Harvest freshwater facility at Slørdal, Sør-Trøndelag. The first case investigates a typical two year production plan, while the other explores the production of 500 grams smolts, which is a new product. These case studies have yielded three core insights. First, they indicate that optimization can have a significant impact on the total costs. A typical two year production plan at the Slørdal facility experienced a total cost reduction of 11%. Second, smolt orders can be fulfilled with reduced water heating compared to today's praxis if the fry are deployed earlier. Third, 500 grams smolt should be delivered during the winter to exploit the natural temperature seasonalities. Smolt delivered during the winter has significantly lower cost and production time compared to other seasons.

Preface

This master thesis is the end of our Master of Science at the Norwegian University of Science and Technology (NTNU) with specialization in Financial Engineering at the Department of Industrial Economics and Technology Management.

First and foremost, we would like to thank our supervisor Professor Asgeir Tomasgard for stimulating discussions and valuable counsel. We would also like to thank Gerardo Alfredo Perez Valdes, postdoc at NTNU, for helping us run the model on the computational cluster. Further, we would like to thank Professor Bo Lundqvist and Professor Håkon Tjelmeland for helpful counsel regarding the statistical methods used.

Our thesis could not have been written without the fruitful collaboration with Marine Harvest. We would like to thank Anders Jon Fjellheim, Ole Christian Norvik and Stian Aspaas for sharing their valuable industry knowledge, and for motivation us by showing patience and encouragement regarding our work.

Lastly, a special thanks to our friends and family for feedback and support with our thesis.

Trondheim, 7. July 2013

Kristian Frøystein | Endre Kure

Abbreviations and definitions

Abbreviations

ACF	-	Autocorrelation Function
AR	-	Autoregressive
ARIMA	-	Autoregressive Integrated Moving Average
ARMA	-	Autoregressive Moving Average
B&B	-	Branch-and-Bound
BC	-	Box-Cox (power transformation)
EVPI	-	Expected Value of Perfect Information
GARCH	-	Generalized Autoregressive Conditional Heteroskedasticity
i.i.d.	-	Independent and identically distributed (random variables)
LP	-	Linear Programming
MA	-	Moving Average
MAB	-	Maximum Allowed Biomass
OLS	-	Ordinary Least Squares
SOS2	-	Special Ordered Set of type 2
PACF	-	Partial Autocorrelation Function
VSS	-	Value of Stochastic Solution

Definitions

Biomass	-	Technical term used to describe the fish mass in a facility.
Culling	-	The act of shaping the fish population by selective slaughter.
Day degrees	-	The product of the average daily temperature and the number of days at that temperature. E.g 10 days of 5°C gives a total day degrees of 50.
Freshwater intake	-	Intake at a nearby lake which supplies water to the smolt facility.

Table of contents

Chapter 1

Introduction to salmon farming

The salmon farming industry is shifting from an entrepreneurial spirit towards industrialization and cost efficiency. Large players are purchasing the small players, and the industry is consolidating. Optimization is therefore becoming more and more relevant.

Farmed salmon spend the first 10-18 months in freshwater before it becomes smolt and is able to survive in seawater. This stage is the foundation for the salmon's growth at sea and is of paramount importance. However, little academic work has been done on optimizing the freshwater part of the value chain.

This thesis presents an optimization model for the freshwater part of the salmon value chain. The model objective is to minimize costs given a requirement to meet smolt orders and considers both the most important costs, decisions and uncertainties faced by a smolt production facility.

First, chapter 1 gives an introduction to salmon farming with a focus on smolt production. The literature relevant to smolt optimization is then reviewed in chapter 2. Chapter 3 defines the model scope and gives an outline of the stochastic programming model. This model is formulated mathematically in chapter 4. The model is parametrized for a Marine Harvest freshwater facility in chapter 5. Chapter 6 presents the results from two case studies along with sensitivity analyses. Lastly, chapter 7 concludes and discusses further work.

1.1 Industry history

Salmon farming began in the late 1960s on an experimental level in Norway and the UK. Chile and North America followed in the 1970s. These regions have since dominated the industry due to the water temperature and sheltered coastal areas. Together, they have a market share of 94% as shown in figure 1.1.

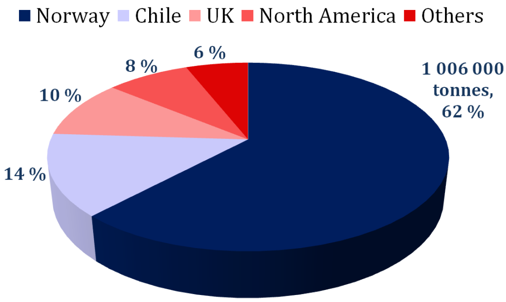

Figure 1.1: The market shares of salmon farming regions (Norwegian Seafood Council, 2012b)

Norway has by far the largest salmon farming industry in the world, with a 62% market share in 2011. During the past two decades, the Norwegian salmon farming industry has grown with a compounded annual rate of 10% (see figure 1.2).

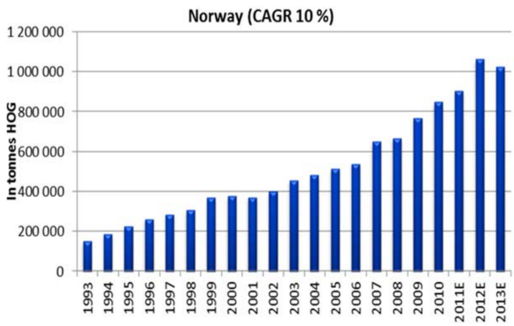

Figure 1.2: The total biomass produced by the Norwegian salmon farming industry (Marine Harvest, 2013c).

Norway is also the world's largest exporter, with 97% of its total production exported (Marine Harvest, 2013c). In 2012 51.6 billion NOK worth of seafood was exported, with 29.6 billion NOK being salmon (Norwegian Seafood Council, 2012a), making seafood the third largest export industry in Norway. It is exported all over the world, with EU, Eastern-Europe and Asia being the largest markets.

1.2 The salmon value chain

The salmon value chain can be split into three stages: freshwater production, seawater production, and sales and distribution, as illustrated in figure 1.3.

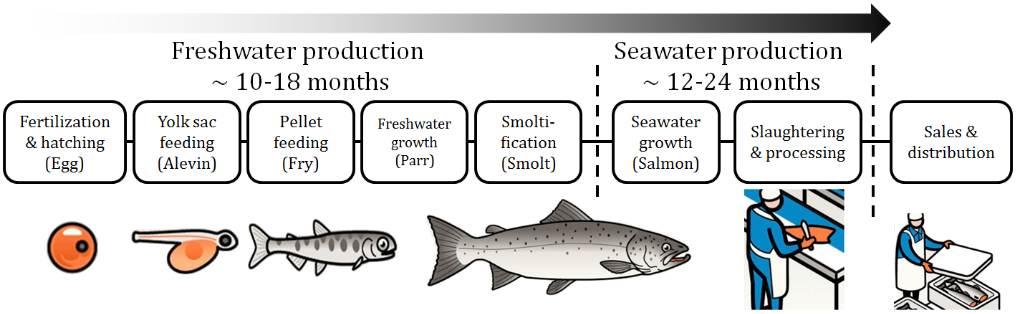

Figure 1.3: The stages of a typical salmon value chain (Marine Harvest, 2012). The name used for the fish at each life stage is shown in parenthesis.

Typically, all of these divisions are vertically integrated and tightly coordinated. This is essential when managing a biological supply chain, where both ethics and freshness are aspects of paramount importance. Based on Marine Harvest (2012) these three production stages are described in greater detail.

1.2.1 Freshwater production

The freshwater part of the production is split into five steps:

1. **Fertilization and Hatching (Egg)** The salmon life-cycle begins with the parent fish, called "*broodstock*". These are selected based on characteristics such as color, growth and overall robustness. When they become sexually mature, the eggs from the female are mixed with the sperm from the males. This produces *fertilized eggs*. The eggs develop a visible embryonic eye, and are therefore referred to as *eyed eggs*. The development rate of the eyed eggs is determined by the temperature in the tanks where they are kept. The industry praxis is to keep this temperature stable at 8°C (Hansen, 1998).

2. **Yolk sac phase (Alevin)** About three months after fertilization, the eggs hatch and tiny fish called *alevins* emerge. The alevins do not feed yet, but get nutrition from a yolk sac attached to them. This phase lasts for several weeks before the alevins start feeding.

3. **Pellet feeding (Fry)** Once the fish start feeding, they are named *fry*. Initially, fry weigh about 0.2 grams and are 2.5 cm long. The fry eat dry pellets, which are designed specially to match their nutritional requirements.

4. **Freshwater growth (Parr)** The fry are moved into larger tanks when they weigh about six grams. The fish will now change into a green-brown color and become *parr*. Under the right temperature conditions, the parr will grow very rapidly. Generally, a temperature of 15-16°C is optimal. When its weight reaches 50-70 grams, the parr is vaccinated to ensure robustness through

resistance to common diseases (Marine Harvest, 2013a). This is also required by law (Ministry of Fisheries and Coastal Affairs, 2008).

5. **Smoltification (Smolt)** Parr is called *smolt* after it has undergone a physiological change which enables it to survive in seawater. This process is called smoltification and can occur when the parr weighs at least 60-80 grams. Smoltification is then governed by the temperature and light conditions that the parr are exposed to (Hansen, 1998). The smolts are then transferred to seawater in large tanks aboard boats or trucks. If the smolt is not moved into seawater within a given time after becoming smolt, it may desmoltify and become parr again. This takes a great toll on the fish, weakens the fish and causes mortality. Desmoltification is therefore undesirable.

1.2.2 Seawater production

Once the smolt is deployed at sea, it is referred to as *salmon*. The salmons are kept in several net cages called *net pens* where they can grow and mature. Separating the fish in several net pens has many advantages. It enables the operator to maintain traceability while deploying smolt of different sizes and batches to a sea facility. Further, sorting and controlled transfers of specific fish types can be performed.

After about a year at sea, the salmon is ready to be slain and gutted. By then it has usually reached the desired market weight of 4.5 to 5.5 kg. The salmon can then be sold with or without head, fresh or frozen, or processed as fillets, steaks or portions.

1.2.3 Distribution and marketplaces

Atlantic salmon products are traded all over the world. The market has grown by 7% annually since 2000 (Marine Harvest, 2013c), and created a boom in both supply and demand. Salmon products are sold as both a commodity and differentiated product, and through both bilateral contracts and marketplaces (e.g. Fish Pool). The marketplace prices are highly volatile, and this risk is of paramount importance to salmon producers (Bergfjord, 2009).

When transporting fish, the priority concern is to maintain freshness. This is done through effective packing and an unbroken cold-chain. A combination of road, rail, ship and air freight is used to achieve this.

1.3 Production cost structure

The costs associated with farming salmon varies between different locations, since both environmental and regulatory conditions differ. A typical cost structure for 1 kg farmed salmon is shown in table 1.1. Feed is by far the largest cost, followed by slaughtering and processing. The third largest are the costs associated with smolt production.

Cost type	**NOK per kg salmon**
Feed	11.65
Slaughtering and processing	2.4
Smolt	**2.11**
Salary	1.38
Well boats	1.03
Maintenance	0.75
Depreciation	0.63
Sales and marketing	0.52
Mortality	0.3
Other	2.56
Total	**23.33**

Table 1.1: A breakdown of cost per kg salmon produced (Marine Harvest, 2013c).

Smolts are usually delivered from many different freshwater facilities, so segmenting the smolt cost is an intricate process. The freshwater facility at Slørdal is considered to have a typical cost segmentation. This breakdown is shown in table 1.2.

Cost type	**% of total costs**
Vaccination	19.8 %
Feed	13.8 %
Energy	10.3 %
Eggs	8.7 %
Oxygen	2.8 %
Insurance	0.6 %
Total variable costs	**56.0 %**
Total fixed costs	**44.0 %**

Table 1.2: A breakdown of total costs for the freshwater facility at Slørdal (Marine Harvest, 2013a).

1.4 Regulatory environment

The fish farming industry is regulated by various laws and governmental agencies to ensure ethical and sustainable operations. The three most important regulatory aspects regarding smolt production are:

1. **Smolt handling**: *Akvakulturdriftforskriften* is the main piece of legislation regulating fish handling at farms in Norway (Ministry of Fisheries and Coastal Affairs, 2008). Some of the aspects are legal smolt weight ranges, vaccination procedures and maximum fish density. The law is administrated by the Ministry of Fisheries and Coastal Affairs, while The Norwegian Food Safety Authority is in charge of ensuring that the actors are law abiding.

2. **Operating licenses**: Licenses for freshwater production are awarded by the Ministry of Fisheries and are administered by the Directorate of Fisheries. These dictate the maximum allowed biomass at a facility, and can be traded between players in the industry (Marine Harvest, 2013c). It takes up to several years to process these applications, making capacity expansion a slow process (Bjørgo, 2013).

3. **Freshwater intake**: Most smolt production facilities have dedicated water intakes from nearby lakes. The amount which can be drawn from these lakes is regulated by The Norwegian Water Resources and Energy Directorate through licenses. Thus, there is an upper bound in the available freshwater flow (Marine Harvest, 2013a).

1.5 Production uncertainty

Smolt production facilities face a wide range of uncertainties. These can be segmented into five categories as shown in figure 1.4. At the core is the *biomass development* in the facility, a term used to describe how the total fish mass evolves.

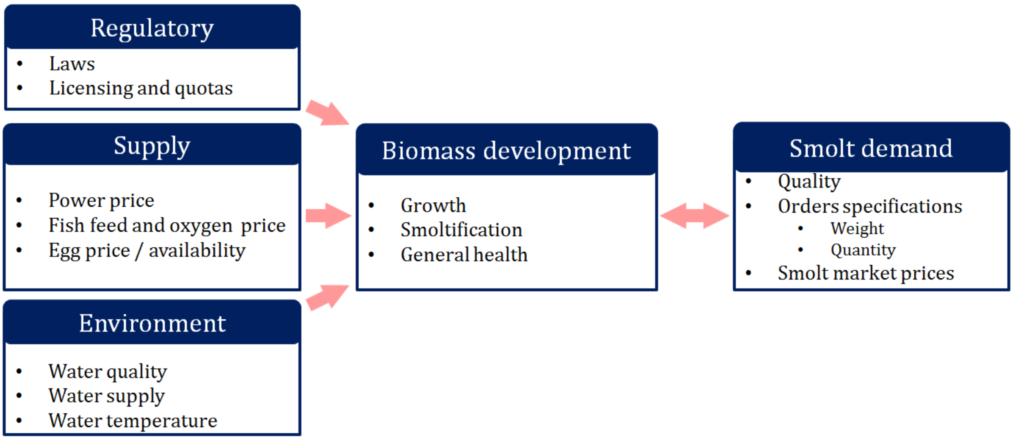

Figure 1.4: An overview of the uncertainty classes affecting smolt production. Red arrows denote causality between these groups.

1.5.1 Regulatory

The main regulatory aspects are new laws, assigning quotas and application processing (Bjørgo, 2013; Marine Harvest, 2013a). Neither are considered a relevant source of uncertainty within the production- and planning horizon of a smolt batch.

1.5.2 Supply

The main supply uncertainties are the power price, feed price, oxygen price, egg price and egg availability.

- **Power price:** Power prices are known to be one the most volatile commodities there is. It is vital to many users and cannot be stored at a reasonable cost. Heating water to ensure good growth conditions is common at smolt facilities and is often done using electricity. As shown in section 1.3 heating costs account for a significant amount ($\sim$10%) of the total costs at a facility, and they are therefore exposed to this risk. By using financial instruments it is possible to hedge this risk by fixing the power price. The uncertainty can therefore be reduced at a cost.

- **Feed and oxygen price:** Fish feed and oxygen are bought from external suppliers. The fish feed market is consolidated around three large actors: BioMar, Skretting and EWOS (Røsstad et al., 2013). Marine Harvest (2013a) has indicated that the feed costs can be considered stable on a time horizon of a couple of years. The same applies to the oxygen prices. So within the horizon of a production plan uncertainty in feed and oxygen prices are near non-existent.

- **Egg price and availability:** The egg price is stable since eggs are produced by the brooding division at Marine Harvest. Recent advances in production techniques makes it possible to deliver eggs throughout the whole year and no uncertainty is therefore related to the price or availability of eggs (Marine Harvest, 2013a).

1.5.3 Environment

The main environmental uncertainties are the freshwater intake temperature, quality and supply.

- **Freshwater quality:** The water quality at the freshwater intake is assessed based on pH-levels, presence of microbes and ion-levels. Strict environmental regulations in Norway and good procedures at Marine Harvest ensures that this is not a significant source of uncertainty (Marine Harvest, 2013a).

- **Freshwater availability:** The freshwater supply is regulated by the Norwegian Water Resources and Energy Directorate by licenses and there is a possibility that the water supply can run short (Marine Harvest, 2013a). Availability of freshwater is an uncertain factor which may impact the production severely. To our knowledge this has not occurred at any Marine Harvest facility and is therefore considered a small source of uncertainty.

- **Freshwater temperature:** Temperature is the main driver behind salmon growth. Since salmon are cold-blooded, the core biological processes, such as food uptake and metabolic rate, are largely governed by the their body temperature. Figure 1.5 shows how the water intake temperature varies significantly between different years. Due to the growth impact and difficulty related to long-term weather forecasting, the freshwater intake temperature is considered a large source of uncertainty.

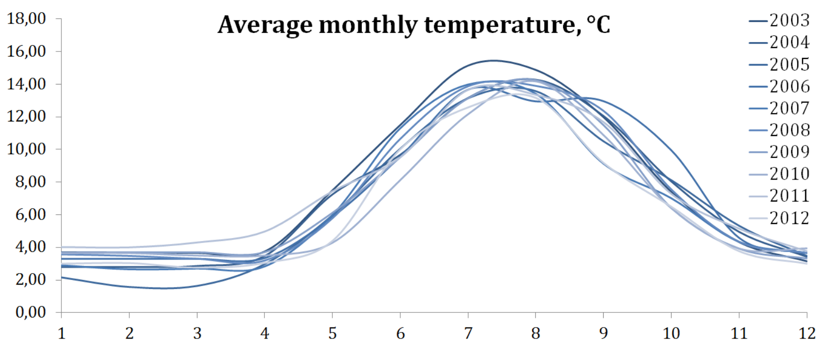

Figure 1.5: The average monthly temperature at the Slørdal freshwater facility in the period of 2003-2012.

1.5.4 Biomass development

The uncertainty in biomass development is relation to growth, smoltification and general fish health.

- **Growth:** Salmon growth is determined by feed availability, oxygen levels, fish handling, general fish health and temperature. It is standard industry praxis to saturate the production tanks with both feed and oxygen, and to keep the fish in general good health. Fish handling consist of sorting, moving and vaccinating. The fish are always starved before any of the mentioned events, thus causing a momentary halt in growth (Marine Harvest, 2013a). Vaccination against common diseases occurring at sea is required by law (Ministry of Fisheries and Coastal Affairs, 2008) and normally done when the fish is between 50 and 70 grams (Marine Harvest, 2013a). Side effects of the vaccination are reduced appetite and infections in the abdomen (Veterinærinstituttet, 2012a,b). It is required by law to keep fish movement and sorting at a minimum (Ministry of Fisheries and Coastal Affairs, 2008). Since all the events can be planned ahead in time, their influence on uncertainty are minimal. Issues affecting the general health, such as diseases and injuries, can reduce the growth significantly and are hard to predict. Thus the main uncertainties in growth are related to temperature and general fish health.

- **Smoltification:** The smoltification process is governed by temperature and light conditions in the tank. The production may be planned fairly accurately such that the parr become smolt within a desired month. This process naturally reverses after some time if the smolt is not moved to sea and the fish desmoltify and become parr again. This takes a great toll on the fish, and can cause death, disease or reduced growth. One way to prevent desmoltification is to have seawater tanks (Marine Harvest, 2013a), but this is not available at all facilities. Desmoltification rarely occurs unless the seawater division changes its orders and smoltification can be considered a low uncertainty factor given that the orders are certain.

- **General health:** If the fish is not healthy, reduced growth or death may occur. During the *hatching period* the water temperature has to be stable. If the temperature rises above 8°C, malformations occur at an increasingly higher rate (Marine Harvest, 2013a). Temperatures below 8°C prolong the hatching period, since the hatching time is governed by the number of *day degrees*. The most common cause of death during hatching are fungal infections such as Saprolegnia (Veterinærinstituttet, 2012a,b). Normally, a 2-3% mortality rate can be expected in the transition from egg to fry (Marine Harvest, 2013a). Due to the highly controlled environment in the hatchery there is very little variation in this rate. The health in the hatching period is therefore not considered to be uncertain (Marine Harvest, 2013a).

  In the period after hatching, disease outbreaks can have a great impact on both mortality and growth. Since the causes of outbreaks are not completely known it imposes a large source of uncertainty. The most common viral diseases are Infectious Pancreatic Necrosis (IPN) (Veterinærinstituttet, 2012b), Heart and Skeletal Muscle Inflammation (HSMI) (Veterinærinstituttet, 2012a) and Haemorrhagic Smolt Syndrome (HSS) (Nylund et al., 2003). The most common bacterial diseases are Flavobacterium Psychrophilum, Pseudomonas Fluorescens, Yersinia Ruckeri and Bacterial Kidney Disease (BKD) (Veterinærinstituttet, 2012a). Outbreaks are known to cause mortality rates as high as 50% (Veterinærinstituttet, 2012b). However, on average outbreaks only cause a 4 -10% mortality rate (Marine Harvest, 2013a).

1.5.5 Smolt demand

Smolt quality, order specifications and market prices are the main risk sources related to demand.

- **Smolt quality:** Smolt quality is hard to quantify as it is a culmination of the general health, size, growth and degree of smoltification. Good standard routines secure that all smolts are of acceptable quality and this is therefore not considered an uncertain factor (Marine Harvest, 2013a).
- **Orders specifications:** Order specifications have historically been stable, with the same amounts, weight classes and delivery dates each year. Events in the seawater division may delay when orders can be delivered and cause complications related to desmoltification (Marine Harvest, 2013a). Events are known to happen and this does therefore impose a medium source of uncertainty.
- **Smolt market prices:** The public smolt market is near non-existent and with market prices being more or less fixed. Marine Harvest produce smolt almost solely for internal use, and have on rare occasions sold spare smolt inventories (Marine Harvest, 2013a). The smolt price is therefore not a significant uncertainty factor.

1.5.6 Uncertainty summary

In collaboration with Marine Harvest the uncertainty sources were prioritized in terms of their impact on production. The *freshwater intake temperature* and *mortality* are considered the main sources of uncertainty.

Chapter 2

Literature review

There is a shortage of academic literature published on optimization of smolt production, but some literature exist on optimization in other biological systems such as ocean fisheries, shrimp farming and the seawater part of fish farming. The aim of this review is to discover techniques relevant for modeling smolt production.

2.1 Optimization of biological production

Clark et al. (1973) investigated how an actor can maximize profits in ocean fishing areas by using differential equations to describe the supply and demand. They derived the *"Fisher rule"*, which suggested that the fish should only be harvested if the growth rate is less than the interest rate. Their approach is not directly applicable to our problem, but it suggests that a micro economic approach could be used.

Bjørndal (1988) argued that fish farming has more in common with forestry and agriculture than with ocean fishing. He used a similar approach as Clark et al. (1973) to investigate the optimal harvesting strategy. Major costs such as feed, insurance and harvesting were included. The paper presents two case studies of salmon and turbot production. For salmon the growth was represented as a third degree polynomial depending solely on time, and was parameterized using regression. He divided the fish into weight classes and showed that optimal harvesting differs between the classes. He assumed a deterministic growth rate and one time investment in fish. For our purpose, this model would not handle interactions between various batches, control of growth, biomass restriction or the life stages of smolt. The approach is therefore viewed as less applicable.

Forsberg (1999) used a linear optimization model to investigate how two harvesting strategies would affect profit in fish farming. In order to do so he introduced two different growth models: one with batch growth and one with gradient growth. The batch growth uses a vector to describe growth for each batch. The gradient growth uses a time-varying Markov process and a transition matrix to model fish growth. The model formulation also includes biomass restrictions, feeding and transport costs. The author also discussed the applicability of dynamic programming and linear programming to a biological system, and concluded that the latter is the best approach. The model captures the most basic features that would also be faced by a freshwater facility. However, it lacks the option of controlling growth through temperature and modeling the life stages of smolt.

Song and Chen (2001) introduced a model where a species have two life stages, mature and immature, where only the mature can be harvested. Using ordinary differential equations they modeled how the size in the immature individuals will affect the size of the mature individuals. Based on boundary restrictions they derive a solution on the annual optimal harvesting. Although the model was developed for aquatic mammals, it is applicable to the two stages smolt experience before and after smoltifications. However, the approach is cumbersome and is therefore viewed as less applicable to larger systems such as a freshwater facility.

Yu et al. (2006) published a model for optimal pond scheduling on a shrimp farm where they consider the problems of stocking and harvesting shrimps. They used linear programming to schedule the use of multiple ponds and shrimp cycles. The model also includes growth and mortality scenarios. To calculate growth rates they used an Artificial Neural Network (ANN) and included a total of 18 input factors. They formulated the scheduling problem as a network flow problem in order to solve it efficiently. The freshwater facilities also face the scheduling problem since they are restricted by the number of tanks that can use. However, Yu et al. (2006) did not enable the shrimps to move between the ponds which would be expected in freshwater facilities, since fish change tanks through their life stages.

A few master theses have studied and optimized salmon farming production with linear programming. The modeling techniques used are highly relevant to a freshwater facility, and are these theses are therefore studied in detail.

Hæreid (2011) investigated the seawater part of the salmon value chain, focusing on harvesting and sales of salmon over a year. Using a stochastic programming model, he derived an optimal harvesting strategy. In the implemented model he only used a stochastic parameter on temperature, keeping mortality and sales price deterministic. The model consists of several weight classes where each class has a certain growth rate and transition rate for each time period. This model shares strong similarities with the gradient model of Forsberg (1999). Also, Hæreid assumed that the growth is only affected by temperature. The growth rate for each weight class is derived from a seawater growth model rate developed by Skretting. However, Hæreid's model does not enable control of the growth by increasing temperature or move fish between tanks, which are important aspects for a smolt producer.

Langan and Toftøy (2011) investigated a similar problem as Hæreid (2011). They postulated an alternative growth model to the one presented by Hæreid (2011) where a list is used to represent the fish weights of a batch. This alternative model is similar to the batch growth model presented by Forsberg (1999). They argue that the original growth model may cause an unnaturally high or low growth if the growth rate and the weight class granulations class are small. However, they acknowledge that the original model gives more flexibility in terms of harvesting possibilities compared to the alternative model. This conclusion was also supported by Forsberg (1999). In their implemented model they used the alternative model, which suffers from the same problem as the model by Hæreid (2011) since growth is pre-generated.

Rynning-Tønnesen and Øveraas (2012) refined the stochastic programming model presented by Hæreid (2011), by including the parts of the freshwater facilities and using rolling horizon. They investigated the optimal harvesting and smolt ordering over a 5 years period using a one year rolling horizon. Time series and moment matching were used to create temperature scenarios. They used the growth model by Hæreid (2011) at the seawater facilities and the alternative growth model by Langan and Toftøy (2011) at the freshwater facilities. It was assumed that growth was only dependent on temperature and they used growth data provided by the fish feed producer Skretting to parameterize the model. The model also considered maximum biomass restrictions at the freshwater facilities and restrictions on smolt deployment at sea. The main shortcoming found in the model by Hæreid (2011) and Langan and Toftøy (2011) is still present since the growth is pre-calculated.

The conclusion based on available literature suggests that linear optimization is the best approach. Two possible methods to model growth have been identified. It can either be based on a weight transition matrix or on a weight vector, but none of them can be used directly since they do not include heating options.

2.2 Salmon growth rate modeling

The presented biological optimization models all agree that growth is driven by temperature. However, they assume various relationships between growth and temperature without sufficiently justifying the choice of method.

Elliott and Hurley (1997) studied the effect of temperature on salmon growth in the stages from egg to parr. They investigated the growth rate of salmon from two river populations in Great Britain. Within highly controlled environments, they found that the growth rate and temperature could be modeled with a piecewise linear model, consisting of two pieces, in the range of 6°C to 22°C, with an optimal temperature (breakpoint) around 16°C. Forseth et al. (2001) repeated this experiment with salmon from Norway with similar results. They tested the model against another model, the Ratkowsky model (Ratkowsky et al., 1983), which is often used to model growth in bacterial cultures. Forseth et al. (2001) concluded that the model presented by Elliott and Hurley (1997) gave the best fit to the data.

A limitation with the model is that it is fitted for fish that have different gene pool and diet than fish in freshwater facilities. The feed producer Skretting has developed a model for daily growth rates at freshwater facilities which depend on the fish weight and temperature. The daily growth rate for each weight class is linearly determined by the tank temperature. For a given temperature, the growth rate is also decreasing in fish weight. These characteristics are similar to the ones found by Elliott and Hurley (1997). Since Marine Harvest uses the Skretting growth rate model and it is consistent with the academic literature, it can be deemed as reliable growth rate.

Chapter 3

Model outline

Modeling smolt production precisely requires a vast degree of detail, and the core aspects and business drivers needs to be extracted in order to achieve both a tractable and believable optimization model. Complex biological processes such as growth, diseases and smoltification have to be considered and related to environmental parameters, facility capacities and cost drivers.

This chapter presents an outline of the stochastic programming model which is formulated mathematically in chapter 4. First, the model scope is defined, specifying which parts and aspects of the salmon value chain that are considered. Second, the conceptual model gives an overview of the core components of the optimization model. Third, the assumptions and design choices regarding decisions variables, constraints, objective function, stochasticity and time horizon are presented. Lastly, possible model applications are presented and discussed.

3.1 Model scope

Marine Harvest defined the scope to only consider the freshwater part of the salmon value chain, which includes the life stages prior to deployment in net pens at sea. Seawater growth, slaughtering, processing and distribution will therefore neither be modeled nor further discussed.

The freshwater part of the value chain consists of a network of smolt facilities. As agreed with Marine Harvest the scope is defined to only investigate one facility. By developing a single facility model it will yield a solid platform for testing, feedback sessions and further work, since there is a need for academic work in this field. By designing a small scale model first it will give valuable insights into the nature of the model before scaling it. Further extensions with multiple factories can be emulated by adding new regions which represent additional facilities. The network aspect can easily be included once a formulation for a single factory has been made.

Marine Harvest has requested a model to support tactical decision making. Thus, operational and risk management aspects are not considered. The value of many strategic decisions will depend on their effect on tactical aspects, for instance investing in new tanks or water heating equipment. The model can therefore be used to support some strategic decision making. Possible model applications are discussed in section 3.4.

Most salmon farming companies, including Marine Harvest, have a vertically integrated smolt production. These are almost solely used for internal production, which makes the public market for smolt very limited. The Marine Harvest freshwater facilities are considered cost centers which have to satisfy the orders from the seawater divisions. Therefore, the model objective was defined to be cost minimization.

3.2 Conceptual model

The conceptual model in the table below highlights the core concepts of the optimization model. It is based on the knowledge gained in the previous chapters and forms a foundation for the mathematical formulation in chapter 4.

Minimize	Total expected costs
Subject to	1. Only the freshwater value chain is included 2. Orders from the seawater divisions must be met 3. Temperature driven growth model 4. Regulatory aspects must be included 5. Facility capacities must be respected 6. Relevant stochastic factors are included 7. Planning horizon suitable for tactical decisions

The model objective is to minimize the total expected costs and all of the costs occurring due to normal operations are included. Those without a driver amongst the decision variables are deemed to be fixed costs. The choice of cost drivers is further discussed in section 3.3.2.

The seawater divisions place smolt orders which must be met, but deviations in both order quantity and weight are allowed within certain ranges. At Marine Harvest smolt weight deviations of -20% and +30% and quantity variations of $\pm 15\%$ are allowed (Marine Harvest, 2013a). The latter was not included since a cost minimizing model will always deliver -15% below the order quantity.

A temperature driven growth model is used in the industry and in the academic work discussed in the literature review in section 2. The characteristics of such a model vary greatly. Important design choices are how the temperature is related to growth and how the fish transition between weight classes. The growth model is based on the work of Hæreid (2011) and Forseth et al. (2001). They defined a set of weight classes with temperature dependent transition rates between these classes. Since water can be heated at freshwater facilities some augmentations were necessary. The most significant change was to let temperature act as a decision variable. This included linearizing a non-analytical function using Specially Ordered Sets of Type 2(SOS2) and establishing that growth is linearly related to the temperature.

Regulatory aspects such as maximum fish density and minimum/maximum water flow are included in the model. These aspects ensure fish quality, general health and protect the environment.

Freshwater facilities have tank capacities, flow capacities and water heating characteristics. These aspects are included in the model as constraints. Many uncertain production factors were highlighted in section 1.5. We determined in collaboration with Marine Harvest to prioritize uncertainties in the water intake temperature. Other uncertainty factors such as mortality, feed and oxygen price, feed and oxygen factor, power prices, order demand are modeled using expected values. Section 3.3.4 discusses the rationale behind these choices. Marine Harvest wanted the model to support tactical decision making, meaning that both the planning horizon and time resolution have to be selected appropriately. This is discussed in section 3.3.5.

3.3 Model design and assumptions

The model design and assumptions are closely related, as both are needed to transform the conceptual model into a mathematical formulation. The design choices and assumptions are segmented into five categories which are decision variables, constraints, objective function, stochasticity and planning horizon.

3.3.1 Decision variables

Smolt facility operators are faced by a vast amount of decisions, and including all are not possible. The most important ones have been selected in collaboration with Marine Harvest's freshwater division. These are:

- **Fry deployment**: Fry deployment is used instead of egg deployment since the temperature, water quality and flow during the transition from egg to fry must follow a very strict schedule. This period is not regarded as variable. The egg deployment in the hatchery will occur about two months prior to the fry deployment into the smallest tanks. Also, the capacity of the hatchery is not a bottleneck in the production, which support this modeling choice (Marine Harvest, 2013a).

- **Tank temperature**: The water temperature is one of the most important drivers of salmon growth. It can be controlled by heating the water flow in some or all of the tanks, and is the second largest variable cost. Including this aspect as a decision variable is therefore paramount.

- **Growth**: A linear relationship between tank temperature and growth is assumed, and thus growth becomes a variable. This decision is based on the findings in chapter 2.

- **When/where to move fish**: When fish reach a certain size they will outgrow the current tank and will have to be moved. Juggling the different fish batches in order to utilize the tank capacity is an important aspect of the facility operations.

- **Smolt delivery**: At most one order per month is assumed, where both the target weight and size are predetermined. Only deviations in terms of delivery weight are allowed.
- **Fish culling**: Fish which fall outside of the desired weight range or create a bottleneck in production can be culled to free capacity or avoid unnecessary costs. This is a decision that is actively used to control the biomass in today's production scheme.

Variables denoting fish quantities are continuous rather than integer variables, which is a reasonable assumption given the large scale which fish farms operate on.

Many relevant decisions at a smolt facility have *not* been included as decision variables in the model. The rationale for omitting these are:

- **Fish quality**: Fish quality is an aspect which is hard to quantify. Marine Harvest's standard procedures secure the robustness and overall health of the smolt delivered and uses fish weight as a proxy for fish quality (Marine Harvest, 2013a). We have adopted this standard. Other intangible quality aspects such as general health, genetics, robustness against diseases, saltine tolerance and appearance were not considered since they are not cost drivers.
- **Water quality**: If the water quality is not maintained, the fish can become ill and die or experience halted growth. Ensuring water quality can be viewed upon as a two stage process. First, the intake water has to be of a certain quality before entering the tank. This is not an issue, as discussed in section 1.5. Second, once the water enters the tank the quality has to be sustained. This is affected by the biological processes in the tank. The more fish and the faster it grows, the more waste products it creates. Water quality is therefore not included as a decision variable but implicitly through biomass restrictions.
- **Smoltification**: The most critical factor determining when smoltification occurs is whether or not the parr has reached a minimum weight. After that stage has been reached, the smoltification process can be accelerated or slowed down by varying the illumination in the tank (Hansen, 1998). Manipulating the smoltification does not involve a significant cost and would be very complex to model. This aspect has not been included as a decision variable in the model.
- **Vaccination**: The parr is vaccinated when it weighs 50-70 grams. As discussed in section 1.3, the vaccination cost is the largest variable cost. Since vaccination is required by law, the question is when it will occur rather than if. Two factors ensure that there is limited flexibility regarding the vaccination timing (Marine Harvest, 2013a). First, the vaccination machines can only be used on fish within a specific weight range. Second, the fish needs some time to recuperate after the vaccination before it can smoltify and be deployed at sea. Thus, vaccination is not a decision variable in the model, but the cost is

instead driven by other decision variables. This is discussed in the following section.

3.3.2 Objective function

The total expected costs include all of the six variable cost components present in section 1.3: vaccination, feed, energy, eggs, oxygen and insurance. The variable costs have to be driven by the decision variables determined in the previous section. The drivers behind each cost are assumed to be the following:

- **Feed cost**: It is assumed that the only variable affecting the feeding costs is growth. This is a well established praxis in the salmon farming industry. Both the feed factor ($\frac{\text{kg of feed}}{\text{kg of growth}}$) and feed prices are assumed to be constant.
- **Energy cost**: The only *variable* affecting the energy costs is the tank temperature. The cost is calculated in two steps. First, the energy required to heat the water flow to the desired tank temperature given the intake temperature is determined using basic thermodynamics. Second, the cost of the required energy is calculated using the heat pump efficiency and power prices.
- **Egg cost**: Since fry deployment is used rather than egg deployment, the cost of purchasing and hatching each egg is assigned to each fry deployed.
- **Oxygen cost**: It is assumed that the only variable affecting the oxygen costs is growth. This is a well established praxis in the salmon farming industry. Both the oxygen factor ($\frac{\text{kg of oxygen}}{\text{kg of growth}}$) and oxygen prices are assumed to be constant.
- **Insurance cost**: It is assumed that a fixed fee is paid per fry deployed to insure a fish throughout its life in freshwater. This is a large simplification of real life, since the insurance products are complex and depend on risk profile and fish characteristics. This approximation is deemed accurate enough since the cost is only 0.6% of the total costs.
- **Vaccination cost**: This is the largest cost associated with smolt production. Since the actual vaccination is not a decision variable it is important that another suitable driver for this cost is found. The following three variables can serve this purpose:
  1. *Fry deployment*: Using fry deployment as a driver would overestimate the amount of fish vaccinated. This is because mortality and culling will occur in the time between fry deployment and vaccination. A possible way to correct for this effect is to reduce unit price per vaccination based on the expected amount of mortality and culling.
  2. *Vaccination weight reached*: This method assigns a vaccination cost whenever the fish reach a certain weight (typically $\sim$50 gram). For some smolt facility types, this can be modeled by linear equations without introducing binary variables. This is done by exploiting that the fish have to be moved between regions when it reaches a certain weight. If moving

back again is not permitted, the vaccination cost can be driven by the amount of fish moved *into* a region. This would be a more accurate driver for the vaccination cost, since the fish can be culled both before and after the vaccination driver occurs.

3. *Delivery*: A vaccination cost can also be assigned to every smolt delivered. This method would underestimate the vaccination cost because fish die between vaccination and delivery.

Approximating the vaccination cost using a vaccination weight is regarded as the best method. Since it requires that the production flow throughout the facility follow a specific pattern, it is not applicable to all freshwater facilities. It is therefore not suitable for a general model formulation. Using delivery as a driver effectively makes the variable costs fixed, since the amount delivered is a parameter determined by the orders placed by the seawater divisions. Thus, fry deployment is chosen as the vaccination cost driver in agreement with Marine Harvest.

3.3.3 Constraints

The constraints of an optimization model specify how the decision variables affect each other and how they are bounded by parameters. There are infinitely many couplings in real life, but only the most important ones are considered for the model to remain tractable. The main constraint groups included in the model are:

- **Growth model**: This set of equations specifies what drives the biomass development in each tank. Factors such as initial biomass, fry deployment, culling, movement and growth are related together to ensure consistency. Even though there are many factors which can affect growth, the environmental temperature is by far the most important one as long as the water quality remains within specified ranges (Hansen, 1998). Temperature is therefore the only growth driver in the model. Also, non-negative growth is assumed, which implies that the fish must always become bigger or remain the same size. This assumption is reasonable because it is standard procedure to feed the fish until saturation (Marine Harvest, 2013a).

- **Loss in production**: Disease, desmoltification and weak individuals are common causes of production loss. In the model, these factors have all been aggregated into an empirically determined mortality rate which is a function of the weight the fish only. This simplification was made due to limitations in available data and in agreement with Marine Harvest (2013a).

- **Maximum allowed biomass**: There are limits to how much biomass each fish tank can contain. This limit is defined both by governmental regulations, internal policies and ethics (Ministry of Fisheries and Coastal Affairs, 2008; Marine Harvest, 2013b). The two main limiting factors are the tank volume and the tank water flow, which are both included in the model. If the fish does not have enough space they can hurt each by accident or become aggressive and attack each other. This aspect is especially important in freshwater facilities, since salmon does not develop a pack behavior until after the smoltification (Marine Harvest, 2013a). The water quality is affected by the metabolism in the fish. Waste products such as feces, urine and CO_2 reduce the water quality in the tank and increase the risk of diseases and poor growth. The flow is the best way to keep the water quality in the tank at a desired level.

- **Heating water**: The tank temperature is an important decision variable not only because it can be a significant cost, but also because it is the driver behind fish growth. A set of restrictions are included to capture how heating and freshwater intake temperature affect the tank temperature.

Many relevant aspects at a smolt facility have *not* been included in the model constraints and the following is the rationale for omitting these:

- **Fallowing of tanks**: Between each batch, the tanks should remain empty to allow cleaning and decontamination in order to prevent diseases. This usually takes about a week for an average sized set of tanks, which is less than each time period used in the model (see section 3.3.5).

- **Separating batches**: A common praxis is to separate batches in order to prevent diseases from spreading and ensure traceability. Traceability is important for seawater facilities, as it identifies the causal relationship between fish mortality at sea and smolt facility conditions. Micromanaging the content of each tank is too granular for the scope of the model.

- **Sorting**: Sorting machines are used to separate fish into different weight classes. Sorting requires starving the fish prior to the event, which reduces growth. In order to simplify, sorting has not been included in the constraint. There is an implicit assumption of perfect sorting in the way the growth model works, since fish within a specific weight class can be culled, delivered or moved to another tank.

3.3.4 Stochasticity

The conclusion of the uncertainty analysis in section 1.5 was that the mortality rate and water intake temperature were the most important factors to consider. Empirical data indicate that the mortality rate and temperature are uncorrelated to each other (Marine Harvest, 2013a). This means that a scenario tree containing both variables will increase drastically in size with each scenario introduced. Thus, including both factors will quickly render the problem computationally intractable. We determined in collaboration with Marine Harvest to prioritize uncertainties in the water intake temperature. A unique aspect of this model is that water heating can be used as a recourse action to counter scenarios with low intake temperature. How this is used in different temperature scenarios may therefore provide valuable business insights.

3.3.5 Planning horizon

The planning horizon of the model is at least two years. This period is chosen for two reasons. First, because deliveries are planned up to 18 months ahead of time (Marine Harvest, 2013a). Second, consistency between the initial biomass and end of horizon biomass can be enforced realistically, and thus avoids the problem of valuing the fish inventory. The length of each time period $t \in \mathcal{T}$ is one month because the model scope is tactical decision making. The time index t denotes the first day of the month and it is assumed that all actions related to decision variables occur on the first day. The model formulation allows for a rolling horizon, which effectively allows the operator to re-run the model every time new information becomes available. Due to the computational power available we have chosen to *exclude* this in this thesis.

3.4 Possible model applications

In a supply chain, decision types can be grouped into three classes: strategic, tactical and operational (Schmidt and Wilhelm, 2000). **Strategic decisions** usually involve a planning period of two years or more, and can be related to decisions such as where to locate facilities, capacity planning or the choice of production technology used. Examples of **tactical decisions** are production levels at facilities, inventory levels and batch sizes. The planning horizon for tactical decisions usually spans 6-24 months. **Operational decisions** have a short planning horizon, and examples can be coordinating network logistics to assure on-time delivery and staff scheduling. Operational decisions are outside of the model scope.

3.4.1 Tactical decisions

The model can be applied as a tool when making the following tactical decisions:

- **Improve current production plans**: The optimal decision variable values could provide key insights into how a smolt production facility should be operated tactically in order to minimize costs. This is therefore *the core* application of the model.

- **Internal pricing scheme**: By varying the order characteristics and logging the costs of producing each product type, the model could be used to generate an internal pricing scheme for smolt. Such a "menu" could be used to properly charge the marginal cost of production from the seawater divisions, and thus better align the incentives in the value chain.

- **Sensitivity analysis**: By varying order characteristics such as delivery time, batch size, target weight etc. the model can be used to determine how sensitive the cost and feasibility of the order is. This allows the model to act as a naïve risk management tool.

3.4.2 Strategic decisions

The value of strategic decisions is driven by their impact on the value of future tactical decisions. For instance, investing in a new tank or heating equipment will impact the future cost efficiency of operating a facility. Understanding this impact is therefore at the core when making the strategic decisions. The model can therefore easily be augmented to go beyond its intended tactical use:

- **Investment analysis**: By changing the facility parameterization or including boolean variables for investing in new equipment, the optimization model can be used to investigate how these new investments will affect the economics of a smolt production facility. This could be used to make decisions regarding how to invest. Relevant investments could be increased heating capacity, new tanks, increased water capacity or optimal design of new facilities.
- **Product specialization**: A smolt product can be defined as a combination of smolt size, delivery time and batch size. An opportunity for product specialization arises when considering a network of smolt facilities, and the supply chain performance could be increased. For instance, some facilities might specialize in delivering large smolts while other specialize in small smolts.
- **Introducing new products**: A relevant example is to determine whether smolt facilities should aim to produce larger smolts, e.g. 500 grams. If new products are to be introduced, the model could also give indications on how these could be produced most cost efficiently.

Chapter 4

Model formulation

A stochastic programming model is formulated in this chapter based on the model outline in chapter 3. The indexes, sets, parameters and variables used in the model are presented in section 4.1, and the objective function is defined in section 4.2. Lastly, the constraints are defined and discussed in section 4.3.

4.1 Sets, parameters and variables

There is a connection in the model formulation between the type of letter and the function of that letter. The following conventions are used throughout this chapter:

- Lower case Latin letters are used to denote variables or indexes.
- Upper case Latin letters are used to denote constants.
- Greek letters are used to denote stochastic parameters.
- Calligraphic letters are used to denote sets.

Accent marks, such as "ˆ", are used on indexes to denote which index describe the *end state* when the respective variables describes transitions. E.g. $x^t_{\hat{f}firs}$ denote the transition rate in percentage from fish class f to fish class $\hat{f}$.

4.1.1 Sets and indexes

Set	**Description**	**Indexes**
$\mathcal{C}$	Set of all smolt classes, $\mathcal{C} \subset \mathcal{F}$.	c, k
$\mathcal{E}$	Set of event nodes, which are located at the splits in the scenario tree.	e
$\mathcal{F}$	Set of all fish classes.	f, $\hat{f}$
$\mathcal{I}$	Set of all fish tanks.	i, $\hat{i}$
$\mathcal{O}$	Set of fish number classes.	o
$\mathcal{R}$	Set of all regions.	r, $\hat{r}$
$\mathcal{S}$	Set of all scenarios.	s
$\mathcal{S}(e)$	Set of scenarios passing through the event node e.	s
$\mathcal{T}$	Set of all time periods.	t
$\mathcal{T}(e)$	Set of time periods prior to the event node e.	t
$\mathcal{X}_{ts}$	Denotes all variables that are dependent on time t and scenario s.	-

4.1.2 Decision variables

Variable	Description
b^t_{irs}	Freshwater intake flow in tank i at region r in scenario s at time t.
$d^{t\hat{i}\hat{r}}_{firs}$	Number of fish in class f moved from tank i at region r to tank $\hat{i}$ at region $\hat{r}$ in scenario s at time t.
e^t_{firs}	Number of fish in class f destroyed in tank i at region r in scenario s at time t.
g^t_{firs}	Growth in grams per time period of fish in class f in tank i at region r in scenario s at time t.
$g^{t,\ daily}_{firs}$	Growth in grams per day of fish in class f in tank i at region r in scenario s at time t. Used when calculating the required water flow in each tank.
h^t_{cirs}	Number of smolts delivered from smolt class c in tank i at region r in scenario s at time t.
$j^t_{\hat{f}firs}$	Number of fish which transition from class f to $\hat{f}$ in tank i at region r in scenario s at time t.
m^t_{irs}	Regulatory minimum flow ($\frac{volume}{time}$) per kg salmon in tank i at region r in scenario s at time t.
n^t_{firs}	Number of fish in class f in tank i at region r in scenario s at time t.
q^{t+}_{irs}	Ratio of heated water to intake water in tank i at region r in scenario s at time t.
q^{t-}_{irs}	Ratio of cooled water to intake water in tank i at region r in scenario s at time t.
v^t_{irs}	Temperature in tank i at region r in scenario s at time t.
w^{t+}_{irs}	Flow of heated water in tank i at region r in scenario s at time t.
w^{t-}_{irs}	Flow of cooled water in tank i at region r in scenario s at time t.
$x^t_{\hat{f}firs}$	Transition rate in percentage from fish class f to fish class $\hat{f}$.
y^t_{firs}	Number of fry added at time t in tank i at region r in scenario s.
z	Total expected costs.
z_{energy}	Total expected energy costs related to heating of water.
z_{fry}	Total expected costs related to fry deployment. It consists of egg cost, hatching cost, vaccination cost and insurance cost.
z_{growth}	Total expected growth costs. It consists of feed cost and oxygen cost.

4.1.3 Deterministic parameters

Parameter	Description
$A_{rf}^{\text{fish allowed}}$	Boolean value denoting whether fish in weight class f is allowed at region r. True suggest that fish weight class f is allowed at region r.
$A_{r\hat{r}}^{\text{movement}}$	Boolean value denoting whether relocation of fish from region r to region $\hat{r}$ is permitted.
A_{r}^{heating}	Boolean value denoting whether water heating is possible at region r.
A_f	Interception parameter in the growth regression model for fish weight class f.
B_f	Slope parameter in the growth regression model for fish weight class f.
C^{EVI}	Cost of one unit of fertilized eggs, insurance and vaccination.
C^{feed}	Cost of one kilogram of fish feed.
C^{oxygen}	Cost of one kilogram of oxygen.
E_{irt}	Energy cost of changing the temperature in the tank flow ($\frac{volume}{time}$) by 1°C in tank i at region r at time t.
F^{feed}	Feed conversion rate, a ratio for the amount of feed needed to produce one kilogram of salmon growth.
F^{oxygen}	Oxygen conversion rate, a ratio for amount of oxygen needed to produce one kilogram of salmon growth.
G_r^{max}	Temperature of water when warmed up to the maximum temperature at region r in scenario s.
G_r^{min}	Temperature of water when cooled down to the minimum temperature at region r in scenario s.
H_c^t	Ordered amount of smolt in smolt class c to be delivered at time t.
I_{fir}	Initial number of fish in weight class f in tank i at region r.
L_{tir}^{+}	Maximum feasible value of q_{irs}^{t+}
L_{tir}^{-}	Maximum feasible value of q_{irs}^{t-}
L_{ir}^{size}	Size of tank i at region r measured in m^3.
N_o	Quantity of fish in number class $o \in \mathcal{O}$. A discretization of the space spanned by n_{firs}^t. Used when linearizing the growth model.
P_s	Probability of scenario $s \in \mathcal{S}$
RD_{ir}	Maximum allowed fish density ($\frac{kg}{m^3}$) in tank i at region r.
S_f^{fish}	Survival rate of fish weight class f per time period.
S^{egg}	Survival rate in the transition from egg to fry.
TL_r^{+}	Maximum amount of flow ($\frac{volume}{time}$) that can be heated at region r.
TL_r^{-}	Maximum amount of flow ($\frac{volume}{time}$) that can be cooled at region r.
U_c^{upper}	Largest allowed smolt class that can be used to satisfy an order of smolt class c.

Continued on next page

Parameter	Description
U_c^{lower}	Smallest allowed smolt class that can be used to satisfy an order of smolt class c.
$VCO2_{ir}$	Reduction in the flow requirement, κ_{irs}^t, in i at region r due to having CO_2-ventilators installed.
V_f	Weight of fish in class f.
W_{ir}	Water flow in tank i at region r.
Z_{fixed}	Total fixed costs. It consists of all facility costs which are unaffected by decision variables.

4.1.4 Stochastic parameters

Parameter	Description
$\iota_{ts}^{hatching}$	Energy cost per fry related to hatching the eggs in the time periods prior to fry deployment at time t in scenario s.
κ_{irs}^t	Minimum required flow ($\frac{volume}{time}$) per kg salmon in tank i at region r in scenario s at time t.
τ_s^t	Freshwater intake temperature in scenario s at time t
$\chi_{\hat{f}fs}^t$	Transition rate from fish class f into fish class $\hat{f}$ in scenario s at time t in. Used in regions where water heating is not available.

4.1.5 Linearization variables

This section defines the variables used to approximate a non-linear function in the model formulation and gives an introduction to the technique used.

Williams (1999) showed that a non-linear function of two variables $z = h(x, y)$ can be approximated using special order sets of type 2 (SOS2). *"An SOS2 is a set of variables within which at most two can be non-zero. The two variables must be adjacent in the ordering given to the set."* (Williams, 1999, page 165) The approximation of $z = h(x, y)$ is split into seven steps:

1. Discretize the x, y and z dimensions: Let $\mathcal{I}$ and $\mathcal{J}$ denote the set of breakpoint indexes of the x and y dimension respectively. x is then discretized using $\{X_i | i \in \mathcal{I}\}$, y is discretized using $\{Y_j | j \in \mathcal{J}\}$ and z is discretized using $\{Z_{i,j} = h(X_i, Y_j) | i \in \mathcal{I} \wedge j \in \mathcal{J}\}$. This results in a piecewise linear surface which approximates z. An illustration of the discretization is shown in figure 4.1.

2. Assign the weighting variable λ_{ij} to each break point $(X_i,\ Y_j)$ in the (x,y)-plane, as illustrated in figure 4.1.

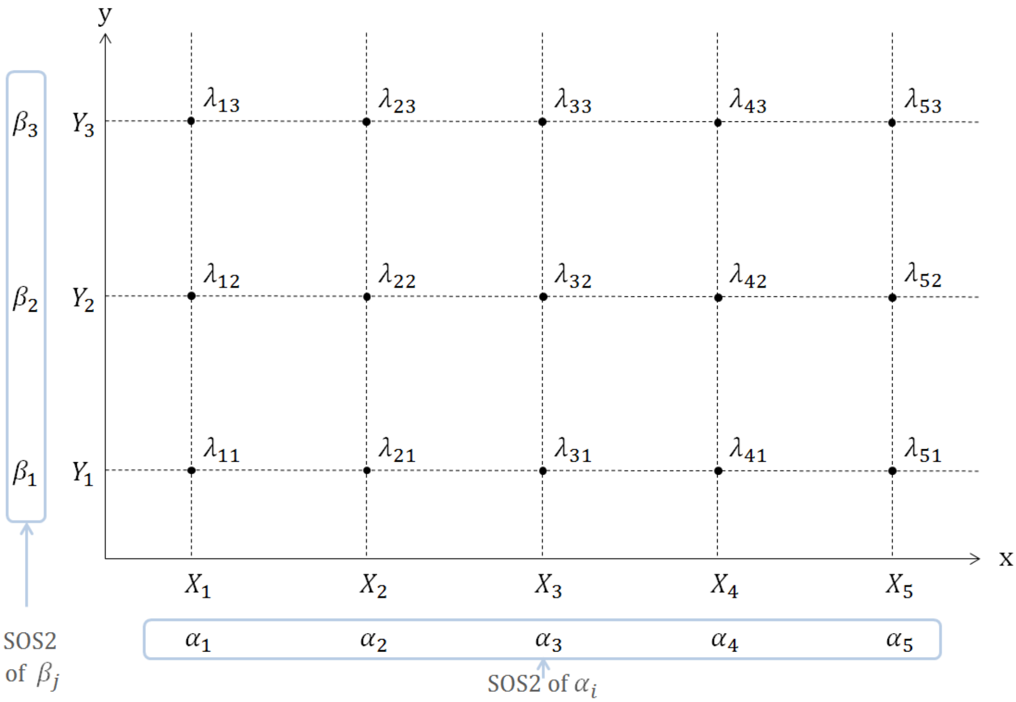

Figure 4.1: The discretized (x, y)-plane with weighting variables λ and the variables of the two SOS2.

3. Express x as a linear combination of the X_i's using λ_{ij}:

$$x = \sum_{\substack{i \in \mathcal{I} \\ j \in \mathcal{J}}} X_i \lambda_{ij} \tag{4.1}$$

4. Express y as a linear combination of the Y_j's using λ_{ij}:

$$y = \sum_{\substack{i \in \mathcal{I} \\ j \in \mathcal{J}}} Y_j \lambda_{ij} \tag{4.2}$$

5. Ensure that the linear combinations are valid by normalizing the $\lambda_{i,j}$'s:

$$1 = \sum_{\substack{i \in \mathcal{I} \\ j \in \mathcal{J}}} \lambda_{ij} \tag{4.3}$$

6. The surface z can be approximated using the $\lambda_{i,j}$'s as:

$$z = \sum_{\substack{i \in \mathcal{I} \\ j \in \mathcal{J}}} Z_{ij} \lambda_{ij} \tag{4.4}$$

7. In addition, it is necessary to impose that "*at most four neighboring λ_{ij}'s are*

non-zero" (Williams, 1999, page 167). This is done in order to ensure that the approximated values of z fits the real surface $h(x, y)$ as tightly as possible. This condition can be imposed by using SOS2 in the following way:

$$
\begin{aligned}
\alpha_i &= \sum_{j \in \mathcal{J}} \lambda_{ij} \quad \text{where } \{\alpha_i | i \in \mathcal{I}\} \text{ is a SOS2} \\
\beta_j &= \sum_{i \in \mathcal{I}} \lambda_{ij} \quad \text{where } \{\beta_j | j \in \mathcal{J}\} \text{ is a SOS2}
\end{aligned}
\tag{4.5}
$$

Their relation to the grid and weighting variables is illustrated in figure 4.1.

An issue with this approximation technique is that *three* equations are used to define a linear combination of *four* variables. Thus, there is one degree of freedom. An optimization model will exploit this in order to get the best objective function value. Williams (1999) proposes that this issue can be addressed by introducing a third SOS2 related to the diagonal of the discretized (x, y)-plane. However, this third SOS2 has not been included in the model formulation. The reason for this is at it reduced the overall model quality. A discussion of this decision is included in appendix A.

Variable	**Description**
$\alpha_{firs}^{t\hat{f}}$	For each combination of starting fish class f, tank i, region r, scenario s and time t, every $\alpha_{firs}^{t\hat{f}}$ corresponding to a destination fish class $\hat{f}$ is part of a SOS2. Mathematically: $\{\alpha_{firs}^{t\hat{f}} \| \hat{f} \in \mathcal{F}\}$ is a SOS2 $\forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}$
β_{firs}^{to}	For each combination of starting fish class f, tank i, region r, scenario s and time t, every β_{firs}^{to} corresponding to a number class o is part of a SOS2. Mathematically: $\{\beta_{firs}^{to} \| o \in \mathcal{O}\}$ is a SOS2 $\forall f \in \mathcal{F}, t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}$
$\lambda_{firs}^{t\hat{f}o}$	Weighting variable used in the approximated growth model where its feasible values are determined by the corresponding $\alpha_{firs}^{t\hat{f}}$ and β_{firs}^{to}.

4.2 Objective function

The model's objective is to minimize the expected total production costs, z, given in equation (4.6).

$$Minimize\ z = z_{energy} + z_{fry} + z_{growth} + Z_{fixed} \tag{4.6}$$

The total costs have been segmented into four categories based on their underlying cost drivers: energy cost (z_{energy}), fry deployment cost (z_{fry}), growth cost (z_{growth}) and fixed cost (Z_{fixed}).

4.2.1 Energy cost

The total expected energy cost, z_{energy}, is related to the heating and cooling of water in the fish tanks. Its underlying cost drivers are the power price, the water flow in the tanks and the temperature difference between the water intake temperature and the desired temperature in the tanks. Other energy cost such as lighting and heating outside the production tank are not driven by the production plans and are therefore assigned to fixed costs. The energy cost function is given in equation (4.7).

$$z_{energy} = \sum_{s \in \mathcal{S}} P_s \left(\sum_{\substack{i \in \mathcal{I} \\ r \in \mathcal{R} \\ t \in \mathcal{T}}} E_{irt} \Big((G_r^{max} - \tau_s^t) W_{ir} q_{irs}^{t+} + (\tau_s^t - G_r^{min}) W_{ir} q_{irs}^{t-} \Big) \right) \tag{4.7}$$

This approach is based on basic thermodynamics as described by for instance Tipler and Mosca (2008). The power price, the efficiency of the heating machine and the heat capacity of the water are merged into one parameter, E_{irt}, with denomination $\frac{\text{cost * time}}{\Delta\text{temperature * volume}}$. Thus, when it is multiplied with the flow, $W_{ir} q_{irs}^{t+}$, with denomination $\frac{\text{volume}}{\text{time}}$ and with the temperature difference, $G_r^{max} - \tau_s^t$, it yields the cost of energy used to heat the water flow. The cost of cooling the water is given by the temperature difference, $\tau_s^t - G_r^{min}$, and water flow, $W_{ir} q_{irs}^{t-}$. Situations where the temperature difference might be negative are discussed in the constraint section. P_s is the probability of scenario s.

The formulation assumes a perfectly insulated system, thus excess heat to/from the surroundings is not accounted for. There is no dependence between tank temperature in various time periods because the water is replaced 200-600 times per month (Marine Harvest, 2013a).

4.2.2 Fry deployment cost

The total expected fry deployment cost, z_{fry}, is given in equation (4.8).

$$z_{fry} = \sum_{s \in \mathcal{S}} P_s \Big(\sum_{\substack{i \in \mathcal{I} \\ r \in \mathcal{R} \\ t \in \mathcal{T} \\ f \in \mathcal{F}}} y^t_{firs} (C^{EVI} + \iota^{hatching}_{ts}) \Big) \tag{4.8}$$

This cost is driven by the number of fry deployed, y^t_{firs}, in fish class f in tank i at region r in scenario s at time t. The constant C^{EVI} is the unit cost related to buying eggs, and vaccinating and insuring the fry as they grow. The stochastic parameter $\iota^{hatching}_{ts}$ represents the hatchery energy cost of deploying fry at time t in scenario s. P_s is the probability of scenario s.

4.2.3 Growth cost

The fish needs both oxygen and feed to grow. The feeding factor, F^{feed}, is a well established metric denoting how much feed is necessary to sustain a growth of one kilogram. The oxygen factor, F^{oxygen}, is the equivalent for oxygen demand. Oxygen have to be added to the water in the tanks, since the natural occurrences are not sufficient to sustain the high fish density and growth present. Both these costs are therefore driven by the total growth in the facility.

The total expected growth costs, z_{growth}, is calculated by multiplying the cost per growth by the total growth and is defined in equation (4.9).

$$z_{growth} = \sum_{s \in \mathcal{S}} P_s \Bigg((C^{feed} F^{feed} + C^{oxygen} F^{oxygen}) \\ \Bigg(\sum_{\substack{i \in \mathcal{I} \\ r \in \mathcal{R} \\ f \in \mathcal{F}}} V_f \Big(n^T_{firs} + \sum_{t \in \mathcal{T}} \big((1 - S^{fish}_f) n^t_{firs} + e^t_{firs} + h^t_{firs} \big) \Big) - \sum_{\substack{i \in \mathcal{I} \\ r \in \mathcal{R} \\ f \in \mathcal{F}}} \Big(V_f I_{fir} + \sum_{t \in \mathcal{T}} V_f y^t_{firs} \Big) \Bigg) \Bigg) \tag{4.9}$$

Non-negative growth is assumed, implying that the total growth is equal to the total biomass gain in the system. The total biomass gain in the system is the sum of the total biomass leaving the system and the end of horizon inventory, corrected for the initial inventory and the start weight of the fry.

There are three ways the fish can leave the system. It can either die of natural causes, $(1 - S^{fish}_f) n^t_{firs}$, be destroyed, e^t_{firs}, or it can be delivered, h^t_{firs}. The end of horizon inventory in fish class f is n^T_{firs}. The total biomass gain in each class f is the sum of all these amounts for all time periods t multiplied by the fish weight V_f. Summing over all tanks i and regions r yields the total biomass gain in the system.

There are two ways the fish can enter the system, it can either be in the initial inventory, I_{fir}, or it can be delivered as fry, y^t_{firs}. Likewise, the biomass leaving the system, the entering amount of fish is also multiplied with fish weight V_f for each weight class f.

The cost per biomass gain is calculated using the cost per kg feed and oxygen, C^{oxygen} and C^{feed}, and the feed and oxygen factors, F^{feed} and F^{oxygen}.

An alternative way to calculate the total biomass gain in the system would be to sum the number of fish in each fish class f, n^t_{firs}, multiplied by the growth in that class, g^t_{firs}. However, this formulation would be non-linear and therefore not used in this model.

4.2.4 Fixed cost

The fixed cost, Z_{fixed}, is a constant which represents costs not impacted by any decision variables. It is included in the model to make it easier to compare the model's total expected cost, z, to historical values. Including a fixed cost will neither affect the optimal solution nor solution time of the model.

4.3 Constraints

The production of smolt is limited by growth behavior of fish, technical equipment and business logic. This sections will present the constraints outlined in chapter 3. In some cases the of the constraints are nonlinear. However, only the linear approximations are included in the final model.

4.3.1 Growth model

The growth model is one of the core components of the optimization model, and determine how the fish in the system grows and transitions between weight classes. It is based on the work of Hæreid (2011), but is augmented to allowing water heating. This implies that growth is a decision variable rather than a parameter.

The growth model is presented in two parts. First, a non-linear formulation is defined. Second, the growth model is linearized with SOS2 using an approach similar to the one described in section 4.1.5. This formulation yields approximately the same behavior as the non-linear formulation, but is more computationally tractable and is therefore used in the optimization model.

4.3.1.1 Non-linear growth model

The non-linear growth model consists of three components. First, the *balance equations* which ensures consistency between variables that affect the amount of fish in each weight class. Second, the *transition rate equations* allocates fish to a new weight class as it grows. Third, the equations which specify the *relationship between tank temperature and growth*.

Balance equation Equation (4.10) sets the amount of fish $n^t_{\hat{f}irs}$ in class $\hat{f}$ in tank i at region r at time t in scenario s.

$$n^t_{\hat{f}irs} = \sum_{f \leq \hat{f}} (S^{fish}_f x^{t-1}_{\hat{f}firs} n^{t-1}_{firs}) - e^t_{\hat{f}irs} - h^t_{\hat{f}irs} + S^{egg} y^t_{\hat{f}irs} + \sum_{\substack{\hat{i} \in \mathcal{I}, \\ \hat{r} \in \mathcal{R}}} (d^{tir}_{\hat{f}\hat{i}\hat{r}s} - d^{t\hat{i}\hat{r}}_{\hat{f}irs}) \tag{4.10}$$
$$\forall \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}$$

This amount is determined by five factors: First, the amount of fish in class f, n^{t-1}_{firs}, in the previous time period $t-1$ that has grown to class $\hat{f}$. This amount is determined by the transition rate $x^{t-1}_{\hat{f}firs}$ and survival rate S^{fish}_f. Second, the amount of fish which is destroyed, $e^t_{\hat{f}irs}$. Third, the amount of fish delivered to fulfill an order, $h^t_{\hat{f}irs}$. Fourth, the amount of fry deployed time t in fish class f, $y^t_{\hat{f}irs}$. This amount is affected by the survival rate from egg to fry, S^{egg}. Fifth, the net amount of fish moved to tank $\hat{i}$ at region $\hat{r}$ from tank i at region r, $(d^{tir}_{\hat{f}\hat{i}\hat{r}s} - d^{t\hat{i}\hat{r}}_{\hat{f}irs})$.

Transition rate The transition rate, $x^t_{\hat{f}firs}$, is determined uniquely by the current weight, V_f, and the growth in the period, g^t_{firs}. Since the fish weight is discretized into weight classes, the new fish weight, $V_f + g^t_{firs}$, may not necessarily fit directly into a new weight class. Therefore, the amount of fish grown will have to be allocated to the two nearest weight classes. These two nearest weight classes, $\underline{f}$ and $\overline{f}$, are the two adjacent classes which satisfy $V_{\underline{f}} \leq V_f + g^t_{firs} \leq V_{\overline{f}}$. Figure 4.2 illustrates this allocation.

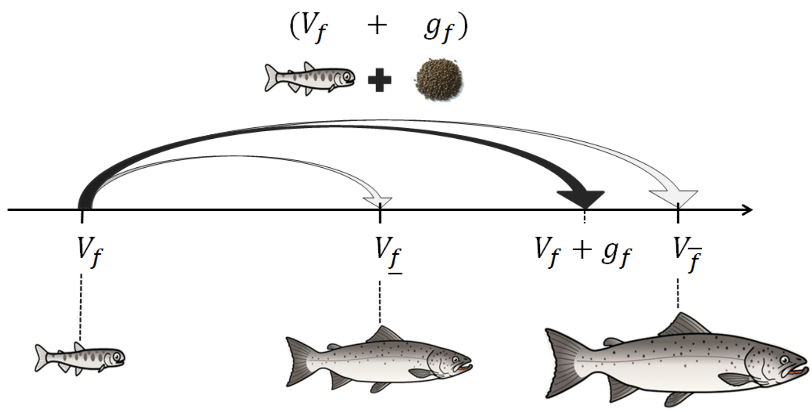

Figure 4.2: Shows how the fish transition between weight classes.

The percentage split between $\underline{f}$ and $\overline{f}$ is determined by equation (4.11) and (4.12) respectively.

$$x^t_{\underline{f}firs} = \frac{V_{\overline{f}} - (V_f + g^t_{firs})}{V_{\overline{f}} - V_{\underline{f}}} \tag{4.11}$$

$$x^t_{\overline{f}firs} = \frac{(V_f + g^t_{firs}) - V_{\underline{f}}}{V_{\overline{f}} - V_{\underline{f}}} \tag{4.12}$$

The transition rate, $x^t_{\hat{f}firs}$, from f to class $\hat{f}$ in tank i at region r in scenario s at time t is defined in equation (4.13). Variables not satisfying the logical criteria are set equal to zero.

$$x^t_{\hat{f}firs} = \begin{cases} \frac{(V_f + g^t_{firs}) - V_{\underline{f}}}{V_{\overline{f}} - V_{\underline{f}}} & \text{if } (\hat{f} = \overline{f}) \wedge (V_{\overline{f}} \geq V_f + g^t_{firs} \geq V_{\underline{f}}) \wedge (*) \\ \\ \frac{V_{\overline{f}} - (V_f + g^t_{firs})}{V_{\overline{f}} - V_{\underline{f}}} & \text{if } (V_{\overline{f}} \geq V_f + g^t_{firs} \geq V_{\underline{f}}) \wedge (\underline{f} = \hat{f}) \wedge (*) \\ \\ 0 & \text{else} \end{cases}$$
$$\forall f, \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}$$
$$*\overline{f} \text{ and } \underline{f} \text{ are adjacent weight classes} \tag{4.13}$$

Relationship between growth and temperature The growth of the fish, g^t_{firs}, is assumed to be determined by a linear relationship with the tank water temperature, v^t_{irs}. This relationship is specified in equation (4.14), where A_f and B_f are constants.

$$g^t_{firs} = A_f + B_f v^t_{irs}$$
$$\forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.14}$$

4.3.1.2 Linearization of the growth model

The current formulation of the growth model is non-linear due to two factors. First, the balance equation (4.10) contains a product of two variables, $x^t_{\hat{f}firs} n^t_{firs}$. This product denotes the number of fish which transition to weight class $\hat{f}$ from class f in tank i at region r in scenario s at time t. Second, the transition rate $x^t_{\hat{f}firs}$ is set in equation (4.13) on the basis of the growth, g^t_{firs}, using deductive tests (if-then tests). Thus, the transition rate is a function $x_{\hat{f}}(g^t_{firs})$ which is neither linear nor analytic. That is, the function is not continuously differentiable (Gonchar, A.A. and Shabat, B.V., 2011).

A general function $h(x, y)$ can be linearized using SOS2, as explained in section 4.1.5. However, the function to be linearized have to be defined such that it includes both the transition rate function and the product in the balance equation. This function is defined in equation (4.15).

$$
\begin{gathered}
h_{\hat{f}}(V_f + g^t_{firs}, n^t_{firs}) = x^t_{\hat{f}firs} n^t_{firs} \\
\forall \hat{f} \in \mathcal{F}, f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}
\end{gathered}
\tag{4.15}
$$

$h_{\hat{f}}(V_f + g^t_{firs}, n^t_{firs})$ denotes the number of fish which transition from weight class f to $\hat{f}$. This function is uniquely determined by the initial amount, n^t_{firs}, in class f and the after-growth weight, $V_f + g^t_{firs}$. The after-growth weight is used as a functional argument because it is desirable to use fish weights rather than the growth. The reason for this becomes apparent in the linearization.

Each function $h_{\hat{f}}$ can now be linearized using the approach described in section 4.1.5. However, there is a characteristic of this function which provides an opportunity to use fewer SOS2.

There are $|\mathcal{F}|$ functions for each $(V_f + g^t_{firs}, n^t_{firs})$-tuple. This is because there is an amount transitioned function, $h_{\hat{f}}$, for each destination class $\hat{f} \in \mathcal{F} = \{1, 2, ..., |\mathcal{F}|\}$. There is a coupling between the functions since they have to sum up to the initial amount and only two neighboring functions in terms of $\hat{f}$ can be non-zero. This knowledge can be exploited when linearizing $h_{\hat{f}}$.

For the sake of notational brevity, assume that f, i, r, t and s are all fixed. That is, we are considering the function $h_{\hat{f}}(V + g, n)$. It denotes how many fish should transit to a target weight class $\hat{f}$ given an initial amount n and after-growth weight $V + g$. This function can be linearized in the following seven step process:

1. Discretize the dimensions spanned by the variables $V + g$, n and $h_{\hat{f}}(V + g,\ n)$. When discretizing the weight dimension spanned by the after-growth weight $V + g$, it is convenient to use the fish weight $V_{\hat{f}}$ of the destination classes $\hat{f} \in \mathcal{F} = \{1, 2, ..., |\mathcal{F}|\}$. The amount of fish dimension spanned by n_f is discretized by N_o where $o \in \mathcal{O} = \{1, 2, ..., |\mathcal{O}|\}$. This means that a set of grid points $(V_{\hat{f}},\ N_o)$ is defined, as illustrated in figure 4.3. The space spanned by $h_{\hat{f}}(V + g,\ n)$ is discretized by $h_{\hat{f}}(V_{\hat{f}},\ N_o)$. This results in a piecewise linear surface which approximates $h_{\hat{f}}(V + g,\ n)$.
2. Assign the weighting variable $\lambda^{\hat{f}o}$ to each break point $(V_{\hat{f}},\ N_o)$ in the $(V+g, n)$-plane. An illustration of this is shown in figure 4.3.

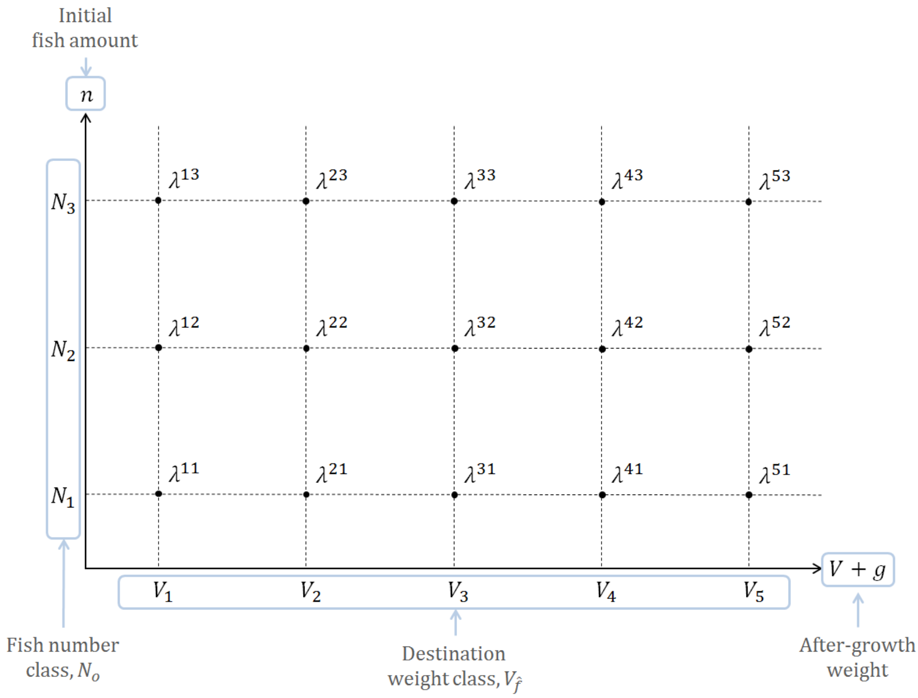

Figure 4.3: The discretized $(V + g, n)$-plane with weighting variables $\lambda^{\hat{f}o}$ assigned to each breakpoint $(V_{\hat{f}},\ N_o)$.

3. Express $V + g$ as a linear combination of the destination fish class weights, $V_{\hat{f}}$, using $\lambda^{\hat{f}o}$:

$$V + g = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F}}} \lambda^{\hat{f}o} V_{\hat{f}} \tag{4.16}$$

4. Express n as a linear combination of the number class quantities, N_o, using $\lambda^{\hat{f}o}$:

$$n = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F}}} \lambda^{\hat{f}o} N_o \tag{4.17}$$

5. Ensure that the linear combinations are valid by normalizing the sum of all the $\lambda^{\hat{f}o}$:

$$1 = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F}}} \lambda^{\hat{f}o} \tag{4.18}$$

6. The surface $h_{\hat{f}}(V + g, n)$ can be approximated using the $\lambda^{\hat{f}o}$'s, but differently than described in section 4.1.5. The $h_{\hat{f}}(V + g, n)$ surface is discretized as $h_{\hat{f}}(V_{\hat{f}}, N_o)$. However, when $V + g = V_{\hat{f}}$ the function h will allocate all of the fish, N_o, to the weight class $\hat{f}$ since the new weight fits the target class perfectly. Thus, the transition rate $x_{\hat{f}}(g) = 1$ and $h_{\hat{f}}(V_{\hat{f}}, N_o) = N_o$.

The $h_{\hat{f}}$ functions are coupled because they have to sum up to the initial amount, and only two neighboring functions in terms of $\hat{f}$ can be non-zero. Using this knowledge, we can infer that if all of the $\lambda^{\hat{f}o}$'s related to a given end weight class $\hat{f}$ are summed, this yields the fraction of the total number of fish starting in class f which will transition into the class $\hat{f}$. These insights mean that $h_{\hat{f}}(V+g,n)$ can be approximated as:

$$h_{\hat{f}}(V+g,n) = \sum_{o\in\mathcal{O}} \lambda^{\hat{f}o} N_o \qquad \hat{f} \in \mathcal{F}, \tag{4.19}$$

7. In addition, it is necessary to impose that at most four neighboring $\lambda^{\hat{f}o}$'s are non-zero. This condition can be imposed by using SOS2 in the following way:

$$\begin{aligned} \alpha^{\hat{f}} &= \sum_{o\in\mathcal{O}} \lambda^{\hat{f}o} \qquad \text{where } \{\alpha^{\hat{f}} | \hat{f} \in \mathcal{F}\} \text{ is a SOS2} \\ \beta^{o} &= \sum_{\hat{f}\in\mathcal{F}} \lambda^{\hat{f}o} \qquad \text{where } \{\beta^{o} | o \in \mathcal{O}\} \text{ is a SOS2} \end{aligned} \tag{4.20}$$

How these two sets are related to the grid and weighting variables is illustrated in figure 4.4.

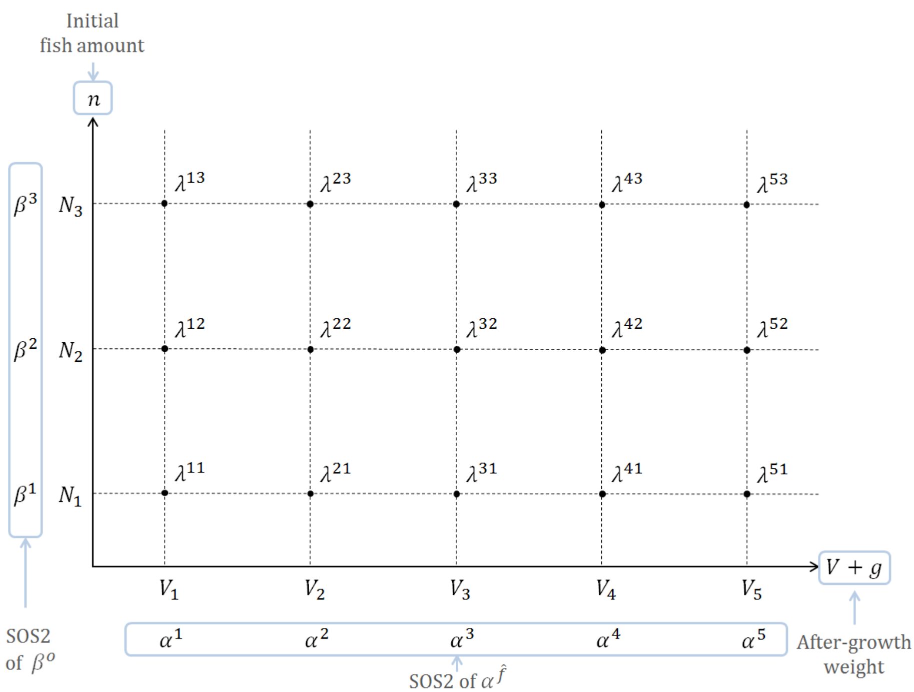

Figure 4.4: The discretized $(V+g,n)$-plane with an $\alpha^{\hat{f}}$ assigned to each breakpoint $V_{\hat{f}}$ and a β^o to each breakpoint N_o.

Equations (4.16)-(4.20) illustrate the core of the growth model linearization. However, three aspects are included to this equation set before they are used in the model formulation. These are:

- The equations will be defined for all combinations of start fish classes f, tanks i, regions r, scenarios s and points in time t.

- Let the variable $j^t_{\hat{f}firs}$ denote the amount of fish which transition from class f to $\hat{f}$. Then we know that:

$$\begin{gathered} j^t_{\hat{f}firs} = h_{\hat{f}}(V_f + g^t_{firs}, n^t_{firs}) = x^t_{\hat{f}firs} n^t_{firs} \\ \forall \hat{f} \in \mathcal{F}, f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{gathered} \tag{4.21}$$

  This variable is introduced to avoid functional notation in the model.

- The range of the summations of destination weight classes has been changed from $\hat{f} \in \mathcal{F}$ to $\hat{f} \in \mathcal{F} | \hat{f} \geq f$. This is done to reduce the number of elements in the SOS2. This is valid because non-negative growth is assumed.

Taking these three aspects into account, we get the linearized growth model as equations (4.22) - (4.27). Equation (4.22) and (4.23) sets $\lambda^{t\hat{f}o}_{firs}$ according to the weight after growth $V_f + g^t_{firs}$ and the initial amount of fish n^t_{firs}. Equation (4.24) sets the amount of fish, $j^t_{\hat{f}firs}$, which transition from f to $\hat{f}$ according to $\lambda^{t\hat{f}o}_{firs}$.

$$\begin{gathered} V_f + g^t_{firs} = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} V_{\hat{f}} \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{gathered} \tag{4.22}$$

$$\begin{gathered} n^t_{firs} = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} N_o \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{gathered} \tag{4.23}$$

$$\begin{gathered} j^t_{\hat{f}firs} = \sum_{o \in \mathcal{O}} \lambda^{t\hat{f}o}_{firs} N_o \\ \forall \hat{f}, f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{gathered} \tag{4.24}$$

Equation (4.25) normalizes the grid point weighting variables $\lambda^{t\hat{f}o}_{firs}$. Equations (4.26) - (4.27) ensures that only neighboring weighting variables, $\lambda^{t\hat{f}o}_{firs}$, are non-zero.

$$\sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} = 1 \qquad \forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.25}$$

$$\alpha^{t\hat{f}}_{firs} = \sum_{o \in \mathcal{O}} \lambda^{t\hat{f}o}_{firs} \quad \text{where } \{\alpha^{t\hat{f}}_{firs} | \hat{f} \in \mathcal{F}\} \text{ is a SOS2} \quad \forall f, \hat{f} \in \{\mathcal{F} | \hat{f} \geq f\}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.26}$$

$$\beta^{to}_{firs} = \sum_{\hat{f} \in \mathcal{F} | \hat{f} \geq f} \lambda^{t\hat{f}o}_{firs} \quad \text{where } \{\beta^{to}_{firs} | o \in \mathcal{O}\} \text{ is a SOS2} \quad \forall f \in \mathcal{F}, o \in \mathcal{O}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.27}$$

The balance equation (4.10) in the growth model can now be linearized by substituting $j^t_{\hat{f}firs} = x^t_{\hat{f}firs} n^t_{firs}$. This yields equation (4.28).

$$n^t_{\hat{f}irs} = \sum_{f \leq \hat{f} | f \in \mathcal{F}} (S^{fish}_f j^{t-1}_{\hat{f}firs}) - e^t_{\hat{f}irs} - h^t_{\hat{f}irs} + S^{egg} y^t_{\hat{f}irs} + \sum_{\substack{\hat{i} \in \mathcal{I}, \\ \hat{r} \in \mathcal{R}}} (d^{tir}_{\hat{f}\hat{i}\hat{r}s} - d^{t\hat{i}\hat{r}}_{\hat{f}irs}) \qquad \forall \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.28}$$

4.3.2 Initial biomass

To account for the initial inventory at $t = 1$ and that no growth occur in periods prior to $t = 1$, the balance equation can be simplified to equation (4.29), where $I_{\hat{f}ir}$ denotes the initial inventory of fish in weight class $\hat{f}$. The balance equation is therefore defined by equation (4.29) for $t = 1$ and equation (4.30) for $t > 1$.

$$n^t_{\hat{f}irs} = I_{\hat{f}ir} - e^t_{\hat{f}irs} - h^t_{\hat{f}irs} + S^{egg} y^t_{\hat{f}irs} + \sum_{\substack{\hat{i} \in \mathcal{I}, \\ \hat{r} \in \mathcal{R}}} (d^{tir}_{\hat{f}\hat{i}\hat{r}s} - d^{t\hat{i}\hat{r}}_{\hat{f}irs}) \qquad \forall \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t = 1 \tag{4.29}$$

$$n^t_{\hat{f}irs} = \sum_{f \leq \hat{f} | f \in \mathcal{F}} (S^{fish}_f j^{t-1}_{\hat{f}firs}) - e^t_{\hat{f}irs} - h^t_{\hat{f}irs} + S^{egg} y^t_{\hat{f}irs} + \sum_{\substack{\hat{i} \in \mathcal{I}, \\ \hat{r} \in \mathcal{R}}} (d^{tir}_{\hat{f}\hat{i}\hat{r}s} - d^{t\hat{i}\hat{r}}_{\hat{f}irs}) \qquad \forall \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} | t > 1 \tag{4.30}$$

4.3.3 End of horizon biomass

The end of horizon ($t = T$) biomass, $n^T_{\hat{f}irs}$, for each fish class $\hat{f}$ must be equal to the initial biomass, $I_{\hat{f}ir}$ in the system. This condition must hold for each tank i at region r in every scenario s.

$$\begin{aligned} n^T_{\hat{f}irs} &= I_{\hat{f}ir} \\ \forall \hat{f} \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} \end{aligned} \tag{4.31}$$

4.3.4 Fry deployment

Equation (4.32) governs which weight class the hatchery can deliver fry, y^t_{firs}, to. For most freshwater facilities the hatchery only deliver fry to the smallest weight class, thus restricting y^t_{firs} to be zero for all other weight classes.

$$\begin{aligned} y^t_{firs} &= 0 \\ \forall f \in \mathcal{F} | f \neq 1, t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} \end{aligned} \tag{4.32}$$

4.3.5 Smolt delivery

The fish goes through a physiological change before it becomes smolt and can live in seawater. It is assumed that this only happens with fish in the largest weight classes and that the process can be fully controlled by the facility operators. Equation (4.33) ensures that the delivered amounts, h^t_{firs}, from tank i at region r at time t can only occur from smolt weight class c where $c \in \mathcal{C}$ and $\mathcal{C} \subset \mathcal{F}$.

$$\begin{aligned} h^t_{firs} &= 0 \\ \forall f \in \mathcal{F} | f \notin \mathcal{C}, t \in \mathcal{T}, s \in \mathcal{S}, r \in \mathcal{R} \end{aligned} \tag{4.33}$$

The amount and weight of the delivered smolt determine whether an order has been fulfilled. Equation (4.34) ensures that the order amount, H^t_c, of smolt in class c at time t are satisfied.

$$\begin{aligned} \sum_{\substack{i \in \mathcal{I}, \\ r \in \mathcal{R}, \\ k \in \{U^{lower}_c \ldots U^{upper}_c\}}} h^t_{kirs} &= H^t_c \\ \forall c \in \mathcal{C}, t \in \mathcal{T}, s \in \mathcal{S} \end{aligned} \tag{4.34}$$

An order for smolt in class c can be satisfied by the sum of amounts, h^t_{kirs}, from nearby smolt classes, $k \in \{U^{lower}_c \ldots U^{upper}_c\}$, in all tanks i and all regions r. Feasible deviations in the delivery weight class is determined by the highest acceptable weight class U^{upper}_c and lowest acceptable weight class U^{lower}_c.

All of the smolt classes k which cannot be used to satisfy the order for class c at time t must be set equal to zero. This is enforced in equation (4.35).

$$\begin{aligned} h_{kirs}^t = 0 \\ \forall c \in \mathcal{C} | H_c^t > 0, k \in \mathcal{C} | k \notin \{U_c^{lower} \dots U_c^{upper}\}, \\ i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{aligned} \tag{4.35}$$

All delivery variables are set to zero if there are no orders at time t. This is enforced in equation (4.36).

$$\begin{aligned} h_{cirs}^t = 0 \\ \forall c \in \mathcal{C} | \sum_{c \in \mathcal{C}} H_c^t = 0, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{aligned} \tag{4.36}$$

The harvesting formulation assumes that there is at most one order per time unit and that the order contains one weigh class, as specified in section 3.3. Thus all delivery variables, h_{kirs}^t, that are not within the acceptable weight delivering interval for a given order is set to zero.

4.3.6 Tank temperature

The temperature, v_{irs}^t, in each tank i at region r in scenario s at time t is set in equation (4.37).

$$\begin{aligned} v_{irs}^t = \tau_s^t \frac{b_{irs}^t}{b_{irs}^t + w_{irs}^{t+} + w_{irs}^{t-}} + G_r^{max} \frac{w_{irs}^{t+}}{b_{irs}^t + w_{irs}^{t+} + w_{irs}^{t-}} + G_r^{min} \frac{w_{irs}^{t-}}{b_{irs}^t + w_{irs}^{t+} + w_{irs}^{t-}} \\ \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{aligned} \tag{4.37}$$

It is a function of the freshwater intake temperature, τ_s^t, the temperature of the heated water, G_r^{max}, and the temperature of the cooled water, G_r^{min}. The flow of the water intake, b_{irs}^t, warm water, w_{irs}^{t+}, and cold water, w_{irs}^{t-}, in relation to the total flow $b_{irs}^t + w_{irs}^{t+} + w_{irs}^{t-}$ sets the temperature in each tank.

Equation (4.37) is non-linear, which makes it computationally undesirable. However, it can be linearized if a constant flow of W_{ir} is assumed. By doing this, w_{irs}^{t+}, can be described as a fraction, q_{irs}^{t+}, of the constant flow. Similarly, w_{irs}^{t-} can be described by q_{irs}^{t-}. Also, the fraction of intake water is $1 - q_{irs}^{t-} - q_{irs}^{t+}$. The tank temperature, v_{irs}^t, can then be set according to equation (4.38). Since the model minimizes total costs, it is unnecessary to restrict only q_{irs}^{t-} or q_{irs}^{t+} to be non-zero.

$$\begin{aligned} v_{irs}^t = \tau_s^t + (G_r^{max} - \tau_s^t) q_{irs}^{t+} + (G_r^{min} - \tau_s^t) q_{irs}^{t-} \\ \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \end{aligned} \tag{4.38}$$

The processing capacity (heating or cooling) may be smaller than the constant water intake flow in a tank. It is therefore necessary to set an upper limit on the on the fraction of warmed water, L^{+}_{tir}, and cooled water, L^{-}_{tir}. This is enforced in equation (4.40) and (4.39).

$$\begin{gathered} q^{t+}_{irs} \leq L^{+}_{tir} \leq 1 \\ \forall t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} \end{gathered} \tag{4.39}$$

$$\begin{gathered} q^{t-}_{irs} \leq L^{-}_{tir} \leq 1 \\ \forall t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} \end{gathered} \tag{4.40}$$

In some cases the water intake temperature, τ^{t}_{s}, can be higher than the temperature of the heated water, G^{max}_{r}, or lower than the temperature of the cooled water, G^{min}_{r}. Restrictions are therefore necessary to ensure a consistent use of the heated and cooled water. For instance, it would be impossible to use "heated" water to lower the water intake temperature if the facility does not have access to cooling equipment. This aspect is enforced in equation (4.41) and (4.42).

$$\begin{gathered} q^{t+}_{irs} = 0 \\ \forall t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} | \tau^{t}_{s} \geq G^{max}_{r} \end{gathered} \tag{4.41}$$

$$\begin{gathered} q^{t-}_{irs} = 0 \\ \forall t \in \mathcal{T}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S} | \tau^{t}_{s} \leq G^{min}_{r} \end{gathered} \tag{4.42}$$

All tanks in a region share a given capacity to heat water, TL^{+}_{r}, and cool water, TL^{-}_{r}. The total flow of heated water, $q^{t+}_{irs} W_{ir}$, and cooled water, $q^{t-}_{irs} W_{ir}$, cannot surpass these regional limits. This is enforced in equation (4.43) and (4.44).

$$\begin{gathered} \sum_{i \in \mathcal{I}} W_{ir} q^{t+}_{irs} \leq TL^{+}_{r} \\ \forall t \in \mathcal{T}, r \in \mathcal{R}, s \in \mathcal{S} \end{gathered} \tag{4.43}$$

$$\begin{gathered} \sum_{i \in \mathcal{I}} W_{ir} q^{t-}_{irs} \leq TL^{-}_{r} \\ \forall t \in \mathcal{T}, r \in \mathcal{R}, s \in \mathcal{S} \end{gathered} \tag{4.44}$$

4.3.7 Maximum allowed biomass

The maximum allowed biomass in a tank is restricted by two factors: the fish density and the water flow.

4.3.7.1 Maximum fish density

The fish density in a tank is a function of the tank size, L_{ir}^{size}, and the total biomass, $\sum_{f\in\mathcal{F}} n_{firs}^{t} V_f$. Guidelines at Marine Harvest (2013b) and regulators (Ministry of Fisheries and Coastal Affairs, 2008) sets an upper limit on the density in a tank, RD_{ir}, as specified in equation (4.45).

$$\frac{\sum_{f\in\mathcal{F}} V_f n_{firs}^{t}}{L_{ir}^{size}} \leq RD_{ir} \qquad \forall t\in\mathcal{T}, i\in\mathcal{I}, r\in\mathcal{R}, s\in\mathcal{S} \tag{4.45}$$

4.3.7.2 Minimum water flow

The faster the fish grows, the more waste products it creates. The only practical way to remove these waste products is by having a sufficiently high water flow. Thus, when the water flow, W_{ir}, in each tank is assumed to be constant (as specified in section 4.3.6), it puts a limit on the amount of biomass in each tank.

The minimum required flow per kilogram of fish, m_{irs}^{t}, depends on the metabolic activity (growth) and can be approximated through the use of the "7 *flow factor method*" as shown in equation (4.46) (Marine Harvest, 2013b).

$$m_{irs}^{t} = VCO2_{ir}\frac{100}{7}\sum_{f\in\mathcal{F}}\frac{g_{firs}^{t,\,daily}}{V_f}\frac{n_{firs}^{t}}{\sum_{\hat{f}\in\mathcal{F}} n_{\hat{f}irs}^{t}} \qquad \forall i\in\mathcal{I}, r\in\mathcal{R}, s\in\mathcal{S}, t\in\mathcal{T} \tag{4.46}$$

This method uses the average *daily* growth in percentage, $\sum_{f\in\mathcal{F}}\frac{g_{firs}^{t,\,daily}}{V_f}\frac{n_{firs}^{t}}{\sum_{\hat{f}\in\mathcal{F}} n_{\hat{f}irs}^{t}}$, for a given tank i at region r to find the minimum required flow. V_f, $g_{firs}^{t,\,daily}$ and n_{firs}^{t} is the weight, daily growth and amount respectively of the fish class f. CO_2 ventilators, if available, reduces the flow requirement with a given fraction, $VCO2_{ir}$. This method is further explained in the standard procedures of Marine Harvest (2013b).

Equation (4.47) enforces that the flow in the tank, W_{ir}, is greater than or equal to the required flow in the tank, $m_{irs}^{t}\sum_{f\in\mathcal{F}} n_{firs}^{t} V_f$.

$$m_{irs}^{t}\sum_{f\in\mathcal{F}} V_f n_{firs}^{t} \leq W_{ir} \qquad \forall i\in\mathcal{I}, r\in\mathcal{R}, s\in\mathcal{S}, t\in\mathcal{T} \tag{4.47}$$

Equation (4.47) is non-linear, and have to be linearized in order to make the optimization model tractable. The approach used is to pre-calculate the flow requirement m^t_{irs} as the stochastic parameter κ^t_{irs}. The new flow restriction is given in equation (4.48).

$$\kappa^t_{irs} \sum_{f \in \mathcal{F}} V_f n^t_{firs} \leq W_{ir} \qquad \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{4.48}$$

This design choice was done in collaboration with Marine Harvest, and how κ^t_{irs} is pre-calculated while maintaining accuracy is described in section 5.5.3.

4.3.8 Non-anticipativity

Non-anticipativity constraints are used to ensure that the model does not optimize on unrevealed information (Birge and Louveaux, 2011). The non-anticipativity constraints can be represented compactly as in equation (4.49).

$$\frac{1}{|S(e)|} \sum_{s' \in S(e)} \mathcal{X}_{ts'} = \mathcal{X}_{ts} \qquad e \in \mathcal{E}, s \in \mathcal{S}(e), t \in \mathcal{T}(e) \tag{4.49}$$

Event nodes $e \in \mathcal{E}$ occur whenever the scenario tree splits. $\mathcal{S}(e)$ denotes all of the scenarios which go through event node e. $\mathcal{T}(e)$ is the set of all time periods prior to e and $\mathcal{X}_{ts}$ denotes all variables depending on time t and scenario s.

4.3.9 Non-negativity

Equation (4.50) ensures that the intended variables are non-negative.

$$\alpha^{t\hat{f}}_{firs}, \beta^{to}_{firs}, \lambda^{t\hat{f}o}_{firs}, d^{t\hat{i}\hat{r}}_{firs}, e^t_{firs}, g^{t,\,daily}_{firs}, g^t_{firs}, h^t_{cirs}, j^t_{\hat{f}firs}, n^t_{firs}, q^{t+}_{irs}, q^{t-}_{irs}, v^t_{irs}, y^t_{firs} \geq 0 \qquad \forall f, \hat{f} \in \mathcal{F}, c \in \mathcal{C}, i, \hat{i} \in \mathcal{I}, r, \hat{r} \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}, o \in \mathcal{O} \tag{4.50}$$

4.4 Overview of the stochastic programming model

The following table summarizes the equations which are included in the model.

Description	**Equation number**
Objective function	(4.6)
- Energy cost - Fry deployment cost - Growth cost	(4.7) (4.8) (4.9)
Linearized growth model	
- Balance equation - Transition amount - Relationship between growth and temperature	(4.29) and (4.30) (4.22)-(4.27) (4.14)
End of horizon biomass	(4.31)
Fry deployment	(4.32)
Smolt delivery	(4.33)-(4.36)
Tank temperature	(4.38)-(4.44)
Maximum fish density	(4.45)
Minimum water flow	(4.48)
Non-anticipativity	(4.49)
Non-negativity	(4.50)

Chapter 5

Model parameterization

This chapter describes in detail how the model formulated in chapter 4 has been parametrized for the Marine Harvest freshwater facility located at Slørdal, Sør-Trøndelag. The facility is used in the case studies in this thesis. This chapter contains five parts.

First, characteristics of the case study facility can be exploited in order to reduce the number of variables. Section 5.1 explains and justifies these simplifications.

Second, the parameters requiring a significant amount of preprocessing are presented in section 5.2-5.4. Section 5.2 describes and evaluates how temperature scenario trees have been generated. The relationship between temperature and growth is parametrized in section 5.3. Section 5.4 describes the statistical method used to derive the survival rates of eggs and fish.

Third, the remaining parameters not requiring significant pre-processing are presented in section 5.5. This section includes the technical parameters of the Slørdal facility, the facility aggregation into tanks and regions, boolean parameters denoting valid facility operations, parameters used to set the biomass restrictions, fish weight and number classes, allowed smolt order deviations and cost data.

Fourth, the system design of the implemented model is discussed in section 5.6.

Lastly, extensive testing was conducted in order to ensure model quality, an accurate parameterization and a correct implementation. Section 5.7 presents these activities along with a discussion of their results.

In order to ensure a model quality, an accurate parameterization and a correct implementation

5.1 Variable reduction

Based on three characteristics of the freshwater facility at Slørdal, a significant reduction in the number of variables can be obtained when implementing the model. First, not all regions allow water heating. Second, there are restrictions on valid regions for certain fish weight classes. Third, there are restrictions on the movement between the different regions. Section 5.1.1- 5.1.2 investigate how these characteristics can be exploited.

The number of variables is also related to the implementation of the scenario tree. Section 5.1.3 shows that by explicitly building the scenario tree rather than using non-anticipativity constraints, the number of variables can be reduced.

5.1.1 Valid regions for fish weights and water heating

Physical constraints such as tank size and water flow capacity determine the eligible fish sizes in the various regions. This knowledge can be used to avoid declaring variables and constraints. Let $A_{rf}^{\text{fish allowed}}$ be a boolean parameter which denotes whether fish weight class f is allowed in region r. $A_r^{heating}$ is a boolean parameter which denotes whether water can be heated at region r.

None of the regions at the Slørdal facility can cool water, and only one region can heat water. Outside this region the tank water temperature, v_{irs}^t, is determined solely by the intake water temperature, τ_{ts}. The consequence is that neither the growth, g_{firs}^t, nor the transition rate, $x_{\hat{f}f\hat{i}\hat{r}s}^t$, are variables. They can instead be pre-calculated and treated as stochastic parameters. Let $\chi_{\hat{f}fs}^t$ denote the stochastic transition rate, which can be pre-calculated using equation (5.1). This equation is closely related to equation (4.13) presented in the model formulation.

$$\chi_{\hat{f}fs}^t = \begin{cases} \frac{(V_f+g_{firs}^t)-V_{\underline{f}}}{V_{\overline{f}}-V_{\underline{f}}} & \text{if } (\hat{f}=\overline{f}) \wedge (V_{\overline{f}} \geq V_f + g_{firs}^t \geq V_{\underline{f}}) \wedge (*) \\ \\ \frac{V_{\overline{f}}-(V_f+g_{firs}^t)}{V_{\overline{f}}-V_{\underline{f}}} & \text{if } (V_{\overline{f}} \geq V_f + g_{firs}^t \geq V_{\underline{f}}) \wedge (\underline{f}=\hat{f}) \wedge (*) \\ \\ 0 & \text{else} \end{cases}$$
$$\forall f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T},$$
$$\forall r \in \mathcal{R} |\ (A_{rf}^{\text{fish allowed}} = \text{true}) \wedge (A_r^{\text{heating}} = \text{false})$$
$$*\overline{f} \text{ and } \underline{f} \text{ are adjacent weight classes} \tag{5.1}$$

The growth, g_{firs}^t, is either determined by the variable tank temperature or fresh-water intake temperature as shown in equation (5.2). This equation is similar to equation (4.14) presented in the model formulation.

$$g_{firs}^t = \begin{cases} A_f + B_f v_{irs}^t & \text{if } * \\ A_f + B_f \tau_{ts} & \text{else} \end{cases}$$
$$\forall f \in \mathcal{F}, i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T}$$
$$* = \ (A_{rf}^{\text{fish allowed}} = \text{true}) \wedge (A_r^{\text{heating}} = \text{true}) \tag{5.2}$$

By using the stochastic transition rate $\chi_{\hat{f}fs}^t$ when applicable and only declaring constraints for fish classes f valid in a region r, both the number of constraints and SOS2 can be significantly reduced. Specifically, equation (4.22)-(4.27) in the model formulation can be rewritten as equation (5.3)-(5.8).

$$\begin{gathered} V_f + g^t_{firs} = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} V_{\hat{f}} \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \wedge (A^{\text{heating}}_{r} = \text{true}) \end{gathered} \tag{5.3}$$

$$\begin{gathered} n^t_{firs} = \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} N_o \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \wedge (A^{\text{heating}}_{r} = \text{true}) \end{gathered} \tag{5.4}$$

$$\begin{gathered} j^t_{\hat{f}firs} = \begin{cases} \sum_{o \in \mathcal{O}} \lambda^{t\hat{f}o}_{firs} N_o & \text{if } (A^{\text{heating}}_{r} = \text{true}\ ) \\ \chi^t_{\hat{f}fs} n^t_{firs} & \text{else} \end{cases} \\ \forall \hat{f}, f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \end{gathered} \tag{5.5}$$

$$\begin{gathered} \sum_{\substack{o \in \mathcal{O} \\ \hat{f} \in \mathcal{F} | \hat{f} \geq f}} \lambda^{t\hat{f}o}_{firs} = 1 \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \wedge (A^{\text{heating}}_{r} = \text{true}) \end{gathered} \tag{5.6}$$

$$\begin{gathered} \alpha^{t\hat{f}}_{firs} = \sum_{o \in \mathcal{O}} \lambda^{t\hat{f}o}_{firs} \\ \text{where } \{\alpha^{t\hat{f}}_{firs} | \hat{f} \in \mathcal{F}\} \text{ is a SOS2} \\ \forall \hat{f}, f \in \{\mathcal{F} | \hat{f} \geq f\}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \wedge (A^{\text{heating}}_{r} = \text{true}) \end{gathered} \tag{5.7}$$

$$\begin{gathered} \beta^{to}_{firs} = \sum_{\hat{f} \in \mathcal{F} | \hat{f} \geq f} \lambda^{t\hat{f}o}_{firs} \\ \text{where } \{\beta^{to}_{firs} | o \in \mathcal{O}\} \text{ is a SOS2} \\ \forall f \in \mathcal{F}, i \in \mathcal{I}, s \in \mathcal{S}, t \in \mathcal{T}, o \in \mathcal{O} \\ \forall r \in \mathcal{R} |\ (A^{\text{fish allowed}}_{rf} = \text{true}\ ) \wedge (A^{\text{heating}}_{r} = \text{true}) \end{gathered} \tag{5.8}$$

Equation (5.9) ensures that the fish of a given weight class f is not allowed to be in a region r if the boolean $A_{rf}^{\text{fish allowed}}$ is false.

$$\begin{aligned} n_{firs}^{t} &= 0 \\ \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t &\in \mathcal{T} \\ \forall f \in \mathcal{F} | (A_{rf}^{\text{fish allowed}} &= \text{false}) \end{aligned} \tag{5.9}$$

Since only one out of three regions at the Slørdal facility can heat water, a 67% reduction in SOS2 can be achieved. This is because the approximated growth model only have to be declared for one region. Section 5.5.1 explains how the Slørdal facility is aggregated into tanks and regions, and section 5.5.2 explains how the boolean parameters are set.

Also, only the smallest weight classes are allowed in this region. Given the weight classes chosen in section 5.5.4, the approximate growth model only have to be declared for 9 out of 23 fish classes. The number of SOS2 are therefore further reduced by 61%. In total, these two characteristics reduced the number of SOS2 in the model by 87%. The number of continuous variables and constraints were also reduced, but the reduction in SOS2 had by far the largest impact on the solution time.

5.1.2 Valid movement of fish

Moving the fish causes distress and reduced growth as the fish have to be starved before it can be moved. Fish are therefore not moved unnecessarily back and forth between regions. $A_{r\hat{r}}^{\text{movement}}$ is a boolean parameter used to capture this praxis. Equation (5.10) ensures that if $A_{r\hat{r}}^{\text{movement}}$ is false then movement of fish, $d_{firs}^{t\hat{i}\hat{r}}$, from region r to region $\hat{r}$ is illegal.

$$\begin{aligned} d_{firs}^{t\hat{i}\hat{r}} &= 0 \\ \forall f \in \mathcal{F}, \hat{i}, i \in \mathcal{I}, s \in \mathcal{S}, t &\in \mathcal{T}, \\ \forall \hat{r}, r \in \mathcal{R} |\ A_{r\hat{r}}^{\text{movement}} &= \text{false} \end{aligned} \tag{5.10}$$

5.1.3 Explicitly building the scenario tree

Scenario trees can be represented either explicitly or implicitly, as shown in figure 5.1. The explicit formulation has only one variable, x, prior to the split, which ensures non-anticipativity. The implicit formulation has several variables, x_s, before the split, and non-anticipativity constraints have to be included to ensure consistency amongst them. Birge and Louveaux (2011) showed that an optimization model using either an implicit or explicit scenario tree yield the same optimal solution. However, the implicit trees will contain redundant variables. Using explicitly formed trees therefore requires fewer variables. How much depends on where the splits occur in the scenario tree. For instance, a reduction of 42% occurs in the two-stage scenario tree used in the case study in chapter 6. This has a large impact on the model's tractability, and we have therefore used explicit scenario trees in the *implementation*.

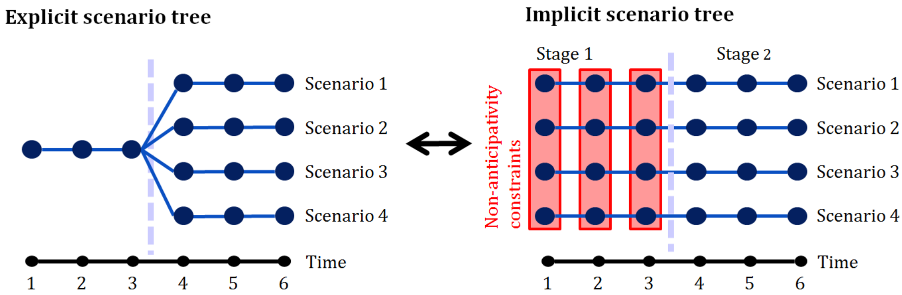

Figure 5.1: Implicit and explicit scenario trees are equivalent. Blue circles denotes variables at a given time, blue lines denotes connections between variables in a scenario and red squares denotes non-anticipativity constraints.

5.2 Generating water temperature scenarios

The water intake temperature scenarios and tree structure used for the stochastic programming model can have a large impact on the optimal solution found. Kaut and Wallace (2007), and Zenios (2008) mention several important aspects to consider when generating scenarios:

1. The scenario set should be derived from a correct underlying model. Correctness is measured by how well the model represents distributional moments, seasonality etc.

2. The discretization of the underlying variable should accurately approximate the underlying distribution of the random variable.

3. The values of the variables in the scenarios have to be internally consistent. Since we are only dealing with one time series, autocorrelation is the relevant aspect to consider.

4. The optimization model using the scenario generation procedure should be *stable*. Simply put, this requirement implies that different scenario trees generated by the same procedure and input should yield approximately the same optimal objective value (see section 5.7.3).

5. The model should be able to produce a *tractable* number of scenarios for the optimization model.

We chose a scenario generation approach based on a time series model, due to the mentioned aspects. The approach consists of splitting the temperature series into two components: one deterministic and one stochastic. The deterministic component is represented through a time series model, and the stochastic component is the residual distribution of this time series model. To generate temperature scenarios, the residual distribution is discretized into a finite number of equally likely realizations, that is, *residual scenarios*. The residual scenarios are then combined with the time series model to generate a scenario tree.

The following approach is used in section 5.2.1-5.2.5 to generate temperature scenarios:

1. **Ensure stationarity in the temperature time series.**
2. **Identify candidate time series models.**
3. **Parametrize the candidate models and select the best one.**
4. **Create residual term scenarios.**
5. **Generate temperature scenarios from the time series model and residual scenarios.**

Step 1 through 3 is based on the Box-Jenkins methodology (Box et al., 2011) for fitting a model to a time series. Step 4 is done using the moment-matching heuristic developed by Høyland et al. (2003). The quality of the water temperature scenario set will be evaluated and tested in sections 5.2.6 and 5.7.3.

5.2.1 Ensure stationarity in the temperature time series

A stochastic process $\{Z_t\}$ is stationary if its joint probability distribution $f(Z_t, ..., Z_{t+k})$ $\forall k \in \mathbb{Z}^+$ does not change when shifted in time. This implies that core statistical parameters, such as mean and variance, does not change over time. Stationarity is a core assumption of many time series models, and is in many cases necessary for the model analysis and results to be mathematically valid (Wei, 2006). This section will therefore present the method employed to test for and ensure stationarity in the water temperature time series.

A common way to deal with non-stationarity is to transform the time series in such a way that the resulting time series is stationary. Some non-stationary time series, e.g. prices of some financial products, can be made stationary by differencing them (Wei, 2006). However, many time series are stationary in the mean but non-stationary in the variance. A stochastic process with time varying variance is called *heteroskedastic*. Conversely, a process without time varying variance is called *homoskedastic*. In these cases, a variance stabilizing transformation can be applied in order to get a stationary time series.

The power transformation described by (5.11) can be used as a variance stabilizing technique and was introduced by Box and Cox (1964). Here, Z_t is a random variable and λ is the transformation parameter.

$$BC(Z_t, \lambda) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln Z_t & \text{if } \lambda = 0 \end{cases} \tag{5.11}$$

Since the analysis of variance is unchanged by a linear transformation, equation (5.11) is equivalent to (5.12).

$$BC(Z_t, \lambda) = \begin{cases} Z_t^\lambda & \text{if } \lambda \neq 0 \\ \ln Z_t & \text{if } \lambda = 0 \end{cases} \tag{5.12}$$

The best suited value for λ can be determined from the time series $\{Z_t\}$ using maximum likelihood estimation. The details of the mathematics underlying the parameter estimation process is discussed by Box and Cox (1964).

Stationarity has to be ensured before a time series model can be estimated on the water intake temperature series. The first step is to test for stationarity in mean and variance. Figure 5.2 shows that the time series does not appear to have a trend, just seasonal effects. This indicates that the time series is stationary in mean.

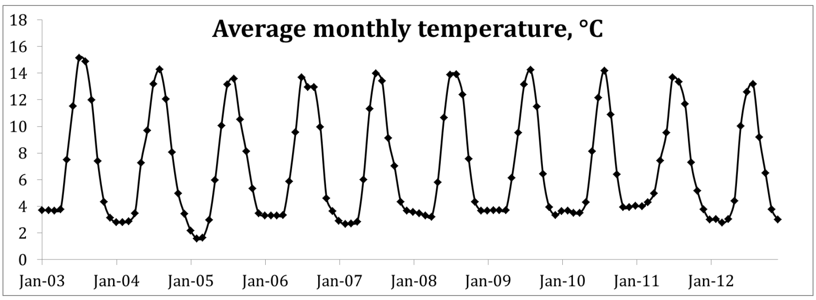

Figure 5.2: Average monthly water intake temperature at the Slørdal facility in the period 2003-2012.

A formal test for this is the Augmented Dickey-Fuller test (Dickey and Fuller, 1981). It tests the null-hypothesis, H_0, of non-stationarity (unit root present) versus the alternative hypothesis, H_1, of stationarity (no unit root present) in the mean. A summary of the results from of this test is shown in table 5.1, and the null-hypothesis of non-stationarity in mean is rejected at a 5% significance level.

	Statistic	**5% Critical Value**	**Reject H_0 at 5%?**
Augmented Dickey-Fuller (H_0=non-stat. in mean)	-7.971	-2.89	Yes
Hetero-X (H_0=stat. in variance)	2.9143	1.81	Yes

Table 5.1: Stationarity test statistics for the freshwater temperature series.

The temperature appears to vary less during the winter than the other seasons. This could indicate that the time series is *heteroskedastic*. In order to test for this, a Hetero-X test has been used. This is an F-test for heteroskedasticity in the residual process of an auxiliary regression which is based on White's test (White, 1980). The difference is that the Hetero-X test is also able to capture cross-correlations between regressors as well, thus making it a more "thorough" test. The null-hypothesis for this test is that the residual process is homoskedastic. Table 5.1 shows a summary of this test and shows that at a 5% significance level, the null-hypothesis of homoskedasticity is rejected.

To summarize, the water intake temperature series is stationary in mean but not in variance. A variance stabilizing transformation such as the Box-Cox power transformation is therefore a natural choice. The first step necessary to transform the series using equation (5.11) is to perform a Box-Cox test in order to determine the optimal transformation parameter λ in terms of maximum likelihood. Figure 5.3 shows the log-likelihood values for various λ's. $\lambda = 0.5$ is the optimal choice, and thus the power transformation in (5.12) can be simplified to (5.13).

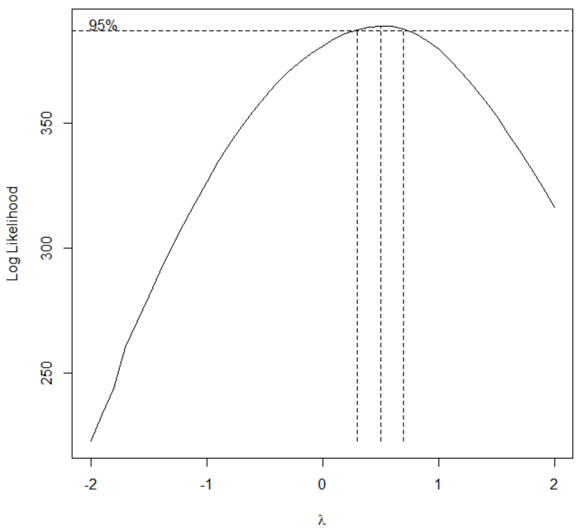

Figure 5.3: The log-likelihood for λ values estimated on the water temperature series. The optimal transformation parameter is $\lambda = 0.5$.

$$BC(Z_t, \lambda = 0.5) = \sqrt{Z_t} \tag{5.13}$$

A common technique for validating the effectiveness of the transformation, is to apply the Box-Cox test again on the transformed series. Figure 5.4 show that the optimal transformation parameter $\lambda = 1$ and no further transformations are necessary.

A time series model can therefore be estimated on the transformed temperature series $\{\sqrt{Z_t}\}$.

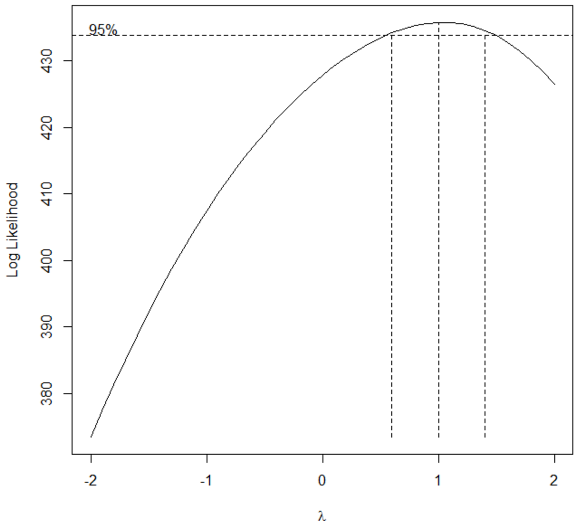

Figure 5.4: The log-likelihood for λ values estimated on the transformed water temperature series. The optimal transformation parameter is $\lambda = 1$

5.2.2 Identify candidate time series models.

There are two basic time series models, autoregressive models (AR) and moving average models (MA). By combining them one can create popular models such as ARMA and ARIMA which can be further evolved into more advanced models like GARCH models or State Space models (Wei, 2006). Since the models are closely related it is important to identify candidate models before testing for a possible fit between the chosen model and the time series. Not doing so might result in the use of a "suboptimal" model. Since the models are related, many model classes are likely to fit but not necessarily be the best fit.

The process of finding a candidate model consists of using both mathematical tools and human interpretation. The core mathematical tools used are the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The ACF(p) gives the correlation between the time series Z_t and the p lags shifted time series Z_{t+p}. The PACF(p) gives the correlation between Z_t and Z_{t+p} given that Z_{t+1} to Z_{t+p-1} are known.

The ACF or PACF is said to *tail off*, if there is a gradual reduction in value for each incrementation of the lag, and to *cut off*, if there is a distinct reduction in the value. Thus, the nominal value of the ACF or PCF at a given lag is not important, but rather its relative size compared to the other lags. Determining a good candidate time series model from the ACF and PACF is based on these guidelines:

1. If the ACF tails off and the PACF cuts off at lag p, there is strong indication of an AR(p) model.

2. If the PACF tails off and the ACF cuts off at lag q, there is a strong indication of a MA(q) model.

3. If the ACF tails off after lag (q - p) and the PACF tails off after lag (p - q), there is a strong indication of an ARMA(p,q) model.

Wei (2006) presents a mathematical proof for these guidelines. It should be noted that these guidelines assumes that the residuals are i.i.d and normally distributed. It is therefore important that these assumptions are checked once the model is estimated. The process of finding candidate time series model using the PACF and ACF can be split into two steps:

1. **Inspect the plot** of the stationary time series and look for trends or seasonality.

2. **Inspect the ACF and PACF** to find candidate models. If necessary remove seasonal effects by using explanatory variables or differencing the time series, and inspect the plots again.

5.2.2.1 Inspect the plot of the BC-transformed temperature series

A visual inspection of the time series in figure 5.5 indicates a seasonal pattern, with peaks during the summer and bottoms during the winter. The ACF and PACF cannot be used directly to identify the candidate time series model if the time series is seasonal (Wei, 2006). These effects should be removed before proceeding.

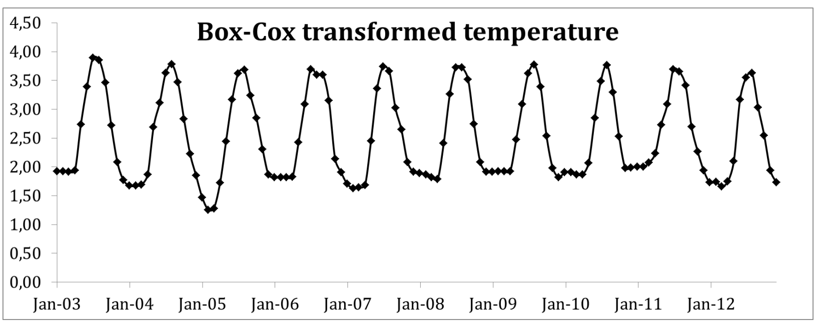

Figure 5.5: The Box-Cox transformed temperature series.

5.2.2.2 Inspect the ACF and PACF to find candidate models

The ACF and PACF is shown in figure 5.6. The ACF is not decaying, indicating that the seasonal effects should be removed.

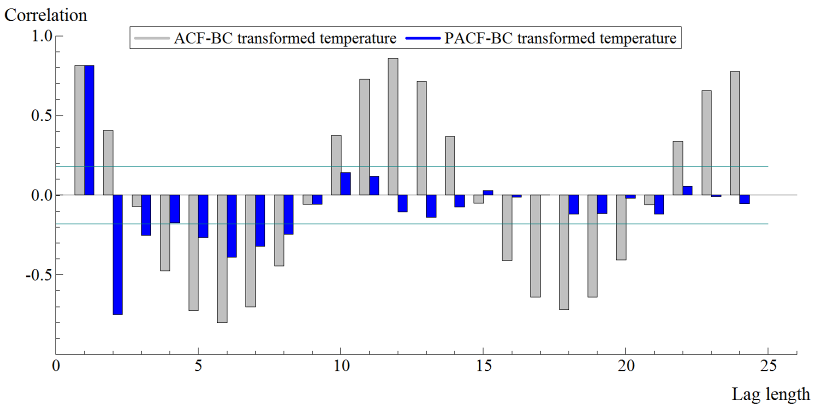

Figure 5.6: The ACF and PACF for the BC-transformed time series.

Brooks (2008) explains how seasonal effects can be removed by introducing *binary dummy variables*, $\delta_{d,t}$. It denotes if time t is part of the seasonal period d (month, quarter, winter, summer etc.). The binary variable is 1 if time t is in the period d and 0 else. For instance $\delta_{d,t} = 1$ if $t =$ January and $d =$ First Quarter. These dummy variables are used as explanatory variables in a linear regression model, and the residual time series, $\{\epsilon_t\}$, is the seasonally adjusted time series. Equation (5.14) shows this regression model.

$$\sqrt{Z_t} = \mu + \sum_{d \in D} \alpha_d \delta_{d,t} + \epsilon_t \tag{5.14}$$

In addition, one binary variable is removed from each set of seasonal factors (one of the months, one of the quarters etc.) to avoid multicollinearity (Alexander, 2009).

When removing the seasonal effects in the BC-transformed water intake series, one dummy variable per month except January has been used to adjust the BC-transformed water temperature series. Figure 5.7 shows that seasonally adjusted series has lost its seasonality effects. Figure 5.8 shows the corresponding ACF and PACF. The ACF is tailing off and the PACF cuts off at lag 1. An MA model can therefore be excluded, and with it also ARIMA and ARMA models.

The PACF indicates that the AR(1) is the only suited candidate model. Since the time series was seasonally adjusted, the suited candidate model is a **seasonal AR(1) model**. However, both goodness-of-fit and model assumptions (independence, normality, homoskedasticity) have to be verified through statistical analyses before a final conclusion can be made. This is performed in section 5.2.6.

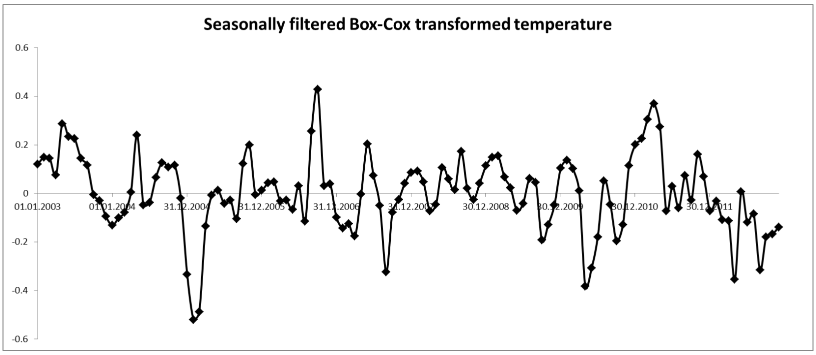

Figure 5.7: A plot of the seasonally adjusted BC-transformed temperature series.

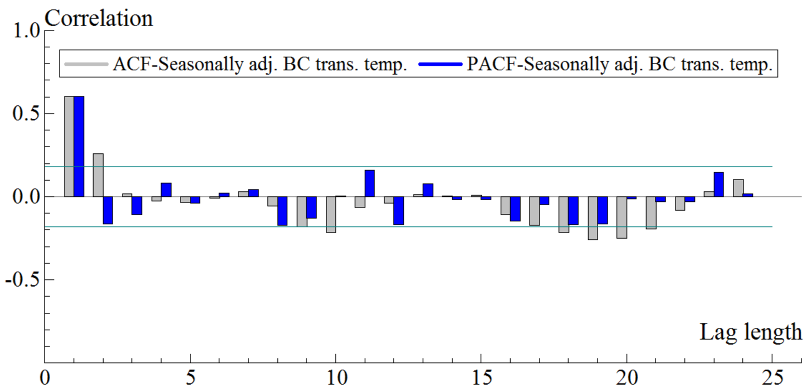

Figure 5.8: The ACF and PACF of the seasonally adjusted BC-transformed temperature series

5.2.3 Parametrize the seasonal AR(1) model

The general seasonal AR(1) model is defined in equation (5.15). In addition to the seasonal dummy variables used previously, the BC-transformed temperature from the previous period, $\sqrt{Z_{t-1}}$, is included as a regressor.

$$\sqrt{Z_t} = \mu + \sum_{d \in D} \alpha_d \delta_{d,t} + \beta \sqrt{Z_{t-1}} + \epsilon_t \quad ; \quad \epsilon_t \sim N(0, \sigma) \tag{5.15}$$

ϵ_t is a normally distributed residual term, μ is a constant, $d \in D$ is the set of seasonal effects, α_d are the seasonal effect coefficients and β is the coefficient of the 1 lag transformed temperature.

The coefficients, μ, α_d, and β, are estimated using Ordinary Least Squares (Alexander, 2009) and the historical water temperatures at Slørdal from the period 2003-2012. The resulting coefficients are summarized in table 5.2.

	Coefficient	**t-probability**
μ	0.650	0.000
β	0.610	0.000
α_{feb}	0.024	0.691
α_{mar}	0.037	0.539
α_{apr}	0.131	0.032
α_{may}	0.668	0.000
α_{jun}	1.012	0.000
α_{jul}	1.090	0.000
α_{aug}	0.827	0.000
α_{sep}	0.430	0.006
α_{oct}	0.037	0.771
α_{nov}	-0.203	0.022
α_{dec}	-0.066	0.291

Table 5.2: The coefficients of the seasonal AR(1) model.

The adjusted R^2 is 0.969 for the seasonal AR(1) model indicating that it explains a very large part of the total variation in the time series and that the model fit is very good. However, the mathematical tools used to select, estimate and evaluate this model depends on certain assumptions (independence, normality, homoskedasticity). These aspects of the estimated seasonal AR(1) model are therefore evaluated in section 5.2.6.

5.2.4 Create residual term scenarios

The residual time series $\{\epsilon_t\}$ from the seasonal AR(1) model is the stochastic component that is used to create scenarios for the optimization model.

Høyland et al. (2003) presented an approach to represent a multivariate probability distribution through a finite number of scenarios. This is called the moment-matching method, and it finds a scenario set which approximately matches the first four moments of each marginal distribution and the correlation matrix between the random variables. We have used this approach and the software ScenGen (Kaut, 2003) to estimate scenario sets from the residual distribution.

A summary of statistics for the residual distribution is shown in table 5.3 and a plot of a normal distribution with the residual distribution in figure 5.9. The excess kurtosis is a bit larger than zero, meaning that the distribution has a higher peak and fatter tails than a normal distribution. The negative skewness indicate that the distribution is slightly asymmetric and "leaning" towards the right.

Observations	120
Mean	0
Std.Devn.	0.12648
Skewness	-0.37019
Excess Kurtosis	0.63622
Minimum	-0.38829
Maximum	0.32556

Table 5.3: Descriptive statistics for the residual time series.

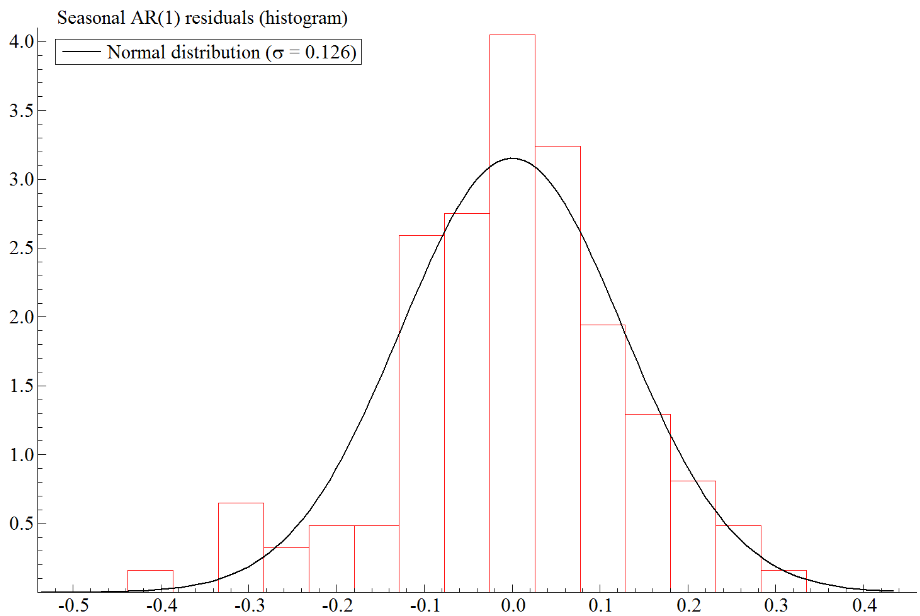

Figure 5.9: Shows the residual distribution vs. a normal distribution.

Matching the first four moments of the residual distribution, as shown in table 5.3, generates a set of equally likely residual scenarios $\{\epsilon^s\}$ where $s \in \mathcal{S}$ and the set size $|\mathcal{S}| \in \{4, 5, ..., \infty\}$.

By employing an assumption of normally distributed residuals, which can be justified by the analysis in section 5.2.6, equally likely residual scenarios of set size $|\mathcal{S}| \in \{2, 3, ..., \infty\}$ can be generated.

5.2.5 Generate temperature scenarios

The seasonal AR(1) model can be used to create scenarios. This is done by letting the time series method output a forecast, $\sqrt{Z_t}$, of the BC-transformed temperature at any future point in time t. $\sqrt{Z_t}$ is given by equation (5.16).

$$\sqrt{Z_t} = \mu + \sum_{d \in D} \alpha_d \delta_{d,t} + \beta \sqrt{Z_{t-1}} \tag{5.16}$$

Elements from the residual scenario set $\{\epsilon^s\}$ are added only when the scenario tree splits, making the forecast at scenario splits $\sqrt{Z_t} + \epsilon^s$. This process is illustrated in figure 5.10. Each scenario generated by this procedure is then inverse-transformed using the Box-Cox power transformation. That is, each scenario is squared.

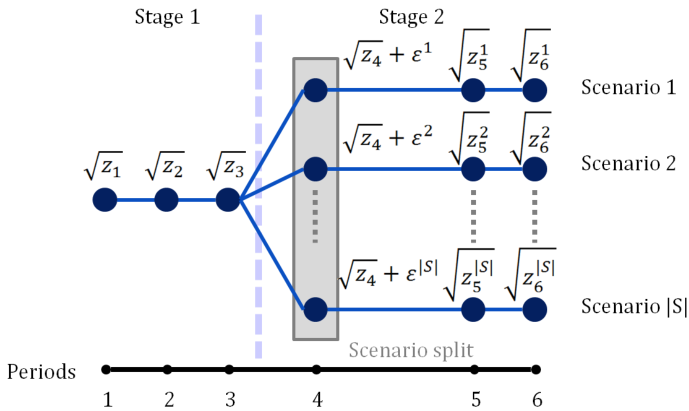

Figure 5.10: Shows how time series modeling and scenario tree generation can be combined.

5.2.6 Time series model evaluation

This section will evaluate the seasonal AR(1) model estimated on the transformed temperature series, to ensure that it conforms with model assumptions and desired characteristics. A more holistic approach to evaluating the scenario generation procedure together with the optimization model is applied in section 5.7.3 when the *stability* is assessed.

When estimating a linear regression model it is desirable that the residual process satisfy the following three properties:

1. Independence
2. Homoskedasticity
3. Normality

When these properties are satisfied, the estimators are deemed to be "best linear unbiased estimators" (BLUE). That is, they have the lowest prediction error amongst the class of all unbiased estimators. This has important implications and not fulfilling these criteria *could* render the model results and test statistics biased and useless. Thus, the time series model estimated in section 5.2.3 will in this section be evaluated based on these criteria.

	Distribution	Statistic	p-value	Reject H_0 at 5%?
Jarque-Bera	$\chi^2(2)$	4.2438	0.1198	No
Hetero-X	F(13,106)	1.5553	0.1100	No
Ljung-Box (12)	$\chi^2(12)$	18.471	0.1021	No

Table 5.4: Test statistics for the residual time series.

Independence When testing the residuals for independence, both the *individual* autocorrelation and the *joint* autocorrelation have been tested for the first 12 lags.

The ACF for the residuals is shown in figure 5.11. For all lag lengths except 10, the autocorrelation is not significantly different from 0 at a 5% level. When testing the joint autocorrelation of the residual series, the Q-statistic (Ljung and Box, 1978) has been used for 12 lags. Table 5.4 shows that the null-hypothesis of no autocorrelation cannot be rejected at a 5% significance level.

The implication of significant individual autocorrelation is that the regression estimators remain unbiased but are no longer BLUE (Brooks, 2008).

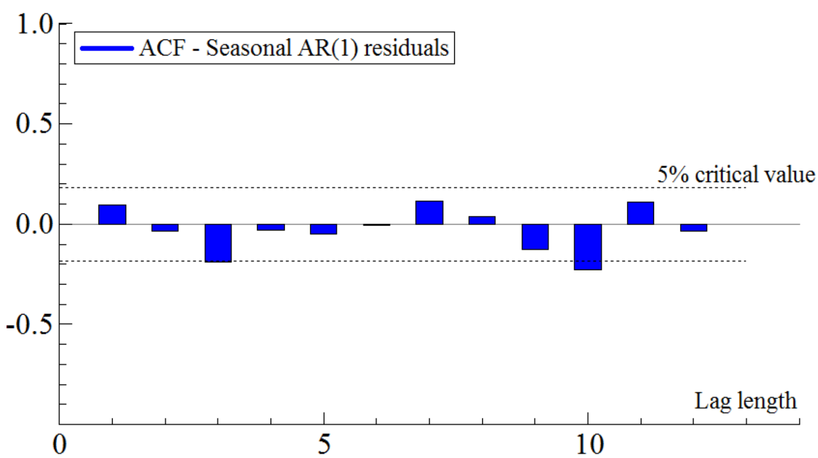

Figure 5.11: The ACF for the residual distribution versus the 5% critical value.

Heteroskedasticity The Hetero-X test is used to test for unconditional heteroskedasticity. As explained in section 5.2.1, this is a modified version of White's heteroskedasticity test (White, 1980). Table 5.4 shows that the null-hypothesis of unconditional homoskedasticity cannot be rejected at a 5% significance level. The implication is that the regression estimators remain BLUE.

Normality The Jarque and Bera (1980) statistic has been used to test for normality. This is a test for the null-hypothesis of a normal distribution, and looks at the skewness, and excess kurtosis in relation to the number of observations. Table 5.4 shows this test statistic for the residual series, and at a 5% significance level the null-hypothesis of normality cannot be rejected. The implication is that the regression estimators remain BLUE.

Summary From evaluating the seasonal AR(1)model, we have observed one limitation: there is some significant autocorrelation present at lag length 10. One possible approach to deal with this is to include a 10 period lagged term as a regressor. We tried this, and the regression coefficient was not significantly different from 0 at a 5% level. We have therefore chosen to keep the seasonal AR(1) model. We recognize that this leaves us with a model where the estimators are unbiased, but are not guaranteed to be the ones with the lowest error in the class of all unbiased estimators. However, taking into account an adjusted R^2 of 0.969, the model fit is considered good enough for this purpose even though there might exist estimators with a better fit. A possible approach to find optimal estimators in the presence of autocorrelation would be to use a Generalized Least Squares approach rather than the Ordinary Least Squares when finding the parameters (Alexander, 2009).

5.3 Growth regression model

The feed producer Skretting has extensively researched the relationship between growth and temperature for farmed Atlantic salmon. A product of this is the freshwater growth rate model with *daily* resolution shown in table 5.5.

	Temperature (°C)						
Weight (g)	2	3	4	...	13	14	15
0.1	0.81	1.22	1.62	...	5.28	5.69	6.09
0.2	0.80	1.20	1.60	...	5.22	5.62	6.02
...	...	...	...	...	...	...	...
20.0	0.29	0.45	0.62	...	2.10	2.26	2.42
22.5	0.27	0.43	0.59	...	2.01	2.17	2.33
...	...	...	...	...	...	...	...
90.0	0.10	0.18	0.26	...	0.97	1.05	1.13
100.0	0.09	0.16	0.24	...	0.89	0.97	1.04

Table 5.5: Daily growth in percentage for a given weight class and temperature. (Skretting, 2009)

As discussed in section 4.3, the *monthly* growth of the fish, g^t_{firs}, is a linear function of tank temperature, v^t_{irs}. Assuming that region r, tank i, time t and scenario s is fixed, this relationship is defined in equation (5.17).

$$g_f = A_f + B_f v \quad \forall f \in \mathcal{F} \tag{5.17}$$

The parameters of this monthly linear growth model, A_f and B_f, are estimated using the daily Skretting growth rate model in a three step process:

1. Continuously compound the growth rates in the Skretting model to change the period from 1 day to 1 month: $r^{monthly} = \exp^{30r^{daily}} - 1$
2. Calculate the monthly growth in grams: $g_f = V_f r_f^{monthly}$
3. For each weight class f, estimate the linear model using the monthly table in grams. The temperatures in the range $\tau_{min} = 2°C$ through $\tau_{max} = 15°C$ in the table are explanatory variables. This is the minimum and maximum water intake temperature observed at the Slørdal facility. The monthly growth in grams, g_f, for each temperature are the dependent variables. The parameter estimation is done using a variation of ordinary least squares (OLS) (Alexander, 2009). In order to avoid negative growth, the linear model is required to fit perfectly (zero error) at $\tau_{min} = 2°C$.

Figure 5.12 shows the monthly linear growth model estimated for two weight classes, 0.1 grams and 100 grams. We can see that the fit is better for larger weight classes than smaller weight classes, and that the regression model tends to overestimate the growth at low temperatures and underestimate it at high temperatures.

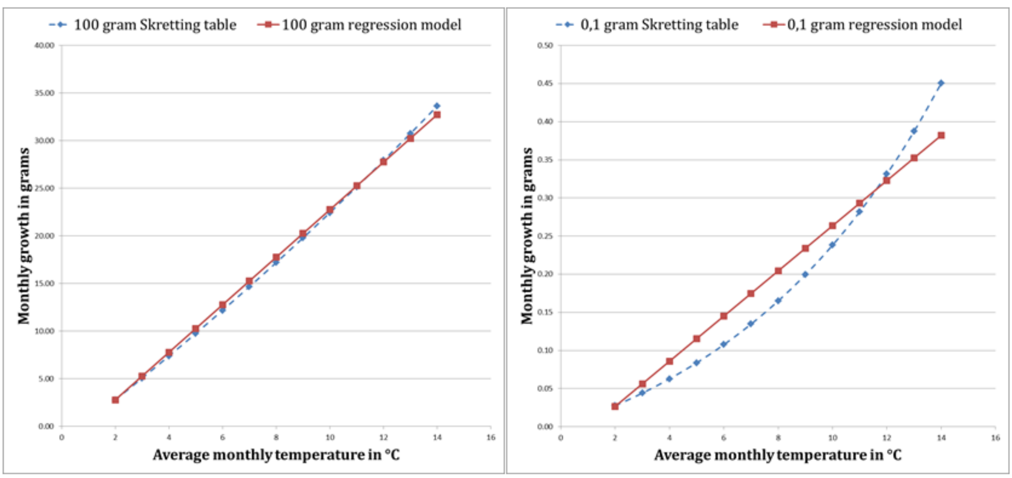

Figure 5.12: Growth regression model fit for the 0.1 gram (right) and 100 gram (left) weight class.

5.4 Survival rates

Section 1.5 thoroughly describes common reasons why fish and eggs die in the freshwater facility. This section quantifies the survival rate of eggs, S^{egg}, and the survival rate of fish, S_f^{fish}. The analysis is based on privately known data sets provided by Marine Harvest as well as expert opinions from their biologists.

5.4.1 Egg survival rate

Marine Harvest did not have any data sets regarding mortality rates for fertilized eggs to fry in freshwater facilities. Instead we consulted with one of Marine Harvest's freshwater biologists, Dr. Anders Fjellheim. According to him the mortality rate is between 2-3 % with the current production scheme. Also, they have not seen any variation in these numbers suggesting that the mortality rate is not affected by seasonality. A mortality rate of 2.5 % is deemed to be a good approximation leaving the survival rate to be $S^{egg} = (1 - 2.5\%) = 97.5\%$

5.4.2 Fish survival rate

Marine Harvest Freshwater does not currently possess a survival rate model (Marine Harvest, 2013a). However, they do have large amounts of raw data available regarding the mortality rates of fish in freshwater facilities. To be used in the optimization model, it needs to be analyzed in a way which is both mathematically valid and yields a desired output format. A methodology had to be developed because previous academic work on this topic is sparse. This was done in collaboration with Professor Bo Henry Lindqvist, which is an expert in statistical survival analysis at NTNU. This section presents this methodology and the resulting survival rates.

It is assumed in chapter 4 that the survival rate of fish, S_f^{fish}, depends only on the fish weight class f. This assumption is deemed reasonable by Marine Harvest (2013a). Marine Harvest provided a total of 235 data series from all smolt batches produced in Norway in the period of 2009-2011. Each series consists of weekly measurements of the average weight and amount of fish. The length of each series is in the range of 1 to 79 weeks. The methodology used to calculate the survival rate, S_f^{fish}, can be split into five steps:

1. Filter the data set to remove batches without sufficient data quality.
    - Some of the batch data series were missing data points or had outliers and erroneous observations. Due to this, 27 batch data series were removed, leaving the data set consisting of 208 batch data series.
2. Calculate the weekly mortality rate for each batch.
    - The weekly mortality rate for each batch, $R^{\text{Mortality per week}}$, was calculated as in equation (5.18), where N_t denotes the number of fish in a batch at

a given point in time. This approach assumes that the mortality rate is continuously compounded.

$$R^{\text{Mortality per week}} = \ln \frac{N_{\text{this week}}}{N_{\text{last week}}} \tag{5.18}$$

3. Define fish weight classes, and allocate all of the mortality rates calculated in step (2) to a class. Within each weight class, calculate the mean and variance of the mortality rates.
    - The weight classes used are those presented in section 5.5.4. Figure 5.13 shows the mean weekly mortality rate, $R_f^{\text{Mortality per week}}$, for each weight class f. Figure 5.14 shows the variance in the data set for each weight class f.
4. Rescale the weekly mortality rates to monthly mortality rates, $R_f^{\text{Mortality per month}}$, using continuous compounding as specified in equation (5.19).
    - Figure 5.13 shows the resulting monthly mortality rates per weight class. The rescaling assumes that each month consists of 4 weeks and that the fish population experiences the same mortality rates each of those weeks.

$$R_f^{\text{Mortality per month}} = e^{4R_f^{\text{Mortality per week}}} - 1 \qquad \forall f \in \mathcal{F} \tag{5.19}$$

5. Calculate the survival rate for each fish class as specified in equation (5.20).

$$S_f^{fish} = R_f^{\text{Mortality per month}} + 1 \qquad \forall f \in \mathcal{F} \tag{5.20}$$

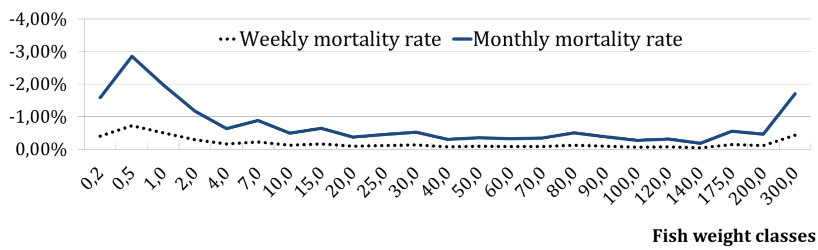

Figure 5.13: The calculated monthly ($R_f^{\text{Mortality per month}}$) and weekly ($R_f^{\text{Mortality per week}}$) mortality rates.

The cumulative mortality rate for the various types for fish batches were used to test the validity of the mortality rates. The case study results presented in chapter 6 shows that the cumulative mortality rates of the model are similar to the actual ones at the Slørdal facility.

The variance of the weekly mortality rates in figure 5.14 show that there is a difference between the weight classes. This effect could be reduced by using more advanced statistical techniques such as Kernel smoothing (Friedman et al., 2001). After discussing this issue with Bo Henry Lindqvist it was concluded further investigation would not yield any large improvements and that current methodology appeared to be good enough for the intended model.

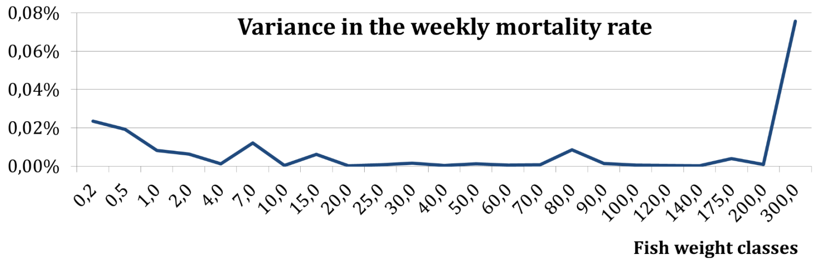

Figure 5.14: The variance in the weekly mortality rate.

5.5 Data sets

This section contains the technical parameters of the Slørdal facility, the facility aggregation into tanks and regions, boolean parameters denoting valid facility operations, parameters used to set the biomass restrictions, fish weight and number classes, allowed smolt order deviations and cost data.

5.5.1 Technical parameters

The technical parameters of the Slørdal facility are presented in this section along with the aggregation of the facility. The latter was done to reduce the size of the problem.

5.5.1.1 Facility characteristics

The Marine Harvest freshwater facility at Slørdal is located in Snillfjorden, Sør-Trøndelag. It is in close proximity to the sea and its freshwater source, Slørdalsvatnet. The facility consists of 87 tanks with different dimensions and water flow capacities as shown in table 5.6. The tanks are spread out across five sections, which all share the same water supply capacity. An overview of the facility layout is shown in figure 5.15. These five section are used for specific purposes:

- The hatchery is used to hatch fertilized eggs. Equipment to both cool and heat water is installed, enabling a strict control of the environment.

- The start feeding section is indoors and contains fry which have recently left the hatchery. The fish in this section typically weigh 0.2-20 grams. Equipment to heat the water to 14°C is installed.
- The 0-yearing section is indoors and typically contains fish weighing 15-80 grams. The vaccination machines are also located in this section.
- The tent section is outdoors and typically contains fish weighing 15-80 grams.
- The 15 meter section is outdoors and solely used for the largest fish weighing 50+ grams.

The water flow throughout the facility is constrained by both the technical equipment and the quotas assigned by the Norwegian Water Resources and Energy Directorate. The technical equipment is the limiting factor at the Slørdal facility (Marine Harvest, 2013a).

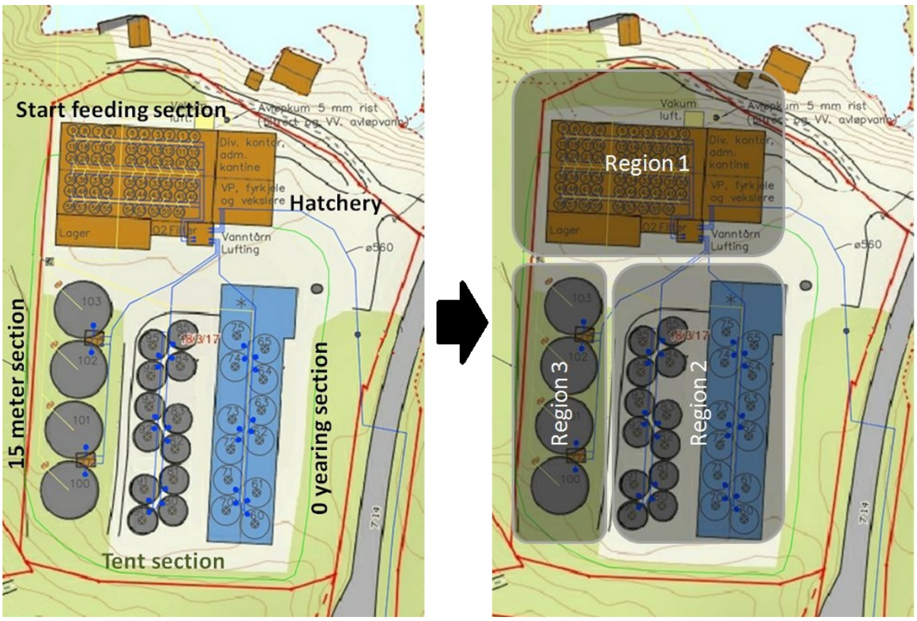

Figure 5.15: A map showing how the Marine Harvest freshwater facility at Slørdal has been aggregated.

Section	**Number of tanks**	**Tank size** m^3	**Flow limit** $\frac{\text{liter}}{\text{minute}}$	**Heating limit** $\frac{\text{liter}}{\text{minute}}$	**Heating temp.** °C
Start feeding	59	9	10 000	5000	14
0-yearing	12	110	11 000	N/A	N/A
Tent	12	65	11 000	N/A	N/A
15 meter	4	720	28 000	N/A	N/A

Table 5.6: A summary of the technical specifications of the Slørdal facility.

5.5.1.2 Facility aggregation

The facility has been aggregated into regions and tanks in order to reduce the problem size. The aggregation was done in the following manner:

- The hatchery is disregarded in the optimization model, as discussed in section 3.3.1.
- The start feeding section is divided into two equally sized tanks. This is done in order to enable heating all of the water flow in one tank to the maximum temperature. In order to keep the energy costs of at a realistic level, one of the two tanks operate at a the minimum allowed water flow of 1500 $\frac{liter}{minute}$. This choice is a result of the feedback sessions described in section 5.7. The aggregated start feeding section is referred to as **Region 1**.
- The tent section and the 0-yearing section is aggregated into **Region 2** with one tank representing each section.
- The 15-meter section is referred to as **Region 3** and contains two equally sized tanks.

The aggregation was done in collaboration with Marine Harvest and is shown in figure 5.15. Table 5.7 defines the maximum fraction of warm water in each tank (L^{+}_{tir}), the maximum fraction of warm water in each region (TL^{+}_{r}) , the size of each tank (L^{size}_{ir}) and the amount of water flow to each tank (W_{ir}). Since water cannot be cooled in any region, L^{-}_{tir} and TL^{-}_{r} are both set to zero.

Region name	**Tank name**	$\mathbf{L^{size}_{ir}}$	$\mathbf{W_{ir}}$	$\mathbf{L^{+}_{tir}}$	$\mathbf{TL^{+}_{r}}$	$\mathbf{G^{max}}$
Region 1	Tank 1	265	1500	1	0,5	14
Region 1	Tank 2	265	5000	1	0,5	14
Region 2	Tank 1	1320	11 000	0	0	N/A
Region 2	Tank 2	780	11 000	0	0	N/A
Region 3	Tank 1	1440	14 000	0	0	N/A
Region 3	Tank 2	1440	14 000	0	0	N/A

Table 5.7: The parameters assigned to each aggregated tank.

5.5.2 Valid facility operations

The valid facility operations parameters are used to reduce the number of variables and constraints as described in section 5.1. There are three parameters types which were set in collaboration with Marine Harvest (2013a):

1. $A^{\text{fish allowed}}_{rf}$ denotes if fish weight class f is allowed in region r. The values are shown in table 5.8.
2. A^{heating}_{r} denotes if water heating is available in region r. The values are shown in table 5.9.
3. $A^{\text{movement}}_{r\hat{r}}$ denotes if it is allowed to move fish between region r and $\hat{r}$. The values are shown in table 5.10.

Fish weigh in gram	Region 1	Region 2	Region 3
0.2	T	F	F
0.5	T	F	F
1	T	F	F
2	T	F	F
4	T	F	F
7	T	F	F
10	T	F	F
15	T	T	F
20	T	T	F
25	F	T	F
30	F	T	F
40	F	T	F
50	F	T	T
60	F	T	T
70	F	T	T
80	F	T	T
90	F	F	T
100	F	F	T
120	F	F	T
140	F	F	T
175	F	F	T
200	F	F	T
300	F	F	T

Table 5.8: The values of $A_{rf}^{\text{fish allowed}}$ (T = True, F= False).

Region 1	Region 2	Region 3
T	F	F

Table 5.9: The values of A_{r}^{heating} (T = True, F= False).

	From Region 1	From Region 2	From Region 3
To Region 1	T	F	F
To Region 2	T	T	F
To Region 3	F	T	T

Table 5.10: The values of $A_{r\hat{r}}^{\text{movement}}$ (T = True, F= False). The fish can move freely between the tanks within each region.

5.5.3 Maximum allowed biomass

The maximum allowed biomass in the facility is restricted by the fish density and the water flow. This section parametrizes the constraints defined in section 4.3.7.

5.5.3.1 Minimum water flow

Let the stochastic parameter κ^t_{irs} denote the flow requirement in tank i at region r in scenario s at time t. Two assumptions have to be made to pre-calculate κ^t_{irs}. First, choose a weight class $f(r)$ to represent the typical growth in each region r. Second, assume the daily growth, $g^{t,\, daily}_{f(r)irs}$, to be a parameter. The growth is pre-calculated using the stochastic water intake temperature τ^t_s in regions r *without* water heating and the maximum water temperature G^{max}_r in regions *with* heating. Equation (5.21) and (5.22) defines κ^t_{irs} using these assumptions.

$$\kappa^t_{irs} = VCO2_{ir} \frac{100}{7} \frac{g^{t,\, daily}_{f(r)irs}}{V_{f(r)}} \qquad \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{5.21}$$

$$g^{t,\, daily}_{f(r)irs} = \begin{cases} A_{f(r)} + B_{f(r)} G^{max}_r & \text{if } A^{\text{heating}}_r = \text{true} \\ A_{f(r)} + B_{f(r)} \tau_{ts} & \text{else} \end{cases} \qquad \forall i \in \mathcal{I}, r \in \mathcal{R}, s \in \mathcal{S}, t \in \mathcal{T} \tag{5.22}$$

Table 5.11 and 5.12 has been determined in collaboration with Marine Harvest (2013a) and contains the data necessary to calculate κ^t_{irs}. This approach may over-estimate or underestimate the actual flow requirement depending on the reference weight class $f(r)$ chosen. The flow requirement in region 1 is slightly overestimated because the maximum temperature is assumed.

Region name	**Interception, $\mathbf{A_{f(r)}}$**	**Slope, $\mathbf{B_{f(r)}}$**	**Reference class, $f(r)$**	**Temperature**
Region 1	-0.034	0.221	7 gram	Fixed at 14°C
Region 2	-0.047	0.125	40 gram	Stochastic
Region 3	-0.054	0.073	120 gram	Stochastic

Table 5.11: The parameters used to calculate the flow factors for each region.

$\mathbf{VCO2_{ir}}$	**Tank 1**	**Tank 2**
Region 1	0.7	0.7
Region 2	0.7	1
Region 3	0.7	0.7

Table 5.12: The reduction, $VCO2_{ir}$, in flow factor requirements due to CO_2 ventilators. Tank 2 in Region 2 (tent section) does not have CO_2-ventilation equipment installed.

5.5.3.2 Maximum fish density

The density restriction, RD_{ir}, in tank i at region r is governed by the internal procedures at Marine Harvest and depends on the weight range of the fish in the tank (Marine Harvest, 2013b). These parameters are shown in table 5.8.

$\mathbf{RD_{ir}}$	**Tank 1**	**Tank 2**
Region 1	35	35
Region 2	50	50
Region 3	70	70

Table 5.13: Density restriction, RD_{ir}, in $\frac{Kg}{m^3}$ for the various tanks at the Slørdal facility.

5.5.4 Fish weight and number classes

Table 5.14 shows the fish classes, $f \in \mathcal{F}$, and smolt classes, $c \in \mathcal{C} \subset \mathcal{F}$, used in the model.

Fish class, f	**Weight in grams,** V_f	**Smolt class,** c
1	0.2	No
2	0.5	No
3	1.0	No
4	2.0	No
5	4.0	No
6	7.0	No
7	10.0	No
8	15.0	No
9	20.0	No
10	25.0	No
11	30.0	No
12	40.0	No
13	50.0	No
14	60.0	No
15	70.0	Yes
16	80.0	Yes
17	90.0	Yes
18	100.0	Yes
19	120.0	Yes
20	140.0	Yes
21	175.0	Yes
22	200.0	Yes
23 (sink class)	300.0	Yes

Table 5.14: The fish and smolt weight classes used.

23 weight classes between 0.2 grams and 300 grams have been used. 0.2 grams is the lowest fry weight and 200 grams is the largest weight which is commonly delivered. 300 grams is a "sink state" which prevents fish from weighting too much, and fish stop growing when reaching this class. The granularity of the classes was chosen to preserve growth accuracy while achieving model tractability. All weight classes below 100 grams are defined in the Skretting growth rate model described in section 5.3. We determined in collaboration with Marine Harvest (2013a) to use the growth rate for 100 grams fish for the larger fish classes. All fish larger than 70 grams is assumed to be smolt.

Table 5.15 shows the number classes, $o \in \mathcal{O}$, used. Only two have been included since all valid fish quantities are a linear combination of these two classes. We found that adding more number classes did not change the behavior of the linearized growth model, only increased the solution time.

Number class, o	**Value of number class, N_o**
1	0
2	10000

Table 5.15: The number classes used.

5.5.5 Smolt delivery

The model formulation in chapter 4 defines a lower, U_c^{lower}, and an upper, U_c^{upper}, smolt class which can be used to satisfy an order for class c. (Marine Harvest, 2013a) considers smolt 20% below and 30% above the ordered weight, V_c, to be acceptable. Since the smolt weight is discretized, we have chosen therefore to use the *nearest* weights to these percentage limits. 70 grams is the smallest smolt that can be delivered. The upper and lower limits for a given order are defined in table 5.16.

$\mathbf{U_c^{lower}}$	**Order weight class, c**	$\mathbf{U_c^{upper}}$
70 gram	**70 gram**	90 gram
70 gram	**80 gram**	100 gram
70 gram	**90 gram**	120 gram
80 gram	**100 gram**	120 gram
100 gram	**120 gram**	175 gram
120 gram	**140 gram**	175 gram
140 gram	**175 gram**	200 gram
175 gram	**200 gram**	300 gram

Table 5.16: The upper and lower weight class that is accepted at a given order weight class.

5.5.6 Cost data

Costs are approximated based on data provided by Marine Harvest and consist of feed cost, oxygen cost, egg cost, vaccination cost, insurance cost and energy cost.

5.5.6.1 Feed and oxygen cost

Table 4.2.3 shows the feed factor (F^{feed}), oxygen factor (F^{oxygen}), feed price (C^{feed}) and oxygen price (C^{oxygen}) provided by Marine Harvest (2013a).

	Factor, $\frac{\text{kg oxygen or feed}}{\text{kg growth}}$	**Price, $\frac{\text{NOK}}{\text{kg}}$**
Oxygen	0.75	2
Feed	1	10

Table 5.17: The parameters used to calculate the feed and oxygen cost.

5.5.6.2 Cost of egg, vaccination and insurance

Table 5.18 shows the cost of insurance, fertilized eggs and vaccination per fry deployed. These sum up to C^{EVI} which is used in the model formulation. The vaccination cost per fry is calculated by taking the actual vaccination cost per unit (1.48 NOK), and correcting for the expected mortality and culling with the current production praxis (30%). This was done in collaboration with Marine Harvest (2013a).

	Price, $\frac{\text{NOK}}{\text{fry}}$
Vaccination	1.14
Egg	0.46
Insurance	0.03
Total ($\mathbf{C^{EVI}}$)	**1.63**

Table 5.18: The parameters used to calculate fry deployment cost.

5.5.6.3 Energy cost

The model formulation uses the parameter E_{irt} to calculate the energy costs, as described in section 4.2.1. This parameter denotes the cost per temperature difference in the water flow. Using basic thermodynamics, as described by Tipler and Mosca (2008), this parameter can be calculated as shown in table 5.19.

P_t denotes the power price at time t. The average monthly power price from NordPoolSpot (2013) the last seven years in Sør-Trøndelag has been used to calculate P_t. For instance, if the month of period t is January, P_t is the average power price in January the last 7 years. The average monthly power prices are shown in figure 5.16.

Description	**Denomination**	**Value**
Heat pump efficiency	-	20 %
Mass of 1 liter water	$\frac{kg}{liter}$	* 1
Heat capacity	$\frac{J}{kg*K}$	* 4185
Joule per kWh	$\frac{kWh}{j}$	* $\frac{1}{3{,}6*10^6}$
Minutes per month	minute	* 43200
Power price	$\frac{NOK}{kWh}$	* P_t
Product	$\frac{\textbf{NOK * minute}}{\textbf{K * liter}}$	**= E_{irt}**

Table 5.19: How the energy constant E_{irt} has been calculated.

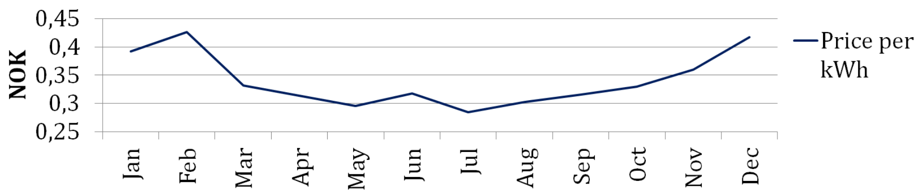

Figure 5.16: The average monthly power price, P_t, used to calculate the energy cost (NordPoolSpot, 2013)

Hatchery energy cost The energy cost of having a constant temperature in the hatchery, $\iota_{ts}^{hatching}$, is given in equation (5.23).

$$\iota_{ts}^{hatching} = \frac{\left(|8 - \tau_s^{t-1}|E_{t-1} + |8 - \tau_s^{t-2}|E_{t-2}\right)600}{2\ 000\ 000} \qquad \forall t \in \mathcal{T} s \in \mathcal{S} \tag{5.23}$$

The eggs spend two months in the hatchery before being deployed as fry. The energy cost during this period is determined by three factors. First, the flow in the hatchery which is constant at 600 $\frac{liter}{minute}$. Second, the cost per flow, $|8 - \tau_s^{t-1}|E_{t-1}$ and $|8 - \tau_s^{t-2}|E_{t-2}$, to maintain a constant temperature of 8°C in the two previous months. Third, divided by the average number of fry in the hatchery (2 000 000). This amount was determined based on the historical egg amounts kept in the hatchery (Marine Harvest, 2013a).

E_t denotes the same as E_{irt} in the mathematical formulation, but is not indexed on region or tank since the hatchery is not modeled explicitly. E_t and E_{irt} are identical for the Slørdal facility.

5.6 System design of implementation

The model runs were performed on a computational cluster owned by the Department of Industrial Economics and Technology Management at NTNU. The technical information about hardware and software used is given in table 5.20.

Hardware - Solstorm		
Component	**information**	
24 identical computers: - Model - Memory - CPU	 HP BL686 G7 128 Gb RAM, 300 GB SAS 15000rpm 4 x AMD Opteron 6274 2,2 GHz	
Software		
Program	**Distributor**	**Used for**
Rocks Cluster Distribution - Unix OS	National Partnership for Advanced Computational Infrastructure	Operating system at Solstorm
Xpress Mosel v.7.5.0	FICO inc.	Optimization
OxMetrics	Timberlake Consultants	Scenario generation
ScenGen	Michael Kaut	Scenario generation
Excel 2010	Microsoft	Storing and processing of data

Table 5.20: Information about tools used to implement the model.

The system design consists of several components as shown in figure 5.17. OxMetrics and ScenGen were used for scenario generation and gave input to the preprocessing. Excel was used for data storage and to perform the preprocessing. The memory requirement exceeded the capabilities of a standard PC and the computational cluster Solstorm was therefore used. Solstorm does not support Excel and the data were translated over into text tables before it was uploaded. Excel was used to process the model results.

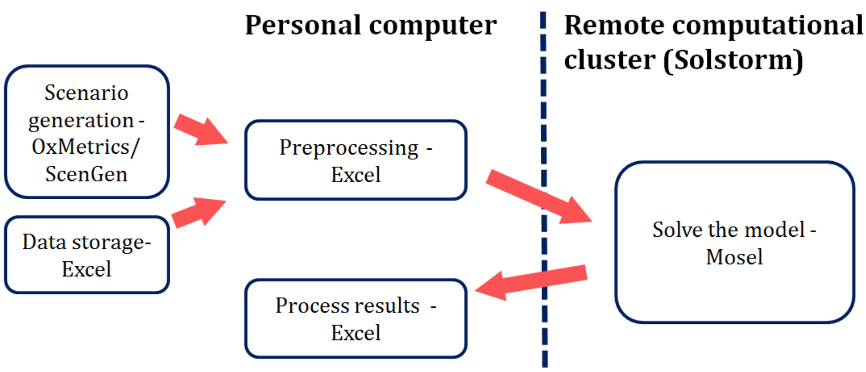

Figure 5.17: Flow chart of the system design.

5.7 Testing

In order to ensure model quality, an accurate parameterization and a correct implementation black box inspired testing was used. It ensures that the model acts in a desired manner by looking at model output in relation to input (Braude, 2000). Three types of black box tests were performed:

1. **Constraint tests** consisted of several data sets that were designed to provoke a certain behavior in the model. These tests check the basic features of the model, e.g. the growth model.

2. **Realism tests** were used to investigate if the model results were corresponding to the real world. Simplifications and assumptions have been made both in the formulation and parameterization of the model. These tests consisted of investigating and adjusting both such that the model acts in a realistic and believable manner.

3. **Stability testing** may be viewed upon as robustness tests of the optimization model regarding how it is affected by the scenario generation procedure.

5.7.1 Constraint testing

A baseline data set, which contained data similar to the parameterization described in this chapter, was used as a starting point for all the test cases. The tests indicates that the growth model, the movement of fish, the biomass restrictions and the scenario tree structure were correctly implemented. Appendix B gives an overview of the test performed and their results.

5.7.2 Realism testing

Going from model to executed production plan is complex when dealing with biological production. This is because there are many intangible aspects to consider when dealing with living animals. Traditional realism testing approached such as locking decision variables and inspecting output are not able to capture many of these aspects. Feedback sessions with biologists at Marine Harvest were therefore conducted in order to test the realism of the model. The feedback sessions consisted of reviewing the results from case studies together with these experts. Both the model formulation and parameterization were fine tuned iteratively until the model produced results that the biologist viewed as possible valid answers.

The feedback sessions were conducted with Anders Fjellheim and Ole Christian Norvik from Marine Harvest's regional main office in Trondheim on the:

- 22^{nd} of April
- 2^{nd} of May (only with Ole Christian Norvik)
- 8^{nd} of May
- 21^{st} of May

Several important model decisions were made as a consequence of these sessions:

- The energy costs were deemed too high given the tank temperature profile. This was addressed by lowering the constant flow in one tank in region 1, since the praxis is to reduce the flow when there is little biomass in the tank.
- The choice of cost drivers and split between variable and fixed costs was refined.
- The fish weight classes used in the growth model were determined. Also, the reference weight classes used to set the MAB defined by tank flow was determined.

Also, the feedback sessions discovered some limitations regarding the data set used:

- **Too optimistic growth**: The biologist have trouble maintaining the growth rate given by the Skretting model. This is due to starvation and haltered growth in relation to vaccination, moving of fish and sorting. However, the Skretting growth rate model is the best one available and was therefore used. The growth rate used in this model is therefore optimistic compared to experienced growth rate.
- **Different production scheme**: At Marine Harvest the freshwater facilities follows a biomass maximization scheme, thus the biologist had only this as a reference point when evaluation the feasibility of the model's results. The biggest difference was that Marine Harvest typically heat water as much as possible while the model uses heating more restrictive. Through adjusting the energy cost, we were not able to produce a result that replicated the experience use of heating. This is also tightly related to the difference in growth rates in the Skretting model and growth rate experienced by the biologist.

5.7.3 Stability testing

A model is said to be *stable* if several different scenario trees, generated using the same method and input, yields approximately the same optimal solution value in the optimization model (Kaut and Wallace, 2007). Thus, a stability test may be viewed upon as a robustness test of the model regarding how it is affected by the scenario generation method. As a glossary note, the term *model* denotes the system including both the optimization model and scenario generation procedure.

There are two main types of stability. These are called *in-sample stability* and *out-of-sample stability*, and addresses quite different aspects of the model. An actual test for in-sample stability is performed in section 6.1.8, since stability tests are done on the actual model used in the case study. Out-of-sample stability testing is omitted due to the computational platform available. The reasoning behind this and how such a test can be performed is discussed in appendix C.

5.7.3.1 In-sample stability

An in-sample stability test investigates how the discretization of the random variables in the scenario tree affects the objective function value. It is a test of the model's robustness towards the discretization procedure. This can be viewed as a test of the internal consistency of the model (King and Wallace, 2012). The test can be performed by solving the model on different scenario sets with the same statistical properties (e.g. mean and variance). Each of these runs should produce approximately the same optimal objective function value.

King and Wallace (2012) describes two approaches to test the in-sample stability. The choice of these depends on whether or not the scenario generation procedure is random or deterministic. A scenario generation procedure is said to be random if it *does not* produce the same scenario tree with exactly the same input data. Conversely, it is deterministic if it produces the same scenario tree.

Although the scenario generation procedure described in section 5.2 uses random sampling for the residual distribution (Høyland et al., 2003), it will converge to the same discretization. It should therefore be considered deterministic. The argument is that the residual distribution is assumed to follow a univariate normal distribution. ScenGen discretizes this distribution using scenarios with the same probability. Thus, there is only one way to discretize the distribution such that the moments are matched. Even though random sampling is used, it will converge onto the same discretization. Experimenting with ScenGen have yielded results consistent with this argument.

King and Wallace (2012) describes that an in-sample stability test on a deterministic scenario generation procedure can be conducted by creating scenario trees of different sizes and comparing the objective function values from model runs with these trees. These trees should have the same statistical properties and the model is in-sample stable if the objective functions are approximately equal to each other. We have used six scenario trees of sizes ranging from 4 through 16 to test the in-sample stability. The results are presented in section 6.1.8.

Chapter 6

Results, discussion and sensitivity analyses

This chapter contains the results from two case studies based on the Slørdal freshwater facility. The first case study considers a typical 24 month production plan and is used to investigate the current production. Typical orders for Marine Harvest on a two year perspective were used as input. The optimization model yields a minimum cost production plan which can satisfy these orders. The setup and results for this case study is discussed in section 6.1. Key parameters are also highlighted and investigated further in four sensitivity analyses in section 6.2.

The second case study considers the introduction of a new product which is a smolt of 500 grams. Smolt this large are not currently being produced at the Slørdal facility, but Marine Harvest is considering it. Section 6.3 discuss how it can be produced in a cost efficient manner.

6.1 Case study 1: Typical 24 month production

Marine Harvest does not currently use cost minimization at the Slørdal facility. Their production plans aim at produce as much biomass as possible and have a high facility utilization. Elements of the optimized production plan presented in this case study deviates from this praxis, as it is based on a cost minimizing model. The most significant deviations impacting the production costs are reduced water heating and fewer fish culled.

6.1.1 Case study setup

The planning horizon of the case study is two years. This period is chosen for two reasons. First, deliveries are planned up to 18 months ahead (Marine Harvest, 2013a). Second, equality between the initial biomass and end of horizon biomass can be enforced realistically.

The period between 1. February 2011 and 31. January 2013 is used as a reference when determining smolt orders, temperature profiles, initial fish inventory and fixed costs.

Smolt orders: The orders placed by the seawater division are given in figure 6.1. These are based on the orders at the Slørdal facility during a typical year and were provided by Marine Harvest (2013a).

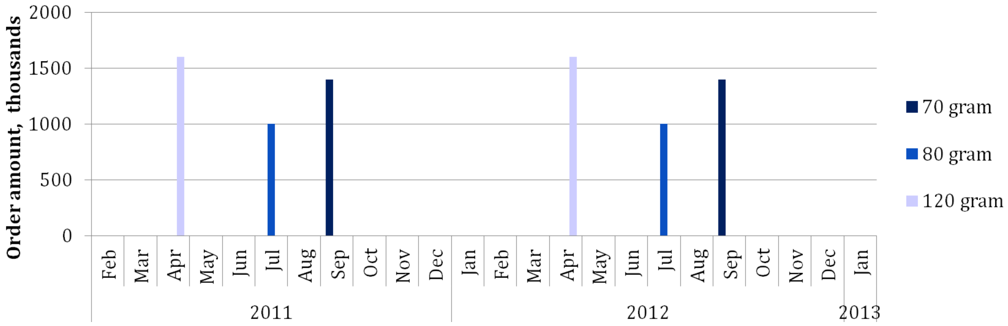

Figure 6.1: The smolt orders placed by the seawater division.

Initial and end of horizon inventory: The initial model inventory levels are based on the actual inventory of February 2011. Three changes were made to the actual levels to ensure both realism and model feasibility. First, the initial amount of fry has been set equal to zero to let the model decide for itself if it is optimal to deploy. Second, fish larger than 140 grams were not included since there were no orders for fish this large. Third, enough fish were added such that the first two orders could be met since there is not sufficient time for the model to produce fish for these orders. Figure 6.2 shows the initial inventory used.

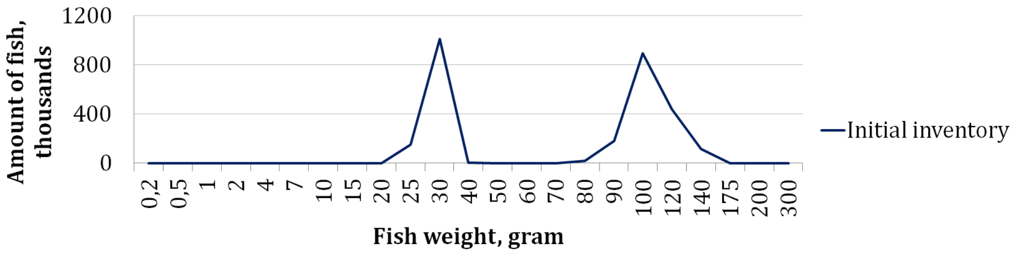

Figure 6.2: The initial inventory and required end of horizon inventory used.

Fixed costs: Based on accounting data for 2012 provided by Marine Harvest the total fixed cost were set to 13.6 million NOK per year (Marine Harvest, 2013a).

Temperature scenario tree: The temperature scenario tree was generated with the seasonal AR(1) model estimated in section 5.2.3. The actual water intake temperature in February 2011 was used as the initial condition. Figure 6.3 shows the forecasted values along with the scenario split.

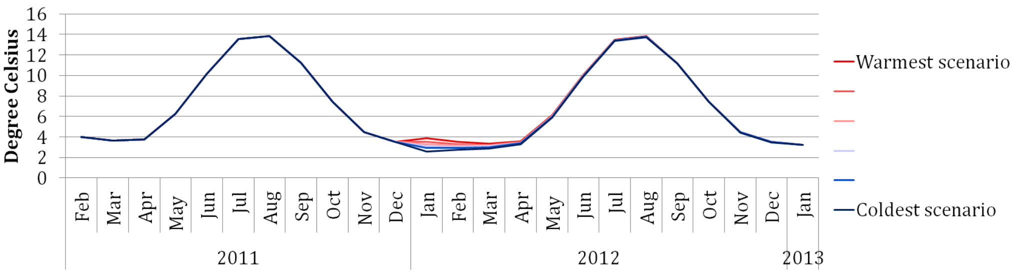

Figure 6.3: The temperature scenarios used.

Only one scenario split has been included in order to keep the computational complexity tractable. That is, a two-stage scenario tree has been used. The consequence of this is that the different scenarios converge together again after the split, as shown in figure 6.4.

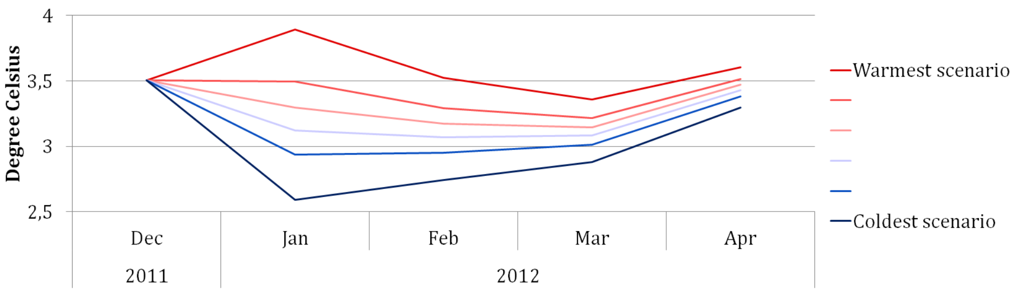

Figure 6.4: The scenario split and subsequent convergence in the temperature scenarios used.

Gap: Branch-and-Bound (B&B) is a commonly employed solution method for Mixed-Integer Problems (MIP), and a version has been used to solve the models in this thesis. This method keeps track of the best integer solution found so far, which provides an upper bound in a minimization problem. The LP-relaxation solutions found provide a lower bound on the objective function value. A commonly used metric for the quality of a B&B solution is the *gap*, which is defined as:

$$\text{Gap} = 1 - \frac{\text{Best lower bound}}{\text{Best upper bound}} \tag{6.1}$$

As the gap decreases, the range wherein the optimal solution lies becomes smaller. When the gap is zero, the best integer solution found is the optimal integer solution. The lower the gap requirement on an integer solution, the longer time it takes to solve the problem.

A 10% gap requirement has been used in the case study. This number is chosen because tests have shown that the model typically converge around this gap. That is, running the model several days extra does not result in a significant gap reduction.

Solution time: Solution time increases significantly with the number of scenarios as shown in figure 6.5. This is because the SOS2 are implemented using binary variables, and the number of binary variables increase the solution time exponentially. We therefore chose to use six scenarios. This yields a good trade-off between tractable solution time and granularity in the scenarios.

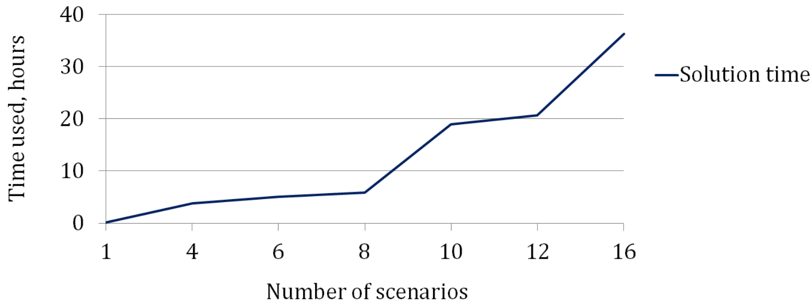

Figure 6.5: The time used to reach a gap lower than 10% with various number of scenarios.

6.1.2 Expected costs

Cost per smolt is a commonly used performance measurement within the salmon industry (Marine Harvest, 2013a). It is therefore used in the comparison between the modeled costs and costs from the actual production. The costs from the actual production have been adjusted to match the model production amount of 8 million smolt, and comparing the cost per smolt and total cost are therefore equivalent. The model cost structure compared to actual cost is shown in figure 6.6.

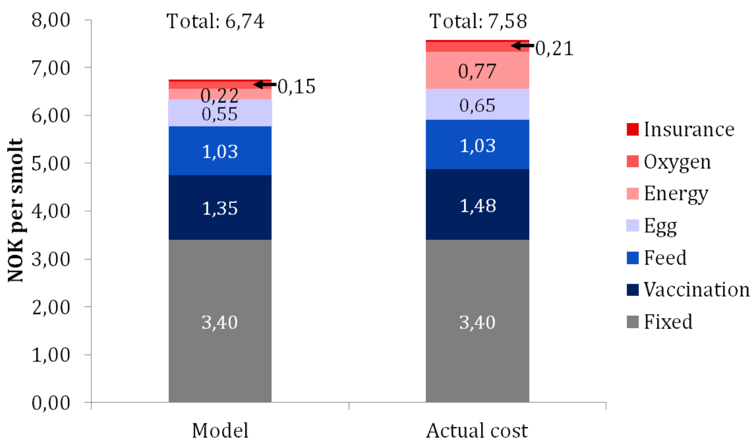

Figure 6.6: A breakdown of the total costs per smolt delivered.

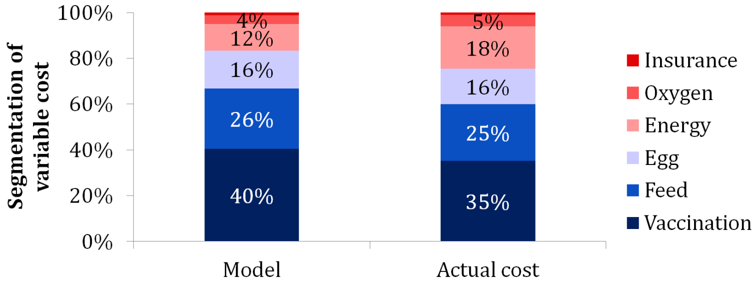

Figure 6.7: A percentage breakdown of the variable costs per smolt.

The model cost per smolt is 6.74 NOK while the actual cost was 7.58 NOK. This is a difference of 11%. Since both have the same fixed costs, the model variable costs are 0.84 NOK (20%) lower than the actual variable costs. This is a significant difference which is caused mainly by three key factors: energy, vaccination and egg costs. The remainder of this section investigates what decisions are driving this variable cost difference:

- **Energy cost**: The model energy cost is 71% lower than the actual cost. This could be explained by four interesting aspects of the model results. First, the tanks in region 1 in the model is running at a total flow rate of 6500 l/min rather than 10 000 l/min which is the maximum capacity. This reduced the energy cost and was done on the basis of the feedback session described in section 5.7.2. The impact of the flow through region 1 in the model is investigated further in section 6.2.3.
  Second, the temperature in region 1, shown in figure 6.8, is generally lower than what is common at the Slørdal facility (Marine Harvest, 2013a). This was discussed in section 5.7.2, and the conclusion was that the Skretting model overestimates the growth and the facility can therefore operate at a slightly lower temperature.
  Third, during the second year barely any water is heated. This is related to the chosen fry deployment time and is discussed in the section 6.1.3.
  Lastly, average power prices were used. Since electricity markets are very volatile, the difference could be attributed this modeling choice.
- **Vaccination cost**: This cost is 9% lower in the model than the actual one. This effect could appear because the amount of fry deployed is the vaccination cost driver and the unit cost per fry has not been properly approximated. Two alternative cost drivers, amount reached vaccination weight and delivery amount, are discussed in section 3.3.2. The cost driver impact is investigated in the sensitivity analysis in section 6.2.4.
- **Egg cost**: The egg cost is 15% lower in the model than the actual cost. This appears to be caused by 11% fewer fry being deployed in the model than in the actual production. The difference in fry amount is mainly due to fewer fish culled.
- **Feed and oxygen cost**: The feed cost is the same in the model as the actual

cost, while the oxygen cost is 29% lower. This is interesting, because they both have growth as the cost driver. This appears to be an issue with the feed and oxygen cost per unit of growth, since the amount of fry deployed in the model is lower. Both the feeding factor and oxygen factor are affected by the temperature and fish weight, but are assumed to be constant. This aspect could explain some of the difference.

- **Insurance cost**: Both the absolute size and difference of this cost is negligible.

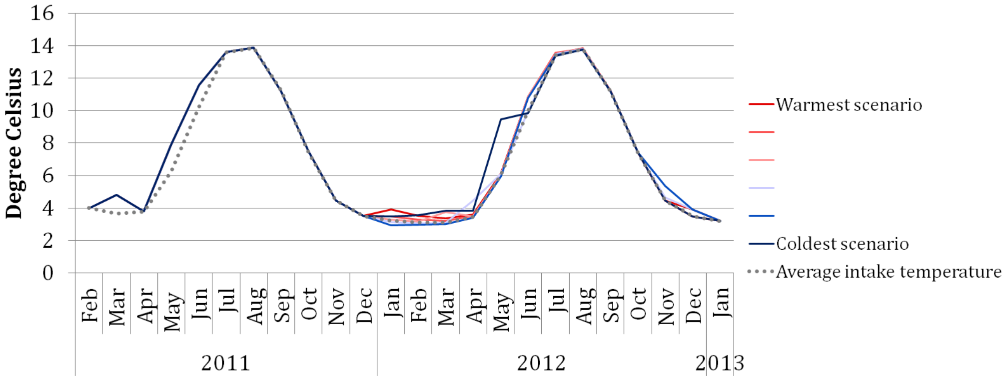

Figure 6.8: The temperature in region 1 after heating in each scenario (solid lines) along with the *average* intake temperature (dotted line).

6.1.3 Fry deployment

The fry deployment is shown figure 6.9. The optimization model splits the fry deployment with 75% in 2011 and 25% in 2012. The batch intended to satisfy the delivery in September is deployed in October rather than February. With an extra 4 months in the facility, it can grow to the desired weight with far less water heating as shown in figure 6.8. This could explain part of the difference between the actual and model energy cost. This deviates from the current best praxis at Marine Harvest, and could be a valuable business insight.

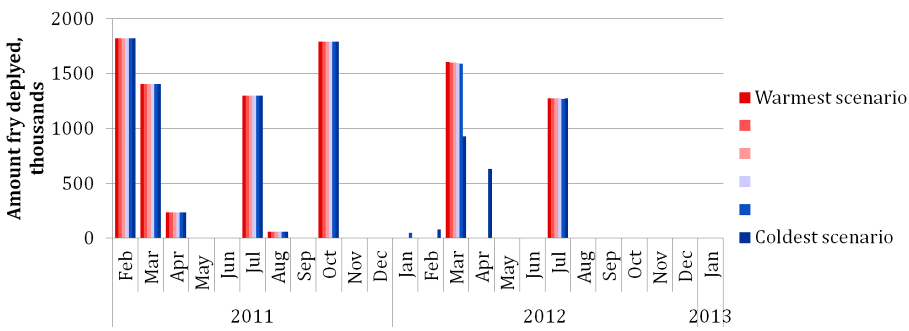

Figure 6.9: Fry deployment in the various scenarios.

6.1.4 Biomass at the facility

Figure 6.10 shows that the maximum allowed biomass (MAB) for the facility as a whole is not a binding constraint. However, the biomass is at a maximum in some of the regions during the period. This is shown in figures 6.11 - 6.13.

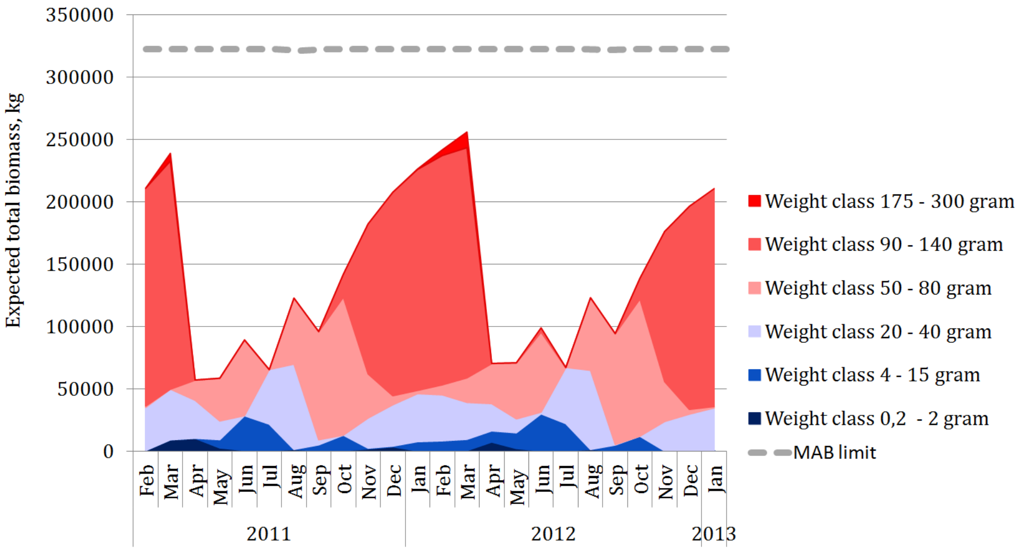

Figure 6.10: The *expected* biomass at the facility along with the MAB. The expected biomass is the weighted average biomass for all scenarios.

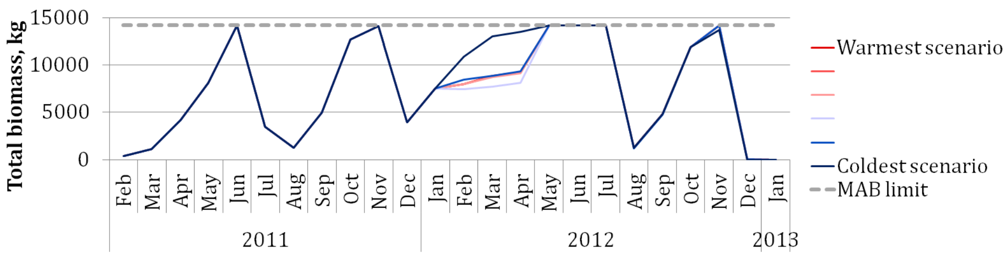

Figure 6.11: The biomass development in **region 1** along with the MAB limit.

As expected, the biomass difference between the various scenarios is significant as temperature is the main driver behind growth. This makes different regions bottlenecks at different points in time. Bottlenecks also impact the chosen fry deployment time. The deployment summary in figure 6.9 shows that in the coldest scenario half of the fry deployment has to be delayed from March until April. Figure 6.11 indicates that this is because the biomass capacity in region 1 is reached.

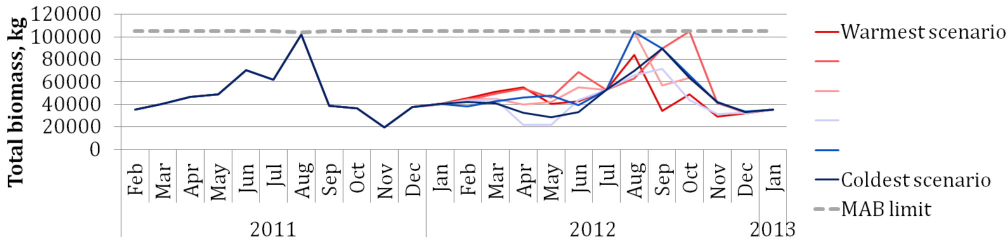

Figure 6.12: The biomass development in **region 2** along with the MAB limit.

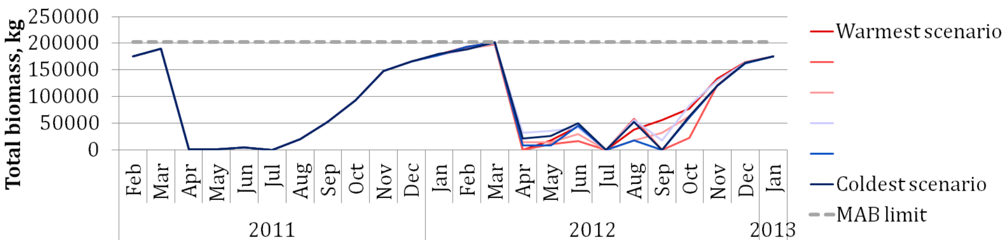

Figure 6.13: The biomass development in **region 3** along with the MAB limit.

An interesting observation is that it is always the density restriction which is binding on the biomass at the facility, never the flow restriction. These restrictions are defined in 4.3.7. A sensitivity analysis in section 6.2.1 investigates further how the tank size in each region affects the optimal production plan.

6.1.5 Mortality and culling

Fish at a freshwater facility can die due to either natural causes or culling. Natural causes are for instance diseases and culling is the act of shaping the fish population by selective slaughter.

The mortality due to natural causes is shown in figure 6.14. The average cumulative mortality rate in the model was 10.9%, which is 0.5% above the average for the Slørdal facility and 1.5% above the average for Marine Harvest during 2009-2011. The reason this could be that the fish spend more time in the facility due to a generally lower tank temperature and that the cumulative mortality increases with the time spent.

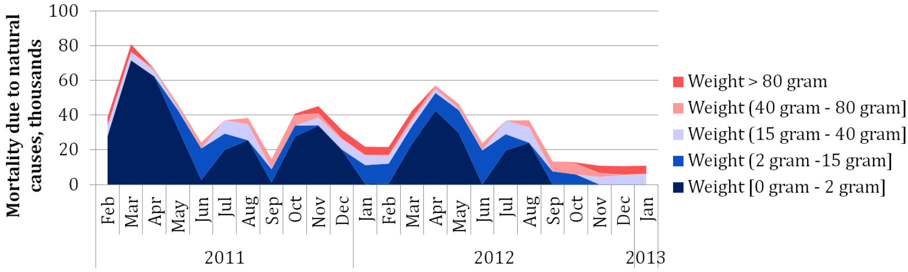

Figure 6.14: The amount of fish which has died of natural causes.

The culling rate in the model is 4.9%, which is only a third the actual rate of ∼15.0% at the Slørdal facility (Marine Harvest, 2013a). The difference may be due to the current Marine Harvest praxis of having a safety stock available in case of emergency. Most of the culling happens in May each year, as shown in figure 6.15. However, the average weight is higher in the months where fewer fish are culled as shown in figure 6.16.

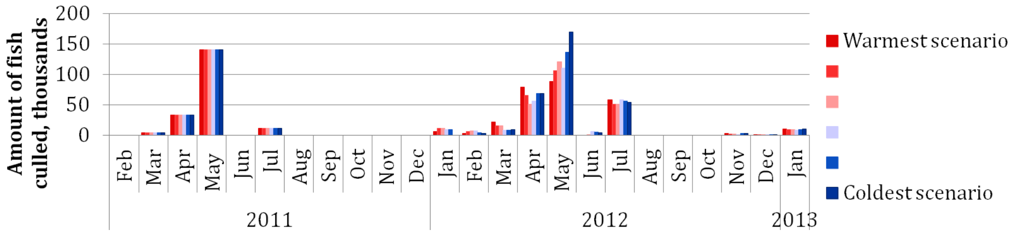

Figure 6.15: The amount fish has been culled in the various scenarios.

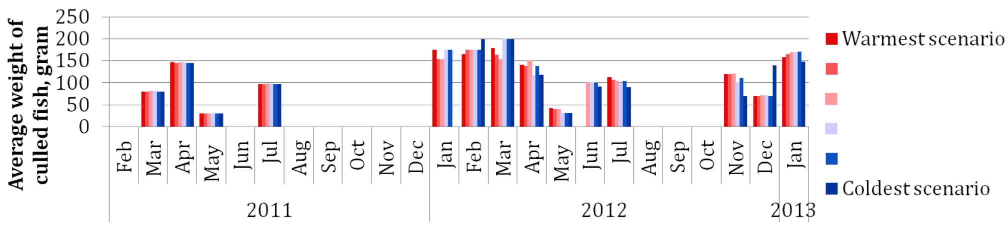

Figure 6.16: The average weight of the fish culled in the various scenarios.

The praxis at Marine Harvest is to cull fish in the lower fish weight classes in order to shape the biomass distribution. The model culling follows this praxis, however it also culls fish in the larger fish weight classes which falls outside the delivering range of the orders. This motivated the sensitivity analysis presented in section 6.2.2. Since the vaccination cost is tied to fry deployment this could also affect the model's culling strategy. This motivated the alternative cost function formulation presented in section 6.2.4.

6.1.6 Segmentation of fish delivered

Figure 6.17 shows that all of the orders except for the July order in 2011 have an average weight above the order weight. However, the delivery weight differs significantly between the various temperature scenarios. In April and July in 2012 there is a 5% difference between the average delivery weight in the warmest and coldest scenario.

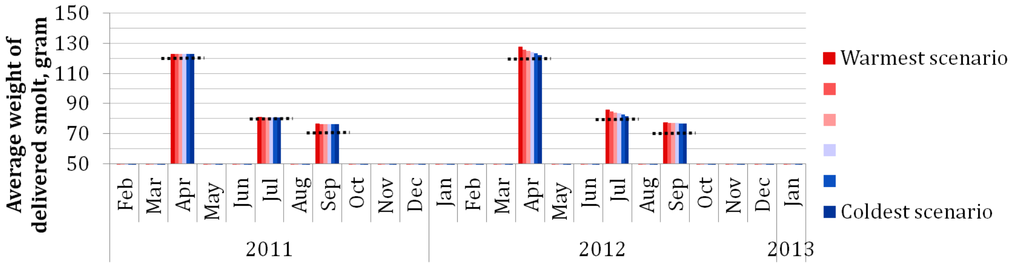

Figure 6.17: Shows the average delivery weight used to satisfy the orders in the various scenarios. The black dotted line denotes the order weight.

6.1.7 Evaluating the model's stochasticity

By introducing stochasticity the model gets access to recourse actions. They allow the model to compensate for either warm or cold temperature scenarios, and include culling, fry deployment and heating. However, the model becomes more computationally intensive to solve because the problem size increases. It is therefore important to evaluate the gain from introducing stochasticity into a model.

Two methods are commonly used in order evaluate stochastic programming models. The **Value of Stochastic Solution (VSS)** measures the objective function value of using a stochastic model versus a model where the expected value of the stochastic parameters is used. It effectively measures the impact that access to recourse actions has on the expected objective function value. The **Expected Value of Perfect Information (EVPI)** measures how much a decision maker would be willing to pay for perfect information about the future of the stochastic parameters. For instance, knowing for certain what the water temperature profile will be the next two years. A more thorough presentation of these concepts is given by Birge and Louveaux (2011).

However, these measures cannot be used to evaluate stochastic MIP models with a non-zero gap. The reason for this is that neither VSS nor EVPI can be negative (Birge and Louveaux, 2011). This is because more information or access to recourse action can only improve the objective value optimization model setting. When comparing models with gap there is a range wherein the optimal objective solution lies and comparing two models at either the upper or lower bound can yield negative VSS or EVPI. For instance, we got an EVPI of -643 117 when using the best upper bound on the objective function values.

Running the model until a gap of 0% was not computationally feasible. We are therefore not able to prove the value of stochasticity in the model using these methods. However, the case study results show little difference in the decisions taken in the various scenarios. This indicates that recourse actions might not have a large value in different temperature scenarios. The scenario tree structure could also explain the small difference in decisions. The temperature scenarios converge together after the scenario split, since only a two-stage tree is used. This means that the temperature difference between the scenarios for the 2 year period seen as a whole is low. A multi-stage scenario tree could be used to test this, but it would be computationally intractable.

6.1.8 In-sample stability

Using the approach described in section 5.7.3 we have conducted an in-sample stability test. A model is in-sample stable if scenario trees of different sizes yield approximately the same objective function value, but the presence of gaps complicates the comparison. We have therefore chosen to compare the different model runs at both the upper and lower bound.

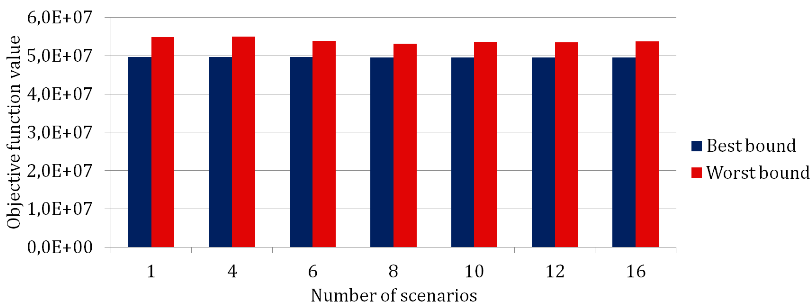

Figure 6.18: The worst and best bound of the problem with various scenarios. Worst bound is based on the best integer solution, while the best bound is the LP-relaxation solution.

Figure 6.18 shows that the worst bound solution is fairly stable with a difference of 3.68% between the minimum and maximum value. The best bound is also stable with a difference of 0.2% between minimum and maximum value. Therefore, the model used for case study 1 appears to be in-sample stable.

6.2 Case study 1: Sensitivity analyses

Sensitivity analyses in mathematical programming investigate how parameter changes impact the solution. The analyses in this section focus on how the objective function changes. This is done by comparing the objective function value of the original run and the run with changed parameters. However, the gaps in the model runs decreases the accuracy of the sensitivity analyses since only the upper and lower bound on the optimal value are known. We have chosen to compare the models at the upper bound, which is the best integer solution found. The gaps may distort the results and the results in this section can only give an indication of the model's behavior.

6.2.1 Investment in +50% tank capacity

The discussion of biomass capacity utilization in 6.1.4 showed that each region acts as a bottleneck at some point during the production. We performed three additional model runs in order to investigate the impact the regional capacity has on the objective function value. In each model run, the tank size in one of the regions was increased by +50%. This emulated an investment in additional tank capacity in each region. Figure 6.19 shows the total costs from these three runs compared to the original model costs.

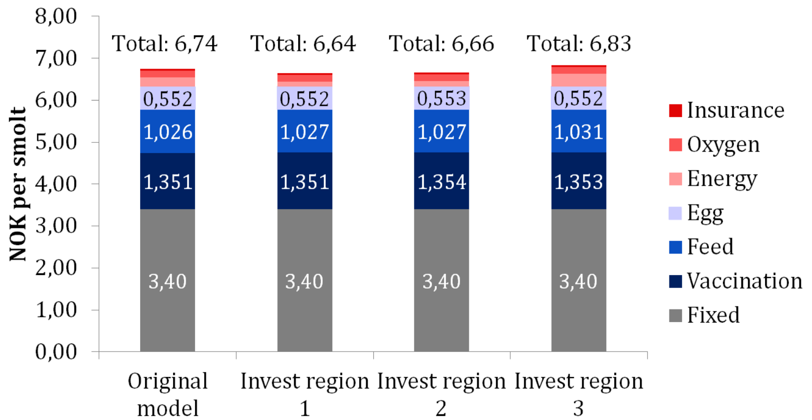

Figure 6.19: The breakdown of cost per smolt with an investment of +50% tank capacity respectively in each region.

Only investments in region 1 and 2 results in a lower cost and the cost reduction is only 1.5% and 1.25% respectively. For investments in region 3 the solution cost is actually 1.3% higher than the original. We believe this is due to the gap in the model runs, since more tank capacity cannot result in a worse solution. This sensitivity analysis indicates that the Slørdal facility is properly scaled to deliver the current orders, since no significant cost benefits arise from increasing the tank capacity.

6.2.2 No upper limit on delivery weight class

The discussion regarding fish culling in section 6.1.5 highlighted that fish in heavy weight classes were being culled. This is because they cannot be used to satisfy orders. To investigate the impact culling has on the solution we changed the delivery weight range such that all weight classes heavier than the order weight can be used to satisfy an order. Figure 6.20 shows the cost breakdown for this run compared to the original run.

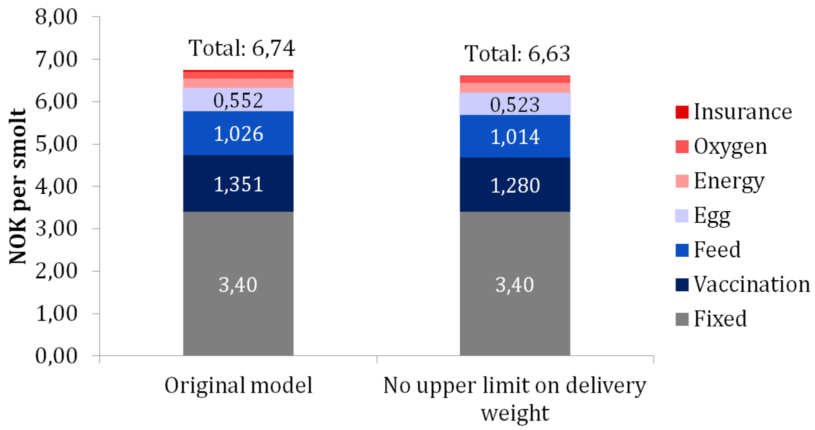

Figure 6.20: The cost breakdown for a run with no upper bound on delivery weight class compared to the original model run.

The cost decrease from removing the upper limit is 1.6% and is mainly driven by decreased vaccination and egg cost, because fewer fry are deployed. In this solution only 0.2% of the fish are culled in comparison to 4.9% in the original solution.

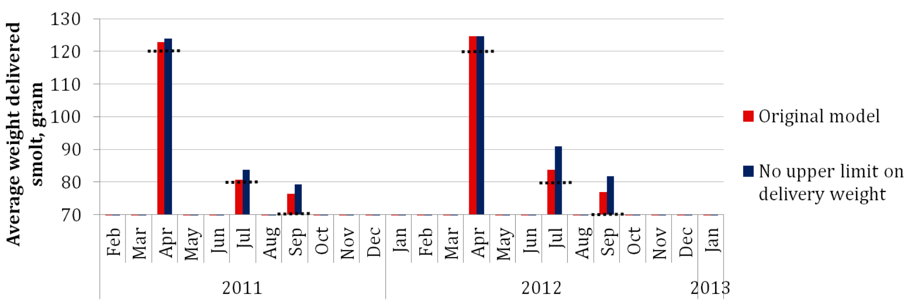

Figure 6.21: The average delivery weight for the two models. The black dotted line denotes the order weight.

Figure 6.21 shows that removing the upper limit on the delivery weight significantly increases the average delivery weight. The months with the largest change is July and September. Larger fish are generally more robust when deployed at sea, and an intangible aspect of removing the upper limit could be higher quality on the smolt delivered.

6.2.3 Increased flow in region 1

Section 6.1.2 highlighted that the model energy cost were lower than the actual one. The water flow parameter in region 1 was hypothesized as one of the causes of this. In order to test this, we performed a model run with the total flow rate increased from 6500 liter/min to the maximum of 10 000 liter/min. Figure 6.22 shows how the costs of the two model runs compare.

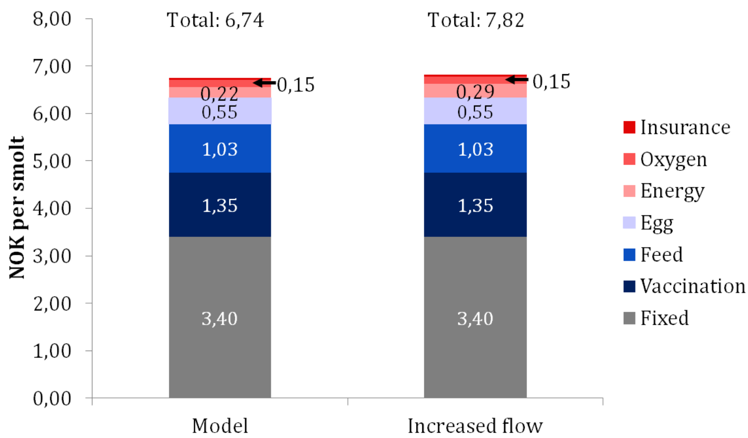

Figure 6.22: The segmented cost per smolt with increased flow compared to the original model cost.

The total cost increased by 1.1%. This cost increase is due to an increased energy cost, since more water has to be heated in order to get the same tank temperatures. The energy costs increased by 31% from 0.224 to 0.293. However, the energy cost increase is relatively small given that the flow has increased by 54%. This could be accounted for by the objective function gaps and that slightly less heating is used.

6.2.4 Vaccination cost driven by tank transitions

The cost analysis in section 6.1.2 showed that the model vaccination cost was lower than the actual one. Due to the layout of the Slørdal freshwater facility we can investigate this by changing the objective function through the approach discussed in section 3.3.2. All fish have to pass through region 2 and are moved there when they reach a weight of 15-20 gram. This is below the normal vaccination weight of 50-70 gram, but it gives the opportunity to cull fish before the vaccination cost incurs. The current fry driven vaccination cost does not enable this. It might therefore prove to be a more accurate representation of the vaccination costs.

The new vaccination cost component, $z_{vaccination}$, is defined in equation(6.2).

$$z_{vaccination} = \sum_{s \in \mathcal{S}} P_s \Big(\sum_{\substack{i \in \mathcal{I} \\ \hat{i} \in \mathcal{I} \\ t \in \mathcal{T} \\ f \in \mathcal{F}}} C^{vaccination} d_{fi1s}^{t\hat{i}2} \Big) \tag{6.2}$$

This cost is driven by the variable, $d_{firs}^{t\hat{i}\hat{r}}$, which denotes the movement of fish in class f from tank t in region $r = 1$ to tank $\hat{t}$ in region $\hat{r} = 2$ in scenario s at time t. $C^{vaccination}$ is a parameter denoting the cost per vaccination and is equal to 1.48 NOK per fish. The total expected costs, z, can now be defined in equation 6.3.

$$Minimize\ z = z_{energy} + z_{fry} + z_{growth} + z_{vaccination} + Z_{fixed} \tag{6.3}$$

Have in mind that the variable costs driven by the amount of fry deployed, z_{fry}, has been changed so that it no longer contains the vaccination cost. Figure 6.23 shows the breakdown of cost per smolt for the new objective function, the original model and the actual costs at the Slørdal facility in the period.

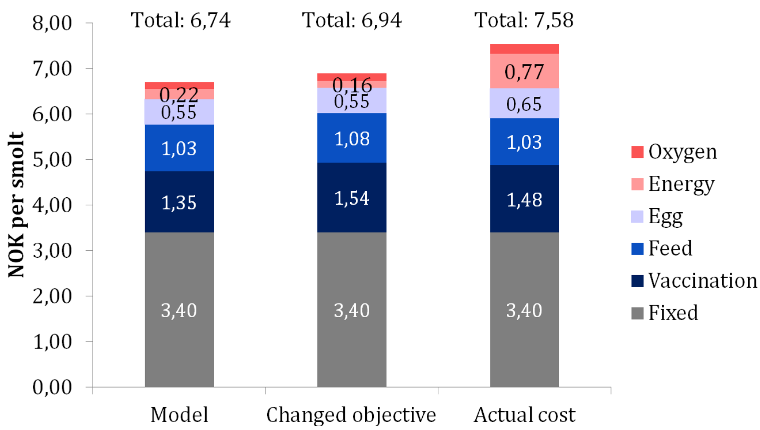

Figure 6.23: The segmented cost per smolt of the three model runs.

The new vaccination cost is 14% higher than in the original model. However, it is closer to the actual vaccination cost which may indicate that this new formulation is a more accurate approximation. The slight overestimation is because mortality and culling occurs after moving the fish and before the *actual* vaccination weight is reached.

A factor complicating this analysis is the gap. The run with changed objective function did not achieve a gap lower than 19.9% after running for 59 hours. The increase in solution time could be due to more variables in the objective function. In comparison, the original model had a gap just below 10%. This difference makes it hard to conclusively compare the different model runs.

6.3 Cast study 2: 500 grams smolt production.

The regulatory environment in Norway has recently opened for production of larger smolts in freshwater facilities. The benefit is hypothesized to be more robust smolt since they are larger when deployed at sea. The Slørdal facility is not currently producing smolt as heavy as this, and it is therefore interesting to investigate how such a product can be made at a minimum cost.

During which season the fish should be delivered remains an open question, and in order to compare the different options we did four model runs. These have deliveries of 50 000 smolt in the spring (April), summer (July), fall (October) and winter (January) respectively. The results from these four runs are compared in order to determine both the preferred season and the optimal production plan for this season.

In order to produce 500 grams smolts, the dataset had to be changed and augmented in the following manner:

- The time horizon is 3 years starting February 2011 to January 2014.
- New weight classes are added at 225, 250, 275, 350, 400, 450 and 500 grams, with a new "sink weight class" at 900 grams.
- The weight limits in region 3 is changed into order to accommodate the largest weight classes. The limits in region 1 and 2 remain unchanged.
- Skretting's seawater growth model is used to determine the relationship between temperature and growth in weight classes above 200 grams.
- The mortality rate of 200 grams is used for all weight classes above 200 grams.
- The order amount is for 50 000, and the order can be satisfied with fish in the range 400 grams to 900 grams.
- There are no initial fish inventory levels.
- The scenario tree split is put halfway thought the time period (July 2012).
- The fish will smoltify prior to reaching 500 grams, and will therefore have to be put into seawater tanks. It is assumed that the tanks at Slørdal enable this, and the freshwater intake temperature at the Slørdal facility is used as a proxy for the seawater temperature. This is because historical seawater temperature data is not available.
- A minimum gap requirement of 10 % gap is used.

Table 6.1 summarizes the results from the four model runs. It shows that delivering 500 grams smolt during the winter has by far the lowest cost. There is also a significant difference in the time spent in the facility by the different fish batches.

This is because the average water temperature and growth varies significantly. The required weight can therefore be reached faster. The average delivery weight is significantly larger for the delivery during the summer than the others. This is because the smolt has high water temperatures during the last months prior to delivery which makes them grow a lot.

Delivery season	**Months in the facility**	**Variable cost, NOK**	**Average delivery weight, grams**	**Gap**
Spring	22	21.8	478	3%
Summer	24	47.6	604	7%
Fall	19	17.0	527	2%
Winter	20	14.7	497	2%

Table 6.1: A summary of four optimal 500 grams smolt with deliveries in each season respectively.

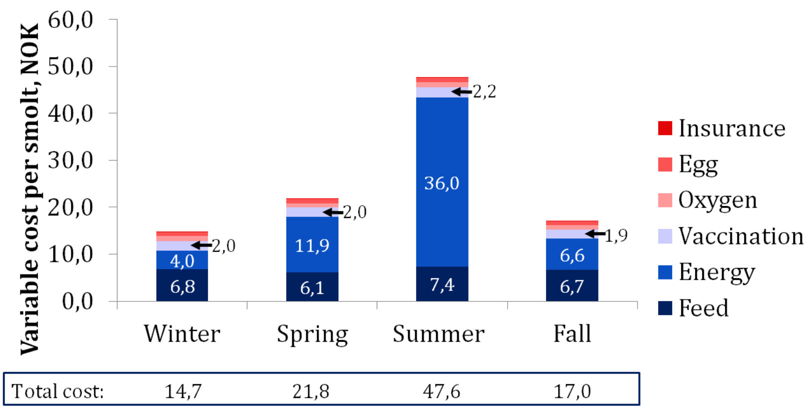

Figure 6.24: A breakdown of variable costs for deliveries in each season.

The breakdown of the variable costs in figure 6.24 shows that the cost difference is dominated by the energy cost, and this cost is lowest for the winter batch. Energy has such a large impact because the batch size is a lot smaller than in case study 1 (50 000 vs. 8 million). The amount of water heated is comparable to case study 1, but there are fewer fish to allocate the costs to in case study 2. The second biggest cost is feeding. The spring delivery has the lowest feeding cost, and for a larger batch size this would have a larger impact on the total cost.

The case study results indicate that the 500 grams smolt should be delivered in January since this involves the lowest total cost. Fry deployment should occur 20 months earlier, which is in May. However, the desired batch size and other products made in parallel at the facility should be considered since this will greatly affect the energy costs per smolt delivered.

Chapter 7

Conclusion

The main result is four key findings on smolt production based on two case studies. The main academic contribution is a novel stochastic programming model for cost minimization of smolt production. To our knowledge, this is the first published model considering this area of research.

Our model is inspired by profit maximization models used for the seawater phase of salmon production. One of our innovations is a linear formulation using SOS2 which enables water temperature to be a decision variable. The main uncertainty, freshwater intake temperature, is modeled through the use of scenarios. Discrete scenarios have been generated using a seasonal AR(1)-model and moment-matching. A computational cluster has been used to run the model due to its size. Testing and feedback sessions with biologists were conducted to ensure model quality.

Based on the case studies, sensitivity analyses and feedback sessions with Marine Harvest we have identified four key findings regarding smolt production:

- **Cost minimization can have a significant impact on the total costs.** A typical two year production plan at the Slørdal facility experienced a total cost reduction of 11%.
- **Smolt orders can be fulfilled with reduced water heating.** Smolt can be produced with far less water heating compared with today's praxis if the fry are deployed earlier. This has a significant impact on the production costs.
- **Cost minimization will not necessarily reduce smolt quality.** Weight is commonly used as a proxy for quality and all orders in the main case study are delivered above the order weight. This is especially evident for orders delivered during the summer and autumn months. However, since there is no incentive to deliver heavier smolt in the model this statement is not guaranteed to hold for all case studies.
- **500 grams smolt should be delivered during the winter.** The case study indicates that there are significant cost and production time differences between delivering 500 grams smolt in the summer compared to the winter. This is because the temperature seasonalities are utilized to produce growth in a better way. However, these results should be challenged in a more holistic case study which considers several batches and products in parallel such that energy costs may be shared.

Computational resources are limited and aspects included in a model must be prioritized such that the model quality is as high as possible. Analysis of the case studies have indicated modeling aspects which should be reconsidered to increase the model quality:

- **The gain of modeling heating relative to the cost of introducing SOS2 is low.** The case study results shows that little water heating is used and energy makes up only $\sim$7% of the model variable costs. The computational resources may therefore be better spent in terms of solution quality by including more temperature scenarios, other uncertainty factors such as mortality scenarios or considering several facilities in a value chain.
- **Including temperature uncertainty does not necessarily yield a better solution.** The case study results show little difference in the decisions taken in the various scenarios. This indicates that recourse actions might not have a large value in different temperature scenarios. Representing other uncertain factors as stochastic parameters (e.g. mortality) might therefore have a higher value.

Since we present a new model, there are multiple aspects of the model which could be further analyzed or extended. An interesting analysis could be to run the model using a rolling horizon and consider the long term effects of production plans. However, the available computational platforms were not compatible for this kind of analysis.

We believe that it would be fruitful to extend the model outside of the defined scope. It could be beneficial to consider freshwater production in a holistic value chain and to include model risk management aspects, such as rewarding flexibility and having buffer fish levels. The latter would represent important aspects of today's praxis.

Also, as a general note, a model cannot be better than the available data. We therefore recommend that Marine Harvest develop a growth model which approximates the growth experienced at their freshwater facilities more accurately. This is because the feedback sessions revealed that the Skretting growth model currently used exaggerates the growth.

Reference list

Alexander, C. (2009). *Market Risk Analysis, Quantitative Methods in Finance*, volume 1. Wiley. pages 143–185.

Bergfjord, O. J. (2009). Risk perception and risk management in Norwegian aquaculture. *Journal of Risk Research*, pages 91–104.

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science+ Business Media. pages 21–26, 40–46 and 163–176.

Bjørndal, T. (1988). Optimal harvesting of farmed fish. *Marine Resource Economics*, 5:139–159.

Bjørgo, S. (2013). En fylkeskommune med et utvidet ansvar? The Norwegian Seafood Federation, annual conference. Presentation - 13.02.2013.

Box, G. E. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 211–252.

Box, G. E., Jenkins, G. M., and Reinsel, G. C. (2011). *Time series analysis: forecasting and control*, volume 734. Wiley.

Braude, E. J. (2000). *Software engineering: an object-oriented perspective*. John Wiley & Sons, Inc. pages 40–41.

Brooks, C. (2008). *Introductory econometrics for finance*. Cambridge university press. pages 451–486.

Clark, C., Edwards, G., and Friedlaender, M. (1973). Beverton-Holt model of a commercial fishery: optimal dynamics. *Journal of the Fisheries Board of Canada*, 30(11):1629–1640.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, pages 1057–1072.

Elliott, J. and Hurley, M. (1997). A functional model for maximum growth of Atlantic salmon parr, Salmo salar, from two populations in northwest England. *Functional Ecology*, 11(5):592–603.

Forsberg, O. I. (1999). Optimal harvesting of farmed Atlantic salmon at two cohort management strategies and different harvest operation restrictions. *Aquaculture Economics & Management*, 3(2):143–158.

Forseth, T., Hurley, M., Jensen, A., and Elliott, J. (2001). Functional models for growth and food consumption of Atlantic salmon parr, Salmo salar, from a Norwegian river. *Freshwater biology*, 46(2):173–186.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer Series in Statistics. pages 191–216.

Gonchar, A.A. and Shabat, B.V. (2011). Analytic function. `http://www.encyclopediaofmath.org/index.php/Analytic_function`. Date retrieved: 6. June, 2013.

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

Hæreid, M. B. (2011). Allocating Sales in the Farming of Atlantic Salmon: Maximizing Profits Under Uncertainty. Master's thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management.

Hansen, T. (1998). *Oppdrett av laksesmolt*. Landbruksforlaget. pages 1–225.

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

Jarque, C. and Bera, A. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Kaut, M. (2003). ScenGen. `http://work.michalkaut.net/`. Program used for scenario generation - Date retrieved: 1 February. 2013.

Kaut, M. and Wallace, S. W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2):257–271.

King, A. J. and Wallace, S. W. (2012). *Modeling with stochastic programming*. Springer. pages 83–88.

Langan, T. B. and Toftøy, T. (2011). Produksjonsoptimering innenfor lakseoppdretten - planlegging under usikkerhet. Master's thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management.

Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Marine Harvest (2012). Seafood value chain. `http://www.marineharvest.com/en/Seafood-Value-Chain1/`. Date retrieved: 6. June, 2013.

Marine Harvest (2013a). Data received from Marine Harvest. Contact: Anders Jon Fjellheim, Ole Christian Norvik, Stian Aspaas, Ståle Eilertsen and Lisbeth Løvmo Martinsen.

Marine Harvest (2013b). Kvalitetshåndbok Kontroll av produksjonsmiljø – tetthet, vannstrøm og spesifikt vannforbruk. Received by email from Marine Harvest.

Marine Harvest (2013c). The Marine Harvest Salmon Industry Handbook. http://hugin.info/209/R/1698446/559980.pdf. Data retrieved: 6. June,2013.

Ministry of Fisheries and Coastal Affairs (2008). FOR 2008-06-17 nr 822: Forskrift om drift av akvakulturanlegg. http://www.lovdata.no/for/sf/fi/xi-20080617-0822.html#46. Date retrieved: May 8, 2013.

Misener, R. and Floudas, C. (2010). Piecewise-linear approximations of multidimensional functions. *Journal of optimization theory and applications*, 145(1):120–147.

NordPoolSpot (2013). Monthly Elspot Prices. http://www.nordpoolspot.com/Market-data1/Downloads/Historical-Elspot-prices/Monthly-Elspot-prices/. Date retrieved: 29 April. 2013.

Norwegian Seafood Council (2012a). Årstallspresentasjon 2012. http://www.seafood.no/content/download/78831/949842/file/sjomatradet_aarstallpresentasjon2012_highgres.pdf. Date retrieved: 3. June, 2013.

Norwegian Seafood Council (2012b). Seafood - Key figures. http://en.seafood.no/content/download/34512/619870/file/Compressed_English_Generell%20Norges%20sjomatrad%20og%20norsk%20sjomat%202012.pptx. Date retrieved: 3. June, 2013.

Nylund, A., Plarre, H., Hodneland, K., Devold, M., Aspehaug, V., Aarseth, M., Koren, C., and Watanabe, K. (2003). Haemorrhagic smolt syndrome (HSS) in Norway: pathology and associated virus-like particles. *Diseases of aquatic organisms*, 54(1):15.

Ratkowsky, D., Lowry, R., McMeekin, T., Stokes, A., and Chandler, R. (1983). Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *Journal of Bacteriology*, 154(3):1222–1226.

Røsstad, J. S., Bjordal, K., and Granfeldt, C. (2013). Dialog med våre tre fôrselskap. The Norwegian Seafood Federation, annual conference. Presentation - 13.02.2013.

Rynning-Tønnesen, C. and Øveraas, S. (2012). Production Optimization in the Salmon Farming Industry: Ordering Smolt Under Uncertainty. Master's thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management.

Schmidt, G. and Wilhelm, W. E. (2000). Strategic, tactical and operational decisions in multi-national logistics networks: a review and discussion of modeling issues. *International Journal of Production Research*, 38(7):1501–1523.

Skretting (2009). Skretting growth rate model - freshwater (revised Sept. 2009). Not published - received by email from Marine Harvest.

Song, X. and Chen, L. (2001). Optimal harvesting and stability for a two-species competitive system with stage structure. *Mathematical biosciences*, 170(2):173–186.

Tipler, P. and Mosca, G. (2008). *Physics for Scientists and Engineers sixth edition*. W. H. Freeman and Company. pages 591-665.

Veterinærinstituttet (2012a). Fiskehelserapporten 2011. Technical report, Veterinærinstituttet. Olsen, A. B. (ed.) and Hellberg, H. (ed.).

Veterinærinstituttet (2012b). Fiskehelserapporten 2012. Technical report, Veterinærinstituttet. Johansen, R. (ed.).

Wei, W. (2006). *Time series analysis: univariate and multivariate methods*. Pearson Addison Wesley. pages 33–52, 108–119, 160–182, 368–373 and 464–470.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, pages 817–838.

Williams, H. P. (1999). *Model building in mathematical programming*. Wiley. pages 144–185.

Yu, R., Leung, P., and Bienfang, P. (2006). Optimal production schedule in commercial shrimp culture. *Aquaculture*, 254(1):426–441.

Zenios, S. A. (2008). *Practical financial optimization*. Blackwell Pub. pages 225–250.

Appendix A

Degrees of freedom in SOS2 linearization

An issue with the linearization technique using SOS2 described in section 4.1.5 is that *three* equations are used to define a linear combination of *four* variables. Thus, there is one degree of freedom which an optimization model will exploit in order to improve the objective function value.

Williams (1999) proposed a solution which introducing a third SOS2 related to the diagonal of the discretized (x, y)-plane. This eliminates the single degree of freedom. Misener and Floudas (2010) explored this approach further by showing how the third SOS2 can be applied either as a downward or upward slope in a two dimensional plane.

Figure A.1 shows how a third SOS2 can be applied as an upward slope to the generic linearization approach discussed section 4.1.5. In this approach, one weighting variable ω_k is assigned to each diagonal k. The ω_k's are a SOS2, and thus only two neighboring variables can be non-zero.

We tried using both upward and downward sloping SOS2 in the model implementation. Both resulted in unnatural large amounts of culling and a significantly increased solution time. This observed behavior lead to the conclusion that applying an extra SOS2 did not improve the model quality. A better approach is to increase the number of fish weight and number classes. Other authors, such as Gunnerud et al. (2012) came to a similar conclusion when using SOS2 to linearize an oil optimization model.

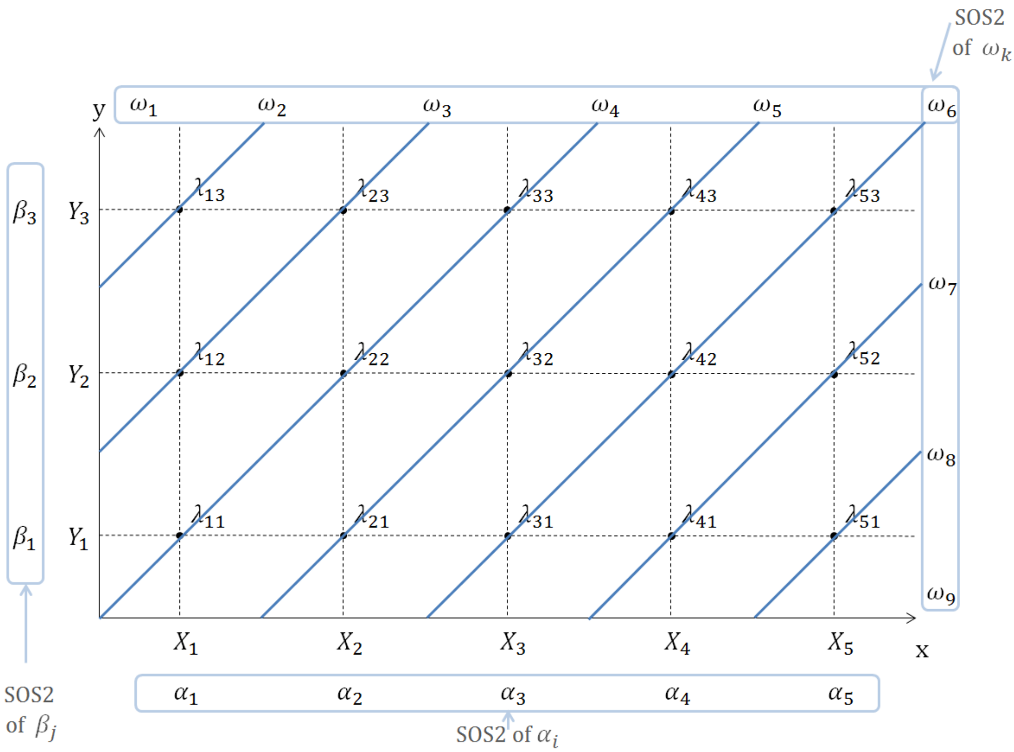

Figure A.1: A third upwardly sloping SOS2 set is added. There is one weighting variable, ω_k, for each upward sloping diagonal, and only two neighboring variables can be non-zero.

Appendix B

Constraint testing performed

Test number	Purpose of test	Test description	expected results	Results
Test 1	Test active and passive growth	Use one tank in one region. Let the power price and the temperature profile be flat. Let the passive growth have a jump to the delivering class from weight class X, so that it will always outperform the active growth for any fish weight class larger than X. Perform two runs where: 1. Active growth is used in the tank. 2. Passive growth is used in the tank.	With the first run (active growth) jumps in weight larger than one fish weight class should occur. With the second run (passive growth) the fish weight should only jump once the fish weight class X has been reached.	The model reacted as expected
Test 2	Test active and passive growth between regions	Same setup as Test 1, but with two regions and one tank in each region. Let one region have active growth and the ability to take in fry. Let the other region have passive growth and the ability to deliver fish on an order.	Fish should start in the region with active growth and stay there until they reach fish class X. They should then move over to the other region.	The model reacted as expected
Test 3	Test movement of fish	Same setup as Test 2, but the region with active growth is only allowed to have fish from the lowest weight class to class X. Let the area with passive growth be able to have fish from class X to the delivery class. Perform two runs where: 1. It is legal to move fish between regions. 2. It is not legal to move fish between regions.	If movement is legal, fish should be delivered from the region with passive growth. If movement is not legal, the model should be infeasible.	The model reacted as expected
Test 4	Test Maximum Allowed Biomass (MAB)	Same setup as Test 3, but this time include the maximum allowed biomass (MAB) constraints. Do two runs where: 1. Order size is set such that MAB have to be violated in order to fulfill it. 2. Order size is set such that upon delivery the biomass is close to the MAB.	In the first run where MAB is set too strict, the problem should be infeasible. In the second run, the problem should have a solution.	The model reacted as expected
Test 5	Scenarios	Same setup as Test 4. Do three model runs where: 1. Only one temperature scenario where the order time is set so that heating is needed. 2. Multiple scenarios where each scenario is equal to the scenario in run 1. 3. Multiple scenarios, as in run 2, but with one scenario having a elevated temperature in each time period.	There should be no difference between run 1 and run 2, while run 3 should have a slightly better result (lower objective function)	The model reacted as expected.

Table B.1: Constraint tests performed on the implemented model. *Active growth* denotes the use of growth approximation (water heating allowed). *Passive growth* denotes the use of pre-calculated growth paths (no water heating allowed).

Appendix C

Out-of-sample stability testing

Out-of-sample stability is related to model quality and how well the model, using the scenario generation procedure, approximates the *actual* expected objective function value. *"Out-of-sample stability means that if you calculate the true objective function value corresponding to the solutions coming from different scenario trees, you get (about) the same value."* (King and Wallace, 2012, page 85). In contrast, in-sample stability regards the model's robustness and does not concern the actual quality of the solution as long as it is not sensitive to the scenario generation procedure (King and Wallace, 2012).

Ideally, we would like to be able to assess whether or not the scenario generation method is able to produce robust solutions with respect to the true distribution of the random variables. However, since this true distribution rarely is observable, we have to approximate it. A possible approach is to generate a scenario tree with 100 scenarios, and deem this a sufficient approximation of the *true* temperature distribution.

Based on the approach described by King and Wallace (2012) and discussions with Professor Asgeir Tomasgard, an out-of-sample stability test can be conducted in the following manner:

1. Solve the optimization model with a scenario tree with X scenarios. X could for instance be 4
2. Solve the optimization model with a scenario tree with Y scenarios. Y could for instance be 10
3. Use first stage decisions from a scenario tree with X=4 scenarios, and solve the 100 second stage optimization problems corresponding to each branch in the *true* scenario tree. The approximated true objective function value corresponding to the solution from tree X is the expected value of these 100 objective values weighted by the scenario probability.
4. Use first stage decisions from a scenario tree with Y=10 scenarios, and solve the 100 second stage optimization problems corresponding to each branch in the *true* scenario tree. The approximated true objective function value corresponding to the solution from tree Y is the expected value of these 100 objective values weighted by the scenario probability.
5. The approximated true objective function values corresponding to the decisions from tree X and Y respectively should have about the same value. The model is then deemed out-of-sample stable.

However, we have chosen to omit out-of-sample stability testing in this thesis for three reasons: First, the computational platform available (the computational cluster Solstorm) does not readily allow automation for solving 100 second-stage problems. Second, the solution time per second stage problem is relatively long (about 1000 seconds each). Third, MIP problems give upper and lower bounds on the optimal objective value which complicates calculating, comparing and making inferences based on the objective function values.