Fallacy of a correction factor in automated computer vision-based lice counting

Application of mixed-effects models

Marlin Firveld Andersen

Master's thesis in Statistics

Data Analysis

Department of Mathematics

University of Bergen

November 2023

Abstract

Parasitic salmon lice (*Lepeophtheirus salmonis*, Krøyer 1837) are a threat to the health and welfare of farmed salmonids and to the sustainability of wild salmon populations in Norway. Accordingly, the opportunities for growth of the farming industry are strictly regulated by the Norwegian government based on the relative abundance of salmon lice in commercial sea cages. Farmers are obliged to disclose weekly lice counts per location and to perform delousing treatments when preset thresholds are exceeded. Reported lice numbers are by default based on manual counts that are obtained by physical inspection of a limited sample of fish. In recent years, advanced camera technology and artificial intelligence have made it possible to obtain image-based lice counts based on representative sample sizes and without stressful handling of the fish. However, camera-based solutions do not usually represent the entire fish. In this thesis, a correction factor between image-based and manual parasite counts in Atlantic salmon (*Salmo salar*, Linnaeus 1758) aquaculture is therefore being explored. The objective was to determine a factor that accurately adjusts image-based parasite counts to account for lice on the opposite site of the fish. For this purpose, simple linear models and linear mixed-effects models were developed and validated, to model manual lice counts that were reported to the Norwegian authorities. Image-based and human-verified lice counts from Stingray Marine Solutions AS of equivalent sample sizes were used as the main predictor variable. Additional predictors included farmer-reported production data such as fish counts, average fish weight, delousing treatments, and sea temperature, as well as operational data from the Stingray system, which is designed to detect, target, and eliminate sea lice on freely swimming fish by means of machine vision and laser technology. Various candidate models were compared and tested on unseen data. Results did not confirm the hypothesis that image-based counts require upward adjustment to faithfully represent the accuracy of manual counts. Instead, the resulting correction factor suggested that the opposite was the case. Statistical and procedural explanations were explored by means of simulated lice counts and are discussed with regard to current regulations.

Acknowledgements

I would like to express my deepest appreciation to my supervisors professor Jan Bulla from University of Bergen, as well as Martin Worm and Benedikt Frenzl from Stingray Marine Solutions AS, later referred to as Stingray. Your help and guidance have been extremely helpful. Professor Jan Bulla has long experience with practical statistics, and he is one of the most capable statisticians I have met. He has been pushing me to work harder as well as sharing good knowledge about structure and statistical models to mention some, I am forever grateful he was my supervisor. Martin and Benedikt both hold a PhD in biology, and have long experience in the field of fish farming and specifically salmon lice. Their input, especially on the biology parts, have been invaluable. I would also like to extend my appreciation to Stingray in general that shared data with me, and to everyone involved who provided their knowledge, experience and support. Thank you to my friends and family for supporting me and being patient throughout this project, I am grateful to you all, especially to George.

Contents

1 Introduction

Aquaculture is the husbandry and propagation of aquatic organisms such as plants and animals in all types of water environments. According to Norway's Directorate of Fisheries, referred to as Fiskeridirektoratet in Norwegian, approximately 1,551,972 tons of Norwegian produced salmon were sold in 2022. In addition 84,928 tons of rainbow trout were sold the same year. Fiskeridirektoratet is the advisory and executive body within fisheries and aquaculture management in Norway (*fiskeridir.no*).

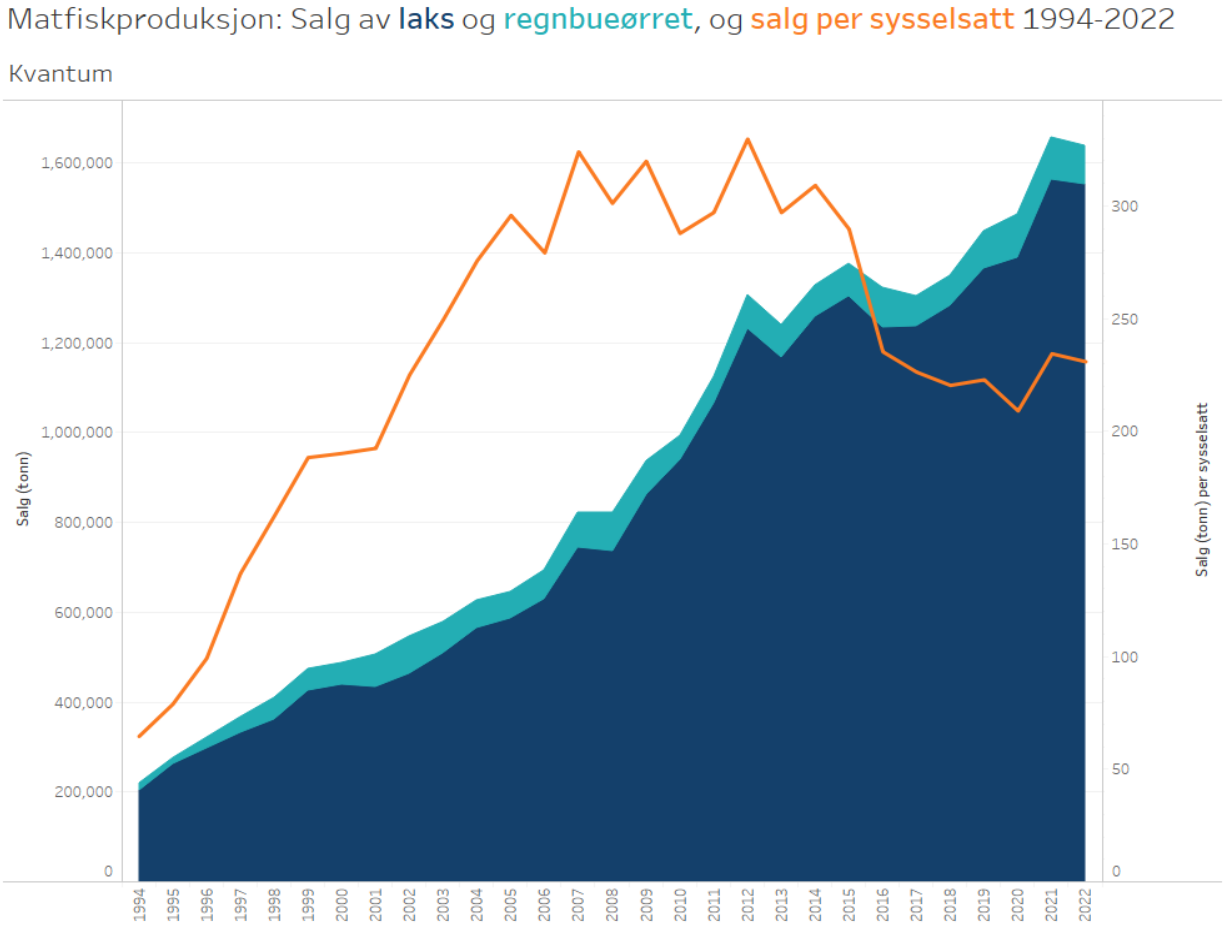


Figure 1: *This figure shows quantity (tons) of salmon (dark blue) and rainbow trout (light blue) sold in Norway between 1994 and 2022 (primary y-axis). The orange line shows sale in tons per employee (secondary y-axis). This figure is adapted from Norwegian. Source: Fiskeridirektoratet, https://www.fiskeridir.no/Akvakultur/Tall-og-analyse/Akvakulturstatistikk-tidsserier/Laks-regnbueoerret-og-oerret/salg-av-laks-og-regnbueorret. Last accessed 05.11.2023*

Figure 1 shows steady increase in quantity of sold trout and salmon since 1994 and reflects the expansive nature of the aquaculture in Norway. The export of Atlantic salmon made up over 60% of the total seafood export value in Norway in 2021 (Johnsen et al., 2021). Therefore, salmon farming can be considered one of the most important industries in Norway (F. Johnsen et al., 2022). Salmon farming in Norway has been established along the entire West coast and is structured in thirteen salmon production areas (Produksjonsområdeforskriften, 2017, §3). The high density of salmon, as well as the close proximity of salmon farms to each other, lead to a large biomass of fish. As a result, similar to all intensive farming practices (Mennerat et al., 2010), disease outbreaks must be expected. The most common parasitic disease for salmonids in Norway is the salmon louse (*Lepeophtheirus salmonis,* Krøyer), a parasitic temperate marine copepod species.

Salmon lice are parasites that feed off the skin, mucus and blood of salmonid hosts (Torrissen et al., 2013; Grimnes and Jakobsen, 1996). Low to medium louse levels can lead to injured and sick fish with reduced appetite and growth. High infection levels can lead to increased mortality in fish (Costello, 2006; Hamre et al., 2009). However all sea louse grazing on the host's body can cause secondary viral, bacterial or fungal infections detrimental to the fish (Ugelvik and Dalvin, 2022). In addition to being a threat to the welfare of the farmed salmonids, and a potential threat to wild salmon populations (Taranger et al., 2015), salmon lice cost the salmon industry economically around 5 billion NOK per year in Norway in 2017 (Iversen et al., 2017). This includes direct costs for treatments against salmon lice as well as indirect costs related to monitoring the sea louse levels. In Norway, the government and its governing body, the Norwegian food safety authority, referred to as Mattilsynet in Norwegian, have enforced a strict sea louse monitoring program (*mattilsynet.no*; Forskrift om lakselusbekjempelse, 2013). All salmon farms have to count and report sea louse abundance to Mattilsynet every week. The Norwegian government have fixed limits on allowed amount of adult female lice per fish. When exceeding the sea louse limit, treatment against the parasite has to be initiated (Forskrift om lakselusbekjempelse, 2013).

The threshold for allowed sea louse abundance is set between 0.2 and 0.5 adult female sea lice per fish with limits depending on location and time of the year (Forskrift om lakselusbekjempelse, 2013). The abundance of salmon lice in a fish farm is measured by calculating the number of lice

per fish on a minimum of ten or twenty fish per pen, depending on time of the year. The average is calculated for all sea louse life cycle stages, namely adult females, mobile stages and fixed stages, but only increased levels in adult females can trigger a legal response. The different louse stages will be explained in further detail in Section 2. As of today, sea lice are counted by physical handling and visually inspecting the fish, unless the farmer has obtained a dispensation from this obligation by Mattilsynet (Mattilsynet, 2020). This counting method is not ideal due to practical restraints such as rough weather, or the difficulty to obtain a representative sample of fish (Thorvaldsen et al., 2019). In addition, physical handling is stressful for the fish (Delfosse et al., 2021) and it has been shown that manual counts can be biased downwards (Elmoslemany et al., 2013; Godwin et al., 2021) and prone to human counting error. The main problem might, however, be the physical limitations of manual counting since it is only practically possible to count a small amount of fish each week.

In recent years, the Norwegian aquaculture industry has been pushing to implement novel technology to streamline all aspects of salmonid production. This push includes upgrades of infrastructure, such as hybrid barges for clean and stable electricity, 4G and 5G internet coverage and the use of robotics and artificial intelligence for welfare and behaviour logging (Føre et al., 2018). Stingray Marine Solutions AS specialises in sea louse treatment using a method referred to as optical delousing (*stingray.no*). Stingray uses a guided laser pulse and machine learning to recognise and shoot sea lice on fish in salmon or trout pens.

Additionally the Norwegian government has encouraged the use of modern technology to monitor sea louse levels. Stingray, among others, offers image-based lice counts to improve fish health and fish welfare monitoring. Image-based counts are counts performed on random images of fish. These counts are analysed by a machine learning detector for sea lice and can be verified by a human operator. Image-based counts are approved for commercial use on a case to case basis, granted by the authorities (Mattilsynet, 2020). In addition, Norway's standardisation organisation, referred to as Standard Norge in Norwegian, has been appointed to establish a standard for automated louse counts in Norway to fully replace manual, physical lice counts with image detection technology (Standard Norge).

The advantages of incorporating modern machine vision, neural networks and detector technology in aquaculture as compared to manual human counting and analysis must be understood in a statistical context. It is now possible to monitor sea lice more frequently and at much higher sample sizes. Theoretically there is no limit to both parameters, but practically constraints are imposed by data volume and up/download capabilities. Whereas it is possible to fully examine a fish manually, using a combination of human senses for a full 3-dimensional check, images only offer a 2-dimensional view, without the possibility of physically handling the fish. Thus any camera-recorded image will not be able to present the whole fish, from every angle. This problem can be solved by correcting for the lee side of the fish, which is not captured by the camera. Therefore image-based sea louse counts need to be subjected to a mathematical algorithm which estimates the total amount of sea lice on the whole fish from the number of sea lice counted in the picture. This algorithm is referred to as a correction factor for image-based lice counting. It is thought that the automatic counts will underestimate the true louse abundance, hence a correction is needed to get closer to a correct sample estimate. To my knowledge, there exists no published information on finding a correction factor in this context. However, statistically, manual sea louse counts cannot be taken as the ground truth for comparison due to a large sample size needed in a pen with up to 200,000 individuals (Departementenes sikkerhets- og serviceorganisasjon Teknisk redaksjon, 2023). This thesis will highlight ways, possibilities and challenges to implement a correction for Stingray's image-based sea louse counting method. This author will treat manual louse counts as the ground truth, and the dependent variable in the mixed-effects models throughout the analysis. Furthermore, I will explore why finding a correction factor is such a challenging task given the data we have access to today, by means of a simulation study. The word "count" will be used frequently throughout this thesis, but this is not to be confused with discrete counts. When the word "count" is used in this thesis, we are referring to average salmon lice counts, which are continuous values.

2 Biological background

In this section, some biological information about salmon lice and treatment against them will be provided to better the understanding of the content of this thesis. A salmon louse is a copepod crustacean from the caligid family. It is a marine parasitic crustacean on salmonids and damages the fish by eating mucus, skin and blood (Torrissen et al., 2013; Grimnes and Jakobsen, 1996). The infected fish host can die if infection levels are high, have reduced growth of the host, osmoregulatory problems, severe skin wounds and secondary infections (Costello, 2006; Hamre et al., 2009).

2.1 Life cycle of the salmon louse

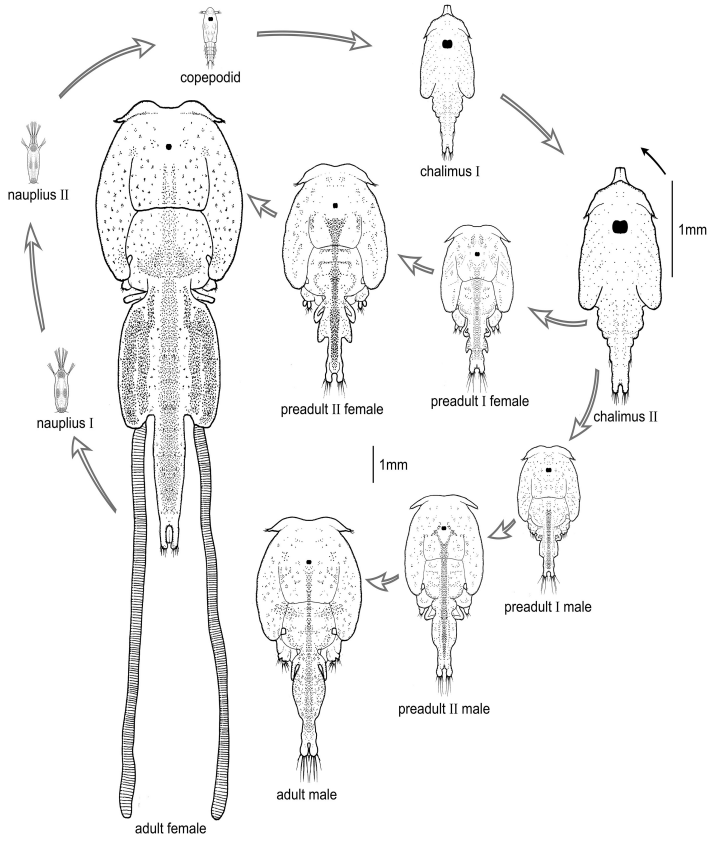


Figure 2: *Life cycle of the salmon louse (Lepeophtheirus salmonis). Once a salmon louse has hatched from an egg, it has a life cycle consisting of eight stages, each separated by a molt. This illustration shows an obvious difference in appearance and size for male and female lice in the final three stages of their life cycle. Illustration: Sea Lice Research Centre, 2020, "SLRC - Life cycle of the salmon louse (Lepeophtheirus salmonis)", https://doi.org/10.18710/GQTYYL, DataverseNO, V1, last accessed 23.04.2022.*

Figure 2 shows the life cycle of a salmon louse, consisting of eight stages. The first two stages are called the nauplius stages. At this stage the lice are planktonic, orient themselves through visual cues and light and float freely in the water column. The copepodid stage is the infective stage when the salmon louse attaches itself to a salmonid host, often on the fins and the skin on the side of the fish (Tucker et al., 2002). If failing to attach itself to a host, the copepodid will not be able to complete its life cycle (Nordtug et al., 2021). After attaching, the louse anchors itself to the host and starts to feed on the fish (Grimnes and Jakobsen, 1996). Once fully attached, the sea louse will enter the first of two chalimus stages (Hamre et al., 2013). The chalimus stages are also referred to as fixed stages because they remain attached in a fixed position and do not move around on the fish.

The preadult and adult stages of the salmon louse are mobile, which means that they can move around freely on the fish (Igboeli et al., 2014). Since adult female lice give raise to new offspring, they are of particular importance when managing salmon lice infestation in fish farms (Forskrift om lakselusbekjempelse, 2013). It is therefore commonly distinguished between adult females and all other mobile stages (Forskrift om lakselusbekjempelse, 2013). Mobiles include all lice in the preadult I, preadult II and adult male stages, whereas adult females are represented in their own category. After mating, an adult female louse produces pairs of egg strings from her abdomen, which might contain between seventy and 290 eggs per string (Heuch et al., 2000). The female carries these eggs until they hatch. Salmon lice larvae are hatched into the planktonic nauplius I stage, repeating the cycle.

2.2 Reproduction time

Salmon lice transition through the different stages of their life cycle depending on temperature and sex (Hamre et al., 2019). Salmon lice generally have greater infestation success, transit faster through the different stages of their life cycle, reaching sexual maturity at a greater rate, and reproduce faster with higher sea temperatures. Sea temperatures along the Norwegian coastline vary a lot depending on season and geography. The sea temperatures in the north of Norway are generally lower than the sea temperatures further south all year round. For reference, one of the locations used in this assignment is located in the western part of Norway and reached a maximum temperature of 16.5 °C in the data collection period. The rest of the six locations used in this assignment,

located in the north of Norway, never reached a temperature higher than 13.3 °C, having an average maximum temperature of 12.6 °C in the same period. The temperatures in the north of Norway usually increase later in the season, giving the northernmost fish farm locations a shorter period of time with higher sea temperatures. Because of higher temperatures, longer periods with higher temperatures and a greater density of fish farms in the western part of Norway, sea louse infestation levels are usually greater here than further north in the country (Hoddevik, 2021).

Males	RA_{frac}	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Ch1 all	0.2	23.4	17.3	13.5	10.9	9	7.6	6.5	5.7	5	4.4	3.9	3.5	3.2	2.9	2.7	2.4	2.2	2.1	1.9
Ch2 early	0.35	41	30.3	23.6	19.1	15.8	13.3	11.4	9.9	8.7	7.7	6.9	6.2	5.6	5.1	4.6	4.3	3.9	3.6	3.4
Ch2 all	0.4	46.8	34.6	27	21.8	18	15.2	13.1	11.3	10	8.8	7.9	7.1	6.4	5.8	5.3	4.9	4.5	4.1	3.8
Pa1 early	0.55	64.4	47.6	37.1	29.9	24.8	20.9	17.9	15.6	13.7	12.1	10.8	9.7	8.8	8	7.3	6.7	6.2	5.7	5.3
Pa1 all	0.6	70.3	52	40.5	32.7	27	22.8	19.6	17	14.9	13.2	11.8	10.6	9.6	8.7	8	7.3	6.7	6.2	5.8
Pa2 early	0.72	82	60.6	47.2	38.1	31.5	26.6	22.8	19.8	17.4	15.4	13.8	12.4	11.2	10.2	9.3	8.5	7.9	7.3	6.7
Pa2 all	0.8	93.7	69.3	54	43.5	36.1	30.4	26.1	22.7	19.9	17.6	15.8	14.2	12.8	11.6	10.6	9.7	9	8.3	7.7
Adult early	0.87	101.9	75.3	58.7	47.4	39.2	33.1	28.4	24.7	21.7	19.2	17.1	15.4	13.9	12.7	11.6	10.6	9.8	9	8.4
Adult all	1	117.1	86.6	67.5	54.4	45.1	38.1	32.6	28.4	24.9	22.1	19.7	17.7	16	14.5	13.3	12.2	11.2	10.4	9.6
Planktonic development: N1+N2	0.134	15.7	11.6	9	7.3	6	5.1	4.4	3.8	3.3	2.9	2.6	2.4	2.1	1.9	1.8	1.6	1.5	1.4	1.3
Females	RA_{frac}	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Ch1 all	0.16	23.5	17.3	13.5	10.9	9.1	7.7	6.6	5.8	5.1	4.5	4	3.6	3.3	3	2.7	2.5	2.3	2.2	2
Ch2 early	0.29	42.5	31.4	24.5	19.8	16.4	13.9	12	10.4	9.2	8.2	7.3	6.6	6	5.4	5	4.6	4.2	3.9	3.6
Ch2 all	0.36	52.8	39	30.4	24.6	20.4	17.3	14.9	13	11.4	10.1	9.1	8.2	7.4	6.8	6.2	5.7	5.2	4.9	4.5
Pa1 early	0.48	70.4	52	40.5	32.8	27.2	23.1	19.8	17.3	15.2	13.5	12.1	10.9	9.9	9	8.2	7.6	7	6.5	6
Pa1 all	0.56	82.1	60.6	47.3	38.2	31.8	26.9	23.1	20.2	17.8	15.8	14.1	12.7	11.5	10.5	9.6	8.8	8.2	7.5	7
Pa2 early	0.66	96.7	71.5	55.7	45.1	37.4	31.7	27.3	23.8	20.9	18.6	16.6	15	13.6	12.4	11.3	10.4	9.6	8.9	8.3
Pa2 all	0.76	111.4	82.3	64.2	51.9	43.1	36.5	31.4	27.4	24.1	21.4	19.2	17.3	15.6	14.3	13.1	12	11.1	10.2	9.5
Adult early	0.87	127.5	94.2	73.5	59.4	49.3	41.8	35.9	31.3	27.6	24.5	21.9	19.8	17.9	16.3	14.9	13.7	12.7	11.7	10.9
Adult all	1	146.6	108.3	84.5	68.3	56.7	48	41.3	36	31.7	28.2	25.2	22.7	20.6	18.8	17.2	15.8	14.6	13.5	12.5
First egg string all	1.3	190.5	140.8	109.8	88.8	73.7	62.4	53.7	46.8	41.2	36.6	32.8	29.5	26.8	24.4	22.3	20.5	18.9	17.5	16.3
Days between egg batches	0.25	36.6	27.1	21.1	17.1	14.2	12	10.3	9	7.9	7	6.3	5.7	5.1	4.7	4.3	3.9	3.6	3.4	3.1
Planktonic development: N1+N2	0.106	15.5	11.5	9	7.2	6	5.1	4.4	3.8	3.4	3	2.7	2.4	2.2	2	1.8	1.7	1.5	1.4	1.3

Table 1: *Table showing the amount of days before the early developers and the amount of days before all (above 85% of all lice within each group) have reached a given stage for both sexes, for different water temperatures in a controlled environment. RA stands for relative age and frac stands for fraction. Ch is Chalimus, Pa is Pre-adult, Adult is the adult stage and N stands for Nauplius. The table, exactly as it is presented here, is Table 4 in "Development of the salmon louse Lepeophtheirus salmonis parasitic stages in temperatures ranging from 3 to 24 °C" (Hamre et al., 2019).*

Table 1 shows that salmon lice transition faster through their respective life stages with increas-

ing temperature within their physiological optimum between 3 °C and 21 °C. Adult female sea lice transition through the different stages at an 80% slower rate than the males. At 21 °C, the amount of days from infection to the first eggs produced was 16.3 days. At 3 °C, the amount of days it took before the first egg production was 190.5 days. Typically in Norway, sea temperatures can drop below 4 °C in the winter time. There are special laws in place to account for the lower reproduction rate and slower transition time of salmon lice when that is the case. More about laws around salmon lice in Norwegian fish farms will be presented in Section 3.

2.3 Delousing treatments

Treatments against sea lice can be categorised as medicinal and non-medicinal treatments. This section will provide a short overview of the different delousing methods used in Norway between 2014 and 2019.

Delousing category	Frequency	Per cent (%)
Thermal	2692	57.97
Mechanical	619	13.33
Hydrogen peroxide	445	9.58
Medicinal bath	198	4.26
Freshwater bath	172	3.70
Combination	518	11.15
Total	4644	100

Table 2: *Table showing different types of delousing methods, the frequency of them and the overall percentage of all treatments conducted in Norway between 2014 and 2019. Combination is a combination of medicinal baths, or a combination of hydrogen peroxide and medicinal bath. The values in the table are taken from Table 2 in "Estimating cage-level mortality distributions following different delousing treatments of Atlantic salmon (Salmo salar) in Norway" (Sviland Walde et al., 2021)*

Table 2 shows the treatment methods used in Norway between 2014 and 2019. Thermal delousing is an invasive, non-medicinal method, and is performed by crowding fish, pumping them into a vessel and exposing them to approximately 28-38 °C, for approximately thirty seconds (Folkedal

et al., 2021). Thermal delousing was the most common delousing method in the given observation period. Mechanical delousing is an invasive, non-medicinal delousing method, where lice get removed by flushing, brushing or turbulence (Østevik et al., 2022). Hydrogen peroxide is a substance used as a medicinal delousing method. Fish get exposed to hydrogen peroxide within well-boats or as regular bath treatments using tarpaulin enclosed cages (Overton et al., 2018). Medicinal bath treatments are performed in the same way as the hydrogen peroxide method, but other substances may be used (Gautam et al., 2017). Freshwater baths are non-medicinal treatments that are performed in a similar manner to other bath treatments, but the fish are exposed to freshwater instead of medicinal substances for a period approximately between three and five hours (Thompson et al., 2023).

An increased mortality rate has been shown after all of the delousing methods described in Table 2 (Sviland Walde et al., 2021). The accumulated mortality of the standing stock of salmonid fish increased from 2014 to 2018, from 14.3% to 16.8% (Bang Jensen et al., 2020). Resistance to different medicines has become an increasing problem, and was acknowledged as a serious problem in the last years prior to 2008 (Myhre Jensen et al., 2020). This study also concluded that it is important to avoid heavy reliance on one or a few substance groups, as it has been shown to be a catalyst for the evolution of resistances. The resistance problem is likely the reason for an increased focus on other treatment methods, but the increased mortality rate and fish welfare issues following the alternative methods (Thompson et al., 2023) make the salmon louse situation a complex problem.

3 Methods

3.1 Data sets

This section will provide an overview of where data was collected from, how it was collected and why these specific data sets were included in the analysis to construct a correction factor. A total of twenty-one data sets have been collected across multiple platforms. Stingray has provided image-based lice counts, number of laser pulses, number of laser units and number of fish passing the lasers. Selected Stingray customers have provided fish weight and number of fish at the different locations. Manual lice counts, sea temperature and louse treatments have been collected from the governmental website *barentswatch.no*, where all this information is open access for everyone. In addition, the most important Norwegian laws regarding salmon lice have been included since regulations may explain missing data points as well as standard, established sea louse counting practices.

3.1.1 Data origin

The initial data sets contained data collected from fish farms at six different locations in Norway. Five of these locations are located in Northern Norway, whilst the last one is close to Florø. Different locations have different numbers of pens and numbers of fish per pen. A pen is an enclosure for fish. The pens at the locations reviewed in this assignment contain anything between zero and 200,000 individuals. The pens usually have one to two lasers working at all times, this varies from pen to pen. Pens with two lasers will cover a larger area than the pens with only one laser.

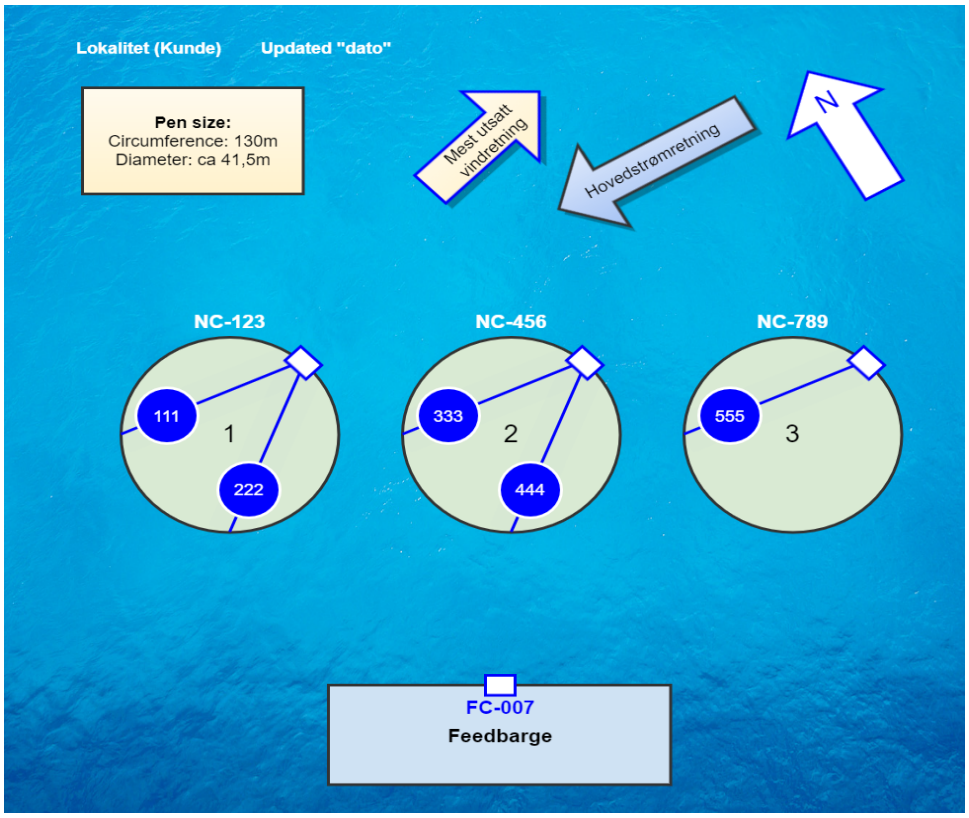


Figure 3: *Figure showing an example fish farm location utilising the Stingray technology in all its pens. Cabinets are denoted "NC-123", "NC-356", "NC-789" and "FC-007", and is shown as white rectangles with blue edges. The green circles, denoted "1", "2" and "3" are pens, the blue circles within the green circles denoted "111", "222", "333", "444" and "555" are nodes, or laser setups. Pen 1 and pen 2 contain two lasers, pen 3 contains one laser. Feedbarge is the fleet where the fish are fed from. The example feedbarge has a cabinet. The blue lines stretching across the pens from the cabinets, "through" the nodes, and stopping at the edge of the pen are cables. Furthermore, the map contains arrows containing information about currents, winds and what direction North is in relation to the pens.*

Figure 3 illustrates an example of a laser setup at a location. It gets clear from Figure 3 that pen 3 with only one laser covers a smaller area as the lasers can only be moved along the cable as well as up and down. All locations have a cabinet set up at the feedbarge. In Figure 3, that cabinet is denoted "FC-007". The cabinet on the feedbarge is the distribution point for the other cabinets. The distribution point cabinet supplies the other cabinets with internet. All ingress and egress traffic goes through the distribution point.

The data sets in this assignment have each been collected from different locations. To protect the privacy of the customers, all data sets have been anonymised and will be referred to as Location 1 through 6.

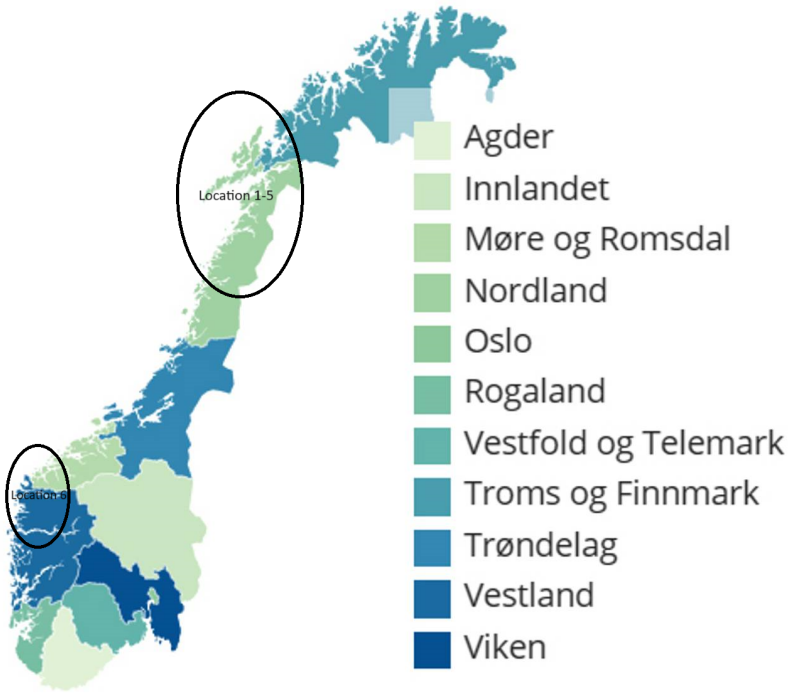


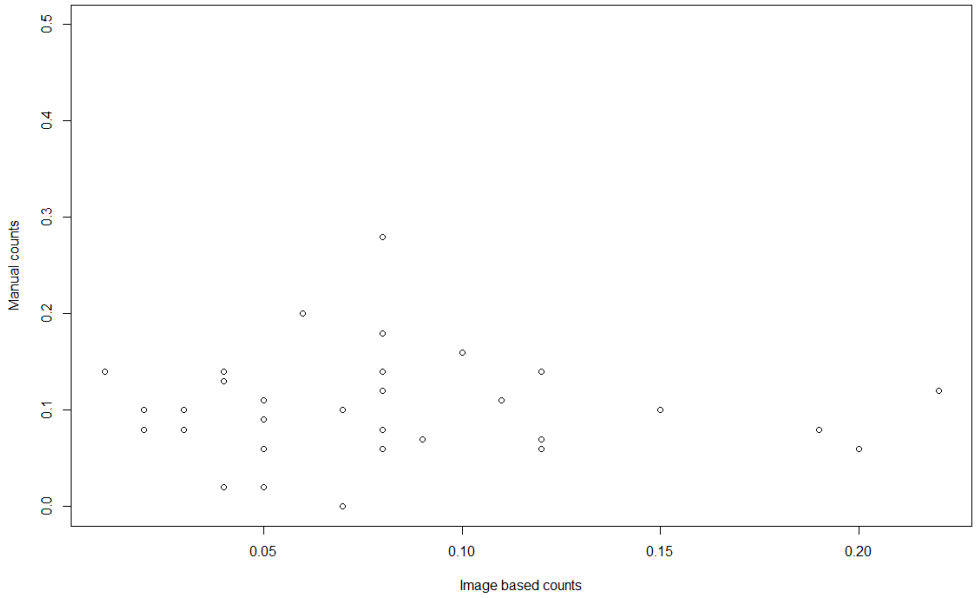
Figure 4: *Map with zones outlined where the fish farms are located. Location 1-5 are located in the northernmost circle and Location 6 is within the bottom circle. Norwegian counties are outlined with colours on the right side. Image taken from kommunaldepartementet.*

Figure 4 shows roughly where the fish farm locations analysed are located. Location 1 to 5 are close to each other in the north, whereas Location 6 is further south, on the Western coast of Norway.

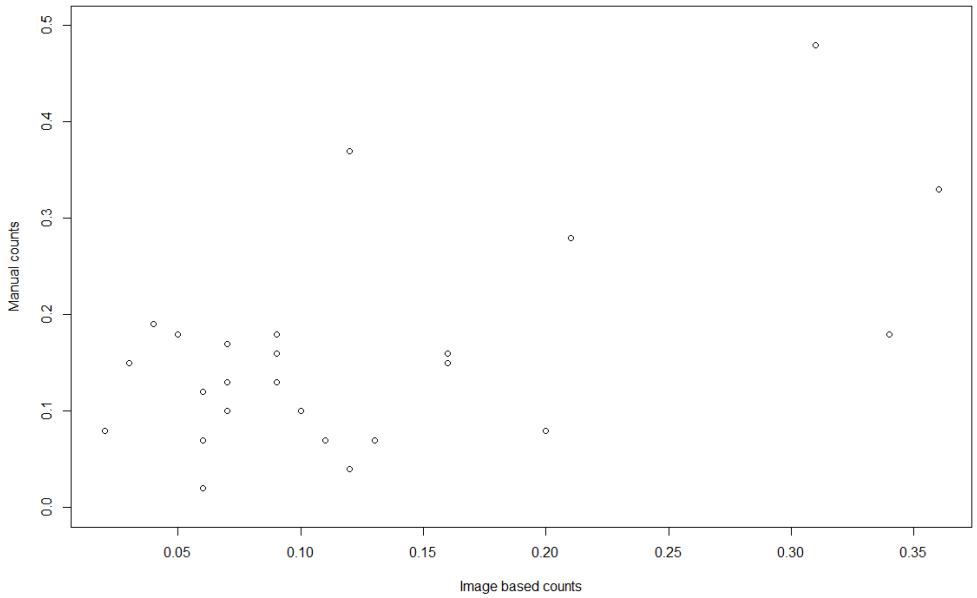
Location	Collection period	Production area	Population size	Lasers	Pens
Location 1	04.01.2021-23.08.2021	Close to Mo i Rana	383.640 fish	5	4
Location 2	04.01.2021-23.08.2021	Close to Lofoten	1.375.978 fish	12	10
Location 3	04.01.2021-02.08.2021	Close to Lofoten	817.230 fish	6	5
Location 4	04.01.2021-23.08.2021	North of Harstad	941.055 fish	17	10
Location 5	04.01.2021-19.07.2021	Close to Lofoten	310.374 fish	3	6
Location 6	04.01.2021-23.08.2021	Close to Florø	1.494.980 fish	16	8

Table 3: *The table shows the time periods the data has been collected in, production areas of the locations and the population sizes on location level (accumulated over all pens in that location). In addition, information about how many lasers each location had at the beginning of the data collection period, as well as how many pens each location consisted of, has been included.*

Table 3 shows general information about the locations that provided the data foundation for this assignment. The locations had varying numbers of fish, lasers, and pens. Location 5 had the fewest fish and the lowest number of lasers, and Location 6 had the most fish and the largest amount of lasers of all the locations. Location 5 was the only location with fewer lasers than of pens. However, after inspecting the relationships between manual counts and image-based counts, it was evident that none of the locations were displaying obvious relationships or trends between the two variables, except from at Location 6. Additional observations were therefore needed and collected later on for all of the locations, with some exceptions. The fish in Location 3 and 5 were harvested in week 35 in 2021, it was therefore impossible to gather more data for these two locations. Harvest found place at Location 6 in week 38, but the louse numbers were generally higher for this location and it was possible to somewhat observe a trend between manual and image-based counts. Hence, it was possible to use this location for analysis without additional observations.



(a) *Image-based counts plotted against manual counts for the available data in Location 3.*



(b) *Image-based counts plotted against manual counts for the available data in Location 5.*

Figure 5: *Plots of image-based counts against manual counts for Location 3 and 5 with the available observations, both with image-based counts on the x-axis and image-based counts on the y-axis. It was not possible to obtain more data for the two locations than what is displayed here.*

Figure 5 shows little to no trend in the plots of Panel (a) and (b). Based on this, the decision to remove Location 3 and 5 for further analysis was made. The locations and time periods used for further analysis were therefore:

- *Location 1*: 04.01.2021-27.12.2021
- *Location 2*: 04.01.2021-20.12.2021
- *Location 4*: 04.01.2021-15.11.2021
- *Location 6*: 04.01.2021-06.09.2021.

Lastly, a test data set was included to test how well the final models of the analysis performed on unseen data. The test data set included manual counts, image-based counts, time of treatments and normalised laser pulses for Location 1, 2, 4 and 6. Observations in the test data were collected for all four locations between 03.01.2022 and 23.10.2023. Many observations were missing. Rows with missing values were deleted. The test data set with missing values removed consisted of 113 observations in total.

3.1.2 Laws about salmon lice in Norwegian fish farms

Strict laws are in place for controlling salmon louse populations in Norwegian fish farms. All fish farms have to report louse numbers and sea temperatures measured 3m below the surface to Mattilsynet every seven days (Forskrift om lakselusbekjempelse, 2013). There are two exceptions to the law of having to report louse numbers once every week:

1) The sea temperature measured 3m below the surface is below 4 °C.
2) Harvest of the fish is to happen within fourteen days after the lice count was meant to be conducted.

Fish farms are only required to report louse numbers once every fourteen days when the sea temperature is below 4 °C. They are not required to report louse numbers if exception 2) is the case due to fish welfare concerns when counting fish in cold temperatures. Because of the two exceptions, missing data values can occur during winter time and close to harvesting time. If the fish are harvested in the winter at low temperatures, lice counts can be missing for as much as four weeks

before the fish are being harvested.

The reports of manual counts need to categorised by life cycle stage:

1) fixed stages

2) mobile stages

3) adult females lice

All stages listed here are described in Section 2.1. A sample size of a minimum of ten fish from each pen of a fish farm location needs to be counted. Exceptions to this rule are described under (Forskrift om lakselusbekjempelse, 2013):

- Exception 1: starting at Monday week 14 until and including Sunday week 21, a sample size of twenty fish per pen is required at all locations from Nord-Trøndelag and southwards.
- Exception 2: starting at Monday week 19 to and until and including Sunday week 26, the same rules as exception 1 applies to all locations in Nordland, Troms and Finnmark.

Exception 1 applies for Location 6, whereas exception 2 applies for the remaining locations.

Delousing

Strict rules for sea louse thresholds are enforced by Norwegian law (Forskrift om lakelusbekjempelse, 2013). A limit of 0.5 adult female lice per fish is the common limit for most locations most times of the year. Stricter limits are in place during spring when wild salmon smolts migrate from their natal freshwater habitats into the open ocean (Hatlem). Norway is divided into geographical areas when deciding general louse limits (Forskrift om lakselusbekjempelse, 2013):

- In Nord-Trønderlag and southwards, the upper louse limit is an average of 0.2 adult female lice per fish from and with Monday week 16, to and with Sunday week 21. An upper limit of an average of 0.5 adult female lice per fish is valid the rest of the year.
- In Nordland, Troms and Finnmark, a limit of 0.2 adult female lice per fish is set from and with Monday week 21, to and with Sunday week 26. The 0.5 limit is valid for the rest of the year.

If the number of adult female lice in a fish farm location exceeds any of these limits, the fish farm is required by law to use approved delousing methods to reduce louse numbers. If failing to do so, fines and potentially jail time are possible punishments, depending on the severity.

3.1.3 Data collection

In this section, all variables used in the analysis of this thesis will be described. How they were collected, why they may be relevant to describe manual counts, and missing values and how they were dealt with are included.

Barentswatch data

Manual counts

The relevant number of fish to be examined for manual counts from each pen are collected with special nets and collected in such a manner that the sample is as representative of the population as possible. All of the collected fish are then quickly sedated to make it as stress-free as possible for the fish. When sedated, the counting teams will count lice of the different stage groups on each fish and add them together separately.

Manual counts can be performed in different ways, but the principles are the same. First, the counters have to collect the fish. A surrounding net or lift net is used to collect a sub sample of fish from a pen. Fish feed is often used as a way of attracting fish before collecting the subset. From this subset, a handheld net is used to collect ten to twenty fish, depending on the time of the year. The fish caught with a handheld net are then transferred to an anaesthetising tub to sedate the fish prior to sampling. When sufficiently sedated, the fish, one by one, are being inspected by trained site staff. The lice are sedated in the process as well, and some fall off the fish host into the tub. By law, the lice in the tub are to be counted as well. The process of counting the lice in the tub varies. Some have a white cloth at the bottom of the tub so they can count the lice directly in the tub they fell into. Others have a type of sieve that they pour the liquid through, for then to count the lice that are left in the sieve. The manual count procedure as described here has been gathered from a SINTEF 2018 report (Torvaldsen et al., 2018). Some manual counts were missing in the data set.

Location	Amount missing	Week number	Sea temperature [°C]	Observations left
1	0	NA	NA	52
2	1	12	5	50
4	3	11, 13, 15	All < 4	45
6	3	15, 36, 37	6.4, 13.1, 11.9	34

Table 4: *Missing values in the manual count data. Amount missing is the number of missing values in total for each location. Week number is the week numbers of the weeks the values were missing for. Sea temperatures for the weeks where observations were missing is included as well, in the same order as the week numbers are displayed in the table from left to right. Observations left is the number of observations left for each location, after removing the missing values.*

Table 4 shows the amount of missing values. For Location 4, all missing values were due to the exceptions for manual counts when sea temperatures are below 4 °C as mentioned in Section 3.1.2. For Location 6, the observations for week 36 and 37 were missing because of the upcoming harvest in week 38, allowing them to not count the last two weeks prior to harvest. The observations missing for week 12 at Location 2, and for week 15 at Location 6, are missing for unknown reasons. Removing values in time-series data may be problematic, because important correlation structures on a week-to-week basis may get lost or altered. However, both image-based and manual counts were very small for several weeks before and after the missing values occurred, except from week 36 and 37 at Location 6. The low values over several weeks indicated that there was little change in louse pressure in the time periods where observations were missing. The risk of losing or changing important time-related correlation structures was therefore smaller than if the lice numbers were higher and changing significantly every week. All rows containing the week numbers in 4 were thus deleted. The missing observations for week 36 and 37 at Location 6 could be removed without issues, because they were the end points of the data collection period for that location.

Sea temperature

The sea temperatures measured 3m below the surface have been gathered from the governmental website *barentswatch.no*. The sea temperature gets registered together with sea lice abundance estimates each week from fish farmers, and is public information that can be accessed through this

website. No missing values occurred for the sea temperatures. Sea temperatures affect the development speed of lice as seen in Table 1. Therefore, sea temperatures are of importance for lice counts, because increased sea temperatures may explain increased manual counts. The sea temperatures measured 3m below the surface at the different fish farm locations in this assignment range between the lower temperatures of Table 1 and 17 °C.

Louse treatments

Louse treatments in this context are the week numbers of when an anti sea louse treatment has been initiated at the relevant fish farm locations in the data collection periods. The customers of Stingray register the time of treatment in the Stingray system, and the dates have therefore been gathered from Stingray. Time of treatments can also be found on *barentswatch.no*, information from the Stingray platform was checked on *barentswatch.no* as well. After a treatment, the louse numbers are thought to "reset" back to zero or close to zero, and it is possible to see a clear drop in lice numbers directly after treatments. This variable is logical, where TRUE indicates that a treatment found place for a specific week, whereas FALSE means that no treatment was initiated that week.

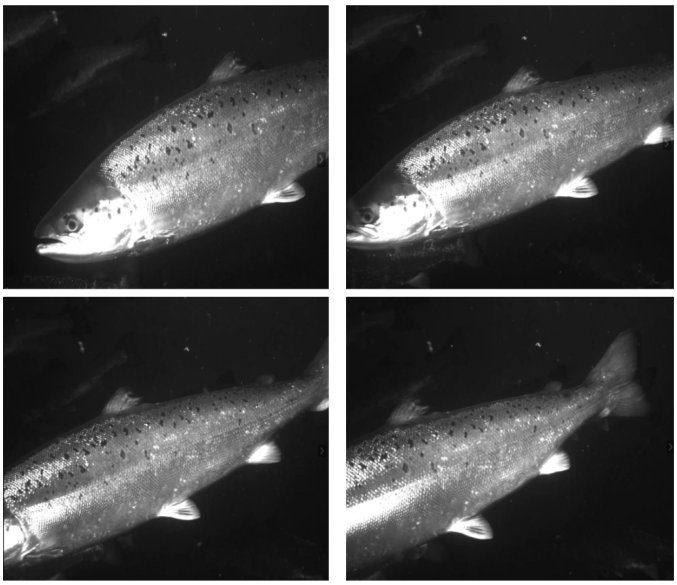
Stingray data

Image-based counts

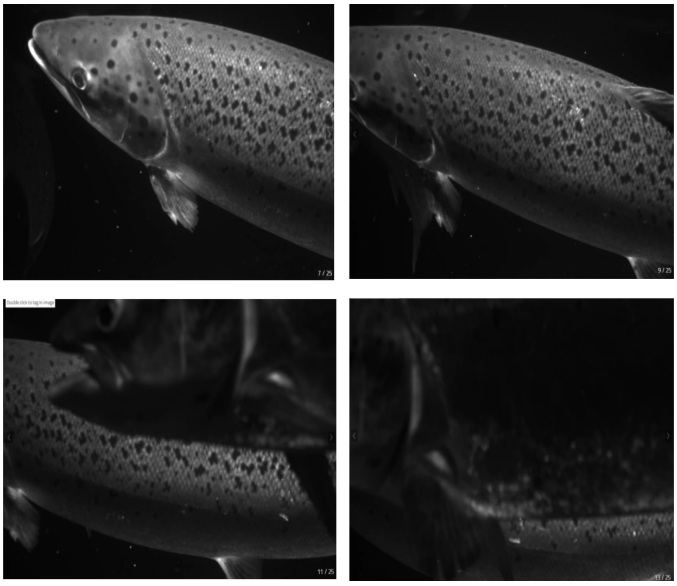
Image-based louse counts in Stingray are based on human inspections of sequences of images. This type of counting is performed by image analysts, who are trained to correctly identify sea lice on images. Unlike the manual counts, image-based counts account for only two groups of salmon lice: adult females and mobiles. The Stingray system uses advanced computation, stereo cameras, and an optical rig with a laser to kill sea lice and collect data about fish, sea lice, and the number of laser pulses shot. These data are stored and sent back to servers in the headquarters in Oslo. Random sequences of images are taken in this way and stored. A minimum of twenty random sequences are analysed per pen at each location, per week. There are two main requirements that need to be fulfilled for a sequence of images to be used as a sample:

1) The sequence needs to show a fish from head to tail, including the end of the tail and the snout of the fish, clearly showing all parts of the fish.

2) The quality of the images needs to hold a certain standard to be used, such that the sea lice can be identified as accurately as possible.



(a) *A short sequence of images of a fish that gets approved as an adequate sample.*



(b) *A short sequence of images of a fish that does not get approved as an adequate sample. The numbers in the bottom right corners show what images from the real sequence of twenty-five images have been chosen to make the simplified illustration.*

Figure 6: *Figure showing an approved sample to be used for image-based lice counts (Panel (a)), and one that gets rejected as an appropiate sample (Panel (b)).*

Figure 6 (a) shows a sequence of four images and is a qualified sequence to be used as a sample. A sequence usually consists of more than four images, but the sequences shown here are shorter to be used as illustrations. A full sequence consists of up to twenty-five images. The quality of the images of Panel (a) is adequate and the whole fish is visible.

Panel (b) shows an example of a sample that gets discarded. The quality of the images is good, but a fish is swimming in front of the fish in focus and it is impossible to see the whole fish. When a sample gets discarded, the analyser will get another random sequence to analyse. When twenty sequences have been analysed, the count is complete, inline with guidelines for purely manual counting (Forskrift om lakselusbekjempelse, 2013).

Location	Amount missing	Week number	Observations left
1	0	NA	52
2	0	NA	50
4	2	22, 29	43
6	1	15	34

Table 5: *Missing values in the image-based count data. Amount missing is the number of missing values in total for each location. Week number is the week numbers of the weeks the values were missing for. Observations left is the number of observations that are left for each location after removing the rows with the missing values (including the missing values from the manual counts).*

Table 5 shows that three observations were missing in total of the image-based counts, two for Location 4 and one for Location 6. The rows in the data set for the week numbers listed were removed. For Location 4, the number of observations was reduced from forty-five to forty-three observations, and for Location 6, the number of observations remained unchanged. The reasoning behind the removal of the observations were the same for manual and image-based counts, and can be found under missing data for the manual counts.

Laser pulses shot

Every time the Stingray laser unit fires a laser pulse, the pulse is registered. Laser pulses shot is the

number of times the Stingray unit pulses in a given time interval. The number of laser pulses has been collected from the Stingray platform *stingrayonline.no* and are weekly sums. The amount of shots can be affected by how many fish there are in a fish-farming location as well as how many lasers there are in each location. This variable has therefore been altered to account for that. The new variable will later be referred to as normalised pulses.

$$Normalised pulses = laser pulses / number of fish / amount of lasers.$$

The amount of lasers has been gathered manually, by reading from the most currently updated maps from each location for the relevant time periods, of the type in Figure 3. There were no missing values of the laser pulses at any of the locations. The normalised laser pulses are of interest in relation to a correction factor mainly for two reasons. Firstly, laser pulses shot are usually going up when there are more lice present. Therefore, it might be a relationship between increased number of pulses and increased manual counts. Secondly, more shots fired may have an effect on manual counts, because lice are being shot and killed.

Fish passings

Fish passings are the number of fish that are swimming past the Stingray cameras. Passings are weekly estimates and are the number of fish swimming past the Stingray system per week. The way fish passings are estimated are through tracking fish eyes within a specified distance from the cameras. There were no missing values of the fish passings at any of the locations in the relevant time periods. This variable can give an indication of how well the laser units are positioned in a pen. If there are almost no fish swimming in front of the cameras over time, it can have an effect on the relationship between image-based counts and manual counts because it may affect the sampling process of the image-based counts.

Stingray customer data

Fish weight

Fish weight has been collected by each of the remaining four customers that helped provide data for this assignment. An estimate of the average of the fish weight has been used. The weight estimates are measured weekly. The weight is estimated by calculating a theoretical growth of the fish by the

amount of fish feed fed. This value is typically provided by the fish feed manufacturers in so called feeding tables. No missing values occurred for this variable. Fish weight is relevant for describing the manual counts because bigger fish may have more lice. Therefore, an increase in weight may affect louse numbers.

Number of fish

Number of fish is a measurement provided by the customers. The number of fish are weekly estimates. The fish farmers estimate number of fish by deducting all fish losses from a starting number of fish when first stocking a pen at the start of a production of a production cycle. This variable was used as it is as a variable, and as a way to normalise the laser pulses. No missing values occurred for this variable. Number of fish in a location may affect lice numbers because more fish can have more lice. On the other hand, lice counts are reported as lice per fish, it might therefore be more lice per fish observed when there are less fish.

3.1.4 Estimates of salmon louse populations

Both image-based and manual louse counts are estimating the louse populations based on averages from counts. The values in the data sets for both image-based and manual counts are reported as lice per fish. Both counting methods are on location level, not on pen level. Fish are being counted from all pens, but the population mean in this case is the total mean for the whole location.

For the image-based counts, the analyser will count sea lice of different groups as described in Section 3.1.3. Total number of the different louse stage groups are then divided by the sample size to get the number of lice per fish, of each group. This method is used on all the pens at each location, where a sample of a minimum of twenty sequences of images are drawn from each pen. The averages of each pen are then added together and divided by the number of pens, to get a total average of each group. The averages of adult females are the measurements used in the data sets for both manual and image-based counts.

Similarly, the manual estimates are conducted by averaging the averages of each louse stage group from each pen over the whole location. The counting teams will sedate the minimum of ten random

fish and count lice of the three different groups described in Section 2.2 in all pens at each location. The total numbers of each group in every pen are divided by sample sizes, to get an estimate of lice per fish. Averages of every pen are then added together and divided by the number of pens to get an estimate of the louse situation in the whole location.

3.2 Mixed-effects models

Mixed-effects models are powerful and versatile tools for analysing grouped data. Grouped data appears in many fields like medicine, biology and economics. Longitudinal data being measured at multiple locations is an example of grouped data, which is the type of data analysed in this assignment. Mixed-effects models are practical tools when within-group correlation in a grouped dataset is high and the between-group correlation is low. High within-group correlation between a response variable Y and an independent variable X indicates that X can predict Y, but not necessarily across the groups (Marzban et al. 2013). The manuscript of *Mixed-Effects Models in S and S-PLUS* (Pinheiro and Bates, 2006) has been heavily referenced in other books and articles dealing with mixed-effects modeling. In addition, the `nlme` library in R is based on a lot of information from that book, which is the library used for analysing grouped data in this assignment. Because of the many citations, the theory in the Method section will be heavily based on Pinheiro and Bates (2000). The sources of information from other literature will be mentioned or outlined within parentheses.

Linear Mixed-Effects Models (LMMs) are a versatile statistical framework used to analyse data with complex structures, such as repeated measurements, hierarchical data, or clustered observations. These models extend traditional linear regression by incorporating both fixed effects, which capture population-level trends, and random effects, which account for within-group variability. In this section, we explore the key components and principles of LMMs as outlined by Pinheiro and Bates (2000), both the theoretical approach and practical implementation.

3.2.1 Linear mixed-effects models

Theoretical implementation

At the core of LMMs is the linear relationship between the response variable ($\mathbf{y}$) and a combination

of fixed and random effects:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{X}_i\beta + \mathbf{Z}_i\mathbf{b}_i + \varepsilon_i, \qquad i = 1, ..., M, \\ \mathbf{b}_i &\sim N(\mathbf{0}, \mathbf{\Sigma}), \qquad \varepsilon_i \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I}) \end{aligned} \tag{1}$$

where:

- $\mathbf{y}_i$ is the n_i-dimensional vector of observed responses for the ith group.
- $\mathbf{X}_i$ is the known $n_i \times p$ regressor matrix for so-called fixed effects, and $\boldsymbol{\beta}$ represents the p-dimensional vector of fixed effect coefficients.
- $\mathbf{Z}_i$ is the known $n_i \times q$ regressor matrix for so-called random effects, and $\boldsymbol{b}_i$ represents the q-dimensional vector of random effects. The random effects $\boldsymbol{b}_i$ are assumed to be normally distributed with mean $\mathbf{0}$, and is characterised by by its variance-covariance matrix $\mathbf{\Psi}$.
- $\boldsymbol{\varepsilon}_i$ represents the n_i-dimensional vector of within-group errors, assumed to follow a spherical Gaussian distribution with mean zero and a covariance matrix as shown in Equation (1). We describe the terms fixed/random effects in the following sections.

Fixed Effects ($\boldsymbol{\beta}$)

Fixed effects capture the systematic and population-level variation in the data. The model assumes that the fixed effects are constants that apply to the entire population. The goal is to estimate the values of $\boldsymbol{\beta}$ that best explain the observed variation in the response variable $\mathbf{y}$.

One estimation method to estimate the fixed effects involves minimizing the residual sum of squares (RSS) to find the values of $\boldsymbol{\beta}$ that minimize the discrepancy between the observed data and the model predictions:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 \right\}$$

However, for this assignment, maximum likelihood and restricted maximum likelihood estimation will be used for this purpose. Both methods will be described in detail in two separate sections.

Random Effects ($\boldsymbol{b}_i$)

Random effects account for the unexplained variability within groups or clusters in the data. Unlike fixed effects, random effects are assumed to follow a multivariate normal distribution with a mean vector of zeros and a covariance matrix that characterizes the within-group variability.

The estimation of random effects involves finding the values of $\boldsymbol{b}_i$ that maximize the likelihood of the observed data, given the estimated fixed effects $\hat{\boldsymbol{\beta}}$:

$$\hat{\boldsymbol{b}_i} = \arg\max_{\boldsymbol{b}_i} \left\{ L(\boldsymbol{b}_i | \mathbf{y}_i, \mathbf{X}_i, \mathbf{Z}_i, \hat{\boldsymbol{\beta}}) \right\}$$

By estimating fixed effects ($\boldsymbol{\beta}$) and random effects ($\boldsymbol{b}_i$) while considering the underlying assumptions of multivariate normality and covariance structure, LMMs provide a valuable tool for statistical modeling and inference.

Practical implementation

The `nlme` library in R has been used to fit the linear mixed-effects models. The specific function used for this is the `lme(fixed, data, random, correlation, weights, method, subset,...)` function. It is a generic function that fits a linear mixed-effects model as described in the theoretical implementation of the mixed-effects model part. Fixed effects can be specified as a two-sided linear formula object, with the response on the left of a '$\sim$' operator. The independent variables are separated by '+' operators on the right side. Random effects can be specified as a one-sided formula of the form '`random = x1 + x2 +...+ xn | g1/g2/.../gm`', where `x1 +...+ xn` specify the model for the random effects and `g1,..., gm` are the grouping structure. A random intercept model could for example be expressed as `1|g`, where `g` is just a one-level grouping factor in this case.

To extract valuable information, like the coefficient estimates of the fixed effects or the standard deviation estimates for the random effects, the function `summary(object,...)` is useful. When the `object` is a LME model, the function outputs valuable information like mentioned above as

well as AIC, BIC and log-likelihood values, p-values corresponding to the fixed effects and estimates of the additional parameters of the alternative correlation and variance-covariance structures when these are specified.

3.2.2 Maximum likelihood estimation

Estimating the model parameters for LMMs, both fixed and random, is crucial for making inferences and predictions. Maximum Likelihood Estimation (MLE) is a widely used method for parameter estimation in LMMs, and in this section, we delve into the details of MLE in the context of LMMs.

Theoretical implementation

Maximum Likelihood Estimation seeks to find the parameter values that maximize the likelihood function given the observed data. In the context of LMMs, we assume that the observations $\mathbf{y}$ follow a multivariate normal distribution with mean $\mathbf{X}\beta$ and covariance matrix $\mathbf{R}+\mathbf{ZDZ}^T$, where $\mathbf{D}$ represents the covariance structure of the random effects.

The likelihood function for LMMs is:

$$L(\beta,\mathbf{u},\boldsymbol{\varepsilon}|\mathbf{y},\mathbf{X},\mathbf{Z}) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}+\mathbf{ZDZ}^T|^{1/2}} \times \exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{X}\beta-\mathbf{Zu})^T(\mathbf{R}+\mathbf{ZDZ}^T)^{-1}(\mathbf{y}-\mathbf{X}\beta-\mathbf{Zu})\right),$$

where:

- n is the number of observations.
- $|\cdot|$ denotes the determinant of a matrix.

MLE for Fixed Effects (β)

To estimate the fixed effects β, we maximize the likelihood function by finding the values of β that maximize $L(\beta,\mathbf{u},\boldsymbol{\varepsilon}|\mathbf{y},\mathbf{X},\mathbf{Z})$. This is typically done using numerical optimization algorithms like

the Newton-Raphson method or the Expectation-Maximization (EM) algorithm.

The MLE for β is the set of values that maximizes the likelihood:

$$\hat{\beta} = \arg\max_{\beta} L(\beta, \hat{\mathbf{u}}, \hat{\boldsymbol{\varepsilon}} | \mathbf{y}, \mathbf{X}, \mathbf{Z})$$

MLE for Random Effects ($\mathbf{u}$*)*
Estimating the random effects $\mathbf{u}$ is more challenging because they are not directly observed. Instead, we estimate them as part of the joint optimization problem. The MLE for $\mathbf{u}$ can be found through the restricted profile likelihood:

$$L_R(\mathbf{u}|\mathbf{y}, \mathbf{X}, \mathbf{Z}, \hat{\beta}) = L(\hat{\beta}, \mathbf{u}, \boldsymbol{\varepsilon}|\mathbf{y}, \mathbf{X}, \mathbf{Z}),$$

where $\hat{\beta}$ is the MLE of the fixed effects obtained in the previous step.

The key steps involve of maximum likelihood estimation is maximizing the likelihood function with respect to fixed effects β and estimating random effects $\mathbf{u}$ using restricted profile likelihood.

Practical implementation
Within the `lme()` function, the parameter estimation method can be specified to maximum likelihood by writing '`method = "ML"`'. The default is restricted maximum likelihood, hence it is important to specify the method if another estimation procedure is to be used.

3.2.3 Restricted maximum likelihood estimation

Restricted maximum likelihood Estimation (REML) is a statistical method used for parameter estimation in linear mixed-effects models. In this section, we explore the principles and techniques of REML estimation within the context of LMMs.

Theoretical implementation

REML estimation is a variant of maximum likelihood estimation (MLE) that focuses on estimating the variance-covariance parameters of the random effects while "marginalizing out" the fixed effects. Unlike MLE, which estimates both fixed and random effects jointly, REML only estimates the parameters of the random effects. The key idea is to construct a likelihood function that conditions on the observed data and the estimated fixed effects, resulting in a "restricted" likelihood function.

The REML likelihood function is constructed by integrating out the fixed effects, yielding a likelihood function for the random effects parameters only:

$$L_{\text{REML}}(\mathbf{u}|\mathbf{y},\mathbf{X},\mathbf{Z}) = \int L(\beta,\mathbf{u},\boldsymbol{\varepsilon}|\mathbf{y},\mathbf{X},\mathbf{Z})\, d\beta$$

Estimating $\mathbf{u}$ in this context provides a measure of the "conditional" variance-covariance parameters of the random effects, which are adjusted for the uncertainty in the fixed effects estimates.

REML Estimation Procedure

The REML estimation procedure involves the following steps:

1. Specify the LMM, including the fixed and random effects structures.
2. Condition on the observed data and estimate the fixed effects $\hat{\beta}$.
3. Compute the restricted likelihood $L_{\text{REML}}(\mathbf{u}|\mathbf{y},\mathbf{X},\mathbf{Z})$ by integrating out the fixed effects using numerical techniques.
4. Maximize $L_{\text{REML}}(\mathbf{u}|\mathbf{y},\mathbf{X},\mathbf{Z})$ to obtain the REML estimates of the random effects $\hat{\mathbf{u}}$.

Restricted Maximum Likelihood Estimation (REML) is a valuable method for estimating the variance-covariance parameters of random effects in Linear Mixed Effects Models (LMMs). By focusing on the conditional likelihood of the random effects while marginalizing out the fixed effects, REML provides robust parameter estimates that are especially useful when the primary interest is in modeling the within-cluster variability.

Practical implementation

Restricted maximum likelihood estimation is the default within the `lme` function. If you want to specify it either way, it can be done by specifying the method as '`method = "REML"`'.

3.2.4 Correlation and variance-covariance structures

The relevant correlation and variance-covariance structures to account for within-group errors, as well as heteroskedasticity and unequal variance in the model residuals across groups in the data in this thesis, will be presented in the following section. Both theoretical and practical implementation of the structures are included.

Theoretical implementation

Correlation structures

Correlation structures in linear mixed-effects models are used to model the dependence among within-group errors. We assume here that within-group errors ε_{ij} are associated with integer scalar position vectors $\boldsymbol{p}_{ij}$. The correlation between two within-group errors, ε_{ij} and $\varepsilon_{ij'}$, is assumed to depend on the corresponding position vectors $\boldsymbol{p}_{ij}$ and $\boldsymbol{p}_{ij'}$ through some distance between them, and not through the exact values of ε_{ij} and $\varepsilon_{ij'}$. The distance can be denoted $d(\boldsymbol{p}_{ij}, \boldsymbol{p}_{ij'})$. The general within-group correlation structure for single-level grouping is expressed by Pinheiro and Bates (2000) as

$$cor(\varepsilon_{ij}, \varepsilon_{ij'}) = h[d(\boldsymbol{p}_{ij}, \boldsymbol{p}_{ij'}), \boldsymbol{\rho}], \quad i = 1, ..., M \quad j, j' = 1, ..., n_i, \tag{2}$$

where $\boldsymbol{\rho}$ is a vector containing correlation parameters and $h()$ is a correlation function that can take on a value between -1 and 1. A value close to 1 or -1 indicates a strong linear relationship whereas a value close to 0 indicates a weaker relationship. The function $h()$ is assumed to be continuous in $\boldsymbol{\rho}$, in such a way that $h(0, \boldsymbol{\rho}) = 1$. The 0 in $h(0, \boldsymbol{\rho})$ means that the distance between the positions of ε_{ij} and $\varepsilon_{ij'}$ is 0, and they must therefore be the same observation and their correlation must be equal to 1. Now let $r_{ij} = \frac{y_{ij} - \hat{y}_{ij}}{\hat{\sigma}_{ij}}$ be the standardised residuals of a fitted linear mixed-effects model. The empirical autocorrelation function at lag l is then defined as

$$\hat{\rho}(l) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{n_i - l} r_{ij} r_{i(j+l)} / N(l)}{\sum_{i=1}^{M} \sum_{j=1}^{n_i} r_{ij}^2 / N(0)}, \tag{3}$$

where $\sigma_{ij}^2 = Var(\varepsilon_{ij})$ and $N(l)$ is the number of residual pairs used in the numerator. The empirical autocorrelation function is a nonparametric estimate of the autocorrelation function.

The data in this assignment is time-series data of longitudinal form. The observations have been gathered weekly and are indexed by a one-dimensional position vector, where its values correspond to the week numbers. Serial correlation structures can be used to model dependence in time-series data. Autoregressive-moving average models can be useful tools to model such dependence, and consist of three different classes of linear stationary models: autoregressive models, moving average models and a mixture of the two. They all assume that data is observed at discrete time points, for example every week. Let x_t denote an observation. The time step, or the lag, between two observations x_t and x_s is given by $|t-s|$. The definition of an autoregressive process of order p, as described in Compertwait and Metcalfe (2009), is the following:

The series $\{x_t\}$ is an autoregressive process of order p, abbreviated to $AR(p)$, if

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + ... + \alpha_p x_{t-p} + w_t, \tag{4}$$

where $\{w_t\}$ is white noise and the α_i are the model parameters with $\alpha_p \neq 0$ for an order p process.

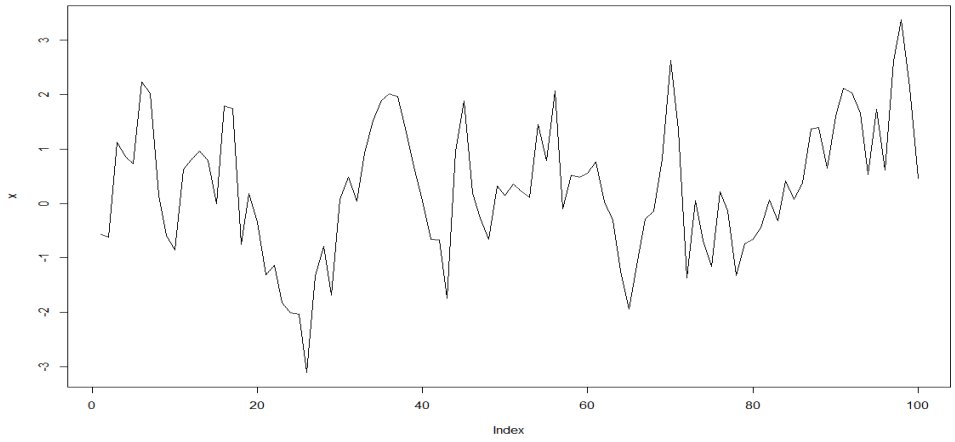
An $AR(1)$ model can hence be expressed as

$$x_t = \alpha x_{t-1} + w_t. \tag{5}$$

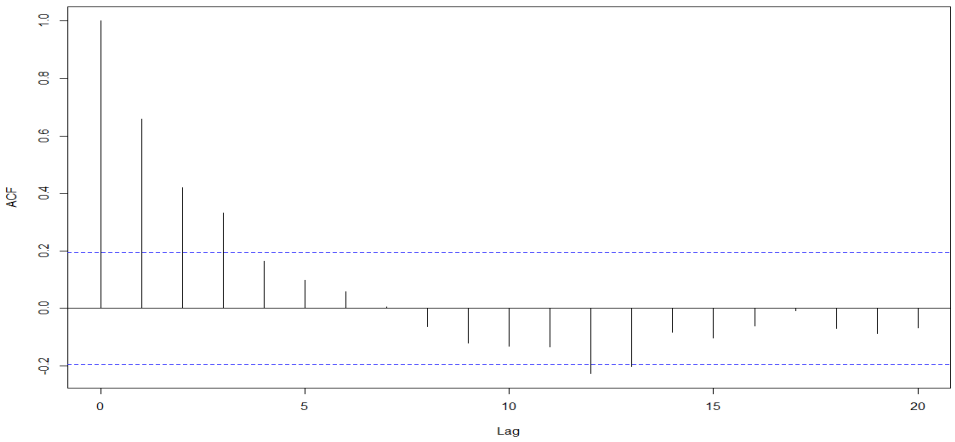
The white noise term w_t has mean 0 and is assumed independent of the previous observations. The correlation function of the $AR(1)$ process is given by

$$h(\alpha, k) = \alpha^k, \quad k = 1, 2, ... \tag{6}$$

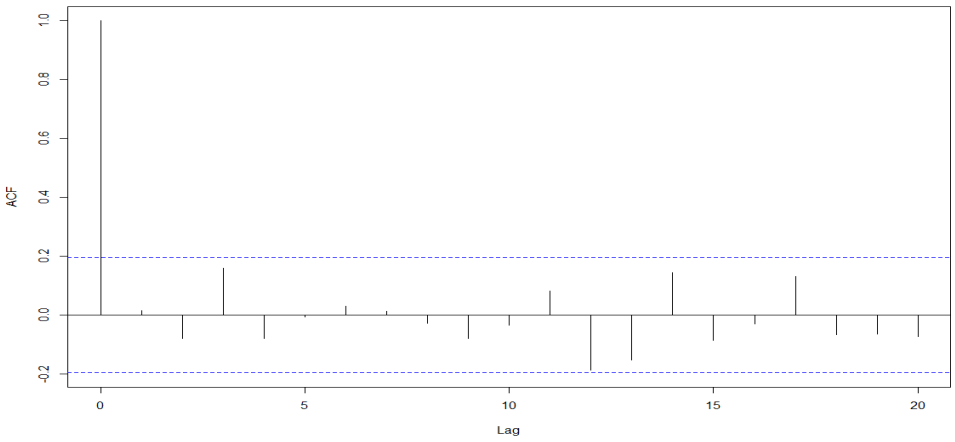
The function $h(\alpha, k)$ decreases in absolute value as the lags increases and decays to 0 quicker with smaller α. A time series is informally said to be stationary if it does not have a trend or a seasonal effect.



(a) *Simulated AR(1) time series, called x on the y-axis.*



(b) *Empirical auto correlation plot of the raw simulated AR(1) time series data in (a).*



(c) *Empirical auto correlation plot of the residuals of an AR(1) model fitted on the simulated data.*

Figure 7

Figure 7 shows three plots. Panel (a) shows a simulated stationary time series x with the following parameterisation:

$$x_t = 0.7x_{t-1} + w_t,$$

where the initial x_0 is equal to -0.560 and $w_t \sim N(0,1)$. Panel (b) shows the empirical auto correlation function of the data in Panel (a). There are significant values at several lags (significant values are values above or below the dashed blue lines), whereas the greatest value is observed at lag 1. Panel (c) shows the first twenty lags of the empirical auto correlation function of the residuals of a fitted $AR(1)$ model, fitted on the simulated data as described above. The plot in Panel (c) shows no significant values at any lags > 0 as expected. This indicates that the residuals of the $AR(1)$ model are uncorrelated, and that an $AR(1)$ model might be a good fit, as it should be.

Variance-covariance structures

It is possible to relax the assumption for the variance-covariance structure for the within-group errors in Equation (1). If the variance is different between groups, or heterogeneity is present in the standardised residuals of an LMM, it might be helpful with alternative variance structures. For the case when variance is different between groups, we may express the error term in Equation (1) as $\varepsilon_{ij} \sim N(0, \sigma_j^2)$. This allows j different variances when there are j groups.

The variance can also be modeled as the power of a covariate. This allows modeling heteroscedasticity of residuals that might be related to the level of some covariate. The theoretical formula of such a variance structure can be expressed as the following:

$$Var(\varepsilon_{ij}) = \sigma^2 |covariate_{ij}|^{2\delta},$$

where $covariate_{ij}$ is the value of the covariate for the i-th subject at the j-th level of the grouping factor. The delta δ is the power parameter that determines how much the variance of the residuals is affected by the covariate. This gets estimated from the data. It is possible to let the δ vary across groups.

Practical implementation

The within-group errors are allowed to have different correlation and variance-covariance structures, and are easily incorporated in the 'correlation' and 'weights' within the lme function. There are multiple different correlation and variance-covariance structures to choose between that are inbuilt functions to be used within the lme function. The relevant structures for this thesis are however the functions 'corAR1()' for the correlation structure, and 'varIdent()' and 'varPower()' for the variance-covariance structures, as described in the theoretical part. The correlation function is simply incorporated as 'correlation = corAR1()' within lme.

For the case where we simply want to allow the variance to be different between groups, we can implement this as 'weights = varIdent(form = $\sim$ 1|Location)', where Locations are our groups. This specification allows for different variance between Locations, and is one of the variance structures we use in the model building process in this thesis. The other structure we use, is the power function. We can specify the weights as 'weights = varPower(form = normPulses|Location)' within the lme function. In this case, the variable normPulses is the covariate, and it is allowed to be different between Locations.

3.2.5 Model selection

It is important to evaluate models and assess whether or not they are a good fit. This section will be about the methods we will use for evaluating candidate models later on. Both theoretical and practical implementation in R, of the methods in the following section will be included. A significance level of 0.05 will be used to evaluate p-values in relation to all hypothesis tests in this thesis (coefficient estimates, likelihood ratio tests etc).

Theoretical implementation

AIC and BIC

Analysing data using different modeling techniques usually includes making multiple candidate models to evaluate for the best fit. Two methods for evaluating the better fit of several candidate models are by using Aikake information criterion (AIC) (Sakamoto et al., 1986) or Bayesian information criterion (BIC) (Schwarz, 1978). Lower value of either AIC or BIC indicates a better

fitting model. A low value in itself does not reveal anything about the model fit in general, but it can assist in choosing the best model of multiple candidates relative to each other. AIC and BIC can be expressed:

$$AIC = 2k - 2ln(\hat{L})$$

$$BIC = kln(N) - 2ln(\hat{L})$$

where k is the number of estimated parameters and $\hat{L}$ is the maximum value of the likelihood function for the model. BIC includes the number of observations N of the observed data. It is important to specify the method used for estimating parameters in a model. When comparing models using AIC or BIC, the models need to be fit using the ML method. It is possible to calculate the AIC and BIC with the REML method, but it is only possible to compare models with the same fixed effects structure using this method. The author of (Zuur et al., 2009) uses ML estimation in all models until a proper model is picked. They change the method to REML after the model selection process to obtain the final parameter estimates, which is what will be done in this thesis as well.

Likelihood ratio test

The likelihood ratio test can be used to assess the best fit of two candidate models relative to each other, based on the ratio of their likelihoods. The likelihood ratio statistic λ_{LR} is given by

$$\lambda_{LR} = 2log(\frac{L_2}{L_1}) = 2[log(L_2) - log(L_1)]. \tag{7}$$

It is here assumed that L_2 is the likelihood of the more general model of the two, whilst L_1 is the likelihood of the restricted model, meaning $L_2 > L_1$ and λ_{LR} is positive. It is not uncommon to see λ_{LR} expressed as $-2[log(L_1) - log(L_2)]$, which is the same expression. The likelihood ratio statistic statistic λ_{LR} follows the χ^2-distribution with $d_2 - d_1$ (numbers of parameters in the models) degrees of freedom under the null hypothesis that the simpler model is satisfactory, when parameters are fitted with maximum likelihood. A low p-value (usually below 0.05) associated with the likelihood ratio statistic therefore suggests that the null hypothesis should be rejected.

Backward and forward selection

Building a model includes excluding non-significant predictors based on some criteria. The goal is to build a model with as few attributes as possible, without compromising the predictive ability of the model. In this thesis, BIC will be used as an exclusion criteria. Methods to determine what variables that are insignificant include backward and forward selection, or a combination of the two. In this section, backward and forward selection will be described.

Backward selection starts with a full model. In this case, that is a model that includes image-based counts, normalised laser pulses, sea temperature, fish weight, number of fish and relevant interaction terms between two or more variables. From the full model, one predictor is removed at the time, based on what variable removal that yields the biggest improvement in BIC values. This process is repeated until removing a variable no longer favourably changes the BIC value. When removing variables no longer result in lower BIC, the backward selection process stops.

Forward selection operates the opposite way of backward selection. In the forward selection process, the initial model contains an intercept only. One predictor is added to the model at the time, based on what variable yields the biggest improvement in BIC values. When adding more variables no longer result in a lower BIC value, the forward selection process stops.

Both backward and forward selection were used as methods to help determine what variables to exclude from the models in this thesis. The results were the same for both methods. Therefore, only backward selection is included in the Result section.

Practical implementation

In R, the AIC, BIC and likelihood ratio test can be carried out by the `anova()` function. This function takes two or more objects containing the results returned by a model fitting function such as linear models (lm). The output of the `anova` function shows AIC and BIC values, and the statistics needed to carry out the likelihood ratio test. The AIC and BIC values can also be calculated directly for a model through the `AIC()` and `BIC()` functions.

3.2.6 Fitted values

Theoretical implementation

Fitted values are the predicted values of the observed responses under the fitted model. They can be used for evaluating the general goodness of fit for a model. A well fitted model will also predict new observations well. In mixed-effects models, fitted values and predictions may be obtained on different nesting levels or on the population level. As described by Pinheiro and Bates (2000), letting $\boldsymbol{x}_h$ be a vector of fixed covariates, the marginal expected value of the corresponding response y_h (at the population level) is given by:

$$E[y_h] = \boldsymbol{x}_h^T \boldsymbol{\beta}. \tag{8}$$

The conditional expectation of the response given the random effects at levels $\leq k$ are estimates of the predicted values at the kth level of nesting. At the first level of nesting, let $\boldsymbol{z}_h(i)$ denote a covariance vector corresponding to the random effects at the ith group. The conditional expectation is then given by:

$$E[y_h(i)|\boldsymbol{b}_i] = \boldsymbol{x}_h^T \boldsymbol{\beta} + \boldsymbol{z}_h(i)^T \boldsymbol{b}_i. \tag{9}$$

The best linear unbiased predictors can be obtained by replacing the $\boldsymbol{\beta}$ and $\boldsymbol{b}_i$ in Equation (2) and (3) with $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ for $\boldsymbol{\beta}$ and $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$ for $\boldsymbol{b}_i$. The unknown parameter vector $\boldsymbol{\theta}$ can be replaced by its MLE or REML estimate, which produces estimated best linear unbiased predictors for the expected values.

Practical implementation

Fitted values of a model can easily be extracted for simple linear models through the '`$`'-operator in R. By writing '`model$fitted`', where model is some specified model, fitted values can be obtained. For mixed-effects models, it is possible to extract predictions by excluding or including random effects. This can be specified through the `'level'` function within the `predict()` function. Level can be set to zero or one, where zero excludes the random effects, and one includes them.

3.3 Goodness of fit

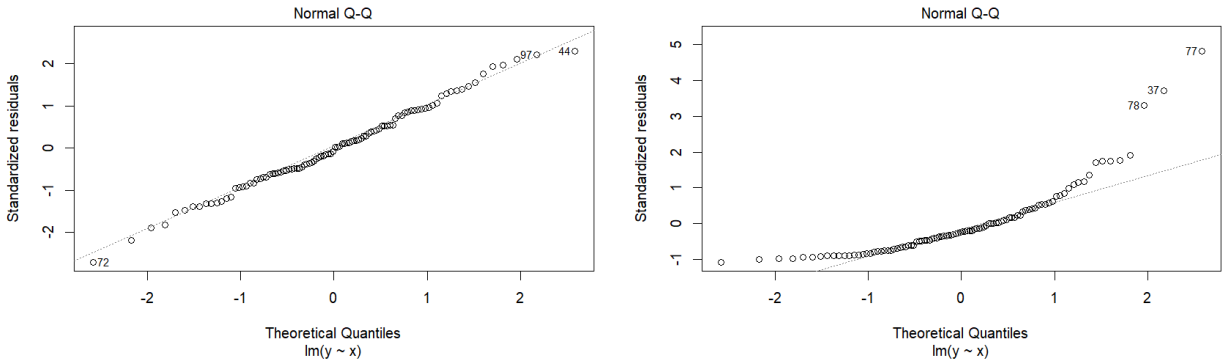
Theoretical implementation

Assumptions in linear mixed-effects models

Making inferences about a fitted model is meaningless if the model is ill-fitting. It is possible to get a long way in evaluating the goodness of fit of a model by making diagnostic plots and visualising results. There are numerous tests and more formal ways to test that some criteria for well-fitted models are met as well. Certain assumptions about the underlying distribution of the data need to be valid for a properly fitted model. There are two assumptions for the simple linear mixed-effects model:

- Assumption 1: the first assumption is made about the within-group errors. They are assumed to be normal i.i.d.d. with mean zero and common variance σ^2, and being independent of the random effects.
- Assumption 2: the second assumption is about the random effects. The random effects are assumed to be following a normal distribution with mean zero and covariance matrix $\mathbf{\Psi}$, and being independent for each group.

There are formal hypothesis tests suitable for checking the validity of these assumptions, but a more intuitive approach may be to plot residuals, fitted values and the estimated random effects and inspect them all. Plotting the residuals for each group can give a quick indication of whether or not the means of the residuals are centered around 0, and whether or not the residuals of each group have constant variance. Making a plot, for example a box plot, can be a powerful tool for visualisation, but if there are few observations in each group, strong conclusions about the variances can not be drawn from plots alone. If the number of observations are reasonably large, and a clear pattern is found in the residuals of each group, for example being larger for two locations than any other locations, the variance pattern could be accounted for in the model through a variance function. Residuals centered around a value $\neq 0$ can be shifted by adding a constant to the fixed effect term.



(a) *A normal Q-Q plot with a good fit.*

(b) *A normal Q-Q plot with a bad fit.*

Figure 8: *Panel (a) is a Q-Q plot of a well-fitting linear model, with normally distributed residuals. Panel (b) is an ill-fitting linear model with non-normal residuals. The dotted line is a reference line. The data used to produce the plots was from self-simulated variables in R.*

Figure 8 shows two panels, both containing Normal Q-Q plots of simulated data. Normal plots like a Q-Q plot of the estimated random effects for each group can be used to evaluate whether the random effects are following a normal distribution or if they are far away from that. Figure 8 shows two normal Q-Q plots for illustrative purposes. Both plots show normal Q-Q plots, which in this case are plots of standardised residuals of a fitted linear model versus the theoretical quantiles of a standard normal distribution. A Q-Q plot of a well-fitting linear model should have points laying close to the reference line. Panel a) shows a Q-Q plot of a well-fitting linear model, fitted on simulated normally distributed data. The points are mainly laying close to the dotted line, with some outliers. Panel b) on the other hand shows a Q-Q plot of an ill-fitting linear model, where the response variable follows an exponential distribution (simulated exponentially distributed data) and has non-normal residuals. The points in Panel 8 b) are not mainly on the reference line, and we can already get a hint of the distribution of the response variable from the plot. It is the same concept for Q-Q plots of estimated random effects.

More formal tests for testing the assumption of normality of the estimates of the random effects exist. They can often be complicated, and the importance of visualising data still stands, as showed

by Anscombe's quartet, amongst other.

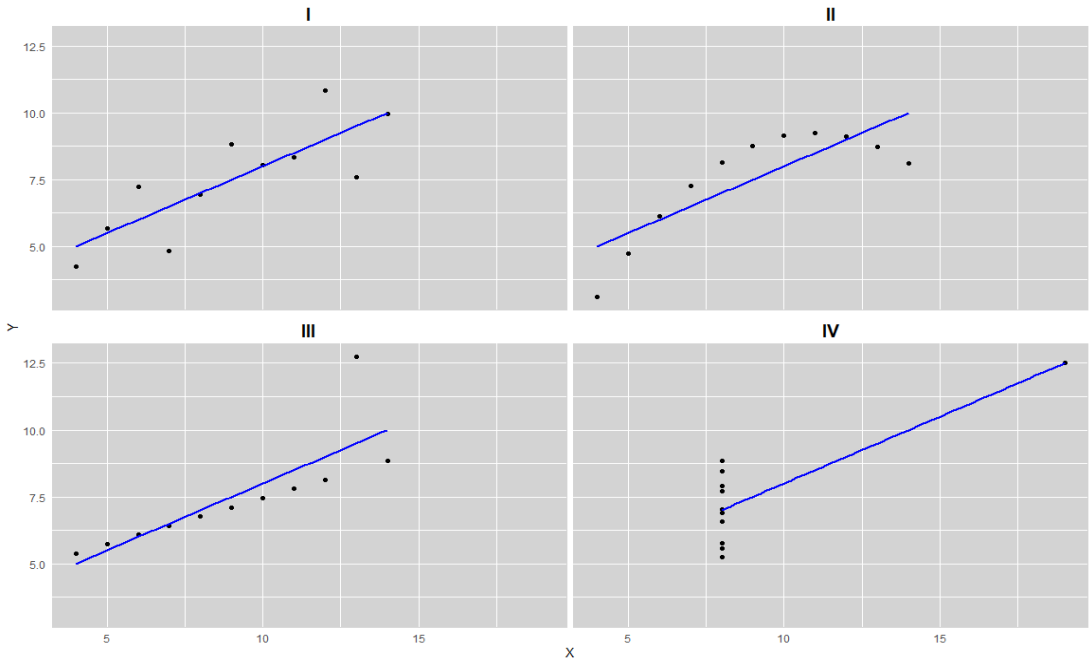


Figure 9: *The Anscombe's quartet plots. The Anscombe data set can be found in the R library, and the plots have been produced with this data set.*

Anscombe's quartet as shown in Figure 9 is a series of four very different data sets with different underlying distributions, but with the same summary statistics after fitting a linear model to them all.

Property	Value	Accuracy
Mean of x ($\bar{x}$)	9	Exact
Standard deviation of x (s_x^2)	11	Exact
Mean of y ($\bar{y}$)	7.50	To three decimal places
Standard deviation of y (s_y^2)	4.125	$\pm$ 0.003
Linear regression line	0.30 + 0.500x	To two and three decimal places
R^2	0.66	To two decimal places

Table 6: *Summary statistics of all of the Anscombe data sets. Property is the type of values being measured, value is the value of each property and the accuracy displays to what degree the values for each data set agreed with each other. The R^2 is the coefficient of determination.*

Table 6 shows that only studying summary statistics of data sets may be problematic. The summary statistics for four different data sets are almost identical, although three out of four data sets should not be modelled through linear regression. Plots can be problematic when there are few levels of a group, say if the observations are gathered from four locations and we want to plot the intercepts for models of each level. A plot consisting of four points will struggle to relay a proper pattern, and the pattern will be heavily influenced by each individual point. Inferences should not be drawn from plots alone, but should be reviewed as an aid for model selection and indications of potential variance heterogeneity.

Coefficient of determination and Mean squared error

The coefficient of determination also called R-squared, or R^2, is a statistical measure that measures the proportion of variance in a dependent variable that is being explained by an independent variable. This measure can take a value between zero and one, where zero means that the independent variable explains nothing of the variance, and a value of one means that all the variance is explained by the independent variable. The R^2 value is a useful metric to evaluate goodness of fit in a linear model. The R^2 coefficient is calculated by looking at the relationship between the residual sum of squares (SSR) and total sum of squares (SST). The SSR is defined for a model with one independent variable by:

$$SSR = \sum (y_i - \hat{y}_i)^2,$$

the SST is defined by:

$$SST = \sum (y_i - \bar{y})^2,$$

where y_i is the dependent variable, $\hat{y}_i$ is the predicted value of y_i, and $\bar{y}$ is the mean of y. The R^2 coefficient can be calculated by

$$R^2 = 1 - \frac{SSR}{SST}.$$

If SSR is equal to zero, the fraction will become equal to zero, and the R^2 is equal to one. On the other hand, if the SSR is equal to the SST, the expression will be equal to one, and the R^2 coefficient is equal to zero.

The mean squared error (MSE) measures the average of the errors in predictions squared. The MSE in the context of predictions can be used as a tool to determine how well a model performed on unseen data. A high value of MSE indicates poor predictions in a sample. Mean squared errors can be calculated as

$$MSE = \frac{1}{n}\sum(y_i - \hat{y}_i)^2,$$

where y_i are the true values of the dependent variable and $\hat{y}_i$ are the predicted values of y_i.

Practical implementation

Normal Q-Q plots can be produced in R by using the `qqnorm()` and `qqline()` functions. The `qqnorm()` function produces the Normal Q-Q plot for a variable, whereas `qqline()` produces a reference line. The R^2 and MSE coefficients can be calculated as described above, with standard operators in R. The concentration can be calculated with the `mean()` function and the predicted values can be found using the `predict()` function.

3.4 Bootstrap-type Monte Carlo simulation

According to the literature and observations from Stingray louse data, there is evidence that the distribution of number of lice per fish in a pen is positively skewed. The consensus is that the negative binomial distribution describes the distribution of number of sea lice between fish well. The Norwegian veterinary institute, called Veterinærinstituttet in Norwegian, is a biomedical institute that delivers research based knowledge and contingency support in fields of fish health amongst other. The veterinary institute produced a series of reports in 2016, 2017, and 2018 (van Son et al., 2016; Viljugrein and Helgesen, 2017; Helgesen and Kristoffersen, 2018) that suggested solutions to deal with uncertainty in lice counts as they are being conducted in Norway today. In these reports the authors estimated the parameters of the negative binomial distribution, by fitting a negative binomial model to count data that they gathered from manual counts of individual fish from counts with varying louse pressure. The negative binomial distribution was the model that best described the number of lice between fish out of a set of candidate models. The candidate models were:

- Poisson distribution

- Zero-inflated Poisson distribution
- Negative binomial distribution
- Zero-inflated negative binomial distribution.

Knowledge about the distribution of lice between fish can be used to better understand the uncertainty that is presented when counting lice on very few fish from a large population. Monte Carlo simulation, similar to the simulation used in the veterinary institute reports, will be used to illustrate the potential uncertainty in the counting results from lice counts in Norwegian fish farms today. We have not estimated the parameters of the negative binomial model for this assignment, and will use veterinary institute estimates for the simulations. Currently in Norway, the average of adult female lice counted on ten to twenty fish are used as an estimate for the population mean in a pen. In this section, the simulation method of how to construct theoretical confidence intervals around different population means, when the sample sizes are twenty fish (pen level) and 120 fish (location level), will be described. We assume that the theoretical underlying distribution correctly describes the distribution of lice between fish. For the location level, we assume that all pens have equal louse pressure. The number 120 was chosen because the average of operational pens in a fish farm location in Norway in 2021 was six pens. We calculated this average from the operational locations and pens overview per county between 2005 and 2023, within the biomass statistics page on *fiskeridir.no*. Six pens multiplied by twenty, which is the required sample size per pen, is equal to 120.

3.4.1 Negative binomial distribution

The negative binomial distribution is often used to model count data where the variance and mean of the data are significantly different from each other. The parameterisation of the negative binomial distribution used for the simulation study is given by:

$$NegBin(y|\mu,\theta) = \binom{y+\theta-1}{y}\left(\frac{\mu}{\mu+\theta}\right)^{y}\left(\frac{\theta}{\mu+\theta}\right)^{\theta}, \tag{10}$$

where E[Y] = μ and Var[Y] = $\mu + \frac{\mu^2}{\theta}$. The parameter theta ($\theta$) is often referred to as the dispersion parameter. If the dispersion parameter is very large and the concentration μ is small, the

second term of the variance will get closer to zero and the distribution will get closer to the Poisson distribution. On the other hand, the variance will become greater with decreasing θ and fixed μ, hence the dispersion of the data depends on θ. In this parameterisation, the success probability p is expressed as

$$p = \frac{\theta}{\mu + \theta}. \tag{11}$$

The veterinary institute used the parameterisation in Equation (10) and estimated the dispersion parameter θ using this. In the programming language R, the `rnbinom(n, size, prob, mu)` function can be used to produce n random values drawn from the negative binomial distribution and has Equation (10) as a default if you only specify size as the dispersion parameter, set `mu` as the concentration and leave `prob` blank within the function.

3.4.2 Constructing confidence intervals

The following section will describe in detail how confidence intervals around specified population means have been calculated. The intervals have been calculated through the steps:

1. create 10,000 samples of size twenty or 120 by drawing twenty or 120 random values from Equation (10),
2. calculate the means of the 10,000 samples,
3. set the lower limit as the 2.5th percentile and the upper limit of the confidence interval as the 97.5th percentile of the calculated means.

The dispersion parameter of Equation (10) is set according to what was found in the veterinary institute reports. Confidence intervals have been calculated around population means ranging from zero to 1.5, with 0.01 increments.

4 Results

4.1 Plots and figures for adult female lice counts

Image-based counts

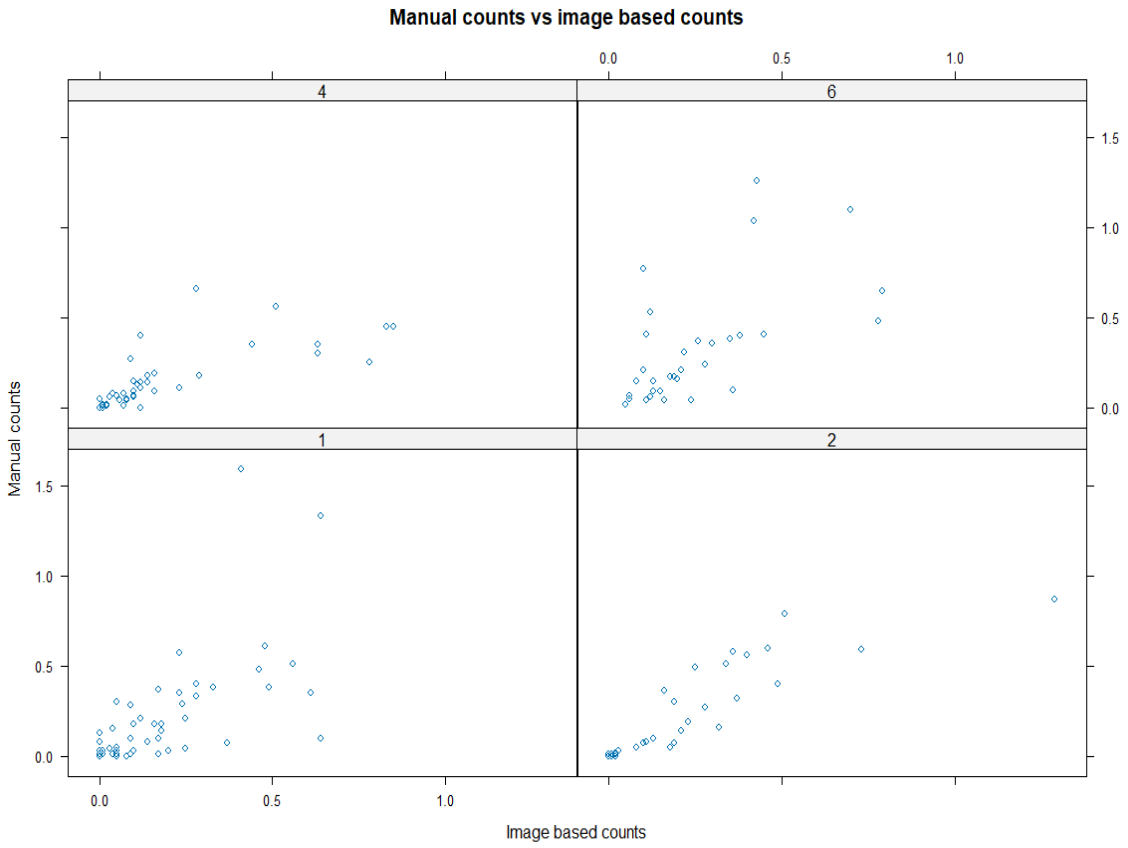


Figure 10: *Relationship between image-based counts on the x-axis and manual counts on the y-axis at the four different locations denoted by a number.*

Figure 10 shows the relationship between manual counts and image-based counts of adult female lice per fish at four locations. All plots show a positive trend with different levels of dispersion. The plot of Location 1 potentially shows two significant outliers, and there are a few potential outliers at Location 4. The plot of Location 6 seem to have a heteroskedastic trend. The plot of Location 2 has the least dispersion. Location 1 has two extreme observations of the manual counts compared to image-based counts, whereas Location 2 have a couple of observations with higher image-based count compared to the manual count.

Normalised laser pulses

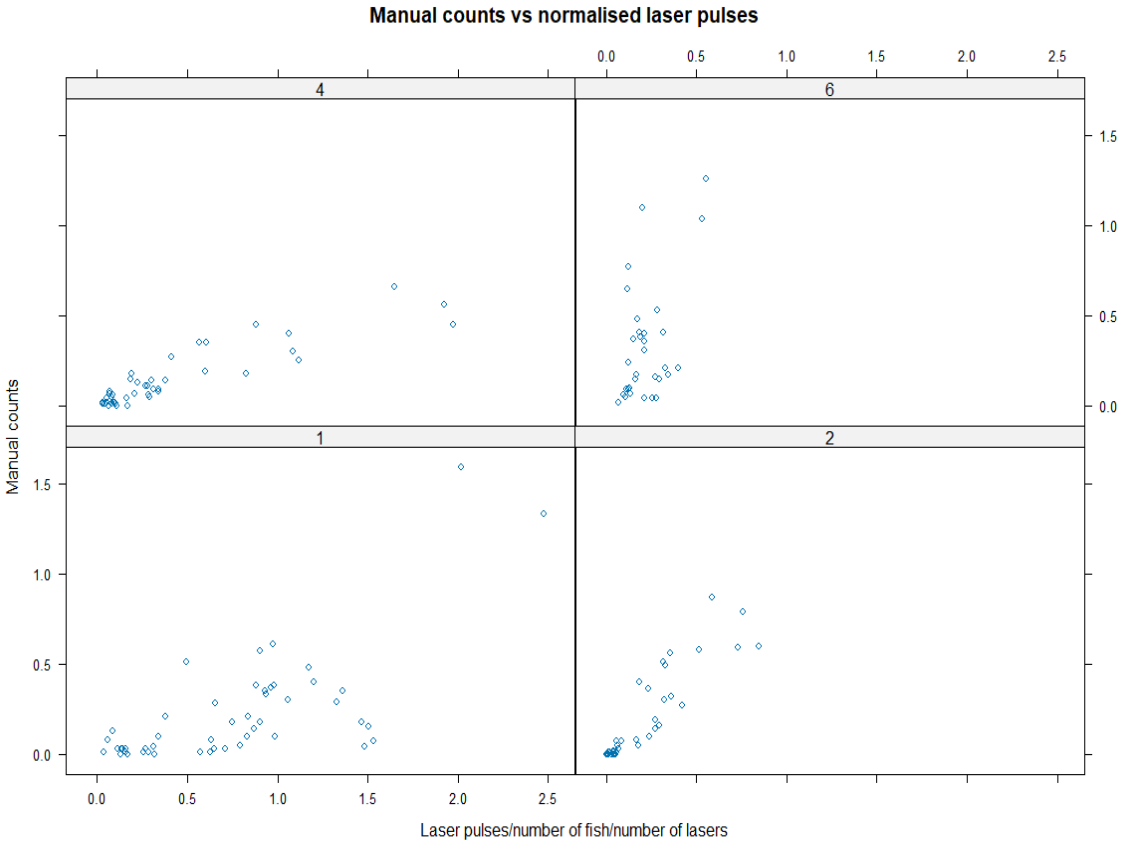


Figure 11: *Plots of different locations with normalised laser pulses (laser pulses / number of fish / number of lasers) on the x-axis and manual counts on the y-axis at the four different locations denoted by a number.*

Figure 11 shows the relationships between normalised laser pulses shot and manual counts. All plots have positive trends, indicating that there are higher manual counts when there are higher numbers of laser pulses shot per fish per laser unit. The plots of Location 2 and 4 potentially show positive linear trends. The plot of Location 1 may show a linear relationship between manual counts and normalised laser pulses. The plots of Location 1 and 6 appear more dispersed than the other two locations, and had the highest observed normalised laser pulses of the locations.

Average fish weight

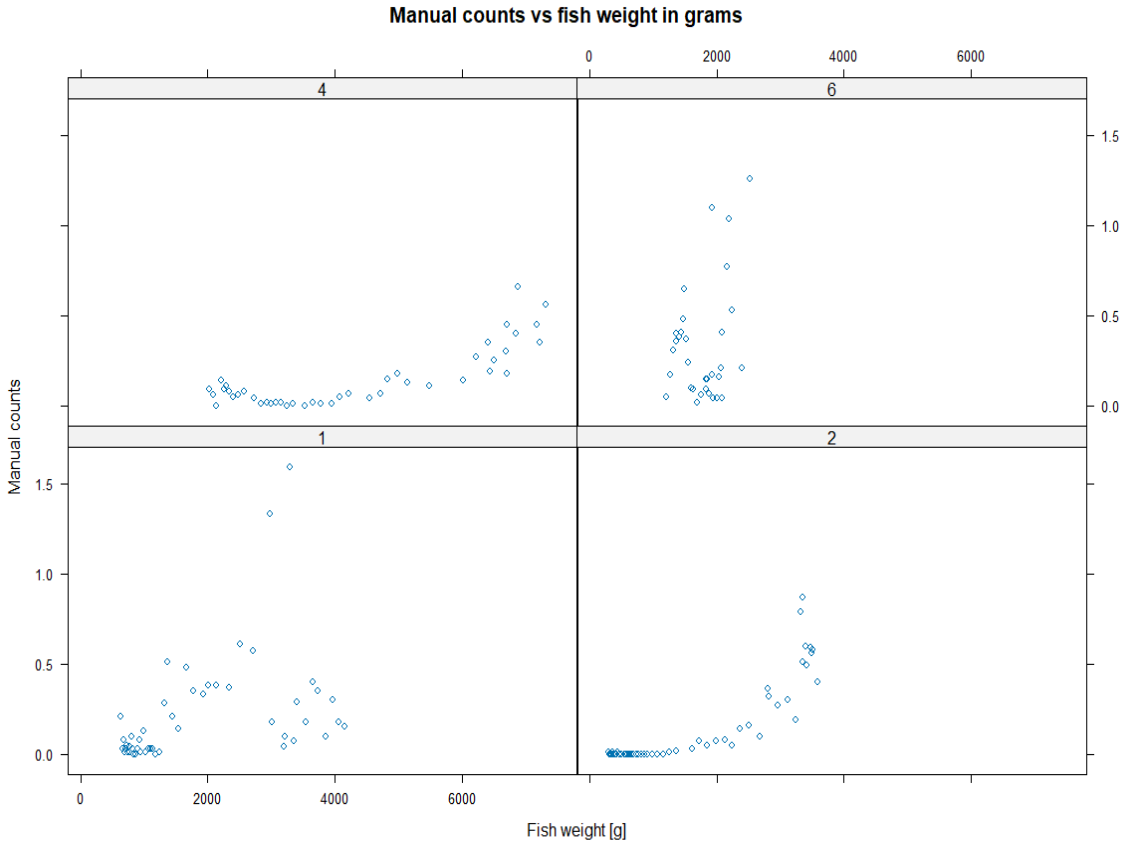


Figure 12: *Plots of different locations with fish weight on the x-axis and manual counts on the y-axis at the four different locations denoted by a number.*

Figure 12 shows that Location 2 and 4 have a positive non-linear trend. The plots at Location 1 and 6 are harder to interpret, but they both seem to have a positive trend. Location 4 has a decline in manual louse counts with increasing weight to begin with, before it starts to increase when average fish weight is around 3.5 kg. This may be due to missing data points for that particular location. Location 4 has the biggest and fastest growing fish throughout the observation period, despite covering a shorter observation period than Location 1 and 2. The plots of Location 1 and 2 show that the initial fish weight at the beginning of observation period was very small, well below one kg. The average weight of the fish of Location 6 does not change much throughout the observation period. Two significant outliers of Location 1 are standing out in this plot as well.

Number of fish

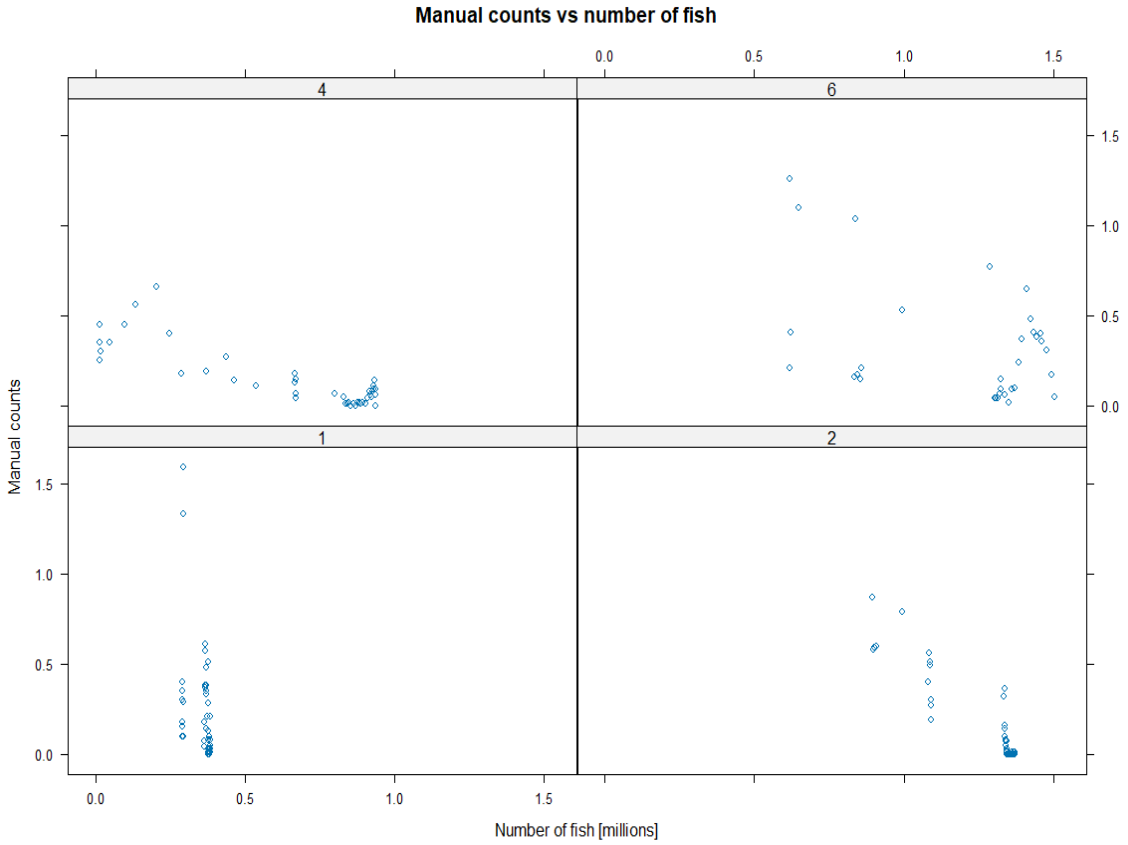


Figure 13: *Plots of different locations with number of fish in millions on the x-axis and manual counts on the y-axis at the four different locations denoted by a number.*

Figure 13 shows that the number of fish varies considerably between locations. All of the plots suggest a negative trend, indicating that there are less lice per fish with more fish. Location 6 has the greatest amount of dispersion. Location 4 appear to have increasing variance with increasing number of fish. Location 4 and 6 have the greatest reduction in fish populations, whereas the fish population in Location 1 barely changes at all. At the end of the collection period, Location 4 had a reduced fish population consisting of less than 100,000 fish. The number of fish is declining gradually in Location 2, with two bigger jumps when the fish population was about 1.3 million and just over 1 million fish.

Sea temperature

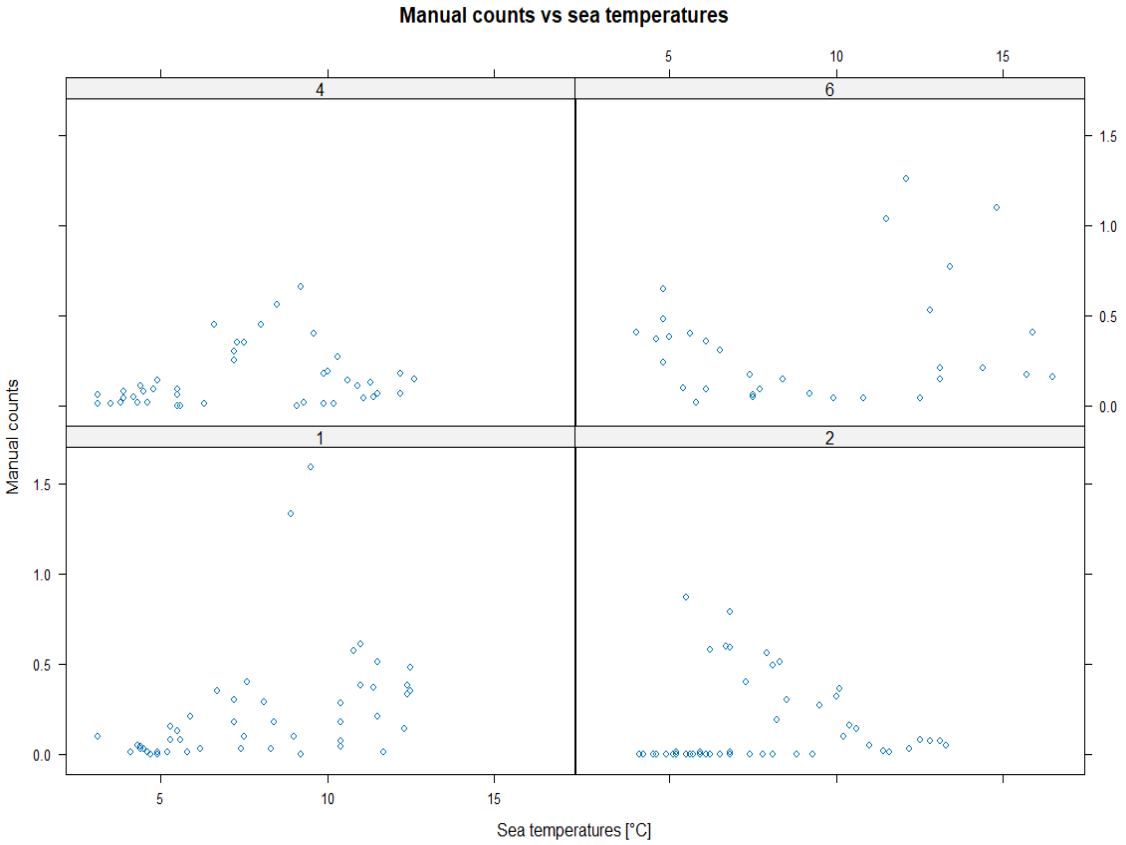


Figure 14: *Plots of different locations with sea temperature on the x-axis and manual counts on the y-axis at the four different locations denoted by a number.*

Figure 14 shows the relationships between sea temperature and manual counts. Location 1 and 4 have positive trends whereas the plots of Location 2 and 6 have more unusual patterns. Location 6 has the highest sea temperatures, with a maximum temperature of over 16 °C. Location 2 and 4 have the lowest sea temperatures, reaching a maximum temperature of around 12.5 °C.

4.2 Linear models

The figures of Section 4.1 show that there are tendencies of linear relationships between the manual counts and image-based counts, and between normalised laser pulses and number of fish. The plots of manual counts versus fish weight, and manual counts versus sea temperatures, did not show any clear sign of linear relationships. In this section, I continue to look further into these relationships

by fitting multiple linear models with manual counts as the responses accompanied by every variable looked at in Section 4.1, one by one. The reason for making such simple isolated models to begin with is to help determine the random structure of the mixed-effects models later on, if in need of a random structure at all, to account for between-group variation.

Image-based counts

Location	Coefficients	Estimate	Std.error	t.value	p.value	R^2
1	intercept	0.0291	0.0383	0.760	0.448	0.396
2	intercept	0.0153	0.0317	0.482	0.631	0.798
4	intercept	0.0533	0.0366	1.456	0.147	0.549
6	intercept	0.0768	0.0552	1.391	0.166	0.366
1	imageFem	0.974	0.144	6.751	< 0.0001	0.396
2	imageFem	0.888	0.114	7.821	< 0.0001	0.798
4	imageFem	0.515	0.124	4.146	< 0.0001	0.549
6	imageFem	0.982	0.171	5.756	< 0.0001	0.366

Table 7: *Summary statistics of the linmodImFem model, with manual counts as dependent variable and image-based counts as independent variable. Locations are denoted under Location by their respective numbers, coefficients are the names of the coefficients, whereas the values under estimate are the coefficient estimates for the intercept and image-based count variables. The std.error stands for standard error. The t-value is a statistic that measures the number of standard errors that the estimate is away from zero. The p-value can aid in determining if a coefficient is significant or not. The R^2-values for each model are added twice, but it is important to note that it is only one unique R^2-value per model, they were added twice to make the table look better.*

Table 7 shows the summary of the linear models where manual counts is a function of image-based counts. The estimates of the intercepts at each location are close to zero and each of them have insignificant p-values. The estimates of the slope coefficients vary between locations. For Location 1 and 6, the estimates are 0.974 and 0.982 respectively which indicates that image-based counts and manual counts are quite similar. Location 2 has a slope estimate relatively close to one. In contrast, the model of Location 4 has a slope estimate of 0.515 which is lower than the other

locations. All of the slope estimates are below one, which means that image-based counts tend to be slightly higher than manual counts. All p-values of the slope estimates are far under the significance level of 0.05. The R^2 coefficients are highest for Location 2 and 4, with the highest value being 0.798 for Location 2. This indicates that image-based counts explain a high degree of the variance at Location 2. On the other hand, the R^2 values of the models at Location 1 and 6 indicate that image-based counts explain less of the variation in manual counts. We plot the confidence intervals of the intercept and slope estimates:

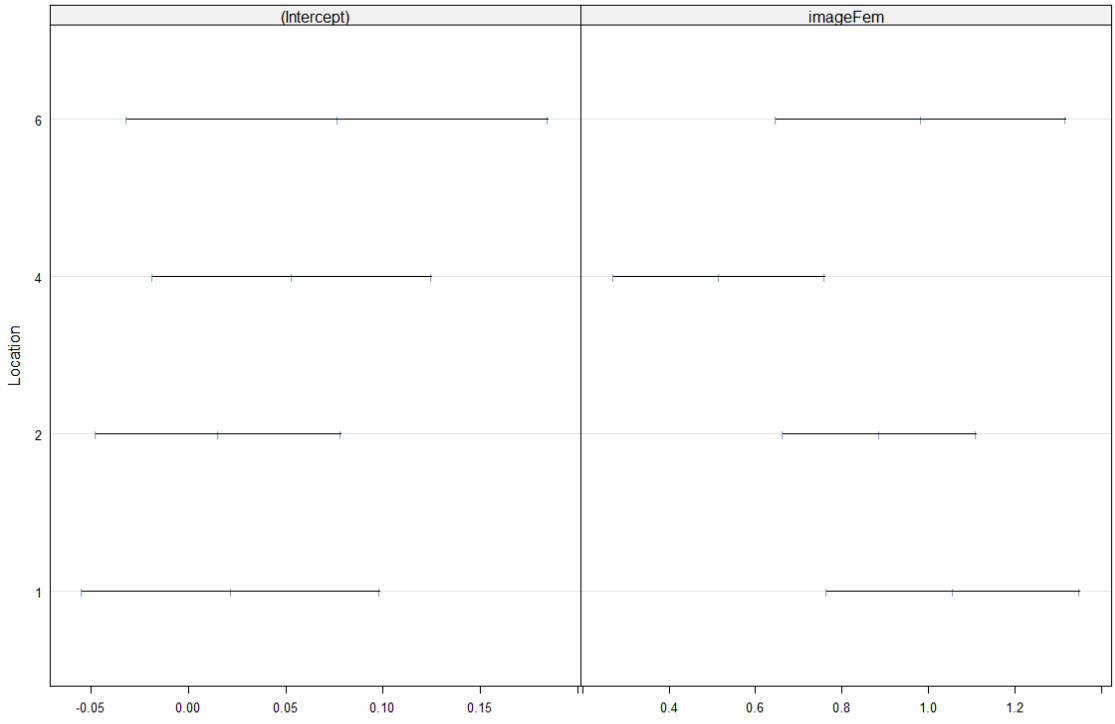


Figure 15: *Plots of 95% confidence intervals on the intercept (left) and slope estimates (right) at the locations denoted by a number on the left for the linear model with manual counts as dependent variable, and image-based counts as independent variable. The x-axis shows coefficient estimates for each coefficient.*

All of the intervals in both the intercepts and slopes are overlapping. An overlap indicates that random effects might not be necessary to adjust for varying intercepts and slopes. The slope of Location 4 has the most unique interval of all of the locations as we also saw in the summary of the linear model with image-based counts as independent variable, where Location 4 had a lower slope estimate.

Normalised laser pulses

Location	Coefficients	Estimate	Std.error	t.value	p.value	R^2
1	intercept	-0.0904	0.0450	-1.788	0.076	0.488
2	intercept	-0.0100	0.0324	-0.309	0.758	0.833
4	intercept	0.0231	0.0363	0.636	0.526	0.811
6	intercept	0.0306	0.0695	0.440	0.661	0.239
1	normPulses	0.401	0.0488	8.219	< 0.0001	0.488
2	normPulses	0.994	0.120	8.302	< 0.0001	0.833
4	normPulses	0.297	0.0565	5.254	< 0.0001	0.811
6	normPulses	1.353	0.279	4.848	< 0.0001	0.239

Table 8: *Summary statistics of the linmodNormPulses model, with manual counts as dependent variable and normalised laser pulses as independent variable. Locations are denoted under Location by their respective numbers, coefficients are the names of the coefficients, whereas the values under estimate are the coefficient estimates for the intercept and normalised laser pulses variables. The Std.error stands for standard error. The t-value is a statistic that measures the number of standard errors that the estimate is away from zero. The p-value can aid in determining if a coefficient is significant or not. The R^2-values for each model are added twice, but it is important to note that it is only one unique R^2-value per model, they were added twice to make the table look better.*

Table 8 shows that intercept estimates are close to zero for all locations. Location 1 and 2 have negative values whereas Location 4 and 6 have positive values. All corresponding p-values are large and above the significance level. The slope estimates of all the locations have positive values. Positive values indicate a positive relationship between normalised laser pulses (laser pulses / number of fish / number of lasers) and manual counts. The slope estimates vary between locations, where Location 6 has the highest slope estimate of 1.353 and Location 4 has the lowest slope estimate of 0.297. All slope estimates have significant p-values. The R^2 coefficients are highest for the model at Location 2 and 4. The highest R^2 value of 0.833 was observed at Location 2, and the lowest R^2 value of 0.239 was observed at Location 6.

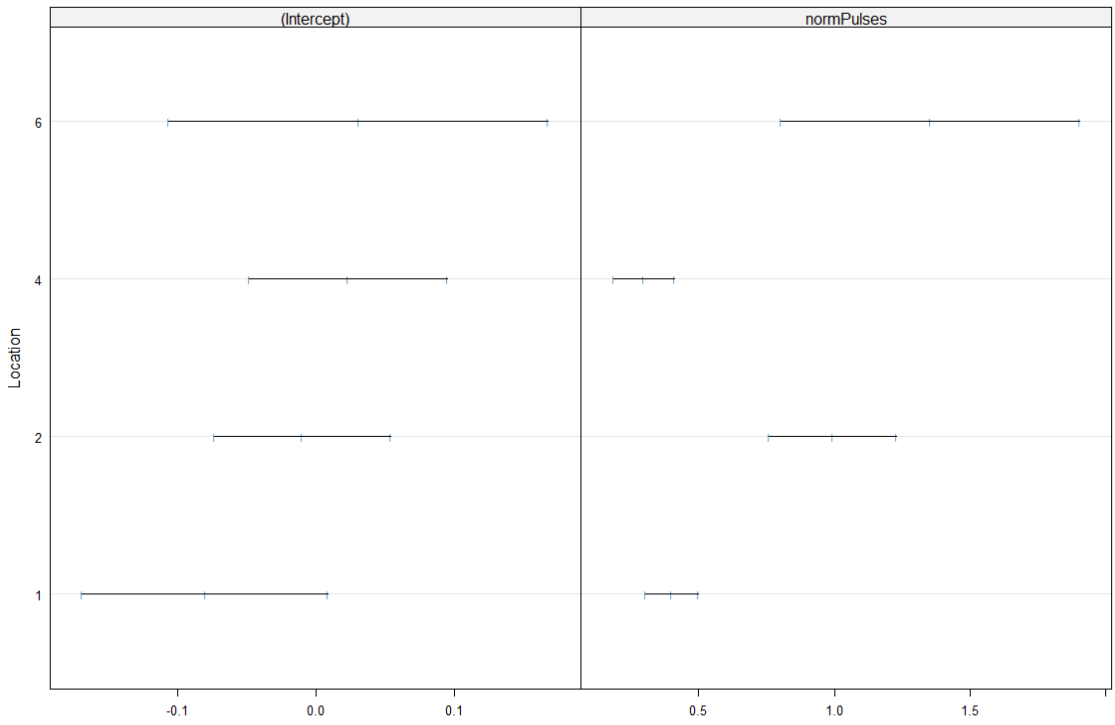


Figure 16: *Plots of 95% confidence intervals on the intercept (left) and slope estimates (right) at the locations denoted by a number on the left. The intervals are based on the linear model with manual counts as dependent variable and normalised laser pulses as independent variable. The x-axis shows coefficient estimates for each coefficient.*

Figure 16 shows that the confidence intervals for the intercepts are overlapping for all of the locations. The interval of the intercept at Location 6 is wider than the rest which is expected because of the imbalance in the data. Location 6 has the highest slope estimate, and the confidence interval of its estimate is only overlapping with the confidence interval of Location 2. A random effect to account for the varying slopes might be needed for the normalised pulses.

Fish weight

Location	Coefficients	Estimate	Std.error	t.value	p.value	R^2
1	intercept	0.0161	0.058	0.280	0.780	0.174
2	intercept	-0.130	0.0525	-2.474	0.0143	0.728
4	intercept	-0.170	0.0874	-1.940	0.0541	0.632
6	intercept	-0.195	0.204	-0.955	0.341	0.0983
1	weight	0.000107	2.615e-05	4.077	< 0.0001	0.174
2	weight	0.000170	2.596e-05	6.536	< 0.0001	0.728
4	weight	0.0000725	1.859e-05	3.899	0.000139	0.632
6	weight	0.000294	1.123e-04	2.615	0.00972	0.0983

Table 9: *Summary statistics of the linmodWeight model, with manual counts as dependent variable and fish weight as independent variable. Locations are denoted under Location by their respective numbers, coefficients are the names of the coefficients, whereas the values under estimate are the coefficient estimates for the intercept and fish weight variables. The std.error stands for standard error. The t-value is a statistic that measures the number of standard errors that the estimate is away from zero. The p-value can aid in determining if a coefficient is significant or not. The R^2-values for each model are added twice, but it is important to note that it is only one unique R^2-value per model, they were added twice to make the table look better.*

Table 9 shows that all the locations except from Location 1 have negative intercept estimates. The corresponding p-values of the intercept estimates of all locations are insignificant. The slope estimates are positive and very small values, indicating a positive relationship between average fish weight and manual counts. All corresponding p-values are below 0.05. The R^2 coefficients are again higher for Location 2 and 4, and lower for Location 1 and 6. Low R^2 values indicate that fish weight struggle to describe the variance in the manual counts. The highest R^2 value was observed at Location 2.

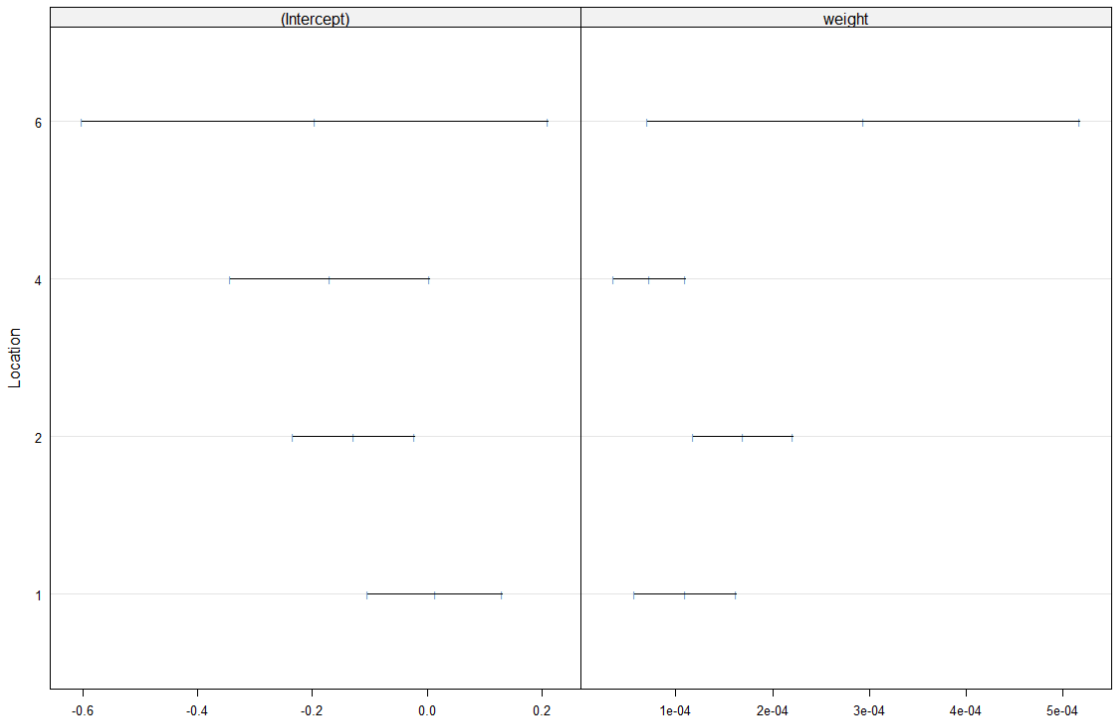


Figure 17: *Plots of 95% confidence intervals on the intercept (left) and slope (right) estimates at the locations denoted by a number on the left. The intervals are based on the linear model with manual counts as dependent variable and fish weight as independent variable. The x-axis shows coefficient estimates for each coefficient.*

Figure 17 shows an overlap between all intercept intervals and all slope intervals, indicating that there would be no need for a random effect. Location 1, 2 and 4 have a higher intercept compensated by lower slopes. Location 6 has wide intervals covering most of the widths of the plots both for the intercept and slope.

Number of fish

Location	Coefficients	Estimate	Std.error	t.value	p.value	R^2
1	intercept	1.585	0.295	5.381	< 0.0001	0.195
2	intercept	1.985	0.249	7.982	< 0.0001	0.839
4	intercept	0.415	0.0685	6.0573	< 0.0001	0.723
6	intercept	0.879	0.149	5.882	< 0.0001	0.185
1	numFish	-3.837	0.822	-4.667	< 0.0001	0.195
2	numFish	-1.441	0.194	-7.427	< 0.0001	0.839
4	numFish	-0.419	0.0950	-4.409	< 0.0001	0.723
6	numFish	-0.460	0.121	-3.794	0.000206	0.185

Table 10: *Summary statistics of the linmodNumFish model, with manual counts as dependent variable and number of fish as independent variable. Locations are denoted under Location by their respective numbers, coefficients are the names of the coefficients, whereas the values under estimate are the coefficient estimates for the intercept and number of fish variables. The std.error stands for standard error. The t-value is a statistic that measures the number of standard errors that the estimate is away from zero. The p-value can aid in determining if a coefficient is significant or not. The R^2-values for each model are added twice, but it is important to note that it is only one unique R^2-value per model, they were added twice to make the table look better.*

The intercepts for Location 1 and 2 are high compared to Location 4 and 6. All corresponding p-values of the intercept estimates indicate a significant fit. The slope estimates are negative and greater for Location 1 and 2 as well, which means that the two locations have a higher starting point and steeper decline. Location 4 and 6 have negative values, but their slopes are not as steep, indicating a more gradual decline. The p-values corresponding to the slope estimates are significant. The negative values indicate more lice with fewer fish, as we saw in Figure 13. The R^2 coefficients are higher for Location 2 and 4, with the highest value being equal to 0.839 at Location 2. At this location, the number of fish variable explains a lot of the variance in manual counts. The lowest R^2 value of 0.185 was observed at Location 6, indicating that number of fish to a small degree manages to describe the variance of the manual counts at this location.

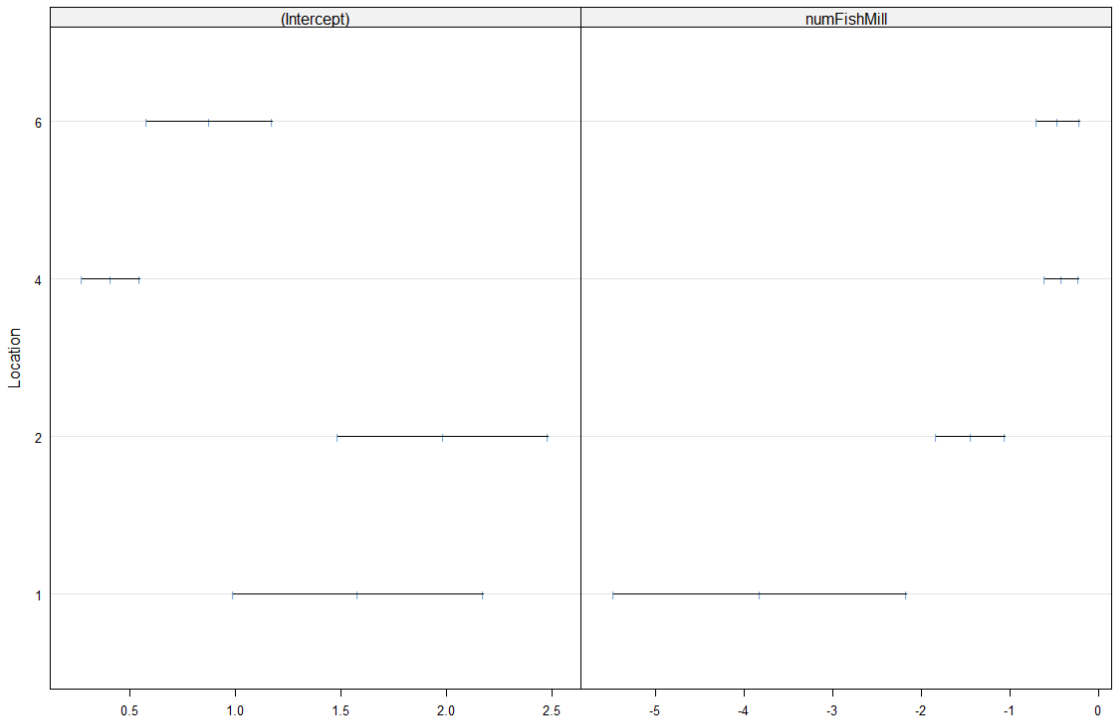


Figure 18: *Plots of 95% confidence intervals on the intercept (left) and slope estimates (right) at the locations denoted by a number on the left. The intervals are based on the model with manual count as dependent variable and number of fish as independent variable. The x-axis shows coefficient estimates for each coefficient.*

Figure 18 shows that the intercepts and slopes have some overlap. The intercept estimate for Location 4 has the lowest value, and the confidence interval overlaps slightly with the interval of Location 6. The intercept estimate of Location 6 is low compared to Location 1 and 2, and the confidence interval overlaps with Location 1. The intercept of Location 2 only overlaps with the interval of Location 1. A random effect might be needed to compensate for the varying intercept. Location 1 has the lowest slope estimate followed by Location 2. Neither of the two estimates overlap with each other or with the other locations. Location 4 and 6 have higher slope estimates and their confidence intervals overlap with each other. Random effects might be needed to compensate for the varying slope estimates.

Sea temperature

Location	Coefficients	Estimate	Std.error	t.value	p.value	R^2
1	intercept	-0.126	0.105	-1.199	0.232	0.174
2	intercept	0.146	0.112	1.307	0.193	0.000305
4	intercept	0.050	0.106	0.473	0.637	0.0530
6	intercept	0.183	0.115	1.587	0.114	0.0370
1	seaTemp	0.0427	0.0123	3.463	0.000676	0.174
2	seaTemp	0.000518	0.0136	0.0382	0.970	0.000305
4	seaTemp	0.0127	0.0131	0.968	0.334	0.0530
6	seaTemp	0.0157	0.0114	1.376	0.171	0.0370

Table 11: *Summary statistics of the linmodSeaTemp model, with manual counts as dependent variable and sea temperature as independent variable. Locations are denoted under Location by their respective numbers, coefficients are the names of the coefficients, whereas the values under estimate are the coefficient estimates for the intercept and sea temperature variables. The std.error stands for standard error. The t-value is a statistic that measures the number of standard errors that the estimate is away from zero. The p-value can aid in determining if a coefficient is significant or not. The R^2-values for each model are added twice, but it is important to note that it is only one unique R^2-value per model, they were added twice to make the table look better.*

Table 11 shows that Location 1 has a negative intercept whereas Location 2, 4 and 6 have positive intercepts. Location 2 and 6 have the highest estimates of intercepts. The intercept of Location 4 is close to 0. None of the corresponding p-values are significant. The slope estimates are quite low, with the lowest being that of Location 2. All slope estimates are positive which indicate more lice with warmer water. The p-values of the slope estimates are insignificant at Location 1 and 4, and significant at Location 2 and 6. The R^2 coefficients are very low for all of the locations, with the highest value observed being 0.174 at Location 1. The low values indicate that a linear relationship between sea temperatures and manual counts are not appropriate. Confidence intervals will not be plotted as there are strong indications from Table 11 that sea temperature and manual counts do not have a linear relationship. The relationships between manual counts and sea temperatures as shown in Figure 14 generally displays different behaviour between locations.

The purpose of conducting the preliminary analysis described in the last two subsections has been to get an impression of the relationships between manual counts and different variables of interest. More information is needed to conclude whether or not linear models are sufficient to describe these relationships, but we will not draw further inferences at this point.

4.3 Linear mixed-effects models

The previous steps of the analysis have provided an idea of what the final model will look like. In this section we will fit linear mixed-effects models. An interaction term has been included between treatment and image-based counts in the initial model. The previous steps of the analysis have shown indications that random effects may be needed. We start off the second part of the analysis with a backward selection process, where multiple models with and without random effects are fitted, and compare them with AIC and BIC values.

Model	Covariates (fixed)	Random	AIC	BIC
linModFull	intercept + IF + NP + ST + W + NFM + IF:treat	none	-97.925	-72.653
linMod2	intercept + IF + NP + ST + NFM + IF:treat	none	-98.210	-76.093
linMod3	intercept + IF + NP + ST + NFM	none	-96.645	-77.554
mlmeFull	intercept + IF + NP + ST + W + NFM + IF:treat	intercept	-95.925	-67.494
mlme2	intercept + IF + NP + ST + NFM + IF:treat	intercept	-96.444	-71.172
mlme3	intercept + IF + NP + ST + IF:treat	intercept	-97.500	-75.387
mlme4	intercept + IF + NP + ST	intercept	-95.743	-76.789
mlme5	intercept + IF + NP	intercept	-94.875	-78.966

Table 12: *Table showing the backward model selection process of simple linear and linear mixed-effects models. All models have manual counts as dependent variable. The variable names have been shortened to fit in the table. IF stands for image-based counts for female lice, NP stands for normalised pulses, ST is sea temperature, W is weight, NFM is number of fish in millions, and IF:treat is the interaction between louse treatment and image-based counts. Names in the Model column are model names, Covariates (fixed) are the fixed effects in each model. Random is the random effects where that applies, and AIC and BIC are the calculated AIC and BIC values from each model. The model with lowest BIC is coloured in gray.*

Table 12 shows the results of a backward selection process for both simple linear models and linear mixed-effects models. The model with an intercept, image-based counts and normalised pulses as fixed effects, as well as a random intercept, had the lowest BIC of all the candidates, with a small margin. It was made an attempt on fitting random slope models as well, but this was not possible due to convergence issues.

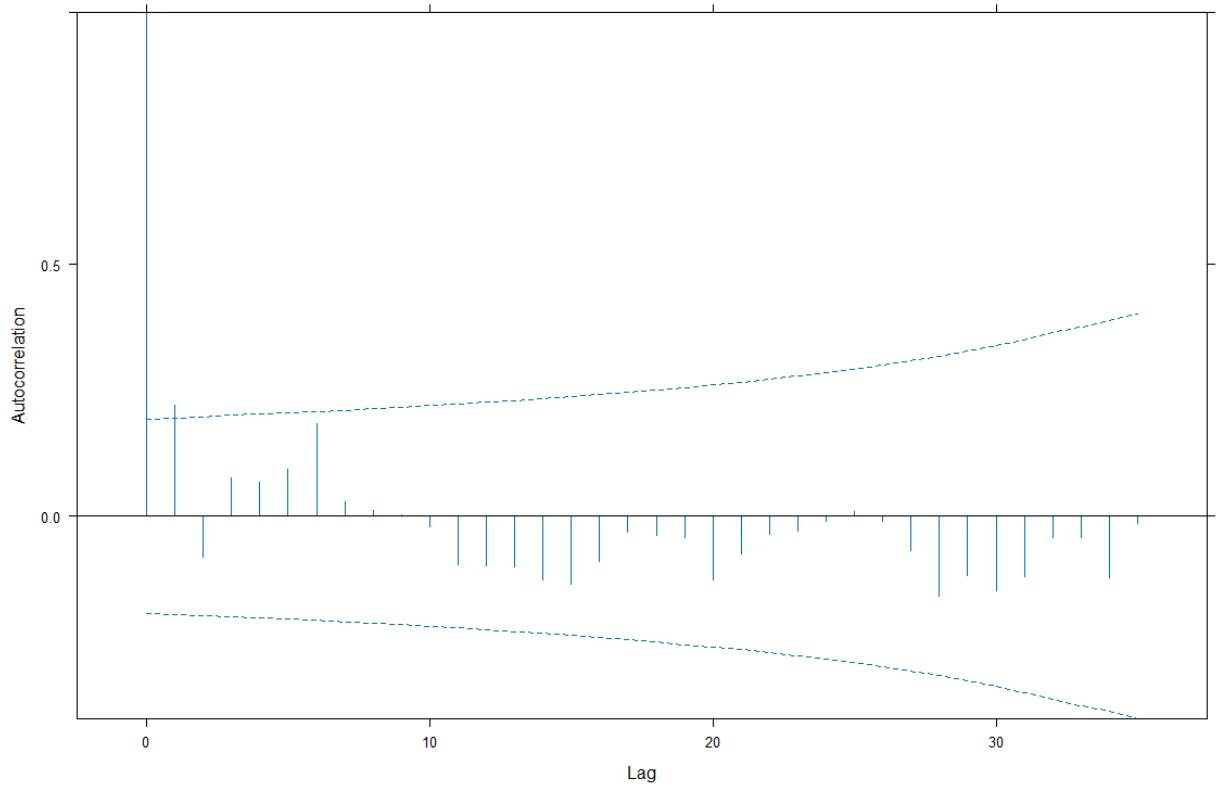


Figure 19: *Empirical autocorrelation function (ACF) corresponding to the standardised residuals of the mlme5 model for the first thirty-five lags. The dotted lines are significance levels.*

Figure 19 shows a significant spike at lag 1, indicating a potential autocorrelation issue. The spikes are generally lower for higher lags. The plot indicates that an alternative correlation structure might be needed in the model.

	call	df	AIC	BIC	logLik	Test	L.Ratio	p-value
mlme5	1	5.00	-94.87	-78.97	53.44			
mlme5Cor	2	6.00	-104.98	-85.89	58.49	1 vs 2	12.10	5e-04

Table 13: *Anova test on the mlme5 model in Table 12 versus the same model, with an AR(1) correlation structure added. The numbers under call are the order of the models, df is degrees of freedom, AIC and BIC are values calculated for each model, logLik is the log likelihood. Test shows what models are being compared, L.Ratio stands for likelihood ratio, and the p-value is the significance level of the likelihood ratio test.*

Table 13 shows the result of a likelihood ratio test performed on the winning model of Table 12 versus the alternative model with an AR(1) correlation structure. The result shows a very low p-value in the test as well as lower AIC and BIC values for the more complex model. This indicates strong evidence that the more complex model with the AR(1) correlation structure is a better fit.

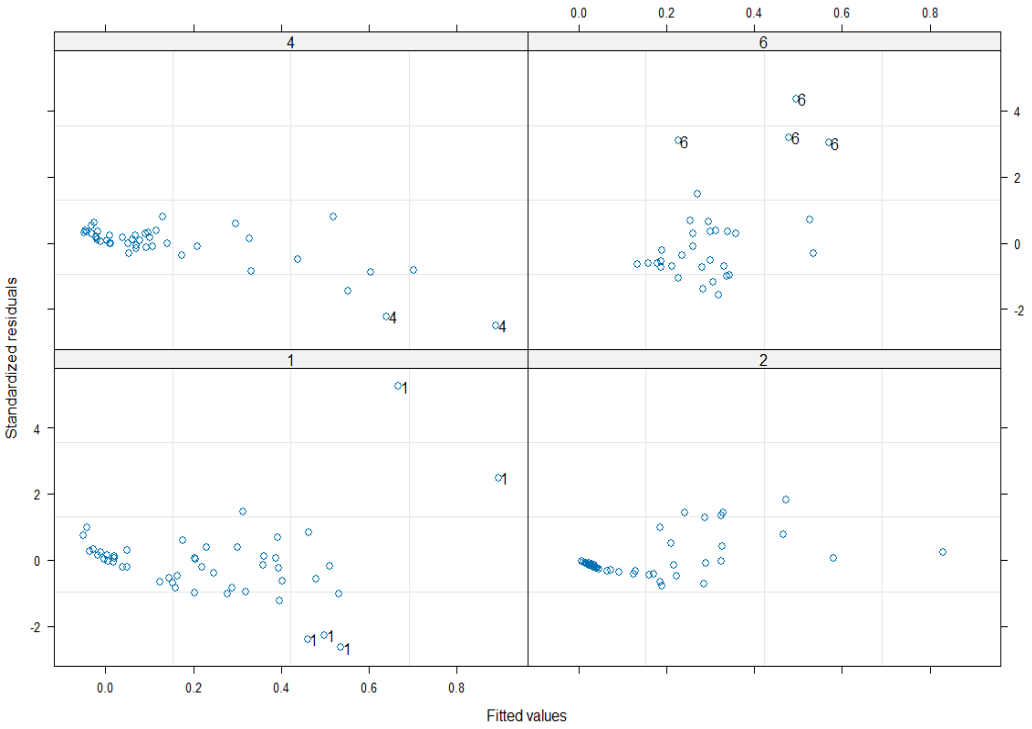


Figure 20: *Location-wise standardised residuals of the mlme5Cor model. The y-axis shows standardised residuals, and the x-axis shows fitted values of the model with manual counts as dependent variable, image-based counts and normalised laser pulses as independent variables, and an AR(1) correlation structure.*

Figure 20 shows that the standardised residuals of all the locations were mostly centered around zero, with different levels of dispersion. In addition, the residuals of the different locations showed signs of heteroskedasticity.

Model	Covariates (fixed)	Variance	AIC	BIC
mlme5Cor	intercept + IF + NP	none	-104.980	-85.889
mlme5V1	intercept + IF + NP	varIdent	-165.320	-136.684
mlme5V2	intercept + IF + NP	varPower (normPulses)	-362.924	-334.493

Table 14: *Table showing a model selection process. All models have manual counts as dependent variable, as well as an AR(1) correlation structure and random intercept. The variable names have been shortened to fit in the table. IF stands for image based counts for female lice, NP stands for normalised pulses and ST is sea temperature. Names in the Model column are model names, covariates (fixed) are the fixed effects in each model. Random is the random effects where that applies, Variance is the variance-covariance structure, and AIC and BIC are the calculated AIC and BIC values from each model. varIdent and varPower are variance functions as described in the Methods section. The model with lowest BIC is coloured in gray.*

In the final model selection part, different variance-covariance structures were added to the winning model in Table 12, in addition to the AR(1) correlation structure. The final model of this model selection process was the mixed-effects model, with an intercept, image-based counts and normalised laser pulses as fixed effects, a random intercept, a power function to model the variance of the residuals and an AR(1) structure to model the correlation in the residuals. The coefficients of the final model were estimated again with the restricted maximum likelihood estimation method to obtain the final coefficient estimates.

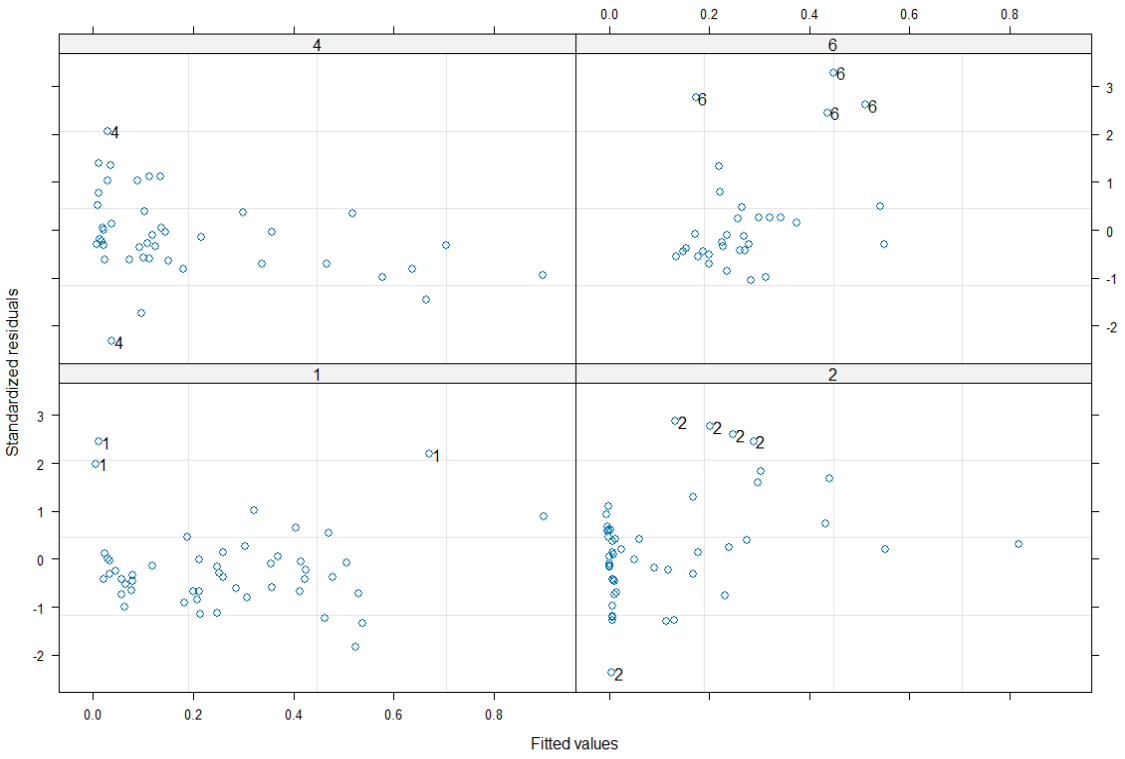


Figure 21: *Location-wise residuals of the mlme5V2 model. The y-axis shows standardised residuals, and the x-axis shows fitted values of the model. Each location is denoted by a number within each plot.*

Figure 21 shows that the residuals seem to mostly be centered around zero, and display no apparent trend. The location-wise residuals of the final model have several outliers. To test model robustness, all residuals more than two standard deviations away from the residual mean were removed (in total twelve), and the coefficients were estimated again without the outliers. The main objective with this thesis is to explore the relationship between manual counts and image-based counts. For the next part, a simple linear model with only image-based counts as a predictor, and a linear model with image-based counts and an interaction term between image-based counts and delousing treatments as predictors, fitted with the original data set and the data set without outliers, were added.

Model robustness

Model	Covariates (fixed)				Random		Data set
	Intercept	IF	IF:treat	NP	Intercept	Residuals	
linModIFA	0.0413	0.853	NA	NA	NA	NA	Original
linModIFB	0.0354	0.787	NA	NA	NA	NA	Altered
linModIFTA	0.0416	0.877	-0.238	NA	NA	NA	Original
linModIFTB	0.03518	0.774	0.155	NA	NA	NA	Altered
mlme5V2AltA	-0.00684	0.548	NA	0.228	2.451e-08	0.263	Original
mlme5V2AltB	-0.00797	0.661	NA	0.166	2.860e-07	0.209	Altered

Table 15: *Estimates of the coefficients of the different models. Model is model name. Covariates (fixed) are the fixed effects in each model, where Intercept is the intercept, IF is image-based counts, IF:treat is an interaction between image-based counts and treatment, and NP is normalised laser pulses. Random is random effects, where the values under Residuals and Intercept are the estimates for their standard deviations. The values under Dataset states whether the models were fitted with the original data set or the altered data set with outliers removed. Models fitted with the altered data are coloured in gray.*

Table 15 shows that the image-based coefficient estimates were not very robust. The estimates changed significantly for image-based counts, the interaction between image-based counts and treatment, and the normalised pulses. The interaction term changed sign in the model fitted with the altered data, whereas the other coefficient estimates kept the same direction. For the random effects, the standard deviations for the random intercepts were small for both models containing random effects, compared to the residual estimates. The big difference indicates that the random intercept did not explain the variation in the data well. The estimates of the random intercepts were all very close to zero and varied little between locations.

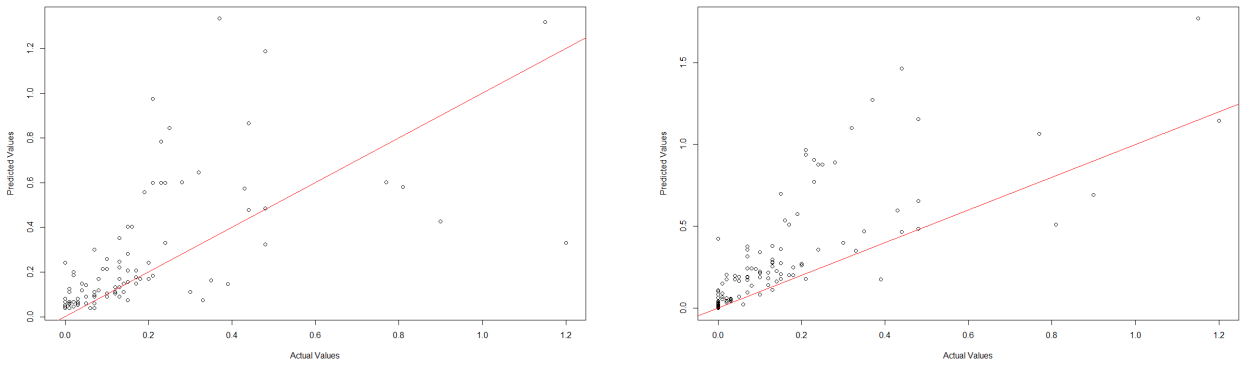
Predictions on unseen data

As a final evaluation of the models in Table 15, all the models were used to predict manual counts on a test data set. The mean square errors (MSE) were calculated for each model.

Model	MSE
linModIFA	0.0675
linModIFB	0.0491
linModIFTA	0.0665
linModIFTB	0.0488
linMod5V2AltA	0.0845
linMod5V2AltB	0.0746

Table 16: *Results of the calculated MSE values of predictions versus actual values of a test data set, for each model of Table 15. The bottom two models were fitted with the fixed effects only, because the random intercept estimates are very close to zero. Models fitted with the altered data set are coloured in gray.*

Table 16 shows that the simple linear model, with manual counts as dependent variable, and image-based counts and an interaction term between treatments and image-based counts, fitted on the altered data set, had the lowest MSE values of the candidate models. The more complex winning model of the model selection process, with an AR(1) structure to model correlation and a power structure of the normalised laser pulses to model the heteroskedasticity and different variance in the residuals between Locations, had the highest MSE values of all the models.



(a) *Model with the lowest MSE values (linModIFTB).* **(b)** *Model with the highest MSE values (linMod5V2AltA).*

Figure 22: *Figure showing two panels (a) and (b), where both plots have predicted values on the y-axis, and actual values on the x-axis. The red line shows where the points should lie for a perfect fit.*

Figure 22 shows that both the model with lowest MSE, and the complex model with higher MSE have a poor fit. They tended to overestimate the manual counts on the unseen data.

4.4 Simulation study

In this section, theoretical confidence intervals are being constructed around means of populations thought to follow the negative binomial distribution. The confidence intervals will be constructed on pen level (twenty fish) and location level (120 fish). For details, see the Methods section. The resulting confidence intervals will be presented in Table 17 and Table 18. The objective of this section is to show why, in theory, we are dealing with a great amount of uncertainty in our manual count and image-based count data. Uncertainty, or a great amount of variation, has been evident in the model selection process of the previous sections.

Confidence intervals, sample size (20)				
Population mean	Lower limit	Upper limit	Margin of error	Dispersion parameter
0.10	0	0.30	-100%, +200%	0.4
0.30	0.05	0.65	-83%, +120%	0.5
0.50	0.20	0.90	-60%, +80%	1.1
0.70	0.30	1.20	-57%, + 71%	1.1
0.90	0.40	1.50	-56%, +67%	1.1
1.10	0.50	1.80	-55%, +64%	1.1
1.30	0.65	2.10	-50%, +62%	1.1

Table 17: *Table showing a selection of 95% confidence intervals calculated around the means listed under population mean. Population mean is the mean of the theoretical negative binomial distributed population. Lower and upper limit are the resulting limits of the confidence intervals. Margin of error is the calculated size of the margin of error of each interval (rounded to the closest whole number). The dispersion parameter column shows the dispersion parameters that have been used in the simulations ran to calculate the confidence intervals. The confidence intervals were calculated based on twenty samples (pen level), with replacement.*

Table 17 shows seven different confidence intervals around different population means. The confidence level is set to be 95% for all of them. Each interval represents a large margin of error. The margins of error get relatively narrower as the population mean increases. It is evident from

the table that the confidence intervals around multiple different population means are overlapping. An observed value of 0.53 adult female lice can for example fall within the upper and lower bounds of 95% confidence intervals around population means ranging from 0.3 to as much as 1.1 in Table 17. Such big intervals indicate a lot of uncertainty in louse counting data, when the average of lice is used as a measurement to determine how much lice there are in a pen, and the measurement is extracted from a very small sample.

Confidence intervals, sample size (120)				
Population mean	Lower limit	Upper limit	Margin of error	Dispersion parameter
0.10	0.0417	0.167	-58%, +67%	0.4
0.30	0.183	0.433	-39%, +44%	0.5
0.50	0.350	0.667	-30%, +33%	1.1
0.70	0.517	0.900	-26%, +29%	1.1
0.90	0.683	1.133	-24%, +26%	1.1
1.10	0.850	1.375	-23%, +25%	1.1
1.30	1.0167	1.608	-22%, +24%	1.1

Table 18: *Table showing a selection of 95% confidence intervals calculated around the means listed under population mean. Population mean is the mean of the theoretical negative binomial distributed population. Lower and upper limit are the resulting limits of the confidence intervals. Margin of error is the calculated size of the margin of error of each interval (rounded to the closest number). The dispersion parameter column shows the dispersion parameters that have been used in the simulations ran to calculate the confidence intervals. The confidence intervals were calculated based on 120 samples (location level), with replacement.*

Table 18 shows the confidence intervals calculated, based on 120 samples. The margins of error are smaller in this table than in Table 17. All of the confidence intervals are overlapping, with the exception of the the interval around 0.1, this interval does not overlap with the other confidence intervals listed.

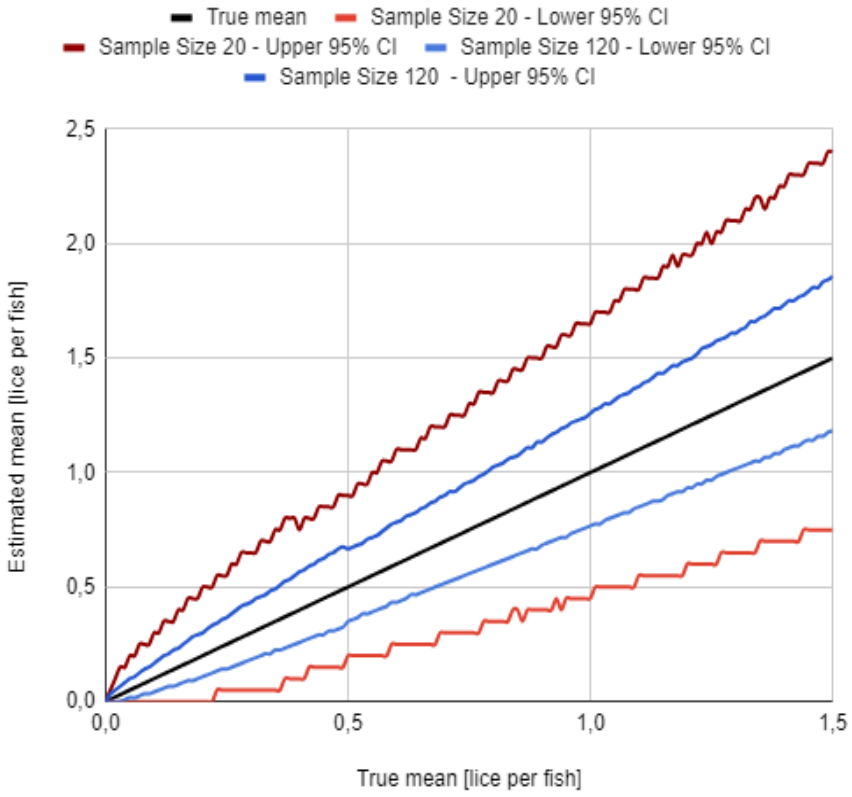


Figure 23: *95% confidence intervals around different population means when the sample sizes are twenty and 120 fish, with replacement. The upper darker red line is the upper bounds of the intervals, whereas the lower lighter line is the lower bounds of the intervals based on twenty samples. Equivalently the lower blue line is the lower bounds and the upper blue line is the upper bounds of the intervals based on 120 samples. The black line in the middle is the population means (True mean in the figure). The x-axis shows the population means and the y-axis shows the estimated mean lice per fish. The intervals have been calculated for theoretical population means ranging from zero to 1.5, with 0.01 step increments.*

Figure 23 shows a visual representation of the size of the confidence intervals around different population means. The black line is where the estimated means should lie if they perfectly align with the true population means. The figure shows that the confidence intervals are not symmetric around the population means, but have a wider margin of error above the population mean. The asymmetric confidence intervals suggest that it is not unlikely to observe a value that is relatively much greater than the population mean in question, whereas it is less likely to observe a value as extreme in relation to the population mean below the respective population mean.

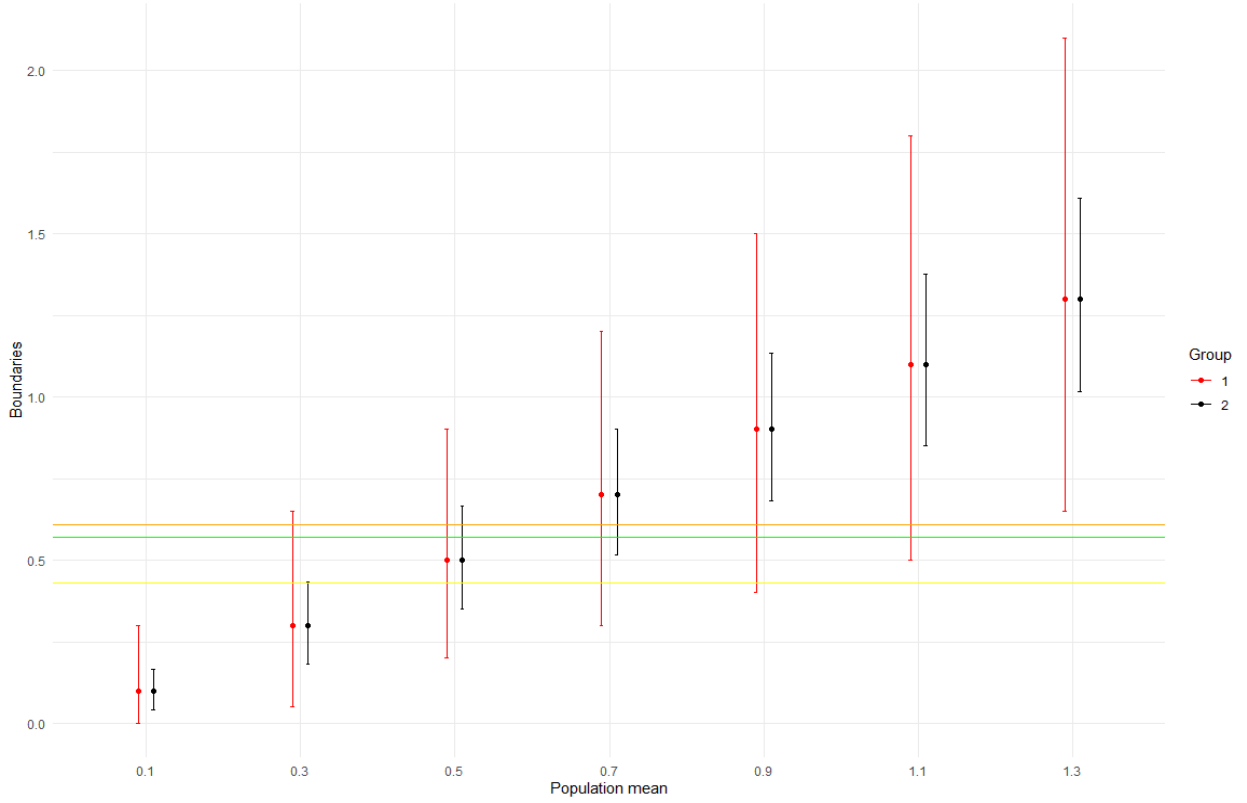


Figure 24: *95% confidence intervals around different population means when the sample size is twenty (Group 1 in red) and 120 (Group 2 in black), with replacement. The top of the vertical lines are the upper limits of each interval, the bottom of the lines are the lower limits. The red and black dots in the middle of each vertical line are the population means. The three horizontal lines in orange (0.61), green (0.57) and yellow (0.43) are examples of possible observed counting estimates to illustrate that the same values can fall within the bounds of multiple confidence intervals around theoretical population means.*

Figure 24 further highlights the overlapping confidence intervals (CIs). It is possible to see from this that an observed mean can fall within confidence intervals around multiple population means. For instance, the observed value of 0.57 (green line) can fall within a 95% CI when the population mean is 0.3, as well as a 95% CI when the population mean is 1.1, when the sample size is twenty fish. In other words, it is not unlikely to observe an average of 0.57 lice per fish when the population mean in reality is any of the values between 0.3 and 1.1 lice per fish. The value of 0.57 falls within the bounds of the confidence intervals around 0.5 and 0.7 when the sample size is 120 fish.

5 Discussion and conclusion

In this thesis we have used linear mixed-effects models to derive a correction factor between manual and image-based lice counts, with the influence of other variables, that compensates for missed lice on the side of the fish that faces away from the camera. The model selection process has mainly been based on comparison through AIC and BIC values, and backwards selection for variable selection. Additional correlation and variance structures of the residuals were added after inspecting the residual plots of the final model, after the variable selection process. This improved the fit of the final model in terms of AIC and BIC values. In cases where AIC and BIC values disagreed with each other, lower BIC was preferred. Throughout the model selection process, the estimation method for the fixed effects in the mixed-effects models were maximum likelihood estimation. When the final model was selected, restricted maximum likelihood was used to obtain the coefficient estimates. The final model was implemented for counts of adult female lice. Manual counts have been treated as the variable we want to model. The results in Table 12 show that a random intercept model with a complex structure of the residuals, is better than a simple linear regression model due to lower AIC and BIC values. However, when tested on unseen data, the complex model did not perform well, and the simple linear model with image-based counts and an interaction term between image-based counts and treatments, performed better on unseen data in terms of MSE.

Model interpretation

The model produced in this thesis was designed to correct the image-based counts of adult female lice, describe the relationship between manual counts and image-based counts, and investigate to what extent such relationships are impacted by other variables, such as normalised laser pulses, fish weight, number of fish, fish passings, sea temperature and louse treatments. The final model was a random intercept model, with image-based counts and normalised pulses as independent variables. An AR(1) correlation structure as well as a power structure of the residuals were included in this model.

The coefficient estimates of the fixed effects were positive, except from the intercept which was negative and close to zero. For image-based counts, the coefficient estimate was equal to 0.548,

which is smaller than one. This indicates that image-based counts tended to be higher than manual counts, and therefore adjust manual counts down. The normalised pulses had a coefficient estimate equal to 0.228, which indicates a positive relationship between normalised pulses and manual counts. The coefficient estimate of image-based counts did not confirm prior expectations regarding their relationship with manual counts. The object of this thesis was to find a correction factor to adjust for the part of a fish we cannot see, meaning we are looking to adjust image-based counts upwards. The resulting model however, adjusts the manual counts down. Image-based counts are systematically higher than manual counts, which was not initially expected. The relationship between manual counts and normalised pulses makes sense, as we expect the lasers to shoot more when there are more lice present.

The standard deviation of the random intercept in the final model is small compared to the standard deviation of the residuals. In addition, the random intercept estimates for each location were very small and close to zero. This is an indication that the random intercept failed to capture the variance in the data, although the random intercept model beat the simpler linear models in terms of BIC values at the beginning of the model selection process.

When fitting the same model on an altered data set, where the worst outliers of the residuals of the first model were removed, the fixed coefficient estimates changed significantly. The coefficient estimates went up to 0.661 for image-based counts and down to 0.166 for normalised pulses. For the simple linear models in Table 15, the coefficient estimates for image-based counts got reduced when fitted on the altered data. The interaction term between image-based counts and treatments changed sign, which indicates that whenever a treatment happened, an additional increase in manual counts would be applied. Such a change make no logical sense, and is probably due to important drops after treatments in lice numbers being removed with the outliers. The original model fitted on the unchanged data had a coefficient estimate of the interaction equal to -0.238, which indicates that whenever a treatment happened, manual counts were adjusted down for that particular week, which is more in line with what would be expected. Removing twelve outliers from such a small data set is somewhat extreme, a change in estimates could be expected. However, all models fitted on the altered data had lower MSE of the predictions on the unseen data, compared to the identical

models fitted on the original data. This indicates that the outliers were influential on the original models, and removing them with care may still be a good idea.

Simulation study and uncertainty in the data

There are several statistical uncertainties, biological and operational effects and biases, which may account for these results. Both image-based counts and manual counts are estimates based on the legally required sample size of twenty fish per pen. However, the distribution of lice between fish is heavily skewed to the right (Baillie et al., 2009; Jeong and Revie, 2020) and is considered to follow a negative binomial distribution rather than a normal distribution, particularly at low levels of infestation. We illustrated in Table 17 and Table 18 in Section 4.4 how counting very few fish from a large population affects the confidence intervals around the theoretical population means from such skewed distributions, on pen level and location level (Helgesen and Kristoffersen, 2018). The confidence intervals calculated on location level assume equal louse pressure in all pens. This is not always the case. It has been shown that the louse pressure can vary between pens (Revie et al., 2005, 2007). The intervals as shown in the simulation study of this thesis are therefore potentially narrower than they should be on location level, if difference in louse pressure between pens was accounted for. However, Figure 23 shows that the margins of error are wide for counts on location level as well. The situation is even more extreme on pen level, when only twenty fish are being counted. All count estimates used in this assignment are based on the legally required sample sizes of twenty fish per pen, which cannot be considered to be numerically representative for a population of up to 200,000 fish per pen. In addition, another statistical uncertainty stems from the fact that lice counts are reported on location level, even though the statistically meaningful level of analysis would be on pen level (Revie et al., 2007). This makes it challenging to generate a well-fitting model from the available data.

However, small sample sizes and resulting statistical uncertainty alone may explain a lack of fit in the resulting model, but they do not explain why image-based counts generally were higher than manual counts for all locations. Biological, environmental and operational factors can affect and indeed support this finding. Simple human error, biases, or insufficient training in both sampling methods will have significant effects on the final results (Thorvaldsen et al., 2019). Biological fac-

tors, such as feeding dominance when using feed to catch fish leads to only dominant bigger fish to be counted. When not feeding the fish, you may only catch subordinate or even sick fish who tend to hang around near the surface. There are more salmon lice in the upper water columns (Frenzl et al., 2014; Oppedal et al., 2011). In addition, we know that salmon lice do not infect salmonid hosts randomly, and there are fewer fish with a lot of lice, and a lot of fish with little to no lice in a pen (Heuch et al., 2011). There are also indications that both fish size (Folkedal et al., 2012) and sea lice infestation level (Bui et al., 2016) have an influence on swimming depth of farmed salmon. Operational factors include placement of cameras during the sampling process, as well as how fish are physically collected for manual counts. For instance, some lice can be lost if the fish are crowded tightly together and they rub against each other physically, dislodging lice. The manual counts are restricted because of physical limitations. It is only a limited number of ways it is possible to catch fish with a handheld net, and it has been demonstrated by (Nilsson and Folkedal, 2019) that sampling Atlantic salmon from sea cages does not return unbiased results, at least not with respect to fish size. Although Stingray units register the positions of the cameras within a pen during the sampling process, this information was not available during the data collection period of this thesis and is thus not included. For future analysis, the position of the camera whilst collecting samples should be included. However, the Stingray units do not underlie the same physical limitations as the manual sampling process, as they are mobile and are continuously and actively placed at more varied depths and widths in the pen in order to optimize the amount of fish passing in front of the cameras. Because of the skewed distribution of lice between fish, getting an unbiased and representative sample for any counting method may be challenging, especially for small samples.

During manual lice counts, fish may get stressed, and they are being crowded for a period of time to get collected for counting. Lice can fall off during crowding as fish are rubbing together, during the crowding period as the fish get stressed, during the handheld netting process as the lice might fall off in the process, and during the sedating process. Salmon lice get sedated as well as the fish, and may end up in the tub that the fish are kept in during the counting process (Torvaldsen et al., 2018). The lice in the tub are also counted after the fish themselves have been counted, but some can potentially be lost when the liquid in the tub is poured through a sieve (Torvaldsen et al., 2018). Based on such procedural limitations, it may not be surprising that image-based counts tend to be

higher than manual counts. Image-based counts do not face the same limitations when it comes to escaping salmon lice, as the images are being taken in real time, while the fish are swimming by undisturbed.

The fact that image-based counts are performed by humans on images can contribute to higher lice estimates. Another parasite similar to the salmon louse is *Caligus elongatus* (von Nordmann 1832). *Caligus* is not of significance when it comes to reporting lice numbers in a fish farm, but they could mistakenly be reported as salmon lice when analysed on images. It could be challenging to, for example, distinguish between a mobile salmon louse and a *caligus* on an image. On the other hand, the analysis of this thesis has been conducted on counts of adult female salmon lice, which are much bigger than the *caligus* adult females, and it should be easier to distinguish these. *C. elongatus* are mostly found further north in the country, and are not a big problem for the fish farms in the western part of the country.

Method

Manual counts have been used as the dependent variable for modeling in this assignment. However, there are several reasons to assume that they do not necessarily represent the true lice situation in a pen at any given time. Using manual counts as the ground truth is therefore problematic (Jeong et al., 2023). In addition to the constraints on sample size and sampling procedure that apply for mandatory reporting of weekly lice numbers to the Norwegian government, there are procedural aspects that may influence or bias these data (Thorvaldsen et al., 2019). I will argue that the manual counts were still the right alternative to represent the population mean when constructing a correction factor. Manual counts are based on inspection of the entire fish and are performed by trained people. If it was possible to remove all potential biases, these estimates should be closer to the true population mean than the image-based counts. The Norwegian government has regulated salmon farmers based on manual lice counts for years, and the reported counts constitute the input to several sea lice dispersion models, which in turn inform the so-called 'traffic light system' that is used to regulate the opportunities for growth of the entire industry by production area (Sandvik et al., 2021; Myksvoll et al., 2018). It was thus a reasonable starting point to assume that manual counts correctly reflect the population mean.

Image-based lice counts in the Stingray system are performed on sequences of images that make the fish accessible for counting from different perspectives. The attachment distribution of salmon lice on the host fish does not occur at random (Bui et al., 2020). Instead, several authors have identified preferred attachment areas for adult female and mobile lice, particularly on the operculum and along the dorsal- and ventroposterior midlines (Jaworski and Holm, 1992; Todd et al., 2000). This means that a significant proportion of lice is attached close to the sagittal midline of the host fish, where they may be visible from both sides and do not need to be corrected for.

In the preliminary part of the analysis, we observed high R^2 values for simple linear models produced, at Location 2 and 4. These two locations stood out in terms of higher R^2 values for all of the simple linear models as shown in Table 7 through 10. A potential improvement on the final models could perhaps be made if Location 1 and 6 were removed.

Conclusion

In conclusion, different lice-counting methods are subject to different biases and constraint that may influence the respective estimates in opposite directions. These effects may be difficult to disentangle based on empirical data where sample sizes are constrained by operational concerns. In the absence of an objectively verifiable ground truth about the true lice situation in a pen or location, deriving a correction factor for only one of these constraints through linear mixed-effects modeling, remains a non-trivial problem.

Future work

The analysis done in this thesis may be improved by simply having bigger sample sizes for both manual and image-based counts, and performing it on pen level. The manual counts to be used for analysis should be counted by independent professional teams, and the louse estimates should not be reported to the government. This way, some of the potential ”reporting” bias may be reduced. However, counting on potentially hundreds of fish is physically challenging, if not impossible. Additionally, as we have seen in this thesis, collecting representative fish for counting is challenging in itself. Physically counting that many fish pose an arguably unnecessary risk to the welfare of

the salmon that are being counted. Handling of the fish should be restricted to a minimum. The benefit of conducting such an experiment may not be worth it. By including more samples, the gap between manual counts and image-based counts should however get smaller.

Even with more samples, the problem of not knowing ”the ground truth” still remains. We do not know what the population mean is in a pen. The fact that we do not know this, makes it even harder to determine what a good estimate is. Controlled experiments may be a possibility, if it is possible to mark the fish somehow and contain them in a smaller pen. That would however not remove the problem of having to handle the same fish multiple times. In addition, performing a controlled experiment like this would be extremely challenging.

The authors of (Jeong and Revie, 2020) discuss the use of prevalence to predict abundance. Abundance is the average of sea lice per fish, whereas prevalence is defined as the proportion of fish with one or more sea lice. It would be interesting to conduct further research on this topic in relation to computer vision-based counting, because it might solve the correction factor problem. If the proportion of fish with one or more lice can describe abundance, we do not need to see the whole fish. The risk of missing a louse on the side of a fish that can not be seen on images is still present, but the likelihood of that happening is smaller.

Correction factors may not be the way to go. It is still undoubtedly important to monitor sea lice severity in fish farms, and the goal should be to get as good estimates of the louse situation in a pen as possible. However, we might need to change our perspective on the issue concerning that we can not see a whole fish on images. The concentration of salmon lice may not be a good measure of the louse situation in a fish farm at all, independent of counting method. It may be possible to use the change in lice counts per some time unit, for example per day, to predict how much lice there is in a pen. For computer vision-based counting, a combination of prevalence-predicted abundance of large sample sizes and the rate of increase in louse numbers may perhaps be possible to use to predict louse abundance in the future.

References

M Baillie, F Lees, G Gettinby, and CW Revie. The use of prevalence as a measure of lice burden: a case study of lepeophtheirus salmonis on scottish atlantic salmon, salmo salar l., farms. *Journal of Fish Diseases*, 32(1):15–25, 2009.

Britt Bang Jensen, Lars Qviller, and Nils Toft. Spatio-temporal variations in mortality during the seawater production phase of atlantic salmon (salmo salar) in norway. *Journal of fish diseases*, 43(4):445–457, 2020.

Samantha Bui, Frode Oppedal, Lars Stien, and Tim Dempster. Sea lice infestation level alters salmon swimming depth in sea-cages. *Aquaculture Environment Interactions*, 8:429–435, 2016.

Samantha Bui, Frode Oppedal, Velimir Nola, and Luke T Barrett. Where art thou louse? a snapshot of attachment location preferences in salmon lice on atlantic salmon hosts in sea cages. *Journal of fish diseases*, 43(6):697–706, 2020.

Mark J Costello. Ecology of sea lice parasitic on farmed and wild fish. *Trends in parasitology*, 22 (10):475–483, 2006.

Cyril Delfosse, Patrick Pageat, Céline Lafont-Lecuelle, Pietro Asproni, Camille Chabaud, Alessandro Cozzi, and Cécile Bienboire-Frosini. Effect of handling and crowding on the susceptibility of atlantic salmon (salmo salar l.) to lepeophtheirus salmonis (krøyer) copepodids. *Journal of Fish Diseases*, 44(3):327–336, 2021.

Departementenes sikkerhets- og serviceorganisasjon Teknisk redaksjon. Helhetlig forvaltning av akvakultur for bærekraftig verdiskaping. Technical report, Departementenes sikkerhets- og serviceorganisasjon Teknisk redaksjon, 04 2023.

Ahmed Elmoslemany, Shona K Whyte, Crawford W Revie, and K Larry Hammell. Sea lice monitoring on atlantic salmon farms in n ew b runswick, c anada: comparing audit and farm staff counts. *Journal of Fish Diseases*, 36(3):241–247, 2013.

Per Fredrik F. Johnsen, Lotte Leming Rognsås, Jonas Erraia, Grønvik Oddbjørn, Sveinung Fjose, Atle Blomgren, Øystein Fjelldal, Roy Robertsen, Audun Iversen, and Thomas Nyrud. RAPPORT

RINGVIRKNINGER AV SJØMATNÆRINGEN I 2021. Technical report, Menon economics, 10 2022.

O Folkedal, SO Utskot, and J Nilsson. Thermal delousing in anaesthetised small atlantic salmon (salmo salar) post-smolts: A case study showing the viability of anaesthesia prior to delousing for improved welfare during treatment for salmon lice. *Animal Welfare*, 30(2):117–120, 2021.

Ole Folkedal, Lars Helge Stien, Jonatan Nilsson, Thomas Torgersen, Jan Erik Fosseidengen, and Frode Oppedal. Sea caged atlantic salmon display size-dependent swimming depth. *Aquatic Living Resources*, 25(2):143–149, 2012.

Martin Føre, Kevin Frank, Tomas Norton, Eirik Svendsen, Jo Arve Alfredsen, Tim Dempster, Harkaitz Eguiraun, Win Watson, Annette Stahl, Leif Magne Sunde, et al. Precision fish farming: A new framework to improve production in aquaculture. *biosystems engineering*, 173:176–193, 2018.

Benedikt Frenzl, Lars Helge Stien, David Cockerill, Frode Oppedal, RH Richards, AP Shinn, JE Bron, and Herve Migaud. Manipulation of farmed atlantic salmon swimming behaviour through the adjustment of lighting and feeding regimes as a tool for salmon lice control. *Aquaculture*, 424:183–188, 2014.

R Gautam, R Vanderstichel, AS Boerlage, CW Revie, and KL Hammell. Evaluating bath treatment effectiveness in the control of sea lice burdens on atlantic salmon in new brunswick, canada. *Journal of Fish Diseases*, 40(7):895–905, 2017.

Sean C Godwin, Martin Krkošek, John D Reynolds, and Andrew W Bateman. Bias in self-reported parasite data from the salmon farming industry. *Ecological Applications*, 31(1):e02226, 2021.

A Grimnes and PJ Jakobsen. The physiological effects of salmon lice infection on post-smolt of atlantic salmon. *Journal of Fish Biology*, 48(6):1179–1194, 1996.

Lars A Hamre, Kevin A Glover, and Frank Nilsen. Establishment and characterisation of salmon louse (lepeophtheirus salmonis (krøyer 1837)) laboratory strains. *Parasitology international*, 58 (4):451–460, 2009.

Lars A Hamre, Christiane Eichner, Christopher Marlowe A Caipang, Sussie T Dalvin, James E Bron, Frank Nilsen, Geoff Boxshall, and Rasmus Skern-Mauritzen. The salmon louse lepeophtheirus salmonis (copepoda: Caligidae) life cycle has only two chalimus stages. *PloS one*, 8(9): e73539, 2013.

Lars Are Hamre, Samantha Bui, Frode Oppedal, Rasmus Skern-Mauritzen, and Sussie Dalvin. Development of the salmon louse lepeophtheirus salmonis parasitic stages in temperatures ranging from 3 to 24 c. *Aquaculture Environment Interactions*, 11:429–443, 2019.

Terje Hatlem. Nye regler for lusegrenser om våren. *fisk*. URL `https://fisk.no/oppdrett/6441-nye-regler-for-lusegrenser-om-varen`.

Kari Olli Helgesen and Anja Bråthen Kristoffersen. Telling av lakselus - hvordan forstå og håndtere usikkerheten i telleresultatene. *Veterinærinstituttet*, 2018.

PA Heuch, JR Nordhagen, and TA Schram. Egg production in the salmon louse [lepeophtheirus salmonis (krøyer)] in relation to origin and water temperature. *Aquaculture research*, 31(11): 805–814, 2000.

Peter A Heuch, George Gettinby, and Crawford W Revie. Counting sea lice on atlantic salmon farms—empirical and theoretical observations. *Aquaculture*, 320(3-4):149–153, 2011.

Beate Hoddevik. Modeller: Viser hva som påvirker mengden lakselus. Technical report, Havforskningsinstituttet, 12 2021.

Okechukwu O Igboeli, John F Burka, and Mark D Fast. Lepeophtheirus salmonis: a persisting challenge for salmon aquaculture. *Animal Frontiers*, 4(1):22–32, 2014.

A Iversen, Ø Hermansen, R Nystøyl, and E Junge Hess. Cost development in farming of norwegian salmon. *Tromso: Nofima Report*, 46, 2017.

A Jaworski and J Chr Holm. Distribution and structure of the population of sea lice, lepeophtheirus salmonis krøyer, on atlantic salmon, salmo salar l., under typical rearing conditions. *Aquaculture Research*, 23(5):577–589, 1992.

Jaewoon Jeong and Crawford W Revie. Appropriate sampling strategies to estimate sea lice prevalence on salmon farms with low infestation levels. *Aquaculture*, 518:734858, 2020.

Jaewoon Jeong, Gabriel Arriagada, and Crawford W Revie. Targets and measures: Challenges associated with reporting low sea lice levels on atlantic salmon farms. *Aquaculture*, 563:738865, 2023.

Per Fredrik F Johnsen, Jonas Erraia, Oddbjørn Grønvik, Sveinung Fjose, Atle Blomgren, Øystein Fjelldal, Roy Robertsen, Audun Iversen, and Thomas Nyrud. Ringvirkninger av sjømatnæringen i 2020. *Rapport: MENONPUBLIKASJON NR*, 105:2021, 2021.

Mattilsynet. VEILEDNING OM KRAVENE TIL SØKNADER OM Å BRUKE NYE METODER FOR Å TELLE OG RAPPORTERE LAKSELUS. Technical report, Mattilsynet, 04 2020.

Adele Mennerat, Frank Nilsen, Dieter Ebert, and Arne Skorping. Intensive farming: evolutionary implications for parasites and pathogens. *Evolutionary biology*, 37:59–67, 2010.

Elena Myhre Jensen, Tor Einar Horsberg, Sigmund Sevatdal, and Kari Olli Helgesen. Trends in de-lousing of norwegian farmed salmon from 2000–2019—consumption of medicines, salmon louse resistance and non-medicinal control methods. *PLoS One*, 15(10):e0240894, 2020.

Mari Skuggedal Myksvoll, Anne Dagrun Sandvik, Jon Albretsen, Lars Asplin, Ingrid Askeland Johnsen, Ørjan Karlsen, Nils Melsom Kristensen, Arne Melsom, Jofrid Skardhamar, and Bjørn Ådlandsvik. Evaluation of a national operational salmon lice monitoring system—from physics to fish. *PLoS One*, 13(7):e0201338, 2018.

Jonatan Nilsson and Ole Folkedal. Sampling of atlantic salmon salmo salar from tanks and sea cages is size-biased. *Aquaculture*, 502:272–279, 2019.

Trond Nordtug, Bjarne Kvæstad, and Andreas Hagemann. Responses and preferences of salmon louse (lepeophtheirus salmonis krøyer 1836) copepodids to underwater artificial light sources. *Aquaculture*, 532:736036, 2021.

Frode Oppedal, Tim Dempster, and Lars H Stien. Environmental drivers of atlantic salmon behaviour in sea-cages: a review. *Aquaculture*, 311(1-4):1–18, 2011.

Liv Østevik, Marit Stormoen, Øystein Evensen, Cheng Xu, Kai-Inge Lie, Ane Nødtvedt, Hamish Rodger, Andreas Skagøy, Farah Manji, and Marta Alarcón. Effects of thermal and mechanical delousing on gill health of farmed atlantic salmon (salmo salar l.). *Aquaculture*, 552:738019, 2022.

Kathy Overton, Francisca Samsing, Frode Oppedal, Sussie Dalvin, Lars H Stien, and Tim Dempster. The use and effects of hydrogen peroxide on salmon lice and post-smolt atlantic salmon. *Aquaculture*, 486:246–252, 2018.

José Pinheiro and Douglas Bates. *Mixed-effects models in S and S-PLUS*. Springer science & business media, 2006.

Crawford W Revie, C Robbins, G Gettinby, L Kelly, and JW Treasurer. A mathematical model of the growth of sea lice, lepeophtheirus salmonis, populations on farmed atlantic salmon, salmo salar l., in scotland and its use in the assessment of treatment strategies. *Journal of Fish Diseases*, 28(10):603–613, 2005.

Crawford W Revie, E Hollinger, G Gettinby, F Lees, and PA Heuch. Clustering of parasites within cages on scottish and norwegian salmon farms: alternative sampling strategies illustrated using simulation. *Preventive Veterinary Medicine*, 81(1-3):135–147, 2007.

Yosiyuki Sakamoto, Makio Ishiguro, and Genshiro Kitagawa. Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel*, 81(10.5555):26853, 1986.

Anne D Sandvik, Samantha Bui, Mats Huserbråten, Ørjan Karlsen, Mari S Myksvoll, Bjørn Ådlandsvik, and Ingrid A Johnsen. The development of a sustainability assessment indicator and its response to management changes as derived from salmon lice dispersal modelling. *ICES Journal of Marine Science*, 78(5):1781–1792, 2021.

Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.

Standard Norge. Automatisk lusetelling i fiskeoppdrett - ny standardiseringskomité. *Standard Norge*. URL `https://standard.no/`

kurs-og-radgiving/standard-morgen/avholdte-standard-morgen/automatisk-lusetelling-i-fiskeoppdrett---ny-standardiseringskomite/.

Cecilie Sviland Walde, Britt Bang Jensen, Jostein Mulder Pettersen, and Marit Stormoen. Estimating cage-level mortality distributions following different delousing treatments of atlantic salmon (salmo salar) in norway. *Journal of fish diseases*, 44(7):899–912, 2021.

Geir Lasse Taranger, Ørjan Karlsen, Raymond John Bannister, Kevin Alan Glover, Vivian Husa, Egil Karlsbakk, Bjørn Olav Kvamme, Karin Kroon Boxaspen, Pål Arne Bjørn, Bengt Finstad, et al. Risk assessment of the environmental impact of norwegian atlantic salmon farming. *ICES Journal of Marine Science*, 72(3):997–1021, 2015.

Cameron RS Thompson, Angelico Madaro, Jonatan Nilsson, Lars Helge Stien, Frode Oppedal, Øyvind Øverli, Wayne J Korzan, and Samantha Bui. Comparison of non-medicinal delousing strategies for parasite (lepeophtheirus salmonis) removal efficacy and welfare impact on atlantic salmon (salmo salar) hosts. *Aquaculture International*, pages 1–29, 2023.

Trine Thorvaldsen, Kevin Frank, and Leif Magne Sunde. Practices to obtain lice counts at norwegian salmon farms: status and possible implications for representativity. *Aquaculture Environment Interactions*, 11:393–404, 2019.

Christopher D Todd, Alan M Walker, Jane E Hoyle, Sally J Northcott, Andrew F Walker, and Michael G Ritchie. Infestations of wild adult atlantic salmon (salmo salar l.) by the ectoparasitic copepod sea louse lepeophtheirus salmonis krøyer: prevalence, intensity and the spatial distribution of males and [2pt] females on the host fish. *Hydrobiologia*, 429:181–196, 2000.

Ole Torrissen, Simon Jones, Frank Asche, Atle Guttormsen, Ove Tommy Skilbrei, Frank Nilsen, Tor Einar Horsberg, and D Jackson. Salmon lice–impact on wild salmonids and salmon aquaculture. *Journal of fish diseases*, 36(3):171–194, 2013.

Trine Torvaldsen, Kevin Frank, and Leif Magne Sunde. Lusetellingsmetoder i lakseoppdrett. Technical report, SINTEF, 04 2018.

CS Tucker, C Sommerville, and R Wootten. Does size really matter? effects of fish surface area on

the settlement and initial survival of lepeophtheirus salmonis, an ectoparasite of atlantic salmon salmo salar. *Diseases of Aquatic Organisms*, 49(2):145–152, 2002.

Mathias Stølen Ugelvik and Sussie Dalvin. The effect of different intensities of the ectoparasitic salmon lice (lepeophtheirus salmonis) on atlantic salmon (salmo salar). *Journal of Fish Diseases*, 45(8):1133–1147, 2022.

Thijs Christiaan van Son, Anja Bråthen Kristoffersen, Hildegunn Viljugrein, Kari Olli Helgesen, Lars Qviller, and Peder Jansen. Forslag til håndtering av falske positive og negative lusetellinger ved lave lusetall. *Veterinærinstituttet*, 2016.

Hildegunn Viljugrein and Kari Olli Helgesen. Forslag til håndtering av falske positive og negative lusetellinger ved en lusegrense på 0,1. *Veterinærinstituttet*, 2017.

Alain F Zuur, Elena N Ieno, Neil J Walker, Anatoly A Saveliev, Graham M Smith, et al. *Mixed effects models and extensions in ecology with R*, volume 574. Springer, 2009.