Aquaculture Data Analysis

Konrad Krzysztof Jarczyk
University of Stavanger, Norway
kk.jarczyk@stud.uis.no

Sarvin Torkamanzadeh
University of Stavanger, Norway
s.torkamanzadeh@stud.uis.no

Olufisayo Akindele Togun
University of Stavanger, Norway
267994@uis.no

ABSTRACT

This research primarily focuses on implementing advanced data analysis techniques including code interpreter to explore potential links between environmental factors and aquaculture health in Norwegian waters. Utilizing extensive datasets from multiple fish farms, we apply coding methodologies, including decision trees and polynomial regression, to investigate the relationships between water temperature, sea salinity, and the prevalence of fish lice and two notable fish diseases. Despite the comprehensive data and our advanced analytical approaches, our findings reveal no significant correlations, highlighting the intricate nature of aquatic ecosystems. This study not only showcases the application of data analysis in marine biology but also points to the need for further research to unravel the complex dynamics affecting fish health and lice infestations in aquaculture settings.

KEYWORDS

Aquaculture, salmon lice, fish disease analysis

1 INTRODUCTION

Salmon aquaculture, an industry that has grown rapidly since the late 1970s, now produces over 1 million tonnes of salmon annually. This industry, primarily based on open net pens in coastal areas along wild salmon migratory routes, has greatly contributed to the global seafood supply. Despite its benefits, it poses significant ecological challenges, particularly in Norway, the world's leading salmon producer. This report investigates these challenges, with a special emphasis on the role of data analysis and machine learning in addressing the issues of salmon lice (Lepeophtheirus salmonis) and their impact on both farmed and wild salmon populations.

1.1 Salmon and the Looming Threat of Lice

Salmon lice, ectoparasitic sea creatures, thrive in dense aquaculture environments, preying on salmon and causing substantial health and economic impacts. [5]. They attach themselves to salmon, feasting on their mucus, skin, and blood. The infestation pressure originating from salmon farms poses a significant threat to both farmed and wild salmonids [5]. For farmed salmon, severe infestations can lead to reduced welfare and economic losses. For wild salmon populations, the situation is even more dire, with population declines attributed to infestations from salmon farms [21].

Efforts to control salmon lice at the scale of modern commercial salmon farms are logistically challenging, with up to 200,000 salmon in a single cage. Traditional chemotherapeutants have faced issues of drug resistance and environmental contamination, leading to their reduced usage. Alternative methods like mechanical and thermal delousing, though less harmful to fish, still have their drawbacks, including elevated mortality rates. Amid these challenges, biological control using 'cleaner fish' has emerged as a promising solution [21].

1.2 Cleaner Fish as a Solution

The use of cleaner fish (species such as lumpfish, wrasse, and others) for biological control of salmon lice marks a unique approach in the aquaculture industry. These cleaner fish opportunistically consume lice stages in tank and cage environments, offering a potential environmentally friendly solution. However, this method raises ethical considerations and requires meticulous behavioral acclimation and monitoring of cleaner fish within sea cages. While cleaner fish show promise, their efficacy on a commercial scale and cost-effectiveness still require further investigation [21].

1.3 The Norwegian Perspective

Norway, as the epicenter of the salmon aquaculture industry, faces unique challenges due to its vast production. With salmon farms often situated along wild salmon migration routes, infestation pressure on out-migrating smolts is a pressing concern. Furthermore, the constant need for delousing in farms has led to rapid resistance development in lice populations. In response, the industry has embraced innovations such as in-cage technologies to mitigate infestations [21].

1.4 Data Analysis and Machine Learning for a Sustainable Future

The complex interplay between farmed and wild salmon populations, cleaner fish management, and evolving lice resistance patterns calls for a holistic approach. Data analysis and machine learning techniques offer the potential to better understand these hidden dynamics. By harnessing data from various sources, including environmental factors, and treatment outcomes, predictive models can be developed to optimize cleaner fish deployment, delousing strategies, and environmental manipulation methods. These advancements are key to reducing the ecological footprint of salmon aquaculture and safeguarding the future of both farmed and wild salmon populations.

In this report, we explore the intersection of salmon aquaculture, salmon lice infestations, and the transformative potential of data analysis and machine learning. Drawing from existing research and insights, we delve deeper into the challenges and opportunities presented by these interconnected elements. As we navigate the landscape of Norwegian aquaculture, we use data-driven solutions that can shape a more sustainable future and solutions for this vital industry and the precious salmon it produces.

Supervised by Nejm Saadallah and Ruth Beatriz Pincinato.

Project in Computer Science (DAT620), IDE, UiS
2023.

1.5 Outlines

Section 2 outlines the background of aquaculture, focusing on its evolution and challenges, particularly in Norway, and also illustrates the ecosystem of the aquaculture industry. Section 3 elaborates on our methodology, detailing the datasets used and the integration of the Code Interpreter in data analysis. We then explore the relationship between fish lice, diseases, sea temperature and salinity, applying machine learning to find potential correlations. Our findings and discussions in Section 4 show the results and the complexities in managing salmon lice, and finally, the report concludes in Section 5 with reflections on our study and avenues for future research.

2 BACKGROUND

Aquaculture is a dynamic and multidisciplinary industry, serving diverse objectives such as producing protein-rich food, enhancing natural stocks, and recycling organic waste. It plays a crucial role in addressing global food demand, boasts a rich historical legacy, and operates across a spectrum of aquatic systems and production scales. Economic factors are pivotal, and practical field trips offer invaluable real-world insights [1].

In this report, we delve into the multifaceted world of aquaculture, with a particular focus on fish lice. These parasites pose significant challenges to the industry, and our analysis centers on the experiments conducted on relevant datasets. Before we delve into the details of our analysis, let's provide a brief, self-contained summary of the essential background information.

The aquaculture industry comprises a complex supply chain with various key components defined by[22] are:

- **Breeding Farm:** This is where juveniles (young fish) are produced.
- **Ongrowing Farm:** Juveniles are transferred here to be reared to specific weights and sizes before being transferred to commercial cages, where they remain isolated until harvest.
  Harvest and Transport: Harvested fish are stored in insulated tubs filled with seawater and ice for transport.
- **Processing Plant:** At processing plants, fish are received, processed, sorted, and packed into boxes with ice. These operations are registered in an Enterprise Resource Planning (ERP) system, and the products are stored until they are delivered to customers via trucks.

Our comprehensive analysis within this industry will concentrate on the management of fish lice infestations during the Ongrowing phase and their implications. This focus is vital considering that the Ongrowing stage is a crucial period for maintaining the health and welfare of the fish, which includes managing challenges such as lice infestations.

2.1 Aquaculture Business Ecosystem

In Figure 1 we tried to depict the industries in the aquaculture supply chain. some of the items are taken from the research [10] but some others are added from other investigations. We gave a brief description of each of the items [18] in the following:

Farm Sites[10]: These are designated areas where aquatic organisms are raised. They can be in natural water bodies or in man-made systems like ponds, cages, or tanks.
Monitoring and Controlling Fish Health: This involves regular checks and management practices to ensure the health of the aquatic organisms. It includes disease prevention, diagnosis, treatment, and water quality management.
Fish Food Producers: Companies or facilities that produce feed specifically formulated for aquatic species in aquaculture. The feed is designed to provide a balanced diet necessary for growth.
Sea-based Marine Infrastructure: Structures and technologies deployed in marine environments to support aquaculture operations, such as cages, nets, and feeding systems.
Hoist and Handling: Equipment and processes used for moving and harvesting aquatic organisms from their growing environments. This includes nets, cranes, and other mechanical systems.
Hatcheries[10]: Facilities where eggs are hatched and juvenile fish (fry) are reared until they can be transferred to on-growing systems or farm sites.
Energy: Represents the power supply needed for various operations in aquaculture, from pumping water to powering feeding mechanisms and maintaining optimal environmental conditions.
Skills and Training Education [10]: Educational programs and training sessions designed to equip workers with the necessary skills and knowledge for effective aquaculture practices.
Transportation: The logistics involved in moving aquatic organisms from one stage of the supply chain to another, such as from hatcheries to farm sites, or from farm sites to processing facilities.
Processing [10]: Facilities where harvested aquatic organisms are cleaned, filleted, packaged, and prepared for sale.

Each part is integral to the aquaculture system, ensuring the efficient and sustainable production of aquatic organisms.

Thus, the ecosystem includes biological components like the fish (within "Farm sites" and "Hatcheries") and their health management ("Monitoring and Controlling Fish Health"), Physical components such as the "Sea-based Marine Infrastructure" and "Energy" needed to maintain the system, Human components which cover "Skills and Training Education", "Processing", "Transportation", and the management of the entire system and finally the Economic interactions with "Fish Food Producers" and "Hoist and Handling", depicting the industry and market aspects.

2.2 Development of Aquaculture in Norway

The aquaculture industry in Norway has shown incredible development and innovation, positioning it as a global leader. This section examines the main characteristics of Norway's aquaculture development, including its rapid expansion, technological breakthroughs, and flexibility in response to market demands according to [12].

Industry Resilience and Growth: The Norwegian aquaculture industry demonstrated remarkable resilience with a strong recovery in 2021, achieving record-high revenues. This rebound from

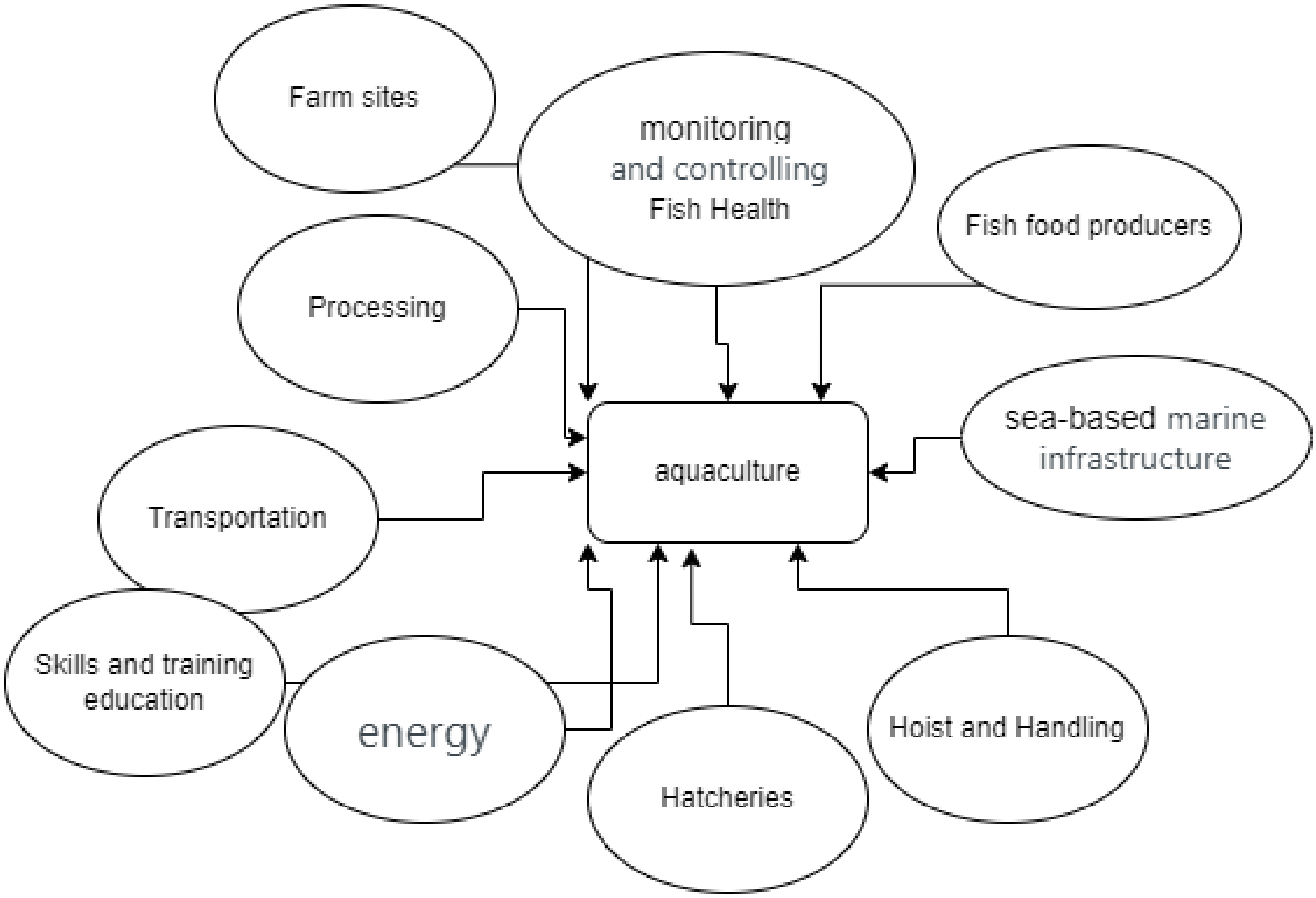

Figure 1: Aquaculture business Ecosystem

the pandemic-induced slowdown in 2020 highlights the industry's robustness and adaptability.

Technical Solutions: Driving Industry Innovation The technical solutions segment within aquaculture has been a bedrock of innovation and growth. Despite some challenges, this segment has consistently shown double-digit revenue growth, underlining its importance in advancing aquaculture technologies and practices.

Biotechnology: The biotechnology segment, encompassing feed production and cleaner fish companies, has also played a vital role in the industry's development. While facing some margin pressures, this segment reached an all-time high in revenues, indicating its significance in supporting sustainable aquaculture practices.

Market Dynamics and Consumer Demand The Norwegian aquaculture industry has adeptly navigated market dynamics, with salmon remaining a highly demanded product both domestically and internationally. The increased home consumption of salmon during the pandemic and the subsequent high demand post-pandemic underpin the industry's capacity to meet changing consumer preferences [12].

2.3 Salmon lice

Norway produces the most salmon worldwide. Since the inception of commercial salmon farming in Norway in the late 1960s, the industry has grappled with the pathogenic marine parasite, salmon louse (Lepeophtheirus salmonis). This parasite, known for attaching to the skin of salmon and feeding on mucus, blood, and skin, poses significant challenges to the health and sustainability of aquaculture operations [20]

Salmon lice infestations can have detrimental effects on salmon, including causing physical damage, and skin erosion, and leading to complications such as osmoregulatory failure, secondary infections, immunosuppression, and chronic stress. These health issues underscore the importance of effective management strategies in salmon farming[20]

2.3.1 Lifecycle and Types of Salmon Lice in Aquaculture. Salmon lice undergo a complex lifecycle comprising eight developmental stages, each of which presents distinct challenges for management in aquaculture. Understanding these stages is crucial for implementing effective control strategies.

Egg Stage: After mating, female lice carry eggs that eventually hatch into the water column.

Nauplius Stages: The first two stages are nauplius I and II. These are un-infective, free-living stages where the lice drift in the plankton.

Copepodid Stage: This is the infective stage where the lice are still free-living. Copepodids must find and attach to a host before their yolk reserve is depleted, which is a critical phase for successful development.

Chalimus Stages: Once attached to a host, the lice enter the sessile chalimus stages. During these stages, the lice are fixed in place on the host and undergo further development.

re-adult and Adult Stages: In these later stages, the lice become motile. They can move on and between hosts, making them more challenging to manage. These stages include sexually mature adults capable of reproduction, thus perpetuating the cycle [8].

2.3.2 Salmon Lice Regulations in Norwegian Aquaculture. Norwegian aquaculture is governed by strict regulations to control salmon lice (Lepeophtheirus salmonis), a critical step in safeguarding both farmed and wild salmon populations. The core aspects of these regulations are:

Lice Count Thresholds The Norwegian government has set specific lice count limits. For instance, in the most sensitive period for wild salmon smolt migration (usually spring), the maximum allowed average of adult female lice vary between different sources. BarensWatch says that is set at 0.5 [3], while Reuters and Science Norway says it is 0.2 [23][28]. Either way the maximum allowed average is very low, compared to other countries where salmon farming is significant [28]. Also farms are required to conduct weekly lice counts and must implement control measures if they exceed these thresholds.

Integrated Pest Management The regulations promote a combination of preventive and reactive strategies. This includes using cleaner fish, optimizing farming practices to prevent high lice concentrations, and employing non-chemical delousing methods when possible [17].

Spatial Management Farms located in close proximity must coordinate their lice management strategies. This collaborative approach aims to reduce the spread of lice and the emergence of resistant strains.

3 METHODOLOGY

This section delves into the methodologies used in our research, focusing on the application of data analysis within the aquaculture industry. Initially, we provide a detailed explanation of the datasets that form the foundation for the subsequent phases of our study. Our research methods are methodically divided into three distinct parts, each conducted by a different member of our three-person team. A big aspect of our analysis involved the utilization of a Code Interpreter, developed by OpenAI, which significantly aided in data processing and interpretation. The functionalities and contributions of this tool are comprehensively discussed in one subsection. This methodological overview aims to offer a transparent approach to our research into the aquaculture industry.

3.1 Datasets

In conducting our comprehensive analysis, we utilized several datasets, each contributing unique insights and dimensions to our study. Our primary data source, BarentsWatch, is an integral Norwegian platform based in Trømso, that provides extensive information on coastal and marine areas. Hosted at www.barentswatch.no, it offers a diverse range of services crucial for public agencies, industries, and the public, emphasizing sustainable management of sea and coastal regions. This collaborative initiative, steered by the Norwegian Coastal Administration under the Ministry of Trade and Fisheries, involves numerous ministries and research institutes. It's instrumental in our analysis due to its comprehensive data, which is pivotal in understanding and managing Norway's marine environments effectively [4].

- Salmon lice dataset includes the location name of fish farms and their geographical coordinates, the municipality, temperature, year, week, different stages of lice count, etc. It has 659,271 records.
- Fish diseases dataset shows disease information among fish farms. Records present information about which fish farm, their location, what disease, if the disease is confirmed or just suspected, how long it varied, etc. It has 71,102 records.
- Dataset with salinity of the water provided which comprises oceanographic data from the Barents Sea Opening, collected from mooring stations between August 1997 and April 2017. It includes temperature and current measurements to monitor Atlantic Water inflow. The array of moorings, positioned at various latitudes along approximately 20°E, was equipped with instruments like Aanderaa RCM7, RCM9, and Sea Guards, recording data every 20 minutes. This dataset is essential for understanding oceanic and climatic dynamics in the region [14]

Lice types in our dataset. Three distinct stages of salmon lice correspond to the various phases in their lifecycle. These 3 stages are adult female lice, lice in moving stages which corresponds with the pre-adult and adult stage, and attached lice which corresponds to the "Chalimus Stages" of the salmon louse life-cycle. The numbers provided in the table for the lice counts appear to be counts or proportions of salmon lice in different stages of development per fish and it is a value between 0 and 1. For example, a value of 0.02 in the "Attached lice" column could suggest that for every fish sampled, there is an average of 0.02 attached lice, which would mean that not every fish is expected to have an attached louse, reflecting a low level of infestation.

Disease types. Within all rows in the second dataset there are two types of diseases ILA and PD. Infectious Salmon Anemia (ILA) is a highly contagious viral disease affecting salmon, characterized by severe anemia, lethargy, and high mortality, making its early detection and control vital for aquaculture sustainability. Pancreas Disease (PD), caused by the Salmonid alphavirus, leads to significant losses in salmon and trout farming due to its impact on the pancreas and heart, necessitating robust monitoring and vaccination strategies for effective management[6].

Salinity in the dataset. The dataset that contains the salinity of the water is an array type of data, unlike the other 2 datasets which were CSV. We had to use the netCDF4 library to read the file. A netCDF file is commonly used for storing multidimensional scientific data [18]. The area lies along the northern coast of Norway, extending northeastward towards the Arctic Ocean. It represents a critical juncture where the warmer Atlantic waters flow into the colder Barents Sea, making it a key location for studying oceanic and climatic interactions. The placement of these moorings provides valuable insights into the dynamics at the western entrance of the Barents Sea, an area pivotal for understanding the environmental and ecological impacts of ocean currents and temperature changes in the Arctic region. It does not only include temperature and current measurements but also contains salinity data, crucial for understanding the marine environment. Salinity, a measure of the salt concentration in water, is an essential factor in determining the characteristics of seawater, including its density and circulation patterns

3.2 Role of Code Interpreter in Data Analysis

Understanding the Code Interpreter. Code interpreter is a tool developed by Open AI to process and analyze data. The tool is available at ChatGPT, where we can upload dataset in various formats and ask with using prompt to do particular tasks. For example to find how many records are in dataset, which was used in our previous subsection for dataset explanation.

It operates on advanced artificial intelligence, trained initially on a vast range of texts to understand language and code syntax. It is then fine-tuned with targeted datasets and human feedback, improving its ability to interpret and generate code. The model excels in pattern recognition and structure analysis, though it doesn't possess a human-like understanding of code. Its key feature is the ability to translate natural language inputs into programming code, making it accessible to users without coding expertise[9].

Application in our research. In the scope of our research, we used the Code Interpreter for handling and analyzing datasets, particularly to perform various preprocessing tasks, such as data cleaning, normalization, and transformation. We utilized it to determine optimal methods for specific functions and to arrange data in a particular order. For instance, one task involved entering specific commands to translate column names from Norwegian to English. Figure 2 illustrates an example of this process, showing how the Code Interpreter facilitated efficient data manipulation.

Limitations. Despite its many advantages, we encountered some limitations with the Code Interpreter. At times, when executing simple operations or functions, it faced difficulties in loading data, resulting in errors. Typically, a refresh or initiating a new chat and re-uploading the dataset resolved this issue. Another common limitation was observed when attempting more complex data manipulations; the Interpreter would either not complete the task or abruptly stop. For instance, general prompts like "Find correlations in the dataset" often failed to yield meaningful results. Success was more likely with specific and concise prompts, outlining exact correlation types and data operations required. However, even then, if the task involved multiple operations, the code often broke midway. Additionally, there were instances where the Interpreter applied incorrect methods, leading to inaccurate results. This underscores the importance of having a solid understanding of syntax and data analysis to verify the operations performed by the AI.

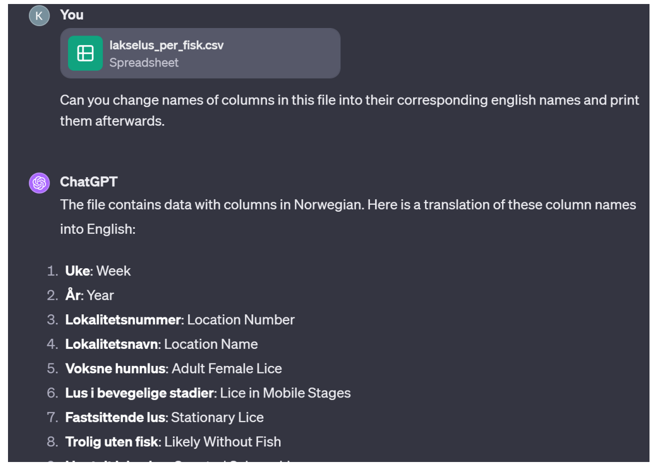

Figure 2: Example of our usage of Code Interpreter

Summary. The Code Interpreter significantly improves data analysis, effectively doubling the speed of traditional programming. Its ability to facilitate a more interactive or efficient engagement with data is a big help. However, while it streamlines many aspects of data processing, a foundational knowledge in programming and data analysis is crucial for its effective usage. This ensures that the user can accurately guide the tool and critically assess its outputs and thereby maximizing the benefits of processing data analysis.

3.3 Applied Research Methods

3.3.1 Diseases vs fish lice.

Problem formation. Regarding the formation of this problem, we began by examining the presented datasets: first about fish lice and second about diseases in fish farms. This led us to an intriguing question: Is there a connection between fish lice infestations and disease outbreaks in fish farms? Intuitively, one might expect there to be a link. To explore this further, we analyzed the data, considering the possibility that different stages of lice infestation might influence the likelihood of certain diseases. This research aims to uncover any patterns or correlations between types of fish lice and the diseases they may promote. By doing so, we hope to better understand how lice infestations can affect the health of fish in aquaculture settings and possibly contribute to disease spread.

Preprocessing. To make easier a broad comparison between the datasets on fish lice and fish farm diseases, we selected and cleansed key fields such as type of lice, location, time period, and disease characteristics. In the fish diseases dataset, we noted that each record included 'from date' and 'to date' columns, indicating the actual duration of the disease outbreak on the farm. This pivotal information enabled us to classify disease occurrences into specific quarters of the year, based on their reported durations. Figure 3 shows how it works. This classification system ensures that each disease event is accurately represented within a defined time frame, aligning with the seasonal variations that are important in understanding disease patterns in aquaculture [27].

ChatGPT

Sure, let's go through some examples to demonstrate how the updated function classifies dates into quarters based on the majority duration. We will consider different scenarios:

1. **Scenario 1**: A span that fits entirely within one quarter.
   - **Dates**: From `2023-01-10` to `2023-03-20`.
   - **Expected Classification**: This span is entirely within Q1 of 2023 (January to March).
2. **Scenario 2**: A span that crosses two quarters but with the majority in one quarter.
   - **Dates**: From `2023-03-20` to `2023-05-15`.
   - **Expected Classification**: This span starts in Q1 but the majority falls in Q2 of 2023 (April to June). Hence, it should be classified in Q2.
3. **Scenario 3**: A span that crosses multiple quarters evenly.
   - **Dates**: From `2023-02-01` to `2023-08-01`.
   - **Expected Classification**: This span covers parts of Q1, Q2, and Q3 of 2023. The majority of February is in Q1, the entirety of May, June, and July falls in Q2 and Q3. So, it should be classified in Q1, Q2, and Q3.
4. **Scenario 4**: A span that is very short and falls at the end of a quarter.
   - **Dates**: From `2023-03-29` to `2023-03-31`.
   - **Expected Classification**: This short span is entirely within the last few days of Q1 of 2023. It should be classified in Q1.

Figure 3: Quarter classification

	LocalityNumber	Quarter	Disease	Adult Female Lice	Movable Lice	Attached Lice
1	12067	2019-Q3	PD	0.01	0.12	0.03
2	12067	2019-Q3	PD	0.03	0.09	0.03

Figure 4: Merged dataframe

Similarly, in the dataset concerning fish lice, we paid special attention to the 'week' and 'year' fields. These were used to determine the corresponding quarter of the year for each lice infestation record.

Now that we have categorized both datasets into quarters, we are in a position to merge these two data frames into one data frame. The merging is conducted based on two key parameters: the quarter of the year and the location number. An important aspect of this process is handling cases where the duration of disease outbreaks extends beyond a single quarter. In such instances, we employ a strategy of duplicating the relevant record across each quarter that the outbreak spans. This method ensures that our merged dataset accurately reflects the prolonged nature of certain disease occurrences. On the Figure 4 we can see how merged data frame looks. It contains the location number, quarter of year, type of disease and fish lice information.

Data Normalization and Correlation Analysis Using Apriori Algorithm. In our data analysis phase, we begin by normalizing lice counts into five distinct categories: Very Low, Low, Medium, High, and Very High. This categorization is performed using the sklearn library, specifically employing the minmaxScaler function. The categorization into bins is achieved using the 'cut' command. Notably, this binning process is value-based rather than using preset limits, which means that the bins may not contain equal numbers of records but are instead tailored to the distribution of the data [25].

Following this, we apply the Apriori algorithm to uncover potential correlations between the categorized sea lice levels and various fish diseases. The Apriori algorithm is a classic data mining tool used in transactional database analysis, known for its efficiency in finding frequent itemsets and relevant association rules. It operates on the principle of identifying items that frequently occur together, in our case, the occurrence of certain lice levels alongside specific diseases. This method is particularly effective in our study, as it allows us to discover if certain levels of lice infestation are consistently associated with specific diseases, which can be helpful for preventive strategies in aquaculture [18].

Moreover, the Apriori algorithm's ability to handle large datasets and identify significant correlations makes it an ideal choice for our research. By setting appropriate support and confidence thresholds, we can filter out the most meaningful and robust associations, thereby focusing on the most impactful lice-disease correlations.

All the methodologies and techniques employed in this analysis, including the use of the sklearn library for normalization and the application of the Apriori algorithm, are comprehensively documented and illustrated in our code repository. This can be found in reference [15] on GitHub, under the file name 'disease_vs_lice'.

3.3.2 Sea temperature and fish lice count. Our analysis of the correlation between water temperature and fish lice acknowledges foundational research like Nilsson et al. (2023) [16]. They established a positive correlation between higher water temperatures and increased salmon lice detachment rates using hierarchical logistic regression models in a controlled setup. Their insights into optimal thermal treatment conditions for lice in aquaculture, balancing delousing efficiency and fish welfare, provide crucial context for our study. Drawing from this, we analyzed temperature and lice count data from 1508 fish farm locations, with a focused case study on Smørdalen fish farm, which has 524 records. Smørdalen is located in Masfjorden, on the west coast of Norway. It's one of the sites operated by the Institute of Marine Research for Atlantic salmon research farming. Masfjorden is situated in the Hordaland county, which is part of the Vestland region. This area is known for its fjords and is a significant region for Norwegian aquaculture, specifically salmon farming. When mentioning Smørdalen in your report, you can refer to it as a research farming site in Masfjorden, Vestland, Norway, operated by the Institute of Marine Research [13]. Our methodology diverges as we employed heat maps for visualization, examining location-specific relationships between temperature and lice counts across multiple farms, including an in-depth analysis of 1,103 records from various farms We tried to find out if the same direct relation can be found from the dataset of salmon lice. We used the attribute temperature in the dataset with their corresponding salmon fish lice count. In order to get a more limited contributing factor and make the analysis more pointed to a specific location we decided to select one fish farm location. The dataset has about 1508 fish farm locations number but the selection of the one fish farm was based on the fish farm having the most records. Among all of them, the fish farm with the most records was named Smørdalen. Smørdalen has 524 records of lice count with corresponding temperature data. We used a heat map to see the correlation between the temperature and the number of different lice-type counts. To get a broader result that included more farms than only Smørdalen we did the same analysis of location-specific relation between temperature and lice type count between all the unique fish farms found in the dataset. We refined the dataset for enhanced accuracy in our correlation study. For example, we excluded fish farms with only a single entry and those with no variation in key metrics like sea temperature and lice counts, as these could skew our results. This careful selection process allowed us to focus on more reliable data, ensuring our correlation analysis between sea temperature and lice counts across various fish farms was grounded in robust and meaningful data. The inclusion of location details further contextualized our findings, vital for understanding the environmental impact on aquaculture. Eventually, we found the correlation for 1,103 records of different fish farms. These approaches can be found in [15] on GitHub, under the file name 'correlation-for-all-farms.ipynb' and 'smørdalen.ipynb'.

3.3.3 Salinity, Sea Temperature, and fish lice . The analysis of sea temperature and fish lice count was conducted, prior to this it was observed that cleaner fish data was unavailable on [4] from 2018 until date. The data set for fish lice count was cleaned and a further analysis was conducted on the whole data to determine the correlation between Adult female fish lice count and temperature, attached lice and temperature, and temperature and lice in the mobile stage and temperature. Some visualization was done to determine which regions along the coast of Norway had adult female fish lice counts greater than 0.5 for the years 2012 and 2016.

Machine learning algorithms were trained using the sea lice dataset, predictions were made and the accuracy of the prediction was measured. The machine learning algorithm regressive decision tree and the random were trained to achieve good results.

Utilizing the Decision Tree. A decision tree is a popular supervised machine learning algorithm used for both classification and regression tasks. It models decisions as a tree-like structure, where each internal node represents a feature or attribute, each branch represents a decision rule, and each leaf node represents a class label (in classification) or a numerical value (in regression)[26]. The regression decision tree algorithm was used to predict the number of adult female fish lice counts and extrapolate the salinity to the data from Barenswatch. The decision tree was trained using a maximum depth of 5

Utilizing the Random Forest: Random Forest is a versatile and powerful machine-learning algorithm used for both classification and regression tasks. It operates by constructing a multitude of decision trees during training and outputting the class that is the mode of the classes (for classification) or mean prediction (for regression) of the individual trees. By averaging multiple deep decision trees, trained on different parts of the same training set, Random Forest mitigates the risk of overfitting. It also offers insights into the importance of each feature in making predictions, which is beneficial for feature selection. Despite its effectiveness, it can be computationally intensive and less interpretable compared to simpler models like single decision trees. The random forest was trained to predict adult female fish lice count using a number of estimators of 20 and a depth of 4. This can be found in reference [15] on GitHub, under the file name 'Salinity, Sea Temperature and fish lice.ipynb'.

Sea salinity vs fish lice. The salinity data was extracted and it was found to have seven longitude and latitude locations, the data contained pressure measurements, salinity measurements and pressure measurements over varying distances, the data was time-calibrated and cleaned crudely using Python code. The salinity data was extrapolated unto data from barenwatchs using longitude, latitude, week and year. The extrapolation was limited to the years between 2012 to 2017 and between a latitude between 67 and 73 and a longitude between 17 and 20. This can be found in reference [15] on GitHub, under the file name 'extract salinity and store as csv.ipynb'.The decision tree, linear regression and polynomial regression were used to extrapolate data unto the data from barenswatch and the accuracy was calculated.

Utilizing the Linear Regression: Linear regression is a fundamental statistical and machine learning technique used for modelling the relationship between a dependent variable (target) and one or more independent variables (features) by fitting a linear equation. It aims to find the best-fit line (or hyperplane in higher dimensions) that minimizes the sum of squared differences between the predicted and actual values. The linear regression algorithm is used to extrapolate the salinity to the data from Barenswatch[29].

Utilizing the Polynomial Regression: Polynomial regression is an extension of linear regression that allows for modelling nonlinear relationships between variables. Instead of fitting a straight line, it fits a polynomial curve to the data. This is achieved by adding higher-order polynomial terms (e.g., quadratic, cubic) to the regression equation. Polynomial regression is particularly useful when the relationship between variables cannot be adequately captured by a linear model. However, choosing an appropriate degree of the polynomial is important to avoid overfitting. Polynomial regression finds applications in fields like physics, engineering, and economics where nonlinear relationships are common[19]. The Polynomial regression is used to extrapolate the salinity to the data from Barenswatch. The polynomial features of a degree of 3 and an include_bias gave the best optimal results while training this machine learning Algorithm. This can be found in reference [15] on GitHub, under the file name 'Salinity, Sea Temperature and fish lice.ipynb'.

4 RESULTS AND DISCUSSION

In this section, we present the findings from our research in the aquaculture field.

4.1 Diseases vs fish lice

In analyzing the relationship between sea lice and fish diseases, our results revealed some notable correlations, particularly in the case of Infectious Salmon Anemia (ILA). We observed a moderate presence of adult female lice in instances where ILA was prevalent. This suggests that occurrences of ILA are often accompanied by increased numbers of adult female lice. For the other fields there was always very low amount of fish lice for both ILA and Pancreas Disease (PD). On Figure 5 we can see the observation of this correlation. This observation is supported by our analysis of the normalized bins for all three types of lice, where the majority of records fell into the 'Very Low' category. Figure **??** illustrates this trend, showcasing the distribution of records across different bins for adult female lice and attached lice, offering a clearer understanding of the lice distribution during disease outbreaks.

Furthermore, in our application of the Apriori model to identify potential associations, we experimented with various configurations of minimum threshold and support levels. Despite the adjustments in these parameters, the results consistently pointed towards

	Disease	Adult Female Lice Category	Movable Lice Category	Attached Lice Category
0	ILA	Medium	Very Low	Very Low
1	PD	Very Low	Very Low	Very Low

Figure 5: Disease vs Lice Correlation

```
Value counts for Adult Female Lice Category:
Adult Female Lice Category
Very Low     355203
Low          322083
Medium       298847
High         254895
Very High     64298
Name: count, dtype: int64


Value counts for Movable Lice Category:
Movable Lice Category
Very Low     325346
Low          324403
Medium       322369
High         258702
Very High     64506
Name: count, dtype: int64
```

Figure 6: Distribution for adult female lice and attached lice

	Positive	Negative
Adult female lice	632	471
Lice in moving stages	618	485
with Attached lice	628	475

Table 1: Number of positive and negative correlation between lice and temperature in entire dataset

the same conclusions. This consistency in findings, regardless of the parameter tuning, underscores the robustness of the correlations we identified.

4.2 Fish lice and sea temperature

The result we saw from Smørdalen fish farm is in contrast with the knowledge that we already had according to [16]. The result here shows that the sea temperature has a negative correlation with the number of fish lice count especially in the moving stage Figure 7. This could imply the fact that the data we are using to analyze the relation between temperature does not include all the influencing factors in the Smørdalen farm. Figure 1 shows this relation in the timeline several years.

Table 1 shows the number of positive and negative correlations between temperature and different fish lice type counts in 1,103 records of the different fish farms.

Out of the 1,103 records of different fish farms and their correlation, most correlations between the temperature and the different lice types are positive. The overall positivity is not as strong as we expected but still, we can say that the average is a positive number. The figure 8 shows that the averages stand at 0.03 and 0.02 positions for different types of lice in overall fish farm locations. The figure shows a histogram of different correlations in all fish farms and the red line shows the average number. It is almost flat since it is very close to 0 but it is still on the positive side.

4.3 Salinity, Sea Temperature, and Sea lice

Machine learning methods were applied to predict the fish lice count and extrapolate salinity data to the already existing data set. Some further analysis is done on sea lice data and sea temperature.

Impact of Sea Temperature on Sea Lice Infection and Development. Studies have shown that temperature significantly affects the developmental rates, infectivity, and survival of Lepeophtheirus salmonis. It has been highlight that sea temperature plays a critical role in the production and spread of copepodids of L. salmonis, influencing egg production, larval development, and larval survival[11]. This aligns with the observed correlation between adult female lice and temperature, where a relatively low correlation value (0.0066) as shown in Table 2 suggests that while temperature has an effect, other factors may also play a significant role[11].

Additionally, temperature and salinity are identified as major environmental factors affecting the developmental rate of salmon lice[11]. This is consistent with the slight negative correlation observed between salinity (determined via a decision tree) and adult female lice (-8.8947e-05), indicating a complex interaction between environmental factors and lice prevalence.

Temperature-Dependent Infection Success. The infection success of L. salmonis varies significantly with temperature. It was found that infection success was only 2% at 5°C[24], increasing to 40-50% at higher temperatures like 15 and 20°C. This variation in infection success across temperatures is critical for understanding the dynamics of sea lice infestations in different sea temperature conditions. It is particularly relevant when considering the correlations in varying years, such as the correlation between adult female lice and temperature in 2017 (0.0045) as shown in Table 2, suggesting a possible fluctuation in infection success rates over time.

Infection Success and Parasite Loss at Different Temperatures. [11] also report that temperatures below 10°C significantly increase the infection success of salmon lice, with negligible parasite loss at intermediate temperatures between 6 and 21°C. This finding is crucial in explaining the observed correlations and underscores the impact of temperature on the lifecycle of L. salmonis. The study also indicates a positive correlation between temperature and Combined Infection Success and Survival (CISS), especially from 6°C to 15°C.

4.3.1 Fish lice and Salinity. [7] investigates the impact of salinity on the survival, infectivity, and behavior of the sea louse Lepeophtheirus salmonis. It was discovered that the survival of L. salmonis copepodids reduces substantially in salinity levels below 29 parts per thousand (ppt)[7]. This survival decrease is more severe as salinity further drops. Additionally, the study shows that copepodids'

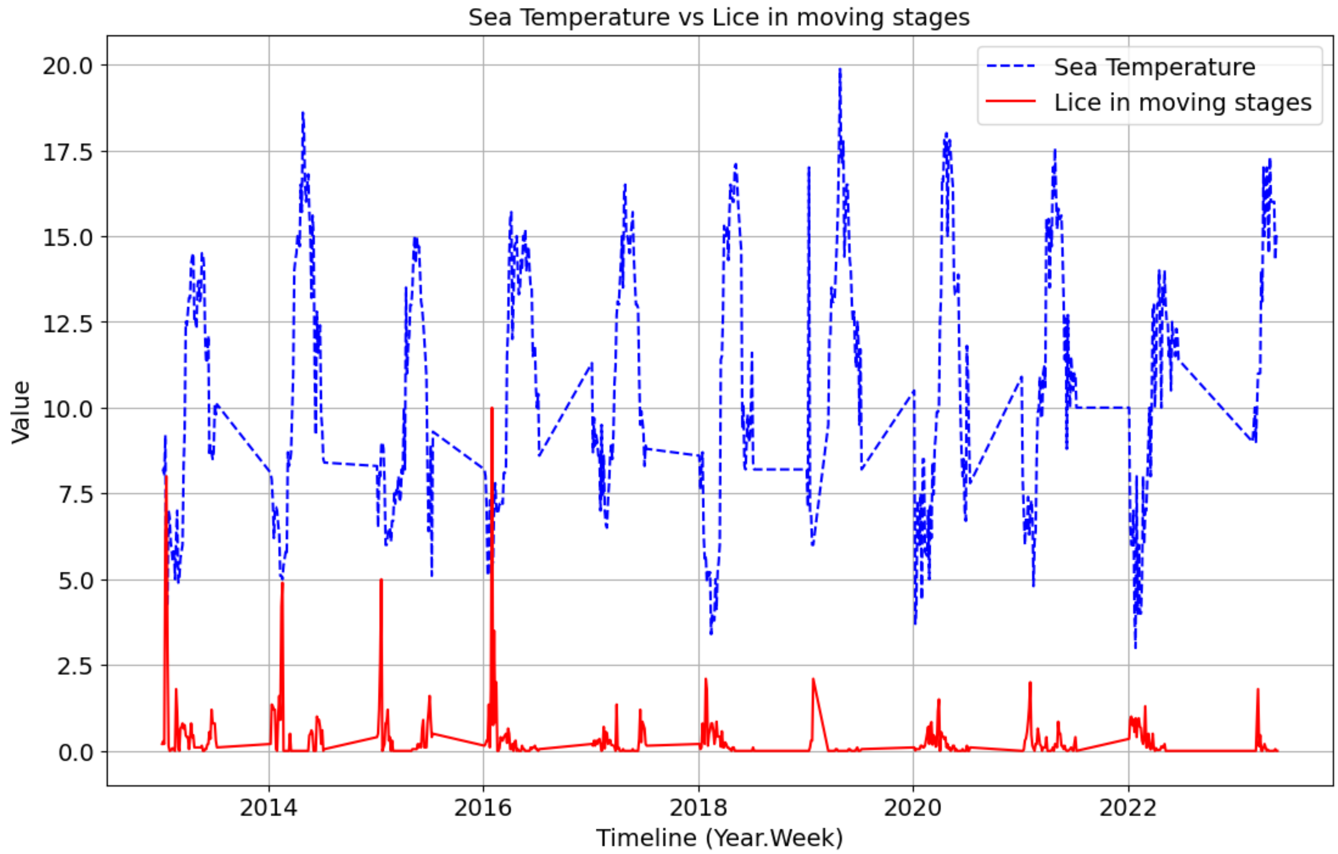

Figure 7: Smørdalen temperature and lice in moving stage trend in time

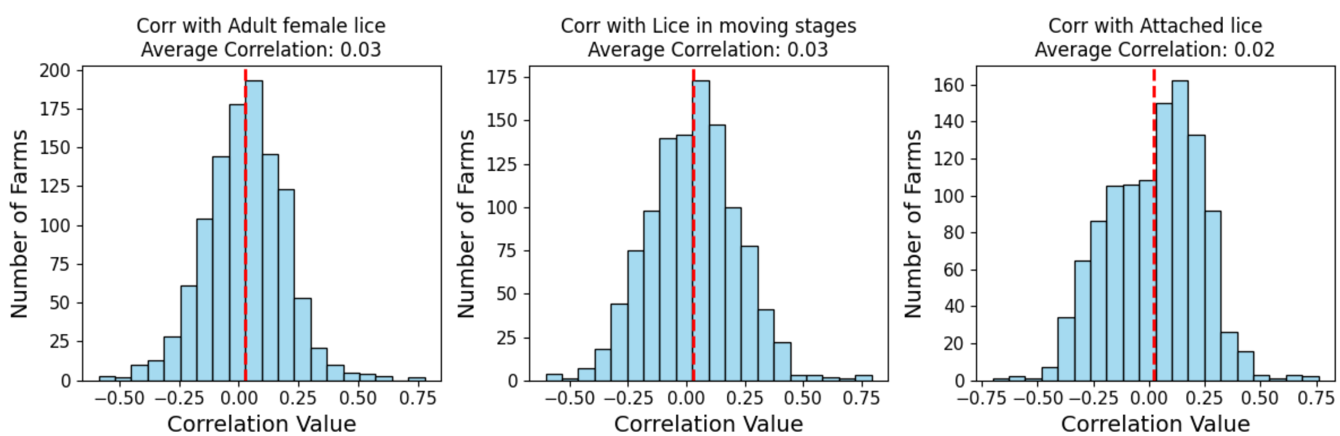

Figure 8: Average of all the correlations

capacity to infect hosts decreases in low salinity settings, which may indicate a problem with their awareness of or reaction to host stimuli. These results suggest that salinity and other environmental conditions can have a major impact on the dynamics of sea lice infestations in aquatic environments..

This investigation is crucial to validate the machine learning finding of a -0.000089 correlation between sea salinity and adult female lice. The study shows that decreased salinity levels have a detrimental effect on sea lice survival and infectivity, these facts are supported by the tiny but negative correlation discovered by

Table 2: Correlation Analysis of Various Factors with Temperature and Salinity

Correlation Factor	Correlation Value
Correlation between Adult Female Lice and Sea Temperature	0.0066
Correlation between Lice in Mobile Stage and Sea Temperature	0.0319
Correlation between Attached Female Lice and Sea Temperature	0.0041
Correlation between Adult Female Lice and Sea Temperature (2017)	0.0046
Correlation between Salinity (Decision Tree) and Adult Female Lice	-0.000089

the machine learning algorithm. This is consistent with the knowledge that the life cycle and behavior of marine parasites such as L. salmonis are significantly influenced by environmental factors. The machine learning analysis's connection can be understood to suggest a small but significant impact of salinity on the prevalence of lice.

We utilized three different models – Decision Tree, Linear Regression, and Polynomial Regression – to extrapolate salinity data from measurements taken at a depth of 50 meters, using columns such as longitude, latitude, year, and week. The primary dataset for temperature was sourced from a depth of 4 meters. The accuracy of these models varied, with the Decision Tree model yielding the most realistic and consistent results.

The Decision Tree model demonstrated remarkably high accuracy, with train and test data accuracies of approximately 99.99% as shown in Table 3. This high performance can be attributed to the model's ability to handle non-linear relationships and interactions between the features effectively. The decision tree likely captured the complex dependencies of salinity on spatial and temporal variables with high precision. Additionally, the depth disparity between the salinity and temperature measurements might not have significantly impacted the model's performance, suggesting robustness in diverse environmental conditions.

In contrast, Linear Regression showed relatively low accuracy, around 29%. This suggests that the relationship between the chosen features and salinity is not linear or that linear regression is too simplistic to capture the complexities of oceanographic data. The substantial difference in the depths of salinity and temperature measurements could have introduced variables that linear regression failed to account for, leading to lower accuracy.

Polynomial Regression, while more sophisticated than linear regression, yielded moderate accuracies (around 74%). The improved performance over linear regression indicates that higher-order terms helped in modelling the non-linear relationships. However, the model may have struggled with overfitting or failed to capture all complexities in the data. The significant difference in the depth of salinity and temperature readings could have introduced complexities that polynomial regression partially addressed but did not fully resolve. The mention of very large numbers in the polynomial regression outputs suggests that the model may have been overly sensitive to certain data points or feature interactions, leading to unrealistic extrapolation in some cases.

The depth disparity between salinity (50m) and temperature (4m) measurements is a critical aspect to consider. This difference introduces a vertical stratification challenge, as oceanographic parameters can vary significantly with depth. The high accuracy of the decision tree model suggests it could discern patterns despite this stratification, while the linear and polynomial models may have been misled by the depth-related variations.

4.3.2 Predicting Sea lice. In the task of predicting adult female sea lice, both the Decision Tree and Random Forest models showcased similar levels of accuracy, as evident in the training and testing phases. The Decision Tree model exhibited a training accuracy of 72.38% and a testing accuracy of 72.48%, while the Random Forest model showed a slightly lower training accuracy of 72.03% but a comparable testing accuracy of 72.40%.

The close performance of these two models is noteworthy, especially considering the inherent differences in their computational complexity and mechanisms as shown in Table 4. The Decision Tree is a simpler model, easier to interpret, and less computationally demanding. It works by creating a tree-like model of decisions based on the input features. On the other hand, Random Forest, an ensemble method, builds multiple decision trees and merges them to get a more accurate and stable prediction. It's typically more computationally intensive due to the creation and combination of multiple trees.

4.4 Reflection on results

In reflecting on the challenges posed by sea lice in aquaculture, it becomes evident that a multitude of factors contribute to the escalating infestations. Notably, the proximity of aquaculture farms to the migratory routes of wild salmon significantly increases the risk of lice transmission, with fish farms providing ideal conditions for lice breeding. This issue is exacerbated by the unusual commingling of wild and farmed salmon populations in these areas, disrupting the natural separation that typically limits lice spread. Additionally, the large concentration of hosts in fish farms amplifies the infestation rates, affecting areas up to 40 miles from the farms. The survivability of sea lice outside their hosts furthers this spread, bridging the gap between wild and farmed populations. Concerningly, the potential development of resistance to common treatments like the veterinary drug Slice poses a significant challenge in managing these infestations. Moreover, the particular vulnerability of juvenile salmon, especially scale-less species like pink and chum, heightens the impact of these infestations, as they suffer more severe health consequences. These factors collectively underscore the complexity and severity of sea lice challenges in aquaculture, necessitating a multi-faceted approach to management and mitigation [2].

Table 3: Accuracy of Models for Extrapolating Salinity Data

Model	Train Data Accuracy	Test Data Accuracy
Decision Tree	0.99993	0.99987
Linear Regression	0.29261	0.29351
Polynomial Regression	0.74473	0.74219

Table 4: Accuracy of Models for Predicting Female Sea Lice

Model	Train Data Accuracy	Test Data Accuracy
Decision Tree	0.72384	0.72482
Random Forest	0.72030	0.72399

5 CONCLUSION AND FURTHER WORK

Our investigation, utilizing sophisticated data analysis models, revealed intricate dynamics within the aquatic ecosystem. A key finding was the lack of a pronounced correlation between water temperature and the prevalence of sea lice, suggesting the complex interplay of environmental factors in these ecosystems. Our comprehensive dataset, although extensive, may not have captured all these nuanced variables, highlighting the dynamic nature of aquaculture settings.

In our analysis about fish lice and fish diseases, we observed a moderate presence of adult female lice in cases of Infectious Salmon Anemia (ILA). This pattern indicates a potential relationship between adult female lice and ILA occurrences. However, for both ILA and Pancreas Disease (PD), we found that the levels of attached and motile lice were predominantly low. This observation suggests a nuanced interaction between different types of lice and fish diseases. Our results emphasize the importance of considering the specific type of lice when studying their impact on fish health.

The intricate relationship between environmental factors like sea salinity and temperature with the life cycle of sea lice, particularly L. salmonis, was also evident in our findings. Temperature influences key aspects such as the developmental rate, infection success, and survival of sea lice, adding layers of complexity to our study. These findings underscore the need for a holistic approach that considers multiple environmental variables in the management of sea lice in marine ecosystems.

Our study's use of decision tree models showcased their high accuracy in navigating complex environmental data, especially in scenarios involving non-linear interactions and varying depths. In contrast, linear and polynomial regression models had limitations, likely due to their inability to grasp the multifaceted relationships within oceanographic data, which is often influenced by vertical stratification.

Looking ahead, future research should focus on depth variability and the potential for non-linear interactions in oceanographic datasets. This approach will aid in selecting the most appropriate modeling techniques. Our study revealed that simpler models, like Decision Trees and Random Forests, can perform comparably to more complex models in predicting female sea lice, highlighting the significance of model selection based on dataset characteristics and computational resources. This insight is particularly valuable for practitioners and researchers, as it suggests that simpler models may offer adequate effectiveness, especially where interpretability and computational efficiency are priorities.

Further improvements to our work could include obtaining a dataset with salinity measurements along the coast of Norway to enhance the accuracy of salinity extrapolation. Additionally, employing polynomial regression to predict salinity based on temperature, precipitation, and evaporation could yield more precise results. Exploring other machine learning models, like linear regression, convolutional neural networks, and time series models such as Long Short-Term Memory (LSTM), could also provide further insights into predicting adult female lice. The Sparse Identification of Non-linear Dynamics (SINDy) model is another promising avenue, potentially offering a novel approach to modeling the dynamics of our data.

REFERENCES

[1] [n. d.]. Introduction to Aquaculture. https://www.fao.org/3/n9317e/n9317e03.htm [n.d.].

[2] Alaska Department of Fish and Game. 2023. What Are Sea Lice? www.adfg.alaska.gov/index.cfm?adfg=wildlifenews.view_article&articles_id=388. Accessed: 2023-12-18.

[3] BarentsWatch. [n. d.]. Salmon Lice. https://www.barentswatch.no/en/articles/Salmon-lice/. ([n. d.]). Accessed: 19-December-2023.

[4] BarentsWatch. 2023. BarentsWatch: Information on Norwegian Coastal and Marine Areas. http://www.barentswatch.no. Accessed: 2023-03-01.

[5] Luke T. Barrett, Kathy Overton, Lars H. Stien, Frode Oppedal, and Tim Dempster. 2020. Effect of cleaner fish on sea lice in Norwegian salmon aquaculture: a national scale data analysis. *International Journal for Parasitology* 50, 10 (2020), 787–796. https://doi.org/10.1016/j.ijpara.2019.12.005 Special issue on 'Fish Parasitology'.

[6] BioMar. [n. d.]. Viral diseases. https://www.biomar.com/en/uk/fish-health/health-challenges/viral-diseases. Accessed: 2023-11-28.

[7] Ian R Bricknell, Sarah J Dalesman, Bríd O'Shea, Campbell C Pert, and A Jennifer Mordue Luntz. 2006. Effect of environmental salinity on sea lice Lepeophtheirus salmonis settlement success. *Diseases of aquatic organisms* 71, 3 (2006), 201–212.

[8] A. Coates et al. 2021. Evolution of salmon lice in response to management strategies: A Review. *Reviews in Aquaculture* 13, 3 (2021), 1397–1422. https://doi.org/10.1111/raq.12528

[9] Cointelegraph. [n. d.]. What is Code Interpreter. https://cointelegraph.com/explained/what-is-openai-code-interpreter-and-how-does-it-work. Accessed: 2023-11-28.

[10] Barry Costa-Pierce. 2021. The Principles and Practices of Ecological Aquaculture and the Ecosystem Approach to Aquaculture. *World Aquaculture* 52 (06 2021), 25–31.

[11] Sussie Dalvin, Lars Are Hamre, Rasmus Skern-Mauritzen, Tone Vågseth, Lars Stien, Frode Oppedal, and Samantha Bui. 2020. The effect of temperature on ability of Lepeophtheirus salmonis to infect and persist on Atlantic salmon. *Journal of Fish Diseases* 43, 12 (2020), 1519–1529.

[12] EY Aquaculture and Fisheries. 2022. The Norwegian Aquaculture Analysis 2022. Report. www.oceanspacemedia.com/files/2023/03/01/The%20Norwegian%20Aquaculture%20Analysis%202022%20(1).pdf Accessed: 2023-11-16.
[13] Per Fjelldal, Samantha Bui, Tom Hansen, Frode Oppedal, Gunnar Bakke, Lea Hellenbrecht, Sofie Knutar, and Abdullah Madhun. 2021. Wild Atlantic salmon enter aquaculture sea-cages: A case study. *Conservation Science and Practice* 3 (02 2021). https://doi.org/10.1111/csp2.369
[14] INTAROS Data Catalogue. 2023. Mooring Data from the Barents Sea Opening – Atlantic Water Inflow. https://catalog-intaros.nersc.no/zh_CN/dataset/mooring-data-from-the-barents-sea-opening-atlantic-water-inflow?fbclid=IwAR290n0AbL90sjIFzscv0HP5hrF--nRSZLVcumLX1S4_nsAKdCRaSl5Qfk8. Accessed: 12 December 2023.
[15] Konrad Jarczyk, Sarvin Torkamanzadeh, and Olufisayo Togun. 2023. Aquaculture Analysis. GitHub repository. https://github.com/jarczykkonrad/Aquaculture
[16] J. Nilsson et al. 2023. Effect of water temperature and exposure duration on detachment rate of salmon lice (Lepeophtheirus salmonis); testing the relevant thermal spectrum used for delousing. *Aquaculture* 562 (2023), 738879. https://doi.org/10.1016/j.aquaculture.2022.738879
[17] Science Norway. [n. d.]. Norwegian fish farmers reprimanded for poor treatment of cleaner fish. https://www.sciencenorway.no/fish-farming-agriculture-and-fisheries/norwegian-fish-farmers-reprimanded-for-poor-treatment-of-cleaner-fish/1800510. ([n. d.]). Accessed: 19-December-2023.
[18] OpenAI. 2023. ChatGPT. https://chat.openai.com. Large language model.
[19] Eva Ostertagová. 2012. Modelling using polynomial regression. *Procedia Engineering* 48 (2012), 500–506.
[20] K. Overton, T. Dempster, F. Oppedal, T. S. Kristiansen, K. Gismervik, and L. H. Stien. 2018. Salmon lice treatments and salmon mortality in Norwegian aquaculture: A Review. *Reviews in Aquaculture* 11, 4 (2018), 1398–1417. https://doi.org/10.1111/raq.12299
[21] Kathy Overton, Tim Dempster, Frode Oppedal, Tore S. Kristiansen, Kristine Gismervik, and Lars H. Stien. 2019. Salmon lice treatments and salmon mortality in Norwegian aquaculture: a review. *Reviews in Aquaculture* 11, 4 (2019), 1398–1417. https://doi.org/10.1111/raq.12299 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/raq.12299
[22] Antonio Parreño-Marchante et al. 2014. Advanced traceability system in Aquaculture Supply Chain. *Journal of Food Engineering* 122 (2014), 99–109. https://doi.org/10.1016/j.jfoodeng.2013.09.007
[23] Reuters. [n. d.]. Norway introduces new salmon farm regulation to combat sea lice. https://www.reuters.com/article/norway-salmon-idUSL5N1GL3JU. ([n. d.]). Accessed: 19-December-2023.
[24] Francisca Samsing, Frode Oppedal, Sussie Dalvin, Ingrid Johnsen, Tone Vågseth, and Tim Dempster. 2016. Salmon lice (Lepeophtheirus salmonis) development times, body size, and reproductive outputs follow universal models of temperature dependence. *Canadian Journal of Fisheries and Aquatic Sciences* 73, 12 (2016), 1841–1851.
[25] Sklearn. 2023. sklearn.preprocessing.MinMaxScaler. http://www.https://scikit-learn.org/. Accessed: 2023-12-01.
[26] Yan-Yan Song and LU Ying. 2015. Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry* 27, 2 (2015), 130.
[27] Alison L Morgan Remi M L Gratacap Sami Nikoskelainen Tim J Bowden, Kim D Thompson. 2007 Jun. Seasonal variation and the immune response: a fish perspective. *Fish Shellfish Immunol* (2007 Jun).
[28] Anna Valberg. [n. d.]. Norway has the strictest aquaculture regulations in the world. Will other countries follow suit? https://partner.sciencenorway.no/aquaculture-fish-farm-fish-farming/norway-has-the-strictest-aquaculture-regulations-in-the-world-will-other-countries-follow-suit/2095202. ([n. d.]). Accessed: 19-December-2023.
[29] Sanford Weisberg. 2005. *Applied linear regression*. Vol. 528. John Wiley & Sons.