CLASSIFYING HORIZONTAL SWIMMING BEHAVIOR OF ATLANTIC SALMON USING DEMING REGRESSION

by

Sayeh Sarkarhosseini

Submitted in partial fulfillment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
August 2022

I dedicate this to my family and friends who have always supported and encouraged me in spite of the challenges along the way.

Table of Contents

List of Figures . . . . . **v**

Abstract . . . . . **viii**

List of Abbreviations Used . . . . . **ix**

Acknowledgements . . . . . **x**

Chapter 1 Introduction **1**

1.1 Contributions 3

1.2 Outline 3

Chapter 2 Related Work And Background **5**

2.1 Machine Vision 5

2.2 Bio-Logging And Acoustic Telemetry 8

2.3 Deming Regression 11

Chapter 3 Data and Methods **13**

3.1 Study Site And Datasets 13

3.2 Data Prepocessing 15

3.3 Overview 21

3.4 Proposed Method 24
3.4.1 Recentering 24
3.4.2 Unwrapping Angles 26
3.4.3 Determining AV and SER 28

3.5 Setting Thresholds And Parameters Of The Model 33

Chapter 4 Experiments and Results **37**

4.1 Classifying Fish Swimming Behaviour 37

4.2 Relations Between Swimming Classes and External Variables 42
4.2.1 Circling Classes VS Dissolved Oxygen 46
4.2.2 Circling Classes VS Water Temperature 50

4.2.3 Circling Classes VS Wind Speed 52
4.2.4 Circling Classes VS Feeding 55
4.2.5 Circling Classes VS Weather Condition 57
4.2.6 Circling Classes VS Water Level 58

Chapter 5 Conclusion and Future Work 60

5.1 Summary 60

5.2 Discussion 60

5.3 Limitations And Future Work 62

Bibliography 64

List of Figures

2.1	Swimming pattern with regard to water current	7
2.2	Average accelerations during the hours of the day for six weeks	10
3.1	Scheme of the study site.	14
3.2	Relationship between HPE and HPEm of sync-tags	17
3.3	HPE vs. HPEm for all the calculated sync-tag positions.	18
3.4	RMSE vs. HPEm for all the sync-tags.	19
3.5	HPE vs. HPEm of the sync-tags after applying RMSE and HPE cut-offs.	20
3.6	HPE Distribution of all the tag positions	21
3.7	Examples of swimming classes	22
3.8	The relationship between polar and Cartesian coordinates.	23
3.9	Examples of Time-Theta plots	23
3.10	An example of theta values without and with recentering x-y positions	25
3.11	An example to illuminate the process of Algorithm 2	28
3.12	Examples of how Algorithm 3 works	30
3.13	An example of visualizing errors for topsoil thickness in cm interpolated using kriging method.	31
3.14	AV and SER.	32
3.15	AV and SER Example.	32
3.16	Comparing $w = 12$ and $w = 20$ for a specific example	35
3.17	Comparing $w = 5$ and $w = 12$ for a specific example.	36
4.1	Comparing angles, unwrapped angles, SER and AV values of slow, fast and no circling swimming behaviors.	38
4.2	Comparing the percentage of every class for the fish population during 17 weeks for seven different combinations of thresholds.	40

4.3	Histogram of angular velocities of circling periods.	40
4.4	Classification visualization during the study period.	41
4.5	An example of fish swimming classes.	41
4.6	Comparing the slow, fast, no circling, and inconclusive periods percentage of all the fish during each week for 5 different thresholds.	42
4.7	Comparing AV and SER values and swimming classes during August and November.	43
4.8	Weekly and daily Comparison of swimming classes during (A),(C) daytimes and (B),(D) nights.	44
4.9	Percentage of each class for the fish population during hours of the day in Atlantic Daylight Time (ADT) (A) before October 18^{th}, (B) after October 18^{th}, (C) during the study period.	45
4.10	Dissolved oxygen VS temperature.	47
4.11	Dissolved Oxygen mg/L.	47
4.12	The percentage of each swimming class for dissolved oxygen below and above than 7 mg/L.	48
4.13	Correlation plots for dissolved oxygen and swimming classes	49
4.14	Water temperature (℃).	50
4.15	Correlation plots for water temperature and swimming classes.	51
4.16	Wind speed (km/h) at Shelburne.	52
4.17	The percentage of each swimming class during wind speed below and above than 28 km/h	53
4.18	The percentage of each swimming class during wind speed intervals.	54
4.19	Correlation plots for wind speed and swimming classes.	55
4.20	Feeding times.	56
4.21	The percentage of each swimming class during feeding and not feeding	56
4.22	Weather conditions.	57

4.23 The percentage of each swimming class during different weather conditions. 58
4.24 Water levels in meters. 59
4.25 The percentage of each swimming class during different water level intervals. 59

Abstract

As aquaculture production of fish is progressively increasing, we need to understand the relationships between their swimming behavior and external variables in order to improve their welfare. Fish in sea cages normally exhibit circular swimming patterns, however, an efficient model to automatically identify circling as a frequent and no circling as an abnormal behavior is lacking. In this study, we used acoustic telemetry data to classify the swimming behavior of Atlantic salmon into three classes: slow circling, fast circling, and no circling. This was achieved by developing an algorithm that uses a locally weighted form of Deming regression. We also developed an interactive visualization tool by representing every fish with a horizontal stacked bar colored according to swimming classes which enables the analysis of individual and population-level swimming behavior. Finally, the impact of several external variables such as natural light, temperature, dissolved oxygen, water levels, weather conditions, and wind speed on each swimming class is analyzed.

List of Abbreviations Used

2DRMS twice the Distance Root Mean Square

AV Angular Velocity

HPEm Measured Horizontal Position Error

HPE Horizontal Position Error

HSV Hue, Saturation, Value

RMSE Root Mean Square Error

SER Standard Error of Regression

VPS VEMCO Positioning System

Acknowledgements

I would like to express my deepest gratitude to my supervisor Dr. Dirk Arnold who has patiently supported me throughout this research with his insightful ideas and suggestions. I wish to thank DeepSense for funding my thesis and providing me with tremendous support. I am also genuinely thankful to my family and friends for their guidance and encouragement in every challenging moment of my life.

Chapter 1

Introduction

Recent years have seen a growing interest in the study of animal movements. Advancements in technology have allowed scientists to collect more data regarding animal movements at sea, on land, and in the air. These data can be used to investigate the impact of internal and external variables on animal movement patterns at various spatial and temporal scales. A common challenge in movement data analysis is identifying frequent or abnormal patterns in spatio-temporal data. To address this, several computational methods have been developed to detect pattern structures and classify distinctions among movement patterns. For instance, Beyan and Fisher [9] used a clustering algorithm involving hierarchy classification and hidden Markov models, to classify fish trajectories into normal and abnormal classes. Furthermore, the classified data can be analyzed further by human analysts to interpret patterns in the movement behaviors using interactive visual analytic tools. Andrienko et al., [4] extracted movement events from trajectories and performed a density-based clustering to find densely concentrated events. The clusters are then visualized using spatio-temporal envelopes to help analysts identify potential traffic jams when multiple cars move at low speeds. Visualization or visual analytics can also be used to evaluate the plausibility of the results by an expert using visual displays.

The movement of salmon in open net pens can be affected by several environmental variables such as light, temperature, dissolved oxygen, water currents, tide, and wind. In response to a sub-optimal environment and stressors, fish may change their swimming behavior and lose their appetite [36]. For instance, when the temperature rises, the metabolic rate of fish increases whereas the oxygen solubility in water decreases, and co-occurrence of other factors such as low tidal current, light, wind, or water flow may intensify the effect of low oxygen levels and lead to detrimental hypoxic conditions [36]. In response to reduced levels of dissolved oxygen, the fish may decrease or increase their swimming speeds [30], [69]. Therefore, an abnormally low

or high swimming speed can be used as an indication of hypoxia [50]. To overcome the effect of hypoxic conditions, several aquaculture systems supplement pure oxygen in the sea cages to maintain dissolved oxygen at levels that increase the growth rate and optimize fish production [72]. Hence, understanding the influence of environmental variables on salmon behavior and movement patterns within sea cages may allow achieving improved welfare outcomes by adjusting the sea-cages [57].

Growing interest in fish health has enabled the development of intelligent methods for monitoring fish behavior and stress assessment. As a noninvasive and automated method, machine vision can be used to analyze fish behavior under stress [59]. Other methods such as acoustic telemetry, underwater sonar, and echo sounders have been used to track and monitor the fish positions despite low water visibility [47]. While individual fish tags do not capture the whole fish population in a sea cage, average individual behaviors are correlated with group behaviors during the short and long term, demonstrated by evidence such as similar diel cyclical movement patterns or avoidance of warm and cold water [57].

A typical characteristic of fish movement in a sea cage is forming a circular swimming pattern during the day [57], [56]. The study of horizontal swimming behavior of the fish in sea cages using acoustic telemetry has been limited to measuring the distance from the center, swimming angle and direction and visualizing the horizontal distribution of the fish [68], [77], [37]. In a study conducted by Narazaki et al., they used biologging data to reconstruct the 3D movements of different tagged marine megafauna (sharks, turtles, penguins, marine mammals, etc.) in the open ocean and identify their circling behaviors [54]. However, they don't provide any information about the no circling behavior or the changes in the circling speed in response to environmental variables. Motivated by identifying circling and no circling behavior of cultured fish in sea cages and exploring the impact of external variables on each behaviour, we propose a method to detect circling and no circling swimming behavior of the fish besides determining the circling speed using angular velocities. In addition, we examine the correlations between several external variables and the circling behavior of the fish.

We start by switching from Cartesian to polar coordinates on account of the circular nature of the polar coordinate system. Then we use the linear pattern of angles during the circling and the scattered pattern of the angles during no circling to separate circling from no circling. In this regard, a Deming regression model is used to overcome the effect of position and time measurement errors. Another challenge in this approach is the periodicity of the angle values that results in Deming regression as a linear model not being immediately applicable to this data type. Consequently, we addressed this issue by applying a locally weighted form of Deming regression to unwrap the angles. We also use Deming regression to calculate the regression error and angular velocity at each point. When a point is surrounded by scattered points the error of the linear model would be high which indicates a no circling behavior. On the other hand, if the points around a given point form a linear shape, the error of the linear model would be low, suggesting a circling behavior. Furthermore, the slope of the line can be used to determine the angular velocity and separate slow and fast circling by selecting appropriate thresholds. Finally, we analyze the relationships between each swimming class and external variables such as light, hours of the day, dissolved oxygen, water temperature, wind speed and direction, water level, weather conditions, and feeding.

1.1 Contributions

The main contributions of this work are proposing an innovative method to classify the horizontal movement of fish in a sea-cage into three classes of slow, fast and no circling using high resolution acoustic telemetry positions and exploring the relationship between the fish swimming classes and several external variables. We also, create a novel interactive visualization tool for the swimming classes of multiple fish to reveal the patterns in the individual and group swimming behaviors.

1.2 Outline

The rest of this thesis is organized as follows. In Chapter 2, we explain the related work and the background needed in this thesis. Next, in Chapter 3 after describing the datasets and data preprocessing methods, we expatiate on the process of the

proposed algorithm. Then, in Chapter 4, we determine suitable thresholds for classifying fish swimming behavior and analyze the relationships between swimming classes and several environmental factors. Finally, in Chapter 5, we discuss the conclusion, limitations, and future work.

Chapter 2

Related Work And Background

In this chapter, first, we go over related work that has been proposed for analyzing fish behavior in Sections 2.1 and 2.2. To monitor and analyze the fish behavior in aquaculture, three common methods have been used by researchers: machine vision, bio-logging technology and acoustic telemetry [1]. Machine vision approaches are discussed in Section 2.1; bio-logging and acoustic telemetry ones are specified in Section 2.2. Then we discuss Deming regression as a computational tool used in our algorithm in Section 2.3.

2.1 Machine Vision

Video monitoring has been extensively used in analyzing fish swimming behavior. In machine vision method, data is stored in the form of optical images and the images are analyzed using image-processing algorithms.

Oppedal et al. [56], used underwater video cameras to monitor the behavior of Atlantic salmon during four periods of winter (15^{th} January to 25^{th} February), early spring (26^{th} February to 7^{th} April), late spring (8^{th} April to 19^{th} May), summer (20^{th} May to 23^{rd} June). They observed the fish behavior randomly for two hours, usually during the daytime and five times during the study period every four hours of 24 hours cycle in addition to at dusk and dawn. The observation scores, were pooled for each period during day and night where the number of observations was between 33 to 8 on days and 20 to 3 on nights. They used these observations to classify the fish swimming behaviors into different classes of horizontal distribution as even (fish occupying most of the available area) or uneven (fish using only parts of the area) and swimming speed based on simultaneous observations between groups over time as fast normal and stationary. Their results show that for all the periods, the horizontal distribution of most observations was uneven during the daytime and the fish was swimming in a circular and polarized (in the same direction) pattern, avoiding the

center and walls of the cage. Moreover, at night the fish were evenly distributed and more dispersed in winter, early spring, and late spring and unevenly distributed in summer, maintaining the circular and polarized school. The swimming speed during all periods in the daytime was normal for most observations while it was significantly more stationary at night in winter and early spring and normal for most observations in late spring and summer.

In a study conducted by Pinkiewicz et al. [62], they presented a computer vision tracking system to automatically detect and analyze changes in the swimming behavior of the fish in sea cages. The fish in the video footage is detected frame by frame by the system using image segmentation techniques and tracked through successive video frames using a Kalman filter. The system stores velocities of valid targets and calculates their average swimming speed and direction (the angle in degrees relative to the x-axis of the sea cage) every 30 seconds. Circular statistics [7] were applied to directions to identify activities within 30 seconds sample intervals. Then they analyzed the effect of feeding and environmental variables such as time of the day and tidal currents on swimming speed and direction of the fish. Their results show that the swimming speed increased and direction slightly changed $(30 - 40^{\circ})$ during the feeding. Moreover, the fish swimming direction changed by around 180° during the days with low tide currents while no clear relationship was found between the swimming speed and tide.

In another paper, Johansson et al. [35], classified the fish swimming behavior based on the water current speed at a commercial salmon farm with exposed conditions. They used two cameras to monitor fish behavior during 48 hours and three swimming patterns were identified which are shown in Figure 2.1.

- A- Circle: Swimming in circular patterns at low current velocities (≈ 20 cms^{-1}).
- B- Mixed: Both circling and remaining on the current with increasing current velocities (≈ 35 cms^{-1}).
- C- On Current: Standing on the currents with increasing current velocities to ≈ 47 cms^{-1}.

Winthereig-Rasmussen et al. [75] reported that the current velocity was the weakest within the middle cage in the farm. Since our study cage is in the middle of the

farm and surrounded by three other cages it is likely that water current speed is low at the site.

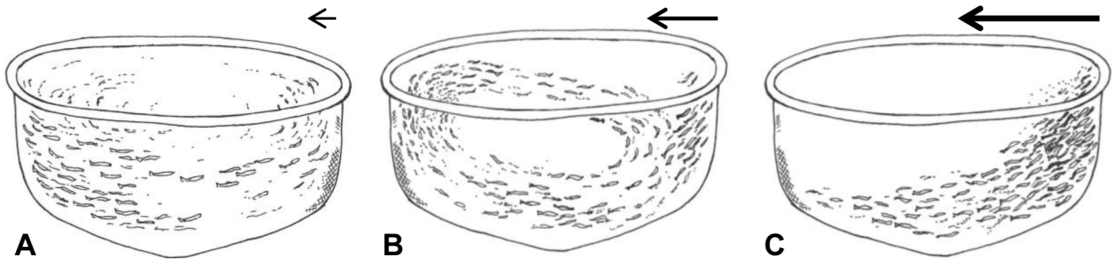


Figure 2.1: Swimming structure with regard to water current. Strength and direction of the water current are shown with arrows. (figure taken from [35])

Wang et al. [73] used a computer vision system and deep learning to propose a method for classifying fish schools' behavior in a recirculating aquaculture system. Their results show that, this method is able to identify five behavioral classes of the fish school, including feeding, hypoxia, hypothermia, frightened, and normal behavior with a high accuracy. This study was done in a closed system with high water quality where barrier technologies ensure that there is no contact between farmed fish and the natural environment. However, in response to external environmental variables, fish might show similar behavior patterns (e.g. fish show a rapid movement during both feeding and frightened behaviors) which results in a high chance of misclassification. Consequently, extracting deep features and accurately classifying the fish behaviors in different environments would be challenging.

Using machine vision, for detecting and analyzing the fish behavior in underwater environments is a challenging task due to illumination, weather change, algae on camera lens, low resolution, complex background, high-density scene, color similarity between fish, etc [78]. Acoustic telemetry method on the other hand, is not affected by light illumination or many other challenges of machine vision and has extensive application possibilities.

2.2 Bio-Logging And Acoustic Telemetry

Bio-logging and acoustic telemetry both monitor fish behavior remotely. However, they differ in the way of collecting the information. In bio-logging, an animal-mounted device is used to collect and store the data which can be downloaded when the logger is retrieved [15]. In a study by Komeyama et al. [41] micro data loggers were used to measure horizontal and vertical speeds and heading angle of one reared Pacific bluefin tuna during one day in a submerged aquaculture net cage. Then they estimated the swimming path of the fish using the dead-reckoning technique and a linear detrending was applied to remove the accumulated error from trajectories. Since most of the reconstructed trajectories drifted nonlinearly, using Fourier analysis Komeyama et al. [42], found that the main source of errors in trajectories was caused by ocean currents. Thus, they applied a high-pass filter to horizontal speed to remove any frequencies lower than the estimated cut-off frequency. Consequently, they determined the 3D trajectories of the circling fish based on the frequency of the circling and current velocity time series, with higher accuracy compared to their previous study.

Narazaki et al. [54], used data loggers to record the depth, temperature, swim speed and 3-axis magnetism, and 3-axis acceleration of tiger sharks, a whale shark, green turtles, king penguins, Antarctic fur seals, and a Cuvier's beaked whale. Then using the dead-reckoning method and swimming speed, acceleration, and magnetism data, they reconstructed the 3D movements of each tagged animal. They defined a circling event as when the animal circled in the same direction more than two times in a row. To identify circling behavior, first, they calculated angular speed (deg/s) using smoothed heading data. Circling began when angular speed exceeded the mean angular speed plus/minus standard deviation multiplied by $1.5 - 2$ for 30 seconds and it ended when the sign of angular speed changed and mean angular speed reached zero for the next 30 seconds. Then they visually examined all of the detected circling behaviors to eliminate false positives. Their results show that foraging, social interactions, capturing prey, and navigation could be responsible for circling behavior.

In acoustic telemetry, a device that is implanted in the animal (transmitter) sends the information signal to a receiver and the position of the transmitter is estimated using mathematical equations [15]. In a study conducted by Anras and Lagardère [5], acoustic telemetry was used to monitor the swimming behavior of nine rainbow trout

for 48 hours in three different culture stocking densities (three fish in each culture density). The fish swimming pattern was analyzed during day and night and three different patterns were identified using turning angles:

1. High turning angles indicating being stationary.

2. Diverse turning angles suggesting chaotic swimming.

3. Average turning angle of 60° showing circling behavior.

The results show circular swimming patterns during the daytime while this pattern was not observed at night.

In another study, Muñoz et al. [53] used acoustic tags equipped with depth pressure and acceleration sensors to record the depth and acceleration values of ten Gilthead seabream for one month (May) in a sea cage. The data was hourly binned and four variables including the average depth, the average acceleration, the interquartile range (IQR) of depth and the 3D space use of each fish were compared during days and nights. Their results show that after adding structural enrichment to the sea cage environment, during the night space use increased and a wider spatial distribution was observed. During the first two weeks, without environmental enrichment, acceleration values were similar in days and nights and a W-shaped pattern was detected with maximum mean values occurring at night and in the afternoon. No significant pattern was found between IQR depth during day and night while the average depth decreased during the night and increased during the day.

In a similar study, Palstra et al. [58] used accelerometer tags on thirty Gilthead seabream in a cubic experimental sea cage which recorded gravity forces and three-dimensional motions of the fish every minute for six weeks (from November to December) and calculated the fish accelerations. A one-way ANOVA test with Tukey post-hoc correlation was used to analyze the variations in activity patterns. They observed more active periods from 6:00 to 14:00 and 18:00 to 00:00 and less active periods from 00:00 to 6:00 and 14:00 to 18:00 as well as the same W-shaped acceleration pattern as [53] during the study period in Figure 2.2. Moreover, there was a positive correlation between acceleration and temperature and the activity increased just before feeding and not during it which is due to anticipating food.

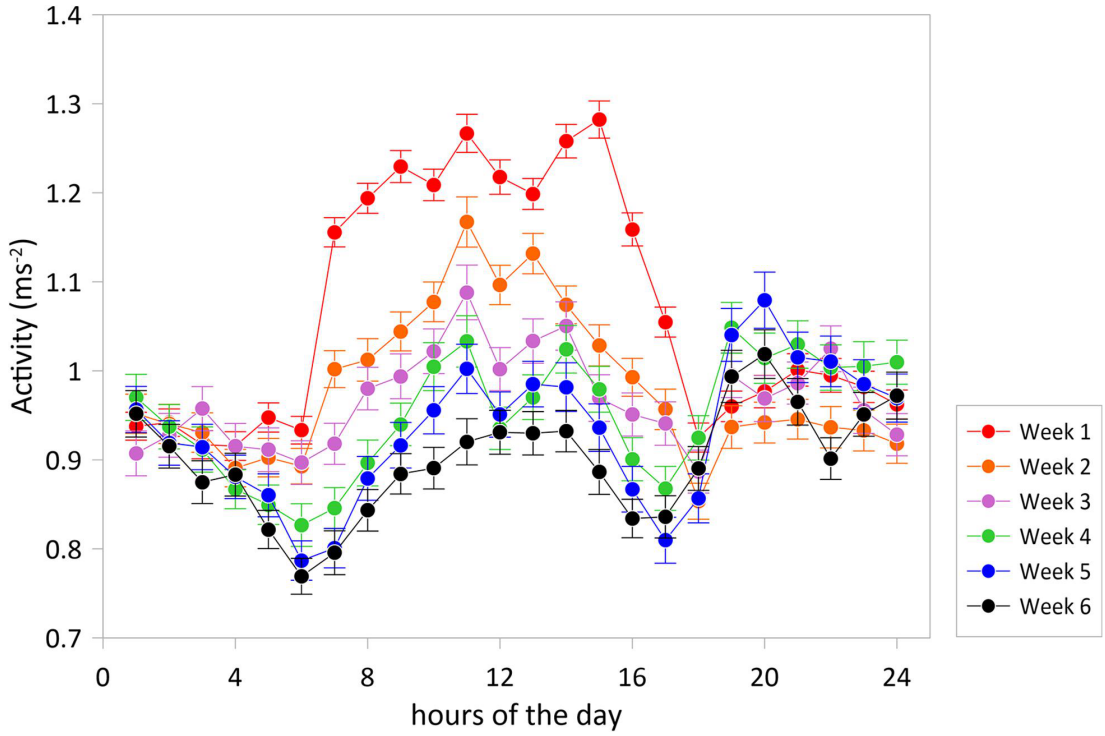


Figure 2.2: Average accelerations during the hours of the day for six weeks. (figure taken from [58])

Using acoustic telemetry Stockwell et al. [68], analyzed the swimming behaviour of ten farmed Atlantic salmon for one month (October) during different environmental conditions such as a storm identified by high wind speeds and fall overturn characterized with a significant difference in the temperature and dissolved oxygen between the first and second half of the month. For every fish position, four variables of distance from the center of the cage, swimming depth, swimming speed, and turning angle were defined and compared before and after the fall overturn and before and during the feeding periods using a Wilcoxon signed-ranks test and a Friedman test was used to compare the variables before, during, after the storm. Their results show that a 24-hour cycle was detected in both swimming depth and speed with higher velocities and greater depths during the day, compared to the night. After the fall overturn, 80% of the fish moved slightly towards the inner part of the net pen and 90% of the fish had a substantial velocity reduction. Also, changes in depth and turning angle were not significant. During the feeding, 80% of the fish moved towards the bottom part of the net pen (depth increased) and the speed of all the fish increased and turning angles decreased indicating straighter paths towards the food. Moreover, the

variation in distances from the center was not significant. After comparing all four variables, before, during, and after the storm, no significant differences were found.

All of the mentioned approaches have aimed to study the swimming behavior of the fish. However, only a few of them attempted to study circling behavior of the marine animals and the impact of environmental variables on circling behavior of the fish. Therefore, in this study, we develop a method to detect the circling and non-circling swimming behavior of the fish in addition to determining the speed of the circling. Then we compare the effect of several external variables on the swimming classes with the existing literature.

2.3 Deming Regression

Essentially, Deming regression [19] is a type of errors-in-variables model which is suitable for when there are random errors in the measurements of both independent and dependent variables. We adapted the equations from [76] with respect to our problem. Assume that each observed value of (t_i, θ_i), $i = 1, 2, ..., n$ has the true value of (t_i^*, θ_i^*) where t and θ represent time and polar angle respectively. The observed values are derived from true values with error measures of (ε_i, η_i) such that we have $t_i = t_i^* + \varepsilon_i$ and $\theta_i = \theta_i^* + \eta_i$. We assume that error values ε_i and η_i are independent and the ratio of their variances $\delta = \frac{\sigma_{\eta\theta}^2}{\sigma_{\varepsilon_t}^2}$ is known. Having the error variance of the dependent variable in the numerator and the error variance of the independent variable in the denominator is the most common way to calculate the δ ratio [24] and sometimes it is calculated inversely [52], [49]. The objective is to find the best fit for $\theta^* = \beta_0 + \beta_1 t^*$ such that the weighted Sum of Squared Residuals (SSR) of the model

$$\text{SSR} == \sum_{i=1}^{n} \left(\frac{\eta_i^2}{\sigma_{\eta\theta}^2} + \frac{\varepsilon_i^2}{\sigma_{\varepsilon_t}^2} \right) = \frac{1}{\sigma_{\eta\theta}^2} \sum_{i=1}^{n} \left((\theta_i - \beta_0 - \beta_1 t_i^*)^2 + \delta \, (t_i - t_i^*)^2 \right)$$

is minimized. To minimize SSR, we compute the partial derivatives and equate them to zero and after solving the equations we have:

$$\left.\begin{array}{l} \dfrac{\partial \mathrm{SSR}}{\partial \beta_1} = 0 \\ \dfrac{\partial \mathrm{SSR}}{\partial \beta_0} = 0 \\ \dfrac{\partial \mathrm{SSR}}{\partial t_i^*} = 0 \end{array}\right\} \begin{array}{l} \hat{\beta}_1 = \dfrac{S_{\theta\theta} - \delta S_{tt} + \sqrt{\left(S_{\theta\theta} - \delta S_{tt}\right)^2 + 4\delta S_{t\theta}^2}}{2S_{t\theta}} \\ \hat{\beta}_0 = \bar{\theta} - \hat{\beta}_1 \bar{t} \\ \hat{t}_i^* = t_i + \dfrac{\hat{\beta}_1}{\hat{\beta}_1^2 + \delta}\left(\theta_i - \hat{\beta}_0 - \hat{\beta}_1 t_i\right) \end{array}$$

Where,

$$S_{tt} = \frac{1}{n}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)^2$$

$$S_{\theta\theta} = \frac{1}{n}\sum_{i=1}^{n}\left(\theta_i - \bar{\theta}\right)^2$$

$$S_{t\theta} = \frac{1}{n}\sum_{i=1}^{n}\left(t_i - \bar{t}\right)\left(\theta_i - \bar{\theta}\right)$$

n is the number of observations and bar notations indicate mean values.

One way of estimating δ is using multiple measurements of the variables; however, there are cases where data is based on single measurements, it can not be used to estimate δ [48]. In some studies with single measurements, if quality-control data is available, δ can be described using the ratio between recorded squared analytical standard deviations [48]. Although, in general, such information is not included in historical data, and sometimes, by default, the simple orthogonal case of $\delta = 1$ is used [52]. Even if δ is not accurately specified, Deming regression is often less biased than the maximum observed bias for ordinary least-squares regression and consequently has more accurate results [48], [76].

Chapter 3

Data and Methods

In this chapter, first, we describe the study site and datasets in Section 3.1. Then, we specify the data preprocessing and cleaning techniques that are used in Section 3.2. Next, we give an overview of the general method structure in Section 3.3. Then we elaborate on the proposed method in Section 3.4. Finally, we discuss the thresholds and the parameters of the model in Section 3.5.

3.1 Study Site And Datasets

The data used in this thesis was shared with us by the authors of [68] from the Department of Oceanography at Dalhousie University. Therefore, we summarize the description of the study site and collected data in this section from that reference. The study site is a salmon farm (Kelly Cove Salmon) located on the north side of McNutts Island, Nova Scotia, Canada. The farm contained 19 active net pens during the study period, as shown in Figure 3.1. Each net-pen had a diameter of approximately 32 meters, a depth of 11 meters, and an average stocking density of 12.09 kgm^{-3} ($\sim 20,000$ individuals). Acoustic tags (Vemco V9P-180, InnovaSea Systems Inc, Bedford, Nova Scotia, Canada) were surgically implanted in fifteen adult Atlantic salmon (73–78 cm length and 3 kg weight) from net-pen 6. However, the data from only ten salmon were usable because two of them died and three of the tags failed to function normally. Fish tags were programmed to transmit the x-y positions every 3 seconds, and z was returned by a pressure sensor in the tags. Four HR2 receivers (InnovaSea Systems Inc.) at 2 m and 9 m depth (eight receivers in total) were deployed on the North, South, West, and East axes outside the net pen. Furthermore, stationary transmitters called sync-tags were deployed at known locations to correct the time drift in receiver clocks and synchronize them. Sync-tags are also used to calibrate the positioning of receivers and measure the effect of positioning errors [66]. The data was collected from the third week of August 2018

until the receivers ran out of battery before the fourth week of December 2018 (17 weeks). In addition to fish positions, temperature and dissolved oxygen were recorded every 5 minutes by four sensors at 2 m and four sensors at 8 m depth in North, South, West, and East, and one sensor at 4 m depth in the center of the cage. Moreover, feeding times were reported, which were completed twice a day.

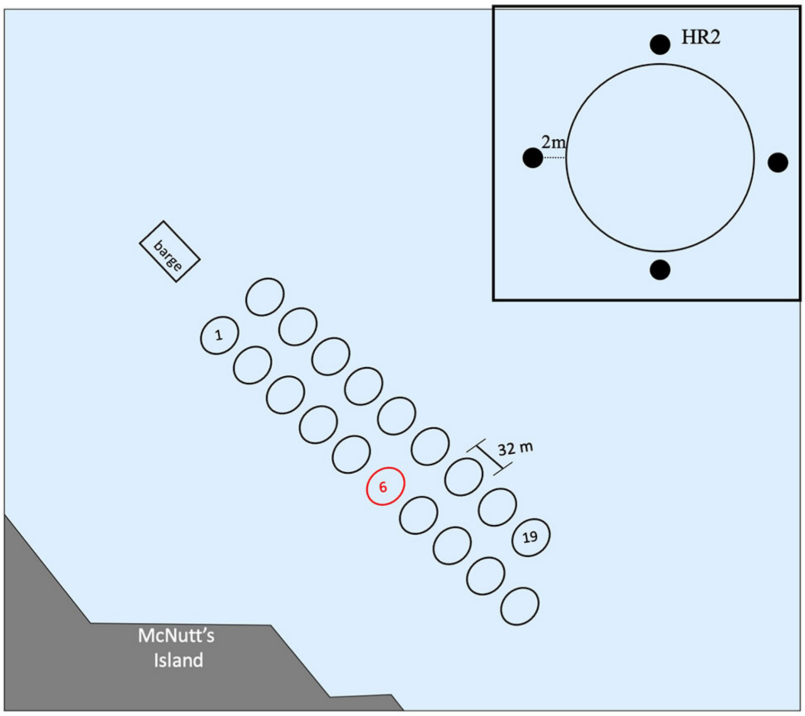


Figure 3.1: Scheme of the study site. The study cage is highlighted with red color and on the top right of the picture the scheme of the acoustic receivers from the top view is shown (figure from [68]).

Furthermore, the water level was acquired every 5 minutes from a station located at Shelburne Harbour from the Department of Fisheries and Oceans Canada. Hourly wind speed and direction data were retrieved from an Environment Canada weather station near Sandy Point, about 7 km from the study site. Additionally, hourly weather conditions and wind data was obtained from BACCARO PT, CA station (about 29 km from the study site) from the Visual Crossing website[1].

[1]https://www.visualcrossing.com/

3.2 Data Prepocessing

In this section, we explain VEMCO Positioning System (VPS) and sources of errors in this system based on a guide by Frank Smith [66] at VEMCO. VPS is an ultrasonic system for tracking aquatic animals under the water, based on the Time Difference of Arrival (TDOA) of a signal at three or more time-synchronized receivers. Since this system assumes ideal propagation in that the speed of a transmitted signal is constant, time is proportional to distance, and the detection time difference is converted into distance difference, referred to as range difference, which indicates the distances of the transmitter to the receivers that detected a signal from it. In this system, errors in the receiver positions, transmitter depth, and range difference can affect the accuracy of the results. Some of the sources of these errors are poor transmitter and receiver geometry, unexpected receiver movements leading to unmeasured receiver positions, errors in detection time of measurement due to the non-linear transmission path, etc.

In this study, two error values of Horizontal Position Error (HPE) and Root Mean Square Error (RMSE) are provided in the data for each fish position, and we explain them in detail in this section. The primary use of RMSE is to find potential positioning errors due to multipath signals which HPE cannot capture. Moreover, the Measured Horizontal Error (HPEm) for each sync-tag was included in the data to evaluate the system's performance. In addition, we have access to HPE and RMSE values of the sync-tags from the data.

HPE is a dimensionless and unitless measure that estimates the relative precision of calculated positions by indicating how sensitive they are to measurement errors. The description of HPE is adapted from [66]. The first step for determining HPE is to calculate basic positions of each tag which are the positions that are calculated using triangles of receivers picked from the set of receivers that detect a signal. VPS then approximates the error sensitivity of each basic position by adding errors in the range differences and depth of transmitters obtained from dataset-specific features that can affect the speed of the sound in water, such as temperature, salinity, and potential depth range, in addition to other elements like the distance between transmitters and receivers. Since the calculated error sensitivity is not calibrated, it should not be interpreted as a distance. Next, VPS determines HPEm values by calculating the horizontal distance (in meters) between calculated sync-tag positions (using the

same method as fish tags) and their known GPS locations. Finally, HPE is obtained by finding a relationship between the HPEm and error sensitivity of sync-tags using a weighted linear regression model. This model is then used to derive HPE from the calculated error sensitivities of the fish tags. HPE should be considered unitless because it is calculated using basic position error sensitivities, which are not calibrated through time.

In many studies that use VPS data, choosing a suitable HPE cut-off value to filter out unreliable positions is challenging. This paragraph presents some of the methods used in previous studies for selecting an appropriate HPE cut-off. Since HPE of sync-tags and fish-tags are calculated using the same method by VPS software, the same level of accuracy is expected for both of them and the data can be filtered past a threshold obtained from the relationship between HPEm and HPE for sync tags (if there is any) [14]. Due to the fact that HPE attempts to predict the relative precision or spread of calculated positions and does not provide an exact error value for a specific position, in many studies, no relationship was found between HPE and HPEm [66]. Consequently, instead of a direct comparison, it would be more beneficial to compare HPE with HPEm statistically using binning technique [66]. In several studies that used VPS data, HPEm values are binned (with bin size of 1) based on the range of HPE values and the relationship between Twice the Distance Root Mean Squared (2DRMS) of errors and the average HPE value of each bin used to construct a linear regression model and find a suitable HPE cut-off [51], [10], [8], [55] and [26]. Instead of 2DRMS, statistic measures such as mean, median, 95th percentile, etc. have also been used [18], [65], [32]. In another study for different HPE thresholds, they looked at the number of calculated positions and mean positioning measurement errors and selected a threshold that minimized the positioning error and loss of data points as much as possible [63]. In another paper, they did not find any correlation between the HPE and HPEm and removed the positions with the top 10% of HPE values [6].

In our dataset, 98.1% of sync-tag and 99.9% of fish-tag positions have HPE values between 0 and 2. We bin the HPEm of sync-tags based on their HPE (with bin size of 0.1) and construct a linear regression model between each bin's average HPE value and 2DRMS. The 2DRMS is calculated using $2 \times \sqrt{\sigma_x^2 + \sigma_y^2}$ formula, where σ_x and σ_y are the standard deviations of the x and y components of the errors within each bin.

As a result, we found a very strong relationship (HPEm $\sim 0.18887 + 2.9448 \times$ HPE; $r^2 = 0.957, P < 0.0001$) which is shown in Figure 3.2. In order to find a prominent threshold, we plot HPEm values against HPE values of all the calculated sync-tag positions as well as each bin's mean, 95th percentile, and 2DRMS in Figure 3.3. When HPE decreases from 0.8 to 0.7, the 95th percentile of HPEm significantly declines (from 5 to 3 meters). Also, when HPE is less than 0.8, all of the statistics only slightly change (less than 1 meter). These features distinguish 0.8 from the other values; therefore, we select it as the cut-off value for HPE. Using the equation obtained from linear regression, fish-tag positions with an HPE less than 0.8 are estimated to have an actual positioning error of smaller than 2.55 meters. By removing the fish-tag positions with HPE greater than 0.8, we preserved 97.56% of the positions and eliminated potentially inaccurate data.

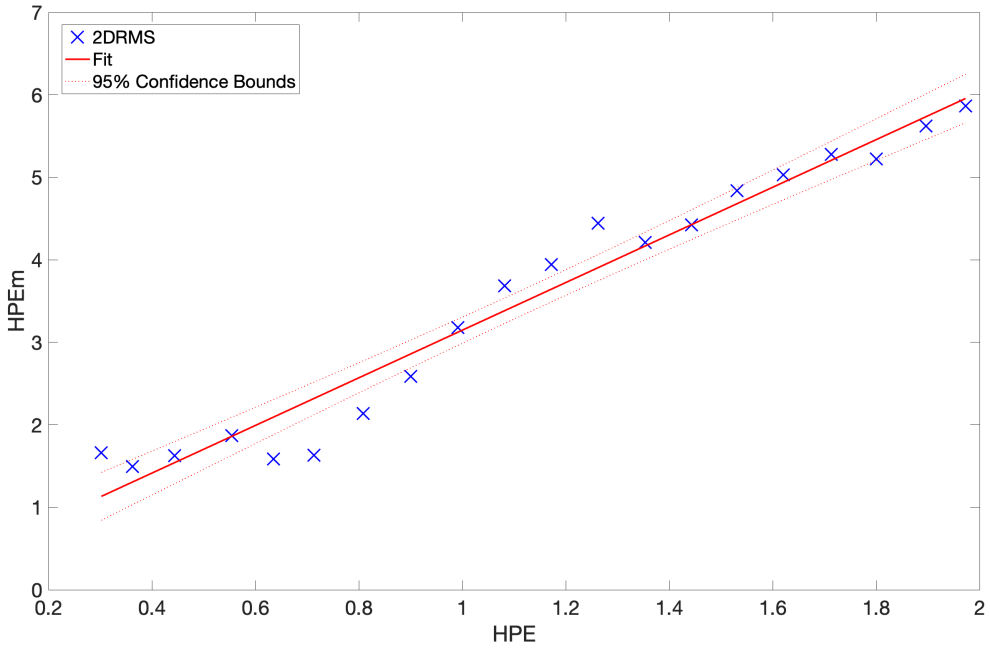


Figure 3.2: Relationship between HPE and HPEm of sync-tags.

In Figure 3.3 we see some points with low HPE that have very high HPEm values due to multipath signals. Most of these errors are addressed by filtering RMSE, which is explained in the next paragraph.

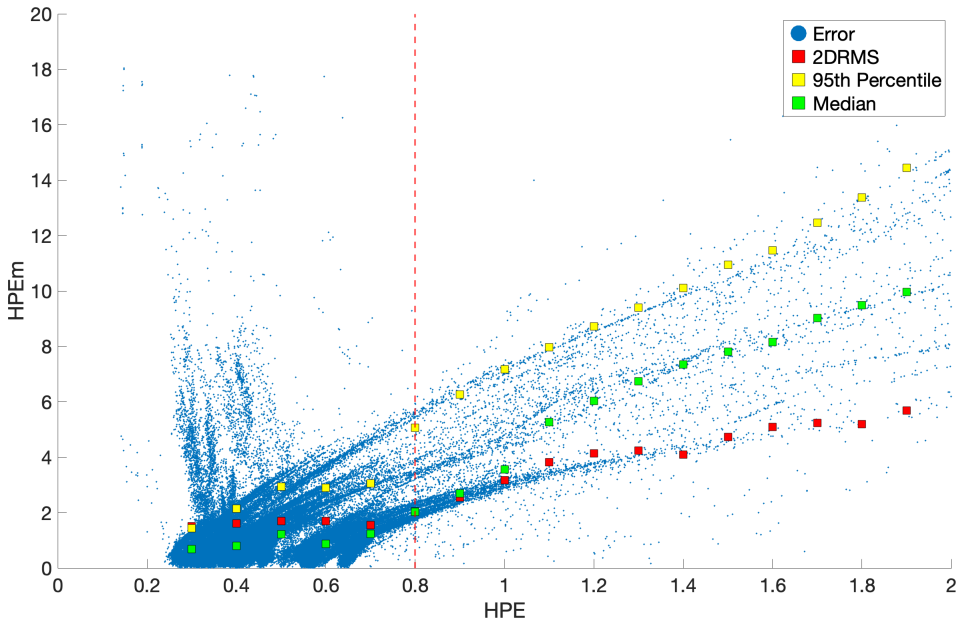


Figure 3.3: HPE vs. HPEm for all the calculated sync-tag positions. Blue dots represent HPE versus HPEm values for estimated sync-tag positions. Red squares show the 2DRMS values. Yellow and green squares represent the median and 95th percentile of HPEm values of each bin.

RMSE of each position is defined as a measure of detection time error, with units in milliseconds [71]. On account of the reflective water surface, ocean floor, rock formations, etc., sometimes transmitters propagate a reflected or diffracted signal instead of the direct and straight signal to the receivers [71]. Consequently, the actual detection time of the transmitter will not match the theoretical detection time, which results in errors when calculating the distance between transmitter and receiver [71]. According to the VPS results description of our data provided by VEMCO, RMSE is calculated as follows. For every detection, the distance between the receiver and the estimated position is measured and divided by the sound speed to estimate the signal's propagation time from the transmitter to the receiver. The propagation time is then subtracted from the detection time at the receiver, and a new estimate of transmission time is calculated. Finally, the difference between the estimated time the signal left the tag and the new estimation of transmission time gives the time error value and the square root of the average squared time error values is equal to RMSE. When only three receivers detect a transmission, RMSE is zero because redundant

measurements are required to calculate RMSE, whereas detection by three receivers is the bare minimum for positioning. In our dataset, 99% of sync-tag and 99.9% of fish-tag positions have RMSE values between 0 and 3. We binned the HPEm of sync-tags based on their RMSE (with bin size of 0.1) and could not find a linear relationship between average RMSE and the 2DRMS or other statistics of HPEm. To determine an appropriate RMSE cut-off, we look at RMSE versus HPEm values of sync-tags in addition to the mean, 95th percentile, and 2DRMS of each bin Figure 3.4. When RMSE decreases from 0.75 to 0.65, the 95th percentile of HPEm considerably drops (from 6.5 to 4.3 meters), and for the smaller values of RMSE, we observe a small change among the statistics (variations less than 1 meter). Thus, we select 0.75 as a cut-off value for RMSE, and it seems to remove most of the points with high HPEm values without losing too many data points. By applying this threshold, we maintain 97.42% of the fish-tag positions and increase the accuracy.

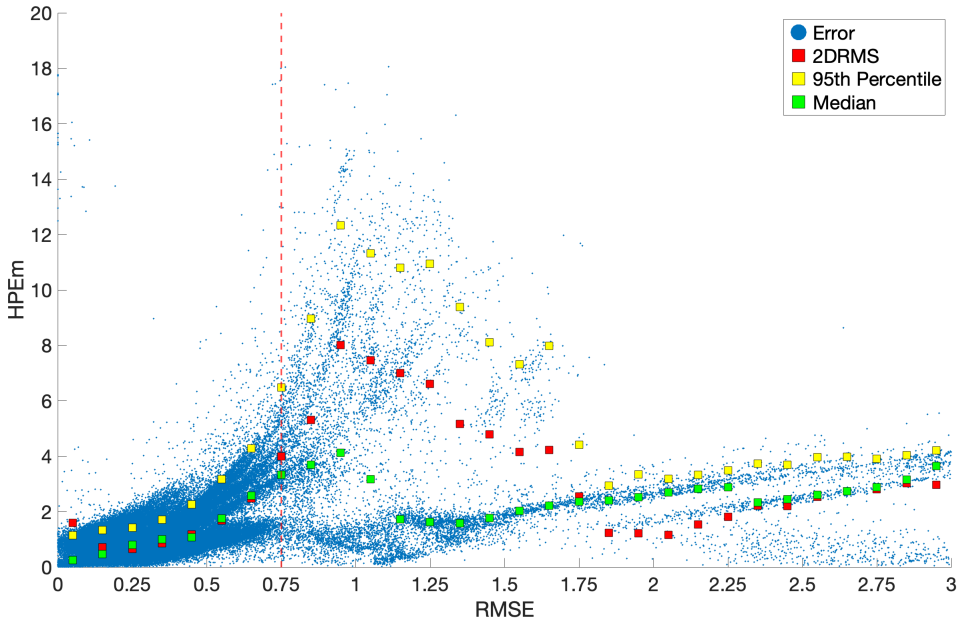


Figure 3.4: RMSE vs. HPEm for all the sync tags. Blue dots represent RMSE versus HPEm values for estimated sync-tag positions. Red squares show the 2DRMS values. Yellow and green squares represent the median and 95th percentile of HPEm values in each bin.

Figure 3.5 shows HPE versus HPEm of the sync-tags after applying RMSE and HPE cut-offs, and, as we can see, the number of points with low-HPE and high HPEm

remarkably reduced by removing high RMSE values. After applying cut-off values of 0.8 to HPE and 0.75 to RMSE values of fish-tag positions, we preserve 95% of data points and eliminate the potentially erroneous 5% compared to [68] that only removed the top 10% of HPE values as outliers.

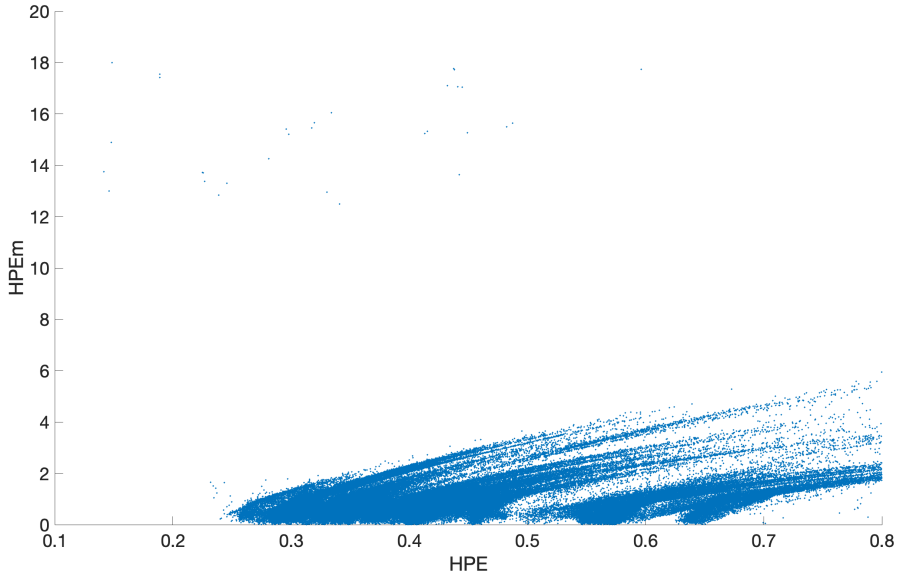


Figure 3.5: HPE vs. HPEm of the sync-tags after applying RMSE and HPE cut-offs.

In Figure 3.6 we look at the positions of all fish tags before and after filtering errors. When a tag is detected in an area with high error sensitivity (high HPE), the calculated position can have a high error. In this case, it will seem to be transmitted from a completely different location compared to where it originated [66] (e.g., red points in Figure 3.6.A that are outside of the cage). HPE is lowest around the center of the cage, and positions that are far from the array of three receivers have high HPE values. We can see a few calculated positions that, despite being inside of a well-formed triangle, have high HPE values. This could be due to the transmitters being detected by only poorly formed triangles due to acoustics or collisions [66]. In addition, after pruning the data as described above, we remove the remaining horizontal positions with an absolute value greater than 21 (199 points that are 0.0044% of data). The maximum absolute value of horizontal position should be $\sim$ 16 (the radius of the cage), but because of the water flow and the flexibility of the cage, we expect it to expand [68].

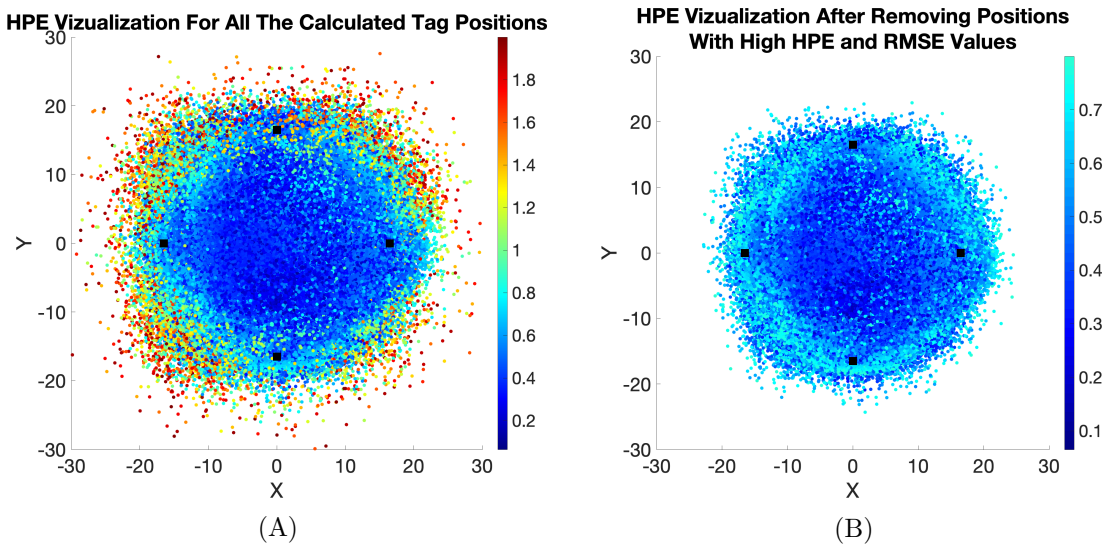


Figure 3.6: HPE Distribution of all the tag positions. (A) All the tag positions before filtering the errors and black squares show the position of the receivers. HPE is lower at the center of the cage and increases as we get closer to the receivers. (B) All the tag positions after filtering the HPE and RMSE.

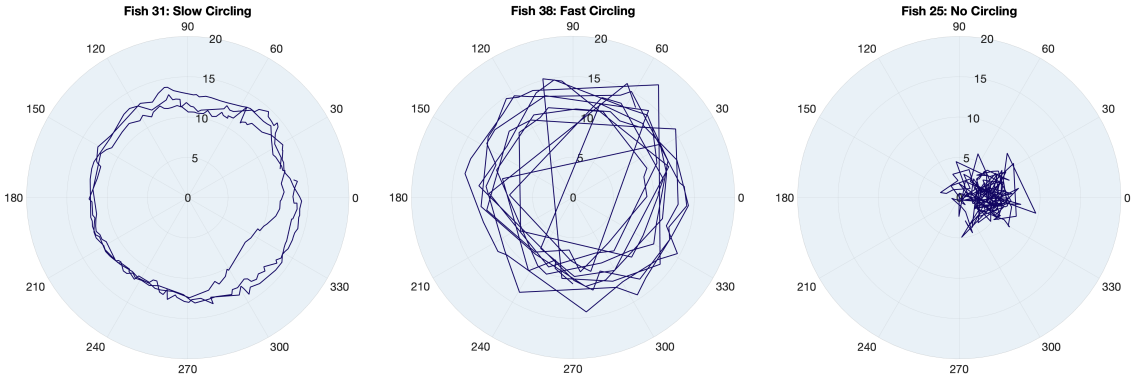
3.3 Overview

This section presents some examples of the identified behaviors in the data and briefly goes over the required steps for detecting these behaviors. Looking at the horizontal swimming behavior of different fish during several periods, we distinguish three distinct behaviors: A. slow circling, B. fast circling, and C. no circling (Figure 3.7).

We define slow circling as when the fish is swimming slowly in circles and completes only a few cycles in a fixed period (Figure 3.7.A), fast circling as when it is quickly swimming in circles and completes several cycles (Figure 3.7.B) and no circling as when the fish only swim back and forth from one side of the cage to another side or in a small area and take lots of turns (Figure 3.7.C).

Since our goal is to identify the circling movement and due to the circular nature of the polar coordinate system, using polar coordinate makes it simpler and more intuitive than Cartesian coordinates. Therefore, we convert each Cartesian coordinate (x, y) to polar coordinate (r, θ) where r is the radial distance from the central point ($r \geq 0$), and θ is the counterclockwise angle from the x-axis (θ is between $-\pi$ and π). We use the $r = \sqrt{x^2 + y^2}$ and $\theta = \text{atan2}\,(y, x)$ formulas to do the conversion and

the relationships between polar and Cartesian coordinates are shown in Figure 3.8. For reasons discussed in Section 3.4.1, we need to recalculate the centers and 2D positions according to the new centers for a fixed rolling window before converting the Cartesian to polar coordinate. Then, we choose to ignore radii since they do not contribute to distinguishing circling from no circling. Classification of fish behaviour is imperfect for fish that swim in a zigzagged pattern for extended period of time near the centre of the cage, but this affects less than 1% of the total time intervals.



(A) (B) (C)

Figure 3.7: Examples of swimming classes. In each plot, we can see the fish swimming behavior during 20 minutes intervals (A) Slow Circling: Fish 31 on November 3rd, 2018, from 1:00 to 1:20 am. (B) Fast Circling: Fish 38 on September 19th, 2018, from 2:20 to 2:40 am. (C) No Circling: Fish 25 on August 26th, 2018, from 12:00 to 12:20 pm.

Next, we need to determine the Angular Velocity (AV), which is defined as the rate of change in the angular position of a rotating object. Looking at the Time-Theta plots of the examples from Figure 3.7, we realize that when the fish is circling, theta values are shaped like an oblique or slanting line during each cycle (Figure 3.9.A and 3.9.B) and when the fish is not circling it is hard to identify any lines (Figure 3.9.C). Thus, a linear model would be beneficial in identifying the lines and the slope of the line gives us the AV.

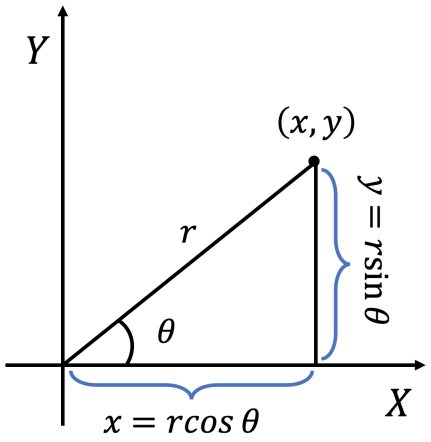


Figure 3.8: The relationship between polar and Cartesian coordinates.

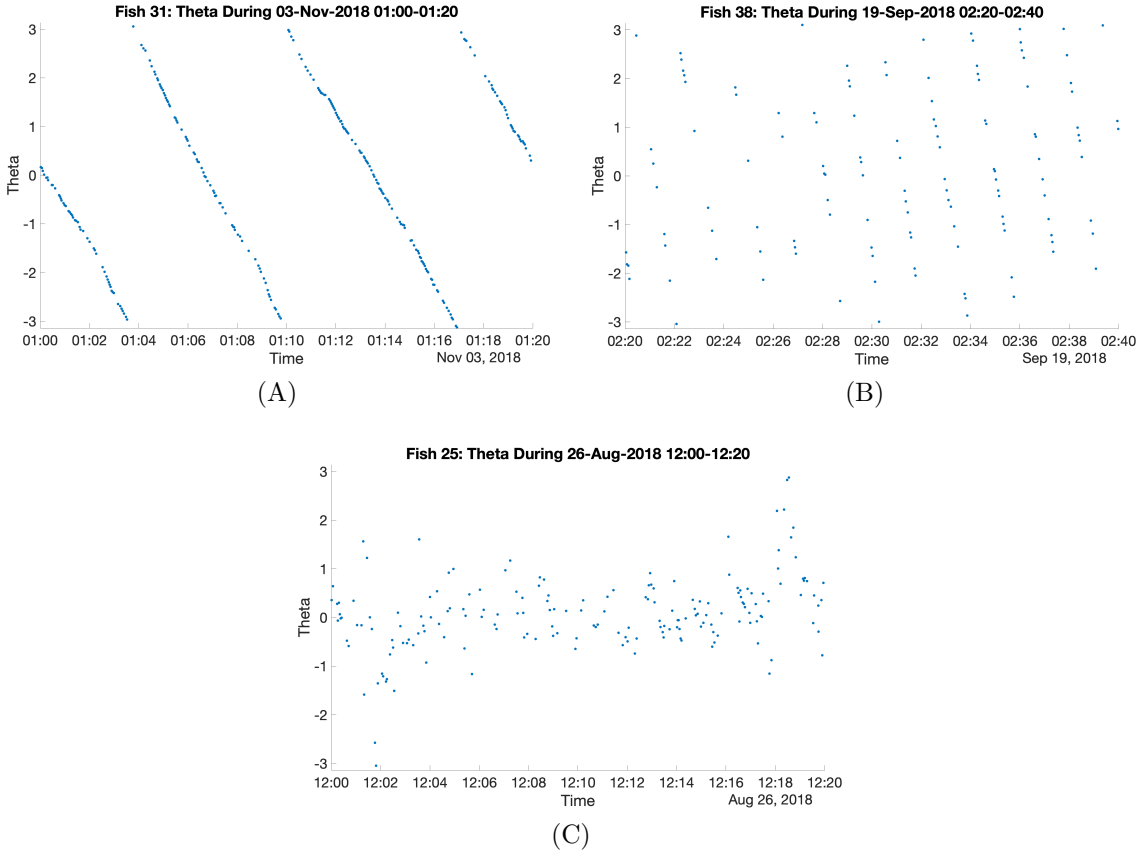


Figure 3.9: Examples of Time-Theta plots. In each plot, theta values during 20 minute intervals are shown. (A) During this slow circling period, we can see four slanting lines. (B) During this fast circling period, we observe twelve slanting lines. (C) lines are hard to identify during a no circling period.

However, this is not a straightforward task because both position and time measurements contain errors. Consequently, theta and time values are also subject to errors. To overcome this problem, we use a Deming regression model. However, we cannot immediately apply the Deming regression to the data as it uses a linear model while our Time-Theta values periodically change between $-\pi$ and π. We address this issue using a locally weighted form of Deming regression which produces a united line that preserves the data characteristics. We name this process "unwrapping the angles". After unwrapping theta values, we use the Standard Error of Regression (SER) to discriminate circling from no circling and the coefficient of the model gives us AV, which is used to distinguish slow circling from fast circling. We discuss the details of the proposed algorithm in Section 3.4.

3.4 Proposed Method

In this section, we delve into the details of the proposed algorithm. For better comprehension, we split the algorithm into three parts. In the first part (Section 3.4.1), we describe how to calculate the theta values and explain why the x-y coordinates should be recentered. Then in the second part (Section 3.4.2), we discuss the proposed algorithm for unwrapping the angles. Finally, in the third part (Section 3.4.3), we demonstrate how to calculate the AV and SER values of each point using the returned unwrapped theta values from the second part.

3.4.1 Recentering

When the center of the circling behavior is significantly different from the center of the cage, theta values are scattered instead of being straight and will not be unwrapped using the linear model. For instance Figure 3.10.A shows example data points where the origin of the coordinate system does not coincide with the center of the circles resulting in irregular theta values in Figure 3.10.B. To overcome this problem, we recalculate the centers and positions based on the new centers using a rolling window in Figure 3.10.C and then theta values are calculated based on the recentered positions in Figure 3.10.D. Algorithm 1 shows this process in depth.

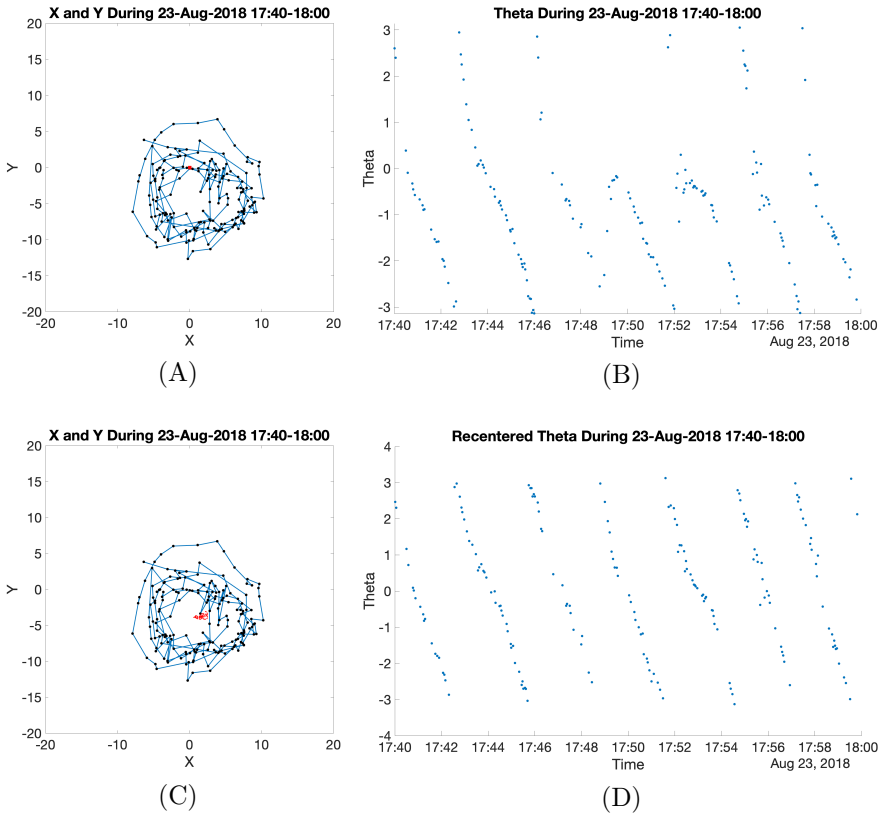


Figure 3.10: An example of theta values without and with recentering x-y positions (A) Sample x-y positions where the center of the cage (red square) is far from the center of the circles. (B) Theta values corresponding to each position before recalculating the centers. (C) x-y positions after recalculating the centers and the red squares represent the positions of the recalculated centers. (D) Theta values after recalculating the centers.

In Algorithm 1, for each fish, given the horizontal positions set $P = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the average of x and y values is calculated for rolling windows of $m = 1200$ seconds centered at each point to find the center of circling behavior in that interval. Then accordingly, x and y values are recalculated based on the new centers. Finally, the new x and y values are switched to polar coordinates and radii are ignored.

Algorithm 1 Recentering Positions and Calculating Angles

Input: $\{x_1, x_2, \ldots, x_n\}$: Trajectory x positions , $\{y_1, y_2, \ldots, y_n\}$: Trajectory y positions, $\{t_1, t_2, \ldots, t_n\}$: Time in seconds, n: Number of data points, m: Rolling window size in seconds to recalculate centers and x-y positions

Output: $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$: Polar angles

1: $\Theta \leftarrow \{\}$
2: **for** $i \leftarrow 1$ **to** n **do**
3: $\quad index = \{j \mid t_i - \frac{m}{2} \leq t_j < t_i + \frac{m}{2}\}$
4: $\quad x' \leftarrow x_i - mean(\{x_j \mid j \in index\})$
5: $\quad y' \leftarrow y_i - mean(\{y_j \mid j \in index\})$
6: $\quad \theta_i \leftarrow \text{atan2}\,(y', x')$
7: **end for**

3.4.2 Unwrapping Angles

In this section, we aim to unwrap the resulting angles from Algorithm 1. First, we build the set $T\Theta = \{(t_1, \theta_1), (t_2, \theta_2), \ldots, (t_n, \theta_n)\}$. Then, in Algorithm 2, starting from the last point in $T\Theta$ we fit a Deming regression line (as described in Section 2.3) with a rolling window of $w = 12$ points to $\{(t_n, \theta_n), (t_{n-1}, \theta_{n-1}), \ldots, (t_{n-w+1}, \theta_{n-w+1})\}$. Then using β_0, β_1 from the Deming regression line, we find an integer $k \in \{0, \pm 1, \pm 2\}$ such that the squared residual (SR) of the point $(t_{n-w}, \theta_{n-w} + 2k\pi)$ is minimized. SR is calculated using the following formula:

$$\text{SR} = (\theta_{n-w} + 2k\pi - \hat{\beta}_0 - \hat{\beta}_1 t^*_{n-w})^2 + \delta(t_{n-w} - t^*_{n-w})^2$$

where,

$$t^*_{n-w} = t_{n-w} + \frac{\hat{\beta}_1}{\hat{\beta}_1^2 + \delta}(\theta_{n-w} + 2k\pi - \hat{\beta}_0 - \hat{\beta}_1 t_{n-w})$$

If $k \neq 0$, we add $2k\pi$ to the preceding angles $\theta_{n-w}, \theta_{n-w-1}, \ldots, \theta_1$. Then the window shifts one point and the process is repeated for the next w points. The algorithm terminates after checking θ_1 and the unwrapped theta values $\left\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n\right\}$ are returned. This process is illustrated in Figure 3.11 using an example.

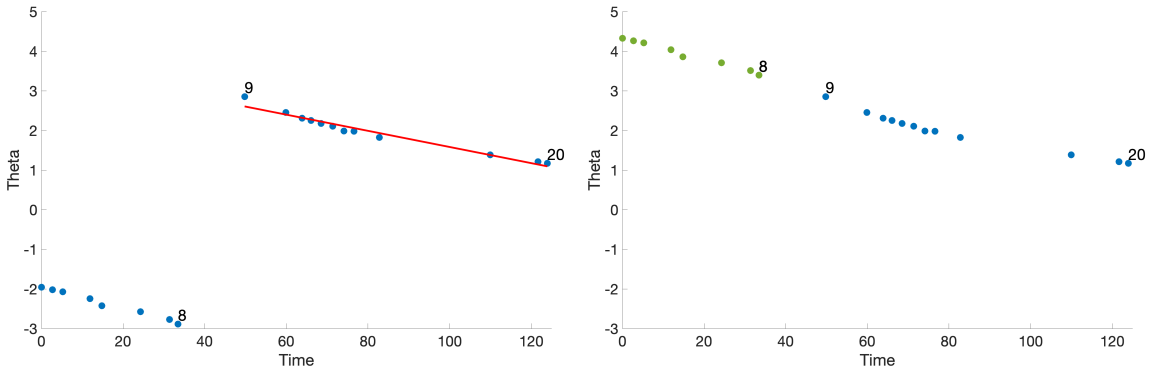
Algorithm 2 Unwrapping Angles

Input: $\{\theta_1, \theta_2, \ldots, \theta_n\}$: Polar angles, $\{t_1, t_2, \ldots, t_n\}$: Time in seconds, n: Number of data points, w: Rolling window size based on number of data points for fitting Deming regression, δ: The ratio between variances of the variables errors

Output: $\hat{\Theta} = \left\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n\right\}$: Unwrapped angles

$\hat{\Theta} \leftarrow \Theta$ // $\hat{\Theta}$ becomes equal to $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$
for $i \leftarrow n - w$ **to** 1 **do**
  $T' \leftarrow \{t_{i+w}, t_{i+w-1}, \ldots, t_{i+1}\}$
  $\Theta' \leftarrow \left\{\hat{\theta}_{i+w}, \hat{\theta}_{i+w-1}, \ldots, \hat{\theta}_{i+1}\right\}$
  $(\hat{\beta}_0, \hat{\beta}_1) \leftarrow Deming\big(T', \Theta', \delta\big)$
  for $j \leftarrow 1$ **to** 5 **do**
    $k = j - 3$
    $t_i^* = t_i + \frac{\hat{\beta}_1}{\hat{\beta}_1^2 + \delta}\left(\hat{\theta}_i + 2k\pi - \hat{\beta}_0 - \hat{\beta}_1 t_i\right)$
    $\mathrm{sr_j} = \left(\hat{\theta}_i + 2k\pi - \hat{\beta}_0 - \hat{\beta}_1 t_i^*\right)^2 + \delta\,(t - t_i^*)^2$
  end for // $\mathrm{SR} = \{\mathrm{sr}_1, \mathrm{sr}_2, \ldots, \mathrm{sr}_5\}$ is calculated
  $\hat{k} = \underset{k \in \{-2,-1,0,1,2\}}{\operatorname{argmin}} \mathrm{SR}(k)$
  if $\hat{k} \neq 0$ **then**
    $\{\hat{\theta}_j \mid j \in i, i-1, \ldots, 1\} \leftarrow \{\hat{\theta}_j + 2\hat{k}\pi \mid j \in i, i-1, \ldots, 1\}$
  end if
end for

In Figure 3.11.A, we show theta values during 125 seconds of a typical slow circling period. The algorithm fits a Deming regression line to the last 12 points (points 20 to 9) to determine the value of point 8 and the preceding points accordingly. Using β_0 and β_1 from the Deming line, we calculate $\mathrm{SR}(k = -2) = 338.25$, $\mathrm{SR}(k = -1) = 146.65$, $\mathrm{SR}(k = 0) = 33.96$, $\mathrm{SR}(k = 1) = 0.21$, $\mathrm{SR}(k = 2) = 45.37$. Since SR is minimum for $k = 1$, $2k\pi = 2(1)\pi = 2\pi$ is added to theta values of point 8 and the rest of the points (Figure 3.11.B). Then we shift one point and use the points 19 to 8 to check the point 7. The algorithm terminates after checking point number 1.



(A) (B)

Figure 3.11: An example to illuminate the process of the Algorithm 2. (A) Blue circles represent theta values during 125 seconds of a slow circling period and the red line represents a Deming regression line fitted to the last 12 points. (B) Green circles represent the resulting theta values after the first iteration.

3.4.3 Determining AV and SER

This section strives to determine AV and SER of every observation. First, we use the resulting unwrapped angles from Algorithm 2 and time values, to construct the set $T\hat{\Theta} = \left\{\left(t_1, \hat{\theta}_1\right), \left(t_2, \hat{\theta}_2\right), \ldots, \left(t_n, \hat{\theta}_n\right)\right\}$. In Algorithm 3, for each point in $T\hat{\Theta}$ we fit a Deming regression line to the points in its past and next $h = 120$ seconds (an interval of 240 seconds centered at each point). The slope of the line gives us AV and the standard error of the line is SER of each point. Also, if the number of data points in 240 seconds is less than or equal to two, AV and SER are considered as missing values. In Figure 3.12 we show how the AV and SER of a point is calculated for the three examples in Figure 3.9 after unwrapping theta values using Algorithm 2.

In Figure 3.12 we fit a Deming regression line (red line) to the points in the window of 240 seconds centered at the selected 50^{th} point (green circle) in each plot. The slope of the line is equal to AV, and the standard error of the line is equal to the SER of point 50. This process is repeated for every point to calculate its corresponding AV and SER. However, for clarity, here, we only show it for one point in each example.

Algorithm 3 Determining AV and SER

Input: $\left\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_n\right\}$: Unwrapped angles, $\{t_1, t_2, \ldots, t_n\}$: Time in seconds, n: number of data points, h: Window size in seconds for calculating AV and SER, δ: The ratio between variances of the variables errors

Output: AV $= \{\mathrm{av}_1, \mathrm{av}_2, \ldots, \mathrm{av}_n\}$: Angular Velocity, SER $= \{\mathrm{ser}_1, \mathrm{ser}_2, \ldots, \mathrm{ser}_n\}$: Standard Error of Regression

AV $\leftarrow \{\}$
SER $\leftarrow \{\}$
for $i \leftarrow 1$ **to** n **do**
  $index = \{j \mid t_i - h \leq t_j < t_i + h\}$
  if $length(index) > 2$ **then**
    $T' \leftarrow \{t_j \mid j \in index\}$
    $\Theta' \leftarrow \{\hat{\theta}_j \mid j \in index\}$
    $(\hat{\beta}_0, \hat{\beta}_1, \hat{t}^*) = Deming\big(T', \Theta', \delta\big)$
    $\mathrm{av_i} \leftarrow \hat{\beta}_1$
    $\sigma_t^2 \leftarrow \dfrac{\sum_{j \in index} (\theta_j - \beta_0 - \beta_1 t_j^*)^2 + \delta(t_i - t_j^*)^2}{2\delta(n-2)}$
    $\sigma_{\hat{\theta}}^2 \leftarrow \delta \sigma_t^2$
    $\mathrm{ser_i} \leftarrow \sqrt{\sigma_t^2 + \sigma_{\hat{\theta}}^2}$
  else
    $\mathrm{av_i} \leftarrow \{\}$
    $\mathrm{ser_i} \leftarrow \{\}$
  end if
end for

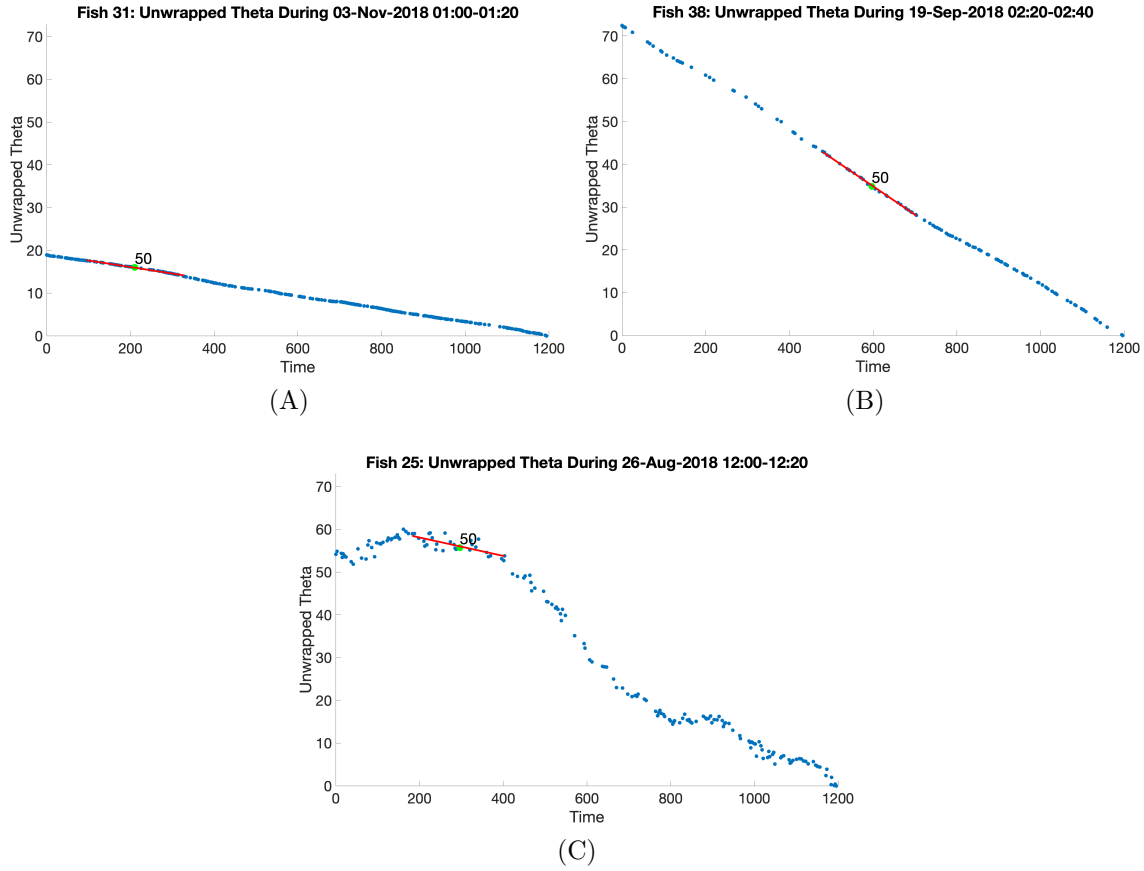


Figure 3.12: Examples of how Algorithm 3 works. In each plot, point number 50 (green circle) is selected to fit a Deming regression line (red line) with a window size of 240 seconds. Unwrapped theta values are shown during the (A) slow circling where $AV_{50} = 0.0153$ rad/s and $SER_{50} = 0.0758$, (B) fast circling where $AV_{50} = 0.0665$ rad/s and $SER_{50} = 0.1510$, (C) no circling where $AV_{50} = 0.0215$ rad/s and $SER_{50} = 1.2447$.

We strive to visualize the behavior of AV and SER over time in a single graph. The problem of visualizing two-dimensional data has been addressed in several context. In an attempt to visualize spatially varying two-dimensional data, bivariate color mapping method has been used to plot two continuous variables on one map, and a two-dimensional color legend is used to display their covariance [16]. Using this technique, Teuling et al., combined temperature and precipitation variables onto a single map [70]. Motivated by the study of Hengl et al. [28], which represented prediction errors by the amount of whiteness, Kaye et al. [39], proposed a white-mixing bivariate technique to visualize temperature and its uncertainty using hue and color saturation or whiteness. Figure 3.13 shows a visualization example of

uncertainty for topsoil thickness in cm, interpolated using kriging method where hue is used to depict predictions, and the spatial prediction error is blended in by adding proportionate amounts of white color to the actual error [29].

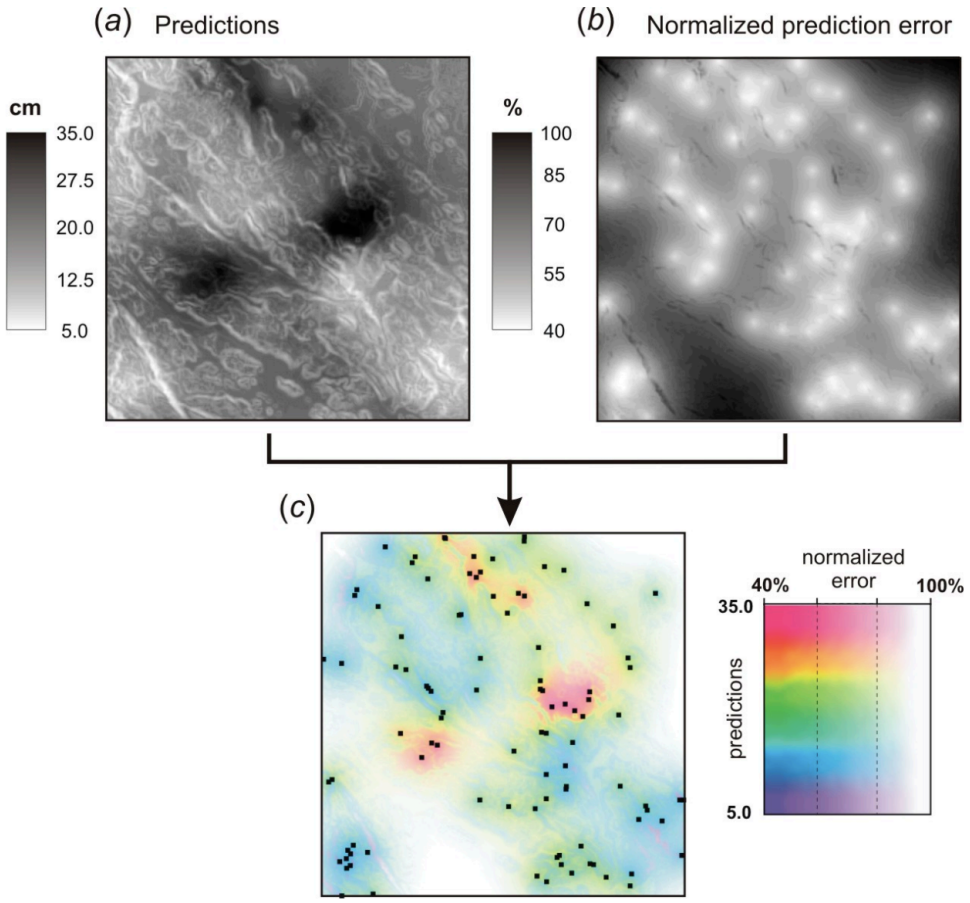


Figure 3.13: An example of visualizing errors for topsoil thickness in cm interpolated using kriging method. (a) predictions, (b) prediction errors and (c) Mix of predictions and errors in two-dimensional legend. (figure taken from [29]).

We adopt the technique used by [29] and visualize the AV and SER values of each fish averaged over 20-minute intervals during the study period in Figure 3.14. We represent each fish with a horizontal stacked bar where every segment's hue encodes the value of AV, and its saturation encodes the value of SER. An HSV (hue, saturation, and brightness value) color model, is used to map the AV and SER values to colors such that every segment's color corresponds to a combination of these variables. In addition, due to the complexity of a 2-dimensional legend, we use separate color bars to denote hue and saturation.

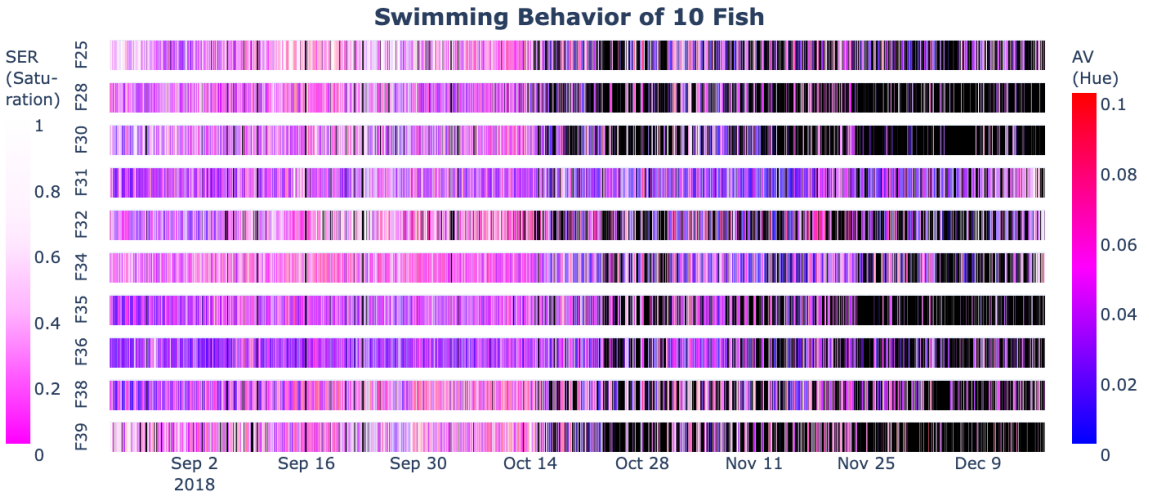


Figure 3.14: AV and SER. Each row represents a fish during the study period. Low saturation indicates high SER values where the white bars show SER values greater than 1. High saturated blue and dark purple colors indicate low AV and high saturated red and pink colors indicate high AV. Black bars indicate insufficient data.

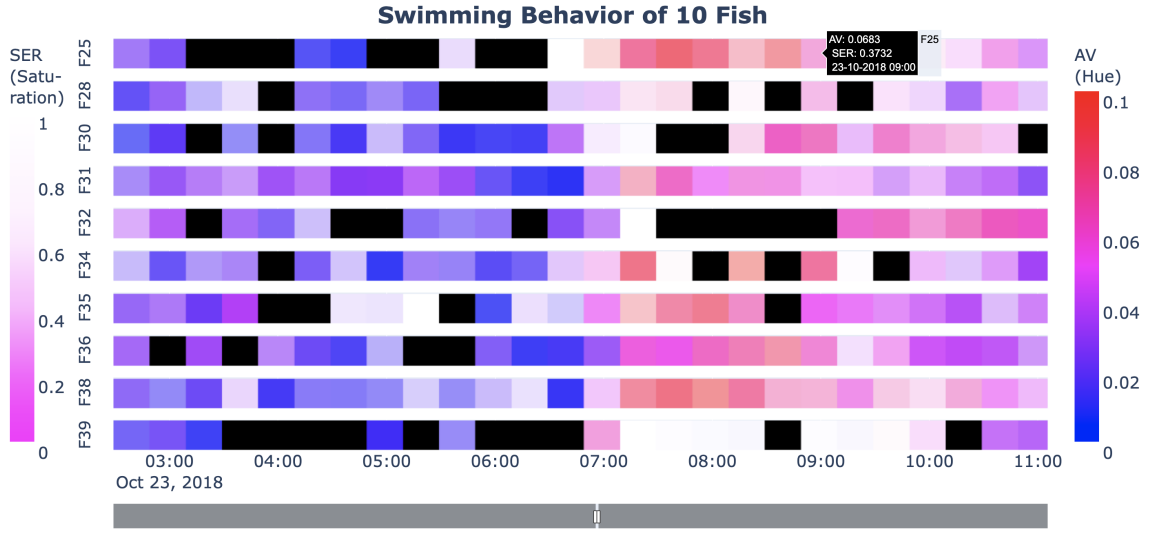


Figure 3.15: AV and SER Example. An example of averaged AV and SER values over 20 minute intervals is shown for every fish during a selected period in October. A range slider on the bottom of the figure is used to look at any interval during the study period.

In Figure 3.15 we provide an example during a smaller period. A slider on the bottom allows us to look at any desired interval during the study period. AV and SER and time and date values of a segment are displayed by moving the mouse cursor

over it. Interactive visualization enables users to comprehend the relationships in the data by actively clicking and filtering; it also provides users with a broad picture and in-depth insights to detect patterns or abnormalities in the data and develop hypotheses [61]. For instance, in Figure 3.15 we use the range slider to look at the AV and SER values of all the fish during 12 hours on October 23$^{\text{rd}}$. A pattern of low AV before 7 am and high AV after it is observed for all fish, plus high SER values during the day are identified for fish 39 by white bars as abnormal behavior.

Theoretically, given the tags' temporal resolution of 3 seconds, we expect 400 observations every 20 minutes. However, in practice, the average number of data points every 20 minutes before October 18$^{\text{th}}$ is 77 before data preprocessing and 73 after data preprocessing and declined after October 18$^{\text{th}}$ to 36 before data preprocessing and 35 after data preprocessing. The notable degradation in the performance of acoustic receivers could be due to factors such as sudden changes in environmental variables (e.g., water temperature, salinity, wind speed, waves height, and current speed), boat traffic, and multi-signal collisions caused by multiple tags or sync-tags present in the same area [23]. Having a small number of data points in an interval can affect the accuracy of the classification. To overcome this problem, if the number of data points in an interval is less than 20, we exclude it from AV and SER calculations, and ensure that the number of data points in each interval is not too small. In Figure 3.14 and 3.15, we use black bars to represent these periods with less than 20 data points and call them inconclusive periods.

3.5 Setting Thresholds And Parameters Of The Model

In this section, we intend to justify the choices made for the value of thresholds and parameters of the proposed method in Section 3.4.

Depending on the speed of circling (if there is any), it takes a fish between roughly 60 and 600 seconds to complete a circle. For recentering in Section 3.4.1, we would like to cover at least 2 circles. Consequently, a window size of $m = 1200$ seconds was selected to ensure that at least 2 circles are used for recalculating the centers and x-y positions.

To determine δ parameter of Deming regression in Section 2.3, first, we need to estimate the time and angle errors. In our study, the time measurement errors are provided in the data in milliseconds (RMSE). Since the measurement of the time data in Deming regression is in seconds, we need to convert RMSE measurements to seconds. For estimating angle errors, first, we calculate HPEm of the fish-tags using the obtained relationship HPEm $\sim 0.18887 + 2.9448 \times$ HPE in Section 3.2. Then we use the radii from the polar coordinates after recentering in Section 3.4.1, to estimate Angle Error $= \frac{\text{HPEm}}{2\pi \times \text{radii}}$. RMSE values range between 0 and 0.00075 seconds and 97% of angle errors range between 0 and 0.09. Consequently, angle errors are approximately a factor of 100 bigger than time errors and their variance ratio would be roughly 10000. Therefore, we consider $\delta = 10000$ in this thesis.

When calculating AV and SER values of each point in Section 3.4.3, we need to ensure that the number of points is adequate to generate good estimates. If there are too many points in an interval, it will be problematic when the fish behavior changes over time. Also, the smaller the value of h, the higher the number of intervals with less than 3 data points and the higher number of missing AV and SER values. We tried intervals with radius of $h = 60, 120, 180$ seconds. The average number of points in the selected interval for all the fish is less than 10 when $h = 60$, more than 20 when $h = 180$, and between 10 and 20 when $h = 120$. Therefore, we use $h = 120$ seconds since it has a reasonable average number of data points for all fish.

The number of measurements per cycle approximately ranges between 3 and 200 data points. For unwrapping the theta in Section 3.4.2, we tried different values of $w = 5, 12, 20$ points as the size of the rolling window for fitting Deming regression. We observed unrealistically high SER values in some samples with too small or too big w even though the fish is circling. In Figure 3.16 we show a specific example where using $w = 20$ results in overfitting of the model and high SER values while the fish is circling with a few U-turns or missing data points. Figure 3.17 represents an example where using $w = 5$ leads to underfitting of the model and high SER values, whereas the fish is circling with approximately 180° change in direction at point 13. Hence, we pick $w = 12$ as it often provides more accurate results.

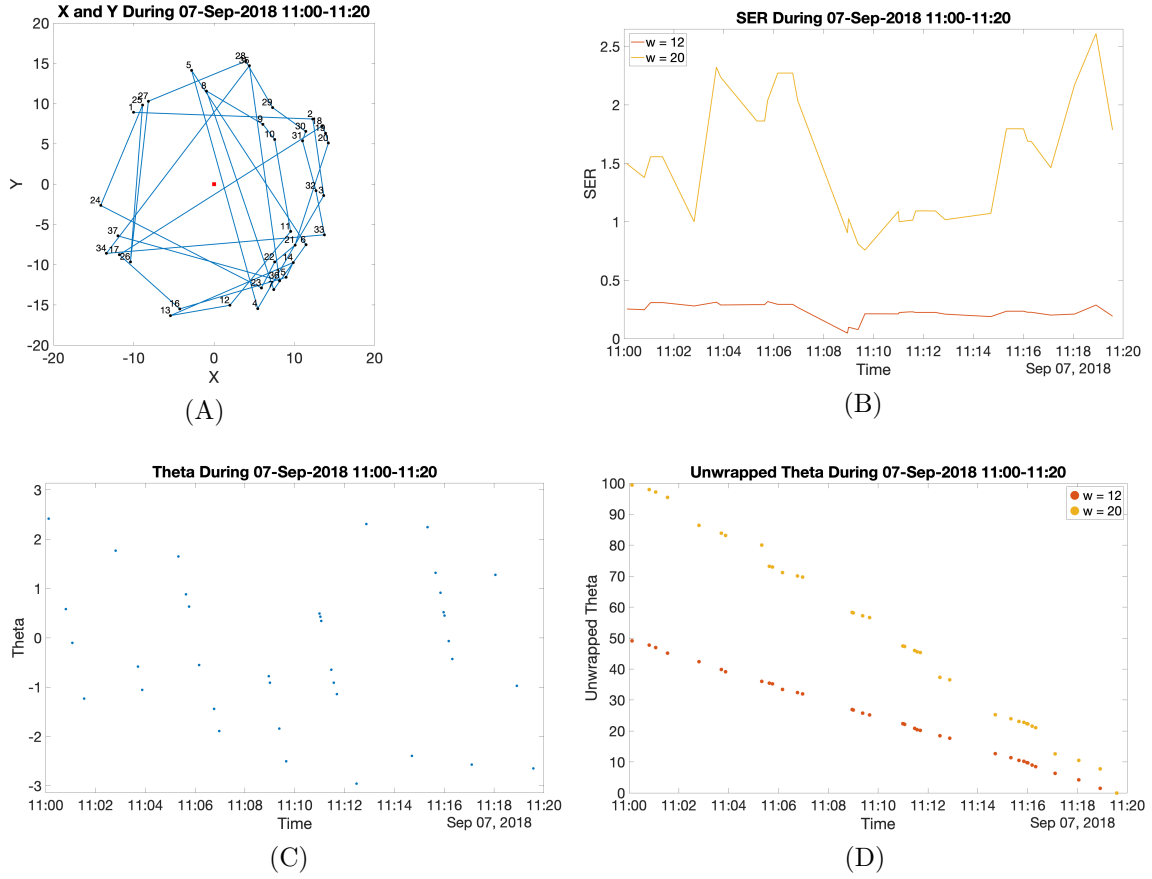


Figure 3.16: Comparing $w = 12$ and $w = 20$ for a specific example. (A) x-y positions of fish 39 on September 7^{th} from 11:00 to 11:20 are shown. The number next to each point indicates the order of the occurrence. (B) SER values for $w = 12$ and $w = 20$ are shown with orange and yellow lines respectively. (C) Theta values. (D) Unwrapped theta values for $w = 12$ and $w = 20$ are shown with orange and yellow circles respectively.

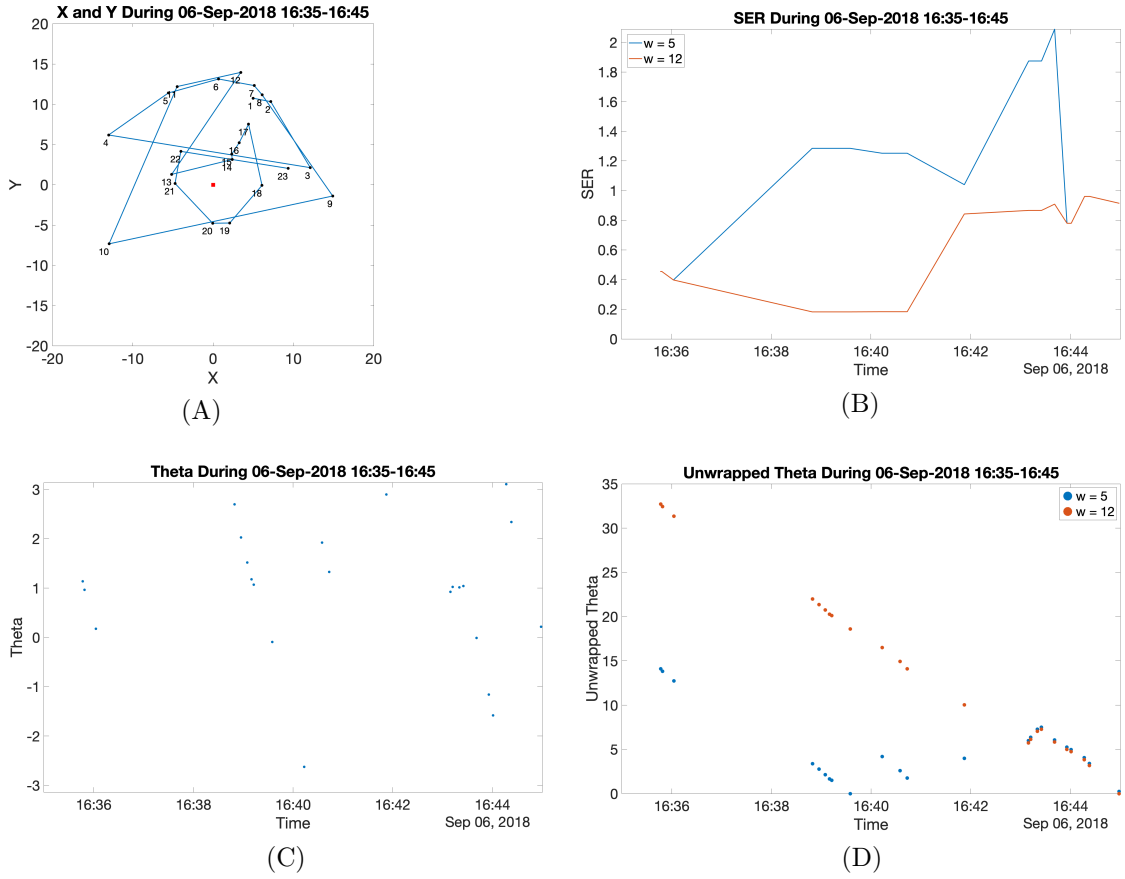


Figure 3.17: Comparing $w = 5$ and $w = 12$ for a specific example. (A) x-y positions of fish 39 on September 6^{th} from 16:35 to 16:45 are shown. The number next to each point indicates the order of the occurrence. (B) SER values for $w = 5$ and $w = 12$ are shown with blue and orange lines respectively. (C) Theta values. (D) Unwrapped theta values for $w = 5$ and $w = 12$ are shown with blue and orange circles respectively.

Chapter 4

Experiments and Results

In this chapter, we use the obtained AV and SER values from Chapter 3 to classify the fish circling behavior. In Section 4.1, we identify suitable thresholds for the classification. Then in Section 4.2, we analyze the correlations or relationships between each swimming class and external variables such as light, time of the day, dissolved oxygen, water temperature, wind speed, feeding, weather conditions, and water level.

4.1 Classifying Fish Swimming Behaviour

In Chapter 3, we have introduced a method that associates AV and SER values with each data point. We now strive to use those values to distinguish between different swimming behaviors. High SER values indicate Deming regression yielding poor fits, and thus AV values being unreliable. This section attempts to determine appropriate thresholds for classifying fish swimming behaviors.

First, in Figure 4.1 we look at the theta, unwrapped theta, SER, and AV values of the three swimming behaviors in Figure 3.7. When the fish is circling slowly (first column in Figure 4.1), we can see about four cycles in 20 minutes by looking at the Time-Theta plot. Moreover, theta values are perfectly unwrapped, which accounts for the low SER values and the alignment of unwrapped theta values has a gradual slope resulting in small AV values. When the fish is circling fast (second column in Figure 4.1), there appear to be about twelve cycles in 20 minutes. Theta values are unwrapped, and SER values are slightly higher than slow circling. The unwrapped theta values are aligned with a steep slope causing relatively high AV values. When the fish is not circling (third column in Figure 4.1), theta values are scattered, and cycles cannot be identified. Consequently, they cannot be unwrapped, which leads to very high SER values and meaningless AV values since there is no circular motion.

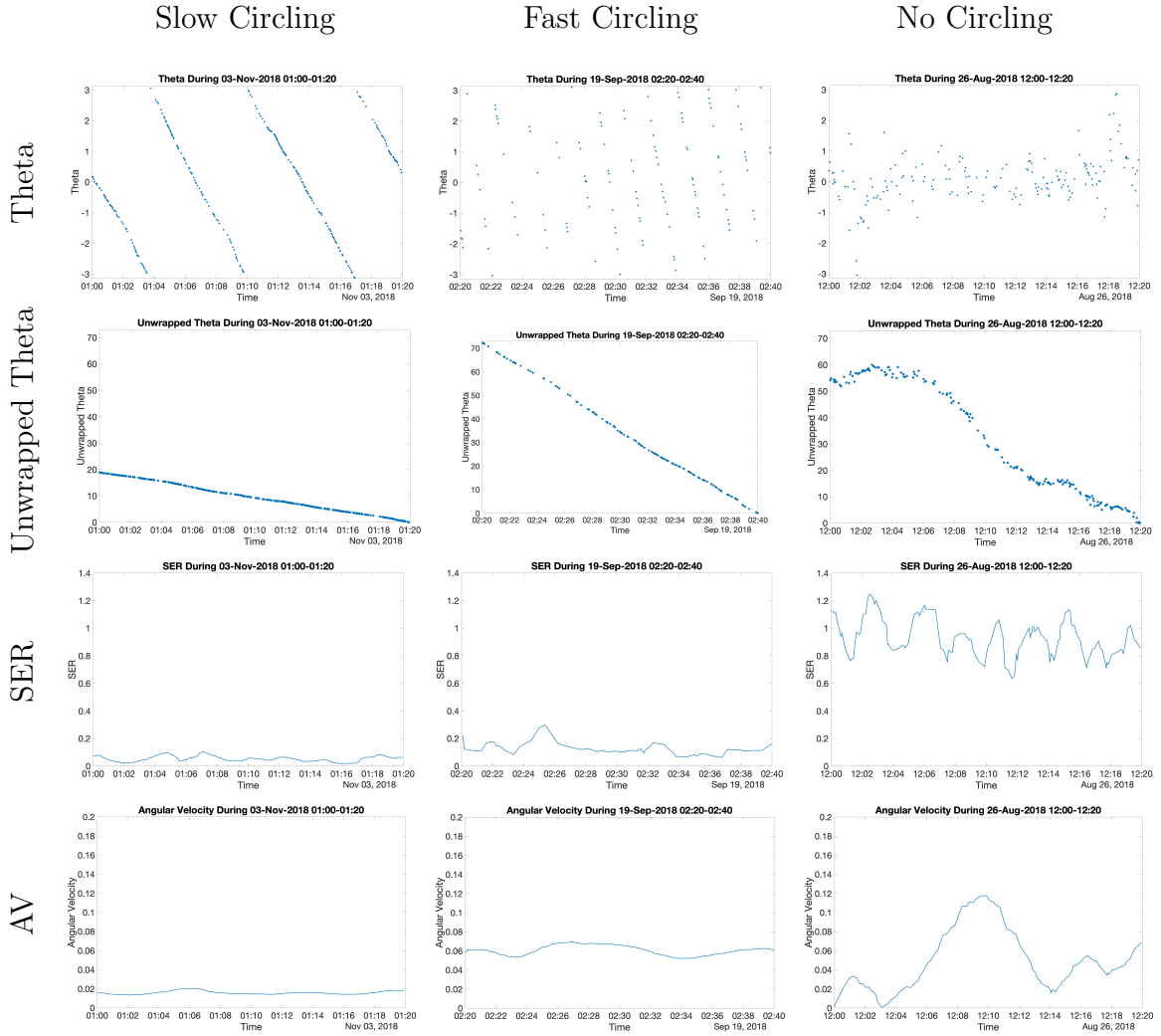


Figure 4.1: Comparing angles, unwrapped angles, SER and AV values of slow, fast and no circling swimming behaviors.

To find thresholds that separate these behaviors, first, we looked at several circling and no circling intervals and found that most of the no circling samples had an average SER above 0.4. When SER is between 0.4 to 0.55, the fish usually does a mix of circling and no circling with a couple of turns, and for average SER greater than 0.55, circling behavior was not observed. Therefore, a suitable threshold for average SER to distinguish circling from no circling should be between 0.4 and 0.55. Next, we define slow circling as when we observe less than six cycles and fast circling as when we observe more than or equal to six cycles during 20 minutes. Looking at the average AV and SER of several circling periods, we observed that the average AV

of slow circling periods was under 0.04, while for fast circling periods, it was above 0.03. Consequently, the appropriate threshold for AV to discriminate between slow and fast circling should be between 0.03 and 0.04.

In the previous paragraph we identified suitable ranges of thresholds to classify the fish swimming behavior. We now aim to show that the observations are not very sensitive to limited changes in the thresholds by visualizing the population behavior. To that end, we divide the study period into 17 weeks and compare the percentage of each class for the fish population for seven combinations of thresholds in Figure 4.2. Since the general patterns which are discussed in detail in Section 4.2 are not affected by changing thresholds, we set 0.035 to AV and 0.45 to SER (the pink line in Figures 4.2.A and 4.2.B and 4.2.C). Moreover, as a further justification for the choice of AV threshold we look at the histogram of angular velocities of circling periods in Figure 4.3. We can see that 0.035 lies within the highest rectangle which corresponds to the modal class of the AV values.

A storm identified with high wind and wind guest speed from October 25^{th} to 28^{th} could be responsible for the spike in inconclusive data in the last week of October (Figure 4.2.D). In November, several days with high percentages of inconclusive periods were observed during or one day before or after the days with high average wind speeds. Additionally, due to higher percentages of inconclusive periods during December, in what follows, we exclude the data in the last three weeks which are shown with red circles in Figure 4.2.D. After setting the AV and SER thresholds, we classify the fish swimming behavior of each fish every 20 minutes during the study period and visualize them in Figure 4.4. In Figure 4.5 we provide an example during a smaller period.

We look at the behavior of every individual fish to see if they follow the same pattern in Figure 4.6. All the fish except fish 36 have lower percentages of slow circling and higher percentages of fast circling before October 18^{th} than after it (Figure 4.6.A and 4.6.B). Fish 25 and 39 have significantly higher no circling percentages in August than the rest of the fish, and except those, the rest of the fish seem to follow the same pattern in Figure 4.6.C. After October 18^{th}, the weekly percentage of inconclusive data significantly increased. The daily percentage of inconclusive data for the population is between $55-70\%$ on October 26^{th} and 27^{th} and November 16^{th}, 26^{th}, 27^{th}, 28^{th} and

29^{th} and we observed high wind speeds during, one day before or after the mentioned days.

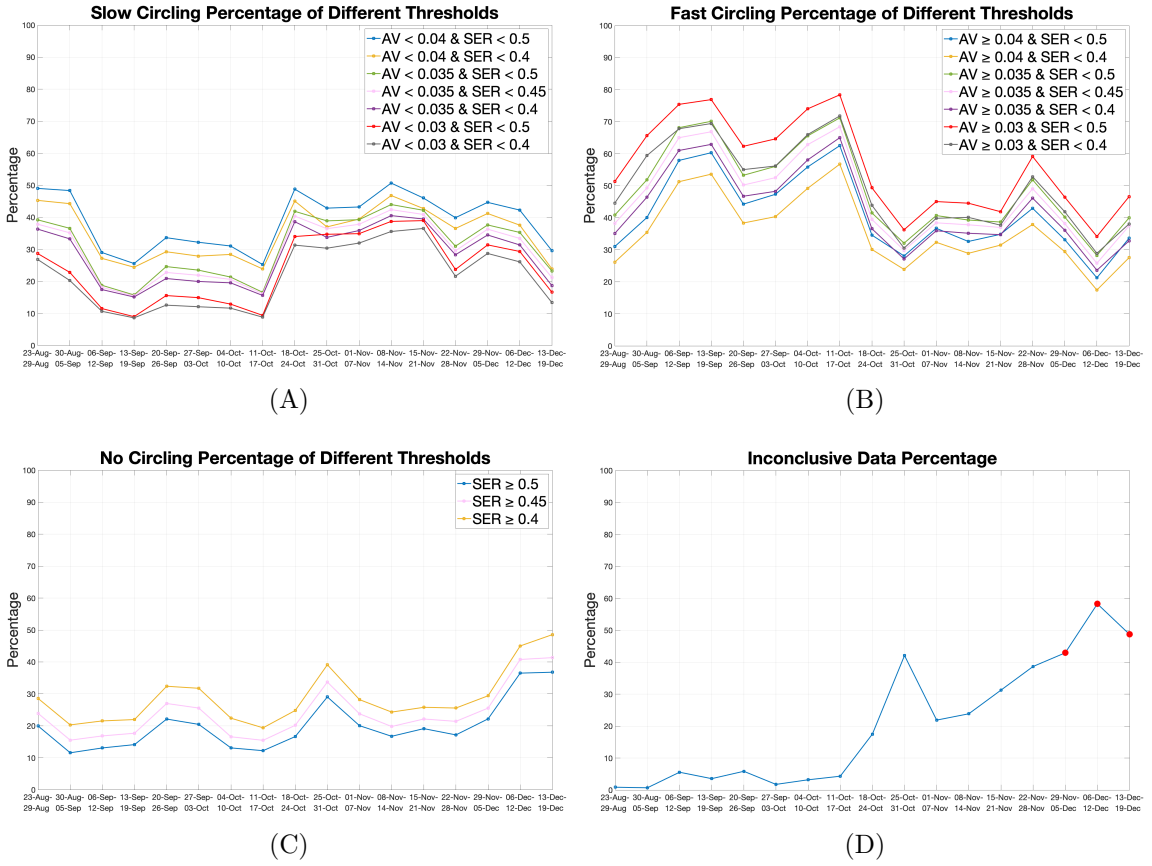


Figure 4.2: Comparing the percentage of every class for the fish population during 17 weeks for seven different combinations of thresholds. (A) Slow circling. (B) Fast circling. (C) No circling. (D) Inconclusive periods. Weeks with high percentages of inconclusive periods are shown with red circles.

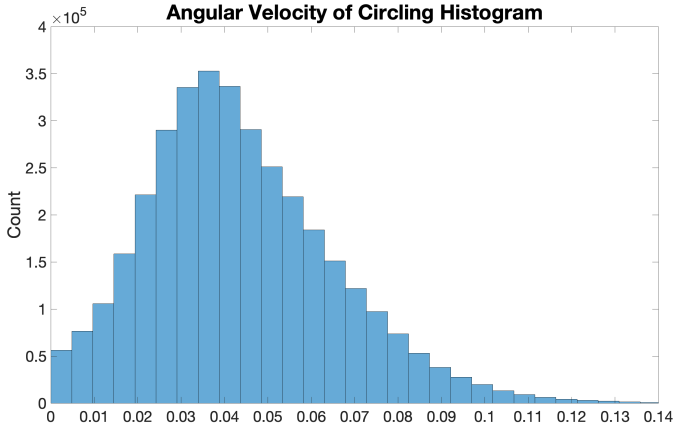


Figure 4.3: Histogram of angular velocities of circling periods.

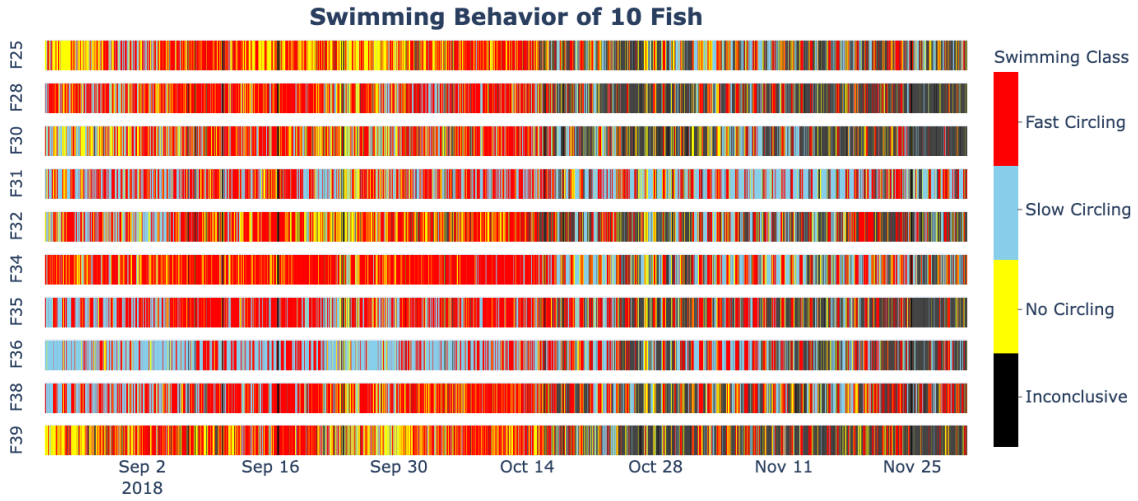


Figure 4.4: Classification visualization during the study period. Fast Circling is marked with red bars where the Average AV ≥ 0.035 & Average SER < 0.45 in a period of 20 minutes. Blue bars represent slow circling where the Average AV < 0.035 & Average SER < 0.45 in a period of 20 minutes. No Circling is represented with yellow bars where the Average SER ≥ 0.45 in a period of 20 minutes. Inconclusive periods are shown with black bars where there are less than 20 data points in a period of 20 minutes.

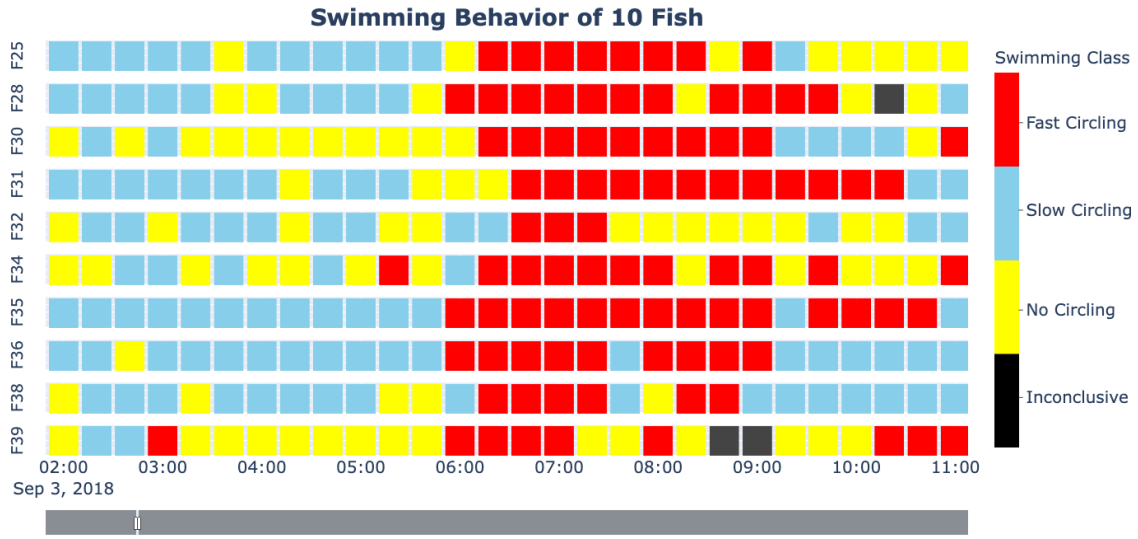


Figure 4.5: An example of fish swimming classes during September 3$^{\text{rd}}$. A range slider on the bottom of the figure is used to look at any interval during the study period.

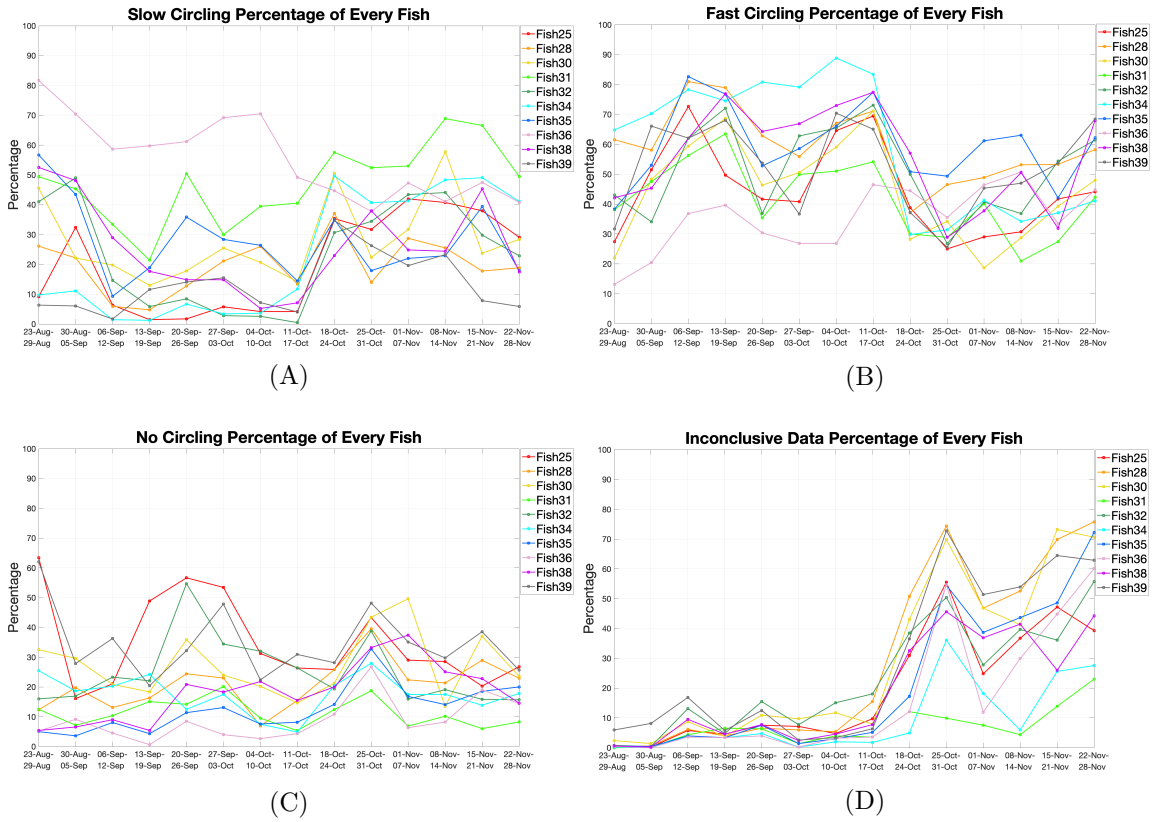


Figure 4.6: Comparing the percentage of every class for the individual fish during 14 weeks. (A) Weekly slow circling percentage of each fish. (B) Weekly fast circling percentage of each fish. (C) Weekly no circling percentage of each fish. (D) Weekly percentage of inconclusive periods for each fish.

4.2 Relations Between Swimming Classes and External Variables

In Section 4.1 we have classified the fish swimming behavior into three classes; slow circling, fast circling and no circling. In this section, we analyze the effect of changes in environmental variables such as wind speed and direction, dissolved oxygen, temperature, water level, weather conditions and feeding on each swimming class.

Many studies pointed to differing behaviours at daytime versus night [57], [12]. To examine this, in Figure 4.7 we look at the fish swimming behavior during days and nights. In Figure 4.7.A, we observe a pattern of low AV during the day and high AV during the night in the last week of August and in Figure 4.7.C we see a pattern of high AV during the day and low AV during the night after mid-October. Figure 4.7.B and

4.7.D show the AV and SER values during a selected day in August and November to provide a more explicit insight into the change in swimming pattern. Looking at the circling classes, we observe an opposite circling behavior during the day and night for most of the fish when comparing a selected day in August (Figure 4.7.E) and November (Figure 4.7.F). Noticing these behavior switches led us to analyze the relationship between fish swimming classes with external variables during days and nights separately.

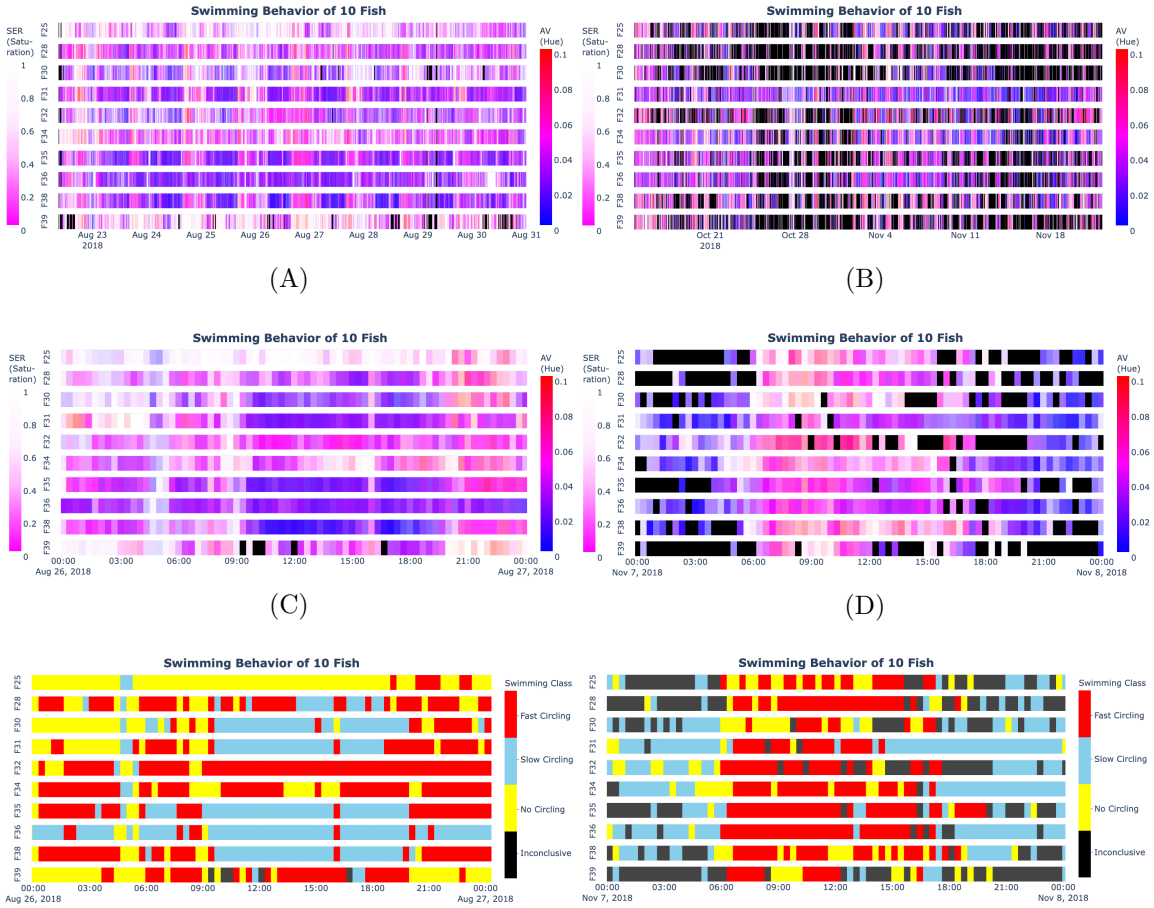


Figure 4.7: Comparing AV and SER values during August and November. AV and SER values are shown (A) during August, (B) from mid-October to November, (C) on August 26$^{\text{th}}$, (D) on November 7$^{\text{th}}$. Swimming classes are presented on (E) August 26$^{\text{th}}$ and (F) November 7$^{\text{th}}$.

Figure 4.8: Weekly and daily comparison of swimming classes during (A),(C) daytimes and (B),(D) nights.

Several studies suggested that salmon swimming speed is usually faster during the daytime than night [43], [56], [20]. Furthermore, the reduction of swimming speed at night is significantly higher during the winter and early spring compared to the late spring and early summer [56], [27]. On each day we define daytime as any time between the sunrise and sunset and define night as any time after the sunset and before the sunrise. As shown in Figure 4.8, after October 18^{th}, slow circling significantly increased during the night and slightly decreased during the daytime while fast circling remarkably decreased during the night and moderately increased during the daytime. Moreover, no circling during the daytime increased especially during the week with the storm. Daily percentage of swimming classes in Figure 4.8.C and 4.8.D, show the same pattern as the weekly plots. Before October 18^{th}, out of 57 days, for 37 days (65% of days), the percentages of fast circling during the night are higher than the daytime and for 41 days (72% of days) the percentages of slow

circling during the nights are lower than the daytimes. In summary, after October 18th circling speeds decrease at night, but not during the day. Furthermore, the overall percentage of circling does not change much.

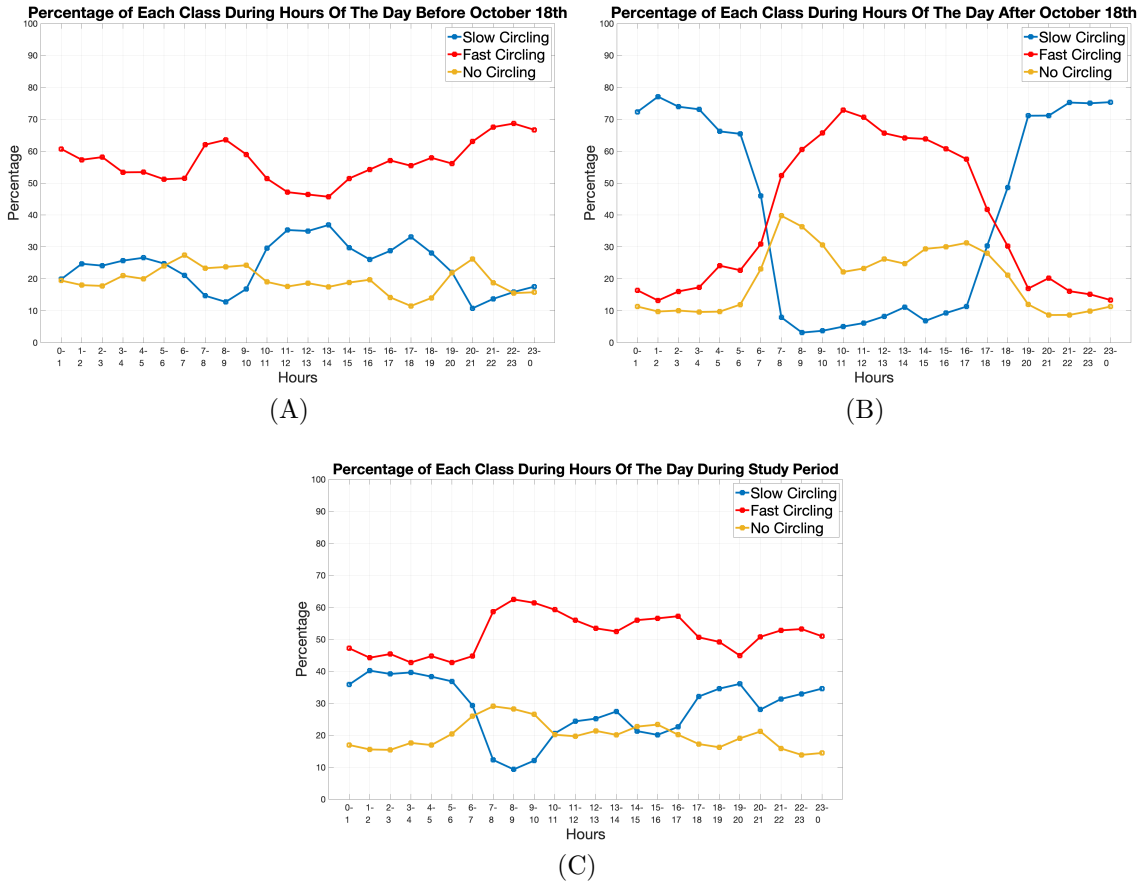


Figure 4.9: Percentage of each class for the fish population during hours of the day in Atlantic Daylight Time (ADT) (A) before October 18th, (B) after October 18th, (C) during the study period.

We look at the hourly percentage of each class during different periods for the fish population in Figure 4.9 (times are in Atlantic Daylight Time (ADT)). Fast circling in Figure 4.9.C has a somewhat similar W-shaped pattern as [58] and [53] in Figure 2.2. During the study period, the percentage of fast circling is higher from 7:00 to 17:00 and slow circling percentages are higher from 00:00 to 07:00 and 17:00 to 00:00 (Figure 4.9.C). The swimming behavior before October 18th (Figure 4.9.A) and after October 18th (Figure 4.9.B) are noticeably different. After October 18th, slow circling from 00:00 to 7:00 and 18:00 to 00:00 significantly increased and fast circling

during these hours considerably decreased. 51% of feeding times were between 8:00 to 11:00 and 40% of them were between 14:00 to 18:00 which could be responsible for the increase in fast circling percentages during these hours in Figure 4.9.C.

4.2.1 Circling Classes VS Dissolved Oxygen

Since the dissolved oxygen and temperature data from the different locations in the cage follow similar patterns, we chose the one from the cage's center at 4 meters depth (red line in Figure 4.10) for further analysis. In Figure 4.11 we show the average dissolved oxygen every 20 minutes during the study period. Dissolved oxygen notably increased from an average of 6.21 mg/L before October 18^{th} to 8.06 mg/L after October 18^{th}. The oxygen levels in the sea-cages are influenced by several environmental factors such as water current, wind speed and light [36]. Hypoxia occurs when fish have to make a physiological adjustment to maintain oxygen levels in their tissues, due to the absence of sufficient dissolved oxygen [12]. For Atlantic salmon, hypoxia considered as when the oxygen level is below 6 mg/L [17], [12]. Pedersen [60] showed that growth rates of rainbow trout, Salmo gairdneri, decreased when dissolved oxygen was less than 7 mg/L. In response to hypoxia, there is a trade-off between reducing swimming speed as a result of reduced metabolic scope to decrease oxygen requirements and increasing swimming speed to elevate the chance of finding a more favorable environment represented by more oxygenated water and lower temperature, which declines the metabolic activity and the lethal oxygen level of the fish [22], [46], [21]. Consequently, contrasting responses to hypoxia have been found in different studies [13], [22], [64], [25]. Kolarevic et al., [40] reported that the hypoxic conditions in the recirculating aquaculture systems resulted in a significant decrease in the total activity of Atlantic salmon measured by acoustic acceleration transmitter tags; in addition to a sharp increase in swimming activity up to four hours after modifying the oxygen levels in the tanks from optimal to hypoxic. In this study, we compare the percentage of each swimming class for dissolved oxygen below and above 7 mg/L in Figure 4.12. Pearson's chi-squared test is used to determine whether there is a statistically significant relationship between slow and fast circling when dissolved oxygen is less and more than 7 mg/L ($\chi^2 = 7.1556, p = 0.0075, df = 1$). When the dissolved oxygen is lower than 7 mg/L, the percentage of fast circling is

significantly higher and the percentage of slow circling is notably lower than when the dissolved oxygen is greater than 7 mg/L. This suggests that when the oxygen levels are low, the fish spends more time swimming fast which could be an indication of increased activity.

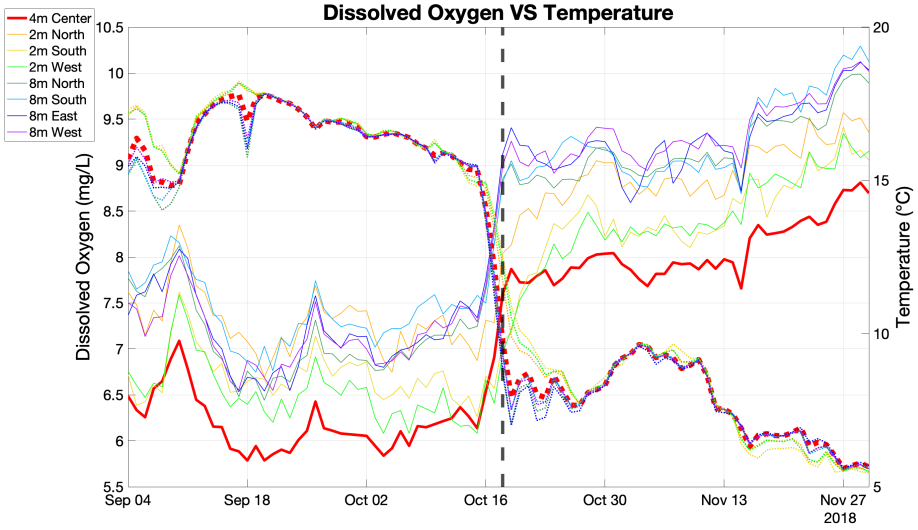


Figure 4.10: Dissolved oxygen VS temperature. Daily average dissolved oxygen (mg/L) and temperature (°C) are shown at different locations of the cage with solid lines and dotted lines respectively. Vertical black dashed line separates before and after October 18th (adapted from [68]).

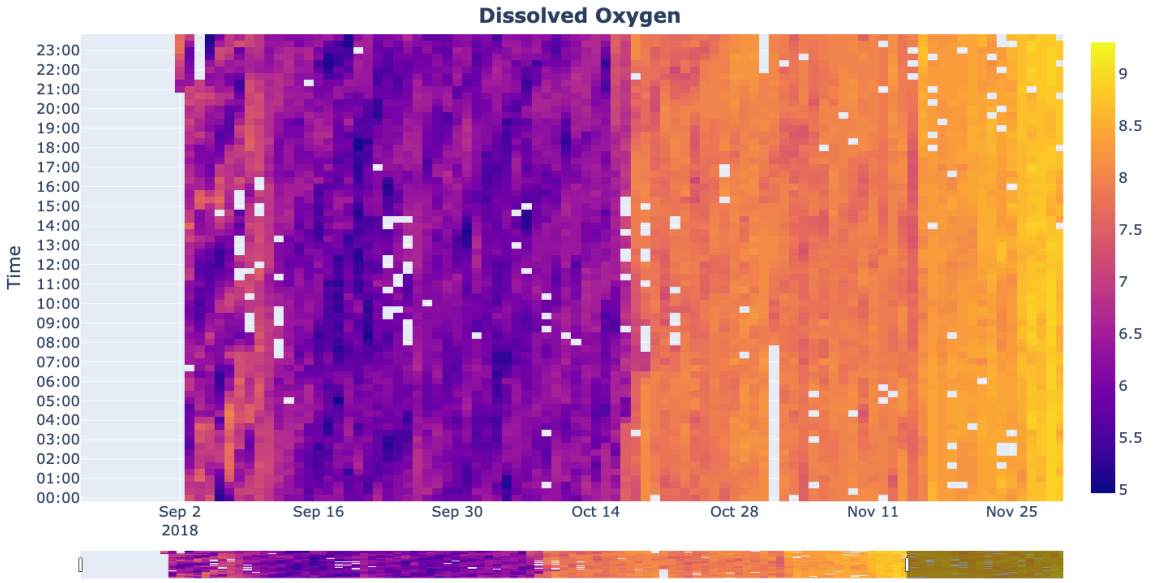


Figure 4.11: Dissolved Oxygen mg/L.

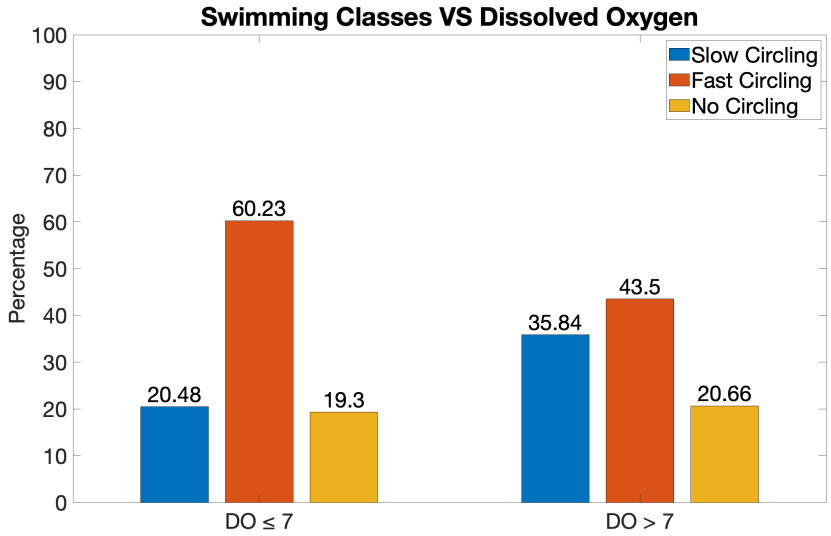


Figure 4.12: The percentage of each swimming class for dissolved oxygen below and above than 7 mg/L.

Furthermore, we examine correlations between swimming classes and dissolved oxygen during days and nights in Figure 4.13. After performing Pearson correlation at a 0.05 confidence interval on the percentage of each swimming class and average dissolved oxygen during daytimes and nights, an opposite correlation direction is observed during daytime and night. A strong negative correlation between the average dissolved oxygen and slow circling during the daytime and a very strong positive correlation during the night is found; whereas between fast circling and dissolved oxygen, there is a weak positive correlation during the daytime and a very strong negative correlation during the night. There is also a moderate positive correlation between no circling and average dissolved oxygen during the daytime, and a week negative correlation between them during the night. The negative correlation between slow circling and dissolved oxygen during the day indicates that slow circling decreases as dissolved oxygen increases during the day and the positive correlation between slow circling and dissolved oxygen during the night suggests that slow circling increases as dissolved oxygen increases at night. In conclusion, the relationship between swimming classes and dissolved oxygen is different during the days and nights.

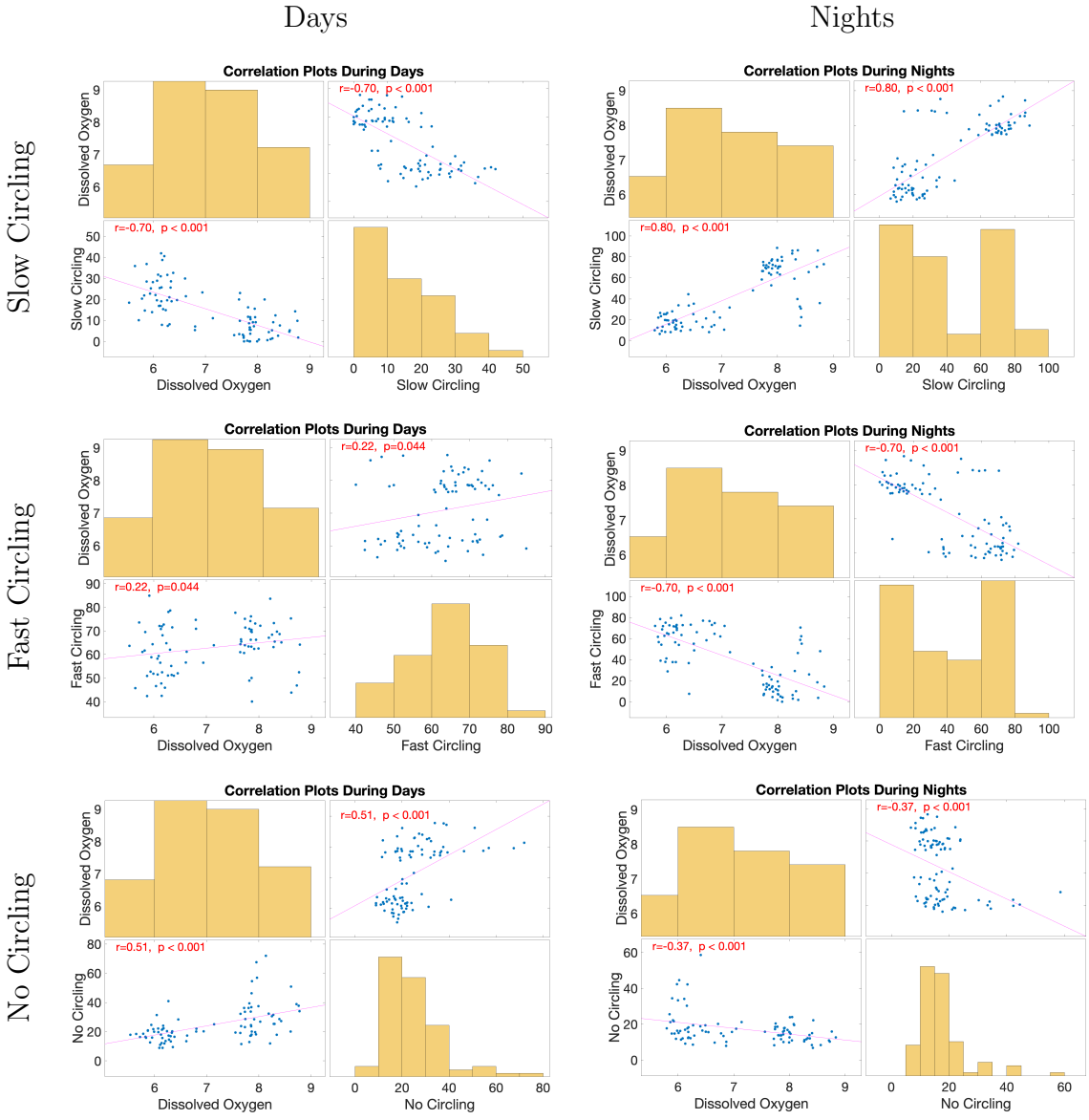


Figure 4.13: Correlation plots for dissolved oxygen and swimming classes. Each point represents the average percentage of a swimming class for fish population and the average dissolved oxygen, during the daytime in the first column and during the night in the second column.

4.2.2 Circling Classes VS Water Temperature

Since temperature and dissolved oxygen are highly correlated, the observations in this section closely mirror those in Section 4.2.1. As temperature rises, the metabolic rate of the fish increases, leading to increased oxygen consumption, while oxygen solubility reduces, resulting in a decrease in oxygen supply [67]. Previous studies have suggested that the swimming speed of the salmon increases with the increase in temperature [56], [33], [68]. In Figure 4.14 we visualize the average water temperature every 20 minutes during the study period. The average water temperature considerably decreased from 16.37 ℃ before October 18^{th} to 7.77 ℃ after October 18^{th}. Pearson correlation analysis was performed at a 0.05 confidence interval on the percentage of swimming classes and average water temperature during daytimes and nights. We observe an opposite correlation direction during daytime and night (Figure 4.15).

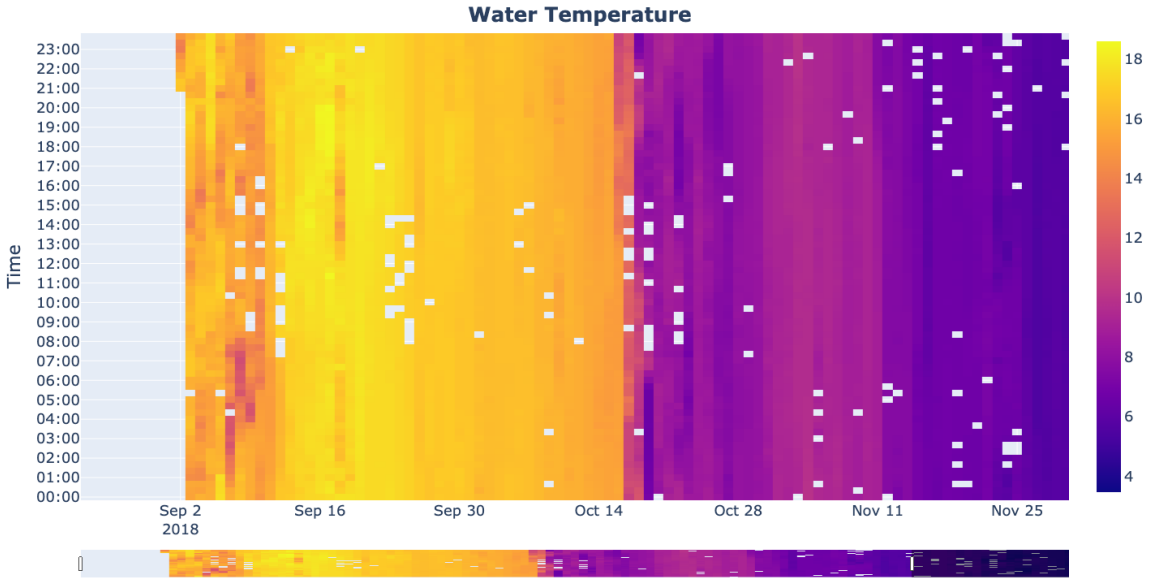


Figure 4.14: Water temperature (℃).

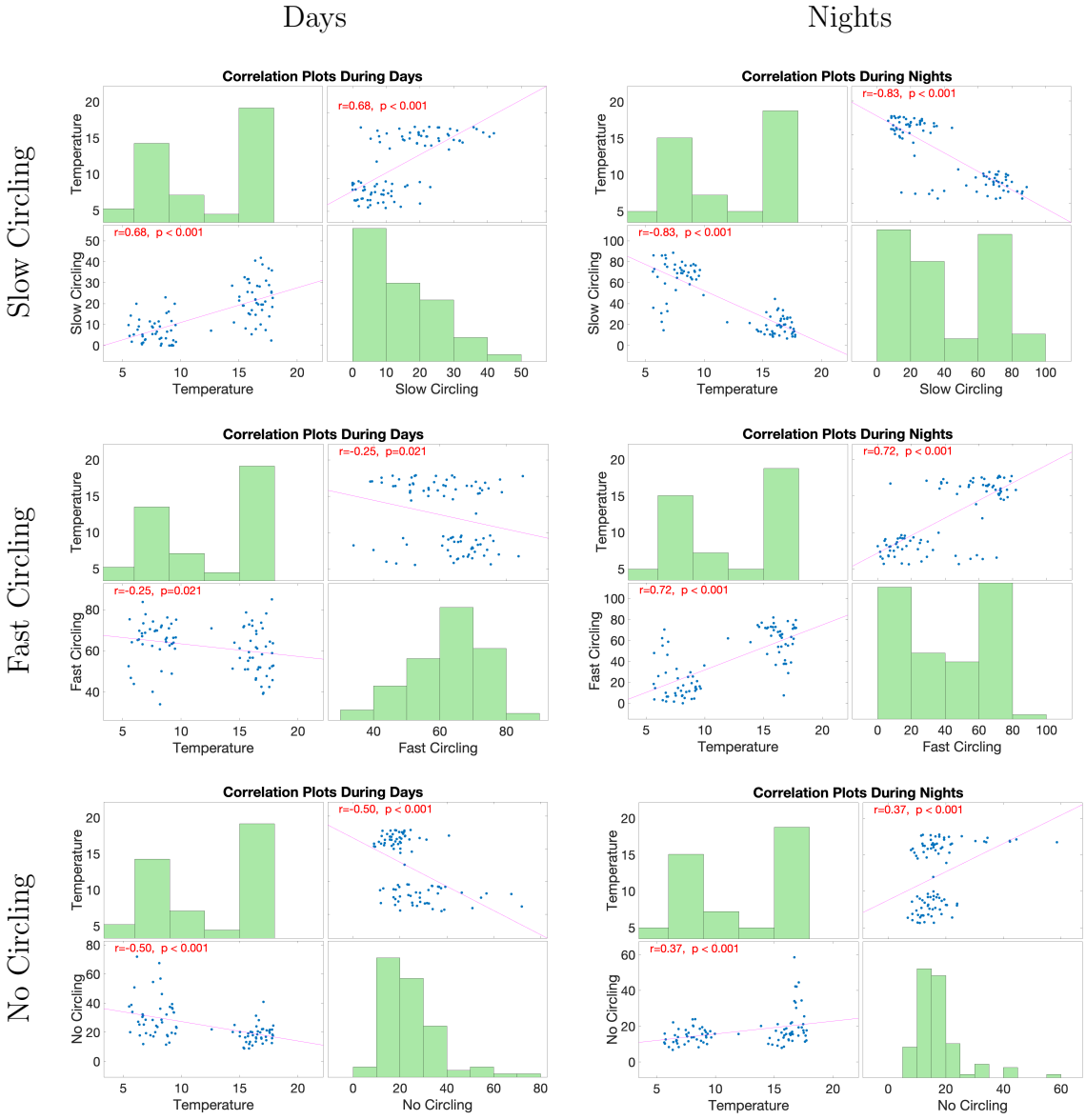


Figure 4.15: Correlation plots for water temperature and swimming classes. Each point represents the average percentage of a swimming class for fish population and the average temperature, during the daytime in the first column and during the night in the second column.

Looking at Figure 4.15, there is a strong positive correlation between the average water temperature and slow circling during the daytime and a very strong negative correlation during the night; while between water temperature and fast circling, we observe a weak negative correlation during the daytime and a strong positive correlation during the night. Furthermore, between no circling and average water

temperature, a moderate negative correlation during the daytime and a weak positive correlation during the night is observed. The positive correlation between slow circling and temperature during the day indicates that slow circling increases as temperature increases during the day and the negative correlation between slow circling and temperature during the night suggests that slow circling decreases as temperature increases during the nights. In summary, the relationship between swimming classes and water temperature is different during the days and nights.

4.2.3 Circling Classes VS Wind Speed

The main water movement is caused by wind, tide, and current residuals [74]. Lagardère et al., found a significant increase in the swimming speed of Sole (Solea solea) during the strong winds (> 8 m/s) with an elliptical-shaped swimming pattern [44]. Another study suggested that during a hailstorm, the swimming activity of the Sole significantly decreased and weaker transmitter signals made it hard to locate the fish [45]. In addition, during calm weather or low currents, Atlantic salmon swim at their preferred speed in a circular pattern [34]. We visualize the wind speed measured in km/h at Shelburne in Figure 4.16. A storm in mid-October disrupted the station at Shelburne and resulted in losing data for one week.

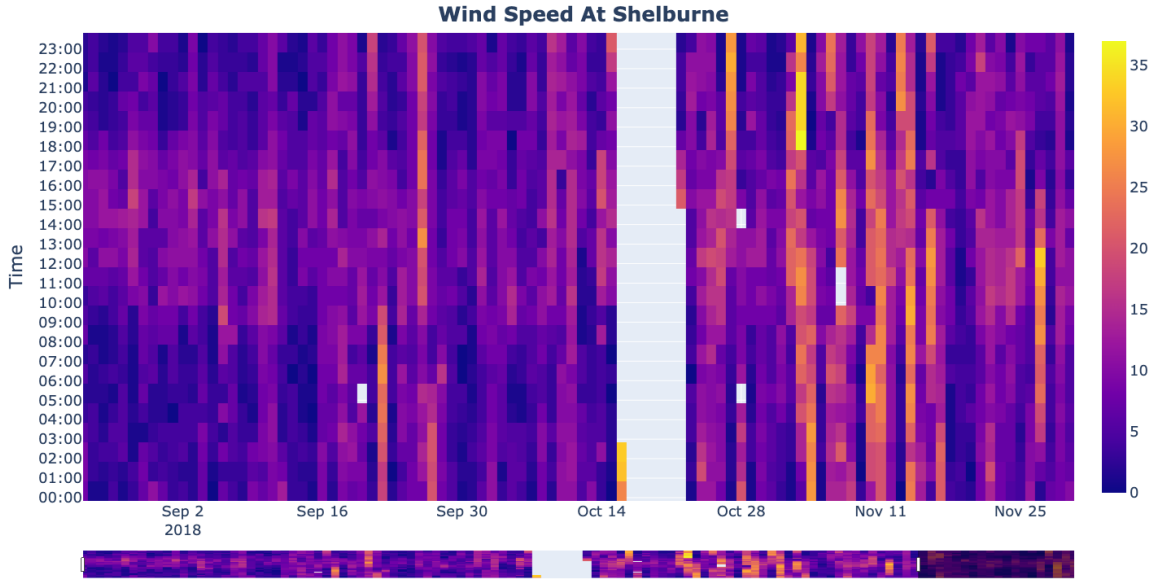


Figure 4.16: Wind speed (km/h) at Shelburne.

When wind speed is greater than 28 km/h, the swimming activity of the fish switched from higher percentage of fast circling and lower percentage of slow circling to higher percentage of slow circling and lower percentage of fast circling (Figure 4.17). Pearson's chi-squared test is used to determine the significance of the relationship between slow and fast circling during wind speed less and higher than 28 km/h ($\chi^2 = 10.4625, p = 0.0012, df = 1$). Figure 4.18 shows that the percentage slow circling is higher than fast circling for wind speed greater than 28 km/h. This indicates that during stormy weather fish tend to circle slowly more often, while in a previous study no significant changes in the swimming speed of Atlantic salmon were observed two days before, during, and two days after a storm identified by wind speed 30 ± 8 km/h [68].

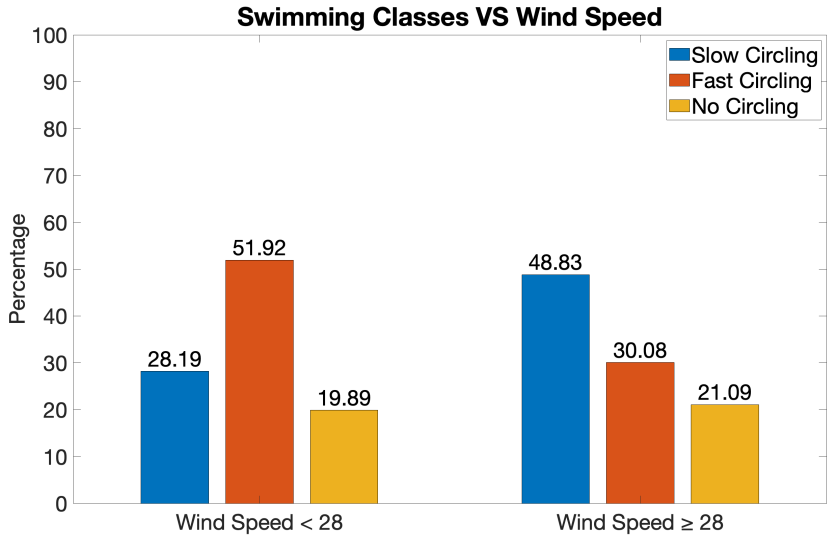


Figure 4.17: The percentage of each swimming class during wind speed less and higher than 28 km/h

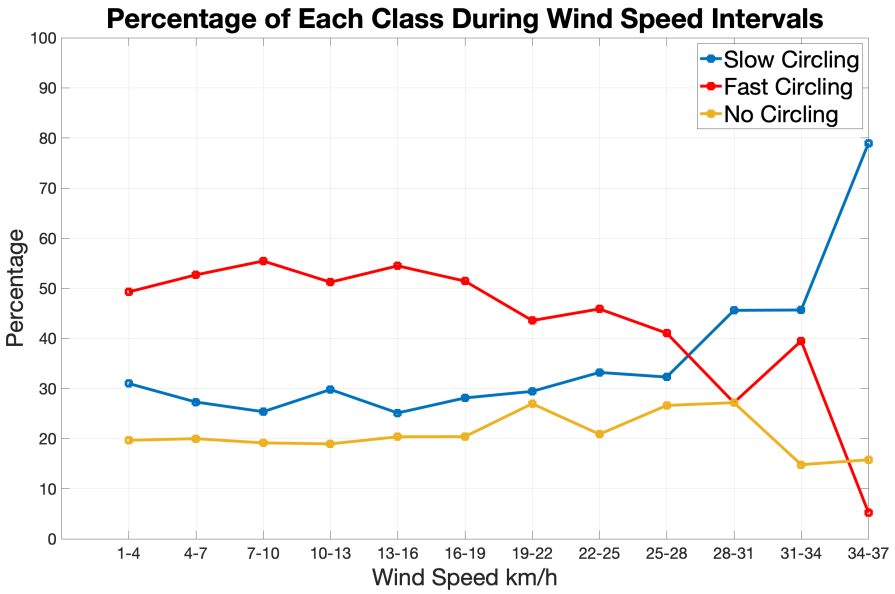


Figure 4.18: The percentage of each swimming class during wind speed intervals.

We examine the correlations between swimming class percentages and average wind speeds during daytimes and nights using Pearson correlation analysis at a 0.05 confidence interval and show the results in Figure 4.19. We observe that correlations during daytime and night have opposite directions. Between the average wind speed and slow circling, there is a weak negative correlation during the daytime and a weak positive correlation during the night. An opposite pattern is observed between fast circling and wind speed during days and nights, where there is a weak positive correlation between them during the daytime and a weak negative correlation during the night. There is also a weak positive correlation between no circling and average wind speed during the days, and there is a weak negative correlation between them during the nights.

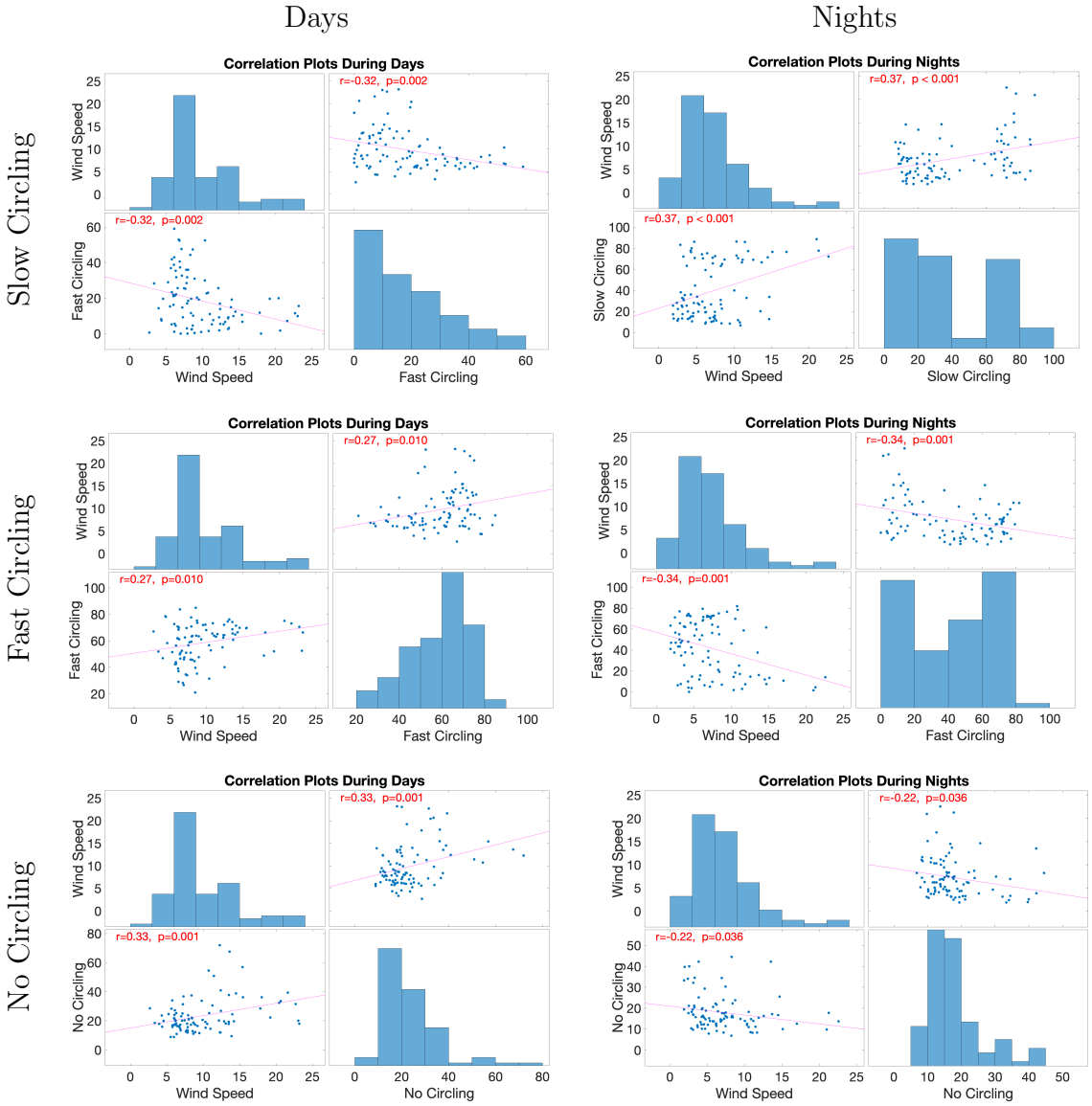


Figure 4.19: Correlation plots for wind speed and swimming classes. Each point represents the average percentage of a swimming class for fish population and the average wind speed, during the daytime in the first column and during the night in the second column.

4.2.4 Circling Classes VS Feeding

Several studies found that fish swimming speed or activity increased during the feeding [50], [3], [40], [68]. This behavior could be due to food deprivation or expanding the search area and increasing the chance of finding food [2]. The feeding times during this study are shown in Figure 4.20. We define feeding as the time that food was

available to the fish and not feeding as other times that fish was not being fed. Then, we look at the percentage of each swimming class during feeding and nonfeeding periods in Figure 4.21. Pearson's chi-squared test is used to determine the significance of the relationship between slow and fast circling during feeding and nonfeeding periods ($\chi^2 = 4.4258, p = 0.0354, df = 1$). Figure 4.21 shows that During the feeding, the percentage of slow circling significantly decreased while the percentage of fast and no circling increased. This indicates that the circling speed increased during the feeding which is in accordance with the literature.

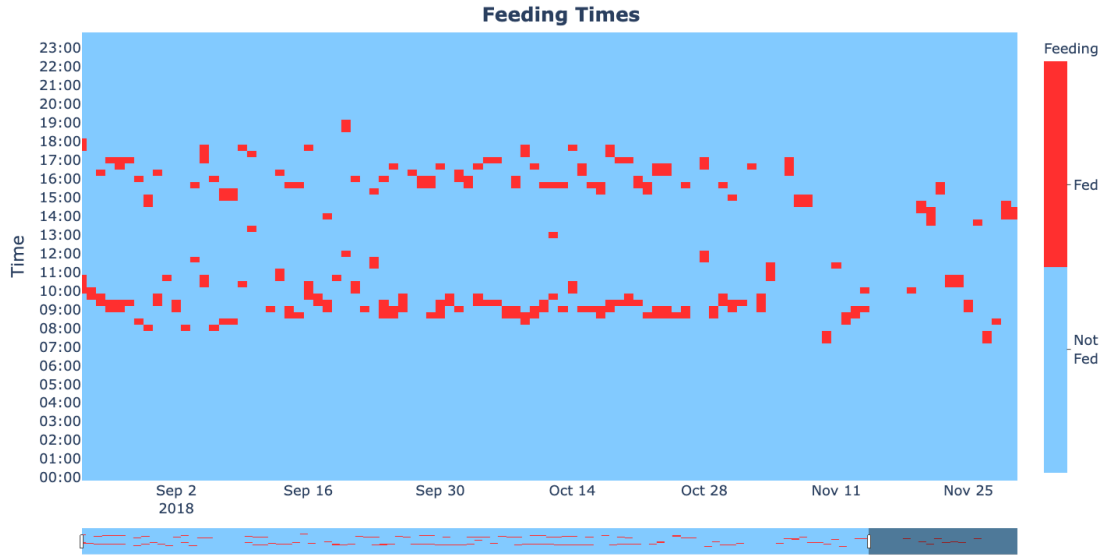


Figure 4.20: Feeding times.

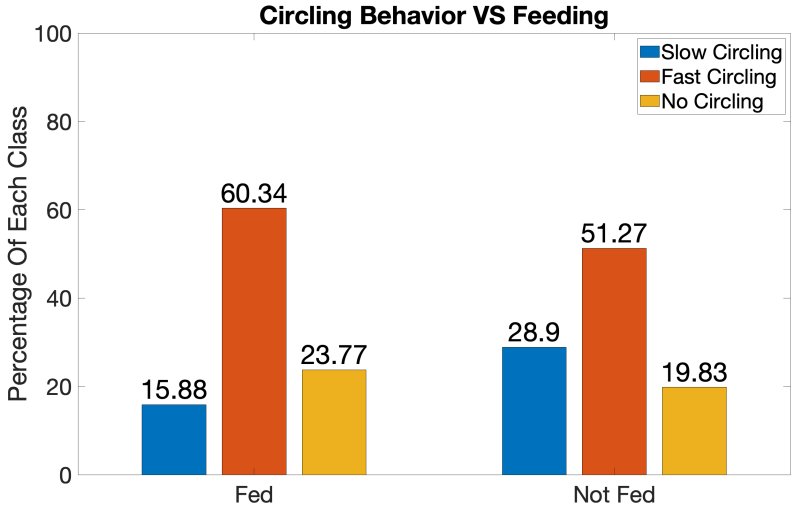


Figure 4.21: The percentage of each swimming class during feeding and not feeding

4.2.5 Circling Classes VS Weather Condition

Holm et al., observed that the swimming activity of Atlantic halibut was higher during cloudy days compared to sunny days [31]. At night and on cloudy days as the clouds prevent the light to reach the water surface, oxygen consumption by planktonic, algae respiration and diminishing photosynthesis activity lead to a decline in oxygen levels. [79]. We visualize the weather conditions during the study period in Figure 4.22. The percentages of each class during different weather conditions of "clear", "rain", "overcast", "partially cloudy" and "partially cloudy, and rain" are shown in Figure 4.23. Pearson's chi-squared test is used to determine whether there is a statistically significant relationship between slow and fast circling during different weather classes ($\chi^2 = 37.6435, p < 0.001, df = 4$). Overcast weather has the highest percentage of fast circling and the lowest percentages of slow and no circling (an almost similar pattern is observed during partially cloudy). The percentages of all the classes are approximately similar during the rain and clear weather. Rain, partially cloudy weather condition, has the highest percentage of slow circling and the lowest percentage of fast circling, and this distinct observation could be due to insufficient data for this weather. The significant increase in the fast circling behavior during the overcast and partially cloudy weather could be due to low oxygen levels during these conditions as we observed in Section 4.2.1 that the activity of salmon increased when oxygen levels dropped.

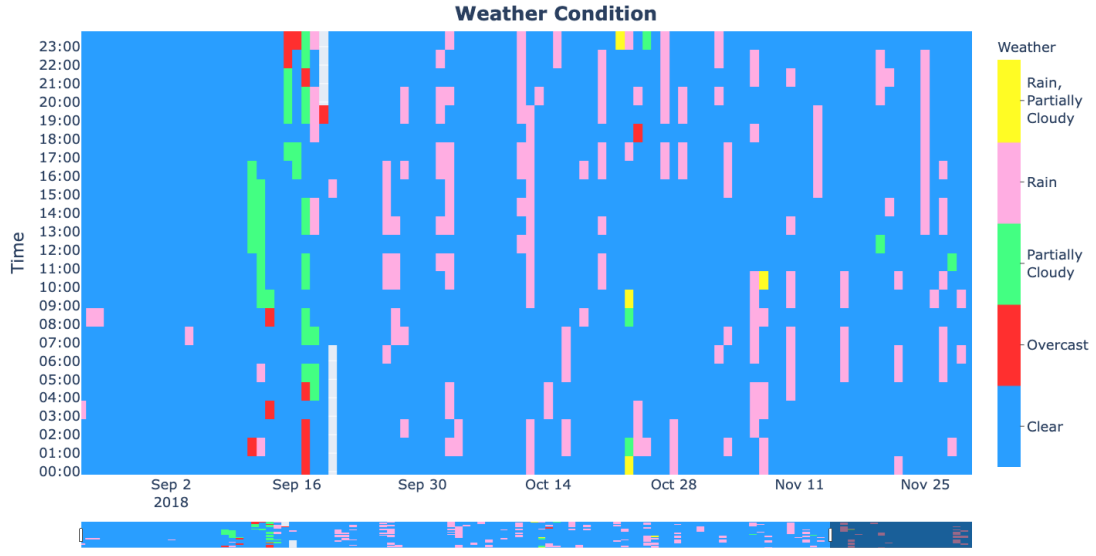


Figure 4.22: Weather conditions.

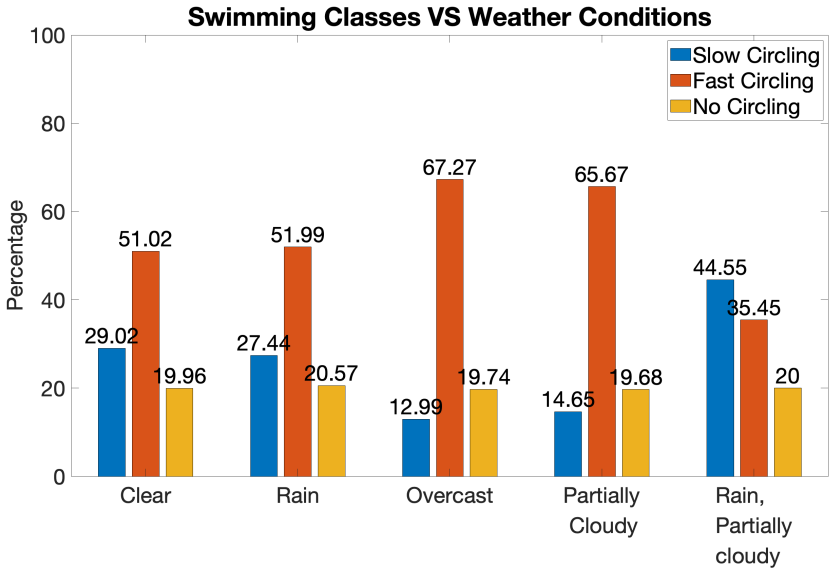


Figure 4.23: The percentage of each swimming class during different weather conditions.

4.2.6 Circling Classes VS Water Level

Kadri et al., found a weak positive correlation between the swimming speed of Atlantic salmon and low water levels [38]. In another study, bonefish were found more active during low tides and less active during the high tides [11]. We show the water level averaged over 20 minutes during the study period in Figure 4.24. Water level ranged from 0.14 to 2.58 meters. We compare the percentages of each swimming class during water level less than 0.5 and greater than 0.5 meters. Then we use Pearson's chi-squared test to examine if there is a statistically significant relationship between slow and fast circling when the water level is less and higher than 0.5 meters ($\chi^2 = 3.9651, p = 0.0465, df = 1$). Looking at Figure 4.24, when the water level is below 0.5 meters, the percentage of fast circling is significantly higher and the percentages of slow is significantly lower than when water level is greater then 0.5 meters. The increase in fast circling when the water level is less than 0.5 meter could indicate that the salmon swimming speed increases when the water level is very low.

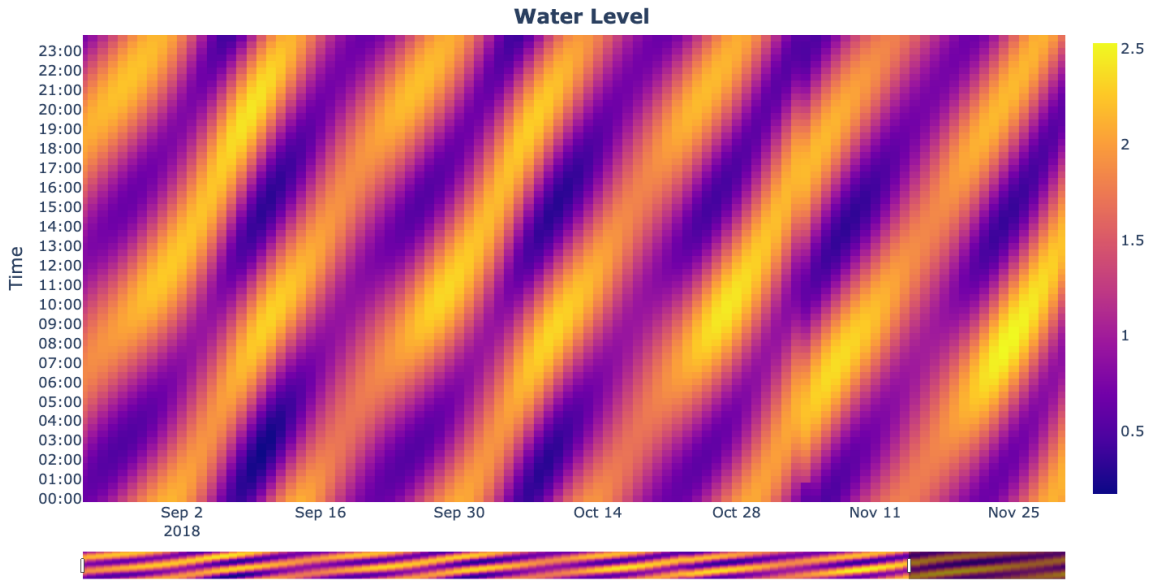


Figure 4.24: Water levels in meters.

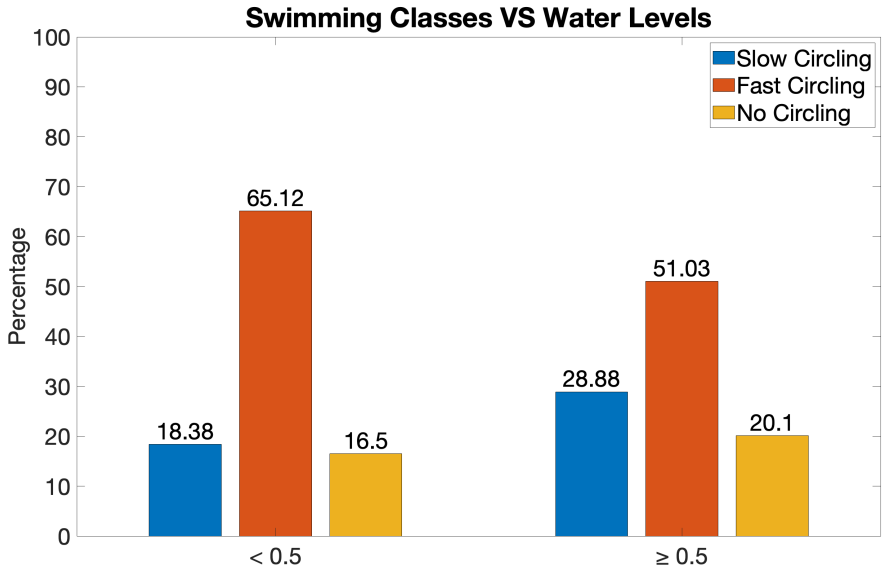


Figure 4.25: The percentage of each swimming class during water level less and greater than 0.5 meters.

Chapter 5

Conclusion and Future Work

In this chapter, we briefly go over what we covered in this thesis and explain the limitations and future work to enhance this approach.

5.1 Summary

The main objective of this thesis was to develop a tool to detect circling and non-circling behavior of farmed fish and determine the speed of the circling. To that end, from the movement trajectories of the fish, we calculated polar coordinates and used the resulting angles for distinguishing circling from no circling behaviors. Due to the linear arrangement of the angles when the fish is circling, this behavior can be detected using a linear model and on account of the presence of the errors in both time and positions, we used a Deming regression model to extract the linear feature of the angles. Since the angles periodically change between $-\pi$ and π through time, a locally weighted form of Deming regression is used to unwrap the angles and eliminate the discontinuities. SER values from the Deming regression line are used to distinguish circling from no circling and the coefficients of the model give us AV values which are used for identifying slow and fast circling. Then we analyzed the relationships between several environmental variables and identified swimming class.

5.2 Discussion

In this study, we created an interactive bivariate visualization of AV and SER values for every fish with the ability to look at any intervals during the study period plus a tooltip feature that presents the exact AV, SER, time, and date values. The crucial benefit of compressing the data from multiple plots into one using colors is allowing direct and intuitive evaluation of relationships between variables and their distributions [70]. However, interpreting bivariate maps can be more challenging than

univariate ones due to their complexity. Nevertheless, utilizing bivariate maps would be more suitable, because we are constrained by space on the screen and visualizing twenty subplots (seperate AV and SER plots for ten fish) would be inefficient. Another drawback of using color channels is since the capacity of humans to distinguish between distinct colors is restricted, tiny color shifts may not be apparent to color-blind people and sometimes interesting patterns may remain undetected. To overcome these problems, we provided the user with several interactive operations to zoom in and observe the actual AV and SER value of the desired segments. Also, a color blindness simulation software can be used to select color blind friendly colors [39]. We used AV and SER plots for a preliminary exploration of the patterns in the data. Then we defined three swimming classes to facilitate further visualization and identifying patterns in the group and individual behaviors.

Furthermore, we analyzed the impact of changes in external variables on each swimming class. Our results show that the fish tends to be fast circling more or as often at night compared to daytime during the late summer and early fall, while fast circling significantly decreased at night after October 18^{th}, during the mid and late fall. This can be a complementary to a previous study where the salmon swimming speed at night significantly decreased only during the winter and early spring and not during the late spring and early summer [56]. Hourly distribution of the fish swimming classes highly differs before and after October 18^{th} where from 00:00 to 7:00 and 18:00 to 00:00 slow circling significantly increased and fast circling considerably decreased after October 18^{th}. When dissolved oxygen is lower than 7 mg/L, fast circling significantly increased and slow circling is significantly reduced. This could be on account of increasing the chance of finding a more oxygenated water and lower temperature [46] which is contrary to [40], where the swimming activity of Atlantic salmon significantly decreased during hypoxic conditions. The correlation between each swimming class and dissolved oxygen, water temperature, and wind speed has an opposite direction during daytime and night-time. When the wind speed is greater than 28 km/h, slow circling significantly increased and fast circling decreased. This is as opposed to the behavior of Sole (Solea solea), which increased their swimming speed during strong winds ($>$ 8 m/s) [44]. In addition, during the feeding activity, overcast and partially cloudy weather, and when the water level is less than 0.5

meters we observed an increase in fast circling and a decrease in slow circling. An increase in swimming speed during the feeding was observed in the previous studies as well [50], [68] which could be caused by food deprivation or elevating the chance of finding food by expanding the search area [2]. Increasing swimming activity in response to cloudy weather was observed in Atlantic halibut [31] which could be as a result of low oxygen levels owing to the lack of light during the cloudy days. [79]. Last but not least, our result during the low water level is in accordance with previous studies on Atlantic salmon and bonefish were they found to be more active during low tides [38], [11]. By providing appropriate tools and insights on the circling movement of the farmed fish and analyzing the impact of external variables on their behavior, we strive to aid oceanographers to improve understanding of fish behaviour.

5.3 Limitations And Future Work

In the following lines, we discuss limitations of this approach and explain different ideas for expanding and improving this study in future work.

- The effect of environmental variables on the swimming classes can be analyzed and compared for various fish species during different months of the year and locations.

- The impact of different external variables such as water currents, salinity and stocking density that are not covered in this thesis can be analyzed in future studies.

- Depending on the quality of the data, smaller intervals than 20 minutes can be used to classify the swimming behavior of the fish.

- The AV and SER thresholds that are used in this study to classify the swimming behavior of the fish may vary in other studies depending on the different data characteristics. We determined these thresholds by observing the swimming behavior of the fish during several periods which can be avoided by developing a method that can extract the features of different behaviors automatically and classify them accordingly.

- Other swimming classes such as stationary swimming behavior can be defined using angular velocities in future studies.

Bibliography

[1] Dong An, Jinze Huang, and Yaoguang Wei. A survey of fish behaviour quantification indexes and methods in aquaculture. *Reviews in Aquaculture*, 13(4):2169–2189, 2021.

[2] JE Andrew, J Holm, Sunil Kadri, and Felicity A Huntingford. The effect of competition on the feeding efficiency and feed handling behaviour in gilthead sea bream (Sparus aurata L.) held in tanks. *Aquaculture*, 232(1-4):317–331, 2004.

[3] JE Andrew, C Noble, S Kadri, H Jewell, and FA Huntingford. The effect of demand feeding on swimming speed and feeding responses in Atlantic salmon Salmo salar l., gilthead sea bream Sparus aurata L. and European sea bass dicentrarchus labrax l. in sea cages. *Aquaculture Research*, 33(7):501–507, 2002.

[4] Gennady Andrienko, Natalia Andrienko, Christophe Hurter, Salvatore Rinzivillo, and Stefan Wrobel. Scalable analysis of movement data for extracting and exploring significant places. *IEEE Transactions on Visualization and Computer Graphics*, 19(7):1078–1094, 2012.

[5] Marie-Laure Bégout Anras and Jean Paul Lagardère. Measuring cultured fish swimming behaviour: first results on rainbow trout using acoustic telemetry in tanks. *Aquaculture*, 240(1-4):175–186, 2004.

[6] NC Armansin, KA Lee, Charles Huveneers, and RG Harcourt. Integrating social network analysis and fine-scale positioning to characterize the associations of a benthic shark. *Animal Behaviour*, 115:245–258, 2016.

[7] E. Batschelet. *Circular Statistics in Biology*. Mathematics in biology. Academic Press, 1981.

[8] Sarah L Becker, John T Finn, Ashleigh J Novak, Andy J Danylchuk, Clayton G Pollock, Zandy Hillis-Starr, Ian Lundgren, and Adrian Jordaan. Coarse-and fine-scale acoustic telemetry elucidates movement patterns and temporal variability in individual territories for a key coastal mesopredator. *Environmental Biology of Fishes*, 103(1):13–29, 2020.

[9] Cigdem Beyan and Robert B Fisher. Detection of abnormal fish trajectories using a clustering based hierarchical classifier. In *British Machine Vision Conference (BMVC)*, pages 1–11, 2013.

[10] Jacob W Brownscombe, Lucas P Griffin, Tyler O Gagne, Christopher R Haak, Steven J Cooke, John T Finn, and Andy J Danylchuk. Environmental drivers of habitat use by a marine fish on a heterogeneous and dynamic reef flat. *Marine Biology*, 166(2):1–13, 2019.

[11] Jacob W Brownscombe, Lee FG Gutowsky, Andy J Danylchuk, and Steven J Cooke. Foraging behaviour and activity of a marine benthivorous fish estimated using tri-axial accelerometer biologgers. *Marine Ecology Progress Series*, 505:241–251, 2014.

[12] Kim Burt, Dounia Hamoutene, Gehan Mabrouk, Chris Lang, Thomas Puestow, Dwight Drover, Randy Losier, and Fred Page. Environmental conditions and occurrence of hypoxia within production cages of Atlantic salmon on the south coast of Newfoundland. *Aquaculture Research*, 43(4):607–620, 2012.

[13] Lauren J Chapman and David J Mckenzie. Behavioral responses and ecological consequences. In *Fish Physiology*, volume 27, pages 25–77. Elsevier, 2009.

[14] Julia H Coates, Kevin A Hovel, John L Butler, A Peter Klimley, and Steven G Morgan. Movement and home range of pink abalone Haliotis corrugata: implications for restoration and population recovery. *Marine Ecology Progress Series*, 486:189–201, 2013.

[15] Steven J Cooke. Biotelemetry and biologging in endangered species research and animal conservation: relevance to regional, national, and IUCN red list threat assessments. *Endangered Species Research*, 4(1-2):165–185, 2008.

[16] Jane E Darbyshire and Bernhard Jenny. Natural-color maps via coloring of bivariate grid data. *Computers & Geosciences*, 106:130–138, 2017.

[17] John C Davis. Minimal dissolved oxygen requirements of aquatic life with emphasis on Canadian species: a review. *Journal of the Fisheries Board of Canada*, 32(12):2295–2332, 1975.

[18] Micah J Dean, William S Hoffman, Douglas R Zemeckis, and Michael P Armstrong. Fine-scale diel and gender-based patterns in behaviour of Atlantic cod (Gadus morhua) on a spawning ground in the western Gulf of Maine. *ICES Journal of Marine Science*, 71(6):1474–1489, 2014.

[19] W Edwards Deming. Statistical adjustment of data. *New York: Dover*, 1964.

[20] Tim Dempster, Øyvind Korsøen, Ole Folkedal, Jon-Erik Juell, and Frode Oppedal. Submergence of Atlantic salmon (Salmo salar l.) in commercial scale sea-cages: a potential short-term solution to poor surface conditions. *Aquaculture*, 288(3-4):254–263, 2009.

[21] Paolo Domenici, John F Steffensen, and Stefano Marras. The effect of hypoxia on fish schooling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1727):20160236, 2017.

[22] Paolo Domenici, John Fleng Steffensen, and Robert S Batty. The effect of progressive hypoxia on swimming activity and schooling in atlantic herring. *Journal of Fish Biology*, 57(6):1526–1538, 2000.

[23] Mario Espinoza, Thomas J Farrugia, Dale M Webber, Frank Smith, and Christopher G Lowe. Testing a new acoustic telemetry technique to quantify long-term, fine-scale movements of aquatic animals. *Fisheries Research*, 108(2-3):364–371, 2011.

[24] Bernard G Francq and Bernadette B Govaerts. Measurement methods comparison with errors-in-variables regressions. from horizontal to vertical OLS regression, review and new perspectives. *Chemometrics and Intelligent Laboratory Systems*, 134:123–139, 2014.

[25] Manuel Gesto, Walter Zupa, Sébastien Alfonso, Maria Teresa Spedicato, Giuseppe Lembo, and Pierluigi Carbonara. Using acoustic telemetry to assess behavioral responses to acute hypoxia and ammonia exposure in farmed rainbow trout of different competitive ability. *Applied Animal Behaviour Science*, 230:105084, 2020.

[26] Matthew M Guzzo, Paul J Blanchfield, Andrew J Chapelsky, and Peter A Cott. Resource partitioning among top-level piscivores in a sub-Arctic lake during thermal stratification. *Journal of Great Lakes Research*, 42(2):276–285, 2016.

[27] Tom Johnny Hansen, Per Gunnar Fjelldal, Ole Folkedal, Tone Vågseth, and Frode Oppedal. Effects of light source and intensity on sexual maturation, growth and swimming behaviour of Atlantic salmon in sea cages. *Aquaculture Environment Interactions*, 9:193–204, 2017.

[28] Tomislav Hengl. Visualisation of uncertainty using the hsi colour model: computations with colours. In *Proceedings of the 7th International Conference on GeoComputation*, pages 8–17. University of Southampton Southampton, United Kingdom, 2003.

[29] Tomislav Hengl and Norair Toomanian. Maps are not what they seem: representing uncertainty in soil-property maps. In *Proc. Accuracy*, pages 805–813. Citeseer, 2006.

[30] NA Herbert and JF Steffensen. The response of Atlantic cod, Gadus morhua, to progressive hypoxia: fish swimming speed and physiological stress. *Marine Biology*, 147(6):1403–1412, 2005.

[31] Jens Christian Holm, Stig Tuene, and Jan Erik Fosseidengen. Halibut behaviour as a means of assessing suitability of ongrowth systems. *International Council for the Exploration of the Sea (ICES)*, 1998.

[32] Eric R Huber and Stephanie M Carlson. Environmental correlates of fine-scale juvenile steelhead trout (Oncorhynchus mykiss) habitat use and movement patterns in an intermittent estuary during drought. *Environmental Biology of Fishes*, 103(5):509–529, 2020.

[33] Malthe Hvas, Ole Folkedal, Albert Imsland, and Frode Oppedal. The effect of thermal acclimation on aerobic scope and critical swimming speed in Atlantic salmon, Salmo salar. *Journal of Experimental Biology*, 220(15):2757–2764, 2017.

[34] Malthe Hvas, Ole Folkedal, and Frode Oppedal. What is the limit of sustained swimming in Atlantic salmon post smolts? *Aquaculture Environment Interactions*, 13:189–198, 2021.

[35] David Johansson, Frida Laursen, Anders Fernö, Jan Erik Fosseidengen, Pascal Klebert, Lars Helge Stien, Tone Vågseth, and Frode Oppedal. The interaction between water currents and salmon swimming behaviour in sea cages. *PloS one*, 9(5):e97635, 2014.

[36] David Johansson, Kari Ruohonen, Anders Kiessling, Frode Oppedal, Jan-Erik Stiansen, Mark Kelly, and Jon-Erik Juell. Effect of environmental factors on swimming depth preferences of Atlantic salmon (Salmo salar L.) and temporal and spatial variations in oxygen levels in sea cages at a fjord site. *Aquaculture*, 254(1-4):594–605, 2006.

[37] Jon-Erik Juell and Håkan Westerberg. An ultrasonic telemetric system for automatic positioning of individual fish used to track Atlantic salmon (Salmo salar l.) in a sea cage. *Aquacultural Engineering*, 12(1):1–18, 1993.

[38] Sunil Kadri, Neil B Metcalfe, Felicity A Huntingford, and John E Thorpe. Daily feeding rhythms in Atlantic salmon in sea cages. *Aquaculture*, 92:219–224, 1991.

[39] NR Kaye, A Hartley, and D Hemming. Mapping the climate: guidance on appropriate techniques to map climate variables and their uncertainty. *Geoscientific Model Development*, 5(1):245–256, 2012.

[40] J Kolarevic, Øyvind Aas-Hansen, Å Espmark, G Baeverfjord, B Fyhn Terjesen, and B Damsgård. The use of acoustic acceleration transmitter tags for monitoring of atlantic salmon swimming activity in recirculating aquaculture systems (RAS). *Aquacultural Engineering*, 72:30–39, 2016.

[41] Kazuyoshi Komeyama, Minoru Kadota, Shinsuke Torisawa, Katsuya Suzuki, Yuichi Tsuda, and Tsutomu Takagi. Measuring the swimming behaviour of a reared Pacific bluefin tuna in a submerged aquaculture net cage. *Aquatic Living Resources*, 24(2):99–105, 2011.

[42] Kazuyoshi Komeyama, Minoru Kadota, Shinsuke Torisawa, and Tsutomu Takagi. Three-dimensional trajectories of cultivated Pacific bluefin tuna thunnus orientalis in an aquaculture net cage. *Aquaculture Environment Interactions*, 4(1):81–90, 2013.

[43] Øyvind J Korsøen, Tim Dempster, Per Gunnar Fjelldal, Frode Oppedal, and Tore S Kristiansen. Long-term culture of Atlantic salmon (salmo salar l.) in

submerged cages during winter affects behaviour, growth and condition. *Aquaculture*, 296(3-4):373–381, 2009.

[44] JP Lagardère, ML Bégout, JY Lafaye, and JP Villotte. Influence of wind-produced noise on orientation in the sole (Solea solea). *Canadian Journal of Fisheries and Aquatic Sciences*, 51(6):1258–1264, 1994.

[45] JP Lagardère, JJ Ducamp, L Frikha, and M Spèrandio. Ultrasonic tracking of common sole juveniles (Solea vulgaris Quensel, 1806) in a saltmarsh: methods and fish response to some environmental factors. *Journal of applied ichthyology*, 4(2):87–96, 1988.

[46] C Lefrançois, RS Ferrari, J Moreira Da Silva, and P Domenici. The effect of progressive hypoxia on spontaneous activity in single and shoaling golden grey mullet liza aurata. *Journal of Fish Biology*, 75(7):1615–1625, 2009.

[47] Daoliang Li, Guangxu Wang, Ling Du, Yingying Zheng, and Zhenhu Wang. Recent advances in intelligent recognition methods for fish stress behavior. *Aquacultural Engineering*, page 102222, 2021.

[48] Kristian Linnet. Performance of deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clinical Chemistry*, 44(5):1024–1031, 1998.

[49] Kristian Linnet. Necessary sample size for method comparison studies based on regression analysis. *Clinical Chemistry*, 45(6):882–894, 1999.

[50] Catarina IM Martins, Leonor Galhardo, Chris Noble, Børge Damsgård, Maria T Spedicato, Walter Zupa, Marilyn Beauchaud, Ewa Kulczykowska, Jean-Charles Massabuau, Toby Carter, et al. Behavioural indicators of welfare in farmed fish. *Fish Physiology and Biochemistry*, 38(1):17–41, 2012.

[51] Trevor D Meckley, Christopher M Holbrook, C Michael Wagner, and Thomas R Binder. An approach for filtering hyperbolically positioned underwater acoustic telemetry data with position precision estimates. *Animal Biotelemetry*, 2(1):1–13, 2014.

[52] Nikolaos Mittas, Makrina Viola Kosti, Vasiliki Argyropoulou, and Lefteris Angelis. Modeling the relationship between software effort and size using deming regression. In *Proceedings of the 6th International Conference on Predictive Models in Software Engineering*, PROMISE '10, New York, NY, USA, 2010. Association for Computing Machinery.

[53] Lidia Muñoz, Eneko Aspillaga, Miquel Palmer, João L Saraiva, and Pablo Arechavala-Lopez. Acoustic telemetry: a tool to monitor fish swimming behavior in sea-cage aquaculture. *Frontiers in Marine Science*, page 645, 2020.

[54] Tomoko Narazaki, Itsumi Nakamura, Kagari Aoki, Takashi Iwata, Kozue Shiomi, Paolo Luschi, Hiroyuki Suganuma, Carl G Meyer, Rui Matsumoto, Charles A Bost, et al. Similar circling movements observed across marine megafauna taxa. *Iscience*, 24(4):102221, 2021.

[55] Ashleigh J Novak, Sarah L Becker, John T Finn, Clayton G Pollock, Zandy Hillis-Starr, and Adrian Jordaan. Scale of biotelemetry data influences ecological interpretations of space and habitat use in Yellowtail Snapper. *Marine and Coastal Fisheries*, 12(5):364–377, 2020.

[56] F Oppedal, J-E Juell, GL Tarranger, and T Hansen. Artificial light and season affects vertical distribution and swimming behaviour of post-smolt Atlantic salmon in sea cages. *Journal of Fish Biology*, 58(6):1570–1584, 2001.

[57] 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.

[58] Arjan P Palstra, Pablo Arechavala-Lopez, Yuanxu Xue, and Ana Roque. Accelerometry of seabream in a sea-cage: is acceleration a good proxy for activity? *Frontiers in Marine Science*, 8:144, 2021.

[59] Vassilis M Papadakis, Ioannis E Papadakis, Fani Lamprianidou, Alexios Glaropoulos, and Maroudio Kentouri. A computer-vision system and methodology for the analysis of fish behavior. *Aquacultural Engineering*, 46:53–59, 2012.

[60] Christian Lentz Pedersen. Energy budgets for juvenile rainbow trout at various oxygen concentrations. *Aquaculture*, 62(3-4):289–298, 1987.

[61] Lisa Maria Perkhofer, Peter Hofer, Conny Walchshofer, Thomas Plank, and Hans-Christian Jetter. Interactive visualization of big data in the field of accounting: A survey of current practice and potential barriers for adoption. *Journal of Applied Accounting Research*, 2019.

[62] TH Pinkiewicz, GJ Purser, and RN Williams. A computer vision system to analyse the swimming behaviour of farmed fish in commercial aquaculture facilities: a case study using cage-held Atlantic salmon. *Aquacultural Engineering*, 45(1):20–27, 2011.

[63] Romain Roy, Jeremy Beguin, Christine Argillier, Laurence Tissot, Frank Smith, Stephanie Smedbol, and Eric De-Oliveira. Testing the VEMCO positioning system: spatial distribution of the probability of location and the positioning error in a reservoir. *Animal Biotelemetry*, 2(1):1–7, 2014.

[64] HENRIK Schurmann and J Steffensen. Spontaneous swimming activity of atlantic cod gadus morhua exposed to graded hypoxia at three temperatures. *The Journal of Experimental Biology*, 197(1):129–142, 1994.

[65] Daniel J Skerritt, Peter A Robertson, Aileen C Mill, Nicholas VC Polunin, and Clare Fitzsimmons. Fine-scale movement, activity patterns and home-ranges of European lobster homarus gammarus. *Marine Ecology Progress Series*, 536:203–219, 2015.

[66] Frank Smith. Understanding HPE in the VEMCO positioning system (VPS). *Available: http://www.oceans-research.com/wp-content/uploads/2016/09/understanding-hpe-vps.pdf*, 2013.

[67] Kilian M Stehfest, Chris G Carter, Jaime D McAllister, Jeff D Ross, and Jayson M Semmens. Response of Atlantic salmon Salmo salar to temperature and dissolved oxygen extremes established using animal-borne environmental sensors. *Scientific Reports*, 7(1):1–10, 2017.

[68] Caitlin L Stockwell, Ramón Filgueira, and Jon Grant. Determining the effects of environmental events on cultured Atlantic salmon behaviour using 3-dimensional acoustic telemetry. *Frontiers in Animal Science*, 2:26, 2021.

[69] Ming Tang and Daniel Boisclair. Relationship between respiration rate of juvenile brook trout (Salvelinus fontinalis), water temperature, and swimming characteristics. *Canadian Journal of Fisheries and Aquatic Sciences*, 52(10):2138–2145, 1995.

[70] AJ Teuling, R Stöckli, and Sonia I Seneviratne. Bivariate colour maps for visualizing climate data. *International journal of climatology*, 31(9):1408–1412, 2011.

[71] Jenna Vergeynst, Thomas Vanwyck, Raf Baeyens, Tom De Mulder, Ingmar Nopens, Ans Mouton, and Ine Pauwels. Acoustic positioning in a reflective environment: going beyond point-by-point algorithms. *Animal Biotelemetry*, 8(1):1–17, 2020.

[72] Thomas Waldrop, Steven Summerfelt, Patricia Mazik, and Christopher Good. The effects of swimming exercise and dissolved oxygen on growth performance, fin condition and precocious maturation of early-rearing Atlantic salmon Salmo salar. *Aquaculture Research*, 49(2):801–808, 2018.

[73] Guangxu Wang, Akhter Muhammad, Chang Liu, Ling Du, and Daoliang Li. Automatic recognition of fish behavior with a fusion of RGB and optical flow data based on deep learning. *Animals*, 11(10):2774, 2021.

[74] DJ Wildish, PD Keizer, AJ Wilson, and JL Martin. Seasonal changes of dissolved oxygen and plant nutrients in seawater near salmonid net pens in the macrotidal bay of fundy. *Canadian Journal of Fisheries and Aquatic Sciences*, 50(2):303–311, 1993.

[75] Heini Winthereig-Rasmussen, Knud Simonsen, and Øystein Patursson. Flow through fish farming sea cages: comparing computational fluid dynamics simulations with scaled and full-scale experimental data. *Ocean Engineering*, 124:21–31, 2016.

[76] Cheng Wu and Jian Zhen Yu. Evaluation of linear regression techniques for atmospheric applications: the importance of appropriate weighting. *Atmospheric Measurement Techniques*, 11(2):1233–1250, 2018.

[77] Ahmad Mohamadi Yalsuyi, Abdolmajid Hajimoradloo, Rasul Ghorbani, Vallyallah Jafari, Marko D Prokić, and Caterina Faggio. Behavior evaluation of rainbow trout (Oncorhynchus mykiss) following temperature and ammonia alterations. *Environmental Toxicology and Pharmacology*, 86:103648, 2021.

[78] Ling Yang, Yeqi Liu, Huihui Yu, Xiaomin Fang, Lihua Song, Daoliang Li, and Yingyi Chen. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: A review. *Archives of Computational Methods in Engineering*, 28:1–32, 09 2020.

[79] Takashi Yoshikawa, Osamu Murata, Ken Furuya, and Mitsuru Eguchi. Short-term covariation of dissolved oxygen and phytoplankton photosynthesis in a coastal fish aquaculture site. *Estuarine, Coastal and Shelf Science*, 74(3):515–527, 2007.