NORROY Tom
M1 Agronome

Internship Report:

Analysis of flat oyster heart rate and valve gape data

Photo: Oscar Bos/ Wageningen Marine Research (WUR website, 2017)

From the 4th of September 2023 to the 1st of December 2023

Supervisor: Mrs. Dr. Pauline KAMERMANS

Referent teacher: Mr. Jean-Louis MARCHAND

Host institution:
Wageningen Marine Research
Yerseke, Netherlands

TABLE OF CONTENT

ACKONWLEDGEMENTS

First, I would like to express my deep gratitude to Mrs. Pauline Kammermans, my supervisor at Wageningen Marine Research, for her unfailing support, invaluable guidance, and expertise throughout my placement. Her enlightened advice, attentive listening, patience, and kindness contributed greatly to my personal and professional development.

Then, I would like to express my deep gratitude to Mr. Hein Sas and the NORA Foundation for giving me the unique opportunity to attend NORA Conference No. 5 on flat oyster restoration in Europe for three days. This immersion in an international environment and in the world of research enabled me to consolidate my professional and personal interests.

I would also like to express my gratitude to the whole team at Wageningen Marine Research for their warm welcome, their invaluable collaboration and the moments shared with you, which have helped me a great deal to integrate.

I would like to express my sincere thanks to Mr. Jean Louis Marchand, my tutor, for his understanding and support during the difficult periods I went through. His empathy and kindness enabled me to tackle the challenges calmly and achieve my goals despite the circumstances. I am grateful for the extra time allowed to complete this report, which enabled me to produce it with serenity despite personal constraints.

My most sincere thanks go to all the people who contributed in any way to the success of this work placement and this report, in particular Ms. Justine et Morea and Ms. Valérie Rouchetet who provided me with support and advice throughout this period.

Glossary

AIC	Akaike Information Criterion
AMP	Amplitude
BPM	Beat per minute
GAM	General Additive Model
HW	heatwave
IR	Infrared
NGO	Non-Governmental Organization
WMR	Wageningen Marine Research

ABSTRACT

The restoration of flat oysters in Europe is one of the themes of the European FutureMARES project, which is studying the effects of climate change on marine ecosystems. In this context, it is important to anticipate the inevitable effects of climate change when it comes to restoring a threatened species in European seas. In this report, we try to find out where the oysters that respond best to heat waves, which are more frequent with global warming, come from.

In this report, we are only interested in the reactions and signals that would indicate a sign of stress in oysters, i.e. the heartbeat and the opening of the valves. To do this, we measured the heartbeat of 10 oysters from 3 different origins (Croatia, Norway, and the Netherlands) by subjecting them to heat waves for 6 weeks. We are also studying the opening of the valves of 9 different oysters from the same 3 origins over a period of 6 weeks by subjecting them to different temperatures.

The data collected by the sensors was imported and analyzed in R in order to establish a statistical model showing significant differences in the heartbeat measurements of the different oysters according to temperature and origin.

The results are not conclusive, and the model does not validate the initial hypotheses. However, it is possible that the data set is too poor to produce usable results, and the experiments would have to be repeated in order to be able to draw conclusions about validate or not the initial hypotheses.

1 - INTRODUCTION

Climate change threatens marine ecosystems and biodiversity by influencing various physic-chemical environmental parameters such as temperature, pH, salinity etc. (Rizzi et al., 2016) It was from this observation that the European FutureMARES project was born. This research project brings together universities, laboratories, and NGOs (Non-Governmental Organization), "to provide socially and economically viable actions and strategies in support of nature-based solutions and nature-inclusive harvesting for climate change adaptation and mitigation" (*FutureMARES*)
This project takes place in different regions around the globe, in Europe and in America (Figure 1).

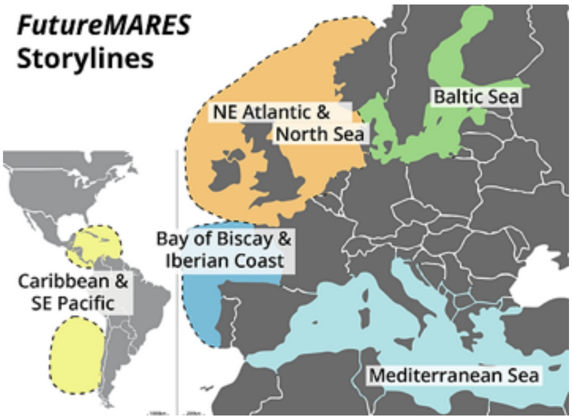


Figure 1: *Map of the areas covered by the FutureMARES research project* (*FutureMARES*)

Wageningen Marine Research (WMR) in Yerseke (Zeeland, Netherlands) is a research department of Wageningen University in the Netherlands, specialized in shellfish and coastal ecosystems, which is taking part in the project by studying the effective restoration of shellfish reefs. In fact, oyster reef is one of the most endangered habitat on earth, with a 85% estimated global loss, even greater than the losses reported for other important habitats including coral reefs, mangroves, and seagrasses (Airoldi, Laura et al., 2009)
Oysters are considered as ecosystem engineers because they form structures that influence the environment around them in ways that are beneficial to other species (Cobacho et al., 2020). An oyster reef is a self-sustaining habitat for marine biodiversity and functions as a nursery for other species (Walles et al., 2015). It provides different ecosystem services such as absorbing wave energy and reducing erosion from boat wakes, sea level rise and storms (Chowdhury, 2019). It also has a beneficial impact to commercial fish stocks (Gilby et al., 2018).

Restoration projects of oyster reefs are numerous and take place in various places on the planet, especially in the North Sea and the Northeast part of the Atlantic Ocean. In the Dutch North Sea, since 2018, pilot projects have started to restore European flat oysters (*Ostrea edulis*) at suitable locations, such as offshore windfarms or natural reefs, which are protected from bottom trawling (Bos et al., 2023).
These projects need to take global change into account for restoration to be truly effective and sustainable. Future increases in sea temperature are likely to cause a shrinkage in the spatial extent of most suitable oyster sites (Aura et al., 2017). Temperature is the main driver of climate change and is being accompanied by increasingly frequent marine heatwaves. Recent marine

heatwaves revealed the high vulnerability of marine ecosystems to such extreme climate event (Frölicher et al., 2018).
The restoration projects must also deal with the parasite *Bonamia ostrae,* introduced in the 1980's and causing catastrophic mortalities in the population of *O.edulis*. Growing efforts to restore flat oysters in Europe should avoid further introduction and spread of this parasite (Sas et al., 2020). For this, oyster from Norway can be introduced in suitable oyster sites in the North Dutch Sea because they are Bonamia-free (Mortensen et al., 2020).

The research project I joined during this internship is part of FutureMARES and explores solutions to manage a sustainable restoration of flat oysters considering climate change and disease. The project is titled "Local adaptation potential of key coastal species" (Arenas et al., 2023). One of the studies was carried out at WMR and concerned the effect of marine heat waves on the physiological responses of the European flat oyster *Ostrea edulis*: regional differences (Chapter 1 in Arenas et al., 2023). The main objective of this study is "to understand thermal impacts on oyster populations and evaluate potential regional variation in response to these extreme thermal events." The oysters used in the experiments originated from 3 different geographical area: the Dutch North Sea, the Norwegian North Sea, and the Mediterranean Sea in order to gain information about the adaptative capacity, vulnerability, and potential management strategies of this species under changing climatic conditions. Different metabolic aspects were monitored during experiments such as the respiration rate, the clearance rate, the growth, the valve gape, and the heart rate. All those parameters will be used to determine the effect of marine heatwaves on physiological responses, that can indicate stress for the European flat oyster.

My work in this study was to extract the data from the heart rate and the valve gape experiments (for details about these experiments, see 2 - Material and Methods part) to make them usable with the programing language R and to realize a statistical analysis with the data.

Our hypothesis is that oysters from Croatia should show fewer signs of stress when exposed to heat waves since they come from a southern location. Norwegian oysters, coming from a northern location should show the greatest signs of stress. The oysters from the North Sea area should show stress levels in between the two other origins.
We expect to observe a significant difference of the heart rate between the Croatian oysters and the Norwegian oysters. Norwegian oysters should have higher value values of heart rate and should be closed for a longer time than the Croatian oysters, especially during heatwave periods.

2- MATERIAL AND METHODS

2.1 - Experimental set-up

First, I will describe the material and methods used for the heart rate experiment and the valve gape experiment. Next, I will describe how I extracted the data and how I selected them for analysis.

European flat oysters were exposed to heat stress during a six-week laboratory experiment and their physiological response were measured. Oysters was separated by origin (Dutch, Norwegian and Croatian) in 10L containers, and those containers were placed in larger tanks of 500L each. Each containers housed 5 oysters and 3 containers per origin were placed in each of the three tanks. Each tank is corresponding to a temperature regime, maintained au Bain-Marie with heaters and thermostat: constant 20°C, constant 25°C and alternated temperature between 25°C and 20°C each week (Heatwave). Containers and tanks were filled with 1µm filtered seawater and aeration was provided to each container. Oysters were continuously fed at a rate of 0.77 mL min^{-1} with live *Rhodomonas salina* algae.

Oysters from Norway were collected in Hafrsfjord, oysters from the Netherlands were collected in Oosterschelde estuary and flat oysters from Croatia were collected in Maliston Bay.
After being brought in the laboratory, an acclimatation period started. Temperature was raised by 1 degree each day from the collection temperature to 20°C. Sea water in the tanks was replaced three times a week by new water acclimated to the temperature regime.
This is the experimental set up for every physiological parameter measurement, but I will only describe the monitoring of the survival rate, the condition index calculation, the heartbeat, and the valve gape of the oysters.

2.2 - Survival and condition index

Survival was determined by counting the live oysters at each cleaning event (3 times per week). The dead oysters were replaced by new oysters from the same origin previously acclimatized.
At the end of the experiment the condition index was calculated with the following equation (Equation 1):

1. $$\frac{AFDW}{DW_{shell}} * 100$$

Ash-Free Dry Weight (AFDW = DW – AW) was determined by removing the flesh from the shell and dried at 70°C until a constant weight to determine dry weight (DW) in g. Then, the flesh was combusted at 540°C until weight constancy to determine the ash-weight (AW) in g.
The dry weight of the shell (DW_{shell}) in g was determined after drying at 70°C until weight constancy.

2.3 - Heart rate measurement and analysis

The heart rate was measured for only 10 oysters in total because the laboratory had access to only one sensor with 10 plugs. There were three oysters from Norway, two from Netherlands and two from Croatia in the heatwave treatment. Three Norwegian oysters, in the 25°C treatment, were also logged with the sensor.
The sensor is the PULSE V2 sensor developed by Electric Blue, which is a non-profit technology transfer startup and spin-off company from the FutureMARES partner CIBIO-Biopolis.

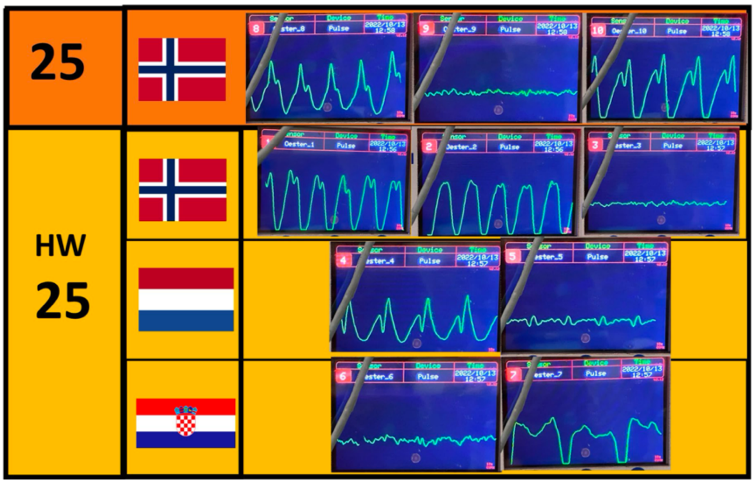


Figure 2: *Snapshot of recordings with PULSE V2 sensors attached to three oysters from Norway, two oysters from The Netherlands and two from Croatia in the heat wave treatment and to three oysters from Norway in the 25 ºC treatment.*

It uses infrared (IR) sensors glued to the exoskeleton of the animals above their heart. Changes in the shape or volume of the circulatory structures during a heart contraction, or heartbeat, change the amount of IR light reflected from the animal's internal anatomy back to each IR detector (ELECTRICBLUE).
These changes in reflected IR light, transduced to changes in electrical current, are then electronically amplified, filtered, and the signal is logged into a memory card by a microprocessor.
Data were recorded continuously every 0.05 seconds throughout the experimental period for each of the ten oysters. Data were placed in a .csv file every hour, so in total there are 984 files for the 6-week experiment.
A R script was made to import the data from these files to R and to create a data frame for each 1-hour file.
The amount of data obtained from the experiment had to be reduced in order to carry out the beat per minute (BPM) and amplitude (Amp) calculations over time for the 10 oysters so the data from the first minute of every hour was selected to effectuate calculations.
Another R script was made to calculate the BPM and the average Amp of each oyster through the time. The rolling average of the signal, with a moving window size of 601 values, was used to count every intersection point between the signal curve smoothed by the rolling average and the straight line from the equation y = 0.
The period of the signal from an oyster during a one-minute interval of time was determined with this equation (Equation 2):

$$2.\quad period = \frac{\frac{Number\ of\ intersection\ point}{2}}{Interval\ of\ time\ between\ the\ first\ and\ the\ last\ intersection\ point}$$

Then, the BPM was obtained by dividing 60 by the period.
The number of intersection point is divided by 2, because two intersection points count for one pulse. The following figure shows the way the period was calculated for every time interval (Figure2):

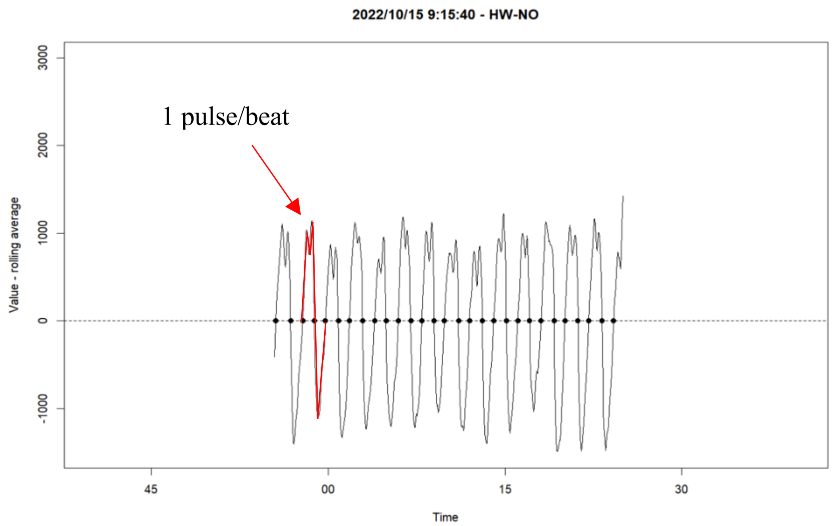


Figure 3: *Rolling average of the signal from a Norwegian oyster with a heatwave treatment during a 30 second interval*

On this figure, there is 29 intersection points so there are 14 pulses during an interval of 30 seconds. According to the Equation 2, the period is 0.48 s^{-1} so the BPM in this case is 124 beats per minute.

The average amplitude of each one-minute interval was calculated by finding the maximum or the minimum value of the signal in each interval delimited by the intersection point. Averages of every maximum and minimum were calculated and the difference between the two averages was the Amp average of the signal during this minute.

The selection of one minute of data every hour permitted to reduce the running time of the code to around ten hours. At the end, the final data frame contained for every oyster: the BPM and Amp measurement for every hour of the six weeks, the temperature through the time, the origin of the oyster, the oyster number (1 to 10) and the condition index.

2.4 - Valve gape measurement and analysis

For this experiment, one oyster of each container was randomly selected to be used for the valve gape measurement. This resulted in three oysters for each origin with the three different temperature treatments. The valve gape was determined using the Biophys sensor developed by the Royal Netherlands Institute for Sea Research (NIOZ) (Figure 4).

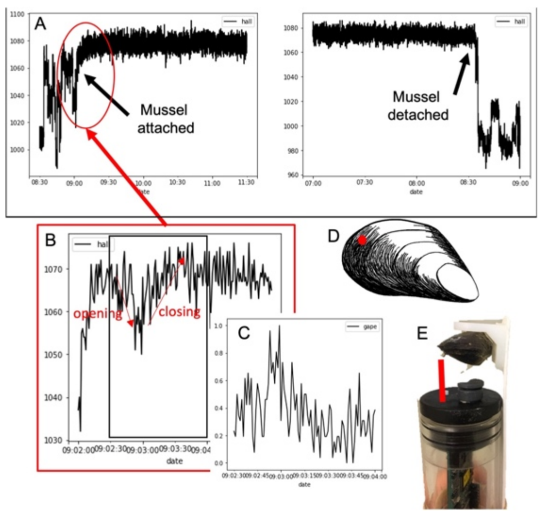


Figure 4: *Processing methods for the gaping data. (A) identification of the moment where mussel was attached to the sensor, jumping occurs because of the presence of the magnet attached to the shell and located directly above the sensor (D, E). Red line in (E) shows the distance being measured to quantify gaping. (B) and (C) represent data series processing descripted in methodology. (Bertolini et al., 2021)*

Small magnets were glued with a waterproof glue on the dry shell of the live oysters. Then, they were put in 7°C in a fridge during the night to allow the glue to harden. The "hall function" was programmed to measure every 10 seconds. Oysters were then attached to the sensor with a clamp, and the clamps were then secured to the wall of their container.

This sensor can measure different parameters such as the temperature but only the "hall function" was used to preserve battery. An external thermometer was used with the same sampling frequency as the sensor.
The "hall function" is the measurement of the distance between a magnet and the sensor (Bertolini et al., 2022).

At the end, the data were downloaded and converted from .BIO file to .txt file. They were imported and plotted in R to determine the direction of hall number changes (dependent upon magnet direction, Figure 4B). The valve gape at each sample was calculated as (Equation 3):

$$3.\quad Gape = \frac{|hall - \max(hall)|}{\max(hall) - \min(hall)}$$

(if the hall number decreased with opening)

Or (Equation 4):

$$4.\quad Gape = \frac{|hall - \min(hall)|}{\max(hall) - \min(hall)}$$

(if the hall number increased with opening) (Figure 4C).
The data set was then averaged over an hour to remove potential noise and obtain a manageable time series.

2.5 - Descriptive and statistical analysis

To analyze the processed data from the heartbeat experiment obtained using the method described above, the dataset was imported, time was converted into year/month/day - hour/minute/second with the package *lubridate* and into numeric values to be used as a quantitative variable in statistical models.

Origins of the oysters were defined as a factor and temperatures were divided into 2 classes, 'cold' and 'warm', with cold corresponding to temperatures below 22.5°C. Temperature has not continuous values because most of the time, it is around 20°C and 25°C (cold and warm). Those two classes were defined as factor to be used as qualitative variables.

The data have been scaled to reduce the variability between individuals and to mitigate the effect of extreme values, which are mostly outliers due to disturbances or experimental problems. The scaling method used was a Min-Max transform between 0 and 1 (Equation 5):

$$5.\ data\$BPM\_scaled \leftarrow \frac{data\$BPM - \min(data\$BPM)}{\max(data\$\text{BPM}) - \min(data\$BPM)}$$

Another dataset has been created from the original dataset, regrouping only the oysters treated with the HW treatment.

R was used to estimate the mean and median BPM of each oyster during the 6-week experiment, the mean BPM during the hot and cold treatments, and the mean BPM during the warm and cold treatments.
In a demarch of reducing the amount of data, BPM during night period and day period was compared because there were perturbations during the day comparing to the night (light, noise, human presence in the room, maintenance). The results were plotted in a boxplot graph with the package ggplot2.

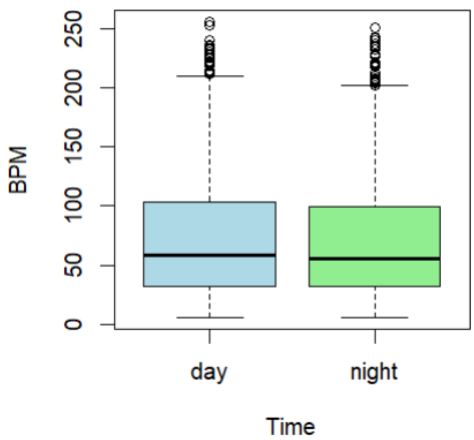


Figure 5: *Boxplot representing the BPM depending on the time of the day (day or night)*

No significative differences between the BPM during the day and during the night, as the one-factorial analysis of variance shows it (Figure 6):

```
> res <- aov(BPM_scaled2 ~ periode, data = data)
> summary(res)
              Df Sum Sq Mean Sq F value Pr(>F)
periode        1    0.1 0.06884   2.125  0.145
Residuals   9838  318.7 0.03239
>
```

Figure 6: *Anova of the BPM and the day period and results of the analysis*

For the statistical analyses, different models were applied to this dataset. The results will be presented in Section 3 - Results and the reasoning behind the choice of these models to interpret the data will be developed in Section 4 - Discussion. In this section, only the different models used, and their formulas will be presented.
Firstly, linear models for each oyster were elaborated to analyze the data. It calculates the effect of the time on the BPM.

linear_model1 ← lm (BPM_scaled ~ DateTime, data = dataOyster1)
.
.
.
.
linear_model10 ← lm (BPM_scaled ~ DateTime, data = dataOyster10)

Then, linear models were applied to every oyster together and effect of the temperature category was added to the model, the data are now regrouping every oyster of the HW treatment:

linear_model ← lm (BPM_scaled ~ DateTime + cat_temp, data = dataHW)
Linear models will be not sufficient to describe well the data, so a General Additive Model (GAM) was chosen to analyze this dataset. This is the first formula tried with the data, it will receive a lot of improvement to fit better with the data:

model_gam ← gam (BPM_scaled ~ s(num_time) + cat_temp,
data = datafiltered, method = "REML")

Results used for the analysis come from the formula below, elaborated by studying the results, the summaries, the plots, and Akaike Criterion (AIC) of the improved models. The effects on the BPM of the temperature, the origin of the oysters, the condition index and the average amplitude of the signal was progressively added to the model.

model <- gam (BPM_scaled2 ~ s (num_time, k = 20) + s (Amplitude_scaled) + cat_temp + Condition.index + Origin, data = datafiltered)

For the valve gape analysis only, a descriptive analysis was processed, due to a lack of time. Descriptive statistics (medians, means) were estimated on gaping, time spent open or closed. Several graphs have been produced to give an initial idea of the behaviour of oysters in response to changes in temperature, and to try to identify avenues for further analysis.

3 - RESULTS

3.1 - Preliminary results

Looking at a graph of the BPM values of a Dutch oyster with HW treatment gives a better idea of the data and the models that could be applied to it.
A quick glance at the graph, shows a non-linear evolution of the BPM over the time, so a linear model may not be the most appropriate model for this data set. (Figure 7):

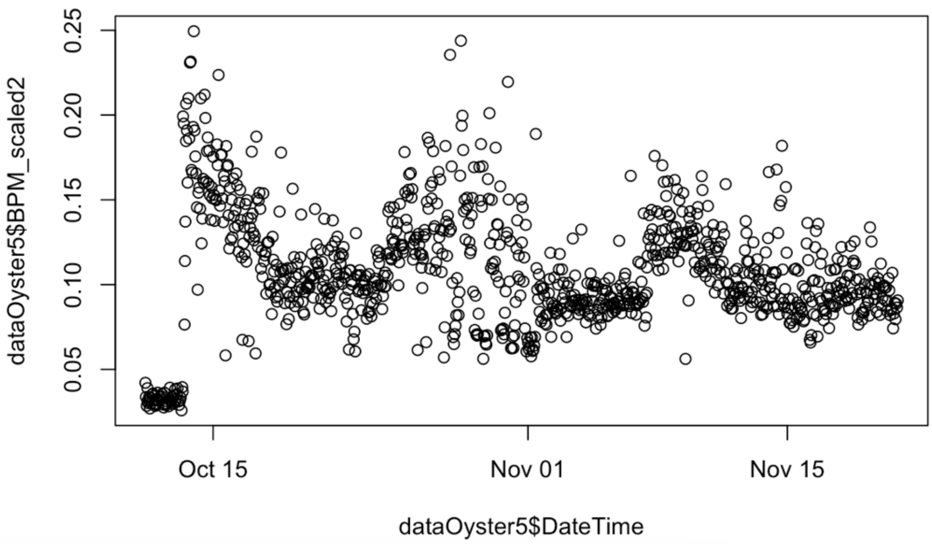


Figure 7: *Plot of the BPM over time for the oyster n°5 (HW-NL)*

The BPM seems to oscillate over the time. Outliers can be observed before 15 October 2022, surely corresponding to an adaptation period for the oysters or a period for setting up the sensor. This part of the data was deleted so the dataset starts now from 00:00 the 16th of October 2022.

Results of the mean of the BPM for each origin and for each temperature category shows that the BPM is higher for the warm category, but there is a very small difference between the cold and the warm categories (less than 3.6 %). Netherlands has the lowest BPM mean and Croatia the highest. (Figure 8)

```
> print(means_BPM_Origin)
       Origin      BPM
1     Croatia 82.86613
2 Netherlands 65.99003
3      Norway 70.91700
```

```
> print(means_BPM)
  cat_temp      BPM
1     cold 70.72533
2     warm 73.37938
```

Figure 8: *Mean of the BPM for each origin and each temperature categories obtained with R*

The first linear model on the filtered data shows a significative effect of the time on the BPM with a very low but negative coefficient so the BPM of the oysters decrease significatively during the 6 weeks. Temperature has not a significative effect on the BPM (Figure 7):

```
Call:
lm(formula = BPM_scaled2 ~ DateTime + cat_temp, data = datafiltered)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.29194 -0.14431 -0.04407  0.11496  0.73763 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.529e+01  5.007e+00   9.045   <2e-16 ***
DateTime     -2.700e-08  3.003e-09  -8.993   <2e-16 ***
cat_tempwarm  8.012e-03  5.502e-03   1.456    0.145    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1771 on 4332 degrees of freedom
Multiple R-squared:  0.01916,	Adjusted R-squared:  0.01871 
F-statistic: 42.31 on 2 and 4332 DF,  p-value: < 2.2e-16
```

Figure 8: *Summary of the linear model (lm(BPM_scaled2 ~ DateTime + cat_temp , data = datafiltered)) obtained with R*

The multiple R-squared is 0.019, which means that this model explains around 1.9% of the total variation in the BPM. The F-statistic is 42.31, with a very low p-value. This indicates that the model is significantly better than the null model (model without predictors), although the contribution of individual predictors may be small.

Another linear model with more predictors gave better results, 22% of the total variation in the BPM is explained. Predictors such as the origin and the condition index of the oysters have a significatively effect on the BPM (Figure 9):

```
Call:
lm(formula = BPM_scaled2 ~ num_time + cat_temp + Origin + Condition.index, 
    data = datafiltered)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.33934 -0.11615 -0.03220  0.08076  0.73516 

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)    
(Intercept)        5.182e+01  3.828e+00  13.535  < 2e-16 ***
num_time          -3.132e-08  2.298e-09 -13.632  < 2e-16 ***
cat_tempwarm       1.186e-02  4.108e-03   2.888  0.00389 ** 
OriginNetherlands -7.405e-02  5.358e-03 -13.820  < 2e-16 ***
OriginNorway       8.500e-02  5.882e-03  14.452  < 2e-16 ***
Condition.index    4.117e-01  1.054e-02  39.049  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1577 on 6063 degrees of freedom
Multiple R-squared:  0.2272,	Adjusted R-squared:  0.2266 
F-statistic: 356.6 on 5 and 6063 DF,  p-value: < 2.2e-16
```

Figure 9: *Summary of the linear model (lm (BPM_scaled ~ num_time + cat_temp + Origin + Condition.index, data = datafiltered)) obtained with R*

The coefficient for "OriginNorway" is estimated at 0.467. This means that, on average, oysters from Norway have a BPM value 0.467 higher than the reference ("OriginCroatia") when all other variables in the model are held constant. This suggests that oysters from Norway tend to have higher BPM than the other origins included in the model. The temperature is almost significatively effective on the BPM, but the p value is still bigger than the 0.05 trust value.
Comparing the two models presented with the AIC method, the first (simpler) model has a lower AIC than the second, so it is preferable to choose the first model. (Figure 10):

```
> AIC(linear_model,linear_model_updated)
                     df       AIC
linear_model          4 -3693.199
linear_model_updated  7 -5186.323
```

Figure 10: *AIC of the linear models presented above*

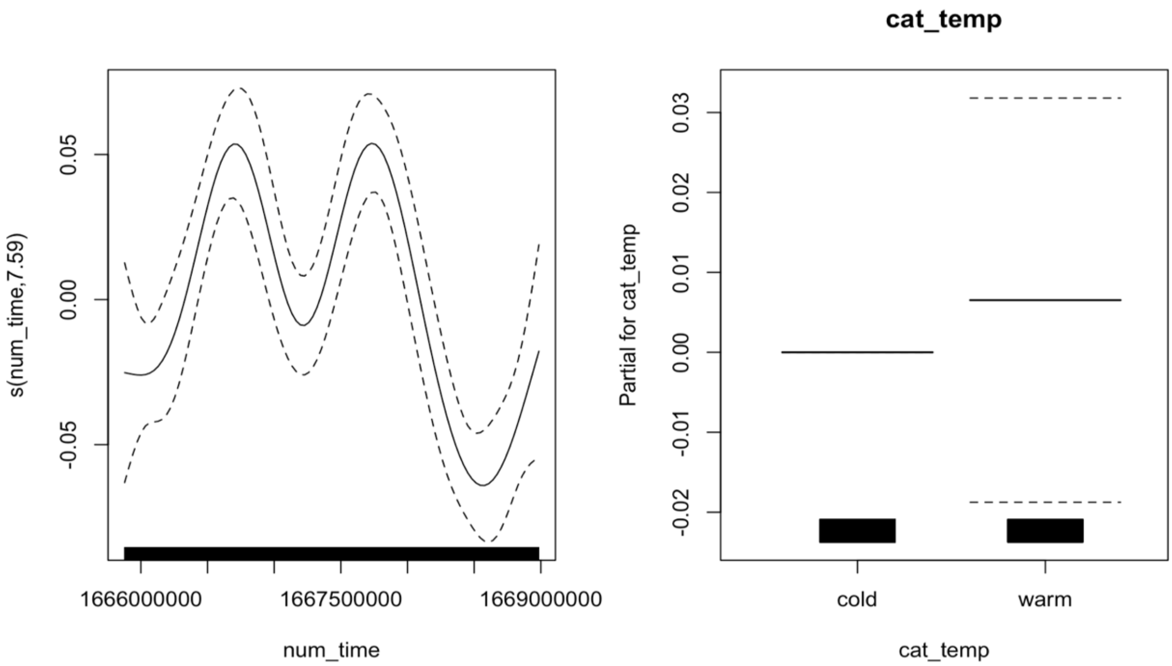
The variables added in the *linear_model_updated* improved significatively the fit of the model. But there is a limit to the linear model, which does not fit the data well enough to explain it in a globality, even with many explanatory and significative variables. The added variables have a significant effect on the BPM, but the model is still not well adapted to the data due to the global shape of the data and the non-linear correlation existing between some of the variables.

3.2 - Generalized Additive Models

The results come from a generalized additive model whose parameters are time and the temperature category qualitative variable. The time variable is a non-linear qualitative variable, as shown by the command: *summary(model_gam)$s.table* function.
In fact, the effective degrees of freedom (EDF) of the time variable are 7.5, i.e. quite far from 1 (Figure 11). It is therefore smoothed in the model as specified in the formula: s(num_time).

```
> summary(model_gam)$s.table
                 edf   Ref.df        F p-value
s(num_time) 7.594723 8.537807 15.55492       0
```

Figure 11: *EDF of the variable s(num_time), variable of the time but converted in numeric values to be an argument in the GAM*



```
> summary(model_gam)

Family: gaussian
Link function: identity

Formula:
BPM_scaled2 ~ s(num_time) + cat_temp

Parametric coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.266920   0.006330  42.166   <2e-16 ***
cat_tempwarm 0.006519   0.012642   0.516    0.606
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
              edf Ref.df     F p-value
s(num_time) 7.595  8.538 15.55  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.045   Deviance explained = 4.77%
GCV = 0.03134  Scale est. = 0.031241  n = 3035
```

Figure 12: *Results from the GAM, plots of the smoothed term time and temperature category and summary of the model*

The num_time term is represented using a spline function. The F-statistic is 15.55 with a very low p-value, indicating that this term is highly significant. The graph shows oscillations in the variation of the BPM as a function of time. During certain periods corresponding to approximately one week (according to the graduations of the abscissa axis, containing 6 portions) the BPM is significantly higher than during the following week. This corresponds to the temperature variation intervals of the HW treatment.
However, the cat_temp variable had no significant effect on BPM in this model, as shown by the p-value of 0.606 >> 0.05. This model explains only 4.5% of the variance of the response variable, so the results do not really explain all the phenomena responsible for the variation in BPM.
An AIC of the previous models gives an idea of the relevance of the GAM compared to linear models. (Figure 12):

```
> model_gam <- gam(BPM_scaled2 ~ s(num_time) + cat_temp , data = datafiltered)
> AIC(model_gam, linear_model_updated, linear_model)
                           df       AIC
model_gam            11.28493 -3947.100
linear_model_updated  7.00000 -5186.323
linear_model          4.00000 -3693.199
```

Figure 12: *AIC of the linear model, linear model updated (with more explicative variables) and of the GAM*

The AIC is better comparing with the linear_model but the lack of explicative variables makes this model less good than the *linear_model_updated.*
The base dimension k defines the number of base functions used to create a smooth function. By default, the model uses 10. If this value is doubled, so k = 20, the model is slightly improved, the variance explained is 6.59% and the AIC decreases slightly.

Updated GAM is inspired by the updated linear model, origin and condition index is added as explicative variables. Results of this model are compared with the previous one. This new model explains 26.2% of the deviance and the AIC is lower than the AIC of the first GAM. The added variables have both a significative effect on the BPM with a strong p-value (Figure 13):

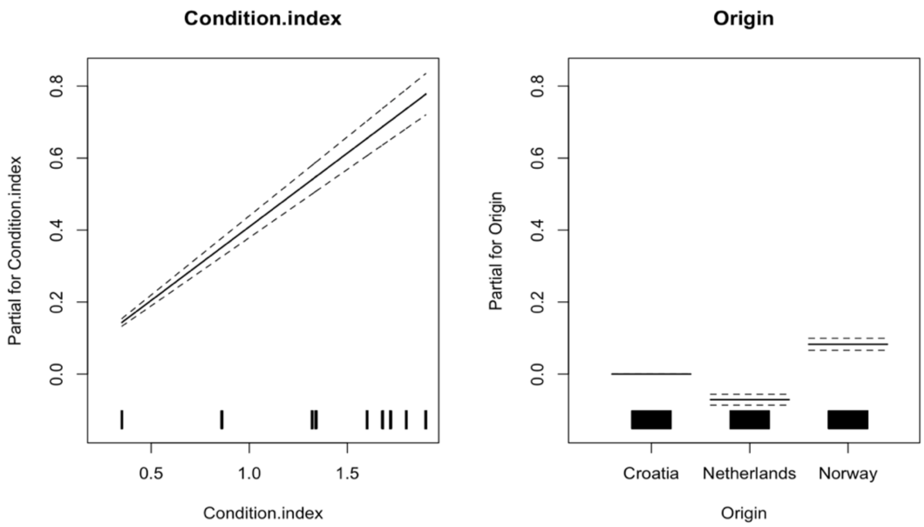
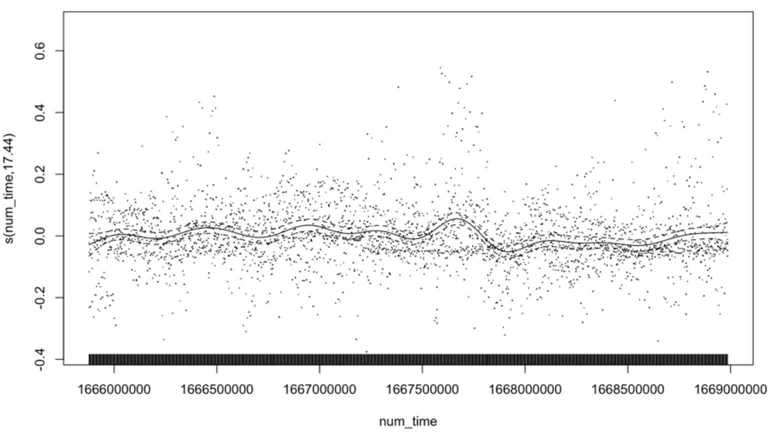


Figure 13: *Estimated effect on the BPM of the variables Condition.index and Origin*

The coefficient of the condition index is 0.409 so as the condition index increases, so does the BPM. An oyster with the origin “Norway” has a significatively higher BPM than oysters with the origin “Croatia” or “Netherlands”. The temperature still does not have a significative effect on the BPM variation in this model.

The last tested GAM takes the amplitude as non-linear variable. This variable is smoothed in the formula and added with the other variables previously there. This model fits better with the data, explaining 67.3% of them with the lowest AIC between all the tested models. (Figure 14):




```
Family: gaussian
Link function: identity

Formula:
BPM_scaled2 ~ s(num_time, k = 20) + s(Amplitude_scaled, k = 10) +
    cat_temp + Condition.index + Origin

Parametric coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        0.125278   0.014367   8.720  < 2e-16 ***
cat_tempwarm      -0.031934   0.007890  -4.048 5.24e-05 ***
Condition.index    0.099759   0.007946  12.555  < 2e-16 ***
OriginNetherlands  0.004876   0.003652   1.335   0.1819
OriginNorway      -0.007824   0.004076  -1.920   0.0549 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                     edf Ref.df      F p-value
s(num_time)        18.52 18.967  16.24  <2e-16 ***
s(Amplitude_scaled)  8.31  8.869 841.35  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.672   Deviance explained = 67.3%
GCV = 0.010616  Scale est. = 0.01056   n = 6069
```

Figure 14: *Estimated variation of the BPM over the time with residuals and summary of the GAM*

On the graph, the estimated curve has less amplitude in the variation of the BPM over the week. Residuals are near the curve and seem to follow its global shape.
The smoothed variables are highly significant in the model, the addition of the amplitude as an explicative variable strongly improved the percentage of the variation explained by the model. The temperature is now significative, but the coefficient is now negative, the category “warm” has a significative lower BPM than the category “cold”. Origin is not significative variables anymore.

3.3 - Valve gape descriptive results

The graphs plotted are corresponding to the data of the Dutch oysters with the three different treatments. The temperature is also on the graph to easily observe if there is a link between the temperature and the gaping. On the ordinate axis, the values go from 0 to 1, 1 representing the maximum gape between the valve (opened) and 0 the minimum gape between the valve (closed). (Figure 15):

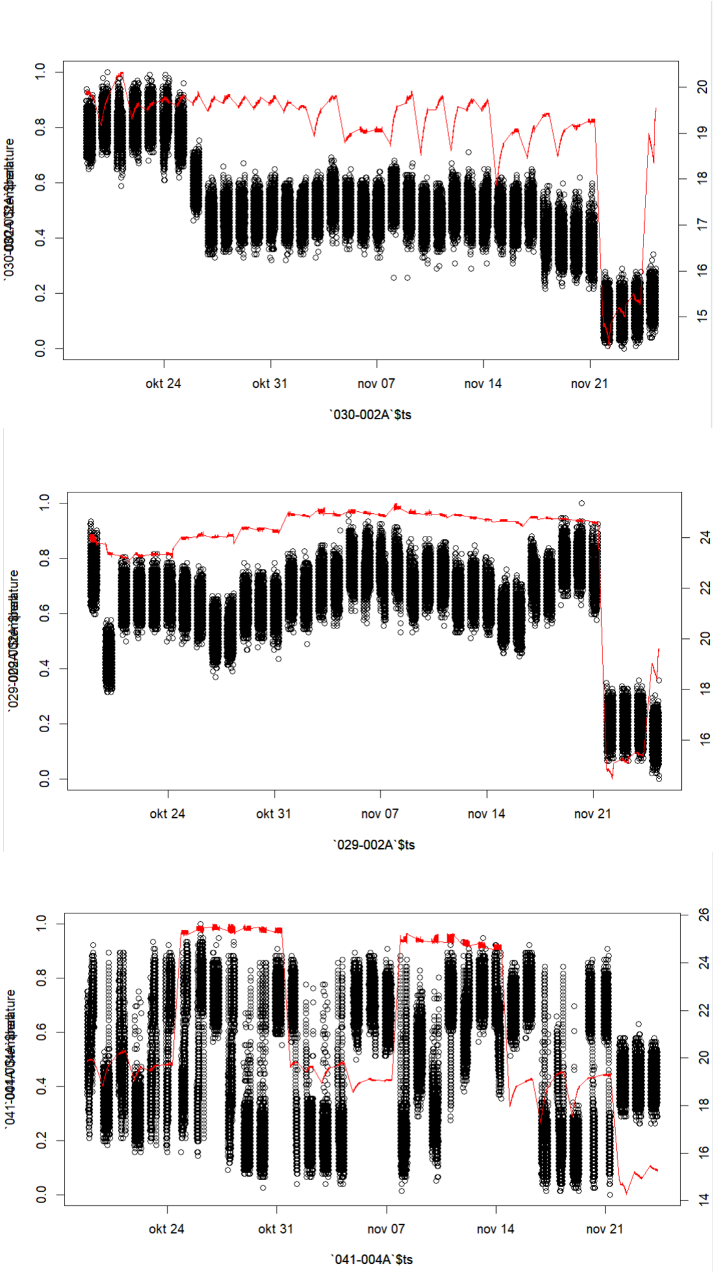


Figure 15: *Valve gape over the time for Dutch oysters at 20°C (1st graph), at 25°C (2nd graph) and with the HW treatment (3rd graph)*

The HW oyster gaping variates a lot over the time, the oyster seemed to be very perturbated. By looking at the graphs, no clear relation between the temperature and the gaping can be observed, a precise statistical analysis is required to complete those results.
The meaning gaping of those three oysters was used to briefly compare them. (Figure 16):

```
> mean_20 <- mean(`030-002A`$hall)
> mean_25 <- mean(`029-002A`$hall)
> mean_HW <- mean(`041-004A`$hall)
> print(c(mean_20,mean_25,mean_HW))
[1] 0.4966504 0.6230046 0.4946754
```

Figure 16: *Mean of the gaping of the Dutch oysters*

The 25°C oyster is wider open than the others with a meaning gaping of 0.62.
A decrease of the temperature appears during the last week, this was not in the protocol but the reaction of every oyster to this brutal decrease in temperature was to close their valves.

4 - DISCUSSION

4.1 - Selection of data and models used

Different methods were used to reduce the amount of data of the initial dataset. Firstly, the three Norwegian oyster at 25°C were removed from the data. These oysters are the only 3 with this treatment, so it is preferable to analyze them separately. In fact, given the very small number of individuals, the results are very much influenced by the presence of these oysters, which are not under the same experimental conditions. In addition, there were no oysters with a constant temperature of 20°C, which would have enabled a longer-term comparison of the effects of temperature on BPM (compared with the HW treatment in which the temperature was changed every week). The final dataset used for every model contained data from the first 7 oysters, which were the ones exposed to the HW treatment, and the data at the start from the 11 October 2022 to the 15 October 2022 has been also removed.

The choice of using a linear model was made as part of a data discovery approach, observing preliminary results before looking for other models that might be more appropriate. The linear model was also used to determine whether the dataset was correctly constructed and could therefore be used for further statistical analysis. It was also used to establish a benchmark: knowing full well that this model was not the most appropriate, new models were then compared with it to see if they were more appropriate to the data and more relevant than the linear model.

A generalized additive model (GAM) is a type of statistical model used when the data present complex and non-linear relationships between the independent variables and the dependent variable, when the residuals are not normally distributed, when data include categorical variables and there are many explicative variables. Furthermore, this type of model is well adapted to modelling spatial or temporal data (Wood, 2017).

The normal distribution of the residuals can be checked with the Kolmogorov-Smirnov test. If the p-value is < 0.05, this suggests that the residuals do not follow a normal distribution. (Figure 17):

```
> resid_gam <- residuals(model)
> ks.test(resid_gam, "pnorm")

        Asymptotic one-sample Kolmogorov-Smirnov test

data:  resid_gam
D = 0.39729, p-value < 2.2e-16
alternative hypothesis: two-sided
```

Figure 17: *Kolmogorov-Smirnov test for the GAM*

The p-value is well below 0.05 so the residuals distribution is not normal. The relation between the dependent variable and the independent variables can be visualized using the function *geom_smooth ()* from the package *ggplot2* (Figure 18).

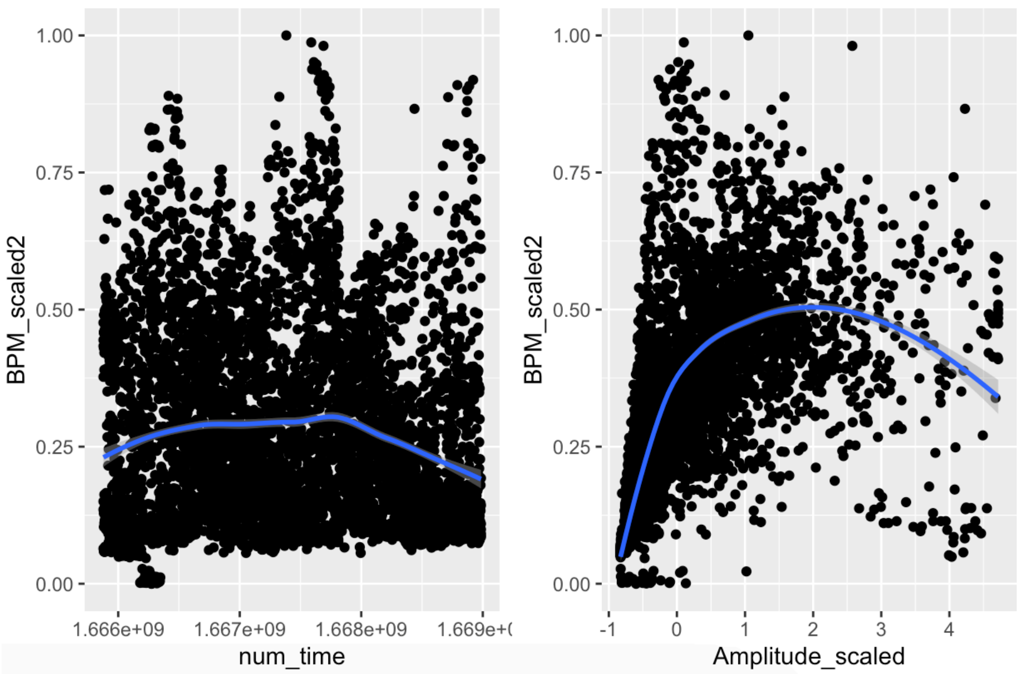


Figure 18: *Relation between the dependent variable BPM_scaled2 and the independent variables num_time and Amplitude_scaled*

These two graphs show no linear relationship between the variables. The dataset and the variables therefore fulfil the main criteria for the use of a GAM. This type of model with optimum parameters explains a maximum of 70% of the variation of the BPM. No other model tested fit the data more than this percentage.

BPM measurements are very much influenced by individuality, with the condition index being a parameter with a very significant effect on BPM. The small number of individuals greatly increases the variability of the data, there are large differences between individuals which makes the data difficult to analyze by any model. Maybe adding a random effect from the oyster in question would be useful for analyzing the data.

4.2 – Expected results and improvement of the experiment

The initial idea of the experiment was to use the bpm as an indicator of stress in the oysters. By subjecting these oysters to different levels of temperature (at abnormally high levels such as 25°C) with a frequency of change of one week, the expected response was an increase in the stress of the oysters due to the significant thermal variations in the environment, but conversely, as shown by the results of the final model, the data show that the BPM tends to be significantly lower when the temperature is higher (as shown by the negative coefficient associated with the cat_temp variable).

The results for the origin variable, which was the second variable of interest in the hypotheses formulated at the outset, are not very significant in the final model. The differences between the BPM values of the different origins are too minimal, the number of individuals being too small, the random effect due to individuality is more important on the BPM variations than the effect of the origin of the oysters.

The amplitude contributed a lot to improve the model, there is probably a correlation between the variable BPM and the variable Amplitude, however it is not possible to interpret the results from this variable because the values are extremely different between individuals. The sensor's signal intensity is influenced by where the oyster body is in its shell. The correlation with BPM could be explained by the fact that the amplitude is related to the quality of the received signal, and therefore the calculation of BPM on a better-quality signal is more accurate and more likely.

None of the hypothesis were verified by this statistical analysis. It may be necessary to repeat this experiment with more individuals to minimize the effect of individualities and measurement errors that may hide a potential trend in the evolution of BPM due to temperature variations. Some oysters presented the expected results, including Oyster 5 (HW, NL), whose BPM increases as the temperature increases and decreases as the temperature decreases.

To decrease the effect of the condition index and remove the abnormal data from the beginning, it would be necessary to increase the length of the period of adaptation of the oysters.

Finally, to improve the data itself, a better algorithm for calculating BPM from recorded signals would be needed. Indeed, the received signals are difficult to read so may be that errors are present in the data because of poor signal analysis.

CONCLUSION

The aim of this project and this mission was to show that heat waves due to global warming are a future obstacle to the restoration of flat oysters in Europe. To anticipate this phenomenon, it would be necessary to use oysters of the same species from warmer seas, such as oysters from Croatia. The results do not support this hypothesis, as the BPM of the Norwegian oysters throughout the experiment was almost significantly lower than that of the Croatian oysters. However, it is impossible to conclude on these results, as the experiment presents numerous experimental biases and requires a new version containing more individuals. We also need to find a more accurate way of calculating BPM from the signals recorded by the sensors. These two changes should greatly improve the quality of the data and therefore of the resulting statistical model.

REFERENCES

Airoldi, Laura, Beck M. W., R. Brumbaugh, A. Carranza, & L. Coen. (2009). *Shellfish Reefs at Risk: A Global Analysis of Problems and Solutions*. https://cris.unibo.it/handle/11585/77343

Aura, C. M., Musa, S., Osore, M. K., Kimani, E., Alati, V. M., Wambiji, N., Maina, G. W., & Charo-Karisa, H. (2017). Quantification of climate change implications for water-based management: A case study of oyster suitability sites occurrence model along the Kenya coast. *Journal of Marine Systems*, *165*, 27–35. https://doi.org/10.1016/j.jmarsys.2016.09.007

Bertolini, C., Capelle, J., Royer, E., Milan, M., Witbaard, R., Bouma, T. J., & Pastres, R. (2022). Using a clustering algorithm to identify patterns of valve-gaping behaviour in mussels reared under different environmental conditions. *Ecological Informatics*, *69*, 101659. https://doi.org/10.1016/j.ecoinf.2022.101659

Bertolini, C., Rubinetti, S., Umgiesser, G., Witbaard, R., Bouma, T. J., Rubino, A., & Pastres, R. (2021). How to cope in heterogeneous coastal environments: Spatio-temporally endogenous circadian rhythm of valve gaping by mussels. *Science of The Total Environment*, *768*, 145085. https://doi.org/10.1016/j.scitotenv.2021.145085

Bos, O. G., Duarte-Pedrosa, S., Didderen, K., Bergsma, J. H., Heye, S., & Kamermans, P. (2023). Performance of European oysters (Ostrea edulis L.) in the Dutch North Sea, across five restoration pilots. *Frontiers in Marine Science*, *10*, 1233744. https://doi.org/10.3389/fmars.2023.1233744

Chowdhury, M. S. N. (2019). *Ecological engineering with oysters for coastal resilience: Habitat suitability, bioenergetics, and ecosystem services*. https://doi.org/10.18174/466205

Cobacho, S. P., Wanke, S., Konstantinou, Z., & El Serafy, G. (2020). Impacts of shellfish reef management on the provision of ecosystem services resulting from climate change in the Dutch Wadden Sea. *Marine Policy*, *119*, 104058. https://doi.org/10.1016/j.marpol.2020.104058

ELECTRICBLUE. (n.d.). *ElectricBlue*. ELECTRICBLUE. Retrieved 23 November 2023, from https://electricblue.eu/pulse

Frölicher, T. L., Fischer, E. M., & Gruber, N. (2018). Marine heatwaves under global warming. *Nature*, *560*(7718), Article 7718. https://doi.org/10.1038/s41586-018-0383-9

FutureMARES. Retrieved 22 November 2023, from https://www.futuremares.eu/about

Gilby, B. L., Olds, A. D., Peterson, C. H., Connolly, R. M., Voss, C. M., Bishop, M. J., Elliott, M., Grabowski, J. H., Ortodossi, N. L., & Schlacher, T. A. (2018). Maximizing the benefits of oyster reef restoration for finfish and their fisheries. *Fish and Fisheries*, *19*(5), 931–947. https://doi.org/10.1111/faf.12301

Mortensen, S., Skår, C. K., & Sælemyr, L. (2020). Summarizing the screening for Bonamia ostreae in Norwegian populations of flat oysters, Ostrea edulis. *17*. https://imr.brage.unit.no/imr-xmlui/handle/11250/2686816

Rizzi, J., Torresan, S., Critto, A., Zabeo, A., Brigolin, D., Carniel, S., Pastres, R., & Marcomini, A. (2016). Climate change impacts on marine water quality: The case study of the Northern Adriatic sea. *Marine Pollution Bulletin*, *102*(2), 271–282. https://doi.org/10.1016/j.marpolbul.2015.06.037

Sas, H., Deden, B., Kamermans, P., zu Ermgassen, P. S. E., Pogoda, B., Preston, J., Helmer, L., Holbrook, Z., Arzul, I., van der Have, T., Villalba, A., Colsoul, B., Lown, A., Merk, V., Zwerschke, N., & Reuchlin, E. (2020). Bonamia infection in native oysters (Ostrea edulis) in relation to European restoration projects. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *30*(11), 2150–2162. https://doi.org/10.1002/aqc.3430

Walles, B., Mann, R., Ysebaert, T., Troost, K., Herman, P. M. J., & Smaal, A. C. (2015). Demography of the ecosystem engineer Crassostrea gigas, related to vertical reef accretion and reef persistence. *Estuarine, Coastal and Shelf Science*, *154*, 224–233. https://doi.org/10.1016/j.ecss.2015.01.006

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

WUR website. (2017, September 26). *Restoration of European flat oyster reefs in the North Sea and Wadden Sea*. WUR. https://www.wur.nl/en/article/restoration-of-european-flat-oyster-reefs-in-the-north-sea-and-wadden-sea.htm