Research Project

SOUND BASED RESEARCH ON CLOCK-SHIFTING THE MANX SHEARWATER

UNIVERSITEIT VAN AMSTERDAM

Julian Main (10541578), Lucas van Berkel (10747958),
Yorick de Boer (10786015) and Amor Frans (10897844)
Amsterdam, University of Amsterdam Januari 2016

Contents

Abstract

The purpose of this research is to gather more information about a seabird called the *Manx shearwater*. We attempted to gather information about the biological clock of the bird, and also wanted to determine if the bird navigates by using the sun. In order to examine the birds' way navigating, an attempt was made to shift the its biological clock by manipulating the light to which the birds were exposed to inside its burrows. Using audio recordings from a microphone which was placed inside the birds' burrow, we have tried to find patterns in the its behaviour. This was done by using various machine learning algorithms to recognize, and label the recorded sounds by the time of occurrence, and type of sound. The labeled data before and after the manipulation of the visible light in the burrow was compared in order to see if the pattern of activity had shifted. Although it seemed as if the birds had a pattern of activity, the results were not significant enough to prove a clock-shift.

Introduction

The purpose of this research is to determine if a clock-shift has occurred on 8 Manx shearwaters in an experiment held last summer in 2015. The experiment manipulated 8 Manx shearwaters by attempting to shift its biological clock. The research conducted here is commissioned by the Institute for Biodiversity and Ecosystem Dynamics. The institute is represented by Oliver Padget, a PhD student of the university of Oxford, helped by the dutch associate dr. ir. Emiel van Loon, Assistant Professor on statistical ecology. The same project was conducted last year by a different group, but the results were unable to give a satisfying answer to the main question of the institute.

The Manx Shearwater is a black and white seabird which nests in burrows. This bird nests on islands in the northeast Atlantic. In July, after the breeding period, the birds will migrate to the South-Atlantic and will return again in March the following year. During the breeding period one of the parents will stay with the eggs during a period of 7 to 12 days, while the other parent will search for food. After these 7-12 days the parents will swap roles.

In order to test if the Manx shearwater uses the sun as a compass while migrating over sea, the tested birds were manipulated, changing their endogenous clock four hours forwards or backwards. For the clock-shift the burrows were closed for nine days and an artificial light was placed in the burrow. For the first six days the time when the light was on, would match the time when the sun was up. After the six days the clock-shift was executed and the time when the light was on would differ four hours from the time when the sun was up. The clock-shift was chosen as test, because, if the Manx shearwater uses the sun as a compass, then the bird needs a precise endogenous clock to navigate. If the sunlight to which the birds were exposed was altered, this would mean the the endogenous clock would also be altered. After the clock-shifting has occurred, the birds were released some distance from their burrows on sea with GPS-tracking, to see what kind of path the birds took. To prove that the clock-shifting has successfully occurred, some birds were manipulated the same way, but with a microphone in the burrow. If the Manx shearwater was clock-shifted, the researchers presumed its sound-pattern would change the same way. Therefore the main question of this research was to prove whether it was possible to show that the bird was successfully manipulated through the audio recordings.

Last years group calculated the most dominant frequency, the activity over a period of time using the amplitude as threshold, the average amplitude over a period of time and the band power. Five minutes was the chosen step time, so every interval provided enough information to calculate each feature. The previous research concluded that the results were not reliable enough to determine whether the birds were clock-shifted.

Problem description

The main subject is to determine if the clock-shift had worked. Recordings were made in 8 burrows. In each of these burrows a clock-shift was performed. Two types of clock-shifts were used: a 4 hour forward shift(shift-fast), and a 4 hour backward shift(shift-slow). Audio recordings of 8 different birds were provided. Each microphone of each bird produced 9 audio files. One audio file for every day of the experiment. For the first six days the time when the light was on, would match the time when the sun was up. In the last three days the clock would be shifted backwards or forward. The recordings were not complete for all burrows.

burrow	16/6	17/6	18/6	19/6	20/6	21//6	22/6	23/6	24/6
F:B73	✓	✓	✓*	✓	✓	✓*	✓	✓	✓
S:B151	×	×	×	×	×	×	×	×	✓
F:B174	✓	✓	✓*	✓	✓	−	−	✓	✓
S:B179	✓	✓	−	✓+	✓	✓	✓	✓	✓
F:DB4	✓	✓	✓	✓	✓	Mic	✓	✓	✓
S:DB12	×	×	×	×	×	×	×	✓	✓
F:DB20	✓	−	✓	✓	✓	✓−	✓	✓	✓
S:DB30	✓	✓	✓	✓	−	✓	✓	✓	×

Table 1: Table of the data given by Oliver Padget, where ✓ means full 24 hours, ✓ −means between 18 and 24 hours, ✓ * means less then 18 hours. × means no recording. Mic means that microphone was not located properly. F means forward shift, S means backward shift.

The presumption was that the clock-shift would not work immediately, comparable with a jet-lag for human beings. Therefore the expectation was that the clock-shift would work, but in steps of 1 hour per day.

To validate the main goal the following sub-questions were attempted to answer:
(I) What is the activity pattern of the Manx shearwater during 24 hours in the audio recordings?
(II) What is the difference between the non-shifted days and the shifted days in the activity pattern?

Strategy

To determine the activity-pattern, every activity in each recording was labeled. Each burrow had 24 hours of recordings for 9 days, in total approximately 216 hours of audio recording for all days. There were nine burrows, so a total of 1944 hours of recordings.

We decided to use the following steps the analyze the data:

1) Preparing data
2) Train model to classify sounds
3) Abstract results from classification to usable data
4) Analyze results

The following chapter is a step-by-step description of how we handled these steps.

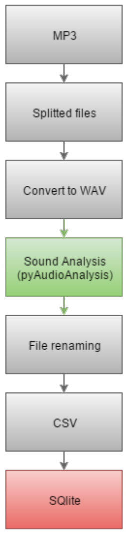

Figure 1: Flow chart

Method

Preparing data

Since the provided audio files were MP3 encoded, we had to decode them to WAV format to process them. In order to make the files more manageable, each recording was split into pieces of exactly one hour long. This resulted in files of approximately 1.5GB. The analyzation process of the file keeps the whole file in memory, so they could not be too big. Next, the files had to be renamed so the filename represented the real date and time of the recording.

Train model for classification

While listening to some audio files, multiple sounds could clearly be classified. The most interesting sound was when the bird moved, this sound was named 'shuffle'. The sound that mostly identifies the Manx shearwater is its call, but previous research showed that the bird calls mostly in anticipation of other birds[1]. Given that the call would not have been a good identifier of the activity pattern of the bird. Therefore we chose 'shuffling' as our identifier for the activity pattern. The constant background noise was classified as silence. We also added a class named environment noise so that we were sure that the noises from other birds would not be classified as shuffles or calls by the bird in the burrow. Therefore we have the classifications: shuffling, calling, silence and environment noise.

The training data was collected by searching in the audio files for shuffles, silent fragments, calls and environment noises by ear. The group of last year combined all the data of the different birds. The result of this strategy was not so good, because every burrow can differ in size and therefore the recordings can be different. Also the place of the microphone can differ for each burrow. Therefore we chose to collect training data for each bird separately.

After collecting the training data we let an algorithm learn to recognize these sounds and make a vector over 24 hours of this classification. An open source library for the programming language Python was used (pyAudioAnalysis)[3] for this. This library uses 34 different short-period features, in order to determine if the sound tested, matches the sound classified using the K-nearest neighbours algorithm. Some of the features the program uses are the Zero Crossing Rate, Entropy of Energy and Spectral Spread. A complete overview of every feature is provided along with the code. We chose for a stepsize of 1 second so it returns a classification of every second in the tested audio file.

A preliminary test of detecting 'shuffles' in a time frame of 241 seconds concluded that 90 percent of the 'shuffles' was detected in that time frame.

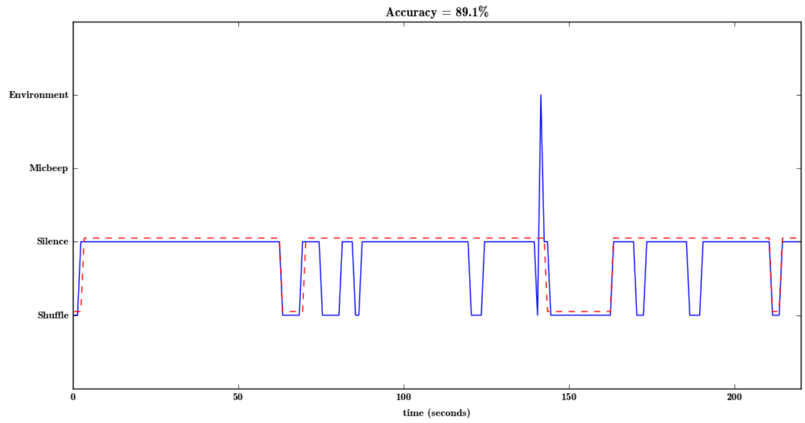

Figure 2: A preliminary accuracy test

Abstract to usable data

Each audio file, approximately $9 \cdot 8 \cdot 24 = 1728$ files in total, was run through the audio analysis, delivering an array of 3600 entries long per file. Since the microphones of two birds did not work properly for 8 days (see table 1), the results of these two birds had to be discarded.

After the classification was done, every second of the audio recording then was labeled with its class and put into an array. Next every class was labeled with a timestamp and put into CSV-files (Comma-separated values files) for every bird. The CSV-files were later put into a sqlite-database (Structured Query Language Database), in order to make it easier to retrieve data from certain time points during the analysis.

Analyze results

After the data was usable we plotted the results. In addition to this, two methods were used to analyze the data: an interval based test and an analysis based on anomaly detection.

Interval based test

To validate if there was a clock-shift, the shifted days had to be compared with the non-shifted days. A vector was made of the amount of shuffles for each day. The vector contained 24 entries, each entry was the amount of shuffles at a certain hour of the day. In total there were 9 vectors for each burrow: 6 for the first 6 days, before the clock-shift had occurred; and 3 for the last 3 days, after the clock-shift had occurred.

The idea was to separate every hour and make an interval for every hour using the data from the non-shifted days. And then look if the same hour in the shifted day falls in that interval. After this procedure the program would shift the vector of the shifted day and repeat the procedure.

The programs repeats this procedure for every hour. At the end the program computes a percentage of the hours that falls in the interval. After this whole procedure we shifted the time-line of the shifted day by an hour and we did this four times. This is because the birds possibly did not shift their activity immediately. The expectation is that the birds will shift by ca. 1 hour per day so we wanted to see that the percentage of hours that falls in the interval would increase by the hours of shifting. If that was the case then we could conclude that the clock-shift has worked.

Anomaly detection

The idea of anomaly detection is to detect anomalous points or anomalous vectors given the data. As anomaly detection algorithm we chose to use the least-squares kernel based method. If there are two vectors x and y then a kernel is the space of all vectors in x that will denote in the zero-vector in y. The least squares equation is an equation to compute the distance between the estimation and the real value. The this is to minimize this distance using the algorithms and kernels. For an extensive explanation of this algorithm see [2].

For this research the training data was the data of the amount of shuffles for each hour in all non-shifted days, these are six vectors which each represent a non-shifted day. Then a shifted day was added to the data. We did this for each shifted day. We hoped to see that this shifted vector would be flagged as anomalous before the shift, because it should not fit in the non-shifted day pattern. Also we wanted to see if this shifted vector was flagged as not anomalous after the shift of 4 hours. When this would be the case, we could conclude that the clock-shift worked.

Results

Classification

After plotting the classified data, we saw that, although there seems to be a pattern of activity for some of the birds, other birds did not show a clear daily pattern. Below are two of the more interesting looking plots are displayed.

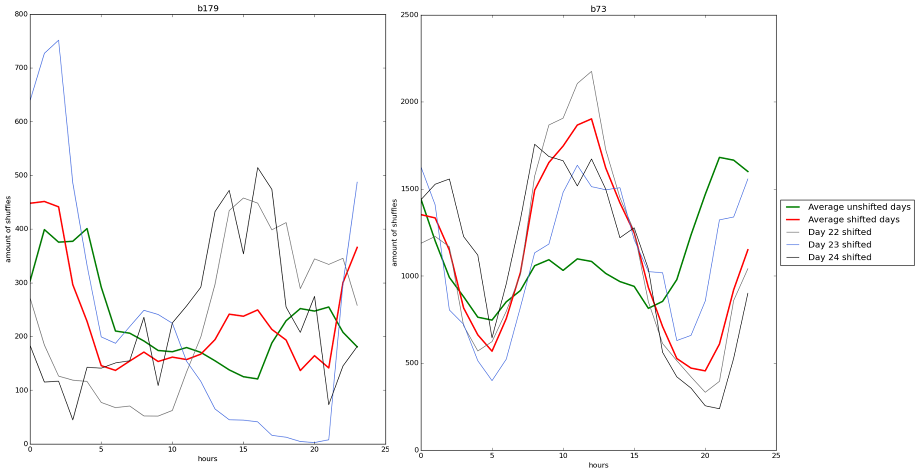

Figure 3: Comparison of shifted vs non-shifted days of bird b179 and b73

To smooth the results we used a running mean. Also the plots were normalized so the surface under the plots is equal for each graph. This was done so the pattern among the graphs could be compared better.

Especially the pattern shift of the b179 bird is interesting, note that this is a slow shifted bird.

To find out how the average of the non-shifted days represents the the real value, a min-max plot was made.

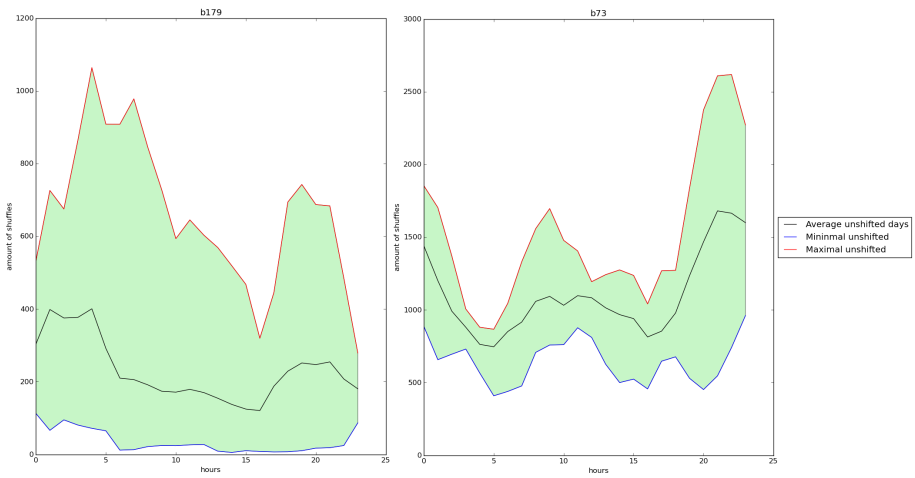

Figure 4: Minimal and maximal values of non-shifted days bird b179 and b73

It can be seen that there is a significant difference between these two birds. While the minimum and maximum values from the b73 bird are quite close to the the average, the values of the b179 bird are much further apart.

Interval based test

For our interval based approach we ran a test to validate whether the pattern from the shifted days and non-shifted days differ significantly. In *Table 2* are the results of the interval based test.

	Day 22					**Day 23**				
hours of shifting	*0*	*1*	*2*	*3*	*4*	*0*	*1*	*2*	*3*	*4*
b73	54.167%	41.67%	45.83%	37.5%	45.83%	29.17%	41.67%	45.83%	37.5%	33.3%
b174	0.0%	0.0%	8.33%	12.5%	8.33%	50.0%	50.0%	50.0%	45.83%	45.83%
b179	75.0%	75.0%	70.83%	83.33%	87.50%	79.17%	70.83	79.17%	83.33%	70.83%
DB4	50.0%	50.0%	58.33%	62.5%	66.67%	37.50%	37.5%	50.0%	54.17%	50.0%
DB20	100.0%	100.0%	91.67%	87.5%	83.33%	95.83%	91.67%	95.83%	95.83%	95.83%
DB30	50.0%	45.83%	50.0%	54.17%	54.17%	25.0%	25.0%	25.0%	25%	25%
	Day 24									
hours of shifting	*0*	*1*	*2*	*3*	*4*					
b73	25.0%	25.0%	25.0%	37.5%	41.67%					
b174	62.5%	62.5%	58.33%	66.67%	62.5%					
b179	83.33%	83.33%	87.5%	91.67%	83.33%					
DB4	58.33%	54.17%	54.17%	62.5%	54.17%					
DB20	87.50%	87.50%	83.33%	83.33%	79.17%					
DB30	8.33%	8.33%	8.33%	8.33%	8.33%					

Table 2: Results of the interval test. *Green: percentage of the hours that falls in the interval increases, red: percentage of the hours that falls in the interval decreases, Yellow: percentage of the of the hours that falls in the interval stays the same*

Because, presumably, the birds will not change their behaviour immediately after the clock shift the columns shifted-hours 1, 2 and 3 are more important for day 22. In *table 3* the percentage increased for day 22 for the columns *shifted-hours* 1, 2 or 3, which are bird b174, b179 and DB4. For day 23 the shifted hours 2, 3, 4 are more important, assuming the bird has adapted to the clock-shift more. Here we can see that 3 of the 6 birds percentage increased at one of the three important hours.

For the last day the important columns are the shifted-hours 3 and 4, assuming again that the bird has now adapted fully to the clock-shift. Remark that day 24 is the most important day to compare because at this day we will assume that the bird has fully adapted to the clock-shift. Interesting to see is that the percentage of four birds (b73, b174, b179 and DB4) increased when we shifted the timeline for 3 hours. So it looks like that these four birds are successfully clock-shifted for 3 hours. Only bird b73 looks like it has been clock-shifted for 4 hours.

Four of the six birds showed good results, so there are signs that the clock-shift worked, but it is not significant enough.

Anomaly detection

Anomaly detection will flag vectors as being anomalous if they do not fit a certain group of data(cluster). If a vector of the shifted day will be flagged as anomalous means that the pattern of activity differs. The activity vector of a shifted day will be compared with the cluster of the non-shifted days using the least squared kernel-based method. If the compared vector exceeds a threshold than the vector will be flagged as anomalous. Due to the chance that the birds clock will not shift immediately we defined the clock-shift by steps of 1 hour. The results of the anomaly detection are in *Table 3*

	Day 22					**Day 23**					**Day 24**				
hours of shifting	*0*	*1*	*2*	*3*	*4*	*0*	*1*	*2*	*3*	*4*	*0*	*1*	*2*	*3*	*4*
b73															
b174															
b179															
DB4															
DB20															
DB30															

Normality	
Anomalous	

Table 3: Results of the anomaly detection test

Each row in *Table 3* represents a bird and the test is done on all three shifted days, which are day 22, 23 and 24. The numbers 0-4 for each day represents the hours of shifting, this is done because the birds presumably could not adapt immediately to the clock-shift. If the cell is green it means that the hour is a normality which means that it fits in the pattern of the non-shifted days. If the cell is red it means that it is flagged as an anomalous and the pattern differs significantly.

To validate whether the clock-shift has worked the shifted day must be flagged as anomalous if we not have shifted the timeline yet (shifted hours=0) and be flagged as normality in at least one of the shifted hours (shifted hours = 1, 2, 3, 4). This pattern appeared only 3 times in the table as we can see in table 3.

Discussion

Conclusion

Following from the results, the conclusion of the research is that the results are too inconclusive to prove that the birds were successfully manipulated into shifting its biological clock. The graphs show that there is a difference between the shifted and non-shifted days. It also shows that the birds follow the normal day night cycle and that there is a slight shift in the pattern, but the difference is not monotonous enough to mark a significant change.

To answer the sub questions, there is an activity pattern but then again, the pattern is not clearly the same for all the birds. The difference in the pattern between the non-shifted and shifted days is also not significant and monotonous. On the other hand, the main question was to see if it was possible to determine if a shift in the birds biological clock has occurred. Determing this shift seems possible, but we were not able to successfully determine the prove this, given only these audio recordings.

Evaluation

The surroundings of this research were very good. The previous group stated that the audio recordings were incomplete, unclear and not named properly. This year, with new audio recordings, these issues were much less present. Little gaps and a few major gaps in the audio recordings were still present, but enough data was offered to create a baseline for each bird, left out the two birds with the broken microphones. The quality of the recordings were probably the same as last year, but this research classified the noise in the recording as silence. This problem of the unclear audio from the previous group was bypassed by explicitly searching for activity of the bird, the shuffling sound. The file names of the audio recordings were nearly perfect, only one of the 54 audio recordings had to be discarded since the given time stamp did not fit into the timeline(The first audio recording of DB30 was too long for the given start time, which created a overlap at the end of the recording). Still some issues have to be discussed. Since the microphones layed in the burrow for nine days under the bird, it was inevitable that the conditions of the microphone changed over time. Dirt could have covered it, which would have an effect on the audio recordings. Also it is not ruled out that the birds daily pattern is influenced by surrounding birds. So the biological clock of the bird is possibly shifted, but it would not be detectable on the audio recordings. At last the preliminary research of the small audio recording tested, gave 90 percent accuracy. The test was conducted multiple times, but on the same recording. It is unclear if our program would have performed differently on a different time of day.

Future research

The research still has flaws which could be improved in future research. Having an artificial burrow, better microphone and batteries that run for the entire nine days would majorly improve the quality of the data. Also the shifting of the biological clock could be spotted differently in future research. This research was conducted with audio recordings, but with a temperature reader or heart rate monitor the data would be more clear to prove that the biological clock of the bird is shifted. But of course these are all very expensive options. To conclude everything, the birds biological clock are most likely shifted, but it could not be proved.

References

[1] M de L Brooke. "Sexual differences in the voice and individual vocal recognition in the Manx shearwater (Puffinus puffinus)". In: *Animal behaviour* 26 (1978), pp. 622–629.

[2] Yaakov Engel, Shie Mannor, and Ron Meir. "The kernel recursive least-squares algorithm". In: *Signal Processing, IEEE Transactions on* 52.8 (2004), pp. 2275–2285.

[3] Theodoros Giannakopoulos. "pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis". In: *PloS one* 10.12 (2015), e0144610.

Appendix

Source code

All source code used for this project can be found here:
`https://github.com/manxshearwater-clockshift`

Log

Week 1

During the first meeting Oliver Padget and Emiel van Loon described the problem to us. Contact details and further appointments were made. We decided that we would meet about two times each week and more if necessary. Oliver Padget is our first contact, and Emiel Van Loon is the secondary contact. We agreed to to send updates of our progress regularly.For communication between group members we made a Whats-App group chat, and also created a Github repository where the code and other files of our project would be stored. Our strategy was formed, and tasks were assigned to each group member.

Last year a group of students also tried to solve the same problem, the report of this group has been sent to us. According to their results they have not been able to successfully determine whether there was a clock-shift. So we decided to try a different approach.

The group of last year separated the data in smaller segments so the size of the files would be manageable. After the separation they classified each file and predicted whether the bird at that time perceived the moment as day or night. They used 3 features(amplitude, frequency and activity).

Before we determined our strategy we investigated which environment has the most potential to solve this problem. We chose to program in Python because we found an extensive library to analyze audio files (pyAudioAnalysis). The machine learning module in this library uses 34 features to analyze an audio file/segment. This is a big improvement to the 3 features of last year.

Before we started programming and analyzing the audio files we first determined a strategy. We had 3 options in mind:

1. The same strategy as last year: separate 24-hour fragments into smaller segments and classify them into day or night activity using very rough features. *In the report of last year we read that they faced many problems and the result were not as good as they hoped, so we decided to not try this option again although our data is better regarding to our client.*

2. Activity detector: Implement a threshold that is depending on features. If the audio file exceeds this threshold then this segment/point will be returned on the time-lines. then compare the time-lines of the normal clock and the shifted clock. If the shifted-clock time line differ from the normal then there is a signal that the clock-shift worked.

3. Supervised isolating distinctive sounds: Collect training data of distinctive sounds such as shuffling, calling and environment noise. Then we would let our program learn to recognize these sounds and make a vector over 24 hours of these classifications. Then we would compare the normal vectors with the clock-shifted vectors.

We decided to choose for option 3, since this would be a more precise way than option 2.

When the strategy was chosen we started to collect the training data. We searched for distinctive sounds such as shuffling, calling, environment and silence. These fragments were mostly between 3 and 6 seconds long. We tested on two audio files:
- b73 the audio file of the first day, where there is no clock-shift.
- b73 the audio file of the last day, where the clock is shifted

We did this because in this way we also could test the compare method after the classification, so we could see if there were signs of the clock shift having worked. Because of the large files it is not possible to analyze more audio files before we were sure that the code worked well.

We started by converting the mp3 to wav. When that was done we started testing pyAudioAnalyis to see if it could successfully distinguish and classify different sounds in the audio files. We used the pyAudioAnalysis library with the k-nearest neighbor algorithm. pyAudioAnalysis uses the values of 64 features at each point and classifies that as shuffling or silence. We started to classify only these two to see if our strategy has potential. Our first test was to let the program recognize the sound of shuffling. We started with an accuracy of 33% but after some fine tuning we reached an accuracy of 89%.

After some testing and collecting training data we discovered that the mp3 files contained errors. We decided to cut the audio mp3 audio files in segments of 1 hour and then convert these smaller files to wav.

Emile Van Loon advised us to look for a better algorithm than k-nearest neighbor. So we investigated which other algorithms can fit our program and our problem. This could be a hard thing to change, because we would have to expand the pyAudioAnalysis library which would take a lot of work and time. It is a complex library which lacks proper documentation.

In our meetings this week with Oliver Padget we described our progress and our strategy. Padget approved our strategy and understood why we chose this strategy. He also recommended to analyze the difference between the day and night patterns between the normal and the clock-shifted audio files. We can do that using the information of the latitude and longitude of the location of the birds, the date/time the audio recorded started and information when on that particular day at that location sunset and sunrise took place.

In week 2 we planned to optimize our program trying other learning algorithms and run the program on a larger scale. We also want to choose a good strategy to compare the results. A possibility is using machine learning but because there are not many audio files it may be possible do it manually and visually. But that depends on the difference between the results of the normal files and the clock-shifted files.

Week 2

Our main goal of week 2 was to get results on a single test bird/burrow b73. Last years group main research was done on with all the data from all burrows merged in one set. The results of this strategy were not conclusive. Because this approach did not work out so well, we decided to research each burrow separately. We investigated if we had to get individual training data from each bird or that the same training data could be used for all birds. It turned out that separated training set worked better because each burrow can differ in size, the position of the microphone and other dependencies and therefore the recording of the sounds can be different.

After trying several tests running on only one hour of the recording we decided to run the program on the 24 hours recordings of all days of burrow b73. The results were transcribed to CSV, so the results could be visualized. The code for this was written this week.

During the last meeting of week 2, we wanted to show the results of burrow b73. We finished the code for the complete run on b73 and also implemented the code to plot and read the data from the CSV file. We also added the plot of the last day because the bird would probably not immediately adapt to the full 4 hour clock-shift after 1 day. The expectation is that the biological clock will shift 1 hour per day.

Oliver Padget was satisfied with the results and saw signs that the clock shift has worked. Because b73 has been clock-shifted forward he expected to see the activity pattern shift to the right. He said that you can see the first peak shift to the right and therefore the results are hopeful. He advised us to normalize the data and also use average smoothing to get cleaner results and to see a better pattern/shift.

At the end of this week Oliver Padget returned to the UK and we agreed to communicate further by email and also continue our weekly meetings on Skype.

We discovered some errors regarding the test results. Some of the recordings were not the full 24 hours, therefore some time segments returned no data. We had taken this into account but there appeared some error with our implementation. After this fix our program was ready to run on all burrows.

For each burrow the run-time is approximately five to six hours, so we planned to run the computer the whole weekend and have the results at the beginning of week 3.

In week 3 we hope to have the complete results and conclude whether the clock-shift has worked or not.

Week 3

We ran our application during the weekend in order to get the results of the classification for all birds. This process took a very long time because of the the amount of data. It eventually took approximately 36 hours to analyse all the audio files. We now have the CSV files of all birds containing classifications of every second. So every second will be classified as shuffling, environment noise, calling or silence.

We found out that there were errors in the script we wrote to rename the filenames to represent a useful date and time, causing some CSV files to contain more than 24 hours of data. This problem must be corrected before we continue the analysis. The biggest problem with these renamings is that the date tag in the file names does not represent the day on which the recording started. Instead it is set to the day where most of the recording took place on.

After this problem was fixed we began writing code to plot the graph for a particular bird and day.

We also had a meeting with our project assistant Ysbrand Galema, we described our progress. He gave us advise about how to analyze the plots and the results of the classification. His advice was to use statistics to validate whether the clock-shift worked or not. The advice was to make a probability density function of the amount of shuffling for every hour for all normal days. Then look if the amount of shuffles of the shifted days fall into a reasonable percentage of the probability density function of the normal days. We expect that the percentage will be higher if we shift the amount of shuffles back or forward on the timeline, depending on what kind of shift was tested on the bird.

At the end of this week we accomplished to finish the code for the statistical tests. We checked for every hour in a shifted day whether it falls in a interval. The interval is made using statistics and using the data from the non-shifted days. The results were not as good as we expected. We expected that the percentage of the hours that falls in the interval will increase if we shift the hours in a shifted day. But this did not happen for all the birds and all the days, so the results were not good enough to conclude whether the clock-shift has worked or not.

In week 4 we will try another test to validate the clock shift and analyze the graphs. Also we will work on the paper and the presentation that is upcoming next week.

Week 4

In the first days of week we were developing another test to see if the clock-shift worked. We chose to use an anomaly detection algorithm. The idea is to make a cluster from the data of the non-shifted days and then compare this cluster with a shifted day. If it exceeds a certain threshold, the day will be flagged as an anomaly. We hoped to see that the shifted days would be flagged as anomalous before the shift and be flagged as normality after the shift. Unfortunately the results were not what we hoped for. Only for bird b73 the results seemed good. For the other birds we could not extract much from the results.

We tried two tests and we could not validate whether the clock-shift worked or not. There were signs that the clock-shift worked based on some graphs, but because the sample size is small we could not validate it with statistics. Because this was the last week of this project, we spend the rest of the days writing the paper and preparing for the presentation.