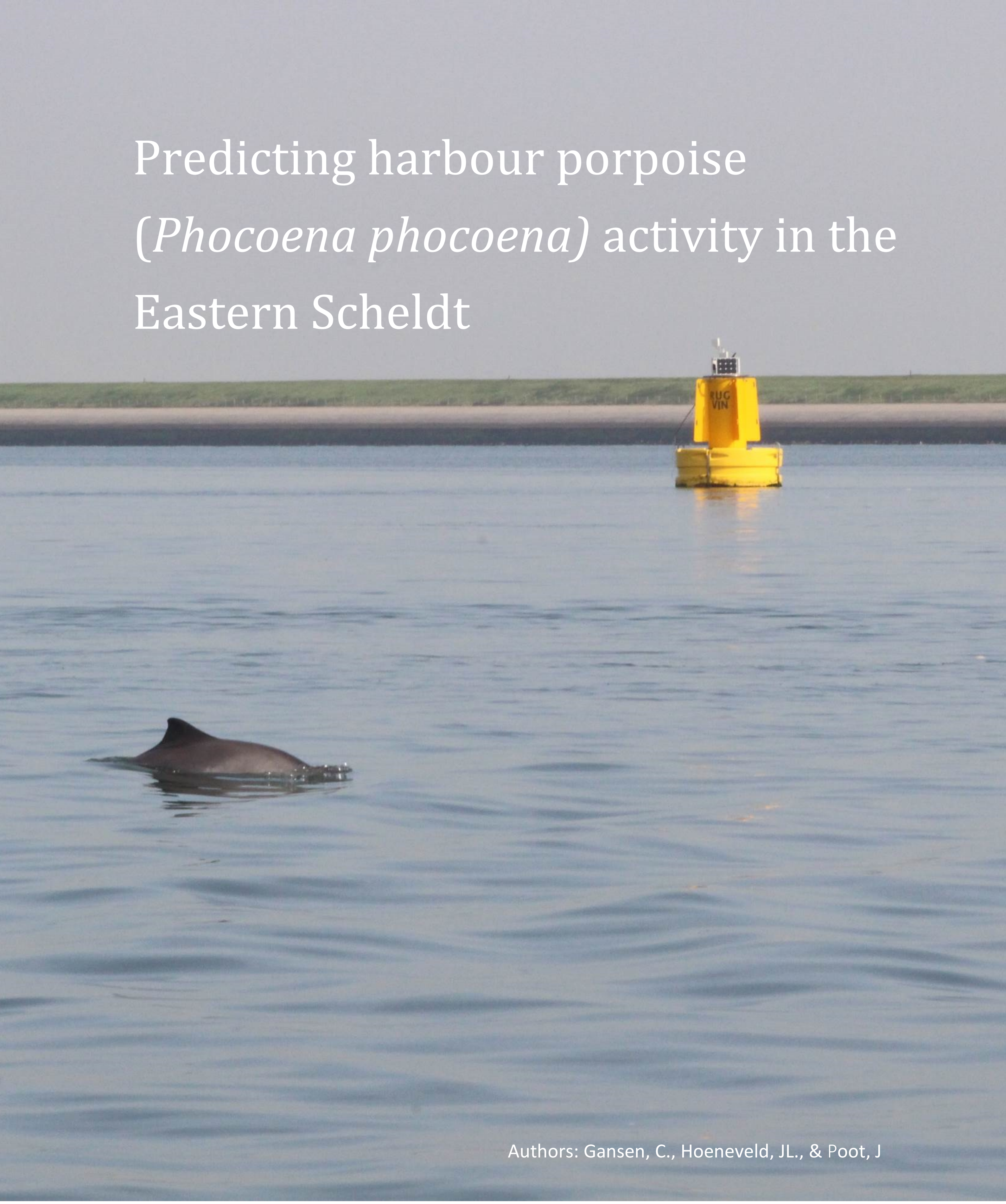
Predicting harbour porpoise (*Phocoena phocoena)* activity in the Eastern Scheldt



Authors: Gansen, C., Hoeneveld, JL., & Poot, J

Predicting harbour porpoise *(Phocoena phocoena)* activity in the Eastern Scheldt

The relation between abiotic factors and harbour porpoise activity around Studio Bruinvis in the Eastern Scheldt and possibilities for developing a mobile predictability application for porpoise watching

Authors

Carine Gansen,

Jordan-lee Hoeneveld

Jesse Poot

Supervisors

Okka Bangma - University of Applied Sciences Van Hall Larenstein

Henry Kuipers - University of Applied Sciences Van Hall Larenstein

Assigned by

Frank Zanderink - Rugvin Foundation

Van Hall Larenstein University of Applied Sciences

June 2018, Leeuwarden

cover photo – Hoeneveld, 2017

Acknowledgement

Three different students, from different countries, but all three with a great interest in cetaceans of the big blue. We are very grateful and happy that we have been given the opportunity to work together on this unique and interesting project.

Special thanks to Frank Zanderink, coordinator of the Rugvin Foundation, for allowing us to work on this project and set up this research. Also many thanks to the WaterProof crew, who helped us to get familiar with the software used during this research. Studio Bruinvis and this project is financially supported by World Wildlife Fund.

Good supervision from the university during this study was given by Okka Bangma and Henry Kuipers. Even though supervision from long distances can be difficult and sometimes it was hard to arrange meetings, we are very thankful for their contribution on this project. Okka always provided us with good advises and Henry was our much needed advisor for the statistical part of the research.

Carine Gansen, Jordan-lee Hoeneveld and Jesse Poot, Leeuwarden, 1th of June, 2018.

Abstract

The harbour porpoise (Phocoena phocoena) is a small ceteacean species and is common in Dutch coastal waters. The Eastern Scheldt, an estuary in the Dutch province of Zeeland, housed a minimum resident population of about 34 porpoises in 2017. This population is monitored and studied by the Rugvin Foundation. The Rugvin Foundation recently installed 'Studio Bruinvis' on one of the quays near Zierikzee. This quay is regularly visited by locals or tourists through bicycle and hiking paths. But more importantly, it is a well-known hotspot for harbour porpoise watching. It is thought that the high number of sightings is linked to prey fish that are attracted by a former and deep underwater ammunition depot located at this hotspot. Studio Bruinvis consists of a mono hydrophone attached to a buoy that is wirelessly linked to a listening post on the quay where visitors can listen live to sounds (clicks) produced by porpoises. The listening post is also equipped with recording equipment. These recordings provide a good opportunity for monitoring porpoise activity around the buoy of Studio Bruinvis. This study aims at gaining more insight on porpoise activity by studying the relation between abiotic factors (i.e. time, water temperature, wind speed, wind direction, tidal state and water height) and porpoise activity

This knowledge can be used for scientific goals, but also to promote porpoise watching. Nowadays many mobile applications (apps) are available to help people spot and identify wildlife which can help to create awareness. Therefore, a second aim of this study is to develop a mobile predictability application for porpoise watching by studying how to develop such an app. This app can help locals or tourists to achieve a higher chance of spotting harbor porpoises at Studio Bruinvis. To achieve these aims, the following research question had to be answered: *"To what extent is harbour porpoise (Phocoena phocoena) click activity influenced by water temperature, water height, tidal water flow, wind direction, wind speed, and time of day around a commonly used hotspot in the Eastern Scheldt and how can this information be integrated in a mobile application to predict harbour porpoise activity?"*

To achieve the first set aim, data was collected from August 15th, 2017 until October 5th, 2017. Porpoise activity data was obtained from Studio Bruinvis and consisted of 1248 audio 1-hour-files. Distortion and stratified sampling resulted in a total sample size of 657 audio files. Pamguard was used to detect the number of produced clicks for each audio file. Data for abiotic variables were obtained from Rijkswaterstaat. Binary logistic regression was used to assess the relation between harbour porpoise activity and the independent abiotic factors. For the dependent variable, porpoise activity, a baseline was set to 400 clicks/hour. Samples with values below this baseline, suggest low porpoise activity. Samples with values above the baseline, suggest high porpoise activity, thus a high chance of spotting harbour porpoises at Studio Bruinvis.

The results of this research showed that all variables (i.e. time of day, tide, water temperature, water height, wind speed and wind direction) were found to have a significant effect on porpoise activity. Tide was found to be the most important predictor followed by wind direction, time of day, wind speed, water height and lastly water temperature. Chance of high activity seemed to increase with increased wind speed, increased water temperature and increased water height. Furthermore, outgoing-to-low tide and incoming tides as well as winds from 225º-270º (west-south-west direction), and time of day between 8:00-16:00hours showed to best predict porpoise click activity at Studio Bruinvis. The equation generated by the regression model was used as the basis for an app to calculate the likelihood of harbour porpoise activity. The model proved to be able to predict activity levels of ≥400 clicks per hour with around 74% certainty. The model was used to form an equation which was then used for the development of the porpoise prediction app.

In order to develop the porpoise prediction app, a literature review about app development was conducted. With the use of Mockflow and Marvelapps, the Studio Bruinvis app was created. This app consists of several screens with information regarding Studio Bruinvis, harbour porpoises and the Rugvin Foundation. But most

importantly, the equation retrieved from the model was used to create a predictive porpoise activity calculator. This calculator informs users whether the likelihood of porpoise activity, and therefore chance to observe porpoises, is high or low.

Despite the outcome of this study, there were some points of discussion. As independent variables in the model, only abiotic factors were used. After a literature study, it was found that these factors are thought to be indirectly related to porpoise activity. The literature study suggests that the used factors can influence prey densities and create foraging opportunities for harbour porpoises, increasing porpoise activity. Due to the short study period and some restrictions that occurred (i.e. distortion), the data did not cover all seasonal conditions. Therefore, the results of this research might not be extrapolated to conditions outside the study period. Furthermore, the chosen baseline of 400 clicks was not validated during this study. Meaning if detected clicks/hour are >400, it is not yet certain if porpoise activity is high at Studio Bruinvis.

The results of this research showed that all variables were found to have a significant effect on porpoise activity. The model proved to be able to predict high porpoise activity (activity levels of ≥400 clicks/h) with around 75% certainty. With this data, the Studio Bruinvis app was created. The results and findings of this study can be seen as a first try in getting a better understanding about factors that influence porpoise activity at Studio Bruinvis and a first step in developing a porpoise activity predictability application.

Samenvatting

De bruinvis (*Phocoena phocoena*) is een veelvoorkomend klein zeezoogdier in de kustwateren van Nederland. In de Oosterschelde, een voormalige zeearm van de Noordzee, leeft een vaste populatie bruinvissen van minimaal 34 bruinvissen, geteld in 2017. Deze populatie wordt gemonitord door Stichting Rugvin. In de zomer van 2017, heeft Stichting Rugvin het nieuwe project Studio Bruinvis gestart op een kade vlakbij Zierikzee. Deze kade wordt regelmatig bezocht door fietsende of wandelende toeristen. Het is dan ook een bijzondere locatie dat bekend staat als hotspot voor het spotten van bruinvissen. Men denkt dat dit komt door een diepe en voormalige munitiedepot. Deze diepe onderwaterput zou allerlei prooivissen van bruinvissen aantrekken. Studio Bruinvis bestaat uit een boei waaraan een hydrofoon hangt. Deze hydrofoon pikt bruinvisgeluiden op en zend deze geluiden draadloos naar een luisterpaal op de kade. Bij deze luisterpaal kunnen bezoekers live luisteren naar bruinvisgeluiden, ook wel clicks genoemd. In de luisterpaal is ook opname apparatuur geplaatst. Met Studio Bruinvis, hoopt Stichting Rugvin meer inzicht te krijgen over de activiteit van bruinvissen. Dit door onderzoek te doen naar het effect van abiotische factoren (tijd, watertemperatuur, windkracht, windrichting, waterhoogte en getij) op bruinvis activiteit bij Studio Bruinvis.

Deze kennis kan zowel gebruikt worden voor wetenschappelijke als educatieve doeleinden. Tegenwoordig wordt veelal gebruik gemaakt van mobiele applicaties (apps) bij het observeren en identificeren van wilde dieren of voor natuureducatie. Stichting Rugvin hoopt meer inzicht te krijgen over de mogelijkheden voor het ontwikkelen van een mobiele voorspellingsapplicatie dat gebruik maakt van informatie over bruinvis activiteit. Om deze doelen te bereiken is de volgende hoofdvraag opgesteld: *"Tot in welke mate wordt bruinvis (Phocoena phocoena) activiteit beïnvloedt door watertemperatuur, waterhoogte, getij, wind richting, windkracht en tijd bij een bruinvis hotspot in de Oosterschelde en hoe kan deze informatie worden gebruikt voor het ontwikkelen van een mobiele voorspellingsapplicatie?".*

Om het eerste doel te bereiken, is data verzameld in de periode van 15 augustus 2017 tot 5 oktober 2017. Data voor bruinvis activiteit is verkregen van Studio Bruinvis en bestond uit 1248 1-uur-audio bestanden. Ruis en een gestratificeerde random steekproef resulteerde 657 audio bestanden. Om deze audio bestanden te analyseren op bruinvis clicks is gebruik gemaakt van Pamguard. Data voor alle abiotische variabelen is verkregen via Rijkswaterstaat. Om de relatie tussen de abiotische variabelen en bruinvis activiteit te analyseren is gebruik gemaakt binaire logistische regressie. In dit onderzoek wordt bruinvis activiteit gezien als de afhankelijke variabele. Een sample met minder dan 400 bruinvis clicks suggereert weinig bruinvis activiteit. Een sample met meer dan 400 bruinvis clicks suggereert veel bruinvis activiteit en dus een grote kans om bruinvissen te spotten.

De resultaten van het model laten zien dat alle abiotische variabelen (tijd, getij, waterhoogte, watertemperatuur, windsnelheid en windrichting) een significant effect hebben op bruinvis activiteit. Getij is de belangrijkste voorspeller voor bruinvis activiteit, daaropvolgend komt windrichting, tijd, windsnelheid, waterhoogte en als laatste watertemperatuur. De kans op veel activiteit van bruinvissen wordt hoger naarmate windsnelheid, watertemperatuur en waterhoogte toeneemt. Tussen 08:00-16:00 uur en met de omstandigheden inkomend tij, inkomend naar laagtij en windrichting uit West-Zuid-West richting, is kans op veel bruinvis activiteit hoger. Het model laat zien dat veel (>400 clicks/uur) of weinig (<400 clicks/uur) bruinvis activiteit met ongeveer 74% zekerheid kan worden voorspeld. Met de uitkomsten van het model zijn een formule gemaakt. Deze formule maakt het mogelijk bruinvis activiteit te voorspellen door gebruik te maken van abiotische data.

Voor het ontwikkelen van een bruinvis voorspellingsapp was eerst literatuuronderzoek over app ontwikkeling nodig. Voor het ontwikkelen van de app, is gebruik gemaakt van Mockflow en Marvelapps. Dit resulteerde in de Studio Bruinvis app. De app bestaat uit meerdere schermen met informatie over Stichting Rugvin, bruinvissen en Studio Bruinvis. De belangrijkste functie van de app is de mogelijkheid om de kans op bruinvis activiteit te berekenen. Voor deze functie is de formule vanuit het model gebruikt. Met deze functie kunnen gebruikers hun kansen optimaliseren voor het spotten van bruinvissen bij Studio Bruinvis.

Tijdens dit onderzoek, kwamen een aantal discussie punten naar voren. Zo is er voor dit onderzoek enkel gebruik gemaakt van een aantal abiotische factoren. Uit de literatuur blijkt dat de gebruikte factoren een indirect effect hebben op bruinvis activiteit en een direct effect hebben op de aanwezigheid van prooidieren. Door de korte onderzoeksperiode en ruis op de audio bestanden, was het niet mogelijk data te verzamelen van alle omstandigheden die gedurende een heel jaar kunnen voorkomen. Daarom zijn de resultaten van dit onderzoek niet representatief voor omstandigheden buiten de onderzoeksperiode. Verder is tijdens dit onderzoek de gekozen baseline van 400 cliks/uur niet gevalideerd. Het is daarom niet zeker dat meer dan 400 clicks per uur, daadwerkelijke leid tot een grote kans op het waarnemen van bruinvissen bij Studio Bruinvis.

De resultaten laten zien dat alle variabele een significant effect hebben op aantal clicks/bruinvis activiteit. Op basis van het model, was het mogelijk bruinvis activiteit met ongeveer 75% zekerheid te voorspellen. De uitkomsten van het model en een literatuur onderzoek resulteerde in de Studio Bruinvis app. Alle resultaten en bevindingen van dit onderzoek zijn een eerste stap in het meer inzicht krijgen in het effect van abiotische factoren op bruinvis activiteit bij Studio Bruinvis. Ook is het een eerste stap in de mogelijkheden voor het ontwikkelen van een mobiele voorspellingsapplicatie voor bruinvis activiteit.

CONTENTS

1. INTRODUCTION

The harbour porpoise (*Phocoena phocoena*) is a cosmopolitan cetacean species that is distributed in temperate waters (figure 1) and reported present in at least 36 countries in the Northern Hemisphere (Lockyer, 2003; Hammond *et al.*, 2008). Four subspecies have been identified, all present in different waters across their range: *P.p. vomerina* and an un-named subspecies in the Eastern and Western Pacific respectively, the *P.p. relicta* in the Black Sea, and the *P.p. phocoena* in the North-Atlantic (Hammond *et al.*, 2008). In the western North-Atlantic Ocean porpoises are present all along the coast of the United States of America and Canada, reaching as far as the west-coast of Greenland and in the eastern part of the Atlantic they range from Senegal to Nova Zembla (Lockyer, 2003; Hammond *et al.*, 2008). Here, they are mainly found in shallow waters of the continental shelf but they are sometimes also sighted in more pelagic settings (Hammond *et al.*, 2008; Brookes *et al.,* 2013). In European waters of the North Sea they are also a common species, with abundances estimated to be around 345.000 (Hammond *et al.,* 2002; 2017). The harbour porpoise is also one of only two resident cetacean species in Dutch waters (the Dutch North Sea), the other being the white-beaked dolphin (*Lagenorhynchus albirostris*) (van der Meij & Camphuysen, 2006; Camphuysen & Siemensma, 2011).

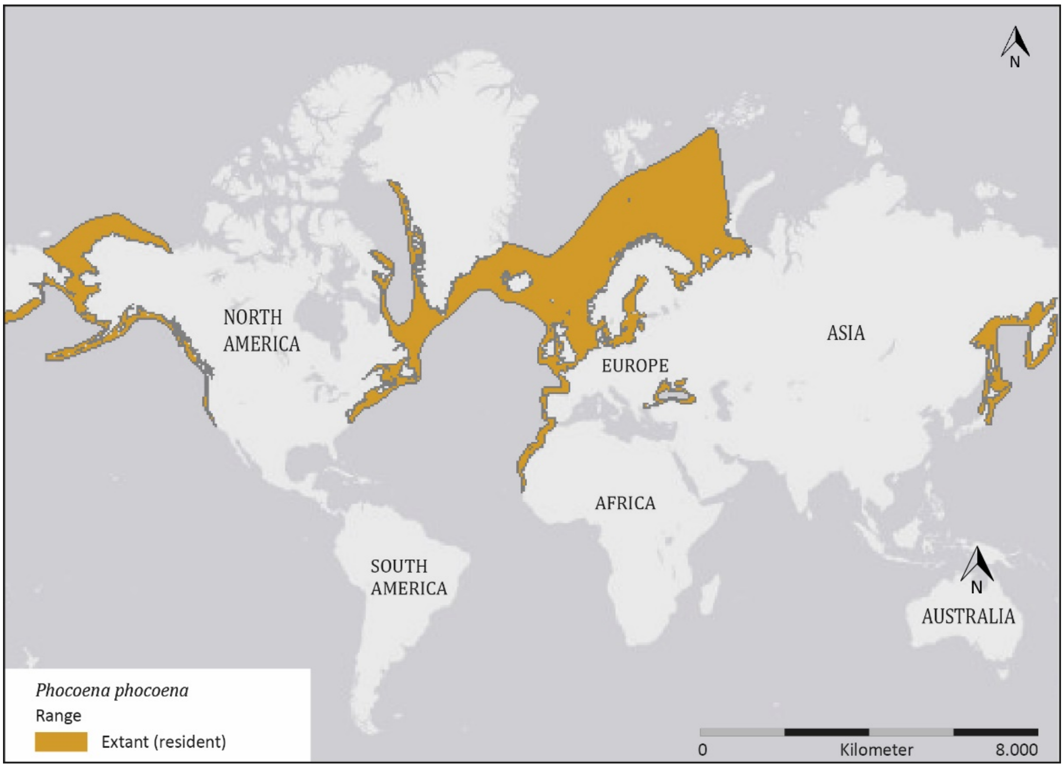


Figure 1: Global residential presence of harbour porpoise (Phocoena phocoena) (Hammond et al., 2008)

On the International Union for Conservation of Nature's (IUCN) red list of threatened species the harbour porpoise is listed as being of *Least Concern* (Hammond *et al.,* 2008). However, in the Netherlands, porpoises are strictly protected under the Nature Protection Act's, habitats directive appendix IV (*NL: wet natuurbescherming, habitat richtlijnen bijlage IV*), meaning that protective measures need to be taken in order to protect the species and its habitat and therefore monitoring of the species is of importance (Ministerie van Economische Zaken, 2017). For this reason, harbour porpoise strandings are closely monitored along the Dutch coast and cadavers are examined by the Faculty of Veterinary Medicine of Utrecht University to determine the cause of the stranding (Keijl *et al.,* 2016).

The National park the Eastern Scheldt (*NL: Oosterschelde*), an estuary in the Dutch province of Zeeland, houses a minimum resident population of about 34 porpoises in 2017, with the population fluctuating between 15 and 61 (figure 2) since the annual counting has started in 2009 (Stichting Rugvin, 2017).

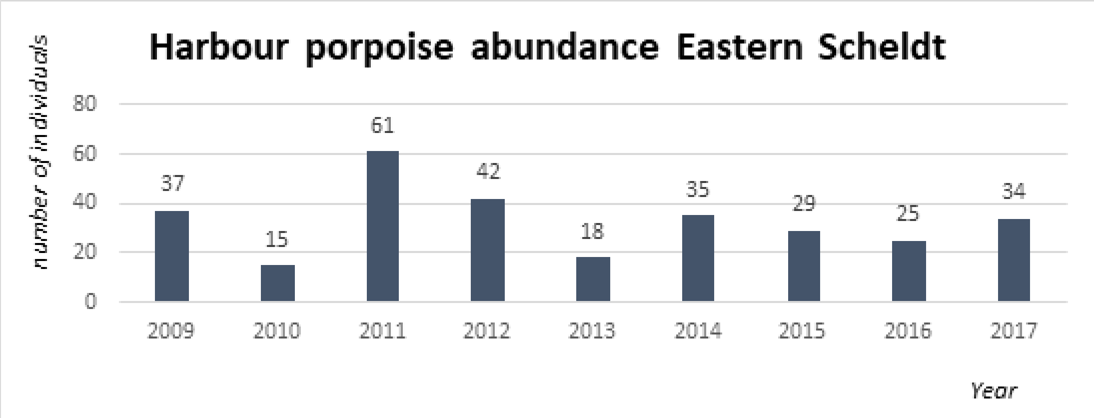


Figure 2: Harbour Porpoise abundance in the Eastern Scheldt (Stichting Rugvin, 2017).

Due to a storm surge barrier (located between the Eastern Scheldt and the North Sea), harbour porpoises are separated from the North Sea to a certain degree. Strong tidal currents are causing high frequency underwater sounds by water flowing/rubbing against the pillars, therefore the openings of the barrier can be difficult for harbour porpoises to pass through. However, migration between the Eastern Scheldt and the North Sea is not impossible, and has been recorded for the resident harbour porpoises, though not very frequently (Korpelshoek, 2011).

With 10.1 million stays in 2016, Zeeland is the most popular province of the Netherlands to be visited by tourists (Kenniscentrum Kusttoerisme, 2017). Their main reasons for visiting Zeeland is to experience nature, food and visit the coast (Toekomst Schouwen-Duiveland, 2017). The Rugvin Foundation is a Dutch foundation that monitors the harbour porpoise population in the Eastern Scheldt. In cooperation with National Park the Eastern Scheldt and the World Wildlife Fund (WWF) in the Netherlands, the Rugvin Foundation recently installed 'Studio Bruinvis' on one of the quays in Zierikzee, which is regularly visited by tourists (Stichting Rugvin, 2017). Studio Bruinvis consists of a mono hydrophone attached to a buoy at a designated porpoise feeding hotspot that is wirelessly linked to a listening post on the quay where visitors can listen to any harbour porpoise activity (clicks). This spot is known for its old and deep underwater ammunition depot which is thought to attract fish which serve as prey for the harbour porpoises (Stichting Rugvin, 2017; Eck *et al.*, 2001). The hydrophone records continuously with a range of up to 500m, providing good opportunity for monitoring harbour porpoise activity around the buoy (Zanderink, pers. comm., 2017).

Distribution and activity of species is often highly dependent on environmental conditions and resources (i.e. food abundance, temperature ranges, barriers, etc.) (Miller & Spoolman, 2009; Bearzi *et al.,* 2012). Monitoring such variables along with a species' presence and absence provides the opportunity to find correlations between these and creates the possibility to predict this. Previous studies have successfully produced models for cetacean distribution, abundance or possible suitable habitat using species distribution models or habitat suitability models (Markoglou *et al.*, 2015; Redfern *et al.,* 2006; Cañadas *et al.,* 2002; 2005; Cañadas & Hammond*,* 2006; Giannoulaki *et al.,* 2016; Gilles *et al.,* 2011). These models have incorporated many variables and tested the effects they have on the presence of a species. These include variables such as depth, slope, sediment type, tidal range, water currents, water temperature, salinity, and chlorophyll concentrations as well as biotic variables such as prey species abundance.

Models for harbour porpoises, in open waters such as the North Sea, have found that distribution varies per season and is affected by factors including steep bottom topography, total nitrogen & chlorophyll concentrations, sediment types, distance from shore, depth, water temperature and residual currents (Gilles *et al.,* 2011; 2016; Brookes *et al.,* 2013; Williamson *et al.,* 2017; Camphuysen & Siemensma, 2011; Macleod *et al.,* 2007). Despite the results of these models, such models were not yet created for the harbour porpoise population in the Eastern Scheldt. The Eastern Scheldt differs in environmental conditions in comparison to open waters such as the North Sea, which suggests a different way of habitat use by harbour porpoises (Jansen *et al.,* 2013).

Nowadays, many mobile applications (app) are available that help people to spot and identify wildlife in the field (Klopstra, 2013). Apps on mobile devices provide advantages to their users such as portability, location awareness, and accessibility. Furthermore, the low price encourages the purchase of such products (Nayebi, *et al.*, 2012). Many spotting or tracking apps let the users contribute to a wildlife database (citizen science) by adding information such as a photo, time and location of a spotted individual (Hann *et al*., 2018). Some apps are even used to predict the weather, or high and low tide. Surfers, for example, can then use this information to find the best waves in their area. These data are retrieved directly from an online source (Bryant *et al*., 2016). However, an app that combines weather and water conditions and calculates the chance of encountering a certain species at any given time hasn't been developed yet. The Rugvin Foundation wanted to know what the best time and environmental conditions were for actually spotting harbour porpoises at the designated hotspot to promote tourism, and an app predicting porpoise activity at Studio Bruinvis can act as a helpful tool to maximise the chance of encounters at Studio Bruinvis by visitors and to plan their trip accordingly.

An app that calculates the chance of encountering a harbour porpoise at Studio Bruinvis was created with the use of a prediction model. The Rugvin Foundation can use this app to advice and educate visitors of the national park when chances are highest for spotting porpoises at Studio Bruinvis. To make the app accurate and easy to use, the model needed to use variables that were accessible to the public and be kept up to date.

Based on previous studies, the relatively smaller scale of the present study and the data accessible for the use in an app this study has used the abiotic variables *water temperature*, *water height*, *tidal water flow*, *wind direction*, *wind speed*, and *time of day*. The study had two goals. The first goal was to find out how *water temperature*, *water height*, *tidal water flow*, *wind direction*, *wind speed*, and *time of day* affected harbour porpoise activity at Studio Bruinvis and to find out if it was possible to predict porpoise activity level (high or low) based on these variables. The second goal was to find out how this gained knowledge could be used for a mobile application and how to develop this app so that it was easy to use for tourists. To achieve these goals, the following research question had been formulated:

"To what extent is harbour porpoise (*Phocoena phocoena*) click *activity influenced by water temperature, water height, tidal water flow, wind direction, wind speed, and time of day around a commonly used hotspot* in the Eastern Scheldt and how can this information be integrated in a mobile application to predict harbour porpoise activity?"

This question has been divided in several sub-questions:

1. *What is the relation between harbour porpoise click activity and abiotic variables around Studio Bruinvis?*
2. *In what way can this information about porpoise activity be integrated into a user-friendly predictability application?*

This report provides the Rugvin foundation with a better understanding about harbour porpoise presence in the Eastern Scheldt. It also provides the Rugvin Foundation all needed information to build a mobile application to promote porpoise watching for tourists by presenting under which circumstances harbour porpoise activity is the highest, increasing the likelihood to observe them.

2. METHODOLOGY

2.1 Study area

The focus of this study lies within a 500m radius of the Studio Bruinvis buoy, as this is the limit for the hydrophone to record any harbour porpoise clicks. This buoy is located in the Eastern Scheldt near the harbour entrance of Zierikzee, (figure 3). Many recreational boats and fishing boats pass by this location. The listening post of Studio Bruinvis, that receives the audio from the hydrophone, is located on the nearby quay (figure 4). Many tourists pass by on foot or bike via a bicycle path and visit Studio Bruinvis. The location is known to be a hotspot for encountering harbour porpoises. It is also known for its old underwater ammunition depot and it is thought that this deep (approximately 53 meters, see appendix I) underwater pit attracts prey fish species, which in turn attract harbour porpoises (Zanderink, pers. comm., 2017).

The Eastern Scheldt (figure 3) is the largest national park of the Netherlands consisting of estuaries, dikes, and swamps; and is part of the Natura2000 network of protected areas of the EU (Nationaal Park Oosterschelde, 2017; European Commission, 2017). The area is versatile and consists of many ecosystems due to actions of salt and fresh water fluxes, currents and tides (Nationaal Park Oosterschelde, 2017). Besides harbour porpoises, the area is also an important habitat for other marine mammal species such as the grey seal (*Halichoerus grypus*) and the harbour seal (*Phoca vitulina*), which use the sandbanks as haul-out sites. The area is also a very important foraging and breeding site for many species of birds, particularly migratory wader birds such as the bar-tailed godwit (*Limosa lapponica*) and the red shank (*Tringa totanus*) (Nationaal Park Oosterschelde, 2017; Dekker, 2016).

Figure 3: The location of Studio Bruinvis pole and buoy in the Eastern Scheldt near Zierikzee

The difference in water height between tidal phases in the Eastern Scheldt can be around 285cm. (Scheijgrond *et al.,* 2000). Tidal differences can cause current speeds of up to 5m/s in coastal zones and between islands (Scheijgrond *et al.,* 2000). Studio Bruinvis is located in the city of Zierikzee (figure 3). In Zierikzee the coldest month (January) usually has mean temperatures between 1°C and 6°C and between 12°C and 23°C in its warmest months (July & August) (Meteoblue, 2017). Mean monthly precipitation is between 38-67mm, and wind speeds between 8-36km/h (Meteoblue, 2017).

2.2 Study population

The study population consists of the clicks produced by the harbour porpoises in the Eastern Scheldt, which are recorded by Studio Bruinvis per hour. As described previously, the most recently counted harbour porpoise population in the Eastern Scheldt consists of at least 34 individuals, but annual numbers tend to fluctuate (Stichting Rugvin, 2017). The hydrophone detects porpoise clicks within a radius of 500m around the buoy (Zanderink, pers. Comm., 2017). It does not distinguish between individuals, meaning that the hydrophone cannot connect clicks to individual harbour porpoises, nor can it tell how many porpoises are vocalizing.

2.3 Data sampling

This research is a longitudinal observational survey. The focus lies on using collected data without intrusion of the subject sample and only searches for correlations between the variables and not causalities (i.e. experimental surveys). Since the beginning of August 2017, activity data of harbour porpoises is continuously being recorded by Studio Bruinvis. Data was collected from August 15th, 2017 until October 5th, 2017.

2.4 Data collection

2.4.1 Activity data of harbour porpoises

In the context of observing harbour porpoises, presence does not equal animal observability, since the harbour porpoise is notoriously elusive (Camphuysen, 2004). Therefore, it was chosen to make use of activity level, with the theory that higher activity leads to better chance of observability (Nuuttila *et al*., 2013; Pierpoint, 2008). Activity level was defined by the number of clicks produced by harbour porpoises, with high numbers of clicks suggesting high activity (Nuuttila et al., 2013; Todd et al., 2009). Click data of harbour porpoises was collected using the equipment of Studio Bruinvis. Studio Bruinvis consists of a buoy equipped with a hydrophone, click detector, antenna and solar panels. The hydrophone is programmed in a way that it only collects audio data within a certain frequency range and within a maximum distance of 500m. The frequency range is set at a frequency (between 100 – 140 kHz) that harbour porpoises use for echolocation to navigate, forage or communicate (Rodrigues, 2014). Audio data collected by the hydrophone is transmitted to a listening post that is located on a nearby quay (figure 4 and 5). Inside this column, there are several devices (installed by Waterproof B.V.) which store audio files 24 hours a day, 7 days a week. Data is saved on a 256GB USB flash drive

Figure 4: Listening post for tourists of Studio Bruinvis and Nationaal Park Oosterschelde that holds recording equipment inside the post (Hoeneveld, 2017)

inside the listening post. Once a month, this USB flash drive is replaced by an empty flash drive. Every day, 24 audio files are saved that last 59 minutes and 59 seconds.

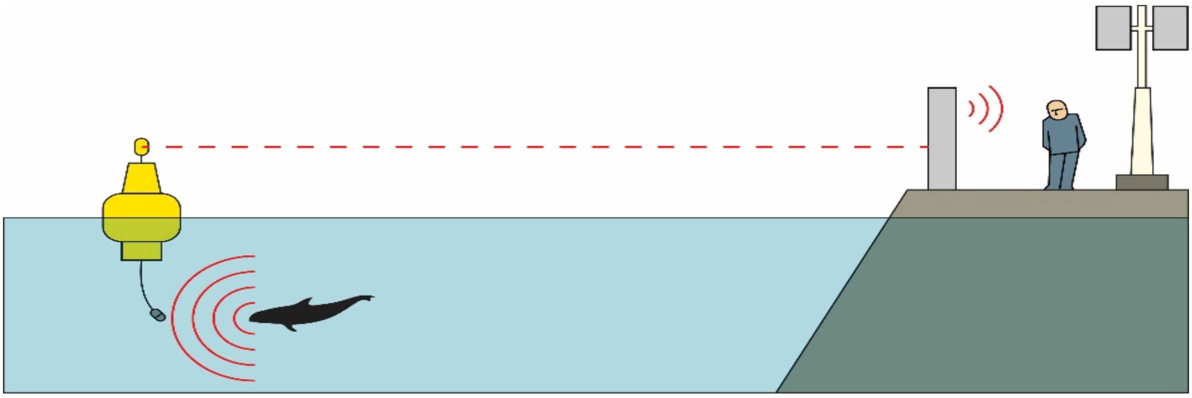


Figure 5: Schematic overview of Studio Bruinvis (Stichting Rugvin, 2017)

2.4.2 Time of day

A study done by Williamson *et al.* (2017) in Scotland, found that harbour porpoises migrate between areas depending on the time of day. Each audio file that was recorded by the Studio Bruinvis hydrophone was saved with a code that indicates the starting time of that audio file. For example, audio file name 20171908101523, started recording at 10:15:23 on the 19th of august in 2017. Because audio files were stored in files lasting 59:59minutes, it was chosen to use hour-blocks. For example, 11:00 o'clock till 12:00 o'clock is one-hour block. Audio file names are used to determine which recorded audio belongs to which hour block.

2.4.3 Water temperature

In Western Scottish and Danish waters, water temperature has shown to correlate with presence of harbour porpoises and is thought to affect prey distribution in warmer waters (Macleod *et al.,* 2007; Sveegaard, 2011). Data on water temperature was obtained from Rijkswaterstaat (Table 1). In the Eastern Scheldt, Rijkswaterstaat has placed devices at a depth of 4m below the Amsterdam Ordnance Datum (*NL: Nieuw Amsterdams Peil*) that measure the temperature of the water with an accuracy of one decimal degree every ten minutes (Rijkswaterstaat (b), 2017). Water temperature was measured at three locations (figure 6, locations with number 3). Distance from all three measure locations to study site were almost equal, but temperatures differed slightly. To get the most accurate temperature measurements for Studio Bruinvis, it has been chosen to use the mean of all three locations.

2.4.4 Water height & tide

Tide phases and water height have also been found to correlate with harbour porpoise densities in the Bay of Fundy in Canada, with higher densities during flood phases (Johnston *et al.,* 2005). It is suggested that this correlation is also likely due to changes in prey distribution (Johnston *et al.,* 2005; Embling *et al*., 2010; IJsseldijk, 2013). Currents have shown to influence the distribution of harbour porpoises directly or indirectly (Mikkelsen *et al*., 2016; Sveegaard, 2011; Waggitt *et al*., 2017). Johnston *et al.* (2005) found that areas with strong currents aggregate fish species that could serve as prey for harbour porpoises. Furthermore, the direction of currents influences the characteristics of the water as this could come from the North Sea or from runoff waters, which could also affect prey distribution and abundance due to a variation in nutrient supply (Cyprus and Blaber, 1992). Although the Eastern Scheldt has been closed off from the North Sea by a storm surge barrier, it is still influenced by tidal changes and currents through open gates (Rijkswaterstaat (a), 2017). Data on water height was obtained from Rijkswaterstaat (table 1). Water height was measured near Stavenisse (figure 6, location 2), approximately 9km from the study site, for every ten minutes in whole centimetres in reference to the Amsterdam Ordnance Datum. Data of water height was also used to determine tidal phase and categorized accordingly (see paragraph 2.3 Data preparation).

2.4.5 Wind speed & direction

Winds can affect currents: they can push water in certain directions (i.e. wave currents), affecting its flow directions and speed (van Ettinger & de Zeeuw, 2010). Currents in turn, may affect porpoises directly or indirectly by affecting prey distribution as described above (Gilles *et al.,* 2011). Wind data was also obtained from Rijkswaterstaat (table 1) and was collected near the Zeelandbrug/Zierikzee (figure 6, location 1), approximately 1.5km from the study site. Wind speed data was collected every ten minutes in meters per second with an accuracy of 2 decimals. Wind direction was also recorded every ten minutes and noted in degrees with reference to the North with an accuracy of one decimal degree.

Table 1 Overview of collected data used in this research

What/variable	When	Where
1 Number of clicks	Constantly from 15/8/2017 till 5/10/2017	At the buoy of Studio Bruinvis (see figure 6, Studio Bruinvis)
2 Time of the day	Constantly from 15/8/2017 till 5/10/2017	At the buoy of Studio Bruinvis (see figure 6, Studio Bruinvis)
3 Water temperature	Every 10 minutes from 15/8/2017 till 5/10/2017	At Plaat van Oude Tonge, Oosterschelde 4 (Storm surge barrier) and Marollegat (see figure 6, locations with number 3)
4 Water height (tide)	Every 10 minutes from 15/8/2017 till 5/10/2017	At Stavenisse (see figure 6, location 2)
5 Wind speed	Every 10 minutes from 15/8/2017 till 5/10/2017	At the Zeelandbrug (see figure 6, location 1)
6 Wind direction	Every 10 minutes from 15/8/2017 till 5/10/2017	At the Zeelandbrug (see figure 6, location 1)

Figure 6: Location of data collection points in the Eastern Scheldt (Rijkswaterstaat, 2018). Points represent measure locations for porpoise activity data (Studio Bruinvis), wind speed & direction (1), water height (2) and water temperature (3). (Google, 2018)

2.4.6 Data collection for the app

Information about other apps and which designs work best were collected through literature review (Bergner & Leonhardt, 2017; Bryant *et al*., 2016; Gazdecki, 2018; Hann *et al*., 2018, Klopstra, 2013; Küchler, 2016; Luchtmeijer, 2017) and the exploration of other apps (Appendix II). An app consists of two main components, the front-end and the back-end (Küchler, 2016). The front-end is what the app displays to the users, while the back-end is the program that runs in the background. The main idea, layout and design (front-end) lie in the hands of the authors. A good app should provide several features. The app should be the solution to a problem (Luchtmeijer, 2017). In this case the goal was to give the visitors of Studio Bruinvis information on the chance of encountering an active harbour porpoise at the hotspot. To give the user a good overview and experience the app design should be simple and user friendly (Luchtmeijer, 2017). Therefore, the app only contains important features and additional explanations if necessary.

To program an app, it is important to be familiar with such software. The help forum of Appinventor2 and many YouTube videos were used to program an app in appinventor2. In order to connect some buttons to websites, the respective links were included in the app. The websites of the Rugvin Foundation, their Facebook, twitter and Instagram page and the website of Rijkswaterstaat were used as links in the app to acquire further information. The Rugvin Foundation website and their complementary social media sites give the opportunity to explore and follow the Rugvin Foundation and their activities, whereas Rijkswaterstaat provides the measurements which are used for the calculation in the app (wind speed, wind direction, water temperature, water height and the tides). In addition, the information presented in the app was collected from the website of the Rugvin Foundation (www.rugvin.nl). General information about the harbour porpoise was based on publications by Hammond *et al.*, 2008 and Shirihai & Jarret, 2008. Illustrations were created, and photos were provided by the Rugvin Foundation. An audio file from the dataset was used to demonstrate sounds recorded by Studio Bruinvis.

2.5 Data preparation

The audio files obtained from Studio Bruinvis were analysed using different software to obtain the number of clicks per hour-block. First, all audio files were visually and audibly inspected using the software program Audacity to find possible distortions that might interfere with click detection. Audacity is a simple audio software program that can be used to listen, explore and edit audio files (Audacity, 2018). Secondly, the files were analysed using the program Pamguard. This software is used for acoustic monitoring of whale and dolphin species (Pamguard, 2017). Special settings to detect harbour porpoises were used to detect clicks at frequencies that harbour porpoises use (Appendix III). These settings were provided by the company WaterProof B.V. Unfortunately, Pamguard could not distinguish between porpoise clicks and environmental or marine vessel distortion (false clicks) because the decibels (dB) of these sounds were too similar (figure 7). By setting the dB level in Pamguard to 11 dB it was possible to avoid some unwanted distortion noise. This however did not solve the problem completely and Pamguard still detected high volume noises as porpoise clicks. Further attempts to filter out unwanted noises in Pamguard were unsuccessful. Therefore, files containing brief periods of interfering distortion were cut out of the files using Audacity to avoid further analysis of the files to lead to the count of false clicks. This resulted in audio files with a shorter duration. To compensate for this, the average number of clicks per minute was calculated for edited files and added to the total number of clicks according to the amount of time removed. Audio files that were deemed too distorted to edit (figure 7), were excluded from the dataset.

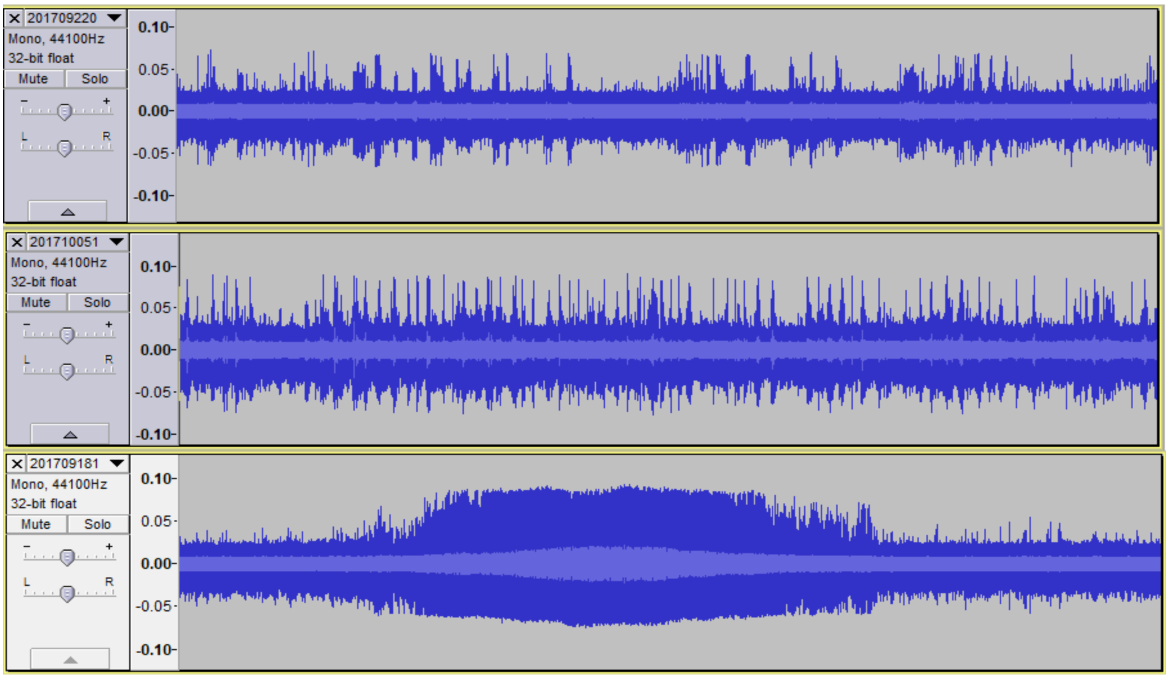


Figure 7: Top image shows an audio file containing harbour porpoise clicks causing high decibel sounds indicated by peaks. Vertical scale indicates decibels. Second image shows an audio file with strong winds causing high decibel sounds indicated by peaks. Bottom image shows passing motor boat causing high decibel sounds indicated by broad peak.

Several data sheets were made using Excel. The first sheet (appendix IV) consisted of data of all the measured variables (for every ten minutes) as described in data collection. This was used to calculate all the hourly means for each variable. The hourly average water temperature was calculated using the three measuring locations (figure 6). Since the distance between the measure locations and to Studio Bruinvis was relatively equal, no weighted average was used.
Water height data in this sheet was also used to determine tidal phase. In total, 3 types of tide variables were created (4 categories; 8 categories, and 14 categories) to see which variable scored best on the predictive model performance according to the Akaike's Information Criterion (AIC). The AIC is a measure of how well a model fits a dataset, while adjusting for the ability of that model to fit any dataset whether or not it's related (Claeskens & Hjort, 2008). Tides where divided into several variables based on where within the hour the tide shifted (Appendix V). For example, if the tidal phase in an hour consisted entirely of outgoing water it was categorized as "1", if the hour consisted entirely of incoming water it was categorized as "2". For tides with 14 categories, if tides shifted within the first 10 minutes of the hour it was categorized as "3"; if the tidal phase shifted within the 10-20 minute mark it was categorized as "4"; etc. (Appendix V). This was done due to the time span of low and high tide, which usually lasted for around 10 minutes. The same was done for tides with 8 categories, but with a 20 minute timeframe instead of 10 minutes. The simplest variable was the tides divided into 4 categories, which consisted of incoming and outgoing water, as well as hours where high tide or low tide occurred. These last 2 variables were created to increase observations for categories that possibly contained too few, which could improve the model performance. The middle of the approximate 10 minute duration of high or low tide was used to define in which part of the hour it was. For example, low tide occurred from 13:14 to 13:24, the middle of this timeframe would be at 13:19 which would be the mark used to assign it to a category.
For wind direction, 2 variables of different categories were also created to see the effects of different number of categories on model performance. The first contained 4 categories of 90°, and the second contained 8 categories of 45°. Lastly, time of day was categorized into variables containing 24 categories of separate hours, 12 categories of 2 hours, and 6 categories of 4 hours each to compare which had the best effect on model predictions.

The second sheet (Appendix IV) consists of the data output from Pamguard and contains each detected click along with the exact time and date of recording. This sheet was used to assign clicks to their recorded hour-blocks whenever audio files overlapped consecutive hours in a day, and also to compensate missing time due to distortion. The third sheet was used to combine sheet 1 and 2 into a single file with all the data for each hour-block (Appendix IV). Appendix V shows an overview and code book of all used variables of the third sheet.

A sample size of approximately 75% of the total dataset was used for the model, leaving 25% of the dataset for validation of the model (Snyder, 2015). To avoid possible effects of consecutive hour blocks, for securing the independency of data, and to make sure all hour-blocks were represented equally within the sample, a stratified random sample was taken. From all samples of each hour-block, 27 random samples were taken.

The final step for preparing data for data analysis was setting the baseline for number of detected clicks that determines harbour porpoise activity. Three baselines were chosen, 400 clicks, 1000 clicks and 1500 clicks to compare model predictions.

2.6 Data analysis

2.6.1 Data exploration

During the study period (from 15th of August in 2017 until the 5th of October in 2017) a total of 1248 audio files, meaning 1248 hours of recorded audio, were obtained from Studio Bruinvis. Analysis of a single audio file took between 5-25 minutes depending on the number of clicks in each audio file. During the exploration of the audio files, it became apparent that around 30% of the audio files were distorted with noise, mostly caused by wind or passing boats with a running motor. This resulted in 834 suitable audio files that were analysed with Pamguard. Stratified random sampling resulted in a total sample size of N=657, which was used as the dataset to build the predictive regression model and to answer the research questions.

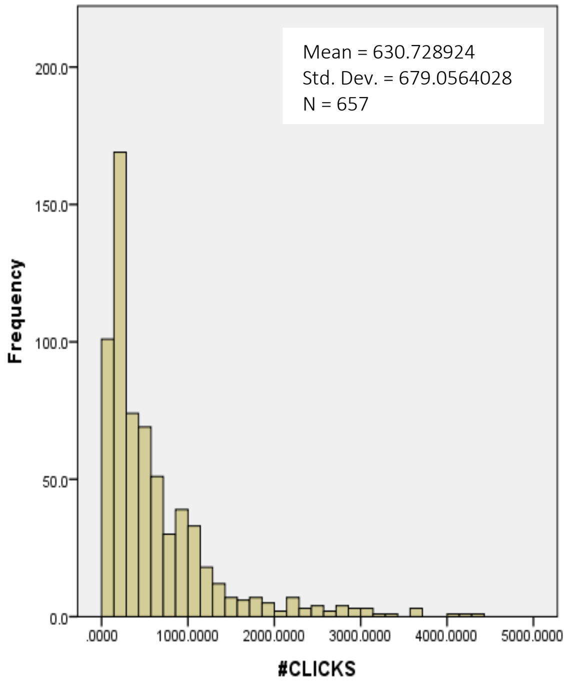


Figure 8: Frequency graph of detected harbour porpoise clicks

The remaining data, a total of 177 hour-blocks, was used afterwards to validate the model's predictive power. The number of clicks ranged between 43 and 4418 clicks, with a mean of 630 clicks per hour (figure 8). Other frequency graphs of variables over time (i.e. the entire study period and means per hour) and mean clicks against variables can be found in appendix VI.

Data from the third Excel sheet as described previously was imported into IBM SPSS v.20. Firstly, the sample dataset was explored using frequency bar charts and boxplots to check distribution over time and distribution of clicks per variable (Appendix VI) (Zuur *et al.,* 2010). Water height and water temperature showed a bimodal distribution over time. Wind speed and wind direction showed skewed distributions. To remove outliers, the square root of wind speed was taken and used as a variable (Lewis-Beck, 1995).

To check for multicollinearity between independent variables, all variables were explored using scatterplots and boxplots (Zuur *et al.,* 2010). In addition, collinearity diagnostics' Variance Inflation Factor (VIF) values for all continuous covariates were calculated. VIF measures the impact of collinearity among covariates in a regression model and is always ≥1. Values of >3 suggest possible collinearity issues, which would make it difficult to assign independent covariate effects (Zuur *et al.,* 2010; Dormann *et al.,* 2013). The collinearity diagnostics test showed low VIF scores for all variables (VIF <1.156), which implied no multicollinearity between variables in this study. Therefore, all tested variables were included in building the predictive model.

2.6.2 Generalized Linear Model

A binary logistic regression was used to assess the relation between harbour porpoise clicks per hour together with *water temperature*, *water height*, *tidal water flow*, *wind direction*, *wind speed*, and *time of day* using IBM SPSS v.20. The independent variables in this model were water temperature (continuous covariate), water height (continuous covariate), tide (categorical covariate), hour (categorical covariate), wind direction (categorical covariate) and the square root of wind speed (continuous covariate). The set baseline, representing harbour porpoise activity (0=low activity, 1= high activity), served as the dependent variable. Several baselines were tested to build a model with high predictability. The first baseline was set to 400 clicks. All samples below 400 clicks were given the value 0, meaning low porpoise activity, thus low chance of spotting harbour porpoises at the study site. All samples above 400 clicks were given the value 1, meaning high porpoise activity, thus high chance of spotting harbour porpoises at the study site. This baseline was chosen based on 50% of samples. To test if the model improved by using higher baselines for activity level, the baseline for the model was also tested at 1.000 and 1.500 clicks.

The model's cut-value, which determines the test values as being high activity or not (Unal, 2017), was determined using the Receiver Operating Characteristic (ROC)-Curve for optimal sensitivity and specificity. Sensitivity is the model's ability to predict true positives (i.e. the predicted high activity levels are actual observed high activity levels), and specificity is the model's ability to predict true negatives (i.e. the predicted low activity levels are actual observed low activity levels) (Chan, 2004). Finally, the best model was chosen by Akaike's Information Criterion (AIC) values, which shows how well the model would operate outside of the used dataset (Claeskens & Hjort, 2008). The final model would produce a formula for calculating the probability of activity (<400 or ≥400 clicks) under various conditions for the included variables as follows:

$$1 / (1+e^{-z})$$

with *Z* being the sum of all included variables that were calculated for by multiplying model scores with logistic regression coefficients (Chan, 2004). For example, in a certain scenario, water temperature would be 10° C and water height would be 1,2m above NAP. If the model would only include these 2 variables, it would look like this:

Z= Constant + B (score for water temperature) x 10 + B (score for water height) x 1.2

The logistic regression coefficients (B-Values) for each variable are used in the equation, and made it possible to calculate and predict the chance of high or low activity (<400 or ≥400 clicks). This formula was then used as the basis for the mobile app.
The Nagelkerke pseudo R-square, was used to test the model fit: it calculates how the variation in the outcome variable (Clicks ≥400) is explained by the model (Chan, 2004). The Hosmer and Lemeshow goodness of fit test was used to shows how much the predicted events matched the observed events (Chan, 2004). A pairwise comparison was used to find differences in effect between individual categories of the categorical covariates, and the estimated marginal means were calculated to see the chance of observations with ≥400 clicks for each category.

2.6.3 Model baseline choice

When comparing models, which included all variables but with different baselines, the model with the baseline set to 400 clicks proved to have the best predictive quality. This model had approximately equal scores for sensitivity (accurately predicting negatives) and specificity (accurately predicting positives). Furthermore, all variables included in the model were found to have significant effect on porpoise click activity (P<0.05). Models with a baseline set to 1000 and 1500 had higher specificity and overall predictive scores, but extremely poor sensitivity scores. Since the ultimate aim of this research was to predict presence (true positives), models with very low sensitivity were not desirable. Additionally, very few variables in these models were found to be significant ($P<0.05$): the 1000 clicks baseline included wind speed, time of day and tides, while the 1500 clicks baseline only included tides as a significant variable. Therefore, the baseline of ≥400 clicks was chosen to indicate high activity level and <400 clicks would indicate a low activity level.

2.7 Creating an app to predict harbour porpoise presence at Studio Bruinvis

Despite popular believes, app development begins on paper. The first design ideas were sketched, including the buttons that link the different screens with each other or to reveal other functions, such as playing a sound, opening a website or displaying a keyboard. A literature review and the exploration of other (similar) apps gave an idea on what kind of apps are already out there and which designs are used. Inspired by some designs, the sketches were adjusted accordingly. To create the user interface (UI) in a more presentable way, the website mockflow (https://mockflow.com/) was used. The UI is part of the front-end of the app, thus the part the user sees and clicks. In the wireframe, the storyboard of the app, each page and buttons are defined on paper or digitally. Based on the wireframe a mock-up version was designed using marvelapps (https://marvelapp.com/) that allows the switch from one page to another by clicking the action areas that serve as buttons. The app was programmed with MIT's appinventor2 (http://ai2.appinventor.mit.edu/) which gave the possibility to run the app via an app simulator on Android phones and adjust the design and functions. The main design of the calculator screen was found quickly and only the details needed to be adjusted according to the outcome of the prediction model. The first design for the main menu, which was composed of tabs, seemed to be confusing, overloaded with text and difficult to program. It was changed to a design that provided a picture for each section (Studio Bruinvis, Rugvin Foundation, harbour porpoise calculator, harbour porpoise) and revealed the information on another screen after being clicked. Creating an app is a process that circles back and forth between planning, testing implementing and adjusting the front-end as well as the back-end (figure 9). Especially the programming part took a lot of time. It is important to program and name each button and link it to a specific task. In addition, the task performed by the app has to be built in a specific order with the building blocks. Furthermore, illustrations were designed and photos from the Rugvin Foundation website were used as graphic elements and descriptions were added.

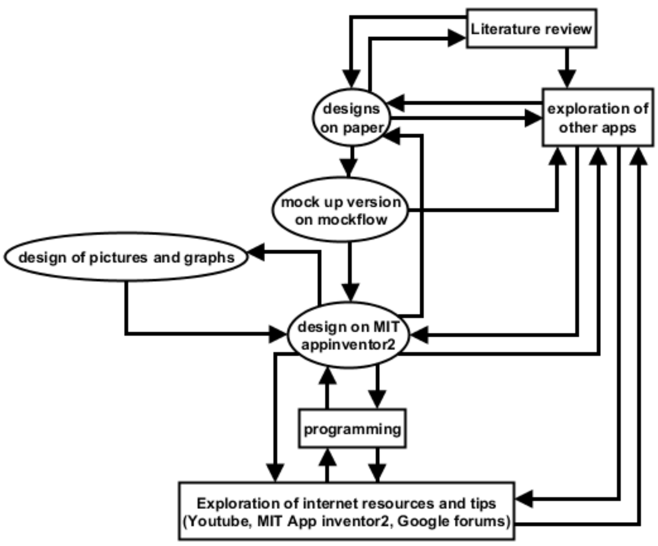


Figure 9: The app creation process is a workflow that goes back and forth. Input is displayed as rectangles. The ellipses represent the output which can also serve as input for the next step.

3. RESULTS

As described above, a binary logistic regression model was created to test which variables were significant in affecting the chance of high activity (clicks ≥400), which would be the best opportunity to observe harbour porpoises at Studio Bruinvis. Furthermore, the variables were tested individually to see their relation to activity level.

3.1 Variable effects on porpoise activity levels

The outcome of the binary regression can be found in Appendix VIII, which shows the variables included in the model with their respective values. Akaike's Information Criterion (AIC) showed that the best model (AIC: 730.399) included the tides variable categorized into 8 categories, time of day categorized into 6 categories of 4 hours, and wind direction categorized into 8 categories of 45°. Appendix VIII shows that all variables (apart from several separate categories) were significant (P<0.05) for the prediction of clicks observed ≥400, and thus porpoise activity. Therefore, all variables were included into the model.

The Wald Chi-squared test, which expresses the relative importance of variables as a value, found that the most important variable for predicting porpoise activity was Tide (Wald χ^2= 85.281, df=7, P=0.001). The estimated marginal means (EMM) shows the chance of observations with ≥400 clicks for each category with their standard errors. The highest chance for ≥400 clicks were outgoing-to-low tide (category 5) and incoming tide (2), followed by incoming-to-high tide (8), outgoing-to-low-to-incoming tide (4), low-to-incoming tide (3), outgoing tide (1) & incoming-to-high-to-outgoing tide (7), and lastly high-to-outgoing tide (6) (figure 10; Appendix X). A pairwise comparison between categories showed significant factor differences between tide category 1 and 2 (P=0.001); 1 and 5 (P=0.001); 2 and 3 (P=0.004); 2 and 6 (P=0.002); 2 and 7 (P=0.001); 3 and 5 (P=0.024); 5 and 6 (P=0.017); 5 and 7 (P=0.012). Chances for high

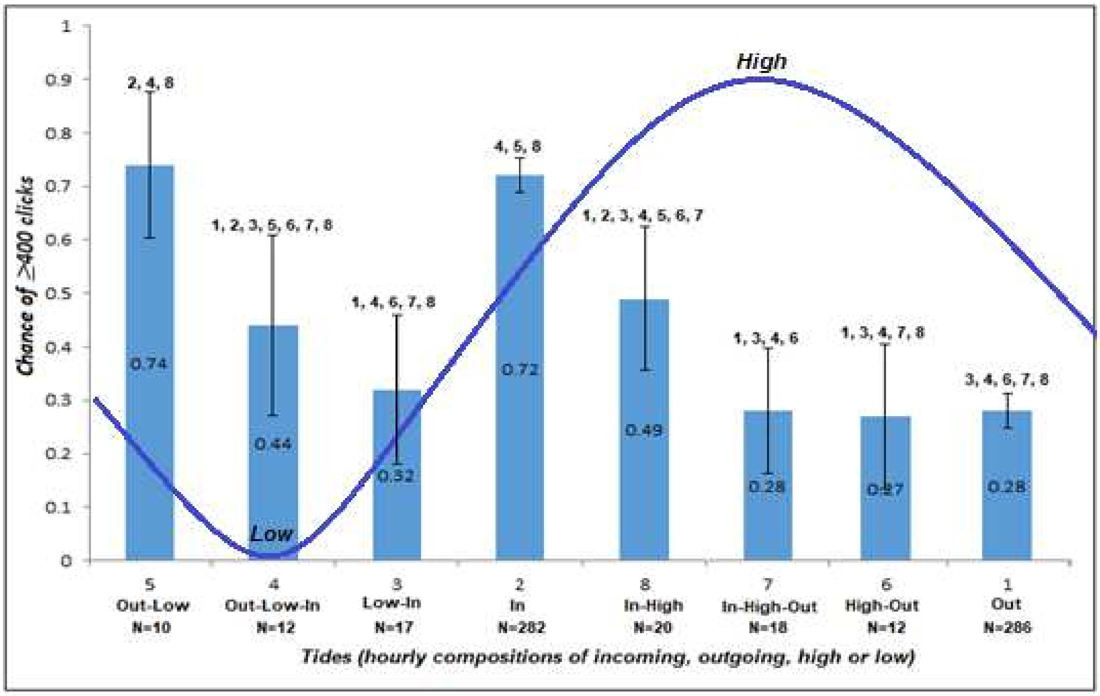


Figure 10: Estimated marginal means for the chance of ≥400 clicks in an hour-block for the 8 Tides categories with their standard errors and number of observations (N). The means are determined with a square root of Wind speed of 2.13 m/s; a Water temperature of 18.5°C; and a Water height of 2.98m.Numbers above bars indicate no significant differences between those categories and the respective category. The line indicates water height for the tidal phases.

activity were spread somewhat between categories; however, this could be due to a relatively low number of observations (N) of some categories. In general, incoming tide showed a high chance for high activity levels, while outgoing tide showed a much lower chance.

The second highest predictor variable was wind direction (Wald χ^2= 33.807, df=7, P=0.001). EMM shows the highest chance for observations with ≥400 clicks from winds coming from 335°-270° (category 6), 0°-45° (1), 90°-135° (3), 45°-90° (2), 270°-315 (7), 180°-225° (5), 315°-360° (8), and 135°-180° (4) (figure 11; Appendix X). Pairwise comparisons between categories showed significant differences between categories 1 and 4 (P=0.008); 1 and 5 (P=0.032); 1 and 8 (P=0.039); 2 and 6 (P=0.008); 3 and 4 (P=0.039); 3 and 6 (P=0.032); 4 and 6 (P=0.001); 5 and 6 (P=0.001); 6 and 7 (P=0.001); 6 and 8 (P=0.001). It is puzzling that winds coming from 335°-270° are found to have the highest chance of high activity, while winds coming from surrounding directions are found to have the lowest chance.

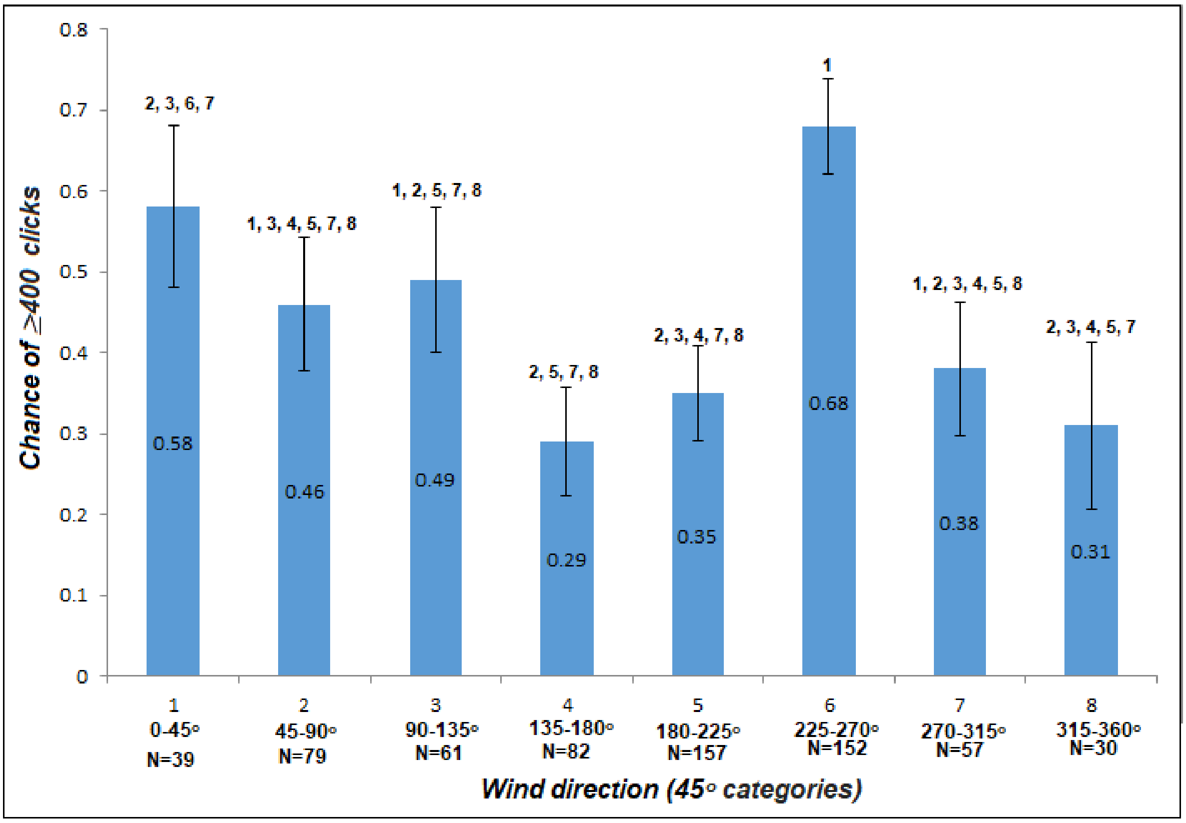


Figure 11: Estimated marginal means for the chance of ≥400 clicks in an hour-block for the 8 Wind direction categories with their standard errors and number of observations (N). The means are determined with a square root of Wind speed of 2.13 m/s; a Water temperature of 18.5°C; and a Water height of 2.98m.Numbers above bars indicate no significant differences between those categories and the respective category.

The third highest predictor was time of day (Wald χ^2= 27,274, df=5, P=0.001). The EMM shows that the highest chance of observations with ≥400 clicks are between 12:00-16:00h (category 4), 8:00-12:00h (3), 4:00-8:00 (2), 0:00-4:00h (1) and 16:00-20:00h (5), and lastly between 20:00-24:00h (6) (figure 12; Appendix X). Pairwise comparisons between categories showed significant differences between categories 1 and 3 (P=0.002); 1 and 4 (P=0.001); 2 and 3 (P=.034); 2 and 4 (P=0.019); 2 and 6 (P=0.011); 3 and 5 (P=0.001); 3 and 6 (P=0.001); 4 and 5 (P=0.002); 4 and 6 (P=0.001). The chances of high activity levels were thus predicted to be highest between 8:00-16:00h and lowest at night between 20:00-24:00h.

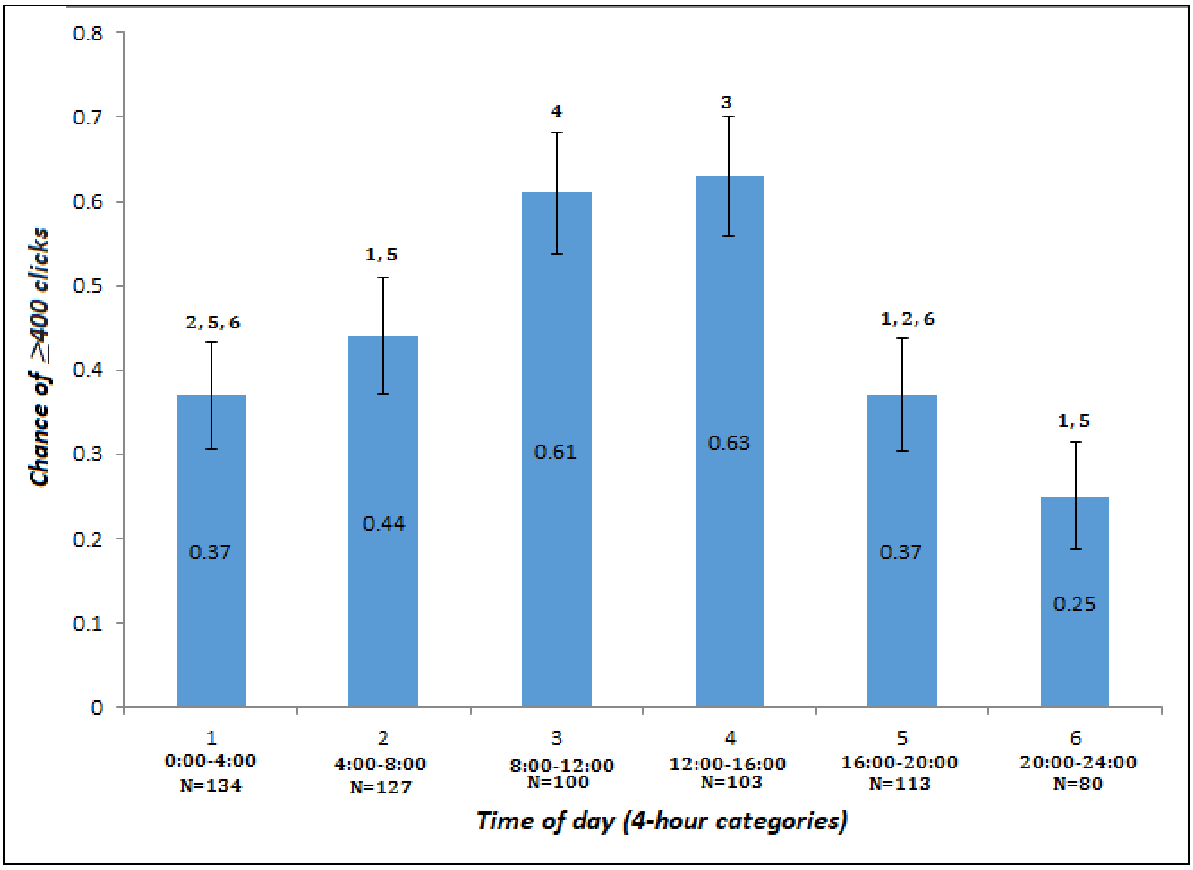


Figure 12: Estimated marginal means for the chance of ≥400 clicks in an hour-block for the 6 Time of day categories with their standard errors and number of observations (N). The means are determined with a square root of Wind speed of 2.13 m/s; a Water temperature of 18.5°C; and a Water height of 2.98m.Numbers above bars indicate no significant differences between those categories and the respective category.

The square root of wind speed was the next highest predictor (Wald χ^2= 25,451, df=1, P=0.001). The logistic coefficients (B=1.063) showed that there is a positive relation between wind speed and chance of high activity. The odds ratio (Exp(B)=2.896) showed that for 1 unit of wind speed (meter/second) increase, the odds of predicting high activity would increase with a factor of 2.896 (95% C.I. lower: 1.916, and upper: 4.378) (Appendix VIII). Water height (Wald χ^2= 13,808, df=1, P=0.001) also had a positive effect (B=.189), with an odds ratio of Exp(B)=1.004 (95% C.I. lower: 1.002 and upper: 1.007)per 1 cm height increase. Lastly water temperature (Wald χ^2= 10,543, df=7, P=0.001) had a positive effect (B=.004) on the chance of high activity, with an odds ratio of Exp(B)=1.208 (95% C.I. lower: 1.078, and upper: 1.353) per 1 degrees of Celsius increase (Appendix VIII).

The ROC curve, used to find the optimal cut-value for model predictive power, found that the optimal sensitivity and specificity were both at around 0.75. In the coordinates of the curve table this corresponded to a cut-value of 0.483, which was then chosen as the cut-value for the model (Appendix VII). The final model showed a specificity of 73.9%, a sensitivity of 74.1%, and an overall predictive percentage of 74% (table 2) meaning that 73.9% of all negatives (clicks <400) and 74.1% of all observed positives (clicks ≥400) were accurate, leading to an overall correct prediction of 74% of all cases. Since the model's base predictions based on the cut-value were 48.3%, the addition of the variables in the model boosted predictions with 25.7%.

Table 2: Observed and predicted values from the Logistic Regression Model

		Predicted Clicks		Percentage Correctly Predicted
		Clicks <400	Clicks ≥400	
Observed Clicks	Clicks <400	243	86	73.9%
	Clicks ≥400	85	243	74.1%
Overall		329	328	74.0%
The cut value is.483				

The final model generated the following equation for calculating the chance of clicks ≥ 400: $1 / (1+e^{-z})$ with *z* being:

z = -7.252 (constant) + 1.063 the square-root of Wind speed + .189*Water temperature + .004*Water height + 1.125*Wind direction category 1 + .627*Wind direction category 2 + .752*Wind direction category 3 + -.102*Wind direction category 4 + .206*Wind direction category 5 + 1.574*Wind direction category 6 + .313*Wind direction category 7 + -.898*Tide category 1 + 1.003*Tide category 2 + -.717*Tide category 3 + -.189*Tide category 4 + 1.116*Tide category 5 + -.924*Tide category 6 + -.888*Tide category 7 + .563*Hour category 1 + .867*Hour category 2 + 1.538*Hour category 3 + 1.618*Hour category 4 + .580*Hour category 5.*

3.2 Model reliability and validation

To test model reliability, the Nagelkerke and Hosmer & Lemeshow test were performed. The Nagelkerke pseudo R-square, which tests the model fit, showed that about 39% of the variation in the outcome variable (Clicks ≥ 400) is explained by the model (Appendix IX, table A) (Chan, 2004). Hosmer and Lemeshow goodness of fit test was at 63% (Appendix IX, table B), which shows how much the predicted events match the observed events (Chan, 2004).

As mentioned previously, validation of the model was done by using approximately 25% of the dataset, a set of 177 hour blocks, that was not included in the sample and had the amount of clicks for each hour-block removed without editing or removing the variables. The equation produced by the model was then tested to predict <400 clicks (0) or ≥400 (1) based on the variable data in the dataset. These predictions were then compared to the actual number of clicks of their respective hour-blocks to see the percentage of correctly predicted clicks. Of the 177 hour-blocks in the dataset 134 (75.71%) were correctly predicted while 43 (24.29%) were incorrect (Table 3). Of the 81 negative predicted values (NPV) with clicks<400, 21 were incorrectly predicted. Of the 96 positively predicted values (PPV) with clicks≥400, 22 were incorrectly predicted. Therefore, NPV: 60/81=74.07% and PPV: 74/96=77.08%, with an overall correctly predicted percentage of 75.58% (Table 3).

Table 3: Validation percentage PPV & NPV

Validation set size: 177		Predicted	
		Negative	Positive
Observed	Negative	60	22
	Positive	21	74
Percentage Correctly Predicted		74.07% (NPV)	77.08% (PPV)
Overall correctly predicted		75.58%	

Average confidence interval lower and upper bounds were 2.46% lower and 3.79%

higher, respectively. Meaning that, the percentage of chance of click activity being ≥400 calculated by the model equation above may actually be 2.46% lower or 3.79% higher than predicted.

3.3 Mobile application

Figure 13 Logo Rugvin Foundation and the app icon

The logo of the Rugvin Foundation was used as an inspiration for the app icon (figure 13). In general, all “back” buttons in this app return to the previous page when clicked. Furthermore, each screen with more information is arranged is a vertical scroll arrangement and can be viewed in horizontal and vertical orientation. Even though, the view is optimal at the vertical orientation in which most smartphones are used. External links are displayed as in *italic*, apart from the social media icons. To close the app, the “close” button or the home button of the smartphone can be used.

The first screen is simple with three buttons that allows the user to choose between three languages: Dutch, English and German (figure 14). Since Zierikzee is a popular tourist attraction, the languages reflect the most common tourist languages and of course Dutch for the locals.

Figure 14 First screen of the app. Choose a language

After choosing a language, the menu screen appears (figure 15). Within this menu, there are seven buttons to choose from. The four main buttons fill the screen in width, have a keyword as text and most of them are underlined by a complementary picture. The Studio Bruinvis button leads to a page that explains what Studio Bruinvis is and how it works (figure 16). The Harbour porpoise activity calculator leads to the core of the app, where the calculation takes place (figure 19). The Harbour porpoise button leads to a page where more information about harbour porpoises in general can be found (figure 17). The Rugvin button opens the screen with information about the Rugvin Foundation (Stichting Rugvin) (figure 18). Icons of three social media providers (twitter, facebook and Instagram) are located on the bottom of the page. These are also buttons which open the respective webpages of the Rugvin Foundation. In addition, the copyright information was placed at the bottom of the page.

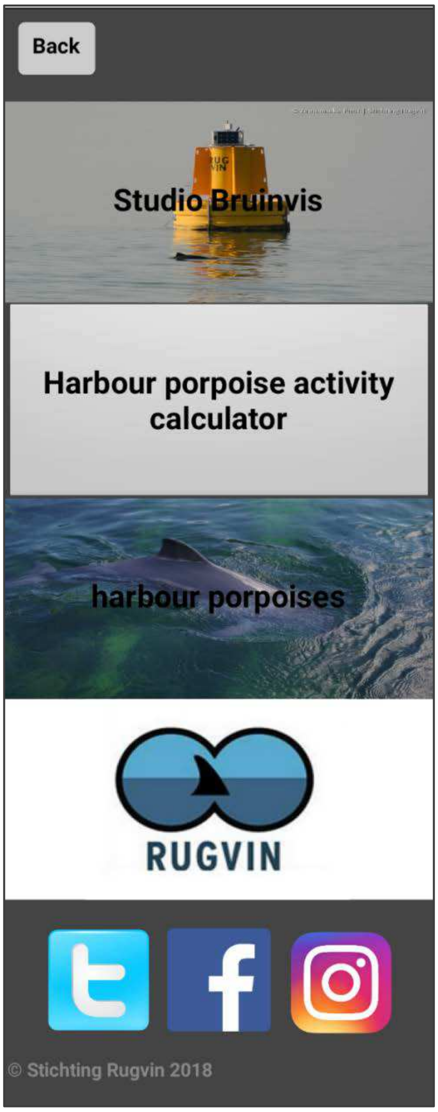


Figure 15 menu

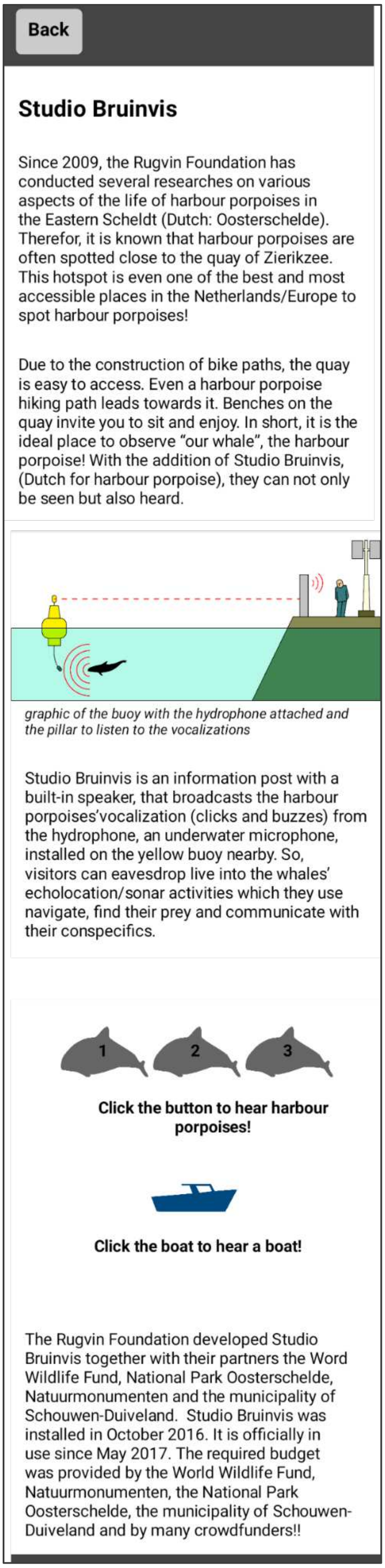


Figure 16: Studio Bruinvis information

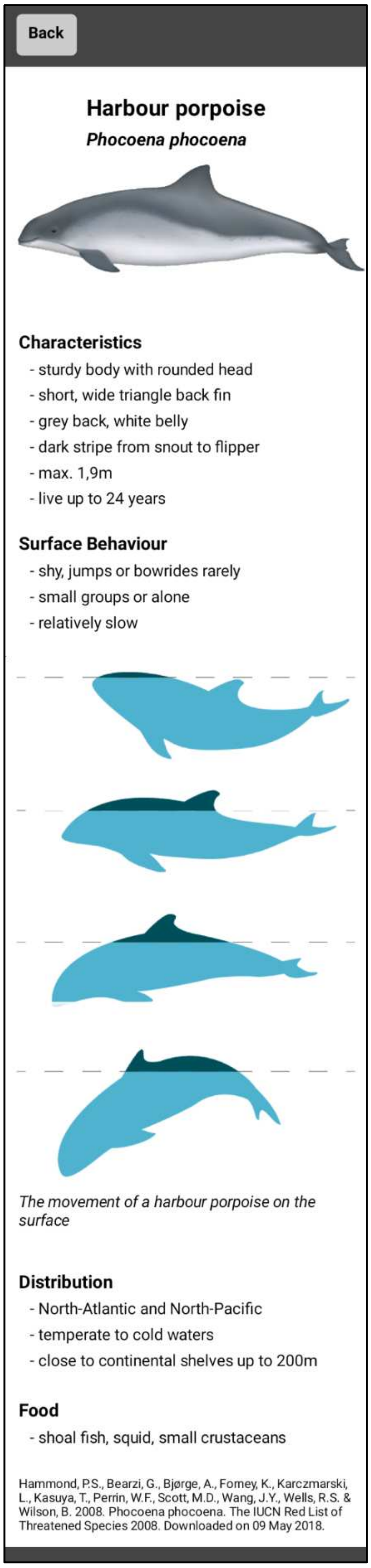


Figure 17: Harbour porpoise screen. How to recognise a harbour porpoise on the surface

Apart from Information about Studio Bruinvis, there are four additional buttons on this page. Each porpoise silhouette plays an example of a harbour porpoise vocalization. By clicking on the little boat, the sound of a boat engine as recorded by the hydrophone can be heard.

By clicking on the harbour porpoise button on the menu another screen opens (figure 17). Here, general information about the harbour porpoise can be found, including their main characteristics, size and diet. Movement patterns of harbour porpoises on the surface are included as these are important to recognize in order to spot and identify porpoises from a distance.

Apart from providing information about the Rugvin Foundation (figure 18), a link at the bottom of the page opens the homepage of the Rugvin Foundation in the Browser.

By opening the harbour porpoise activity calculator page, a screen displaying the variables included in the model appears (figure 19). To predict the activity in the future or under certain circumstances, the variables have to be entered manually.

Figure 18: The Rugvin Foundation

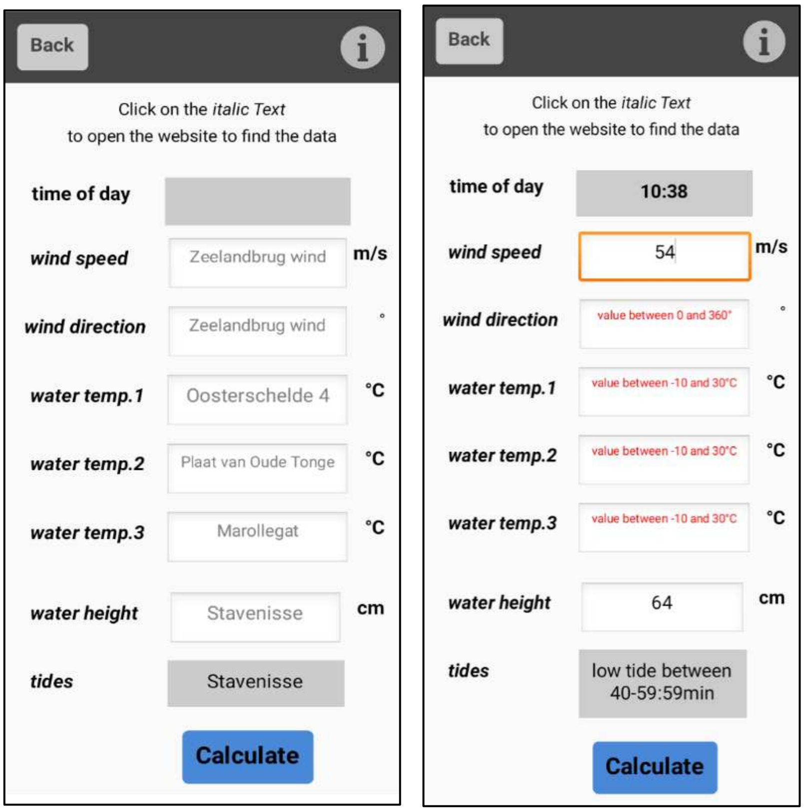


Figure 19: fill in the variable boxes to calculate harbour porpoise activity level in the Eastern Scheldt. If the values are invalid the textboxes display a warning and the range of the values.

Text displayed in *italic* lead to an external website of Rijkswaterstaat where the values for that specific variable can be found. The grey field next to "time of day" opens a time picker which sets the chosen time in an hh:mm format for a 24-hour mode. The button wind speed and wind direction open the website of Rijkswaterstaat that displays a map with the wind measuring points where the direction and the speed can be found in degrees (°) and meter per second (m/s), respectively. Water height and water temperature open a site provided by Rijkswaterstaat which measure the wind direction and the water height in the Eastern Scheldt. The user can fill in the number in centimetres (cm) and degrees Celsius (° C), respectively. Water temperature is calculated as the mean of three different measurement points. By clicking the tides' textbox, a multiple-choice list appears where one of the eight tide classes used in the model can be chosen. By clicking on the calculate button, a new screen opens displaying the result of the calculation (figure 20).

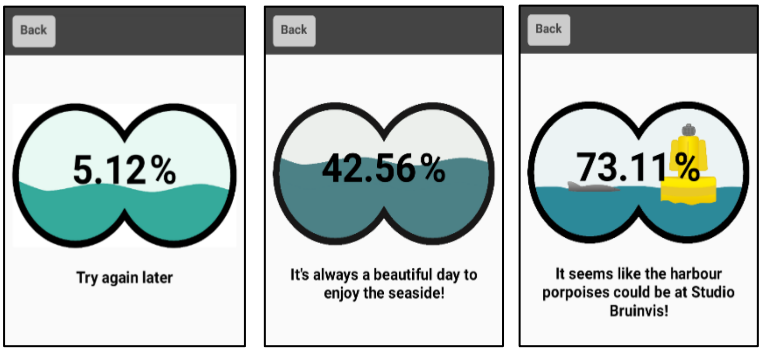


Figure 20: The three possible outcomes of the calculation

Depending on the score, the sentence under the result and the picture will change. For further explanation on the location of the measurement point and the navigation on the Rijkswaterstaat website, there is a little info button, the "i" in the top right corner, which opens the screen shown in figure 21.

Figure 22 offers an overview of the whole app layout. This layout was designed Android operating systems. The design, mostly the position of some buttons, might need to be adjusted slightly for apple and windows products.

In order to make the app popular under like-minded people, the channels, social media accounts and cooperation of the Rugvin Foundation are used. Furthermore, once the app is released it will be adjusted according to feedback and frequently extended and updated with innovative ideas such as an interactive map where citizen scientist can mark their cetacean sightings or a live stream from Studio Bruinvis.

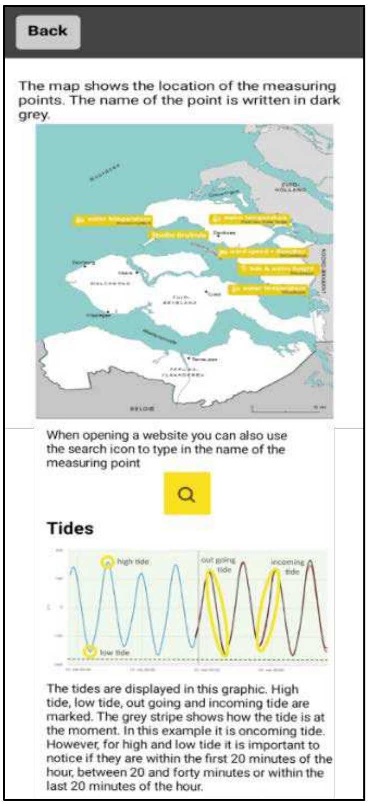


Figure 21: By clicking on the info button this page with further explanation opens

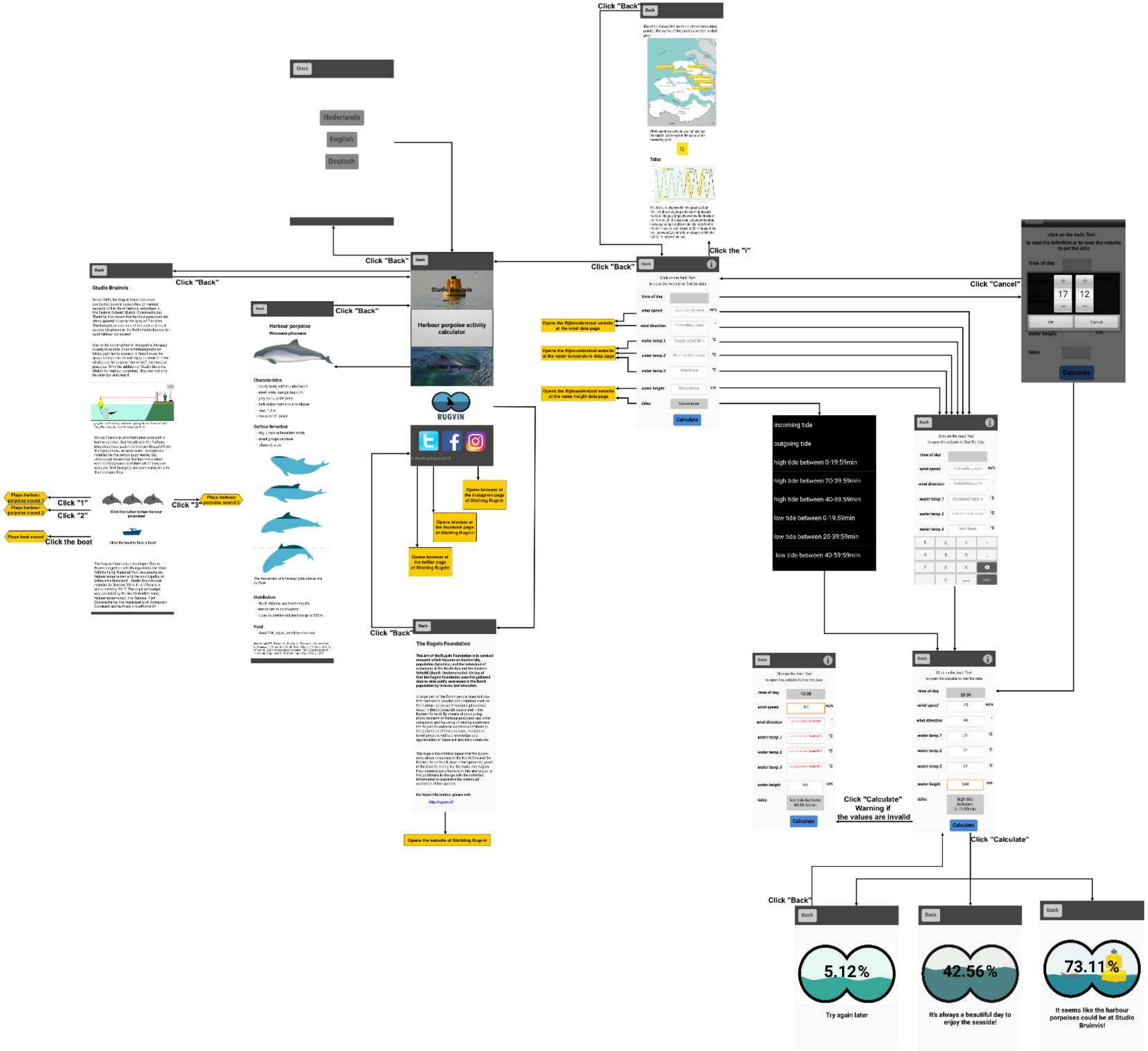


Figure 22: Overview of the app layout. Yellow Hexagon: plays sound; rounded rectangle: opens a webpage; Ellipse: result of the calculation. The arrows display the direction and order of the screens. Only the process of the English screens is displayed since it is the same in any language.

4. DISCUSSION

The first aim of this study was to find out how water temperature, water height, tidal water flow, wind direction, wind speed, and time of day affect the activity of harbour porpoise at Studio Bruinvis and if it was possible to predict their activity. The second aim was to find out how this gained knowledge can be used for a mobile predictability application and how to develop such a mobile application in a user-friendly way. To achieve these aims, most of the time was spend on collecting and preparing data. Gained knowledge of this study provides a better insight of harbour porpoise activity at Studio Bruinvis and functions as a guide for developing a predictability mobile application. All findings of this study can be used for future research, especially those focused on site specific and acoustic marine mammal research.

4.1 Predictability model

The results of the predictability model show that all variables had a significant role in predicting high or low activity of harbour porpoises. According to the Wald scores, the effect of tides was the strongest predictor. Despite this, individual tide categories scored relatively low in contributing to predicting activity. Strongest of the individual categories according to the estimated marginal means were *outgoing-to-low tide* and *incoming tide* followed by *incoming-to-high tide* and *outgoing-to-low-to-incoming tide*. A study done by Pierpoint (2008) in south-west Wales found that tidal currents combined with steep bottom topography create foraging opportunities for harbour porpoises, especially during low tidal phase. Other studies found similar results, wherein porpoise distribution and activity were explained by tidal phase, usually combined with steep bottom topography (Embling *et al.*, 2010; Marubini *et al.,* 2009; Isojunno *et al.*, 2012). As Studio Bruinvis is located on top of a deep underwater storage depot, tidal currents might have a similar effect by creating foraging opportunities. In contrast to this, IJsseldijk (2013) found that more harbour porpoises in the North of The Netherlands were detected during high tide, to which she describes porpoises as having a preference for high tide. Additionally, a study done in the United Kingdom by Goodwin (2008) found significant differences in porpoise behaviour between tidal phases in one area, but not in another. This leads to the conclusion that on a geographical scale, population can differ considerably, which may make interpopulation comparison difficult.

For wind direction, which scored highest after tides as a predictor for activity, individual categories also scored relatively low. Category 6, which covered winds of 245°-270°, or South-West-West winds, was shown to be the strongest predictor. Unexpectedly, surrounding categories 4, 5, and 7 (winds ranging from 135°-225° and from 270°-315°) scored very poorly within the model. After category 6, the categories 1, 3, and 2 scored highest, respectively. Collectively, these categories covered winds of 0°-135°, which are winds coming from North to South-eastern direction. In relation to Studio Bruinvis, categories 1, 2 and 3 are all winds coming from the direction of the Schouwen-Duiveland island, while winds from categories 4, 5, and 7 are from the direction of the water. In general, it could be hypothesized that in this case winds coming from landmasses act as stronger predictors for high activity compared to winds coming from the direction of water bodies. However, this would not explain the strongest predictor (category 6) also coming from the direction of a water body. As stated previously, it is believed that wind direction only indirectly affects porpoise activity as it may influence water currents, which in turn might influence harbour porpoises directly or indirectly (Oijen, 2009). Therefore, the varying results for different wind directions may actually be attributed to an unknown factor.

Time of day was the third strongest predictor for activity, of which categories 4 and 3, which collectively covered 8 hours from 8:00-16:00hour were highest. The reason for this was thought to be related to variables not included in this study, because no strong relation between hours with any variables was found. A study done by IJsseldijk *et al.* (2015) in the Netherlands found time of day to be one of the best describing variables for harbour porpoise distribution in the Marsdiep. Their highest encounter rate was, however, found in early mornings and decreased during the day, which does not entirely align with the results of this study. IJsseldijk *et al.* also discussed that this time of day effect might actually be due to collinearity between time of day and tidal phase. Regardless, time of day might actually play a role in prey availability and distribution, which often follows some sort of diel behaviour (Michalsen *et al.,* 1996; Petrakis *et al.,* 2001).

Wind speed was found to be the next highest predicting factor in the model and was found to have a positive relation with the chance to predict high activity levels. Wind speed might indirectly affect porpoise activity by affecting water currents both vertically and horizontally (Oijen, 2009), which in turn could affect porpoises or their prey species (Benjamins *et al.,* 2016).

Water height and water temperature showed to be predictors in a lesser degree, though still significant. Water height was found to be positively related to high activity, meaning that when water height increased, so would the chance of predicting high activity. This relation is supported by IJsseldijk (2013) who found that harbour porpoises had a preference for high tidal water.
Water temperature also had a positive effect, meaning that if temperature increased, so would the chance of predicting high porpoise activity. This complies with results from IJsseldijk (2013) studying porpoise presence in the Marsdiep, The Netherlands. In IJsseldijk's study, water temperature was thought to be related to water flux from tidal phases: incoming waters bring colder sea water in, while outgoing waters are generally warmer. Because of this, it was suggested that harbour porpoise activity was mainly affected by tidal phase and that temperature might only play a role in smaller time scales (IJsseldijk, 2013). However, no sign of collinearity between water temperature and tidal phase was detected in this study, implying that water temperature could have some other effect on harbour porpoises. This for instance, could be due to the effects of water temperature on currents for example (Oijen, 2009).

As mentioned before, it is important to consider that the activity level chosen in this research, based on ≥400 or <400 clicks, only suggests that there is a higher chance of observing porpoises. The level of clicks emitted by porpoises only suggests the level of activity, which in turn, could affect the frequency of harbour porpoises surfacing, allowing them to be observed or not (Nuuttila *et al*., 2013; Pierpoint, 2008). However, porpoises are notoriously elusive and difficult to spot on the surface because of their relatively small body size and inconspicuous surfacing behaviour (Camphuysen, 2004). Therefore, under certain circumstances there might be a chance that harbour porpoises are very active at Studio Bruinvis, but the sea state might still obscure them from being visible. A future study could examine the predicted chance of activity versus the actual observability of harbour porpoises at Studio Bruinvis.

Lastly, it should also be reminded that the Studio Bruinvis buoy can only detect clicks and cannot distinguish between individuals or estimate the number of individuals. It could be the case that several harbour porpoises are present at the hotspot and only produce a low number of clicks or, in contrast, a single individual could be present producing a high number of clicks.

4.2 Mobile application

The design of the app aimed to be simple and not overloaded with pictures and information. The vertical scroll arrangement favored the portrait orientation of smartphones. However, it adjusted automatically to the landscape orientation as well. Further information via external links was provided.

For now, the app is only available on Android devices, due to the limitations of appinventor2 (Bergner & Leonhardt, 2017). However, the written code is not only available in the appinventor's block format but also in written text which gives an experienced developer the option to partly convert the code to an iOS (apple) or windows compatible product. Unfortunately, apps have to be rebuilt from scratch in order to be compatible with different platforms. Although slight design changes might have to be taken into account due to design differences such as the positioning of buttons and design restrictions (Bondarenko, 2017).

In order to raise awareness for the app, a marketing strategy is needed. As a first instance, the channels of the Rugvin Foundation (website, Facebook, twitter, Instagram,) should be used. This way, parties already interested in the Rugvin Foundation, will become aware of the app's existence. Furthermore, other networks, and collaborations can follow (e.g. national park de Oosterschelde) (Stichting Rugvin, 2018). So far, the app is available for free, without additional payments or advertisements since sponsorships are not part of the goal. Part of the marketing is the compatibility and accessibility of the app. However, before an app can be downloaded from the play store or app store (after conversion to iOS), the guidelines of the respective shops have to be fulfilled. In order to circumnavigate this problem in the first run, the app can be distributed by a third-party website (like the Rugvin Foundation). However, smartphones users need to allow downloads from unknown sources on their device (Bergner & Leonhardt, 2017).
It should be taken into account that, since the app is based on the regression model, predictions will only prove correct about 74% of all cases. Due to model predictions having an overall correctly predicted value of 74% and validation of the model proving to correctly predict about 75%, the app will not predict with complete certainty.

4.3 Study restrictions

Despite the outcomes, some complications occurred during the study, which were unavoidable. Firstly, the equipment of Studio Bruinvis that is used to monitor porpoise activity at the designated hotspot can only be adjusted by the field expert C. Menhennet. The expert lives abroad, meaning the recording quality of audio could not be adjusted on site when needed. Unfortunately, it turned out that the audio quality of around 30% was inadequate due to distortion caused by winds or passing boats. This lead to a smaller sample size as well as gaps in the data. These gaps were mostly linked to strong winds which resulted in lesser samples with strong wind speeds (approximately >9m/s) compared to samples with weak wind speeds (approximately 0-9 m/s). Meaning that, environmental conditions with strong wind speeds are not as representative within the dataset as much as environmental conditions with weaker wind speeds. Adjusting or improving the monitoring equipment could solve this limitation in future research.

A second limitation of this study is the short study period. The study period started at August 15th 2017, as this was the day Studio Bruinvis started monitoring audio continuously after running some tests. The study period ended on October 5th 2017 as this was the day Studio Bruinvis stopped recording due to vandalism and/or bad weather conditions, which led to damaged equipment. Repairs could not be done on short notice and resulted in a loss of 10 days' worth of data.

Furthermore, due to the study period not continuing for an entire year, some used variables such as water temperature, wind speed and wind direction, could not be measured for their complete potential ranges within a year. This prevented seasonal effects to be explored within the model. As a result, the developed predictability model does not accurately portray conditions for the entire year, and thus cannot be extrapolated to a different timeframe. In order to predict the activity throughout the year, the data collection should be continued after this study.

Another limitation of this study was that the model only includes variables that are publicly available. For this reason, it was chosen to use only abiotic factors as variables that could easily be obtained from online sources and be linked to the mobile application. Most studies concerning the distribution of marine mammals also focus on biotic factors such as prey abundance and distribution. For example, a study of Sveegaard *et al*. (2012) found that local prey abundance is the main driver for harbour porpoise distribution/presence. This study further suggested that the local prey abundance can be linked with season, suggesting that season can be seen as a proxy for prey abundance. This may imply that all used abiotic variables in this study can also be seen as indirect factors or proxy for harbour porpoise activity.

At this point, Studio Bruinvis consist of one buoy with monitoring equipment. Therefore, the study area was bound to the location of the buoy from Studio Bruinvis and its detection range. This can also be considered a limitation. As described earlier in paragraph *2.1 Study area*, Studio Bruinvis is located at a porpoise watching hotspot near the harbour of Zierikzee. This location is also known for its old underwater ammunition depot and it is thought that this deep (approximately 53 meters, see appendix I) underwater pit attracts prey fish species (Zanderink, pers. comm., 2017). This implies that the study area is unique and might not be representative for other porpoise watching hotspots.

Despite the predictability model for the first monitoring location of Studio Bruinvis, predictors might not be reliable for other porpoise hotspots with or without different environmental conditions. More monitoring locations of Studio Bruinvis could help in getting a better understanding in why these so-called hotspots attract harbour porpoises.

5. CONCLUSIONS

The results of this research showed that all variables (i.e. time of day, tide, water temperature, water height, wind speed and wind direction) had a significant effect on predicting the number of clicks ≥400, and thus porpoise activity. Tide was found to be the most important predictor followed by wind direction, time of day, wind speed, water height and lastly water temperature. Chance of high activity seemed to increase with increased wind speed, increased water temperature and increased water height. Furthermore, outgoing to low tide and incoming tides as well as winds from 225°-270° (west-south-west direction), and time of day between 8:00-16:00hours showed to best predict porpoise click activity at Studio Bruinvis. Wind speed, water temperature and water height showed to be significant predictors to a lesser degree.

The equation generated by the regression model was used as the basis for an app to calculate the likelihood of harbour porpoise activity. The model proved to be able to predict activity levels of ≥400 clicks per hour with around 75% certainty. A desk study provided the means to develop the app to be user-friendly and using the latest measured data for the variables included in the model equation created a way to predict activity at short notice. Combining the model equation with the knowledge from a desk study about mobile applications development, the Studio Bruinvis app was created: an informative and useful application that informs users whether the likelihood of porpoise activity, and therefore chance to observe porpoises, is high or low.

The results and findings of this study can be seen as a first try in getting a better understanding about factors that influences porpoise activity at a specific location, such as Studio Bruinvis.

6. RECOMMENDATIONS

Recommendations for future research:

The baseline of 400 clicks resulted from the mean clicks of the data set. This baseline should still be validated by comparing if harbour porpoise activity at such a level leads to more frequent of clear observability. The study period should be extended to cover at least all seasons throughout the year in order to gather data on more suitable days and provide a bigger data set. Furthermore, the study should be prolonged to a multi-year study in order to measure the variability between years and season and to identify possible trends. The Studio Bruinvis hydrophone settings should be adjusted to avoid distortions in the recordings and provide more useable data.

To possibly improve the model, it is advised to include additional variables that are important for harbour porpoises' activity; such as water currents and season.

Measurement points that are situated closer to Studio Bruinvis could provide a more accurate picture of the conditions in the study area, resulting in more accurate predictions.

Recommendations for the app include:

In order to facilitate the input for the users of the variables it is desirable that the values can be retrieved automatically from the respective websites.

Adding an interactive map that displays the last sightings of harbour porpoises in the area and give citizen scientists the opportunity to log their own sightings. This way, locals and visitors can contribute to the data set and are encouraged to learn more about harbour porpoises.

In order to improve the app, it has to become known first. This could be done by the channels of Stichting Rugvin (the website, twitter, facebook, Instagram) to attract people that are already interested in such a topic and want to give their feedback. Lastly, it would be beneficial to make the app compatible with iOS and windows operating systems in order to reach and encourage more people to learn about harbour porpoises and visit Studio Bruinvis.

REFERENCES

Audacity (2018). *Audacity*. Available from: https://www.audacityteam.org/

Bearzi, G., Fortuna, C., Reeves, R.R. (2012). *Tursiops truncatus (Mediterranean subpopulation).* The IUCN Red List of Threatened Species 2012: e.T16369383A16369386.

Benjamins, S., Dale, A., van Geel, N., Wilson, B. (2016). *Riding the tide: use of a moving tidal-stream habitat by harbour porpoises*. Marine Ecology Progress Series 549:275-288. https://doi.org/10.3354/meps11677

Bergner, N., Leonhardt, T. (2017). *Apps maken voor kids voor dummies*. BBNC-uitgevers, Amersfoort, NL.

Bondarenko, A. (2017). *How to convert Android app to IOS app (and vice versa)*. Stormotion. Found 11-05- 2018 on https://stormotion.io/blog/how-to-convert-android-app-to-ios-app-and-vice-versa/

Brookes, K.L., Bailey, H., Thompson, P.M. (2013). *Predictions from harbour porpoise habitat association models are confirmed by long-term passive acoustic monitoring.* Journal of Acoustical Society America 134 (3), Pt. 2, September 2013 0001-4966/2013/134(3)/2523/11

Bryant, M.J., Smart, W.J., Wilde, S.J. (2016). *Swipe right with a chance of rain: weather app usage on smartphones.* Advances in Buisness-Related Scientific Research Journal, 7 (2).

Camphuysen, K.C.J. (2004). *The return of the harbour porpoise (Phocoena phocoena) in Dutch coastal waters.* Lutra 2004 47 (2): 113-122

Camphuysen, K.C.J., Siemensma, M.M.L. (2011). *Conservation plan for the Harbour Porpoise (Phocoena phocoena) in The Netherlands: towards a favourable conservation status.* Report 2011-07, Royal Netherlands Institute for Sea Research, Texel.

Cañadas, A., Hammond, P. S. (2006). *Model-based abundance estimates for bottlenose dolphins off southern Spain: implications for conservation and management.* Journal of Cetacean Research Management 8(1):13–27, 2006

Cañadas, A., Sagarminaga, R., García-Tiscar, S. (2002). *Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain.* Deep Sea Research part I. 49: 2053–2073

Cañadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E. and Hammond, P.S. (2005). *Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters.* Aquatic Conservation: Marine and Freshwater Ecosystems., 15: 495–521. doi:10.1002/aqc.689

Chan, Y.H. (2004). Biostatistics 202: *Logistic regression analysis*. Singapore Med J 2004 Vol. 45(4):149.

Claeskens, G., Hjort, N.L. (2008). *Model Selection and Model Averaging.* Cambridge University Press.

Kenniscentrum Kusttoerisme. (2017). *Kerncijfers vrijetijdseconomie Zeeland 2016.*

Cyprus, D.P., Blaber, S.J.M. (1992). *Turbidity and salinity in a tropical northern Australian estuary and their influence on fish distribution.* Estuarine, Coastal and Shelf Science. Volume 35, Issue 6, December 1992, P. 545-563

Dekker, D.H.J. (2016). *De verstoringsafstanden van rustende zeehonden op de Roggenplaat in de Oosterschelde - De reacties van rustende zeehonden op een menselijke benadering in het voorjaar van 2016.* Rijkswaterstaat Zee en Delta

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., Lautenbach, S. (2013). *Collinearity: a review of methods to deal with it and a simulation study evaluating their performance.* Ecography, 36: 27-46. doi:10.1111/j.1600-0587.2012.07348.

Eck, dr. G.Th.M., Holland, A.M.B.M., Pagee, ir. J.A. (2001). *Risicobeoordeling Munitiestort Oosterschelde.* Rijksinstituut voor Kust en Zee/RIKZ

Embling, C.B., Gillibrand, P.A., Gordon, J., Shrimpton, J., Stevick, P.T., Hammond, P.S. (2010). *Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (Phocoena phocoena).* Biological Conservation 143:267-279.

Ettinger, H.D. van, Zeeuw, R.C. (2010). *Prognose van de zwemveiligheid op en rond de Zandmotor.* Projectbureau Pilot Zandmotor

European Commission. (2017) *Natura 2000.* Found on 24-05-2018, on: http://ec.europa.eu/environment/nature/natura2000/index_en.htm

Giannoulaki, M., Markoglou, E., Valavanis, V.D., Alexiadou, P., Cucknell, A., Frantzis, A. (2016). *Linking small pelagic fish and cetacean distribution to model suitable habitat for coastal dolphin species, Delphinus delphis and Tursiops truncatus, in the Greek Seas (Eastern Mediterranean).* Aquatic Conservation: Marine And Freshwater Ecosystems 2016. DOI: 10.1002/aqc.2669

Gazdecki, A. (2018). *How to build a mobile app in 12 easy steps*. Biznessapps. Found on 24-05-2018, on: https://www.biznessapps.com/blog/how-to-build-a-mobile-app-in-12-easy-steps/

Gilles, A., Adler, S., Kaschner, K., Scheidat, M., Siebert, U. (2011). *Modelling harbour porpoise seasonal density as a function of the German Bight environment: implications for management.* Endangered Species Research Vol. 14: 157–169, 2011 doi: 10.3354/esr00344

Gilles, A., Viquerat, S., Becker, E.A., Forney, K.A., Geelhoed, S.C.V., Haelters, J., Nabe-Nielsen, J., Scheidat, M., Siebert, U., Sveegaard, S., van Beest, F.M., van Bemmelen, R., Aarts, G. (2016). *Seasonal habitat-based density models for a marine top predator, the harbor porpoise, in a dynamic environment.* Ecosphere 7(6):e01367. 10.1002/ecs2.1367

Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.P., Heimlich, S., Hiby, A.R., Leopold, M.F., Øien, N. (2002). *Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters.* Journal of Applied Ecology, 39: 361–376. doi:10.1046/j.1365-2664.2002.00713.x

Hammond, P.S., Bearzi, G., Bjørge, A., Forney, K., Karczmarski, L., Kasuya, T., Perrin, W.F., Scott, M.D., Wang, J.Y., Wells, R.S. & Wilson, B. (2008). *Phocoena phocoena.* The IUCN Red List of Threatened Species 2008: e.T17027A6734992.

Hammond P.S., Lacey, C, Gilles, A., Viquerat, S., Börjesson, P., Herr, H., Macleod, K., Ridoux, V., Santos, M.B., Scheidat, M., Teilmann, J., Vingada, J., Øien, N. (2017). *Estimates of cetacean abundance in European Atlantic waters in summer 2016 from the SCANS-Ⅲ aerial and shipboard surveys.* SCANS-Ⅲ, 1-39

Hann, C.H., Stelle, L.L, Szabo, A., Torres, L.G. (2018). *Obstacles and opportunities of using a mobile app for marine mammal research*. ISPRS International Journal of Geo-Information, 7(5), 169; doi:10.3390/ijgi7050169

IJsseldijk, L. (2013). *Tidal influences on the distribution of the Harbour Porpoise Phocoena in the Marsdiep area, The Netherlands.* Imares.

IJsseldijk, L., Camphuysen, K.C.J., Nauw, J.J., Aarts, G. (2015). *Going with the flow: Tidal influence on the occurrence of the harbour porpoise (phocoena phocoena) in the Marsdiep area, The Netherlands.* Journal of Sea Research. 103, 129-137

Isojunno, S., Matthiopoulos, J., Evans, P.G.H. (2012). *Harbour porpoise habitat preferences: robust spatio-temporal inferences from opportunistic data.* Marine Ecology Progress Series Vol. 448 (February 23 2012), pp. 155-170

Jansen, O.E., Aarts, G.M., Reijnders, P.J.H. (2013). *Harbour Porpoises Phocoena phocoena in the Eastern Scheldt: A Resident Stock or Trapped by a Storm Surge Barrier?* PLoS ONE 8(3): e56932. https://doi.org/10.1371/journal.pone.0056932

Johnston, D. W., Westgate, A. J., Read, A. J. (2005). *Effects of fine-scale oceanographic features on the distribution and movements of harbour porpoises Phocoena phocoena in the Bay of Fundy*. Marine Ecology Progress Series 295:279-293.

Keijl, G.O., Begeman, L., Hiemstra, S., IJsseldijk, L.L., Kamminga, P. (2016). *Cetaceans stranded in the Netherlands in 2008-2014.* Lutra 59 (1-2): 75-107.

Klopstra, A. (2013). *Groene apps voor meer verbinding met natuur.* InnovatieNetwerk.

Korpelshoek, L.D. (2011). *Resident harbour porpoises Phocoena phocoena in the Oosterschelde (Netherlands): their behaviour compared to the behaviour of migratory harbour porpoises in the southern North Sea.* Institute of Environmental Sciences (CML), Leiden University.

Küchler, M. (2016). *Kleines Lexikon der App-Entwicklung*. FLYACTS. Die App Agentur. Found on 24-05-2018, on: http://www.flyacts.com/blog/kleines-lexikon-der-app-entwicklung/

Lewis-Becl, M.S. (1995). *Data analysis: an introduction.* Sage University Paper series on Quantative Applications in the Social Sciences, 07-103.

Lockyer, C. (2003). *Harbour porpoises (Phocoena phocoena) in the North Atlantic: Biological parameters.* NAMMCO Sci. Publ. 5:71-90.

Luchtmeijer, D. (2017). *10 tips voor een succesvolle mobiele app.* Available from https://www.digibilities.nl/10-tips-voor-een-succesvolle-mobiele-app

Macleod, C.D., Weir, C.R., Pierpoint, C., Harland, E.L. (2007). *The Habitat Preference Of Marine Mammals West Of Scotland (UK).* Journal of the Marine Biological Association of the United Kingdom (2007) 87, 157–164

Markoglou, E., Frantzis, A., Valavanis, V.D., Alexiadou, P., Kalaitzidis, C., Cucknell, A.C., Giannoulaki, M. (2015). *Habitat Suitability Maps of Bottlenose Dolphin In The Greek Seas.* 11th Panhellenic Symposium on Oceanography and Fisheries, Mytilene, Lesvos island, Greece 2015

Marubini, F., Gimona, A., Evans, P.G.H., Wright, P.J., Pierce, G.J. (2009). *Habitat preferences and interannual variability in occurrence of the harbour porpoise Phocoena phocoena off northwest Scotland.* Marine Ecology Progress Series Vol. 381 (April 17 2009), pp. 297-310

Meteoblue (2017). *Klimaat Zierikzee.* Found on 24-05-2018, on: https://www.meteoblue.com/nl/weer/voorspelling/modelclimate/zierikzee_nederland_2743913

Michalsen, K., Godø, O.R., Fernö, A. (1996). *Diel variation in the catchability of gadoids and its influence on the reliability of abundance indices.* ICES Journal of Marine Science, Volume 53, Issue 2, 1 April 1996, Pages 389–395, https://doi.org/10.1006/jmsc.1996.0054

Mikkelsen, L., Rigét, F.F., Kyhn, L.A., Sveegaard. S., Dietz, R., Tougaard, J. (2016). *Comparing Distribution of Harbour Porpoises (Phocoena phocoena) Derived from Satellite Telemetry and Passive Acoustic Monitoring.* PLoS ONE 11(7): e0158788. doi:10.1371/journal.pone.0158788

Miller, G.T., Spoolman, S.E. (2009). *Essentials of Ecology.* 5e. © 2009, 2007 Brooks/Cole, Cengage Learning, p.108

Ministerie van Economische Zaken (2017). *Habitatrichtlijnen.* Found on 24-05-2018, on: http://minez.nederlandsesoorten.nl/content/habitatrichtlijn

Nationaal Park Oosterschelde (2017). *Natuur.* Found on 24-05-2018, on: http://www.np-oosterschelde.nl/over-het-park/natuur.htm

Nuuttila, K.H., Meier, R., Evans, P.G.H., Turner, J.R., Bennell, J.D., Hiddink, J.G. (2013). *Identifying Foraging Behaviour of Wild Bottlenose Dolphins (Tursiops truncatus) and Harbour Porpoises (Phocoena phocoena) with Static Acoustic Dataloggers.* Aquatic Mammals 2013, 39(2), 147-161, DOI 10.1578/AM.39.2.2013.147

Oijen, T. van (2009). *Dynamiek van water en sediment in de Ostfriese Waddenzee.* Found on 11-05-2018, on https://www.waddenvereniging.nl/wadweten/1150-dynamiek-van-water-en-sediment-in-de-ostfriese-waddenzee.html

Pamguard (2017*). Pamguard, open source software for passive acoustic monitoring.* Found on 24-05-2018, on: https://www.pamguard.org/12_MarineMammalAcoustics.html

Perrin, W.; Würsig, B.; Thewissen, J. (2008). *Encyclopedia of Marine Mammals.* Second Edition. Elsevier. USA.

Petrakis, G., MacLennan, D.N., Newton, A.W. (2001). *Day–night and depth effects on catch rates during trawl surveys in the North Sea.* ICES Journal of Marine Science, Volume 58, Issue 1, 1 January 2001, Pages 50–60, https://doi.org/10.1006/jmsc.2000.0989

Pierpoint, C. (2008). *Harbour porpoise (Phocoena phocoena) foraging strategy at a high energy, near-shore site in south-west Wales, UK*. Journal of the Marine Biological Association of the United Kingdom, 88(6), 1167-1173. doi:10.1017/S0025315408000507

Redfern, J. V., Ferguson, M. C. , Becker, E. A., Hyrenbach, K. D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M. F., Forney, K. A., Ballance, L. T., Fauchald,P., Halpin, P., Hamazaki, T., Pershing, A. J., Qian, S. S., Read, A., Reilly, S. B., Torres L., Werner, F. (2006). *Techniques for cetacean–habitat modeling.* Marine Ecology Progress Series Vol. 310: 271–295, 2006

Rijkswaterstaat (a) (2017). *Getij.* Found on 24-05-2018, on: https://www.rijkswaterstaat.nl/water/waterdata-en-waterberichtgeving/waterdata/getij

Rijkswaterstaat (b) (2017). *Monitoring van de temperatuur van rivierwater*. Found on 24-05-2018, on: https://www.rijkswaterstaat.nl/water/waterbeheer/waterkwaliteit/indicatoren-voor-waterkwaliteit/aqualarm/meetsystemen/chemisch-fysisch/temperatuur.aspx

Rijkswaterstaat (2018). *Rijkswaterstaat*. Found on 10-05-2018, on: https://www.rijkswaterstaat.nl/english/water-systems/protection-against-water/storm-surge-barriers.aspx

Rodrigues, J.M.G. (2014). *Echolocation activity harbour porpoise Phocoena phocoena in the Eastern Scheldt estuary, (The Netherlands) and the North Sea.* The Rugvin Foundation. 1-38.

Scheijgrond, P.C., Schaap, A., Sjerps-Koomen, E.A. (2000). *Kansen voor energiewinning uit getijden in de Oosterschelde.* ECOFYS Energie winning uit getijden in de Oosterschelde / E 45074.1

Shirihai, H., Jarrett, B. (2008). *Meeressäuger alle 129 Arten weltweit*. Franckh-Kosmos Verlags-GmbH & Co. KG, Stuttgart

Snyder, S. (2015). *Validation of stepwise regression using full model and 75 percent training sample approach*. Found on 25-05-2018, on: https://www.youtube.com/watch?v=mAMqxpKbOt8

Stichting Rugvin (2017). *Studio Bruinvis.* Found on 24-05-2018, on: http://rugvin.nl/studio-bruinvis/

Stichting Rugvin (2018). *Rugvin*. Found on 24-05-2018, on: http://rugvin.nl/

Sveegaard, S. (2011). *Spatial and temporal distribution of harbour porpoises in relation to their prey*. PhD thesis. Department of Arctic Environment, NERI. National Environmental Research Institute, Aarhus University, Denmark. 128 pp.

Sveegaard, S., Andreasen, H., Mouritsen, K.N., Jeppesen J.P., Teilmann, J., Kinze, C.C. (2012). *Correlation between the seasonal distribution of harbour porpoises and their prey in the Sound, Baltic Sea.* Marine Biology 159: 1029.

Tibbs, C. (2005). *Weather Handbook. Northern Hemisphere Edition*. The Royal Yacht Association (RYA). Southampton, UK.

Todd, V.L.G., Pearse, W.D., Tregenza, N.C., Lepper, P.A., Todd, I.B. (2009). *Diel echolocation activity of harbour porpoises (Phocoena phocoena) around North Sea offshore gas installations.* ICES Journal of Marine Science, 66(4), 734-745.

Todd, V.; Todd, I., Gardiner, J.; Morrin, E. (2015). *Marine Mammal Observer & Passive Acoustic Monitoring Handbook*. Pelagic Publishing. Exeter, UK.

Toekomst Schouwen-Duiveland. (2017). *De sterkste natuurbeleving.* Found on 24-05-2018, on: https://toekomstschouwen-duiveland.nl/Stellingen_toekomst_toerisme/ArticleID/216/Stelling-3

UCLA (2017). *Logistic Regression - Spss Annotated Output*. UCLA: Statistical Consulting Group. Found on 10-05-2018, on https://stats.idre.ucla.edu/spss/output/logistic-regression/

Unal, I. (2017). *Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 3762651, 14 pages, 2017. https://doi.org/10.1155/2017/3762651.

Van Dam, S. Solé, L., Ijsseldijk, L.L. , Begeman, L., Leopold, M.F. (2017). *The semi-enclosed tidal bay Eastern Scheldt in the Netherlands: porpoise heaven or porpoise prison?* Lutra 60 (1): 5-18

Van der Meij, S.E.T., Camphuysen, K.C.J. (2006). *The distribution and diversity of whales and dolphins (Cetacea) in the southern North Sea: 1970-2005.* Lutra 2006 49 (1): 3-28

Waggitt, J.J., Dunn, H.K., Evans, P.G.H., Hiddink, J.G., Holmes, L.J., Keen, E., Murcott, B.D., Piano, M., Robins, P.E., Scott, B.E., Whitmore, J., Veneruso, G. (2017). *Regional-scale patterns in harbour porpoise occupancy of tidal stream environments.* ICES Journal of Marine Science. 1-10.

Williamson, L.D., Brookes, K.L., Scott, B.E., Graham, I.M., Thompson, P.M. (2017). *Diurnal variation in harbour porpoise detection – potential implications for management.* Marine Ecology Progress Series, 223-232

Zuur, A., Ieno, E.N., Elphick, C.S. (2010). *A protocol for data exploration to avoid common statistical problems*. Methods in Ecology and Evolution 2010, 1, 3–14

GLOSSARY

Abiotic variables: non-living chemical and physical factors such as salinity, light and temperature (Perrin et.al, 2008)

Absence: Harbour porpoises are considered absent when no clicks were recorded

Activity: presence of harbour porpoises based on click data

Baseline: Threshold between high and low click activity. For this project the baseline of 400 clicks/hour were chosen as a baseline.

Beaufort Sea State (BSS): A wind force scale from 0 (calm) to 12 (hurricane) (Tibbs, 2005)

Biotic variables: living organisms which interact with other organisms (Perrin et al. 2008)

Cetacean: Marine mammals of the order Cetacea which include whales, dolphins and porpoises (Perrin et al, 2008)

Click activity describes the hourly click rate of harbour porpoises around Studio Bruinvis. high: >400 clicks/hour low: <400 clicks/hour The more clicks are emitted the more likely it is to encounter harbour porpoise(s) in favourable (calm) weather conditions

Click: a sound of small duration between 110 and 150 kHz produced by odontocete (toothed) cetaceans (Perrin et al. 2008)

Collinearity Different variables influence each other which could make it difficult to assign independent variable effects (Dormann *et al.* 2013).

Distortions Distorting noises in the audio files, mostly caused by marine traffic or strong winds, which obscure clicks and can cause overestimation of total amount of clicks (false clicks).

Echolocation: Emitting acoustic signals in order to obtain a sense of an organism's surroundings from the echoes it receives. Toothed whales use this for orientation and foraging (Todd et al., 2015)

False clicks: Noise that the software PAMGuard mistakenly counted as a porpoise click

Frequency: The number of periods per second. The unit is the hertz (Hz) where 1Hz equals one cycle per second (Todd et al., 2015)

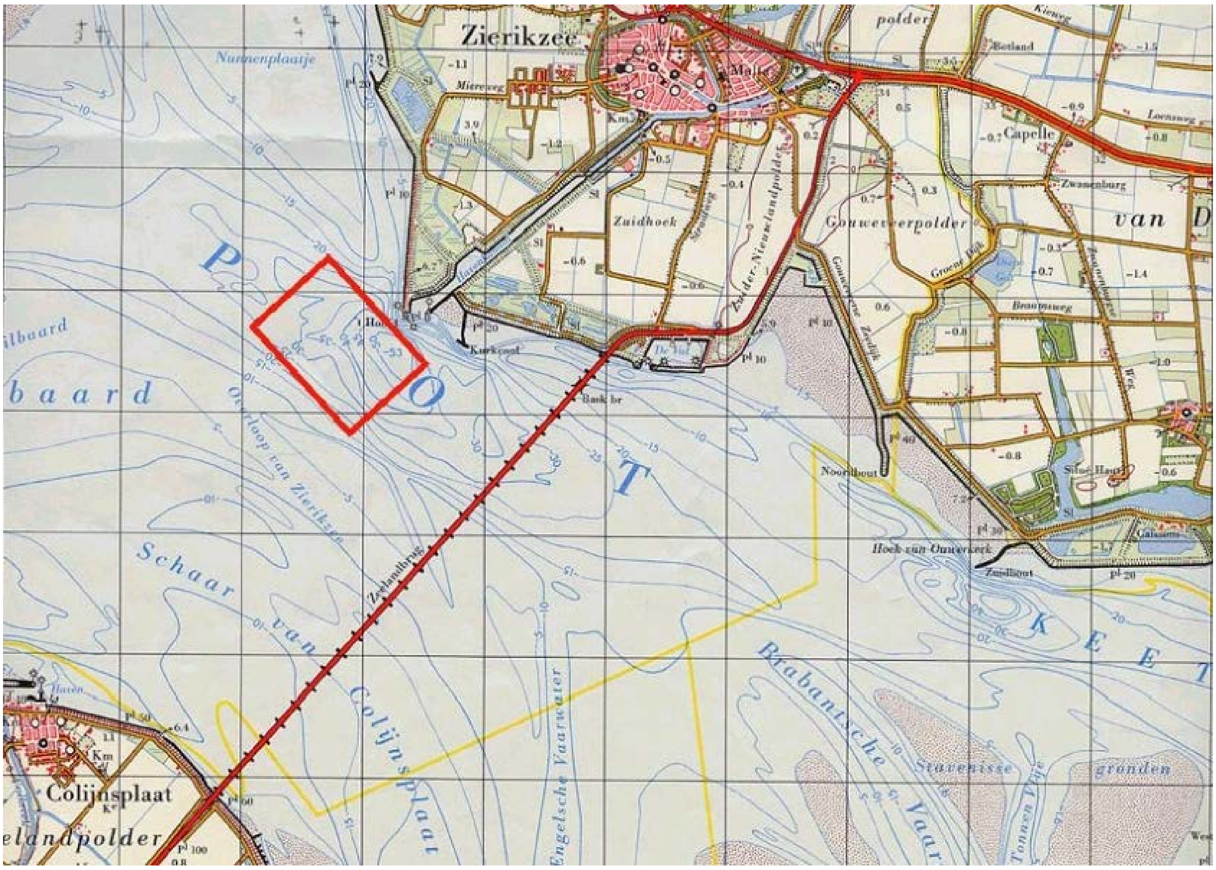
Multi collinearity see collinearity

Presence: At least one harbour porpoises was considered present when at least one click was recorded

Storm surge barrier: Moveable flood barriers at the mouths of rivers, tidal inlets and estuaries which close when water levels are extremely high. The barrier in the Eastern Scheldt (Oosterschelde) is with 9km the largest structure of the Delta works. (Rijkswaterstaat, 2018)

Tides: Vertical movements of water. When the water is rising, the tide is coming in (flooding). When the water level decreases, the tide is going our (ebbing)

Appendix I: Hydrographical map study area



Appendix II: Apps used as inspiration for app development

Table A: Apps used as an inspiration for the Stichting Rugvin app

Name of app	Developer	Version	Released	Updated on
Het Getij – Waterstanden	Surfcheck	3.3	16th June, 2013	22nd February, 2018
Surfcheck	Surfcheck	5.2	20th July, 2010	29th June, 2017
WCA Responsible	World Cetacean Alliance	1.0.1	22nd June, 2017	22nd June, 2017
Center for Whale Research	Webappclouds.com	1.11	4th September, 2017	12th September, 2017
Whale Alert	Conserve.io	1.11	6th June, 2015	1st March, 2018
Wild About Whales	Office of Environment and Heritage	2.0.4	19th August, 2011	2nd August, 2017

Appendix III: Waterproof BV Pamguard settings

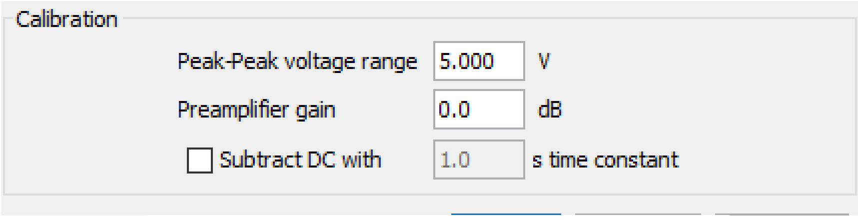


Figure A: Sound Acquisition Setting by Waterproof BV

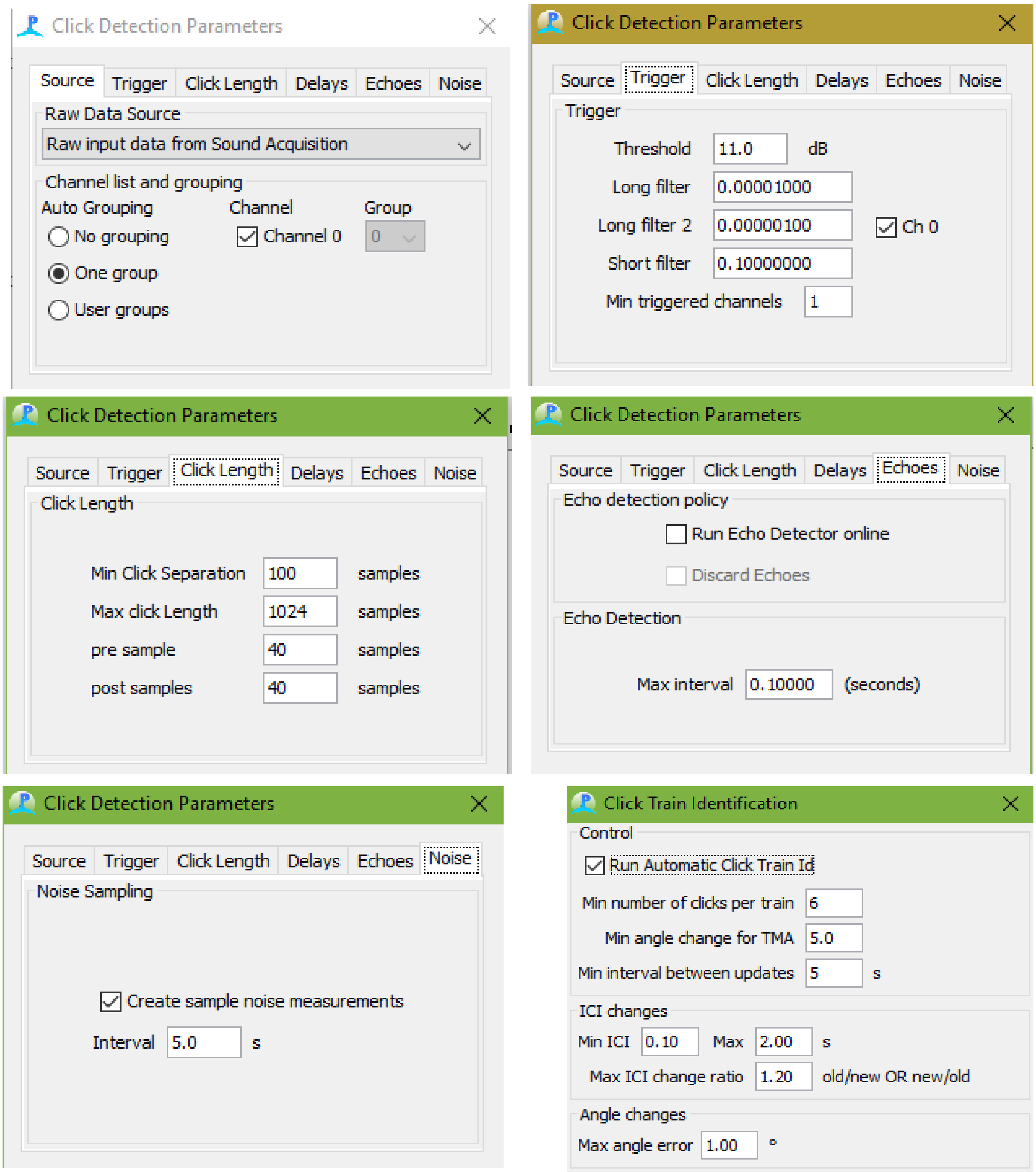


Figure B: Click train detection settings by Waterproof BV

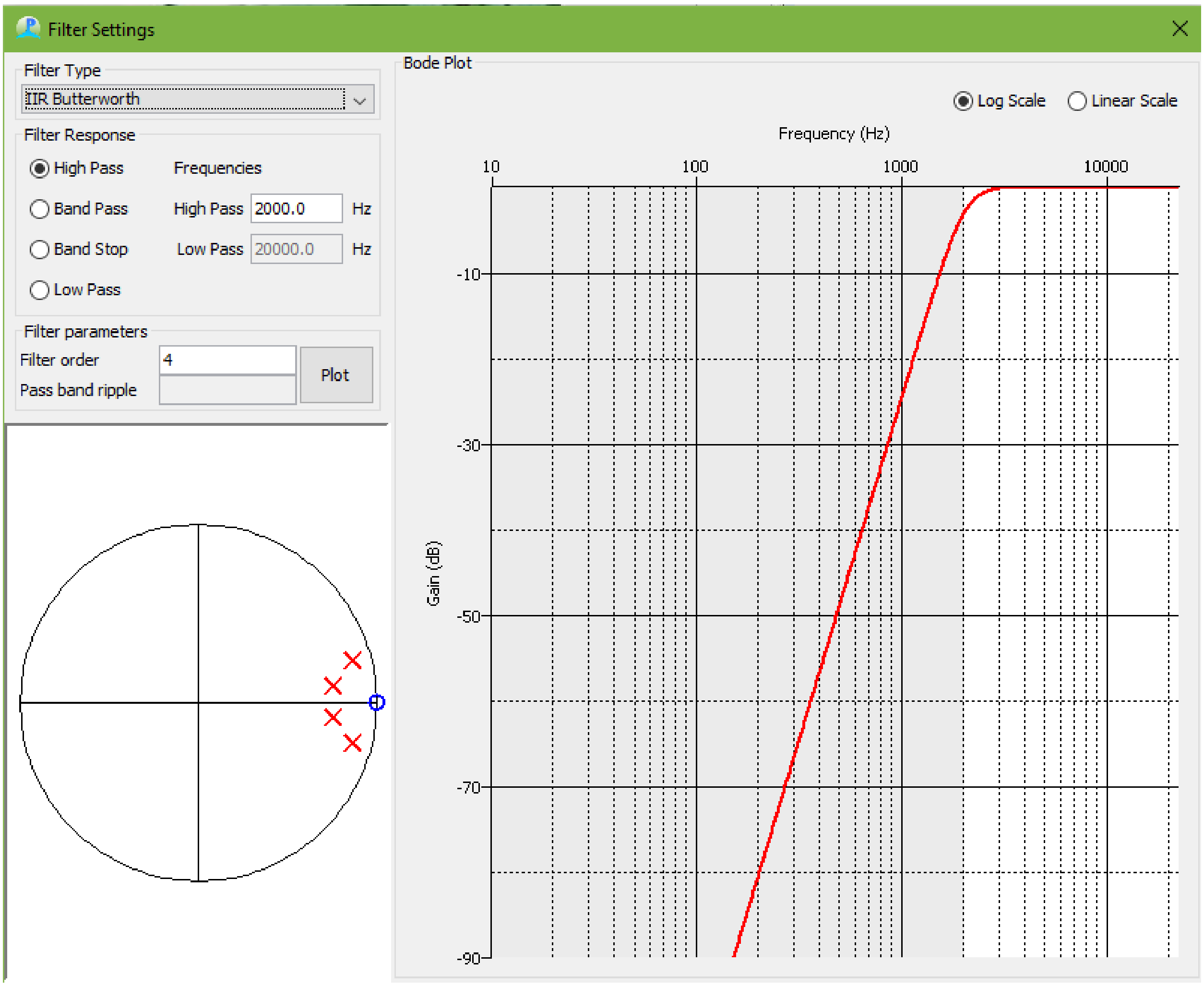


Figure C: Filter Settings by Waterproof BV

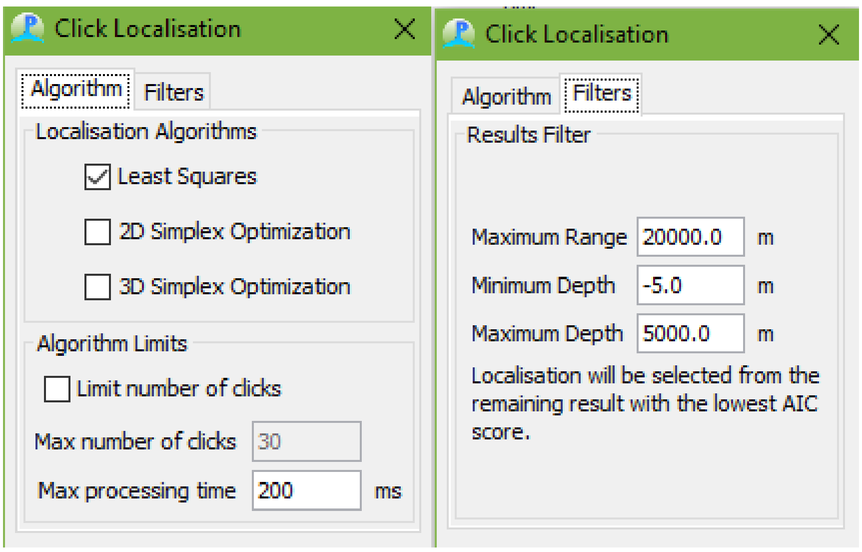


Figure D: Click Localisation setting by Waterproof BV

Appendix IV: Excel data sheet 1, 2 and 3.

	DATUM	TIJD	WATERHOOGTE	TIDES	WT Marollegat	WT Oude Tonge	WT O.K	GEM WT OS	Graden t.o.v Noorden	Windsnelhei m/s
92	15-aug-17	15:00	-117	1	20,6	20	19,9	20,2	267	2,016666667
93	15-aug-17	15:10	-114	1	20,7	20	19,9	20,2	263	2,02
94	15-aug-17	15:20	-109	1	20,7	19,9	19,9	20,2	266	2,016666667
95	15-aug-17	15:30	-103	1	20,8	20	20	20,3	266	2,026666667
96	15-aug-17	15:40	-98	1	20,8	20	20	20,3	267	2,026666667
97	15-aug-17	15:50	-92	1	20,8	20	20	20,3	269	2,026666667
98	15-aug-17	16:00	-86	1	20,8	20	20	20,3	270	2,026666667
99	15-aug-17	16:10	-80	1	20,8	20	20	20,3	267	2,026666667
100	15-aug-17	16:20	-73	1	20,8	19,9	20	20,2	263	2,023333333
101	15-aug-17	16:30	-66	1	20,8	20,1	20	20,3	276	2,03
102	15-aug-17	16:40	-59	1	20,8	20,1	20	20,3	270	2,03
103	15-aug-17	16:50	-52	1	20,8	20,1	20	20,3	279	2,03
104	15-aug-17	17:00	-44	1	20,7	20,1	20	20,3	287	2,026666667
105	15-aug-17	17:10	-36	1	20,6	20,1	20	20,2	289	2,023333333
106	15-aug-17	17:20	-28	1	20,6	20,1	19,9	20,2	285	2,02
107	15-aug-17	17:30	-19	1	20,6	20,1	19,9	20,2	285	2,02
108	15-aug-17	17:40	-10	1	20,6	20	19,9	20,2	280	2,016666667
109	15-aug-17	17:50	-1	1	20,6	20	19,9	20,2	277	2,016666667
110	15-aug-17	18:00	10	1	20,6	20,1	19,9	20,2	287	2,02
111	15-aug-17	18:10	20	1	20,5	20,1	19,9	20,2	294	2,016666667
112	15-aug-17	18:20	32	1	20,5	20,1	19,9	20,2	293	2,016666667
113	15-aug-17	18:30	44	1	20,5	20,1	19,9	20,2	287	2,016666667

	A	B	C
1	File name + click ID	Date + Time	
2	20170918202127_B		
3	1	2017-09-18 20:21:36.892	
4	2	2017-09-18 20:21:36.893	
5	3	2017-09-18 20:21:36.894	
6	4	2017-09-18 20:21:36.895	
7	5	2017-09-18 20:21:36.896	
8	6	2017-09-18 20:21:36.897	
9	7	2017-09-18 20:21:36.898	
10	8	2017-09-18 20:21:36.899	
11	9	2017-09-18 20:21:36.900	
12	10	2017-09-18 20:21:36.901	
13	11	2017-09-18 20:21:36.902	
14	12	2017-09-18 20:21:36.903	
15	13	2017-09-18 20:21:36.904	
16	14	2017-09-18 20:21:36.905	
17	15	2017-09-18 20:21:36.906	
18	16	2017-09-18 20:21:36.907	
19	17	2017-09-18 20:21:36.908	
20	18	2017-09-18 20:21:36.909	
21	19	2017-09-18 20:21:36.910	
22	20	2017-09-18 20:21:36.911	
23	21	2017-09-18 20:21:36.912	

	A	B	C
615	613	2017-09-18 21:04:33.733	
616	614	2017-09-18 21:04:40.208	
617	615	2017-09-18 21:05:05.276	
618	616	2017-09-18 21:05:22.470	
619	617	2017-09-18 21:07:00.639	
620	618	2017-09-18 21:07:06.259	
621	619	2017-09-18 21:07:06.935	
622	620	2017-09-18 21:07:35.574	
623	621	2017-09-18 21:08:00.789	
624	622	2017-09-18 21:08:44.136	
625	623	2017-09-18 21:08:44.160	
626	624	2017-09-18 21:08:53.460	
627	625	2017-09-18 21:09:10.313	
628	626	2017-09-18 21:09:44.697	
629	627	2017-09-18 21:09:49.962	
630	628	2017-09-18 21:10:10.328	
631	629	2017-09-18 21:10:33.653	
632	20170918212128		
633	1	2017-09-18 21:22:39.023	
634	2	2017-09-18 21:22:39.710	
635	3	2017-09-18 21:24:09.707	
636	4	2017-09-18 21:24:13.040	
637	5	2017-09-18 21:24:16.855	

	ID	DATE	HOUR	#CLICKS	WIND DIRECTION (degrees)	WIND SPEED (m/s)	BEAUFORT	WATERTEMPERATURE	WATERHEIGHT	TIDE (0=OUT,1=IN,2=LT, 3=HT)	TIDE (14)	TIDE (8)
2	1	8/15/2017	0	148	143.00	3.43	3	20.00	-104.17	0	0	0
3	2	8/15/2017	1	93	150.50	3.43	3	20.05	-154.50	0	0	0
4	3	8/15/2017	2	49	152.00	3.57	3	20.07	-171.67	2	3	2
5	4	8/15/2017	3	140	146.33	3.55	3	20.07	-147.33	1	1	1
6	5	8/15/2017	4	549	131.33	3.15	2	20.04	-103.67	1	1	1
7	6	8/15/2017	5	976	125.17	3.12	2	19.98	-45.67	1	1	1
8	7	8/15/2017	6	1765	124.17	2.75	2	20.00	37.67	1	1	1
9	8	8/15/2017	7	1894	165.00	3.08	2	19.89	106.17	1	1	1
10	9	8/15/2017	8	868	216.67	5.00	3	19.82	140.00	3	12	7
11	10	8/15/2017	9	926	305.33	2.05	2	19.84	127.33	0	0	0
12	11	8/15/2017	10	415	70.33	4.23	3	19.87	73.67	0	0	0
13	12	8/15/2017	11	1005	116.33	4.57	3	19.85	13.17	0	0	0
14	13	8/15/2017	12	3250	151.50	5.32	3	19.89	-51.17	0	0	0
15	14	8/15/2017	13	141	247.17	1.70	2	19.96	-102.00	0	0	0
16	15	8/15/2017	14	557	286.50	1.70	2	20.00	-119.83	2	5	3
17	16	8/15/2017	15	1164	270.17	5.18	3	19.99	-105.50	1	1	1
18	17	8/15/2017	16	1331	286.00	4.15	3	19.96	-69.33	1	1	1
19	18	8/15/2017	17	765	274.17	5.28	3	19.99	-23.00	1	1	1
20	19	8/15/2017	18	450	260.50	5.97	4	19.99	38.83	1	1	1
21	20	8/15/2017	19	370	256.67	6.22	4	19.92	110.67	1	1	1
22	21	8/15/2017	20	863	253.67	5.45	3	19.92	150.00	1	1	1
23	22	8/15/2017	21	498	247.50	2.57	2	19.91	149.67	3	13	7
24	23	8/15/2017	22	175	233.17	2.52	2	19.89	106.50	0	0	0
25	24	8/15/2017	23	181	231.17	3.10	2	19.87	42.50	0	0	0
26	25	8/16/2017	0	173	226.67	4.20	3	19.87	-24.00	0	0	0
27	26	8/16/2017	1	98	228.17	3.75	3	19.90	-86.50	0	0	0
28	27	8/16/2017	2	78	211.33	4.13	3	19.92	-130.17	0	0	0
29	28	8/16/2017	3	90	203.83	3.77	3	19.94	-141.67	2	3	2
30	29	8/16/2017	4	183	204.17	3.97	3	19.99	-114.50	1	1	1
31	30	8/16/2017	5	280	196.50	3.57	3	19.91	-71.83	1	1	1
32	31	8/16/2017	6	391	187.33	3.65	3	19.86	-18.67	1	1	1
33	32	8/16/2017	7	868	185.50	3.33	2	19.83	54.00	1	1	1
34	33	8/16/2017	8	1215	177.67	2.98	2	19.77	114.67	1	1	1
35	34	8/16/2017	9	715	153.50	3.15	2	19.74	143.00	3	12	7
36	35	8/16/2017	10	407	160.17	4.58	3	19.79	130.50	0	0	0
37	36	8/16/2017	11	84	158.17	2.88	2	19.78	79.67	0	0	0
38	37	8/16/2017	12	181	176.00	3.37	2	19.84	17.50	0	0	0
39	38	8/16/2017	13	84	196.83	3.90	3	19.93	-45.67	0	0	0

Appendix V: Data preparation code book

Variable/Factor	Description	Scale	Values SPSS
Time of day 1	24 Hour blocks ranging from 0 to 23	Nominal	
Time of day 2	Hours classed into 12 pairs	Nominal	0= 0:00-1:59:59 1=2-3:59:59 2=4-5:59:59 3=6-7:59:59 4=8-9:59:59 5=10-11:59:59 6=12-13:59:59 7=14-15:59:59 8=16-17:59:59 9=18-19:59:59 10=20-21:59:59 11=22-23:59:59
Time of day 3	Hours classed into 6 groups of 4 hours	Nominal	0= 0:00-3:59:59 1=4-7:59:59 2=8-11:59:59 3=12-15:59:59 4=16-19:59:59 5=20-23:59:59
Clicks per hour	Number of detected clicks per hour	Ratio	
Water height	Water height in centimetres (NAP)	Ratio	
Tide 1	High tide, low tide, incoming tide and outgoing tide	Nominal	0=incoming tide 1=outgoing tide 2=high tide within the hour 3=low tide within the hour
Tide 2	High tide, low tide, incoming tide and outgoing tide	Nominal	0= Incoming Tide 1= Outgoing Tide 2= High Tide between 0-19:59min 3= High Tide between 20-39:59min 4= High Tide between 40-59:59min 5= Low Tide between 0-19:59min 6= Low Tide between 20-39:59min 7= Low Tide between 40-59:59min
Tide 3	High tide, low tide, incoming tide and outgoing tide	Nominal	0= Incoming Tide 1= High Tide between 0-9:59min 2= High Tide between 10-19:59min 3= High Tide between 20-29:59min 4= High Tide between 30-39:59min 5= High Tide between 40-49:59min 6= High Tide between 50-59:59min 7= Outgoing Tide 8= Low Tide between 0-9:59min 9= Low Tide between 10-19:59min 10= Low Tide between 20-29:59min 11= Low Tide between 30-39:59min 12= Low Tide between 40-49:59min 13= Low Tide between 50-59:59min
Water temperature	Water temperature in Celsius	Ratio	
Wind speed	Wind speed in m/s	Ratio	
Wind direction 1	Wind direction in classes of 45°	Nominal	0= 0 - 44,99 ° 1= 45 – 89,99 ° 2= 90 – 134,99 ° 3= 135 – 179,99 ° 4= 180 – 224,99 ° 5=225 – 269,99 ° 6= 270 – 314,99 ° 7= 315 - 359,99 °
Wind direction 2	Wind direction in degrees 90°	Nominal	0= 0 - 89,99 ° 1= 90 – 179,99 ° 2= 180 – 269,99 ° 3= 270 – 359,99 °

Appendix VI: Sample data exploration graphs

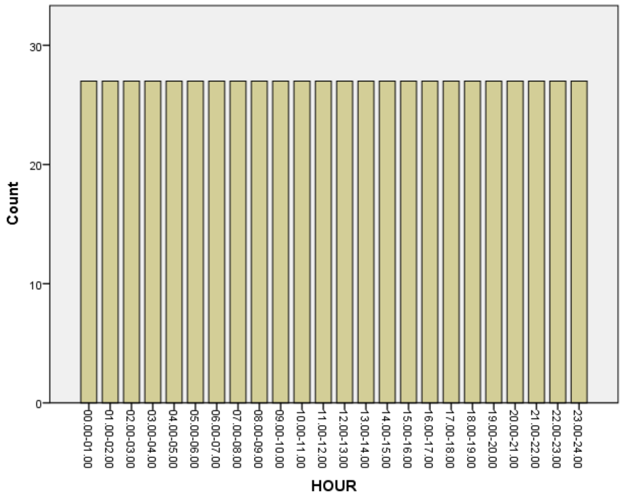


Figure E: count of individual hours included in the sample

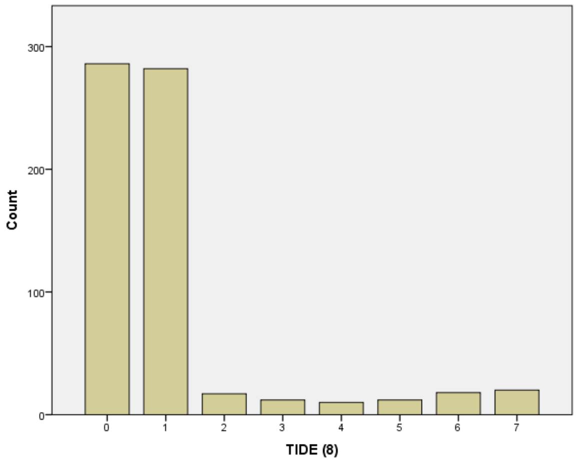


Figure F: count of tide classes included in the sample

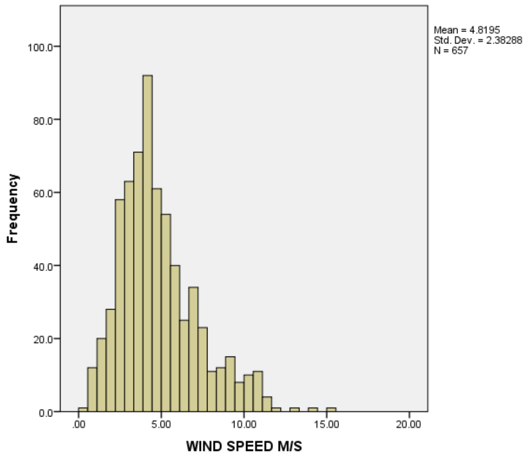


Figure G: count of wind speed (meters per second) frequencies included in the sample

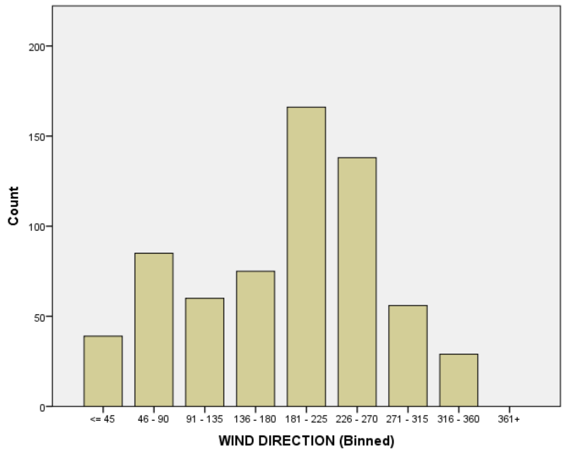


Figure H: count of wind direction classes (degrees) included in the sample

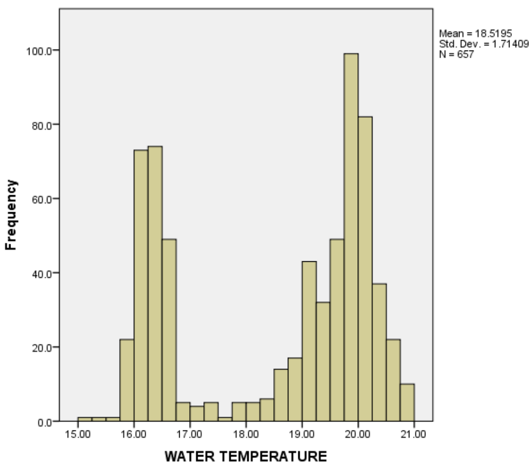


Figure I: frequencies of water height (meters) included in the sample

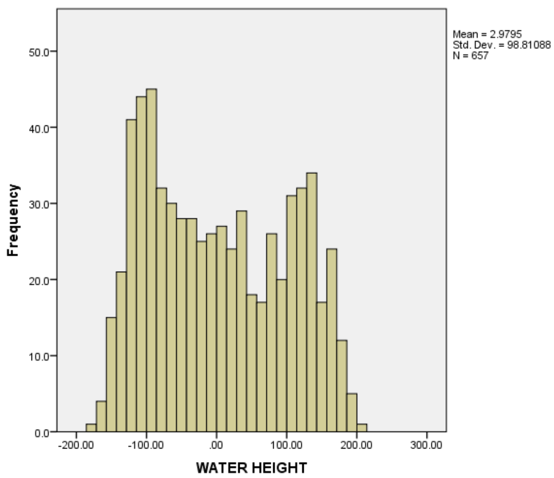


Figure J: Frequencies of water temperatures (degrees Celsius) included in the sample

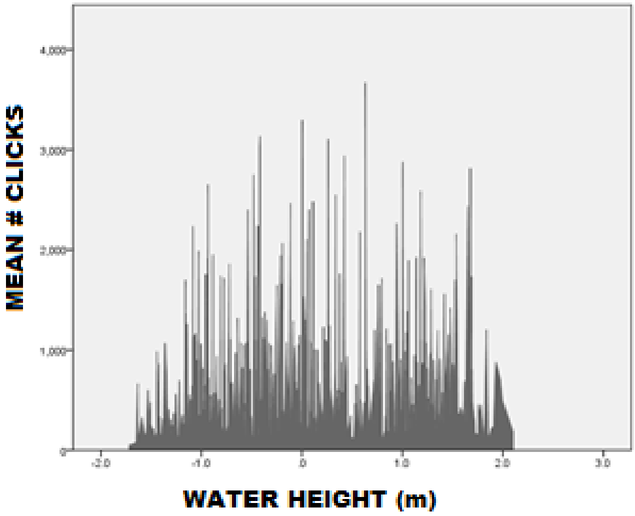


Figure K: Mean Clicks against Water Height

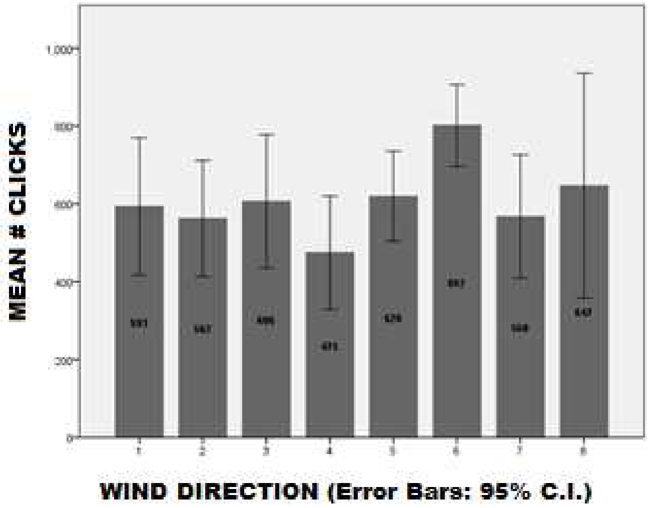


Figure L: Mean Clicks against wind direction (8 classes)

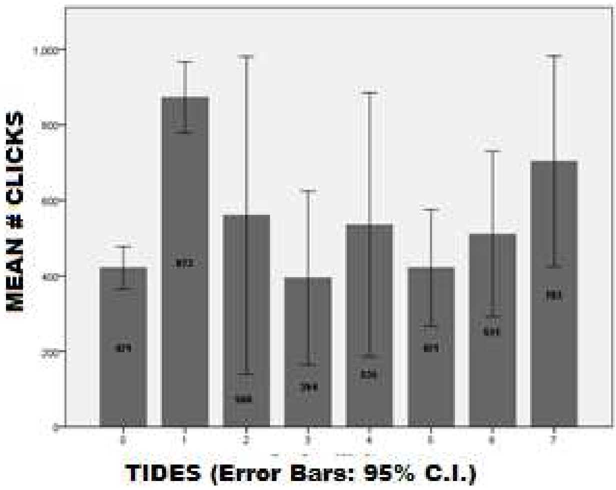


Figure M: Mean Clicks against tides (8 classes)

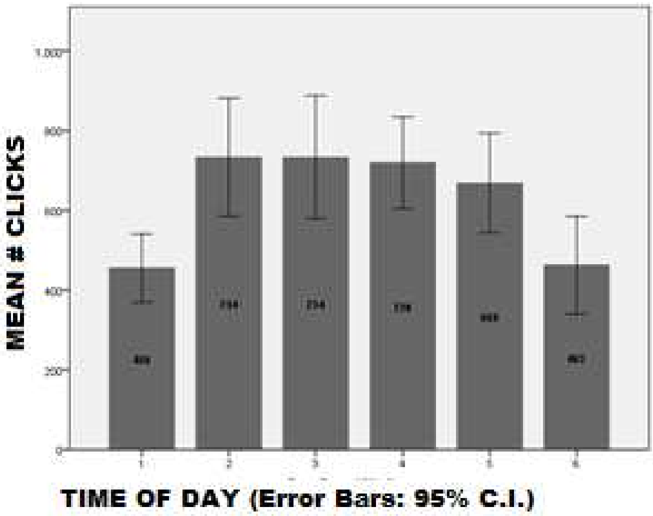


Figure N: Mean Clicks against Time of day (6 classes)

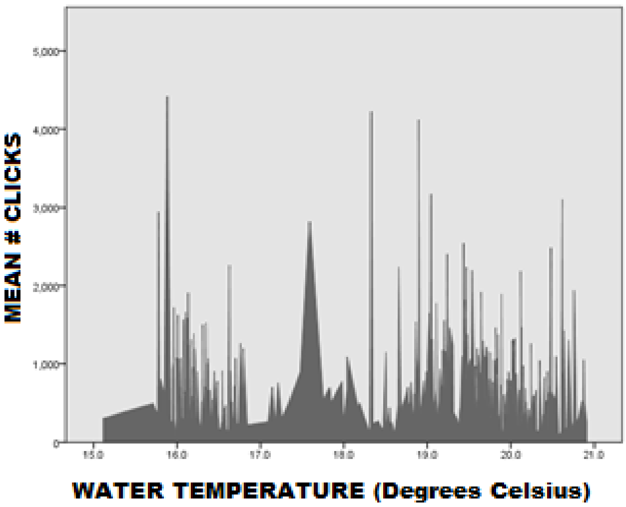


Figure O: Mean Clicks against water temperature

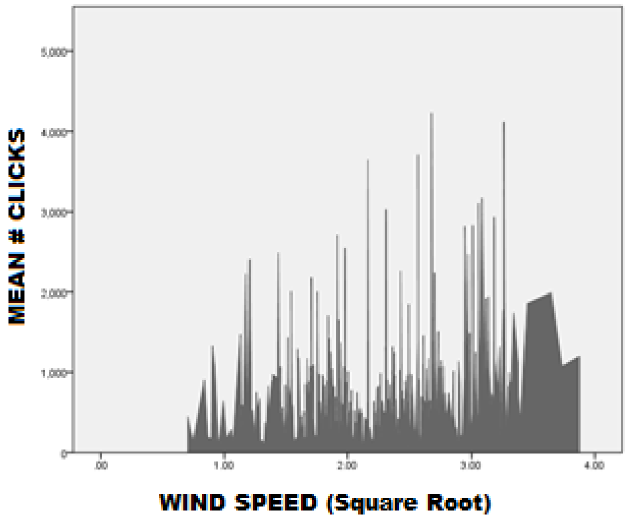
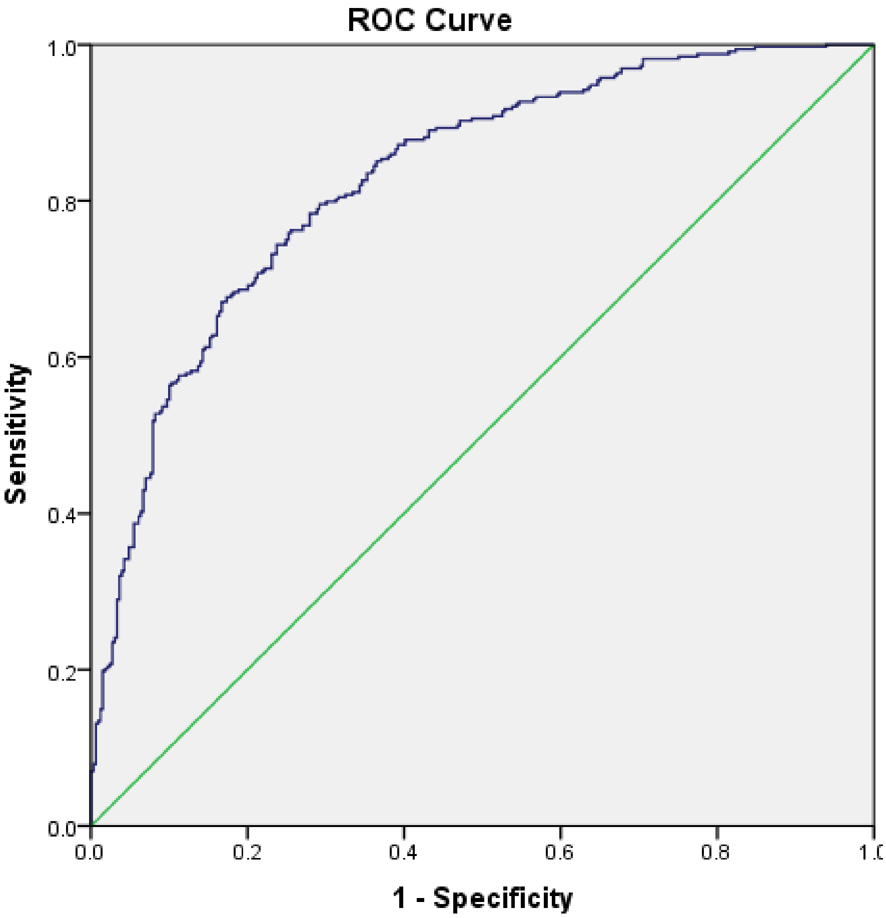


Figure P: Mean Clicks against wind speed (Square root)

Appendix VI: ROC-curve and coordinates table



Coordinates of the Curve

Test Result Variable(s): Predicted probability

Positive if Greater Than or Equal To[a]	Sensitivity	1 - Specificity
0E-7	1.000	1.000
.0072066	1.000	.997
.0159339	1.000	.994
.0241585	1.000	.991
.0245438	1.000	.988
.0249190	1.000	.985
.0263205	1.000	.982
.0303452	1.000	.979
.4756188	.762	.258
.4770777	.762	.255
.4779802	.759	.255
.4794557	.759	.252
.4805889	.756	.252
.4811954	.753	.252
.4823444	.750	.252
.4829711	.750	.249
.4833281	.747	.249
.4885250	.744	.249
.4945319	.744	.246
.4964052	.744	.243
.4972800	.744	.240
.4976258	.744	.237
.4988838	.741	.237

Appendix VIII: Logistic regression model outcome

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Wind Speed (square root)	1.063	.211	25.451	1	.000	2.896	1.916	4.378
Water Temperature	.189	.058	10.543	1	.119	1.208	1.078	1.353
Water Height	.004	.001	13.808	1	.015	1.004	1.002	1.007
Hour			27.274	5	.000			
Hour class 1 (0-4h)	.156	.361	2.432	1	.000	1.755	.865	3.560
Hour class 2 (4-8h)	.867	.358	5.879	1	.117	2.380	1.181	4.799
Hour class 3 (8-12h)	1.538	.384	16.060	7	.000	4.654	2.194	9.874
Hour class 4 (12-16h)	1.618	.386	17.553	1	.052	5.044	2.366	10.754
Hour class 5 (16-20h)	.580	.370	2.455	1	.238	1.786	.864	3.691
Wind direction			33.807	7	.168			
Wind direction class 1 (0-45)	1.125	.579	3.770	1	.849	3.080	.989	9.585
Wind direction class 2 (45-90)	.627	.531	1.391	1	.673	1.872	.660	5.304
Wind direction class 3 (90-135)	.752	.545	1.901	1	.002	2.121	.728	6.174
Wind direction class 4 (135-180)	-.102	.534	.036	1	.562	0.903	.317	2.570
Wind direction class 5 (180-225)	.206	.488	.179	1	.000	1.229	.472	3.199
Wind direction class 6 (225-270)	1.574	.499	9.954	1	.001	4.827	1.815	12.834
Wind direction class 7 (270-315)	.313	.540	.337	1	.000	1.368	.475	3.939
Tide			85.281	7	.000			
Tide class 1 (Outgoing)	-.898	.549	2.682	1	.101	.407	.139	1.193
Tide class 2 (Incoming)	1.003	.559	3.221	1	.073	2.727	.912	8.157
Tide class 3 (Low-Incoming)	-.711	.860	.684	1	.408	.491	.091	2.651
Tide class 4 (Outgoing-Low-Incoming)	-.189	.900	.044	1	.834	.828	.142	4.830
Tide class 5 (Outgoing-Low)	1.116	.908	1.511	1	.219	3.053	.515	18.099
Tide class 6 (High-Outgoing)	-.924	.834	1.226	1	.268	.397	.077	2.037
Tide class 7 (Incoming-High-Outgoing)	-.888	.751	1.397	1	.237	.412	.094	1.794
Constant	-7.252	1.506	23.182	1	.000	.001		

Appendix IX: Nagelkerke and Hosmer & Lemeshow results

Table A: Nagelkerke results

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	684.399[a]	.291	.389

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Table B: Hosmer and Lemeshow results

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.155	8	.630

Appendix X: Estimated Marginal Means

Estimates

Time of Day (6 classes)	Mean	Std. Error	95% Wald Confidence Interval	
			Lower	Upper
1	.37	.063	.25	.50
2	.44	.069	.31	.58
3	.61	.072	.46	.74
4	.63	.072	.48	.75
5	.37	.067	.25	.51
6	.25	.064	.14	.39

Covariates appearing in the model are fixed at the following values: WSsqrt=2.1293; WATERTEMP=18.519559; WATERHEIGHT=2.979452

Figure Q: EMM for time of day (Hours)

Estimates

Wind Direction (8 classes)	Mean	Std. Error	95% Wald Confidence Interval	
			Lower	Upper
1	.58	.100	.38	.76
2	.46	.082	.31	.61
3	.49	.089	.32	.65
4	.29	.068	.17	.44
5	.35	.058	.25	.47
6	.68	.059	.56	.79
7	.38	.082	.24	.55
8	.31	.103	.15	.53

Covariates appearing in the model are fixed at the following values: WSsqrt=2.1293; WATERTEMP=18.519559; WATERHEIGHT=2.979452

Figure R: EMM for Wind direction (degrees)

Estimates

Tide (8 classes)	Mean	Std. Error	95% Wald Confidence Interval	
			Lower	Upper
0	.28	.032	.22	.34
1	.72	.032	.65	.78
2	.32	.138	.12	.62
3	.44	.169	.17	.75
4	.74	.137	.42	.92
5	.27	.136	.09	.59
6	.28	.117	.11	.55
7	.49	.134	.25	.73

Covariates appearing in the model are fixed at the following values: WSsqrt=2.1293; WATERTEMP=18.519559; WATERHEIGHT=2.979452

Figure S: EMM for Tides