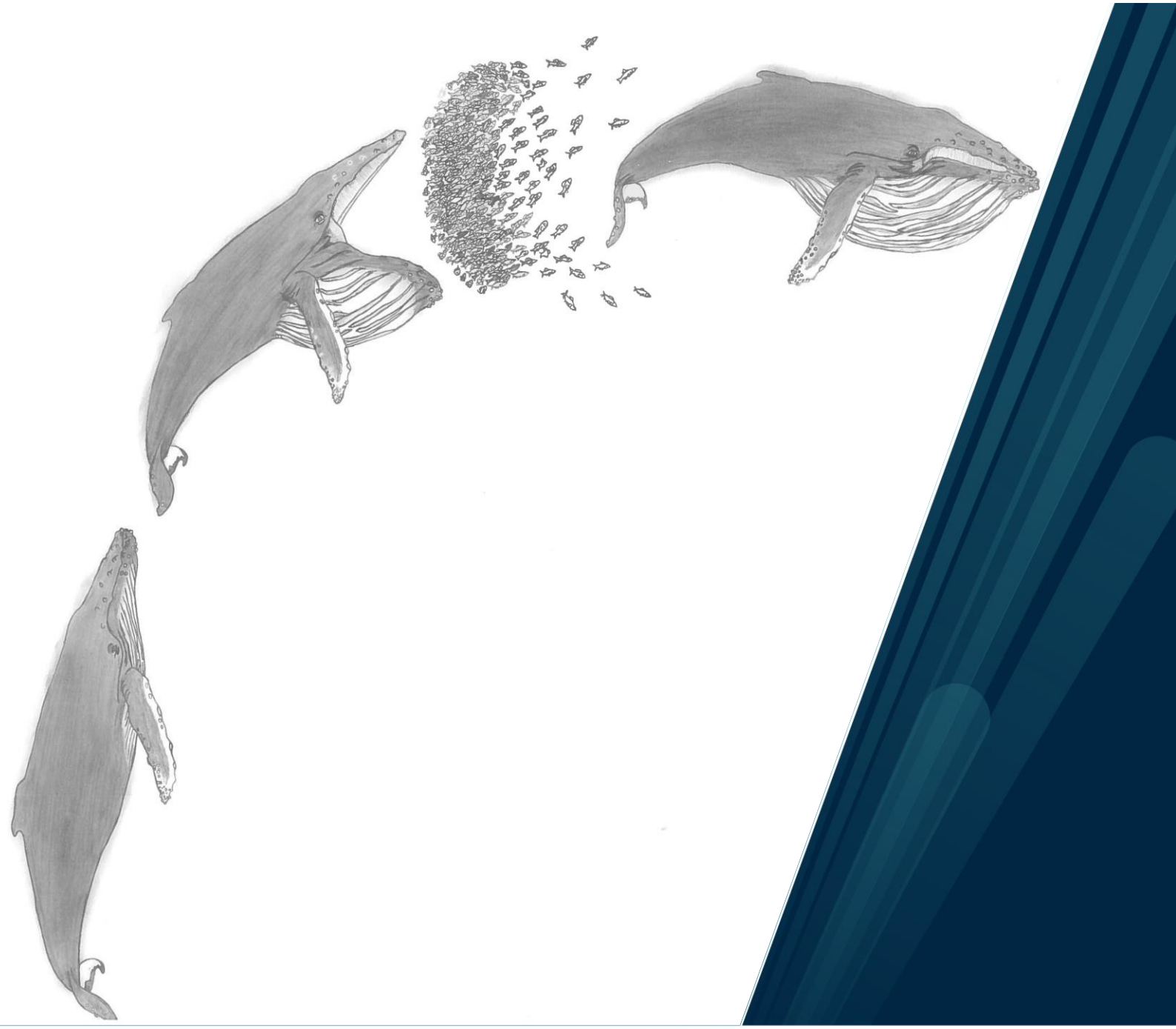
UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries and Economics
Department of Arctic and Marine Biology

Foraging behaviour of humpback whales (*Megaptera novaeangliae*): Automatic detection of feeding lunges from two-dimensional data

Maren Andrea Pedersen

Master's thesis in Biology, Bio-3950, August 2020



Foraging behaviour of humpback whales (*Megaptera novaeangliae*): Automatic detection of feeding lunges from two-dimensional data

Maren Andrea Pedersen
Master of Science in Biology – Marine Ecology and Resource Biology
August 2020

Supervisors
Ulf Lindstrøm: Institute of Marine Research, UiT – The Arctic University of Norway
Martin Biuw: Institute of Marine Research

Front page: Lunging humpback whale. Illustration by Maren Andrea Pedersen

Table of Contents

1 Abstract 1

2 Introduction 3

3 Materials & Method 6

3.1 Study site 6

3.2 Instrumentation 6

3.2.1 Tags 6

3.2.2 Tag deployment 9

3.3 Data analysis 10

3.3.1 Data calibration 10

3.3.2 Three-dimensional lunge detector 10

3.3.3 Two-dimensional lunge detection 11

3.3.4 Validation of the two-dimensional detector 14

3.4 Ethical statement 17

4 Results 18

4.1 Tag deployment 18

4.2 Two-dimensional lunge detector 18

5 Discussion 26

6 Conclusion 31

Works cited 33

Appendix 40

List of Tables

Table 1 The number of whales tagged per year in the instrumentation period, and the tags used. 7

Table 2 Definitions of the seven parameters used to identify lunges in the 2D lunge detector. 12

Table 3 The confusion matrix output from running the 2D detector algorithm, with predicted and observed lunge detections given as True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN). 15

Table 4 Observed and predicted detections, given as true positive (TP), false negative (FN), false positive and true negative (TN). The accuracy and precision of detections, as well as the rates of false and true positive and negative detections are also given. 16

Table 5 The default settings of the parameters, and parameter range tested for the optimisation. 18

Table 6 Parameter settings for optimisation of each individual whale tagged with accelerometers. Also, the two parameter settings that performed best overall for all the whales are annotated as Standardised 1 and 2. 19

Table 7 The results of the AUC analysis for the default and optimised settings for each individual whale. The AUC for the standardised settings run on each whale. 20

List of Figures

Figure 1 Map of Kaldfjorden in relation to Tromsø. The study area is marked with a white square on the map in the lower left-hand corner. 6

Figure 2 Humpback whale equipped with a LL data recorder. The three axes, at which acceleration is recorded in the accelerometer is illustrated along with the movements (heading, roll and pitch). Illustration by Maren Andrea Pedersen. 8

Figure 3 Depth (a), depth change (b), weighted power (c) and reconstructed signal (d) of a single dive, number 173, found in the dive profile of mn16_Jan25b. Potential lunges, here four, are indicated across all four panels by the grey dashed lines. 13

Figure 4 The wavelet power spectrum of the dive in Figure 3, with three potential lunges (black dotted lines). The user specified Fourier period ranging from 8 to 32 seconds (horizontal dashed lines) with the peak at 16 (horizontal solid line). The power of the signal is represented by the colour index, the index value is relative to the data. 14

Figure 5 Two composite ROC curves for the results of all 16 whales simultaneously, with each line corresponding to an individual whale. The ROC was run based on the Default (a) and Standardised 2 (b) settings. The diagonal line in both plots represents an AUC of 0.5..... 22

Figure 6 Dive profiles and lunge detections (2D: red circles and 3D: green circles) of the 16 whales instrumented with accelerometers using the Standardised 2 settings. The dive profiles are created based on time-depth data from the multisensory tags (LL and DTAG). 23

Figure 7 Dive profiles and lunge detections (green circles) of 36 whales instrumented with TDRs. 25

Acknowledgements

Firstly, I would like to thank my supervisors, Ulf Lindstrøm and Martin Biuw, for the time they dedicated to help me with my thesis and allowing me to work on such an interesting project. Thank you for the meetings and read-throughs, for the guidance and help with quarrelsome R. And lastly, thank you so much for your understanding.

I would also like to show my gratitude to Takashi Iwata and Rene Swift who helped with the processing and analysis of some of the accelerometer data. The time you took talking me through your methods and furthering my understanding of these data was invaluable. A thank you to Lars Kleivane and Audun Rikardsen for participating in the data collection and contributing some of the tags used. I wish to give an acknowledgement to the weShare project, Fram Fjord og Kyst flagship and UiT- The Arctic University of Norway for funding the data collection. Furthermore, I want to acknowledge Akvaplan-niva (Tromsø, Norway), the Institute of Marine Research (Tromsø, Norway), the Sea Mammal Research Unit (University of St. Andrews, Scotland) and the Nationl Institute of Polar Research (Tokyo, Japan) who collected the data used in this thesis.

Last, but not least, a big thank you to my friends and family for their patience and support throughout this process. These years have been stressful, but it was always made better by your smiles and encouragement.

1 Abstract

Rorqual whales (Family Balaenopteridae) forage almost exclusively by performing so-called feeding lunges. It is difficult to study rorqual foraging behaviour through direct observation because most of the lunges are carried out deep in the water column. The introduction of high-resolution digital tags recording three-dimensional (3D) acceleration has allowed for the collection of complex movement data, increasing our understanding of their foraging behaviour. Lunges can be detected from specific movement signals in the 3D acceleration data. However, there are still datasets obtained using simpler tags, such as time-depth recorders (TDR) that have yet to be analysed, and there is a lack of automated methods for analysis. In this study, an algorithm allowing for automatic detection of lunges from these two-dimensional (2D) depth-time diving profiles is developed. This detector was applied on 16 humpback whales (*Megaptera novaengliae*) tagged with high-resolution multisensory tags. The data was subset into a simpler 2D format and then validated against lunge detections on the same data using 3D detectors. Optimisation of the 2D detector was done by manually changing the algorithm parameter settings, and then using ROC and AUC to find the best possible settings. The optimisation found much individual variation, with optimised settings resulting in AUCs ranging from 0.499 to 0.805. The detector was then run on data from 36 TDR tagged whales. The detector performed relatively well and have the potential to help with decreasing time and increasing standardisation of dive data analyses. Also, this method can be useful for adding knowledge in relation to rorqual whale foraging behaviour, especially when combined with additional data such as ecological information.

Key words: biologging, accelerometer, TDR, automation, lunge, ROC, humpback whale

2 Introduction

Movement is the fundamental behavioural response by animals to their internal body requirements and external environment. All movement requires energy and allocating energy to a specific act has consequences for lifetime reproductive success and therefore, ultimately, natural selection (Brown et al., 2013). Detailed understanding of an animal's foraging ecology and food requirements depends on information on prey searching strategies, prey types and estimates of prey consumption rates (Carroll et al., 2014; Akiyama et al., 2019). However, it can be difficult to measure and monitor feeding behaviour, especially of aquatic animals that feed in deep waters or areas that are difficult to reach (Johnson & Tyack, 2003; Cooke et al., 2004; Watanabe et al., 2005; Friedleander et al, 2009; Broell et al., 2013; Brown et al., 2013; Womble et al., 2013; Allen et al., 2016; Goldbogen et al., 2017; Cox et al., 2018).

Diving air-breathing animals, such as baleen whales (*Mysticeti*), face the trade-off of balancing their metabolic energy demands associated with prey searching and capture, against the needs to conserve oxygen to avoid entering anaerobic metabolism and incurring an oxygen debt (Acevedo-Gutiérrez, Croll & Tershy, 2002; Halsey et al., 2011; Miller et al., 2012; Hazen et al., 2015). Because oxygen is a limiting factor in marine mammals, they have evolved physiologically to maximise energy storage and minimise oxygen consumption (Acevedo-Gutiérrez et al., 2002; Akiyama et al., 2019). Optimal foraging theory (OFT) predicts that animals should select the foraging strategy or prey that maximise their net energy gain (Charnov, 1976; Doniol-Valcroze et al., 2011; Carroll et al., 2014; Cade et al., 2016; Akiyama et al., 2019). Foraging dive efficiency is determined by the ratio of energy intake through prey ingestion to the energy expended through locomotion, prey capture and basal metabolism (Adachi et al., 2014; Goldbogen et al., 2015; Hazen et al., 2015; Cade et al., 2020). Thus, the time spent foraging depends on prey density and the energetic cost of the dive (Acevedo-Gutiérrez et al., 2002; Goldbogen et al., 2011; Miller et al., 2012; Adachi et al., 2014).

Baleen whales are large-bodied toothless predators, with bilaterally symmetric racks of keratinised plates known as baleen (Bannister, 2009; Goldbogen et al., 2017). They feed on small, mid-trophic level organisms, such as pelagic schooling fish and euphausiids (Bannister, 2009; Hazen et al., 2015; Goldbogen & Madsen, 2018). These filter feeders have evolved highly efficient foraging strategies, that is further aided by their engulfment capacity (Goldbogen, 2010; Goldbogen et al., 2013; Payne et al., 2014). Baleen whales engulf water by elevating their skull while depressing the mandibles, and some increase the amount of water engulfed through extending the ventral groove blubber (VGB) (Brodie, 1993; Bannister, 2009; Goldbogen, 2010;

Goldbogen, Potvin & Shadwick, 2010; Goldbogen et al., 2015; Hazen et al., 2015; Cade et al., 2016; Goldbogen et al., 2017). The feeding efficiency is increased by targeting ocean features and processes that creates high-density, or more easily captured, prey aggregations (Friedleander et al., 2006; Friendlaender et al., 2009; Hazen et al., 2009). Efficient feeding allows these whales to quickly build up lipid reserves required for long-distance migration and fasting periods in association with breeding (Goldbogen, Pyenson & Shadwick, 2007; Friedlaender et al., 2013; Hazen et al., 2015; Goldbogen et al., 2017; Goldbogen & Madsen, 2018; De Weerdt & Ramos, 2020). However, this engulfment capacity introduces a cost through increased drag (Goldbogen et al., 2013; Hazen et al., 2015).

There are three major modes of filter feeding performed by baleen whales: continuous filter feeding, performed by *Balaenidae* (such as bowhead and right whales), suction feeding, a speciality of *Eschirichtiidae* (grey whale), and lunge feeding, which is performed by *Balaenopteridae* (rorquals, e.g blue and humpback whales) (Goldbogen et al., 2013; Goldbogen et al., 2017; Goldbogen & Madsen, 2018). It has been argued that lunge feeding has the highest energy expenditure of these feeding modes (Brodie, 1993; Goldbogen et al., 2007; Doniol-Valcroze et al., 2011), due to the acceleration required for successful lunging (Acevedo-Gutiérrez et al., 2002; Simon et al., 2009; Goldbogen, 2010; Hazen et al., 2015; Cade et al., 2016; Goldbogen et al., 2017). The effect of this is that rorqual dives are much shorter than expected for whales of their sizes (Goldbogen et al., 2007; Goldbogen et al., 2010). Furthermore, the total cost of a feeding dive depends on the maximum speed reached prior to mouth opening, and the volume of the engulfed water mass (Goldbogen & Madsen, 2018). Despite high energy cost, lunge feeding is an efficient foraging mode due to the large amounts of prey that can potentially be a single lunge, with the potential of increasing the prey consumption rate within individual feeding bouts (Goldbogen et al., 2010; Goldbogen et al., 2011; Doniol-Valcroze et al., 2011; Goldbogen et al., 2015).

The humpback whale is a relatively well-studied species, foraging on a wide range of relatively mobile prey (Hazen et al., 2015; Cade et al., 2016; Cade et al., 2020), from euphausiids to small schooling fish such as herring, mackerel, sand lance and capelin (Clapham & Mead, 1999; Clapham, 2018). Dives with overall high feeding rates are characterised by steep ascent and descent angles and an increased bottom phase (Goldbogen et al., 2015). This whale exhibits complex foraging behaviours that often include high-speed bursts and acrobatic manoeuvres, facilitated by their large high aspect ratio flippers and low aspect ratio flukes (Clapham & Mead, 1999; Woodward, Winn & Fish, 2006; Goldbogen et al., 2015; Cade et al., 2016; Clapham, 2018). Through the production of large thrust forces, the large flippers allow

for high manoeuvrability and the large tail area supports rapid speed and acceleration (Woodward et al., 2006).

Biologging is the collection of data from free-ranging animals using onboard sensors and data loggers (Cooke et al., 2004; Payne et al., 2014; Allen et al, 2016; Cox et al., 2018). The information one might gain from animal-borne tags depends on the type of sensors they are equipped with and the time resolution at which sensor data are recorded (Cooke et al., 2004; Goldbogen et al., 2013; Goldbogen et al., 2017). Sensors include accelerometers, magnetometers, audio, video and environmental sensors. Digital tags can potentially give detailed high-resolution information about environmental conditions driving behaviour and physiology (Friedleander et al., 2009; Hazen et al., 2009; Brown et al., 2013; Womble et al., 2013; Carroll et al., 2014; Cox et al., 2018). High-resolution multisensory tags and acoustic measurements can help quantify foraging preferences across depth and prey density gradients (Shepard et al, 2008; Brown et al., 2013; Payne et al., 2014; Hazen et al., 2015; Goldbogen et al., 2017; Cox et al., 2018). One way of studying fine-scale movement, such as that related to feeding, is using acceleration data (Miller et al., 2004; Brown et al., 2013; Goldgoben et al., 2013; Carroll et al., 2014; Vivant et al., 2014), which can give important information about biologically and ecologically significant events and periods (Broell et al., 2013; Brown et al., 2013). However, these multisensory tags are quite expensive, and have only recently been introduced to marine mammal researchers, so an alternative is using simpler time-depth recorders (TDR) (Acevedo-Gutierrez et al., 2002). These tags are equipped with a pressure sensor and, like multisensory tags, are programmed to record depth at pre-determined intervals (Fedak, Lovell & Grant, 2001; Womble et al., 2013). These simpler tags have been available for much longer, and consequently many datasets have already been collected and analysed. However, interpretation of these simple two-dimensional (2D) data records has been limited to relatively simple indirect indices of foraging behaviour, such as lower resolution ‘wiggles’ (i.e. vertical changes in depth).

The objective of this study was to evaluate the ability of using simple time-depth data (2D) to correctly detect feeding lunges in humpback whales, which involved parameterising a detection algorithm. The 2D approach was applied to medium-resolution time-depth data obtained from high-resolution time-depth and acceleration data (3D) and evaluated against the high-resolution data where lunges had previously been detected using purpose-built methods. After evaluation of the detection algorithm, the algorithm is applied to a larger dataset obtained from 2D dataloggers, TDRs, to detect lunges on a larger set of deployments.

3 Materials & Method

3.1 Study site

A total of 16 humpback whales were tagged in Kaldfjord and surrounding areas in Northern Norway (Figure 1) in the period 2013-2018. The tagging took place during the winter seasons between 2013 and 2018. In 2013 and 2014, the whales were tagged in November and December, while tags were deployed from November to February in 2015 and in January and February in 2016 to 2018. This study site was chosen because many humpbacks aggregated to feed on large aggregations of overwintering Norwegian spring-spawning herring, which provided a rich and easily accessible food resource for large number of humpback and killer whales.

Figure 1 Map of Kaldfjorden in relation to Tromsø. The study area is marked with a white square on the map in the lower left-hand corner.

3.2 Instrumentation

3.2.1 Tags

The humpback whales instrumented with multisensory tags in this time period (Table 1), were used for further analyses. Out of these whales, 13 were tagged with Little Leonardo (LL) data loggers and three with digital acoustic recording tags (DTAGs) (Table 1). The LLs deployed in 2013 to 2015 were UWE-3MPD3GT (W2000, 30 mm in diameter, 175 mm in length, 140 g in

air; Little Leonardo Corp., Tokyo, Japan). During this period, three whales were tagged using DTAG-3 prototypes, built by the Sea Mammal Research Unit (75 mm in diameter, 175 mm in length, 325 g in air). These tags are a development of the original DTAG described in Johnson & Tyack, 2003. The LL model used between 2016 and 2018 was W1000 (W1000, 26 mm in diameter, 175 mm in length, 140 g in air; Little Leonardo Corp., Tokyo, Japan). These LL models are different, in that W1000 is resistant up to 1000m with a resolution of 0.25m, while W2000 is resistant to 2000m and have a resolution is 0.5m. Both the LLs and DTAGs were equipped with 3-axes accelerometers and magnetometers, pressure, and temperature sensors. Additionally, the LL tags were equipped with a speed sensor and a GPS sensor. Depth data were recorded every second. The accelerometer and magnetometer data were recorded continuously at sampling rates of 20Hz for DTAGs and 32Hz for LL. This sampling-frequency is sufficient to record even quick movement relevant to larger animals such as cetaceans (Broell et al., 2013).

Table 1 The number of whales tagged per year in the instrumentation period, and the tags used.

Year	Number of whales	Tags
2013	2	LL (W2000)
	8	TDR
2014	4	LL (W2000), DTAG
	9	TDR
2015	2	DTAG
	17	TDR
2016	4	LL (W1000)
	2	TDR
2017	3	LL (W1000)
2018	1	LL (W1000)

Accelerometers use a spring-like piezoelectric sensor that becomes deformed by gravity and movements, generating an acceleration signal that represent both static orientation and dynamic movements (Adachi et al., 2014). They usually measure acceleration on three orthogonal axes. These are a) pitch; acceleration along the lateral axis of the animals between ±90 degrees, b) roll; along the longitudinal axis measured between ±180 degrees and c) heading, or heave, corresponding to the dorsal-ventral axis between ±360 degrees (Aoki et al., 2012;

Carroll et al., 2014, Figure 2). By looking at the values of these three axes the body posture of the animal can be identified. The signal for each axis is separated into a low-frequency component, representing the static orientation (posture), and a high-frequency component, representing the dynamic movements. By knowing the animal's posture, the sensor's ability to measure changes in velocity, and locating peaks in the acceleration data, fine-scale movements can be identified (Brown et al., 2013; Adachi et al., 2014). Furthermore, accelerometers can record data continuously at a defined frequency or at pre-defined time averages (Broell et al., 2013).

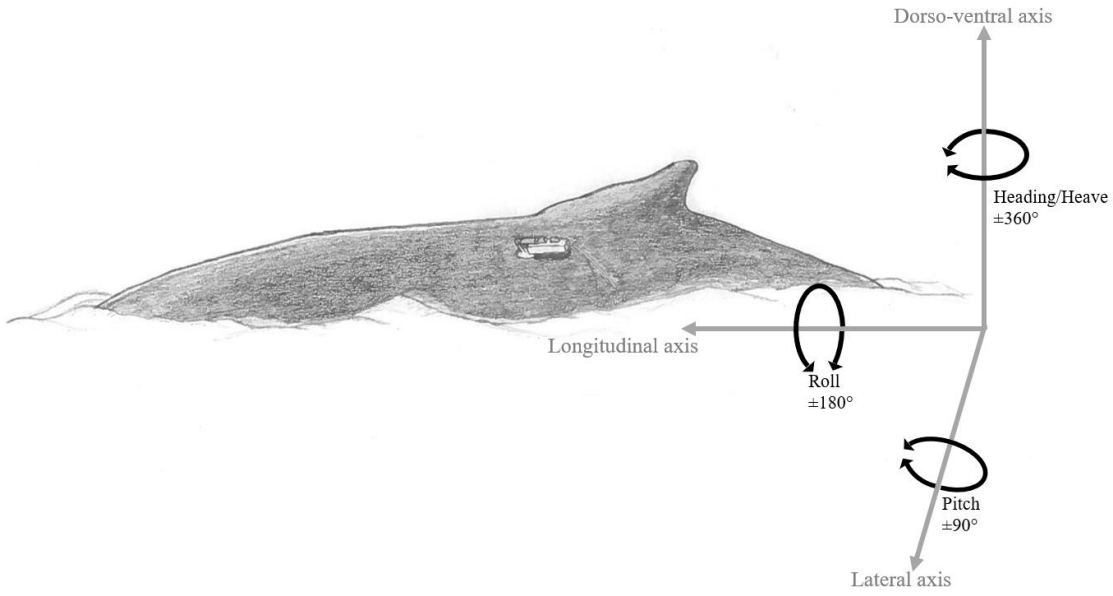


Figure 2 Humpback whale equipped with a LL data recorder. The three axes, at which acceleration is recorded in the accelerometer is illustrated along with the movements (heading, roll and pitch). Illustration by Maren Andrea Pedersen.

In addition to these multisensory tags, several TDR were also deployed from 2013 to 2016 (Table 1), 36 of which will be used for further analyses in this thesis. These tags record time and have a pressure sensor (Fedak et al., 2001; Doniol-Valcroze et al., 2011; Womble et al., 2013) as well as a sensor to detect whether the tag is in the water or air. Furthermore, it is possible to geolocate them. Like the more complex tags discussed previously, TDRs can also be deployed for short or longer periods of time. Also, depending on how long the planned deployment is meant to be, these instruments can record at a continuous or predefined time interval. The model of TDR used was TDR10-F-297C (Wildlife Computers, Redmond, WA, United States) and the instruments record depth every second. The TDRs deployed here were a part of a bigger tag package called a Horizontal-Vertical Tag (HVTag), which also contains GPS and VHF sensors (LKARTS-Norway, Skutvik, Norway). The data used in thes thesis

comes from the TDR component, and from this point on, only the TDR will be specifically referred to.

There is a trade-off to keep in mind for both the multisensory tags and the TDRs, which is that logging such high-resolution data limits the time of data recording. However, as the focus of the study was to get detailed data on diving and feeding behaviour, and suction cups were used as the attachment method, it is not an issue that the data is limited to a shorter time period.

3.2.2 Tag deployment

The multisensory, archival tags were attached to the whales using one (LL) or four (DTAGs) silicone suction cups. Suction cups were also used to attach the TDRs. This attachment is the most commonly used method for attaching tags for short-term, high-resolution feeding studies (Johnson & Tyack, 2003; Johnson, de Soto & Madsen, 2009; Canning et al., 2011, Aoki et al., 2012). While any attachment of objects to the skin of an animal may influence the animal's behaviour, suction cups are thought to be less invasive and have negligible effects on the behaviour (Canning et al., 2011, Johnson & Tyack, 2003, Johnson et al., 2009). Tag deployment was done from a small open power boat (6-8m in length), and a hand-held 6m, carbon fibre pole was used to attach the tags to the whales. Whales were generally approached at slow speed from behind, at an angle of approximately 120-160 degrees relative to the nose of the whale. Tags were generally attached beside the dorsal hump of the whale on either the left or the right side. The majority of the TDRs were deployed using an ARTS whale tagger (LKARTS-Norway, Skutvik, Norway). This whale tagger deploys tags using an air-pressure system and allows for deployment at distances two to three times exceeding that of poles. As all these tags were archival, to get the recorded data they were retrieved after detaching from the whale. Retrieval was done using VHF transmitters (Advanced Telemetry Systems, USA) on the tags, and a R-1000 radio telemetry receiver linked to an AF Antronics F150-3FB three element folding yagi antenna (Communication Specialists Inc, USA). Tags were also equipped with a satellite transmitter (SPOT5: Wildlife Computers, USA), to enable tracking in case whales left the fjord for open water while still carrying the tag. Following retrieval, the data stored in the instrument was downloaded to a computer. Tags could then be redeployed on another animal once the batteries had been recharged.

3.3 Data analysis

3.3.1 Data calibration

All statistical analysis was carried out using R (Version 3.6.0, Foundation for Statistical Computing, Vienna, Austria). After retrieving the tags, data had to be calibrated before any analyses could be run. These calibrations were done using the package `tagtools` (DeRuiter, 2018). Attachment of tags with accelerometers should be stable and in a known orientation relative to the whale's body so the sensor recordings can be considered stable and unbiased (Johnson et al. 2009; Halsey et al., 2011; Brown et al., 2013), but this orientation is not known exactly upon attaching. To correct for any offset in tag placement relative to the major axes of the animal, the axes data is converted in from its original "tag frame" to the corrected "whaleframe" using built-for-purpose functions in `tagtools`. Additionally, data were cropped to remove recordings taking place prior to tag deployment and following detachment from the whale. Dives were defined using standard functions in `tagtools`, and minimum depth required for a dive to be recognised was set to four meters as this depth has been used in several dive studies (e.g. Aoki et al., 2012). Similar calibrations are not needed for the TDR data.

3.3.2 Three-dimensional lunge detector

Lunges had already been detected in the 3D acceleration and speed data for 12 of the 16 datasets. In the tag data collected from 2013 to 2015, lunges were detected using speed through standardised approach in the `tagtools` package. The lunges from data collected between 2016 and 2018 were detected using a method based both on speed and jerk, the rate of change in acceleration. In this approach, the moving Coefficient of Values for speed (CVS) and jerk (CVJ) were calculated and positive peaks predefined minimum intervals, sizes and lengths were detected. If the peak of the largest CVS and CVJ matched up within a given time window of each other, it was considered a lunge candidate (LC). Furthermore, if these LCs happened in the context of specified speed and duration it was identified as a lunge event (Iwata, Unpublished). For the four remaining tags, one LL and three DTAGs, the three-dimensional (3D) lunge detection was done using jerk and the detect_peaks function in the `tagtools` package. Like dive detections, the minimum depth for detection of lunges was set to four meters.

3.3.3 Two-dimensional lunge detection

Although there are several methods for automatic detection of lunges in acceleration data, to date no equivalent method exists for simpler time-depth data. The same dataset for which lunges had already been detected from 3D acceleration and/or speed data was used here for testing a novel algorithm for automatic detection of lunges in two-dimensional data. However, instead of acceleration, information from time and pressure sensors was analysed, since data collected by standard TDRs were limited to these two data types. The algorithm, which was implemented in the `dWave` package developed by Martin Biuw at the Institute of Marine Research. Detection was done by identifying signals in the dive profile which could be defined as a lunge in the context of seven parameters (Table 2). The procedure was as follows:

1. A gaussian (normal) weighted running mean was fitted through the time series of depth changes between consecutive time points. The standard deviation of this gaussian smoothing window was set by the user or optimised through the optimisation routine.
2. A wavelet decomposition analysis was run on the residual data, after removal of the smoothed running mean. This used underlying code and methods available in the `WaveletComp` package for R (Roesch & Schmidbauer, 2018).
3. The mean power of the signal at each time point was calculated, weighted across a user-specified (or optimised) range of Fourier periods (Roesch & Schmidbauer, 2018), using a modified gaussian weighting function to account for the logarithmic power scale. Initial lunge candidates were detected as local peaks in this power curve. Outliers could be detected, and flagged in the data, based on the peak powers. For instance, if one peak is significantly lower than the median of all peaks in a dive, it is flagged as a potential outlier. Similarly, peaks near the edges of data (i.e. near the start and end of a dive) can also be flagged (default) or deleted.
4. From experience, detected peaks sometimes tended to lag in time behind the 'true' lunges. To correct for this, the detector also reconstructs the signal from the fitted wavelet object, using data only within the user-specified Fourier period ranges. The resulting data have peaks in the near vicinity of the peaks in the weighted mean power curve. The closest peak preceding a peak detected in the power curve is identified, and its timestamp is included as an alternative detection point.

The settings for some of the parameters were based on the biologically realistic restrictions of the study animal (Goldbogen et al., 2008; Ware et al., 2011; Simon et al., 2012), while others were defined more subjectively and through testing on single dives. The parameter min.pdist is based on the humpback's inter-lunge interval (ILI), which Ware et al. (2011) calculated to range

between 39.5 to 48.7 seconds in their study. Furthermore, another study found the mean ILI to be about 46 seconds (Simon et al., 2012). Here these numbers were rounded to 40 and 49 seconds.

Table 2 Definitions of the seven parameters used to identify lunges in the 2D lunge detector.

Parameters	Definition
match.window	Size of the time-window within which a peak and a lunge is considered a match
smooth.window	Initial window of smoothing for the gaussian run mean filter
pper	Peak wavelet period for weighted power analysis
rper	The range of all periods to be considered in weighted power analysis
min.pdist	Minimum time period between consecutive peaks
min.ppow	Minimum amplitude for a peak to be considered significant
pk.e	Look at width around the peak that should still be considered as a part of the peak.

The figure below (Figure 3) can better illustrate what these parameters refer to. The top panel (Figure 3a) shows the original depth profile of this dive, where the 3D detected 'true' lunges are represented by green dots. Furthermore, the grey dashed lines seen can be used to trace these lunges over the following three panels. This also relates to match.window, as the lunges detected in the 2D detector must occur within a given timeframe from these lines. Figure 3b. is a representation of the smooth.window parameter. The grey line and dots show the difference in depth between consecutives time points, with the gaussian smoothed running mean, the smoothing, indicated by the red line. The higher the value used for smooth.window, the smoother the red line becomes, making the dive profile more even and potentially removing some lunges. However, making the number too small could introduce additional noise to the data, making it more difficult to identify lunges. The weighted power curve across the user-specified Fourier periods, with initially detected lunges represented by the blue dots is shown in the third panel (Figure 3c). Take note that the dots are on the peaks themselves, but within a given period, there are also solid blue lines that still indicate a potential lunge. This graph relates to pper and rper, where the pper are represented by the blue dots, and rper are the blue dotted line. Figure 3c. can also help explain min.pdist and min.ppow. Whether these blue dots will be placed within the Fourier periods (blue dashed lines), depends on if peaks fall outside the set

minimum time interval between consecutive lunges. Furthermore, whether a peak will be detected at all relates to how tall this peak is compared to the average of the curve. The minimum threshold for this is defined by min.ppow. The last panel, Figure 3d, illustrates the reconstructed signal from the wavelet object, across the previously specified Fourier regions. Here, the red dots are the updated lunge positions.

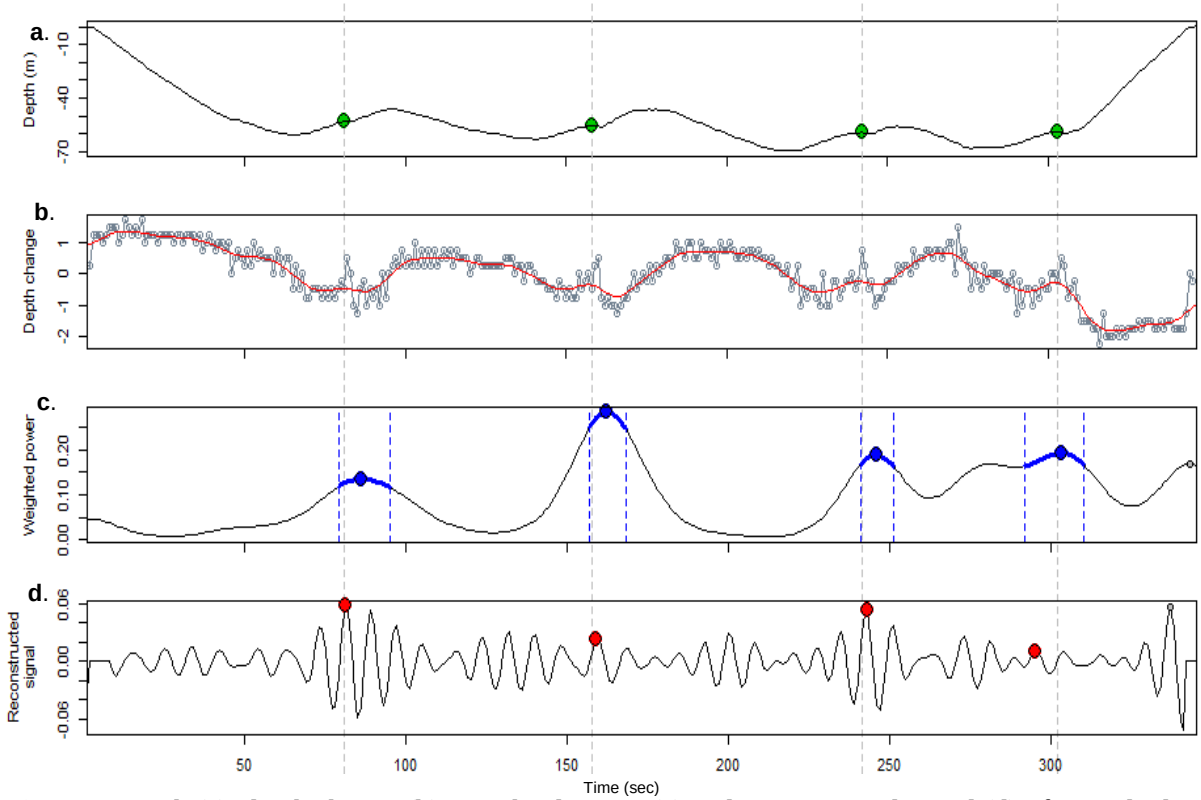


Figure 3 Depth (a), depth change (b), weighted power (c) and reconstructed signal (d) of a single dive, number 173, found in the dive profile of mn16_Jan25b. Potential lunges, here four, are indicated across all four panels by the grey dashed lines.

For additional information regarding some of these parameters, the wavelet power spectrum can be used (Figure 4). The black vertical dotted lines represent the potential lunges, the black solid horizontal line represents the peak, while the horizontal dashed lines are the lower and upper Fourier periods, which together are used to define the Fourier region. This parallels to the blue dots and dashed lines seen in Figure 3c. The colour index in the wavelet graph is an indicator of the power of the signal. The signal is weighted within the user specified Fourier period, and is considered a potential lunge if the peak in the power curve happens within this range.

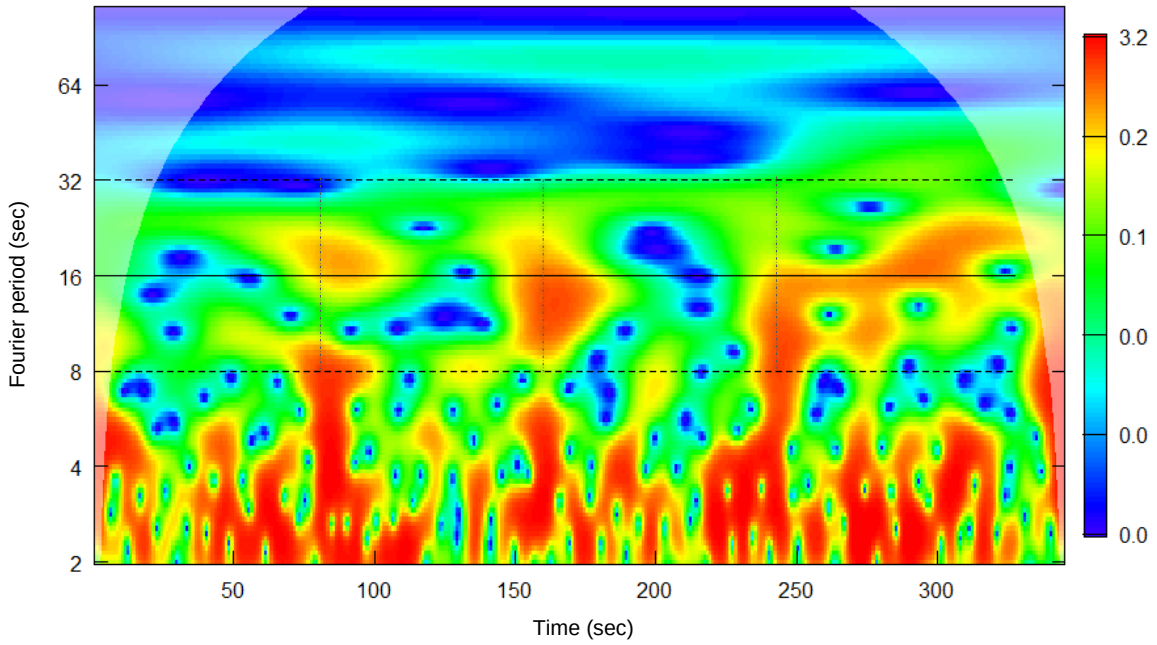


Figure 4 The wavelet power spectrum of the dive in Figure 3, with three potential lunges (black dotted lines). The user specified Fourier period ranging from 8 to 32 seconds (horizontal dashed lines) with the peak at 16 (horizontal solid line). The power of the signal is represented by the colour index, the index value is relative to the data.

The datasets are too large to allow for statistical optimisation at this stage, due to amount of processing time this would take, so the optimisation of the 2D algorithm was done manually. The first step of this optimisation process was changing the parameter values (Table 2) within reasonable ranges and using the values from the confusion matrix to compare the output to the default settings. The tests that yielded the best true positive and false positive rates (TPR and FPR) were kept while other tests were discarded from future use. To get a better idea of the confusion matrix see the Appendix (Table S3) for the outputs. The parameter changes were done for every parameter, both alone and in all possible combinations with the other parameters. For each whale, at least 1000 tests were run. In the longer run, we are aiming to develop more efficient routines for performing automatic optimisations on these data. At this point there was still quite a few numbers of parameter settings that were considered to give good detector outputs, so further optimisation, and validation, was needed.

3.3.4 Validation of the two-dimensional detector

To decide which of the parameter settings found yielded the best 2D detector, the lunge detections based on time and depth was compared to 3D detectors detections based on acceleration. This was done using Receiver Operated Characteristics (ROC) and Area Under the Curve (AUC) analyses.

Table 3 The confusion matrix output from running the 2D detector algorithm, with predicted and observed lunge detections given as True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN).

		Predicted	
		True	False
Observed	True	2D and 3D detected lunge (True Positive (TP))	3D detected lunge (False Negative (FN))
	False	2D detected lunge (False Positive (FP))	No lunge detected (True Negative (TN))

The validity of the optimisation was tested using the ROC analysis, which is a common method for binary classification problems. The ROC analysis was run with the use of the R package `ROCR` (Sing et al., 2005). The analysis was done for the default settings and the optimised settings, for each individual whale. It was also planned to run all accelerometer whales together to get one AUC value and one ROC curve for the default and standardised settings, but due to the size of the data this was not feasible. The dive-profile was also checked manually; and true and false positives and negatives, as well as lunges that both detectors missed and lunges where detections did not match up, were recorded. Below is a list with the relevant metrics used to inform the ROC analyses, as defined by Schrynemackers et al. (2013). These metrics are also presented in the outputs of the 2D detector confusion matrix in Table 4.

- True positive rate (TPR), also be called the *recall* or the *sensitivity*, is equal to the number of true positives divided by the number of actual positives $\frac{TP}{TP+FN}$
- True negative rate (TNR), also called *specificity*, is equal to the number of true negatives divided by the number of actual negatives $\frac{TN}{FP+TN}$
- False positive rate (FPR), corresponding to 1-*specificity*, is equal to the number of false positives divided by the number of actual negatives $\frac{FP}{FP+TN}$
- False negative rate (FNR), also called the *miss*, is equal to the number of false negative divided by the number of actual positives $\frac{FN}{TP+FN}$
- *Precision* is equal to the number of true positives divided by the number of predicted positives $\frac{TP}{TP+FP}$

Table 4 Observed and predicted detections, given as true positive (TP), false negative (FN), false positive and true negative (TN). The accuracy and precision of detections, as well as the rates of false and true positive and negative detections are also given.

		Predicted		
	All detections	3D Detected	3D Not detect	
Observed	2D Detected	TP	FN (Type II Error)	Accuracy $=\frac{\sum TP+\sum TN}{\sum All\ detections}$
	2D Not detect	FP (Type I Error)	TN	Precision $=\frac{\sum TP}{\sum Predicted\ P}$
		TPR $=\frac{\sum TP}{\sum Actual\ P}$	FPR $=\frac{\sum FP}{\sum Actual\ N}$	
		FNR $=\frac{\sum FN}{\sum Actual\ P}$	Specificity $=\frac{\sum TN}{\sum Actual\ N}$	

As Schrynemackers et al. (2013) explained in their paper, achieving a low FPR is desirable as it means the number of FP most likely is smaller than TP, and can allow for decent precision. In the type of data used here, the recording frequency of the instruments corresponds to the placements of data points. In this case, as the tags recorded every second, data points were also collected every second, and as such the number of true negatives is much larger than the number of true positives. To put this into more biological terms, the focus was put on the 2D algorithm to not misidentify signal in the that as a lunge, rather than detecting all feeding events. In energetic or behavioural analyses including feeding rates, this data is often subdivided into "feeding" and "non-feeding" periods. Missing lunges during a feeding period will likely have a negligible effect on the energetics calculations. As such, having many FPs would have a larger effect on the results of analysis, compared to decreasing TP (Allen et al., 2016).

The efficiency of the detector is higher the more convex the ROC curve is (Fawcett, 2006), and simple ROC curves created using a function in the `dWave` package are used to visually evaluate the detector performance. Additionally, the AUC was used to provide a single value representing the performance of the lunge detector, so this could be compared to other possible parameter combinations. The aim of this analysis was to see if specific detector settings performed better than random guessing. This was assumed to be true if the AUC was higher than 0.5 (Fawcett, 2006). Even after running the AUC there were some whales that had several

parameter settings with the same AUC. For these whales, it was decided to report the optimised settings with the least parameter changes from the default. The main goal with this optimisation was to find a setting that allows for the best detector performance across a range of whales, not only perform well for individual whales.

Following the manual optimisation process, the two-dimensional detector was then run on data collected using TDR tags, to detect lunges in these datasets. No validation of these detections could be done since no 3D data were available with which to conduct a 3D detection.

3.4 Ethical statement

All fieldwork involving direct interaction with animals were carried out in accordance with regulations as stipulated by the Norwegian Food Health Authority, as authorised under permit number FOTS-8165.

4 Results

4.1 Tag deployment

For the 16 whales instrumented with accelerometers, the deployment period varied from approximately 2 hours 56 minutes to almost 15 hours (Table S1), with a total deployment time of 123 hours and 32 minutes. The temporal coverage of the data was good i.e. the data from the 16 whales covered the entire day. The deepest dive of these whales was to 283.3m (Table S1). Some of the datasets consisted mainly of shallow dives, which resulted in a poor depth resolution and the dive profiles became more difficult to read in relation to lunges. Additionally, similar unevenness could be seen in some of the deeper dive profiles as well. This was possibly related to the sampling resolution of the tags.

The 36 whales tagged with TDRs had deployment times ranging from close to 14 minutes (Whale_2013Dec06) to almost 64 hours (Whale_2015Nov22, Table S2). The total duration for TDR deployment was 448 hours and 44 min. As with whales equipped with multisensory tags, the TDR data cover all times of the day. For these whales, the deepest dive registered was 265.5m (Table S2). As with the dive profiles of the aforementioned whales, some of these datasets were also noisy, which often related to the shallower dive profiles.

4.2 Two-dimensional lunge detector

The default parameter values as well as those used for the optimisation of the detector can be seen in Table 5. The default values were determined during package development, based on testing on data from whale mn16_Jan25b.

Table 5 The default settings of the parameters, and parameter range tested for the optimisation.

Parameters	Default	Range
match.window	15 s	18 s
smooth.window	4	2-6
pper	16	10-20
rper	20	5-10 & 28-30
min.pdist	40 s	40-49 s
min.ppow	0.05	0.02-0.09
pk.e	0.85	-

The parameter settings that yielded the best results can be seen in Table 6 below. For some of the whales, where match.window is 15, changing this parameter did not have any effect on the matrix output. Due to this 18 was used for the standard settings. This parameter is not carried over to lunge detection in the TDR tagged whales, as these data were not used for validation and therefore lack lunge detections to be compared against.

Table 6 Parameter settings for optimisation of each individual whale tagged with accelerometers. Also, the two parameter settings that performed best overall for all the whales are annotated as Standardised 1 and 2.

Whale ID	match. window	smooth. window	pper	rper	min. pdist	min. ppow	pk.e
Mn13_340a	15	2	18	5	40	0.09	0.85
Mn13_341a	15	3	19	5	43/45	0.07	0.85
Mn14_325a	18	6	19	30	40	0.05	0.85
Mn14_334a	18	4	19	20	40	0.07	0.85
Mn14_335a	15	4	20	20	45	0.05	0.85
Mn14_350a	18	4	20	20	40	0.09	0.85
Mn15_339a	15	6	19	28	45	0.02	0.85
Mn15_341a	18	5	20	16	40	0.09	0.85
Mn16_Jan19a	18	4	19	5	45	0.05	0.85
Mn16_Jan25a	18	4	19	10	45	0.05	0.85
Mn16_Jan25b	18	4	16	10	40	0.05	0.85
Mn16_Jan26	18	2	10	5	40	0.09	0.85
Mn17_022LLa	18	4	16	5	40	0.09	0.85
Mn17_022LLb	18	2	10	5	45	0.09	0.85
Mn17_026LLa	15	4	16	10	40	0.09	0.85
Mn18_013LLa	15	2	16	5	45	0.05	0.85
Standardised 1	18	4	16	5	45	0.05	0.85
Standardised 2	18	4	16	5	40	0.09	0.85

The AUC outputs matching the parameter settings reported above, as well as the AUC results for the default setting, can be seen in Table 7. The Standardised 1 and 2 settings are the settings that performed best for all the whales overall. Apart from mn16_Jan19a, the algorithm was returned with AUC greater than 0.5 for all datasets, meaning that the detector performed

better than randomly allocating lunges. It can be noted that not only was there a clear individual variation in the optimum parameter settings, but there was also a range in how the detector performed when a standardised algorithm was run on all whales (Table 7). Based on the ROC curves (Figure 5) and the AUC results (Table 7) I decided to use the setting for Standardised 2 (Table 6), as it generally performed slightly better than the default setting.

In addition to AUC, the TPR and FPR was calculated for all the whales combined for Standardised 2. This returned an overall TPR of 68.4% and an FPR of 0.025%. Although the FPR seems good the TN in this study were large (Table S3), potentially skewing the result. This was concluded from the TPR and the AUC and ROC outputs.

Table 7 The results of the AUC analysis for the default and optimised settings for each individual whale. The AUC for the standardised settings run on each whale.

Whale ID	AUC Default	AUC Optimised	AUC Standardised 1	AUC Standardised 2
Mn13_340a	0.727	0.869	0.725	0.727
Mn13_341a	0.792	0.867	0.797	0.797
Mn14_325a	0.614	0.616	0.576	0.581
Mn14_334a	0.582	0.612	0.571	0.575
Mn14_335a	0.682	0.738	0.688	0.691
Mn14_350a	0.767	0.765	0.677	0.683
Mn15_339a	0.651	0.679	0.601	0.637
Mn15_341a	0.548	0.639	0.636	0.639
Mn16_Jan19a	0.487	0.535	0.487	0.488
Mn16_Jan25a	0.586	0.709	0.627	0.599
Mn16_Jan25b	0.843	0.843	0.814	0.805
Mn16_Jan26	0.555	0.787	0.743	0.745
Mn17_022LLa	0.596	0.732	0.729	0.732
Mn17_022LLb	0.556	0.681	0.662	0.664
Mn17_026LLa	0.696	0.698	0.697	0.698
Mn18_013LLa	0.68	0.742	0.742	0.682

The Figure 5 shows all ROC curves for the 16 whales instrumented with accelerometers for both default (Figure 5a) and the standardised 2 settings (Figure 5b). These two graphs clearly illustrate the point made earlier of the clear individual variation that can be seen in the

detector performance. When running ROC based on the individual optimised parameter settings, there was much less variation between the whales and the curves did not generally cross below the diagonal line. The graph showing ROC based on their individual optimisation can be seen in the Appendix (Figure S1). Both default and standardised 2 settings resulted in some ROC curves beneath the diagonal line symbolising an AUC of 0.5, suggesting the detector at times performs no better than random. It should be noted, however, that standardised 2 have fewer curves crossing this line, and only one whale that have an AUC below 0.5. The two whales with the poorest ROC curves for the default settings, mn16_Jan26 (light blue curve) and mn17_022LLa (pink curve), are closer to and even above the diagonal line for Standardised 2. This indicates that standardised 2 allows for the detector to work slightly better compared to the default settings. However, 5b also have lines further away from the upper right-hand corner compared to 5a.

The lines in the ROC graphs vary from relatively smooth to more jerky lines with obvious steps. Each step represents detected lunges, and smoother lines correspond to dive profiles with a high number of lunges, which can be seen when comparing Figure 5 to Figure 6. The more even lines are more clustered together and closer to the diagonal line and tend to not go below this line. The smoothest lines correspond to mn14_325a (dark blue), mn14_334a (light blue), mn14_335a (pink), mn15_339a (grey) and mn16_Jan25b (dark blue). Shorter dive records show a greater spread in relation to the diagonal line and come closer to the top left- and bottom right-hand corner. Jerkier lines were more spread out from the diagonal, and the lines with the highest performance, green and grey, corresponds to whale datasets with few lunge detections, mn13_340a and mn17_026LLa (Figure 6).

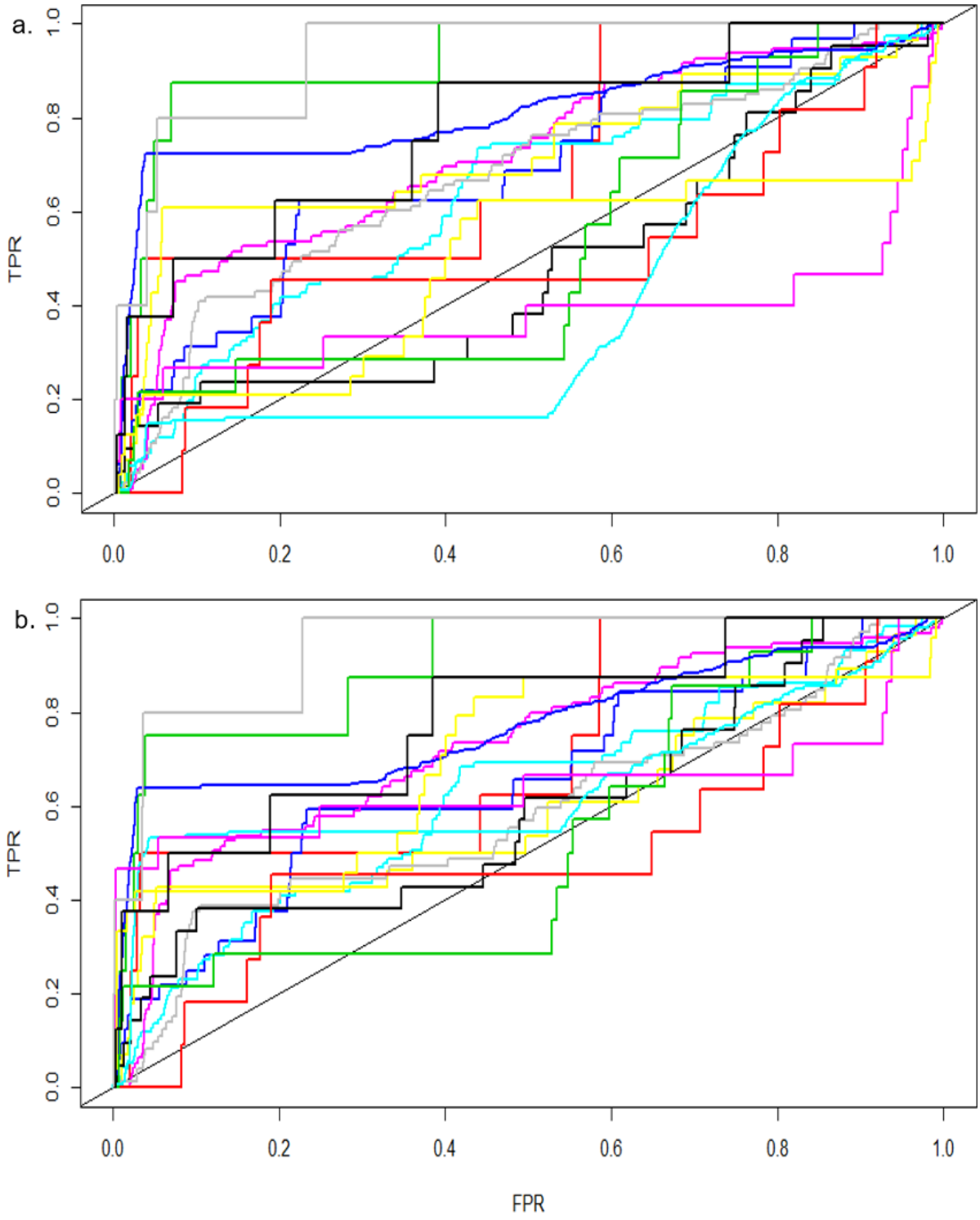


Figure 5 Two composite ROC curves for the results of all 16 whales simultaneously, with each line corresponding to an individual whale. The ROC was run based on the Default (a) and Standardised 2 (b) settings. The diagonal line in both plots represents an AUC of 0.5.

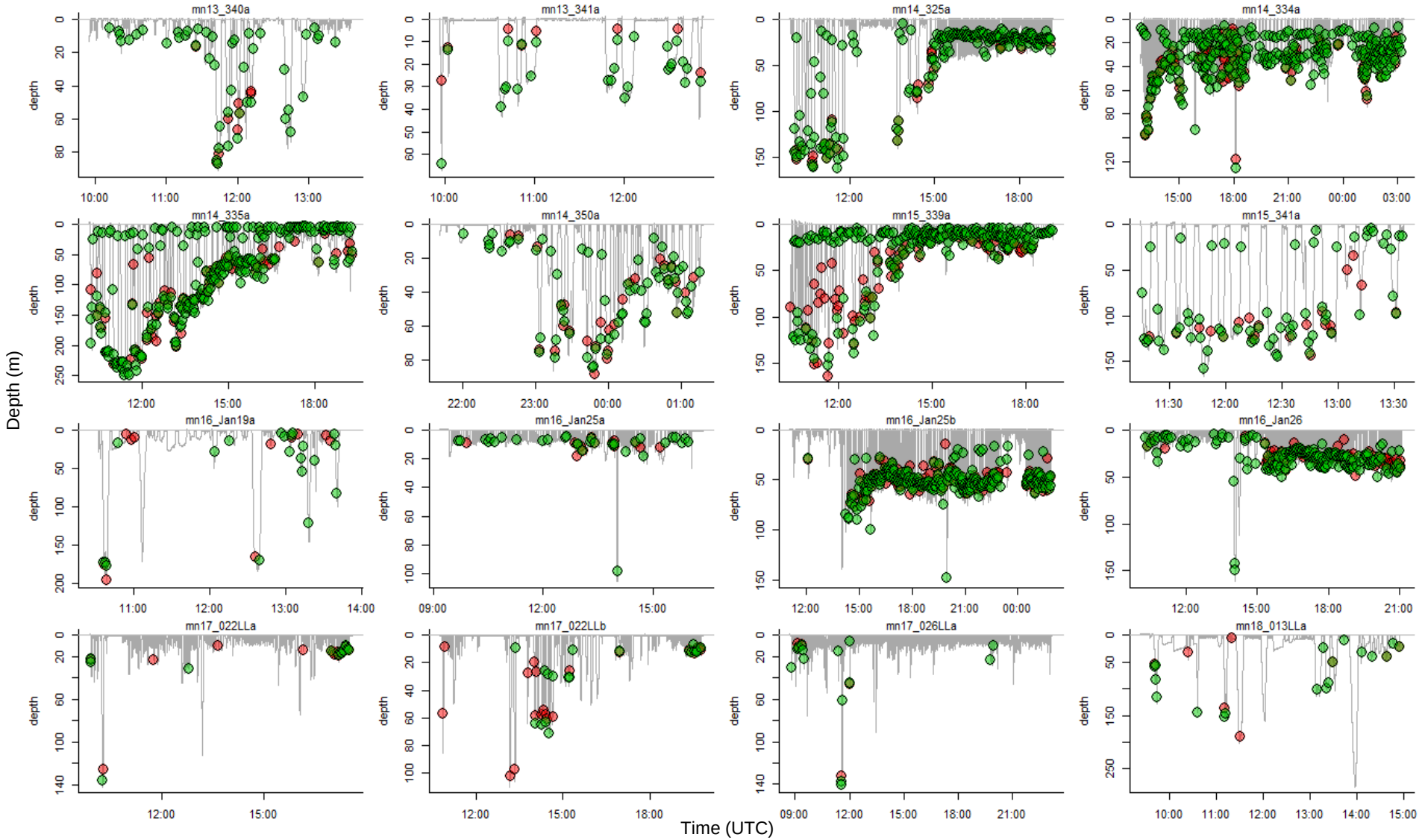


Figure 6 Dive profiles and lunge detections (2D: red circles and 3D: green circles) of the 16 whales instrumented with accelerometers using the Standardised 2 settings. The dive profiles are created based on time-depth data from the multisensory tags (LL and DTAG).

The dive profiles for all 16 accelerometer whales showing both 3D and 2D detectors using the Standardised 2 settings are shown in in Figure 6. The red dots symbolise the 3D detected lunges, while the green dots represent lunges found by the 2D detector. Where these two circles overlap, both detectors have detected a lunge. These graphs also illustrate the substantial individual variation. Figure 6 shows individual variation in number of lunges, dives and the depth at which the lunges takes place. The ROC analysis performed well on data from mn13_340a and mn17_026LLa, but both these records have a small number of lunges detected by the 3D detectors. Mn13_340a only stayed attached to the animal was about three hours. For mn17_026LLa, the tag was deployed for over 14 hours. Comparing Figure 6 to Figure 5, the five whales with the smoothest ROC curves have dive profiles lasting longer than six hours with a higher number of 3D detections.

Figure 7 presents results from applying the optimised 2d detector on the 36 TDR data records. These 36 whales also show variation between individuals, both in depth of dives and lunge activity. TDR deployment happened in November and December in 2013 and 2014, November to February in 2015 and January in 2016. Additionally, based on Figure 7, more lunges seem to have been detected in periods of 12PM to 12AM. Both Figures 6 and 7 illustrate that dives with longer bottom times generally have more lunges detected as well. Duration of deployment should be kept in mind when interpreting these figures.

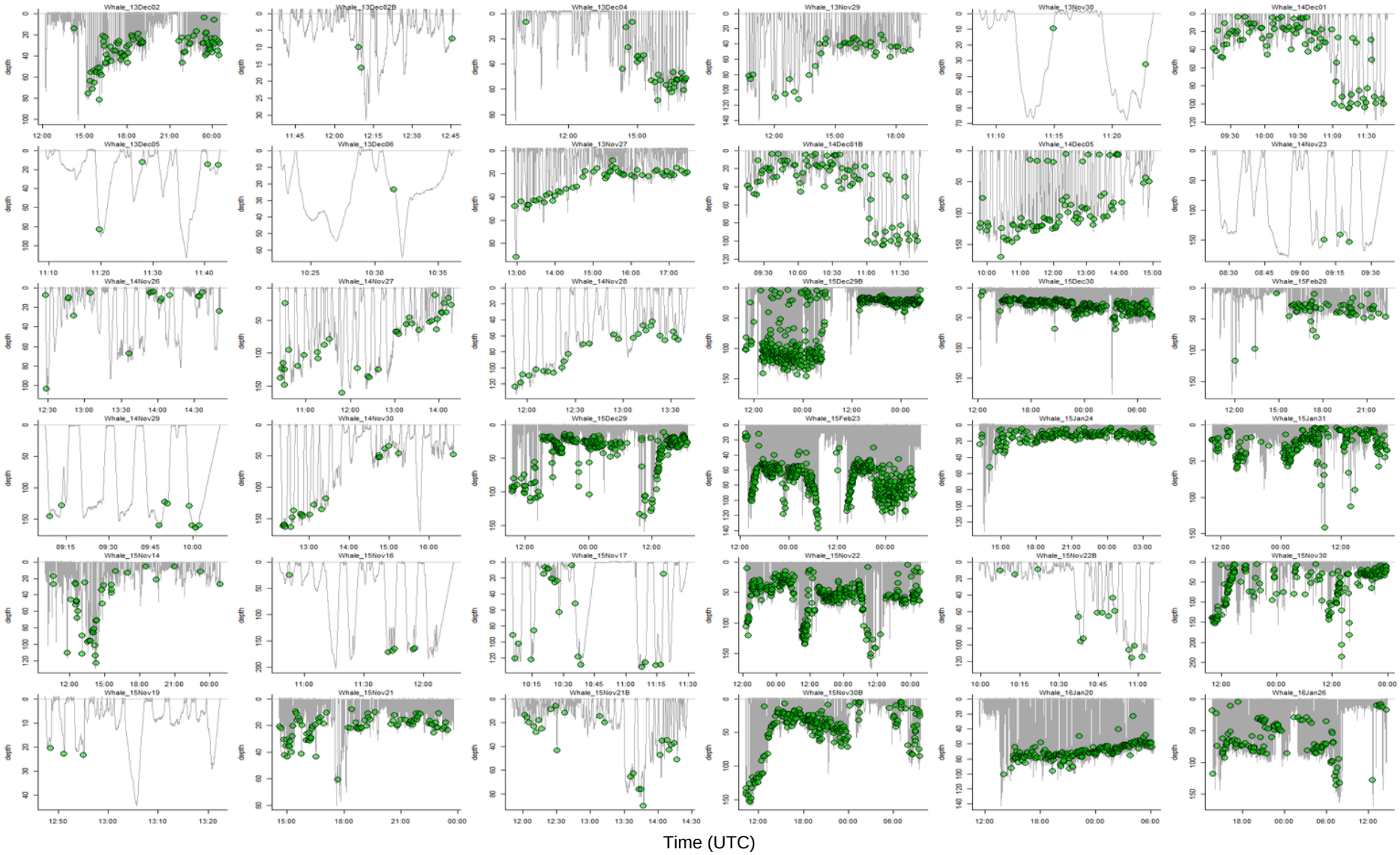


Figure 7 Dive profiles and lunge detections (green circles) of 36 whales instrumented with TDRs.

5 Discussion

The 2D detector performed relatively well despite large individual variations regarding AUCs, ranging from 0.499 to 0.805, a TPR of 68.4% and an FPR of 0.025%. A similar study by Allen et al., (2016), developed a detector based on DTAG data from fin whales that accurately detected 92.8% of the lunges with an FPR of 9.5%, but also had substantial individual variation. As mentioned in the results, our FPR should be interpreted with caution because of the high TN in this study. Therefore, it is difficult to compare the FPR to the value presented by Allen et al (2016). However, the TPRs can be compared between the studies, with the 3D detector performing better than the 2D detector presented here. The results presented here match better with the TPR found by Owen et al. (2016) in their study on surface feeding in humpback whales, which was approximately 70%. The most challenging aspect of this study was to decrease the FP, and even after optimisation 12 of the whales had a higher FP compared to TP (Table S3).

The detector performs better on longer datasets with higher rates of lunge activity, possibly due to the ratio of TP to FP increasing. As mentioned, the FP was still high following optimisation, and datasets with little to no feeding activity had a high proportion of FPs. This suggests that the detector might not operate with a low degree of specificity, and data signals that do not correspond to a lunge could be defined as such. Furthermore, with such short recording periods, the data can potentially be affected by the animal being disturbed for a while after deployment. Thus, results based on short datasets should be treated with some caution. Additionally, the algorithm appears to perform better for animals performing generally deeper dives, which can be linked to less disturbance for the pressure sensor and the resolution of depth data. Individual variability can be linked to deployment duration or be due to physiological and behavioural factors (Allen et al. 2016). There is also a seasonal difference in deployment ranging from November to February, and it is plausible that lunging activity may change from November to February, with whales being hungrier upon arrival and satiated by the end of the feeding season. This is not evident in this study (see Figures 6 and 7), however since no statistical analysis was performed it cannot be ruled out.

To run optimally, the algorithm must be adjusted to the biological restrictions and lunge kinematics of the species it is applied to. Rorqual whale species do not all lunge the same. They vary in ILI, lunge duration and approach, as have been shown by Cade et al. (2016). Therefore, the parameters of the detector must be adjusted to the species it is run on. Additionally, the tag settings for each specific deployment, referring to pre-programming such as the sampling

frequency of the tag need to be constant as having a tag with a 1-s sampling frequency will have different records compared to one with a 5-s frequency for instance.

Automatic identification of important signals in data from animal-borne instruments is a major goal, as the amount of data from movement tags has increased, both in terms of tags deployed (Allen et al., 2016) and the number of species carrying them (Brown et al., 2013). While TDRs gather simpler data compared to multisensory tags such as LL, DTAGs and CATS tags, the datasets recorded are still large, and have been gathered over a much longer time period since the introduction of these tags, but the lower resolution and less multidimensional data makes interpretation challenging (Fedak et al., 2001). The development of detailed multisensory tags allowed for more information to be extracted from older tag data, with varying degree of confidence. Therefore, with re-analyses of a potentially large number of older low-resolution datasets, automatic methods of analysis are required. Having automatic approaches for analysing dive data also allows for a uniform application of selected criteria across all data, where pure human observation might introduce subjective variations in results and possible biases (Allen et al., 2016). Furthermore, the method must be standardised for the automatic detector to be consistent and comparable across studies (Fedak et al., 2001; Womble et al., 2013).

As presented in Materials & Methods, there is no objective way of optimising the detector parameters because the statistical optimisation is presently not feasible due to time constraints. Statistical optimisation is based on running wavelet analyses for every dive, with all potential combinations of the parameters at once. When doing this manually, running one test may take about 30 seconds, so the statistical method could potentially take many months, something that was beyond the scope of this thesis. Choosing to either restrict the number of parameter values used each time in the optimisation or by choosing a random subset of dives to run for each individual are potential methods for reducing the time required for statistical optimisation. Alternatively, a high-performance computer cluster could also reduce analysis time.

Lunges are generally performed during the bottom phase and ascent portions of a dive (Acevedo-Gutiérrez et al., 2002; Goldbogen et al., 2006; Ware et al., 2011; Akiyama et al., 2019), but the 2D detector also detects lunges in the descent portions of the dives (Figures 6 and 7). One way of possibly decreasing the number of FP detections could be to restrict the algorithm to search for lunges during the ascent portions of a dive. This also has a potential drawback of introducing biases as the algorithm is already set to detect specific behaviours and one specific part of the dataset is excluded. Additionally, although depth filtration to remove

lunges shallower than four meters have been done manually in this study, in the future this should be included as a separate parameter.

For future analysis of these types of data, it would be valuable to include tissue samples and information about each of the whales tagged, such as size and sex, to see if biological factors have any effect on the performance of the 2D detector. This information was collected for only a small number of the whales in this study. Size impacts the way an animal moves, with smaller animals moving at higher frequencies, meaning that the parameter settings need to be adjusted to optimise the performance of the detector. If this data was available, it could be used as one explanatory variable in an analysis exploring the causes behind different individuals having different optimal settings.

The kinematic data gathered from deployed instruments could be combined with other sensor data, such as temperature, salinity, GPS location or light levels to place identified behaviours in a broader environmental and biological context (Brown et al., 2013; Allen et al., 2016). Furthermore, video loggers can give insight into how the presence of other animals may affect movement of a target individual. For example, a study by Akiyma et al. (2019) found that humpback whales decreased their feeding time when other animals were present. A goal for this detector is to use it to investigate consumption rates of humpback whales inside a fjord system during a herring superabundance event, by using abundance estimates of the Norwegian spring spawning herring (NSSH). Calculating consumption based on lunge detections has been done previously, but most references found on this focused on accelerometer data (Watanabe & Takahashi, 2013; Carroll et al., 2014; Akiyama et al., 2019). The small number of deployments using accelerometers can be augmented with the larger number of deployments using TDR-type tags to have a larger and more representative sample potentially representing more population-level estimates. This would give insight to the effects prey distribution could have on these animals' behaviour.

Small-scale spatial patterns in prey density, such as fish schools, are likely to change rapidly (Haury et al. 1978; Hazen et al., 2009), and hunting predators like baleen whales depending on high density aggregations for efficient feeding must be able to detect this change (Acevedo-Gutiérrez et al., 2002; Friedlaender et al., 2006; Bannister, 2009; Goldbogen et al., 2011; Goldbogen et al. 2015; Akiyama et al., 2019; De Weerdt & Ramos, 2020). Previous studies show that baleen whales display non-linear threshold aggregative responses towards preferred prey (Brodie, Sameoto & Sheldon, 1978; Piatt & Methven, 1992; Friedlaender et al., 2006; Keen, 2017), implying that they aggregate in areas above a minimum prey density level likely set by some long-term energy intake rate (Marginal Value Theorem, MVT; Charnov, 1976).

MVT predicts that foragers should remain a shorter time in patches with little food. However, given that baleen whales need to come up to the surface to breathe and diving costs energy, the trade-off between prey depth, patch density and profitability has to be considered (Charnov, 1976; Thompson & Fedak, 2001; Doniol-Valcroze et al., 2011; Hazen et al., 2015). A study on seals by Thompson and Fedak (2001), found that dive durations are ultimately constrained by oxygen balance, but also influenced by the seal's assessment of prey patch quality. For baleen whales, the number of feeding events during a dive is assumed to reflect this trade-off between prey patch quality and its vertical position in the water column (Doniol-Valcroze et al., 2011; Goldbogen et al., 2015; Akiyama et al., 2019). Krill-eating humpbacks increase the number of lunges per dive with increasing depth (Ware et al., 2011), and further analyses with the 2D detector can be done to see if a similar, clear, distribution of lunges can be found in fish-eating humpbacks. For example, it would be worthwhile investigating if our data support conclusions by Goldbogen et al. (2015), which found higher feeding rates in humpback whale dives with steeper ascent and descent, as well as extended bottom time. Kinematic data from feeding dives has shown that successful feeding lunges are slower and have shorter ILI compared to unsuccessful ones (Cade et al., 2016). The 2D detector does not currently make this distinction.

Several studies have found a correlation between time of day, or light levels, and feeding activity in marine mammals (Bennett, McConnell & Fedak, 2001; Friedleander et al., 2009; Biuw et al., 2010; Friedleander et al., 2013). A study on southern elephant seals by Bennett et al. (2001) found that seals increased diving duration in pelagic waters, with the longest dives happening in the austral midwinter. Furthermore, another study on the same species of seal found clear DVM taking place in pelagic waters during summer in the Eastern Weddell Sea (Biuw et al., 2010). Friedleander et al. (2013) found that krill-eating humpback whales in the Antarctic Peninsula follow the depth distribution of their prey even with little variation in daily light levels. Furthermore, there is evidence that fish-feeding humpbacks change foraging behaviour in relation to light and prey conditions (Friedleander et al., 2009). It is known that herring perform diel vertical migration (Misund, Melle & Fernö, 1997; Huse & Korneliussen, 2000), and that diurnal pattern in humpback feeding can be found in low light levels, so whether similar daily rhythms can be found in herring-feeding humpbacks in Northern Norway should be investigated. From looking at Figures 6 and 7, humpbacks seem to display some diurnal pattern in feeding activity, but statistical analysis needs be done to see if this is the case.

Biologging allows behavioural data to be gathered with minimal human interaction (Halsey et al., 2011; Carroll et al., 2014; Allen et al., 2016). And the use of accelerometers to measure dynamic movements can be used to estimate energy expenditure (Halsey et al., 2011; Broell et

al., 2013; Cox et al., 2018). Understanding feeding rates and energy expenditure can help with identifying and quantifying human impacts and can be used to gain information needed for the maintenance and recovery of endangered species, such as many rorqual species. (Bannister, 2009; Allen et al., 2016; De Weerdt & Ramos, 2020).

While biologging has allowed for much new knowledge, multisensory tag use has been limited by data storage, battery life and, in the case of cetaceans, attachment longevity (Johnson & Tyack, 2003; Broell et al., 2013; Allen et al., 2016). With the continuous development of this technology, data storage and battery life now allow long-term deployments (Allen et al., 2016). This development has helped advance our understanding of marine mammal behaviour, physiology and ecology, as it decreased periods of unknown activity and increase the detection of trends (Cooke et al., 2004; Allen et al, 2016). However, use still limited to species that can be reliably rediscovered (Allen et al., 2016). Furthermore, the tag placement on the animal's body can impact which behaviours can be identified (Shepard et al., 2008; Brown et al., 2013; Carroll et al., 2014; Allen et al., 2016). Being as consistent as possible with where the instrument is placed on the different animals can help minimise interpretation error (Shepard et al., 2008; Brown et al., 2013; Allen et al., 2016).

Although biologging has allowed for increased information regarding the feeding ecology and foraging behaviour of marine mammals, this has been mainly limited to feeding at depth (Goldbogen et al., 2006; Doniol-Valcroze et al., 2011; Ware et al., 2011; Wiley et al., 2011; Simon et al., 2012; Friedlaender et al., 2013; Goldbogen et al., 2013; Goldbogen et al., 2015; Allen et al., 2016; Owen et al., 2016). This is due to dramatic pressure changes occurring at the air-water interface disturbing the signals in the pressure sensor. Being near the surface, an animal will encounter additional forces associated with the air-water interface and decreased space for manoeuvring, causing the kinematics of the feeding to change (Allen et al., 2016). Corroborating whether a signal found near the surface is feeding or non-feeding is difficult, but the best option is video loggers. At this point in time, several studies have looked into this for 3D tags (Kot et al., 2014; Allen et al., 2016; Owen et al., 2016). However, it is unlikely that similar methods can be developed for TDR data.

In this study, accelerometer data were used to validate the 2D detector and although this is a previously used method of validation (Vivant et al., 2014; Cox et al., 2018), visual observation through the use of video footage is the ideal (Watanabe & Takashi, 2013; Carroll et al., 2014; Cade et al., 2016; Cox et al., 2018; Akiyama et al., 2019). In addition to providing information about foraging behaviour, video loggers can quantify changes in the prey density following a lunge (Akiyama et al., 2019), and identify prey species.

6 Conclusion

The development of digital tags has led to scientists acquiring new knowledge regarding lunging kinematics and foraging behaviour in rorqual whales. Behaviours that were poorly understood due to the difficulty, even impossibility, of observing them was revealed using high-resolution multisensory tags recording 3D acceleration. The introduction of these instruments also opened for the re-analysis of older datasets gathered from simpler tags such as the TDR. As these instruments have been around for longer, and are still in use today, the amount of data available for analysis is huge. Automatic methods of examining these data are sought after as it reduces the time required for examination and increases standardisation of the results. A detector capable of identifying feeding lunges from only time-depth data was developed in this study and performed at a level comparable to some previous research on developing automatic analysis methods for detecting foraging behaviours. The detector is transferrable to other rorqual species as well if it is adjusted to the specific species lunge kinematics. In the future, this detector can help decrease the knowledge gaps related to the foraging behaviour of rorqual whales. This can be done in combination with metadata of the tagged whales, such as size, sex, and tissue samples, as well as environmental data and information about biomass and distribution of prey species. Furthermore, there are steps that can improve upon the detector. This includes creating a more objective method of optimisation, as well as finding ways to further decrease the FP and making it more specific.

Works cited

Acevedo-Gutiérrez A., Croll D. A., & Treshy B. R. (2002). High feeding costs limit dive time in the largest whales. *The Journal of Experimental Biology, 205,* 1747-1753.

Adachi, T., Maresh, J. L., Robinson, P. W., Peterson, S. H., Costa, D. P., Naito Y., Watanabe Y. Y., & Takasashi A. (2014). The foraging benefits of being fat in a high migratory marine mammal. *Proceedings of the Royal Society B: Biological Sciences, 281*(1797), 1-9. doi: 10.1098/rspb.2014.2120

Akiyama, Y., Akamatsu, T., Rasmussen, M. H., Iversen, M. R., Iwata, T., Goto, Y., Aoki, K., & Sato, K. (2019). Leave or stay? Video-logger revealed foraging efficiency of humpback whales under temporal change in prey density. *PLOS ONE, 14*(2). doi: 10.1371/journal.pone.0211138

Allen, A. N., Goldbogen, J. A., Friedlaender A. S., & Calambokidis J. (2016). Development of an automated method of detecting stereotyped feeding events in multisensor data from tagged rorqual whales. *Ecology and Evolution, 6*(20), 7522-7535. doi: 10.1002/ece3.2386

Aoki, K., Amano, M., Mori, K., Kourogi A., Kubodera, T., & Miyazaki, N. (2012). Active hunting by deep-diving sperm whales: 3D dive profiles and maneuvers during bursts of speed. *Marine Ecology Progress Series, 444,* 289-301. doi: 10.3354/meps09371

Bannister, J. (2009). Baleen whales (Mysticetes). *Encyclopedia of Marine Mammals,* 80-89. doi: 10.1016/B978-0-12-373553-9.00024-9

Bennett, K., McConnell, B. J., & Fedak, M. A. (2001). Diurnal and seasonal variations in the duration and depth of the longest dives in southern elephant seals (*Mirounga leonina*): Possible physiological and behavioural constraints. *The Journal of Experimental Biology, 24,* 649-662.

Biuw, M., Nøst, O. A., Stien, A., Zhou Q., Lydersen, C., & Kovacs, K. M. (2010). Effects of hydrographic variability on the spatial, seasonal and diel diving patterns of southern elephant seals in the eastern Weddell Sea. *PLoS ONE, 5*(11). doi: 10.1371/journal.pone.0013816

Broell, F., Noda, T., Wright, S., Domenici, P., Steffensen, J. F., Auclair, J. P., & Taggart, C. T. (2013). Accelerometer tags: detecting and identifying activities in fish and the effect of sampling frequency. *Journal of Experimental Biology, 216*(7), 1255-1264. doi: 10.1242/jeb.077396

Brodie, P. F., Sameoto, D. D., & Sheldon, R. W. (1978). Population densities of euphausiids off Nova Scotia as indicated by net samples, whale stomach contents, and sonar. *Limnology and Oceanography, 23*(6), 1264-1267. doi: 10.4319/lo.1978.23.6.1264

Brodie, P. F. (1993). Noise generated by the jaw actions of feeding fin whales. *Canadian Journal of Zoology, 71*(2), 2546-2550. doi: 10.1139/z93-348

Brown, D. D., Kays, R., Wikelski, M., Wilson, R., & Klimley, A. P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry, 1*(1). doi: 10.1186/2050-3385-1-20

Cade, D. E., Friedlaender, A. S., Calambokidis, J., & Goldbogen. J. A. (2016). Kinematic Diversity in Rorqual Whale Feeding Mechanisms. *Current Biology, 26,* 2617-2624. Retrieved from http://dx.doi.org/10.1016/j.cub.2016.07.037

Cade, D. E., Carey, N., Domenici, P., Potvin, J., & Goldbogen, J. A. (2020). Predator-informed looming stimulus experiments reveal how large filter feeding whales capture highly maneuverable forage fish. *PNAS, 117*(1), 472-478. Retrieved from: https://www.pnas.org/cgi/doi/10.1073/pnas.1911099116

Canning, C., Crain, D., Eaton, T. S. Jr., Nuessly, K., A., Friedlaender, A. S., Hurst, T., Parks, S., Ware, C., Wiley, D., & Weinrich, M. (2011). Population-level lateralized feeding behaviour in North Atlantic humpback whales, *Megaptera novaeangliae. Animal Behaviour, 82,* 901-909. doi: 10.1016/j.anbehav.2011.07.031

Carroll, G., Slip, D., Jonsen I., & Harcourt, R. (2014). Supervised accelerometery analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology, 217*(24), 4295-4302. doi: 10.1242/jeb.113076

Charnov, E. L. (1976). Optimal Foraging, the Marginal Value Theorem. *Theoretical Population Biology, 9*(2), 129-134.

Clapham, P. J., & Mead, J. (1999). Novaengliae megaptera. *American Society of Mammalogists, 604,* 1-9. Retrieved from www.jstor.org/3504352Accessed:18-01-201609:44UTC

Clapham, P. J. (2018). Humpback whale. *Encyclopedia of Marine Mammals (3rd Edt),* 489-492. Elsevier, Academic Press

Cooke, S. J., Hinch, S. G., Wikelski, M., Andrews, R. D., Kuchel, L. J., Wolcott, T. G., & Butler, P. J. (2004). Biotelemetry: A mechanistic approach to ecology. *Trends in Ecology and Evolution, 19*(6), 334-343. doi: 10.1016/j.tree.2004.04.003

Cox, S. L., Orgeret, F., Gesta, M., Rodde, C., Heizer, I., Weimerskirch, H., & Guinet, C. (2018). Processing of acceleration and dive data on-board satellite relay tags to investigate

diving and foraging behaviour in free-ranging marine predators. *Methods in Ecology and Evolution, 9*(1), 64-77. doi: 10.1111/2041-210X.12845

DeRuiter, S. (2018). tagtools: Tolls for Working with Data from High-Resolution Biologging Tags. Rpackage version 0.0.0.9000

De Weerdt, J., & Ramos, E. (2020). Feeding of humpback whales (*Megaptera novaeangliae*) on the Pacific coast of Nicaragua. *Marine Mammal Science, 36*(1), 285-292. doi: 10.1111/mms.12613

Doniol-Valcroze, T., Lesage, V. Giard, J., & Michaud, R. (2011). Optimal foraging theory predicts diving and feeding strategies of the largest marine predator. *Behavioural Ecology, 22*(4), 880-888. doi: 10.1093/beheco/arr038

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27,* 861-874. doi: 10.1016/j.patrec.2005.10.010

Fedak, M. A., Lovell, P., & Grant, S. M. (2001). Two approaches to compressing and interpreting time-depth information as collected by time-depth recorders and satellite-linked data recorders. *Marine Mammal Science, 17*(1), 94-110. doi: 10.1111/j.1748-7692.2001.tb00982.x

Friedlaender, A. S., Halpin, P. N., Qian, S. S., Lawson, G. L., Wiebe, P. H., Thiele, D., & Read A. J. (2006). Whale distribution in relation to prey abundance and oceanographic processes in shelf water of the Western Antarctic Peninsula. *Marine Ecology Progress Series, 317,* 297-310. doi: 10.3354/meps317297

Friedlaender, A. S., Hazen E. L., Nowacek, D. P., Halpin, P. N., Ware, C., Weinrich, M. T., Hurst, T., & Wiley, D. (2009). Diel changes in humpback whale *Megaptera novaeangliae* feeding behavior in response to sand lance *Ammodytes* spp. behavior and distribution. *Marine Ecology Progress Series, 395,* 91-100. doi: 10.3354/meps08003

Friedlaender, A. S, Tyson, R. B., Stimpert, A. K., Read, A. J., & Nowacek D. P. (2013). Extreme diel variation in the feeding behavior of humpback whales along the western Antarctic Peninsula during autumn. *Marine Ecology Progress Series, 494,* 281-289. doi: 10.3354/meps10541

Goldbogen, J. A., Calambokidis, J., Shadwick, R. E., Oleson, R. M., McDonald, M. A., & Hildebrand, J. A. (2006). Kinematics of foraging dives and lunge feeding in fin whales. *Journal of Experimental Biology, 209*(7), 1231-1244. doi: 10.1242/jeb.02135

Goldbogen, J. A., Pyenson, N. D., & Shadwick, R. E. (2007). Big gulps require high drag for fin whale lunge feeding. *Marine Ecology Progress Series, 349,* 289-301. doi: 10.3354/meps07066

Goldbogen, J. A., Calambokidis J., Croll, D. A., Harvey, J. T., Newton, K. M., Oleson E. M., Schorr G., & Shadwick, R. E. (2008). Foraging behavior of humpback whales: kinematic and respiratory patterns suggest a high cost for a lunge. *Journal of Experimental Biology, 211*(23), 3712-3719. doi: 10.1242/jeb.023366

Goldbogen, J. A. (2010). The Ultimate Mouthful: Lunge Feeding in Rorqual Whales: The ocean's depths have long shrouded the biomechanics behind the largest marine mammals' eating methods, but new devices have brought them to light. *American Scientist, 98*(2), 124-131. Retrieved from https://www.jstor.org/stable/27859477

Goldbogen, J. A., Potvin, J., & Shadwick R. E. (2010). Skull and buccal cavity allometry increase mass-specific engulfment capacity in fin whales. *Proceedings of the Royal Society B: Biological Sciences, 277*(1683), 861-868. doi: 10.1098/rspb.2009.1680

Goldbogen, J. A., Calambokidis, J., Oleson, E., Potvin, J., Pyenson, N. D., Schorr, G., & Shadwick, R. E. (2011). Mechanics, hydrodynamics and energetics of blue whale lunge feeding: efficiency dependence on krill density. *The Journal of Experimental Biology, 214*, 131-146. doi: 10.1242/jeb.048157

Goldbogen, J. A., Friedlaender, A. S., Calambokidis, J., McKenna, M. F., Simon, M., & Nowacek, D. P. (2013). Integrative Approaches to the Study of Baleen Whale Diving Behavior, Feeding Performance, and Foraging Ecology. *BioScience, 63*(2), 90-100. doi: 10.1525/bio.2013.63.2.5

Goldbogen, J. A., Hazen, E. L., Friedlaender, A. S., Calambokidis, J., DeRuiter, S. L., Stimpert, A. K., & Southall, B. L. (2015). Prey density and distribution drive the three-dimensional foraging strategies of the largest filter feeder. *Functional Ecology, 29* (7), 951-961. doi: 10.1111/1365-2435.12395

Goldbogen, J. A., Cade, D. E., Calambokidis, J., Friedlaender, A. S., Potvin, J., Segre, P. S., & Werth, A. J. (2017). How Baleen Whales Feed: The Biomechanics of Engulfment and Filtration. *Annual Review of Marine Science, 9,* 367-386. doi: 10.1146/annurev-marine-122414-033905

Goldbogen, J. A., & Madsen, P. T. (2018). The evolution of foraging capacity and gigantism in cetaceans. *Journal of Experimental Biology, 221*. doi: 10.1242/jeb.166033

Halsey, L. G., Shepard, E. L. C., & Wilson, R. P. (2011). Assessing the development and application of the accelerometry technique for estimating energy expenditure. *Comparative Biochemsitry and Physiology – A Molecular and Integrative Physiology, 158*, 305-314. doi: 10.1016/j.cbpa.2010.09.002

Haury, L. R., McGowan, J. A., & Wiebe, P. H. (1978). Patterns and processes in the time-space scales of plankton distribution. In: J. H. Steele (Eds.), *Spatial Pattern in Plankton Communities. NATO Conference Series, IV Marine Sciences,* 277-327. doi: 10.1007/978-1-4899-2195-6_12

Hazen, E. L., Friedlaender, A. S., Thompson, M. A., Ware, C. R., Weinrich, M. T., Halpin, P. N., & Wiley, D. N. (2009). Fine-scale prey aggregations and foraging ecology of humpback whales *Megaptera novaeangliae. Marine Ecology Progress Series, 395,* 75-89. doi: 10.3354/meps08108

Hazen, E. L., Friedlaender A. S., & Goldbogen J. A. (2015). Blue whales (Balaenoptera musculus) optimize foraging efficiency by balancing oxygen use and energy gain as a function of prey density. *Science Advances, 1*(9), 1-7. doi: 10.1126/sciadv.1500469

Huse, I., & Korneliussen, R. (2000). Diel variation in acoustic density measurements of overwintering herring (Clupea harengus L.). *ICES Journal of Marine Science, 57,* 903-910. doi: 10.1006/jmsc.2000.0577

Johnson, M. P., & Tyack, P. L. (2003). A digital acoustic recording for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering, 28*(1), 3-12. doi: 10.1109/JOE.2002.808212

Johnson, M., de Soto, N. A., & Madsen, P. T. (2009). Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: A review. *Marine Ecology Progress Series, 395,* 55-73. doi: 10.3354/meps08255

Keen, E. M. (2017). Aggregative and feeding thresholds of sympatric rorqual whales within a fjord system. *Ecosphere, 8*(3). doi: 10.1002/ecs2.1702

Kot, B. W. Sears, R., Zbinden, D., Borda, E., & Gordon, M. S. (2014). Rorqual whale (Balaenopteridae) surface lunge-feeding behaviors: Standardized classification, repertoire diversity, and evolutionary analyses. *Marine Mammal Science, 30*(4), 1335-1357. doi: 10.1111/mms.12115

Miller, P. J. O., Johnson, M. P., Tyack, P. L., & Terray, E. A. (2004). Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus. Journal of Experimental Biology, 207*(11), 1953-1967. doi: 10.1242/jeb.00993

Miller, P. J. O., Biuw, M., Watanabe, Y. Y., Thompson, D., & Fedak, M. A. (2012). Sink fast and swim harder! Round-trip cost-of-transport for buoyant divers. *The Journal of Experimental Biology, 215*, 3622-3630. doi: 10.1242/jeb.070128

Misund, O. A, Melle, W., & Fernö, A. (1997). Migration behaviour of Norwegian spring spawning herring when entering the cold front in the Norwegian Sea. *Sarsia North Atlantic Marine Science, 82,* 107-112. doi: 10.1080/00364827.1997.10413644

Owen, K., Dunlop, R. A., Monty, J. P., Chung, D., Noad, M. J., Donnelly, D., Goldizen, A. W., & MacKenzie, T. (2016). Detecting surface-feeding behavior by rorqual whales in accelerometer data. *Marine Mammal Science, 32*(1), 327-348. doi: 10.1111/mms.12271

Payne, N. L., Taylor, M. D., Watanabe, Y. Y., & Semmens, J. M. (2014). From physiology to physics: are we recognizing the flexibility of biologging tools? *Journal of Experimental Biology, 217*(3), 317-322. doi: 10.1242/jeb.093922

Piatt, J. F., & Methven, D. A. (1992). Threshold foraging behavior of baleen whales. *Marine Ecology Progress Series, 84,* 205-210.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienne, Austria. Retrieved from https://www.R-project.org/

Roesch, A., & Schmidbauer H. (2018). WaveletComp: Computational Wavelet Analysis. R package version 1.1. Retrieved from https://CRAN.R-project.org/package=WaveletComp

Schrynemackers, M., Küffer, R., & Geurts, P. (2013). On protocols and measures for the validation of supervised methods for the inference of biological networks. *Frontiers in Genetics, 4,* 1-16. doi: 10.3389/fgene.2013.00262

Shepard, E. L. C., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., Myers, A. E., Newman, C., & Macdonald, D. W. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research, 10,* 47-60. doi: 10.3354/esr00084

Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCR: visualizing classifier performance in R. Bioinformatics:, *21*(20), 7881. Retrieved from http://rocr.bioinf.mpi-sb.mpg.de

Simon, M., Johnson, M., Tyack P., & Madsen P. T. (2009). Behaviour and kinematics of continuous ram filtration in bowhead whales (*Balaena mysticetes*). *Proceedings of the Royal Society B: Biological Sciences, 276*(1674), 3819-3828. doi: 10.1098/rspb.2009.1135

Simon, M., Johnson, M., & Madsen P. T. (2012). Keeping momentum with a mouthful of water: behavior and kinematics of humpback whale lunge feeding. *Journal of Experimental Biology, 215*(21), 3786-3798. doi: 10.1242/jeb.071092

Thompson, D., & Fedak, M. A (2001). How long should a dive last? A simple model of foraging decisions by breath-hold divers in a patchy environment. *Animal Behaviour, 61,* 287-296. doi: 10.1006/anbe.2000.1539

Vivant, M., Monestiez, P. & Guinet, C. (2014). Can We Predict Foraging Success in a Marine Predator from Dive Patterns Only? Validation with Prey Capture Attempt Data. *PLOS ONE, 9*(3). doi: 10.1371/journal.pone.0088503

Ware. C., Friedlaender, A. S., & Nowacek, D. P. (2011). Shallow and deep lunge feeding of humpback whales in fjords of the West Antarctic Peninsula. *Marine Mammal Science, 27*(3), 387-605. doi: 10.1111/j.1748-7692.2010.00427.x

Watanabe, S., Izawa, M., Kato, A., Ropbert-Coudert, Y., & Naito, Y. (2005). A new technique for monitoring the detailed behaviour of terrestrial animas: A case study with the domestic cat. *Applied Animals Behaviour Science, 94*(1-2), 117-131. doi: 10.1016/j.applanim.2005.01.010

Watanabe, Y. Y., & Takahashi, A. (2013). Linking animal-borne video to accelerometers reveals prey capture variability. *Proceedings of the National Academy of Sciences, 110*(6), 2199-2204. doi: 10.1073/pnas.1216244110

Wiley, D. N., Ware, C., Bocconcelli, A., Cholewiak, D. M., Friedlaender, A. S., Thompson, M., & Weinrich, M. (2011). Underwater components of humpback whale bubble-net feeding behaviour. *Behaviour, 148*(5-6), 575-602. doi: 10.1163/000579511X570893

Womble, J. N., Horning, M., Lea, M. A., & Rehberg, M. J. (2013). Diving into the analysis of time-depth recorder and behavioural data records: A workshop summary. *Deep-Sea Research II, 88-89,* 61-64. doi: 10.1016/j.dsr2.2012.07.017

Woodward, B. L., Winn, J. P., & Fish, F. E. (2006). Morphological specializations of baleen whales associated with hydrodynamic performance and ecological niche. *Journal of Morphology, 267*(11), 1284-1294. doi: 10.1002/jmor.10474

Appendix

Table S1 The time of attachment, duration of tag attachment and maximum depth of dives for each whale tagged with accelerometers. Total duration of deployment was 123hr, 32min, 50sec.

Whales	Tag time (UTC)	Duration	Max depth (m)
Mn13_340a	2013-12-06, 09:55:01	3hr 17min 13sec	91.07
Mn13_341a	2013-12-07, 09:56:02	2hr 56min 9sec	67.25
Mn14_325a	2014-11-20, 09:54:06	9hr 44min 36sec	164.6
Mn14_334a	2014-11-29, 12:52:54	10hr 34min 26sec	127.8
Mn14_335a	2014-12-01, 10:09:48	9hr 9min 34sec	251.1
Mn14_350a	2014-12-16, 21:41:00	3hr 33min 57sec	89.21
Mn15_339a	2015-12-05, 10:27:18	8hr 24min 26sec	164.4
Mn15_341a	2015-12-07, 11:14:30	2hr 20min 14sec	166.4
Mn16_Jan19a	2016-01-19, 10:24:55	3hr 18min 9sec	196.8
Mn16_Jan25a	2016-01-25, 09:09:56	7hr 11min 23min	105.3
Mn16_Jan25b	2016-01-25, 11:04:21	14hr 55min 14sec	151.5
Mn16_Jan26	2016-01-26, 10:02:51	11hr 7min 3sec	161.3
Mn17_022LLa	2017-01-22, 09:51:10	7hr 45min	141.8
Mn17_022LLb	2017-01-22, 10:44:43	9hr 5min 23sec	110.0
Mn17_026LLa	2017-01-26, 08:42:32	14hr 32min 29sec	142.3
Mn18_013LLa	2018-01-13, 09:22:38	5hr 37min 34sec	283.3

Table S2 The times of attachment, the duration of tag attachment and maximum depth of dives for each whale tagged with TDR. Total duration of deployment is 448hr, 44min, 33sec.

WhaleID	Tag time (UTC)	Duration	Max depth (m)
Whale_2013Dec02	2013-12-01, 12:12:46	12hr 22min 9sec	101
Whale_2013Dec02B	2013-12-03, 11:38:35	1hr 7min 19sec	31.5
Whale_2013Dec04	2013-12-05, 09:34:16	7hr 42min 41sec	84.5
Whale_2013Dec05	2013-12-06, 11:09:18	33min 19sec	112.5
Whale_2013Dec06	2013-12-07, 10:22:29	13min 44sec	64.5
Whale_2013Nov27	2013-11-28, 12:53:17	4hr 34min 44sec	92
Whale_2013Nov29	2013-11-30, 10:34:42	8hr 38min 48sec	140
Whale_2013Nov30	2013-12-01, 11:08:31	15min 4sec	68
Whale_2014Dec01	2014-12-02, 09:13:57	2hr 33min 50sec	118
Whale_2014Dec01B	2014-12-02, 09:13:56	2hr 33min 51sec	118
Whale_2014Dec05	2014-12-06, 09:44:53	5hr 18min 3sec	170.5
Whale_2014Nov23	2014-11-24, 08:22:45	1hr 14min	179
Whale_2014Nov26	2014-11-27, 12:27:51	2hr 43min 15sec	109
Whale_2014Nov27	2014-11-28, 10:23:11	3hr 57min 17sec	163
Whale_2014Nov28	2014-11-29, 11:50:36	2hr 49min 49sec	133.5
Whale_2014Nov29	2014-11-30, 09:07:31	1hr 2min 11sec	169.5
Whale_2014Nov30	2014-12-01, 12:13:56	4hr 13min 41sec	170.5
Whale_2015Dec29	2015-12-30, 09:34:43	35hr 5min 59sec	159
Whale_2015Dec29B	2015-12-30, 10:04:03	42hr 41min 22sec	176.5
Whale_2015Dec30	2015-12-31, 12:08:10	19hr 42min 5sec	177.5
Whale_2015Feb20	2015-02-21, 10:27:27	11hr 56min 56sec	171
Whale_2015Feb23	2015-02-23, 13:23:40	43hr 13min 47sec	141.5
Whale_2015Jan24	2015-01-25, 13:09:44	14hr 46min 33sec	131.5
Whale_2015Jan31	2015-02-01, 10:16:28	34hr 52min 13sec	147
Whale_2015Nov14	2015-11-15, 09:56:09	14hr 56min 49sec	129.5
Whale_2015Nov16	2015-11-17, 10:47:04	1hr 28min 19sec	202.5
Whale_2015Nov17	2015-11-18, 10:05:28	1hr 24min 10sec	133.5
Whale_2015Nov19	2015-11-20, 12:47:21	35min 8 sec	44.5
Whale_2015Nov21	2015-11-20, 14:33:38	9hr 13min 29sec	80.5
Whale_2015Nov21B	2015-11-22, 11:50:25	2hr 35min 54sec	90
Whale_2015Nov22	2015-11-23, 12:55:48	63hr 42min 34sec	174.5
Whale_2015Nov22B	2015-11-23, 09:59:24	1hr 6min 31sec	128.5
Whale_2015Nov30	2015-12-01, 10:02:26	37hr 51min 25sec	265.5
Whale_2015Nov30B	2015-12-01, 10:17:40	23hr 22min 40sec	159
Whale_2016Jan20	2016-01-21, 11:21:49	6hr 58min 13sec	142.5
Whale_2016Jan26	2016-01-26, 13:43:18	23hr 59min 56sec	168.5

Table S3 Confusion matrices produced for all 16 accelerometer instrumented whales, for both the Default and the Standardised 2 settings.

			Default		Standardised 2		
			Predicted				
			True	False		True	False
Mn13_340a	**Observed**	True	8	1	True	8	1
		False	38	13259	False	38	13259
Mn13_341a	**Observed**	True	7	1	True	7	1
		False	26	10574	False	19	10581
Mn14_325a	**Observed**	True	24	8	True	19	13
		False	174	33181	False	184	33171
Mn14_334a	**Observed**	True	69	49	True	56	62
		False	364	51453	False	247	51570
Mn14_335a	**Observed**	True	62	33	True	61	34
		False	191	32689	False	175	32705
Mn14_350a	**Observed**	True	24	4	True	22	6
		False	56	12940	False	43	12953
Mn15_339a	**Observed**	True	58	35	True	42	30
		False	187	29987	False	182	30013
Mn15_341a	**Observed**	True	17	4	True	15	6
		False	65	8329	False	48	8346
Mn16_Jan19a	**Observed**	True	1	10	True	2	6
		False	20	12441	False	18	12443
Mn16_Jan25a	**Observed**	True	6	8	True	6	8
		False	71	25813	False	26	25858
Mn16_Jan25b	**Observed**	True	254	35	True	232	57
		False	79	53347	False	51	53375
Mn16_Jan26	**Observed**	True	35	120	True	120	35
		False	89	39780	False	71	39798
Mn17_022LLa	**Observed**	True	4	11	True	8	7
		False	14	27872	False	5	27881
Mn17_022LLb	**Observed**	True	7	17	True	10	14
		False	26	32674	False	9	32691
Mn17_026LLa	**Observed**	True	5	0	True	5	0
		False	25	52326	False	9	52342
Mn18_013LLa	**Observed**	True	4	4	True	4	4
		False	20	20231	False	14	20237

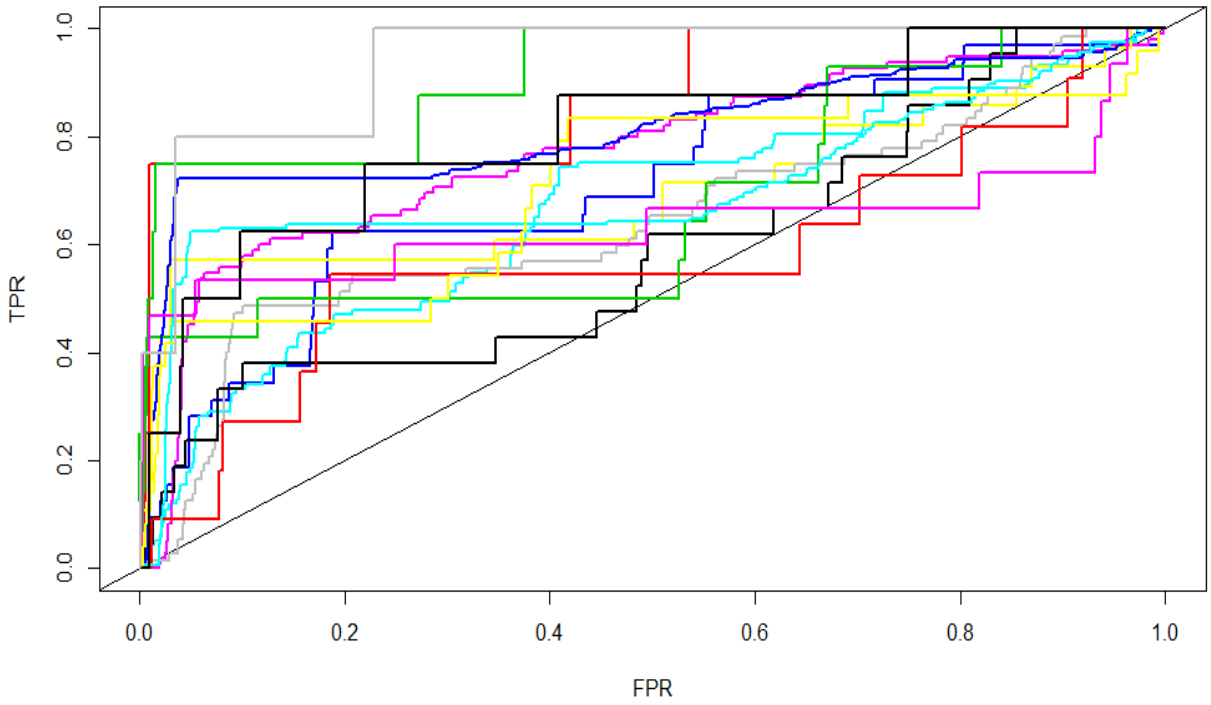


Figure S1 A composite ROC curve showing the results for all 16 whales simultaneously, with the ROC based on the individual optimised settings. Each line corresponds to an individual whale, and the diagonal line across the graph equals an AUC of 0.5.