Temporal patterns in Pacific white-sided dolphin pulsed calls at Barkley Canyon, with implications for multiple populations

by

Kristen Samantha Jasper Kanes

Bachelor of Science (Co-operative Education), University of Victoria, 2013

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the School of Earth and Ocean Sciences

Supervisory Committee

Supervisory Committee

Dr. Stan Dosso (School of Earth and Ocean Sciences)
Co-Supervisor

Dr. Svein Vagle (School of Earth and Ocean Sciences)
Co-Supervisor

Dr. Lucinda Leonard (School of Earth and Ocean Sciences)
Departmental Member

Dr. Tania Lado Insua (Ocean Networks Canada, School of Earth and Ocean Sciences)
Additional Member

Abstract

Evaluation of diel and seasonal patterns in offshore marine mammal activity through visual data collection can be impaired by poor weather and light limitations and by the requirement for costly ship time. As a result, relatively little is known about the diel patterns of wild dolphins. Pacific white-sided dolphins north of Southern California are particularly under-researched. Collecting acoustic data can be a cost-effective approach to evaluating activity patterns in offshore marine mammals. However, manual analysis of acoustic data is time-consuming, and impractical for large data sets. This study evaluates diel and seasonal patterns in Pacific white-sided dolphin communication through automated analysis of one year of continuous acoustic data collected from the Barkley Canyon node of Ocean Networks Canada's NEPTUNE observatory, offshore Vancouver Island, British Columbia, Canada. In this study, marine mammal acoustic signals are manually annotated in a sub-set of the data, and used to train a random forest classifier targeting Pacific white-sided dolphin pulsed calls. Marine mammal vocalizations are classified using the resultant classifier, manually verified, and examined for seasonal and diel patterns. Pacific white-sided dolphins are shown to be vocally active during all diel periods in the spring and summer, but primarily at dusk and night in the fall and winter. Additionally, the percentage of time they are detected drops significantly in the fall and remains low during the winter. This pattern suggests that a group of day-active dolphins, possibly a unique population, leaves Barkley Canyon in the fall and returns in the spring. It is hypothesized that this group may be following the Pacific herring, which are present at the surface during the day at Barkley Canyon in the spring and summer, and migrate inshore for the fall and winter.

Table of Contents

List of Tables

List of Figures

Acknowledgments

This work would not have been possible without the support of many who have contributed to my work and to my development as a scientist. In particular, I would like to thank:

My committee Stan Dosso, Tania Lado Insua, Svein Vagle, and Lucinda Leonard for their guidance, patience, advice, and support. This project took many unexpected turns, and their creative brainstorming, patient teaching, and unyielding faith in my ability to complete this work regardless of the direction it took have been paramount to its success.

Ocean Networks Canada for providing the data for this project as well as other support through funding applications, travel funding, and staff expertise.

JASCO Applied Sciences for allowing me use of their acoustic analysis software.

Xavier Mouy for guiding me through using JASCO's software and teaching me about the mechanics, training, and assessment of random forest classifiers.

George Tzanetakis for his advice about classifier training and validation, which was critical to the success of my classifier.

Tom Dakin for taking a chance on me by hiring a biologist to do acoustics, for teaching me and providing countless opportunities to learn and grow as an acoustic scientist, and for his unwavering faith in my ability take on new disciplines.

My parents for their upbringing and continued support.

My community who saw me through every challenge with grace.

Funding for this work was provided by the Natural Sciences and Engineering Research Council of Canada – Collaborative Research and Development Grant project CRDPJ 500069-16 (PI: Rosaline Canessa; industry partners: JASCO Applied Sciences, IBM).

Chapter 1 Introduction

1.1 Overview

Cetacean behaviour has been a topic of concentrated scientific interest since the aquarium captures of the 1960s and 1970s. Studying cetacean behaviour can provide insight into their habitat use, ecology, and social structures (e.g., Ford, 1991; Hanson & Defran, 1993; Geise *et al.*., 1999), and improve our ability to monitor and protect at-risk cetacean species. While much behavioural research has been conducted visually, such research is hindered by lack of daylight, poor weather, the limited time cetaceans spend at the surface, and the prohibitive costs of vessel time. Passive acoustic monitoring is becoming a popular alternative to visual surveys due to its relatively low cost and independence of light and weather conditions, and the fact that it allows some questions that were previously unanswerable to be investigated.

Some cetaceans are so challenging to study visually for various reasons relating to habitat, behaviour, or distribution that we know very little about them. One example is beaked whales, which live offshore, can take dives of up to an hour or more, and are difficult to spot when they surface. Their behaviour can still be studied acoustically, with recent research showing behavioural changes in relation to mid-frequency active sonar and vessel noise (Tyack *et al.*, 2011; Pirotta *et al.*, 2012), nocturnal foraging behaviour at seamounts (Johnston *et al.*, 2008; McDonald *et al.*, 2009), and identification of previously unknown beaked whale habitat (Yack *et al.*., 2013). There has also been a concentrated effort in recent years to develop techniques for estimating density and population sizes of difficult-to-study cetaceans using both single- and multi-point hydrophone installations (e.g., Marques *et al.*, 2009; Moretti *et al.*., 2010; Küsel *et al.*., 2011; Marques *et al.*, 2011; Marques *et al.*., 2013).

Short- and long-term temporal trends in the behaviour of offshore and other difficult-to-study species that are impractical to investigate visually are also being investigated acoustically. Long-term hydrophone deployments are revealing seasonal patterns in the presence, behaviour, and habitat use of various cetacean species (e.g., Burtenshaw *et al.*, 2004; Verfuß *et al.*., 2007; Munger *et al.*, 2008; Klinck *et al.*., 2012; Dede *et al.*, 2014). Diel patterns, which necessitate night-time observation to evaluate effectively, have also become a topic of interest in cetacean research (e.g., Wiggins *et al.*., 2005; Soldevilla *et al.*., 2010; Baumann-Pickering *et al.*., 2015).

However, passive acoustic research presents its own challenges. Without real-time monitoring of the study site, researchers cannot know when species of interest are present and therefore often collect some quantity of data that does not contain signals of interest. Long-term acoustic deployments can produce very large data sets, and manual methods for finding signals of interest can be time-consuming and impractical. As a result, automated classification of cetacean signals is becoming more common in passive acoustic research (e.g., Deecke *et al.*., 1999; Hannay *et al.*, 2013; Binder & Hines, 2014). However, very few out-of-the-box classifiers exist for cetacean signals, and those that do exist are specific to particular species and regions. Most researchers must choose and train a classifier themselves to distinguish between the signals of interest to them and other signals within the noise conditions of their specific data set, and there are many machine learning classification algorithms with various strengths and weaknesses to choose from.

Data storage space is another issue, since high-frequency acoustic data collection produces very large files compared to more traditional forms of cetacean data collection (photographs, spreadsheets, etc). Some researchers collect data on a duty cycle to reduce the amount of data stored and extend the time period over which data can be collected and stored (e.g., Cerchio *et al.*, 2010; Širović *et al.*, 2013; Williams *et al.*, 2013). While this can be helpful for assessing long-term trends in marine mammal activity, duty cycling may result in some acoustic events being left out of the data set and reduces the ability to monitor short-term changes or overall habitat use (Riera *et al.*, 2013). Other researchers collect continuous data to ensure that nothing is missed, but storing multi-year continuous data sets is expensive and often impractical. Several countries, including Canada, the United States of America, Norway, Japan, Ireland and others, have invested in installing cabled underwater observatories for long-term oceanographic data collection and storing these data on shore. Hydrophones deployed on these observatories can be a good solution for acoustic research focussing on species within the range of the sensors, and some of these observatories make their data readily available to researchers, including Ocean Networks Canada (ONC). ONC operates a network of cabled underwater observatories collecting various types of continuous oceanographic data, including acoustic data. These data are freely provided to researchers and the public. The hydrophone ONC has deployed at Barkley Canyon off the west coast of Vancouver Island frequently records marine mammal vocalizations (unpubl. data), and is a good candidate for use in acoustics-based cetacean research.

Pacific white-sided dolphin (*Lagenorhynchus obliquidens*) vocalizations are particularly common in the acoustic data that ONC has collected from Barkley Canyon (unpubl. data). These pelagic dolphins produce both pulsed calls, which are used for communication, and echolocation clicks, which are used for foraging and navigation in their low visibility marine environment (Janik, 2009; Figure 1.1). Unlike many tropical dolphins, Pacific white-sided dolphins rarely produce whistles (Oswald *et al.*, 2008).

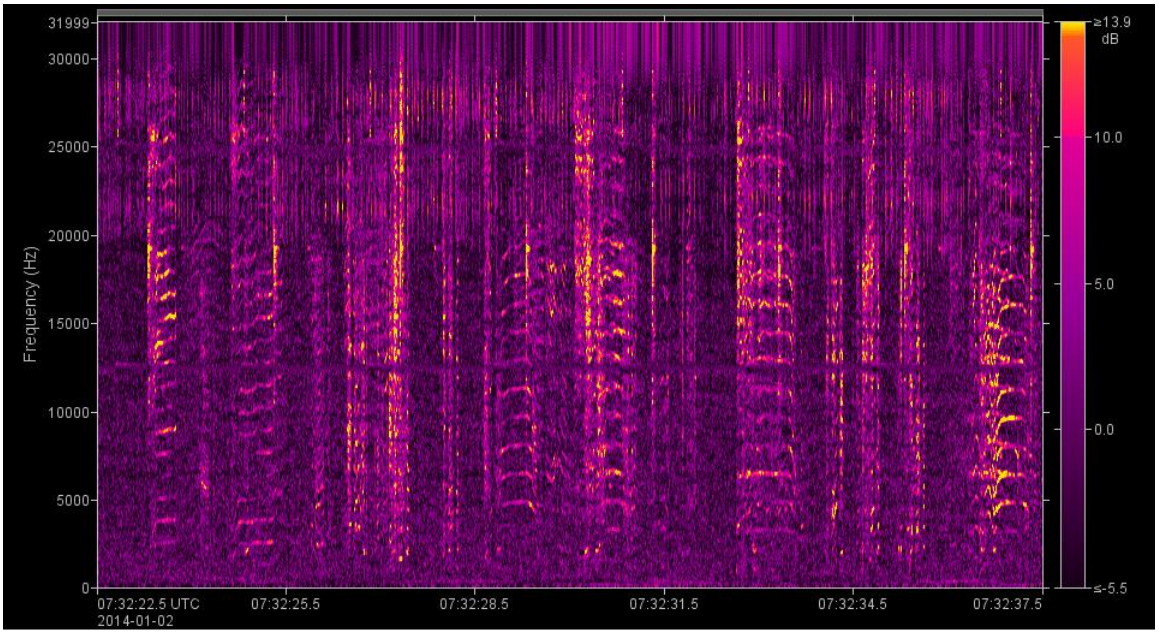


Figure 1.1 Pacific white-sided dolphin pulsed calls and echolocation clicks recorded by an Ocean Sonics icListen HF hydrophone sampling at 64 kHz and deployed at Barkley Canyon by Ocean Networks Canada. The spectrogram was produced in JASCO Applied Sciences' PAMLab software with a 1 Hz frequency step, 0.01 s time step, and 0.01 s frame length.

Pacific white-sided dolphins are very gregarious, with typical group sizes ranging from 40 to 1,000 or more individuals (Stacey & Baird, 1991; Heise, 1996), and sometimes associate with other species of delphinids (Soldevilla *et al.*, 2010). While they are sometimes seen in much smaller groups, a lone Pacific white-sided dolphin is unusual

(Morton, 2000). They are quite vocal, typically remaining silent only while resting (Goley, 1999), and they use communication signals during all non-rest behaviour states (Henderson *et al*., 2011). The meanings of their various calls are unclear, but they likely play a role in social interaction and maintaining group cohesion (Henderson *et al*., 2011; Rehn *et al*., 2007), and may also facilitate co-operative foraging in regions where this behaviour is observed (Heise, 1996; Van Opzeeland *et al.*, 2005; Vaughn-Hirshorn *et al.*, 2012; Eskelinen *et al*., 2016). They live for approximately 45 years, grow up to 2.5 m long, and are sexually monomorphic (Heise, 1996). They are broadly distributed, occurring throughout the temperate North Pacific Ocean (Leatherwood *et al.*, 1984; Stacey and Baird, 1991). While accurate population estimates are difficult to achieve for this species due to their attraction to boats, they are thought to be very abundant, with a population of approximately 223,400 individuals (Heise, 1996). Their diet is broad, including various species of pelagic fishes, bottom fishes, cephalopods, and jellyfish, though the specifics of their diet varies regionally (Black, 1994; Heise, 1996; Morton, 2000). Two populations have been reported in the eastern North Pacific Ocean south and north of the Southern California Bight, where genetic and morphological evidence suggest that a Baja California and a California/Oregon/Washington population overlap (Walker *et al.*, 1986; Lux *et al.*, 1997). It is generally assumed that Pacific white-sided dolphins in Canada are of the same population as those in the northern California/Oregon/Washington population, although Lux *et al.* (1997) suggested that a third North American population inhabiting Canada and Alaska may exist.

There is evidence that Pacific white-sided dolphins in the northeast Pacific exhibit both north-south seasonal movement (Leatherwood *et al.*, 1984; Green *et al.*, 1992; Green *et al.*., 1993; Forney *et al.*, 1995; Forney & Barlow, 1998) and inshore-offshore seasonal movement (Stacey & Baird, 1991; Morton, 2000), though data for Pacific white-sided dolphins in Canada and Alaska are sparse. Similarly, diel patterns in vocal activity have been demonstrated in both the southern and northern populations overlapping in the Southern California Bight (Soldevilla *et al.*, 2010; Henderson *et al.*, 2011), but no such analysis has been conducted north of California. This thesis investigates both seasonal and diel patterns in Pacific white-sided dolphin pulsed calls inhabiting Canadian waters through analysis of one year of near-continuous acoustic data collected from ONC's Barkley Canyon node.

1.2 Thesis Outline

The body of this thesis consists of three chapters covering three mostly independent bodies of work, two of which make use of data sets and tools produced in the previous chapter, followed by a general summary of the conclusions from these chapters. They are intended to be relatively self-contained, and so there is some repetition of introductory material and terms. The following outline summarizes the work presented in each chapter.

Chapter 2 describes the Barkley Canyon research site and the data collected at this site by ONC, which are used in subsequent chapters. This chapter also describes in detail the manual annotation process used to create a data set of acoustic signals of known origin, for use in training a machine learning classifier.

Chapter 3 begins with a general overview of different machine learning classifiers, performance metrics and validation methods before proceeding to describe the selection, training, and validation of a random forest classifier targeting Pacific white-sided dolphin pulsed calls, using the manual annotation data set produced in Chapter 2. Ranges of suitable values for minimum leaf size, confidence threshold, and forest size are selected from boxplots illustrating how altering these values affected the performance of both multiclass and binary random forest classifiers. Classifiers trained using every permutation of multiclass/binary condition and selected values for minimum leaf size, confidence threshold, and forest size are compared through repeated, nested 10-fold cross-validation, and these values are demonstrated to have little effect on the performance of random classifiers on this data set. A multiclass classifier with the default values suggested by Breiman (2001) is trained and validated using repeated, un-nested 10-fold cross-validation. The classifier's precision, recall, and F1-score for classification of non-mammal sounds and of humpback whale, sperm whale, orca, and Pacific white-sided dolphin vocalizations are reported.

Chapter 4 describes a statistical analysis of diel and seasonal patterns in Pacific white-sided dolphin vocal activity in Barkley Canyon. Acoustic signals in a full year of acoustic data collected at ONC's Barkley Canyon node are automatically classified by the classifier trained in Chapter 3. The Pacific white-sided dolphin classifications are manually verified prior to the statistical analysis. Diel patterns over the year and within each season are compared statistically, as are overall seasonal patterns and seasonal patterns within each diel period. Correlations with ambient noise in acoustic bands dominated by vessel noise are evaluated to rule out seasonal and diel patterns resulting from masking effects. The analysis reveals clear seasonal patterns, as well as season-specific diel patterns. A hypothesis explaining the season-specific diel patterns as evidence of different populations inhabiting the Barkley Canyon region is presented.

Chapter 5 summarizes the work of the previous three chapters.

Chapter 2 Creating a Manually Annotated Data Set

2.1 Data and Site Description

Ocean Networks Canada (ONC) is a not-for-profit organization focused on ocean science and monitoring. ONC operates multiple underwater observatories on the west, east and north coasts of Canada and provides open ocean data to users worldwide. Their North East Pacific Time-series Underwater Networked Experiments (NEPTUNE) observatory is based on an 840 km loop of fibre optic telecommunications cable extending from the west shore of Vancouver Island to the spreading ridge between the Juan de Fuca and Pacific plates. The observatory has six nodes, five of which are instrumented with a variety of oceanographic sensors collecting continuous data (Figure 2.1). Data are uploaded to ONC's online database, Oceans 2.0 (dmas.uvic.ca), in near real time and made freely accessible to scientists and the general public.

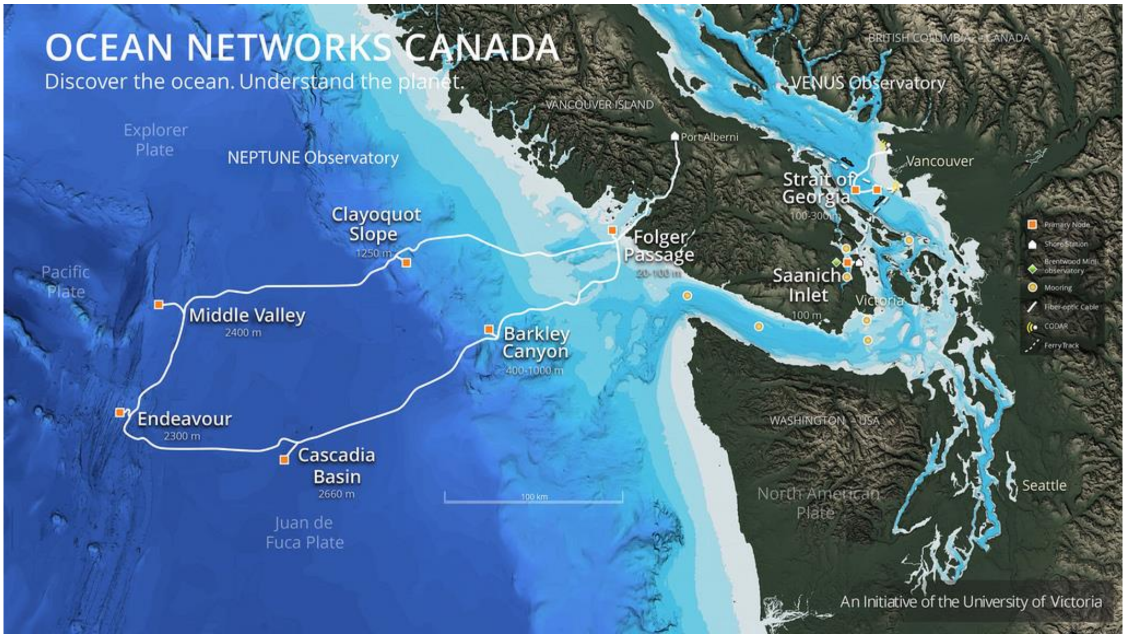


Figure 2.1 ONC's NEPTUNE observatory. Source: www.oceannetworks.ca

Acoustic data used in this project were collected from a hydrophone deployed at the Barkley Canyon node of the NEPTUNE observatory, located on the shelf break at 48° 25.6457’N 126°10.4799’W, approximately 60 km southwest of Vancouver Island. The node has eight instrumented platforms distributed from the canyon’s 400 m deep upper slope to its 985 m deep axis (Figure 2.2). Data were collected with an Ocean Sonics icListen hydrophone sampling continuously at a rate of 64 kHz with 24-bit depth, deployed 70 m from the upper slope instrument platform. Data considered here were recorded at a depth of 392 m from May 11, 2013 to May 3, 2014, and at a depth of 391 m from May 7, 2014 to January 12, 2015. Most of my analysis focuses on data collected in 2014.

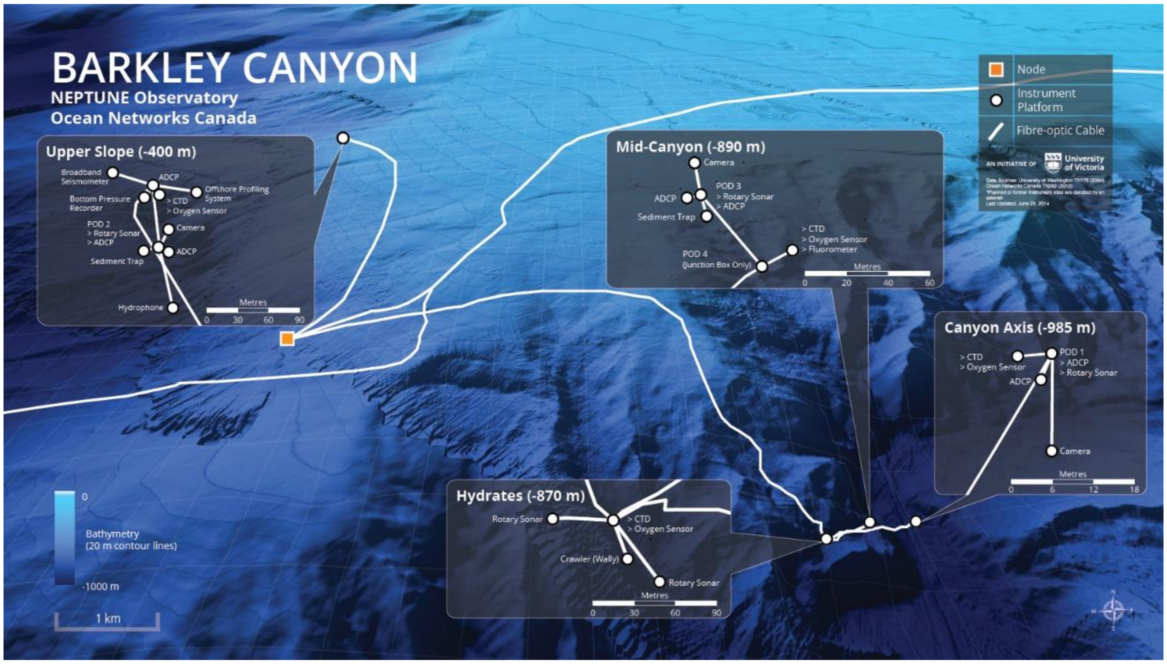


Figure 2.2 ONC’s Barkley Canyon node deployments. Source: www.oceannetworks.ca

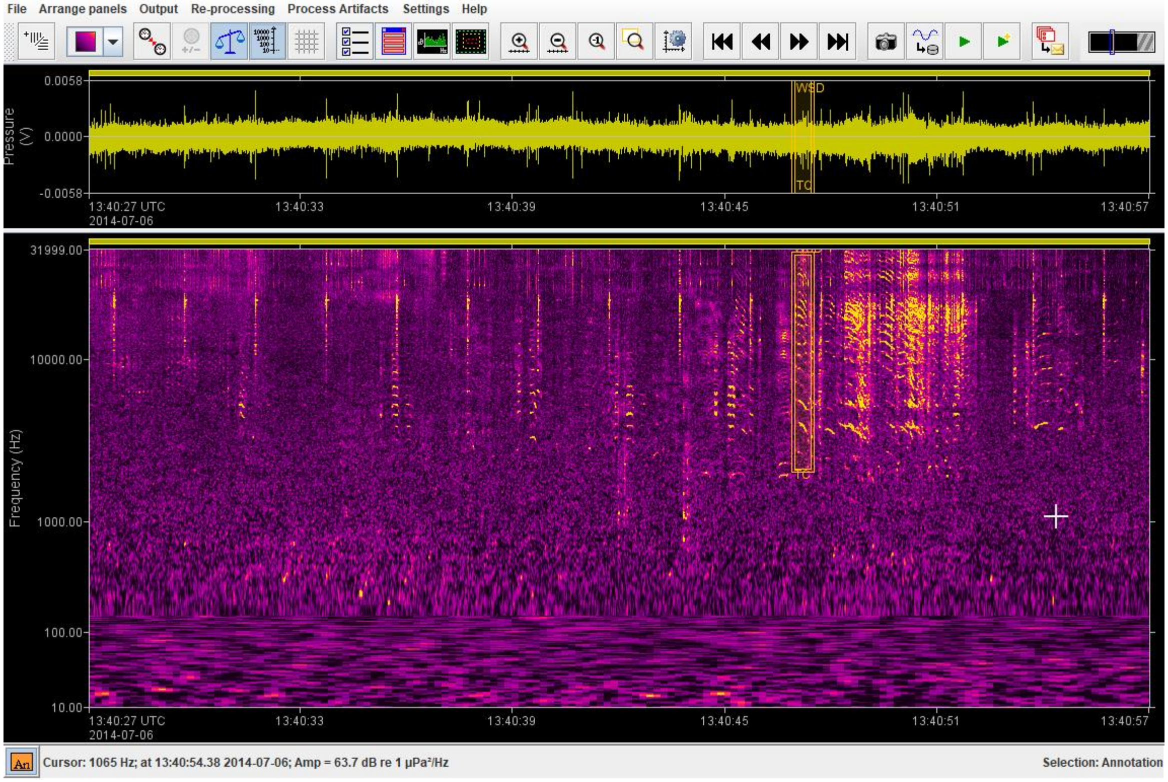
Though considered small relative to other submarine canyons, at only 13 km length and 6 km width, Barkley Canyon has been shown to substantially influence local water properties and currents, enhancing upwelling and forming cyclonic eddies capable of affecting zooplankton movement and trapping plankton near the canyon (Allen *et al.*., 2001; Mackas & Coyle, 2005). The increased availability of nutrients for primary productivity associated with enhanced upwelling is further increased in this region by nutrient input from the Vancouver Island Coastal Current (Allen *et al.*., 2001), and by nutrient outflow from the Strait of Juan de Fuca (Stefánsson & Richards, 1963). One would expect the nutrient availability in the photic zone in this region to result in high levels of primary productivity, reflected by the presence of primary and secondary consumers. This expectation is confirmed by the euphausiid and hake aggregations that occur over and around the canyon during the summer (Mackas *et al.*, 1997).

The regional productivity at Barkley Canyon is expected to also draw larger species, like cetaceans, to the canyon. Hake and euphausiids are predated upon directly by several local cetacean species, including Pacific white-sided dolphins, blue whales (*Balaenoptera musculus*), fin whales (*Balaenoptera physalus*), and humpback whales (*Balaenoptera novaeangliae*), and also may serve as food sources for squids and larger fishes that other odontocetes prey upon (Walker *et al.*, 1986; Schoenherr, 1991; Clapham *et al.*, 1997). Barkley Canyon is also located within the migration routes of several species of baleen whale (Norris *et* al, 1999; Burtenshaw *et al.*, 2004) and within the overlapping habitats of offshore, inner coast Bigg's, outer coast Bigg's, northern resident and southern resident orcas (*Orcinus orca*). Given that all these cetacean species are vocally active, Barkley Canyon is an ideal site for acoustic-based cetacean research, and was selected as the research site for this project.

2.2 Manual Annotation

Continuous acoustic data recorded by the hydrophone deployed at ONC’s Barkley Canyon node during the year 2014 were downloaded through the Oceans 2.0 data portal as five-minute long waveform audio file format (WAV) files. Exploratory analysis of files from this site revealed that, while many cetacean vocalization events lasted under ten minutes, most, if not all, were longer than five minutes. Alternate five-minute files were selected for manual analysis to minimize effort while maintaining a high likelihood of capturing most vocalization events in the manual analysis. To capture seasonal variation, files from the first four days of each month over the course of the year were included in the manual analysis. Thus, the duty cycle for manual analysis was alternate five-minute intervals over the first four days of each month of 2014, and included approximately 7% of the data in the full data set – a typical percentage of the data for this type of analysis (e.g., Ross & Allen, 2014). This effort took approximately 430 hours to complete.

Data were analyzed manually using PAMlab acoustic analysis software (JASCO Applied Sciences), which can produce a customized spectrogram display of acoustic files and allows manipulation and playback of the original data file (Figure 2.3). Acoustic data were viewed over the full 32 kHz bandwidth of the files on a logarithmic scale to enable visualization of cetacean sounds over the full bandwidth. Spectrograms were composed using Hamming windows, and different time-frequency resolutions were used for different bandwidths to maximize visibility of signals in different frequency bands (Table 2.1).



Cursor: 1065 Hz; at 13:40:54.38 2014-07-06; Amp = 63.7 dB re 1 μPa²/Hz

Figure 2.3 Pacific white-sided dolphin vocalizations viewed in PAMLab (JASCO Applied Sciences) with the settings used here for manual annotation. The yellow box is an annotation.

Table 2.1 Time-frequency characteristics of the multi-part spectrogram at three different bandwidths produced in PAMlab for manual annotation of bioacoustics data.

Frequency Band (Hz)	**Frequency Resolution (Hz)**	**Time Step (s)**	**Frame Length (s)**
0 - 150	0.5	0.2	1
150 – 15,000	10	0.01	0.05
15,000 – 32,000	100	0.005	0.01

Annotations were carried out in PAMlab by using the cursor to draw a box around the signal of interest and selecting the species from a resulting drop-down menu. Information about the annotation was automatically saved in an external log file, including the filename, details about the hydrophone deployment, and the start time, end time, upper frequency, lower frequency, root mean square (RMS) sound pressure, sound pressure level (SPL), sound exposure level (SEL), species, and call type.

One call per species per file was annotated to maximize the variation in noise conditions, except for northern right whale dolphins (*Lissodelphis borealis*), for which all calls in the same file were annotated due to their relative rarity in the data set. This effort produced 8,764 annotations. Of these, 2,889 annotated calls were ambiguous or of unclear origin, and were not included in further analysis. Sounds were considered ambiguous or of unclear origin if they were an unfamiliar sound within the phonic range of multiple species and in the absence of familiar call types, an unfamiliar sound occurring only once, or a sound with a signal to noise ratio (SNR) too low to classify with confidence. Delphinid echolocation clicks were also considered ambiguous because delphinid echolocation click frequencies (approximate bandwidth of 20 kHz to 100 kHz) far exceed the bandwidth of the data analyzed here (1 Hz to 32 kHz), and the frequencies that would show differences between the clicks of local delphinid species were not represented in the data (Soldevilla *et al.*, 2008). As a result, delphinid echolocation was not included in the analyses presented in this thesis.

In all, 75.7% of the files analyzed (4,857 of 6,418) contained marine mammal vocalizations, and 3,929, or 61.2%, contained vocalizations from multiple species. Sounds of non-mammalian origin were not annotated manually, as it was not possible to predict ahead of time which non-mammal sounds the feature extraction algorithm would detect. Instead, all detections from files not containing marine mammal signals or possible marine mammal signals were considered to be of non-mammalian origin in future analysis. These detections include anthropogenic signals such as the clicks of an acoustic Doppler current profiler (ADCP) installed nearby and vessel noise, and may include seismic sounds, deployment self-noise, and benthic organisms interacting with the platform.

Annotations from this analysis and annotations produced by six other analysts at the Department of Fisheries and Oceans and JASCO Applied Sciences studying five other sites in the Northeast Pacific Ocean were used by JASCO Applied Sciences to develop feature extraction and classification packages (Mouy *et al.*, 2015). While the feature extraction algorithms associated with the low- and high-frequency classification packages are non-specific, extracting features for all transient signals within a pre-defined bandwidth, the classification packages focused only on blue whales, fin whales, humpback whales, orcas and Pacific white-sided dolphin signals. JASCO did not use the northern right whale dolphin, sperm whale (*Physeter macrocephalus*), or fish sounds from this annotation effort in their classifier development. The breakdown of annotations by species is given in Table 2.2.

Table 2.2 Summary of manual annotation results by species from Barkley Canyon annotation effort, and total annotations used in JASCO classifier development from this and other annotation efforts.

Species	Annotations from Barkley Canyon	Total Annotations for Feature Extractor and Classifier Development
Blue whale	217	1,890
Fin whale	2,298	4,532
Humpback whale	2,730	6,418
Northern right whale dolphin	14	-
Orca	37	2,917
Pacific white-sided dolphin	194	231
Sperm whale	410	-
Fish	5	-

The feature extractor associated with the high-frequency classification package operates on signals between 20 Hz and 8 kHz and produces 94 features for each detected transient signal. A list of features is given in Appendix A. The high-frequency classification package has 5 classes: humpback whale, orca, Pacific white-sided dolphin, fish, and other. This package was run on the full-year continuous acoustic data set from 2014 and used to extract features for transient signals in the 20 Hz to 8 kHz bandwidth, as well as to classify all signals detected.

The classification package was developed using data from a variety of locations with the intent that it be used as a broad tool rather than for targeting a particular species. However, this project seeks to classify signals from one species and one deployment with high accuracy. Without modification, such a classifier may not perform as well as one designed for targeted species identification and trained with site-specific data if the local noise conditions differ significantly from the noise conditions of the training data, or if there are signals common to the region that were not present in the training data set. Hence, the classifier's performance was assessed independently so that its classification parameters could be optimized for use in the circumstance of the present study. Signals from each category were selected randomly for manual inspection to verify classification accuracy. This process revealed that the high-frequency classification package misclassifies sperm whale signals (a class the classifier was not trained to identify) as Pacific white-sided dolphins with high confidence when used to classify data from Barkley Canyon. It is therefore an inappropriate classifier to use at this location. Instead, the features extracted for each detection from this classification effort were used to develop a new classifier, described in Chapter 3.

To develop the new classifier, the features associated with manually-annotated calls were isolated to produce a manual-annotation data set. Spectrograms were produced for detected signals with start and end times identified within ± 1 s and minimum and maximum frequencies within ± 1 kHz of annotated call parameters, and with tags containing the time and frequency boundaries of the detected signals and the species name of the associated annotation. The spectrograms were visually inspected so that signals from sources other than the annotated species with similar time and frequency characteristics could be removed from the manual-annotation data set.

Detections associated with spectrograms containing a signal of the expected species with time and frequency characteristics closely aligning with the detection time and frequency characteristics given in the tags were accepted. Detections associated with spectrograms containing signals originating from other sources with similar time and frequency characteristics were rejected. Table 2.3 summarizes the quantities of detections verified as originating from annotated calls.

Table 2.3 Summary of manual annotations for species vocalizing above 20 Hz, and automated detections manually verified as the same signal. Non-mammal signals were not manually annotated; signals detected in files not containing marine mammals were considered to be non-mammal sounds.

Species	**Manual Annotations**	**Verified Detections**
Humpback whale	2,730	512
Northern right whale dolphin	14	9
Orca	37	12
Pacific white-sided dolphin	194	105
Sperm whale	410	373
Non-mammal	-	139,657

The sample size for all classes was reduced through the verification process, for several reasons. Some annotated calls were not detected by the automated detector. The detector used here is an updated version of that used in Hannay *et al.* (2013), who found that human analysts can detect signals at lower SNRs than the automated detector, resulting in fewer automated detections than manual annotations. As humpback whales are the most-frequently observed species in the data set, with calls that can be detected at a greater distance than other classes, lack of detection due to poor SNR disproportionately affected this class. Furthermore, humpback vocalizations were frequently noted to be distant in the manual-annotation data set. Other calls were actively removed from the manual-annotation data set because they overlapped in both time and frequency with calls from different species phonating in the same frequency band.

The numbers of Pacific white-sided dolphin, northern right whale dolphin, and orca calls detected and verified as aligning with manual annotations were not sufficient for classifier development. Hence, the files from the original manual annotation effort were re-analyzed to produce larger data sets for these species. All orca calls found in the manual annotation data set were annotated, and 2 to 20 Pacific white-sided dolphin calls per file were annotated as well. This effort increased the Pacific white-sided dolphin sample size to 486, and the orca sample size to 47. Signals from only 2 (southern resident and inner coast Bigg’s orcas) out of the 5 orca ecotypes present in this region were found in this stage of annotation.

Many orca and Pacific white-sided dolphin calls have similar time and frequency characteristics, and can be difficult to distinguish using automated techniques. Northern right whale dolphin calls are also similar to Pacific white-sided dolphin calls and are difficult to distinguish accurately. To maximize the likelihood that classifiers developed in this analysis would be able to distinguish between these two categories, a further effort was applied to orca and northern right whale dolphin annotations. In this effort, all files annotated as containing or possibly containing orcas or northern right whale dolphins in the ONC annotations database from the Barkley Canyon hydrophone deployments from May 11, 2013, to May 3, 2014, and May 7, 2014, to January 12, 2015, were downloaded and annotated manually.

This additional effort increased the orca sample size to 436 samples. The signal diversity in the orca data set was enriched through this effort as well, which should improve the classifier's ability to correctly classify a variety of orca calls. While the original manual-annotation data set contained only southern resident and inner coast Bigg's orca signals, the expanded annotation data set contained these as well as northern resident, outer coast Bigg's and offshore orca signals. All signals potentially attributable to northern right whale dolphins were ambiguous (the few calls found had very low SNR and were too similar to Pacific white-sided dolphin calls to be certain of species of origin), and were not added to the manual annotation data set. Training data set composition pre- and post-data set improvement are summarized in Table 2.4.

Table 2.4 Manual annotation data set composition by species before and after an additional annotation effort to include more orca, Pacific white-sided dolphin, and northern right whale dolphin signals in the manual annotation data set. This effort involved further annotation within the original data set, and manual annotation of files from outside the manual annotation data set but within the same hydrophone deployment previously tagged as containing marine mammal signals. Species affected by this effort are bolded.

Species	Verified Detections	Verified Detections-post data set improvement
Humpback whale	512	512
Northern right whale dolphin	9	9
Orca	**12**	**436**
Pacific white-sided dolphin	**105**	**486**
Sperm whale	373	373
Non-mammal	139,657	139,657

Chapter 3 Random Forest Classification of Pacific White-Sided Dolphin Pulsed Calls

3.1 Introduction

Passive acoustic monitoring has some advantages over visual data collection when studying coastal and offshore marine mammal species. It can be used to monitor much wider areas than can be monitored visually, it is not limited by visibility conditions dependent on daylight or weather, and cabled or autonomous systems can be deployed in areas that prove logistically difficult to access regularly, such as protected or deep sea areas. However, while visual methods result in data being collected only while species of interest are present, acoustic methods produce large quantities of continuous or duty-cycled data that must be analyzed after collection to determine which files contain acoustic signals from the species of interest. Manually completing this analysis is time consuming and can be impractical for large data sets. Automated classification can be an efficient alternative to manual analysis of large acoustic data sets. To use automated classification, only a subset of the full data set needs to be analyzed by hand. The manually-annotated data are used to train a classification algorithm to distinguish between the different acoustic classes present in the data set, such as different species or call types, anthropogenic sounds, or environmental sounds. The classification algorithm is then applied to the full data set to detect and classify the acoustic signals of interest.

3.1.1 Performance Metrics

There are several metrics that can be used as a measure of classifier performance, all of which are calculated from a confusion matrix which indicates the proportion of correct and incorrect classifications in each class, as illustrated in Table 3.1.

Table 3.1 Layout of a binary confusion matrix.

	Predicted: Yes	**Predicted: No**
Actual: Yes	True Positive (TP)	False Negative (FN)
Actual: No	False Positive (FP)	True Negative (TN)

Accuracy is often assumed to be a useful metric for assessing the performance of a classifier. It ranges from 0 to 1 and is calculated as

$$\boldsymbol{Accuracy = \frac{TP+TN}{TP+FN+FP+TN}} \qquad \textbf{(1)}$$

However, accuracy can say more about the composition of the annotation data set than the quality of the classifier. For example, if 90% of the data in a two-class (binary) classification problem are of one class and 10% are of the other, the accuracy of a classifier that classifies all data presented to it as belonging to the majority class will be high (0.9), but the classifier is of no practical value. Furthermore, in non-binary classification problems, accuracy cannot give information on the performance of the classifier on each class, and instead gives an estimate of classifier performance based on a summation of true positives and true negatives from all classes. Provost *et al.* (1998) argue that accuracy is also an inappropriate measure when assessing data sets with unknown distributions, or data sets with unequal misclassification costs for different classes.

It is also common to look at the true positive rate (*TPR*), also called recall or sensitivity, and false positive rate (*FPR*), which Provost *et al.* (1998) put forth as alternatives to accuracy. Both of these measures range from 0 to 1, and are calculated as

$$\boldsymbol{TPR = recall = sensitivity = \frac{TP}{TP+FN}} \tag{2}$$

$$\boldsymbol{FPR = \frac{FP}{FP+TN}} \tag{3}$$

These metrics allow for a more meaningful assessment of classifier performance than accuracy, and can be used in receiver operating characteristic (ROC) plots to optimize different parameters of the classification algorithm. This method plots the true positive rate against the false positive rate as some parameter of interest is varied across its range of definition (Figure 3.1). The greater the area under the ROC curve, or the closer the curve approaches the upper left corner of the plot, the better the performance of the classifier. Like accuracy, ROC curves are inappropriate performance measures when class data are unbalanced (Davis & Goadrich, 2006). ROC curves are also challenging to extend to multiclass problems, as they are necessarily pairwise comparisons.

Recall and precision have become popular alternatives to accuracy and ROC curves for assessing classifier performance (e.g., Kotsiantis *et al.*, 2006; Gillespie *et al.*, 2013; Hannay *et al.*, 2013; Jacobson *et al.*, 2013; Liu *et al.*., 2013; Ross & Allen, 2014). Precision is the proportion of results classified as positive that are truly positive, giving a measure of the effect of negative examples being misclassified as positive. Precision ranges from 0 to 1 and is calculated as

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}} \tag{4}$$

Precision and recall can be plotted against one another to produce a curve similar to ROC curves, but with a different shape (Figure 3.1).

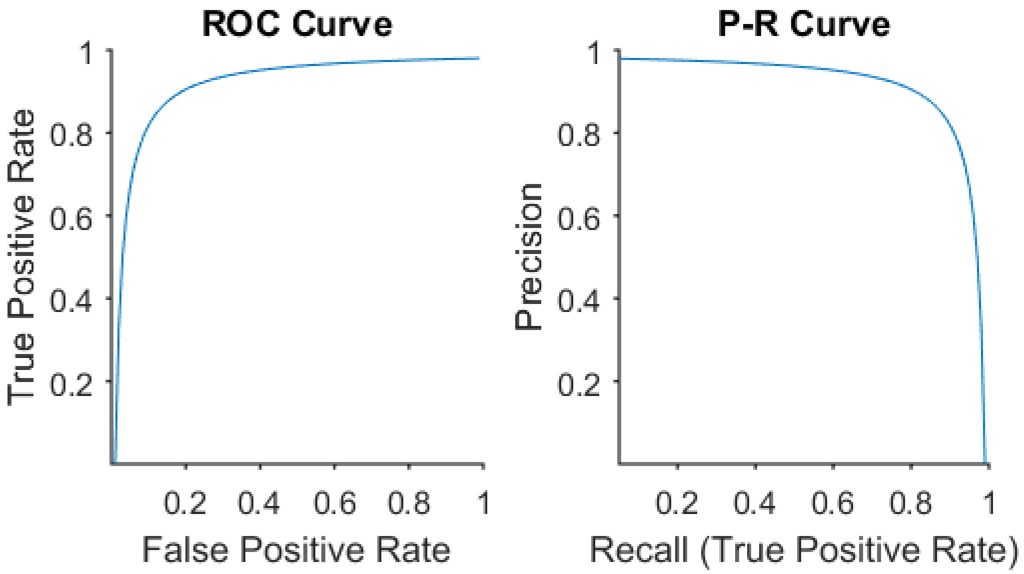


Figure 3.1 Comparison of idealized receiver operating characteristic (ROC) and precision-recall (P-R) curves.

ROC curves start near the lower left of the plot, rise, and then approach horizontal. The ROC curve is optimized where the curve is closest to the upper left corner. Precision-recall curves begin high on the y axis and nearly parallel with the x axis, and then descend as they approach the end of the x axis. They are optimized where the curve is closest to the upper right corner. Classifiers that are optimized in precision-recall space are also optimized in ROC space, but classifiers that are optimized in ROC space can be sub-optimal in precision-recall space (Davis & Goadrich, 2006). Precision and recall can also be used to calculate the F1-score, which can be used as an overall indicator of classifier performance. The F1-score is the harmonic mean of precision and recall, and is calculated

$$\boldsymbol{F1 = 2 \times \frac{precision \times recall}{precision + recall} = 2 \times \frac{\left(\frac{TP}{TP+FP}\right) \times \left(\frac{TP}{TP+FN}\right)}{\left(\frac{TP}{TP+FP}\right) + \left(\frac{TP}{TP+FN}\right)}} \tag{5}$$

Maximizing the F1-score also optimizes precision and recall. Precision, recall and F1-score are used as the performance metrics for this study because they are more robust metrics than accuracy and are more informative metrics than ROC curves.

3.1.2 Validation and Model Selection Methods

Classifier performance must be evaluated using data that were not used in training the classifier. The simplest approach to this is the hold-out method, in which a portion of the manually-annotated data set is kept separate from the training set and used exclusively for testing. However, holding data out from the training set can decrease the robustness of the classifier by limiting the diversity of the data set used for training, as key samples for differentiating between classes may not be included in training. Rather than use a classifier trained on only a portion of the data available, classifiers can be trained on the full data set and their performance approximated by cross-validation. In leave-one-out cross-validation, a classifier is trained using all but one of the samples, and tested on that single sample. This process is repeated for each sample in the data set, such that each sample is left out of the training set and used for testing once. The results from each training and testing iteration are summed and used as an approximation of the performance of a classifier trained on the full data set. While there can be considerable variation in error estimates from cross-validation, the average cross-validation error tends to converge to the correct error value with enough repetitions (Hastie *et al.*, 2001).

While leave-one-out cross-validation allows for the development of a more robust classifier than the hold-out method, it is very computationally intensive. *N*-fold cross-validation is a compromise between the hold-out and leave-one-out methods. In *N*-fold cross-validation, the manual annotation data set is divided into *N* distinct subsets or folds. *N*-1 folds are used to train a classifier and one fold is used to test it. This process is repeated *N* times, such that each fold is left out and used to test once. *N*-fold validation is less computationally intensive than leave-one-out cross-validation while still providing the advantage of allowing for all training samples to be used in the development of the classifier it is validating. However, *N*-fold cross-validation can be more biased than leave-one-out cross-validation due to fewer data being used for training in each iteration, and also has high variance resulting from splitting the data into folds.

Bootstrap validation, where training sets of equal size to the data set are created through random selection with replacement and tested on the data that were not selected for training, is sometimes used instead of cross-validation because bootstrap validation has lower variance than both leave-one-out and *N*-fold cross-validation, and may perform better on small data sets (Efron, 1979; Efron, 1983; Braga-Neto & Dougherty, 2004). However, the variance of *N*-fold cross-validation can be reduced through repetition (Kim, 2009), and bootstrap validation suffers from considerable bias. Even the bias-corrected .632+ bootstrap method proposed by Efron and Tibshirani (1997), which seeks to account for the bias resulting from each bootstrap sample having only 0.632*n* distinct observations where *n* is the size of the data set, exhibits bias when used on small and large data sets. When comparing *N*-fold cross-validation and bootstrap validation using various values for the adjustable parameters and fundamentally different classification methods on several data sets of varied size and composition, Kohavi (1995) found that the bias of the .632+ bootstrap method could be predictably manipulated, rendering the accuracy estimates achieved using this method meaningless. Kohavi (1995) recommend using 10-fold cross-validation for model selection, as cross-validation was less biased than bootstrap validation and moderate numbers of folds (10-20) resulted in lower variance than using either high or low numbers of folds. They also found that 10-fold cross-validation's bias could be further reduced by using stratified folds, where each fold contains the same proportions of samples from each class as is found in the full data set. A similar analysis that assessed repeated *N*-fold cross-validation, repeated hold-out, and .632+ bootstrap found that, while the performance of different estimators was somewhat dependent on sample-size and classifier type, repeated 10-fold cross-validation was the most robust estimator (Kim, 2009).

All of the aforementioned validation methods can be used for model selection. However, using the same data to validate and select a model can optimistically bias performance estimates (Varma & Simon, 2006; Krstajic *et al.*., 2014). Combining either cross-validation or bootstrap validation with the hold-out method, such that a model is selected using cross-validation or bootstrap validation and validated on a held-out data set, is an intuitive way to handle this problem, but suffers the same drawbacks of using the hold-out method as a validation technique – the model cannot be trained using all of the data available. Nested *N*-fold cross-validation allows for all data to be used in training and testing of the model, without biasing the performance estimate for the selected model (Varma & Simon, 2006; Krstajic *et al.*., 2014). In nested *N*-fold cross-validation, an inner cross-validation loop is run using *N*-1 folds. Models with all possible combinations of the hyperparameters of interest, such as number of trees and tree depth for a random forest, are trained using *N*-2 folds and tested on the one remaining inner fold, with the final fold held out for testing a model trained on the outer loop with the hyperparameters that won the inner loop. The process is repeated *N*-2 times, such that each inner fold has been used to test the models once, and a model is selected based on the results. A model with the selected hyperparameters is then trained on all *N*-1 folds from the inner loop and tested on the remaining fold. The entire process is repeated *N*-1 times, such that each fold has been held out from the model selection process and used to estimate the performance of the model selected by the inner loop once. This process is illustrated in Figure 3.2. If model selection is stable, meaning the same or similar hyperparameters are selected by each inner fold, then nested cross-validation closely approximates the true performance of a model, even when used to both select and evaluate the model (Varma & Simon, 2006). Repeated, nested 10-fold cross-validation was chosen as the model selection and validation method for this study.

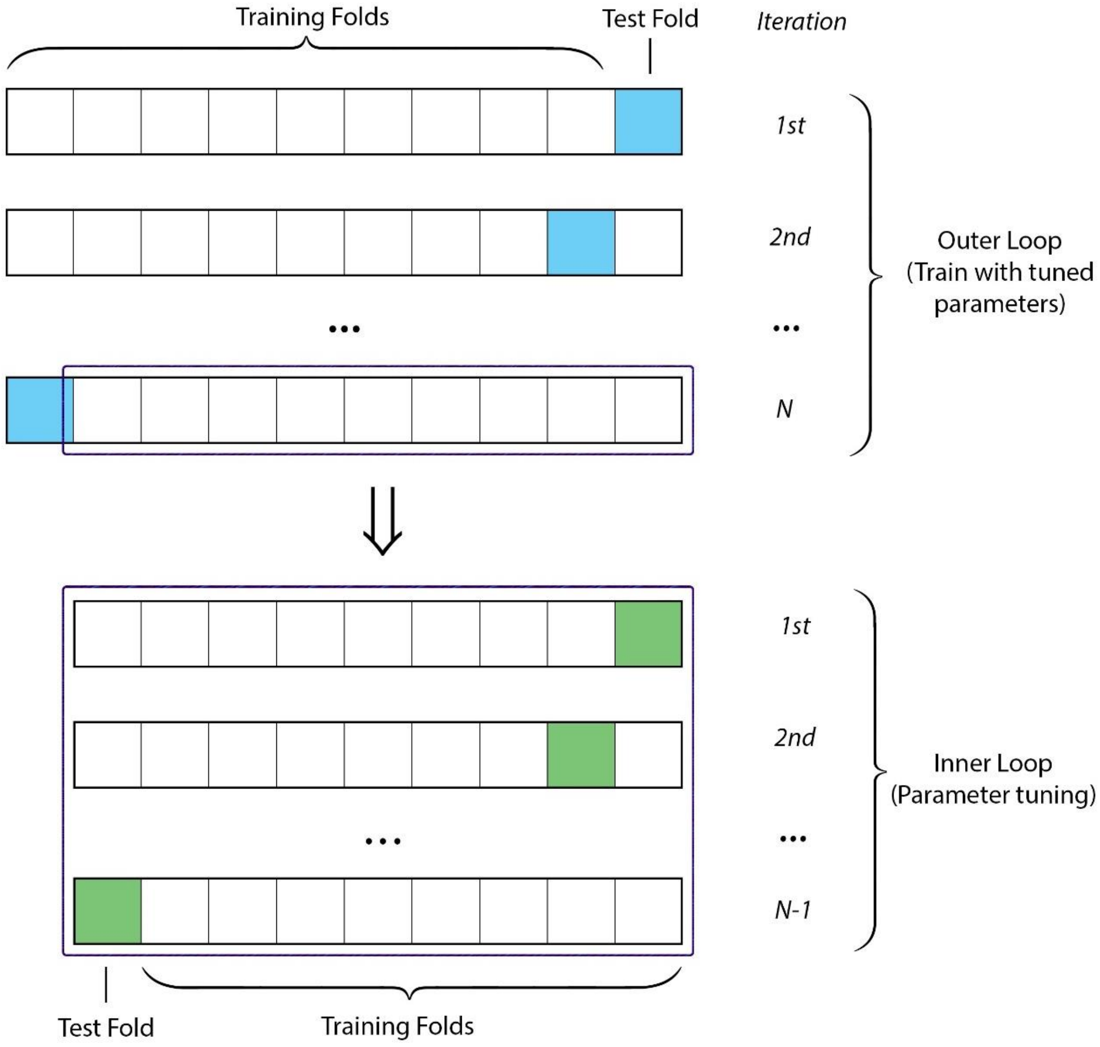


Figure 3.2 Diagram illustrating nested N-fold cross-validation.

3.1.3 Types of Classifiers

At their most basic level, classifiers are methods of categorizing data into pre-defined classes. They do so by identifying patterns in data, and drawing boundaries between parameters relating to different classes so that the classes can be differentiated from one another. While these decision boundaries can be set manually, to do so when there are many parameters to consider is impractical. It is also difficult to determine *a priori* which parameters will prove most useful for classification, or whether relationships between different parameters might prove more useful than each parameter on its own. Supervised machine learning can be a much faster and more effective method of developing classifiers, and is commonly applied to bioacoustic classification problems. Supervised machine learning algorithms use manually annotated data to create decision boundaries that optimize a performance metric (Alpaydin, 2010).

Many machine learning classification algorithms exist, each with its strengths and weaknesses. Early bioacoustic classification was typically done using neural network classifiers (e.g., Gaetz *et al.*., 1993; Au *et al.*, 1995; Mercado & Kuh, 1998; Murray *et al.*., 1998; Deecke *et al.*., 1999), which had originally been developed in the pursuit of artificial intelligence. Neural networks are machine learning classifiers designed to mimic the decision processes of the brain, using new annotated data presented to them during training to modify their pattern recognition boundaries and become more accurate (Rosenblatt, 1961; Rumelhart *et al.*., 1986; Rumelhart *et al.*, 1988). Neural networks can be very effective at differentiating between highly similar classes, such as the same call type given by different orca matrilines (Deecke *et al.*, 1999), or different call types produced by the same species (Deecke & Janik, 2005). However, they are difficult to optimize and require a great deal of analyst oversight and manual tuning relative to newer automated classification techniques.

Support vector machines (SVM) have steadily gained popularity since their initial development in the mid 1990s (Cortes & Vapnik, 1995). In this method, a kernel chosen *a priori* is used to transform the feature data into a high-dimensional feature space, such that a linear or quadratic decision surface dividing the class features can be created and used to classify novel data (Cortes & Vapnik, 1995; Dietrich *et al.*., 1999; Wu & Zhou, 2005). SVMs are simpler to train than neural networks, and can perform as well as or better for some problems, explaining their popularity (Dietrich *et al.*, 1999; Byvatov *et al.*, 2003; Turesson *et al.*., 2016).

More recently, Breiman (2001) developed the increasingly-popular random forest algorithm. This algorithm creates an ensemble of independent decision trees that make binary decisions based on the values of a single variable at each tree node. Each tree in the ensemble, or forest, is trained with a bootstrapped sample of the data, and may not include all the features available for decision making. When presented with novel unclassified data, each tree in the ensemble votes on a class, and the sample is classified as the class that receives the most votes. While individual decision trees generally make poor classifiers, ensembles of decision trees can be very effective classifiers. Random forest is one of the easiest machine learning algorithms to implement, and in many cases it performs as well as or better than both neural network and SVM classifiers used on the same classification problems (Madjarov *et al.*, 2012; Liu *et al.*, 2013; Ness, 2013; Esfahanian *et al.*, 2015). While including too many trees can result in overfitting the data (Kertész, 2016), random forest classifiers are generally very robust compared to other machine learning algorithms (Breiman, 2001; Díaz-Uriarte & de Andrés, 2006). Random forest has become one of the most popular classifiers for bioacoustic classification problems, with excellent results (e.g., Barkley *et al.*, 2011; Briggs *et al.*, 2012; Ross & Allen, 2014; Muoy, 2015). Given its simplicity and effectiveness compared to other machine learning classifiers popular in the bioacoustics field, Breiman (2001)'s random forest algorithm was selected as the machine learning algorithm to be used in this study.

3.2 Materials and Methods

The automated detectors and feature extractor used in this thesis to extract features of manually annotated calls for classifier development were developed by JASCO Applied Sciences as part of their proprietary high-frequency classification package, SpectroDetector (Maloney *et al.*, 2014, Mouy *et al.*, 2015). The feature extractor operates on detections of transient signals between 20 Hz and 8000 Hz, and extracts 94 parameters, which include a variety of both time-frequency and spectral features (see Appendix A).

The manual-annotation data set used for classifier development was created using acoustic data collected in 2014 from the Barkley Canyon site on ONC's NEPTUNE cabled underwater observatory. Details of the manual-annotation data set creation process are given in Chapter 2. The composition of the data set is shown in Table 2.4.

Analysis was completed using MATLAB R2016a software and the Statistics and Machine Learning Toolbox Version 10.2 (The Mathworks, Inc., 2016). The *TreeBagger* function, which is derived from Breiman's random forest classification algorithm (Breiman, 1984; Loh & Shih, 1997; Breiman, 2001; Loh, 2002; Meinshousen, 2006) was used for classifier training. The manual-annotation data set developed in Chapter 2 was divided into separate classes based on the manual classifications (see Table 2.4). As an initial exploratory approach, several hyperparameters that may influence classifier performance were varied separately, used to train a classifier, and evaluated using un-nested, repeated 10-fold cross-validation, to determine appropriate value ranges to use in this approach. These hyperparameters include sample size, forest size (number of trees), minimum leaf size (tree depth), confidence threshold (fraction of trees that must vote for the dolphin class before the sample is classified as dolphin), and binary *versus* multiclass classification. The selected ranges are given in Table 3.2. Forest size and minimum leaf size ranges were selected based on boxplots of the F1-score for Pacific white-sided dolphin classification as that parameter is changed, such that the selected range would include the inflection point and continue just past the point at which increasing that hyperparameter no longer increases the F1-score. Sample size ranges start with the sample size of the smallest class and increase to past the point at which increasing the sample size no longer shows a change in F1-score on boxplots. The selected confidence threshold range is the full range of possible confidence intervals.

Table 3.2 Hyperparameter values used for training and selecting an optimal random forest classifier using repeated, nested 10-fold cross-validation.

	Binary	**Multiclass**
Sample Size (per class)	unbalanced, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000	unbalanced, 350, 400, 450, 500, 550
Forest Size	10, 15, 20, 25, 30	50, 100, 150, 200
Min Leaf Size	1, 3, 7	1, 3, 7
Confidence Threshold	0.5, 0.6, 0.7, 0.8, 0.9, 1	0.2, 0.4, 0.6, 0.8, 1

In the multiclass condition, the original classes were Pacific white-sided dolphin, orca, sperm whale, humpback whale, northern right whale dolphin, and non-mammal. However, the northern right whale dolphin category proved to be too small to be a class on its own during this preliminary analysis, as it was consistently misclassified in every testing condition. Since most northern right whale dolphin vocalizations were consistently misclassified as non-mammal, the northern right whale dolphin calls were included in the non-mammal category for this classifier development.

The data were divided into 10 approximately-stratified folds. Given the disproportionate size of the non-mammal data set, 500 non-mammal samples were randomly selected for use in training, selecting and validating models. While most implementations of cross-validation advocate for either random or pseudorandom fold construction, folds were constructed by hand for this project to compensate for the lack of independence between samples resulting from the relatively small size of the data set (14,173 sounds, 1,816 of which were produced by marine mammals, from 6,418 files). The presence of samples from the same file, or even from the same encounter, in both the training and testing data sets could leak knowledge to the model selection and testing algorithm and optimistically bias model evaluation. To prevent this, folds were carefully constructed such that no fold contained data from the same event as any other fold, taking into account overlapping events from different species. Since many examples of Pacific white-sided dolphins and orcas came from just a few events, it was not possible to create perfectly stratified folds, and some folds contained more or fewer examples of these classes.

Table 3.3 Distribution of data classes in the 10 folds used for training, selecting, and validating an optimal random forest classifier with repeated, nested 10-fold cross-validation.

Fold	Humpback	Orca	Pacific white-sided dolphin	Sperm Whale	Non-Mammal
1	53	94	43	38	50
2	52	21	42	37	50
3	51	104	43	37	50
4	51	93	48	38	50
5	51	18	41	38	50
6	51	31	41	37	50
7	51	18	97	37	50
8	51	20	43	37	50
9	51	19	44	37	50
10	50	19	44	37	50
Total	512	437	486	373	500

The folds listed in Table 3.3 and the hyperparameters used in Table 3.2 were used in repeated, nested 10-fold cross-validation. Random forest classifiers were trained using the *TreeBagger* function in MATLAB. For sample sizes larger than the complete data set for a particular class, the training folds from that class were concatenated and the training data set was upsampled to *N*–2*N/10* in the inner fold, and *N–N/10* in the outer fold, such that the training sample size was equivalent to the desired sample size minus the held-out testing folds. If the training set for a class was larger than the desired size, it was downsampled without replacement to maintain the greatest possible sample diversity.

While fold construction was not repeated in each iteration so that independence of samples between folds could be maintained, data within the training fold were randomly shuffled in each repetition to account for variance caused by the order in which samples are presented to the training algorithm (Mouy *et al.*, 2013). Each inner loop and each outer loop were repeated 10 times, totalling 100 repetitions per hyperparameter setting. Hyperparameters were evaluated via grid search, such that each of the 2043 possible hyperparameter permutations was used to train a classifier 10 times in each inner fold. Models were selected based on a modified F1-score for Pacific white-sided dolphin classification,

$$\boldsymbol{modified\ F1 = F1 - \sigma} \qquad \textbf{(6)}$$

where σ is the standard deviation. The modified F1-score was used instead of the raw F1-score for model selection because the optimal model is that which has both a high F1-score and a low variance (X. Muoy, pers. comm.). A model with high variance across folds and repetitions will not perform consistently on novel data, so standard deviation was subtracted from the F1-score for model selection to decrease the likelihood that a model with both a high F1-score and high variance would be selected over a model with a nearly as high F1-score and a much lower variance.

3.3 Results

Repeated, nested 10-fold cross-validation gave precision, recall, and F1-score values of 0.7937, 0.7696, and 0.7717, respectively, for classification of Pacific white-sided dolphin pulsed calls. However, model selection through this process was unstable;

models with different hyperparameters were selected by each inner loop. The models selected in each inner loop are given in Table 3.4.

Table 3.4 Hyperparameters selected by each loop of repeated, nested 10-fold cross-validation. *n* represents the number of samples per class in a balanced model.

Fold	Class Division	Balance	*n*	Forest Size	Min Leaf Size	Confidence Threshold	Modified F1-score
1	multiclass	unbalanced	-	200	1	0.8	0.8419
2	multiclass	unbalanced	-	200	1	0.8	0.8311
3	binary	unbalanced	-	25	3	0.7	0.8133
4	multiclass	unbalanced	-	200	3	1	0.7697
5	binary	unbalanced	-	30	5	0.5	0.7430
6	multiclass	balanced	500	150	1	1	0.7383
7	multiclass	balanced	350	200	5	0.8	0.7280
8	binary	unbalanced	-	15	7	0.5	0.7327
9	multiclass	balanced	500	200	3	0.2	0.7444
10	binary	unbalanced	0	30	7	0.9	0.7553

Substantial differences between inner and outer loop scores would indicate a major failure of the model selection process. In this case, modified F1-scores from the outer loop test folds differed from inner loop scores at most by 0.0867, with a mean decrease in modified F1-score of 0.0104 from the outer loop to the inner loop across folds, indicating that the model selection process was sound despite instability. Further investigation revealed that model selection was unstable because the performances of most models were approximately equal; the modified F1-scores of all models in all folds ranged from 0.6363 to 0.8419, with a maximum within-fold range of 0.1289. Up to 155 models within a given fold had modified F1-scores within one tenth of that of the selected model, with an average of 70.5 models per fold yielding modified F1-scores within one tenth of that fold's selected model.

Given that most models performed similarly, a multiclass random forest classifier with default values of 100 trees, minimum leaf size of 1, and confidence threshold of 0.2 was selected based on general recommendations in Breiman (2001) and trained using the unbalanced data set. This model was validated with 100 repetitions of un-nested 10-fold cross-validation, which yielded precision, recall and F1-score values of 0.7903, 0.7986, and 0.7906 respectively. The F1-score estimated for this classifier was in close agreement with that of the repeated, nested 10-fold cross validation, differing by only ~2%. A confusion matrix containing the pooled classification results from all repetitions and folds is given in Table 3.5.

Table 3.5 Confusion matrix of classifications given during 100 repetitions of un-nested 10-fold cross-validation by a random forest classifier with 100 trees, minimum leaf size of 1, and confidence threshold of 0.2, trained using unbalanced data.

		Predicted				
		Humpback Whale	**Orca**	**Non-Mammal**	**Sperm Whale**	**Pacific White-Sided Dolphin**
Actual	**Humpback Whale**	49401	684	881	101	133
	Orca	577	34800	3172	525	4526
	Non-Mammal	2646	2851	33309	6169	5025
	Sperm Whale	0	128	6881	29811	480
	Pacific White-Sided Dolphin	0	3412	5596	423	39169

3.4 Conclusions

The investigation described in this chapter aimed to develop a high-performance random forest classifier targeting Pacific white-sided dolphin vocalizations. Various balanced and unbalanced, binary and multiclass models with different class sample sizes, forest sizes, minimum leaf sizes, and confidence thresholds were validated and compared using repeated, nested 10-fold cross-validation. While similar F1-scores between inner and outer loops indicated that the validation process and annotated data set were sound, the performance of most models was roughly equivalent and resulted in unstable model selection. A multiclass random forest classifier was trained for use in Chapter 4 using the unbalanced data set and default values for forest size, minimum leaf size, and confidence threshold, and evaluated using repeated, un-nested 10-fold cross-validation. This classifier had precision, recall, and F1-score of 0.7903, 0.7986, and 0.7906 respectively, when used to classify Pacific white-sided dolphin calls. These are good scores for marine mammal classification (e.g., Gillespie *et al*., 2013; Jacobson *et al*,, 2013; Mouy *et al*., 2015). The classifier also performed well for the other marine mammal classes, with similar scores for orca and sperm whale classification and very high scores for humpback whale classification (Table 3.6).

Table 3.6 Precision, recall, and F1-score for classification of humpback whale, orca, sperm whale, and Pacific white-sided dolphin vocalizations, and non-mammal sounds using a random forest classifier.

	Precision	**Recall**	**F1-Score**
Humpback Whale	0.9649	0.9388	0.9516
Orca	0.7982	0.8310	0.8143
Non-Mammal	0.6662	0.6683	0.6673
Sperm Whale	0.7992	0.8051	0.8021
Pacific White-Sided Dolphin	0.7903	0.7986	0.7906

Chapter 4 Temporal Patterns in Pacific White-Sided Dolphin Pulsed Calls

4.1 Introduction

Many organisms, from a variety of taxa and inhabiting a variety of habitats, exhibit diel patterns in their behaviour. These diel patterns may be tied to biologically-driven circadian rhythms (Oshima *et al.*, 1989; Sánchez-Vásquez & Tabata, 1998), or to ecological factors affecting survivorship, such as hunting (Beier *et al.*, 1995; Andrews *et al.*, 2009) or predator avoidance (Nelson & Vance, 1979; Hays, 2003). Investigations into animal diel patterns can deliver new information about habitat use and predator-prey relationships, and can inform conservation efforts targeting at-risk species.

Some cetaceans have been demonstrated to exhibit diel patterns in their behaviour (e.g., Todd *et al.*, 2005; Wiggins *et al.*, 2005; Baumann-Pickering *et al.*, 2015). However, cetacean diel patterns have not been thoroughly researched. Previous investigations into the nocturnal habits of cetaceans have been carried out either in captivity (e.g. McBride & Herb, 1948; Saayman *et al.*, 1973; Therrien *et al.*., 2012) or using tags (e.g., Baird *et al.*, 2001; Klatsky *et al.*, 2007; Scott & Chivers, 2008). However, the generalizability of the results of such studies may be limited. Studies of captive cetacean behaviour may not be representative of the behaviour of wild cetaceans, because the movements of captive animals are limited, they do not have to hunt, and they live in unnatural social groupings. Their diel patterns may also be influenced by the diel patterns of their caregivers, as was demonstrated by Therrien *et al.* (2012) in their research on the diel patterns of eight captive bottlenose dolphins (*Tursiops truncatus*) at the Brookfield Zoo. While the use of tags is more likely to represent behaviour typical of wild cetaceans, attaching tags necessitates disturbance of subjects prior to data collection through approaching and tagging the animal, which may alter their subsequent behaviour. Suction-cup style tags are relatively short lasting, typically remaining attached to the subject for only a few hours to a few days (e.g., Lerczak *et al.*, 2000; Amano & Yoshioka, 2003), and require agreeable weather conditions and potentially costly vessel time to approach and tag research subjects, and to follow them post-tagging. Other types of tags require invasive attachment hardware that causes local trauma at the attachment site and can leave the tagged animal susceptible to infection and further injury from drag, or from bumping or snagging the tag. Further, invasive tags can alter the post-tagging behaviour of subjects (Geertsen *et al.*, 2004; Balmer *et al.*, 2010), and are generally not permitted for use on large numbers of at-risk animals. These aspects of tag-based research typically make for small sample sizes and can make studying offshore species especially challenging.

With the growing popularity of passive acoustic monitoring as a data collection strategy, which can be used instead of or in conjunction with visual data collection methods, the question of whether cetaceans exhibit diel patterns in their behaviour can now be addressed non-invasively *in situ,* over long time periods and with large sample sizes. Many organisms, both terrestrial and marine, exhibit changes in vocal behaviour associated with diel behaviour changes (e.g., Walsh & Inglis, 1989; Wiggins *et al.*, 2005; Locascio & Mann, 2008; Todd & Pearse, 2009; Sosa-López & Mennil, 2014), so passive acoustic monitoring may prove useful for demonstrating diel patterns in delphinid behaviour. However, passive acoustic monitoring is a relatively new area of research with relatively few studies using passive acoustic technology to investigate diel patterns in delphinid behaviour. While several studies found that dolphins are most vocally active at night (Soldevilla *et al.*, 2010; Hodge *et al.*, 2013; Dede *et al.*, 2014; Deconto *et al.*, 2016), others have found that dolphins are most vocally active during the day (Soldevilla *et al.*, 2010) or during twilight hours (Goold, 2000), and one study found no observable diel patterns in dolphin vocalization rates (Philpott *et al.*., 2007).

Soldevilla *et al.* (2010) found two distinct diel acoustic behaviour patterns for Pacific white-sided dolphins in the Southern California Bight, with different patterns attributable to dolphins producing different click types. These click types exhibited spectral peaks and notches at different frequencies, with click type A having higher frequency peaks than click type B. Dolphins producing click type A produced the most clicks at night, with peak click bout occurrence at sunset. Dolphins producing click type B vocalized the most during the day. It is thought that these two dolphin groups are from either different populations of the same species or are morphologically identical but genetically distinct cryptic species (a northern population producing click type A and a southern population producing click type B) feeding on different prey, with their habitats overlapping in the Southern California region (Soldevilla *et al.*, 2008; Soldevilla *et al.*, 2010; Henderson *et al.*, 2011). Two populations of Pacific white-sided dolphins overlapping in this region, one from Baja California and another from California/Oregon/Washington, were previously identified from morphological evidence (Walker *et al.*., 1984; Walker *et al.*., 1986), and both exhibit north-south and inshore-offshore seasonal movements (Leatherwood *et al.*, 1984; Green *et al.*, 1992; Green *et al.*., 1993; Forney *et al.*, 1995; Forney & Barlow, 1998). Preliminary genetic evidence suggests there may also be a third population inhabiting Canada and Alaska (Lux *et al.*, 1997). Whether the Pacific white-sided dolphins in the Barkley Canyon region are members of the California/Oregon/Washington population identified in these studies or members of a different population is not known. There are also no studies to the author's knowledge investigating the diel patterns of Pacific white-sided dolphins in Barkley Canyon or other places in the Northeast Pacific Ocean. This study seeks to use passive acoustic techniques to determine whether Pacific white-sided dolphins in the Barkley Canyon region exhibit a diel pattern in their vocal behaviour.

4.2 Materials

Features of transient sounds in a full year of 24-bit WAV data collected nearly continuously at 64 kHz sampling rate by an Ocean Sonics icListen HF hydrophone deployed at the Barkley Canyon node of ONC's NEPTUNE cabled underwater observatory in 2014 were extracted using an automated detector and feature extractor created by JASCO Applied Sciences as part of a proprietary acoustic classification package (Maloney *et al.*, 2014; Mouy *et al.*, 2015). This package extracts 94 time-frequency and spectral features of detected transient signals (see Appendix A). Extracted sounds were classified by the random forest classifier trained in Chapter 3 using the *TreeBagger* algorithm, based on Breiman's random forest classification algorithm (Breiman, 1984; Loh & Shih, 1997; Breiman, 2001; Loh, 2002; Meinshousen, 2006) in the Statistics and Machine Learning Toolbox Version 10.2 in MATLAB R2016a (The Mathworks, Inc., 2016). This software and tool box were also used for data manipulation and statistical analyses in this chapter.

The solar calculator spreadsheet publicly provided online by the Earth System Research Laboratory Global Monitoring Division of the National Oceanic and Atmospheric Administration (NOAA) (www.esrl.noaa.gov) was used to produce a time series of solar elevation angles relative to the horizon with 6-minute resolution that could be used to divide the acoustic data by diel period. The solar elevation angles were corrected for atmospheric refraction based on mean conditions (1010 hPa pressure, 10°C temperature, yellow light). The spreadsheet is based on the equations provided by Jean Meeus in *Astronomical Algorithms* (Meeus, 1991), and can be used to calculate the position of the sun for any time and date, accurate to within one minute, for any latitude within ± 72°. The spreadsheet is accurate for dates between 1901 and 2099, although predictions may vary somewhat from observed conditions due to local changes in temperature, pressure, and other weather conditions.

An additional sound pressure level data set was created using JASCO Applied Sciences' PAMlab acoustic analysis software to examine temporal patterns in ocean noise that could contribute to masking effects. If acoustic absence of Pacific white-sided dolphins is correlated with increased vessel and wind noise, which dominate low frequency bands in the 10 to 500 Hz range but can be present up to 100 kHz (Dahl *et al*., 2007), then patterns found in Pacific white-sided dolphin vocal activity could be patterns in ambient noise rather than true patterns in vocal behaviour (e.g. Helbe, 2013). While noise generated by rain also overlaps the bandwidth in which Pacific white-sided dolphins vocalize, it is higher frequency than the bandwidth analyzed here and so was not evaluated for its effect on Pacific white-sided dolphin vocalization detectability (Dahl *et al*., 2007).

4.3 Methods

Transient sounds in the 20 – 8000 Hz band were extracted from all unaltered full-band data from 2014 and classified using the random forest classifier trained in Chapter 3, totalling 86,917 files analyzed and 8,985,064 sounds automatically classified. The 9,958 files collected by the military during diversions (as mandated by their memorandum of understanding with ONC), and subsequently sent to ONC as high- and low-pass filtered files were not included in this analysis, as the filtered data are missing a portion of the band the classifier analyzes. To ensure the validity of the data set, all files flagged by the classifier as containing Pacific white-sided dolphin calls were evaluated manually by an expert analyst, and files containing only misclassifications were removed from the Pacific white-sided dolphin data set. Classification data were coalesced into a presence data set with 5-minute resolution.

Presence data were aligned with solar elevation angle data and with sound pressure level (SPL) data based on time and date, and tagged with their season, defined as the time between the equinoxes and solstices, and with their diel period. Diel periods included dawn, day, dusk, and night, and were defined as the times between and during nautical twilight, where nautical twilight is the time when the sun is between 6° and 12° below the horizon. Since these data sets had different temporal resolutions (presence and SPL data had 5-minute resolution and solar elevation angle data had 6-minute resolution), presence and SPL data were attributed to whichever diel period the majority of the acoustic file these data were derived from was recorded during. A second diel split of day, twilight, and night (DTN), where dusk and dawn have been pooled into one period to increase the twilight sample size and look for crepuscular patterns, was also analyzed.

Diel periods were compared based on the fraction of files during each diel period of each day that contained Pacific white-sided dolphin calls, hereafter referred to as the fractional presence. Similarly, seasons were compared based on the fractional presence for each day in each season. Fractional presence was used instead of raw presence data to account for variation in the number of files per day and within each diel period, since some files are missing from the data set due to military intervention or maintenance, and the numbers of files per diel period are inherently unbalanced. Further segmented data sets, including diel patterns within each season, and seasonal patterns within each diel period, were also analyzed. Data sets were compared using the Kruskal-Wallis test because all temporal data splits failed the Lilliefors test (a robust normal distribution test that does not assume which normal distribution would be the underlying distribution of the data), most data sets also failed the Brown-Forsythe test (a homogeneity of variance test that does not assume normality), and data transformations did not result in normally distributed data with homogenous variances (Whitlock & Schluter, 2009). The Kruskal-Wallis test is similar to ANOVA but is non-parametric and compares medians and distributions instead of means. It is based on the Wilcoxon rank-sum test which assigns each datum a rank based on its value, sums the ranks within each category, and compares sample-size-adjusted rank-sums. Data splits that showed significant differences were further analyzed with pairwise Wilcoxon rank-sum tests to determine which categories were different.

To account for potential masking effects from vessel and wind noise, average SPLs per 5-minute file in the 1/3 octave bands with centre frequencies of 100 Hz, 250 Hz, and 500 Hz were compared as well for all diel periods and seasons. These bands were selected because they are strongly affected by vessel and wind noise but are low enough in frequency to not be affected by Pacific white-sided dolphin vocal activity, and should therefore be positively correlated with vessel and wind noise rather than dolphin noise. All ambient noise data sets failed the Lilliefors test for normality, and most failed the Brown-Forsythe test for homogeneous variances, so the seasonal and diel SPL data were compared using the Kruskal-Wallis test. The non-parametric Spearman's rank correlation coefficient was used to test for correlations between hourly SPL data and hourly Pacific white-sided dolphin fractional presence data.

4.4 Results

Automated classification yielded 647,074 sounds classified as Pacific white-sided dolphin in 38,168 files. Of these files, 4,612 were confirmed to contain Pacific white-sided dolphin pulsed calls through manual spectrogram analysis. While it may seem that a disproportionate number of files flagged as containing Pacific white-sided dolphins did not, this number makes sense when considering that each file contained several hundred sounds, so even a small percentage of misclassifications could flag a file as containing Pacific white-sided dolphin pulsed calls. Also, some discarded files contained delphinid echolocation and were temporally adjacent to files containing pulsed calls; these files may have contained pulsed calls with low signal-to-noise ratio that were correctly classified but went unnoticed during manual analysis. The number of files per season and diel period containing Pacific white-sided dolphin pulsed calls is given in Table 4.1.

Table 4.1 Summary of the total number of files, and number of files containing Pacific white-sided dolphin (PWSD) pulsed calls, per diel period and per season in one year of acoustic data collected from Barkley Canyon.

	Winter		**Spring**		**Summer**		**Fall**	
	PWSD	**Total**	**PWSD**	**Total**	**PWSD**	**Total**	**PWSD**	**Total**
Dawn	195	614	323	758	450	743	150	438
Day	3410	10403	6485	15150	8701	15391	2907	8182
Dusk	229	579	312	756	383	15391	198	431
Night	4282	11627	2858	6530	3736	6340	3576	8226
Totals	8116	23223	9978	23194	13270	23223	3576	8226

4.4.1 Ambient Noise

There were significant differences between seasons in the daily average SPLs for the 100 Hz, 250 Hz, and 500 Hz 1/3 octave bands ($p=6.11\times10^{-17}$, 5.89×10^{-10}, and 1.36×10^{-11}, respectively), with highest SPLs in winter, spring and fall, and lowest in summer. In pairwise tests, differences were found between seasons in all three 1/3 octave bands for winter-summer (100 Hz: $p=1.43\times10^{-8}$; 250 Hz: $p=8.09\times10^{-8}$; 500 Hz: $p=9.56\times10^{-9}$) and spring-summer (100 Hz: $p=2.85\times10^{-14}$; 250 Hz: $p=9.91\times10^{-10}$; 500 Hz: $p=1.19\times10^{-8}$), in two 1/3 octave bands for summer-fall (250 Hz: $p=1.05\times10^{-6}$; 500 Hz: $p=6.82\times10^{-10}$), and only in the 100 Hz band for winter-fall ($p=1.01\times10^{-6}$) and spring-fall ($p=1.21\times10^{-10}$). No significant differences were found for winter-spring in any band. There were no significant differences in SPLs between the DTN or DDDN diel periods.

Moderate to weak correlations between fractional hourly Pacific white-sided dolphin vocal presence and mean hourly SPLs from all three bands were found, but the strength of correlation decreased with increasing frequency and the direction of correlation was inconsistent (100 Hz: $r=0.3471$, $p=7.8957\times10^{-207}$; 250 Hz: $r=0.2745$, $p=4.4396\times10^{-127}$; 500 Hz: $r=-0.2080$, $p=1.4491\times10^{-72}$). Note that this test had the highest sample size for all statistical comparisons presented here (7,341 hours of fractional presence and SPL data) resulting in greater statistical power than other tests presented.

4.4.2 Seasonal Analysis

There were significant differences in the fractional presence of Pacific white-sided dolphins between seasons ($\chi^2=31.96$, $p=5.34\times10^{-7}$). These differences are illustrated by Figure 4.1. Pairwise comparisons revealed differences between every season pair except for winter-fall and spring-fall (winter-spring: $p=0.0226$; winter-summer: $p=2.85\times10^{-7}$; winter-fall: $p=0.3177$; spring-summer: $p=5.25\times10^{-4}$; spring-fall: 0.2633; summer-fall: $p=3.85\times10^{-5}$).

These differences were still significant when broken down by DTN diel period (Figure 4.2), with day showing the most pronounced differences (day: $\chi^2=20.68$, $p=1.0\times10^{-4}$; twilight: $\chi^2=16.69$, $p=8.0\times10^{-4}$; night: $\chi^2=34.05$, $p=1.9\times10^{-7}$); however, only winter-summer (day: $p=5.85\times10^{-7}$; twilight: $p=1.1210^{-4}$; night: $p=1.60\times10^{-5}$) spring-summer (day: $p=0.0015$; twilight: $p=0.0104$ night: $p=0.0026$), and summer-fall (day: $p=0.0039$; twilight: $p=0.0072$ night: $p=1.83\times10^{-5}$) showed significant differences within all three diel periods. Differences in winter-spring and spring-fall pairings were significant only in the day (p=0.0031, and p=0.0293, respectively), and winter and fall did not differ significantly in any diel category. When twilight was further divided into dawn and dusk, dawn showed significant seasonal differences while dusk did not (χ^2=27.84, p=3.92×10^{-6}, and χ^2=6.62, p=0.0852, respectively). Pairwise comparisons for dawn revealed significant differences for the winter-spring, winter-summer, spring-summer, spring-fall, and summer-fall pairs (p=0.0184, 3.06, 0.0035, 6.5×10^{-2} and 1.10×10^{-4}, respectively). No differences were found for winter-fall.

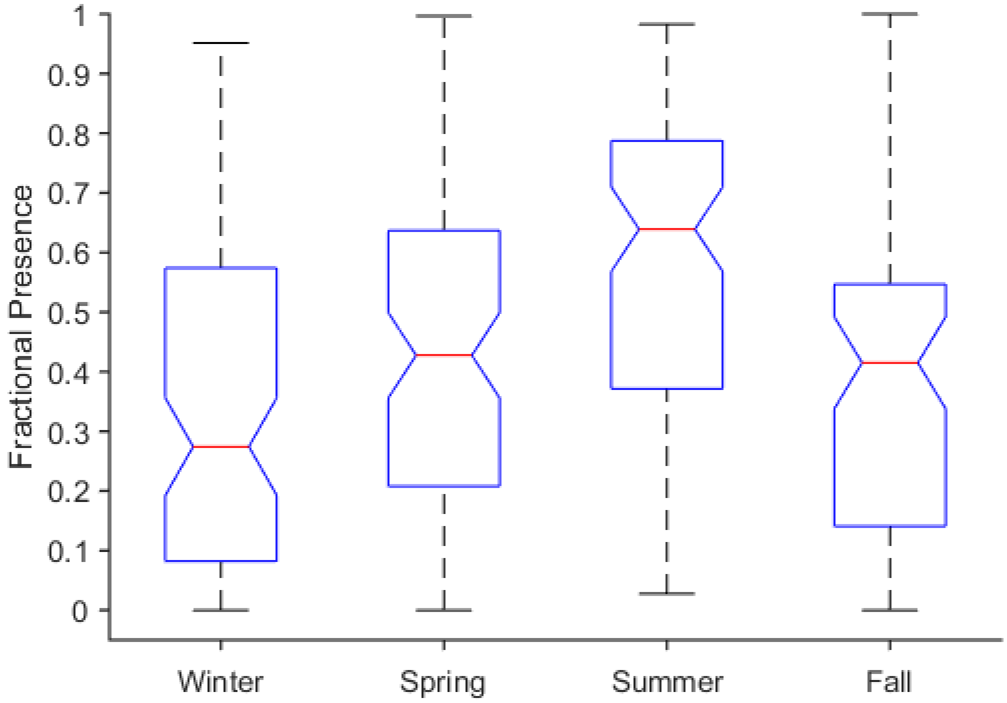


Figure 4.1 The seasonal fractional presence of Pacific white-sided vocal behaviour at Barkley Canyon, where fractional presence is calculated as the fraction of files containing Pacific white-sided dolphin pulsed calls. The blue boxes represent the 25th to 75th percentiles, and the red lines indicate the medians. Whiskers are calculated as ±2.7σ, where σ is the standard deviation.

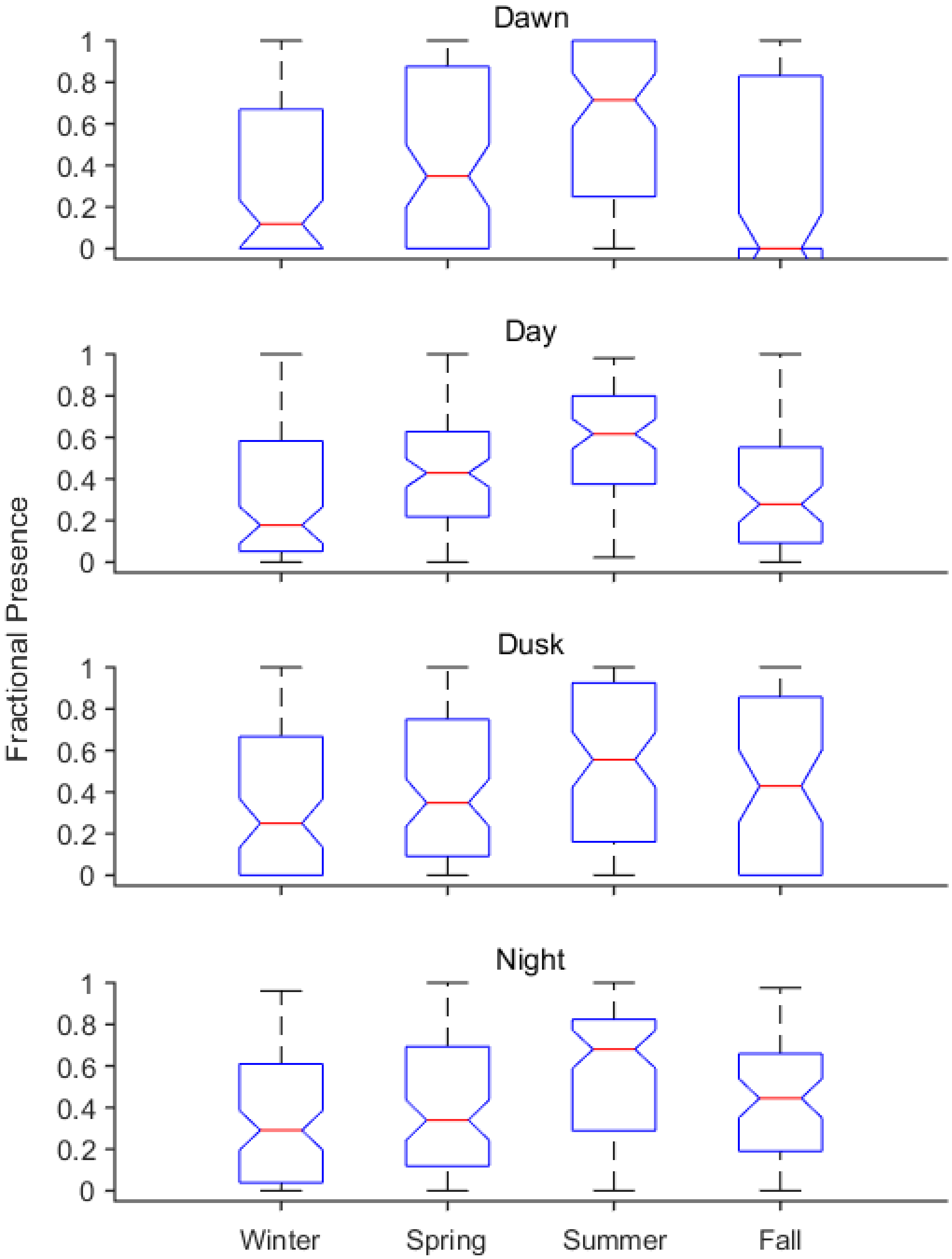


Figure 4.2 Seasonal fractional presence of Pacific white-sided dolphin vocal behaviour at Barkley Canyon for each diel period, where fractional presence is calculated as the fraction of files per season containing Pacific white-sided dolphin pulsed calls. The blue boxes represent the 25th to 75th percentiles, and the red lines indicate the medians. Whiskers are calculated as ±2.7σ, where σ is the standard deviation.

4.4.3 Diel Analysis

No significant differences were found between the dawn, day, dusk or night periods in the full year data set (Figure 4.3). Pooling dawn and dusk into one twilight category did not change the result.

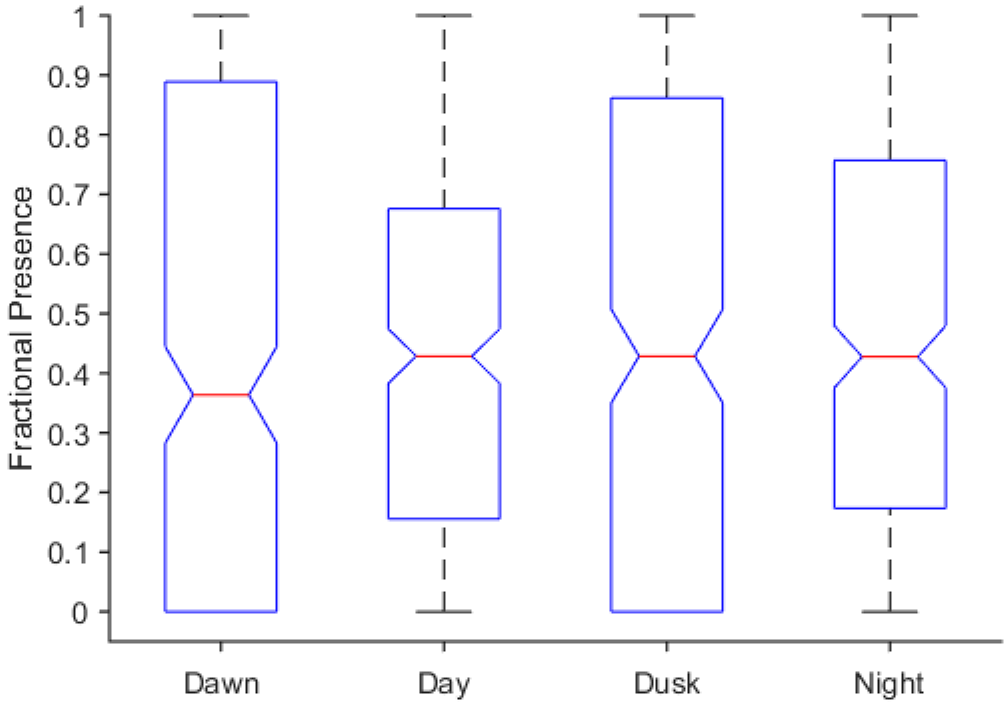


Figure 4.3 Distribution of Pacific white-sided dolphin fractional vocal presence at Barkley Canyon during the dawn, day, dusk, and night diel periods over a year, where fractional presence is calculated as the fraction of files containing Pacific white-sided dolphin pulsed calls. The blue boxes represent the 25th to 75th percentiles, and the red lines indicate the medians. Whiskers are calculated as ±2.7σ, where σ is the standard deviation.

No differences were found in the DTN data set when analyzed by season. However, day shows the lowest fractional presence in both the fall and winter data sets, while fractional presence in spring and summer appears the same across diel periods. This pattern is clearer when dawn and dusk are analyzed separately (Figure 4.4).

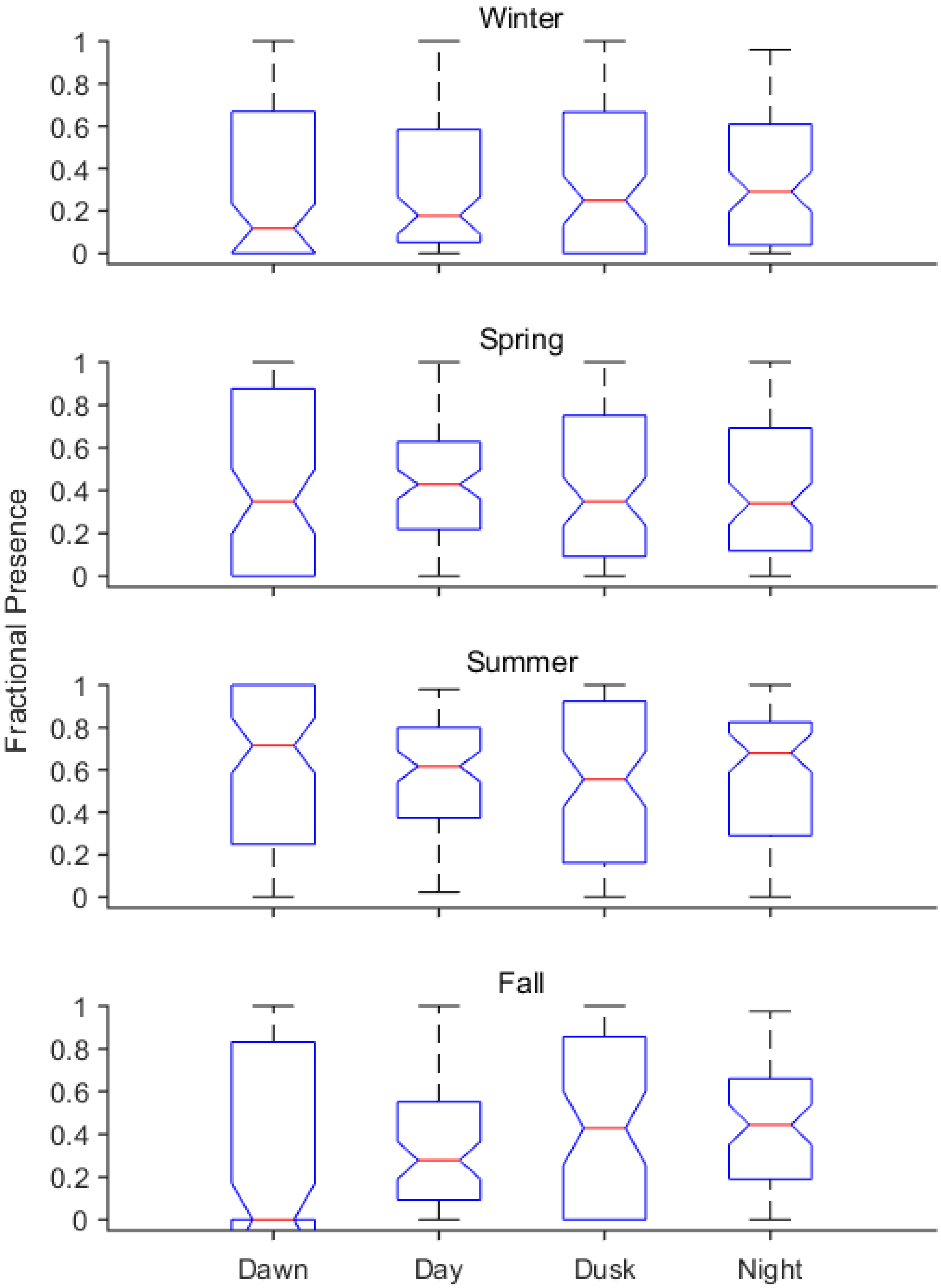


Figure 4.4 Diel fractional presence of Pacific white-sided dolphin vocal behaviour at Barkley Canyon during each season, where fractional presence is calculated as the fraction of files per diel period containing Pacific white-sided dolphin pulsed calls. The blue boxes represent the 25th to 75th percentiles, and the red lines indicate the medians. Whiskers are calculated as ±2.7σ, where σ is the standard deviation.

While no significant differences were found in the dawn-day-dusk-night (DDDN) data set when analyzed by season, a trend appeared in the fall and winter data sets, where pulsed call activity is lowest at dawn and increases throughout the day, with maximum fractional presence during dusk and night (Figure 4.4). These data sets also had the fewest dolphin encounters as a result of being smaller data sets and having the lowest fractional presence of dolphin vocal activity, so winter and fall, which did not differ from one another within any diel period, were combined to increase the sample size and re-analyzed. When combined, significant differences were found between diel periods for winter-fall (χ^2=12.3, p=0.0063), with dawn differing significantly from every other diel period (dawn-day: p=0.007, dawn-dusk: p=0.0363, dawn-night: p=0.02×10^{-4}). No significant differences were found between diel periods when spring and summer data sets were combined.

4.5 Discussion

The significant differences found in seasonal Pacific white-sided dolphin communication activity align with other literature on Pacific white-sided dolphin seasonality, which found seasonal north-south movement in Washington, Oregon, and California (Leatherwood *et al.*, 1984; Green *et al.*, 1992; Green *et al.*., 1993; Forney *et al.*, 1995; Forney & Barlow, 1998), and inshore-offshore seasonal movement in British Columbia and California (Stacey & Baird, 1991). The seasonal pattern identified here differed from the seasonal pattern in the ambient noise data, which were weakly correlated with Pacific white-sided dolphin vocal activity and showed different directions of correlation in different bands, suggesting that masking effects were not a strong explanatory factor for the patterns found here. Morton (2000) found peak encounter rates of Pacific white-sided dolphins in the Inside Passage in November, with dolphin presence increasing abruptly in October and remaining relatively high throughout the winter, tapering off as spring approached. When combined with the data presented here, which found peak vocal activity rates at Barkley Canyon in the spring and summer, a sudden drop in vocal activity in the fall, and a relatively low rate of vocal activity throughout the winter, this suggests that a large portion of the Pacific white-sided dolphins present on the continental shelf west of Vancouver Island in the spring and summer may migrate inshore during fall and overwinter in coastal waters. However, it is worth noting that Pacific white-sided dolphin distribution at the population scale has been linked to sea surface temperature, even when controlling for other temperature-dependent factors like prey (Selzer & Paine, 1988; Salvadeo *et al.*., 2010). An anomalous mass of warm surface water in the North Pacific Ocean basin, referred to as "The Blob," was discovered in 2013 and persisted into 2016, making 2014 an unusually warm year in this region with above average sea surface temperature (Peterson *et al.*, 2015; Peterson *et al.*, 2016). Seasonal patterns illustrated here may not be representative of years with more typical oceanographic conditions.

The increased Pacific white-sided dolphin vocal activity in the spring and summer at Barkley Canyon aligns with seasonal peaks in euphausiids (Robinson & Ware, 1994; Mackas *et al.*, 1997; Thomson & Allen, 2000), and the associated seasonal peak in Pacific herring (*Clupea pallasi*) and Pacific hake (*Merluccius productus*), which have been demonstrated to be the dominant fish species during spring and summer off southwest Vancouver Island (e.g., Tanasichuk *et al.*, 1991; Robinson & Ware, 1994; Ware & McFarlane, 1995; Mackas *et al.*, 1997; Tanasichuk, 1997; Robinson, 2000). While Pacific white-sided dolphins in the Inside Passage were observed feeding primarily on herring (Heise, 1996; Morton, 2000), Pacific white-sided dolphins are thought to be generalists based on data collected from pelagic regions, and stomach contents from Pacific white-sided dolphins inhabiting the open Pacific often contained hake and other benthic or deep water vertical migrators (Heise, 1996). The non-invasive prey-collection technique used in the Inside Passage by both Morton (2000) and Heise (1996) is limited to daylight hours and would not reflect the presence of diel migrators in Pacific white-sided dolphin diet. It is therefore likely that the large concentration of hake at Barkley Canyon in the spring and summer months draws Pacific white-sided dolphins to the region and comprises a significant portion of their spring and summer diet. While only communication signals were analyzed here, they were often seen in the presence of echolocation clicks, which are used primarily for navigation and foraging (Janik, 2009). Barkley Canyon is a very productive site that draws many species in high trophic levels, so it seems likely that dolphins are foraging there at least some of the time they are present. Heise (1996) found individual foraging activity in the Inside Passage to be very rare, and frequently observed co-ordinated efforts of dolphins that corralled fish schools while foraging. This behaviour is thought to suffocate fish trapped in the oxygen-reduced centre of the school, increasing foraging efficiency by rendering active pursuit of alert fish unnecessary (McFarland & Moss, 1967; Heise, 1996). Communication may play a role in co-ordinating their movements when foraging co-operatively (Van Opzeeland *et al.*, 2005; Vaughn-Hirshorn *et al.*, 2012; Eskelinen *et al.*, 2016). While exclusive foraging is the only non-rest behaviour state during which Pacific white-sided dolphins in Southern California produced pulsed calls infrequently, pulsed calls were very common during mixed foraging (Henderson *et al.*., 2011). Additionally, the role communication plays in foraging may vary regionally due to differences in prey type and foraging strategies (Black, 1994; Heise, 1996; Van Opzeeland *et al.*, 2005). Even if vocal activity does not necessarily indicate foraging, it does indicate wakefulness (Goley, 1999); given that Pacific white-sided dolphins have few local predators and are able to modify their diel patterns to match external patterns of importance (Therrien *et al.*, 2012), their diel rhythms probably align with those of their prey.

Pacific white-sided dolphins were vocally active during all diel periods in the spring and summer. Schooling fish are present at the surface during all hours in the spring and summer at Barkley Canyon, with Pacific hake migrating from mid-waters to the surface at night to feed on migrating euphausiids, and Pacific herring aggregations remaining near the surface throughout the day (Thomson & Allen, 2000). The abrupt drop in Pacific white-sided dolphin vocal activity, and associated change from consistent vocal activity to exhibiting a primarily nocturnal diel pattern, is likely related to the fall depletion of euphausiids, which constitute a significant proportion of the Pacific herring and Pacific hake diets (Tanasichuk *et al.*, 1991; Ware & McFarlane, 1995; Mackas *et al.*, 1997; Robinson, 2000; Ianson & Flostrand, 2010). Pacific hake seek other sources of food after the euphausiid decline (Thomson & Allen, 2000), and turn to herring as an alternate food source — adult herring biomass in the region starkly declines in late summer and early fall due to predation by hake (Tanasichuk *et al.*., 1991; Robinson & Ware, 1994). The migratory hake population also declines rapidly in early fall and moves offshore or disperses in search of other food sources (Robinson & Ware, 1994; Thomson & Allen, 2000). Herring may also move off the shelf break and closer to shore during the fall in preparation for their migration to coastal regions where they spawn in early winter and spring (Taylor, 1964; Hay, 1985; Tanasichuk, 1997; Ianson & Flostrand, 2010). Given the sudden decline in overall Pacific white-sided dolphin vocal activity in the fall and possible inshore-offshore seasonal movement described above, it seems likely that groups of Pacific white-sided dolphins feeding on herring during the spring and summer follow them inshore when they leave the canyon after the euphausiid decline (Figure 4.5). The seasonal changes in diurnal vocal activity and lack of seasonal change in nocturnal vocal activity further support the notion that a diurnal group moves into and out of the region seasonally.

Dolphins that remain at Barkley Canyon after the euphausiid decline and associated dispersion of fish prey sources are likely feeding on other diel migrators, as their vocal activity peaks at dusk and night, and sharply declines at dawn. Vermeer *et al.* (1989) found several species of squids in the stomachs of fulmars collected nearby, suggesting that squids may be a readily available prey source for overwintering Pacific white-sided dolphins. Along with squids, Heise (1996) found shrimp and sablefish (*Anaplopoma fimbria*) in the stomach contents of Pacific white-sided dolphins stranded in British Columbia, both of which may also be diel migrators (Barr, 1970; Kikuchi & Omori, 1985; Beamish & McFarlane, 1988; Sogard & Olla, 1998; Chapman, *et al.*, 2012; Goetz *et al.*, 2018). While Pacific white-sided dolphins sometimes eat deep-water and benthic species, and the sea floor at Barkley Canyon Upper Slope is well within their diving range, Pacific white-sided dolphins typically take shallow foraging dives regardless of habitat depth and prefer to wait for diel migrators to approach the surface rather than dive deeper to catch deep-water and benthic species (Black, 1994; Heise, 1996). Dolphins overwintering at Barkley Canyon may be subsisting solely on squids or may include partial diel migrators, such as sablefish, in their diet as well.

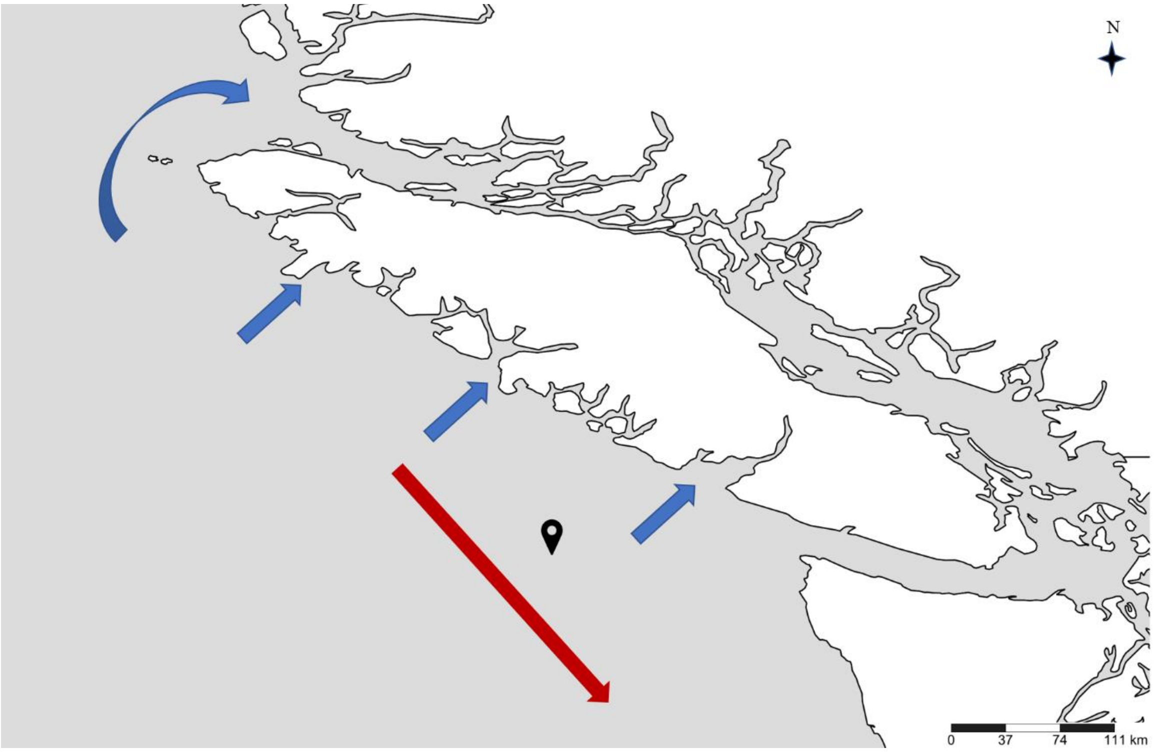


Figure 4.5 Proposed fall/winter migrations of two groups of Pacific white-sided dolphins based on seasonal changes to diel patterns in vocal activity at Barkley Canyon (location icon). It is hypothesized that the group represented by the red arrow are of the California/Oregon/Washington population feeding nocturnally on diel migrators, and the group represented by blue arrows are of a day-active Canada/Alaska population following the Pacific herring.

It is not possible to tell with certainty from the data analyzed here whether the same dolphins are feeding in all diel periods in the spring and summer, or there exists a group feeding in the day on schooling fish and another group feeding at night on diel migrators. There is a precedent for using acoustics to answer this question – Soldevilla *et al.* (2010) identified two unique Pacific white-sided dolphin click-types with different associated diel patterns and geographic distributions in the Southern California Bight. Type A click bouts were most frequent during dusk and night, with some daytime activity, and were present across all sites, while Type B click bouts were most frequent during the day and present only at the two southernmost sites. The two click types also exhibited different seasonal distributions. Their data suggest two distinct populations of Pacific white-sided dolphins in the region, and are corroborated by morphological evidence (Walker *et al.*., 1984; Walker *et al.*., 1986). Walker *et al.* (1986) found a southern population inhabiting waters from Mexico to Southern California, and a northern population inhabiting Northern California to Washington and possibly beyond, overlapping in the Southern California Bight. This also suggests that Pacific white-sided dolphins may not be as generalist as assumed thus far and are instead divided into different populations with distinct prey-type preferences. Indeed, the diet of dolphins in different regions has been demonstrated to differ, even when a species making up much of the diet of dolphins in one region is present in a different region where dolphins do not predate that food item (Heise, 1996). Given the similar diel vocal behaviour pattern and morphological evidence of distinct populations, it seems likely that the overwintering dolphins at Barkley Canyon are of this same northern population identified acoustically by Soldevilla *et al.* (2010).

While it is possible that the dolphins at Barkley Canyon are simply generalists, feeding on diel migrators year-round and taking advantage of the herring schools near the surface during the day in spring and summer, there may be a third population of Pacific white-sided dolphins in the Northeast Pacific Ocean that accounts for these seasonal diel patterns. While their Canada/Alaska sample was small and there are better methods for evaluating population structure available today (e.g. Hayano *et al.*, 2004), preliminary mitochondrial DNA analysis presented in Lux *et al.* (1997) suggested distinct populations in Baja California, California/Washington/Oregon, and Canada/Alaska. Lux *et al.* (1997) did not find genetic differences between inshore and offshore dolphins, so it is likely that shifting population densities inshore/offshore in different seasons are due to migration of one population rather than the existence of separate coastal and offshore populations. Both the Baja California and California/Oregon/Washington Pacific white-sided dolphin population distributions are shifting northward with warming sea surface temperatures (Salvadeo *et al.*., 2010), so it is possible that the California/Washington/Oregon population is present year-round at Barkley Canyon now. If the evidence of different populations preferring different prey in Southern California can be used as an indication that the Canada/Alaska and California/Washington/Oregon populations also differ in their prey preference and diel behaviour (Walker *et al.*., 1984; Walker *et al.*., 1986; Soldevilla *et al.*, 2010, Henderson *et al.*., 2011), then it seems likely that the nocturnal Pacific white-sided dolphins observed at Barkley Canyon are of the California/Washington/Oregon population, and a day-active Canada/Alaska population feeding primarily on schooling fish near the surface exists. This population may follow the migratory Pacific herring, feeding offshore during the spring and summer and inshore during the fall and winter. This hypothesis fits with seasonal patterns observed in British Columbia which suggest an inshore-offshore seasonal migration (Stacey & Baird, 1991; Heise, 1996; Morton, 2000), as well as the observation that Pacific white-sided dolphins overwintering in the Inside Passage feed almost exclusively on herring (Heise, 1996; Morton, 2000). This hypothesis also aligns with the sparse Alaskan data suggesting a peak in inshore Pacific white-sided dolphin presence in the spring (Leatherwood *et al.*, 1984; Dahlheim *et al.*., 1994; Dahlheim *et al.*., 2009), since herring in more northern latitudes spawn later, and would be present inshore in Alaska later than in British Columbia as a result (Hey, 1985). This hypothesis has implications for species management if true as the sizes of these two populations are unknown, their effective sizes could be significantly smaller than abundance estimates suggest if the populations do not interbreed, and they would have different requirements for habitat and resource management. Unfortunately, this hypothesis could not be investigated acoustically with the data presented here, as it has been in Southern California, because the bandwidth of collected data was not broad enough to include the full Pacific white-sided dolphin click bandwidth and distinguish between click types. This analysis considered data with a useable bandwidth of 1 Hz to 32 kHz, whereas Soldevilla *et al.* (2010) found differences between Pacific white-sided dolphin click types in the 25 kHz to 40 kHz band.

4.6 Conclusions

Pacific white-sided dolphins at Barkley Canyon exhibited a strong seasonal pattern in vocal presence, with most vocal activity occurring in the spring and summer, and vocal presence dropping dramatically in the fall. This seasonal pattern aligns with that of Pacific hake and Pacific herring, which form large aggregations at Barkley Canyon to feed on the spring and summer euphausiid population and abruptly disappear or disperse when the euphausiid population has been depleted, strongly suggesting that these fish species are a significant part of the Pacific white-sided dolphin diet in this region. When looked at in context of the seasonal inshore presence of Pacific white-sided dolphins in British Columbia, the seasonal presence of dolphins on the shelf break in the spring and summer suggests an inshore-offshore seasonal migration for this population of dolphins.

Diel patterns were not found in the spring and summer, with dolphins remaining vocal through all diel periods. In the fall and winter, Pacific white-sided dolphins were most active at dusk and night, with the lowest amount of vocal activity at dawn. This suggests that overwintering dolphins in this region could be feeding primarily on diel migrators, such as squids and sablefish. While it is possible that dolphins in this region are simply opportunistic foragers, feeding on diel migrators in the winter and both diel migrators and aggregations of Pacific herring that remain near the surface during the spring and summer, the nocturnal and daytime active dolphins may be of different populations, as has been found with the two Pacific white-sided dolphin populations overlapping in the Southern California Bight. The exclusive presence of daytime-active Pacific white-sided dolphins in spring and summer while Pacific herring and Pacific hake aggregate in the region, and the sudden switch to nocturnal vocal patterns and simultaneous decrease in overall Pacific white-sided vocal activity in the fall, seem to suggest that there may be two different populations of Pacific white-sided dolphins in British Columbia that overlap at Barkley Canyon in the spring and summer. The nocturnal vocal behaviour of the overwintering dolphins matches that described for the California/Washington/Oregon population of Pacific white-sided dolphins. While a

Canada/Alaska population has been suggested in the literature, it has not been conclusively identified and its diel vocal behaviour has not been investigated. If this population does exist, it may be the population exhibiting inshore-offshore seasonal movement previously described in the literature, and following the migratory Pacific herring population.

Future acoustic research on this topic should focus on echolocation clicks to determine whether two distinct populations exist in Barkley Canyon and the surrounding region, as hypothesized here, by looking for distinct click types with associated diel and seasonal patterns. If different click types are found, collocated click-type and prey-type analyses north of Vancouver Island or in the Inside Passage and south of Barkley Canyon between California and Washington could be used to confirm the attribution of different click types to a diurnal fish-eating Canada/Alaska population and a nocturnal California/Oregon/Washington population predating diel migrators. Additional genetic analysis with larger sample sizes and more modern techniques than were used in previous research should also be conducted to determine whether these populations are genetically isolated, as this would have implications for species management.

Chapter 5 Conclusions

The research in this thesis sought to evaluate long- and short-term temporal patterns in Pacific white-sided dolphin communication, using a random forest classifier to identify Pacific white-sided dolphin pulsed calls, in one year of acoustic data collected at Ocean Networks Canada’s Barkley Canyon node. During model selection using nested 10-fold cross-validation, the performance of random forest classifiers with different minimum leaf size, confidence threshold, number of trees, data balancing, and binary *versus* multiclass condition did not differ significantly. A classifier with default parameters was trained using the unbalanced, multiclass data set and performed well for classifying the 5 classes it was trained to distinguish based on un-nested 10-fold cross-validation, with F1-scores exceeding 0.79 for all mammal classes (humpback whale, orca, sperm whale, Pacific white-sided dolphins). The final model had precision, recall, and F1-score of 0.7903, 0.7986, and 0.7906, respectively, when used to classify Pacific white-sided dolphin calls.

The classifier was used to identify files containing Pacific white-sided dolphin pulsed calls. The files flagged by the classifier were validated through manual spectrogram analysis to produce a data set of Pacific white-sided dolphin presence/absence at 5-minute resolution. This data set was evaluated for seasonal and diel patterns in Pacific white-sided dolphin presence at Barkley Canyon. Pacific white-sided dolphin vocal activity was found to be highest in the spring and summer, with a sudden decrease in the fall, and remained relatively low throughout the winter. The fall decrease in Pacific white-sided dolphin activity was associated with a change in the diel pattern of their vocal behaviour, with dolphins being vocally active during all diel periods in the spring and summer, and primarily active only during dusk and night in the fall and winter. This change occurs around the time that Pacific herring in this region begin to migrate inshore, and just before the presence of Pacific white-sided dolphins dramatically increases in coastal waters (Morton *et al.*., 2000). Given that herring are present near the surface during the day at Barkley Canyon (Thomson & Allen, 2000), and that the dolphins that appear in inshore waters in the fall feed on herring almost exclusively (Morton, 2000), the patterns found here in Pacific white-sided dolphin vocal presence suggest that there may be a daytime-active population that follows the migratory herring, possibly the Canada/Alaska population suggested by Lux *et al.*. (1997), as well as the nocturnal California/Oregon/Washington population described as feeding on diel migrators by Soldevilla *et al.* (2010) that appears to be present at Barkley Canyon year-round. Future research on this topic could look for diel patterns in different click types at various locations, similar to Soldevilla *et al.* (2010), for further acoustic evidence of different Pacific white-sided dolphin populations overlapping in Barkley Canyon. Additional genetic analysis is also suggested to confirm different populations.

Bibliography

Allen, S. E., Vindeirinho, C., Thomson, R. E., Foreman, M. G. G., & Mackas, D. L. (2001). Physical and biological processes over a submarine canyon during an upwelling event. *Canadian Journal of Fisheries and Aquatic Sciences, 58,* 671-684.

Alpaydin, E. (2010). *Introduction to Machine Learning*. MIT Press.

Andrews, K. S., Williams, G. D., Farrer, D., Tolimieri, N., Harvey, C. J., Bargmann, G., & Levin, P. S. (2009). Diel activity patterns of sixgill sharks, *Hexanchus griseus*: the ups and downs of an apex predator. *Animal Behaviour*, *78*(2), 525-536.

Amano, M., & Yoshioka, M. (2003). Sperm whale diving behavior monitored using a suction-cup-attached TDR tag. *Marine Ecology Progress Series*, *258*, 291-295.

Au, W. W., Andersen, L. N., Rasmussen, A. R., Roitblat, H. L., & Nachtigall, P. E. (1995). Neural network modeling of a dolphin's sonar discrimination capabilities. *The Journal of the Acoustical Society of America*, *98*(1), 43-50.

Baird, R. W., Ligon, A. D., Hooker, S. K., & Gorgone, A. M. (2001). Subsurface and nighttime behaviour of pantropical spotted dolphins in Hawai'i. *Canadian Journal of Zoology*, *79*(6), 988-996.

Balmer, B. C., Schwacke, L. H., & Wells, R. S. (2010). Linking dive behavior to satellite-linked tag condition for a bottlenose dolphin (*Tursiops truncatus*) along Florida's northern Gulf of Mexico coast. *Aquatic Mammals*, *36*(1), 1.

Barkley, Y., Oswald, J. N., Carretta, J. V., Rankin, S., Rudd, A., & Lammers, M. O. (2011). *Comparison of real-time and post-cruise acoustic species identification of dolphin whistles using ROCCA (real-time odontocete call classification algorithm).* (Report No. NOAA-TM-NMFS-SWFSC-473). La Jolla, CA: National Oceanic and Atmospheric Administration.

Barr, L. (1970). Diel vertical migration of Pandalus borealis in Kachemak Bay, Alaska. *Journal of the Fisheries Board of Canada*, *27*(4), 669-676.

Baumann-Pickering, S., Roch, M. A., Wiggins, S. M., Schnitzler, H. U., & Hildebrand, J. A. (2015). Acoustic behavior of melon-headed whales varies on a diel cycle. *Behavioral Ecology and Sociobiology*, *69*(9), 1553-1563.

Beamish, R. J., & McFarlane, C. A. (1988). Resident and dispersal behavior of adult sablefish (*Anaplopoma fimbria*) in the slope waters off Canada's west coast. *Canadian Journal of Fisheries and Aquatic Sciences*, *45*(1), 152-164.

Beier, P., Choate, D., & Barrett, R. H. (1995). Movement patterns of mountain lions during different behaviors. *Journal of Mammalogy*, *76*(4), 1056-1070.

Binder, C. M., & Hines, P. C. (2014). Automated aural classification used for inter-species discrimination of cetaceans. *The Journal of the Acoustical Society of America*, *135*(4), 2113-2125.

Black, N. A. (1994). *Behavior and ecology of Pacific white-sided dolphins (Lagenorhynchus obliquidens) in Monterey Bay, California* Master's thesis, San Francisco State University.

Braga-Neto, M., & Dougherty, E. R. (2004). Is cross-validation valid for small-sample microarray classification? *Bioinformatics*, *20*(3), 374-380.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Briggs, F., Lakshminarayanan, B., Neal, L., Fern, X. Z., Raich, R., Hadley, S. J. K., Hadley A. S., & Betts, M. G. (2012). Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach. *The Journal of the Acoustical Society of America*, *131*(6), 4640-4650.

Burtenshaw, J. C., Oleson, E. M., Hildebrand, J. A., McDonald, M. A., Andrew, R. K., Howe, B. M., & Mercer, J. A. (2004). Acoustic and satellite remote sensing of blue whale seasonality and habitat in the Northeast Pacific. *Deep Sea Research Part II: Topical Studies in Oceanography*, *51*(10), 967-986.

Byvatov, E., Fechner, U., Sadowski, J., & Schneider, G. (2003). Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. *Journal of Chemical Information and Computer Sciences, 43*(6), 1882-1889.

Cerchio, S., Collins, T., Strindberg, S., Bennett, C., & Rosenbaum, H. (2010). Humpback whale singing activity off northern Angola: an indication of the migratory cycle, breeding habitat and impact of seismic surveys on singer number in Breeding stock B1. *International Whaling Commission Scientific Committee, Agadir, Morocco*.

Chapman, B. B., Skov, C., Hulthén, K., Brodersen, J., Nilsson, P. A., Hansson, L. A., & Brönmark, C. (2012). Partial migration in fishes: definitions, methodologies and taxonomic distribution. *Journal of Fish Biology*, *81*(2), 479-499.

Clapham, P. J., Leatherwood, S., Szczepaniak, I., & Brownell, R. L. (1997). Catches of humpback and other whales from shore stations at Moss Landing and Trinidad, California, 1919–1926. *Marine Mammal Science*, *13*(3), 368-394.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273-297.

Dahl, P. H., Miller, J. H., Cato, D. H., & Andrew, R. K. (2007). Underwater ambient noise. *Acoustics Today*, *3*(1), 23-33.

Dahlheim, M. E., & Towell, R. G. (1994). Occurrence and distribution of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*) in southeastern Alaska, with notes on an attack by killer whales (*Orcinus orca*). *Marine Mammal Science*, *10*(4), 458-464.

Dahlheim, M. E., White, P. A., & Waite, J. M. (2009). Cetaceans of Southeast Alaska: distribution and seasonal occurrence. *Journal of Biogeography*, *36*(3), 410-426.

Davis, J., & Goadrich, M. (2006, June). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning*, 233-240.

Deconto, L. S., & Monteiro-Filho, E. L. (2016). Day and night sounds of the Guiana dolphin, *Sotalia guianensis* (*Cetacea: Delphinidae*) in southeastern Brazil. *Acta Ethologica*, *19*(1), 61-68.

Dede, A., Öztürk, A. A., Akamatsu, T., Tonay, A. M., & Öztürk, B. (2014). Long-term passive acoustic monitoring revealed seasonal and diel patterns of cetacean presence in the Istanbul Strait. *Journal of the Marine Biological Association of the United Kingdom*, *94*(06), 1195-1202.

Deecke, V. B., Ford, J. K., & Spong, P. (1999). Quantifying complex patterns of bioacoustic variation: use of a neural network to compare killer whale (*Orcinus orca*) dialects. *The Journal of the Acoustical Society of America*, *105*(4), 2499-2507.

Deecke, V. B., & Janik, V. M. (2006). Automated categorization of bioacoustic signals: avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America*, *119*(1), 645-653.

Díaz-Uriarte, R., & De Andrés, S. A. (2006). Gene selection and classification of microarray data using random forest. *BioMed Central Bioinformatics*, *7*(1), 3.

Dietrich, R., Opper, M., & Sompolinsky, H. (1999). Statistical mechanics of support vector networks. *Physical Review Letters*, *82*(14), 2975.

Drummond, C., & Holte, R. C. (2003, August). C4. 5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. In *Workshop on Learning from Imbalanced Datasets II* (Vol. 11).

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, *7*(1), 1-26.

Efron, B. (1983). Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American Statistical Association, 78*(382), 316-331.

Efron, B., & Tibshirani, R. (1997). Improvements on cross-validation: the .632+ bootstrap method. *Journal of the American Statistical Association, 92* (438), 548-560.

Esfahanian, M., Zhuang, H., Erdol, N., & Gerstein, E. (2015, August). Comparison of two methods for detection of North Atlantic Right Whale upcalls. In *2015 23rd European Signal Processing Conference (EUSIPCO)*, 559-563.

Eskelinen, H. C., Winship, K. A., Jones, B. L., Ames, A. E. M., Kuczaj, S. A. II (2016). Acoustic behaviour associated with cooperative task success in bottlenose dolphins (*Tursiops truncates*). *Animal Cognition, 19,* 789-797.

Ford, J. K. (1991). Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Canadian Journal of Zoology*, *69*(6), 1454-1483.

Forney, K. A., & Barlow, J. (1998). Seasonal patterns in the abundance and distribution of California cetaceans, 1991–1992. *Marine Mammal Science*, *14*(3), 460-489.

Forney, K. A., Barlow, J., & Carretta, J. V. (1995). The abundance of cetaceans in California waters. Part 2: Aerial surveys in winter and spring of 1991 and 1992. *Fishery Bulletin*, *93*(1), 15-26.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The Elements of Statistical Learning* (Vol. 1). Springer, Berlin: Springer series in statistics.

Gaetz, W., Jantzen, K., Weinberg, H., Spong, P., & Symonds, H. (1993, October). A neural network method for recognition of individual *Orcinus orca* based on their acoustic behaviour: Phase 1. In *OCEANS'93. Engineering in Harmony with Ocean. Proceedings* (pp. I455-I457). IEEE.

Geertsen, B. M., Teilmann, J., Kastelein, R. A., Vlemmix, H. N. J., & Miller, L. A. (2004). Behaviour and physiological effects of transmitter attachments on a captive harbour porpoise (*Phocoena phocoena*). *Journal of Cetacean Research and Management*, *6*(2), 139-146.

Geise, L., Gomes, N., & Cerqueira, R. (1999). Behaviour, habitat use and population size of *Sotalia fluviatilis* (Gervais, 1853) (*Cetacea, Delphinidae*) in the Cananéia estuary region, São Paulo, Brazil. *Revista Brasileira de Biologia*, *59*(2), 183-194.

Gillespie, D., Caillat, M., & Gordon, J. (2013). Automatic detection and classification of odontocete whistles. *The Journal of the Acoustical Society of America*, *134*(3), 2427-2437.

Goetz, F. W., Jasonowicz, A. J., & Roberts, S. B. (2018). What goes up must come down: Diel vertical migration in the deep-water sablefish (*Anoplopoma fimbria*) revealed by pop-up satellite archival tags. *Fisheries Oceanography*, *27*(2), 127-142.

Goley, P. D. (1999). Behavioural aspects of sleep in Pacific white-sided dolphins (*Lagenorhynchus obliquidens,* Gill 1965). *Marine Mammal Science, 15*(4), 1054-1064.

Goold, J. C. (2000). A diel pattern in vocal activity of short-beaked common dolphins, *Delphinus delphis*. *Marine Mammal Science*, *16*(1), 240-244.

Green, G. A., Brueggeman, J. J., Grotefendt, R. A., Bowlby, C. E., Bonnell, M. L., & Balcomb, K. C. III. (1992). Chapter I: Cetacean distribution and abundance off Oregon and Washington, 1989-90. *Oregon and Washington: Marine Mammal and Seabird Survey* (pp. xii-100). Los Angeles, California: Pacific OCS Region, Minerals Management Service, US Dept. of the Interior.

Green, G. A., Grotefendt, R. A., Smultea, M. A., Bowlby, C. E., & Rowlett, R. A. (1993). *Delphinid aerial surveys in Oregon and Washington offshore waters*. (Report No. 50ABNF200058). Seattle, WA: National Marine Mammal Laboratory, National Marine Fisheries Service.

Hanson, M. T., & Defran, R. H. (1993). The behavior and feeding ecology of the Pacific coast bottlenose dolphin, *Tursiops truncatus*. *Aquatic Mammals*, *19*, 127-127.

Hannay, D. E., Delarue, J., Mouy, X., Martin, B. S., Leary, D., Oswald, J. N., & Vallarta, J. (2013). Marine mammal acoustic detections in the northeastern Chukchi Sea, September 2007–July 2011. *Continental Shelf Research*, *67*, 127-146.

Hay, D. E. (1985). Reproductive biology of Pacific herring (*Clupea harengus pallasi*). *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(S1), s111-s126.

Hayano, A., Yoshioka, M., Tanaka, M., & Amano, M. (2004). Population differentiation in the Pacific white-sided dolphin *Lagenorhynchus obliquidens* inferred from mitochondrial DNA and microsatellite analyses. *Zoological Science*, *21*(9):989-999.

Hays, G. C. (2003). A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. In *Migrations and Dispersal of Marine Organisms* (pp. 163-170). Springer Netherlands.

Heise, K. A. (1996). *Life history parameters of the Pacific white-sided dolphin (Lagenorhynchus obliquidens) and its diet and occurrence in the coastal waters of British Columbia* (Master's thesis, University of British Columbia).

Helbe, T. A., D'Spain, G. L., Campbell, G. S., & Hildebrand, J. A. (2013). Calibrating passive acoustic monitoring: correcting humpback whale call detections for site-specific and time-dependent environmental characteristics. *The Journal of the Acoustical Society of America*, *134*(5), EL400-EL406.

Henderson, E. E., Hildebrand, J. A., & Smith, M. H. (2011). Classification of behavior using vocalizations of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*) a. *The Journal of the Acoustical Society of America*, *130*(1), 557-567.

Hay, D. E. (1985). Reproductive biology of Pacific herring (*Clupea harengus pallasi*). *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(S1), s111-s126.

Hodge, L. E., Bell, J. T., Kumar, A., & Read, A. J. (2013). The influence of habitat and time of day on the occurrence of odontocete vocalizations in Onslow Bay, North Carolina. *Marine Mammal Science*, *29*(4), E411-E427.

Ianson, D., & Flostrand, L. (2010). Ecosystem status and trends report: coastal waters off the west coast of Vancouver Island, British Columbia. *Fisheries & Oceans Canada*, CSAS 2010/046

Jacobson, E. K., Yack, T. M., & Barlow, J. (2013). *Evaluation of an automated acoustic beaked whale detection algorithm using multiple validation and assessment methods.* (Report No. NOAA-TM-NMFS-SWFSC-509). La Jolla, CA: National Ocean and Atmospheric Administration.

Janik, V. M. (2009). Acoustic communication in delphinids. In M. Naguib, V. M. Janik, K. Zuberbühler & N. S. Clayton (Eds.), *Advances in the Study of Behaviour Volume 40* (pp. 123-157). New York, NY: Elsevier Inc.

Johnston, D. W., McDonald, M., Polovina, J., Domokos, R., Wiggins, S., & Hildebrand, J. (2008). Temporal patterns in the acoustic signals of beaked whales at Cross Seamount. *Biology Letters*, *4*(2), 208-211.

Kertész, C. (2016). Rigidity-Based Surface Recognition for a Domestic Legged Robot. *IEEE Robotics and Automation Letters, 1*(1), 309-315.

Kikuchi, T., & Omori, M. (1985). Vertical distribution and migration of oceanic shrimps at two locations off the Pacific coast of Japan. *Deep Sea Research Part A. Oceanographic Research Papers*, *32*(7), 837-851.

Kim, J-H. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics and Data Analysis, 53*(2009), 3735-3745.

Klatsky, L. J., Wells, R. S., & Sweeney, J. C. (2007). Offshore bottlenose dolphins (*Tursiops truncatus*): Movement and dive behavior near the Bermuda Pedestal. *Journal of Mammalogy*, *88*(1), 59-66.

Klinck, H., Nieukirk, S. L., Mellinger, D. K., Klinck, K., Matsumoto, H., & Dziak, R. P. (2012). Seasonal presence of cetaceans and ambient noise levels in polar waters of the North Atlantic. *The Journal of the Acoustical Society of America*, *132*(3), EL176-EL181.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence, 14*(2), 1137-1145.

Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, *30*, 25-36.

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics 6*(1), 10.

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., & Ward, J. (2011). Cetacean population density estimation from single fixed sensors using passive acoustics. *The Journal of the Acoustical Society of America*, *129*(6), 3610-3622.

Leatherwood, S., Reeves, R. R., Bowles, A. E., Stewart, B. S., & Goodrich, K. R. (1984). Distribution, seasonal movements, and abundance of Pacific white-sided dolphins in the eastern North Pacific. *The Scientific Reports of the Whales Research Institute*, *35*, 129-157.

Lerczak, J. A., Shelden, K. E., & Hobbs, R. C. (2000). Application of suction-cup-attached VHF transmitters to the study of beluga, *Delphinapterus leucas*, surfacing behavior in Cook Inlet, Alaska. *Marine Fisheries Review*, *62*(3), 99-111.

Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*, *177*, 970-980.

Locascio, J. V., & Mann, D. A. (2008). Diel periodicity of fish sound production in Charlotte Harbor, Florida. *Transactions of the American Fisheries Society*, *137*(2), 606-615.

Loh, W. Y. (2002). Regression tress with unbiased variable selection and interaction detection. *Statistica Sinica*, *12*(2), 361-386.

Loh, W. Y., & Shih, Y. S. (1997). Split selection methods for classification trees. *Statistica sinica*, *7*(4), 815-840.

Lux, C. A. (1997). Mitochondrial DNA population structure of the Pacific white-sided dolphin. *Report of the International Whaling Commission*, *47*, 645-652.

Mackas, D. L., & Coyle, K. O. (2005). Shelf-offshore exchange processes, and their effects on mesozooplankton biomass and community composition patterns in the northeast Pacific. *Deep-Sea Research II, 52,* 707-725.

Mackas, D. L., Kieser, R., Saunders, M., Yelland, D. R., Brown, R. M., & Moore, D. F. (1997). Aggregation of euphausiids and Pacific hake (*Merluccius productus*) along the outer continental shelf off Vancouver Island. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(9), 2080-2096.

Madjarov, G., Kocev, D., Gjorgjevikj, D., & Džeroski, S. (2012). An extensive experimental comparison of methods for multi-label learning. *Pattern Recognition*, *45*(9), 3084-3104.

Marques, T. A., Munger, L., Thomas, L., Wiggins, S., & Hildebrand, J. A. (2011). Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research*, *13*(3), 163-172.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D. & Tyack, P. L. (2013). Estimating animal population density using passive acoustics. *Biological Reviews*, *88*(2), 287-309.

Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., & Tyack, P. L. (2009). Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. *The Journal of the Acoustical Society of America*, *125*(4), 1982-1994

McDonald, M. A., Hildebrand, J. A., Wiggins, S. M., Johnston, D. W., & Polovina, J. J. (2009). An acoustic survey of beaked whales at Cross Seamount near Hawaii. *The Journal of the Acoustical Society of America*, *125*(2), 624-627.

McFarland, W. N., & Moss S. A. (1967). Internal behaviour in fish schools. *Science*, *156*(3772), 260-262.

Meeus, J. H. (1991). *Astronomical Algorithms*. Willmann-Bell, Incorporated.

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, *7*(Jun), 983-999.

Mercado, E., & Kuh, A. (1998, May). Classification of humpback whale vocalizations using a self-organizing neural network. In *1998 IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36227)*, *2*, 1584-1589.

Moloney, J., Hillis, C., Mouy, X., Urazghildiiev, I., & Dakin, T. (2014). Autonomous Multichannel Acoustic Recorders on the VENUS Ocean Observatory. *Oceans-St. John's, 2014*, 1-6.

Moretti, D., Marques, T. A., Thomas, L., DiMarzio, N., Dilley, A., Morrissey, R., McCarthy, E., Ward, J. & Jarvis, S. (2010). A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a Mid-Frequency Active (MFA) sonar operation. *Applied Acoustics*, *71*(11), 1036-1042.

Morton, A. (2000) Occurrence, photo-identification, and prey of Pacific white-sided dolphins (*Lagenorhyncus obliquidens*) in the Broughton Archipelago, Canada 1984–1998. *Marine Mammal Science*, *16*(1), 80-93.

Mouy, X., Ford, J., Pilkington, J., Kanes, K., Riera, A., Dakin, T., & Mouy, P. A. (2015). Automatic Marine Mammal Monitoring off British Columbia, Canada. *7th International DCLDE Workshop*.

Mouy, X., Oswald, J., Leary, D., Delarue, J., Vallarta, J., Rideout, B., Mellinger, D., Erbe, C., Hannay, D., Martin, B. (2013). Passive acoustic monitoring of marine mammals in the Arctic. In O. Adam & F. Samaran (Eds.), *Detection, Classification and Localisation of Marine Mammals Using Passive Acoustics, 2003-2013: 10 Years of International Research* (pp. 185-224). Paris, France: Dirac NGO.

Munger, L. M., Wiggins, S. M., Moore, S. E., & Hildebrand, J. A. (2008). North Pacific right whale (*Eubalaena japonica*) seasonal and diel calling patterns from long-term acoustic recordings in the southeastern Bering Sea, 2000–2006. *Marine Mammal Science*, *24*(4), 795-814.

Murray, S. O., Mercado, E., & Roitblat, H. L. (1998). The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. *The Journal of the Acoustical Society of America*, *104*(6), 3626-3633.

McBride, A. F., & Hebb, D. O. (1948). Behavior of the captive bottle-nose dolphin, *Tursiops truncatus*. *Journal of Comparative and Physiological Psychology*, *41*(2), 111.

Nelson, B. V., & Vance, R. R. (1979). Diel foraging patterns of the sea urchin *Centrostephanus coronatus* as a predator avoidance strategy. *Marine Biology*, *51*(3), 251-258.

Ness, S. (2013). *The Orchive: A system for semi-automatic annotation and analysis of a large collection of bioacoustic recordings* Doctoral dissertation, University of Victoria.

Norris, T. F., Mc Donald, M., & Barlow, J. (1999). Acoustic detections of singing humpback whales (*Megaptera novaeangliae*) in the eastern North Pacific during their northbound migration. *The Journal of the Acoustical Society of America*, *106*(1), 506-514.

Oshima, I., Yamada, H., Goto, M., Sato, K., & Ebihara, S. (1989). Pineal and retinal melatonin is involved in the control of circadian locomotor activity and body temperature rhythms in the pigeon. *Journal of Comparative Physiology A*, *166*(2), 217-226.

Oswald, J. N., Rankin, S., & Barlow, J. (2008). To whistle or not to whistle? Geographic variation in the whistling behavior of small odontocetes. *Aquatic Mammals*, *34*(3), 288.

Peterson, W., Robert, M., & Bond, N. (2015). The warm blob-Conditions in the northeastern Pacific Ocean. *PICES Press*, *23*(1), 36.

Peterson, W., Bond, N., & Robert, M. (2016). The Blob (Part Three): Going, going, gone? *PICES Press*, *24*(1), 46.

Philpott, E., Englund, A., Ingram, S., & Rogan, E. (2007). Using T-PODs to investigate the echolocation of coastal bottlenose dolphins. *Journal of the Marine Biological Association of the United Kingdom*, *87*(01), 11-17.

Pirotta, E., Milor, R., Quick, N., Moretti, D., Di Marzio, N., Tyack, P., Boyd, I. & Hastie, G. (2012). Vessel noise affects beaked whale behavior: results of a dedicated acoustic response study. *PLoS One*, *7*(8), e42535.

Provost, F., Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In *Proceedings of the 15th International Conference on Machine Learning (IMLC-98)*, *98*, 445-453.

Rehn, N., Teichert, S., & Thomsen, F. (2007). Structural and temporal emission patterns of variable pulsed calls in free-ranging killer whales (*Orcinus orca*). *Behaviour*, *144*(3), 307-329.

Riera, A., Ford, J. K., & Ross Chapman, N. (2013). Effects of different analysis techniques and recording duty cycles on passive acoustic monitoring of killer whales. *The Journal of the Acoustical Society of America*, *134*(3), 2393-2404.

Robinson, C. L. (2000). The consumption of euphausiids by the pelagic fish community off southwestern Vancouver Island, British Columbia. *Journal of Plankton Research*, *22*(9), 1649-1662.

Robinson, C. L., & Ware, D. M. (1994). Modelling pelagic fish and plankton trophodynamics off southwestern Vancouver Island, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*, *51*(8), 1737-1751.

Rosenblatt, F. (1961). *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms* (Report No. VG-1196-G-8). Buffalo, NY: Cornell Aeronautical Laboratory.

Ross, J. C., & Allen, P. E. (2014). Random forest for improved analysis efficiency in passive acoustic monitoring. *Ecological Informatics*, *21*, 34-39.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-538.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. In J. L. McClelland & D. E. Rumelhart (Eds.) *Parallel Distributed Processing* (pp. 318-362). Cambridge, MA: MIT Press.

Saayman, G. S., Tayler, C. K., & Bower, D. (1973). Diurnal Activity Cycles in Captive and Free-Ranging Indian Ocean Bottlenose Dolphins (*Tursiops Ad Uncus Ehrenburg*). *Behaviour*, *44*(3), 212-233.

Salvadeo, C. J., Lluch-Belda, D., Gómez-Gallardo, A., Urbán-Ramírez, J., & MacLeod, C. D. (2010). Climate change and a poleward shift in the distribution of the Pacific white-sided dolphin in the northeastern Pacific. *Endangered Species Research*, *11*(1), 13-19.

Sánchez-Vázquez, F. J., & Tabata, M. (1998). Circadian rhythms of demand-feeding and locomotor activity in rainbow trout. *Journal of Fish Biology*, *52*(2), 255-267.

Schoenherr, J. R. (1991). Blue whales feeding on high concentrations of euphausiids around Monterey Submarine Canyon. *Canadian Journal of Zoology*, *69*(3), 583-594.
Scott, M. D., & Chivers, S. J. (2009). Movements and diving behavior of pelagic spotted dolphins. *Marine Mammal Science*, *25*(1), 137-160.

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2008, December). A comparative study of data sampling and cost sensitive learning. In *2008 IEEE International Conference Data Mining Workshops,* 46-52.

Selzer, L. A., & Payne, P. M. (1988). The distribution of white-sided (*Lagenorhynchus acutus*) and common dolphins (*Delphinus delphis*) vs. environmental features of the continental shelf of the northeastern United States. *Marine Mammal Science, 4*(2), 141-153.

Širović, A., Williams, L. N., Kerosky, S. M., Wiggins, S. M., & Hildebrand, J. A. (2013). Temporal separation of two fin whale call types across the eastern North Pacific. *Marine Biology*, *160*(1), 47-57.

Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., & Roch, M. A. (2008). Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks. *The Journal of the Acoustical Society of America*, *124*(1), 609-624.

Soldevilla, M. S., Wiggins, S. M., & Hildebrand, J. A. (2010). Spatio-temporal comparison of Pacific white-sided dolphin echolocation click types. *Aquatic Biology*, *9*(1), 49-62.

Sogard, S. M., & Olla, B. L. (1998). Behavior of juvenile sablefish, *Anoplopoma fimbria* (Pallas), in a thermal gradient: balancing food and temperature requirements. *Journal of Experimental Marine Biology and Ecology*, *222*(1–2), 43-58.

Sosa-López, J. R., & Mennill, D. J. (2014). The vocal behavior of the Brown-throated Wren (*Troglodytes brunneicollis*): song structure, repertoires, sharing, syntax, and diel variation. *Journal of Ornithology*, *155*(2), 435-446.

Stacey, P. J., & Baird, R. W. (1991). Status of the Pacific white-sided dolphin, *Lagenorhynchus obliquidens*, in Canada. *Canadian Field-Naturalist*, *105*(2), 219-232.

Stefánsson U., & Richards, F. A. (1963). Processes contributing to the nutrient distribution off the Columbia River and Strait of Juan de Fuca. *Limnology and Oceanography*, *8*(4), 394-410.

Tanasichuk, R. W. (1997). Influence of biomass and ocean climate on the growth of Pacific herring (*Clupea pallasi*) from the southwest coast of Vancouver Island. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(12), 2782-2788.

Tanasichuk, R. W., Ware, D. M., Shaw, W., & McFarlane, G. A. (1991). Variations in diet, daily ration, and feeding periodicity of Pacific hake (*Merluccius productus*) and spiny dogfish (*Squalus acanthias*) off the lower west coast of Vancouver Island. *Canadian Journal of Fisheries and Aquatic Sciences*, *48*(11), 2118-2128.

Taylor, F. H. C. (1964). Life history and present status of British Columbia herring stocks. *Canadian Bulletin of Fisheries and Aquatic Sciences, 143,* 81.

Therrien, S. C., Thomas, J. A., Therrien, R. E., & Stacey, R. (2012). Time of day and social change affect underwater sound production by bottlenose dolphins (*Tursiops truncatus)* at the Brookfield Zoo. *Aquatic Mammals*, *38*(1), 65.

Thomson, R. E., & Allen, S. E. (2000). Time series acoustic observations of macrozooplankton diel migration and associated pelagic fish abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, *57*(9), 1919-1931.

Turesson, H. K., Ribeiro, S., Pereira, D. R., Papa, J. P., & de Albuquerque, V. H. C. (2016). Machine learning algorithms for automatic classification of marmoset vocalizations. *PloS ONE, 11*(9), e0163041.

The Mathworks, Inc. (2016). MATLAB R2016a and Statistics and Machine Learning Toolbox Version 10.2. Natick, MA.

Todd, V. L., 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*, 734-745.

Tyack, P. L., Zimmer, W. M., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., Clarck, C. W., D'Amico, A., DiMarzio, N., Jarvis, S. & McCarthy, E. (2011). Beaked whales respond to simulated and actual navy sonar. *PloS ONE*, *6*(3), e17009.

Van Opzeeland, I. C., Corkeron, P. J., Leyssen, T., Similä, T., & Van Parijs, S. M. (2005). Acoustic behaviour of Norwegian killer whales, *Orcinus orca*, during carousel and seiner foraging on spring-spawning herring. *Aquatic Mammals*, *31*(1), 110-119.

Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BioMed Central Bioinformatics 7*(1), 91.

Vaughn-Hirshorn, R. L., Hodge, K. B., Würsig, B., Sappenfield, R. H., Lammers, M. O., & Dudzinski, K. M. (2012). Characterizing dusky dolphin sounds from Argentina and New Zealand. *The Journal of the Acoustical Society of America*, *132*(1), 498-506.

Verfuß, U. K., Honnef, C. G., Meding, A., Dähne, M., Mundry, R., & Benke, H. (2007). Geographical and seasonal variation of harbour porpoise (*Phocoena phocoena*) presence in the German Baltic Sea revealed by passive acoustic monitoring. *Journal of the Marine Biological Association of the United Kingdom*, *87*(1), 165-176.

Vermeer, K., Morgan, K. H., Smith, G. J., & Hay, R. (1989). Fall distribution of pelagic birds over the shelf off SW Vancouver Island. *Colonial Waterbirds*, *12*(2), 207-214.

Walker, W. A., Goodrich, K. R., Leatherwood, S., & Stroud, R. K. (1984) *Population biology and ecology of the Pacific white-sided dolphin, Lagenorhynchus obliquidens, in the northeastern Pacific. Part II. Biology and geographical variation.* (Report No. LJ-84-34c). National Marine Fisheries Service.

Walker, W. A., Leatherwood, S., Goodrich, K. R., Perrin, W. F., & Stroud, R. K. (1986). Geographical variation and biology of the Pacific white-sided dolphin, *Lagenorhynchus obliquidens*, in the north-eastern Pacific. In M. M. Bryden & R. J. Harrison (Eds.), *Research on Dolphins* (pp. 441-465). Oxford, England: Oxford University Press.

Walsh, P. B., & Inglis, J. M. (1989). Seasonal and diel rate of spontaneous vocalization in coyotes in south Texas. *Journal of Mammalogy*, *70*(1), 169-171.

Ware, D. M., & McFarlane, G. A. (1995). Climate induced changes in hake abundance and pelagic community interactions in the Vancouver Island Upwelling System. *Climate change and northern fish populations. Edited by RJ Beamish. Canadian Special Publication of Fisheries and Aquatic Sciences*, *121*, 509-521.

Weiss, G. M., McCarthy, K., & Zabar, B. (2007). Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs? *Data Mining*, *7*, 35-41.

Whitlock, M. C., & Schluter, D. (2009). *The Analysis of Biological Data*. Greenwood Village, Colorado: Roberts and Company Publishers.

Wiggins, S. M., Oleson, E. M., McDonald, M. A., & Hildebrand, J. A. (2005). Blue whale (*Balaenoptera musculus*) diel call patterns offshore of Southern California. *Aquatic Mammals*, *31*(2), 161.

Williams, R., Clark, C. W., Ponirakis, D., & Ashe, E. (2014). Acoustic quality of critical habitats for three threatened whale populations. *Animal Conservation*, *17*(2), 174-185.

Wu, Q., & Zhou, D. X. (2005). SVM soft margin classifiers: linear programming versus quadratic programming. *Neural Computation*, *17*(5), 1160-1187.

Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., & Coates, S. (2013). Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat. *The Journal of the Acoustical Society of America*, *134*(3), 2589-2595

Appendix A

Names and descriptions of features extracted by SpectroDetector (JASCO Applied Sciences) and used here for automated classification of Pacific white-sided dolphin pulsed calls. Table was provided by JASCO Applied Sciences.

ID	Name	Description	Frequency envelope			Time envelope		
			sum	max	median	sum	max	median
1	'Fpeak (Hz)'	Peak frequency	X					
2	'Fmedian (Hz)'	Median frequency	X					
3	'Fiqr (Hz)'	Inter-quartile range of frequency envelope	X					
4	'Fassym'	Assymetry of frequency envelope	X					
5	'Fconc (Hz)'	Concentration of frequency envelope	X					
6	'Fmin (Hz)'	Min frequency	X					
7	'Fmax (Hz)'	Max frequency	X					
8	'Fbandwidth (Hz)'	Frequency bandwidth (Fmax-Fmin)	X					
9	'FampDerivpct90'	90th percentile of the first derivative of the frequency envelope	X					
10	'Fkurt'	Kurtosis of the frequency envelope	X					
11	'Fvar'	Variance of the frequency envelope	X					
12	'FmedToPeak (Hz)'	absolute value of median frequency - peak frequency	X					
13	'FM (Hz)'	Hamonicity (i.e. lag of maxmum peak of the autocorrelation function)	X					
14	'FM CorrCoef'	"Strength" of harmonicity (i.e. value of maximum peak in autocorrelation function)	X					
15	'Contour Fmin (Hz)'	Min frequency of frequency contour						

16	'Contour Fmax (Hz)'	Max frequency of frequency contour						
17	'Contour Frange (Hz)'	Bandwidth of frequency contour						
18	'Contour downsweep (%)'	% of frequency contour going down in frequency						
19	'Contour upsweep (%)'	% of frequency contour going up in frequency						
20	'Contour flat (%)'	% of frequency contour staying the same in frequency						
21	'Contour upsweep range (Hz)'	cumulated bandwidth of upsweep portions of the frequency contour						
22	'Contour downsweep range (Hz)'	cumulated bandwidth of downweep portions of the frequency contour						
23	'Contour std'	Standard deviation of the frequency values of the frequency contour						
24	'Contour #inflec pts concave'	Number of concave inflection points in the frequency contour						
25	'Contour #inflec pts convex'	Number of convex inflection points in the frequency contour						
26	'Contour Slope mean (Hz/s)'	mean of the absolute value of the first derivative of the frequency contour						
27	'Contour Slope variance'	standard deviation of the absolute value of the first derivative of the frequency contour						
28	'Contour overall Slope (Hz/s)'	End frequency - start frequency of frequency contour / duration of frequency contour						
29	'Contour Mean - Overall slope (Hz/s)'	feature 26 - feature 28						
30	'Tiqr (s)'	Inter-quartile range of time envelope				X		
31	'Tassym'	Assymetry of time envelope				X		
32	'Tconc (s)'	Concentration of time envelope				X		
33	'Dur (s)'	Duration				X		

34	'Tderiv pct90'	90th percentile of the first derivative of the time envelope				X		
35	'Tvar'	Variance of the time envelope				X		
36	'Tkurt'	Kurtosis of the time envelope				X		
37	'TmedToPeak (s)'	absolute value of median time - peak time				X		
38	'AM (s)'	Amplitude modulation (i.e. lag of maxmum peak of the autocorrelation function)				X		
39	'AM CorrCoef'	"Strength" of amplitude modulation (i.e. value of maximum peak in autocorrelation function)				X		
40	'Mean Energy'	Mean energy (not calibrated)				X		
41	'Fpeak_maxenv (Hz)'	Peak frequency		X				
42	'Fmedian_maxenv (Hz)'	Median frequency		X				
43	'Fiqr_maxenv (Hz)'	Inter-quartile range of frequency envelope		X				
44	'Fassym_maxenv'	Assymetry of frequency envelope		X				
45	'Fconc_maxenv (Hz)'	Concentration of frequency envelope		X				
46	'Fmin_maxenv (Hz)'	Min frequency		X				
47	'Fmax_maxenv (Hz)'	Max frequency		X				
48	'Fbandwidth_maxenv (Hz)'	Frequency bandwidth (Fmax-Fmin)		X				
49	'FampDerivpct90_maxenv'	90th percentile of the first derivative of the frequency envelope		X				
50	'Fkurt_maxenv'	Kurtosis of the frequency envelope		X				
51	'Fvar_maxenv'	Variance of the frequency envelope		X				
52	'FmedToPeak_maxenv (Hz)'	absolute value of median frequency - peak frequency		X				
53	'FM_maxenv(Hz)'	Hamonicity (i.e. lag of maxmum peak of the autocorrelation function)		X				
54	'FM_maxenv CorrCoef'	"Strength" of harmonicity (i.e. value of maximum peak in autocorrelation function)		X				
55	'Tiqr_maxenv (s)'	Inter-quartile range of time envelope					X	
56	'Tassym_maxenv'	Assymetry of time envelope					X	

57	'Tconc_maxenv (s)'	Concentration of time envelope					X	
58	'Tderiv pct90_maxenv'	90th percentile of the first derivative of the time envelope					X	
59	'Tvar_maxenv'	Variance of the time envelope					X	
60	'Tkurt_maxenv'	Kurtosis of the time envelope					X	
61	'TmedToPeak_maxenv (s)'	absolute value of median time - peak time					X	
62	'AM_maxenv (s)'	Amplitude modulation (i.e. lag of maxmum peak of the autocorrelation function)					X	
63	'AM CorrCoef_maxenv'	"Strength" of amplitude modulation (i.e. value of maximum peak in autocorrelation function)					X	
64	'Fpeak_medenv (Hz)'	Peak frequency			X			
65	'Fmedian_medenv (Hz)'	Median frequency			X			
66	'Fiqr_medenv (Hz)'	Inter-quartile range of frequency envelope			X			
67	'Fassym_medenv'	Assymetry of frequency envelope			X			
68	'Fconc_medenv (Hz)'	Concentration of frequency envelope			X			
69	'Fmin_medenv (Hz)'	Min frequency			X			
70	'Fmax_medenv (Hz)'	Max frequency			X			
71	'Fbandwidth_medenv (Hz)'	Frequency bandwidth (Fmax-Fmin)			X			
72	'FampDerivpct90_medenv'	90th percentile of the first derivative of the frequency envelope			X			
73	'Fkurt_medenv'	Kurtosis of the frequency envelope			X			
74	'Fvar_medenv'	Variance of the frequency envelope			X			
75	'FmedToPeak_medenv (Hz)'	absolute value of median frequency - peak frequency			X			
76	'FM_medenv(Hz)'	Hamonicity (i.e. lag of maxmum peak of the autocorrelation function)			X			
77	'FM_medenv CorrCoef'	"Strength" of harmonicity (i.e. value of maximum peak in autocorrelation function)			X			
78	'Tiqr_medenv (s)'	Inter-quartile range of time envelope						X
79	'Tassym_medenv'	Assymetry of time envelope						X

80	'Tconc_medenv (s)'	Concentration of time envelope						X
81	'Tderiv pct90_medenv'	90th percentile of the first derivative of the time envelope						X
82	'Tvar_medenv'	Variance of the time envelope						X
83	'Tkurt_medenv'	Kurtosis of the time envelope						X
84	'TmedToPeak_medenv (s)'	absolute value of median time - peak time						X
85	'AM_medenv (s)'	Amplitude modulation (i.e. lag of maxmum peak of the autocorrelation function)						X
86	'AM CorrCoef_medenv'	"Strength" of amplitude modulation (i.e. value of maximum peak in autocorrelation function)						X
87	'nTones'							
88	'nTonesPerkHz'							
89	'avToneAmp'							
90	'maxToneAmp'							
91	'DetecBox_nTones'							
92	'DetecBox_nTonesPerkHz'							
93	'DetecBox_avToneAmp'							
94	'DetecBox_maxToneAmp'							