Modelling Marine Vessels Engaged in Wildlife-viewing Behaviour using Automatic Identification Systems (AIS)

by

Andrea Nesdoly

B.Sc., University of Victoria, 2019

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

MASTER OF SCIENCE

in the Department of Geography

We acknowledge and respect the lək̓ʷəŋən peoples on whose traditional territory the university stands, and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Modelling Marine Vessels Engaged in Wildlife-viewing Behaviour using Automatic Identification Systems (AIS)

by

Andrea Nesdoly

B.Sc., University of Victoria, 2019

Supervisory Committee

Dr. Christopher Bone, Supervisor
Department of Geography

Dr. David Atkinson, Departmental Member
Department of Geography

Abstract

Observation of marine animals in their environment – whale-watching – has grown greatly in recent years, bringing risk to the animals. Of particular concern are harmful impacts on marine mammals, some of which are endangered. As a result, regulations have been developed for their protection, but these conservation measures require enforcement across a broad geographic region, which is difficult due to limited monitoring resources. A ship-borne information transmission system called AIS – Automatic Identification System – can provide information-rich marine vessel movement data that can be used to passively monitor vessels engaged in viewing wildlife, aiding regulatory bodies with compliance enforcement. Few studies explore the use of AIS data to determine when vessels are engaged in wildlife-viewing, and as such little guidance exists on how to implement classification models appropriately.

The objective of this thesis is to use AIS data to evaluate the accuracy and utility of existing classification models to detect vessels engaged in observing wildlife, and determine whether information about species being observed can be extracted. Using a control set of observed cetacean encounter data, three classification models were statistically assessed. From this, a hidden Markov model was chosen for detailed analysis in the vicinity surrounding Vancouver Island, B.C., Canada. The resulting analysis concluded that a hidden Markov unsupervised classification approach was feasible for detecting vessel behaviours and differentiating species type. These findings suggest AIS can aid managers and the commercial whale-watching industry in making informed decisions regarding conservation regulations and their compliance.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank everyone that has supported me over my time at the University of Victoria, but especially those who have supported me through my MSc the last two years. First and foremost, my mother and sister for being steadfast emotional supports, without you I would not have made it through. To my supervisor, Dr. Chris Bone, thank you for stepping into this role seamlessly, believing in me even when I may not have believed in myself, and supporting me through the completion of my MSc and BSc. Although broad, I want to thank the Department of Geography: the office staff and lab instructors for affording me learning, teaching, and work opportunities, and, specifically, my committee member Dr. David Atkinson for your guidance and encouragement. A big thank you to those in the CORAL group for all the brain storming sessions over coffee, and the SURREAL lab for allowing me to join you and for providing advice and peer support over the last two years. I want to thank Molly Fraser for sharing the data she collected for her MSc project with me, contributing to this research, and be a valuable peer support during turbulent times. Finally, thank you to my co-funders MEOPAR and exactEarth, and to Ocean Networks Canada for providing the data for this study. Without the support and encouragement of all those mentioned this research would not have been possible.

Co-authorship Statement

Chapter 2 and 3 of this thesis will be submitted as separate standalone papers to peer review journals, and were co-authored by Dr. Bone, Dr. McWhinnie (Chapter 3), Molly Fraser, Norma Serra-Sogas, and Dr. Canessa. The research, method development, data preparation, analyses, and writing for these two papers was conducted by me and supported by the co-authors who helped to develop the research objectives, interpret the results, and edit the two papers. Specifically, co-authors provided writing and editing support for the papers, but Dr. Bone also provided extensive guidance in project development, Molly Fraser provided control data for these papers, and Dr. McWhinnie and Molly Fraser provided guidance on interpreting the results for their significance and applications.

Chapter 1 Introduction

1.1 Research Context

A global increase in vessel traffic has contributed to negative impacts on marine mammals and their environment, prompting marine managers to monitor and regulate vessels at sea (Alessandrini et al., 2017; Halpern et al., 2015). Vessel traffic can cause harm to wildlife directly, through disturbances to physical health and behaviour, or indirectly, through habitat impacts from vessel pollution (e.g. air, oil spills, or propulsion system noise), which has promoted research into the management of anthropogenic activities for marine mammals conservation (McWhinnie et al., 2021; Merchant et al., 2014; Robards et al., 2016). In particular, wildlife-viewing traffic has increased with a surge in the global popularity of whale-watching ecotourism (Mallard, 2019), which has impacts on the species being targeted for observation and the coastal ecosystem (Robards et al., 2016). Commercial whale-watching is a service provided by designated ecotourism companies allowing the public to observe wildlife in their natural environment, and non-commercial (i.e. recreational) whale-watching is typically opportunistic wildlife encounters on fishing or pleasure vessels. Although vessel based whale-watching is argued to be a sustainable form of human-wildlife interactions (Mallard, 2019; Schuler et al., 2019), these vessels have the potential to impact the physical health of the cetaceans being observed (Conn & Silber, 2013; Vanderlaan & Taggart, 2007) by causing changes in their behaviour (e.g. foraging, diving, and resting behavioural patterns) (Lusseau, 2006; Schuler et al., 2019; Williams et al., 2002) and altering their habitats (Bejder et al., 2006; Mallard, 2019).

Along the British Columbia (B.C.) and Washington state coastlines, regulations and guidelines for commercial and non-commercial whale-watching are in place for the conservation of marine mammals. They include, but are not limited to, restrictions on vessel speeds while observing and approaching cetaceans, minimum distances at which boats can approach cetaceans, and limits on the number of vessels present at one encounter (Be Whale Wise, 2020; Fisheries and Oceans Canada, 2019, 2020, 2021; Washington State Legislature, 2019). Moreover, spatially explicit conservation measures include voluntary vessel slowdown in high traffic regions (Vancouver Fraser Port Authority, 2020) and sanctuary zones within cetacean critical habitat (Fisheries and Oceans Canada, 2019, 2021).

Compliance of commercial and recreational whale-watching vessels to these regulations and guidelines is directly related to their enforcement (Seely et al., 2017). It was found that commercial ecotourism companies placed higher importance on regulations that were actively enforced to avoid fines and litigation (Amerson & Parsons, 2018), and recreational vessel were more likely to be non-compliant around killer whales and violate minimum distance regulations, which may be due to a lack of awareness of existing policies or complex transboundary regulations leading to confusion (Fraser et al. 2020; Seely et al., 2017). Active regulation enforcement through the presence of enforcement vessels at sea is needed to ensure regulation compliance (Mallard, 2019). Unfortunately, it is difficult for enforcement vessels to provide full time monitoring of cetacean encounters at sea across a geographically vast area because both cetaceans and vessels are highly mobile and often range across political boundaries and a complex spectrum of stakeholders (Giles & Koski, 2012; Higham et al., 2008; Mallard, 2019). Thus, passive methods to monitor commercial and non-commercial whale-watching vessels are needed to properly assess compliance and enforce wildlife observation regulations.

Vessel monitoring systems (VMS) have been utilized for passive monitoring of regulation compliance through the detection of vessel behaviour (i.e. the activity in which a vessel is engaged) in numerous sectors, such as fisheries management (Merten et al., 2016), maritime safety (Wang et al., 2013), and environmental monitoring (e.g. detecting oil spills, pollution, and illegal dumping) (Robards et al., 2016; Schwehr & McGillivary, 2007). Through detecting vessel behaviour, regulatory bodies can directly enforce regulations by sending an enforcement officer to the scene, or retroactively enforce regulations through fines and litigation. Similarly, by utilizing VMS to detect when commercial and non-commercial whale-watching vessels are engaged in wildlife-viewing behaviour, these vessels can be passively monitored to assess compliance to wildlife-viewing regulations and guidelines. Often recreational vessels use commercial whale-watching vessels as their guide for locating cetaceans and determining their distance from cetaceans during an encounter (Kessler & Harcourt, 2013). Due to complex and varying regulations pertaining to the different species being observed and the type of vessel engaged in whale-watching (commercial or recreational), boaters are not always aware of their compliance. As such, passive vessel monitoring would provide assurance to both operators and marine managers that compliance measures are being adequately followed.

Knowing when vessels are specifically engaged in wildlife-viewing behaviour is required to evaluate compliance. Fortunately, many commercial whale-watching vessels along the B.C. and Washington state coastlines require a VMS aboard called an Automatic Identification System (AIS), which is an autonomous vessel reporting system that was originally developed for safety and collision avoidance at sea (IMO, 2003). Suggested in 2019 and enacted in 2021, Washington State requires AIS on all commercial whale-watching vessels (Washington State Legislature, 2021). Since 2019, Canada requires any passenger vessels with 12 or more passengers or longer than 8 m to be equipped with AIS (Canadian Shipping Act, 2019), which encompasses many commercial ecotourism vessels. Unfortunately, AIS cannot capture all whale-watching activity because Canadian commercial whale-watching vessels that do not meet these regulations are not required to carry AIS. Regardless, the recent enactment of these regulations provides a unique opportunity for the study of wildlife-viewing vessel activity using AIS. With a network of terrestrial and satellite AIS receivers, AIS has the potential to passively monitor AIS enabled whale-watching vessels during wildlife encounters, yet research is still needed into how AIS data can be utilized in vessel movement models to classify when commercial whale-watching vessels are engaged in wildlife-viewing behaviour.

Previous research has successfully utilized AIS data in vessel movement models to classify vessel behaviour, with a focus on the fishing and shipping industries (De Souza et al., 2016; Le Tixerant et al., 2018; Sheng et al., 2018). However, there is only one known study (Almunia et al., 2020) that used AIS data to classify when commercial whale-watching vessels were engaged in wildlife-viewing in Spain through thresholds on speed over ground (SOG) and auxiliary bathymetry data. Although these methods were successful in the classification of wildlife-viewing events, there exists a reliance on auxiliary data and possible limitations to applying SOG thresholds. For example, in the bathymetric complexity of the B.C. and Washington state coastal region SOG thresholds may not capture changes in speed during vessel maneuvers while the vessel is not engaged in wildlife-viewing, such as navigating through deep narrow channels, or high traffic areas. Therefore, it is necessary to determine if established classification models can be used to detect wildlife-viewing vessel behaviour, and if these models can take advantage of temporal, spatial, kinematic, or geometric variables derived from the AIS data to reduce reliance on auxiliary data.

Classification methods used to detect vessel behaviour can be approached from statistical, and statistical data mining and machine learning perspectives (Sidibé & Shu, 2017), but the application of these methods utilizing AIS data are predominantly found in marine traffic analyses, classifying abnormal or illegal behaviour, fishing vessel monitoring, or maritime security (Pallotta et al., 2013; Schwehr & McGillivary, 2007). For example, data mining and machine learning approaches were used to classify vessel behaviour using AIS data in studies of vessel traffic flow forecasting (Mone et al., 2010), vessel movement anomaly detection and trajectory prediction (Pallotta et al., 2013; Sidibé & Shu, 2017), and fishing vessel behaviour classification (Charles et al., 2014; De Souza et al., 2016; Peel & Good, 2011). However, to date, there are no known studies applying established data mining and machine learning classification methods to the detection of whale-watching vessel behaviour. Statistical modelling can be used for classification of wildlife-viewing behaviour, but it is typically used for smaller sample sizes, and it sacrifices classification accuracy for interpretability to assess the relationship between the AIS variables and the classification output (Bzdok, 2018). Statistical data mining focuses on accurate classification and utilizes statistical methods with the computer power of machine learning practices to learn the structure and patterns within the large volume of AIS data that represent the activities vessel are engaged in (Ratner, 2011). As precise observed behaviour class data (i.e. wildlife encounter data) for commercial ecotourism vessels are limited, machine learning may be beneficial for wildlife-viewing behaviour classification because it encompasses methods that automatically classify data patterns that can be used to make decisions under uncertainty (Murphy, 2012; Samuel, 1959).

Two main machine learning approaches exist: supervised learning, which uses pre-defined observed behaviour class data to derive model parameters, or unsupervised learning, which obtains the model parameters from patterns within the AIS data without prior knowledge of the observed behaviour classes (Murphy, 2012; Zhang et al., 2005). However, with limited precise observed behaviour class data, there is a need to understand how an unsupervised classification model compares to a supervised model in order to develop suitable methods for the classification of commercial ecotourism vessel behaviour without reliance on observed class data.

1.2 Research Objectives and Structure

Therefore, in light of the negative impacts associated with wildlife-viewing activities and the regulatory regime that has been developed in response, it is important to be able to rapidly evaluate regulation compliance over a large geographical region. While there exist AIS data that can form the basis of an evaluation approach, no classification methods have been shown to reliably classify when commercial ecotourism vessels are engaged in wildlife-viewing activities. Reliable classification methods that are applicable to a commercial ecotourism fleet across a broad geographic region are essential to better understand vessel movement characteristics while observing wildlife. Therefore, the objectives of this research are (1) to evaluate established classification models for their feasibility in detecting wildlife-viewing vessel behaviour using AIS data, and (2) to determine if an unsupervised machine learning approach is comparable to a supervised approach, and to what degree vessel movement information can indicate which cetacean species is being observed based on the models' outputs. These research objectives are addressed in Chapters 2 and 3, respectively, which are formatted as two standalone papers that will be submitted as separate manuscripts for publication in peer-reviewed journals.

Following the introduction, Chapter 2 addresses the first research objective through a systematic evaluation of three established classification methods applied in other marine industries for their feasibility in detecting when commercial ecotourism vessels are engaged in wildlife-viewing. These models include density-based spatial clustering application with noise (DBSCAN), hidden Markov model (HMM), and logistic regression (LR), all of which have been used previously to classify vessel behaviour in industries, such as fishing, shipping, and marine security.

To address the second research objective, Chapter 3 took the best performing classification model from Chapter 2 to determine if an unsupervised learning approach could perform as well as the supervised learning approach without reliance on observed class data. This was done through evaluating the supervised and unsupervised models' statistical performance, and through exploring any cetacean species related vessel movement trends in the classified wildlife-viewing events. The latter of which could only be assessed once these classification models and approaches were validated to ensure all possible species observations were captured by the models prior to associating classified wildlife-viewing events with approximated cetacean sightings based on assumptions. The concluding chapter in this thesis summarizes the findings from Chapters 2 and 3 while outlining potential applications of these methods and suggesting future research directions.

1.5 Literature Cited

Alessandrini, A., Alvarez, M., Greidanus, H., Gammieri, V., Arguedas, V. F., Mazzarella, F., Santamaria, C., Stasolla, M., Tarchi, D., & Vespe, M. (2017). Mining Vessel Tracking Data for Maritime Domain Applications. *IEEE International Conference on Data Mining Workshops*, 361–367. https://doi.org/10.1109/ICDMW.2016.0058

Almunia, J., Delponti, P., & Rosa, F. (2020). Using big data to estimate whale watching effort. *bioRxiv*, 1–19. https://doi.org/10.1101/2020.11.30.403923

Amerson, A., & Parsons, E. C. M. (2018). Evaluating the sustainability of the gray-whale-watching industry along the pacific coast of North America. *Journal of Sustainable Tourism*, 26(8), 1362–1380. https://doi.org/10.1080/09669582.2018.1449848

Be Whale Wise (2020). Marine Wildlife Laws and Guidelines for Boaters, Paddlers and Viewers. https://www.bewhalewise.org/wp-content/uploads/2020/07/Be-Whale-Wise-Brochure-2020-v5.pdf

Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J., Connor, R., Heithaus, M., Watson-Capps, J., Flaherty, C., & Krützen, M. (2006). Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conservation Biology*, 20(6), 1791–1798. https://doi.org/10.1111/j.1523-1739.2006.00540.x

Bzdok, D., Altman, N., Krzywinski, M. (2018). Points of Significance: Statistics versus machine learning. *Nature Methods,* 15, 233–234. https://doi.org/10.1038/nmeth.4642

Canadian Shipping Act (2019). Regulations Amending the Navigation Safety Regulations (Automatic Identification Systems) (SOR/2019-100). *Canada Gazette Part II*, 153(9). https://gazette.gc.ca/rp-pr/p2/2019/2019-05-01/html/sor-dors100-eng.html

Charles, C., Gillis, D., & Wade, E. (2014). Using hidden Markov models to infer vessel activities in the snow crab (*Chionoecetes opilio*) fixed gear fishery and their application to catch standardization. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(12), 1817–1829. https://doi.org/10.1139/cjfas-2013-0572

Conn, P. B., & Silber, G. K. (2013). Vessel speed restrictions reduce risk of collision-related mortality for North Atlantic right whales. *Ecosphere*, 4(4):43. https://doi.org/10.1890/ES13-00004.1

De Souza, E. N., Boerder, K., Matwin, S., & Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PLoS ONE*, 11(7). https://doi.org/10.1371/journal.pone.0158248

Fisheries and Oceans Canada, 2019. Government of Canada outlines its 2019 plan for protecting Southern Resident killer whales. https://www.canada.ca/en/fisheries-oceans/news/2019/05/government-of-canada-outlines-its-2019-plan-for-protecting-southern-resident-killer-whales.html

Fisheries and Oceans Canada. 2020. Watching marine wildlife. https://www.dfo-mpo.gc.ca/species-especes/mammals-mammiferes/watching-observation/index-eng.html

Fisheries and Oceans Canada, 2021. 2021 management measures to protect Southern Resident killer whales. https://www.pac.dfo-mpo.gc.ca/fm-gp/mammals-mammiferes/whales-baleines/srkw-measures-mesures-ers-eng.html

Fraser, M. D., McWhinnie, L. H., Canessa, R. R., & Darimont, C. T. (2020). Compliance of small vessels to minimum distance regulations for humpback and killer whales in the Salish Sea. *Marine Policy*, 121, 104171. https://doi.org/10.1016/j.marpol.2020.104171

Giles, D. A., & Koski, K. L. (2012). Managing vessel-based killer whale watching: A critical assessment of the evolution from voluntary guidelines to regulations in the Salish Sea. *Journal of International Wildlife Law and Policy*, 15(2), 125–151. https://doi.org/10.1080/13880292.2012.678792

Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., Lowndes, J. S., Rockwood, R. C., Selig, E. R., Selkoe, K. A., & Walbridge, S. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature Communications*, 6:7615, 1–17. https://doi.org/10.1038/ncomms8615

Higham, J. E. S., Bejder, L., & Lusseau, D. (2008). An integrated and adaptive management model to address the long-term sustainability of tourist interactions with cetaceans. *Environmental Conservation*, 35(4), 294–302. https://doi.org/10.1017/S0376892908005249

International Maritime Organization (IMO). (2003). Guidelines for the Installation of a Shipborne Automatic Identification System (AIS) (Ref. T2/8.02)

Kessler, M., & Harcourt, R. (2013). Whale watching regulation compliance trends and the implications for management off Sydney, Australia. *Marine Policy*, 42, 14–19. https://doi.org/10.1016/j.marpol.2013.01.016

Le Tixerant, M., Le Guyader, D., Gourmelon, F., & Queffelec, B. (2018). How can Automatic Identification System (AIS) data be used for maritime spatial planning? *Ocean and Coastal Management*, 166, 18–30. https://doi.org/10.1016/j.ocecoaman.2018.05.005

Lusseau, D. (2006). The short-term behavioral reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Marine Mammal Science*, 22(4), 802–818. https://doi.org/10.1111/j.1748-7692.2006.00052.x

Mallard, G. (2019). Regulating whale watching: A common agency analysis. *Annals of Tourism Research*, 76, 191–199. https://doi.org/10.1016/j.annals.2019.04.011

McWhinnie, L. H., O'Hara, P. D., Hilliard, C., Le Baron, N., Smallshaw, L., Pelot, R., & Canessa, R. (2021). Assessing vessel traffic in the Salish Sea using satellite AIS: An important contribution for planning, management and conservation in southern resident killer whale critical habitat. *Ocean and Coastal Management*, 200, 1–17. https://doi.org/10.1016/j.ocecoaman.2020.105479

Merchant, N. D., Pirotta, E., Barton, T. R., & Thompson, P. M. (2014). Monitoring ship noise to assess the impact of coastal developments on marine mammals. *Marine Pollution Bulletin*, 78, 85–95. https://doi.org/10.1016/j.marpolbul.2013.10.058

Merten, W., Reyer, A., Savitz, J., Amos, J., Woods, P., & Sullivan, B. (2016). Global Fishing Watch: Bringing Transparency to Global Commercial Fisheries. *Bloomberg Data for Good Exchange Conference*. http://arxiv.org/abs/1609.08756

Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. In *The MIT Press* (Issue 4).

Pallotta, G., Vespe, M., & Bryan, K. (2013). Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6), 2218–2245. https://doi.org/10.3390/e15062218

Peel, D., & Good, N. M. (2011). A hidden markov model approach for determining vessel activity from vessel monitoring system data. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(7), 1252–1264. https://doi.org/10.1139/f2011-055

Ratner, B. (2011) Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition, *CRC Press LLC*,. *ProQuest Ebook Central*, https://ebookcentral-proquest-com.ezproxy.library.uvic.ca/lib/uvic/detail.action?docID=840391

Robards, M. D., Silber, G. K., Adams, J. D., Arroyo, J., Lorenzini, D., Schwehr, K., & Amos, J. (2016). Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)-A review. *Bulletin of Marine Science*, 92(1), 75–103. https://doi.org/10.5343/bms.2015.1034

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal*, 210–229.

Schuler, A. R., Piwetz, S., Di Clemente, J., Steckler, D., Mueter, F., & Pearson, H. C. (2019). Humpback Whale Movements and Behavior in Response to Whale-Watching Vessels in Juneau, AK. *Frontiers in Marine Science*, 6:710, 1–13. https://doi.org/10.3389/fmars.2019.00710

Schwehr, K. D., & McGillivary, P. A. (2007). Marine ship automatic identification system (AIS) for enhanced coastal security capabilities: An oil spill tracking application. *Oceans Conference Record (IEEE)*. https://doi.org/10.1109/OCEANS.2007.4449285

Seely, E., Osborne, R. W., Koski, K., & Larson, S. (2017). Soundwatch: Eighteen years of monitoring whale watch vessel activities in the Salish Sea. *PLoS ONE*, 12(12), 1–18. https://doi.org/10.1371/journal.pone.0189764

Sheng, K., Liu, Z., Zhou, D., He, A., & Feng, C. (2018). Research on Ship Classification Based on Trajectory Features. *Journal of Navigation*, 71(1), 100–116. https://doi.org/10.1017/S0373463317000546

Sidibé, A., & Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. *Journal of Navigation*, 70(4), 847–858. https://doi.org/10.1017/S0373463317000066

Soe, O. K. M., Chaojian, S., Qinyou, H., & Weintrit, A. (2010). Clustering Analysis and Identification of Marine Traffic Congested Zones. *Zeszyty Naukowe Akademii Morskiej W Gdyni*, 67, 101–113.

Vancouver Fraser Port Authority. (2020). ECHO Program 2019 voluntary vessel slowdown trial in Haro Strait and Boundary Pass: Summary findings.

Vanderlaan, A. S. M., & Taggart, C. T. (2007). Vessel collisions with whales: The probability of lethal injury based on vessel speed. *Marine Mammal Science*, 23(1), 144–156. https://doi.org/10.1111/j.1748-7692.2006.00098.x

Wang, Y., Zhang, J., Chen, X., Chu, X., & Yan, X. (2013). A spatial-temporal forensic analysis for inland-water ship collisions using AIS data. *Safety Science*, 57, 187–202. https://doi.org/10.1016/j.ssci.2013.02.006

Washington State Legislature (2019). Protection of southern resident orca whales—Unlawful activities—Penalty. https://apps.leg.wa.gov/rcw/default.aspx?cite=77.15.740

Washington State Legislature (2021). Commercial whale watching compliance and reporting. https://apps.leg.wa.gov/wac/default.aspx?cite=220-460-140

Williams, R., Trites, A. W., & Bain, D. E. (2002). Behavioural responses of killer whales (Orcinus orca) to whale-watching boats: opportunistic observations and experimental approaches. *Journal of Zoology*, 256, 255–270. https://doi.org/10.1017/s0952836902000298

Zhang, D., Gatica-Perez, D., Bengio, S., & McCowan, L. (2005). Semi-supervised adapted HMMs for unusual event detection. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 611–618. https://doi.org/10.1109/CVPR.2005.316

Chapter 2 Evaluating Models for Classifying Movement of Whale-watching Vessels

Abstract

The Automatic Identification System (AIS) provides information-rich marine vessel movement data that can be used for monitoring and managing vessel traffic. In addition, AIS provides a unique opportunity to understand various aspects of vessel traffic patterns, including analyzing vessel density and route patterns, predicting vessel trajectories, and detecting vessel behaviour. These data can provide a basis for monitoring wildlife-viewing activities. However, minimal research currently exists that utilizes AIS data for such purposes, and as such there is little guidance on the types of classification models that would be appropriate to undertake AIS-focused wildlife-viewing vessel movement analysis. The objective of this study is to evaluate three established classification models for their ability to utilize AIS data to describe wildlife-viewing vessel behaviour. The models include density-based spatial clustering application with noise (DBSCAN), hidden Markov model (HMM), and logistic regression (LR), all of which have been previously used to classify vessel behaviour in industries, such as fishing, shipping, and marine security. A sample of AIS data from whale-watching vessels originating in a harbour in the city of Victoria, British Columbia, Canada was collected for August 2019 and used for model development. The results of each model's classification were validated against observed whale sighting data using statistical performance and accuracy metrics. The findings suggest that all three classification models sufficiently detect wildlife-viewing behaviour, but the HMM and LR had preferable performance metrics compared to DBSCAN. Although LR provides an informative glance at which AIS variables are most important to detecting wildlife-viewing events, the HMM has comparable performance metrics and requires less data processing. With a false positive rate $\leq$ 10% suggesting

minimal misclassification, and a receiver operating characteristic area under the curve score of 0.91suggesting the HMM distinguishes between classes well, this study recommends the use of HMM due to its computational efficiency and because it provides an accurate classification of wildlife-viewing behaviour for whale-watching vessels. The results of this study can be used to support policy decisions, manage regulation compliance, and inform marine conservation initiatives.

2.1 Introduction

The rising popularity of whale-watching has led to a global increase in wildlife-viewing vessel traffic (Mallard, 2019), resulting in growing threats to coastal ecosystems and marine mammals (Robards et al., 2016). Commercial wildlife-viewing vessels provide a service for the public to observe wildlife in their natural environment, and are typically smaller than those involved in other marine activities such as shipping (e.g. cargo ships and tankers) and commercial fishing. However, these small vessels (i.e. <150 tonnes) still have the potential to exert harmful impacts to marine wildlife, such as vessel collisions with whales (Conn & Silber, 2013), acoustic disturbance from noise created by vessel engines (Erbe, 2002; Frankel & Gabriele, 2017; Hermannsen et al., 2019), and water contamination caused by vessel pollution (Lachmuth et al., 2011). Such impacts have the potential to alter the behaviour of wildlife in the short term, such as changes to foraging, diving, and resting patterns (Lusseau, 2006; Schuler et al., 2019), which in the long term can impact biological fitness and alter habitats (Bejder et al., 2006; Mallard, 2019). Therefore, characterization of wildlife-viewing vessel traffic, especially in the context of regulatory adherence, is crucial in aiding marine policy-makers in mitigating impacts to wildlife and the marine environment.

Monitoring wildlife-viewing vessel behaviour can be done using location and time-specific data on vessel movement provided by the Automatic Identification System (AIS), an autonomous vessel reporting system originally developed as a navigational aid for safety and collision avoidance at sea (IMO, 2003; Robards et al., 2016). An AIS transponder aboard a vessel automatically broadcasts vessel information, including identifiers, kinematics (speed, course, etc.), and position at regular time intervals (i.e. typically between 30 seconds and 3 minutes depending on vessel speed) to other AIS vessels and AIS receivers. Applications of these data include, but are not limited to, national security, fisheries management, marine spatial planning (Alessandrini et al., 2017), environmental conservation, irregular or illegal vessel activity detection (Robards et al., 2016), and analyses of marine traffic patterns (Lei et al., 2016). In addition, there is a growing body of research utilizing AIS data to investigate marine vessel compliance to speed restrictions to prevent collisions with wildlife (Silber et al., 2014; Wiley et al., 2011), the impacts of acoustic disturbances from vessels on wildlife (Fouda, 2012; Frankel & Gabriele, 2017; Joy et al., 2019), and the management of marine traffic to mitigate wildlife impacts (Parrott et al., 2011). However, in the investigation of wildlife-viewing vessel behaviour, an underlying challenge is to detect when vessels with AIS are engaged in wildlife-viewing activities accurately because precise observed vessel behaviour data are limited.

While marine vessel behaviour classification using AIS is common in the fishing and shipping industries (De Souza et al., 2016; Le Tixerant et al., 2018; Sheng et al., 2018), and used to classify abnormal vessel behaviour (Sidibé & Shu, 2017), there are few studies that investigate how AIS can be utilized to classify when small marine vessels are observing wildlife. An existing study (Almunia et al., 2020) classified vessels using a vessel speed threshold along with auxiliary bathymetry data to distinguish between those vessels undertaking wildlife-viewing and those that are not. However, speed and bathymetry thresholds may be insufficient in regions with coastal and bathymetric complexity because they do not capture the nuances of the changes in speed when a vessel is maneuvering during non-viewing events. For example, the speed threshold would not be able to differentiate between a wildlife-viewing event and navigating through deep narrow passages, or high traffic areas. Including temporal, spatial, or geometric variables derived from AIS data in the classification method would provide a richer context of vessel movement behaviour while at the same time reducing reliance on auxiliary data. Research is needed to determine if such characteristics can be used in the context of small-vessel classification for wildlife-viewing, and if so, which type of models are most useful at providing appropriate classifications.

With regards to available classification methods, there now exists a range of statistical data mining and machine learning approaches to detect marine vessel activities (Ratner, 2011; Sidibé & Shu, 2017), but many previous applications using these methods are largely in the context of marine traffic analyses, maritime security initiatives, monitoring fishing vessels, or classifying abnormal or illegal behaviour (Pallotta et al., 2013; Schwehr & McGillivary, 2007). For example, density-based spatial clustering application with noise (DBSCAN) is an clustering classification model that has been used in marine traffic flow forecasting (Soe et al., 2010), trajectory prediction (Pallotta et al., 2013), and classification of abnormal vessel behaviour or abnormal AIS reception (Li et al., 2014). The DBSCAN model partitions spatial data into clusters of arbitrary shapes based on sufficiently high density of AIS points that represent marine vessel behaviour (Soe et al., 2010), and provides a means to classify points that are considered noise (Pallotta et al., 2013). These studies revealed that, with domain expert knowledge and correct model parameters, DBSCAN can be adapted to a variety of marine situational awareness applications.

Another modelling approach utilized by vessel researchers is hidden Markov models (HMM), a state-based model that takes advantage of vessel speed over ground (SOG). Traditionally used in ecological movement and pattern detection (Joo et al., 2013), HMM have been used to infer fishing vessel activity and estimate fishing effort (Charles et al., 2014; De Souza et al., 2016; Peel & Good, 2011; Vermard et al., 2010; Walker & Bez, 2010), and to detect vessel moorage (Waterbolk et al., 2019). HMM infer vessel behaviour classes (i.e. hidden states) probabilistically from the observed AIS data sequence (Kouemou, 2011), which assumes behaviour classes (states) are linked by a Markov chain (Markov, 2006) that accounts for transitions between states. That is, the probabilities that make up the HMM determine whether each AIS point is viewing wildlife based only on information observed in each point in the AIS data and the previous AIS point's classification output. Utilizing HMM, these studies successfully classify various marine vessel behaviour with observed vessel data, such as speed, turning angle, position, or distance.

Finally, logistic regression models have been used in studies of marine vessel behaviour to analyze accidents at sea between vessels (Bye & Aalberg, 2018), collisions between vessels and marine mammals (Conn & Silber, 2013), and to differentiate between marine vessel types (Sheng et al., 2018). Logistic regression can be considered a discriminative classification model that can utilize kinematic, geometric, and spatial AIS variables to provide a thorough analysis of the factors contributing to wildlife-viewing vessel behaviour. Logistic regression is a special case of a generalized linear model arising from a Bernoulli conditional probability model for binary classification problems, and have been found to provide a versatile approach for accurately inferring vessel behaviour, and for providing information on variable and class relationships (Murphy, 2012; Schein & Ungar, 2007; Sheng et al., 2018).

While the models mentioned above all hold promise for performing classifications of wildlife-viewing vessel behaviour, research is needed to investigate the application of these models to wildlife-viewing vessels. This will lead to a better understanding of wildlife-viewing activities that can be used to support marine management and aid in wildlife conservation (McWhinnie et al., 2021). As such, the objective of this study is to evaluate DBSCAN, HMM and logistic regression to determine the feasibility of each model for wildlife-viewing behaviour classification. This study focuses on commercial wildlife-viewing (i.e. whale-watching) vessels in the Salish Sea region on the southern coast of Vancouver Island in British Columbia (B.C.), Canada. This does not include recreational or private vessels. Each model is considered in terms of its methodological basis, ease of implementation (e.g. data preprocessing and model execution time), and statistical performance compared to observed vessel behaviour. The aim of performing this analysis is to provide marine managers, conservationists, and researchers an appraisal of the available methods to evaluate small marine vessels behaviour using AIS while observing wildlife.

2.2 Methods

2.2.1 Study Area and Data

This study analyzed a single commercial wildlife-viewing vessel whose port location is in the city of Victoria, B.C. on the west coast of Canada (Fig. 2.1 & 2.2). The study area is contained within what is referred to as the Salish Sea, a region with an abundance of marine mammals and rich coastal biodiversity that shapes a popular commercial ecotourism industry (Giles & Koski, 2012; Seely et al., 2017). Marine mammals observed include killer whales (*Orcinus orca*), humpback whales (*Megaptera novaeangliae*), grey whales (*Eschrichtius robustus*), Harbour seals (*Phoca vitulina*), Steller sea lions (*Eumetopias jubatus*) and California sea lions (*Zalophus californianus*) (Fisheries and Oceans Canada, 2020).

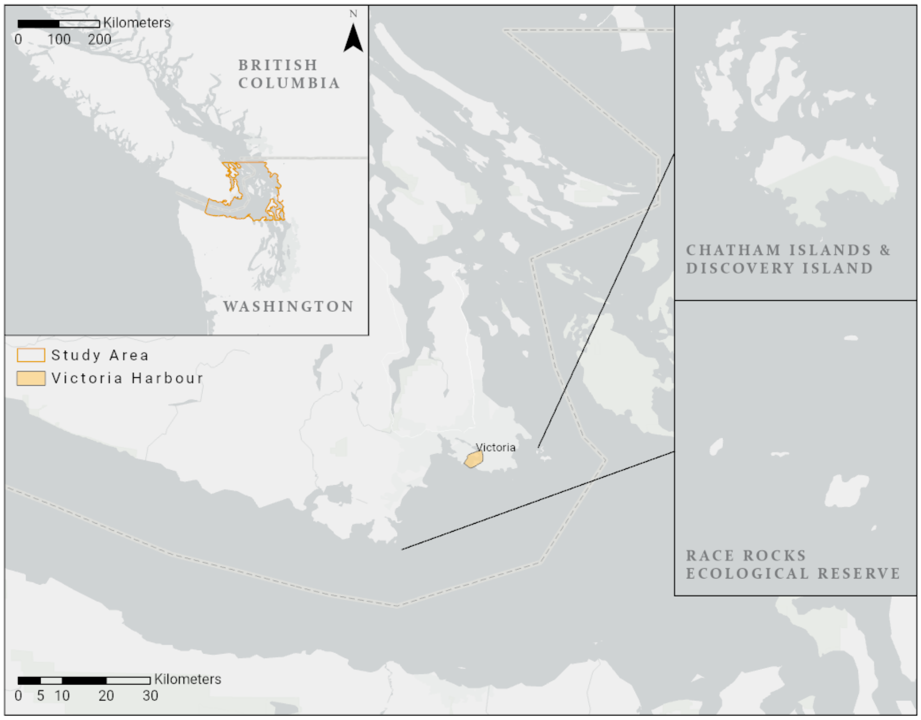

Figure 2.1: Study Area of the Salish Sea and known wildlife observation sites of Race Rocks Ecological Reserve, and the waters between Chatham Islands and Discovery Island.

For this study, observed whale encounter data and AIS data were obtained for the same commercial ecotourism vessel for the month of August 2019. These data were used to define the observed behaviour class for each AIS point, which are then used to derive the classification model parameters and for model validation. The observed whale encounter data were obtained from a previous study by Fraser et al. (2020) that was focused on the compliance of whale-watching vessels to minimum distance restrictions within the Salish Sea. These data were collected aboard a single commercial whale-watching vessel where there were potentially multiple whale encounters observed for each trip. The vessel used to collect the whale encounter data is uniquely identified by its Maritime Mobile Service Identity (MMSI), which is assigned to vessels at sea for automated safety communication as part of the global maritime distress and safety system (Government of Canada, 2019). This unique identifier can be used to obtain AIS data for this vessel.

The AIS data were collected by the Canadian Coast Guard and provided for this study by Ocean Networks Canada (ONC). AIS transponders were made mandatory on small passenger vessels carrying more than 12 passengers or that are 8 m or more in length by the Canadian Government in 2019, which applies to many commercial wildlife-viewing vessels (Canadian Shipping Act, 2019). Wildlife-viewing vessels that do not meet these regulations are not required to carry AIS. The AIS data are encoded in the National Marine Electronics Association format and require decoding to be able to read vessel identifiers (MMSI), kinematics (speed, course, etc.), position, and timestamp. Once decoded, the AIS data for the vessel that collected the whale encounter data were extract from all the AIS vessel data provided by ONC using its MMSI. Individual data points were removed if they represented: (1) a duplicate AIS message received within a 10 second timespan; (2) stationary docked vessel in a harbour; (3) an invalid vessel position report with speed over ground greater than 70 knots; or (4) an invalid vessel position report with a sudden change of direction.

Once the AIS data were obtained for the vessel that collected the observed whale encounter data, the date for each wildlife-viewing trip, the tour departure time, and the time interval for the observed whale encounters were applied to each corresponding point in the AIS data. This process resulted in a sample of 20 wildlife-viewing trips for the month of August 2019 that included 1,854 AIS messages. For this study, a wildlife-viewing trip is considered to begin and end when the vessel left and returned to the harbour, respectively. A harbour dataset was derived representing the spatial extent of the Victoria Inner Harbour based on the narrow harbour opening (Fig. 2.1). A unique trip identifier was assigned to each AIS point within a wildlife-viewing trip. In addition, this study assumes that wildlife-viewing occurred at two known wildlife-viewing sites within the study area at Race Rocks Ecological Reserve and the waters surrounding Chatham Islands and

Discovery Island where commercial ecotourism vessels frequently observed Harbour seals, Stellar and California sea lions, and possibly coastal avian and terrestrial animals (e.g. bald eagles and wolves). This study assumes that wildlife-viewing behaviour includes all species encountered, so any AIS points intersecting these two sites were considered to belong to the observed wildlife-viewing class. This process resulted in 35% AIS points classified as wildlife-viewing and 65% classified as non-viewing in the observed classes (Fig. 2.2). This formed the control dataset against which the automated methods were compared.

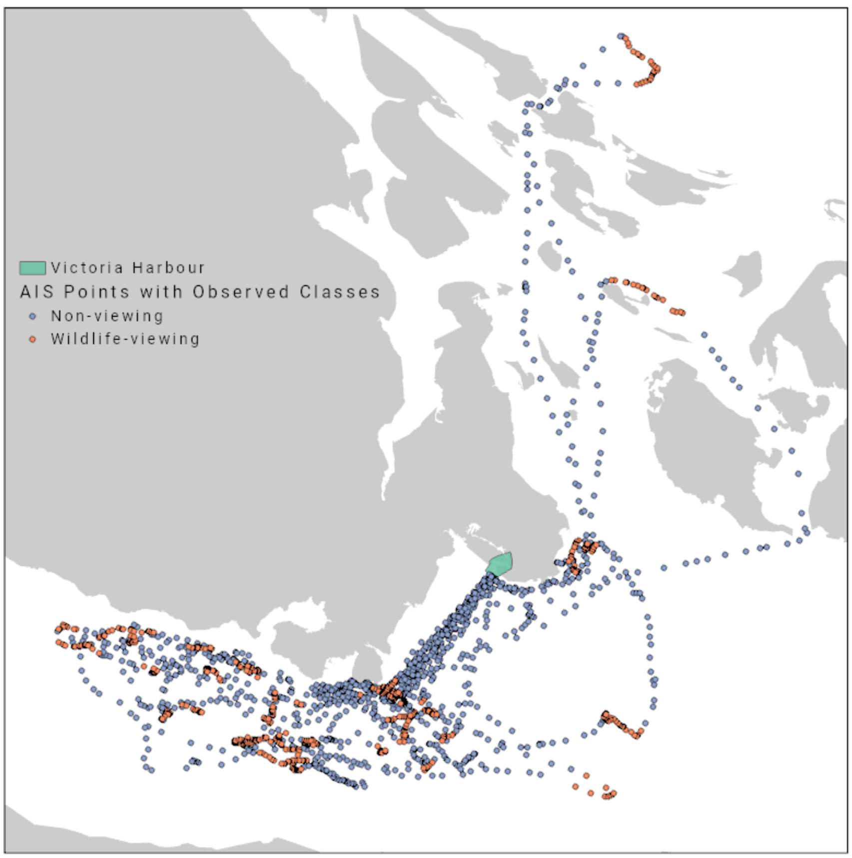

Figure 2.2: AIS data for a single wildlife-viewing vessel out of the Victoria, B.C. harbour with the observed classes manually assigned.

2.2.2 Models

2.2.2.1 Density-Based Spatial Clustering Application with Noise

DBSCAN (Ester et al. 1996) is a non-parametric classification model commonly used to classify anomalous vessel behaviour across different industries (Sidibé & Shu, 2017). This density-based clustering method is computationally simple to employ and is useful because it can find arbitrarily shaped clusters of multi-dimensional data points while detecting and eliminating data points determined to be noise (Li et al., 2018). Although the multi-dimensional capabilities of DBSCAN can incorporate numerous AIS variables, in this study, DBSCAN generated density-based clusters of AIS points based only on geographic location.

DBSCAN is a systematic process that begins by choosing a random point and determining a defined neighbourhood surrounding this point, resulting in an estimated density of points. A nearest neighbour algorithm (NNA) finds all points within the neighbourhood defined by radius (ε). If the neighbourhood contains a minimum number of points (m), then the randomly chosen point is defined as a core point with high-density. If the point is a core point, a new cluster starts and all points within the neighbourhood are assigned to this cluster. The algorithm continues to loop through the unclassified points until every core point and subsequent cluster is found. Any point not assigned to a cluster is classified as noise. This study assumes that the points within clusters represent the wildlife-viewing class, and the noise points represent the non-viewing class.

The observed classes were used to validate revisions to these parameters and evaluate the model using statistical performance metrics (Cazzanti & Pallotta, 2015). First, inferences from the AIS data and prior wildlife-viewing vessel behaviour knowledge were used to determine a reasonable range of values for ε and m (Table 2.1). The optimal parameter values were then defined by finding the combination of ε and m with the highest F-Score (Table 2.2) resulting from an assessment of the F-Score's sensitivity to the changing ε and *m* parameters with each DBSCAN iteration. The F-Score was used because it was found to be a suitable test when optimizing binary classifier parameters (Hossin & Sulaiman, 2015). The optimal NNA was chosen based on its frequency in relation to the optimal ε and *m*, and employs the Euclidean distance metric to calculate the distance between points. (H. Li et al., 2017; Y. Li et al., 2017). The validated parameters were then used to generate the DBSCAN model and classify the AIS points for each wildlife-viewing trip.

Table 2.1: Classification model parameter descriptions and tested values.

Classification Algorithm	**Model Parameters**	**Description**	**Parameter Values Tested**
Density-based Spatial Clustering Application with Noise	Epsilon (ε)	Distance between two points for one to be within the neighbourhood of the other	Range of 0-5 with variable step sizes
	Minimum Samples (*m*)	Minimum number of points for each cluster to form	Range of 0-10 with 1 step
	Distance Metric	Method used to measure distance	Euclidean Distance
	Nearest Neighbour Algorithm	Method used to find the number of points with the closest distance	Brute-Force, KD-tree, Ball-tree
	Leaf Size	The max number of leaves in the KD-tree and Ball-tree	Range of 0-55 with 5 step
Hidden Markov Model	Initial Probabilities	The probability that the first point in a sequence is state 0 or state 1	Derived from the sample data with observed classes
	Emission Probabilities	Probability density functions fit to each state's observed variables	105 established continuous distribution functions for each state
	Transition Probabilities	Conditional probability matrix of the transition between states	Derived from entire sample
	Decoding Algorithm	Decodes the sequence of observations and determines the most likely hidden state sequence	Viterbi Algorithm

Logistic Regression	Solver	Gradient-based optimization solving algorithm	Coordinate descent (liblinear), or Broyden–Fletcher–Goldfarb–Shanno algorithm (lbfgs)
	Penalty	Regularization penalty method	L1-norm, L2-norm
	C	Regularization strength: Inverse of the regularization coefficient	Range of 0-5 with 0.001 step
	Maximum Iterations	Maximum number of iterations for solvers to converge	Range of 50,100-500 with 100 step
	Tolerance	Tolerance for stopping gradient descent near local optimum	Range of 0.00001 - 1 with 0.001 step

2.2.2.2 Hidden Markov Model

The HMM is a stochastic model that represents a probability distribution for a sequence of AIS points. This state-space model is commonly used in marine vessel pattern classification because it is simple to employ and is not computationally intensive (Joo et al., 2013; Walker & Bez, 2010). At a high level, the HMM uses probabilities to determine whether an AIS point belongs to the wildlife-viewing or non-viewing class based only on the observed speed over ground (SOG) for that AIS point and the classification output for the AIS point immediately preceding this. A simplified explanation of the HMM in the context of this study is provided, and a more thorough explanation is found in Rabiner (1989).

The HMM contains hidden variables and observed variables. For this study, the HMM's hidden variables are the behaviour states that represent the wildlife-viewing and non-viewing classes for each AIS point, and the observed variables include both the AIS variables (e.g. SOG) for each point and the observed classes that were assigned to each AIS point based on observed whale encounter data. The behaviour states (classes) are considered hidden because they are not directly observed in the AIS data and must be inferred through Markovian transitions based on the

SOG for a sequence of AIS points. The progression of the wildlife-viewing and non-viewing states form the hidden Markov chain (Vermard et al., 2010). Given a set of hidden states S and a series of known observations O over time t, consider that the current state is S_t, the current observation is O_t, and the previous state is S_{t-1}. Then, the HMM assumes: (1) O_t is generated by hidden state S_t, and (2) given S_{t-1}, S_t is independent of states prior to S_{t-1} (De Souza et al., 2016; Toloue & Jahan, 2018). The HMM determines the probability of the point's state given its observation, factored as:

$$p(S_{1:T}|O_{1:T}) = p(S_1)\, p(O_1|S_1) \prod_{t=2}^{T} p(O_t|S_t)\, p(S_t|S_{t-1}) \tag{2.1}$$

where, $p(S_1)\ p(O_1|S_1)$ are the initial state and the initial probabilities, $p(O_t|S_t)$ are the emission probabilities, $p(S_t|S_{t-1})$ are the transition probabilities, and T is the last SOG observation read from the data (De Souza et al., 2016; Waterbolk et al., 2019).

The initial probabilities are the probability that the first point in a sequence (S_1) belongs to the wildlife-viewing state or non-viewing state, the emission probabilities are the probability density functions belonging to each state's observed variables (O), and the transition probabilities are the conditional probability of moving between states given the previous state (S_{t-1}). The probability computation was performed by the Viterbi algorithm (Grewal et al., 2019; Rabiner, 1989), which decoded O from SOG values to the most likely state sequence using the maximum log-likelihood score. Optimal emission, transition, and initial probabilities were derived from the sample dataset's SOG, timestamp, and observed classes (Table 2.1). To avoid overfitting the emission probabilities to the sample dataset, established continuous probability distribution functions were fit to each of the SOG profiles for both states. The optimal probability distribution functions were fitted by direct numerical minimization of the negative log-likelihood (Zucchini et al., 2016), which translates to the most likely probability distribution functions for these SOG profiles. Conditional probabilities of the observed classes were used to derive the transition probabilities (Kouemou, 2011; Murphy, 2012), and the initial probabilities were defined based on the first AIS point's observed class for each trip. After the emission, transition, and initial probabilities were defined, they were used to create the HMM and classify the points in each wildlife-viewing trip.

2.2.2.3 Logistic Regression

Logistic regression (LR), in the context of this study, is a discriminative classification algorithm widely used in binary classification problems across sectors. In the literature related to marine research, LR has been employed to utilize large AIS databases to classify vessel types (Sheng et al., 2018) and to detect patterns in vessel behaviour (Bye & Aalberg, 2018; Conn & Silber, 2013). A more thorough explanation of LR can be found in Cox (1958) or Schein & Ungar (2007).

LR was used to classify wildlife-viewing and non-viewing vessel behaviour using geometric, kinematic, temporal and geographical AIS variables. The geometric variables of distance between points and turning angle were derived from each AIS point's latitude and longitude position using the Euclidean distance metric and the cosine formula, respectively (Muthu, 2015). All variables required standardization as their units of measurement varied. This was done by scaling the variable to values between 0 and 1 based on each variable's original minimum and maximum values. For example, COG has values between 0° and 360°, so a set of COG values of {0°,180°,360°} would map to the set {0,0.5,1}. A pairwise covariate assessment using the Pearson Correlation coefficient was conducted to assess covariate independence.

Given a training dataset $\{(x_{i1}, x_{i2}, x_{i3}, \ldots, x_{iN}, y_i)\}_{i=1}^{M}$, where the set of variable predictors are represented as an *N*-dimensional vector $X = (x_1, x_2, x_3, \ldots, x_n) \in R^N$, a probability model was defined for the binary outcome y using the sigmoid function $\sigma(z)$, which maps any real value *z* to the interval [0,1] representing genuine probabilities. When *z* > 0.5, the sigmoid curve rapidly approaches one, and when *z* ≤ 0.5 the sigmoid curve rapidly approaches zero. Therefore, when *z* > 0.5 the binary class of *y* = 1 (wildlife-viewing) was detected. Conversely y = 0 for *z* ≤ 0.5. Let $z = (w \cdot X)$, then the probability model is defined as:

$$p(y = 1|X) = \sigma(w \cdot X) = \frac{1}{1 + exp^{-(w \cdot X)}} \tag{2.2}$$

$$p(y = 0|X) = 1 - \sigma(w \cdot X) = \frac{1}{1 + exp^{(w \cdot X)}} \tag{2.3}$$

with the corresponding likelihood function:

$$\begin{aligned} p(y|X) &= \prod_{i=1}^{M} \sigma(w \cdot X_i)^{y_i} \, (1 - \sigma(w \cdot X_i))^{(1-y_i)} \\ &= \prod_{i=1}^{M} \sigma(w \cdot X_i)^{y_i} \, \sigma(-w \cdot X_i)^{(1-y_i)} \end{aligned} \tag{2.4}$$

where, w is the regression coefficients, an *n*-dimensional vector of weights applied to each variable. A maximum likelihood estimation determines the optimal w using the observed classes *y* and *X* with the log-likelihood function:

$$l(w) = \ln p(y|X) = \sum_{i=1}^{M} (y_i \ln[\sigma(X_i)] + (1 - y_i) \ln[1 - \sigma(X_i)]) \tag{2.5}$$

This function is maximized using the gradient descent optimization method to define the most likely value for w, as defined by

$$w^{*} = max_{w}\big(l(w)\big) \tag{2.6}$$

where $w*$ is the optimized w, max_w is the maximization through gradient descent, and $l(w)$ is Equation 2.5. Gradient descent is an unconstrained numerical optimization algorithm to solve for w^{*} (Schein & Ungar, 2007; Sheng et al., 2018). This study employs the coordinate descent algorithm (Wright, 2015), which cyclically iterates through one variable at a time to define the optimal w for that variable. This is done by linearly iterating along the coordinate direction at a step size until convergence to the local minimum, or until the maximum number of iterations is reached (Murphy, 2012; Wright, 2015).

To avoid overfitting the LR model to X, a regularization penalty is commonly applied to w^{*}. This study employs the L1-norm ($\|w\|_1$) regularization penalty which sums the absolute value of w and applies this and the regularization coefficient (λ) to $l(w)$ (Chen et al., 2018; Murphy, 2012). With its application, the optimization problem becomes:

$$w^{*} = max_{w}(l(w) - \lambda\|w\|_1) \tag{2.7}$$

Defining an appropriate regularization strength (C), which is the inverse of λ (Table 2.1), is crucial as a larger λ shrinks w closer to zero leading to underfitting, but a smaller λ may lead to overfitting. Therefore, the observed classes were used to validate revisions to the parameter C. First, the AIS data were used to determine a reasonable range of values for C, and a fivefold cross-validation grid search defined the LR model's optimal parameter values (C, maximum iterations, and tolerance) iteratively based on the highest F-Score (Table 2.1 and 2.2). The F-Score was used as it is a preferred test when optimizing binary classifier parameters (Hossin & Sulaiman, 2015;

Jansche, 2005). The validated parameters were used to generate the LR model, and the model was fit to a 70% stratified subset of the AIS data to derive w^*. The final LR model with optimal w^* then classified each AIS point with the remaining 30% subset used for statistical model validation.

2.2.3 Model Validation and Evaluation

Each of the three models evaluated in this study have their own performance metrics that provide insight into its accuracy. To facilitate comparison, however, an inter-model validation and evaluation framework was adopted from Çiğşar and Ünal (2019) that provides consistent performance metrics for different model types. This validation framework included statistical performance metrics (Table 2.2) of overall accuracy, precision, recall, specificity, receiver operating characteristic area under the curve (ROC AUC), root mean square error (RMSE), F-Score, and Kappa statistic for each model (Hossin & Sulaiman, 2015). A confusion matrix was generated from the observed classes and modelled classes for each model and used to calculate the performance metrics (Table 2.2). More details about these metrics can be found in Çiğşar and Ünal (2019), Cohen (1960), and Hossin and Sulaiman (2015).

Table 2.2: Classifier validation performance metrics, adapted from Hossin and Sulaiman (2015).

Confusion Matrix		Observed Classes	
		Wildlife-viewing	Non-viewing
Modelled Classes	Wildlife-viewing	True positive (tp)	False positive (fp)
	Non-viewing	False negative (fn)	True negative (tn)

Metric	**Formula**	**Focus**
Overall Accuracy	$\frac{tp + tn}{tp + tn + fn + fp}$	Ratio of correct classifications over the total number of observations.

Precision	$\frac{tp}{tp + fp}$	Fraction of correct positive classifications over the total positive observations.
Recall	$\frac{tp}{tp + fn}$	Fraction of correct positive classifications over the total observations in the class.
Specificity	$\frac{tn}{tn + fp}$	Fraction of correct negative classifications.
F-Score	$\frac{2 * (precision * recall)}{precision + recall}$	Harmonic mean between precision and recall.
RMSE	$\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$	Standard deviation of the difference between observed and modelled classes.

Alongside each model's performance metrics, an inter-model evaluation was conducted using the Kappa and McNemar statistics. The Kappa statistic examines the level of agreement between classifiers while accounting for the possibility of random disagreement (Cohen, 1960; Viera & Garrett, 2005). The McNemar statistic evaluates the level of disagreement (i.e. the consistency in the errors) (Dietterich 1998). Please see the Appendix for a more detailed definition of these metrics.

2.3 Results

The final dataset included 19 wildlife-viewing trips consisting of 1,770 AIS points, with 36% belonging to the observed wildlife-viewing class and 64% the non-viewing class. During the process of analyzing the models, an outlier wildlife-viewing trip was identified and removed from consideration (Fig. 2.3). Each of the three models classified a wildlife-viewing event along the transnational marine boundary in the Strait of Juan de Fuca for ~65 minutes with no recorded whale sightings. This event is unlikely to be misclassified due to terrestrial or avian species observations because of its distance from shore, so it was removed from the dataset and excluded from the models.

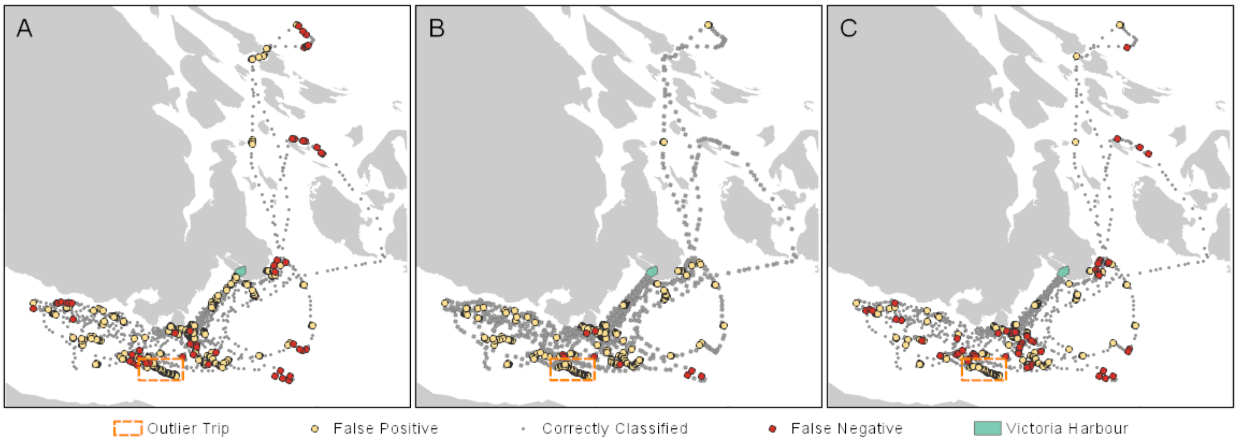

Figure 2.3: Classification results for density-based spatial clustering application with noise (A), hidden Markov model (B), and logistic regression (C).

After removing the outlier, the degree to which each model's performance is sensitive to the assumption that wildlife-viewing events occur at the known wildlife observation sites of Race Rocks Ecological Reserve and the waters between Chatham Islands and Discovery Island was assessed (Fig. 2.1). This was done by redefining each model's optimal parameters based solely on the observed data for whale sightings, excluding the known wildlife-viewing sites, and comparing their statistic validity with and without this assumption.

2.3.1 Density-Based Spatial Clustering Application with Noise

The iterative optimization of ε, *m*, and the NNA, based on the highest F-Score, resulted in an ε parameter value of 0.004°, a *m* of two points, and the brute-force NNA. The parameter ε is measured in fractions of degrees as DBSCAN classifies each AIS point based on its latitude and longitude. Its sensitivity to outliers was apparent when the maximum F-Score changed the optimal ε from 0.003° to 0.004° with their removal (Fig. 2.4). Notably, the optimal minimum sample remained the same (Fig. 2.4). The optimal *m* was chosen based on the highest F-Score and the brute-force NNA, which enumerates through all pairwise distances between AIS points and selects the points within the ε radius, was chosen based on its frequency with respect to ε. The optimal DBSCAN parameters without including the known wildlife-viewing sites were a ε of 0.003°, *m* of two points, and the brute-force NNA. Typically, ε is found by plotting the pairwise distances between points as a curve and selecting the distance at the location where this curve turns sharply upward. This method resulted in an ε of 0.04° and an overall F-Score of 0.50, which was insufficient in classifying wildlife-viewing events.

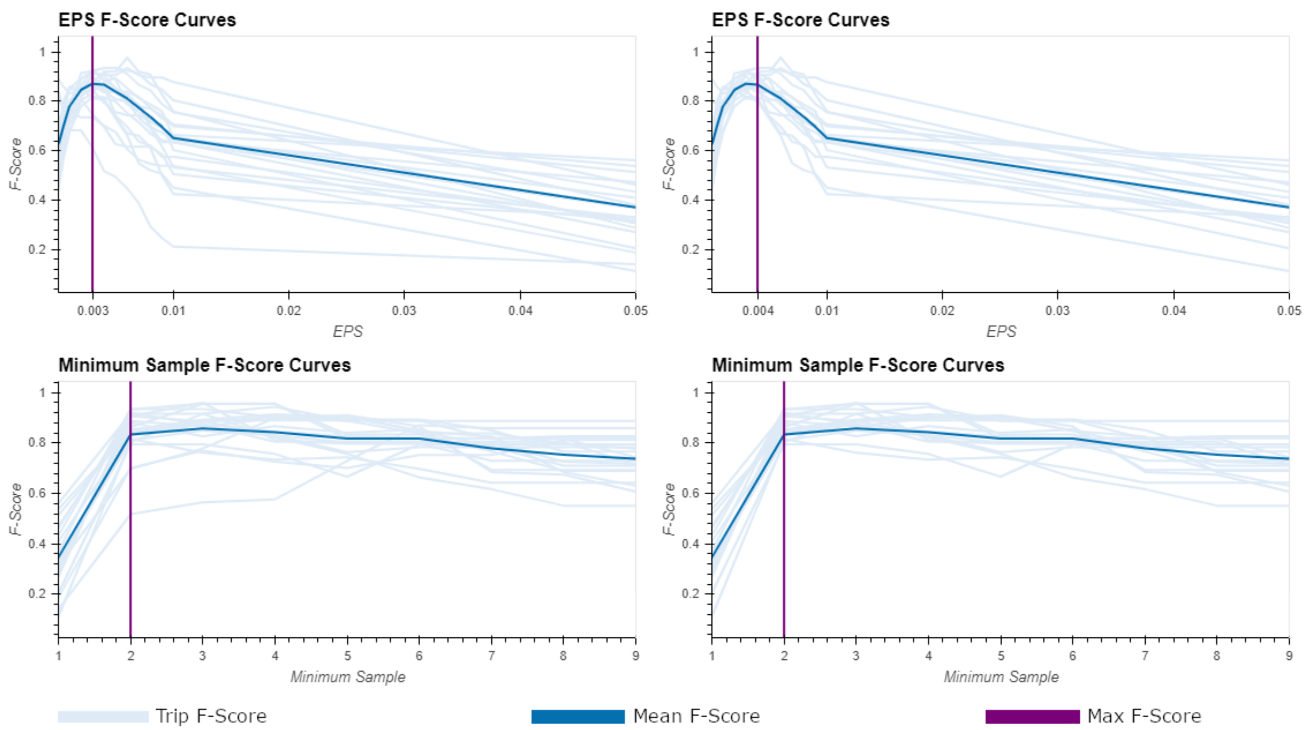

Figure 2.4: F-Score valitaion curves for epsilon (EPS) and the minimum number of samples. Density-based spatial clustering application with noise parameter value selection with outliers (left) and without (right).

The DBSCAN was simple to employ and iteratively generated classes for each individual trip. Due to the heavy clustering of AIS points near the harbour and along the southern tip of Vancouver Island, this classifier was not performed on the dataset in its entirety (Fig. 2.2).

Nonetheless, the DBSCAN generation and classification executed for all 19 trips in 1.06 seconds with an overall accuracy of 0.85, F-Score of 0.82, and ROC AUC score of 0.86. When the wildlife-viewing assumptions at known wildlife-viewing sites (Fig. 2.1) were removed, the DBSCAN was fit with an ε of 0.003°, and the overall accuracy dropped to 0.78, and the F-Score and ROC AUC score dropped to 0.68 and 0.79 respectively.

2.3.2 Hidden Markov Model

The initial, transition, and emission probabilities that create the HMM were derived from the sample data. The initial probabilities were defined from temporally sorted AIS data, and the transition probabilities were derived using conditional probability (Table 2.3). The initial behaviour state (class) for all 19 trips in the sample dataset were observed as non-viewing points resulting in an initial probability of 1.0 for non-viewing and 0.0 for wildlife-viewing.

Table 2.3: Hidden Markov model initial and transition probability matrices.

<table>
<tr><th colspan="2">Initial Probabilities</th><th></th><th colspan="3">Transition Probabilities</th></tr>
<tr><td>Wildlife-viewing</td><td>Non-viewing</td><td></td><td></td><td>Non-viewing</td><td>Wildlife-viewing</td></tr>
<tr><td rowspan="2">0.0</td><td rowspan="2">1.0</td><td rowspan="2"></td><td>Non-viewing</td><td>0.95</td><td>0.05</td></tr>
<tr><td>Wildlife-viewing</td><td>0.09</td><td>0.91</td></tr>
</table>

Comparable to De Souza et al. (2016), the SOG distributions for each class were distinctive (Fig. 2.5). The resulting emission probability density functions are represented as a double gamma distribution function for the non-viewing class, and a generalized inverse Gaussian distribution function for the wildlife-viewing class. The removal of outliers did not result in any changes to the distribution functions, but the removal of assumed wildlife-viewing at known wildlife-viewing sites (Fig. 2.1) resulted in the non-viewing class fitting to a Gauss hypergeometric distribution and the wildlife-viewing class fitting to an exponential modified normal distribution (Fig. 2.6).

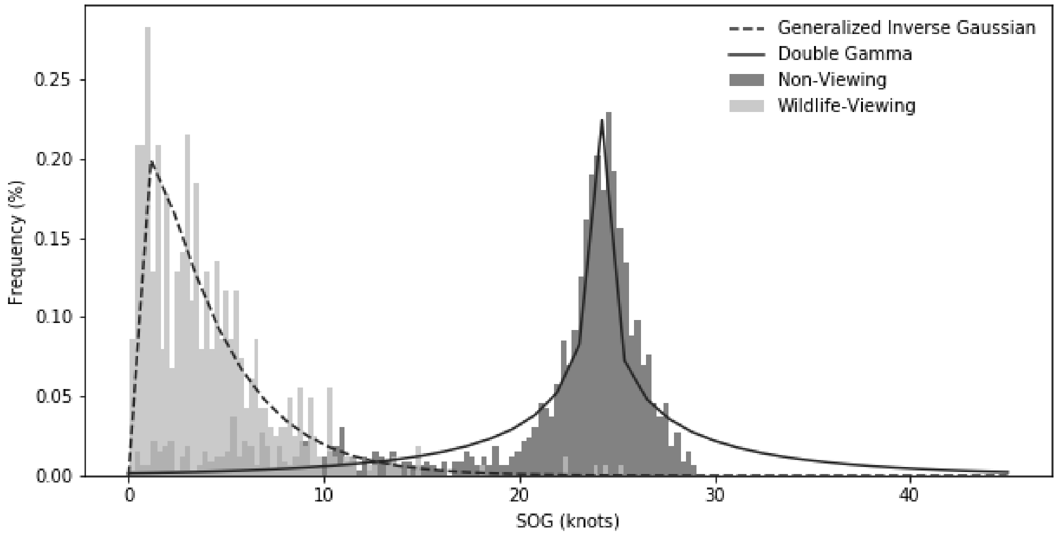

Figure 2.5: Speed profiles. Histogram of speed over ground (SOG) observations with the best fitting probability density functions for each hidden state.

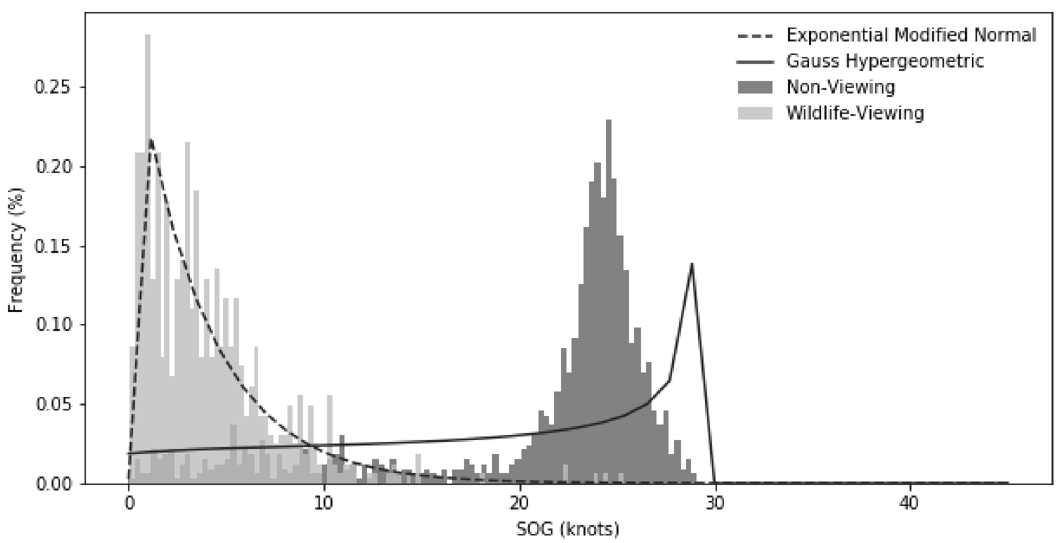

Figure 2.6: Speed profiles with the points assumed to be wildlife-viewing at known wildlife observation sites removed. Histogram of speed over ground (SOG) observations with the best fitting probability density functions for each hidden state.

Although the negative log-likelihood of the Gauss hypergeometric distribution suggests it fits the non-viewing class, without assumptions, the histogram shows this is not the best representation of the data (Fig. 2.6). In addition, for the wildlife-viewing class, the negative log-likelihood of the exponential modified normal distribution was only slightly smaller (-0.13) than the generalized inverse Gaussian distribution. Therefore, the HMM model was implemented with the double gamma distribution and generalized inverse Gaussian distribution for both iterations, with and without assumptions.

The HMM was developed with custom emission probabilities to generate classes for all 19 trips at once, speeding up model generation and classification. It took 4.02 milliseconds to execute the HMM classifier which resulted in an overall accuracy of 0.89, F-Score of 0.87, and ROC AUC score of 0.91. The overall accuracy and ROC AUC score remained high at 0.81 and 0.86 respectively when the wildlife-viewing assumption at known wildlife observation sites (Fig. 2.1) was removed, but the F-Score dropped to 0.74.

2.3.3 Logistic Regression

The iterative optimization of *C*, maximum iterations, and tolerance for stopping, based on the highest F-Score, resulted in a *C* of 1.0, maximum iterations of 100, and stopping tolerance of 0.0001. Notably, the grid search ran multiple times to draw these conclusions and output slightly different values for *C* ranging from 1.0 to 2.5. Thus, the sensitivity of the F-Score to *C* was explored to confirm its selection (Fig. 2.7). Considering a smaller *C* can lead to underfitting, while a larger *C* leads to overfitting, a *C* of 1.0 captures 97% of the variability in the LR model's F-Score without overfitting to the data. This observation holds with and without assuming wildlife-viewing at known wildlife observation sites (Fig. 2.1); therefore, the same optimized parameters were used for both iterations, with and without assumptions.

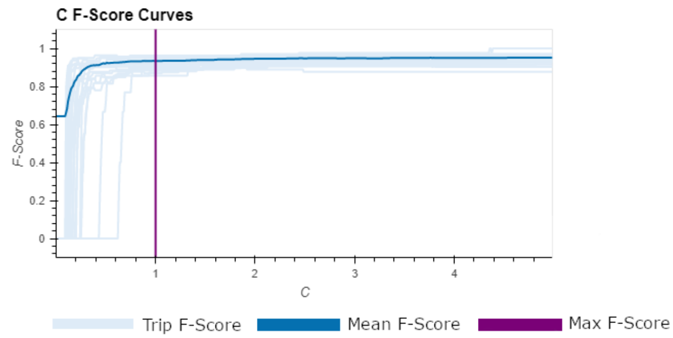

Figure 2.7: F-Score validation curve for regularization strength (*C*). Logisitic regression parameter value selection.

Prior to model optimization, AIS variable selection was assessed using the Pearson correlation coefficient, which resulted in a correlation coefficient between course over ground (COG) and heading > 0.80. The LR model was tested using both the COG and heading, which resulted in nominal differences in overall accuracy, F-Score and ROC AUC score. Therefore, COG was included in the LR model because it relates the marine vessel to the Earth and not the intended vessel direction. The final variables used in the LR classification model were epoch (timestamp converted to Unix time), SOG, COG, latitude, longitude, turning angle, and distance.

The LR model was generated using the optimal parameters and found the regression coefficients through coordinated descent by fitting the stratified training data subset (70%) to the model before classifying all AIS points at once. The regression coefficients (Fig. 2.8) provide an understanding on how influential each variable is on the LR model. Model generation and classification took 2.97 milliseconds and, using the test subset, resulted in an overall accuracy, F-Score, and ROC AUC score of 0.89, 0.86, and 0.90 respectively. When the wildlife-viewing assumptions at known wildlife-viewing sites were removed the overall accuracy, F-Score, and ROC AUC score dropped to 0.85, 0.77, and 0.81.

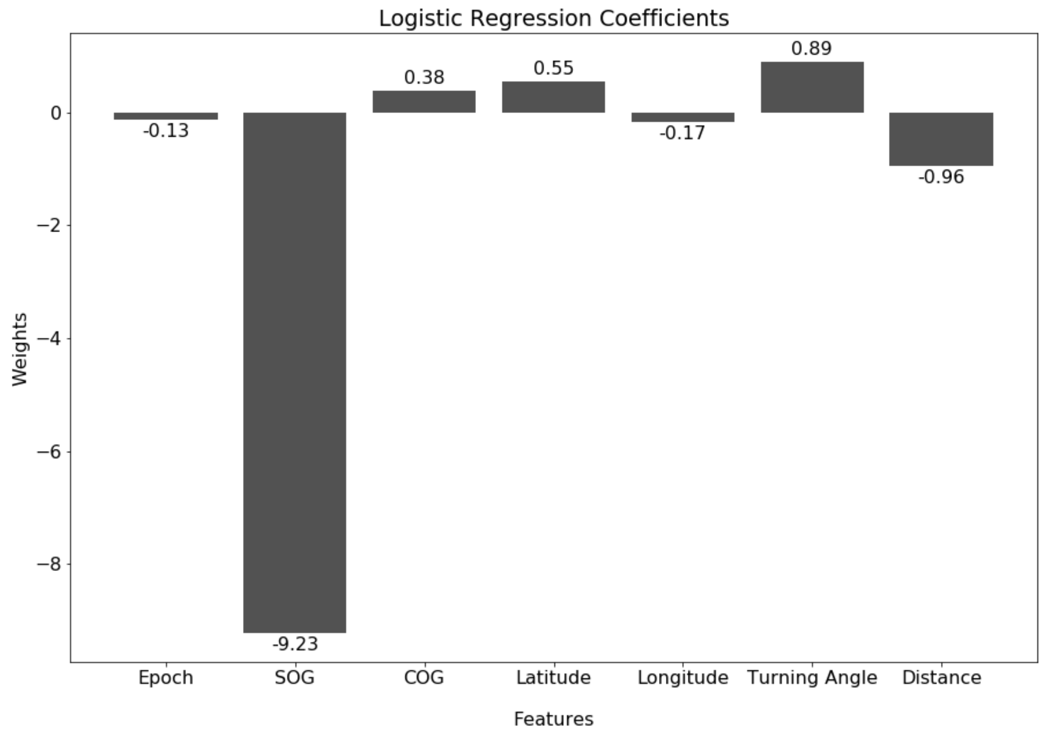

Figure 2.8: Logistic regression coefficients derived using the 70% stratified training data subset.

2.3.4 Model Validation and Evaluation

The individual statistical validation for all three models resulted in overall accuracy, F-score, and ROC AUC score >0.80 (Table 2.4). Although the DBSCAN performed well, its overall accuracy, F-Score, ROC AUC score, and Kappa coefficient were all lower than the HMM and LR model. The HMM and LR outperformed DBSCAN, but the HMM and LR measures differ slightly. The similarities diverge with the HMM having greater recall and the LR being more precise. Additionally, the F-Score, ROC AUC score, and Kappa coefficient favor the HMM. The RMSE was relatively small overall, but HMM and LR had smaller errors than DBSCAN. The individual assessment of the Kappa statistic concluded that there was substantial agreement >0.60 (p-value < 0.05) between the observed classes and the modelled classes for all classification models.

Table 2.4: Individual classifier statistical performance metrics.

				Without Wildlife-viewing Assumption		
	DBSCAN	**HMM**	**LR**	**DBSCAN**	**HMM**	**LR**
Overall Accuracy	0.8492	0.8904	0.8921	0.7847	0.8051	0.8475
Precision	0.7348	0.7749	0.8138	0.5896	0.5973	0.8615
Recall	0.9180	0.9861	0.9133	0.8079	0.9723	0.6935
Specificity	0.8096	0.8354	0.8780	0.7755	0.7383	0.9359
F-Score	0.8162	0.8678	0.8607	0.6817	0.7400	0.7684
ROC AUC Score	0.8638	0.9107	0.8966	0.7917	0.8553	0.8147
Kappa Coefficient	0.6909	0.7765	0.7731	0.5250	0.5982	0.5318
RMSE*	0.3884	0.3311	0.3285	0.4640	0.4305	0.4389

***Minimization of the RMSE is preferred.**

The removal of the assumption that wildlife-viewing occurred at known wildlife-viewing sites (Fig. 2.1) resulted in a drop in most metrics except the specificity and precision of LR. Although the DBSCAN was less successful than the HMM and LR, its overall accuracy, F-Score, and ROC AUC score remained above 0.68. The HMM and LR overall accuracy and ROC AUC scores remained above 0.80, suggesting successful classification, but the F-Score fell below 0.77 for all classifiers. The F-Score now favors LR instead of the HMM when the assumptions are removed. The RMSE for each classifier increased by at least 0.08, with the HMM having the smallest RMSE. The Kappa coefficients dropped below 0.60 (p-value < 0.05), which indicates moderate agreement between the observed classes and modelled classes.

The inter-model evaluation using the Kappa statistic determined there was near perfect agreement between DBSCAN and HMM, and HMM and LR with Kappa coefficients ≥0.81, and that DBSCAN and LR agree substantially with a Kappa coefficient ~0.80 (Table 2.5). Although the Kappa statistic results were significant with p-values < 0.05, the McNemar statistic for the DBSCAN and HMM comparison does not have the same significance (p-value > 0.27). Therefore, it cannot be stated with certainty that the errors rates between the DBSCAN and HMM differ. Conversely, the McNemar statistic comparing the LR model's errors to the DBSCAN, or HMM, resulted in significantly different error rates (p-value < 0.05).

Table 2.5: A matrix of inter-classifier statistical measures for Kappa (Left), McNemar (Right).

Kappa Statistic (p-value)			*McNemar Statistic (p-value)*		
	HMM	**LR**		**HMM**	**LR**
DBSCAN	0.8124 (0.0000)	0.7956 (0.0000)	**DBSCAN**	75 (0.2757)	48 (0.0000)
HMM		0.8729 (0.0000)	**HMM**		7 (0.0000)

An assessment of the spatial trends between classification models was conducted by plotting the misclassified false positive and false negative points and comparing them to the correctly classified points (Fig. 2.9). Common misclassification occurred on the periphery of wildlife sighting events, that is, the points leading up to or following a series of wildlife-viewing points. Most of the misclassifications were false positive (≥ 8% for all models), which are points that were classified as wildlife-viewing when the observed class was non-viewing. Common false positive misclassification occurred near Race Rocks Ecological Reserve (Fig. 2.1) and with two wildlife-viewing events directly south of the Victoria, B.C. harbour. The number of false negative classifications remained lowed, with the HMM having the lowest at 1%. Where the false negative misclassifications occurred varied for each model. Although the trade off between false positive and false negative misclassifications for all models are captured in the F-Score and ROC AUC score, many studies aim to reduce false positive classifications due to their frequency (De Souza, et al., 2016; Sidibé & Shu, 2017). In the context of wildlife-viewing vessel behaviour, false positive classifications may occur during difficult to navigate areas, which could skew vessel movement analyses, or information gathered through application of these models.

Notably, the HMM classified more straight sailing wildlife-viewing points correctly compared to DBSCAN and LR. For example, the two relatively straight sailing wildlife-viewing trips in the northeast quarter of the maps in Figure 2.9 are better classified by the HMM, even though the LR performs similarly in other regions. In particular the higher density of points in narrow channels, near the Victoria Harbour, and along the southern tip of Vancouver Island show the misclassification of points by DBSCAN (Fig. 2.9).

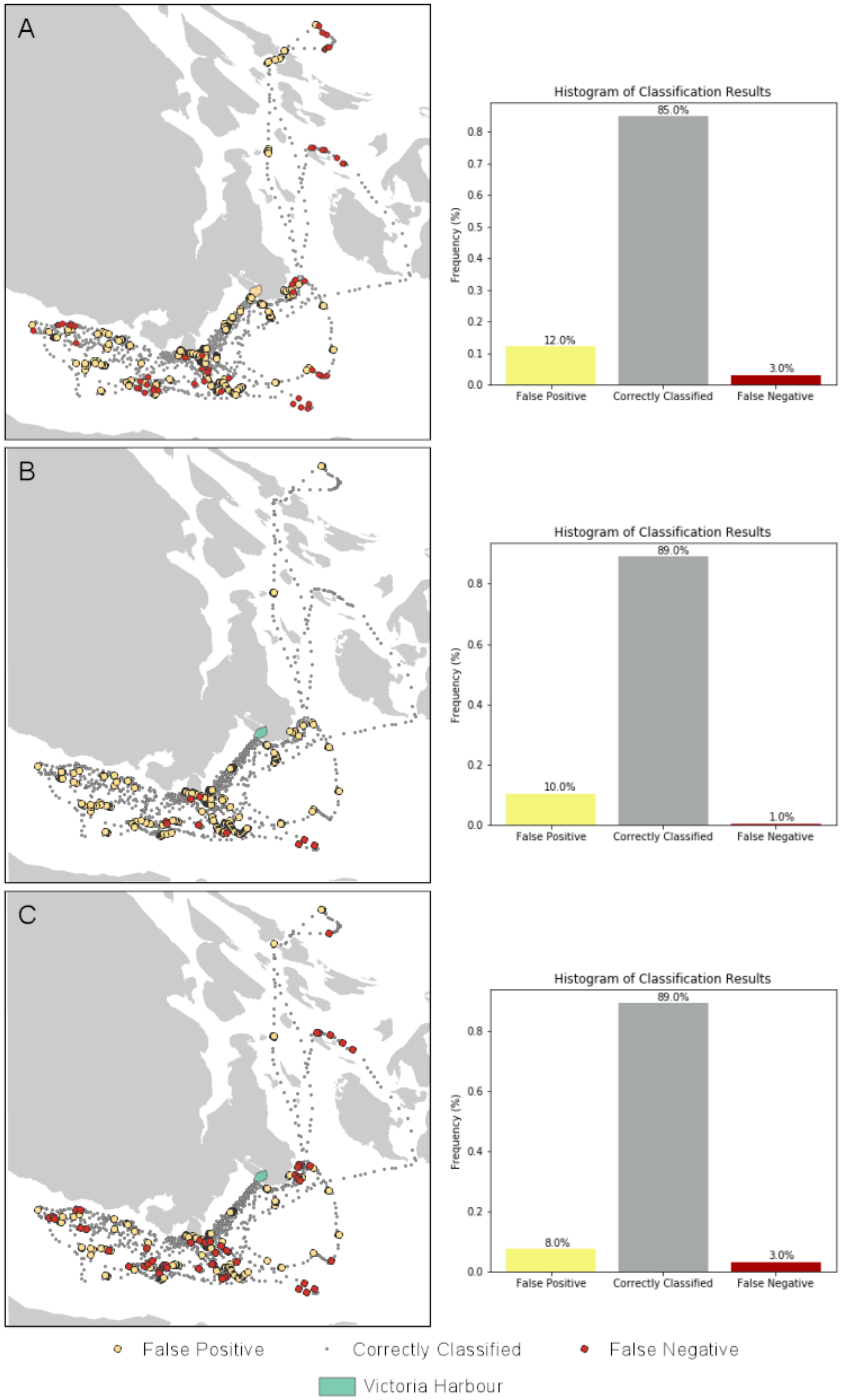

Figure 2.9: Classification results for density-based spatial clustering application with noise (A), hidden Markov model (B), and logistic regression (C) after outlier removal.

2.4 Discussion

The overall objective of this study was to evaluate established classification models for their feasibility of classifying wildlife-viewing behaviour of small marine vessels utilizing AIS data. The findings from this study suggest that all three classification models are suitable approaches to classifying wildlife-viewing events, although they demonstrate differences in model development and classification performance. Model performance and adaptation is highly dependent on parameter optimization (Cao et al., 2016; Hossin & Sulaiman, 2015). Without optimal parameters defined based on the AIS data and region specific domain knowledge, the models run the risk of underperforming and may have misleading results. Specifically, the results of this study demonstrated the model's sensitivity to parameter optimization, model structure, and assumptions, and that regional specific domain knowledge is required for accurate model development.

As the models evaluated in this study have not been previously utilized in the context of wildlife-viewing vessel behaviour, the correct definition of each model's optimal parameters was necessary in performing accurate classifications without overfitting to the sample dataset. In particular, DBSCAN proved sensitive to outliers, which was observed while determining the optimal radius between points and minimum number of points in a cluster (Fig. 2.4). This optimization process was time consuming, computationally complex, and required domain expert knowledge. The optimization method used, F-Score maximization, may have led to overfitting of the DBSCAN model to the sample dataset, but other methods used to determine the optimal radius resulted in the DBSCAN underperforming. The maximization of the F-Score was used for optimization because it was found to be a suitable optimization method for binary classification parameters (Hossin & Sulaiman, 2015), and for consistency across the three models. The limitations of the DBSCAN were amplified when the model required iterative implementation for each trip to separate the points temporally. Regardless of cautious parameter tuning, and these limitations, the DBSCAN was found to be a suitable method for classifying wildlife-viewing events.

The HMM was most sensitive to properly defined emission probability density functions. Considering the SOG probability distributions are distinctive between the two classes, it could be argued that a SOG threshold be simply employed instead (Fig. 2.5 and 2.6). In fact, Almunia et al. (2020) employ a speed threshold along with bathymetry data to determine when whale-watching vessels are observing a single whale species in the open waters off of the Canary Islands, Spain, which was sufficient in the context of their study. However, it is argued here that SOG and bathymetry thresholds do not capture the temporal aspect of each wildlife-viewing trip, or the variability in speed that occurs when the vessel is maneuvering alongside different whale species or through difficult to navigate regions within the coastal and bathymetric complexity of the Salish Sea, such as crossing paths with other vessels (e.g. ferries or shipping vessels), or in deep narrow channels. To account for the variable speed during non-viewing maneuvers a single established continuous emission probability density function for each class was defined that avoided overfitting the model to the sample dataset, but a combination of two or more probability density functions may better represent the probability of the non-viewing class at slower speeds. The Gauss hypergeometric distribution attempts to fit the slower speeds within the non-viewing class without assumptions (Fig. 2.6), but fails to fit to the higher speeds. Due to the small size of the sample dataset, the implementation of multiple probability density functions for the non-viewing class was not explored in this study. Aside from defining the optimal emission, initial, and transition probabilities, the HMM does not require extensive parameter tuning, and was relatively computationally simple. Its use in modelling movement across disciplines and its performance in this study as an accurate classifier proves the HMM's suitability in classifying wildlife-viewing vessel behaviour.

LR historically extends beyond marine vessel research, and comes with the caveat that it is sensitive to how each of the AIS variables are represented. Therefore, extensive variable exploration leading up to model creation and generation was required to ensure the correct data representation, and to analyze which variables influence the classification of wildlife-viewing vessel behaviour. Similar to DBSCAN, LR required time consuming and computationally complex parameter optimization, but, unlike DBSCAN, was simply employed to classify wildlife-viewing events for all trips at once. LR was implemented with geometric, temporal, and kinematic AIS variables in an attempt to provide a robust and extensive classification method. The difference between LR and HMM performance was negligible regardless of the inclusion of additional data. A benefit of LR is that the level of influence each variable has on the classification is output as regression coefficients, providing additional information into which AIS variables are essential in the classification of wildlife-viewing vessel behaviour (Fig. 2.8). The significance of speed to the classification of wildlife-viewing events supports the variable selected for the HMM.

If the goal is accurate classification, the HMM and LR supersede DBSCAN. Considering the data processing requirements leading up to LR, it is argued that the HMM is the easiest to employ, requires only a single AIS variable (SOG), and exceeded the LR model in the F-Score, ROC AUC score, and Kappa statistic. The performance of the HMM declines marginally with the removal of the assumption that wildlife-viewing events occurred in the Race Rocks Ecological Reserve and the waters between Chatham Islands and Discovery Island (i.e. wildlife observation sites frequently visited by commercial ecotourism vessels to observe seals, sea lions, avian or terrestrial [e.g. eagle or wolf] species). Arguably, with and without assumptions, the HMM is most suitable when accurate classification is the goal.

Regional specific domain knowledge and an understanding of AIS is required for model implementation. For example, Almunia et al. (2020) classify whale-watching vessels with a speed threshold of 2.5 knots, but this speed threshold is not applicable to this study. This study determined that wildlife-viewing events may approach ~12.7 knots with most wildlife-viewing occurring at or below 7 knots based on the SOG distributions (Fig. 2.5 and 2.6). As SOG collected by the AIS is based on the vessel's throttle speed, the SOG could be considered to accurately represent guidelines on speed while observing whales (Fisheries and Oceans Canada, 2021; Washington State Legislature, 2021). This finding is important as it confirms that the commercial ecotourism vessel used to develop these models adhered to wildlife-viewing regulations, which speaks to how the methods developed in this study have the potential to inform managers of regulation compliance.

The AIS data required to accurately classify wildlife-viewing vessel behaviour needs to have a sufficiently high spatial and temporal resolution, which requires an AIS receiver network with enough coverage of the Salish Sea. A large spatial or temporal gap in AIS data would skew the DBSCAN clusters, misrepresent the HMM sequence, and provide inaccurate geometric measures for LR. Moreover, AIS transponder antenna height influences AIS message reception (Last et al., 2015). This study considered a single whale-watching vessel with relatively high spatial and temporal resolution, collected through an AIS network with relatively good coverage, few gaps in the data, and a consistent antenna height. The variability of AIS antenna height when these models are applied to other vessels will influence the AIS data quality and resolution (Last et al., 2015). Furthermore, providing recommendations from the results of these models requires region specific domain knowledge of the wildlife observation regulations and guidelines in addition to knowledge on the structure and limitations of AIS data.

Generalized and deployable models for the classification of wildlife-viewing vessel behaviour across a large geographic area may not be feasible; therefore, future work analyzing how these models perform, and how the parameterization of these models change based on geographic region and scale would be valuable. Also worth exploring is model enhancement, such as the implementation of a multidimensional DBSCAN model that incorporates speed, or an HMM with more specific emission probability density functions fit to the non-viewing class. Prior to implementing any of the models explored in this study, it is recommended that region specific domain knowledge be attained, AIS data quality and resolution be assessed, and development time and researcher programming knowledge be considered.

The classification of wildlife-viewing vessel behaviour has various applications that benefit marine management, such as informing policies on marine mammal protections, environmental conservation, and vessel traffic management. This could be done through examining the spatial-temporal characteristics of vessel movement while observing wildlife to determine if these vessels are adhering to marine policies, or to provide recommendations on how these policies may be revised. Consequently, industry specific regulation and guideline compliance and enforcement could be investigated using these classification outputs. Moreover, marine mammal conservationists, researchers and managers may benefit from future studies examining how the classification of wildlife-viewing vessel behaviour may be used as a proxy for the probable whale locations.

This study's appraisal of established models for the classification of wildlife-viewing behaviour using whale-watching vessels aims to contribute to scientific endeavors in whale-watching vessel movement, marine traffic management, impacts to marine ecosystems, wildlife conservation, and marine spatial planning. Analyses of wildlife-viewing vessel movement contribute to efforts for managing and mitigating impacts to the marine environment and marine mammals, to support the preservation of known whale habitats, and may be a proxy for the analysis of non-commercial, or unknown, wildlife-viewing vessels. Furthermore, industry specific applications of these models can provide more informative fleet management and aid in assessing wildlife-viewing regulation compliance.

2.5 Literature Cited

Alessandrini, A., Alvarez, M., Greidanus, H., Gammieri, V., Arguedas, V. F., Mazzarella, F., Santamaria, C., Stasolla, M., Tarchi, D., & Vespe, M. (2017). Mining Vessel Tracking Data for Maritime Domain Applications. *IEEE International Conference on Data Mining Workshops*, 361–367. https://doi.org/10.1109/ICDMW.2016.0058

Almunia, J., Delponti, P., & Rosa, F. (2020). Using big data to estimate whale watching effort. *bioRxiv*. https://doi.org/10.1101/2020.11.30.403923

Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J., Connor, R., Heithaus, M., Watson-Capps, J., Flaherty, C., & Krützen, M. (2006). Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conservation Biology*, 20(6), 1791–1798. https://doi.org/10.1111/j.1523-1739.2006.00540.x

Bye, R. J., & Aalberg, A. L. (2018). Maritime navigation accidents and risk indicators: An exploratory statistical analysis using AIS data and accident reports. *Reliability Engineering and System Safety*, 176, 174–186. https://doi.org/10.1016/j.ress.2018.03.033

Canadian Shipping Act: Regulations Amending the Navigation Safety Regulations (Automatic Identification Systems), SOR/2019-100 (2019). *Canada Gazette Part II*, 153(9). https://gazette.gc.ca/rp-pr/p2/2019/2019-05-01/html/sor-dors100-eng.html

Cao, J., Cambria, E., Lendasse, A., Miche, Y., & Vong, C. M. (Eds.). (2016). Proceedings in Adaptation, Learning and Optimization. *Proceedings of ELM-2016*. https://doi.org/10.1007/978-3-319-57421-9

Cazzanti, L., & Pallotta, G. (2015). Mining maritime vessel traffic: Promises, challenges, techniques. *OCEANS 2015 - Genova*, 1–6. https://doi.org/10.1109/OCEANS-Genova.2015.7271555

Charles, C., Gillis, D., & Wade, E. (2014). Using hidden Markov models to infer vessel activities in the snow crab (*Chionoecetes opilio*) fixed gear fishery and their application to catch standardization. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(12), 1817–1829. https://doi.org/10.1139/cjfas-2013-0572

Chen, Z., Xue, J., Wu, C., Qin, L. Q., Liu, L., & Cheng, X. (2018). Classification of vessel motion pattern in inland waterways based on Automatic Identification System. *Ocean Engineering*, 161, 69–76. https://doi.org/10.1016/j.oceaneng.2018.04.072

Çiğşar, B., & Ünal, D. (2019). Comparison of Data Mining Classification Algorithms Determining the Default Risk. *Scientific Programming*, 2019. https://doi.org/10.1155/2019/8706505

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, XX(1), 37–46.

Conn, P. B., & Silber, G. K. (2013). Vessel speed restrictions reduce risk of collision-related mortality for North Atlantic right whales. *Ecosphere*, 4(4) :43. https://doi.org/10.1890/ES13-00004.1

Cox, D. R. (1958). The Regression Analysis of Binary Sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215–232. https://doi.org/10.1111/j.2517-6161.1958.tb00292.x

De Souza, E. N., Boerder, K., Matwin, S., & Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PLoS ONE*, 11(7). https://doi.org/10.1371/journal.pone.0158248

Erbe, C. (2002). Underwater noise of whale-watching boats and potential effects on killer whales (Orcinus orca), based on an acoustic impact model. *Marine Mammal Science*, 18(2), 394–418.

Ester, M., Kriegel, H.-P., Sander, J., & Xiaowei, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *KDD-96 Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 96, 226–231. https://doi.org/10.1016/B978-044452701-1.00067-3

Feng, J., Xu, H., Mannor, S., & Yan, S. (2014). Robust logistic regression and classification. *Advances in Neural Information Processing Systems*.

Fisheries and Oceans Canada. (2020). Watching marine wildlife. https://www.dfo-mpo.gc.ca/species-especes/mammals-mammiferes/watching-observation/index-eng.html

Fisheries and Oceans Canada. (2021). 2021 management measures to protect Southern Resident killer whales. https://www.pac.dfo-mpo.gc.ca/fm-gp/mammals-mammiferes/whales-baleines/srkw-measures-mesures-ers-eng.html

Fouda, L. (2012). Noisy Neighbours-using Automatic Identification System (AIS) and passive acoustic monitoring data to measure individual vessel source levels in critical whale habitat.

Frankel, A. S., & Gabriele, C. M. (2017). Predicting the acoustic exposure of humpback whales to cruise and tour vessels in Glacier Bay, Alaska. *Endangered Species Research*, 34, 397–415.

Fraser, M. D., McWhinnie, L. H., Canessa, R. R., & Darimont, C. T. (2020). Compliance of small vessels to minimum distance regulations for humpback and killer whales in the Salish Sea. *Marine Policy*, 121, 104171. https://doi.org/10.1016/j.marpol.2020.104171

Giles, D. A., & Koski, K. L. (2012). Managing vessel-based killer whale watching: A critical assessment of the evolution from voluntary guidelines to regulations in the Salish Sea. *Journal of International Wildlife Law and Policy*, 15(2), 125–151. https://doi.org/10.1080/13880292.2012.678792

Government of Canada, 2019. CPC-2-3-07 - Obtaining Identities in the Maritime Mobile Service. https://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/sf01032.html

Grewal, J. K., Krzywinski, M., & Altman, N. (2019). Markov models — hidden Markov models. *Nature Methods*, 16(9), 795–796. https://doi.org/10.1038/s41592-019-0532-6

Hermannsen, L., Mikkelsen, L., Tougaard, J., Beedholm, K., Johnson, M., & Madsen, P. T. (2019). Recreational vessels without Automatic Identification System (AIS) dominate anthropogenic noise contributions to a shallow water soundscape. *Scientific Reports*, 9, 15477. https://doi.org/10.1038/s41598-019-51222-9

Hossin, M., & Sulaiman, M. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1–11. https://doi.org/10.5121/ijdkp.2015.5201

International Maritime Organization (IMO). (2003). Guidelines for the Installation of a Shipborne Automatic Identification System (AIS) (Ref. T2/8.02). https://www.navcen.uscg.gov/pdf/marcomms/imo/Circulars/IMO.SN.Circ.227_AIS_Installation.pdf

Jansche, M. (2005). Maximum expected f-measure training of logistic regression models. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, October*, 692–699. https://doi.org/10.3115/1220575.1220662

Joo, R., Bertrand, S., Tam, J., & Fablet, R. (2013). Hidden Markov Models: The Best Models for Forager Movements? *PLoS ONE*, 8(8): e71246. https://doi.org/10.1371/journal.pone.0071246

Joy, R., Tollit, D., Wood, J., MacGillivray, A., Li, Z., Trounce, K., & Robinson, O. (2019). Potential benefits of vessel slowdowns on endangered southern resident killer whales. *Frontiers in Marine Science*, 6(344), 1–20. https://doi.org/10.3389/fmars.2019.00344

Kouemou, G. L. (Eds.). (2011). History and Theoretical Basics of Hidden Markov Models. *Hidden Markov Models, Theory and Applications*. https://doi.org/10.5772/15205

Lachmuth, C. L., Barrett-lennard, L. G., Steyn, D. Q., & Milsom, W. K. (2011). Estimation of southern resident killer whale exposure to exhaust emissions from whale-watching vessels and potential adverse health effects and toxicity thresholds. *Marine Pollution Bulletin*, 62(4), 792–805. https://doi.org/10.1016/j.marpolbul.2011.01.002

Last, P., Hering-Bertram, M., & Linsen, L. (2015). How automatic identification system (AIS) antenna setup affects AIS signal quality. *Ocean Engineering*, 100, 83–89. https://doi.org/10.1016/j.oceaneng.2015.03.017

Le Tixerant, M., Le Guyader, D., Gourmelon, F., & Queffelec, B. (2018). How can Automatic Identification System (AIS) data be used for maritime spatial planning*? Ocean and Coastal Management*, 166, 18–30. https://doi.org/10.1016/j.ocecoaman.2018.05.005

Lei, P. R., Tsai, T. H., & Peng, W. C. (2016). Discovering Maritime Traffic Route from AIS network. *18th Asia-Pacific Network Operations and Management Symposium*. https://doi.org/10.1109/APNOMS.2016.7737223

Li, H., Liu, J., Liu, R. W., Xiong, N., Wu, K., & Kim, T. H. (2017). A dimensionality reduction-based multi-step clustering method for robust vessel trajectory analysis. *Sensors*, 17(1792), 1-26 https://doi.org/10.3390/s17081792

Li, H., Liu, J., Wu, K., Yang, Z., Liu, R. W., & Xiong, N. (2018). Spatio-temporal vessel trajectory clustering based on data mapping and density. *IEEE Access*, 6, 58939–58954. https://doi.org/10.1109/ACCESS.2018.2866364

Li, X., Wang, J., Song, W., Ma, J., Telesca, L., & Zhang, Y. (2014). Automatic smoke detection in MODIS satellite data based on K-means clustering and fisher linear discrimination. *Photogrammetric Engineering and Remote Sensing*, 80(10), 971–982. https://doi.org/10.14358/PERS.80.10.971

Li, Y., Zhang, Y., & Zhu, F. (2017). The method of detecting AIS isolated information based on clustering and distance. *2016 2nd IEEE International Conference on Computer and*

Communications, ICCC 2016 - Proceedings, 870–873. https://doi.org/10.1109/CompComm.2016.7924827

Lusseau, D. (2006). The short-term behavioral reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Marine Mammal Science*, 22(4), 802–818. https://doi.org/10.1111/j.1748-7692.2006.00052.x

Mallard, G. (2019). Regulating whale watching: A common agency analysis. *Annals of Tourism Research*, 76, 191–199. https://doi.org/10.1016/j.annals.2019.04.011

Markov, A. A. (2006). Classical text in translation: An example of statistical investigation of the text *Eugene Onegin* concerning the connection of samples in chains. *Science in Context*, 19(4), 591–600. https://doi.org/10.1017/S0269889706001074

McWhinnie, L. H., O'Hara, P. D., Hilliard, C., Le Baron, N., Smallshaw, L., Pelot, R., & Canessa, R. (2021). Assessing vessel traffic in the Salish Sea using satellite AIS: An important contribution for planning, management and conservation in southern resident killer whale critical habitat. *Ocean and Coastal Management*, 200, 1–17. https://doi.org/10.1016/j.ocecoaman.2020.105479

Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. In *The MIT Press* (Issue 4).

Muthu, S. S. (2015). *Visualization, Statistical Analysis, and Mining of Historical Vessel Data* (Issue 296). http://www2.unb.ca/gge/Pubs/TR296.pdf

Pallotta, G., Vespe, M., & Bryan, K. (2013). Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6), 2218–2245. https://doi.org/10.3390/e15062218

Parrott, L., Chion, C., Martins, C. C. A., Lamontagne, P., Turgeon, S., Landry, J. A., Zhens, B., Marceau, D. J., Michaud, R., Cantin, G., Ménard, N., & Dionne, S. (2011). A decision support system to assist the sustainable management of navigation activities in the St. Lawrence River Estuary, Canada. *Environmental Modelling and Software*, 26, 1403–1418. https://doi.org/10.1016/j.envsoft.2011.08.009

Peel, D., & Good, N. M. (2011). A hidden markov model approach for determining vessel activity from vessel monitoring system data. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(7), 1252–1264. https://doi.org/10.1139/f2011-055

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2), 257–286. https://doi.org/10.1109/5.18626

Ratner, B. (2011) Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition, *CRC Press LLC,. ProQuest Ebook Central*, https://ebookcentral-proquest-com.ezproxy.library.uvic.ca/lib/uvic/detail.action?docID=840391

Robards, M. D., Silber, G. K., Adams, J. D., Arroyo, J., Lorenzini, D., Schwehr, K., & Amos, J. (2016). Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)-A review. *Bulletin of Marine Science*, 92(1), 75–103. https://doi.org/10.5343/bms.2015.1034

Schein, A. I., & Ungar, L. H. (2007). Active learning for logistic regression: An evaluation. In *Machine Learning*, 68(3). https://doi.org/10.1007/s10994-007-5019-5

Schuler, A. R., Piwetz, S., Di Clemente, J., Steckler, D., Mueter, F., & Pearson, H. C. (2019). Humpback Whale Movements and Behavior in Response to Whale-Watching Vessels in Juneau, AK. *Frontiers in Marine Science*, 6:710, 1–13. https://doi.org/10.3389/fmars.2019.00710

Schwehr, K. D., & McGillivary, P. A. (2007). Marine ship automatic identification system (AIS) for enhanced coastal security capabilities: An oil spill tracking application. *Oceans Conference Record (IEEE)*. https://doi.org/10.1109/OCEANS.2007.4449285

Seely, E., Osborne, R. W., Koski, K., & Larson, S. (2017). Soundwatch: Eighteen years of monitoring whale watch vessel activities in the Salish Sea. *PLoS ONE*, 12(12):e0189764, 1–18. https://doi.org/10.1371/journal.pone.0189764

Sheng, K., Liu, Z., Zhou, D., He, A., & Feng, C. (2018). Research on Ship Classification Based on Trajectory Features. *Journal of Navigation*, 71(1), 100–116. https://doi.org/10.1017/S0373463317000546

Sidibé, A., & Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. *Journal of Navigation*, 70(4), 847–858. https://doi.org/10.1017/S0373463317000066

Silber, G. K., Adams, J. D., & Fonnesbeck, C. J. (2014). Compliance with vessel speed restrictions to protect North Atlantic right whales. *PeerJ*, 2014(1). https://doi.org/10.7717/peerj.399

Sobocinski, K.L. (2021). State of the Salish Sea. G. Broadhurst and N.J.K. Baloy (Contributing Eds.). *Salish Sea Institute, Western Washington University*. https://doi.org/10.25710/vfhb-3a6

Soe, O. K., Chaojian, S., Qinyou, H., & Weintrit, A. (2010). Clustering Analysis and Identification of Marine Traffic Congested Zones. *Zeszyty Naukowe Akademii Morskiej W Gdyni*, 67, 101–113.

Toloue, K. F., & Jahan, M. V. (2018). Anomalous behavior detection of marine vessels based on Hidden Markov Model. *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems*, 10–12. https://doi.org/10.1109/CFIS.2018.8336611

Vermard, Y., Rivot, E., Mahévas, S., Marchal, P., & Gascuel, D. (2010). Identifying fishing trip behaviour and estimating fishing effort from VMS data using Bayesian Hidden Markov Models. *Ecological Modelling*, 221(15), 1757–1769. https://doi.org/10.1016/j.ecolmodel.2010.04.005

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 37(5), 360–363. http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf

Walker, E., & Bez, N. (2010). A pioneer validation of a state-space model of vessel trajectories (VMS) with observers' data. *Ecological Modelling*, 221(17), 2008–2017. https://doi.org/10.1016/j.ecolmodel.2010.05.007

Washington State Legislature, 2021. Protection of southern resident orca whales – Unlawful activities – Penalty. https://apps.leg.wa.gov/rcw/default.aspx?cite=77.15.740

Waterbolk, M., Tump, J., Klaver, R., van der Woude, R., Velleman, D., Zuidema, J., Koch, T., & Dugundji, E. (2019). Detection of Ships at Mooring Dolphins with Hidden Markov Models. *Transportation Research Record*, 2673(4), 439–447. https://doi.org/10.1177/0361198119837495

Wiley, D. N., Thompson, M., Pace, R. M., & Levenson, J. (2011). Modeling speed restrictions to mitigate lethal collisions between ships and whales in the Stellwagen Bank National Marine Sanctuary, USA. *Biological Conservation*, 144, 2377–2381. https://doi.org/10.1016/j.biocon.2011.05.007

Wright, S. J. (2015). Coordinate descent algorithms. *Mathematical Programming*, 151(1), 3–34. https://doi.org/10.1007/s10107-015-0892-3

Zucchini, W., Macdonald, I. L., & Langrock, R. (Eds.) (2017). Hidden Markov models for time series: An introduction using R (2nd ed.). In *Monographs on Statistics and Applied Probability* 150. https://doi.org/10.1201/b20790

Chapter 3 Evaluating Hidden Markov Model Approaches for Classifying Whale-watching Vessel Movement

3.1 Introduction

Starting in the mid-1980s, an international moratorium on commercial whaling has helped the recovery of many cetacean populations around the world, leading to a global increase in recreational whale-watching activity (IWC, 1983; Schuler et al., 2019). Whale-watching is often considered a sustainable, non-consumptive activity for exploiting whales and other wildlife (Mallard, 2019; Schuler et al., 2019). However, these human-wildlife interactions have the potential to impose harm on cetacean species and coastal ecosystems (Robards et al., 2016). Evidence that whale-watching has had short-term impacts on some cetacean populations (Dahlheim & Castellote, 2016; Krasnova et al., 2020; Lundquist et al., 2013; Meissner et al., 2015) and longer-term consequences for the health of marine environments (Lachmuth et al., 2011; Mallard, 2019) continues to emerge. Short-term impacts on cetaceans due to marine vessel disturbances include behavioural changes, such as altering resting, diving, and foraging behavioural patterns (Lusseau, 2006; Schuler et al., 2019; Williams et al., 2002), prohibiting acoustic communication between cetaceans due to the noise created by vessel propulsion (Erbe, 2002; Frankel & Gabriele, 2017; Hermannsen et al., 2019), and vessel collisions with cetaceans (Conn & Silber, 2013; Vanderlaan & Taggart, 2007). Long-term exposure to vessels (although even more challenging to quantify) will also likely have impacts on cetacean populations, such as alterations to or abandonment of habitats, and negative effects on biological fitness (Bejder et al., 2006; Findley & Vidal, 2002; Mallard, 2019). Therefore, understanding the behaviour of vessels while viewing wildlife is a crucial first step for informing the development of mitigation and management strategies that aim to reduce or minimise the potential impacts vessels can have on marine wildlife and in particular cetaceans.

Understanding how vessels and cetaceans interact is imperative for mitigating potential negative impacts particularly for sensitive or threatened species, and requires information on wildlife abundance and distribution to determine if conservation management strategies are effective (Allen & Singh, 2016; Sveegaard et al., 2015; Vinding et al., 2015). Tracking individual animals as they transit an area can be used to model population movement, which can inform studies on prey or foraging ecology (Hazen et al., 2009; Shields et al., 2018), migratory patterns and locations of critical habitats (Calambokidis et al., 2014; Norris et al., 1999; Pirotta et al., 2011), or responses to climate change (Moore & Huntington, 2008; Trites, 2014). However, tracking cetaceans is typically costly and time consuming because they often occur through dedicated observation time on vessels or from land (Olson et al., 2018; Williams et al., 2002), which has led to supplementary monitoring methods such as aerial surveys (Laake et al., 1997), acoustic detection (Gillespie et al., 2020; Hazen et al., 2009; Norris et al., 1999; Pirotta et al., 2011; Van Parijs et al., 2009), or satellite detection (Mate et al., 1997; Sveegaard et al., 2015; Wilson et al., 2007). In most instances, the species or population being tracked is highly mobile with a large geographic range over a prolonged time (Calambokidis et al., 2014; Dahlheim & White, 2010), which often requires the amalgamation of different cetacean detection sources to gather enough data for these models (Pirotta et al., 2011; Sveegaard et al., 2015). There exists the potential to supplement established cetacean tracking methods using vessel monitoring systems aboard vessels associated with marine mammals, such as commercial ecotourism vessels, which requires further exploration.

One such monitoring system is AIS – Automatic Identification System – which is an autonomous vessel reporting system that can provide location and time-specific data on vessels. An AIS transponder aboard a vessel automatically broadcasts vessel information regarding vessel movement (speed over ground, course over ground, heading, etc.), identifiers, and position at regular time intervals (i.e. typically between 30 seconds and 3 minutes depending on vessel speed) to terrestrial and satellite AIS receivers and other vessels carrying AIS. Originally created to enhance safety of vessels while at sea (IMO, 2003), AIS data have been increasingly utilized in studies on the acoustic disturbances that vessels have on marine mammals (Fouda, 2012; Frankel & Gabriele, 2017; Joy et al., 2019), the environmental impacts vessels have on coastal ecosystems (Meijles et al., 2021), and managing marine traffic to mitigate harmful impacts on wildlife (McWhinnie et al., 2021; Parrott et al., 2011).

Recently, in 2019 the Canadian Government made AIS transponders mandatory on all passenger vessels carrying more than 12 passengers or that are 8 m or more in length, which applies to many commercial wildlife-viewing vessels (Canadian Shipping Act, 2019). This has led to the availability of AIS data that can be utilized in studies of commercial ecotourism vessels. Although extensive AIS data for commercial ecotourism vessels exists, these data may not be inclusive as these regulations do not require vessel shorter than 8 m or that carry less than 12 passengers to be AIS enabled. Typically, AIS data are used in the classification of vessel behaviour in the fishing and shipping industries (De Souza et al., 2016; Le Tixerant et al., 2018; Sheng et al., 2018), but there are relatively few studies that focus on using AIS to classify wildlife-viewing behaviour by commercial boats. One such study (Almunia et al., 2020) in Spain used AIS data to classify wildlife-viewing events through applying thresholds on speed over ground (SOG) and auxiliary bathymetry data. However, one possible limitation with using SOG and bathymetry thresholds along the bathymetric complexity of the British Columbia (B.C.) and Washington state coastline is that they may lead to misleading classification results because they do not capture the temporal aspect of wildlife-viewing trips, or the navigational changes in the vessel's speed during maneuvers when the boat is not engaged in wildlife-viewing, such as when travelling through high traffic areas, or deep narrow channels. Therefore, research is needed to determine if AIS data can be employed in an automated approach for the classification of wildlife-viewing behaviour to reduce reliance on auxiliary data.

AIS data have been utilized in the automated classification of marine vessel behaviour in applications, such as inferring fishing vessel activity (De Souza et al., 2016), forecasting marine traffic flow (Mone et al., 2010), predicting vessel trajectories (Pallotta et al., 2013), or to classifying abnormal or illegal behaviour (Schwehr & McGillivary, 2007; Sidibé & Shu, 2017). These applications employ a range of statistical data mining or machine learning approaches (Sidibé & Shu, 2017). One such approach is the hidden Markov model (HMM), which is a probabilistic state-based model traditionally used in ecological movement studies and pattern detection (Joo et al., 2013). Previous studies paired the HMM with AIS data to estimate fishing effort (Charles et al., 2014; Vermard et al., 2010), classify fishing vessel behaviour (De Souza et al., 2016; Peel & Good, 2011), and detect vessel moorage (Waterbolk et al., 2019). These studies were successful in classifying the behaviour of vessels using observed AIS data, suggesting that they may also provide a suitable method for classifying boats demonstrating movement patterns that would indicate that they are likely viewing wildlife (wildlife-viewing behaviour). However, research is needed to determine how to appropriately apply the HMM in a way that provides a suitable classification of wildlife-viewing behaviour for a geographically extensive commercial vessel fleet.

In the context of this study, the HMM is a stochastic model representing the probability distribution of observed data from a sequence of AIS points (Rabiner, 1989). In general, the probabilities that make up the HMM determine whether an AIS point is viewing wildlife based only on information observed within that point and the behaviour of the previous AIS point. The HMM takes advantage of the SOG sequence observed in the AIS data to probabilistically infer vessel behaviour classes (i.e. hidden states) (Kouemou, 2011) with the assumption that the hidden behaviour states form a Markov chain (Markov, 2006) representing the transitions between states. This model is considered a machine learning algorithm that can be applied in a supervised or unsupervised manor (Boussemart et al., 2011; Joo et al., 2013; Zhang et al., 2005). Supervised learning leverages pre-defined observed behaviour classes to determine the most likely set of HMM parameters, whereas the unsupervised HMM learns the most likely set of model parameters based solely on the information contained from the observed SOG sequence (Murphy, 2012; Zhang et al., 2005). The definition of unsupervised learning used here aligns with the Khodabandelou et al. (2014) definition, which supports prior knowledge of the number of behaviour states. With sufficiently large and reliable observed data on behavioural classes, supervised learning typically provides the best HMM performance, and is favoured in the machine learning community because the observed class data increase how much information the HMM captures (Boussemart et al., 2011). When observed class data are unavailable, unsupervised learning is a sufficient alternative that does not require region specific domain knowledge, but may be less computationally efficient than supervised learning and, due to the automatic inference of HMM parameters, is sensitive to the initialization of parameters, which can lead to uncertainty (Boussemart et al., 2011; Khodabandelou et al., 2014).

A study by Zhang et al. (2005) of unusual event classification using the HMM noted that a supervised approach outperformed an unsupervised approach if a sufficient amount of observed class data are available, and that with an insufficient amount of observed class data supervised learning performs worse. Obtaining observed class data for a sufficient amount of the commercial ecotourism fleet is infeasible, so the study conducted here takes advantage of a small sample of observed class data collected aboard a single commercial wildlife-viewing vessel for model validation. Therefore, due to limited observed class data, research is needed to determine how comparable the performance of the unsupervised HMM is to the supervised HMM.

The objective of this study is to determine if an unsupervised HMM is comparable to the supervised HMM in the classification of wildlife-viewing vessel behaviour, and whether outputs from these models can be used to provide vessel movement information that may indicate which target species is being observed. This is performed by assessing supervised and unsupervised HMM performance for a single wildlife-viewing vessel's observed behaviour. Additional information may be derived from the resulting classification dataset, such as species-specific spatial-temporal vessel movement trends. By establishing the utility of the unsupervised approach, the HMM can be employed at a large geographic scale and across an array of different ecotourism vessels to inform the management and study of wildlife-viewing vessel activities. Furthermore, the ability to determine which species is being targeted by commercial ecotourism vessels is useful for studies of cetacean distribution and conservation. The results from this study have the potential to aid the development of near real-time methods and tools that could eventually assist regulators and marine managers in performing systematic evaluations of the behaviours of small vessel observing wildlife, and possibly contribute to existing knowledge and data related to cetacean population distributions and abundance.

3.2 Methods

3.2.1 Study Area and Data

This study analyzed 62 commercial ecotourism vessels with AIS in the waters surrounding Vancouver Island in the province of B.C., Canada between April 1, 2019 and September 30, 2019, equating to 183 days of data collection (Fig. 3.1 and 3.2). These waters include the Salish Sea (Southern Strait of Georgia, Juan de Fuca Strait, Haro Strait, and Puget Sound) containing the Gulf Islands and San Juan Islands, the westerly Barkley Sound, Clayoquot Sound, and Quatsino Sound, and the northerly Queen Charlotte Strait (Fig. 3.1). These nutrient rich coastal waterways are known worldwide for their abundance of marine wildlife where marine mammals are commonly observed, such as killer whales (*Orcinus orca*), humpback whales (*Megaptera novaeangliae*), grey whales (*Eschrichtius robustus*), Harbour seals (*Phoca vitulina*), Steller sea lions (*Eumetopias jubatus*) and California sea lions (*Zalophus californianus*) (Fisheries and Oceans Canada, 2020). Alongside marine mammals, observations of coastal terrestrial and avian species (e.g. wolves, bears, and bald eagles) are also frequently viewed by boats engaged in wildlife tourism, with the particular target species changing seasonally and geographically.

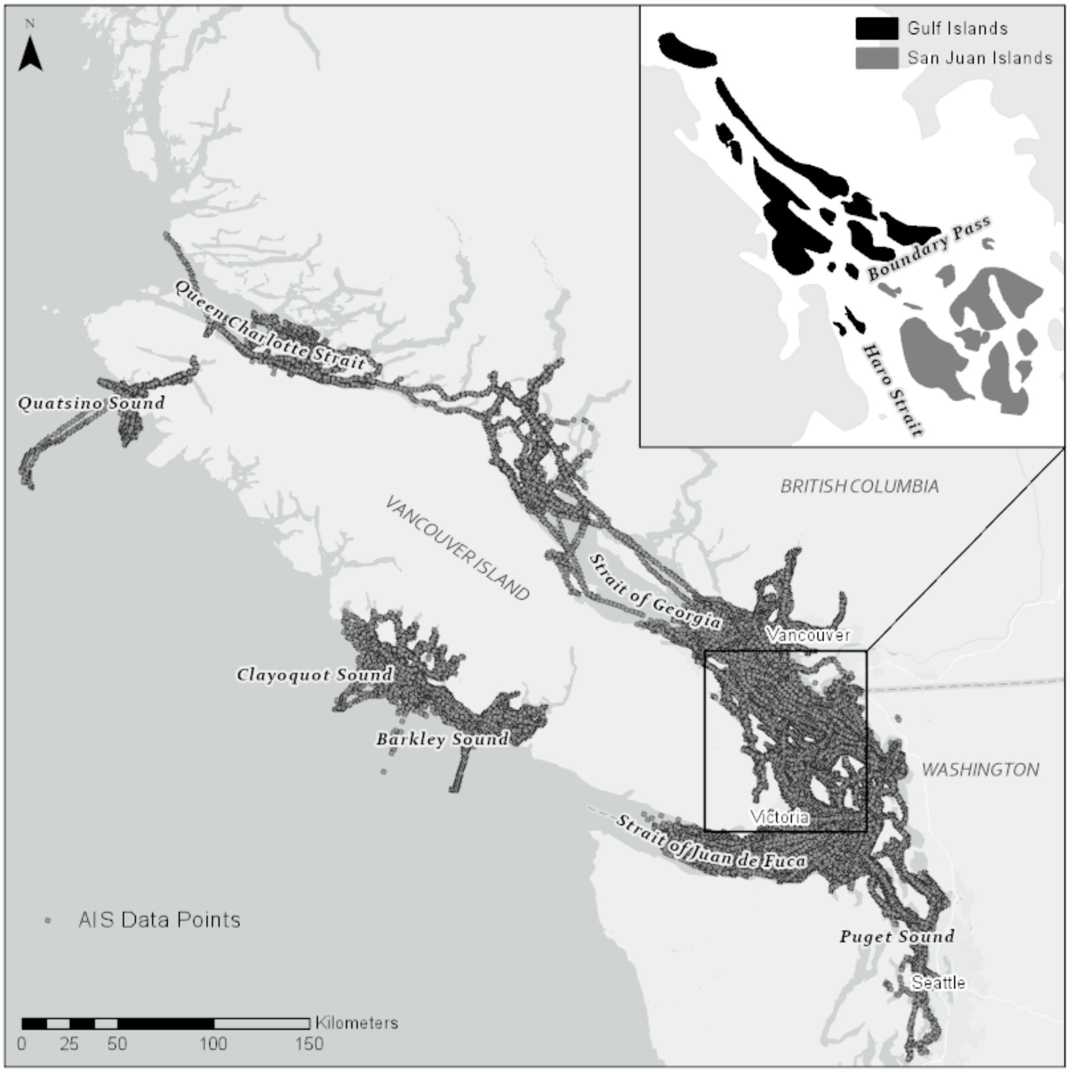

Figure 3.1: Study area of the waters surrounding Vancouver Island, and AIS transmitted commercial ecotourism vessel point locations.

The AIS data for this study were collected by the Canadian Coast Guard and provided by Ocean Networks Canada (ONC). These data include received locations for all vessels surrounding Vancouver Island during the time period of the study and consist of vessel identifiers, position, kinematics (speed, course, etc.), and a timestamp. A total of 62 known wildlife-viewing vessels were extracted from all AIS vessel data based on their Maritime Mobile Service Identity (MSSI), which is a unique identifier assigned to vessels at sea for automated safety communication as part of the global maritime distress and safety system (Government of Canada, 2019). The AIS data were cleaned by removing individual data points if they were a duplicate AIS message (i.e. received within a 10 second timespan), a stationary docked vessel, or an invalid vessel report (i.e. invalid position, a sudden change of direction, or speed over ground greater than 70 knots).

For this study, the duration of a wildlife-viewing trip begins when the vessel leaves a known commercial ecotourism harbour and ends when it returns to a known commercial ecotourism harbour. A harbour dataset was derived by manually delineating harbour boundaries based on provided homeport information for the commercial ecotourism companies in this study, and the accumulation of moored vessels AIS points. This study assumes that observing cetaceans within these harbours is rare. This resulted in 62 wildlife-viewing vessels, a total of 790,679 AIS messages or points, and 11,154 wildlife-viewing trips. All trips with less than two AIS points were removed from the dataset as these do not provide the necessary amount of data required for the modelling process, which totals 10,825 wildlife-viewing trips. One minor limitation with the data is that there is a decrease in AIS data received in the second half of July due to a data storage error. While the impact on the model is assumed to be minimal, this limitation is considered when analyzing the model results.

The first step of this study was to classify the clean AIS data into wildlife-viewing events (class 1) and non-viewing events (class 0). Cetacean sightings data, collected on a single commercial wildlife-viewing vessel, were assigned to each AIS point representing their observed class, which were then utilized in the development of the supervised HMM and the validation of the supervised and unsupervised HMM. These data were acquired during previous research (Fraser et al., 2020) into the compliance of whale-watching vessels to minimum distance restrictions in the Salish Sea. The Fraser et al. (2020) cetacean sightings data provided the date of each wildlife-viewing trip for this vessel between June and September, 2019, the tour departure time, and the time interval for each sighting, which was applied to the vessel's AIS points. This resulted in 27 trips used in the development and statistical validation of the HMMs.

The BC Cetaceans Sightings Network (BCCSN) provided cetaceans sightings data, which were used to assign a probable cetacean species to each HMM classified wildlife-viewing event. These data included a point location, date, time, and probable species that were collected through a crowdsourcing initiative (BCCSN, 2021). Dolphin and porpoise sightings were removed from consideration, as commercial wildlife-viewing vessels do not consistently stop to view these animals unless there is an anomalous event occurring (e.g. large number of dolphin or porpoise at a single location). This process resulted in a set of possible species observed that include fin whale, grey whale, humpback whale, killer whale, minke whale, sperm whale, unidentified baleen whale, and unidentified cetacean. The latter category can include any cetacean species, including dolphin and porpoise; despite that, it was included in this analysis due to this uncertainty.

3.2.2 Hidden Markov Model

The HMM is made up of hidden variables that are the behaviour states, and observed variables that are values directly observed within the AIS data, and, potentially, the associated observed whale encounter data. There are two HMM behaviour states for this study that represent the wildlife-viewing class (i.e. those AIS points that are involved in viewing wildlife at that particular moment), and the non-wildlife viewing class. These two states are considered hidden because they are not directly observable from the AIS data. Instead, they are inferred through Markovian transitions of the observed variable, speed over ground (SOG), for a series of AIS points. The progression of these states form the hidden Markov chain (Vermard et al., 2010). Given a series of known observations (O) over time (T), there exists a set of hidden behaviour states (S). The HMM assumes that the observation at time t (O_t) is generated by a hidden state (S_t), and that, given the previous state (S_{t-1}), S_t is independent of states prior to S_{t-1} (De Souza et al., 2016; Toloue & Jahan, 2018).

The initial probabilities (π), transition probabilities (a), and emission probabilities (b) are the building blocks of the HMM (λ), such that λ = {π, a, b}. The parameter π is the probability that the first point in a sequence belongs to each behaviour state, a is the probability of moving between states, and b is the probability density functions for each state's observed variables (O). The λ is characterized by three fundamental problems (Rabiner, 1989):

1. Evaluation Problem: Given λ and O, what is the model likelihood $P(O \mid \lambda)$?
2. Decoding Problem: Given O and λ, what is the best hidden state sequence S?
3. Learning Problem: Given O and S in λ, how can the HMM's parameters for π, a, and b be adjusted to maximize $P(O \mid \lambda)$?

$P(O \mid \lambda)$ is derived through solving:

$$P(O \mid \lambda) = \sum_{S} \pi_{S_1} \prod_{t=2}^{T} a_{S_{t-1}} S_t \prod_{t=1}^{T} b_{S_t}(O_t) \tag{1}$$

where, S_1 is the initial state and T is the last SOG observation read from the data (De Souza et al., 2016; Waterbolk et al., 2019). This cannot be solved through brute force probabilistic methods that enumerate all possible state sequences because they are computationally complex, requiring N^T number of operations (where N is the number of hidden states). Therefore, the forward-backward algorithm (Rabiner, 1989) is employed to solve for $P(O \mid \lambda)$, which uses induction to derived the forward probabilities, and backward probabilities in two parts. The forward probabilities are the joint probability of the partially observed sequence O_t and S_n up to time t, given λ, and the backward probabilities are the conditional probability of the partially observed sequence from O_{t+1} to O_T, given state S_n at time t (Rabiner, 1989). The forward-backward algorithm is not computationally taxing as it uses recursion, and only the forward part of the algorithm is required to solve the evaluation problem (Rabiner, 1989; Waterbolk et al., 2019). The Viterbi algorithm solves the decoding problem by using the maximum log-likelihood score to decode the best hidden state sequence from O (Grewal et al., 2019; Rabiner, 1989). The learning problem is solved through expectation-maximization using the Baum-Welch algorithm (Baum & Petrie, 1966), which is an iterative procedure involving an E step and M step after the initialization of π, a, and b. The E step takes advantage of the forward-backward algorithm to estimate the conditional expectations of π, a, and b given O (Zucchini et al., 2016). The M step uses re-estimation formulas to update the parameters defining π, a, and b via maximum likelihood (Rabiner, 1989). This algorithm repeatedly refines these parameters until convergence, which is when the change in model likelihood between re-estimations is less than the predefined tolerance.

The distinction between the supervised and unsupervised approaches to the HMM is that the former derives π and a directly from the AIS data for a single wildlife-viewing vessel with observed classes, and b through direct maximum likelihood fit to the SOG data. Therefore, it only needs to solve the decoding problem as the parameters for π, a, and b do not need to be re-estimated. The HMM naturally transitions from supervised to unsupervised with the integration of the expectation-maximization algorithm (Joo et al., 2013; Khodabandelou et al., 2014) for the learning problem to perform the classification of wildlife-viewing vessel behaviour.

3.2.2.1 Supervised Hidden Markov Model

Based on cetacean encounter data collected in previous work by Fraser et al. (2020), observed behaviour classes were assigned to the AIS data, which were used to directly derive the supervised HMM initial, transition, and emission probabilities from the SOG observations for 19 trips aboard one boat known to be undertaking wildlife-viewing. This subset of 19 trips from the 27 total trips for this vessel between June and September 2019 was chosen for model development (i.e. training) to avoid overfitting the model to the sample dataset, with the remaining 8 trips used for model testing. Overfitting the HMM to the sample dataset may bias the model toward this vessel's movement and may not perform as well with the 61 other vessels.

The training subset was used to derive the initial, transition, and emission probabilities from the sample's SOG, timestamp, and observed classes. The first AIS point's observed class for each of these trips were used to define the initial probabilities. The transition probabilities were defined through conditional probability of the sequence of observed classes (Kouemou, 2011; Murphy, 2012). The emission probabilities were defined through maximum likelihood fit of established continuous probability distribution functions to each observed class's SOG profiles through direct numerical minimization of the negative log-likelihood (Zucchini et al., 2016). These probabilities were used to create the HMM and then classify each AIS point belonging to 10,825 wildlife-viewing trips for all 62 vessels surrounding Vancouver Island, which was performed by decoding the most likely hidden state sequence with the Viterbi algorithm (Grewal et al., 2019; Rabiner, 1989).

3.2.2.2 Unsupervised Hidden Markov Model

The unsupervised HMM automatically derived the initial, transition, and emission probabilities from the observed SOG for all AIS points representing the 10,825 wildlife-viewing trips. The unsupervised aspect of this HMM required solving the learning problem prior to decoding the binary classes for each point. Gaussian probability distribution functions were used for the emission probabilities based on the SOG distribution for all 62 wildlife-viewing vessels. The Baum-Welch algorithm re-estimated the means and diagonal covariance matrix for both classes' Gaussian curves, alongside the re-estimation of the initial and transition probabilities, until convergence.

The Baum-Welch algorithm required the definition of parameter values for a convergence threshold, or tolerance for stopping, and the maximum number of iterations to perform if convergence is not reached. In order to best estimate these parameters, the unsupervised HMM was tested through maximizing model likelihood using plausible values for convergence thresholds and maximum iterations (Boussemart et al., 2011). Peel and Good (2011) suggest defining these parameters using similar vessel data if historic data are unavailable. Fortunately, the observed class data for the single vessel used to train the supervised HMM affords a unique opportunity for validation and parameter tuning of the unsupervised HMM, which were used to confirm (not train) these estimated parameters through maximization of the F-Score (Table 3.1). The maximization of F-Score for parameter re-estimation was found to be a suitable method for parameterization of binary classification problems (Hossin & Sulaiman, 2015). Once the convergence parameters were defined, the AIS data for all 62 vessels were fit to the unsupervised HMM before classifying each AIS point.

3.2.2.3 Validation

Statistical validation of unsupervised classification models are typically difficult due to the lack of observed classes (Cazzanti & Pallotta, 2015). Fortunately, the observed classes for a single wildlife-viewing vessel from previous research (Fraser et al., 2020) can be used to apply standard supervised learning performance metrics to validate both the supervised and unsupervised HMM. These performance metrics include overall accuracy, precision, recall, receiver operating characteristic area under the curve (ROC AUC), F-Score (Hossin & Sulaiman, 2015), Kappa statistic (Cohen, 1960), and root mean square error (RMSE) (Çiğşar & Ünal, 2019) referred to in

Table 3.1. Precision is interpreted as the fraction of confirmed wildlife-viewing classifications retrieved by the HMM over all observed wildlife-viewing classifications, and recall is the fraction of confirmed wildlife-viewing classifications retrieved by the HMM over all possible classifications. F-Score is the harmonic mean between precision and recall, which is argued to be a better discriminator for binary classification algorithms compared to overall accuracy (Hossin & Sulaiman, 2015). The same argument has been made for the ROC AUC score, which is an essential evaluation criteria for learning algorithms and is interpreted as the HMM's capability of distinguishing between the wildlife-viewing and non-viewing behaviour states (Çiğşar & Ünal, 2019).

Table 3.1: Classifier validation performance metrics, adapted from Hossin and Sulaiman (2015).

Confusion Matrix		Observed Classes	
		Wildlife-viewing	Non-viewing
Modelled Classes	Wildlife-viewing	True positive (tp)	False positive (fp)
	Non-viewing	False negative (fn)	True negative (tn)

Metric	**Formula**	**Focus**
Overall Accuracy	$\frac{tp + tn}{tp + tn + fn + fp}$	Ratio of correct classifications over the total number of observations.
Precision	$\frac{tp}{tp + fp}$	Fraction of correct positive classifications over the total positive observations.
Recall	$\frac{tp}{tp + fn}$	Fraction of correct positive classification over the total observations in the class.
F-Score	$\frac{2 * (precision * recall)}{precision + recall}$	Harmonic mean between precision and recall.
RMSE	$\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}}$	Standard deviation of the difference between observed and modelled classes where, *n* is the total number of AIS points in the samples, y the observed class for point *i* in *n*, and $\hat{y}_i$ is the modelled class.

Kappa Statistic	$K = \frac{p_o - p_e}{1 - p_e}$	Tests the null hypothesis that the modelled classes were randomly chosen where, p_o is the observed probability of agreement, and p_e is the expected probability given by chance or at random (Cohen, 1960).

As there are a small number of wildlife-viewing trips associated with the observed class data and these are limited to a single vessel, a test of each model's log-likelihood score was conducted, which does not require the observed class data. This test determines which HMM has a higher chance of generating the SOG observed in the AIS data (Khodabandelou et al., 2014; Zucchini et al., 2016).

3.2.3 Species Level Trends

The BCCSN cetacean sightings data were used to assess whether there are spatial, temporal, or kinematic vessel movement characteristics that can distinguish between the probable target species being observed. This was done by evaluating spatial, temporal, or kinematic (i.e. SOG) distribution signatures found in the AIS data associated with target species. The classified wildlife-viewing events were each assigned a probable target species based on spatial and temporal proximity to the BCCSN sightings data. As the BCCSN sightings data are crowd-sourced points, they may not align well with the classified wildlife-viewing events. Therefore, the spatial and temporal proximity these data have to the AIS points were used to evaluate vessel movement characteristics for the classified wildlife-viewing events. Using the observed class AIS data, maximum distance and temporal thresholds between the BCCSN cetacean sightings and the observed wildlife-viewing events were defined. This was done by examining the distribution of the Euclidean distance between the observed wildlife-viewing class AIS points and the BCCSN cetacean sightings locations, and the distribution for the time difference between these points.

Once these thresholds were defined, the BCCSN cetacean sightings data that were closest spatially and temporally, and within the defined thresholds, were assigned to each AIS point that the HMMs classified to be wildlife-viewing for all 62 vessels. Only the classified wildlife-viewing events were included in this evaluation. Due to the dynamic nature of cetaceans and the possible misalignment between when the BCCSN reported sighting occurred and when the commercial ecotourism vessel began wildlife-viewing, non-viewing vessel behaviour could fall within the defined distance and time thresholds, which could create noise in the datasets and skew the distributions. For example, while a commercial ecotourism vessel is transiting to a wildlife-viewing event they can travel at speeds up to 40 knots, which does not constitute wildlife-viewing behaviour.

If there was at least one AIS point with an assigned species, then that cetacean, their location, and time were assigned to that wildlife-viewing event. Consequently, when two or more species were assigned to the same wildlife-viewing event the species that occurred most frequently was applied to the entire wildlife-viewing event. If no BCCSN cetacean sighting fell within the defined thresholds, the wildlife-viewing event was assigned to the ‘no recorded species’ category. The classified wildlife-viewing class for the supervised and unsupervised HMMs were then evaluated based on these assignments and used to inform species level trends within the AIS data.

An evaluation of species level temporal, kinematic, and spatial distributions determined if there existed any distinctive vessel movement characteristics between the probable species observed. This was done using non-parametric Kruskal-Wallis tests to determine how encounter time and SOG varied between species and HMM approaches. The spatial distributions for each species were assessed through kernel density maps of expected wildlife-viewing AIS point counts at a 1 km spatial resolution to estimate vessel frequency and wildlife-viewing effort to a region.

3.3 Results

Of the 11,154 total wildlife-viewing trips, 10,825 trips met the requirement of having a minimum of two AIS points. Of those 10,825 trips, the maximum number of AIS points was 352 (which occurred for one trip), and the mean number of points per trip was 72.89. The AIS data had a mean temporal resolution of 2 minutes and 45 seconds between AIS messages, which is considered an adequate resolution for capturing detailed vessel movement in the context of this study.

3.3.1 Classification Results

3.3.1.1 Supervised Hidden Markov Model

The observed class training data consisting of 19 trips were used to derive the initial, transition, and emission probabilities that create the supervised HMM. The AIS data were sorted temporally to derive the correct SOG sequence, and the initial probabilities were defined from the first point in each trip. The initial behaviour state (class) for all 19 trips in the training dataset were non-viewing points resulting in a probability of 1.0 for the non-viewing state, and 0.0 for the wildlife-viewing state. Conditional probability of the observed class sequence was used to derive the transition probability matrix (Table 3.2), which is interpreted as a 0.95 probability that the current AIS point in the sequence is non-viewing given that the previous point was also non-viewing, and a 0.91 probability that the currently point is wildlife-viewing given the previous point was also wildlife-viewing. Conversely, there is a 0.05 probability that a non-viewing point would transition to a wildlife-viewing state, and a 0.09 probability that a wildlife-viewing point would transition to a non-viewing state.

Table 3.2: Supervised hidden Markov model transition probability matrix.

	Non-viewing	Wildlife-viewing
Non-viewing	0.95	0.05
Wildlife-viewing	0.09	0.91

The emission probabilities were derived from the SOG distributions for each class, which were distinctive (Fig. 3.2). The resulting emission probability density function for the non-viewing class is represented by a double gamma distribution function, and the wildlife-viewing class is represented by the generalized inverse Gaussian distribution function. These probability density functions were chosen based on numerical minimization of the log-likelihood score and visual inspection of how these functions fit the SOG data. Model generation and classification for all wildlife-viewing trips in the dataset took < 8 seconds to complete.

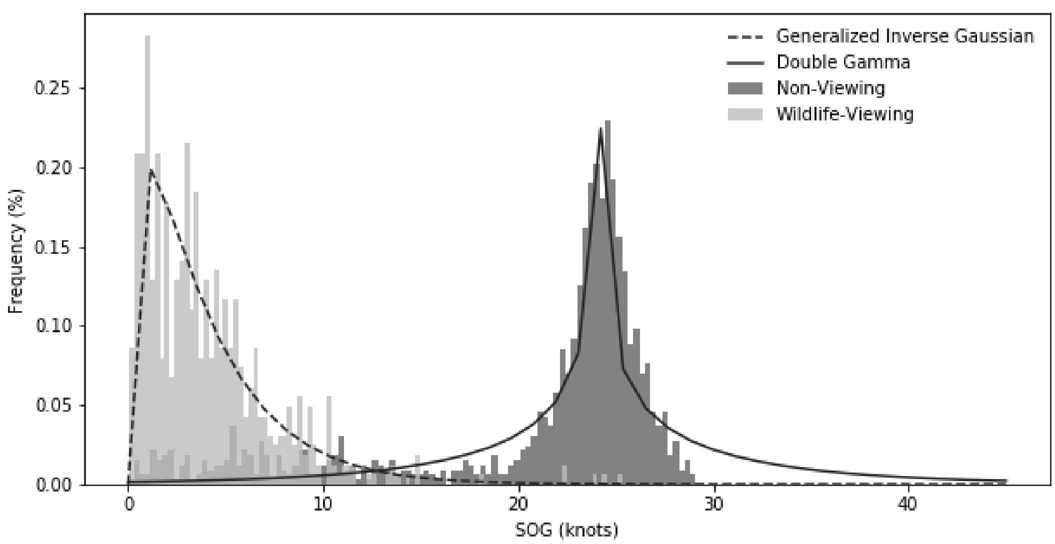

Figure 3.2: Speed profiles for the supervised hidden Markov model. Histogram of speed over ground (SOG) observations with the best fitting probability density functions for each hidden state.

3.3.1.2 Unsupervised Hidden Markov Model

Based on the SOG distributions of the AIS data, Gaussian emission probabilities were employed for the unsupervised HMM. The unsupervised HMM determined the Gaussian probability density functions that best fit the SOG observations by re-estimating their means and diagonal covariance matrix for the two behaviour states (classes) (Fig. 3.3). The initial and transition probabilities were also re-estimated during model fitting (Table 3.3). This resulted in different initial probabilities than the supervised HMM with a 0.35 probability that the first AIS point in a trip is wildlife-viewing and 0.65 probability that the first point is non-viewing. The values in the transition probability matrix can interpreted as a 0.92 probability that the current AIS point is a non-viewing point given the previous state was non-viewing, and a 0.08 probability that the current point is wildlife-viewing given the previous state a was non-viewing. Similarly interpreted, the current point has a 0.91 probability of being wildlife-viewing given the previous state was wildlife-viewing, and a 0.09 probability that the current point is non-viewing given the previous point was wildlife-viewing.

Table 3.3: Unsupervised hidden Markov model transition probability matrix, and Gaussian emission probability density function means and diagonal covariance.

Transition Matrix

	Non-viewing	Wildlife-viewing
Non-viewing	0.92	0.08
Wildlife-viewing	0.09	0.91

Gaussian Probability Density Function Parameters

Diagonal Covariance Matrix

	Non-viewing	Wildlife-viewing
Non-viewing	47.46	0
Wildlife-viewing	0	8.27

Means

Wildlife-viewing	Non-viewing
4.21	23.18

The convergence threshold value (i.e. the minimum change in likelihood for stopping) was 0.01 and the maximum number of iterations was 50 as defined through maximizing model likelihood, although the HMM reached convergence after 22 iterations. Model generation and binary classification for all wildlife-viewing trips in the dataset took < 70 seconds to complete. Manual initialization of the means and diagonal covariance matrix for the Gaussian emission probability density functions were tested, which confirmed that the unsupervised HMM resulted in a more suitable approximation (i.e. best fitting) of the SOG probability density functions and did not require human intervention to reach convergence.

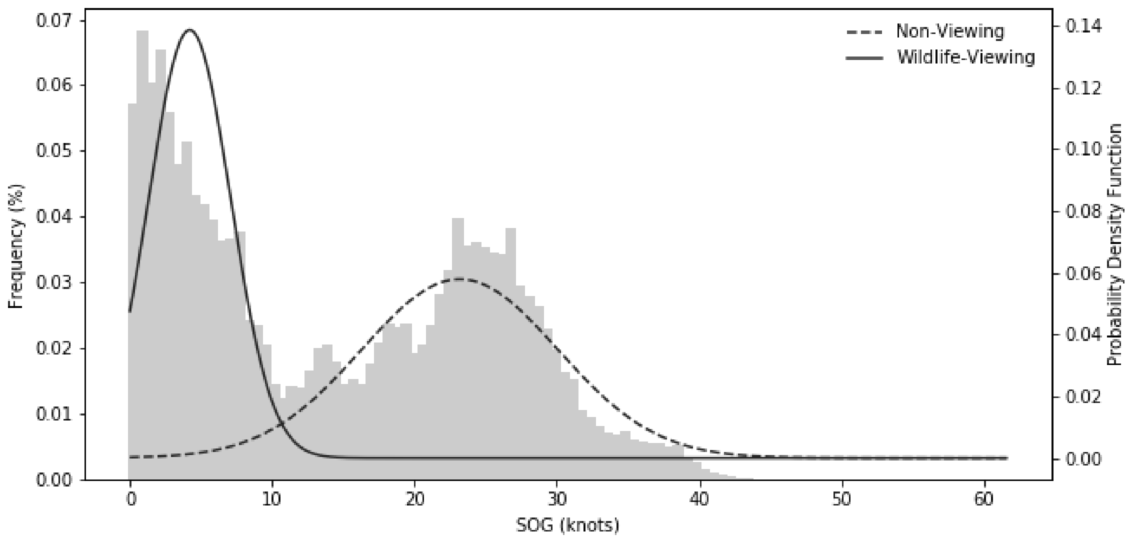

Figure 3.3: Speed profiles for the unsupervised hidden Markov model. Histogram of speed over ground (SOG) observations with Gaussian probability density functions for each hidden state.

3.3.1.3 Model Validation

The HMMs classified the AIS points for all 10,825 wildlife-viewing trips, but only the trips with observed class data could be statistically validated. Specifically, 19 wildlife-viewing trips were used to train the supervised HMM and 8 wildlife-viewing trips were reserved for testing which represents ~0.07% of the total trips. This resulted in 10 wildlife-viewing events made up of 457 AIS points, 22% of which were observed wildlife-viewing points. In comparison, the supervised HMM classified 27,671 wildlife-viewing events and the unsupervised HMM classified 34,082 wildlife-viewing events within the 10,825 wildlife-viewing trips. Importantly, the performance metrics are solely based on observed cetacean sightings, but the models may classify all possible species encountered.

The difference between performance metrics was marginal, but favored the supervised HMM in all categories except the RMSE and Kappa statistic (Table 3.4). Considering these metrics were calculated based solely on cetacean sightings, the unsupervised HMM overall accuracy of 0.75, F-Score of 0.64 and ROC AUS of 0.84 are comparable to the supervised HMM's performance and considered relatively accurate. The ROC AUC values suggest that both HMMs are capable of distinguishing between the wildlife-viewing and non-viewing behaviour classes. The Kappa coefficient concludes that moderate agreement exists between the observed classes and the modelled classes for both HMMs >0.41 (p-value <0.05) (Viera & Garrett, 2005). Although very similar, the log-likelihood scores suggest the supervised HMM has a slightly higher chance of generating the SOG observed in the AIS data.

Table 3.4: Statistical classifier performance metrics.

	Supervised HMM	**Unsupervised HMM**
Overall Accuracy	0.7967	0.7571
Precision	0.5713	0.4737
Recall	0.9920	0.9900
F-Score	0.7250	0.6408
ROC AUC Score	0.8582	0.8409
Kappa Coefficient	0.4826	0. 4897
RMSE*	0.2495	0.2429
Log-likelihood*	$-2.56e^{6}$	$-2.54e^{6}$
***Minimization of these values is preferred.**		

As part of the statistical validation, the supervised and unsupervised HMM modelled classes were individually evaluated for their accuracy using the F-Score (Table 3.5), which is the harmonic mean between precision and recall (Hossin & Sulaiman, 2015). This evaluation determined that the AIS data were skewed toward the non-viewing class. The non-viewing class obtained near perfect precision and was considered accurate (F-Score ~0.82). Conversely, the wildlife-viewing class obtained near perfect recall, which is the fraction of correctly classified wildlife-viewing points retrieved by the HMM over all possible classification output.

Table 3.5: Classification report, statistical performance metrics, for each class.

	Supervised HMM		**Unsupervised HMM**	
	Wildlife-Viewing	Non-Viewing	Wildlife-Viewing	Non-Viewing
Precision	0.47	1.00	0.47	1.00
Recall	1.00	0.68	0.99	0.69
F-Score	0.64	0.81	0.64	0.82
Number of Observations	100	357	100	357

3.3.2 Species Level Trends

The maximum distance and temporal thresholds between the BCCSN cetacean sightings and the observed wildlife-viewing events were 15 km and 75 minutes, respectively. The closest BCCSN cetacean sighting in distance and time within these thresholds was associated with the appropriate HMM classified wildlife-viewing events. This resulted in ~15% of the supervised HMM classified wildlife-viewing events and ~14% of the unsupervised HMM classified wildlife-viewing events connected to a confirmed BCCSN cetacean sighting. Although there were nine possible species categories, only seven species categories were found in these wildlife-viewing events, including grey whale, humpback whale, killer whale, minke whale, unidentified baleen whale, unidentified cetacean, and no recorded species (Table 3.6). Specifically, the no recorded species category represents that there were no observed species associated with a classified wildlife-viewing event, and the unidentified baleen whale and unidentified cetaceans categories represent that there were observed cetaceans present during a wildlife-viewing event, but the species was unknown. Due to the small sample of minke whale, unidentified baleen whale, and unidentified cetaceans sightings associated with wildlife-viewing events, this evaluation only considered the killer whale, humpback whale, and grey whale sightings (Table 3.6).

Table 3.6: Frequency of wildlife-viewing events that are associated with a species category based on the B.C. Cetacean Sightings Network recorded sightings.

	Supervised HMM	**Unsupervised HMM**
No Recorded Species	82.32%	82.74%
Grey Whales	1.60%	1.56%
Humpback Whales	6.73%	6.61%
Killer Whales	9.11%	8.85%
Minke Whales	0.13%	0.12%
Unidentified Baleen Whales	0.01%	0.01%
Unidentified Cetaceans	0.09%	0.1%

The species level trends explored in this study are examined for both the supervised HMM and unsupervised HMM to evaluate vessel movement characteristics. To ensure the most likely wildlife-viewing events were included in the species level analysis, the length of time that each wildlife-viewing trip lasted and wildlife encounter time was explored. It was determined that ~96% of all wildlife-viewing trips were less than 6.6 hours in length, excluding the trips that were less than 15 minutes long, and lasted an average of 3.5 hours. Based on this information, the length of time that each encounter lasted was calculated for each wildlife-viewing event, given that it lasted less than 6.6 hours, but longer than 1 minute. It was discovered that ~95% of the wildlife-viewing encounters lasted less than 90 minutes, and based on the distribution of these data, only the wildlife-viewing events between 1 and 90 minutes were utilized in further analysis. The average encounter length was 22 minutes for the supervised HMM and 20 minutes for the unsupervised HMM. This resulted in 12,382 wildlife-viewing events and 170,337 points for the supervised HMM, and 13,709 wildlife-viewing events and 135, 010 points for the unsupervised HMM.

Through subsetting the wildlife-viewing events based on the species associated with the event, species level temporal, kinematic (SOG), and spatial trends can be observed. A statistical analysis using the Kruskal-Wallis test of the time length and SOG for each encounter was examined for each species (Fig. 3.4 and 3.5). The length of time spent observing the killer whale species tended to be longer than the time spent at humpback whale or grey whale encounters (Table 3.7), which can be seen with the increased frequency between 20 and 60 minutes (Fig. 3.4). On average the length of time spent viewing each species was slightly higher for the supervised HMM, and the wildlife-viewing events lasted longest for killer whale observations (Table 3.7). The Kruskal-Wallis test confirmed that species encounter times are distinctive with p-values <0.05 within each HMM (Table 3.8). Also, the Kruskal-Wallis statistic reported that the grey whale encounter time for the supervised and unsupervised HMM were not significantly different from each other with a p-value of ~0.07 suggesting similarities between the HMM's classifications, but the humpback whale and killer whale encounter times were distinctive between HMMs with p-values < 0.05 (Table 3.8).

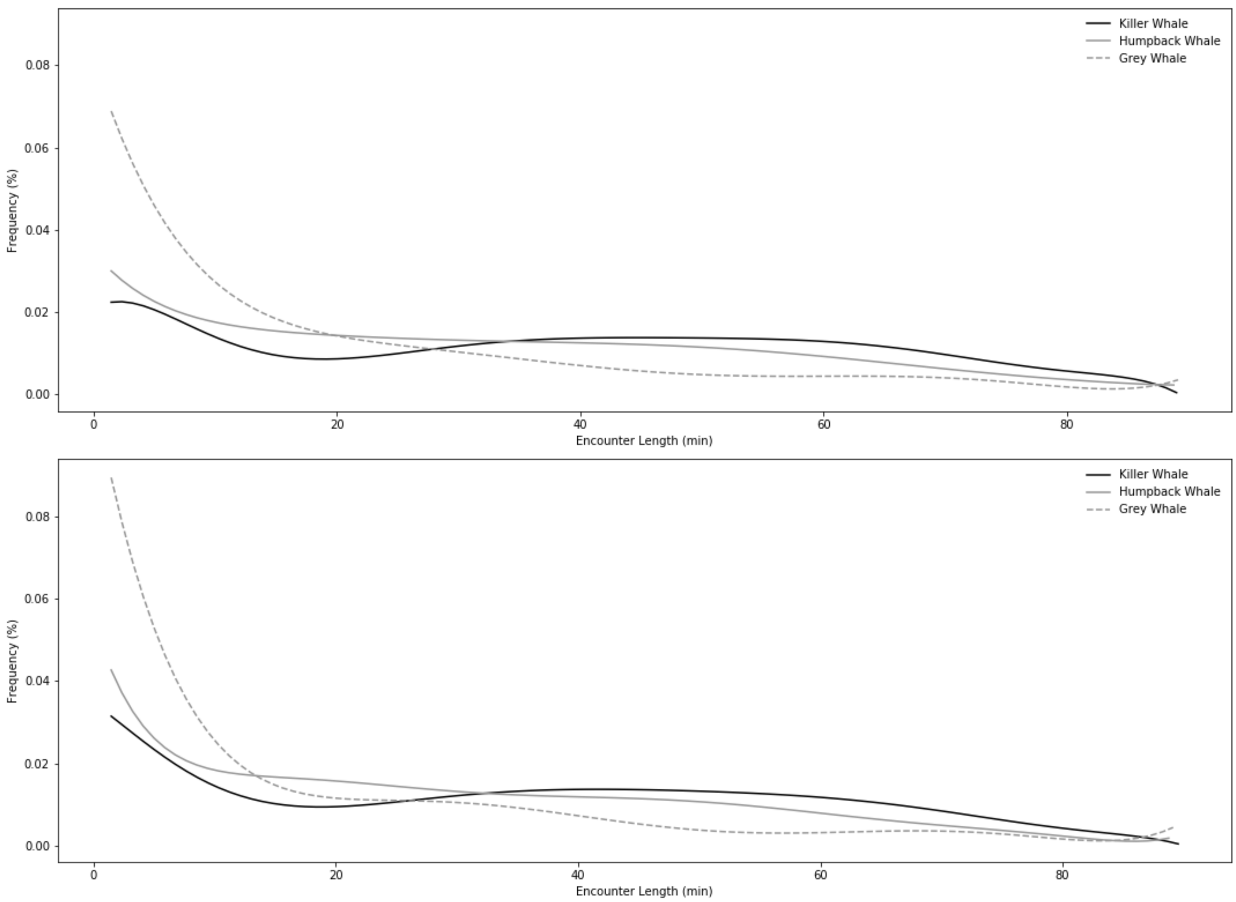

Figure 3.4: Estimated encounter length profiles of each species associated with classifications from the supervised (top) and unsupervised (bottom) hidden Markov models.

Table 3.7: Wildlife-viewing event mean encounter time and speed over ground for the prominent observed species categories.

	Supervised HMM		**Unsupervised HMM**	
	Encounter time (minutes)	Speed over ground (knots)	Encounter time (minutes)	Speed over ground (knots)
Grey Whales	22	3.9	20	3.4
Humpback Whales	33	4.2	30	3.9
Killer Whales	39	4.7	36	4.5

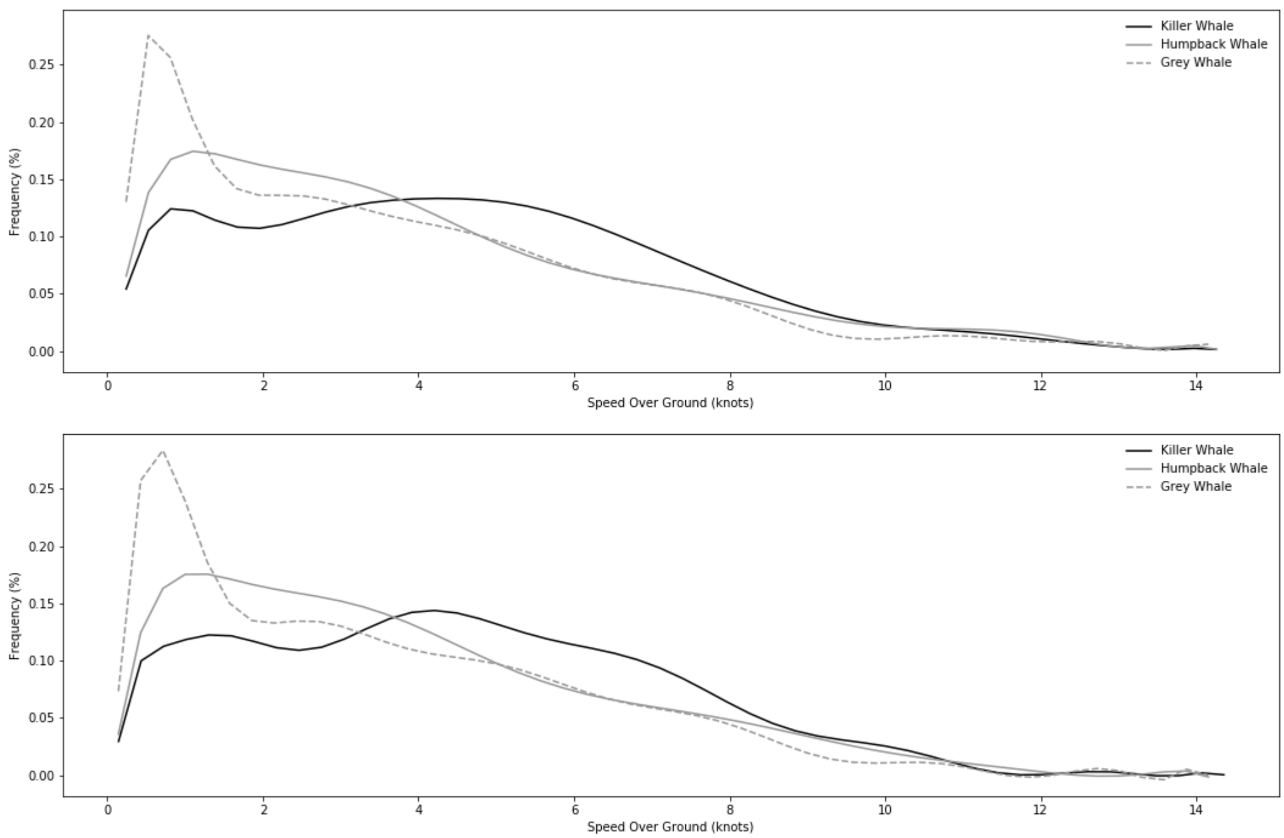

Figure 3.5: Estimated speed over ground profiles of each species associated with classifications from for the supervised (top) and unsupervised (bottom) hidden Markov models.

Table 3.8: Encounter time pairwise H-statistic (p-value) for the Kruskal-Wallis test.

		Supervised HMM		
		Grey Whale	Humpback Whale	Killer whale
Unsupervised HMM	Grey Whale	3.29 (6.94e-2)	83.39 (1.47e-20)	166.74 (3.82e-38)
	Humpback Whale	87.66 (7.75e-21)	18.31 (1.87e-5)	56.78 (4.88e-14)
	Killer whale	180.36 (4.04e-41)	77.09 (1.64e-18)	15.34 (8.99e-5)

Similarly, the SOG distributions for each species subset was explored for any defining qualities (Fig. 3.5). The average SOG for the supervised HMM was slightly faster than the unsupervised HMM (Table 3.7). Both HMMs agree that the average vessel SOG while observing killer whales was faster than the averages for humpback whales and grey whales. The Kruskal-

Wallis test confirmed that the species SOG are distinctive within each HMM, and that each species SOG distributions were distinctive between the supervised and unsupervised HMM approaches with p-values < 0.05 (Table 3.9).

Table 3.9: Speed over ground pairwise H-statistic (p-value) for the Kruskal-Wallis test.

		Supervised HMM		
		Grey Whale	Humpback Whale	Killer whale
Unsupervised HMM	Grey Whale	11.96 (5.42e-4)	76.32 (2.42e-18)	476.63 (1.16e-105)
	Humpback Whale	124.10 (8.02e-29)	35.93 (2.04e-9)	977.63 (1.31e-214)
	Killer whale	646.69 (1.17e-142)	1222.11 (9.55e-268)	32.45 (1.22e-8)

The spatial distribution of wildlife-viewing events was compared between species through kernel density maps of the expected AIS point counts for each HMM (Fig. 3.6), which represent an estimation of vessel frequency for a region. These spatial distributions can infer probable cetacean species distributions that can be compared to their historic distributions to infer HMM performance. It was determined that the vessel frequency of observing humpback whales were most prominent in the Juan de Fuca Strait (south of Victoria) and in the southern Strait of Georgia (west and southwest of Vancouver), which aligns with the humpback whale distributions found in previous studies (Calambokidis et al., 2018; Cominelli et al., 2019; Miller, 2020). The vessel frequency of observing killer whales was more dispersed across the Juan de Fuca Strait and Strait of Georgia with prominent regions near the southern tip of Vancouver Island, in the southern Strait of Georgia near Vancouver, and throughout the Gulf Islands and San Juan Islands. These concentrations of killer whale sightings are indicative of their spatial distribution found in previous studies (Cominelli et al., 2019; Olson et al., 2018). Clayoquot Sound was the prominent region for grey whale observations, but it did not exceed 15 AIS points per 1 km grid cell. This location is a known feeding area for the grey whale during its migration between California and Alaska (Urbán R et al., 2021). Based on these observations, it can be inferred that both the supervised and unsupervised HMM have potential for classifying species-specific wildlife-viewing effort.

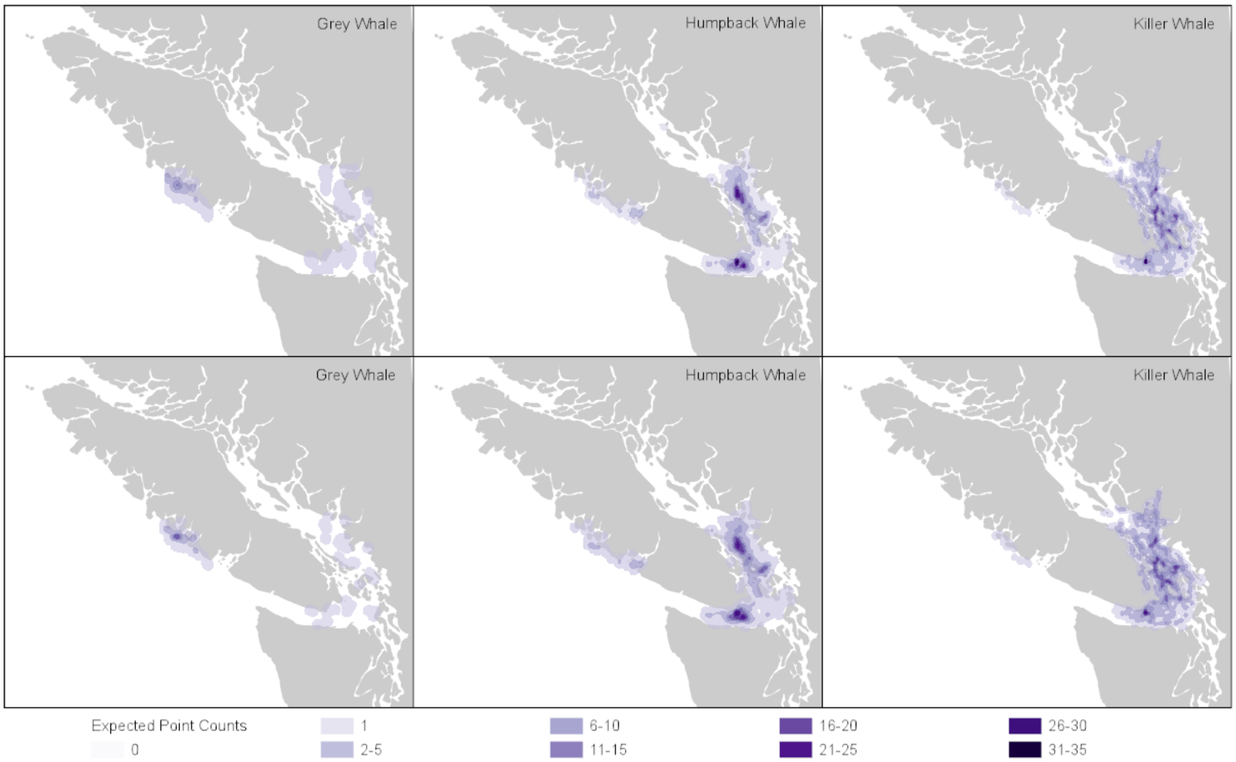

Figure 3.6: Spatial distributions of wildlife-viewing AIS points for the supervised (top) and unsupervised (bottom) hidden Markov models.

3.4 Discussion

The overall objective of this study was to determine if an unsupervised HMM is comparable to a supervised HMM in the classification of wildlife-viewing vessel behaviour, and to determine if the HMMs' classifications could be used to gather vessel movement information that may indicate which target species was being observed. Therefore, their performances were assessed followed by the integration of BCCSN cetacean sightings data for an evaluation of wildlife-viewing event vessel movement characteristics. The findings from this study suggest that the unsupervised approach was comparable to the supervised HMM in classifying wildlife-viewing events, and that vessel movement characteristics were distinctive between species supporting the possibility of a multiclass species level classification model.

This research strived to develop an unsupervised HMM with comparable performance to a supervised HMM that can be employed generally without region specific domain knowledge and dependence on observed class data. Supervised approaches require regional domain specific knowledge, are dependant on observed class data, and require customization of model parameters. For example, the supervised HMM was developed with specific ecotourism knowledge of the vessel's homeport and regular wildlife encounters at known sites. Incorporating region specific assumptions for a large geographic area and commercial ecotourism fleet would be taxing. The use of the unsupervised HMM attempted to circumvent these requirements and was able to do so without compromising performance. The only knowledge required for the unsupervised approach was the number of behaviour states. Through this approach, the HMM was successfully applied to a broad geographic region and a variety of commercial ecotourism vessels, but this generalization is interconnected to model interpretability through data dependence, parameter re-estimation, and model performance (Cazzanti & Pallotta, 2015); fortunately, this study was able to utilize the observed class data to assess the interpretability and performance of the unsupervised HMM.

Although supervised HMMs are dependent on sufficient annotated observed class data, they are still favoured because of their computational efficiency, interpretability, and ability to capture information (Boussemart et al., 2011; Cazzanti & Pallotta, 2015; Zhang et al., 2005). The performance metrics in this study favoured the supervised HMM, but the difference between HMMs was not significant. This, coupled with the fact that ~0.2% of the wildlife-viewing trips had observed class data used for validation suggests that there may be insufficient observed class data for the supervised HMM to outperform the unsupervised HMM. Without access to precise wildlife encounter data for multiple wildlife-viewing vessels, the development of the supervised HMM and the statistical validation of these models may be biased toward the single wildlife-viewing vessel that collected observed cetacean sightings. Nonetheless, these results support the claim that the unsupervised HMM is an accurate classifier of wildlife-viewing vessel behaviour and less data reliant.

As expected, the supervised HMM was more computationally efficient than the unsupervised HMM, taking ~62 seconds less time for model generation and classification. These execution times are relatively quick considering the unsupervised HMM was able to learn the model parameters from the AIS observed SOG data without human intervention. However, the implementation of the supervised HMM required human intervention for observed class data preparation, model development of custom emission probabilities, and parameter re-estimation and testing. The automation, generalization, and data independence the unsupervised HMM affords makes it appealing for further model development, and for its use in studies that inform marine policy.

Species level vessel movement characteristics were explored to assess the HMMs potential for classifying the probable target species being observed based on the AIS data. BCCSN cetacean sightings data provided information on the probable species being observed for ~15% of the classified wildlife-viewing events for each HMM, which were then assessed for variations in vessel movement characteristics. The variations in the encounter time and SOG distributions between killer whale, humpback whale, and grey whale were significantly distinct (Kruskal-Wallis $p\text{-value} < 0.05$). For example, vessels tended to observe killer whales for longer (Fig. 3.4), and the vessels tended to travel at slightly faster speeds (Fig. 3.5). These vessel characteristics may be attributed to the popularity of observing killer whales, their variable spatial distribution (Fig. 3.6), or more frequent surface activities providing more opportunities to be observed (Cominelli et al., 2019; Noren et al., 2009; Olson et al., 2018). For example, Noren et al. (2009) found that when vessels closely approached southern reseident killer whales (SRKW) surface active behaviour was likely to occur. This is indicative to how wildlife-viewing vessels typically move alongside killer whales, but can remain relatively stationary while observing humpback whales because humpback whales are commonly viewed while feeding or travelling, which can include frequent dives resulting in vessels waiting for the animal to surface (Schuler et al., 2019). The distinctive vessel movement characteristics found in this study demonstrate how these methods can utilize AIS data to supplement established cetacean tracking data. Monitoring population trends and movements of highly mobile, migratory species is a daunting and expensive task often requiring dedicated surveys and specialized surveillance equipment. Deriving information from additional data sources such as AIS data could be a relatively opportunistic, timely and cost effective way to provide additional insights into the movement patterns of these animals.

Given the findings that the SOG and encounter time distributions are distinct between species, it is possible that the HMM classification of AIS data can be a proxy for the wildlife being observed. This leads to the potential for the unsupervised HMM to be developed as a multiclass classification method to detect the probable target species being observed. Without access to precise target species encounter data, the supervised HMM does not have the same potential. Future studies could explore this possibility by developing an unsupervised multiclass HMM in a univariate fashion with SOG, a multivariate fashion with SOG and geographic location, or the use of a two-step model that incorporates encounter time length as a secondary phase after the wildlife-viewing classification.

As a demonstrative example, the unsupervised HMM from this study can assist in the assessment of the voluntary slowdown measures that were implemented by the Enhancing Cetacean Habitat and Observation (ECHO) Program (Vancouver Fraser Port Authority, 2020). These measures were implemented from July 5, 2019 to October 15, 2019 in Haro Strait and Boundary Pass (high vessel traffic regions in the Salish Sea) to reduce vessel noise disturbances and protect the SRKW. As such, they are triggered to begin upon the SRKW arrival to this region, but these measures have the capacity to protect other species. For example, vessel speed restrictions can reduce the lethality and number of collisions vessels have with whales (Conn & Silber, 2013; Wiley et al., 2011). Therefore, the slowdown can help protect other cetacean populations (e.g. humpback whales) vulnerable to vessel strikes when they are in the region (Williams & O'Hara, 2010). In addition to the SRKW, the transient killer whales (Bigg's killer whales) are also impacted by vessel noise (Smith, 2017); therefore, transient killer whales in the region are also benefitting from the slowdown measures. Data for the SRKW population was included in the killer whale category in this study, so the killer whale category was more closely examined. The annual temporal trends in wildlife-viewing events were compared to the dates of the voluntary slowdown, which determined that these dates captured ~59% of the total wildlife-viewing events classified for 2019 and ~46% of the wildlife-viewing events associated with killer whale (SRKW and transient) sightings (Fig. 3.7). By examining these annual wildlife-viewing event trends, it can be seen that species other than SRKW could benefit from the slowdown measures, and that the methods developed in this study have the potential to be an early warning system of killer whale presence in the region.

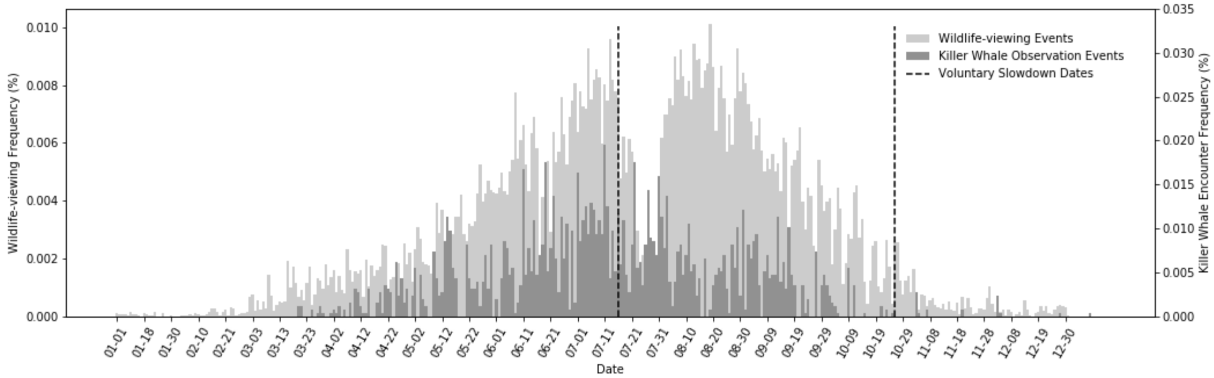

Figure 3.7: Histogram of the frequency of the unsupervised HMM classifed wildlife-viewing events, and killer whale encounters, for each day of 2019 with voluntary slowdown dates displayed. Note the secondary y-axis (right) for the killer whale wildlife encounters.

Additionally, compliance with regulations within Transport Canadas sanctuary zones, which are in place to prohibit vessel movement (Fisheries and Oceans Canada, 2019), can be considered by examining the sanctuary zone boundaries and extracting spatial statistics from the distribution of wildlife-viewing events. A cursory analysis determined that there exists vessel traffic in two of the three sanctuary zones. The Boundary Pass sanctuary zone contained 124 wildlife-viewing points, and the sanctuary zone on the southwest shores of the Pender Islands contained 212 wildlife-viewing points. These findings suggest that, although vessels are prohibited from viewing cetaceans in these areas, wildlife-viewing by some operators within these zones is still occurring. One limitation to applications of these methods is that not all commercial ecotourism vessels are required to carry AIS, so these estimates of vessel frequency to a region may not be complete. Regardless, these applications demonstrate how the methods developed in this study can inform cetacean conservation research and vessel management practices, and have the potential for retroactive regulation enforcement.

Although this study finds that both the supervised and unsupervised HMM are suitable for the classification of wildlife-viewing events, these models are limited by the quality of AIS data being used and the variety of vessels that are represented by the AIS data. High-resolution AIS data are recommended for the implementation of either HMM, and precise wildlife encounter data are recommended for the supervised HMM, model validation, or the further development and exploration of species level vessel behaviour. Due to limitations on the types of commercial ecotourism vessels (i.e. longer than 8m or that carry 12 or more passengers) that are AIS enabled, these models may not capture all small vessel movement. Future studies into how the size of the vessel impacts these models would be valuable.

Additionally, the BCCSN cetacean sightings data are crowdsourced using an online platform and can be collected by professional or recreational wildlife observers at sea or from shore, which needs to be considered when assessing the species level trends in the HMM classes. Therefore, there could be potential issues with these data, such as validating species identification or bias toward more easily identifiable species (e.g. killer whales may be more identifiable by their prominent fin). Because this study's findings suggest that an AIS based multiclass model may be possible without reliance on BCCSN data, future work should explore the implementation of such a model. This would provide an additional source of passive cetacean detection data. Moreover, Cazzanti & Pallotta (2015) argue that trajectory-based machine learning could lead to a “richer knowledge base for downstream analyses”, so it is recommended that future studies explore the classification of commercial ecotourism vessel trajectories. Alongside this work, obtaining more precise wildlife-viewing encounter data could unlock a more powerful HMM. This paired with the addition of multiannual AIS data may be able to provide a framework for studying the classification of recreational wildlife-viewing vessel behaviour.

Overall, this study's evaluation of a supervised and unsupervised HMM for the classification of wildlife-viewing behaviour, and exploration of species level boater behaviour, using commercial whale-watching vessels aims to contribute to scientific endeavours in marine traffic management, wildlife conservation, marine spatial planning, and marine policy studies, ultimately aiding marine managers in decision making. The potential for industry-focused contributions based on these models may provide a means for commercial ecotourism fleet management, such as the potential to reassure operators that their vessels are complying with guidelines, or to verify contested regulation infractions. Additionally, these models, trained using commercial whale-watching vessels, have the potential to be applied to other types of AIS enabled vessels engaged in wildlife-viewing (i.e. recreational whale-watching vessels) to help better understand their behaviour and compliance to conservation regulations.

3.5 Literature Cited

Allen, A. M., & Singh, N. J. (2016). Linking movement ecology with wildlife management and conservation. *Frontiers in Ecology and Evolution*, 3:155, 1–13. https://doi.org/10.3389/fevo.2015.00155

Almunia, J., Delponti, P., & Rosa, F. (2020). Using big data to estimate whale watching effort. *bioRxiv*. https://doi.org/10.1101/2020.11.30.403923

Baum, L. E., & Petrie, T. (1966). Statistical inference for finite state markov chains. *The Annals of Mathematical Statistics*, 37, 1554–1563.

B.C. Cetaceans Sightings Network (BCCSN). (2021). How sightings are used. https://wildwhales.org/sightings/how-sightings-are-used/

Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J., Connor, R., Heithaus, M., Watson-Capps, J., Flaherty, C., & Krützen, M. (2006). Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conservation Biology*, 20(6), 1791–1798. https://doi.org/10.1111/j.1523-1739.2006.00540.x

Boussemart, Y., Cummings, M. L., Las Fargeas, J., & Roy, N. (2011). Supervised vs unsupervised learning for operator state modeling in unmanned vehicle settings. *Journal of Aerospace Computing, Information and Communication*, 8(3), 71–85. https://doi.org/10.2514/1.46767

Calambokidis, J., Flynn, K., Dobson, E., Huggins, J. L., & Perez, A. (2018). Return of the Giants of the Salish Sea: Increased occurrence of humpback and gray whales in inland waters. *Salish Sea Ecosystem Conference*; Extended Abstract.

Calambokidis, J., Laake, J., & Pérez, A. (2014). Updated analysis of abundance and population structure of seasonal gray whales in the Pacific Northwest, 1996-2012. In *Final report to the National Marine Fisheries Service*.

Canadian Shipping Act: Regulations Amending the Navigation Safety Regulations (Automatic Identification Systems), SOR/2019-100 (2019). *Canada Gazette Part II*, 153(9). https://gazette.gc.ca/rp-pr/p2/2019/2019-05-01/html/sor-dors100-eng.html

Cazzanti, L., & Pallotta, G. (2015). Mining maritime vessel traffic: Promises, challenges, techniques. *OCEANS 2015 - Genova*, 1–6. https://doi.org/10.1109/OCEANS-Genova.2015.7271555

Charles, C., Gillis, D., & Wade, E. (2014). Using hidden Markov models to infer vessel activities in the snow crab (*Chionoecetes opilio*) fixed gear fishery and their application to catch standardization. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(12), 1817–1829. https://doi.org/10.1139/cjfas-2013-0572

Çiğşar, B., & Ünal, D. (2019). Comparison of Data Mining Classification Algorithms Determining the Default Risk. *Scientific Programming*, 2019. https://doi.org/10.1155/2019/8706505

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, XX(1), 37–46.

Cominelli, S., Leahy, M., Devillers, R., & Hall, B. G. (2019). Geovisualization tools to inform the management of vessel noise in support of species' conservation. *Ocean and Coastal Management*, 169, 113–128. https://doi.org/10.1016/j.ocecoaman.2018.11.009

Conn, P. B., & Silber, G. K. (2013). Vessel speed restrictions reduce risk of collision-related mortality for North Atlantic right whales. *Ecosphere*, 4(4) :43. https://doi.org/10.1890/ES13-00004.1

Dahlheim, M., & Castellote, M. (2016). Changes in the acoustic behavior of gray whales *Eschrichtius robustus* in response to noise. *Endangered Species Research*, 31(1), 227–242. https://doi.org/10.3354/esr00759

Dahlheim, M. E., & White, P. A. (2010). Ecological aspects of transient killer whales *Orcinus orca* as predators in southeastern Alaska. *Wildlife Biology*, 16(3), 308–322. https://doi.org/10.2981/09-075

De Souza, E. N., Boerder, K., Matwin, S., & Worm, B. (2016). Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PLoS ONE*, 11(7). https://doi.org/10.1371/journal.pone.0158248

Erbe, C. (2002). Underwater noise of whale-watching boats and potential effects on killer whales (Orcinus orca), based on an acoustic impact model. *Marine Mammal Science*, 18(2), 394–418.

Findley, L. T., & Vidal, O. (2002). Gray whale (*Eschrichtius robustus*) at calving sites in the Gulf of California, México. *Journal of Cetacean Research and Management*, 4(1), 27–40.

Fisheries and Oceans Canada. (2019). Government of Canada outlines its 2019 plan for protecting Southern Resident killer whales. https://www.canada.ca/en/fisheries-oceans/news/2019/05/government-of-canada-outlines-its-2019-plan-for-protecting-southern-resident-killer-whales.html

Fisheries and Oceans Canada. (2020) Watching marine wildlife. https://www.dfo-mpo.gc.ca/species-especes/mammals-mammiferes/watching-observation/index-eng.html

Fouda, L. (2012). Noisy Neighbours-using Automatic Identification System (AIS) and passive acoustic monitoring data to measure individual vessel source levels in critical whale habitat.

Frankel, A. S., & Gabriele, C. M. (2017). Predicting the acoustic exposure of humpback whales to cruise and tour vessels in Glacier Bay, Alaska. *Endangered Species Research*, 34, 397–415.

Fraser, M. D., McWhinnie, L. H., Canessa, R. R., & Darimont, C. T. (2020). Compliance of small vessels to minimum distance regulations for humpback and killer whales in the Salish Sea. *Marine Policy*, 121, 104171. https://doi.org/10.1016/j.marpol.2020.104171

Gillespie, D., Palmer, L., Macaulay, J., Sparling, C., & Hastie, G. (2020). Passive acoustic methods for tracking the 3D movements of small cetaceans around marine structures. *PLoS ONE*, 15(5), 1–16. https://doi.org/10.1371/journal.pone.0229058

Government of Canada, 2019. CPC-2-3-07 - Obtaining Identities in the Maritime Mobile Service. https://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/sf01032.html

Grewal, J. K., Krzywinski, M., & Altman, N. (2019). Markov models — hidden Markov models. *Nature Methods*, 16(9), 795–796. https://doi.org/10.1038/s41592-019-0532-6

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

Hermannsen, L., Mikkelsen, L., Tougaard, J., Beedholm, K., Johnson, M., & Madsen, P. T. (2019). Recreational vessels without Automatic Identification System (AIS) dominate anthropogenic noise contributions to a shallow water soundscape. *Scientific Reports*, 9, 15477. https://doi.org/10.1038/s41598-019-51222-9

Hossin, M., & Sulaiman, M. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2), 1–11. https://doi.org/10.5121/ijdkp.2015.5201

International Maritime Organization (IMO). (2003). Guidelines for the Installation of a Shipborne Automatic Identification System (AIS) (Ref. T2/8.02) https://www.navcen.uscg.gov/pdf/marcomms/imo/Circulars/IMO.SN.Circ.227_AIS_Installation.pdf

International Whaling Commission (IWC), 1983. Chairman's Report of the Thirty-Fourth Annual Meeting. *Reports of the International Whaling Commission 33*.

Joo, R., Bertrand, S., Tam, J., & Fablet, R. (2013). Hidden Markov Models: The Best Models for Forager Movements? *PLoS ONE*, 8(8): e71246. https://doi.org/10.1371/journal.pone.0071246

Joy, R., Tollit, D., Wood, J., MacGillivray, A., Li, Z., Trounce, K., & Robinson, O. (2019). Potential benefits of vessel slowdowns on endangered southern resident killer whales. *Frontiers in Marine Science*, 6(344), 1–20. https://doi.org/10.3389/fmars.2019.00344

Khodabandelou, G., Hug, C., Deneckère, R., & Salinesi, C. (2014). Supervised vs. unsupervised learning for intentional process model discovery. *Lecture Notes in Business Information Processing*, 175 LNBIP, 215–229. https://doi.org/10.1007/978-3-662-43745-2_15

Kouemou, G. L. (Eds.). (2011). History and Theoretical Basics of Hidden Markov Models. *Hidden Markov Models, Theory and Applications*. https://doi.org/10.5772/15205

Krasnova, V. V., Prasolova, E. A., Belikov, R. A., Chernetsky, A. D., & Panova, E. M. (2020). Influence of boat tourism on the behaviour of Solovetskiy beluga whales (*Delphinapterus leucas*) in Onega Bay, the White Sea. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30(10), 1922–1933. https://doi.org/10.1002/aqc.3369

Laake, J. L., Calambokidis, J., Osmek, S. D., & Rugh, D. J. (1997). Probability of Detecting Harbor Porpoise from Aerial Surveys: Estimating g(0). *The Journal of Wildlife Management*, 61(1), 63–75.

Lachmuth, C. L., Barrett-lennard, L. G., Steyn, D. Q., & Milsom, W. K. (2011). Estimation of southern resident killer whale exposure to exhaust emissions from whale-watching vessels and potential adverse health effects and toxicity thresholds. *Marine Pollution Bulletin*, 62(4), 792–805. https://doi.org/10.1016/j.marpolbul.2011.01.002

Le Tixerant, M., Le Guyader, D., Gourmelon, F., & Queffelec, B. (2018). How can Automatic Identification System (AIS) data be used for maritime spatial planning*? Ocean and Coastal Management*, 166, 18–30. https://doi.org/10.1016/j.ocecoaman.2018.05.005

Lundquist, D., Sironi, M., Würsig, B., Rowntree, V., Martino, J., & Lundquist, L. (2013). Response of southern right whales to simulated swim-with-whale tourism at Península Valdés, Argentina. Marine Mammal Science, 29(2), 24–45. https://doi.org/10.1111/j.1748-7692.2012.00583.x

Lusseau, D. (2006). The short-term behavioral reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Marine Mammal Science*, 22(4), 802–818. https://doi.org/10.1111/j.1748-7692.2006.00052.x

Mallard, G. (2019). Regulating whale watching: A common agency analysis. *Annals of Tourism Research*, 76, 191–199. https://doi.org/10.1016/j.annals.2019.04.011

Mate, B. R., Nieukirk, S. L., & Kraus, S. D. (1997). Satellite-monitored movements of the northern right whale. *Journal of Wildlife Management*, 61(4), 1393–1405.

Markov, A. A. (2006). Classical text in translation: An example of statistical investigation of the text *Eugene Onegin* concerning the connection of samples in chains. *Science in Context*, 19(4), 591–600. https://doi.org/10.1017/S0269889706001074

McWhinnie, L. H., O'Hara, P. D., Hilliard, C., Le Baron, N., Smallshaw, L., Pelot, R., & Canessa, R. (2021). Assessing vessel traffic in the Salish Sea using satellite AIS: An important contribution for planning, management and conservation in southern resident killer whale critical habitat. *Ocean and Coastal Management*, 200, 1–17. https://doi.org/10.1016/j.ocecoaman.2020.105479

Meijles, E. W., Daams, M. N., Ens, B. J., Heslinga, J. H., & Sijtsma, F. J. (2021). Tracked to protect - Spatiotemporal dynamics of recreational boating in sensitive marine natural areas. *Applied Geography*, 130(102441). https://doi.org/10.1016/j.apgeog.2021.102441

Meissner, A. M., Christiansen, F., Martinez, E., Pawley, M. D. M., Orams, M. B., & Stockin, K. A. (2015). Behavioural effects of tourism on oceanic common dolphins, Delphinus sp., in New Zealand: The effects of markov analysis variations and current tour operator compliance with regulations. *PLoS ONE*, 10(1), 1–23. https://doi.org/10.1371/journal.pone.0116962

Miller, H. (2020). Relating the Distribution of Humpback Whales to Environmental Variables and Risk Exposure. In *University of Washington*.

Moore, S., & Huntington, H. (2008). Arctic Marine Mammals and Climate Change: Impacts and Resilience. *Ecological Applications*, 18(2), S157–S165. https://www-jstor-org.ezproxy.library.uvic.ca/stable/40062163?pq-origsite=summon&seq=1#metadata_info_tab_contents

Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. In *The MIT Press* (Issue 4).

Noren, D. P., Johnson, A. H., Rehder, D., & Larson, A. (2009). Close approaches by vessels elicit surface active behaviors by southern resident killer whales. *Endangered Species Research*, 8(3), 179–192. https://doi.org/10.3354/esr00205

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. https://doi.org/10.1121/1.427071

Olson, J. K., Wood, J., Osborne, R. W., Barrett-lennard, L., & Larson, S. (2018). Sightings of southern resident killer whales in the Salish Sea 1976−2014: the importance of a long-term opportunistic dataset. *Endangered Species Research*, 37, 105–118. https://doi.org/https://doi.org/10.3354/esr00918

Pallotta, G., Vespe, M., & Bryan, K. (2013). Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. *Entropy*, 15(6), 2218–2245. https://doi.org/10.3390/e15062218

Parrott, L., Chion, C., Martins, C. C. A., Lamontagne, P., Turgeon, S., Landry, J. A., Zhens, B., Marceau, D. J., Michaud, R., Cantin, G., Ménard, N., & Dionne, S. (2011). A decision support system to assist the sustainable management of navigation activities in the St. Lawrence River Estuary, Canada. *Environmental Modelling and Software*, 26, 1403–1418. https://doi.org/10.1016/j.envsoft.2011.08.009

Peel, D., & Good, N. M. (2011). A hidden markov model approach for determining vessel activity from vessel monitoring system data. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(7), 1252–1264. https://doi.org/10.1139/f2011-055

Pirotta, E., Matthiopoulos, J., MacKenzie, M., Scott-Hayward, L., & Rendell, L. (2011). Modelling sperm whale habitat preference: A novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436, 257–272. https://doi.org/10.3354/meps09236

Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2), 257–286. https://doi.org/10.1109/5.18626

Robards, M. D., Silber, G. K., Adams, J. D., Arroyo, J., Lorenzini, D., Schwehr, K., & Amos, J. (2016). Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)-A review. *Bulletin of Marine Science*, 92(1), 75–103. https://doi.org/10.5343/bms.2015.1034

Schuler, A. R., Piwetz, S., Di Clemente, J., Steckler, D., Mueter, F., & Pearson, H. C. (2019). Humpback Whale Movements and Behavior in Response to Whale-Watching Vessels in Juneau, AK. *Frontiers in Marine Science*, 6:710, 1–13. https://doi.org/10.3389/fmars.2019.00710

Schwehr, K. D., & McGillivary, P. A. (2007). Marine ship automatic identification system (AIS) for enhanced coastal security capabilities: An oil spill tracking application. *Oceans Conference Record (IEEE)*. https://doi.org/10.1109/OCEANS.2007.4449285

Sheng, K., Liu, Z., Zhou, D., He, A., & Feng, C. (2018). Research on Ship Classification Based on Trajectory Features. *Journal of Navigation*, 71(1), 100–116. https://doi.org/10.1017/S0373463317000546

Shields, M. W., Hysong-Shimazu, S., Shields, J. C., & Woodruff, J. (2018). Increased presence of mammal-eating killer whales in the Salish Sea with implications for predator-prey dynamics. *PeerJ*, 2018(12). https://doi.org/10.7717/peerj.6062

Sidibé, A., & Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. *Journal of Navigation*, 70(4), 847–858. https://doi.org/10.1017/S0373463317000066

Smith, C. E. (2017). Evaluating Social Network Dynamics of Bigg's Killer Whales (*Orcinus orca*) and Vessel Traffic within a Transboundary Region: Implications for Conservation Management. *Dissertations*. 1485. https://search.proquest.com/docview/1976785768?accountid=14169

Soe, O. K. M., Chaojian, S., Qinyou, H., & Weintrit, A. (2010). Clustering Analysis and Identification of Marine Traffic Congested Zones. *Zeszyty Naukowe Akademii Morskiej W Gdyni*, 67, 101–113.

Sveegaard, S., Galatius, A., Dietz, R., Kyhn, L., Koblitz, J. C., Amundin, M., Nabe-Nielsen, J., Sinding, M. H. S., Andersen, L. W., & Teilmann, J. (2015). Defining management units for cetaceans by combining genetics, morphology, acoustics and satellite tracking. *Global Ecology and Conservation*, 3, 839–850. https://doi.org/10.1016/j.gecco.2015.04.002

Toloue, K. F., & Jahan, M. V. (2018). Anomalous behavior detection of marine vessels based on Hidden Markov Model. *2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems*, 10–12. https://doi.org/10.1109/CFIS.2018.8336611

Trites, A. W. (2014). The Marine Mammals. In R. J. Beamish & G. A. McFarlane (Eds.), *The Sea Among Us: The Amazing Strait of Georgia.* Harbour Publishing.

Urbán R, J., Jiménez-López, E., Guzmán, H. M., & Viloria-Gómora, L. (2021). Migratory Behavior of an Eastern North Pacific Gray Whale From Baja California Sur to Chirikov Basin, Alaska. *Frontiers in Marine Science*, 8:619290, 1–7. https://doi.org/10.3389/fmars.2021.619290

Van Parijs, S. M., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., & Van Opzeeland, I. C. (2009). Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Marine Ecology Progress Series*, 395, 21–36. https://doi.org/10.3354/meps08123

Vancouver Fraser Port Authority. (2020). ECHO Program 2019 voluntary vessel slowdown trial in Haro Strait and Boundary Pass: Summary findings.

Vanderlaan, A. S. M., & Taggart, C. T. (2007). Vessel collisions with whales: The probability of lethal injury based on vessel speed. *Marine Mammal Science*, 23(1), 144–156. https://doi.org/10.1111/j.1748-7692.2006.00098.x

Vermard, Y., Rivot, E., Mahévas, S., Marchal, P., & Gascuel, D. (2010). Identifying fishing trip behaviour and estimating fishing effort from VMS data using Bayesian Hidden Markov Models. *Ecological Modelling*, 221(15), 1757–1769. https://doi.org/10.1016/j.ecolmodel.2010.04.005

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 37(5), 360–363.

http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf

Vinding, K., Bester, M., Kirkman, S. P., Chivell, W., & Elwen, S. H. (2015). The use of data from a platform of opportunity (whale watching) to study coastal cetaceans on the southwest coast of South Africa. *Tourism in Marine Environments*, 11(1), 33–54. https://doi.org/10.3727/154427315X14398263718439

Waterbolk, M., Tump, J., Klaver, R., van der Woude, R., Velleman, D., Zuidema, J., Koch, T., & Dugundji, E. (2019). Detection of Ships at Mooring Dolphins with Hidden Markov Models. *Transportation Research Record*, 2673(4), 439–447. https://doi.org/10.1177/0361198119837495

Wiley, D. N., Thompson, M., Pace, R. M., & Levenson, J. (2011). Modeling speed restrictions to mitigate lethal collisions between ships and whales in the Stellwagen Bank National Marine Sanctuary, USA. *Biological Conservation*, 144, 2377–2381. https://doi.org/10.1016/j.biocon.2011.05.007

Williams, R., & O'Hara, P. (2010). Modelling ship strike risk to fin, humpback and killer whales in British Columbia, Canada. *Journal of Cetacean Research and Management*, 11(1), 1–8.

Williams, R., Trites, A. W., & Bain, D. E. (2002). Behavioural responses of killer whales (*Orcinus orca*) to whale-watching boats: opportunistic observations and experimental approaches. *Journal of Zoology*, 256, 255–270. https://doi.org/10.1017/s0952836902000298

Wilson, R. P., Liebsch, N., Davies, I. M., Quintana, F., Weimerskirch, H., Storch, S., Lucke, K., Siebert, U., Zankl, S., Müller, G., Zimmer, I., Scolaro, A., Campagna, C., Plötz, J., Bornemann, H., Teilmann, J., & McMahon, C. R. (2007). All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 54, 193–210. https://doi.org/10.1016/j.dsr2.2006.11.017

Zhang, D., Gatica-Perez, D., Bengio, S., & McCowan, L. (2005). Semi-supervised adapted HMMs for unusual event detection. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, 611–618. https://doi.org/10.1109/CVPR.2005.316

Zucchini, W., Macdonald, I. L., & Langrock, R. (Eds.) (2017). Hidden Markov models for time series: An introduction using R (2nd ed.). In *Monographs on Statistics and Applied Probability* 150. https://doi.org/10.1201/b20790

Chapter 4 Conclusion

Whale-watching is argued to be a sustainable form of human-wildlife interactions (Mallard, 2019; Schuler et al., 2019), but the global increase in whale-watching vessel traffic can still have negative impacts on the species being targeted (Conn & Silber, 2013; Findley & Vidal, 2002; Lusseau, 2006; Williams et al., 2002). Therefore, it is important to marine managers, and whale-watching stakeholders, that wildlife-viewing vessels adhere to the regulations and guidelines in place that aim to mitigate these negative impacts to cetaceans and their habitat. Fortunately, there exists a vessel information transmission system called AIS – Automatic Identification Systems – that can track whale-watching vessel movement. These data can facilitate the classification of wildlife-viewing behaviour, which is required to evaluate to what degree whale-watching vessels are compliant. As such, reliable classification methods that are applicable across a geographically broad commercial whale-watching fleet are essential to better understand vessel movement characteristics while observing wildlife. Therefore, the goal of this research was to evaluate established classification models, and their approaches, for their feasibility in classifying when commercial whale-watching vessels are engaged in wildlife-viewing, and to determine to what degree vessel movement can indicate which cetacean species is being observed.

This research focused on commercial whale-watching vessels to first perform a systematic evaluation of three established classification models for their performance in classifying wildlife-viewing behaviour. These models were implemented and validated with AIS data obtained for a single whale-watching vessel that was used to collect observed cetacean encounter data. Then, the best performing classification model was selected for generalization across an extensive geographic region and its commercial ecotourism fleet. A supervised and unsupervised approach to this model were assessed to ensure that the unsupervised approach had comparable performance to the supervised model, thus reducing reliance on observed behaviour class data. Finally, to determine the extent that AIS data may be used in wildlife-viewing classification, these models' potential for the classification of the species targeted for observation was assessed.

4.1 Classification Models

The preliminary evaluation of classification models for the detection of wildlife-viewing events using AIS data found that, although all three models were accurate classifiers for commercial ecotourism vessel behaviour, logistic regression (LR) and the hidden Markov model (HMM) outperformed density-based spatial clustering application with noise (DBSCAN). This thesis concluded that the HMM was best suited for further study due to its comparable accuracy to LR and ease of implementation, requiring minimal data preparation and a single AIS variable (speed over ground). Fortunately, the use of LR provided insight into which AIS variables were most influential on wildlife-viewing behaviour classification supporting the use of speed over ground (SOG) for the HMM. Generally, the findings from this evaluation determined that these three models were sensitive to parameter optimization and assumptions, and that region specific domain knowledge is required for accurate model development.

The classification models evaluated in this thesis obtained favorable performance metrics, but they each come with limitations and recommendations. The implementation of DBSCAN is sensitive to outliers and parameter estimation, and it requires the model to be implemented iteratively for each wildlife-viewing trip, which cannot be deployed until each wildlife-viewing trip is complete. Since DBSCAN underperformed compared to the HMM and LR the exploration into alternative clustering methods, such as TREAD (Traffic Route Extraction for Anomaly Detection) or OPTICS (Ordering Points to Identify the Clustering Structure), could provide insights into these results or potentially preferable performance (Ankerst et al., 1999; Sidibé &

Shu, 2017). Unlike the DBSCAN, LR can be performed on a short sequence of AIS points, but requires extensive data preprocessing. The initial analysis into the HMM found that it was sensitive to correctly derived emission probability distribution functions. Therefore, it is recommended that these classification models be implemented with precise observed class data and region specific domain knowledge for parameter optimization and correct model implementation. Moreover, sufficiently high resolution AIS data are required to ensure vessel movement characteristics are representative of wildlife-viewing behaviour. Although these requirements extend to the supervised HMM, the unsupervised HMM removes these dependencies, but has a longer processing time and is typically difficult to statistically validate.

The supervised implementation of the HMM required observed cetacean encounter data to directly derive the HMM model parameters – the initial, transition, and emission probabilities. An unsupervised approach takes advantage of machine learning practices to learn the most likely HMM parameters automatically from data patterns in the AIS SOG sequence. This approach avoids dependence on precise observed class data and region specific domain knowledge allowing its application across a larger geographic region and a variety of commercial ecotourism vessels. This research determined that the unsupervised HMM performance was comparable to the supervised HMM and was simpler to implement, although it was less computationally efficient and may contain some uncertainty demonstrated by a marginally lower log-likelihood score. Regardless, the value of an unsupervised HMM for the classification of wildlife-viewing events is apparent in its generalizable application across the study site examined in this research.

Although classic statistical models have the capacity to classify whale-watching vessel behaviour, this thesis focused on statistical data mining and machine learning approaches. Typically, statistical models utilize a smaller cleaned dataset, and the large volume of AIS data used in this research meant that manual exploration of these data may not capture the underlying patterns in the AIS variables related to wildlife-viewing. This is especially true given the control data was collected on a single whale-watching vessel of a certain size with a set antenna height, which, when used in statistical modelling, may not accurately capture all vessel movement. Therefore, data mining and machine learning approaches were used to discover the unknown trends in the AIS data belonging to other vessels that were not used to collect the observed vessel behaviour classes. Additionally, statistical models are used to assess the relationship between variables and the classification output, but this research was not concerned about this relationship, so a variety of AIS variables were explored through data mining and machine learning approaches with a focus on classification accuracy. Finally, there exists potential that machine learning classification models could continue to learn from a live stream of AIS data. This has the potential to provide more accurate and robust classification as more whale-watching vessels become AIS enabled, or account for changes in vessel movement with the addition of new vessels or changes in wildlife observation regulations.

To demonstrate computational efficiency simplistically, this thesis reported each model's execution time in seconds, but this definition of efficiency is not necessarily conclusive. The demonstrated execution times will vary depending on the hardware and software used to execute these models, programmer knowledge, choices made during model implementation, and each model's complexity. For example, this research was conducted locally on a single computer, but utilizing parallel computing power would lower execution times. Therefore, more in-depth measurements and reporting on how the computational complexity of each model changes based on the hardware and software used could be valuable from a computational standpoint. Regardless, this thesis demonstrates that these models can be implemented locally without the need for high computational power.

An assessment of the species level vessel movement characteristics for the supervised and unsupervised HMM classified wildlife-viewing events found the SOG and encounter times to be significantly distinct between grey whales, humpback whales, and killer whales. These findings support the development of a multiclass HMM that only utilizes AIS data to classify the target species being observed without reliance on precise observed cetacean encounter data (e.g. BC Cetacean Sightings Network data). The development and validation of DBSCAN and LR for the entire commercial whale-watching vessel fleet was not explored in this research, and neither was an assessment of species level vessel movement trends in the DBSCAN and LR classified outputs. The inclusion of such a study could provide additional information of species level vessel movement trends, providing a comparison to the species level trends seen in the HMM classified events, possibly providing information on the HMM advantages and disadvantages, or potentially lead to creating more robust classification models. Moreover, the cetacean spatial distributions derived by these methods show their potential to provide a passive means of cetacean detection to compliment already used cetacean detection methods, which is essential for implementation of effective cetacean conservation efforts (Allen & Singh, 2016; Sveegaard et al., 2015).

4.2 Management Implications

The results of this research demonstrate how AIS can be utilized to passively detect when commercial ecotourism vessels are engaged in wildlife-viewing with a vessel tracking system already equipped on many of these vessels (Canadian Shipping Act, 2019; Washington State Legislature, 2021). The need exists for passive methods to monitor compliance and enforce conservation regulations and guidelines because at sea enforcement of these regulations is resource intensive. Full time monitoring of cetacean encounters across a geographically broad area are difficult because both cetaceans and vessels are highly mobile, and regulations may vary based on species, vessel type (i.e. commercial or recreational whale-watching), or political boundaries (Giles & Koski, 2012; Higham et al., 2008; Mallard, 2019). Regulations and guidelines, such as the minimum distance that boats can approach cetaceans, restrictions on vessel speeds while approaching and observing cetaceans, limits to the number of vessel present at one encounter (Be Whale Wise, 2020; Fisheries and Oceans Canada, 2019, 2020, 2021; Washington State Legislature, 2019), voluntary vessel slowdown (Vancouver Fraser Port Authority, 2020), and Transport Canada sanctuary zones (Fisheries and Oceans Canada, 2019, 2021) exist within the study site to mitigate the potentially harmful impacts vessels can have on cetaceans. By automatically detecting when whale-watching vessels are engaged in wildlife-viewing behaviour with the classification methods developed in this thesis, passive monitoring of whale-watching vessels for regulation compliance can provide assurance to operators and marine managers.

Detecting when commercial whale-watching vessels are observing wildlife affords an opportunity to study vessel movement characteristics while in proximity to these animals, which has the potential to be applied to AIS equipped recreational whale-watching vessels. Often recreational vessels use commercial whale-watching vessels as their guide for locating cetaceans and determining their distance from cetaceans during an encounter (Kessler & Harcourt, 2013). Detecting commercial whale-watching behaviour could provide information to better understand recreational whale-watching vessel behaviour and their compliance to conservation regulations. For example, the application of these methods to AIS equipped recreational whale-watching vessels could supplement vessel-based and land-based observations in quantifying the number of vessels at cetacean encounters. As such, industry specific applications exist with the potential to inform regulators where greater enforcement may be needed and marine managers and tour operators where to focus conservation efforts. Moreover, this information has applications in cetacean conservation research, such as studies on acoustic disturbances because a greater number of vessels translates to increased vessel noise, which can impact cetacean communication and behaviour states (Erbe, 2002; Joy et al., 2019). By detecting which cetacean species is being targeted for observation, these methods have the potential to provide additional cetacean tracking data that can supplement existing studies employing vessel-based, land-based, acoustic, aerial, or satellite cetacean detection data (Gillespie et al., 2020; Laake et al., 1997; Mate et al., 1997; Sveegaard et al., 2015).

In general, the methods developed in this research have the potential to inform studies on the possible impacts whale-watching vessels have on marine mammals and the coastal environment, which can be used by managers in their decision making regarding marine policy. Two demonstrative examples on how the methods can be applied to marine policy and conservation measures were explored in Chapter 3, which evaluated seasonal and annual trends in wildlife-viewing behaviour, and the spatial distributions of wildlife-viewing effort. This cursory assessment found that, although regulatory measures were in place, there still existed wildlife-viewing within the sanctuary zones for some operators, and that the slowdown measures could benefit cetacean populations outside of the southern resident killer whales. These examples show how these methods can be used to assess compliance to current policies and inform future implementation of conservation measures.

4.3 Future Research

There are many possible applications of these methods and potential developmental directions, and as such this research is not inclusive. As these methods can estimate wildlife-viewing effort for a region, there exists an opportunity to supplement cetacean tracking data with this information, but further study is needed into how to best integrate these data. From a developmental standpoint, future studies into an unsupervised multiclass HMM for the detection of probable cetacean species being observed without relying on observed class data are recommended. As cetacean regulations and guidelines vary regionally (e.g. between provinces or states, and between countries), vessel movement characteristics may also vary, such as approach speed and encounter times, which are the basis for the species level assessment. Therefore, although these methods were deployed across a large geographic scale, more research is needed into if these methods can be applied at continental or global scale. For the purpose of regulation compliance, future studies into how to apply these methods in near real-time would benefit enforcement agencies and commercial operators. Specifically, studies into how these methods can integrate a live stream of AIS data could provide operators and managers a platform to ensure that whale-watchers are adequately following regulations.

In summary, this thesis presented novel application of established classification models and approaches for the detection of wildlife-viewing behaviour of commercial whale-watching vessels. This research provides insights into how supervised and unsupervised classification models can be used to better understand vessel movement characteristics while observing cetaceans, which has the potential to be applied to AIS equipped non-commercial whale-watching vessels. Although there exists substantial potential for further development of the current methods, the research presented in this thesis demonstrates that these methods can be a powerful tool for scientists, marine managers, and commercial operators to further conservation and vessel impact research, aid in marine policy decision making, and inform vessel management practices.

4.4 Literature Cited

Ankerst, M., Breunig, M., Kriegel, H., and Sander, J. (1999) OPTICS: ordering points to identify the clustering structure. *ACM SIGMOD Record*, 28:2, 49–60, 1999.

Allen, A. M., & Singh, N. J. (2016). Linking movement ecology with wildlife management and conservation. *Frontiers in Ecology and Evolution*, 3:155, 1–13. https://doi.org/10.3389/fevo.2015.00155

Be Whale Wise (2020). Marine Wildlife Laws and Guidelines for Boaters, Paddlers and Viewers. https://www.bewhalewise.org/wp-content/uploads/2020/07/Be-Whale-Wise-Brochure-2020-v5.pdf

Canadian Shipping Act (2019). Regulations Amending the Navigation Safety Regulations (Automatic Identification Systems) (SOR/2019-100). Canada Gazette Part II, 153(9). https://gazette.gc.ca/rp-pr/p2/2019/2019-05-01/html/sor-dors100-eng.html

Conn, P. B., & Silber, G. K. (2013). Vessel speed restrictions reduce risk of collision-related mortality for North Atlantic right whales. *Ecosphere*, 4(4) :43. https://doi.org/10.1890/ES13-00004.1

Erbe, C. (2002). Underwater noise of whale-watching boats and potential effects on killer whales (Orcinus orca), based on an acoustic impact model. *Marine Mammal Science*, 18(2), 394–418.

Findley, L. T., & Vidal, O. (2002). Gray whale (*Eschrichtius robustus*) at calving sites in the Gulf of California, México. *Journal of Cetacean Research and Management*, 4(1), 27–40.

Fisheries and Oceans Canada. (2019). Government of Canada outlines its 2019 plan for protecting Southern Resident killer whales. https://www.canada.ca/en/fisheries-oceans/news/2019/05/government-of-canada-outlines-its-2019-plan-for-protecting-southern-resident-killer-whales.html

Fisheries and Oceans Canada. (2020). Watching marine wildlife. https://www.dfo-mpo.gc.ca/species-especes/mammals-mammiferes/watching-observation/index-eng.html

Fisheries and Oceans Canada. (2021). 2021 management measures to protect Southern Resident killer whales. https://www.pac.dfo-mpo.gc.ca/fm-gp/mammals-mammiferes/whales-baleines/srkw-measures-mesures-ers-eng.html

Giles, D. A., & Koski, K. L. (2012). Managing vessel-based killer whale watching: A critical assessment of the evolution from voluntary guidelines to regulations in the Salish Sea. *Journal of International Wildlife Law and Policy*, 15(2), 125–151. https://doi.org/10.1080/13880292.2012.678792

Gillespie, D., Palmer, L., Macaulay, J., Sparling, C., & Hastie, G. (2020). Passive acoustic methods for tracking the 3D movements of small cetaceans around marine structures. *PLoS ONE*, 15(5), 1–16. https://doi.org/10.1371/journal.pone.022905

Higham, J. E. S., Bejder, L., & Lusseau, D. (2008). An integrated and adaptive management model to address the long-term sustainability of tourist interactions with cetaceans. *Environmental Conservation*, 35(4), 294–302. https://doi.org/10.1017/S0376892908005249

International Maritime Organization (IMO). (2003). Guidelines for the Installation of a Shipborne Automatic Identification System (AIS) (Ref. T2/8.02)

Joy, R., Tollit, D., Wood, J., MacGillivray, A., Li, Z., Trounce, K., & Robinson, O. (2019). Potential benefits of vessel slowdowns on endangered southern resident killer whales. *Frontiers in Marine Science*, 6(344), 1–20. https://doi.org/10.3389/fmars.2019.00344

Kessler, M., & Harcourt, R. (2013). Whale watching regulation compliance trends and the implications for management off Sydney, Australia. *Marine Policy*, 42, 14–19. https://doi.org/10.1016/j.marpol.2013.01.016

Laake, J. L., Calambokidis, J., Osmek, S. D., & Rugh, D. J. (1997). Probability of Detecting Harbor Porpoise from Aerial Surveys: Estimating g(0). *The Journal of Wildlife Management*, 61(1), 63–75.

Lusseau, D. (2006). The short-term behavioral reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Marine Mammal Science*, 22(4), 802–818. https://doi.org/10.1111/j.1748-7692.2006.00052.x

Mallard, G. (2019). Regulating whale watching: A common agency analysis. *Annals of Tourism Research*, 76, 191–199. https://doi.org/10.1016/j.annals.2019.04.011

Mate, B. R., Nieukirk, S. L., & Kraus, S. D. (1997). Satellite-monitored movements of the northern right whale. *Journal of Wildlife Management*, 61(4), 1393–1405.

Schuler, A. R., Piwetz, S., Di Clemente, J., Steckler, D., Mueter, F., & Pearson, H. C. (2019). Humpback Whale Movements and Behavior in Response to Whale-Watching Vessels in Juneau, AK. *Frontiers in Marine Science*, 6:710, 1–13. https://doi.org/10.3389/fmars.2019.00710

Sidibé, A., & Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. Journal of Navigation, 70(4), 847–858. https://doi.org/10.1017/S0373463317000066

Sveegaard, S., Galatius, A., Dietz, R., Kyhn, L., Koblitz, J. C., Amundin, M., Nabe-Nielsen, J., Sinding, M. H. S., Andersen, L. W., & Teilmann, J. (2015). Defining management units for cetaceans by combining genetics, morphology, acoustics and satellite tracking. *Global Ecology and Conservation*, 3, 839–850. https://doi.org/10.1016/j.gecco.2015.04.002

Vancouver Fraser Port Authority. (2020). ECHO Program 2019 voluntary vessel slowdown trial in Haro Strait and Boundary Pass: Summary findings.

Washington State Legislature (2019). Protection of southern resident orca whales—Unlawful activities—Penalty. https://apps.leg.wa.gov/rcw/default.aspx?cite=77.15.740

Washington State Legislature (2021). Commercial whale watching compliance and reporting. https://apps.leg.wa.gov/wac/default.aspx?cite=220-460-140

Williams, R., Trites, A. W., & Bain, D. E. (2002). Behavioural responses of killer whales (*Orcinus orca*) to whale-watching boats: opportunistic observations and experimental approaches. *Journal of Zoology*, 256, 255–270. https://doi.org/10.1017/s0952836902000298

Appendix

A.1 Statistical Tests of Agreement – Kappa and McNemar Statistics

Kappa Statistic: Tests the null hypothesis that the modelled classes were randomly chosen.

$$K = \frac{p_o - p_e}{1 - p_e} \tag{A.1}$$

where, p_o is the observed probability of agreement, and p_e is the expected probability given by chance, or at random (Cohen, 1960). A $K > 0$ signifies agreement between modelled and observed classes, with $K > 0.8$ representing strong agreement, and $K \leq 0$ signifies a random assignment of classes (Viera & Garrett, 2005).

McNemar Statistic: Tests the null hypothesis that the two classifiers have the same error rate (Dietterich, 1998). Consider that *A* and *B* represent two different classifiers, a confusion matrix of the misclassified AIS points between *A* and *B* is used to derive the McNemar statistic:

$$M_{A,B} = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \tag{A.2}$$

where n_{01} represents the number of records misclassified by *A* and not *B*, and n_{10} represents the number of records misclassified by *B* but not *A*.

A.2 Literature Cited

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, XX(1), 37–46.

Dietterich, T. G. (1998). Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. In *Neural Computation*, 10, 1895–1923). https://doi.org/10.1007/978-3-319-50926-6_6

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 37(5), 360–363. http://www1.cs.columbia.edu/~julia/courses/CS6998/Interrater_agreement.Kappa_statistic.pdf