Evaluating whale and vessel detection methods in and around the Saturna Island Interim Sanctuary Zone (ISZ).

by
Mikayla C. R. Young

Bachelor of Science (Honours) in Biology and Environmental & Resource Sciences – Trent University, 2020

Project Submitted in Partial Fulfilment of the
Requirements for the Degree of
Master of Science

in the
Ecological Restoration Program

Faculty of Environment (SFU)
and
School of Construction and the Environment (BCIT)

SIMON FRASER UNIVERSITY
BRITISH COLUMBIA INSTITUTE OF TECHNOLOGY
2024

Declaration of Committee

Name:	**Mikayla Young**
Degree:	**Master of Science**
Title:	**Evaluating whale and vessel detection methods in and around the Saturna Island Interim Sanctuary Zone (ISZ).**
Examining Committee:	**Chair: Dr. Ruth Joy** Supervisor Faculty, SFU
	Dr. Ana Chara-Serna Examiner Faculty, BCIT

Abstract

This research compares whale and marine vessel detection methods through performance metrics adapted from machine-learning models. Monitoring whale habitat use and vessel infractions in exclusion zones can inform adaptive management for whale recovery efforts. Land-based cetacean observation (LBCO) surveys and dedicated vessel surveys (DVS) were conducted during the summer of 2023 and are considered the gold standard methods for this study. Data collected for comparison from alternative detection methods include a citizen science network, thermal imaging, acoustic, radar, and automatic identification systems (AIS). The citizen science network was the most reliable method for whale detection of all species observed. Vessel detection methods demonstrated similar overall detection reliability, as radar consistently had higher recall values while AIS consistently had higher precision values. Differing scenarios where human observation is unlikely to be the gold standard are discussed and are recommended as a topic for continued research.

Keywords: killer whales; humpback whales; marine vessels; detection methods; performance metrics; Salish Sea

Dedication

To my niece Paisley and my parents, Rosalie and Mike, for your constant love and support.

To the whales for giving me a glimpse of their world and with the hope that I can help make it a better place.

Acknowledgements

I respectfully acknowledge that my research took place within the overlapping traditional territories of the original stewards of the lands and waters of Saturna Island, including the W̱SÁNEĆ First Nations, the bands of the Hul’quimi’num Treaty Group, and the Tsawwassen, Semiahmoo and Stz’uminis First Nations.

I have been very fortunate to have much support for my project and many people to thank. I want to express my gratitude to my project partners on Saturna Island, SIMRES and the SGIWSN; thank you for giving me a once-in-a-lifetime experience and making Saturna feel like home. I thank Maureen Welton, Susie Washington-Smyth, and Pam Janszen for opening their homes and providing me with a place to stay. Kathleen Durant, thank you for your hospitality during my visit to the Pender Island ISZ. David Dick, thank you for taking the time to meet with me, providing advice on my project, and for the critical work of your team, the QENTOL, YEN W̱SÁNEĆ Marine Guardians.

This project was only possible with the data, equipment, and funding provided to me. Thank you, Harald Yurk from DFO and Sebastian Richter from Biophysics Group, for providing me with thermal imaging and radar data for my analysis and your willingness to answer my endless questions. Lucy Quayle from DFO, Spyhopper, and Quayle Consulting Ltd., thank you for supplying me with citizen science and AIS data and for your constant and valuable guidance. Sheila Thornton and her team at DFO, thank you for loaning me the theodolite for the summer. Fabio Soares Frazao, thank you for running your killer whale detection model on the acoustic data and allowing me to incorporate it into my study. I am additionally grateful to the Port of Vancouver’s ECHO program and a MITACS NSERC Alliance Grant for funding my research.

I had the most incredible team to assist with data collection in the field, and they have all become my dear friends. A huge thank you to my research partner, Rachel Fairfield Checko, who, despite having her own master’s project, devoted a significant amount of time to helping me collect data for mine. I also want to thank last year’s SFU researchers, Lauren Laturnus and Olivia Murphy, who trained us and came to fill in whenever we needed help. Thanks to you all, the amount of time giggling far surpassed the tears. I want to express my gratitude to Samantha Broadley, Janine McNeilly, the rest of Dr. Joy’s lab “R-pod,” former SFU whale researchers Azadeh Gheibi and Kaitlin Baril, and my MER classmates for their ongoing support and feedback as I progressed through my project. To my furry friends Koda, Tigger, and Dodger, thank you for the much-needed emotional support during all stages of my project.

Last but certainly not least, I want to thank my supervisor, Dr. Ruth Joy. Words cannot truly convey my gratitude for your guidance and the opportunities you have given me. You were always there to support me, even if it meant helping with my R code after midnight on a weekend. None of this would have been possible without you, and I feel so fortunate to have worked with you. Thank you, Ruth.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AIS	Automatic Identification System
BC	British Columbia
BCIT	British Columbia Institute of Technology
BKW	Bigg's killer whale
DANN	Deep Artificial Neural Network
DFO	Fisheries and Oceans Canada
DVS	Dedicated Vessel Survey
ECHO	Enhancing Cetacean and Habitat Observation
FN	False Negative
FP	False Positive
HALLO	Humans and Algorithms Listening for Orcas project
IR	Infrared
ISZ	Interim Sanctuary Zone
LBCO	Land-Based Cetacean Observation
MCC	Matthew's Correlation Coefficient
NPV	Negative Predictive Value
NRKW	Northern Resident killer whale
PDT	Pacific Daylight Time
SARA	Species at Risk Act (Canadian)
SD	Standard Deviation
SFU	Simon Fraser University
SGIWSN	Southern Gulf Islands Whale Sighting Network
SIMRES	Saturna Island Marine Research and Education Society
SRKW	Southern Resident killer whale
TN	True Negative
TP	True Positive
USA	United States of America

A one-year-old Southern Resident killer whale, J59 Sxwyeqólh, breaches out of the water ahead of her mother, J37 Hy’shqa. Photo credit: Rachel Fairfield Checko.

Humpback whale BCX1658 Bullet in Boundary Pass. Photo credit: Rachel Fairfield Checko.

Chapter 1. Introduction

The Salish Sea is a diverse and vital ecosystem for many cetaceans, an order of aquatic mammals including whales, dolphins, and porpoises (Gaydos & Pearson 2011). It is also one of the world's busiest shipping routes. The number of vessels and the amount of time vessels spend on the water in the Salish Sea has been increasing in recent years (Gillespie 2016; McWhinnie et al. 2021; Seely et al. 2017). This study focuses on the waters surrounding Saturna Island in the Salish Sea, including Boundary Pass and the Strait of Georgia which are busy sections of the shipping route and for recreational boaters. Two at-risk cetaceans frequently live, forage, and travel in this area. These include killer whales (*Orcinus orca*), also called orcas, and humpback whales (*Megaptera novaeangliae*).

Killer whales are toothed whales, which are part of the dolphin family but are more commonly called whales. There are two ecotypes of killer whales that are often found in the Salish Sea: Resident killer whales and Bigg's (also known as transient) killer whales. These different killer whale ecotypes are considered the same species but are socially isolated and exhibit differences in morphology, genetics, behaviour, and dietary preferences (Ford et al. 1998, 2000; Barrett-Lennard & Ellis 2001). Due to their distinct life histories, a recent study reviewed data on these differences and suggested that the taxonomy of Bigg's and residents should be revised to be different species rather than different ecotypes (Morin et al. 2024). Morin et al. (2024) suggest that Bigg's should be recognized as *Orcinus rectipinnus* and residents be recognized as *Orcinus ater* if they become recognized as distinct species.

Resident killer whales are an ecotype of killer whales that can be further delineated into two genetically separate clans, the Southern Resident killer whale (SRKW) and the Northern Resident killer whale (NRKW) (Barrett-Lennard & Ellis 2001). The range of NRKWs extends geographically from central Vancouver Island to Alaska and does not overlap with the area of interest for this study (Krahn et al. 2002). The summer core geographic range of SRKWs is well documented and spans from Washington State to central Vancouver Island, including the area of interest (Krahn et al. 2002). Their range in winter is not as well understood; however, between 2018 and 2022, SRKW have been sighted at least one time in each calendar month across all years but have not been sighted in all months within each individual year (Krahn et al. 2002; Shields 2023). This shows an increase in SRKW presence in fall and winter from previous trends (Krahn et al. 2002; Shields 2023). SRKWs are listed as an endangered killer whale population and are federally protected under the Canadian *Species at Risk Act* (*SARA*) (COSEWIC 2008). The population consisted of 98 individuals in 1995 before experiencing a severe decline to 75 individuals as of July 2023 (Center for Whale Research 2023). They live and travel together in 3 matriarchal groups called pods named J, K, and L (Ford 1991). Due to their ecological, economic, and cultural significance, SRKWs are a focus for cetacean conservation efforts. Many of the whale recovery initiatives in the Salish Sea are designed with them in mind.

The Bigg's killer whales (BKW) are listed under *SARA* as Threatened (COSEWIC 2008). The population and presence of BKWs in the Salish Sea have been increasing in recent years (Shields et al. 2018). It is thought this is mainly due to the increases in pinnipeds which are their primary prey source (Roman et al. 2013; Shields et al. 2018). It is believed that the populations of pinnipeds commonly preyed upon by BKWs such as harbour seals (*Phoca vitulina*), Steller Sea Lions (*Eumetopias jubatus*) and California Sea Lions (*Zalophus californianus*) saw a tremendous increase in the Salish Sea due to the implementation of the US *Marine Mammal Protection Act* in 1972 (Roman et al. 2013; Shields et al. 2018). The Canadian government also incorporated marine mammal regulations into the federal *Fisheries Act* in 1993. These acts introduced new laws and restrictions against hunting marine mammals, enabling many populations impacted by human harvest, like pinnipeds, to recover (Roman et al. 2013; Shields et al. 2018). Despite this recent increase, BKWs still faces severe threats, and the population is not yet stable (COSEWIC 2008; Shields et al. 2018).

Humpback whales that transit through the Salish Sea are part of the North Pacific humpback whale population (*Megaptera novaeangliae kuzira*). This population has been listed as a Species of Special Concern under *SARA* since 2013, and its latest conservation status assessment occurred in 2022 (COSEWIC 2022). For hundreds of years, excessive exploitation by the whaling industry threatened humpback whale populations around the world (Clapham 2018). Global commercial whaling of humpbacks was halted in 1966, leading many populations to begin to recover (Clapham 2018). While the North Pacific humpback whale population has increased in recent years, it is not yet stable (COSEWIC 2022).

Vessel traffic in the Salish Sea has negative impacts on humpback whales and both ecotypes of killer whales. As a result, vessels have been identified as a critical threat to cetacean survival in these waters due to vessel-related noise pollution, physical disturbance, emissions, exhaust, and oil spills (COSEWIC 2008; Di Clemente et al. 2018; Harwood et al. 2016; McWhinnie et al. 2021). In 2019, the Government of Canada, the Vancouver Fraser Port Authority and other partners signed the *Species at Risk Act* section 11 conservation agreement to support the recovery of the Southern Resident Killer Whale (Government of Canada 2019a). This agreement resulted in implementing new SRKW population management measures, many of which also benefit BKWs and humpback whales (Government of Canada 2019b). Measures to reduce vessel noise and disturbance include interim sanctuary zones (ISZ), interim speed-restricted zones and increased vessel approach distances for whales (Government of Canada 2019b). With some minor adjustments since 2019, these measures are still in place as of 2023 (Government of Canada 2023). Additionally, in 2017, the Vancouver Port Authority Enhanced Cetacean Habitat and Observation (ECHO) program began a voluntary commercial vessel slowdown initiative for noise reduction, with zones in Boundary Pass, Haro Strait, and Swiftsure Bank (Le Baron et al. 2019).

ISZs are designated short-term marine protected areas where most vessels are prohibited from transiting and fishing (Government of Canada 2023). There are currently two ISZs in the Salish Sea. Both are in the Southern Gulf Islands archipelago and are in effect from June 1 to November 30. One is off the coast of Saturna Island, and is the area of interest for this research, and the second is off the coast of Pender Island (Government of Canada, 2023). These ISZs were established to decrease vessels' acoustic and physical impacts on SRKWs (Government of Canada 2023). The government chose the locations of the ISZs based on their historical importance as crucial foraging areas for SRKWs (Government of Canada 2023). These areas are often frequented by humpbacks and BKWs as well (Quayle 2021).

Vessel exclusion zones can provide safe habitat from vessel strikes. Baleen whales are more likely to be struck by vessels due to their large size and behaviour (Department of Fisheries and Oceans Canada 2017). No current evidence suggests they can echolocate, which means they are also less likely to be aware of the presence of nearby vessels (Frazer & Mercado 2000; COSEWIC 2022). Humpbacks are the cetaceans that are most frequently struck by ships on Canada's Pacific coast, and strikes are likely to increase as vessels get larger, faster, and more numerous (COSEWIC 2022). Vessel strikes have been identified as a critical threat to humpback whale population recovery (Department of Fisheries and Oceans Canada 2017).

As cetaceans frequent the Saturna ISZ, a designated vessel exclusion zone, numerous methods of detecting vessel and whale presence are deployed. For whales, this includes a local volunteer citizen science sighting network called the Southern Gulf Islands Whale Sighting Network (SGIWSN), an infrared temperature sensor for thermal detection, and a hydrophone for acoustic detection. There is a radar device and an antenna automatic identification system (AIS) receiver for vessel detection. The SGIWSN and researchers partnered with the Saturna Island Marine Research and Education Society (SIMRES) also report vessel infractions in the ISZ.

1.1. Goals and Objectives

Adaptive management plays a vital role in ecological restoration efforts, especially in environments with high levels of uncertainty, such as marine ecosystems (Palmer et al. 2016; Wintle 2007; Payne et al. 2016). For adaptive management to succeed, the effectiveness of restoration strategies must be continuously monitored to allow for adjustments that can enhance their impact (Wintle 2007). Continuous monitoring of whales helps to understand how they move within potentially suitable habitats such as the Saturna ISZ and how much time they spend in them. Recent studies suggest that SRKWs may be shifting their temporal patterns in the Salish Sea away from historical trends (Shields 2023). An analysis of their movement patterns in the Salish Sea between 2018 and 2022 suggests that overall, SRKWs are spending less time in the Salish Sea, especially during the spring and summer when compared to their traditional movement patterns derived from historical data (Shields 2023). However, SRKWs have spent more time in the Salish Sea during the fall and winter in recent years (Shields 2023). These shifts correlate with declines in the availability of Chinook salmon (*Oncorhynchus tshawytscha*) from the Fraser River in the summer and an increase in the availability of chum salmon (*Oncorhynchus keta*) in Puget Sound in the fall and winter (Shields 2023). Although there are too many uncertainties to know if these shifts will become long-term, it is essential to consider both long and short-term patterns when implementing policies intended to assist in the long-term recovery of the SRKW population (Murphy et al. 2023; Shields 2023).

During the two decades coinciding with the decline in SRKWs and Chinook populations, vessel activity in the Salish Sea has increased (McWhinnie et al. 2021; Gillespie 2016). These confounding factors, a lack of prey availability, and vessel disturbances, are making it increasingly difficult for SRKWs to forage and may have contributed to the shift in their temporal patterns and reduced presence in the Salish Sea (Joy et al. 2019; McWhinnie et al. 2021). SRKWs are known to change their behaviour in the presence of vessels and will spend significantly less time foraging when a vessel is within 400 m (Joy et al. 2019). These behaviour alterations raise the question of whether effective management of acoustic and physical disturbance from vessels would help restore traditional SRKW foraging habitat, enabling them to resume effective forage behaviours in these waters. Proper enforcement and monitoring are critical for whale habitat restoration efforts, such as the ISZ, to be effective.

The Saturna Island ISZ is one of the few locations in the Salish Sea where many different human and automatic technological methods for monitoring whale presence and vessel compliance are deployed in close proximity. As technology and artificial intelligence can allow for continuous monitoring and be more cost-effective in the long term, many researchers wonder if human presence is still required for monitoring or if we can leave these tasks to technology and artificial intelligence. This study aims to evaluate the performance of the various methods for monitoring vessels and detecting the presence of whales in and around the Saturna Island ISZ. Additionally, the discussion provides specific recommendations to inform adaptive management and monitoring to enhance the protection of marine mammal species. This research intends to do so through the following goals and objectives:

Goal #1: Evaluate the performance of various methods and technologies for monitoring whale presence.

Objective #1.1: Compare and contrast the performance of each monitoring method for whales in Boundary Pass using machine learning performance metrics.

Objective #1.2: Discuss the limitations of each whale detection method and the potential for automatic or real-time reporting.

Goal #2: Evaluate the performance of various methods and technologies for monitoring vessel adherence to the ISZ regulations.

Objective #2.1: Compare and contrast the performance of each monitoring method for vessels using machine learning performance metrics.

Objective #2.2: Discuss the limitations of each monitoring method and the potential for automatic or real-time reporting.

Goal #3: Provide recommendations for using the various methods and technologies when monitoring for adaptive management.

Objective #3.1: Distinguish the value of human observation vs artificial intelligence.

Objective #3.2: Discuss the potential value of using the methods in other places or for different whale recovery management measures, such as speed-restricted and voluntary slowdown zones.

Chapter 2. Methods

2.1. Site Description

Saturna Island is one of the Southern Gulf Islands of British Columbia, Canada, situated in the Salish Sea between Vancouver Island and the mainland of British Columbia and the United States of America (Figure 1). It is within the overlapping traditional territories of the W̱SÁNEĆ First Nations, the bands of the Hul'quimi'num Treaty Group, the Tsawwassen First Nation Treaty, and the Semiahmoo and Stz'uminis First Nations.

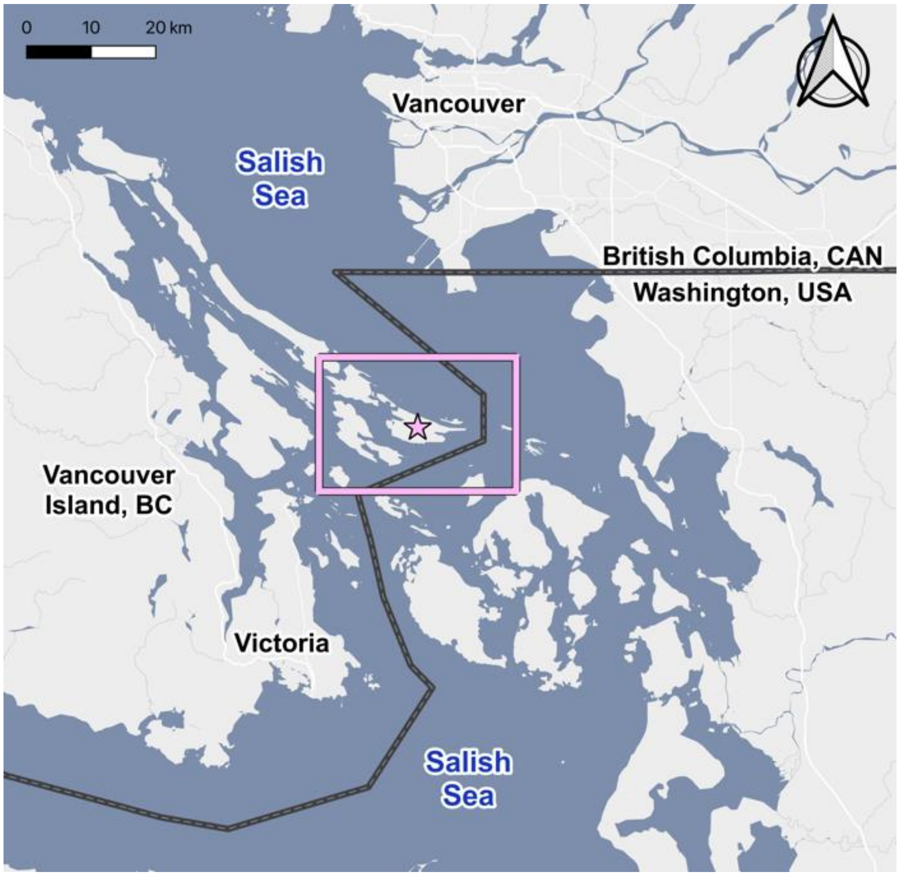


Figure 1. Map of the Southern Salish Sea, with the area of interest outlined in a pink box and a pink star identifying Saturna Island. The Salish Sea lies between British Columbia (BC) in Canada and Washington State in the United States of America (USA).

Boundary Pass borders Saturna Island's southern coast and is a crucial passage linking the Strait of Georgia with Haro Strait (Figure 2). The waters surrounding Saturna Island are within the designated critical habitat of SRKWs under the SARA and the *Endangered Species Act.*

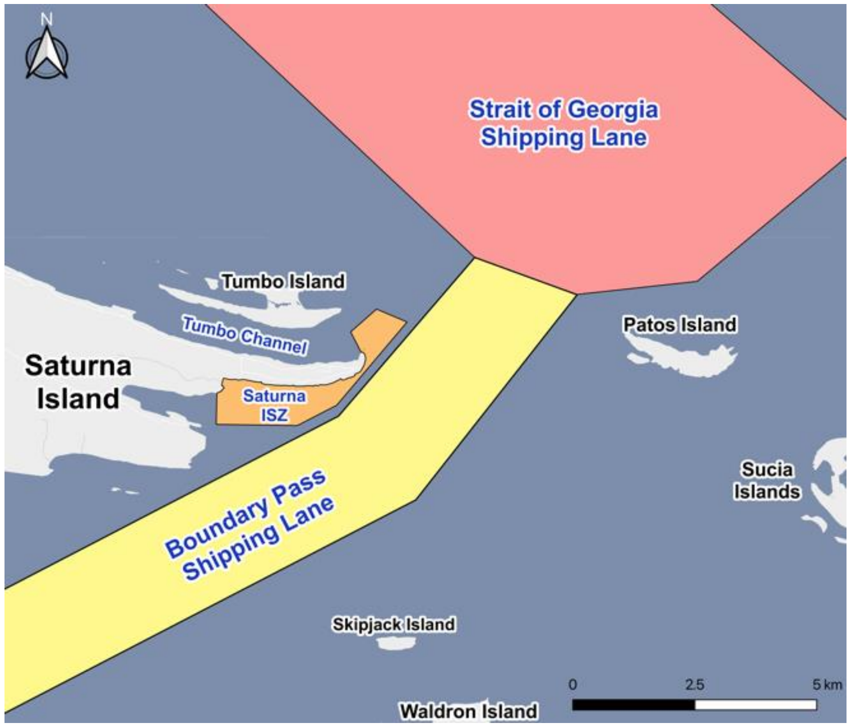


Figure 2. **Map of important features in the area of interest, including the Saturna Island Interim Sanctuary Zone (ISZ) in orange, the Boundary Pass and Strait of Georgia shipping lanes in yellow and pink, respectively, and Tumbo Channel and the islands surrounding Saturna Island each labelled.**

2.2. Evaluating Detection Methods

Machine learning performance metrics were used to assess the performance of various whale and vessel detection methods on Saturna Island. Detection methods were identified as either a gold standard or an alternative method. The gold standard is considered the most reliable method (Hripcsak & Rothschild 2005). In many cases, no gold standard method is perfect, but it should be derived by experts and repeatable (Hripcsak & Rothschild 2005). The detections made by the alternative methods are then compared to the gold standard method to assess their performance.

For this study, the gold standard method for whale detection was the dedicated land-based cetacean observation (LBCO) surveys, and the gold standard method for vessel detection was the dedicated vessel surveys (DVS). The LBCO surveys are based on methods by Lusseau et al. (2009), which was accepted in the peer-reviewed journal *Endangered Species Research*. The DVS is based on methods by Le Baron et al. (2019), a technical report for the Vancouver Port Authority's ECHO program and has been repeated in studies such as Baril (2022). The alternative methods to be evaluated for whale detection capabilities are opportunistic reporting by citizen scientists of the SGIWSN, an acoustic hydrophone killer whale detection model and thermal imaging. A radar device and AIS are the alternative methods to be assessed for vessel detection capabilities. Figure 3 is a visual representation of the hierarchies of the methods of detection as described in this section, and map of the sites of the methods with fixed deployment locations is shown in Figure 4.

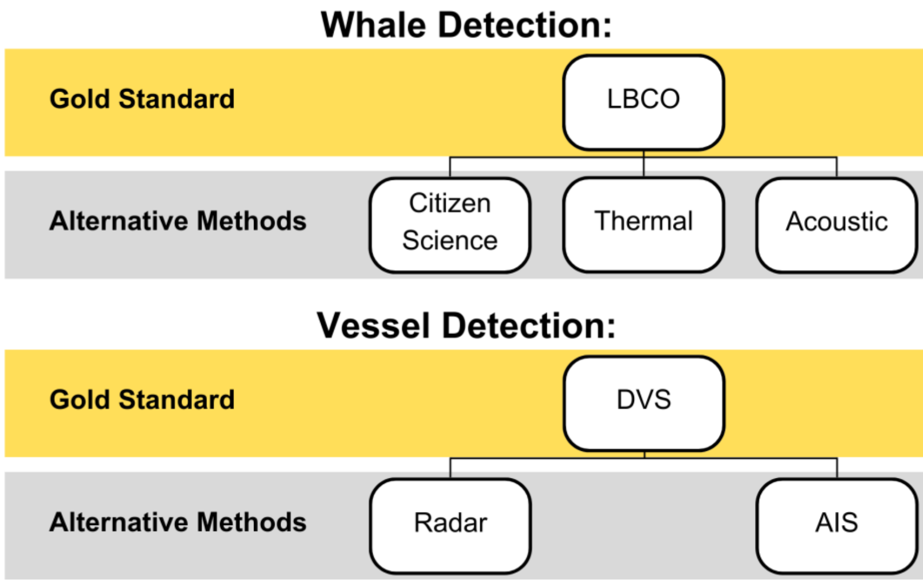


Figure 3. Hierarchies for the methods of whale and vessel detection used in this study.

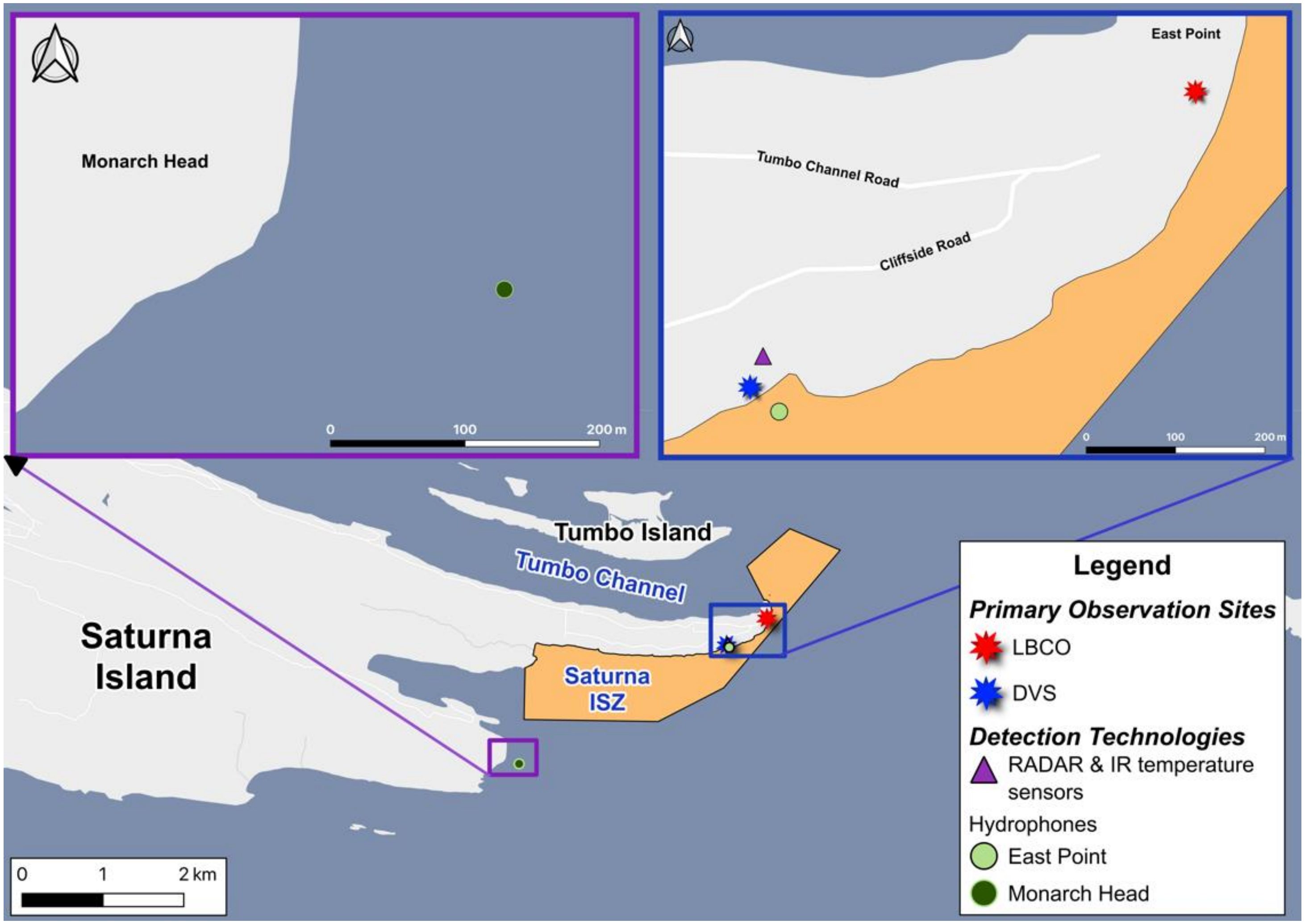


Figure 4. **Map of the locations of the primary observation sites for the land-based cetacean observation (LBCO) surveys and the dedicated vessel surveys (DVS), and the locations for the alternative detection methods with fixed deployments, including the infrared (IR) temperature sensors, radar, and hydrophones.**

Confusion matrices were created to summarize the quality and quantity of detections by an alternative method compared to the gold standard. The matrix is filled out by determining the number of whale or vessel detections that were true positives (TP), false positives (FP), false negatives (FN), and for vessels, true negatives (TN), through comparison to an identified gold standard method (Hripcsak & Rothschild 2005). An outline of a confusion matrix is shown in Table 1. These values are later used in the performance metric calculations described in Section 2.2.1. Individual confusion matrices for each method of detection are included in Appendix B.

Table 1. Outline of a confusion matrix used to summarize the detections of an alternative method compared to the gold standard. Values within the matrix are used for performance metric calculations.

	Whale/vessel detected by the gold standard method	**Whale/vessel not detected by the gold standard method**
Whale/vessel detected by the alternative method	True positive (TP)	False Positive[1] (FP)
Whale/Vessel not detected by the alternative method	False negative (FN)	True negative[2] (TN)

[1] A false positive may be deemed an external positive for some whale detections, as explained further in section 2.4.1.
[2] True negatives are not determined for whale detections, as explained further in this section.

When a target object (whale or vessel) is detected by the gold standard and the detection method under consideration, it is considered a true positive. If a detection method misses a target object detected by the gold standard method, this is regarded as a false negative. When a target object is detected by the detection method but not the gold standard method, this is considered a false positive. This study did not define a true negative whale event as the field of view for each method varies greatly and specific time increments where the recorded presence or absence of a whale was not reported by all methods (i.e., reporting presence or absence at 5-minute intervals). Since the field of view for the methods of vessel detection had more overlap, and vessel presence was recorded in the same time increments, true negatives were defined for vessels. A vessel was considered a true negative when there was no detection at the 5-minute interval by both the alternative method and the gold standard.

2.2.1. Performance Metric Calculations

All machine learning performance metric calculations and figures were conducted in R Studio. Precision and recall were calculated to measure the accuracy and completeness of detections for all detection methods. Recall indicates the ability of the method to detect all target objects, while precision indicates the proportion of positive detections that could be confirmed as true positives. These values were then used to calculate the F1 score, which evaluates the method's reliability for detecting the target object with equal weight given to precision and recall. All these calculations provided a value between 0 and 1, with 0 indicating poor performance and 1 indicating perfect performance.

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

True negatives were only counted for vessels; therefore, negative predictive value (NPV) and Matthew's correlation coefficient (MCC) were calculated only for methods of vessel detection. NPV indicates the proportion of reported negatives that are true negatives based on detections by the gold standard method. NPV also provided a value between 0 and 1, with 0 indicating poor performance and 1 indicating perfect performance. MCC is a statistical tool that considers all entities in the confusion matrix to gauge the agreement between the alternative and gold standard methods. It is a favourable method for evaluating binary classifications where true negatives can be identified, as it considers all entities in the confusion matrix in its calculation (Chicco et al. 2021). It provides a value between -1 and 1, with -1 indicating total disagreement, 0 indicating agreement is no better than random chance, and 1 indicating perfect agreement. A two-sided t-test was performed in R studio on each MCC value to determine significance. The MCC score was significant if the resulting p-value was ≤0.05.

$$NPV = \frac{TN}{TN + FN}$$

$$MCC = \frac{TN * TP - FN * FP}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

2.3. Gold Standard Methods of Detection

The following subsections explain the methods used for the gold standard detection methods. The author and other experienced researchers from Simon Fraser University (SFU) conducted these surveys.

2.3.1. Land-Based Cetacean Observation (LBCO) Surveys

The primary observation site for the dedicated LBCO surveys was at a constant vantage point (48.783722º N, -123.045167º W) in the Gulf Islands National Park Reserve at East Point on Saturna Island (Figure 5).This location was chosen for its optimal visibility of the Saturna Island ISZ and the commercial voluntary slowdown zone in the Boundary Pass shipping lane. A level wooden platform was built at this site to facilitate the use of a theodolite.

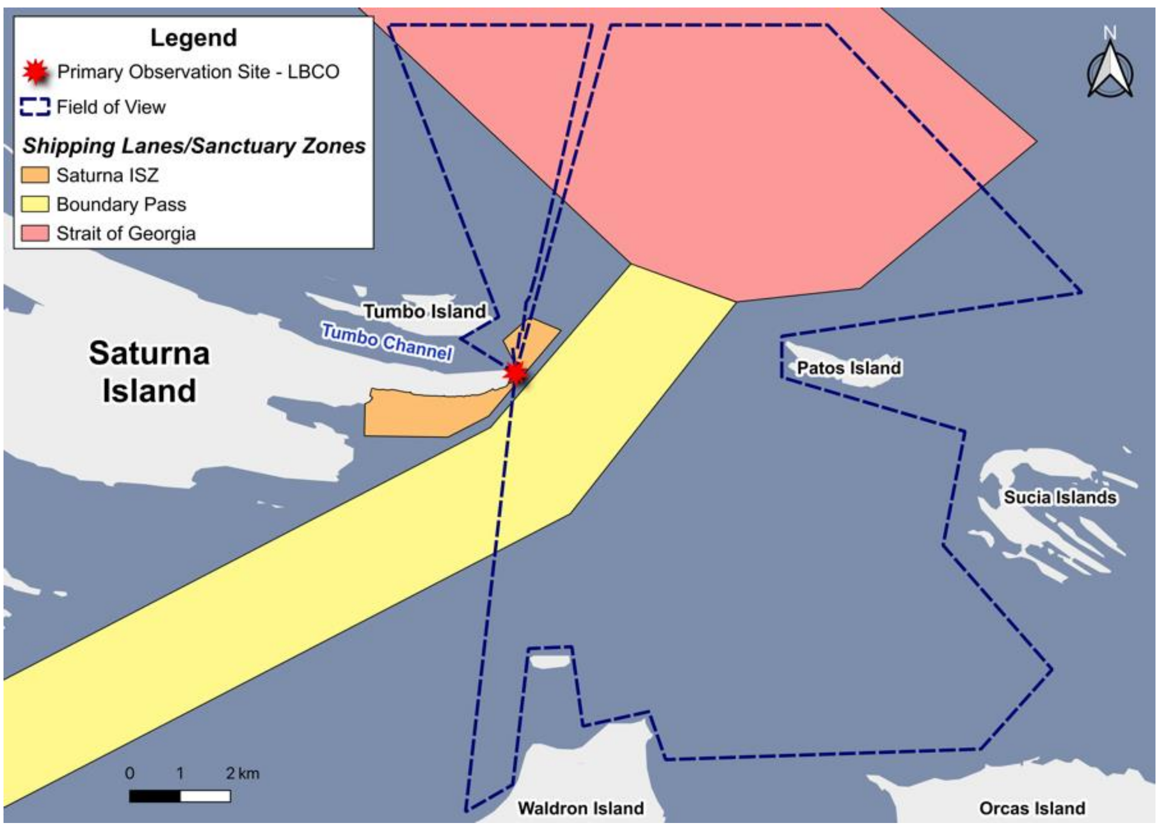


Figure 5. Map demonstrating the primary observation site and the approximate field of view for the dedicated land-based cetacean observation (LBCO) surveys.

LBCO surveys were conducted on 76 dates between June 1 and August 31, 2023, over a total of 494.25 hours. The duration of each survey was a minimum of 3 hours, a maximum of 10.5 hours, and an average of 6.5 hours (*SD* = 1.6). A visual scan with binoculars (Nikon 10 x 42, Zeiss 10 x 42) in a northeast-to-southeast-to-northeast sequence area was done every 15 minutes during survey hours. These methods originated from Lusseau et al. (2009) in Haro Strait and were adapted for Boundary Pass for surveys by Le Baron et al. (2019), Quayle (2021), Gheibi (2022), and Murphy (2023). The survey was conducted at a consistent vantage point at the wooden theodolite platform at the primary observation site. Constant auditory monitoring, where the observer listens for any splash or exhale, was also used to detect whales.

Whales were tracked using a TOPCON DT-200 theodolite (Figure 6) on a tripod and connected to a Panasonic ToughPad with Mysticetus software that can track marine mammals and vessels from land (Mysticetus LLC, 2019). The level wooden platform has permanent points carved into the deck to ensure the three feet of the tripod are placed in the same location each day. Once the tripod was in place, the theodolite was screwed securely onto the mount and levelled, with the plumb bob line hanging from the center of the mount and centred over the mark on the platform that indicates the geo-referenced theodolite location. Coordinates for a visible horizontal reference location across Boundary Pass on Patos Island and a visible vertical reference across Tumbo Channel on Tumbo Island were previously obtained using a handheld GPS device and Google Earth. The theodolite is then calibrated to these coordinates using Mysticetus before beginning each survey. The vertical reference is recalibrated every 15 minutes to account for changes in sea level, and the horizontal reference is recalibrated every hour to account for accidental unnoticed bumps or movements that could offset the calibration. Weather conditions were recorded on Mysticetus and updated whenever there was a change in weather.

Figure 6. Primary observers using the theodolite and mirrorless camera to track a whale event detected in the LBCO survey on June 3, 2023. Photo credit: Lauren Laturnus.

When a whale was detected, this indicated the start of a “whale event.” The start time was recorded, and the primary observer measured the distance from the vantage point and position of the whale with the theodolite by clicking Alt+ on the keyboard, with the crosshairs in the viewfinder of the theodolite positioned at the water line where a whale surfaced. Consecutive points of the same whale could be recorded by clicking

Shit-Alt-+, which would form a track of points connected by a line. A whale event continued until the whale travelled out of view or had been undetected for more than 20 minutes, and the end time was recorded. The species, start and end time, direction of travel, activity, group size, group configuration, and vessel presence were recorded during the event.

A mirrorless camera (Sony α7R IV) with a telephoto lens (Sony 200-600 mm) was used to take photographs to assist with whale identification. Ideally, at least two primary observers were present for an LBCO survey. If only one primary observer could be present at any given time, priority was given to obtaining whale points with the theodolite and collecting whale event data before beginning photography. If the individual whale, pod, or ecotype could be identified in real-time, this was recorded. All identifications were later verified by comparison with whale identification catalogues. Ecotype identifications were verified for all whale events. When the theodolite could not be used during the LBCO survey due to weather restrictions or technical difficulty, whale event distances were estimated relative to visual landmarks and recorded, and an estimated location was reported by the primary observer in the WhaleReport app. A map was created of the points and tracks of each whale species/ecotype recorded in Mysticetus during the LBCO surveys using QGIS software.

The primary observers also tracked vessels without AIS using the theodolite and Mysticetus software to facilitate accurate reports of ISZ infractions and possible marine mammal violations to authorities on behalf of the SGIWSN. This was not included in the comparison of vessel detection methods due to a misalignment between the timing of the radar installation and the LBCO surveys and increased difficulties associated with having different fields of view for methods of vessel detection. Additional details on methods for theodolite vessel tracking during the LBCO surveys are included in Appendix C.

2.3.2. Dedicated Vessel Survey (DVS)

A DVS was conducted seven days between 09:00 and 16:00 throughout September 2023. Methods were adapted from the small vessel surveys by Le Baron et al. (2019) and Baril (2022). The primary observation site was at a consistent vantage point (48.780739º N, -123.051950º W), as shown in Figure 7. A Razor HD4000 range finder was used to measure the distance from the observer to the vessel. A visual scan in a northeast-to-southwest direction was done at 5-minute intervals, and all vessels visible in Boundary Pass within 3 km of the primary observer were recorded.

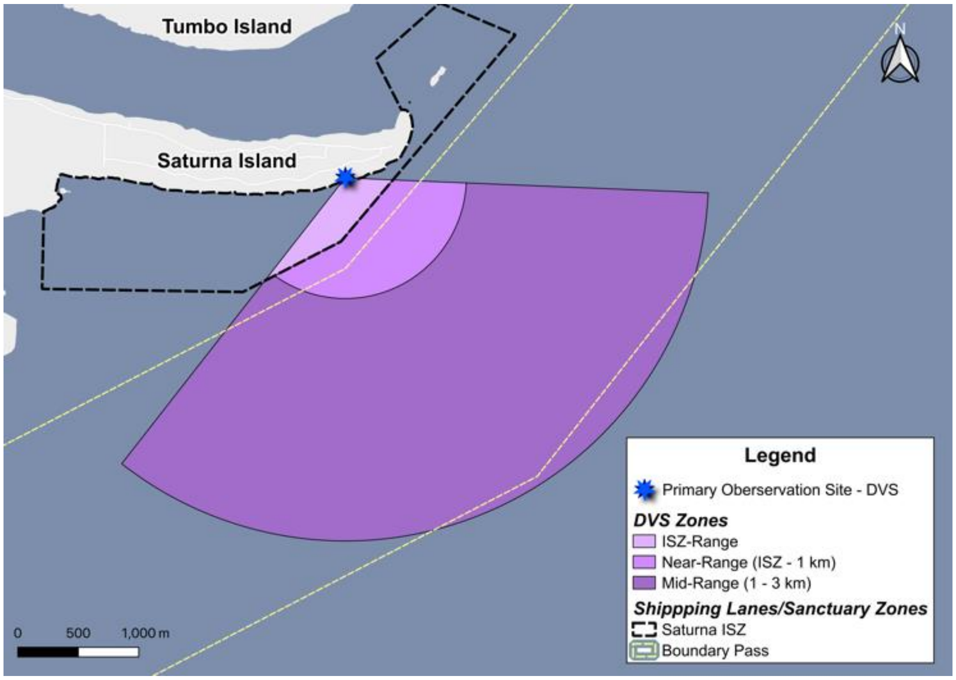


Figure 7. Map demonstrating the primary observation site and designated study zones for the dedicated vessel surveys (DVS).

Boundary Pass was split into three zones: ISZ range (all waters visible to the primary observer within the ISZ), near range (ISZ to 1 km away from the primary observer), mid-range (1-3 km away from the primary observer) (Figure 7). The primary observer calculated the distance from the primary observation site to the edge of the ISZ at angles toward visible reference landmarks across Boundary Pass. This was done to assess whether the measured distance by the range finder from the primary observation site to the detected vessel indicated that it was in the ISZ or the Near Zone (for calculated distances, see Appendix C). When the laser range finder could not pick up a vessel, the detection zone was estimated by comparing it to visual landmarks such as buoys at known distances.

Vessel detections were classified into the following categories: small motorized, sailing vessel (sailing), sailing vessel (motoring), whale watching, fishing, government, large commercial, and other. The number of each type of vessel in each zone was recorded at the start of each 5-minute interval. It would be re-recorded if a vessel were still present by the start of the next 5-minute interval. Any time the primary observer could not perform a scan at the beginning of the 5-minute interval was given a ‘no data’ value. All vessels detected in each zone on each survey day were counted using Microsoft Excel and plotted in a histogram using R Studio.

2.4. Alternative Methods of Detection

The following subsections explain the methods for the alternative methods of detection. Data for these methods were collected from the organizations that run them and are mentioned in each subsection.

2.4.1. Citizen Science

Whale reports from opportunistic sightings by volunteer citizen scientists from the SGIWSN were collected from the Ocean Wise Sightings Network (formerly British Columbia Cetacean Sightings Network) WhaleReport app via their reporting website Spyhopper.ca, hereafter referred to as Spyhopper. The SGIWSN has citizen scientists on Saturna, Pender, Mayne, and Galiano Islands. For this study, only reports from someone on Saturna Island were included. A Discord server is used by the SGIWSN to effectively communicate whale presence and discuss relevant research topics. Volunteers in the SGIWSN are trained on submitting high-quality and consistent whale. Data collected via the WhaleReport app include the date and time of the whale detection, species, ecotype (when species = killer whale), self-identified certainty of their identification, number of animals, direction of travel, estimated location (latitude and longitude), whale behaviour, whether the animal is stressed or deceased, and any additional comments provided by the sighter.

Whale reports by citizen scientists from the SGIWSN undergo quality assurance and verification by an expert human operator before they are reported on Spyhopper (Lucy Quayle, Spyhopper, personal communication, 2023). This study used the raw data for whale reports prior to quality assurance and verification to evaluate the accuracy of real-time reports. Using these data, events that occurred during the on-effort hours of the LBCO surveys were categorized as true positive, false negative, and false positive whale events. One limitation of this comparison results from the different fields of view between the LBCO surveys and the citizen scientists. Therefore, false positives outside of the field of view of the LBCO surveys were investigated and verified through other means, such as identification photos in the SGIWSN Discord or verification by the quality assurance operator for Spyhopper. If the detection could be verified, it was considered an external positive and excluded from the confusion matrix and performance metric calculations. If the event could not be verified it was added as a false positive.

The proportion of correct and incorrect species (killer whale or humpback) and ecotype identification (BKW or SRKW) were counted using Microsoft Excel. Identifications were counted as correct or incorrect based on consistent identification from the associated LBCO survey whale event. The proportion of correct to incorrect identifications was calculated and plotted as stacked histograms using R Studio.

2.4.2. Thermal Imaging

Thermal imaging data from June 1 through October 31, 2023, were collected from two infrared temperature sensors operated by Fisheries and Oceans Canada (DFO) and BioPhysics Group Erlangen. The model for the first infrared temperature sensor, hereafter referred to as Sensor A, is a FLIR thermal camera with a 12.5º horizontal field of view. The model for the second infrared temperature sensor, hereafter referred to as Sensor B, is a FLIR a65 thermal camera with a 25º horizontal field of view. Sensor A was installed and operational before the study period, and Sensor B was installed and began operating on August 18, 2023. Both sensors are stationed 15 m above sea level (48.781016° N, -123.05181° W), approximately 570 m southwest of the primary observation site (Figure 8).

The detection and classification algorithm for thermal imaging detections at this site is explained by Richter et al. (2023) and is based on Zitterbart et al. (2013). Thermal anomalies that are caused by a whale's blow or body when surfacing is automatically detected due to the contrast in their heat compared to the temperature of the surrounding water surface (Richter et al. 2023). These detections are then processed by a machine learning-based classification system to determine the probability of whale detection (Richter et al. 2023). This model can predict whether the detected heat anomaly is from a whale, fish, bird, etc. but cannot differentiate between whale species in real-time (Richter et al. 2023). Possible whale detections are later provided to a human operator as 6-second video snippets for validation and species identification (Richter et al. 2023).

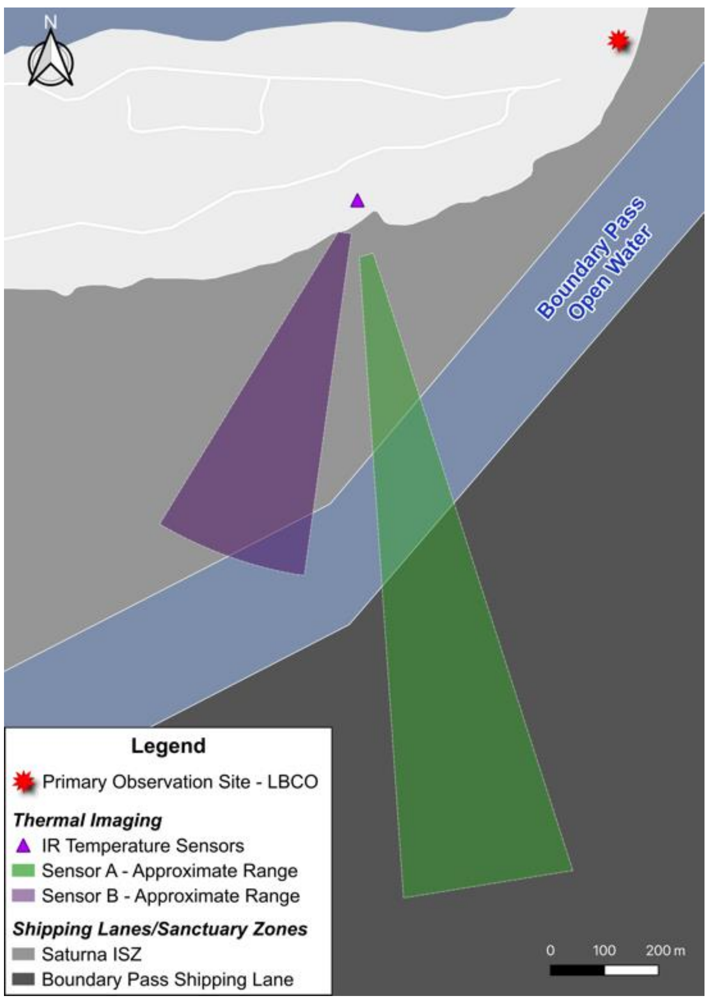


Figure 8. Map demonstrating an approximation of the detection range for Sensor A and Sensor B based on the determined location of all thermal imaging detections in the collected data relative to the primary observation site for LBCO surveys.

To determine the approximate location of a whale, the image data stream from the infrared temperature sensor is evaluated in overlapping segments (Sebastian Richter, BioPhysics Group, personal communication, 2024). The study by Richter et al. (2023) determined a reliable and maximum detection range for Sensor A. However, it has since been discovered that the 25 m altitude used to determine the location of detections was incorrect. The algorithm has since been updated to a more representative value of the actual altitude, 15 m above sea level (Sebastian Richter, BioPhysics Group, Erlangen, personal communication, 2024). The approximated ranges in Figure 8 were determined using the location of all thermal imaging detections in the data collected by each infrared temperature sensor. Short system outages of the infrared temperature sensors are difficult to determine, however, long-term outages are defined by breaks in detections greater than 24 hours (Lucy Quayle, DFO, personal communication, 2024). Any whale events detected during the LBCO surveys within a potential long-term outage would be excluded from performance metric calculations for thermal imaging.

A true positive was counted each time at least one thermal detection coincided with a whale event from the LBCO surveys and could be verified as the same event using the start and end time, direction of travel and verification by the human operator. A false negative was counted every time there were no thermal detections of a whale that coincided with a whale event from the LBCO surveys. If there was a thermal whale detection during the on-effort hours of the LBCO that did not coincide with a whale event detection by the LBCO surveys, they would be verified as external positives or false positives by looking at the corresponding video snippet.

2.4.3. Deep Artificial Neural Network (DANN) Acoustic Detection Model

Acoustic hydrophone data were collected from two Ocean Sonics icListen HF hydrophones operated by SIMRES. Hydrophones pick up underwater sound and transmit the acoustic energy through cables to a receiver or amplifier for listening or running deep learning models on the sound data. They are intended to operate continuously 24/7 and collect consecutive 5-minute audio recordings. The East Point hydrophone is deployed approximately 570 m southwest of the primary observation site (48.767503ºN, -123.08666ºW) while the Monarch Head hydrophone is deployed approximately 3500 m southwest (48.767503ºN, -123.05154ºW) (Figure 4). The East Point hydrophone was operational from June 1 through August 8, 2023, before experiencing technical difficulties. These issues resulted in no recordings collected for the remainder of the study period. Recordings from the Monarch Head hydrophone were used for analysis for the remainder of the study period from August 9 through August 31, 2023. Any time there was a lack of a 5-minute recording due to temporary system outages, any whale events that occurred within that time were excluded from the analysis (for a list of temporary system outages, see Appendix A).

Unlike the other detection methods used in this study, an estimated or actual coordinate for a whale detection is not provided. When multiple hydrophones are synchronized and combined, the probable location of a whale can be calculated

(Gassmann et al. 2013). However, this was outside of the scope of this study as only recordings from a single hydrophone without directional information were collected for any given time (either East Point or Monarch Head). A hydrophone's detection range depends on many specific geographic and weather-based factors (Prior et al. 2011). For example, there is a complex relationship between the depth, temperature, and the speed at which sound can travel through the water, as evident in the sample sound speed profile in Figure 9. Sound speed profiles change across the season, so the profile in Figure 10 would be considered a 'typical' cast but could vary slightly. This relationship also indicates that the depth of the whale at any given time would influence the sound's ability to travel to the hydrophone at a high enough amplitude for detection. Due to these complications, it was outside of the scope of this study to calculate an approximate detection range for the hydrophones. Since the potential detection range is uncertain but likely large, a buffer time of ±15 minutes was allotted to the whale event times from the LBCO surveys when determining the entities of the confusion matrix.

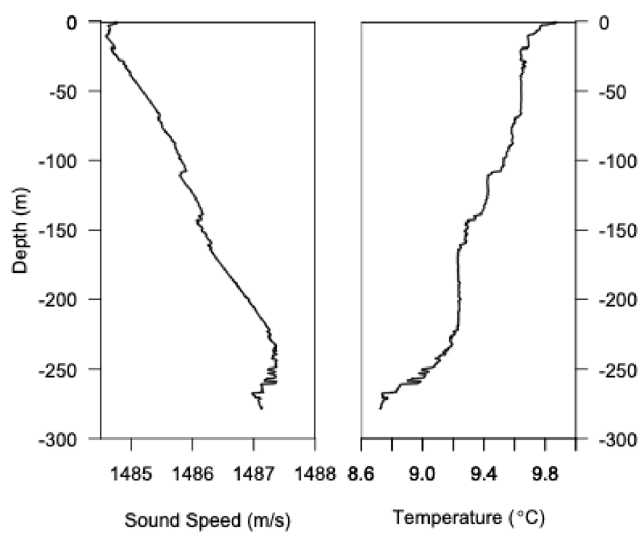


Figure 9. Sample sound speed profile and temperature at various depths of Boundary Pass from a conductivity, temperature, and depth cast on 10-26-2023 at 08:36:18 PST.

Methods for the hydrophone detection model used in this study were explained in personal communications by Fabio Soares Frazao, SFU, 2024 and Ruth Joy, SFU, 2024. The model uses the Mel spectrogram as the baseline input and uses Deep Artificial Neural Networks (DANNs) to devise pattern recognition algorithms that can identify killer whale vocalizations in hydrophone recordings. The model learns to identify spectrogram images that belong to the class ‘killer whale’ and those that do not, as ‘non-killer whale.’ The data stream is then broken into 3-second relative scores that evaluate the likelihood of a killer whale detection. If the score exceeds the ‘killer whale’ threshold, it is labelled as ‘killer whale.’ A ‘peak’ in a DANN is a point where the model's output score surpasses this predefined threshold, indicating that the model has made a ‘killer whale’ detection. The number of peaks can be used as a proxy of acoustic activity, however, a single vocalization can result in multiple peaks, so this is not a direct measurement.

Due to time constraints, the model was run before fine-tuning to the unique underwater soundscape surrounding the SIMRES hydrophones. As a result, the analysis of the performance of this model is only a preliminary test of its transferability to Boundary Pass, and its accuracy could be improved as the model continues to be trained. As the model is still learning to classify ‘killer whales’ from the acoustic stream, it is prone to false positives. Therefore, a minimum threshold of 10 peaks per 5-minute recording was selected by the developer to trigger a positive killer whale detection.

Since there is no detection distance or directional information provided by the model and the potential range of the hydrophone is undetermined, an arbitrary buffer time of 15 minutes was allotted to the recorded start and end time of each killer whale event from the LBCO surveys when counting the entities of the confusion matrix. A true positive was counted every time at least one recording detected by the model had ≥10 peaks that coincided with the buffered killer whale event. A false positive was counted every time at least one recording detected by the model with ≥10 peaks was within the LBCO observation hours but outside of the buffered killer whale event. A false negative was counted every time there was not at least one recording detected by the model that had ≥10 peaks coincided with the buffered killer whale event.

2.4.4. Radar

Radar vessel detection data were collected from a GARMIN Radome GMR^{M} 24 x HD operated by DFO from September 3 through 26, 2023. The device was stationed at 15 m above sea level (48.781016° N, 123.05181° W) and approximately 40 m north-north-east from the primary observation site for the DVS (Figure 10). It has a detection range minimum of 20 m and maximum of 88 km. The maximum detection range can be manually adjusted and was set to detect within 4 km at this site without obstruction from land. Radar works to detect vessels by sending out a radio signal via the attached antenna. When this signal encounters a vessel, it is reflected back to the radar device. The radar software has an algorithm that uses the reflected signal to calculate a distance, estimate the location, and provide a coordinate. Parameters for weather and sea conditions are automatically adjusted in the algorithm. The antenna can rotate at a speed of 24 or 48 rotations per minute, allowing multiple vessels to be tracked at any time. Each vessel is given a unique identification number, and this number is associated with each of its detected locations.

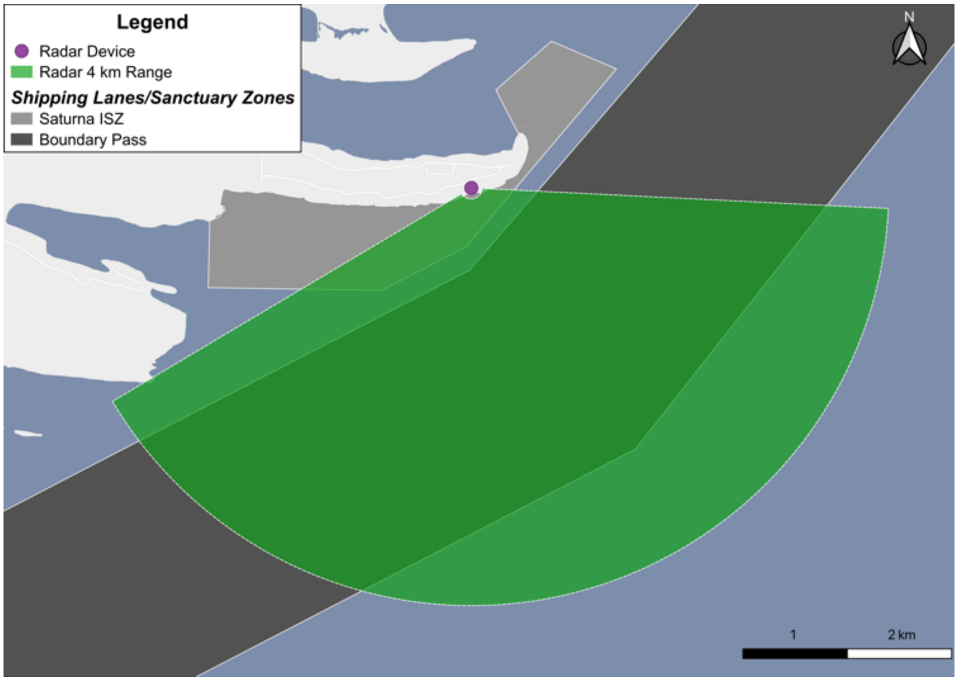


Figure 10. Map demonstrating an approximation of the 4.5 km detection range for the radar device.

The radar vessel detection data were filtered using Microsoft Excel to encompass only the detections on the same 5-minute intervals from the DVS (i.e. 09:00 PDT to 09:00:59 PDT for the 09:00 PDT DVS interval). Coordinates for the remaining detections were plotted as points on a map of the DVS study zones using QGIS. The points in each zone were selected using the select by location feature, and the associated data were separated into new Microsoft Excel sheets categorized by zone. The number of distinct vessel identifications in each zone at each interval was then counted and compared to the vessel detections of the DVS. A true positive was counted when there was a corresponding vessel detection by both radar and the DVS at a given time interval. A false positive was counted for every vessel detection by radar, for which there was no corresponding vessel detection by the DVS at a given time interval. A false negative was counted for every vessel detection by the DVS, for which there was no corresponding detection by radar at a given time interval. A true negative was counted when there were no vessel detections by radar and the DVS at a given time interval. The measure tool on QGIS was also used to estimate the distance between radar and its nearest and furthest detections in all the collected data to compare the set range of 20 to 4000 m to the range of its detections.

2.4.5. Automatic Identification Systems (AIS)

AIS is an onboard navigation safety device that transmits the location and characteristics of the vessel to nearby receivers and other vessels for collision avoidance and safety. Data for vessels equipped with an AIS device were collected by an AIS receiver antenna from Quayle Consulting Ltd. and accessed from AISHub. In Canada, *Navigation Safety Regulations 2020* Sec. 118 (2) states that vessels that are certified to carry more than 12 passengers or eight metres or more in length and carry a passenger must be fitted with either a Class A or B AIS device. Otherwise, AIS is optional, and the AIS receiver will miss vessels without an AIS device. Along with the determined coordinate estimates of a detected vessel, AIS data includes identification information of the vessel, such as the unique Maritime Mobile Services Identity assigned to each vessel with an AIS device.

The AIS vessel detection data were filtered using Microsoft Excel to encompass only detections on the same 5-minute intervals from the DVS (i.e. 09:00:00 PDT to 09:00:59 PDT for the 09:00 PDT DVS interval). Detections at intervals where data was not collected in the DVS were also excluded. Using QGIS coordinates for the remaining detections were plotted as points on a map of the DVS study zones. The points in each zone were selected using the select by location feature, and the associated data were separated into new Microsoft Excel sheets categorized by zone. The number of distinct Maritime Mobile Services Identities in each zone at each interval was compared to the vessel detections in the DVS. A true positive was counted when AIS and the DVS detected a corresponding vessel in the same zone at a given time interval. A false positive was counted for every vessel detection by AIS, for which there was no corresponding vessel detection by the DVS in the same zone at a given time interval. A false negative was counted for every vessel detection by the DVS for which there was no corresponding AIS detection in the same zone at a given time interval. A true negative was counted when there were no vessel detections by both AIS and the DVS in the same zone at a given time interval.

Chapter 3. Results

3.1. Whale Detections

During the 76 LBCO surveys between June 1 through August 31, 2023, 51 whale events were detected. Seven were SRKWs, 22 were BKWs, and 22 were humpback whale events. Points and tracks collected with the theodolite and Mysticetus software or the reported location to the WhaleReport app from these events are featured in Figure 11. Additional details about the whale events are included in Appendix C.

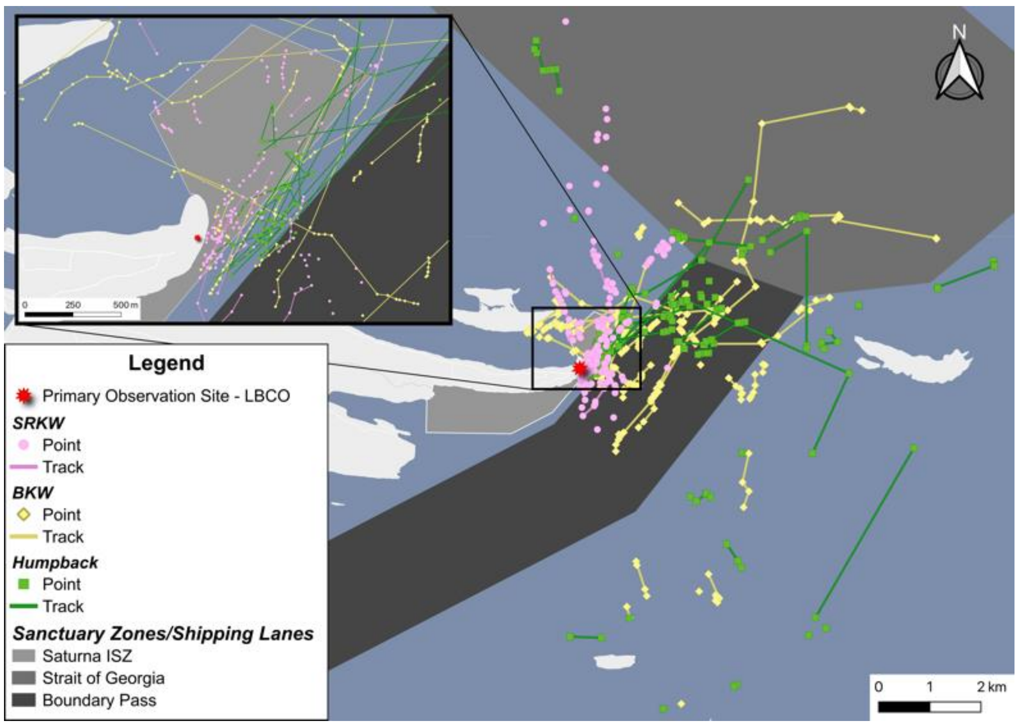


Figure 11. Points and tracks demonstrating the location of Southern Resident killer whales (SRKW), Bigg's killer whales (BKW), and humpback whales detected during the land-based cetacean observation (LBCO) surveys. Surveys occurred on 76 days between June 1 through August 31, 2023. The zoom panel is focused on detections in the Saturna Island Interim Sanctuary Zone (ISZ).

The citizen scientists of the SGIWSN submitted 109 reports of whales from Saturna Island to the WhaleReport app that coincided with the on-effort hours of the LBCO surveys. The estimated location of the whale in the report by the citizen scientist is featured in Figure 12. These reports were determined to be of 43 distinct whale events. Five reports described 3 whale events outside the field of view of the LBCO survey. These events were verified as external positives and excluded from the confusion matrices, leaving 104 reports of 40 distinct whale events.

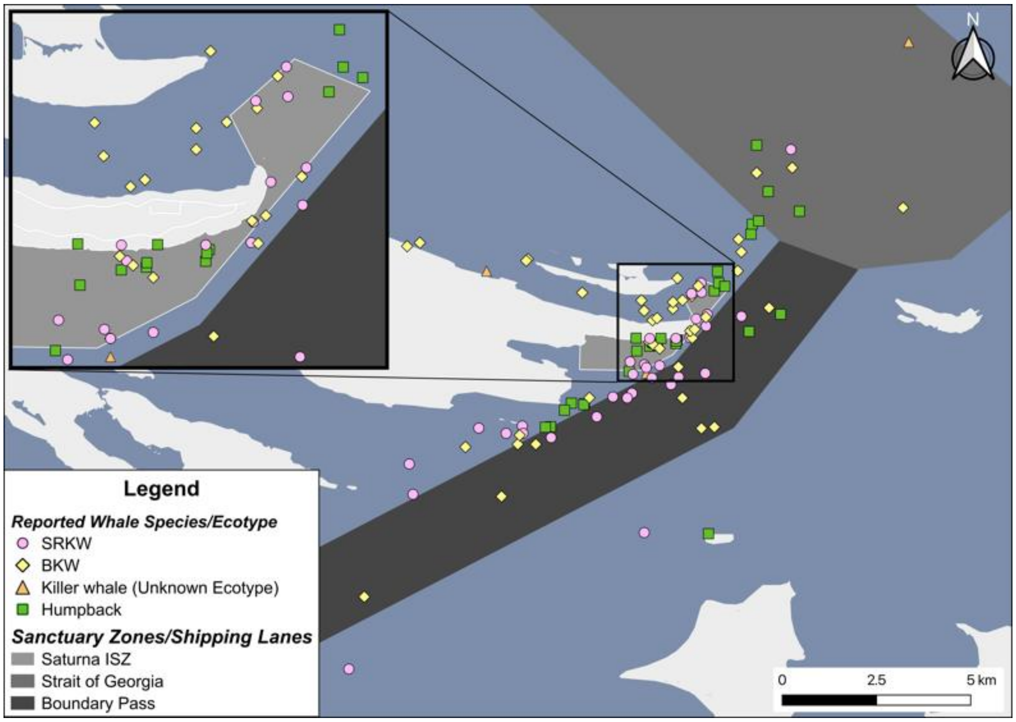


Figure 12. Points demonstrating the location for reported Southern Resident killer whales (SRKW), Bigg's killer whales (BKW), killer whales (unknown ecotype) and humpback whales detected by the citizen scientists during on-effort hours of the land-based cetacean observation (LBCO) surveys. The zoom panel is focused on detections in the Saturna Interim Sanctuary Zone (ISZ).

There were 103 thermal imaging detections during the on-effort hours of the LBCO surveys. Sensor B began operating on August 18, 2018, and had one detection during the on-effort hours of the LBCO surveys. The remaining 102 detections were by Sensor A. All these detections were verified as whales by humans viewing the infrared imagery, and nine whale events from the LBCO surveys were detected. There were no thermal imaging detections to be confirmed external or false positives as they all aligned with the whale event times of the LBCO surveys. There was one potential long-term system outage during the study period in Sensor A (dates and times listed in Appendix B), but the LBCO surveys detected no whale events during this time.

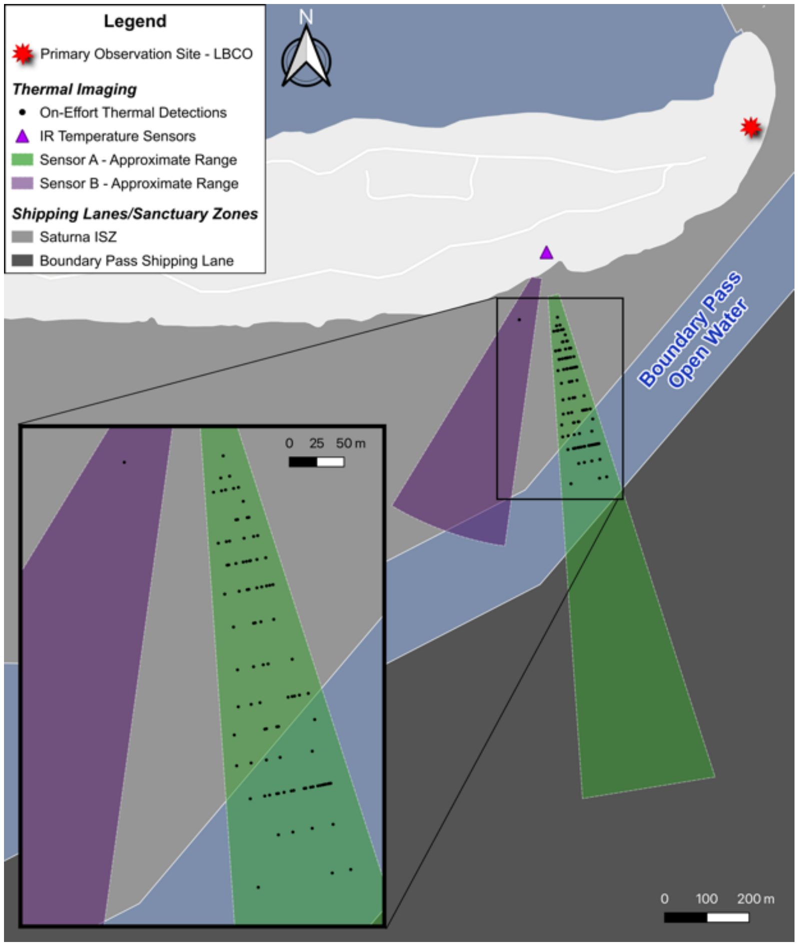


Figure 13. **A map demonstrating the thermal whale detections by Sensor A and Sensor B during the on-effort hours of the land-based cetacean observation (LBCO) surveys. Detections are shown within the approximated range of the sensors. A zoom panel is centred on all on-effort whale detections.**

Sensor A detected all 7 SRKW events, while Sensor B was not yet operational during any of those events. Only one detection was recorded for both BKW and humpback whales by Sensor A and Sensor B, respectively. Sensor B was not operational when the BKW was detected. Figure 14 shows a screen capture from the infrared video snippet of an SRKW detected by Sensor A on July 4, 2023. Figure 15 shows a screen capture from the infrared video snippet of a humpback whale detected by Sensor B on August 18, 2023.

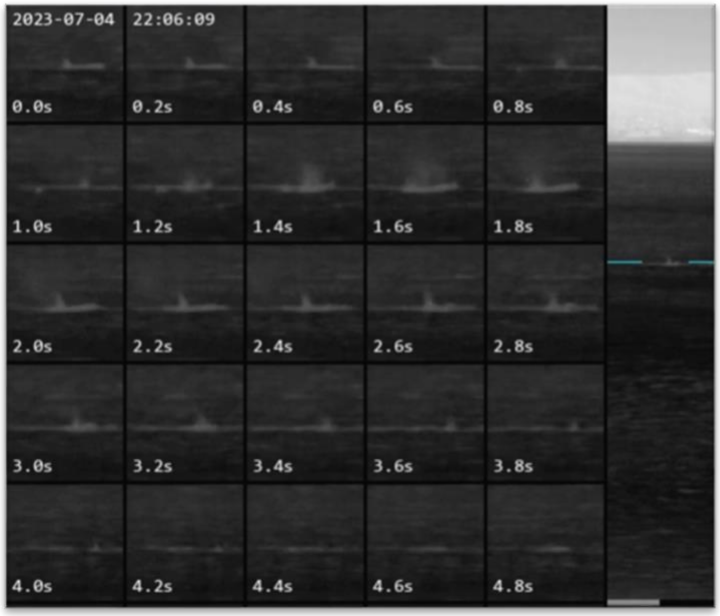
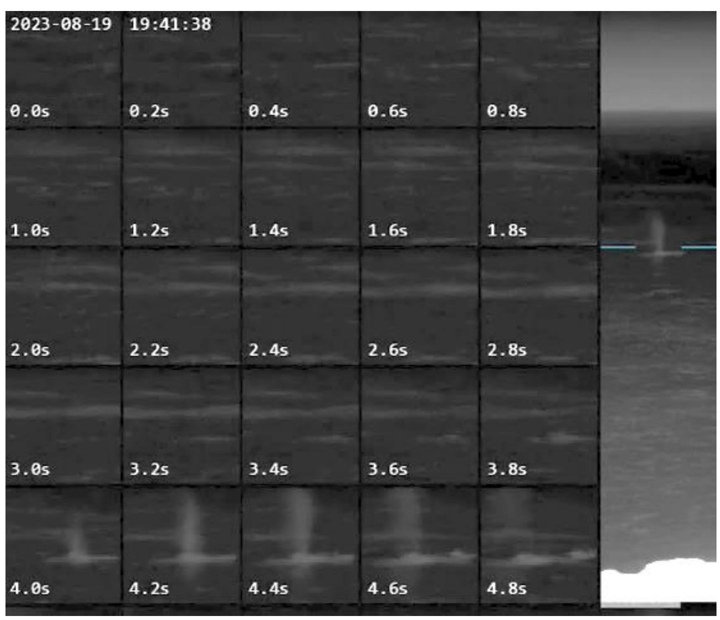


Figure 14. **Sample screen capture of a video snippet of an SRKW detected through thermal imaging by Sensor A on July 4, 2023, at 22:06:09 UTC (15:06:09 PDT).**



Note: The image brightness was manually enhanced for ease of viewing in this report.

Figure 15. **Sample screen capture of a video snippet of a humpback whale detected through thermal imaging by Sensor B on August 18, 2023, at 19:41:38 UTC (12:41:06:38 PDT).**

Between June 1 and August 31, 2023, the DANN model detected 7752 5-minute hydrophone recordings that contained ≥10 peaks, 1570 of which were recorded during on-effort LBCO hours (Figure 16). These recordings represent positive killer whale detections by the model. There were multiple short-term system outages, including one that spanned the entire duration of the BKW event detected in the LBCO surveys on August 27, 2023 (for dates and times of all system outages, see Appendix B). Therefore, this event was excluded from the calculation of performance metrics for the DANN model.

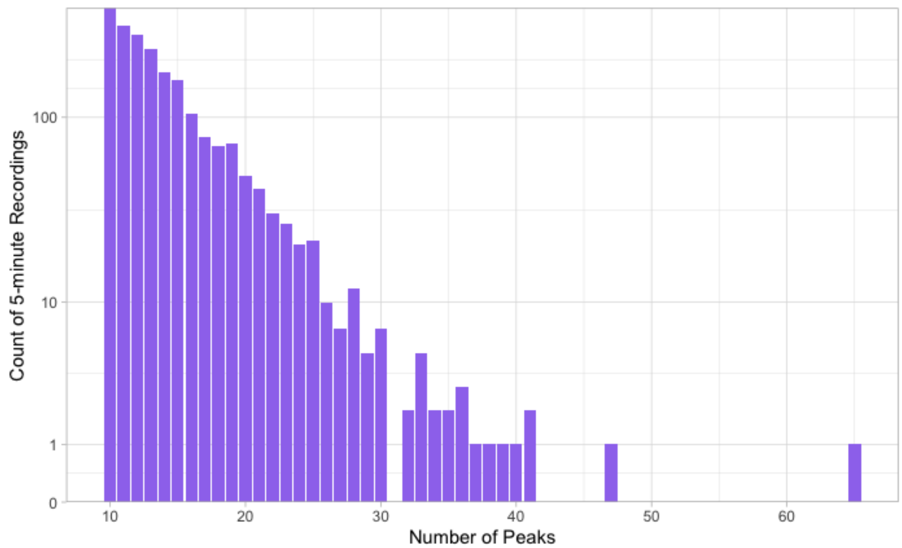


Figure 16. A histogram illustrating the distribution of recordings detected by the DANN model across the number of peaks detected, beginning at 10 peaks. A logarithmic scale is applied to the y-axis.

3.2. Performance Metrics for Methods of Whale Detection

3.2.1. Citizen Science

When comparing whale detections by the citizen scientists to those of the LBCO, the recall was 0.86 for detecting killer whale events, 0.68 for humpback whale events, and 0.78 for events of both species combined. The precision was 1.00 for both species. The proportion of whale events detected during the LBCO surveys that the citizen scientists also detected is shown in Figure 17.

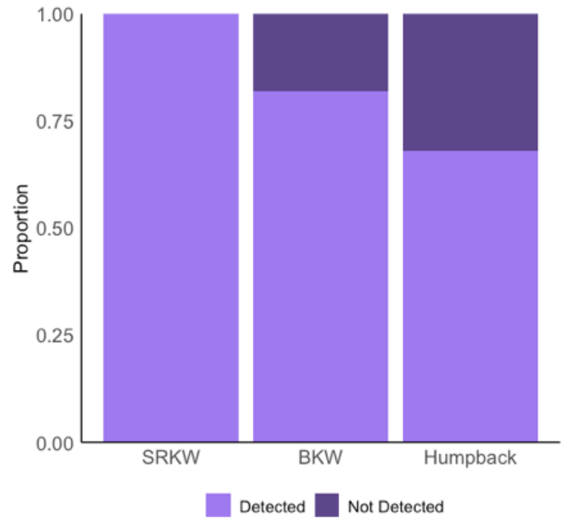


Figure 17. The proportion of whale events for Southern Resident killer whales (SRKW), Bigg's killer whales (BKW), and humpback whales either detected or not detected by citizen scientists from the Southern Gulf Islands Whale Sighting Network (SGIWSN). The data is compared to whale event data collected through the gold standard method, dedicated land-based cetacean observation surveys.

The proportion of correct or incorrect whale species identifications by the citizen scientists is shown in Figure 18. All species identifications made in the whale reports by the citizen scientists was consistent with the LBCO survey identification for both killer and humpback whales.

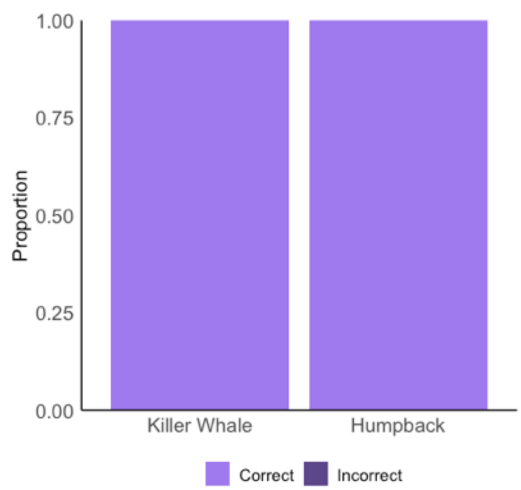


Figure 18. The proportion of correct and incorrect whale species identifications for killer and humpback whales in reports by citizen scientists from the Southern Gulf Islands Whale Sighting Network (SGIWSN). The data is compared to whale event data collected through the gold standard method, dedicated land-based cetacean observation surveys.

The proportion of correct or incorrect killer whale ecotype identifications by the citizen scientists is shown in Figure 19. In 4% of their killer whale reports, citizen scientists did not choose an ecotype that was consistent with the ecotype identified in the LBCO surveys (Figure 19). The ecotype was selected as “Unknown” for 3% of their killer whale reports. The remaining 1% chose an ecotype in their killer whale report that was inconsistent with the identification from the LBCO survey. Specifically, the ecotype was selected as a “Possible Southern Resident” when it was a BKW. This was also the only whale report by the citizen scientists with “Possible” associated with the ecotype selection during the on-effort LBCO survey whale events.

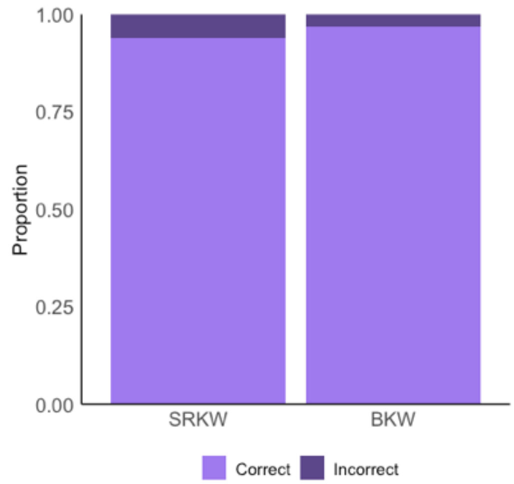


Figure 19. The proportion of correct and incorrectly reported killer whale ecotypes for Southern Resident killer whales (SRKW) and Bigg’s killer whales (BKW) by citizen scientists from the Southern Gulf Islands Whale Sighting Network (SGIWSN). The data is compared to whale event data collected through the gold standard method, dedicated land-based cetacean observation surveys.

3.2.2. Thermal Imaging

When comparing whale detections by thermal imaging to those of the LBCO surveys, the recall was 0.27 for detecting killer whale events, 0.05 for humpback whale events, and 0.18 for events of both species combined. The precision was 1.00 for both species. The proportion of whale events detected during the LBCO surveys that were also detected through thermal imaging is shown in Figure 20.

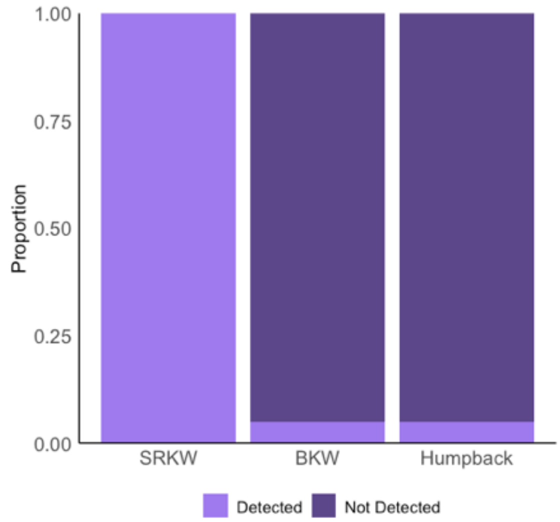


Figure 20. The proportion of whale events for Southern Resident killer whales (SRKW), Bigg's killer whales (BKW), and humpback whales (HBW) that were either detected or not detected by the infrared temperature sensor(s). The data is compared to whale event data collected through the gold standard method, dedicated land-based cetacean observation surveys.

During the LBCO surveys, six BKW and eight humpback whale events were observed in Tumbo Channel or the Strait of Georgia and the whales were not observed to have travelled into Boundary Pass. The performance metrics were calculated again to include only events that occurred in Boundary Pass and the proportions of events detected are shown in Figure 21. The proportion of SRKW events detected was unchanged, but the proportion of BKW increased by 0.01, and the proportion of humpback whales increased by 0.02. The precision stayed the same for all whale events. The recall increased to 0.28 for killer whales, 0.05 for humpback whales, and 0.24 for all whale events.

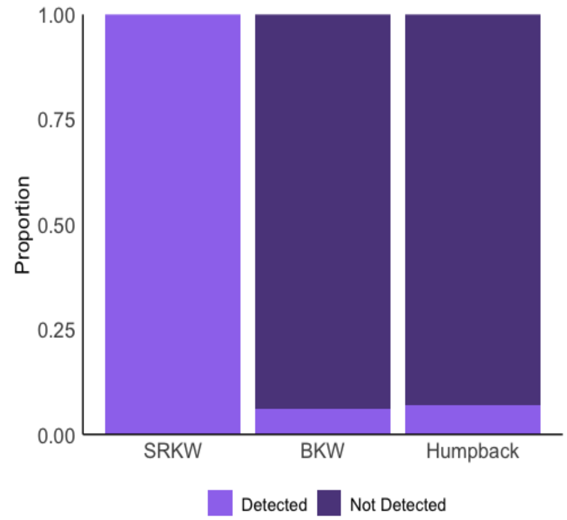


Figure 21. **The proportion of whale events confirmed to be in Boundary Pass for Southern Resident killer whales (SRKW), Bigg's killer whales (BKW), and humpback whales (HBW) either detected or not detected by the Infrared temperature sensors(s). The data is compared to whale event data collected through the gold standard method, dedicated land-based cetacean observation surveys.**

3.2.3. DANN Hydrophone Killer Whale Detection Model

When comparing killer whale detections by the DANN model to those of the LBCO surveys, the recall was 0.96 and the precision was 0.28. Figure 28 shows the proportion of whale events detected by the DANN model.

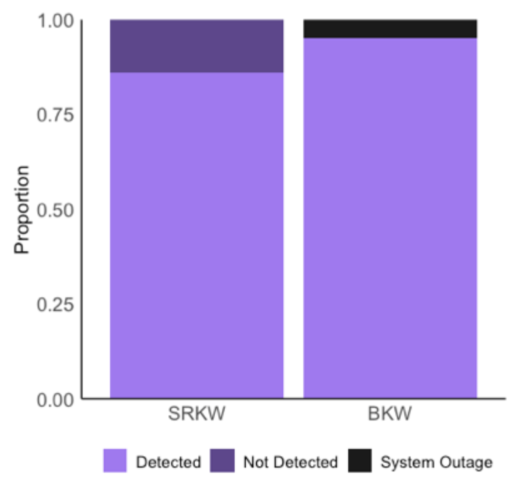


Figure 22. **The proportion of killer whale events either detected or not detected by the DANN model for Southern Resident killer whales (SRKW) and Bigg's killer whales (BKW). The proportion of events during a hydrophone system outage is also included. The data is compared to whale events detected by the gold standard method, the land-based cetacean observation surveys.**

3.2.4. F1 Scores

Table 2 summarizes the F1 scores of the alternative whale detection methods. The citizen scientists had the highest F1 score of all the alternative methods of whale detection. Although thermal imaging exhibited a higher F1 score when only events in Boundary Pass were considered in its performance metric calculations, it matched the DANN model's F1 score when all events were considered. The DANN model does not have an F1 score for humpback whale events or all whale events because it cannot detect humpback whales.

Table 2. F1 score of each whale detection method, calculated based on its ability to detect a killer whale event, a humpback whale event, or all whale events. A higher F1 score indicates a higher level of reliability in the ability to detect a killer whale event.

	F1 Score		
Detection Method	**Killer whale events**	**Humpback whale events**	**All whale events**
Citizen science	0.93	0.81	0.88
Thermal imaging (Events in Boundary Pass)	0.52	0.13	0.39
Thermal imaging (All Events)	0.43	0.09	0.30
DANN Model	0.43	NA[1]	NA[1]

[1]F1 score not calculated as DANN model cannot detect humpback whale events.

3.3. Vessel Detections

A range of 32-65 vessels were counted during each day of the DVS, with an average of 42.86 vessels per day (*SD* = 11.39). The number of vessels counted in each zone during the DVS is shown in Figure 23.

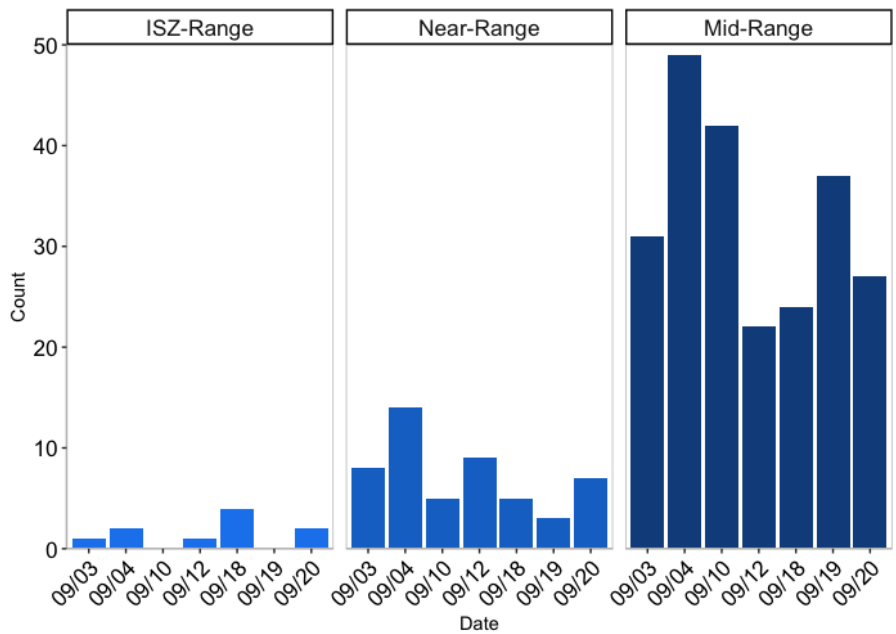


Figure 23. The number of vessels counted in each zone during the dedicated vessel surveys (DVS). Vessels were counted during scans at the beginning of 5-minute intervals between 09:00 and 16:00 PDT.

Radar detected 215 distinct vessels on the 5-minute intervals of the DVS. The count of vessels detected in each zone is shown in Figure 24, and the coordinates of the detected vessels are demonstrated on a map in Figure 25. Total vessels detected per survey day ranged from 21 to 46, averaging 30.71 vessels (*SD* = 9.93). The approximated distance to vessels detected in all the data collected ranged from ~105 m to ~4100 m.

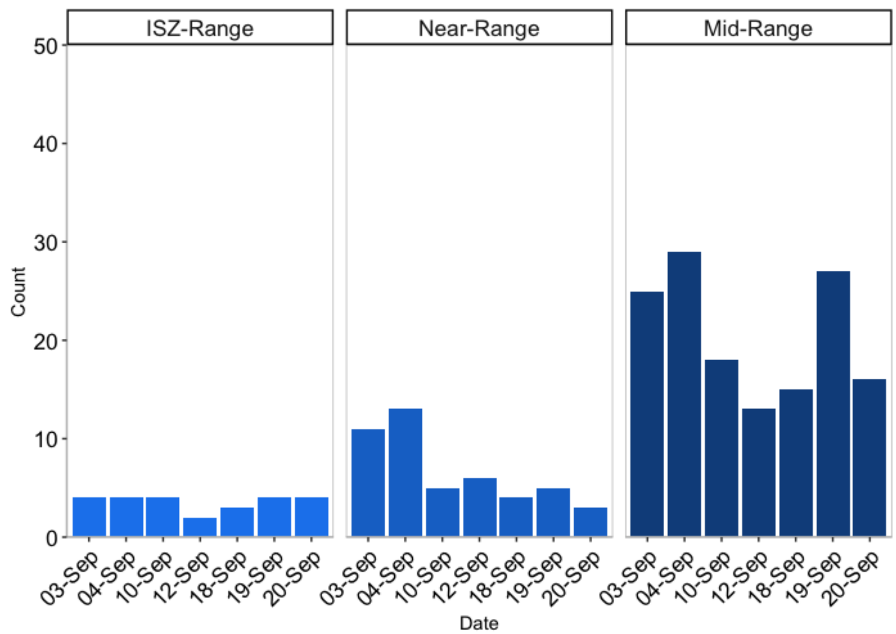


Figure 24. **Number of vessels detected in each zone by radar on the 5-minute intervals during every day of the dedicated vessel survey (DVS) between 09:00-16:00 PDT.**

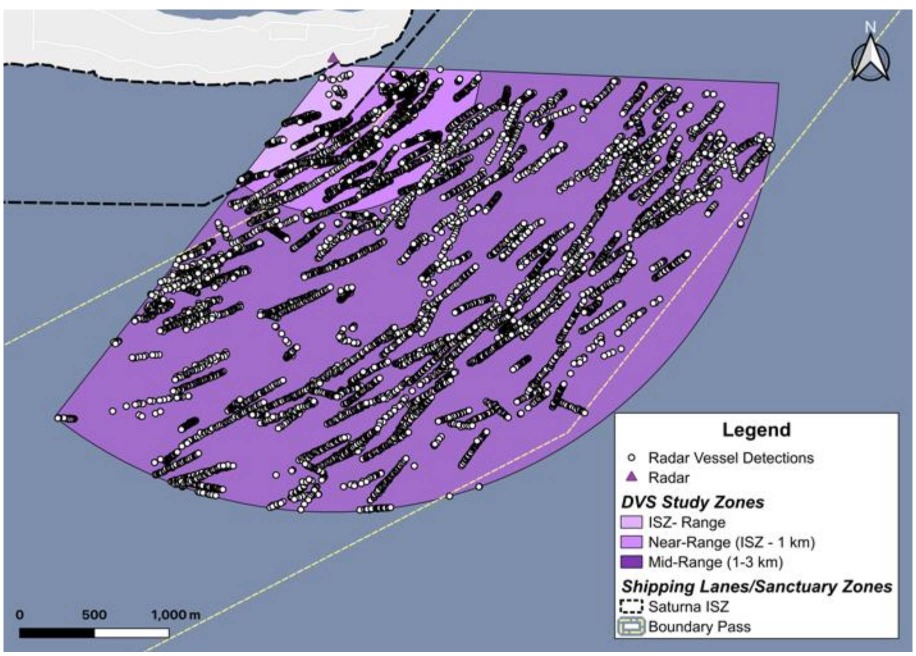


Figure 25. **Map demonstrating radar vessel detections at 5-minute intervals in each zone used for the dedicated vessel surveys (DVS).**

In total, 146 distinct vessels were detected on the 5-minute intervals of the DVS through AIS. The count of vessels detected in each zone is shown in Figure 26, and the coordinates of the detected vessels are demonstrated on a map in Figure 27. The total number of vessels detected daily by AIS ranged from 10 to 28, with an average of 20.86 (*SD* = 6.52).

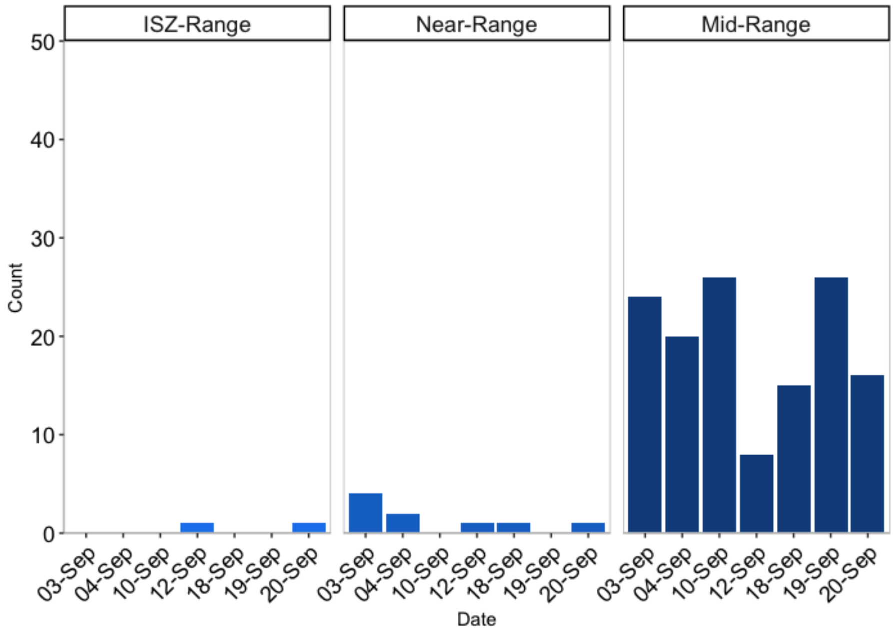


Figure 26. **The number of vessels detected in each zone by AIS at 5-minute intervals during each day of the dedicated vessel survey (DVS) between 09:00 and 16:00 PDT.**

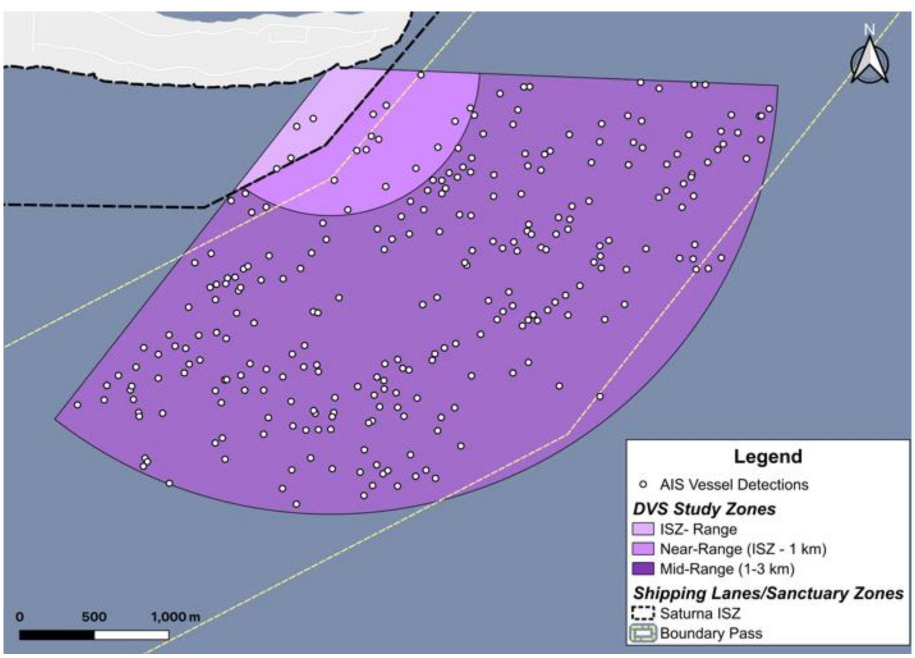


Figure 27. **Map demonstrating AIS vessel detections at 5-minute intervals and in each zone used for the dedicated vessel surveys (DVS).**

3.4. Performance Metrics for Methods of Vessel Detection

3.4.1. Radar

Figure 29 shows the proportion of vessels counted in the DVS that were detected by radar in each zone. Radar successfully detected nearly half of the vessels counted in the DVS, with the highest proportion of detections in the mid-range and the lowest in the ISZ range.

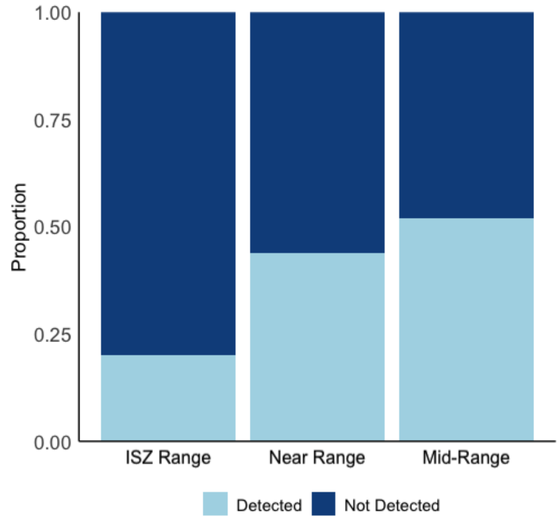


Figure 28. The proportion of vessels detected or not detected by the radar in each study zone. The data is compared to vessel data collected through the gold standard method, dedicated vessel surveys.

Table 3 summarizes the performance metrics for vessel detection by radar. The F1 scores indicate that radar is least reliable for detecting the presence of vessels in the ISZ range and most reliable in the mid-range. The MCCs were all significant and indicated a positive correlation between radar's vessel detections and the vessel detections in the DVS. This correlation was strongest in the near range and when results of all zones were combined and weakest in the ISZ range.

Table 3. Summary of performance metrics for vessel detections by radar compared to the gold standard (most reliable) method, the dedicated vessel surveys (DVS).

	Precision	Recall	F1 Score	NPV	MCC
ISZ	0.22	0.20	0.21	0.99	0.20*
Near	0.65	0.44	0.52	0.95	0.50*
Mid	0.67	0.52	0.58	0.77	0.41*
All Zones	0.65	0.49	0.56	0.91	0.50*

* Significant (p<0.05).

3.4.2. AIS

Figure 30 shows the proportion of vessels counted in the DVS detected by radar in each zone. AIS successfully detected 42% of the total vessels counted in the DVS, with the highest proportion of detections in the mid-range and the lowest in the ISZ range.

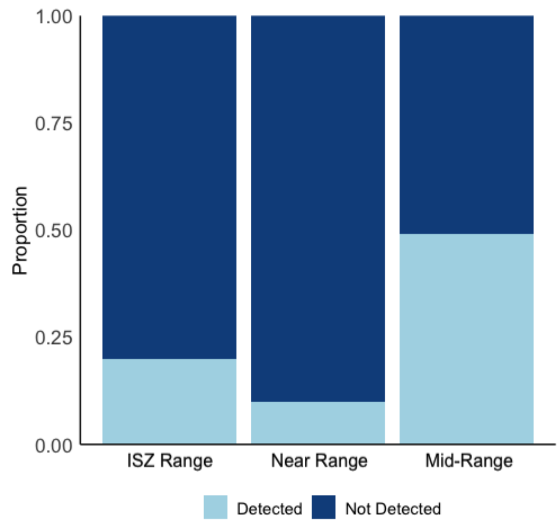


Figure 29. The proportion of vessels detected or not detected by AIS in each study zone. The data is compared to vessel data collected through the gold standard (most reliable) method, dedicated vessel surveys.

Table 4 summarizes the performance metrics for vessel detection by AIS. The F1 scores indicate that AIS is less reliable for detecting the presence of vessels in the near range and most reliable in the mid-range. The MCCs were significant and showed a positive correlation between the AIS and DVS vessel detections. Looking at individual zones, the correlation was strongest in the mid-range and weakest in the near range.

Table 4. Summary of performance metrics for vessel detections by AIS compared to the gold standard (most reliable) method, the dedicated vessel surveys.

DVS Zone	Precision	Recall	F1 Score	NPV	MCC
ISZ	1.00	0.20	0.33	0.99	0.44*
Near	0.56	0.10	0.17	0.92	0.21*
Mid	0.85	0.49	0.63	0.76	0.52*
All Zones	0.84	0.42	0.56	0.90	0.54*

* Significant ($p<0.05$)

3.4.3. F1 Scores and MCC

Table 5 summarizes the F1 scores and the MCCs for the methods of vessel detection. Radar and AIS both had their highest respective F1 scores in the mid-range. This indicates they were most reliable for detecting vessels in the mid-range, but AIS was found to be more reliable in this zone. Radar had its lowest F1 score in the ISZ range. This indicates it was least reliable in its ability to detect vessels in the ISZ range and had a lower score in this zone than AIS. AIS had its lowest F1 score in the near range. This indicates it was least reliable in its ability to detect vessels in the near range and was lower than radar in this zone. When the results of all zones were combined, they had identical F1 scores of 0.56. MCC calculations indicate that detections by AIS were more strongly correlated with detections from the DVS than radar in all zones except for the near range.

Table 5. F1 scores and MCCs of the vessel detection methods for each study zone and overall in all zones. A higher F1 score indicates a higher level of reliability in the ability to detect a vessel. A higher MCC indicates a higher overall agreement between the method and the dedicated vessel surveys.

DVS Zone	F1 Score		MCC	
	Radar	**AIS**	**Radar**	**AIS**
ISZ	0.21	0.33	0.20*	0.44*
Near	0.52	0.17	0.50*	0.21*
Mid	0.58	0.63	0.41*	0.52*
All Zones	0.56	0.56	0.50*	0.54*

* MCC is significant ($p<0.05$)

Chapter 4. Discussion

The performance metrics based on machine learning calculated in this study can be used to compare the performance of the alternative whale and vessel detection methods. Citizen scientists from the SGIWSN emerged with the highest F1 score for both killer whale and humpback whale event detections, indicating that this alternative method has the highest level of reliability for detecting whale events. The preliminary test of the DANN model's transferability to the SIMRES hydrophones resulted in an F1 score of 0.43, the same F1 score for killer whale detections by thermal imaging when all events were considered. When only whale events in Boundary Pass were considered in the performance metric calculations for thermal imaging, it outperformed the DANN model. The score for the DANN model could improve as it is fine-tuned to the unique soundscape of the SIMRES hydrophones.

When detections in all zones of the DVS were considered, radar and AIS had the same F1 score, indicating the same level of reliability for detecting vessels. However, the MCC calculation revealed that AIS had a slightly stronger correlation with the DVS than radar. AIS generally outperformed radar in the ISZ and mid-ranges, and radar outperformed AIS in the near range. AIS consistently had higher precision values, and radar consistently had higher recall values in all zones. This indicates that AIS was less likely to detect a false positive, while radar was more likely to detect a vessel counted in the DVS. The probability of an interval with no alternative method detection being a true negative as determined by the NPVs was similar in all three zones (within ±0.03). It demonstrated the same trend, with NPV being highest in the ISZ range and lowest in the mid-range.

4.1. Bias in the Alternative Methods of Whale Detection

All three methods of whale detection consistently detected most or all the SRKW events detected during the LBCO surveys. The DANN model missed an SRKW event only once, while the others detected all seven SRKW events. BKW events were missed more frequently than SRKW by the citizen scientists and thermal imaging but less regularly by the DANN model. For alternative methods that could detect both killer and humpback whale events, both were more likely to miss humpbacks than killer whales. However, when the killer whale events are further specified by ecotypes, thermal imaging detected the same proportion of BKW events as it did humpback whale events.

Most humpback whale detections by thermal imaging are triggered by their blow, but the blow of a killer whale is smaller and not detected as often (Richter et al. 2023). Thermal detections of killer whales are more often made by heat emitted from their body during surface active behaviours (Richter et al. 2023). In previous LBCO surveys at East Point, SRKWs have been observed to demonstrate social surface-active behaviours, such as full-body breaching, more often than BKWs (Quayle 2021). This is consistent with the behaviour observations recorded during the LBCO surveys (as shown in Appendix C Table C3).

The pod size of killer whales can also impact the likelihood of thermal imaging detections and may contribute to the higher number of SRKW detections (Richter et al. 2023). SRKW events were frequently spread out between the ISZ and the shipping lanes, with upwards of 30 individual whales present during a whale event (Appendix C Table C3). Larger group sizes increase the likelihood that one will surface within the sensor's narrow field of view and be detected (Richter et al. 2023). During the LBCO surveys, a maximum of seven individuals were seen during the BKW events and a maximum of two individuals during humpback whale events (Appendix C Table C3). Therefore, behavioural, and morphological differences could explain the difference in the recall for detections of the different species/ecotypes in this study by thermal imaging.

The furthest thermal imaging detections in the data from June through August of 2023 extend to 549 m from Sensor A, while detections of at least 1274 m away were seen in the data for September and October of 2023. Thermal imaging detections are biased towards colder water temperatures, and as such it can be more challenging to distinguish the temperature anomaly between the whale's body heat and the warmer water (Zitterbart et al. 2013). Therefore, research on the possible influence of distance to the whale on the relationship between detection rates and water temperature in different seasons is recommended for future research. Like methods that rely on human observers, thermal imaging is also known to be limited by weather conditions, such as extreme wind or rain (Richter et al., 2023).

The DANN model exhibits the highest species detection bias as it is limited to killer whales; however, hydrophones are not, and noise from any cetacean can be recorded, including those of humpback whales (Payne & McVay 1971). It should also be noted that even if this model was trained to detect humpback whales, it may not have been reliable at this time of year. The long sequences of structured vocalizations, known as songs, produced by male humpback whales are common in their winter breeding grounds but not their summer feeding grounds (Herman et al. 2013; Ryan et al. 2019). They have been found to make shorter distinctive sounds at their summer feeding grounds, lasting around 1 second or less, and can be heard on a hydrophone called wops and grunts (Stimpert et al. 2011). Hydrophones have also been used to conduct studies on bubble-net feeding behaviours in humpback whales. These events occur when either a single whale or group of humpbacks release bubbles that form nets to trap their prey (Leighton et al.). This behaviour was not observed by the primary observers in the LBCO surveys for this study and has not been reported in previous LBCO surveys at East Point (Quayle 2021; Gheibi 2022). Further research on potential artificial intelligence models to detect these quick sounds emitted by humpbacks in their summer feeding grounds is recommended.

The DANN model was the only method that detected a higher proportion of BKW events than SRKW events (Figure 22). Excluding events occurring during hydrophone system outages, the model detected all BKW events. This finding was unexpected, as SRKWs are known to be highly vocal while BKWs are often not (Myers et al. 2021). BKWs prey on marine mammals that have a high sensitivity to underwater noises in the vocal range of killer whales (Myers et al. 2021). As such, the BKWs will hunt silently and use stealth and passive listening to find their prey rather than echolocation (Myers et al. 2021). As the model was not fine-tuned to the specific soundscape of the hydrophone locations in Boundary Pass, the model had a high proportion of false positive detections. Therefore, it is possible that the detections in these events were from a noise emitted by something other than a killer whale (personal communication, Fabio Soares Frazao, SFU, 2024). Whale vocalizations can be masked by loud vessel activity picked up by the hydrophone (Zitterbart et al. 2013). In contrast, vessel noise does not limit thermal imaging detections and is unlikely to restrict sightings by citizen scientists (Richter et al. 2023; Zitterbart et al. 2013).

While citizen science emerged as the most reliable detection method in this study, it is important to remember that the gold standard method was also human-based. This means that both the LBCO surveys and citizen scientist detections were biased towards times of day and weather conditions suitable for human observation. This could have provided an advantage to citizen science over the other alternative detection methods. Due to time constraints, weather conditions were not considered in performance calculations. The detection probability for methods that rely on human observers decreases in poor weather conditions, often due to reduced visibility and high sea states, making it challenging to detect orally or visually (Richter et al. 2023). Additionally, due to the opportunistic nature of sightings, it can be difficult to quantify observation hours for comparison research. These considerations do not underscore the value of a well-organized and trained citizen science initiative for providing reliable whale sightings data. Citizen science networks like these may be challenging to establish elsewhere, as many homes in the Southern Gulf Islands are near or directly on the shore where frequent observations can be made. However, these networks and initiatives should still be encouraged in all communities that are interested in contributing to whale research.

4.2. Reporting Species or Ecotype in Real-Time

Knowing the species or ecotype present in real-time is irrelevant when informing nearby vessels of whale presence, as both whale species are at risk of vessel strikes (COSEWIC 2008, 2022). However, species and ecotype information are valuable for adaptive management of whale habitat restoration and recovery initiatives specific to individual species. Citizen scientists can report both species and ecotypes in real-time, and this study found that 100% of their species identifications and 96% of the ecotype identifications in their reports were consistent with the ecotype confirmed through the LBCO surveys. Citizen science was the only alternative method in this study to identify ecotypes in real-time.

The thermal imaging model can predict whether the detected heat anomaly is from a whale, fish, bird, etc. but cannot differentiate between whale species in real-time (Richter et al. 2023). In this study, the precision for the thermal imaging was 1.00, indicating no false positive detections of a whale by the model algorithm. The species detected through thermal imaging can often be identified later by a human operator through the image stream. This can be done by the distinct and large blow indicative of a humpback whale or by the characteristic and visible dorsal fins of the killer whales (Graber et al. 2011; Richter et al. 2023). Recent studies suggest that through continued model training using human verifications, it will likely be able to distinguish and differentiate between species shortly (Richter et al. 2023). The video stream recorded of the infrared temperature sensors is quick and often not clear enough to make an ecotype identification with certainty (Richter et al., 2023).

As previously mentioned, the DANN model is not trained to identify humpback whales and could not be used for alerting of humpback presence in real-time. As a binary classifier that determines any noise as either a “killer whale” or “not a killer whale,” the DANN model also does not distinguish between the ecotypes. This is not a limitation for all hydrophone detection methods and models (Myers et al. 2021). Specific SRKW pods can be identified through hydrophone recordings, as they each exhibit vocalizations, such as calls or whistles, that are unique to each individual pod (Ford et al. 2023).

In 2018, the Ocean Wise Sighting Network launched a Whale Report Alert System, which uses real-time reports made by trusted observers in the WhaleReport app to alert pilots of nearby commercial vessels of the whale presence (Ocean Wise 2024). This alert system enables the pilot to take adaptive mitigation such as slowing down or changing course (Ocean Wise 2024). A limitation of the citizen science whale reports is the location accuracy of real-time reports. Occasionally, a citizen scientist will forget to move the pin to the whale's location in their report, meaning that their estimated whale location is reported on land. Examples of these errors can be seen in the zoom panel of Figure 12. These real-time errors are later corrected in the quality assurance process by Spyhopper and adjusted according to the distance recorded by the observer and any relevant comments in the report.

Research is underway on the potential of using detection models for real-time reporting. Dr. Ruth Joy leads the Humans and Algorithms Listening to Orcas (HALLO) project which aims to train artificial intelligence systems to detect underwater whale vocalizations in real-time and develop a forecasting system to warn nearby vessels of whale presence and potential movement (Ruth Joy, SFU, personal communication, 2024). Like the Whale Report Alert System, this would allow vessel pilots to adapt by changing course or slowing down to mitigate noise pollution and prevent whale strikes.

4.3. Bias in the Alternative Methods of Vessel Detection

The F1 scores in each DVS zone indicate that both AIS and radar are most effective at detecting vessels in the furthest zone from the observer. For radar, this could be a result of the minimum distance required for detection. The specifications of the radar device indicate a minimum detection range of 20 m. However, the closest detection reported in the radar data was ~105 m, suggesting the portion of the ISZ directly in front of the radar had no vessel detections. This may have been due to obstructions on shore, the elevation of the mounting location, and the distance required to detect smaller vessels (Sung 2020).

Small vessels are more frequently detected in the ISZ, likely due to the legal restrictions on large commercial vessels restricting them within the shipping lanes (Baril 2022). The Boundary Pass shipping lane crosses the mid-zone and a small portion into the near-zone, and while this could not be investigated due to time constraints, it may have caused some of the true positive detections in this zone. This indicates radar may be more effective for tracking vessels in management zones that occur or extend further offshore and when larger vessels are the target.

A known limitation of using AIS to detect vessels in sanctuary zones in areas frequented by small recreational vessels is that the Canadian *Navigation Safety Regulations* (2020) do not require them all to carry an AIS device. A study on small recreational vessel activity near the Saturna Island ISZ found that 77% of small recreational vessels observed in the study area did not have an AIS device and thus would be missed by this detection method (Murphy et al., 2023). Due to time constraints, the proportion of each vessel type missed by AIS was not calculated in this study. This could, however, be an explanation for some the total 58% of vessels that AIS did not detect.

Radar does not have this limitation and can detect vessels regardless of AIS device installation. The requirement of AIS devices by the Canadian *Navigation Safety Regulations* (2020) to be implemented on all large commercial vessels will likely cause more true positive detections where Boundary Pass and the mid-zone intersect. AIS transmission is usually stronger for large commercial vessels as they must have a Class A device, which can be picked up by the receiver antennae more readily than the Class B devices found on small vessels (Last et al. 2015). Therefore, AIS may be best suited for monitoring management measures targeting large commercial vessels, such as the ECHO program's volunteer slowdown, or ecotourism vessels, which are also required to have an AIS device by the *Navigation Safety Regulations* (2020).

While reviewing the false positive detections by radar in GIS mapping software, it was observed that there were several vessel detections whose travel patterns appeared inconsistent with other vessel travel patterns. Rather than continuing in a line where the travel direction could be followed (like most true positive vessel detections), these detections bounced around within the same area. It is possible that these were caused by a vessel not under power and drifting within the current. However, it is unlikely that the primary observer would miss a relatively stationary vessel within the field of view. Observations by the primary observers and Saturna Island residents indicate that a large commercial vessel transiting quickly or close to shore can cause a lot of wave action within the field of view of this study (Maureen Welton, SIMRES, personal communication, 2023). Clutter, or the presence of an unwanted signal return, can be the source of false positive detections made by radar. Compared to terrestrial environments, radar experiences more large clutter signals when aimed at marine environments because of the sea wave action (Zainuddin et al. 2019). This was especially prevalent in radar systems detecting relatively small vessels (Zainuddin et al. 2019). This is often a consideration when developing the detection algorithm and is a possible source of false positive detections made by radar (Zainuddin et al. 2019). The literature reviewed for this study did not reveal any similar limitations in AIS detections caused by wave.

Radar was more reliable for detecting vessels in the near zone than AIS. Both AIS and radar can be limited by environmental conditions and nearby obstacles or obstructions and this could explain discrepancies in both methods (Last et al. 2015; Sung 2020). In a previous study, radar was considered the gold standard for a test on AIS track accuracy due to the increased frequency of track reports and radar's higher accuracy for reporting location (Jankowski et al. 2021). The authors suggest that AIS has more errors when reporting location (Jankowski et al. 2021). As the area of the near zone is relatively small, it is possible that false positives and false negatives in this study were due to the AIS location being reported in another zone. Unlike radar, which had many tracks to follow for location confirmation, AIS pings less frequently and is more challenging to track. Therefore, radar may have outperformed AIS in the near zone due to higher location accuracy.

4.4. Potential to Assist in the Enforcement of Vessel Management Measures

The AIS receiver antenna run by Quayle Consulting Ltd. is set up to automatically report a vessel in the ISZ to the SGIWSN, who then submit the vessel reports to the relevant authorities. AIS provides valuable information for enforcement, such as vessel identification information that allows warning letters or fines to be sent to the pilot. There is no current equivalent system set up with the radar device. The data provided by the current radar system does not include information that could identify the individual vessels or their pilots. Therefore, this system is more suitable for research on compliance and vessel presence than enforcement. Vessel identification could be accomplished through available radar devices that also contain a high-quality camera that could clearly capture the registration numbers of the vessels. Primary observers for this study, including the author, noted that while many registration numbers could be captured with a camera, there were some cases where the registration number was not legible or was obstructed from view. As such, it is possible that a camera associated with the radar system may not be able to capture a clear image of the registration number. Despite this limitation, the use of a combined system could still allow for more vessels without AIS to be monitored and reported without the presence of a human observer.

In addition to the use of ISZs to mitigate impacts of vessels on SRKWs in their Salish Sea critical habitat, there is an interim speed-restricted zone at Swiftsure Bank and the Vancouver Port Authority ECHO program's voluntary commercial vessel slowdown (Government of Canada 2023). In 2023, a team of researchers at SFU analyzed the current management measures surrounding Saturna Island and provided recommendations to be considered for implementation in 2024 (Murphy et al. 2023). They recommended that the current ISZ should be maintained, and an interim speed-restricted zone should be implemented in Tumbo Channel due to the potential for decreased noise source levels received by SRKWs in areas of high to moderate sighting density (Murphy et al., 2023). The Government of Canada implemented this recommendation in their management measures for 2024 and 2025 but adjusted the speed-restricted zone in Tumbo Channel to be voluntary (Government of Canada 2024).

AIS actively reports the speed of a vessel in real-time at each pinged location which is a valuable tool for insights on speed-related vessel information. The radar data for the system used in this study also includes a speed with each detection of a vessel. However, this value is calculated by the software based on the vessel's entire detected path and does not necessarily reflect its speed at the time of an individual detection. Therefore, it would not report a difference in speed of a vessel outside of a speed-restricted zone versus within the zone and may not be a good representation of a vessel's speed at specific locations and times. This information could be helpful for comparison where a custom detection range of a radar device could be set include the area within the speed-restricted zone and another nearby radar with a custom detection range to include the area outside the speed-restricted zone.

4.5. Integrating Indigenous Science and Traditional Ecological Knowledge

At the time of this study, detection methods were implemented and run without guidance from Indigenous Science or Traditional Ecological Knowledge. This represents a substantial shortfall and limitation, as these valuable types of knowledge are rooted in long-term lived experiences that cannot be accurately replicated through Western scientific methods (Biedenweg et al. 2023). The Salish Sea is home to Indigenous peoples from more than 40 First Nations who speak the many different Coast Salish languages and are the traditional stewards of the land since time immemorial (Miller 2011; Efford et al. 2023). Despite evidence showing that including such longstanding ecological knowledge and the well-being of the communities who hold it is crucial for well-informed conservation decision-making, it has historically been disregarded and continues to be overlooked (Pilbeam et al. 2019; Wheeler & Root-Bernstein 2020; Biedenweg et al. 2023).

Efforts to address this ongoing ignorance have commenced through partnerships between the organizations implementing this study's methods and local First Nations' marine programs. Dr. Joy and David Dick have initiated a collaboration between SFU and their team, the QENTOL, YEN W̱SÁNEĆ Marine Guardians, who were consulted regarding the placement of new hydrophones for the HALLO project (personal communications, Ruth Joy, SFU, 2024). The QENTOL, YEN W̱SÁNEĆ Marine Guardians possess valuable Traditional Ecological Knowledge on local whale activity, in addition to their comprehensive data that is actively gathered through frequent surveillance and Indigenous Science initiatives in the Salish Sea (QENTOL, YEN W̱SÁNEĆ Marine Guardians 2024). The partnership between SFU and the QENTOL, YEN / W̱SÁNEĆ Marine Guardians aims to integrate Indigenous science to enhance the efficiency of whale detection methods and research, with the expectation that the Government of Canada will begin to recognize and acknowledge the significance of Indigenous science (David Dick, QENTOL, YEN W̱SÁNEĆ Marine Guardians, personal communication, 2024).

In 2024, the W̱SÁNEĆ Marine Guardians were also added to the SGIWSN Discord communication network to facilitate the sharing of whale sighting information between organizations. Since knowledge and lived experiences vary among First Nations, it is recommended to continue fostering collaborations throughout the Salish Sea with a diverse range of First Nations invested in whale recovery (Wheeler & Root-Bernstein 2020). Like the methods of whale detection in this study, collaboration through consilience can provide us with increasingly comprehensive knowledge to bolster whale conservation and recovery efforts (Wheeler & Root-Bernstein 2020; Biedenweg et al. 2023).

4.6. Limitations of the Gold Standard Methods

No known whale or vessel detection method can be considered perfect. This study faced limitations with the gold standard method. The LBCO surveys had a different field of view than the alternative methods and were biased towards daylight and weather conditions suitable for human observation. Aspects of the LBCO surveys, such as accuracy of location reporting, were also limited by weather conditions ideal for theodolite use and calibration. The theodolite cannot be calibrated to the horizontal and vertical reference points when visibility is too low. It is also not waterproof and could not be used in heavy precipitation without an appropriate cover. This study did not have access to a proper cover but is recommended for future studies.

A notable limitation of this study is the potential influence of human error in the gold standard methods. Both gold standard methods used in this study are human-based and include human error. These errors could influence the performance metric calculations, as the apparent accuracy of a classification can be different from the truth when an imperfect gold standard method is used (Foody 2023). Human-based observation studies can result in observer fatigue, resulting in possible missed detections or details, which is not a limitation of artificial intelligence detection technology (Richter et al., 2023). A likely limitation involving human error specific to the DVS in this study was challenges with measuring distance for increasingly distant small vessels using the laser range finder. A few small vessels detected near the furthest border of the mid-range zone were not picked up by the laser range finder due to complications with hand-steadiness and the relative size of the target vessel. As a result, the primary observer estimated whether it was in or out of the mid-zone based on visual landmarks such as buoys and islands. Many of the false positives by the alternative methods of vessel detections were of vessels near the far border of the mid-zone, which this could explain. This could also explain some of the false negatives in the mid-zone and the reason for a lower NPV in these zones compared to the others.

4.7. Methods of Detection in the Absence of Humans

The on-effort hours of the gold standard methods were always within daylight hours, between 06:00-19:00 PDT for the LBCO surveys and 09:00-16:00 PDT for the DVS surveys. Therefore, the results of this study are not representative of the performance of the alternative method in detecting whales or vessels at night. However, from the data collected for this study, it was observed that there were numerous detections by all the alternative methods of detection that rely on technology and models throughout all hours of the night. In contrast, citizen scientists reported a noticeable lack of reports at hours after sunset compared to daylight. This is expected as these technologies are intended to run 24/7, including late and dark nighttime hours, while humans would be asleep or have limited visibility. Thermal imaging detection rates are similar between night and day, however, detection performance at night exceeds performance during the day due to the lack of glare and reflections caused by the sunlight (Zitterbart et al. 2013). Additionally, vessel activity is decreased at night, which means there is less vessel noise pollution to mask killer whale vocalizations in the hydrophone recordings (Ogawa & Kimura 2023; Richter et al. 2023). The DANN model could be tested at night to see if the decrease in vessel noise in the hydrophone recordings has any influence on the F1 score.

With all of this considered, methods of detection that rely on the presence of a human observer would likely not be the best method for detection at night. Where possible, deployment of both thermal imaging and hydrophones together in an area can provide complementary detections where the other method may overlook, as hydrophone detections rely on frequent vocalizations and thermal imaging detections rely on surfacing within the narrow field of view (Richter et al. 2023). In vessel-exclusion sanctuary zones, radar detection similar to the one used in this study can provide insight into the number of infractions. As it does not report vessel-specific details, it may be unable to assist in automatic reporting to facilitate fines and warning letters. AIS does report information specific to the vessel, including the Maritime Mobile Services Identity, name, and type of vessel.

Similarly, there are knowledge gaps in whale behaviour and habitat use in remote areas with limited human observers, such as far offshore in the open ocean, and when weather conditions are harsh, such as winter in British Columbia (Pilkington et al. 2023; COSEWIC 2008). Offshore killer whales are another threatened ecotype found off the outer coast of British Columbia that does not frequently enter the Salish Sea(Ford et al. 2000; COSEWIC 2008). As this ecotype spends much of their time far from the coastlines, they are less frequently observed and less is known about their population and range extent compared to the SRKW and BKW ecotypes (COSEWIC 2008). A recent study aimed at filling knowledge gaps associated with remote, inaccessible sites and during the winter season, has encouraged using multiple hydrophones recommends particular configurations based on the target ecotype (Pilkington et al. 2023). As previously discussed, work like the HALLO project has made strides toward artificial intelligence that sends an automatic alert of a whale’s location and forecasts its possible movements to nearby vessels without the assistance of a human observer. The location information collected through such projects can also be used in further research on automated detections of whale activity. Continued research on methods of whale detection at times and places with limited availability of human observers is recommended to enhance our understanding of their habitat use and better inform management strategies for their conservation and recovery.

Chapter 5. Conclusion

This study aimed to compare alternative whale and vessel detection methods in and around the Saturna Island ISZ throughout the summer of 2023. Methods of detection were compared on their performance utilizing machine-learning performance metrics. Findings revealed that whale detection through methods using trained citizen scientist observers outperformed methods that rely on observations by technology and modelled detection algorithms. AIS and radar had similar performance for detecting vessels overall but differed in their ability to detect vessels in the study zones at varying ranges from shore. AIS consistently had higher precision, while radar had higher recall.

Insights from this study led to several recommendations to inform adaptive management for whale recovery management measures. Where possible, integrating multiple detection methods provides an opportunity for increased reliability and accuracy in an area, as they have different strengths and limitations that can be counteracted through their collaboration. While we are not quite ready to leave whale and vessel monitoring to the AI-based systems, they have potential value for increased performance, real-time reporting, and should be a focus of continued research. This is especially true in scenarios where human observation is difficult or lacking. Increased collaboration with Indigenous communities and implementation of Indigenous science and TEK into whale research methods and decision-making is critical for well-informed conservation efforts.

References

Baril K (2022) Marine Restoration in Boundary Pass: Evaluating Management Strategies for Conserving Canada's Coastal Whale Populations. Simon Fraser University

Barrett-Lennard LG, Ellis GM (2001) Population Structure and Genetic Variability in Northeastern Pacific Killer Whales: Towards an Assessment of Population Viability. Fisheries and Oceans Canada

Biedenweg K, Anderson L, Hatfield S, Hollender R, Kintner L, Trimbach DJ (2023) Seeking consilience: Traditional ecological knowledge and Western social science contributions to orca conservation knowledge. *Journal for Nature Conservation* 72:126364

Center for Whale Research (2023) Southern Resident Orca (SRKW) Population. https://www.whaleresearch.com/orca-population (accessed 9 November 2023)

Chicco D, Tötsch N, Jurman G (2021) The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining* 14:13

Clapham PJ (2018) Humpback Whale: *Megaptera novaeangliae*. In: Encyclopedia of Marine Mammals (Third Edition). Würsig, B, Thewissen, JGM, & Kovacs, KM, editors. Academic Press pp. 489–492.

COSEWIC (2022) COSEWIC assessment and status report on the Humpback Whale *Megaptera novaeangliae kuzira*, North Pacific population,in Canada. Committee on the Status of Endangered Wildlife in Canada., Ottawa

COSEWIC (2008) COSEWIC Assessment and Update Status Report on the Killer Whale Orcinus orca Southern Resident population Northern Resident population West

Coast Transient population Offshore population Northwest Atlantic / Eastern Arctic population in Canada.

Department of Fisheries and Oceans Canada (2017) Assessing the risk of ship strikes to humpback (*Megaptera novaeangliae*) and fin (*Balaenoptera physalus*) whales off the west coast of Vancouver Island, Canada.

Di Clemente J, Christiansen F, Pirotta E, Steckler D, Wahlberg M, Pearson HC (2018) Effects of whale watching on the activity budgets of humpback whales, *Megaptera novaeangliae* (Borowski, 1781), on a feeding ground. *Aquatic Conservation: Marine and Freshwater Ecosystems* 28:810–820

Efford M, Taft S, Morin J, George Micheal, George Michelle, Cavers H, Hilsden J, Paskulin L, Loewen D, Zhu J, Christensen V, Speller C (2023) Archaeology demonstrates sustainable ancestral Coast Salish salmon stewardship over thousands of years. *PLOS ONE* 18:e0289797

Foody GM (2023) Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient. *PLOS ONE* 18:e0291908

Ford JK, Ellis GM, Barrett-Lennard LG, Morton AB, Palm RS, Balcomb Iii KC (1998) Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Canadian Journal of Zoology* 76:1456–1471

Ford JKB (1991) Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Canadian Journal of Zoology* 69:1454–1483

Ford JKB, Deecke VB, Pilsington JF (2023) Dialects of Killer Whales in Coastal Northeastern Pacific Waters: A Catalogue of Call Types. Orca Sound Call Catalogue and Annotations https://orca.research.sfu.ca/call-library/?f=eyJwb3B1bGF0aW9uIjpbIlNSS1ciXSwiU1JLVyI6eyJjbGFuIjpbIkoiXSwicG9kIjpbIkoiLCJLIiwiTCJdfX0%3D&p=1&s=call_type&sa=as&ps=24 (accessed 5 April 2024)

Ford JKB, Ellis GM, Balcomb KC (2000) Killer Whales: The Natural History and Genealogy of Orcinus Orca in British Columbia and Washington*. UBC Press*

Frazer LN, Mercado E (2000) A sonar model for humpback whale song. *IEEE Journal of Oceanic Engineering* 25:160–182

Gassmann M, Elizabeth Henderson E, Wiggins SM, Roch MA, Hildebrand JA (2013) Offshore killer whale tracking using multiple hydrophone arrays. *The Journal of the Acoustical Society of America* 134:3513–3521

Gheibi A (2022) Drivers of Humpback Whale Movement in Boundary Pass, British Columbia. Simon Fraser University

Gillespie A (2016) Vulnerability and Response to the Risk of International Shipping: The Case of the Salish Sea. Review of European, Comparative & International Environmental Law 25:317–332

Government of Canada (2019a) A *Species at Risk Act* Section 11 Conservation Agreement to Support the Recovery of the Southern Resident Killer Whale.

Government of Canada (2019b) Government of Canada announces enhanced measures for protecting British Columbia's Southern Resident killer whales. https://www.canada.ca/en/fisheries-oceans/news/2019/05/government-of-canada-announces-enhanced-measures-for-protecting-british-columbias-southern-resident-killer-whales.html (accessed 15 December 2023)

Government of Canada (2020) Summary of input provided on management measures to address key threats to Southern Resident killer whales | Pacific Region | Fisheries and Oceans Canada. https://www.pac.dfo-mpo.gc.ca/consultation/fm-gp/srkw-eprs/index-eng.html (accessed 15 December 2023)

Government of Canada (2023) Management measures to protect Southern Resident killer whales | Pacific Region | Fisheries and Oceans Canada. https://www.pac.dfo-mpo.gc.ca/fm-gp/mammals-mammiferes/whales-baleines/srkw-measures-mesures-ers-eng.html (accessed 9 November 2023)

Government of Canada (2024) Input on recovery of Southern Resident killer whale potential management measures for 2024-25 | Pacific Region | Fisheries and Oceans Canada. https://www.pac.dfo-mpo.gc.ca/consultation/fm-gp/srkw-eprs/2024-srkw-management-gestion-ers-eng.html (accessed 4 April 2024)

Graber J, Thomson J, Polagye B, Jessup A (2011) Land-based infrared imagery for marine mammal detection. University of Washington

Herman LM, Pack AA, Spitz SS, Herman EYK, Rose K, Hakala S, Deakos MH (2013) Humpback whale song: who sings? *Behavioral Ecology and Sociobiology* 67:1653–1663

Hripcsak G, Rothschild AS (2005) Agreement, the F-Measure, and Reliability in Information Retrieval. *Journal of the American Medical Informatics Association* 12:296–298

Jankowski D, Lamm A, Hahn A (2021) Determination of AIS Position Accuracy and Evaluation of Reconstruction Methods for Maritime Observation Data. *IFAC-PapersOnLine* Vol. 54 pp. 97–104.

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

Krahn MM, Wade PR, Kalinowski ST, Dahlheim ME, Taylor BL, Hanson MB, Ylitalo GM, Angliss RP, Stein JE, Waples RS (2002) Status Review of southern resident killer whales (*Orcinus orca*) under the Endangered Species Act. Northwest Fisheries Science Centre

Last P, Hering-Bertram M, Linsen L (2015) How automatic identification system (AIS) antenna setup affects AIS signal quality. *Ocean Engineering* 100:83–89

Le Baron N, McWhinnie LH, Frouin-Muoy H (2019) Boundary Pass 2019 land-based cetacean observations: ECHO 2019 slowdown evaluation.

Leighton T, Finfer D, Grover E, White P An acoustical hypothesis for the spiral bubble nets of humpback whales, and the implications for whale feeding.

Lusseau D, Bain DE, Williams R, Smith JC (2009) Vessel traffic disrupts the foraging behavior of southern resident killer whales Orcinus orca. *Endangered Species Research* 6:211–221

McWhinnie LH, O’Hara PD, 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 & Coastal Management* 200:105479

Miller BG (2011) Be of Good Mind: Essays on the Coast Salish. UBC Press

Morin PA, McCarthy ML, Fung CW, Durban JW, Parsons KM, Perrin WF, Taylor BL, Jefferson TA, Archer FI (2024) Revised taxonomy of eastern North Pacific killer whales (Orcinus orca): Bigg's and resident ecotypes deserve species status. *Royal Society Open Science* 11:231368

Murphy O, Fairfield Checko R, Young M, Laturnus L, Lin T-W, Padovese B, Joy R (2023) Analysis of Southern Resident Killer Whale Management Measures Surrounding Saturna Island. Simon Fraser University https://www.sfu.ca/~rjoy/SaturnaAnalysisOfMeasures_SFU.pdf (accessed 2 April 2024)

Myers HJ, Olsen DW, Matkin CO, Horstmann LA, Konar B (2021) Passive acoustic monitoring of killer whales (Orcinus orca) reveals year-round distribution and residency patterns in the Gulf of Alaska. *Scientific Reports* 11:20284

Ocean Wise (2024) WhaleReport Alert System. Ocean Wise https://ocean.org/whales/wras/ (accessed 3 April 2024)

Ogawa M, Kimura SS (2023) Variations in echolocation click characteristics of finless porpoise in response to day/night and absence/presence of vessel noise. *PLOS ONE* 18:e0288513

Palmer MA, Zedler JB, Falk DA, eds. (2016) Foundations of Restoration Ecology. Island Press/Center for Resource Economics, Washington, DC

Payne MR, Barange M, Cheung WWL, MacKenzie BR, Batchelder HP, Cormon X, Eddy TD, Fernandes JA, Hollowed AB, Jones MC, Link JS, Neubauer P, Ortiz I, Queirós AM, Paula JR (2016) Uncertainties in projecting climate-change impacts in marine ecosystems. *ICES Journal of Marine Science* 73:1272–1282

Payne RS, McVay S (1971) Songs of Humpback Whales. *Science* 173:585–597

Pilbeam V, van Kerkhoff L, Weir T (2019) Conservation Decision-Making in Palau: An Example of the Parallel Working of Scientific and Traditional Ecological Knowledge. *Environmental Management* 64:564–579

Pilkington JF, Stredulinsky EH, Gavrilchuk K, Thornton SJ, Ford JKB, Doniol-Valcroze T (2023) Patterns of winter occurrence of three sympatric killer whale populations off eastern Vancouver Island, Canada, based on passive acoustic monitoring. *Frontiers in Marine Science* 10

Prior MK, Meless O, Bittner P, Sugioka H (2011) Long-Range Detection and Location of Shallow Underwater Explosions Using Deep-Sound-Channel Hydrophones. IEEE *Journal of Oceanic Engineering* 36:703–715

QENTOL, YEN W̱SÁNEĆ Marine Guardians (2024) Annual Report 2023-2024. Brentwood Bay, BC.

Quayle LS (2021) Assessing Vessel Related Threats on Whales in Boundary Pass. Simon Fraser University

Richter S, Yurk H, Winterl A, Chmelnitsky E, Serra N, O'Hara PD, Zitterbart D (2023) Coastal marine mammal conservation using thermal imaging-based detection systems. *bioRxiv* 2023.08.25.554754

Roman J, Altman I, Dunphy-Daly MM, Campbell C, Jasny M, Read AJ (2013) The Marine Mammal Protection Act at 40: status, recovery, and future of U.S. marine mammals. *Annals of the New York Academy of Sciences* 1286:29–49

Ryan JP, Cline DE, Joseph JE, Margolina T, Santora JA, Kudela RM, Chavez FP, Pennington JT, Wahl C, Michisaki R, Benoit-Bird K, Forney KA, Stimpert AK, DeVogelaere A, Black N, Fischer M (2019) Humpback whale song occurrence reflects ecosystem variability in feeding and migratory habitat of the northeast Pacific. *PLOS ONE* 14:e0222456

Shields MW (2023) 2018–2022 Southern Resident killer whale presence in the Salish Sea: continued shifts in habitat usage. *PeerJ* 11:e15635

Shields MW, Hysong-Shimazu S, Shields JC, Woodruff J (2018) Increased presence of mammal-eating killer whales in the Salish Sea with implications for predator-prey dynamics*. PeerJ* 6:e6062

Stimpert AK, Au WWL, Parks SE, Hurst T, Wiley DN (2011) Common humpback whale (*Megaptera novaeangliae*) sound types for passive acoustic monitoring. *The Journal of the Acoustical Society of America* 129:476–482

Sung JJ (2020) A Study on the Radar Operational and Technical Performance Requirements for Vessel Traffic Service. *Journal of Navigation and Port Research* 44:110–118

Wheeler H, Root-Bernstein M (2020) Informing decision-making with Indigenous and local knowledge and science*. Journal of Applied Ecology* 57: 1634-1643

Wintle BA (2007) Adaptive Management, Population Modeling and Uncertainty Analysis for Assessing the Impacts of Noise on Cetacean Populations. International *Journal of Comparative Psychology* 20

Zainuddin S, Pasya I, Rashid NEA, Zaiaami N, Maryam A (2019) Maritime Radar: A Review on Techniques for Small Vessels Detection. 14

Zitterbart DP, Kindermann L, Burkhardt E, Boebel O (2013) Automatic Round-the-Clock Detection of Whales for Mitigation from Underwater Noise Impacts. *PLOS ONE* 8:e71217

Appendix A. Detection System Outages

Table A1. Start dates and times of missing 5-minute recordings indicating temporary system outages by the hydrophones during the study period. Consecutive missing recordings are listed together.

Hydrophone	Dates (mm/dd)	Times (PDT)
East Point	06/03	03:20
	06/04	10:35
	06/10	10:35
	06/11	17:05
	06/20	06:50
	06/28	20:35
	07/07	10:20
	07/16	00:05
	07/24	13:50
	08/09	04:40-16:55
Monarch Head	08/09	21:30
	8/12	09:50
	8/14	03:25
	8/14	11:30
	8/16	17:05
	8/17-8/18	22:30-00:05
	8/21	00:20
	8/25	23:20
	08/26	13:50
	8/27	04:25
	8/27	07:20-16:55
	08/28	17:00-18:35
	08/28	18:55-19:00
	08/29	13:35
	08/29	14:00
	08/29	14:15-14:20
	08/30	09:40
	08/30	11:00
	08/30	12:00
	08/30	12:05
	08/30	14:15
	08/30	14:25

Table A2. Start and end dates and times of potential system outages determined by a break in detections greater than 24 hours for Sensor A during the study period (June 1 – August 31, 2023).

Start Date (mm/dd)	Start Time (PDT)	End Date (mm/dd)	End Time (PDT)
06/08	21:00	06/11	01:00

Appendix B. Confusion Matrices

Table B1. Confusion matrix for whale event detections by citizen scientists of the Southern Gulf Islands Whale Sighting Network (SGIWSN) against the LBCO survey (the Gold standard (most reliable) method of detecting whales).

	Whale event detected by the LBCO survey method	Whale event not detected by the LBCO survey method
Whale event detected by the citizen scientists	40 (TP)	0 (FP)
Whale event not detected by the citizen scientists	11 (FN)	- (TN)

Table B2. Confusion matrix for whale species identification by the Southern Gulf Islands Whale Sighting Network (SGIWSN) with high certainty.

	Whale species identified by the LBCO surveys	Whale species not identified by the LBCO surveys
Whale species identified by the citizen scientists	104 (TP)	0 (FP)
Whale species not identified by the citizen scientists	0 (FN)	- (TN)

Table B3. Confusion matrix for killer whale ecotype identification with high certainty by the Southern Gulf Islands Whale Sighting Network (SGIWSN).

	Ecotype identified by the LBCO surveys	Ecotype not identified by the LBCO surveys
Ecotype identified by the citizen scientists	72 (TP)	1 (FP)
Ecotype not identified by the citizen scientists	2 (FN)	- (TN)

Table B4. Confusion matrix for whale event detections by the Infrared temperature sensor through thermal imaging.

	Whale event detected by the LBCO surveys	Whale event not detected by the LBCO surveys
Whale event detected by thermal Imaging	9 (TP)	0 (FP)
Whale event not detected by thermal imaging	42 (FN)	- (TN)

Table B5. Confusion matrix for whale event detections by the infrared temperature sensor through thermal imaging that were confirmed to be in Boundary Pass (BP) by the gold standard method.

	Whale event detected in BP by the LBCO surveys	Whale event not detected in BP by the LBCO surveys
Whale event detected by thermal imaging	9 (TP)	0 (FP)
Whale event not detected by thermal imaging	28 (FN)	- (TN)

Table B8. Confusion matrix for killer whale (KW) detections by the DANN hydrophone killer whale detection model.

	KW event detected by the LBCO surveys	KW event not detected by the LBCO surveys
KW event detected with by DANN model	27 (TP)	71 (FP)
KW event not detected by DANN model	1* (FN)	- (TN)

* 1 KW event coincided with a hydrophone system outage and was excluded from analysis for this model.

Table B9. Confusion matrix for vessel detections in the ISZ range zone by radar.

	Vessel detected by the DVS	Vessel not detected by the DVS
Vessel detected by radar	2 (TP)	7 (FP)
Vessel not detected by radar	8 (FN)	564 (TN)

Table B10. Confusion matrix for vessel detections in the near range zone by radar.

	Vessel detected by the DVS	Vessel not detected by the DVS
Vessel detected by radar	22 (TP)	12 (FP)
Vessel not detected by radar	28 (FN)	528 (TN)

Table B11. Confusion matrix for vessel detections in the mid-range zone by radar.

	Vessel detected by the DVS	Whale event not detected by the DVS
Vessel detected by radar	120 (TP)	58 (FP)
Vessel not detected by radar	113 (FN)	373 (TN)

Table B12. Confusion matrix for vessel detections in the ISZ range zone by AIS.

	Vessel detected by the DVS	Whale event not detected by the DVS
Vessel detected by radar	2 (TP)	0 (FP)
Vessel not detected by radar	8 (FN)	571 (TN)

Table B13. Confusion matrix for vessel detections in the near range zone by AIS.

	Vessel detected by the DVS	Whale event not detected by the DVS
Vessel detected by radar	5 (TP)	4 (FP)
Vessel not detected by radar	45 (FN)	531 (TN)

Table B14. Confusion matrix for vessel detections in the mid-range zone by AIS.

	Vessel detected by the DVS	Whale event not detected by the DVS
Vessel detected by radar	115 (TP)	20 (FP)
Vessel not detected by radar	118 (FN)	383 (TN)

Appendix C. Supplementary Data and Methods

Table C1. Number of 5-minute recordings detected by the killer whale hydrophone detection model with ≥ 10 peaks during the study period (June 1 – Aug 31, 2023) from the East Point and Monarch Head hydrophones.

Hydrophone	On-Effort	Total
East Point (Jun 1 – Aug 8)	1039	5852
Monarch Head (Aug 9 – Aug 31)	531	1900
Combined (Jun 1 – Aug 31)	1570	7752

Table C2. Distance to the border of the designated ISZ range from the primary observation site for the DVS when aimed at the landmark as a visual reference. "Left" and "Right" are relative to the view of the primary observer at the DVS observation site facing out into Boundary Pass.

Landmark	Distance
Left end Sucia Islands	370 m
Right end of Sucia Islands	345 m
Left end of Orcas Island	315 m
Middle of Orcas Island	310 m
Right end of Orcas Island	325 m
Left end of Waldron Island	345 m
Middle of Skipjack Island	390 m
Right end of Waldron Island	550 m
Left end of Johns Island	730 m
Left end of Stuart Island	980 m

Methods for Tracking ISZ Vessel Infractions with the Theodolite

During the LBCO surveys, vessels without AIS were also tracked by the primary observers using the theodolite and Mysticetus software and ISZ infractions were reported on behalf of the SGIWSN. This was not included in the comparison of vessel detection methods due to the misalignment between the timing of the radar installation and the LBCO surveys, the increased difficulties between the differing fields of view when compared to the differing field of view of the methods of whale detection, and the priority of theodolite tracking given to whales once detected over vessels during a whale event. If possible, vessels were tracked with the theodolite during a whale event if there was a suspected ISZ infraction or marine mammal violation. There is currently no level platform available for the theodolite that provides a similar field of view of the radar, hence the need for a range finder during the DVS. The vessel's location, distance from the vantage point, approximate speed, type of vessel, and activity were recorded. Optimally, 4 data points were collected on the theodolite for each vessel, but at some times was not possible due to speed, equipment malfunction, etc. In those cases, as many points as possible were taken.

Table C3. Supplementary details of whale events detected during the LBCO surveys. The "Near Zone" includes the Saturna Island ISZ to the edge of the Boundary Pass shipping lane. The Far Zone is on the far side of Boundary Pass from the primary observers, near the San Juan Islands.

Date (2023)	Event #	Start time	End time	Species/ Ecotype	Est. # of individuals	Travel Direction	Zone(s)	Behaviour	Vessels Present	Notes
06/01	1	09:49	10:30	SRKW	15-20	S/SW/W	Near Zone, BP, SoG	Travel-Forage-Social	2	J-pod
06/02	2	10:20	11:02	SRKW	10-20	N	Near Zone, BP, SoG	Travel-Forage-Social	3	J-pod
06/03	3	14:13	14:40	Humpback	1	N	SoG, Near Zone	Travel	4	
06/04	4	14:08	14:30	Humpback	1	N	Near Zone, SoG	Travel	0	Graphite
06/04	5	14:32	15:01	BKW	1	NE	Far Zone	Forage-Travel	2	Cooper
06/05	6	07:51	08:11	SRKW	15-20	W	Near Zone, SoG, BP	Travel-Social	0	J-pod
06/05	7	08 :14	08:19	Humpback	1	W	Near Zone	Travel	0	
06/05	8	10:14	10:24	SRKW	4-5	W	Near Zone, SoG, BP	Travel	0	J-pod
06/07	9	10:52	11:00	BKW	3	E	Near Zone, SoG	Travel-Social	0	
06/15	10	13:02	13:18	BKW	4	N	Near Zone, SoG	Travel	7	
06/26	11	13:33	14:13	SRKW	25	S/SW/W	Near Zone, SoG, BP	Travel-Social	3	J-pod
06/29	12	12:38	14:15	Humpback	1	E/N/W/S	Near Zone	Forage-Rest	3	Orion
06/30	13	15:09	16:07	SRKW	25-30	S/SW/W	Near Zone, SoG, BP	Travel-Forage-Social	4	J & L-pod
07/03	14	13:17	13:42	Humpback	1	E/N	Near Zone, SoG	Travel	1	Orion
07/04	15	14:31	15:01	SRKW	15-20	S/SW/W	Near Zone, SoG, BP	Travel-Social	2	J-pod
07/04	16	16:05	17:09	Humpback	1	E/N/W	Near Zone	Travel	0	Orion
07/05	17	15:20	15:37	Humpback	1	NE	Near Zone, SoG	Travel	1	Orion
07/14	18	09:36	09:47	BKW	4	N/NE	Far Zone, BP	Travel	3	
07/14	19	09:52	10:09	BKW	6	NE	Near Zone, BP	Travel	5	
07/14	20	11:45	12:24	BKW	3-4	W	Far Zone	Travel	10	
07/16	21	10:07	10:09	BKW	3	N/NE/NW	Near Zone	Travel-Rest	0	
07/16	22	11:33	11:33	BKW	3-5	SW	Near Zone	Travel	0	
07/16	23	15:56	17:17	BKW	4	N/NW/W	Far Zone, BP, SoG	Travel	10	
07/20	24	16:21	16:31	Humpback	2	E/NE/N	Near Zone, BP, SoG	Travel	5	Ghost & Calf
07/21	25	13:33	14:07	BKW	4-5	NE/E/SE	SoG, Far Zone	Travel	3	
07/22	26	13:00	14:32	BKW	6-7	NE/E/SW	BP, Far Zone	Travel-Forage-Social	11	T065Bs
07/25	27	12:06	12:12	BKW	3-5	E/S	Near Zone	Travel	0	T065Bs

07/28	28	14:48	15:08	BKW	4	NE/E/W/S	Far Zone, BP, SoG	Travel-Forage	4	
07/31	29	12:57	13:12	Humpback	1	E/S	SoG	Travel	2	
08/04	30	09:42	10:52	Humpback	1	NE/E	BP	Travel-Forage	2	
08/04	31	14:22	14:42	Humpback	1	NE	BP	Travel	2	Zephyr
08/04	32	15:36	16:05	BKW	3-4	E	SoG	Travel	4	T037Bs
08/04	33	16:02	16:14	Humpback	1	N	BP	Social	7	Yogi
08/04	34	16:56	17:06	BKW	2	E/NE/N	BP	Travel	7	T137s
08/06	35	15:25	16:00	Humpback	1	E/S/N	Far Zone	Travel	6	
08/07	36	12:05	13:08	Humpback	1	NW/N	Far Zone	Travel	9	
08/08	37	12:08	13:16	Humpback	1	E/W	SoG	Travel-Forage	6	
08/09	38	10:42	10:53	BKW	4-6	SE/SW/W	TC	Forage	4	
08/09	39	12:36	12:52	BKW	6	W/S/SW	TC, Near Zone, BP	Forage	4	
08/10	40	15:46	16:09	Humpback	1	NE	Near Zone, Bp, SoG	Travel	7	Bond
08/12	41	14:27	14:29	Humpback	2	N	SoG	Travel-Social	1	
08/15	42	10:39	11:32	Humpback	1	N/NE/E/S	BP, SoG, Far Zone	Travel	3	
08/15	43	13:18	13:37	Humpback	1	SW	Far Zone	Travel	4	
08/15	44	13:59	14:23	BKW	7	NE, N	BP	Travel	5	
08/15	45	14:38	15:35	Humpback	1	N/NE	Far Zone	Travel	3	
08/16	46	09:14	11:37	Humpback	1	N/NE/E/SW	Near Zone, BP, Far Zone	Travel-Forage	3	Scratchy
08/16	47	12:30	12:57	BKW	4	NE/N	Near Zone	Travel	0	
08/17	48	16:27	16:31	BKW	4	NW/W	Near Zone	Travel	0	T057Bs
08/19	49	12:49	13:01	Humpback	1	NE	Near Zone, SoG	Travel	0	Raptor
08/22	50	16:02	16:23	BKW	4	NE/W	Near Zone, SoG	Travel	9	T037As
08/27	51	15:48	16:00	BKW	5	N/E/W/SW/SE	TC, Near Zone, BP	Travel-Forage-Social	4	T075Bs and T077D