CREEM
Centre for Research into Ecological and Environmental Modelling

Effectiveness of Navy lookout teams in detecting cetaceans

Final report on HDR Task Order 002/1000300001856

Authors:	C. Oedekoven and L. Thomas
Report Code:	CREEM-24289-1
Date:	22 March 2022

Please cite this report as: Oedekoven, C. and L. Thomas (2022) Effectiveness of Navy lookout teams in detecting cetaceans. Report number CREEM-24289-1. Provided to HDR Inc, 22 March 2022 (unpublished).

Document Control

Please consider this document as uncontrolled copy when printed

Rev.	Date.	Reason for Issue.	Prepared by	Checked by
1	10/Dec/2021	Draft sent to HDR for review	CO/LT	LT/CO
2	22/March/2022	Revised version sent to HDR	CO/LT	LT/CO

University of St Andrews

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
22-03-2022

2. REPORT TYPE
Monitoring report

3. DATES COVERED *(From - To)*
2010 - 2019

4. TITLE AND SUBTITLE
EFFECTIVENESS OF NAVY LOOKOUT TEAMS IN DETECTING CETACEANS

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Cornelia Oedekoven
Len Thomas

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Centre for Research into Ecological and Environmental Modelling, The University of St Andrews

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander, U.S.Pacific Fleet, 250 Makalapa Dr. Pearl Harbor, HI

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The United States (US) Navy uses lookouts (LOs) to detect objects in the water in the vicinity of ships. One class of object that LOs are trained to detect is marine mammals ; this forms an important component of the Navy’s procedures for marine mammal mitigation during training activities . As well as dedications of marine mammals by these LOs, detections of marine mammals may also be made by other members of the ship’s crew such as officers on the bridge (“watchstanders”) or sonar technicians, although acoustic detections require visual confirmation. We refer to these personnel together as the “lookout team” (LT). The primary goal of this project was to determine how effective LTs are at detecting marine mammals before they entered a defined set of mitigation ranges during mid-frequency active sonar training activities. These ranges were 200, 500 and 1,000 yards. A secondary goal was to compare this effectiveness with that of trained marine mammal observers (MMOs).

In collaboration with Navy environmental personnel, we developed a field protocol for at-sea experiments, where MMOs set up trials by locating marine mammals around Navy ships training with mid-frequency active sonar and determined whether these animals were detected by the LT. We also developed new analytical methods that allow estimation of the probability of animals approaching to within a specified mitigation range without being detected (probability of remaining undetected, PrU). These methods include a model for the surfacing pattern of animal pods , and for the range-dependent probability of detecting a pod each time it surfaces. Crucially, the methods allow us to account for the possibility that animal pods may remain undetected by both MMOs and the LT. The methods are flexible in allowing for

various patterns of animal surfacing and various experimental configurations (in terms of communication between MMO and LT positions and whether repeat surfacings of the same pod are recorded). They are, however, simplistic in assuming that there is no measurement error (in surfacing location, taxon designation and whether duplicate detections are correctly assigned), that pods only move in the vertical plane (i.e., there is negligible horizontal movement during the period when the pod is within observation range), and that the ship moves at constant speed in a straight line. We tested the new analytical methods using computer simulation and found they generally produce unbiased estimates when the model assumptions are met, although in some circumstances (including those in our at-sea study) it is not possible to estimate both detectability and surfacing pattern; in this situation if the parameters governing surfacing pattern are known then unbiased estimation of detectability, and hence PrU, is possible.

A total of 27 embarks were conducted between 2010 and 2019, mostly on destroyer class ships. These generated 716 valid sightings of animal pods. Each sighting consisted of one or more detection of a marine mammal pod by the MMO and/or LT positions; to be valid there had to be enough information recorded to derive a taxonomic code at the level needed for analysis (see below), pod size and, for each detection, a location (relative to the ship) and an observer position. There were no valid acoustic detections, and so all LT detections were generated by the LOs or watchstanders. Some species of small cetacean are known to approach ships and "bowride"; after discussion with Navy environmental personnel it was decided to exclude detections of pods observed during the sighting to engage in bowriding behavior. There were 46 such sightings, with first detections predominantly made at close ranges. After excluding these, 670 sightings remained.

Our data collection protocol asked MMOs to prioritize new sightings over repeated detections (resights) of an already-sighted school, and so resights were not recorded consistently. We therefore used analytical methods that require data only on the first detection of a pod by each position. Analysis at species level was not possible because of limited sample size, and because many sightings were not identified using a taxonomic code that refers to species, but instead to a higher taxonomic level such as "large whale" or "dolphin" . We therefore divided the data into four groups according to similarity in surfacing pattern and detectability: rorquals (i.e., large baleen whales), sperm whales, small cetaceans in small pods (6 or less) (SCSP) and small cetaceans in large pods (more than 6) (SCLP). We assumed the parameters governing surfacing pattern for each group were known, and we used values derived from the literature. For the sperm whale group, for which there were only two sightings, we used the detectability parameters estimated for rorquals. There were not enough detections of pinnipeds for us to estimate range-dependent probability of detection from the detection data and, unlike sperm whales, we elected not to use the estimated detectability parameters from one of the other groups; hence, our results only cover cetaceans.

Before undertaking the modelling we performed some exploratory analyses, including calculating a simple distance-specific index of effectiveness at 200, 500 and 1,000 yards (yds) for rorquals, SCSP and SCLP. For this analysis, we quantified LT effectiveness as the number of pods detected by the LT before they enter within the mitigation range divided by the total number of pods thought to have entered within the mitigation range (as estimated by the number seen by the LT or MMOs within a given distance of the ship's track). We speculate that this provides an upper bound on absolute effectiveness, because it does not take account of pods that pass through the mitigation zones undetected by either position. Estimated effectiveness was highest for rorquals: 0.35, 0.21 and 0.13 at 200, 500 and 1,000 yds for the LT and 0.74, 0.70 and 0.54 respectively for MMOs. It was lowest for SCSP: 0.03, 0.03 and 0.02 respectively at 200, 500 and 1,000 yds for the LT and 0.25, 0.29 and 0.14 respectively for MMOs. The estimates for SCLP were similar to SCSP for the LT but higher than SCSP for MMOs.

Results from the modelling analysis to obtain PrU are summarized graphically in the figure on the next page. For each group, we estimated PrU at 200, 500 and 1,000 yards (yds). Please note that, although the results are quoted at these ranges, all of the data from each taxonomic group (including data beyond 1,000 yds) was used in deriving the results with these models. For rorquals the estimated PrU at 200 yds for the LT was 0.80 (95% confidence interval (CI) 0.74-0.86), rising to 0.91 (95% CI 0.87-0.94) at 1000 yds. PrU is the complement of effectiveness, so estimated absolute effectiveness was 1-0.80=0.20 at 200 yds and 1-0.91=0.09 at 1,000 yds. As expected, these values are slightly lower than the simple distance-specific index of effectiveness quoted in the previous paragraph (and this pattern held true for all such comparisons). MMOs were estimated to be considerably better, with PrU at 200 yds of 0.49 (95% CI 0.40-0.59) and at 1,000 yds of 0.59 (95% CI 0.51-0.67).

Taking the estimated detectability parameters and applying them to sperm whales, where time spent underwater is considerably higher, led to PrU for the LT of 0.89 (95% CI 0.87-0.92) at 200 yds and 0.95 (95% CI 0.93-0.96) at 1,000 yds. MMO PrU for sperm whales was 0.77 (95% CI 0.74-0.80) at 200 yds and 0.80 (95% CI 0.77-0.84) at 1,000 yds. Hence, in this case the difference between LT PrU and MMO PrU was smaller because the long dive times place an insurmountable constraint on any visual observation position, no matter how good.

Estimated probability that a pod of marine mammals of the taxonomic group shown along the top remains undetected by the Navy lookout team (blue) or marine mammal observers (red) at ranges of 200, 500 and 1000 yards from the ship. Dots show estimates and vertical lines give 95% confidence limits. Note that the sperm whale results assume their detectability while on the surface is the same as rorquals.

For small cetaceans, many of the first detections by both LT and MMO positions were at very close ranges, well within the smallest mitigation range of 200 yds, even after bowriding pods were removed. Because of this, for the SCSP group, the estimated PrU was close to 1 at all mitigation ranges tested and for both positions. We speculate that this result was caused by a combination of (a) genuinely low detectability combined with the surfacing pattern of this group, (b) fast and possibly responsive movement (attraction to the boat) by some pods, which violates a model assumption, (c) some rounding of detection distances and possibly angles, which violates another model assumption. For the SCLP group, which are assumed to have a surfacing pattern that makes them more available for detection, results improved slightly compared to the SLSP group. Estimated LT PrU for this group was 0.94 (95% CI 0.91-1.00) at 200 yds and 0.99 (95% CI 0.99-1.00) at 1,000 yds. The equivalent estimates for MMOs were 0.83 (95% CI 0.74-0.90) at 200yds and 0.97 (95% CI 0.95-0.98) at 1,000yds. Overall, for small cetaceans, we conclude that PrU is high (and hence effectiveness low) across pod sizes, caused by a combination of low detectability of small pods and possibly responsive movement of some taxa within the small cetacean groups.

We summarize our findings as follows:

1. Based on the data and analyses presented here, the ship's lookout team (LT) have approximately an 80% chance of failing to detect a pod of large baleen whales (rorquals) before they come closer than a mitigation range of 200 yards. This probability of a pod remaining undetected (PrU) rises to 85% at 500 yards and 91% at 1,000 yards.
2. The marine mammal observers (MMOs) performed better for this taxonomic group: for example, the PrU at 200 yards was lower at 49%. Note that the MMO team consisted of two dedicated observers while the LT consisted varying number of LOs depending on the type of ship and the training activity the ship was engaged in.
3. For species (sperm whales) with longer dive times but the assumed same detectability as rorquals, the PrU for both LT and MMOs was estimated to be higher (e.g., 89% for LT and 77% for MMOs at 200 yards), with less difference between the LT and MMOs.
4. For small cetaceans the majority of first detections of a pod (particularly those made by the LT) took place at very close range regardless of pod size. Estimated PrU for small pods (1-6 individuals) was close to 100% for any range, while for large pods this probability was lower for 200yds at 94% for the LT and at 83% for MMOs and for 500 yds at 98% for the LT and 93% for the MMOs. Small cetacean pods are genuinely difficult to detect, but in addition a limitation of our model was that it assumed no horizontal movement while some small cetaceans are attracted to ships and can move quickly (although we excluded pods where bowriding behavior was noted explicitly). Despite this it seems clear that PrU is high for small cetaceans.
5. We did not estimate PrU for beaked whales as none were recorded in the surveys. However, given they are not as detectable as sperm whales but have similar dive patterns, we would expect their PrU to be higher than sperm whales.
6. Our analyses assumed that the average surfacing pattern is known for each taxonomic group and used values taken from the literature. In reality, surfacing pattern varies by species and will likely differ from literature values. We undertook some sensitivity analyses and found that results were largely the same, except for sperm whales where assumptions about dive pattern made some difference to the predicted PrU. Overall our findings are unlikely to differ substantially if uncertainty and heterogeneity in surfacing could be included. Deviation of ship trajectory from the straight-line constant-speed assumption will also have some effect on results, but ship trajectory was unknown to us.
7. If further data collection were envisaged in the future, we would encourage further revision and tightening of the data recording procedures, in collaboration with the analysts.
8. Further analytical developments could include incorporation of responsive animal movement, changing ship trajectory and measurement error.

15. SUBJECT TERMS

Monitoring, cetaceans, baleen whales, toothed whales, lookout effectiveness, Hawaii Range Complex, Southern California Range Complex

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			UU	41	Department of the Navy
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPONE NUMBER (*Include area code*)**
Unclassified	Unclassified	Unclassified			808-471-6391

Executive Summary

The United States (US) Navy uses lookouts (LOs) to detect objects in the water in the vicinity of ships. One class of object that LOs are trained to detect is marine mammals[1]; this forms an important component of the Navy's procedures for marine mammal mitigation during training activities[2]. As well as dedications of marine mammals by these LOs, detections of marine mammals may also be made by other members of the ship's crew such as officers on the bridge ("watchstanders") or sonar technicians, although acoustic detections require visual confirmation. We refer to these personnel together as the "lookout team" (LT). The primary goal of this project was to determine how effective LTs are at detecting marine mammals before they entered a defined set of mitigation ranges during mid-frequency active sonar training activities. These ranges were 200, 500 and 1,000 yards. A secondary goal was to compare this effectiveness with that of trained marine mammal observers (MMOs).

In collaboration with Navy environmental personnel, we developed a field protocol for at-sea experiments, where MMOs set up trials by locating marine mammals around Navy ships training with mid-frequency active sonar and determined whether these animals were detected by the LT. We also developed new analytical methods that allow estimation of the probability of animals approaching to within a specified mitigation range without being detected (probability of remaining undetected, PrU). These methods include a model for the surfacing pattern of animal pods[3], and for the range-dependent probability of detecting a pod each time it surfaces. Crucially, the methods allow us to account for the possibility that animal pods may remain undetected by both MMOs and the LT. The methods are flexible in allowing for various patterns of animal surfacing and various experimental configurations (in terms of communication between MMO and LT positions and whether repeat surfacings of the same pod are recorded). They are, however, simplistic in assuming that there is no measurement error (in surfacing location, taxon designation and whether duplicate detections are correctly assigned), that pods only move in the vertical plane (i.e., there is negligible horizontal movement during the period when the pod is within observation range), and that the ship moves at constant speed in a straight line. We tested the new analytical methods using computer simulation and found they generally produce unbiased estimates when the model assumptions are met, although in some circumstances (including those in our at-sea study) it is not possible to estimate both detectability and surfacing pattern; in this situation if the parameters governing surfacing pattern are known then unbiased estimation of detectability, and hence PrU, is possible.

A total of 27 embarks were conducted between 2010 and 2019, mostly on destroyer class ships. These generated 716 valid sightings of animal pods. Each sighting consisted of one or more detection of a marine mammal pod by the MMO and/or LT positions; to be valid there had to be enough information recorded to derive a taxonomic code at the level needed for analysis (see below), pod size and, for each detection, a location (relative to the ship) and an observer position. There were no valid acoustic detections, and so all LT detections were generated by the LOs or watchstanders. Some species of small cetacean are known to approach ships and "bowride"; after

[1] Seals and turtles are also included in mitigation for various activities; however, they were not included in this study.

[2] The Navy's required mitigations for each training activity are described in each Letter of Authorization (LOA), and lookout configurations are dependent on the type of ship and training activity (see AFTT and HSTT Training LOAs; Section 6(a)(2)) (NMFS 2019, 2020).

[3] We use the term "pod" to refer to a group of one or more marine mammals. This term is typically used only for cetaceans but, as we document lower down in the report, there were not enough pinniped detections to include them in the analysis.

discussion with Navy environmental personnel it was decided to exclude detections of pods observed during the sighting to engage in bowriding behavior. There were 46 such sightings, with first detections predominantly made at close ranges. After excluding these, 670 sightings remained.

Our data collection protocol asked MMOs to prioritize new sightings over repeated detections (resights) of an already-sighted school, and so resights were not recorded consistently. We therefore used analytical methods that require data only on the first detection of a pod by each position. Analysis at species level was not possible because of limited sample size, and because many sightings were not identified using a taxonomic code that refers to species, but instead to a higher taxonomic level such as "large whale" or "dolphin"[4]. We therefore divided the data into four groups according to similarity in surfacing pattern and detectability: rorquals (i.e., large baleen whales), sperm whales, small cetaceans in small pods (6 or less) (SCSP) and small cetaceans in large pods (more than 6) (SCLP). We assumed the parameters governing surfacing pattern for each group were known, and we used values derived from the literature. For the sperm whale group, for which there were only two sightings, we used the detectability parameters estimated for rorquals. There were not enough detections of pinnipeds for us to estimate range-dependent probability of detection from the detection data and, unlike sperm whales, we elected not to use the estimated detectability parameters from one of the other groups; hence, our results only cover cetaceans.

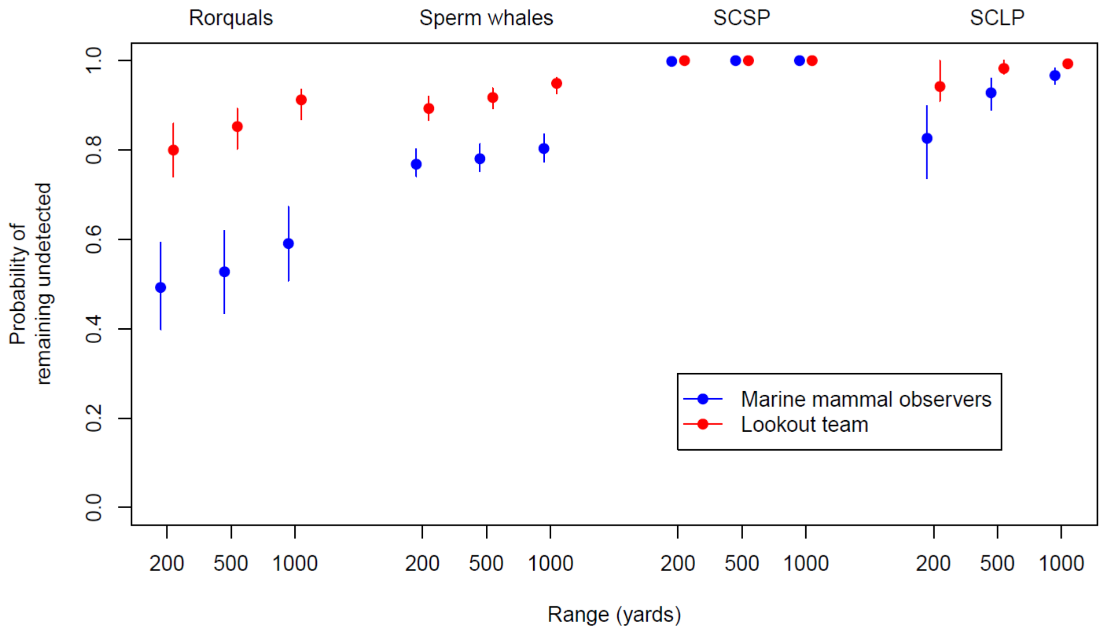
Before undertaking the modelling we performed some exploratory analyses, including calculating a simple distance-specific index of effectiveness at 200, 500 and 1,000 yards (yds) for rorquals, SCSP and SCLP. For this analysis, we quantified LT effectiveness as the number of pods detected by the LT *before* they enter within the mitigation range divided by the total number of pods thought to have entered within the mitigation range (as estimated by the number seen by the LT or MMOs within a given distance of the ship's track). We speculate that this provides an upper bound on absolute effectiveness, because it does not take account of pods that pass through the mitigation zones undetected by either position. Estimated effectiveness was highest for rorquals: 0.35, 0.21 and 0.13 at 200, 500 and 1,000 yds for the LT and 0.74, 0.70 and 0.54 respectively for MMOs. It was lowest for SCSP: 0.03, 0.03 and 0.02 respectively at 200, 500 and 1,000 yds for the LT and 0.25, 0.29 and 0.14 respectively for MMOs. The estimates for SCLP were similar to SCSP for the LT but higher than SCSP for MMOs.

Results from the modelling analysis to obtain PrU are summarized graphically in the figure on the next page. For each group, we estimated PrU at 200, 500 and 1,000 yards (yds). Please note that, although the results are quoted at these ranges, all of the data from each taxonomic group (including data beyond 1,000 yds) was used in deriving the results with these models. For rorquals the estimated PrU at 200 yds for the LT was 0.80 (95% confidence interval (CI) 0.74-0.86), rising to 0.91 (95% CI 0.87-0.94) at 1000 yds. PrU is the complement of effectiveness, so estimated absolute effectiveness was 1-0.80=0.20 at 200 yds and 1-0.91=0.09 at 1,000 yds. As expected, these values are slightly lower than the simple distance-specific index of effectiveness quoted in the previous paragraph (and this pattern held true for all such comparisons). MMOs were estimated to be considerably better, with PrU at 200 yds of 0.49 (95% CI 0.40-0.59) and at 1,000 yds of 0.59 (95% CI 0.51-0.67).

Taking the estimated detectability parameters and applying them to sperm whales, where time spent underwater is considerably higher, led to PrU for the LT of 0.89 (95% CI 0.87-0.92) at 200 yds

[4] A full list of taxonomic codes is given in Appendix A. One reason that identification to species level was sometimes not possible was that, unlike many research cruises, Navy ships did not approach pods in order to confirm species identification.

and 0.95 (95% CI 0.93-0.96) at 1,000 yds. MMO PrU for sperm whales was 0.77 (95% CI 0.74-0.80) at 200 yds and 0.80 (95% CI 0.77-0.84) at 1,000 yds. Hence, in this case the difference between LT PrU and MMO PrU was smaller because the long dive times place an insurmountable constraint on any visual observation position, no matter how good.



Estimated probability that a pod of marine mammals of the taxonomic group shown along the top remains undetected by the Navy lookout team (blue) or marine mammal observers (red) at ranges of 200, 500 and 1000 yards from the ship. Dots show estimates and vertical lines give 95% confidence limits. Note that the sperm whale results assume their detectability while on the surface is the same as rorquals.

For small cetaceans, many of the first detections by both LT and MMO positions were at very close ranges, well within the smallest mitigation range of 200 yds, even after bowriding pods were removed. Because of this, for the SCSP group, the estimated PrU was close to 1 at all mitigation ranges tested and for both positions. We speculate that this result was caused by a combination of (a) genuinely low detectability combined with the surfacing pattern of this group, (b) fast and possibly responsive movement (attraction to the boat) by some pods, which violates a model assumption, (c) some rounding of detection distances and possibly angles, which violates another model assumption. For the SCLP group, which are assumed to have a surfacing pattern that makes them more available for detection, results improved slightly compared to the SLSP group. Estimated LT PrU for this group was 0.94 (95% CI 0.91-1.00) at 200 yds and 0.99 (95% CI 0.99-1.00) at 1,000 yds. The equivalent estimates for MMOs were 0.83 (95% CI 0.74-0.90) at 200yds and 0.97 (95% CI 0.95-0.98) at 1,000yds. Overall, for small cetaceans, we conclude that PrU is high (and hence effectiveness low) across pod sizes, caused by a combination of low detectability of small pods and possibly responsive movement of some taxa within the small cetacean groups.

We summarize our findings as follows:

1. Based on the data and analyses presented here, the ship's lookout team (LT) have approximately an 80% chance of failing to detect a pod of large baleen whales (rorquals) before they come closer than a mitigation range of 200 yards. This probability of a pod remaining undetected (PrU) rises to 85% at 500 yards and 91% at 1,000 yards.
2. The marine mammal observers (MMOs) performed better for this taxonomic group: for example, the PrU at 200 yards was lower at 49%. Note that the MMO team consisted of two dedicated observers while the LT consisted varying number of LOs depending on the type of ship and the training activity the ship was engaged in.
3. For species (sperm whales) with longer dive times but the assumed same detectability as rorquals, the PrU for both LT and MMOs was estimated to be higher (e.g., 89% for LT and 77% for MMOs at 200 yards), with less difference between the LT and MMOs.
4. For small cetaceans the majority of first detections of a pod (particularly those made by the LT) took place at very close range regardless of pod size. Estimated PrU for small pods (1-6 individuals) was close to 100% for any range, while for large pods this probability was lower for 200yds at 94% for the LT and at 83% for MMOs and for 500 yds at 98% for the LT and 93% for the MMOs. Small cetacean pods are genuinely difficult to detect, but in addition a limitation of our model was that it assumed no horizontal movement while some small cetaceans are attracted to ships and can move quickly (although we excluded pods where bowriding behavior was noted explicitly). Despite this it seems clear that PrU is high for small cetaceans.
5. We did not estimate PrU for beaked whales as none were recorded in the surveys. However, given they are not as detectable as sperm whales but have similar dive patterns, we would expect their PrU to be higher than sperm whales.
6. Our analyses assumed that the average surfacing pattern is known for each taxonomic group and used values taken from the literature. In reality, surfacing pattern varies by species and will likely differ from literature values. We undertook some sensitivity analyses and found that results were largely the same, except for sperm whales where assumptions about dive pattern made some difference to the predicted PrU. Overall our findings are unlikely to differ substantially if uncertainty and heterogeneity in surfacing could be included. Deviation of ship trajectory from the straight-line constant-speed assumption will also have some effect on results, but ship trajectory was unknown to us.
7. If further data collection were envisaged in the future, we would encourage further revision and tightening of the data recording procedures, in collaboration with the analysts.
8. Further analytical developments could include incorporation of responsive animal movement, changing ship trajectory and measurement error.

Table of Contents

List of abbreviations

AFTT	Atlantic Fleet Training and Testing
CREEM	Centre for Research into Ecological and Environmental Modelling
DMMO	Data recorder marine mammal observer
EDA	Effective detection area
ESHW	Effective strip half-width
HSTT	Hawaii-Southern California Training and Testing
g(0)	Probability of detection at zero perpendicular distance from the direction of travel
IE	Index of effectiveness of an observation position (LT or MMO)
LMMO	Liaison marine mammal observer
LO	Lookout
LOA	Letter of authorization
m	meter(s)
min	minute(s)
LT	Lookout team
MMO	Marine mammal observer
PrU	Probability of remaining undetected (at some specified mitigation range from the ship)
SCLP	Small cetaceans in large pods (more than 6 individuals)
SCSP	Small cetaceans in small pods (6 or fewer individuals)
SOCAL	Southern California (offshore range)
SMMO	Surveying marine mammal observer
US	United States
yds	yards

Introduction

The United States (US) Navy undertakes mitigation for marine mammals during training activities as part of mitigation procedures designed to minimize risk to these animals. One component of this mitigation is the shipboard lookouts (LOs), who are part of the standard operating procedure that ships use to detect objects (including marine mammals and other animals) around the ship during operations. The LOs are an element of monitoring requirements specified by the National Marine Fisheries Service in the Letters of Authorization (LOAs) issued pursuant to the Marine Mammal Protection Act for the US Navy's training and testing. As well as dedicated LOs, detections of marine mammals may also be made by other members of the ship's crew such as officers on the bridge (watchstanders) or sonar technicians (although in the latter case visual confirmation is required). We refer to all these personnel together as the "lookout team" (LT). The primary goal of this project was to determine how effective the LTs are at detecting marine mammals before they enter a defined set of mitigation ranges. This was achieved undertaking a set of at-sea trials where we could compare LT observations with those made by teams of highly trained civilian or contractor Marine Mammal Observers (MMOs) who were stationed on board Navy ships training with mid-frequency active sonar. This setup enabled a secondary aim of determining how LT effectiveness compared with that of MMO teams.

History of project

This project was initiated in 2010, when researchers from the Centre for Research into Ecological and Environmental Modelling (CREEM) collaborated with marine biologists from the Navy to design a field protocol for the at-sea trials. The protocol was revised slightly (Burt and Thomas 2010) after the first four cruises. In parallel, these initial four cruises (which generated 125 sightings) were used as the basis for initial development of the required modelling approach (Rexstad and Thomas 2010a, Thomas et al. 2011). The new analytical methods constitute a substantial extension of previous methods for analysis of line transect survey data, and in turn formed the basis for two peer-reviewed publications (Langrock et al. 2013, Borchers and Langrock 2015).

Data collection continued, with eight cruises having been completed by Feb 2012, yielding 182 sightings. An analysis of these data was undertaken (Thomas et al. 2012), with the data divided into two functional groups: "large animals" (i.e., large whale species) and "small animals" (mainly dolphins). The main conclusion was that more data was required for reliable results; it was recommended that, if possible, data collection was focused on a single ship type (destroyer) to minimize heterogeneity and location (SOCAL) to maximize detections per cruise. Notwithstanding the small sample size, the preliminary analysis indicated that probability of an animal pod (group of 1 or more animals) coming to within close range of the ship and yet remaining undetected (here denoted "PrU") by the LT may be quite high – for example estimated PrU at 100 meters (m) (109 yards, yds) ranged from 0.74 to 0.94 depending on assumptions about animal surfacing behavior.

Further data collection was undertaken between 2012 and 2019, with the data collection protocol receiving a further round of revisions by Navy personnel (Department of the Navy 2016) – this version is included as Appendix B. There are now 27 embarks complete, with 716 valid sightings (Table 1; a valid sighting is where one or more detection is made by the MMOs and/or LT, and all required information was recorded, such as taxon, distances, etc. – see Methods). The Navy has requested an analysis of observer effectiveness based on these data. After discussion with Navy personnel, the three mitigation ranges for calculating PrU used in this report are 1,000, 500 and 200 yds (914, 457 and 183 m). In other words, we seek to quantify the probability of a marine mammal school coming to within 1,000 yds of the ship without being detected, 500 yds without being detected, and 200 yds without being detected.

Report overview

In the Methods section, we briefly summarize the field survey protocol and give a summary of the modelling approach developed. We have further extended the approach used in Thomas et al. (2012) and Borchers and Langrock (2015) to account for the fact that, in the survey protocol, MMOs are aware of LT detections but not vice versa, and we describe the extension. Full technical details are given in Appendix C. We then summarize the simulation studies used to test these new methods, and how we applied these methods to the complete survey dataset.

In the Results section we present results from the simulation study, exploratory analyses of the survey dataset and our modelling results. Further results are given in Appendix C. We finish by discussing our results.

Methods

Field survey protocol

A detailed field protocol was developed and refined; this is reproduced in Appendix B. Here we give a summary.

During Navy mid-frequency active sonar training activities, dedicated LOs (in varying numbers depending on the type of ship) are positioned at the forward part of the ship, with additional lookouts serving as members of the bridge watch. On destroyers and cruisers, LOs may be located on the bridge wings. The Navy's required mitigations for each training activity are described in each LOA, and lookout configurations are dependent on the type of ship and training activity (see AFTT and HSTT Training LOAs; Section 6(a)(2)) (NMFS 2019, 2020). When sonar technicians are on duty during certain training activities, they report acoustic detections of marine mammals to the bridge. Visual detections are sometimes made by officers on the bridge (as well as other crew members). We refer to all of these observers together as the lookout team (LT). Officers on the bridge or in combat operations center were responsible for entering marine mammal records into a log, but we did not use this log in our experiments as it was not sufficiently detailed for our purpose. Instead, one of the MMO team acted as a liaison with the primary LO (see below) and kept detailed records of the detections made. During the experiments, every effort was made to allow the LT to operate as normal: MMOs were careful not to interfere with the work of the LT, or to cue them when the MMO team made sightings (see below). The only condition under which the MMO would inform the LT of a MMO sighting was when the animal was close aboard and a strike or close encounter might occur.

On each experimental embark, a team of four MMOs were embarked on a Navy ship (frigate, destroyer or cruiser) training with mid-frequency active sonar. The four rotated in duties and collected data during daylight hours and as much of the cruise as possible, including in all sea states and visibility conditions. At any one time during data collection, two MMOs were surveying for animals, one acted as liaison with the LT and one acted as data recorder or was resting. Where it is important to distinguish these duties, we refer to surveying marine mammal observers as SMMOs, the liaison as LMMO and the data recorder as DMMO. All MMOs were in radio contact with one another. MMOs kept standardized records of their survey effort, who was on which station, high-level ship activity (e.g., whether sonar was in use), weather and visibility (Appendix B, Section 4).

The SMMOs were generally stationed on the bridge wings of the ship. Their main functions were to detect and track marine mammals and determine whether detections made by the LT and reported by the LMMO were duplicates with sightings they had made. The SMMOs were careful to operate in such a way as not to cue the LT when SMMOs made detections. The SMMOs searched with naked eye and 7x50 binoculars; the port observer searched from about 5 degrees starboard to abeam on the port side while the starboard observer searched from about 5 degrees port to abeam on the starboard side. On detecting an animal (or pod of animals) a set of information was recorded (see Section 4 of Appendix B) including taxonomic code, range, bearing, pod size and behavior. The taxonomic code allowed for identification at a range of taxonomic levels, from high-level ("large whale", "dolphin", etc.) down to species – a full list is given in Appendix A. One reason that identification to species level was sometimes not possible was that, unlike many research cruises, Navy ships did not purposefully approach marine mammals to confirm species identification. As part of the protocol, SMMOs attempted to track the pod until past abeam and record each resight where possible; however, under the protocol they were asked to give higher priority to detecting new animals and so tracking was not done consistently. The reason for prioritizing sighting new pods is that each new sighting by the SMMOs sets up a new "trial" for the LT, which is either a success (LT detects school) or failure (LT does not detect school) giving direct information relevant to the primary goal of determining observer effectiveness. On the other hand, tracking already-sighted pods helps to identify whether a later LT detection is a match for an already-sighted SMMO pod and so is useful in determining duplicates; consistent tracking data, if gathered, can also be used in the estimation procedure (see below). Overall, prioritizing direct information on observer effectiveness was judged to be more important where a choice had to be made.

The LMMO was stationed on the bridge to observe the OT. Depending upon the configuration of the lookouts, the LMMO may be positioned either inside the pilot house, on one bridge wing, or moving between the bridge wings and the pilot house. Their main function was to record information on the first detections of pods by the LT, particularly the LO. Information on resights were recorded if possible. The LMMO passed relevant information on the detection to the rest of the MMO team as soon as possible – this included the LT's estimation of range, bearing and taxon. In some cases the LT made detections before the SMMOs.

The MMO team were responsible for determining whether a surfacing seen by either observation position (SMMO or LT) was a resight of a previous pod and whether a surfacing seen by the LT was a duplicate of a sighting by the SMMOs. Determinations were classified as either definite (at least 90% likely), probable (50-90% likely) or remote (less than 50% likely – although in practice this category was not used).

In the following we use the term "sighting" to refer to one or more detections of the same pod by either or both positions. We refer to an individual surfacing that is detected by either or both positions as a "detection". After the first detection of a pod by a position (MMO or LT), second and subsequent surfacings that are detected by the same position are referred to as "resights".

Exploratory data analysis

Initial analyses of the data were undertaken several times and are documented in previous reports (Rexstad and Thomas 2010a, b, Thomas et al. 2011, Thomas et al. 2012). Here we undertook the following analyses in advance of modelling to estimate PrU, which is described in subsequent sections.

Note, in this and subsequent sections, unless stated otherwise, we use the term "marine mammal observer" (MMO) to mean the surveying marine mammal observers (SMMOs).

Some data formatting and cleaning was required. Data files often had different structures and column headers. It was not clear how reticle readings were converted into distances (even though the formula was given in most files, the values could not be reproduced). Hence, we used the NOAA conversion to calculate distances from reticles (Kinzey et al. 2000) using the height of the ship given in the data files. Sightings and resights were often not clearly labeled and matches of sightings or individual detections between MMOs and LT were often described verbally in the comment section. This required manual inspection of each record and assigning corrected sighting and resight numbers. Data analysis required unique identifiers for each sighted pod and each resighted surfacing of the respective pod and which observer team had detected these. Information on who may have cued who also needed to be obtained from the comment section.

We removed sightings where it was not clear which position made each detection, where it was not possible to derive the perpendicular or horizontal distance, taxonomic code or pod size (the latter two were often derived from other detections within the same sighting). We removed a single detection where the observer was recorded as "sonar" for which the species was logged as "Biologics" and no corresponding visual confirmation in the same sighting existed which may have aided to improve the species identification. The only detection reported by the sonar technician, logged as an un-identified marine mammal, was also not included in the analysis as this taxonomic code could not be attributed to any of the four groups with similar surfacing behavior described above. We removed one detection (of a humpback whale in Hawaii) where the observer was recorded as "Aerial" because this detection came from a contracted monitoring aircraft conducting a simultaneous survey, not a Navy asset. We also removed any detections made after the pod passed abeam of the ship because the pod was then beyond its closest point of approach; however we did use these detections to help with taxonomic identification of any previous detections made during the same sighting. Remaining sightings and detections-within-sightings are referred to as "valid".

In matching detections to form a single sighting, we used all cases where the MMOs judged the match to be "definite" or "probable". We planned to discard all cases where the match probability was judged "remote", but in practice there were none of these in the dataset.

One piece of information noted by the observers relates to animal behavior, and this included cases where sightings were of pods observed to actively engage in "bowriding" behavior at some point during the sighting. Such pods were typically first observed at close range. After discussion with Navy environmental it was decided to exclude these detections (see Discussion). Other pods noted to be "closing" on the ship were not excluded because it was not known whether they were responding to the ship or would have been travelling in that direction without the ship's presence.

A simple index of the effectiveness (IE) of an observation position is the proportion of pods known to be present that were detected by that position. This is not an absolute measure of effectiveness as it does not account for pods undetected by either position, nor does it incorporate range-dependent detection. Denoting the number of marine mammal sightings by LTs, MMOs and both combined as n_{LT}, n_{MMO} and n respectively, we calculated the indices of effectiveness as

$$IE_{LT} = \frac{n_{LT}}{n}$$
$$IE_{MMO} = \frac{n_{MMO}}{n}$$

As an initial view of sighting location relative to the mitigation ranges (200, 500 and 1,000 yds), we plotted first detection location for each sighting separately for the LT and MMOs. If the majority of first detections are made within the mitigation ranges, this implies low effectiveness to detect pods before they enter these ranges. However, this does not account for pods that are unobserved, which provides motivation for the more sophisticated modeling described in the next section.

To provide a quantitative summary of the information in the plots, we extended the index of effectiveness to incorporate the spatial location of detections and hence derive a simple index of effectiveness in detecting pods before they enter the mitigation ranges. Under the assumption that pods do not change their perpendicular distance from the ship, then any pods first sighted within perpendicular distance x will pass within range x of the ship as the ship moves forward and past the pod. Hence all detections of pods made by either position within perpendicular distance x, denoted n_x, form a set of "trials". The number of these first detected by the LT at ranges r greater than x, denoted $n_{LT,r>x}$, is the number of "success" for the LT because each of these pods was detected before it entered the mitigation range. Similarly, the number of trials first detected at ranges greater than x by the MMOs, $n_{MMO,r>x}$ is the number of MMO "successes". We calculated the distance-specific indices of effectiveness as

$$IE_{LT}(x) = \frac{n_{LT,r>x}}{n_x}$$

$$IE_{MMO}(x) = \frac{n_{MMO,r>x}}{n_x}$$

for $x = 200$, 500 and 1,000 yds. This simple measure is likely an over-estimate of effectiveness because the n_x comprises only the pods known to be present because they were detected by one or both positions – it does not account for pods that were present but not detected by either position. (It does, however, assume that the perpendicular distance a pod is first detected at is their perpendicular distance when the ship passes abeam – see Discussion.) Again, this provides motivation for the methods described next.

Estimating range-dependent probability of remaining undetected (PrU)

Here we give an outline of the approach developed; full details are in Appendix C. Note that the methods are also of potential use in shipboard line-transect surveys for estimating density of marine fauna, where there is also the problem of estimating detectability in the presence of non-continuous availability.

Conceptually, the modeling approach used can be divided into two parts: a process model describing the horizontal and vertical movement of pods relative to the ship transporting the observers, and an observation model describing the way that observers sight and record pods given the pods' horizontal and vertical location. We describe each in turn.

For the vertical movement component of the process model, animals are assumed to be in one of two behavioral states: (1) diving relatively shallowly where they may on occasion be at the surface (e.g., to breathe), or (2) diving relatively deeply such that they will not surface. In state 1, they are assumed to surface at random intervals (drawn from an exponential distribution) with an average rate that is a model parameter, λ; while in state 2, they do not surface. Surfacings are assumed to be instantaneous. The probability of switching from one state to the other at any instant only depends upon the current state, and is governed by two model parameters, q_{12} and q_{21} (the state switching rates). This framework gives the flexibility to allow for clusters of surfacings followed by extended periods underwater. It also includes two special cases: intermittent and continuous availability. Intermittent availability is where one or more member of the pod is on the surface for extended periods between dives, and can be accommodated by assuming pods are at the surface while in state 1; continuous availability is where one or more member of the pod is always at the surface and can be accommodated by additionally removing the state 2 component from the model (or by setting the switching rate q_{21} to a suitably high value, so that negligible time is spent in state 2).

In theory, given sufficient data, it is possible to estimate the availability parameters q_{12}, q_{21} and λ as well as the observation model parameters described below. However, in practice, these parameters are only accurately identifiable in a limited range of circumstances – for example with large datasets with reliable resighting data and an amenable detection process (see Borchers and Langrock 2015). Borchers and Langrock (2015) distinguished between no assumed knowledge of the availability parameters (where all three parameters are estimated), partial knowledge (where q_{12} and q_{21} are assumed known and λ is estimated) and full knowledge (where all three parameters are assumed known). In the work carried out as part of this current contract we assumed full knowledge and used parameter values from the literature (see "Application to survey data", below).

The horizontal component of the process model is simplistic. Schools are assumed not to move horizontally, while the ship is assumed to move in a straight line at a constant speed of 12 knots. This means that pods do not change their location relative to the ship in the *x*-direction (i.e., perpendicular to the direction of travel of the ship), but their position in the *y*-direction (parallel to the direction of travel of the ship) gets smaller at a rate of 12 nautical miles per hour (with zero being abeam).

The observation model assumes that the LT and MMO positions each have separate two-dimensional hazard functions, describing the probability of detecting a surfacing given its (*x*,*y*) location. Animals beyond some specified distance from the observer are assumed to be undetectable, as are animals below the surface. The hazard functions have parameters that fix their shape; the function used in this report, the inverse power hazard function, has three parameters, α, β and γ, and is of the form:

$$h(x,y) = \frac{\alpha\beta^{\gamma}}{(\beta^2 + x^2 + y^2)^{\gamma/2}}$$

(see figures in Results for example fits). (An alternative hazard function, which also has three parameters, is the exponential power function – see Borchers and Langrock 2015.) Hence there are six parameters in the observation model: three-parameters for the LT hazard function and three parameters for the MMO hazard function. It is assumed that the location of sighted animals is recorded without error, that duplicate identification is certain (i.e., that there are no mistakes in determining whether a surfacing was detected by MMO alone, LT alone, or both), and that the object detected has been correctly identified to the taxonomic level used in the analysis. The model also assumes some maximum detection range beyond which there are no detections, and can optionally use a perpendicular truncation distance w (this latter is just like standard line transect analysis).

The model framework is capable of accommodating the first detection of a sighting by either or both observation position and also resights by either or both positions of later surfacings. It is likely that the detection function for an observation position changes significantly after that position becomes aware of a school, and this can be accommodated by allowing the detection function parameters to differ between initial detections and resights. As noted previously, our survey protocol meant that resights were not collected consistently, and so in the data analyses performed here we used a version of the model that utilizes information only from first detections by each position.

Initial development of this model was documented in Thomas et al. (2011, 2012) with more developed versions given in Langrock et al. (2013) (which considered only one observation position) and Borchers et al. (2015) (which included multiple independent observation positions). The latter assumed the two positions operated fully independently, so further development was required to accommodate the one-way dependence in our survey protocol: the MMO position could make detections independently of the LT, but when the LT made a detection the MMO position was informed, creating a dependence. The additional development to accommodate this one-way dependence is described in detail in Appendix C. When only first detection data are used, then each pod encounter can yield one of four kinds of sighting record: (1) MMOs detected the pod but LT did not; (2) MMOs detected the pod and LT detected it on the same surfacing; (3) MMOs detected the pod but LT missed that surfacing; LT detected the pod on a subsequent surfacing (before the pod went abeam); (4) LT detected a pod but MMOs missed it until the LT brought it to their attention.

The methods require specification of a forward truncation distance, beyond which detection probability is assumed to be 0, and a perpendicular distance, w, which functions like the truncation distance in line transect analysis. In the analysis of survey data reported here we used the furthest observed forward distance and furthest observed perpendicular distance as these two truncation points. Hence all valid detections were used in the analysis. Results reported here are insensitive to the choice of truncation distance in the sense that using much larger distances would have yielded the same results (but taken more computer time to run) because the estimated detection hazard at the truncation points was extremely small.

The model is fitted to data using maximum likelihood. This yields parameter estimates and estimates of uncertainty in the model parameters.

Given a fitted model, it is straightforward to estimate the probability of remaining undetected (PrU) to a given set of mitigation ranges. This involves integrating the surfacing and detection models in the forward distance direction from the maximum detection range up to the mitigation range, and in the perpendicular distance direction from the truncation distance w to zero.

The same technique can be used to calculate the probability of detection up to a given forward distance. An important application of this technique in the context of line-transect surveys is to calculate the probability of detecting an animal between the maximum detection distance and a forward distance of 0 (i.e., when the animal goes abeam of the ship, for a given perpendicular distance). This equates to the perpendicular detection probability, $g(x)$. Of particular interest is $g(0)$, the probability of detecting an animal at zero perpendicular distance. We use this as one metric to compare different detection functions in the analyses that follow. Another metric related to line-transect surveys is the effective strip half-width (ESHW, i.e., the perpendicular distance within which as many animals are missed as are detected outside that distance) (Buckland et al. 2001).

Uncertainty in the derived quantities PrU, ESHW and g(0) was obtained from a parametric bootstrap procedure. We sampled 1,000 random parameter estimates of the model, assuming estimates were distributed according to a multivariate normal distribution with mean equal to the maximum likelihood estimates and using the estimated covariance matrix. For each sample, we calculated PrU (at each mitigation range), ESHW and g(0). From these 1,000 estimates of each quantity, we estimated variance as the empirical variance in values and used the percentile method to obtain confidence intervals.

We evaluated the goodness-of-fit of the fitted models by comparing the observed number of detections in a set of distance intervals with those predicted under the model. We made visual plots for both MMOs and the LT showing observed and expected detections in the forward and perpendicular distance directions, and also in a set of range bins. For the latter we also undertook χ^2 goodness-of-fit tests.

Simulation studies

Simulation is a useful way to examine the performance of a new statistical method under conditions where the truth is known. Many random replicate datasets are generated from a model and then analyzed; results can then be used to compute various metrics such as percentage bias (i.e., the mean of the estimates minus the true value, all divided by the true value and multiplied by 100).

A small simulation study of the initial model was documented in Thomas et al. (2012), and more extensive studies given in Langrock et al. (2013) and Borchers and Langrock (2015). These established (among other things) that the analysis method is capable of estimating:

- both availability parameters and ESHW with low bias from single or independent-observer double platform (i.e., two independent observation position) data for a combination of parameters that yield a peak in detections away from $y = 0$;
- both availability parameters and ESHW with low bias from independent-observer double platform data where data are collected on resights, for a combination of parameters that yield a peak in detections close to $y = 0$;
- ESHW and potentially surfacing rate λ with low bias from independent-observer double platform data when only first detection data are collected, for a combination of parameters that yield a peak in detections close to $y = 0$, if the other two availability parameters are known.

Some other scenarios demonstrated considerable bias – for example independent-observer double platform data with only first detections recorded when there is a peak in detections close to y=0.

For this report, we extended the modeling approach to account for one-way independence between MMO and LT positions and using first detections by either observer only; hence we wished to test this approach through simulation, for the scenario where availability parameters are known. We also wished to examine any potential bias in estimation of PrU, which was the main focus of the observer effectiveness study.

To this end, we constructed two simulation scenarios. In both cases the pod availability pattern was taken from a preliminary analysis of the rorqual dataset (see below), yielding $q_{12} = 0.00104$ m^{-1}, $q_{21} = 0.00064$ m^{-1} and $\lambda = 0.0090$ m^{-1} (calculated as 1 over the distance travelled by the ship between events). With an assumed maximum detection range of 15,726 m and ship speed of 12 knots, this yielded an expected x surfacings per pod while the pod was in the visual field of the observers. The detection hazard parameters were also taken from the rorqual analysis, which yielded for the LT position $\theta = (0.05,\ 8103.08, 5.47)$ and for the MMO position $\theta = (0.10, 8103.08, 6.69)$. The resulting detection hazards are shown in Figure 1. In the first scenario, we assumed full independence between positions, and retained all detections. In the second scenario, which is closer to that in our survey data, we assumed one-way independence between the MMO and LT positions, and we retained only first detections for each sighting. The cumulative distribution of first detection distances is shown in Figure 2 (green dashed lines). The interpretation of this plot is that each point gives the probability of detecting a pod at that perpendicular distance *x* by the time it gets to forward distance *y* from the ship – hence the values at *y*=0 give the probability of detecting a pod at perpendicular distance x by the time it passes abeam of the ship. The value at *y*=0 and *x*=0 is what is referred to as *g(0)* in the line transect literature, and in this simulation is approximately 0.52 for the LT position and 0.71 for the MMO position.

To generate a simulated dataset, we repeated the following until we had 300 sightings. We first generated a random perpendicular distance between 0 and $w = 6{,}500$ m. We then generated a random availability pattern and for each surfacing used the detection hazard for each position to determine whether the surfacing was detected. For scenario 1 we retained all detections, while for scenario 2 we retained only the first detections by each position.

Each simulated dataset was analyzed using the independent observer all detections model (Scenario 1) and the one-way independence first detection model (Scenario 2), in both cases assuming the availability parameters were known. We repeated this exercise for 100 replicate simulations and calculated percentage bias in estimates of PrU at 200, 500 and 1,000 yards, as well as ESHW and g(0).

Application to survey data

Previous reports (Rexstad and Thomas 2010a, b, Thomas et al. 2011, Thomas et al. 2012) undertook preliminary analyses of data available at that time. The most recent (Thomas et al. 2012) divided the data taxonomically into large and small cetaceans. For the former, a preliminary version of the analysis methods used here were employed assuming three sets of availability parameters corresponding to long shallow dives, intermediate and long deep dives. Estimated PrU at 100 m (109 yards) for the LT was 0.74, 0.79 and 0.94 for the three scenarios, while for the MMOs the estimated PrU at 100m was 0.38, 0.48 and 0.81. PrU at larger range was lower. Small cetaceans were assumed, optimistically, to be continuously available, and this yielded LT PrU at 100m of 0.75. More realistically (for some species), an instantaneous availability model with assumed expected times in deep and shallow dive states of 2 minutes gave an LT PrU of 0.95.

Given the larger dataset available for this report, we undertook a re-evaluation of the taxonomic level at which observer effectiveness could be evaluated. To this end, we took guidance from the NOAA hierarchical classification system of sighting-categories (KInzey et al. 2000) to construct a similar hierarchy using the taxonomic codes used during the surveys (Appendix A). The majority of whales identified at lower levels than the level 5 category WHALE (unidentified whale) were rorquals (Table 2). The 277 sightings in this group represented 41% of all sightings. Most of these identified to species were humpback whales (84 sightings), but also including Balaenoptera sp. (29), Minke whales (3), blue whales (16) and fin whales (1). Nearly 50% in this group were not identified to species (94 sightings of unidentified whales and 50 sightings of unidentified large whales). As almost all sightings identified to species were rorquals, we grouped all these sightings (with the exception of two sperm whale sightings) into the rorqual group for analysis. Regarding availability pattern for this group, as the majority were classified as humpback whales, we assumed that humpback whale availability parameters would form a reasonable proxy for the rorquals group. We assumed that members of the rorquals group would spend on average 2.4 min near the surface (state 1) and 4.2 min diving (state 2) (Dolphin 1987). While in state 1, members of this group would surface on average once every 0.3 min. Hence, parameters q_{12} and q_{21} were calculated using: $q_{12} = \frac{1}{2.4*12*1852/60}$, $q_{21} = \frac{1}{4.2*12*1852/60}$ and $\lambda_1 = \frac{1}{0.3*12*1852/60}$, respectively.

In grouping rorqual species together, we excluded one sighting with another large whale species: the single sperm whale sighting. As an ESA-listed endangered species it was of interest to obtain an approximate estimate of PrU. On the assumption that the detection hazard for this species was the same as that of the rorquals group, we combined the estimated rorqual detection hazard with availability information derived from Drouot et al. (2004). We assumed that sperm whales on average spend 9.1 min near the surface (state 1) and 44.8 min diving (state 2); while near the surface we assumed they produce on average 4.6 blows per min.

For small cetacean species, one important factor affecting both availability and detectability is pod size. After examining the distribution of observed pod sizes, we divided the small cetacean sightings into those with an estimated pod size of 6 or fewer ("small cetacean small pod", SCSP) and those with a pod size of more than 6 ("small cetacean large school", SCLP). We analyzed these two groups separately, for the SCSP group assuming instantaneous availability and for SCLP assuming intermittent availability. We used the study by Scott and Chivers (2009) to populate the parameters for the availability model: for both groups we assumed that the pods on average spend 0.99 min near the surface (state 1) and 1.26 min diving (state 2). For the SCSP group we assumed that on average every 0.1 min at least one animal of the pod was at the surface while the pod is in state 1 and hence, making the pod available to be detected. For the SCLP group we assumed that at least one member of the pod was at the surface at all times while in state 1.

While the values described above represented the settings for the main analyses, we conducted a sensitivity analysis for each group. As part of this, we used other suitable values from the literature to populate the availability model (see Appendix C for details).

Results

Exploratory analysis of survey data

After data cleaning there were 716 valid sightings in the data set. Of these 46 were recorded as including bowriding behavior and were therefore excluded, giving 670 remaining sightings (Table 2). There were 5 sightings of animals recorded as "closing" but not recorded as bowriding during the sighting; these were retained. The remaining sightings were divided into the four groups described earlier (rorquals, sperm whales, SCSP and SCLP, 544 sightings) and other animal taxa outside these groups, including pinnipeds, turtles and fish (126 sightings) that were not analyzed further.

The rorquals group encompassed 277 sightings (Table 3), comprising 301 detected surfacings by either or both positions (including first detections and subsequent resights). Out of the 277 sightings, 212 were detected by the MMOs only, 21 by the LT only and 44 by both teams. This gives an index of effectiveness for the MMOs of $IE_{MMO} =$ 0.92 and for the LT of $IE_{LT} =$ 0.23, i.e., 4 times lower. Out of the 44 sightings detected by both positions, 19 were detected by the MMOs first, 5 by the LT first and 20 by both positions during the same surfacing. 104 encounters could be identified to species, 29 to genus and the remaining 144 to higher taxonomic levels.

The location of first detections of rorquals relative to the ship is shown in Figure 3. The percentage of first detections made by the LT within 200, 500 and 1,000 yds of the ship was 6.2%, 13.8% and 32.3%; the corresponding percentages for the MMOs were 0.8%, 3.1% and 13.3%. Comparing these values from the two positions, it is clear that the LT were focusing their search effort closer to the ship than the MMOs.

This finding is also reflected in the distance-specific indices of effectiveness (Table 4). Of all pods first detected within 200 yds perpendicular distance from the vessel's track 74% of them were detected by the MMOs outside of the 200 yds mitigation range ($IE_{MMO}(200)$ = 0.74) while 35% were detected by the LT outside this range ($IE_{MMO}(200)$ = 0.35). Similarly, for the 500 yds distance, MMOs detected 70% of pods outside 500 yds range while the LT detected 21%; for the 1000 yds distance, MMOs detected 54% of pods outside 1000 yds range while the LT detected 13%.

As indicated earlier, there were only 2 sperm whale sightings so no exploratory analysis was performed.

The SCSP group included 178 sightings (Table 5), comprising 201 detections made by either or both positions (including first detections or subsequent resights). Out of the 178 sightings, 125 were detected by the MMO only, 20 by the LT only and 33 by both. This gives an index of effectiveness for the MMOs of $IE_{MMO} = 0.88$ and for the LT of $IE_{LT} = 0.30$, i.e., almost 3 times lower. Of the 33 encounters detected by both positions, five were detected first by the MMOs, four first by the LT and 24 by both positions during the same surfacing. 63 of the 178 sightings were identified to species, four to genus and 111 to higher taxonomic levels. We note that 20 sightings were excluded from the analysis as they were recorded as 'bowriding'; hence, the total number of sightings and detections included in the analyses were 178 and 201, respectively.

The location of first detections of SCSP relative to the ship is shown in Figure 4. The first detections were generally much closer to the ship than for rorquals. Indeed, many detections were made at distances that would indicate the pod was at or almost at the bow of the ship when first detected: 13.2% of LT first detections and 8.8% of MMO first detections were given recorded ranges of 10 yds or less. The percentage of first detections made within 200, 500 and 1000 yds of the ship was 52.8, 77.4, 88.7% for the LT and 32.3, 43.7, 74.1% for the MMOs.

The close ranges at which detections were first made meant that the distance-specific indices of effectiveness (Table 6) were lower for SCSP compared with rorquals. Of the pods first detected within 200 yds perpendicular distance from the vessel's track 25% of them were detected by the MMOs outside of the 200 yds mitigation range while only 3% were detected by the LT outside this range. For the 500 yds distance, MMOs detected 29% of pods outside 500 yds range while the LT detected 3%. For the 1000 yds distance, MMOs detected 14% of pods outside 1000 yds range while the LT detected 2%.

The SCLP group included 87 sightings (Table 7) and 136 detections by either or both positions. Out of the 87 sightings, 58 were by the MMOs only, nine by the LT only and 20 by both positions. This gives an index of effectiveness for the MMOs of $IE_{MMO} = 0.90$ and for the LT of $IE_{LT} = 0.33$, i.e., approximately 3 times lower. Of the 20 detected by both, six were detected first by the MMOs, one first by the LT and 13 by both simultaneously. Of the 87 sightings, 49 were identified to species, ten to genus and 28 to higher taxonomic levels. As for the SCSP group, pods recorded as bowriding were excluded from the analyses leaving 87 sightings and 94 detections in the analyzed data set.

The location of first detections of SCLP relative to the ship is shown in Figure 5. The percentage of pods first signed at very close range (10 yds or less) was again high for the LT at 17.2% for the LT but 0% for the MMOs. The percentage of first detections made within 200, 500 and 1000 yds of the ship was 48.3, 69.0 and 86.2% for the LT and 11.5, 24.4 and 46.2% for the MMOs. Combined with the very large proportion of detections made at close ranges, there was a significant tail of detections at larger ranges (Figure 5).

The distance-specific indices of effectiveness (Table 8) were generally somewhat higher for SCLP than the SCSP, although still very low for the LT. Of the pods first detected within 200 yds perpendicular distance from the vessel's track 41% of them were detected by the MMOs outside of the 200 yds mitigation range while only 6% were detected by the LT outside this range. For the 500 yds distance, MMOs detected 33% of pods outside 500 yds range while the LT detected 4%. For the 1000 yds distance, MMOs detected 31% of pods outside 1000 yds range while the LT detected 3%.

Simulation study

Visual comparison of observed vs expected estimates provides an informal means for checking the simulation. The true detection hazard from simulation scenario 2 (which is closer to our real data situation) for MMOs and the LT is shown on the top panel of Figure 1 and an example estimated hazard function from one simulated dataset is shown on the bottom panel. The two appear very similar. Likewise, Figure 2 shows the cumulative distribution of first detections in the 2-d plane as green dashed lines, and the estimated distribution from a single dataset as black lines and shading. They are again similar, although the black lines are within in the green dashed lines, indicating slight underestimation of detection probability at closer ranges. Further diagnostic plots are provided in Appendix C.

A more formal quantification of performance is the percentage bias calculated over the 100 simulated realizations. Bias in estimation of all quantities of interest was very low (<3%) for both scenario 1 (independent observers, all detections) and scenario 2 (one-way independence, first detection only) (Table 9).

Range-dependent probability of remaining undetected (PrU)

Full details of results, including additional diagnostic plots, goodness-of-fit test results and sensitivity analysis are given in Appendix C. An extended summary is given here.

The largest forward detection distance was 15,725 m and the largest perpendicular distance was 13,425 m; these were used as truncation distances in the analysis for all species groups.

The estimated detection hazard for the rorqual group was very flat and near zero in most areas for both positions, only exhibiting a sharp rise within small radii around zero radial distance from the ship (Figure 6). This increase in detection hazard was much more pronounced for MMOs compared to LT. The former also rose to much higher values at x=0, y=0 compared to the latter. The estimated cumulative distribution of first detections is shown in Figure 7. The value shown in Figure 7 at x=0, y=0 is the estimate of *g(0)* for this position, and this is also given in Table 10 together with the estimated effective strip half-widths (ESHWs) and the range-dependent probability of remaining undetected (PrU). For the LT, estimated PrU was 0.80 (95% confidence interval (CI) 0.74-0.86) at 200 yds, 0.85 (95% CI 0.80-0.89) at 500 yds, and 0.91 (95% CI 0.87-0.94) at 1000 yds. PrU for MMOs were lower (i.e., better): 0.49 (95% CI 0.40-0.59) at 200 yds, 0.53 (95%CI 0.43-0.62) at 500 yds and 0.59 (0.51-0.67) at 1000 yds. Overall these are around 1.6 times lower. Results from the χ^2 goodness-of-fit tests indicated that the fit of the rorquals model was poor for both MMO and LT data (Appendix C Table 12); visual inspection of diagnostic plots (Appendix C Figures 11 and 12) showed the model under-predicted MMO detections at close ranges in the perpendicular distance direction and over-predicted in the forward distance direction; for the LT the lack of fit appeared again to be under-prediction at close range in the perpendicular distance direction. The sensitivity analysis, which comprised a second run with alternative availability model parameters, produced almost identical results to the main analysis (Appendix C Table 11).

Taking the estimated detectability parameters for rorquals and applying them to sperm whales, where time spent underwater is considerably higher, led to the estimated cumulative distribution of distances shown in Figure 8, and detection statistics given in Table 10. As would be expected, predicted *g(0)* and ESHW are lower, and the PrUs are higher (Table 10). For the LT estimated PrU was 0.89 (95% CI 0.87-0.92) at 200 yds, 0.92 (0.89-0.94) at 500 yds and 0.95 (0.93-0.96) at 1,000 yds. MMO PrU for sperm whales was 0.77 (95% CI 0.74-0.80) at 200 yds, 0.78 (0.75-0.81) at 500 yds and 0.80 (95% CI 0.77-0.84) at 1,000 yds. Note that the difference between the LT and MMO PrUs was smaller in this case, because a major limitation on PrU comes from the dive behavior and that is the same for both positions. The sensitivity analysis, using two sets of divergent availability model parameters, produced somewhat different results (e.g., estimated PrU at 200 yds varying from 0.60 to 0.87 for MMOs and from 0.79 to 0.95 for the LT) (Appendix C Table 13).

For small cetaceans in small pods (1-6 individuals, SCSP), the estimated detection hazard was extremely "spiked" at small ranges (Figure 9). Most pods were estimated to remain undetected, and of those detected, the estimated cumulative distribution of first detections (Figure 10) showed that these detections are likely to be at very close ranges. The estimated PrUs were close to 1 for both observation positions and all mitigation ranges (Table 10). As with rorquals, the χ^2 goodness-of-fit tests indicated a poor fit for both MMO and LT data (Appendix C Table 17); however the diagnostic plots (Appendix C Figures 16 and 17) did not indicate any obvious problems, and showed that the large spike in detections in forward, perpendicular and radial distances appeared to be well captured by the model. The sensitivity analysis, which included 3 divergent sets of availability parameters, gave almost identical results to the main analysis (Appendix C Table 16).

For small cetaceans in large pods (7+ individuals, SCLP), the estimated detection hazard was also spiked at small ranges (Figure 11), although less so than for SCSP. The estimated cumulative distribution of first detections (Figure 12) indicated that most pods detected are first detected at distances within the mitigation ranges. For the LT estimated PrU was 0.94 (95% CI 0.91-1.00) at 200 yds, 0.98 (0.97-1.00) at 500 yds and 0.99 (0.99-1.00) at 1,000 yds. MMO PrU for SCLP was 0.83 (95% CI 0.74-0.90) at 200 yds, 0.93 (0.89-0.96) at 500 yds and 0.97 (95% CI 0.95-0.98) at 1,000 yds. Once again the χ^2 goodness-of-fit tests indicated a poor fit for both MMO and LT data (Appendix C Table 20); the diagnostic plots (Appendix C figures 20 and 21) showed that although the model did a good job of fitting the spike in detections at small forward and perpendicular distances, generally underfitted for ranges <6000 yds and overfitted the larger ranges. The sensitivity analysis, which comprised one set of alternative availability parameters, gave almost identical results to the main analysis (Appendix C Table 19).

Discussion

The main goal of this project was to quantify the effectiveness of the Navy LOs and other members of the LT in detecting marine mammals before they enter a set of specified mitigation ranges. To achieve this, the Navy expended considerable effort in deploying MMOs on board ships during training exercises, where they have generated data on their own detections and those of the LT. We developed the experimental protocol, and have developed new methods for analysis of the resulting data. One additional benefit of the new analytical methods is that they are applicable to estimating animal density from line transect surveys, which could potentially also aid in producing better estimates of marine mammal density for use in the Navy's Marine Species Density Database.

In analyzing the experimental data, we grouped species along taxonomic lines, aiming to put together taxa with similar sightability and diving behavior. This gave us four groups: rorquals, sperm whales, SCSP and SCLP.

Rorquals are large whales often with conspicuous blows, and can therefore be relatively easy to sight when surfacing compared with small cetaceans – although the mean pod sizes were small (means between 1-2 individuals, depending on the species, Table 3). Given the forward truncation distance used in our analyses (15,725m) and assumed ship speed (12 knots) a patch of water at perpendicular distance x=0 would be in view for 42 mins. During this time, using the assumed dive parameters used in the main analysis for rorquals (2.4 mins in state 1 with inter-surfacing interval 0.4 mins, 4.2 mins in state 2), the expected number of surfacings while in view is 37. We found that, for MMOs, the probability of detecting an animal at x=0, g(0), was 0.53 (95% CI 0.43-0.62). This is lower than the estimate of 0.92 derived by Barlow and Forney (2007) for "large whales (most baleen whales and killer whales)" on NOAA line transect surveys. However, Barlow and Forney acknowledged that their methods tended to produce over-estimates; also NOAA survey data is collected only in good survey conditions while the experimental trials reported here took place in all sea state and sightability conditions.

We found for rorquals in our exploratory analysis that the LT detected only 23% of the surfacings known to have taken place within visual range, while the MMOs detected 93% of them. This leads to a naïve conclusion that the LT is 4 times less effective than the MMOs at detecting roquals. However, this is not directly relevant to mitigation effectiveness as it ignores range (the LT are unlikely to be scanning for objects at large perpendicular distances, for example) and pods unobserved by both teams. A more refined exploratory analysis was undertaken, examining all detections made within 200, 500 and 1,000 yds perpendicular distance from the ship track (assuming the ship continued in a straight line) and determining what proportion of these were detected by the MMOs and the LT before entering the a 200, 500 and 1,000 yard radius from the ship. In every case, the proportion of detections by the MMOs was substantially higher than that for the LT – for example for 500 yds the MMOs detected 70% of pods before they entered this mitigation range while the LT detected 21% – 3 times worse. This appeared to be due to a combination of two factors: the LT detected far fewer pods overall and also tended to make a larger proportion of their detections within the mitigation ranges rather than outside of them, as would be required for effective mitigaton.

The relative performance of MMOs and LT was not of primary interest, partly because the MMO team was typically larger and also they were tasked only with detecting animals in the water, while the LT had other responsibilities. The primary interest was in estimating the absolute effectiveness of the LT, and the simple analyses discussed in the previous paragraph may provide an upper estimate of effectiveness because they do not account for animals missed by both positions. One caveat, however, is that the distance-specific indices do assume that the ship moves in a straight line and that animals do not move in the perpendicular direction – these assumptions are required for it to be true that pods seen at perpendicular distance x will pass within range x of the ship as it passes by. Some animals seen by the MMOs far ahead of the ship may have moved away before the ship passed, and so not entered within mitigation range.

The more sophisticated modelling exercise was intended to address the issue of animals undetected by both positions, and produce absolute estimates of PrU, which can be seen as the complement of effectiveness. For example, the estimate of PrU at 500 yds for of 0.53 for the MMOs and 0.85 for the LT mean that the MMOs are estimated to have had a (1-0.53)x100 = 47% chance of detecting a rorqual pod before it reached the 500 yard boundary while the LT had a (1-0.85)x100 = 15% chance. If the logic of the previous paragraph is correct that the previously-quoted simple distance-based indices of effectiveness provide an upper bound then we would expect these model-based estimates to be lower. Indeed they are in both cases (47% vs 70% for MMOs and 15% vs 21% for the LT), and also for all cases across all three sets of mitigation ranges. This provides some reassurance that the analyses are producing estimates that are, at least approximately, correct.

The analysis leading to estimation of PrU made the following assumptions:

1. the ship travels at 12 knots in a straight line;
2. pods do not move in the horizontal plane;
3. pods are uniformly distributed with respect to perpendicular distance from the ship;
4. time in near-surface and deep dive phases follow exponential distributions with known parameters, as does the inter-surfacing interval while in the near-surface state;
5. dive behavior is not affected by presence of the ship;
6. pod location and taxon are recorded accurately;

7. all first detections of a sighting are recorded, and MMOs do not cue the LT.

The effect of violation of these assumptions on estimation of PrU is typically not intuitively obvious. It could be studied via simulation. In the present study assumption 1 is likely broken. For rorquals assumption 2 may be mildly violated, although pods typically move slowly compared with ship speed; there is in some cases known to be avoidance behavior. There is no reason to suspect assumption 3 is violated for rorquals. For assumption 4, the grouping of species together means there is certainly variation in dive behavior within the group. The sensitivity analysis performed gave very similar results, so it may be that variation in dive behavior is not of primary importance for this group. Assumption 5 is likely violated, and warrants further study. For assumption 6, we have not attempted to look at measurement accuracy but one avenue here would be to compare LT and MMO records. We did encounter some difficulties in processing the data collected, and return to this below. In addition there is some rounding of distances, and possibly angles, evident in the data (see Figure 3). Assumption 7 appears to have been met.

Overall, we judge that our results for rorquals are broadly robust, but further investigations could be taken to assess this. One concern is the poor goodness-of-fit (as with the other two taxa analyzed), and we return to this below.

For the sperm whale taxon, we additionally assumed that the rorqual detection hazard parameters apply to this taxon. This enabled us to derive estimates of PrU, and illustrates an approach that might be applied to other taxa such as beaked whales if a suitable proxy taxon can be identified for which detectability of surfacings is measurable. Our findings that PrU for this taxon is higher than rorquals makes intuitive sense, as does the finding that PrU is more similar between the LT and MMOs because it is dominated for both observation positions by the lower surface availability. The sensitivity analysis showed that estimates of PrU and hence absolute effectiveness were somewhat influenced by the availability parameters, and hence more work is needed to determine what parameters are most reasonable for these long, deep divers.

For SCSP, the great majority of first detections were made within the mitigation ranges (nearly 90% of LT and 74% of MMO first detections within 1,000 yds of the ship). This clearly indicates that neither position were able to detect pods before they entered the mitigation ranges. The distance-specific indices of effectiveness for MMOs ranged from 25% at 200 yds to 14% at 1,000 yds and for the LT from 3% at 200 yds to 2% at 1,000 yds. It is important to note that these results are after excluding pods noted to have been bowriding, the majority of which were first detected well within 200 yds and so would have made these effectiveness figures even lower had they been included.

The model-based estimates of PrU for SCSP were not satisfactory, being unrealistically high (close to 1) for both positions and all mitigation ranges. This was likely caused by the significant proportion of first detections (around 10%) that were recorded at very close ranges (10 yds or less), leading to extremely "spiked" estimated detection hazard functions. The recorded very close ranges may be a combination of fast and responsive movement in some pods, violating assumptions 2 (no horizontal movement) and 3 (uniform distribution with respect to perpendicular distance). It may be that rounding of distances also contributed to the estimation difficulties. Despite these issues, a highly spiked detection hazard may not be unrealistic: for example Roberts et al. (2015 Figure 13) present a perpendicular distance detection function fitted to over 500 sightings of bottlenose dolphin from line transect surveys that exhibits a similarly spiked shape.

For SCLP, the great majority of first detections made by the LT (nearly 90%) were again within the mitigation ranges, while for the MMOs less than half (46.2%) were seen within 1,000 yds. This led to the LT having similar very low distance-specific indices of effectiveness to SCSP (e.g., 6% at 200 yds) but the MMOs having higher estimated effectiveness (e.g., 41% at 200 yds).

The model-based estimates of PrU for SCLP were very high for both platforms, although not much higher than would be indicated from the distance-specific indices for the LT. Like the SCSP, the fit was strongly influenced by the preponderance of first detections recorded at very close ranges, for the LT, where 17% of first detections were recorded as being at 10 yds or less from the ship. Similar problems of fast movement and rounding error are likely present in this group as for the SCSP.

Given the above, we conclude that effectiveness in detecting small cetaceans before they reach even the 200 yard mitigation range is very low for the LT, while for the MMOs it is likely to be low for small pods and somewhat higher for larger pods.

If more quantitative estimates were required, it may be possible to extend the analysis methods to account for animal movement and rounding error. Another factor potentially causing lack of fit is the assumption that detection hazard is the same in all directions. For rorquals at least, there is some evidence from the diagnostic plots in Appendix C that there are more detections at close perpendicular distances and fewer at close forward distances than predicted by the model – implying that observers may be searching further away in the forward distance direction than abeam. This may make sense for observers seeking to detect objects ahead of the ship, and so allowing different detection hazard parameters in the forward and perpendicular distance directions may improve model fit.

In preparing this report, we were required to expend considerable effort on data checking and cleaning due to inconsistencies in data recording. Some data had to be discarded. Thomas et al. (2012) recommended that, in advance of any further data collection, MMOs liaise with our group to discuss experimental data collection and recording protocols and that exploratory analyses are undertaken by way of data validation after every cruise. We again make this recommendation which we believe, if implemented, would reduce data loss and lead to better data quality in future.

Acknowledgements

We thank the dedicated MMOs who collected the field survey data, as well as the ships' commanders and crews who participated in the trials. Jay Barlow, Debbie Palka, and Lance Garrison of the US National Marine Fisheries Service provided helpful advice during construction of the experimental protocols. From within CREEM, David Borchers, Louise Burt, Roland Langrock and Eric Rexstad all provided valuable input into earlier parts of the project, as documented in previous reports.

Literature cited

Barlow J. and K.A. Forney. 2007. Abundance and density of cetaceans in the California Current ecosystem. Fish. Bull. 105: 509-526.

Borchers, D.L., and R. Langrock. 2015. Double-observer line transect surveys with Markov-modulated Poisson process models for animal availability. Biometrics 71 (4): 1060–9.

Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas. 2001. Introduction to Distance Sampling. Oxford University Press.

Burt, M.L. and L. Thomas. 2010. Calibrating US Navy lookout observer effectiveness: information for marine mammal observers, Version 2.1. Document submitted as part of Technical report IAW CDRL A003. University of St Andrews.

Department of the Navy. 2016. U.S. Navy Lookout Effectiveness Study Marine Mammal Observer Survey Protocol. August 2016. (Included as Appendix B of this report.)

Dolphin, W.F. 1987. Ventilation and dive patterns of humpback whales, *Megaptera novaeangliae*, on their alaskan feeding grounds. Canadian Journal of Zoology 65 (1): 83–90.

Drouot, V., A. Gannier, and J. Goold. 2004. Diving and feeding behaviour of sperm whales (*Physeter macrocephalus*) in the northwestern Mediterranean Sea." Aquatic Mammals 30: 419–26.

Kinzey, D., P. Olson and T. Gerrodette. 2000. Marine mammal data collection procedures on research ship line-transect surveys by the Southwest Fisheries Science Center. Southwest Fisheries Science Center, Administrative Report LJ-00-08, 32 p.

Langrock, R., D.L. Borchers, and H.J. Skaug. 2013. Markov-modulated nonhomogeneous Poisson processes for modeling detections in surveys of marine mammal abundance. Journal of the American Statistical Association 108 (503): 840–51.

NMFS. 2019. Letter of Authorization for US Navy training activities in the Atlantic Fleet Training and Testing (AFTT) Study Area.

NMFS. 2020. Letter of Authorization for US Navy training activities in the Hawaii-Southern California Training and Testing (HSTT) Study Area

Rexstad, E., and L. Thomas. 2010a. Preliminary analysis of observer effectiveness data gathered from FFG-A, DDG-A, DDG-B and DDG-C cruises. Technical report IAW CDRL A002. University of St Andrews.

Rexstad, E., and L. Thomas. 2010b. Observer effectiveness study – results from FFG-A, DDG-A, DDG-B and DDG-C cruises – Task C final survey protocol report. Technical report IAW CDRL A003. University of St Andrews.

Roberts J.J., Best B.D., Mannocci L., Fujioka E., Halpin P.N., Palka D.L., Garrison L.P., Mullin K.D., Cole T.V.N., Khan C.B., McLellan W.M., Pabst D.A. and Lockhart G.G. 2015. Density Model for Bottlenose Dolphin (*Tursiops truncatus*) for the U.S. Gulf of Mexico Version 3.3, 2015-10-07, and Supplementary Report. Marine Geospatial Ecology Lab, Duke University, Durham, North CarolinaScott, M., and S. Chivers. 2009. Movements and diving behavior of belagic spotted dolphins. Publications, Agencies; Staff of the US Department of Commerce. 46.

Thomas, L., E. Rexstad and D.L. Borchers. 2011. Calibrating US Navy lookout observer effectiveness – Task D: Final Data Analysis. Technical report IAW CDRL A004. University of St Andrews.

Thomas, L., R. Langrock, E. Rexstad and D.L. Borchers. 2012. Lookout effectiveness analysis development. Final Contract Report. HDR Task Order 002-0000105055. University of St Andrews.

Tables and Figures

Table 1. Cruises included in this study. Ship codes starting with CG are cruisers, DDG are destroyers and FFG are frigates. Study areas included Atlantic Fleet Training and Testing (AFTT), Hawaii Rance Complex (HRC) and the Southern California Range Complex (SOCAL).

Ship	Month	Year	# sightings	Study area
FFG-A	Feb	2010	21	HRC
DDG-A	Mar	2010	11	AFTT
DDG-B	Jun	2010	15	AFTT
DDG-C	Jul	2010	84	SOCAL
CG-A	Nov	2010	7	HRC
DDG-D	Feb	2011	29	HRC
DDG-E	Apr	2011	21	SOCAL
DDG-F	Nov	2011	5	HRC
DDG-G	Feb	2012	13	HRC
FFG-B	May/Jun	2012	24	AFTT
DDG-H	Jul	2012	62	SOCAL
DDG-I	Feb	2013	5	HRC
DDG-J	Aug	2013	2	HRC
DDG-K	Jan	2014	57	HRC
CG-B	Feb	2014	7	HRC
CG-C	Aug	2014	23	AFTT
DDG-L	Feb	2015	34	HRC
DDG-M	Apr	2015	3	AFTT
DDG-N	Feb	2016	12	HRC
DDG-O	Mar/Apr	2016	52	AFTT
DDG-P	Aug	2016	44	AFTT
DDG-Q	Aug	2017	56	AFTT
DDG-R	Feb	2018	22	HRC
DDG-S	Jun	2018	34	AFTT
DDG-T	Feb	2019	30	HRC
CG-D	Mar	2019	15	AFTT
DDG-U	Sep	2019	28	AFTT
Cruises	**27**	**Total**	**716**	

Table 2. Summary of sightings from lookout effectiveness field surveys. The hierarchy of taxon codes is given in Appendix A.

Taxon	Scientific name	Common name	# sightings	Mean pod size	SD pod size	# det by MMO only	# det by LO only	#det by both	# both: MMO first	# both: LO first	# both: same time
BAL--	Balaenoptera sp.	--	29	1.27	0.53	23	4	2.00	1	0	1
BALAC	Balaenoptera acutorostrata	Minke whale	3	1	0	1	0	2.00	1	0	1
BALMU	Balaenoptera musculus	Blue whale	16	2	0.91	12	0	4.00	3	0	1
BALPH	Balaenoptera physalus	Fin whale	1	1	NA	1	0	0.00	0	0	0
BLACK	--	Blackfish	2	2	0	2	0	0.00	0	0	0
CARCA	Caretta caretta	Loggerhead turtle	41	1	0	36	1	4.00	0	1	3
CET--	Unidentified cetacean		9	3.11	2.85	8	1	0.00	0	0	0
CHEMY	Chelonia mydas	Green turtle	12	1.08	0.29	10	0	2.00	0	0	2
DEL--	Delphinus sp.	Unid. common dolphin	11	92.32	109.03	8	0	3.00	2	0	1
DELCA	Delphinus capensis	Long-beaked common dolphin	2	42.5	38.89	1	0	1.00	1	0	0
DELDE	Delphinus delphis	Short-beaked common dolphin	11	27	38	8	0	3	1	0	2
DERCO	Dermochelys coriacea	Leatherback turtle	1	1	NA	0	0	1.00	0	0	1
DOLPH	--	Unid. dolphin	129	3.99	4.04	100	17	12.00	2	2	8
GLO--	Globicephala sp.	Unid. pilot whale	3	3.67	2.08	2	0	1.00	0	0	1
GLOMA	Globicephala macrorhynchus	Short-finned pilot whale	15	12.93	11.96	11	0	4.00	1	2	1
GRAGR	Grampus griseus	Risso's dolphin	7	13.71	12.42	6	0	1.00	0	0	1
LAGAC	Lagenorhynchus acutus	Atlantic white-sided dolphin	1	8	NA	1	0	0.00	0	0	0
LEPKE	Lepidochelys kempii	Kemp's ridley turtle	1	1	NA	1	0	0.00	0	0	0
LGWHA	--	Unid. large whale	50	1.32	0.56	44	3	3.00	1	0	2
MEGNO	Megaptera novaeangliae	Humpback whale	84	1.98	1	58	3	23.00	10	3	10
MIXED	--	--	6	31.37	36.35	2	1	3.00	1	0	2
MOLMO	Mola mola	Pelagic sunfish	1	1	NA	1	0	0.00	0	0	0
ORCOR	Orcinus orca	Killer whale	1	1	NA	1	0	0.00	0	0	0
PHOPO	Phocoena phocoena	Harbor porpoise	1	1	NA	1	0	0.00	0	0	0
PHYMA	Physeter microcephalus	Sperm whale	2	1	NA	2	0	0.00	0	0	0
SMALL	--	Unid. small cetacean	8	6.69	4.17	5	0	3.00	0	0	3

STEAT	Stenella attenuata	Pantropical spotted dolphin	1	35	NA	1	0	0.00	0	0	0
STEBR	Steno bredanensis	Rough-toothed dolphin	1	NaN	NA	1	0	0.00	0	0	0
STECO	Stenella coeruleoalba	Striped dolphin	3	8.83	4.37	2	0	1.00	1	0	0
STEFR	Stenella frontalis	Atlantic spotted dolphin	30	6.77	8.04	20	0	10.00	0	0	10
STELO	Stenella longirostris	Spinner dolphin	1	45	NA	0	0	1.00	0	0	1
TURTL	--	Unid. turtle	33	14.61	43.51	29	3	1.00	0	0	1
TURTR	Tursiops truncatus	Common bottlenose dolphin	24	6.96	9.91	14	1	9.00	2	0	7
UN-MM	--	Unid. marine mammal	3	1.33	0.58	3	0	0.00	0	0	0
UNID	--	Unid. animal	1	1	NA	1	0	0.00	0	0	0
UNOTA	--	Unid. eared seal	5	1.2	0.45	2	1	2.00	0	0	2
UNPIN	--	Unid. pinniped	13	1.5	0.53	8	3	2.00	1	0	1
WHALE	--	Unid. whale	94	1.4	0.72	73	11	10.00	3	2	5
ZALCA	Zalophus californianus	California sea lion	14	1.23	0.6	8	0	6.00	0	0	6
Total			**670**			**507**	**49**	**114**	**31**	**10**	**73**

Table 3. Summary of sightings included in the rorquals group from lookout effectiveness field surveys. The hierarchy of taxon codes is given in Appendix A.

Taxon	# sightings	Mean pod size	SD pod size	# det by MMO only	# det by LO only	#det by both	# both: MMO first	# both: LO first	# both: same time
BAL--	29	1.27	0.53	23	4	2	1	0	1
BALAC	3	1	0	1	0	2	1	0	1
BALMU	16	2	0.91	12	0	4	3	0	1
BALPH	1	1	NA	1	0	0	0	0	0
LGWHA	50	1.32	0.56	44	3	3	1	0	2
MEGNO	84	1.98	1	58	3	23	10	3	10
WHALE	94	1.4	0.72	73	11	10	3	2	5
Total	**277**			**212**	**21**	**44**	**19**	**5**	**20**

Table 4. Distance-specific indices of effectiveness for the rorquals group. The index of effectiveness $IE_p(x)$ is the proportion of pods thought to have entered within mitigation range x, n_x, that were successfully detected by position p before they entered within that range, $n_{p,r>x}$. The calculation of n_x does not account for pods not detected by either position.

	MMO			LT	
	#trials n_x	#successes $n_{MMO,r>x}$	effectiveness $IE_{MMO}(x)$	#successes $n_{LT,r>x}$	effectiveness $IE_{LR}(x)$
x=200 yds	34	25	0.74	12	0.35
x=500 yds	58	41	0.70	12	0.21
x=1000 yds	104	57	0.54	14	0.13

Table 5. Summary of sightings included in the SCSP group from lookout effectiveness field surveys. The hierarchy of taxon codes is given in Appendix A.

Taxon	# sightings	Mean pod size	SD pod size	# det by MMO only	# det by LO only	#det by both	# both: MMO first	# both: LO first	# both: same time
BLACK	2	2	0	2	0	0	0	0	0
DEL--	1	1	NA	1	0	0	0	0	0
DELDE	4	2	1.41	4	0	0	0	0	0
DOLPH	103	2.59	1.57	79	13	11	2	2	7
GLO--	3	3.67	2.08	2	0	1	0	0	1
GLOMA	5	3.5	1.94	3	0	2	1	1	0
GRAGR	2	3.5	2.12	2	0	0	0	0	0
MIXED	2	4.17	2.59	0	1	1	0	0	1
ORCOR	1	1	NA	1	0	0	0	0	0
PHOPO	1	1	NA	1	0	0	0	0	0
SMALL	4	2.88	0.85	1	0	3	0	0	3
STECO	1	4	NA	1	0	0	0	0	0
STEFR	28	2.57	1.75	16	4	8	0	1	7
TURTR	21	3.48	1.78	12	2	7	2	0	5
Total	**178**			**125**	**20**	**33**	**5**	**4**	**24**

Table 6. Distance-specific indices of effectiveness for the SCSP group. The index of effectiveness $IE_p(x)$ is the proportion of pods thought to have entered within mitigation range x, n_x, that were successfully detected by position p before they entered within that range, $n_{p,r>x}$. The calculation of n_x does not account for pods not detected by either position.

		MMO		LT	
	#trials n_x	#successes $n_{MMO,r>x}$	effectiveness $IE_{MMO}(x)$	#successes $n_{LT,r>x}$	effectiveness $IE_{LR}(x)$
x=200 yds	102	25	0.25	3	0.03
x=500 yds	144	42	0.29	5	0.03
x=1000 yds	181	26	0.14	3	0.02

Table 7. Summary of sightings included in the SCLP group from lookout effectiveness field surveys. The hierarchy of taxon codes is given in Appendix A.

Taxon	# sightings	Mean pod size	SD pod size	# det by MMO only	# det by LO only	#det by both	# both: MMO first	# both: LO first	# both: same time
DEL--	10	101.45	110.41	7	0	3	2	0	1
DELCA	2	42.5	38.89	1	0	1	1	0	0
DELDE	7	40.64	42.79	4	0	3	1	0	2
DOLPH	21	11.69	5.15	18	2	1	0	0	1
GLOMA	10	17.64	12.12	8	0	2	0	1	1
GRAGR	5	17.8	12.54	4	0	1	0	0	1
LAGAC	1	8	NA	1	0	0	0	0	0
MIXED	3	49.5	37.49	2	0	1	1	0	0
SMALL	4	10.5	1	4	0	0	0	0	0
STEAT	1	35	NA	1	0	0	0	0	0
STECO	2	11.25	1.77	1	0	1	1	0	0
STEFR	13	15.85	8.36	5	5	3	0	0	3
STELO	2	26	26.87	0	1	1	0	0	1
TURTR	6	24	15.52	2	1	3	0	0	3
Total	**87**			**58**	**9**	**20**	**6**	**1**	**13**

Table 8. Distance-specific indices of effectiveness for the SCLP group. The index of effectiveness $IE_p(x)$ is the proportion of pods thought to have entered within mitigation range x, n_x, that were successfully detected by position p before they entered within that range, $n_{p,r>x}$. The calculation of n_x does not account for pods not detected by either position.

		MMO		LT	
	#trials n_x	#successes $n_{MMO,r>x}$	effectiveness $IE_{MMO}(x)$	#successes $n_{LT,r>x}$	effectiveness $IE_{LR}(x)$
x=200 yds	34	14	0.41	2	0.06
x=500 yds	49	16	0.33	2	0.04
x=1000 yds	74	23	0.31	2	0.03

Table 9. Simulation study results, showing mean percent bias in estimates of the effective strip half-width (ESHW), the trackline detection probability (g(0)) and the probability of remaining undetected (PrU) calculated for 200 yds, 500 yds and 1,000 yds for simulation scenarios 1 (two-way independence, all detections) and 2 (one-way independence, first detections only). MMO is marine mammal observers and LT is lookout team.

	Scenario 1: two-way independence all detections		Scenario 2: one-way independence first detection only	
	MMO	LT	MMO	LT
ESHW	0.1	-0.3	-1.5	-1.2
g(0)	-0.2	-1.9	-0.6	-2.5
PrU 200 yds	0.4	1.9	1.4	2.6
PrU 500 yds	0.3	1.7	1.4	2.4
PrU 1,000 yds	0.2	1.4	1.4	2.0

Table 10. *Survey data results, showing estimated effective strip half-width (ESHW), the trackline detection probability (g(0)) and the probability of remaining undetected (PrU) calculated at 200 yds, 500 yds and 1,000 yds for four cetacean taxa. MMO is marine mammal observers and LT is lookout team. Values in brackets are 95% confidence intervals. (Note, estimated PrUs for small cetaceans are 1.00 when rounded to 2 decimal places, but are denoted 0.99 to indicate that they are not exactly 1.)*

	Rorqual		Sperm whale		Small cetaceans in small pods (6 or fewer)		Small cetaceans in large pods (more than 6)	
	MMO	LT	MMO	LT	MMO	LT	MMO	LT
ESHW	1739 m (1396-2126)	408 m (310-579)	886 m (733-1060)	234 (178-318)	0.66 m (0.26-1.63)	0.192 m (0.069-0.481)	240 m (137-363)	70 m (0-109)
g(0)	0.53 (0.43-0.62)	0.24 (0.16-0.31)	0.24 (0.20-0.27)	0.12 (0.09-0.154)	0.0027 (0.0011-0.0064)	0.0011 (0.0004-0.0027)	0.49 (0.31-0.68)	0.25 (0.00-0.37)
PrU 200 yds	0.49 (0.40-0.59)	0.80 (0.74-0.86)	0.77 (0.74-0.80)	0.89 (0.87-0.92)	0.99 (0.99-0.99)	0.99 (0.99-1.00)	0.83 (0.74-0.90)	0.94 (0.91-1.00)
PrU 500 yds	0.53 (0.43-0.62)	0.85 (0.80-0.89)	0.78 (0.75-0.81)	0.92 (0.89-0.94)	0.99 (0.99-0.99)	1.00 (0.99-1.00)	0.93 (0.89-0.96)	0.98 (0.97-1.00)
PrU 1,000 yds	0.59 (0.51-0.67)	0.91 (0.87-0.94)	0.80 (0.77-0.84)	0.95 (0.93-0.96)	1.0000 (0.99-1.00)	1.00 (1.00-1.00)	0.97 (0.95-0.98)	0.99 (0.99-1.00)

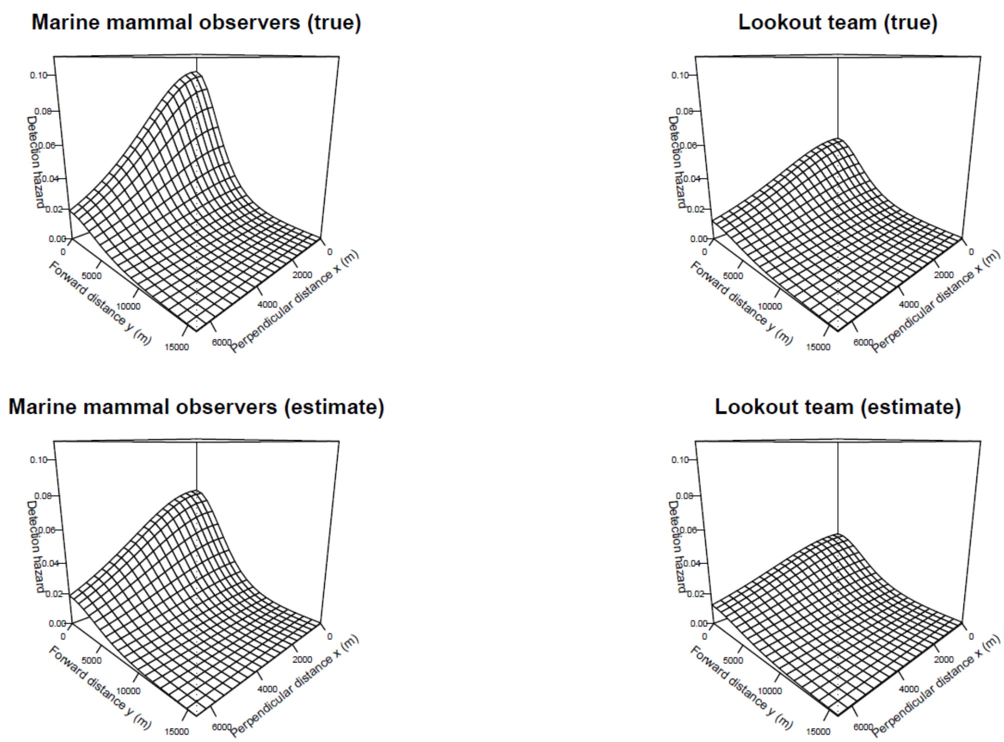


Figure 1. True detection hazard from the simulation study (top) and example estimated detection hazard from one example dataset generated under scenario 2 (bottom).

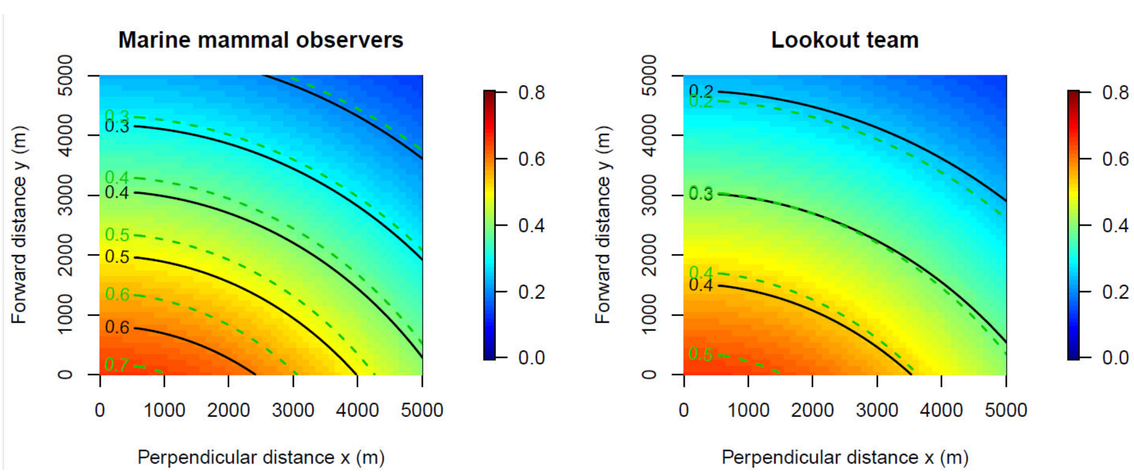


Figure 2. Cumulative distribution of distance to first detection from simulation scenario 2. Green dashed lines are the true distribution (i.e., calculated with the true simulation parameters); black lines and shading show estimated distribution from an analysis of one example dataset generated from the simulation.

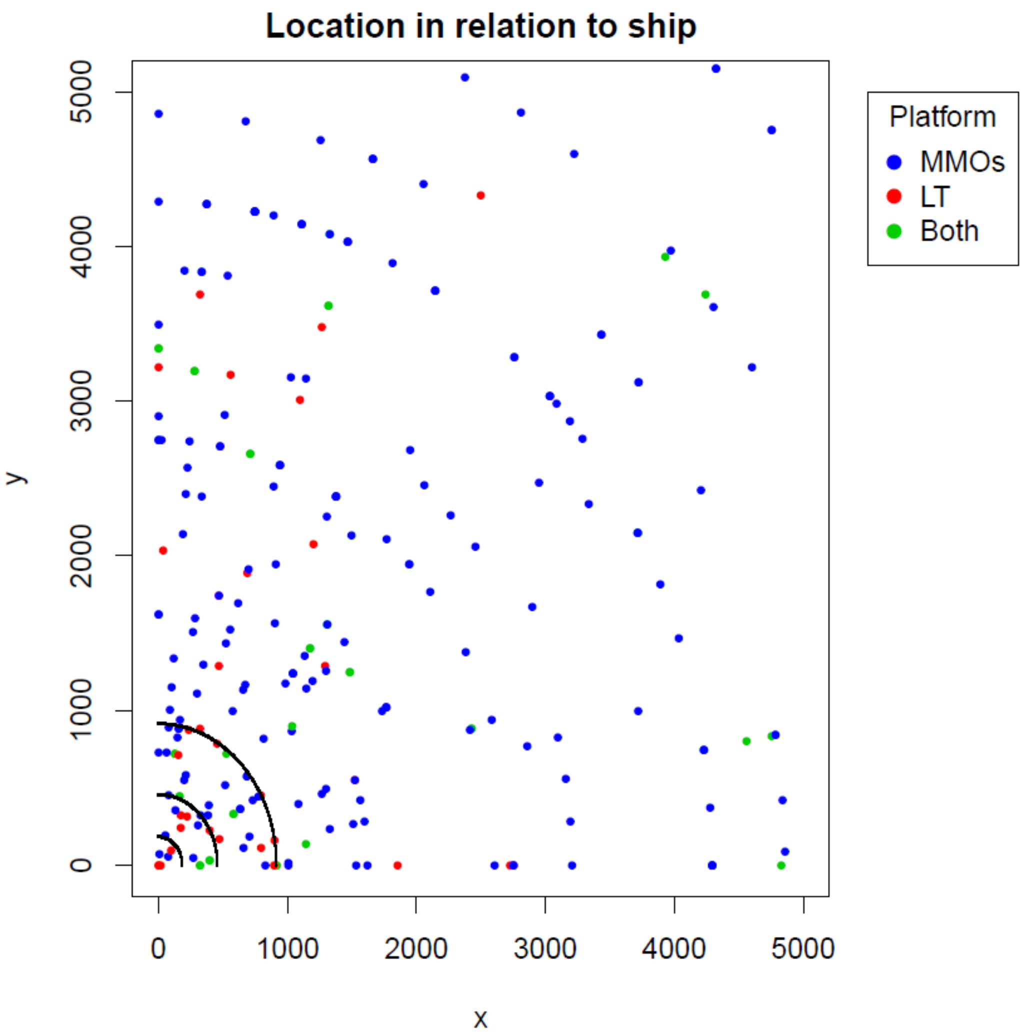


Figure 3. Location of rorqual first detections relative to ship location at (0, 0). Lines represent the three mitigation ranges of 200, 500 and 1,000 yards.

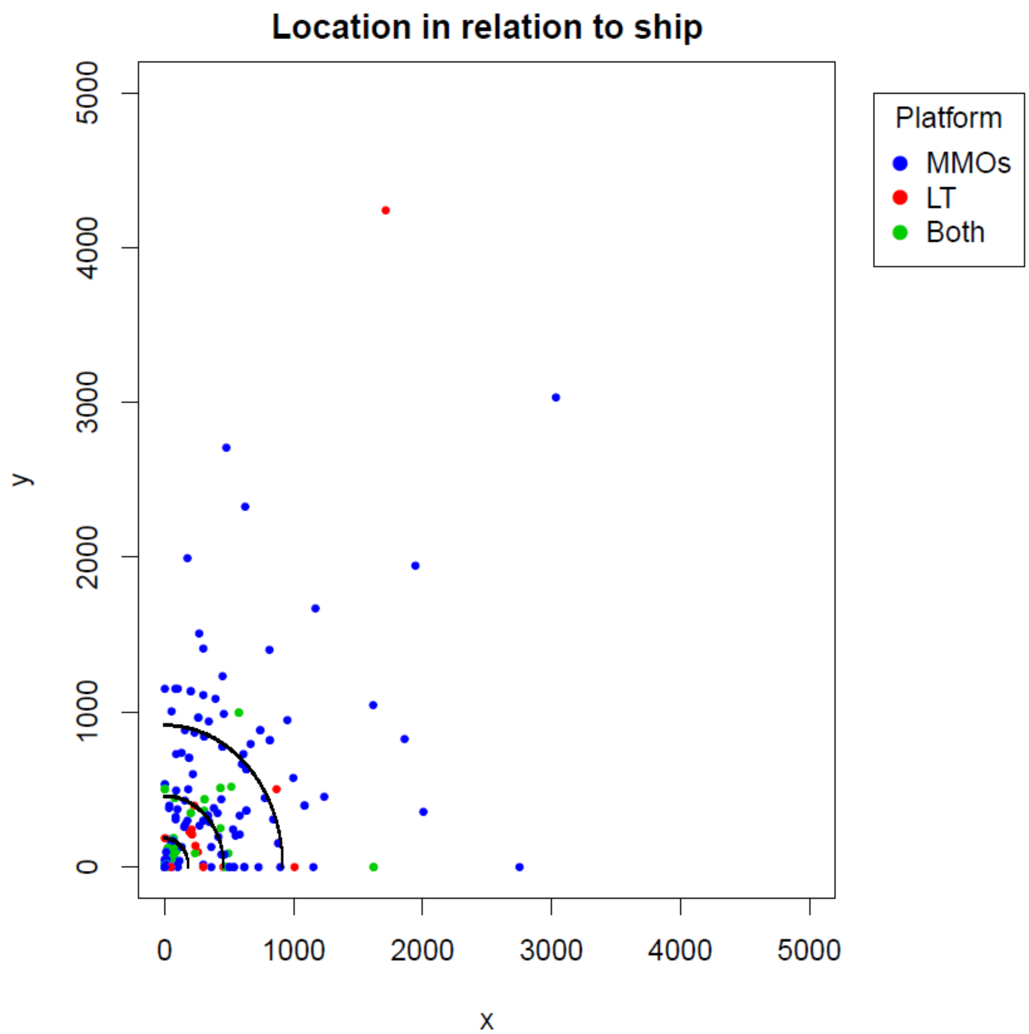


Figure 4. Location of small cetacean small pods (SCSP) first detections relative to ship location at (0, 0). Lines represent the three mitigation ranges of 200, 500 and 1,000 yards.

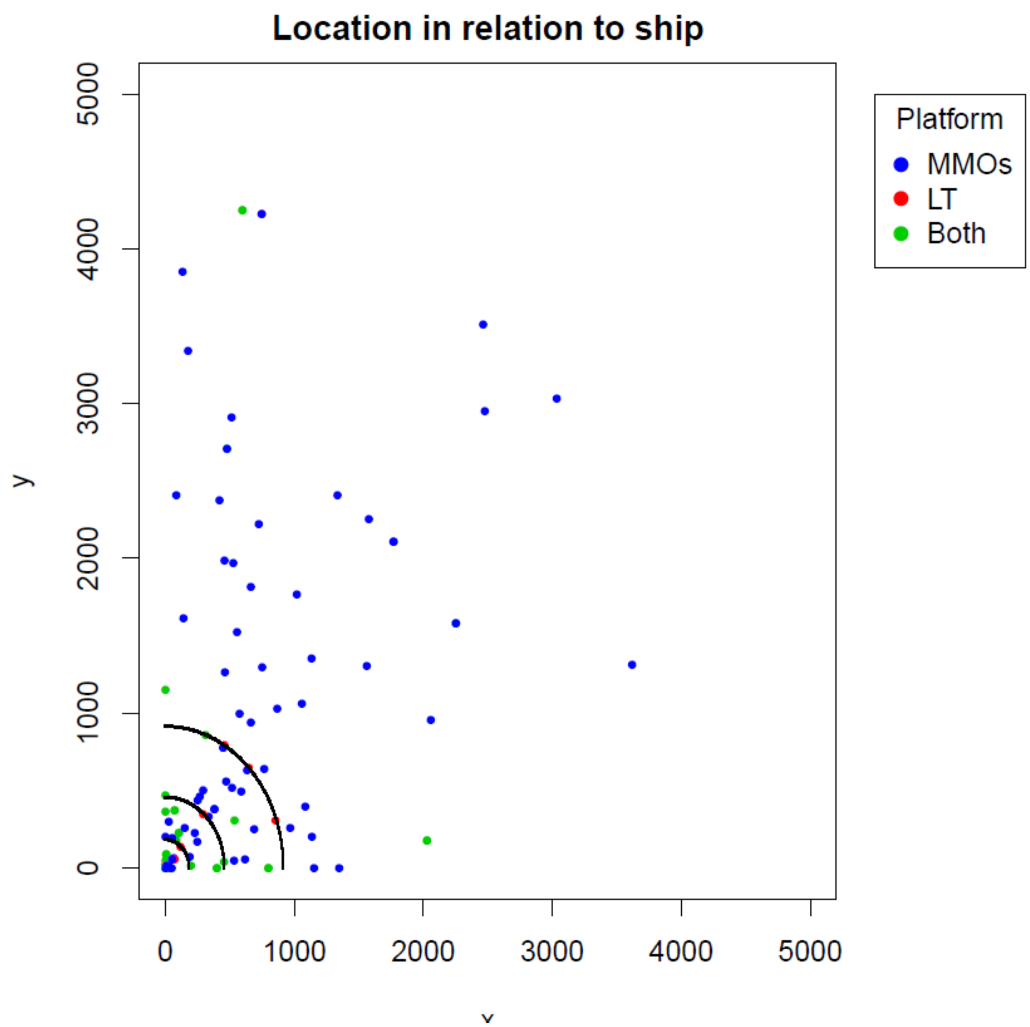


Figure 5. Location of small cetacean large pods (SCLP) first detections relative to ship location at (0, 0). Lines represent the three mitigation ranges of 200, 500 and 1,000 yards.

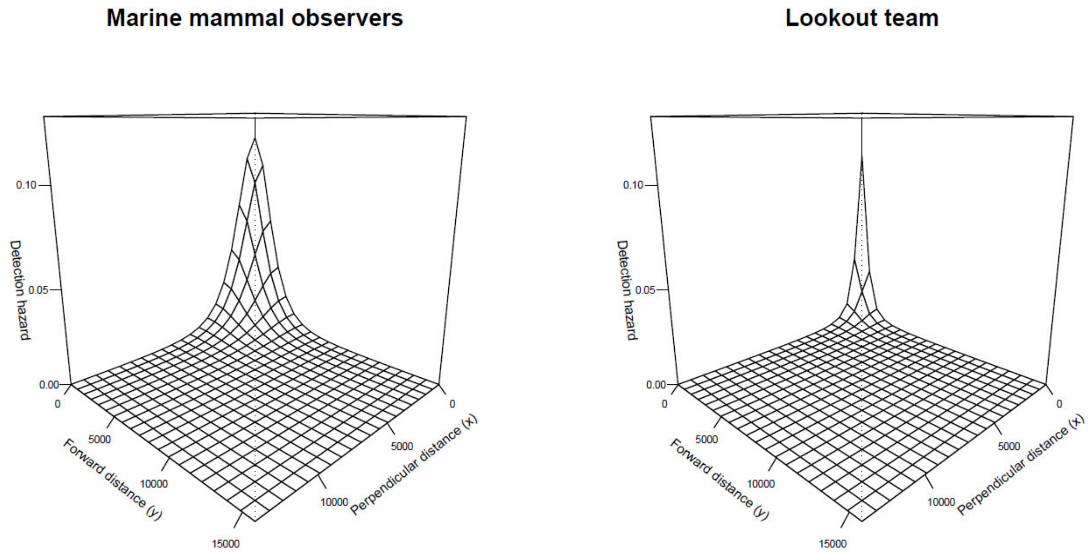


Figure 6. Estimated detection hazard of a surfacing from model fitted to rorqual group.

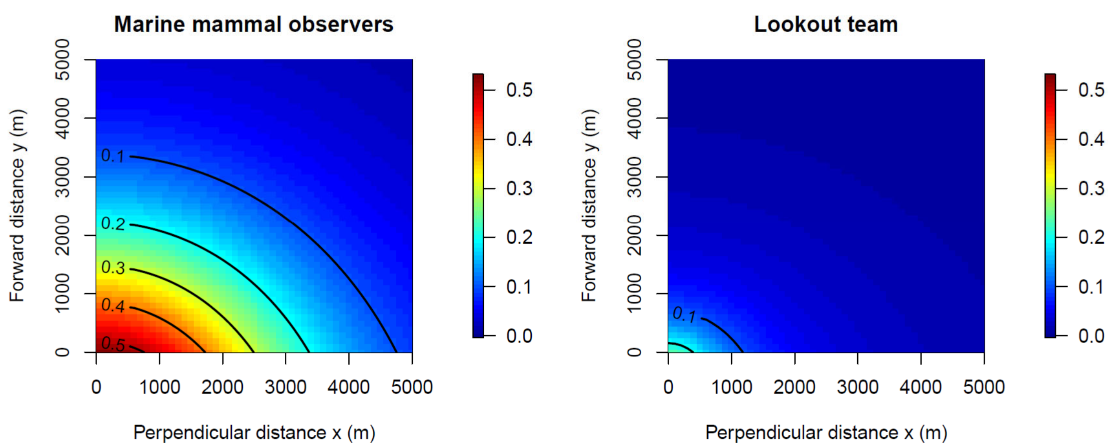


Figure 7. Estimated cumulative distribution of distances to first detection from model fitted to the rorqual group.

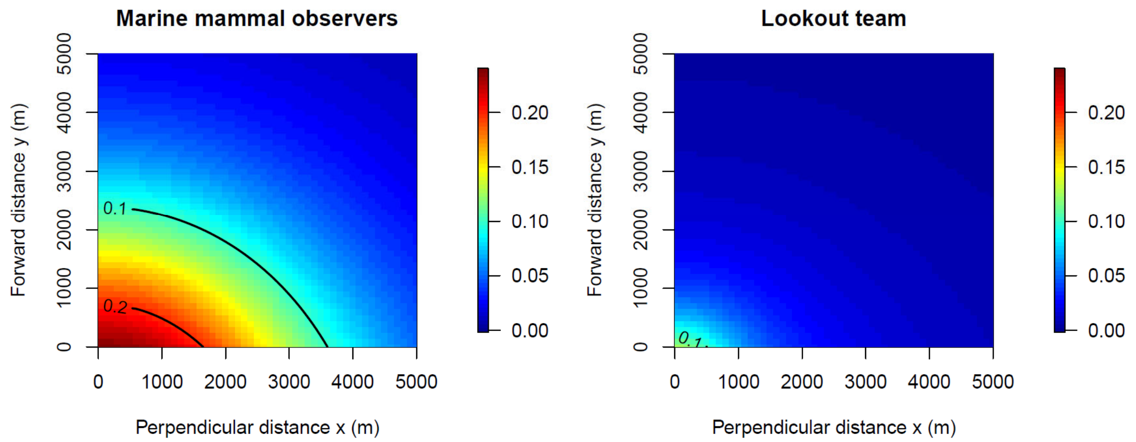


Figure 8. Estimated cumulative distribution of distances to first detection for sperm whales calculated using the estimated detection hazard from the model fitted to the rorqual group.

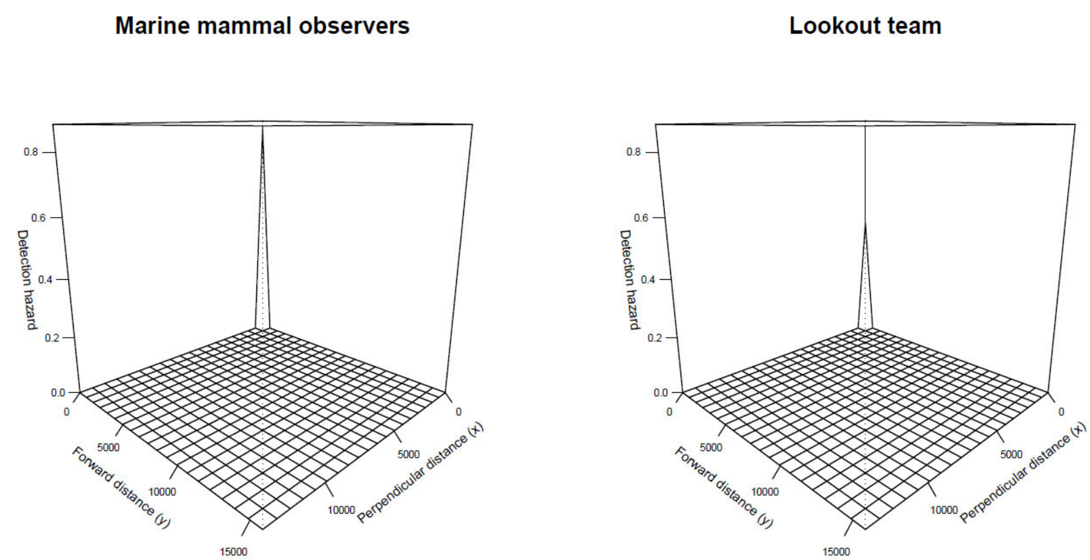


Figure 9. Estimated detection hazard of a surfacing from model fitted to the small cetaceans in small pods (SCSP) group.

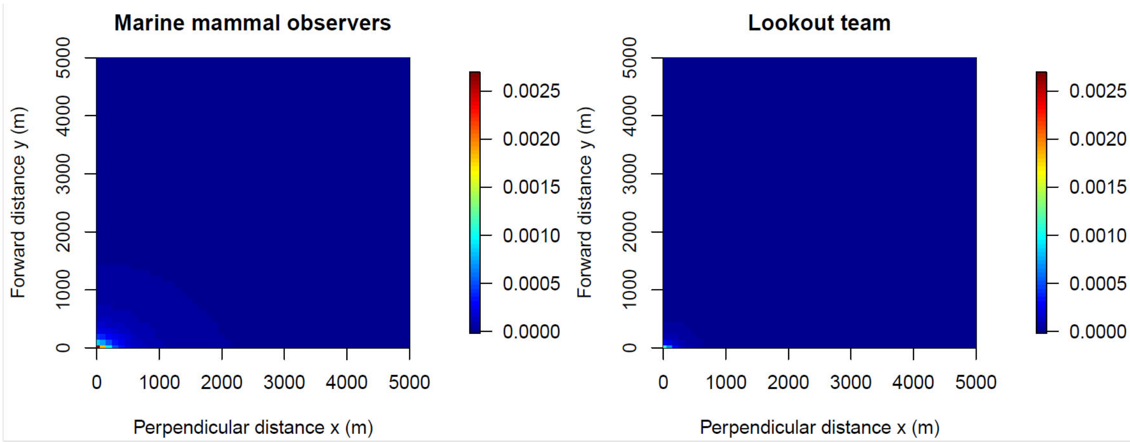


Figure 10. Estimated cumulative distribution of distances to first detection from model fitted to the small cetaceans in small pods (SCSP) group.

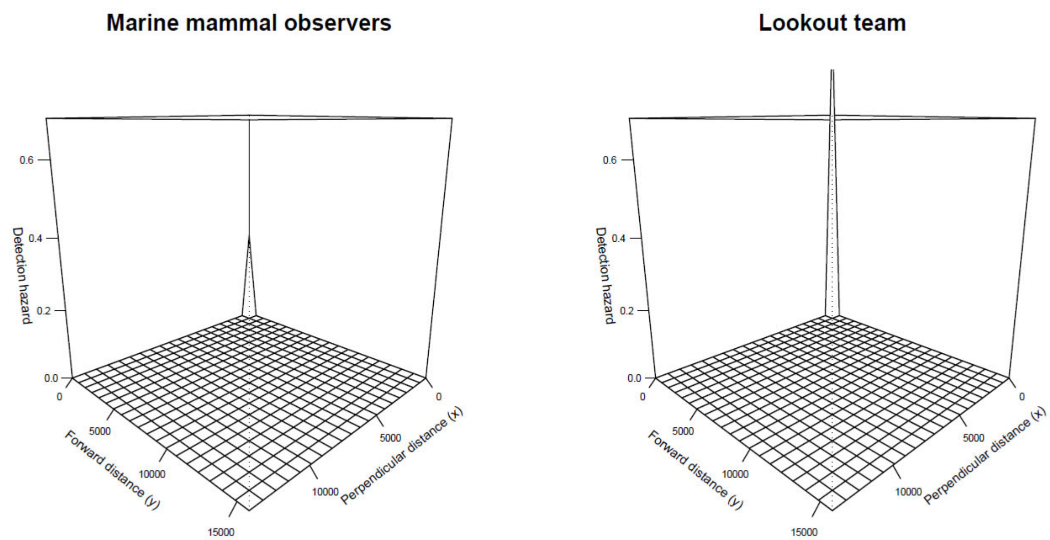


Figure 11. Estimated detection hazard of a surfacing from model fitted to the small cetaceans in small pods (SCLP) group.

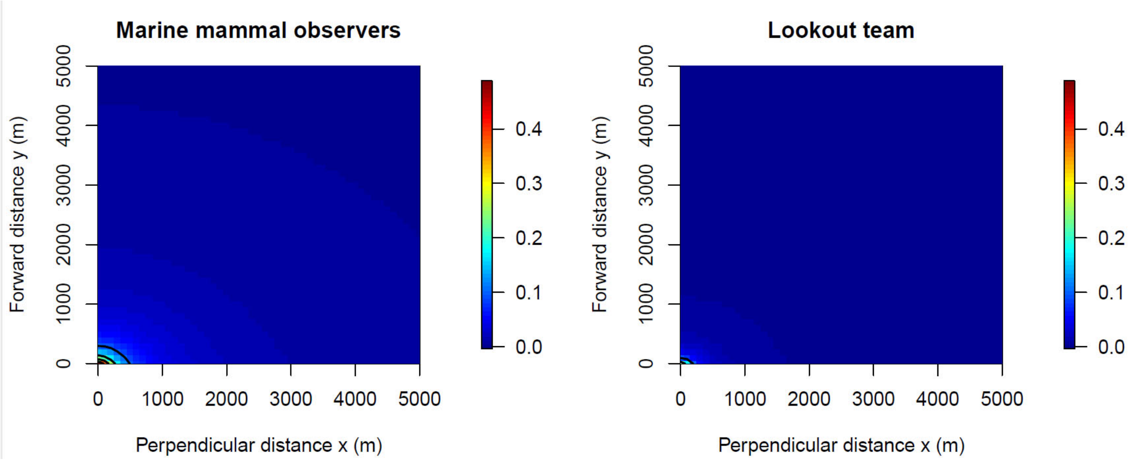


Figure 12. Estimated cumulative distribution of distances to first detection from model fitted to the small cetaceans in small pods (SCLP) group.

Appendix A: Taxonomic codes used in lookout effectiveness surveys.

Common and scientific names for each code is given in Table 4.4 of Appendix B.

Level1	Level2	Level3	Level4	Level5	Level6	Level7
BALAC	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
BALED	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
BALBO	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
BALMU	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
BALPH	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
MEGNO	NA	BAL--	LGWHA	WHALE	CET--	UN-MM
ESCRO	NA	NA	LGWHA	WHALE	CET--	UN-MM
MESDE	MES--	ZIP--	LGWHA	WHALE	CET--	UN-MM
ZIPCA	NA	ZIP--	LGWHA	WHALE	CET--	UN-MM
INDPA	NA	ZIP--	LGWHA	WHALE	CET--	UN-MM
BERBA	NA	ZIP--	LGWHA	WHALE	CET--	UN-MM
PHYMA	NA	NA	LGWHA	WHALE	CET--	UN-MM
KOGBR	KOG--	NA	NA	SMALL	CET--	UN-MM
KOGSI	KOG--	NA	NA	SMALL	CET--	UN-MM
ORCOR	NA	BLACK	NA	SMALL	CET--	UN-MM
PSECR	NA	BLACK	NA	SMALL	CET--	UN-MM
FERAT	NA	BLACK	NA	SMALL	CET--	UN-MM
PEPEL	NA	BLACK	NA	SMALL	CET--	UN-MM
GLOMA	GLO--	BLACK	NA	SMALL	CET--	UN-MM
TURTR	NA	NA	DOLPH	SMALL	CET--	UN-MM
GRAGR	NA	NA	DOLPH	SMALL	CET--	UN-MM
STEFR	STE--	NA	DOLPH	SMALL	CET--	UN-MM
STEAT	STE--	NA	DOLPH	SMALL	CET--	UN-MM
STELO	STE--	NA	DOLPH	SMALL	CET--	UN-MM
STECO	STE--	NA	DOLPH	SMALL	CET--	UN-MM
STEBR	NA	NA	DOLPH	SMALL	CET--	UN-MM
DELDE	DEL--	NA	DOLPH	SMALL	CET--	UN-MM
DELCA	DEL--	NA	DOLPH	SMALL	CET--	UN-MM
LAGHO	NA	NA	DOLPH	SMALL	CET--	UN-MM
LAGOB	NA	NA	DOLPH	SMALL	CET--	UN-MM
LISBO	NA	NA	DOLPH	SMALL	CET--	UN-MM
CHEMY	NA	NA	NA	NA	TURTL	NA
EREIM	NA	NA	NA	NA	TURTL	NA
LEPKE	NA	NA	NA	NA	TURTL	NA
DERCO	NA	NA	NA	NA	TURTL	NA
CARCA	NA	NA	NA	NA	TURTL	NA
LEPOL	NA	NA	NA	NA	TURTL	NA
NEOSC	NA	NA	UNOTA	SEALS	UNPIN	UN-MM
ZALCA	NA	NA	UNOTA	SEALS	UNPIN	UN-MM
PHOVI	NA	NA	UNOTA	SEALS	UNPIN	UN-MM
MIXED	NA	NA	NA	NA	NA	UN-MM
MOLMO	NA	NA	NA	NA	FISH	NA

Appendix B: Calibrating US Navy lookout observer effectiveness. Information for Marine Mammal Observers, Version 2.1.

See attachment.

Appendix C: Markov-modulated Poisson process models for lookout effectiveness data

See attachment.

August 2016

U.S. Navy Lookout Effectiveness Study
Marine Mammal Observer Survey Protocol

Prepared for:
U.S. Fleet Forces Command and Commander U.S. Pacific Fleet

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

ASWO	Anti-submarine Warfare Officer
CG	United States Navy guided missile cruiser
CMDCM	Command Master Chief
CTR	Contractor
DDG	United States Navy guided missile destroyer
DMMO	data marine mammal observer
ESA	Endangered Species Act
GPS	global positioning system
ICMP	Integrated Comprehensive Monitoring Program
LMMO	liaison marine mammal observer
LO	Lookout
MFAS	mid-frequency active sonar
MMO	marine mammal observer
MMPA	Marine Mammal Protection Act
NAVFAC	Naval Facilities Engineering Command
NMFS	National Marine Fisheries Service
OOD	Officer of the Deck
OPS	Operations Officer
POC	Point of Contact
SMMO	survey marine mammal observer
SYSCOM	Systems Command
U.S.	United States
XO	Executive Officer

SECTION 1 INTRODUCTION

In order to train and test with active sonar, the United States (U.S.) Navy has obtained Authorizations and permits from the National Marine Fisheries Service (NMFS) under the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA) for the incidental take of protected species. The Navy conducts monitoring within Navy Range Complexes and testing ranges, guided by the Integrated Comprehensive Monitoring Program (ICMP), as required under the MMPA and the ESA (Department of the Navy 2010).

The ICMP provides the overarching framework for coordination of the U.S. Navy Marine Species Monitoring Program. The ICMP outlines objectives for marine species monitoring and U.S. Navy-funded research relating to the effects of naval training and testing activities on protected marine species. The ICMP includes the following scientific objectives (Department of the Navy 2010):

1. monitor and assess the effects of Navy activities on protected marine species;
2. ensure that data collected at multiple locations is collected in a manner that allows comparison between and among different geographic locations;
3. assess the efficacy and practicality of the monitoring and mitigation techniques; and
4. add to the overall knowledge base of protected marine species and the effects of Navy activities on these species.

In accordance with the third objective, the Navy is conducting a Lookout Effectiveness Study to evaluate the effectiveness of Navy lookout and bridge watch teams at detecting protected marine species during at-sea training and testing events. To conduct the Lookout Effectiveness Study, trained marine mammal observers (MMO) embark on Navy ships to collect data that characterizes the likelihood of detecting marine species in the field from U.S. Navy guided missile destroyers (DDG) or cruisers (CG). The MMO sighting data is then compared to the sighting data collected from the ship's watch team.

SECTION 2 PLATFORMS AND CONFIGURATIONS

2.1. Search Platforms

Data collection takes place onboard either CG or DDG vessels. CGs are multi-mission surface combatants capable of supporting carrier battle groups, amphibious forces, or of operating independently and as flagships of surface action groups. CGs are 567 feet in length. DDGs are also multi-mission surface combatants, but are slightly smaller at 504 feet in length. Both ships are capable of speeds over 30 knots, and have crews of over 300 Sailors. Both of these ships are equipped with mid-frequency active sonar (MFAS) and are the ships most commonly involved in sonar exercises.

The layouts of the bridge wings on both ships are similar and are approximately 6x20 ft. Each bridge wing is equipped with a set of "big eye" binoculars, a Captain's chair, and a Pelorus. Navy personnel on the bridge that may be acting as a lookout, either on the wings or as part of the bridge team, include the Junior Officer of the Watch, Boatswain's Mate of the Watch, sky and surface lookouts, and Quartermaster of the Watch.

2.2. Marine Mammal Observer Configuration

Four MMOs are required to perform lookout effectiveness data collection:

- One MMO is positioned on each bridge wing; these MMOs are called survey MMOs (or SMMOs).
- One MMO is dedicated to recording data that the other MMOs relay (Data MMO or DMMO).
- One MMO is dedicated to liaising (liaison MMO, or LMMO) with the bridge and lookout team, and reporting their sightings to the DMMO.

The MMOs rotate positions hourly, such that no two hours are spent consecutively as SMMOs. For example, the rotation would be port SMMO → DMMO → starboard SMMO → LMMO.

2.3. Navy Lookout and Bridge Team Configuration

Lookouts are stationed in various arrangements depending on ship configuration and training exercise. Ships using low-frequency and hull-mounted MFAS sources associated with anti-submarine warfare and mine warfare activities at sea have a minimum of two Lookouts, one on each bridge wing.

Lookouts on some ships wear headsets for communication with the bridge, which makes determining when they see a surfacing species more difficult. SMMOs need to be flexible in their positioning relative to the lookouts to maximize data collection while minimizing the potential for cueing the lookout to a sighting.

SECTION 3 MARINE MAMMAL OBSERVER PROTOCOL

The goal of the survey is to set up "trials" for the bridge/lookout team, such that the MMO observes an animal before the bridge/lookout, and determines if the bridge/lookout team subsequently detects that animal or not, and at what distance. It is imperative that the MMOs do not cue the bridge or lookouts to any animals; if cueing the bridge/lookout has occurred, the sighting is no longer considered a trial for the purpose of analysis.

A sighting is a trial if:

1. The animal or group is sighted either first by the MMO or the MMO sights the animal or group at the same time as the lookout team.

2. The animal or group is between 270° and 90° relative to the ship.

A sighting is not considered a trial if:

1. The bridge/lookout sights the animal or group before the MMO.

2. The bridge/lookout sights the animal or group and the MMO does not.

3.1 Survey MMO

Survey MMOs are responsible for actively searching for animals from abeam to 5° to the opposite side of the bow (i.e., starboard SMMO surveys 355° to 90° and port SMMO surveys 270° to 5° relative to the bow of the ship). It is imperative that the MMOs do not cue the lookout or bridge team upon sighting an animal so between sightings, SMMOs should occasionally use the binoculars or camera as if they would during a sighting. All effort and environmental data recorded as well as sightings made by the SMMOs are relayed to the DMMO.

For an SMMO sighting:

1. Relay sighting to the DMMO. At the first instance of a sighting, the MMO that saw the animal will say "mark".

2. Provide the following information when prompted by the DMMO: species, bearing, sighting distance, and group size. Information (e.g., cue or behavior) should be collected afterwards if time allows. WARNING: MMOs must not hold button on headset/radio longer than necessary as it would not allow other sightings to be relayed.

3. Inform the lookout/bridge team if the animal is within or enters the shutdown zone during MFAS use (not bow-riding), or if a collision may occur, for appropriate action. The sighting is no longer considered a trial once the lookout/bridge team has been notified of the animal(s).

4. Track the animal, relaying all subsequent surfacing (for example, whale blows), as appropriate. In the case of continuously or almost-continuously available sightings (for example, large dolphin pods), record the beginning & end times of the sighting.

5. Inform the DMMO that you are no longer tracking the animal(s) once the animal(s) has either passed the beam of the ship or can no longer be tracked.
6. If the track of the animal is lost, any additional sightings that can't be confirmed as the same animal(s) previously being tracked would be a new trial.

SMMOs are also responsible for determining if a bridge/lookout sighting is a duplicate (i.e., the same animal) as one that is being tracked.

3.2 Liaison MMO

The LMMO is responsible for relaying all sightings made by the bridge and/or lookout team to the DMMO. Depending upon the configuration of the lookouts, the LMMO may be positioned either inside the pilot house, on one bridge wing, or moving between the bridge wings and the pilot house.

At the beginning of each ship watch team rotation, the lead MMO should introduce themselves to the Officer of the Deck (OOD) and let them know that for the purposes of the study, the LMMO needs to be informed when anyone observes a marine mammal. Also remind them that MMO team is only collecting scientific data and does not replace the lookouts. Therefore, the watch team should observe, report and mitigate as they would if the MMO team were not aboard.

When there is a sighting made by the lookout/bridge team, the LMMO should immediately query the lookout/bridge team to obtain, at minimum, the lookout's estimated distance, bearing, and species group (whale, dolphin, etc.) as quickly as possible. It is necessary for data analysis requirements that the LMMO collects these data in numerical units (e.g. degrees, yards); even though in most cases the lookouts/bridge team may not initially provide these numerical data. The LMMO should also obtain a reticle distance to the animal(s) in addition to reporting the distance estimated by the bridge/lookout team. When the LMMO is reporting information, it needs to be clear whether the information was provided by the bridge/lookout or whether the information is an observation from the LMMO. The sighting information described by the lookout/bridge team should have its own entry on the Sightings datasheet. Additional information such as species identification can be added in the Comments field or under a separate data entry if the animals are re-sighted by an MMO. In addition to real-time data gathering of lookout/bridge team sightings, any bridge logs of animals observed should be copied to ensure all sightings are captured. Additionally, the LMMO may be able to plug into the same communications network as the lookouts (typically sound powered phones). This would improve the ability to obtain the required contact information.

It is important to not cue the bridge/lookout team to the presence of an animal or group. This is because providing any indication of a sighting to the bridge/lookout team would eliminate the sighting as a trial. The LMMO can try to photograph the animals for species identification purposes once the bridge/lookout has sighted an animal. Alternately, if photographs can be taken discretely, taking them is recommended to help in species identification. The LMMO can frequently photograph non-marine mammals, such as birds, water, etc., such that the lookouts desensitize to camera presence.

The LMMO may also operate as an additional SMMO so long as the bridge/lookout is not cued by any sightings. If the LMMO first observes the animal(s), the initial sighting distance/bearing should be provided by the LMMO, and then tracking of the sighting should be passed to the SMMO so that the LMMO can focus on the bridge/lookout team for their observations.

LMMOs may assist the SMMOs in determining if a bridge/lookout sighting is a duplicate (i.e., the same animal) as one that is being tracked.

3.3 Data MMO

The DMMO is responsible for recording all effort and environmental data at the beginning of the shift, as well as all sightings made by the SMMOs and those relayed by the LMMO. The DMMO position is also considered a resting position; when feasible, the DMMO may sit down so long as a waypoint can be immediately recorded upon a sighting being called.

Ideally, the DMMO will be positioned in the pilot house with permission from the Officer of the Deck. This facilitates the collection of ship bearing and ship speed, which can be obtained from displays on the bridge (typically available on the bridge wing as well). The DMMO location is dependent upon bridge wing configuration and pilot house congestion.

At the first instance of a sighting, the MMO that saw the animal will say “mark”. This lets the DMMO know to take a waypoint using a Global Positioning System (GPS) unit to mark the time and location of a sighting. Any GPS devices brought on board (either dedicated devices or enabled on certain camera bodies) must not record the trackline of the vessel since a detailed record of ship movements during exercises would be sensitive information. The DMMO should then immediately note the ship bearing and speed. Next, the DMMO asks the SMMO for the bearing, distance, group size, and species of the animal. The SMMOs/lookouts can report distance in yards, meters, or reticles, but all distance measures should be converted to meters for the final datasheet. These are the most crucial pieces of information. The DMMO should then query the SMMO for the remaining data fields as the SMMO has time/ability to answer. While focused on maintaining sighting of the animal, the SMMO is not expected to remember all of the data fields on the data sheets. Therefore, it is the responsibility of the DMMO to ask the SMMOs for missing entries and ensure that all fields are filled in by the end of the sighting. It is the responsibility of the DMMO to prioritize data collection from the SMMOs when multiple SMMOs have a sighting at the same time. The DMMO should enter the time, latitude, and longitude for each sighting from the GPS onto the data sheets when there is time to do so.

SECTION 4 DATA COLLECTION AND FORMS

4.1. Effort and Environmental Data

At the beginning of each effort period and at each observer rotation, effort and environmental data are recorded. Additionally, any significant change in weather also warrants recording new data. Each field in the Effort and Environmental Data form is described in Table 1. Additional discussion is provided below the table for those fields needing better explanation. An example of a completed Effort and Environmental Data form is provided in Appendix A.

Table 1. Effort and Environmental Data Form Field Descriptions

Field	Description
WP (Waypoint)	Waypoint obtained from the GPS unit
Effort	Whether search effort is on or off
Event*	Event Options: 1. Begin effort 2. End effort 3. Observer rotation 4. Significant weather change 5. Other (for any other waypoints, changes in lookout configuration, etc.)
Time	Time (hh:mm:ss) recorded off GPS waypoint
Ship Latitude and Longitude	Latitude and longitude recorded off GPS waypoint in decimal degrees
MMO Positions (Port, Starboard, Liaison, Data)	Record the (three-letter initials) of the person at each MMO position. If one position is vacant, enter N/A for that field.
Sea State*	Beaufort sea states are provided in the text below and Table 2
Wave Height	Total (swell included) wave height. Light (0 – 3ft) Moderate (4 – 6ft) Heavy (>6ft)
Visibility	Visibility codes: B – Bad (<0.5km) P – Poor (0.5 – 1.5km) M – Moderate (1.5 – 10km) G – Good (10 - 15km) E – Excellent (>15km)
% Glare	Percent glare should be the total for the 180° field of view for both Port and Starboard observers. Each observer will report the % glare in their 90° quadrant.
% Cloud Cover	Percent cloud cover should be taken by each Port and Starboard observer for their 90° quadrant. The % cloud cover for the 180° field of view will be averaged upon data entry.
Sonar on/off	Indicate whether sonar was on or off during an effort period.
Explosives in use?	Yes or no.
LO config*	Identify the locations and numbers of the Navy lookouts.
Comments	Any additional comments relative to the observing session.

* indicates additional description is provided below.

4.1.1. Event

Events 1, 2, and 3 are used to record when effort starts (at the beginning of each day and after breaks), when effort is off for breaks (e.g., lunch), and at each observer rotation, respectively. Event 4 is used to record significant changes in weather. Event 5 is used to record miscellaneous points that are deemed important such as changes in sonar use, changes in lookout configuration, etc.

4.1.2. Sea State

Beaufort sea states should be used when recording data. However, be advised that the Navy has historically used a separate sea state scale, and therefore information obtained from the bridge may not be consistent with the Beaufort sea state scale, so use the Beaufort sea state and not the sea state reported by the bridge. A description of the Beaufort sea states are provided in Table 2 and Figure 1.

Table 2. Beaufort Sea State Descriptions

Beaufort Sea State	Wind speed (kts)	Wind description	Wave height (ft)	Description – Beaufort
0	<1	Calm	0	Calm; like a mirror
1	1-3	Light air	¼	Ripples with appearance of scales; no foam crests
2	4-6	Light breeze	½ – 1	Small wavelets; crests of glassy appearance, not breaking
3	7-10	Gentle breeze	2 – 3	Large wavelets; crests begin to break; scattered whitecaps
4	11-16	Moderate breeze	3 ½ – 5	Small waves, becoming longer numerous whitecaps
5	17-21	Fresh breeze	6 – 8	Moderate waves, taking longer form; many whitecaps; some spray
6	22-27	Strong breeze	9 ½ – 13	Larger waves forming; whitecaps everywhere; more spray
7	28-33	Near gale	13 ½ – 19	Sea heaps up; white foam from breaking waves begins to be blown in streaks
8	34-40	Gale	18 – 28	Moderately high waves of greater length; edges of crests begin to break into spindrift; foam is blown in well-marked streaks
9	41-47	Strong gale	23 – 32	High waves; sea begins to roll; dense streaks of foam; spray may reduce visibility
10	48-55	Storm	29 – 41	Very high waves with overhanging crests; sea takes white appearance as foam is blown in very dense streaks; rolling is heavy and visibility is reduced
11	56-63	Violent storm	39 – 46	Exceptionally high waves; sea covered with white foam patches; visibility still more reduced
12	≥ 64	Hurricane	37 – 52	Air filled with foam; sea completely which with driving spray; visibility greatly reduced


Force 0


Force 1


Force 2


Force 3


Force 4


Force 5

Figure 1. Beaufort Sea State Photographs, Sea States 0 – 7

4.1.3. Lookout Configuration

Identify the positions of the Navy lookouts, i.e. port wing, starboard wing, both wings or bridge team, being careful not to cue the lookouts that you are taking information on them. The DMMO should discreetly gather this information rather than requesting the SMMOs for updates so as to avoid potential signals that these data are being gathered.

4.2. Sighting Data

Each field in the Sighting Data form is described in Table 3. Additional discussion is provided below the table for those fields needing further explanation. An example of a completed Sightings Data form is provided in Appendix B.

Table 3. Sightings Data Form Field Descriptions

Field	Description
Sighting Number*	Sighting numbers are sequential, with each cruise beginning with sighting number 1.
Time (& Ending if continuous)	Time at sighting is to be recorded as HH:MM:SS and is obtained from the GPS unit in local time. For species that are continuously available (such as a large pod of dolphins), the start time would be the initial sighting and the end time would be once the animals pass the beam or are lost from view.
WP (Waypoint)	Waypoint obtained from the GPS unit
Animal Bearing	Animal bearing is estimated using 'the compass' on the big eyes (if calibrated correctly), the gyrocompass repeater, or observer estimate. If based on observer estimate, for ease in the field this can be written as 0 to 90° Port or Starboard (P or S) on the datasheet to account for the position of the animal(s) in the survey quadrant that it is sighted in. .
Animal Distance	Distance is estimated to the single animal or to the geometric center of a group of animals. Distance is recorded as an observer estimate (in yards, nautical miles, meters, etc.) or as a reticle distance. Each demarcation line in the Fuginon binoculars represents 1 reticle.
Species*	Species codes are determined by merging the first three letters of the genus with the first two letters of the species. For example, humpback whale would be MEGNO.
Group Size (min/max/best) & # of Calves	The minimum, maximum, and best estimate for number of animals in a group, as well as the number of calves are to be included.
Ship Latitude and Longitude	Latitude and longitude of the GPS waypoint in decimal degrees.

Ship Bearing	True bearing of the ship needs to be recorded so the true position of the animal can be calculated.
Ship speed (kts)	Ship speed provides an indication of the time available for the bridge/lookout to observe an animal before it passes the beam. To avoid classification issues, categories of ship speeds are provided: < 5 kts, 5 – 10 kts, 11 – 15 kts, 16 – 20 kts, and > 20 kts.
Relative motion*	The animal's motion with respect to the ship.
Observer	The three-letter initials of the MMO, or "BR" for the bridge or "LO" for the lookout.
Sighting cue	Indicate what the sighting cue was. For consistency, the following are provided: Blow, Head slap, Splash Body, Porpoising, Tail slap Breach, Pectoral fin slap, Seabirds/Wildlife Dorsal fin, Slick, footprint, or ring, Other Fluke up
Behavior*	Provide any relevant information on the behavior of the animal or group sighted (traveling, diving, spyhopping, etc).
End of Track*	Reason for which the animal(s) is no longer being tracked.
Mitigation	Indicate N/A if sonar is not on; N if sonar is on and no mitigation was implemented, and Y if sonar is on and the type of mitigation implemented (6 dB down, shut off, ship turned away from animal, etc.)
Trial (Y/N)	Identifies if the sighting is considered a "trial" for the effectiveness study. A sighting is a trial if: 1. The animal or group is sighted either first by the MMO or the MMO sights the animal or group simultaneously as the bridge/lookout team **and** 2. The animal is between 270° and 90° relative to the ship
Comments	All other pertinent data should be included here, such as LO configuration at the time of sighting, and ID numbers of photos taken.

* indicates additional description is provided below.

4.2.1. Sighting Number

Sighting numbers are sequential throughout a cruise; each cruise begins with sighting #1.

Upon the first sighting of an animal, the next sequential number is assigned, for example Sighting # 4.0 in the Tables 4-6 below. Additional sightings of the same animal (e.g., for each surfacing) would be given decimal numbers, such that sightings would be called 4.1, 4.2, 4.3, etc.

In cases where the surfacing of animals are continuously available (e.g. large pod of dolphins), one sighting number is recorded, which includes the start time of the initial sighting and the end time for when the animals pass the beam or is lost from view.

If the bridge or lookout team observes the animal(s) at the same time as the MMO, two entries are recorded on the data form and each entry is given the same sighting number (Table 4 and Table 7). If the bridge/lookout team observes the same animal(s) as the MMO, whether it is a surfacing series or a continuous sighting, but sees it later in time, the same sighting number is used but the later time is recorded (Table 5 and Table 8). If the bridge or lookout team observes a surfacing of the animal(s) within a series and that surfacing is not observed by the MMO, the next sequential decimal number is assigned (Table 6). If the bridge or lookout team observes an initial surfacing of the animal(s) and that surfacing is not observed by the MMO, the next sequential sighting number is assigned (Table 9).

Table 4. Example Sighting – MMO and Lookout Simultaneous Observation During a Surfacing Series

Sighting Number	Time	WP	Observer	Trial	Comments
4.0	08:12:23	10	ABC	Y	
4.1	08:12:57	11	ABC	Y	
4.2	08:13:37	12	ABC	Y	
4.3	08:14:01	13	ABC	Y	LO observed at same time as MMO
4.3	08:14:01	13	LO	N	

Table 5. Example Sighting – Lookout Observes Same Surfacing Series, but Later in Time

Sighting Number	Time	WP	Observer	Trial	Comments
4.0	08:12:23	10	ABC	Y	
4.1	08:12:57	11	ABC	Y	
4.2	08:13:37	12	ABC	Y	
4.3	08:14:01	13	ABC	Y	
4.3	08:14:20	13	LO	N	LO observed same surfacing series as MMO, but slightly later. Same sighting number is used, but a later time is recorded.

Table 6. Example Sighting – Lookout Sighting after Initial MMO Sighting During a Surfacing Series

Sighting Number	Time	WP	Observer	Trial	Comments
4.0	08:12:23	10	ABC	Y	
4.1	08:12:57	11	ABC	Y	
4.2	08:13:37	12	ABC	Y	
4.3	08:14:01	13	LO	N	MMO missed this surfacing.
4.4	08:14:42	14	ABC	N	No longer a trial because the lookout has seen the animal

Table 7. Example of a Continuous Sighting- MMO and Lookout Simultaneous Observation

Sighting Number	Start Time	End Time	WP	Observer	Trial	Comments
1.0	08:12:23	08:14:27	1	ABC	Y	
2.0	08:25:57	08:26:38	9	ABC	Y	
2.0	08:25:57	08:26:38	9	LO	N	LO observed at same time as MMO
3.0	08:37:01	08:38:04	13	ABC	Y	

Table 8. Example of a Continuous Sighting- Lookout Observes Same Surfacing, but Later in Time

Sighting Number	Start Time	End Time	WP	Observer	Trial	Comments
1.0	08:12:23	08:14:27	1	ABC	Y	
2.0	08:25:57	08:27:38	9	ABC	Y	

2.0	08:26:42	08:27:38	10	LO	N	LO observed same group of dolphins as MMO, but slightly later. Same sighting number is used, but a later time is recorded.
3.0	08:37:01	08:38:04	13	ABC	Y	

Table 9. Example of a Continuous Sighting- Lookout Observes Initial Surfacing, but MMO Does Not

Sighting Number	Start Time	End Time	WP	Observer	Trial	Comments
1.0	08:12:23	08:14:27	1	ABC	Y	
2.0	08:25:57	08:27:38	9	LO	N	
2.1	08:26:42	08:27:38	10	ABC	N	LO observed group of dolphins before MMO. A decimal number is used for the sighting, and a later time is recorded.
3.0	08:37:01	08:38:04	13	ABC	Y	

4.2.2. Species

Species codes are determined by combining the first three letters of the genus with the first two letters of the species. For example, humpback whale would be MEGNO for ***MEG****aptera* ***NO****vaeangliae*. Table 10 provides a list of species codes; this list will be updated as necessary.

Table 10. Species Codes

Group	Species Code	Common Name	Scientific Name
Mysticetes	BALAC	Minke whale	*Balaenoptera acutorostrata*
	BALED	Bryde's whale	*Balaenoptera edeni*
	BALBO	Sei whale	*Balaenoptera borealis*
	BALMU	Blue whale	*Balaenoptera musculus*
	BALPH	Fin whale	*Balaenoptera physalus*
	BAL--	Unidentified rorqual	Balaenopteridae
	MEGNO	Humpback whale	*Megaptera novaeangliae*
	ESCRO	Gray whale	*Eschrichtius robustus*
	WHALE	Unidentified whale	
Beaked whales	ZIP--	Unidentified beaked whales	Ziphiidae
	MES--	Unidentified *Mesoplodon*	*Mesoplodon* spp.
	MESDE	Blainville's beaked whale	*Mesoplodon densirostris*
	ZIPCA	Cuvier's beaked whale	*Ziphius cavirostris*
	INDPA	Longman's beaked whale	*Indopacetus pacificus*
	BERBA	Baird's beaked whale	*Berardius bairdii*
Sperm whales	PHYMA	Sperm whale	*Physeter macrocephalus*
	KOGBR	Pygmy sperm whale	*Kogia breviceps*
	KOGSI	Dwarf sperm whale	*Kogia sima*
	KOG--	Unidentified pygmy/dwarf sperm whale	*Kogia* spp.
Blackfish	ORCOR	Killer whale	*Orcinus orca*

Group	Species Code	Common Name	Scientific Name
	PSECR	False killer whale	*Pseudorca crassidens*
	FERAT	Pygmy killer whale	*Feresa attenuata*
	PEPEL	Melon-headed whale	*Peponocephala electra*
	GLOMA	Short-finned pilot whale	*Globicephala macrorhynchus*
	BLACK	Unidentified blackfish	
Dolphins	TURTR	Bottlenose dolphin	*Tursiops truncatus*
	STEAT	Pantropical spotted dolphin	*Stenella attenuata*
	GRAGR	Risso's dolphin	*Grampus griseus*
	STELO	Spinner dolphin	*Stenella longirostris*
	STE--	Unidentified *Stenella*	*Stenella* spp.
	STECO	Striped dolphin	*Stenella coeruleoalba*
	STEBR	Rough-toothed dolphin	*Steno bredanensis*
	LAGHO	Fraser's dolphin	*Lagenodelphis hosei*
	LAGOB	Pacific white-sided dolphin	*Lagenorhynchus obliqidens*
	LISBO	Northern right whale dolphin	*Lissodelphis borealis*
	DOLPH	Unidentified dolphin	
Turtles	CHEMY	Green turtle	*Chelonia mydas*
	EREIM	Hawksbill turtle	*Eretmochelys imbricata*
	LEPKE	Kemp's ridley turtle	*Lepidochelys kempii*
	DERCO	Leatherback turtle	*Dermochelys coriacea*
	CARCA	Loggerhead turtle	*Caretta caretta*
	LEPOL	Olive ridley turtle	*Lepidochelys olivacea*
	TURTL	Unidentified turtle	
Pinnipeds	NEOSC	Hawaiian monk seal	*Neomonachus schauinslandi*
	ZALCA	California sea lion	*Zalophus californianus*
	PHOVI	Harbor seal	*Phoca vitulina*
Unidentified	CET--	Unidentified cetacean	
	LGWHA	Unidentified large whale	
	SMALL	Unidentified small cetacean	

4.2.3. Relative Motion

The relative motion is that of the animal relative to the ship.

- Opening: animal is moving away from the ship
- Closing: animal is moving toward the ship
- Parallel: animal is staying at the same distance from the ship
- None: animal is stationary

If only one surfacing is detected, and the direction of movement of the body is not discerned (eg. only a single blow is seen), enter UNK in the field.

4.2.4. Behavior

Record the initial behavior of the animal or group once sighted. Any other changes in behavior can be recorded in the comments. Examples of behavior include:

- Breaching
- Bowriding
- Feeding
- Fluking
- Flipper slapping
- Milling
- Logging
- Resting
- Traveling
- Tail slap
- Vocalizing
- Other

4.2.5. End of Track

End of Track identifies why an animal or group is no longer being observed. The end of the track is either due to the animal observed passing the beam or the animal is "lost" meaning that sufficient time has passed (species dependent) that another surfacing should have occurred, but was not observed by the MMO. If neither of those cases sufficiently explains the end of track, write a brief reason in the space provided or in the comments.

4.3. Data Storage

Copies of the sighting data sheets, entered data files, and all supplemental data (such as pictures) are submitted to U.S. Fleet Forces Command (N46) or Commander, Pacific Fleet Environmental (N465), depending on the cruise location, and NUWCDIVNPT. NUWCDIVNPT currently formats the data to send to the University of St. Andrews for analysis.

SECTION 5 BRIEFINGS AND INTERACTION WITH OFFICERS AND CREW

5.1. Initial Briefing

It is important that the Commanding Officer (CO), the officers and the crew know that you work for the U.S Navy (as a Government Service or Contractor (CTR), what you need from them and why you are on board. At the discretion of the CO, you may be asked to brief the Wardroom on the first day (or at a presail) or at the first Operations Officer (OPS) brief. A template will be provided by CPF or USFF. The senior MMO, as designated by the Fleet, should be prepared to give the brief if requested by the CO.

5.2. Use the Opportunity Aboard to Inform and Potentially Learn

5.2.1. Inform

- How important implementation of protective measures are (e.g they are legally required to train and test, they protect marine mammals)
- How important the lookouts are (e.g. they are the "front line" of marine mammal protection)
- What the Fleet/Naval Facilities Engineering Command (NAVFAC)/Systems Command (SYSCOM) Environmental programs do (e.g. we do the paperwork to get authorization for them to train and test.)

5.2.2. Learn

- If you work on the EIS, MMPA and ESA authorizations (this does not apply to the CTRs), you may have the need to know. Assuming you have SECRET clearance, at the discretion of the vessel, and if there is time you may request access to the following from your onboard Point of Contact (POC) (usually the OPS) or Anti-submarine Warfare Officer (ASWO) : SONAR to see how they observe/report marine mammal vocalizations to the bridge, see how they track, etc.; COMBAT to observe portions of the training event; NIXIE (Torpedo Countermeasures System) if it is being deployed while you are onboard. Seeing these operations in action will help your support of environmental compliance documents.

5.3. Professionalism

The MMO may be the only U.S. Navy environmental staff that most onboard the ship will ever interact with. This is especially true for the enlisted Sailors. This probably goes without saying, but you are the ambassadors for the at sea program. While what we do is very important, we are visitors aboard the ship. Please be professional at all times.

Remember that all vessel traffic can hear the MMOs radio communications while onboard. Please be sensitive to the information that is being relayed to one another. Limit personal conversations, use professional language, and do not relay information that may be labeled as sensitive such as ship speed or sonar information.

At the beginning of the embark, the senior MMO should introduce themselves and the rest of the MMO team to the CO, Executive Officer (XO) and Command Master Chief (CMDCM) and let them know that for the purposes of the study, they need to be informed when anyone observes a marine mammal. Also remind them that the MMOs are only collecting scientific data and are not replacing the lookouts. Therefore, the lookouts should observe, report and mitigate as they would if the MMOs were not onboard.

SECTION 6 REFERENCES

Department of the Navy (2010). United States Navy Integrated Comprehensive Monitoring Program 2010 Update, 20 December 2010.

APPENDIX A COMPLETED EFFORT AND ENVIRONMENTAL DATA FORM

Example of a completed Effort and Environmental Data Form is provided on the next page.

Event : Form Example 1
Date: 30 February 2011

Vessel Type: SSN - X
Vessel Name: USS Porthole

Page 1 of 1
Location: Hawaii Range Complex

Effort and Environmental Data

WP	Effort	Event	Time	Ship Latitude and Longitude	Port MMO	Stbd MMO	Liaison MMO	Data MMO	Sea State	Sonar on/off	Explosives in use?	Visibility	Wave Height	Swell Direction	% Glare	% Cloud Cover	LO config	Comments
1	on off	1	07:01:23	Lat: Lon:	AMY	JUL	TAJ	RKU	3	on off	Y N	B P M G E	Light Mod Heavy	010	40	75	LOs arrived on wings at 0713	
2	on off	3	08:00:10	Lat: Lon:	TAJ	RKU	JUL	AMY	4	on off	Y N	B P M G E	Light Mod Heavy	170	0	75	1 LO on each wing	
4	on off	5	08:22:46	Lat: Lon:	TAJ	RKU	JUL	AMY	4	on off	Y N	B P M G E	Light Mod Heavy	190	0	75	1 LO on each wing	
9	on off	3	09:05:17	Lat: Lon:	JUL	AMY	RKU	TAJ	3	on off	Y N	B P M G E	Light Mod Heavy	180	0	75	1 LO on each wing	
9	on off	3	10:03:29	Lat: Lon:	RKU	TAJ	AMY	JUL	3	on off	Y N	B P M G E	Light Mod Heavy	30	0	95	1 LO on each wing	
10	on off	5	10:11:33	Lat: Lon:	RKU	TAJ	AMY	JUL	3	on off	Y N	B P M G E	Light Mod Heavy	90	0	100	1 LO on each wing	Rain squall. MMOs and LOs in pilot house observing
11	on off	5	10:46:15	Lat: Lon:	RKU	TAJ	AMY	JUL	4	on off	Y N	B P M G E	Light Mod Heavy	90	0	100	1 LO on each wing	
17	on off	3	11:01:23	Lat: Lon:	AMY	JUL	TAJ	RKU	4	on off	Y N	B P M G E	Light Mod Heavy	90	0	100	1 LO on each wing	
21	on off	2	12:06:01	Lat: Lon:						on off	Y N	B P M G E	Light Mod Heavy					Off effort for lunch. Yum!
	on off		: :	Lat: Lon:						on off	Y N	B P M G E	Light Mod Heavy					

Event: 1 – begin effort; 2 – end effort; 3 – observer rotation; 4 – weather change; 5 – waypoint (sonar change); 6 – high density area; 7 – back to normal

APPENDIX B COMPLETED SIGHTINGS DATA FORM

Example of a completed Sightings Data Form is provided on the next page.

Event: Form Example 2
Date: 20 July 2010

Vessel Type: DDG
Vessel Name: USS Porthole

Page 3 of 15
Location: SOCAL Range Complex

Sighting Number	Time (& ending time if continuous)	WP #	Animal Bearing	Animal Distance	Species	Group Size (min/max/best)	Ship Latitude and Longitude	Ship Bearing	Ship Speed (kts)	Relative Motion	Observer	Sighting cue	Behavior	End of Track	Mitiga-tion	Trial Y/N	Comments
42.0	10: 11 :45 : :	0473	10 P °	2	pinniped	1/ 2 /1 ___ calves	Lat: 33.0034 Lon: 118.9484	359 °	<5 5-10 11-15 16-20 >20	opening closing parallel none	RKU	body	travel	Lost Passed Beam ___	N	Y N	
42.0	10: 11 :55 : :	0473	10 S °	1000 yds	pinniped	1/ 2 /1 ___ calves	Lat: 33.0034 Lon: 118.9484	359 °	<5 5-10 11-15 16-20 >20	opening closing parallel none	LO	body		Lost Passed Beam ___	N	Y N	LO probably cued by MMO
42.1	10: 13 :54 : :	0473	80 S °	3	pinniped	1/ 2 /1 ___ calves	Lat: 33.00678 Lon: 118.9481	359 °	<5 5-10 11-15 16-20 >20	opening closing parallel none	RKU	fin	travel	Lost Passed Beam ___	6 dB down	Y N	Passed 90° at 10:14:59
42.1	10: 13 :54 : :	0474	80 P °	500 yds	pinniped	1/ 2 /1 ___ calves	Lat: 33.00678 Lon: 118.9481	359 °	<5 5-10 11-15 16-20 >20	opening closing parallel none	LO	splash		Lost Passed Beam ___	6 dB down	Y N	
	: : : :		°			/ / ___ calves	Lat: Lon:	°	<5 5-10 11-15 16-20 >20	opening closing parallel none				Lost Passed Beam ___		Y N	
43.0	14: 23 :09 14: 35:00	23	320 S °	1.5	Tursiops	3/ 6 /5 ___ calves	Lat: 32.80225 Lon: 117.6538	270 °	<5 5-10 11-15 16-20 >20	opening closing parallel none	MS	splash	bow riding	Lost Passed Beam ___	NA	Y N	Photos #4826-4986 suggest bottlenose dolphin
	: : : :		°			/ / ___ calves	Lat: Lon:	°	<5 5-10 11-15 16-20 >20	opening closing parallel none				Lost Passed Beam ___		Y N	
	: : : :		°			/ / ___ calves	Lat: Lon:	°	<5 5-10 11-15 16-20 >20	opening closing parallel none				Lost Passed Beam ___		Y N	
	: : : :		°			/ / ___ calves	Lat: Lon:	°	<5 5-10 11-15 16-20 >20	opening closing parallel none				Lost Passed Beam ___		Y N	

Markov-modulated Poisson process models for lookout effectiveness data

Cornelia Oedekoven & Len Thomas

22 March 2022

1 Introduction

Distance sampling (DS) is one of the most commonly used methods for assessing wildlife abundance (e.g., Buckland et al. 2015). Line transect surveys are a form of DS during which an observer moves along predefined lines and records detections to animals – or groups of animals – along with distances to the detections. Perpendicular distances from the line to the detection are required to estimate the average detection provability in the search area of the observer. In this way, DS accounts for the fact that the number of detections made by the observer is a function of two processes: 1. the distribution of the animals and 2. the observation process. In the classical version of DS ("conventional DS"), an unbiased estimate of the average detection probability relies on perfect detection on the line, i.e. all animals directly on the line must be detected with certainty, and is paramount for estimating animal abundance without bias. We distinguish between two biases caused by different reasons why animals may be missed by the observer. 1. availablity bias: animals are not available to be detected, e.g. diving cetaceans, and 2. perception bias: animals are available but missed, e.g. because the observer was looking in the wrong direction.

Mark-recapture (MR) DS alleviates reliance on this assumption by estimating the trackline detection probability, $g0$ (e.g., Burt et al. 2014). Two observers make detections of the same animals as they pass through the search area, creating mark-recapture occasions. We generally distinguish between two scenarios with varying dependences between platforms. In the simpler case, which we refer to as one-way independence, one observer, say observer 2, is independent from the other, say observer 1, i.e. observer 2 makes detections independent from observer 1. The MR component here is: if observer 2 detected the same animal as observer 1, either at the same time or later, the trial was a success, otherwise a failure. If both observers make detections independent from each other we refer to this as two-way independence in the following. Now the MR component also includes the animals first detected by observer 1: if observer 2 detected the same animal as observer 1, either at the same time or later, the trial was a success, otherwise a failure. We note that one-way or two-way independence refer to detections made at a given surfacing. We consider consecutive surfacings and, hence, detections of these as independent. This is different from conventional MRDS methods where generally only first detections are used and for the trial configuration (Burt et al. 2014) only animals are included that were first detected by, using the example above, observer 1.

Non-independence between observers can be caused by reasons other than cueing each other. It may also arise from animals only being stochastically available, e.g. periodically surfacing cetaceans, and being stochastically available to both observers at the same instances (Langrock, Borchers, and Skaug 2013; Borchers and Langrock 2015). Langrock, Borchers, and Skaug (2013) developed methods that deal with this type of non-independence by incorporating a Markov modulated Poisson process (MMPP) as the availability model. Borchers and Langrock (2015) extended these methods by incorporating a MR component. Both versions include fitting a two-dimensional (2D) detection model which uses forward distances between the observer and animal in addition to the perpendicular distances.

When distances to the animals are recorded in the field as radial distance from the observer in combination with the angle between the line and the detection, as is generally the case during ship-board line transect surveys for cetaceans, we can obtain both perpendicular and forward distances. This, in turn, enables us to fit a 2D detection function which, in combination with the MMPP, allows us to separate the models describing the detection and availability processes (Borchers and Langrock 2015). Now the detection function is conditional on availability and models solely the perception bias. For MRDS, on the other hand, these biases cannot readily be separated.

A further advantage of 2D detection functions is that, besides using the information of detecting the animal at specific locations defined by forward and perpendicular distances, they allow us to use the information of *not seeing* the animal between detections or from the horizon until it was detected some distance from the horizon. Borchers and Langrock (2015) developed the joint density of detections of an animal and its capture history that was tracked from the horizon until it passed the observer at zero forward distance. This joint density is based on an MMPP and a detection model. In order to fit such a model, the observers in the field must have tracked each detected animal from the first detection until the animal passed the observer at zero forward distance. These data are often not available as observers commonly only record the first detection for each animal. Even if animals were tracked, problems may arise as detection probabilities may change after initial detection.

In light of these potential issues, we present a modified version of the probability density developed by Borchers and Langrock (2015) in which only first detections are included. This was motivated by a study to investigate how effective lookouts (LO) from the ship's crew are in detecting marine mammals during military exercises on US Navy vessels. Trained marine mammal observers (MMOs) and LOs simultaneously monitored the areas ahead of the ship and recorded all detections of marine mammals. The team of MMOs generally consisted of two surveying MMOs, one liason MMO and one data MMO (see main text for details). As well as dedicated LOs, detections of marine mammals could also be made by other members of the ship's crew such as officers on the bridge or sonar technicians (although acoustic detections require visual confirmation). We refer to these observers together as the "lookout team" (LT). As it was possible for MMOs to be cued by the LT, the data had to be analysed under the trial configuration which we describe in detail below. Further, as it was unclear if resights were recorded consistently, we limited the data to first detections only.

We tested the performance of these methods under simulation and in a real-world case study. The simulation and case study including our findings are described below.

2 Methods

We use the methods first described in Langrock, Borchers, and Skaug (2013) for single-observer data and extended for double-observer data by Borchers and Langrock (2015) as a starting point and modify these according to our needs. Although conceptually the framework can deal with N discrete animal availability states, we limit our methods to two states as for our case study animals are either in state 1 (at the surface) or in state 2 (diving). Borchers and Langrock (2015) described three scenarios with different levels of a priori knowledge of the availability process, e.g. via published literature. We focus on the scenario termed "FKA" (full knowledge of the availability process) in these references which is the preferred method when detection data are limited. It allows populating the MMPP model with parameter values from the literature. Hence, we assume the parameters defining the availability process are known and only estimate detection function parameters.

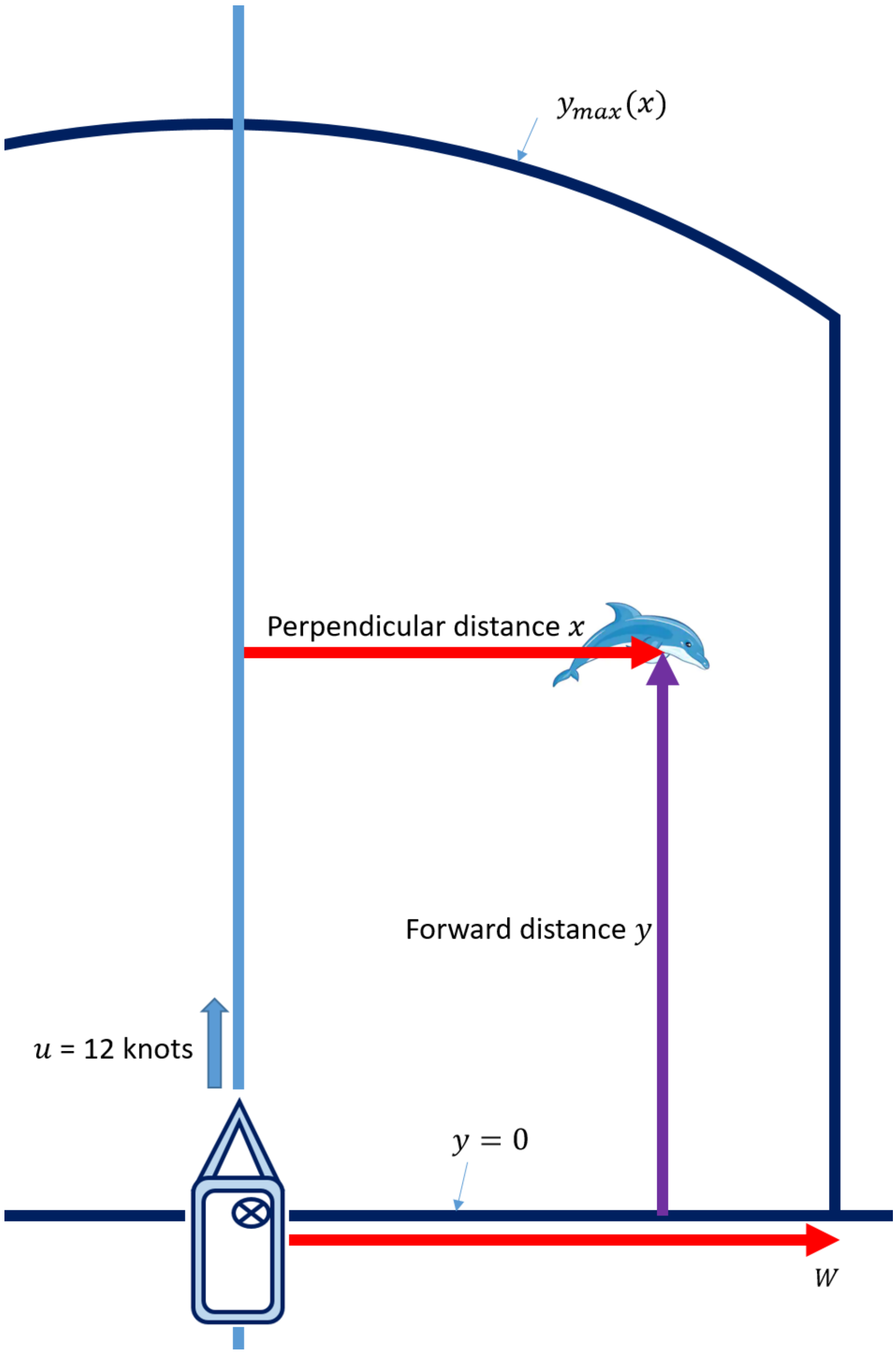


Figure 1: Schematic representation of detection process (redrawn from Borchers and Langrock 2015): as the ship moves along the trackline, animals are detected at perpendicular distance x and forward distance y. W is the furthest perpendicular distance included in the analysis.

Furthermore, Borchers and Langrock (2015) defined a probability density for the detection and capture history of the individual animals that is suitable for data collected under the independent observer (IO) configuration with two-way independence. We describe what changes when analysing data collected with one-way independence.

2.1 The availability process

We begin by describing the process by which animals become available at the surface for detection by the observers. Although other spatial resolutions are possible, we follow the example of Borchers and Langrock (2015) and use a resolution of 1m. We assume that animals only move vertically and not in the horizontal plane. The only movement in the horizontal plane is done by the ship moving along the trackline at a constant speed, say $u = 10$knots, i.e. 18,520m/hr or 309m/min (Fig. 1). Animals can be in one of two states, either near the surface and potentially available for detection (state 1) or diving (state 2). We assume that the duration in each state is exponentially distributed with parameter $-q_{kk}$ with $k = 1, 2$ (and expected value $1/-q_{kk}$). These parameters are expressed as 1/expected distance travelled by the ship while the animal is in the respective state. For example, if a whale spends on average 4min in state 1 and 20min in state 2 (and with the ship moving at 309m/min), parameter $-q_{11} = 1/(4 \times 309)$ and $-q_{22} = 1/(20 \times 309)$.

We model the transition between states using the transition matrix $\mathbf{Q}$ (also referred to as the infinitesimal generator matrix):

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \tag{1}$$

with $-q_{11} = q_{12}$ and $-q_{21} = q_{22}$ and row-sums equal zero.

The initial state distribution $\boldsymbol{\pi}$, i.e., the state the animal group is in when entering the detectable range, is a vector with two elements, the first and second giving the probability that the animal is in state 1 and 2, respectively. We assume that animals enter the visible range at random; hence, $\boldsymbol{\pi}$ is also assumed to be the stationary distribution and is given by $\pi_1 = q_{21}/(q_{12} + q_{21})$ and $\pi_2 = q_{12}/(q_{12} + q_{21})$.

Animals that switch between these two states can either be continuously at the suface and available for detection while in state 1 – which is refered to as intermittent availability (switching between periods of being continuously available and unavailable) – or instantaneously. In the latter case, we model the rate at which the animal is at the surface using an exponential distribution with parameter λ_1. This rate parameter is given as 1/the expected distance travelled by the ship between surfacings while in state 1. In our example, if an animal surfaces on average once every 0.5min, $\lambda_1 = 1/(0.5 \times 30.9)$, i.e. we expect that the animal surfaces once every 15.45m that the ship moves. As animals do not surface while in state 2, $\lambda_2 = 0$. In the case of intermittent availability $\lambda_1 = 1$.

2.2 The detection process

Here we address the question of detection probability given the animal is at the surface. We model this probability using a two dimensional detection function $h(y, x)$, where y and x are the forward and perpendicular distances between the observer and animal respectively (Fig. 1). We use $h(y, x)$ as the generic case of this function and use subscript notation to differentiate between observers. Specifically, we use $h_1(y, x)$, $h_2(y, x)$ and $h.(y, x) = 1 - (1 - h_1(y, x))(1 - h_2(y, x))$ to model the probability of detecting an animal for observer 1, observer 2 and for any observer, respectively. Note that the latter assumes independence between observers. Functions $h_1(y, x)$ and $h_2(y, x)$ can be modelled using e.g. the inverse power function (Borchers and Langrock 2015):

$$h(y, x) = \alpha \left. \quad \frac{\beta}{\sqrt{\beta^2 + x^2 + y^2}} \right)^{\gamma} \tag{2}$$

where parameters $\boldsymbol{\theta} = \alpha, \beta, \gamma$ are estimated for both observers, i.e. $\boldsymbol{\theta}_1 = \alpha_1, \beta_1, \gamma_1$ and $\boldsymbol{\theta}_2 = \alpha_2, \beta_2, \gamma_2$.

Even though $h(y,x)$ is a continuous function, we approximate it numerically (with arbitrary accuracy) by diving the range of forward distances into R intervals, $i_0 - i_1, i_1 - i_2, ..., i_{R-1} - i_R$ and assume $h(y,x)$ has a constant value within each interval. Generally, these are set up from the furthest forward distance at which animals can be detected $i_0 = y_{max}(0)$ to the beam of the ship at $i_R = 0$. We assume that $h(y,x)$ is constant within a given interval and use $h^{(r)}(x)$ to denote the value for the rth interval for short (Borchers and Langrock 2015). This value is generally calculated using the centre point of the interval, e.g., for the first interval using $h^{(1)}(x) = h((i_0 + i_1)/2, x)$.

Each of the N detected animals is detected once or more with $d = 1, 2, ..., D_n$ denoting the individual detections of the nth animal. While perpendicular distances are assumed to be the same for all detections of a given animal, forward distances decrease in terms of y, i.e. $y_1 > y_2 > ... > y_{D_n}$. Each y_d falls into the R_dth interval, $i_{R_d - 1} > y_d \geq i_{R_d}$. Two consecutive detections, e.g., the first and second, may fall into the same interval in which case $R_1 = R_2$. Borchers and Langrock (2015) used the notation $\Lambda(x)^{(r)} = diag(\lambda_1 h^{(r)}(x), \lambda_2 h^{(r)}(x), ..., \lambda_N h^{(r)}(x))$ for (in their notation) the N states. However, we only consider two states and know that the surfacing rate in state 2 equals 0 and, hence, $\lambda_2 h^{(r)(x)} = 0$ for state 2. Thus, our equivalent term is $\Lambda(x)^{(r)} = (\lambda_1 h^{(r)}(x), 0)$ for the rth interval.

2.3 The probability density function for the detections

2.3.1 Using all detections

Now we are in a position to set up the probability density function $f(\mathbf{y}, x)$ for the D_n detections of the nth animal using the components from above. We use subscripts ., 1, 2 to denote which $h(y,x)$ function was used, $h.(y,x)$, $h_1(y,x)$ or $h_2(y,x)$ (see above), for calculating the individual components. For brevity, we ommit subscript n in D_n. The probability density of detections at forward distances $\mathbf{y}$ and perpendicular distance x made by the two observers is:

$$f_y(\mathbf{y}|x) = \boldsymbol{\pi} \prod_{d=1}^{D} \{P.(y_{d-1}, y_d|x)\mathbf{\Lambda}.(x)^{(R_d)}\} P.(y_D, 0|x)\mathbf{1}^t \tag{3}$$

where $y_0 = y_{max}(x)$, $P.(y_{d-1}, y_d|x)$ is the probability that neither observer detected the animal between y_{d-1} and y_d, $\mathbf{\Lambda}.(x)^{(R_d)}$ is the probability that at least one of the observers detected the animal in the R_dth interval and $P.(y_D, 0|x)$ is the probability that neither observer detecteded the animal after the detection at y_D until it passes the beam at $y = 0$.

The $P.()$ components depend on if y_{d-1} and y_d (or y_D and 0) fall into the same interval. Should they fall into the same interval, we use:

$$P.(y_{d-1}, y_d|x) = \exp\left[\mathbf{Q} - \mathbf{\Lambda}.(x)^{(R_d)}(y_{d-1} - y_d)\right] \quad . \tag{4}$$

Should they fall into different intervals, we use:

$$\begin{aligned} P.(y_{d-1}, y_d|x) = &\exp\left[\mathbf{Q} - \mathbf{\Lambda}.(x)^{(R_{d-1})}(y_{d-1} - i_{R_{d-1}})\right] \\ &\times \prod_{r=R_{d-1}+1}^{R_d - 1} \exp\left[\mathbf{Q} - \mathbf{\Lambda}.(x)^{(r)}(i_{r-1} - i_r)\right] \\ &\times \exp\left[\mathbf{Q} - \mathbf{\Lambda}.(x)^{(R_d)}(i_{R_d - 1} - y_d)\right] \quad . \end{aligned} \tag{5}$$

2.3.2 Incorporating the mark-recapture component

2.3.2.1 Two-way independence

As we have two observers, we have a capture history $\mathbf{w}$ for each detected animal and denote the capture history of the individual detections using w_d. When using all detections, each detection represents one of three cases (Borchers and Langrock 2015): only observer 1 detected the animal during the surfacing ($w_d = (1,0)$), only observer 2 detected it ($w_d = (0,1)$) and both observers detected it ($w_d = (1,1)$):

- only observer 1 detected the animal ($w_d = (1,0)$)
- both observers detected the animal ($w_d = (1,1)$)
- only observer 2 detected the animal ($w_d = (0,1)$)

Hence, conditional on detection by either observer, the probability mass $p(w_d|y_d, x)$ of w_d is trinomial:

$$p(w_d|y_d,x) = \begin{cases} \frac{h_1(y_d,x)(1-h_2(y_d,x))}{h.(y_d,x)} & \text{if } w_d = (1,0) \\ \\ \frac{(1-h_1(y_d,x))h_2(y_d,x)}{h.(y_d,x)} & \text{if } w_d = (0,1) \\ \\ \frac{h_1(y_d,x)h_2(y_d,x)}{h.(y_d,x)} & \text{if } w_d = (1,1) \end{cases} \tag{6}$$

Incorporating the mark-recapture component, equation (3) becomes:

$$f_{yw}(\mathbf{y}, \mathbf{w}|x) = \boldsymbol{\pi} \prod_{d=1}^{D} \{P.(y_{d-1}, y_d|x)\boldsymbol{\Lambda}.(x)^{(R_d)} p(w_d|y_d, x)\} P.(y_D, 0|x)\mathbf{1}^t \quad , \tag{7}$$

where each w_d can be one of the three cases from equation (6).

2.3.2.2 One-way independence

For one-way independence, we describe the case that observer 2 is independent of observer 1, but not vice versa. This has consequences for how we define the capture history. The first detection can again be one of the three cases described above. Again, we condition on detection by either observer and use equation (6).

For subsequent detections, the capture history can be one of two cases, conditional on either observer detecting the animal: either observer 1 detected the animal and observer 2 did not ($w_d = (1,0)$) or observer 2 detected the animal ($w_d = (u,1)$, with u = unknown). In the latter case, we cannot distinguish between observer 1 detecting or not detecting the animal, i.e. $h_1(y_d,x)h_2(y_d,x) + (1 - h_1(y_d,x))h_2(y_d,x) = h_2(y_d,x)$, as detections by observer 1 may be cued by observer 2.

$$p(w_d|y_d,x) = \begin{cases} \frac{h_1(y_d,x)(1-h_2(y_d,x))}{h.(y_d,x)} & \text{if } w_d = (1,0); d > 1 \\ \\ \frac{h_2(y_d,x)}{h.(y_d,x)} & \text{if } w_d = (u,1); d > 1 \end{cases} \tag{8}$$

However, the probability density for one-way independence using all detections is given by equation (7).

2.3.3 First detections only

2.3.3.1 Two-way independence

We now consider how equation (7) changes in the case that only first detections were recorded by either observer, as is often the case during line transect surveys. Then, equation (7) can take three possible forms that depend on which observer detected the animal first.

If both observers detected the animal during the same surfacing, we use:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P_{\cdot}(y_{max}(x), y_1|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_1)}p(w_1|y_1, x)\} \tag{9}$$

with $w_1 = (1, 1)$ (equation (6)). Equation (9) uses the information of neither observer detecting the animal before y_1 and both observers detecting the animal at y_1. The component from equation (7) which pertains to not detecting it between y_D and $y = 0$ is ommitted as the animal is not tracked by either observer until it passes the beam.

If the animal is first detected by observer a with $a = 1, 2$ and during a later surfacing by observer b with $b = 1, 2$ and $a \neq b$, we use:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P_{\cdot}(y_{max}(x), y_1|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_1)}p(w_1|y_1, x)\}\{P_b(y_1, y_2|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_2)}p(w_2|y_2, x)\} \quad . \tag{10}$$

The capture history for the first detection is $w_1 = (1, 0)$ or $w_1 = (0, 1)$ (equation (6)), depending on if observer a is observer 1 or 2, respectively. The individual components of equation (10) use the information of neither observer detecting the animal before y_1, observer a detecting the animal at y_1, and observer b not detecting the animal at y_1 or between y_1 and y_2 and detecting it at y_2. We do not know if observer a detected the animal after y_1 as only first detections were recorded; hence, the probability mass for the second detection, w_2 is the sum of two probabilities and can be $w_2 = (0, 1) + (1, 1) = (u, 1)$ or $w_2 = (1, 0) + (1, 1) = (1, u)$ depending on if $a = 1$ or $a = 2$, respectively (equation (6)). Again, the animal is not tracked after y_2 until it passes the beam.

If the animal is only detected by observer a, we use:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P_{\cdot}(y_{max}(x), y_1|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_1)}p(w_1|y_1, x)\}P_b(y_1, 0|x)\mathbf{1}^t \quad , \tag{11}$$

with $w_1 = (1, 0)$ or $w_1 = (0, 1)$ (equation (6)), depending on if a is 1 or 2, respectively, and where the last component uses the information of observer b not detecting the animal from y_1 until it passed the beam.

In the independent observer configuration, either observer can be observer a for a given animal who detects it first.

2.3.3.2 One-way independence

We now describe the probability density for the case that only observer 2 is independent from observer 1 and not vice versa and using first detections only. Again, we have three possibilities for the capture histories of the first detection, each conditional on either observer detecting the animal:

1. both observers detected the animal ($w_1 = (1, 1)$)
2. only observer 1 detected the animal ($w_1 = (1, 0)$)
3. only observer 2 detected the animal ($w_1 = (0, 1)$)

The probability density depends on who detects the animal first and what follows.

We use equation (9) in the case that both observers detect the animal during the same surfacing.

In the case that only observer 1 detected the animal first and observer 2 at a later instance, the probability density is:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P_{\cdot}(y_{max}(x), y_1|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_1)}p(w_1|y_1, x)\}\{P_2(y_1, y_2|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_2)}p(w_2|y_2, x)\} \quad , \tag{12}$$

with $w_1 = (1, 0)$ from equation (6)) and $w_2 = (u, 1)$ from equation (8)).

In the case that only observer 1 detected the animal, we use:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P_{\cdot}(y_{max}(x), y_1|x)\boldsymbol{\Lambda}_{\cdot}(x)^{(R_1)}p(w_1|y_1, x)\}P_2(y_1, 0|x)\mathbf{1}^t \quad , \tag{13}$$

with $w_1 = (1, 0)$ from equation (6)). Again, the last component pertains to observer 2 not detecting the animal between $y = y_{max}(x)$ and $y = 0$.

In the case that observer 2 detected the animal first, the probability density becomes:

$$f_{yw}(\mathbf{y}|x) = \boldsymbol{\pi}\{P.(y_{max}(x), y_1|x)\boldsymbol{\Lambda}.(x)^{(R_1)}p(w_1|y_1, x)\} \quad , \tag{14}$$

with $w_1 = (0, 1)$.

2.4 The likelihood

In order to set up the likelihood for detecting N animals, we need the cumulative distribution of distance y to the first detection (Borchers and Langrock 2015). For detections by either observer this is given by:

$$F_{\cdot y}(y|x) = 1 - \boldsymbol{\pi}P.(y_{max}(x), y|x)\mathbf{1}^t \quad . \tag{15}$$

For detections by observer 1 or 2 this is given by:

$$F_{1y}(y|x) = 1 - \boldsymbol{\pi}P_1(y_{max}(x), y|x)\mathbf{1}^t \ ; \quad F_{2y}(y|x) = 1 - \boldsymbol{\pi}P_2(y_{max}(x), y|x)\mathbf{1}^t. \tag{16}$$

Now we can define the likelihood for the detection function parameters $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$:

$$L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \prod_{n=1}^{N} \frac{f_{yw}(\mathbf{y}, \mathbf{w}|x_n)\pi(x)}{\int_0^W F_{\cdot y}(0|z)\pi(z)dz} \quad , \tag{17}$$

where W is the perpendicular truncation distance and $\pi(x)$ describes the distribution of perpendicular distances (often assumed the be uniform in DS surveys). Which probability density function $f_{yw}(\mathbf{y}, \mathbf{w}|x_n)$ is used depends on whether data were collected under the two-way or one-way independence as well as whether all detections were included and, in the case of first detections only, which observer detected a given animal first (see equations (7) – (11)). In the case of line transect surveys with randomly placed transects $\pi(x) = 1/W$ and cancels in equation (17).

2.5 Quantities of interest

Equations (15) and (16) are also used to obtain an estimate of the detection probability on the line, often referred to as $g0$, or as well as the effective strip half-width ($ESHW$). The former can be obtained, e.g. for both observers combined, using:

$$g.0 = F_{\cdot y}(0|0) \quad . \tag{18}$$

The ESHW is the perpendicular distance beyond which as many animals were detected as were missed within and can be obtained using, e.g. for both observers combined:

$$ESHW. = \int_0^W F_{\cdot y}(0|x)\pi(x)dx \quad , \tag{19}$$

Furthermore, we can use equations (4) and (5) to calculate the probability PrU that an animal remains undetected at a certain radius around the ship. To estimate, e.g. the probability that an animal remains undetected by both observers at a 100m radius around the ship $PrU._{(100m)}$ we use:

$$PrU._{(100m)} = \int_{x=0}^{100} P_{\cdot y}(y_{max}(x), \sqrt{100^2 - x^2}|x) \quad . \tag{20}$$

3 Simulation study

To address the performance of these methods, we conducted a simulation study during which we generated detection data similar to a real-world scenario. We then analysed these data using the functions described above in order to ascertain if parameters and quantities of interest could be estimated reliably without bias. Testing methods via simulation has the advantage of knowing the true values of the parameters – as these were used to generate the data – and quantities of interest – as these can be calculated using the true parameter values.

3.1 Generating detections

To simulate each detection, we began by generating a perpendicular distance x drawn from a uniform distribution with boundaries $\{0, W\}$ where $W = 6,500$m. From this we calculated $y_{max}(x) = \sqrt{y_{max}(0)^2 + W^2}$ where $y_{max}(0) = 15,725$m. We then simulated in several steps a ship passing each of these animals from $y_{max}(x)$m to $y = 0$m forward distance – while x remained fixed – and making detections. These steps comprized simulating the availability of the animal to be detected by generating surfacings and simulating detecting it given it was available at the surface for detection. The surfacings were generated using a combination of random state-switching and random surfacing while in state 1.

For the state-switching, we first drew a sample of which state (1 or 2) the animal was in when entering the visible range at $y_{max}(x)$ with probabilities proportional to the expected time spent in states 1 and 2, which were known. Thereafter, the animal alternated between the two states. For each time the animal was in one of the two states, we drew a random sample z of how long the animal was in this state (expressed in meters travelled by the ship while in this state) from an exponential distribution with parameters $\boldsymbol{\mu_{sim}}$ (1/distance travelled by the ship during the expected time spent in states 1 or 2). For the first state, we added the extra step of drawing a random sample from a uniform distribution ($U(0, z_1)$) to mark the point at which the animal entered the visual horizon of the ship at forward distance $y = y_{max}(x)$.

In the case that the first state was 1, we drew a sequence of random samples from an exponential distribution with parameter λ_1 to determine the location of surfacings, until the animal switched into state 2. Parameter $\lambda_{1_{sim}}$ was 1/distance travelled by the ship during the expected time between surfacings while in state 1.

For each surfacing, we determined whether each observer detected it by drawing random samples from a Bernoulli distribution with probability calculated using equation (2) with known perpendicular and forward distances from the respective surfacing and known $\boldsymbol{\theta}_{1_{sim}}$ and $\boldsymbol{\theta}_{2_{sim}}$.

The above procedure generated a detection history for a single individual (or pod, although detection probability is not pod-size dependent in this simulation study). In some cases individuals passed abeam without a single detection, and these were discarded as they would not form part of an observed dataset. We repeated this procedure until we had generated $N_{sim} = 300$ detection histories each with at least one detection, and this set of 300 detection histories constituted a single dataset.

A new data set was created for each iteration of the simulation. When testing the methods on first detections only, the data set was reduced to first detections by either observer.

3.2 Analysis of simulated data

We generated 100 data sets in the manner described above and completed two analyses on each to estimate detection function parameters ($\boldsymbol{\theta_{1_{sim}}}$ and $\boldsymbol{\theta_{2_{sim}}}$):

1. Two-way independence, using all detections
2. One-way independence, using first detections only

We only kept the full data set for analysis 1 (see previous section). For all analyses the likelihood was based on equation (7). We used the approach "FKA" from Borchers and Langrock (2015) during which the values for the parameters defining the availability model ($\boldsymbol{\mu_{sim}}, \lambda_{1_{sim}}$) were known. Hence, any bias arising in the estimates would be due to bias in the estimated detection function.

Analyses were conducted in R version 4.0.5 (R Core Team 2021) using R packages expm (Goulet et al. 2021) for matrix exponentiating and mvtnorm (Genz and Bretz 2009, Genz et al. (2019)) for random sampling multivariate normal distributions. We built on functions written by Borchers and Langrock (2015) and adjusted these for one-way independence and first detections only scenarios. We used function nlm to minimise the negative log-likelihood to obtain best estimates of detection hazard parameters. These were then used together with the known availability process parameters to obtain estimates of various quantities of interest for each iteration: the effective strip half-width $ESHW$, the trackline detection probability $g0$ and the probability of remaining undetected PrU until entering a certain radial distance in front of the ship, for three distances of interest: 200yds, 500yds and 1,000yds.

We used both visual and quantitative methods to assess performance of the methods. We visually assessed model fit by comparing the true detection hazard with the estimated detection hazard, and comparing the true probability density of distance to first detection ($F_{1y}(y|x)$ and $F_{2y}(y|x)$ from equation (16)) with the estimated equivalent. The results section contains these comparisons for a single example (data set 1 of the simulation and the estimates from the models fitted to this data set).

We calculated the mean percent bias (mean((estimate-true value)/true value) $\times$ 100) across all iterations to quantify how close the estimates were to the true values. True values for each of the quantities ($ESHW$, $g0$ and PrU) were calculated using the known detection function parameters instead of the estimates. If methods performed well, we expected the mean percent bias to be small.

3.3 Results

3.3.1 Visual comparison

Visual comparison of the estimated $\hat{h}_1(x,0)$y and $\hat{h}_2(x,y)$ with the true detection hazards ($h_1(x,y)$ and $h_2(x,y)$) revealed that the estimates from both analyses were close to the truth. The examples depicted in Figures 2 and 3 show that the general shape was very similar although the shapes of the estimated hazards from analyses 1 and 2 were slightly flatter at small x and y and the estimates $\hat{h}_1(0,0)$ and $\hat{h}_2(0,0)$ were lower than $h_1(0,0)$ and $h_2(0,0)$. For this example, these differences were seemingly not more striking for analysis 2 (Figure 3) than for analysis 1 (Figure 2) even though less information was used for the former for fitting the model.

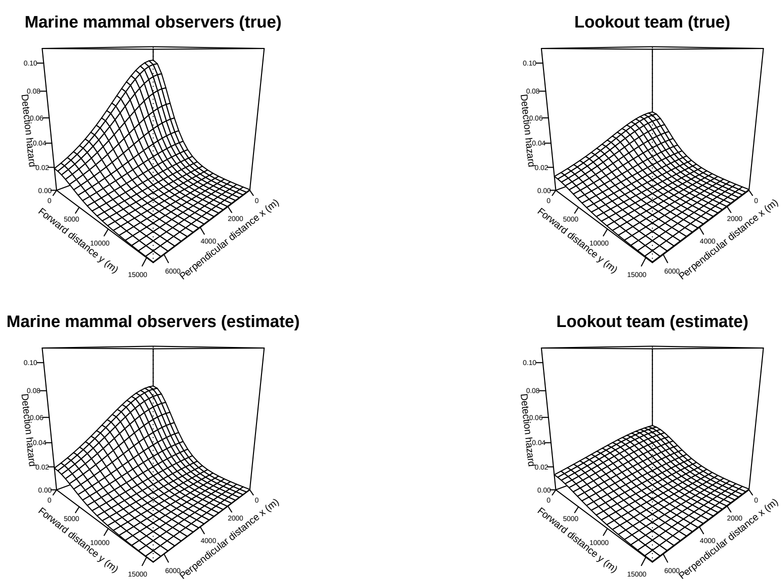


Figure 2: True and one example of the estimated detection hazard shown with respect to perpendicular distances and forward distances from the ship for each platform from analysis 1 with two-way independence and using all detections.

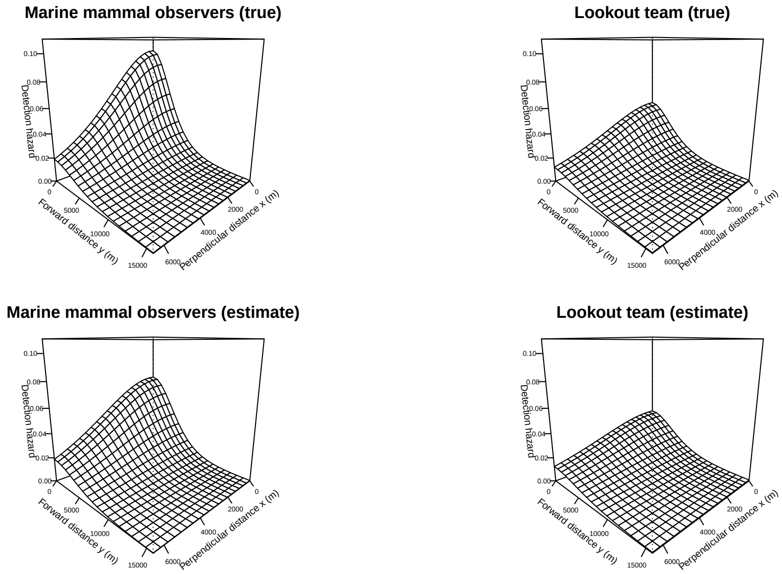


Figure 3: True and one example of the estimated detection hazard shown with respect to perpendicular distances and forward distances from the ship for each platform from analysis 2 with one-way independence and using first detections.

The Figures depicting the estimates of the cumulative distribution of first sightings from this example revealed further slight discrepancies between the estimates and true values (Figures 4) and 5. Note that in both figures, the black contours are generally inside the green (with the exception of lines for 0.2 and 0.3 for the lookout team), indicating that for this example, cumulative distributions were generally slightly overestimated within the range depicted in these plots.

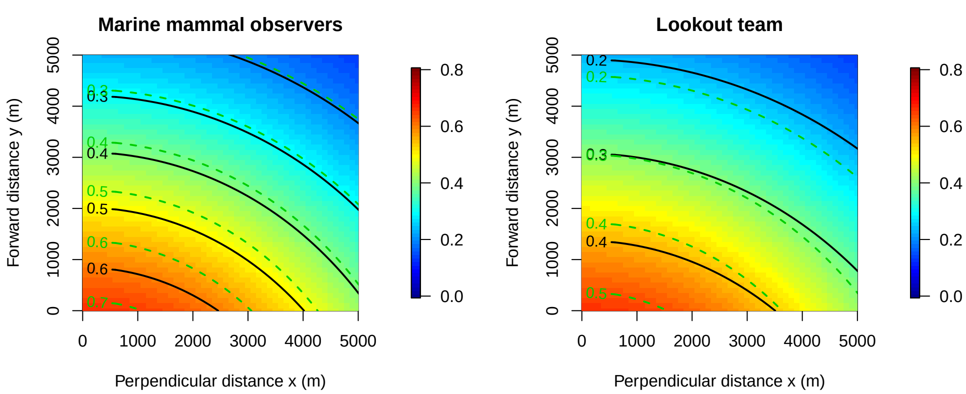


Figure 4: Example of the estimated cumulative distributions of distances to first detection from analysis 1 with two-way independence and using all detections, shown for each platform. Black contours refer to the estimates, green to the equivalent values calculated with the true parameters.

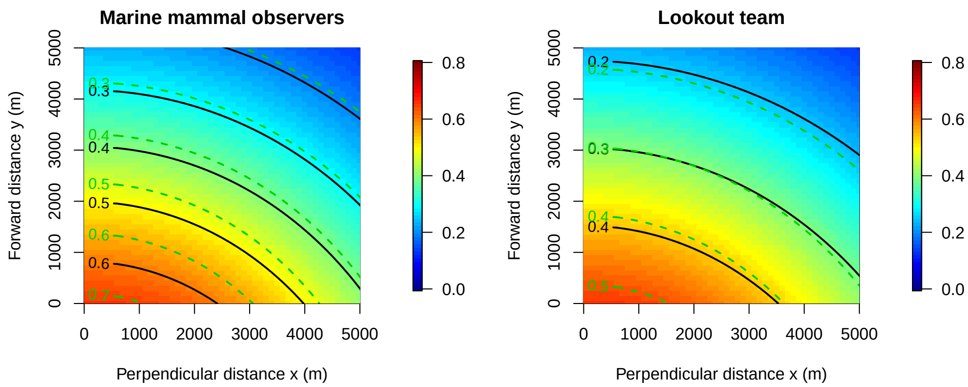


Figure 5: Example of the estimated cumulative distributions of distances to first detection from analysis 2 with one-way independence and using first detections, shown for each platform. Black contours refer to the estimates, green to the equivalent values calculated with the true parameters.

3.3.2 Evaluation of biases

All biases in the estimated quantities from both analyses of the simulated data were minor, i.e. less than three percent (Table 1). Biases were generally slightly higher for analysis 2 with one-way independence and using first detections only. Biases were also generally slightly higher for observer 2. In both cases this was likely due to less detection data when using first detections only (compared to all detections in analysis 1) and for observer 2 compared to observer 1.

Table 1: Mean percent bias in estimates of the effective strip half-width ($ESHW$), the trackline detection probability $g0$ and the probability of remaining undetected PrU calculated for 200yds, 500yds and 1,000yds for analyses 1 and 2 of the 1,000 simulated data sets described above; TWI: two-way independence, OWI: one-way independence, all: using all detections, first: using first detections only, Comb: observers 1 and 2 combined.

	% Bias					
	1: TWI + all			2: OWI + first		
	Comb	Obs 1	Obs 2	Comb	Obs 1	Obs 2
ESHW	0.119	0.112	-0.341	-0.912	-1.458	-1.245
$g0$	-0.337	-0.196	-1.868	-0.611	-0.567	-2.459
PrU_{200yds}	1.464	0.430	1.928	2.804	1.411	2.573
PrU_{500yds}	1.275	0.343	1.720	2.691	1.411	2.350
$PrU_{1,000yds}$	0.945	0.197	1.356	2.504	1.433	1.957

4 Case study

4.1 Data

In this study, we analysed the sightings data collected during 27 cruises between November 2010 and September 2019 (Table 2). We used the furthest forward and perpendicular detection distances as $y_{max}(0)$=15,725m and perpendicular truncation distance W=13,425m.

Table 2: Cruises included in this study, indicated by vessel code (Vessel), month and year of the cruise, and the number of sightings. Study areas included Atlantic Fleet Training and Testing (AFTT), Hawaii Range Complex (HRC) and the Southern California Range Complex (SOCAL).

Vessel	Month	Year	Sightings	Study area
FFG-A	Feb	2010	21	HRC
DDG-A	Mar	2010	11	AFTT
DDG-B	Jun	2010	15	AFTT
DDG-C	Jul	2010	84	SOCAL
CG-A	Nov	2010	7	HRC
DDG-D	Feb	2011	29	HRC
DDG-E	Apr	2011	21	SOCAL
DDG-F	Nov	2011	5	HRC
DDG-G	Feb	2012	13	HRC
FFG-B	May/Jun	2012	24	AFTT
DDG-H	Jul	2012	62	SOCAL
DDG-I	Feb	2013	5	HRC
DDG-J	Aug	2013	2	HRC
DDG-K	Jan	2014	57	HRC
CG-B	Feb	2014	7	HRC
CG-C	Aug	2014	23	AFTT
DDG-L	Feb	2015	34	HRC
DDG-M	Apr	2015	3	AFTT
DDG-N	Feb	2016	12	HRC
DDG-O	Mar/Apr	2016	52	AFTT
DDG-P	Aug	2016	44	AFTT
DDG-Q	Aug	2017	56	AFTT
DDG-R	Feb	2018	22	HRC
DDG-S	Jun	2018	34	AFTT
DDG-T	Feb	2019	30	HRC
CG-D	Mar	2019	15	AFTT
DDG-U	Sep	2019	28	AFTT
Cruises	27	Total	716	

For analyses, the sightings data were split into different groups according to their similarity in times spent in states 1 and 2 and their surfacing rate while in state 1 as well as similarity in detectability, to the extent possible. Data from the same groups were analysed together using a single value for each of these three parameters describing the availability model (times spent in states 1 and 2 and the surfacing rate while in state 1) which were obtained from the literature and detailed below. For both whales and small cetaceans, we formed two groups each which are detailed in the following. No detections on beaked whales were made during the 27 cruises; hence, beaked whales were not considered here. Whales were roughly grouped into rorquals and sperm whales, although the former included all unidentified whales which may have been sperm whales or beaked whales.

Small cetaceans were split into two groups by pod size: 1. pods of 6 or less (this includes pod sizes of one) and 2. pods of more than 6. We assumed that for the former, animals would surface instantaneously at a given rate while in state 1 (similar to the whale groups), while for the latter, at least one animal of a given sighting would be at the surface at all times while in state 1 providing continuous availability to be detected.

4.1.1 Rorquals

The rorquals group included all rorquals, unidentified large whales and unidentified whales with a total of 277 sightings (Table 3).

Table 3: Number of sightings included in the Rorqual group.

Code	Species	Sightings
BAL–	Balaenoptera sp.	29
BALAC	Balaenoptera acutorostrata	3
BALMU	Balaenoptera musculus	16
BALPH	Balaenoptera physalus	1
MEGNO	Megaptera novaeanglia	84
LGWHA	Unidentified large whale	50
WHALE	Unidentified whale	94
	Total	277

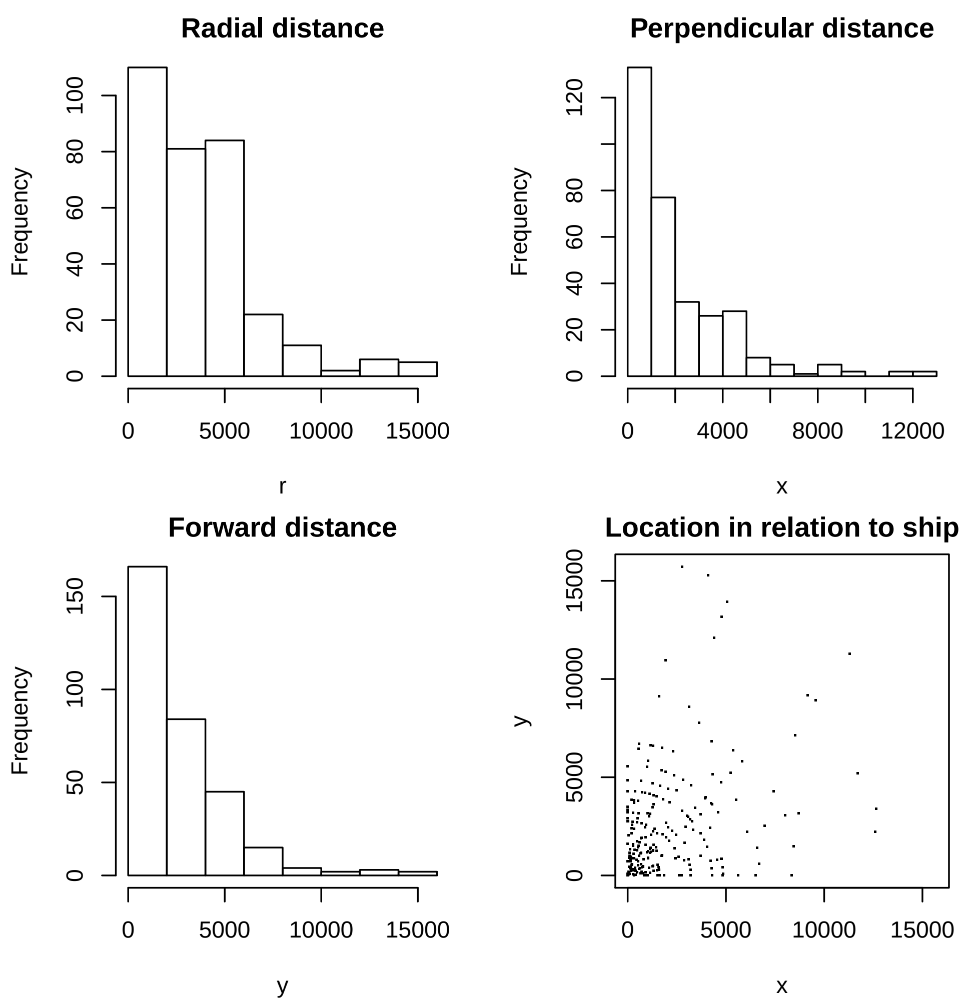


Figure 6: Histograms of radial, perpendicular and forward distances of first detections for the Rorqual group as well as their location in relation to the ship. Detections made on the left side of the ship were folded to the right side.

We used values for the expected times spent in states 1 (2.4min) and 2 (4.2 min) as well as the mean surfacing (blow) interval (0.3min) from Dolphin (1987) who estimated these values for humpback whales, i.e. the species with the highest number of sightings in this group. The values represent the expected times in the respective states for combined non-feeding behaviour. A second analysis was undertaken with alternative values (estimates for overall behaviors) taken from Dophin (1987) in order to evaluate sensitivity to reasonable initial conditions (Table 4).

Table 4: Parameter values used for the availability model for the rorquals group including the expected time spent in states 1 and 2 as well as the interval between surfacings (IBS) while in state 1. Analysis 1 represents the main analysis; analysis 2 was part of the sensitivity analysis.

Analysis	State 1 (sec)	State 2 (sec)	IBS_1 (sec)	Source
1	156	252	18	Dolphin (1987)
2	180	66	18	Dolphin (1987)

4.1.2 Sperm whales

Only two sightings of sperm whales were recorded during all cruises combined. Therefore, we used the estimates of detection function parameters from the rorqual group, assuming that given the animals were at the surface, sperm whales shared the same detection function as rorquals. We used the best estimates from Drouot, Gannier, and Goold (2004) for the expected times spent in states 1 (9.1min) and 2 (44.8min) as well as the mean blow interval (0.217min) as the main analysis (analysis 1, Table 5). Worst case and best case scenarios in terms of being available for detection were explored as part of a sensitivity analysis using minima and maxima from Drouot, Gannier, and Goold (2004). Analyses 2 represented the worst case scenario with minimal time available for detection, analyses 3 the best case scenario with maximum time available for detection.

Table 5: Parameter values used for the availability model for the sperm whales including the expected time spent in states 1 and 2 as well as the interval between surfacings (IBS) while in state 1. Analysis 1 represents the main analysis, analyses 2 and 3 were part of the sensitivity analysis.

Analysis	State 1 (sec)	State 2 (sec)	IBS_1 (sec)	Source
1	546	2688	13.04	Drouot et al. (2004)
2	360	3300	16.37	Drouot et al. (2004)
3	780	1860	10.51	Drouot et al. (2004)

4.1.3 Small cetaceans in small pods (SCSP group)

198 pods were sighted for this group during the 27 cruises (Table 6). These were mostly of various species or higher taxonomic groups of the delphinid family; however, other small cetaceans may have been included via the unidentified small cetaceans. Sightings recorded as bowriding were excluded from the analysis. Hence, the total number of sightings and detections included in the analyses were 178 and 201, respectively.

Table 6: Number of sightings by species included in the group of small cetaceans in small pods.

Code	Species	Sightings
STEBR	Steno bredanensis	1
STECO	Stenella coeruleoalba	1
STEFR	Stenella frontalis	33
STE–	Stenella sp.	3
TURTR	Tursiops truncatus	24
DELDE	Delphinus delphis	5
DEL–	Delphinus sp.	1
GLOMA	Globicephala macrorhynchus	5
GLO–	Globicephala sp.	3
GRAGR	Grampus griseus	2
ORCOR	Orcinus orca	1
BLACK	Blackfish	2
PHOPO	Phocoena phocoena	1
MIXED	–	2
DOLPH	Unidentified dolphin	110
SMALL	small unidentified cetaceans	4
	Total	198

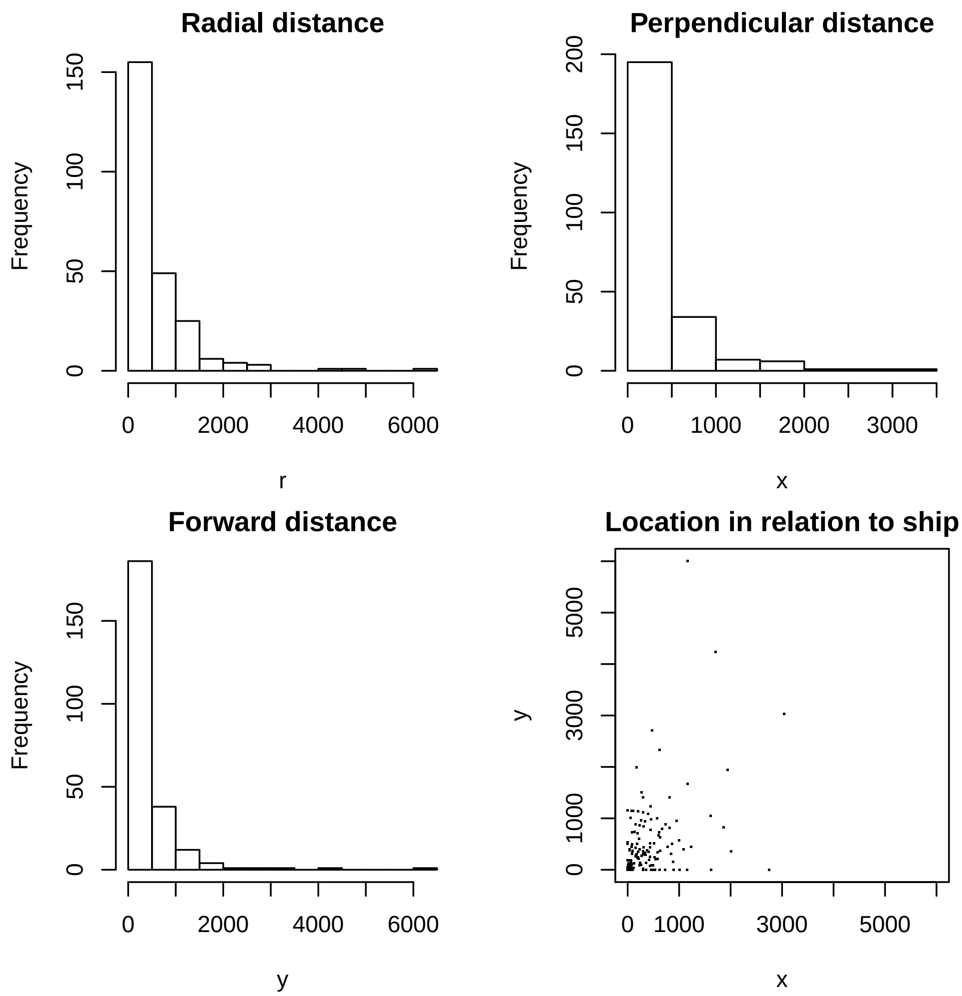


Figure 7: Histograms of radial, perpendicular and forward distances of first detections for the SCSP group as well as their location in relation to the ship. Detections made on the left side of the ship were folded to the right side.

As the genus *Stenella* was predominant in this group, we used best estimates for the expected times spent in states 1 (0.99min) and 2 (1.26min) as well as the mean surfacing interval (0.1min) in state 1 from Scott and Chivers (2009) in analysis 1. Further values were explored as part of a sensitivity analysis: analysis 2 uses the maximum dive time from Scott and Chivers (2009) for state 2. Values for analyses 3 and 4 were taken from Palka et al. (2017) and represent best estimates for *Delphinus sp.* and *Tursiops truncatus*.

Table 7: Parameter values used for the availability model for the SCSP group including the expected time spent in states 1 and 2 as well as the interval between surfacings (IBS) while in state 1. Analysis 1 represents the main analysis; analyses 2-4 were included in the sensitivity analysis. λ_1=1 refers to intermittent availability.

Analysis	State 1 (sec)	State 2 (sec)	IBS_1 (sec)	Source
1	59.4	75.6	6	Scott and Chivers (2009)
2	59.4	258	6	Scott and Chivers (2009)
3	44	59.4	6	Palka et al. (2017)
4	3.03	26.6	NA (λ_1=1)	Palka et al. (2017)

4.1.4 Small cetaceans in large pods (SCLP group)

The 94 sightings of this group also included predominantly species of the delphinid family but also four sightings of small cetaceans which may not have been delphinids (Table 8). Sightings recorded as bowriding were excluded from the analysis leaving a total of 87 sightings and 94 detected surfacings for the analysis.

Table 8: Number of sightings included in the group of small cetaceans in large pods.

Code	Species	Sightings
STEAT	Stenella attenuata	1
STECO	Stenella coeruleoalba	2
STEFR	Stenella frontalis	17
STELO	Stenella longirostris	2
TURTR	Tursiops truncatus	6
DELCA	Delphinus capensis	3
DELDE	Delphinus delphis	7
DEL–	Delphinus sp.	10
GLOMA	Globicephala macrorhynchus	10
GRAGR	Grampus griseus	5
LAGAC	Lagenorhynchus acutus	1
MIXED	–	3
DOLPH	Unidentified dolphin	23
SMALL	small unidentified cetaceans	4
	Total	94

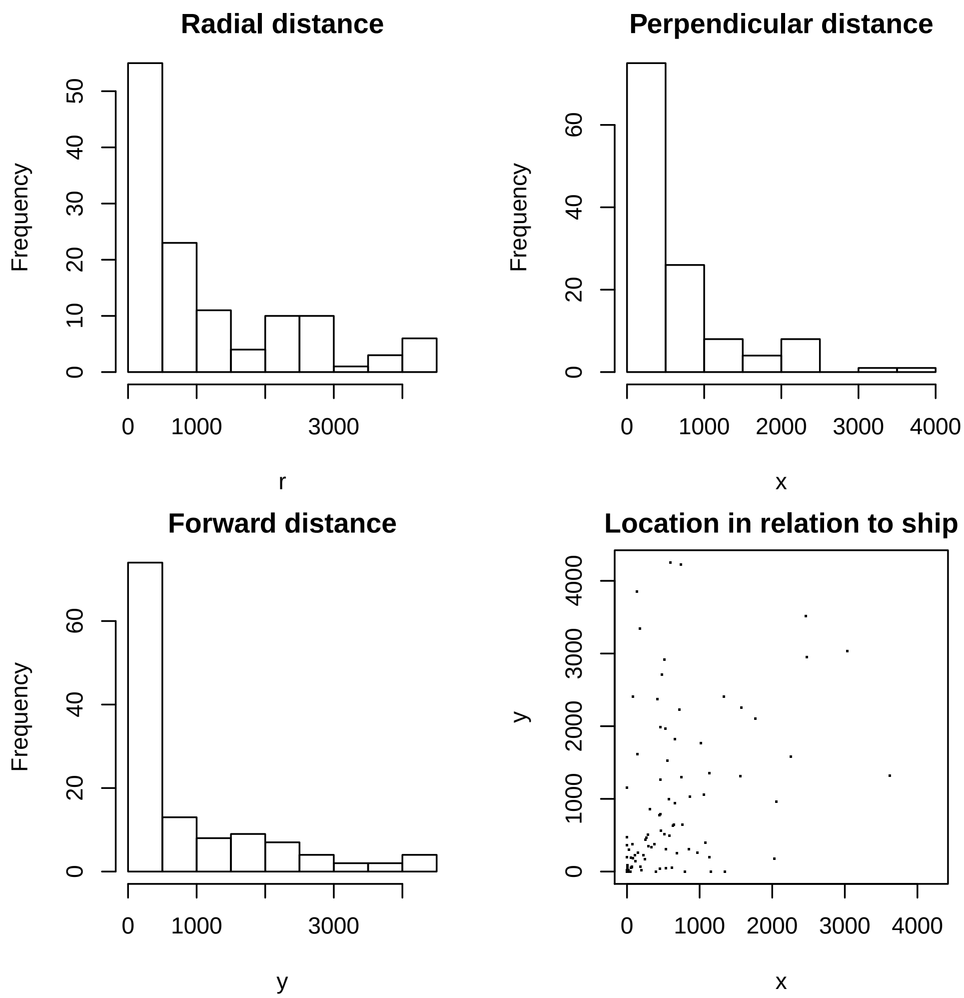


Figure 8: Histograms of radial, perpendicular and forward distances of first detections for the SCLP group as well as their location in relation to the ship. Detections made on the left side of the ship were folded to the right side.

For this group we assumed continuous availability at the surface. In order to use the framework described above, we modeled this using intermittent availability with an expected 60min in state 1 and 1sec in state 2. The parameter pertaining to the surfacing rate λ_1 was set to 1. The only alternative we explored in a sensitivity analysis was increasing the number in state 2 to 1min.

Table 9: Parameter values used for the availability model for the SCLP group including the expected time spent in states 1 and 2 as well as the interval between surfacings (IBS) while in state 1. Analysis 1 is the main analysis, analysis 2 was part of the sensitivity analysis. λ_1=1 refers to intermittent availability.

Analysis	State 1 (sec)	State 2 (sec)	IBS_1 (sec)	Source
1	3600	1	NA (λ_1=1)	NA
2	3600	60	NA (λ_1=1)	NA

4.2 Analyses

Resights were often recorded by the MMOs and sometimes by the LT. However, we only used first detections made by any MMOs or the LT as it was unclear whether this effort was consistent throughout the study or whether it was density dependent or influenced by distractions on or near the platform. Also, as MMOs may have been cued by the LT (as part of the survey protocol, see above), we used the approach described above with one-way independence between observers, i.e. the LT (observer 2) being independent of MMOs (observer 1) but not vice versa. In particular, we used the probability densities from equations (9), (12), (13) and (14) for setting up the likelihood (equation (17)).

Analyses were conducted using the same R packages and functions as for the simulation study. Again, parameter estimates from the fitted models were used to obtain estimates of the strip half-width $ESHW$, the trackline detection probability $g0$ and the probability of remaining undetected PrU until entering a certain radial distance in front of the ship, for three distances of interest: 200yds, 500yds and 1,000yds.

A parametric bootstrap was conducted for the main analysis of each group to quantify uncertainty in the estimates. To this end, we drew 999 new multivariate samples using the parameter estimates and covariance matrices from the fitted models. New samples were used to obtain 999 estimates of the quantities described above amounting to 1,000 estimates including the original estimate. 95% confidence intervals were obtained using the 2.5% and 97.5% quantiles from the 1,000 estimates.

4.2.1 Goodness of fit

We assessed goodness-of-fit using visual depictions of the number of detections which we expect to see under the model in various distance bins vs the number of detections that were seen in those bins. This required obtaining an estimate of the cumulative probability of observing a pod at least by distance y for a given x as well as the probability density function (PDF) of observing a pod at y, x. These can be obtained using the model parameters and the functions defined above.

While $F(y|x)$ (equation (15)) is the cumulative probability, the first derivative of this function, $F'_w(y|x)$ is the weighted probability density function (pdf). Obtaining $F'_w(y|x)$ mathematically is intractable; hence, we used numerical approximation by calculating the slope of $F(y|x)$ at a given y, x using numerical approximation:

$$F'(y|x_i) \approx \frac{F(y-5|x_i) - F(y+5|x)}{10} \quad . \tag{21}$$

These were then weighted using:

$$F'_w(y|x_i) = \frac{F'(y|x_i) \times F(y|x_i)}{\sum_y F'(y|x_i)} \quad . \tag{22}$$

The binning was done in the following manners:

1. along the perpendicular distance axis: from 0 to W
2. along the forward distance axis: from 0 to $y_{max}(0)$
3. using radial distance bins: from 0 to the maximum observed radial distance

For each binning, the number of observed first detections in each bin was determined and then compared with the following. For 1., the comparison was done with $F(y = 0|x)$ calculated for each perpendicular distance bin. For 2., the comparison was done using $\sum_x F'_w(y|x)$ calculated for each forward distance bin. For 3., the comparison was done for using $\sum_r F'_w(y|x)$ each radial distance bin, where r indicates which radial distance bin coordinates y, x fall into.

We assessed goodness-of-fit quantitatively via a χ^2-test using:

$$\chi^2 = \sum_{b=1}^{B} \frac{(n_b - n\pi_b)^2}{n\pi_b} \tag{23}$$

where $b = 1, 2, 3, ..., B$ refer to the radial distance bins with boundaries $i_0, i_1, i_2, ..., i_B$ (with $i_0 = 0$), n_b is the number of first detections that fell within the bth distance bin, π_b is the probability density of observations in distance band b calculated by integrating the pdf over the area covered by the respective distance bin. Here, binning was done in a manner that the expected number of first detections that fell within each bin was approximately equal. Test statistics were calculated individually for MMOs and the LT using $F'_1(y|x)$ and $F'_2(y|x)$. As modeling the detection process required estimating three parameters for each observer, we used $B - 3$ as the degrees of freedom for the tests.

4.3 Results

4.3.1 Rorquals

The estimated detection hazard for the rorqual group was very flat and near zero in most areas for both platforms, only exhibiting a sharp rise within <5,000m around the ship for MMOs and within <2,000m for the LT (Figure 9).

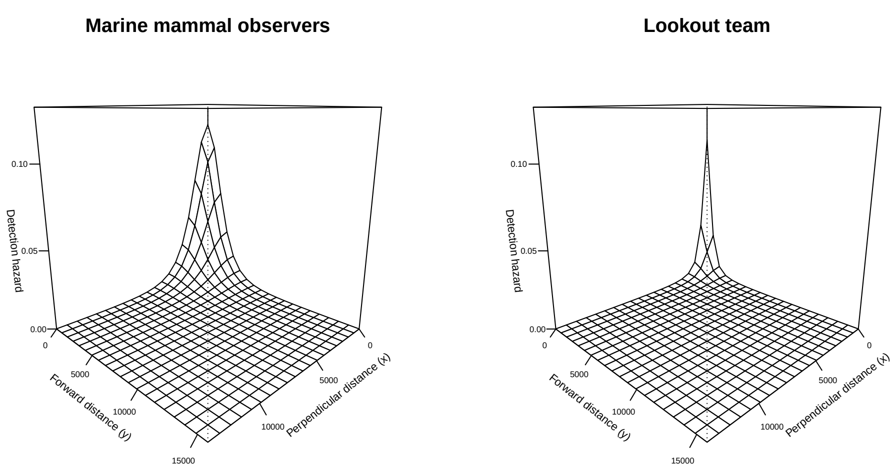


Figure 9: Estimated detection hazard shown with respect to perpendicular distances and forward distances from the ship for each platform from the model fitted to the detection data of the rorquals group assuming one-way independence and using first detections.

The generally higher estimates of the detection hazards for the MMOs was also the cause for the higher estimated cumulative distribution of distance y to first detection (Figure 10). These were much higher out to greater distances for the MMOs compared to the LT.

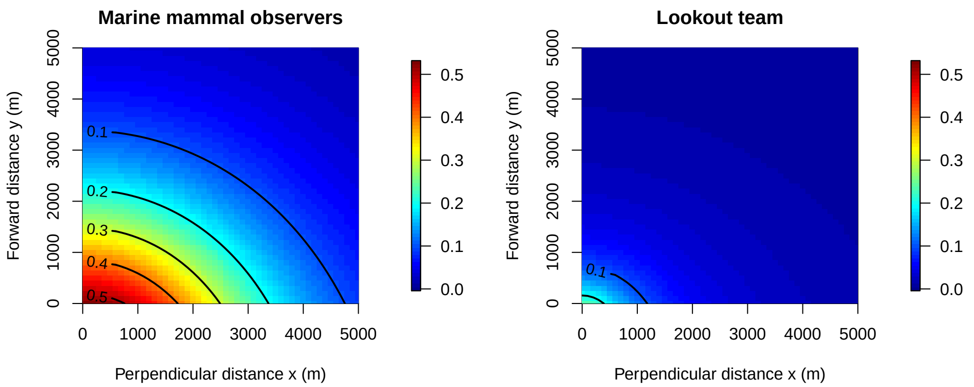


Figure 10: Estimated cumulative distribution of distance y to first detection from the model fitted to the detection data of the rorqual group assuming one-way independence and using first detections, shown for each platform. Black lines indicate contours of the cumulative distribution.

The estimated effective strip half-width and trackline detection probability were much higher for MMOs than the LT (Table 10). The estimated probability of animals of this group entering a 200yds radius around the ship undetected was much higher for the LT at 0.800 (95% CI: 0.740–0.859) compared to MMOs at 0.493 (95% CI: 0.399–0.593). The equivalent probabilities for a 500yds radius were 0.853 (95% CI: 0.803–0.893) for the LT and 0.527 (95% CI: 0.434–0.620) for MMOs. For a 1,000yds radius they were 0.913 (95% CI: 0.869–0.936) for the LT and 0.592 (95% CI: 0.507–0.673) for MMOs.

Table 10: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates and 95% confidence intervals (CIs) for the rorquals group.

	Both platforms combined		MMOs		LT	
	Best	CI	Best	CI	Best	CI
ESHW	1945	(1605 – 2368)	1739	(1396 – 2126)	408	(310 – 579)
g0	0.595	(0.508 – 0.674)	0.528	(0.425 – 0.622)	0.237	(0.161 – 0.314)
PrU_{200yds}	0.432	(0.353 – 0.518)	0.493	(0.399 – 0.593)	0.800	(0.740 – 0.859)
PrU_{500yds}	0.477	(0.395 – 0.561)	0.527	(0.434 – 0.620)	0.853	(0.803 – 0.893)
$PrU_{1,000yds}$	0.555	(0.473 – 0.634)	0.592	(0.507 – 0.673)	0.913	(0.869 – 0.936)

Results from analysis 2, conducted as part of the sensitivity analysis, were similar to results from analysis 1 (Table 11).

Table 11: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates of sensitivity analysis for the rorquals group.

	Analysis 1			Analysis 2		
	both	MMOs	LT	both	MMOs	LT
$ESHW$	1945	1739	408	1982	1775	416
$g0$	0.595	0.528	0.237	0.605	0.539	0.236
PrU_{200yds}	0.432	0.493	0.800	0.423	0.482	0.800
PrU_{500yds}	0.477	0.527	0.853	0.468	0.518	0.851
$PrU_{1,000yds}$	0.555	0.592	0.913	0.547	0.584	0.910

Figure 11 illustrates that the fit of the model with respect to the perpendicular distance was poor for MMOs and the LT – underfitting at distances <1,500m and often overfitting at distances >1,500m. The fit of the model with respect to forward distance was overall better; however, overfitting at distances >2,000m occurred for the MMOs.

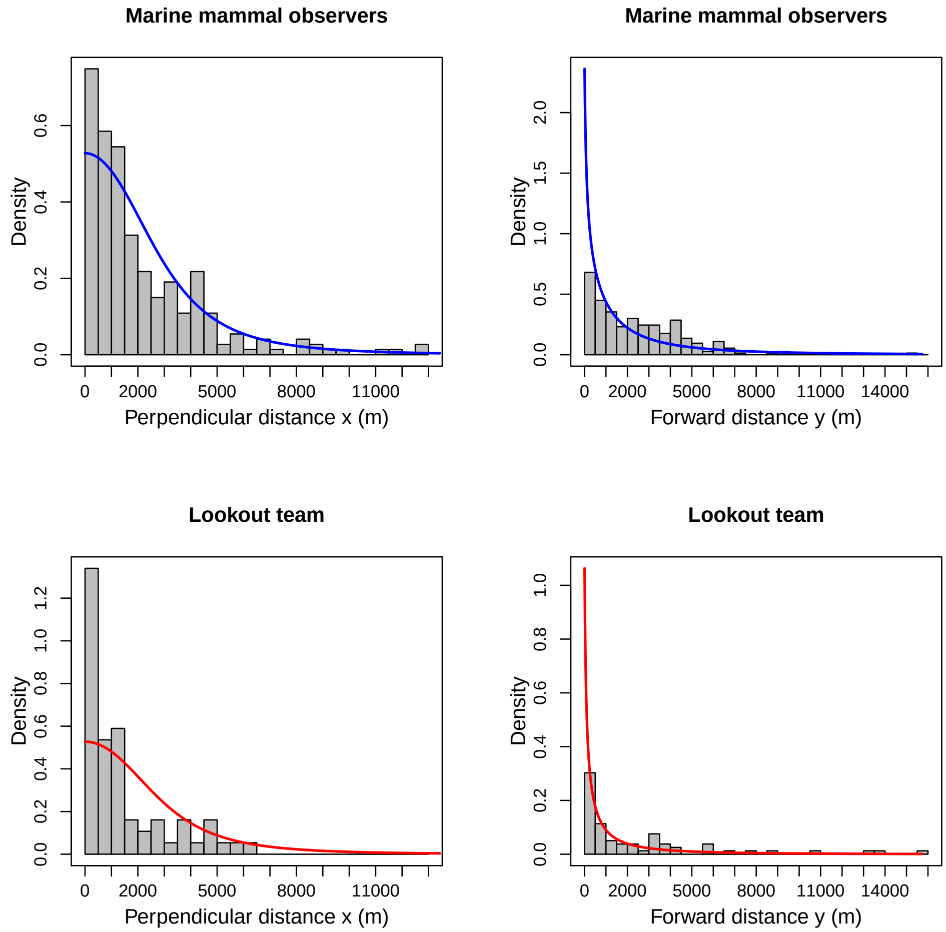


Figure 11: Expected number (lines) versus observed number (histograms) for distance bins along the perpendicular axis (left column) and forward axis (right column) for the rorquals group.

Results from the χ^2-tests indicated that the fit of the rorquals model was poor for MMOs and the LT (Table 12).

Visual inspection of the expected and observed proportions of detections within the 1,500yds radial distance bins confirmed the χ^2-test results. The proportions of detections for the MMOs were irregular, not following the shape of the pdf (Figure 12). These types of irregularities were not as pronounced for the LT, hence, resulting in a slightly better fit of the model (Table 12).

Table 12: Results of the χ^2-tests for the rorquals group including the test statistic (χ^2), degrees of freedom (DF), p-value and number of bins used for the tests. Binning was done using the following upper bin-limits: for MMOs: 340, 690, 1,060, 1,450, 1,870, 2,330, 2,860, 3,490, 4,280, 5,360, 7,030, 10,270, 16,000m; for the LT: 200, 410, 660, 970, 1,370, 1,920, 2,710, 3,920, 5,940, 9,740, 16,000m.

Observer	χ^2	DF	p-value	Bins
MMOs	134.11	10	0.000	13
LT	20.12	8	0.010	11

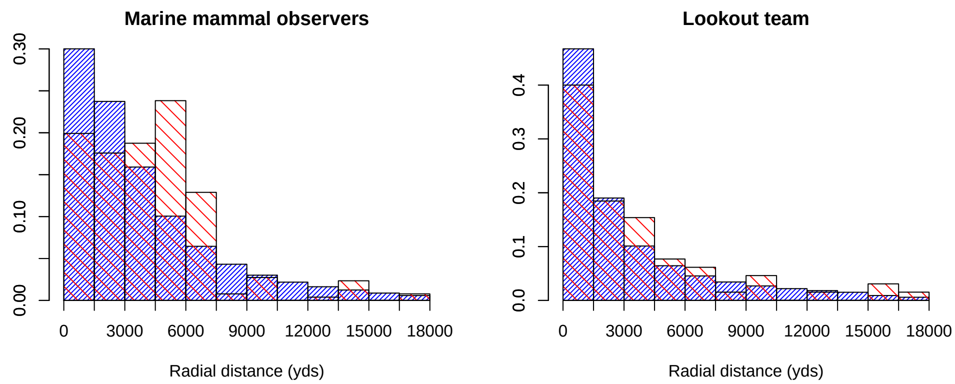


Figure 12: Observed (red) and expected proportions (blue) of detections in each of 12 radial distance bins from the ship for the rorquals group.

4.3.2 Sperm whales

We assumed the same detection hazard for sperm whales as for the rorquals group which is shown in Figure 9.

Figure 13, on the other hand, illustrates that the estimated cumulative distribution of distance y to first detection was much lower for sperm whales compared to the rorquals group (Figure 10) at any given perpendicular distance x or forward distance y. This was due to the much longer times spent in state 2 for this species. Comparing the two platforms, MMOs cumulative distribution values were overall higher – and only reached 0.1 near 2,500m forward distance – compared to those of the LT – which reached 0.1 near 500m forward distance.

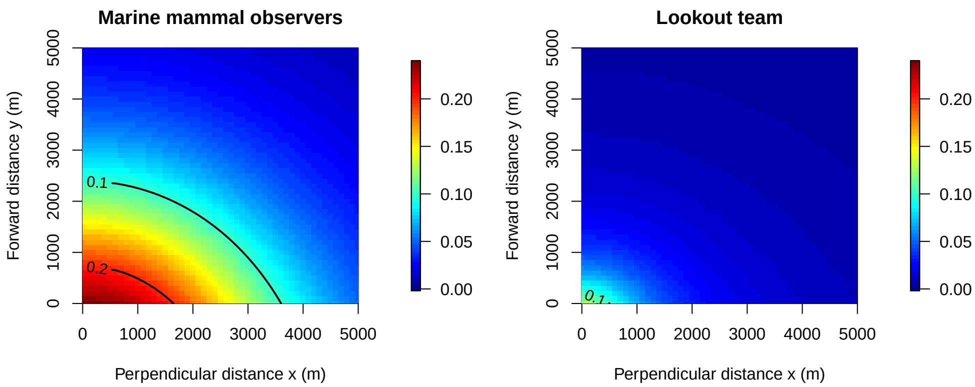


Figure 13: Estimated cumulative distribution of distance y to first detection for sperm whales calculated using the detection hazard parameters from the model fitted to the data of the rorqual group assuming one-way independence and using first detections, and using the availability model parameters for sperm whales (see text), shown for each platform. Black lines indicate contours of the cumulative distribution.

As we used the detection model fitted to the rorquals group data for the sperm whales, we speak of predicted values here instead of estimated.

Similarly to the rorquals group, the predicted probability of sperm whales entering each of the three radii of interest undetected was higher for the LT compared to MMOs (Table 13). For 200yds this probability was predicted at 0.892 (95%CI: 0.867–0.921) for the LT and 0.769 (95%CI: 0.741–0.802) for MMOs. For 500yds this probability was predicted at 0.918 (95%CI: 0.894–0.939) for the LT and 0.780 (95%CI: 0.752–0.814) for MMOs. For 1,000yds this probability was predicted at 0.949 (95%CI: 0.927–0.963) for the LT and 0.803 (95%CI: 0.774–0.836) for MMOs.

Table 13: Predicted effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including expected values (Best) and 95% confidence intervals (CIs) for sperm whales.

	Both platforms combined		MMOs		LT	
	Best	CI	Best	CI	Best	CI
ESHW	980	(827 – 1174)	886	(733 – 1060)	234	(178 – 318)
$g0$	0.259	(0.233 – 0.281)	0.238	(0.204 – 0.266)	0.124	(0.091 – 0.154)
PrU_{200yds}	0.750	(0.727 – 0.777)	0.769	(0.741 – 0.802)	0.892	(0.867 – 0.921)
PrU_{500yds}	0.764	(0.739 – 0.792)	0.780	(0.752 – 0.814)	0.918	(0.894 – 0.939)
$PrU_{1,000yds}$	0.789	(0.762 – 0.819)	0.803	(0.774 – 0.836)	0.949	(0.927 – 0.963)

The sensitivity analysis for sperm whales revealed pronounced differences in results when using different availablity models (Table 14). Decreasing the time spent in state 1 and increasing the time spent in state 2 decreased the $ESHW$ and $g0$ predictions and increased all PrU predictions (analysis 2). Increasing the time spent in state 1 and decreasing the time spent in state 2, on the other hand, had the reverse effect, i.e. increased $ESHW$ and $g0$ predicitons and decreased PrU predictions (analysis 3).

Table 14: Predicted effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) from the sensitivity analysis for sperm whales.

	Analysis 1			Analysis 2			Analysis 3		
	both	MMOs	LT	both	MMOs	LT	both	MMOs	LT
$ESHW$	980	886	234	529	473	114	1819	1657	484
$g0$	0.259	0.238	0.124	0.156	0.139	0.064	0.429	0.406	0.241
PrU_{200yds}	0.750	0.769	0.892	0.851	0.866	0.946	0.581	0.603	0.788
PrU_{500yds}	0.764	0.780	0.918	0.862	0.875	0.959	0.597	0.617	0.834
$PrU_{1,000yds}$	0.789	0.803	0.949	0.881	0.891	0.976	0.628	0.647	0.894

No assessment of goodness of fit could be made for sperm whales as only two sightings were made during the surveys.

4.3.3 Small cetaceans in small pods

Estimated detection hazard for the SCSP group was near zero for most x and y for both platforms with the exception of a very small radius around $x = 0, y = 0$ (Figure 14). At $x = 0, y = 0$, the hazard was higher for MMOs compared to the LT.

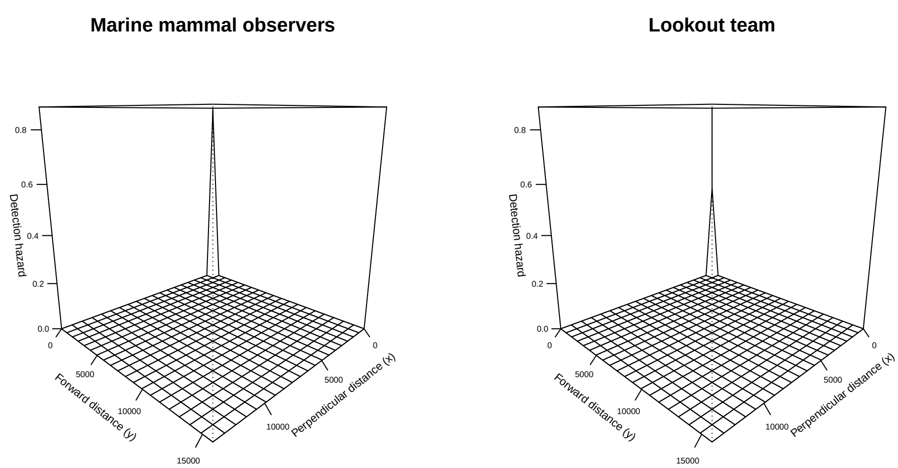


Figure 14: Estimated detection hazard shown with respect to perpendicular distances and forward distances from the ship for each platform from the model fitted to the detection data of the SCSP group assuming one-way independence and using first detections.

Similarly, the estimated cumulative distribution of distance y to first detection was near zero for most x and y with the exception of a small radius around $x = 0, y = 0$. Again, these densities at $x = 0, y = 0$ were higher for MMOs compared to LT.

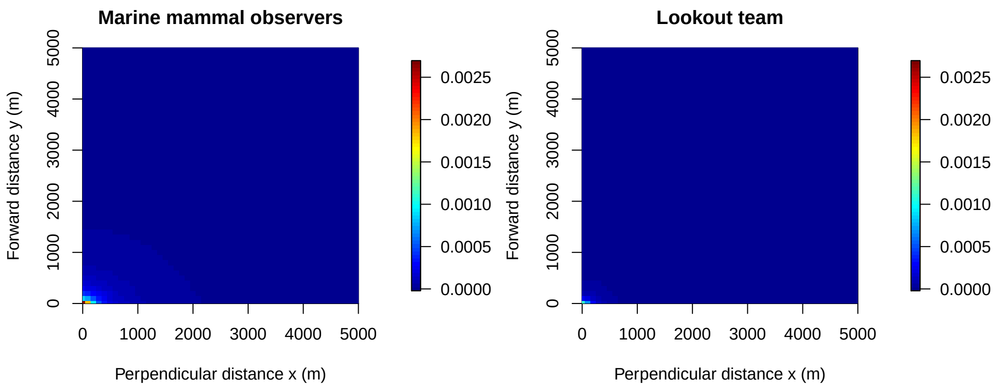


Figure 15: Estimated cumulative distribution of distance y to first detection from the model fitted to the detection data of the SCSP group assuming one-way independence and using first detections, shown for each platform.

Overall, estimates of the probability of entering any of the radii around the ship undetected was very high for the SCSP group (Table 15). If rounded to the second decimal point, estimates of this probability would be 1.00 for the LT, MMOs and both teams combined for each of the three distances considered (200yds, 500yds and 1,000yds).

Table 15: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates and 95% confidence intervals (CIs) for the group of small cetaceans in small pods.

	Both platforms combined		MMOs		LT	
	Best	CI	Best	CI	Best	CI
$ESHW$	0.853	(0.341 – 2.118)	0.661	(0.257 – 1.633)	0.192	(0.069 – 0.481)
$g0$	0.0038	(0.0016 – 0.0088)	0.0027	(0.0011 – 0.0064)	0.0011	(0.0004 – 0.0027)
PrU_{200yds}	0.9993	(0.9983 – 0.9997)	0.9995	(0.9987 – 0.9998)	0.9999	(0.9996 – 1.0000)
PrU_{500yds}	0.9998	(0.9995 – 0.9999)	0.9999	(0.9996 – 0.9999)	1.0000	(0.9999 – 1.0000)
$PrU_{1,000yds}$	0.9999	(0.9998 – 1.0000)	1.0000	(0.9999 – 1.0000)	1.0000	(1.0000 – 1.0000)

The sensitivity analysis revealed that changing the expected time in state 2, decreased the estimated $ESHW$ and $g0$ and increased all estimates of PrU (compare $ESHW$ and $g0$ from analyses 1 and 2, Table 16). Changing both the expected time in state 1 and 2, while maintaing a similar ratio between the two, has negligible influence on the estimates (compare $ESHW$ and $g0$ from analyses 1 and 3). Changing the model more drastically in analysis 4 had a larger effect on the estimates of $ESHW$ and $g0$. All four analysis revealed similar estimates for any of the PrUs.

Table 16: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates of sensitivity analysis for the SCSP group.

	Analysis 1			Analysis 2		
	both	MMOs	LT	both	MMOs	LT
$ESHW$	0.853	0.661	0.192	0.364	0.282	0.082
$g0$	0.00379	0.00267	0.00112	0.00162	0.00114	0.00048
PrU_{200yds}	0.99932	0.99947	0.99986	0.99971	0.99977	0.99994
PrU_{500yds}	0.99982	0.99985	0.99997	0.99992	0.99993	0.99999
$PrU_{1,000yds}$	0.99994	0.99995	0.99999	0.99998	0.99998	1
	Analysis 3			Analysis 4		
	both	MMOs	LT	both	MMOs	LT
$ESHW$	0.827	0.641	0.186	6.369	4.955	1.445
$g0$	0.00368	0.00259	0.00109	0.02828	0.0201	0.00849
PrU_{200yds}	0.99934	0.99948	0.99986	0.99492	0.99597	0.99894
PrU_{500yds}	0.99983	0.99985	0.99997	0.99866	0.99886	0.9998
$PrU_{1,000yds}$	0.99994	0.99995	0.99999	0.99957	0.99961	0.99995

The steep rise in the pdf near distance zero shown in Figure 16 for both platforms and both the perpendicular and forward distances, is in accordance with the pattern seen in Figures 14 and 15. This was caused by the large number of detections at distances <20m (see also Figure 22 below).

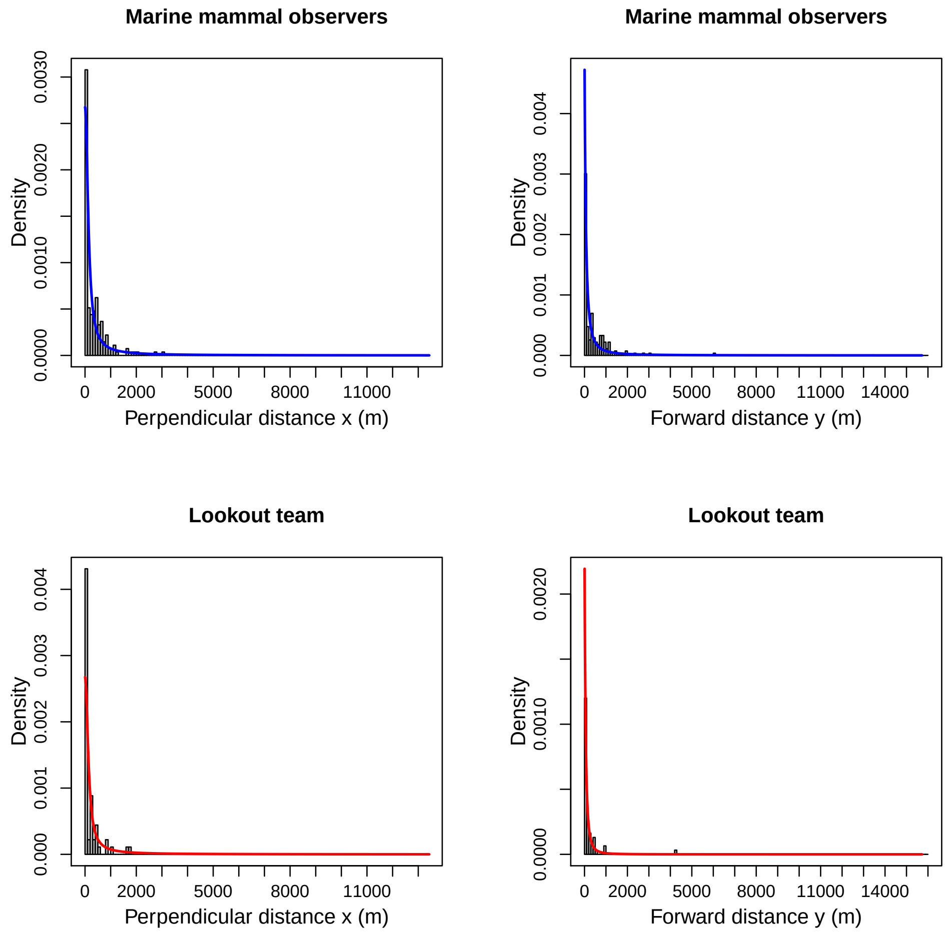


Figure 16: Expected number (lines) versus observed number (histograms) for distance bins along the perpendicular axis (left column) and forward axis (right column) for the SCSP group.

Results of the χ^2-test were significant (Table 17) indicating a poor fit of the model. However, the visual depiction of the expected and observed proportions of detections in radial distance bins indicated a good fit Figure 17.

Table 17: Results of the χ^2-tests for the SCSS group including the test statistic (χ^2), degrees of freedom (DF), p-value and number of bins used for the tests. Binning was done using the following upper bin-limits: for MMOs: 40, 80, 120, 160, 210, 280, 380, 540, 820, 1,410, 3,100, 16,000m; for the LT: 30, 60, 90, 120, 160, 210, 280, 400, 680, 2,760, 16,000m.

Observer	χ^2	DF	p-value	Bins
MMOs	191.68	9	0.000	12
LT	62.91	8	0.000	11

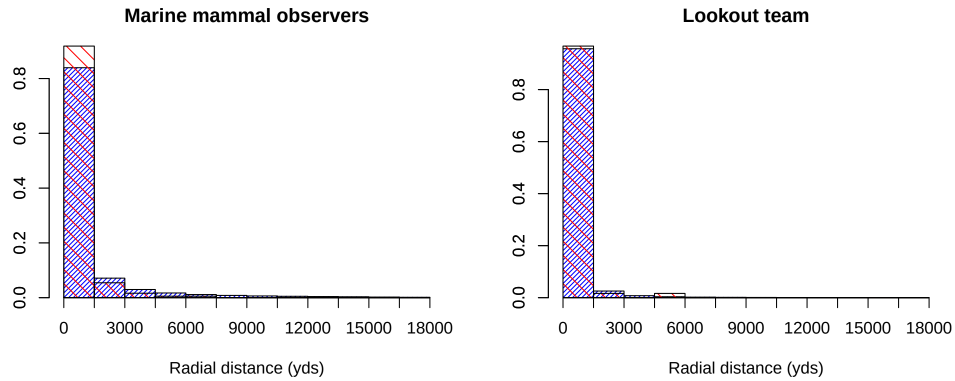


Figure 17: Observed (red) and expected proportions (blue) of detections in each of 12 radial distance bins from the ship for the SCSP group.

4.3.4 Small cetaceans in large pods

Estimated detection hazard for the SCSP group was near zero for most x and y for both platforms with the exception of a very small radius around $x = 0, y = 0$ (Figure 18). At $x = 0, y = 0$, the hazard was higher for LT compared to the MMOs.

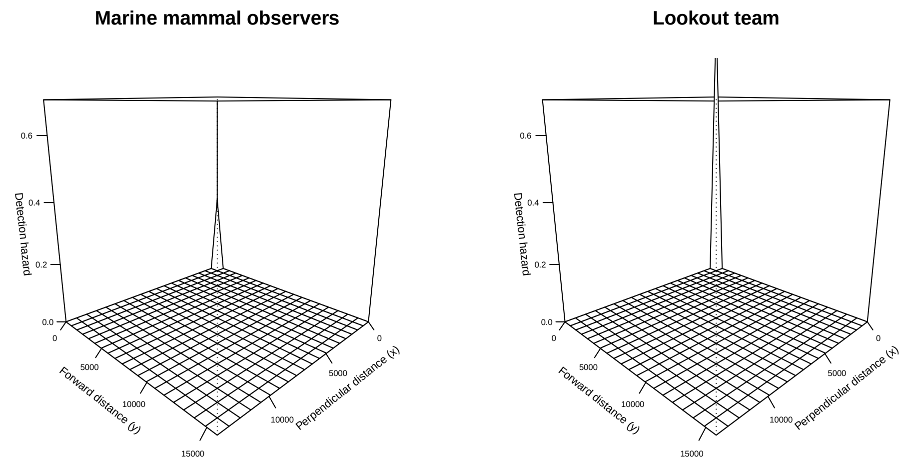


Figure 18: Estimated detection hazard shown with respect to perpendicular distances and forward distances from the ship for each platform from the model fitted to the detection data of the SCLP group assuming one-way independence and using first detections.

Similar to the SCSP group, the estimated cumulative distribution of distance y to first detection was near zero for most x and y with the exception of a small radius around $x = 0, y = 0$. Again, these densities at $x = 0, y = 0$ were higher for MMOs compared to LT.

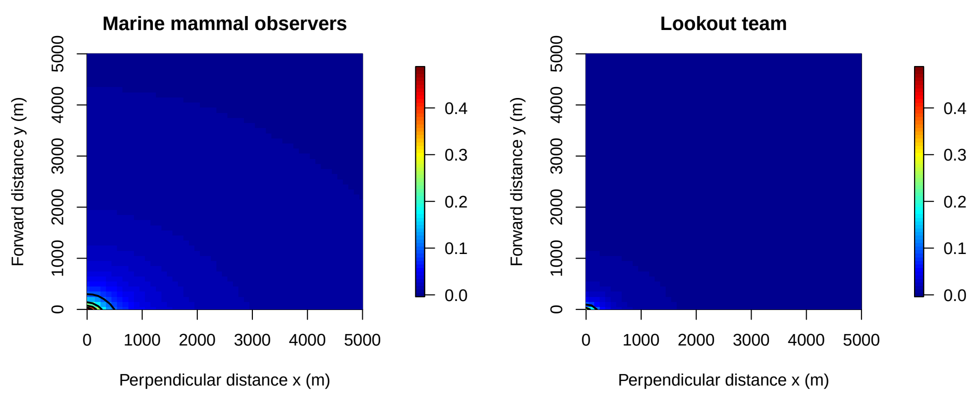


Figure 19: Estimated cumulative distribution of distance y to first detection from the model fitted to the detection data of the SCLP group assuming one-way independence and using first detections, shown for each platform.

Overall, estimates of the probability of entering any of the radii around the ship undetected was high for large pods of small cetaceans (Table 18). For 200yds this probability was estimated at 0.942 (95%CI: 0.910–1.000) for the LT and 0.826 (95%CI: 0.736–0.899) for MMOs. For 500yds this probability was estimated at 0.983 (95%CI: 0.971–1.000) for the LT and 0.928 (95%CI: 0.890–0.961) for MMOs. For 1,000yds this probability was estimated at 0.994 (95%CI: 0.989–1.000) for the LT and 0.968 (95%CI: 0.948–0.983) for MMOs.

Table 18: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates and 95% confidence intervals (CIs) for the group of small cetaceans in large pods.

	Both platforms combined		MMOs		LT	
	Best	CI	Best	CI	Best	CI
$ESHW$	292	(147 – 408)	240	(137 – 363)	70	(0 – 109)
$g0$	0.618	(0.334 – 0.755)	0.488	(0.308 – 0.676)	0.253	(0.000 – 0.371)
PrU_{200yds}	0.778	(0.691 – 0.892)	0.826	(0.736 – 0.899)	0.942	(0.910 – 1.000)
PrU_{500yds}	0.912	(0.873 – 0.958)	0.928	(0.890 – 0.961)	0.983	(0.971 – 1.000)
$PrU_{1,000yds}$	0.962	(0.943 – 0.982)	0.968	(0.948 – 0.983)	0.994	(0.989 – 1.000)

The sensitivity analysis gave nearly identical best estimates for both analyses.

Table 19: Estimated effective strip half-width ($ESHW$), trackline detection probability ($g0$) and probability of entering a radius undetected (PrU) including best estimates of sensitivity analysis for the SCLP group.

	Analysis 1			Analysis 2		
	both	MMOs	LT	both	MMOs	LT
$ESHW$	292	240	70	290	239	69
$g0$	0.618	0.488	0.253	0.613	0.485	0.251
PrU_{200yds}	0.778	0.826	0.942	0.800	0.827	0.942
PrU_{500yds}	0.912	0.928	0.983	0.913	0.929	0.983
$PrU_{1,000yds}$	0.962	0.968	0.994	0.962	0.968	0.994

The models for the SCLP group in Figure 20 revealed patterns similar to those for the SCSP group in Figure 16 with a steep rise in expected values near distance zero. These were likely caused again by the high number of detections within 20m radial distance (see also Figure 23 below).

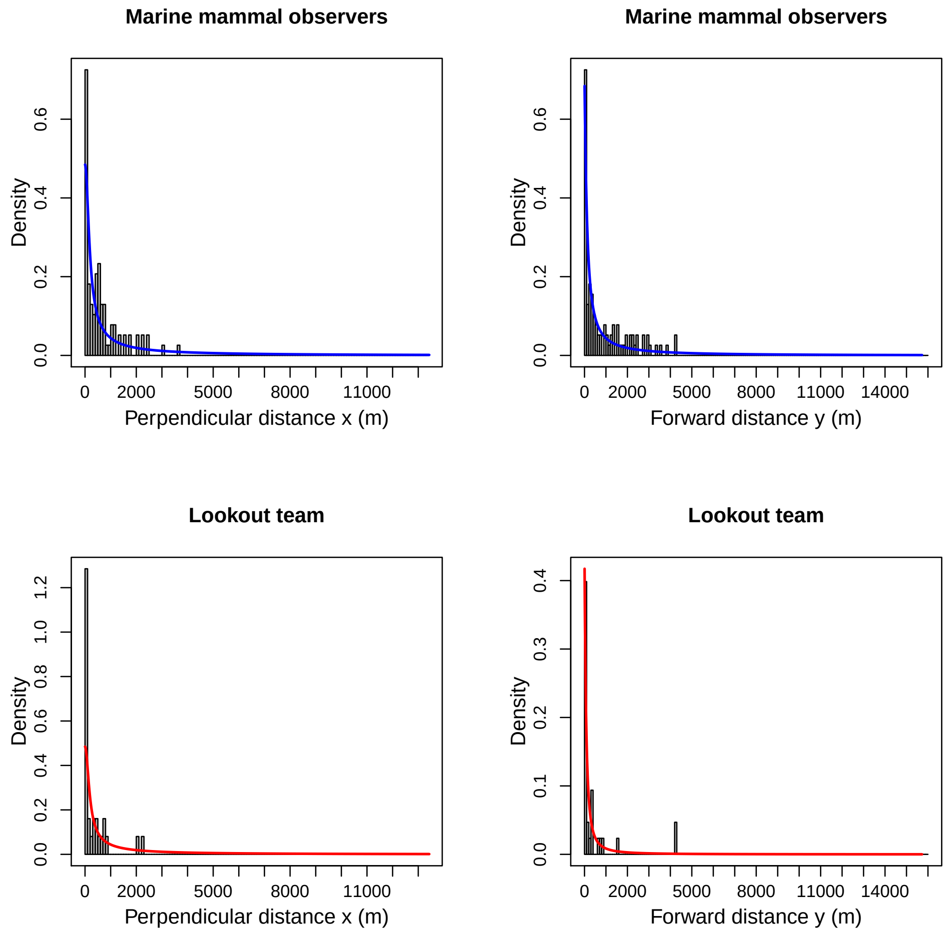


Figure 20: Expected number (lines) versus observed number (histograms) for distance bins along the perpendicular axis (left column) and forward axis (right column) for the SCLP group.

Results of the χ^2-test were significant indicating a poor fit of the model for both the MMOs and the LT (Table 20). Visual inspection of the expected and observed proportions of detections showed that the model for the MMOs underfitted for most distance bins <6,000yds and generally overfitted at distances >6,000yds (Figure 21). However, the model captured the steep spike at the smallest distances well for both the MMOs and the LT.

Table 20: Results of the χ^2-tests for the SCLS group including the test statistic (χ^2), degrees of freedom (DF), p-value and number of bins used for the tests. Binning was done using the following upper bin-limits: for MMOs: 100, 190, 310, 490, 780, 1,250, 2,070, 3,570, 6,510, 16,000m; for the LT: 80, 170, 310, 620, 1,710, 16,000m.

Observer	χ^2	DF	p-value	Bins
MMOs	46.94	7	0.000	10
LT	18.21	3	0.000	6

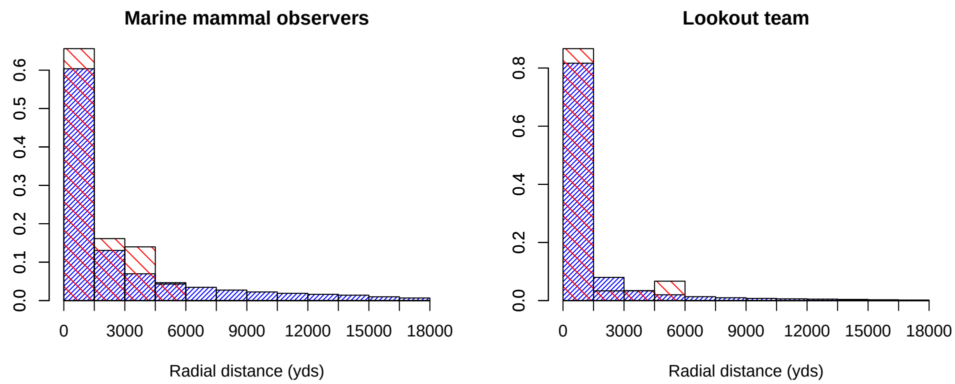


Figure 21: Observed (red) and expected proportions (blue) of detections in each of 12 radial distance bins from the ship for the SCLP group.

5 Discussion

Out of all groups investigated, the rorquals group had the lowest probabilities for entering any of the radii of interest undetected. However, even for this group, these probabilities were high for the LT: 80% for the closest of ranges at 200yds and 91% at 1,000yds. These probabilities estimated for the rorquals group were generally lower for MMOs, ranging between 49% at 200yds and 59% at 1,000yds.

Due to the long dive times of sperm whales, this species had much higher probabilities of entering any of the mitigation ranges undetected by either of the platforms. Although these probabilities were generally lower for MMOs, they were 77% at 200yds and 80% at 1,000yds for this platform. The equivalent estimates for the LT were 89% and 95% for 200 and 1,000yds, respectively.

Extremely high estimates of the probability of entering the mitigation radii of interest for the SCSP group were likely due to most of these groups being first detected after they have already approached the ship – although those pods recorded as bowriding were excluded from analyses. This tendency to approach together with the difficulty in detecting them for either platform makes this group very vulnerable to exposure. A zoom-in of Figure (7) shows that the highest proportion of pods were within 10m radial distance (Figure 22). While we assume that these probabilities would be high, we have some reasons to doubt these results. Firstly, some rounding of distances may have occurred (see spikes at 0, 10, 20, 30, 40, 50, 60 and 100m radial distance in Figure 22). Most importantly, however, when animals approach the ship, the underlying assumption of uniform distribution of animals with respect to perpendicular distance x is violated (see equation (17) and text following this equation). Furthermore, we had assumed that animals do not move in the horizontal plane, i.e. in the x and y dimensions. This assumption was clearly violated, in particular for those pods approaching the ship, likely at high speeds.

Similar issues existed for the SCLP group. Radial distances also seemed to be rounded to prefered values (Figure 23) and most pods were detected near zero forward or zero perpendicular distances.

Although our simulation study provided evidence that our methods provide unbiased estimates, it is important to emphasize that they rely on various assumptions, and that violation of these assumptions (such as from animal movement) can be expected to lead to biased results. Future research could potentially evaluate the extent of any bias (via a simulation study) and extend the methods further to relax assumptions thought to be incorrect for these data. Another potential avenue would be to explore plausible values of PrU given assumed detection hazard functions, as well as assumed diving parameters.

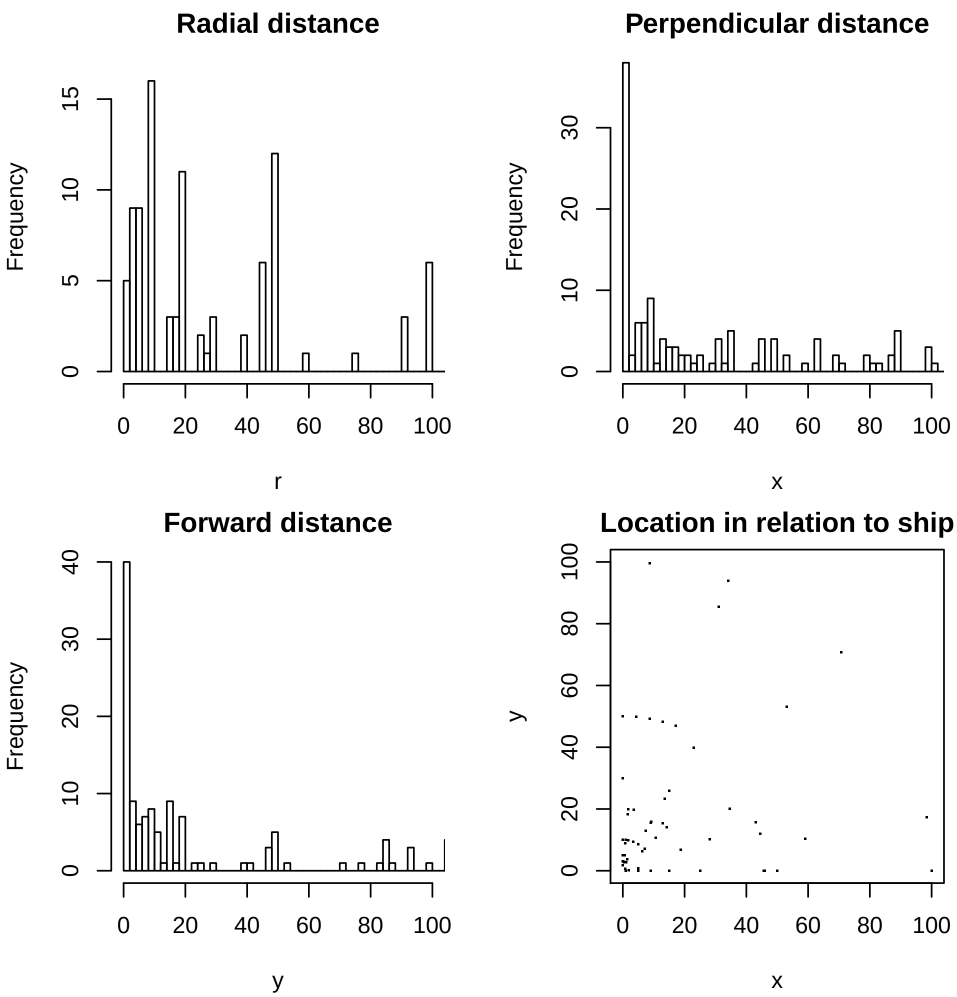


Figure 22: Histograms of radial, perpendicular and forward distances (all in meters) of first detections for the SCSP group as well as their location in relation to the ship within 100m. Detections made on the left side of the ship were folded to the right side.

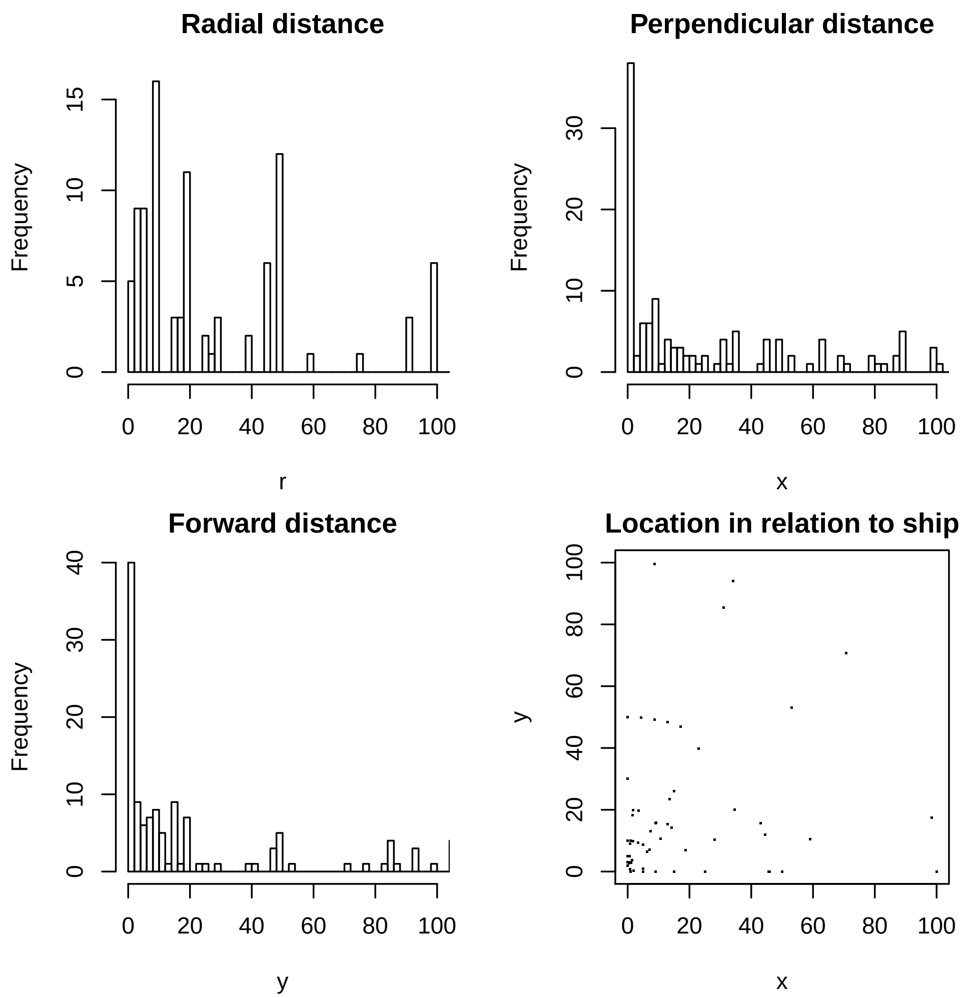


Figure 23: Histograms of radial, perpendicular and forward distances (all in meters) of first detections for the SCLP group as well as their location in relation to the ship within 100m. Detections made on the left side of the ship were folded to the right side.

References

Borchers, D.L., and R. Langrock. 2015. "Double-Observer Line Transect Surveys with Markov-Modulated Poisson Process Models for Animal Availability." *Biometrics* 71 (4): 1060–9.

Buckland, S.T., E.A. Rexstad, T.A. Marques, and C.S. Oedekoven. 2015. *Distance Sampling: Methods and Applications*. Methods in Statistical Ecology. Springer International Publishing.

Burt, Mary Louise, David L. Borchers, Kurt J. Jenkins, and Tiago A. Marques. 2014. "Using Mark–recapture Distance Sampling Methods on Line Transect Surveys." *Methods in Ecology and Evolution* 5 (11): 1180–91.

Dolphin, William Ford. 1987. "Ventilation and Dive Patterns of Humpback Whales, Megaptera Novaeangliae, on Their Alaskan Feeding Grounds." *Canadian Journal of Zoology* 65 (1): 83–90.

Drouot, Violaine, Alexandre Gannier, and John Goold. 2004. "Diving and Feeding Behaviour of Sperm Whales (Physeter Macrocephalus) in the Northwestern Mediterranean Sea." *Aquatic Mammals* 30: 419–26.

Genz, Alan, and Frank Bretz. 2009. *Computation of Multivariate Normal and T Probabilities*. Lecture Notes in Statistics. Heidelberg: Springer-Verlag.

Genz, Alan, Frank Bretz, Tetsuhisa Miwa, Xuefei Mi, Friedrich Leisch, Fabian Scheipl, and Torsten Hothorn. 2019. *mvtnorm: Multivariate Normal and T Distributions*. https://CRAN.R-project.org/package=mvtnorm.

Goulet, Vincent, Christophe Dutang, Martin Maechler, David Firth, Marina Shapira, and Michael Stadelmann. 2021. *Expm: Matrix Exponential, Log, 'Etc'*. https://CRAN.R-project.org/package=expm.

Langrock, R., D.L. Borchers, and H.J. Skaug. 2013. "Markov-Modulated Nonhomogeneous Poisson Processes for Modeling Detections in Surveys of Marine Mammal Abundance." *Journal of the American Statistical Association* 108 (503): 840–51.

Palka, D.L., S. Chavez-Rosales, E. Josephson, H.L Cholewiak D. and. Haas, L. Garrison, L Jones, D. Sigourney, et al. 2017. "Atlantic Marine Assessment Program for Protected Species: 2010-2014." US Dept. of the Interior, Bureau of Ocean Energy Management, Atlantic OCS Region, Washington, DC. OCS Study BOEM 2017-071. 211 pp.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Scott, Michael, and Susan Chivers. 2009. "Movements and Diving Behavior of Pelagic Spotted Dolphins." Publications, Agencies; Staff of the U.S. Department of Commerce. 46.