Combining passive acoustic data from a towed hydrophone array with visual line transect data to estimate abundance and availability bias of sperm whales (*Physeter macrocephalus*)

Douglas B. Sigourney[1], Annamaria DeAngelis[2], Danielle Cholewiak[2] and Debra Palka[2]

[1] Integrated Statistics, Woods Hole, Massachusetts, United States
[2] NOAA Northeast Fisheries Science Center, Woods Hole, Massachusetts, United States

ABSTRACT

Visual line transect (VLT) surveys are central to the monitoring and study of marine mammals. However, for cryptic species such as deep diving cetaceans VLT surveys alone suffer from problems of low sample sizes and availability bias where animals below the surface are not available to be detected. The advent of passive acoustic monitoring (PAM) technology offers important opportunities to observe deep diving cetaceans but statistical challenges remain particularly when trying to integrate VLT and PAM data. Herein, we present a general framework to combine these data streams to estimate abundance when both surveys are conducted simultaneously. Secondarily, our approach can also be used to derive an estimate of availability bias. We outline three methods that vary in complexity and data requirements which are (1) a simple distance sampling (DS) method that treats the two datasets independently (*DS-DS Method*), (2) a fully integrated approach that applies a capture-mark recapture (CMR) analysis to the PAM data (*CMR-DS Method*) and (3) a hybrid approach that requires only a subset of the PAM CMR data (*Hybrid Method*). To evaluate their performance, we use simulations based on known diving and vocalizing behavior of sperm whales (*Physeter macrocephalus*). As a case study, we applied the *Hybrid Method* to data from a shipboard survey of sperm whales and compared estimates to a VLT only analysis. Simulation results demonstrated that the *CMR-DS Method* and *Hybrid Method* reduced bias by >90% for both abundance and availability bias in comparison to the simpler *DS -DS Method*. Overall, the *CMR-DS Method* was the least biased and most precise. For the case study, our application of the *Hybrid Method* to the sperm whale dataset produced estimates of abundance and availability bias that were comparable to estimates from the VLT only analysis but with considerably higher precision. Integrating multiple sources of data is an important goal with clear benefits. As a step towards that goal we have developed a novel framework. Results from this study are promising although challenges still remain. Future work may focus on applying this method to other deep-diving species and comparing the proposed method to other statistical approaches that aim to combine information from multiple data sources.

INTRODUCTION

Worldwide, cetaceans face a growing number of threats presenting challenges for management and conservation (*Avila, Kaschner & Dormann, 2018*). Effective management relies on accurate and precise estimates of abundance (*Hammond et al., 2021*). For most large whale species visual line transect (VLT) surveys are a common method to collect information on density and distribution. Generally, to estimate abundance from VLT surveys distance sampling (DS) techniques are applied (*Buckland et al., 2001*). However, for many species, particularly deep diving cetaceans, VLT surveys alone may not be adequate. Low sample sizes may limit the ability to model ecological relationships (*Barkley et al., 2022*) and to detect trends in abundance (*Taylor et al., 2007*). In addition, failure to account for the proportion of animals that are not at the surface, and therefore, not available to be detected can result in considerable bias (*Laake et al., 1997*, *Borchers et al., 2013*). This bias, commonly referred to as availability bias (*McLaren, 1961*; *Marsh & Sinclair, 1989*), is particularly problematic in deep diving cetaceans and often requires auxiliary data on diving behavior to properly adjust abundance estimates. For these reasons, there is a pressing need to combine VLT surveys with other sources of data, when possible, to achieve accurate and precise estimates of abundance and distribution.

The advent of remote passive acoustic technology to detect and track diving animals offers an opportunity to improve estimates of abundance and estimate availability bias. Passive acoustic monitoring (PAM) can be conducted *via* a variety of platforms, from stationary recorders deployed at fixed sites on the seafloor, to mobile ones, such as towed hydrophone arrays or gliders. For estimating abundance, towed arrays have proven useful because they can be directly incorporated into a line transect sampling design (*Barlow & Taylor, 2005*; *Gerrodette et al., 2011*). However, PAM surveys are also prone to availability bias as animals that are not vocalizing are not available to the PAM survey (*Barlow & Taylor, 2005*). Studies that utilize both platforms simultaneously by, for example, deploying a towed hydrophone array during VLT surveys have the advantage of potentially minimizing or eliminating this bias as animals are likely be available to at least one platform. Importantly, dual survey designs also have the potential to increase sample size and offer an opportunity to estimate availability bias *in situ*.

Combing datasets has many potential benefits such as decreasing bias and increasing precision of demographic parameters (*Pacifici et al., 2017*; *Zipkin & Saunders, 2018*; *Miller et al., 2019*; *Conn et al., 2022*). Although the use of PAM technology is steadily increasing in cetacean research (*Marques et al., 2013*; *Gibb et al., 2019*) and several studies have developed methods for analyzing PAM data (*Whitehead, 2009*; *Barlow et al., 2021*; *Westell et al., 2022*), there have been fewer efforts to develop statistical methods for combining PAM data with other data streams such as VLT data. There have been some attempts to compare results from models built with PAM data to models built with VLT data (*Barlow*

& Taylor, 2005; Williamson et al., 2016). In addition, some studies have combined VLT and PAM data with data using existing statistical methods. For example, Gerrodette et al. (2011) used DS methods to combine an estimate of abundance from a VLT survey in deep water habitat with a PAM estimate of abundance from shallow water habitat to get an overall estimate of abundance of vaquitas (Phocoena sinus). Thompson, Brookes & Cordes (2015) used a complementary approach whereby the probability of presence of dolphins was estimated from PAM data and subsequently classified into species group using visual data. Fleming et al. (2018) and Barkley et al. (2022) developed specific protocols for sampling and post-processing data from dual VLT and PAM surveys such the data could be combined into a single species distribution model analysis using standard statistical methods. Despite these novel uses of PAM data, we are not aware of any formal attempts to develop a statistical framework to integrate VLT and PAM data into one estimate of abundance.

Along with increasing precision and accuracy of abundance estimates integrating VLT with PAM data offers an opportunity to estimate and adjust for availability bias *in situ*. Because of the possibility of large negative bias in abundance estimates, particularly for deep diving cetaceans, there have been numerous attempts to develop methods to account for the diving behavior of animals (Laake et al., 1997; Okamura, 2003; Okamura et al., 2012; Borchers et al., 2013). One common method is to apply a correction factor based on average percent time at the surface (Forcada et al., 2004; Nykänen et al., 2018; Sucunza et al., 2018). However, this method can still result in bias (Borchers et al., 2013). A more sophisticated approach is to integrate information on diving behavior either non-parametrically (Okamura et al., 2012) or through a hidden Markov model approach, but these methods generally require data on the surfacing and diving behavior usually from tagged animals (Borchers et al., 2013; Langrock, Borchers & Skaug, 2013). To date, we are not aware of any studies that attempt to incorporate towed array data into estimates of availability bias. Towed arrays have the advantage of being an independent survey conducted simultaneously with the VLT survey. In addition, unlike tag data where sample sizes are generally low, PAM towed arrays offer the potential for larger samples sizes that could increase precision (*e.g.,* Westell et al., 2022). If the data are sufficiently robust there is also the potential to investigate temporal and spatial variability in availability bias as there may be spatial variability in diving behavior (Watwood et al., 2006).

Although there are many clear benefits to combining VLT and PAM data there are also many challenges. When both surveys are conducted simultaneously one problem is the issue of duplicates. Identifying a sighting as a duplicate is standard protocol for VLT surveys that utilize more than one survey team with the goal of applying mark-recapture distance sampling (MRDS) techniques (Laake & Borchers, 2004; Burt et al., 2014). Although such designs are not uncommon, error associated with duplicate identification is problematic and can lead to bias (Hamilton et al., 2018; Stevenson, Borchers & Fewster, 2019). For surveys that use unmanned aerial platforms it may require novel statistical techniques to objectively define which sightings are duplicates (Stevenson, Borchers & Fewster, 2019; Stevenson, Fewster & Sharma, 2022). With a PAM towed array, this issue has the potential to be particularly challenging because the platforms are surveying animals in

different dive states with an unknown probability of animals transitioning among states. As such, a duplicate detection not only depends on the location and timing of the detection but whether an animal transitions among dive states. With regard to location, there is an additional problem of left-right ambiguity with a linear towed array containing two hydrophones as it is difficult to assign a localized detection to one side of the trackline (*Barlow & Taylor, 2005*; *Barkley et al., 2022*). Finally, communication among platforms is limited as standard line transect theory requires survey teams to be blind to each to avoid bias (*Laake & Borchers, 2004*; *Burt et al., 2014*). These complications make a simple assignment of duplicates challenging just based on the timing and radial distance of a detection.

Sperm whales (*Physeter macrocephalus*) are an acoustically active species that lend themselves to survey methodologies that use data collected visually or *via* PAM. Sperm whales have been studied extensively since the 1950s (*Worthington & Schevill, 1957*; *Watkins & Schevill, 1977*). When they undergo deep foraging dives, they use echolocation to search for and hone in on their prey (*Miller, Johnson & Tyack, 2004*). It is the reliability in the pattern of their echolocation clicks and the need to forage regardless of gender or age (barring nursing calves <1 yr old, *Tønnesen et al., 2018*) that makes sperm whale foraging clicks easy to track while at depth. Foraging clicks (also referred to as 'usual clicks' in the literature) have a high source level (236 db; *Møhl et al., 2003*), can propagate between 5 and 16 km in certain environmental conditions, and have a total vocal active phase of ~35 min, which makes it possible to track diving whales with a hydrophone array towed at the surface. Additionally, the detection range in deep water can exceed that of the known maximum sperm whales' diving depth of 1,330 m (*Madsen, Wahlberg & Møhl, 2002*; *Mellinger, Thode & Martinez, 2002*; *Møhl et al., 2003*), therefore a sperm whale's entire search phase is available to be detected (*Westell et al., 2022*).

In this article we present a statistical method for integrating PAM towed hydrophone array data with VLT data when both surveys are conducted simultaneously. We begin by providing a brief summary of the challenges involved with integrating these two sources of data (see Appendix S1 for a comprehensive list). Next, we describe the statistical framework. Our method combines a capture mark-recapture (CMR) approach with a DS analysis to estimate abundance and surface availability while accounting for transition probabilities among surfacing and diving states. Because processing PAM data can present challenges, we include an alternative approach that requires only a subset of the PAM data be processed into capture histories. We test both methods on simulated data based on diving behavior of sperm whales and compare results to a simpler method that ignores transition probabilities. Finally, we provide a case study from a shipboard survey conducted in the Northwest Atlantic.

METHODS

Challenges

There are several challenges to be considered when analyzing PAM data from a towed array data and combining it with VLT data. For example, a submerged whale may not always be available for detection because it could be located in a blind spot where the

acoustic signal is blocked from the PAM array or could be in a behavioral state where it is not vocalizing. In addition, there are technical limitations such as the ability to process all the data and confidently assign a series of clicks to an individual whale. In Appendix S1 we provide a detailed outline of the challenges we considered.

Methods to integrate PAM with VLT data

Line transect data are commonly analyzed with DS techniques to estimate detection probability on the trackline where distances (y) of animals or groups of animals are observed. If there is only one survey team the canonical equation can be written as.

$$\hat{p} = \frac{\int_0^w g(y)dy}{W}, \tag{1}$$

where W is a truncation distance on the trackline and $g(y)$ is a detection function (*e.g.,* half-normal).

One assumption of Eq. (1) is that the detection probability on the trackline ($g(0)$) is equal to one. Two teams can be introduced to deal with assumptions of detectability on the trackline. In this MRDS design one team can serve as a trial for the other and mark-recapture techniques can be employed to estimate $g(0)$ and adjust the $\hat{p}$ such that.

$$\hat{p}^{\cdot} = \hat{g}(0)\frac{\int_0^w g(y)dy}{W}, \tag{2}$$

A further adjustment can be applied to account for availability bias as

$$\hat{p} = \hat{g}(0)\frac{\int_0^w g(y)dy}{W}\hat{a}, \tag{3}$$

where $\hat{a}$ is the surface availability bias correction typically estimated from auxiliary data such as digital recording tags (DTAGs, *sensu* Johnson & Tyack, 2003). We refer to Eq. (3) as MRDS$_{AV}$.

Dual line transect surveys with a VLT platform and PAM platform offer the opportunity to record vocalizing animals below the surface while simultaneously observing animals visually above the surface. and therefore, may not require an auxiliary correction factor as in Eq. (3). This data collection design offers the opportunity to estimate the abundance below the surface (N_B) and abundance above the surface (N_S) to calculate an estimate of total abundance (N_T). However, when the two data types are collected simultaneously there is some unknown percentage of animals that may be available to both survey platforms. Ignoring these potential "duplicates" could result in a positive bias in the estimate of abundance. Therefore, estimating N_T by combining N_B and N_S requires an adjustment for the total number of duplicates (N_D) such that

$$N_T = N_S + N_B - N_D \tag{4}$$

Below we outline a general framework to analyze the acoustic data and describe how the information from the acoustic analysis is integrated with visual line transect data to derive unbiased estimates of abundance and surface availability. We leverage techniques from DS theory (*Buckland et al., 2001*) and capture mark-recapture (CMR) theory (*Royle &*

Dorazio, 2008). We describe (1) a simple method that ignores duplicates (*DS-DS Method*), (2) a method for the case when all acoustic detections can be fully annotated (*CMR-DS Method*) and (3) a method for when only a subset of the acoustic data can be fully annotated (*Hybrid Method*). For the purpose of this modelling exercise, we assume two behavioral states, a surface state that is available to the VLT platform and a foraging state that is available to the PAM platform. We do not specifically model silent states but will address that assumption within our modelling framework.

DS-DS method

As a default method we first consider a simple approach that ignores duplicates. With this method we simply apply DS methods to both the visual data and the acoustic data and sum the two estimates together for a final abundance estimate (*i.e.*, assume $N_D = 0$). This method is analogous to the method of *Gerrodette et al. (2011)* where the difference is we use a Bayesian framework to estimate abundance and precision and we apply it to data collected simultaneously rather than data collected from two separate locations. Either Eq. (1) or Eq. (2) can be used to estimate N_S. In practice, $g(0)$ is assumed to be 1 for acoustic data so Eq. (1) is used to estimate N_B.

CMR-DS method

For this method we assume that all acoustic detections recorded by the passive acoustic towed array from a vocalizing whale can be annotated into a click train and problems with ambiguous click trains (see Appendix S1) are negligible. For each click train, we divide the forward distances of all recorded clicking events into a series of equally spaced distance bins (see Appendix S2). The first bin is the maximum distance that whales are detected by the towed array, similar to the truncation distance in distance sampling (see Fig. 1 for a conceptual diagram).

Because the towed array continuously tracks the position of a clicking whale over a given time period and distance we can treat information collected from the array as CMR data. We adopt a state-space formulation of the Jolly-Seber model using data augmentation (*Royle & Dorazio, 2008*). The unobserved state of whale i in distance bin j ($z_{i,j}$) represents its position above or below the surface, where $z_{i,j} = 1$ indicates that it is below the surface and available to the towed array and $z_{i,j} = 0$ indicates that is at the surface and not available to the acoustic array. The probability that a whale is available in the first distance bin when it first comes into range of the acoustic array can be modeled as the outcome of a Bernoulli trial as

$$z_{i,1} \sim Bernoulli\ (\gamma_1). \tag{5}$$

Whales that are not present below the surface in the first bin ($z_{i,1} = 0$) can enter the diving state in subsequent bins given that they were not below the surface in any previous bin. For all other bins the state process is conditional on the previous state such that a whale that has not yet entered the diving state can enter with a probability γ_j whereas a whale that has already entered the diving sate can remain in that state with probability ϕ. We model this process as
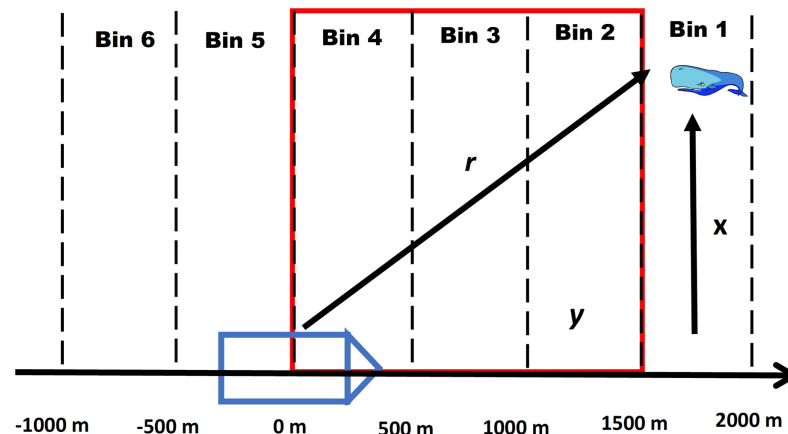
Figure 1 Conceptual diagram of a foraging whale. Conceptual diagram illustrating how a click train of a foraging whale is divided into equally spaced intervals where y represents the forward distance, r is the radial distance and x is the perpendicular distance. The red outline details the area of overlap with the visual team where it is assumed that a whale can be detected by either platform depending on its diving state. Note that the intervals continue behind the ship as the acoustic array is capable of picking up signals both in front of and behind the ship.   Full-size ⊡ DOI: 10.7717/peerj.15850/fig-1

$$z_{ij} \sim Bernoulli\left(\phi z_{i,j} + \gamma_j\left(1 - z_{i,j-1}\right)\right), \tag{6}$$

Equation (3) assumes that the probability of remaining in the foraging state is constant regardless of how long a whale is observed. However, the probability of transitioning out of the foraging state is dependent on how long an animal has been in the diving state. Therefore, we model ϕ as a function of time where $Time_{i,j}$ represents the amount of time whale i has already been in the foraging state when it enters bin j which can be written as

$$Time_{i,j} = Time_{i,1}z_{i,1} + \sum_{j=2}^{J} z_{i,j}T_j, \tag{7}$$

where $Time_{i,1}$ represents the time whale i has already been in the foraging state when it first enters bin 1 and Tj is the amount of time spent in bin j. Because we assume the ship is moving at constant rate then each distance bin represents the same unit of time. Because we do not know how long a whale has been in the foraging state when it first enters the zone of detection we need to assign a prior probability. We use a uniform distribution such that $Time_{i,1} \sim$ uniform(0, maxTime) where maxTime represents the maximum amount of time a whale can be in the foraging state. We then model the probability of remaining in the foraging phase as

$$logit\left(\phi_{i,j}\right) = \alpha_0 + \alpha_T Time_{i,j}. \tag{8}$$

To model detection probability we assume the probability of acoustically detecting a whale is a function of its radial distance from the acoustic array. Therefore, we modeled the detection probability ($p_{i,j}$) of whale i in distance bin j using a logit link function as

$$logit\left(p_{i,j}\right) = \beta_0 + \beta_1 R_{i,j}, \tag{9}$$

where $R_{i,j}$ is the radial distance of whale i from the acoustic array when it is in bin j and β_1 is the effect of radial distance on detection. To calculate $R_{i,j}$ we use the localized perpendicular distance from the trackline for whale i and the midpoint of forward distance of bin j.

Finally, the likelihood of the observed acoustic data was modeled as the outcome of a Bernoulli trial as

$$Y_{ij} \sim Bernoulli\left(p_{i,j}z_{i,j}\right), \tag{10}$$

where $Y_{i,j}$ is an indicator variable that equals 1 when animal i is detected clicking in bin j and 0 otherwise.

By summarizing the unknown states for every whale in each distance bin we estimated a number of derived statistics using data augmentation (*Royle, Dorazio & Link, 2007*; *Royle & Dorazio, 2008*). Data augmentation involves increasing the data set of n individuals to size M where M-n represents the number of "pseudo-individuals" with zero observations that have been added to the original dataset of capture histories. Using this approach, we estimate for each distance bin j the number entering the foraging state (F_j) or number transitioning from the foraging state to the surface state (S_j) (Table 1). By summing F_j and S_j across bins we also estimate the total number transitioning from the surface state to the foraging state (F_T) and number transitioning from the foraging state to the surface state (S_T). In terms of duplicates, we are only interested in the number of animals that transition within a "zone of overlap" between the VLT and PAM platforms (see Fig. 1), and therefore, we limit the bins in the calculation of F_T and S_T to this zone. Finally, we estimate the total number of whales that had been below the surface in the foraging state ($N_B^{(CMR)}$), which is also referred to as the super population using the terminology from *Royle & Dorazio (2008)* (Table 1). To estimate $N_B^{(CMR)}$, F_T and S_T caution needs to be used to determine what distance bins to include.

Zone of Overlap- To accurately estimate F_T and S_T it is necessary to define the appropriate zone of overlap between the VLT and PAM platforms by defining the minimum and maximum forward distance bins where duplicates can occur. In addition, for $N_B^{(CMR)}$ the zone should be small enough to minimize the possibility of counting repeat divers (see Appendix S1). For example, a whale that transitions to the surfacing state behind the ship will not be available to the VLT platform. Because there are silent states at the beginning and end of a dive, the zone of overlap should take this into account. Therefore, we define Y_{max}, Y_{min}, Y_{Max_F}, Y_{Min_F}, Y_{Max_S} and Y_{Min_S} as the maximum and minimum distance bins to include in the calculation of $N_B^{(CMR)}$, F_T and S_T, respectively (see Table 1). We provide a more detailed explanation of defining zones of overlap in Appendix S3.

To integrate the visual and acoustic data we use Eq. (4) to estimate N_T. We calculate the total number of duplicates as $N_D = F_T + S_T$. To estimate N_S from the VLT survey either Eqs. (1) or (2) can be used depening on the survey design. We estimate $N_B^{(CMR)}$ using the PAM data and the above analysis CMR method.

Finally, because we define availability bias (a_s) as the proportion of whales that were available to be detected at the surface, it is simply calculated as.

Table 1 Parameter and definitions for derived parameters for the *CMR-DS Method*.

Parameter	Equation	Definition
F_j	$F_j = \sum\limits_{i=1}^{M} \left(1 - z_{1j-1}\right) * z_{ij}$	Number entering the foraging state in forward bin j.
F_T	$F_T = \sum\limits_{j=Y_{Max_D}}^{j=Y_{Min_D}} F_j$	Total number entering the foraging state in within the window of overlap
S_j	$S_j = \sum\limits_{i=1}^{M} z_{ij-1} * \left(1 - z_{ij}\right)$	Number of individuals entering the surfacing state in forward bin j
S_T	$S_T = \sum\limits_{j=Y_{Max_S}}^{j=Y_{Min_S}} S_j$	Total number entering the surfacing state in within the window of overlap
$N_B^{(CMR)}$	$N_B^{(CMR)} = \sum\limits_{i=1}^{M} I\left\{ \sum\limits_{j=Y_{Max}}^{j=Y_{Min}} z_{ij} > 0 \right\}$	Total number ever in the foraging state over all sampling occasions. Analogous to the superpopulation from a Jolly-Seber analysis.

$$a_S = \frac{N_S}{N_T}. \tag{11}$$

Hybrid method

Problems with ambiguity and processing time may preclude the ability to develop full capture histories for all click train events detected by the towed array. However, there may be enough information from a subset of fully annotated click trains to estimate the transition probabilities among states. These estimates can be used to estimate the proportion of whales below the surface that likely transitioned while in the zone of overlap and therefore are duplicates. Additionally, despite ambiguity it may be possible to localize all acoustically detected whales and assign a perpendicular distance. For this scenario, we developed a hybrid approach between applying DS and CMR to analyze the acoustic data and correct for the number of duplicates.

We premise this method on the assumption that the number of whales entering the foraging state or transitioning to the surface state is approximately constant among all distance bins. As a corollary, the *per capita* rate (*i.e.*, percent of the total abundance) transitioning to the foraging state (P_F) or surface state (P_S) within a given bin is also assumed constant. We also assume that all whales detected by the acoustic array can be localized and assigned a perpendicular distance from the trackline, and therefore, we can estimate abundance below the surface using standard DS techniques as previous studies have done (*Barlow & Taylor, 2005*; *Westell et al., 2022*). Finally, we assume that the subset of available annotated click trains is representative of all click trains, including those not annotated. Under these assumptions, we model the number of whales transitioning to the foraging state ($F_j^{(H)}$) or the surfacing state ($S_j^{(H)}$) as a proportion of the total number below where

$$F_j^{(H)} = N_B^{(DS)} P_F \tag{12}$$

and

$$S_j^{(H)} = N_B^{(DS)} P_S \qquad (13)$$

where $N_B^{(DS)}$ is the total number below estimated from applying DS to the localized acoustic detections. To calculate $F_T^{(H)}$ and $S_T^{(H)}$ we multiply $F_j^{(H)}$ and $S_j^{(H)}$ by the total number of distance bins in the zone of overlap, respectively (see Table 2). For example, if we estimate an average of five whales transitioning from the foraging state to the surface forage state ($F_j^{(H)} = 5$) and there are a total of 10 bins in the zone of overlap then we would estimate a total of 50 duplicates.

To estimate P_F and P_S we use the subset of click trains that were processed into capture histories. Using this subset of capture histories we estimate the number entering the foraging state $f_j^{(H)}$ and surfacing state $s_j^{(H)}$ and the superpopulation $n_B^{(H)}$ where the lowercase letters indicate that these quantities are estimated from a subset of the PAM data (see Table 2). We provide more detail and an example application of this method in the Case Study section.

Simulations

Our current method makes a number of simplifying assumptions. For example, it assumes a constant probability of entering the foraging state and that remaining in the state is a linear function of time spent foraging. In addition, it does not include individual variation in dive behavior and does not correct for changes in depth. To test the robustness of these methods we designed a simulation framework with the intention of replicating sperm whale diving behavior as closely as possible. We based our simulation on data from *Watwood et al. (2006)* who used data collected from DTAGs to characterize the dive cycle and vocalizing behavior of sperm whales. We used maximum depth and the percentage of time in various diving states reported in *Watwood et al. (2006)* to simulate sperm whale diving behavior. Our simulations include an initial silent state at the start of a dive, an acoustically active foraging state and an extended silent state during the ascent. We also included individual variation in dive cycles and maximum depth. We used half-normal functions to simulate the detection process for both the VLT and PAM data. For more details see Appendix S4.

We simulated 100 datasets and applied all three methods to each dataset to estimate abundance and surface availability as well as coefficients of variation (CV) for each estimate. We calculated relative bias as $(\Theta_{ESTIMATE} - \Theta_{TRUE})/\Theta_{TRUE}$ where Θ_{TRUE} represents the true abundance or availability bias and $\Theta_{ESTIMATE}$ represents the estimate from the model.

Case study with sperm whales
Data collection

We applied this method on sperm whale data collected during the Northeast Fisheries Science Center's (NEFSC) shipboard survey conducted from 4 July–17 August 2013 off of the northeast US coast (*Palka et al., 2017*). During this survey, passive acoustic data and visual line transect data were collected simultaneously. A two team visual sampling approach was used such that $g(0)$ could be estimated directly from the visual data.

Table 2 Parameter and definitions for derived parameters from the *Hybrid Method*.

Parameter	Equation	Definition
$F_j^{(H)}$	Eq. (9) methods	Number entering the foraging state in forward bin j
$S_j^{(H)}$	Eq. (10) methods	Number entering the surfacing state in forward bin j
$F_T^{(H)}$	$F_j^{(H)} Zone_F$	Total number entering the foraging state within a defined zone of overlap ($Zone_F$)
$S_T^{(H)}$	$S_j^{(H)} Zone_S$	Total number entering the surface state within a defined zone of overlap ($Zone_S$)
$f_j^{(H)}$	Same as F_j (see Table 1)	Subset of the number entering the foraging state in bin j estimated from a subset of acoustic capture histories
$s_j^{(H)}$	Same as S_j (see Table 1)	Subset of the number entering the surfacing state in bin j estimated from a subset of acoustic capture histories
$n_B^{(H)}$	Same as $N_B^{(CMR)}$ (see Table 1)	Subset of the number ever in the foraging state over all sampling occasions estimated from a subset of acoustic capture histories
P_F	$P_F = \dfrac{\overline{f}^{(H)}}{\overline{n}_B^{(H)}}$	*Per capita* number entering the foraging state per forward bin estimated from the *Hybrid Method* where $\overline{f}^{(H)}$ is the average number entering the foraging sate per bin
P_S	$P_S = \dfrac{\overline{s}^{(H)}}{\overline{n}_B^{(H)}}$	*Per capita* number entering the surface state per forward bin estimated from the *Hybrid Method* where $\overline{s}^{(H)}$ is the average number entering the foraging sate per bin

One team was situated in front of the bridge approximately 11.8 m above the water level and the other team was positioned 15.1 m above the water level. Each visual observer searched the waters in front of the ship for groups of whales using high powered 150 powered binoculars. Radial distance and angle were recorded. For each sighting, the data collected to calculate the perpendicular distance was the distance and angle between the observer and center of the group. In addition, each observer estimated the group size of each group. Information on detection covariates collected including Beaufort sea state, visibility sighting cue and magnitude of glare. Further details can be found in *Palka et al. (2017)*.

We conducted PAM surveys concomitantly with VLT surveys using a towed hydrophone array. We typically deployed the towed array in waters of 100 m or greater and approximately 300 m behind the ship during normal visual survey operations. The array was comprised of two oil filled, modular sections, separated by 30 m of cable. There were 8 hydrophones in total, each pair spaced 1 meter apart. Only the last two hydrophones in the array were used as they were the farthest from the ship and therefore recorded less noise from the ship. The hydrophones used in this study were APC International, Ltd. (Mackeyville, PA, model 42-1021; uncalibrated; 36 dB gain added) with a sampling rate of 192 kHz. *Palka et al. (2017)* provides further details of the array design and subsequent data collection.

Data processing

Recordings were collected and then analyzed using the acoustic software PAMGuard v.1.12.02 (*Gillespie et al., 2009*). PAMGuard's click detector was run across all the data on the last hydrophone pair in the array. Click trains that were received on similar bearings were grouped into "events", where each event in this analysis is defined as a two-dimensional track of a single whale, represented in PAMGuard as a continuous

change of received bearings. These events were localized using the Target Motion Analysis 2D Simplex Optimisation algorithm within PAMGuard. Each event was assigned a latitude and longitude position, a timestamp and bearing angle for each click in the event and an ID number. Using this information and the GPS location of the ship, we calculated a radial distance for each annotated click.

We processed a total of 155 acoustic events that could be localized and used in the analysis. Because of the difficulty in annotating all clicks in each event we were only able to fully annotate clicks for a subset of the data. We used a combination of annotated clicks and capture histories at the 1 min level for time periods that had too many whales clicking simultaneously. To develop capture histories we divided the events into one minute time bins and assigned a 1 if any clicks were observed within that time bin and a 0 if otherwise. This approach precluded the need to annotate every click within an event. We converted these time bins into forward distance intervals based on the forward distance of the first recorded click and the constant speed of the ship (see Appendix S2 for more details).

Analysis

Because we only had a subset of click trains in which all of the clicks could be annotated within an event, we could only apply the *Hybrid Method* to this dataset. To estimate N_S from the visual data we used a MRDS approach (*Laake & Borchers, 2004*). To estimate N_B from the acoustic data we used Eq. (1) with a hazard rate function for $g(y)$. For the CMR data we used capture histories developed from the data processing. For a comparison to more traditional methods, we also applied MRDS with an availability bias correction factor to the VLT data where we used an estimate of surface availability from *Palka et al. (2017)*. For a more detailed summary of the sperm whale analysis see Appendix S5.

Model implementation

We performed all analyses in a Bayesian framework using the software R and JAGS (*Plummer, 2003*). We assumed vague priors for all parameters. We used a burn-in period of 10,000 iterations followed by an additional 15,000 iterations at a thinning rate of 15. Convergence was checked by examining trace plots and calculating Gelman—Rubin statistics, with convergence presumed when R < 1.1 (*Brooks & Gelman, 1998*).

RESULTS

Simulations

Results from the simulations demonstrated that the *DS-DS Method* which ignores duplicates had a relatively high positive bias for abundance (+54.4%) and relatively high negative bias for surface availability (−33.0%) (Fig. 2). The *CMR-DS Method* had the least bias in terms of abundance (−1.2%) and surface availability (−0.2%) where the bias for the *Hybrid Method* was 3.2% and −3.4% for abundance and surface availability, respectively (Fig. 2). In terms of precision, the *CMR-DS Method* was more precise than the *Hybrid Method*. Coefficients of Variation (CVs) for abundance estimates ranged from 0.03 to 0.06 for the *CMR-DS Method* and 0.05 to 0.15 for the *Hybrid Method*. For surface availability
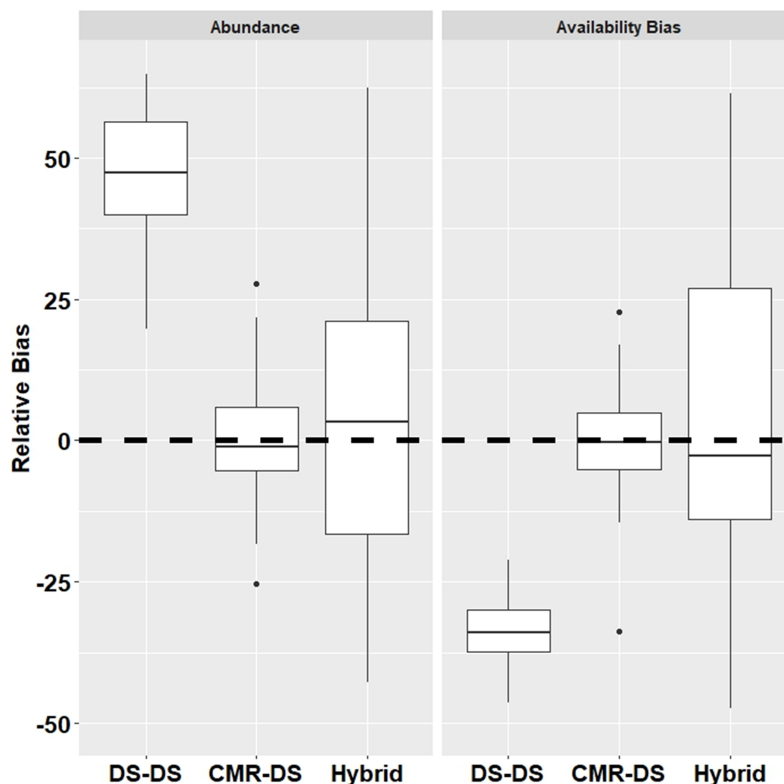
Figure 2 Relative bias from simulations by method for abundance and availability bias.
Full-size ⬓ DOI: 10.7717/peerj.15850/fig-2

CVs ranged from 0.03 to 0.07 for the *CMR-DS Method* and 0.05 to 0.18 for the *Hybrid Method*.

Case study

Gelman-Rubin statistics indicated adequate convergence of all parameters (Table S1) and the CMR component of the *Hybrid Method* provided an adequate fit to the PAM data as evidenced by a Bayesian *p*-value of 0.28. From the MRDS analysis of the visual data using the *Hybrid Method* we estimated an effective strip width (ESW) of approximately 3.8 km (CV = 0.21) and a $g(0)$ of 0.80 (CV = 0.10). From the DS analysis of the acoustic data, we estimated an ESW of approximately 4.1 km (CV = 0.08). The estimate of abundance at the surface (N_S) was approximately 6% higher than the estimate of abundance below the surface (N_B) but with a high overlap in credible intervals (Table 3).

The final estimate of abundance from the *Hybrid Method* combining VLT and PAM survey data was approximately 17% lower than the estimate of abundance from the $MRDS_{AV}$ analysis using only VLT survey data but with a CV that was approximately 50% lower (Table 3). The estimate of surface availability was 0.72 which was similar to the estimate of 0.61 from *Palka et al. (2017)* with a CV of 0.08 which is lower than the CV of 0.25 reported in *Palka et al. (2017)*.

Table 3 Posterior parameter estimates of abundance.

Parameter	Method	Estimate	CV	Upper	Lower
N_S	Hybrid	306	0.38	668	216
N_B	Hybrid	289	0.08	335	249
N_T	Hybrid	426	0.29	814	332
N_T	MRDS$_{AV}$	516	0.58	1,398	294

Note:
Posterior parameter estimates of abundance at the surface (N_S), abundance below the surface (N_B) from the *Hybrid Method* and total abundance (N_T) from both the *Hybrid Method* and MRDS$_{AV}$ method for sperm whales. The coefficient of variation (CV) and the upper and lower 95% credible interval are provided for each estimate.

DISCUSSION

In this study, we explored the challenges of combining PAM towed array data with VLT survey data when the surveys are conducted simultaneously and developed a statistical framework to estimate abundance and surface availability bias. In particular, we focused on the challenge of correcting for duplicate detections. Using simulations and a case study of sperm whales, we explored several methods that range in complexity and data requirements. Our results demonstrate that (1) duplicates can lead to significant bias if ignored (2) modeling transition probabilities can greatly reduce bias in estimates of surface availability and abundance compared to a more simplistic approach and (3) combining these two sources of data can increase precision of estimates of abundance and surface availability.

To evaluate our methodology, we developed simulations based on known diving behavior of sperm whales. Using this simulation design we could assess, firstly, whether or not the process of transitioning among different states in the dive cycle could cause significant bias if ignored, and secondly, if our proposed framework can adequately capture this dynamic and adjust for this bias appropriately. We found significant bias in the *DS-DS Method* suggesting that the process of transitioning among diving states during a survey can result in a significant number of duplicate detections. The comparatively large reduction in bias exhibited by the *CMR-DS Method* suggests our framework can capture this dynamic and adjust for this bias. In addition, the *Hybrid Method* proved to be a viable alternative when the ability to fully annotate all click train events is not feasible. However, it is important to note that the *Hybrid Method* is less precise than the *CMR-DS Method* and, the level of precision will be partly influenced by the sample size of click trains that can be fully annotated. Finally, our simulations demonstrated that PAM data has the potential to be used to estimate availability bias.

Simulations can be a powerful tool to test new statistical methods (*DiRenzo, Hanks & Miller, 2023*). Despite their usefulness, many studies either do not include simulations or use the statistical model as the data generating model in the simulations precluding a true assessment of model misspecification (*DiRenzo, Hanks & Miller, 2023*). Our goal with this simulation design was to establish the utility of our methodology beyond a theoretical proof of concept. Our simulations were designed to mimic realistic diving behavior in sperm whales to address the challenges we outlined in the Methods. For example, we included silent states, possibility of the same animal being assigned to two different events

(double divers) and individual variation in dive behavior (*e.g.*, maximum depth, time at the surface, *etc.*,). Although our simulations do not capture all the complexity of sperm whale diving and vocalizing behavior, they provide a basis from which to assess the robustness of our methods. The ability of both methods to decrease bias and produce reasonable results lends credence to this framework as a valid statistical tool capable of producing unbiased and precise estimates in the face of complex diving behaviors.

Analysis of the sperm whale dataset with the *Hybrid Method* produced an estimate of abundance that was similar to an estimate from a MRDS$_{AV}$ analysis using only the VLT survey data but with considerably higher precision. For deep diving cetaceans, abundance estimates from VLT data alone can result in lower precision for two reasons. One reason is simply that there are few detections at the surface resulting in low sample sizes for conventional DS analyses. A second reason is that the correction factor for availability bias is often based on a limited amount of data as well which can result in a large amount of uncertainty being propagated to the final estimate (*Sigourney et al., 2020*). In addition to low precision, if the data used to estimate availability bias is limited to, for example, a small sample of tagged individuals it may also lead to bias if the sample is not truly representative of average surfacing time. Studies that use only VLT data often need to pool data among species or surveys to have enough data to estimate abundance (*Barlow & Forney, 2007*). Pooling across species may introduce bias in the detection function and pooling across years may limit the ability to detect trends in abundance. Using the proposed framework we were able to markedly increase the sample of detections by including the PAM data while still taking advantage of the VLT data to derive a more precise estimate of abundance. In our example with field data, we used only 1 year of survey data and were able to achieve a high level of precision. This level of precision can have important management implications, particularly when trying to detect trends (*Taylor et al., 2007*).

Concomitant with continued technological advancement there has been increasing progress towards developing statistical methods that meld together disparate datasets providing better estimates of demographic parameters to inform management and conservation efforts (*Pacifici et al., 2017*; *Miller et al., 2019*). For example, there have been important advancements with species distribution models that use sophisticated integrated modeling techniques (*Miller et al., 2019*). However, even a simple average among abundance surfaces can yield improved results (*Conn et al., 2022*). In our example we are adding data sets together but harnessing the information in the PAM data to develop a correction factor for duplicates. By utilizing the PAM data we can increase precision while our method to model transitions can adjust for bias. However, it is also important to note this result may not be realized in all cases. For example, if PAM detections are limited such that there is large variance in both the PAM estimate of abundance and duplicates then applying the proposed methods may decrease precision over a traditional estimate. Moreover, there may be little to be gained from PAM data if, for example, the sample size for the VLT is already large and either an accurate and precise estimate for surface availability bias already exists or availability bias is negligible.

In addition to abundance our approach demonstrates a novel use of PAM data to estimate availability bias. For a comparison, we contrasted our estimate from the *Hybrid*

Method to an estimate reported in *Palka et al. (2017)*. *Palka et al. (2017)* used DTAG data of sperm whales and the method of *Laake et al. (1997)* to estimate a correction factor for availability bias. Because DTAG data are limited, the estimate from *Palka et al. (2017)* was limited to a small sample of whales. We found close agreement between this estimate and the estimate from the *Hybrid Method*, but the estimate from the *Hybrid Method* resulted in a CV that was approximately three times lower. By combining information from the PAM and VLT surveys we were able to increase the amount of information to estimate availability bias. This increase in data also allows the opportunity to estimate availability bias at finer scales. Studies of diving behavior of deep diving cetaceans are often based on a small sample size precluding the ability to characterize temporal or spatial scale variation in diving behavior. *Watwood et al. (2006)* managed to combine information from several tagging studies which allowed them to characterize diving behavior at regional scales. With our method, potentially finer scales could be examined within a region or study area. Ideally, this may allow for more fine-tuned estimates of availability bias to be applied spatially or temporally for more accurate estimates of abundance.

Because of the potential for large bias in abundance estimates, there have been numerous attempts to develop statistical methods to estimate and adjust for availability bias. These methods either develop an estimator to estimate availability bias and apply it as a correction factor (see *Laake et al., 1997*) or embed information directly into the detection function estimator (*Okamura et al., 2012*; *Borchers et al., 2013*; *Langrock, Borchers & Skaug, 2013*). Most of these approaches require some auxiliary information that is used to correct or adjust the estimator. In our framework, we approach the estimation of availability bias indirectly. By defining availability bias as the proportion of animals that are available at the surface, we can then derive an estimate by combining the estimate of total abundance with an estimate of abundance at the surface from the VLT survey. Our framework does not require auxiliary data on focal follow surveys or tagged animals as input to the model, however, some knowledge of the diving cycle is required to inform the zone of overlap where duplicates can occur. In addition, because the estimate is based on user-defined inputs such as the maximum perpendicular and forward distance it is somewhat bespoke and not directly transferable to other platforms such as aerial surveys.

Limitations and future work

Although initial results are promising, there remain several caveats that could be explored in future extensions of our framework. For example, our methods currently only model transitions into a foraging state. In addition, we do not include an adjustment for depth which may explain the difference between the higher VLT only estimate of abundance and the estimate from the *Hybrid Method* in our case study. We also do not include group size in the PAM estimate of abundance which may also partly explain the lower estimate of abundance from the *Hybrid Method*. Another factor that we did not assess in this study is possible error in the location data of localized whales. Finally, we use relatively simple, linear functions to model the PAM data.

Despite these limitations our flexible framework should allow for multiple extensions. For example, information from DTAG data could be used to better model the probability

of being in a silent state. Also, new techniques may allow for estimating depth from the PAM towed array such that depth could be included as another source of information (*DeAngelis et al., 2017*; *Barkley, Nosal & Oleson, 2021*; *Westell et al., 2022*). In addition, information on group size is available from the VLT surveys and could be used to inform the PAM data (see *Barlow et al. (2021)* for an example). Other extensions could also include modelling error in the location data (see *Borchers et al., 2010*). Finally, more flexible functions could easily be substituted into the CMR component to better capture the dynamics, particularly regarding diving behavior.

Another data limitation in the current analysis is that we only modelled foraging clicks. Typically, non-foraging clicks such as codas occur during ascents and descents of a deep foraging dive (*Frantzis & Alexiadou, 2008*), therefore the dive would be encapsulated if just marking usual clicks. Slow clicks are also emitted within minutes of usual clicks (*Oliveira et al., 2013*), thus the individual would be included in the analysis. These clicks could be added to the analysis as they can also be localized and processed (see *Barkley et al., 2022*). Future modeling attempts, however, should consider using a separate detection function as the detection probability from these events likely differs from foraging clicks.

Along with modifications to the statistical model, further development of the simulation design could be beneficial particularly if our methods are to be applied to other species. With the current simulations we primarily focused on silent states during foraging dives with simulated whales being silent during the initial descent phase and again during the ascent phase. *Isojunno & Miller (2015)*, however, described a more complex array of diving behaviors where they identified six potential behavioral states. Their detailed analysis of these behaviors includes parameter estimates that could be incorporated into future simulations. To extend this simulation design to other species such as beaked whales a different set of behavioral dynamics would also need to be considered. For example, beaked whales engage in shallow non-foraging dives in between longer foraging dives (*Tyack et al., 2006*) and their echolocation is highly directional (*Johnson et al., 2008*) which will influence both their overall availability and detectability. A more intricate simulation design which includes several species could potentially be a useful tool for testing and comparing new methods. In addition, acoustic simulations can also play a role in assisting study design (*Peel et al., 2014*), and therefore, can serve as an important stand-alone tool beyond testing statistical methods.

CONCLUSIONS

The use of PAM survey technology is a rapidly growing area in ecological studies creating a fundamental need for analytic tools to analyze and integrate these data with other data streams (*Gibb et al., 2019*). As more datasets become available it will be necessary to continue to develop a diverse statistical toolbox. As a step towards this end, we focused on the challenges of integrating PAM towed array data with VLT data when the surveys are conducted simultaneously. Our results demonstrate the value of combining these data for abundance estimates and a novel use of PAM data to estimate availability bias for sperm whales. Although results are promising, our methods are not intended as a panacea and clear challenges remain. With the rapid growth in statistical tools, we anticipate continued

progress in this arena. Going forward, we recommend rigorous development of simulations to accompany tool development to fully assess the robustness of methods.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

Douglas B. Sigourney is employed by Integrated Statistics.

Author Contributions

- Douglas B. Sigourney conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Annamaria DeAngelis conceived and designed the experiments, performed the experiments, prepared figures and/or tables, authored or reviewed drafts of the article, annotated and processed acoustic data, and approved the final draft.
- Danielle Cholewiak conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Debra Palka conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Data and code are available at GitHub and Zenodo: https://github.com/NEFSC/READ-PSB-DSIGOURNEY-AcousticDataIntegration.

Douglas Sigourney. (2023). Acoustic Data Integration. https://doi.org/10.5281/zenodo.7566500.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.15850#supplemental-information.

REFERENCES

Avila IC, Kaschner K, Dormann CF. 2018. Current global risks to marine mammals: taking stock of the threats. *Biological Conservation* **221**:44–58 DOI 10.1016/j.biocon.2018.02.021.

Barkley YM, Nosal E, Oleson ER. 2021. Model-based localization of deep-diving cetaceans using towed line array acoustic data. *Journal of the Acoustical Society of America* **150**:1120–1132 DOI 10.1121/10.0005847.

Barkley YM, Sakai T, Oleson EM, Franklin EC. 2022. Examining distribution patterns of foraging and non-foraging sperm whales in hawaiian waters using visual and passive acoustic data. *Frontiers in Remote Sensing* **3**:940186 DOI 10.3389/frsen.2022.940186.

Barlow J, Forney KA. 2007. Abundance and population density of cetaceans in the California current ecosystem. *Fishery Bulletin* **105**:509–526.

Barlow J, Fregosi S, Thomas L, Harris D, Griffiths ET. 2021. Acoustic detection range and population density of Cuvier's beaked whales estimated from near-surface hydrophones. *The Journal of the Acoustical Society of America* **149(1)**:111–125 DOI 10.1121/10.0002881.

Barlow J, Taylor BL. 2005. Estimates of sperm whale abundance in the northeastern temperate pacific from a combined acoustic and visual survey. *Marine Mammal Science* **21(3)**:429–445 DOI 10.1111/j.1748-7692.2005.tb01242.x.

Borchers D, Marques T, Gunnlaugsson T, Jupp P. 2010. Estimating distance sampling detection functions when distances are measured with errors. *Journal of Agricultural, Biological, and Environmental Statistics* **15(3)**:346–361 DOI 10.1007/s13253-010-0021-y.

Borchers DL, Zucchini W, Heide-Jørgensen MP, Cañadas A, Langrock R. 2013. Using hidden markov models to deal with availability bias on line transect surveys. *Biometrics* **69(3)**:703–713 DOI 10.1111/biom.12049.

Brooks SP, Gelman A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7(4)**:434–455 DOI 10.2307/1390675.

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L. 2001. *Introduction to distance sampling: estimating abundance of biological populations*. New York (NY): Oxford University Press.

Burt ML, Borchers DL, Jenkins KJ, Marques TA. 2014. Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution* **5(11)**:1180–1191 DOI 10.1111/2041-210X.12294.

Conn PB, Ver Hoef JM, McClintock BT, Johnson DS, Brost B. 2022. A GLMM approach for combining multiple relative abundance surfaces. *Methods in Ecology and Evolution* **13**:2236–2247 DOI 10.1111/2041-210X.13948.

DeAngelis AI, Valtierra R, van Parijs SM, Cholewiak D. 2017. Using multipath reflections to obtain dive depths of beaked whales from a towed hydrophone array. *The Journal of the Acoustical Society of America* **142(2)**:1078–1087 DOI 10.1121/1.4998709.

DiRenzo GV, Hanks E, Miller DAW. 2023. A practical guide to understanding and validating complex models using data simulations. *Methods in Ecology and Evolution* **14(1)**:203–217 DOI 10.1111/2041-210X.14030.

Fleming AH, Yack T, Redfern JV, Becker EA, Moore TJ, Barlow J. 2018. Combining acoustic and visual detections in habitat models of dall's porpoise. *Ecological Modelling* **384(4)**:198–208 DOI 10.1016/j.ecolmodel.2018.06.014.

Forcada J, Gazo M, Aguilar A, Gonzalvo J, Fernandez-Contreras M. 2004. Bottlenose dolphin abundance in the NW Mediterranean: addressing heterogeneity in distribution. *Marine Ecology Progress Series* **275**:275–287 DOI 10.3354/meps275275.

Frantzis A, Alexiadou P. 2008. Male sperm whale (Physeter macrocephalus) coda production and coda-type usage depend on the presence of conspecifics and the behavioural context. *Canadian Journal of Zoology* **86(1)**:62–75 DOI 10.1139/Z07-114.

Gerrodette T, Taylor BL, Swift R, Rankin S, Jaramillo-Legorreta AM, Rojas-Bracho L. 2011. A combined visual and acoustic estimate of 2008 abundance, and change in abundance since 1997, for the vaquita, *Phocoena* sinus. *Marine Mammal Science* **27(2)**:E79–E100 DOI 10.1111/j.1748-7692.2010.00438.x.

Gibb R, Browning E, Glover-Kapfer P, Jones KE. 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution* **10(2)**:169–185 DOI 10.1111/2041-210X.13101.

Gillespie D, Mellinger DK, Gordon J, McLaren D, Redmond P, McHugh R, Trinder P, Deng XY, Thode A. 2009. PAMGUARD: semiautomated, open source software for real-time acoustic detection and localization of cetaceans. *Journal of the Acoustical Society of America* **125**:2547 DOI 10.1121/1.4808713.

Hamilton ONP, Kincaid SE, Constantine R, Kozmian-Ledward L, Walker CG, Fewster RM. 2018. Accounting for uncertainty in duplicate identification and group size judgements in mark-recapture distance sampling. *Methods in Ecology and Evolution* **9(2)**:354–362 DOI 10.1111/2041-210X.12895.

Hammond PS, Francis TB, Heinemann D, Long KJ, Moore JE, Punt AE, Reeves RR, Sepúlveda M, Sigurðsson GM, Siple MC, Víkingsson G, Wade PR, Williams R, Zerbini AN. 2021. Estimating the abundance of marine mammal populations. *Frontiers in Marine Science* **8**:735770 DOI 10.3389/fmars.2021.735770.

Isojunno S, Miller PJO. 2015. Sperm whale response to tag boat presence: biologically informed hidden state models quantify lost feeding opportunities. *Ecosphere* **6(1)**:1–46 DOI 10.1890/ES14-00130.1.

Johnson M, Hickmott LS, Aguilar Soto N, Madsen PT. 2008. Echolocation behaviour adapted to prey in foraging blainville's beaked whale (*Mesoplodon densirostris*). *Proceedings of the Royal Society B: Biological Sciences* **275(1631)**:133–139 DOI 10.1098/rspb.2007.1190.

Johnson MP, Tyack PL. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering* **28(1)**:3–12 DOI 10.1109/JOE.2002.808212.

Laake JL, Borchers DL. 2004. Methods for incomplete detection at distance zero. In: Buckland S, Anderson D, Burnham K, Laake J, Borchers D, eds. *Advanced Distance Sampling*. Oxford: U.K. Oxford University press, 108–189.

Laake JL, Calambokidis J, Osmek SD, Rugh DJ. 1997. Probability of detecting harbor porpoise from aerial surveys: estimating g(0). *Journal of Wildlife Management* **61(1)**:63–75 DOI 10.2307/3802415.

Langrock R, Borchers DL, Skaug HJ. 2013. Markov-modulated nonhomogeneous poisson processes for modeling detections in surveys of marine mammal abundance. *Journal of the American Statistical Association* **108(503)**:840–851 DOI 10.1080/01621459.2013.797356.

Madsen PT, Wahlberg M, Møhl B. 2002. Male sperm whale (Physeter macrocephalus) acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology* **53**:31–41 DOI 10.1007/s00265-002-0548-1.

Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Moretti DJ, Harris D, Tyack PL. 2013. Estimating animal population density using passive acoustics. *Biological Reviews* **88(2)**:287–309 DOI 10.1111/brv.12001.

Marsh H, Sinclair DF. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* **53**:1017–1024 DOI 10.2307/3809604.

McLaren IA. 1961. Methods of determining the numbers and availability of ring seals in the Eastern Canadian Arctic. *Arctic* **14**:162–175 DOI 10.14430/arctic3671.

Mellinger DK, Thode AM, Martinez A. 2002. Passive acoustic monitoring of sperm whales in the Gulf of Mexico, with a model of acoustic detection distance. In: *Proceedings: Twenty-First Annual Gulf of Mexico Information Transfer Meeting,* 493–501.

Miller PJO, Johnson MP, Tyack PL. 2004. Sperm whale behaviour indicates the use of echolocation click buzzes "creaks" in prey capture. *Proceedings of the Royal Society B: Biological Sciences* **271(1554)**:2239–2247 DOI 10.1098/rspb.2004.2863.

Miller DAW, Pacifici K, Sanderlin JS, Reich BJ. 2019. The recent past and promising future for data integration methods to estimate species' distributions. *Methods in Ecology and Evolution* **10(1)**:22–37 DOI 10.1111/2041-210X.13110.

Møhl B, Wahlberg M, Madsen PT, Heerfordt A, Lund A. 2003. The monopulsed nature of sperm whale clicks. *The Journal of the Acoustical Society of America* **114(2)**:1143–1154 DOI 10.1121/1.1586258.

Nykänen M, Jessopp M, Doyle TK, Harman LA, Cañadas A, Breen P, Hunt W, Mackey M, Cadhla O, Reid D, Rogan E. 2018. Using tagging data and aerial surveys to incorporate availability bias in the abundance estimation of blue sharks (*Prionace glauca*). *PLOS ONE* **13(9)**:e0203122 DOI 10.1371/journal.pone.0203122.

Okamura H. 2003. A line transect method to estimate abundance of long-diving animals. *Fisheries Science* **69**:1176–1181 DOI 10.1111/j.0919-9268.2003.00743.x.

Okamura H, Minamikawa S, Skaug HJ, Kishiro T. 2012. Abundance estimation of long-diving animals using line transect methods. *Biometrics* **68(2)**:504–513 DOI 10.1111/j.1541-0420.2011.01689.x.

Oliveira C, Wahlberg M, Johnson M, Miller PJO, Madsen PT. 2013. The function of male sperm whale slow clicks in a high latitude habitat: communication, echolocation, or prey debilitation? *The Journal of the Acoustical Society of America* **133(5)**:3135–3144 DOI 10.1121/1.4795798.

Pacifici K, Reich BJ, Miller DAW, Gardner B, Stauffer G, Singh S, McKerrow A, Collazo JA. 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. *Ecology* **98(3)**:840–850 DOI 10.1002/ecy.1710.

Palka DL, Chavez-Rosales S, Josephson E, Cholewiak D, Haas HL, Garrison L, Jones M, Sigourney D, Waring G, Jech M, Broughton E, Soldevilla M, Davis G, DeAngelis A, Sasso CR, Winton MV, Smolowitz RJ, Fay G, LaBrecque E, Leiness JB, Dettloff Warden, Murray M, Orphanides C. 2017. *Atlantic marine assessment program for protected species:*

2010–2014. Washington, DC: US Dept. of the Interior, Bureau of Ocean Energy Management, 211.

Peel D, Miller BS, Kelly N, Dawson S, Slooten E, Double MC. 2014. A simulation study of acoustic-assisted tracking of whales for mark-recapture surveys. *PLOS ONE* **9(5)**:e95602 DOI 10.1371/journal.pone.0095602.

Plummer M. 2003. JAGS: a program for analysis of bayesian graphical models using gibbs sampling. In: *Proceedings of the Third International Workshop on Distributed Statistical Computing*. Vienna, Austria: R Project for Statistical Computing.

Royle JA, Dorazio RM, Link WA. 2007. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* **16(1)**:67–85 DOI 10.1198/106186007X181425.

Royle JA, Dorazio RM. 2008. *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Oxford, UK: Academic Press.

Sigourney DB, Chavez-Rosales S, Conn PB, Garrison L, Josephson E, Palka D. 2020. Developing and assessing a density surface model in a Bayesian hierarchical framework with a focus on uncertainty: insights from simulations and an application to fin whales (Balaenoptera physalus). *PeerJ* **8**:e8226 DOI 10.7717/peerj.8226.

Stevenson BC, Borchers DL, Fewster RM. 2019. Cluster capture-recapture to account for identification uncertainty on aerial surveys of animal populations. *Biometrics* **75(1)**:326–336 DOI 10.1111/biom.12983.

Stevenson BC, Fewster RM, Sharma K. 2022. Spatial correlation structures for detections of individuals in spatial capture-recapture models. *Biometrics* **78(3)**:963–973 DOI 10.1111/biom.13502.

Sucunza F, Danilewicz D, Cremer M, Andriolo A, Zerbini AN. 2018. Refining estimates of availability bias to improve assessments of the conservation status of an endangered dolphin. *PLOS ONE* **13(3)**:e0194213 DOI 10.1371/journal.pone.0194213.

Taylor BL, Martinez M, Gerrodette T, Barlow J, Hrovat YN. 2007. Lessons from monitoring trends in abundance of marine mammals. *Marine Mammal Science* **23(1)**:157 DOI 10.1111/j.1748-7692.2006.00092.x.

Thompson PM, Brookes KL, Cordes LS. 2015. Integrating passive acoustic and visual data to model spatial patterns of occurrence in coastal dolphins. *ICES Journal of Marine Science* **72**:651–660 DOI 10.1093/icesjms/fsu110.

Tyack PL, Johnson M, Soto NA, Sturlese A, Madsen PT. 2006. Extreme diving of beaked whales. *Journal of Experimental Biology* **209(21)**:4238–4253 DOI 10.1242/jeb.02505.

Tønnesen P, Gero S, Ladegaard M, Johnson M, Madsen PT. 2018. First-year sperm whale calves echolocate and perform long, deep dives. *Behavioral Ecology and Sociobiology* **72**:165 DOI 10.1007/s00265-018-2570-y.

Watkins WA, Schevill WE. 1977. Sperm whale codas. *Journal of the Acoustical Society of America* **62(6)**:1485–1490 DOI 10.1121/1.381678.

Watwood SL, Miller PJO, Johnson M, Madsen PT, Tyack PL. 2006. Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*). *Journal of Animal Ecology* **75(3)**:814–825 DOI 10.1111/j.1365-2656.2006.01101.x.

Westell A, Sakai T, Valtierra R, van Parijs SM, Cholewiak D, DeAngelis A. 2022. Sperm whale acoustic abundance and dive behaviour in the western North Atlantic. *Scientific Reports* **12(1)**:21 DOI 10.1038/s41598-022-20868-3.

Whitehead H. 2009. Estimating abundance from one-dimensional passive acoustic surveys. *Journal of Wildlife Management* **73(6)**:1000–1009 DOI 10.2193/2008-303.

Sigourney et al. (2023), *PeerJ*, DOI 10.7717/peerj.15850

22/23

Williamson LD, Brookes KL, Scott BE, Graham IM, Bradbury G, Hammond PS, Thompson PM. 2016. Echolocation detections and digital video surveys provide reliable estimates of the relative density of harbour porpoises. *Methods in Ecology and Evolution* 7:762–769 DOI 10.1111/2041-210X.12538.

Worthington LV, Schevill WE. 1957. Underwater sounds heard from sperm whales. *Nature* **180(4580)**:291 DOI 10.1038/180291a0.

Zipkin EF, Saunders SP. 2018. Synthesizing multiple data types for biological conservation using integrated population models. *Biological Conservation* **217**:240–250 DOI 10.1016/j.biocon.2017.10.017.