



Effect of matching uncertainty on population parameter estimation in mark-recapture analysis of photo-identification data

Erin Ashe^{1,2} · Philip S. Hammond¹

Received: 13 October 2021 / Accepted: 23 January 2022 / Published online: 13 May 2022
© The Author(s) 2022, corrected publication 2022

Abstract

Quantifying and dealing with uncertainty are key aspects of ecological studies. Population parameter estimation from mark-recapture analyses of photo-identification data hinges on correctly matching individuals from photographs and assumes that identifications are detected with certainty, marks are not lost over time, and that individuals are recognised when they are resighted. Matching photographs is an inherently subjective process. Traditionally, two photographs are not considered a “match” unless the photo reviewer is 100% certain. This decision may carry implications with respect to sample size and the bias and precision of the resultant parameter estimates. Here, we present results from a photo-identification experiment on Pacific white-sided dolphins to assign one of three levels of certainty that a pair of photographs represented a match. We then illustrate how estimates of abundance and survival varied as a function of the matching certainty threshold used. As expected, requiring 100% certainty of a match resulted in fewer matches, which in turn led to higher estimates of abundance and lower estimates of survival than if a lower threshold were used to determine a match. The tradition to score two photographs as a match only when the photo reviewer is 100% certain stems from a desire to be conservative, but potential over-estimation of abundance means that there may be applications (e.g., assessing sustainability of bycatch) in which it is not precautionary. We recommend exploring the consequences of matching uncertainty and incorporating that uncertainty into the resulting estimates of abundance and survival.

Keywords Abundance · Capture-recapture · Photo-identification · Precautionary principle · Survival

Introduction

A variety of methods are available for estimating the abundance of marine mammals (and other species) including correcting and extrapolating counts, transect sampling, spatial

modelling, and mark-recapture approaches (Hammond et al. 2021). The individual recognition data obtained by identifying and following individual animals used in mark-recapture approaches to estimate abundance can also be used to estimate survival rates (e.g., Lebreton et al. 1992; Ramp et al. 2014; Arso Civil et al. 2019) and reproductive rates (e.g., Barlow and Clapham 1997; Arso Civil et al. 2017; Coxon et al. 2022), which are essential parameters when modelling the dynamics and assessing the conservation status of animal populations.

Photo-identification has become widely used to follow marine mammals since researchers first noticed that some individuals possessed naturally occurring, identifiable, and persistent features. The unique markings of bottlenose dolphins (*Tursiops truncatus*) were recorded and tracked as early as the 1950s (Caldwell 1955). Photo-identification of killer whales in the northeastern Pacific Ocean began in the 1970s (Bigg 1982), and the resulting demographic records now span decades. Other examples of photo-identification studies that have generated long-term datasets include

Handling editors: Stephen C.Y. Chan and Leszek Karczmarski.

This article is a contribution to the special issue on “Individual Identification and Photographic Techniques in Mammalian Ecological and Behavioural Research – Part 1: Methods and Concepts” — Editors: Leszek Karczmarski, Stephen C.Y. Chan, Daniel I. Rubenstein, Scott Y.S. Chui and Elissa Z. Cameron.

✉ Erin Ashe
ea84@st-andrews.ac.uk

¹ Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY16 8LB, Scotland, UK

² Present Address: Oceans Initiative, 117 E Louisa St #135, Seattle, WA 98102, USA

bottlenose dolphins in Sarasota Bay, Florida (Wells and Scott 1990), North Atlantic right whales (*Eubalaena glacialis*; Pace et al. 2017), North Atlantic humpback whales (*Megaptera novaeangliae*; Stevick et al. 2003), blue (*Balaenoptera musculus*) and fin (*B. physalus*) whales in the Gulf of St Lawrence (Ramp et al. 2006, 2014); and southern right whales off Argentina (*Eubalaena australis*; Agrelo et al. 2021).

Data used in mark-recapture analysis must meet a number of assumptions if reliable parameter estimates are to be made: (1) marks are unique; (2) marks cannot be lost or missed; (3) all marks are correctly recorded and reported (Hammond 2018). The purpose of this study is to explore the variability inherent in correctly matching individuals among sampling occasions.

Errors in individual identification are known to occur (Payne et al. 1983; Langtimm et al. 2004), and the few studies that have explored effects of misidentification have found that, even at small rates, errors in identification can bias parameter estimates (Stevick et al. 2001; Lukacs and Burnham 2005; Yoshizaki et al. 2009). Misidentification involves many factors, but a recurring theme involves the importance of choosing the right features to use as a natural mark in order to satisfy assumption 2, above. Anatomical features should be chosen so that the natural markings used in a mark-recapture experiment will last longer than the experiment and should not change in such a way that might affect the ability to recognize it in future (Wilson et al. 1999). For killer whales, the shape of the dorsal fin and patterns in the saddle patch are most often used as natural marks (Bigg 1982; Kuningas et al. 2014). For humpback whales, pigmentation patterns on the underside of the flukes, as well as the edge of the flukes themselves, are used to identify individuals (Stevick et al. 2001, 2003). In this study, naturally occurring nicks and notches in a Pacific white-sided dolphin's (*Lagenorhynchus obliquidens*; Fig. A1) dorsal fin were used as natural marks, such that it could be recognized from both left- and right-side photographs.

Observers tend to conflate photo-quality with animal distinctiveness because a well-marked individual is more easily recognized than a subtly marked individual in a poor-quality photograph (Urian et al. 2015). As a result, previous cetacean studies have relied on strict protocols when gauging whether two photographs are a match (Wilson et al. 1999; Read et al. 2003). However, the final dataset used to estimate population parameters is still subject to human error, because it is dependent on a somewhat subjective decision about whether a human observer is convinced that a pair of photographs represent two encounters of the same individual or two different individuals. Little attention has been paid to the process by which researchers reach a final decision about whether two photographs constitute a match, but a survey has shown that researchers vary widely in their approach

to defining a match (Urian et al. 2015). Historically, cetacean studies use “conservative” protocols and, after seeking advice from experienced colleagues in the case of any ambiguity, only score two photographs as a match if there is consensus among observers (Friday et al. 2000; Stevick et al. 2001; Urian et al. 2015). Most protocols reviewed by Urian et al. (2015) are inherently averse to false positives; the corollary to this is that false negatives will arise as a result (Stevick et al. 2001). Not all researchers use protocols that are averse to false positives. Urian and colleagues reported “an unsettling degree of variation among researchers in the evaluation of image quality, distinctiveness, images selected and matches. Participants from the same institution generally had similar results, suggesting that most variation was due to the different methods used by each laboratory.” Many researchers may be trained to quantify their level of certainty that two photographs do or do not represent a match, but there is little guidance from statisticians about how to incorporate that uncertainty into the binary framework of conventional mark-recapture models.

Erring on the side of false negatives is not always a precautionary approach. Deciding always to call ambiguous matches a non-match will cause recapture rates to be biased low, which will cause estimates of abundance to be positively biased and estimates of survival rates to be negatively biased (Hammond 1986; Hammond et al. 1990; Friday et al. 2008). For management procedures that set allowable harm limits based on abundance (e.g., Wade 1998; Winship et al. 2006; Genu et al. 2021) a positively biased abundance estimate could lead to overexploitation. The extent to which this is a problem for real-world conservation and management decisions is case-specific, but few studies have estimated the magnitude of bias in abundance and survival estimates depending on matching uncertainty.

There are two primary reasons for misidentification errors: (1) errors in identification due to changes in the natural markings; and (2) misidentification as a result of variation at the level of the matching process. The first can occur if individuals acquire new marks such as scars or damage due to predation or intra-specific interactions (Gordon 1987; Steiger et al. 2008), or if marks such as scratches or pigmentation patterns heal and subsequently disappear (Dufault and Whitehead 1995). Dufault and Whitehead (1995) found that mark acquisition occurred at a higher rate than mark loss. Mark acquisition is presumed less likely to cause misidentification, especially in small populations (Urian et al. 2015), but it is easy to imagine a scenario in which mark acquisition may lead to changes that are substantial enough for false negative errors to occur. For larger, wide-ranging populations, it is recommended that mark acquisition rates are estimated and that strict animal distinctiveness criteria that rely on markings that are unlikely to change over time are used (Urian et al. 2015).

The second misidentification process, errors that occur at the level of the matching process, has received comparatively little attention. Previous analyses have shown that conflating photo-quality with individual distinctiveness biases the matching process and, subsequently, the parameter estimates from mark-recapture analyses (Arnbom 1987; Friday et al. 2000, 2008). False rejections of true matches, and field protocols that photograph individuals using non-symmetrical markings on left and right sides, can result in a dataset containing multiple encounter histories for an individual (Hiby et al. 2012). In addition, many animals may simply have similar markings. As the number of individuals in a population increases, so too does the difficulty in distinguishing individuals with similar natural markings. The extent to which matching uncertainty biases resulting estimates of population parameters requires investigation for each study. Differences in protocols among individuals and laboratories will likely result in different biases in parameter estimates, as long as protocols require investigators to force an inherently subjective matching process into a binary (match/not-a-match) outcome (Urian et al. 2015). This issue has become increasingly important as long-term cetacean studies have switched from film to digital photography, which may introduce heterogeneity in matches (Urian et al. 2015). Ideally, the level of uncertainty associated with any given match should be quantified and incorporated into resulting population parameter estimates (Urian et al. 2015).

Acknowledging explicitly the uncertainty in the photo matching process, the aim of this study was to use 6 years of photographic data on Pacific white-sided dolphins to quantify the extent to which matching uncertainty affects the bias and precision of abundance and survival estimates. The study also explores the challenges inherent in datasets with relatively low rates of recapture and the effect of matching uncertainty in these cases.

Methods

Study area

The study took place in the waters between northeastern Vancouver Island, British Columbia (BC), Canada and the Broughton Archipelago and Knight Inlet on BC's mainland coast. The study area is characterized by a complex geography of numerous islands, narrow inlets, and fjords (Fig. 1).

Data collection

To ensure consistency in data collection with an existing, long-term catalogue and to maximise the number of long-term resightings, field protocols in the current study

followed those of a previous study as closely as possible (Morton 2000). Photo-identification surveys for Pacific white-sided dolphins in the Broughton Archipelago were conducted from 2008 to 2013. Photo-ID effort was distributed throughout the year but was restricted by weather conditions. Groups of dolphins were found using a combination of boat-based searches and from radio reports and communication from local mariners. Reports from a stationary hydrophone network (OrcaLab) monitored 24 h/day (Morton and Symonds 2002; Deecke et al. 2010), were used to direct dolphin searches. Searches and photographic encounters were limited to sea conditions of a maximum Beaufort scale = 2 for reasons of safety and sightability.

For each encounter, a GPS position and an estimate of group size was made in the field and recorded. Total group size was estimated by tallying the number of individuals in smaller subgroups (typically 2–8 individuals) at intervals throughout the encounter (Morton 2000). A group was defined as all of the dolphins encountered in a discrete location in a day. Finer scale information (e.g., groups defined using a 15 m 'chain rule'; Smolker et al. 1992) on group composition was collected from 2011 onwards to inform studies of sociality. Encounters with dolphins lasted a minimum of 20 min during which the following data were recorded: an estimate of group size (minimum, maximum, and best estimate); location; predominant group activity state (although scan-sample data were collected at 5-min intervals during longer encounters), and number of calves in the group (minimum, maximum, and best estimate). Photographs were collected with digital SLR cameras.

Groups of dolphins were approached slowly in an effort to reduce the probability of bow-riding behaviour, which brings some individuals, especially juveniles, very close to the boat and makes other individuals less available for photographic capture. Large groups were generally traversed in two passes to try to obtain photographs from both sides. In the first pass, individuals were photographed in sub-groups as each sub-group came into photographing range until the far edge of the group was reached. In the second pass, the group was traversed in the opposite direction and at the same angle as the first pass and individuals were photographed in sub-groups in the same manner as the first pass. Dolphins were photographed almost exclusively while engaged in slow, milling (non-directional) behaviour in tight groups (behaviour typically observed following medium to high speed travel). The non-directional/milling behaviour facilitated photographing both right and left sides of the dorsal fin.

Photo-ID efforts ended when dolphins engaged in activity states (e.g., high-speed travel) that resulted in water splashing around the dorsal fin, which results in poor quality photographs. An encounter ended when all of the members of the group had been approached, if weather conditions changed

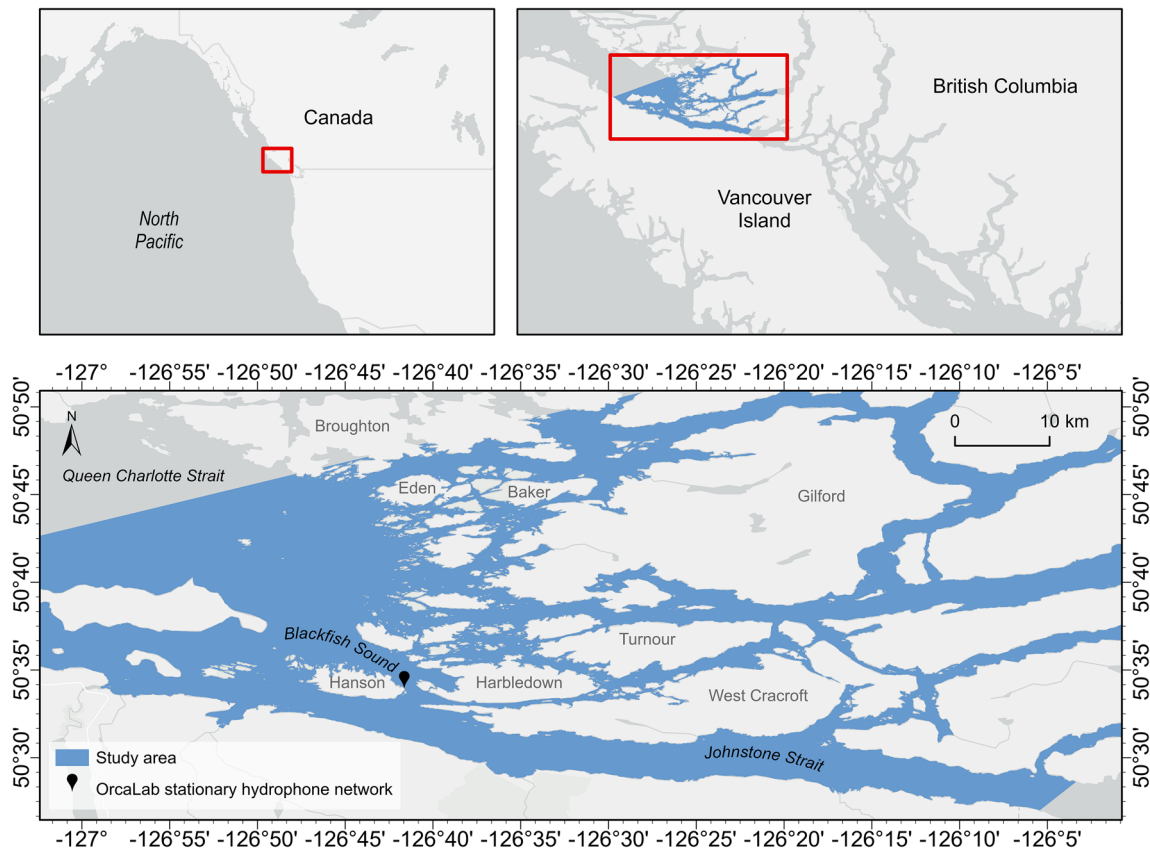


Fig. 1 Map of the study area. The blue polygon represents the study area of the Broughton Archipelago, British Columbia, Canada and adjacent waters

or when the time limit of the research permit (30 min per sub-group) was reached.

Photo-processing methods

All photographs of a dorsal fin were graded for quality of the image and distinctiveness of the markings in two independent stages (Urian et al. 2015). Information on quality, distinctiveness and other attributes were entered into Photo Mechanic 5 (Camera Bits) photo-processing software.¹ First, photographs were graded for photographic quality using a standardized set of photographic quality criteria ranging from 1 (poor quality) to 3 (high quality) following the image quality scoring criteria used in studies of bottlenose dolphins in Scotland (Wilson et al. 1999). Dorsal fins of Pacific white-sided dolphins varied among individuals from extremely well marked with nicks and scars, to completely clean, unmarked fins. Thus, not all dolphins were distinctive enough to be included in mark-recapture analyses. A separate photo reviewer scored each quality 3 photograph

to grade the distinctiveness of each individual. The distinctiveness score ranged from D1 (Highly distinctive) to D4 (Unmarked). A distinctiveness score of D2 (Moderately distinctive) included fins with intermediate features such as a small nick, or many small nicks that are detectable from both sides and D3 (Somewhat distinctive) included fins with subtle features such as black scratches or other long-lasting distinguishing marks that are only identifiable from one side. The D3 score category does not include nicks and notches on the trailing edge of the dorsal fin. A separate set of photo reviewers conducted the matching step (see below).

Photo-matching

A team of six photo reviewers (including EA) conducted the photographic matching of the current study's catalogue in Photo Mechanic 5. The pattern of nicks on the trailing edge of the dorsal fin was the primary mode of identification. Fin shape provided a secondary indicator. Dorsal notches had to match in size, angle of tear and other details. The definition of a match allowed for acquisition of marks over time, but no loss of nicks; that is, if there was an additional notch on the more recent photo,

¹ <https://store.camerabits.com/products/photo-mechanic-version-5>.

but the original nick or notch was present in both photographs, then this was scored as a potential match. Nicks, notches and tears in the fin, along with the shape of the fin itself, are detectable from both sides, so the decision to include these features in the distinctiveness scoring and thresholds meant that both left- and/or right-side photographs could be used to identify individuals.

Only photographs of quality 3 and with a distinctiveness score of D1 or D2 (i.e., symmetrical markings that would be recognized from both sides) were included in the analysis. This protocol was chosen to reduce misidentification errors, while allowing both left- and right-side photographs to be included in the analysis. Photographs of individuals believed to be calves (i.e., small, ruffled dorsal fin, orange colouration, foetal fold marks on the body, photographed alongside mother) were excluded from the analysis. The high-quality subset of photographs was then matched within each photographic encounter, and each individual was assigned a preliminary identification code. Identified individuals within an encounter were matched and a certainty score of “Certain” (100% confident), “Likely” (< 100% but ≥ 90% confident), or “Possible” (< 90% but ≥ 50% confident) was assigned to putative matches between pairs of photographs based on the degree of confidence in each match.

An encounter history of 1 s and 0 s, corresponding to whether a putative individual was or was not detected (i.e., captured) during each sampling encounter, respectively, was created for each individual for each matching certainty level for analysis.

Available data

The number of sampling occasions and the months in which sampling took place each year varied widely throughout this study. Between 2008 and 2013, a total of 34 photographic encounters with dolphins occurred. Of these, 32 encounters contained photographs of sufficient quality and distinctiveness to enter the analysis for the “Certain” and “Likely” certainty levels, whereas all 34 encounters contained photographs that were of sufficient quality to create encounter histories at the “Possible” certainty level. The frequency of capture of individuals for the three certainty levels is shown in Fig. 2. The single encounter from 2008 was not included in the analysis due to low sample size (only 2 individuals were identified).

Estimation of abundance

The encounter histories for each of the three certainty levels were analysed to produce three estimates of abundance. Chapman’s modification to the Lincoln–Petersen

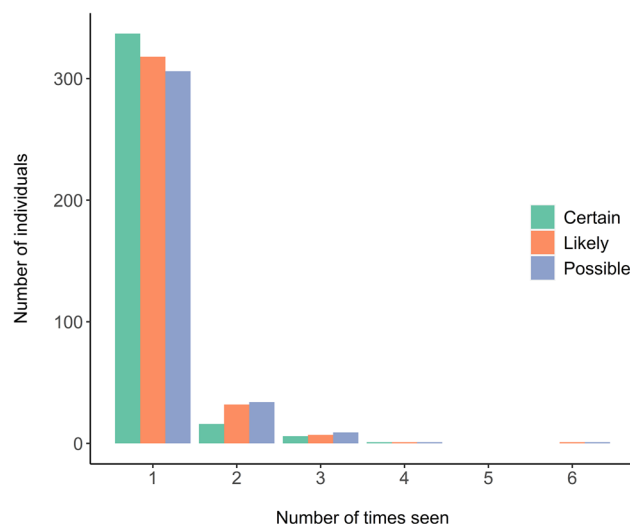


Fig. 2 Number of times an individual was seen, tallied for each certainty level

two-sample estimator to account for small sample bias was used to estimate abundance (Hammond 1986; Seber 2002).

$$\hat{N} = \frac{(n_1 + 1)(n_2 + 1)}{(m_2 + 1)} - 1,$$

where \hat{N} is the abundance estimate; estimate of population size, n_1 is the number of individuals captured during the first sampling occasion, n_2 is the number of individuals captured during the second sampling occasion, m_2 is the number of individuals recaptured. That is, the number of animals captured during the first sampling occasion that were also captured during the second sampling occasion.

For this analysis, each year was treated as a sampling occasion, and recaptures were restricted to individuals seen in adjacent pairs of years. Given the low number of recaptures in adjacent pairs of years and the comparatively large number of photographs taken in 2010, a separate within-year analysis was conducted for 2010 (see “Results”).

Variance was estimated as:

$$\text{var}(\hat{N}) = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m_2)(n_2 - m_2)}{(m_2 + 1)^2(m_2 + 2)}.$$

Log-normal 95% confidence intervals were calculated (Borchers et al. 2002) as \hat{N}/d to $\hat{N}*d$, where

$$d = e^{\left(z_{\alpha} \sqrt{\text{var}[\ln(\hat{N})]}\right)},$$

$z_{0.025} = 1.96$ for a 95% CI.

and

$$\text{var} \left[\ln(\hat{N}) \right] = \ln \left(1 + \frac{\text{var}(\hat{N})}{\hat{N}^2} \right)$$

The simple two-sample estimator used makes a number of assumptions regarding population closure and capture probabilities (Hammond 2018), violation of which can introduce bias in abundance estimates. However, the objective of this exercise was to assess the relative importance of matching uncertainty on the resulting abundance estimates, not to generate a robust abundance estimate for use in decision-making. Any such violations should affect estimates in a comparable way and therefore are unlikely to compromise these results.

Estimating adult annual survival rate

Annual encounter histories from 2008 to 2013 were created for each certainty level to estimate annual apparent survival rate using a Cormack-Jolly-Seber model (Cormack 1964; Jolly 1965; Seber 1965; Amstrup et al. 2005). Models were explored that allowed survival and recapture probability to vary over time or to be constant, and the model with the lowest value of Akaike's Information Criterion (AIC) was selected as that which had the most support from the data. Analysis was carried out in software MARK (White and Burnham 1999) version 6.1. No goodness of fit tests were conducted to test for lack of fit. However, similar to

estimating abundance, the objective of this exercise was to assess the relative importance of matching uncertainty on the resulting estimates of survival, rather than generating the best estimates for wider use, and any assumption violations are unlikely to compromise these results.

Results

The number of individual dolphins photographed and recaptured in each sampling period and at each matching certainty level is shown in Table 1. Including less certain matches resulted in fewer individual dolphin identifications overall, because a more permissive matching threshold will decrease the number of putative individuals in n_1 and n_2 , and increase the number of matches in m_2 . Sampling effort was greatest during 2010 and high in 2011, resulting in the greatest number of recaptures between these years. The lack of recaptures in pairs of years not including 2010 preclude estimation of abundance. Consequently, data from 2012 to 2013 were pooled to boost the number of recaptures with 2011 (Table 1). Compared to other years, 2010 had a substantially higher number of within-year recaptures, so data were also analysed using two sampling occasions within 2010 (2010a: April–June; 2010b: July–November; Table 1).

Two-sample estimates at each matching certainty level produced abundance estimates of individually identifiable dolphins in the population for 2009–2010, 2010–2011, 2011–2012 + 2013 (data pooled for 2012 and 2013), and for time periods within 2010 (Table 2). Abundance of marked dolphins at the “Certain” matching level, ranged from a low of 985 (CV = 0.55) in the “2011–2012 + 2013” sample, to a high of 2005 (CV = 0.31) in 2010 (Table 2). As expected, the “Certain” matching level produced abundance estimates

Table 1 Number of individual dolphins photographed in the two sampling periods (n_1 , n_2), and the number of matches between these (m_2), for each of three matching certainty levels

| Matching certainty level | 2008–2009 | 2009–2010 | 2010–2011 | 2011–2012 | 2012–2013 | 2011–2012 + 2013 | 2010a–2010b |
|------------------------------------|-----------|-----------|-----------|-----------|-----------|------------------|-------------|
| Certain | | | | | | | |
| n_1 | 1 | 37 | 247 | 33 | 2 | 33 | 106 |
| n_2 | 37 | 247 | 33 | 2 | 55 | 57 | 149 |
| m_2 | 0 | 4 | 5 | 0 | 0 | 1 | 7 |
| Certain + Likely | | | | | | | |
| n_1 | 1 | 37 | 245 | 33 | 2 | 33 | 140 |
| n_2 | 37 | 245 | 33 | 2 | 55 | 51 | 244 |
| m_2 | 0 | 4 | 5 | 0 | 0 | 2 | 18 |
| Certain + Likely + Possible | | | | | | | |
| n_1 | 1 | 37 | 239 | 33 | 2 | 33 | 137 |
| n_2 | 37 | 239 | 33 | 2 | 55 | 57 | 241 |
| m_2 | 0 | 5 | 6 | 0 | 0 | 2 | 24 |

Data are summarised for consecutive years in the period 2008–2013, and additionally for 2011 and 2012–2013, and for two time periods within 2010 (2010a: April–June; 2010b: July–November)

Table 2 Two-sample estimates of abundance (\hat{N}) for pairs of years from 2009 to 2011, within 2010 (2010a: April–June; 2010b: July–November), and for 2011 with pooled data for 2012 and 2013

| Matching certainty level | \hat{N} | 95% CI | CV |
|-----------------------------|-----------|--------------|------|
| 2009–2010 | | | |
| Certain | 1884 | (922, 3848) | 0.38 |
| Certain + Likely | 1869 | (915, 3816) | 0.38 |
| Certain + Likely + Possible | 1513 | (787, 2907) | 0.34 |
| 2010–2011 | | | |
| Certain | 1404 | (736, 2680) | 0.34 |
| Certain + Likely | 1393 | (730, 2659) | 0.34 |
| Certain + Likely + Possible | 1159 | (640, 2103) | 0.31 |
| 2010a–2010b | | | |
| Certain | 2005 | (1103, 3645) | 0.31 |
| Certain + Likely | 1817 | (1233, 2678) | 0.20 |
| Certain + Likely + Possible | 1335 | (962, 1852) | 0.17 |
| 2011–2012 + 2013 | | | |
| Certain | 985 | (359, 2701) | 0.55 |
| Certain + Likely | 656 | (276, 1564) | 0.47 |
| Certain + Likely + Possible | 656 | (275, 1564) | 0.47 |

Table 3 Apparent survival rate estimates derived from three different levels of certainty in photographic matching from 2008 to 2013

| Matching certainty level | Survival | SE | 95% CI |
|-----------------------------|----------|-------|----------------|
| Certain | 0.458 | 0.288 | (0.183, 1.000) |
| Certain + Likely | 0.460 | 0.287 | (0.184, 1.000) |
| Certain + Likely + Possible | 0.468 | 0.276 | (0.191, 1.000) |

that were greater than the estimates at the “Certain + Likely” or “Certain + Likely + Possible” levels in the same pairs of samples (Table 2). In 2010–2011, for example, abundance was estimated as 1404 (CV = 0.34) at the highest level of certainty (“Certain”) and at 1159 (CV = 0.31) at the “Certain + Likely + Possible” matching level. The greater the matching certainty level, the lower the precision of the abundance estimates (Table 2).

Estimates of annual survival rate

Annual apparent survival rate for well-marked adult dolphins was estimated using a CJS model at each matching certainty level for the period 2008–2013. The model for constant survival and time-varying recapture probability had the best support from the data at all certainty levels. Annual survival for the “Certain” matching level was estimated as 0.458 (SE = 0.288) rising slightly to 0.468 (SE = 0.276) for the “Certain + Likely + Possible” matching level (Table 3). Thus, as less certain matches were included in the analyses,

the apparent survival estimates increased slightly and the precision of the estimates also increased slightly. The very low estimates of apparent survival are likely a result of emigration out of the study area over the period of the study.

Discussion

The study demonstrates that matching uncertainty has the potential to introduce substantial degrees of both bias and uncertainty in a real-world photo-identification study of a dolphin species with a low capture probability. The extent to which this translates into conservation or management risk hinges on the extent to which the less-than-certain matches are actually true matches rather than false positives. If all the less-than-certain matches are false positives, then using the highest certainty level as a threshold to define a match will give the least biased abundance estimate. But if all or some of the less-than-certain matches are actually true matches, then using the highest certainty level as the threshold to define a match means that the abundance estimates are positively biased.

This pattern shows up in our abundance estimates (Table 2) and, with a smaller effect, in our survival estimates (Table 3). Abundance estimates were found to vary among years and matching certainty levels. Because of inter-annual variation in effort and the low rate of recapture in 2009 and 2011–2013, it is most informative to focus on a comparison of within-year estimates from 2010. Within 2010, the abundance estimates ranged from 2005 (95% CI 1103–3645) for “Certain” matches to 1335 (95% CI 962–1852) for “Certain + Likely + Possible” matches. The “Possible” and “Likely” categories may contain false positive matches, which will cause negative bias in abundance estimates (Yoshizaki et al. 2009). However, false positive errors typically arise from inclusion of poor-quality photographs (Stevick et al. 2001; Friday et al. 2008; Barlow et al. 2011) and only the highest quality photographs were included in this analysis.

Notably, there is still much variation in abundance estimates depending on matching certainty level, despite restricting our analyses to photographs of the highest quality. While restricting analysis to the best quality photographs of marked individuals, there may still be substantial false positive errors in species in which individuals may share similar marks. As the number of individuals in the population increases, so too will the probability of seeing two dolphins with very similar markings on their dorsal fins. As more tenuous matches were categorised as recaptures in the analyses, the abundance estimates decreased (i.e., assumed to become negatively biased if the lower certainty level matches were not true matches) and the apparent precision increased. But if some of these less than certain matches actually were

matches, the estimates from the “Certain + Likely” and “Certain + Likely + Possible” scenarios are less biased than the “Certain” scenarios, and the abundance estimate from the “Certain” scenarios were biased high. Although the number of recaptures in this study is low, this effect was most obvious in the within-2010 analysis, when effort was highest and there were a substantial number of matches. In terms of survival, estimates of survival rate and their precision increased only slightly as less certain matches were categorized as a match (Table 3).

Although the direction of these trends can be predicted from first principles, the magnitude of the effect found in our study was unexpectedly high, at least for abundance. The “Certain” abundance estimates were ~50% higher than estimates derived from “Certain + Likely + Possible” matches in the “Within 2010” and the “2011–2012 + 2013” scenarios (Table 2). The “Certain” abundance estimates were 25 and 21% higher than estimates derived from “Certain + Likely + Possible” matches in the “2009–2010” and “2010–2011” scenarios, respectively (Table 2). While reliance on absolute certainty is often described as a “conservative” and “recommended” feature of any photo matching protocol (e.g., Friday et al. 2008), our results suggest that use of an overly conservative threshold to define a match could result in substantial bias (20–50%) in abundance estimates.

The trade-off illustrated in this study exemplifies the need for researchers to decide whether it is better to include uncertain matches to increase sample size to get a more precise (but potentially biased) estimate, or to prioritise accuracy over precision. This decision is inherently case-specific. If monitoring for overall trends in abundance (Wilson et al. 1999; Gerrodette and Forcada 2005), precision may be paramount and bias less of a concern, as long as the bias remains consistent over the period of interest (Taylor and Gerrodette 1993; Taylor et al. 2007; Williams et al. 2016a).

Pacific white-sided dolphins are known to be caught in salmon gillnet fisheries in British Columbia, and our ability to assess the sustainability of that bycatch hinges on improving the accuracy and precision of both the estimates of dolphin abundance and bycatch rate (Williams et al. 2008). A new trade rule requires countries wishing to export seafood to US markets to demonstrate that their management schemes are comparable in effectiveness to those under the US Marine Mammal Protection Act (Williams et al. 2016b). We anticipate investments in filling data gaps in many understudied species and regions to facilitate compliance with this new rule (Ashe et al. 2021a,b; Hammond et al. 2021; Punt et al. 2021). Given the potential positive bias in abundance estimates using overly strict matching criteria, assuming that some likely or possible matches were actually true matches (Table 2), it will be useful to investigate the potential for matching uncertainty to bias abundance estimates in new studies, because positively biased abundance estimates could

lead to overestimates of allowable harm limit for assessing sustainability of bycatch (Wade 1998; Punt et al. 2020).

When the trade-off makes the difference between having a biased estimate versus no estimate at all, then it is better to report an estimate, as long as the sources of bias are acknowledged. One could potentially quantify that bias using an approach such as the one presented here. Selecting an acceptable trade-off between bias and precision may be more challenging to consider with small sample sizes.

It is a well-known problem that traditional mark-recapture methods are sensitive to misidentification of animals that are recognised from natural markings (Link et al. 2010). Previous studies have considered effects of photo-quality and animal distinctiveness on bias and precision in abundance and survivorship estimates (Friday et al. 2000; Stevick et al. 2001). Although matching certainty is clearly confounded with photographic quality and animal distinctiveness, the current study examined matching uncertainty with a dataset of high-quality photographs of well-marked individual dolphins. The aim was to evaluate the effect on estimates of population parameters caused by uncertainty in individual identification. Results showed that while this issue had relatively modest impact on estimates of apparent survival, abundance estimates could vary by 20–50% as a result of this source of uncertainty (Table 2). It is difficult to evaluate the effect of matching uncertainty on precision of estimates, however, due to low number of recaptures in our study. Nevertheless, it is worth noting that in the within-2010 abundance estimates, which had the largest number of recaptures, there is a big difference in CV among certainty levels. The next problem, of course, is how to resolve this.

This study has looked at misidentification by examining the impact on abundance and survival estimates that arises at the processing level. Conventionally, photo reviewers processing ID photographs are instructed to assert that two photographs are or are not a match. Depending on the protocols used by a given research team, most researchers will be averse to false positives and will default to a non-match in the case of <100% certainty, whereas others may be equally averse to false negatives (Urian et al. 2015). Nonetheless, the conventional mark-recapture models require a binary decision to be made. The current study shows that there is value to having a number of photo reviewers, matchers, and experienced researchers record their level of certainty that two photographs represent a match, because there is useful information contained in that matching certainty level (Urian et al. 2015). The size of effect due to matching uncertainty found here, as high as a third on abundance, may not be typical, but the approach could easily be incorporated in other photo-ID studies (Urian et al. 2015). In many cases, the effect on estimates of abundance or survival of matching uncertainty may be negligible, but it will be impossible to know this unless matching protocols instruct matchers

to record their level of confidence in a match so that this information can be used at the analysis stage.

Our study confirms much what has been said in other studies of misidentification in mark-recapture arising from issues of data quality, and the advice for coping with the source of misidentification. Mark misidentification can arise when the sampling (e.g., photo-ID) introduces heterogeneity in the observer's ability to recognise marks. One study that identified a minimum quality level when including photographic data in a mark-recapture analysis using two kinds of tags (i.e., photo-ID and genetics) placed bounds on the uncertainty and incorporated that uncertainty in a bootstrap estimate of the variance around the abundance estimate (Stevick et al. 2001). Conceptually, the recommendation of Stevick et al. (2001) could apply here as well: encounter histories corresponding to the three matching certainty levels could be resampled via a bootstrap, to incorporate this source of uncertainty in estimates of abundance, survival and their variances. This will be more pragmatic in the short term than the suggestion to use genetic double-tagging to minimise or avoid misidentification in the first place (Lukacs and Burnham 2005; Link et al. 2010). For some cetacean species, it may be possible to use natural markings on two morphological features as another way to investigate misidentification, e.g., northern bottlenose whales (*Hyperoodon ampullatus*, Gowans and Whitehead 2001) and bottlenose dolphins (*Tursiops truncatus*, Genov et al. 2018).

Misidentification can also arise when marks change through time. With a sufficiently large number of individuals followed through time, it may be possible to build mechanistic models to understand how marks evolve. Quantifying this effect could then allow to account for misidentification in the resulting demographic parameters. Simulation studies suggest that these mechanistic models will only work when capture probability is higher than that observed during this study (i.e., > 0.2) and when the absolute number of resightings is sufficiently large to have enough data to estimate demographic parameters and changes in marks simultaneously (Yoshizaki et al. 2009).

Statistically, the process of misidentification that this study discusses is quite challenging to address using traditional likelihood methods, but could be handled in a straightforward way using Bayesian methods (Link et al. 2010). Bayesian mark-recapture methods have been an active research area for some time (Schofield et al. 2009). As Bayesian survival estimators become more commonly used and associated code made accessible to population ecologists (Gimenez et al. 2009), addressing misidentification in a Bayesian framework may be the next logical step. By integrating all known sources of uncertainty into a single analytical framework, Bayesian methods offer the potential to assess whether misidentification (or any other source of uncertainty) has the potential to affect what is perhaps the

most important question of all—does any of this uncertainty affect the ultimate category of risk to which we assign a species (Brooks et al. 2008)?

Appendix



Fig. A1 Two Pacific white-sided dolphins (*Lagenorhynchus obliquidens*) leap during a bout of foraging activity near the Broughton Archipelago, British Columbia, Canada. (Photo: E. Ashe, under research permit)

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42991-022-00236-4>.

Acknowledgements The authors wish to thank Alexandra Morton and Rob Williams for advice and logistical support that facilitated Erin's Ph.D. thesis, in which an earlier version of this paper formed a chapter. We thank Paul Spong and Helena Symonds (OrcaLab) for sharing sightings, and Don Willson for field support. This study involved processing large numbers of photographs for image quality, animal distinctiveness, and potential matches. We thank Melissa Boogaards, Nicole Koshure, Leila Fouda, Marie Fournier, Emily Hague, Christie McMillan, and Alyssa Rice for help with that effort. Thanks also to the SeaDoc Society, Molly and John Bailey, the Richardson family, Sarah Haney (Canadian Whale Institute), National Geographic, and Beto Bedolfe at the Marisla Foundation for financial support over the years, and to The Willow Grove Foundation for supporting the Knight Inlet expedition that made 2010 so productive. Erin thanks Air Canada's Aeroplan Beyond Miles program for travel support, and Prof. Sascha Hooker and Dr. Jaume Forcada for examining her PhD. Their feedback, along with advice from the editors, Kimberly Nielsen, and two anonymous reviewers, improved this manuscript. Erin was a beneficiary of a writing retreat for women in science supported by Lyda Hill Philanthropies and the National Geographic Explorers Program. This paper represents the third thesis chapter to be published with that support for travel, accommodation, and childcare.

Author contributions EA: conceptualization, data collection, data analysis and writing—original draft preparation. PH: supervision, writing—reviewing and editing.

Declarations

Conflict of interest On behalf of all authors, the corresponding author state that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agrelo M, Daura-Jorge FG, Rowntree VJ, Sironi M, Hammond PS, Ingram SN, Marón CF, Vilches FO, Payne R, Simões-Lopes PC (2021) Ocean warming threatens southern right whale population recovery. *Sci Adv* 7:ebh2823. <https://doi.org/10.1126/sciadv.abh2823>
- Amstrup CS, McDonald TL, Manly BFJ (2005) Handbook of capture-recapture analysis. Princeton University Press, Princeton
- Arnborn T (1987) Individual identification of sperm whales. *Rep Int Whal Commn* 37:201–204
- Arso Civil M, Cheney B, Quick NJ, Thompson PM, Hammond PS (2017) A new approach to estimate fecundity rate from inter-birth intervals. *Ecosphere* 8:e01796. <https://doi.org/10.1002/ecs2.1796>
- Arso Civil M, Quick NJ, Cheney B, Islas-Villanueva V, Graves JA, Janik V, Thompson PM, Hammond PS (2019) Variations in age- and sex-specific survival rates could explain population trend in a discrete marine mammal population. *Ecol Evol* 9:533–544. <https://doi.org/10.1002/ece3.4772>
- Ashe E, Williams R, Clark C, Erbe C, Gerber LR, Hall AJ, Hammond PS, Lacy RC, Reeves R, Vollmer NL (2021a) Minding the data-gap trap: exploring dynamics of abundant dolphin populations under uncertainty. *Front Mar Sci* 8:606932. <https://doi.org/10.3389/fmars.2021.606932>
- Ashe E, Williams R, Morton A, Hammond PS (2021b) Disentangling natural and anthropogenic forms of mortality and serious injury in a poorly studied pelagic dolphin. *Front Mar Sci* 8:606876. <https://doi.org/10.3389/fmars.2021.606876>
- Barlow J, Clapham PJ (1997) A new birth-interval approach to estimating demographic parameters of humpback whales. *Ecology* 78(2):535–546. [https://doi.org/10.1890/0012-9658\(1997\)078\[0535:ANBIAT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[0535:ANBIAT]2.0.CO;2)
- Barlow J, Calambokidis J, Falcone EA, Baker CS, Burdin AM, Clapham PJ, Ford JKB, Gabriele CM, LeDuc R, Mattila DK, Quinn TJ, Rojas-Bracho L, Straley JM, Taylor BL, Urbán RJ, Wade P, Weller D, Witteveen BH, Yamaguchi M (2011) Humpback whale abundance in the North Pacific estimated by photographic capture-recapture with bias correction from simulation studies. *Mar Mammal Sci* 27:793–818. <https://doi.org/10.1111/j.1748-7692.2010.00444.x>
- Bigg M (1982) An assessment of killer whale (*Orcinus orca*) stocks off Vancouver Island, British Columbia. *Rep Int Whal Commn* 32:655–666
- Borchers DL, Buckland ST, Zucchini W (2002) Estimating animal abundance: closed populations. Springer Verlag, London
- Brooks SP, Freeman SN, Greenwood JJ, King R, Mazzetta C (2008) Quantifying conservation concern-Bayesian statistics, birds and the red lists. *Biol Conserv* 141:1436–1441. <https://doi.org/10.1016/j.biocon.2008.03.009>
- Caldwell DK (1955) Evidence of home range of an Atlantic bottlenose dolphin. *J Mammal* 36:304–305. <https://doi.org/10.2307/1375913>
- Cormack RM (1964) Estimates of survival from the sighting of marked animals. *Biometrika* 51:429–438. <https://doi.org/10.2307/2334149>
- Coxon J, Arso Civil M, Claridge D, Dunn C, Hammond PS (2022) Investigating local population dynamics of bottlenose dolphins in the northern Bahamas and the impact of hurricanes on survival. In: Karczmarski L, Chan SCY, Rubenstein DI, Chui SYS, Cameron EZ (eds) Individual identification and photographic techniques in mammalian ecological and behavioural research - Part 2: Field studies and applications. *Mamm Biol (Special Issue)*, 102(3). <https://doi.org/10.1007/s42991-021-00208-0>
- Deecke VB, Barrett-Lennard LG, Spong P, Ford JKB (2010) The structure of stereotyped calls reflects kinship and social affiliation in resident killer whales (*Orcinus orca*). *Naturwissenschaften* 97:513–518. <https://doi.org/10.1007/s00114-010-0657-z>
- Dufault S, Whitehead H (1995) An assessment of changes with time in the marking patterns used for photoidentification of individual sperm whales, *Physeter macrocephalus*. *Mar Mammal Sci* 11:335–343. <https://doi.org/10.1111/j.1748-7692.1995.tb00289.x>
- Friday N, Smith TD, Stevick PT, Allen J (2000) Measurement of photographic quality and individual distinctiveness for the photographic identification of humpback whales, *Megaptera novaeangliae*. *Mar Mammal Sci* 16:355–374. <https://doi.org/10.1111/j.1748-7692.2000.tb00930.x>
- Friday NA, Smith TD, Stevick PT, Allen J, Fernald T (2008) Balancing bias and precision in capture-recapture estimates of abundance. *Mar Mammal Sci* 24:253–275. <https://doi.org/10.1111/j.1748-7692.2008.00187.x>
- Genov T, Centrih T, Wright AJ, Wu GM (2018) Novel method for identifying individual cetaceans using facial features and symmetry: a test case using dolphins. *Mar Mam Sci* 34:514–528. <https://doi.org/10.1111/mms.12451>
- Genu M, Gilles A, Hammond PS, Macleod K, Paillé J, Paradinas I, Smout S, Winship A, Authier M (2021) Evaluating strategies for managing anthropogenic mortality on marine mammals: an R implementation with the package RLA. *Front Mar Sci* 8:795953. <https://doi.org/10.3389/fmars.2021.795953>
- Gerrodette T, Forcada J (2005) Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. *Mar Ecol Prog Ser* 291:1–21. <https://doi.org/10.3354/meps291001>
- Gimenez O, Bonner SJ, King R, Parker RA, Brooks SP, Jamieson LE, Grosbois V, Morgan BJT, Thomas L (2009) WinBUGS for population ecologists: Bayesian modeling using Markov chain Monte Carlo methods. In: Thomson DL, Cooch EG, Conroy MJ (eds) Modeling demographic processes in marked populations. *Environ Ecol Stat*, vol 3. Springer, Boston, pp 883–915. https://doi.org/10.1007/978-0-387-78151-8_41
- Gordon JCD (1987) Behaviour and ecology of sperm whales off Sri Lanka. PhD Dissertation, University of Cambridge.
- Gowans S, Whitehead H (2001) Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): sources of heterogeneity from natural marks. *Mar Mammal Sci* 17:76–93. <https://doi.org/10.1111/j.1748-7692.2001.tb00981.x>
- Hammond PS (2018) Mark-recapture. In: Würsig B, Thewissen JGM, Kovacs K (eds) Encyclopedia of marine mammals, 3rd edn. Academic Press/Elsevier, San Diego, pp 580–584

- 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. *Front Mar Sci* 8:735770. <https://doi.org/10.3389/fmars.2021.735770>
- Hammond PS, Mizroch SA, Donovan GP (1990). Individual recognition of cetaceans: use of photo-identification and other techniques to estimate population parameters. *Rep Int Whal Commn (Spec Iss 12)*, IWC, Cambridge
- Hammond PS (1986) Estimating the size of naturally marked whale populations using capture-recapture techniques. *Rep Int Whal Commn 1986(Special Issue 8)*:253–282
- Hiby L, Paterson WD, Redman P, Watkins J, Twiss SD, Pomeroy P (2012) Analysis of photo-id data allowing for missed matches and individuals identified from opposite sides. *Methods Ecol Evol* 4:252–259. <https://doi.org/10.1111/2041-210x.12008>
- Jolly GM (1965) Explicit estimates from capture-recapture data with both death and immigration—a stochastic model. *Biometrika* 52:225–247. <https://doi.org/10.2307/2333826>
- Kuningas S, Similä T, Hammond PS (2014) Population size, survival and reproductive rates of northern Norwegian killer whales (*Orcinus orca*) in 1986–2003. *J Mar Biol Assoc UK* 94:1277–1291. <https://doi.org/10.1017/S0025315413000933>
- Langtimm CA, Beck CA, Edwards HH, Fick-Child KJ, Ackerman BB, Barton SL, Hartley WC (2004) Survival estimate for Florida manatees from the photo-identification of individuals. *Mar Mammal Sci* 20:438–463. <https://doi.org/10.1111/j.1748-7692.2004.tb01171.x>
- Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol Monogr* 62:67–118. <https://doi.org/10.2307/2937171>
- Link WA, Yoshizaki J, Bailey LL, Pollock KH (2010) Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. *Biometrics* 66:178–185. <https://doi.org/10.1111/j.1541-0420.2009.01244.x>
- Lukacs PM, Burnham KP (2005) Review of capture-recapture methods applicable to noninvasive genetic sampling. *Mol Ecol* 14:3909–3919. <https://doi.org/10.1111/j.1365-294X.2005.02717.x>
- Morton A (2000) Occurrence, photo-identification and prey of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*) in the Broughton Archipelago, Canada 1984–1998. *Mar Mammal Sci* 16:80–93. <https://doi.org/10.1111/j.1748-7692.2000.tb00905.x>
- Morton AB, Symonds HK (2002) Displacement of *Orcinus orca* (L.) by high amplitude sound in British Columbia. *Canada ICES J Mar Sci* 59:71–80. <https://doi.org/10.1006/jmsc.2001.1136>
- Pace RM, Corkeron PJ, Kraus SD (2017) State-space mark-recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecol Evol* 7:8730–8741. <https://doi.org/10.1002/ece3.3406>
- Payne R, Brazier O, Dorsey EM, Perkins JS, Rowntree VJ, Titus A (1983) External features in southern right whales (*Eubalaena australis*) and their use in identifying individuals. In: Payne R (ed) *Communication and behavior of whales*. AAAS selected symposium 76. Westview Press, Colorado, pp 371–445
- Punt AE, Siple M, Francis TB, Hammond PS, Heinemann D, Long KJ, Moore JE, Sepúlveda M, Reeves RR, Sigurðsson GM, Víkingsson G, Wade PR, Williams R, Zerbini AN (2020) Robustness of potential biological removal to monitoring, environmental, and management uncertainties. *ICES J Mar Sci* 77:2491–2507. <https://doi.org/10.1093/icesjms/fcaa096>
- Punt AE, Siple MC, Francis TB, Hammond PS, Heinemann D, Long KJ, Moore J, Sepúlveda M, Reeves RR, Sigurðsson GM, Víkingsson G, Wade PR, Williams R, Zerbini AN (2021) Can we manage marine mammal bycatch effectively in low-data environments? *J Appl Ecol* 58:596–607. <https://doi.org/10.1111/1365-2664.13816>
- Ramp C, Bérubé M, Hagen W, Sears R (2006) Survival of adult blue whales *Balaenoptera musculus* in the Gulf of St. Lawrence. *Canada Mar Ecol Prog Ser* 319:287–295. <https://doi.org/10.3354/meps319287>
- Ramp C, Delarue J, Bérubé M, Hammond PS, Sears R (2014) Fin whale survival and abundance in the Gulf of St. Lawrence, Canada. *Endanger Species Res* 23:125–132. <https://doi.org/10.3354/esr00571>
- Read AJ, Urian KW, Wilson B, Waples DM (2003) Abundance of bottlenose dolphins in the bays, sounds, and estuaries of North Carolina. *Mar Mammal Sci* 19:59–73. <https://doi.org/10.1111/j.1748-7692.2003.tb01092.x>
- Schofield MR, Barker RJ, MacKenzie DI (2009) Flexible hierarchical mark-recapture modeling for open populations using WinBUGS. *Environ Ecol Stat* 16:369–387. <https://doi.org/10.1007/s10651-007-0069-1>
- Seber GAF (1965) A note on multiple recapture census. *Biometrika* 52:249–259. <https://doi.org/10.2307/2333827>
- Seber GAF (2002) *The estimation of animal abundance and related parameters*. Blackburn Press, Blackburn
- Smolker RA, Richards AF, Connor RC, Pepper JW (1992) Sex differences in patterns of association among Indian Ocean bottlenose dolphins. *Behaviour* 123:38–69
- Steiger GH, Calambokidis J, Straley JM, Herman LM, Cerchio S, Salden DR, Urbán-R J, Jacobsen JK, von Ziegesar O, Balcomb KC, Gabriele CM, Dahlheim ME, Uchida S, Ford JKB, Ladrón de Guevara-P P, Yamaguchi M, Barlow J (2008) Geographic variation in killer whale attacks on humpback whales in the North Pacific: implications for predation pressure. *Endanger Species Res* 4:247–256. <https://doi.org/10.3354/esr00078>
- Stevick P, Palsboll PJ, Smith TD, Bravington MV, Hammond PS (2001) Errors in identification using natural markings: rates, sources, and effects on capture-recapture estimates of abundance. *Can J Fish Aquat Sci* 58:1861–1870. <https://doi.org/10.1139/f01-131>
- Stevick PT, Allen J, Clapham PJ, Friday N, Katona SK, Larsen F, Lien J, Mattila DK, Palsbøll PJ, Sigurjónsson J, Smith TD, Øien N, Hammond PS (2003) North Atlantic humpback whale abundance and rate of increase four decades after protection from whaling. *Mar Ecol Prog Ser* 258:263–273. <https://doi.org/10.3354/meps258263>
- Taylor BL, Gerrodette T (1993) The uses of statistical power in conservation biology: the vaquita and northern spotted owl. *Conserv Biol* 7:489–500. <https://doi.org/10.1046/j.1523-1739.1993.07030489.x>
- Taylor BL, Martinez M, Gerrodette T, Barlow J, Hrovat YN (2007) Lessons from monitoring trends in abundance of marine mammals. *Mar Mammal Sci* 23:157–175. <https://doi.org/10.1111/j.1748-7692.2006.00092.x>
- Thompson PM, Tollit D, Wood D, Corpe HA, Hammond PS, MacKay A (1997) Estimating harbour seal abundance and status in an estuarine habitat in north-east Scotland. *J Appl Ecol* 34:43–52. <https://doi.org/10.2307/2404846>
- Urian K, Gorgone A, Read A, Balmer B, Wells RS, Berggren P, Durban J, Eguchi T, Rayment W, Hammond PS (2015) Recommendations for photo-identification methods used in capture-recapture models with cetaceans. *Mar Mammal Sci* 31:298–321. <https://doi.org/10.1111/mms.12141>
- Wade PR (1998) Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar Mammal Sci* 14:1–37. <https://doi.org/10.1111/j.1748-7692.1998.tb00688.x>
- Wells RS, Scott MD (1990) Estimating bottlenose dolphin population parameters from individual identification and capture-recapture techniques. *Rep Int Whal Commn 1990(Special Issue 12)*:407–415.

- White GC, Burnham KP (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study* 46:120–139. <https://doi.org/10.1080/00063659909477239>
- Williams R, Hall A, Winship A (2008) Potential limits to anthropogenic mortality of small cetaceans in coastal waters of British Columbia. *Can J Fish Aquat Sci* 65:1867–1878. <https://doi.org/10.1139/F08-098>
- Williams R, Burgess MG, Ashe E, Gaines SD, Reeves RR (2016a) US seafood import restriction presents opportunity and risk. *Science* 354:1372–1374. <https://doi.org/10.1126/science.aai8222>
- Williams R, Moore JE, Gomez-Salazar C, Trujillo F, Burt L (2016b) Searching for trends in river dolphin abundance: designing surveys for looming threats, and evidence for opposing trends of two species in the Colombian Amazon. *Biol Conserv* 195:136–145. <https://doi.org/10.1016/j.biocon.2015.12.037>
- Wilson B, Hammond P, Thompson PM (1999) Estimating size and assessing trends in a coastal bottlenose dolphin population. *Ecol Appl* 9:288–300. [https://doi.org/10.1890/1051-0761\(1999\)009\[0288:ESAATI\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1999)009[0288:ESAATI]2.0.CO;2)
- Winship AJ, Berggren P, Hammond PS (2006) Management procedures, simulation testing, and data sources for determining appropriate limits to the bycatch of small cetaceans in the European Atlantic and North Sea. Paper SC/58/SM16 presented to the IWC Scientific Committee. IWC, Cambridge.
- Yoshizaki J, Pollock KH, Brownie C, Webster RA (2009) Modeling misidentification errors in capture-recapture studies using photographic identification of evolving marks. *Ecology* 90:3–9. <https://doi.org/10.1890/08-0304.1>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.