ICES Journal of Marine Science, 2024, Vol. 10, Issue 0, 1–10
https://doi.org/10.1093/icesjms/fsad202
Received: 31 August 2023; **revised:** 28 November 2023; **accepted:** 29 November 2023
Original Article

ICES CIEM
International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

Impact of fishery observer protocol on estimated bycatch rates of marine mammals

Kristin Precoda[1] and Christopher D. Orphanides[2,*]

[1]Integrated Statistics, 16 Sumner St., Woods Hole, MA 02543, United States
[2]NOAA Fisheries, Northeast Fisheries Science Center, 28 Tarzwell Dr, Narragansett, RI 02882, United States
*Corresponding author. NOAA Fisheries, Northeast Fisheries Science Center, 28 Tarzwell Dr, Narragansett, RI 02882, USA. E-mail: chris.orphanides@noaa.gov

Abstract

Managing fishing operations' threat to marine mammal populations hinges on accurate bycatch estimates, often derived from fishery observer or monitoring programmes. Much global marine mammal bycatch occurs in gillnets, and observer protocols that do not include watching the haulback of gillnets may miss animals that drop out of the net. We investigated whether trips using a fish-focused observer protocol (no requirement to watch the haulback) in US northwestern Atlantic gillnet fisheries from 1994 to 2019 had different observed bycatch rates from trips under a mammal-focused observer protocol (watching the haulbacks) for grey seals (*Halichoerus grypus atlantica*), harbour seals (*Phoca vitulina vitulina*), and harbour porpoise (*Phocoena phocoena phocoena*). We found that observer protocol was likely to affect observed drop-out and bycatch rates. Under the fish-focused protocol, the ratio of animals removed from the net to those that fell from the net was generally higher than under the mammal-focused protocol, suggesting fish-focused observers missed bycatch that fell. Bycatch rates of animals removed from the net by fishers differed significantly between observer protocols for seals, but not for harbour porpoise, perhaps because of differences in entanglement and manner of decomposition. We estimate bycatch was underreported by 3–25% because of unobserved drop-outs on fish-focused observer protocols.

Keywords: bycatch; fishery observer programme; bias; gillnet; marine mammals; bycatch estimation

Introduction

Fisheries bycatch, i.e. the incidental catch of unintended species, is a significant threat to marine mammals worldwide (Read et al. 2006, Read 2008, Avila et al. 2018), with gillnets capturing the majority of marine mammal bycatch (Read 2008, Reeves et al. 2013). National and international regulations such as the US Marine Mammal Protection Act and the EU Marine Strategy Framework Directive (2008/56/EC, EC 2008) are designed to limit incidental takes of marine mammals, and increased attention to mammal bycatch is reflected in the Marine Stewardship Council Fisheries Standard 3.0 (Marine Stewardship Council 2022) and the Fish and Fish Product Import Provisions of the US Marine Mammal Protection Act. However, the ability of governments to estimate and effectively limit bycatch varies (Williams et al. 2016).

Precise and unbiased estimates of injured or dead bycatch allow for better management of marine mammal populations through improved understanding of fishing-related population impacts (Wade et al. 2021), and can also reduce unnecessary burden on human use of the ocean, e.g. in the form of precautionary regulation of fishing activities in light of uncertain impacts. Yet many areas of the world lack robust observer programmes to calculate precise and unbiased estimates of bycatch (Gilman et al. 2012). To support the development of such programmes, Moore et al. (2021) provided an overview of concepts and best management practices for estimating marine mammal bycatch.

One phenomenon that could affect the accuracy of bycatch estimates is 'drop-outs' or 'drop-offs', or animals caught in a fishing net that drop from it during retrieval of the net. Fisheries observers may not notice drop-outs if their duties and space requirements prevent them from watching the retrieval of the net, as can occur when fisheries observer programmes have goals other than observing and sampling protected species bycatch. In the US northwestern Atlantic, two protocols for observing a gillnet trip exist, one focused on documenting fish catches and the other focused on observing marine mammal bycatch (NEFOP 2016). Australia has used similar gillnet protocols for observing either fish catch or sea lion bycatch (Australian Fisheries Management Authority 2010, Goldsworthy et al. 2010, p. 188).

Indications that missed drop-outs could be an issue appeared when harbour porpoise bycatch was investigated early in the US northwestern Atlantic observing programme. One study reported that of 36 bycaught harbour porpoise, 15 reached the deck of the fishing vessel, and 21 fell out before reaching the deck (Bravington and Bisack 1996). Similarly, significant fractions of drop-outs were observed in work on harbour porpoise and common dolphin (*Delphinus delphis*) bycatch in the Celtic Sea (Tregenza et al. 1997a, b). Ten of twelve bycaught sea lions (*Neophoca cinerea*) observed in an Australian gillnet fishery dropped out when hauled above water, and sea lion bycatch estimates based on fish-focused observer data were therefore assumed to be underestimates (Hamer et al. 2013). The review of best practices by Moore et al. (2021) noted the issue of drop-outs and discussed using marine mammal strandings as a measure of unobserved bycatch events. However, the review did not propose a direct means of estimating the number of drop-outs or a way to deal with them during bycatch estimation.

In this study, we explored drop-outs and their potential effect on bycatch estimates using observations of marine mammal bycatch in gillnets on trips completed by US Northeast Fisheries Science Center observer programmes under either protocol used in the US northwestern Atlantic. Both observer protocols require all observed marine mammal bycatch to be recorded. One protocol is focused on kept and discarded catch ('fish-focused'), and the other focuses on marine mammal bycatch and requires the observer to watch the haulback of the net ('mammal-focused'). Both mammal-focused and fish-focused observations are used for marine mammal bycatch estimation to maximize the available data.

The goals of this study were to determine whether the observer protocols affect observed bycatch rates of grey seals (*Halichoerus grypus atlantica*), harbour seals (*Phoca vitulina vitulina*), and harbour porpoise (*Phocoena phocoena phocoena*) and, if so, whether the number of unobserved drop-outs can be estimated. To support a detailed investigation of the primary question, we also examined the relationship between net soak duration and decomposition of bycaught animals. Some animals likely disappear underwater before being observed (Hamer et al. 2013), and scavenging, decomposition, or the end of rigor mortis during longer soaks may allow a bycaught animal to fall from the net during haulback. Therefore, to illuminate possible differences between observer protocols and facilitate their comparison, we examined whether more decomposed animals were more likely to fall out of the net during haulback, and whether animals that were removed from the net by the crew—rather than falling out—were observed at the same rate under either observer protocol.

Methods

We performed several steps to assess the possible impact of observer protocols on seal and harbour porpoise bycatch rates. After compiling the observer data, we investigated the relationships between decomposition and gillnet soak duration and between decomposition and the rate at which bycatch fell from gillnets on mammal-focused trips. This investigation preceded the comparison of observer protocols, as we anticipated the protocols might differ in their probability of detecting drop-outs and we needed to better understand factors influencing drop-out rates. We then examined whether observers were equally likely under either protocol to see animals that were actively removed from the net (i.e. did not fall from the net). Lastly, we assessed whether observer protocols could be associated with differences in how often observers saw bycatch that fell out of the net as it was being hauled back.

Observer data

The US Northeast Fisheries Science Center estimates marine mammal bycatch using data from two observer programmes, the Northeast Fisheries Observer Program (NEFOP) and At-Sea Monitoring (ASM). Under these programmes, an average of 212 gillnet vessels and 1184 trips were observed each year from 1989 to 2019. Vessels had a median length of ~12.2 m (40 feet), and two-thirds of observed trips carried 2 or 3 crew, including the captain. The gillnet strings were composed of an average of 10 monofilament nets, each usually 91.4 m (300 feet) long and 2.7–3.7 m (9–12 feet) high, with 75% of mesh sizes falling between 9.5 and 30.5 cm (3.75 and 12 inches). The most commonly targeted species were monkfish (*Lophius americanus*), Atlantic cod (*Gadus morhua*), spiny dogfish (*Squalus acanthias*), pollock (*Pollachius virens*), Atlantic croaker (*Micropogonias undulatus*), winter skate (*Leucoraja ocellata*), and a mixture of species referred to as groundfish, together targeted by >65% of the trips observed. More details on the fishery can be found in Orphanides and Palka (2013). NEFOP, which was standardized in 1994, collects information on many fisheries characteristics, including the vessel, trip, haul, gear, target catch, and bycatch. This observer programme uses one of two sampling protocols when collecting data on gillnet fishing trips, with the same protocol applied throughout the trip. One observer protocol, the fish-focused protocol, concentrates on collecting fish discard information. Observers following this protocol are often processing fish from a previous haul as the net is hauled and are therefore not able to watch incoming nets. Furthermore, the observer's position on deck is determined by safety and available space, and thus the observer may not have a good view of the net being hauled. The second, mammal-focused observer protocol requires the observer to watch the net on the haulback for bycatch of marine mammals and other protected species, provided the observer can do so safely. Observers record all observed bycatch on both protocols. ASM trips, whose primary purpose is fish quota monitoring, began in 2010 and use only the fish-focused protocol. Gillnet observer records were gathered from both NEFOP and ASM programmes. Over the period 1994–2019, 26 007 trips were observed under the NEFOP programme and 6451 were observed under the ASM programme. Of these, 1499 NEFOP trips and 579 ASM trips had bycatch of the species studied. Harbour seal bycatch was analysed with data from 1994 to 2019. Observer data from 2004 to 2019 were used for grey seals, as there was little grey seal bycatch before 2004. For harbour porpoise, observer data from 2000 to 2019 were used because acoustic deterrent device use, which can influence bycatch rates, was not recorded before 2000.

Bycatch rate per species was defined as the number of animals divided by fishing effort. Two alternative measures of fishing effort were used: the exposure to nets during a fishing trip, and the total weight of fish landed (for consistency with how bycatch estimates are calculated in this region). The main text contains results using exposure; the supplementary material presents results with weight of fish landed. Exposure was calculated as the product of soak duration, number of nets hauled, net length, and effective net height. Soak duration was recorded in hours or tenths of hours for durations shorter than a day, and generally in multiples of 24 h for durations of a day or longer. Effective net height was the height of the net if tie-downs were not used, or the lesser of the length of the tie-downs and the recorded height of the net. Tie-downs affect the vertical profile of the net: straight if tie-downs are not used, or more C-shaped if they are used. We included hauls that used tie-downs on either all nets or no nets in a string.

A small number of hauls were discarded which had missing or unusually large values recorded for some variables known to impact bycatch (total landings, soak duration, number of nets, net length, and mesh size); together these constituted 0.38% of all hauls. Hauls with mesh sizes <11.4 cm (4.5″; 17.91% of all hauls since 1994) were also removed, as bycatch in those meshes is rare (only one bycatch event since 1994). Missing values for net height and twine size were filled in with medians of hauls on the same trip, on the same vessel using the same gear in the same year and month, on the same

Table 1. Methods for combining variables across hauls to obtain trip-level values.

Variable	Method for combining across hauls on a single trip
Tie-downs	Hauls with (without) tie-downs were kept if the majority of hauls on that trip used (did not use) tie-downs. For example, if most hauls on a trip used tie-downs but one haul did not, that one haul was discarded.
Nets set on same or on previous trip	Hauls whose nets were set on the same (previous) trip were kept if the majority of hauls on that trip hauled nets that were set on the same (previous) trip. Any hauls in the minority were discarded.
Depth (fathoms)	Median depth of hauls on the trip.
Sea surface temperature (°C)	Median temperature of hauls on the trip.
Exposure	Sum over hauls on the trip.
Date	Median over hauls on the trip.
Mesh size (small/large)	Mesh sizes on each haul were weighted by the net area on each haul on the trip and averaged, then each trip's average was categorized as small (at least 11.4 cm [4.5″] and <22.9 cm [9″]) or large (22.9 cm [9″] or larger). Bycatch rarely occurs in meshes <11.4 cm, and 22.9 cm was a natural breakpoint in the distribution of mesh size.
Twine size (mm)	Average of the twine sizes on each haul, weighted by the net area on each haul on the trip.
Latitude/longitude	Median over each haul on the trip.
Pingers	Fraction of hauls on a trip that used pingers.

vessel using the same gear in the same year, or on the same vessel in the same year. This method filled in net height for 12.8% of hauls and twine size for 25.0% of hauls. Hauls that were missing either depth or the number of nets were excluded.

Observers recorded a bycaught animal's condition as alive (0.3%), freshly dead (83.5%), moderately decomposed (10.7%), severely decomposed (2.8%), or other less common conditions which we discarded (2.7%). Observers were formally trained to classify body conditions in a standardized manner using a set of characteristics that distinguish freshly dead, moderately decomposed, and severely decomposed conditions (NEFSC 2021, p. 105). The characteristics pertain to overall appearance, odour, degree of bloating, tissue firmness and colour, intactness or tendency to disintegrate, skin condition, and other features. Observers were also required to write comments giving at least three reasons supporting their classification of the condition (NEFSC 2021). The characteristics used are reproduced from NEFSC (2021) in Supplementary Table S1 in the supplementary material. Observers also noted whether an animal fell from the net on its own or was removed from the net by crew.

Hauls from the same trip were combined because they were not expected to be independent, whereas trips were considered to be independent conditional on characteristics such as date, location, gear, and environmental factors. Table 1 shows how variables were combined across hauls to derive a trip-level value.

Decomposition and soak duration

We tabulated decomposition state and median and minimum soak durations for hauls with bycatch to determine if soak duration influences whether an animal is observed. We chose species-specific thresholds to cap soak duration when calculating exposure for analyses that included only live or freshly dead animals. The soak duration thresholds were intended to relate to the length of time for an animal to transition from freshly dead to moderately decomposed, which varied by species. When calculating the bycatch rate, a capped soak duration is a more relevant duration for a freshly dead animal, which was necessarily recently bycaught, than total soak duration is. We examined the relation between state of decomposition and soak duration only for hauls observed under a mammal-focused protocol, to avoid possible bias from unobserved drop-outs under a fish-focused protocol, since degree of decomposition may relate to the probability of dropping out. This examination only included dead bycatch because so few were observed alive that they would have provided little information.

Decomposition and falling from the net

We suspected that more decomposed animals may be increasingly likely to fall from the net rather than be removed by the crew, in part because the bodies may be more flaccid. To assess this, we calculated the proportion of animals that fell from the net and 95% Clopper-Pearson confidence bounds for the proportion using the ci_proportion() function in the R confintr package (Mayer 2020). The proportion was calculated for each observer protocol and decomposition state (freshly dead, moderately decomposed, and severely decomposed).

Observer protocol and bycatch rate of animals removed from net

We hypothesized that under either protocol, observers are likely to see animals that the crew removes from the net. To test this, we compared bycatch rates of animals that were removed from the net for the two observer protocols. For this analysis, we created pairs of trips that differed in observer protocol and that were matched as closely as possible for date, location, and relevant gear and environmental characteristics; trips that didn't have closely comparable counterparts under the other protocol were discarded. The matching process produces post-hoc paired control and treatment groups with similar covariate distributions; this approach supports causal inference in observational studies (Rubin 1973, Rosenbaum and Rubin 1983, Ho et al. 2007). Matching methods can provide more robust modelling than multiple regression when an outcome of interest (e.g. bycatch) is rare because matching methods allow estimation of the relationship between covariates and treatment (e.g. observer protocol) instead of the relationship between covariates and outcome that is the focus of multiple regression (Fu et al. 2019).

Trips were discarded before matching if they were missing a value for any of the characteristics to be used in matching. Nearest-neighbour matching was performed using Mahalanobis distance with calliper values for most covariates (Table 2), using the matchit() function in the MatchIt pack-

Table 2. Matching pairs of trips for observer protocol: calliper values for pairing criteria, number of matched trips, and number of matched trips with bycatch.

Species	Max days apart	Max depth difference (ftm)	Max twine size difference (mm)	Max lat/lon degrees apart	Same mesh size category and usage of tiedowns	Max surface temperature difference (°C)	Nets set & hauled	Max pinger difference	Total trips paired	Matched trips with bycatch
Grey seal	34	17 (31 m)	0.16	0.32, 0.33	Yes	6.67	Yes	NA	374	211
Harbour seal	40	16 (29 m)	0.16	0.32, 0.33	Yes	5	Yes	NA	380	209
Harbour porpoise	36	15 (27 m)	0.17	0.32, 0.33	Yes	5.56	Yes	0.2	274	141

'Nets set & hauled' indicates that for both members of a pair, either all the nets were hauled on the same trip as they were set on, or they were all hauled on a later trip from the trip when they were set. Pairing criteria marked NA were not used for a species.

age (Ho et al. 2011) in R (R Core Team 2020). A calliper value for a variable represents a threshold on what trips can be matched; trips that differ on that variable by more than the threshold cannot be considered a match. Matching criteria were explored by a grid search through reasonable ranges of values to determine calliper values. The calliper values were chosen by balancing similarity between matched trips against the number of matched trips with bycatch. If the callipers are too restrictive, trips will be closely matched but there will be fewer matched trips and fewer matched trips with bycatch. The balance of the covariates was assessed before and after matching by plotting the standardized mean differences and by examining the variance ratios, empirical cumulative distribution function (eCDF) mean, and eCDF max values for the covariates. After matching, the standardized mean differences between groups for all covariates were <0.10, and most were <0.05 [below the thresholds for assessing balance of 0.25 and 0.10 suggested by Ho et al. (2007) and Stuart (2010)]. These values indicated the matching procedure balanced the groups adequately. The mean absolute differences on each covariate between paired trips are given in the supplementary material (Supplementary Tables S2, S3, and S4, together with the number of matched and unmatched trips in Supplementary Tables S5, S6, and S7).

The differences in bycatch rate of removed animals between the mammal- and fish-focused trips within each pair were examined with a one-tailed Wilcoxon signed-rank test. The hypothesis was that the bycatch rates of removed animals on mammal-focused trips exceeded the rates on the corresponding paired fish-focused trips. During matching, first mammal-focused trips with bycatch were matched where possible with fish-focused trips; then matches for any still-unpaired fish-focused trips with bycatch were sought with mammal-focused trips. The decision to match trips with bycatch whenever possible was based on the fact that the Wilcoxon signed-rank test discards pairs with tied values; as most trips had no bycatch, matching trips into pairs where neither had bycatch is not useful.

Observer protocol and live or freshly dead animals falling from net

We hypothesized that mammal-focused observers are more likely to see animals that fall from the net than are fish-focused observers. This analysis compared the relative frequency under the two observer protocols of bycaught animals that fell from the net. Only live or freshly dead animals were considered, to avoid any interaction between decomposition and observer protocol. Net characteristics could influence whether an animal falls out of the net, for example by entangling the animal more or less tightly, and therefore matching was again useful. The same matching process and criteria were used as for the comparison of bycatch rates (Table 2), with one additional criterion, namely exposure or weight of fish landed. In the previous comparison, fishing effort was incorporated by calculating the bycatch rate; here, trips were matched to have similar fishing effort. Further, because this comparison only considered live or freshly dead animals, which were caught within a short time before haulback, the soak duration used in calculating exposure was capped to not exceed a threshold of recency. A two-by-two contingency table of the counts of bycatch falling or being removed from the net for each observer protocol was analysed with a one-tailed Fisher's exact test.

Finally, we estimated the odds of a live or freshly dead bycaught animal falling from the net, versus being removed from it, on a mammal-focused trip. The mammal-focused trips were those from the pairs used for the comparison between protocols of proportions of live and freshly dead animals falling or being removed from the net. 95% confidence intervals around the odds ratios were calculated with the adjusted bootstrap percentile (BCa) method in the R boot package v1.3–27 (Davison and Hinkley 1997, Canty and Ripley 2021).

Results

Decomposition and soak duration

The minimum soak duration at which an animal was recorded was greater for more decomposed conditions, and decomposed conditions were generally associated with a longer median soak duration (Table 3). Based on these observations, we chose species-specific maxima to cap soak duration when calculating exposure for use in the analysis of live and freshly dead bycatch. Maxima were picked to try to separate freshly dead animals from decomposed animals; the maxima chosen were 60 h for grey seals, 24 h for harbour seals, and 36 h for harbour porpoise.

Decomposition and falling from the net

We found that a larger proportion of moderately decomposed grey and harbour seals fell from the net than of freshly dead animals, although the confidence intervals around the proportions were fairly wide (Fig. 1a, b, Supplementary Tables S8 and S9). The numbers of severely decomposed seals that were observed were quite small, resulting in very wide confidence intervals that are difficult to draw inferences from. Data on harbour porpoise exhibited a different pattern, with

Table 3. State of decomposition and median and minimum soak durations for hauls with known soak durations on which dead bycaught animals were observed on mammal-focused trips.

Species	State of decomposition	Count of animals	Minimum soak duration (h)	Median soak duration (h)
Grey seal	Freshly dead	453	2.4	120
	Moderately decomposed	39	48.0	144
	Severely decomposed	3	142.0	144
Harbour seal	Freshly dead	332	4.9	96
	Moderately decomposed	80	23.0	120
	Severely decomposed	20	48.0	131
Harbour porpoise	Freshly dead	175	2.5	96
	Moderately decomposed	41	18.7	144
	Severely decomposed	29	48.0	120

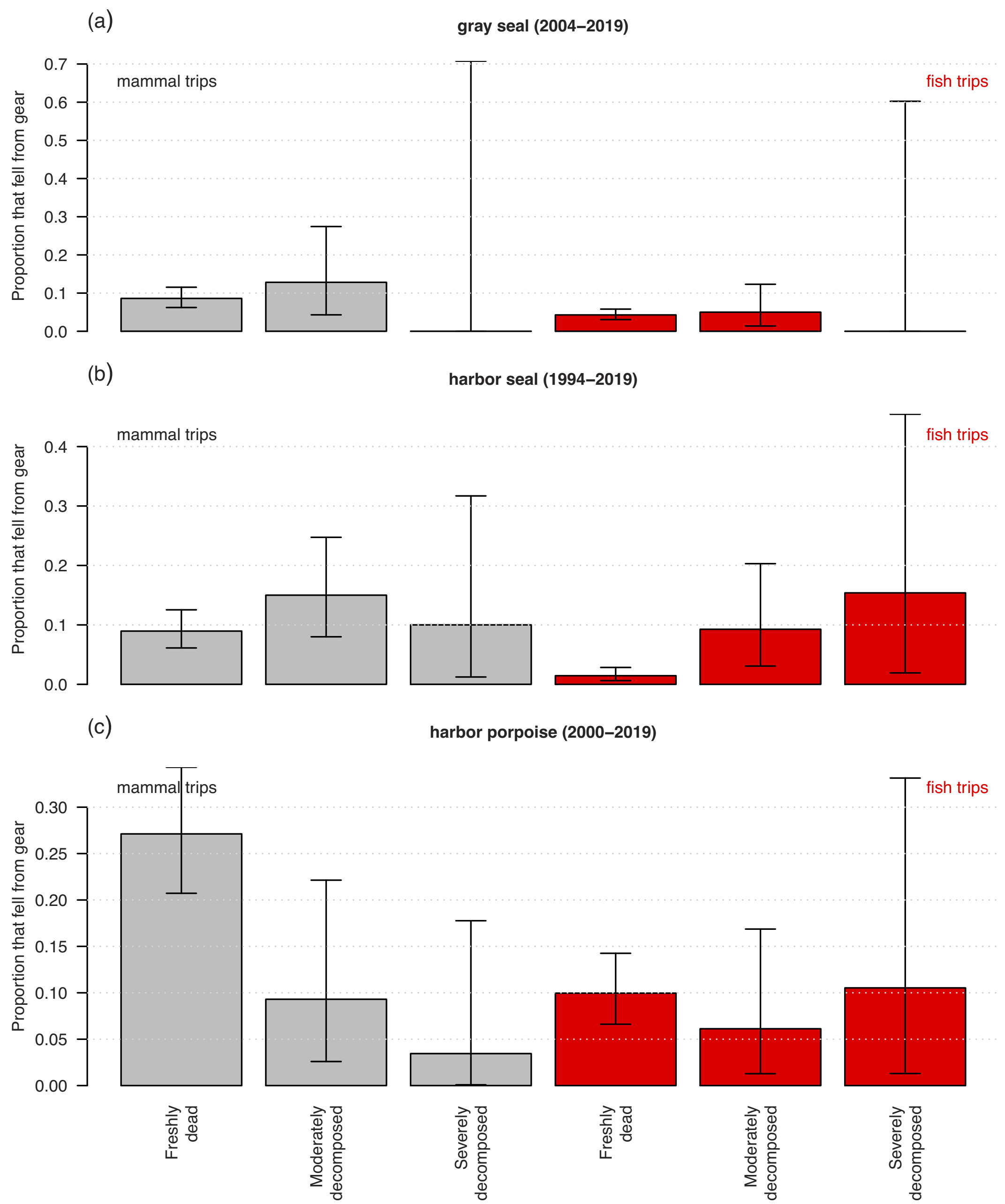

Figure 1. Proportion of animals that fell from gear, by decomposition state and observer protocol, with 95% confidence intervals for the proportions. Proportions should not be compared across protocols because the trips are not matched for environmental or fishing characteristics.

Table 4. Results of analysis of whether bycatch rate of animals removed from the net was higher on mammal-focused observer protocol trips than on fish-focused protocol trips.

Species	One-tailed Wilcoxon signed-rank test of differences in bycatch rate
Grey seal	$P = 0.0189$
Harbour seal	$P = 0.000\ 294$
Harbour porpoise	$P = 0.512$

The bycatch rate per trip was the number of animals that were removed from the net by the crew, divided by exposure as the unit of effort.

a greater proportion of freshly dead animals falling from the net than of moderately decomposed animals (Fig. 1c, Supplementary Table S10). In these results, the proportions are not easily comparable across protocols because the mammal-focused and fish-focused trips may differ in important characteristics.

Observer protocol and bycatch rate of animals removed from net

Observer protocol had a significant effect for harbour seals, such that the bycatch rate was higher when the trip was assigned to a mammal-focused observer protocol, with a mean of 0.18 animals per million square-foot hours of net exposure compared to a mean of 0.11 on fish-focused trips (Table 4). The protocol may have had a significant effect in the same direction for grey seals (mean of 0.27 animals per million square-foot hours of net exposure vs. 0.18 on fish-focused trips), while for harbour porpoise, there was insufficient evidence to conclude the bycatch rate was higher on mammal-focused trips (mean of 0.18 vs. 0.13 animals per million square-foot hours of net exposure). The bycatch rates included animals that had to be removed from the net, in any decomposition state, and excluded animals that fell from the net. Results were quite similar whether bycatch rate was calculated per exposure (Table 4) or per weight of fish landed (Supplementary Table S11).

Observer protocol and live or freshly dead animals falling from net

Grey seals and harbour seals were likely to be reported as having fallen from the net less often by fish-focused observers than mammal-focused observers (though no adjustment to the *p*-values has been made for the use of multiple tests; see Table 5 for results with the bycatch rate calculated per exposure and Supplementary Table S12 with bycatch rate calculated per weight of fish landed). The proportion of harbour porpoise that fell from the net shows the same tendency as for the seal species but was more similar for the two observer protocols. Harbour porpoise had the smallest sample and highest proportions of fallen animals, and an effect might be detected with more confidence in a larger data set. Using only the mammal-focused trip data, the odds for an animal falling from the net on a mammal-focused trip were similar for grey and harbour seals, and higher for harbour porpoise (Table 6, Supplementary Table S13).

Estimation of bias in bycatch estimate

If data from both types of protocol are available and if the sampling designs associated with each observer protocol are or can be made comparable, an estimate of unseen fallen animals can be calculated using the rates of animals removed from the net and of animals observed to fall from the net. If the true ratio on mammal-focused trips of animals falling from the net to those removed from the net is assumed equal to the true ratio on fish-focused trips, one can estimate the number of animals $\hat{F}_f$ that are expected to fall from the net on a fish-focused trip, including those that fall unobserved (assume $F_m/R_m = F_f/R_f$, so $\hat{F}_f = \hat{F}_m \times \hat{R}_f/\hat{R}_m$, where F is the number of animals that fell from the net, R is the number of animals that were removed from the net, with subscripts f and m representing observations under fish- and mammal-focused protocols, respectively; in calculations, the observed counts $\hat{F}_m$, $\hat{R}_f$, and $\hat{R}_m$ can be used as estimates of the true quantities F_m, R_f, and R_m). The total estimated fallen animals can then be added to the observed number removed on fish-focused trips. In practice, the bycatch rate for seals removed from nets on fish-focused trips may have been lower than that on mammal-focused trips (Table 4), so the result could still represent an underestimate. If we assess the bias thusly for fish-focused trips in 2019 using 16–26 years of historical mammal-focused trips, assuming that the fishery is sampled in the same way for both protocols, we roughly estimate that on fish-focused trips grey seal bycatch would increase by 3%, harbour seals would increase by 6%, and harbour porpoise would increase by 25% (Table 7; Supplementary Table S14 illustrates the variability in the estimated bias, with % increases in bycatch estimates over 2015–2019 ranging from 0% to 6% for grey seals, 0% to 10% for harbour seals, and 25% to 40% for harbour porpoise).

Discussion

This study is among the first to examine the effect of fisheries observer protocols on observed marine mammal bycatch and drop-out rates. We tested the hypothesis that protocols affected observed bycatch and drop-out rates of grey seals, harbour seals, and harbour porpoises, and tested whether protocol was related to bycatch rate for animals that did not

Table 5. One-tailed Fisher's exact test of independence of observer protocol from whether live or freshly dead bycatch fell from the net or was removed from the net by the crew.

Species	Observer protocol	Animal fell from net	Animal removed from net	Fisher's exact test
Grey seal	Fish-focused	12	232	$P = 0.0181$
	Mammal-focused	26	227	
Harbour seal	Fish-focused	1	69	$P = 0.0143$
	Mammal-focused	16	143	
Harbour porpoise	Fish-focused	11	67	$P = 0.0607$
	Mammal-focused	21	63	

The criteria used in pairing fish-focused and mammal-focused trips included exposure.

Table 6. Estimate and 95% confidence interval for the odds of a live or freshly dead bycaught animal falling from the net, versus being removed from it, on a mammal-focused trip.

Species	Odds of falling from the net on a mammal-focused trip	Confidence interval for the odds
Grey seal	0.11	0.07–0.16
Harbour seal	0.11	0.06–0.18
Harbour porpoise	0.33	0.18–0.50

Odds were calculated from the mammal-focused trips used in the paired analysis in Table 5.

fall from the nets on their own. Observer protocol affected the reported bycatch and drop-out rates, at least for some species. To better understand what influences drop-outs, we performed exploratory data analysis to roughly gauge the length of time for a bycaught animal to decompose in the study region and to examine the relationship between an animal's state of decomposition and falling from the net. Soak duration was related to decomposition state, and decomposition state related to drop-out rates, though the exact relationships varied by species. Formally testing the hypotheses suggested by the explorations of decomposition and drop-out rate would require a new, independent data set. The results of this study should inform observer programme operations, ultimately supporting better marine mammal bycatch estimates and improved management of protected species. This is important because of the significant threat bycatch poses to marine mammal populations worldwide, and because of US import standards for fish products requiring strong bycatch controls.

Marine mammal decomposition increased with soak duration. For seals, greater decomposition corresponded to more observed drop-outs from gillnets. We might expect that as animals decompose, the carcass would fall from the net more easily as the flesh softens or is removed by predators. This suggests that the true bycatch rate for seals, approximated by the observed bycatch per unit of effort, is underestimated because increasing decomposition results in more drop-outs, not all of which are observed, even by conscientious and well-placed observers. Some drop-outs may occur underwater from scavenging or decomposition while the net is soaking, or the force of the net haulback may cause a carcass to come loose before it is visible to an observer. The bycatch rate would be underestimated even more if data from fish-focused observers were treated as equal to data from mammal-focused observers, because our data suggest that observers focused on documenting marine mammal bycatch tended to notice more seal drop-outs than did observers focused on documenting fish catch.

Because of the relationships between soak duration, decomposition, and dropping out of the nets, the observed number of bycaught animals in hauls with long soak durations may reflect both accumulation and disappearance, and reported bycatch could even decrease with long soaks. Therefore, the rate of bycatch can most directly be estimated by considering only recently caught animals, or as an approximation to recency, live and freshly dead bycatch. The length of time that can be considered 'recent' differs by species, as smaller animals may decompose more quickly and softer-bodied animals might be scavenged more rapidly. Recency also likely depends on environmental characteristics such as water temperature, net depth, and scavenger presence.

Harbour porpoise data showed some of the same trends as seals, but had some important differences. As with seals, greater decomposition was associated with longer soak duration (Table 3). However, harbour porpoise showed a different relationship from seals between state of decomposition and the number of animals falling from the net, with fewer porpoise falling from the net on mammal-focused trips as decomposition increased (Fig. 1). Possible explanations for why this might be the case can be found in the supplementary material.

The bycatch rate for harbour porpoise removed from the net was not significantly different between observer protocols (Table 4); however, there was uncertainty whether the proportion of animals that fell from the net differed between protocols ($P = 0.06$, Table 5). As described above, harbour porpoise that were removed from the net were likely firmly entangled, which may make it more likely that both fish- and mammal-focused observers would notice them. Consequently, we would not expect a difference in fish- and mammal-focused bycatch rates of harbour porpoise removed from the net. By contrast, seals that need to be removed from the net are anecdotally reported to be easier to disentangle, and so may be easier to miss if an observer is not focused on marine mammal bycatch. Harbour porpoise fell at about three times the rate of seals (Table 6), which may increase the chance of drop-outs that a fish-focused observer misses. The effect size for the harbour porpoise observer protocol bycatch rate difference was smaller than for seals and had a P-value of 0.06 (Table 5); this could be partly related to the relative rarity of harbour porpoise bycatch.

Our research indicates the influence of fisheries observer protocols on bycatch rates is species-specific and complicated by soak duration, which relates to decomposition and hence to the proportion of animals that fall. Bravington and Bisack (1996), who did not explicitly account for soak duration and state of decomposition, estimated differences in harbour porpoise bycatch rates between observers watching the gear and water during haulback and observers attending to other duties, and found that the relative efficiencies of the two protocols varied. Northridge et al. (2014, 2015) likewise reported larger numbers of all mammals and of cetaceans per observed haul when observers were dedicated to monitoring bycatch of protected species. While the magnitude of impact of observer protocol on bycatch rate may differ by species and by fishery, it generally appears that when observers are not focused on recording marine mammal bycatch, estimated bycatch rates are lower than when observers are focused on mammal bycatch.

The results of this study are instructive for organizations responsible for developing or managing observer programmes (e.g. 2008/56/EC, EC 2008; South Pacific Regional Fisheries Management Organisation, SPRFMO 2022). Our results indicate that to best estimate marine mammal bycatch in gillnets, an observer should be dedicated to watching the haulback and surrounding waters. If electronic monitoring (EM) by video is used, the camera field of view should include the haulback of the nets. Reviewers of EM footage watching for marine mammal bycatch could note whether the field of view includes the nets, giving context for the use of the data and for the interpretation of the results. Alternative approaches designed to meet multiple objectives in a cost-effective manner include randomly selecting a subset of hauls on a trip for mammal-focused observation (ICES 2019). Advantages of such an approach include using a single sampling design for all obser-

Table 7. Example calculation of unseen animals that fell from the gear, using bycatch observed on fish-focused trips in 2019 and $\hat{F}_f = \hat{F}_m \times \hat{R}_f / \hat{R}_m$.

Species	Historical mammal-focused trips		2019 fish-focused trips				
	Removed from gear ($\hat{R}_m$)	Fell from gear ($\hat{F}_m$)	Removed from gear ($\hat{R}_f$)	Fell from gear (Observed)	Estimated total fallen animals ($\hat{F}_f$)	Estimated unseen animals that fell	Percent increase in bycatch
Grey seal	425	40	130	8	12.2	4.2	3%
Harbour seal	305	30	26	1	2.6	1.6	6%
Harbour porpoise	130	51	17	2	6.7	4.7	25%

The odds ratios of removed to fallen animals are calculated using live or freshly dead bycatch observed on mammal-focused trips over 2004–2019 for grey seals, 1994–2019 for harbour seals, and 2000–2019 for harbour porpoise.

vations, which would facilitate comparison of the protocols. If both fish- and mammal-focused observations are available, several approaches could be taken to improve bycatch estimates. If there are sufficient mammal-focused observations, estimates could be calculated without fish-focused observations. Many regions may not have the choice of only using mammal-focused observer data to estimate bycatch, either because those data do not exist, or because removing fish-focused observations would reduce the available data so much as to lead to poor estimates. In such cases, fish-focused observations could be used to develop a lower-bound estimate, or rates from the two observer protocols could be considered both separately and together to get a fuller picture of the potential range of bycatch.

Bycatch rate differences observed in this study may not apply directly to other fisheries because the results vary at least by taxon, likely by species, and possibly also by body size and gear. Rates of decomposition and, thus, drop-out rates are probably related to local environmental characteristics such as water temperature, and consequently bias may also vary with the local environment. Gear characteristics such as mesh size and twine size are known to influence bycatch (Orphanides 2009, Northridge et al. 2017) and likewise could influence drop-out rates. Another consideration is the unit of effort used in the bycatch estimation process, although we found that using an alternate metric, namely fisheries landings, had little impact on the results (see Supplementary Material). Whether the biases estimated here play a significant role in population assessments depends on the status of the stock, the size of the fishery, and the proportions of fish- and mammal-focused observer records. Apart from that, if the protocols are applied to different stratifications of the fishery, the degree of underestimation may vary across strata. In theory, one could also compensate for animals that were never recorded because they disappeared over the course of longer soaks. However, calculating this would require data from underwater monitoring of nets, which are currently unavailable, or observations monitoring bycaught animals that are left in place over time.

To achieve unbiased estimates of marine mammal bycatch, the use of mammal-focused observer protocols should be incorporated into the best management practices suggested by Moore et al. (2021), and observers should provide information on the decomposition state of bycatch according to standardized definitions. If fish-focused observer protocols are used to estimate marine mammal bycatch for gillnet fishing, mortality is likely underestimated, and consideration should be given to how to best evaluate the underestimation and its potential impact on affected marine mammal populations (Wade et al. 2021). We have outlined a way to evaluate bias and shown the relevance of species and decomposition state, and our estimates may serve as a useful reference point and comparison for other fisheries.

Acknowledgements

Thanks for very useful discussions are due to Andrea Bogomolni, Stephanie DePasquale, Greg Early, Jordan Katz, Marjorie Lyssikatos, Kimberly Murray, Debra Palka, and Frederick Wenzel. Three anonymous reviewers provided helpful comments that improved the manuscript.

Conflict of interest: The authors have no conflicts of interest to declare.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Supplementary material available at *ICESJMS* online includes a list of characteristics used by observers to classify marine mammal body condition; tables showing the mean absolute differences in matched characteristics between matched pairs of trips; sample sizes after matching; tables giving the estimated proportion of bycaught animals that fell from gillnets, 2004–2019, and confidence bounds on the proportion, by decomposition state and observer protocol; the results of analyses using weight of fish landed as a measure of fishing effort; and example estimates over 5 years of unseen freshly dead animals on fish-focused trips. Also included is a discussion of why the state of decomposition and the proportion of drop-outs might have a different relationship for harbour porpoise than for seals.

Author contributions

C.D.O. conceptualized the project. K.P. developed the methodology and performed the formal analysis. Both authors contributed to writing and revising the manuscript.

Data availability

Data underlying the analyses of decomposition and soak duration, decomposition and falling from the net, and observer protocol and live or freshly dead animals falling from the net is available from https://doi.org/10.6084/m9.figshare.24646026, together with code. Data underlying the analysis of observer protocol and bycatch rate of animals removed from the net can only be shared in aggregate or summary form because the US Magnuson-Stevens Fishery Conservation and Management Act (MSA) requires NOAA to maintain the confidentiality of any information submitted in compliance with the

Act and of any observer information. Because the analysis requires detailed, unaggregated information, the data used in this analysis are not provided. Release of specific information associated with incidental takes of marine mammals is exempt from MSA confidentiality restrictions due to requirements under the US Marine Mammal Protection Act.

References

Australian Fisheries Management Authority, Australian Government. Australian Sea Lion Management Strategy. 2010; https://www.fish.gov.au/Archived-Reports/2012/reports/Documents/AFMA_2010b.pdf (5 May 2022, date last accessed).

Avila IC, Kaschner K, Dormann CF. Current global risks to marine mammals: taking stock of the threats. *Biol Conserv* 2018;**221**:44–58. https://doi.org/10.1016/j.biocon.2018.02.021

Bravington MV, Bisack KD. Estimates of harbour porpoise bycatch in the Gulf of Maine sink gillnet fishery, 1990–1993. *Rep Intern Whaling Comm* 1996;**46**:567–74.

Canty A, Ripley B. boot: bootstrap R (S-plus) functions. R package version 1.3–27. 2021. https://CRAN.R-project.org/package=boot

Davison AC, Hinkley DV. *Bootstrap Methods and their Applications*. Cambridge: Cambridge University Press, 1997.

EC. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). 2008. http://data.europa.eu/eli/dir/2008/56/2017-06-07

Fu EL, Groenwold RH, Zoccali C *et al.* Merits and caveats of propensity scores to adjust for confounding. *Nephrol Dial Transplant* 2019;**34**:1629–35. https://doi.org/10.1093/ndt/gfy283

Gilman E, Passfield K, Nakamura K. *Performance Assessment of Bycatch and Discards Governance by Regional Fisheries Management Organizations*. Gland: IUCN, 2012, ix + 484.

Goldsworthy SD, Page B, Shaughnessy PD *et al.* Mitigating Seal Interactions in the SRLF and the Gillnet Sector SESSF in South Australia: Report to the Fisheries Research and Development Institute. South Australian Research and Development Institute (Aquatic Sciences), Adelaide. SARDI Publication No. F2009/000613-1. SARDI Research Report Series No. 405. 2010. https://pir.sa.gov.au/__data/assets/pdf_file/0013/232006/No_405_Mitigating_Seal_Interactions_in_the_SRLF_and_the_Gillnet_Sector_SESSF_in_SA.pdf (5 May 2022, date last accessed).

Hamer DJ, Goldsworthy SD, Costa DP *et al.* The endangered Australian sea lion extensively overlaps with and regularly becomes bycatch in demersal shark gill-nets in South Australian shelf waters. *Biol Conserv* 2013;**157**:386–400. https://doi.org/10.1016/j.biocon.2012.07.010

Ho DE, Imai K, King G *et al.* Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Pol Analysis* 2007;**15**:199–236. https://doi.org/10.1093/pan/mpl013

Ho DE, Imai K, King G *et al.* MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw* 2011;**42**:1–28. https://doi.org/10.18637/jss.v042.i08

ICES. 2019. Joint WGBYC-WGCATCH Workshop on sampling of bycatch and PET species (WKPETSAMP), 24–26 April 2018, SLU Aqua, Lysekil, Sweden. *ICES CM 2018/EOSG:35* 76pp. https://doi.org/10.17895/ices.pub.8183

Marine Stewardship Council. MSC Fisheries Standard v3.0. 2022. https://www.msc.org/docs/default-source/default-document-library/for-business/program-documents/fisheries-program-documents/msc-fisheries-standard-v3-0.pdf?sfvrsn=53623a3_31

Mayer M. confintr: confidence intervals. R package version 0.1.1, 2020. https://CRAN.R-project.org/package=confintr

Moore JE, Heinemann D, Francis TB *et al.* Estimating bycatch mortality for marine mammals: concepts and best practices. *Front Mar Sci* 2021;**8**:752356. https://doi.org/10.3389/fmars.2021.752356

NEFOP. Northeast Fisheries Observer Program, Fisheries Sampling Branch Observer Operations Manual. United States, National Marine Fisheries Service; Northeast Fisheries Science Center (U.S.). 2016. https://doi.org/10.25923/wnxt-9840.

Northeast Fisheries Science Center (NEFSC). Observer Operations Manual 2021. Fisheries Monitoring and Research Division, Northeast Fisheries Science Center, National Marine Fisheries Service, NOAA Fisheries Service, U.S. Department of Commerce. 2021. https://web.archive.org/web/20231012151611/https://www.nafo.int/Portals/0/PDFs/fc/proc/USA_2021ObserverOperationsManual.pdf. (12 October 2023, date last accessed).

Northridge S, Coram A, Kingston A *et al.* Disentangling the causes of protected-species bycatch in gillnet fisheries. *Conserv Biol* 2017;**31**:686–95. https://doi.org/10.1111/cobi.12741

Northridge S, Kingston A, Thomas L. Annual report on the implementation of Council Regulation (EC) No 812/2004 during 2013. 2014. https://web.archive.org/web/20231016162952/https://www.ascobans.org/sites/default/files/document/NSG4_Inf_3.1.1.b_UK_812report2013.pdf. (16 October 2023, date last accessed).

Northridge S, Kingston A, Thomas L. Annual report on the implementation of Council Regulation (EC) No 812/2004 during 2014. 2015. https://web.archive.org/web/20231016160953/https://research-repository.st-andrews.ac.uk/bitstream/handle/10023/6855/12530_UK812Report2015on2014Final.pdf. (16 October 2023, date last accessed).

Orphanides CD. Protected species bycatch estimating approaches: estimating harbor porpoise bycatch in US northwestern Atlantic gillnet fisheries. *J Northwest Atlantic Fish Sci* 2009;**42**:55–76. https://doi.org/10.2960/J.v42.m647

Orphanides CD, Palka DL. Analysis of harbor porpoise gillnet bycatch, compliance, and enforcement trends in the US northwestern Atlantic, January 1999 to May 2010. *Endang Species Res* 2013;**20**:251–69. https://doi.org/10.3354/esr00499

R Core Team. *R Foundation for Statistical Computing*. Vienna: R Core Team, 2020. https://www.R-project.org/

Read AJ. The looming crisis: interactions between marine mammals and fisheries. *J Mammal* 2008;**89**:541–8. https://doi.org/10.1644/07-MAMM-S-315R1.1

Read AJ, Drinker P, Northridge S. Bycatch of marine mammals in U.S. and global fisheries. *Conserv Biol* 2006;**20**:163–9. https://doi.org/10.1111/j.1523-1739.2006.00338.x

Reeves RR, McClellan K, Werner TB. Marine mammal bycatch in gillnet and other entangling net fisheries, 1990 to 2011. *Endang Species Res* 2013;**20**:71–97. https://doi.org/10.3354/esr00481

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**:41–55. https://doi.org/10.1093/biomet/70.1.41

Rubin DB. Matching to remove bias in observational studies. *Biometrics* 1973;**29**:159. https://doi.org/10.2307/2529684

SPRFMO. CMM 16-2022: conservation and management measure establishing the SPRFMO observer programme, 2022. https://www.sprfmo.int/assets/Fisheries/Conservation-and-Management-Measures/2022-CMMs/CMM-16-2022-Observer-Programme-7Mar22.pdf

Stuart EA. Matching methods for causal inference: a review and a look forward. *Statis Sci* 2010;**25**:1. https://doi.org/10.1214/09-STS313

Tregenza NJC, Berrow SD, Hammond PS *et al.* Common dolphin, *Delphinus delphis* L., bycatch in bottom set gillnets in the Celtic Sea. *Rep Intern Whaling Comm* 1997a;**47**:835–9.

Tregenza NJC, Berrow SD, Hammond PS *et al.* Harbour porpoise (*Phocoena phocoena* L.) by-catch in set gillnets in the Celtic Sea. *ICES J Mar Sci* 1997b;**54**:896–904. https://doi.org/10.1006/jmsc.1996.0212

Wade PR, Long KJ, Francis TB *et al*. Best practices for assessing and managing bycatch of marine mammals. *Front Mar Sci* 2021;**8**: 757330. https://doi.org/10.3389/fmars.2021.757330

Williams R, Burgess MG, Ashe E *et al*. U.S. seafood import restriction presents opportunity and risk. *Science* 2016;**354**:1372–4. https://doi.org/10.1126/science.aai8222.

Handling Editor: Anita Gilles