PAPER

# Long-Term Effectiveness, Failure Rates, and "Dinner Bell" Properties of Acoustic Pingers in a Gillnet Fishery

AUTHORS
**James V. Carretta**
**Jay Barlow**
Southwest Fisheries Science Center,
National Marine Fisheries Service,
National Oceanic and
Atmospheric Administration

ABSTRACT

The long-term effectiveness of acoustic pingers in reducing marine mammal bycatch was assessed for the swordfish and thresher shark drift gillnet fishery in California. Between 1990 and 2009, data on fishing gear, environmental variables, and bycatch were recorded for over 8,000 fishing sets by at-sea fishery observers, including over 4,000 sets outfitted with acoustic pingers between 1996 and 2009. Bycatch rates of cetaceans in sets with ≥30 pingers were nearly 50% lower compared to sets without pingers ($p = 1.2 \times 10^{-6}$), though this result is driven largely by common dolphin (*Delphinus delphis*) bycatch. Beaked whales have not been observed entangled in this fishery since 1995, the last full year of fishing without acoustic pingers. Pinger failure (≥1 nonfunctioning pingers in a net) was noted in 3.7% of observed sets. In sets where the number of failed pingers was recorded, approximately 18% of deployed pingers had failed. Cetacean bycatch rates were 10 times higher in sets where ≥1 pingers failed versus sets without pinger failure ($p = 0.002$), though sample sizes for sets with pinger failure were small. No evidence of habituation to pingers by cetaceans was apparent over a 14-year period of use. Bycatch rates of California sea lions in sets with ≥30 pingers were nearly double that of sets without pingers, which prompted us to examine the potential "dinner bell" effects of pingers. Depredation of swordfish catch by California sea lions was not linked to pinger use—the best predictors of depredation were total swordfish catch, month fished, area fished, and nighttime use of deck lights on vessels.
Keywords: acoustic pingers, bycatch, marine mammals, habituation, gillnets

## Introduction

Acoustic pingers are effective at reducing the bycatch of a wide variety of marine mammal species (Kraus et al., 1997; Trippel et al., 1999; Gearin et al., 2000; Bordino et al., 2002; Barlow & Cameron, 2003; Palka et al., 2008). With few exceptions (Kraus, 1999; Cox et al., 2001; Palka et al., 2008), the long-term efficacy of pingers has seldom been addressed, particularly with respect to the potential for "habituation" by animals to acoustic devices. Even fewer studies have examined the potential "dinner bell" effect of pingers, whereby marine mammals are attracted by pingers to fishing nets, resulting in depredation of catch (Dawson, 1994; Kraus et al., 1997).

Barlow and Cameron (2003) reported that acoustic pingers significantly reduced cetacean and pinniped bycatch in the drift gillnet fishery for swordfish and thresher shark in California (hereafter referred to as "the fishery") during a controlled experiment in 1996 and 1997. At that time, conclusions about pinger effectiveness in reducing bycatch were limited to short-beaked common dolphins (*Delphinus delphis*) and California sea lions (*Zalophus californianus*), due to small sample sizes for other species. With nine additional years of observer data from the fishery, Carretta et al. (2008) showed that acoustic pingers apparently eliminated beaked whale bycatch. Acoustic pingers have been utilized in the fishery for 14 consecutive years (1996–2009), with 4,238 sets outfitted with pingers. This extensive dataset allowed us to assess the long-term performance of pingers beyond the experimental results reported by Barlow and Cameron (2003) and to address the following questions: Are observed data consistent with gear compliance regulations outlined in the Pacific Offshore Cetacean Take Reduction Plan (Federal Register, 1997) implemented in 1997? Have pingers remained effective at reducing bycatch over the period 1996–2009, or has "habituation" occurred? Does the failure of a few pingers in a given fishing set affect bycatch? Are pingers linked to pinniped depredation of catch in the fishery?

## Methods

Fishery observers were placed onboard fishing vessels to collect data on incidental entanglement and mortality of protected species, along with data on the gear characteristics of each set fished (net length, number of pingers, extender length, pinger functionality) and on the catch of fish species. From 1990 to 2009, over 8,000 fishing sets were observed (Table 1). An attempt was made to sample at least every fifth vessel trip, with an overall goal of 20% observer coverage in the fishery (Julian & Beeson, 1998; Carretta et al. 2004). It is not practical to observe every vessel in the fishery, because some smaller vessels lack berthing space for observers. Nets in this fishery are approximately 1,800-m (1 nautical mile) long and 65-m deep, with mesh sizes ranging from 35 to 60 cm. Nets are fished for approximately 12 h from dusk until dawn and are suspended from floats so that the tops of the nets are at 11-22-m depth and the bottoms are at 75-90-m depth. Fishing regulations require that acoustic pingers be attached every 91 m along the floatline and leadline of the net and that nets be fished at a minimum depth of 10.9 m with the use of "extenders" (Federal Register, 1997). Thus, the average 1,800-m net contains approximately 40 pingers, with floatline and leadline pingers spatially "staggered" to provide acoustic coverage over the entire area of the net. Pingers emit pulsed tones with source levels of 135 dB RMS; re: 1 µPa @ 1 m, fundamental operating frequencies of 10-12 kHz (with harmonics to 80 kHz), a pulse duration of 300 ms, and a pulse interval of 4 s. Additional pinger details have previously been described by Barlow and Cameron (2003).

Pinger efficacy on bycatch reduction was evaluated by comparing proportions of fishing sets with and without bycatch for a variety of gear and set situations. The characteristics of the gear and set variables used in our analyses are summarized in Table 2, and abbreviations for all variables are used throughout this paper. Statistical comparisons of set proportions with and without bycatch in this paper are based on Fisher's exact test, with 2 × 2 contingency tables (no bycatch versus ≥1 bycatch events per set). The proportion of sets with and without bycatch was compared for sets fished without pingers and sets with ≥30 pingers for the years 1990 through 2009. Occasionally, sets of less than 1,500 m in length were fished, with fewer than 30 pingers (referred to as "short

**TABLE 1**

Summary of sets observed and estimated fishing effort in the California drift gillnet fishery, 1990–2009.

| Year | Observed (No Pingers) | Observed Sets (With Pingers) | Estimated Total Sets Fished (and Fraction Observer Coverage) | Fraction of Observed Sets With Pingers |
|---|---|---|---|---|
| | | | | |
| 1990 | 178 | n/a | 4,078 (0.043) | n/a |
| 1991 | 470 | n/a | 4,778 (0.098) | n/a |
| 1992 | 596 | n/a | 4,379 (0.136) | n/a |
| 1993 | 728 | n/a | 5,442 (0.133) | n/a |
| 1994 | 759 | n/a | 4,248 (0.178) | n/a |
| 1995 | 572 | n/a | 3,673 (0.155) | n/a |
| 1996 | 275 | 146 | 3,392 (0.124) | 0.346 |
| 1997 | 304 | 388 | 3,039 (0.227) | 0.560 |
| 1998 | 14 | 573 | 3,353 (0.175) | 0.976 |
| 1999 | 2 | 524 | 2,634 (0.199) | 0.996 |
| 2000 | 0 | 444 | 1,936 (0.229) | 1.00 |
| 2001 | 1 | 338 | 1,665 (0.203) | 0.997 |
| 2002 | 0 | 360 | 1,630 (0.220) | 1.00 |
| 2003 | 0 | 298 | 1,467 (0.203) | 1.00 |
| 2004 | 0 | 223 | 1,084 (0.205) | 1.00 |
| 2005 | 0 | 225 | 1,075 (0.209) | 1.00 |
| 2006 | 0 | 266 | 1,433 (0.185) | 1.00 |
| 2007 | 1 | 203 | 1,241 (0.164) | 0.995 |
| 2008 | 0 | 149 | 1,103 (0.135) | 1.00 |
| 2009 | 0 | 101 | 761 (0.132) | 1.00 |
| All years | 3,900 | 4,238 | 52,411 (0.155) | 0.52 |

Pingers were first utilized in the fishery in 1996.

## TABLE 2

Variables used in the prediction of depredation (damage to swordfish catch) events in the drift gillnet fishery.

| Variable Name | Variable Description | Range of Values |
|---|---|---|
| *TotCatch* | Total catch of swordfish, number of individuals | 1-21 |
| *Month* | Month of set[a] | 1-6 |
| *Lat* | Latitude | ≤34.5 |
| *Lon* | Longitude | ≤120.00 |
| *DeckLght* | Were the main deck lights left on all night? 0 = No, 1 = Yes | 0, 1 |
| *Soak* | Number of hours that net was left to soak overnight (time fished) | 2-62 |
| *DepthMesh* | Number of meshes from top to bottom of net | 36-160 |
| *LengthNet* | Total length of net in meters | 914-1,828 |
| *Genr* | Was the generator engine left on all night? 0 = No, 1 = Yes | 0, 1 |
| *Mesh* | Mesh size in cm | 40-55 |
| *Random* | Random integer | 1-6 |
| *Main* | Was the main engine left on all night? 0 = No, 1 = Yes | 0, 1 |
| *Sonr* | Was the vessel's sonar left on all night? 0 = No, 1 = Yes | 0, 1 |
| *Patl* | Did the vessel patrol the length of net while it soaked, or did vessel remain stationary at one end of net? | 0, 1 |
| *Beau.Pul* | Beaufort sea state when the net was retrieved | 0-6 |
| *NumPing* | Number of acoustic pingers attached to the net | 0 or 30-42 |
| *Extnd* | Length (in m) of the line which joins the cork line and surface floats (how deep below the surface was the net fished?) | 4-18 |
| *NumLght* | Number of lightsticks attached to the net | 0-25 |
| *DepthWater* | Water depth in meters when net was retrieved | 100-1,902 |
| *Set* | Sequential set fished during a vessel's fishing trip | 1-9 |

Variables are ranked in order of importance as determined by the algorithm Random Forest for the year 1997, shown in Figure 8.

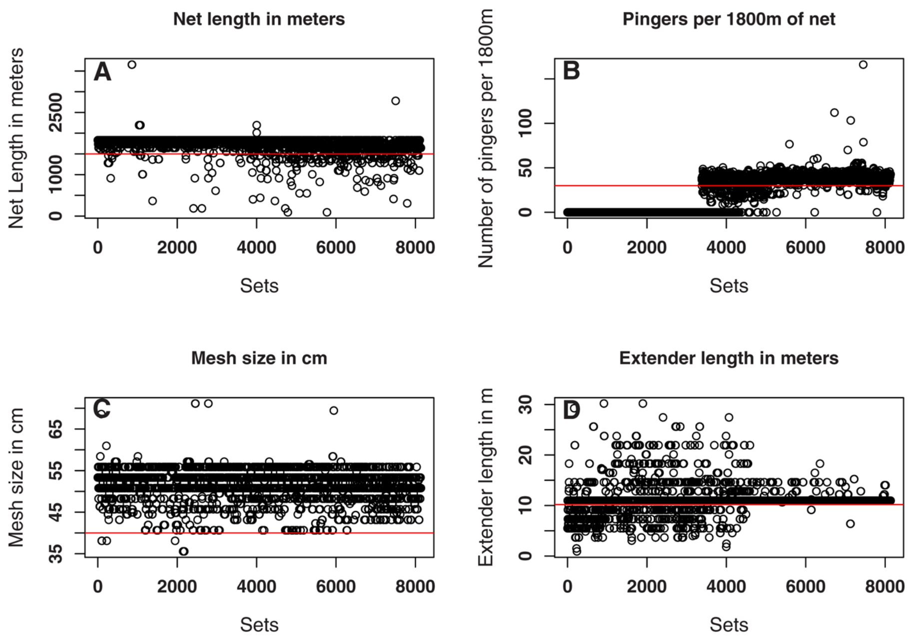
[a]The variable "*Month*" was recoded to correct for circularity of data values and represents the sequential month of fishery activity from August (=1) to January (=6) during the fishing season.

sets"). Our analyses of sets with pingers include only those that were a minimum of 1,500 m in length with ≥30 pingers. Based on the observed variability of gear variables for all sets (Figure 1), we further pared data by including only sets with the following criteria: net soak time (*Soak*) ≥8 and ≤20 h, extender lengths (*Extnd*) ≥10.9 m, and mesh size (*Mesh*) ≥40 cm. After omitting sets not meeting these criteria, 4,073 sets remained for analysis (1,281 sets without pingers and 2,792 sets with ≥30 pingers). To avoid confounding effects of spatial changes in fishing effort during the period of pinger use, we excluded sets fished inside of and north of a time/area closure implemented in 2001 to protect leatherback turtles (Figure 2). We tested the alternative hypothesis (one-tailed) that the proportion of sets with bycatch was lower in sets with ≥30 pingers. The species categories *all cetaceans*, *all pinnipeds*, and the following individual species were tested: short-beaked common dolphin (*Delphinus delphis*), northern right whale dolphin (*Lissodelphis borealis*), Pacific white-sided dolphin (*Lagenorhynchus obliquidens*), California sea lion (*Zalophus californianus*), and northern elephant seal (*Mirounga angustirostris*). We excluded species with fewer than 10 total bycatch events. Beaked whale species of the genera *Berardius*, *Mesoplodon*, and *Ziphius* are included in the results for the species category *all cetaceans* but were not tested separately, as results for this group are reported in Carretta et al. (2008) with larger sample sizes. Pingers were not used in this fishery prior to the 1996–1997 experiment (Barlow & Cameron, 2003), and pingers have been used in >99% of all observed sets since 1998 (Table 1; Figure 3). For these reasons, we are unable to assess the potential *year effect* on bycatch rates and pinger effectiveness (see Discussion).

We evaluated whether habituation to pingers may have occurred using the 2,792 sets with ≥30 pingers described above. Sets were divided chronologically into two time periods representing *early* and *late* periods of pinger use, resulting in 1,396 sets from the *early* years (1996–2001) and 1,396 sets from the *late* years (2001–2009). Sets that were examined in the overlap year of 2001 were independent (early 2001 vs. late 2001). We tested the null hypothesis (two-tailed) that the proportion of sets with bycatch was equal for *early* and *late* periods for the species categories *all cetaceans* and *all pinnipeds* as well as short-beaked common dolphins and California sea lions. We also estimated the variability of bycatch rates between early and late periods for *all cetaceans* with a

## FIGURE 1

(A-D) Gear characteristics for 8,138 observed drift gillnet sets fished between 1990 and 2009. Horizontal lines mark the minimum thresholds for data included in bycatch analyses (see text). Individual set data are shown in chronological order along the *x* axis, ordered by year and month.



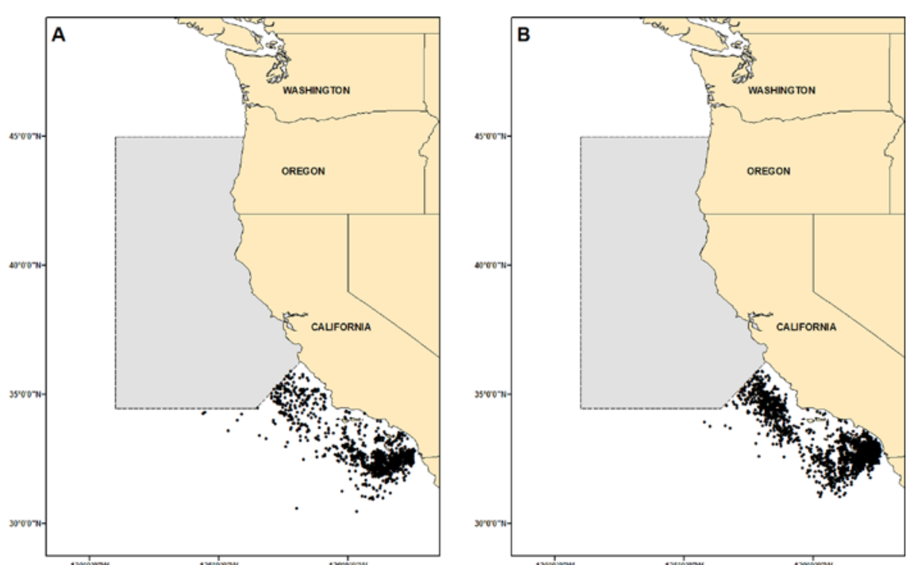

nonparametric bootstrap. Sets from each time period were sampled with replacement (using the observed effect size of 1,396 sets from each period) and a bootstrap estimate of the bycatch rate was calculated. This was done 1,000 times to provide a distribution of "pseudo-bycatch rates" for each period. We did not determine whether or not the bootstrap bycatch rates were significantly different between the two periods, as this is addressed in the results of the Fisher's exact test and a simple visual inspection of the bootstrap estimates (Figure 4).

Pinger failure sometimes occurs in the fishery, for reasons including expired batteries, water intrusion, and physical damage from fishing operations. Beginning in 2001, fishery observers were instructed to listen to each pinger during the first set retrieval of a fishing trip. If all pingers were functioning, pinger functionality was coded as "Yes", otherwise "No" if one or more nonfunctioning pingers were found. Observers also recorded notes for sets with nonfunctioning pingers, including a count of the number of failed pingers and their relative locations on the net (e.g., floatline vs. leadline). The effect of pinger failure on bycatch was evaluated by comparing proportions of sets with bycatch for sets with all pingers functioning versus sets with ≥1 nonfunctioning pingers. Between 2001 and 2009, there were 502 observed sets with ≥30 pingers where pinger functionality was recorded. Comparisons were limited to the species category *all cetaceans* because sample sizes were too small for other species/categories.

Depredation of swordfish catch by California sea lions in the fishery has sometimes been blamed on attraction to acoustic pingers, otherwise known as the "dinner bell effect" (Dawson, 1994). Depredation of swordfish catch was infrequently observed between 1991 and 1996 (<5% of sets), but there has been a marked increase in depredation since 1997 (>15-20% of sets), coinciding with the second

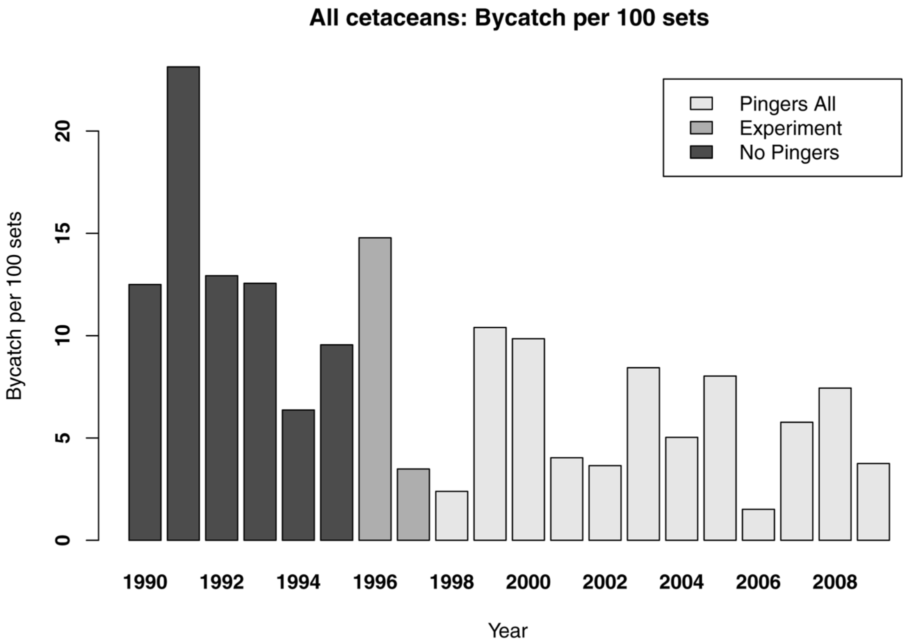
## FIGURE 2

Locations of observed drift gillnet fishing sets 1990–2009, used in pinger efficacy analyses. Shown are locations of 1,281 sets fished without pingers (A) and 2,792 sets fished with ≥30 pingers (B). Gray region represents leatherback turtle conservation area closed to fishing between 15 August and 15 November since 2001.



## FIGURE 3

Annual rates of cetacean bycatch, 1990–2009. Three periods of pinger use in the fishery are shown: 1990–1995 (no pingers), 1996–1997 (experimental pinger use), and 1998–2009 (mandatory pinger use on all sets). Only sets fished outside of the closure area implemented in 2001 (see text) are included in this figure.

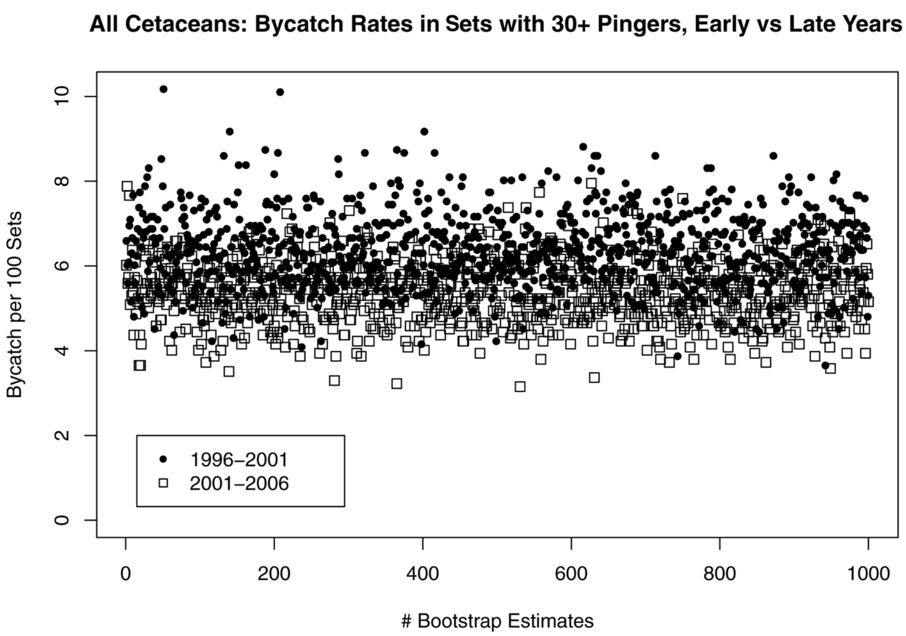



year of experimental pinger use (Figure 5). Pinger use has been mandatory in the fishery since late 1997 and has essentially been constant since that time. Thus, the effect of pingers on depredation is difficult to assess without examining sets fished both *with* and *without* pingers. For this reason, we examined sets fished in 1997 ($n$ = 193), during the second year of the pinger experiment (Barlow & Cameron, 2003). We chose 1997 because it provided an adequate sample of sets with observed depredation, in addition to a sufficient number of sets with and without pingers. In order to be able to examine the effects of variables other than pingers on depredation, we also examined a larger sample of observed sets ($n$ = 1,357) fished during a period of mandatory pinger use from 1998 through 2009, where all sets utilized ≥30 pingers. The depredation metric investigated was "mammal damage" (Y/N) to catch of broadbill swordfish (*Xiphias gladius*). Fishery observers distinguish between mammal and shark damage to catch based on differences in damage characteristics. Shark damage is characterized by discrete, semi-circular, clean bites out of the body of the fish, while mammal damage by pinnipeds is characterized by shredding of the body of the fish. Initial examination of the fishery observer data revealed that most cases of mammal depredation on swordfish catch occurred in the southern part of the fishery area, where California sea lions are most abundant. Due to the observed geographic bias in depredation, we selected a subset of data within the southern end of the fishery area for analysis (Figures 6 and 7).

The effect of pingers on depredation was evaluated by two methods. First, we tested the null hypothesis (two-tailed) that the proportion of sets with depredation was equal for sets without pingers and sets with ≥30 pingers during 1997, using Fisher's exact test. Sets in the 1997 depredation analysis included those without pingers ($n$ = 69) and sets with ≥30 pingers

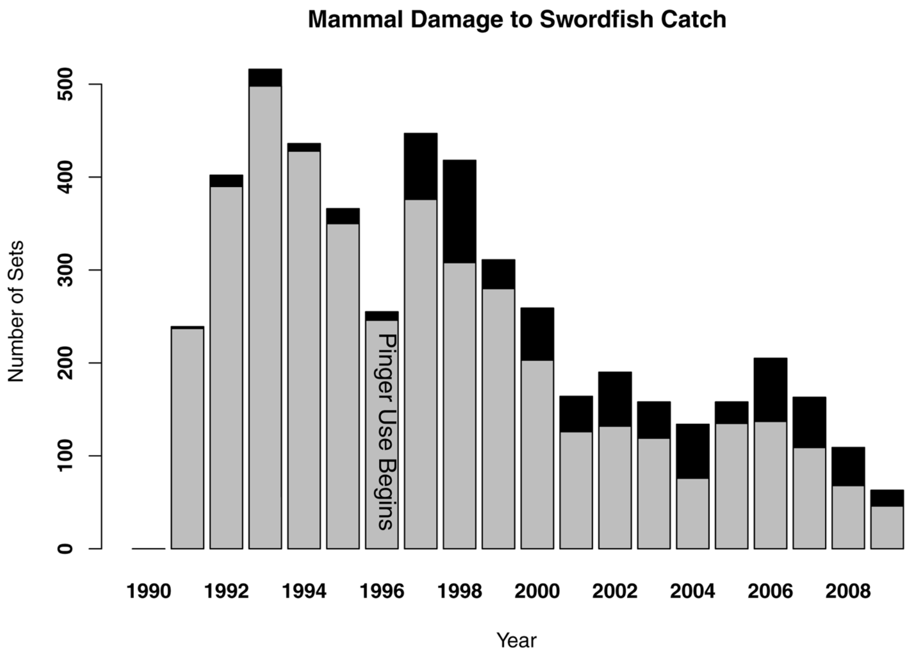
## FIGURE 4

Bootstrap estimates of cetacean bycatch rates during early (1996–2001) and late (2001–2009) periods of pinger use in the fishery.



## FIGURE 5

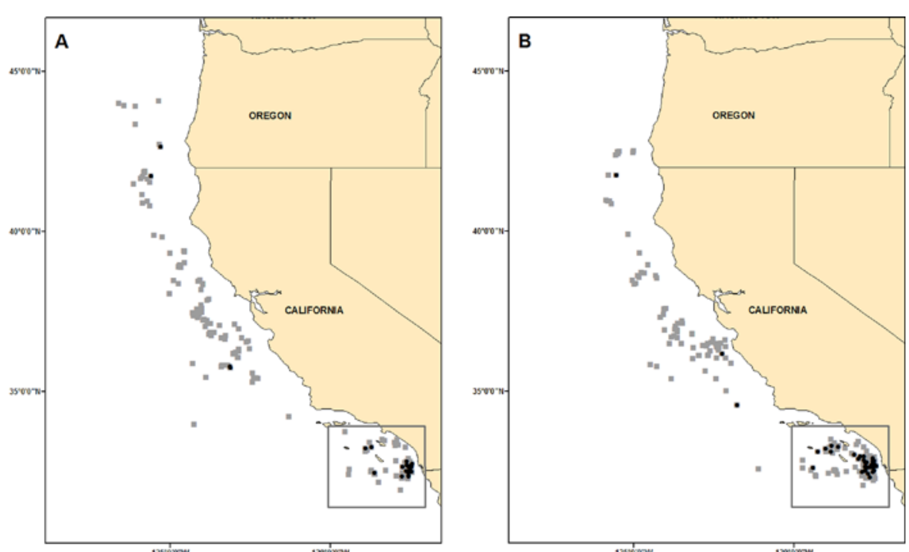
Observed occurrence of mammal damage (depredation) to swordfish catch in the drift gillnet fishery, 1990–2009. Light bars represent sets without mammal damage and dark bars represent sets with mammal damage. Damage status was not recorded in 1990.




($n$ = 124) (Figure 6). We also investigated the effects of vessel, gear, and environmental variables on depredation for the 1997 experimental year and the period of mandatory pinger use from 1998 to 2009, using the machine-learning method *Random Forest* (Breiman, 2001). Random Forest is an extension of the classification and regression tree (CART) method of Breiman et al. (1984), which we implemented in the programming language R (Liaw & Wiener, 2002; R Development Core Team, 2006). The method creates multiple bootstrap trees (a forest) to provide consensus predictions for novel input data. Our goal was to test a suite of variables to assess if they were individually or collectively useful in predicting depredation. We examined 20 variables, including the number of acoustic pingers (*NumPing*), latitude (*Lat*), longitude (*Lon*), total swordfish catch (*TotCatch*), and a random integer (*Random*) as a calibration of variable importance (Table 2). We treated depredation as a two-class prediction problem, where the classes to be predicted were depredation = Y/N. A forest of 1,000 classification trees was built from fishing sets inside the box shown in Figures 6 and 7. Each tree was constructed using two thirds of the available sets (randomly selected, without replacement) and cross-validation of each tree was accomplished by predicting the depredation status for the one third of the sets not used in tree construction (referred to as the "out-of-bag" sample). The out-of-bag sets are introduced to each tree, predictions are made, and an overall forest error rate is calculated as the average error rate of all individual trees. Tree construction was accomplished by randomly sampling (without replacement) an equal number of sets ($n$ = 15) with and without depredation. This effectively made the prediction task equivalent to predicting the flip of a fair coin if all variables were uninformative. Our samples (individual sets) may represent multiple sets within a single

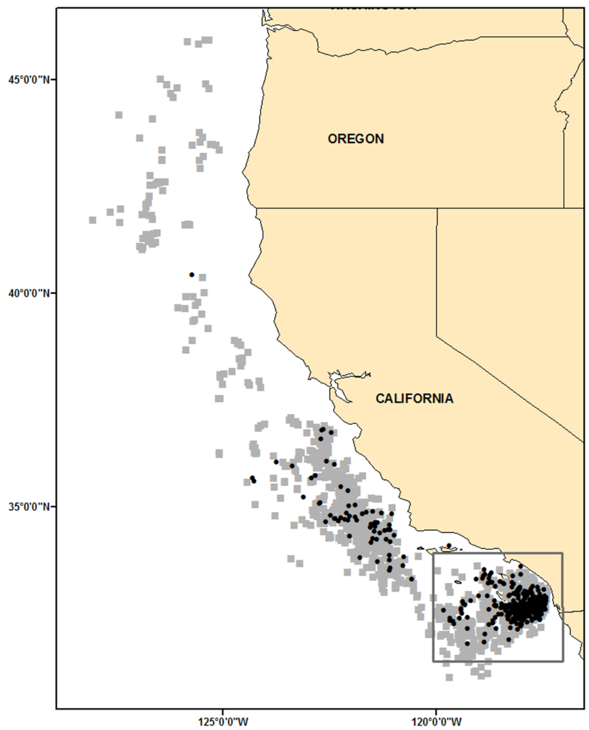
## FIGURE 6

Locations of sets fished without pingers (A) and with ≥30 pingers (B), where the status of mammal damage to swordfish catch was recorded during 1997. Gray squares represent observed sets without mammal damage to swordfish catch and dark circles represent sets with mammal damage to swordfish catch. The rectangle bounds those sets that were included in the depredation analysis.



## FIGURE 7

Locations of sets fished from 1998 to 2009 where ≥30 pingers were fished and the status of mammal damage to swordfish catch was recorded. Gray squares represent observed sets without mammal damage to swordfish catch and dark circles represent sets with mammal damage to swordfish catch. The rectangle bounds those sets that were included in the depredation analysis.




fishing trip, with potential correlation between depredation events within a trip. To eliminate trip correlation, we created separate Random Forests from odd- and even-numbered fishing trips, respectively. The "odd trip" forest was used to predict "even trip" data, and vice versa. Error rates for each depredation category were calculated as the aggregate error rate of both forest predictions on novel data and summarized as a confusion matrix. Variable importance was assessed within Random Forest through a routine that randomizes (swaps) variable values between records. Variables are randomized one at a time, trees are built from the randomized data, and out-of-bag error rates are generated as described above. Variables are then "ranked" by importance, with the "most important" variables represented by the greatest decline in predictive performance under the condition of randomization.

## Results

Following gear regulations in 1997 requiring pinger use and minimum extender lengths, fishermen have been largely compliant in meeting these requirements (Table 1; Figure 1). Over 99% of all observed sets since 1998 have utilized the required number of pingers per length of net and have adhered to minimum extender length requirements. Compliance is based on observed vessels only, as some smaller vessels are "unobservable" because they lack berthing space for observers (see Discussion).

Although there has been considerable interannual variability in bycatch rates in the fishery, it is apparent that bycatch rates in sets with pingers are considerably lower than in sets without (Figure 3). The proportion of sets with cetacean bycatch was significantly lower ($p = 6.7 \times 10^{-7}$) in sets with ≥30 pingers (4.4% of sets with bycatch) than in sets without pingers (8.4% of sets) (Table 3). Among the individual cetacean species tested, only short-beaked common dolphin ($n$ = 164 sets with bycatch) had significantly lower bycatch ($p = 2.0 \times 10^{-4}$) in sets with ≥30 pingers (3.2% of sets) than in sets without pingers (5.7% of sets). Consistent with the findings of Barlow and Cameron (2003), bycatch rates of northern right whale dolphin ($n$ = 19 bycatch events, $p$ = 0.893) and Pacific white-sided dolphin ($n$ = 14, $p$ = 0.115) were not significantly different between sets without pingers and sets with ≥30 pingers, possibly due to small sample sizes. Beaked whale bycatch has not been observed in this fishery since 1995, the last full year of fishing without acoustic pingers. Over 4,000 fishing sets with pingers have been observed since 1996 without beaked whale bycatch, compared with 33 beaked whale entanglements in 3,300 fishing sets without pingers between 1990 and 1995 (Carretta et al., 2008). Pinniped bycatch was not significantly different ($p$ = 0.141) between sets with ≥30 pingers (3.1% of sets) and sets without pingers (3.8% of sets). However, opposite patterns were observed for California sea lions and northern elephant seals. Sea lions were entangled more frequently in sets with ≥30 pingers (2.6% of sets), compared to sets without pingers (1.6% of sets, $p$ = 0.988).

## TABLE 3

Number of sets with and without bycatch for selected species/species groups.

| Species | No Pingers | | | ≥30 Pingers | | | |
|---|---|---|---|---|---|---|---|
| | No Bycatch | Bycatch ≥ 1 | Individuals per 100 Sets | No Bycatch | Bycatch ≥ 1 | Individuals per 100 Sets | Fisher's Exact Test *p* Value |
| California sea lion *Zalophus californianus* | 1,261 | 20 | 1.5 | 2,719 | 73 | 2.9 | 0.988 |
| Northern elephant seal *Mirounga angustirostris* | 1,250 | 31 | 2.4 | 2,779 | 13 | 0.46 | $1.1 \times 10^{-7}$ |
| Common dolphin, short-beaked *Delphinus delphis* | 1,208 | 73 | 7.6 | 2,701 | 91 | 4.0 | $2.0 \times 10^{-4}$ |
| Pacific white-sided dolphin *Lagenorhynchus obliquidens* | 1,274 | 7 | 0.62 | 2,785 | 7 | 0.35 | 0.115 |
| Northern right whale dolphin *Lissodelphis borealis* | 1,277 | 4 | 0.31 | 2,777 | 15 | 0.75 | 0.893 |
| All cetaceans | 1,173 | 108 | 11.4 | 2,667 | 125 | 5.9 | $6.7 \times 10^{-7}$ |
| All pinnipeds | 1,232 | 49 | 4.0 | 2,705 | 87 | 3.5 | 0.141 |

Sets are divided among those without pingers and those where ≥30 pingers were used. The Fisher's exact test significance level for the one-tailed alternative hypothesis (sets with ≥30 pingers have lower proportions of bycatch) is given in the last column. The species categories "all cetaceans" and "all pinnipeds" include bycaught animals identified to species or genera in this table, and other species for which fewer than 10 total bycatch events were recorded (e.g., unidentified cetacean, Risso's dolphin, unidentified pinniped).

Northern elephant seals were entangled far less frequently in sets with ≥30 pingers (0.5% of sets) than in sets without pingers (2.4% of sets, $p = 1.1 \times 10^{-7}$).

Habituation to pingers is not apparent in this fishery: the proportion of sets with bycatch was not significantly different between *early* and *late* periods of pinger use for *all cetaceans* ($p = 0.583$), *all pinnipeds* ($p = 0.827$), short-beaked common dolphin ($p = 0.522$), and California sea lions ($p = 0.235$) (Table 4; Figure 4). Bycatch rates of cetaceans and pinnipeds were lower in the *late* period of pinger use (Table 4; Figure 4), although bycatch rates of California sea lions were 18% higher during the *late* period.

Pinger failure was recorded in 19 of the 502 sets (3.7%) examined from

## TABLE 4

Number of sets with and without bycatch for *early* (1996–2001) and *late* (2001–2009) pinger periods.

| Species | 1996–2001 (*n* = 1,396) | | | 2001–2009 (*n* = 1,396) | | | |
|---|---|---|---|---|---|---|---|
| | *Early* Sets No Bycatch | *Early* Sets Bycatch ≥ 1 | Individuals per 100 Sets | *Late* Sets No Bycatch | *Late* Sets Bycatch ≥ 1 | Individuals per 100 Sets | Fisher's Exact Test *p Value* |
| Common dolphin, short-beaked *Delphinus delphis* | 1,347 | 49 | 4.5 | 1,354 | 42 | 3.5 | 0.522 |
| California sea lion *Zalophus californianus* | 1,365 | 31 | 2.7 | 1,354 | 42 | 3.2 | 0.235 |
| All cetaceans | 1,330 | 66 | 6.5 | 1,337 | 59 | 5.3 | 0.583 |
| All pinnipeds | 1,354 | 42 | 3.5 | 1,351 | 45 | 3.4 | 0.827 |

The Fisher's exact test significance level for the two-tailed null hypothesis (proportions of sets with bycatch are equal for early and late periods of pinger use) is given in the last column. The species categories "all cetaceans" and "all pinnipeds" include bycaught animals identified to species below, all beaked whales, and other species for which fewer than 10 total bycatch events were recorded (e.g., unidentified cetacean, Risso's dolphin, unidentified pinniped).

## TABLE 5

Summary of sets with and without cetacean bycatch for 502 sets where pinger functionality was recorded.

| | Pinger Failure Occurred | | | Pingers Functional | | | |
|---|---|---|---|---|---|---|---|
| | No Bycatch | Bycatch ≥ 1 | Bycatch per 100 Sets (Individuals) | No Bycatch | Bycatch ≥ 1 | Bycatch per 100 Sets (Individuals) | Fisher's Exact Test *p Value* |
| All cetaceans | 12 | 4 | 50.0 | 471 | 15 | 4.7 | 0.002 |

The Fisher's exact test significance level for the one-tailed alternative hypothesis (sets with pinger failure have a higher proportion of cetacean bycatch) is given in the last column.

2001 to 2009 (Table 5). In sets with ≥1 nonfunctioning pingers, the proportion of sets with cetacean bycatch (4/16, 25% of sets) was significantly higher ($p$ = 0.002, one-tailed Fisher's exact test) than in sets where all pingers were functioning (15/486, 3.0% of sets). Observer notes on pinger failure were not systematically recorded, but for 12 sets with sufficient documentation, observers indicated a range of 3-26 pingers as nonfunctioning (median failure rate = 4, mean failure rate = 6.8, mean number of pingers deployed = 37). There was no apparent pattern of failure on the floatline or on the leadline in these sets. Among sets with pinger failure where the number of failed pingers was recorded, approximately 18% of deployed pingers failed. There were 32 bycatch events for which the functional status of the pinger nearest to the entangled cetacean was recorded. In these sets, the adjacent pinger was fully functional in 27 of 32 cases (84%) and nonfunctioning in 5 cases (18%).

Sea lion depredation of swordfish catch increased in 1997, coincident with the second year of the pinger experiment, when the number of nets outfitted with pingers more than doubled (Table 1; Figure 5). However, there was no difference in the proportion of sets depredated between sets without pingers and sets with ≥30 pingers (two-tailed Fisher's exact test, $p$ = 0.742; Table 6). Random Forest correctly predicted the depredation status for 63.7% (123/193) of sets observed in 1997 (Table 7) and 66.3% (899/1,357) of sets observed during 1998–2009 (Table 8). Variable importance rankings (in order of importance) returned by Random Forest indicate that in 1997, the variables *TotCatch*, *Month*, *Lat*, *Lon*, and *DeckLght* provided the most predictive power (Figure 8). The variable *NumPing* ranked 16$^{th}$ in importance out of 20 variables, less important than the variable *Random* and outperforming only the variables *Extnd, NumLght, DepthWater,* and *Set*. Among the sets with positive depredation in 1997, the vessel's deck lights remained on all night in 45 of 57 sets (79%), while sets without depredation had all-night deck light use in 86 of 136 (63%) of sets. In the larger 1998–2009 data set where *all sets* were outfitted with ≥30 pingers, the variables *Lat*, *Lon*, *TotCatch*, *Month*, and *DepthWater* provided the most predictive power (Figure 9). During 1998–2009, the variable *DeckLght* "lost importance" relative to 1997 and the rate of all-night deck light use during 1998–2009 was nearly equal among sets without depredation (83%) and sets with positive depredation (85%). Although the variable *NumPing* ranked sixth in importance in the

## TABLE 6

Summary of depredation status on swordfish catch in 1997 for sets without pingers and sets with ≥30 pingers.

| | Pingers = 0 | | | Pingers ≥30 | | | |
|---|---|---|---|---|---|---|---|
| | Depredation (Y) | Depredation (N) | Fraction Sets Depredated | Depredation (Y) | Depredation (N) | Fraction Sets Depredated | Fisher's Exact Test *p Value* |
| Sets | 19 | 50 | 0.275 | 38 | 86 | 0.306 | 0.742 |

Sets where ≥30 pingers were fished are divided among those sets where fishery observers noted any pinger failure and those where all pingers tested were functional. The Fisher's exact test significance level for the two-tailed alternative hypothesis (sets with ≥30 pingers and sets without pingers have different proportions of depredation) is given in the last column.

## TABLE 7

Prediction of depredation status from Random Forest analysis for 1997, where sets were fished either without pingers (*n* = 69) or with ≥30 pingers (*n* = 124).

| | Predicted Yes | Predicted No | % Correct Classified |
|---|---|---|---|
| Observed Yes | **38** | 19 | 66.7% |
| Observed No | 51 | **85** | 62.5% |
| All observations | | | **63.7%** |

Correct predictions are shown in bold font. Table values represent correct classification percentages for novel data, based on Random Forest algorithms described in the text.

## TABLE 8

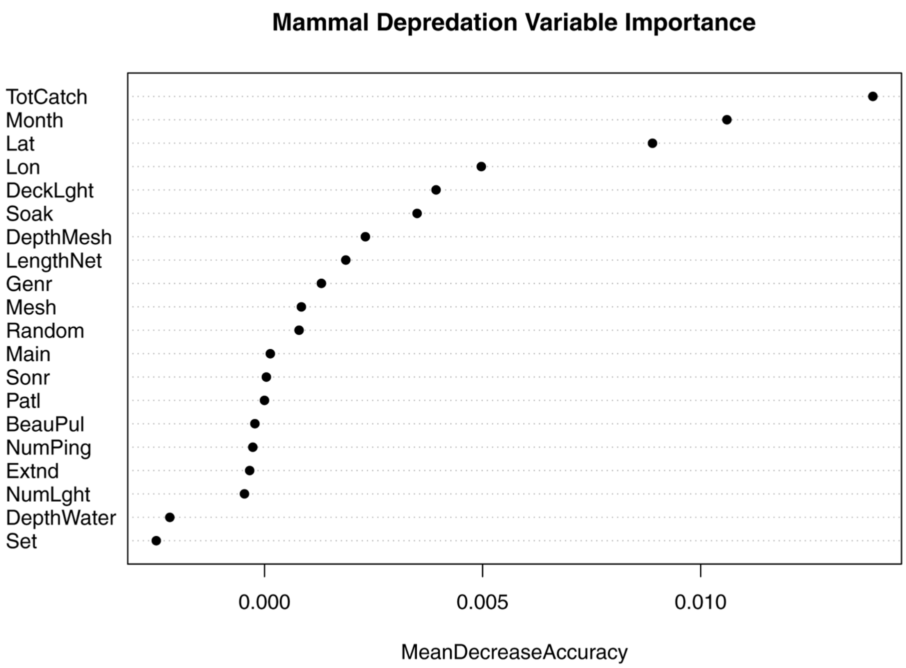
Prediction of depredation status from Random Forest analysis for 1998 to 2009, where all sets were fished with ≥30 pingers (*n* = 1,357).

| | Predicted Yes | Predicted No | % Correct Classified |
|---|---|---|---|
| Observed Yes | **317** | 147 | 68.4% |
| Observed No | 311 | **582** | 65.2% |
| All observations | | | **66.3%** |

Correct predictions are shown in bold font. Table values represent correct classification percentages for novel data, based on Random Forest algorithms described in the text.

## FIGURE 8

Variable importance measures from a Random Forest classification tree algorithm used to predict the status (Yes/No) of mammal damage to swordfish catch in 1997. Variable descriptions are provided in Table 2.




1998 to 2009 set data, only a negligible decrease in predictive accuracy was observed (<0.1%) when this variable was randomized. Overall correct classification rates were also negligibly changed when *NumPing* was omitted from analyses of both the 1997 and 1998–2009 set data.

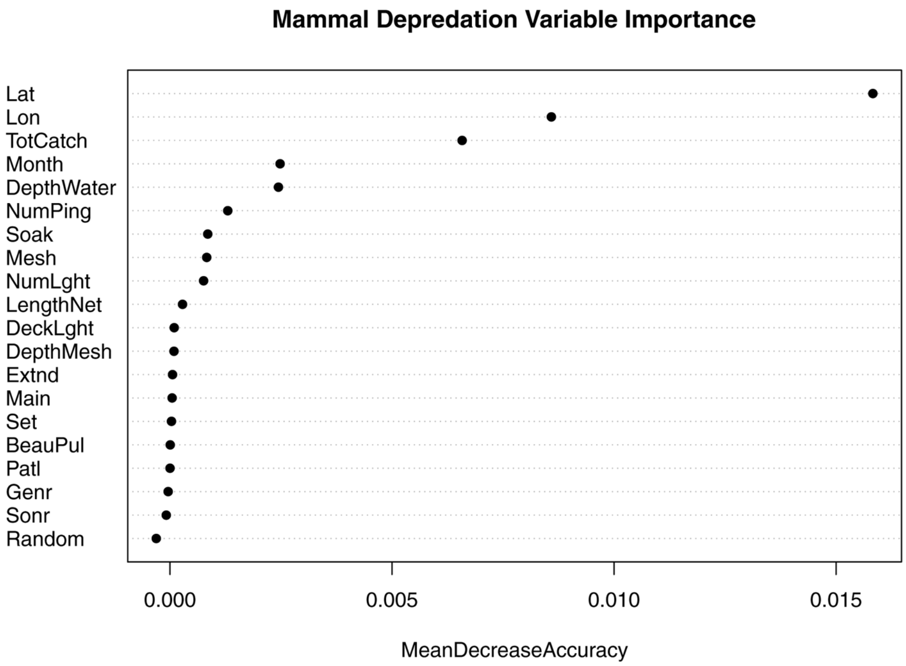
## Discussion

While pinger and extender length gear compliance for *observed* vessels is high, an increasing fraction of fishing effort in this fishery is conducted by vessels too small to accommodate observers. In 2009, 34 vessels participated in the fishery, 11 of which were unobservable. Total estimated fishing effort for the unobservable vessels in 2009 was 368 sets, or 48% of all estimated fishing effort (Carretta & Enriquez, 2010). While unobservable vessels are occasionally boarded by Coast Guard personnel to check for gear compliance, the frequency is too rare to draw conclusions from. Therefore, we cannot evaluate pinger and other gear compliance for the unobserved portion of this fishery.

Pingers continue to be effective at reducing cetacean bycatch in this fishery, though this conclusion is largely driven by short-beaked common dolphin results. Pinger effects on the bycatch of Pacific white-sided dolphin and northern right whale dolphin are unclear, as these species are infrequently entangled in the fishery. The magnitude of common dolphin bycatch reduction we report is approximately 50% for sets with pingers, which is less dramatic than the 80% reduction reported by Barlow and Cameron (2003) in the 1996–1997 experiment. The level of bycatch reduction we report is still highly significant and we do not know the reasons for the apparent dif-

## FIGURE 9

Variable importance measures from a Random Forest classification tree algorithm used to predict the status (Yes/No) of mammal damage to swordfish catch from 1998 to 2009. Variable descriptions are provided in Table 2.




ference in pinger effectiveness between the 1996–1997 experimental years and the 1998–2009 mandatory pinger years. There are sample size differences to consider, with many more sets being evaluated in the nonexperimental years (Table 1). Observed differences in pinger effectiveness between experimental and operational fishery periods was also reported by Palka et al. (2008), who attributed those differences to differences in mesh size fished. We attempted to standardize the sets used in our analysis by setting *a priori* boundaries for variables such as mesh size, extender length, net length, and soak time. Observed differences in bycatch rates between experimental and operational fishery periods in sets with pingers for our fishery could be due to variability in pinger functionality, as observers only began systematic testing of pinger functionality in 2001, several years after the pinger experiment began. In retrospect, there is no way to test this, other than noting that 2 years with full pinger use (1999 and 2000) had the highest cetacean bycatch rates during the era of mandatory pinger use and occurred prior to systematic pinger checks by observers (Figure 3). It is also worth noting that cetacean bycatch rates were lower (but not significantly so) during the late period of pinger use (2001–2009) when compared with the early period of pinger use (1996–2001) (Table 4, Figure 4).

Bycatch rates of California sea lion were higher with pinger use, but pingers do not appear to be responsible for this increase. A more likely explanation is continuing increases in California sea lion numbers (Carretta et al., 2009) coincident with a decline in fishing effort. Northern elephant seal bycatch significantly declined with pinger use, which is interesting because so little is known about the hearing capabilities of these animals. One confounding factor in assessing long-term pinger effectiveness is that the year effect on entanglement rates is unknown, because only 2 years (1996 and 1997) are characterized by a sufficient number of sets with and without pingers. This potential effect could be better addressed if the experimental design of Barlow and Cameron (2003) were applied every year, but the desire to reduce absolute bycatch levels necessitates using pingers on all sets.

Habituation to pingers by cetaceans or pinnipeds is not apparent in this fishery, based on comparisons of set proportions with bycatch for early and late periods of pinger use. For cetaceans, this conclusion is largely driven by the relatively large numbers of short-beaked common dolphin entanglements. Increases in California sea lion bycatch rates in recent years are not likely due to pinger habituation or the dinner bell effect (see below).

Failure of ≥1 pingers in 19 of 502 observed sets (3.7%) provides one measure of the minimum fraction of sets where some pinger failure may be expected in this fishery. The true rate of pinger failure is probably higher because observers may sometimes fail to detect nonfunctioning pingers. Mean pinger failure in the 19 sets where it was observed was 6.8 per set, or approximately 15–20% of the usual number (35–40) fished per set. This failure rate appears to have a significant impact on the probability of cetacean bycatch and is probably related to resulting gaps in acoustic coverage of the net. Pinger failure rates have not been published for most fisheries, but Palka et al. (2008) reported that 13% of tested pingers were nonfunctional in an Atlantic gillnet fishery during years of high pinger use.

Our assumption is that depredation of swordfish catch is caused by California sea lions, which are the most abundant pinniped in California waters (Carretta et al., 2009) and are known to depredate swordfish catch in this fishery (Miller et al., 1983). It is unlikely that cetaceans depredate catch in this fishery, as most cetaceans entangled in the fishery feed on small schooling fishes or squid too small to be entangled in drift gillnets. Increases in depredation rates in 1997 coincide with the second year of the pinger experiment and the onset of a major El Niño event (Enfield, 2001). Reduced prey availability for California sea lions associated with El Niño events (DeLong et al., 1991) may increase the likelihood of depredation on gillnets and perhaps contributed to the relatively high depredation rates seen in 1997. Depredation rates have remained high since 1997 (Figure 5), which may reflect learned behavior by sea lions and increases in their population size since that time (Carretta et al., 2009). However, pingers do not appear to be linked to depredation, based on nearly equal depredation rates in sets with and without pingers and variable importance measures from Random Forest analysis. The most important variables, in order of importance, were *TotCatch*, *Month*, *Lat*, *Lon*, and *DeckLght*, with three of five related to the timing and location of fishing activity. The importance of *TotCatch* may reflect that sea lions are attracted to nets with greater numbers of entangled swordfish, while *DeckLght* importance suggests sea lions use vessel lights as visual cues to locate nighttime fishing activity. For the larger data set of 1998–2009, the variable *DeckLght* "loses importance." For unknown reasons, fishermen began using deck lights at much higher rates beginning in 1999 (in this case *DeckLght* behaves more like a constant than a variable). The year 1997 was characterized by low rates of deck light use compared to subsequent years and use of deck lights was not recorded prior to 1996 when the pinger experiment began. Thus, outside of 1997, it is difficult to assess the importance of deck lights on depredation. Although Random Forest provides measures of variable importance, no single variable may be "statistically significant" in the traditional sense. More often, there are ensembles of "weak predictors" with *collective* predictive power, as is the current case. The variable importance score for *TotCatch* reflects a ~1% decline in predictive accuracy after randomization (Figure 8), which would not be *statistically significant* in most types of analyses. However, in the framework of prediction, "significance" is based on the *aggregate predictability of an event*, with respect to the prior probability of success if none of the variables are informative. Recall that our Random Forest was constructed with equal numbers of Y/N depredation events, reducing the problem to a binomial one, with a 0.5 probability of success if all of the variables were uninformative. In that context, the probability of correctly predicting the depredation status of at least 64% of sets is <0.005.

## Acknowledgments

We would like to thank Lyle Enriquez for managing the drift gillnet observer program. This work would not have been possible without the support of the National Marine Fisheries Service Southwest Regional Office and the cooperation of the California drift gill net vessel operators. The manuscript was improved by the reviews of Susan Chivers, Jeff Moore, and two anonymous reviewers.

## Authors:

James V. Carretta and Jay Barlow
U.S. National Marine Fisheries Service
Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration,
3333 North Torrey Pines Court,
La Jolla, CA 92037
Email: Jim.Carretta@noaa.gov;
Jay.Barlow@noaa.gov

## References

**Barlow**, J., & Cameron, G.A. 2003. Field experiments show that acoustic pingers reduce marine mammal bycatch in the California drift gillnet fishery. Mar Mammal Sci. 19:265-83. http://dx.doi.org/10.1111/j.1748-7692.2003.tb01108.x.

**Bordino**, P., Kraus, S., Albareda, D., Fazio, A., Palmerio, A., Mendez, M., & Botta, S. 2002. Reducing incidental mortality of franciscana dolphin *Pontoporia blainvillei* with acoustic warning devices attached to fishing nets. Mar Mammal Sci. 18(4):833-42. http://dx.doi.org/10.1111/j.1748-7692.2002.tb01076.x.

**Breiman**, L. 2001. Random Forests. Mach Learn. 45:5-32. http://dx.doi.org/10.1023/A:1010933404324.

**Breiman**, L., Friedman, J., Olshen, R., & Stone, C. 1984. Classification and Regression Trees. Pacific Grove, CA: Wadsworth Publishing Co., 358 p.

**Carretta**, J.V., Barlow, J., & Enriquez, L. 2008. Acoustic pingers eliminate beaked whale bycatch in a gill net fishery. Mar Mammal Sci. 24(4):956-61.

**Carretta**, J.V., & Enriquez, L. 2010. Marine mammal and sea turtle bycatch in the California/Oregon swordfish and thresher shark drift gillnet fishery in 2009. Southwest Fisheries Science Center Administrative Report LJ-10-03. 11 p.

**Carretta**, J.V., Forney, K.A., Lowry, M.S., Barlow, J., Baker, J., Johnston, D., … Carswell, L. 2009. U.S. Pacific Marine Mammal Stock Assessments: 2009. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-453. 326 p.

**Carretta**, J.V., Price, T., Petersen, D., & Read, R. 2004. Estimates of marine mammal, sea turtle, and seabird mortality in the California drift gillnet fishery for swordfish and thresher shark, 1996–2002. Mar Fish Rev. 66(2):21-25.

**Cox**, T.M., Read, A.J., Solow, A., & Tregenza, N. 2001. Will harbour porpoises (*Phocoena phocoena*) habituate to pingers? J Cetacean Res Manage. 3(1):81-6.

**Dawson**, S.M. 1994. The potential for reducing entanglement of dolphins and porpoises with acoustic modifications to gillnets. Rep Int Whal Commn Special Issue. 15:573-8.

**DeLong**, R.L., Antonelis, G.A., Oliver, C.W., Stewart, B.S., Lowry, M.S., & Yochem, P.K. 1991. Effects of the 1982–1983 El Niño on several population parameters and diet of California sea lions on the California Channel Islands. In: Pinnipeds and El Niño: Responses to environmental stress, eds. Trillmich, F., & Ono, K.A., pp. 166-72. Berlin: Springer-Verlag.

**Enfield**, D.B. 2001. Evolution and historical perspective of the 1997–1998 El Niño–Southern Oscillation Event. Bull Mar Sci. 69(1):7-25.

**Federal Register.** 1997. Taking of marine mammals incidental to commercial fishing operations; Pacific offshore cetacean take reduction plan regulations. Final Rule. 62(192):51805-14.

**Gearin**, P.J., Gosho, M.E., Leake, J.L., Cooke, L., DeLong, R.L., & Hughes, K.M. 2000. Experimental testing of acoustic alarms (pingers) to reduce bycatch of harbour porpoise, *Phocoena phocoena*, in the state of Washington. J Cetacean Res Manage. 2(1):1-9.

**Julian**, F., & Beeson, M. 1998. Estimates of marine mammal, turtle, and seabird mortality for two California gillnet fisheries: 1990–95. Fishery Bulletin, U.S. 96:271-84.

**Kraus**, S. 1999. The once and future ping: Challenges for the use of acoustic deterrents in fisheries. Mar Technol Soc J. 33(2):90-3.

**Kraus**, S.D., Read, A., Anderson, E., Baldwin, K., Solow, A., Spradlin, T., & Williamson, J. 1997. Acoustic alarms reduce porpoise mortality. Nature. 388:525. http://dx.doi.org/10.1038/41451.

**Liaw**, A., & Wiener, M. 2002. Classification and regression by randomForest. R News. 2(3):18-22.

**Miller**, D.J., Herder, M.J., & Scholl, J.P. 1983. California marine mammal fishery interaction study, 1979–81. NMFS Southwest Fish. Cent., Administrative Report LJ-83-13C. 233 p.

**Palka**, D.L., Rossman, M.C., VanAtten, A.S., & Orphanides, C.D. 2008. Effect of pingers on harbour porpoise (*Phocoena phocoena*) bycatch in the U.S. Northeast gillnet fishery. J Cetacean Res Manage. 10(3):217-26.

**R Development Core Team.** 2006. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

**Trippel**, E.A., Strong, M.B., Terhune, J.M., & Conway, J.D. 1999. Mitigation of harbour porpoise (*Phocoena phocoena*) by-catch in the gillnet fishery in the lower Bay of Fundy. Can J Fish Aquat Sci. 56:113-23. http://dx.doi.org/10.1139/cjfas-56-1-113.