SCIENTIFIC COMMITTEE
THIRD REGULAR SESSION

13-24 August 2007
Honolulu, United States of America

EFFECTS OF GEAR CHARACTERISTICS ON THE PRESENCE OF BIGEYE TUNA (*THUNNUS OBESUS*) IN THE CATCHES OF THE PURSE-SEINE FISHERY OF THE EASTERN PACIFIC

WCPFC-SC3-FT SWG/IP-1

Paper prepared by

Cleridy E. Lennert-Cody[1,], Jason J. Roberts[2] and Richard J. Stephenson[3]

1: IATTC, La Jolla, California, USA
2: Duke University Marine Geospatial Ecology Laboratory, Durham, North Carolina, USA
3: contact by way of primary author

SUBMITTED TO ICES JOURNAL OF MARINE SCIENCE, JUNE 2007

DO NOT CITE WITHOUT AUTHOR'S PERMISSION

(Revised version of SAR-8-09c presented at IATTC SAWG, May 2007**)**

Effects of gear characteristics on the presence of bigeye tuna (*Thunnus obesus*) in the catches of the purse-seine fishery of the eastern Pacific Ocean

Cleridy E. Lennert-Cody, Jason J. Roberts, and Richard J. Stephenson

Overfishing of bigeye tuna in the eastern Pacific Ocean has motivated a search for pratical means of reducing bigeye catch. We develop a classification algorithm for the presence/absence of bigeye in purse-seine sets on floating objects, the dominant mode of purse-seining for bigeye, using the tree-based method random forests to explore the effects of gear characteristics. Although the location of the set was the strongest determinant of the presence of bigeye catch with these data, in some areas, bigyeye were more likely to be caught on floating objects with greater underwater depths and with deeper purse-seine nets. Misclassified sets were found to be concentrated within certain vessels, suggesting that the existence of additional 'vessel effects' on the presence of bigeye which may be amenable to further study. Results indicate that fishermen can avoid catching bigeye in some areas by changing the depth of the material hanging from the floating object and the actual fishing depth of the net, or by moving to other fishing areas. However, we believe that the complex nature of gear and environmental interactions, and the impact of gear restrictions on the catches of tuna species other than bigeye, argue

against the feasibility of fishery-wide gear restrictions.

Keywords: bigeye, classification, gear effects, purse-seine, random forest, tuna

Cleridy E. Lennert-Cody: Inter-American Tropical Tuna Commission, 8604 La Jolla Shores Drive, La Jolla, California USA. Jason J. Roberts: Duke University Marine Geospatial Ecology Laboratory, Durham, North Carolina, USA. Richard J. Stephenson: owner/fishing captain of the purse-seiner F/V *Connie Jean*, San Diego, California, USA. Correspondence to Cleridy E. Lennert-Cody: tel: +1 858 546 7190; fax: +1 858 546 7133; e-mail:clennert@iattc.org.

Introduction

Despite recent efforts to improve the status of bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean (EPO), the most recent stock assessment (Maunder and Hoyle, 2007) indicates that fishing mortality remains too high to be sustainable. Management of the bigeye tuna population in the EPO is complicated (Maunder and Harley, 2006). Bigeye tuna is caught predominantly by longline and purse-seine gears, with approximately half the catch of bigeye occurring in purse-seine sets made on floating objects (IATTC, 2006a). However, the dominant tuna catch in these floating object sets is skipjack tuna (*Katsuwonus pelamis*), and the skipjack population in the EPO is estimated to be healthy. Thus, one of the management goals has been to find means of reducing catches of bigeye tuna in the purse-seine fishery on floating objects while minimizing losses of skipjack catch. Implementation of seasonal closures since 2000 affecting both purse-seine and longline fisheries have not provided an adequate reduction in bigeye tuna fishing mortality. Given that operationally-feasible time-area closures are unlikely to result in a sustainable bigeye fishery (Harley and Suter, 2007), other options, including gear modifications, are currently being explored (Maunder, 2006).

Floating object sets that capture bigeye tuna appear to be concentrated within vessels. All three tuna species, bigeye, skipjack and yellowfin (*Thunnus albacares*), that are targeted by the purse-seine fishery in the EPO, are caught in floating object sets. Based on data collected by Inter-American Tropical Tuna Commission (IATTC) observers during 2001-2005 (see Data section), approximately 54% of floating object sets by large vessels (> 363 mt fish-carrying capactiy) yielded catches of bigeye tuna, compared to 81% for yellowfin tuna and 93% for skipjack tuna. Yet total catches of bigeye tuna on floating objects are greater than those of yellowfin tuna on floating objects. Of the 158 vessels represented by these data, 28% did not catch any bigeye tuna. Floating object sets, regardless of the catch, tended to be concentrated within vessels. However, even accounting for this, the relationship between numbers of sets on floating objects and numbers of sets on floating objects that caught bigeye, by vessel, is not linear (Figure 1). This is contrary to what would be expected if sets per vessel that caught bigeye tuna were proportional to total sets. Although the specific vessels making the most sets on bigeye tuna changes from year to year, one possible explanation for this pattern is that each year some vessels increase their chances of catching bigeye tuna by where and when they fish, the gear that they use, and/or combinations of these options.

It has been suggested that some gear characteristics affected tuna catches during the

1993-1998 period of this fishery (Lennert-Cody and Hall, 2000). However, this was a time when the fishery on floating objects was in transition from a fishery on flotsam (*e.g.*, tree limbs) in nearshore areas to a largely fish-aggregating device (FAD)-associated fishery further offshore. Although partial confounding of gear and environmental factors is to be expected with fishery-dependent data, particularly in the EPO where there exist strong environmental gradients (e.g., Kessler, 2006), the need to find operationally-feasible means of reducing fishing mortality of bigeye tuna in the current FAD-dominated fishery, the availability of more comprehensive environmental data (*e.g.*, ocean color), and the availability of improved descriptive statistical techniques for large data sets (e.g., Berk, 2006) suggests that gear effects warrant further study. Because vertical stratification of tuna species around floating objects has been noted by fishermen and fisheries observers, and identified through research (Schaefer and Fuller, 2002), with bigeye deeper in the water column than skipjack, the current analysis focuses on aspects of the fishinging gear that might interact with vertical structure of the object-associated community, thereby affecting catch composition.

In this manuscript we present an analysis of the presence/absence of bigeye tuna catch in purse-seine sets on floating objects for the 2001-2005 period. Given the results of the most recent stock assessments for bigeye tuna (Maunder and Harley, 2006; Maunder and Hoyle, 2007), and the fact that almost half of floating object sets caught no bigeye, we focus on understanding processes that led to any amount of bigeye catch. The tree-based method random forests (Breiman, 2001) was used to build a classification algorithm for sets with and without bigeye tuna catch, placing more emphasis on correctly predicting the presence of catch. With this analysis, we attempt to determine: 1) how well the presence of bigeye catch can be described by characteristics of the environment, and the fishing operation and gear; 2) whether there is spatial structure in any gear effects; and, 3) the extent to which there may exist additional ‘vessel effects’ on the presence of bigeye catch beyond the explanatory ability of the predictors included in this analysis.

Data

Data used in this analysis are from purse-seine sets on floating objects collected by IATTC observers aboard large vessels (> 363 mt fish-carrying capacity) between 2001 and 2005. The IATTC sampled over 67% of all fishing trips of large vessels over this five-year period (e.g., IATTC, 2006b), amounting to over 75% IATTC observer coverage of floating object sets of large vessels. The IATTC onboard observer program is described in Bayliff (2001). Data were limited to sets that caught some amount of at least one of the three target species to avoid observations for which the fish escaped capture. Repeated sets on the same floating object, where they could be identified, were excluded (> 85% of sets were deemed ‘first’ sets). Data collected prior to 2001 were not included in this analysis to avoid potential trends in biases in tuna species identification. In particular, in 2000 the IATTC implemented a system for tracking tuna catch as part of the AIDCP ‘dolphin safe’ certification. As part of this process, the observer may discuss catch composition with the vessel’s fishing captain. In addition, in 2000 the IATTC passed a resolution encouraging vessels to retain all tuna catch (IATTC, 2000). This resolution has been renewed annually, but the degree of compliance is unclear (IATTC, unpublished data). Because tunas found in association with floating objects can be of small size, and hence less marketable, strict compliance with the resolution might affect a vessel's decision as to whether to initiate a set. After data processing, a total of 10,421 floating object sets was available for analysis.

Over 85% of the floating objects set upon during this five-year period were estimated to have been FADs (IATTC, 2006a). FADs may be constructed of a variety of materials, but the most typical construction is a raft (often of bamboo) with old purse-seine netting hanging underneath. FADs often carry some form of locating device (*e.g.*, radio transmitter, satellite transmitter).

To describe variability in the occurrence of bigeye tuna catch, 22 predictors were considered in this analysis (see Table 1 for details). These predictors can be divided roughly into three groups: those describing aspects of fishing operations and gear, those describing the environment, and miscellaneous predictors. There were eight predictors included to describe aspects of the fishing operations and gear, five of which relate directly to the fishing depth of the purse-seine net or the underwater depth of the floating object. The actual in-water depth of both will vary depending on a number of factors, including winds and currents. Moreover, the fishing depth of the net is determined not only by its hanging depth, but also the rate at which it descends. For a given set of environmental conditions, the descent rate of the net is a function of mesh size, dolphin safety panel use, the 'hang-in' (number of meshes per unit length along the cork line), and the weight of the purse cable and chain. Data were available on the hanging depth of the net, its mesh size, presence of a safety panel, and the duration of the period over which the net descends to its fishing depth (Green, 1969). Data were also available on the maximum underwater depth of the floating object. Environmental predictors included relate to measures of upper-ocean circulation (e.g., major currents, eddies), stratification and productivity. With the exception of sea surface temperature, environmental predictors represent climatologies estimated at set locations and dates. Location and date of the set were also included as proxies for local environmental conditions not captured by the other predictors. The two miscellaneous predictors were a proxy for the non-tuna community size at the object, and a proxy for the local floating object density.

As would be anticipated, given the opportunistic nature of the data collection process, the inshore-offshore orientation of the fishery (Figure 2), and the gradients in the oceanographic environment (Kessler, 2006), several of these predictors are partially correlated (Table 2). For example, correlation between environmental predictors and predictors such as percent fouling likely result because floating objects will have a tendency to drift offshore in many areas of the EPO. The oceanography and bathymetry of the EPO result in correlations between latitude and longitude, and many environmental predictors such as sea surface temperature, chlorophyll-a density, and mixed layer depth. In addition, some gear and operational predictors are inherently correlated. For example, larger vessels (greater fish-carrying capacity) can carry larger nets which may have greater hanging depths than smaller nets. Larger vessels can fish further offshore due to their greater fish- and fuel-carrying capacities. Examples of the spatial dependence of several gear predictors are shown in Figure 3.

The classification of each set as to the presence/absence of bigeye tuna catch was based on the catch weights. Both catch weights (loaded weights plus discards) and loaded weights are estimated by observers. We use catch weights because they may more closely reflect the ecological relationship between the object-associated community and the environment and fishing gear.

Methods of analysis

With this analysis, we want to determine: how well the presence of bigeye tuna catch can be described by characteristics of the environment, fishing operations and gear; whether there is spatial structure in any gear effects; and, whether there may exist additional ‘vessel effects’. Towards this end, the ensemble method ‘random forests’ (Breiman, 2001; Berk, 2006) was used to build a classification algorithm for the presence/absence of bigeye tuna catch. The random forest method has been demonstrated to build better classification algorithms than other methods (Breiman, 2001). In addition, the estimates of misclassification errors provided by the random forest method are true forecasting errors, and the relative costs of the two types of mistakes that can be made (predicting bigeye catch when none occurred – ‘false positive’; predicting no bigeye catch when in fact there was catch – ‘false negative’) can be easily specified. Our overall approach is similar to that of Lennert-Cody and Berk (2007).

Random forests is a tree-based algorithm that builds on the classical Classification and Regression Tree approach (CART; Breiman et al., 1984). It can be described in three conceptual steps. First, a large number of CART-like trees (a ‘forest’) are constructed, each on a different randomly selected sample from the original data. Observations not included in a particular random sample are referred to as ‘out-of-bag’ or ‘OOB.’ Second, each tree in the forest is built in a manner that is similar to a CART tree, but with two important differences: the candidate predictors that are available to define each node in the tree are a randomly selected subset of all predictors, drawn anew for each node; and, the resulting tree is not pruned. Finally, the predicted class of an observation by the forest is determined by majority vote among the individual trees for which the observation was OOB. (The predicted class of an observation from an individual tree in the forest is the dominant class on the relevant terminal node.) Details of the random forest algorithm can be found in Breiman (2001) and Berk (2006).

We use the **R** statistical computing (R Core Development Team, 2005) package *randomForest* (Liaw and Wiener, 2002) to build a random forest classification algorithm for these data. The data set was randomly divided (by year) into two parts: a training data set with 5,210 sets (2,827 sets with bigeye, 2,383 without) and a test data set with 5,211 sets (2,844 sets with bigeye, 2,367 without). All classification algorithms were built on the training data set. The test data set was used to explore 'vessel effects' as described below. In the context of the current problem, it seems reasonable to place equal, if not added, emphasis on correctly predicting the presence of bigeye tuna catch when it occurred. Thus, we consider two different relative costs: equal relative costs of false negatives and false positives, and the relative cost of false negatives three times that of false positives. The different relative costs were achieved by building forests on data sets with different proportions of presence and absence observations (*sampsize* option in the *randomForest* package). Each classification algorithm was based on 5,000 trees.

Within the combinations of environmental conditions, locations, and fishing dates in the data set, we summarize the effects of gear characteristics on the presence of bigeye catch in several ways. The relative importance of each predictor was computed as the average percent decrease in prediction accuracy on the OOB data when the predictor's values were scrambled (Liaw and Wiener, 2002; Berk, 2006). In addition, the relationship of each of the most influential gear predictors to the occurrence of bigeye catch were summarized by plotting the logit of the proportion if trees in the forest that voted for the presence of bigeye tuna versus the predictor (a form of 'partial dependence;' e.g., Hastie et al., 2001). This provides an estimate of the effect of a particular predictor on the 'probability' that a set was classified as having caught bigeye tuna, taking into account the average effects of the other predictors. To look for spatial structure in these relationships, these partial depedence plots were also constructed within each of 40 rectangular areas (10° longitude by 2.5° latitude, between 90-140°W and 12.5°S-7.5°N). The size of the rectangular areas and the overall region were selected according to the large-scale circulation patterns of the EPO (Kessler, 2006) and the spatial distribution of the floating object fishery (Figure 2).

To explore 'vessel effects,' beyond what can be described by the available predictors, we compared observed and reported set classifications of the test data set. We focus on false negatives, bigeye tuna caught but none predicted, because this type of error may indicate alternative fishing strategies that were successful with respect to bigeye tuna. Because there are different numbers of sets per vessel in the data set (Figure 1), we compare the number of misclassifications of sets that caught bigeye to that which would be predicted from a binomial distribution. The binomial parameter was taken to be the false negative rate. In other words, for each vessel we computed the probability that out of n sets that caught bigeye tuna there would have been r or more sets for which no bigeye tuna were predicted. We refer to these probabilities as 'per-vessel' probabilities. There is no convincing way to assess the extent that observations within vessels are independent, and thus we use the per-vessel probabilities as a relative measure of 'vessel effects;' the smaller the probability, the more unusual the data of that vessel with respect to the data of other vessels.

Results

The random forest classifier was reasonably successful at predicting the occurrence of sets with bigeye tuna catch (Table 3). Misclassification errors at equal relative costs for false negatives and false positives were 15% for sets that caught bigeye tuna and 18% for sets that did not (Table 3a). When emphasis was placed on the correct classification of sets with bigeye tuna (relative costs of false negatives three times that of false positives), the false negative rate decreased from 15% to 8%, while the false positive rate increased from 18% to 29% (Table 3b). (Achieving a false negative error rate of less than 8% would require higher relative costs, which may not be acceptable.) When location and date predictors were not included in the classifier, but the requirement of three to one relative costs was maintained, the false negative rate increased by 2%, while the false positive rate increased by about 6% (Table 3c).

Predictor importance shows indication of strengths among some gear and environmental predictors, even though the location of the set appeared to be the most influential in determining whether a set caught bigeye (Figure 4). Of the gear and environmental predictors included in this analysis, object depth, chlorophyll-a density, bathymetry, mixed layer depth, and sea surface temperature appeared to be the most useful for predicting the presence of bigeye tuna catch with this data set. The relative dominance of gear and environmental predictors remained largely the same when location and date predictors were not included in the classification algorithm, except that object depth became slightly less important while net depth became slightly more important (Figure 4). The overall weak levels of variable importance would be anticipated given the correlations between predictors (Table 2); when a specific predictor is not selected to define a node of a given tree, some of its predictive ability may be captured by other predictors with which it is correlated.

The marginal effects of object depth and net depth, the two most important gear predictors that directly relate to the in-water depth of the gear (Figure 4), are shown in Figure 5 for the classification algorithm with three to one relative costs (Table 3b). Overall, sets were more likely to be classified as having caught bigeye tuna the greater the object depth and the greater the net depth. However, the marginal effects decreased somewhat on the deepest objects and with the deepest nets. Within the region of 90°-140°W and 12.5°S-7.5°N, the greatest marginal effects of object depth were found between 100°-130°W along the equator and in the southern area of the fishery, and offshore north of the equator between 2.5°-5.0°N (Figure 6). Object depth appeared to have little effect on whether a set was classified as having caught bigeye tuna in the inshore areas and in the northern most areas. Similar but less pronounced spatial structure is evident in the marginal effect of net depth (Figure 7). Also evident in Figures 6-7 is the influence of set location. For example, marginal effects in the inshore areas are clearly less than further offshore, regardless of latitude and gear.

In the test data set, there were 105 vessels over the five-year period that made at least one set catching bigeye tuna. The frequency distribution of per-vessel probabilities computed for these vessels using a binomial parameter of 0.08 (Table 3b) is shown in Figure 8. Per-vessel probabilities at or close to 1.0 correspond to vessels with relatively few sets for which bigeye tuna was caught but none was predicted. These are vessels for which the relationships captured by the random forest classifier adequately describe the occurrence of bigeye. Per-vessel probabilities close to 0.0 correspond to vessels with a larger number of false negatives, relative to the number of sets in which these vessels caught bigeye tuna. The random forest classifier failed to capture some of the important aspects of the data of these vessels with the available predictors. Within this group, the data of those vessels making the most sets on bigeye tuna might prove useful for exploring the possibility of other fishing strategies.

Discussion

In this manuscript we have developed a classification algorithm for the presence/absence of bigeye tuna catch in floating object sets to explore the effects of gear characteristics on the occurrence of bigeye catch. The presence of bigeye tuna catch could be reasonably predicted from information on the set location, the environment, and the fishing operation and gear. Among the gear characteristics studied that directly relate to the in-water depth of the floating object and the purse-seine net, the maximum depth of the object below the surface and the hanging depth of the net had the greatest effect on whether bigeye tuna were caught, with catches more likely on deeper objects and with deeper nets. These gear effects were most pronounced near the equator and in the southern area of the fishery. Nonetheless, the location of the set (latitude, longitude) was the strongest determinant with this data set for the presence of bigeye tuna. False negatives (bigeye tuna caught but none predicted) were found to be concentrated to some extent within certain vessels suggesting that these vessels may also catch bigeye tuna in ways different from most of the fleet, *i.e.*, in ways poorly described by the predictors included in this analysis. This represents a form of a ‘vessel effect’ that could be amenable to further study.

Although results of this analysis are consistent with fishermen’s experience that deeper objects and deeper nets may be more likely to lead to catch of bigeye tuna in some areas, the details indicate that gear effects are complex. For example, comparison of Figures 5-7 suggests that some of the decrease in marginal effects on the deepest objects and with the deepest nets likely reflects spatial structure in gear effects. However, particularly for net depth, an additional possibility is that this decrease may be indicative of a different fishing strategy used by some larger vessels. Larger vessels (greater fish-carrying capacity), which will tend to have deeper nets (Table 2), will also have greater fishing range, and can carry a greater number of FADs. Instead of waiting for the optimal conditions to make a set so as to maximize catch on a particular FAD, these vessels may set on objects as they are encountered, a strategy made economically viable by the number of the FADs that can be placed and the period of time for which the vessels can remain at sea. In addition, many of the environmental predictors (including latitude and longitude) were correlated with gear characteristics, making it impossible to estimate gear effects independent of environmental conditions.

The results presented in this manuscript suggest that the presence of bigeye tuna catch in floating object sets exhibits characteristics consistent with some level of fishermen control. The importance of location as a predictor indicates that the presence of bigeye tuna in the catch is not a spatially random event (Figure 4). In addition, similarilites between the spatial distribution of object depth (Figure 3) and that of its marginal effect on the presence of bigeye catch (Figure 6), suggest some degree of planning on the part of fishermen. Thus, results of this study indicate that fishermen have several options available to them to try to avoid catching bigeye tuna, including changing the depth of the material hanging below the floating object and the fishing depth of the purse-seine net in certain areas of the fishery, and changing their overall fishing location.

Given the current status of bigeye tuna populations (Maunder and Harley, 2006; Maunder and Hoyle, 2007) and the operational infeasibility of spatial-temporal closures (Harley and Suter, 2007), gear restrictions might seem a reasonable option for reducing fishing mortality of bigeye tuna. However, gear restrictions would affect all vessels and all areas of the fishery, perhaps reducing catches of other tuna species, such as skipjack. Previous studies (Lennert-Cody and Hall, 2000) found some indication that catch per set of skipjack tuna increased with the hanging depth of the net. Fishery-wide restrictions on hanging depth might reduce catches of skipjack across a broad segment of the fishery, a seemingly unnecessary outcome given the focused nature of the fishery for bigeye. Many factors combine to determine the actual fishing depth of a net in a given set of environmental conditions. For this, and other reasons (Branch et al. 2006), restrictions on the set-up of fishing gear would seem ill-advised. On the other hand, gear research directed towards improving acoustic techniques for the characterization of object-associated communities (e.g., Shaefer and Fuller, 2005; Doray et al. 2006) would seem beneficial. The ability to accurately assess the compostion of object-associated tuna schools in areas where bigeye catch is likely would clearly improve imformation available to fishermen for making choices.

One benefit of the approach used in this analysis is that it identifies ‘unusual’ sets (fishing trips) of specific vessels through accumulation of misclassifications that could be subject to further analysis. Results of this approach could also be used to create categories of vessels for the purpose of estimating conventional vessel effects (or skipper effects; e.g., Ruttan and Tyedmers, 2007), for example, by defining categories of vessels (or skippers) based on the magnitude of per-vessel probabilities.

ACKNOWLEDGEMENTS

We thank Nicholas Vogel for data base assistance, and Jorge Parraga, Marlon Román-Verdesoto, and Alex Urdiales for helpful discussions on fishing practices that they observed while working as observers. We also thank Robin Allen and William Bayliff for comments and suggestions that improved this manuscript.

REFERENCES

Bayliff, W.H. 2001. Organization, functions, and achievements of the Inter-American Tropical Tuna Commission. Special Report 13. Inter-American Tropical Tuna Commission, La Jolla, California. http://www.iattc.org/PDFFiles2/IATTC_Special_Report13_ENG.pdf

Berk, R.A. 2006. An introduction to ensemble methods for data analysis. Sociological Methods and Research 34:263-295.

Breiman, L. 2001. Random Forests. Machine Learning 45:5-32.

Breiman, L, Friedman, J.H., Olshen, R.A., Stone, C.J. 1984. *Classification and Regression Trees*. Chapman and Hall.

Casey, K.S. and Evans, R. 2006. Global AVHRR 4 km SST for 1985-2005, Pathfinder v5.0, NODC/RSMAS. NOAA National Oceanographic Data Center, Silver Spring, Maryland. URL http://pathfinder.nodc.noaa.gov

Cayula, J.-F. and Cornillon, P. 1992. Edge detection algorithm for SST images. J. Atmos. Oceanic Technol. 9:67-80.

CLS. 2006. SSALTO/DUACS User Handbook: (M)SLA and (M)ADT Near-Real Time and Delayed Time Products. SALP CLS-DOS-NT-06.034. CLS, Agne, France. URL http://www.aviso.oceanobs.com/

Doray, M., Josse, E., Gervain, P., Reynal, L., Chantrel, J. 2006. Acoustic characterization of pelagic fish aggregations around moored fish aggregating devices in Martinique (Lesser Antilles). Fisheries Research 82:162-175.

Feldman, G.C. and McClain, C.R. 2005. Ocean Color Web, SeaWiFS Reprocessing 5.1.

NASA Goddard Space Flight Center. Eds Kuring, N. and Bailey, S.W. URL http://oceancolor.gsfc.nasa.gov

Green, R.E. 1969. Depth-time sequential analysis of the operation of two California tuna purse seines. Fishery Industrial Research 5 (5):191-201.

Harley, S.J. and Suter, J.M. 2007. The potential use of time-area closures to reduce catches of bigeye tuna (*Thunnus obesus*) in the purse-seine fishery of the eastern Pacific Ocean. Fishery Bulletin 105: 49-61.

Hastie, T., Tibshirani, R., Friedman, J. 2001. *The Elements of Statistical Learning; Data mining, Inference and Prediction*. Springer.

IATTC, 2000. Inter-Amer. Trp. Tuna Comm. Resolution on Bycatch C-00-08. http://www.iattc.org/PDFFiles/C-00-08%20Bycatch%20resolution%20Jun%2000.pdf

IATTC, 2006a. Tunas and billfishes in the eastern Pacific Ocean in 2005. Inter-Amer. Trop. Tuna Comm. Fishery Status Report No. 4.

IATTC, 2006b. Annual Report of the Inter-American Tropical Tuna Commission 2004. Inter-Amer. Trop. Tuna Comm. La Jolla, California, USA.

Kessler, W.S. 2006. The circulation of the eastern tropical Pacific: A review. Progress in Oceanography 69:181-217.

Kilpatrick, K.A., Podesta, G.P., Evans, R. 2001. Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. Journal of Geophysical Research-Oceans 106(C5):9179-9197.

Lennert-Cody, C.E. and Berk, R.A. 2007. Statistical learning procedures for monitoring regulatory compliance: an application to fisheries data. Journal of the Royal Statistical Society Series A 170 Part 3: 1-19.

Lennert-Cody, C.E. and Hall, M.A. 2000. The development of the purse-seine fishery on drifting fish aggregating devices in the eastern Pacific Ocean: 1992-1998. *In*: Péche thonière et dispositifs de concentration de poisons. Le Gall J.-Y., Cayré P., Taquet M. (eds.) Éd. Ifremer, Actes Colloq., 28, 78-107.

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

Marks, K.M. and Smith, W.H.F. 2006. An evaluation of publicly available global bathymetry grids. Marine Geophysical Researches 27(1):19-34.

Maunder, M.N. 2006. Workshop on Management Strategies. Compiled by Maunder, M. October 17-20, 2006. http://www.iattc.org/PDFFiles2/Management-strategies-WS-Oct-06-ReportENG.pdf

Maunder, M.N. and Harley, S.J. 2006. Evaluating tuna management in the eastern Pacific Ocean. Bulletin of Marine Science 78(3):593-606.

Maunder, M.N. and Hoyle, S.D. 2007. Status of bigeye tuna in the eastern Pacific Ocean in 2005 and outlook for 2006. Inter-Amer. Trop. Tuna Comm., Stock Assess. Rep. 7:119-248.

McClean, C.R, Feldman, G.C., Hooker, S.B. 2004. An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. Deep Sea Res. II 51:5-42.

Monterey, G. and Levitus, S. 1997. Seasonal variability of mixed layer depth for the World Ocean. NOAA Atlas, NESDIS 14, 100 pp., Washington, D.C.

R Development Core Team 2005. 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.

Rio, M.-H. and Hernandez, F. 2004. A mean dynamic topography computed over the world ocean from altimetry, in-situ measurements, and a geiod model. J. Geophys. Res. 109. C12032. doi:10.1029/2003JC002226.

Roberts, J.J. 2005. A 21-year global database of sea surface temperature fronts. Unpublished data. Duke University Marine Lab, Beaufort, North Carolina.

Ruttan, L.M. and Tyedmers, P.H. 2007. Skippers, spotters and seiners: Analysis of the "skipper effect" in US menhaden (*Brevoortia* spp.) purse-seine fisheries. Fisheries Research 83:73-80.

Schaefer, K.M., Fuller, D.W. 2005. Behavior of bigeye (*Thunnus obesus*) and skipjack (*Katsuwonus pelamis*) tunas within aggregations associated with floating objects in the equatorial eastern Pacific. Marine Biology 146:781-792.

Schaefer, K.M. and Fuller, D.W. 2002. Movements, behavior, and habitat selection of bigeye tuna (*Thunnus obesus*) in the eastern equatorial Pacific, ascertained through archival tags. Fishery Bulletin 100:765-788.

Smith, W. 2004. 1-minute global bathymetry data. Unpublished data.

Watters, G.M. 1999. Geographical distributions of effort and catches of tunas by purse-seine vessels in the eastern Pacific Ocean during 1965-1998. Data Report 10. Inter-American Tropical Tuna Commission, La Jolla, California.

TABLE 1. Predictors used to describe the presence/absence of bigeye tuna catches.

Predictor	**Additional details and range of values (minimum, median, maximum)**
Gear and operational predictors	
Vessel fish-carrying capacity (‘vessel capacity’)	Metric tons. (397; 1,089; 2,833)
Hanging depth of the purse-seine (‘net depth’)	Counted in strips and converted to meters (1 strip ≈ 11 m). Actual fishing depth of net not measured. (132; 219; 329)
Size of the mesh in the net (‘mesh size’)	Stretch measurement in inches. (3.5; 4.25; 12.0)
Dolphin safety panel	Presence/absence. The safety panel has a stretch mesh size of 1.25 in.
Maximum depth of floating object below water’s surface (‘object depth’)	Estimated in meters by the observer. Actual depth of object below surface not measured. (0.01; 18.1; 130)
Duration of encirclement and pursing (‘encirclement’)	Time (decimal hours) between the departure of the net skiff from the seiner and point at which the bottom of the net has been pursed. (0.27; 0.52; 2.43)
Percent of the object covered with fouling organisms (‘percent fouling’)	Used as a proxy for time the object spent in the water (i.e., soak time), although the relationship between fouling and actual soak time may be compromised by the fact that vessels may set upon/use objects belonging to other vessels, and it is not possible to track individual objects across vessel trips. (0; 40; 100)
Start time of the set (‘set time’)	Local time (decimal hours) of the release of the net skiff from the purse-seine vessel. This predictor was included because bigeye tuna have been shown to exhibit diel variability in their depth distribution when associated with floating objects (Schaefer and Fuller, 2002). (4.75; 6.68; 19.0)
Environmental predictors	
Sea surface temperature (‘SST’)	Measured *in-situ* by the observer in °C. (13.0; 26.1; 31.4)
Probability of a sea surface temperature front (‘SST front’)	Estimated at set locations using the NOAA National Oceanographic Data Center 4 km Advanced Very High Resolution Radiometer (AVHRR) SST data (Kilpatrick et al., 2001; Casey and Evans, 2006). The location of SST fronts were identified in the daytime AVHRR images from 1985-2005 by the presence of bimodal distributions in local SST (Cayula and Cornillon, 1992; Roberts, 2005). For each month, the proportion of images that a pixel contained a front and was cloud free is the estimate of the probability of a front. (0; 0.008; 0.07)
Mixed layer depth (‘MLD’)	Meters. Monthly average by 1° area. The MLD was defined as depth at which the temperature falls to 0.5°C below the surface temperature (data from the World Ocean Database 1998; estimates courtesy of Pacific Fisheries Environmental Laboratory, N.M.F.S., Pacific Grove, California, as outlined in Monterey and Levitus (1997)). (0.7; 35.1; 414.0)
Depth of the sea floor below the surface (‘bathymetry’)	Meters. Sampled from the “S2004” 1-minute global bathymetry data base (Smith, 2004) at the set location. See also Marks and Smith (2006). (-6,265; -3,935; -114)

Predictor	**Additional details and range of values (minimum, median, maximum)**
Strong currents	Presence/absence (estimated subjectively by the observer).
Sea surface height anomaly (‘SSH’)	Centimeters. Sampled at the set location and date from the Aviso 1/3° weekly Delayed Time Mean Sea Level Anomaly “Reference” data (DT-MSLA Ref) (CLS 2006), which estimated the height difference from the 1993-1999 mean surface (Rio and Hernandez, 2004). The altimeter products were produced by SSALTO/DUACS and distributed by Aviso, with support from CNES. (-21.25; 0.94; 34.12)
Slope of the sea surface height anomaly (‘SSH slope’)	Unitless (see SSH above for more details). (3.8×10^{-7}; 2.3×10^{-5};1.5×10^{-4})
Chlorophyll-a density (‘chlorophyll’)	mg/m^3. Sampled at the set location and date from monthly climatologies of chlorophyll density estimated by NASA Goddard Space Flight Center from 1998-2005 SeaWiFS ocean color measurements (Feldman and McClain 2006; McClain et al. 2004). (0.06; 0.17; 2.63)
Latitude (and latitude^2)	Decimal degrees.
Longitude (and longitude^2, longitude·latitude)	Decimal degrees (negative). ‘longitude·latitude’ indicates the product of longitude and latitude. Higher-order terms were included to help capture spatial structure.
Month	Categorical (1-12).
Miscellaneous predictors	
Year	Categorical (2001-2005).
Proxy for floating object density (‘object density’)	The number of unique object numbers within a 5° area around the set location and one month prior to the set date. Ideally, the number of unique objects in a given area and time window would be computed. However, this was not possible because the data do not allow objects to be tracked across vessel trips, nor do the data identify objects shared with /stolen by other vessels. (0; 29; 584)
Proxy for size of the non-tuna object-associated community (‘non-tuna bycatch’)	Natural logarithm of the observer’s count of the number of animals (other than tuna) that were brought onto the vessel’s deck dead. (0; 4.29; 11.06)

TABLE 2. Spearman's rank correlation coefficient between continuous predictors described in Table 1.

	Vessel capacity	Net depth	Mesh size	Object depth	Encircle -ment	Fouling	Set time	Latitude	Longitude	SST	SST fronts	MLD	Bathy- metry	SSH	SSH slope	Chloro- phyll	Non-tuna bycatch
Vessel capacity																	
Net depth	0.53																
Mesh size	0.42	0.47															
Object depth	-0.02	0.09	0.01														
Encirclement	0.31	0.20	0.23	-0.08													
Fouling	0.19	0.18	0.12	0.12	0.01												
Set time	-0.06	-0.11	-0.01	-0.12	-0.05	-0.11											
Latitude	0.10	-0.03	0.05	-0.06	0.10	-0.11	-0.02										
Longitude	-0.44	-0.33	-0.19	-0.18	-0.14	-0.25	0.20	-0.39									
SST	0.22	0.12	0.14	-0.11	0.14	-0.01	0.02	0.45	-0.21								
SST fronts	-0.06	-0.06	-0.03	-0.02	-0.04	-0.08	0.04	-0.07	0.03	-0.24							
MLD	0.28	0.27	0.13	0.23	0.06	0.29	-0.21	0.09	-0.68	0.04	-0.14						
Bathymetry	-0.31	-0.22	-0.14	0.01	-0.08	-0.13	0.02	0.29	0.23	-0.11	-0.03	-0.16					
SSH	0.11	0.07	0.06	0.06	0.01	0.08	-0.07	0.16	-0.25	0.18	<|.01|	0.25	-0.13				
SSH slope	0.08	<|.01|	0.01	-0.05	0.05	-0.03	0.04	0.25	-0.10	0.19	-0.02	<|.01|	-0.07	0.02			
Chlorophyll	-0.33	-0.30	-0.16	-0.19	-0.07	-0.32	0.16	0.05	0.63	-0.06	0.20	-0.65	0.18	-0.06	-0.05		
Non-tuna bycatch	0.04	-0.01	-0.01	0.04	-0.01	<|.01|	<|.01|	0.41	-0.15	0.15	-0.07	0.12	0.14	0.13	0.08	0.02	
Object density	-0.23	-0.15	-0.11	-0.02	-0.11	-0.18	0.11	-0.28	0.59	-0.15	0.04	-0.37	0.04	-0.09	-0.08	0.43	-0.06

TABLE 3. Confusion tables for: (a) the classification algorithm with approximately equal relative costs of false negatives and false positives (436/438 = 0.995); (b) the classification algorithm with the relative costs of false negatives set at approximately three times that of false positives (228/685 = 0.333); and (c) the classification algorithm *without* location and date predictors at approximately three to one relative cost of false negatives to false positives. The initial classification algorithm fit to the data (i.e., without setting the relative costs of the two types of errors) yielded misclassification errors of approximately 11% for sets that caught bigeye and 22% for sets that did not.

	Observed class	Predicted class		Misclassification error
		0 (no bigeye)	1 (bigeye)	
(a)	0 (no bigeye)	1945	438	0.184
	1 (bigeye)	436	2391	0.154
(b)	0 (no bigeye)	1698	685	0.287
	1 (bigeye)	228	2599	0.081
(c)	0 (no bigeye)	1539	844	0.354
	1 (bigeye)	281	2546	0.099

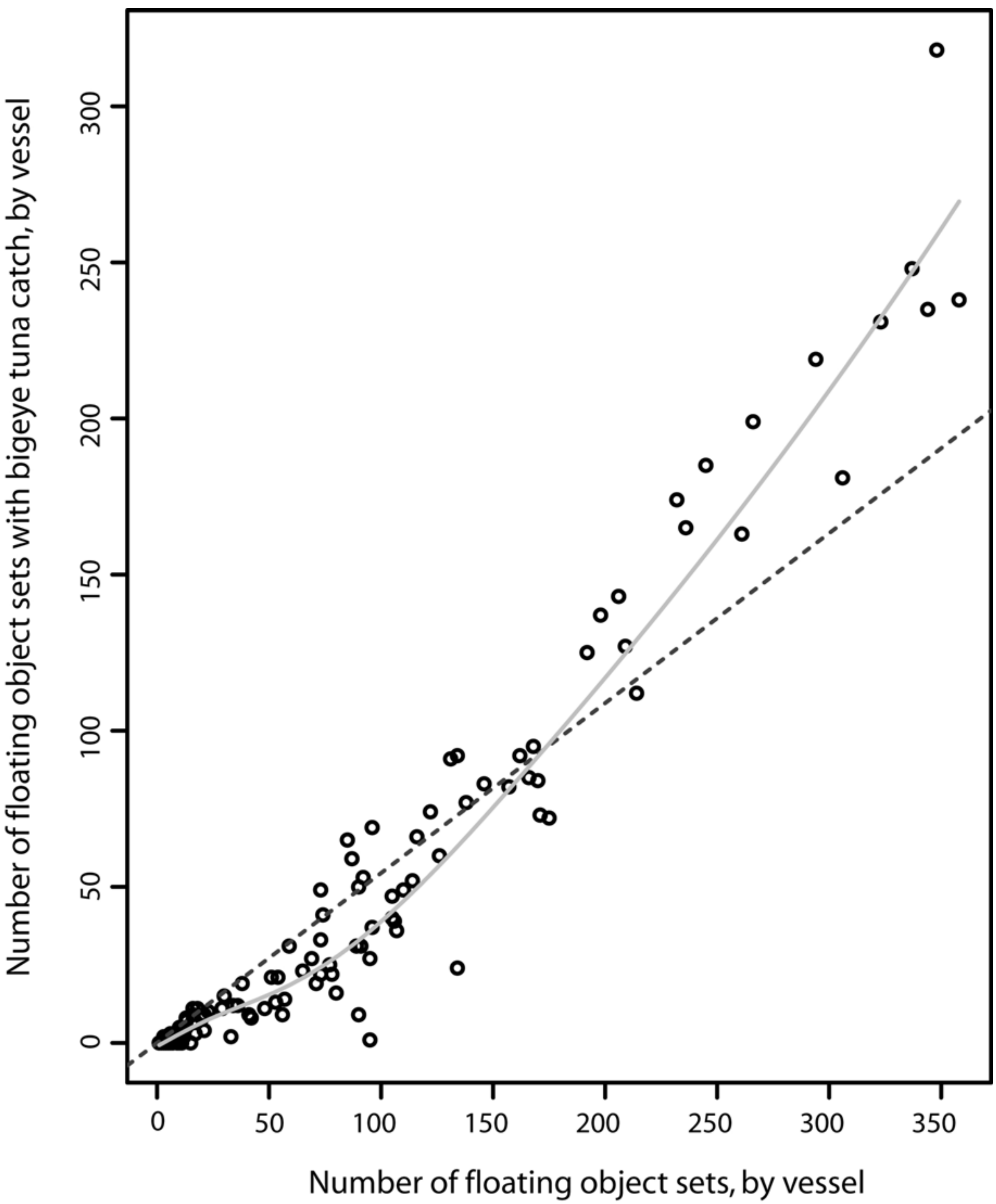

FIGURE 1. Number of purse-seine sets on floating objects *versus* number of purse-seine sets on floating objects that caught bigeye tuna, by vessel. The dashed line is the overall proportion of sets that caught bigeye tuna (0.54) multiplied by the number of sets on floating objects per vessel. The solid gray curve is a locally-weighted regression smooth of the data points.

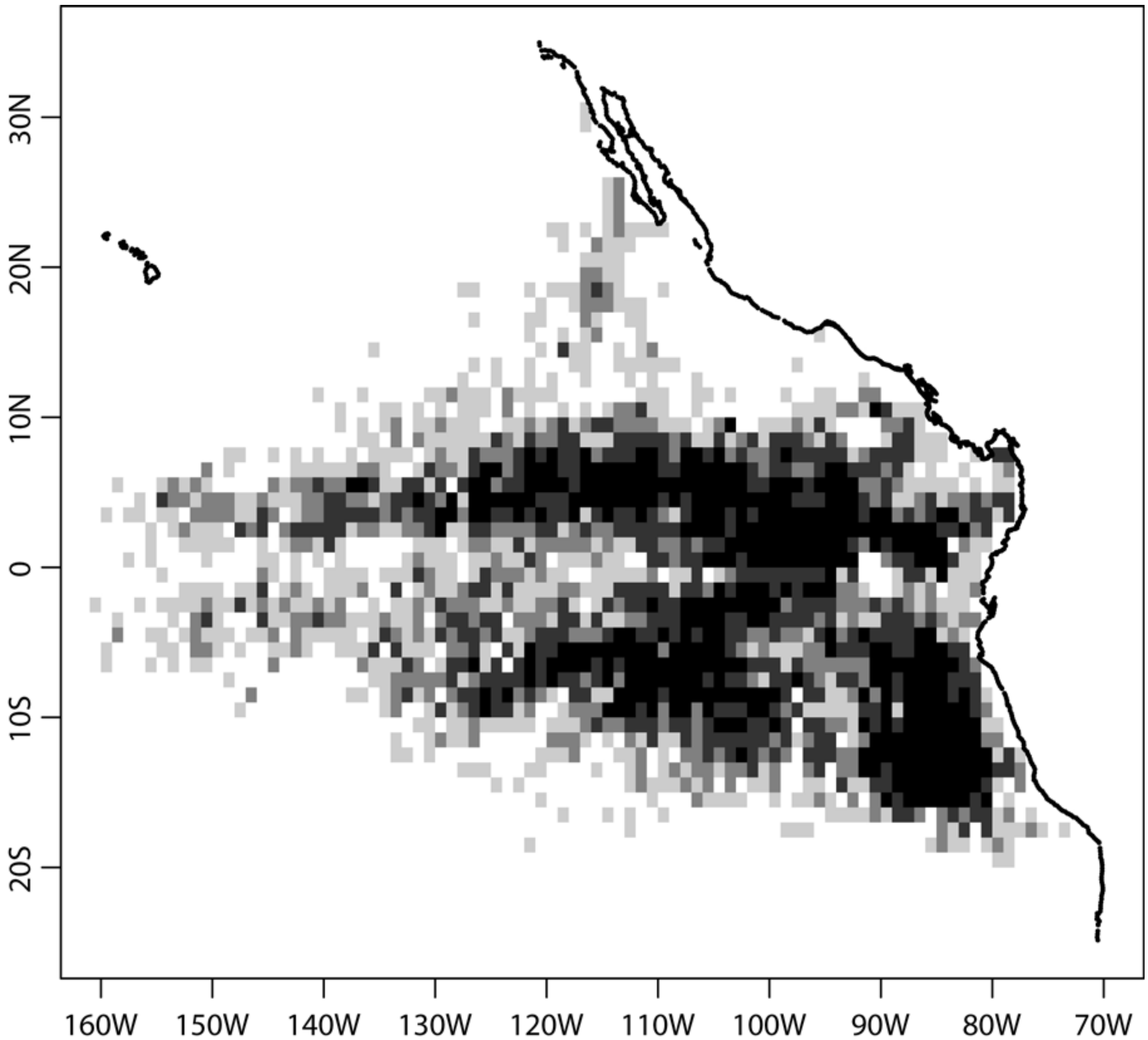

FIGURE 2. Number of sets on floating objects by 1° area, 2001-2005. The darker the area, the more sets (lightest gray: 1-2; medium gray: 3-4; dark gray: 5-9; black: ≥ 10).

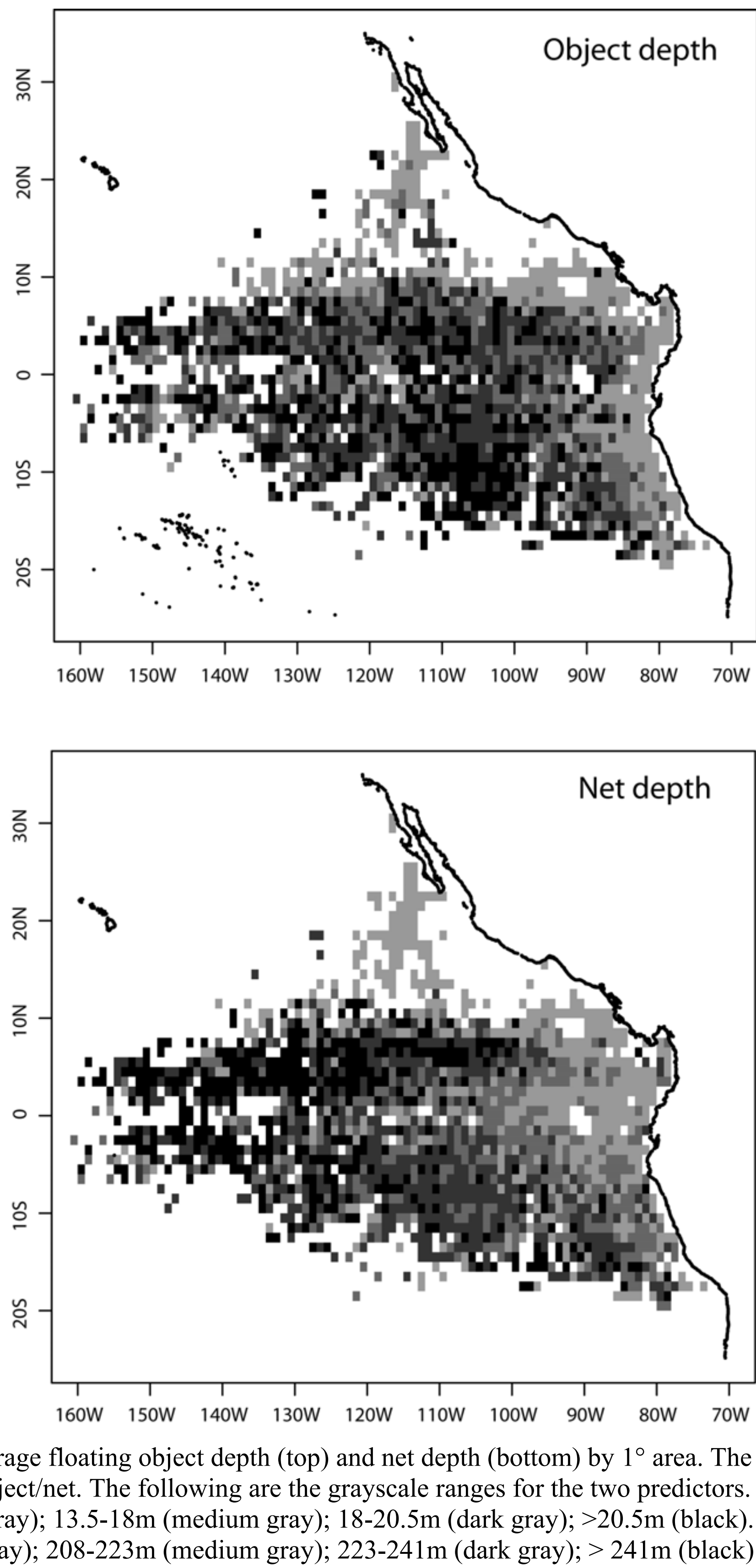

FIGURE 3. Average floating object depth (top) and net depth (bottom) by 1° area. The darker the square, the deeper the object/net. The following are the grayscale ranges for the two predictors. Object depth: ≤ 13.5m (lightest gray); 13.5-18m (medium gray); 18-20.5m (dark gray); >20.5m (black). Net depth: ≤ 208m (lightest gray); 208-223m (medium gray); 223-241m (dark gray); > 241m (black).

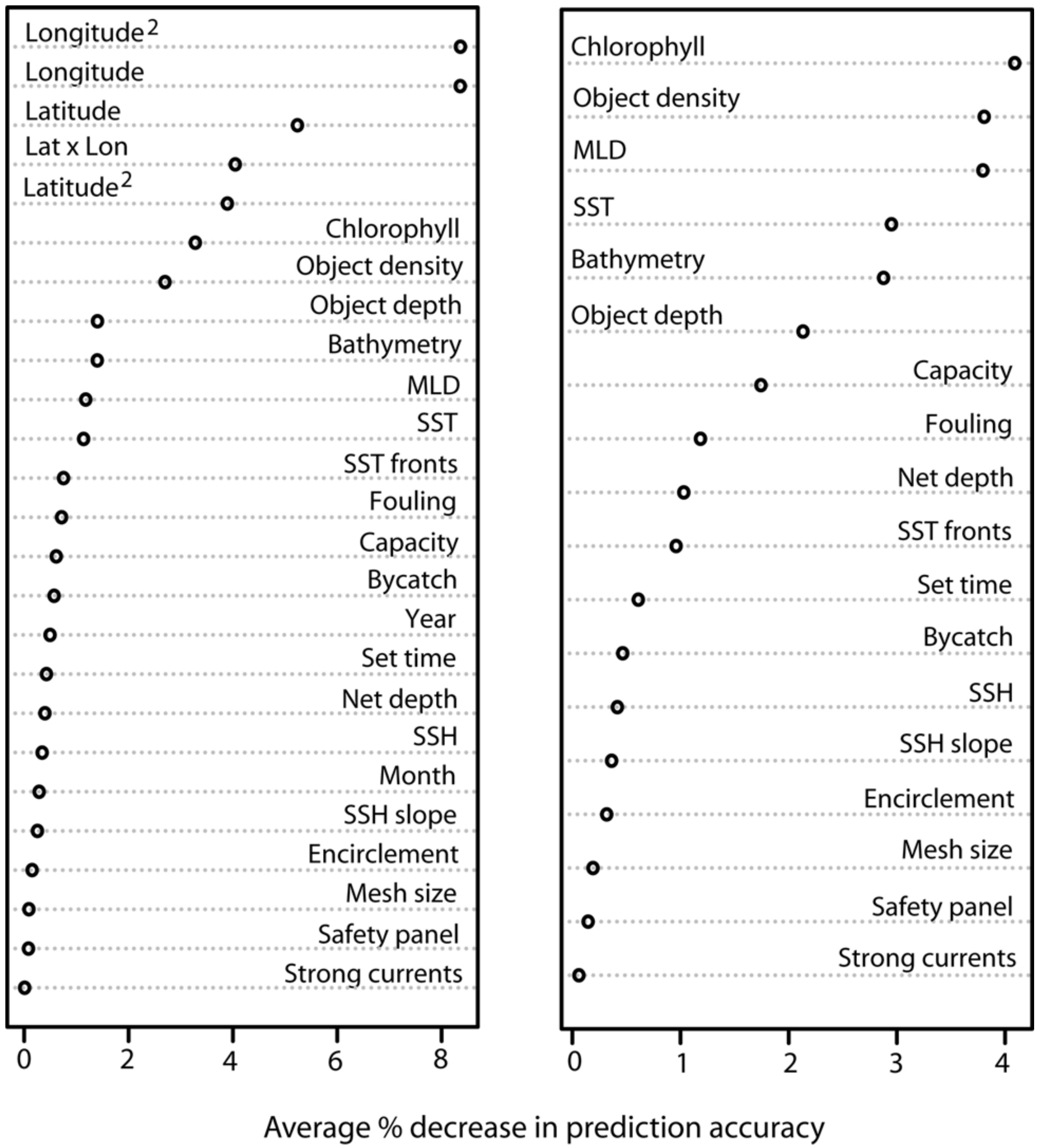

FIGURE 4. Predictor importance (average percent decrease in prediction accuracy when variable values were scrambled; based on the OOB data) for the classification algorithms with approximately three to one relative costs, with (left) and without (right) location and date predictors (Tables 3b-c). 'Lat x Lon' indicates the predictor constructed by taking the product of latitude and longitude.

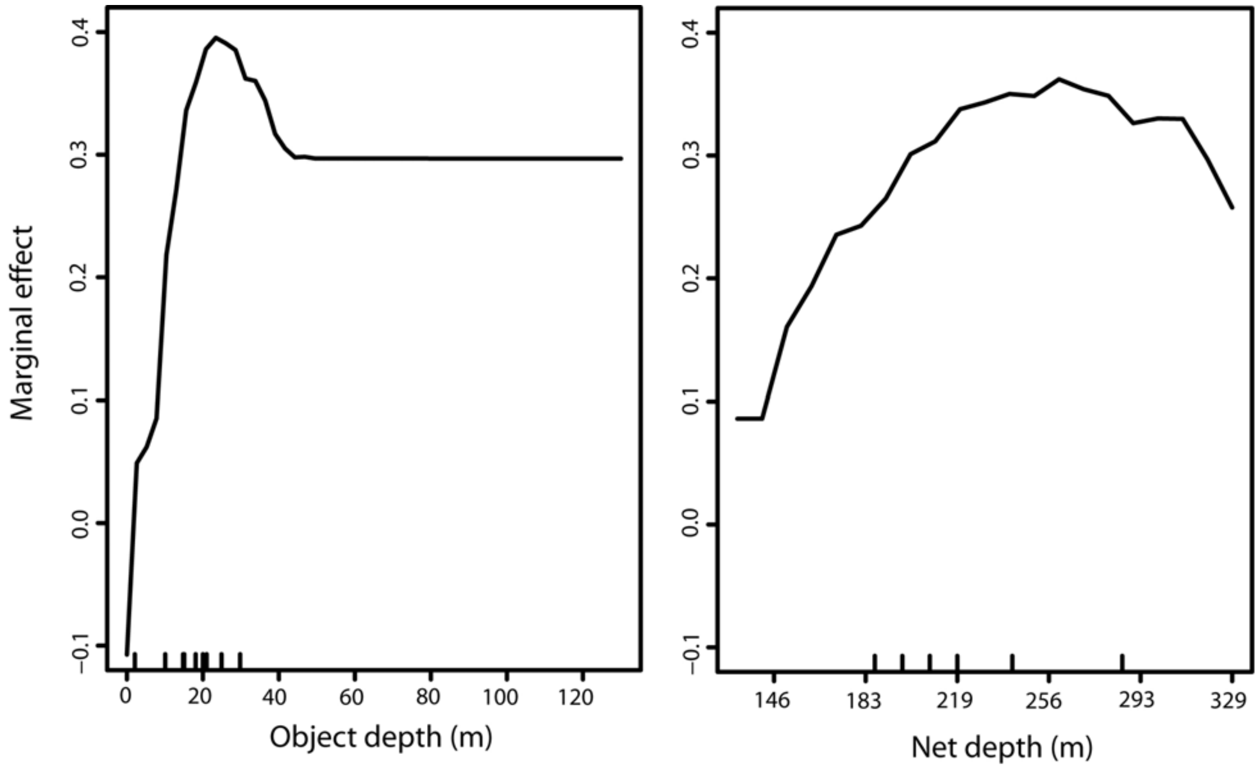

FIGURE 5. Marginal effect of object depth (left) and net depth (right) on the 'probability' that a set was classified as having caught bigeye tuna. The marginal effect is proportional to the average (over observations with a given object depth/net depth) of the logit of the proporiton of trees in the forest voting for the presence of bigeye tuna (i.e., the log of the fraction of trees voting for presence of bigeye). The 'rug' at the bottom of each graph shows the deciles of the values of object depth/net depth (if decile values are the same, hash marks will lay on top of one and other, resulting in fewer than nine hash marks).

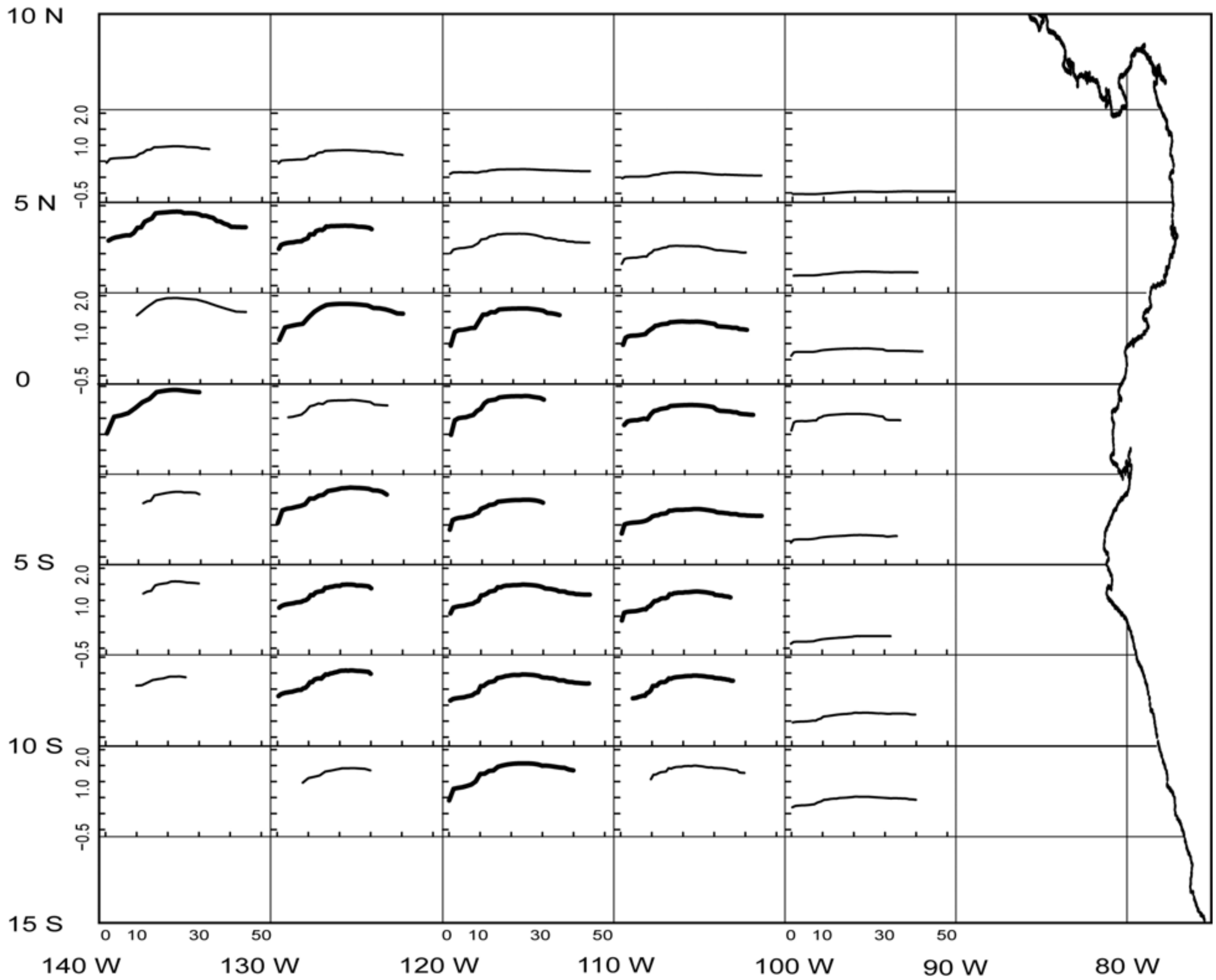

FIGURE 6. Marginal effect of object depth on the 'probability' that a set was classified as having caught bigeye tuna, by area, bewteen 90°-140°W and 12.5°S-7.5°N. Thick black lines indicate those areas with greater than averge increase in the marginal effect (increase in the marginal effect was computed by rectangular area as the maximum value of the marginal effect minus the minimum value). Marginal effect defined in Figure 5.

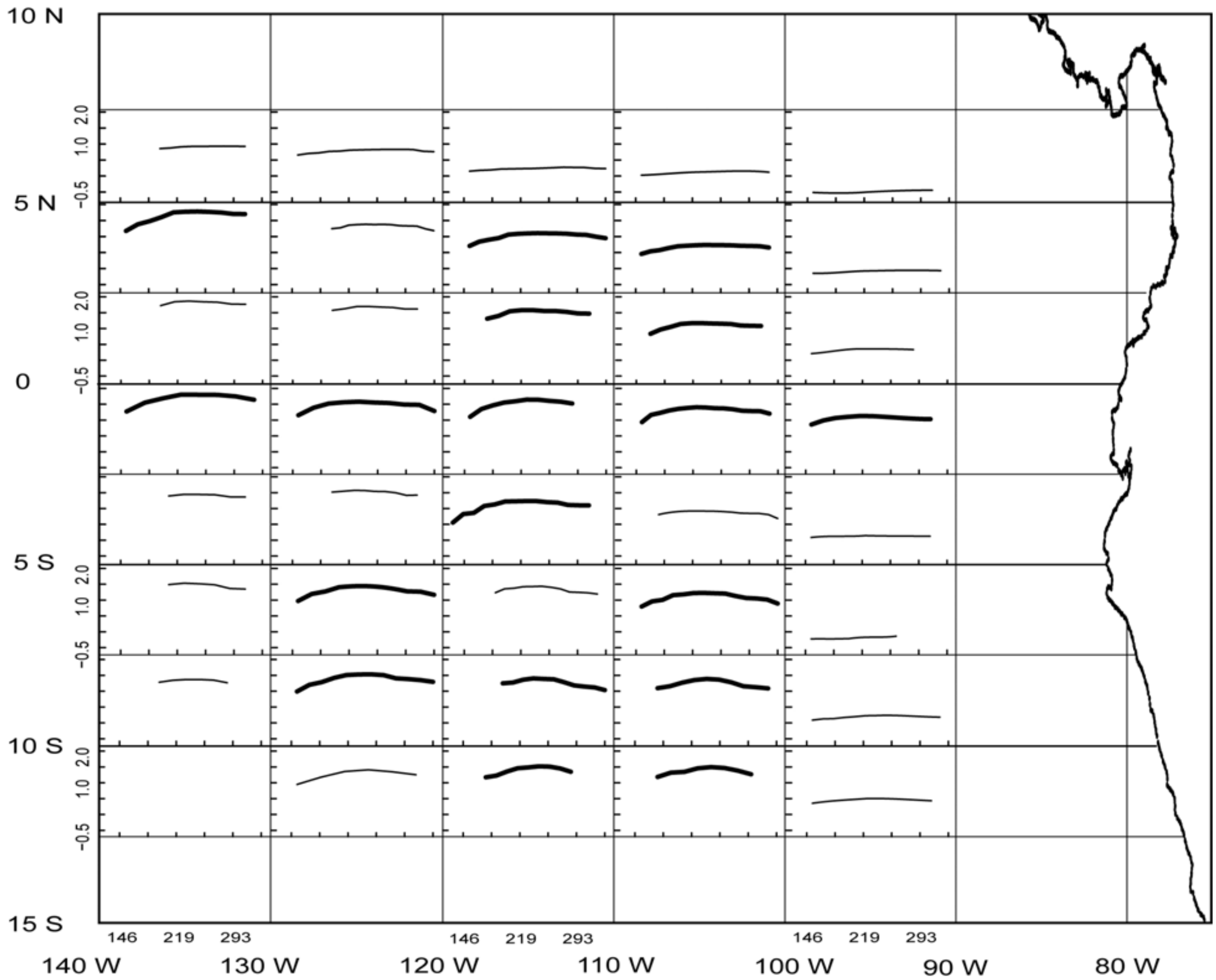

FIGURE 7. Marginal effect of net depth on the ‘probability’ that a set was classified as having caught bigeye tuna, by area, bewteen 90°-140°W and 12.5°S-7.5°N. Thick black lines indicate those areas with greater than averge increase in the marginal effect (increase in the marginal effect was computed by rectangular area as the maximum value of the marginal effect minus the minimum value). Marginal effect defined in Figure 5.

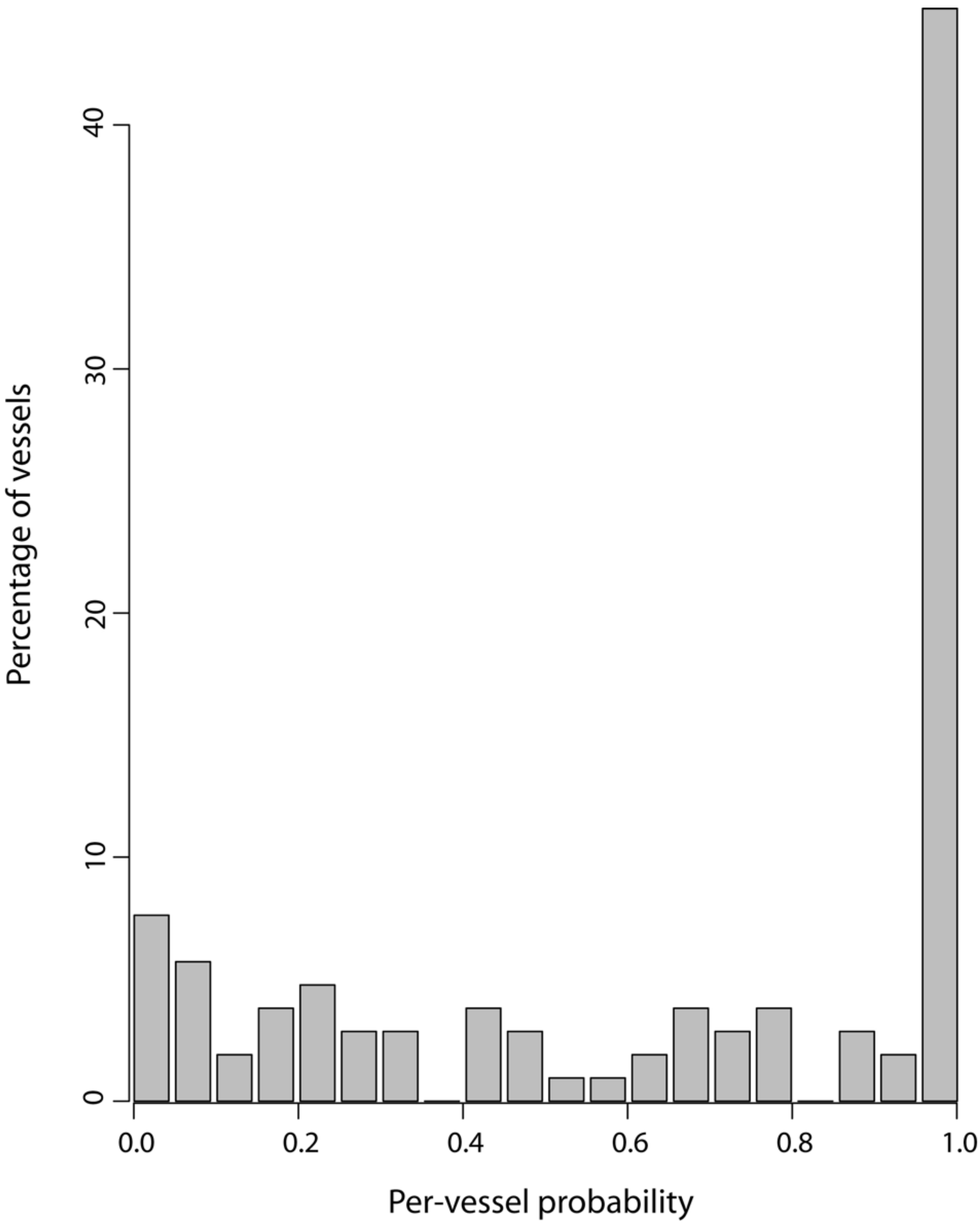

FIGURE 8. Frequency distribution of per-vessel probabilities (in terms of percentage of vessels). Values close to 1.0 indicate vessels with relatively few sets for which bigeye tuna was caught but none was predicted. Values close to 0.0 indicate vessels with a larger number of false negatives, relative to the number of sets in which these vessels caught bigeye tuna.