Machine learning to detect bycatch risk: Novel application to echosounder buoys data in tuna purse seine fisheries

Laura Mannocci, Yannick Baidai, Fabien Forget, Mariana Travassos Tolotti, Laurent Dagorn, Manuela Capello

Machine learning to detect bycatch risk: novel application to echosounder buoys data in tuna purse seine fisheries

Laura Mannocci[1*], Yannick Baidai[1,2], Fabien Forget[1], Mariana Travassos Tolotti[1],

Laurent Dagorn[1], Manuela Capello[1]

[1] MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Avenue Jean Monnet CS 30171 34203 Sète cedex, France

[2] Centre de Recherches Océanologiques (CRO), 29, Rue des pêcheurs, BPV 18, Abidjan, Côte d'Ivoire

*Corresponding author: laura.mannocci@gmail.com

Acknowledgments

Competing interest statement

The authors have no competing interest to declare.

1 Machine learning to detect bycatch risk: novel application to

2 echosounder buoys data in tuna purse seine fisheries

3 Abstract

4 The advent of big data and machine learning offers great promise for addressing conservation and

5 management questions in the oceans. Yet, few applications of machine learning exist to mitigate the

6 overexploitation of marine resources. Tropical tuna purse seine fisheries (TTPSF) are distributed

7 worldwide and account for two thirds of the global tuna catch. In these fisheries, the use of Drifting

8 Fish Aggregating Devices (DFADs)—man-made floating objects massively deployed by fishers to

9 increase their tuna catches—results in the incidental catch of non-target species, termed bycatch.

10 We explored the possibility of applying machine learning on echosounder buoys attached to DFADs

11 — representing an unprecedented source of big data – for identifying high bycatch risk at DFADs. We

12 trained random forests algorithms to differentiate between high and low bycatch occurrence based

13 on matched echosounder and onboard observer data for the same DFADs (representing sample sizes

14 of 838 and 2144 in the Atlantic and the Indian Ocean, respectively). Algorithms showed a better

15 performance in the Atlantic Ocean (accuracy of 0.66 *versus* 0.58 in the Indian Ocean) and were best

16 at detecting the "high bycatch" occurrence class. These results unravel the potential of machine

17 learning applied to fishers' buoys data for bycatch reduction and improved selectivity in one of the

18 largest fisheries worldwide.

19 **Keywords**: **Atlantic Ocean; drifting fish aggregating devices; echosounder buoys; Indian Ocean;**

20 **random forests; tropical tuna purse seine fisheries**

Introduction

The past decade has seen an unprecedented rise in data collections on aquatic ecosystems (Durden et al 2017). In the oceans, emerging sources of datasets include automatic vessel tracking systems (Kroodsma et al. 2018) and autonomous underwater vehicles (Sahoo et al. 2019). To accelerate the processing of these massive and complex datasets often referred to as "big data", machine learning algorithms have emerged as powerful tools, with promising applications in marine sciences (Beyan & Browman 2020). Together, big data and machine learning have a huge potential for addressing conservation and management issues in the oceans.

With over half of the world's oceans subject to industrial fishing (Kroodsma et al. 2018), the overexploitation of marine resources is a pervasive threat to marine biodiversity. Pursuant to the Sustainable Development Goal 14 of the United Nations, achieving sustainable use of ocean resources is a major societal challenge. Machine learning tools provide new opportunities for improving the management of fisheries by allowing the automated monitoring of catches for both target and non-target species (Bradley et al. 2019; Malde et al. 2020). There is a growing need for pragmatic and illustrative cases for how these tools can be used to promote sustainable fishing practices at sea.

The global catch of tropical tuna is close to 5 million tons per year and tropical tuna purse seine fisheries (TTPSF) account for two thirds of this global catch (ISSF 2019). In these fisheries, the use of Drifting Fish Aggregating Devices (DFADs)—man-made floating objects deployed by fishers to increase their tuna catches—results in the incidental catch of other fish species, termed bycatch (Dagorn et al. 2013). Though bycatch in TTPSF is relatively lower than in some other fisheries (with a reported discards rate of 3.9% in TTPSF *versus* 12.3% in pelagic longlines fisheries, for example) (Roda et al. 2019), the sheer magnitude of TTPSF means that the induced mortality on non-target species is not negligible. Moreover, the strong associations of some species with DFADs make them particularly

45   vulnerable to bycatch (Forget et al. 2015). Decreasing overall bycatch volumes in TTPSF is central to

46   the application of an ecosystem approach to fisheries management (Pikitch et al. 2004).

47   Between 81,000 and 121,000 DFADs are deployed every year by fishers across the globe, and this

48   number is likely to have increased since this estimate was made (Gershman et al. 2015). Most DFADs

49   are equipped with satellite-linked echosounder buoys that automatically provide remote information

50   on their geolocation and fish presence. These buoys represent an unprecedented source of big data

51   for scientists who can use this constellation across the oceans as a scientific platform for observing

52   pelagic biodiversity (Moreno et al. 2016). While echosounder buoys have been designed to provide

53   fishers with information on tuna presence, we postulate that data from these buoys can also inform

54   on the presence of bycatch. In this study, we explored the potential of processing echosounder buoy

55   data with machine learning to assess bycatch risk at DFADs prior to fishing sets. Our novel approach

56   opens the door to the development of new machine learning tools to aid fishers in reducing bycatch,

57   thereby improving selectivity in TTPSF.

58   Methods

59   Database description

60   We obtained echosounder data from Marine Instruments "M3I" buoys attached to DFADs deployed

61   by French purse seiners in the Atlantic and Indian oceans for the 2013-2018 period. Each buoy is

62   equipped with a GPS and an echosounder operating at 50 Khz. Simplified acoustic profiles are stored

63   every 2 h and transmitted by satellite every 12 h (by default) in the form of 50 integer acoustic scores

64   (ranging from 0 to 7) indicating the acoustic energy recorded within 3-m depth layers over a total

65   detection range of 150 m (Figure S1) (Baidai et al. 2020). We obtained bycatch data from scientific

66   observer programs onboard French purse seiners in the Atlantic and Indian oceans implemented

67   under the EU Data Collection Framework (DCF) and the French OCUP program (Observateur

68   Commun Unique et Permanent). Bycatch data were used to associate acoustic profiles with the

69   actual bycatch biomass, based on the assumption that the entire aggregation is encircled and

70   captured. For each DFAD fishing set, dedicated observers recorded species-specific bycatch in

71   numbers of individuals, or in weight when the number of individuals was too high to be reliably

72   counted. Numbers were converted to weight based on the mean recorded individual length and

73   published length-weight relationships. Total bycatch biomass was then calculated by summing

74   species-specific biomasses. Observers also recorded the date, time, GPS location and the buoy

75   unique identifier associated with each set.

Acoustic data pre-processing

77   We pre-processed acoustic data for standardization and reduction of dimensionality as illustrated in

78   Figure S1 (detailed in Baidai et al. 2020). We excluded the two layers corresponding to the transducer

79   blanking zone (from 0 to 6 m depth). Acoustic scores were first aggregated temporally by averaging

80   them over 4-h slots. Next, acoustic scores were aggregated vertically based on a cluster analysis

81   (Murtagh & Legendre 2014), which identified 6 homogeneous groups of depth layers in each ocean

82   (Baidai et al. 2020). For each homogeneous depth group, acoustic scores recorded for each of the

83   layers constituting the group were summed and rescaled to obtain a single score per depth group

84   and time interval (Baidai et al. 2020). This pre-processing led to daily acoustic matrices of 6 x 6

85   dimension providing synthetic acoustic profiles.

Machine learning

87   The originality of our approach lies in the ability of machine learning to detect bycatch occurrence at

88   DFADs based on patterns on echosounder buoy outputs when bycatch species were present, with no

89   required a priori knowledge on these patterns, nor on species-specific acoustic responses. We

90   considered as bycatch species all non-tuna teleosts (Table S2). Because these species have swim

91   bladders and usually form dense schools, they are highly detectable by echosounder buoys. In

92   contrast, sharks generally occur in low numbers around DFADs (between 2 and 6 individuals on

93   average in the Atlantic and Indian oceans, respectively), have no swim bladder and do not usually

4

94 form dense schools. Hence, owing to their relatively insignificant biomass, they were less likely to be

95 detected by the echosounders and were excluded from our analyses.

96 We obtained learning datasets by matching bycatch biomasses reported by observers with daily

97 acoustic matrices in each ocean, based on buoy identifiers, and dates and times of the sets. We

98 discarded aberrant buoy identifiers for which positions reported by observers were inconsistent with

99 GPS positions. We used daily acoustic matrices corresponding to the day preceding the set in order

100 to avoid potential disturbances to fish aggregations induced by the fishing operation. Because very

101 few sets contained no bycatch (only 5.4% in the Atlantic Ocean and 6.2% in the Indian Ocean), we

102 used a cutoff for categorizing sets into "high" and "low" bycatch, and performed a binary

103 classification. We defined the best cutoff among values ranging from 0.1 to 1 t in each ocean as the

104 one leading to the highest classification accuracy (see below).

105 In each ocean, we applied a random forest (RF) classification algorithm (Breiman 2001) to

106 discriminate between high and low bycatch occurrence using the R package "randomForest" (Liaw &

107 Wiener 2002). A random forest consists of an ensemble of independent decision trees leading to a

108 more accurate prediction than that of any individual tree (Breiman 2001). Candidate predictor

109 variables were the elements of the 6 x 6 daily acoustic matrices. To deal with the imbalanced number

110 of observations in high and low bycatch classes, we resampled the dominant size class to make its

111 frequency closer to the rarest class (Kuhn & Johnson 2013). We grew three thousand trees in each

112 ocean. We assessed the number of predictor variables randomly sampled at each split (denoted as

113 "*mtry*") through a grid-search strategy implemented with the R package "caret" (Kuhn et al. 2019).

114 We selected the best "*mtry*" as the value generating the lowest classification error rate. The

115 importance of predictors was assessed using the mean decrease accuracy (i.e., the increase of

116 prediction error after permuting each predictor, leaving all other predictors unchanged) (Breiman

117 2001).

118   We performed model training and evaluation through a hold-out validation (corresponding to setting

119   a portion of the data aside to evaluate model performance) repeated 10 times. In each replica, we

120   divided the original dataset into training and validation datasets, representing 75% and 25% of the

121   data, respectively. We then derived the model accuracy (proportion of correct predictions) and the

122   kappa coefficient (a reliability index) (Cohen 1968) on the validation dataset. Finally, sensitivity (true

123   positive rate), specificity (true negative rate), and precision (positive predictive value) were derived

124   from the confusion matrix.

125   Results

126   In both oceans, we obtained the best RF accuracy with a cutoff of 0.2 t between low and high

127   bycatch classes (Figure S3-1 and S3-2). Learning datasets represented 838 and 2144 data points in

128   the Atlantic and Indian Ocean, respectively (Figure 1, Table S4-1). In the Atlantic Ocean (AO), the high

129   bycatch class corresponded to acoustic scores that were higher in intermediate depth layers (from 21

130   to 45 m), but slightly lower in shallow layers during the day (Figure S4-1). In the Indian Ocean (IO),

131   differences were less clear, with a high bycatch class associated with slightly higher acoustic scores in

132   shallow layers (from 6 to 18m) during the day.

133   The RF algorithm was better at discriminating high and low bycatch classes in the AO than in the IO,

134   with respective accuracies of 0.66 and 0.58, and a kappa coefficient twice higher in the AO (Table 1).

135   In both oceans, the sensitivity was higher than the specificity, indicating a better ability at detecting

136   the high bycatch class (Table 1, Figure 2).

137   In the IO, the most important predictors corresponded to shallow layers (6-15 m) during the day

138   (Figure 3). Conversely, in the AO, deep layers (27-150 m) during the day, and shallow layers (9-21 m)

139   around dawn, appeared most important.

140   Discussion

141   Fishing on DFADs accounts for more than half of the tuna landings of TTPSF (Restrepo et al. 2017),

142   but also leads to large overall bycatch volumes (Roda et al. 2019). We capitalized on the massive data

6

143   available from fishers' buoys analyzed through machine learning to assess whether they could inform

144   potential bycatch risks at DFADs. RF algorithms showed promise to discriminate between high and

145   low levels of teleost bycatch, particularly in the AO, and better detected the high bycatch class.

146   Implemented in real time, these algorithms could help fishers to avoid fishing sets on DFADs

147   associated with high risks of non-target species catch, thereby improving selectivity in TTPSF. Though

148   additional research is needed to increase the algorithms' accuracy, our approach shows great

149   potential for progressing towards an ecosystem approach to fisheries (Pikitch et al. 2004). Indeed,

150   our approach focuses on non-tuna teleost species impacted by fisheries and could help design spatio-

151   temporal management measures that could lead to a reduction in overall bycatch volume, thereby

152   complementing mitigation measures already in place for vulnerable species (e.g., Poisson et al.

153   2014). Regardless of the management strategy, knowledge on bycatch risk prior to fishing is key to

154   support best fishing practices.

155   As bycatch species are generally known to occur shallower than tuna species (Forget et al. 2015;

156   Macusi et al. 2017), we expected acoustic scores to be higher in shallower layers in the high bycatch

157   class. In the AO, acoustic scores in shallower layers were instead higher in the low bycatch class. We

158   also found a pronounced geographical structure, with 65% of our low teleost bycatch class

159   originating from coastal Angola and Gabon (Appendix S5) (an area that is nevertheless characterized

160   by high shark bycatch (Lopez et al. 2020)). In this region, the higher acoustic scores detected near the

161   surface despite low teleost bycatch could be explained by an increased abundance of tuna in shallow

162   layers usually occupied by bycatch species. RF algorithms further implemented at the regional scale

163   had lower performances than the Atlantic-scale algorithm (Appendix S5). These results suggest that

164   the higher performance in the AO compared to the IO could be related to the stronger

165   geographical/environmental structure in the AO. Remarkably, these contrasted patterns between

166   oceans and the importance of deep layers as predictor variables stress the need to reconsider the

167   fixed 25 m depth layer previously assumed as adequate for discriminating bycatch and tuna species

168   at DFADs (Orue et al. 2019).

169   Discrepancies between bycatch detection on echosounder profiles and in the catches may negatively

170   affect the performance of algorithms by introducing noise. For example, small schooling fish like

171   *Decapterus spp* and the blue runner *Caranx crysos* are characterized by a strong acoustic back-

172   scatter, but are not always retained by the net. The buoy's blanking zone and its smaller acoustic

173   cone in shallow layers also prevent the detection of bycatch species that occupy these positions in

174   the water column during the night (Forget et al. 2015), leading to an overall underestimation of

175   bycatch biomass from echosounders. Observer-derived bycatch estimates may be prone to

176   estimation errors, especially in such large volume fisheries, but remain the most reliable data source

177   of bycatch. In the future, observer data could be combined with complementary data sources, such

178   as electronic monitoring, in an attempt to improve bycatch estimates (Ruiz et al. 2015). Finally,

179   echosounder buoys cannot be used to detect sharks despite their vulnerability to bycatch and DFAD

180   entanglement (Dagorn et al. 2013; Filmalter et al. 2013). A different machine learning approach

181   applied to video data obtained from autonomous cameras attached to DFADs could be more

182   appropriate for detecting shark species.

183   Echosounder buoys initially developed as fishing tools represent an unprecedented and massive

184   source of information for characterizing patterns in pelagic fish occurrence (Moreno et al. 2016), but

185   challenges remain for their utilization in direct fisheries management applications. Despite the large

186   amounts of echosounder buoy data, matches with actual catches are limited (as shown by our

187   modest learning datasets). Expanded coverage of observer programs and better reporting of buoy

188   identifiers would help increase the size of learning datasets in the future. The potential for deriving

189   accurate estimates of fish biomass from the current echosounder buoys is also limited, as illustrated

190   by the poor performance of RFs for quantifying tuna biomass (Baidai et al. 2020). Finally, the unique

191   sampling frequency of the echosounder buoys does not appear as the best technological approach to

192   differentiate species. Nevertheless, our machine learning approach is applicable to any echosounder

193   buoy type, including multi-frequency buoys that have revealed helpful for refining species

194   differentiation (Moreno et al. 2019).

195   In conclusion, machine learning applied to massive data collected through fishers' buoys offers

196   promise for assessing bycatch risk at DFADs and could help designing spatio-temporal measures for

197   the reduction of bycatch in TTPSF. Implemented in real-time, such algorithms have the potential to

198   help fishers avoid bycatch at DFADs by providing them with information prior to the fishing sets,

199   paving the way for more sustainable fishing practices. Given the magnitude and extensive

200   distribution of TTPSF worldwide, this novel approach could contribute to reduce the impact of fishing

201   on pelagic marine ecosystems.

Literature cited

203   Baidai Y, Dagorn L, Amandé M, Gaertner D, Capello M. 2020. Machine learning for characterizing
204   tropical tuna aggregations under drifting fish aggregating devices (FADs) from commercial
205   echosounder buoys data Yannick Baidai; Laurent Dagorn; Monin J Amandé; Daniel Gaertner;
206   Manuela Capello. Fisheries Research **In press**.
207   Beyan C, Browman HI. 2020. Setting the stage for the machine intelligence era in marine science.
208   ICES Journal of Marine Science **77**:1267–1273. Oxford Academic.
209   Bradley D, Merrifield M, Miller KM, Lomonico S, Wilson JR, Gleason MG. 2019. Opportunities to
210   improve fisheries management through innovative technology and advanced data systems.
211   Fish and Fisheries **20**:564–583.
212   Breiman L. 2001. Random Forests. Machine Learning **45**:5–32.
213   Cohen J. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or
214   partial credit. Psychological Bulletin **70**:213–220.
215   Dagorn L, Holland KN, Restrepo V, Moreno G. 2013. Is it good or bad to fish with FADs? What are the
216   real impacts of the use of drifting FADs on pelagic marine ecosystems? Fish and Fisheries
217   **14**:391–415.
218   Filmalter JD, Capello M, Deneubourg J-L, Cowley PD, Dagorn L. 2013. Looking behind the curtain:
219   quantifying massive shark mortality in fish aggregating devices. Frontiers in Ecology and the
220   Environment **11**:291–296.
221   Forget FG, Capello M, Filmalter JD, Govinden R, Soria M, Cowley PD, Dagorn L. 2015. Behaviour and
222   vulnerability of target and non-target species at drifting fish aggregating devices (FADs) in the
223   tropical tuna purse seine fishery determined by acoustic telemetry. Canadian Journal of
224   Fisheries and Aquatic Sciences **72**:1398–1405.
225   Gershman D, Nickson A, O'Toole M. 2015. Estimating The Use of FADs Around the World: An updated
226   analysis of the number of fish aggregating devices deployed in the ocean Gershman, D.,
227   Nickson, A., O'Toole, M.,. The Pew Charitable Trusts. Available from
228   https://www.pewtrusts.org/en/research-and-analysis/reports/2015/11/estimating-the-use-
229   of-fads-around-the-world.
230   ISSF. 2019. Status of the world fisheries for tuna. Oct. 2019. ISSF Technical Report 2019-12.
231   International Seafood Sustainability Foundation, Washington, D.C., USA.
232   Kroodsma DA et al. 2018. Tracking the global footprint of fisheries. Science **359**:904–908.
233   Kuhn M et al. 2019. caret: Classification and Regression Training. Available from https://CRAN.R-
234   project.org/package=caret (accessed January 20, 2020).
235   Kuhn M, Johnson Kj. 2013. Applied Predictive ModelingSpringer. New York. Available from
236   https://www.springer.com/gp/book/9781461468486 (accessed January 20, 2020).
237   Liaw A, Wiener M. 2002. Classification and Regression by randomForest **2**:6.

238 Lopez J, Alvarez-Berastegui D, Soto M, Murua H. 2020. Using fisheries data to model the oceanic
239   habitats of juvenile silky shark (Carcharhinus falciformis) in the tropical eastern Atlantic
240   Ocean. Biodiversity and Conservation **29**:2377–2397.

241 Macusi ED, Abreo NAS, Babaran RP. 2017. Local Ecological Knowledge (LEK) on Fish Behavior Around
242   Anchored FADs: the Case of Tuna Purse Seine and Ringnet Fishers from Southern Philippines.
243   Frontiers in Marine Science **4**. Frontiers. Available from
244   https://www.frontiersin.org/articles/10.3389/fmars.2017.00188/full (accessed July 22,
245   2020).

246 Malde K, Handegard NO, Eikvil L, Salberg A-B. 2020. Machine intelligence and the data-driven future
247   of marine science. ICES Journal of Marine Science **77**:1274–1285.

248 Moreno G, Boyra G, Sancristobal I, Itano D, Restrepo V. 2019. Towards acoustic discrimination of
249   tropical tuna associated with Fish Aggregating Devices. PLOS ONE **14**:e0216353. Public
250   Library of Science.

251 Moreno G, Dagorn L, Capello M, Lopez J, Filmalter J, Forget F, Sancristobal I, Holland K. 2016. Fish
252   aggregating devices (FADs) as scientific platforms. Fisheries Research **178**:122–129.

253 Murtagh F, Legendre P. 2014. Ward's Hierarchical Agglomerative Clustering Method: Which
254   Algorithms Implement Ward's Criterion? Journal of Classification **31**:274–295.

255 Orue B, Lopez J, Moreno G, Santiago J, Soto M, Murua H. 2019. Aggregation process of drifting fish
256   aggregating devices (DFADs) in the Western Indian Ocean: Who arrives first, tuna or non-
257   tuna species? PLOS ONE **14**:e0210435.

258 Pikitch EK et al. 2004. Ecosystem-Based Fishery Management. Science **305**:346–347. American
259   Association for the Advancement of Science.

260 Poisson F, Filmalter JD, Vernet A-L, Dagorn L. 2014. Mortality rate of silky sharks (Carcharhinus
261   falciformis) caught in the tropical tuna purse seine fishery in the Indian Ocean. Canadian
262   Journal of Fisheries and Aquatic Sciences **71**:795–798.

263 Restrepo V, Dagorn L, Itano D, Justel-Rubio A, Forget F, Moreno G. 2017. A Summary of Bycatch
264   Issues and ISSF Mitigation Initiatives To-Date in Purse Seine Fisheries, with emphasis on
265   FADs. ISSF Technical Report 2017-06. International Seafood Sustainability Foundation,
266   Washington, D.C., USA. Available from https://www.wcpfc.int/node/30729 (accessed March
267   11, 2020).

268 Roda MAP, Gilman E, Huntington T, Kennelly SJ, Suuronen P, Chaloupka M, Medley PA. 2019. A third
269   assessment of global marine fisheries discards. Food and Agriculture Organization of the
270   United Nations.

271 Ruiz J, Batty A, Chavance P, McElderry H, Restrepo V, Sharples P, Santos J, Urtizberea A. 2015.
272   Electronic monitoring trials on in the tropical tuna purse-seine fishery. ICES Journal of Marine
273   Science **72**:1201–1213.

274 Sahoo A, Dwivedy SK, Robi PS. 2019. Advancements in the field of autonomous underwater vehicle.
275   Ocean Engineering **181**:145–160.

276

Tables

Table 1: Performances of the random forest classifications derived from a hold-out validation in each ocean. Mean and standard deviation values (in brackets) of evaluation metrics are provided.

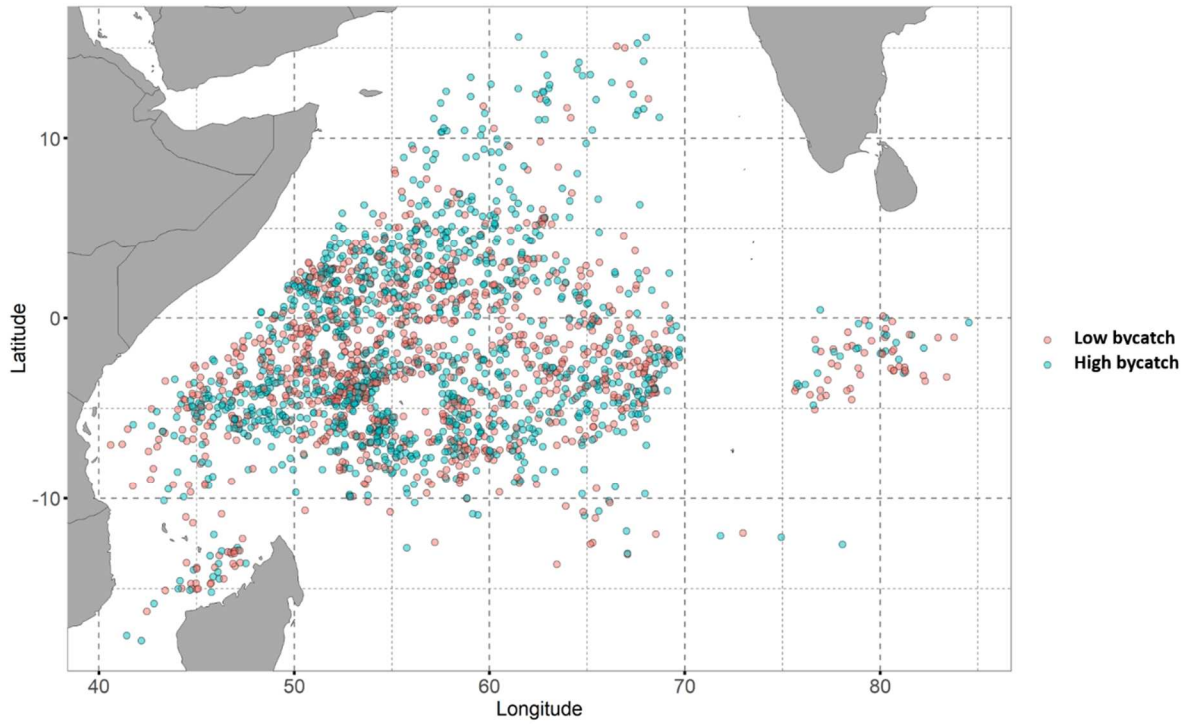| Evaluation Metrics	Atlantic Ocean	Indian Ocean
Accuracy	0.66 (0.04)	0.58 (0.02)
Kappa	0.32 (0.07)	0.16 (0.03)
Sensitivity	0.74 (0.05)	0.63 (0.03)
Specificity	0.59 (0.05)	0.53 (0.03)
Precision	0.64 (0.03)	0.57 (0.02)

Figures
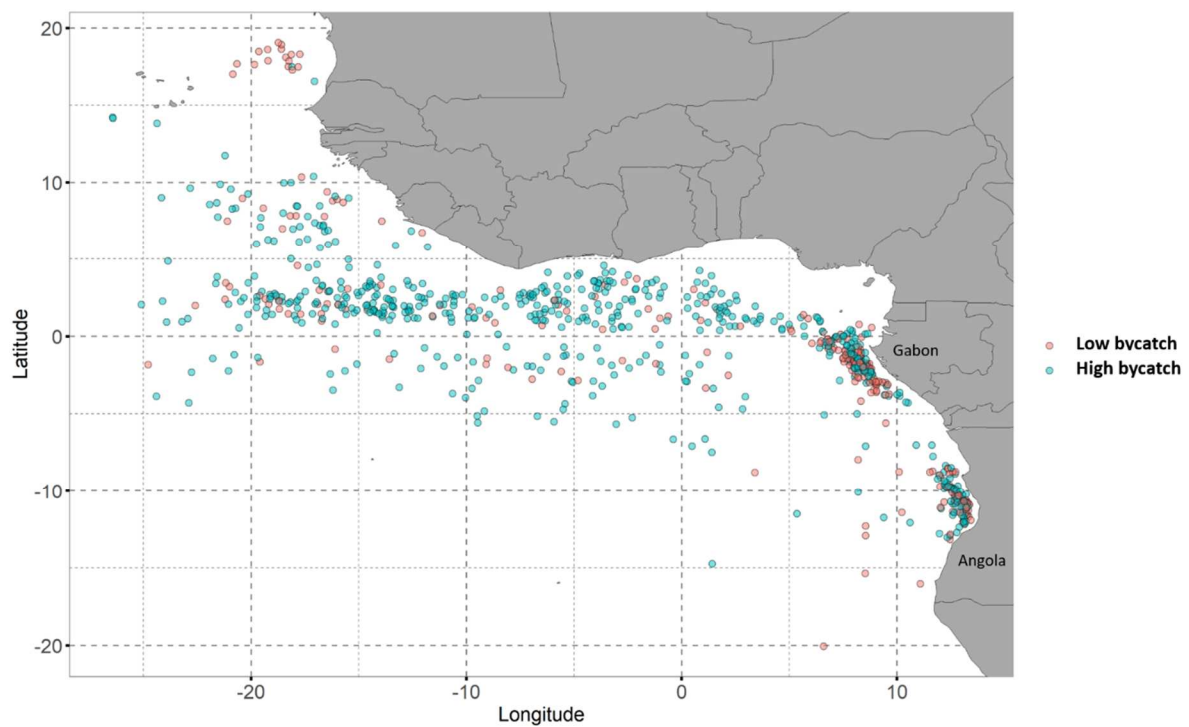
Figure 1: Geographical span of the learning datasets in the Atlantic (n=838 sets) and Indian oceans

(n=2144 sets). Each point corresponds to a fishing set on DFAD and the colors represent high and low
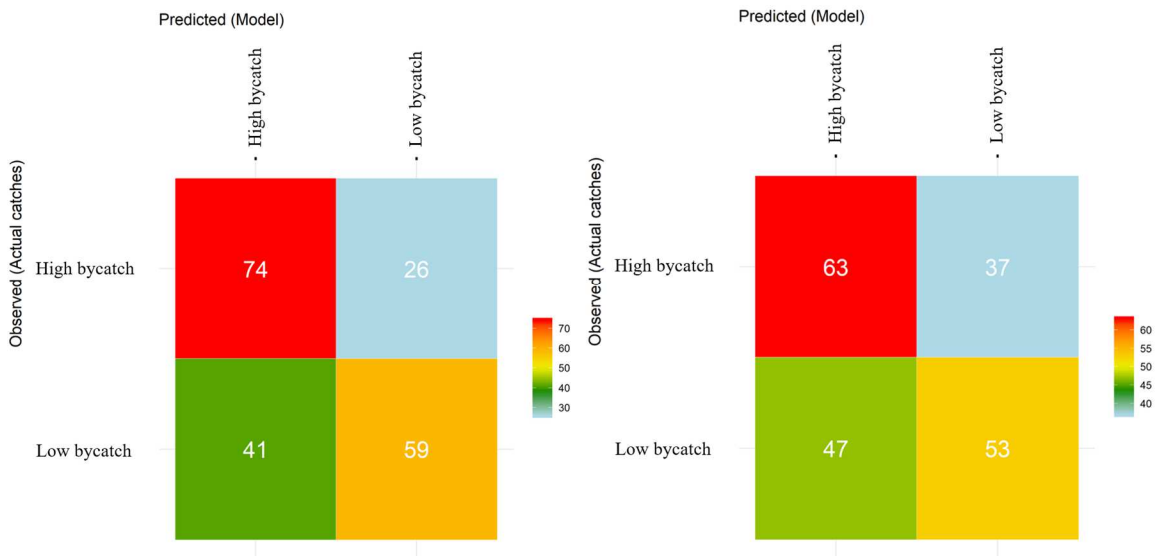
bycatch classes.

Figure 2: Confusion matrices standardized per row (i.e., with respect to observations) for random

forest classifications in the Atlantic Ocean (left) and the Indian Ocean (right). The diagonal elements

represent the percentage of data points for which the predicted class is equal to the observed class, while off-diagonal elements are those that are misclassifed by the random forest. The color scale represent the percentage of data points.
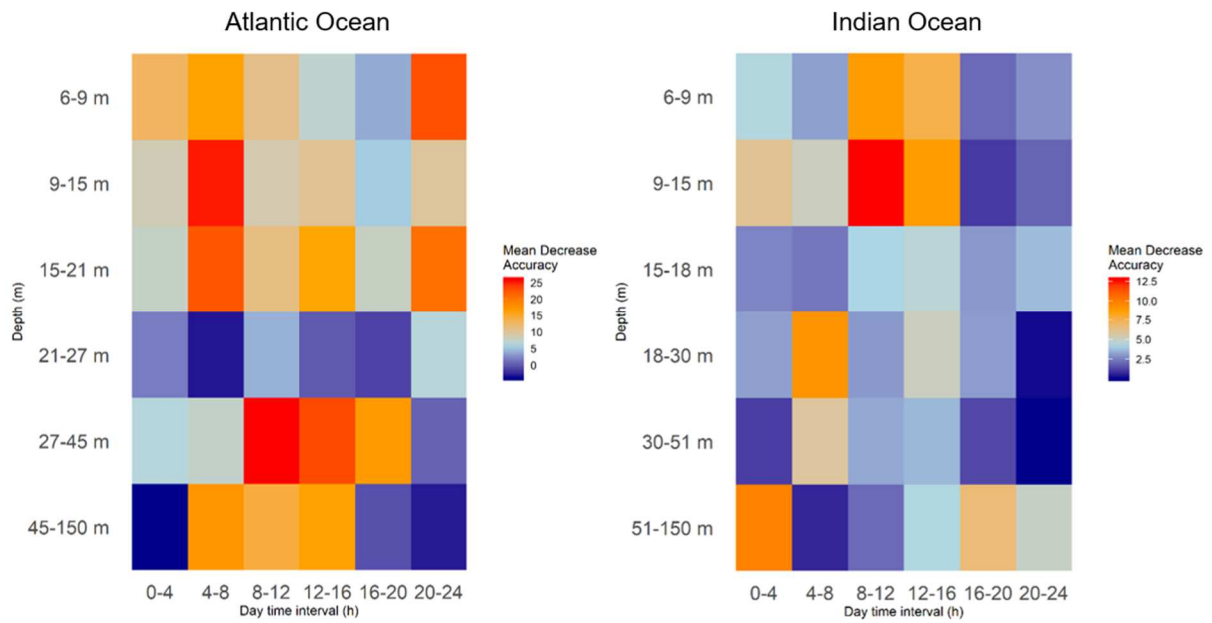

Figure 3: Importance of random forests predictors in the bycatch classification in both oceans assessed through the mean decrease accuracy (the mean decrease accuracy corresponds to the increase of prediction error after permuting each predictor, leaving all other predictors unchanged).