RECENT ADVANCES ON THE USE OF SUPERVISED LEARNING ALGORITHMS FOR DETECTING TUNA AGGREGATIONS UNDER FADS FROM ECHOSOUNDER BUOYS DATA

Y. Baidai[12], M.J. Amande[2], D. Gaertner[1], L. Dagorn[1] and M. Capello[1]

[1] MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France

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

Abstract

Assessing the accuracy of biomass estimates obtained through echosounder buoys and improving the current algorithms used for estimating the associated biomass is a key step towards the derivation of fisheries-independent abundance indices for tropical tuna. Recent results obtained through supervised learning algorithms on M3I buoys, one of the main buoy models deployed by the French tuna purse-seiners, demonstrate a good accuracy for assessing the presence and absence of tuna under FADs, regardless of the ocean. However, these algorithms (and buoy model) are less accurate in determining the size of tuna aggregations. In this paper we investigated possible ways of improving the classification of tuna aggregation sizes by accounting for the species composition constituting the aggregation. Also, we inspected how environmental variables (sea-surface temperature and chlorophyll-a) can affect the accuracy of the biomass estimates. Our results demonstrate that accounting for the species composition of tuna aggregation, sea-surface temperature and chlorophyll-a does not improve significantly the accuracy of biomass estimates with this buoy model.

Introduction

Echosounder buoys are used worldwide by purse seiners to remotely assess the amount of tuna aggregated around drifting FADs (DFADs). The number and wide spatial distribution of DFADs deployed (Maufroy et al. 2015, Moreno et al. 2016), coupled with their constantly evolving technology (Lopez et al. 2014, Moreno et al. 2016), grant to these fishing tools the status of a privileged platform for the observation of the pelagic communities that constitute the aggregation (Moreno et al. 2016; Brehmer et al. 2018). Recent studies proposed to use this new data for building fisheries-independent abundance indices for tropical tuna (Santiago et al. 2016, Capello et al. 2016, Baidai et al. 2017). Assessing the accuracy of biomass estimates obtained through the echosounder buoys and improving the current algorithms used for estimating the associated biomass is a key step towards the derivation of such indices.

In a recent paper «SUPERVISED LEARNING APPROACH FOR DETECTING PRESENCE-ABSENCE OF TUNA UNDER FAD FROM ECHOSOUNDER BUOYS DATA. » by Baidai et al (see the Annex), we proposed a new method for processing the echosounder data collected by the M3I Marine Instruments echosounder buoys, one of the main buoy models which equip the DFADs deployed by tuna purse-seiners. Thanks to the availability of a large acoustic database (all echosounder buoys deployed by the French fleet) and observers' data, we could obtain biomass estimates using an empirical approach, based on a supervised learning algorithm (random forest) trained on actual catch data reported by on-board observers on the same aggregations. The accuracy of the method was evaluated in the Atlantic and Indian Ocean and showed that the approach had very good efficiency for recognition of presence/absence of tuna but was less accurate for estimating the aggregation sizes. Among the possible causes of the loss of accuracy in the multiclass classification, we identified the influence of environmental conditions and the multispecific nature of tuna aggregations. Environmental conditions is well known to affect both accuracy of acoustic measurements (Bamber and Hill 1979, Straube and Arthur 1994) and tuna abundance (Boyce et al. 2008), and fishermen usually integrate various environmental factors in their empirical interpretation of echosounder buoys data.

Moreover, skipjack has a lower target strength than yellowfin and bigeye tuna (Bertrand et al. 1999, Josse and Bertrand 2000, Boyra et al. 2018)and thus the same acoustic signal could correspond to very different aggregation sizes, depending on the species composition of the aggregation. Similarly, large bigeye and yellowfin tuna could provide a different acoustic response relative to small individuals, due to the fact that they occupy lower depths, where the buoy may be less efficient in detecting the backscattered acoustic energy.

This paper constitutes an update of the recent results obtained by Baidai et al (2018). Based on the same data, we investigated how to improve the supervised learning multiclass classification algorithm to account for the species composition and size of the individuals constituting the aggregation and we analyzed the contribution of two environmental parameters (chlorophyll-a and sea surface temperature) in the algorithm accuracy.

DATA AND METHODS

Species composition of the aggregations

We considered the observers data reporting the amount of tuna caught at each FAD having a unique buoy ID present in the echosounder buoy database (567 sets in the Indian Ocean and 942 sets in the Atlantic Ocean, see details in Baidai et al 2018 in the Annex). First, we characterized the catch-per-set composition of our database, considering the amount of skipjack (SKJ), small yellowfin and bigeye (s-YFT-BET, individuals <= 10 Kg) and large yellowfin and bigeye (l-YFT-BET, individuals > 10 Kg) tuna caught per set. Globally, Figure 1 demonstrate that the catch data present in our database are representative of the multi-specific nature of tuna aggregations caught at FADs where, despite the majority of sets showed a larger percentage of SKJ, as expected, there is a significant amount of s-YFT-BET and l-YFT-BET, particularly in the Indian ocean.

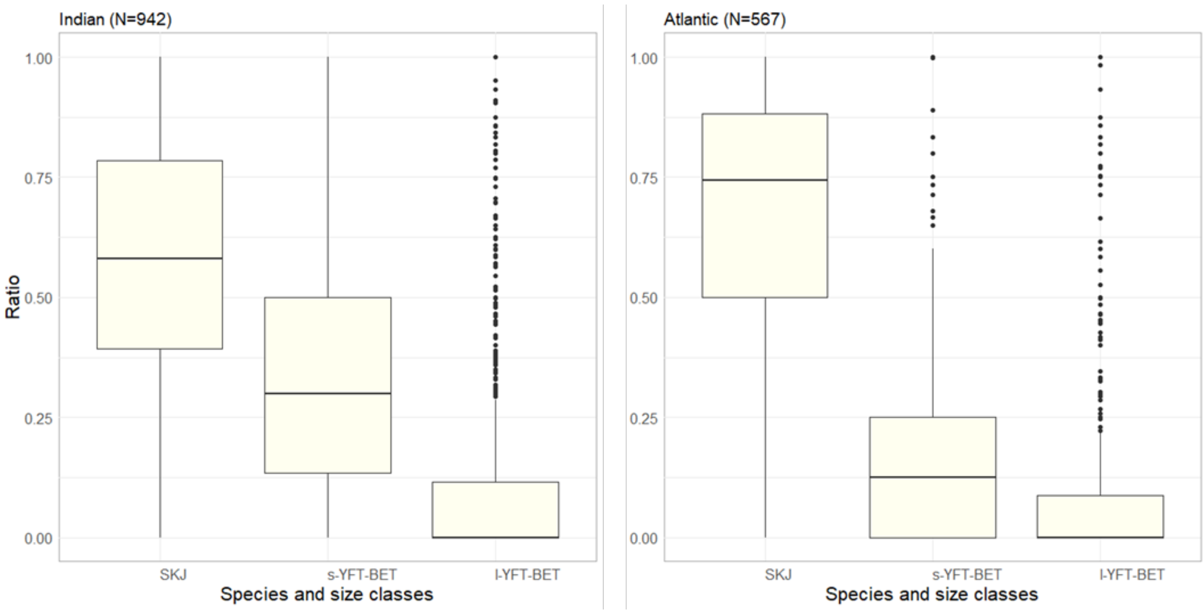

Figure 1: Catch-per-set composition for each ocean (SKJ : Skipjack tuna; s-YFT-BET: small yellowfin and bigeye and l-YFT-BET : large yellowfin and bigeye tuna). The ratio indicates the tonnage reported for each species/size class relative to the total catch per set (total tons of tuna).

Reweighted catch data accounting for the species composition of the aggregation in the classification algorithm

In order to test the hypothesis (H0) that accounting for the catch composition improves the quality of the classification algorithm, the multiclass random forest algorithm described in Baidai et al 2018 (Annex), was run on the catch and echosounder buoys dataset, with the catch data reweighted as follows:

$$\text{Catch\_eff}^{\text{s-YFT-BET}} = [\text{s-YFT-BET}] + \alpha\,[\text{SKJ}] + \beta\,[\text{l- YFT-BET}] \qquad \text{(Equation 1)}$$

were the brackets denote, for each species/size-class, the reported catch per set (in tons) and α, β are two multiplicative constants. The above equation, for α and β in the interval [0,1), accounts for the fact that the buoy underestimates the actual amount of [SKJ] and [l- YFT-BET] relative to [s-YFT-BET], due to the lower target strenght of SKJ and the lower detectability of l-YFT-BET. In the following, H0 was tested by comparing the accuracy of the classification algorithm for different values of α, β in [0,1) relative to the case $\alpha=1, \beta=1$. If H0 holds, an improved accuracy of the classification algorithm would be expected for $\alpha<\beta<1$.

Environmental parameters

The random forest multi-class classification algorithm was run integrating the sea surface temperature (SST) and chlorophyll-a (Chla), as classification features. SST and Chla data were obtained from the daily data collected by MODIS (Moderate Resolution Imaging Spectroradiometer) global 9 km resolution product, from 2014 to 2017 (ORNL DAAC, 2018). The missing values for a defined time and position have been replaced by their monthly average at the same position. Figures 3 and 4 show an example of the values of

SST and Chla recorded during one month in the proximity of the buoy positions constituting our database in the Indian and Atlantic Ocean, respectively. The boxplot of Figure 5 shows the overall values of SST and Chla recorded in the proximity of all the buoys constituting our database for the two oceans.

RESULTS

Environmental parameters

Overall performances of the multi-class classification algorithm integrating the SST and Chla as classification features are shown in Table 1 and 2 for the two oceans. Although the identification of tuna absence and large tuna aggregations under FADs (more than 25 tons) remain relatively acceptable (precision respectively 0.80 and 0.44 in the Atlantic Ocean, and 0.83 and 0.59 in the Indian ocean, the classification algorithms are still less effective in identifying small (less than 10 tons) and intermediate tuna aggregation (between 10 and 25 tons). However, the contribution of SST and Chla in the classification process appears higher than that of several acoustic predictors (Figure 6).

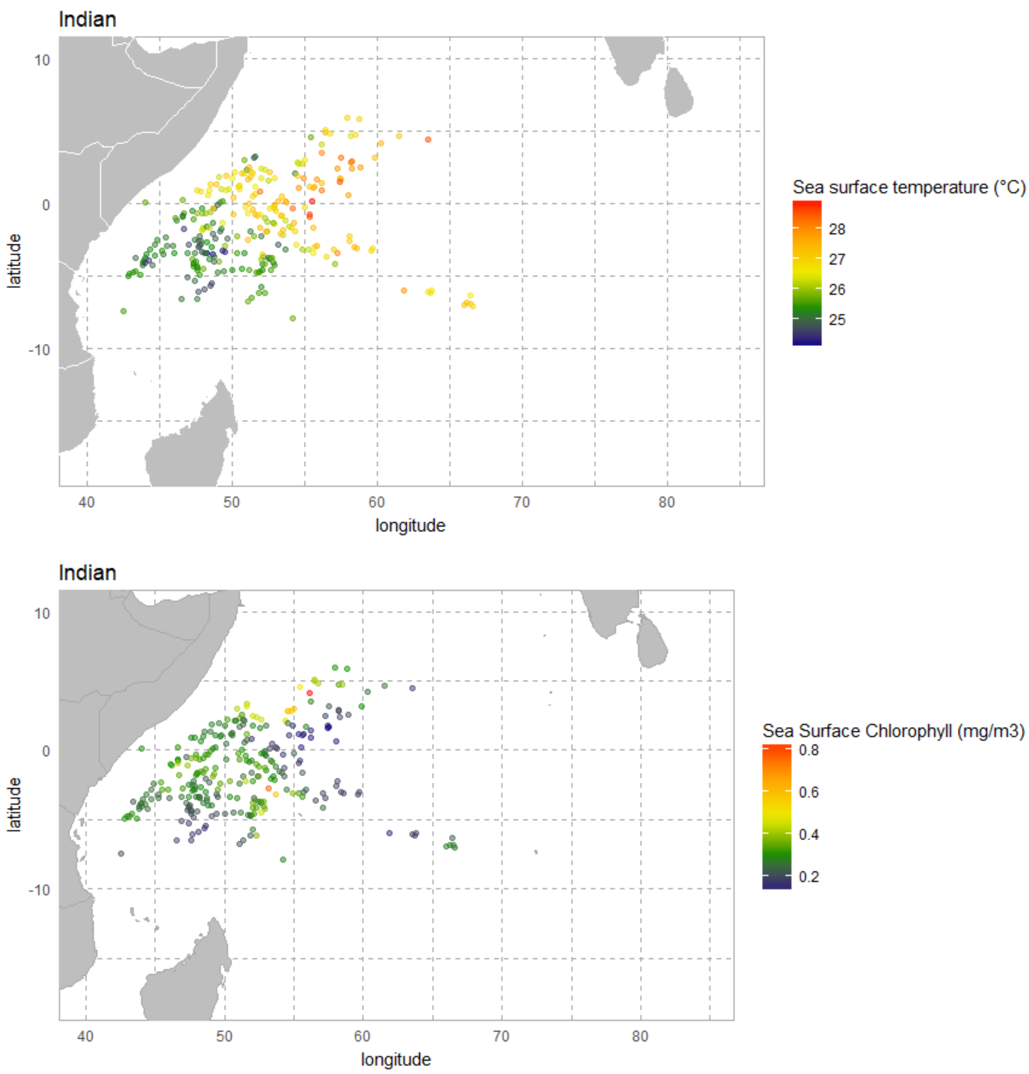

Figure 2: Examples of sea surface temperature and sea surface chlorophyll located at the position of 277 buoys present within our database (acoustic data + catch data) during August 2016 in the Indian Ocean.

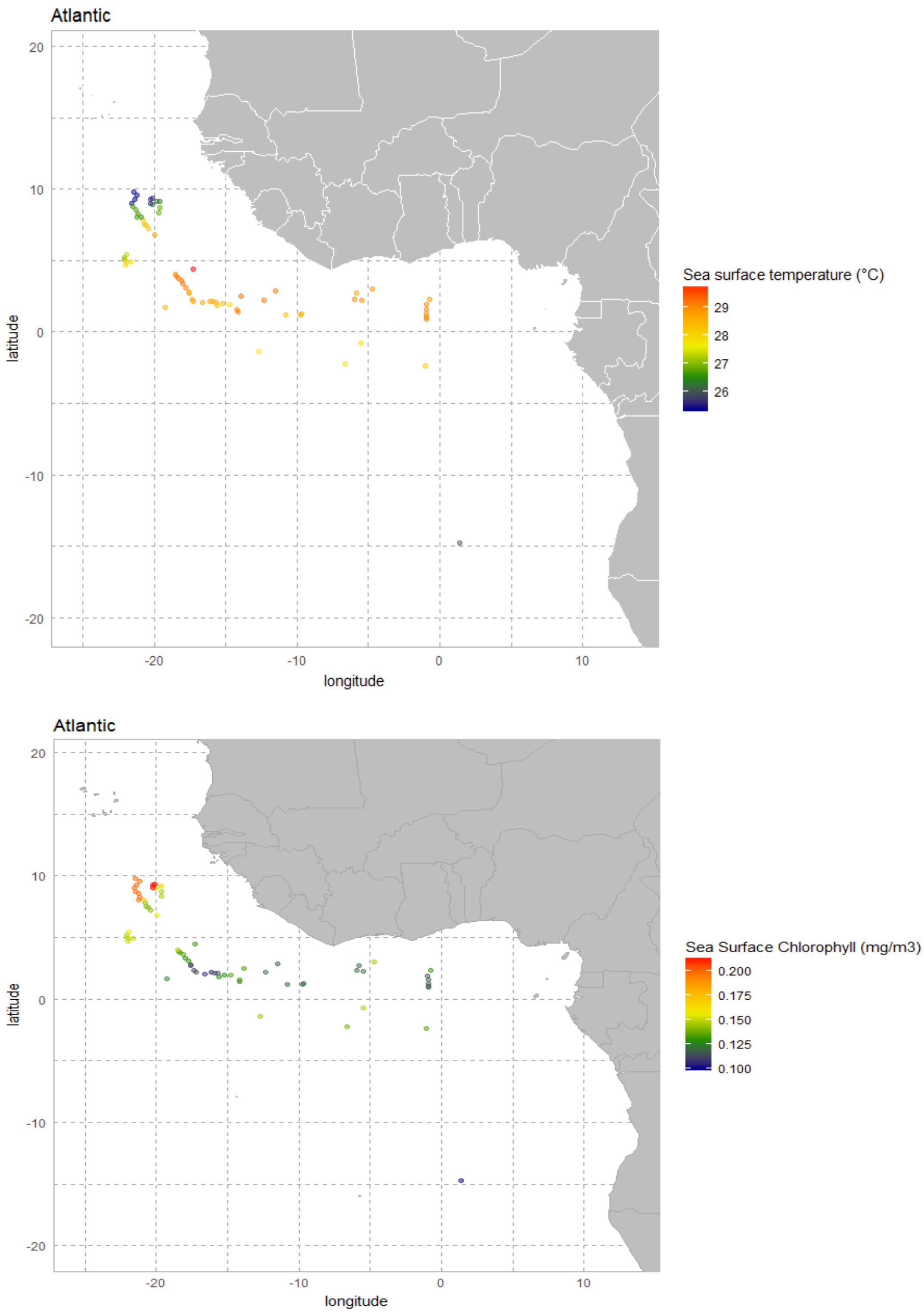

Figure 3: Examples of sea surface temperature (top) and sea surface chlorophyll a (bottom) located at the position of 70 buoys present within our database (acoustic data + catch data) during Mars 2016 in the Atlantic Ocean.

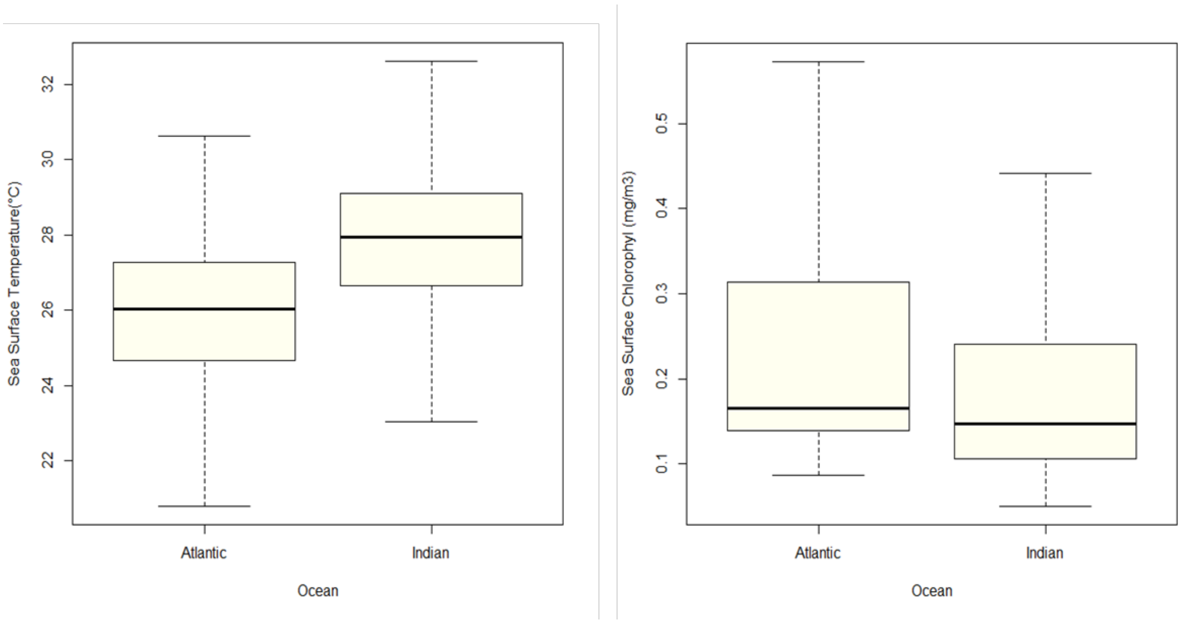

Figure 4: Boxplot of sea surface temperature and chlorophyll-a recorded at buoys positions in the Atlantic and Indian ocean. The outliers are omitted for obtaining a more visible plot.

Table 1: Summary of multiclass classification performance for the Indian Ocean with the SST and Chla as predictor variables. Means and standard deviations (in bracket) of evaluation metrics by classes.

Evaluation Metric	**Indian**				
	No tuna	<10 tons	[10, 25 tons]	> 25 tons	**Average**
Sensitivity	0.95 (0.01)	0.13 (0.05)	0.21 (0.02)	0.39 (0.05)	0.42
Specificity	0.58 (0.03)	0.95 (0.01)	0.95 (0.006)	0.96 (0.01)	0.86
Precision	0.85 (0.01)	0.13 (0.03)	0.36 (0.03)	0.55 (0.06)	0.47
F_1 score	0.90 (0.01)	0.13 (0.04)	0.27 (0.02)	0.45 (0.05)	0.44
Accuracy			0.76 (0.03)		
Kappa			0.40 (0.01)		

Table 2: Summary of multiclass classification performance for the Atlantic Ocean with the SST and Chla as predictor variables. Means and standard deviations (in bracket) of evaluation metrics by classes.

Evaluation Metric	**Atlantic**				
	No tuna	<10 tons	[10, 25 tons]	> 25 tons	**Average**
Sensitivity	0.85 (0.03)	0.45 (0.06)	0.35 (0.08)	0.38 (0.07)	0.49
Specificity	0.83 (0.05)	0.88 (0.02)	0.85 (0.02)	0.920(0.02)	0.86
Precision	0.81 (0.05)	0.41 (0.05)	0.32 (0.07)	0.51 (0.11)	0.49
F_1 score	0.83 (0.03)	0.43 (0.04)	0.33 (0.07)	0.43 (0.08)	0.49
Accuracy			0.60 (0.04)		
Kappa			0.42 (0.03)		

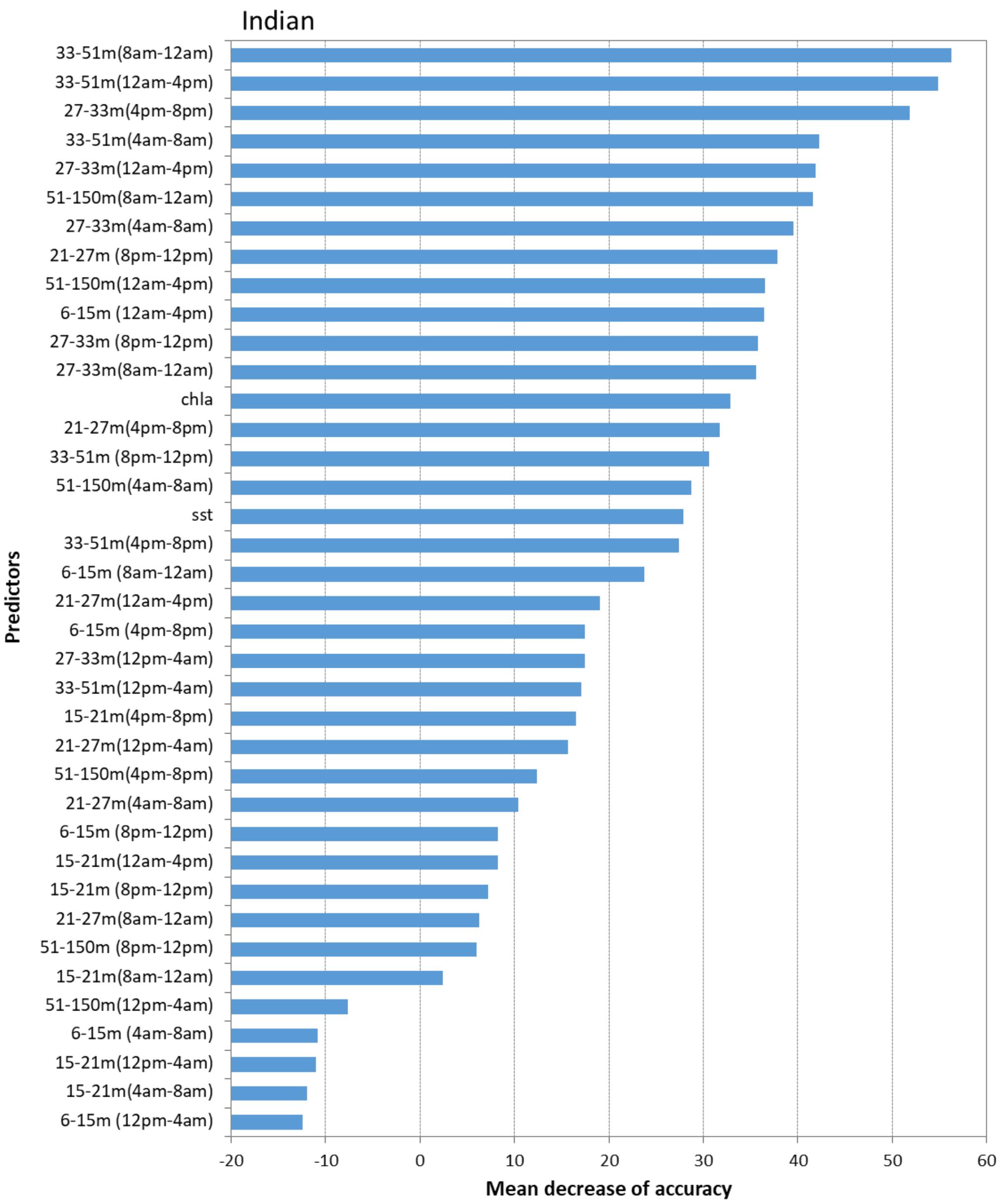

Figure 6: Predictors importance from random forest multiclass classification algorithm with the SST and Chla as predictor variables for the Indian Ocean. Acoustic predictors refer to acoustic values recorded in a depth layer at a specific day period (see Annex).

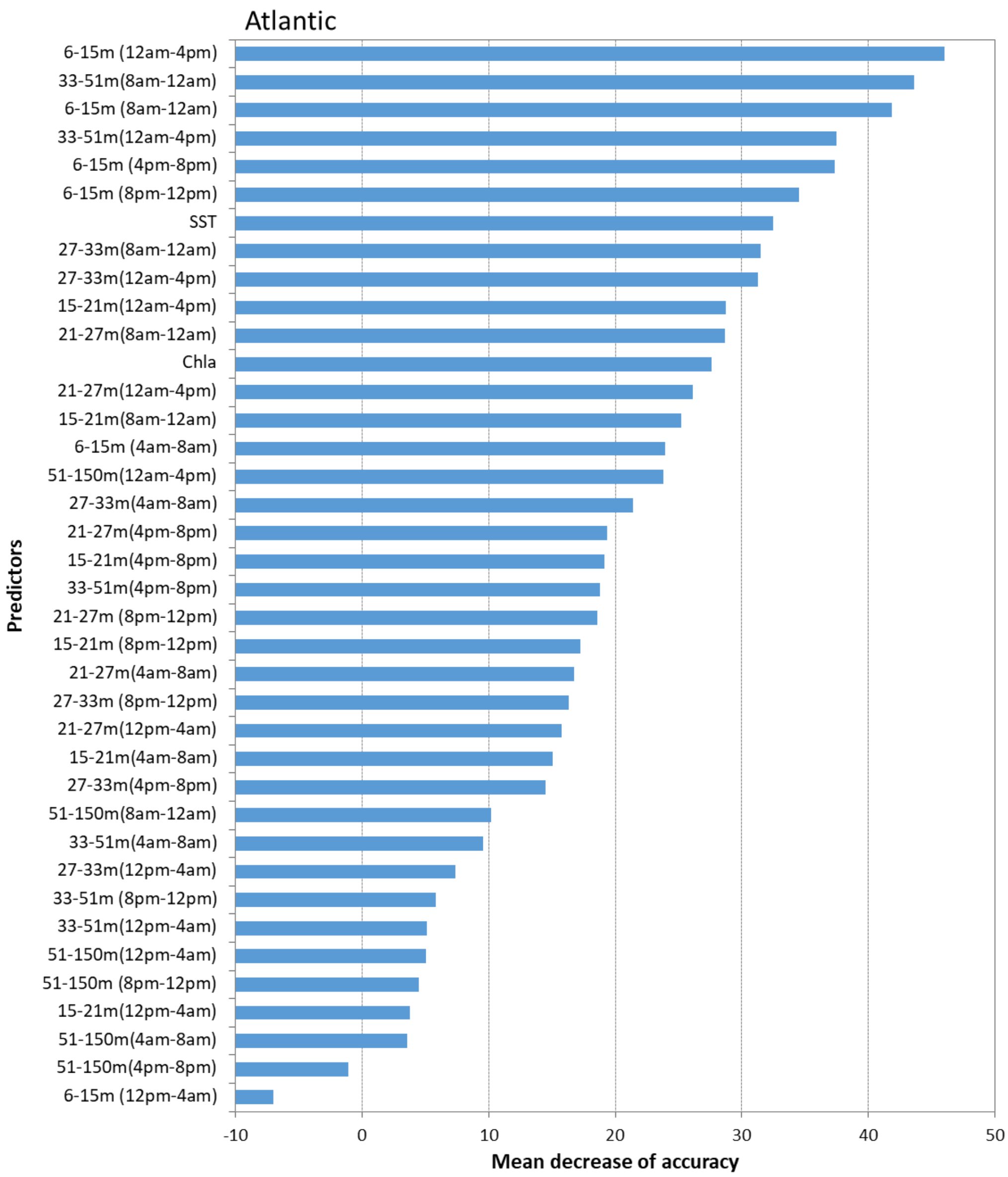

Figure 7: Predictors importance from random forest multiclass classification algorithm with the SST and Chla as predictor variables for the Atlantic Ocean. Acoustic predictors refer to acoustic values recorded in a depth layer at a specific day period (see Annex).

Catch composition

Figures 8 and 9 shows the kappa and accuracy values (performance index of the classification algorithm, see Annex) of the classification algorithm for different values of α, β indicate that there is little improvement in the multiclass classification when a lower weight is assigned to the catches of SKJ and l-YFT-BET composing the aggregation.

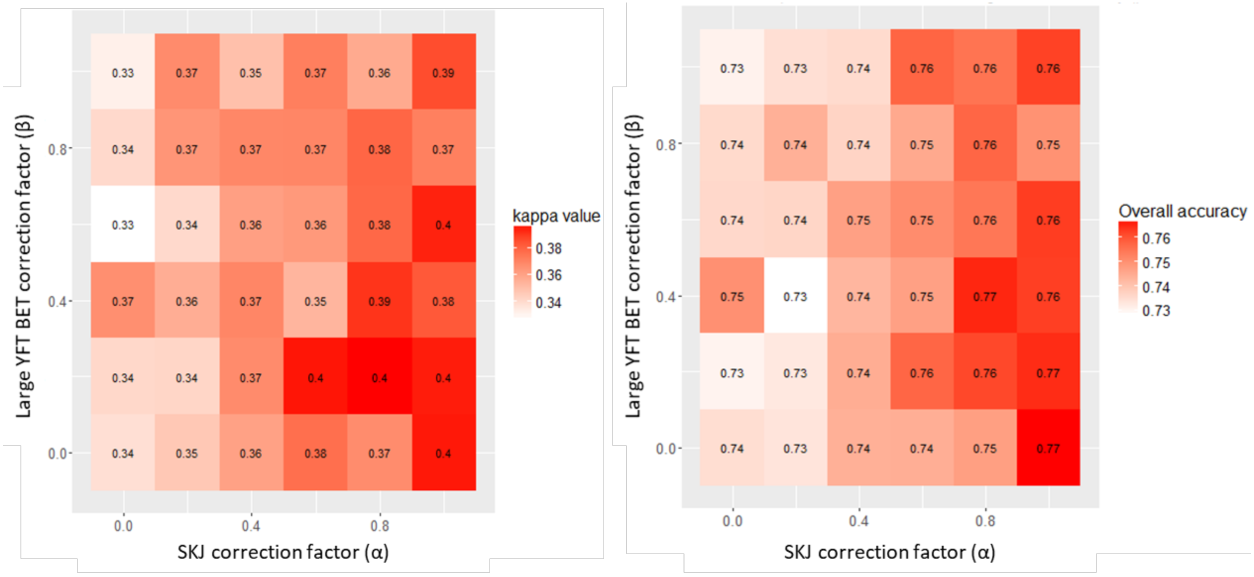

Figure 8. Heat map of kappa and accuracy values for different combinations of α, β in the Indian ocean.

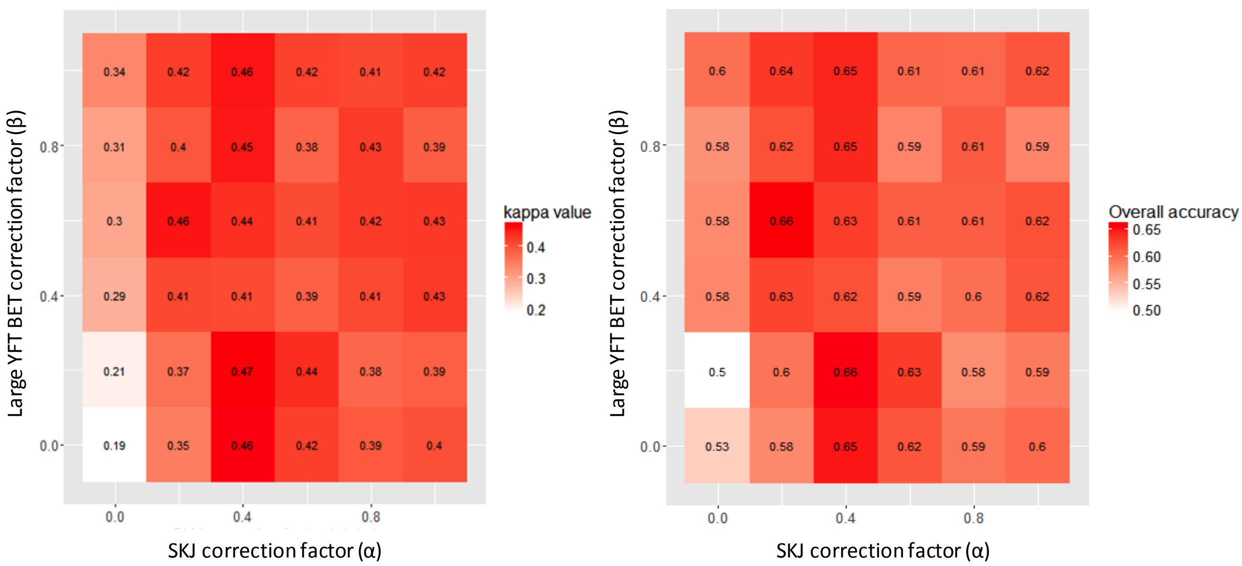

Figure 9. Heat map of kappa and accuracy values for different combinations of α, β in the Atlantic Ocean.

Conclusion

This study shows that, for the echosounder data obtained from the M3I buoys, accounting for two environmental parameters (SST and Chla) and for the species composition of the aggregation have little effects on the accuracy of biomass estimates. The analysis of predictors' importance in the multiclass classification shows that SST and Chla constitute more relevant predictors than some of the acoustic values recorded in the water column (namely acoustic data collected at surface layers at night). However, their integration in the classification process did not induce a significant improvement of model performance. Further analyses could focus on the integration of other environmental factors, like the mixed layer depth, that may be more relevant than SST and Chla. Finally, accounting for the catch composition of the aggregation did not demonstrate to improve the quality of the classification. Future attempts could be made by reducing the number of classes used for the aggregation sizes (like considering only <25t and >25t classes). Also, the application of this algorithm on other more recent buoy models, like the M3I+ Marine Instruments buoys, that work on double frequencies and has an improved hardware/software, may provide more accurate results for the size of the aggregation.

Acknowledgements

This project was co-funded by Ob7 and by the ANR project BLUEMED (ANR-14-ACHN-0002). The authors are grateful to ORTHONGEL and its contracting parties (CFTO, SAPMER, SAUPIQUET) for providing the echosounder buoys data. The authors also thank all the skippers who gave time to share their experience and knowledge on the echosounder buoys. The authors sincerely thank the contribution of the staff of the Ob7 "Observatoire des Ecosystèmes Pélagiques Tropicaux exploités" for providing the echosounder buoys and observers' data. We are also deeply grateful to the buoy manufacturers for their useful advice and information on their echosounder buoys.

References

Baidai, Y., M. Capello, N. Billet, L. Floch, M. Simier, P. Sabarros, and L. Dagorn. 2017. Towards the derivation of fisheries-independent abundance indices for tropical tunas: progress in the echosounders buoys data analysis Y. BAIDAI. 2017–WPTT19–22_Rev1, Mahé.

Bamber, J. C., and C. R. Hill. 1979. Ultrasonic attenuation and propagation speed in mammalian tissues as a function of temperature. Ultrasound in Medicine and Biology 5(2):149–157.

Bertrand, A., E. Josse, and J. Massé. 1999. In situ acoustic target-strength measurement of bigeye (Thunnus obesus) and yellowfin tuna (Thunnus albacares) by coupling split-beam echosounder observations and sonic tracking. ICES Journal of Marine Science 56(1):51–60.

Boyce, D., D. Tittensor, and B. Worm. 2008. Effects of temperature on global patterns of tuna and billfish richness. Marine Ecology Progress Series 355:267–276.

Boyra, G., G. Moreno, B. Sobradillo, I. Pérez-Arjona, I. Sancristobal, and D. A. Demer. 2018. Target strength of skipjack tuna (Katsuwanus pelamis) associated with fish aggregating devices (FADs). ICES Journal of Marine Science 75(5):1790–1802.

Brehmer, P., G. Sancho, V. Trygonis, D. Itano, J. Dalen, A. Fuchs, A. Faraj, and M. Taquet. 2018. Towards an Autonomous Pelagic Observatory: Experiences from Monitoring Fish Communities around Drifting FADs. Thalassas: An International Journal of Marine Sciences:1–13. Thalassas: An International Journal of Marine Sciences.

Capello, M., J. L. Deneubourg, M. Robert, K. N. Holland, K. M. Schaefer, and L. Dagorn. 2016. Population assessment of tropical tuna based on their associative behavior around floating objects. Scientific Reports 6(1):36415.

Josse, E., and A. Bertrand. 2000. In situ acoustic target strength measurements of tuna associated with a fish aggregating device. ICES Journal of Marine Science 57(4):911–918.

Lopez, J., G. Moreno, I. Sancristobal, and J. Murua. 2014. Evolution and current state of the technology of echo-sounder buoys used by Spanish tropical tuna purse seiners in the Atlantic, Indian and Pacific Oceans. Fisheries Research 155(March):127–137.

Maufroy, A., E. Chassot, R. Joo, and D. M. Kaplan. 2015. Large-scale examination of spatio-temporal patterns of drifting Fish Aggregating Devices (dFADs) from tropical tuna fisheries of the Indian and Atlantic Oceans. PLoS ONE 10(5):1–21.

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

ORNL DAAC. 2018. MODIS and VIIRS Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed March 28, 2018. Subset obtained for MOD13Q1 product at various sites in Spatial Range: N=90N, S=90N, E=180W, W=180W, time period: 2014-01-01 to 2017-12-01, and subset size: 9 x 9 km. https://doi.org/10.3334/ORNLDAAC/1379.

Santiago, J., J. Lopez, G. Moreno, I. Quincoces, M. Soto, and H. Murua. 2016. Towards a Tropical Tuna Buoy-derived Abundance Index (TT-BAI). Col. Vol. Sci. Pap. ICCAT 72(3):714–724.

Straube, W. L., and R. M. Arthur. 1994. Theoretical estimation of the temperature dependence of backscattered ultrasonic power for noninvasive thermometry. Ultrasound in Medicine & Biology 20(9):915–922.

Annex: Paper "SUPERVISED LEARNING APPROACH FOR DETECTING PRESENCE-ABSENCE OF TUNA UNDER FAD FROM ECHOSOUNDER BUOYS DATA." Y. Baidai, M. Capello, M.J. Amande, D. Gaertner and L. Dagorn.

SUPERVISED LEARNING APPROACH FOR DETECTING PRESENCE-ABSENCE OF TUNA UNDER FAD FROM ECHOSOUNDER BUOYS DATA.

Y. Baidai[12], M. Capello[2], M.J. Amande[1], D. Gaertner[2] and L. Dagorn[2]

SUMMARY

This study presents a new methodology for analyzing acoustic data from commercial echosounder buoys that equip the drifting FADs (DFADs) deployed by tuna purse-seiners. Our approach is based on specific processing of acoustic information, combined with machine learning methods, to translate the raw outputs provided by the buoys into metrics of tuna presence and abundance. The classifications were built from a training dataset constituted from cross-referencing of acoustic and catch data recorded on the same schools, considered as tuna occurrences, and acoustic data recorded a few days after a deployment of new DFADs considered as tuna absences. Our results demonstrate that the detection of tuna aggregations from echosounder buoys was typically more effective during daytime periods and at ocean-specific depths. The approach has very good efficiency for pattern recognition of presence and absence of tuna aggregation under DFADs regardless of the ocean (84 and 87 % of correct predictions respectively in Atlantic and Indian Ocean), but is less accurate for estimating the precise range of aggregation sizes. Future research lines for improving models performance are discussed. This work constitutes the first steps towards the development of novel fishery-independent indices of abundance for tropical tuna based on acoustic data.

KEYWORDS

Echosounder, Fish detection, Fishing buoys, Floating structures, Purse seining, Tuna fisheries

1. Introduction

A wide diversity of marine fauna is known to aggregate naturally under floating objects. Although still poorly explained, this behavior is widely exploited by the fishermen, who deploy artificial floating objects (called fish aggregating devices or FADs) worldwide to improve their catches (Kojima 1956 ; Biais and Taquet 1990 ; Kakuma 2000 ; Taquet 2004). In the tropical tuna purse-seine fishery, the use of artificial drifting devices was first introduced in the late 1980s in the Eastern Pacific ocean by the US purse seine fleet (Lennert-Cody and Hall 2000) and was later extended to all oceans and fleets from the 1990s (Fonteneau et al. 2013). These devices are referred to as Drifting Fish Aggregating Devices (DFAD), as opposed to anchored FADs, used near the coasts and mainly operated by artisanal fisheries (Itano et al. 2004; Scott et al. 2014). DFADs significantly improve the probability of success of tuna seiners' fishing sets (Fonteneau et al. 2000) and, currently, over half of the yearly tuna catches worldwide originate from DFADs (Miyake et al. 2010; Fonteneau et al. 2013). The instrumentation of DFADs with GPS beacons and echosounder buoys, which occurred in the mid and late 2000s, respectively, led to major changes in the fishing strategies and behavior of the purse-seine fleet (Lopez et al. 2014). By providing purse seiners with real-time remote information on the precise location of DFADs and the estimated size of the aggregation, echosounder buoys allow the elaboration of more productive exploration pathways, thereby contributing to the fishing efficiency (Suuronen et al. 2012). On the other hand, instrumenting DFADs with a constantly evolving technology (Lopez et al. 2014), grant them the status of a privileged platform for observing the epipelagic communities and scientists recently proposed to use the data collected at DFADs for scientific purposes (Moreno et al. 2016; Brehmer et al. 2018).

One of the major opportunities offered by DFADs in scientific research is the possibility of using the biomass estimates provided by the echosounder buoys that equip them to provide fisheries-independent abundance indices for tropical tuna (Santiago et al. 2016, Capello et al 2016, Baidai et al. 2017). However, the direct use of the biomass estimates provided by the buoys is limited by the reliability and variability of the information provided, that depend on the hardware and software characteristics of the buoys and can vary depending on the buoy manufacturers (Lopez et al. 2014; Moreno et al. 2016, Santiago et al 2016). As a result, the data provided

[1] MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France yannick.baidai@ird.fr (corresponding author)
[2] Centre de Recherches Océanologiques (CRO), DRAV, 29, Rue des pêcheurs, BPV 18, Abidjan, Côte d'Ivoire

by echosounder buoys are of heterogeneous types, natures and formats and little studies provided an assessment of their accuracies to the purpose of exploiting them for scientific purposes (Lopez et al. 2014, Baidai et al 2017). Moreover, in order to limit the volume of satellite communications, only simplified values of the acoustic energy detected by the buoys at different layers are transmitted to the satellite. The loss of raw acoustic data is a major limitation to the use of conventional echo-integration techniques for scientific purposes and requires the development of dedicated approaches to estimate the presence and size of aggregations.

This paper proposes a new method for processing the echosounder data collected by the M3I Marine Instruments echosounder buoys, one of the main buoy models which equip the DFADs deployed by tuna purse-seiners. Thanks to the availability of a large acoustic database (all echosounder buoys deployed by the French fleet), we could obtain biomass estimates using an empirical approach, based on a supervised learning algorithm trained on actual catch data reported by on-board observers on the same aggregations. The accuracy of the method could be evaluated in the Atlantic and Indian Ocean. This method offers the advantage of being easily transferable and adaptable to other existing buoy models and brands.

2. Material and Methods

2.1. Database description

2.1.1. Echosounder buoy database

We considered the echosounder buoys database hosted by Ob7/IRD, constituting the totality of the Marine Instruments echosounder buoys deployed by the French fleet between 2010 and nowadays. Our study focused on the period 2014-2017. Four different buoy models constituted the database: MSI, M3I, M4I and M3I+ (Figure 1). These buoys are all equipped with GPS positioning devices, and differ mainly in the presence and specifications of their echosounders (Table 1). Of the 37526 unique buoy identifiers in the database from 2014 to 2017, the M3I model was the most represented (Figure 1).This model is equipped with an echosounder powered by solar panels operating at a frequency of 50 KHz, with a power of 500 W, and a beam angle of 42° (Figure 2).

In its default operating mode, the M3I buoy is programmed to sample the water column at 5-minutes intervals (288 pings per day). The acoustic data are processed by an internal buoy module that saves them, selects and transmits the best recorded sample (namely the sample that corresponds to the highest biomass) every 2-hour (default mode). Data exports are then carried out by satellite communications (6 samples per message transmitted every 12 hours in the default mode), during which GPS coordinates are also acquired. In addition to acoustic and geolocation data, these buoys also provide information on water temperature and battery level.

The acoustic data generated by M3I buoys over the 2-hours sampling period correspond to 50 integer scores ranging from 0 to 7, representing the intensity of the acoustic signal detected at layers of 3-meters depth over a total detection range of 150 m (Figure 2). In the following, these 50 discrete scores are referred to as “acoustic sample”. Each acoustic sample is also associated with a biomass index calculated by an internal buoy algorithm.

2.1.2 Observers data

We considered observers’ data recorded during 2014-2017 from the observers programs “OCUP” and “DCF” (IRD) conducted in the Atlantic and Indian oceans. Two types of data were used to build the training dataset used in the analysis: DFAD deployment data and catch data for sets conducted on DFADs. The former consisted on the deployment of a new instrumented DFAD in the water, reporting the date and time of the deployment, GPS position and buoy unique ID. The catch data consisted of the estimated catch size for the tuna and bycatch species caught at DFADs, together with the GPS location of the set and the unique identifier of the buoy equipping the DFAD. Only the data for which the unique buoy identifier corresponded to a buoy identifier present in the echosounder buoys database were retained in the analysis. For catch data, null fishing sets and fishing sets for which the positions reported by the buoy and the observer were not consistent were excluded from the analysis. The final database consisted of 1509 catch data and 4506 deployment data (see Table 2). Catch data were also categorized into three classes of 0-10 tons, 10-25 tons and >25 tons, by summing the recorded catches of the target tuna species (yellowfin tuna, bigeye tuna and skipjack tuna), see Table 3.

2.2. Acoustic data pre-processing

In order to derive a buoy biomass index on a daily basis, we considered, for each buoy and day, a 50 × 12 matrix, reporting the acoustic scores recorded at different depths (50 rows) and different times of the day (every 2 hours in the default operating mode of the buoy, corresponding to 12 columns). To reduce the dimensionality of the problem, this matrix was pre-processed and the temporal and spatial information aggregated. First, in order to aggregate the acoustic samples over time, the 24-hours data were aggregated over 6 slices of 4 hours each, each slice containing the acoustic sample whose sum of scores corresponded to the maximum recorded over the period considered. This resulted in a matrix of 6 columns (one for each time slot), and 50 rows for the different depth layers. Secondly, clustering methods were used for identifying groups of homogeneous layers. The cluster analysis was based on a dissimilarity matrix computed from Euclidean distance and Ward's method for merging clusters (Murtagh and Legendre 2014)[3]. In each of the two oceans considered (Atlantic and Indian), excluding the first two layers (corresponding to the transducer blanking zone), 6 groups of layers were identified (Figure 3 and 4). The acoustic scores (denoted below as s_i^G) recorded for each of the *i* layers constituting a group G were summed and scaled to obtain a single score (S'_G) per group, ranging between 0 and 1, according to Eq.1.

$$S'_G = \frac{(\sum_{i=1}^{n_G} s_i^G)}{\text{maxs } n_G} \quad \text{(Eq. 1)}$$

where the sum runs over the n_G layers belonging to group G and *maxs* is a constant denoting the maximum value of the score (7 in the case of M3I buoys). This pre-processing allowed obtaining a daily matrix of 6 rows (groups of layers) and 6 columns (time slots), referred in this study as "daily acoustic matrix (Figure 5).

2.3. Supervised learning classification algorithm

We applied a random forest classification algorithm (Breiman 2001)[4], considering each ocean separately. The training datasets for each ocean were constructed from the cross-matching of observers data (see paragraph 2.1.2) and the daily acoustic matrices corresponding to the same buoy ID (see paragraph 2.2). Two types of classification algorithms were considered: a binary classification algorithm describing the absence or presence of tuna, and a multiclass classification considering different sizes of aggregations under DFAD (no tuna, less than 10 tons, between 10 and 25 tons, more than 25 tons). The acoustic data recorded 5 days after a new DFAD deployment were used to build the “no tuna” or “absence” training database. The rationale for choosing a 5-days period after deployment is related to the need of accounting for the acoustic signal produced by the bycatch species within the “no tuna” class.

A stratified down-sampling procedure was applied for building the training dataset, for dealing with the imbalanced number of observations in the different classes. The down-sampling procedure consisted in resampling the majority class to make their frequencies closer to the rarest class (Kuhn and Johnson 2013). Classifiers were represented by daily acoustic matrix values and predicted classes were "No tuna" and "Tunas" for presence-absence classification, and "No tuna", "less than 10", "between 10 and 25 tons" and "more than 25 tons", for multiclass classification.

The importance of the different predictors in the classification process (combination between layers group and day period) in each ocean was assessed through analysis of the mean decrease accuracy (increase of prediction error after permutation of a variable while all others are left unchanged, during tree construction), in the random forest model (Breiman, 2001).

Model training and evaluation were performed through a two-fold cross-validation replicated 10 times (Dietterich 1998). The original dataset was divided into two subsets, in each of the 10 replica: the training set (75% of the initial data volume) used for training and building the classification models, and the validation dataset (remaining 25%) used for model performance evaluation.

2.4. Model evaluation

Several metrics were considered to evaluate the performance of the classification algorithm. The overall accuracy (proportion of correct prediction) and the kappa coefficient (Cohen 1968) were used to assess the overall performance of both binary and multiclass classifications. Kappa coefficient is a reliability index varying between 0 and 1, estimated according to (Eq. 2)

---

[3] The cluster analysis was conducted using the R package “cluster” (Maechler et al, 2018)

[4] It was conducted with the R package "random forest" (Liaw and Wiener 2014).

$$kappa = \frac{Pr(a) - Pr\ (e)}{1 - Pr\ (e)} \qquad \text{(Eq. 2)}$$

where *Pr(a)* is the total proportion of agreement between the two classifications and *Pr(e)* is the theoretical proportion of agreement expected by chance. The closer this ratio is to 1, the better the classification performed.

In presence-absence classification, sensitivity, specificity, precision and F1 score were also evaluated from confusion matrix, using Eqns. 3-6:

$$Sensitivity = \frac{TP}{TP+FN} \qquad \text{(Eq. 3)}$$

$$Specificity = \frac{TN}{FP+TN} \qquad \text{(Eq. 4)}$$

$$Precision = \frac{TP}{TP+FP} \qquad \text{(Eq. 5)}$$

$$F1 \ = \ 2\left(\frac{precision * sensitivty}{precision + sensitivty}\right) \qquad \text{(Eq. 6)}$$

where TP (true positive) and TN (true negative) are the proportions of presence (respectively absence) correctly classified; FN (false negative) and FP (false positive) are the proportions of absence (respectively presence) incorrectly predicted.

Sensitivity (recall or true positive rate) measures the efficiency of the algorithm in correctly classifying positive cases, and specificity (or true negative rate) measures the efficiency of the algorithm in correctly classifying negative cases. Precision (also called positive predictive value) is the fraction of correctly predicted presence among tuna presence prediction. The F1 score is the harmonic average of the precision and recall, where an F1 score reaches its best value at 1 (perfect precision and recall) and worst at 0. The same evaluation metrics recalculated at the level of classes, were also considered for the multiclass classification. These validation metrics were computed from the 10-fold cross-validation, with the R package “caret” (Kuhn, 2018).

3. Results

3.1. Aggregation of the depth layers sampled by the buoy

The 6 groups of depth layers identified from cluster analysis during data pre-processing (Figure 3) are characterized by significantly different score values in their constituent layers (p-value at Kruskall-Wallis test < 0.001 for both Indian and Atlantic Oceans). These values decline strongly with depth. The deepest group of layers (which is also the one that aggregates the largest number of depth layers, in the Atlantic and Indian oceans), exhibits the lowest acoustic values, with average score values in its constituent layers, almost equal to zero (Figure 4).

3.2. Presence/absence classification

The results for the metrics obtained for the binary classification (tuna presence or absence) are shown in Table 4. The random forests algorithm shows a good performance in successfully discriminating the presence and absence of tuna, in both oceans, with an accuracy of 84% and 87% in the Atlantic and Indian ocean, respectively. However, while in the Atlantic ocean the correct classification of positive (“Tuna”) and negative instances (“no tuna”) remains broadly similar (about 0.8 for sensitivity and specificity), in the Indian Ocean, tuna absence under FAD is more easily detected than its presence as shown by the high specificity (0.95) and low sensitivity (0.64) for this ocean.

3.3. Multiclass classification

Multi-class classification performs less than binary classification (Table 5 and 6), and in both oceans, classification algorithms are not sufficiently effective to correctly identify small (less than 10 tons) and medium tunas aggregation (between 10 and 25 tons). Nevertheless, the identification of absence and large aggregations of tunas appears relatively acceptable (precision respectively 0.80 and 0.44 in Atlantic, and 0.83 and 0.59 in Indian). The comparison of classification performance between the two oceans shows that, although the overall efficiency of multiclass classifications remains similar (kappa ≈ 0.40), the classification of aggregation sizes in the intermediate classes (less than 25 tons) appears more accurate in the Atlantic than in the Indian Ocean.

3.4. Predictors' importance

The importance of the predictors shows different patterns depending on the ocean, whereas, for the same ocean, it does not significantly differ between the binary and multiclass classification, see Figure 6. In both oceans, the acoustic data recorded during the daytime appear to be more important to assess tuna presence/absence. In the Atlantic Ocean, the most relevant time periods to detect and estimate tuna aggregations appear to be from 8 am to 4 pm and the most important depths range between 6 to 50 m depth. By contrast, in the Indian Ocean, the main predictors correspond to deeper layers (from 21 meters to 150 meters) over a more spread period of time (from 4 am to 4 pm), see Figure 6 and 7.

4. Discussion

To the purpose of providing abundance indicators for large pelagic fish species like tropical tuna, whose habitat is ocean-wide, the large amount of acoustic data collected by echosounder buoys is of undeniable value. However, this data still remains difficult to exploit directly. Although many buoy models process the data internally, and generate an index that is supposed to assess the abundance of fish under the buoy, previous studies demonstrated that these indices have a poor reliability (Lopez et al 2014). Lopez and co-authors indicated that most fishers pay little attention to these indices and their exploitation of the buoy data is rather based on the interpretation of the acoustic scores recorded at different depths and times. Our approach, based on a supervised learning algorithm trained on a set of catch and deployment data, can be assimilated to the fishermen's approach that learn how to interpret the acoustic scores based on their experience.

The pre-processing methodology presented in this paper, combined with the use of supervised learning algorithms, may constitute a valuable approach for the exploitation of buoy acoustic data, as evidenced by the satisfactory results obtained in the discrimination of the presence/absence of tuna aggregations under DFADs. The potential of automated classification for unconventional acoustic data treatment has been highlighted by the recent work of Uranga et al. (2017). These authors proposed a very efficient method for detecting the presence-absence of bluefin tuna schools (kappa values ranged between 0.74 and 0.79 for the different machine learning algorithms used), from echograms provided by the commercial medium-range sonar used on fishing vessels.

In the construction of our training dataset for tuna presence/absence, we made two main assumptions: (i) that newly deployed DFADs corresponded to tuna absences and (ii) that the DFADs associated to a fishing set corresponded to tuna presence. For the new deployments, we took the acoustic scores recorded 5 days after deployment as a signature of a DFAD without tuna. In this respect, it is important to notice that we limited the database to the deployments of new DFADs only, excluding other possible deployment events, like the addition of a buoy to an existing log/DFAD or its reinforcement. This restriction ensured that a virgin floating object, without an existing aggregation, was considered in the "tuna absence" database. The choice of considering acoustic data after a time delay of 5 days for newly deployed FADs was based on previous field studies that showed the FAD colonization of bycatch species occurs during the first days after deployment (Taquet 2004; Nelson 2003; Moreno et al 2007a). This choice allowed separating the acoustic signature of bycatch species (that may already occupy the FADs after 5 days), from the characteristic signal of tuna in the fished DFADs. Further sensitivity analyses should be conducted, by varying the size of the time window after deployment, to evaluate the sensitivity of results relative to this variable. Secondly, the assumption that a fished DFAD corresponded to tuna presence was driven by the fact that the captain's choice of fishing on a DFAD is guided by different proxies of tuna presence, not only the echosounder buoy data. For example, purse-seiners dispose of on-board echosounders and sonars, much more powerful and accurate than those of the buoys, and it is very likely that a fishing operation takes place when the on-board instruments confirmed the tuna presence at the DFADs.

On the other hand, the performance of our algorithms remains more limited for the discrimination of tuna aggregation sizes under DFADs. One possible cause may be the imbalanced percentage of data samples in the different size classes, (for example, of the four classes defined for multiclass classifications, the "No tuna" class alone represents about 80% and 50% of the training data used respectively in the Indian and the Atlantic Ocean). This imbalance in the data is known to produce undesirable results such as a lower performance on both test and training data (Provost 2000; Japkowicz and Stephen 2002; He and Garcia 2010). Despite the sampling methodology used to offset its effects, this fact could be one of the main factors involved in the decrease in efficiency observed for multiclass classification. Future research actions testing other learning techniques that deal with unbalanced data such as oversampling or SMOTE (Synthetic Minority Oversampling Technique; Chawla et al. 2002), and other classification algorithms such as neural networks (Bishop 2007), SVM (Cortes et Vapnik 1995), or gradient boosting machine (Friedman 2001), could improve the efficiency of multiclass classification. Another possible source of bias for the multiclass classification may rely in the different species composition of the aggregations considered in each class. Considering that skipjack and yellowfin/bigeye have a different acoustic response (Bertrand et al. 1999 ; Josse and Bertrand 2000 ; Boyra et al. 2018), a total aggregation size of, for example, 10 tons would provide a different acoustic signature depending on the percentage of each species. Moreover, the analysis of scores in layer groups identified by the cluster analysis reveals that layers below 50 meters are characterized by very low score values (Figure 4). Several studies on vertical species distribution under DFADs, evidenced that tuna clearly evolve below this depth (Dagorn et al. 2007a ; Dagorn et al. 2007b ; Moreno et al. 2007b ; Forget et al. 2015; Matsumoto et al. 2016 ; Lopez et al. 2017). Thus, the surprisingly low values obtained for this layer groups with this buoy model might potentially constitute a severe limitation for the accurate estimation of aggregation sizes under FADs through M3I buoys, especially for tuna species which are vertically distributed beyond this limit. Finally, catch data could not reflect the full size of the aggregation and part of the school may have escaped from the net, thus inducing an underestimation for certain catch sizes that thus do not reflect the acoustic data of the buoy.

The analysis of the importance of the predictive factors in the random forest classifications has shown that, regardless of the ocean, the daytime period appears to be the most relevant to distinguish the presence of tuna schools from other acoustic targets. This outcome could be linked to the behavior of fish schools and their spatial and temporal distribution around DFAD. Indeed, the accuracy of the measurement made by the echosounder remains highly dependent on the characteristics and the position of fish schools in relation to that of the measuring instrument. Sonar surveys conducted on FADs in the Indian ocean, revealed that fish form a large number of small and dispersed schools during nighttime, and a few and large schools during daytime (Trygonis et al. 2016), more easily detectable at this period. However, as evidenced by Lopez et al. (2017), this behavioral patterns are highly variable and appear to be both species and region specific. Another possible reason for the higher importance of daytime detection could be related to the influence of deep scattering layers (Hobert 1962; Chapman and Marshall 1966; Robinson and Gómez-Gutiérrez 1998), whose nycthemeral upwards migration may affect the acoustic signal. Also, this study highlighted the differences in the relevance of the water column strata for the detection of tuna schools between the Atlantic and Indian Ocean. At first sight, one would be tempted to explain these variations by a vertical stratification of species, differing according to the oceans (deeper and more segregated tuna schools in the Indian than in the Atlantic Ocean, linked to the lower depth of the thermocline in the Indian Ocean). Further studies, comparing the presence-at-depth profiles of tuna species in the two oceans, should be conducted to confirm these findings.

Finally, environmental factors could affect the acoustic signal detection and fish behavior and could thus have an effect on the classification of the aggregation size. For example, it is known that water temperature has an effect on both the acoustic signal (Straube and Arthur 1994; Bamber and Hill 1979) and the abundance of tuna (Boyce et al. 2008). The integration of these parameters into the classification algorithm may eventually lead to a better algorithm performance.

5. Conclusion

The methodology developed in this study provides accurate results for the assessment of presence/absence of tuna schools at DFADs in both oceans. This approach can be applied to other buoy models and particularly to the recent M3I+ buoy model, a novel multi-frequency buoy model that is being widely adopted in recent years. Despite further improvements of the algorithm can be still operated, particularly for the assessment of the aggregation size, the accurate assessment of presence/absence of tuna schools at the DFADs opens the way towards the exploitation of echosounder buoy data to provide novel and robust indicators for the management of FAD fisheries in future years.

6. Acknowledgements

This project was co-funded by Ob7 and by the ANR project BLUEMED (ANR-14-ACHN-0002). The authors sincerely thank the contribution of the staff of the Ob7 "Observatoire des Ecosystèmes Pélagiques Tropicaux exploités" for providing the echosounder buoys and observers' data. We are also deeply grateful to the buoy manufacturers for their useful advice and information on their echosounder buoys.

References

Baidai, Y., M. Capello, N. Billet, L. Floch, M. Simier, P. Sabarros, and L. Dagorn. 2017. Towards the derivation of fisheries-independent abundance indices for tropical tunas: progress in the echosounders buoys data analysis Y. BAIDAI. 2017–WPTT19–22_Rev1, Mahé.

Bamber, J. C. and Hill, C. R. 1979. Ultrasonic attenuation and propagation speed in mammalian tissues as a function of temperature. Ultrasound Med. Biol. 5, 149–157.

Bertrand, A., E. Josse, and J. Massé. 1999. In situ acoustic target-strength measurement of bigeye (Thunnus obesus) and yellowfin tuna (Thunnus albacares) by coupling split-beam echosounder observations and sonic tracking. ICES Journal of Marine Science. 56(1):51–60.

Biais G. and Taquet M. 1990. Dispositifs de concentration de poisons à la Reunion. Equinoxe 34, 20-26.

Bishop, C. M. 2007. Pattern Recognition and Machine Learning. J. Electron. Imaging 16, 049901.

Boyce, D., Tittensor, D. and Worm, B. 2008. Effects of temperature on global patterns of tuna and billfish richness. Mar. Ecol. Prog. Ser. 355, 267–276.

Boyra, G., G. Moreno, B. Sobradillo, I. Perez-Arjona, I. Sancristobal, and Demer D.. 2018. Target strength of skipjack tuna (*Katsuwonus pelamis*) associated with fish aggregating devices (FADs). doi: 10.1093/icesjms/fsy041.

Brehmer, P., G. Sancho, V. Trygonis, D. Itano, J. Dalen, A. Fuchs, A. Faraj, and M. Taquet. 2018. Towards an Autonomous Pelagic Observatory: Experiences from Monitoring Fish Communities around Drifting FADs. Thalassas: An International Journal of Marine Sciences. Thalassas: An International Journal of Marine Sciences, 1-13.

Breiman, L. 2001. Random forests. Machine Learning 45(1):5–32.

Chapman, R. P. and Marshall, J. R. 1966. Reverberation from Deep Scattering Layers in the Western North Atlantic. J. Acoust. Soc. Am. 40, 405–411.

Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P. 2002. SMOTE: Synthetic Minority Over-sampling Technique. J. Artif. Intell. Res. 16, 321–357.

Cohen, J. 1968. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin 70(4):213–220.

Cortes C, Vapnik V. 1995. Support-vector networks. Machine learning. 20(3):273–297

Dagorn, L., Holland, K. N. and Itano, D. G. 2007a. Behavior of yellowfin (Thunnus albacares) and bigeye (T. obesus) tuna in a network of fish aggregating devices (FADs). Mar. Biol. 151, 595–606.

Dagorn, L. et Pincock D., Girard C., Holland K., Taquet M., Sancho G., Itano D. and Aumeeruddy R. 2007b. Satellite-linked acoustic receivers to observe behavior of fish in remote areas. Aquat. Living Resour. 20, 307–312.

Dietterich, T. G. 1998. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. Neural Computation 10(7):1895–1923.

Fonteneau, A., E. Chassot, and N. Bodin. 2013. Global spatio-temporal patterns in tropical tuna purse seine fisheries on drifting fish aggregating devices (DFADs): Taking a historical perspective to inform current challenges. Aquatic Living Resources 26(01):37–48.

Fonteneau, A., P. Pallarés, and R. Pianet. 2000. A worldwide review of purse seine fisheries on FADs. Pêche thonière et dispositifs de concentration de poissons, Caribbean-Martinique, 15-19 Oct 1999 (1):15–35.

Forget, F. G., Capello M., Filmalter J. D., Govinden R., Soria M., Cowley P.D., and Dagorn L. 2015. Behaviour and vulnerability of target and non-target species at drifting fish aggregating devices (FADs) in the tropical tuna purse seine fishery determined by acoustic telemetry. Can. J. Fish. Aquat. Sci. 72, 1398–1405.

Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. Ann. Stat. 29, 1189–1232.

He, H. and Garcia, E. A. 2009. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 21, 1263–1284.

Itano, D., S. Fukofuka, and D. Brogan. 2004. The development, design and recent status of anchored and drifting FADs in the WCPO. 17th Meeting of the Standing Committee on Tuna and Billfish (August):1–25.

Japkowicz, N. and Stephen, S. 2002. The class imbalance problem: A systematic study. Intell. Data Anal. 6, 429–449.

Josse, E., and A. Bertrand. 2000. In situ acoustic target strength measurements of tuna associated with a fish aggregating device. ICES Journal of Marine Science. 57(4):911-918.

Kakuma, S. 2000. Synthesis on moored FADs in the North West Pacific region:1–3.

Kojima S. 1956. Fishing for dolphins in the western part of the Japan Sea. II. Why do the fish take shelter under floating materials? Bull. Jap. Soc. Sci. Fish. 21, 1049-1052

Kuhn, M., and K. Johnson. 2013. Applied predictive modeling (Vol. 26). Springer. New York. 620pp

Kuhn M. Contributions from Wing J., Weston S., Williams A., Keefer C., Engelhardt A., Cooper T., Mayer Z., Kenkel B., the R Core Team, Benesty M., Lescarbeau R., Ziem A., Scrucca L., Tang Y., Candan C. and Hunt T. 2018. caret: Classification and Regression Training. R package version 6.0-80. https://CRAN.R-project.org/package=caret

Lennert-Cody, C. E., and M. A. Hall. 2000. The development of the purse seine fishery on drifting Fish Aggregating Devices in the Eastern Pacific Ocean: 1992-1998. Pêche thonére et dispositifs de concentration de poissons: Colloque Caraïbe-Martinique, Trois-Ilets:78–107.

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

Lopez, J., G. Moreno, I. Sancristobal, and J. Murua. 2014. Evolution and current state of the technology of echo-sounder buoys used by Spanish tropical tuna purse seiners in the Atlantic, Indian and Pacific Oceans. Fisheries Research 155(March):127–137.

Lopez, J., Moreno, G., Ibaibarriaga, L. and Dagorn, L. 2017. Diel behaviour of tuna and non-tuna species at drifting fish aggregating devices (DFADs) in the Western Indian Ocean, determined by fishers' echo-sounder buoys. Mar. Biol. 164, 44.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. 2018. cluster: Cluster Analysis Basics and Extensions. R package version 2.0.7-1.

Matsumoto, T., Satoh, K., Semba, Y. and Toyonaga, M. 2016. Comparison of the behavior of skipjack (Katsuwonus pelamis), yellowfin (Thunnus albacares) and bigeye (T. obesus) tuna associated with drifting FADs in the equatorial central Pacific Ocean. Fish. Oceanogr. 25, 565–581.

Miyake, M. P., P. Guillotreau, C.-H. Sun, and G. Ishimura. 2010. Recent developments in the tuna industry stocks, fisheries, management, processing, trade and markets. FAO Fisheries and Aquaculture Technical Paper 543 FAO Technical Paper 543.

Moreno, G., Dagorn, L., Sancho, G., García, D. and Itano, D. 2007a. Using local ecological knowledge (LEK) to provide insight on the tuna purse seine fleets of the Indian Ocean useful for management. Aquat. Living Resour. 20, 367–376.

Moreno, G., Josse, E., Brehmer, P. and Nøttestad, L. 2007b. Echotrace classification and spatial distribution of pelagic fish aggregations around drifting fish aggregating devices (DFAD). Aquatic Living Resources 20, 343–356.

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

Murtagh, F. and Legendre, P. 2014. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms

Implement Ward's Criterion? J. Classif. 31, 274–295 .

Nelson, P. A. 2003. Marine fish assemblages associated with fish aggregating devices (FADs): effects of fish removal, FAD size, fouling communities, and prior recruits. Fish. Bull. 101, 835–850.

Provost, F. 2000. Machine learning from imbalanced data sets 101. Proc. AAAI'2000 Work. imbalanced data sets. 1–3. doi:10.1.1.33.507

Robinson, C. J. and Gómez-Gutiérrez, J. 1998. Daily vertical migration of dense deep scattering layers related to the shelf-break area along the northwest coast of Baja California, Mexico. J. Plankton Res. 20, 1679–1697.

Santiago, J., J. Lopez, G. Moreno, I. Quincoces, M. Soto, and H. Murua. 2016. Towards a Tropical Tuna Buoy-derived Abundance Index (TT-BAI). Col. Vol. Sci. Pap. ICCAT 72(3):714–724.

Scott, Gerald P.;Lopez, J. 2014. The use of FADs in tuna fisheries. Page European Union. Directorate General for Internal Policies. 70p

Straube, W. L. and Arthur, R. M. 1994. Theoretical estimation of the temperature dependence of backscattered ultrasonic power for noninvasive thermometry. Ultrasound Med. Biol. 20, 915–922.

Suuronen, P., F. Chopin, C. Glass, S. Løkkeborg, Y. Matsushita, D. Queirolo, and D. Rihan. 2012. Low impact and fuel efficient fishing-Looking beyond the horizon. Fisheries research 119 : 135-146.

Taquet, M. 2004. Le comportement agrégatif de la dorade coryphène (Coryphaena hippurus) autour des objets flottants. Thèse de Doctorat de l'Université de Paris 6, Océanologie biologique, 168p.

Trygonis, V., Georgakarakos, S., Dagorn, L. and Brehmer, P. 2016. Spatiotemporal distribution of fish schools around drifting fish aggregating devices. Fish. Res. 177, 39–49.

Uranga J, Arrizabalaga H, Boyra G, Hernandez MC, Goñi N, Arregui I, Fernandes J. A., Yurramendi Y. and Santiago J. 2017. Detecting the presence-absence of bluefin tuna by automated analysis of medium-range sonars on fishing vessels. PLoS ONE 12(2): 1–18. doi:10.1371/journal.pone.0171382

Table 1 : Main technical specifications of Marine Instruments buoys

	MSI	*M3I*	*M4I*	*M3I+*
Year	2010	2010	2012	2016
Satellite GPS :	Yes	Yes	Yes	Yes
Echo-sounder :	No	Yes	Yes	Yes
Frequency :	-	50 kHz	50, 120, 200 KHz	50 and 200 KHz
Power :	-	500 W	500 W	500 W
Resolution per layer :	-	3 m	3 m	3 m
Range :	-	150 m	150 m	150 m
Blind area :	-	6 m	6 m	6 m
Soundings :	-	each 5 minutes	each 5 minutes (in three frequencies)	each minute (in two frequencies)
TVG correction :	-	No	No	Yes

Table 2 : Number of catch and deployment data used in the presence-absence classification for the Atlantic and Indian ocean.

	Catch data	**Deployment data**
Atlantic	567	562
Indian	942	3944

Table 3: Number of catch and deployment data used in the multiclass classification for the Atlantic and Indian ocean.

	No tuna	**< 10 tons**	**[10, 25 tons]**	**> 25 tons**
Atlantic	562	176	217	174
Indian	3944	233	346	363

Table 4 : Summary of tuna presence/absence classification performances for the Atlantic and Indian ocean: mean and standard deviation values (in bracket) of evaluation metrics.

Evaluation Metric	Atlantic	Indian
Accuracy	0.84 (0.005)	0.87 (0.002)
Kappa	0.67 (0.03)	0.63 (0.02)
Sensitivity	0.81 (0.02)	0.64 (0.02)
Specificity	0.87 (0.02)	0.95 (0.005)
Precision	0.88 (0.02)	0.80 (0.01)
F_1 score	0.84 (0.01)	0.71 (0.01)

Table 5 : Summary of multiclass classification performances for the Atlantic Ocean. Mean and standard deviation (in bracket) of evaluation metrics.

Evaluation Metric	Atlantic				
	No tuna	<10 tons	[10 , 25 tons]	> 25 tons	**Average**
Sensitivity	0.86 (0.04)	0.37 (0.05)	0.35 (0.06)	0.39 (0.09)	0.49
Specificity	0.82 (0.02)	0.88 (0.02)	0.85 (0.03)	0.89 (0.03)	0.86
Precision	0.80 (0.03)	0.35 (0.07)	0.38 (0.05)	0.44 (0.08)	0.49
F_1 score	0.83 (0.03)	0.36 (0.05)	0.36 (0.05)	0.41 (0.08)	0.49
Accuracy			0.60 (0.05)		
Kappa			0.40 (0.04)		

Table 6: Summary of multiclass classification performance for Indian Ocean. Means and standard deviations (in bracket) of evaluation metrics by classes.

Evaluation Metric	Indian				
	No tuna	<10 tons	[10 , 25 tons]	> 25 tons	**Average**
Sensitivity	0.95 (0.005)	0.13 (0.037)	0.21 (0.03)	0.40 (0.03)	0.42
Specificity	0.54 (0.02)	0.96 (0.004)	0.95 (0.006)	0.97 (0.004)	0.85
Precision	0.83 (0.01)	0.17 (0.03)	0.35 (0.05)	0.59 (0.004)	0.48
F_1 score	0.89 (0.008)	0.14 (0.03)	0.27 (0.03)	0.47 (0.03)	0.44
Accuracy			0.75 (0.01)		
Kappa			0.39 (0.02)		

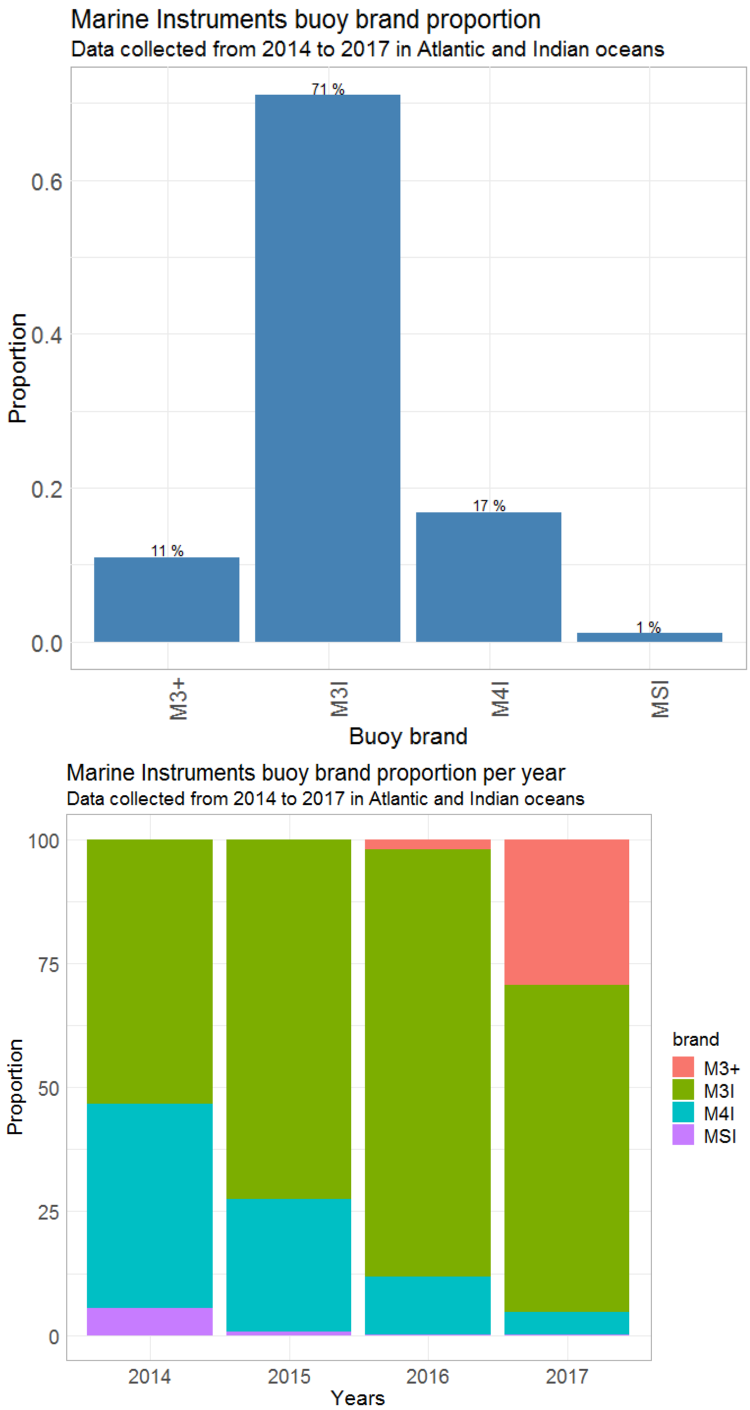

Figure 1: Characteristics of the echosounder buoys database from 2014 to 2017.

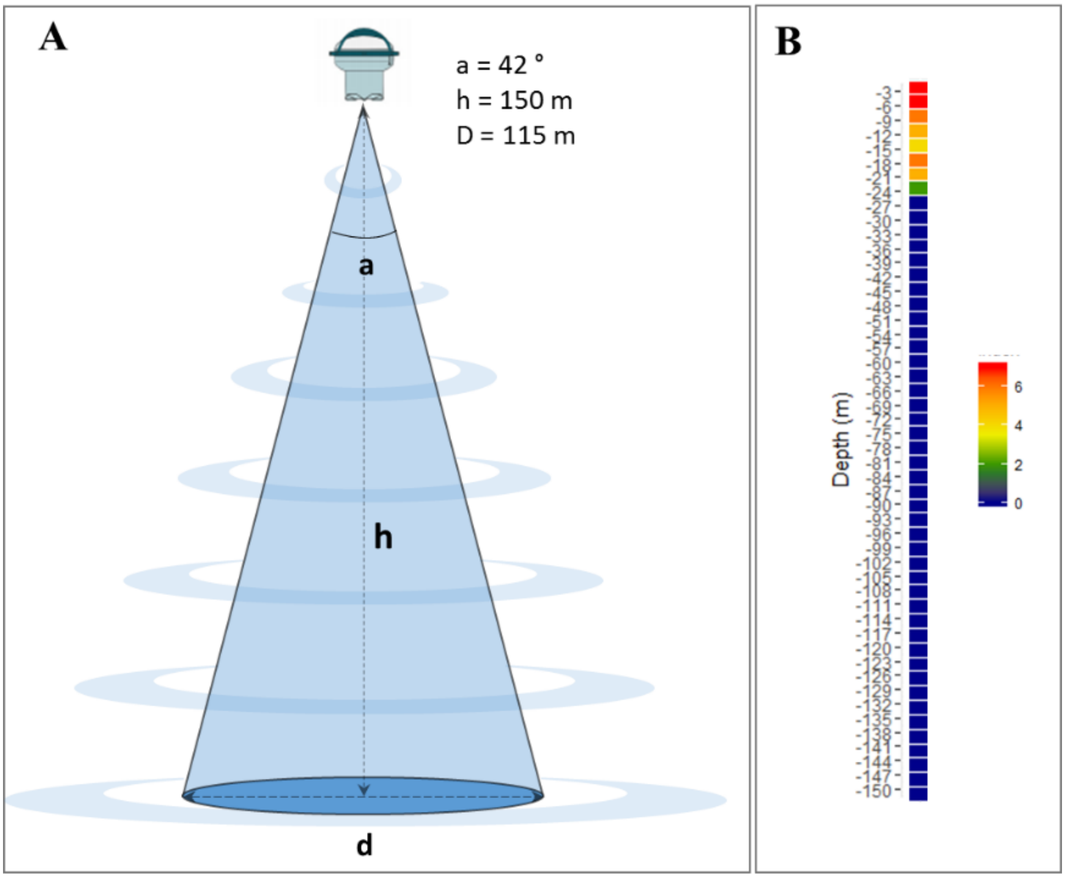

Figure 2: (A) Beam width (or angle) (a), depth range (h), and diameter (d) at 150 m of the Marine Instruments M3I echosounder buoy; (B) Example of an acoustic sample

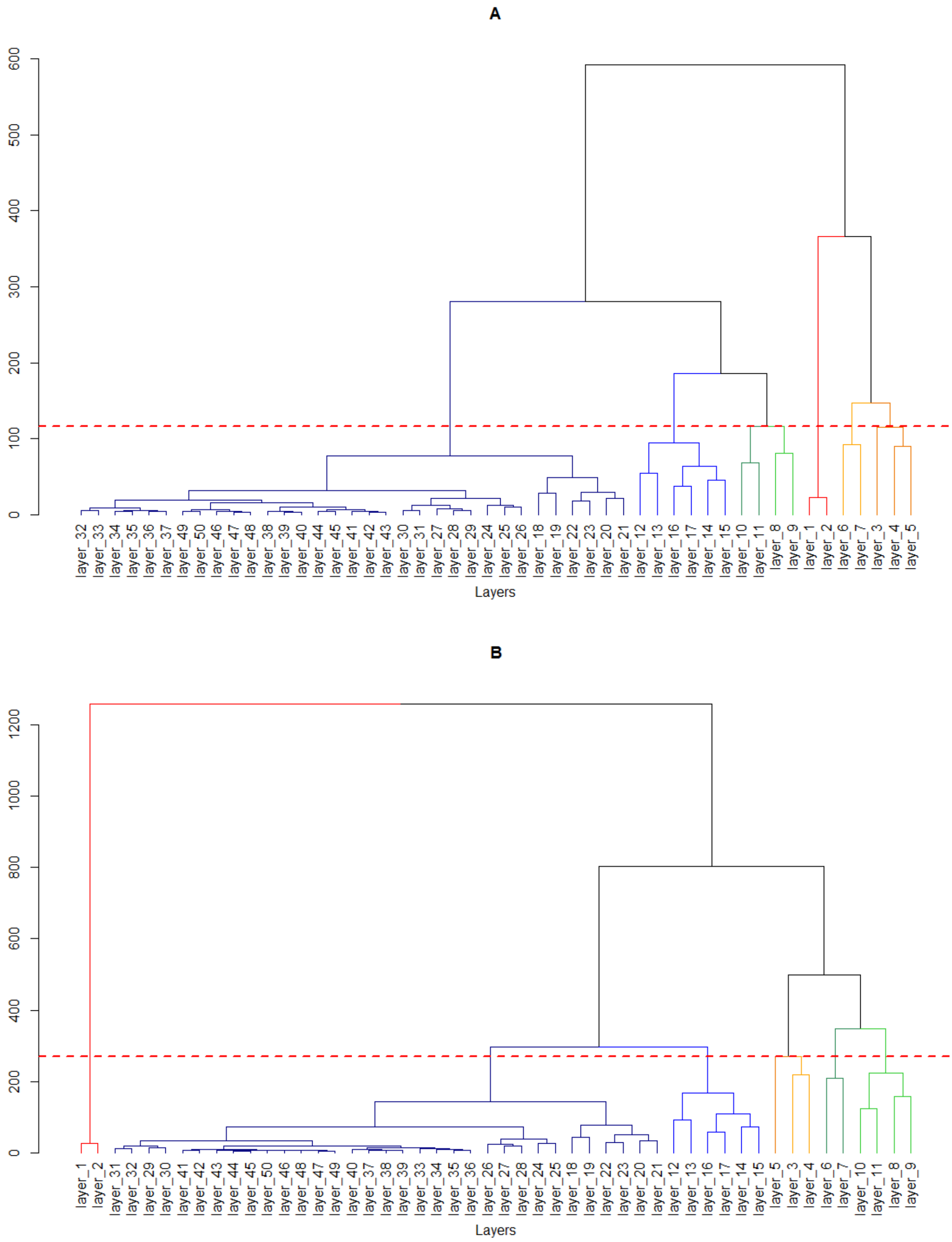

Figure 3: Dendrogram resulting from the cluster analysis of raw acoustic data for the Atlantic (A) and (B) Indian ocean. The red horizontal line indicates the height at which the dendrogram was sliced to create the 6 groups of layers. Colors identify the different groups of depth layers sampled by the buoy.

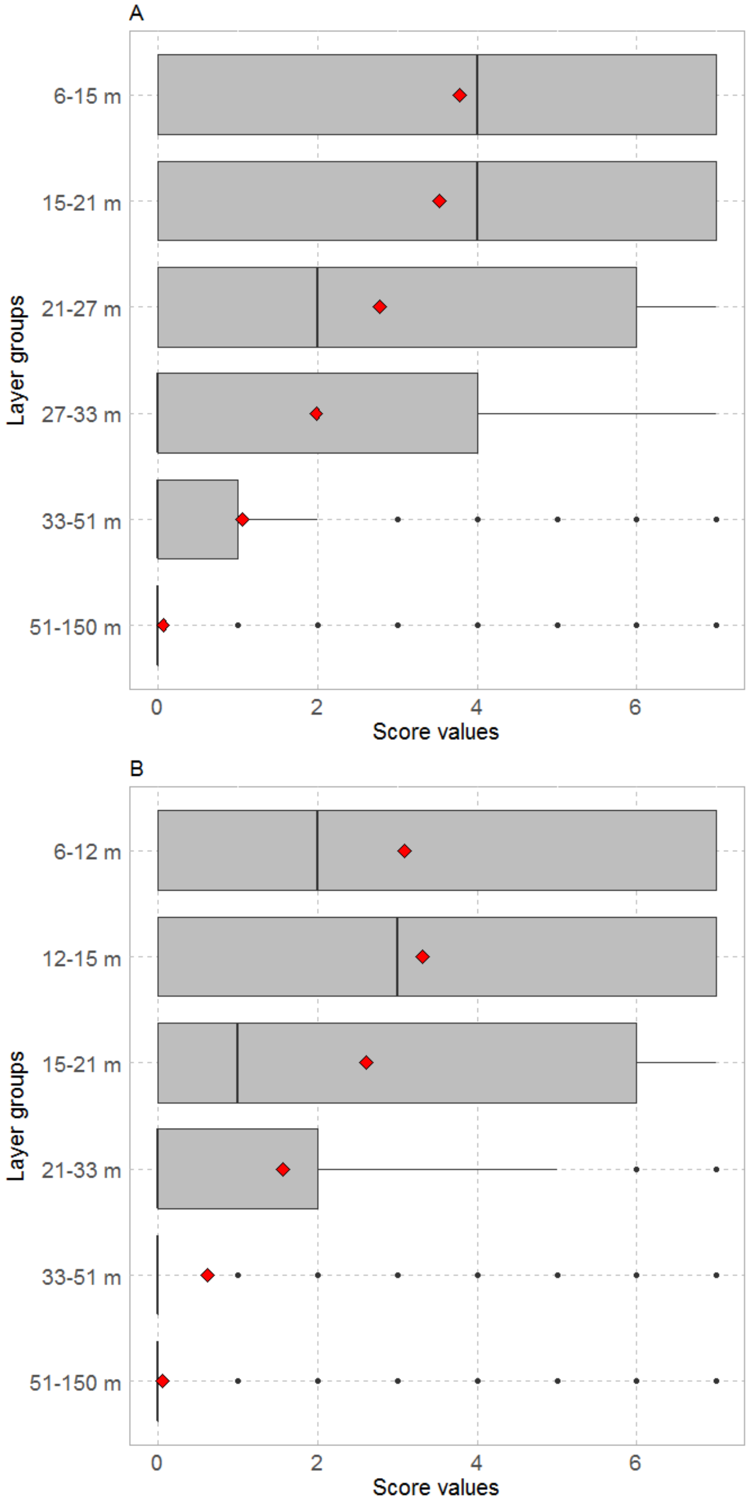

Figure 4: Boxplot of score values in the layer groups identified by the cluster analysis (with exclusion of transducer blanking zone 0-6m), for the (A) Atlantic, and (B) Indian oceans. Red squares represents mean value of scores in each layer group.

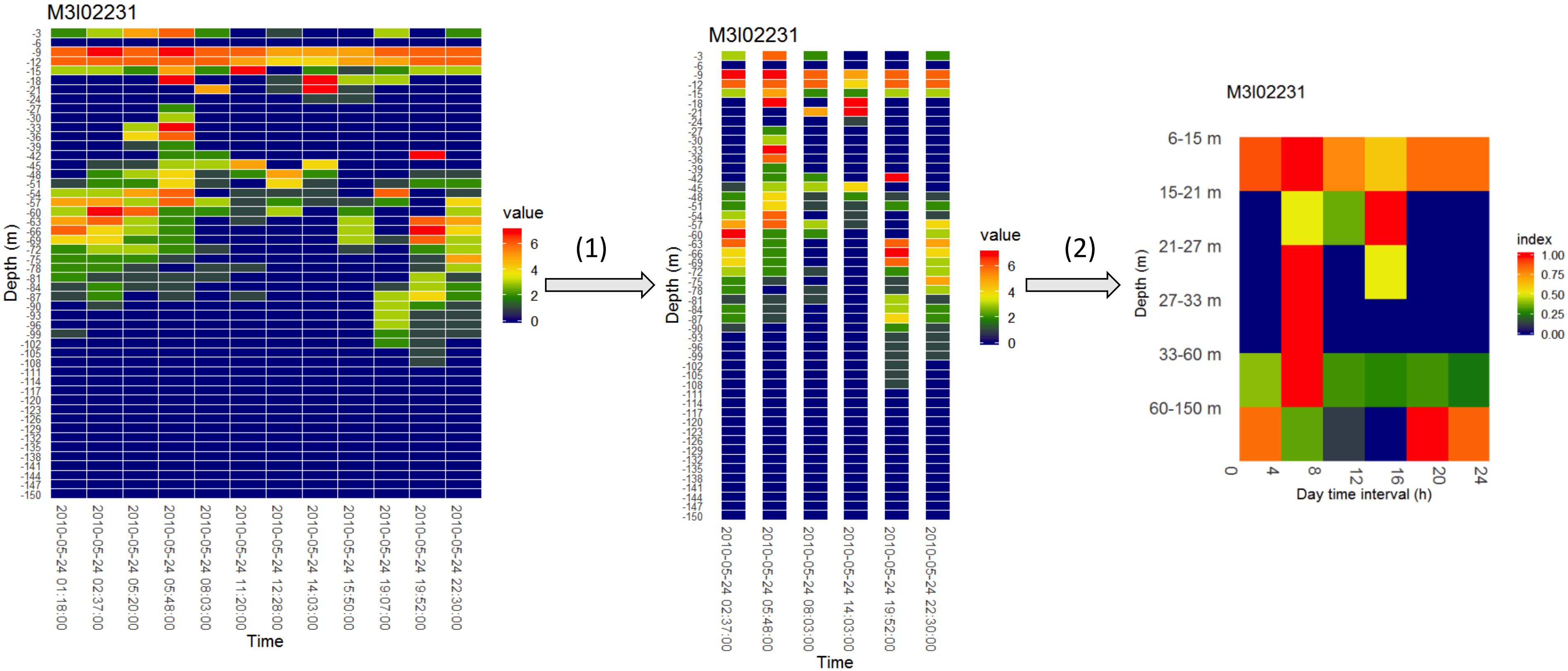

Figure 5: Acoustic data pre-processing. (1) Temporal resolution reduction, selecting the highest echo over a 4 hours period on the whole sampling day. (2) Layers aggregation gathering the 50 vertical layers into 6 layers based on cluster analysis. The final output is a 6x6 matrix (daily acoustic matrix) summarizing the aggregation on a full sampling day.

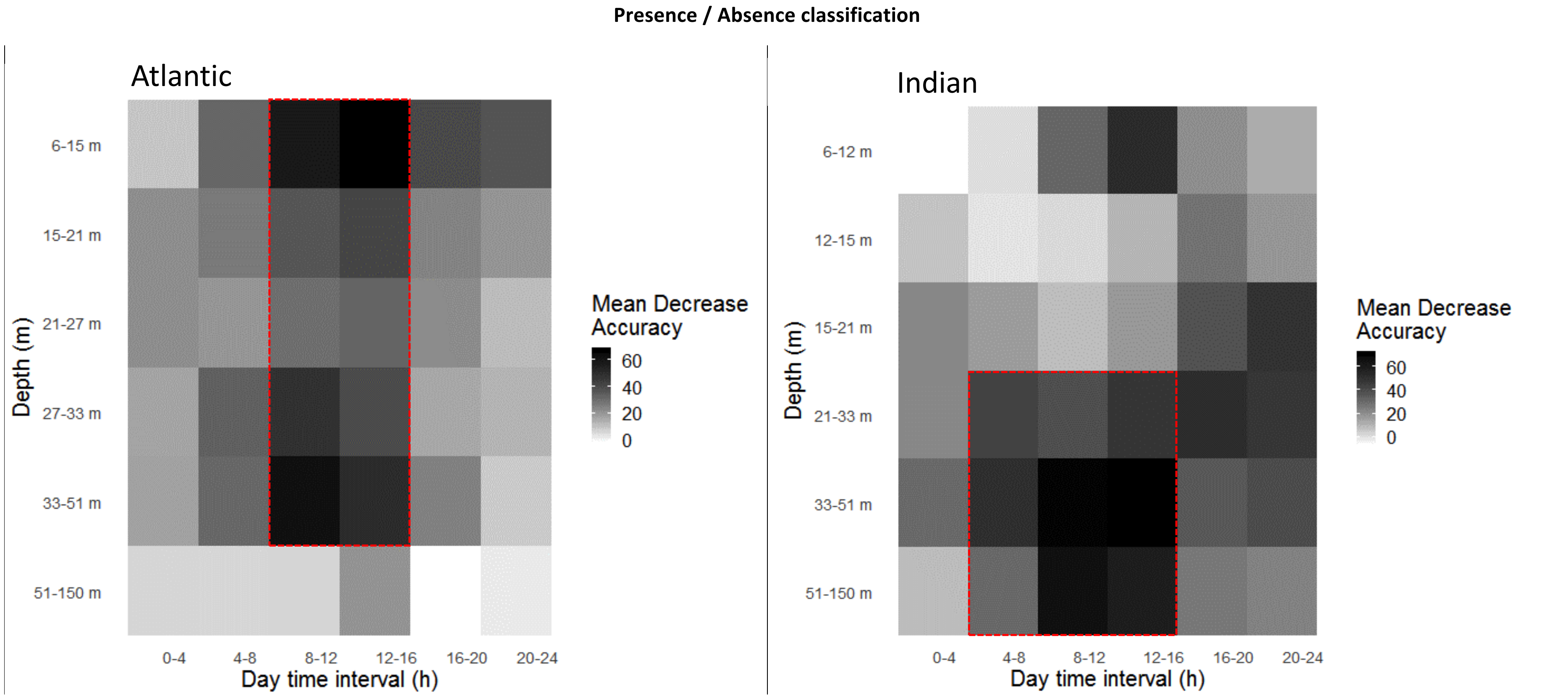

Figure 6 : Predictors importance from random forest algorithm for the Atlantic and Indian in presence/absence classification. Each cell represents a combination between a depth and a time period. Cells brightness indicates the relevance of the predictor in the classification. Red box indicates groups of relevant predictors in classification.

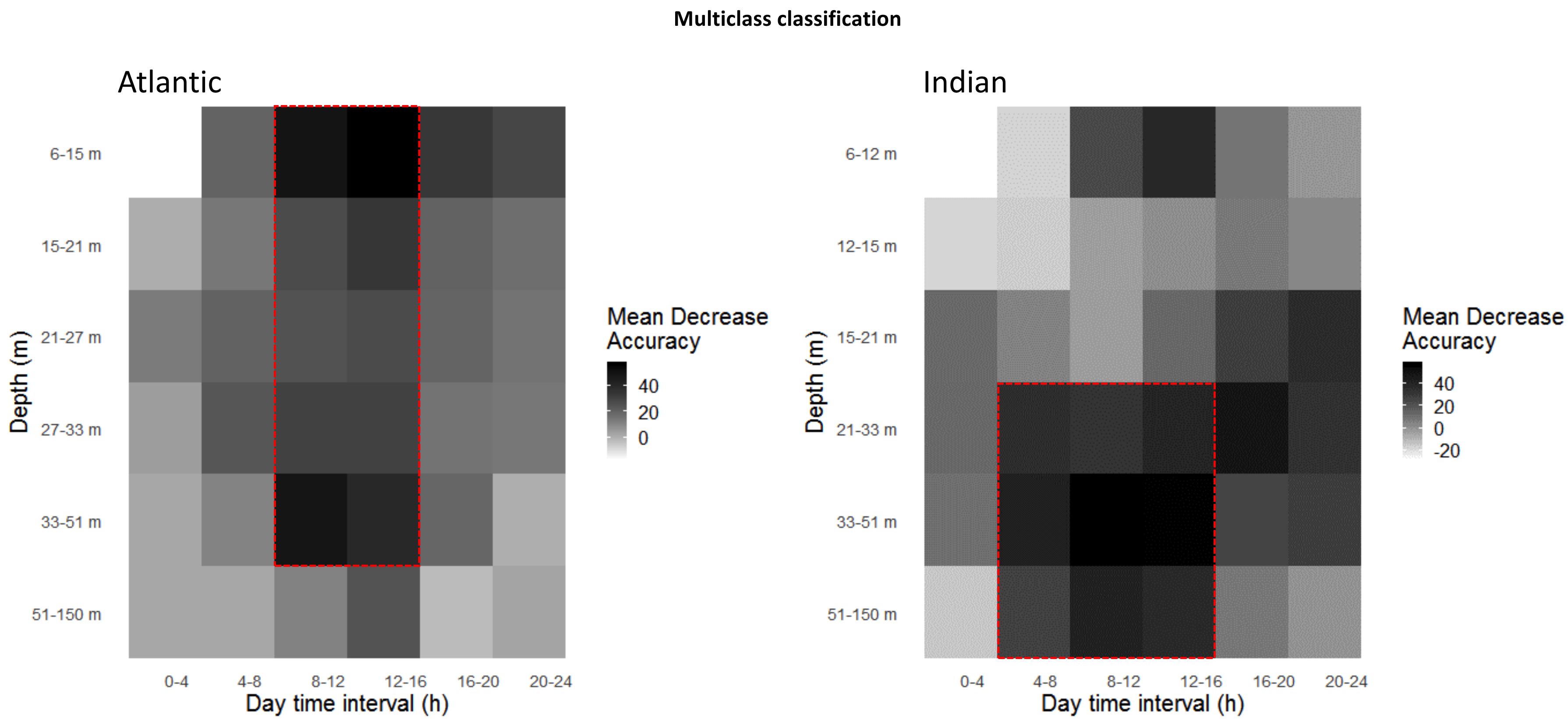

Figure 7: Predictors importance from random forest algorithm for the Atlantic and Indian in multiclass classification. Each cell represents a combination between a depth and a time period. Cells brightness indicates the relevance of the predictor in the classification. Red box indicates groups of relevant predictors in classification.