Received : 16 October 2017
IOTC-2017-WPTT19-22 Rev1

Towards the derivation of fisheries-independent abundance indices for tropical tunas: progress in the echosounders buoys data analysis

Y. BAIDAI [a,b], M. CAPELLO [a], J. AMANDÉ [b], N. BILLET[a], L. FLOCH [a], M. SIMIER [a], P. SABBAROS[a], L. DAGORN [a]

[a] Institut de Recherche pour le Développement (IRD), UMR MARBEC, B.P. 171, Av. Jean Monnet, 34203 Sète Cedex, France;
[b] Centre de Recherches Océanologiques (CRO), DRAV, 29, Rue des pêcheurs, BP V 18, Abidjan, Côte d'Ivoire

Abstract

Currently, the whole of the drifting FADs deployed by the purse seiners are equipped with echosounder buoys to remotely locate the FADs and to assess the amount of associated tuna. The acoustic signals provided by the buoys constitute an unprecedented, wide-scale database that can potentially provide real-time indicators on the dynamics of tuna populations, as well as local and regional abundance indices of tropical tunas for their stock assessments. We present the current progress in the treatment of the echosounder buoys database provided by the French fleet. We show results obtained through two novel algorithms developed in order to *(i)* filter-out erroneous and non-valid data from the echosounder buoys database (wrong positions, wrong biomass estimation, on-board positions); *(ii)* improve biomass estimates for tropical tuna species at the FAD level, by comparing their outputs with the data collected from onboard observers.

1. Introduction

The use of Fish Aggregating Devices (FAD) by commercial tuna fisheries has been growing continuously since the late 1980s (Fonteneau *et al.*, 2000). Currently, the whole of the FADs deployed by the purse seiners are instrumented with echosounder buoys that can remotely provide their GPS position and estimate the amount of tuna that is associated. On one side, the massive deployment of instrumented FADs alerts scientists and tuna RFMOs on their impacts on tuna populations and, more generally, on the pelagic ecosystem (Dagorn, *et al.*, 2013; Gregory, *et al.*, 2013; Hallier & Gaertner, 2008). . On the other side, the collection of such data across the tropical oceans offers an unprecedented opportunity for observing the pelagic realm, providing real-time data on the dynamics of tuna populations aggregating under floating objects (Moreno, *et al.*, 2016). So far, few pioneering works (Lopez, *et al.*, 2014 ; Lopez, et al., 2016) have been conducted in order to exploit the data produced by the buoys for scientific purposes.

This data are key for developing novel fishery-independent abundance indices for tropical tuna (Capello, *et al*., 2016) that can help the management of such fisheries.

The scientific exploitation of these data remains resolutely dependent on the essential steps of processing, cleaning, control and validation, since these data come from instruments designed exclusively for commercial purposes. This paper describes a processing approach of this database, allowing the production of scientifically exploitable data. It also presents preliminary results and methods for estimating fish biomass (and its uncertainty) from the comparison of catch data with the acoustic measures recorded by echosounder buoys.

2. Echosounder database

The echosounder database is currently made up of data recorded by buoys from the manufacturer "Marine Instruments", equipping the French fleet in the Indian and Atlantic oceans, during the period from 2010 to the present day. The data were made available as part of an ORTHONGEL / IRD confidentiality agreement concluded in May 2016.

The dataset corresponds to four different models (Table 1). Three of them, representing more than 97% of all the buoys in this database (and about 100% since 2015), are equipped with an echosounder device. In addition to the usual buoy GPS location, they provide acoustic data of different characteristics depending on the buoy model (Table 1).

Tableau 1 : Main technical specifications of marine instruments buoys

	MSI	M3I	M4I	M3I+
Occurrence in database	3%	77%	19%	1%
Satellite GPS :	Yes	Yes	Yes	Yes
Echo-sounder :	No	Yes	Yes	Yes
Operating frequency :	-	50 kHz	50, 120, 200 KHz	50 & 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	each minute
TVG correction :	-	No	No	Yes

Buoys have different modes of acquisition and transmission of data, all based on a similar principle. In the default mode (standby mode), acoustic pings are continuously made by the buoy at constant time intervals (every 1 or 5 minutes depending on buoy model). The best echo, corresponding to the highest acoustic energy measured over a period of 2 hours is saved and exported to servers during satellite communications (every 3 or 12 hours depending on the buoy model). At each satellite communication, location data are also acquired. Thus, per day, 6 (or more) soundings data are sent back to servers. In addition, buoys also provide various information ranging from water surface temperature to instantaneous speed and turning angle, as well as the voltage level of its battery.

3. Database processing

The processing of the raw data, focused on five main steps:

(i). Filtering data recorded on buoy with low voltage conditions;
(ii). Filtering erroneous location data, related to failures in satellite communication and location data acquisition;
(iii). Identifying buoys on land positions;
(iv). Identifying buoys with bathymetry lower than 150 meters (detection threshold of the buoy echosounder device);
(v). Identifying buoys data recording on-board positions.

Furthermore, the database processing was governed by two constraints. The first one is the need of modularity between all data filtering processes so that different filters can be activated depending on the user purposes (ex: for buoys trajectory analyses there is no need to account for bathymetry, etc.). The second is an algorithmic performance sufficiently high to allow for the management of ever-increasing amount of data over the years (Figure 1).

3.1. Low voltage filtering

According to the buoys manufacturer, data provided under a voltage of 11.5 V, present poor reliability (in term of location and acoustic measures). Thus, these data (less than 1% of whole dataset) have been removed from the database (Figure 2).

Received : 16 October 2017
IOTC-2017-WPTT19-22 Rev1

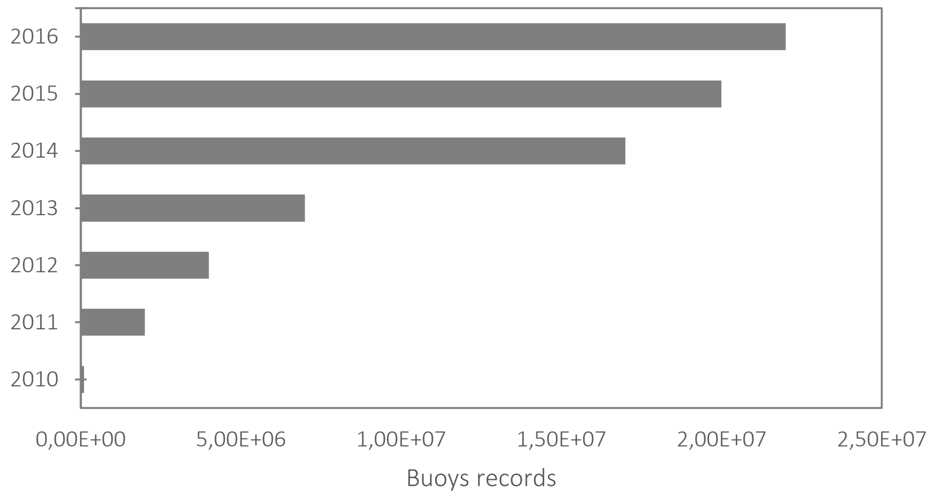

Figure 1 : Marine Instruments buoys records over years

3.2. Erroneous location filtering

Erroneous location data mainly resulting from failures in satellite communication and GPS data acquisition were also retrieved from the database (Figure 2). These data can be grouped into three categories:

(i). Isolated GPS data (data with a gap higher than 48h from other location records);

(ii). Data with multiple positions for the same timestamp and data with aberrant buoy speed (greater than 35 knots);

(iii). Single location data: which designates a buoy which occurs only once in the database (resulting from buoys transmission tests).

Globally, the erroneous location data corresponded to less than 1 % of the whole database.

3.3. Land positions and bathymetry filtering

Buoys with continental positions was flagged, using the NOAA - Global Self-consistent, Hierarchical, High-resolution Shoreline (Wessel & Smith, 1996), considering a 5 kilometers buffer around the shoreline. Echosounder data recorded at depth of less than 150 meters were also reported in the database, using the ETOPO1 global relief model of Earth's surface; (ETOPO1; Amante & Eakins, 2009) with a resolution of 1 arc-minute. Continental positions and flagged bathymetric data represented together around 6 % of the whole dataset (Figure 2).

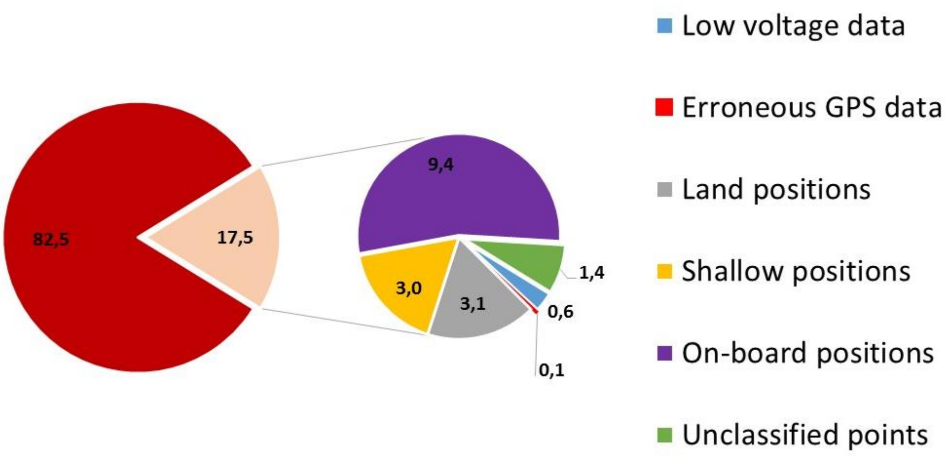

Figure 2 : Proportion of processed data

3.4. On-board positions filtering

We developed a classification algorithm for discriminating buoys emitting on board the fishing vessels from those located at sea. One important requirement for our algorithm was that it could deal with very large (and still growing) datasets. As such, rather than using the same classification methods as in (Maufroy, *et al*., 2015), this algorithm was based exclusively on the analysis of the buoy speed and its variations along the buoy trajectory.

For each buoy, the first step of the algorithm was to associate to each recorded position the corresponding value of mean speed (from the distance between two subsequent GPS positions) and mean acceleration (estimated from the difference in speed between two consecutive points of the trajectory).

Then, the algorithm relied on two main classification rules:

(i) Positions with associated mean speed greater than 6 knots were considered "on-board".

(ii) Positions with associated mean speed lower than 6 knots during a continuous period of 3 days (before the position to be assigned) were considered as emitting "at-sea".

The above rules led to the initial classification of a sub-portion of the buoy trajectory. Over this sub-portion, we identified the sequences of the buoy trajectory with no state variation (i.e. portions of the trajectory presenting "on-board – on-board" or "at-sea – at-sea" states only, herein called *constant sequences*) and the sections of buoy trajectory where the buoy state changed from one to another state (i.e. "on-board" – "at-sea" or vice versa, herein called *transition sequences*). The mean accelerations estimated over constant sequences turned out to be much smaller than those estimated for constant sequences (Figure 3).

Finally, the points of the trajectory that were still unclassified (positions having speed lower than 6 knots and not fulfilling rule (ii)) were assigned to the “on-board” or “at-sea” class by comparing their mean acceleration with the distributions of mean accelerations estimated for constant and transition sequences (using the t-test of comparison at confidence level of 0.95). Such classification was conducted along the full buoy trajectory in both directions.

The "on-board" data represent about 10% of the total volume of data processed (Figure 2). Nevertheless less than 2% of the points could not be classified by this algorithm (mainly resulting from trajectories not having enough data).

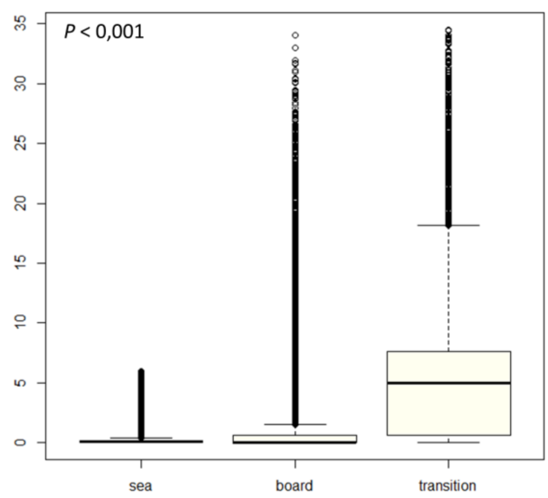

Figure 3: Boxplots of the mean acceleration (absolute values) for constant (sea and board) and transition states with p-value at ANOVA test (p<0.001)

The quality of this classification algorithm was evaluated by comparison with a classification made exclusively from observer data. Indeed, onboard observers perform buoy data collection (model and serial number of the buoy associated with the FAD) at each operation on a FAD (visit, retrieving or deployment of FADs by the vessel). Based on these data, we could classify “on-board” and “at-sea” states for a subset of the buoys trajectories that were present within our database (around 27000 positions) (Figure 4). Comparisons of the two classifications through calculation of proportion of concordant pairs through simple matching coefficient

Received : 16 October 2017
IOTC-2017-WPTT19-22 Rev1

(Sokal & Michener, 1958) revealed a good similarity between them (more than 90% of similarities between observer-based and speed-based approaches).

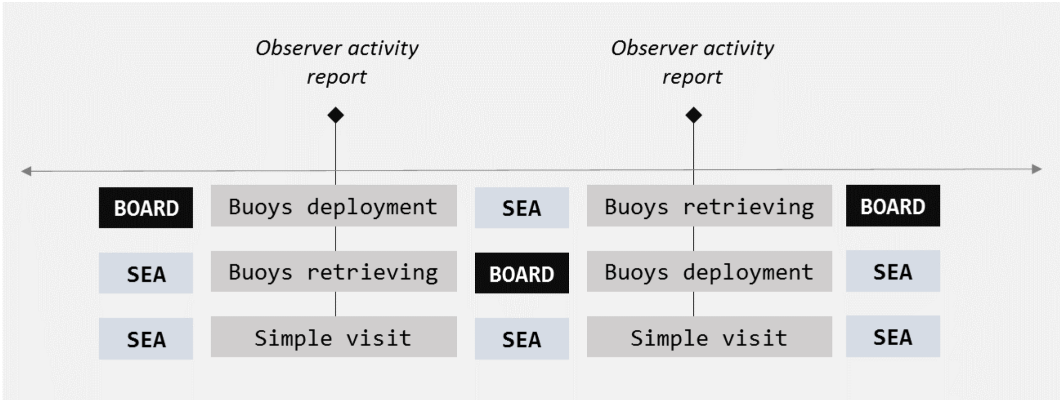

Figure 4 : Buoy states classification based on observer data

4. Biomass estimation through echosounder buoys data at the FAD level

4.1. Echosounder buoy data

For each frequency, the echosounder buoys data appear as discrete indices varying between 0 to 7 (0 to 15 for the most recent M3+ buoy model), reflecting the strength of the signals received per layer of 3 meters (Figure 5). Only the values of these indices are made available, which makes it difficult to use conventional echo-integration methods to transform the signal into biomass. Additionally, the buoys provide a biomass estimate calculated through an algorithm provided by the constructor.

4.2. Observers 'data

Catches data were collected by observers from 2014 to 2015, over the course of 410 trips on French purse seiners (185 and 225 respectively in Indian and Atlantic oceans), corresponding to 4582 fishing sets on floating objects (53 % and 47% respectively in Atlantic and Indian oceans).

4.3. Biomass estimates

The comparison between the biomass estimate provided by the constructor and the catches reported by the observers is shown in Figure 6. In order to improve these estimates, we have employed an empirical approach that combines the acoustic data recorded at each layer with the catch data. Our approach expresses the biomass using the following equation:

Received : 16 October 2017
IOTC-2017-WPTT19-22 Rev1

$$B = \sum_{i=1}^{n} \beta_i f(X_i, W)$$

Where B represents the catch (biomass), i denotes the layer, n is the number of depth layers, X_i is the value of the acoustic index at layer i and β_i denotes the (unknown) coefficients for each depth layer. The function $f(X_i, W)$ denotes either the mean, or the maximum value if the index X_i estimated over a time window W before the set and the β_i coefficients are optimized through a least squares fit. Figures 7 and 8 show the comparison between the catch and the estimated biomass obtained from this approach, considering $f(X_i, W)$=max (X_i, W) and $f(X_i, W)$=mean (X_i, W), respectively, and a time window W=48h.

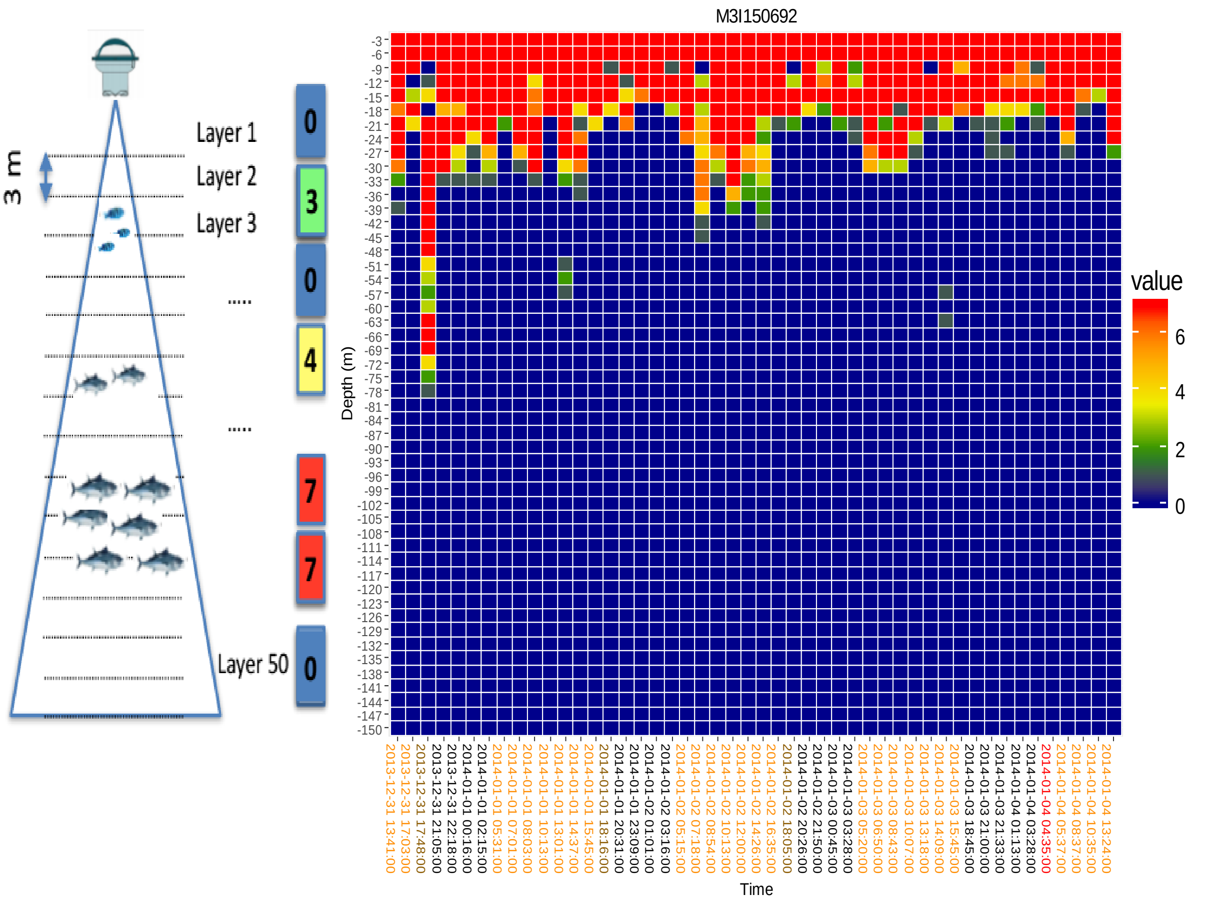

Figure 5: View of acoustic data recorded by "Marine Instruments" echosounder buoys (example of echosounder data over one day for a single buoy)

Received : 16 October 2017
IOTC-2017-WPTT19-22 Rev1

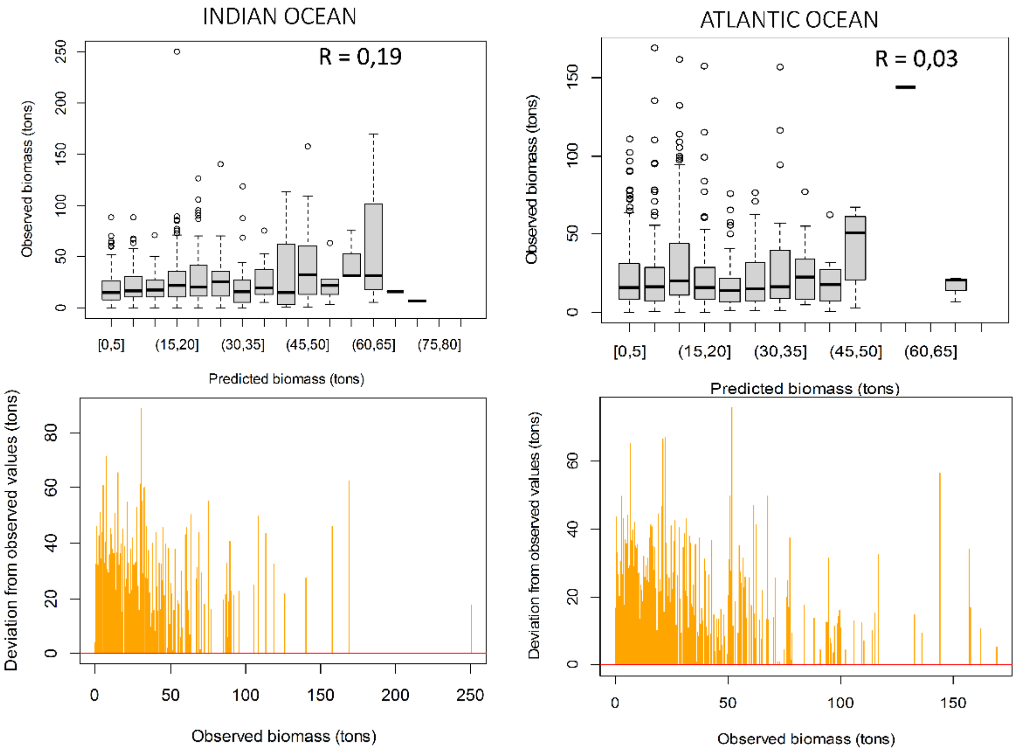

Figure 6: Biomass estimates in metric tons (numbers in black) for total catch, based on manufacturer's algorithms (maximum of buoy estimation over 48h) in Indian and Atlantic Oceans (N= 832 and 953 sets respectively for Indian and Atlantic; R=Pearson correlation coefficient).

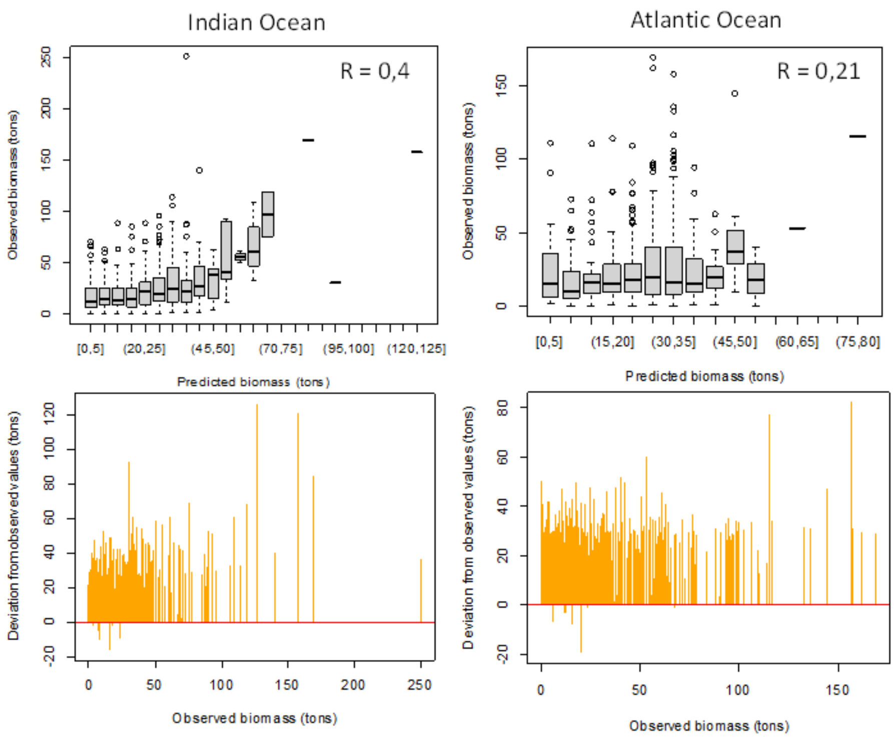

Figure 7: Biomass estimates in metric tons for total catch, through least squares optimization from the maximum of the echosounder data recorded 48 h before the fishing set. (N= 832 and 953 sets respectively for Indian and Atlantic; R=Pearson correlation coefficient)

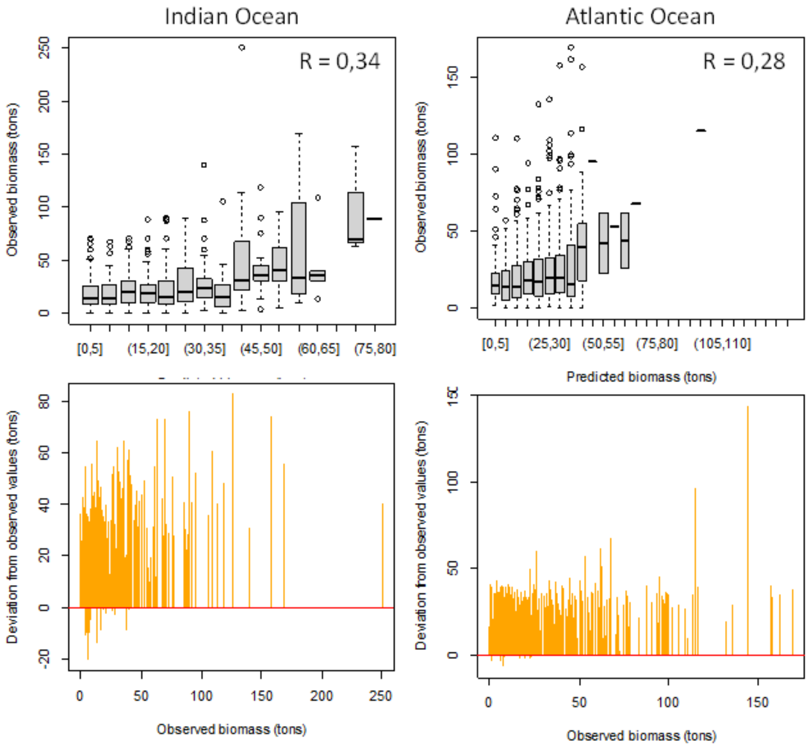

Figure 8 : Biomass estimates in metric tons for total catch, through least squares optimization from the average of the echosounder data recorded 48 h before the fishing set (N= 832 and 953 sets respectively for Indian and Atlantic; R=Pearson correlation coefficient).

5. Discussion

Developing a relevant indicator for the aggregated biomass under FADs from echosounder buoy data faces many challenges. The first is related to the format of the acoustic data recorded by the buoys that appears as discrete indices and thus cannot be treated using the standard acoustic data analysis techniques. Another difficulty relies on the limited cone of the buoy that can only sample a subportion of the aggregation. Moreover, the daily, horizontal and vertical migrations of tunas (Doray et *al.*, 2006 ; Josse & Bertrand, 2000) can lead to a high variability in the acoustic responses obtained.

The derivation of an appropriate method that best estimates the aggregation below the FAD from the acoustic data provided by the buoy is key. Also, it is very important to quantify which is the uncertainty for a biomass estimate and how this uncertainty depends on the type of the buoy, the buoy location and the species forming the aggregation. Previous works considered a different buoy brand and cannot be applied to this database. In this respect, this work constitutes the first step of such assessment for the buoy brand used by the French fleet. Thanks to the availability of both echosounder buoy data and catch data we could demonstrate that the correlation between the estimated biomass provided by the buoy and the observed catch is low (Figure 6). The least-squares algorithm that we implemented provides better correlations (Figure 7 and 8). However, we still need to refine our approach considering how far extrinsic covariables (such as the location of the buoy, the time window before the set, the environmental variables) can play a role in the large variability in the accuracy of the biomass estimates that is observed. Moreover, we aim at including also complementary data such as the visits without a catch, since this would increase our database for those buoys that presented little biomass (and no catch) and thus allow a better assessment of the fit parameters.

Acknowledgements

The authors sincerely thank the contribution of the staff of the “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

Amante, C., & Eakins, B. (2009). *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis.* NOAA, National Geophysical Data Center. NOAA Technical Memorandum NESDIS NGDC-24. doi:10.7289/V5C8276M

Capello, M., Deneubourg, J., Robert, M., Holland, K., Schaefer, K., & Dagorn, L. (2016). Population assessment of tropical tuna based on their associative behavior around floating objects. *Scientific report, 6, 36415*. doi:10.1038/srep36415

Dagorn, L., Holland, K., Restrepo, V., & Moreno, G. (2013). Is it good or bad to fish with FADs? What are the real impacts of the use of drifting FADs on pelagic marine ecosystems? *Fish and Fisheries, 14*, 391-415. doi:10.1111/j.1467-2979.2012.00478.x

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

Fonteneau A., Pallares P., & Pianet P. (2000). A worldwide review of purse seine fisheries on FADs. *Le Gall J.-Y., Cayre P. and Taquet M. (eds), Peche thoniere et dispositifs de concentration de poisons, 28(15-34)*. Ed. Ifremer, Actes Colloq.

Gregory, S., Dagorn, L., Robert, M., & Deneubourg, J. (2013). Impact of increasing deployment of artificial floating objects on the spatial distribution of social fish species. *Journal of Applied Ecology, 50*, 1081-1092.

Hallier, J., & Gaertner, D. (2008, January 17). Drifting fish aggregation devices could act as an ecological trap for tropical tuna species. *Marine Ecology Progress Series, 353*, 255-264.

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

Lopez, J., Moreno, G., Boyra, G., & Dagorn, L. (2016). A model based on data from echosounder buoys to estimate biomass of fish species associated with fish aggregating devices. *Fishery Bulletin, 114*, 166-178.

Lopez, J., Moreno, G., Sancristobal, I., & Murua, J. (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*(0), 127-137.

Maufroy, A., Joo, R., Chassot, E., & Kaplan, D. (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), e0128023. doi:10.1371/journal.pone.0128023

Moreno, G., Dagorn, L., Capello, M., Lopez, J., Filmater, J., Forget, F., . . . Holland, K. (2016, June). Fish aggregating devices (FADs) as scientific platforms. *178*, 122-129. Retrieved from https://doi.org/10.1016/j.fishres.2015.09.021

Sokal, R., & Michener, C. (1957). A Quantitative Approach to a Problem in Classification. (S. f. Evolution, Ed.) *Evolution, 11*(2), 130-162. doi:10.2307/2406046

Sokal, R., & Michener, C. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin, 39*, 1409-1438.

Wessel, P., & Smith, W. (1996). A Global, Self-consistent, Hierarchical, High-resolution Shoreline Database. *J. GeoPhys. Res.101(B4: 8741-8743., 101*(B4), 8741-8743.