AGGREGATION TIMES OF TUNA SCHOOLS TO FADS ESTIMATED BY ECHOSOUNDER DATA

Manuel Navarro-García[1,2], Daniel Precioso[3], Kathryn Gavira-O'Neill[4], Alberto Torres-Barrán[2], David Gordo[2], Víctor Gallego-Alcalá[2], and David Gómez-Ullate[*3,5]

[1]Universidad Carlos III de Madrid, Spain
[2]Komorebi AI Technologies, Madrid, Spain
[3]Department of Computer Science, Higher School of Engineering, Universidad de Cádiz, Spain
[4]Satlink S.L., Madrid, Spain
[5]On leave of absence from Department of Theoretical Physics, Universidad Complutense de Madrid, Spain.

September 20, 2021

ABSTRACT

We perform a systematic study of aggregation and disaggregation times of tuna schools to drifting Fish Aggregating Devices (dFADs), using the signal provided by the echo-sounder buoys attached to dFADs deployed across all major oceans in the period 2018-2020. The tuna biomass estimation for each day in the time series has been obtained by applying the TUN-AI Machine Learning model (Precioso et al., 2021), which incorporates oceanographic information and hourly echo-sounder data in 10 depth layers on a time window of 72 hours prior to the prediction. We preprocess the data collected from the buoys to select around $10\,000$ series with daily estimations where no human intervention has occurred. A statistical analysis of these time series with different smoothing techniques shows that tuna schools remain aggregated to dFADs for a median time of 3-9 days, and that the aggregation and disaggregation processes are symmetrical.

***K*eywords** Tropical tunas · Echo-sounder buoys · Fish aggregating device · Associative behaviour · Ecological trap

1 Introduction

Floating objects drifting on the ocean's surface have been known to attract a number of fish species for centuries, including tropical tunas such as skipjack tuna (*Katsuwonus pelamis*), yellowfin tuna (*Thunnus albacares*) and bigeye tuna (*Thunnus obesus*) (Castro et al., 2002; Maufroy et al., 2015). As fishermen have noticed this behavior, they have used both natural and man-made floating objects, or drifting Fish Aggregating Devices (dFADs), as a tool for finding and catching tropical tunas. The use of dFADs in tuna purse-seine fisheries has gradually increased since the 1980s to nowadays, where vessels using dFADs now contribute to 36% of the world's total tropical tuna catch (Davies et al., 2014; Wain et al., 2021; ISSF, 2021). These wide-spread changes in tuna purse-seine fisheries have highlighted the need to better understand the potential ecological effects of dFADs on tuna ecology and the marine environment, in order to ensure adequate management of fish stocks and dFAD usage.

The reasons driving tuna's tendency to aggregate around dFADs are poorly understood. A number of hypotheses have been suggested, but none have been universally accepted by the scientific community (for a comprehensive review, see Fréon and Dagorn (2000); Dempster and Taquet (2004); Castro et al. (2002)). However, two hypotheses explaining tuna aggregation to naturally floating objects and dFADs have gained traction in recent years: "indicator-log" hypothesis (Hall et al., 1992) and the "meeting-point" hypothesis (Fréon and Dagorn, 2000). The "meeting-point" hypothesis considers that dFADs facilitate the encounter between individuals or schools, thus constituting larger schools that could benefit survival rates (Castro et al., 2002). On the other hand, the "indicator log" hypothesis suggests that some fish

*Corresponding author: david.gomezullate@uca.es

displaying this behavior, including tunas, may be safeguarding the survival of their eggs, larvae and juvenile stages by using drifting objects as indicators of convergence zones, thus obtaining short-term trophic benefits by helping them to remain in areas where plankton and food is readily available.

Based on the "indicator-log" hypothesis, some authors have ventured that the massive use of dFADs in tuna fisheries, especially in areas where natural floating objects are not usually present, could have detrimental effects on tuna populations by creating a so-called "ecological trap" (Marsac et al., 2000; Hallier and Gaertner, 2008). This hypothesis, postulates that tuna's attraction to dFADs could lead them to remain associated to dFADs even as these drift into areas that could negatively affect the tuna's behavior and biology (Marsac et al., 2000; Hallier and Gaertner, 2008). However, after reviewing the scientific literature available on the "ecological trap" hypothesis in the context of dFADs and tuna, Dagorn et al. (2012) state that no conclusion can be made as to whether or not the "ecological trap" is taking place. Small scale studies using acoustic or electronic tags to monitor tuna movements around floating objects (both anchored and drifting) have shown that the time tunas remain associated, though variable, is generally less than 10 days, and tunas alternate between associated and non-associated phases (Ohta and Kakuma, 2004; Schaefer and Fuller, 2005; Dagorn et al., 2007; Schaefer and Fuller, 2010; Robert et al., 2014; Govinden et al., 2021), contrasting with the perception that the association of tuna to floating objects is "fast, strong and long-lasting" as the "ecological trap" hypothesis would imply (Marsac et al., 2000).

Today, the dFADs used by tropical tuna purse-seine fisheries are generally deployed with satellite-linked instrumented buoys equipped with one or more echo-sounders, which provide fishermen with dFAD geolocation as well as estimates of aggregated tuna biomass (Davies et al., 2014; Wain et al., 2021). Given the number and large-scale distribution of dFADs in recent years, the information provided by these buoys has caught the attention of several authors, who highlight their potential to provide insights in tuna migration and behaviour on a global scale (Santiago et al., 2016; Baidai et al., 2020a; Orue et al., 2019a; Lopez et al., 2016). Recent studies have begun to model and process the echo-sounder data provided by these buoys as a means to remotely assess tuna presence or absence, or investigate patterns in tuna aggregation around dFADs (Baidai et al., 2020b; Precioso et al., 2021; Orue et al., 2019b).

In this context, the current study aims to complement the small scale studies carried out on tuna aggregation (for example, Schaefer and Fuller (2010); Govinden et al. (2021)) by using one such model (Precioso et al., 2021) and a historic database of biomass estimates provided by echo-sounder buoys of the AGAC[2] tuna purse seine fleet, deployed across all major oceans. Through our analysis, we will assess temporal patterns in aggregation and disaggregation, and time spent by tunas under dFADs. We will also examine the nature of these aggregations and disaggregation, to evaluate whether these processes occur gradually or suddenly.

The rest of this paper is organized as follows. Section 2 outlines the sources of information and the data preprocessing performed and describes the algorithm followed to estimate time spent by school tunas associated to dFADs. Subsequently, Section 3 reports the resulting time distributions for the raw and smoothed biomass estimation series. Finally, it concludes with some final remarks in Section 4.

2 Material and methods

2.1 Database description

Our study draws from three sources of information: activity data on FADs, echo-sounder buoy data, and oceanography data.

2.1.1 FAD logbook data

The first database is the FAD logbooks of the Spanish tropical tuna purse seine fleet operating in the Atlantic, Indian and Pacific Oceans (2018 - 2020, AGAC ship owner's association data). These logbooks contain almost 66 000 interactions between fishing boats and Satlink echo-sounder buoys. Every interaction within the dataset is linked to a specific buoy, via the ID and model of the buoy attached to each dFAD. Each record contains information about the date, time and GPS coordinates where the interaction occurred, as well as the kind of interaction. These interactions are listed and described in Ramos et al. (2017). For this study we will only consider disruptive interactions, understood as those that affect the echo-sound recording process or the fishes' behaviour. Disruptive events are: deployments, sets, retrievals at sea or at port, and dFAD losses. Interactions of non-disruptive nature, such as visits and modifications over previous objects, are ignored for the purposes of this work.

[2]Asociación de Grandes Atuneros Congeladores

2.1.2 Echo-sounder buoy data

The echo-sounder buoy data was collected from $15\,497$ Satlink buoys[3] for which there were registered interactions in the FAD logbook data (see Section 2.1.1). There were over 68 million records contained in this database, reported by buoys attached to dFADs scattered over the Atlantic, Indian and Pacific Oceans. Each record corresponds to a specific buoy ID and timestamp, ranging from 2018 to 2020, and contains biomass estimates and GPS coordinates of the dFAD's last known position at the time of measuring. This study used three buoy models (ISL+, SLX+ and ISD+), the details of which are described in Table 1. The observation range of the echo-sounder is from 3m to 115m depth, split into ten layers, each with a resolution of 11.2m (see Figure 1). Biomass estimates (in metric tons, t) are obtained from acoustic samples taken periodically throughout the day (see Table 1). The average back-scattered acoustic response is converted into estimated tonnage, based on the target strength of Skipjack tuna (*Katsuwonus pelamis*). Out of all the records registered within an hour, only the one with the highest estimated tonnage is sent to the satellite, reducing the amount of data transmitted. Thus, the final temporal resolution of echo-sounder records for each buoy is 1h. The buoys have an internal detection threshold of $1\mathrm{t}$, which means that a biomass estimation below $1\mathrm{t}$ is not transmitted, and interpreted as a zero-reading.

Model	**Echo-sounder**	**Freq. (kHz)**	**Beam Angle (°)**	**Measuring rate**
ISL+	ES12	190.5	20	Every 15 minutes
SLx+	ES16	200	23	From sunrise to sunset: every 5 minutes From sunset to sunrise: every 60 minutes
ISD+	ES16x2	200 and 38[a]	23 and 33	From sunrise to sunset: every 5 minutes From sunset to sunrise: every 60 minutes

[a] Biomass estimates are calculated according to the acoustic response registered by the 200kHz echo-sounder, allowing data to be comparable across all buoy models.

Table 1: Buoy models and characteristics

2.1.3 Oceanography data

Oceanographic data was downloaded from the global ocean model (products GLOBAL-ANALYSIS-FORECAST-PHY-001-024, $1/12°$ resolution; and GLOBAL-ANALYSIS-FORECAST-BIO-001-028, $1/4°$ resolution) provided by the EU Copernicus Marine Environment Monitoring Service[4] (Global Monitoring and Forecasting Center, 2018). For each record contained in the echo-sounder buoy data (see Section 2.1.2), we downloaded several oceanographic variables at surface level (depth = 0.494 m). This data set will be used by the model from Section 2.2.1 to estimate the size of the tuna school aggregated under each dFAD.

2.2 Data preprocessing

2.2.1 TUN-AI estimations

We will use the estimations of tuna schools from the TUN-AI model (Precioso et al., 2021). This model is based on a Gradient Boosting (GB) algorithm (Friedman, 2001). TUN-AI uses information from the acoustic records, buoy location, and oceanography, to estimate the tuna biomass below the dFAD, with an average error of ± 8.54 tons. The model is trained using set and deployment events from the FAD logbook. Set events include the amount of tuna caught, which was assumed to be the whole biomass under the dFAD. On the other hand, for deployment events it is assumed that no tuna had yet associated to the dFAD. For a detailed explanation of TUN-AI, please refer to Precioso et al. (2021).

To predict the size of the tuna school for a given day, TUN-AI requires an input echo-sounder window of the previous 3 days, containing one acoustic record per hour. By decreasing the frequency of the biomass estimations from hourly to daily, the day-night oscillations on the biomass are avoided, thus making the results more robust. Also, TUN-AI has proven to be more accurate when estimating the tuna under the dFAD than simply aggregating the raw acoustic

[3] SATLINK, Madrid, Spain, www.satlink.es

[4] http://marine.copernicus.eu/

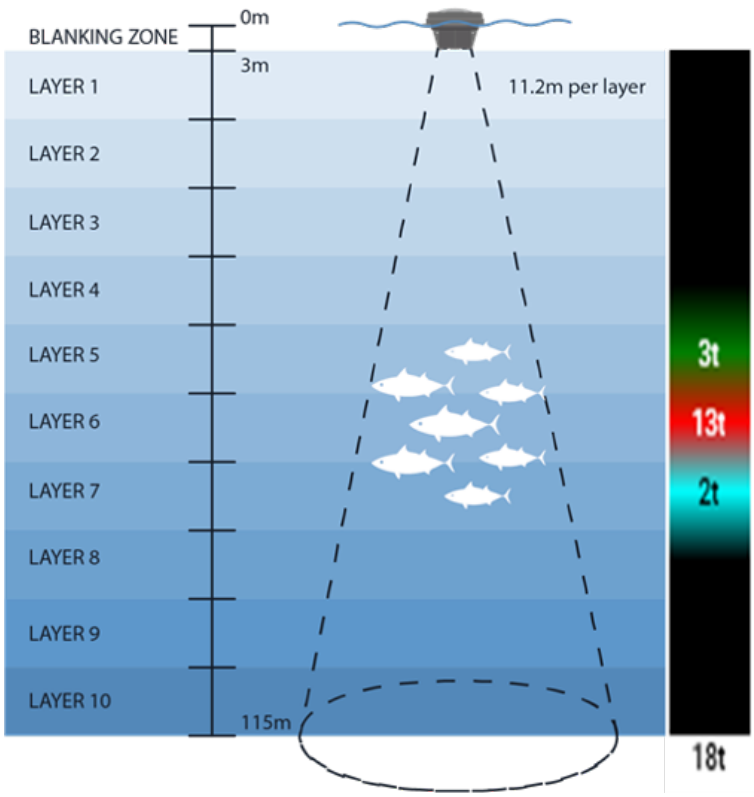


Figure 1: Left: Depth layer configuration and set-up of the Satlink echo-sounder buoys. Right: example of the biomass estimates (in metric tons) and echogram display available to buoy users. Raw acoustic backscatter is converted into biomass estimates based on the target strength of skipjack tuna *(Katsuwonus pelamis)* using manufacturer's algorithms.

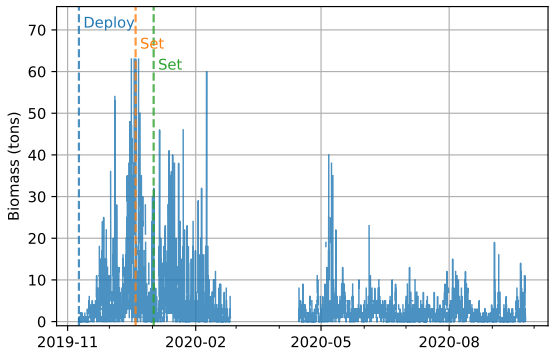
signal (Precioso et al., 2021). We applied TUN-AI to each dFAD in our dataset, sliding the echo-sounder window sequentially to obtain one biomass estimation per day for each buoy in the dataset.

2.2.2 Generating splits

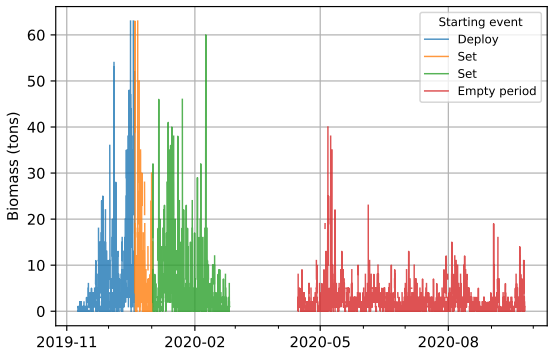
Many buoys from the echo-sounder data set are involved in multiple events along the historical series of acoustic data, which may disturb the biomass under the buoys. Moreover, it sometimes happens that a buoy does not emit any signal for several days in a row, so it is uncertain to determine the aggregation time in these situation. Therefore, to properly study the aggregation time of the tunas to the dFADs, it is necessary to divide the complete series into smaller partitions where echo-sounder data is available for the most part of the interval and no human activity was performed during this range. Only in this situation where tuna schools receive no external interaction makes sense to estimate the aggregation times.

As we have previously stated, generating the splits from historical series of a buoy requires both the echo-sounder and the FAD logbook data sets, since the logbook contains the disruptive events (human interactions with the dFAD) that can be used to divide the series. We first merge both data sets using the date, the buoy ID and the event that took place as columns to join on. Figure 2a shows the complete historical acoustic series of a buoy together with human interactions, depicted in coloured dashed lines. There are ten values of acoustic records per hour (one for each layer depth); but to simplify the visualization of Figures 2a to 2c we only plot the maximum value per hour. In this paradigmatic example, it is shown that long time intervals without any acoustic signal recorded may happen as we pointed out before, so special treatment needs to be developed for this casuistry.

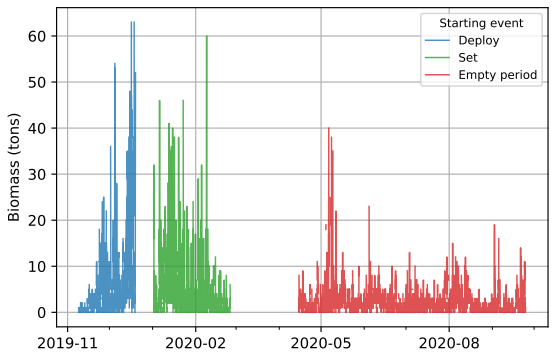
At this point, we are able to divide each historical series using as disruptive events the non acoustic observations that appear now in the merged data set. Moreover, we also split the series when there were no echo-sounder data for 6 days in a row, even if there was not any human activity. These empty intervals are either caused by the absence of tuna, or by buoy inactivity. Without a method to discriminate between these two causes, the study drops wide periods of no echo-sound signal. Lesser empty intervals (up to 5 days) are expected to happen, specially after a disruptive event, because it has been observed that tuna takes at least 5 days to aggregate under a FAD (Orue et al., 2019b), and thus are



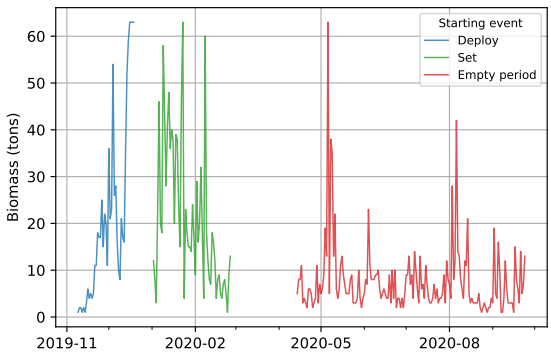
(a) Historical biomass series under a sample dFAD (in blue), together with the disruptive events that happen within this interval (marked as dashed lines whenever they happen).



(b) The series is sliced when a disruptive event happens, or after a long period without acoustic records. Each color represents an individual split.



(c) Splits shorter than 40 days are dropped.



(d) TUN-AI is applied to estimate the corrected biomass.

Figure 2: Biomass evolution under a sample dFAD.

allowed within the series. Therefore, a partition of the global series of acoustic data may have as a triggering event one of the following: a human activity collected on the FAD logbook, a sequence of missing echo-sounder values during more than 6 days in a row or an unknown event (since it can happen that acoustic data is available in the echo-sounder data set with no previous event recorded in the FAD logbook). Figure 2b shows the splits defined on the sample dFAD biomass evolution series, where each split is plotted in a different colour for better visualization.

In many situations, series are truncated during the process of aggregation to the dFAD. This may happen due to fishermen who lay their nets and capture some of the biomass, a buoy stops sending acoustic data. Therefore, to safely compute the time a tuna school is aggregated to a dFAD and not being influenced by the events that happen at the edges, we will not consider the first and the last 20 days of a given split. This large time interval is a safe choice and, although we lose some valid splits to include in the following studies, this approach ensures that no external interaction is produced and still remains a large sample of biomass estimation series to perform the forthcoming analyses. This decision constrains the splits we computed to last at least 40 days, so every division shorter than that is removed from the data set. Figure 2c shows the biomass evolution of the sample dFAD once the shortest split (which is shorter than 40 days) is dropped.

To end the split generation process, we apply the TUN-AI model to each series. As the model returns daily estimations of the tuna biomass, the new splits end up having daily frequency, which is preferable for our study. This bundling meets the need of computing properly the elapsed time which tuna is aggregated to a dFAD and, as some authors have pointed out before, the presence of tuna depends strongly on solar time (Escalle et al., 2019). Thus, having hourly frequency may cause undesired effects since, in general, biomass measurements are very low at night. Figure 2d shows the sample dFAD biomass estimation of TUN-AI, showing that the noise is reduced and most points are greater than zero, contrary to what happened before.

2.2.3 Smoothing the signal and detecting peaks

To compute the amount of time a tuna school remains associated to a dFAD, we need to define the events when we consider the aggregation begins and ends. The approach we propose is to focus on finding local maxima of the biomass series and then finding the left and right points where the signal has decayed from the maximum to a certain proportion of its value.

As shown in Figure 2d, the raw signal series presents very steep peaks that can entail unrealistic distribution for the elapsed time that tuna school are aggregated. To circumvent this problem, these series are smoothed using both a rolling mean algorithm and a more sophisticated penalized regression smoothing splines approach to reduce the noise and properly estimate the trend of the data. In particular, $P-$splines (Eilers and Marx, 1996) are used to smooth the data due to its good numerical properties. Figure 3 shows the rightmost split shown in Figure 2d together with the smoothing raw signals. Clearly, both smoothings are less influenced by steep behaviour of the acoustic signal and better estimate the trend of the data, making them more appropriate to compute the aggregation times under the dFADs.

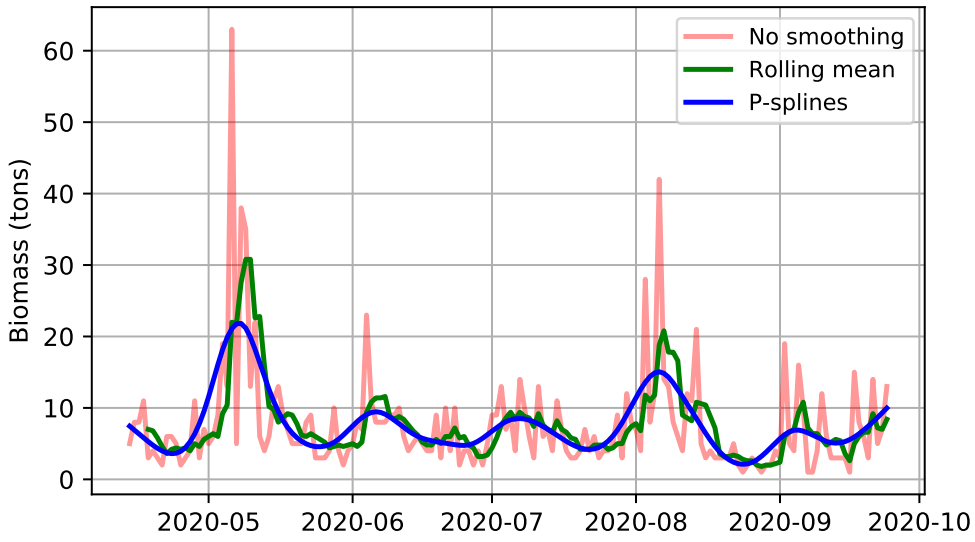


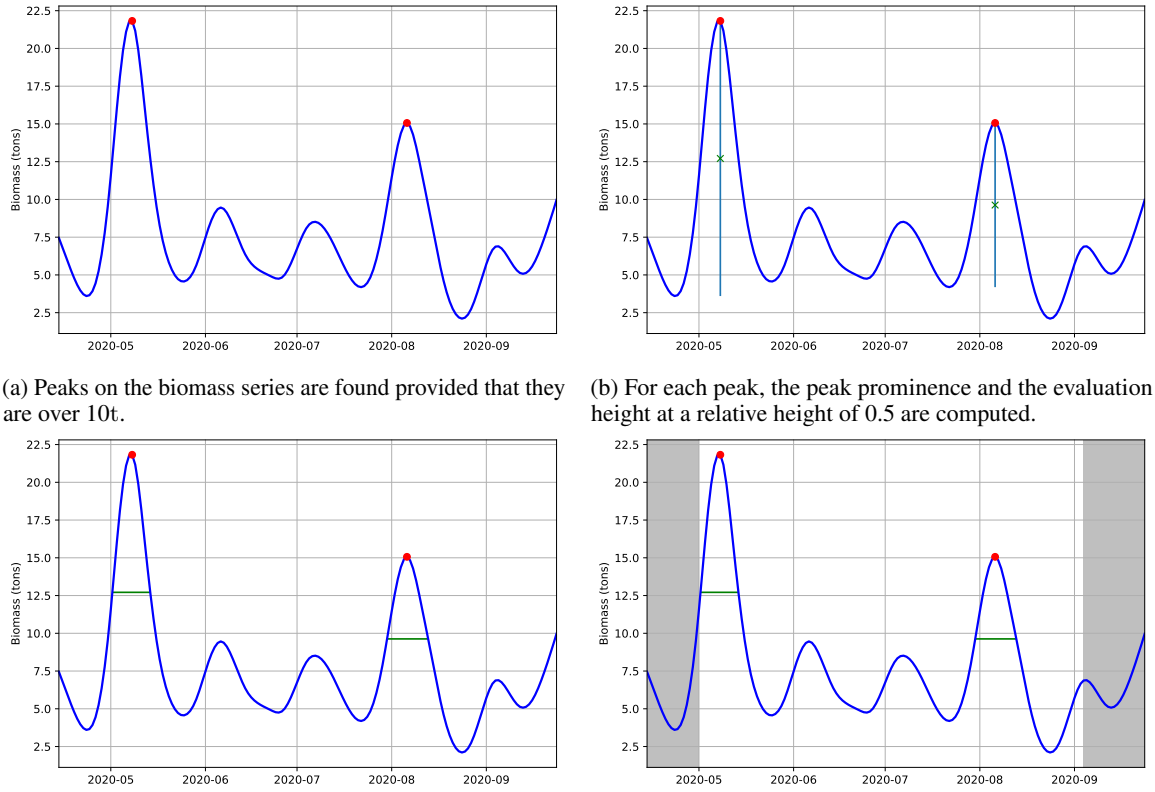
Figure 3: Biomass estimation by TUN-AI model (Precioso et al., 2021) (in pink), together with a rolling mean (in green) and a $P-$splines (in blue) smoothings. 5 days were used in the rolling mean calculation and 40 interior knots were employed in the $P-$splines smoothing.

After smoothing the acoustic series, we proceed by finding the maximum and determining the amount of time tuna schools remain aggregated to the dFADs. The procedure described above can be applied both to the raw biomass series and the smoothings performed and it is illustrated in Figure 4, where the process is carried out for the $P-$splines smoothing.

The algorithm to find peaks is the one implemented in `scipy.signal.find_peaks`[5] (Virtanen et al., 2020), where they are defined as any sample whose two direct neighbours have a smaller amplitude. This is shown in Figure 4a. We also require a minimal biomass estimation of 10t to classify a sample as a peak. This constraint avoids considering small peaks on the final statistics, since they may be caused by tuna school that are not really aggregated to a dFAD. Once all the relevant peaks are located, we compute the elapsed time tuna school are aggregated to dFADs following the steps below:

- Compute the peak prominence, given by the vertical distance between the peak and its lowest contour line. The peak prominences are depicted in light blue solid lines in Figure 4b.
- Calculate the evaluation height, defined as the difference between the peak height and the peak prominence multiplied by 0.5. The evaluation heights on the split used to illustrate the process are marked as green $\times$ in Figure 4b.

[5]Documentation: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html`



(a) Peaks on the biomass series are found provided that they are over 10t.

(b) For each peak, the peak prominence and the evaluation height at a relative height of 0.5 are computed.

(c) The peak width is calculated sketching an horizontal line from the evaluation height of the peak.

(d) Edge peaks are dropped to avoid facing truncated series.

Figure 4: Procedure used to estimate the time tuna school remains aggregated to a dFAD.

- Subsequently, an horizontal line at the evaluation height is drawn from the peak's position until reaching a slope or a signal border. The time we consider the tuna school to be aggregated to the dFAD is given by the width between the two intersecting points. These widths are plotted as solid green lines in Figure 4c.
- Finally, to avoid undesired effects due to the human interaction that may occur at the edges of the split series, peaks found in the first and the last 20 days are removed from the statistics used to estimated the distribution of the elapsed times tuna school remains aggregated. These 20-day intervals are coloured as grey rectangles in Figure 4d.

This definition for the elapsed time of aggregation allow us to define also the aggregation and disaggregation times: the first is given by the difference between the peak and the left intersecting point and the second as the difference between the right intersecting point and the peak.

3 Results

The data preprocessing detailed in Section 2.2.2 reduces the initial dataset to 1 069 063 daily biomass estimations that are divided in 9696 splits. As we previously stated, the biomass estimation on each of these splits is now smoothed using two smoothing techniques: rolling mean and cubic $P-$splines.

In this setting, choosing the right parameters in both procedures is essential for properly determining the time distributions. Indeed, a very long period in the moving average or a low number of interior knots in $P-$splines may underfit the data, which entails a major problem due to the steep behaviour of the tuna estimation series. Therefore, we analyse the time distribution for a variety of parameters: the period for the moving average calculation is set as 2, 5 and 8 days and the number of equidistant interior knots in the $P-$splines algorithm is fixed as 20, 40 and 60. Finally, $P-$splines also requires to select the smoothing parameter, which has a crucial effect on the quality of the fit. In this study, the smoothing parameter is chosen based on minimizing Generalized Cross-Validation (GCV) (Golub et al., 1979) over the set 2^n with $n \in \{-15, -13, -11, \ldots, 5, 7\}$.

Table 2 contains the number of peaks found using the procedure described in Figure 4 by each type of smoothing. Moreover, many statistical indicators describing the distribution of the elapsed time that a tuna school was aggregated to a dFAD are added in each case. Results show that the maximum number of peaks is attained when no smoothing is performed and that the amount of peaks found decreases as the time period (number of knots) increases (decreases) in the rolling mean ($P-$splines) algorithm, as expected. In addition, we observe that even in the most extreme case for $P-$splines with 20 interior knots (when data may be underfitted, and therefore the peak widths is overestimated), the time difference between the 5th and the 95 percentile is slightly above a month.

Smoothing	Parameter	Nr. of peaks	Median (days)	IQR (days)	5th-95th percentile dist. (days)
No smoothing	-	128 276	1	2	9
Rolling mean	2 days	86 311	2	3	13
	5 days	61 586	3	5	18
	8 days	52 939	2	7	21
$P-$splines	60 knots	46 214	4	6	19
	40 knots	37 497	5	7	23
	20 knots	22 748	9	10	34

Table 2: Number of peaks for the distribution of the elapsed time that a tuna school was aggregated to a dFAD. The median, the interquartile distance and the difference between the 5th and the 9th percentile of the distribution (in days) are included.

To better illustrate the time distribution that a tuna school remains associated to a dFAD, Figure 5 represents the violin plot of this time distribution for the three smoothings performed using $P-$splines. Violin plots are similar to box plots, but additionally they contain a rotated kernel density plot on each side which will improve the understanding of the statistical indicators shown in Table 2. We chose to plot $P-$spline results since they reported the largest records of these indicators. From Figure 5, we see the kernel density plots associated to $P-$splines smoothings with 40 and 60 interior knots are really similar and present a high level of positive skewness, while the remaining density show the same behaviour but not so left-tailed.

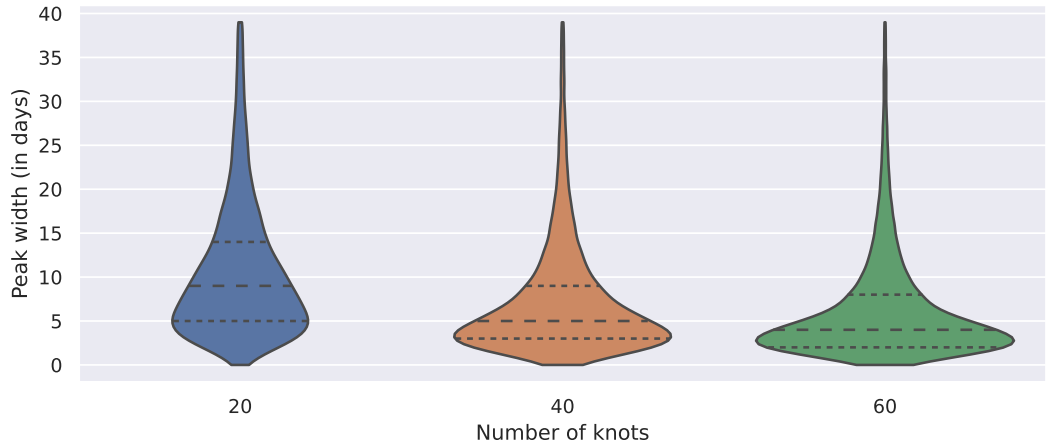


Figure 5: Violin plot of the time distribution that a tuna school remains associated to a dFAD using $P-$splines smoothing. For each distribution, dashed lines show the position of the quartiles.

Finally, we address the issue of determining the time distributions of the aggregation and disaggregation processes, that is to say, the times between the moment when it is considered that tuna start to be aggregated to the dFAD and the maximum value the biomass estimation attains, and from this point and the moment when it is deemed that the tuna school disaggregates from the buoy, respectively. Although we have determined using a wide range of smoothings the time distribution that tuna schools remains associated to a dFAD, it is not straightforward that the aggregation and dissagregation times should be equally distributed.

To explore this question, we estimate the aggregation and disaggregation times for each peak on the $P-$splines smoothings, which yield wider distributions as is shown in Table 2. The resulting aggregation and disaggregation distribution are displayed in Figure 6. From this figure, we can conclude that the distribution of both processes are mostly symmetrical and no remarkable differences are appreciated.

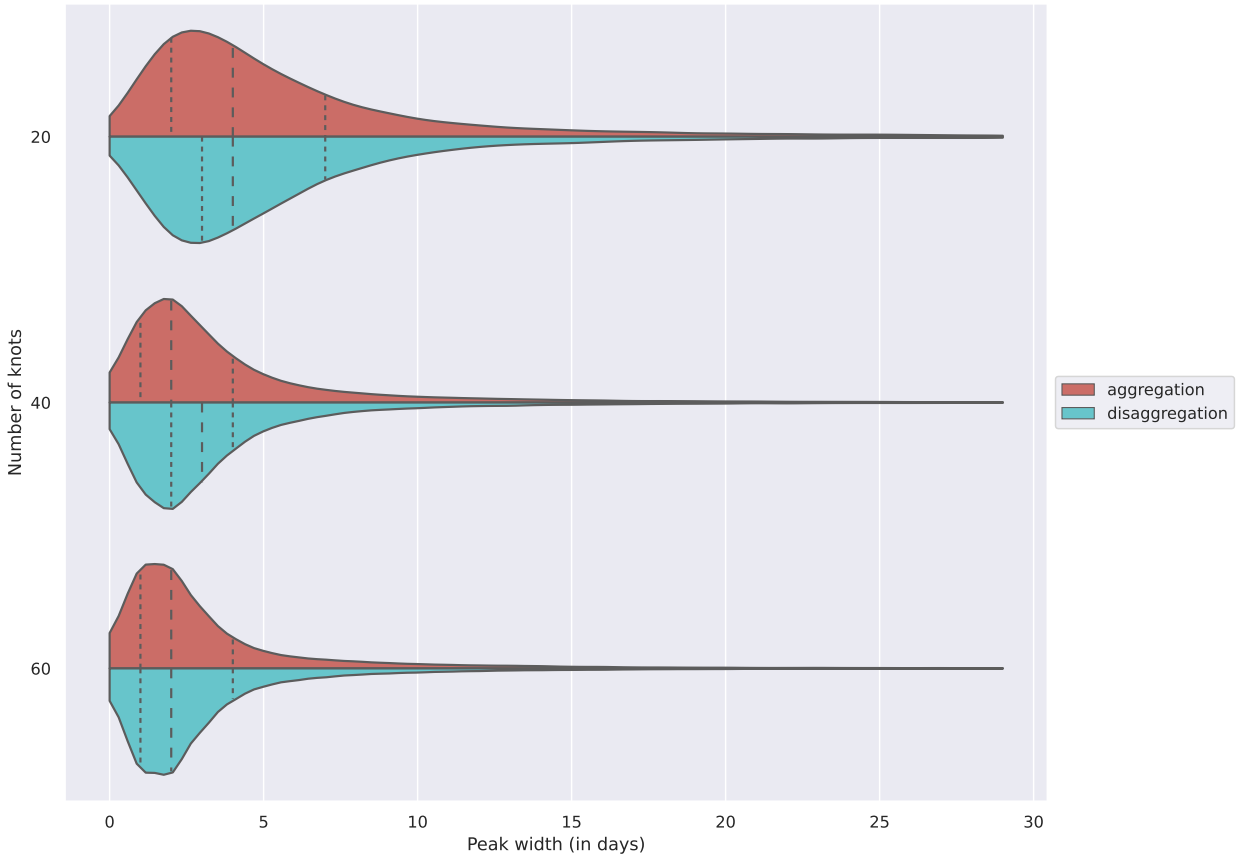


Figure 6: Violin plot of the aggregation (in blue) and disaggregation (in orange) processes time distribution.

4 Discussion

The current study contribute to improve our knowledge on the temporal aggregation dynamics of tropical tuna around dFADs, through the data provided by echo-sounder buoys deployed across all major oceans. In general, our results show that the process of aggregation of tunas to dFADs is symmetrical, with aggregation times within the range of disaggregation times, and that both of these processes tend to occur gradually, taking several days (Figure 6). Even with the most conservative smoothing method ($P-$splines, 20 knots), the median time spent by tunas aggregating at dFADs was less than 10 days, while other less conservative methods showed even shorter time ranges, with a median of 5 days or less (Table 2). Considering IQR we see similar results, with aggregated times between 2 and 10 days (Table 2).

These results are in accordance to those found in several small scale studies on tuna behaviour around floating objects. For instance, Govinden et al. (2021) studied the associative behaviour of tuna species around dFADs in the Indian Ocean, and found that bigeye tuna showed the longest continuous residence time, at 7.59 days, followed by yellowfin (6.64 days) and skipjack (4.58 days). Others have found mean residence times of 5 to 8 days for yellowfin and bigeye (Dagorn et al., 2007). These results are in the range of the current study's findings, where median times of aggregation were between 1 and 9 days according to the processing applied to the TUN-AI biomass estimates (Precioso et al., 2021). Likewise, the maximum continuous residence time found by Govinden et al. (2021) was 26.72 days, which would fall within the ranges of our most conservative analysis ($P-$splines, 20 knots), with 5th-95th percentile distance of 34 days.

It is worth noting that the biomass estimates used in the TUN-AI model (Precioso et al., 2021), based on measurements provided by the echo-sounder buoys, are calculated using the target strength of skipjack tuna, so differences in the

associative behaviour of the three main tropical tuna species are not evident. However, recent buoy models, such as the ISD+ buoy included in this study, use multiple echo-sounders to distinguish species composition as well as tonnage estimation. Though this technology isn't yet predominant, it holds potential for studying the large-scale aggregation patterns of different tuna species in the future.

One notable aspect to consider when examining our results, is the fact that tuna aggregation and disaggregation have been evaluated on an entire school basis. This differs from other small-scale studies, which generally work with individually tagged tuna to evaluate their movements around dFADs (Ohta and Kakuma, 2004; Schaefer and Fuller, 2005; Dagorn et al., 2007; Schaefer and Fuller, 2010; Robert et al., 2014; Govinden et al., 2021). One such study found that tagged fish at a dFAD typically departed in small groups, while others remained at the dFAD, concluding the tuna associated to dFADs were generally composed of multiple "sub-schools" rather than one single cohesive unit (Dagorn et al., 2007). Our results would seem to indicate similar patterns, as peak biomass isn't suddenly registered by the echo-sounder buoys, but rather increases over the course of a few days, as occurs for disaggregation (Figure 6). These results could support the "meeting point" hypothesis (Fréon and Dagorn, 2000), by which dFADs allow tuna to encounter other individuals of their same species and form larger schools, providing temporary protection from predators.

Although our research provides some insight into the associative behaviour of tuna around dFADs at a temporal scale, future studies should look into these processes on a spatial scale. This would help identify differences between oceans, but also provide further insight into whether dFADs in areas with favorable oceanographic conditions show a prevalence of tuna aggregation, versus those in areas with non-favorable conditions, as would be expected if the "indicator-log" hypothesis were true (Hall et al., 1992). Studying the permanence of those aggregations as the dFADs drift out of favorable conditions could also serve to evaluate whether dFADs constitute an "ecological trap" (Marsac et al., 2000; Hallier and Gaertner, 2008). These sort of analyses are only now possible at a much larger scale than previously, thanks to the data collected by echo-sounder buoys deployed with dFADs across all major oceans which are now being recognized as an unprecedented resource for scientific research into the behavioural patterns of pelagic species such as tuna (Orue et al., 2019a; Santiago et al., 2020; Baidai et al., 2020a).

Acknowledgments

This study has been conducted using E.U. Copernicus Marine Service Information. We also thank AGAC for providing the logbook data used in the analysis and the helpful comments about the manuscript. The authors would also like to thank Carlos Roa for rendering available the Satlink echosounder dataset. The research of DGU has been supported in part by the Spanish MICINN under grants PGC2018-096504-B-C33 and RTI2018-100754-B-I00, the European Union under the 2014-2020 ERDF Operational Programme and the Department of Economy, Knowledge, Business and University of the Regional Government of Andalusia (project FEDER-UCA18-108393).

References

Baidai, Y., Dagorn, L., Amande, M. J., Gaertner, D., and Capello, M. (2020a). Machine learning for characterizing tropical tuna aggregations under Drifting Fish Aggregating Devices (DFADs) from commercial echosounder buoys data. *Fisheries Research*, 229:105613.

Baidai, Y., Dagorn, L., Amandè, M. J., Gaertner, D., Capello, M., and Proud, R. (2020b). Tuna aggregation dynamics at Drifting Fish Aggregating Devices: a view through the eyes of commercial echosounder buoys. *ICES Journal of Marine Science*, 77(7-8):2960–2970.

Castro, J. J., Santiago, J. A. J., and Santana-Ortega, A. T. (2002). A general theory on fish aggregation to floating objects: An alternative to the meeting point hypothesis. *Reviews in Fish Biology and Fisheries*, 11(3):24.

Dagorn, L., Holland, K. N., and Itano, D. G. (2007). Behavior of yellowfin (Thunnus albacares) and bigeye (T. obesus) tuna in a network of fish aggregating devices (FADs). *Marine Biology*, 151(2):595–606.

Dagorn, L., Holland, K. N., Restrepo, V., and Moreno, G. (2012). 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(3):391–415.

Davies, T. K., Mees, C. C., and Milner-Gulland, E. J. (2014). The past, present and future use of drifting fish aggregating devices (FADs) in the Indian Ocean. *Marine Policy*, 45:163–170.

Dempster, T. and Taquet, M. (2004). Fish aggregation device (FAD) research: gaps in current knowledge and future directions for ecological studies. *Reviews in Fish Biology and Fisheries*, 14:21–42.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties. Technical Report 2.

Escalle, L., Heuvel, B. V., Clarke, R., Brouwer, S., Pilling, G., Lauriane Escalle, Heuvel, B. V., Clarke, R., Brouwer, S., and Pilling, G. (2019). Report on preliminary analyses of FAD acoustic data. *Western and Central Pacific Fisheries Commission*, 53(9):17.

Fréon, P. and Dagorn, L. (2000). Review of fish associative behaviour: toward a generalisation of the meeting point hypothesis. *Reviews in Fish Biology and Fisheries*, 10(2):183–207.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232.

Global Monitoring and Forecasting Center (2018). Operational Mercator global ocean analysis and forecast system, E.U. Copernicus Marine Service Information. *https://resources.marine.copernicus.eu (Accessed: 15th January 2021)*.

Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized Cross-Validation a s a Method f or Choosing a Good Ridge Parameter. Technical Report 2.

Govinden, R., Capello, M., Forget, F., Filmalter, J. D., and Dagorn, L. (2021). Behavior of skipjack (Katsuwonus pelamis), yellowfin (Thunnus albacares), and bigeye (T. obsesus) tunas associated with drifting fish aggregating devices (dFADs) in the Indian Ocean, assessed through acoustic telemetry. *Fisheries Oceanography*, (00):1–14.

Hall, M., Lennert-Cody, C., Garcia, M., and Arenas, P. (1992). Characteristics of floating objects and their attractiveness for tunas. In *Proceedings of the International Workshop on the Ecology and Fisheries for Tunas Associated with Floating Objects*, pages 396–446.

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

ISSF (2021). Status of the World Fisheries for Tuna. Mar 2021. *ISSF Technical Report 2021-10*, March 2021(March):1–120.

Lopez, J., Moreno, G., Boyra, G., and 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(2):166–178.

Marsac, F., Fonteneau, A., and Ménard, F. (2000). Drifting FADs used in tuna fisheries: an ecological trap? *Biology and behaviour of pelagic fish aggregations*, (July 2015):17.

Maufroy, A., Chassot, E., Joo, R., and Kaplan, D. M. (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.

Ohta, I. and Kakuma, S. (2004). Periodic behavior and residence time of yellowfin and bigeye tuna associated with fish aggregating devices around Okinawa Islands, as identified with automated listening stations. *Marine Biology*, 146(3):581–594.

Orue, B., Lopez, J., Moreno, G., Santiago, J., Boyra, G., Uranga, J., and Murua, H. (2019a). From fisheries to scientific data: A protocol to process information from fishers' echo-sounder buoys. *Fisheries Research*, 215(February):38–43.

Orue, B., Lopez, J., Moreno, G., Santiago, J., Soto, M., and Murua, H. (2019b). Aggregation process of drifting fish aggregating devices (DFADs) in the Western Indian Ocean: Who arrives first, tuna or non-tuna species? *PLOS ONE*, 14(1):e0210435.

Precioso, D., Navarro-García, M., Gavira-O'Neill, K., Torres-Barrán, A., Gordo, D., and Gómez-Ullate, D. (2021). Tuna-AI : tuna biomass estimation with Machine Learning models trained on oceanography and echosounder FAD data. pages 1–18.

Ramos, M. L., Báez, J. C., Grande, M., Herrera, M. A., López, J., Justel, A., Pascual, P. J., Soto, M., Murua, H., Muniategi, A., and Abascal, F. J. (2017). Spanish FADs logbook: solving past issues, responding to new global requirements. *1st Ad-Hoc IOTC Working Group on FADs*, 2017(April):1–24.

Robert, M., Dagorn, L., Bodin, N., Pernet, F., Arsenault-Pernet, E.-J., Deneubourg, J. L., and Rose, K. (2014). Comparison of condition factors of skipjack tuna (Katsuwonus pelamis) associated or not with floating objects in an area known to be naturally enriched with logs. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(3):472–478.

Santiago, J., Lopez, J., Moreno, G., Murua, H., Quincoces, I., and Soto, M. (2016). Towards a Tropical Tuna Buoy-Derived Abundance Index (TT-BAI). *Collective Volume of Scientific Papers ICCAT*, 72:714–724.

Santiago, J., Uranga, J., Quincoces, I., Orue, B., Grande, M., Murua, H., Merino, G., Urtizberea, A., Pascual, P., and Boyra, G. (2020). A Novel Index of Abundance of Juvenile Yellowfin Tuna in the Atlantic Ocean Derived from Echosounder Buoys. *Collective Volume of Scientific Papers ICCAT*, 76:321–343.

Schaefer, K. M. and Fuller, D. W. (2005). Behavior of bigeye (Thunnus obesus) and skipjack (Katsuwonus pelamis) tunas within aggregations associated with floating objects in the equatorial eastern Pacific. *Marine Biology*, 146(4):781–792.

Schaefer, K. M. and Fuller, D. W. (2010). Vertical movements, behavior, and habitat of bigeye tuna (Thunnus obesus) in the equatorial eastern Pacific Ocean, ascertained from archival tag data. *Marine Biology*, 157(12):2625–2642.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). {SciPy} 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

Wain, G., Guéry, L., Kaplan, D. M., Gaertner, D., and O'Driscoll, R. (2021). Quantifying the increase in fishing efficiency due to the use of drifting FADs equipped with echosounders in tropical tuna purse seine fisheries. *ICES Journal of Marine Science*, 78(1):235–245.