A NOVEL INDEX OF ABUNDANCE OF JUVENILE YELLOWFIN TUNA IN THE ATLANTIC OCEAN DERIVED FROM ECHOSOUNDER BUOYS

Josu Santiago[1], *Jon Uranga*[2], *Iñaki Quincoces*[1], *Blanca Orue*[2], *Maitane Grande*[2], *Hilario Murua*[2], *Gorka Merino*[2], *Guillermo Boyra*[2]

SUMMARY

The collaboration with the Spanish vessel-owners associations and the buoy-providers companies, has made it possible the recovery of the information recorded by the satellite linked GPS tracking echosounder buoys used by the Spanish tropical tuna purse seiners and associated fleet in the Atlantic since 2010. These instrumental buoys inform fishers remotely in real-time about the accurate geolocation of the FAD and the presence and abundance of fish aggregations underneath them. Apart from its unquestionable impact in the conception of a reliable CPUE index from the tropical purse seine tuna fisheries fishing on FADs, echosounder buoys have also the potential of being a privileged observation platform to evaluate abundances of tunas and accompanying species using catch-independent data. Current echosounder buoys provide a single acoustic value without discriminating species or size composition of the fish underneath the FAD. Therefore, it has been necessary to combine the echosounder buoys data with fishery data, species composition and average size, to obtain a specific indicator. This paper presents a novel index of abundance of juvenile yellowfin tuna in the Atlantic Ocean derived from echosounder buoys for the period 2010-2017.

1. Introduction

Fishery stock assessment models are demographic analyses designed to determine the effects of fishing on fish populations and to evaluate the potential consequences of alternative harvest policies (Methot & Wetzel, 2012). Quantification of fish populations is the central part of any fish stock assessment and it is commonly the most difficult task. This is even more complicated in the case of highly migratory fish stocks, such as tuna, were conventional fishery-independent surveys are in general not practicable. And, in the absence of fishery-independent information, most of the abundance indices used in fish stock assessments are derived from estimates of Catch per Unit Effort (CPUE), the number or biomass of fish caught as a function of effort (Quinn & Deriso, 1999).

Relative abundance indices based on CPUE data are notoriously problematic (Maunder et al., 2006), as catch data is usually biased by fishing effort, coverage, and other limiting factors of fishery data. The primary assumption behind a CPUE-based abundance index is that changes in the index are assumed to be proportional to changes in the actual stock abundance (Maunder & Punt, 2004), being catchability (q) -the portion of the stock captured by one unit of effort - the coefficient of proportionality. One of the associated difficulties is that q is rarely constant and depends on several different components, such as those related to changes in the fishing efficiency and dynamics of the fleet.

Tropical tuna purse seining is one of such fisheries where both factors, fishing efficiency and dynamics of the fleet, are evolving very rapidly due to the fast technological development (Torres-Irineo et al, 2014) and the sharp increase of the use of Fish Aggregating Devices (FADs) (Scott & Lopez, 2014). This fact makes it difficult to obtain reliable CPUE indices for tropical tunas from purse fisheries fishing with DFADs. Recent initiatives such as the EU funded projects RECOLAPE, CECOFAD-1 and CECOFAD-2 are focusing on the understanding of the use of FADs in tropical purse seine tuna fisheries and to try to provide reliable estimates of abundance indices (Gaertner et al., 2016). And science-industry collaboration in the context of these and other projects is clearly improving the understanding of the FAD use but also the availability of data with great potential for

[1] AZTI, Marine Research Division, Txatxarramendi ugartea z/g, 48395 Sukarrieta, Basque Country (Spain)

[2] AZTI, Marine Research Division, Herrera Kaia-Portu aldea z/g, 20110 Pasaia, Basque Country (Spain)

improving CPUE indices and for developing new novel abundance indicators.

The collaboration with the Spanish vessel-owners associations (ANABAC and OPAGAC) and the buoy-providers companies (Marine Instruments, Satlink and Zunibal), has made it possible the recovery of the information recorded by the satellite tracking echosounder buoys used by the Spanish tropical tuna purse seiners and associated fleet in the Atlantic for the period 2010-2018. These instrumental buoys inform fishers remotely in near real-time about the accurate geolocation of the FAD and the presence and abundance of tuna aggregations underneath them.

Apart from its unquestionable impact in the conception of a reliable CPUE index from the purse seine tropical tuna fisheries fishing on FADs, echosounder buoys have also the potential of being a privileged observation platform to evaluate abundances of tunas and accompanying species using catch-independent data (Dagorn et al., 2006; Lopez et al., 2014; Santiago et al., 2016).

Current echosounder buoys provide a single acoustic value without discriminating species or size composition of the fish underneath the FAD. Therefore, it has been necessary to combine the echosounder buoys data with fishery data, species composition and average size, to obtain a specific indicator. This paper presents a novel index of abundance of juvenile yellowfin tuna in the Atlantic Ocean derived from echosounder buoys for the period 2010-2017.

2. Material and methods

2.1 The acoustic data

Acoustic data from echosounder buoys used in this analysis have been provided by the company Satlink. This type of buoy is equipped with a sounder, which operates at a frequency of 190.5 kHz with a power of 100 W. The range extends from 3 to 115 m, with a transducer blanking zone running from 0 to 3 m. At an angle of 32°, the cone of observation under the buoy has a diameter of 78.6 m at a depth of 115 m. The echosounder provides acoustic information in 10 different vertical layers, each with a resolution of 11.2 m. During the period analysed three different buoy models have been used by the fleet: DS+, DSL+ and ISL+. These three buoy models work with similar beam angel, frequency and power, and with the above mention vertical stratification. DSL+ and DS+ obtain three acoustic records per day, i.e. before dawn, at dawn and after dawn in the default mode. ISL+ has the capacity to sample along the day each 15 minutes, transmitting the signal if the value recorded for a 24 hours period is larger than the previous record.

The fishing companies belonging to ANABAC and OPAGAC that have provided acoustic information from their echosounder buoys were: Albacora SA, Atunera Sant Yago SA, Atunsa, Calvopesca El Salvador SA de CV, Cantabrica de Tunidos SAU, Icube Tuna Fisheries NV, Inpesca Fishing Belize Ltd, Integral Fishing Service INC, Intertuna, NV and Overseas Tuna Company NV. This adds up to a total of 22 purse seine vessels from 7 different flags (Belize, Cabo Verde, Curacao, El Salvador, Spain, Guatemala and Panama) operating in the ICCAT convention area.

The database of acoustic information of the Atlantic Ocean from Satlink buoys comprises around 5 million of records from over 36,000 buoys for the period from January 2010 to December 2017. The information on buoy positions and acoustic information is received in two different data-sets with the following fields:

- Data-set on buoy positions
  - *Date*: Date of the last position of the day
  - *Time*: Hour (GMT)
  - *Buoy code*: Unique identification number of the buoy, given by the model code (D+, DS+, DL+, DSL+, ISL+, ISD+ followed by 5-6 digits.
  - *Latitude:* Latitude of the last position of the day (in decimals)
  - *Longitude*: Longitude of the last position of the day (in decimals)
  - *Velocity*: v calculated from the distance/time between the last position of the day and the last position of the previous day.
  - *Notes:* Empty column

- Data-set on acoustic records
  - *Name*: Unique identification number of the buoy, given by the model code (D+, DS+, DL+, DSL+, ISL+, ISD+ followed by 5-6 digits.
  - *OwnerName*: Name of the buoy owner assigned to a unique purse seine vessel

- *MD*: Message descriptor (160, 161 and 162 for position data, without sounder data, and 163, 168, 169 and 174 for sounder data)
- *StoredTime*: Date (dd/mm/yyyy) and hour (H:MM) of the echosounder record
- *Latitude, Longitude*: Not provided (this information is provided in the position data-set)
- *Bat*: Not provided. (Charge level (in percentage). Except for the D+ and DS+ in voltage)
- *Temp*: Temperature (Not provided)
- *Speed*: Speed in knots (Not provided)
- *Drift*: bearing in degrees (Not provided)
- *Layer1-Layer10*: Depth observation range extends from 3 to 115 m, which is split in ten homogeneous layers, each with a resolution of 11.2 m. The buoy has also a blanking zone (a data exclusion zone to eliminate the near-field effect of the transducer between 0 and 3 m. 32 pings are sent from the transducer and an average of the backscattered acoustic response is computed and stored in the memory of the buoy. Manufacturer´s method converts raw acoustic backscatter into biomass in tons, using a depth layer echo-integration procedure based exclusively on an algorithm based on the TS and weight of skipjack tuna.
- *Sum*: Sum of the biomass estimated at each layer
- *Max*: Maximum biomass estimated at any layer
- *Mag1, Mag3, Mag5 and Mag7*: Magnitudes corresponding to the counts of detected targets according to the TS of the detection peak.

2.2 From acoustic data to a species-specific abundance indicator

To calculate the biomass aggregated under a dFAD from the acoustic signal, Satlink uses the density of one species, skipjack, to provide the biomass in tons, biomass data from Satlink was converted to decibels reversing their formula for the biomass computation. Then we recomputed biomass using standard abundance estimations equations (Simmonds & MacLennan, 2005):

$$Biomass_i = \frac{s_V \cdot Vol \cdot p_i}{\sum_i \sigma_i \cdot p_i}$$

where *Vol* is the sampled volume and p_i and σ_i are the proportion and linearized target strength of each species *i* respectively. Species proportions in weight were extracted from ICCAT T2 data of the EU fleet for each 1ºx1º and month stratum, as explained below. Mean fish lengths (L_i) used for SKJ, BET and YFT were 48, 47 and 46 cm respectively[3], and weights were obtained using weight-length relationships (ICCAT conversion factors). Then, the following TS-length relationships were used to obtain linearized target strength per kilogram:

$$\sigma_i = \frac{10^{(20\log(L_i)+b_{20,i})/10}}{w_i}$$

Where w_i is the mean weight of each species. Given that each brand uses different operating frequencies, we used different b20 values for each. For Satlink, the b20 values were obtained from (Boyra et al., 2018) for SKJ and (Oshima, 2008) for YFT and BET.

Since all acoustic data is not covered with a corresponding species distribution data from the T2, assignment of the distribution were done following a three steps approach: first, buoy distributions and species composition at the same 1x1 grid, year and month were filtered and assigned; then, to fill uncovered acoustic data from first step, data at the same 1x1 grid and trimester were filtered and assigned; and finally, remaining uncovered acoustic data took the distribution of 5 large regions (region 1:longitude<35W and latitude>25N; region 2: longitude >35W and latitude >10N; region 3: longitude <35W and latitude <=25N; region 4: longitude >35W and latitude <=10N; and region 5: latitude <10S) over the Atlantic ocean for each trimester. Steps 1 to 3 provided species characterization for 6%, 59% step 2 and 35% of the acoustic data, respectively.

[3] Mean length values used at this stage are preliminary. In further revisions of the document the idea is to work with mean lengths of the PS FAD fishery catches by finer spatial-temporal strata.

2.3 Acoustic data cleaning and filtering

Data cleaning included the removal of records without acoustic information (records with only position, speed and velocity), outliers (invalid, impossible or extreme values) related to bad geolocation, time, or other general variables. And aside of the regular exclusions due to this type of inconsistencies, the following considerations were also taken into account for accepting the data for the standardization analysis:

- Vertical boundary between tuna and non-tuna species: acoustic information from the shallower layers, <25m, was not considered for the analysis. According to (Lopez et al., 2017; Robert et al., 2013), the vertical boundary between non-tuna species and tunas can be considered at about 25 m. Excluding the first layers, we try to eliminate noise from the non-tuna species associated to the FAD.
- Bottom depth: Using high resolution bathymetry data (British Oceanographic Data Centre, UK, www.gebco.net), acoustic records from buoys located in areas with a bottom depth shallower than 200 m were excluded. The rational of this exclusion is not to incorporate acoustic records of FADs that have drifted to coastal areas were tunas are less likely to be present.
- Acoustic measurements at sea: Buoys are normally turned on before deployment, so some records may correspond to onboard buoys. To deal with this issue we developed a Random Forest (RF) model (Orue et al., 2019) to classify the buoys at sea and onboard using information from Zunibal buoys. These buoys have the capability to identify between positions at sea or onboard, using a conductivity sensor. The sensor measures the current between two electrodes and through a simple algorithm it determines whether the Zunibal buoy is inside or outside the water. Those records classified as onboard were excluded.
- Time of the day: Only those samples obtained around sunrise, between 4 a.m. and 8 a.m., were considered for the analysis. These samples are supposed to capture the echosounder biomass signals that better represents the abundance of fish under the FADs, as this is the time when tuna is observed to be more closely aggregated around the FADs (Brill et al., 1999; Josse et al., 1998; Moreno et al., 2007). For the specific case of comparing the acoustic data with abundance it is important that the echosounder measurements are received when the signal is more representative of the biomass around the FAD model (Orue et al., 2019).
- Days since deployment: The objective of this selection criterium was to consider those acoustic records that were more likely associated to FAD trajectory, termed “virgin segments”.
  A virgin segment is defined as the segment of a buoy trajectory whose associated FAD likely represents a new deployment which has been potentially colonized by tuna and not already fished. Orue et al. (2019) concluded that tuna seemed to arrive at DFADs in 13.5±8.4 days and, thus, we consider as virgin segments (i.e. when tuna has aggregated to DFAD) those segments of trajectories from 20-35 days at sea.
  In order to identify and separate those segments and their acoustic samples, the overall trajectories of the entire life-time of each buoy were fractioned in smaller sequences corresponding to periods where they could have been attached to different FADs. A new sequence of a buoy was considered to occur, and hence an attachment to a new FAD, when the difference between two consecutive observations of the same buoy was larger than 30 days. Each sequence was assigned with a “new trajectory code” that included the code of the buoy plus the consecutive number of the sequence of each buoy. A deployment/redeployment of a buoy was considered to occur when the “new trajectory code” appears for the first time in the database.
  Sequences with less than 30 observations were excluded from the analysis. Sequences having a time difference between any of the consecutive observations longer than 4 days during the first 35 days were also excluded.
  Figure 1 shows a diagram with an example of “virgin” segments used for the calculation of the BAI index.
- Detection threshold: Acoustic records equal or less than 0,1 tonnes were considered zeros. This is a conservative preliminary value since further validation is needed.

2.4 The BAI index: Buoy-derived Abundance Index

The estimator of abundance BAI was defined as the 0.9 quantile of the integrated acoustic energy observations in each of the "virgin" sequences. A high quantile was chosen because the large values are considered to be likely produced by tuna (in opposition to plankton or bycatch species). This assumption is followed by all the buoy

brands in the market, which use the maximum value as the summary of each time interval. In our case we selected a high quantile instead of the maximum to try to provide a more robust estimator by avoiding eventual outlier values. We did this to avoid taking into account the expected lowest values that might appear after eventual hauls occurring along the sequence. The total number of "virgin" sequences analysed, and hence the number of observations in the model, rose to 49,880, of which 43,323 (86.85%) were positives.

2.5 Covariates

Covariates included year-quarter (yyqq), and 5°x5º ICCAT areas fitted as categorical variables. Other variables used in the standardization process included velocity of the buoy, FAD densities and a set of environmental variables. They were chosen for potential effects on the horizontal-vertical distribution of tunas and their association to FADs (FAD density, mixed layer height, sea surface temperature, chlorophyll concentration and detected fronts in sea surface temperature and chlorophyll daily datasets computed with Belkin and O'Reilly method) or on the quality of the echosounder measurements (buoy velocity). These variables were incorporated in the model as continuous variables.

A proxy of 1ºx1º and monthly FAD densities were calculated as the average number of buoys over each month by summing up the total number of active buoys recorded per day over the entire month and dividing by the total number of days.

The environmental variables evaluated in the model were:

- Ocean mixed layer thickness: defined as the depth where the density increase compared to density at 10 m depth corresponds to a temperature decrease of 0.2°C in local surface conditions (θ10m, S10m, P0= 0 db, surface pressure).
    - Source: Copernicus Marine Environment Monitoring Service (http://marine.copernicus.eu)
    - Product: GLOBAL_REANALYSIS_PHY_001_030
    - Update frequency: Yearly
    - Available time series: 04/12/1992 to 27/12/2017
    - Temporal resolution: daily mean
    - Horizontal resolution: 1/12 ° (equirectangular grid)
    - Units: [m]
- Chlorophyll: Mass concentration of chlorophyll a in sea water (depth = 0).
    - Source: Copernicus Marine Environment Monitoring Service (http://marine.copernicus.eu)
    - Product: GLOBAL_REANALYSIS_BIO_001_029
    - Available time series: 1993/01/01 up to 2017/12/31
    - Temporal resolution: daily mean
    - Horizontal resolution: 1/4 ° (equirectangular grid)
    - Units: [mg.m-3]
- Sea Surface Temperature (SST):
    - Source: Multi-scale Ultra-High-Resolution Sea Surface Temperature (https://mur.jpl.nasa.gov)
    - Product: JPL_OUROCEAN-L4UHfnd-GLOB-G1SST
    - Available time series: 2010/01/01 up to present
    - Target delivery time: daily
    - Temporal resolution: daily mean
    - Horizontal resolution: Regular 0.01-degree grid
    - Units: Kelvin degrees

SST and Chlorophyll fronts: Oceanographic front detection was performed using the "grec" package for R for each daily dataset, that provides algorithms for detection of spatial patterns from oceanographic data using image processing methods based on Gradient Recognition (Belkin & O'Reilly, 2009).

Figure 2 shows an example of one month of data, January 2010, of buoy trajectories, density of buoys by 1ºx1º rectangles and environmental variables incorporated in the GLM analysis.

2.6 The model

The model we propose is based in an assumption very similar to the fundamental relationship among CPUE and abundance widely used in quantitative fisheries analysis. In our case we built the index based on the assumption that the signal from the echosounder is proportional to the abundance of fish.

$$BAI_t = \varphi \, . B_t$$

where BAI_t is the Buoy-derived Abundance Index and B_t is the abundance in time t (Santiago et al., 2016).

Although it would appear to be obvious, there is not plenty of literature on the relationship between acoustic indicators and fishing performance. It is assumed that acoustic echo integration is a linear process, i.e., proportional to the number of targets (Simmonds & MacLennan, 2005) and has been experimentally proven to be correct with some limitations (Foote, 1983; Røttingen, 1976). Therefore, acoustic data (echo integration) is commonly taken as an estimator of abundance and is thoroughly applied to provide acoustic estimation of abundance of many pelagic species (e.g., Hampton, 1996; ICES, 2015; Massé, Uriarte, Angélico, & Carrera, 2018). So, does large biomass indicated by acoustic data consistently result in large catches when sets follow soon after? A recent study has found a positive significant correlation between echosounder acoustic energy and catches of tropical tuna around FADs (Moreno et al., in press)The study was based on data collected in two surveys conducted onboard commercial purse seiners during regular fishing activity in the Atlantic and Pacific Oceans in years 2014 and 2016. Simrad EK60 echosounders with split beam transducers at 38, 120 and 200 kHz were used to collect acoustic data. Generalized linear models were run between acoustic backscattering energy (NASC, Maclennan et al., 2002) and catches of the three main tuna species found at FADs (skipjack, bigeye and yellowfin) providing positive significant relationships with total catches (in weight) as well as catches for each tuna species.

As with the catchability, the coefficient of proportionality φ is not constant for many reasons. In order to ensure that φ can be assumed to be constant (i.e. to control the effects other than those caused by changes in the abundance of the population) a standardization analysis should be performed aiming to remove factors other than changes in abundance of the population. This can be performed standardizing nominal measurements of the echosounders using a Generalized Linear Mixed Modelling approach.

Because of the significant proportion of records with zero abundance (13.15%), delta lognormal analyses (Lo et al., 1992) were carried out using generalized linear models. Delta lognormal analysis estimates the predicted abundances as the result of two processes: i) the probability of encounter tropical tuna in the acoustic observations (proportion of positives) and, ii) the mean relative abundance given that a positive observation has been realized. Then the estimated Buoy-derived Abundance Index (BAI) is the product of these two processes.

For each component, a stepwise regression was applied to the model with all the explanatory variables and interactions in order to determine those that significantly contributed to explain the deviance of the model. For this, deviance analysis tables were created for the proportion of positive observations (e.g., positive acoustic records/total records), and for the positive acoustic records. Final selection of explanatory factors was conditional to: a) the relative percentage of deviance explained by adding the factor in evaluation (normally factors that explained more than 5% were selected), and b) The Chi-square ($\chi 2$) significance test. Those factors that explained less than 5% of the variability of the model were not considered.

Interactions of the temporal component (year-quarter) with the rest of the variables were also evaluated. If an interaction was statically significant, it was then considered as a random interaction(s) within the final model (Maunder & Punt, 2004).

Lastly, the selection of the final mixed model was based on the Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC), and a Chi-square ($\chi 2$) test of the difference between the log-likelihood statistic of different model formulations. Once having a final model selected, the relative indices for the Delta lognormal model formulation were calculated as the product of the year-quarter effect least square means (LSmeans) from the binomial and the lognormal model components (Punt et al., 2000). These LSmeans estimates use a weighted factor of the proportional observed margins in the input data to account for the non-balance characteristics of the data. The LSMeans of the lognormal component were bias corrected for the logarithm transformation algorithms using Lo et al (1992). All analyses were done using the lme4 package for R (Bates et al., 2015).

3. Results

A total of 5 million of records from over 36,000 buoys for the period from January 2010 to December 2017 were integrated into 49,880 observations for the GLM analysis. Each observation was calculated as the 90% percentile of a "virgin" segment of buoy trajectories. A virgin segment was defined as the segment of a buoy trajectory from 20-35 days at sea, so that the associated FAD likely represents a new deployment which has been potentially colonized by tuna and not already fished.

In this analysis we have obtained from the acoustic signal of the echosounder buoys associated to FADs the biomass of yellowfin tuna aggregated under a dFAD. The aggregations of yellowfin tuna associated with floating objects are mostly composed of small individuals (FL around 46cm). Therefore, the Buoy-derived Abundance Index (BAI) would represent an indicator of YFT juvenile; a modal size of 46cm would correspond to around 1 year of life.

Figure 3 shows the histograms of the BAI and log transformed BAI nominal values. Log transformation makes the data to follow a normal distribution, as shown in the left panel of Figure 3. Figure 4 shows the spatial distribution [5ºx5º] of the number of "virgin" sequences of buoy trajectories that have been used in the GLM analysis. The quarterly evolution of the number of observations on a 5ºx5º grid is shown in Figure 5. The number of observations available has grown considerably over the years, especially since 2013. Only the area between 5ºN and 5ºS and east of 10ºW shows a similar pattern of evolution of the number of observations; the rest of 5ºx5º areas show, in general, non-comparable and more irregular patterns.

Figure 6 shows a tableplot of the variables used in the GLM analysis. Independent variables tested in the GLM were year-quarter (yyqq), 5ºx5º area (area), buoy model (model), buoy velocity (vel), FAD density (den), chlorophyll concentration (chl), detected fronts in chlorophyll (chlfront), sea surface temperature (sst), detected fronts in sst (sstfront) and mixed layer height (mid). The dependent variable for the lognormal component was the 0.9 quantile of the integrated acoustic energy observations in the "virgin" sequence (index); and for the binomial component, the proportion of positives (posit).

Figure 7, Figure 8 and Figure 9 show boxplots of log (index) and proportion of positives for each of the independent variables, expressed as categorical. Figure 10 and Figure 11 show the quarterly evolution of the log BAI index and the proportion of positive observations by squares of 5x5 degrees from 2010 to 2017.

The results of the deviance analysis are shown in Table 1. The most significant explanatory factors for the binomial model on the proportion of positives included year-quarter and buoy model. In the binomial component of the model, fitted probabilities of 0 or 1 may occur. This is known as 'perfect separation' and is problematic for uncertainty estimation (Venables & Ripley 2002). Perfect separation should not bias prediction of the year-quarter effects, as long as a category at which separation occurs is not used in the predictor (Hoyle et al., 2019)

The most significant explanatory factors for the lognormal model on the positive records were year-quarter, 5ºx5º area and the interaction year-quarter*area that was considered as random interaction. No significant residual patterns were observed for either the lognormal or the binomial model Figure 12.

The estimates of the delta lognormal are provided in Table 2 and Figure 13. Most of the nominal values are embedded within the confidence interval of the standardized BAI index. The BAI index shows a general decreasing trend at the beginning of the series, from 2010 to 2012; then a stabilization period at a low level from 2013 to 2016, followed by an increasing trend in 2017 to levels of the beginning of the series. The CVs remain relatively stable (between 10-16%) during the whole time series.

Acknowledgements

This work was supported by the Government of the Basque Country and the EU funded projects RECOLAPE and CECOFAD. We thank the Spanish vessel owner associations ANABAC and OPAGAC and those companies that have provided acoustic information from their echosounder buoys were: Albacora SA, Atunera Sant Yago SA, Atunsa, Calvopesca El Salvador SA de CV, Cantabrica de Tunidos SAU, Icube Tuna Fisheries NV, Inpesca Fishing Belize Ltd, Integral Fishing Service INC, Intertuna, NV and Overseas Tuna Company NV. We are also thankful to the three buoy providers companies Marine Instruments, Satlink and Zunibal.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Belkin, I. M., & O'Reilly, J. E. (2009). An algorithm for oceanic front detection in chlorophyll and SST satellite imagery. *Journal of Marine Systems*, *78*, 319–326. https://doi.org/10.1016/j.jmarsys.2008.11.018

Boyra, G., Moreno, G., Sobradillo, B., Pe, I., Sancristobal, I., & Demer, D. A. (2018). Target strength of skipjack tuna (Katsuwanus pelamis) associated with fish aggregating devices (FADs). *ICES Journal of Marine Science*, *75*, 1790–1802. https://doi.org/10.1093/icesjms/fsy041

Brill, R. V., Block, B. A., Boggs, C. H., Bigelow, K., Freund, E. V., & Marcinek, D. J. (1999). Horizontal movements and depth distribution of large adult yellowfin tuna (Thunnus albacares) near the Hawaiian Islands , recorded using ultrasonic telemetry : implications for the physiological ecology of pelagic fishes. *Marine Biology*, *133*, 395–408.

Dagorn, L., Holland, K., Puente, E., Taquet, M., Ramos, A., Brault, P., … Aumeeruddy, R. (2006). *FADIO (Fish Aggregating Devices as Instrumented Observatories of pelagic ecosystems): a European Union funded project on development of new observational instruments and the behavior of fish around drifting FADs* (No. IOTC-2006-WPTT-16).

Foote, K. G. (1983). Linearity of fisheries acoustics, with addition theorems. *The Journal of the Acoustical Society of America*, *73*(6), 1932–1940. https://doi.org/10.1121/1.389583

Gaertner, D, Ariz, J., Bez, N., Clermidy, S., Moreno, G., Murua, H., & Soto, M. (2016). Objectives and first results of the CECOFAD project. *Collect. Vol. Sci. Pap. ICCAT*, *72*(2), 391–405.

Gaertner, Daniel, Ariz, J., Bez, N., Clermidy, S., Moreno, G., Murua, H., … Marsac, F. (2016). *Results achieved within the framework of the EU research project: Catch, Effort, and eCOsystem impacts of FAD-fishing (CECOFAD)* (No. IOTC-2016-WPTT18-35).

Hampton, I. (1996). Acoustic and egg-production estimates of South African anchovy biomass over a decade: comparisons, accuracy, and utility. *ICES Journal of Marine Science*, *53*(2), 493–500. https://doi.org/10.1006/jmsc.1996.0071

Hoyle, S. D., Huang, J. H., Kim, D. N., Lee3, M. K., Matsumoto, T., & III, J. W. (2019). Standardization of bigeye tuna CPUE in the Atlantic Ocean by the Japanese longline fishery which includes cluster analysis. *Collect. Vol. Sci. Pap. ICCAT*, *75*(7), 2098–2116.

ICES. (2015). *Manual for International Pelagic Surveys (IPS)*. Copenhagen.

Josse, E., Bach, P., & Dagorn, L. (1998). Simultaneous observations of tuna movements and their prey by sonic tracking and acoustic surveys. In *Advances in Invertebrates and Fish Telemetry* (pp. 61–69). https://doi.org/10.1007/978-94-011-5090-3_8

Lo, N. C., Jacobson, L. D., & Squire, J. L. (1992). Indices of Relative Abundance from Fish Spotter Data based on Delta-Lognornial Models. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*(12), 2515–2526.

Lopez, J., Moreno, G., Ibaibarriaga, L., & 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. *Marine Biology*, *164*(3), 44. https://doi.org/10.1007/s00227-017-3075-3

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*, 127–137. https://doi.org/10.1016/j.fishres.2014.02.033

Maclennan, D., Fernandes, P. G., & Dalen, J. (2002). A consistent approach to definitions and symbols in fisheries acoustics. *ICES Journal of Marine Science*, *59*(2), 365–369. https://doi.org/10.1006/jmsc.2001.1158

Massé, J., Uriarte, A., Angélico, M. M., & Carrera, P. (2018). *Pelagic survey series for sardine and anchovy in ICES subareas 8 and 9 – Towards an ecosystem approach*. ICES CRR 332

Maunder, M. N., & Punt, A. E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research*, *70*(2–3), 141–159. https://doi.org/10.1016/j.fishres.2004.08.002

Maunder, M. N., Sibert, J., Fonteneau, A., Hampton, J., Kleiber, P., & Harley, S. J. (2006). Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES Journal of Marine Science*,

63(8), 1373–1385. https://doi.org/10.1016/j.icesjms.2006.05.008

Methot, R. D., & Wetzel, C. R. (2012). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*. https://doi.org/10.1016/j.fishres.2012.10.012

Moreno, G., Boyra, G., Sancristobal, I., & Restrepo, V. (in press). Towards acoustic discrimination of tropical tunas at FADs. *PLOS ONE , Accepted in April 2019*.

Moreno, Gala, Dagorn, L., Sancho, G., & Itano, D. (2007). *summary of Can . J . Fish . Aquat . Sci . In press Fish behaviour from fishers ´ knowledge : the case study of tropical tuna around drifting fish aggregating devices (DFADs)*.

Orue, B., Lopez, J., Moreno, G., Santiago, J., Soto, M., & Murua, H. (2019). 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. https://doi.org/10.1371/journal.pone.0210435

Oshima, T. (2008). *Target strength of Bigeye, Yellowfin and Skipjack measured by split beam echo sounder in a cage.* (No. WPTT-22).

Punt, A. E., Walker, T. I., Taylor, B. L., & Pribac, F. (2000). Standardization of catch and effort data in a spatially-structured shark fishery. *Fisheries Research*, *45*(2), 129–145. https://doi.org/10.1016/S0165-7836(99)00106-X

Quinn, T. J., & Deriso, R. B. (Richard B. . (1999). *Quantitative fish dynamics*. Retrieved from https://books.google.es/books/about/Quantitative_Fish_Dynamics.html?id=5FVBj8jnh6sC&redir_esc=y

Robert, M., Dagorn, L., Lopez, J., Moreno, G., & Deneubourg, J.-L. (2013). Does social behavior influence the dynamics of aggregations formed by tropical tunas around floating objects? An experimental approach. *Journal of Experimental …*, *440*, 238–243. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022098113000099

Røttingen, I. (1976). On the relation between echo intensity and fish density. *FiskDir. Skr. Ser. Havunders.*, *16*(9), 301–314. Retrieved from https://brage.bibsys.no/xmlui/handle/11250/114400

Santiago, J., Lopez, J., Moreno, G., Murua, H., Quincoces, I., & Soto, M. (2016). Towards a tropical tuna buoy-derived abundance index (TT-BAI). *Collect. Vol. Sci. Pap. ICCAT*, *72*(3), 714–724.

Scott, G. P., & Lopez, J. (2014). *The use of FADs in Tuna Fisheries*. European Paliament. Policy Department B: Structural and Cohesion Policies: Fisheries

Simmonds, E. J., & MacLennan, D. N. (2005). *Fisheries acoustics : theory and practice*. Blackwell Science.

Torres-Irineo, E., Gaertner, D., Chassot, E., & Dreyfus-León, M. (2014). Changes in fishing power and fishing strategies driven by new technologies: The case of tropical tuna purse seiners in the eastern Atlantic Ocean. *Fisheries Research*, *155*, 10–19. https://doi.org/10.1016/j.fishres.2014.02.017

Binomial component

Variable	Df	Deviance	Resid..Df	Resid..Dev	Pr..Chi.	Dev..Exp
NULL	NA	NA	49828	38807	NA	NA
yyqq	31	2496	49797	36311	0.000	6.43 %
area	37	387	49760	35923	0.000	1 %
model	2	1457	49758	34466	0.000	3.75 %
den	1	2	49757	34465	0.190	0 %
chl	1	0	49756	34465	0.751	0 %
chlfront	1	27	49755	34438	0.000	0.07 %
sst	1	0	49754	34437	0.569	0 %
sstfront	1	5	49753	34432	0.024	0.01 %
mld	1	10	49752	34422	0.001	0.03 %
yyqq:model	27	143	49725	34279	0.000	0.37 %
yyqq:chl	31	131	49694	34148	0.000	0.34 %
yyqq:sst	31	246	49663	33902	0.000	0.63 %
yyqq:sstfront	31	106	49632	33796	0.000	0.27 %
yyqq:mld	31	133	49601	33663	0.000	0.34 %

Lognormal component

Variable	Df	Deviance	Resid..Df	Resid..Dev	F	Pr..F.	Dev..Exp
NULL	NA	NA	43271	73385	NA	NA	NA
yyqq	31	5537	43240	67848	150	0e+00	7.54 %
area	37	9527	43203	58322	217	0e+00	12.98 %
model	2	542	43201	57780	228	0e+00	0.74 %
den	1	197	43200	57583	166	0e+00	0.27 %
chl	1	1162	43199	56421	977	0e+00	1.58 %
chlfront	1	71	43198	56351	59	0e+00	0.1 %
sst	1	19	43197	56332	16	1e-04	0.03 %
sstfront	1	221	43196	56111	186	0e+00	0.3 %
mld	1	19	43195	56092	16	1e-04	0.03 %
yyqq:area	1087	5191	42108	50902	4	0e+00	7.07 %
yyqq:model	29	171	42079	50731	5	0e+00	0.23 %
yyqq:den	31	228	42048	50503	6	0e+00	0.31 %
yyqq:chl	31	183	42017	50319	5	0e+00	0.25 %
yyqq:chlfront	31	145	41986	50175	4	0e+00	0.2 %
yyqq:sst	31	147	41955	50028	4	0e+00	0.2 %
yyqq:sstfront	31	135	41924	49892	4	0e+00	0.18 %
yyqq:mld	31	94	41893	49799	3	0e+00	0.13 %

Table 1. Deviance tables for the binomial (top) and the lognormal (bottom) components of the Delta-lognormal model of the 2010-20177 period. Significant (p<0.05) factors and interactions explaining >4% of total deviance are highlighted.

Quarter	BAI nominal	BAI Index	BAI se	BAI cv
10Q1	0.582	0.442	0.068	0.153
10Q2	0.640	0.439	0.064	0.145
10Q3	0.627	0.412	0.065	0.158
10Q4	0.912	0.628	0.097	0.155
11Q1	0.656	0.452	0.070	0.155
11Q2	0.636	0.511	0.077	0.152
11Q3	0.533	0.373	0.060	0.160
11Q4	0.435	0.330	0.051	0.155
12Q1	0.304	0.234	0.035	0.150
12Q2	0.472	0.344	0.052	0.150
12Q3	0.234	0.220	0.034	0.155
12Q4	0.171	0.165	0.025	0.150
13Q1	0.126	0.122	0.017	0.143
13Q2	0.164	0.168	0.023	0.139
13Q3	0.185	0.168	0.022	0.129
13Q4	0.277	0.217	0.028	0.130
14Q1	0.229	0.166	0.021	0.126
14Q2	0.201	0.180	0.023	0.127
14Q3	0.276	0.219	0.025	0.116
14Q4	0.297	0.217	0.025	0.117
15Q1	0.189	0.147	0.018	0.119
15Q2	0.208	0.174	0.021	0.121
15Q3	0.280	0.224	0.021	0.093
15Q4	0.344	0.215	0.022	0.102
16Q1	0.178	0.139	0.016	0.113
16Q2	0.186	0.194	0.024	0.122
16Q3	0.283	0.220	0.026	0.119
16Q4	0.310	0.205	0.022	0.108
17Q1	0.231	0.174	0.021	0.119
17Q2	0.293	0.244	0.027	0.111
17Q3	0.401	0.342	0.037	0.109
17Q4	0.704	0.463	0.051	0.109

Table 2. Nominal and standardized Buoy-derived Abundance Index for the period 2010-2017. Standard errors and coefficient of variations of the standardized series are also included.

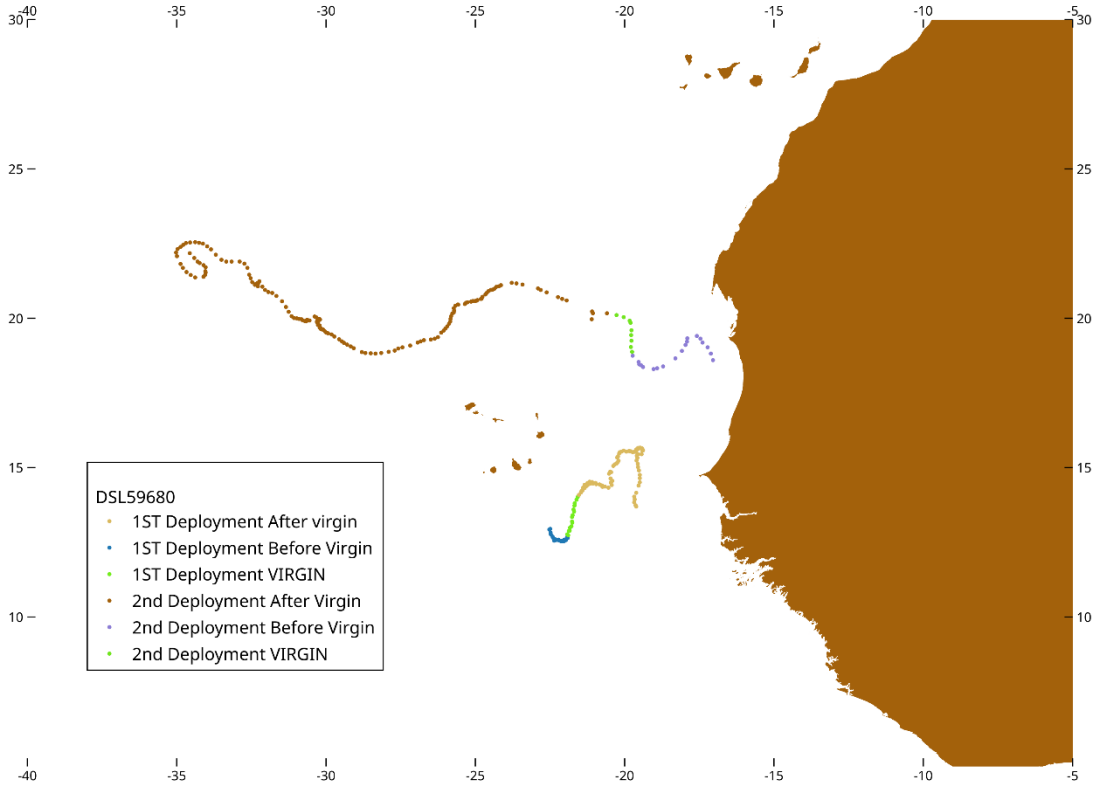
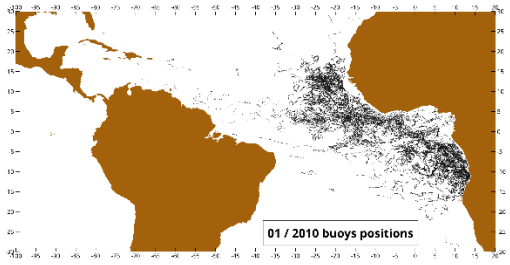
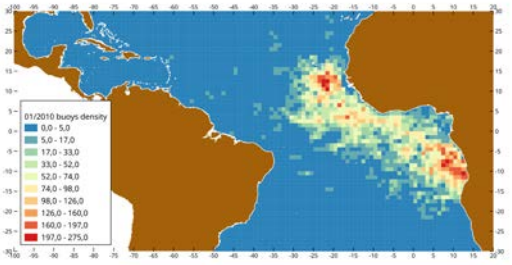


Figure 1. Example of "virgin" segments used for the calculation of the BAI index. Trajectories correspond to buoy DSL59680 with two different paths representing drifts of different FADs. A virgin segment is defined as the segment of a buoy trajectory whose associated FAD likely represents a new deployment, which has been potentially colonized by tuna and not already fished. We consider as virgin segments (i.e. when tuna has aggregated to DFAD) those segments of trajectories from 20-35 days at sea. "Virgin" segments are shown in green in the Figure.

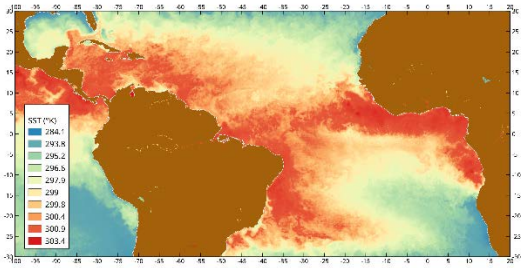
Trajectories



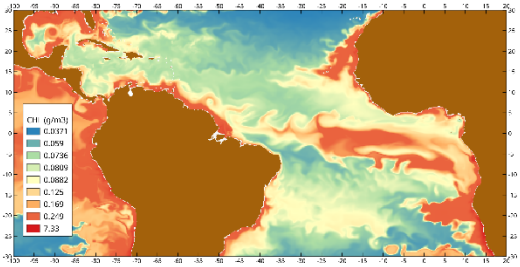
Densities



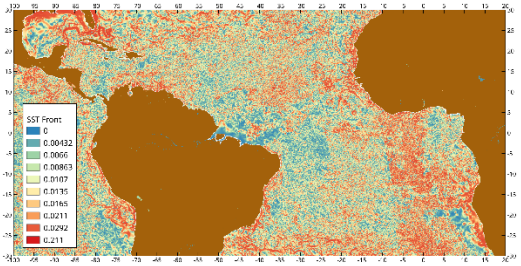
Sea surface temperature (SST)



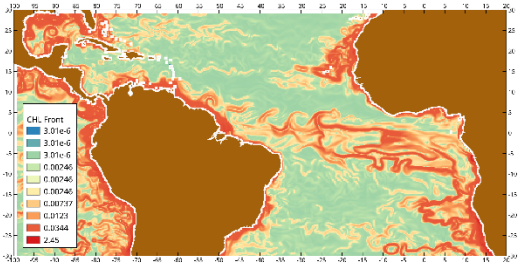
Chlorophyll concentration



SST fronts



Chlorophyll fronts



Mixed layer height

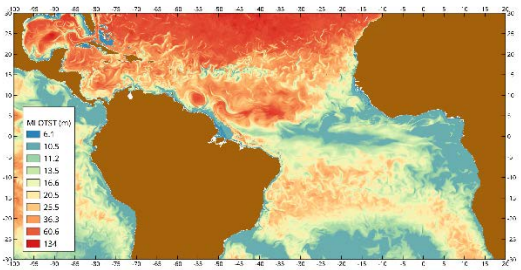


Figure 2. Example of one month of data, January 2010, of buoy trajectories, density of buoys by 1ºx1º and environmental variables incorporated in the GLM analysis: sea surface temperature (SST), chlorophyll concentration, detected fronts in SST, detected fronts in chlorophyll and mixed layer height.

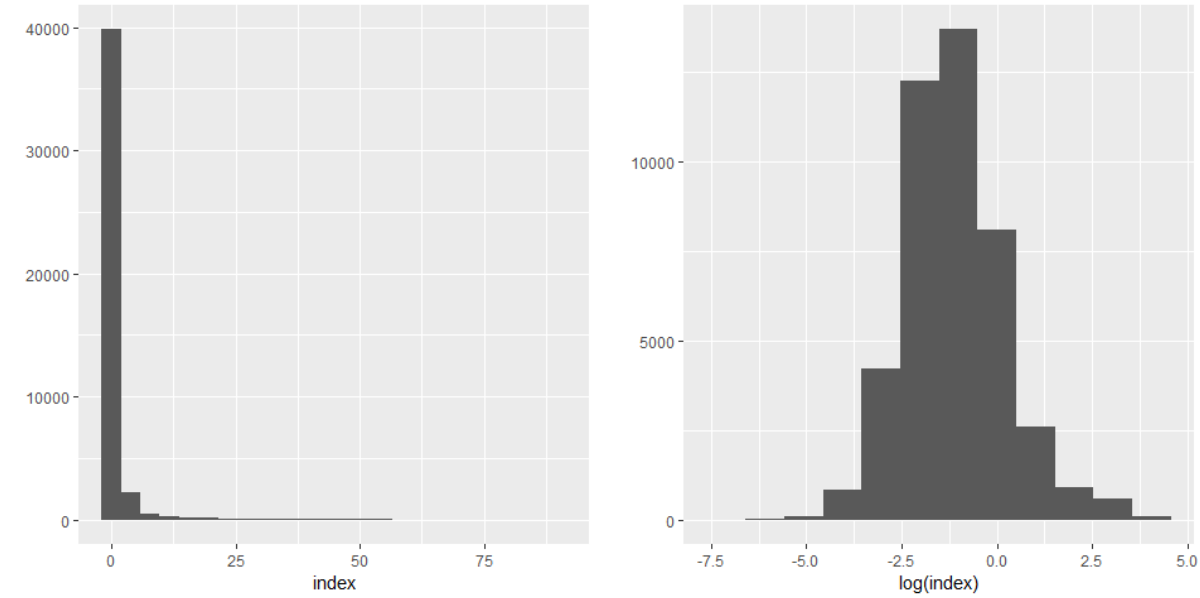


Figure 3. Histograms of the nominal values (left) and the log transformed nominal values (right) of the Buoy-derived Abundance Index (0.9 quantile of the integrated acoustic energy observations in "virgin" sequences).

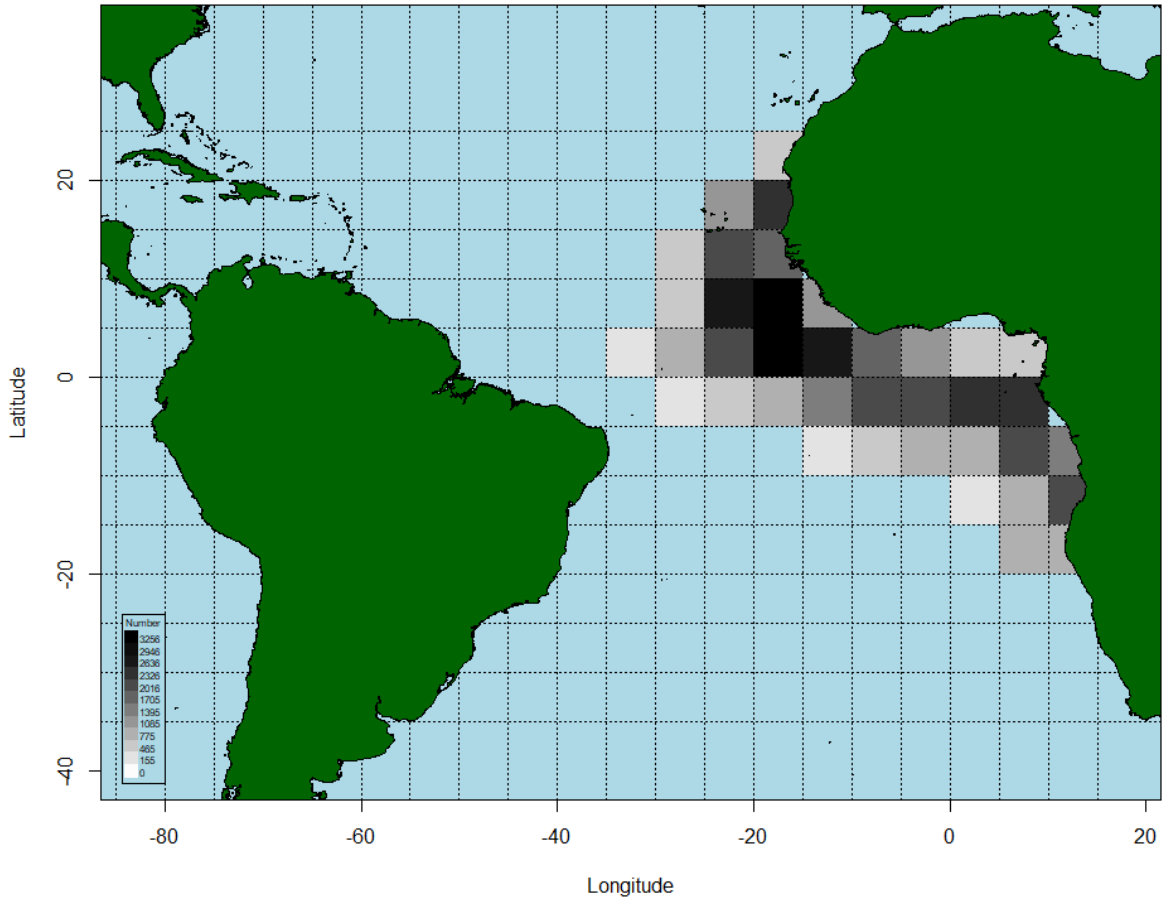


Figure 4. Spatial distribution [5ºx5º] of the "virgin" sequences of buoy trajectories that have been used in the GLM analysis.

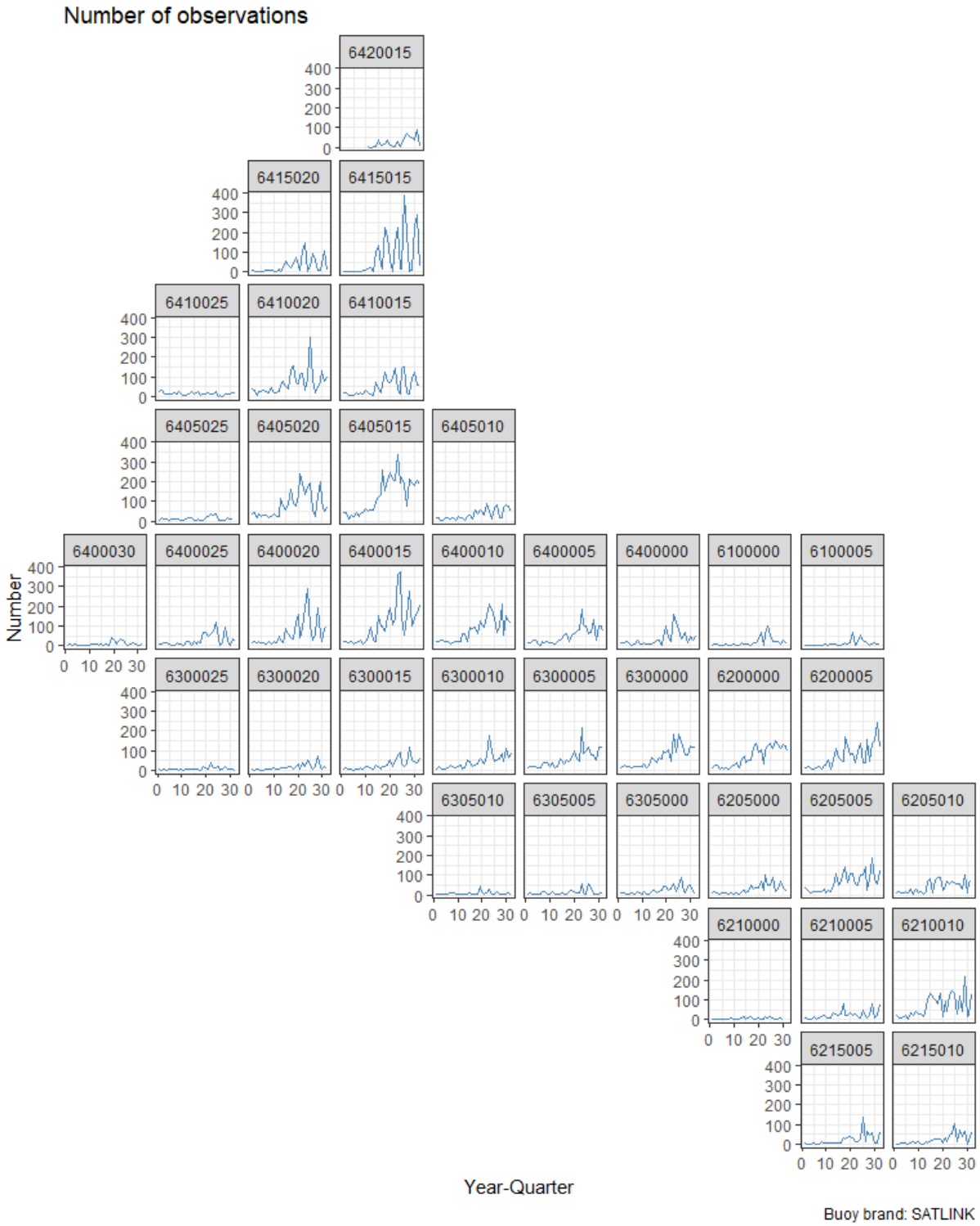


Figure 5. Quarterly evolution of the number of observations ("virgin" sequences of buoy trajectories) on a 5ºx5º grid.

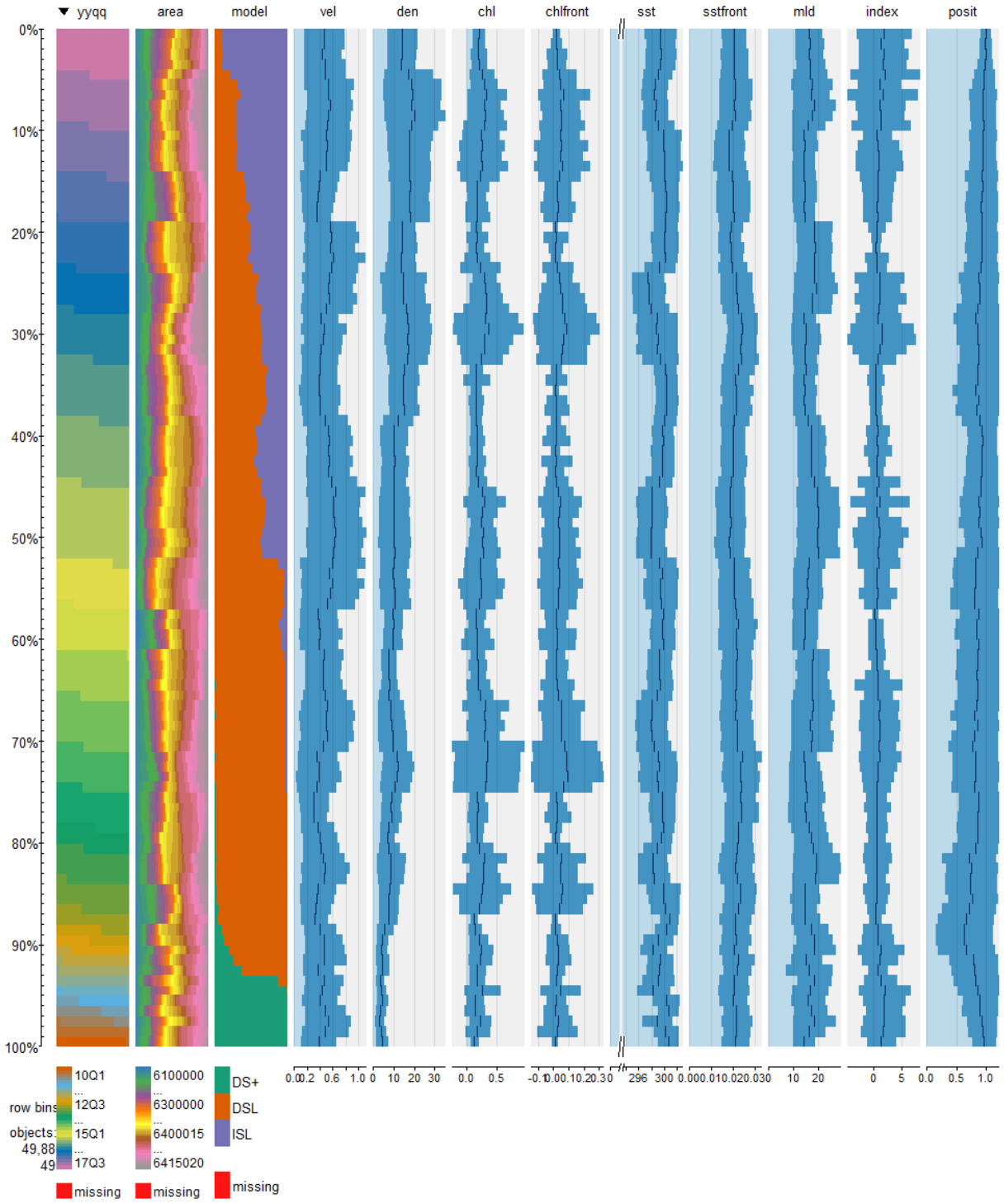
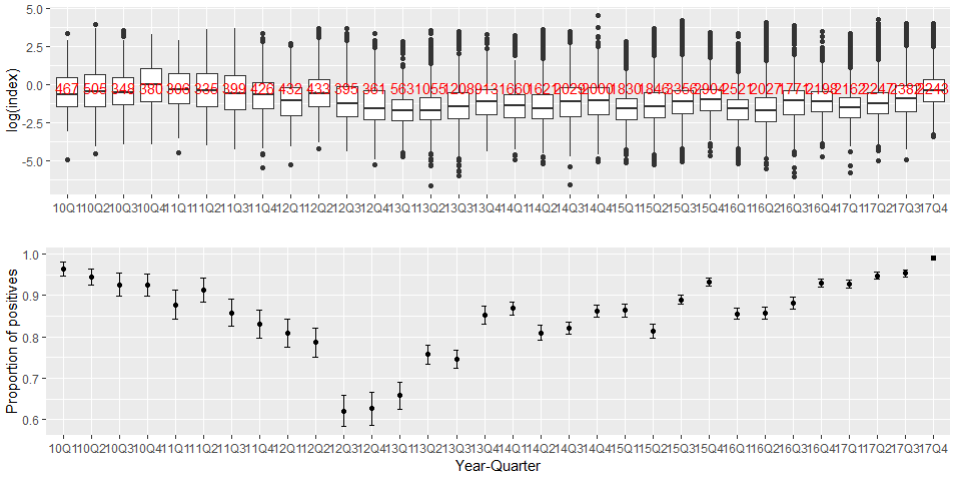
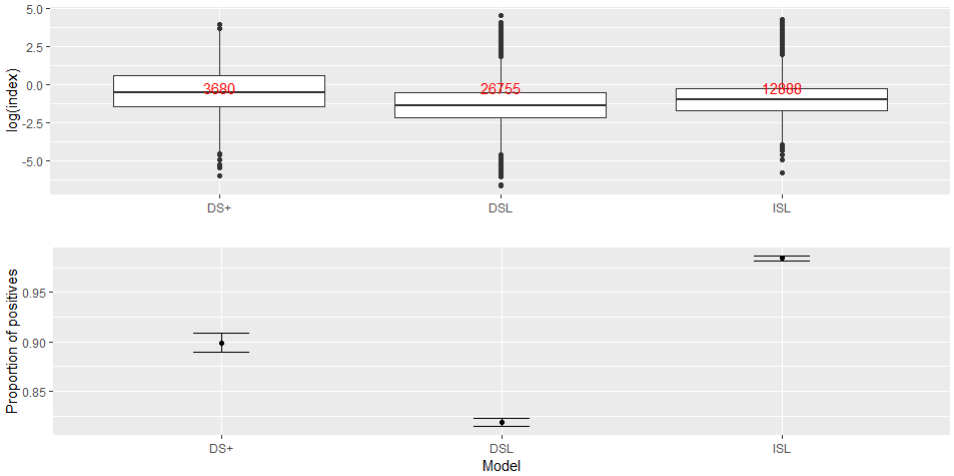


Figure 6. Tableplot of the variables used in the GLM analysis. Each column represents a variable and each row bin is an aggregate of a certain number of records. For numeric variables, a bar chart of the mean values is depicted. For categorical variables, a stacked bar chart is depicted of the proportions of categories. Independent variables tested in the GLM were year-quarter (yyqq), 5ºx5º area (area), buoy model (model), buoy velocity (vel), FAD density (den), chlorophyll concentration (chl), detected fronts in chlorophyll (chlfront), sea surface temperature (SST), detected fronts in SST (sstfront) and mixed layer height (mid). The dependent variable for the lognormal component was the 0.9 quantile of the integrated acoustic energy observations in the "virgin" sequence (index); and for the binomial component, the proportion of positives (posit).

A) Year-quarter



B) Buoy model



C) Buoy speed

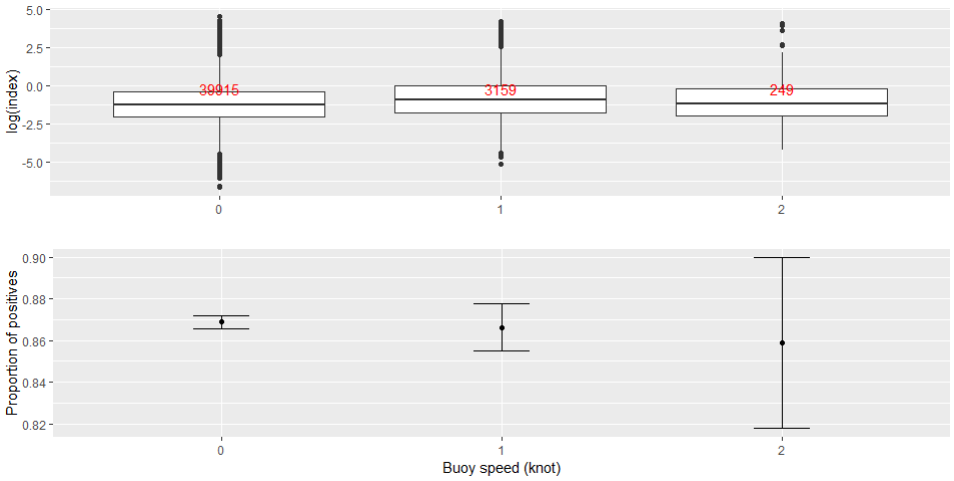
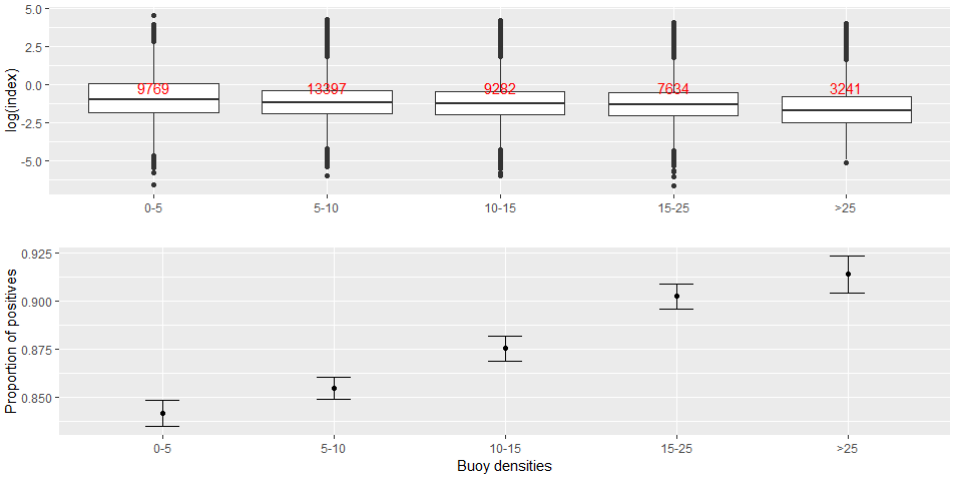
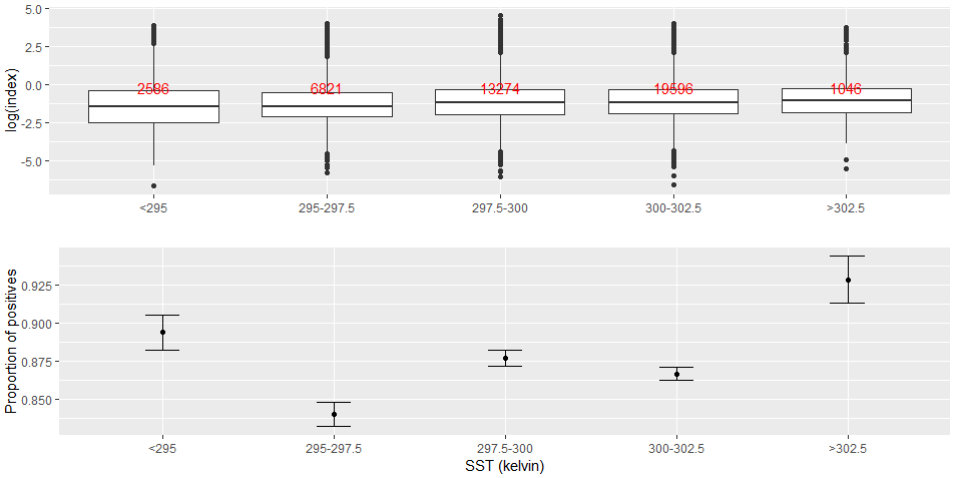


Figure 7. Boxplot of log (BAI) and proportion of BAI positives for year-quarter, buoy model and buoy speed (expressed as categorical). Number of observations for each categorical value is shown in red.

A) Buoy densities



B) Sea surface temperature



C) Chlorophyll concentration

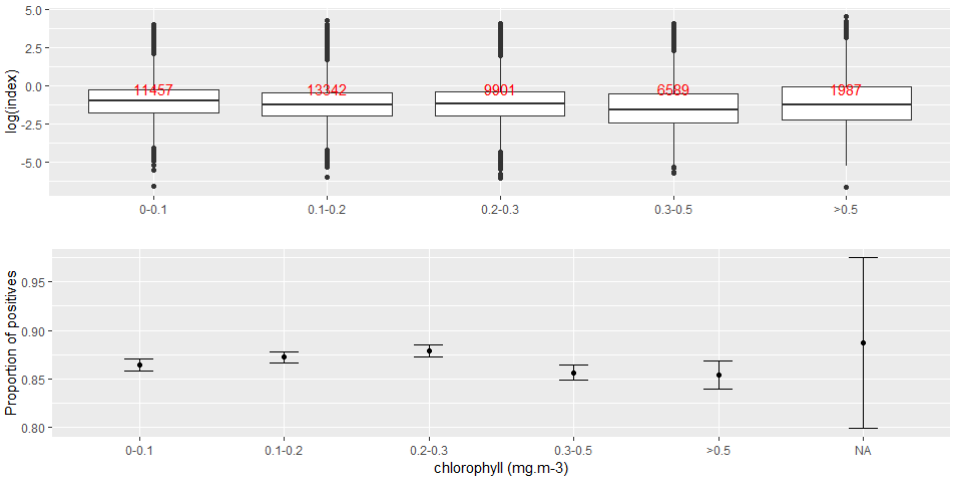
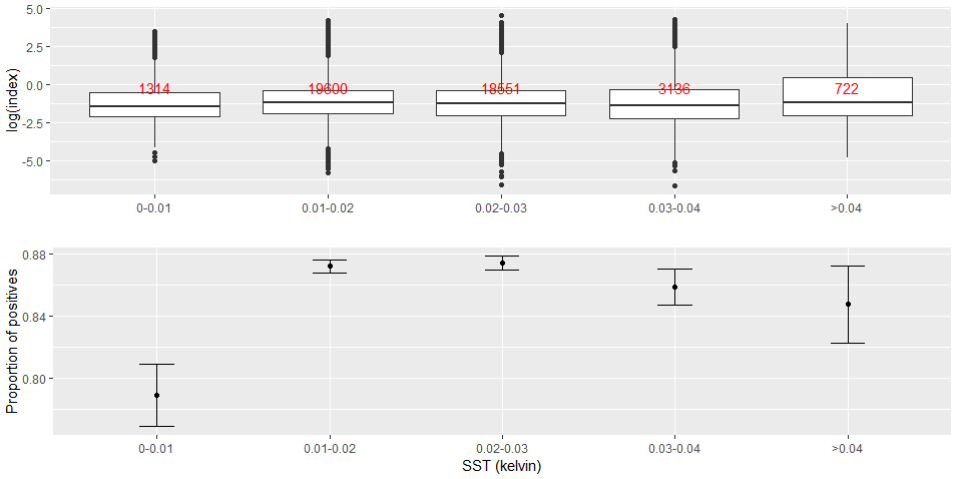
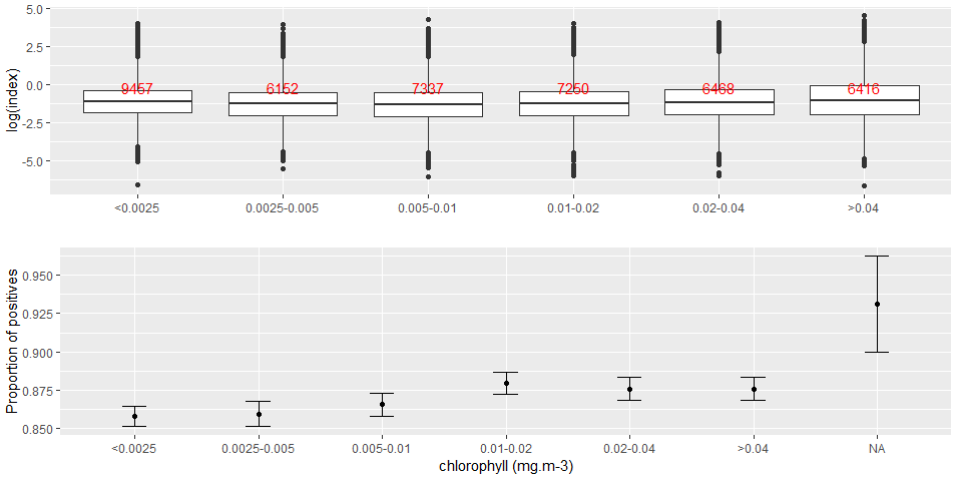


Figure 8. Boxplot of log (BAI) and proportion of BAI positives for buoy densities, sea surface temperature and chlorophyll concentration (expressed as categorical). Number of observations for each categorical value is shown in red.

A) Fronts of sea surface temperature



B) Fronts of chlorophyll concentration



C) Mixed layer height

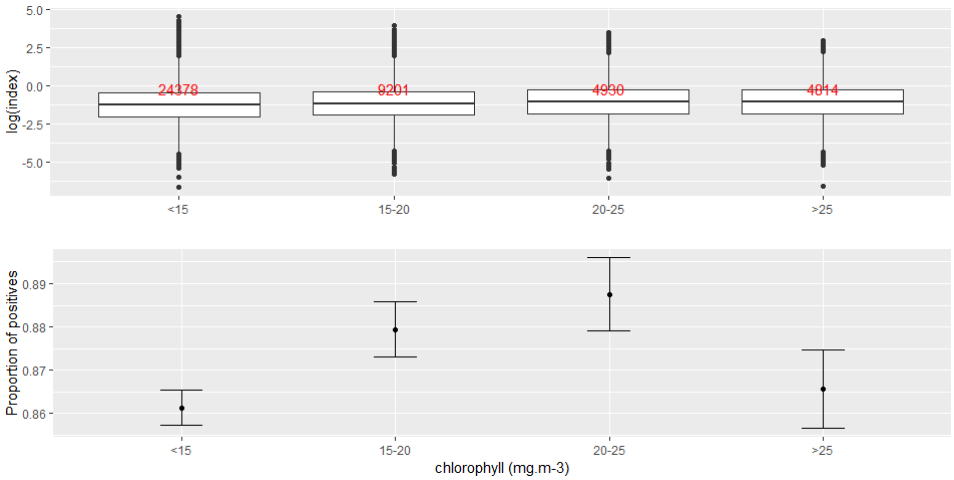


Figure 9. Boxplot of log (BAI) and proportion of BAI positives for fronts of sea surface temperature, fronts of chlorophyll concentration and mixed layer height (expressed as categorical). Number of observations for each categorical value is shown in red.

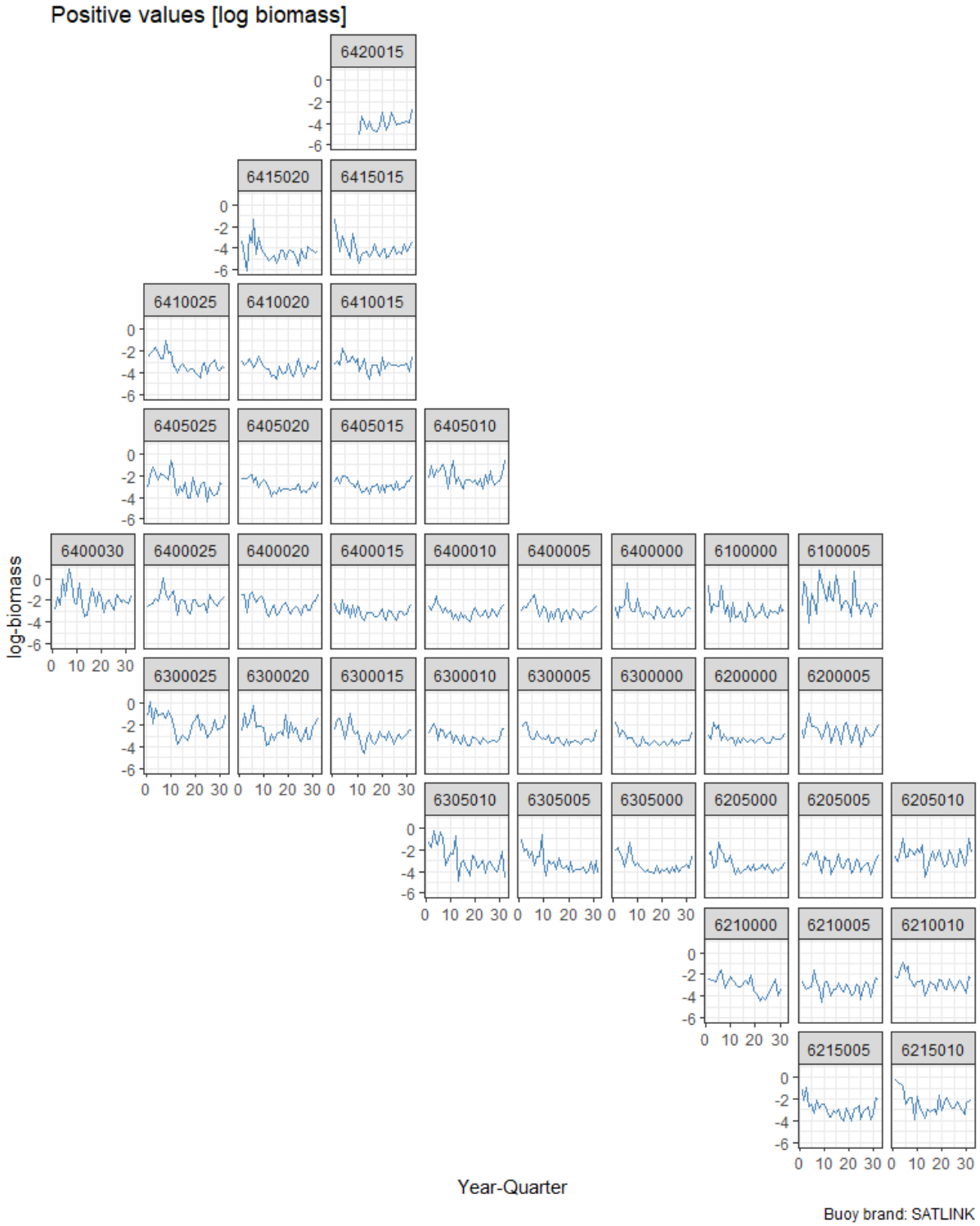


Figure 10. Quarterly evolution of the log BAI index in the Atlantic Ocean by squares of 5x5 degrees from 2010 to 2017.

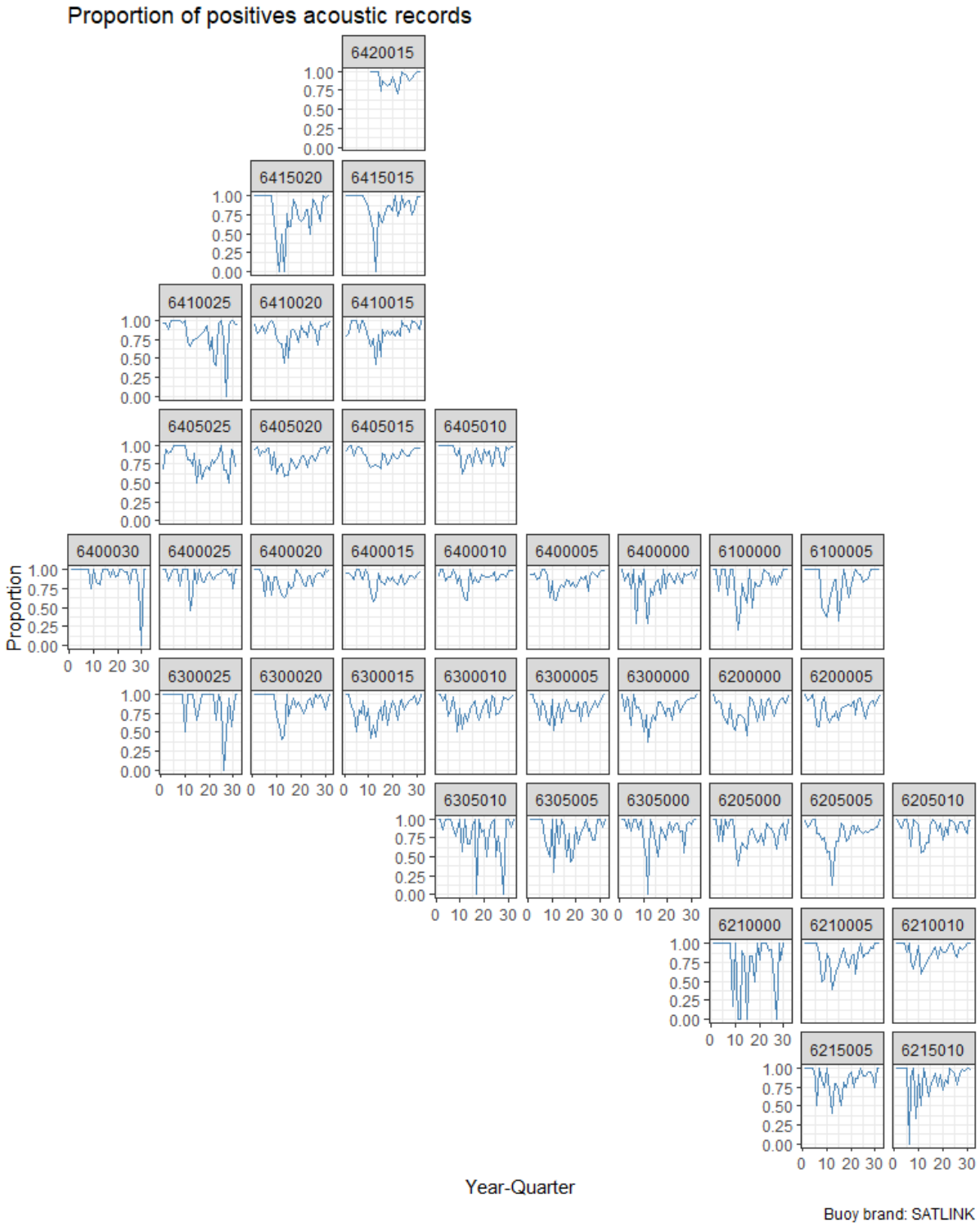


Figure 11. Quarterly evolution of the proportion of BAI positives in the Atlantic Ocean by squares of 5x5 degrees from 2010 to 2017.

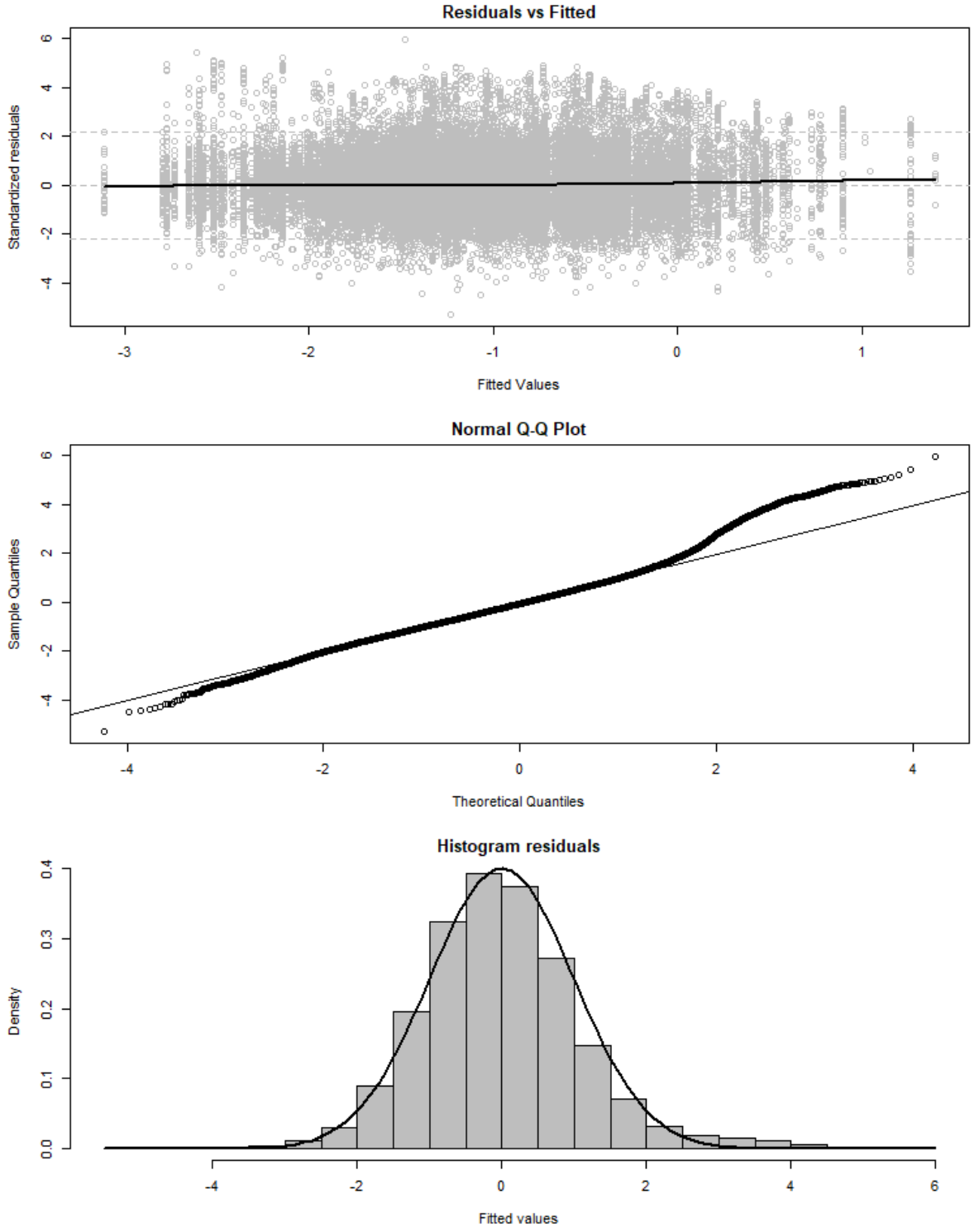


Figure 12. Diagnostics of the lognormal model selected for the period 2010-2017: residuals vs fitted, Normal Q-Q plot and frequency distributions of the residuals.

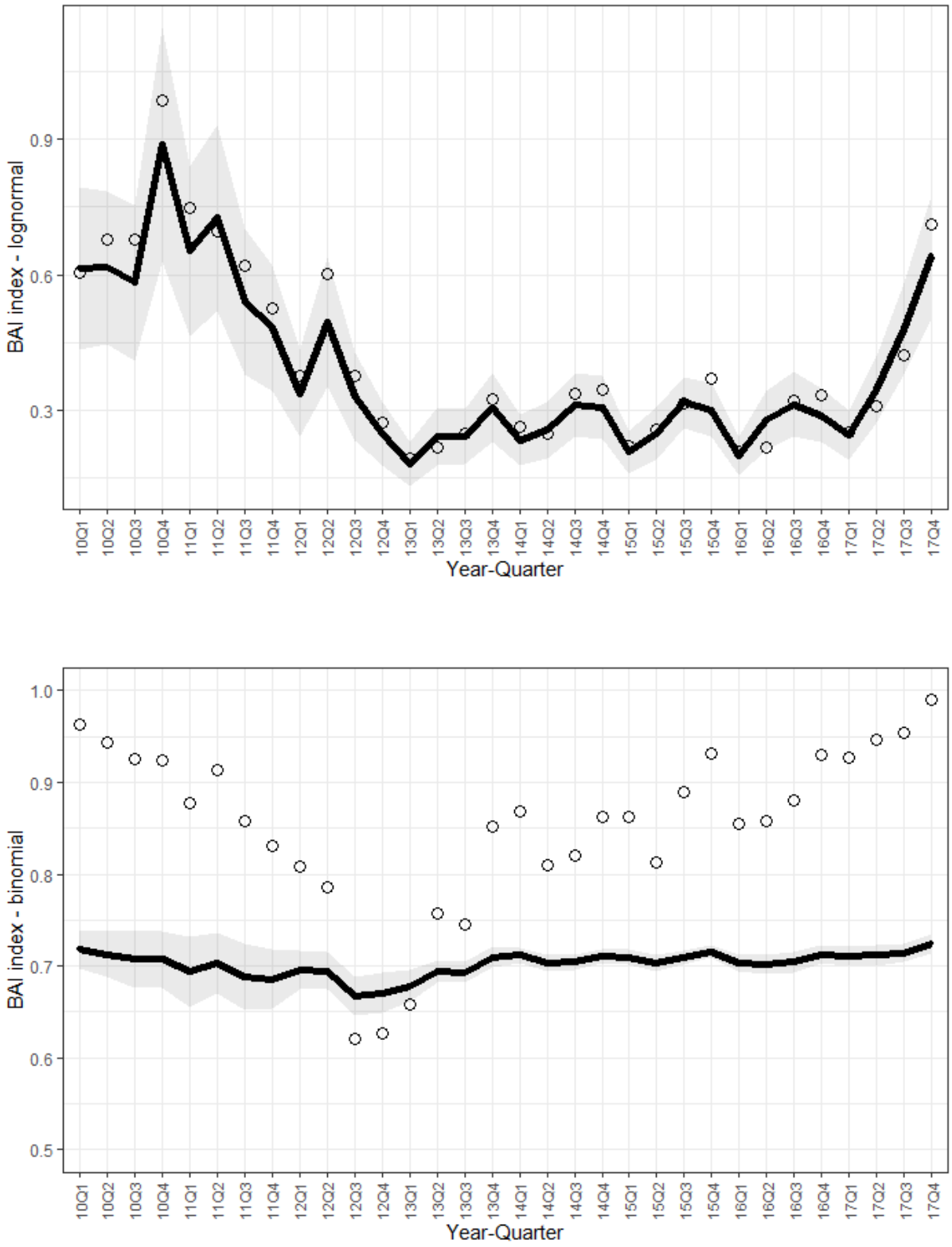


Figure 13. Time series of nominal (circles) and standardized (continuous line) values of the lognormal (upper panel) and binomial (lower panel) components of the Buoy-derived Abundance Index for the period 2010-2017. The 95% upper and lower confidence intervals of the standardized components of the BAI index are shown.

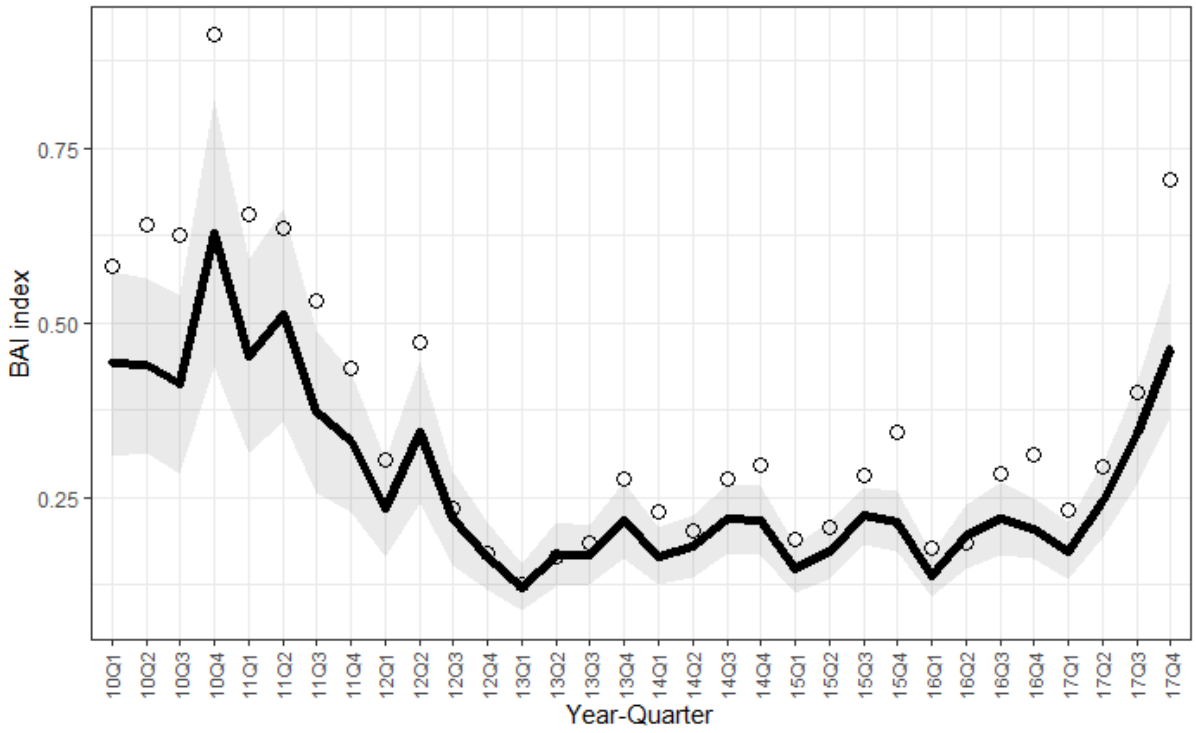


Figure 14. Time series of nominal (circles) and standardized (continuous line) Buoy-derived Abundance Index for the period 2010-2017. The 95% upper and lower confidence intervals of the standardized BAI index are shown.