Seasonal and spatial variability in the vertical distribution of pelagic forage fauna in the southwest Pacific

Aurore Receveur, Christophe Menkes, Valérie Allain, Anne Lebourges-Dhaussy, David Nerini, Morgan Mangeas, Frédéric Ménard

Seasonal and spatial variability in the vertical distribution of pelagic forage fauna in the southwest Pacific

Aurore Receveur[a], Christophe Menkes[b], Valérie Allain[a], Anne Lebourges-Dhaussy[c], David Nerini[d], Morgan Mangeas[e], Frédéric Ménard[d].

a - OFP/FEMA, Pacific Community, 95 Promenade Roger Laroque, BP D5, 98848 Noumea, New Caledonia

b - ENTROPIE, UMR 9220, IRD, Univ. de la Réunion, CNRS, 101 Promenade Roger Laroque, 98800 Noumea, New Caledonia

c - IRD, Univ. Brest, CNRS, Ifremer, LEMAR, Campus Ifremer, BP70, 29280 Plouzané, France

d - Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, 13288 Marseille, France

e - ESPACE DEV, IRD, Univ. La Réunion, Univ. Guyane, Univ. des Antilles, Univ. Montpellier, 101 Promenade Roger Laroque, 98800 Noumea, New Caledonia

Contributions

A.R analyzed the data and interpreted the results with the help of all authors.
A.R. performed the statistical analysis with the help of M.M. and D.N.
A.R., C.M. and V.A. wrote the manuscript with the help of F.M. and A.L.D.
All authors contributed to and provided feedback on various drafts of the paper.

Corresponding authors

Correspondance to Aurore Receveur, aurorer@spc.int, 95 Promenade Roger Laroque, BP D5, 98848 Noumea, New Caledonia

Abstract

Acoustic data are an invaluable source of information for characterizing the distribution and abundance of mid-trophic level organisms (MTLOs) in the ocean. These organisms play a key role in the ecosystem as prey of top predators and as predators of lower trophic level organisms, as well as in carbon export from the surface into deeper waters. This study used 38 kHz-EK60 acoustic echo sounder data from six cruises spanning 2011-2017 to explore the seasonal and spatial variability in the vertical distribution of MTLOs' from 10-600m in the New Caledonian (South Pacific) Exclusive Economic Zone. A total of 16715 acoustic vertical profiles of acoustic backscattering strength were clustered into homogeneous groups. Two small shallow scattering layers (SSLs) between 0 and 100m, and one large deep scattering layer (DSL) at around 550m depth characterized the mean vertical distribution of MTLOs. A machine-learning model (eXtreme Gradient tree Boosting algorithm, XGBoost) was fitted to explain the acoustic profile clusters with environmental variables as predictors. Sun inclination was the most important factor in structuring the vertical profile shapes due to the diel vertical migration signal, followed by the mean oxygen value of the top 600m. Bathymetry, euphotic depth, 0-600m mean temperature and SST were the next most significant variables. Isotherm depth, surface chlorophyll-*a*, wind, and mean salinity had a lower influence on the shape of the vertical profiles. The model was then used to construct vertical echograms at the scale of the New Caledonian EEZ, showing an accuracy up to 87% in cross validation. Across the EEZ, the shape of vertical acoustic profiles were comparable, though layer echo intensities varied spatially with a marked north-south gradient that remained relatively constant seasonally. The vertically-averaged acoustic values were characterized by a maximum to the south of the EEZ in summer, mainly driven by high oxygen values as well as shallow euphotic depth. We also estimated a migrant proportion between day DSL and night SSL of about 78%. Our methodology offers a promising approach for analyzing the control of the environment on the vertical distribution of MTLOs for other oceanic provinces, while also providing a framework to investigate the corresponding trophic interactions between MTLOs and their predators feeding at different depths and times. Moreover, our findings stress the need to consolidate knowledge on species composition in order to optimize acoustic data interpretation.

Key words*: Micronekton, echo sounder, Pacific Ocean, mesopelagic zone, sound scattering layer, environment*

1. Introduction

In pelagic ecosystems, mid-trophic level organisms (MTLOs), also referred to as micronekton, are composed of crustaceans, molluscs, gelatinous organisms and fish with size ranging from 1 to 20cm long (Bertrand et al., 2002; Young et al., 2015). MTLOs play an important role as intermediate components between lower trophic levels (phytoplankton and zooplankton) and predators, including commercially targeted fish species (Bertrand et al., 2002; Duffy et al., 2017; Olson et al., 2014) as well as emblematic endangered marine species (Lambert et al., 2014; Miller et al., 2018). The feeding habitats and vertical behaviors of predators through the water column are very diverse (e.g. Benoit-Bird and McManus, 2012; Choy et al., 2017). Moreover, the habitat depth range of a specific predator may change spatially, as a function of prey distribution or due to physiological tolerance to environmental parameters (Houssard et al., 2017; Schaefer and Fuller, 2010, 2007). Ecosystem Based Fishery Management (EBFM) aims to develop relevant knowledge on ecological mechanisms and processes that shape such predator-prey interactions (Christensen et al., 1996; Koslow, 2009). To date, most of the studies assessing the influence of prey distribution on predator distribution at regional scales have used ecosystem models (e.g. Lambert et al., 2014; Miller et al., 2018). Observations and data on the vertical distribution of prey are still lacking, although they could greatly contribute to calibrate state-of-the-art ecosystem models that inform EBFM (Fulton et al., 2005; Lehodey et al., 2010; Maury, 2010; Pauly et al., 2000; Shin and Cury, 2001).

MTLOs are usually aggregated into layers, which are present in all ocean basins between the surface and 2000m depth (Opdal et al., 2008). The thickness of a single layer ranges from a few meters to tens of meters, and the layer can horizontally spread over hundreds of kilometers (Benoit-Bird et al., 2017). Layers of MTLOs residing in the epipelagic zone (0-200m) are referred to as shallow scattering layers (SSLs) and those in the mesopelagic zone (200-1000m) as deep scattering layers (DSLs). The aggregation of pelagic organisms into scattering layers is a highly organized process of many individuals reacting to predation pressure as well as to environmental resources such as food availability, temperature, or oxygen concentration (Benoit-Bird et al., 2017; Cade and Benoit-Bird, 2015; Ritz et al., 2011). Scattering layer characteristics (depth, echo intensity, composition and number of layers) vary geographically and seasonally (e.g. Escobar-Flores et al., 2018a). DSL depth has been linked to various environmental variables such as seawater density (Godo et al., 2012) or oxygen concentration (Bianchi et al., 2013a; Klevjer et al., 2016). Primary production and sea

temperature also affect the DSL and SSL echo intensity (Escobar-Flores et al., 2013; Irigoien et al., 2014), and DSLs are sometimes split into more than one layer comprising different species (Ariza et al., 2016a; Benoit-Bird and Au, 2004).

Day DSLs and night SSLs are connected through diel vertical migration (DVM), a well-known phenomenon observed at the global scale (Bianchi and Mislan, 2016; Klevjer et al., 2016). recognized as the world's largest animal migration (Hays, 2003). DVM patterns relate to population-wide movements in the water column, with ascents and descents of a large proportion of the MTLOs from the mesopelagic zone, where they remain during day time, toward the more productive epipelagic zone (0-200m) where they feed during the night (Pearre, 2003). By migrating between surface and deep waters, MTLOs actively contribute to the downward flux of nutrients and particulate organic matter via their respiration and excretion processes (Ariza et al., 2015; Drazen and Sutton, 2017). Quantifying the proportion of MTLOs performing DVM and identifying the environmental drivers can thus contribute to a better understanding of the overall role of DVM in the global carbon cycle (Aumont et al., 2018; Belcher et al., 2019).

Data from scientific calibrated echo sounders can provide a proxy of the vertical distribution of SSLs and DSLs (Kloser et al., 2002). Single-frequency acoustic data from echo sounders at 38 kHz and lower frequencies can typically describe both SSLs and DSLs down to 1000m, encompassing the entire DVM. Because they vary widely in two dimensions (depth and time/distance), echograms are complex to analyze in relation to a multivariate environment. Most studies simplify information contained in the depth profile of an echogram through few metrics and analyze them through time together with environmental variables. For instance, acoustic backscatter has been studied using invariant depth-averaged vertical layers (e.g. 0-200m and 200-1000m) (*e.g.* Bedford et al., 2015; Behagle et al., 2014; Doray et al., 2009). Other studies extracted schools or layers and studied these layers' depth, thickness and echo intensity (Burgos and Horne, 2008; Proud et al., 2018a). These methods provide information on the layer echo intensity variability but not on the vertical structure variability. Behagle et al. (2016) and Boersch-Supan et al. (2017) classified vertical acoustic profiles but they did not test the influence of environmental parameters on their classification results.

To author's knowledge, there is not any robust method that statistically links the complete vertical distribution of scattering layers to environmental variables. We attempted to fill part of this gap by designing a method to link the vertical distribution of MTLOs to oceanographic

conditions; and using this method to predict vertical distributions in un-sampled areas with similar environmental conditions. We used acoustic vertical profiles as sampling units from six cruises, and classified them into clusters to first describe the main vertical profile modes. We then modeled, with a machine learning algorithm, clusters as a function of environmental variables to understand the main links between oceanographic factors and vertical distribution. We finally predicted vertical echograms and migrant proportion between SSL and DSL at the scale of the New Caledonian EEZ in un-sampled regions where oceanographic data were available.

2. Material and Methods

Our study area fell within the New Caledonian Exclusive Economic Zone (EEZ), a region of more than 1.4 million km². Recent studies have provided an overview of the physical and biological oceanographic context in the New Caledonian EEZ (Ceccarelli et al., 2013; Menkes et al., 2015). Studies specifically focusing on micronekton have explored species richness and diversity in the region, identifying more than 480 MTLO species (e.g. Grandperrin, 1975; Payri et al., 2019), as well as the spatial-temporal distributions of MTLOs averaged in the 20-120m layer (Receveur et al., submitted). We focused on the MTLOs vertical distribution in the present study.

2.1. Acoustic data

We gathered data from six cruises (Nectalis 1-5 referred to as N1 to N5, and Puffalis) on board the R/V Alis in the New Caledonian EEZ, covering the area between 156°E–175°E and 14°S–27°S over the period 2011 to 2017 (Figure 1, Table 1). During the cruises, *in situ* acoustic data were recorded continuously using an EK60 echo sounder (SIMRAD Kongsberg Maritime AS, Horten, Norway) connected to four split-beam transducers at 38, 70, 120 and 200 kHz. EK60 calibration was performed according to Foote et al. (1987) for each cruise. In the present study, we used 38 kHz only. The hull-mounted transducer was 4m below the surface and shallower than 6m below the transducer face was deleted from the records (data collection started at 10m below the surface). The maximum detection range was 800m for all the surveys except for N1 cruise, where the records were limited to <600m depth. For consistency, the analyses were thus limited to 600m.

All raw acoustic data were processed with the open-source Matecho software (Perrot et al., 2018). A first cleaning step removed ghost bottom echoes. Then, four semi-automatic

cleaning filters were applied to: (i) remove acoustic device interference ('un-parasite' Matecho filter), (ii) remove attenuated signals ('white pings' filter), (iii) remove elevated signals ('deep spike' filter) and (iv) reduce background noise (De Robertis and Higginbottom, 2007). Details of filter parameters can be found in Behagle et al. (2016) and Perrot et al. (2018). After data cleaning, the echo-integration was done on cells of 1m-deep and 0.1nm-long, providing volume backscattering strength S_v data ($dB.re.1.m^{-1}$) (hereafter referred to as echo intensity), the linear measure of the volume backscattering strength $\boldsymbol{s_v}$ (m^{-1}) $\left(\boldsymbol{s_v} = 10^{\frac{S_v}{10}}\right)$ and the area backscattering strength S_a ($dB.re.1.m^2m^{-2}$), a proxy for the MTLOs' biomass (Irigoien et al., 2014) (Maclennan et al., 2002) for each cell. We used the linear form $\boldsymbol{s_v}$ when arithmetic operations were necessary.

Vertical profiles were smoothed using a locally polynomial quantile regression (Koenker, 2004) to remove high-frequency peaks (e.g. interferences or very small schools that create peaks in an acoustic profile) that were considered non-interpretable in the present study. Each vertical profile ranging from 10 to 600m was averaged in 4m vertical bins keeping the 0.1nm horizontal resolution. Correlations between consecutive vertical profiles were high for distances ranging from 0.1 to 0.4 nm and decreased after. We then selected one profile out of four to limit autocorrelation effects. The final dataset was composed of 16715 vertical profiles.

2.2. Environmental data

Table A1 displays the environmental variables selected to explore the physical drivers of the MTLOs' vertical distribution. For each vertical profile, environmental data were extracted at the dates and positions of the acoustic samples.

Bathymetry data were extracted from the ZoNéCo database at a 500m spatial resolution (ZoNéCo, 2013). Sun inclination was calculated as a function of spatial position and date, with negative values for nights and positive values for days (Blanc and Wald, 2012; Michalsky, 1988). Twilight periods (*i.e.* dawn and dusk periods) were defined as the periods when sun inclination was in the range -10° and 10°. During these periods, as organisms actively swim up or down due to DVM, their orientations change, creating strong variability in backscatter (McGehee et al., 1998; Zedel et al., 2005), and were removed from the final dataset. Hence, migration vertical profiles are highly changeable.

2.2.1. Inter-annual surface variables

Sea Surface Temperature (SST) was taken from the NOAA OI SST High Resolution Dataset at a daily resolution (Reynolds et al., 2007). The depth of the 20°C isotherm and the surface geostrophic ocean velocity amplitude were extracted from the Armor3D dataset (Guinehut et al., 2012) available at a weekly time scale. The depth of the euphotic zone was extracted from the MERCATOR GLORYS2V4 reanalysis (Garric et al., 2017) at a weekly resolution. Surface wind amplitudes were obtained from Cross-Calibrated Multi-Platform (CCMP-v2, Wentz et al., 2015) datasets at a weekly resolution. Surface chlorophyll-*a* was extracted from GLOBCOLOUR (Saulquin et al., 2009) at a daily resolution. All interannual variables were extracted on a ¼° spatial grid and were included as environmental covariates for modeling acoustic profiles (see section 2.3).

2.2.2. Subsurface datasets

In addition to surface values, Armor3D provided an ocean reanalysis of observed vertical profiles of ocean temperature (T) and salinity (S) (Guinehut et al., 2012). Armord3D was used rather than CTD field data, because CTD casts were only taken at a limited number of sampling stations (156 stations, Figure 1). However we systematically checked the relationship between CTD and co-located Armor3D data (correlation of 0.99 for temperature values and 0.96 for salinity values). We extracted the oxygen (O_2) vertical distribution from the climatological dataset CARS (Ridgway et al., 2002) as inter-annual data are not available. We also checked the relationship between co-located CTD and CARS data and we found a correlation of 0.74. We used the 6-606m monthly averages of seasonal temperature, salinity and oxygen (30m vertical resolution) at a 1/2° spatial resolution as environmental variables for modeling acoustic profiles (see section 2.3).

2.2.3. Water masses

Water masses describe bodies of water with homogenous physical properties, and constitute a synthetic way of understanding the physical oceanography. Water masses can be defined in terms of temperature, salinity (hence density) and oxygen values and have been described in the south west Pacific (Gasparin et al., 2014; Germineaud et al., 2016). We pooled temperature, salinity and oxygen values for all depths in the top 606m as derived from Armor3D and CARS (Figure A1). We then classified data with a k-means algorithm (Hartigan and Wong, 1979) and identified five distinct water masses corresponding to those

identified by Gasparin et al. (2014) and Germineaud et al. (2016) (water masses' full description in Figure A1 and Table A2). The cluster results were transformed into water mass covariates by calculating, for each acoustic profile, the vertical proportion in depth occupied by the corresponding water mass in the water column.

2.3. Statistical methods

Figure 2 displays the schematic framework of the analyses, considering one acoustic vertical profile as the sampling unit (one 'observation' hereafter). We first reduced the vertical dimension by principal component analysis and then classified the acoustic profiles in homogenous groups using their principal coordinates as variables. In the last step, we fitted a machine-learning type model to link vertical profile clusters to environmental variables.

2.3.1. Noise reduction and classification

A Principal Component Analysis (PCA, Jolliffe, 2011) allowed us to reduce the dimensions of observations (Figure 2, left panel, step 1). Vertical acoustic profiles were then grouped using a model-based clustering (MBC) (Figure 2, left panel, step 2). As PCA brings similar observations close, we performed clustering based on the density of observations in PCA space (or similar). Each cluster was centered around points (e.g. the clusters' center) where the point density was the highest in the PCA space (Fraley and Raftery, 2002). We maximized the Bayesian Information Criteria (BIC) (Raftery, 1995) to select the appropriate number of clusters. BIC values as a function of the number of classes were plotted, and we added the BIC values derivative to better identify discontinuities.

2.3.2. Metrics on vertical profiles

Acoustic metrics were calculated using $\boldsymbol{s_v}$ (linearized backscatter) according to Urmy et al. (2012). We calculated the mean backscatter value over the entire vertical profile (called 'density') and a mean depth location calculated by the average sampled depths weighted by their $\boldsymbol{s_v}$ values (called 'center of mass'). We calculated a proxy of the acoustic aggregation rate over the water column: a high value corresponds to high backscatter concentrated over short depth ranges in the vertical profiles (called 'aggregation'). These metrics are detailed in Table 2 (see also Urmy et al. 2012).

2.3.3. Environmental factors driving the acoustic clusters

A "machine-learning" model was fitted to link the acoustic clusters to environmental covariates (Figure 2, left panel, step 3). We used the XGBoost algorithm (eXtreme Gradient tree Boosting), which is an optimized distributed gradient boosting designed to be highly efficient, flexible and portable (Chen and Guestrin, 2016). XGBoost uses machine-learning algorithms under the Gradient Boosting framework. The basic idea is to incrementally create new sub-models that predict the residuals or errors of prior sub-models, and then merge sub-models together to make the final prediction. Gradient boosting uses a gradient descent algorithm to minimize the loss when adding new models. We used cross-validation to tune parameters (Browne and Cudeck, 1989), with the proportion of well-classified observations in the validation dataset as the criterion. The model was first fitted on a training dataset (75% of randomly selected profiles) and then tested on a validation dataset (the remaining 25% of data). To prevent overfitting, XGBoost parameters were set equal to 0.3 for the learning rate η and to six maximum tree depths. Environmental variables listed in section 2.2 were included as covariates.

To rank the importance of covariates, SHapley Additive exPlanation (SHAP) values were computed (Lundberg et al., 2018) for the overall model and for each cluster. SHAP values indicated how much a given covariate value could change the predicted value compared to the prediction done without this covariate (Lundberg and Lee, 2017). For instance, a high SHAP value for a given covariate value indicates a strong significance in the prediction. For a given prediction, the difference between the value predicted by the model (e.g. the set of probabilities to be in each acoustic cluster) and the predicted value without one covariate was calculated. To take into account the integration order of the remaining covariates, all possible orders of covariate inputs were tested for predictions. Then, all differences were added to calculate SHAP values. Following the same process, SHAP values were calculated for other covariates. By averaging SHAP values by covariate across all the observations, we could rank the explanatory variables in the final model.

In the same way, SHAP values could be averaged by covariates across groups of observations (for example acoustic clusters) to determine the importance of each explanatory variable for those groups. To visualize the importance of variables by cluster in the predicted observations, we first normalized and centered the covariates. Then we plotted, by acoustic cluster and for each covariate, SHAP values for each predicted value associated with that covariate, with color coding for the normalized covariate value (green to yellow, see Figure 8 and section 3.2 for a complete interpretation).

2.3.4. Model predictions

We next used the model as a predictive tool (Figure 2, right panel) based on the climatology of all explanatory variables computed in a given spatial cell (latitude and longitude resolution of ¼°). The distribution of prediction dataset values were similar to observation dataset values (Figure A2). For each spatial cell, the model predicted the probability of belonging to each acoustic cluster (Figure 2, right panel). Two alternative approaches were considered to finally allocate one acoustic profile per cell: 1) we selected the acoustic cluster with the highest probability; and 2) we calculated each mean vertical profile by cluster, and then we averaged the mean vertical profiles of clusters weighted by their predicted probability (Figure 2, right).

The second option allowed us to predict acoustic values for all EEZ cells by month, during day and night, and at each depth. We then estimated echograms for the whole EEZ and produced maps of the integrated 10-600m acoustic value by season. Finally, we quantified the proportion of migrant MTLOs (%) with:

$$Mp = \frac{s_v^N - s_v^D}{s_v^N}, \qquad \text{Eq. (1)}$$

with Mp representing the proportion of migrants, $\boldsymbol{s_v^N}$ the mean $\boldsymbol{s_v}$ for a given vertical layer (e.g. 10-200m) during the night in m^{-1} and $\boldsymbol{s_v^D}$ the mean $\boldsymbol{s_v}$ for the same layer during the day.

Statistical analyses were performed using R (Core Team, 2018) version 3.5.0. Classification was carried out using the library "`mclust`" (Scrucca et al., 2016) with the 'VVV' option. Extreme gradient boosting tree was carried out with the "`xgboost`" package (Chen et al., 2018).

3. Results

The six cruises provided a dataset covering the two mains seasons as well as most of the New Caledonian EEZ. Cruises N2, N3 and N5 were carried out during the warm season (December-May), and N1 and N4 during the cold season (June-November) (Table 1). The New Caledonian EEZ was reasonably well sampled, with cruise tracks for N1 and N2 covering the northern region, N3 the west, N4 the south-west, N5 the south-east, and Puffalis close to the coast (Figure 1). The full dataset encompassed more than 17500km (e.g. about 9500nm), including 16715 vertical profiles each with 146 depth points in the 10-600m depth range.

3.1. Main patterns of MTLOs vertical distribution

The first two axes of the PCA accounted for 64.3% of the variability and revealed two high density regions of acoustic profiles well separated on the first axis. Within these regions a secondary maximum appeared, separated on the second axis (Figure 3A). The first axis split night from day profiles (50%), and the second axis split the vertical profiles geographically: those located in the north from those located in the south of the EEZ (14.3%). The cumulative variance explained by the axes increased relatively quickly (Figure 3B). We thus kept the first 11 PCA components for the MBC classification step, as these 11 dimensions (instead of the initial 146 depths) contributed to 90% of the vertical profile shapes. Based on the BIC curve and its derivative function, 10 clusters were chosen rather than two or four which are the three first peaks highlighted by the derivative curve. Two or four clusters appeared too low to correctly represent the high diversity of acoustic vertical profile shapes observed among the 16715 observations (Figure 3C). Moreover, the BIC increased quickly between one and 10 clusters, after which the rate of increase was smaller. Finally, 10 clusters allowed us to keep the number of clusters interpretable.

Day and night profiles were almost perfectly separated into different acoustic clusters. Six clusters were mainly composed of day profiles (light grey bar on Figure 4A, referred as ‘day group’ hereafter) and four clusters were composed mainly of night profiles (dark grey bar on Figure 4A, referred as ‘night group’ hereafter). The number of acoustic profiles per cluster ranged from 277 to 2065. Cluster 10 contained less than 300 vertical profiles, while the other clusters described frequent features with more than 1000 profiles per cluster. The spatial distribution of the acoustic clusters indicated a north-south separation for both day and night groups (Figure 4B) with clusters 1, 6 (day) and 5 (night) in the north, and clusters 7 (day), 4 and 8 (night) in the south of the EEZ.

Among the six day clusters (Figure 5), we observed persistent detections at 20-80m, which were composed of non-migrant MTLOs staying within the upper 150m zone during the day. DSLs were located between 450 and 600m depth. Cluster 10 displayed an intermediate layer in the 350-400m range. Cluster 9 had the highest density and cluster 6 the lowest, showing a very flat profile indicative of a near empty water column (Table 3). The center of mass of the clusters varied according to the echo intensity of the SSL and DSL: the shallowest center of mass (338.2m) of cluster 2 was due to a strong SSL, while cluster 10 had the deepest mass center (408.5m) due to an intense DSL. For other day clusters, the center of mass varied

between 340 and 385m depth, indicating an almost equivalent ratio between DSL and SSL intensities. Clusters 7 and 10 were the most aggregated clusters (aggregation index greater than 1.5). Indeed, these two clusters showed a narrower DSL than the other clusters. Cluster 9 showed a more gradual change in DSL intensity than other clusters (Figure 5) and a very small aggregation index (Table 3), indicating a diffuse vertical distribution through the water column. In addition, profiles of cluster 9 were mainly located at the beginning or at the end of transects (Figure 4B).

The shape variability of the night vertical profiles among clusters (Figure 6) was concentrated on the epipelagic zone. Two clusters had two well-marked SSLs (clusters 4 and 8) and other clusters had one high peak only (cluster 3) (Figure 6). Clusters 4 and 8 had the highest densities (Table 3) and were found in the south (Figure 4B). By contrast, cluster 5 had the smallest densities and was mainly found in the north. The deepest center of mass was 187m for cluster 8 and the shallowest was 144m for cluster 3, which had a very intense SSL. Clusters 4 and 8 in the south of the EEZ had, on average, a deeper mass center than other night clusters. Cluster 8 was the most aggregated, and cluster 5 ranked second. Clusters 3 and 4 were less aggregated.

3.2. Environmental influence on the vertical distribution

The relationships between acoustic clusters and environmental covariates were examined using XGBoost modelling (Figure 2, left panel, step 3). Among the 16 explanatory variables, the least important covariates were the proportion of the five water masses and ocean currents. They were removed from the final model as all six together increased the success rate of the model (*i.e.*, the rate of the well-classified profiles from the validation dataset) by only 1%. The success rate of the most parsimonious model reached 87%.

The most influential variable on the profile shape was sun inclination (Figure 7). This major effect was due to the DVM signal. The second most important variable was oxygen followed by bathymetry, 20°C isotherm depth, mean temperature over 6-606m and euphotic depth. Wind, chlorophyll-a, and SST ranked then. Mean salinity was the last one.

As the sun inclination influence is obvious by comparing day and night clusters, we removed it from Figures 8 and 9 to clarify and simplify them. The SHAP values of sun inclination can be found in Figure A3.

For acoustic cluster 1, the significant covariates were bathymetry, mean temperature, euphotic depth and SST (Figure 8). The high values of mean temperature and SST had a strong influence on this cluster (green and yellow colors indicating high temperature values together with high SHAP value), as well as values of deep bathymetry. High oxygen, bathymetry, deep euphotic depth and weak wind contributed significantly to the profiles of cluster 2. For cluster 6, low oxygen and shallow euphotic depth were important, as well as mixed bathymetry values. The shallowest values of bathymetry, relatively deep values of euphotic depth, low chlorophyll-a and high oxygen shaped vertical profiles of cluster 7. Profiles of cluster 9 were mainly influenced by intermediate values of sun inclination (Appendix A3) as well as low mean oxygen, bathymetry, deep euphotic depth, and strong wind. Finally, profiles of cluster 10 were mostly influenced by deep 20°C isotherm depth.

For the night group, oxygen level influenced all clusters except cluster 8, with low oxygen values for cluster 5, and high oxygen values for clusters 3 and 4 (Figure 9). The bathymetry shaped all clusters except cluster 9 with a large diversity of values. Deep 20°C isotherm depth acted on cluster 3. Quite warm temperatures were important for cluster 5 and extremely cold temperatures for clusters 3 and 4. The deep euphotic depth impacted cluster 8 and strong winds clusters 5 and 8. Very low chlorophyll-a concentration had a strong influence on cluster 8 and relatively high chlorophyll-a concentration on cluster 5. Quite warm SST drove cluster 4. Finally, mean salinity did not influence any cluster.

Generally, low oxygen values influenced clusters in the north of the EEZ (clusters 1, 5, 6), and high oxygen and low chlorophyll-*a* were significant for clusters in the south (clusters 2, 4, 7 and 8). Bathymetry and oxygen influenced almost all clusters, while the impact of other covariates was more variable among clusters.

3.3. Prediction at a larger spatial scale

3.3.1. Acoustic regionalization

Figure 10 displays the spatial distribution of the most probable acoustic clusters by season and by day and night. Day distribution was patchier than night distribution. During the day, cluster 1 dominated in the north during the warm season with a southward extension of its spatial range during the cold season. Cluster 2 occurred in the south during the two seasons. The southwest corner was partially invaded by cluster 7 during the warm season. Cluster 10 was present during the two seasons, but scattered in isolated patches in the south. At night, cluster

5 dominated in the north of the EEZ, and cluster 4 in the south. The cluster distribution patterns were very similar during both seasons; however, a small southward extension of cluster 5 was predicted during the warm season. Cluster 3 and 8 patches were present in the south of Bellona (see Figure 1 for location) during the warm season.

3.3.2. Vertical predictions of MTLO distributions

Following section 2.3.4, for each spatial cell, we calculated an average vertical acoustic profile by weighting each mean clusters' vertical profile by the probability of cluster occurrence predicted by the model. After estimating the quality of predictions for a given transect, we predicted acoustic vertical profiles at the scale of the New Caledonian EEZ by month. Then we averaged values for the 10-600m vertical layer by season. In addition, we quantified the proportion of migrant MTLOs in the 10-200m layer using Eq. (1).

For illustration, we selected the track of N4 that encompassed 1034 observations. A visual comparison of the predicted reconstructed echogram versus the observed echogram indicated that the method could reproduce the main patterns of the observed echogram (Figures 11A and 11B). Dynamics of some small layers were replicated, as in box (2) where the shallowest SSL became more intense, or in box (3) where the shallowest SSL connected with the deepest SSL. However, some other features were not well reproduced, as for box (1) or (5) where predicted values did not replicate observed changes. Finally in box (4), there was an observed intensification of the deepest SSL whereas the model predicted an intensification of the shallowest SSL. The high correlation between observed and predicted S_v values pooled for all depths for the N4 cruise (Figure 11C, correlation = 0.88, p-value < 0.0001) indicated again that the methodology could be used to predict echograms in non-sampled areas if the range of environmental variables was similar to sampled data.

By averaging predicted acoustic values in the whole water column (10-600m), we proposed an integrated view of the spatial and seasonal variations of the MTLOs' distribution (see Figure 12). The mean backscatter maximum always occurred in the south of the EEZ, extending toward the north during the cold season.

The proportion of MTLOs migrating within the epipelagic (10-200m) during the night showed a larger part of migrant population below 20°S, especially in the southeast (Figure 12B). Migrant proportion varied spatially spanning a range from 75% in the north to 85% in the south with mean values around 78%.

4. Discussion

By analyzing six cruises of EK60 vertical profiles, we provide new insights into the spatiotemporal variability in the vertical distribution of MTLOs in the New Caledonian EEZ. We proposed a statistical framework to link MTLO vertical distributions to oceanographic conditions. This framework allowed us to predict acoustic vertical distribution with some success in un-sampled areas. Sun inclination and the mean oxygen concentration were the main factors driving the acoustic vertical profile shape. Three homogeneous acoustic based regions, which spatial extent seasonally moved, were identified in the New Caledonian EEZ: north of 20°S; south of 20°S and west of 165°E; and south of 20°S and east of 165°E. The northern mean vertical distribution was characterized by weak echo intensities of DSL and SSL and low mean oxygen values as well as warm SST mainly influenced this vertical distribution. The vertical distributions in the southwest corner showed strong DSL and SSL and deep center of gravity; high mean oxygen values and deep euphotic depth drove it. Finally, southeast corner vertical distributions revealed strong SSL, especially between 0 and 50m and was impacted by shallow euphotic depth and high mean oxygen values. Finally, due to this spatial pattern of vertical distributions, there were generally more MTLOs in the south of the New Caledonian EEZ than in the north and the proportion of vertical migrants was about 75%.

4.1. Methodological framework

Behagle et al. (2016) performed classification of acoustic vertical profiles but they did not statistically link clusters to environmental covariates. Proud et al. (2018) developed an innovative approach and classified acoustic layers below 200m to investigate variability in the vertical distribution of MTLOs. Using a 38 kHz global dataset, they identified six spatially-coherent regional clusters using estimated probability distributions of local SSL depth and of echo intensity. They mapped the clusters at a global scale and matched them with Longhurst's provinces (Longhurst, 1995, 2007). In our work, we fitted relationships between mean MTLO vertical distribution and environmental covariates and were able to predict vertical acoustic profiles in un-sampled areas and so at a larger spatial scale than cruise tracks.

Our proposed methodological framework treated the vertical acoustic profiles as the sampling unit. We thus kept all the shape information contained by the profiles; echo intensity and location of intermediate layers or moderate peaks were taken into account in the analyses. The

PCA allowed us to represent the profile data in a lower-dimensional space, reducing the degrees of freedom while keeping 90% of the variability of the raw dataset. This lower-dimensional space was then grouped into 10 clusters that captured the large variabilities of vertical shapes and identify three large homogeneous regions. Clusters were then used in a machine learning algorithm, an approach still underutilized in analyses of marine data despite their efficiency (De'ath, 2007; Elith et al., 2008). These approaches are often seen as 'black boxes' and the interpretation of the results remains challenging. Here, we obtained a high performance level (87% in cross validation) for the XGBoost modelling. At the same time, the importance of each covariate was ranked, and we evaluated how they influenced each acoustic cluster through SHAP values that have been shown to be reliable indicators of covariate influence (compare to Gain, split count or Saabas) (Lundberg and Lee, 2017). However, the use of SHAP values did not allow us to explore the degree of the interactions among covariates.

The XGBoost model predicted the probability of belonging to a given acoustic cluster based on a vector of covariate values when the range of covariate values used for the model prediction were similar to the range covered by the initial data used to build the model. Except for extremely low values of temperature, oxygen and salinity and for shallow waters, covariates ranges used for prediction were well sampled (Figure A2). The number of predicted cells with covariate vectors out of the covariate range sampled were relatively small, mainly around the Main Island, and extreme northern and southern parts of the EEZ in the cold season (Figure 10, black points). The narrow range of un-sampled covariate values even for the two widely different seasons suggests that observational sampling was sufficient for the purpose of this study.

Using the predicted probabilities at a given point, we reconstructed an acoustic profile by weighting the mean cluster profiles by these probabilities for each point, leading to a complete 3-D reconstruction of acoustic profiles in the EEZ. Despite that correlation between pooled observed and reconstructed backscatter values was high, the variability inside the scatter plot remained relatively high (Figure 11C). Predicted values could vary from +- 10dB for an observed value, pointing out to the limits of the model in terms of reproducing the exact variability. A change of a few dB in the backscatter values may result in large biomass changes (Proud et al., 2018b). Indeed, for layers dominated by Myctophids, a 10dB change with a 38 kHz frequency could translate into a three-fold increase in animal density (Benoit-

Bird, 2009). Yet, the main strength of the model was to reproduce the MTLOs' vertical distribution patterns in space and time rather than predict echo-intensities per se.

In addition, the use of the 38kHz frequency could help in detecting organisms with a swimbladder (mostly fish) while excluding other organisms without gas-filled swimbladders (Davison et al., 2015). Foote (1980) even showed that more than 95% of the organisms' backscatter at 38 kHz was produced by gas-filled swimbladders of fish and gas-filled pneumatophores of siphonophores. One major limitation of our study was the inability to determine if changes in layer intensity and depth position were due to changes in the quantity or in the community/aggregation composition of MTLOs. Developing a similar method including species composition would be of great interest. Some efficient algorithms already exist (e.g. Ariza et al., 2016b; Behagle et al., 2017; Kloser et al., 2016; Korneliussen et al., 2008) based on two or three frequencies that allow for distinguishing among different types of organisms in echograms. Going further into such analysis would require an extensive program of *in situ* sampling with appropriate trawls to identify the species composition of the different layers. So far, with the *in situ* samplers available and given the species diversity already observed (Ceccarelli et al., 2013; Payri et al., 2019), it is unlikely that such a goal can be met in the near future in the New Caledonian region or in the Coral Sea in general.

4.2. Acoustically based regionalization

Studies of biogeography require methods that partition large areas into distinct regions with homogeneous biological and/or physical oceanographic conditions (e.g. Longhurst, 1995, 2007). Our findings provide evidence that new statistical tools used on acoustic recordings are valuable for conducting regionalization that take into account the distribution of MTLOs in the water column. Several studies already include information on mesopelagic organisms (Sutton et al., 2017), and even SSL depth and echo intensity (Proud et al., 2017, 2018). Our methodological framework is reproducible at a global scale, and contributes to the general effort for partitioning the mesopelagic domain using acoustic information on the vertical distribution of MTLOs. Such an expansion would require the acquisition of acoustic data on a large scale and on a 'routine' basis, as is done for other variables on ships of opportunity. Extending the present analysis using a global dataset (e.g. Malaspina 2010 Circumnavigation Expedition, Irigoien et al., 2014; Klevjer et al., 2016) would allow for the detection of higher layer depth variability across the world's ocean.

By predicting and integrating acoustic values for the whole EEZ, we offered a method to map out MTLO spatial distribution and a migrant proportion proxy between day DSL and night SSL. Our range of integrated backscatter 10-600m values was relatively small (-77 to -73dB). We assumed that this small range was because we averaged DSL and SSL. We showed previously that echo-intensity layers change differently for the SSL and for the DSL. Yet, the 10-600m integrated values erased echo-intensity changes. The water column was, on average, denser in MTLOs in the southern part of the EEZ and characterized by higher oxygen concentrations. To calculate a proportion of migrants, our method assumed that mean backscatter values for the epipelagic layer during the night and the day were comparable, e.g. that MTLO community acoustic responses were similar. Because of DVM, we acknowledge that this hypothesis was strong and that results need to be considered with some caution. The 75% proportion of migrants between day DSL and night SSL we estimated is similar to the 62% reported by Klevjer et al. (2016) for the South Pacific. The large proportion of migrants between the night DSL and the day SSL points out the essential role of DVM in the carbon cycle, as already demonstrated (e.g. Davison et al., 2013; Hidaka et al., 2001; Schukat et al., 2013). However, our dataset is limited to the upper 600m of the water column, missing a part of the DSL. Thus, having data incorporating depth layers down to 1000m would help to clarify migration processes as well as organism identification – information that could aid in distinguishing migrant from non-migrant organisms.

Very few biogeochemical models constructed to date have included DVM processes (Ariza et al., 2015). Bianchi et al. (2013b) implemented a DVM on a size-structured NPZD model (Nutrient, Phytoplankton, Zooplankton and Detritus). They reported a migrants' proportion range from 31% (158°W-22.5°N, ALOHA station) to 63% (161°E-47°N, K2 station) between the night DSL and the day SSL. By including for the first time DVM in an end-to-end ecosystem model, Aumont et al. (2018) estimated a contribution of migratory meso- and macro-zooplankton organisms (e.g. smaller than organisms detected by our 38 kHz) to the total epipelagic biomass of about 50% around New Caledonia. Along the same lines, including proportions of migrant and non-migrant MTLOs in models would certainly improve the understanding of carbon cycling in such lower trophic ecosystem models.

4.3. Environmental drivers of the MTLOs vertical distribution

Physiological tolerance varies by species (Duffy et al., 2017). Hence, direct relationships between acoustic echo-intensities and environmental covariates are complex to understand

due to the high species diversity including MTLOs. In this section, we made some hypothesis based on the literature about environmental effects on MTLO dynamics without analyzing them by species.

4.3.1. Role of the most significant covariates

Sun inclination was by far the most important covariate influencing acoustic vertical profiles. Indeed, the DVM pattern drastically affects the shape of the vertical profiles at night and during the day. Ascents and descents of MTLOs depend strongly on light intensity, with migrating organisms generally going up to the surface at night to feed, and returning back to deep layers during the day to avoid visual predation (Benoit-Bird and Au, 2004; Hays, 2003).

The oxygen concentration averaged over the 6-606m water column was the second most important covariate influencing the vertical profiles. We highlighted a positive influence of oxygen concentration on the mean echo intensity, indeed we found that vertical profiles with high mean density were influenced by high oxygen values. Previous studies demonstrated the influence on MTLOs' residing depth. Indeed, in the South Pacific, the lower vertical expansion of daytime and nighttime SSLs is limited by the depth of the oxygen minimum (Bertrand et al., 2010), and Bianchi et al. (2013b) demonstrated at a global scale that the higher the oxygen concentration, the deeper the DSL daytime depth. Our results are in contrast with the widespread hypotheses that low oxygen concentrations provide a refuge from large visual predators and so enhance acoustic values (Bianchi et al., 2013a; Steinberg et al., 2008). As there is no oxygen minimum zone in the region, MTLOs are different from organisms present in regions with oxygen minimum zone. High oxygen at depth in New Caledonia EEZ may then help support higher aggregations of respiring organisms at depth during the day. As oxygen minimum zones are predicted to expand both spatially and vertically in the future (Keeling et al., 2010), the direct impacts on predator distribution (Stramma et al., 2012) are likely to be heightened by changes in prey distribution.

Bathymetry was the third most important variable influencing vertical profile characteristics (Figure 7). The topography of the deep ocean floor is complex around New Caledonia (Gardes et al., 2014), with the presence of three ridges, numerous seamounts with a high shape diversity, one trench and some sedimentary basins. The presence of seamounts, distance to the coast, differences in the topography of the ocean floor between the north and south New Caledonian EEZ affected MTLO vertical distribution. The highest 10-600m mean predicted

acoustic values occurred in the southeast corner, which was characterized by a highly variable bathymetry (Figure 12). Shallow waters offer lower-quality habitat for mesopelagic organisms and may induce increased density in scattering layers of MTLOs in the upper layers of the water column (Escobar-Flores et al., 2018b). In addition, the southeast corner of the EEZ has a high density of seamounts that can enhance the mean backscatter values. Seamount impacts on MTLO dynamics are a function of the type of seamount (e.g. isolated or not, shape, depth, upwelling, currents) and of the organisms that aggregate around seamounts (e.g. Drazen et al., 2011; Morato et al., 2010, 2008; Rogers, 2018).

4.3.2. Other environmental drivers of MTLOs vertical distribution

A deep 20°C isotherm strongly impacted profiles of clusters in the south during the warm season (cluster 3, Figure 10). The depth of the 20°C isotherm is a proxy of thermocline location: a deeper 20°C isotherm is associated with higher stratification and limits nutrient inputs in the euphotic surface layer (Kessler and Cravatte, 2013; Le Borgne et al., 2011). Conversely, a 20°C isotherm close to the surface boosts vertical transport of nutrients to the euphotic layer, and could potentially stimulate the trophic web through enhancing phytoplankton, zooplankton and micronekton production (Benoit-Bird and McManus, 2012; Lebourges-Dhaussy et al., 2014). We emphasized that the link to the isotherm depth was not causal (isotherm depths do not act on animals) but instead reflected different oceanographic conditions. Isotherm depth can also be considered as a proxy for the presence of eddies. Keppler et al. (2018) identified two areas in the New Caledonian EEZ with distinct eddy characteristics: a smaller number of eddies with a longer lifetime in the southern portion of the EEZ part compared to the northern part, with a clear limit at 20°S that corresponds with the spatial delimitation of Figure 10.

Mean 0-600m temperature also impacted acoustic clusters. Spatial patterns of mean temperature showed a strong north-south gradient, with one front well delimited around 20°S (Menkes et al., 2015). This north-south separation coincides with the spatial distribution of the predicted acoustic clusters (Figure 10). We hypothesized that mean temperature was simply a measure of hydrographic features that likely reflected different oceanographic conditions. Therefore, mean temperature appeared to delimit large homogeneous biogeographic regions (the north from the south of the EEZ) rather than having a direct impact on organisms.

The deep euphotic zone had a strong influence on clusters with a deeper center of mass (clusters 7, 8 and 9). A deep euphotic zone indicates a deep chlorophyll maximum characterizing oligotrophic areas dominated by vertical processes with nutrients coming from deep waters. Given that three clusters showed more intense deep SSL compare to other clusters, we assume that a deep euphotic zone allowed ecosystem organization around the chlorophyll maximum with deep zooplankton maxima and, in turn, deep SSL.

High values of surface chlorophyll-*a* concentration had a strong influence on vertical profiles, with the highest echo intensities at the surface layer (daytime cluster 2, Figure 5). We suggest that the enhancement of primary production at the surface during the day could be associated with enhanced zooplankton production, leading to an increase in MTLO density in that part of the water column. The trophic link of a positive relationship between primary production and mesopelagic organisms had been found previously (Escobar-Flores et al., 2013; Irigoien et al., 2014), but at a much wider scale.

Finally, we did not find any effect of the water mass proportions in the New Caledonian EEZ while previous studies (Behagle et al., 2016; Sutton and Beckley, 2017; Jungblut et al., 2017) documented this factor as significant. Water masses are often linked to biogeographical provinces (Briggs and Bowen, 2012) at a wide spatial scale. We hypothesized that the size of New Caledonian EEZ and the lack of contrasting water masses precluded the detection of a significant effect in our case.

4.4. Conclusion and perspectives

While acoustics provides a useful avenue for routine measurement of crucial trophic level organisms, it remains that analyses of such datasets depend on knowledge of the species giving rise to the signal. Indeed, it is crucial to increase our understanding of the actual relationships between acoustic output and true species composition and density. So far, in our region, the lack of adequate *in situ* sampling is a major limitation to our understanding of the ecological processes based on acoustic data.

Our findings highlight the importance of the environmental variables characterizing the structure of the water column, such as the mean oxygen, the euphotic depth (as a proxy for the vertical structure of primary production) and the 20°C isotherm depth (as a proxy for the functioning of ocean dynamics such as upwelling or down welling). Investigating the links between the 3-D oceanographic conditions and the dynamics of MTLO distributions requires

information on the physics, biogeochemistry and biology of the upper (< ~1000m) water column. Accessing 3-D high-resolution oceanographic data or models to understand complex interactions at the acoustic data acquisition scale is not yet possible. For instance, we did not find a satisfying 3-D biogeochemical model estimating the values of primary production that matched with our in situ data. Consolidating and validating coupled dynamical-biogeochemical 3-D models will strongly contribute to a better comprehension of the pelagic ecosystem.

Variability in scattering layer depths during day and night impacts predator-prey interactions, including the predators' energy budget allocated to feeding. For active vertically migrating predators, the possibility to feed in shallow, warm, and rich water brings energetic savings (Hazen et al., 2015). Acoustic data are extremely helpful for analyzing predator-prey interactions on a wide scale (Bertrand et al., 2003; Koslow, 2009). Indeed, acoustics recorded continuously along vessels' tracks allow building datasets with better coverage than trawl data or predators' stomach content, for example. However, the area sampled along cruise tracks remains small compared to the area in which predators feed. Our model provides the possibility to fill gaps around acoustic cruise tracks while keeping all the information contained on acoustic profiles. Then, by averaging vertical predictions on a specific vertical layer (e.g. 0-30m to link to seabirds foraging compartment), we offer an innovative method to predict a proxy of prey biomass (through echo intensity) that could be used in predator niche modeling on large spatial and temporal scales (Briand et al., 2011; Lambert et al., 2014; Miller et al., 2018).

Further, our method may provide a valuable contribution to assess the climate change impact on MTLOs, and consequently on pelagic ecosystems as a whole. The increase in ocean temperature, the extension of low-oxygen zones (Bindoff et al., 2007; Doney et al., 2012) as well as the possible decrease of the primary production (Bopp et al., 2013) predicted under future scenarios of environmental change will have dramatic effects on the distribution of MTLOs (i.e. layer position, abundance and biomass). Our model, including predictive oceanographic variables could help to predict changes in MTLO vertical distributions (i.e. layer position and echo intensity) for the next 100 years using climate change scenarios as Proud et al. (2017) did.

Acknowledgements.

This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of A. Receveur and can under no circumstance be regarded as reflecting the position of European Union. We thank R/V ALIS officers and crews and science parties who participated to the cruises which data are included in the present paper. This work was supported by the French national program LEFE/INSU. We thank Jed MacDonald for his constructive comments that improve the earlier version of the manuscript. Finally, we gratefully thank the two reviewers for their useful comments.

5. References

Ariza, A., Garijo, J.C., Landeira, J.M., Bordes, F., Hernández-León, S., 2015. Migrant biomass and respiratory carbon flux by zooplankton and micronekton in the subtropical northeast Atlantic Ocean (Canary Islands). Prog. Oceanogr. 134, 330–342. https://doi.org/10.1016/j.pocean.2015.03.003

Ariza, A., Landeira, J.M., Escánez, A., Wienerroither, R., Aguilar de Soto, N., Røstad, A., Kaartvedt, S., Hernández-León, S., 2016a. Vertical distribution, composition and migratory patterns of acoustic scattering layers in the Canary Islands. J. Mar. Syst. 157, 82–91. https://doi.org/10.1016/j.jmarsys.2016.01.004

Ariza, A., Landeira, J.M., Escánez, A., Wienerroither, R., Aguilar de Soto, N., Røstad, A., Kaartvedt, S., Hernández-León, S., 2016b. Vertical distribution, composition and migratory patterns of acoustic scattering layers in the Canary Islands. J. Mar. Syst. 157, 82–91. https://doi.org/10.1016/j.jmarsys.2016.01.004

Aumont, O., Maury, O., Lefort, S., Bopp, L., 2018. Evaluating the potential impacts of the diurnal vertical migration by marine organisms on marine biogeochemistry 74.

Bedford, M., Melbourne-Thomas, J., Corney, S., Jarvis, T., Kelly, N., Constable, A., 2015. Prey-field use by a Southern Ocean top predator: enhanced understanding using integrated datasets. Mar. Ecol. Prog. Ser. 526, 169–181. https://doi.org/10.3354/meps11203

Behagle, N., Cotté, C., Lebourges-Dhaussy, A., Roudaut, G., Duhamel, G., Brehmer, P., Josse, E., Cherel, Y., 2017. Acoustic distribution of discriminated micronektonic organisms from a bi-frequency processing: The case study of eastern Kerguelen oceanic waters. Prog. Oceanogr. 156, 276–289. https://doi.org/10.1016/j.pocean.2017.06.004

Behagle, N., Cotté, C., Ryan, T.E., Gauthier, O., Roudaut, G., Brehmer, P., Josse, E., Cherel, Y., 2016. Acoustic micronektonic distribution is structured by macroscale oceanographic processes across 20–50°S latitudes in the South-Western Indian Ocean. Deep Sea Res. Part Oceanogr. Res. Pap. 110, 20–32. https://doi.org/10.1016/j.dsr.2015.12.007

Behagle, N., du Buisson, L., Josse, E., Lebourges-Dhaussy, A., Roudaut, G., Ménard, F., 2014. Mesoscale features and micronekton in the Mozambique Channel: An acoustic approach. Deep Sea Res. Part II Top. Stud. Oceanogr. 100, 164–173. https://doi.org/10.1016/j.dsr2.2013.10.024

Belcher, A., Saunders, R.A., Tarling, G.A., 2019. Respiration rates and active carbon flux of mesopelagic fishes (Family Myctophidae) in the Scotia Sea, Southern Ocean. Mar. Ecol. Prog. Ser. 610, 149–162. https://doi.org/10.3354/meps12861

Benoit-Bird, K.J., 2009. The effects of scattering-layer composition, animal size, and numerical density on the frequency response of volume backscatter. ICES J. Mar. Sci. J. Cons.

Benoit-Bird, K.J., Au, W.W.L., 2004. Diel migration dynamics of an island-associated sound-scattering layer. Deep Sea Res. Part Oceanogr. Res. Pap. 51, 707–719. https://doi.org/10.1016/j.dsr.2004.01.004

Benoit-Bird, K.J., McManus, M.A., 2012. Bottom-up regulation of a pelagic community through spatial aggregations. Biol. Lett. 8, 813–816. https://doi.org/10.1098/rsbl.2012.0232

Benoit-Bird, K.J., Moline, M.A., Southall, B.L., 2017. Prey in oceanic sound scattering layers organize to get a little help from their friends: Schooling within sound scattering layers. Limnol. Oceanogr. https://doi.org/10.1002/lno.10606

Bertrand, A., Ballón, M., Chaigneau, A., 2010. Acoustic Observation of Living Organisms Reveals the Upper Limit of the Oxygen Minimum Zone. PLoS ONE 5, e10330. https://doi.org/10.1371/journal.pone.0010330

Bertrand, A., Bard, F.-X., Josse, E., 2002. Tuna food habits related to the micronekton distribution in French Polynesia. Mar. Biol. 140, 1023–1037. https://doi.org/10.1007/s00227-001-0776-3

Bertrand, A., Josse, E., Bach, P., Dagorn, L., 2003. Acoustics for ecosystem research: lessons and perspectives from a scientific programme focusing on tuna-environment relationships. Aquat. Living Resour., Acoustics in Fisheries and Aquatic Ecology. Part 2 16, 197–203. https://doi.org/10.1016/S0990-7440(03)00018-4

Bianchi, D., Galbraith, E.D., Carozza, D.A., Mislan, K.A.S., Stock, C.A., 2013a. Intensification of open-ocean oxygen depletion by vertically migrating animals. Nat. Geosci. 6, 545–548. https://doi.org/10.1038/ngeo1837

Bianchi, D., Mislan, K.A.S., 2016. Global patterns of diel vertical migration times and velocities from acoustic data: Global patterns of diel vertical migration. Limnol. Oceanogr. 61, 353–364. https://doi.org/10.1002/lno.10219

Bianchi, D., Stock, C., Galbraith, E.D., Sarmiento, J.L., 2013b. Diel vertical migration: Ecological controls and impacts on the biological pump in a one-dimensional ocean model. Glob. Biogeochem. Cycles 27, 478–491. https://doi.org/10.1002/gbc.20031

Bindoff, N.L., Willebrand, J., Artale, V., Cazenave, A., Gregory, J.M., Gulev, S., Hanawa, K., Le Quere, C., Levitus, S., Nojiri, Y., Shum, CK., Talley, LD., Unnikrishnan, AS., 2007. Observations: Oceanic Climate Change and Sea Level, in: Climate Change 2007: The Physical Science Basis: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S Solomon, D Qin, M Manning, Z Chen, M Marquis, et al., pp. 385–432.

Blanc, Ph., Wald, L., 2012. The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period. Sol. Energy 86, 3072–3083. https://doi.org/10.1016/j.solener.2012.07.018

Boersch-Supan, P.H., Rogers, A.D., Brierley, A.S., 2017. The distribution of pelagic sound scattering layers across the southwest Indian Ocean. Deep Sea Res. Part II Top. Stud. Oceanogr., The Pelagic Ecology of Seamounts of the South West Indian Ocean 136, 108–121. https://doi.org/10.1016/j.dsr2.2015.06.023

Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., Vichi, M., 2013. Multiple stressors of

ocean ecosystems in the 21st century: projections with CMIP5 models. Biogeosciences 10, 6225–6245. https://doi.org/10.5194/bg-10-6225-2013
Briand, K., Molony, B., Lehodey, P., 2011. A study on the variability of albacore (Thunnus alalunga) longline catch rates in the southwest Pacific Ocean: Albacore longline fishery and environment. Fish. Oceanogr. 20, 517–529. https://doi.org/10.1111/j.1365-2419.2011.00599.x
Briggs, J.C., Bowen, B.W., 2012. A realignment of marine biogeographic provinces with particular reference to fish distributions: Marine biogeographic provinces. J. Biogeogr. 39, 12–30. https://doi.org/10.1111/j.1365-2699.2011.02613.x
Browne, M.W., Cudeck, R., 1989. Single Sample Cross-Validation Indices for Covariance Structures. Multivar. Behav. Res. 24, 445–455. https://doi.org/10.1207/s15327906mbr2404_4
Burgos, J.M., Horne, J.K., 2008. Characterization and classification of acoustically detected fish spatial distributions. ICES J. Mar. Sci. 65, 1235–1247.
Cade, D.E., Benoit-Bird, K.J., 2015. Depths, migration rates and environmental associations of acoustic scattering layers in the Gulf of California. Deep Sea Res. Part Oceanogr. Res. Pap. 102, 78–89. https://doi.org/10.1016/j.dsr.2015.05.001
Ceccarelli, D.M., McKinnon, A.D., Andréfouët, S., Allain, V., Young, J., Gledhill, D.C., Flynn, A., Bax, N.J., Beaman, R., Borsa, P., Brinkman, R., Bustamante, R.H., Campbell, R., Cappo, M., Cravatte, S., D'Agata, S., Dichmont, C.M., Dunstan, P.K., Dupouy, C., Edgar, G., Farman, R., Furnas, M., Garrigue, C., Hutton, T., Kulbicki, M., Letourneur, Y., Lindsay, D., Menkes, C., Mouillot, D., Parravicini, V., Payri, C., Pelletier, B., Richer de Forges, B., Ridgway, K., Rodier, M., Samadi, S., Schoeman, D., Skewes, T., Swearer, S., Vigliola, L., Wantiez, L., Williams, Alan, Williams, Ashley, Richardson, A.J., 2013. The Coral Sea, in: Advances in Marine Biology. Elsevier, pp. 213–290. https://doi.org/10.1016/B978-0-12-408096-6.00004-3
Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. ACM, New York, NY, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785
Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., implementation), Xgb. contributors (base Xgb., 2018. xgboost: Extreme Gradient Boosting.
Choy, C.A., Haddock, S.H.D., Robison, B.H., 2017. Deep pelagic food web structure as revealed by *in situ* feeding observations. Proc. R. Soc. B Biol. Sci. 284, 20172116. https://doi.org/10.1098/rspb.2017.2116
Christensen, N.L., Bartuska, A.M., Brown, J.H., Carpenter, S., D'Antonio, C., Francis, R., Franklin, J.F., MacMahon, J.A., Noss, R.F., Parsons, D.J., Peterson, C.H., Turner, M.G., Woodmansee, R.G., 1996. The Report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. Ecol. Appl. 6, 665–691. https://doi.org/10.2307/2269460
Davison, P.C., Checkley Jr., D.M., Koslow, J.A., Barlow, J., 2013. Carbon export mediated by mesopelagic fishes in the northeast Pacific Ocean. Prog. Oceanogr. 116, 14–30.
Davison, P.C., Koslow, J.A., Kloser, R.J., 2015. Acoustic biomass estimation of mesopelagic fish: backscattering from individuals, populations, and communities. ICES J. Mar. Sci. 72, 1413–1424. https://doi.org/10.1093/icesjms/fsv023
De Robertis, A., Higginbottom, I., 2007. A post-processing technique to estimate the signal-to-noise ratio and remove echosounder background noise. ICES J. Mar. Sci. 64, 1282–1291.

De'ath, G., 2007. Boosted Trees for Ecological Modeling and Prediction. Ecology 88, 243–251. https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N., Sydeman, W.J., Talley, L.D., 2012. Climate Change Impacts on Marine Ecosystems. Annu. Rev. Mar. Sci. 4, 11–37. https://doi.org/10.1146/annurev-marine-041911-111611

Doray, M., Petitgas, P., Nelson, L., Mahévas, S., Josse, E., Reynal, L., 2009. The influence of the environment on the variability of monthly tuna biomass around a moored, fish-aggregating device. ICES J. Mar. Sci. 66, 1410–1416.

Drazen, J.C., De Forest, L.G., Domokos, R., 2011. Micronekton abundance and biomass in Hawaiian waters as influenced by seamounts, eddies, and the moon. Deep Sea Res. Part Oceanogr. Res. Pap. 58, 557–566. https://doi.org/10.1016/j.dsr.2011.03.002

Drazen, J.C., Sutton, T.T., 2017. Dining in the Deep: The Feeding Ecology of Deep-Sea Fishes. Annu. Rev. Mar. Sci. 9, 337–366. https://doi.org/10.1146/annurev-marine-010816-060543

Duffy, L.M., Kuhnert, P.M., Pethybridge, H.R., Young, J.W., Olson, R.J., Logan, J.M., Goñi, N., Romanov, E., Allain, V., Staudinger, M.D., Abecassis, M., Choy, C.A., Hobday, A.J., Simier, M., Galván-Magaña, F., Potier, M., Ménard, F., 2017. Global trophic ecology of yellowfin, bigeye, and albacore tunas: Understanding predation on micronekton communities at ocean-basin scales. Deep Sea Res. Part II Top. Stud. Oceanogr. 140, 55–73. https://doi.org/10.1016/j.dsr2.2017.03.003

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Escobar-Flores, P., O'Driscoll, R., Montgomery, J., 2013. Acoustic characterization of pelagic fish distribution across the South Pacific Ocean. Mar. Ecol. Prog. Ser. 490, 169–183. https://doi.org/10.3354/meps10435

Escobar-Flores, P.C., O'Driscoll, R.L., Montgomery, J.C., 2018a. Spatial and temporal distribution patterns of acoustic backscatter in the New Zealand sector of the Southern Ocean. Mar. Ecol. Prog. Ser. 592, 19–35. https://doi.org/10.3354/meps12489

Escobar-Flores, P.C., O'Driscoll, R.L., Montgomery, J.C., 2018b. Predicting distribution and relative abundance of mid-trophic level organisms using oceanographic parameters and acoustic backscatter. Mar. Ecol. Prog. Ser. 592, 37–56. https://doi.org/10.3354/meps12519

Foote, K.G., 1980. Importance of the swimbladder in acoustic scattering by fish: A comparison of gadoid and mackerel target strengths. J. Acoust. Soc. Am. 67, 2084–2089. https://doi.org/10.1121/1.384452

Foote, K.G., Knudsen, H.P., Vestnes, G., 1987. Calibration of acoustic instruments for fish density estimation : a practical guide. Coop. Res. Rep. Int. Counc. Explor. Sea 144.

Fraley, C., Raftery, A., 2002. Model-Based Clustering, Discriminant Analysis, and Density Estimation. J. Am. Stat. Assoc. 97, 611–631. https://doi.org/10.1198/016214502760047131

Fulton, E.A., Smith, A.D.M., Punt, A.E., 2005. Which ecological indicators can robustly detect effects of fishing? ICES J. Mar. Sci. 62, 540–551. https://doi.org/10.1016/j.icesjms.2004.12.012

Gardes, L., Tessier, E., Allain, V., Alloncle, N., Baudat-Franceschi, J., Butaud, J.-F., Collot, J., Etaix-Bonnin, R., Hubert, A., Jourdan, H., Loisier, A., Menkes, C., Payri, C., Rouillard, P., Samadi, S., Vidal, E., Yokohama, Y., 2014. Analyse stratégique de l'Espace maritime de la Nouvelle-Calédonie - vers une gestion intégrée. Agence des aires marines protégées / Gouvernement de la Nouvelle-Calédonie éditeurs.

Garric, G., Parent, L., Greiner, E., Drévillon, M., Hamon, M., Lellouche, J.-M., Régnier, C., Desportes, C., Le Galloudec, O., Bricaud, C., Drillet, Y., Hernandez, F., Le Traon, P.-Y., 2017. Performance and quality assessment of the global ocean eddy-permitting physical reanalysis GLORYS2V4. Presented at the EGU General Assembly Conference Abstracts, p. 18776.

Gasparin, F., Maes, C., Sudre, J., Garcon, V., Ganachaud, A., 2014. Water mass analysis of the Coral Sea through an Optimum Multiparameter method. J. Geophys. Res. Oceans 119, 7229–7244. https://doi.org/10.1002/2014JC010246

Germineaud, C., Ganachaud, A., Sprintall, J., Cravatte, S., Eldin, G., Alberty, M.S., Privat, E., Germineaud, C., Ganachaud, A., Sprintall, J., Cravatte, S., Eldin, G., Alberty, M.S., Privat, E., 2016. Pathways and Water Mass Properties of the Thermocline and Intermediate Waters in the Solomon Sea. Httpdxdoiorg101175JPO--16-01071. https://doi.org/10.1175/JPO-D-16-0107.1

Godo, O.R., Samuelsen, A., Macaulay, G.J., Patel, R., Hjøllo, S.S., Horne, J., Kaartvedt, S., Johannessen, J.A., 2012. Mesoscale Eddies Are Oases for Higher Trophic Marine Life. PLoS ONE 7, e30161. https://doi.org/10.1371/journal.pone.0030161

Grandperrin, R., 1975. Structures trophiques aboutissant aux thons de longue ligne dans le Pacifique sud-ouest tropical (PhD). Aix Marseille Universite.

Guinehut, S., Dhomps, A.-L., Larnicol, G., Le Traon, P.-Y., 2012. High Resolution 3-D temperature and salinity fields derived from in situ and satellite observations. Ocean Sci. Discuss. 9, 1313–1347. https://doi.org/10.5194/osd-9-1313-2012

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. J. R. Stat. Soc. Ser. C Appl. Stat. 28, 100–108. https://doi.org/10.2307/2346830

Hays, G.C., 2003. A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations, in: Migrations and Dispersal of Marine Organisms. Springer, pp. 163–170.

Hazen, E.L., Friedlaender, A.S., goldbogen, J.A., 2015. Blue whales (Balaenoptera musculus) optimize foraging efficiency by balancing oxygen use and energy gain as a function of prey density. Sci. Adv. 1.

Hidaka, K., Kawaguchi, K., Murakami, M., Takahashi, M., 2001. Downward transport of organic carbon by diel migratory micronekton in the western equatorial Paci"c: its quantitative and qualitative importance. Deep-Sea Res. I 48, 1923–1939.

Houssard, P., Lorrain, A., Tremblay-Boyer, L., Allain, V., Graham, B.S., Menkes, C.E., Pethybridge, H., Couturier, L.I.E., Point, D., Leroy, B., Receveur, A., Hunt, B.P.V., Vourey, E., Bonnet, S., Rodier, M., Raimbault, P., Feunteun, E., Kuhnert, P.M., Munaron, J.-M., Lebreton, B., Otake, T., Letourneur, Y., 2017. Trophic position increases with thermocline depth in yellowfin and bigeye tuna across the Western and Central Pacific Ocean. Prog. Oceanogr. 154, 49–63. https://doi.org/10.1016/j.pocean.2017.04.008

Irigoien, X., Klevjer, T.A., Røstad, A., Martinez, U., Boyra, G., Acuña, J.L., Bode, A., Echevarria, F., Gonzalez-Gordillo, J.I., Hernandez-Leon, S., Agusti, S., Aksnes, D.L., Duarte, C.M., Kaartvedt, S., 2014. Large mesopelagic fishes biomass and trophic efficiency in the open ocean. Nat. Commun. 5. https://doi.org/10.1038/ncomms4271

Jolliffe, I., 2011. Principal Component Analysis, in: Lovric, M. (Ed.), International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1094–1096. https://doi.org/10.1007/978-3-642-04898-2_455

Keeling, R.F., Körtzinger, A., Gruber, N., 2010. Ocean Deoxygenation in a Warming World. Annu. Rev. Mar. Sci. 2, 199–229. https://doi.org/10.1146/annurev.marine.010908.163855

Keppler, L., Cravatte, S., Chaigneau, A., Pegliasco, C., Gourdeau, L., Singh, A., 2018. Observed Characteristics and Vertical Structure of Mesoscale Eddies in the Southwest Tropical Pacific. J. Geophys. Res. Oceans 123, 2731–2756. https://doi.org/10.1002/2017JC013712

Kessler, W.S., Cravatte, S., 2013. Mean circulation of the Coral Sea. J. Geophys. Res. Oceans 118, 6385–6410. https://doi.org/10.1002/2013JC009117

Klevjer, T.A., Irigoien, X., Røstad, A., Fraile-Nuez, E., Benítez-Barrios, V.M., Kaartvedt., S., 2016. Large scale patterns in vertical distribution and behaviour of mesopelagic scattering layers. Sci. Rep. 6, 19873. https://doi.org/10.1038/srep19873

Kloser, R.J., Ryan, T., Sakov, P., Williams, A., Koslow, J.A., 2002. Species identification in deep water using multiple acoustic frequencies. Can. J. Fish. Aquat. Sci. 59, 1065–1077. https://doi.org/10.1139/f02-076

Kloser, R.J., Ryan, T.E., Keith, G., Gershwin, L., 2016. Deep-scattering layer, gas-bladder density, and size estimates using a two-frequency acoustic and optical probe. ICES J. Mar. Sci. J. Cons. 73, 2037–2048. https://doi.org/10.1093/icesjms/fsv257

Koenker, R., 2004. Quantile regression for longitudinal data. J. Multivar. Anal., Special Issue on Semiparametric and Nonparametric Mixed Models 91, 74–89. https://doi.org/10.1016/j.jmva.2004.05.006

Korneliussen, R.J., Diner, N., Ona, E., Berger, L., Fernandes, P.G., 2008. Proposals for the collection of multifrequency acoustic data. ICES J. Mar. Sci. J. Cons. 65, 982–994.

Koslow, J.A., 2009. The role of acoustics in ecosystem-based fishery management. ICES J. Mar. Sci. 66, 966–973.

Lambert, C., Mannocci, L., Lehodey, P., Ridoux, V., 2014. Predicting Cetacean Habitats from Their Energetic Needs and the Distribution of Their Prey in Two Contrasted Tropical Regions. PLoS ONE 9, e105958. https://doi.org/10.1371/journal.pone.0105958

Le Borgne, R., Allain, V., Griffiths, S.P., Matear, R.J., McKinnon, A.D., Richardson, A.J., Young, J.W., 2011. Vulnerability of open ocean food webs in the tropical Pacific to climate change, in: Vulnerability of Tropical Pacific Fisheries and Aquaculture to Climate Change. Secretariat of the Pacific Community, New Caledonia.

Lebourges-Dhaussy, A., Huggett, J., Ockhuis, S., Roudaut, G., Josse, E., Verheye, H., 2014. Zooplankton size and distribution within mesoscale structures in the Mozambique Channel: A comparative approach using the TAPS acoustic profiler, a multiple net sampler and ZooScan image analysis. Deep Sea Res. Part II Top. Stud. Oceanogr. 100, 136–152. https://doi.org/10.1016/j.dsr2.2013.10.022

Lehodey, P., Murtugudde, R., Senina, I., 2010. Bridging the gap from ocean models to population dynamics of large marine predators: A model of mid-trophic functional groups. Prog. Oceanogr. 84, 69–84. https://doi.org/10.1016/j.pocean.2009.09.008

Longhurst, A., 1995. Seasonal cycles of pelagic production and consumption. Prog. Oceanogr. 36, 77–167. https://doi.org/10.1016/0079-6611(95)00015-1

Longhurst, A.R., 2007. Ecological geography of the sea. Academic Press, Amsterdam; Boston, MA.

Lundberg, S.M., Erion, G.G., Lee, S.-I., 2018. Consistent Individualized Feature Attribution for Tree Ensembles. ArXiv180203888 Cs Stat.

Lundberg, S.M., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions, in: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 30. Curran Associates, Inc., pp. 4765–4774.

Maclennan, D.N., Fernandes, P.G., Dalen, J., 2002. A consistent approach to definitions and symbols in fisheries acoustics. ICES J. Mar. Sci. 59, 365–369. https://doi.org/10.1006/jmsc.2001.1158

Maury, O., 2010. An overview of APECOSM, a spatialized mass balanced "Apex Predators ECOSystem Model" to study physiologically structured tuna population dynamics in their ecosystem. Prog. Oceanogr., Special Issue: Parameterisation of Trophic Interactions in Ecosystem Modelling 84, 113–117. https://doi.org/10.1016/j.pocean.2009.09.013

McGehee, D.E., O'Driscoll, R.L., Traykovski, L.V.M., 1998. Effects of orientation on acoustics scattering from Antarctic krill at 120 kHz. Deep Sea Res. Part II 45, 1273–1294.

Menkes, C.E., Allain, V., Rodier, M., Gallois, F., Lebourges-Dhaussy, A., Hunt, B.P.V., Smeti, H., Pagano, M., Josse, E., Daroux, A., Lehodey, P., Senina, I., Kestenare, E., Lorrain, A., Nicol, S., 2015. Seasonal oceanography from physics to micronekton in the south-west Pacific. Deep Sea Res. Part II Top. Stud. Oceanogr. 113, 125–144. https://doi.org/10.1016/j.dsr2.2014.10.026

Michalsky, J.J., 1988. The Astronomical Almanac's algorithm for approximate solar position (1950–2050). Sol. Energy 40, 227–235. https://doi.org/10.1016/0038-092X(88)90045-X

Miller, M., Carlile, N., Scutt Phillips, J., McDuie, F., Congdon, B., 2018. Importance of tropical tuna for seabird foraging over a marine productivity gradient. Mar. Ecol. Prog. Ser. 586, 233–249. https://doi.org/10.3354/meps12376

Morato, T., Hoyle, S.D., Allain, V., Nicol, S.J., 2010. Seamounts are hotspots of pelagic biodiversity in the open ocean. Proc. Natl. Acad. Sci. 107, 9707–9711.

Morato, T., Varkey, D., Damaso, C., Machete, M., Santos, M., Prieto, R., Pitcher, T., Santos, R., 2008. Evidence of a seamount effect on aggregating visitors. Mar. Ecol. Prog. Ser. 357, 23–32. https://doi.org/10.3354/meps07269

Olson, R., Duffy, L., Kuhnert, P., Galván-Magaña, F., Bocanegra-Castillo, N., Alatorre-Ramírez, V., 2014. Decadal diet shift in yellowfin tuna Thunnus albacares suggests broad-scale food web changes in the eastern tropical Pacific Ocean. Mar. Ecol. Prog. Ser. 497, 157–178. https://doi.org/10.3354/meps10609

Opdal, A.F., Godø, O.R., Bergstad, O.A., Fiksen, ø., 2008. Distribution, identity, and possible processes sustaining meso- and bathypelagic scattering layers on the northern Mid-Atlantic Ridge. Deep Sea Res. Part II Top. Stud. Oceanogr. 55, 45–58. https://doi.org/10.1016/j.dsr2.2007.09.002

Pauly, D., Christensen, V., Walters, C., 2000. Ecopath, Ecosim, and Ecospace as tools for evaluating ecosystem impact of fisheries. ICES J. Mar. Sci. 57, 697–706. https://doi.org/10.1006/jmsc.2000.0726

Payri, C.E., Allain, V., Aucan, J., David, C., David, V., Dutheil, C., Loubersac, L., Menkes, C., Pelletier, B., Pestana, G., Samadi, S., 2019. Chapter 27 - New Caledonia, in: Sheppard, C. (Ed.), World Seas: An Environmental Evaluation (Second Edition). Academic Press, pp. 593–618. https://doi.org/10.1016/B978-0-08-100853-9.00035-X

Pearre, S., 2003. Eat and run? The hunger/satiation hypothesis in vertical migration: history, evidence and consequences. Biol. Rev. Camb. Philos. Soc. 78, 1–79. https://doi.org/10.1017/S146479310200595X

Perrot, Y., Brehmer, P., Habasque, J., Roudaut, G., Behagle, N., Sarré, A., Lebourges-Dhaussy, A., 2018. Matecho: An Open-Source Tool for Processing Fisheries Acoustics Data. Acoust. Aust. 8.

Proud, R., Cox, M.J., Brierley, A.S., 2017. Biogeography of the Global Ocean's Mesopelagic Zone. Curr. Biol. 27, 113–119. https://doi.org/10.1016/j.cub.2016.11.003

Proud, R., Cox, M.J., le guen, C., Brierley, A.S., 2018a. Fine-scale depth structure of pelagic communities throughout the global ocean based on acoustic sound scattering layers. Mar. Ecol. Prog. Ser. 598, 35–48. https://doi.org/10.3354/meps12612

Proud, R., Handegard, N.O., Kloser, R.J., Cox, M.J., Brierley, A.S., Handling editor: David Demer, 2018b. From siphonophores to deep scattering layers: uncertainty ranges for the estimation of global mesopelagic fish biomass. ICES J. Mar. Sci. https://doi.org/10.1093/icesjms/fsy037

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A., 1995. Bayesian Model Selection in Social Research (with Discussion by Andrew Gelman & Donald B. Rubin, and Robert M. Hauser), Peter V. Marsden. ed, Sociological Methodology. Blackwells, Oxford, U.K.

Receveur, A., Kestenare, E., Allain, V., Menard, F., Cravatte, S., Lebourges-Dhaussy, A., Lehodey, P., Mangeas, M., Smith, N., Radenac, M.H., Menkes, C., Submitted. Micronekton distribution in the southwest Pacific (New Caledonia) inferred from Shipboard-ADCP backscatter data. Deep Sea Res. Part Oceanogr. Res. Pap.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. J. Clim. 20, 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Ridgway, K.R., Dunn, J.R., Wilkin, J.L., 2002. Ocean Interpolation by Four-Dimensional Weighted Least Squares—Application to the Waters around Australasia. J. Atmospheric Ocean. Technol. 19, 1357–1375. https://doi.org/10.1175/1520-0426(2002)019<1357:OIBFDW>2.0.CO;2

Ritz, D.A., Hobday, A.J., Montgomery, J.C., Ward, A.J.W., 2011. Chapter Four - Social Aggregation in the Pelagic Zone with Special Reference to Fish and Invertebrates, in: Lesser, M. (Ed.), Advances in Marine Biology, Advances in Marine Biology. Academic Press, pp. 161–227. https://doi.org/10.1016/B978-0-12-385529-9.00004-4

Rogers, A.D., 2018. The biology of seamounts: 25 Years on. Adv. Mar. Biol. 79.

Saulquin, B., Gohin, F., Garello, R., 2009. Regional objective analysis for merging MERIS, MODIS/Aqua and SeaWiFS Chlorophyll-a data from 1998 to 2008 on the European Atlantic Shelf at a resolution of 1.1Km. Oceans 2009 - Eur. 1, 1165–1174. https://doi.org/10.1109/OCEANSE.2009.5278165

Schaefer, K.M., 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. Mar. Biol. 157, 2625–2642. https://doi.org/10.1007/s00227-010-1524-3

Schaefer, K.M., Fuller, D.W., 2007. Vertical movement patterns of skipjack tuna (Katsuwonus pelamis) in the eastern equatorial Pacific Ocean, as revealed with archival tags. Fish. Bull. 105, 379–389.

Schukat, A., Bode, M., Auel, H., Carballo, R., Martin, B., Koppelmann, R., Hagen, W., 2013. Pelagic decapods in the northern Benguela upwelling system Distribution, ecophysiology and contribution to active carbon flux. Deep-Sea Res. I 75, 146–156.

Scrucca, L., Fop, M., Murphy, T.B., Raftery, A.E., 2016. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. R J. 8, 289–317.

Shin, Y.-J., Cury, P., 2001. Exploring fish community dynamics through size-dependent trophic interactions using a spatialized individual-based model. Aquat. Living Resour. 14, 65–80. https://doi.org/10.1016/S0990-7440(01)01106-8

Steinberg, D.K., Cope, J.S., Wilson, S.E., Kobari, T., 2008. A comparison of mesopelagic mesozooplankton community structure in the subtropical and subarctic North Pacific Ocean. Deep Sea Res. Part II Top. Stud. Oceanogr., Understanding the Ocean's Biological Pump:results from VERTIGO 55, 1615–1635. https://doi.org/10.1016/j.dsr2.2008.04.025

Stramma, L., Prince, E.D., Schmidtko, S., Luo, J., Hoolihan, J.P., Visbeck, M., Wallace, D.W.R., Brandt, P., Körtzinger, A., 2012. Expansion of oxygen minimum zones may

reduce available habitat for tropical pelagic fishes. Nat. Clim. Change 2, 33–37. https://doi.org/10.1038/nclimate1304

Sutton, A.L., Beckley, L.E., 2017. Vertical structuring of epipelagic euphausiid assemblages across a thermohaline front in the south-east Indian Ocean. J. Plankton Res. 1–16. https://doi.org/10.1093/plankt/fbx012

Sutton, T.T., Clark, M.R., Dunn, D.C., Halpin, P.N., Rogers, A.D., Guinotte, J., Bograd, S.J., Angel, M.V., Perez, J.A.A., Wishner, K., Haedrich, R.L., Lindsay, D.J., Drazen, J.C., Vereshchaka, A., Piatkowski, U., Morato, T., Błachowiak-Samołyk, K., Robison, B.H., Gjerde, K.M., Pierrot-Bults, A., Bernal, P., Reygondeau, G., Heino, M., 2017. A global biogeographic classification of the mesopelagic zone. Deep Sea Res. Part Oceanogr. Res. Pap. 126, 85–102. https://doi.org/10.1016/j.dsr.2017.05.006

Urmy, S.S., Horne, J.K., Barbee, D.H., 2012. Measuring the vertical distributional variability of pelagic fauna in Monterey Bay. ICES J. Mar. Sci. 69, 184–196. https://doi.org/10.1093/icesjms/fsr205

Wentz, F.J., Scott, R.H., Leidner, M., Atlas, R., Ardizzone, J., 2015. Remote Sensing Systems Cross-Calibrated Multi-Platform (CCMP) 6-hourly ocean vector wind analysis product on 0.25 deg grid, Remote Sensing Systems. Santa Rosa, CA.

Young, J.W., Hunt, B.P.V., Cook, T.R., Llopiz, J.K., Hazen, E.L., Pethybridge, H.R., Ceccarelli, D., Lorrain, A., Olson, R.J., Allain, V., Menkes, C., Patterson, T., Nicol, S., Lehodey, P., Kloser, R.J., Arrizabalaga, H., Anela Choy, C., 2015. The trophodynamics of marine top predators: Current knowledge, recent advances and challenges. Deep Sea Res. Part II Top. Stud. Oceanogr. 113, 170–187. https://doi.org/10.1016/j.dsr2.2014.05.015

Zedel, L., Patro, R., Knutsen, T., 2005. Fish behaviour and orientation-dependent backscatter in acoustic Doppler profiler data. ICES J. Mar. Sci. 62, 1191–1201.

ZoNéCo, 2013. L'atlas bathymétrique de la Nouvelle-Calédonie. ZoNéCo.

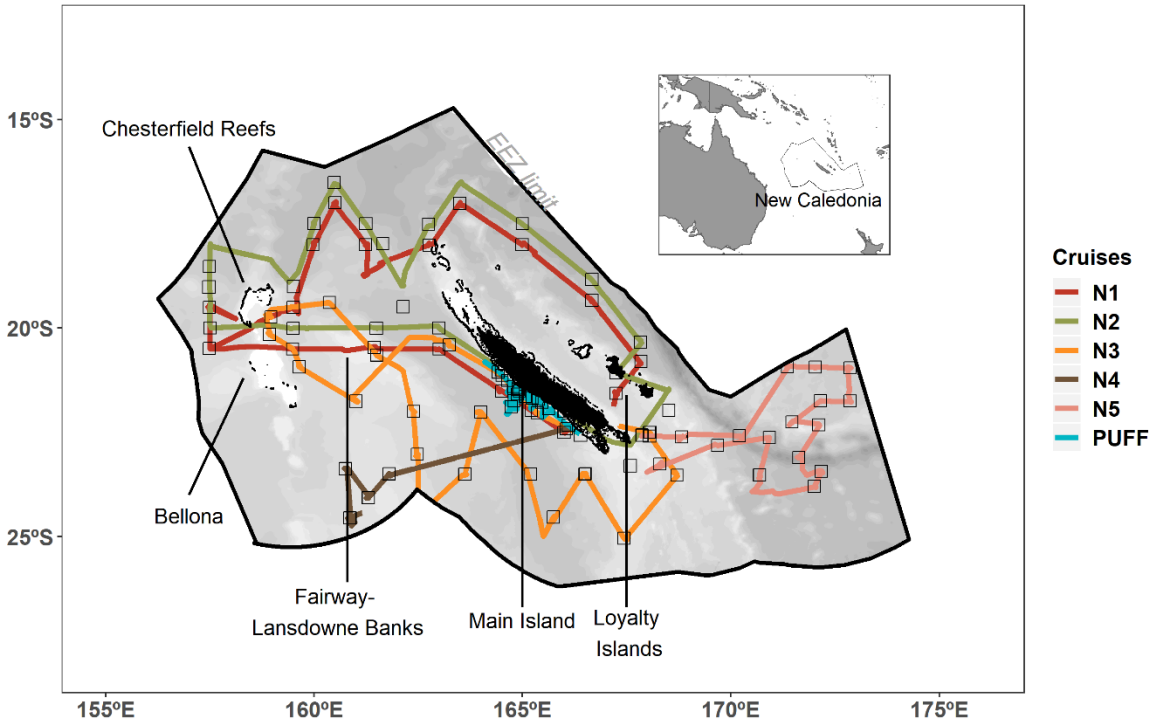


Figure 1: Cruise tracks of the R/V Alis with EK60 echosounder (colored lines) in the New Caledonian Exclusive Economic Zone. Black boxes show CTD stations. The background grey colors represent the relative seabed depth (where lighter colors are shallower). Note that N1 and N2 tracks partially overlap but N2 track has been slightly shifted to the north for visualization purposes.

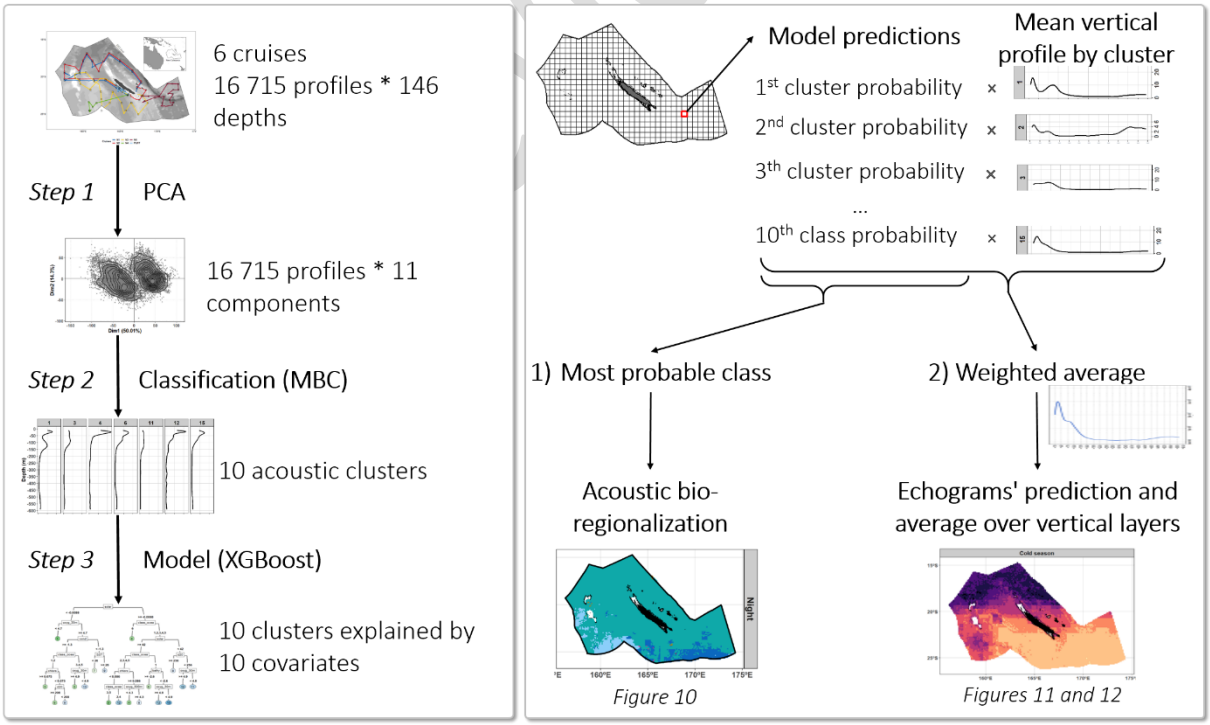


Figure 2: Diagram explaining the different steps of the analysis. Details of the approach are provided in the text.

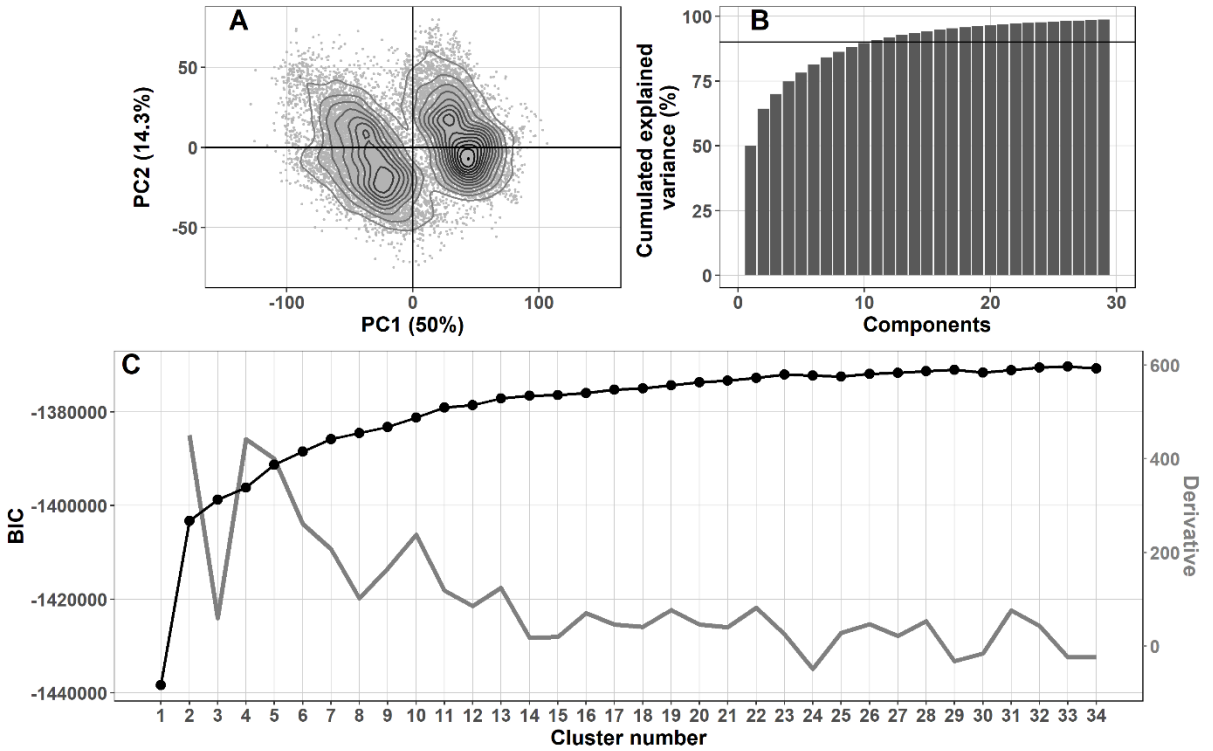


Figure 3: PCA results with the two first axes (panel A) and the cumulative variance explained by the PCA dimensions (panel B). MBC classification results: the BIC (Bayesian Information Criterion) is represented as a function of the potential number of classes in black, with its derivative curve in grey (panel C).

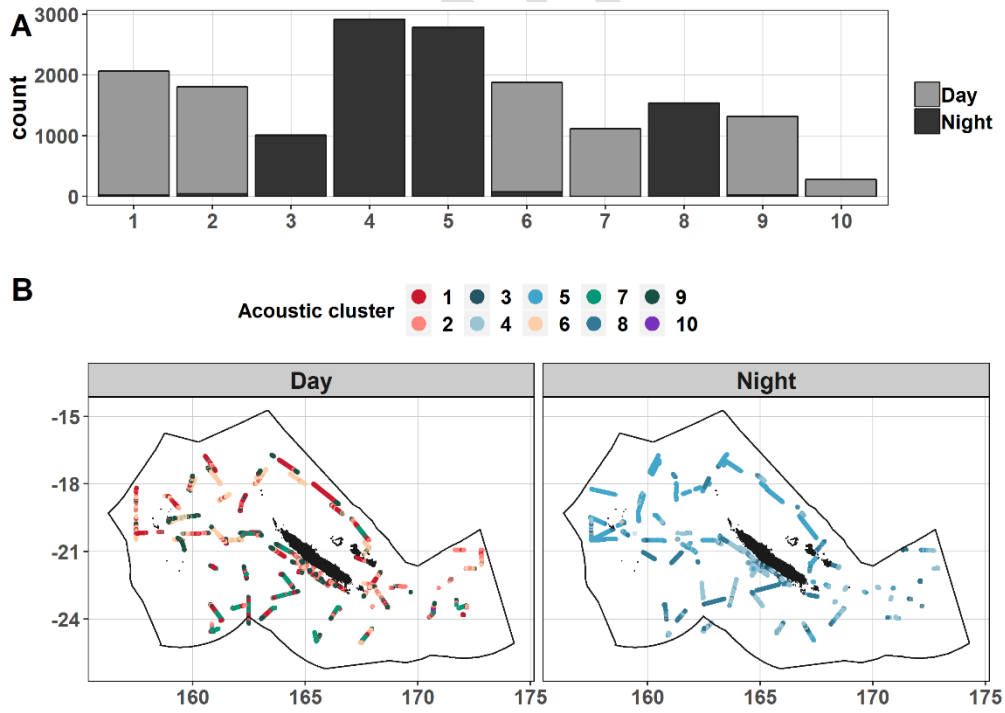


Figure 4: Classification results with the number of day and night vertical profiles in each acoustic cluster (panel A) and the spatial position of all vertical profiles colored by the acoustic cluster they belong to by day (left) and by night (right) (panel B).

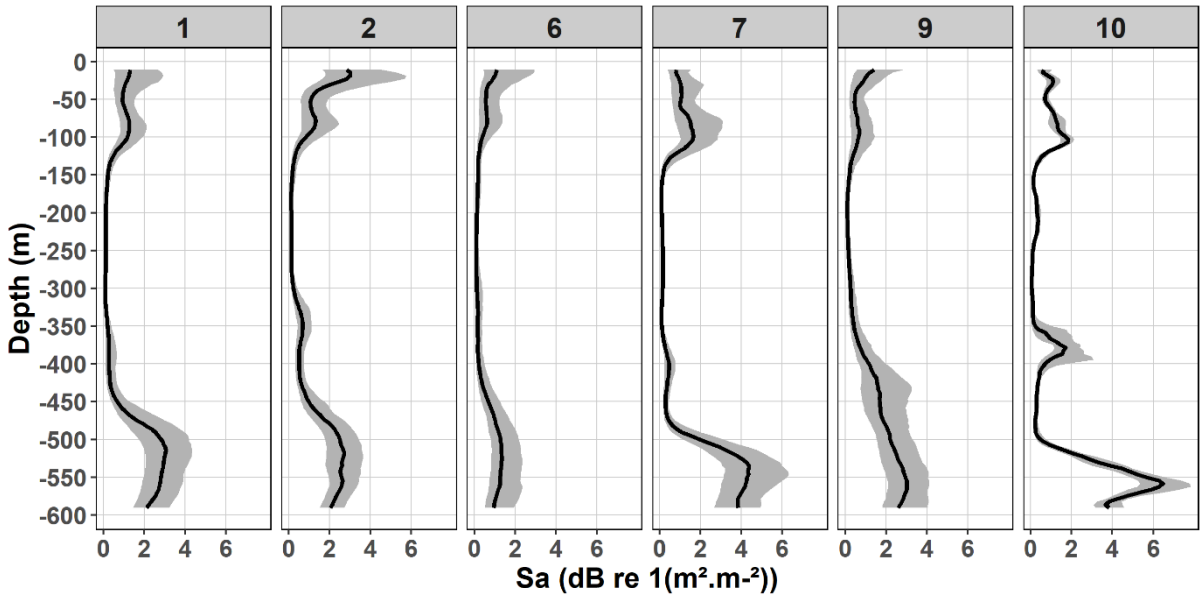


Figure 5: Vertical profile medians for each day acoustic class. The grey ribbon is the interquartile range.

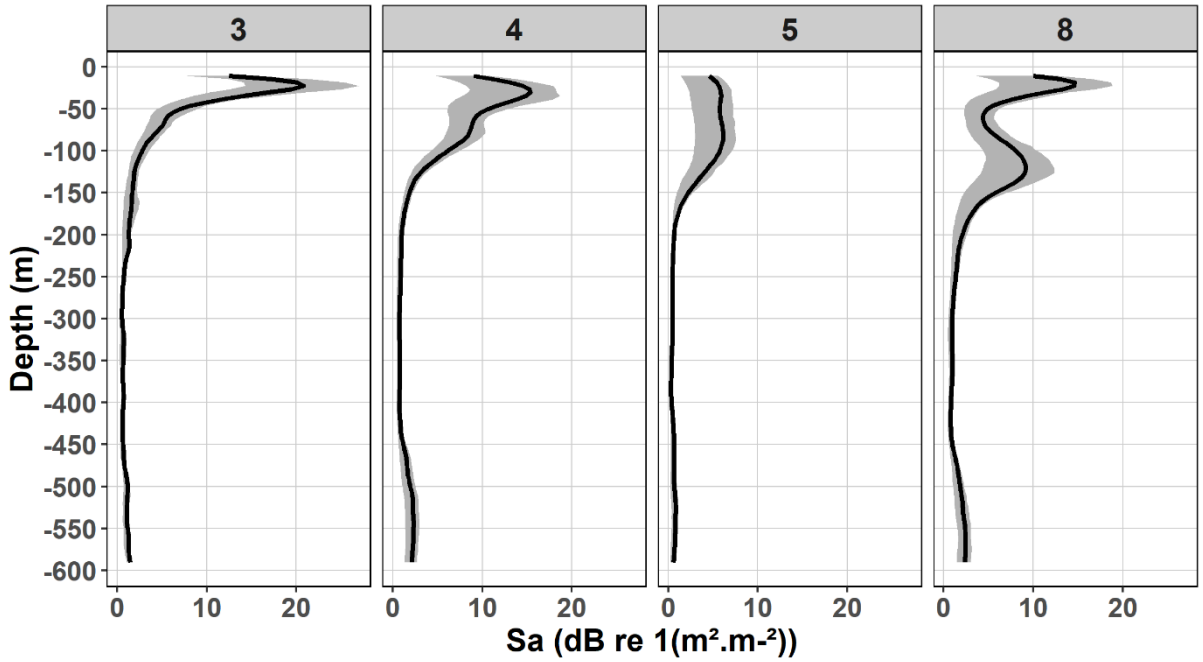


Figure 6: Vertical profile means for each night acoustic class. The grey ribbon is the interquartile range.

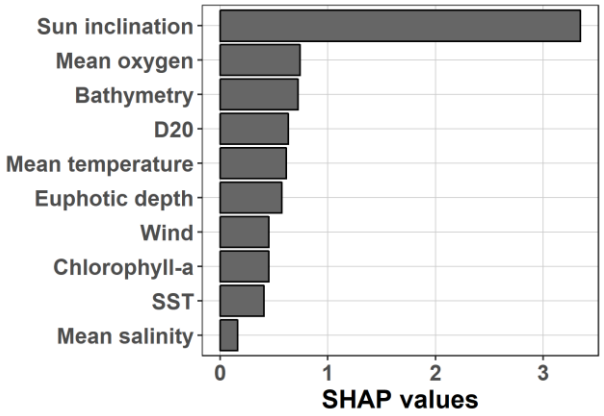


Figure 7: Mean SHAP values for the predictions by each environmental covariate (y-axis).

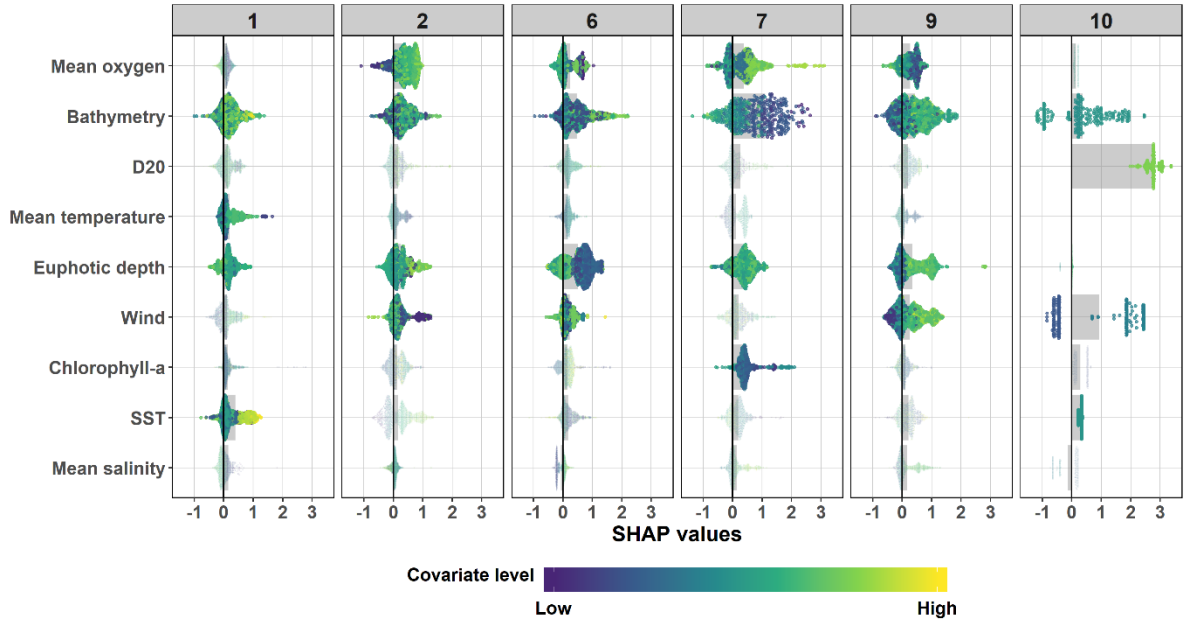


Figure 8: SHAP (SHapley Additive exPlanation) values (*x-axis)* by covariate (*y-axis*) for each day cluster (*columns*). Every observation is one dot on each row. The SHAP value *(x-axis)* represents the influence of a given covariate on the prediction. The dot color represents the covariate normalized value/level: yellow for high value (high normalized SST for example) and dark blue for low value (low normalized SST for example). The height of one patch (the violin shape) gives an indication of the dot density. Grey rectangles by row and by column show the mean SHAP value by cluster and by covariate. Based on these grey rectangles, dots of the four most important covariates by cluster are plotted in brighter colors.

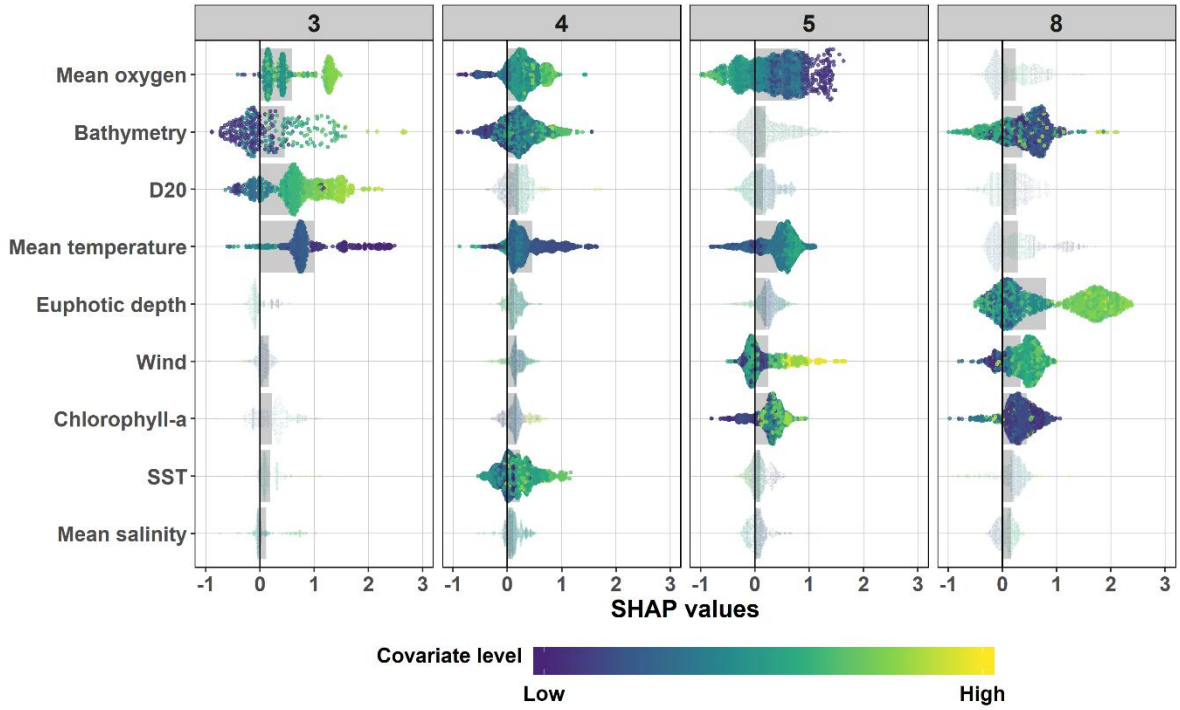


Figure 9: Same as Figure 8 but for the night classes.

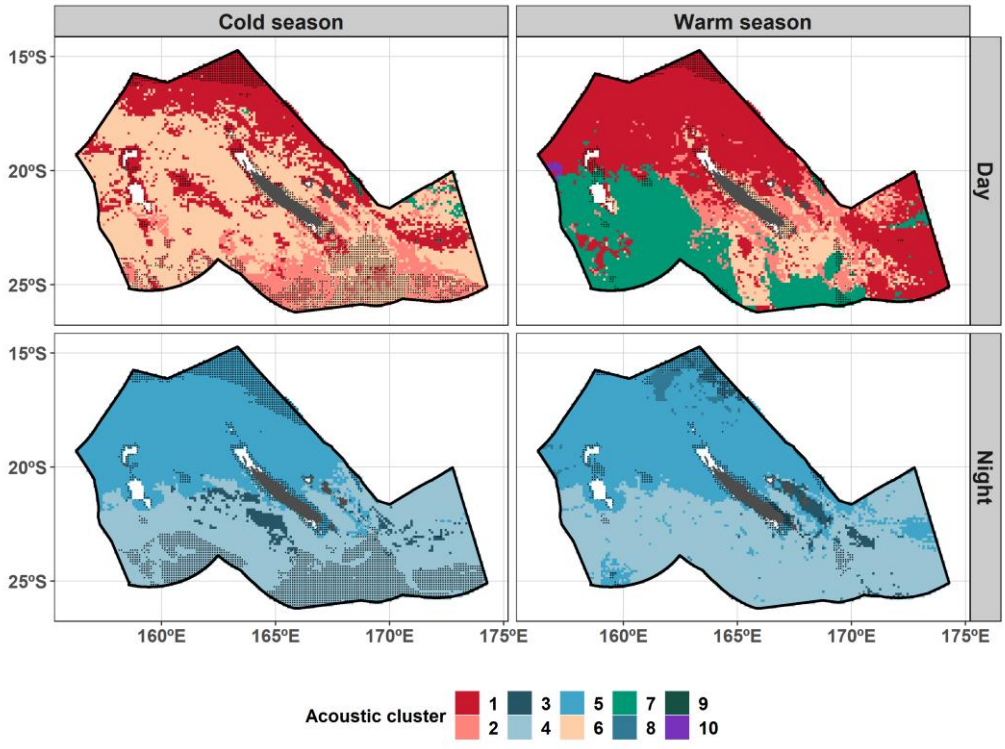


Figure 10: Main cluster predicted for day (1st row) and night (2nd row) during the cold season (left column) and the warm season (right column). Small black dots identify extrapolated points (i.e. where predictions were made with at least one covariate value falling outside of the sampled range). White areas represents non-predicted regions.

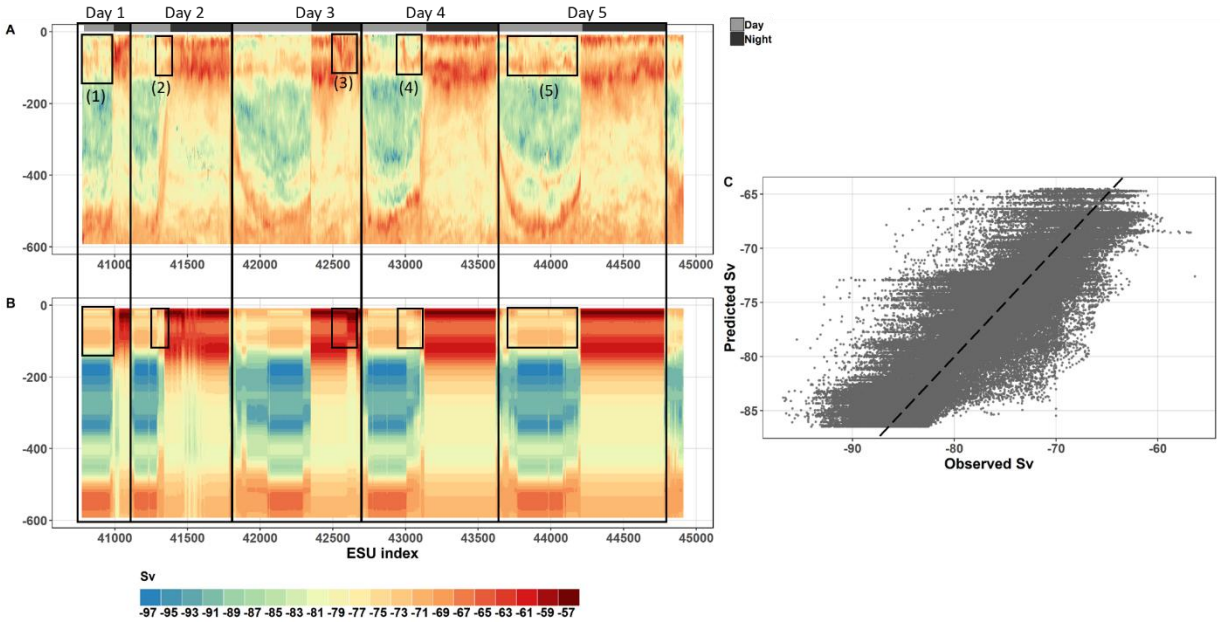
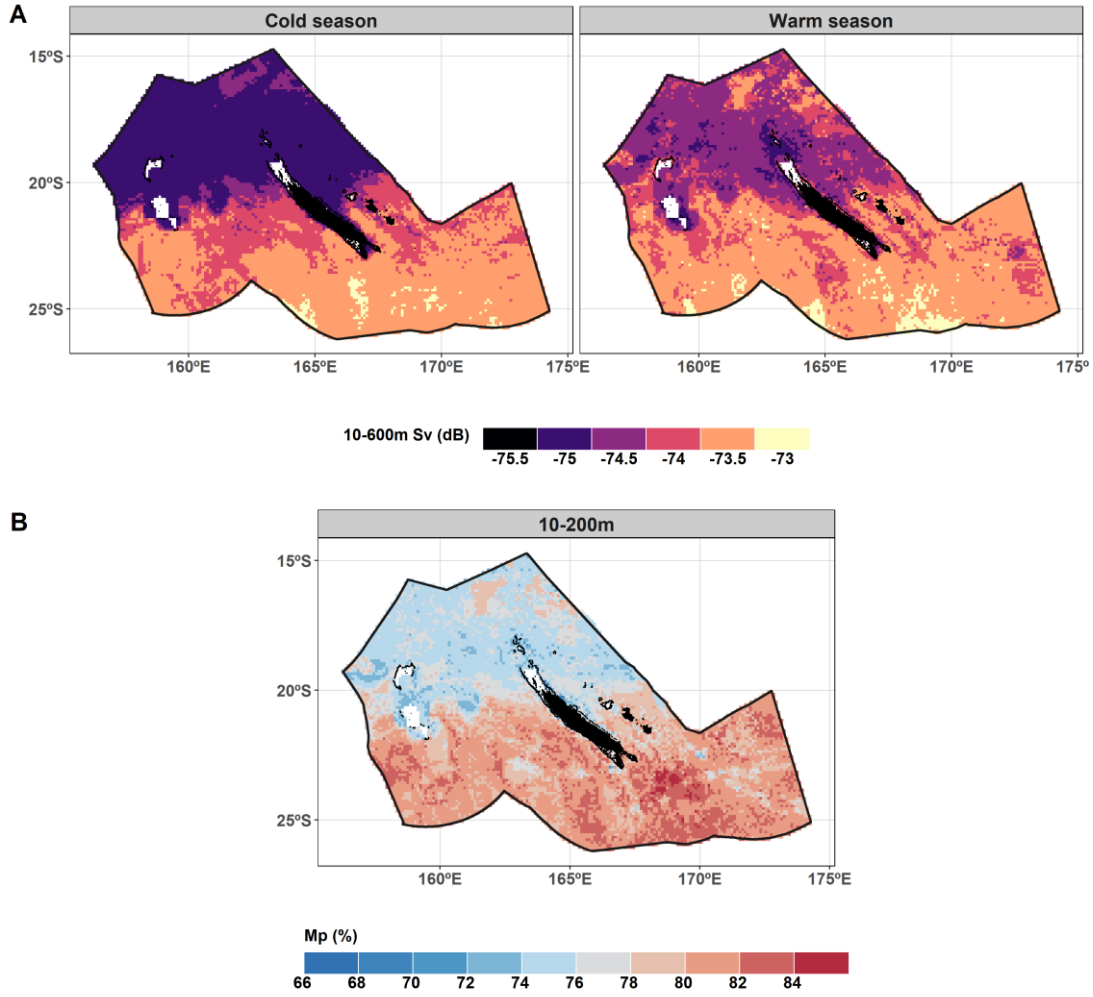


Figure 11: N4 echogram observed (panel A) and predicted (panel B). The scatter plot of predicted values as a function of observed values with the y = x dashed line over all data of N4 (panel C). Boxes drawn on the plots are discussed in the main text as box (1), (2), etc.



1150 Figure 12: Predictions of the S_v averaged over the day and the night and through the entire water
1151 column (10-600m) for the cold season (left) and the warm season (right) (panel A). The ratio of
1152 migrants (%) between the night and the day for the epipelagic layer (10-200m) (panel B).

1153 Table 1: Cruise details, with the cruise name, dates, the number of 0.1nm bins per cruise, and
1154 the DOI of each cruise.

Cruise name	Start	End	Number of 0.1nm bins	DOI
Nectalis 1 (N1)	30/07/2011	15/08/2011	3681	10.17600/11100050
Nectalis 2 (N2)	26/11/2011	14/12/2011	2896	10.17600/11100070
Nectalis 3 (N3)	21/11/2014	08/12/2014	3617	10.17600/14004900
Nectalis 4 (N4)	19/10/2015	25/10/2015	1034	10.17600/15004000
Nectalis 5 (N5)	23/11/2016	06/12/2016	3989	10.17600/16004200
Puffalis (PUFF)	18/03/2017	31/03/2017	1498	10.17600/17003300

1155

1156

1157 Table 2: Details of parameters and formulas used for metric calculations. s_v is the linear
1158 measure of the volume backscattering strength (m^{-1}), z is the depth (m) and all integrals are
1159 calculated between the first depth level (10m) and the deepest depth (600m).

Name	Metric	Formula	Parameters	Unit
Density	Mean volume backscattering strength	$10.log_{10}\left(\frac{\int s_v(z)dz}{H}\right)$	$H = 146$	dB re 1 m^{-1}
Center of mass	Mean vertical localization	$\frac{\int z.s_v(z)dz}{\int s_v(z)dz}$	-	m
Aggregation	Index of aggregation	$\frac{\int s_v(z)^2dz}{(\int s_v(z)dz)^2}$	-	m^{-1}

1160

1161 Table 3: Parameters for profiles. Details of calculations are given in Table 3.

	Day 1	Day 2	Night 3	Night 4	Night 5	Day 6	Day 7	Night 8	Day 9	Day 10
N observations	2065	1805	1010	2919	2787	1878	1116	1538	1320	277
Density (dB)	-75.6	-75.2	-72.7	-71.7	-74.0	-76.8	-75.3	-71.4	-75.0	-75.7
Center of mass (m)	369.3	338.2	144.3	174.7	152.7	339.6	375.0	187.0	385.2	408.5
Aggregation (m^{-1})	1.28	1.18	2.55	1.72	1.63	1.02	1.63	1.36	0.96	1.84

1162

1163