Standardized catch per unit effort of bigeye tuna in the Indian Ocean for the European purse seine fleet operating on floating objects

Giancarlo M. Correa[1,✉], David M. Kaplan[2,3], Jon Uranga[4], Maitane Grande[4], Taha Imzilen[2,3], Gorka Merino[4], and María Lourdes Ramos Alonso[5]

[1] AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Txatxarramendi ugartea z/g, 48395 Sukarrieta (Bizkaia), Spain
[2] MARBEC (Univ. Montpellier, CNRS, Ifremer, IRD), 87 av. Jean Monnet, CS30171, 34203 Sète, France
[3] Institut de Recherche pour le Développement (IRD), 87 av. Jean Monnet, 34203 Sète, France
[4] AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Herrera Kaia, Portualdea z/g, 20110 Pasaia (Gipuzkoa), Spain
[5] Instituto Español de Oceanografía, Centro Oceanográfico de Málaga, Puerto Pesquero, s/n Apdo., 29640, Fuengirola, Málaga, Spain

✉ Correspondence: Giancarlo M. Correa <gmoron@azti.es>

SUMMARY

Indian Ocean EU purse seine floating object (FOB) school catches of big-eye tuna (BET; Thunnus obesus) per fishing set for the period 1991-2023 were standardized with two geostatistical spatiotemporal modelling approaches using the sdmTMB R package. One approach considered only the recent time period 2010-2023, but included detailed covariates describing intensity and use patterns of drifting fish aggregating devices (dFADs) by the fleet. The second approach considered the full time period 1991-2023, but was limited to standardization for vessel size, identifier and mixed layer depth. In both cases, a generalized Gamma model was chosen for modeling catches as this distribution family had the lowest AIC. Predictions were made on an extrapolation area for every time step (year-quarter). To calculate the standardized CPUE index, we aggregated the spatial predictions based on an area-weighting approach. We also presented influence plots to explore the impacts of the model components on the standardized CPUE index. The FOB index from this study showed a long-term negative temporal trend, though over the most recent period (>2010), estimated abundance is more or less stable with a noticeable increase in abundance over the period 2021-2022. The index provided here can be incorporated into the 2025 bigeye stock assessment model to inform changes in biomass of juvenile BET.

KEYWORDS

Mathematical models, Stock assessment, Random effects, Mixed layer depth, Environmental variability, DFADs, Catchability

1. Introduction

An abundance index is a key data input in stock assessment models that can inform fluctuations in population abundance or biomass (Magnusson and Hilborn, 2007). Typically, an abundance index is obtained from fishery-independent (e.g., scientific surveys) and dependent sources. For highly migratory and large pelagic fishes (e.g., tunas), performing a scientific survey is impractical given the large extent of their distribution. Therefore, fishery-dependent abundance indices such as catch per unit effort (CPUE) are primarily used (Hoyle et al., 2024). Using nominal CPUE is inappropriate since it is normally biased due to the spatial heterogeneity of fish populations, environmental factors, the behavior of fishers, and features of fishing vessels (Wilberg et al., 2009). These factors may produce a disparity between the nominal CPUE and true population abundance trends. For this reason, a CPUE standardization process needs to be performed in order to remove the impact of external factors that can influence catch rates (Maunder and Punt, 2004).

The European (EU) tropical tuna purse seine fishery in the Indian Ocean has operated since the 1980's and has high quality, species-specific catch-effort data since 1991. The fleet has experienced significant technological developments in recent years, which have increased its efficiency in locating and catching tunas (Torres-Irineo et al., 2014; Wain et al., 2021). The EU purse seine fleet primarily uses two fishing modes: 1) targeting free-swimming schools, and 2) fishing on schools associated with floating objects (FOBs). The latter category initially used natural objects (e.g., logs) that occurred naturally in the ocean; however, they now use artificial buoys known as fishing aggregating devices (a.k.a. FADs) with incorporated technology (e.g., satellite tracks, echo-sounders) (Lopez et al., 2014). The EU purse seine fleet principally targets three tropical tuna species: yellowfin (*Thunnus albacares*), bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*).

Bigeye tuna (BET; *Thunnus obesus*) is widely distributed in the tropical and subtropical Indian Ocean from Ireland to South Africa in the east and from southern Canada to northern Argentina in the west (Fonteneau et al., 2005; Kaplan et al., 2014). BET remains within the surface layer (50 m) during the night and can dive to depths of 500 m at sunrise (Brill et al., 2005). Spawning takes place in tropical waters when the environment is favorable. From nursery areas in tropical waters, juvenile fish tend to diffuse into temperate waters as they grow. Small bigeye tuna are caught using purse seine gear operating on FOBs in the warm equatorial surface waters, whereas most adults are caught by longliners, indicating vertical stratification in which adult bigeye tuna schools remain at greater depths than do juvenile bigeye tuna schools (Fonteneau and Pallares, 2005). Lan et al. (2018) showed that the depth of the 20°C isotherm, sea surface temperature, and sea height deviation are climate variables significantly related to bigeye abundance in the Atlantic Ocean. Mixed layer depth (MLD) has also been generally shown to impact catchability of tropical tunas by its impact on the vertical distribution of tunas (Kaplan et al., 2024; Lopez et al., 2017). The size of BET caught varies among fisheries: medium to large fish for the longline fishery and purse seine free school sets, small to large for subtropical baitboat fisheries, and small for tropical baitboat, western handline and purse seine FOB fisheries.

The Stock Synthesis assessment platform was used in the 2022 BET assessment (Fu et al., 2022). There were four axes of uncertainty in the final uncertainty grid, which included growth, natural mortality, selectivity, and steepness assumptions. The results of the assessment show that, in 2021, the Indian BET stock was most likely overfished and subject to overfishing. The main index of abundance used in the 2022 BET assessment was derived from the catch and effort data from the longline fishery. A quarterly BET catch per purse-seine FOB set index covering the period 2010-2021 and built using the *VAST* R package (Thorson, 2019a) was also employed in sensitivity runs (Akia et al., 2022).

In this study, we present two new (juvenile) BET abundance indices using data from the EU tropical tuna purse-seine fishery operating on FOBs, one covering the time period 1991-2023 and the other covering the recent time period 2010-2023. The reason for this separation into two indices is that for the recent time period we have access to detailed information on dFAD use that was not available before 2010. Both indices are derived from geostatistical spatiotemporal models built using the *sdmTMB* R package (Anderson et al., 2022). These indices can inform juvenile BET abundance in the assessment process and help to improve the stock assessment model estimates.

2. Methods

2.1 Data

We used logbook data from the EU purse seine fleet (Spain and France) targeting tropical tunas and operating on floating objects in the Indian Ocean from 1991 to 2023. The logbook data sets are managed by the Tuna Observatory (Ob7) and the Spanish Institute of Oceanography (IEO) for the French and Spanish fleets, respectively. The raw logbook data (Level 0) produced by the skippers were corrected in terms of total catch (t) per set to account for the difference between reported catch at sea and landed catch. Likewise, the species composition per set was corrected based on port size sampling and the T3 methodology (Pallarés and Hallier, 1997) to generate a Level 1 logbook data set.

Filtering of the dataset differed somewhat for the two time periods considered in this study due primarily to differences in the variables used for the standardization process and the need to characterize core fishing effort areas over quite different time periods.

2.1.1 2010-2023

For data limited to the time period 2010-2023, we applied the following filters to our initial dataset:

- We excluded null sets (i.e., sets with zero catch of tropical tunas).
- We excluded observations from vessels with less than 5 years of activity in the study period. This filter removed information from 7 vessels (out of 38 vessels), which corresponded to 1781 observations (5.49%).
- We removed observations with missing information for any of the used variables, as well as duplicated rows. This filter excluded 412 rows (< 1%).
- We removed data in areas east of $67.9°$E to avoid a large hole in the distribution due to the closure of the Chagos Archipelago to fishing in 2010. This filter excluded 411 rows (< 1%).
- We removed data in areas north of $15°$N and south of $15°$S. This filter excluded 3985 rows (~ 6%).
- Observations from fishing sets that operated in areas ($1° \times 1°$) that were not fished for less than four years during the studied period were excluded in order to retain areas constantly sampled. This filter removed 316 rows (< 1%).

After applying these filters, we retained 63,823 observations.

2.1.2 1991-2023

For the full dataset (1991-2023), we applied the following filters to our initial dataset:

- We excluded null sets (i.e., sets with zero catch of tropical tunas).
- We limited data to minimum set of vessels representing at least 95% of all FOB fishing sets over the time period. After this filter, the dataset consisted of 133267 FOB fishing sets by 64 distinct fishing vessels, each representing between 677 and 5527 FOB fishing sets.
- Data was then limited to areas west of the Chagos Archipelago ($67.9°$E) to avoid a large hole in the spatio-temporal distribution of fishing effort due to the closure of the Chagos Archipelago to fishing in 2010, and to areas between $15°$N and $15°$S. This filter excluded 10,124 fishing sets (8.2%).
- The resulting dataset was limited to $1° \times 1°$ cells where FOB fishing occurred over at least 10 different years between 1991 and 2023. This filter eliminated 1786 fishing sets (1.5%).
- Finally, we removed any sets outside the spatial zone for which environmental data is available. This filter eliminated 15 fishing sets (0.01%).

The final 1991-2023 dataset consisted of 121342 fishing sets.

2.2 Spatial indicators

Using the observed data, we calculated six indicators to summarize the spatial behavior of the fleet during the studied periods. Diverse spatial indicators have previously been used for fishery-dependent (Kaplan et al., 2021; Russo et al., 2013; Sosa-López and Manzo-Monroy, 2002) and independent (Woillez et al., 2009; Woillez et al., 2007) sources to increase the chance of picking up changes in critical fleet-related factors over time. We calculated the following spatial indicators, which were calculated by year-quarter:

1. *Clark-Evans*: It is an index of point spatial aggregation (Clark and Evans, 1954), here represented by fishing sets, and provides information on how spatially clustered the fishing sets were. Smaller values indicate higher spatial clustering of fishing sets.
2. *Covered area* (km^2): Represents the spatial extent of the fishing sets. It was calculated assuming that each fishing set has an area of influence of 1 km^2, and then calculating the spatial union of those areas.
3. *Center of gravity (lon)*: Indicates the longitude where the BET catches per set were centered.
4. *Center of gravity (lat)*: Indicates the latitude where the BET catches per set were centered.

5. *Moran's autocorrelation coefficient*: Measure of spatial autocorrelation of the BET catch per set (Gittleman and Kot, 1990).
6. *Gini coefficient*: It is a measure of inequality (Cowell, 2011) among BET catch per set values.

2.3 Spatiotemporal model

We used a geostatistical spatiotemporal approach to perform the CPUE standardization. Geostatistical spatiotemporal models (a.k.a. spatiotemporal generalized linear mixed models-GLMM) can account for unmeasured variables (e.g., population biomass) that cause observations (e.g., catch) to be correlated over space and time through random effects (Anderson et al., 2022). A Gaussian random field (GRF) is a multidimensional spatial process where the random effects that describe the spatial pattern follow a multinomial distribution with mean $\mu = [\mu(s_1), \ldots, \mu(s_n)]$ and spatially structured covariance matrix Σ (Blangiardo and Cameletti, 2015).

The *sdmTMB* R package (Anderson et al., 2022) can implement spatiotemporal GLMMs in TMB (Kristensen et al., 2016) for model fitting. *sdmTMB* approximates the GRF by relying on the Stochastic Partial Differential Equation (SPDE) approach using the Integrated Nested Laplace Approximation in *R-INLA* (Bakka et al., 2018) to reduce computational costs. The first step in using the SPDE approach is to construct the mesh, which is composed of triangles covering the studied area with a minimum allowed triangle edge length (*cutoff*) of 1.5 degrees. Following Anderson et al. (2022), our model can be mathematically represented as:

$$\mathbb{E}[y_s] = \mu_s = g^{-1}(\eta_s)$$

$$\eta_s = \mathbf{X_s}\boldsymbol{\beta} + \nu_v + \omega_s + \varepsilon_{s,y} \qquad (1)$$

Where the expected value $\mathbb{E}[.]$ of an observation y (BET catch per set) at coordinates in space s is equal to mean μ_s, which is the result of an inverse link function g^{-1} applied to a linear predictor η_s. β is the vector of coefficients of fixed effects, ν_v is the vessel effect (*numbat*) treated as random effects, and **X** is the model matrix. ω_s is the spatial random field, which is constant across time and represents the effect of latent spatial variables that are not otherwise accounted for in the model:

$$\omega \sim MVN(0, \Sigma_\omega)$$

Where Σ is the covariance matrix of the multivariate normal (MVN) distribution and is constrained by a Matérn function and the spatiotemporal random fields ε_{sy} represent latent variables causing spatial correlation that changes with each time step (i.e., year). ε_{sy} was assumed to be iid (i.e., independent at each year):

$$\varepsilon_y \sim MVN(0, \Sigma_\varepsilon)$$

There is evidence that large BET may perform seasonal migrations for spawning (Cayre et al., 1991), which may impact the spatial pattern of ω_s within a year. Although migrations of small bigeye have not been thoroughly documented yet, a similar behavior might be found. Therefore, we allowed the spatial random field ω_s to vary by quarter q (ω_{sq}) to approximate this behavior.

Since we had a very low number of sets with BET catch equal to zero that remained roughly constant over time (Figure 1), we replaced the zero values with the minimum BET catch value found in our dataset (0.08 t). Then, we tested four probability distributions for the response variable: lognormal, gamma, Tweedie (Shono, 2008), and generalized gamma (Dunic et al., 2025), which were then compared through AIC. We selected the family with the lowest AIC. For the long time series model (1991-2023), the Tweedie model did not converge within a reasonable amount of time so it was discarded as an option for that model.

2.4 Covariates

Table 1 shows all the covariates explored in this study. The source for potential sea surface temperature ($°C$) and mixing layer depth (m) data was the Copernicus-Global Ocean Physics Reanalysis based on the current available real-time global forecasting CMEMS system, having a $0.083° \times 0.083°$ spatial resolution and monthly temporal resolution. Depth-integrated (0-100 m) net primary production ($mg/m^3/day$) was downloaded from Copernicus-

PICES biogeochemical global hindcast with a $0.25^{\circ} \times 0.25^{\circ}$ spatial resolution and monthly temporal resolution. Depth of the 20°C isotherm (m) was downloaded from the NCEP Global Ocean Data Assimilation System with a $0.33^{\circ} \times 1^{\circ}$ spatial resolution and monthly temporal resolution. Covariates associated with the buoy features were calculated from Vessel Monitoring System (VMS) information and labelled position data from acoustic drifting buoys, including vessel and company names, as well as acoustic buoy specifications.

Climatologies and anomalies for environmental variables were calculated by averaging monthly environmental variable maps across years for the entire time series of data available (1993-2023 for SST, MLD and NPPV; 1990-2023 for depth of the 20°C isotherm) and then subtracting that climatology from the time varying maps. These separate variables were used for long time series models due to the fact that Copernicus environmental variables are not available for the period 1991-1992, making the climatology the best available estimate of environmental conditions for these years.

Thorson (2019b) distinguishes between 'catchability' and 'density' covariates: both are included in the linear predictor to explain catch-and-effort data, but only density (and not catchability) covariates are conditioned upon when predicting densities across space (see more details in Section 2.6). This distinction controls for the effect of catchability covariates (i.e., filters out these components of covariation) and conditions upon the effect of habitat covariates (i.e., uses information about habitat covariates to improve performance when predicting population density). We consider that all the oceanographic variables were density covariates except for MLD and MLD anomaly as MLD is known to impact catchability (Kaplan et al., 2024; Lopez et al., 2017). Variables associated with the number of buoys and the buoy's echosounder capacity were considered catchability covariates (see Table 1). Due to large differences in scale among covariates and in order to avoid convergence issues in the model, for the short time series model, we standardized continuous catchability and density covariates to a mean of 0 and standard deviation of 1 (z-score standardization), whereas for the long time series model we divided environmental covariates into climatology and anomaly. Also, we calculated variance inflation factors (VIF) from a simple linear regression model to explore multicollinearity among the candidate covariates. We removed covariates with a VIF larger than 5 from our analysis.

2.5 Model selection

To choose the distribution families for model residuals, we first ran simple models with just year and quarter (*yr_qtr*) as an explanatory variable (in addition to a quarterly-varying spatial random field and a yearly spatiotemporal random field), one such model for each potential distribution family and time series (long and short). The AIC of these models was calculated and the model with the lowest AIC was chosen for all future modeling work.

We next ran a *full configuration* model, which included all the covariates assuming a linear relationship with the response variable. We did not explore nonlinear relationships since an exploratory analysis did not show evident nonlinear associations between the response variable and the continuous covariates. Then, we identified those covariates that did not have a significant effect (p-value > 0.05) and removed them from the model. Then, the model was run again only including the significant covariates. The interaction between year and quarter (*yr_qtr*) was the only explanatory variable that was never removed.

Once the final model was identified, we then used the *DHARMa* R package (Hartig, 2022) to evaluate the model residuals. Standard raw residuals are not always appropriate when using generalized linear models, and other types of residuals are commonly used instead. *DHARMa* uses a simulation-based approach to create readily interpretable scaled (quantile) residuals for generalized linear mixed models. We analyzed two plots produced by *DHARMa*: 1) the QQ plot residuals, which detects overall deviations from the expected distribution, and 2) the residual vs. predicted plot, which detects trends in residuals along model predictions and simulation outliers. We also evaluated significant spatial autocorrelation of MCMC-based randomized quantile residuals using the Moran's I statistic (Moran, 1950).

2.6 Standardized CPUE calculation

Previous index standardization methods have typically involved fitting a regression model including a year intercept and covariates, and then treating the year intercept as the abundance index. This approach implicitly treated all covariates as "catchability covariates", even when these variables were likely associated with increases in local population density (Thorson, 2019b). In this study, we calculated the standardized CPUE index by year-quarter using the "predict-then-aggregate" approach (Hoyle et al., 2024). This approach consists of making predictions over an extrapolation area composed of $1° \times 1°$ grids, which was kept constant for every time step. For catchability covariates, we fixed continuous covariates at their mean values (0 in the case of standardized values and for MLD anomaly), while discrete covariates were fixed at the level with the largest sample size. For density covariates, we assigned values at every grid centroid across the modeled spatial (i.e., extrapolation area) and temporal domain.

Then, we aggregated the predicted values $\widehat{CPUE}_{y,q,a}$ (a corresponds to grids in the extrapolation area) over space using an area-weighted approach:

$$\widehat{CPUE}_{y,q} = \sum_{a} A_a \times \widehat{CPUE}_{y,q,a} \qquad (2)$$

Where A_a is the ocean area (km^2) of grid a. For the short time series, indices were scaled to mean 1, whereas for the long time series indices were scaled so that they represent catch in tonnes per set.

2.7 Influence analysis

We used influence plots to understand how the covariates and their values affect the calculated CPUE index. Bentley et al. (2012) initially proposed these exploratory plots for CPUE standardizations using generalized linear models. Then, Hsu et al. (2022) adapted these plots to explore the influence of the spatial (ω) and spatiotemporal (ϵ) terms in geostatistical spatiotemporal models. For the spatial term, we first calculated the normalized coefficient (ρ_ω):

$$\rho_\omega = \frac{\sum_i \omega_i}{n} \qquad (3)$$

Where ω_i is the estimated spatial term value corresponding to observation i and n is the number of observations. Then, the mean difference between the coefficients corresponding to all observations in year y was calculated:

$$\delta_{\omega_y} = \frac{\sum_{i=1}^{n_y} (\omega_i - \rho_\omega)}{n_y} \qquad (4)$$

Where n_y is the number of observations in year y. Then, since the log-link function was used, the annual influence value in year y can be calculated: $I_{\omega_y} = exp\left(\delta_{\omega_y}\right)$. These influence metrics can also be calculated over space. However, since there were > 200 knots in the used mesh, calculating a coefficient by knot would be impractical. For that reason, we identified 8 knot groups with similar ω quarter-1 values using a partitioning around medoids clustering (Hennig and Liao, 2013), which considered spatial proximity. Then, the normalized coefficient and influence for ω per knot group k (ρ_ω and I_{ω_k}) were calculated using the steps described above. A similar procedure was followed to calculate the influence of the spatiotemporal term.

In order to evaluate the influence of catchability and density covariates, we plotted the changes in the CPUE index produced by the inclusion of each significant covariate ('step plot'). Whereas for the short time series, all covariates were included in this stepwise process, including the vessel random effect and spatial and spatiotemporal random fields, for the long time series, only fixed model covariates (i.e., country, vessel capacity and environmental covariates) were included due to the long run times needed for this model.

3. Results

The number of sets per quarter included in our models increased from 1991 to ~2017, but has stabilized somewhat in recent years (Figure 1; results for the short time series, not shown, were similar). The values of BET catch per set were skewed to the left, with values generally smaller than 15 tons and rarely above 30 tonnes (Figure 2). In the log scale, we noticed that the BET catch per set values remained roughly stable over the years, with somewhat larger values in the 1990's and since 2020. When comparing among quarters, we noticed that quarter 1 generally had larger catch per set values (Figure 2). The proportion of sets with BET catch equal to zero remained below 1%, except for the first year of the time series (1991) and a few exceptional quarters over the 33 year time series (Figure 2).

The fishing sets occur primarily around the equator, between 5°S and 5°N (Figure 3 & Figure 4). We did not observe a clear spatial pattern in average BET catch per set values, although they were generally larger in areas far from the coast (Figure 5 & Figure 6). We did not observe a clear spatial pattern in the proportion of null sets (Figure 7 & Figure 8).

3.1 Spatial indicators

The covered area (within the core purse-seine fishing area considered for this standardization) expanded progressively over the years up to ~2018, as is to be expected given the increase in the number of FOB sets per year over time (Figure 9; indices for the short time series, not shown, showed similar patterns for the overlap between the two series). The Clark-Evans indices suggest that the clustering of fishing sets decreased (larger Clark-Evans index) up until about 2010, after which time it stabilized (Figure 9). The center of gravity (longitude) has moved notably eastward since ~2010 (Figure 9). In terms of latitude, the center of gravity for quarters 1 and 2 has moved steadily northwards since ~2010, but remained stable for quarters 3 and 4 (Figure 9). The Moran index indicated that the BET catch per set values did not largely change their spatial autocorrelation over the years except for a few exceptional quarters 1 and 2 in the early 2000's (Figure 9), potentially driven by small-scale targetting during the "golden years" (Fonteneau et al., 2008). Finally, the Gini index indicated that the heterogeneity of BET catch per set values remained roughly stable over time except for a few exceptional values, particularly for quarter 2, before 2005 (Figure 9).

3.2 Spatiotemporal model

3.2.1 Short time series (2010-2023)

When including all the candidate covariates in a simple linear model for the short time series (2010-2023), we found that the variance inflation factors associated with each of them were not larger than 5, which suggests that multicollinearity was not an issue (Figure 10). However, we found that the total net primary production (*nppv*) and sea surface temperature were highly correlated (Figure 11), therefore, we did not use *nppv* in the model selection process. The nodes of the defined mesh for the spatiotemporal model are shown in Figure 12, which covered the entire distribution of the observations. When comparing the AIC among different statistical families, we found that the generalized gamma family had the best performance (Table 2).

We found that all the tested variables had a significant, although small, effect on the response variable (Table 3). The variable with the largest effect was the set time from sunrise (*t_sunrise*), which was negative and suggests that catch rates are higher when fishing closer to sunrise. Also, the *country* variable had a relatively important effect, indicating that catch rates are higher for the Spanish fleet.

Although the KS and dispersion tests were not passed, the pattern of simulation residuals did not show large deviations from the expected distribution (Figure 13). Simulation outliers were not observed. Using randomized quantile residuals, the Moran's I p-values suggest that there was no spatial autocorrelation in residuals (Figure 14).

The spatial term showed larger values in zones far from coast and negative values in northern areas and south of the equator closer to coast, a pattern that was present for all the quarters (Figure 15). This pattern was similar among quarters. The spatiotemporal term is shown in Figure 16 and generally showed, like the spatial term, negative values in northern areas and larger values in areas far from coast.

The extrapolation area is shown in Figure 17. We predicted CPUE values for every grid in this area and quarter by fixing the catchability covariates. Generally, the predicted CPUE shows larger values in areas far from coast in the southern hemisphere (Figure 18). We aggregated these spatial predictions to obtain a quarterly index of abundance. The CPUE index did not show a trend over the years (Figure 19 and Table 4) and with standardized values very similar to the nominal CPUE. There were years with particularly high values, such as 2010-Q1, 2011-Q1, or 2021-Q1. Generally, we noticed that the first quarter produced larger values compared to quarters 2-4.

When evaluating the influence of the spatial and spatiotemporal terms on the model, we found that the spatial term for knot groups 2 and 3 (Figure 20, which coincided with those areas with generally higher predicted CPUE) had a positive influence on predicted values. Conversely, knot group 5 (northern area) had a negative influence on predicted values. Over the years, we only found a large influence coefficient for the spatiotemporal term for 2010-2011 (Figure 20). The step plot helped us to understand the impacts of each covariate on the quarterly CPUE index. Overall, we did not find very large impacts of any covariate (Figure 21). Also, we noticed that the inclusion of the spatial and spatiotemporal terms slightly impacted the predicted CPUE index for some quarters.

3.2.2 Long time series (1991-2023)

When including all candidate covariates in a simple linear model for the long time series (1991-2023), we found that the variance inflation factors associated with each of them were not larger than 5, which suggests that multicollinearity was not an issue (Figure 23). However, as with the short time series, we found that the total net primary production (*nppv*) and sea surface temperature were highly correlated (Figure 24), therefore, we did not use *nppv* in the model selection process. We also found that the depth of the 20°C isotherm had important correlations with the other environmental variables (Figure 24). Given these correlations and the relatively long run times required for the long time series model, we decided not to include the depth of the 20°C isotherm in long time series models.

The nodes of the defined mesh for the spatiotemporal model are shown in Figure 25, which covered the entire distribution of the observations. When comparing the AIC among different statistical families, we found that the generalized gamma family had the best performance (Table 5). Note that the long time series model based on the Tweedie distribution did not converge in a practical amount of time, so this distribution was not considered for our analysis.

As with the short time series model, we found that all the tested variables had a significant effect on the response variable (Table 6). Although the KS and dispersion tests were not passed, the pattern of simulation residuals did not show large deviations from the expected distribution (Figure 26). Simulation outliers were not observed. Using randomized quantile residuals, the Moran's I p-values suggest that there was no spatial autocorrelation in residuals (Figure 27).

The spatial term showed similar patterns to those of the short time series model, with larger values in zones far from coast and in the center of the domain, and negative values in northern areas and south of the equator closer to coast (Figure 28). This pattern was similar among quarters. The spatiotemporal term is shown in Figure 29 and generally showed, like the spatial term, negative values in northern areas and larger values in areas far from coast.

The extrapolation area for the long time series model is shown in Figure 30. We predicted CPUE values for every grid in this area and quarter by fixing the catchability covariates. We aggregated these spatial predictions to obtain a quarterly index of abundance. The CPUE index showed a long-term decreasing trend over the years (Figure 31, Figure 32 and Table 7), though as with the short time series, the index is roughly stable over the most recent 13 years. Values are similar to the nominal CPUE, though there are some important corrections 1995-2003. The long and short times series indices are very close over the years that they overlap (Figure 33).

When evaluating the influence of the spatial and spatiotemporal terms on the model, we found little strong trend in the influence of the spatial term for the different knot groups (Figure 34). Knot groups 3 and 4, corresponding to the northern extremes of the model domain, are relatively poorly represented in the early parts of the time series. The influence of the spatiotemporal terms has varied over the years, but appears to have a long term tendency towards smaller values (Figure 34). Overall, we did not find very large impacts of any covariate on the index (Figure 35, Figure 36 and Figure 37).

3.2.3 Online availability of indices

The short and long time series indices presented in this paper can be downloaded at:

- Quarterly short time series index: https://drive.ird.fr/s/MaKtm4kDEDPmozY
- Quarterly long time series index: https://drive.ird.fr/s/3DX3xxpJZDwsHRc
- Annual long time series index: https://drive.ird.fr/s/LnnD5soatW9cZNo

4. Discussion

In this study, we used a geostatistical spatiotemporal model to standardize the BET catch rates of the EU purse seine fleet operating on floating objects (FOB) in the Indian Ocean. There is evidence that spatiotemporal models like the one implemented here outperform alternative modelling strategies (Grüss et al., 2019), and their use has shown potential for CPUE standardization for tropical tunas using data from purse seine sets associated with FOBs (Xu and Lennert-Cody, 2022) and dolphins (Xu et al., 2019). We tested different model configurations and catchability and density covariates. Our final models showed good performance and produced indices with a long-term negative trend over the years since 1991, but approximate stability over the most recent 13 year period. These indices can help inform the abundance of juvenile BET in stock assessment models.

We have produced two time series, one for the time period since 2010 that includes sophisticated covariates accounting for the use of dFADs by vessels, and another since 1991 for which these covariates are unfortunately not available. In our model, we included a covariate (follow) associated with the echosounder capacity of the buoys, which has been shown to increase the fishing power of this fleet (Wain et al., 2021). We found that dFAD use covariates generally had significant impacts on BET, but there impacts were often contradictory and quite small. One possible explanation for this is that while echosounder-equipped tracking buoys have been shown to increase dFAD catch per set, they have also been shown to favor increased catch of skipjack over other species (Wain et al., 2021). Though the magnitude of this latter effect is small, BET is also the rarest of the three major tunas in FOB catches, perhaps explaining the equivocal impact of dFAD use on their catches.

Regarding density covariates, SST was found to be the most important environmental explanatory variable for both the long and the short time series. Though temperature is known to impact the horizontal and vertical distributions of BET (Cai et al., 2020; Hino et al., 2019), the relatively minor impact of MLD is potentially linked to the confounding effects of set time of day, with even juvenile BET known to perform deep vertical migrations during daylight hours (Hino et al., 2019).

One variable that might be included in future standardizations is the number of days at sea of the buoy on which the fishing set was performed. Previous studies have shown that tuna colonize dFADs during the first 16 days after release in the ocean (Baidai et al., 2020). Therefore, fishing on a dFAD that has been just released might produce lower catch rates, and the opposite effect might be expected when fishing on a fully-colonized dFAD. However, accurately calculating dFAD time at sea is currently quite difficult given that only dFAD tracking buoys, but not dFADs themselves, have a tracking system, and access to dFAD trajectory data remains limited. Furthermore, it is possible that the effect of time at sea for *fished dFADs* may have little impact given that purse-seine vessels only fish on floating objects at which tunas are present (i.e., the object is already colonized).

Finally, we also compared the FOB index obtained in this study with the longline indices used in the 2022 BET stock assessment in areas 1N and 1S, which correspond to the western Indian Ocean (Figure 39). Overall, we noticed a similar trend between both indices (FOB and LL), especially the decrease from the late 1990s to mid 2000s. Then, both indices showed an increase in CPUE values around 2010, and then a decrease until 2020 approximately.

5. References

Akia, S., Guéry, L., Grande, M., Kaplan, D.M., Baéz, J.C., Ramos, M.L., Uranga, J., Abascal, F.J., Santiago, J., Merino, G., Gaertner, D., 2022. European purse seiners CPUE standardization of Big Eye tuna caught under dFADs (No. IOTC-2022-WPTT24-12). IOTC 24th Working Party on Tropical Tunas (WPTT) Data Preparatory Meeting, Online/Virtual.

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A.K., 2022. sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. https://doi.org/10.1101/2022.03.24.485545

Baidai, Y., Dagorn, L., Amandè, M.J., Gaertner, D., Capello, M., 2020. Tuna aggregation dynamics at Drifting Fish Aggregating Devices: A view through the eyes of commercial echosounder buoys. ICES Journal of Marine Science 77, 2960–2970. https://doi.org/10.1093/icesjms/fsaa178

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D., Lindgren, F., 2018. Spatial modeling with R-INLA: A review. WIREs Computational Statistics 10, e1443. https://doi.org/10.1002/wics.1443

Bentley, N., Kendrick, T.H., Starr, P.J., Breen, P.A., 2012. Influence plots and metrics: Tools for better understanding fisheries catch-per-unit-effort standardizations. ICES Journal of Marine Science 69, 84–88. https://doi.org/10.1093/icesjms/fsr174

Blangiardo, M., Cameletti, M., 2015. Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley and Sons, Inc, Chichester, West Sussex.

Brill, R.W., Bigelow, K.A., Musyl, M.K., Fritsches, K.A., Warrant, E.J., 2005. Bigeye tuna (Thunnus obesus) behavior and physiology and their relevance to stock assessments and fishery biology. Col Vol Sci Pap ICCAT 57, 42–161.

Cai, L., Xu, L., Tang, D., Shao, W., Liu, Y., Zuo, J., Ji, Q., 2020. The effects of ocean temperature gradients on bigeye tuna (*Thunnus Obesus*) distribution in the equatorial eastern Pacific Ocean. Advances in Space Research 65, 2749–2760. https://doi.org/10.1016/j.asr.2020.03.030

Cayre, P., Kothias, J.B.A., Diouf, T., Stretta, J.M., 1991. Biologia de los atunes (No. 37). ICCAT (International Commission for the Conservation of Atlantic Tunas).

Clark, P.J., Evans, F.C., 1954. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. Ecology 35, 445–453. https://doi.org/10.2307/1931034

Cowell, F., 2011. Measuring Inequality. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199594030.001.0001

Dunic, J.C., Conner, J., Anderson, S.C., Thorson, J.T., 2025. The generalized gamma is a flexible distribution that outperforms alternatives when modelling catch rate data. ICES Journal of Marine Science 82, fsaf040. https://doi.org/10.1093/icesjms/fsaf040

Fonteneau, A., Ariz, J., Delgado, A., Pallares, P., Pianet, R., 2005. A comparison of bigeye stocks and fisheries in the Atlantic, Indian and Pacific Oceans. Col Vol Sci Pap ICCAT 57, 41–66.

Fonteneau, A., Lucas, V., Tewkai, E., Delgado, A., Demarcq, H., 2008. Mesoscale exploitation of a major tuna concentration in the Indian Ocean. Aquatic Living Resources 21, 109–121. https://doi.org/10.1051/alr:2008028

Fonteneau, A., Pallares, P., 2005. Tuna natural mortality as a function of their age: The bigeye tuna (Thunnus obesus) case. ICCAT Col. Vol. Sci. Pap 57, 127–141.

Fu, D., Merino, G., Winker, H., 2022. Preliminary Indian Ocean bigeye tuna stock assessment 1950-2021 (stock synthesis) (No. IOTC–2022–WPTT24–10). IOTC 24th Working Party on Tropcial Tunas, Victoria, Seychelles.

Gittleman, J.L., Kot, M., 1990. Adaptation: Statistics and a Null Model for Estimating Phylogenetic Effects. Systematic Zoology 39, 227. https://doi.org/10.2307/2992183

Grüss, A., Walter, J.F., Babcock, E.A., Forrestal, F.C., Thorson, J.T., Lauretta, M.V., Schirripa, M.J., 2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fisheries Research 213, 75–93. https://doi.org/10.1016/j.fishres.2019.01.008

Hartig, F., 2022. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.

Hennig, C., Liao, T.F., 2013. How to Find an Appropriate Clustering for Mixed-Type Variables with Application to Socio-Economic Stratification. Journal of the Royal Statistical Society Series C: Applied Statistics 62, 309–369. https://doi.org/10.1111/j.1467-9876.2012.01066.x

Hino, H., Kitagawa, T., Matsumoto, T., Aoki, Y., Kimura, S., 2019. Changes to vertical thermoregulatory movements of juvenile bigeye tuna (Thunnus obesus) in the northwestern Pacific Ocean with time of day, seasonal ocean vertical thermal structure, and body size. Fisheries Oceanography 28, 359–371. https://doi.org/10.1111/fog.12417

Hoyle, S.D., Campbell, R.A., Ducharme-Barth, N.D., Grüss, A., Moore, B.R., Thorson, J.T., Tremblay-Boyer, L., Winker, H., Zhou, S., Maunder, M.N., 2024. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research 269, 106860. https://doi.org/10.1016/j.fishres.2023.106860

Hsu, J., Chang, Y.-J., Ducharme-Barth, N.D., 2022. Evaluation of the influence of spatial treatments on catch-per-unit-effort standardization: A fishery application and simulation study of Pacific saury in the Northwestern Pacific Ocean. Fisheries Research 255, 106440. https://doi.org/10.1016/j.fishres.2022.106440

Kaplan, D., Baez, J.C., Pascual, P., Vidal, T., 2021. Temporal trends and variability in the spatial distribution of European tropical tuna purse-seine fishing in the Atlantic and Indian Oceans (No. IOTC-2021-WPTT23-20_Rev1). Indian Ocean Tuna Comission.

Kaplan, D.M., Chassot, E., Amandé, J.M., Dueri, S., Demarcq, H., Dagorn, L., Fonteneau, A., 2014. Spatial management of Indian Ocean tropical tuna fisheries: Potential and perspectives. ICES Journal of Marine Science 71, 1728–1749. https://doi.org/10.1093/icesjms/fst233

Kaplan, D.M., Correa, G.M., Ramos Alonso, M.L., Duparc, A., Uranga, J., Santiago, J., Floch, L., Baez, J.C., Rojo Méndez, V., Pascual Alayón, P.J., Merino, G., 2024. Standardized CPUE abundance indices for adult yellowfin tuna caught in free-swimming school sets by the European purse-seine fleet in the Indian Ocean, 1991-2022 (No. IOTC-2024-WPTT26(DP)-13rev2). IOTC Working Party on Tropical Tunas (WPTT26) Data Preparatory Meeting, Online.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B.M., 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70. https://doi.org/10.18637/jss.v070.i05

Lan, K.-W., Lee, M.-A., Chou, C.-P., Vayghan, A.H., 2018. Association between the interannual variation in the oceanic environment and catch rates of bigeye tuna (Thunnus obesus) in the Atlantic Ocean. Fisheries Oceanography 27, 395–407. https://doi.org/10.1111/fog.12259

Lopez, J., Moreno, G., Lennert-Cody, C., Maunder, M., Sancristobal, I., Caballero, A., Dagorn, L., 2017. Environmental preferences of tuna and non-tuna species associated with drifting fish aggregating devices (DFADs) in the Atlantic Ocean, ascertained through fishers' echo-sounder buoys. Deep Sea Research Part II: Topical Studies in Oceanography, Future of oceanic animals in a changing ocean 140, 127–138. https://doi.org/10.1016/j.dsr2.2017.02.007

Lopez, J., Moreno, G., Sancristobal, I., Murua, J., 2014. Evolution and current state of the technology of echo-sounder buoys used by Spanish tropical tuna purse seiners in the Atlantic, Indian and Pacific Oceans. Fisheries Research 155, 127–137. https://doi.org/10.1016/j.fishres.2014.02.033

Magnusson, A., Hilborn, R., 2007. What makes fisheries data informative? Fish and Fisheries 8, 337–358. https://doi.org/10.1111/j.1467-2979.2007.00258.x

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

Moran, P.A.P., 1950. Notes on Continuous Stochastic Phenomena. Biometrika 37, 17. https://doi.org/10.2307/2332142

Pallarés, P., Hallier, J.P., 1997. Analyse du schéma d'échantillonnage multispécifique des thonidés tropicaux (Rapport {{Scientifique}} No. OEP/ORSTOM, Programme n° 95/37 réalisé avec le soutien financier de la Commission des Communautés Européennes). Madrid, Spain.

Russo, T., Parisi, A., Cataudella, S., 2013. Spatial indicators of fishing pressure: Preliminary analyses and possible developments. Ecological Indicators 26, 141–153. https://doi.org/10.1016/j.ecolind.2012.11.002

Shono, H., 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. Fisheries Research 93, 154–162. https://doi.org/10.1016/j.fishres.2008.03.006

Sosa-López, A., Manzo-Monroy, H., 2002. Spatial patterns of the yellowfin tuna (Thunnus albacares) in the Eastern Pacific Ocean: An exploration of concentration profiles. Ciencias Marinas 28, 331–346. https://doi.org/10.7773/cm.v28i4.241

Thorson, J.T., 2019b. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210, 143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Thorson, J.T., 2019a. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fisheries Research 210, 143–161. https://doi.org/10.1016/j.fishres.2018.10.013

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

Wain, G., Guéry, L., Kaplan, D.M., Gaertner, D., 2021. Quantifying the increase in fishing efficiency due to the use of drifting FADs equipped with echosounders in tropical tuna purse seine fisheries. ICES Journal of Marine Science 78, 235–245. https://doi.org/10.1093/icesjms/fsaa216

Wilberg, M.J., Thorson, Linton, and Berkson, J., 2009. Incorporating Time-Varying Catchability into Population Dynamic Stock Assessment Models. Reviews in Fisheries Science 18, 7–24. https://doi.org/10.1080/10641260903294647

Woillez, M., Poulard, J.-C., Rivoirard, J., Petitgas, P., Bez, N., 2007. Indices for capturing spatial patterns and their evolution in time, with application to European hake (Merluccius merluccius) in the Bay of Biscay. ICES Journal of Marine Science 64, 537–550. https://doi.org/10.1093/icesjms/fsm025

Woillez, M., Rivoirard, J., Petitgas, P., 2009. Notes on survey-based spatial indicators for monitoring fish populations. Aquatic Living Resources 22, 155–164. https://doi.org/10.1051/alr/2009017

Xu, H., Lennert-Cody, C.E., 2022. Standardizing the purse-seine indices of abundance and associated length compositions for skipjack tuna in the eastern Pacific Ocean (No. SAC-13 INF-K). Inter-American tropical tuna commission scientific advisory committee, Online/Virtual.

Xu, H., Lennert-Cody, C.E., Maunder, M.N., Minte-Vera, C.V., 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus Albacares*) in the eastern Pacific Ocean. Fisheries Research 213, 121–131. https://doi.org/10.1016/j.fishres.2019.01.013

6. Tables

Table 1: Code, description, class, and type of candidate explanatory variables. The time period series indicates which of the time series (long, short or both) for which the variable was considered. Catchability covariates were fixed at a specified value when predicting, while density covariates varied over space and time.

Code	Description	Class	Time series	Type
yr_qtr	Year-Quarter interaction	Factor (levels: 2010-1,…,2023-4)	Both	-
country	Fleet country	Factor (levels: France, Spain)	Both	Catchability
capacity	Vessel capacity in cubic meters	Numeric	Long	Catchability
numbat	Vessel identifier code	Factor (levels meaningless)	Short	Catchability
follow	Followed a FAD with echosounder capacity?	Factor (levels: No, Yes_No-Echo, Yes_Echo)	Short	Catchability
n_buoys20	Number of buoys within 20 nm	Numeric	Short	Catchability
n_buoys250	Number of owned buoys within 250 km	Numeric	Short	Catchability
t_sunrise	Set time since sunrise	Numeric	Short	Catchability
mld	Mixing layer depth	Numeric	Short	Catchability
mld_clim	Mixing layer depth climatology (1993-2023)	Numeric	Long	Density
mld_anom	Mixing layer depth anomaly	Numeric	Long	Catchability
sst	Potential sea surface temperature	Numeric	Short	Density
sst_clim	Potential sea surface temperature climatology (1993-2023)	Numeric	Long	Density
sst_anom	Potential sea surface temperature anomaly	Numeric	Long	Density
nppv	Total net primary production 0-100m depth	Numeric	Short	Density
nppv_clim	Total net primary production 0-100m depth, climatology (1993-2023)	Numeric	Long	Density
nppv_anom	Total net primary production 0-100m depth, anomaly	Numeric	Long	Density
iso20	Depth of 20C isotherm	Numeric	Short	Density
iso20_clim	Depth of 20C isotherm, climatology (1990-2023)	Numeric	Long	Density
iso20_anom	Depth of 20C isotherm, anomaly	Numeric	Long	Density

Table 2: AIC values for short time series (2010-2023) models assuming different probability distributions for the response variable.

Family	AIC
Lognormal	206441.6
Gamma	207312.7
Generalized gamma	203761.9
Tweedie	212077.8

Table 3: Summary table of the final model (generalized gamma) for the short time series (2010-2023). Coefficients associated with the year-quarter interaction are not shown.

term	estimate	std.error	p_value
countrySpain	0.14	0.04	<0.01
followYes_No-echo	-0.13	0.04	<0.01
followYes_Echo	0.01	0.01	0.28
n_buoys20	-0.02	0.00	<0.01
n_buoys250	0.02	0.01	<0.01
t_sunrise	-0.17	0.00	<0.01
sst	-0.14	0.01	<0.01
mld	-0.05	0.01	<0.01
iso20	-0.03	0.01	<0.01

Table 4: Standardized CPUE index for the short time series (2010-2023) and 95% confidence interval obtained from the model using the quarter-specific spatial term configuration.

year	quarter	est	lwr	upr	cv
2010	1	4.760	4.309	5.259	0.051
2010	2	2.262	2.001	2.558	0.063
2010	3	2.592	2.347	2.862	0.051
2010	4	2.066	1.879	2.272	0.048
2011	1	6.340	5.743	7.000	0.051
2011	2	1.666	1.488	1.865	0.058
2011	3	1.673	1.487	1.883	0.060
2011	4	1.403	1.281	1.536	0.046
2012	1	2.409	2.152	2.696	0.058
2012	2	1.897	1.682	2.139	0.061
2012	3	1.873	1.676	2.093	0.057
2012	4	1.942	1.757	2.148	0.051
2013	1	4.481	4.093	4.907	0.046
2013	2	2.639	2.362	2.948	0.057
2013	3	3.713	3.334	4.134	0.055
2013	4	3.346	3.073	3.644	0.043
2014	1	3.430	3.083	3.814	0.054
2014	2	1.965	1.752	2.203	0.059
2014	3	2.660	2.381	2.970	0.056
2014	4	1.817	1.642	2.009	0.051
2015	1	4.223	3.798	4.695	0.054
2015	2	1.773	1.571	2.001	0.062
2015	3	2.154	1.900	2.442	0.064
2015	4	1.654	1.508	1.814	0.047
2016	1	3.134	2.858	3.436	0.047
2016	2	1.859	1.686	2.050	0.050
2016	3	2.312	2.105	2.539	0.048
2016	4	2.341	2.165	2.531	0.040
2017	1	3.040	2.783	3.321	0.045
2017	2	1.602	1.440	1.782	0.054
2017	3	1.930	1.732	2.152	0.055
2017	4	2.544	2.302	2.812	0.051
2018	1	3.231	3.005	3.474	0.037
2018	2	1.952	1.786	2.133	0.045
2018	3	1.710	1.566	1.867	0.045
2018	4	1.884	1.744	2.035	0.039
2019	1	2.420	2.259	2.593	0.035
2019	2	1.819	1.637	2.021	0.054
2019	3	2.040	1.842	2.259	0.052

2019	4	1.281	1.177	1.394	0.043
2020	1	1.124	1.030	1.226	0.044
2020	2	1.151	1.038	1.276	0.053
2020	3	2.215	2.020	2.429	0.047
2020	4	1.547	1.420	1.684	0.043
2021	1	5.027	4.507	5.606	0.056
2021	2	2.970	2.601	3.392	0.068
2021	3	2.291	1.991	2.635	0.072
2021	4	1.989	1.773	2.231	0.059
2022	1	3.262	2.957	3.597	0.050
2022	2	3.189	2.839	3.582	0.059
2022	3	3.822	3.372	4.333	0.064
2022	4	4.516	4.033	5.058	0.058
2023	1	3.329	3.042	3.642	0.046
2023	2	2.318	2.049	2.622	0.063
2023	3	1.301	1.151	1.471	0.063
2023	4	1.587	1.465	1.719	0.041

Table 5: AIC values for long time series (1991-2023) models assuming different probability distributions for the response variable.

Family	AIC
Lognormal	450054.8
Gamma	461206.7
Generalized gamma	444553.8

Table 6: Summary table of the final model (generalized gamma) for the long time series (1991-2023). Coefficients associated with the year-quarter interaction are not shown. Note that coefficients have units of *tonnes/UNIT*, where *UNIT* is the unit of the corresponding predictor variable.

term	estimate	std.error	p_value
countrySpain	0.15371	0.02830	0.00000
capacity	0.00017	0.00002	0.00000
sst_clim	-0.11790	0.00822	0.00000
sst_anom	-0.12828	0.00989	0.00000
mld_clim	-0.00950	0.00083	0.00000
mld_anom	-0.00195	0.00078	0.01236

Table 7: Annual standardized CPUE index for the long time series (1991-2023) and 95% confidence interval obtained from the model using the quarter-specific spatial term configuration. The index is in units of tonnes/set.

year	est	lwr	upr	cv
1991	4.82	4.18	5.58	0.07
1992	3.50	3.11	3.97	0.06
1993	2.83	2.52	3.20	0.06
1994	3.53	3.18	3.95	0.06
1995	5.33	4.79	5.95	0.06
1996	4.79	4.36	5.27	0.05
1997	6.83	6.22	7.53	0.05
1998	4.25	3.81	4.77	0.06
1999	5.47	5.05	5.94	0.04
2000	4.29	3.96	4.67	0.04
2001	4.38	3.99	4.82	0.05
2002	4.88	4.52	5.28	0.04
2003	3.09	2.84	3.38	0.04
2004	3.85	3.55	4.20	0.04
2005	2.54	2.36	2.75	0.04
2006	2.60	2.44	2.78	0.03
2007	2.61	2.46	2.79	0.03
2008	3.47	3.20	3.78	0.04
2009	3.76	3.51	4.06	0.04
2010	2.75	2.57	2.97	0.04
2011	2.64	2.46	2.85	0.04
2012	1.92	1.79	2.07	0.04
2013	3.43	3.22	3.66	0.03
2014	2.35	2.17	2.56	0.04
2015	2.22	2.06	2.41	0.04
2016	2.22	2.09	2.39	0.03
2017	2.09	1.94	2.26	0.04
2018	2.07	1.96	2.20	0.03
2019	1.86	1.74	2.00	0.04
2020	1.41	1.31	1.54	0.04
2021	2.87	2.57	3.21	0.06
2022	3.42	3.12	3.76	0.05
2023	2.01	1.87	2.18	0.04

7. Figures

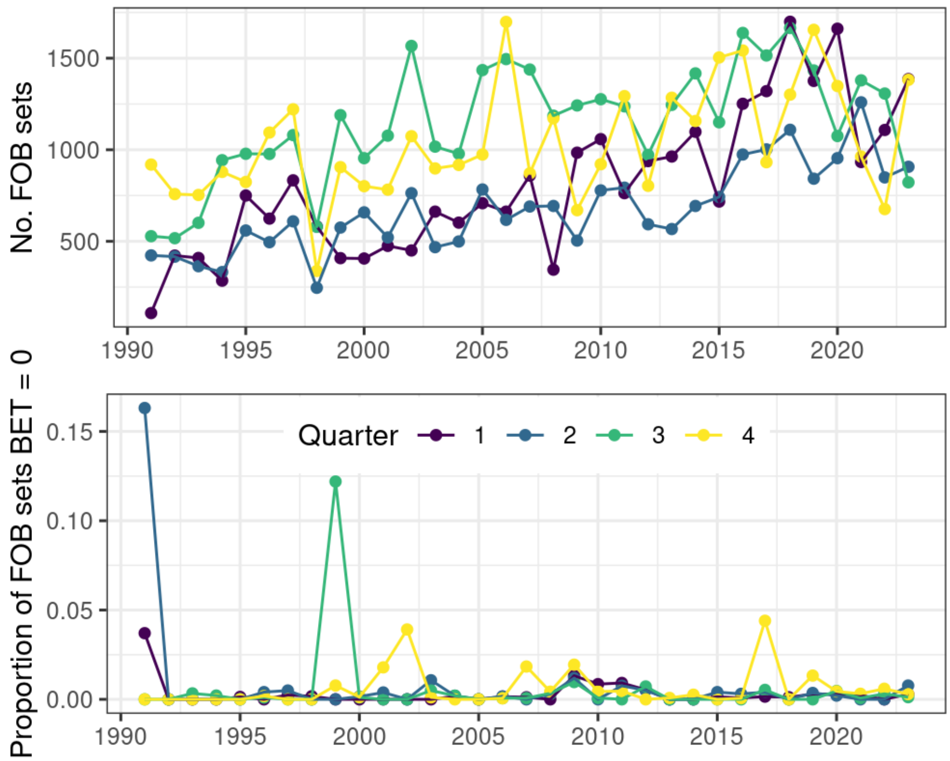


Figure 1: Number of non-null FOB sets in the data (top) and proportion of sets with BET catch equal to zero (bottom) by year and quarter for the long series (1991-2023).

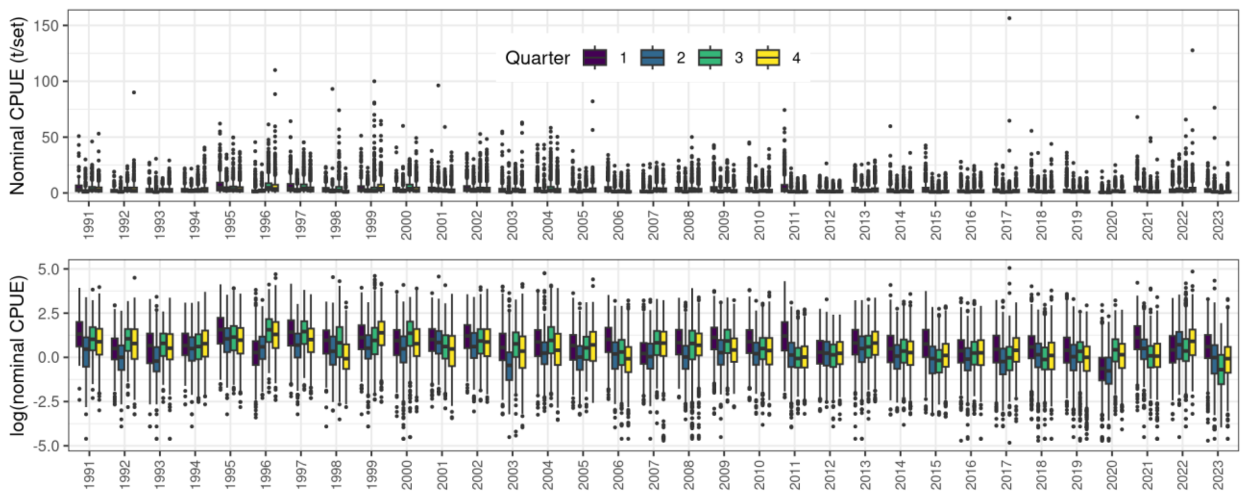


Figure 2: Distribution of BET catch per set, original (top) and log-transformed (bottom), by year and quarter for the long time series (1991-2023).

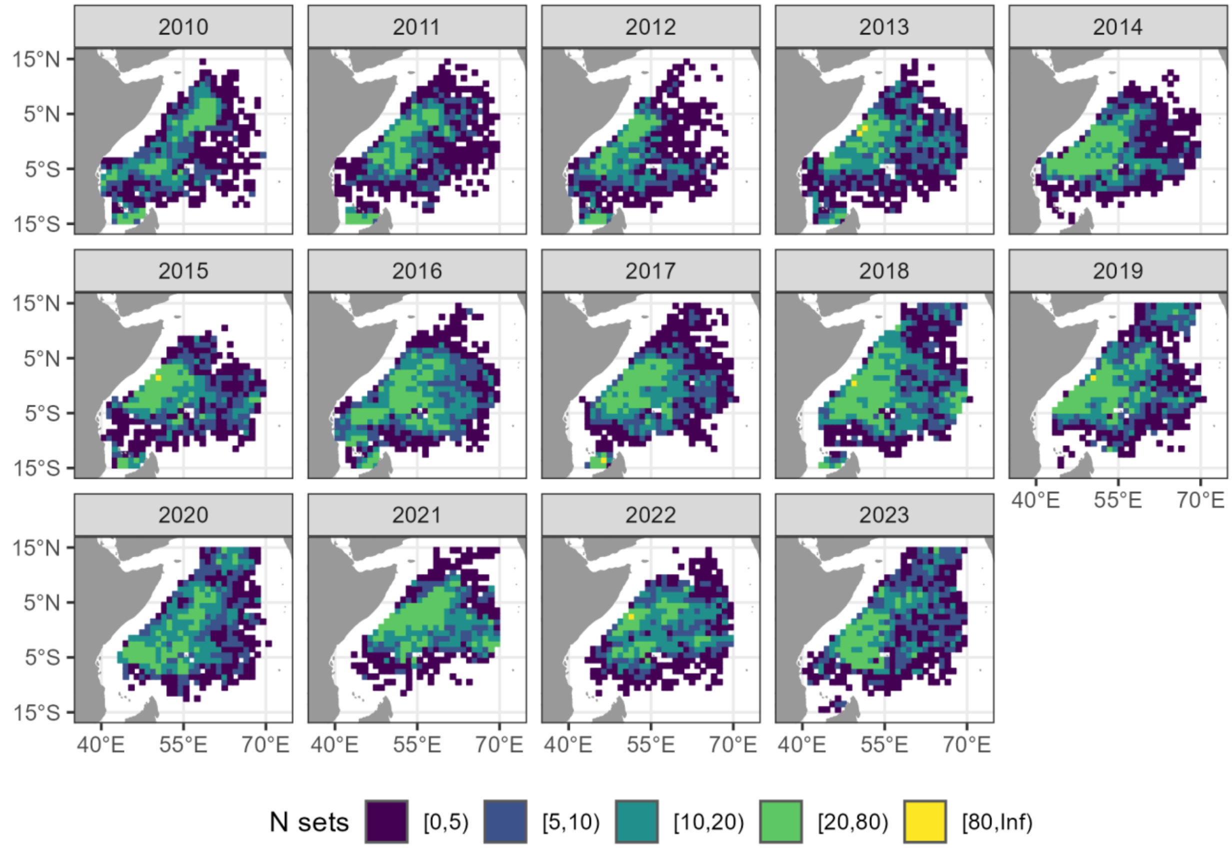


Figure 3: Number of FOB fishing sets (effort) per 1° × 1° grid cell and year for the short time series (2010-2023).

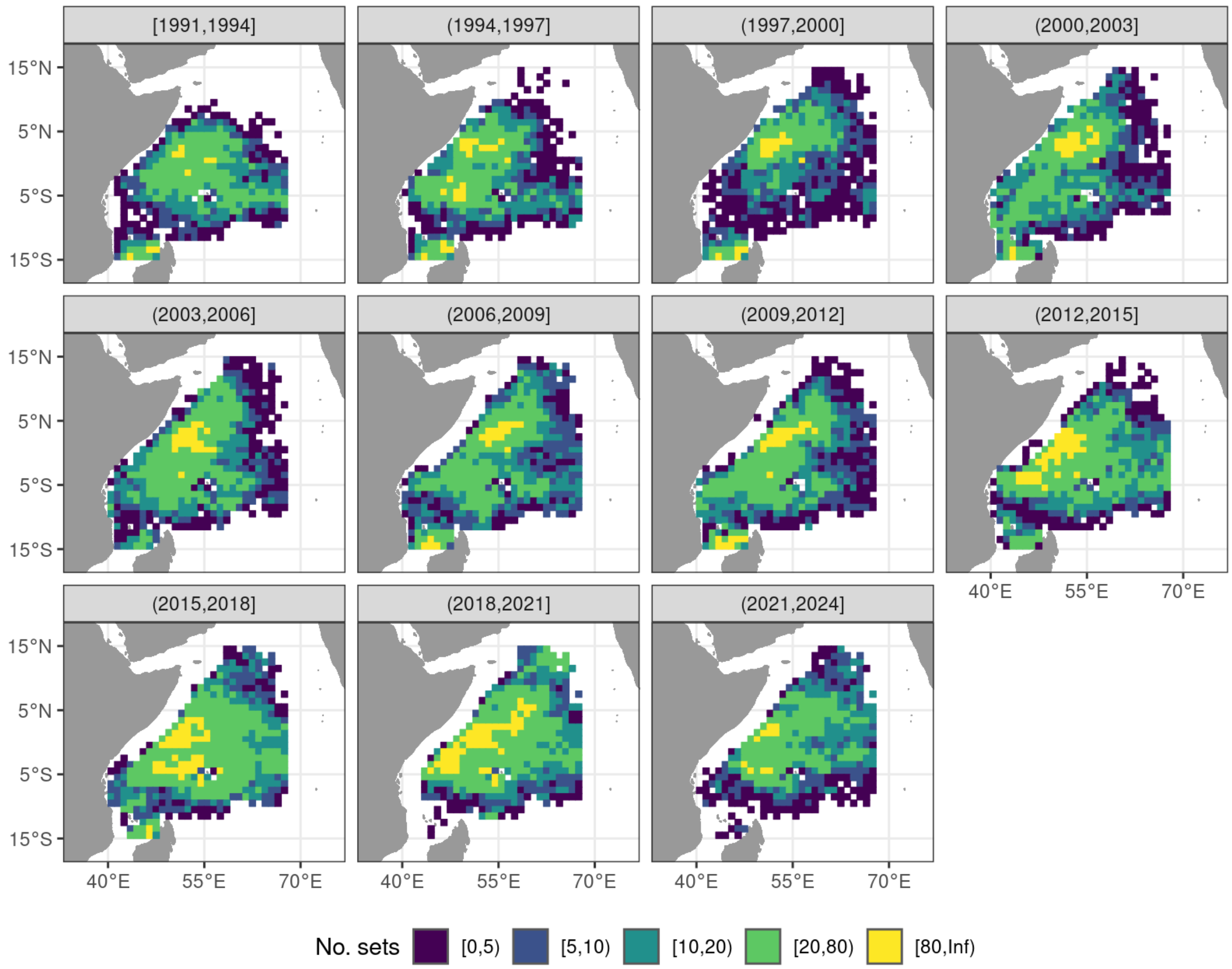


Figure 4: Number of FOB fishing sets (effort) per 1° × 1° grid cell and 3-year time period for the long time series (1991-2023).

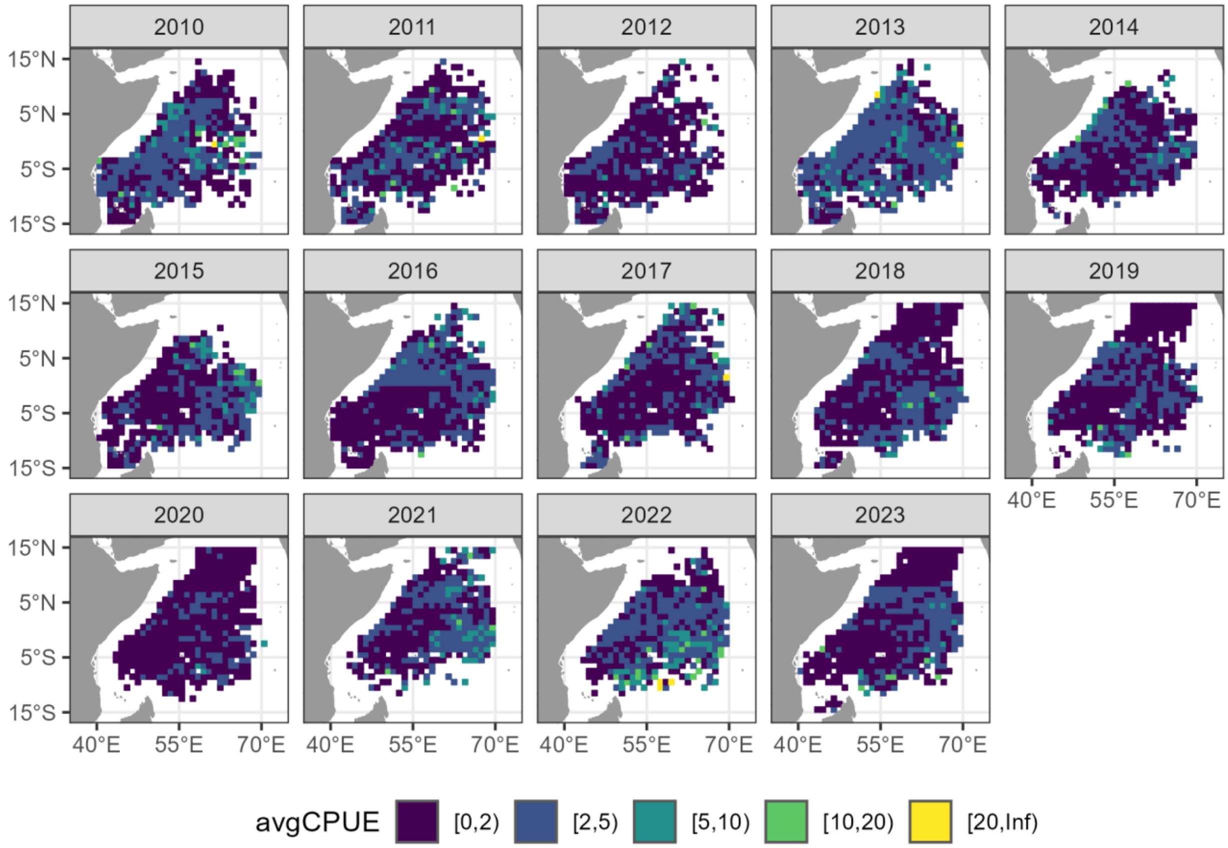


Figure 5: Observed average catch per 1° × 1° grid cell and year for the short time series (2010-2023).

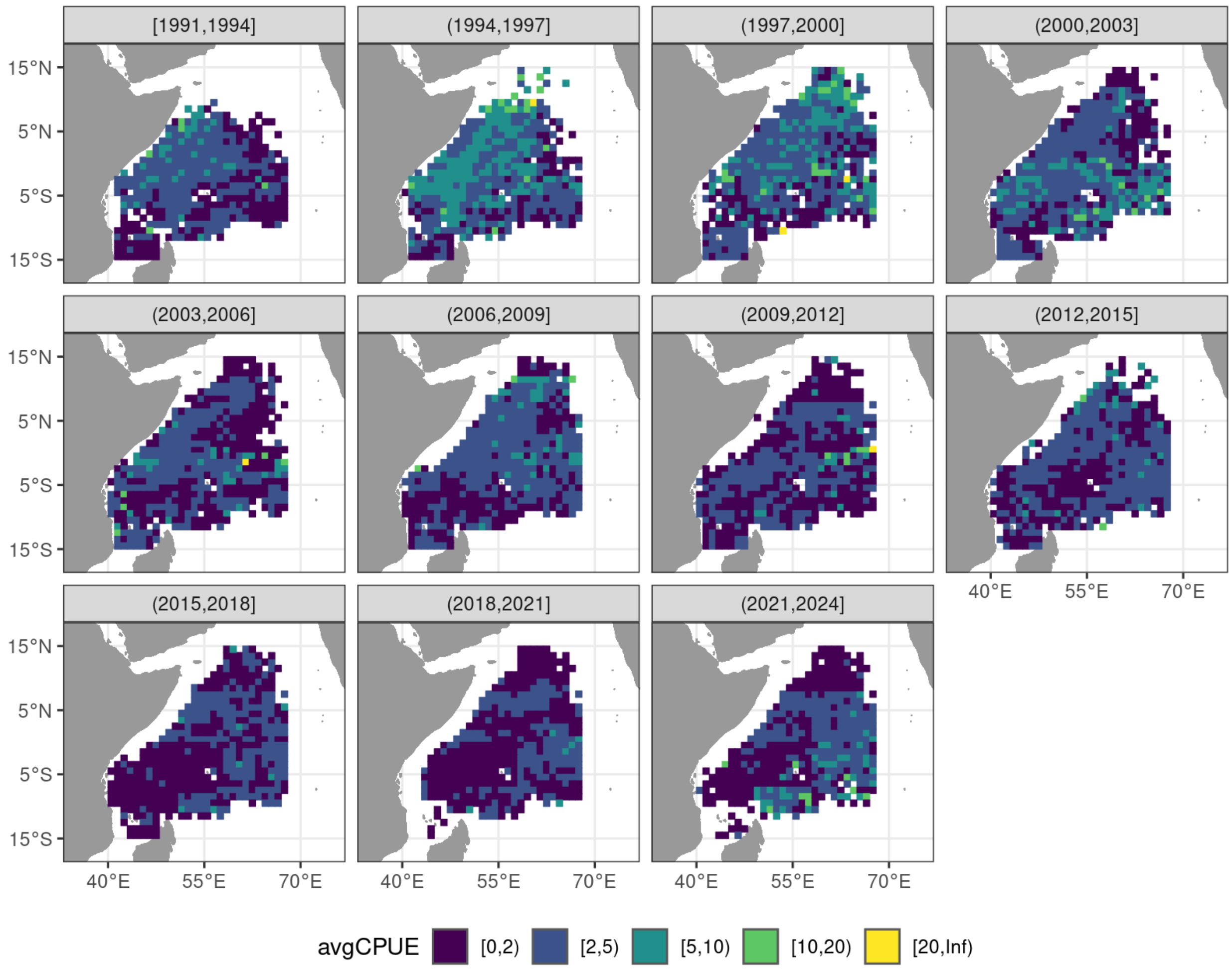


Figure 6: Average catch per set (CPUE) per 1° × 1° grid cell and 3-year time period for the long time series (1991-2023).

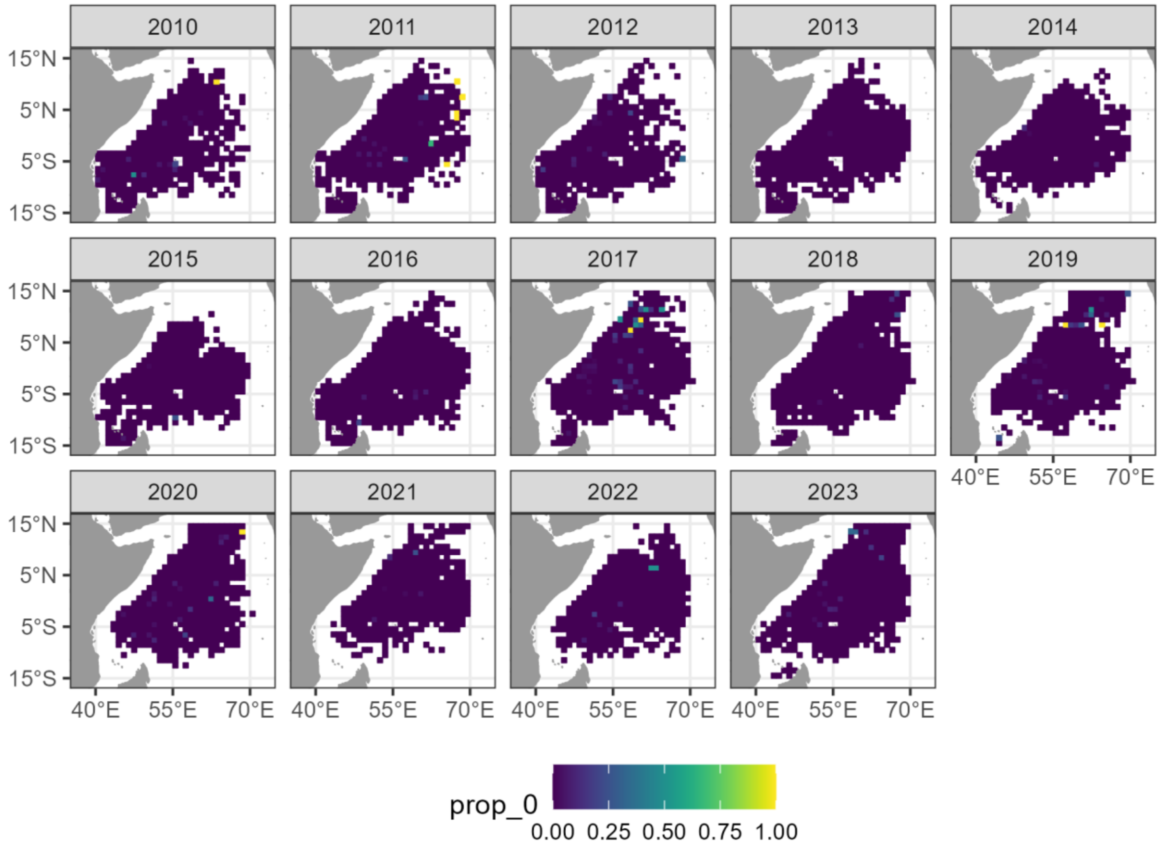


Figure 7: Observed proportion of sets with zero BET catch per 1° × 1° grid cell and year in the short time series (2010-2023).

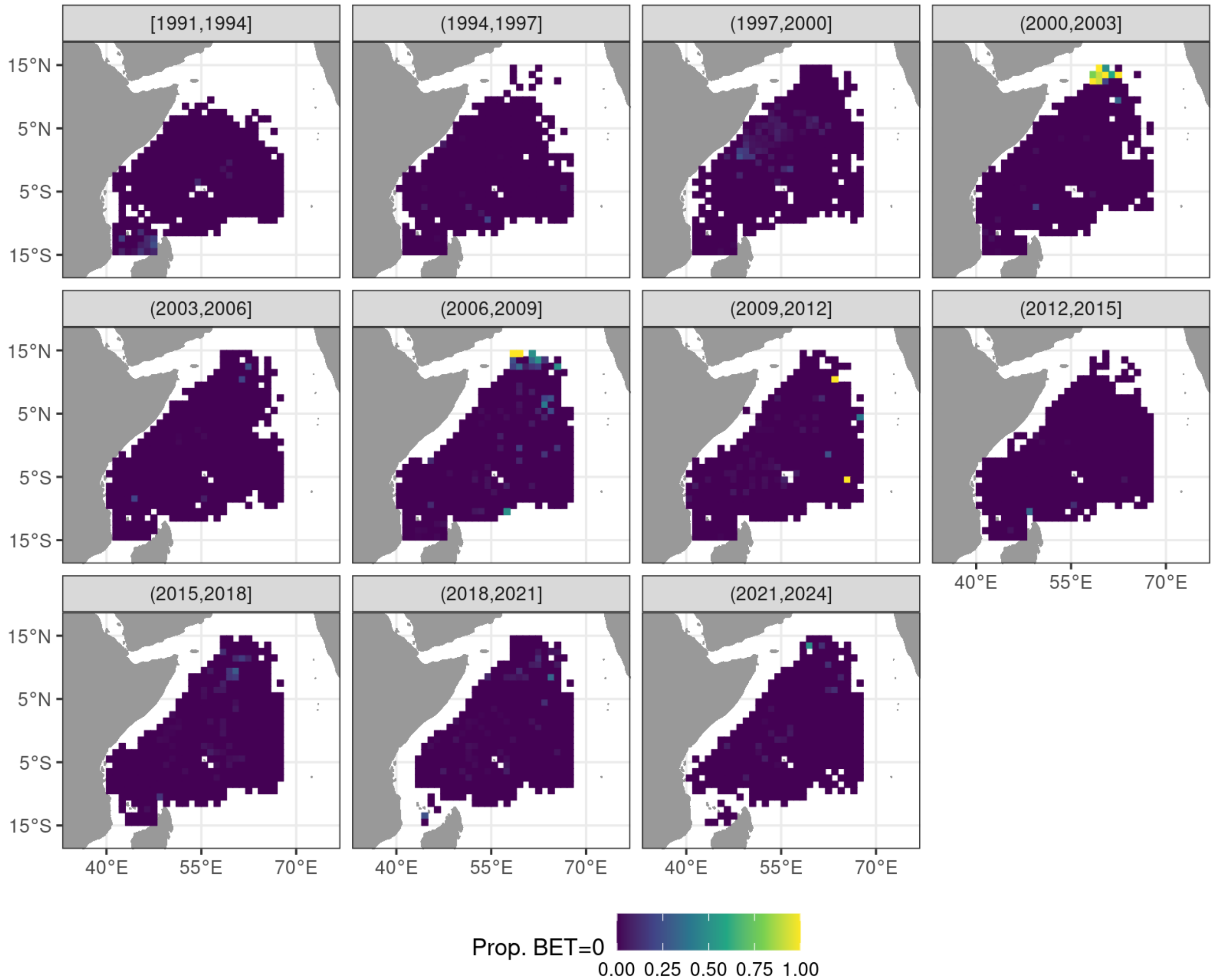


Figure 8: Observed proportion of sets with zero BET catch per 1° × 1° grid cell and 3-year time period for the long time series (1991-2023).

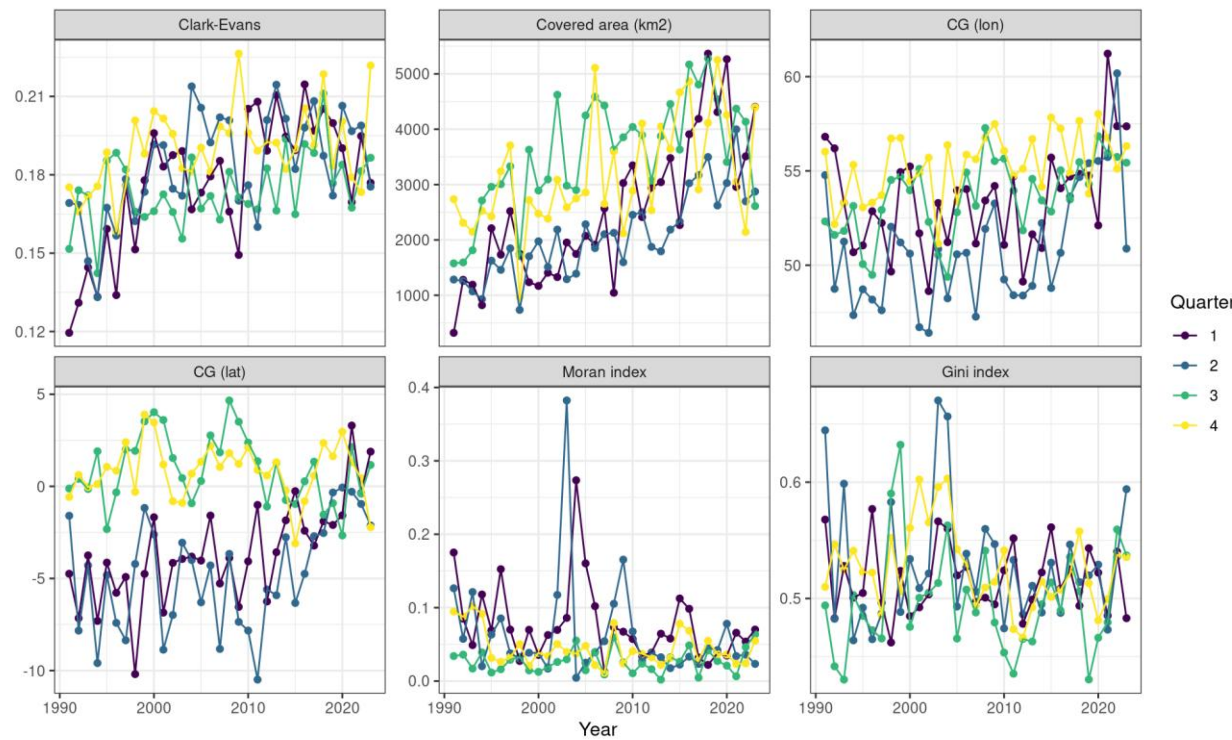


Figure 9: Spatial indicators calculated by quarter for the long time series (1991-2023).

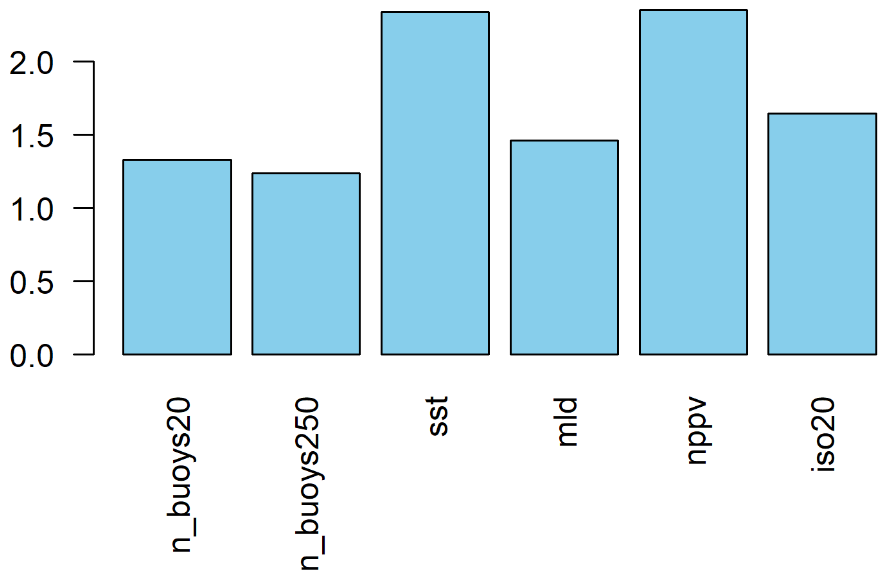


Figure 10: Variance inflation factors (VIF) by covariate obtained when fitting a simple linear model with short time series data (2010-202#).

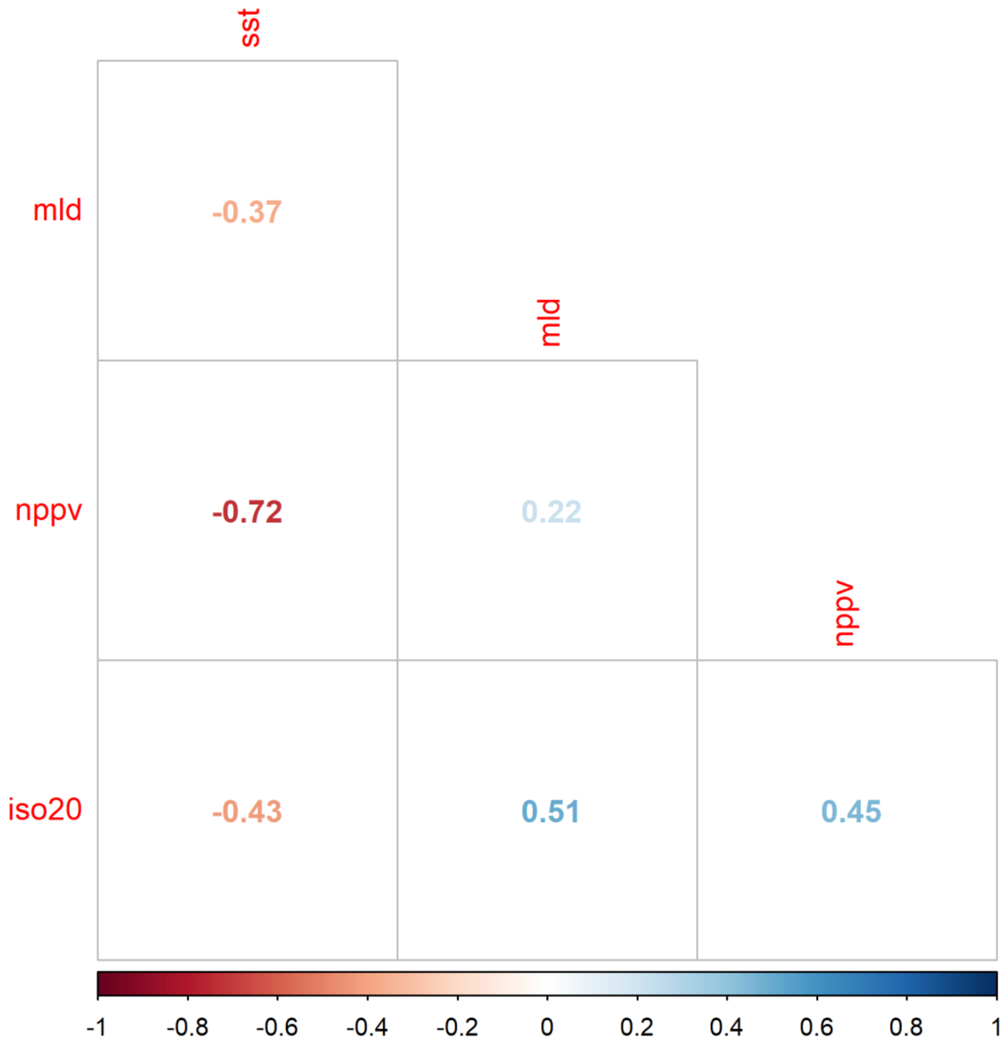


Figure 11: Correlation matrix among environmental covariates for the short time series (2010-2023).

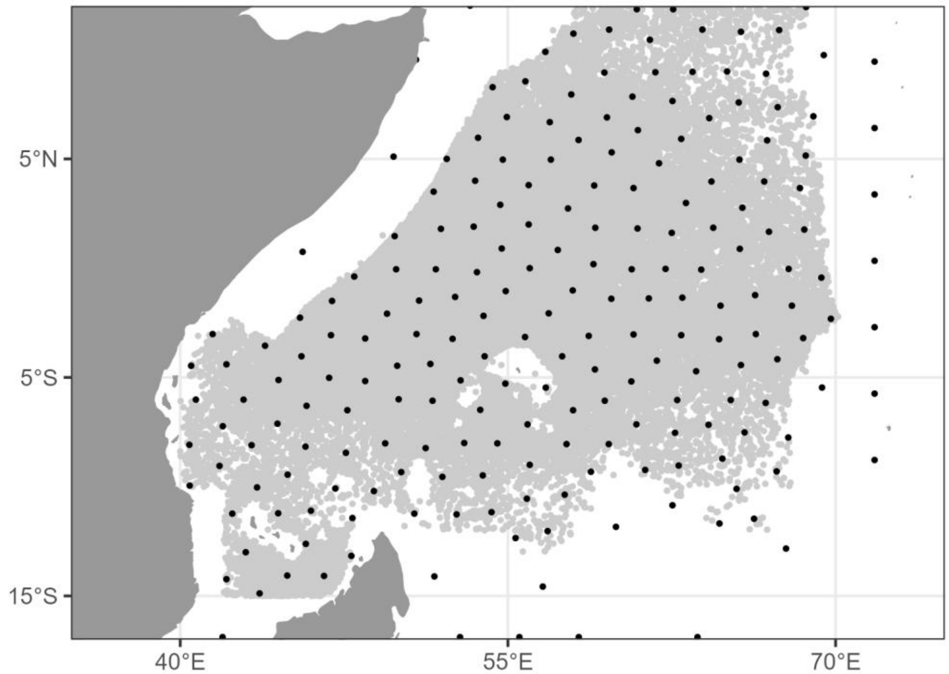


Figure 12: Mesh nodes (black dots) used in the spatiotemporal model and observations (gray dots) for the short time series (2010-2023).

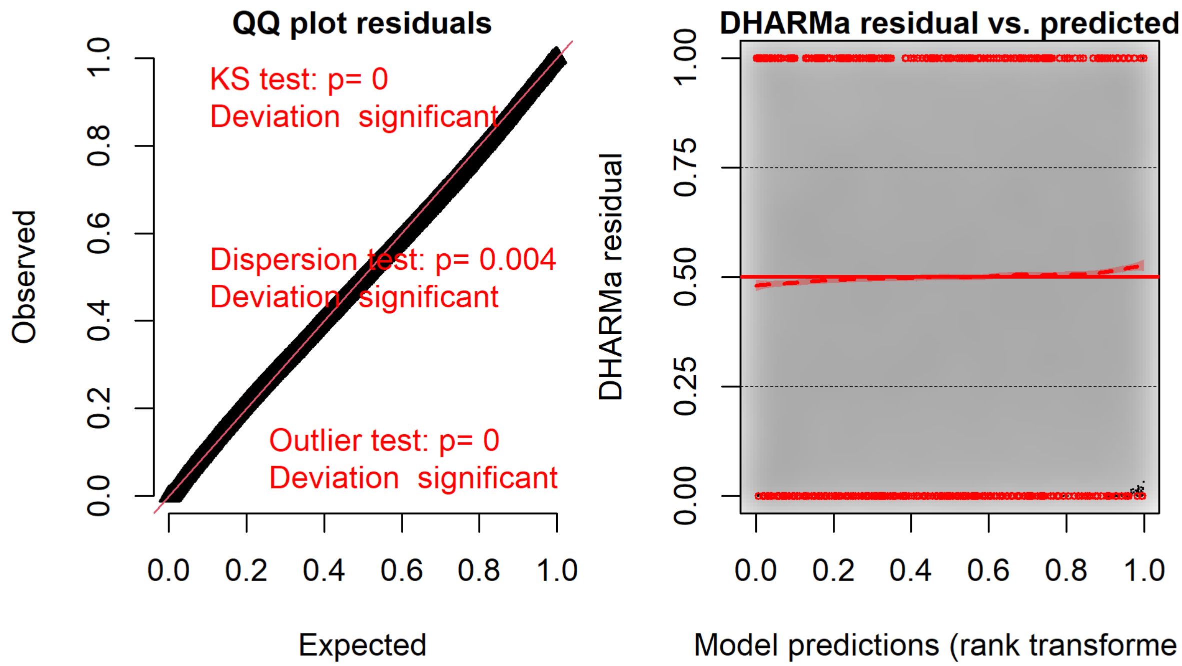


Figure 13: Simulation-based randomized-quantile residuals for the short time series model (2010-2023). QQ-plot (left) detects overall deviations from the expected distribution, by default with added tests for correct distribution (KS test), dispersion and outliers. Residual plot (right) shows the residuals against the predicted value.

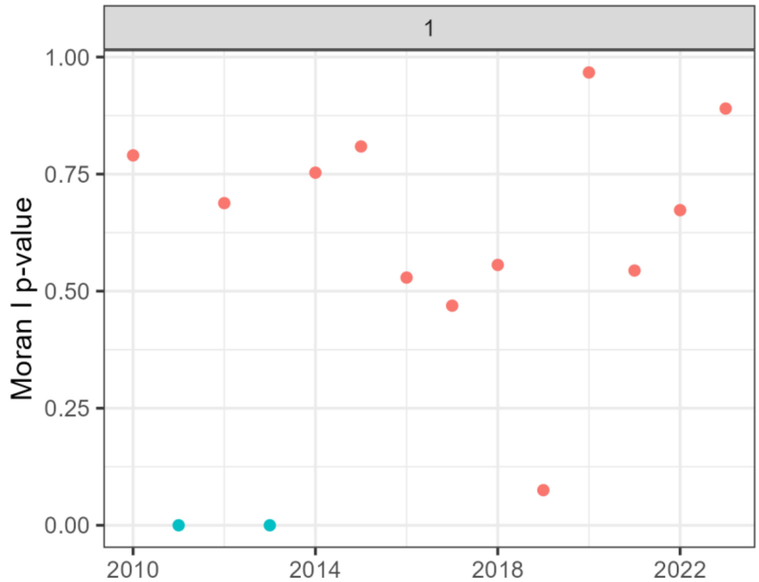


Figure 14: Moran I’s p-value of the randomized quantile residuals for standardization model of short time series (2010-2023). Blue points represent years with significant spatial autocorrelation in residuals.

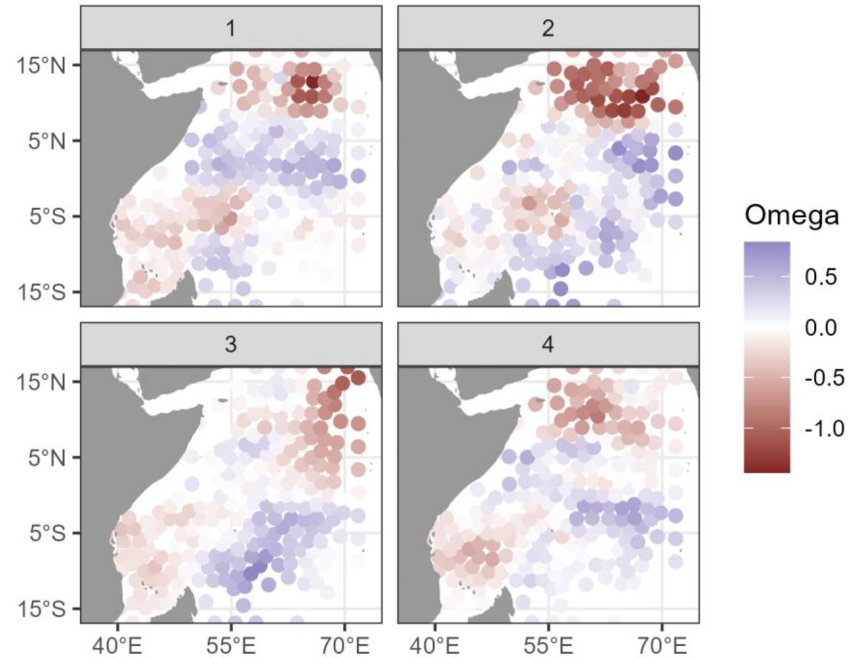


Figure 15: Quarter-specific spatial term for short time series model (2010-2023).

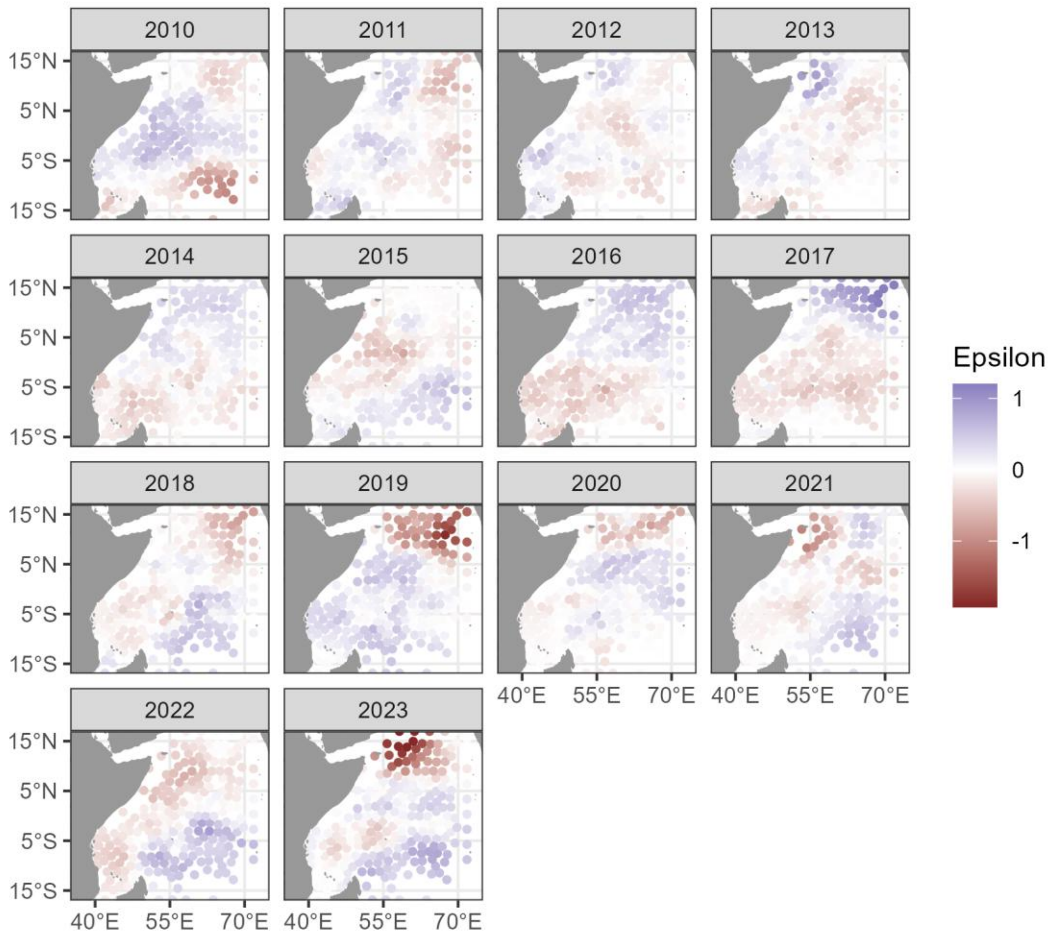


Figure 16: Spatiotemporal term for short time series model (2010-2023). To reduce the number of panels, the spatiotemporal term has been averaged over periods of three years.

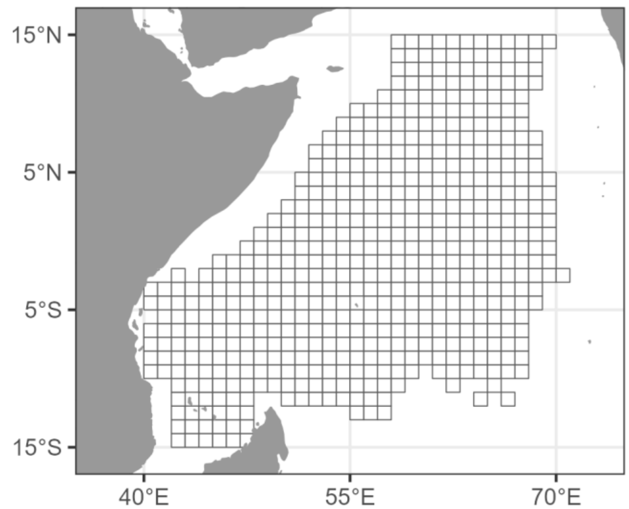


Figure 17: Extrapolation area composed by 1° × 1° grids. This area was used when predicting CPUE values over space and time for the short times series (2010-2023).

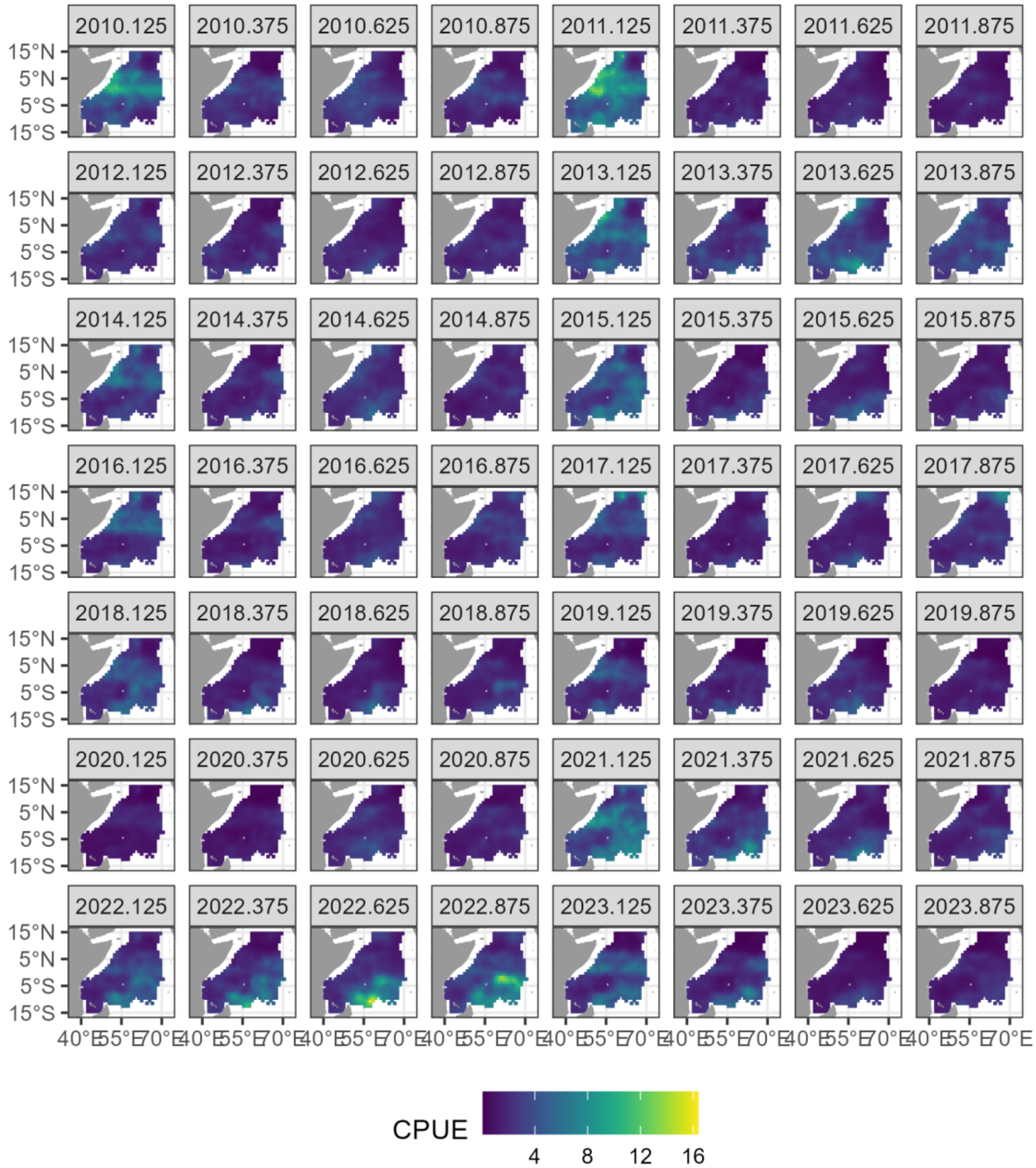


Figure 18: CPUE predictions per quarter over the extrapolation grid for the short time series (2010-2023).

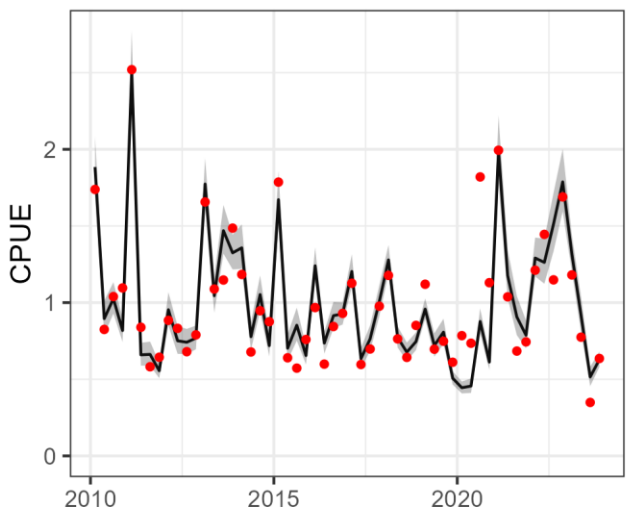


Figure 19: Standardized CPUE index for the short time series (2010-2023). Gray area represents the 95% confidence interval. Red dots represent the nominal CPUE. The index was rescaled to a mean of 1.

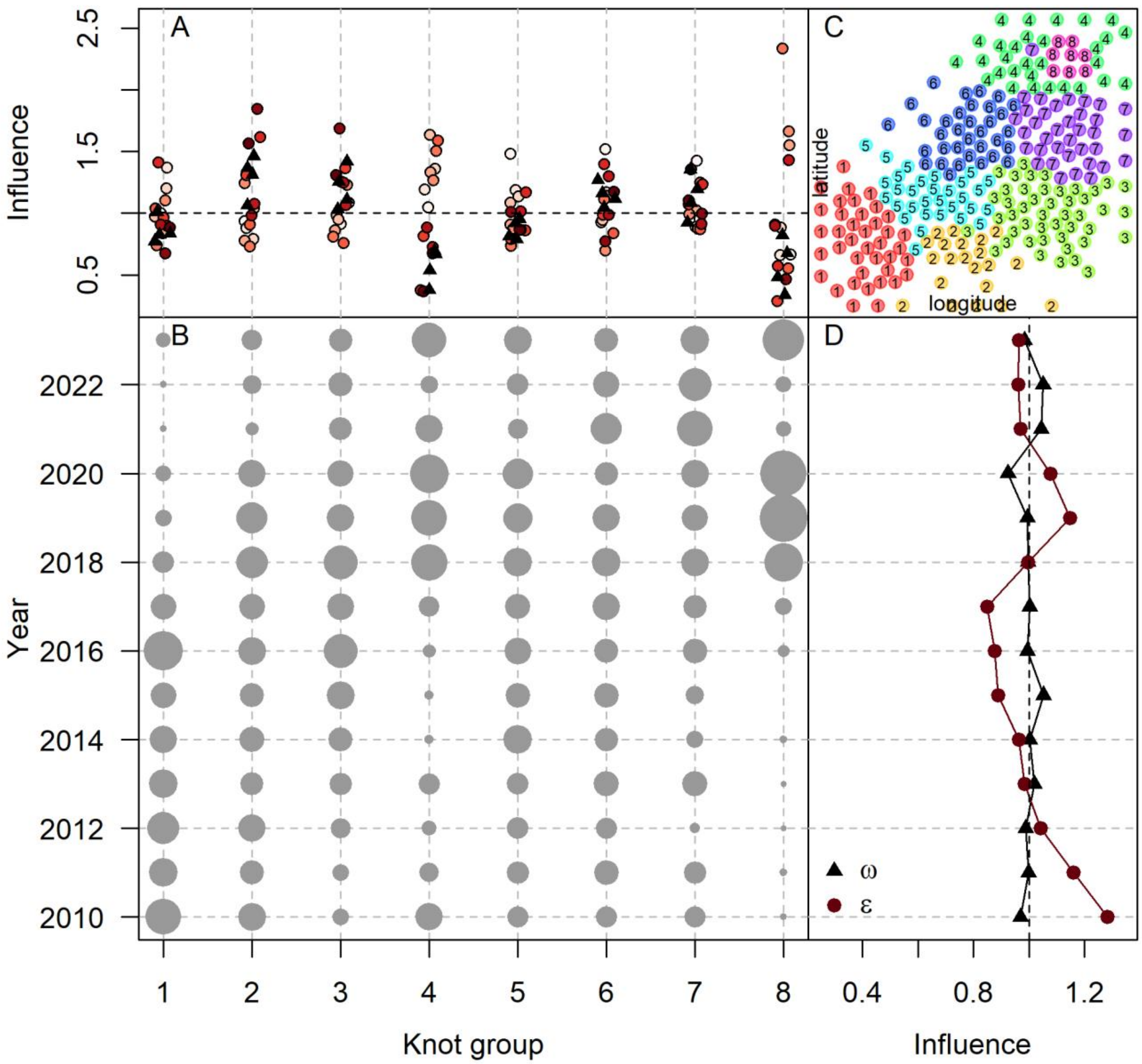


Figure 20: Coefficient-distribution-influence plot for the short times series (2010-2023). Figure A shows the coefficient by knot group for the spatial (black triangles, each triangle represents a quarter) and spatiotemporal (red dots, each dot represents a year) term. The knot groups are shown in Figure C. Figure B shows the number of observations per year and knot group. Figure D shows the influence coefficient per year for the spatial and spatiotemporal term.

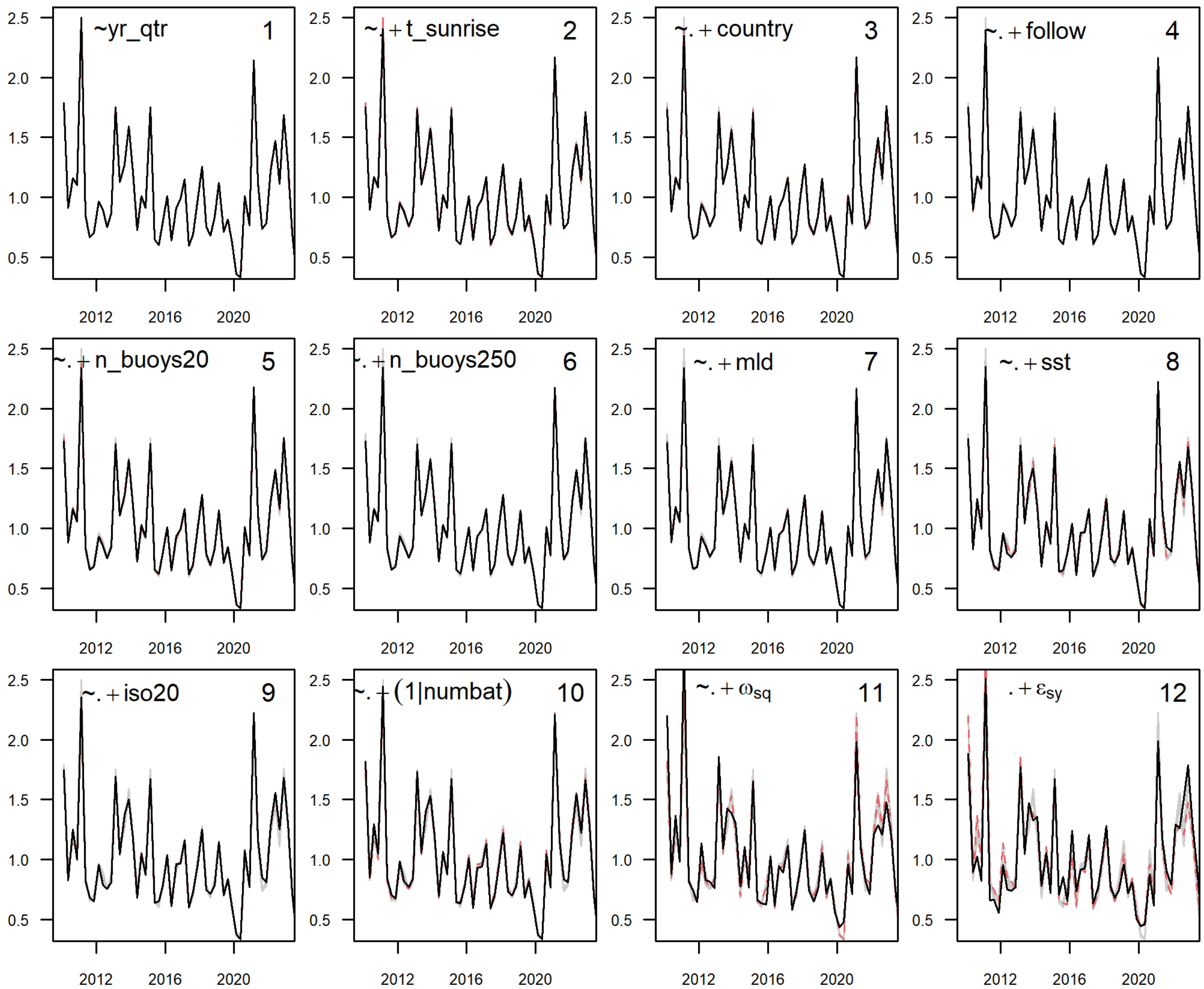


Figure 21: Step plot to evaluate the effect of adding a new covariate on the CPUE index for the short time series (2010-2023). The numbers indicate the steps, and the formula indicates the covariate that was added. The black line represents the CPUE index at that step, the red line is the CPUE index from the previous step, and the gray lines indicate all the previous CPUE indices.

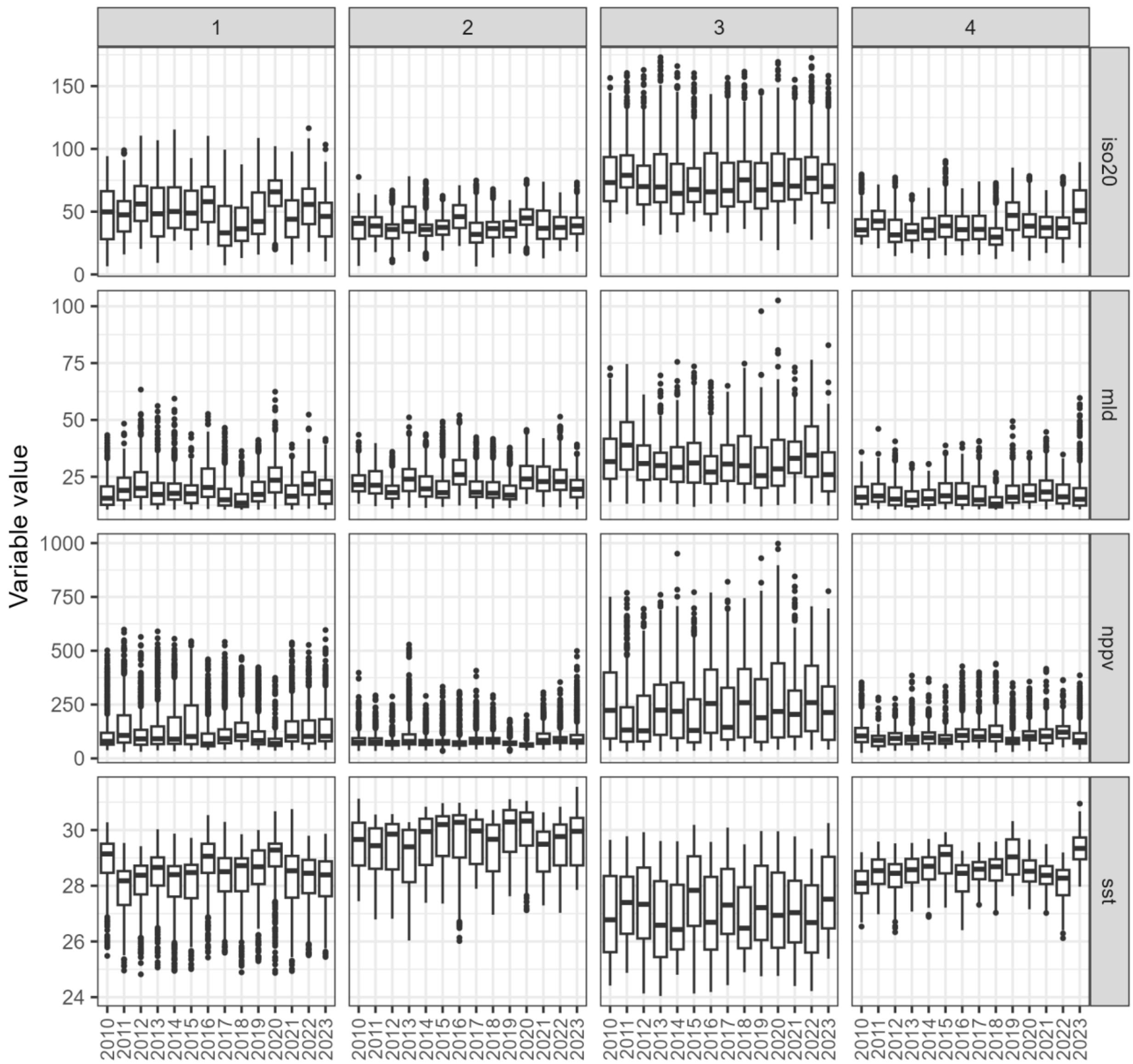


Figure 22: Temporal trends in oceanographic conditions. Boxplots are composed of environmental information per grid of the extrapolation area for the short itme series (2010-2023).

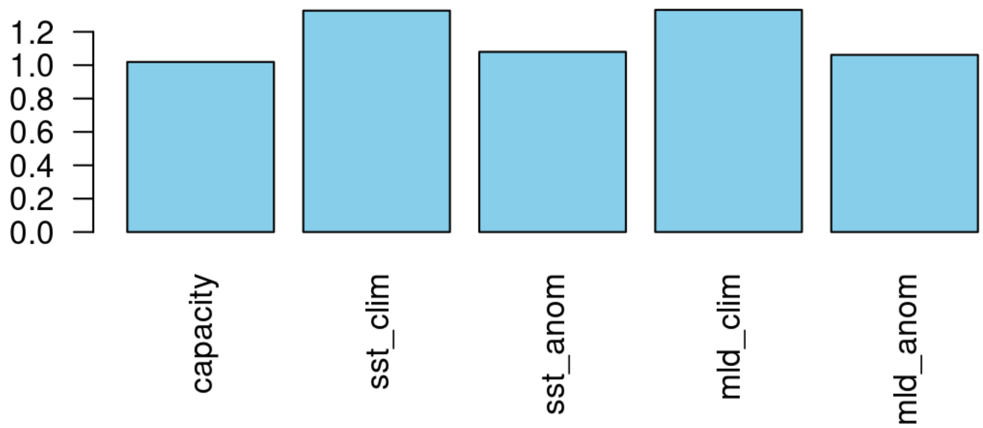
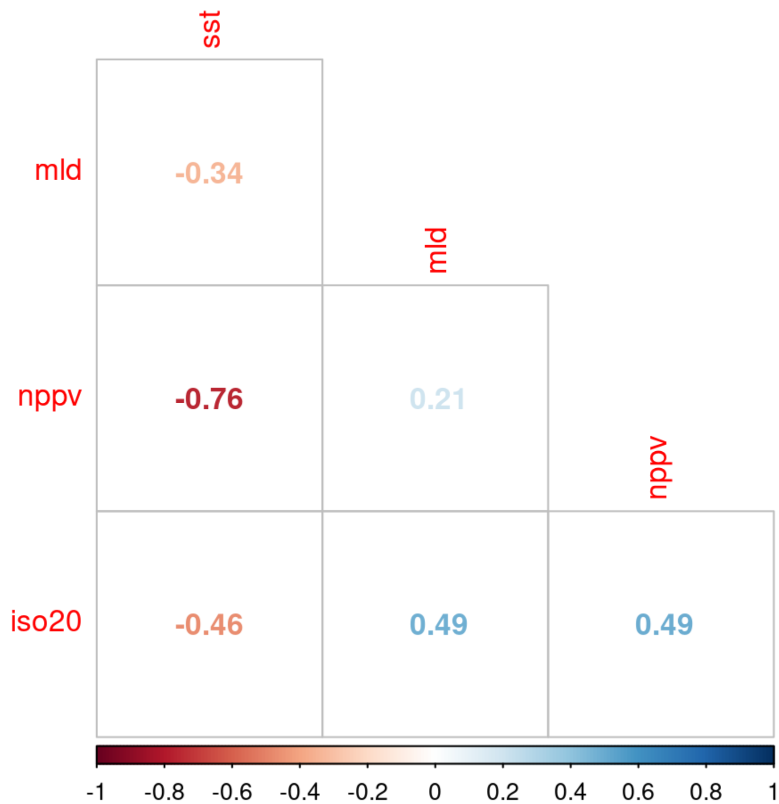


Figure 23: Variance inflation factors (VIF) by covariate obtained when fitting a simple linear model with long time series data (1991-2023).


sst
mld
nppv
mld
-0.34
nppv
-0.76
0.21
iso20
-0.46
0.49
0.49
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

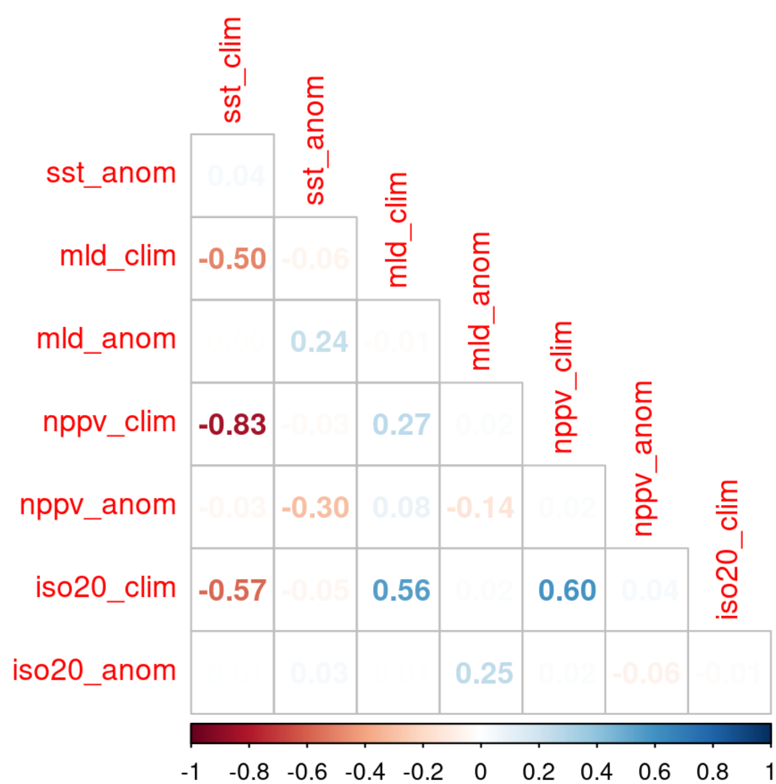

sst_clim
sst_anom
mld_clim
mld_anom
nppv_clim
nppv_anom
iso20_clim
sst_anom
mld_clim
mld_anom
nppv_clim
nppv_anom
iso20_clim
iso20_anom
-0.50
0.24
-0.83
0.27
-0.30
-0.14
-0.57
0.56
0.60
0.25
-0.06
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 24: Correlation matrix among environmental covariates for the long time series (1991-2023). The top panel shows original environmental covariates before decomposition into climatology and anomaly, whereas the bottom panel shows the climatology and anomaly variables that were used in long time series models.

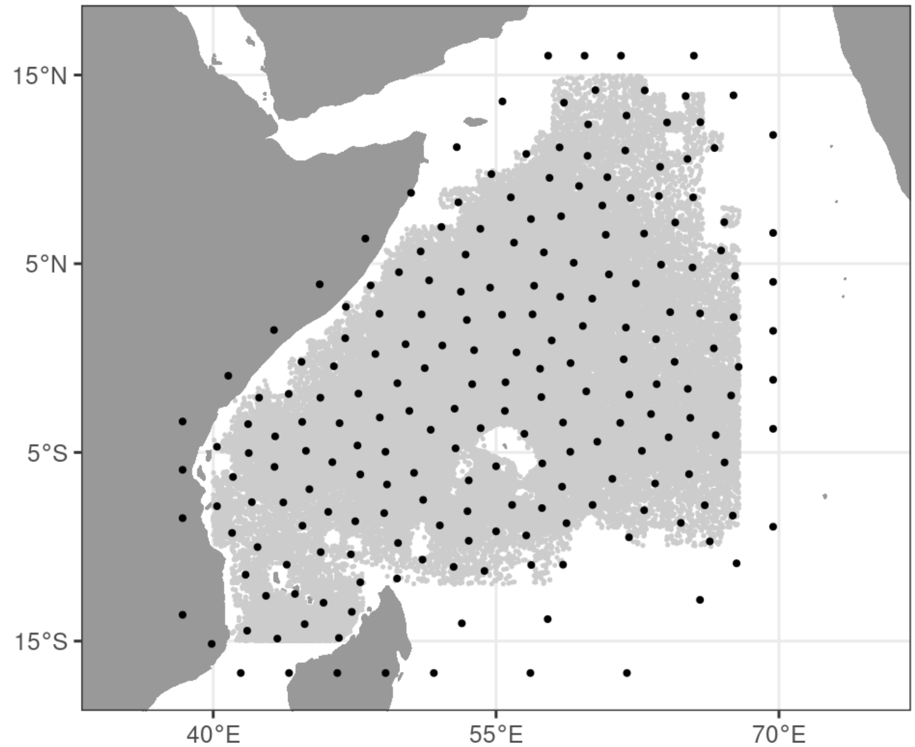


Figure 25: Mesh nodes (black dots) used in the spatiotemporal model and observations (gray dots) for the long time series (1991-2023).

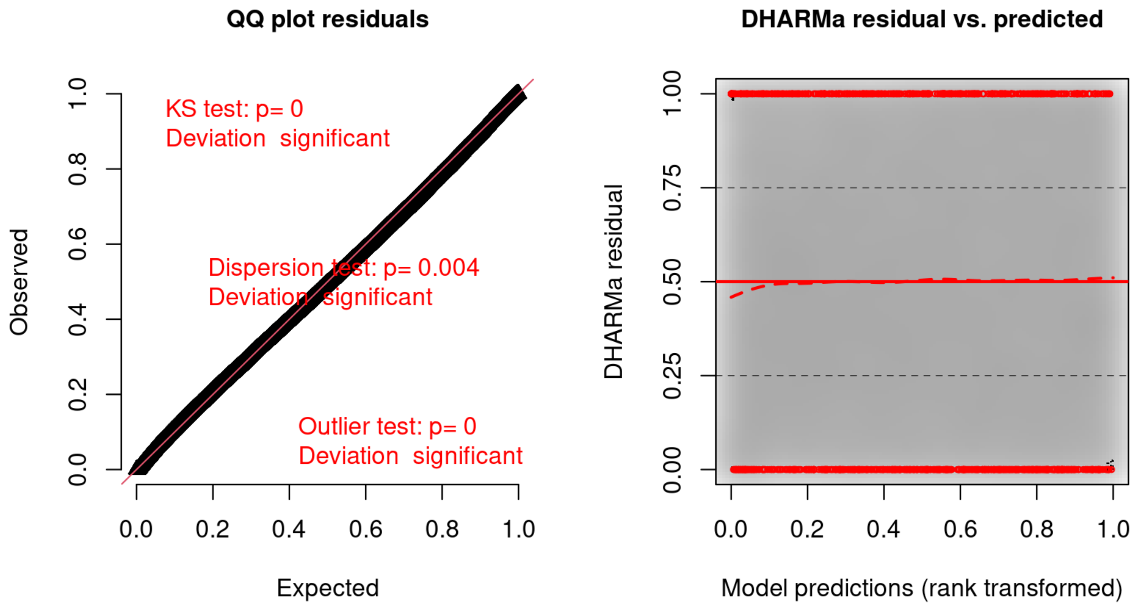


Figure 26: Simulation-based randomized-quantile residuals for the long time series model (1991-2023). QQ-plot (left) detects overall deviations from the expected distribution, by default with added tests for correct distribution (KS test), dispersion and outliers. Residual plot (right) shows the residuals against the predicted value.

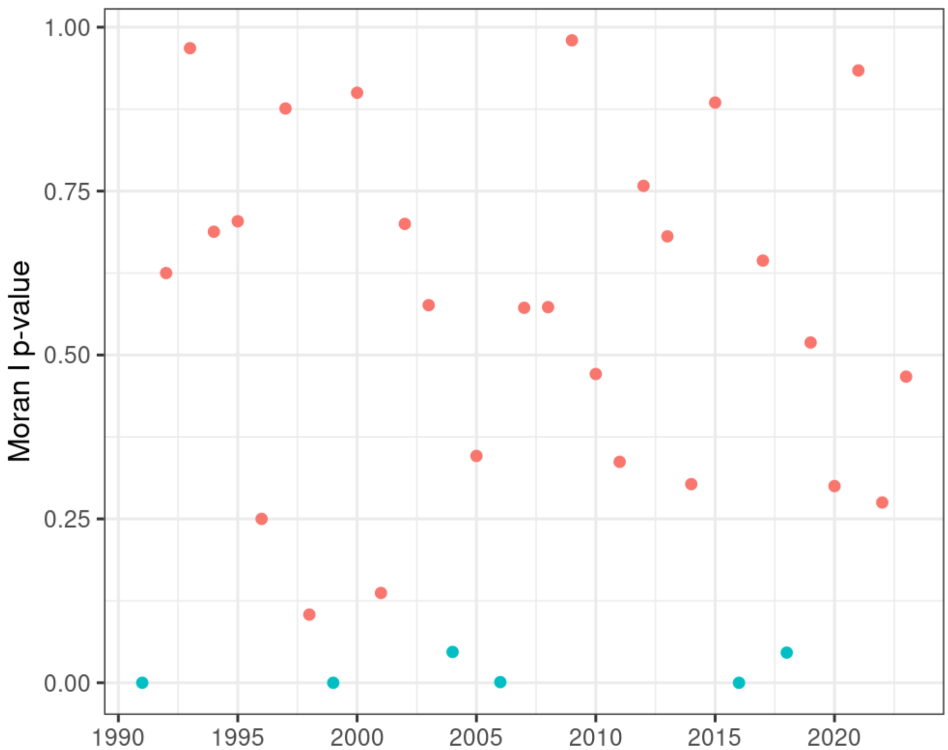


Figure 27: Moran I's p-value of the randomized quantile residuals for standardization model of long time series (1991-2023). Blue points represent years with significant spatial autocorrelation in residuals.

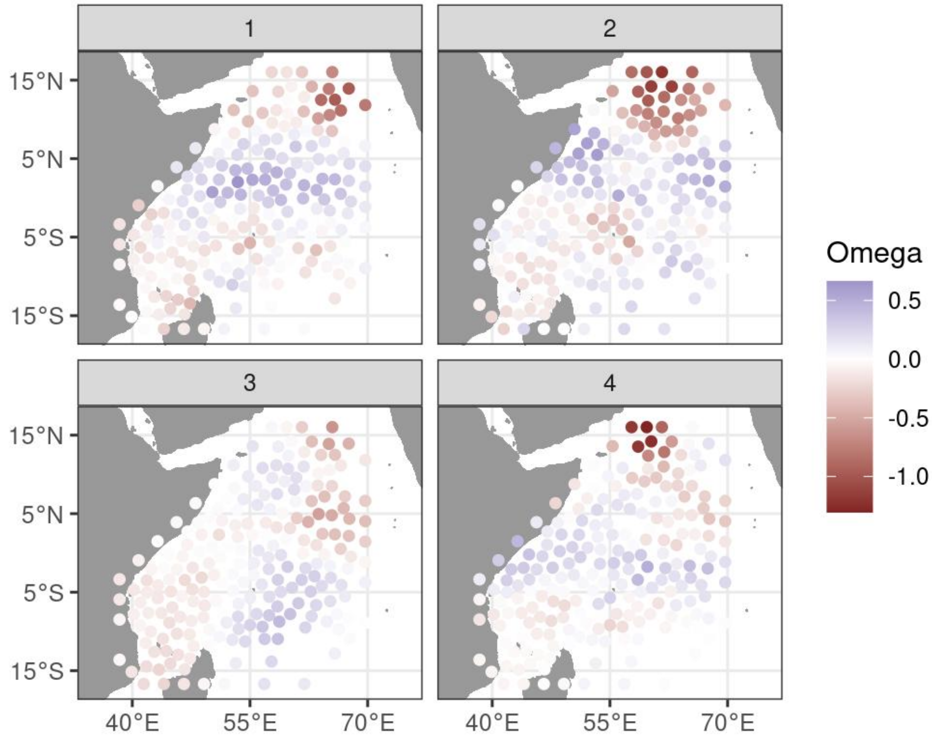


Figure 28: Quarter-specific spatial term for long time series model (1991-2023).

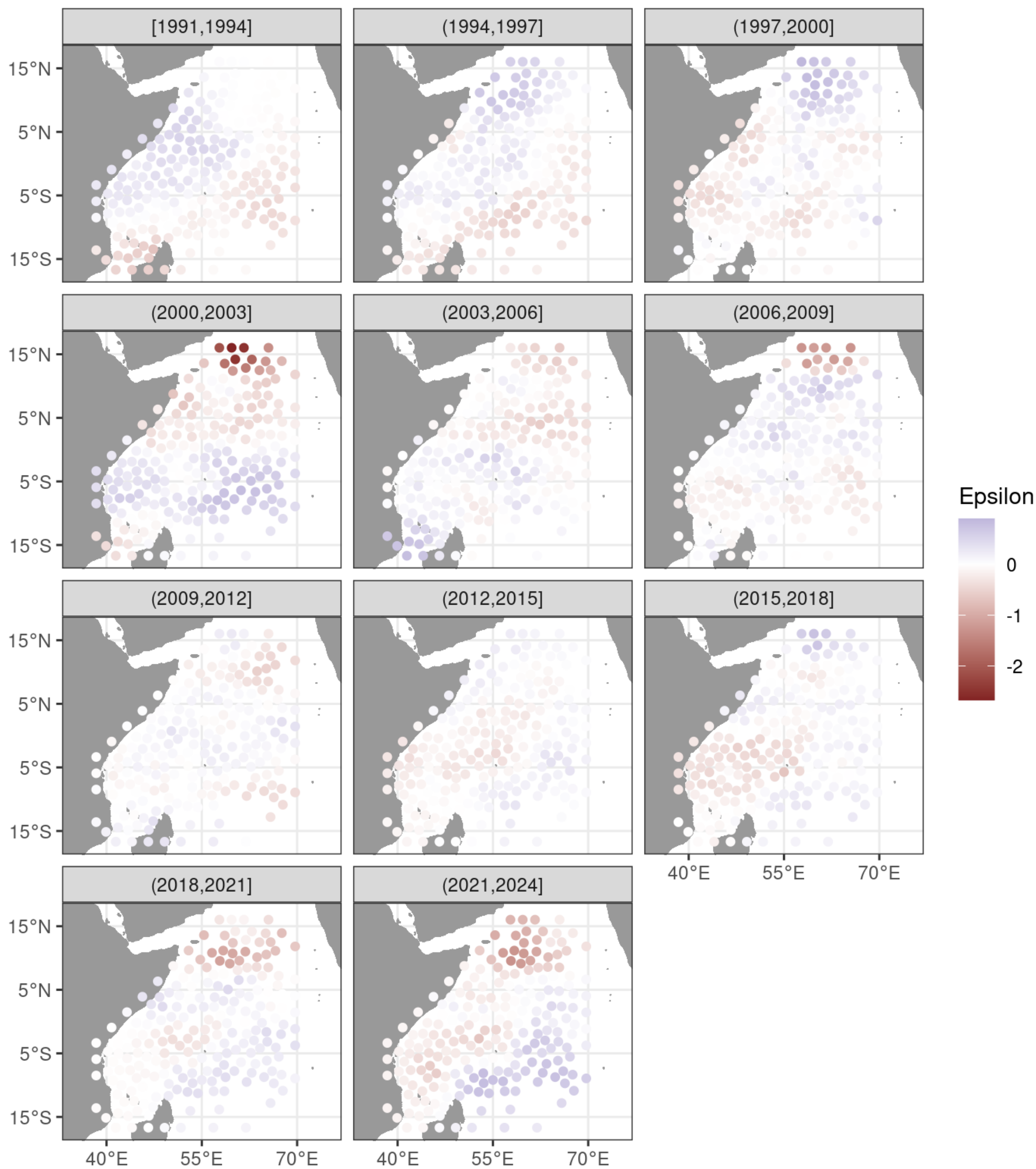


Figure 29: Spatiotemporal term for long time series model (1991-2023). Values have been meaned over each 3-year time period.

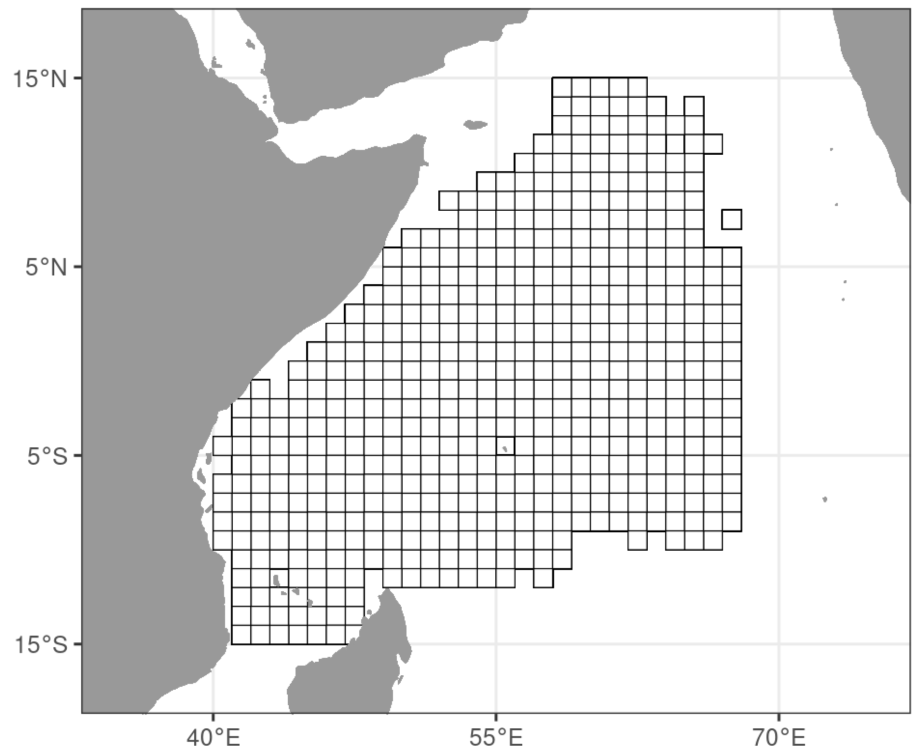


Figure 30: Extrapolation area composed by 1° × 1° grids. This area was used when predicting CPUE values over space and time for the long times series (1991-2023).

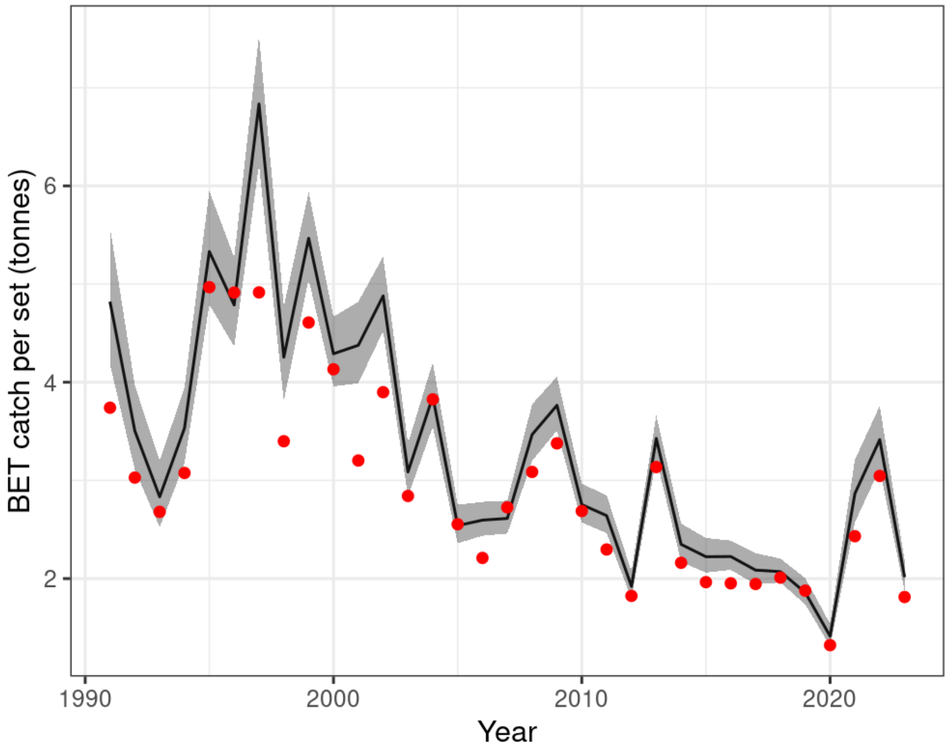


Figure 31: Annual standardized CPUE index for the long time series (1991-2023). Gray area represents the 95% confidence interval. Red dots represent the nominal CPUE.

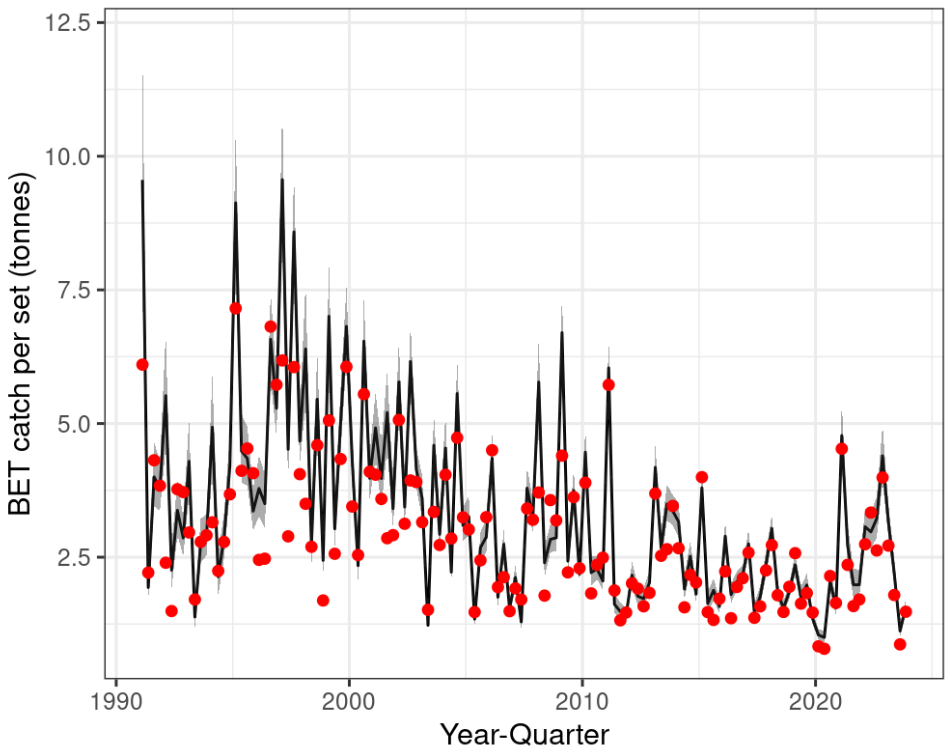


Figure 32: Quarterly standardized CPUE index for the long time series (1991-2023). Gray area represents the 95% confidence interval. Red dots represent the nominal CPUE.

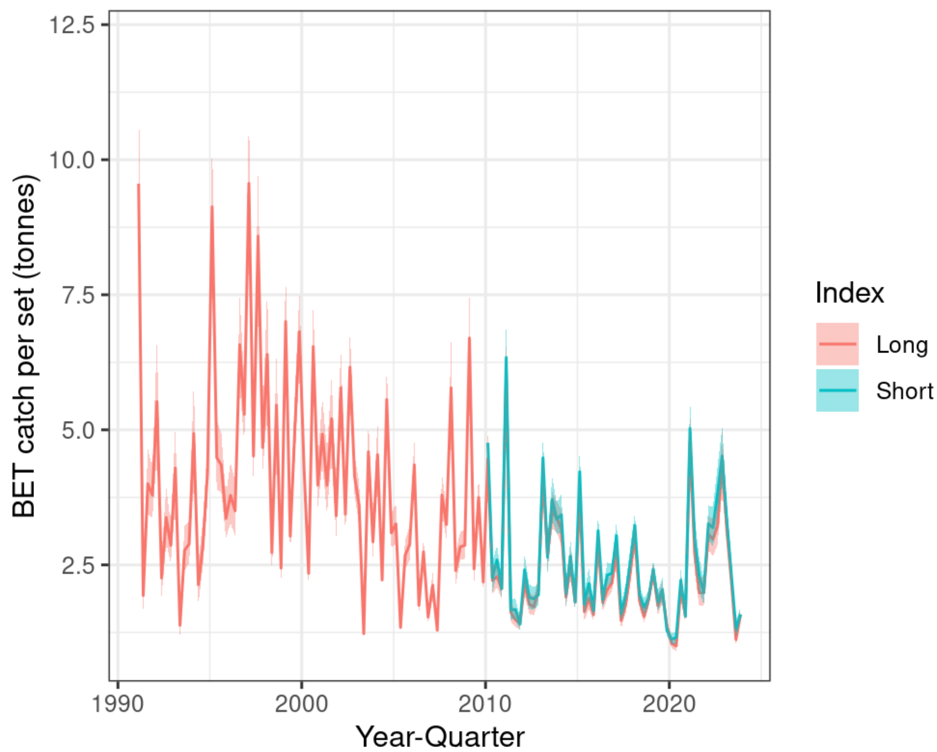


Figure 33: Quarterly standardized CPUE index for the long and short time series superposed.

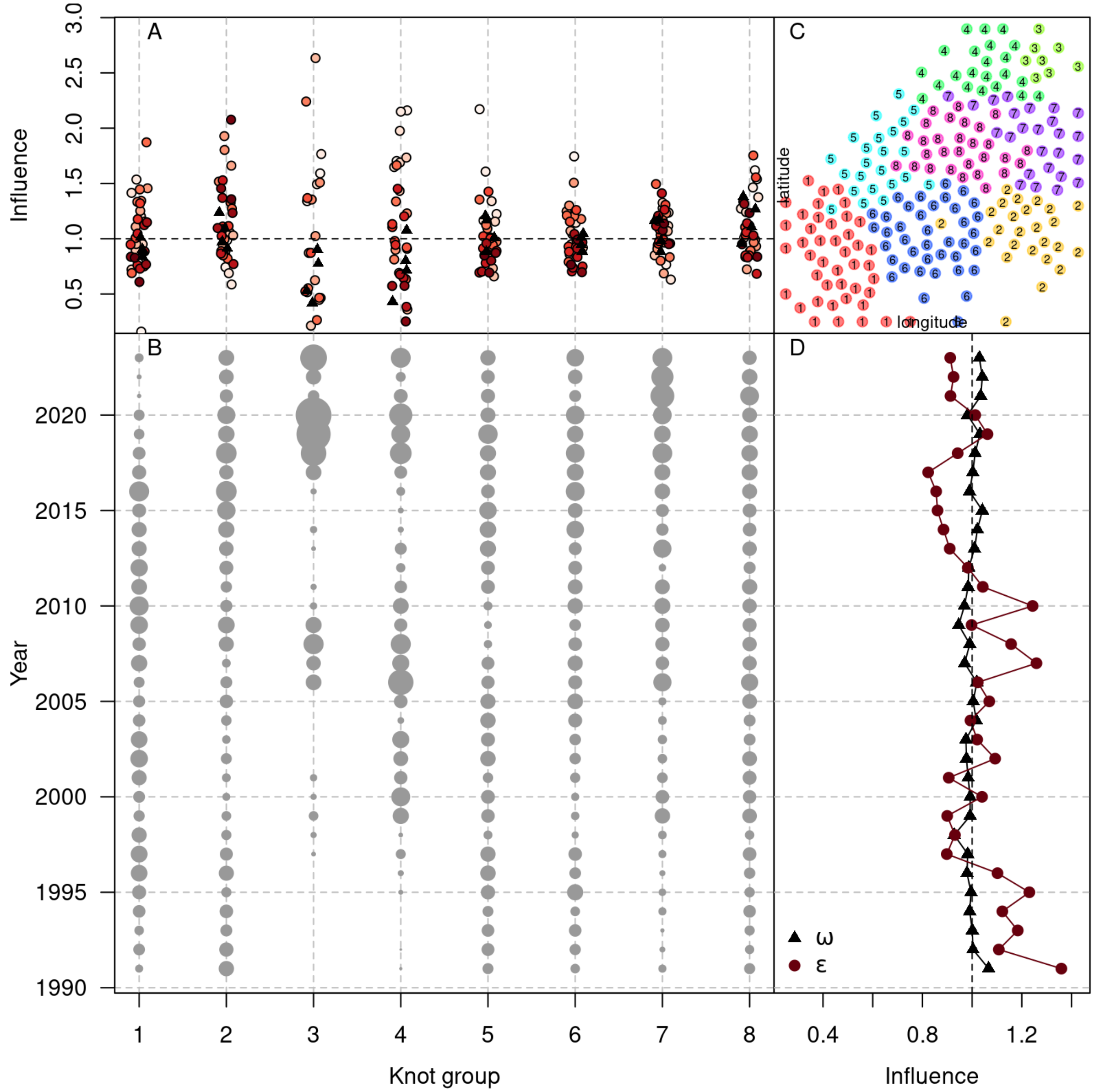


Figure 34: Coefficient-distribution-influence plot for the long times series (1991-2023). Figure A shows the coefficient by knot group for the spatial (black triangles, each triangle represents a quarter) and spatiotemporal (red dots, each dot represents a year) term. The knot groups are shown in Figure C. Figure B shows the number of observations per year and knot group. Figure D shows the influence coefficient per year for the spatial and spatiotemporal term.

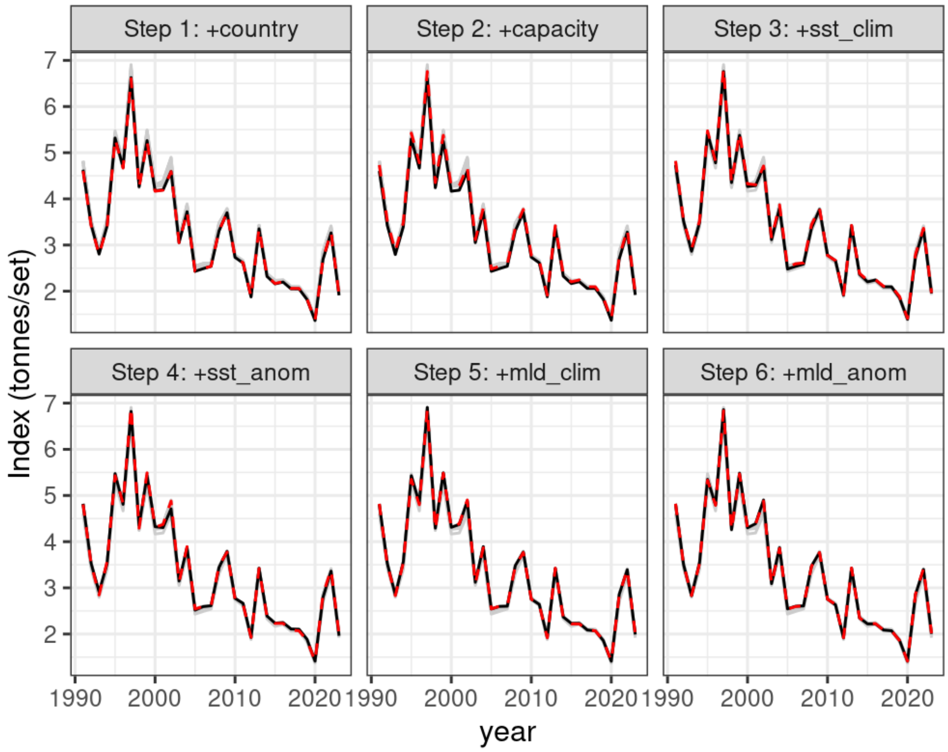


Figure 35: Step plot to evaluate the effect of adding a new covariate on the annual standardized CPUE index for the long time series (1991-2023). The numbers indicate the steps, and variable name indicates the covariate that was added at that step. The black line represents the CPUE index at that step, the dashed red line is the CPUE index from the previous step, and the gray lines indicate all the CPUE indices from all steps. Note that year-quarter, vessel random effect (`(1|numbat)`), spatial (ω_{sq}) and spatio-temporal (ϵ_{st}) terms were included in all models.

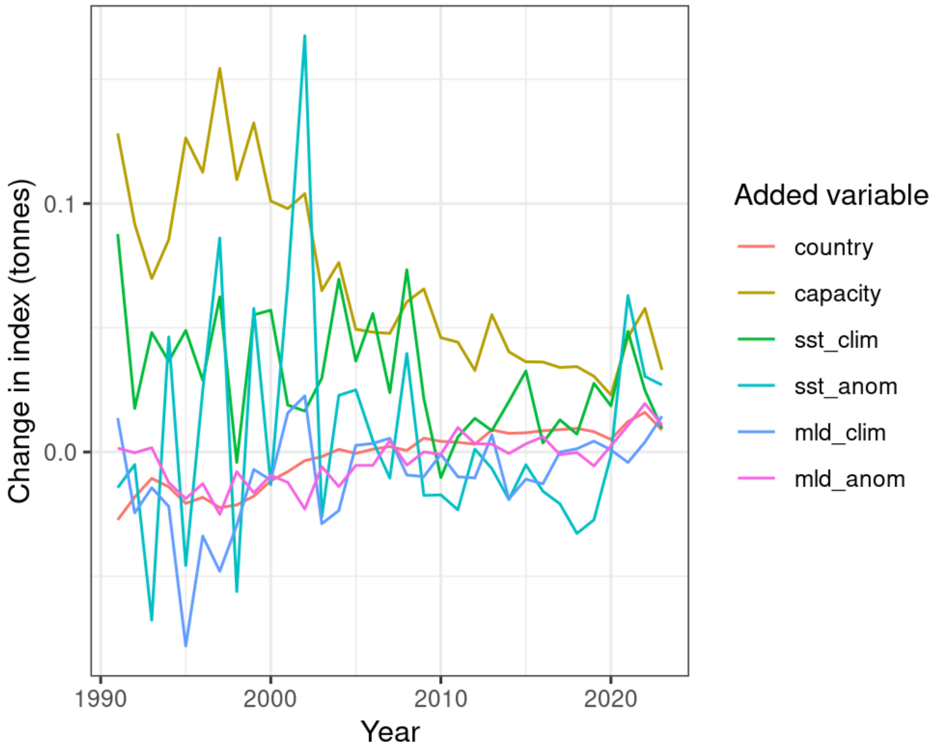


Figure 36: Change in annual standardized CPUE index for the long time series (1991-2023) due to adding a new covariate to the model formula. Color indicates which covariate was added to the formula and the covariates are presented in the legend in the order they were added to the model. Note that year-quarter, vessel random effect (`(1|vessel)`), spatial (ω_{sq}) and spatio-temporal (ϵ_{st}) terms were included in all models.

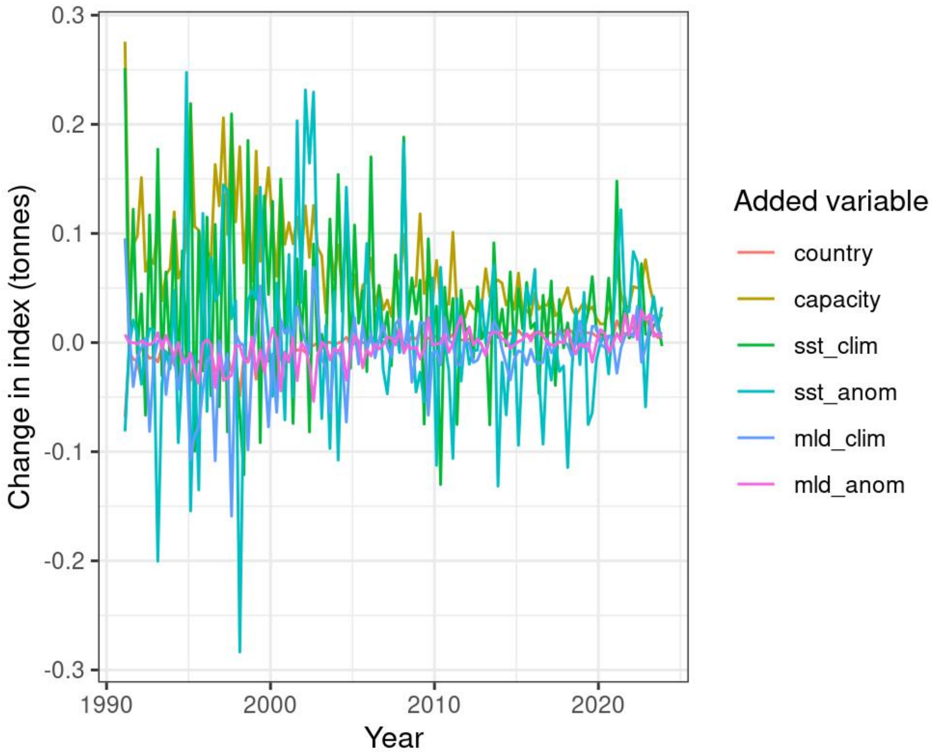


Figure 37: Change in quarterly standardized CPUE index for the long time series (1991-2023) due to adding a new covariate to the model formula. Color indicates which covariate was added to the formula and the covariates are presented in the legend in the order they were added to the model. Note that year-quarter, vessel random effect (`(1|vessel)`), spatial (ω_{sq}) and spatio-temporal (ϵ_{st}) terms were included in all models.

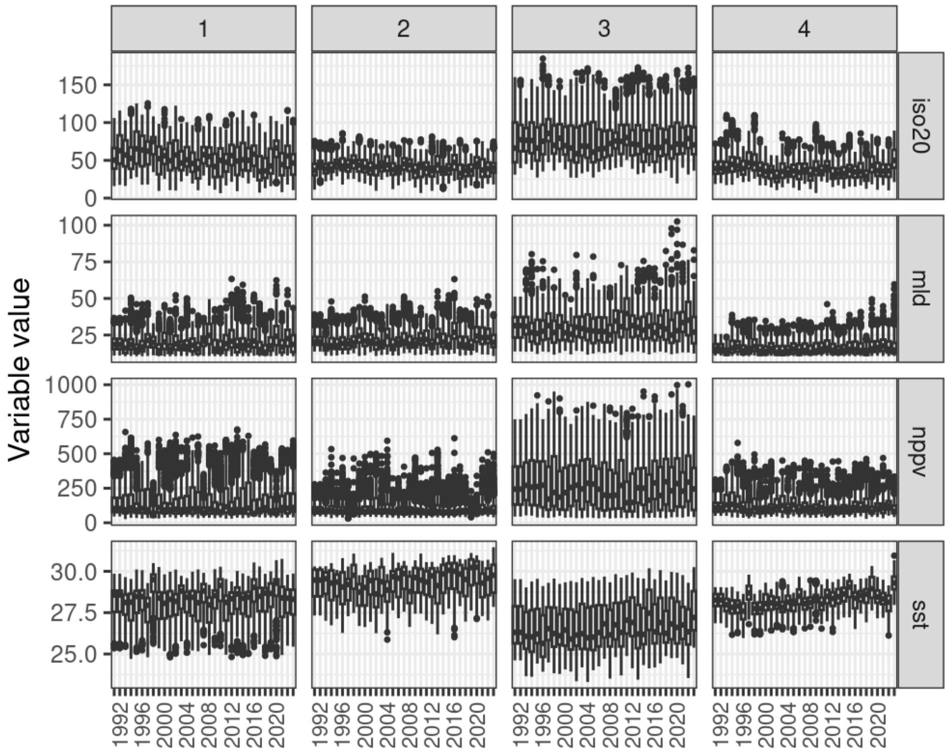


Figure 38: Temporal trends in oceanographic conditions. Boxplots are composed of environmental information per grid of the extrapolation area for the long time series (1991-2023).

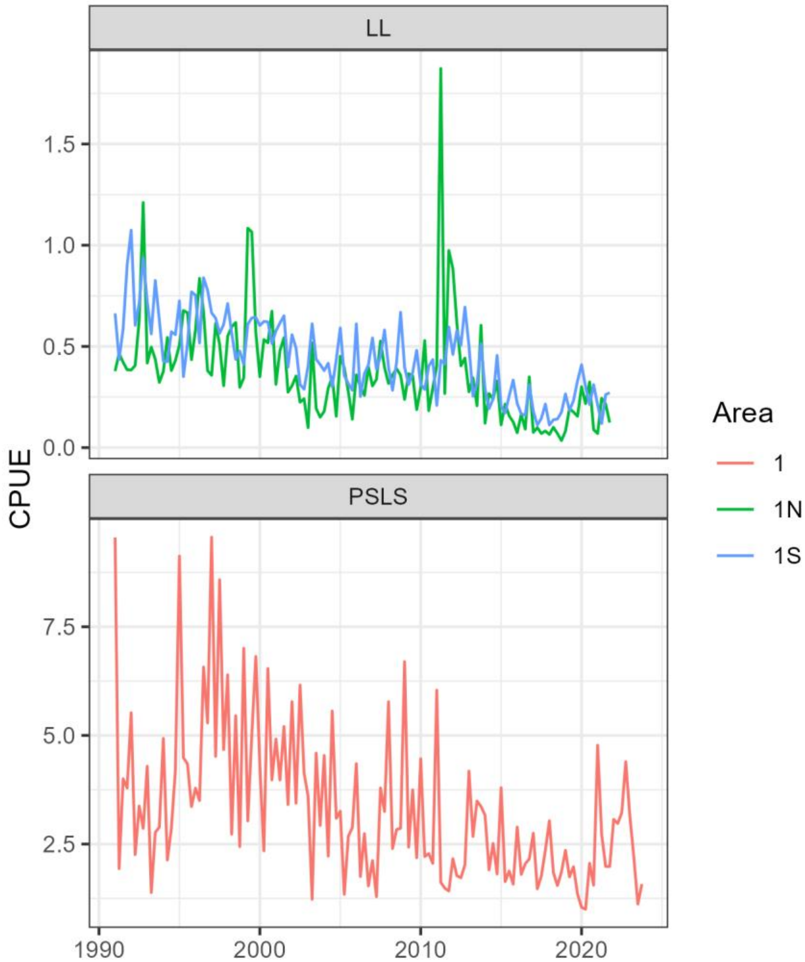


Figure 39: Comparison between the long FOB index (PSLS) obtained in this study with the longline indices for areas 1N and 1S used in the 2022 BET stock assessment.