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

Giancarlo M. Correa[1,✉], David M. Kaplan[2,3], Maitane Grande[4], Jon Uranga[4], Gorka Merino[4], Taha Imzilen[2,3], Antoine Duparc[2,3], Laurent Floc'h[2,3], José Carlos Báez Barrionuevo[5], Julissa de la Rosa[5], and María Lourdes Ramos Alonso[6]

[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] Centro Oceanográfico de Málaga, Instituto Español de Oceanografía, Consejo Superior de Investigaciones Científicas, Puerto Pesquero, s/n Apdo., 29640, Fuengirola, Málaga, Spain
[6] Centro Oceanográfico de Canarias, Instituto Español de Oceanografía, Consejo Superior de Investigaciones Científicas, Calle Farola del Mar No. 22-Dársena Pesquera 38180, Santa Cruz de Tenerife, Spain

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

SUMMARY

Indian Ocean EU purse seine floating object (FOB) school catches of skipjack tuna (SKJ; Katsuwonus pelamis) per fishing set for the period 1991-2025 were standardized with two geostatistical spatiotemporal modelling approaches using the sdmTMB R package. One approach considered only the recent time period 2010-2025, 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 1991-2025, 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 large, multi-year fluctuations with a relatively weak long-term negative temporal trend, though over the most recent period (>2010), estimated abundance is more or less stable with an increase in abundance over the period 2021-2022 followed by somewhat lower, but not historically unusual, values for the period 2023-2025. The index provided here can be incorporated into the 2026 skipjack stock assessment model to inform changes in biomass of skipjack tuna in the Western Indian Ocean.

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*).

Skipjack tuna is a tropical and subtropical pelagic fish characterized by rapid growth, early maturation, and high fecundity, which underpin its resilience to exploitation. It is widely distributed across the Atlantic, Indian, and Pacific Oceans, typically inhabiting warm surface waters (~ 15–30 °C) and forming large schools often associated with productive oceanographic features such as fronts and upwelling zones (Artetxe-Arrate et al., 2021). Skipjack exhibits opportunistic feeding behavior, preying primarily on small pelagic fish, crustaceans, and cephalopods, and plays a key role as both predator and prey in pelagic ecosystems. Reproduction is characterized by batch spawning throughout the year in equatorial waters (with seasonal patterns at higher latitudes), high fecundity, and pelagic eggs and larvae (Grande et al., 2014). Its migratory behavior is largely driven by the search for favorable thermal and trophic conditions, linking feeding and spawning habitats across broad oceanic regions. Combined with a relatively short lifespan (~6–12 years) and fast turnover, these life-history traits support high population productivity and dominance in global tuna catches.

The Stock Synthesis assessment platform was used in the 2023 SKJ assessment (Fu, 2023). There were four axes of uncertainty in the final uncertainty grid, which included two sets of growth parameters, two levels of effort creep in abundance indices, three steepness values, and the inclusion of three sets of CPUE indices. The results of the assessment show that, in 2022, the Indian SKJ stock was most likely not overfished and not subject to overfishing. The main indices of abundance used in the 2023 SKJ assessment were derived from the catch and effort data from the pole and line fishery and the purse seine fishery operating on dFADs.

In this study, we present two new (juvenile) SKJ abundance indices using data from the EU tropical tuna purse-seine fishery operating on FOBs, one covering the time period 1991-2025 and the other covering the recent time period 2010-2025. 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., 2025). These indices can inform juvenile SKJ 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 2025. 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 2010-2025

For data limited to the time period 2010-2025, 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.
- We removed observations with missing information for any of the used variables, as well as duplicated rows.
- We removed data in areas east of 73°E to avoid a large hole in the distribution due to the closure of the Chagos Archipelago to fishing in 2010.
- We removed data in areas north of 20°N and south of 20°S.
- 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.
- We excluded sets with a SKJ catch lower than 1 ton due to large uncertainty in small SKJ catch per set. We also eliminated a single set from 2022 that nominally consisted of more than 1000 tonnes of tropical tunas as we determined this to be a data entry error.

After applying these filters, we retained 72,993 observations.

2.1.2 1991-2025

For the full dataset (1991-2025), we applied the following filters to our initial dataset:

- We excluded null sets (i.e., sets indicated in log books as having not managed to capture the targeted fishing school).
- We excluded any fishing sets consisting of less than 1 tonne of the major tropical tunas (SKJ, YFT, BET). This was done because we considered sets with such small catches of the major tunas to be highly uncertain and potentially misclassified as positive sets instead of null sets. We also eliminated a single set from 2022 that nominally consisted of more than 1000 tonnes of tropical tunas as we determined this to be a data entry error. This step eliminated 603 fishing sets corresponding to 0.4% of the initial dataset.
- 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 142294 FOB fishing sets by 65 distinct fishing vessels, each representing between 676 and 6015 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 11,231 fishing sets (8.6%).
- The resulting dataset was limited to $1° \times 1°$ cells where FOB fishing occurred over at least 10 different years between 1991 and 2025. This filter eliminated 1555 fishing sets (1.2%).
- Finally, we removed any sets outside the spatial zone for which environmental data is available. This filter eliminated 16 fishing sets (0.01%).

The final 1991-2025 dataset consisted of 129492 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 SKJ catches per set were centered.
4. *Center of gravity (lat)*: Indicates the latitude where the SKJ catches per set were centered.
5. *Moran's autocorrelation coefficient*: Measure of spatial autocorrelation of the SKJ catch per set (Gittleman and Kot, 1990).
6. *Gini coefficient*: It is a measure of inequality (Cowell, 2011) among SKJ 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., 2025). 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., 2025) 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. (2025), our model can be mathematically represented as:

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

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

Where the expected value $\mathbb{E}[.]$ of an observation y (SKJ 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, v_v is the vessel effect (*numbat*) treated as random effects, and $\mathbf{X}$ 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 SKJ 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 SKJ catch equal to zero that remained roughly constant over time (Figure 1), we replaced the zero values with the minimum SKJ catch value found in our dataset (0.08 t). Then, we tested two probability distributions for the response variable that had in previous years been effective for tropical tuna CPUE standardization: lognormal and generalized gamma (Dunic et al., 2025), which were then compared through AIC. We selected the family with the lowest AIC.

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 daily temporal resolution. Depth-integrated (0-100 m) net primary production ($mg/m^3/day$) was downloaded from Copernicus-PICES biogeochemical global hindcast with a $0.25° \times 0.25°$ spatial resolution and daily temporal resolution. Depth of the 20°C isotherm (m) was downloaded from the NCEP Global Ocean Data Assimilation System with a $0.33° \times 1°$ 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 (1993-2025 for SST, MLD and NPPV; 1991-2025 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 (2019) 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, we standardized continuous catchability and density covariates to a mean of 0 and standard deviation of 1 (z-score standardization). 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, 2019). 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').

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 SKJ catch per set

were skewed to the left, with values generally smaller than 50 tons and rarely above 100 tonnes (Figure 2). In the log scale, we noticed that the SKJ catch per set values remained roughly stable over the years, with somewhat larger values in the late 1990's and early 2000's. When comparing among quarters, we noticed that quarter 1 generally had larger catch per set values (Figure 2). The proportion of sets with SKJ catch equal to zero was generally below 5%, except for a few exceptional quarters in the mid-2000's (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 SKJ 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 northwards since ~2010, but remained stable for quarters 3 and 4 (Figure 9). The Moran index indicated that the SKJ 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 targeting during the "golden years" (Fonteneau et al., 2008). Finally, the Gini index indicated that the heterogeneity of SKJ catch per set values increased during the mid-2000's and has shown a longterm decrease since that time (indicating increased consistency of catches across sets; Figure 9).

3.2 Spatiotemporal model

3.2.1 Short time series (2010-2025)

When including all the candidate covariates in a simple linear model for the short time series (2010-2025), 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 northern areas, especially in quarters 3 and 4; however, this pattern was less evident for quarters 1 and 2 (Figure 15). The spatiotemporal term is shown in Figure 16 and generally showed, like the spatial term, larger positive values in northern areas for most years.

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 northern areas (Figure 18), with some remarkable high values in some years (e.g., 2018, 2022). We aggregated these spatial predictions to obtain a quarterly index of abundance. The CPUE index showed an increase around 2018, and remained roughly

stable up to 2023 (Figure 19 and Table 4). After 2023, we observed a decrease in 2024 and a tendency to increase during the last quarters of 2025.

When evaluating the influence of the spatial and spatiotemporal terms on the model, we found that the spatial term for knot group 4 (Figure 20, which coincided with the northern zone which generally had higher predicted CPUE) had a positive influence on predicted values. Over the years, we did not find large influence of the spatiotemporal term (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 (1991-2025)

When including all candidate covariates in a simple linear model for the long time series (1991-2025), we found that the variance inflation factors associated with each of them were not larger than 5, but those for *sst_clim* and *nppv_clim* were very close to 5 (Figure 23). 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*, either climatology or anomaly, in the model selection process. We also found that the depth of the 20°C isotherm (*iso20*) had important correlations with the other environmental variables, particularly mixed layer depth (*mld*; Figure 24). Given these correlations, the uncertainty regarding whether *iso20* should be treated as a density or catchability covariate and the potential redundancy with *mld* in terms of impacts on abundance and catch rates, we decided not to include the depth of the 20°C isotherm in long time series models. After removing *nppv* and *iso20* covariates from the simple linear model, all VIF values for the remaining covariates were well below 5.

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).

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 major spatial autocorrelation in residuals (Figure 27).

The spatial term showed similar patterns to those of the short time series model, with larger values in the north of the model domain, particularly in quarters 3 and 4 (Figure 28). The spatiotemporal term was more variable over time, but generally had positive values in the north-west of the model domain and negative values in the south (Figure 29).

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 high levels of multi-year variability, with notable peaks in the early 1990's and early 2000's and, to a lesser degree, the period 2018-2022 (Figure 31, Figure 32 and Table 7). Values are similar to the nominal CPUE, though there are some important corrections 2002-2006 and 2014-2018. The long and short times series during the period of temporal overlap between the two showed similar patterns, though in absolute terms the short time series consistently estimated somewhat higher catch per set values (Figure 33). After scaling the short time series to have the same mean value as the long time series for the period of temporal overlap, the two time series are seen to be very close to each other (Figure 34).

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 35). Knot group 3, corresponding to the northern extremes of the model domain, is relatively poorly represented in the early parts of the time series and had a positive influence on estimated CPUE, similar to the pattern for the short time series. The influence of the spatiotemporal terms has varied over the years, but did not have strong or consistent impacts on model results (Figure 35). Overall, we did not find very large impacts of any covariate on the index (Figure 36, Figure 37 and Figure 38).

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/Tjbajx9BKjXPywM
- Quarterly long time series index: https://drive.ird.fr/s/BxSKZNWKBYD72kY
- Annual long time series index: https://drive.ird.fr/s/WREkBjRn5aM7yYB

4. Discussion

In this study, we used a geostatistical spatiotemporal model to standardize the SKJ 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 significant multi-year variability in abundance estimates. Typical standardized CPUE values and variability in CPUE were somewhat greater in the first half of the long time series (i.e., before ~2005) than in the second half, but there was no strong negative trend and since ~2010 abundance estimates have varied up and down with no obvious tend. Over the period of temporal overlap, the short time series index including factors accounting for dFAD use showed a very similar temporal trend to the long time series index that did not include these factors. These indices can help inform the abundance of juvenile SKJ 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 SKJ, but there impacts were often contradictory and quite small. One possible explanation for this is that our data on whether or not a vessel has access to tracking and tuna abundance information provided by dFAD tracking buoys is not fine-grained at the level of vessels and therefore we were forced to assume that all vessels from the same fishing company had access to tracking information, which may not be the case.

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). This may nullify the impact of time at sea on realized catches (as opposed to catches that would occur if vessels randomly sampled dFADs at various times since deployment).

5. References

Anderson, S.C., Ward, E.J., English, P.A., Barnett, L.A.K., Thorson, J.T., 2025. sdmTMB: An R Package for Fast, Flexible, and User-Friendly Generalized Linear Mixed Effects Models with Spatial and Spatiotemporal Random Fields. Journal of Statistical Software 115. https://doi.org/10.18637/jss.v115.i02

Artetxe-Arrate, I., Fraile, I., Marsac, F., Farley, J.H., Rodriguez-Ezpeleta, N., Davies, C.R., Clear, N.P., Grewe, P., Murua, H., 2021. A review of the fisheries, life history and stock structure of tropical tuna (skipjack Katsuwonus pelamis, yellowfin Thunnus albacares and bigeye Thunnus obesus) in the Indian Ocean. Advances in Marine Biology 88, 39–89. https://doi.org/10.1016/bs.amb.2020.09.002

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.

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., 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

Fu, D., 2023. Indian Ocean Skipjack tuna stock assessment 1950-2022 (Stock Synthesis) (No. IOTC-2023-WPTT25-09). Indian Ocean Tuna Comission.

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

Grande, M., Murua, H., Zudaire, I., Goñi, N., Bodin, N., 2014. Reproductive timing and reproductive capacity of the Skipjack Tuna (Katsuwonus pelamis) in the western Indian Ocean. Fisheries Research 156, 14–22. https://doi.org/10.1016/j.fishres.2014.04.011

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

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., 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

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

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., 2019. 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,…,2025-4)	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 (1993-2025)	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 (1993-2025)	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 (1993-2025)	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 (1990-2025)	Numeric	Long	Density
iso20_anom	Depth of 20C isotherm, anomaly	Numeric	Long	Density

Table 2: AIC values for short time series (2010-2025) models assuming different probability distributions for the response variable.

Family	AIC
Lognormal	574678.1
Generalized gamma	574361.3

Table 3: Summary table of the final model (generalized gamma) for the short time series (2010-2025). Coefficients associated with the year-quarter interaction are not shown.

term	estimate	std.error	p_value
countrySpain	0.18	0.04	<0.01
followYes_No-echo	-0.12	0.03	<0.01
followYes_Echo	0.02	0.01	0.01
n_buoys20	-0.02	0.00	<0.01
n_buoys250	0.02	0.00	<0.01
t_sunrise	-0.18	0.00	<0.01
sst	-0.12	0.01	<0.01
mld	-0.04	0.00	<0.01
iso20	-0.04	0.01	<0.01

Table 4: Standardized CPUE index for the short time series (2010-2025) and 95% confidence interval obtained from the model using the quarter-specific spatial term configuration.

year	quarter	est	lwr	upr	cv
2010	1	27.879	25.487	30.495	0.046
2010	2	21.103	19.051	23.377	0.052
2010	3	29.657	26.964	32.619	0.049
2010	4	21.181	19.416	23.106	0.044
2011	1	32.555	29.762	35.611	0.046
2011	2	18.851	17.023	20.875	0.052
2011	3	20.650	18.646	22.869	0.052
2011	4	18.634	17.190	20.199	0.041
2012	1	22.231	20.226	24.433	0.048
2012	2	14.150	12.799	15.644	0.051
2012	3	19.099	17.250	21.147	0.052
2012	4	15.204	13.896	16.636	0.046
2013	1	20.031	18.435	21.765	0.042
2013	2	16.544	15.080	18.151	0.047
2013	3	18.694	16.968	20.596	0.049
2013	4	20.060	18.542	21.701	0.040
2014	1	24.660	22.501	27.026	0.047
2014	2	15.810	14.317	17.460	0.051
2014	3	26.235	23.758	28.971	0.051
2014	4	15.950	14.598	17.426	0.045
2015	1	27.875	25.508	30.461	0.045
2015	2	15.049	13.677	16.557	0.049
2015	3	22.850	20.586	25.362	0.053
2015	4	15.474	14.253	16.800	0.042
2016	1	20.006	18.557	21.569	0.038
2016	2	19.364	17.848	21.008	0.042
2016	3	26.458	24.329	28.773	0.043
2016	4	19.207	17.912	20.596	0.036
2017	1	22.155	20.588	23.841	0.037
2017	2	18.259	16.813	19.830	0.042
2017	3	28.618	26.252	31.196	0.044
2017	4	31.495	29.123	34.060	0.040
2018	1	35.903	33.523	38.451	0.035
2018	2	23.632	22.171	25.189	0.033
2018	3	40.884	37.915	44.085	0.038
2018	4	24.849	23.359	26.434	0.032
2019	1	26.252	24.558	28.062	0.034
2019	2	20.659	19.097	22.348	0.040
2019	3	29.702	27.496	32.085	0.039

2019	4	28.698	26.838	30.688	0.034
2020	1	21.008	19.570	22.552	0.036
2020	2	17.364	16.194	18.619	0.036
2020	3	33.766	31.214	36.527	0.040
2020	4	24.981	23.289	26.796	0.036
2021	1	28.795	26.567	31.211	0.041
2021	2	23.211	21.483	25.078	0.039
2021	3	37.500	34.348	40.940	0.045
2021	4	34.561	31.866	37.484	0.041
2022	1	28.742	26.397	31.296	0.043
2022	2	24.586	22.495	26.871	0.045
2022	3	36.031	32.711	39.687	0.049
2022	4	30.612	27.879	33.613	0.048
2023	1	27.749	25.751	29.901	0.038
2023	2	22.753	20.982	24.674	0.041
2023	3	25.709	23.803	27.768	0.039
2023	4	22.616	21.213	24.112	0.033
2024	1	20.304	18.764	21.970	0.040
2024	2	21.745	19.992	23.653	0.043
2024	3	23.489	21.291	25.915	0.050
2024	4	20.557	18.864	22.402	0.044
2025	1	23.582	21.545	25.811	0.046
2025	2	23.865	21.904	26.001	0.044
2025	3	27.972	25.531	30.647	0.047
2025	4	31.579	29.223	34.126	0.040

Table 5: AIC values for long time series (1991-2025) models assuming different probability distributions for the response variable.

Family	AIC
Lognormal	1056780
Generalized gamma	1033147

Table 6: Summary table of the final model (generalized gamma) for the long time series (1991-2025). Coefficients associated with the year-quarter interaction are not shown. Note that continuous predictors have been standardized to have unit variance so that coefficients have units of *tonnes* and are comparable across predictors.

term	estimate	std.error	p_value
countrySpain	0.05728	0.02792	0.04024
capacity	0.08319	0.01503	0.00000
sst_clim	-0.09234	0.00873	0.00000
sst_anom	-0.05264	0.00472	0.00000
mld_clim	-0.07134	0.00541	0.00000
mld_anom	-0.02437	0.00355	0.00000

Table 7: Annual standardized CPUE index for the long time series (1991-2025) 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	se_natural	cv
1991	24.45	22.14	27.01	23.89	0.05
1992	24.63	22.52	26.93	24.08	0.05
1993	24.35	22.16	26.76	23.80	0.05
1994	25.86	23.67	28.24	25.30	0.04
1995	21.92	20.08	23.93	21.45	0.04
1996	19.06	17.59	20.64	18.70	0.04
1997	16.13	15.02	17.31	15.85	0.04
1998	15.19	13.94	16.55	14.86	0.04
1999	25.84	24.21	27.59	25.33	0.03
2000	27.56	25.73	29.51	26.99	0.03
2001	24.42	22.84	26.11	23.93	0.03
2002	30.75	28.81	32.82	30.19	0.03
2003	29.33	27.43	31.35	28.74	0.03
2004	25.36	23.49	27.38	24.85	0.04
2005	24.62	22.96	26.41	24.16	0.04
2006	23.69	22.40	25.05	23.30	0.03
2007	15.91	15.01	16.85	15.66	0.03
2008	19.09	17.97	20.29	18.79	0.03
2009	20.71	19.51	21.97	20.35	0.03
2010	20.02	18.85	21.26	19.68	0.03
2011	18.18	17.12	19.30	17.88	0.03
2012	14.32	13.39	15.32	14.08	0.03
2013	15.53	14.67	16.44	15.30	0.03
2014	16.45	15.32	17.66	16.16	0.04
2015	16.05	14.95	17.22	15.78	0.04
2016	17.12	16.21	18.07	16.87	0.03
2017	20.12	18.98	21.32	19.79	0.03
2018	24.39	23.21	25.64	24.03	0.03
2019	20.85	19.70	22.07	20.52	0.03
2020	18.93	17.89	20.04	18.65	0.03
2021	24.39	22.75	26.15	23.93	0.04
2022	23.98	22.31	25.77	23.52	0.04
2023	18.91	17.92	19.95	18.62	0.03
2024	16.92	15.85	18.07	16.63	0.03
2025	20.60	19.21	22.09	20.21	0.04

7. Figures

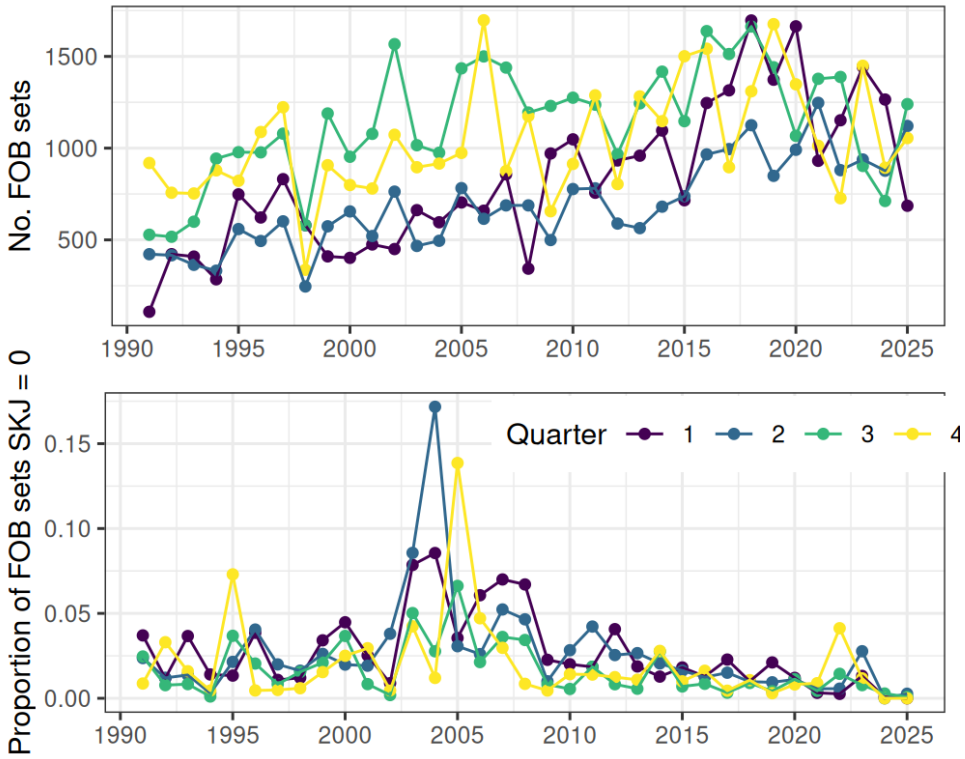


Figure 1: Number of non-null FOB sets in the data (top) and proportion of sets with SKJ catch equal to zero (bottom) by year and quarter for the long series (1991-2025).

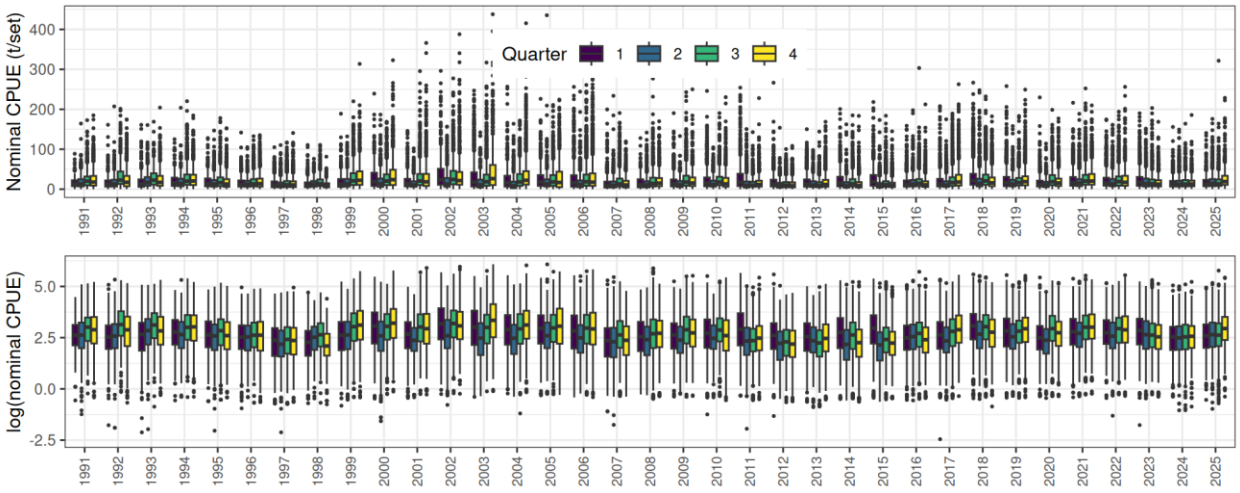


Figure 2: Distribution of SKJ catch per set, original (top) and log-transformed (bottom), by year and quarter for the long time series (1991-2025).

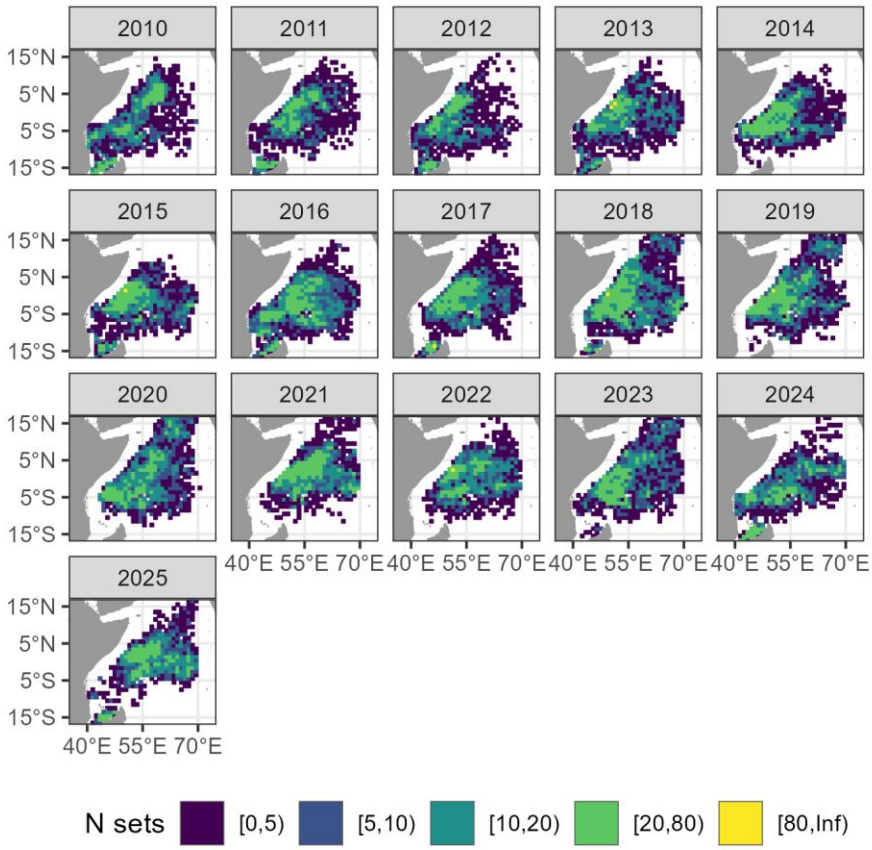


Figure 3: Number of FOB fishing sets (effort) per 1° × 1° grid cell and year for the short time series (2010-2025).

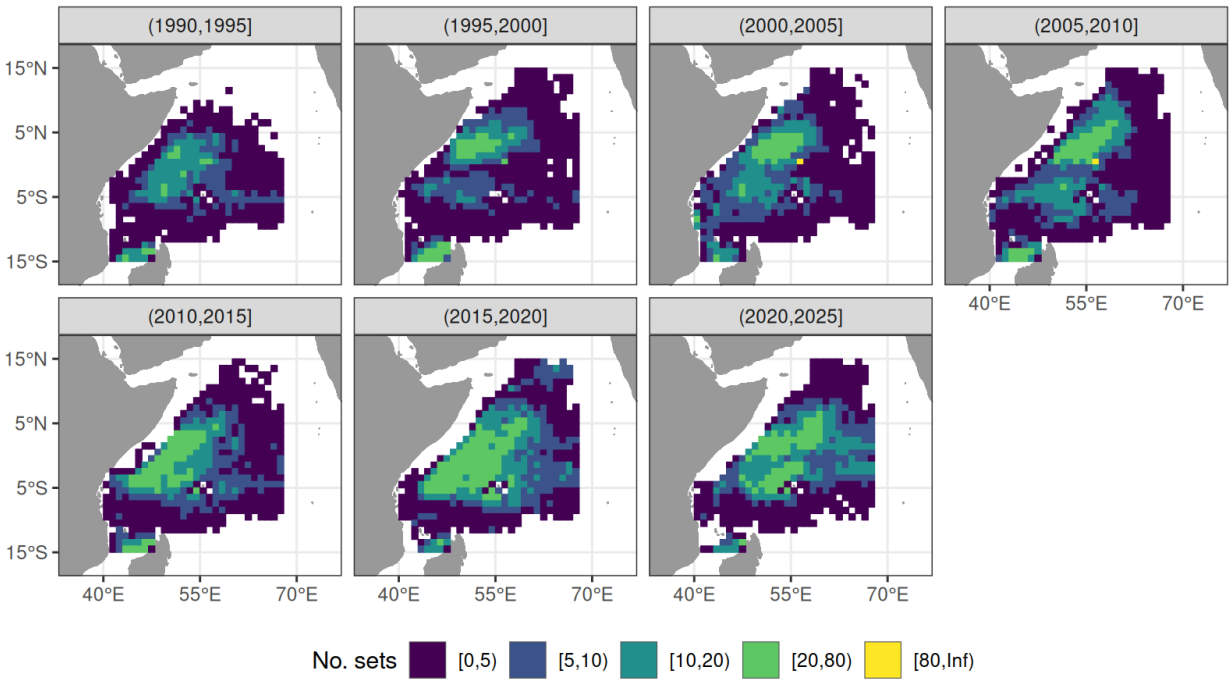


Figure 4: Average number of FOB fishing sets (effort) per 1° × 1° grid cell and per year for each 5-year time period for the long time series (1991-2025).

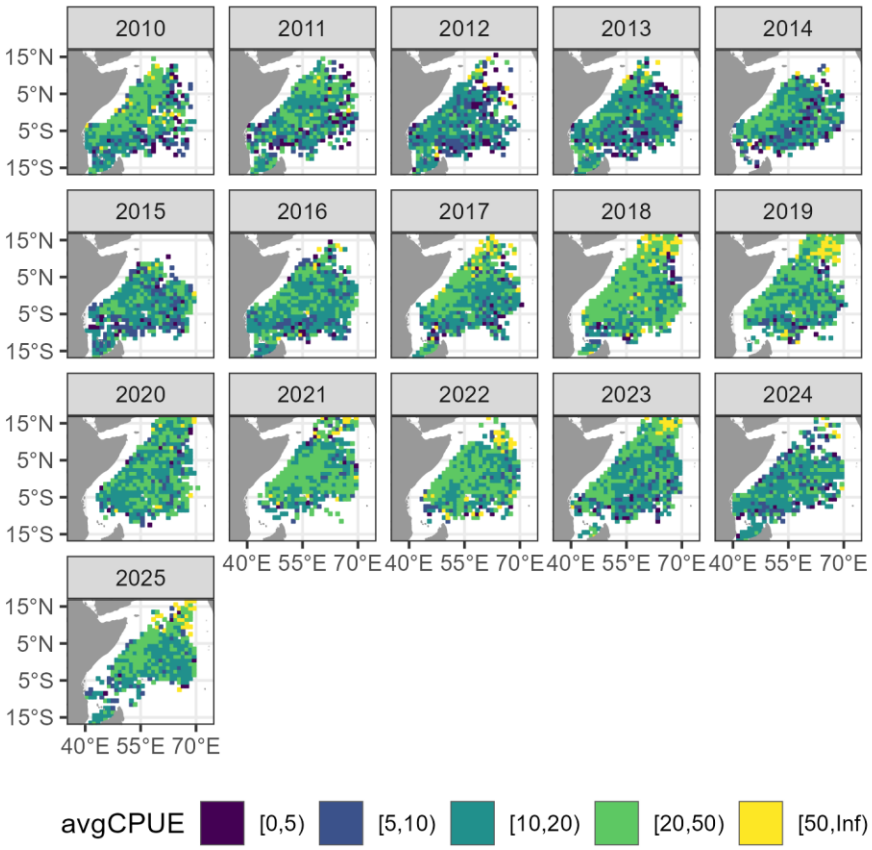


Figure 5: Observed average catch per set for each 1° × 1° grid cell and year for the short time series (2010-2025).

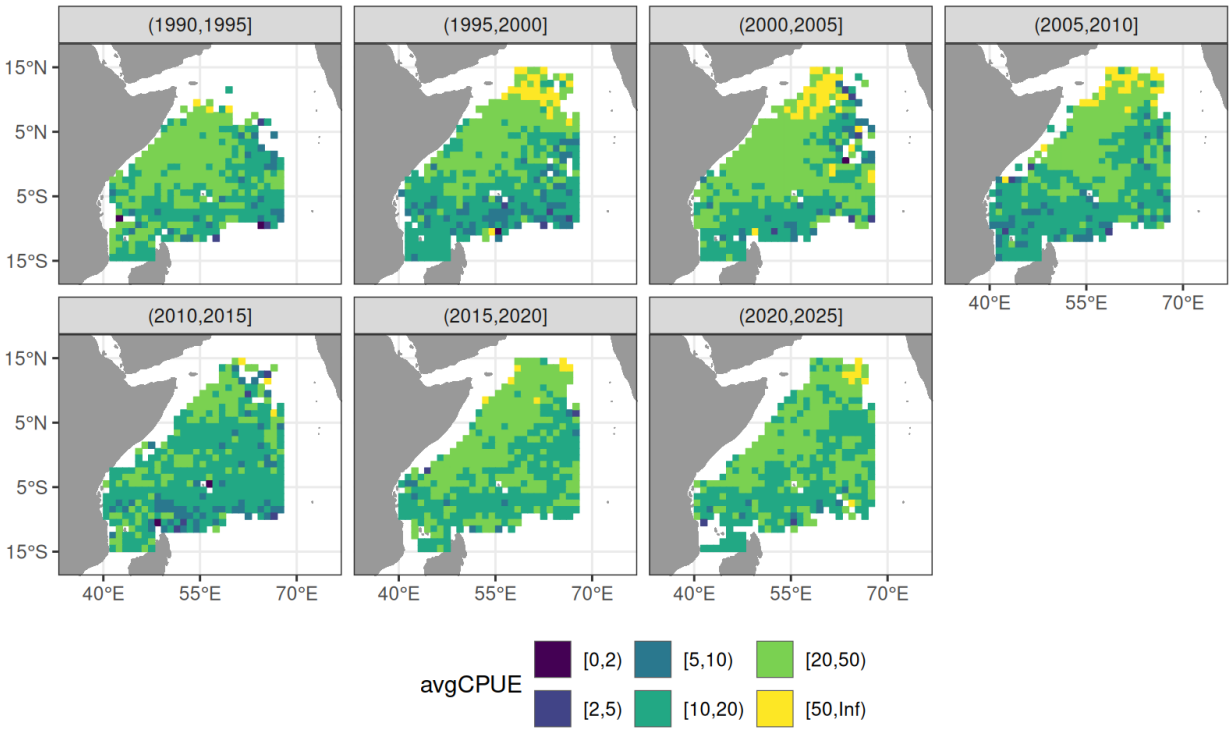


Figure 6: Average catch per set (CPUE) for each 1° × 1° grid cell and 5-year time period for the long time series (1991-2025).

Figure 7: Observed proportion of sets with zero SKJ catch per 1° × 1° grid cell and year in the short time series (2010-2025).

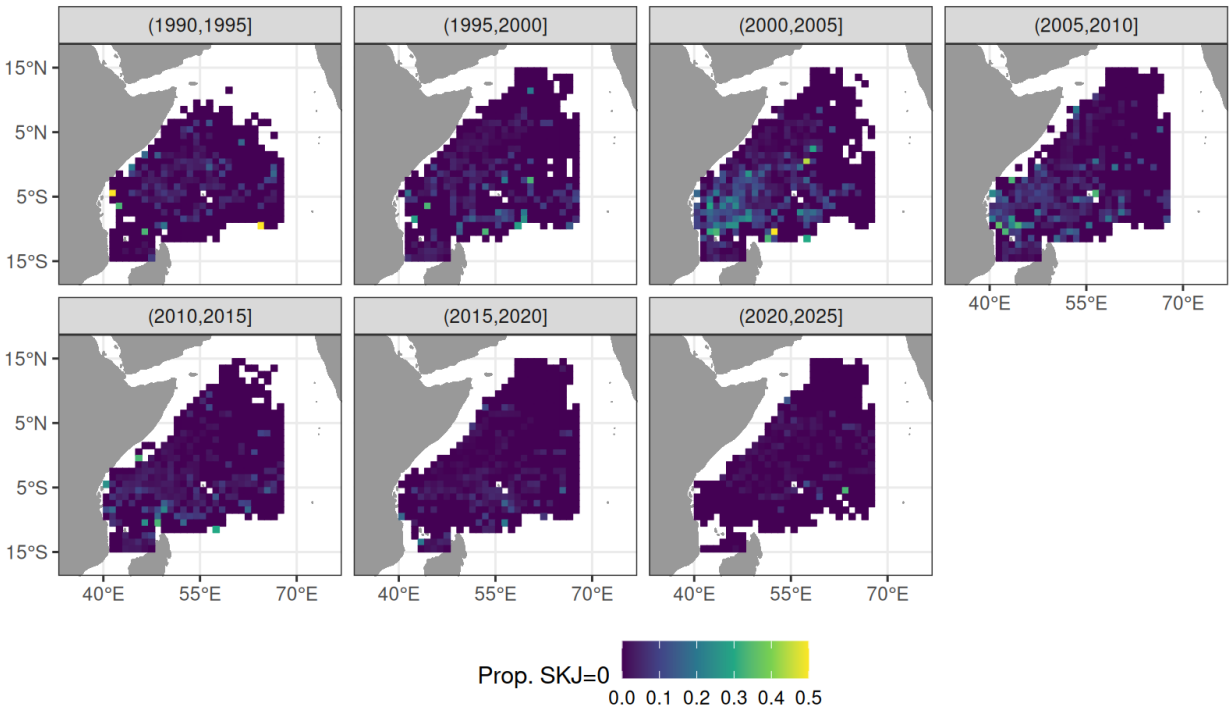


Figure 8: Observed proportion of sets with zero SKJ catch per 1° × 1° grid cell and 5-year time period for the long time series (1991-2025).

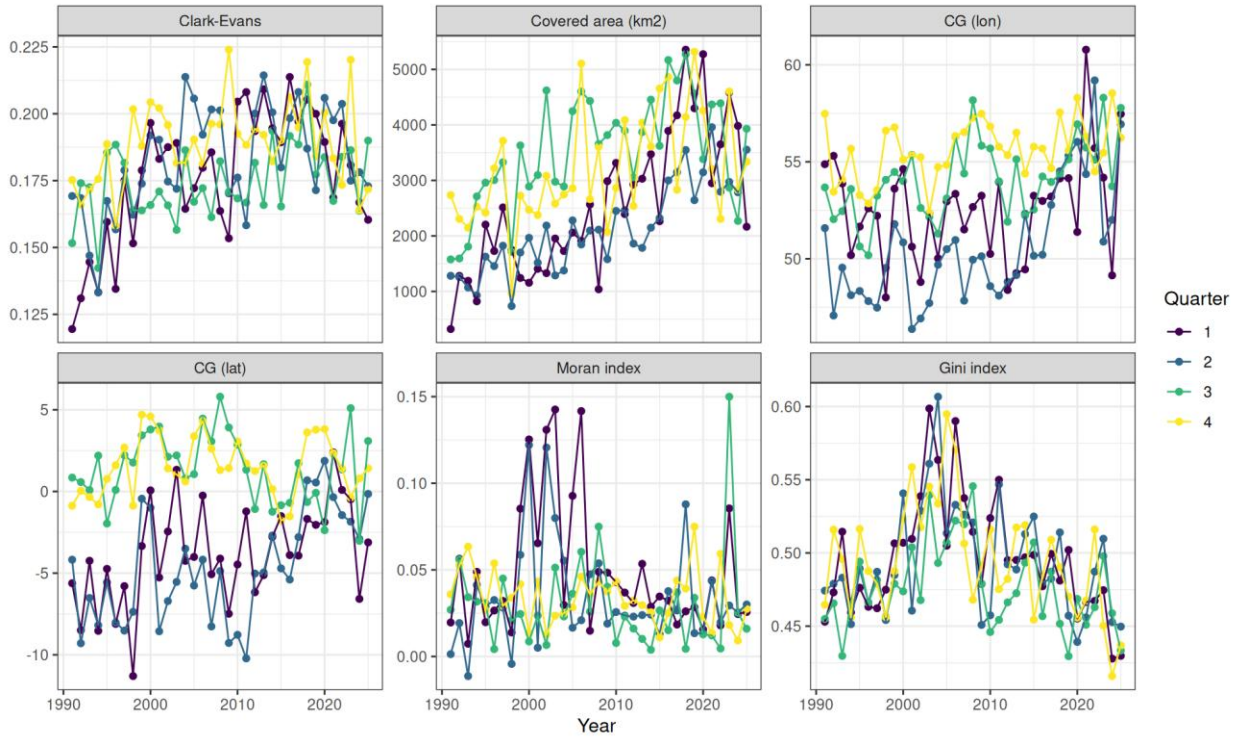


Figure 9: Spatial indicators calculated by quarter for the long time series (1991-2025).

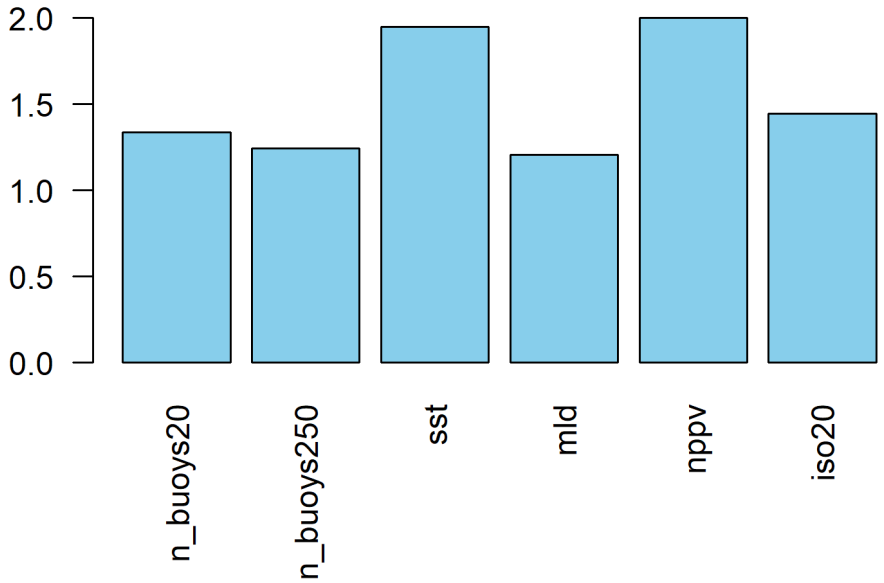


Figure 10: Variance inflation factors (VIF) by covariate obtained when fitting a simple linear model with short time series data (2010-202#).

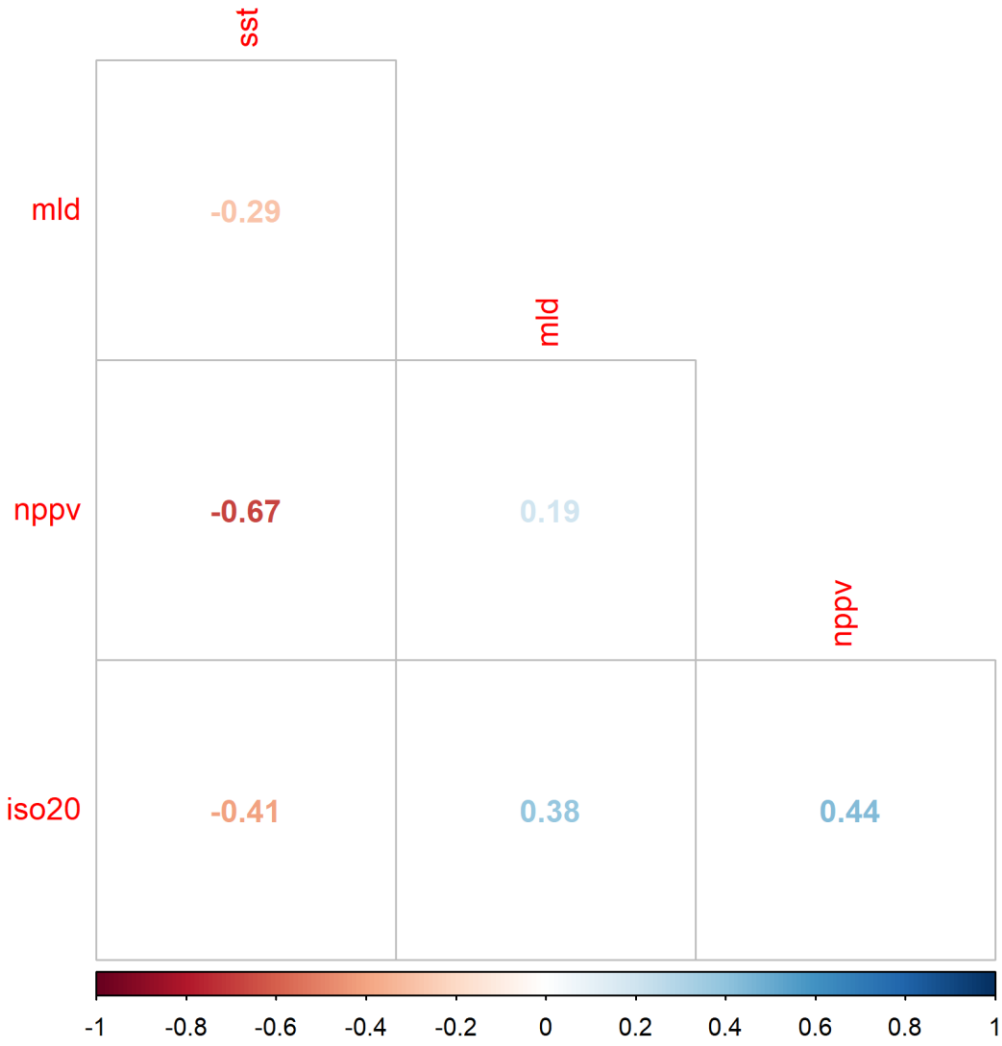


Figure 11: Correlation matrix among environmental covariates for the short time series (2010-2025).

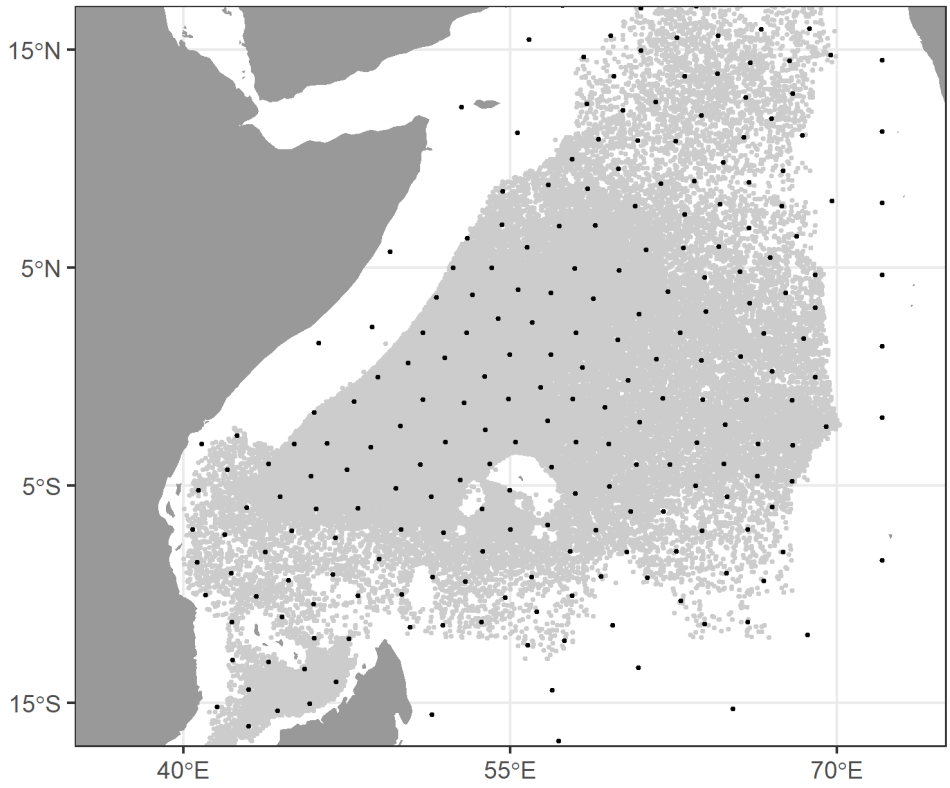


Figure 12: Mesh nodes (black dots) used in the spatiotemporal model and observations (gray dots) for the short time series (2010-2025).

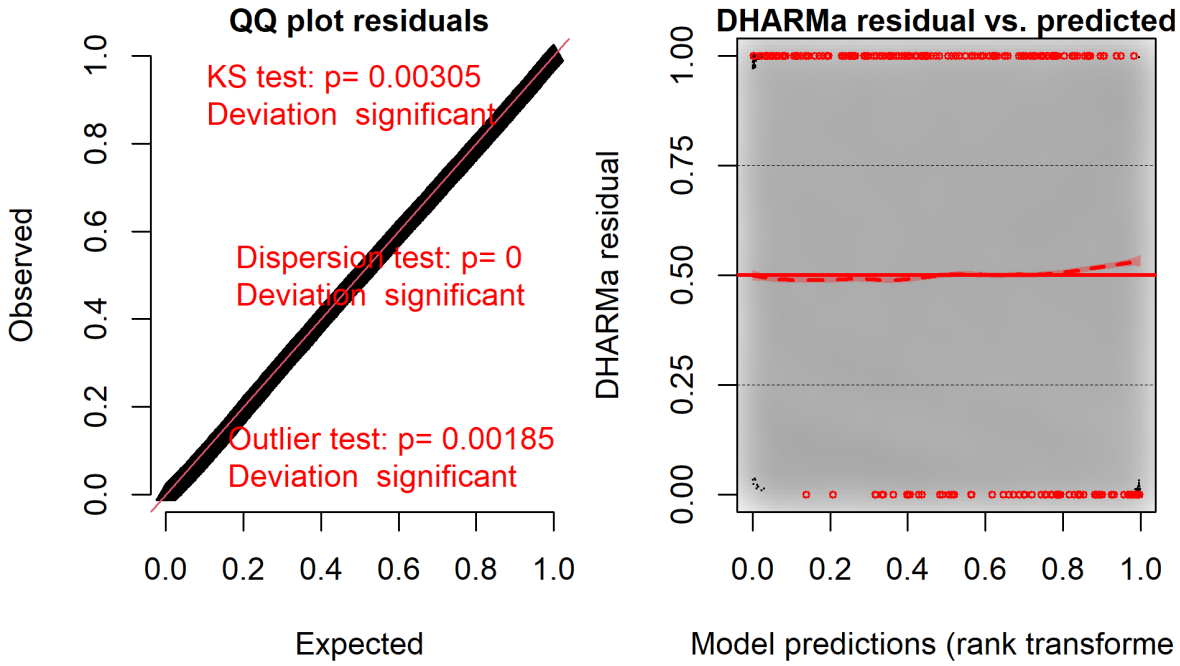


Figure 13: Simulation-based randomized-quantile residuals for the short time series model (2010-2025). QQ-plot (left) detects overall deviations from the expected distribution, by default with added tests for correct distribution (KS test), dispersion and outliers. Residual plot (right) shows the residuals against the predicted value.

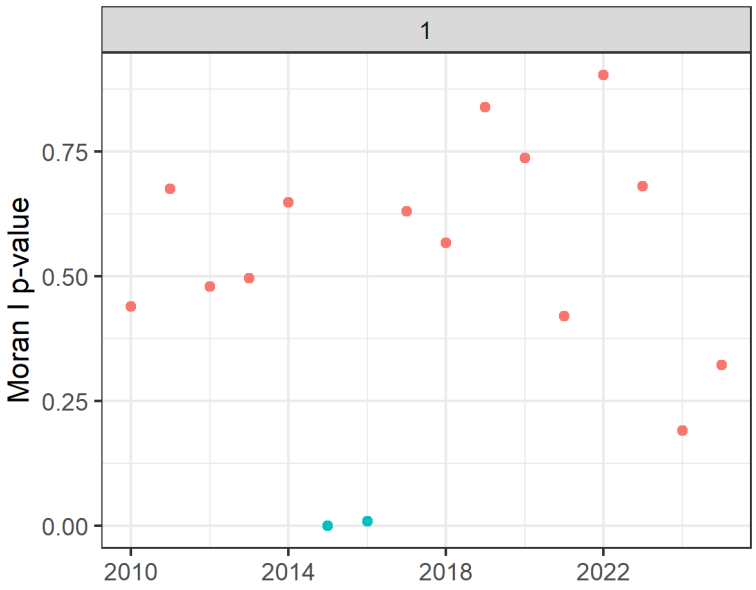


Figure 14: Moran I's p-value of the randomized quantile residuals for standardization model of short time series (2010-2025). Blue points represent years with significant spatial autocorrelation in residuals.

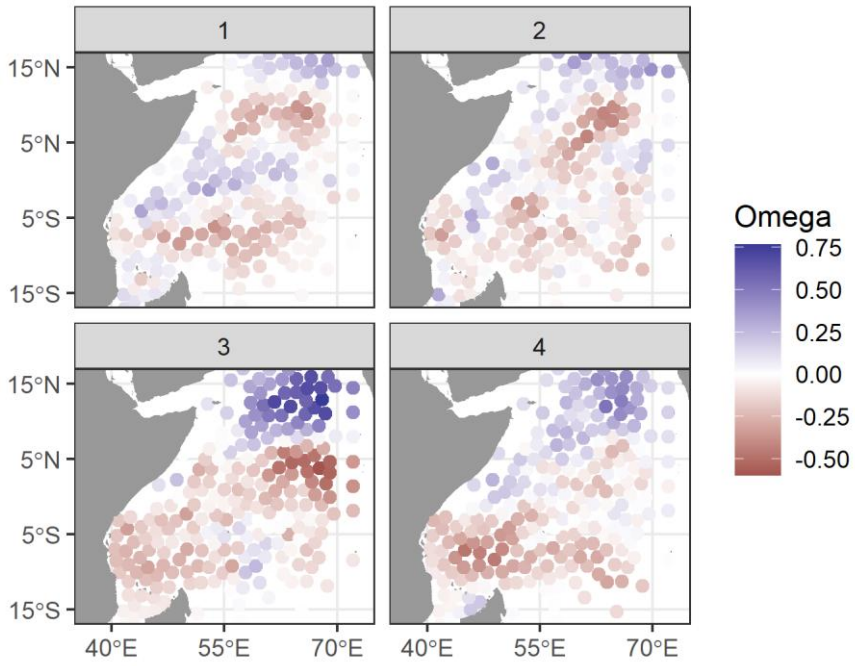


Figure 15: Quarter-specific spatial term for short time series model (2010-2025).

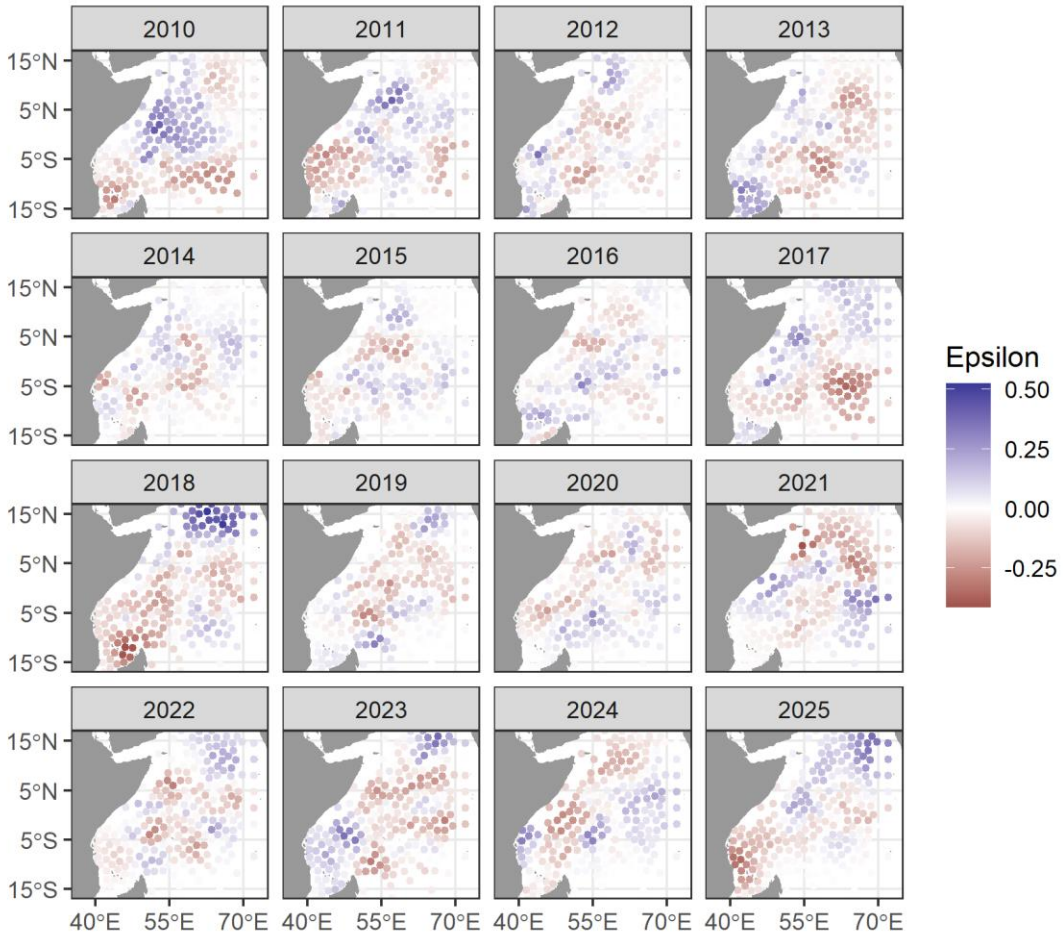


Figure 16: Spatiotemporal term for short time series model (2010-2025). To reduce the number of panels, the spatiotemporal term has been averaged over periods of three years.

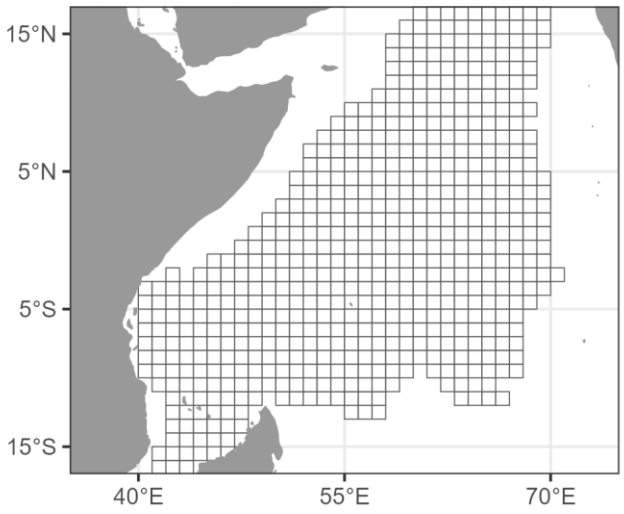


Figure 17: Extrapolation area composed by 1° × 1° grids. This area was used when predicting CPUE values over space and time for the short times series (2010-2025).

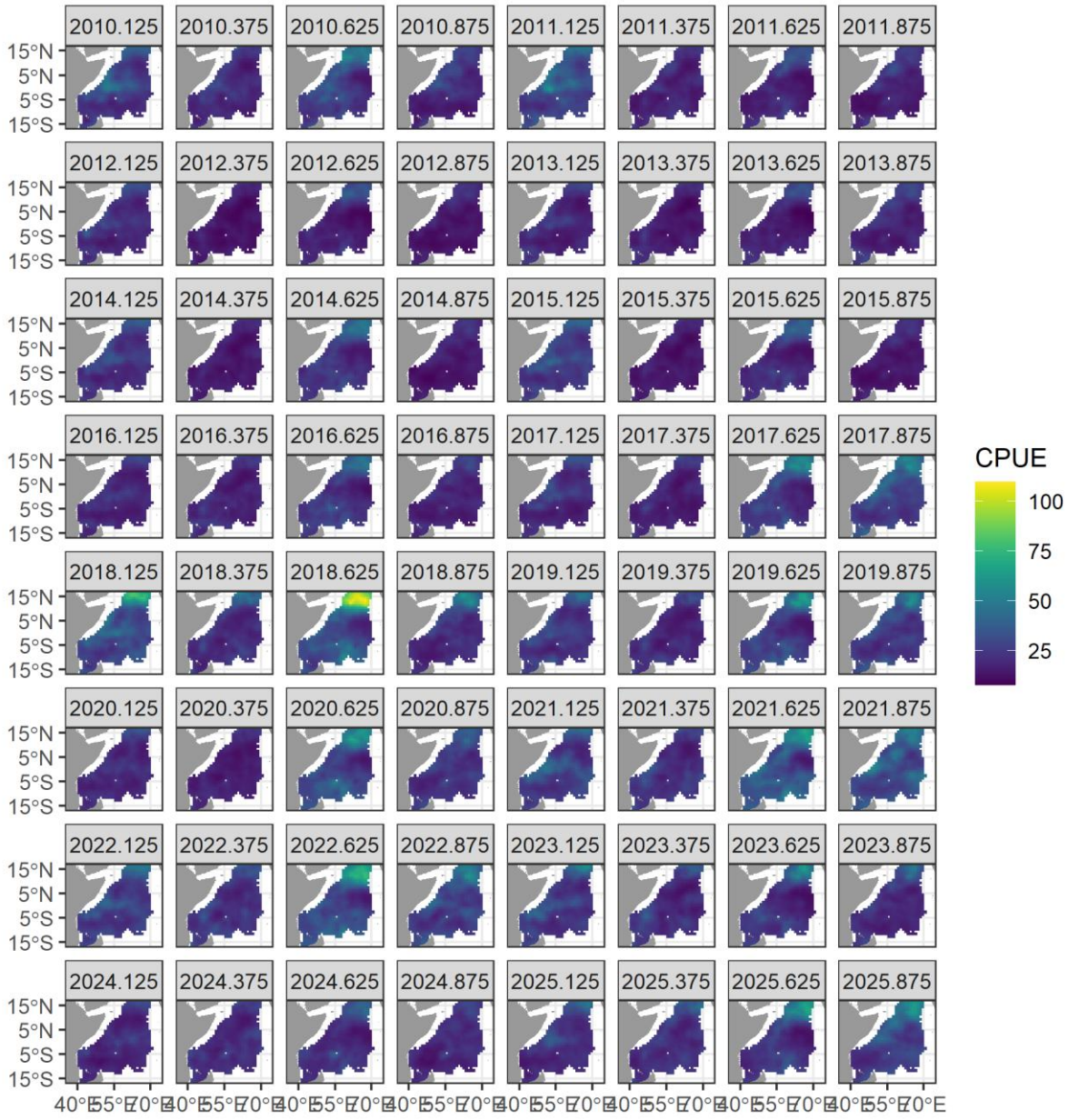


Figure 18: CPUE predictions per quarter over the extrapolation grid for the short time series (2010-2025).

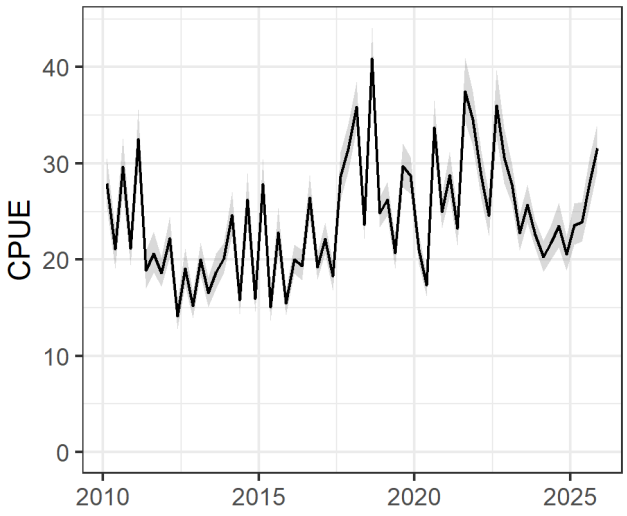


Figure 19: Standardized CPUE index for the short time series (2010-2025). Gray area represents the 95% confidence interval.

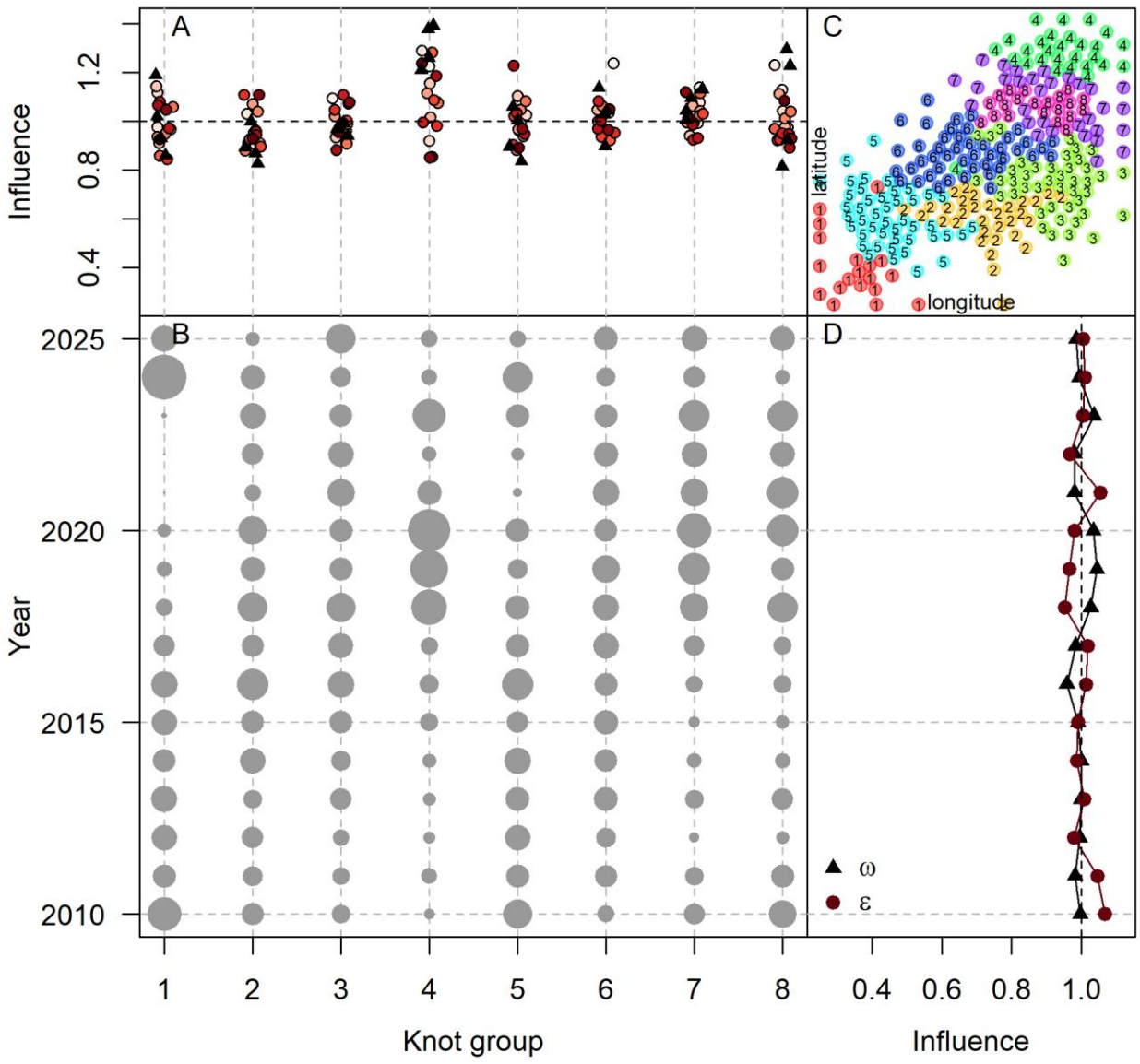


Figure 20: Coefficient-distribution-influence plot for the short times series (2010-2025). Figure A shows the coefficient by knot group for the spatial (black triangles, each triangle represents a quarter) and spatiotemporal (red dots, each dot represents a year) term. The knot groups are shown in Figure 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.

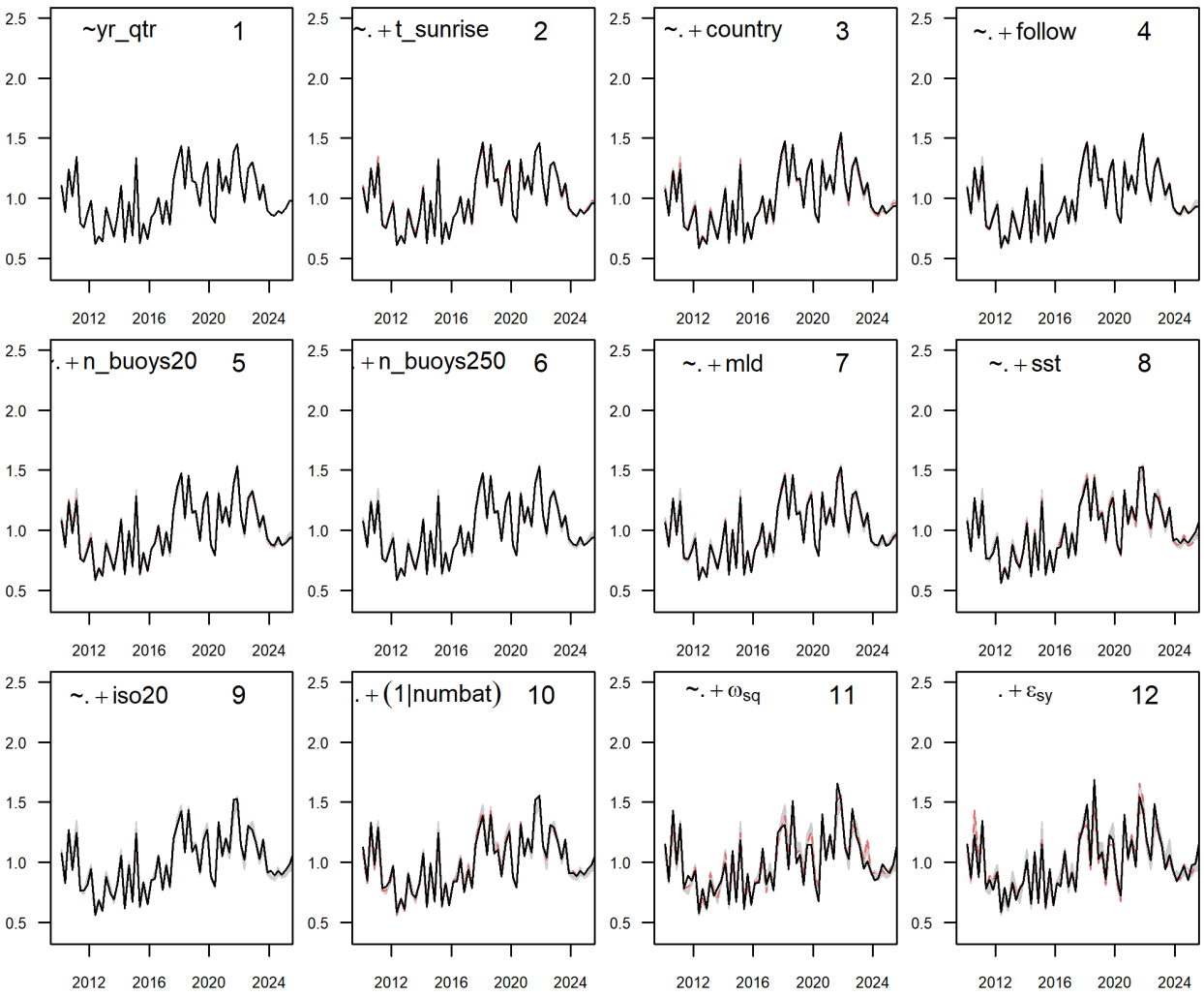


Figure 21: Step plot to evaluate the effect of adding a new covariate on the CPUE index for the short time series (2010-2025). The numbers indicate the steps, and the formula indicates the covariate that was added. The black line represents the CPUE index at that step, the red line is the CPUE index from the previous step, and the gray lines indicate all the previous CPUE indices.

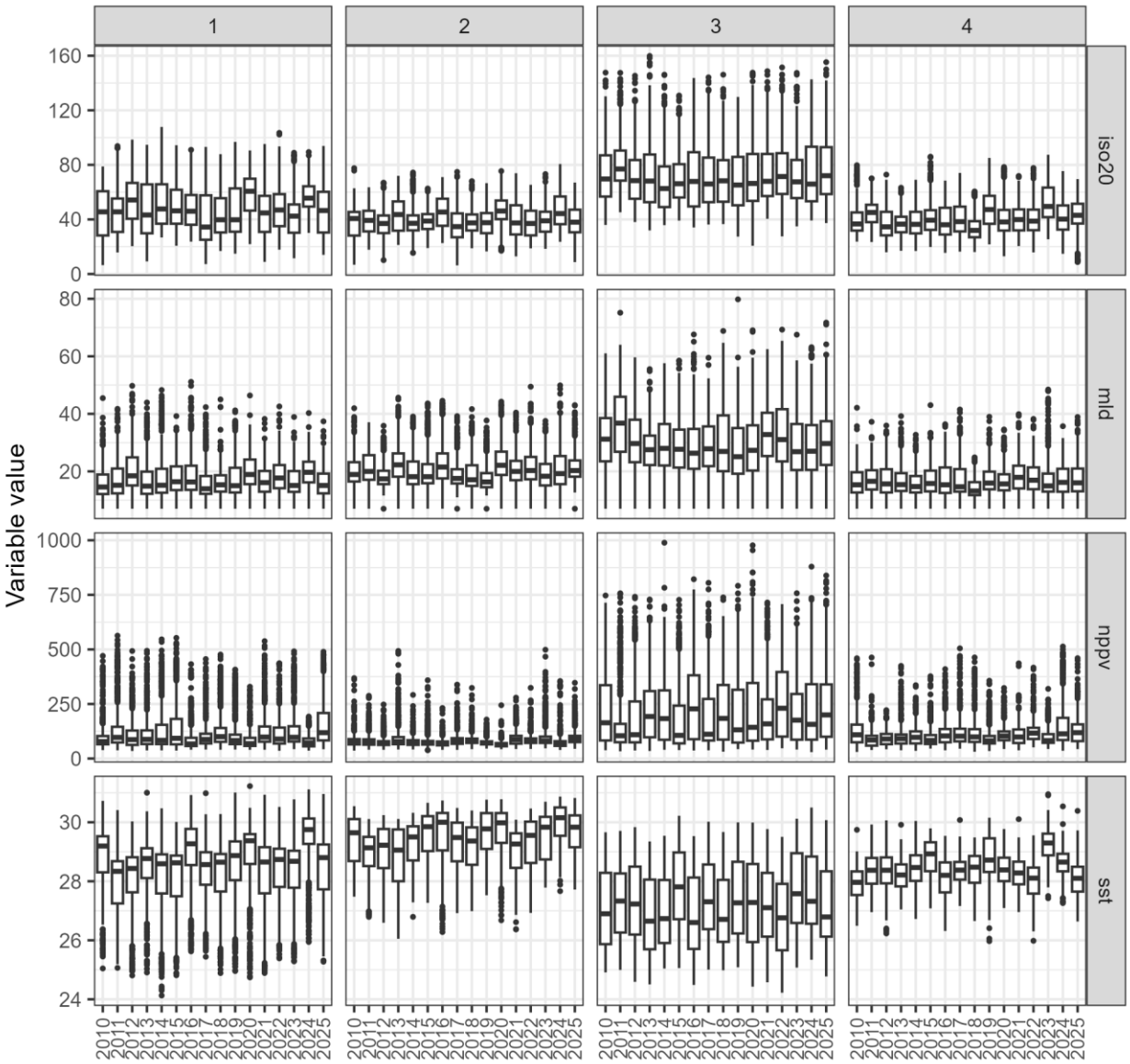


Figure 22: Temporal trends in oceanographic conditions. Boxplots are composed of environmental information per grid of the extrapolation area for the short itme series (2010-2025).

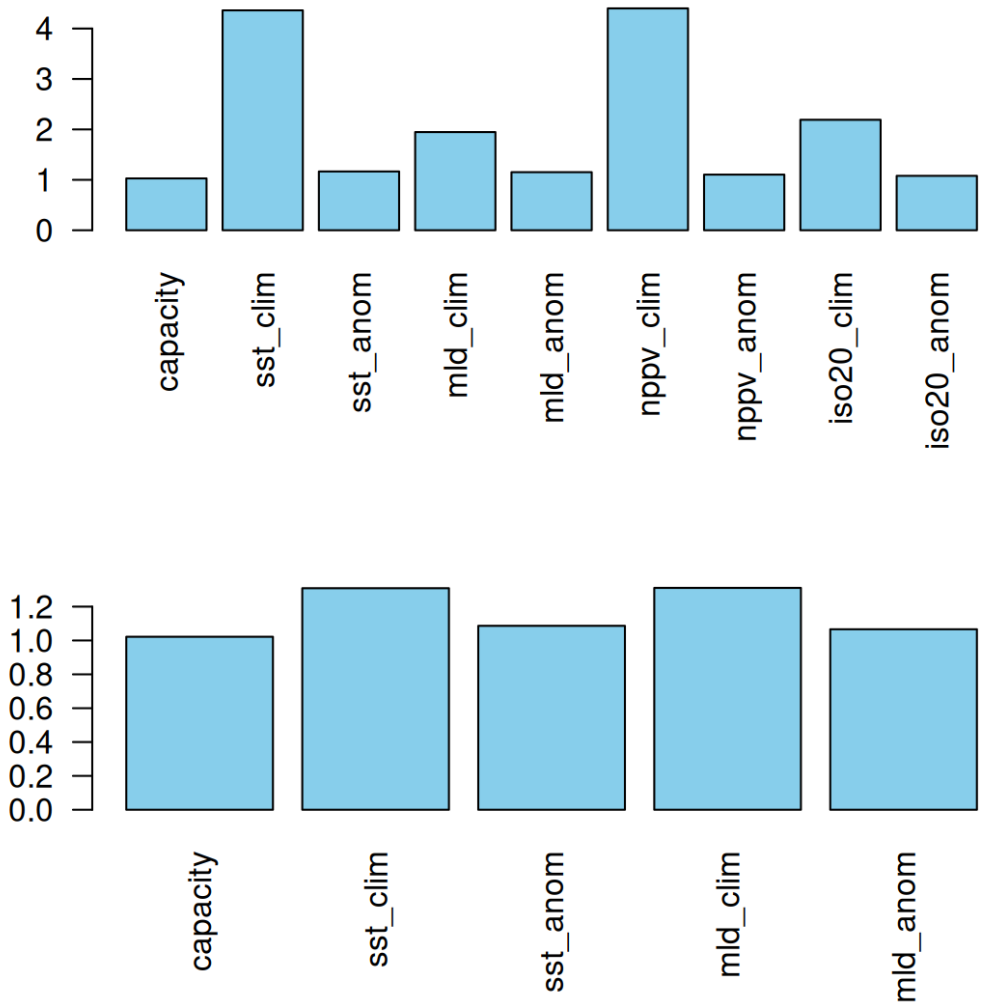
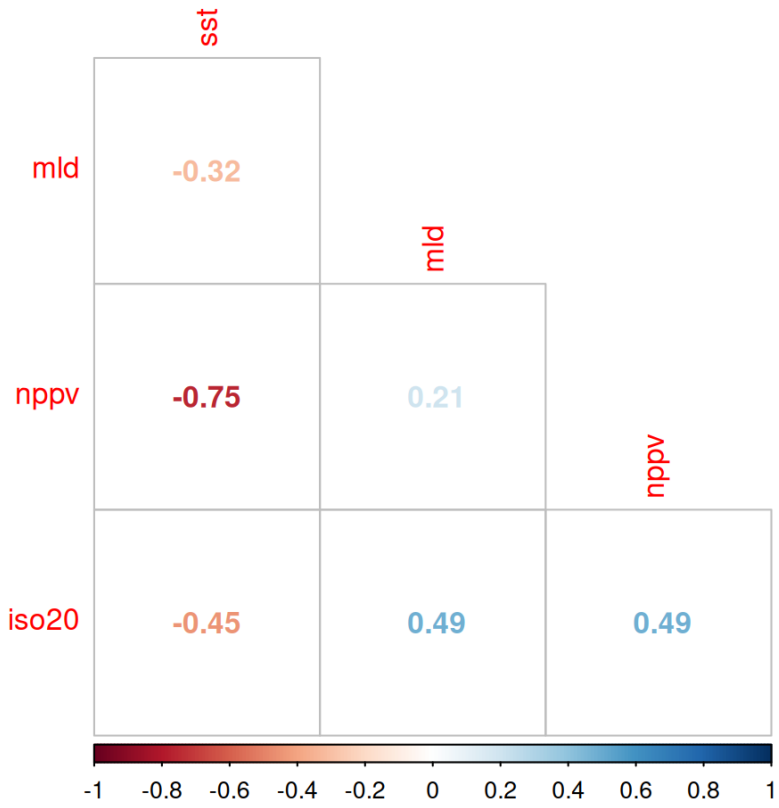


Figure 23: Variance inflation factors (VIF) by covariate obtained when fitting a simple linear model with long time series data (1991-2025) with (top) and without (bottom) variables related to NPPV and the depth of the 20°C isotherm (ISO20).


sst
mld
nppv
mld
-0.32
nppv
-0.75
0.21
iso20
-0.45
0.49
0.49
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

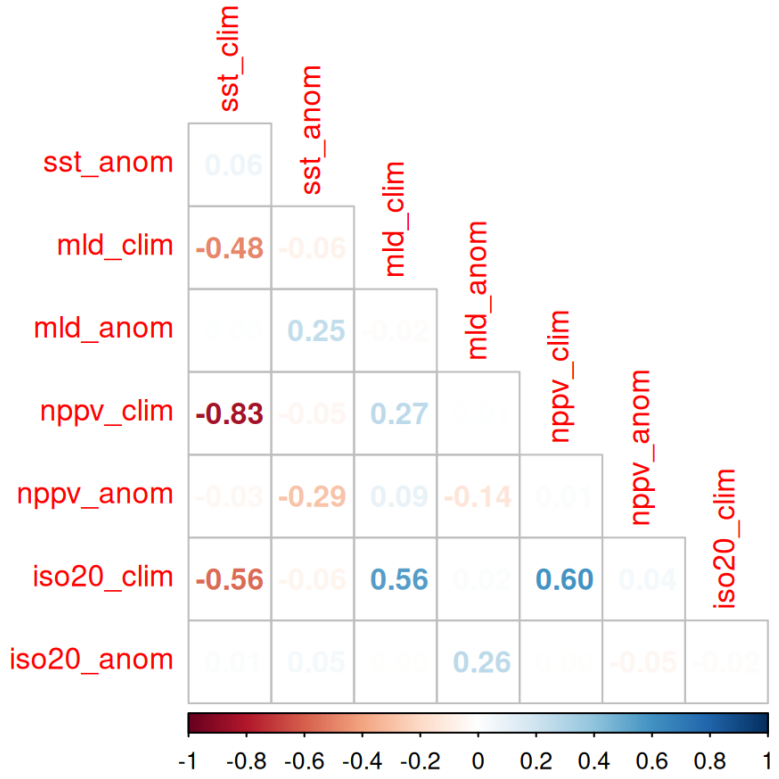

sst_clim
sst_anom
mld_clim
mld_anom
nppv_clim
nppv_anom
iso20_clim
sst_anom
mld_clim
-0.48
mld_anom
0.25
nppv_clim
-0.83
0.27
nppv_anom
-0.29
0.09
-0.14
iso20_clim
-0.56
0.56
0.60
0.04
iso20_anom
0.26
-0.05
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

Figure 24: Correlation matrix among environmental covariates for the long time series (1991-2025). 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.

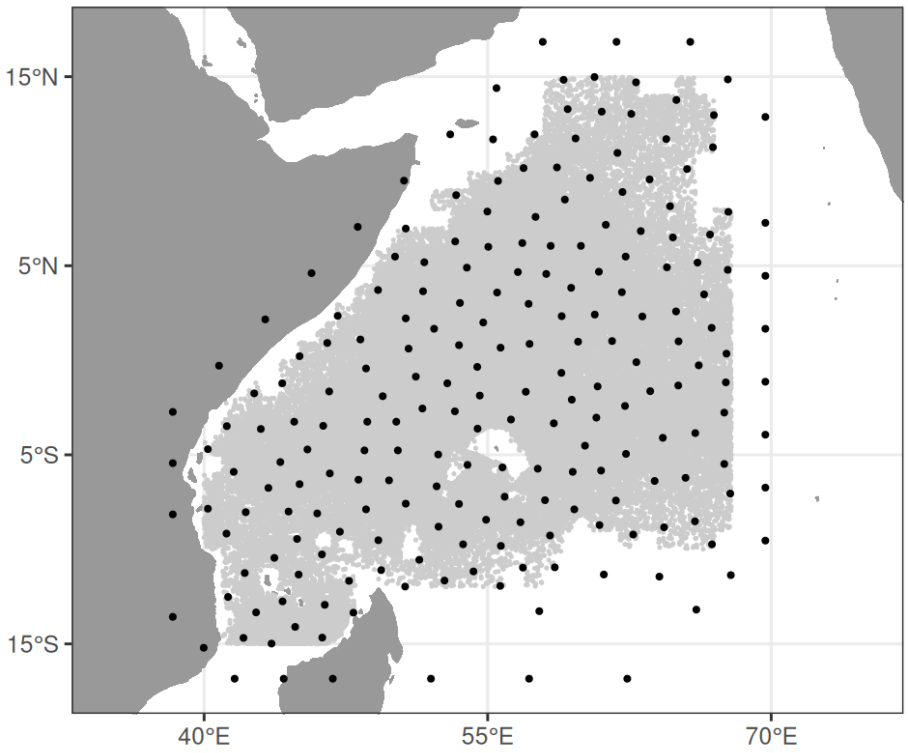


Figure 25: Mesh nodes (black dots) used in the spatiotemporal model and observations (gray dots) for the long time series (1991-2025).

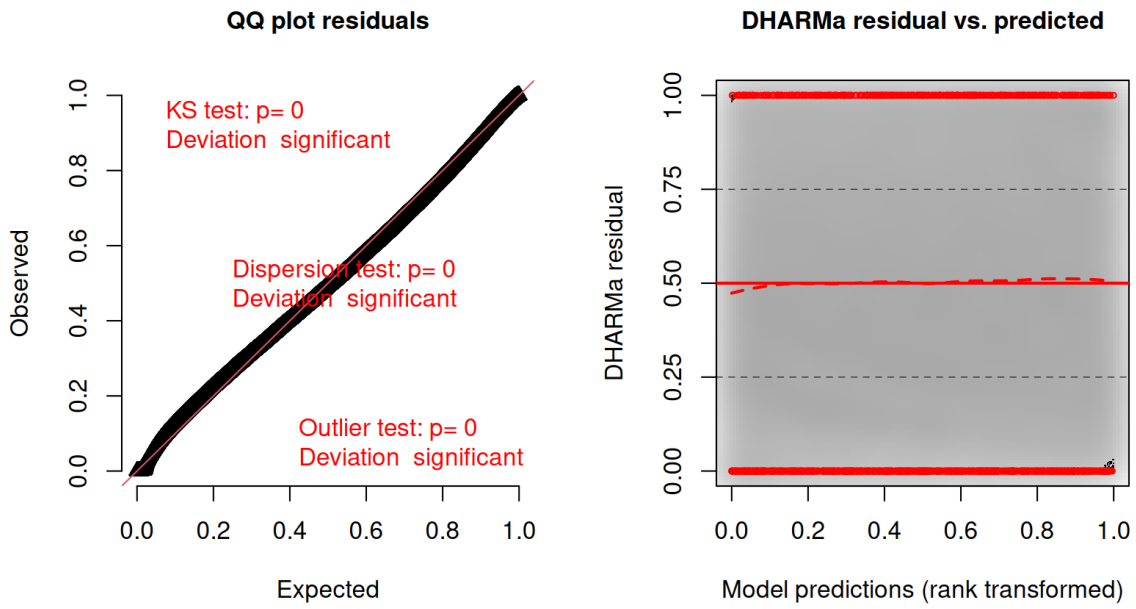


Figure 26: Simulation-based randomized-quantile residuals for the long time series model (1991-2025). QQ-plot (left) detects overall deviations from the expected distribution, by default with added tests for correct distribution (KS test), dispersion and outliers. Residual plot (right) shows the residuals against the predicted value.

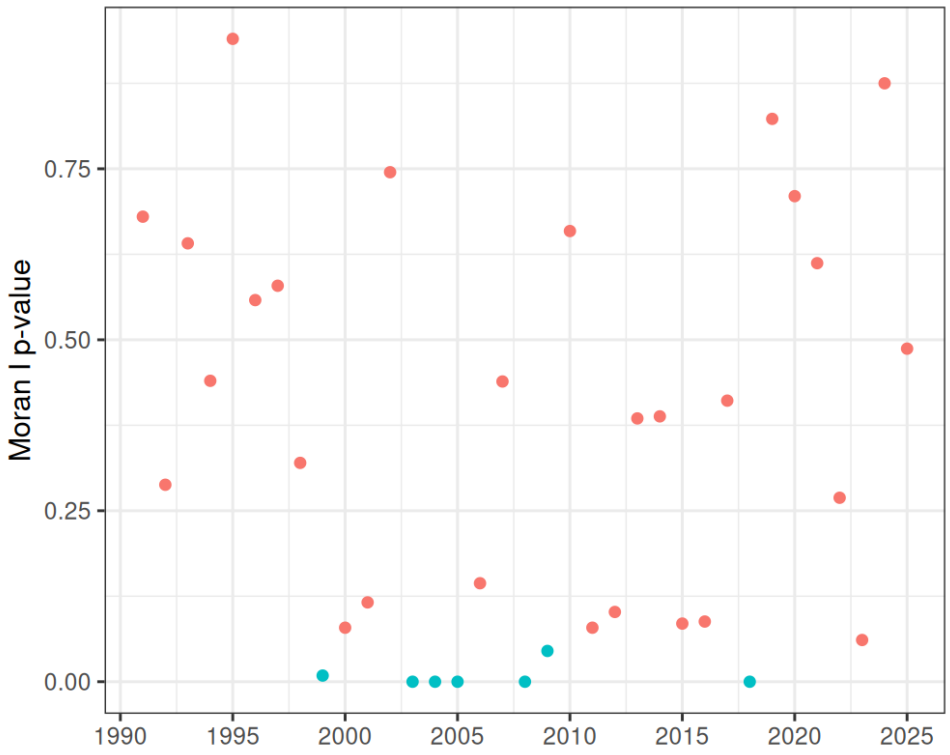


Figure 27: Moran I's p-value of the randomized quantile residuals for standardization model of long time series (1991-2025). Blue points represent years with significant spatial autocorrelation in residuals.

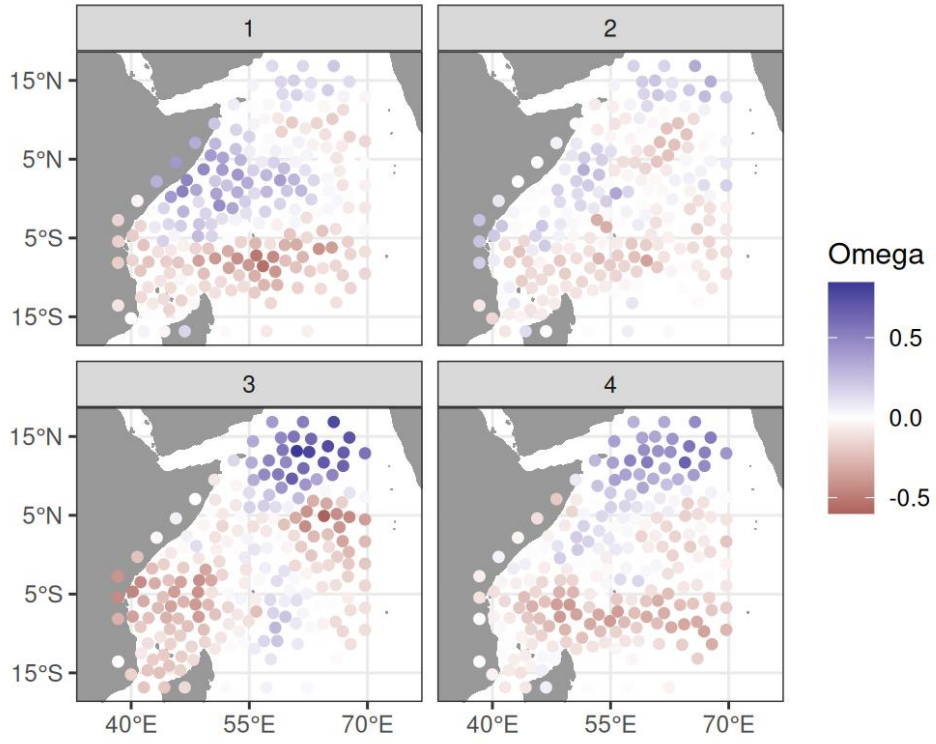


Figure 28: Quarter-specific spatial term for long time series model (1991-2025).

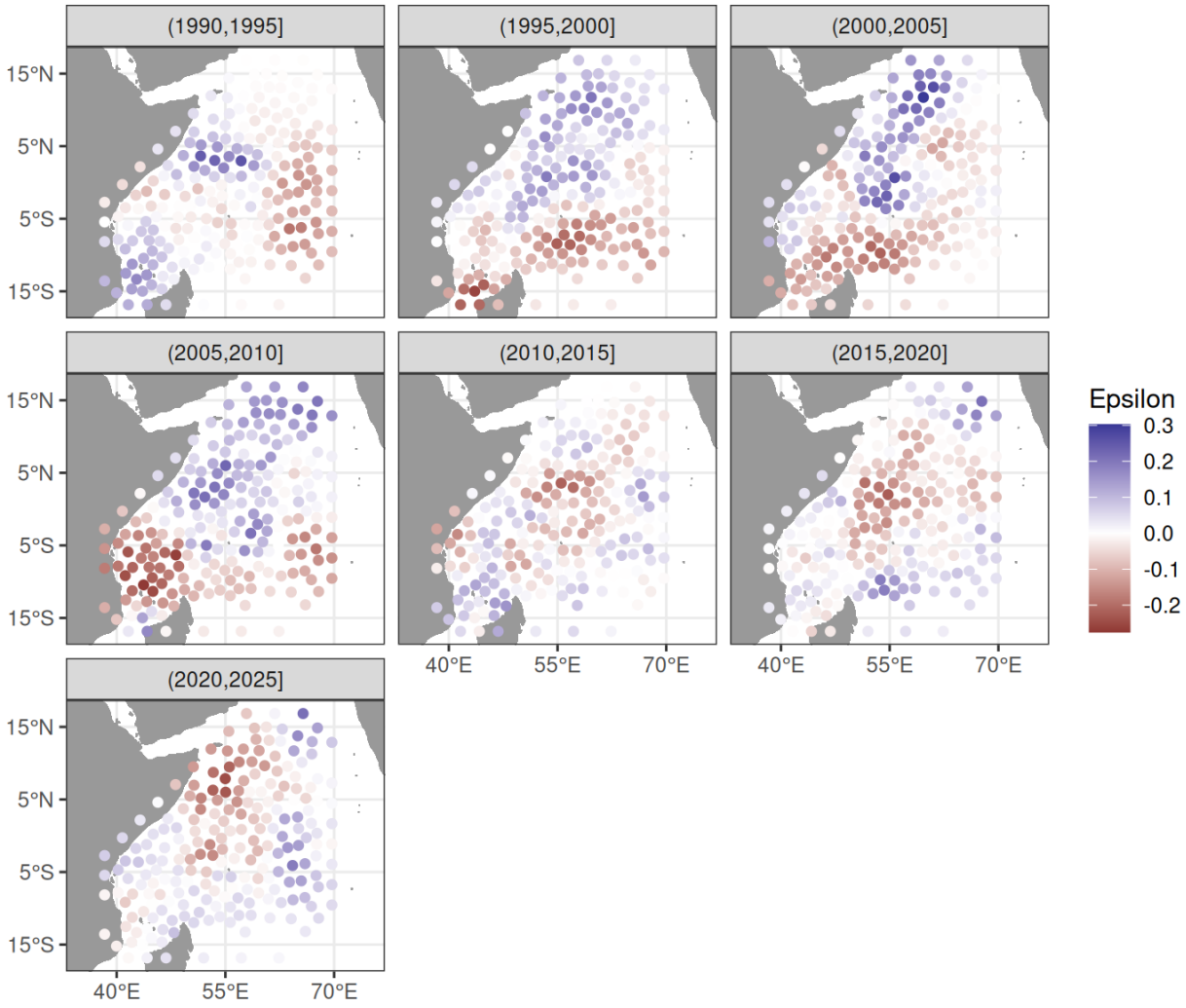


Figure 29: Spatiotemporal term for long time series model (1991-2025). Values have been meaned over each 5-year time period.

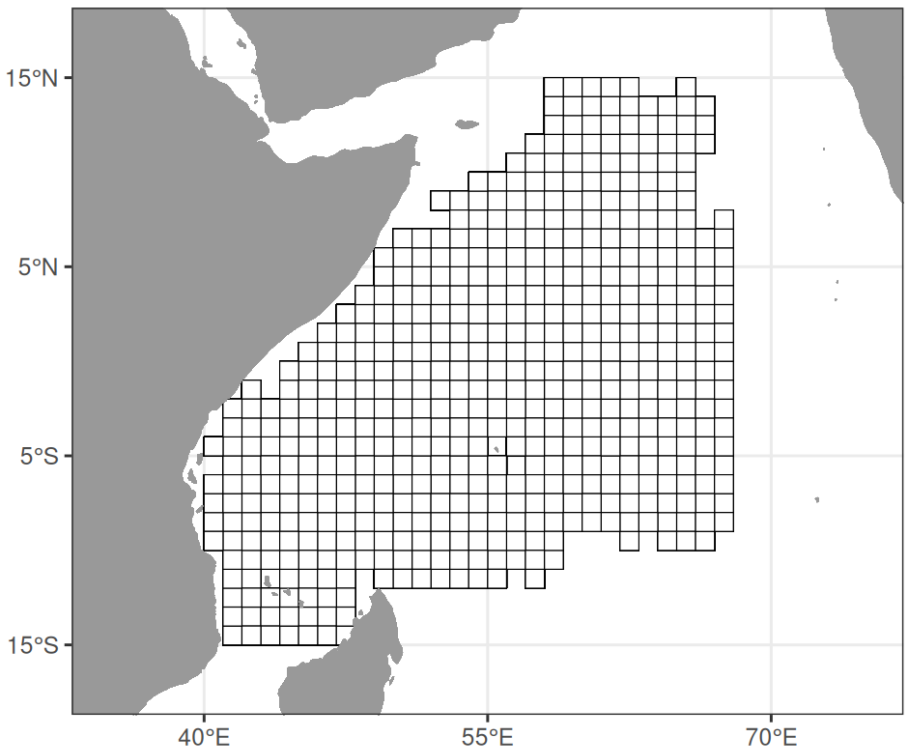


Figure 30: Extrapolation area composed by 1° × 1° grids. This area was used when predicting CPUE values over space and time for the long times series (1991-2025).

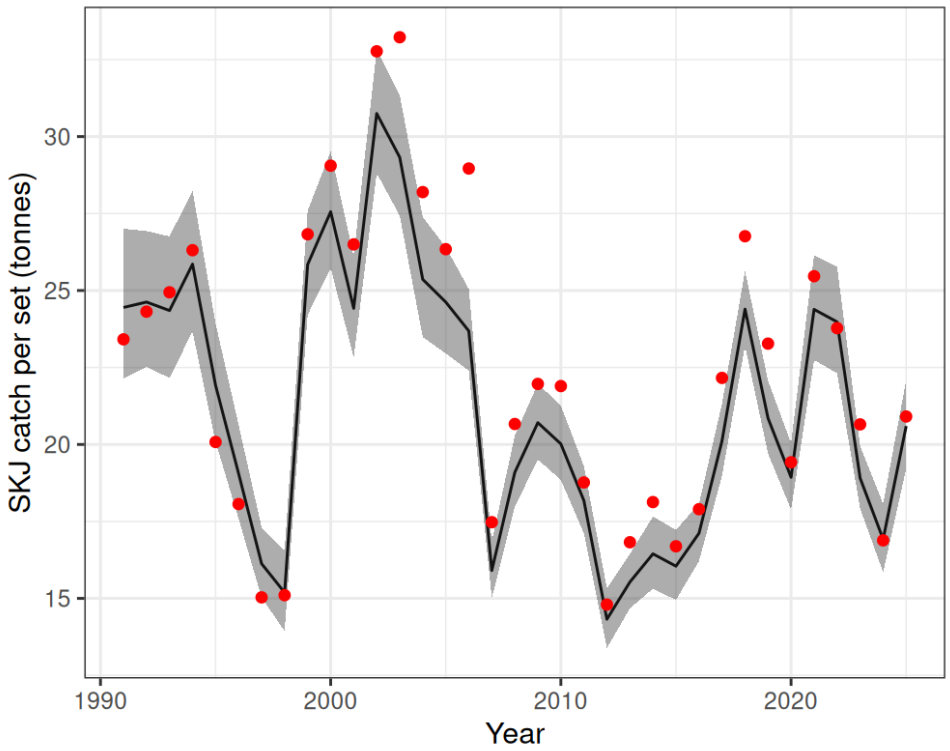


Figure 31: Annual standardized CPUE index for the long time series (1991-2025). Gray area represents the 95% confidence interval. Red dots represent the nominal CPUE.

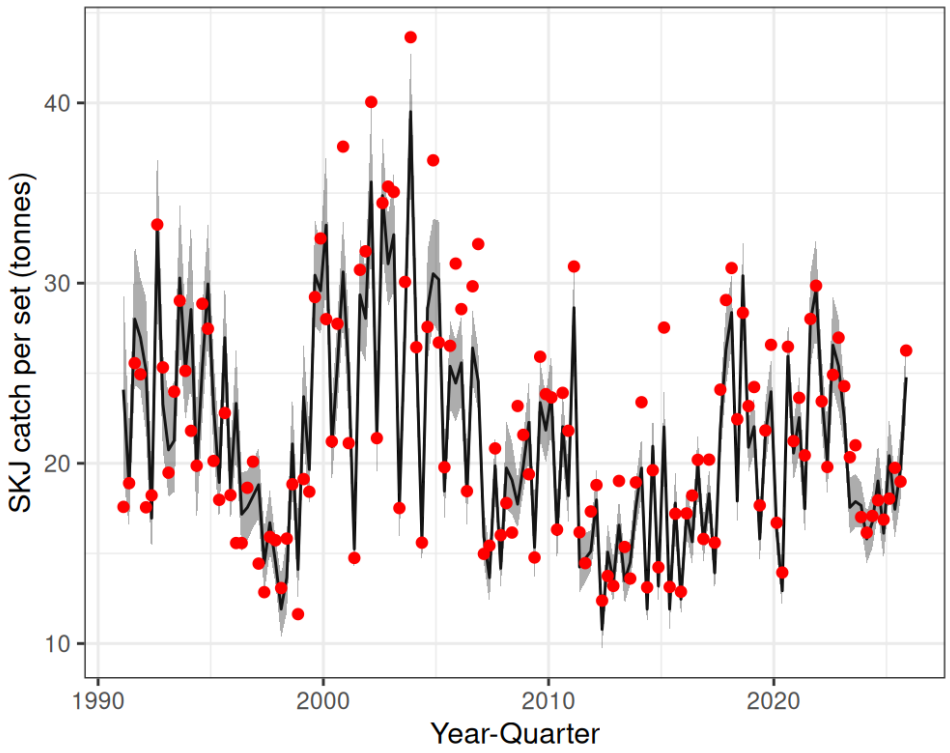


Figure 32: Quarterly standardized CPUE index for the long time series (1991-2025). Gray area represents the 95% confidence interval. Red dots represent the nominal CPUE.

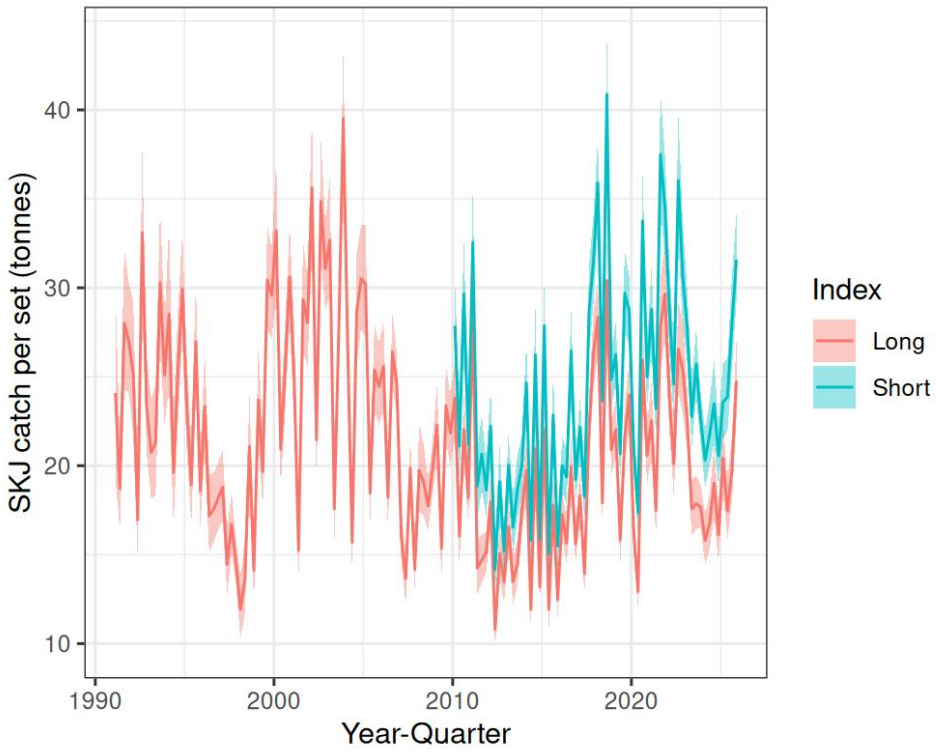


Figure 33: Quarterly standardized CPUE index for the long and short time series superposed.

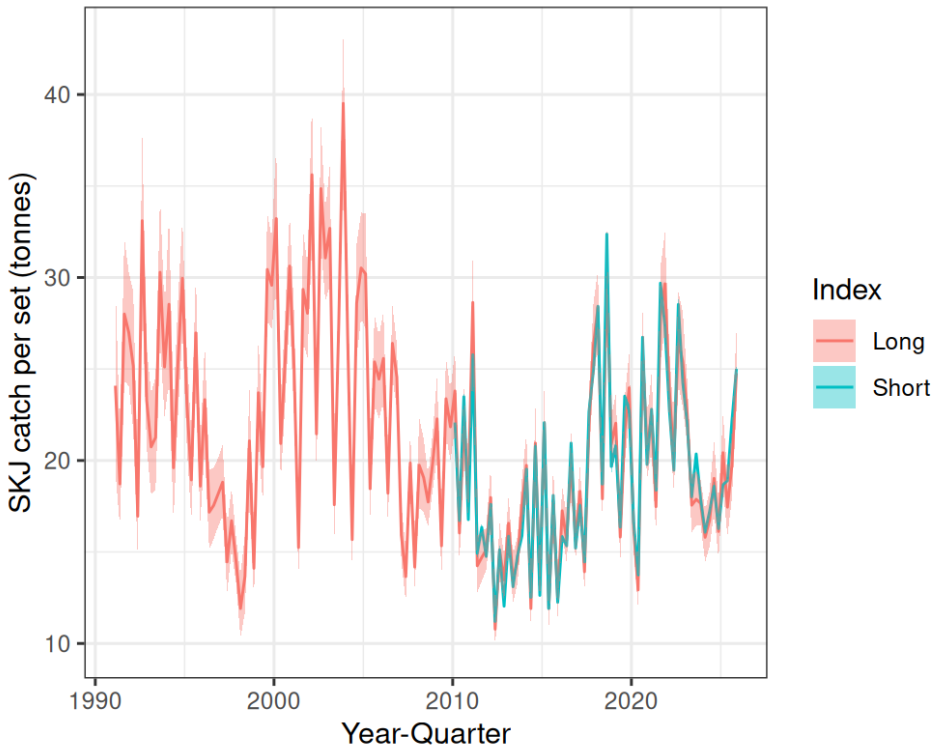


Figure 34: Quarterly standardized CPUE index for the long and short time series superposed after having scaled the short time series to have the same mean as the long time series over the period of overlap.

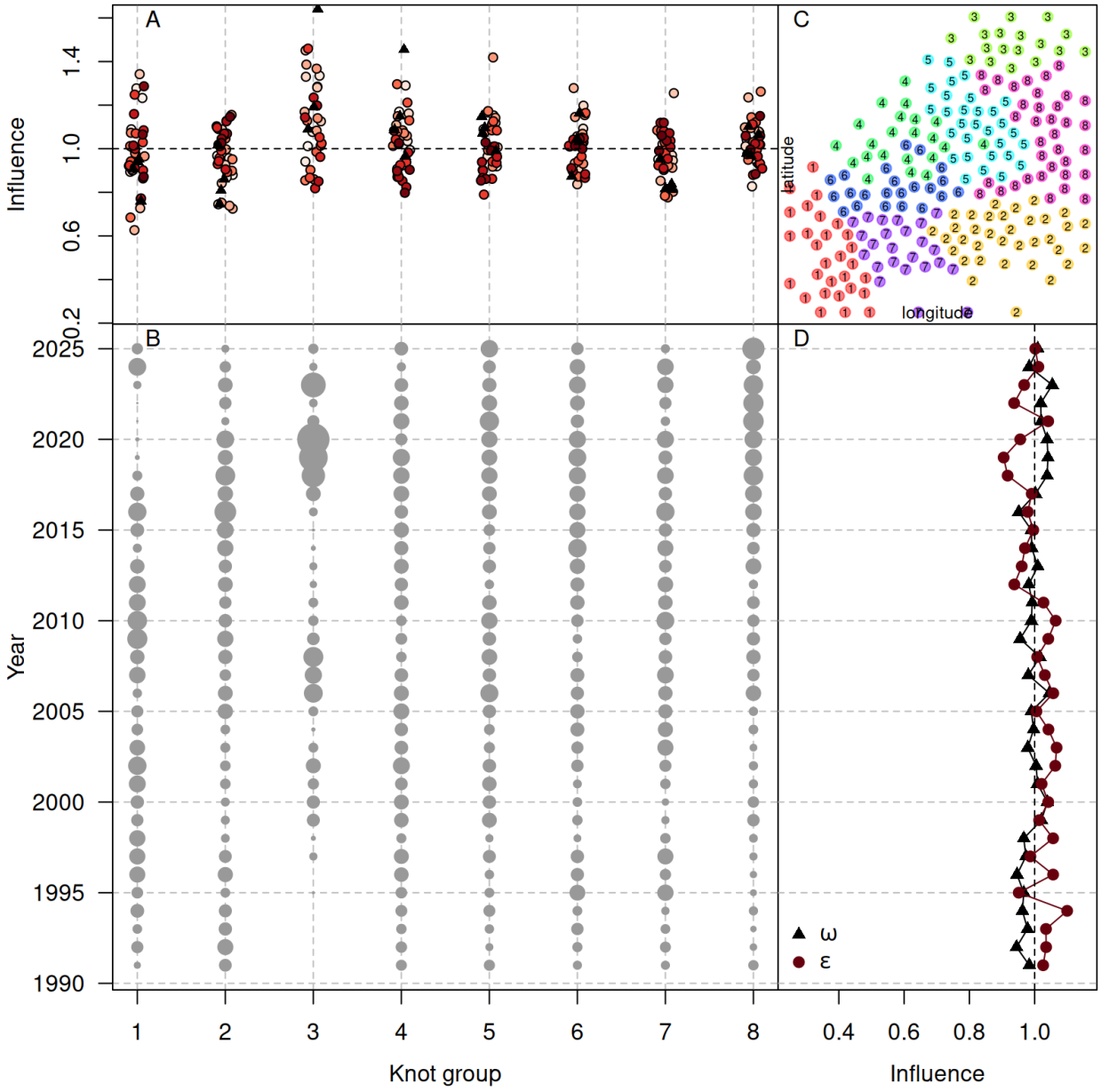


Figure 35: Coefficient-distribution-influence plot for the long times series (1991-2025). Figure A shows the coefficient by knot group for the spatial (black triangles, each triangle represents a quarter) and spatiotemporal (red dots, each dot represents a year) term. The knot groups are shown in Figure 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.

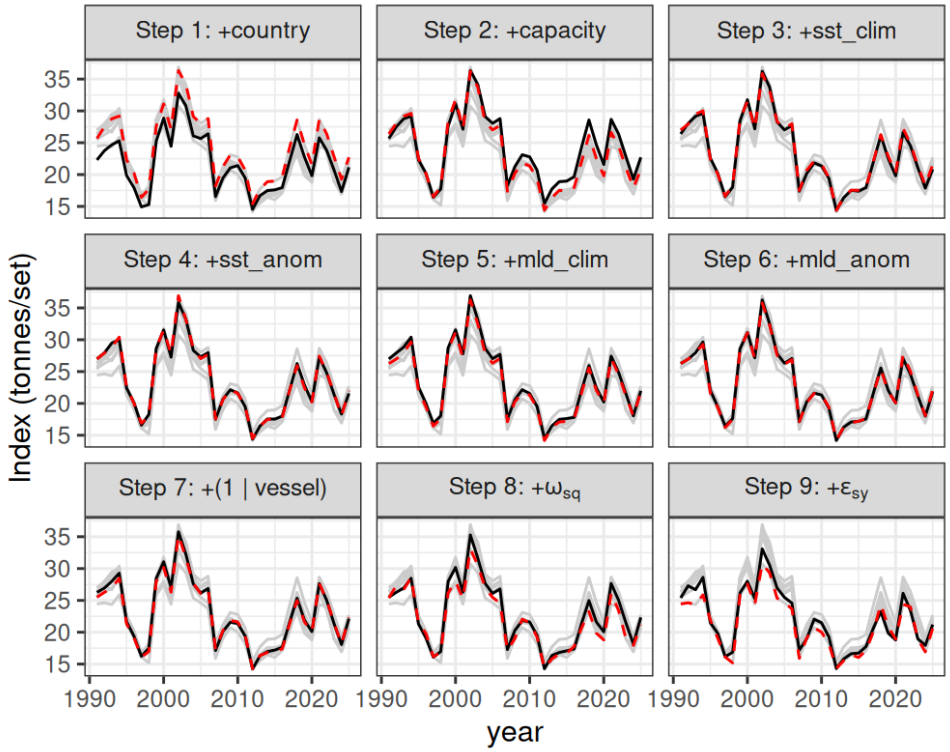


Figure 36: Step plot to evaluate the effect of adding a new covariate on the annual standardized CPUE index for the long time series (1991-2025). The numbers indicate the steps, and variable name indicates the covariate that was added at that step. The black line represents the CPUE index at that step, 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.

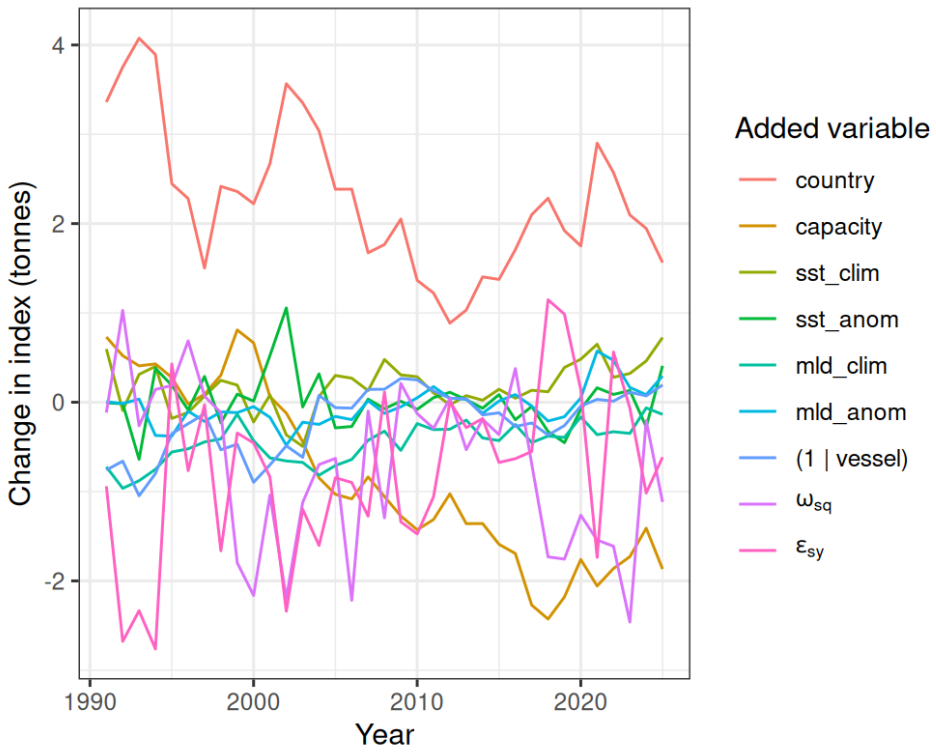


Figure 37: Change in annual standardized CPUE index for the long time series (1991-2025) due to adding a new covariate to the model formula. Color indicates which covariate was added to the formula and the covariates are presented in the legend in the order they were added to the model. Note that year-quarter, vessel random effect (`(1|vessel)`), spatial (ω_{sq}) and spatio-temporal (ϵ_{st}) terms were included in all models.

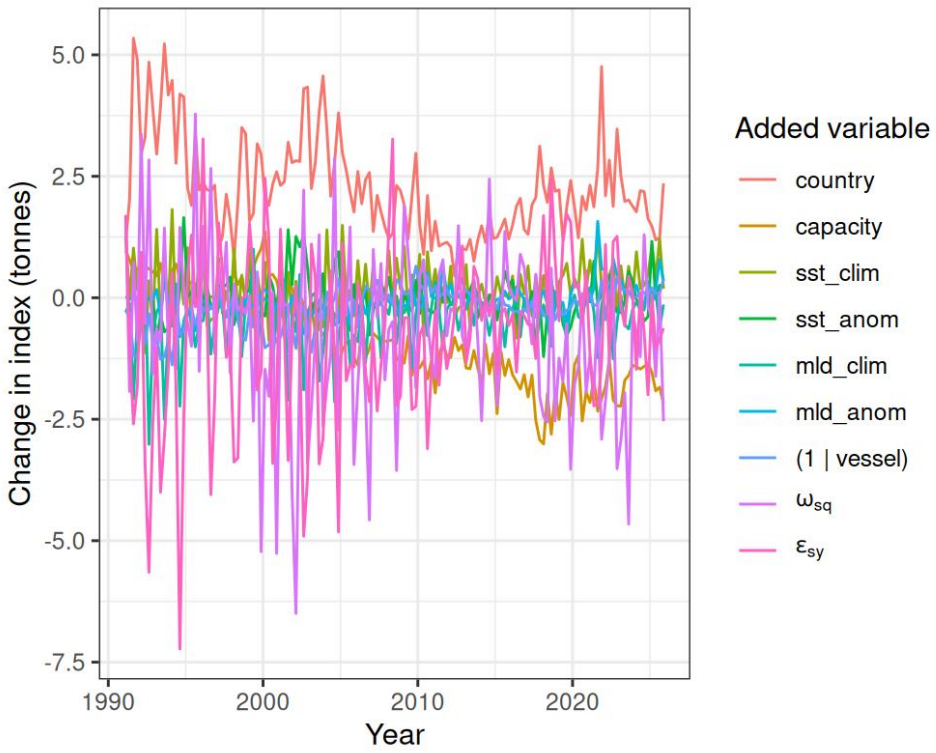


Figure 38: Change in quarterly standardized CPUE index for the long time series (1991-2025) due to adding a new covariate to the model formula. Color indicates which covariate was added to the formula and the covariates are presented in the legend in the order they were added to the model. Note that year-quarter, vessel random effect (`(1|vessel)`), spatial (ω_{sq}) and spatio-temporal (ϵ_{st}) terms were included in all models.

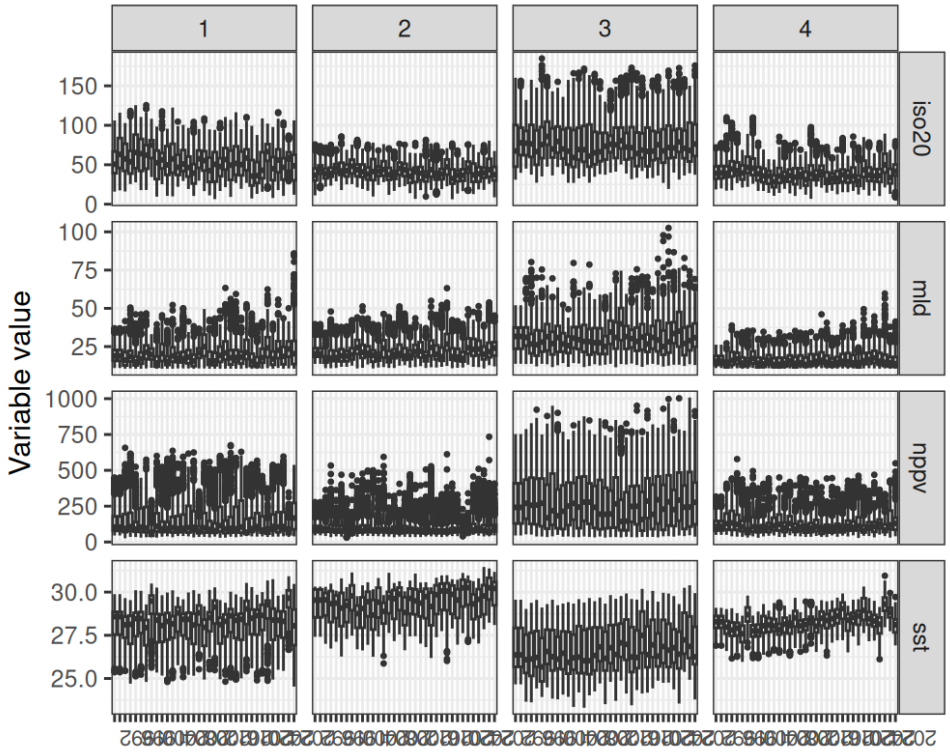


Figure 39: Temporal trends in oceanographic conditions. Boxplots are composed of environmental information per grid of the extrapolation area for the long time series (1991-2025).