

Standardised CPUE indices of abundance for pelagic sharks, mako shark (*Isurus oxyrinchus*) and blue shark (*Prionace glauca*), off South Africa

Dawit Yemane¹, Charlene da Silva¹, and Sven Kerwath^{1,2}

2025-02-21 22:13:41

1 - Department of Forestry, Fisheries & the Environment (DFFE)

2 - Department of Biological Sciences, University of Cape Town

Abstract

This report provides preliminary results for standardized Catch per Unit Effort indices based on catches of the large pelagic longline fishery to track abundance of two pelagic shark stocks off South Africa: blue shark (*Prionace glauca*) and mako sharks (*Isurus oxyrinchus*). Given the spatio-temporal nature of the data, the standardized index of abundance was generated based on a model that takes advantage of this information to learn about the long-term trend in the abundance of modelled stock, accounting both for catchability and abundance covariates. Data from both indicator vessels (former shark longline vessels that continue to catch a significant proportion of sharks) and from the entire large pelagic longline fleet were considered. This fleet targets multiple tuna species, thus, to account for changes in targeting, a multivariate index of species composition of the catch was included in the model. A spatio-temporal Generalized Linear Mixed Effect Model (GLMM) was applied, accounting both for catchability and abundance covariates. Multiple models were fitted of which the best model was selected based on information theoretic approach using the AIC. The standardized indices of abundance for both mako and blue sharks were then calculated from the best model.

1. Background

Indices of abundance derived from standardized fisheries Catch per Unit Effort (CPUE) data are an important input in the assessment of the two pelagic sharks, blue and mako sharks, caught by the South African pelagic longline fishery operating in the south Atlantic- and Indian ocean, across the boundary of ICCAT and IOTC. The previous approaches used to generate indices of abundance for blue and mako sharks considered models that range in complexity on how they handled 1) the spatial component; 2) targeting; 3) temporal components. For example (Fernández-Costa *et al.*, 2023) and (Coelho *et al.*, 2023) modelled the potential effects of targeting as what they termed as *ratio* (ratio of sword fish catch to the sum of sword and blue shark catch); they incorporated spatial effects at much broader scales (Areas 6 to 11) both as main and interaction effects. For temporal effects year and quarter effects were included as a factor. In the ICCAT report by Kai (2023) a more complex spatio-temporal model was implemented in the R package VAST (Thorson, 2019). Specifically, they applied delta-negative binomial models with the probability of occurrence modelled as binomial model (although fixed at a constant due to high percentage

of positive catches > 91%), and a positive component modelled as negative binomial. Spatial and spatio-temporal components were modelled as random fields. In addition the model also included vessel as a random effect and numbers of hooks as measure of fishing effort. In this report we applied a similar approach to that in Kai (2023), but instead of using the implementation of the spatio-temporal *GLMM* model in *VAST* we used a similar (and recommended Anderson *et al.* (2024)) implementation in the R package *sdmTMB* (Anderson *et al.*, 2024). In this report, preliminary results from the spatio-temporal Generalized Linear Mixed Effect Model is presented.

2. Methods

A map of the study region, the southern Atlantic and Indian ocean of South Africa, is shown in Figure 1.

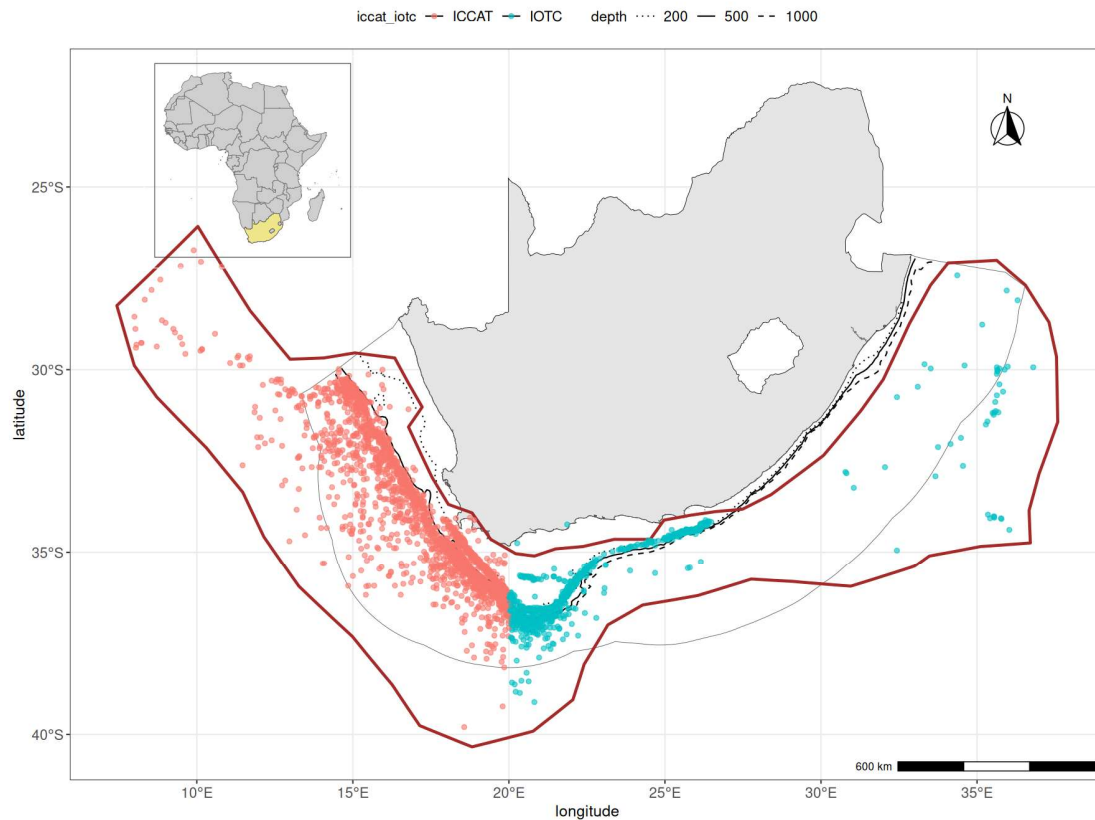


Figure 1: Map of the South Africa showing the typical fishing region in the south Atlantic and Indian ocean. Points represent fishing locations from indicator vessels. Brown lines represent the boundary of the prediction grid for model based on data from indicator vessels.

Before the application of any data filtration/exclusion the total numbers of records for the indicator vessels and that for entire pelagic long-line was 3380 and 45259 respectively.

Prior to the modelling fitting process, multiple data filtration/exclusions were undertaken so as to retain only reliable sets of catch records:

- 1st step filtering:
 - exclude records with missing coordinates
 - exclude records with negative/zero effort (hooks used)
 - exclude records with missing catch for species for either mako or blue sharks
 - exclude records from years earlier than 2000 (due to paucity records)
 - exclude fishing sets/outing that are far outside the common fishing ground in the south Atlantic and Indian oceans (east of 5° *East* and north of –45° *South*) resulting in loss of about (~ 0.7% the records when using all the data).

Addition data filtration were also undertaken:

- 2nd step filtering:
 - exclude sets that fall outside the predictions grid (generated for main fishing grounds)
 - exclude records with zero total catch (sum of catch of all tuna and/or sharks)

After the application of the above-described data filtration/exclusion steps the numbers of records remaining in the indicator vessel and entire pelagic longline fishery were 3378 and 42739 respectively.

2.1 Fleet dynamics of South African Pelagic longline fishery

Domestic commercial longlining for tuna has been documented from the early 1960s, when the fishery landed approximately 2 000 t of tuna. The tuna longline fishery declined rapidly in the mid-1960s due to an insufficient market for low quality bluefin and albacore tuna which constituted the bulk of the catch. However, in 1995 an interest in this fishery was expressed when profits were made by a joint venture with a Japanese longlining vessel that targeted tuna and swordfish within South African waters. Consequently, 31 experimental longline permits were issued in 1997 and the first commercial long-term rights were allocated in 2005, expiring and reverting back to the State on 31 December 2015. The allocation policy stipulated that the TAE be set at 20 swordfish-directed vessels and 30 tuna-directed vessels to avoid depleting local swordfish resources. In the allocation process 18 Rights were issued for the swordfish-directed fishery and 26 for the tuna-directed fishery (1 Right = 1 vessel). A Precautionary Upper Catch Limit (PUCL) of 2 000 t dressed weight of all sharks was set. If this limit was reached within a given season, the large-pelagic fishery would be closed. The PUCL was calculated as the highest fishing level which could be considered under the guise of meeting the 2005 policy objective (50 rights holders targeting tuna and swordfish with no 10% bycatch limit in place).

The pelagic shark longline sector, which operated under exemption from 2005, was absorbed into the large pelagic (tuna and swordfish) longline fishery in 2011. This increased the total number of Large Pelagic Longline Rights to 50, with a split of 21 swordfish-directed and 29 tuna-directed Rights. Joint venture (foreign-flagged vessels) was prohibited from landing more than 10% sharks of total dressed weight of tuna species per season.

In 2013, pelagic sharks were designated as bycatch in the policy for this fishery. In 2015 a decision was made to manage the fishery collectively as the Large Pelagic Longline fishery, with no distinction between tuna-directed and swordfish-directed vessels or permits. Sixty-three commercial fishing Rights were allocated in January 2017 , expiring and reverting back to the State on 31 January 2032. .

When the PUCL was first proposed within DFFE, the intention was to reduce the PUCL to 370 t over a 5-year period. This amount was calculated based on the maximum extrapolated catch of this fishery operating under the 10% by-catch limit. The PUCL remained at 2 000 t and was never decreased. In 2013, wire traces were prohibited for the original tuna and swordfish fleet. As of 2015, the use of wire traces has been banned in the South African Large Pelagic Longline sector, as is the use of stainless-steel hooks. Furthermore, shark fins must be naturally attached to the body or tethered when landed. As of 2015, the targeting of sharks (defined as 50% or more sharks per fishing season by mass) was prohibited and if quarterly landings exceeded 60% shark, the vessel was required to have 100% observer coverage for the next quarter. As of 2013, South Africa banned the retention of specific shark species: thresher sharks (2017) (genus *Alopias*), hammerhead sharks (genus *Sphyrna*) (2017), oceanic whitetip sharks (2014) (*Carcharhinus longimanus*), porbeagle sharks (*Lamna nasus*) (2014), silky sharks (*Carcharhinus falciformis*) (2014), and dusky sharks (*Carcharhinus obscurus*) (2014), often misidentified as silky sharks.

Results from initial exploration of blue and mako sharks catch over the entire time series, to summarise the overall spatial pattern in the catch location, are shown in Figure 2 before processing the data, and the corresponding pattern after processing the data are shown in Figure 3.

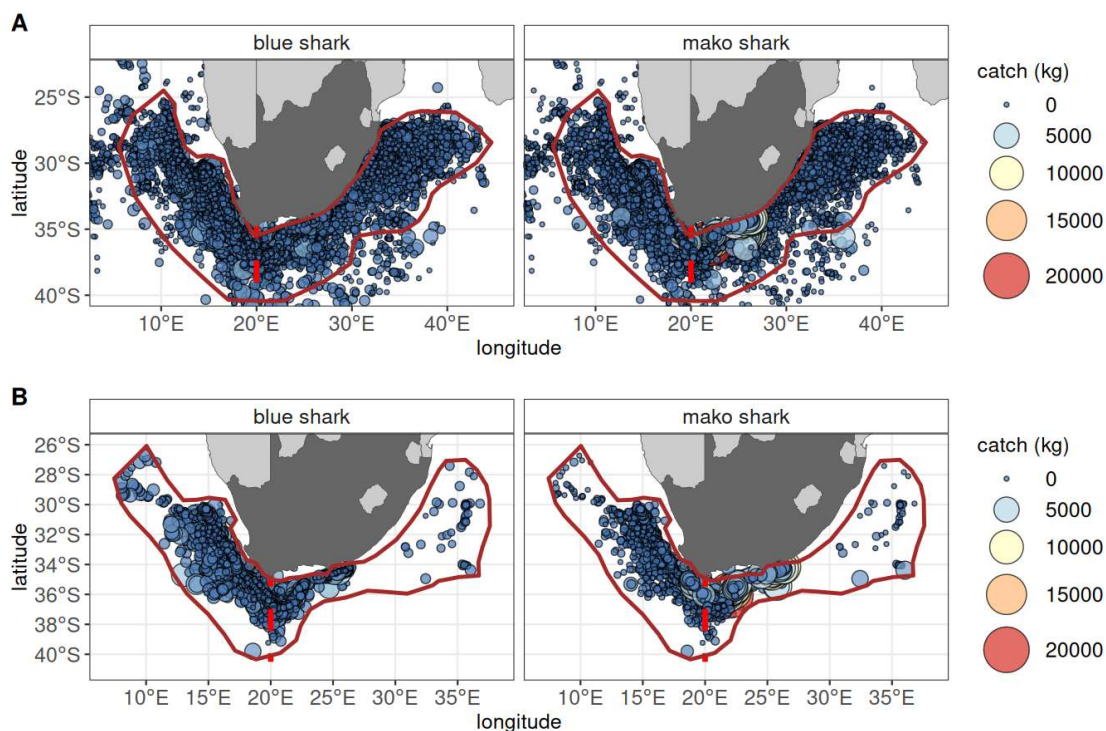
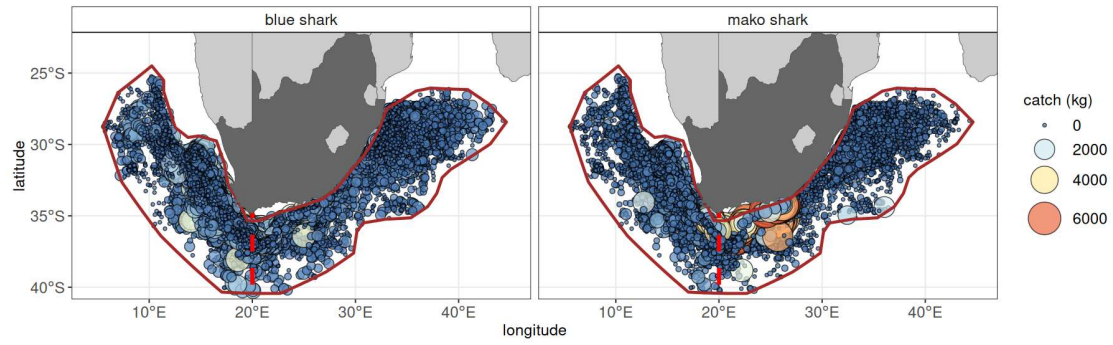


Figure 2: Distribution of raw catches of blue and mako sharks caught in the southern Indian and Atlantic oceans. A) data from all vessels B) data from indicator vessels. Brown lines denote the boundary of the prediction grid. Red dashed lines denote ICCA-IOTC boundary off South Africa.

A



B

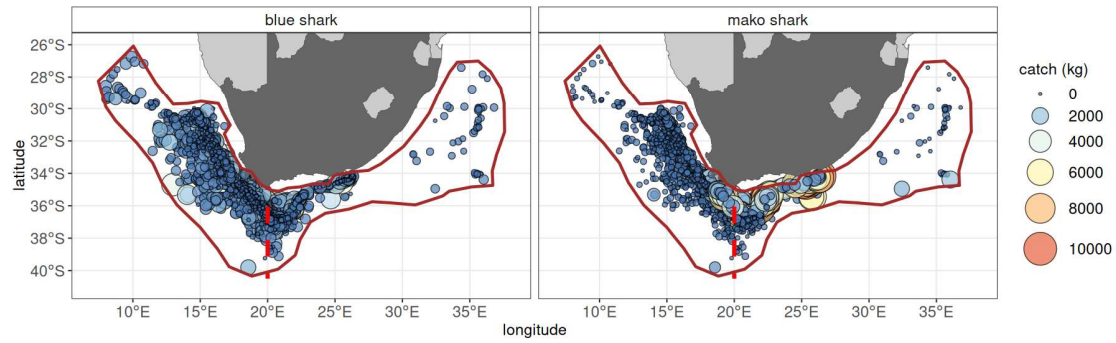


Figure 3: Distribution of processed catches of blue and mako sharks caught in the southern Indian and Atlantic oceans. A) data from all vessels B) data from indicator vessels. Brown lines denote the boundary of the prediction grid. Red dashed lines denote ICCAT-IOTC boundary off South Africa.

The prediction grid used to generate the indices of abundance for both blue and mako sharks when using data from all vessels and indicator vessels is shown in Figure 4.

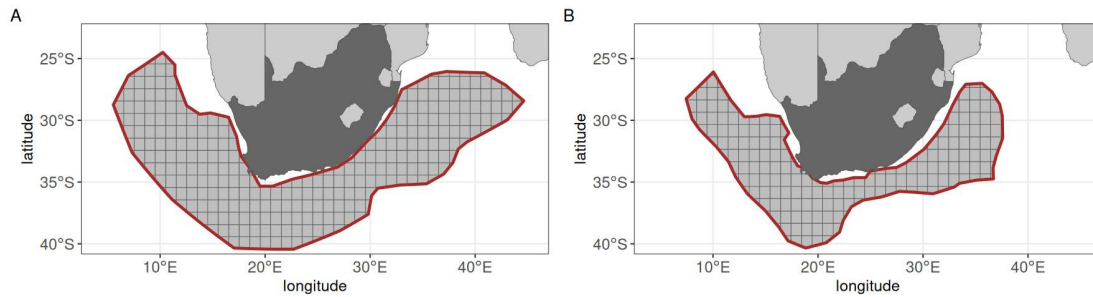


Figure 4: Plot of the prediction grid used to generate index of abundance for blue and mako sharks. A) when using all the data, B) when using data from indicator vessels.

2.2 Modelling the effect of targeting

There are multiple approaches used when developing standardizing cpue from fisheries dependent data. As noted above some of the previous work on blue shark attempted to address the issue of targeting by incorporating what is termed as ratio (ratio of catch of blue sharks to total catch), this raises some issues including potential confounding from having the response on both side of the equation. Previous works on the multispecies linefishery off South Africa has shown the utility of multi-species measure of species composition, the axis scores from one or multiple principal components in accounting for switch in targeting (Winker *et al.*, 2014). In the context of this study a similar approach was adopted. Result from the PCA analysis is shown Figure 5.

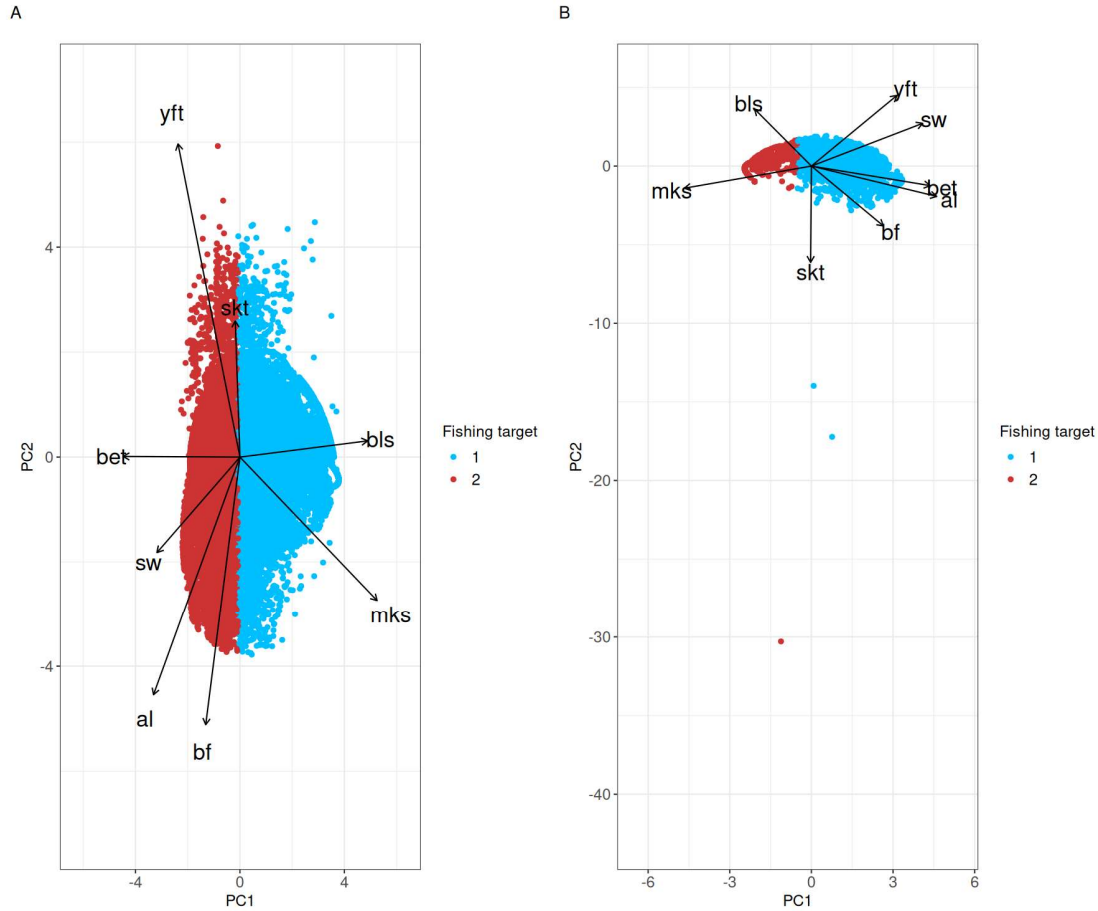


Figure 5: Bi-plot from PCA on double-square root transformed species composition data. The length and direction of vector plot shows the strength and the direction of association of the species with the first and second principal components. The colors denote the membership of axis score to the optimal clusters identified. A) when using all the data, B) when using data from indicator vessels.

2.3 CPUE standardization

Given the spatio-temporal nature of the data, any potential method to be used to generate a standardized index of abundance needs to account for the effect of location of the catches and how these catch locations vary over time. In addition, the relative influence of additional covariates, e.g month, fishing depth, and other relevant variables needs to be included in the model. Although there are multiple modelling frameworks that can be used, for the purpose of this report a spatio-temporal GLMM implemented in the R package *sdmTMB* (Anderson *et al.*, 2024) was used. Although **sdmTMB**, as the name suggests species distribution model implemented in TMB, was initially intended to be used for modelling species distribution, it has since been extended to be used in a range context including generation of indices of abundance both from fisheries dependent and independent sources (Anderson *et al.*, 2024). To model the spatial process the coordinates (longitude and latitude) were projected to UTM zone 34, the zone in which most observations fall (spatial processes are modeled with respect to distance). The standard GLMM with covariate effects and spatial and spatio-temporal component takes the form:

$$\begin{aligned}\mathbb{E}[y_{s,t}] &= \mu_{s,t} \\ \mu_{s,t} &= f^{-1}(\mathbf{X}_{s,t}\boldsymbol{\beta} + \omega_s + \epsilon_{s,t})\end{aligned}$$

where $\mathbb{E}[y_{s,t}]$ is the expected value of the observation, in this case catch at location s and time t ; $\mathbf{X}_{s,t}$ is the design matrix for the main effect (e.g. covariates, fixed effect of year); $\boldsymbol{\beta}$ is the vector of coefficients for the main effects; $\mu_{s,t}$ the mean of the observation, here catch, at location s , and t ; f^{-1} is the link function linking the mean response to the predictors (it allows to model response in multiple space: *logit*, *log*, *inverse*, and *identity*).

$$\omega_s \sim \text{MVN}(\mathbf{0}, \boldsymbol{\Sigma}_\omega)$$

$$\epsilon_{s,t} \sim \text{MVN}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$$

where ω_s is the spatial random effect, and $\boldsymbol{\Sigma}_\omega$ is the covariance of the spatial random field; $\epsilon_{s,t}$ is the spatio-temporal random field and $\boldsymbol{\Sigma}_{\epsilon_{s,t}}$ is the covariance of the spatio-temporal random field.

The six models considered for each species and the two data sets are shown below. For each of the models two variants were considered with and without anisotropy. In addition two distribution family were considered (*tweedie* and *delta-Gamma*). Thus in total 48 models were considered for each species or a total of 96 models.

Model	formula
<i>Model</i> ₁	<i>catch</i> ~ <i>Year</i> + <i>s(depth)</i> + <i>month</i> + <i>s(PC1)</i> + (<i> vessel</i>) + <i>s(slope)</i> + <i>offset(log(Hooks))</i>
<i>Model</i> ₂	<i>catch</i> ~ <i>Year</i> + <i>s(depth)</i> + <i>month</i> + <i>s(PC1)</i> + (<i> vessel</i>) + <i>offset(log(Hooks))</i>
<i>Model</i> ₃	<i>catch</i> ~ <i>Year</i> + <i>month</i> + <i>s(PC1)</i> + (<i> vessel</i>) + <i>offset(log(Hooks))</i>
<i>Model</i> ₄	<i>catch</i> ~ <i>Year</i> + <i>month</i> + <i>s(PC1)</i> + <i>offset(log(Hooks))</i>
<i>Model</i> ₅	<i>catch</i> ~ <i>Year</i> + <i>month</i> + <i>offset(log(Hooks))</i>
<i>Model</i> ₆	<i>catch</i> ~ <i>Year</i> + <i>offset(log(Hooks))</i>

The proportion of zero values over the entire time series and for the data from all vessels was 50% for blue shark ~ 44% for mako shark. The proportion of zeros for data from the indicator vessels was 21% for blue shark and 17% for mako sharks. Although not specifically a major issue for the data from indicator vessels, given that it is not zero dominated, when one is dealing with observations that have substantial amount of zeros, the most commonly followed approaches are either to use hurdle model, where two sub-models are fitted to the data (modelling probability of encounter/occurrence and modelling positive observations) and combine them, or to use Tweedie distribution. In the current release of *sdmTMB*, multiple types of hurdle models are implemented including delta-lognormal and delta-gamma.

Spatio-temporal GLMMs with spatial and spatio-temporal random effects can account for unmeasured variable(s), leading to observations to be correlated in space and space-time. The most common approach to model these spatial and spatio-temporal random effects is to use Gaussian Random Fields (*GRF*), where the spatial random effects describing the spatial pattern are drawn from a multivariate normal distribution constrained by spatial covariance functions (e.g exponential, spherical or in the current implementation in *sdmTMB* from Matern family), but these models are computationally demanding and can be difficult to fit for big data. Multiple workarounds were proposed to deal with computation difficulties such as the stochastic partial differential equation *spde* approximation to *GRF* implemented in the R package

INLA (**R-INLA**) that is also implemented in TMB, making it accessible to sdmTMB. Implementation of the spatio-temporal GLMM requires construction of a mesh containing triangulation and projection matrices that are needed for the *spde* approach.

The meshes used to model the spatial random fields for the data from all vessels and that from indicator vessels are shown in Figure 6 and Figure 7 respectively.

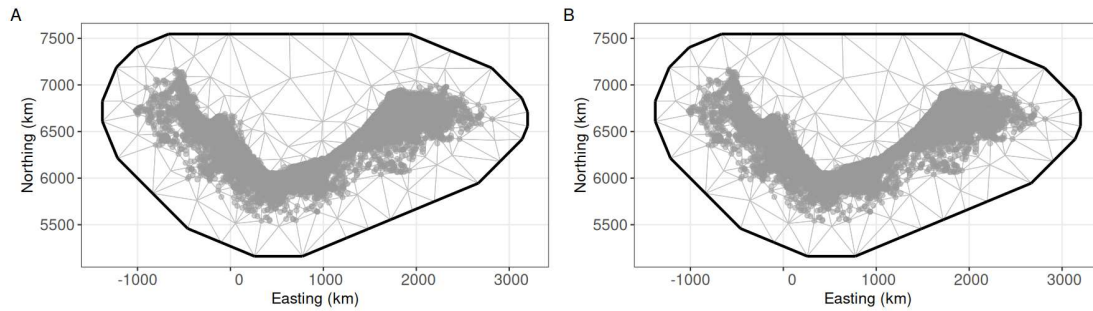


Figure 6: Plot of the mesh used to model the spatial random fields for the model based on data from all the vessels in pelagic longline fishery. A) for mako shark B) for blue shark.

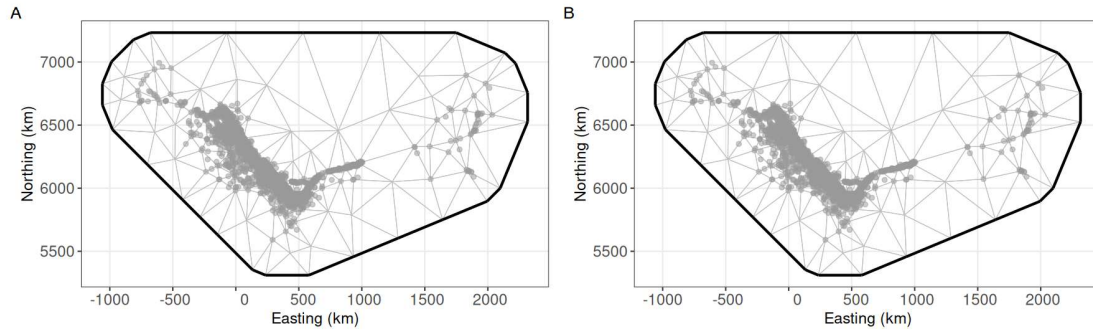


Figure 7: Plot of the mesh used to model the spatial random fields for the model based on indicator vessels in the pelagic longline fishery. A) for mako shark B) for blue shark.

All the analysis, visualisation and report generation were conducted in R (R Core Team, 2024). Multiple R packages were utilised for data processing, visualization, analysis and summary of results including (Allaire *et al.*, 2024; Anderson *et al.*, 2024; Letaw, 2015; Maechler *et al.*, 2024; Pebesma, 2024; Raiche and Magis, 2022; Robinson *et al.*, 2024; Spinu *et al.*, 2024; Wickham *et al.*, 2023, 2024; Wickham and Henry, 2023; Wood, 2023; Xie, 2024).

3. Results

3.1 Model selection

Multiple models were considered for the purpose of this report (as indicated above 48 for each species - although not all combinations converged): The form of the spatial variability (anisotropic vs omnidirectional); distribution family; and the covariates sets were considered. The performance of the models were assessed based on information theoretic approach using AIC. Model performance table are shown in Table 1 and Table 2 for blue and mako sharks respectively.

Table 1: Comparison of model performance, based on AIC, models with different fixed effect structure and distribution family for blue shark. Using all the data, and data from indicator vessels.

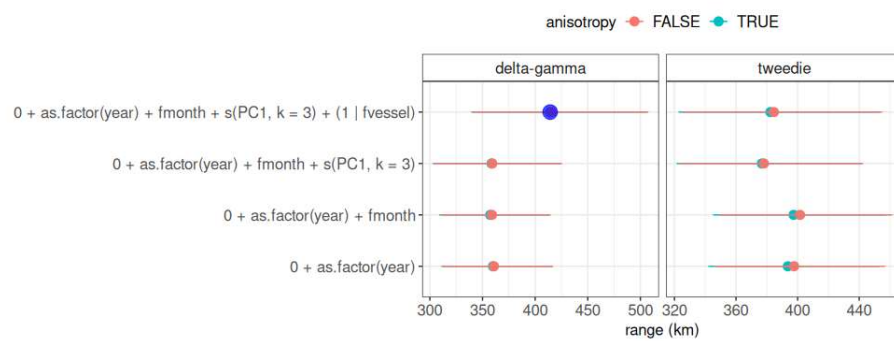
data source	family	formulas	anisotropy	AIC
All vessels	tweedie	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	278,063.31
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	278,064.86
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	285,219.62
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	285,221.76
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	300,269.05
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	300,278.80
		$0 + \text{as.factor}(\text{year})$	TRUE	300,763.84
		$0 + \text{as.factor}(\text{year})$	FALSE	300,773.05
	delta-gamma	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	266,468.04
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	266,474.92
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	275,461.61
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	275,463.33
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	293,302.01
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	293,304.96
		$0 + \text{as.factor}(\text{year})$	TRUE	293,977.24
		$0 + \text{as.factor}(\text{year})$	FALSE	293,979.95
Indicator vessels	tweedie	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	TRUE	38,732.19
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	FALSE	38,728.66
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	38,729.02
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	38,725.41
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	38,895.75
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	38,891.94
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	39,176.35
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	39,174.64
		$0 + \text{as.factor}(\text{year})$	TRUE	39,306.39
		$0 + \text{as.factor}(\text{year})$	FALSE	39,303.14
	delta-gamma	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	TRUE	37,919.51
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	37,913.46
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	37,913.26
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	38,316.37
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	38,312.90
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	38,650.78
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	38,647.29
		$0 + \text{as.factor}(\text{year})$	TRUE	38,825.88
		$0 + \text{as.factor}(\text{year})$	FALSE	38,825.08

Table 2: Comparison of model performance, based on AIC, models with different fixed effect structure and distribution family for mako shark. Using all the data, and data from indicator vessels.

data source	family	formulas	anisotropy	AIC
All vessels	tweedie	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3) + \text{s}(\text{slope}, k = 3)$	TRUE	295,337.21
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3) + \text{s}(\text{slope}, k = 3)$	FALSE	295,349.29
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	TRUE	295,381.91
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	FALSE	295,394.14
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	295,413.86
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	295,426.05
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	304,738.03
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	304,765.31
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	318,165.37
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	318,216.18
		$0 + \text{as.factor}(\text{year})$	TRUE	318,532.98
		$0 + \text{as.factor}(\text{year})$	FALSE	318,586.44
	delta-gamma	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	TRUE	285,996.84
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	FALSE	286,008.57
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	286,022.73
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	286,034.38
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	293,626.27
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	293,629.88
		$0 + \text{as.factor}(\text{year})$	TRUE	312,657.72
		$0 + \text{as.factor}(\text{year})$	FALSE	312,662.39
Indicator vessels	tweedie	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel}) + \text{s}(\text{depth_scaled}, k = 3)$	FALSE	38,966.74
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	39,000.44
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	39,982.50
		$0 + \text{as.factor}(\text{year})$	TRUE	40,176.90
		$0 + \text{as.factor}(\text{year})$	FALSE	40,190.81
	delta-gamma	$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	TRUE	38,231.48
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3) + (1 \text{fessel})$	FALSE	38,232.97
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	TRUE	38,401.69
		$0 + \text{as.factor}(\text{year}) + \text{fmonth} + \text{s}(\text{PC1}, k = 3)$	FALSE	38,404.88
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	TRUE	39,514.02
		$0 + \text{as.factor}(\text{year}) + \text{fmonth}$	FALSE	39,525.39
		$0 + \text{as.factor}(\text{year})$	TRUE	39,639.64
		$0 + \text{as.factor}(\text{year})$	FALSE	39,647.44

Visual summary of estimated range from the different models are shown in Figure 8 and Figure 9 for blue- and mako- sharks respectively. The range, distance after which observations can be considered independent (specifically here it refers to the distance where correlation drops to 10% of estimated correlation ρ), was relatively comparable among the different models, with range from the different models having overlapping estimates of range.

A



B

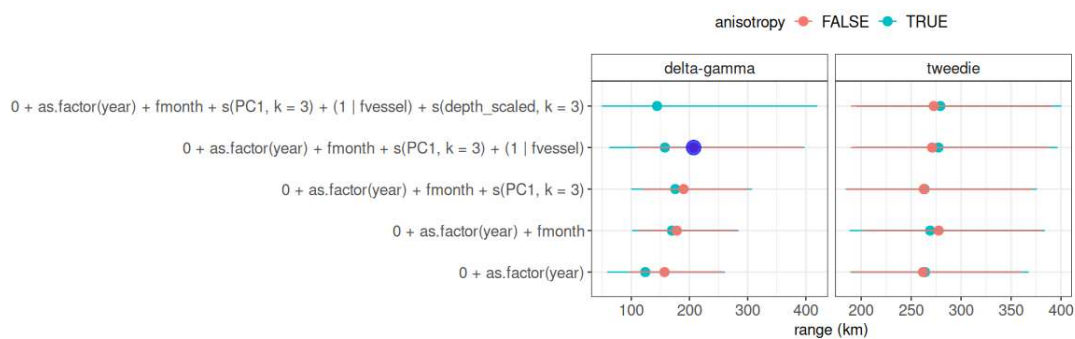
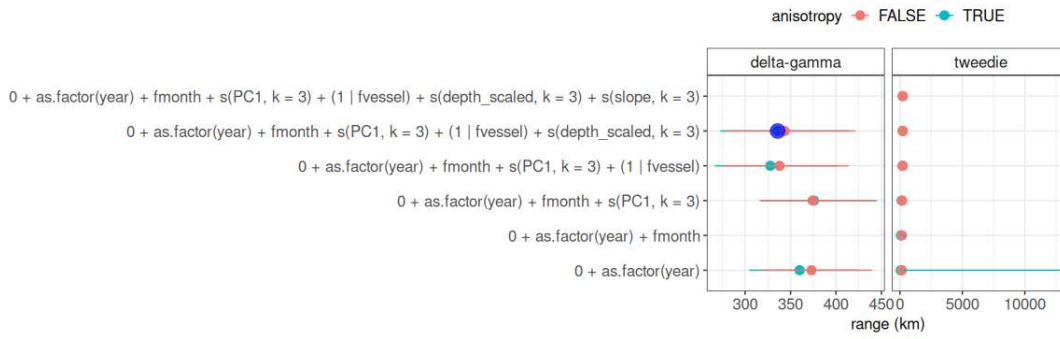


Figure 8: Summary of the estimated range from the different models (Models with different fixed effect structure) for blue sharks. Estimated range from the best model is denoted with blue shade. A) for models that use all the data B) for models that uses only indicator vessel.

A



B

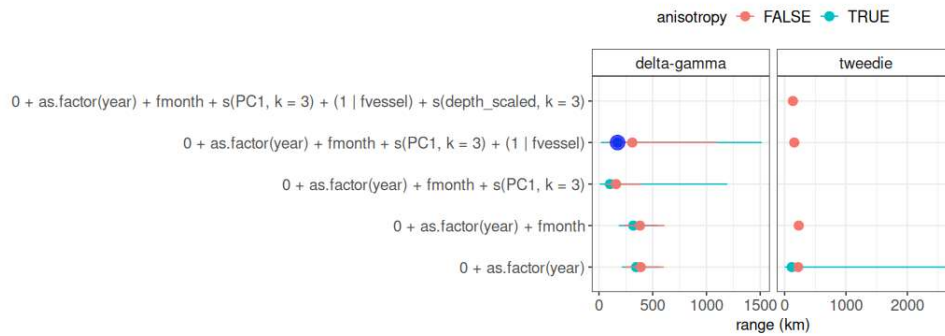


Figure 9: Summary of the estimated range from the different models (Models with different fixed effect structure) for mako sharks. Estimated range from the best model is denoted with blue shade. A) for models that use all the data B) for model that uses only indicator vessel.

3.2 Partial effects

The partial effects of the covariates, depth (depth_scaled) - z-transformed bottom depth, PC1 (first principalcomponent axis score), and month. The partial effects of the covariate on both the positive and binomial components of the delta-Gamma model are summarised below. The partial effect of bottom depth on mako sharks probability of occurrence and catches are shown in Figure 10. Although characterized by wider confidence interval the probability of occurrence of mako shark appears to increase with increase in depth on the other hand positive catches were largely flat.

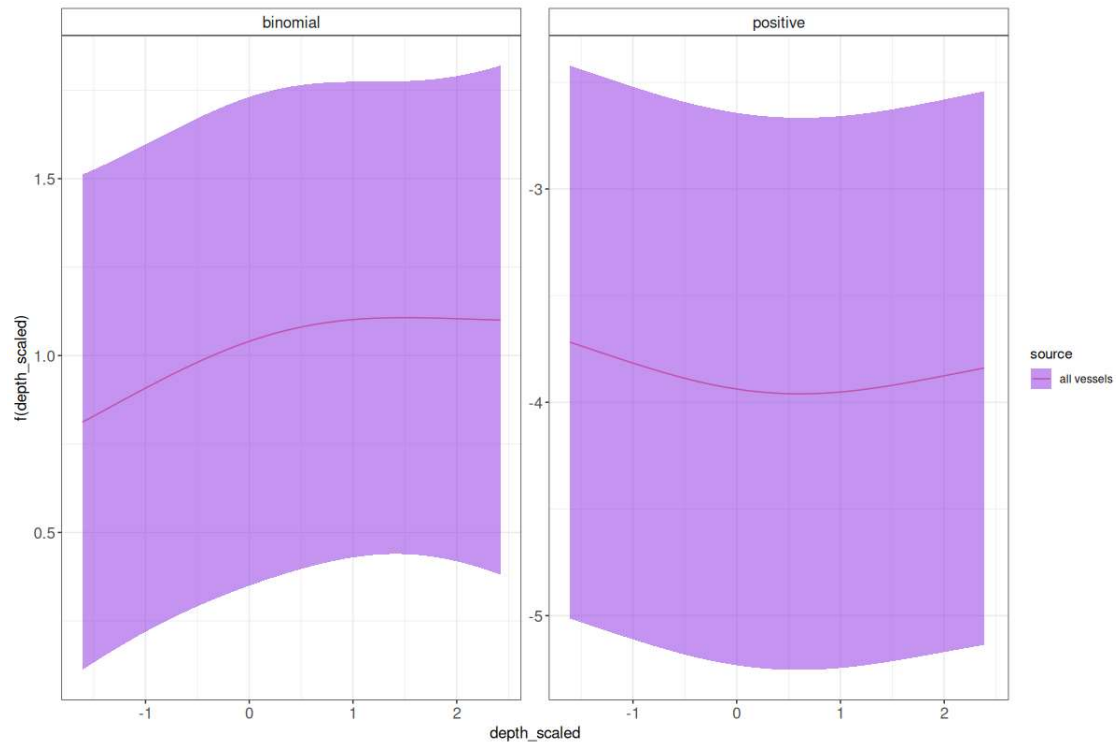


Figure 10: Partial effects of depth on the distribution and abundance of mako sharks for the model based on the data from all vessels.

The effect of the index of targeting, PC1, appears to be largely similar between blue and mako sharks (Figure 11). The form and the direction of the effect of PC1 appear to vary between model based on data from indicator vessels and those based on data from all vessels that landed either of the two pelagic sharks. This contrasting form appears to hold true for both the binomial and positive components of the delta-Gamma model.

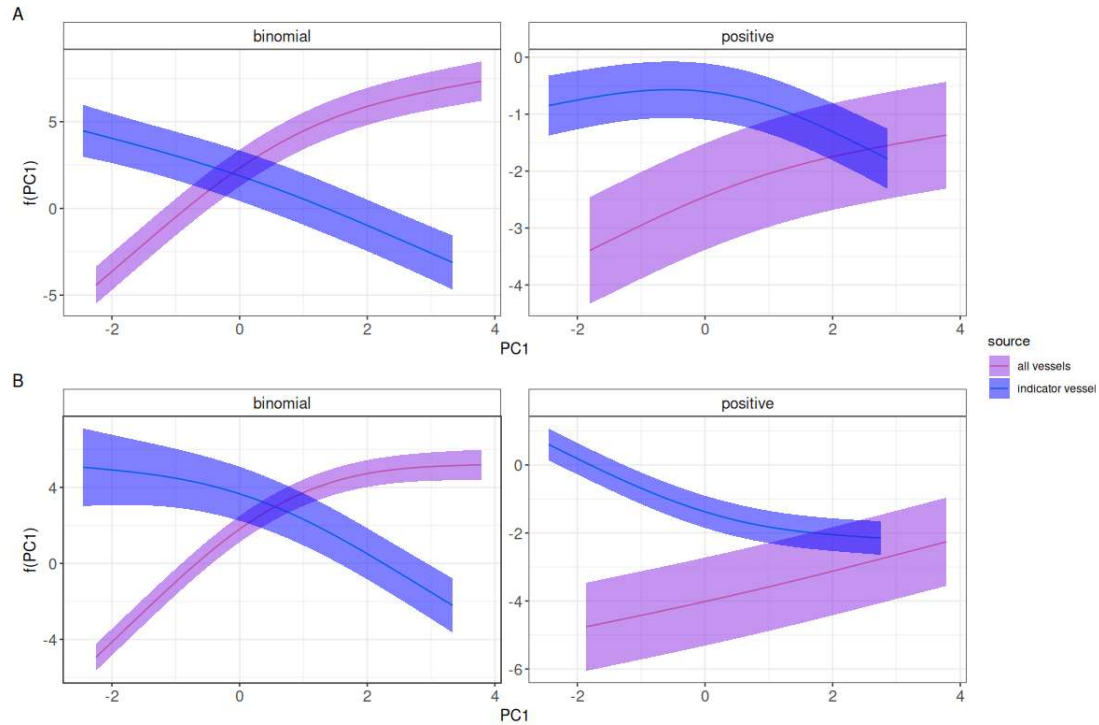


Figure 11: Partial effects of PC1, index of species composition used as proxy for targeting, on the distribution and abundance of A) blue shark B) mako sharks.

Similar to that observed for the partial effect of $PC1$ the partial effect of month on the blue and mako sharks appear to be relatively similar Figure 12. In addition, the forms and direction of the effect of month appears to be similar across the data source (data from all vessels vs data from indicator vessels) for the binomial components. For the positive components, peak in blue shark catch appears to occur late in the year or the start of the year, whereas for mako shark catch appears to peak around the middle of the year.

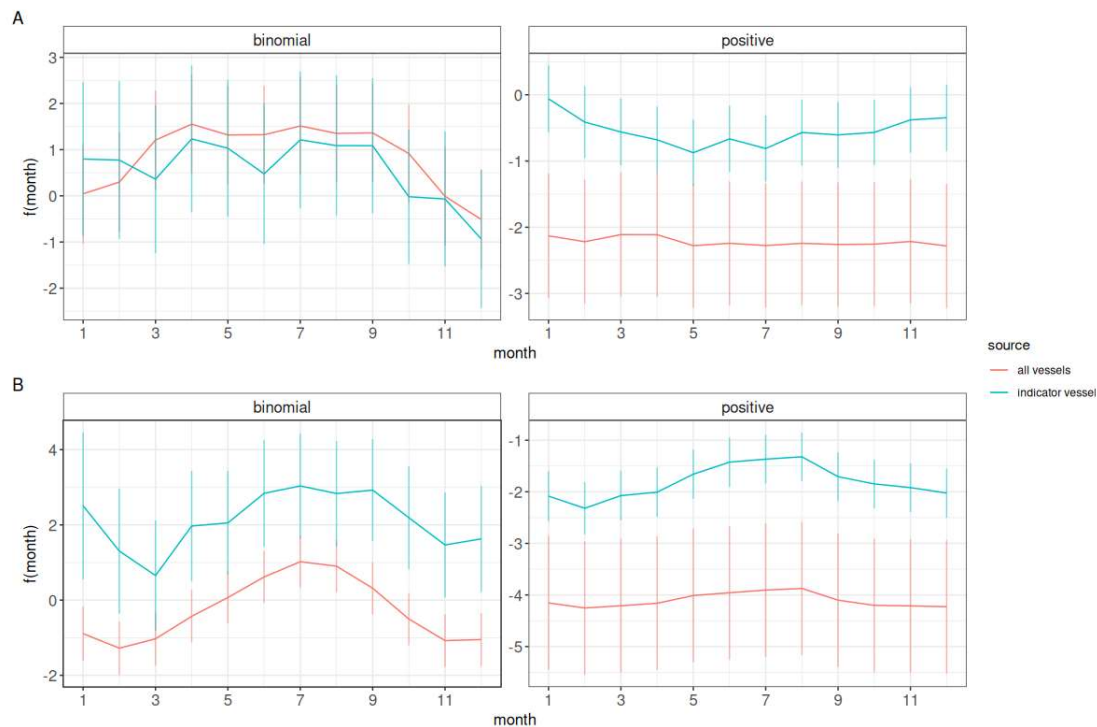


Figure 12: Partial effects of month on the distribution and abundance of A) blue shark B) mako sharks.

3.3 Model diagnostics

Standard model diagnostics, described below, have been checked. To save space, only those corresponding to the model that used the entire data set are presented below. The quantile-quantile plots of residuals, randomized quantile residuals (if the data is consistent with the model residuals should be distributed $N(0,1)$), for the blue shark model that uses all the data is shown in Figure 13 whereas that for mako shark is shown in Figure 14. The quantile-quantile plots of residuals for the blue shark model based on indicator vessels is shown in Figure 15 and that for mako sharks is shown in Figure 16.

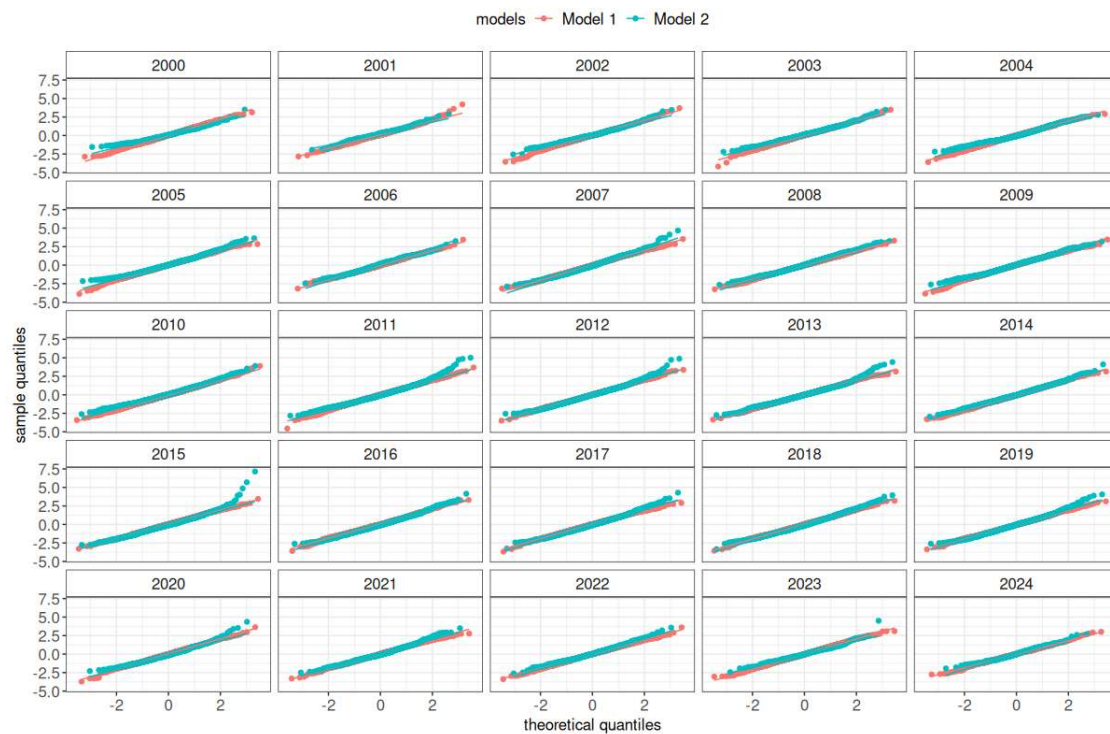


Figure 13: Quantile-quantile plots residuals for the best model for blue sharks that uses data from all vessels.

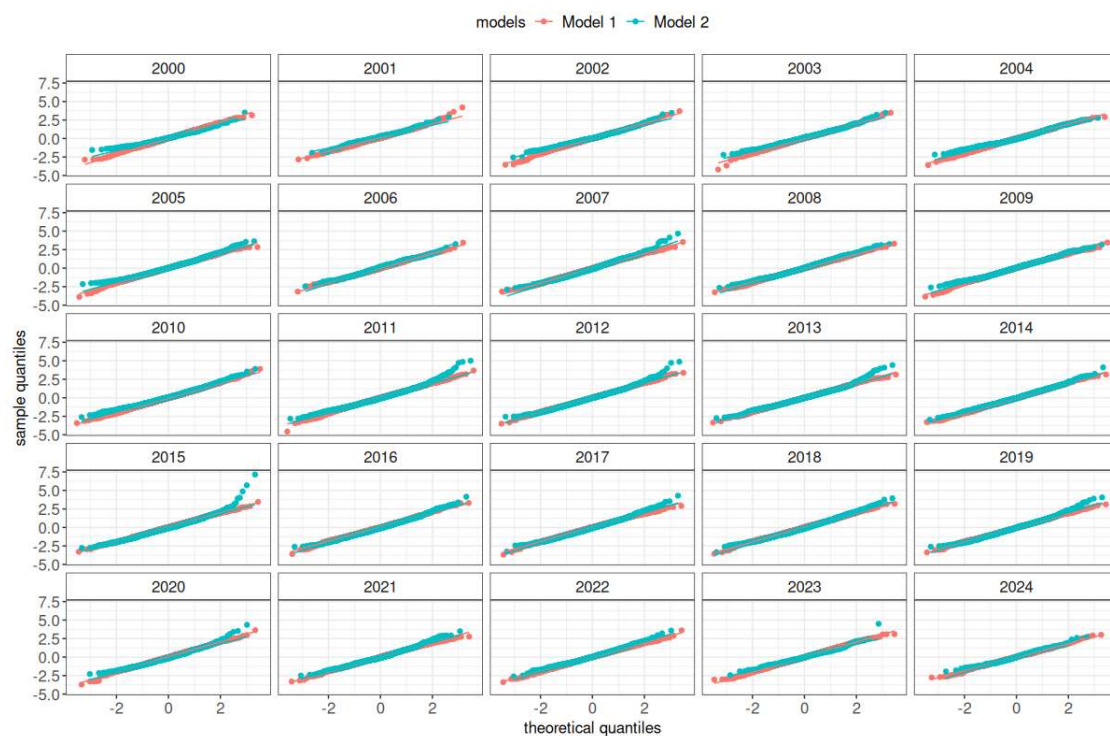


Figure 14: Quantile-quantile plots residuals for the best model for mako sharks that uses data from all vessels.

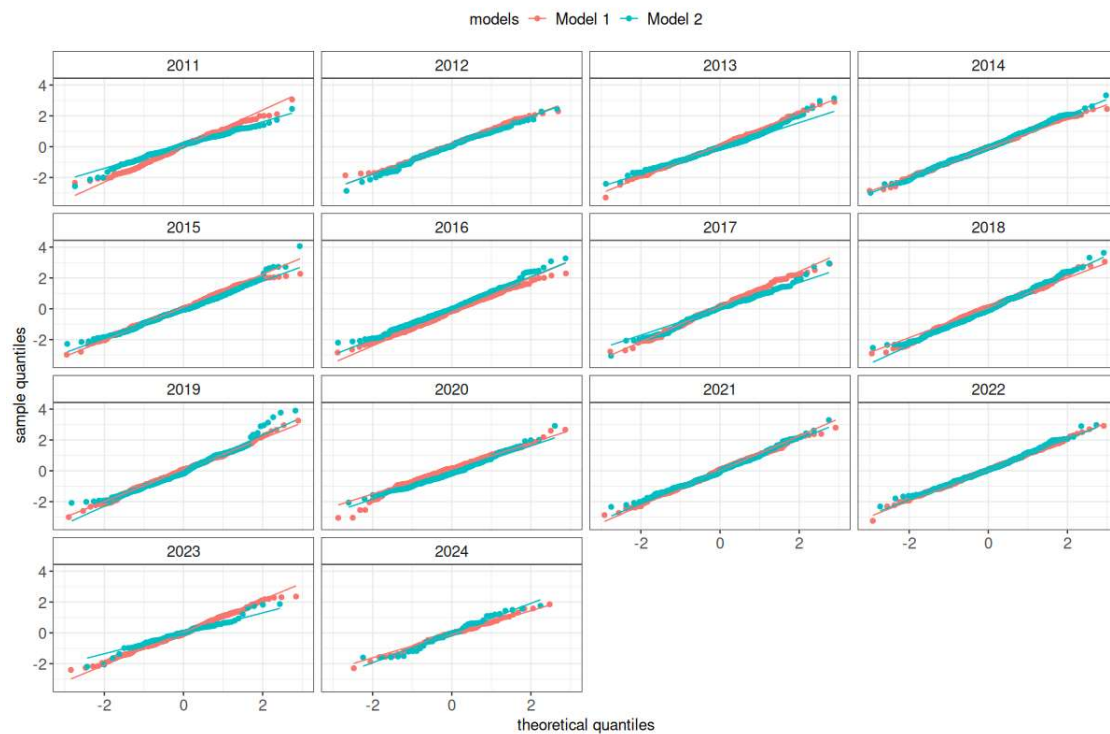


Figure 15: Quantile-quantile plots residuals for the best model for blue sharks that uses data from indicator vessels.

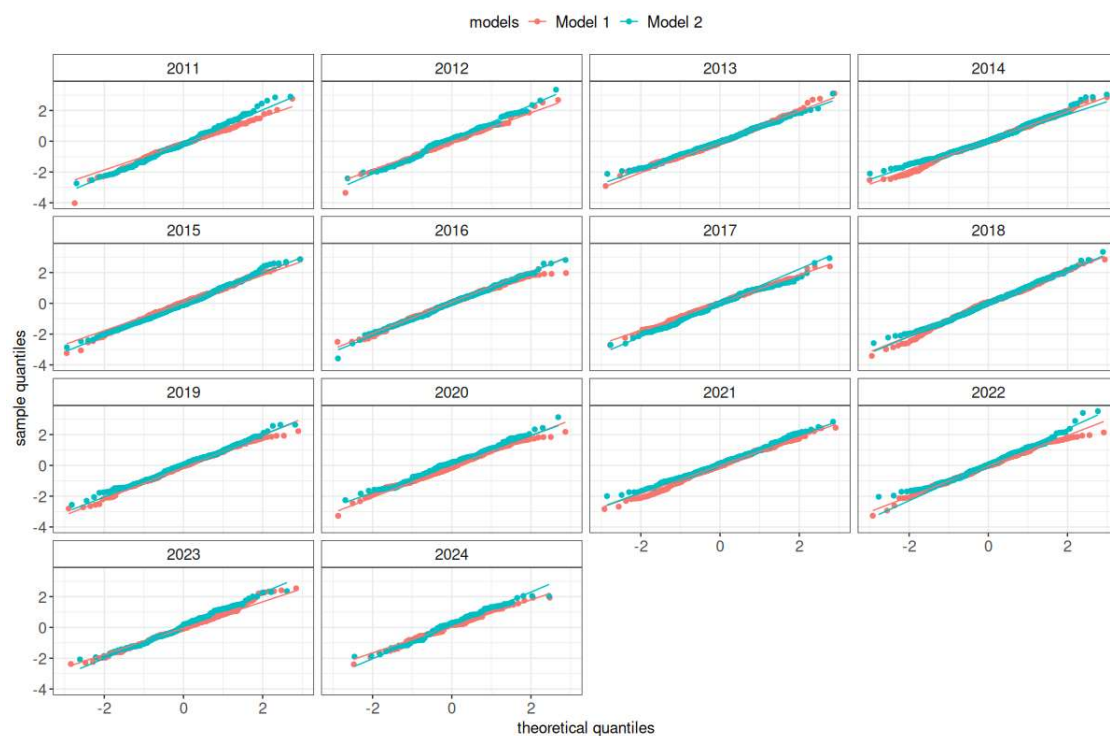


Figure 16: Quantile-quantile plots residuals for the best model for mako sharks that uses data from indicator vessels.

3.4 Spatio-temporal pattern in predicted catch

The predicted catch for the best model that uses all the data are shown in Figure 17 and Figure 18 for blue and mako sharks respectively.

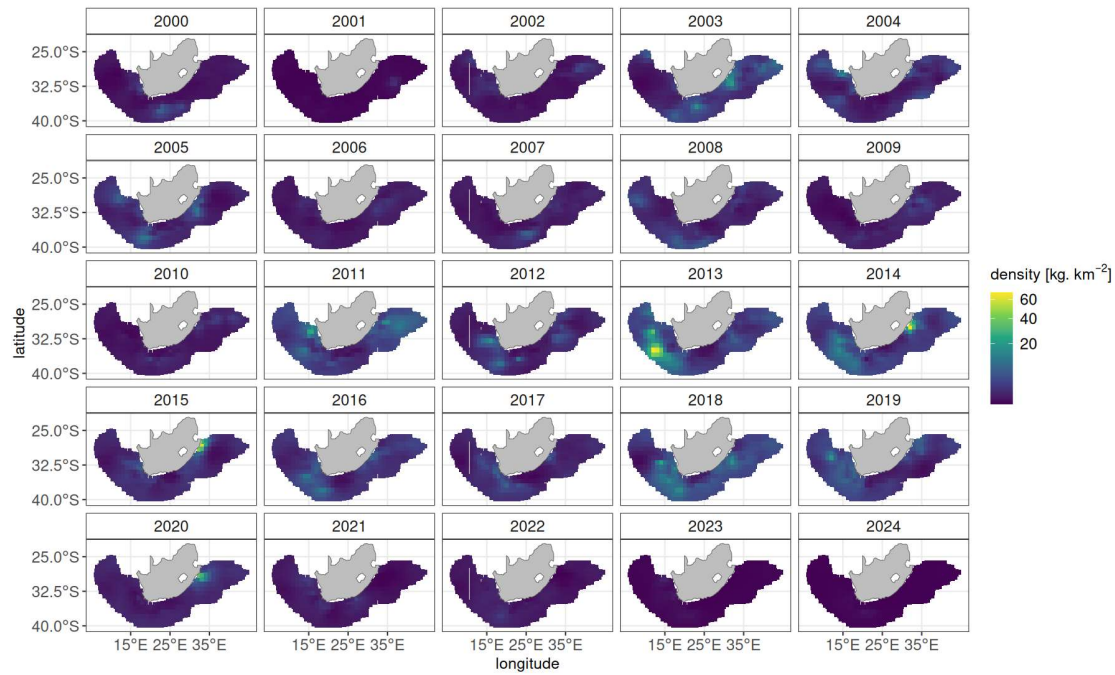


Figure 17: Predictions of catch for blue sharks from the best model that uses data from all vessels operating in the south Atlantic and Indian oceans.

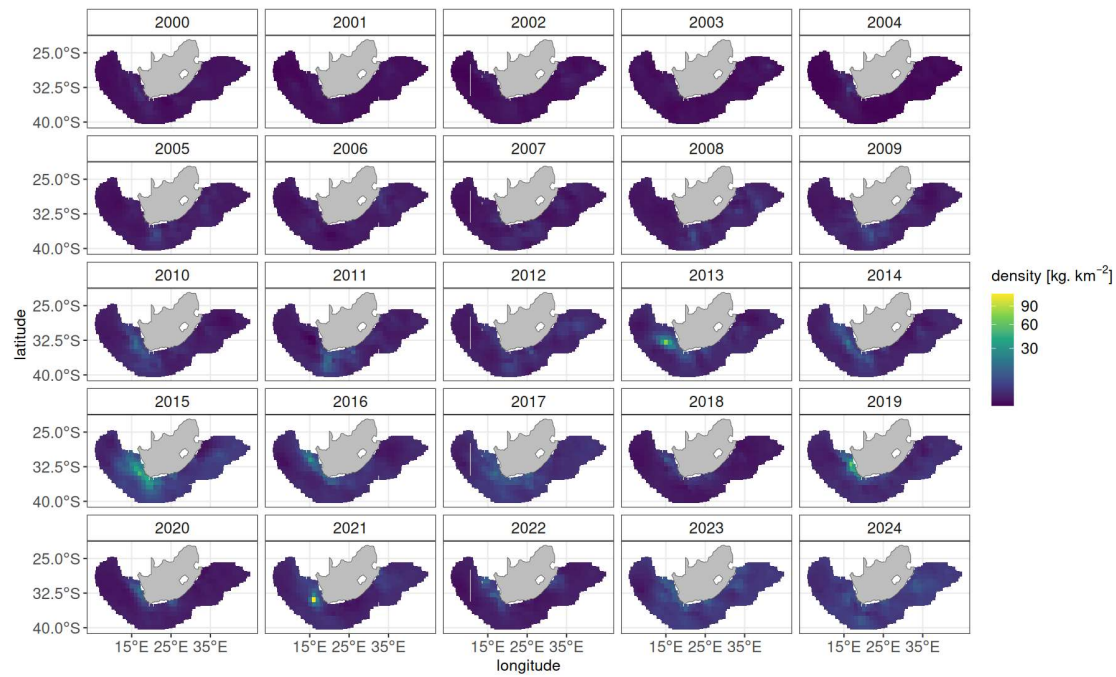


Figure 18: Predictions of catch for mako sharks from the best model that uses data from all vessels operating in the south Atlantic and Indian oceans.

The corresponding predicted catch for the best models that used data from indicator vessels are shown in Figure 19 and Figure 20 for blue and mako sharks respectively.

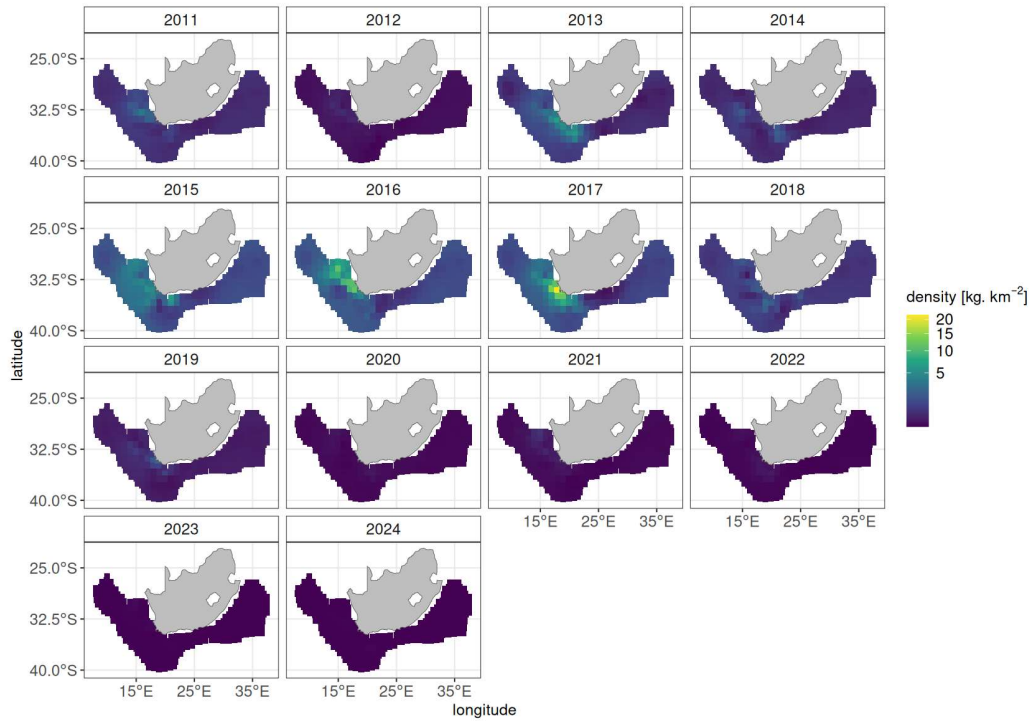


Figure 19: Predictions of catch for blue sharks from the best model that uses data from indicator vessels in the pelagic longline vessels operating in the south Atlantic and Indian oceans.

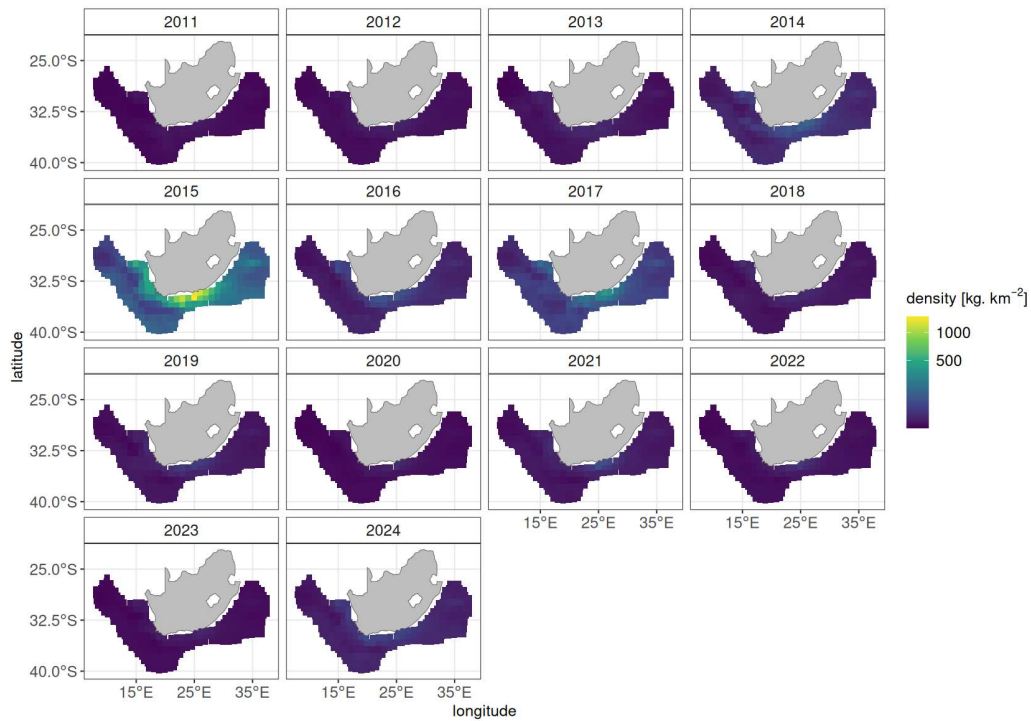


Figure 20: Predictions of catch for mako sharks from the best model that uses data from indicator vessels operating in the south Atlantic and Indian ocean.

3.5 Standardized indices of abundance

The standardized indices from the best model are shown in Figure 21. As it can be seen in the index from indicator vessels are not that different from that based on data from all vessels.

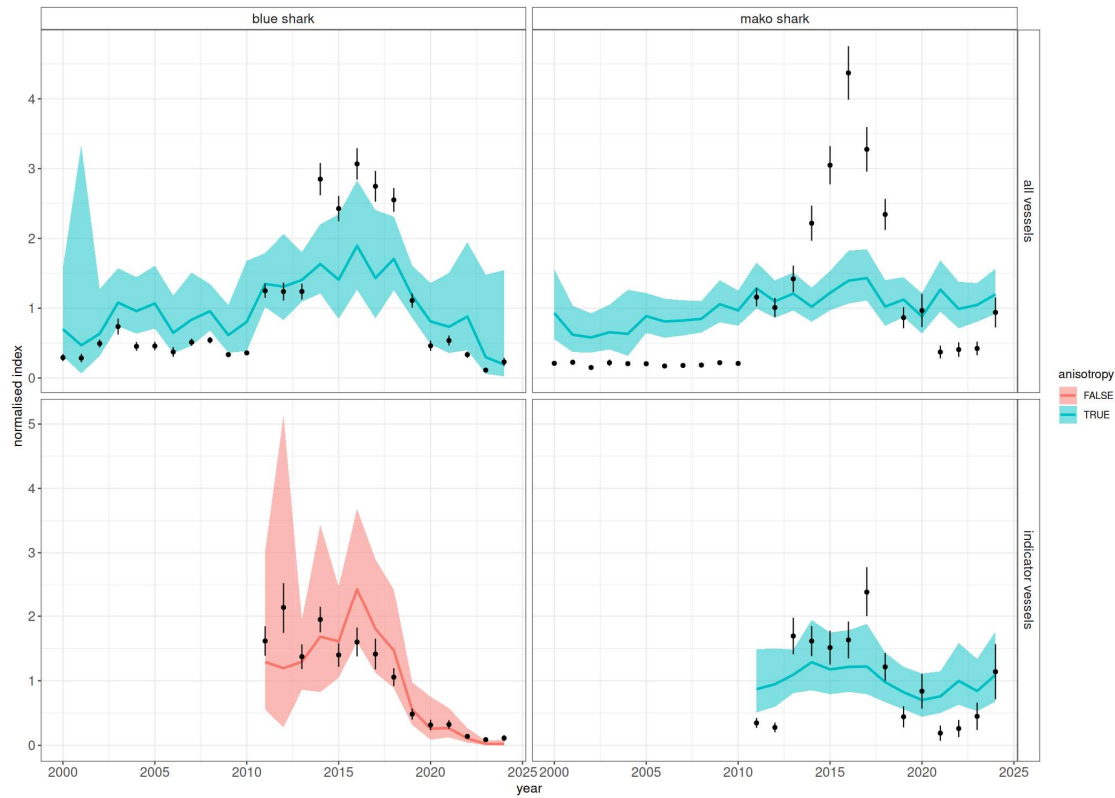


Figure 21: Standardized index from the best model for blue and mako sharks. Filled circles denote the nominal indices.

4. Discussion

The result from this work shows the value of spatio-temporal GLMM in generating standardized *CPUE* for the pelagic-long line fishery operating in the south Atlantic and Indian ocean landing blue and mako sharks.

Appendix

Spatial random effects

Relevant output on model diagnostics is presented below. Estimated spatial random effects for the best model based on data from all vessels are presented in Figure 22 for the binomial component and positive components for blue shark. The corresponding spatial random effect for mako shark based on data from all vessels are shown in Figure 23.

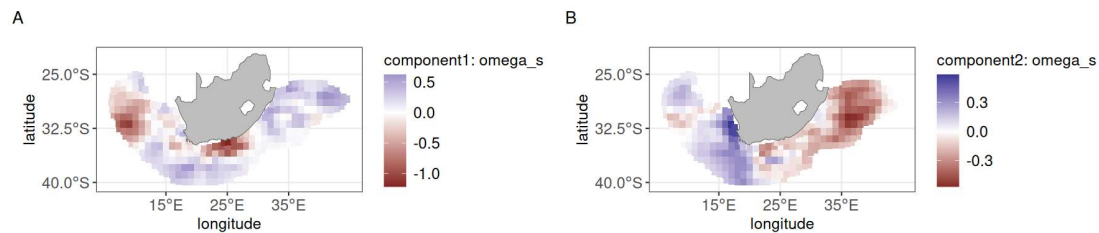


Figure 22: Spatial random effects (omega_s) from the best model for blue sharks, when using data from all vessels in the pelagic longline fishery. A) for binomial component, B) for the positive component. The spatial random effect is expected to account for time invariant effects (both biotic and abiotic) that are not taken into account by the current fixed effect structure.

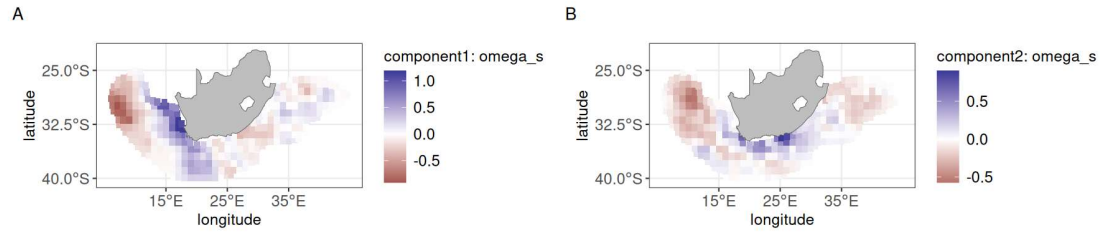


Figure 23: Spatial random effects (omega_s) from the best model for mako sharks, when using data from all vessels in the pelagic longline fishery. A) for the binomial component, B) for the positive component. The spatial random effect is expected to account for time invariant effects (both biotic and abiotic) that are not taken into account by the current fixed effect structure.

Similar to that done for model based on data from all vessels the spatial random effect for the model based on indicator vessels are summarised below. blue and mako sharks spatial random effect based on data from indicator vessels are presented in Figure 24 and Figure 25.

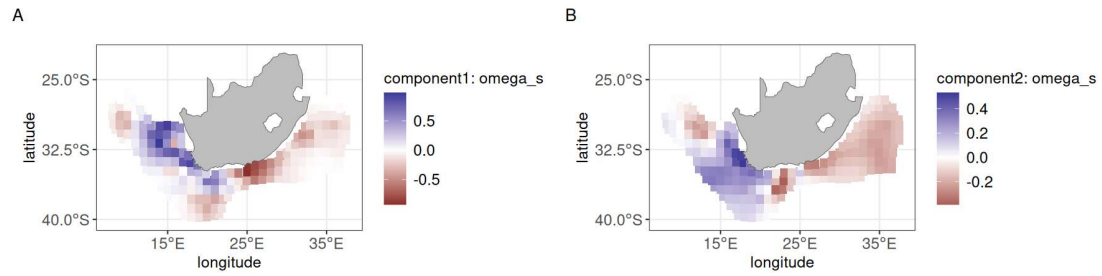


Figure 24: Spatial random effects (ω_s) from the best model for blue sharks, when using data from indicator vessels in the pelagic longline fishery. A) for binomial component, B) for the positive component. The spatial random effect is expected to account for time invariant effects (both biotic and abiotic) that are not taken into account by the current fixed effect structure.

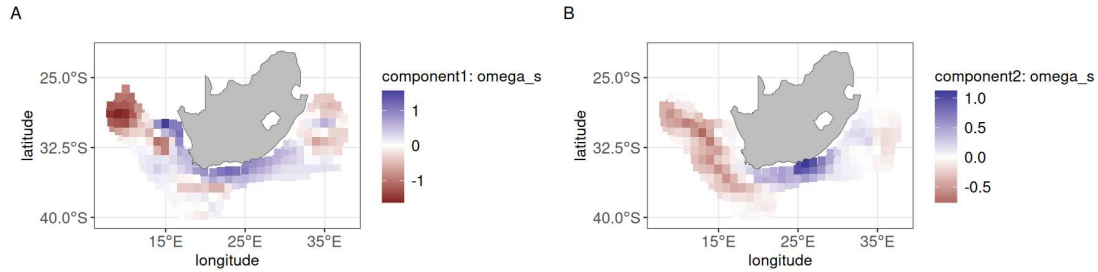


Figure 25: Spatial random effects (omega_s) from the best model for mako sharks, when using data from indicator vessels in the pelagic longline fishery. A) for the binomial component, B) for the positive component. The spatial random effect is expected to account for time invariant effects (both biotic and abiotic) that are not taken into account by the current fixed effect structure.

Spatio-temporal random effect

The spatio-temporal random effect for the best model based on the data from all vessel for the binomial and positive components for blue shark are shown in Figure 26 and Figure 27 respectively. Figure 28 and Figure 29 summarise the spatio-temporal random effect for the binomial and positive components of the delta-Gamma model based on all data for mako sharks.

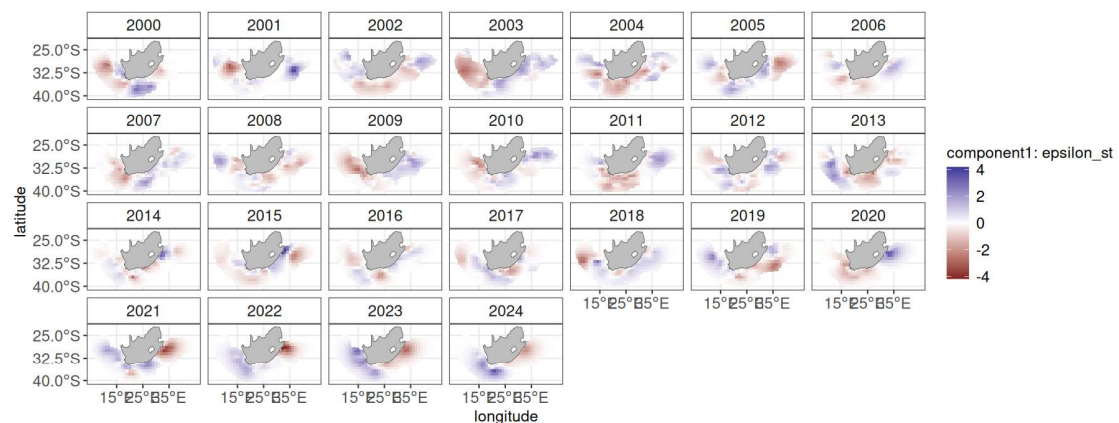


Figure 26: Spatio-temporal random effects (epsilon_st), for the binomial component, accounting for deviation from the fixed effect prediction and spatial random effect for blue sharks for model based on data from all vessels. These represent temporally varying biotic and abiotic effects.

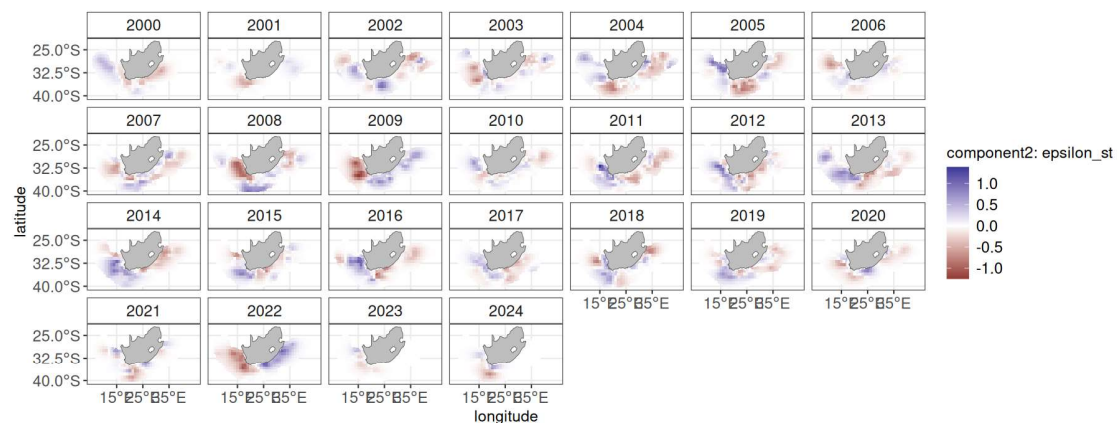


Figure 27: Spatio-temporal random effects (epsilon_st), for the positive component, accounting for deviation from the fixed effect prediction and spatial random effect for blue sharks for model based on data from all vessels. These represent temporally varying biotic and abiotic effects.

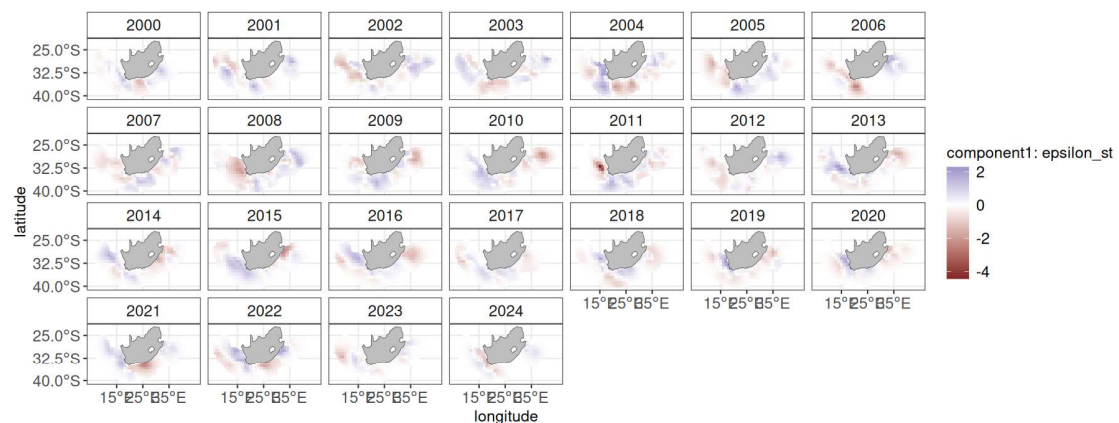


Figure 28: Spatio-temporal random effects (epsilon_st), for the binomial component, accounting for deviation from the fixed effect prediction and spatial random effect for mako sharks for model based on data from all vessels. These represent temporally varying biotic and abiotic effects.

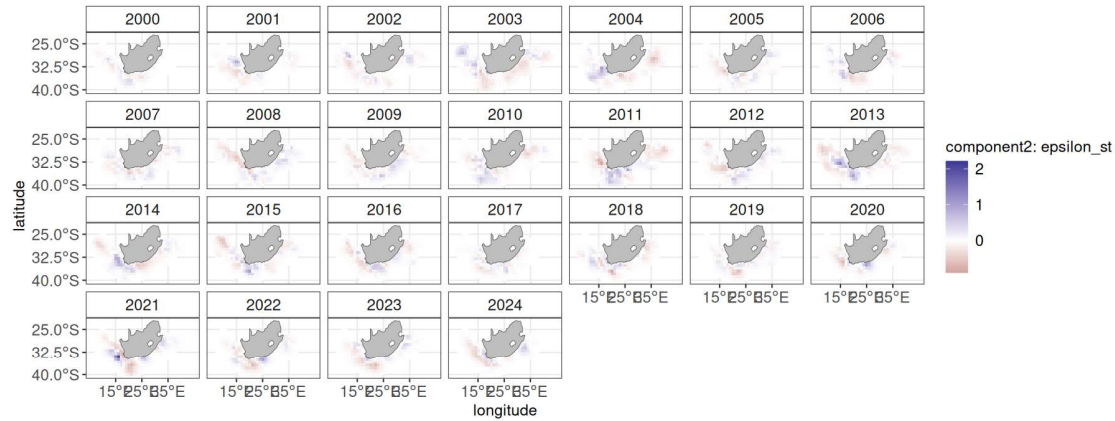


Figure 29: Spatio-temporal random effects (ϵ_{st}), for the positive component, accounting for deviation from the fixed effect prediction and spatial random effect for mako sharks for model based on data from all vessels. These represent temporally varying biotic and abiotic effects.

For the models based on data from indicator vessels the spatio-temporal random effect are summarised below. Figure 30 and Figure 31 show the spatio-temporal random effect for the binomial and positive components for blue sharks. Spatio-temporal random effect for the binomial and positive components for mako sharks are presented in Figure 32 and Figure 33 respectively.

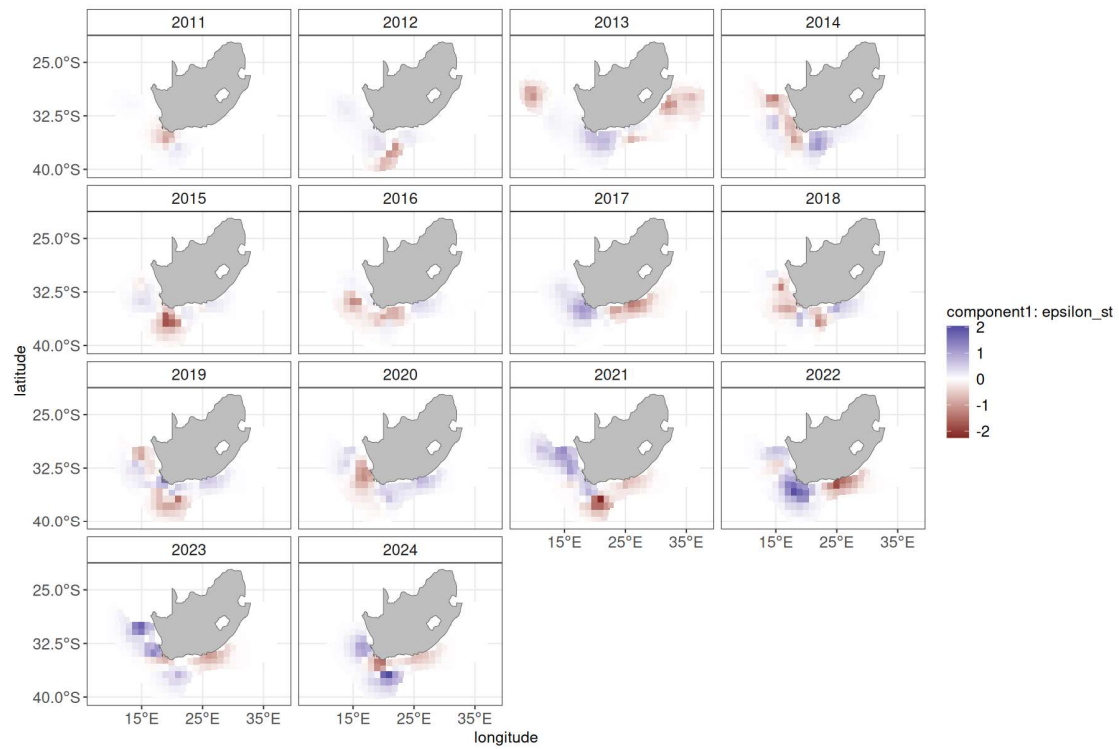


Figure 30: Spatio-temporal random effects (epsilon_st), for the binomial component, accounting for deviation from the fixed effect prediction and spatial random effect for blue sharks for model based on data from indicator vessels. These represent temporally varying biotic and abiotic effects.

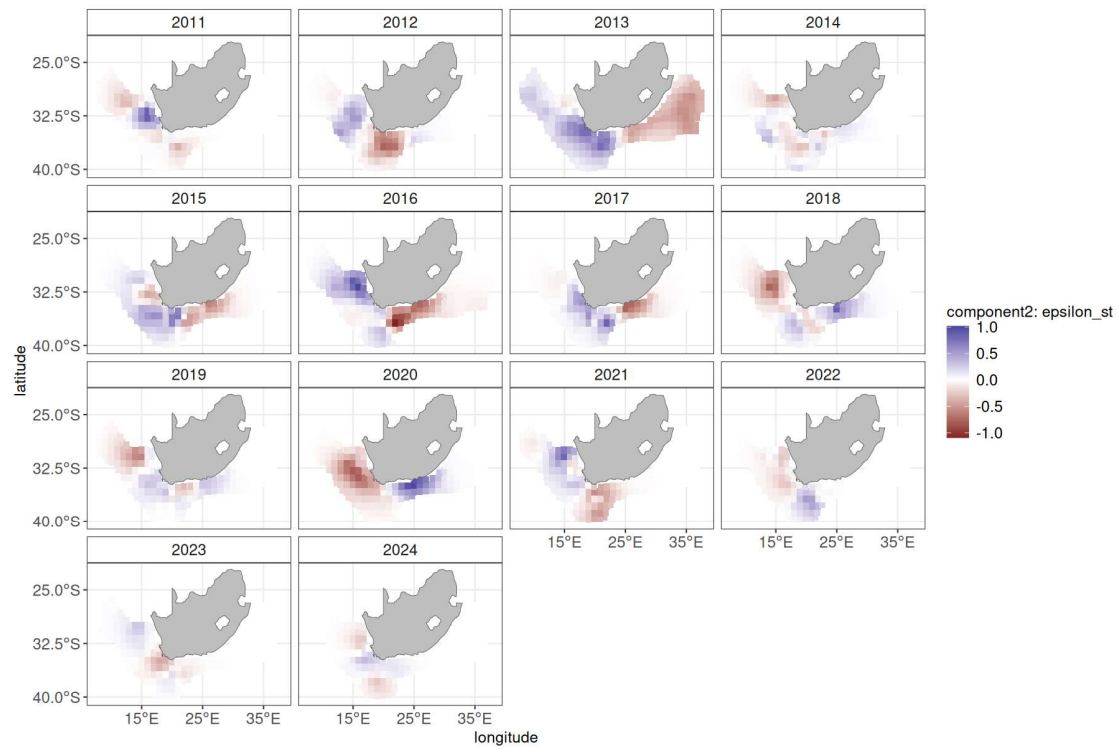


Figure 31: Spatio-temporal random effects (epsilon_st), for the positive component, accounting for deviation from the fixed effect prediction and spatial random effect for blue sharks for model based on data from indicator vessels. These represent temporally varying biotic and abiotic effects.

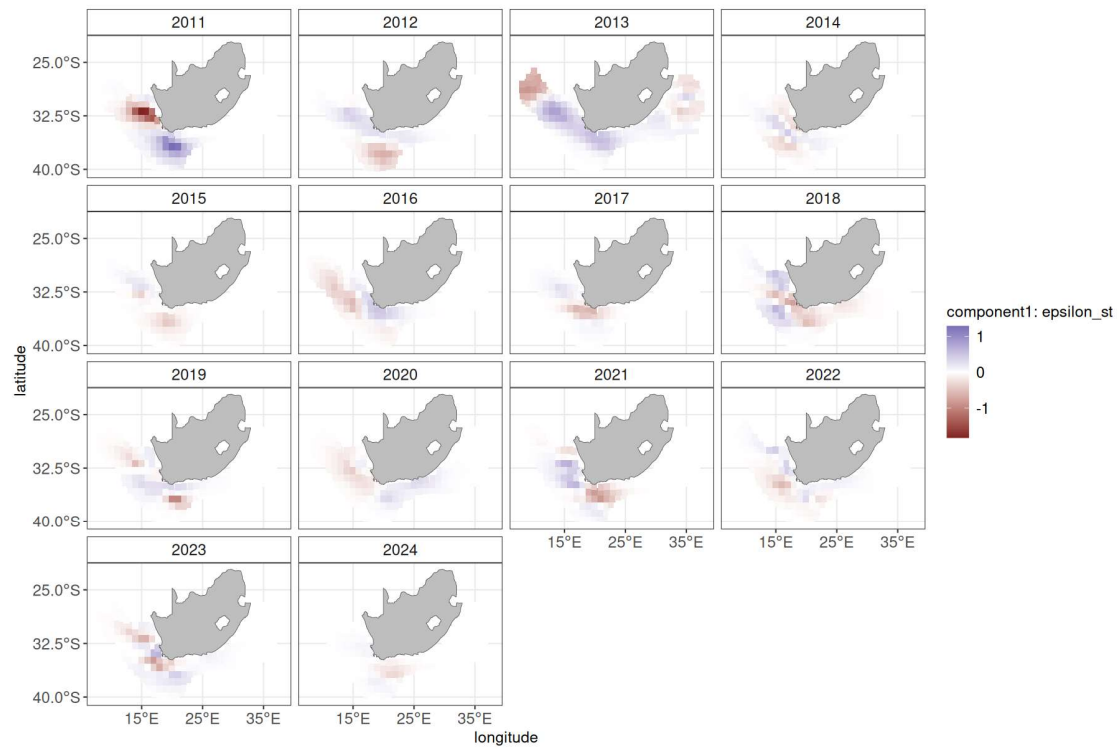


Figure 32: Spatio-temporal random effects (epsilon_st), for the binomial component, accounting for deviation from the fixed effect prediction and spatial random effect for mako sharks for model based on data from indicator vessels. These represent temporally varying biotic and abiotic effects.

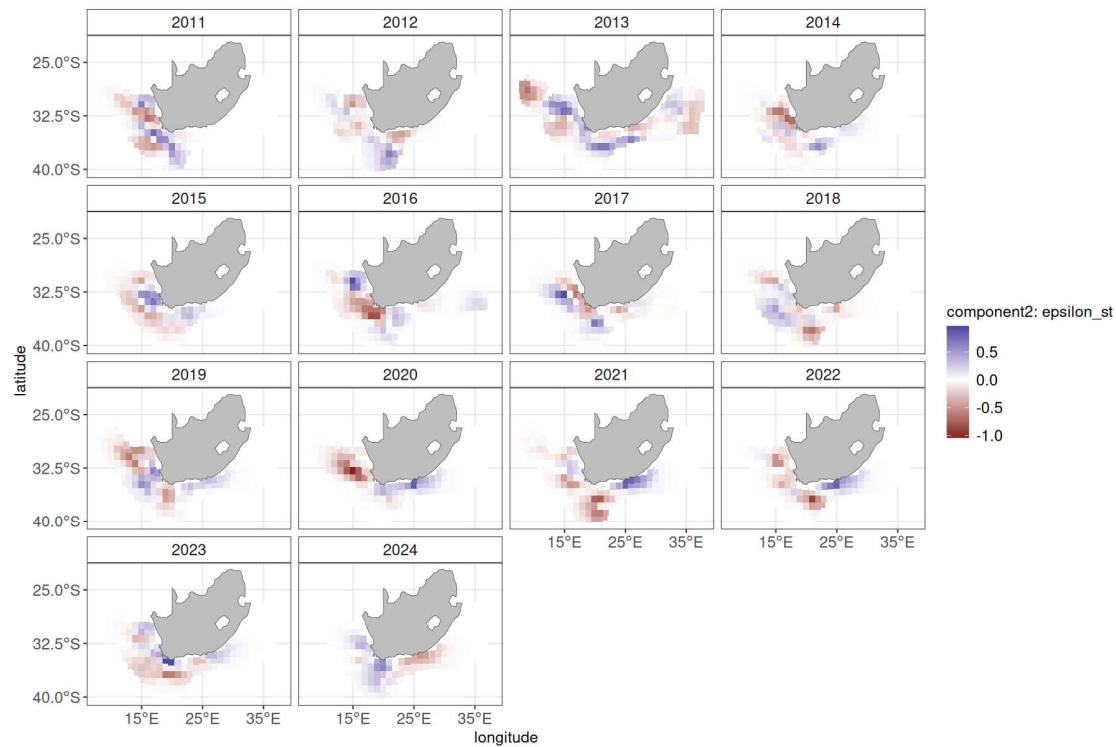


Figure 33: Spatio-temporal random effects (ϵ_{st}), for the positive component, accounting for deviation from the fixed effect prediction and spatial random effect for mako sharks for model based on data from indicator vessels. These represent temporally varying biotic and abiotic effects.

Spatial pattern of residuals

Spatio-temporal pattern in the distribution of residuals for the positive components of the best model based on data from all vessels for blue sharks are shown in Figure 34 and that for mako sharks are shown in Figure 35. The spatio-temporal pattern in the distribution of residuals for the positive components of the best mode based on data from indicator vessels are shown in Figure 36 and Figure 37 for blue and mako sharks respectively.

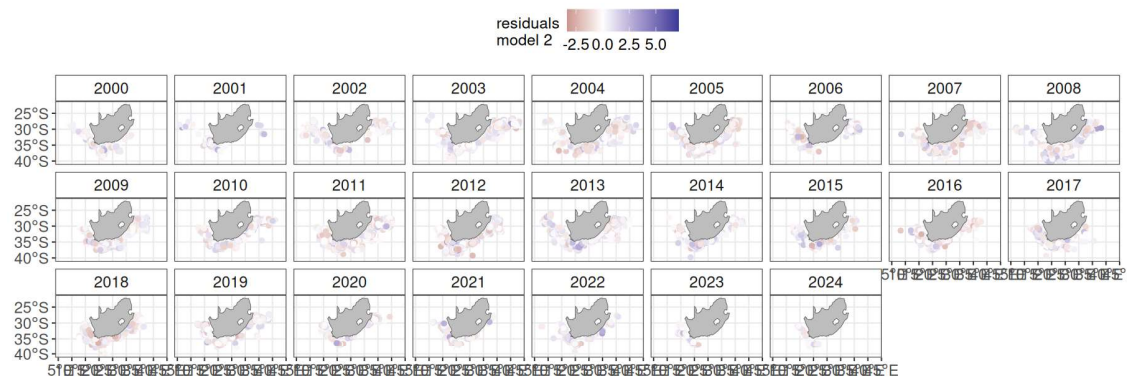


Figure 34: Spatio-temporal pattern in the residuals for the positive components for blue sharks based on the data from all vessels in the pelagic longline fishery.

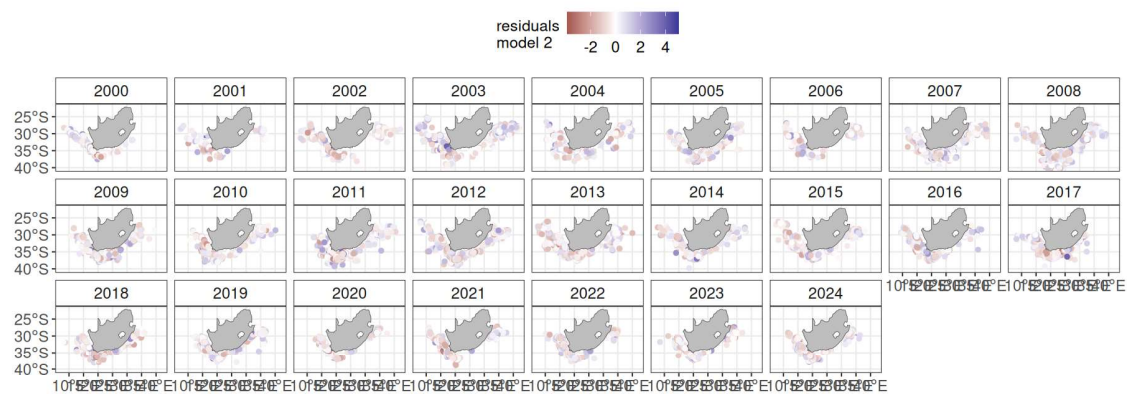


Figure 35: Spatio-temporal pattern in the residuals for the positive components for mako sharks based on the data from all vessels in the pelagic longline fishery.

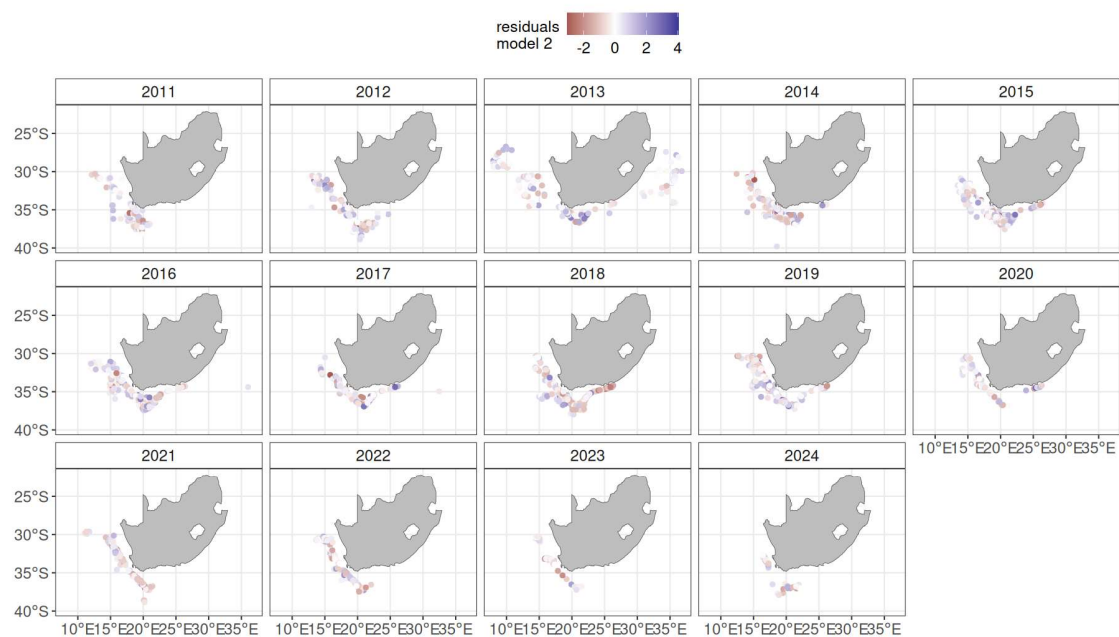


Figure 36: Spatio-temporal pattern in the residuals for the positive components for blue sharks based on the data from indicator vessels in the pelagic longline fishery.



Figure 37: Spatio-temporal pattern in the residuals for the positive components for mako sharks based on the data from indicator vessels in the pelagic longline fishery.

Spatial pattern in prediction uncertainty

Spatio-temporal pattern in prediction uncertainty cv for blue sharks based on data from all vessels is shown in Figure 38, based on indicator vessels is presented in Figure 39. Prediction uncertainty when using data from indicator vessels for blue sharks and mako sharks is presented in Figure 40 and Figure 41, respectively.

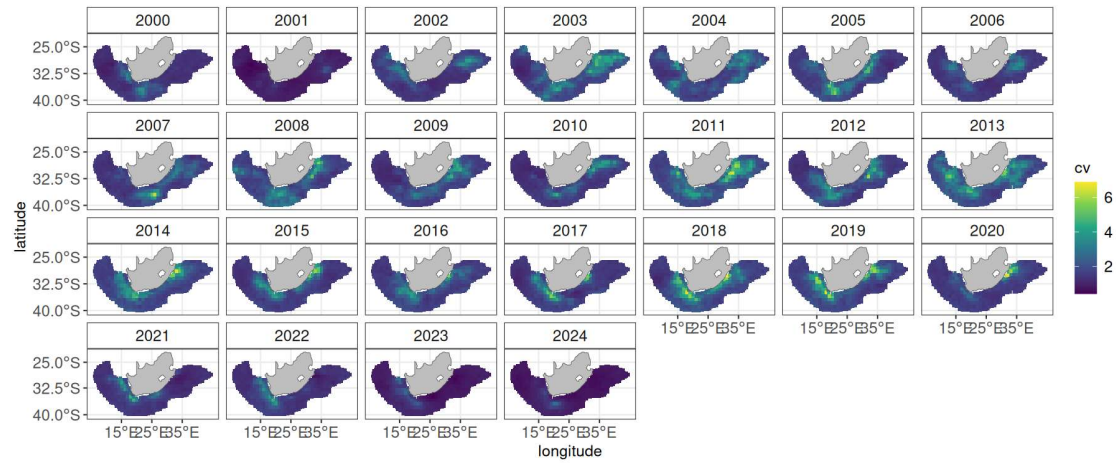


Figure 38: Prediction error from the best model for blue sharks based on the data from all vessels in the pelagic longline fishery.

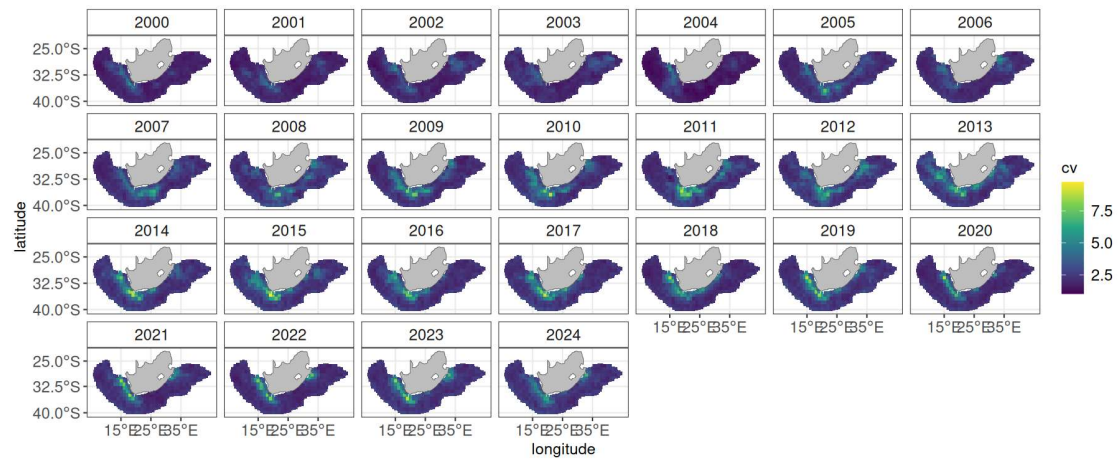


Figure 39: Prediction error from the best model for mako sharks based on the data from all vessels in the pelagic longline fishery.

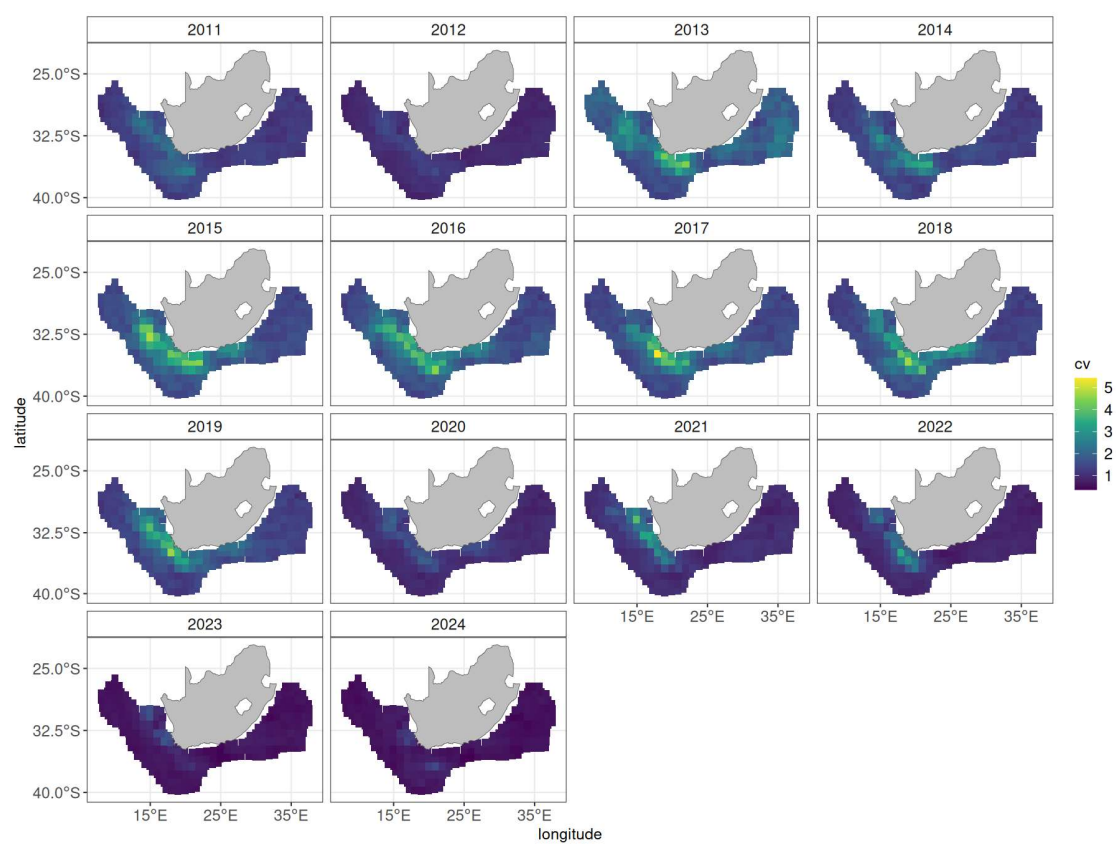


Figure 40: Prediction error from the best model for blue sharks based on the data from all vessels in the pelagic longline fishery.

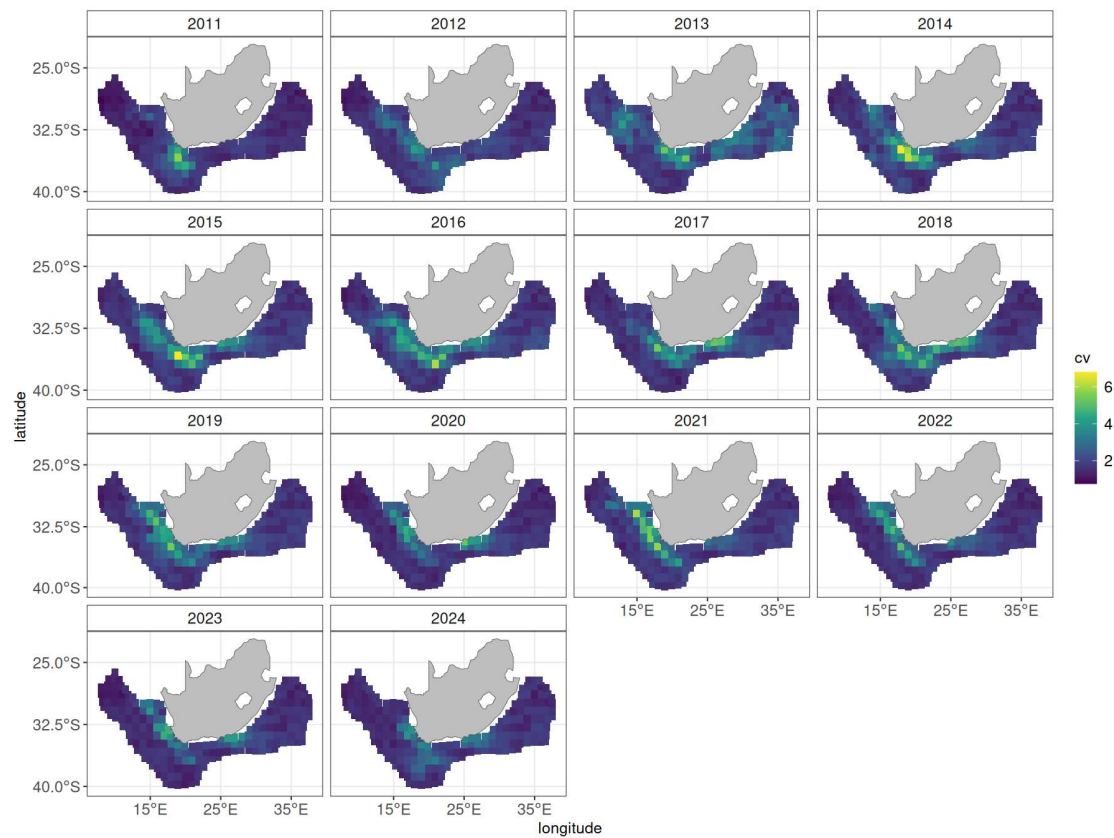


Figure 41: Prediction error from the best model for mako sharks based on the data from all vessels in the pelagic longline fishery.

Session information

System and session info for reproducibility

Setting

Value

version

R version 4.4.2 (2024-10-31)

os

Ubuntu 24.04.2 LTS

system

x86_64, linux-gnu

ui

X11

language

(EN)

collate

en_US.UTF-8

ctype

en_US.UTF-8

tz

Africa/Johannesburg

date

2025-02-21

pandoc

3.2 @ /usr/lib/rstudio/resources/app/bin/quarto/bin/tools/x86_64/ (via rmarkdown)

R packages for reproducibility

	Package	Loaded version	Date
1	broom	1.0.7	2024-09-26
2	captioner	2.2.3.9000	2024-09-11
3	cluster	2.1.8	2024-12-11
4	cowplot	1.1.3	2024-01-22
5	devtools	2.4.5	2022-10-11
6	dplyr	1.1.4	2023-11-17
7	flextable	0.9.7	2024-10-27
8	fmeshier	0.2.0	2024-11-06
9	forcats	1.0.0	2023-01-29
10	ggplot2	3.5.1	2024-04-23
11	ggrepel	0.9.6	2024-09-07
12	inlabru	2.12.0	2024-11-21
13	kableExtra	1.4.0	2024-01-24
14	knitr	1.49	2024-11-08

	Package	Loaded version	Date
15	lattice	0.22-5	2023-10-24
16	lubridate	1.9.4	2024-12-08
17	mgcv	1.9-1	2023-12-21
18	nFactors	2.4.1.1	2022-10-10
19	nlme	3.1-167	2025-01-27
20	patchwork	1.3.0	2024-09-16
21	purrr	1.0.2	2023-08-10
22	readr	2.1.5	2024-01-10
23	scales	1.3.0	2023-11-28
24	sf	1.0-19	2024-11-05
25	stringr	1.5.1	2023-11-14
26	tibble	3.2.1	2023-03-20
27	tidyr	1.3.1	2024-01-24
28	tidyverse	2.0.0	2023-02-22
29	usethis	3.1.0	2024-11-26

References

- Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., *et al.* 2024. Rmarkdown: Dynamic documents for r. <https://github.com/rstudio/rmarkdown>.
- Anderson, S. C., Ward, E. J., English, P. A., Barnett, L. A. K., and Thorson, J. T. 2024. sdmTMB: Spatial and spatiotemporal SPDE-based GLMMs with TMB. <https://pbs-assess.github.io/sdmTMB/>.
- Coelho, R., Rosa, D., Santos, C., and Lino, P. 2023. UPDATED STANDARDIZED CPUES OF BLUE SHARK IN THE PORTUGUESE PELAGIC LONGLINE FLEET OPERATING IN THE NORTH ATLANTIC. Collect. Vol. Sci. Pap. ICCAT, 80: 147–177.
- Fernández-Costa, J., Ramos-Cartelle, A., García-Cortés, B., and Mejuto-García, J. 2023. Update standardized catch rates in biomass for the north atlantic stock of blue shark (*prionace glauca*) from the spanish surface longline fleet for the period 1997-2021. International Commission for the Conservation of Atlantic Tunas.
- Kai, M. 2023. Spatio-temporal model for CPUE Standardization: Application to blue shark caught by the Japanese Tuna longline fishery in the South Atlantic from 1994 to 2021. Collect. Vol. Sci. Pap. ICCAT, 80: 222–237.
- Letaw, A. 2015. Captioner: Numbers figures and creates simple captions. <https://github.com/adletaw/captioner>.
- Maechler, M., Rousseeuw, P., Struyf, A., and Hubert, M. 2024. Cluster: "Finding groups in data": Cluster analysis extended rousseeuw et al. <https://svn.r-project.org/R-packages/trunk/cluster/>.
- Pebesma, E. 2024. Sf: Simple features for r. <https://r-spatial.github.io/sf/>.
- R Core Team. 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Raiche, G., and Magis, D. 2022. nFactors: Parallel analysis and other non graphical solutions to the cattell scree test. <https://CRAN.R-project.org/package=nFactors>.
- Robinson, D., Hayes, A., and Couch, S. 2024. Broom: Convert statistical objects into tidy tibbles. <https://broom.tidymodels.org/>.
- Spinu, V., Golemund, G., and Wickham, H. 2024. Lubridate: Make dealing with dates a little easier. <https://lubridate.tidyverse.org>.
- 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. Elsevier.
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., *et al.* 2024. ggplot2: Create elegant data visualisations using the grammar of graphics. <https://ggplot2.tidyverse.org>.
- Wickham, H., François, R., Henry, L., Müller, K., and Vaughan, D. 2023. Dplyr: A grammar of data manipulation. <https://dplyr.tidyverse.org>.

Wickham, H., and Henry, L. 2023. Purrr: Functional programming tools. <https://purrr.tidyverse.org/>.

Winker, H., Kerwath, S. E., and Attwood, C. G. 2014. [Proof of concept for a novel procedure to standardize multispecies catch and effort data](#). Fisheries Research, 155: 149–159.

Wood, S. 2023. MgcV: Mixed GAM computation vehicle with automatic smoothness estimation. <https://CRAN.R-project.org/package=mgcv>.

Xie, Y. 2024. Knitr: A general-purpose package for dynamic report generation in R. <https://yihui.org/knitr/>.