Environmental Effects on the Spatiotemporal Variability of Sardine Distribution Along the Portuguese Continental Coast

Daniela SILVA, Raquel MENEZES, Ana MORENO,
Ana TELES-MACHADO, and Susana GARRIDO

Scientific tools capable of identifying distribution patterns of species are important as they contribute to improve knowledge about biodiversity and species dynamics. The present study aims to estimate the spatiotemporal distribution of sardine (*Sardina pilchardus*, Walbaum 1792) in the Portuguese continental waters, relating the spatiotemporal variability of biomass index with the environmental conditions. Acoustic data was collected during Portuguese spring acoustic surveys (PELAGO) over a total of 16,370 hauls from 2000 to 2020 (gap in 2012). We propose a spatiotemporal species distribution model that relies on a two-part model for species presence and biomass under presence, such that the biomass process is defined as the product of these two processes. Environmental information is incorporated with time lags, allowing a set of lags with associated weights to be suggested for each explanatory variable. This approach makes the model more complete and realistic, capable of reducing prediction bias and mitigating outliers in covariates caused by extreme events. In addition, based on the posterior predictive distributions obtained, we propose a method of classifying the occupancy areas by the target species within the study region. This classification provides a quite helpful tool for decision makers aiming at marine sustainability and conservation.

Supplementary materials accompanying this paper appear on-line.

Key Words: Environmental effects; Geostatistics; *Sardina pilchardus*; Spatiotemporal Species distribution model; Species occupancy areas.

D. Silva (✉) Centre of Mathematics (CMAT), Minho University, Braga, Portugal (E-mail: *danyela-sylva2@gmail.com*). R. Menezes, Centre of Mathematics (CMAT), Minho University, Guimarães, Portugal (E-mail: *rmenezes@math.uminho.pt*). A. Moreno, A. Teles-Machado, S. Garrido, Division of Modelling and Management of Fishery Resources, Portuguese Institute for the Sea and Atmosphere (IPMA), Lisbon, Portugal (E-mail: *amoreno@ipma.pt*)

A. Teles-Machado (E-mail: *ana.machado@ipma.pt*)

S. Garrido (E-mail: *susana.garrido@ipma.pt*).

A. Teles-Machado, Instituto Dom Luiz (IDL), Campo Grande, Lisbon, Portugal.

1. INTRODUCTION

Models describing the distribution patterns of species are important scientific tools for improving our understanding of biodiversity and species abundance, thereby enabling informed decisions for sustainable biodiversity management. This significance is particularly pronounced in the marine environment, where numerous species, habitats, and ecosystems have experienced substantial declines (McCauley et al. 2015), also due to climate change (Hoegh-Guldberg and Bruno 2010; Pires et al. 2018).

To effectively monitor fishery resources, spatiotemporal analyses of abundance indexes have been developed [e.g., Paradinas et al. (2017); Pennino et al. (2020); Dambrine et al. (2021)]. Particularly, species distribution modeling (SDM) has been employed to comprehend species dynamics, identify and assess the impacts of climate change, define species habitats, and design protected areas (Martínez-Minaya et al. 2018).

Several approaches have been established in the realm of SDM over the past few decades. Indeed, SDMs can be employed to establish connections between species occurrence and/or abundance processes and environmental conditions. They also aid in identifying species' ecological niches and drawing inferences regarding species distribution. Leathwick et al. (2005) employed multivariate adaptive regression splines to elucidate nonlinear relationships between environmental factors and species in New Zealand's freshwater ecosystems. Reiss et al. (2011); Attorre et al. (2011) conducted comparative studies involving diverse SDM techniques, encompassing regression, classification, envelope models, machine learning approaches, and spatial interpolation models. Their studies evaluated the environmental impacts on marine benthic species in the North Sea and tree species in the Italian Peninsula, respectively. To accommodate zero-inflated data within six bird species abundance indexes, Joseph et al. (2009) utilized zero-inflated N-mixture models. The growing recognition of the importance of integrating prior knowledge about species ecology and complex dynamics into the modeling process has driven an increased adoption of Bayesian approaches (Martínez-Minaya et al. 2018). Notably, negative binomial hidden Markov models, as well as hierarchical Bayesian models, were applied by Spezia et al. (2014); Paradinas et al. (2017); Adde et al. (2020), respectively.

Species distribution data often exhibits residual spatial autocorrelation, indicating that observations are not conditionally independent in space. The spatial autocorrelation frequently arises due to the oversight of significant environmental factors, such as climate conditions that influence species distribution, as well as intrinsic factors like competition, dispersal, and aggregation (Miller 2012; Guélat and Kéry 2018). Indeed, contrasting the application of spatial and non-spatial methods on the same dataset can yield different conclusions (Kühn 2007; Kneib et al. 2008). Moreover, a simulation study conducted by Guélat and Kéry (2018) revealed that neglecting spatial autocorrelation can lead to erroneous predictions, even when dealing with extensive datasets. Hence, it is imperative to account for spatial autocorrelation in SDM (Martínez-Minaya et al. 2018). However, certain case studies may involve regions of interest with distinctive shapes or boundaries, such as coastlines, introducing physical barriers. In such cases, the application of classical spatial models like the Matérn field may be misleading due to inherent smoothing over these physical barriers (Bakka et al. 2019).

In addition to the spatial distribution, the temporal scale constitutes an important component that merits consideration within the modeling process since species abundance exhibits variability both in time and space (Hefley and Hooten 2016) and its temporal evolution represents an ecological interest (Paradinas et al. 2017; Martínez-Minaya et al. 2018). Time series analysis operates on a similar principle as spatial statistics, where observations closer in time are more closely related than distant ones.

The objective of the present study is to estimate the spatiotemporal distribution of the European sardine (*Sardina pilchardus*, Walbaum 1792) in the Northern part of the Canary upwelling system (Portuguese shelf). The study seeks to pinpoint the environmental drivers influencing sardine spatial dynamics, comprehend sardine dynamics across both time and space, and characterize the study region based on sardine occupancy.

The European sardine, a small pelagic fish, inhabitats the eastern North Atlantic Ocean (from the North Sea to Senegal), the Mediterranean Sea, the Sea of Marmara, and the Black Sea (Whitehead 1985). It holds socioeconomic importance for Portugal and Spain, emerging as a key species for the purse-seine fishery (Mendes et al. 2019). The late 1990 s witnessed a positive phase in the sardine stock, the past decade recorded persistently low biomass levels (ICES 2017; Pennino et al. 2020; Izquierdo et al. 2022), whereas recent years have witnessed a revival (Cabrero et al. 2019). These fluctuations may be attributed to oceanographic conditions variability, including climate-driven changes that have been demonstrated to affect small pelagic fish generally (Checkley et al. 2017), and the dynamics of the European sardine population specifically (Cabrero et al. 2019; Garrido et al. 2017).

Numerous studies have been undertaken to unravel the intricate aspects of sardine distribution, its habitat preferences, and its interplay with the marine ecosystem. Noteworthy examples include investigations into the effects of environmental conditions on sardine abundance and distribution in Mediterranean waters (Bellido et al. 2008; Voulgaridou and Stergiou 2003; Gordó-Vilaseca et al. 2021), the northwest coast of Africa witnessed (Bacha et al. 2017), the Alboran Sea (Jghab et al. 2019), the Bay of Biscay (Alvarez and Chifflet 2012), and the Portuguese continental coast and the northern Spanish waters (Santos et al. 2012). Despite the socioeconomic importance of this species, the understanding of sardine distribution on the Portuguese shelf remains limited. While Santos et al. (2012); Zwolinski et al. (2010); Rodríguez-Climent et al. (2017) delved into the temporal distribution of sardine and its association with environmental conditions, the spatial dimension was regrettably overlooked. In a different vein, Izquierdo et al. (2022) employed a Bayesian spatiotemporal approach to model the standardized sardine catch-per-unit-effort (CPUE) along the west coast of Portugal. Nonetheless, strong relationships between sardine CPUE and environmental conditions remained elusive.

This paper presents a proposal for hierarchical spatiotemporal SDM designed to estimate the sardine distribution. The modeling of spatiotemporal dynamics for species distribution poses significant challenges, stemming from the growing complexity of the data, data collection methodologies, and the specific characteristics and evolutionary patterns of each species. Moreover, addressing hurdles such as the semi-continuous nature of the response variable, excessive zero values, discrepancies between presence and biomass processes, and the complex geomorphology of the study region adds further complexity. To effectively address these challenges, we propose a two-part model that decomposes the problem

into a series of interconnected levels through probability functions. Our approach adopts a Bayesian framework for modeling species distribution, with the inference process facilitated by the integrated nested Laplace approximation (INLA) methodology (Rue et al. 2009). INLA enables the approximation of posterior marginals of latent Gaussian fields (GFs). Additionally, we utilized the stochastic partial differential equation (SPDE) technique to approximate a Barrier spatial GF (Bakka et al. 2019) with Matérn covariance functions into a discretely indexed Gaussian Markov random field (GMRF) (Lindgren et al. 2011). This choice is made due to the computationally intensive nature of factorizing dense covariance matrices. Moreover, our proposed model encompasses spatiotemporal and temporal effects, incorporating environmental conditions from the recent past. This facet permits the suggestion of different time lags, each accompanied by associated weights for every covariate. Finally, the introduced framework facilitates the construction of maps of species occupancy that can provide valuable insights for decision makers striving for marine sustainability.

This manuscript is structured into four main sections. Section 2 outlines the data, the study area, and the methodology employed. In Sect. 3, we present the obtained results, followed by a comprehensive discussion of the findings in Sect. 4.

2. MATERIAL AND METHODS

2.1. Data

The spatial distribution of sardine biomass was evaluated based on Spring annual acoustic surveys from the Portuguese sprinc acoustic (PELAGO) series, conducted by the Portuguese Institute for the Sea and Atmosphere (IPMA) in Continental Portuguese waters (Fig. 1). Positioned at the northern extent of the Canary upwelling ecosystem, this marine region exhibits distinct topographic and oceanographic features along its western (west of 9°W) and southern (east of 9°W) coasts.

On the west coast, the presence of coastal upwelling dominates during spring/summer, characterized by a southward flowing upwelling jet over the shelf, cold water fronts, and filaments. These features arise due to the persistent northerly winds (Relvas et al. 2007; Teles-Machado et al. 2016). In contrast, winter on the west coast is dominated by the Iberian Poleward current, transporting warmer and saltier waters northward. The season is marked by warm fronts, eddies (Teles-Machado et al. 2016), and buoyant plumes of lower salinity arising from enhanced winter runoff (Peliz et al. 2002).

The southern coast is demarcated by Cape Santa Maria, a geographical boundary segregating distinct oceanographic regimes. To the west, a cyclonic cell, representative of the spring/summer circulation, propels west coast upwelling waters eastward into the Gulf of Cadiz (García-Lafuente et al. 2006).

The main objective of this survey series is to monitor the spatial distribution of abundance, biomass, and various biological parameters of sardine and other small pelagic fish. The survey design entails continuous daytime acoustic measurements along parallel transects, facilitated by a calibrated 38-kHz echosounder. To process the data, the resulting backscatter from the water column is integrated and averaged over 1nm intervals, expressed as nautical area-scattering coefficients [NASC; S_A (in m^2nm^{-2})]. The inter-transect distance varies is
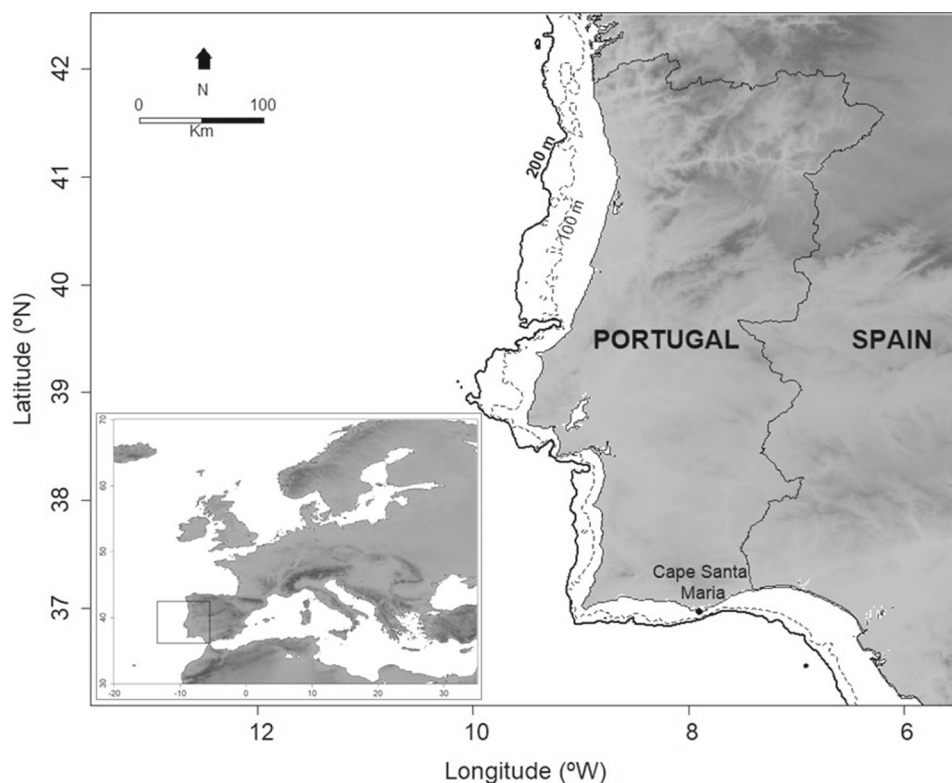
Figure 1. Study area map. Dashed and black lines indicate bathymetric contours at 100 m and 200 m depths, respectively.

6nm. The methodology underpinning the PELAGO series is detailed in Doray et al. (2021). Throughout the analyzed timeframe, 2000–2020 (with a gap in 2012), a total of 16,370 sardine NASC values were collected (refer to Table 1 of Supplementary Material S.1). Each NASC value, serving as a proportionate representation of fish density, is adopted as a proxy for biomass henceforth for a specific pair of coordinates (longitude and latitude). A visual representation of the data's spatial distribution is available in Fig. 2.

To explore the relation between sardine distribution and environmental variables, a comprehensive dataset of daily environmental information was procured from the COPERNICUS server (https://resources.marine.copernicus.eu/products), encompassing the study region and time frame. Specifically, the dataset includes satellite-derived sea surface temperature (SST) measured in degrees Celsius (ESA SST CCI and C3S reprocessed sea surface temperature analyses https://doi.org/10.48670/moi-00169), as well as chlorophyll-a concentration in $mg\ m^{-3}$ (Global Ocean Colour project https://doi.org/10.48670/moi-00281), bathymetry in meters, intensity in $m\ s^{-1}$, and direction in degrees of ocean currents (Atlantic-Iberian Biscay Irish-Ocean Physics Reanalysis https://doi.org/10.48670/moi-00029).

The Portuguese coastline undergoes an abrupt change in direction, forming a "L" shape. This geographical feature introduces significant biomass index variability, particularly in the meridional (N-S) orientation along the west coast and in the zonal (E-W) orientation along
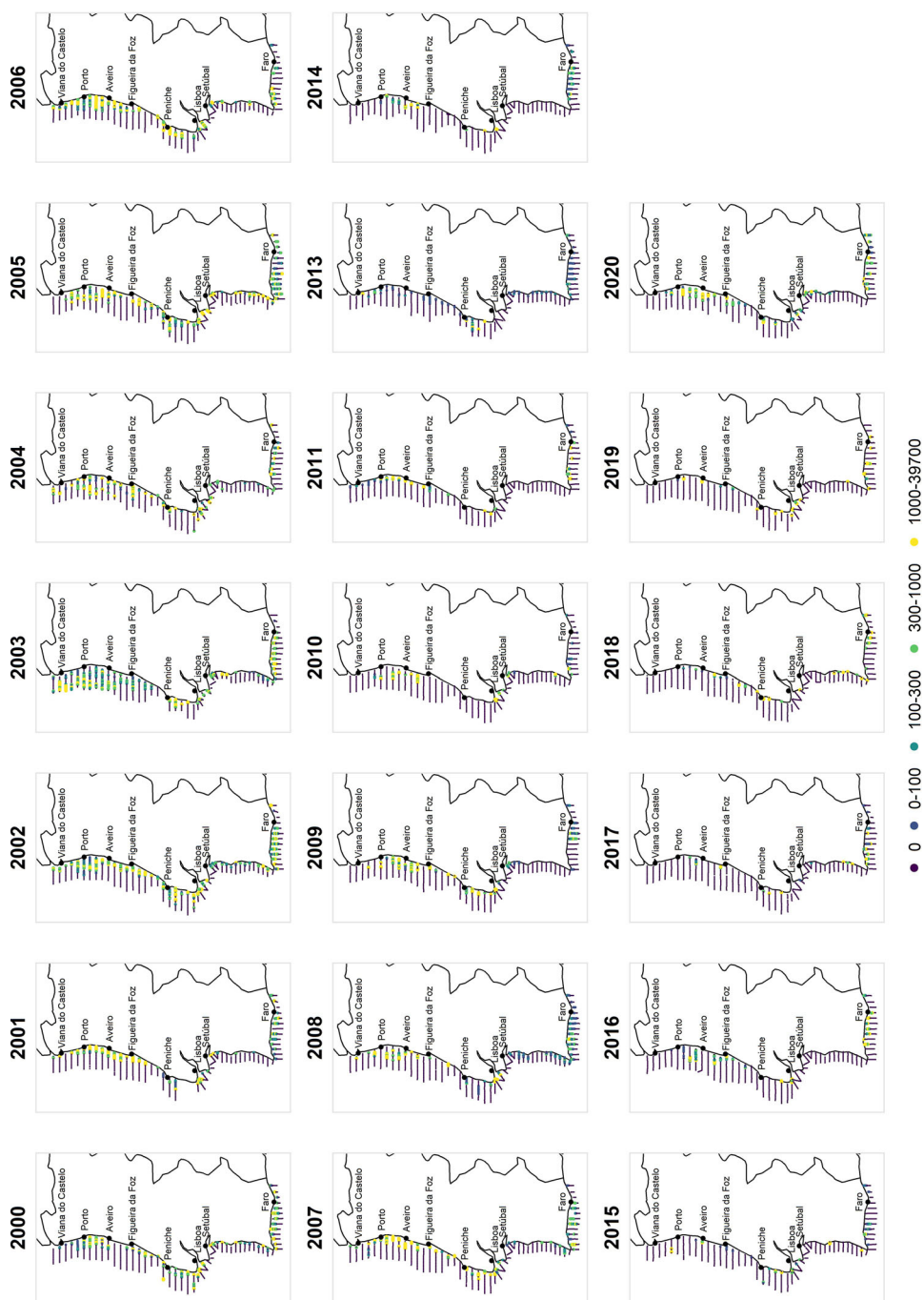
Figure 2. Maps depicting the sardine biomass index (NASC, $n^2 nm^{-2}$) for each annual PELAGO survey conducted between 2000 and 2020 in Portuguese waters.

south coast (refer to Fig. 2). This distinct pattern prompted the incorporation of a binary covariate, representing the coast with values "south" and "west" to capture discrepancies not accounted for by the environmental conditions.

2.2. SPATIOTEMPORAL SPECIES DISTRIBUTION MODEL

2.2.1. Two-Part Model

The species biomass distribution is articulated as the outcome of two distinct components: the presence distribution and the biomass distribution under presence (Eq. 1).

Let Y_{st} be the spatiotemporal distributed biomass process at year $t = 1, \ldots, T$ and location $\mathbf{s} \in \mathfrak{D} \subset \mathbb{R}^2$ where $\mathfrak{D}$ represents the study region. Z_{st} denotes the presence sub-process, taking the binary value 0 if no species was observed at location $\mathbf{s}$ and year t, and 1 otherwise. I.e., NASC $= 0$ indicates absense, while positive NASC values signify presence. $Y_{st}|(Z_{st} = 1)$ takes the positive value of biomass index (i.e., positive values of NASC) observed at location $\mathbf{s}$ and year t. Consequently, the distribution of (represented by [.] in Eq. 1) the process of interest, species biomass index, is given by:

$$[Y_{st}] = [Z_{st}][Y_{st}|(Z_{st} = 1)] = \begin{cases} 1 - \pi_{st}, & y_{st} = 0 \\ \pi_{st}[Y_{st}|(Z_{st} = 1)], & y_{st} > 0 \end{cases} \quad (1)$$

Considering the semi-continuous nature of the data, the process governing presence is assumed to come from a Bernoulli distribution with probability π_{st}. For the biomass process under the presence, a continuous distribution is required such as Gamma or log-Normal. In our case, we employed the Gamma distribution, parameterized by shape (a_{st}) and scale (b_{st}) parameters.

The proposed hierarchical model can be written as follows:

$$Z_{st} \sim Bernoulli(\pi_{st})$$

$$logit(\pi_{sti}) = \alpha' + \beta' X_{Csti} + \sum_{j}^{p'} f'_j(K(X'_{jsti}, c, l)) + \gamma'_t + W_{st}$$

$$Y_{st}|(Z_{st} = 1) \sim Gamma(a_{st}, b_{st})$$

$$log(\mu_{sti}) = log(a_{sti}/b_{sti}) = \alpha + \beta X_{Csti} + \sum_{j}^{p} f_j(K(X_{jsti}, c, l)) + \gamma_t + kW_{st} \quad (2)$$

where i designates the ith day of the survey during year t. For modeling the biomass index μ_{sti} under presence, we adopted the logarithmic link function, whereas the probability of presence π_{st} was represented through the logistic function (Eq. 2). The functions $f_j(.)$ and $f'_j(.)$ denote smoother functions, particularly B-splines, thin plate, and cubic regression splines. The intercepts α and α' serve as regression constants. X_{Csti} signifies the binary covariate with a value of 0 assigned to the west part of the study region and 1 to the south part. Consequently, β and β' capture the impact of the south coast in comparison to the west coast for each respective process.

$K(\cdot)$ represents a weighted average of environmental covariates X_{jsti} observed at location $\mathbf{s}$ and day i of year t such that:

$$K(X_{jsti}, c, l) = \sum_{q=-l}^{l} w_{c-q} X_{jst(i-(c-q))}. \tag{3}$$

The weights w_{c-q} are determined through a Gaussian kernel density (GKD) function, as proposed by Sheather (2004). In this context, c represents the central day, indicating the mode of the GKD, for a time interval prior to day i when the process of interest was observed. Additionally, l denotes the distance between the mode and the minimum (or maximum) of the GKD. Notably, the constraint $0 \leq l \leq c$ is upheld. The GKD function necessitates the bandwidth parameter for its computation. In our study, this bandwidth parameter was estimated using the maximum likelihood cross-validation (MLCV) method (Habbema et al. 1974; Duin 1976). This approach involves estimating the log-likelihood of the density at the kth observation, with all observations except the kth being considered. The utilization of the $K(\cdot)$ function stems from the requirement to account for the effects of environmental conditions with a temporal delay relative to the date when the species biomass index was either estimated or observed. In essence, this function permits the incorporation of covariates from specific past moments and their amalgamation, assigning distinct degrees of importance to different past moments. Supplementary Materials S.3 provide illustrative examples of this concept.

W_{st} denotes a progressive spatiotemporal phenomenon (Paradinas et al. 2017). Underlying this phenomenon is a precision matrix $\mathbf{Q}$, governed by the structure matrix $\mathbf{R}_W$, where $\mathbf{Q} = \tau_W \mathbf{R}_W$ and τ_W is an unknown scalar. Consequently, the structure matrix delineates the nature of temporal and/or spatial interdependencies among elements of $\mathbf{W}$, and it can be factorized as the Kronecker product of structure matrices of the corresponding interacting main effects. In our analysis, the phenomenon exhibits temporal change following a first-order autoregressive process, as defined by the equation:

$$W_{st} = \delta W_{s(t-1)} + \xi_{st} \tag{4}$$

where $t = 2, \ldots, T$, $|\delta| < 1$, and $W_{s1} \sim N(0, \sigma_W^2/(1-\delta^2))$. The selection of the first-order autoregressive model was motivated by the limited temporal span of the data (20 years) and the evident rapid shifts in species distribution, driven largely by climatic changes. Meanwhile, ξ_{st} denotes a zero-mean GF, assumed to be temporally independent. Thus, $Cov(\xi_{st}, \xi_{uh}) = Cov(\xi_s, \xi_u)$ for $t = h$ and $\mathbf{s} \neq \mathbf{u}$.

To address discrepancies between the west and south coasts and the unique geomorphology of the Portuguese continental coast (as depicted in Fig. 2), ξ_s embodies a Barrier spatial GF (Bakka et al. 2019). This approach is rooted in the Matérn covariance function, albeit diverging from conventional methods of computing distances between two points. The methodology accommodates two distinct sets of paths: one confined solely to the study region and the other traversing a physical barrier. For the former set, a stationary Matérn field with marginal variance σ^2 and a range denoted by ϕ is utilized. The latter set involves a Matérn field with the same σ^2 and a range that approximates zero. Defining this range

close to zero effectively nullifies correlation over the physical barrier, which is extraneous to the region of interest.

The interconnection between the two linear predictors is performed by a shared spatiotemporal latent field, W_{st}, governed by a scaling parameter k. This parameter serves to assume the same spatiotemporal correlation structure, while accommodating distinct covariance structures for each process. The proper interpretation of k within the presence probability linear predictor warrants attention. A value of $k = 1$ corresponds to equivalent covariance structures. Moreover, $|k| < 1$ signifies lower variability in presence process when compared to species biomass process, while $|k| > 1$ denotes a higher variability in presence process. In the context of INLA, this approach can be effectively implemented through the copy model, as elucidated by Krainski et al. (2019).

Additionally, temporal dynamics are accounted for through unstructured temporal effects, denoted by γ_t and γ'_t, characterized by a Gaussian exchangeable prior with a mean of zero and precisions τ_γ and $\tau_{\gamma'}$, respectively. These effects are introduced to model variability that cannot be explained by environmental conditions or the spatiotemporal structure. Instead, they encompass variations stemming from survey-related factors, such as disparities in survey timing or the utilization of different vessels.

In order to maintain minimal informativeness, prior distributions were assigned as follows: α, $\alpha' \sim N(0, \infty)$, β, $\beta' \sim N(0, 1000)$, $k \sim N(1, 0.1)$, $log(\tau_\gamma), log(\tau_{\gamma'}) \sim log\text{-}Gamma(1, 0.00005)$ and $log\left(\frac{1+\delta}{1-\delta}\right) \sim N\left(\frac{20}{3}\right)$. For σ^2 and ϕ, the prior specification adhered to the PC prior framework (Simpson et al. 2017). A sensitivity analysis about the priors of parameters was performed (see Supplementary Material S.6).

2.2.2. Model Evaluation

Concerning the evaluation and comparison of models, established metrics rooted in goodness of fit and complexity were employed to guide the selection of covariates and optimal parameter combinations for c and l in Eq. (3). Different combinations were examined for each covariate, with the exception of bathymetry, as it was considered a static covariate across time.

One metric frequently used for assessing model fit within the Bayesian framework is the deviance information criterion (DIC), introduced by Spiegelhalter et al. (2002). This criterion serves to identify models that best explain observed data, minimizing uncertainty concerning observations generated under similar conditions and at the same temporal juncture. The DIC comprises two components, one appraising model fit and the other gauging model complexity.

The conditional predictive ordinate (CPO) represents a metric rooted in the posterior predictive distribution and derived through cross-validation methods (Roos and Held 2011). Specifically, the CPO index is constructed using the posterior predictive distribution at the kth observation, wherein the data is generated while excluding the kth observation itself. In the context of two-part models, the CPO encompasses posterior predictive distributions of both estimated processes, denoted as $P(\tilde{z}|z)$ and $P((\tilde{y}|\tilde{z})|(y|z))$. A global measure of model fitting is given by the log-conditional predictive ordinates (LCPO).

Consequently, the selection of covariates, smoother terms, and c and l parameter combinations was guided by the pursuit of minimized DIC and LCPO values. The decision-making process is elucidated in detail in Supplementary Material S.4.

2.2.3. Areas of Occupancy

Given the acknowledged impact of ecosystem conditions on species distribution and behavior, the characterization of the study area in terms of species occupancy holds significant value for marine ecology and fisheries management. Four distinct categories of biomass index were delineated: *rare*, *low occasional*, *high occasional*, and *preferred*. The classification process was based on predictions of the biomass index and its associated uncertainty. Biomass index predictions were determined through the median of posterior predictive distribution, $F_{\tilde{Y}_{st}}^{-1}(0.5)$, derived for each new location using Equation (2) of Supplementary Material S.2. Concurrently, the level of uncertainty was defined by the uncertainty linked with the posterior distribution of the interest process, $P\left(|W_{st}| > \frac{4}{5}\sigma\right)$, since W_{st} encapsulates unexplained variability inherent to the interest process, beyond the scope of considered covariates. σ denotes the estimated posterior mean of the marginal standard deviation, while $\frac{4}{5}$ serves as a scaling parameter for the explained variability encapsulated by the spatial structure, the marginal standard deviation itself.

Two distinct thresholds were established, one for the biomass index and another for the uncertainty measure. These thresholds were informed by the annual median of the predicted biomass values within the study area, as well as the median of the uncertainty support - fixed at 0.5. Consequently, *rare* areas are characterized by biomass values falling below the biomass threshold and possessing low uncertainty (below the uncertainty threshold). *low occasional* areas feature low biomass values coupled with high uncertainty (above the uncertainty threshold). In contrast, *high occasional* areas are marked by elevated biomass values exceeding the biomass threshold, accompanied by high uncertainty. *Preferred* areas are distinguished by high biomass values and low uncertainty.

Furthermore, the exploration of the temporal evolution of these distinct areas can facilitate the identification of both favorable and unfavorable habitats—an integral aspect in defining species habitat and comprehending species dynamics. In this particular study, *favorable* areas are defined as those exhibiting a sustained period of *preferred* occupancy, while *unfavorable* areas are characterized by *rare* and persistent species occupancy. For the purposes of this study, persistence was defined by a duration of 3 years due to the limited time series. Two-part model was performed using the INLA approach in R software. The R code corresponding to the creation ok kernel time-lagged covariates, the fit of Barrier and the definition of occupancy areas is available on Github (https://github.com/SilvaPDaniela/Environmental-Effects-on-the-Spatio-Temporal-Variability-of-Sardine-Distribution).

3. RESULTS

Bathymetry, chlorophyll-a concentration, and ocean current intensity surfaced as explanatory variables for both sardine presence and the biomass index, whereas SST exclusively contributed to explaining the species' biomass index (Table 1, Fig. 3).

Table 1. Illustration of the most effective sardine biomass model for the Portuguese continental coast based on performance metrics, DIC and LCPO

Presence process Z_{st}	Biomass process under presence $Y_{st}(Z_{st} = 1)$
Coast + K(CHL,2,2)+ Bathymetry + K(INT,24,4) + **Struc$_{st}$** + **Unstruc$_t$**	K(SST,0,0) + K(CHL,17,0) + Bathymetry + K(INT,10,6) + **Struct$_{st}$** + **Unstruc$_t$**

Key components highlighted in bold include random factors, spatiotemporal effect (**Struc$_{st}$**), and unstructured annual effect (**Unstruc$_t$**). Specific covariates are indicated by acronyms (SST: sea surface temperature, CHL: chlorophyll-a concentration, INT: intensity of ocean currents). The "Coast" variable signifies the coast indicator (0 for West, 1 for South)

The best model embraced purely unstructured annual effects, symbolized by γ_t and γ_t', for the biomass and presence processes (Table 1), respectively. Although structured annual effects, characterized by linear effects or autoregressive processes with varying orders, were also considered, their incorporation did not yield a discernible enhancement in model performance.

Notably, the coastal indicator solely demonstrated relevance in explaining the species' presence, with the corresponding parameter yielding a posterior mean of 0.89 (Figure 3 of Supplementary Material S.5). This finding indicates that, on average, the probability of sardine occurrence is more than twice as high in the southern region compared to the western coast. Chlorophyll-a observed in the four days leading up to, and including, the day of biomass index determination emerged as influential factors in explaining sardine presence. Interestingly, the effect of chlorophyll-a exhibited a pronounced inflection point at $0.45\,\text{mg}\,\text{m}^{-3}$, with a decreasing impact up to this threshold followed by an ascending effect beyond it. Noteworthy patterns also emerged in relation to water depth, where the probability of presence exhibited a peak at 30 m before declining. Bathymetry demonstrated limited influence, becoming irrelevant beyond a depth about 615 m. The probability of sardine presence was positively correlated with lower current intensity values observed within time interval from 20 to 28 days prior to biomass index determination (peaking around $0.08\,\text{m}\,\text{s}^{-1}$, $0.16\,\text{m}\,\text{s}^{-1}$, and $0.22\,\text{m}\,\text{s}^{-1}$) while exhibiting a negative relationship beyond $0.26\,\text{m}\,\text{s}^{-1}$, though the intensity effect became irrelevant beyond a threshold of $0.37\,\text{m}\,\text{s}^{-1}$.

The sardine biomass index showcases a distinctive pattern in relation to SST, where an ascending trend is evident up to 14°C, followed by a relevant decline starting at 15.3°C with increasing SST. The biomass index displays an increasing trend aligned with chlorophyll-a levels, observed 17 days prior to sardine biomass estimation, up to $1.14\,\text{mg}\,\text{m}^{-3}$. Subsequently, a prominent decrease manifests up to $1.25\,\text{mg}\,\text{m}^{-3}$, followed by another relevant upsurge within the range of 10 to $30\,\text{mg}\,\text{m}^{-3}$. The relationship between sardine biomass and bathymetry also reveals noteworthy insights, with a positive correlation observed. This influence is most pronounced at depths of up to 125 m, beyond which its effect becomes marginal, ultimately tapering off at depths exceeding 420 m. The impact of intensity, within the temporal interval spanning from 4 to 16 days prior to biomass index determination. On sardine biomass index, it is characterized by oscillating patterns, featuring four distinctive peaks of positive effect at 0.02, 0.1, 0.18, and $0.67\,\text{m}\,\text{s}^{-1}$. Beyond the threshold of $0.7\,\text{m}\,\text{s}^{-1}$, the influence of ocean current intensity becomes negligible.
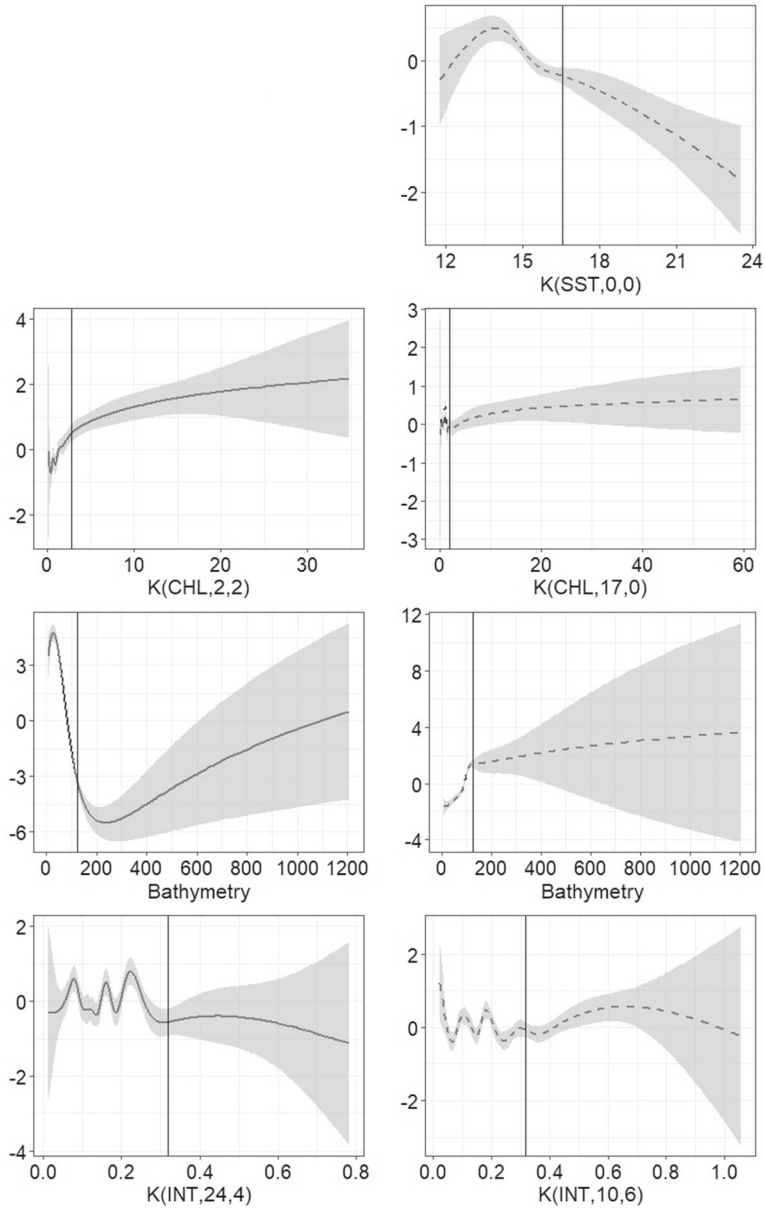
Figure 3. Fixed (environmental) effects for sardine presence (first column) and sardine biomass (second column) derived from biomass index modeling during the PELAGO surveys conducted between 2000 and 2020. Certain covariates are represented by acronyms (SST: sea surface temperature, CHL: chlorophyll-a concentration, INT: intensity of ocean currents). Vertical lines depict the 80% quantiles for each observed covariate, and $K()$ refers to the function defined in Eq. (3).

The obtained posterior predictive distributions for the spatial covariance parameters as well as for the time correlation parameter were relevant (Figure 4 Supplementary Material S.5). The spatial autocorrelation was almost null from about 86 Km, while the mean annual dependence was estimated at approximately 0.18. The unstructured annual effects conveyed an annual variability of 2.38 for presence and 0.41 for the biomass index.

Both processes hinged on an identical spatiotemporal structure, scaled by the parameter k. The estimated mean of k stood at -0.48, indicating that the remaining unexplained variance in the biomass process under presence diverges notably from that in the presence process.

Upon completing the modeling process, we gained the capacity to map the mean of the posterior predictive distributions of presence probability ($E\left[\tilde{Z}_{\mathbf{s}t}\right]$) (Fig. 4) and the median of the posterior predictive distribution of the biomass index ($F_{\tilde{Y}_{\mathbf{s}t}}^{-1}(0.5)$) (Fig. 5) for specific days. In this endeavor, we opted to employ the central days corresponding to the surveys under investigation, while also designating May 10, 2012, as a representative day for the omitted year within the dataset. This choice emerged from its positioning midway between the chosen days for 2011 and 2013. An examination of the resulting maps reveals predominant instances of low presence probability across the majority of the study area, particularly concentrated between Peniche and Lisboa. Remarkably, regions situated between Viana do Castelo and Porto, as well as those stretching from Aveiro to Figueira da Foz, often manifested elevated biomass values.

Figure 6 enables the identification of sustained sardine distribution areas over time. Noteworthy is the observation of a *preferred* occupancy of the species along the coastal expanse between Porto and Aveiro. Conversely, deeper zones along the west coast and along with coastal areas at the south coast, typified *occasional* occupancy of sardine.

During the time span of 2000–2002, a predominance of *favorable* zones contrasted with *unfavorable* counterparts, exhibiting a converse trend during the period of 2018–2020 (Fig. 7). In the first period, most of *favorable* zones were concentrated north of Lisboa, while these zones were more dispersed in the latter period. Despite distinct persistent zones being identified within each period, certain similarities can be discerned. An area located proximate to the coast and situated to the south of the western coastal region was consistently deemed *unfavorable* in both epochs. Upon juxtaposing the two periods, substantial changes become evident in the western extent of the study region, particularly in the northern vicinity of Lisboa. Furthermore, a proliferation of *unfavorable* zones materialized in recent years, with two distinguishable regions emerging—one between Viana do Castelo and Porto, and another between Peniche and Lisboa. The paucity of persistent zones aligns with the marginal value of the time correlation parameter (Figure 4 of Supplementary Material S.5).

4. DISCUSSION

Within the realm of SDM, a central aim lies in devising comprehensive methodologies that encapsulate the intricate interplay between the species and its ecosystem. However, reliance solely on environmental factors might fall short in capturing the nuanced relationship between species and environment, particularly in the case of small pelagic species, known for their heightened sensitivity to environmental fluctuations (Schickele et al. 2021).
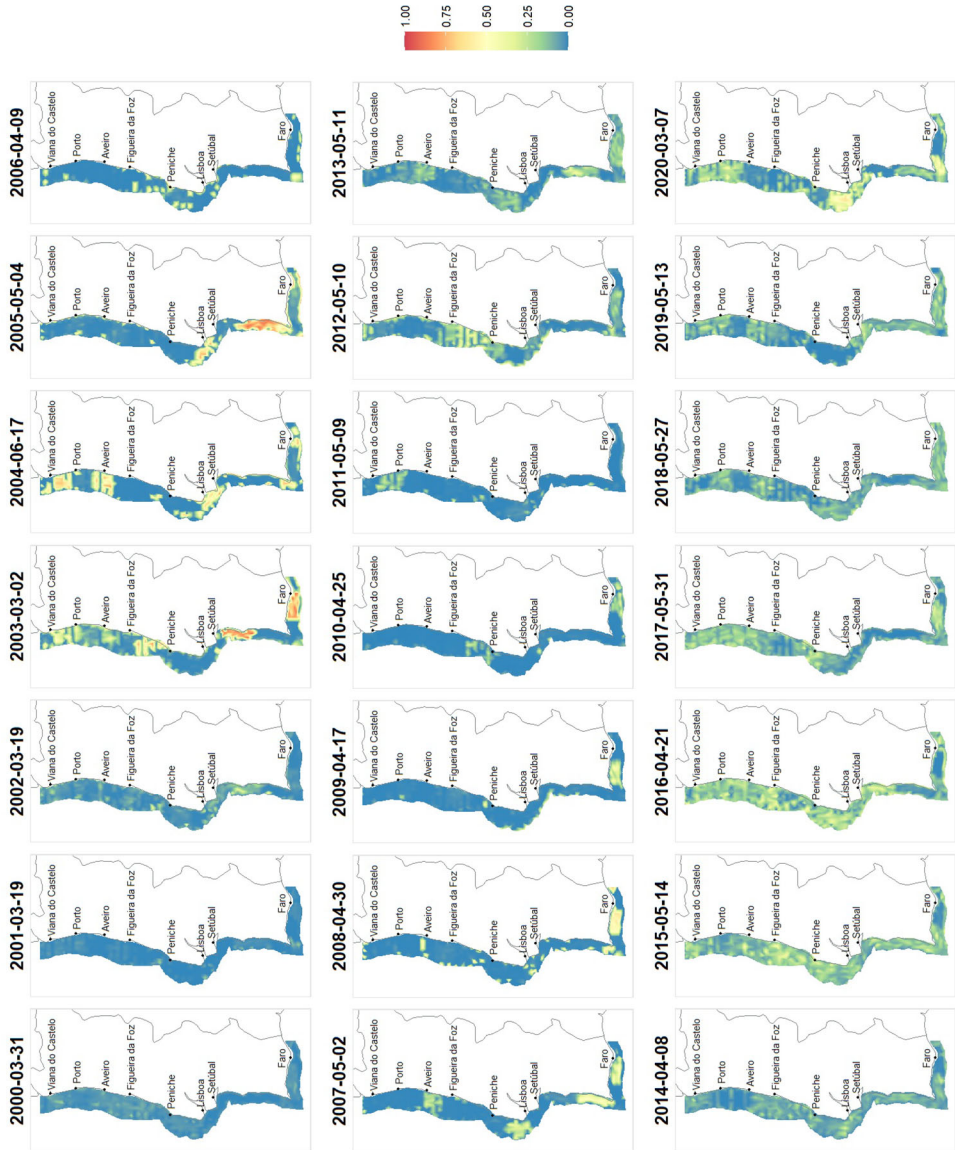
Figure 4. Map depicting the mean of the posterior predictive distribution for the probability of sardine presence ($E[\tilde{Z}_{st}]$) on the representative day of each survey, obtained from the modeling of the biomass index during the PELAGO surveys conducted between 2000 and 2020.

In this study, we put forth a dynamic spatiotemporal SDM, tailored to predict sardine distribution across the waters of the Portuguese shelf and relate the sardine behavior with environmental conditions in a flexible way. We used a hierarchical Bayesian spatiotemporal approach capable of dealing with the complex spatiotemporal dynamics underlying biological phenomena. Environmental information was incorporated referenced in both time and space, in a temporally lagged way, to relate the ecosystem conditions with the process of interest.
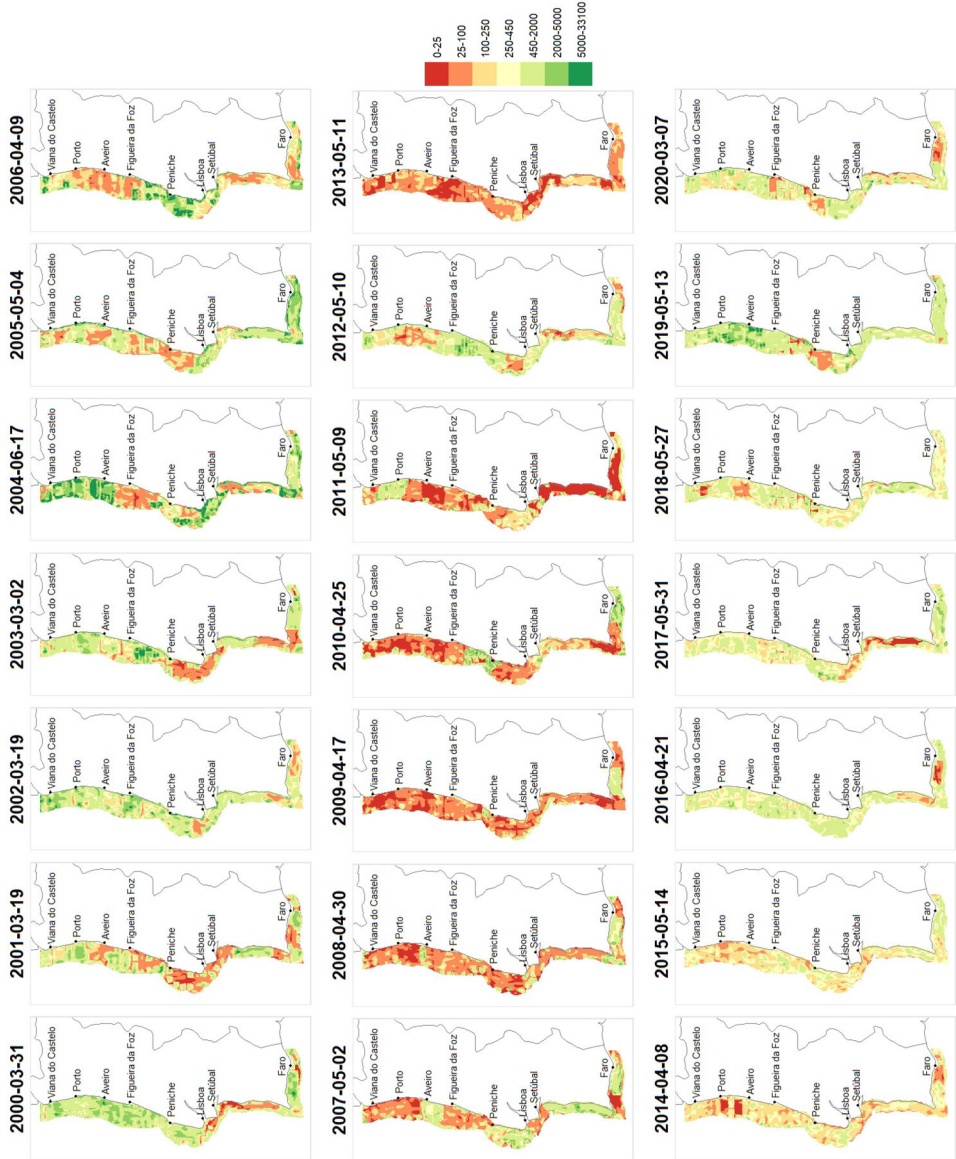
Figure 5. Map depicting the median of the posterior predictive distribution for the biomass index, $F_{\bar{Y}_{st}}^{-1}$, $(m^2\,nm^{-2}$, on the representative day of each survey, obtained from the modeling of the biomass index during the PELAGO surveys conducted between 2000 and 2020.

Paradinas et al. (2017) and Izquierdo et al. (2022) have made substantial contributions to our current comprehension of the two-part model framework within the realm of spatiotemporal SDM. However, a crucial distinction emerges: both models lack the capability to accommodate the effects of environmental conditions with temporal lags. This omission becomes particularly significant due to the pivotal role of temporal lags in shaping envi-
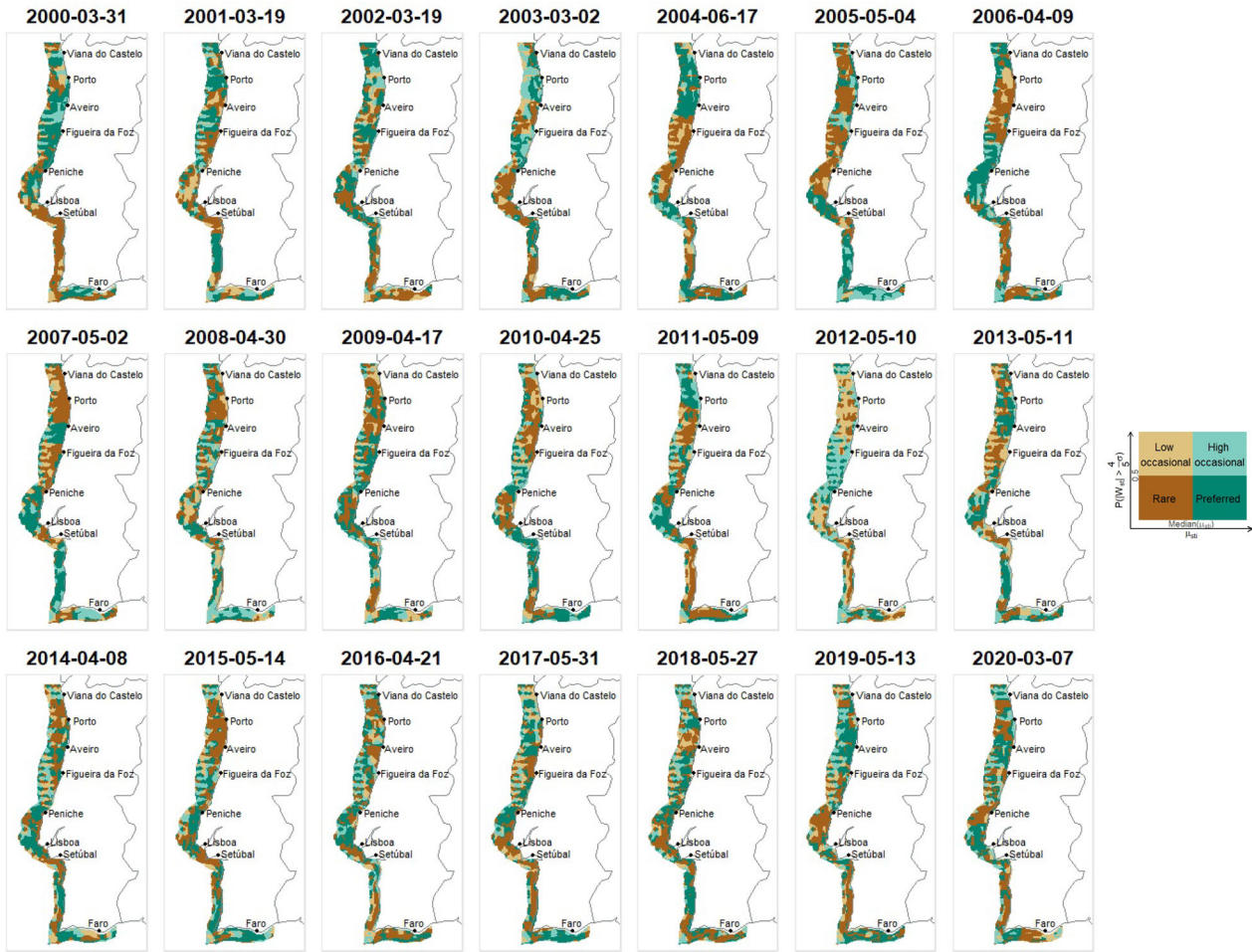
Figure 6. Maps illustrating the sardine occupancy areas categorized as *rare*, *low occasional*, *high occasional*, and *preferred* areas, based on the representative day of each survey.
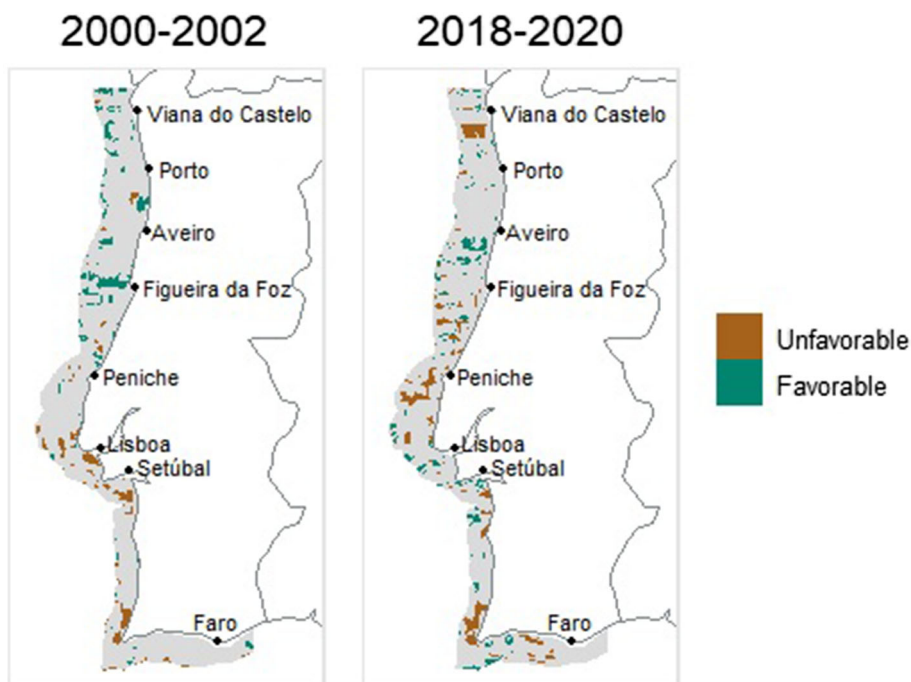
Figure 7. Representation of *favorable* (*preferred* and persistent) and *unfavorable* (*rare* and persistent) sardine zones for two distinct 3-year periods: 2000–2002 (left) and 2018–2020 (right).

ronmental influences. Therefore, we elevated its capacity to encapsulate the intricacies of ecological marine dynamics.

Several studies have employed distributed lag approaches in their analyses. For instance, Gordó-Vilaseca et al. (2021) incorporated covariates with time lags to elucidate the spawning patterns of sardine and anchovy, while Tredennick et al. (2016) utilized time-lagged covariates to investigate the impact of climate change on plant populations. When comparing with these approaches, our proposed methodology gains dynamism, as it employs shorter time lags (measured in days rather than months) and permits a diverse array of time period combinations, each with distinct weights. In essence, our approach affords the flexibility for time lags to encompass intervals, rendering the model a more comprehensive and realistic framework capable of mitigating prediction bias and addressing outliers stemming from extreme events. Heaton and Peng (2012) proposed a generalized linear model that incorporates temporal lags to capture the effects of heat on mortality in various US metropolitan areas. This approach necessitates the consideration of temporal lags over a wide interval, encompassing recent and distant past lags that exhibit decreasing correlation as the lag duration increases. However, recent past lags may not always be pertinent, as illustrated by the effect of chlorophyll-a on fish distribution and abundance. In this case, favorable chlorophyll-a conditions may require several weeks to influence zooplankton abundance and consequently attract fish to the area (Bellido et al. 2008). Pugh et al. (2019) introduced a spatiotemporal model for crop yield, incorporating soil water content as a weighted average observed at different temporal lags. The weighting was determined by a spatiotemporal

kernel utilizing separable effects, which were discretized over a grid of time periods and locations due to computational constraints. However, this approach may prove restrictive when applied to species distribution data, considering that sampling locations evolve over time, and nonlinear effects of covariates need to be accounted for. Our proposed approach circumvents these limitations by accommodating a broader range of lag intervals and treating the spatial domain as continuous. It is important to highlight that applying time lags to static covariates may lack conceptual relevance in certain contexts.

Effective fish distribution monitoring is integral to harmonizing fisheries with marine spatial planning (Janßen et al. 2018). Successful fisheries management hinges on comprehending fish population spatial dynamics. In this context, we introduced a classification of the study region based on model outputs, offering an accessible alternative to complex distribution maps. This classification, simpler to interpret yet equally essential, aids conservation decisions and targeted management strategies in two scenarios. Firstly, it supports the protection of one species while pursuing another, optimizing catch outcomes. The proposed classification scheme identifies potential no-take zones, focusing on areas where protected species tend to gather, thereby minimizing their capture while permitting other species' capture within the fishery. Secondly, our classification framework proves invaluable for conserving juvenile fish populations. By designating areas as preferred or even preferred plus high occasional for juveniles, these regions can be established as no-take zones within the stock's habitat. Such spatial insight is pivotal for creating effective Marine Protected Areas, as underlined by Grüss et al. (2019).

Our classification process centers on predicted biomass estimates and associated uncertainties. To quantify these uncertainties, we examine the latent field probability in relation to a chosen scale of the marginal standard deviation. For our investigation, we adopted a scale of 4/5, albeit determined empirically. However, we advocate for a decision driven by biological insights and expert judgment to ensure credible and meaningful outcomes. Furthermore, exploring the persistence of *rare* and *preferred* zones offers valuable insights into species distribution shifts. Although our delineation is temporally limited due to a short time series, it is conceivable to extend this definition by incorporating multiple consecutive time periods. Such adaptations guided by biological insights or empirical testing offer promising avenues for further exploration.

Within the context of equivalent environmental conditions, a conspicuous divergence in sardine presence emerged, with diminished occurrences along the west coast as compared to the southern counterpart. This divergence could potentially be attributed to a higher incidence of species absences observed in the west comparing to the south during most surveys. The zenith of sardine biomass index was concentrated in the northern vicinity near the coast, emblematic of more prolific zones attributed to the presence of river mouths and estuaries (Monteiro 2017; Zwolinski et al. 2010). The perpetuation of certain ecological niches over time was particularly conspicuous along the west coast. Remarkably, the influence of depth exhibited a nonlinear relationship with both occurrence and biomass, revealing sardine's affinity for shallow waters (up to 85 m). This affinity for shallow depths has been corroborated by other studies, including those conducted in Spanish Mediterranean waters (Bellido et al. 2008), the Northwestern Mediterranean Sea (Saraux et al. 2014), and the western Iberian coast (Zwolinski et al. 2010). While the inclusion of current direction yielded minimal

enhancement to the model, the incorporation of current intensity demonstrated discernible improvement. Conversely, Izquierdo et al. (2022) noted a slight enhancement in their model's performance with the inclusion of current intensity, although it was excluded from their final model. In our point of view, the ocean current effect on sardine biomass is not easy to understand and deserves further investigation, to try to identify the oceanographic processes signaled by current velocity that might be affecting sardine distributions, such as, e.g., the effect of the ocean fronts (Palomera 1992). Sardine biomass index exhibited pronounced elevation in regions characterized by SST ranging between 13.1 and 14.6°C, in line with existing literature on the distribution of sardine eggs (Coombs et al. 2006), sardine larvae (Garrido et al. 2016), and sardine reproduction (Zwolinski et al. 2010). A similar trend in the SST effect on sardine was discerned in the NW Mediterranean Sea (Gordó-Vilaseca et al. 2021), albeit the optimal temperature was estimated around 13.5°C, during a period of lower temperatures. Chlorophyll-a, an indicator of primary productivity, emerged as a pivotal determinant for both processes which resonates with studies by Bacha et al. (2017); Garrido et al. (2017); Gordó-Vilaseca et al. (2021); Izquierdo et al. (2022). The temporal lag between chlorophyll-a and sardine biomass potentially mirrors the lag between phytoplankton and zooplankton blooms—a primary source of sustenance for sardines—spanning from larvae to mature individuals (Garrido et al. 2008).

The spatiotemporal structure contributed to a modest annual dependency in the sardine biomass index. Santos et al. (2012) discerned that sardine recruitment in a specific year is contingent upon the recruitment observed in the preceding year. On the other hand, Borges et al. (2003) conducted a time series analysis of sardine catch data spanning from 1946 to 1990, revealing that sardine catch is influenced by catches recorded up to 14 years earlier. Indeed, the last decades have been characterized by more recurrent and impactful environmental changes (Tang 2019) which makes it difficult to detect strong annual dependencies.

Studying the dependency between the fish abundance and environmental variables augments our understanding of abundance dynamics, facilitates habitat delineation, and enhances our capacity to prognosticate marine species trends. Our study underscored the utility of Bayesian spatiotemporal modeling in elucidating species distribution patterns. While our investigation focused on a specific species inhabiting the Portuguese shelf, the proposed approach can be readily extrapolated to diverse species sets and geographical domains. We advocate for the exploration of various time lag combinations for covariates, guided by prior knowledge (e.g., ecological insights), given the inherent computational challenges of Bayesian methodologies, especially in higher-dimensional datasets.

REFERENCES

Adde A, Darveau M, Barker N, Cumming S (2020) Predicting spatiotemporal abundance of breeding waterfowl across Canada: a Bayesian hierarchical modelling approach. Biodiversity Res 26(10):1248–1263

Alvarez P, Chifflet M (2012) The fate of eggs and larvae of three pelagic species, mackerel (Scomber scombrus), horse mackerel (Trachurus trachurus) and sardine (Sardina pilchardus) in relation to prevailing currents in the Bay of Biscay: Could they affect larval survival? Sci Mar 76(3):573–586

Attorre F, Alfò M, Sanctis MD, Francesconi F, Valenti R, Vitale M, Bruno F (2011) Evaluating the effects of climate change on tree species abundance and distribution in the Italian peninsula. Appl Veg Sci 14(2):242–255

Bacha M, Jeyid MA, Vantrepotte V, Dessailly D, Amara R (2017) Environmental effects on the spatio-temporal patterns of abundance and distribution of Sardina pilchardus and sardinella off the Mauritanian coast (North-West Africa). Fish Oceanogr 26(3):282–298

Bakka H, Vanhatalo J, Illian JB, Simpson D, Rue H (2019) Non-stationary gaussian models with physical barriers. Spatial Stat 29:268–288

Bellido J, Brown A, Valavanis V, Giráldez A, Pierce G, Iglesias M, Palialexis A (2008) Identifying Essential Fish Habitat for small pelagic species in Spanish Mediterranean waters. Hydrobiologia 612:171–184

Borges M, Santos AM, Crato N, Mendes H, Mota B (2003) Sardine regime shifts off Portugal: a time series analysis of catches and wind conditions. Sci Mar 67:235–244

Cabrero Á, González-Nuevo G, Gago J, Cabanas JM (2019) Study of sardine (Sardina pilchardus) regime shifts in the Iberian Atlantic shelf waters. Fish Oceanogr 28(3):305–316

Checkley DM, Asch RG, Rykaczewski RR (2017) Climate, Anchovy, and Sardine. Ann Rev Mar Sci 9(1):469–493

Coombs S, Smyth T, Conway D, Halliday N, Bernal M, Stratoudakis Y, Álvarez P (2006) Spawning season and temperature relationships for sardine (Sardina pilchardus) in the eastern North Atlantic. J Mar Biol Assoc 86:1245–1252

Dambrine C, Woillez M, Huret M, de Pontual H (2021) Characterising Essential Fish Habitat using spatio-temporal analysis of fishery data: a case study of the European seabass spawning areas. Fisheries Oceanography

Doray M, Boyra G, van der Kooij J (2021) ICES Survey Protocols - Manual for acoustic surveys coordinated under In: ICES working group on acoustic and egg surveys for small pelagic fish (WGACEGG) (Version 1). ICES Techniques in Marine Environmental Science (TIMES), 64:100pp

Duin (1976) On the choice of smoothing parameters for parzen estimators of probability density functions. IEEE Trans Comput C-25(11):1175–1179

García-Lafuente J, Delgado J, Criado-Aldeanueva F, Bruno M, del Río J, Miguel Vargas J (2006) Water mass circulation on the continental shelf of the Gulf of Cádiz. In: Deep sea research part II: topical studies in oceanography, 53(11):1182–1197. The Gulf of Cadiz Oceanography: A Multidisciplinary View

Garrido S, Ben-Hamadou R, Oliveira P, Cunha M, Chícharo M, van der Lingen C (2008) Diet and feeding intensity of sardine Sardina pilchardus: correlation with satellite-derived chlorophyll data. Mar Ecol Prog Ser 354:245–256

Garrido S, Cristóvão A, Caldeira C, Ben-Hamadou R, Baylina N, Batista H, Saiz E, Peck M, Re P, Santos AM (2016) Effect of temperature on the growth, survival, development and foraging behaviour of *Sardina pilchardus* larvae. Mar Ecol Progress Ser, 559

Garrido S, Silva A, Marques V, Figueiredo I, Bryère P, Mangin A, Santos AMP (2017) Temperature and food-mediated variability of European Atlantic sardine recruitment. Prog Oceanogr 159:267–275

Gordó-Vilaseca C, Pennino MG, Albo-Puigserver M, Wolff M, Coll M (2021) Modelling the spatial distribution of Sardina pilchardus and Engraulis encrasicolus spawning habitat in the NW Mediterranean Sea. Mar Environ Res 169:105381

Grüss A, Biggs CR, Heyman WD, Erisman B (2019) Protecting juveniles, spawners or both: a practical statistical modelling approach for the design of marine protected areas. J Appl Ecol 56(10):2328–2339

Guélat J, Kéry M (2018) Effects of spatial autocorrelation and imperfect detection on species distribution models. Methods Ecol Evol 9:1614–1625

Habbema JDF, Hermans J, Vandenbroek K (1974) A stepwise discriminant analysis program using density estimation. In Bruckmann, G (eds) COMPSTAT 1974:101–110

Heaton MJ, Peng RD (2012) Flexible distributed lag models using random functions with application to estimating mortality displacement from heat-related deaths. J Agric Biol Environ Stat 17:313–331

Hefley TJ, Hooten MB (2016) Hierarchical species distribution models. Curr Landscape Ecol Rep 1:87–97

Hoegh-Guldberg O, Bruno JF (2010) The impact of climate change on the World's marine ecosystems. Science 328(5985):1523–1528

ICES (2017) Report of the working group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). Technical report, ICES CM 2017/ACOM:17, Bilbao, Spain. 640 pp

Izquierdo F, Menezes R, Wise L, Teles-Machado A, Garrido S (2022) Bayesian spatio-temporal CPUE standardization: case study of European sardine (*Sardina pilchardus*) along the western coast of Portugal. Fish Manag Ecol

Janßen H, Bastardie F, Eero M, Hamon KG, Hinrichsen H-H, Marchal P, Nielsen JR, Le Pape O, Schulze T, Simons S, Teal LR, Tidd A (2018) Integration of fisheries into marine spatial planning: Quo vadis? Estuar Coast Shelf Sci 201:105–113

Jghab A, Vargas-Yañez M, Reul A, Garcia-Martínez M, Hidalgo M, Moya F, Bernal M, Ben Omar M, Benchoucha S, Lamtai A (2019) The influence of environmental factors and hydrodynamics on sardine (Sardina pilchardus, Walbaum 1792) abundance in the southern Alboran Sea. J Mar Syst 191:51–63

Joseph LN, Elkin C, Martin TG, Possingham HP (2009) Modeling abundance using N-mixture models: the importance of considering ecological mechanisms. Ecol Appl 19(3):631–642

Kneib T, Müller J, Hothorn T (2008) Spatial smoothing techniques for the assessment of habitat suitability. Environ Ecol Stat 15:343–364

Krainski E, Gómez-Rubio V, Bakka H, Lenzi A, Castro-Camilo D, Simpson D, Lindgren F, Rue H (2019) Advanced spatial modeling with stochastic partial differential equations using R and INLA. Chapman & Hall/CRC Press, Boca Raton, FL

Kühn I (2007) Incorporating spatial autocorrelation may invert observed patterns. Diversity Distrib 13(1)

Leathwick JR, Rowe DK, Richardson J, Elith J, Hastie TJ (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. Freshw Biol 50:2034–2052

Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian mMarkov random fields: the stochastic partial differential equation approach. Stat Methodol 73(4):423–498

Martínez-Minaya J, Cameletti M, Conesa DV, Pennino MG (2018) Species distribution modeling: a statistical review with focus in spatio-temporal issues. Stoch Environ Res Risk Assess 32:3227–3244

McCauley DJ, Pinsky ML, Palumbi SR, Estes JA, Joyce FH, Warner RR (2015) Marine defaunation: animal loss in the global ocean. Science 347(6219):1255641

Mendes C, Costa-Dias S, Teixeira C, Afonso A, Bordalo A (2019) An overview of small-scale fisheries in the Northern Portuguese coast. Front Mar Sci, 6

Miller JA (2012) Species distribution models: spatial autocorrelation and non-stationarity. Prog Phys Geogr: Earth Environ 36(5):681–692

Monteiro PV (2017) The purse seine fishing of sardine in Portuguese Waters: a difficult compromise between fish stock sustainability and fishing effort. Rev Fisheries Sci Aquaculture 25(3):218–229

Palomera I (1992) Spawning of anchovy Engraulis encrasicolus in the Northwestern Mediterranean relative to hydrographie features in the region. Mar Ecol Prog Ser 79(3):215–223

Paradinas I, Conesa D, López-Quílez A, Bellido JM (2017) Spatio-temporal model structures with shared components for semi-continuous species distribution modelling. Spatial Stat 22:434–450. Spatio-temporal Statistical Methods in Environmental and Biometrical Problems

Peliz A, Rosa TL, Santos AMP, Pissarra JL (2002) Fronts, jets, and counter-flows in the Western Iberian upwelling system. J Mar Syst 35(1):61–77

Pennino MG, Coll M, Albo-Puigserver M, Fernández-Corredor E, Steenbeek J, Giráldez A, González M, Esteban A, Bellido JM (2020) Current and future influence of environmental factors on small pelagic fish distributions in the Northwestern Mediterranean Sea. Front Mar Sci 7:622

Pires APF, Srivastava DS, Marino NAC, Macdonald AAM, Figueiredo-Barros MP, Farjalla VF (2018) Interactive effects of climate change and biodiversity loss on ecosystem functioning. Ecology 99(5):1203–1213

Pugh S, Heaton M, Svedin J, Hansen N (2019) Spatiotemporal lagged models for variable rate irrigation in agriculture. J Agric Biol Environ Stat, 24

Reiss H, Cunze S, König K, Neumann H, Kröncke I (2011) Species distribution modelling of marine benthos: a North Sea case study. Mar Ecol Prog Ser 442:71–86

Relvas P, Barton E, Dubert J, Oliveira PB, Álvaro Peliz da Silva J, Santos AMP (2007) Physical oceanography of the western Iberia ecosystem: latest views and challenges. Progr Oceanogr, 74(2):149–173. Ecological Functioning of the Iberian Seas: A synthesis of GLOBEC Research in Spain and Portugal

Rodríguez-Climent S, Angélico MM, Marques V, Oliveira P, Wise L, Silva A (2017) Essential habitat for sardine juveniles in Iberian waters. Sci Mar 81(3):351–360

Roos M, Held L (2011) Sensitivity analysis in Bayesian generalized linear mixed models for binary data. Bayesian Anal 6(2):259–278

Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian Models by using integrated nested Laplace approximations. J R Stat Soc B 71:319–392

Santos MB, González-Quirós R, Riveiro I, Cabanas JM, Porteiro C, Pierce GJ (2012) Cycles, trends, and residual variation in the Iberian sardine (Sardina pilchardus) recruitment series and their relationship with the environment. ICES J Mar Sci 69(5):739–750

Saraux C, Fromentin J-M, Bigot J-L, Bourdeix J-H, Morfin M, Roos D, Van Beveren E, Bez N (2014) Spatial structure and distribution of small pelagic fish in the northwestern mediterranean sea. PLoS ONE 9(11):e111211

Schickele A, Goberville E, Leroy B, Beaugrand G, Hattab T, Francour P, Raybaud V (2021) European small pelagic fish distribution under global change scenarios. Fish Fish 22(1):212–225

Sheather SJ (2004) Density estimation. Stat Sci 19(4):588–597

Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH (2017) Penalising model component complexity: a principled, practical approach to constructing priors. Stat Sci 32(1):1–28

Spezia L, Cooksley S, Brewer M, Donnelly D, Tree A (2014) Modelling species abundance in a river by negative binomial hidden markov models. Comput Stat Data Anal 71:599–614

Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD (2002) Bayesian measures of model complexity and fit. J R Stat Soc Ser B Sat Methodol 64(4):583–639

Tang KHD (2019) Climate change in Malaysia: trends, contributors, impacts, mitigation and adaptations. Sci Total Environ 650:1858–1871

Teles-Machado A, Peliz Álvaro, McWilliams JC, Couvelard X, Ambar I (2016) Circulation on the Northwestern Iberian Margin: vertical structure and seasonality of the alongshore flows. Prog Oceanogr 140:134–153

Tredennick AT, Hooten MB, Aldridge CL, Homer CG, Kleinhesselink AR, Adler PB (2016) Forecasting climate change impacts on plant populations over large spatial extents. Ecosphere 7(10):e01525

Voulgaridou P, Stergiou KI (2003) Trends in various biological parameters of the European sardine, Sardina pilchardus (Walbaum, 1792), in the Eastern Mediterranean Sea. Sci Mar 67(S1):269–280

Whitehead PJP (1985) *FAO species catalogue. V. 7: Clupeoid fishes of the world (suborder Clupeoidei). An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, shads, anchovies and wolf-herrings. Part 1. Chirocentridae, Clupeidae and Pristigasteridae*. FAO

Zwolinski JP, Oliveira PB, Quintino V, Stratoudakis Y (2010) Sardine potential habitat and environmental forcing off western Portugal. ICES J Mar Sci 67(8):1553–1564