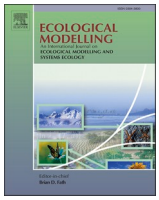
Ecological Modelling 490 (2024) 110632

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Pan-Atlantic 3D distribution model incorporating water column for commercial fish

Mireia Valle [a,*], Eduardo Ramírez-Romero [b], Leire Ibaibarriaga [a], Leire Citores [a], Jose A. Fernandes-Salvador [a], Guillem Chust [a]

[a] AZTI, Marine Research, Basque Research and Technology Alliance (BRTA), Txatxarramendi ugartea z/g, 48395 Sukarrieta, Spain
[b] Institute of Marine Sciences of Andalusia (ICMAN – CSIC), C/ República Saharaui, 4, 11519 Puerto Real, Cádiz, Spain

ARTICLE INFO

Keywords:
Species distribution models3
Ecological niche
Macroecology
Depth
Sc-gam
Gam-niche

ABSTRACT

Fisheries have a crucial contribution, with animal protein supply and economic income, to the subsistence and blue economy of several human societies of the Atlantic Ocean, the second largest water body in the planet. However, an accurate distribution of commercial fish across the Atlantic and through the water column is still unknown. The wide use of Species Distribution Models (SDMs) for marine fish mapping generally faces two shortcomings: (i) ignoring the vertical dimension of the ocean; and (ii) ignoring the ecological niche theory in the model fitting. Our aim is to develop 3D habitat models of the main commercial fishes across the Atlantic Ocean, accounting for 67 % of the total biomass catches, to provide an enhanced spatial representation of the environmental niche of the fish species. In particular, here we (1) explore the macroecological patterns testing if latitudinal-vertical distribution of main commercial fish species follows the isothermal distribution across the Atlantic ocean; (2) apply a novel 3D modelling approach incorporating depth dimension into the environmental data and based exclusively on public species occurrence data; (3) use Shape-Constrained Generalized Additive Models (SC-GAMs) to build SDMs in accordance with the ecological niche theory (GAM-NICHE model), avoiding potential model overfitting and hence allowing automatic model selection; and (4) estimate potential fish catch biomass in the 3D space based on the species probability of occurrence. Our results indicated that latitudinal-vertical distribution follows the prevailing isothermal distribution in the ocean, confirming that an accurate representation of stock distributions needs 3D modelling and incorporate explicitly depth dimension into the environmental data. The species response curves to 3D environmental gradients for the 30 main commercial fish species of the Atlantic yielded very good model accuracy performance (78–98 %). The developed 3D models of fish occurrence probability have the capability to be improved with the updates of new data for data-poor species, and to be projected under climate change scenarios. The obtained 3D maps conform useful and new knowledge that may help policy makers to balance the need for environmental protection with sustainable marine resource exploitation of the Atlantic Ocean.

1. Introduction

The Atlantic Ocean, the second largest water body in the planet, covers all the oceanic ecosystems, from oligotrophic subtropical gyres to eastern boundary upwellings and subpolar waters in both hemispheres (Longhurst, 2007), and provides essential ecosystems services (Magalhães Filho et al., 2022), from highly biodiverse ecosystems such as coral reefs and mangroves to very productive and highly exploited North Sea and European shelves (Emeis et al., 2015). Atlantic fisheries have a crucial contribution, with animal protein supply and economic income, to the subsistence and blue economy of several human societies (FAO 2022). Currently, there is a pressing need of sustainably managing marine resources and preserving associated biodiversity due to the complex array of threats derived from anthropogenic activities (e.g., overfishing, pollution, ocean acidification, deoxygenation, invasive species) (Halpern et al., 2019). Climate change is putting an extra pressure to marine resources and traditional fisheries management. Climate change can modify the species abundance and distribution (Cooley et al., 2022; Pecl et al., 2017; Dell'Apa et al., 2023), and global ensemble projections have revealed trophic amplification of ocean

* Corresponding author.
E-mail address: mvalle@azti.es (M. Valle).

https://doi.org/10.1016/j.ecolmodel.2024.110632
Received 16 May 2023; Received in revised form 24 January 2024; Accepted 26 January 2024
Available online 14 February 2024

animal biomass declines with climate change (Lotze et al., 2019; Thompson et al., 2023; Tittensor et al., 2021). Assessing the relationship between fisheries and the marine environment, including species and habitat, emerge as key to reduce uncertainty on current fisheries (Pauly and Zeller, 2016; Worm et al., 2006) and future stocks under climate change scenarios (du Pontavice et al., 2023; Erauskin-Extramiana et al., 2023; Gaines et al., 2018; Payne et al., 2021; Predragovic et al., 2023).

Marine ecosystems cannot be accurately understood without regarding the vertical domain of the ocean (Sayre et al., 2017). For instance, thermal patterns in the Atlantic ocean involve very particular features such as the gradient in surface temperature of 10 °C that could extend 2000 km in latitude in open ocean (Aiken et al., 2000) and 100 km in continental shelves (Morey and O'Brien, 2002), whilst typically extends 200 m in the vertical, due to markedly stratified waters and associated thermocline. Nevertheless, deep basins of the whole ocean are filled by cold and sub-polar water masses without marked horizontal gradients, due to thermohaline circulation. These physical patterns shape species biogeography. One interesting pattern known since Darwin times, but seldom tested, is the "equatorial submergence" (Close et al., 2006) which establishes that the latitudinal-vertical distribution of some marine species follows the isothermal distribution in the ocean, from deep and intermediate waters at low latitudes towards shallow layers at mid-high latitudes (Trubovitz et al., 2020). As a response to temperature-respiration constraints driving the species thermal niche (Pauly et al., 1998), individuals of a species can be found at different depth ranges depending on the latitude. In some cold-water affinity species, transequatorial migration through deep layers maintain both sub-polar Atlantic populations connected (Møller et al., 2003), while in other species, tropical waters have disconnected populations promoting allopatric speciation (termed anti-tropical species) (Ludt, 2021).

Species distribution models (SDMs) are widely used as a tool for understanding species spatial ecology (Guisan et al., 2013; Robinson et al., 2017). They link species occurrence or abundance with environmental features of the location, via statistical modelling (Elith and Leathwick, 2009). SDMs for commercial fishes have been developed to capture the spatiotemporal patterns and project distribution shifts under climate change (Bruge et al., 2016; Erauskin-Extramiana et al., 2019; Erauskin-Extramiana et al., 2020; Erauskin-Extramiana et al., 2019; Hobday, 2010; Maynou et al., 2020; Schickele et al., 2021). However, these approaches usually do not address the water column explicitly. Surface environmental variables are usually paired with deep-water occurrence records despite the disparity of environmental conditions at the surface and at depth (Duffy and Chown, 2017), which is a consequence of data limitation related to species occurrences that are not always vertically informed (in most of cases) or/and the availability of depth-specific environmental data. Ignoring depth in SDMs could lead to misleading outcomes. Therefore, some authors have proposed to include proxies of the real 3D seascape with e.g., water column physical features (Brodie et al., 2018), and use a 2.5D approach with multiple 2D layers to approximate 3D systems (Duffy and Chown, 2017). Taking advantage of the rising biological and environmental available data resources, new efforts should try to fully implement 3D SDMs (Bentlage et al., 2013; Schwing, 2023).

Another general problem of the most widely-applied species modelling techniques is that species response curves are fit statistically without any assumption or restriction, which sometimes do not respect the ecological niche theory (Hutchinson, 1957). According to ecological niche theory, species distributions should provide unimodal relationships with respect to environmental gradients (Hutchinson, 1957). Key stages of the life cycle (feeding, growth, and reproduction) are determined by the physiological range of tolerance of the species and are affected by unfavourable environmental conditions, resulting in lower presence of the species (Austin and Heyligers, 1989; Helaouët and Beaugrand, 2009). In this context, it has been claimed that species distribution models need a stronger theoretical background (Elith and Leathwick, 2009). Excessively flexible SDM algorithms and parametrizations such as machine learning can lead to overfitted models where resulting patterns can be spurious and affected by noise, and predictions based on such models can be biased and unreliable (Burnham and Anderson, 2002). SDMs should therefore consider theoretical background such as the ecological niche theory and pursue the unimodality of the response curve with respect to environmental gradients (Citores et al., 2020).

The overall aim of this work is to improve the understanding of the fish macroecological patterns and develop 3D habitat models of the main commercial fish caught in the Atlantic Ocean. We first explore the macroecological patterns testing if latitudinal-vertical distribution of commercial fish species follows the isothermal distribution across the Atlantic Ocean. Then, based exclusively on public species occurrence data, we apply a novel 3D modelling approach incorporating depth dimension into the environmental data. In particular, to build SDMs in accordance with the ecological niche theory (Hutchinson, 1957), we use Shape-Constrained Generalized Additive Models (SC-GAMs) (Pya and Wood, 2015) that impose unimodality of the response curve with respect to environmental gradients (Citores et al., 2020) which in turn avoids potentially model overfitting and hence allows automatic model selection and fitting. Finally, we estimate potential catch biomass in the 3-dimensional space based on the modelled species probability of occurrence. Our results may provide valuable information to support fisheries managers and policy makers to balance the need for environmental protection with sustainable marine resource exploitation in the Atlantic Ocean.

2. Material and methods

The novel 3D modelling approach incorporating water column data followed five main steps (Fig. 1): 1) selection of main commercial fish species and compilation of 3D occurrence data; 2) selection and set up of 3D environmental data; 3) building and cross-validation of species distribution 3D models under the ecological niche theory (see https://gam-niche.azti.es/ developed by Valle et al. (Valle et al., 2023) for detailed methodology); 4) prediction of suitable habitat along the water column; and 5) estimation of potential catch biomass within the 3-dimensional space. All analysis were conducted using R (R Core Team 2023).

2.1. Study area

The Atlantic Ocean is the second largest water body and has the largest drainage area in the planet, including semi-enclosed areas such as the Baltic and the Mediterranean Sea (Longhurst, 2007). The largest ocean-scale oceanographic features in this ocean are wind-driven currents and the Atlantic Meridional Overturning Circulation, which is the ocean conveyor that distributes heat and energy and regulates the climate. These currents are directly linked and controlled by basin-scale forcing and dynamics at inter-annual to decadal timescales (Biastoch et al., 2021). Main Atlantic oceanographic gyres shape the wide variety of productivity areas, from low-productive subtropical areas to high-productive subpolar areas (Longhurst, 2007). Overall, the large differences in seafloor depth (shelves, slopes, abyssal plains), coastline topography, plankton productivity, coastal vs oceanic areas, and temperature gradient offer diverse environments to the large variety of fish species living in the Atlantic (Castilho et al., 2013; Ormond et al., 1997).

We established our study area selecting the Food and Agriculture Organization (FAO) areas that correspond to the Atlantic Ocean (21, 27, 31, 34, 37, 41, 47, 48), excluding Black Sea subarea (37.4) (Fig. 2).

2.2. Fish catches data

We ranked main commercial fish species for the Atlantic Ocean in terms of catches (tons/year), based on spatially allocated catch sourced from Global fisheries landings (V4) database (Watson, 2020), in order to

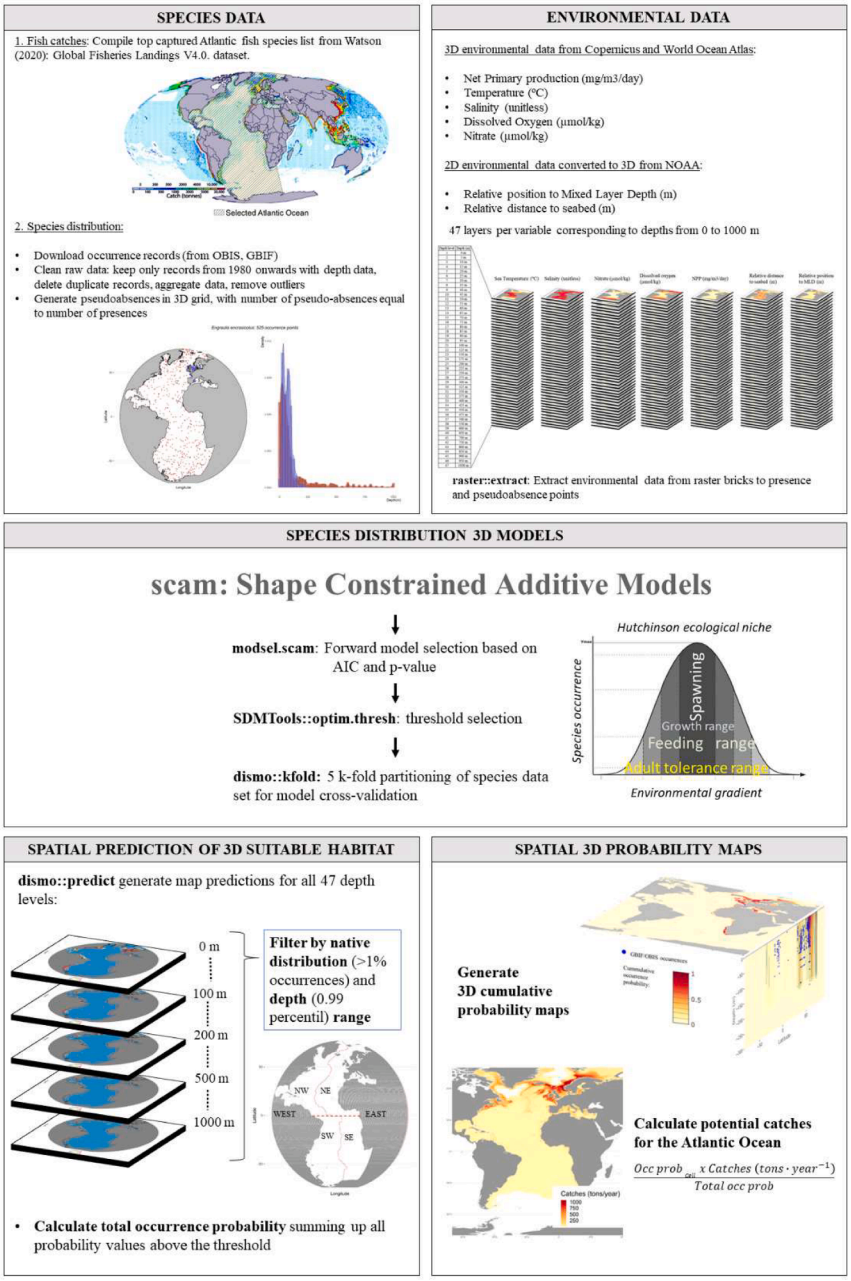


Fig. 1. Flow chart showing the novel 3D modelling approach applied to main Atlantic commercial fish.

select those species to be modelled, and in addition, we included 5 species of tunas regarding their economic interest. Watson (Watson, 2020) database brings a high spatial resolution (cell 30 × 30′) and an accurate taxonomy of species of marine origin, comprising and integrating several sources; furthermore, catches came from industrial and non-industrial fleets, also including illegal and unreported catches. We extracted the yearly catches for the period 2010–2015 and the associated taxa information from the cells falling within our study area. Subsequently, yearly catches from all cells were summed up and mean catches over 2010–2015 period were calculated and ranked only including fish species.

2.3. Species occurrences data

We retrieved species occurrence data for the selected 25 commercial fish and the 5 tuna species (30 species in total) from global public databases. To do so, we first, validated species scientific name from our species list using "rfishbase" R package (Boettiger et al., 2012). Available occurrence data with associated information (depth and date of collection) were retrieved from global open-access datasets Ocean Biodiversity Information System (OBIS 29 July 2022) and Global Biodiversity Information Facility (http://www.gbif.org) from 1980 to July 2022 using online queries via the R packages "robis" (Provoost, 2023) and "rgbif" (Chamberlain et al., 2023). Subsequently, we removed duplicated records based on associated coordinates and date of collection and we discarded records with missing sampling depth information. To avoid outliers, occurrences deeper than 1000 m were also removed due to their scarce presence. Additionally, we also removed records that were flagged as outliers in geographic space according to the distance method (minimum absolute distance = 1000 km) using the R package "CoordinateCleaner" (Zizka et al., 2019). By applying this method, we identified and removed out records if the minimum distance to the next record of the species was larger than 1000 km. Public data repositories often present geographic outliers due to erroneous coordinates, for

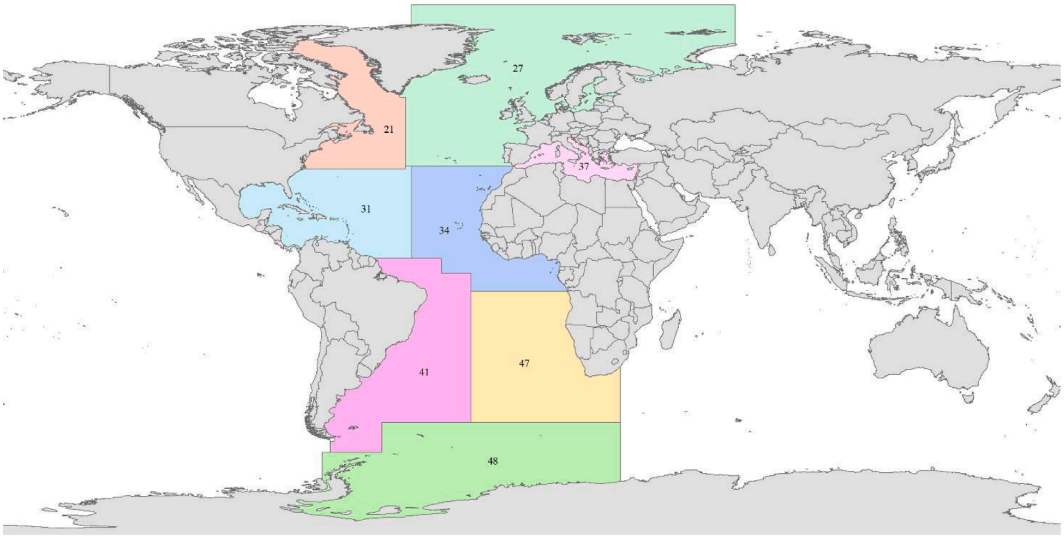


Fig. 2. Selected study area based on FAO areas.

example due to data entry errors or imprecise geo-references, thus applying this standardized procedure to avoid outliers in the geographic space is highly recommended. See dataset key details of the records we kept in the derived dataset (GBIF, 2024).

Finally, we created a 3D grid within the study area of 1/4° resolution (692 × 666 cells) and with 47 standard depth levels from the surface to a maximum depth of 1000 m, to match the environmental grid (see Section 2.4) with same spatial resolution and depth levels used at World Ocean Atlas, WOA 2018 (Locarnini et al., 2018) and then aggregated the occurrences keeping only one record when several occurrences fell in the same grid cell. Depths values associated with each standard level in our 3D grid are distributed as follows: they increase every 5 m until 100 m depth, then they increase every 25 m until 500 m depth, and after 500 m depth the values increase every 50 m. Thus the 47 depth levels values of our 3D grid are (in meters): 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000.

2.4. Environmental data

We selected a set of potential environmental drivers aiming to capture the basic niche of the species, with a particular focus on the vertical patterns. Two types of environmental data were collated (Table 1): three-dimensional variables (3D), which were depth-specific; and two-dimensional (2D), that were transformed into three-dimensional (see below).

Three-dimensional variables consisted of (1) temperature, (2) salinity, (3) dissolved oxygen, (4) nitrate, and (5) Net Primary Production (NPP). Selected environmental variables are known to capture the main ecological niche of the fish species in terms of thermal and ecophysiological boundaries (salinity and oxygen), and are proxies of indirect food availability (i.e., NPP and nitrate as limiting nutrient for phytoplankton growth) and of the trophic status of the ecosystem, oligotrophic vs productive waters (NPP) (Arrizabalaga et al., 2015; Kesner-Reyes et al., 2020). The World Ocean Atlas 2018 (WOA, https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/) was the main data source of depth-specific variables, specifically we used "Objectively analyzed climatologies" which consisted of interpolated mean fields for oceanographic variables at standard depths (Locarnini et al., 2018). For temperature and salinity variables, we downloaded the data at 1/4° resolution and selected the 1981–2010 decadal period, whilst for dissolved oxygen and nitrate, data were available at 1° for 1900 to 2020. All variables presented 102 vertical standard depth levels, from surface to 5500 m depth, and were downloaded in *netcdf* format. NPP data were downloaded from the Global Ocean biogeochemistry hindcast, an ocean product available at Marine Copernicus (https://resources.marine.copernicus.eu/products) for the time period 1993–2010 at 1/4° and on 75 vertical levels. Depth levels in NPP variable were not equally distributed as depth levels from our 3D grid, thus we needed to apply a linear interpolation of NPP values across depth levels to match the standard depths. Then NPP values were transformed to logarithmic scale as follows: $\log(NPP + 1)$.

Two-dimensional variables included: (6) seabed depth and (7) Mixed Layer Depth (MLD). Seabed depth is a proxy of fish habitat preference (neritic vs oceanic species), while MLD is a proxy of thermal water stratification (Arrizabalaga et al., 2015; Kesner-Reyes et al., 2020). Seabed depth was extracted from NOAA server (Amante and Eakins, 2009) using *getNOAA.bathy* function from "marmap" R package (Pante and Simon-Bouhet, 2013); and MLD data was downloaded in *netcdf*

Table 1
Summary of the environmental variables, type of data, initial horizontal resolution, period of the climatology and source.

	Variable	Type of data	Initial Horizontal Resolution	Period	Source
1	**Temperature (°C)**	3D	0.25°	1981–2010	World Ocean Atlas 2018 (Locarnini et al., 2018)
2	**Salinity (unitless)**	3D	0.25°	1981–2010	World Ocean Atlas 2018 (Zweng et al., 2018)
3	**Dissolved Oxygen (μmol/kg)**	3D	1°	1900–2020	World Ocean Atlas 2018 (Garcia et al., 2018)
4	**Nitrate (μmol/kg)**	3D	1°	1900–2020	World Ocean Atlas 2018 (Garcia et al., 2018)
5	**Net Primary production (mg/m^3/day)**	3D	0.25°	1993–2010	Marine Copernicus. GLOBAL_MULTIYEAR_BGC_001_029
6	**Mixed Layer Depth (m)**	2D	0.25	1981–2010	World Ocean Atlas 2018 (Boyer et al., 2018)
7	**Seabed depth (m)**	2D	0.25°	–	NOAA- ETOPO1 ("marmap" R package) (Pante and Simon-Bouhet, 2013)

format from WOA for 1981–2010 decadal period and integrated objectively analysed mean fields for ocean mixed layer thickness at same standard depth levels as WOA oceanographic variables. To transform the 2D variables into 3D, we first created a data frame using the 3D grid as a template with 47 depth levels. Seabed depth was converted to depth-specific *relative distance to seabed* variable by subtracting the meters at which each depth level is located (Depth value) to the seabed depth value at each cell (Eq. (1)). NA (not available or missing value) was assigned to those cells with depth values higher than seabed depth.

$$\textit{Relative distance to seabed}_{\textit{Depth level}}\,(m) = \textit{Seabed depth value}\,(m) - \textit{Depth value}_{\textit{Depth level}}(m) \quad (1)$$

Following a similar procedure, MLD variable data was converted to depth-specific *relative distance to MLD* variable by subtracting the meters at which each depth level is located (Depth value) to the MLD value at each cell (Eq. (2)).

$$\textit{Relative Position to MLD}_{\textit{Depth level}}\,(m) = \textit{MLD}\,(m) - \textit{Depth value}_{\textit{Depth level}}(m) \quad (2)$$

Occurrences above MLD will present positive values of this variable indicating preference to the mixed zone where warmer, lower density and mixed upper waters are found; whilst occurrences below MLD will present negative values of relative distance to MLD indicating preference to colder and denser deeper waters.

All variables were resampled to match the same horizontal resolution as the 3D grid (1/4° resolution (692 × 666 cells)) using R package "raster" (resample) with a bilinear method (Hijmans, 2022). Resample is applied to transfer values between non-matching raster objects. Here we applied the bilinear method for resampling, which is one of the most common resampling techniques and calculates values of a grid using the four nearest neighbouring cells, the bilinear interpolation assigns the output cell value by taking the weighted average.

Prior to model building, we calculated variance inflation factors (VIF) to test collinearity between environmental variables for each depth level (water column layers sets) and excluded those variables with VIF values above 5 (Zuur et al., 2009). Finally, values of the selected environmental variables along the depth range from 0 to 1000 m were extracted to species presence locations using *extract* function from "raster" R package (Hijmans, 2022). The method used to extract values from the environmental raster brick was "simple" method, so the value returned was the value of the cell where the point fell. This allowed us to generate our occurrence dataset comprising location of the occurrence (longitude, latitude, and depth values) and the value of each environmental variable (sea temperature, salinity, nitrate, net primary production, relative position to mixed layer depth and relative position to seabed).

2.5. Exploring macroecological patterns: vertical-latitudinal distribution

We explored the macroecological patterns testing if latitudinal-vertical distribution of main commercial fish species follows the isothermal distribution in the ocean. This hypothesis testing contributes also to empirically confirm the importance of considering explicitly the 3D aspect on the modelling approach for improving model realism (Levins, 1966). To address this, we employed two complementary methods.

In the first method, we fitted linear models for Northern and Southern hemisphere subsets of the occurrence points (> 50 occurrences) of each of the selected species and modelled depth as a function of the latitude where species occurs. These linear models allowed us to test our hypothesis, i.e., species tend to occur in upper layers of the water column poleward and hence at deeper layers towards the equator. Thus, according to our hypothesis, we expect positive slope in linear models along the latitudinal gradient for Northern hemisphere species, and negative slope for species occurrences at the Southern hemisphere.

For each species *i*, depth was modelled as a linear function of latitude as follows Eq. (3):

$$D_{i,j} = \alpha_i + \beta_i\, Lat_{i,j} + \varepsilon_{i,j} \quad (3)$$

where $D_{i,j}$ and $Lat_{i,j}$ denote the depth and the latitude of occurrence *j* for species *i*. For each species α_i and β_i are the model parameters and $\varepsilon_{i,j}$ follows a normal distribution with mean 0 and variance σ_i^2.

In the second method, we considered non-linear response and all species simultaneously because isothermal distribution in the ocean is not linear with respect to latitude and presents a slight inverted pattern in tropical waters, i.e. shallower poleward but with a secondary minimum close to equator (Palmer et al., 2019; Stewart, 2008). In particular, we tested the hypothesis in the overall community using a Generalized Additive Mixed Model (GAMM) (Zuur et al., 2009; Wood, 2017), where depth was modelled as a smooth function of the fixed effect of latitude where species occur, and species were considered a random effect. Our hypothesis will be supported if the fitted curve with GAMMs follows the latitudinal isothermal patterns. GAMMs were fitted, assuming Gaussian error distribution, using *bam* function from "mgcv" R package (Wood, 2011), which allows fitting GAM models to very large data sets.

In this second approach, the following model was fitted to all species simultaneously Eq. (4):

$$D_{i,j} = \alpha_i + \beta_i\, s\left(Lat_{i,j}\right) + \varepsilon_{i,j} \quad (4)$$

where

$$\alpha_i \sim Normal\left(0,\ \sigma_\alpha^2\right)$$

2.6. Species distribution 3D models

2.6.1. Shape-Constrained generalized additive models

According to ecological niche theory, species response curves are unimodal with respect to environmental gradients (Hutchinson, 1957). While a variety of statistical methods have been developed for species distribution modelling, a general problem with most of these habitat modelling approaches is that the estimated response curves can display biologically implausible shapes which do not respect ecological niche theory. This is because species response curves are fit statistically with any assumption or restriction, which sometimes do not respect the ecological niche theory. To better understand species response to environmental changes, SDMs should consider theoretical background such as the ecological niche theory (Elith and Leathwick, 2009) and pursue the unimodality of the response curves with respect to environmental gradients (Citores et al., 2020). Shape-Constrained Generalized Additive Models (SC-GAMs) have been pointed to be an effective alternative to fitting nonsymmetric parametric response curves, while retaining the unimodality constraint, required by ecological niche theory, for direct variables and limiting factors (Citores et al., 2020). Thus, here we selected SC-GAMs to build the 3D distribution models of the main commercial fish species of the Atlantic Ocean. SC-GAMs (Pya and Wood, 2015) allow incorporating monotonicity and concavity shape-constraints in the component functions of the environmental predictors of the GAMs (Wood, 2017) and avoid overfitting. A single-species and 2D approach of the SC-GAMs is available as open-access in the GAM-NICHE tutorial at: https://gam-niche.azti.es/ (Valle et al., 2023).

Since this type of models with binomial distribution need absences, we randomly generated (Barbet-Massin et al., 2012) an overall set of 500,000 pseudo-absence points through our 3D grid which covers the whole Atlantic Ocean. We selected this number of pseudo-absence points as a balance between having enough data, available cells in the grid, and computational limitations. Generated pseudo-absences points were then aggregated keeping only one record when several points fell in the same grid cell, and values of each of the environmental variables were extracted to pseudo-absence points using *extract* function and

simple method from "raster" R package (Hijmans, 2022) as we done to generate the occurrence dataset. For each species, we selected from the overall set of pseudo-absences an equal number of pseudo-absences to the number of occurrences (prevalence = 0.5) (Liu et al., 2005), avoiding pseudo-absences where species presences have been recorded.

We built the SC-GAMs using the R package "scam" (version 1.2–12) (Pya, 2021), and we selected binomial distribution and logit link to model species presence and absence. Following Citores et al. (Citores et al., 2020), we fixed the smoothing parameter (with the argument sp) to 10^{-5} and controlled the smoothness with a fixed number of knots in the construction of the model ($k = 8$). Node equal to 8 was selected after some trials to balance high vs low smoothing fitting, since it was considered a good compromise between complexity and model fit.

2.6.2. Model selection and cross-validation

To automatize the model selection, we created a function that generated a subset of models with different combinations of environmental variables (only including smooths with p-value < 0.05) and performed a forward model selection based on AIC (Akaike Information Criterion) (Sakamoto et al., 1986). The model selection process was stopped when the AIC improvement was lower than 2, following the rule of thumb according to which models with an AIC difference less than 2 are considered to have substantial support (Burnham and Anderson, 2002). The selected final models, or best models, were those constructed by adding significant variables one by one that lead to a reduction of the AIC higher than 2.

We used "SDMTools" R package (VanDerWal et al., 2019) for model evaluation and validation. As model performance, we have used the explained deviance, 1 - (residual deviance)/(null deviance), which is the equivalent to R^2 in least-squares models (Guisan and Zimmermann, 2000). We validated the models based on the cross-validation resampling procedure, which uses independent datasets for model building and model validation (Burnham and Anderson, 2002). The comparison between the accuracy of the model (that using all observations to build the model) and that of cross-validated permits the detection of model overfitting, which highly reduce the use of such models for extrapolation. The accuracy of the model has been evaluated using Area Under the Receiver Operating Characteristic—ROC—curve (AUC; (Fielding and Bell, 1997; Lobo et al., 2008)) and accuracy indices derived from the confusion matrix (VanDerWal et al., 2019) To this end, first, the modelled probability of species presence was converted to either presence or absence using probability thresholds. Threshold value for presence-absence classification for each model was obtained using *optim.thresh* function that estimates optimal threshold values given different accuracy measures. Following Jiménez-Valverde and Lobo (Jiménez-Valverde and Lobo, 2007), we selected sensitivity–specificity sum maximiser criteria for conversion of probability of species presence to either presence or absence. After selecting the thresholds, we validated the presence-pseudo-absence models via cross-validation method (Burnham and Anderson, 2003), with 5-fold equally sized sub-datasets (Hijmans, 2012) using *kfold* function from "dismo" R package (Hijmans, 2023). The data sets were divided into 5 groups or folds and each group was used iteratively to validate the model fitted to the remaining data (i.e., 80 % of randomly selected observations were used to fit the model and the remaining 20 % observations were used for validation). We computed five measures of accuracy for each k random subset and then averaged the results: 1) AUC, 2) omission rates (false predicted presences), 3) sensitivity (true predicted presences), 4) specificity (true predicted absences), and 5) proportion correctly identified.

Additionally, we explored how sensitive these measures were to the extent of the study region. To do so, we first generated Repeatable Training Regions for each species as a wrapper around each species occurrences. We used the function *marineBackground* from "voluModel" package (Owens, 2023) which generates background sampling regions by fitting an alpha hull polygon around an occurrence point dataset. Specifically, we applied the function without suppling the buff argument, thus our Training Regions were calculated by taking the mean between the 10th and 90th percentile of horizontal distances between occurrence points. Once we defined the species-specific training regions, we converted each of them into a 3D grid following the same method we applied to the pseudo-absences generated over the entire Atlantic Ocean, and then we randomly generated pseudo-absences points through the species-specific 3D grids. As we did for the original pseudo-absences set, we aggregated the points to the grid cells, we removed NA values, and we selected an equal number of pseudo-absences to the number of occurrences for each species avoiding pseudo-absences where species presences have been recorded. Having generated the species-specific training datasets for each species, we followed two different approaches to explore the influence of the extend of the study area in our models results. In the first approach we compared the resulting values of accuracy measures and the response curves, for the models built with pseudo-absence selected from species-specific training regions and the models built using pseudo-absence selected from the entire Atlantic. In the second approach we analysed how well were the models built using pseudo-absence selected from the entire Atlantic, predicting over the species-specific presence-pseudoabsence datasets.

2.6.3. Spatial prediction of 3D suitable habitat and catch distribution

To be able to extrapolate the models spatially, we first generated 47 water column layers sets, one for each depth standard level from surface up to 1000 m, and then we estimated the probability of occurrence for each water column level using the *predict* function from "scam" R package (Pya, 2021). Biogeographic processes such as dispersal limitation among others (Ludt, 2021) constrain the region used by the species, thus there can be suitable areas that are not occupied because species cannot disperse. Here we decided to mask the spatial predictions for each water column layer by the native distribution of the species to avoid including those areas. Native distribution was assigned to each species based on the occurrence points used in the models and the catches distribution from Watson (Watson, 2020). We divided the Atlantic Ocean into west and east creating a transect following the Mid-Atlantic ridge, whereas north and south divisions were based on latitude (>0 north, <0 south) (Fig. 3). A species was considered to be native of one of the divisions (north, south, east, west, north-east, north-west, south-east or south-west) if occurrences and catches found at that division were > 1% of the total occurrences. The species were

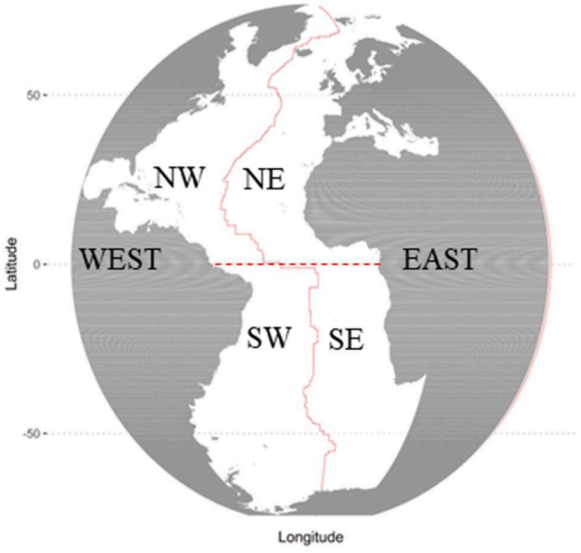


Fig. 3. Atlantic Ocean divisions defined to assign native distribution to modelled species. NW: North-West, NE: North-East, SW: South-West, SE: South-East.

assigned to be pan-Atlantic if they were present (> 1 % of the total occurrences) in all defined divisions. The threshold of 1 % was selected to prevent including occurrences data with geographic errors.

Spatial predictions were also vertically masked considering only the predictions until the maximum depth value according to the occurrence points (0.99 percentile). Probabilities were then summed up (and scaled to 0–1 range) along the water column and the latitude to get the final 3D cumulative probabilities maps. In order to get a proxy of the predictions confidence, we computed Multivariate Environmental Similarity Surfaces (MESS) analysis over all depth layers from each species using the observations datasets used to build the models following Elith et al. (Elith et al., 2010). Computing MESS analysis we got an index representing how similar a point in space was to a reference set of points, with respect to the set of predictors variables used to build the models. Resulting MESS maps were reclassified keeping only positive values and summed up through the water column (until the depth corresponding to 0.99 quantile depth) as done for the probability maps. As a result, we got one map for each species showing the summed up number of areas where no extrapolation is happening, these are areas where the model can be projected (i.e., the environment is well-represented by the set of conditions used to calibrate the model). Finally, to generate the potential catch biomass predictions, we estimated the total occurrence probability by summing up the probability of occurrence of all suitable cells above the threshold. This total occurrence probability was used to distribute the overall biomass catches of the given species across the 3D potential species habitat suitability as follows Eq. (5):

$$Catches_{Cell} = \frac{Occurrence\ probability_{\ Cell}\ x\ Species\ catches\ (tons{\cdot}year^{-1})}{Total\ occurrence\ probability} \quad (5)$$

Once we obtained the value of catches per cell, we summed it up along the water column. Resulting maps were used to represent species potential catch biomass into the 3D space.

3. Results

3.1. Fish species selection and occurrence across the water column

We selected the main 25 commercial species according to the catch amount for period 2010–2015. The 13 first selected commercial fish species summed up the 90 % of the total catches with reference to the 30 species, being the Atlantic herring (*Clupea harengus*), the Atlantic cod (*Gadus morhua*), the Atlantic mackerel (*Scomber scombrus*), the capelin (*Mallotus villosus*), and the blue whiting (*Micromesistius poutassou*) the most captured species in terms of biomass (Table 2). Among the tuna species, Skipjack tuna (*Katsuwonus pelamis*) was the species with documented higher mean biomass catches along 2010–2015 period (Table 2). Overall, the 30 species account for the 67 % of the total fish biomass catches.

The overall number of aggregated occurrences ranged from 504 (*Molva molva*) to 28,475 (*Xiphias gladius*) (Table 3). Widely distributed species such as tunas show the highest number of occurrences followed by top captured species (Table 3). Native distribution of modelled species was mainly assigned to north division of the Atlantic Ocean (12 out of 30 species, 40 %), followed by north-east and pan-Atlantic (8 species for each, 27 %) and east Atlantic Ocean (2 species out of 30, 6 %) (Table 3).

We found three main groups of fish according to depth data associated to occurrences retrieved from the global open-access datasets (Table 3 and Fig. 4). The first group includes shallow water species (maximum depth value less than 300 m); the second group is for species reaching maximum depths below 300 m with mean depths between 75 and 170 m; and the third group has species reaching 1000 m deep with mean depths deeper than 200 m.

Table 2
List of selected 30 species for modelling. Catches refer to the mean biomass over 2010–2015 period in the Atlantic Ocean.

Selected groups		Common name	Scientific name	Catch (tons/year)	% of catches
Commercial species	1	Atlantic herring	*Clupea harengus*	2303,212	23.44
	2	Atlantic cod	*Gadus morhua*	1516,384	15.43
	3	Atlantic mackerel	*Scomber scombrus*	1223,354	12.45
	4	Capelin	*Mallotus villosus*	936,920	9.54
	5	Blue whiting	*Micromesistius poutassou*	851,248	8.66
	6	Haddock	*Melanogrammus aeglefinus*	466,179	4.74
	7	European sprat	*Sprattus sprattus*	399,989	4.07
	8	Saithe	*Pollachius virens*	431,277	4.39
	9	European anchovy	*Engraulis encrasicolus*	249,791	2.54
	10	Atlantic horse mackerel	*Trachurus trachurus*	117,176	1.19
	11	European hake	*Merluccius merluccius*	104,306	1.06
	12	Greenland halibut	*Reinhardtius hippoglossoides*	101,513	1.03
	13	European plaice	*Pleuronectes platessa*	86,837	0.88
	14	Norway pout	*Trisopterus esmarkii*	60,517	0.62
	15	Deepwater redfish	*Sebastes mentella*	56,008	0.57
	16	Ling	*Molva molva*	49,577	0.50
	17	Swordfish	*Xiphias gladius*	42,352	0.43
	18	Whiting	*Merlangius merlangus*	35,551	0.36
	19	Largehead hairtail	*Trichiurus lepturus*	36,624	0.37
	20	Common sole	*Solea solea*	27,366	0.28
	21	Flounder	*Platichthys flesus*	18,031	0.18
	22	John dory	*Zeus faber*	17,389	0.18
	23	Angler	*Lophius piscatorius*	14,657	0.15
	24	Common dolphinfish	*Coryphaena hippurus*	13,332	0.14
	25	European conger	*Conger conger*	11,364	0.12
Tuna species	26	Skipjack tuna	*Katsuwonus pelamis*	387,407	3.94
	27	Yellowfin tuna	*Thunnus albacares*	144,013	1.47
	28	Bigeye tuna	*Thunnus obesus*	64,919	0.66
	29	Albacore	*Thunnus alalunga*	49,415	0.50
	30	Atlantic bluefin tuna	*Thunnus thynnus*	9084	0.09
Total				9825,792	100.00

3.2. Macroecological patterns for main commercial fish species

Results from the fitted linear models to test if latitudinal-vertical distribution of main commercial fish species follows the general pattern of isothermal distribution in the ocean, showed significant response ($p < 0.001$) for 48.48 % of the cases that followed the pattern and 18.18 % of the cases that followed opposite pattern; whilst not significant responses ($p < 0.001$) were found for 21.21 % of the cases that followed the expected pattern, and 12.12 % that presented the opposite pattern (Supplementary Figures 1 & 2 in Supplementary Material 1). These results support our hypothesis, indicating that most of the analysed species tend to occur in the upper ocean layers poleward.

Resulting GAMM fitted using data from the overall communities from the northern hemisphere allowed us to relate the depth at which

Table 3
Number of occurrences, depth values (mean, maximum, and 99% quantile), and assigned native distribution of the main commercial fish from the Atlantic Ocean.

Scientific name	n occurrences	Mean depth	Max depth	99% quantile depth	Atlantic division
Molva molva	504	123.79	500	423.5	north
Engraulis encrasicolus	525	65.92	225	200	east
Trichiurus lepturus	571	155.75	750	650	pan-Atlantic
Conger conger	596	169.61	850	700	north-east
Solea solea	615	47.72	225	171.5	north-east
Platichthys flesus	647	43.32	450	125	north-east
Zeus faber	846	107.49	450	350	east
Micromesistius poutassou	917	216	1000	700	north
Lophius piscatorius	1209	139.9	750	600	north-east
Trisopterus esmarkii	1236	92.2	425	225	north-east
Sprattus sprattus	1239	56.33	225	150	north-east
Merluccius merluccius	1255	141.69	750	623	north-east
Pleuronectes platessa	1640	58.75	250	175	north
Pollachius virens	1705	104.28	600	400	north
Merlangius merlangus	1887	67.4	275	200	north
Katsuwonus pelamis	1913	144.02	1000	950	pan-Atlantic
Sebastes mentella	2123	339.83	1000	900	north
Mallotus villosus	2618	142.54	1000	650	north
Reinhardtius hippoglossoides	2679	342.23	1000	950	north
Melanogrammus aeglefinus	3274	91.17	650	350	north
Trachurus trachurus	3418	125.33	750	400	north-east
Thunnus thynnus	4125	98.65	1000	850	pan-Atlantic
Scomber scombrus	4782	125.47	600	275	north
Clupea harengus	6443	75.7	550	325	north
Gadus morhua	6705	122.43	900	550	north
Thunnus alalunga	9201	100.68	1000	750	pan-Atlantic
Thunnus obesus	15,345	102.43	1000	800	pan-Atlantic
Coryphaena hippurus	17,077	91.54	1000	300	pan-Atlantic
Thunnus albacares	22,251	121.27	1000	900	pan-Atlantic
Xiphias gladius	28,475	113.83	1000	850	pan-Atlantic

species occur as a function of latitude (deviance explained = 14 %, $p < 0.0001$) (Fig. 5; Supplementary Figures 3 & 4 in Supplementary Material 1) and confirmed that commercial fish species at northern hemisphere followed a similar pattern of isothermal distribution of the ocean.

3.3. 3D species distribution modelling incorporating water column data

According to the correlation assessment between environmental variables performed to each water column layer set using VIF, dissolved oxygen variable was highly correlated to temperature (Supplementary Figure 1 in Supplementary Material 2) and was removed from the environmental variables. SC-GAM models were built using the other 6 environmental variables (see the corresponding statistics for all variables within the occurrence points and pseudo-absence points in Supplementary Table 1). SC-GAM models performed well according to the explained deviance (ranging from 0.39 to 0.93, with mean = 0.75) and *R*-squared (0.44–0.96, mean = 0.80) (Table 4) (see also Supplementary Figure 2 in Supplementary Material 2). Relative distance to seabed was included in 27 of the 30 models and was the variable explaining most of the variation in 18 cases. Temperature and salinity were included in 24 out of 30 models, being temperature, the variable explaining most of the variation in 6 cases and salinity in 1 case. Temperature response curves fitted with SC-GAMs showed the large variation in optimal and width thermal niche across commercial fishes (Fig. 6). Partial response curves for the different environmental variables from species occurrence probability models also showed a large variation in the optimal ranges (see Supplementary Figures from 3 to 7 in Supplementary Material 2). Relative distance to MLD was included in 20 out of the 30 models and explained most of the variation in 5 cases. NPP and Nitrate were included in 17 and 14 models, respectively, and were not within the most explaining variables in any case.

Tuna species were those species presenting the lowest explained deviance (Table 4) and the lowest values in the accuracy measurements, except for omission errors where the models presented the highest values (Table 5). In addition to the 5 tuna species there are also 2 species (*Xiphias gladius* and *Coryphaena hippurus*) which are widely distributed species (native range = Pan-Atlantic) that also presented low explained deviance and low accuracy measurements values (Table 4 & Table 5).

All models showed good predictive powers with mean values of all accuracy indices (AUC, sensitivity, specificity, and proportion of presences and absences correctly identified) above 90 % (Table 5). All models mean omission rate was 0.05.

Results from the additional exploration that was carried out to assess how sensitive were the models results to the extent of the study region showed that, although resulting AUC values obtained for the models built using species-specific training regions were lower to the models built using the pseudo-absence selected from the entire Atlantic, the performance of the models built using species-specific training regions was consistent with that from the models built using pseudo-absences selected from the entire Atlantic. AUC values for the species-specific training regions models ranged from 0.71 to 0.98 whilst those for the models built using pseudo-absences selected from the entire Atlantic ranged from 0.78 to 0.98 (Supplementary Table 1 in Supplementary Material 3). Similar differences were found in the rest of measures: Omission error (underprediction, false predicted presences), Sensitivity (proportion of correctly predicted presence records) and Specificity (proportion of correctly predicted absences) and Prop. Correct (proportion correctly identified) measures (Supplementary Table 1 in Supplementary Material 3). While comparing the response curves of the models, we found that overall, curves were keeping the similar shapes and when changes occurred, those occurred specially at the lowest and largest values of the covariables, due to the influence of the range of values where the pseuso-absences were created (Supplementary Figures 1 to 4 in Supplementary Material 3). In some cases, the response curves were basically the same (e.g. *Engraulis encrasicolus* response curve of the variable relative distance to seabed), whereas in some other cases the response curves shapes were different (e.g. *Merluccius merluccius* response curve of the variable relative position to mixed layer depth). Finally, when assessing how well were the models built using pseudo-absence selected from the entire Atlantic, predicting over the species-specific presence-pseudoabsence datasets, we found an accurate prediction (Supplementary Table 2 in Supplementary Material 3).

Based on the fitted models, we predicted species occurrence probability along the water column for the 30 main commercial fish species of the Atlantic Ocean. Resulting probability 3D maps are freely available for download (see Data Availability section). Several models predicted high probability of occurrence in areas where the species are not currently present because their native distribution corresponds only to one hemisphere or to the east of west of the Atlantic Ocean. Thus, masking the predictions to native distribution of the species allowed

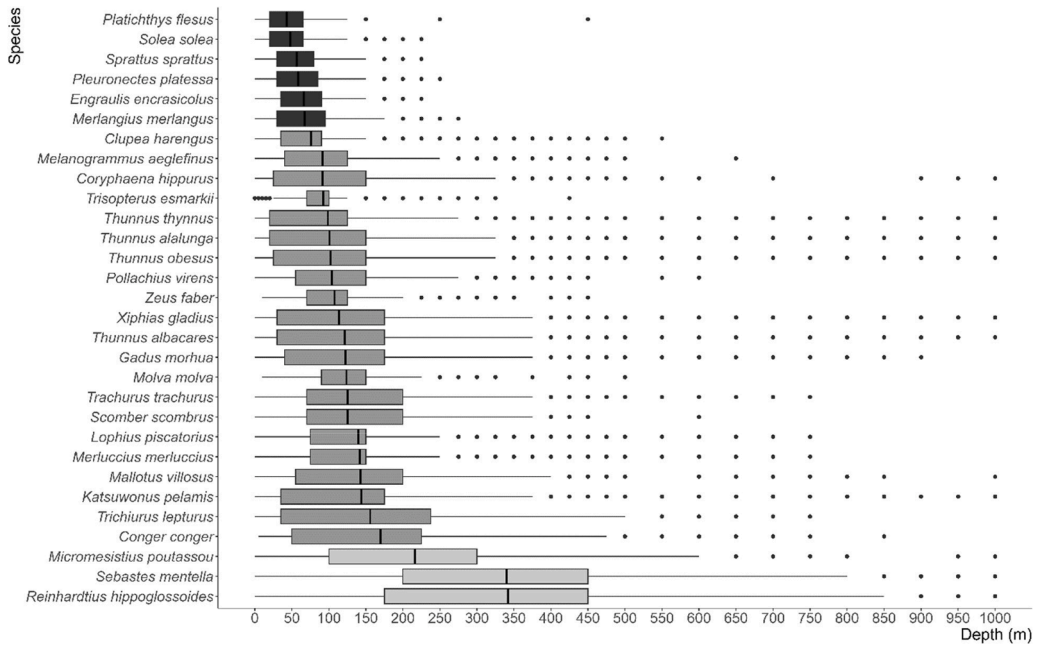


Fig. 4. Distribution of the species occurrences from the along the depth gradient from 0 to 1000 m. Lower and upper box boundaries are 25th and 75th percentiles, the line inside the box represents the mean and the lower and upper error whiskers are the 10th and 90th percentiles, respectively. The filled circles represent data falling outside 10th and 90th percentiles. Boxplots are coloured in greyscale based on the three main fish groups according to depth.

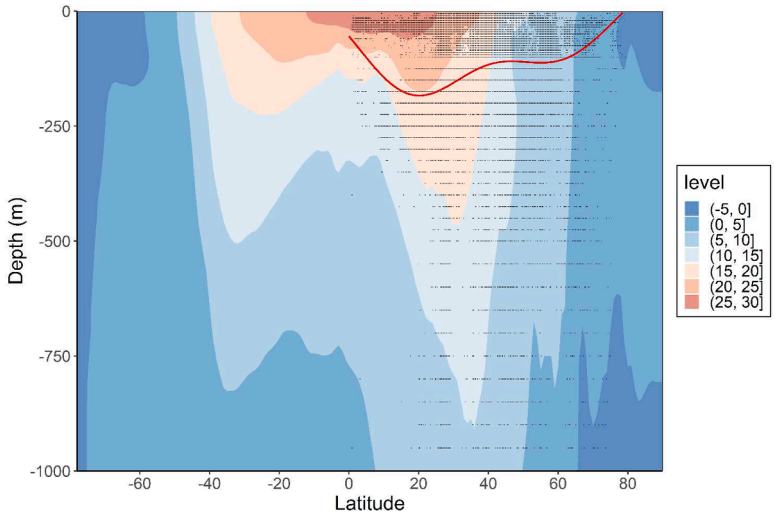


Fig. 5. Mean ocean temperature profile along the latitudinal gradient with overlapped species occurrence points for the overall fish community from the Atlantic Northern hemisphere and the fitted GAMM curve (in red) of depth as a function of latitude with species as a random effect. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

avoiding unrealistic estimation of catches and applying threshold level to filter the predictions allowed obtaining more accurate estimations. The 3D cumulative probability maps are shown for *E. encrasicolus* and *T. alalunga* as illustration examples (Fig. 7, Fig. 8 and conceptual versions in semi-3D in Supplementary Figures 38 & 39 in Supplementary Material 2), but are available for all species together with the cumulative positive MESS maps in the Supplementary Material 2 (Supplementary Figures from 8 to 37). The SC-GAM model for European anchovy was fitted including temperature and relative distance to seabed variables. The model explained 0.87 of the deviance (Table 4) and an AUC of 0.97 (Table 5). Cumulative occurrence probability 3D map (Fig. 7) showed high occurrence probability in the Northeast Atlantic region at shallow depths.

The SC-GAM model for albocore was fitted including all variables. This tuna species occurs in tropical and temperate waters of all oceans including the Mediterranean Sea. The model had a low predictive power in comparison to models from other species with a more restricted distribution range, it explained 0.4 of the deviance (Table 4) and presented an AUC of 0.78 (Table 5). The cummulative probability 3D map (Fig. 8) showed widespread occurrence probabilty for the species between 60°N to 60°S latitudes, indicating that species required environmental conditions are almost found elsewhere in that area of the Atlantic ocean.

Resulting map of cumulative potential biomass catches along the water column for the Atlantic Ocean, are shown for the 25 main commercial species and for the 5 tuna species (Fig. 9). Individual maps for each species are freely available for download (see Data Availability section). Total potential catch biomass for the 30 species, summed up 9825,698 tons/year, with maximum values in the North Sea and the

Table 4
Explained deviance (ED) and *R*-squared (r.sq) values for the model built for each species. Chi squares values for each variable included in the model (those variables with higher chi square values are explaining most of the variation). temp: temperature; sal: salinity, nit: nitrate, dist: relative distance to seabed; MLD: Relative Position to Mixed Layer Depth; and NPP: Net Primary Production.

Scientific name	ED	r.sq	temp	sal	nit	dist	MLD	NPP
Clupea harengus	0.78	0.83	748.58	209.12	118.14	**1206.09**	–	16.23
Conger conger	0.83	0.88	13.68	15.13	–	**125.13**	–	–
Coryphaena hippurus	0.52	0.58	1313.70	71.53	41.33	2364.27	**2479.75**	758.61
Engraulis encrasicolus	0.87	0.91	**25.65**	–	–	20.23	–	–
Gadus morhua	0.8	0.85	110.24	200.54	278.70	216.11	**445.08**	–
Katsuwonus pelamis	0.62	0.7	195.99	40.62	–	**561.86**	194.15	28.22
Lophius piscatorius	0.93	0.95	32.05	**35.50**	–	26.31	9.71	–
Mallotus villosus	0.76	0.82	–	38.77	50.48	**208.71**	65.10	45.35
Melanogrammus aeglefinus	0.83	0.88	**127.56**	40.69	201.76	13.57	111.14	29.48
Merlangius merlangus	0.81	0.85	**173.17**	20.12	–	–	73.11	–
Merluccius merluccius	0.88	0.91	**112.52**	7.33	16.72	43.49	19.05	–
Micromesistius poutassou	0.83	0.87	17.62	24.65	–	**174.68**	–	20.07
Molva molva	0.77	0.83	–	30.53	–	**42.90**	–	–
Platichthys flesus	0.9	0.93	–	–	–	**51.63**	–	–
Pleuronectes platessa	0.83	0.88	**156.12**	24.19	–	–	34.62	–
Pollachius virens	0.85	0.89	19.30	12.13	–	**186.99**	–	–
Reinhardtius hippoglossoides	0.84	0.88	–	45.85	25.49	**222.95**	69.95	13.87
Scomber scombrus	0.79	0.85	946.04	–	–	**1012.24**	247.87	195.12
Sebastes mentella	0.9	0.94	62.52	13.99	–	**96.07**	60.92	46.76
Solea solea	0.85	0.89	–	–	8.13	**38.77**	–	–
Sprattus sprattus	0.93	0.96	17.14	30.36	–	**45.29**	–	–
Thunnus alalunga	0.4	0.44	377.97	22.68	376.75	603.56	**1662.39**	544.62
Thunnus albacares	0.5	0.57	2505.31	105.71	141.22	**3687.56**	3136.95	1103.01
Thunnus obesus	0.39	0.44	598.45	16.42	642.43	1123.29	**1790.40**	585.58
Thunnus thynnus	0.53	0.59	270.15	45.93	111.84	**1004.24**	471.42	27.45
Trachurus trachurus	0.84	0.88	259.75	–	69.96	**599.46**	191.76	59.21
Trichiurus lepturus	0.87	0.92	26.89	11.08	–	**99.81**	37.79	–
Trisopterus esmarkii	0.75	0.81	–	17.71	–	**94.85**	–	33.19
Xiphias gladius	0.41	0.46	1293.15	185.54	738.53	2867.94	**3673.79**	1204.49
Zeus faber	0.66	0.7	**133.64**	–	–	–	67.61	63.18
All models mean	0.75	0.80						
All models SD	0.16	0.16						

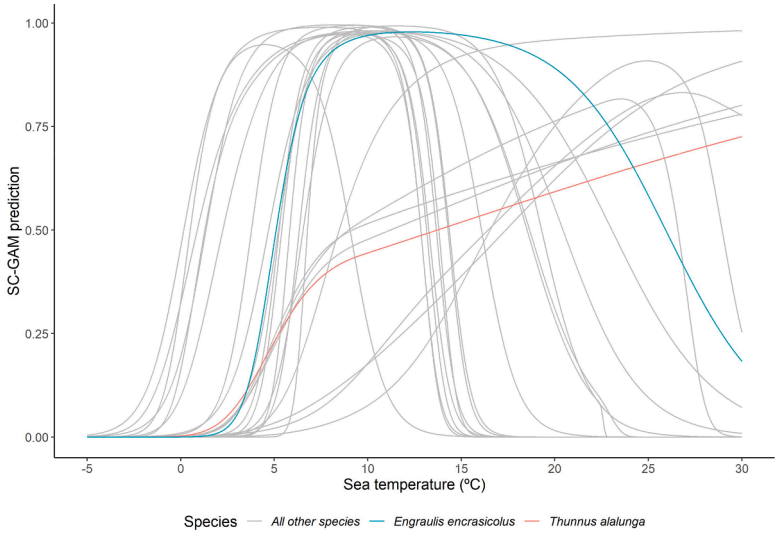


Fig. 6. Partial response curves for sea temperature from species occurrence probability models. Each line corresponds to different Atlantic commercial fish species, *E. encrasicolus* and *T. alalunga* are highlighted in colours.

Norwegian Sea. Both coastal areas, eastern and western, were found to be most productive in comparison with open ocean.

4. Discussion

4.1. Macroecological patterns: vertical-latitudinal distribution

The distribution of most of the commercial fish species studied here follows the general latitudinal-vertical pattern of isothermal distribution in the ocean, i.e., species occur at shallower layers at mid to high latitudes, whilst they shift towards deeper and intermediate waters at lower latitudes, except for an inverted pattern close to equator. Thus, our results confirmed the expected hypothesis of "equatorial submergence" (Close et al., 2006), showing that Atlantic commercial fish species avoid the warmer surface waters towards the equator by diving down to deeper waters and tend to occur in shallower layers when moving poleward (Reygondeau et al., 2012; Trubovitz et al., 2020). The equatorial submergence phenomenon was known since Charles Darwin, who

Table 5
Mean accuracy measures for the SC-GAM models built for each species: mean AUC, mean omission, mean sensitivity, mean specificity, and mean proportion correctly identified.

Scientific name	mean AUC	mean Omission	mean Sensitivity	mean Specificity	mean Prop Correct
Clupea harengus	0.95	0.03	0.97	0.93	0.95
Conger conger	0.96	0.04	0.96	0.96	0.96
Coryphaena hippurus	0.86	0.14	0.86	0.85	0.86
Engraulis encrasicolus	0.97	0.01	0.99	0.94	0.97
Gadus morhua	0.96	0.03	0.97	0.94	0.96
Katsuwonus pelamis	0.91	0.07	0.93	0.88	0.91
Lophius piscatorius	0.98	0.01	0.99	0.98	0.98
Mallotus villosus	0.94	0.01	0.99	0.9	0.94
Melanogrammus aeglefinus	0.96	0.02	0.98	0.95	0.96
Merlangius merlangus	0.95	0.04	0.96	0.94	0.95
Merluccius merluccius	0.97	0.02	0.98	0.97	0.97
Micromesistius poutassou	0.96	0.03	0.97	0.95	0.96
Molva molva	0.95	0.06	0.94	0.96	0.95
Platichthys flesus	0.97	0.02	0.98	0.96	0.97
Pleuronectes platessa	0.96	0.03	0.97	0.96	0.96
Pollachius virens	0.97	0.03	0.97	0.96	0.97
Reinhardtius hippoglossoides	0.97	0.03	0.97	0.97	0.97
Scomber scombrus	0.95	0.02	0.98	0.91	0.95
Sebastes mentella	0.97	0.02	0.98	0.96	0.97
Solea solea	0.97	0.01	0.99	0.94	0.97
Sprattus sprattus	0.98	0	1	0.96	0.98
Thunnus alalunga	0.78	0.03	0.97	0.6	0.78
Thunnus albacares	0.86	0.11	0.89	0.82	0.86
Thunnus obesus	0.8	0.14	0.86	0.73	0.8
Thunnus thynnus	0.86	0.13	0.87	0.84	0.86
Trachurus trachurus	0.97	0.02	0.98	0.95	0.97
Trichiurus lepturus	0.97	0.03	0.97	0.97	0.97
Trisopterus esmarkii	0.94	0.03	0.97	0.92	0.94
Xiphias gladius	0.8	0.18	0.82	0.78	0.8
Zeus faber	0.9	0.08	0.92	0.87	0.9
All models mean	0.93	0.05	0.95	0.91	0.93
All models SD	0.06	0.05	0.05	0.08	0.06

described how the Arctic fishes disappear in the seas of Japan and of northern China and reappear on the coast of Tasmania, southern New Zealand and the Antarctic islands (Pauly, 2004). The occurrence of this phenomenon can be extended to any type of marine animals (Close et al., 2006; Ekman, 1967), as found in radiolarians for instance (Trubovitz et al., 2020). This macroecological pattern is probably a response related to temperature-respiration constraints shaping the species thermal niche (Close et al., 2006; Pauly et al., 1998). Equatorial submergence is caused by the same physiological constraints that also determine the latitudinal range of species (Close et al., 2006), and therefore its shift due to global warming are expected to vary accordingly. This has relevant implications since those species that cannot track optimal temperatures on a global scale might undergo local extinction where local environmental thresholds are exceeded (Trubovitz et al., 2020). As fish latitudinal-vertical distribution follows the prevailing isothermal distribution in the ocean, an accurate representation of species distributions requires 3D modelling with incorporation of explicit depth dimension into the environmental data.

4.2. Species distribution 3D models: contributions and limits

Species distribution models at the marine realm have been widely used as a tool for understanding species spatial ecology (Robinson et al., 2017), but scarcely addressed the water column explicitly. In comparison with SDMs for terrestrial organisms, the three-dimensional habitat of pelagic organisms implies consideration of depth information (Dambach and Rödder, 2011). Ignoring depth in SDMs leaves out crucial information and could lead to misleading outcomes (Bentlage et al., 2013). Some studies use ocean depth as a covariate in the model as an attempt to account for the third dimension and other use simplified at-depth variables (e.g. surface temperature and bottom temperature) to represent marine environments (Duffy and Chown, 2017), however final prediction is always a 2D representation. Several authors have already highlighted the particular relevance and challenge of the three-dimensionality in modelling marine environments (Dambach and Rödder, 2011; Duffy and Chown, 2017; Robinson et al., 2011). In fact, a recent comprehensive review of ecological niche models and SDMs in marine environments concluded that the additions of a third dimension has yet to be incorporated (Melo-Merino et al., 2020). This is a consequence of data limitation related to species occurrences that are not always vertically informed (e.g., fisheries-derived information) or/and the availability of depth-specific environmental data (e.g., environmental predictors). Dambach and Rödder (Dambach and Rödder, 2011) proposed novel techniques for addressing specific needs in a 3D context and some authors have attempted to consider the three-dimensionality of the marine environment, Bentlage et al. (2013) decomposed 3-dimensional structure of marine environmental into a single, continuous 2D grid that replicated global oceans multiple times allowing to evaluate all the ecological information contained in distinct depth layers at the same time. Duffy and Chown (Duffy and Chown, 2017) used multiple 2D layers to approximate the 3-dimension as proposed by Dambach and Rödder (Dambach and Rödder, 2011). Recently, Pérez-Costas et al. (Pérez-Costas et al., 2019) developed a modelling procedure called NOO3D available in the ModestR software, which can be used to estimate the 3D distribution of species using 3D occurrence samples and 3D environmental datasets.

Taking advantage of the rising available data resources, here we adapted the methodology from Duffy and Chown (Duffy and Chown, 2017) to estimate the species habitat suitability along the water column. Similarly to us, Owens and Rahbek, (2023) have also developed a way forward to model species distribution into the 3 dimensional space by generating species distribution models based on environmental data extracted at the depths where individuals were observed, and calibrated with three-dimensional sampling of pseudo-absences. The novelty of our approach is that based on the modelled species probability of occurrence along the water column, we estimate the potential catch biomass for the Atlantic Ocean from surface to 1000 m depth. Resulting maps showed that most of the potential catch biomass for 25 main non-tuna commercial species is mainly accumulated at the northern hemisphere, while potential tuna catches mainly occur at tropical and subtropical regions. However, due to the limitations of the input data used to model species distribution (see below) results should not be over-interpreted in a certain area and species.

Our models were based exclusively on public species occurrence data retrieved from the global repositories GBIF and OBIS. These public data repositories usually present geographic biases, such as inaccurate geolocation and spatial autocorrelation among occurrence points, as well as species misidentifications (Melo-Merino et al., 2020). In addition, many occurrence records held in these repositories lack associated depth data and must, therefore, be excluded from 3D analyses (Duffy and Chown, 2017). Our approach included a data cleaning procedure to address these issues which removes outliers from the geographic space and validates species names before including the occurrence points of each species data set. Through this data cleaning procedure, those records lacking the depth information were removed as recommended by Duffy

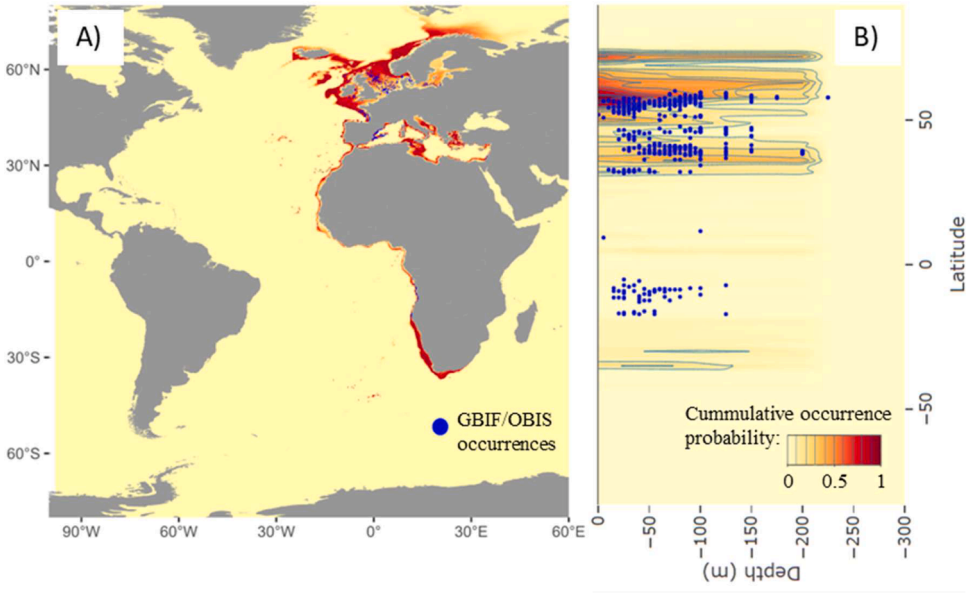


Fig. 7. 3D cumulative probability map for European anchovy (*Engraulis encrasicolus*), A) summed up probabilities along the water column and B) summed up probabilities through the latitudinal gradient.

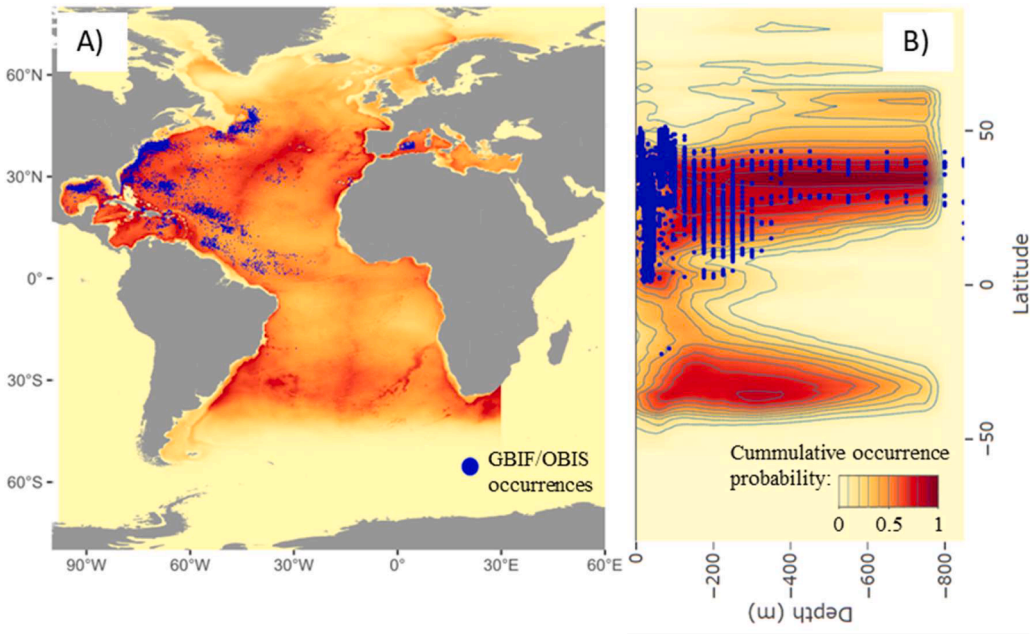


Fig. 8. 3D cumulative probability map for Albacore (*Thunnus alalunga*), A) summed up probabilities along the water column and B) summed up probabilities through the latitudinal gradient.

and Chown (Duffy and Chown, 2017). Any improvement in the quantity and quality of information of these databases, especially regarding the depth dimension, would directly lead to better and more reliable 3D models. Published occurrence datasets also lack occurrence data in some regions (Gaiji et al., 2013), this might imply an incomplete characterization of the niche when building models based on this type of data and might cause inaccuracies in the spatial predictions in particular areas. This is the case of *Sprattus sprattus*, where South Mediterranean areas are not being predicted as suitable (Supplementary Figure 14 in Supplementary Material 2). A large proportion of the occurrence data we retrieved from the global repositories was fishery-dependent in conjunction with a smaller proportion of fishery-independent data. Although fishery dependent data are inherently biased, they can still be useful to supplement other data sources, if the SDMs account properly for preferential sampling and other potential bias sources (Karp et al., 2023).

We developed here a methodological approach to generate 3D fish distribution based exclusively on public species occurrence data and environmental correlates that conform with ecological niche theory, using shape-constrained GAMs. Fitted SC-GAMs resulted in a good balance between goodness of fit and agreement with ecological niche theory. Models were selected and fitted automatically through our model selection function. Automatic selection of models was found to be very advantageous when modelling Multiple species, however, influence of the conditions fitted for variable selection (p-value and AIC criterion) should be considered according to research aims. Within the 30

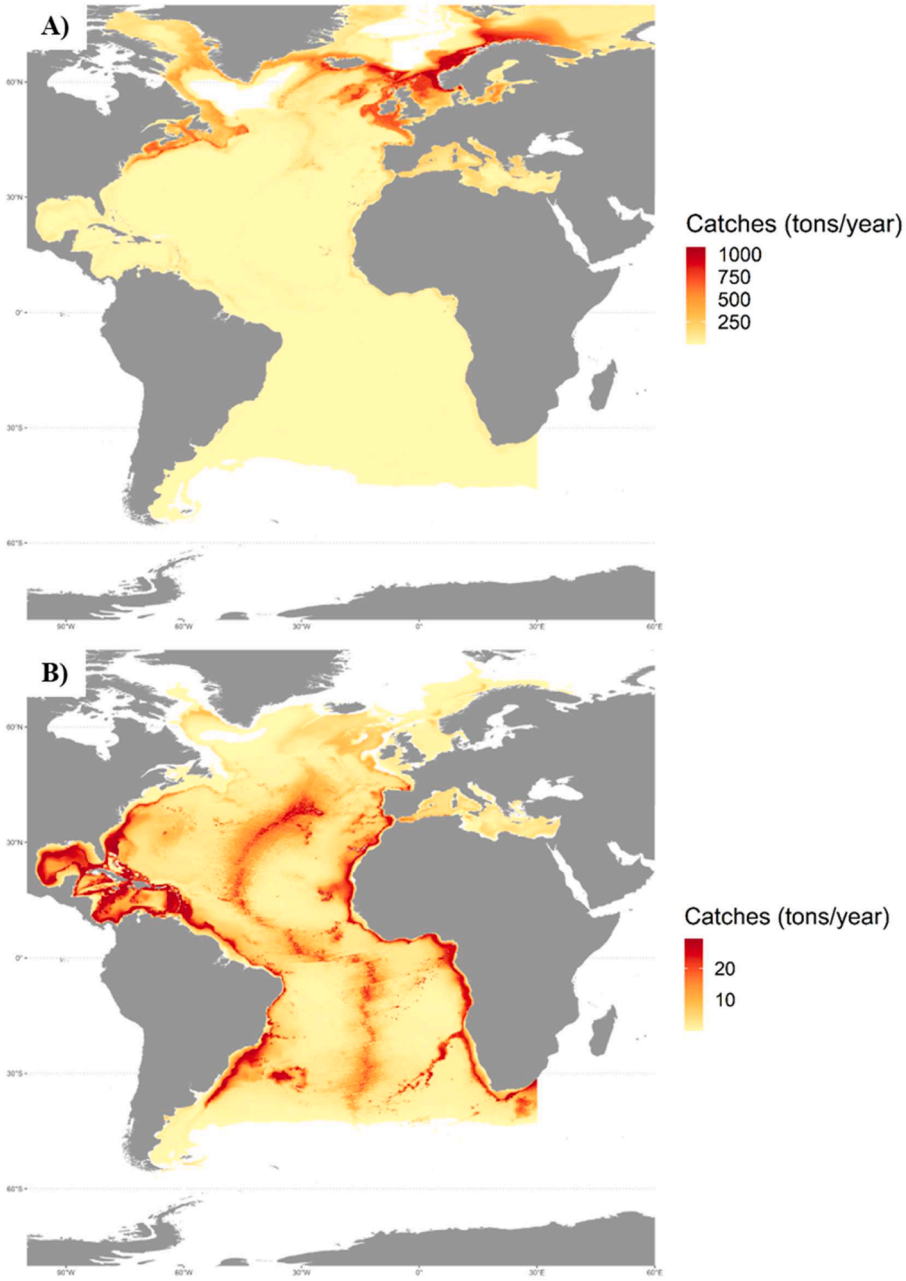


Fig. 9. Modelled potential biomass catches (tons/year) per surface unit (0.25° latitude x 0.25° longitude) summing up catches along the water column for A) 25 main commercial species, and B) 5 Tuna species.

modelled species we found that spatial predictions for *Solea solea* (Supplementary Figure 27 in Supplementary Material 2) and *Platichthys flesus* (Supplementary Figure 28 in Supplementary Material 2) were overestimating suitability to the northernmost region of the Northeast Atlantic Ocean. This could be related to the fact that the final model for these species did not include sea temperature as predictor variable. As mentioned above, results from our study were useful to reach our goal of getting an overall view of potential catch biomass based on the probability of occurrence modelled M. species but should not be over-interpreted in a certain area and species.

The developed 3D models of fish occurrence probability have the capability to be improved with the updates of new data for data-poor species, and to be projected under climate change scenarios. Theoretical background such as the ecological niche theory (Hutchinson, 1957) taken into account here avoids overfitting-derived artifacts in climate change projections. The use of SDMs in marine environment is increasing (Melo-Merino et al., 2020) because of its relatively easiness to build and apply; however, SDM methods should be improved to overcome some common limitations. One limitation we encountered here was the need of presence and absence data to fit SC-GAMs with depth distribution. We addressed the lack of real absence data randomly generating pseudo-absence points through the 3D grid. Although selection of pseudo-absences points was species-specific in number (we selected an equal number of pseudo-absences to the number of occurrences), we did not make the selection spatially specific and therefore the extent of the study area was the entire Atlantic Ocean for all the species. The influence of the extent of the study area have been found to strongly affect the results of SDMs (Acevedo et al., 2012; Barve et al., 2011; Lobo et al., 2008). Different approaches have been proposed for delimiting the geographical background in species distribution modelling, such as marine training regions (Owens and Rahbek 2022) or trend surface analysis (TSA) (Acevedo et al., 2012). By restricting the extent

using a geographical criterion, model performance in the core area of the species distribution can be significantly improved (Acevedo et al., 2012). However, when predicting to areas that have environmental values that are beyond those found in the training region, prediction uncertainty increases (Elith et al., 2010). Here we kept the extent of the geographical background to the entire Atlantic Ocean according to our goal, nevertheless, we calculated multivariate environmental similarity surfaces to get a proxy of predictions' confidence and masked our predictions considering that there are biogeographic processes such as dispersal limitation among others (Ludt, 2021) that constrain the region used by the species. As pointed out by Lobo et al. (2008), AUC values might be affected by the extent of the study area. Here we found lower AUC values for the models built using species-specific training regions, however the additional exploration we carried out to assess how sensitive were the models results to the extent of the study region did not show substantial differences across settings (dataset built using pseudo-absences selected from the entire Atlantic vs dataset built selecting pseudo-absence from species-specific training regions) and predictions were found to be consistent. In addition, our models explained deviance for tuna species (ranging from 0.39 to 0.62) is comparable to that found by Erauskin-Extramiana et al., (2019) (ranging from 0.35 to 0.62).

It must be taken into account that in this case study we did not include any consideration regarding the fishing activity, such as the gear employed, the effort or the fleet characteristics. Another limitation of our work is the use of climatological 3D fields of the ocean, i.e., not considering the temporal variability scales of the ocean into the SDM. This might be taken into account in further steps for migrating species, i.e., with different seasonal foraging or spawning habitat. Limitations that our approach have not addressed are the population dynamics, dispersal movements, multi-habitat use of species, and explicit species interactions. Interspecific interactions, for instance, can strongly affect the biogeography of species beyond local extents (Hui, 2016; Tikhonov et al., 2020). Some tools such as joint SDMs have recently been developed to combine both climatic and species interactions (Hui, 2016; Tikhonov et al., 2020).

4.3. Perspectives

Our approach could be applied to other areas, different target species, future scenarios, and help identifying the most fish productive areas. This information could be also useful to support policy makers to balance the need for environmental protection with sustainable marine resource exploitation (i.e., spatial prioritization, design of marine protected areas, and marine spatial planning). Using the resulting habitat suitability maps biodiversity assessments and global change impact assessments could be performed. Furthermore, species models could be projected under climate change scenarios to help reducing the uncertainty of the climatic impacts on fisheries (e.g., Erauskin-Extramiana et al. (Erauskin-Extramiana et al., 2019)). A better understanding of the 3D distribution of the commercial species can help to resolve conflicts when new human activities are planned (Coccoli et al., 2018; Queirós et al., 2021). This 3D models can also help to adapt fisheries to climate change and maybe an opportunity for the industry to mitigate climate change by reducing emissions (Granado et al., 2021). Climate change is affecting species distribution and biodiversity patterns (Baudron et al., 2020; Chust et al., 2022; Erauskin-Extramiana et al., 2020), and it is expected to continue in the future (Erauskin-Extramiana et al., 2023; Fernandes et al., 2020). Furthermore, climate-adaptation objectives in fisheries management are largely missing (Bryndum-Buchholz et al., 2021). Industry can adapt to these changes by using operational species distribution forecasting tools which are starting to emerge (Honarmand Ebrahimi et al., 2021; Rubio et al., 2021). These operational forecasting tools may also be used to identify avoidance zones, e.g. to reduce bycatch (Howell et al., 2008). Another potential use is to support the optimization of fishing operations to reduce their fuel consumption (reduce operational costs) and the consequent reduction in emissions (climate change mitigation) (Basurko et al., 2022; Bell et al., 2017; Parker et al., 2018).

5. Concluding remarks

The accurate distribution of main commercial fish over the Atlantic and through the water column is still unknown. Here, we explored fish macroecological patterns and concluded that latitudinal-vertical distribution follows the prevailing isothermal distribution in the ocean, confirming that a reliable representation of distributions needs 3D modelling and explicit incorporation of depth dimension into the environmental data. Thus, we developed a methodological approach to model 3D species distribution based exclusively on public species occurrence data and environmental correlates that conform with ecological niche theory, using shape-constrained GAMs. The ecological models are, in this way, selected and fitted automatically. The species response curves to 3D environmental gradients (sea water temperature, salinity, nitrate, net primary productivity, distance to seafloor, and relative distance to mixed layer depth) for the 30 main commercial species of the Atlantic yielded very good model accuracy performance (78–98 %). The developed 3D models of fish occurrence probability have the capability to be improved with the updates of new data for data-poor species, and to be projected under climate change scenarios. The obtained 3D maps of estimated fish distribution conform useful and new knowledge that may help policy makers to balance the need for environmental protection with sustainable marine resource exploitation of the Atlantic Ocean.

CRediT authorship contribution statement

Mireia Valle: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Eduardo Ramírez-Romero:** Writing – review & editing, Formal analysis. **Leire Ibaibarriaga:** Writing – review & editing, Methodology, Formal analysis. **Leire Citores:** Writing – review & editing, Methodology. **Jose A. Fernandes-Salvador:** Writing – review & editing. **Guillem Chust:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Detailed methods, code, intermediate data, and output data for this analysis are freely available at https://github.com/missionatlantic/3D_SDMs_AtlanticFish (Valle, 2024).

Species probability map sets (one map for each depth layer) and cumulative potential catches along the water column maps for each species are freely available in netcdf format into the Mission Atlantic folder from Marine Data Archive (see https://www.vliz.be/en/imis?module=dataset&dasid=8057; and https://www.vliz.be/en/imis?module=dataset&dasid=8461, respectively) and can be visualized in the Mission Atlantic GeoNode (https://geonode.missionatlantic.eu/) and in the European Marine Observation and Data Network (EMODnet) central portal website (https://emodnet.ec.europa.eu/en).

Funding

This study has been supported by the European Union's Horizon 2020 research and innovation programme under grant agreements No

862428 (MISSION ATLANTIC project) and No 869342 (SusTunTech project).

Acknowledgments

We acknowledge Ocean Biodiversity Information System (OBIS; https://obis.org/) and the Global Biodiversity Information Facility (GBIF; https://www.gbif.org/) for providing fish global occurrences studied here. We also thank to publicly available datasets Copernicus (http://marine.copernicus.eu/services-portfolio/access-to-products/) and the World Ocean Atlas (https://www.ncei.noaa.gov/products/world-ocean-atlas) for providing environmental data. The comments of two anonymous reviewers have improved considerably the first manuscript draft. This paper is contribution n° 1206 from AZTI, Marine Research, Basque Research and Technology Alliance (BRTA).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2024.110632.

References

Acevedo, P., Jiménez-Valverde, A., Lobo, J.M., Real, R., 2012. Delimiting the geographical background in species distribution modelling. J. Biogeogr. 39, 1383–1390.

Aiken, J., Rees, N., Hooker, S., Holligan, P., Bale, A., Robins, D., Moore, G., Harris, R., Pilgrim, D., 2000. The Atlantic Meridional Transect: overview and synthesis of data. Prog. Oceanogr. 45, 257–312.

Amante, C., Eakins, B., 2009. ETOPO1 1 Arc-Minute Global Relief Model: procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24 19.

Arrizabalaga, H., Dufour, F., Kell, L., Merino, G., Ibaibarriaga, L., Chust, G., Irigoien, X., Santiago, J., Murua, H., Fraile, I., Chifflet, M., Goikoetxea, N., Sagarminaga, Y., Aumont, O., Bopp, L., Herrera, M., Marc Fromentin, J., Bonhomeau, S., 2015. Global habitat preferences of commercially valuable tuna. Deep Sea Research Part II: Topical Studies in Oceanography 113, 102–112.

Austin, M.P., Heyligers, P.C., 1989. Vegetation survey design for conservation: gradsect sampling of forests in North-eastern New South Wales. Biol. Conserv. 50, 13–32.

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution 3, 327–338.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J., Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecol Modell 222, 1810–1819.

Basurko, O.C., Gabiña, G., Lopez, J., Granado, I., Murua, H., Fernandes, J.A., Krug, I., Ruiz, J., Uriondo, Z., 2022. Fuel consumption of free-swimming school versus FAD strategies in tropical tuna purse seine fishing. Fish. Res. 245, 106139.

Baudron, A.R., Brunel, T., Blanchet, M.-A., Hidalgo, M., Chust, G., Brown, E.J., Kleisner, K.M., Millar, C., MacKenzie, B.R., Nikolioudakis, N., Fernandes, J.A., Fernandes, P.G., 2020. Changing fish distributions challenge the effective management of European fisheries. Ecography 43, 494–505.

Bell, J.D., Watson, R.A., Ye, Y., 2017. Global fishing capacity and fishing effort from 1950 to 2012. Fish and Fisheries 18, 489–505.

Bentlage, B., Peterson, A.T., Barve, N., Cartwright, P., 2013. Plumbing the depths: extending ecological niche modelling and species distribution modelling in three dimensions. Global Ecology and Biogeography 22, 952–961.

Biastoch, A., Schwarzkopf, F.U., Getzlaff, K., Rühs, S., Martin, T., Scheinert, M., Schulzki, T., Handmann, P., Hummels, R., Böning, C.W., 2021. Regional imprints of changes in the Atlantic Meridional Overturning Circulation in the eddy-rich ocean model VIKING20X. Ocean Sci 17, 1177–1211.

Boettiger, C., Lang, D.Temple, Wainwright, P., 2012. rfishbase: exploring, manipulating and visualizing FishBase data from R. J. Fish Biol.

Boyer, T., Garcia, H., Locarnini, R., Zweng, M., Mishonov, A., Reagan, J., Weathers, K., Baranova, O., Seidov, D., Smolyar, I., 2018. Mixed Layer Depth. NOAA National Centers for Environmental Information. Dataset. https://www.ncei.noaa.gov/archive/accession/NCEI-WOA18. Accessed [May 2022].

Brodie, S., Jacox, M.G., Bograd, S.J., Welch, H., Dewar, H., Scales, K.L., Maxwell, S.M., Briscoe, D.M., Edwards, C.A., Crowder, L.B., Lewison, R.L., Hazen, E.L., 2018. Integrating Dynamic Subsurface Habitat Metrics Into Species Distribution Models. Front Mar Sci 5.

Bruge, A., Alvarez, P., Fontán, A., Cotano, U., Chust, G., 2016. Thermal Niche Tracking and Future Distribution of Atlantic Mackerel Spawning in Response to Ocean Warming. Front Mar Sci 3.

Bryndum-Buchholz, A., Tittensor, D.P., Lotze, H.K., 2021. The status of climate change adaptation in fisheries management: policy, legislation and implementation. Fish and Fisheries 22, 1248–1273.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd edn. Springer, Berlin. edition.

Burnham, K.P., Anderson, D.R., 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York.

Castilho, R., Grant, W.S., Almada, V.M., 2013. Biogeography and phylogeography of the Atlantic. Front. Biogeogr. 5, 5–7.

Chamberlain, S., V. Barve, D. Mcglinn, D. Oldoni, P. Desmet, L. Geffert, and K. Ram. 2022. rgbif: interface to the Global Biodiversity Information Facility API.2023 R package version 3.7.2. https://CRAN.R-project.org/package=rgbif.

Chust, G., González, M., Fontán, A., Revilla, M., Alvarez, P., Santos, M., Cotano, U., Chifflet, M., Borja, A., Muxika, I., Sagarminaga, Y., Caballero, A., de Santiago, I., Epelde, I., Liria, P., Ibaibarriaga, L., Garnier, R., Franco, J., Villarino, E., Irigoien, X., Fernandes-Salvador, J., Uriarte, A., Esteban, X., Orue-Echevarria, D., Figueira, T., Uriarte, A., 2022. Climate regime shifts and biodiversity redistribution in the Bay of Biscay. Sci Total Environ 803, 149622.

Citores, L., Ibaibarriaga, L., Lee, D.J., Brewer, M.J., Santos, M., Chust, G., 2020. Modelling species presence–absence in the ecological niche theory framework using shape-constrained generalized additive models. Ecol Modell 418, 108926.

Close, C.H., W. Cheung, S. Hodgson, V. Lam, R. Watson, and D. Pauly. 2006. Distribution ranges of commercial fishes and invertebrates. Pages 27-37 Palomares, M.L.D., Stergiou, K.I., Pauly, D. (eds.), Fishes in Databases and Ecosystems. Fisheries Centre Research Reports.

Coccoli, C., Galparsoro, I., Murillas, A., Pınarbaşı, K., Fernandes, J.A., 2018. Conflict analysis and reallocation opportunities in the framework of marine spatial planning: a novel, spatially explicit Bayesian belief network approach for artisanal fishing and aquaculture. Mar Policy 94, 119–131.

Cooley, S., Schoeman, D., Bopp, L., Boyd, P., Donner, S., Ghebrehiwet, D.Y., Ito, S.-I., Kiessling, W., Martinetto, P., Ojea, E., Racault, M.-F., Rost, B., Skern-Mauritzen, M., et al., 2022. Oceans and Coastal Ecosystems and Their Services. editors. In: Pörtner, H.-O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., et al. (Eds.), Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel On Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 379–550.

Dambach, J., Rödder, D., 2011. Applications and future challenges in marine species distribution modeling. Aquatic Conservation: Marine and Freshwater Ecosystems 21, 92–100.

Dell'Apa, A., Boenish, R., Fujita, R., Kleisner, K., 2023. Effects of climate change and variability on large pelagic fish in the Northwest Atlantic Ocean: implications for improving climate resilient management for pelagic longline fisheries. Front Mar Sci 10.

Duffy, G.A., Chown, S.L., 2017. Explicitly integrating a third dimension in marine species distribution modelling. Mar. Ecol. Prog. Ser. 564, 1–8.

du Pontavice, H., Gascuel, D., Kay, S., Cheung, W.W.L., 2023. Climate-induced changes in ocean productivity and food-web functioning are projected to markedly affect European fisheries catch. Mar. Ecol. Prog. Ser. 713, 21–37.

Ekman, S., 1967. Zoogeography of the Sea. Sidgwick & Jackson, London.

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods in Ecology and Evolution 1, 330–342.

Elith, J., Leathwick, J.R., 2009. Species Distribution Models: ecological Explanation and Prediction Across Space and Time. Annual Review of Ecology Evolution, and Systematics 40, 677–697.

Emeis, K.-C., van Beusekom, J., Callies, U., Ebinghaus, R., Kannen, A., Kraus, G., Kröncke, I., Lenhart, H., Lorkowski, I., Matthias, V., Möllmann, C., Pätsch, J., Scharfe, M., Thomas, H., Weisse, R., Zorita, E., 2015. The North Sea — A shelf sea in the Anthropocene. Journal of Marine Systems 141, 18–33.

Erauskin-Extramiana, M., Alvarez, P., Arrizabalaga, H., Ibaibarriaga, L., Uriarte, A., Cotano, U., Santos, M., Ferrer, L., Cabré, A., Irigoien, X., Chust, G., 2019a. Historical trends and future distribution of anchovy spawning in the Bay of Biscay. Deep Sea Research Part II: Topical Studies in Oceanography 159, 169–182.

Erauskin-Extramiana, M., Arrizabalaga, H., Cabré, A., Coelho, R., Rosa, D., Ibaibarriaga, L., Chust, G., 2020. Are shifts in species distribution triggered by climate change? A swordfish case study. Deep Sea Research Part II: Topical Studies in Oceanography 175, 104666.

Erauskin-Extramiana, M., Arrizabalaga, H., Hobday, A.J., Cabré, A., Ibaibarriaga, L., Arregui, I., Murua, H., Chust, G., 2019b. Large-scale distribution of tuna species in a warming ocean. Glob Chang Biol 25, 2043–2060.

Erauskin-Extramiana, M., Chust, G., Arrizabalaga, H., Cheung, W.W.L., Santiago, J., Merino, G., Fernandes-Salvador, J.A., 2023. Implications for the global tuna fishing industry of climate change-driven alterations in productivity and body sizes. Glob Planet Change 222, 104055.

FAO, 2022. The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation. FAO, Rome.

Fernandes, J.A., Rutterford, L., Simpson, S.D., Butenschön, M., Frölicher, T.L., Yool, A., Cheung, W.W.L., Grant, A., 2020. Can we project changes in fish abundance and distribution in response to climate? Glob Chang Biol 26, 3891–3905.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24, 38–49.

Gaiji, S., Chavan, V., Ariño, A.H., Otegui, J., Hobern, D., Sood, R., Robles, E., 2013. Content Assessment of the Primary Biodiversity Data Published through GBIF Network: status, Challenges and Potentials. Biodiversity Informatics 8.

Gaines, S.D., Costello, C., Owashi, B., Mangin, T., Bone, J., Molinos, J.G., Burden, M., Dennis, H., Halpern, B.S., Kappel, C.V., Kleisner, K.M., Ovando, D., 2018. Improved fisheries management could offset many negative effects of climate change. Sci Adv 4, eaao1378.

Garcia, H., Weathers, K., Paver, C., Smolyar, I., Boyer, T., Locarnini, R., Zweng, M., Mishonov, A., Baranova, O., Seidov, D., Reagan, J., 2018a. World Ocean Atlas. In:

Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation, 3. NOAA Atlas NESDIS, p. 38. *A. Mishonov Technical Ed.*pp.
Garcia, H., Weathers, K., Paver, C., Smolyar, I., Boyer, T., Locarnini, R., Zweng, M., Mishonov, A., Baranova, O., Seidov, D., Reagan, J., 2018b. World Ocean Atlas. In: Dissolved Inorganic Nutrients (phosphate, Nitrate and nitrate+nitrite, Silicate), 4. NOAA Atlas NESDIS, p. 35. *A. Mishonov Te*chnical Ed.pp.
GBIF, 2024. Derived dataset GBIF.org (25 January 2024) Filtered export of GBIF occurrence data. https://doi.org/10.15468/dd.hnnamj.
Granado, I., Hernando, L., Galparsoro, I., Gabiña, G., Groba, C., Prellezo, R., Fernandes, J.A., 2021. Towards a framework for fishing route optimization decision support systems: review of the state-of-the-art and challenges. J Clean Prod 320, 128661.
Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H. P., Buckley, Y.M., 2013. Predicting species distributions for conservation decisions. Ecol. Lett. 16, 1424–1435.
Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol Modell 135, 147–186.
Halpern, B.S., Frazier, M., Afflerbach, J., Lowndes, J.S., Micheli, F., O'Hara, C., Scarborough, C., Selkoe, K.A., 2019. Recent pace of change in human impact on the world's ocean. Sci Rep 9, 11609.
Helaouët, P., Beaugrand, G., 2009. Physiology, Ecological Niches and Species Distribution. Ecosystems 12, 1235–1245.
Hijmans, R. 2022. raster: geographic Data Analysis and Modeling. R package version 3.5-15. https://CRAN.R-project.org/package=raster.
Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. Ecology 93, 679–688.
Hijmans, R.J., S. Phillips, J.R. Leathwick, and J. Elith. 2021. dismo: species Distribution Modeling. R package version 1.3-5.2023 https://CRAN.R-project.org/package=dismo.
Hobday, A.J., 2010. Ensemble analysis of the future distribution of large pelagic fishes off Australia. Prog. Oceanogr. 86, 291–301.
Honarmand Ebrahimi, S., Ossewaarde, M., Need, A., 2021. Smart Fishery: a Systematic Review and Research Agenda for Sustainable Fisheries in the. Age of AI. Sustainability 13, 6037.
Howell, E.A., Kobayashi, D.R., Parker, D.M., Balazs, G.H., 2008. TurtleWatch: a tool to aid in the bycatch reduction of loggerhead turtles Caretta caretta in the Hawaii-based pelagic longline fishery. Endanger Species Res 5, 267–278.
Hui, F.K.C., 2016. boral – Bayesian Ordination and Regression Analysis of Multivariate Abundance Data in r. Methods in Ecology and Evolution 7, 744–750.
Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harb. Symp. Quant. Biol. 22, 415–427.
Jiménez-Valverde, A., Lobo, J.M., 2007. Threshold criteria for conversion of probability of species presence to either–or presence–absence. Acta Oecologica 31, 361–369.
Karp, M.A., Brodie, S., Smith, J.A., Richerson, K., Selden, R.L., Liu, O.R., Muhling, B.A., Samhouri, J.F., Barnett, L.A.K., Hazen, E.L., Ovando, D., Fiechter, J., Jacox, M.G., Buil, M.Pozo, 2023. Projecting species distributions using fishery-dependent data. Fish and Fisheries 24, 71–92.
Kesner-Reyes, K., C. Garilao, K. Kaschner, J. Barile, and R. Froese. 2020. AquaMaps: algorithm and data sources for marine organisms.*in* R. F. D. P. (Eds.), editor. FishBase. https://www.fishbase.org, version (10/2019).
Levins, R., 1966. THE STRATEGY OF MODEL BUILDING IN POPULATION BIOLOGY. Am. Sci. 54, 421–431.
Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28, 385–393.
Lobo, J.M., Jiménez-Valverde, A., Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17, 145–151.
Locarnini, R., Mishonov, A., Baranova, O., Boyer, T., Zweng, M., Garcia, H., Reagan, J., Seidov, D., Weathers, K., Paver, C., Smolyar, I., 2018. World Ocean Atlas. In: A. Mishonov Technical Ed.; NOAA Atlas NESDIS, 1. Temperature, p. 52 pp.
Longhurst, A.R., 2007. Chapter 9 - THE ATLANTIC OCEAN. editor. In: Longhurst, A.R. (Ed.), ***Ecological Geography of the Sea***, 2nd Edition. Academic Press, Burlington, pp. 131–273.
Lotze, H.K., Tittensor, D.P., Bryndum-Buchholz, A., Eddy, T.D., Cheung, W.W.L., Galbraith, E.D., Barange, M., Barrier, N., Bianchi, D., Blanchard, J.L., Bopp, L., Büchner, M., Bulman, C.M., Carozza, D.A., Christensen, V., Coll, M., Dunne, J.P., Fulton, E.A., Jennings, S., Jones, M.C., Mackinson, S., Maury, O., Niiranen, S., Oliveros-Ramos, R., Roy, T., Fernandes, J.A., Schewe, J., Shin, Y.-J., Silva, T.A.M., Steenbeek, J., Stock, C.A., Verley, P., Volkholz, J., Walker, N.D., Worm, B., 2019. Global ensemble projections reveal trophic amplification of ocean biomass declines with climate change. Proceedings of the National Academy of Sciences 116, 12907–12912.
Ludt, W.B., 2021. Missing in the Middle: a Review of Equatorially Disjunct Marine Taxa. Front Mar Sci 8.
Magalhães Filho, L., Roebeling, P., Villasante, S., Bastos, M.I., 2022. Ecosystem services values and changes across the Atlantic coastal zone: considerations and implications. Mar Policy 145, 105265.
Maynou, F., Sabatés, A., Ramirez-Romero, E., Catalán, I.A., Raya, V., 2020. Future distribution of early life stages of small pelagic fishes in the northwestern Mediterranean. Climatic Change 161, 567–589.
Melo-Merino, S.M., Reyes-Bonilla, H., Lira-Noriega, A., 2020. Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. Ecol Modell 415, 108837.
Møller, P.R., Nielsen, J.G., Fossen, I., 2003. Patagonian toothfish found off Greenland. Nature 421, 599.
Morey, S.L., O'Brien, J.J., 2002. The spring transition from horizontal to vertical thermal stratification on a midlatitude continental shelf. Journal of Geophysical Research: Oceans 107, 12. -11-12-12.
OBIS. 2022. Ocean Biodiversity Information System. Intergovernmental Oceanographic Commission of UNESCO. [29 July 2022].
Ormond, R., Gage, J., Angel, M., 1997. *Marine Biodiversity: Patterns and Processes*. Cambridge University Press, Cambridge.
Owens, H., and C. Rahbek. 2022. voluModel: modeling Species Distributions in Three Dimensions 2023.
Owens, H.L., Rahbek, C., 2023. voluModel: modelling species distributions in three-dimensional space. Methods in Ecology and Evolution 14, 841–847.
Palmer, M.D., Durack, P.J., Chidichimo, M.P., Church, J.A., Cravatte, S., Hill, K., Johannessen, J.A., Karstensen, J., Lee, T., Legler, D., Mazloff, M., Oka, E., Purkey, S., Rabe, B., Sallée, J.-B., Sloyan, B.M., Speich, S., von Schuckmann, K., Willis, J., Wijffels, S., 2019. Adequacy of the Ocean Observation System for Quantifying Regional Heat and Freshwater Storage and Change. Front Mar Sci 6.
Pante, E., Simon-Bouhet, B., 2013. marmap: a Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R. PLoS ONE 8 (9), e73051.
Parker, R.W.R., Blanchard, J.L., Gardner, C., Green, B.S., Hartmann, K., Tyedmers, P.H., Watson, R.A., 2018. Fuel use and greenhouse gas emissions of world fisheries. Nat Clim Chang 8, 333–337.
Pauly, D., 2004. ***Darwin's Fishes: An Encyclopedia of Ichthyology, Ecology and Evolution***. Cambridge University Press, Cambridge.
Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., Torres, F., 1998. Fishing Down Marine Food Webs. Science 279, 860–863.
Pauly, D., Zeller, D., 2016. Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. Nat Commun 7, 10244.
Payne, M.R., Kudahl, M., Engelhard, G.H., Peck, M.A., Pinnegar, J.K., 2021. Climate risk to European fisheries and coastal communities. Proceedings of the National Academy of Sciences 118, e2018086118.
Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.-C., Clark, T. D., Colwell, R.K., Danielsen, F., Evengård, B., Falconi, L., Ferrier, S., Frusher, S., Garcia, R.A., Griffis, R.B., Hobday, A.J., Janion-Scheepers, C., Jarzyna, M.A., Jennings, S., Lenoir, J., Linnetved, H.I., Martin, V.Y., McCormack, P.C., McDonald, J., Mitchell, N.J., Mustonen, T., Pandolfi, J.M., Pettorelli, N., Popova, E., Robinson, S.A., Scheffers, B.R., Shaw, J.D., Sorte, C.J.B., Strugnell, J.M., Sunday, J. M., Tuanmu, M.-N., Vergés, A., Villanueva, C., Wernberg, T., Wapstra, E., Williams, S.E., 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being Science 355, eaai9214.
Pérez-Costas, E., Guisande, C., Vilas, L.G., Roselló, E.G., Heine, J., Dacosta, J.G., Lobo, J. M., 2019. NOO3D: a procedure to perform 3D species distribution models. Ecol. Informatics 54.
Predragovic, M., Cvitanovic, C., Karcher, D.B., Tietbohl, M.D., Sumaila, U.R., Horta e Costa, B., 2023. A systematic literature review of climate change research on Europe's threatened commercial fish species. Ocean Coast Manag 242, 106719.
Provoost, P., and S. Bosch. 2021. robis: ocean Biodiversity Information System (OBIS) Client. R package version 2.8.2. 2023 https://CRAN.R-project.org/package=robis.
Pya, N. 2021. scam: shape Constrained Additive Models. R package version 1.2-12. https://CRAN.R-project.org/package=scam.
Pya, N., Wood, S.N., 2015. Shape constrained additive models. Stat Comput 25, 543–559.
Queirós, A.M., Talbot, E., Beaumont, N.J., Somerfield, P.J., Kay, S., Pascoe, C., Dedman, S., Fernandes, J.A., Jueterbock, A., Miller, P.I., Sailley, S.F., Sará, G., Carr, L.M., Austen, M.C., Widdicombe, S., Rilov, G., Levin, L.A., Hull, S.C., Walmsley, S.F., Nic Aonghusa, C., 2021. Bright spots as climate-smart marine spatial planning tools for conservation and blue growth. Glob Chang Biol 27, 5514–5531.
R Core Team, 2023. R: A language and Environment For Statistical Computing.In R. F. f. S. Computing, Editor., Vienna, Austria.
Reygondeau, G., Maury, O., Beaugrand, G., Fromentin, J.M., Fonteneau, A., Cury, P., 2012. Biogeography of tuna and billfish communities. J. Biogeogr. 39, 114–129.
Robinson, L.M., Elith, J., Hobday, A.J., Pearson, R.G., Kendall, B.E., Possingham, H.P., Richardson, A.J., 2011. Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Global Ecology and Biogeography 20, 789–802.
Robinson, N.M., Nelson, W.A., Costello, M.J., Sutherland, J.E., Lundquist, C.J., 2017. A Systematic Review of Marine-Based Species Distribution Models (SDMs) with Recommendations for Best Practice. Front Mar Sci 4.
Rubio, I., Hobday, A.J., Ojea, E., 2021. Skippers' preferred adaptation and transformation responses to catch declines in a large-scale tuna fishery. Ices Journal of Marine Science 79, 532–539.
Sakamoto, Y., Ishiguro, M., Kitagawa, G., Reidel, D., 1986. ***Akaike Information Criterion Statistics***. Dordrecht, The Netherlands, p. 81.
Sayre, R., D. Wright, S. Breyer, K. Butler, K. Van Graafeiland, M. Costello, P. Harris, K. Goodin, J. Guinotte, Z. Basher, M. Kavanaugh, P. Halpin, M. Monaco, N. Cressie, P. Aniello, C. Frye, and D. Stephens. 2017. A Three-Dimensional Mapping of the Ocean Based on Environmental Data. Oceanography 30.
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 and Fisheries 22, 212–225.
Schwing, F.B., 2023. Modern technologies and integrated observing systems are "instrumental" to fisheries oceanography: a brief history of ocean data collection. Fish Oceanogr 32, 28–69.
Stewart, R.H. 2008. Introduction To Physical Oceanography.

Thompson, M.S.A., Couce, E., Schratzberger, M., Lynam, C.P., 2023. Climate change affects the distribution of diversity across marine food webs. Glob Chang Biol 29, 6606–6619.
Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehikoinen, A., de Jonge, M.M.J., Oksanen, J., Ovaskainen, O., 2020. Joint species distribution modelling with the r-package Hmsc. Methods in Ecology and Evolution 11, 442–447.
Tittensor, D.P., Novaglio, C., Harrison, C.S., Heneghan, R.F., Barrier, N., Bianchi, D., Bopp, L., Bryndum-Buchholz, A., Britten, G.L., Büchner, M., Cheung, W.W.L., Christensen, V., Coll, M., Dunne, J.P., Eddy, T.D., Everett, J.D., Fernandes-Salvador, J.A., Fulton, E.A., Galbraith, E.D., Gascuel, D., Guiet, J., John, J.G., Link, J. S., Lotze, H.K., Maury, O., Ortega-Cisneros, K., Palacios-Abrantes, J., Petrik, C.M., du Pontavice, H., Rault, J., Richardson, A.J., Shannon, L., Shin, Y.-J., Steenbeek, J., Stock, C.A., Blanchard, J.L., 2021. Next-generation ensemble projections reveal higher climate risks for marine ecosystems. Nat Clim Chang 11, 973–981.
Trubovitz, S., Lazarus, D., Renaudie, J., Noble, P.J., 2020. Marine plankton show threshold extinction response to Neogene climate change. Nat Commun 11, 5069.
Valle, M., 2024. missionatlantic/3D_SDMs_AtlanticFish: 0.0. In Ecological modelling. Zenodo. https://doi.org/10.5281/zenodo.10637351.
Valle, M., Citores, L., Ibaibarriaga, L., Chust, G., 2023. GAM-NICHE: Shape-Constrained GAMs to Build Species Distribution Models under the Ecological Niche Theory. AZTI.
VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., Storlie, C., 2019. SDMTools: species Distribution Modelling Tools: tools for processing data associated with species distribution modelling exercises. R package version, 1.1-221.2. https://CRAN.R-project.org/package=SDMTools.
Watson, R. 2020. Global Fisheries Landings V4.0.
Wood, S., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society 73, 3–36.
Wood, S.N. 2017. Generalized Additive Models: an Introduction with R (2nd ed.). Chapman and Hall/CRC.
Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S., Jackson, J.B. C., Lotze, H.K., Micheli, F., Palumbi, S.R., Sala, E., Selkoe, K.A., Stachowicz, J.J., Watson, R., 2006. Impacts of Biodiversity Loss on Ocean Ecosystem Services. Science 314, 787–790.
Zizka, A., D. Silvestro, T. Andermann, J. Azevedo, C.Duarte Ritter, D. Edler, H. Farooq, A. Herdean, M. Ariza, R. Scharn, S. Svanteson, N. Wengstrom, V. Zizka, and A. Antonelli. 2019. CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. R package version 2.0-20. Methods in Ecology and Evolution.
Zuur, A., E. Ieno, N. Walker, A. Saveliev, and G. Smith. 2009a. Mixed effects modes and extensions in ecology with R.
Zweng, M., Reagan, J., Seidov, D., Boyer, T., Locarnini, R., Garcia, H., Mishonov, A., Baranova, O., Weathers, K., Paver, C., Smolyar, I., 2018. World Ocean Atlas. In: Salinity. A. Mishonov Technical Ed.; NOAA Atlas NESDIS, 2, p. 50 pp.