Robust identification of potential habitats of a rare demersal species (blackspot seabream) in the Northeast Atlantic

Lola de Cubber, Verena Trenkel, Guzman Diez, Juan Gil-Herrera, Ana Maria Novoa Pabon, David Eme, Pascal Lorance

▶ To cite this version:

Lola de Cubber, Verena Trenkel, Guzman Diez, Juan Gil-Herrera, Ana Maria Novoa Pabon, et al.. Robust identification of potential habitats of a rare demersal species (blackspot seabream) in the Northeast Atlantic. Ecological Modelling, 2023, 477, pp.110255. 10.1016/j.ecolmodel.2022.110255 . hal-04028923

HAL Id: hal-04028923

https://hal.inrae.fr/hal-04028923v1

Submitted on 8 Jan 2025

Robust identification of potential habitats of a rare demersal species (blackspot seabream) in the Northeast Atlantic

Lola De Cubber[a,b,*], Verena M. Trenkel[a], Guzman Diez[c], Juan Gil-Herrera[d], Ana Maria Novoa Pabon[e], David Eme[a,f], Pascal Lorance[a]

[a] *DECOD (Ecosystem Dynamics and Sustainability), IFREMER, INRAe, Institut-Agro - Agrocampus Ouest, France, Nantes, France*
[b] *Institut de Recherche pour le Développement (IRD), MARBEC (IRD, Ifremer, CNRS, Univ Montpellier), Av. Jean Monnet - Sète, France*
[c] *AZTI, Marine Research, Basque Research and Technology Alliance (BRTA). Txatxarramendi ugartea z/g, 48395 Sukarrieta - Bizkaia, Spain*
[d] *Centro Oceanográfico de Cádiz (IEO-CSIC), Muelle de Levante s/n, 11006, Cádiz, Spain*
[e] *Departamento de Oceanografia e Pescas, Universidade dos Açores, 9901-862 Horta, Portugal*
[f] *RiverLY Research Unit, National Research Institute for Agriculture Food and Environment (INRAE), Villeurbanne, France*

Abstract

Species distribution models (SDM) are commonly used to identify potential habitats. When fitting them to heterogeneous, opportunistically collated presence/absence data, imbalance in the number of presence and absence observations often occurs, which could influence results. To robustly identify potential habitats for blackspot seabream (*Pagellus bogaraveo*) throughout its distribution area in the Northeast Atlantic and the western Mediterranean Sea, we used an ensemble species distribution modelling (eSDM) approach, modelling gridded presence-absence data with environmental predictors for two types of occurrence data sets. The first data set displayed the observed unbalanced spatially heterogeneous presence/absence ratio and the second a balanced presence/absence ratio. The data covered the full distribution area, including the European Atlantic shelf, the Azorean region and the Western Mediterranean Sea. Across these regions, populations display variable status. The main environmental predictors for potential habitats were bathymetry and annual maximum SST. The fitted ensemble compromise (eSDM) was projected over the whole grid to create a habitat suitability map. This map exhibited higher probabilities of presence for the balanced-ratio data set. A binary presence-absence map was then generated using optimised presence probability thresholds for four validation indices. Using the true skill statistic to optimise the threshold, the surface areas of the binary presence-absence map was 53% smaller for the balanced data set than for the observed unbalanced data set. However, the choice of validation index had an even greater impact (up to 15 000 %). This indicates that studies using opportunistic data for SDM fitting need to pay attention to the effects of presence/absence data imbalance and the choice of validation index to fully evaluate uncertainty.

Keywords

Pagellus bogaraveo, species distribution models, ensemble modelling, heterogeneous data set, presence-absence imbalance

---

*Corresponding author: lola.decubber@gmail.com

1. Introduction

Actual and potential areas of species distribution can be investigated via ecological niche modeling (Soberon and Nakamura, 2009). A species' niche is defined as a subset of environmental conditions under which populations of a species have positive growth rates (Soberon and Nakamura, 2009). The habitat is then the geographical translation of these environmental conditions. The fundamental niche is the theoretical combination of environmental variable that allows for physiological processes (feeding, growth, reproduction) to take place (Hutchinson, 1978). Essential fish habitats, defined as areas or volumes of water and bottom substrates that provide the most favourable habitats for fish populations to spawn, feed and mature throughout their full life cycle, are thus the geographical translation of the optimal part of the fundamental niche of a species (Helaouet and Beaugrand, 2009; Valanis et al., 2008). The realised niche is the subset of environmental conditions the species is actually using (Soberon and Nakamura, 2009). Species may occur outside the fundamental niche during migrations. In contrast, the realised niche might be reduced when densities are low because of intensive predation or fishing (Helaouet and Beaugrand, 2009). The realised habitat of a species can then be defined as the geographical translation of the realised niche of a species. It differs from the species' distribution since all locations displaying the environmental conditions of the realised niche might not be occupied simultaneously, especially if the species' distribution is wide.

Species distribution models (SDMs) have been used in conservation biology to describe the habitat distribution of organisms in both marine and terrestrial systems (Laman et al., 2018; Elith & Leathwick, 2009; Valanis et al., 2008). They are grounded in the concept of ecological niche (Hutchinson, 1957). They have been widely used since 2005 and have reached high statistical sophistication in recent years (Schickele et al., 2020; Jiménez & Soberón, 2020; Robinson et al., 2017). Ecological assumptions implied when using SDMs are that there is niche conservatism (Crisp et al., 2009) and unlimited dispersal abilities (Wiens et al., 2009) and that biotic interactions do not influence large-scale distributions (Gleason, 1926; Guisan and Thuiller, 2005; Wiens et al., 2009; Schickele et al, 2020). Among the numerous statistical SDMs approaches developed to map fish habitats, ensemble species distribution modelling (eSDM), also referred to as ensemble niche modelling (Thuillier et al, 2016), which combines the use of several SDM categories, appears to be a good compromise in terms of programming skills required, computation time and consistency of the results (Schickele et al., 2020; Mateo et al., 2009).

Data availability is often opportunistic, so that neither the fundamental habitat nor the realised species' habitat is entirely represented by SDMs. Indeed, the theoretical entire range of fundamental environmental conditions of a species is never fully known and available presence records will never cover the full habitat. Ecologists thus generally refer to SDM output as potential niche and habitat of the species of interest (Schickele et al., 2020; Helaouet and Beaugrand, 2009). Several data filtration and selection processes, as well as physiological prospects (for example, optimal environmental ranges for spawning or egg development) can then help approaching the species' realised or essential habitat (Schickele et al., 2020; Helaouet and Beaugrand, 2009).

Implementing SDMs, especially in the case of widely distributed species such as the blackspot seabream *Pagellus bogaraveo* (Brünnich, 1768), often requires combining heterogeneous multiple data sets (Schickele et al., 2020; Fithian et al., 2015). In the case of presence/absence data, two types of biases have then to be taken into account. First, detectability might vary among sampling techniques used to collect data (Kellner and Swihart, 2014). Second, variations in prevalence (i.e. the number of presence records among sampled points) might reflect primarily variations in abundance rather than habitat suitability. When data are missing on the detection probability of sampling techniques, taking into account detection might not always improve SDM performance (Welsh et al., 2013), and these two biasing effects (detectability and variations in prevalence) might be difficult to disentangle. In the case of presence-only data, a common practice is to generate pseudo-absence data (Schickele et al., 2020). In this case, the number of generated pseudo-absences is generally set equal to the number of presences (Montgomery, 2005). For actual presence-absence data, prevalence will vary in space, in particular for large study areas. This raises questions given spatial predictions from SDMs are known to be sensitive to sample prevalence (Jimenez-Valverde et al., 2021).

The general aim of this study was to investigate the potential habitat of the blackspot seabream and its occupancy level in three regions in the Northeast Atlantic: Atlantic European shelf, the Azorean region and the Mediterranean Sea. Occupancy levels were presumed to differ between regions because of the contrasted population status and variable degree of fishery exploitation. To evaluate the impact of heterogeneous prevalence in the data and obtain robust results we compared eSDM models using 1) all available presence/absence records, i.e. prevalence varying over the distribution area of blackspot seabream; 2) the same number of presence and absence record, i.e. constant prevalence over the distribution area.

2. Material and methods

2.1. Case study

The blackspot seabream used to be a widely distributed and abundant species of the North Eastern Atlantic shelf from the Faroe Islands down to Gibraltar, the Azores and the Western Mediterranean Sea (Desbrosses, 1932; Sanz-Fernandez et al., 2019; Pinho et al., 2014; Erzini et al., 2005; Spedicato et al., 2002). In fact, it was also referred to as "la dorade commune" (understand "the common seabream") by French authors in the early 1900s (Desbrosses, 1932; Olivier, 1928). Increase in fishing effort in the Bay of Biscay (North Eastern Atlantic shelf) in the 1960s linked to stock declines of other species of fisheries interest such as hake, associated to its susceptibility to overexploitation, led to a brutal collapse of this blackspot seabream stock 20 years later in 1975-1985 and low stock size ever since (see Fig.1a, Lorance, 2011; Guichet et al., 1971; Dardignac, 1988).

Blackspot seabream displays three characteristics that make it susceptible to over-exploitation (Francis and Clark, 2005). First, its biological productivity is low, individuals reaching 70 cm long in 25 to 30 years and females being mostly the older individuals since the species is hermaphroditic protandrous, with changing sex from male to female (Guéguen, 1969, Lorance, 2011). Second, blackspot seabream is easy to capture during its seasonal migrations because of its aggregative behaviour (Afonso et al., 2012; 2014).

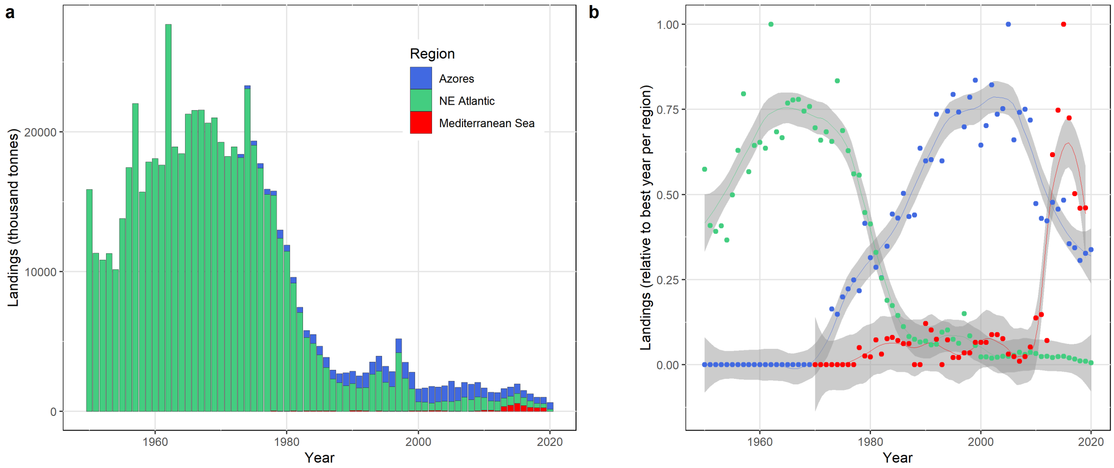

Figure 1: Commercial landings of blackspot seabream between 1950 and 2020 in the three regions investigated in this study as total catch (a) and relative to the highest year per region (b), where dots are individual values and lines are smoothed time trends with 95% confidence intervals. Data before 2000 from Lorance (2011) and from ICES and FAO catch statistics thereafter.

Indeed, adults carry out geographic and depth seasonal migrations from coastal waters, where they reproduce and where juveniles are found, down to 700 m and up to several hundreds of km away from the coast (Morato et al., 2001; Mytilineou et al., 2005). Accordingly, in the Bay of Biscay, individuals present to the West of Brittany (48°N) were found to overwinter in the Cantabrian Sea (43°N) (Guéguen, 1974). In Azorean waters, juveniles which are only found in coastal areas migrate to isolated seamounts when reaching adult stage, sometimes more than 400 km away (Hareide and Garnes, 2003). Lastly, the species has a high commercial value owing to its organoleptic quality, comparable to gilthead seabream and sea bass (Rincon et al. 2016).

Overall, available stock assessments and landings of commercial fisheries suggest that populations from the European shelf are at low level with small recent catches (2018-2021) compared to past levels (anterior to 1980), while populations from the Azorean region are in better condition with current fisheries being sustainable (ICES, 2021, Fig. 1). Historically (before 1980), catches from the Northeast Atlantic shelf constituted the bulk of landings, reaching up to more than 20000 t per year (Fig. 1a), with the Bay of Biscay being the main fishing area. In comparison, levels of gilthead seabream catches were similar back then, while currently, hake is the most fished species in the Bay of Biscay with around 30000 t landed per year, followed by monkfish (8000 t per year), sole (3000 t per year) and seabass (2000 t per year) (Official Nominal Catches 2006-2019. Version 15-10-2021. Accessed 05-05-2022 via https://ices.dk/data/dataset-collections/Pages/Fish-catch-and-stock- assessment.aspx., ICES). From the 1990s, catches from the Northern Atlantic came mostly from the Iberian coast and the Strait of Gibraltar and were at similar level than catches from the Azorean area (Fig. 1a). Reported catches from the Mediterranean Sea are probably not realistic, because in this region 5000 to 10000 tonnes of fish have been landed as unidentified sparid fish or similar labelling and this might have comprised catch statistics of blackspot seabream (FAO-GFCM, 2021). Therefore, the increased reported landings in recent years (Fig. 1b) may be due to improved reporting of landings by species. Quotas as well as other management measures such as minimum landing size and closed fishing seasons are implemented in all areas (Pinho et al., 2014; Lorance, 2011). Indeed, fishing has been shown to be the main factor accounting for variations in the species' stock abundance, with values reaching up to 73% of the variations in stock abundance nowadays around Gibraltar (Sanz-Fernandez et al., 2019). In the case of poor stock status, the species distribution might contract within its essential habitats, that thus needs to be identified to enable targeted conservation management measures to be implemented.

2.2. Data

2.2.1. Species observations

Presence/absence records of blackspot seabream were compiled from trawling and longline scientific surveys (EVHOE, SP-NORTH, SP-ARSA, PT-IBTS, MEDITS, ARQDAÇO) available on the DATRAS portal (https://datras.ices.dk/Data_products/Download/Download_Data_public.aspx) or held by national research Institutes, from commercial fisheries data from the Voracera fleet in Gibraltar and from on-board observations of fishing activities in the Bay of Biscay and the Mediterranean Sea, as well as from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/) (see Fig.2).

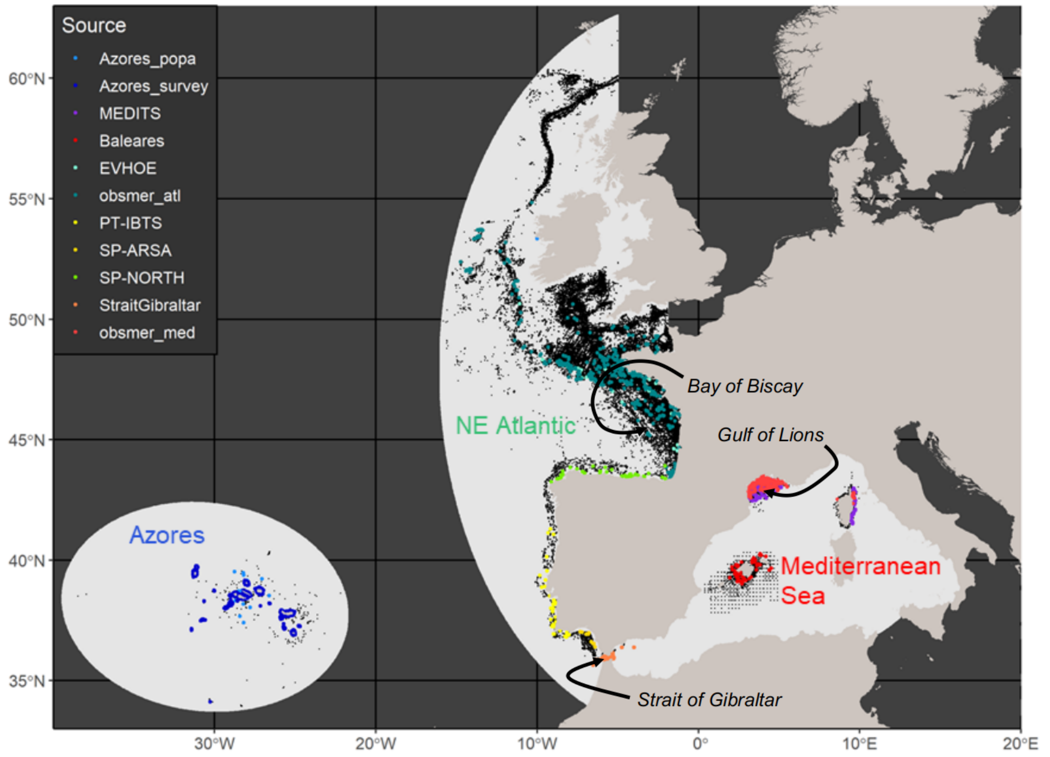

Figure 2: Presence-absence data for blackspot seabream compiled in this study. Black dots represent absence records. Coloured dots correspond to presence records from different data sets: Azores_popa (Fisheries Observer Program, 1998-2013), Azores_survey (ARQ-DAÇO longline survey, 1996-2013), Baleares (Marine Biodiversity Atlas of the Balearic Sea, GBIF, 2001-2008), EVHOE (scientific bottom-trawl survey, 1997-2019), MEDITS (scientific bottom-trawl survey, 2004-2019), obsmer_atl and obsmer_med (French onboard observation program), PT-IBTS (scientific bottom-trawl survey, 2002-2017), SP-ARSA (scientific bottom-trawl survey, 1996-2019), SP-NORTH (scientific bottom-trawl survey, 2001-2019), StraitGibraltar (Commercial fisheries data, 2009-2011). light grey area is the model domain composed of three regions: Azores, NE Atlantic region and Mediterranean Sea.

2.2.2. Environmental data

Environmental variables consisted of topographic data, sea bottom type and seawater parameters. We extracted bathymetry at a 0.0003° resolution from GEBCO (https://www.gebco.net /data_and_products/gridded_bathymetry_data) and the R terrain function (raster package, Hijmans et al., 2011) enabled the calculation of bottom slope. Seabed habitat data were extracted from EMODnet (https://www.emodnet.eu/en/seabed-habitats) at a 250 m resolution. Homogenization of substrate type according to EMODnet categories among all regions led to 14 sea bottom type categories: unknown, rock or other hard substrata, coarse substrate, coarse and mixed sediment, mixed sediment, sediment, sand, sandy mud, muddy sand, sandy mud or muddy sand, fine mud or sandy mud or muddy sand, fine mud, *Posidonia oceanica*, and dead mattes of *Posidonia oceanica*. Monthly values of Sea Surface Temperature (SST), bottom temperature, surface current velocity and salinity between January 1994 and December 2018 at a 0.083° resolution were extracted from Copernicus Marine Service (GLOBAL_REA NALYSIS_PHY_001_030 product, https://resources.marine.co pernicus.eu/). An overview of environmental variables investigated is provided in Sup. Mat. 1. The mean, maximum, minimum and standard deviation of environmental variables were computed for each grid cell ($n = 6465$).

2.3. Methods

The general workflow used in this study is presented in Fig. 3. Model reporting was done following recommendations by Zurell et al. (2020).

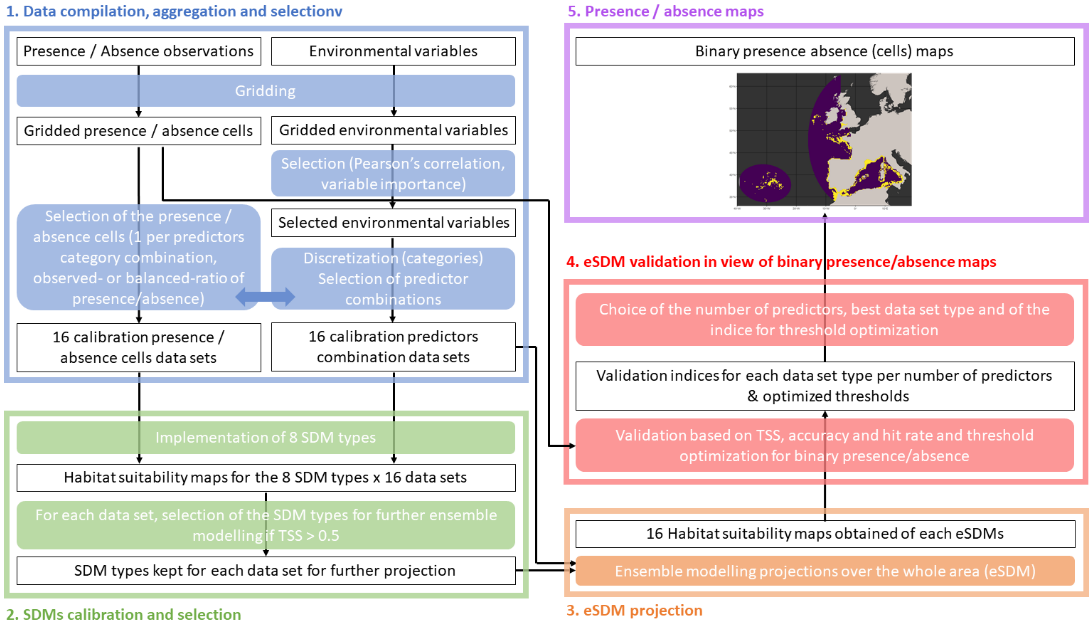

Figure 3: General workflow of the procedure used for identifying blackspot seabream habitats: (1) compilation, gridding and selection of environmental and occurrence data, (2) statistical Species Distribution Models (SDM) calibration (for each category) and selection, (3) Projection of the ensemble (eSDM) model (habitat suitability map), (4) Validation of projections based on maximization of various indices and determination of threshold value for (5) binary presence/absence mapping.

2.3.1. Study area

Three regions were modelled corresponding to the three main areas of distribution of the blackspot seabream: The Northeast Atlantic shelf, the Azorean region and the western Mediterranean basin (Fig. 2). A spatial grid of 0.1° x 0.1° cells (44851 in total) was created covering the three regions. This resolution was chosen accounting for the need for local and general overviews of species habitat, as well as computation time.

2.3.2. Occurrence data compilation, gridding and selection

For each grid cell with data (n = 6465), the number of presence and absence records was counted. The occurrence data set displayed over- and under-sampled areas (Fig. 2). In order to take into account this heterogeneity, occurrence records were compiled into presence/absence records: cells with at least one presence were considered as presence cells, cells with only absence records were considered as absence cells and the remaining cells were considered as non-sampled cells and were not used for model fitting. In order to maintain an homogeneous distributions of sampled cells along environmental gradients, continuous environmental variables were discretised into 60 classes each (which appeared to be a good compromise to create relevant classes for all environmental variables) over the whole grid domain. When a given combination of environmental variable classes corresponded to several presence cells, only one presence cell was kept, similarly for absence cells. When a given combination corresponded to several presence and absence cells, one presence and one absence cell was kept. Thus the number of data points used for model fitting depended on the environmental predictors included in a particular model (Table 2).

Since the ratio between the number of presence and absence grid cells still varied among regions after the gridding process, two approaches were tested. In the first case (observed-ratio data set), all available presence / absence grid cells were kept in the next steps of the analysis. In the second case (balanced-ratio data set), the same number of absence and presence grid cells was used for all regions by randomly selecting absence grid cells among all available absence cells. The second approach corresponds to the common practice for presence-only data for which pseudo-absence data are created (Montgomery, 2005)(See Fig. 3, step 1).

2.3.3. Selection of environmental predictors

Given the observation of heterogeneous responses to some of the environmental variables according to the region (Azores, Atlantic shelf and Mediterranean Sea, see Sup. Mat. 2), a categorical predictor was added for region. To identify the most parsimonious environmental data set explaining the blackspot seabream distribution, pairwise correlations between all environmental variables were investigated with a Pearson's correlation test using as correlation threshold $r > 0.7$ (Schickele et al., 2020; Dormann et al. 2012). When several environmental variables were highly correlated, we retained the environmental variable with the highest relative importance (Schickele et al., 2020; Leroy et al., 2014)(Sup. Mat. 3). As the previous selection step led to a high number (9) of remaining environmental predictors for both occurrence data sets, multiple models were fitted with decreasing number of predictors (9 to 2 predictors), removing sequentially the predictor with the smallest relative importance to balance model fit and model complexity (Meynard et al., 2019) (See Fig. 3, step 1). Next, for each of the eight combinations of predictors and each occurrence data set type (observed-ratio and balanced-ratio), one occurrence data set was created for calibration (16 data sets).

2.3.4. Species distribution model categories and settings

The following eight SDM categories were implemented in R using the BIOMOD2 package (Thuillier et al., 2016) with default parameter settings: generalised linear model (GLM), generalised boosting model (GBM), generalized additive model (GAM), artificial neural network (ANN) model, flexible discriminant analysis (FDA), random forest (RF), classification tree analysis (CTA) and surface range envelope (SRE) model (Valanis et al., 2008; Thuiller et al., 2009; Albouy et al., 2012; Clairbaux et al., 2019; Pecchi et al., 2019). For each SDM category and data set, a 3-fold cross validation procedure was performed.

2.3.5. SDM category selection and compromise

We used the True Skill Statistics index (TSS, Allouche et al., 2006) to quantify the performance of each of the eight fitted SDMs for each calibration data set (See Fig. 3, step 2). It was calculated as TSS = sensitivity (proportion of presence cells classified correctly) + specificity (proportion of correct absence cells) - 1. A SDM was selected for ensemble modelling if it had TSS > 0.5. A compromise (the ensemble Species Distribution Model, eSDM) of the presence probability was then calculated as the mean of probabilities of retained SDMs weighted by their TSS value. Presence probability uncertainty was quantified using the coefficient of variation from cross-validation results. Individual SDM response curves to environmental predictors as well as the eSDM resulting presence probabilities according to the different predictors (Schickele at al., 2020) are presented in Sup. Mat. 4 and 5.

2.3.6. Binary habitat maps and predictors selection

To define habitat suitability maps, for each of the 16 calibrated ensemble models (2 data sets x 8 predictors combination with 2 to 9 predictors), presence probabilities were projected over the whole domain, including the cells not included in the calibration process (extrapolation for the non sampled cells). Then, in order to create binary presence/absence maps, the habitat suitability maps (presence probabilities) were compared to all observed presence and absence compiled cells (not only those used for model fitting). Threshold values for binary projections of presence and absence were calculated based on a set of indices as recommended by Robinson et al. (2017), namely the hit rate (proportion of correctly classified presence cells), the True Skill Statistics (TSS, Allouche et al., 2006), the CBI (computed with the ecospat.boyce function of the ecospat package on R, Hirzel et al. 2006), and the overall accuracy (sum of the proportion of correctly classified presence and absence cells, Allouche et al., 2006). More precisely, each validation index was maximized varying threshold values for binary projections of presence and absence (from 0 to 1 with a 0.001 interval) with the optimize function in R. Each maximized index value thus corresponded to a distinct optimized threshold value.

For each data set type, the best calibrated eSDM was chosen as the one showing the lowest number of predictors and the highest validation index values. To combine validation index values, their values were re-scaled between 0 (lowest index value obtained across eSDM outputs with 2-9 predictors) and 1 (highest index value). In the case of the balanced-ratio data set, as absence cells selection might have impacted ensemble model performance in the previous steps, 10 data sets with the chosen number of predictors and varying absence cells were randomly selected and the average of habitat suitability maps (re-calibration) was used for further validation. Lastly, for each data set type and validation index, the surface of potential habitat was calculated.

3. Results

3.1. Data characteristics

In total, 106 457 occurrence records were compiled, among which 6465 presence records, corresponding to 782 cells where the species was present and 5683 cells where it was recorded as absent (Fig. 2, Table 1).

Table 1: Summary of blackspot seabream occurrence data by region. Domain per region as in Fig. 2. Records correspond to point observations of presence or absence of the species while cells correspond to grid cells of the domain where one or several presence or absence point observation was made. One presence observation was sufficient to qualify as presence cells.

	Total area			Atlantic			Azores			Mediterranean		
	records	cells	rec./cell	records	cells	rec./cell	records	cells	rec./cell	records	cells	rec./cell
N presence	6928	782	8.8	639	389	1.6	4872	165	29.5	1417	228	6.2
Proportion		1.7 %			1.7 %			1.8 %			1.9 %	
N absence	99529	5683	17.5	74556	4626	16.1	13906	376	37	11067	681	16.3
Proportion		12.7 %			20 %			4%			5.6 %	
N sampled	106457	6465	16.5	75196	5015	15	18778	541	34.7	12484	909	13.7
Proportion		14.4 %			21.6 %			5.8 %			7.4%	
N presence / N sampled	6.5 %	**12 %**		0.8 %	**7.8 %**		25.9 %	**30.5 %**		11.4 %	**25 %**	
N cells per region		44851			23209			9392			12250	

The proportion of presence cells among sampled cells was highly heterogeneous between regions, as well as the number of records per cell (respectively 7.8 to 30.5 % and 14 to 35 records per cell, see Table 1).

3.2. eSDM outputs and habitat suitability maps

After predictor selection using pairwise correlation analysis, nine predictors were retained for the ensemble modelling procedure: region (NE Atlantic, Azores, Mediterranean Sea), sea bottom type, mean and standard deviation of bathymetry (m), annual maximum Sea Surface Temperature (SST, °C), annual mean and minimum absolute current velocity ($m.s^{-1}$), annual standard deviation of salinity (‰) and annual standard deviation of bottom temperature (°C) (Sup. Mat. 3). Mean bathymetry and maximum SST had the highest explanatory power for both data sets, both variables contributing equally (Table 2). The main difference between ensemble models for the two data set types was that for the observed-ratio data set (heterogeneous ratios of presence and absence over the model domain), the categorical predictor "region" had a high relative importance, which was not the case for the balanced-ratio data set (balanced number of presence and absence grid cells in each region) (Table 2, Sup. Mat. 6).

SDM categories ANN, GBM and RF were selected in most cases ($TSS > 0.5$), while the selection of the other SDM categories varied according to the number of predictors and the data set type used (Table 2, Sup. Mat. 6). TSS values were higher for the observed-ratio data set for all models (two to nine predictors), with values around 0.7 for the balanced-ratio data set, and around 0.9 for the observed-ratio data set (Table 2, Sup. Mat. 6). As a result of the data selection procedure, the data set size decreased with decreasing number of predictors, leading to low data set sizes for models with less than five predictors (Table 2). Comparing projected habitat suitability maps (presence probabilities per grid cell over the whole domain) between data set types, it appeared that a balanced ratio between presence and absence data led to overall higher presence probabilities and hence a much wider potential habitat in the NE Atlantic region and in the Mediterranean Sea region compared to the results obtained with the observed-ratio data set, while the projected presence probabilities appeared rather similar for the Azores region (Figs. 4 and 5). The visual difference was confirmed by the mean projected presence probability over the whole domain being 0.16 ± 0.19 for the balanced-ratio data set and only 0.07 ± 0.09 for the observed-ratio data set (Fig. 4).

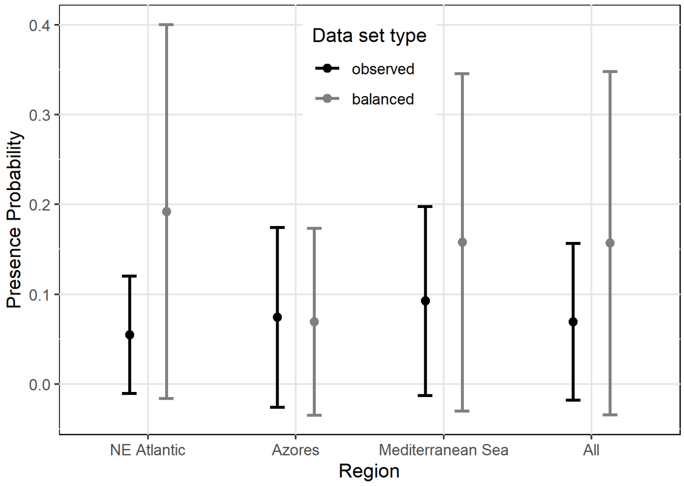

Figure 4: Presence probabilities (mean and sd) of the blackspot seabream per region and on the whole grid according to the data set type used for ensemble species distribution modelling with 6 predictors.

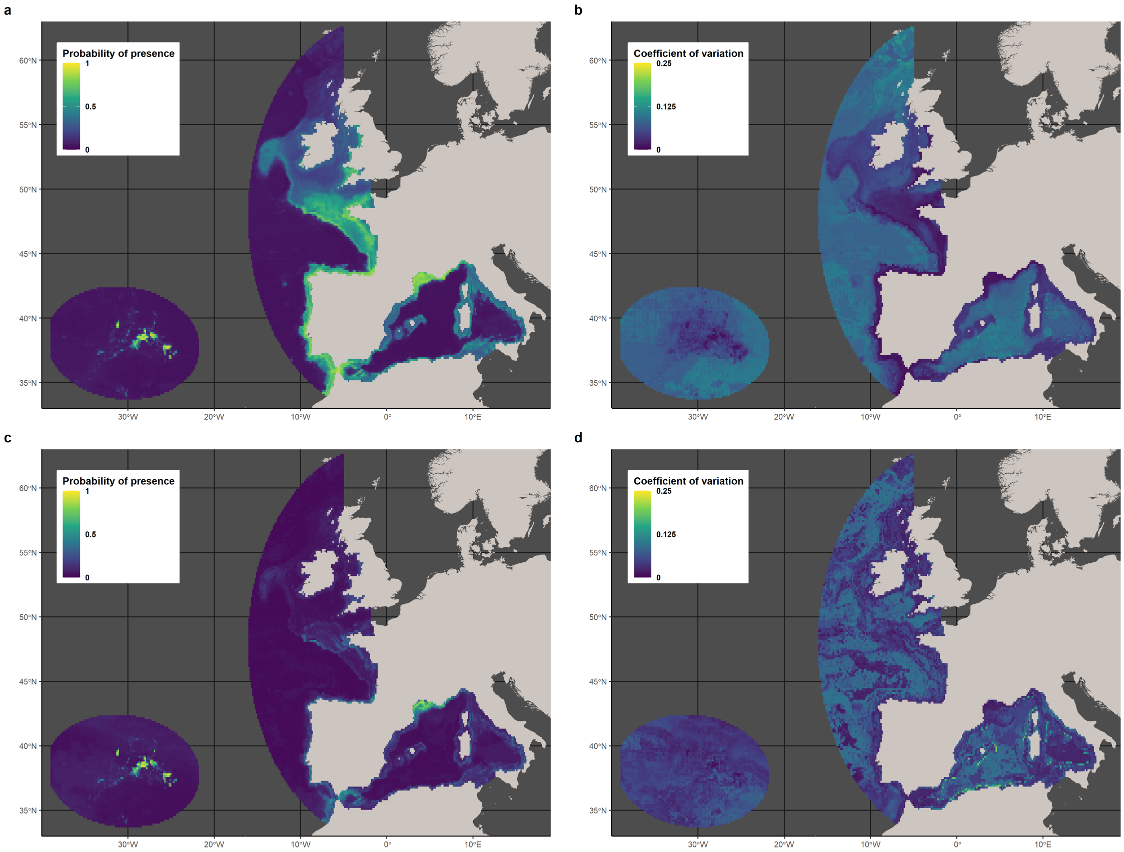

Figure 5: Habitat suitability maps for blackspot seabream for the main distribution area (a, c) and associated coefficients of variation (b, d) obtained with an ensemble species distribution model performed using a data set with a balanced number of presence and absence data (balanced-ratio, a, b) and a data set with the observed and heterogeneous number of presences and absences (observed-ratio, c, d), for models including six predictors (see Table 2 and Sup. Mat. 6).

Table 2: Summary of the 16 calibrated ensemble species distribution models implemented in this study with 8 different numbers of predictors for the observed-ratio data set type: predictors used and their relative importance (proportion of explained variance), characteristics of the data set used, SDMs included in eSDM and overall performance.

	2 predictors	3 predictors	4 predictors	5 predictor	6 predictors	7 predictors	8 predictors	9 predictors
Environmental variables' importance								
Mean bathymetry	50%	49 %	50%	37%	30%	31%	26%	27 %
Max SST	50%	34%	18%	25%	31%	25%	27%	26%
Sd bathymetry	-	17%	18%	20%	18%	16%	12%	11%
Sd bottom temperature	-	-	-	8%	8%	9%	7 %	8%
Sd salinity	-	-	-	-	5%	6 %	5 %	5%
Min current velocity	-	-	-	-	-	3%	4%	2%
Mean current velocity	-	-	-	-	-	-	-	2 %
Sea bottom type	-	-	-	-	-	-	1%	1%
Region	-	-	11%	13 %	9%	10%	18%	17%
Occurrence data set								
Data set size	474	1510	1696	2620	3280	3357	3752	4163
Number of presence cells	8	161	161	372	470	475	536	576
Number of absence cells	466	1349	1535	2248	2810	2882	3216	3587
SDM types and eSDM performance								
Selected models (TSS > 0.5)	ANN, FDA, GAM, GBM, GLM, RF	ANN, CTA, FDA, GAM, GBM, GLM, RF	ANN, CTA, FDA, GBM, GLM, RF, SRE	ANN, GBM, GLM, RF	ANN, GBM, RF	ANN, GBM, RF	ANN, GBM, RF	ANN, GBM, RF
True Skill Statistic (mean ± sd across SDMs)	0.996 ± 0.003	0.791 ± 0.013	0.826 ± 0.013	0.867 ± 0.033	0.916 ± 0.030	0.895 ± 0.039	0.904 ± 0.043	0.894 ± 0.053

Differences between projected presence probabilities obtained with the balanced-ratio data set and the observed-ratio data set were the highest in the NE Atlantic (with respective values of 0.19 ± 0.21 and 0.06 ± 0.07) and the lowest in the Azores (with both values equal to 0.07 ± 0.10) (Fig. 4). For both data set types, coefficients of variation of presence probabilities did not exceed 0.25 (Fig. 5).

Overall, the choice of data set type impacted presence probability values, but did not impact consistently which predictors were selected, except for the region predictor, nor the general shape of the response curve for each predictor, except for maximum annual SST for which the relationship was dome-shaped for balanced-ratio data set and more flat for the observed-ratio data set (Table 2 and Sup. Mat. 4).

3.3. Binary habitat maps

3.3.1. Predictors

Comparison of the three validation index values between eSDMs using 2 to 9 predictors normalized between 0 an 1 (0 being the smallest index value across the set of of predictors and 1 being the highest) led to select the eSDM with six predictors for both data set types. Indeed, these models displayed among the best index values for the smallest number of predictors (Fig. 6).

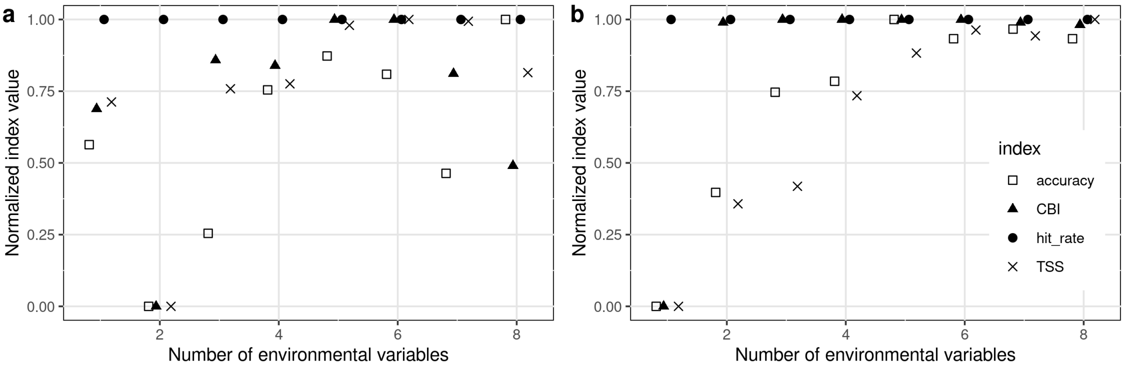

Figure 6: Normalized values of validation indices (accuracy, Allouche et al., 2006; hit rate, Allouche et al., 2006; True Skill Statistics, Allouche et al., 2006, Continuous Boyce Index, Hirzel et al. 2006) of the models according to the number of predictors used, for the balanced-ratio data set type (same number of absences and presences cells) (a) and the observed-ratio data set type (all records) (b).

The best predictors for the balanced-ratio data set were mean bathymetry, annual maximum SST, standard deviation (sd) of bathymetry, sd of mean annual bottom temperature, sd of mean annual bottom salinity and minimum annual absolute current velocity. For the observed-ratio data set, predictor region was selected instead of minimum annual absolute current velocity (Table 2, Sup. Mat. 6). The region predictor had a 9% contribution to the explained variance and, compared to the balanced-ratio data set, contribution of other variables was lower for bathymetry and higher for all other variables.

3.3.2. Indices and threshold optimization

Validation index values (accuracy, hit rate, CBI and TSS) were slightly higher for the observed-ratio data set than for the balanced-ratio data set (Sup. Mat. 8). Variations in estimated habitat area between validation indices were higher than between data set types (Fig. 7). Overall, the choice of threshold value used for transforming presence probabilities into binary habitat maps strongly influenced results for the whole area and for each of the regions (Fig. 8). For regional habitat area estimation, separate threshold values were obtained by maximising index values regionally. Maximizing the hit rate led to the lowest threshold values for the whole area and each region, and subsequently the largest habitat areas.

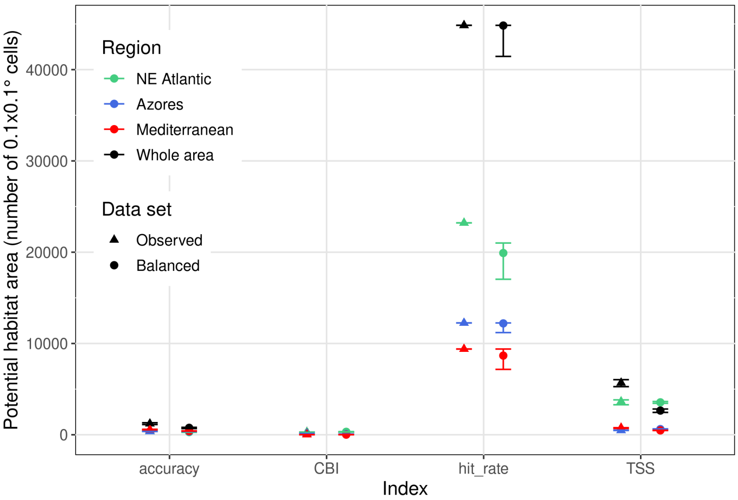

Figure 7: Blackspot seabream potential habitat area using the best model (6 predictors) and different validation indices for setting the presence/absence threshold value: accuracy, hit rate, Continuous Boyce Index and True Skill Statistics. Results for observed- and balanced-ratio data sets, when models were optimized for the whole model domain (black), the NE Atlantic region (green), the Azores region (blue), the Mediterranean Sea region (red).

CBI and accuracy led to the smallest habitat areas, especially for the NE Atlantic region for which the observed presence / absence ratio was low. For the whole area, averaged across the two data types, the habitat areas for hit rate and TSS were 14849% and 1280% larger than for CBI respectively. Comparing results between the two data sets showed that the balanced data set let to smaller habitat areas for the majority of indices for the whole area and each of the three regions (Figs. 7 and 8). For the whole area, the difference ((balanced-observed)/observed) ranged between -53% for TSS and 4% for CBI. For the Azores the difference ranged between -3% for accuracy and 27% for TSS, for the Mediterranean Sea between -36% for TSS and 0% for CBI, and for the NE Atlantic region between -28% for accuracy and 23 % for CBI. Thus, the choice of validation index had a much greater impact than the data set type, though the later was also important for certain indices.

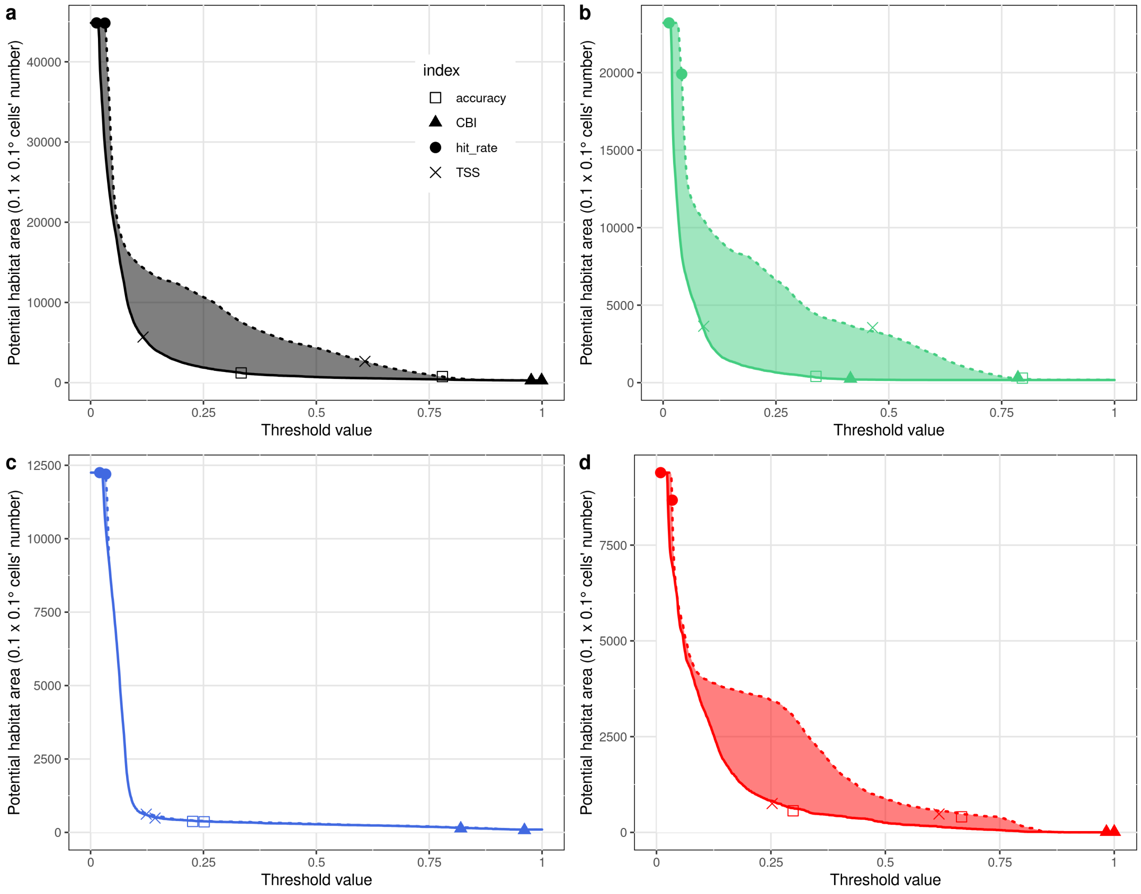

Figure 8: Area of the potential habitat of the blackspot seabream as a function of the threshold value used for binary presence/absence predictions for the balanced-ratio data set (dotted line) and the observed-ratio data set (solid line) and results obtained using different validation indices (accuracy, hit rate, CBI and TSS) for the whole area (a), NE Atlantic region (b), Azores region (c) and Mediterranean Sea region (d).

3.3.3. Binary habitat maps

Visual comparison of binary (presence / absence) habitat maps (threshold optimized according to TSS index) between data set types showed that for the balanced-ratio data set some locations from which the species had been reported were missed, particularly to the West of Ireland and along the Mediterranean coast (Figs. 2 and 9). Overall, it appears that the potential habitat of the species covers a large area around seamounts in the Azores region, a wide area on the NE Atlantic shelf and a narrower are on the Mediterranean shelf (Fig. 9).

4. Discussion

In this study, to obtain robust estimates of the potential habitat of blackspot seabream across its wide distribution area exhibiting varying exploitation status, an ensemble species distribution modelling approach and two data sets with different prevalence levels were used. The identified potential habitats included islands contours and seamounts in the Azores region, the NE Atlantic shelf south of 48°N, with smaller areas further North, and the Northern shores of the western Mediterranean Sea, with more extended areas in the Strait of Gibraltar, in the Gulf of Lions and along the Italian coast.

Potential habitats of blackspot seabream were best explained by bathymetry (down to 700 - 1000 m) and SST (annual maximum SST generally greater than 16°C), as well as bottom temperature, salinity and region as secondary predictors, independent of the prevalence level in the data set except for region. Further, the general shape of the response curves for each predictor were similar for the two data sets except for max SST (see response curves in Sup. Mat. 4). Other studies have reported the same main factors influencing the species' distribution and abundance, with occurrences reported in areas with bottom depths between 100 m and 700 m (Santos et al., 2019; Mytilineou et al., 2014; Burgos et al., 2013; Menezes et al., 2013; D'Onghia et al., 2010; Gueguen, 1974) and environmental conditions linked to temperature and salinity influencing stock variations (Sanz-Fernandez et al., 2019).

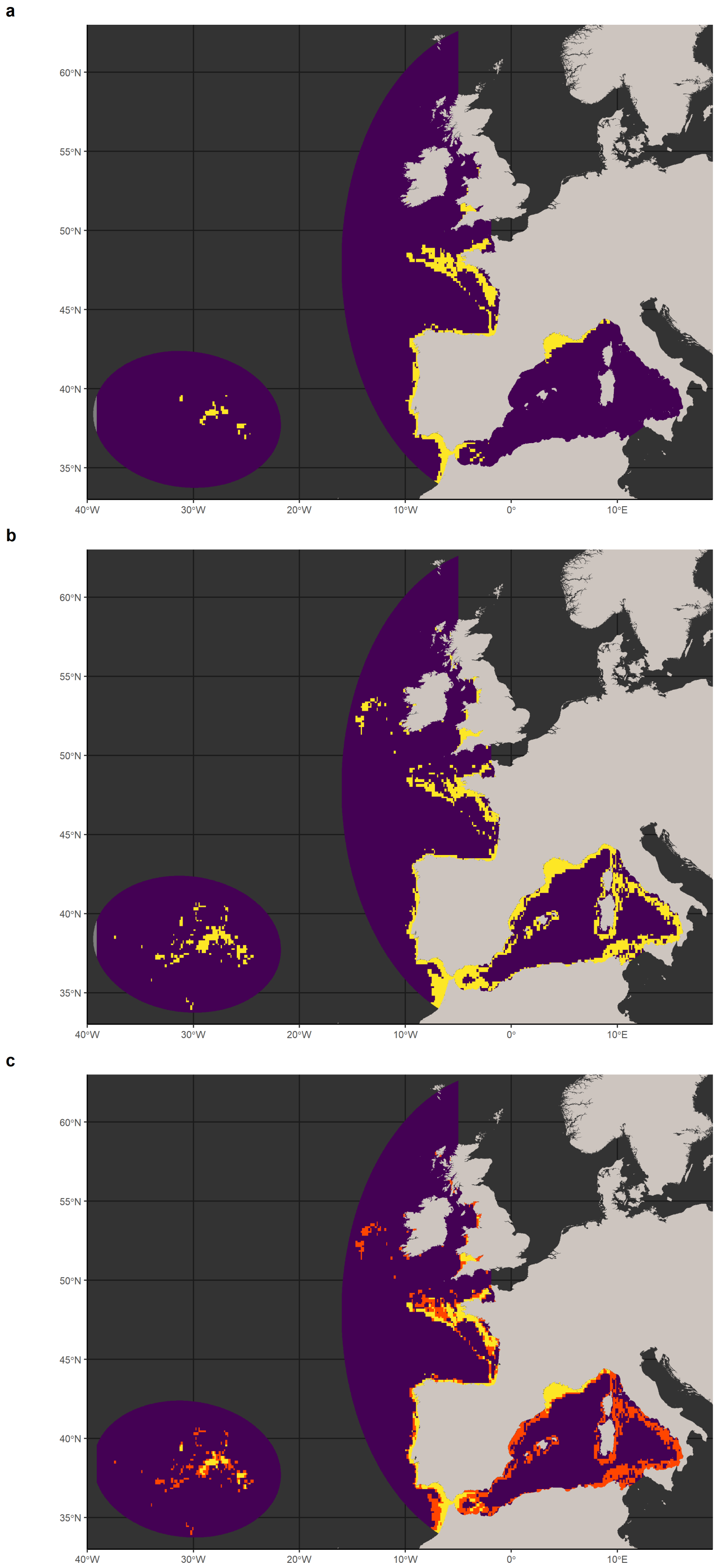

Figure 9: Map of estimated potential habitats with six predictors (yellow) for blackspot seabream over its whole area of distribution using the balanced-ratio data (a) or the observed-ratio data set (b) maximizing the TSS index for presence threshold estimation, and difference between the two potential habitat maps (c, red).

è The general importance of the region predictor can be explained by regional differences in biophysical conditions while the difference in importance of this predictor for the two prevalence data sets is most likely caused by differences in the prevalence in the observed-ratio data sets (Table 2). These regional discrepancies in prevalence might be explained by the differences in population status, with the Northeast Atlantic population being most depleted. It is less likely caused by differences in detectability by the employed sampling methods, which could however be explored by comparing several methods for the same area and using models integrating detectability (Fithian et al., 2015; Fletcher et al., 2016; 2019).

In the Bay of Biscay, which is in the center of the Northeast Atlantic region, the eSDM suggested that the potential habitat covered a large part of the continental shelf. Indeed, the species used to occur over most of this shelf before the population collapsed in the late 1970s. In the Azorean region, which has been termed an oceanic seamount ecosystem (Silva and Pinho, 2007), the species occurs around islands, with juveniles distributed in near coast habitats as in the Bay of Biscay and adults spread from the coast to island slopes and isolated seamounts; juveniles never occur at sea mounts (Pinho et al., 2014). Depth was identified as the main factor for explaining blackspot seabream abundance and size composition on Azorean seamounts (Santos et al., 2021). In the same area, Morato et al. (2001) showed that the species feeds on both pelagic and benthic preys and suggested that its food supply on seamounts may depend upon oceanic production that drifts past seamounts, which in turn makes bottom slope a factor for the species distribution as current strength and therefore the amount of prey advected increase with slope. At the scale of the environment perceived by individual fish, similar hydrological conditions, in particular strong and variable tidal currents resulting from the interaction of oceanic water masses (Koslow, 1996; Lorance et al. 2002) prevail at seamounts and along the continental slope, where blackspot seabream also occurs. At this fine scale, the species shows diel vertical migrations (Afonso et al., 2012), which may suggest a behavior similar to that of deeper living "seamount aggregators" which share with blackspot seabream a high lipid content and strong swimming abilities, which are both related to high metabolism (Koslow, 1996). Overall, the habitat characteristics of fish displaying this type of behavior include sloping sea bottom and related variations in current speed, which generate also temperature variations and is in-line with habitats variables found to be predictors of *blackspot seabream* presence-absence in this study.

The ratio between presence and absence cells in the data impacted strongly the projected probabilities of presence. Imposing a balanced-ratio between the number of presence and absence cells for model calibration led to higher presence probabilities on average compared to using the observed-ratio data set with a large majority of absence cells (0.16 ± 0.19 and 0.07 ± 0.09 respectively). Thus, including primarily absence cells in model calibration reduced estimated presence probabilities.

The habitat suitability map for the balanced-ratio data set indicated wide potential habitats over the NE Atlantic shelf and along the western Mediterranean and Azores coasts, while potential habitats were smaller using the observed-ratio data set. The habitat size was reversed between the two data sets for the corresponding presence/absence maps, with larger binary habitats for the observed-ratio data set due to a smaller optimized threshold value. These binary maps obtained by applying a optimized presence probability threshold are probably closer to the realised habitat of the species as it involved the full presence/absence data to set the threshold values. The binary presence habitat was notably smaller over the European shelf, where one regional population is depleted. The past high abundance of the species throughout the Bay of Biscay (Olivier, 1928; Desbrosses, 1932; Guichet et al., 1971) indicates that this area was suitable for the species 100 to 50 years ago. The increase of fishing effort on the species at the same period (Lorance, 2011) must indeed have been the main trigger for the species' decreased abundance, leading to a concentration of the remaining individuals into the most suitable habitats in this region (along the coast around isobath -100m and on the continental slope). The present study suggests that current environmental factors remain suitable for the species, although environmental changes have occurred since, notably a temperature rise of 0.2 °C/decade for the period 1965 - 2004 in the 0 - 200 m water column layer (Michel et al., 2009, Valencia et al., 2019). The observed presence-absence ratio data set had higher validation indices for the fitted model compared to the balanced ratio data set. Further, the corresponding binary habitat map included more areas outside the sampled area. For example, it predicted the species' presence close to the West of Ireland, where the species did indeed occur in the past (Guégen, 1974) and more widely along the western Mediterranean coast where it is present nowadays (Spedicato et al., 2002; Lechekhab et al., 2010).

The use of different validation indices (hit rate, TSS, CBI and accuracy) led to different threshold values for transforming the probability of presence into binary habitat and subsequently different habitat surface areas. Overall, the difference in the size of estimated habitat area was larger between validation indices when between data set types. Differences in predicted habitat areas according to the selected thresholding method have been reported by various authors (Nenzen and Araujo, 2011; Jimenez-Valverde and Lobo, 2007; Liu et al., 2005). In our case, given the species' low prevalence, absence cell records had a large impact on the estimated threshold when maximizing using the accuracy index, leading to habitats mostly restrained to the cells where species observations had been made, hence being closer to the (only partial) observed distribution. Although its use has been recommended in the case of unbalanced prevalence (Leroy et al., 2018), the use of the CBI index led to the same effect. In contrast, using the TSS index for setting the probability threshold value involved balancing correct predictions of both presence and absence cells, and hence seemed more likely to lead to binary habitats closer to the species' potential habitats.

The difference between habitat areas derived using TSS optimized threshold values for balanced and observed unbalanced data sets depended strongly on the region used for optimization. The largest negative difference was observed for the whole area (-53%), while it was negligible (-2%) for the NE Atlantic region and positive (27%) for the Azores. Unbalanced prevalence has been reported to artificially increase the TSS value (Leroy et al., 2018), which implies that the balanced data set should have provided a better understanding of blackspot seabream potential habitats. Contrary to this expectation it seems that in our case using a bigger data set with unbalanced presence/absence cells was more informative than using a smaller data set with a selected number of balanced presence/absence cells.

Several studies have shown that the use of presence-true absence data are generally better than presence-only data with or without using pseudo absences, and that presence-only models generally under-estimate the species' presence in locations where is has not been sampled (Dorazio, 2014; Meynard et al., 2019; Wisz and Guisan, 2009). Our study is in accordance with this general result. In addition it showed that the proportion of absences data matters. The difference between the estimated area of the blackspot seabream potential habitat obtained with the observed-ratio data set and the balanced-ratio data set for the same threshold value increased with the decreasing proportion of presence data in the observed data, from the Azores to the NE Atlantic, which in turn corresponds to decreasing stock status of blackspot seabream. In summary, the results of this study provided evidence that when using opportunistic data for SDM fitting attention needs to be paid to the effects of presence/absence data imbalance as well as the choice of validation indices to fully evaluate uncertainty of estimated habitat maps.

Acknowledgement

The study received financial support from France Filière Pêche (project DynRose) and the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 773713 (PANDORA).

References

Afonso, P., Graça, G., Berke, G., Fontes, J., 2012. First observations on seamount habitat use of blackspot seabream (*Pagellus bogaraveo*) using acoustic telemetry. J. Exp. Mar. Bio. Ecol. 436–437, 1–10. https://doi.org/10.1016/j.jembe.2012.08.003

Afonso, P., McGinty, N., Graça, G., Fontes, J., Inácio, M., Totland, A., Menezes, G., 2014. Vertical migrations of a deep-sea fish and its prey. PLoS One 9. https://doi.org/10.1371/journal.pone.0097884

Albouy, C., Guilhaumon, F., Araújo, M.B., Mouillot, D., Leprieur, F., 2012. Combining projected changes in species richness and composition reveals climate change impacts on coastal Mediterranean fish assemblages. Glob. Chang. Biol. 18, 2995–3003. https://doi.org/10.1111/j.1365-2486.2012.02772.x

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: How, where and how many? Methods Ecol. Evol. 3, 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Burgos, C., Gil, J., Del Olmo, L.A., 2013. The Spanish blackspot seabream (*Pagellus bogaraveo*) fishery in the Strait of Gibraltar: Spatial distribution and fishing effort derived from a small-scale GPRS/GSM based fisheries vessel monitoring system. Aquat. Living Resour. 26, 399–407. https://doi.org/10.1051/alr/2013068

Chebel, F., Mezedjri, L., Boulahdid, M., 2018. Reproductive cycle of the gilthead sea bream *Sparus aurata* Linnaeus, 1758 (Pisces Perciformes Sparidae) in the Gulf of Skikda (Algerian East coast). Biodivers. J. 9, 19–24.

Clairbaux, M., Fort, J., Mathewson, P., Porter, W., Strøm, H., Grémillet, D., 2019. Climate change could overturn bird migration: Transarctic flights and high-latitude residency in a sea ice free Arctic. Sci. Rep. 9. https://doi.org/10.1038/s41598-019-54228-5

D'Onghia, G., Maiorano, P., Carlucci, R., Capezzuto, F., Carluccio, A., Tursi, A., Sion, L., 2012. Comparing Deep-Sea Fish Fauna between Coral and Non-Coral "Megahabitats" in the Santa Maria di Leuca Cold-Water Coral Province (Mediterranean Sea). PLoS One 7. https://doi.org/10.1371/journal.pone.0044509

D'Onghia, G., Maiorano, P., Sion, L., Giove, A., Capezzuto, F., Carlucci, R., Tursi, A., 2010. Effects of deep-water coral banks on the abundance and size structure of the megafauna in the Mediterranean Sea. Deep. Res. Part II Top. Stud. Oceanogr. 57, 397–411. https://doi.org/10.1016/j.dsr2.2009.08.022

Dardignac, J., 1988. Les pêcheries du Golfe de Gascogne, bilan des connaissances. Plouzané, IFREMER, 204 pp.

Desbrosses, P., 1932. Poissons de chalut - la dorade commune (*Pagellus centrodontus* Delaroche) et sa pêche. Rev. des Trav. l´Office des Pêches Marit. 5, 167–222.

Dorazio, R.M., 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Glob. Ecol. Biogeogr. 23, 1472–1484. https://doi.org/10.1111/geb.12216

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography (Cop.). 36, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Elith, J., Leathwick, J.R., 2009. Species distribution models: Ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Erzini, K., Salgado, M., Castro, M., 2006. Dynamics of black spot sea bream (*Pagellus bogaraveo*) mean length: Evaluating the influence of life history parameters, recruitment, size selectivity and exploitation rates. J. Appl. Ichthyol. 22, 183–188. https://doi.org/10.1111/j.1439-0426.2006.00702.x

FAO-GFCM, 2021. Fisheries and aquaculture statistics. GFCM Catches 1970-2019 (FishstatJ). In: FAO Fisheries Division [on-line]. Rome. 2021 update. https://www.fao.org/fishery/statistics/software/fishstatj/en

Fithian, W., Elith, J., Hastie, T., Keith, D.A., 2015. Bias correction in species distribution models: Pooling survey and collection data for multiple species. Methods Ecol. Evol. 6, 424–438. https://doi.org/10.1111/2041-210X.12242

Fletcher, R.J., Hefley, T.J., Robertson, E.P., Zuckerberg, B., McCleery, R.A., Dorazio, R.M., 2019. A practical guide for combining data to model species distributions. Ecology 100, 1–15. https://doi.org/10.1002/ecy.2710

Fletcher, R.J., McCleery, R.A., Greene, D.U., Tye, C.A., 2016. Integrated models that unite local and regional data reveal larger-scale environmental relationships and improve predictions of species distributions. Landsc. Ecol. 31, 1369–1382. https://doi.org/10.1007/s10980-015-0327-9

Francis, R.I.C.C., Clark, M.R., 2005. Sustainability issues for orange roughy fisheries. Bull. Mar. Sci. 76, 337–351.

Gueguen, J., 1974. Précisions sur les migrations de la dorade rose *Pagellus Bogaraveo* (Brunnich 1768). Sci. Pêche 237, 1–10.

Guichet, R., Guéguen, J., Guillou, A., 1971. La pêche du merlu et de la dorade à La Rochelle, analyse des statistiques d'effort de pêche et de production des années 1966, 1967 et 1968. Rev. Trav. Inst. Pèches marit. 35, 239–286.

Hareide, N.R., Garnes, G., 2001. The distribution and catch rates of deep water fish along the Mid-Atlantic ridge from 43 to 61 °N. Fish. Res. 51, 297–310. https://doi.org/10.1016/S0165-7836(01)00253-3

Hattab, T., Albouy, C., Lasram, F.B.R., Somot, S., Le Loc'h, F., Leprieur, F., 2014. Towards a better understanding of potential impacts of climate change on marine species distribution: A multiscale modelling approach. Glob. Ecol. Biogeogr. 23, 1417–1429. https://doi.org/10.1111/geb.12217

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecol. Modell. 199, 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

ICES. 2021. Working Group on the Biology and Assessment of Deep-sea Fisheries Resources (WGDEEP). ICES Scientific Reports. 3:47. 944 pp. http://doi.org/10.17895/ices.pub.8108

Jiménez-Valverde, A., 2021. Prevalence affects the evaluation of discrimination capacity in presence-absence species distribution models. Biodivers. Conserv. 30, 1331–1340. https://doi.org/10.1007/s10531-021-02144-4

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. https://doi.org/10.1016/j.actao.2007.02.001

Jiménez, L., Soberón, J., 2020. Leaving the area under the receiving operating characteristic curve behind: An evaluation method for species distribution modelling applications based on presence-only data. Methods Ecol. Evol. 2020, 1–16. https://doi.org/10.1111/2041-210X.13479

Kellner, K.F., Swihart, R.K., 2014. Accounting for imperfect detection in ecology: A quantitative review. PLoS One 9. https://doi.org/10.1371/journal.pone.0111436

Koslow, J. A. 1996. Energetic and life-history patterns of deep-sea benthic, benthopelagic and seamount-associated fish. J. Fish Biol. 49:54-74.

Laman, E.A., Rooper, C.N., Turner, K., Rooney, S., Cooper, W., Zimmermann, M., n.d. Using species distribution models to describe essential fish habitat in Alaska. Can. J. Fish. Aquat. Sci.

Leroy, B., Bellard, C., Dubos, N., Colliot, A., Vasseur, M., Courtial, C., Bakkenes, M., Canard, A., Ysnel, F., 2014. Forecasted climate and land use changes, and protected areas: The contrasting case of spiders. Divers. Distrib. 20, 686–697. https://doi.org/10.1111/ddi.12191

Leroy, B., Delsol, R., Hugueny, B., Meynard, C.N., Barhoumi, C., Barbet-Massin, M., Bellard, C., 2018. Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. J. Biogeogr. 45, 1994–2002. https://doi.org/10.1111/jbi.13402.

Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography (Cop.). 28, 385–393. https://doi.org/10.1111/j.0906-7590.2005.03957.x

Liu, C., Newell, G., White, M., 2016. On the selection of thresholds for predicting species occurrence with presence-only data. Ecol. Evol. 6, 337–348. https://doi.org/10.1002/ece3.1878

Lorance, P., 2011. History and dynamics of the overexploitation of the blackspot sea bream (*Pagellus bogaraveo*) in the Bay of Biscay. ICES J. Mar. Sci. 68, 290–301. https://doi.org/10.1093/icesjms/fsq072

Lorance, P., F. Uiblein, and D. Latrouite. 2002. Habitat, behaviour and colour patterns of orange roughy *Hoplostethus atlanticus* (Pisces: Trachichthyidae) in the Bay of Biscay. J. Mar. Biolog. Assoc. 82:321-331.

Mateo, R., Aroca-Fernández, M.J., Gastón, A., Gómez-Rubio, V., Saura, S., García-Viñas, J.I., 2019. Looking for an optimal hierarchical approach for ecologically meaningful niche modelling. Ecol. Modell. 409. https://doi.org/10.1016/j.ecolmodel.2019.108735

Meynard, C.N., Leroy, B., Kaplan, D.M., 2019. Testing methods in species distribution modelling using virtual species: what have we learnt and what are we missing? Ecography (Cop.). 42, 2021–2036. https://doi.org/10.1111/ecog.04385

Michel, S., Vandermeirsch, F., Lorance, P., 2009. Evolution of upper layer temperature in the bay of biscay during the last 40 years. Aquat. Living Resour. 22, 447–461. https://doi.org/10.1051/alr/2009054

Montgomery, D.C., 2005. Design and Analysis of Experiments. Wiley

Morato, T., Solà, E., Grós, M.P., Menezes, G., 2001. Feeding habits of two congener species of seabreams, *Pagellus bogaraveo* and *Pagellus acarne*, off the Azores (northeastern Atlantic) during spring of 1996 and 1997. Bull. Mar. Sci. 69, 1073–1087.

Mytilineou, C., Politou, C.Y., Papaconstantinou, C., Kavadas, S., D'Onghia, G., Sion, L., 2005. Deep-water fish fauna in the Eastern Ionian Sea. Belgian J. Zool. 135, 229–233.

Mytilineou, C., Smith, C.J., Anastasopoulou, A., Papadopoulou, K.N., Christidis, G., Bekas, P., Kavadas, S., Dokos, J., 2014. New cold-water coral occurrences in the Eastern Ionian Sea: Results from experimental long line fishing. Deep. Res. Part II Top. Stud. Oceanogr. 99, 146–157. https://doi.org/10.1016/j.dsr2.2013.07.007

Nenzén, H.K., Araújo, M.B., 2011. Choice of threshold alters projections of species range shifts under climate change. Ecol. Modell. 222, 3346–3354. https://doi.org/10.1016/j.ecolmodel.2011.07.011

Olivier, R., 1928. Poissons de chalut, la dorade (Pagellus centrodontus). Rev. des Trav. l´Office des Pêches Marit. 1, 5–32.

Pecchi, M., Marchi, M., Burton, V., Giannetti, F., Moriondo, M., Bernetti, I., Bindi, M., Chirici, G., 2019. Species distribution modelling to support forest management. A literature review. Ecol. Modell. 411, 108817. https://doi.org/10.1016/j.ecolmodel.2019.108817

Penã, M., Ratilal, P., 2019. Mesopelagic fish avoidance from the vessel dynamic positioning system. ICES J. Mar. Sci. 76, 734–742. https://doi.org/10.1093/icesjms/fsy157

Pinho, M., Diogo, H., Carvalho, J., J., P.G., 2014. Harvesting juveniles of blackspot sea bream (*Pagellus bogaraveo*) in the Azores (Northeast Atlantic): biological implications, management, and life cycle considerations. ICES J. Mar. Sci. 71, 2448–2456. https://doi.org/10.4135/9781412953924.n678

Pinsky, M.L., Selden, R.L., Kitchel, Z.J., 2020. Climate-Driven Shifts in Marine Species Ranges: Scaling from Organisms to Communities. Ann. Rev. Mar. Sci. 12, 153–179. https://doi.org/10.1146/annurev-marine-010419-010916

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., Levin, S.A., 2013. Marine taxa track local climate velocities. Science (80-.). 341, 1239–1242. https://doi.org/10.1126/science.1239352

Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K., Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C. V., O'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F., Thompson, S.A., Richardson, A.J., 2013. Global imprint of climate change on marine life. Nat. Clim. Chang. 3, 919–925. https://doi.org/10.1038/nclimate1958

Rincón, L., Castro, P.L., Álvarez, B., Hernández, M.D., Álvarez, A., Claret, A., Guerrero, L., Ginés, R., 2016. Differences in proximal and fatty acid profiles, sensory characteristics, texture, colour and muscle cellularity between wild and farmed blackspot seabream (*Pagellus bogaraveo*). Aquaculture 451, 195–204. https://doi.org/10.1016/j.aquaculture.2015.09.016

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, 1–11. https://doi.org/10.3389/fmars.2017.00421

Santos, R., W. Medeiros-Leal, A. Novoa-Pabon, H. Silva, and M. Pinho. 2021. Demersal fish assemblages on seamounts exploited by fishing in the Azores (NE

Atlantic). J. Appl. Ichthyol. 37:198-215.

Sanz-Fernández, V., Gutiérrez-Estrada, J.C., Pulido-Calvo, I., Gil-Herrera, J., Benchoucha, S., el Arraf, S., 2019. Environment or catches? Assessment of the decline in blackspot seabream (*Pagellus bogaraveo*) abundance in the Strait of Gibraltar. J. Mar. Syst. 190, 15–24. https://doi.org/10.1016/j.jmarsys.2018.08.005

Schickele, A., Leroy, B., Beaugrand, G., Goberville, E., Hattab, T., Francour, P., Raybaud, V., 2020. Modelling European small pelagic fish distribution: Methodological insights. Ecol. Modell. 416, 108902. https://doi.org/10.1016/j.ecolmodel.2019.108902

Silva, H. M., Pinho, M. R. (2007). Small-Scale Fishing on Seamounts. In T. J. Pitcher, T. Morato, P. J. B. Hart, M. R. Clark, N. Haggan, and R. S. Santos (Eds.), Seamounts: Ecology, Fisheries and Conservation (pp. 335–360). Blackwell Publishing.

Spedicato, M.T., Greco, S., Sophronidis, K., Lembo, G., Giordano, D., Argyri, A., 2002. Geographical distribution, abundance and some population characteristics of the species of the genus *Pagellus* (Osteichthyes: Perciformes) in different areas of the Mediterranean. Sci. Mar. 66, 65–82. https://doi.org/10.3989/scimar.2002.66s265

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - A platform for ensemble forecasting of species distributions. Ecography (Cop.). 32, 369–373. https://doi.org/10.1111/j.1600-0587.2008.05742.x

Thuillier, W., Georges, D., Gueguen, M., Engler, R., Breiner, F., 2003. Biomod2: Ensemble platform for species distribution modeling. R Packag. version 3.3-7.

Valanis, D. V., Pierce, J.P., Zuur, A.F., Palialexis, A., Saveliev, A., Katara, I., Wang, J., 2008. Modelling of essential fish habitat based on remote sensing, spatial analysis and GIS. Hydrobiologia 612, 5–20.

Valencia, V., Fontán, A., Goikoetxea, N., Chifflet, M., González, M., López, A., 2019. Long-term evolution of the stratification, winter mixing and Θ-S signature of upper water masses in the southeastern Bay of Biscay. Cont. Shelf Res. 181, 124–134. https://doi.org/10.1016/j.csr.2019.05.010

Welsh, A.H., Lindenmayer, D.B., Donnelly, C.F., 2013. Fitting and Interpreting Occupancy Models. PLoS One 8. https://doi.org/10.1371/journal.pone.0052015

Wisz, M.S., Guisan, A., 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecol. 9, 1–13. https://doi.org/10.1186/1472-6785-9-8

Zurell, D., Franklin, J., König, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J.J., Leitão, P.J., Park, D.S., Peterson, A.T., Rapacciuolo, G., Schmatz, D.R., Schröder, B., Serra-Diaz, J.M., Thuiller, W., Yates, K.L., Zimmermann, N.E., Merow, C., 2020. A standard protocol for reporting species distribution models. Ecography (Cop.). 43, 1261–1277. https://doi.org/10.1111/ecog.04960

Zurell, D., Pollock, L.J., Thuiller, W., 2018. Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments? Ecography (Cop.). 41, 1812–1819. https://doi.org/10.1111/ecog.03315