# Predicting marine species distributions: complementarity of food-web and Bayesian hierarchical modelling approaches

M. Coll*[1,2], M. Grazia Pennino*[3,4,5], J. Steenbeek[1,2], J. Sole[1], J.M. Bellido[5,6]

*Authors share first co-authorship.

[1] Institut de Ciències del Mar (CMIMA-CSIC), P. Marítim de la Barceloneta, 37-49, 08003 Barcelona, Spain (current address).

[2] Ecopath International Initiative Research Association, Barcelona, Spain.

[3] Fishing Ecology Management and Economics (FEME) - Universidade Federal do Rio Grande do Norte – UFRN. Depto. de Ecologia, Natal (RN), Brazil.

[4] Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Vigo, Subida a Radio Faro 50-52, 36390 Vigo, Pontevedra, Spain.

[5] Statistical Modeling Ecology Group (SMEG). Departament d'Estadística i Investigació Operativa, Universitat de València. C/Dr. Moliner 50, Burjassot. 46100 Valencia, Spain.

[6] Instituto Español de Oceanografía, Centro Oceanográfico de Murcia. C/Varadero 1, San Pedro del Pinatar. 30740 Murcia, Spain.

Corresponding author: Marta Coll. E-mail: mcoll@icm.csic.es; marta.coll.work@gmail.com

**Keywords:** spatial ecology, species distribution models, Bayesian model, food-web model, Ecospace, commercial species, Mediterranean Sea.

**Abstract**

The spatial prediction of species distributions from survey data is a significant component of spatial planning and the ecosystem-based management approach to marine resources. Statistical analysis of species occurrences and their relationships with associated environmental factors is used to predict how likely a species is to occur in unsampled locations as well as future conditions. However, it is known that environmental factors alone may not be sufficient to account for species distribution. Other ecological processes including species interactions (such as competition and predation), and the impact of human activities, may affect the spatial arrangement of a species. Novel techniques have been developed to take a more holistic approach to estimating species distributions, such as Bayesian Hierarchical Species Distribution model (B-HSD model) and mechanistic food-web models using the new *Ecospace* Habitat Foraging Capacity model (E-HFC model). Here we used both species distribution and spatial food-web models to predict the distribution of European hake (*Merluccius merluccius*), anglerfishes (*Lophius piscatorius* and *L. budegassa*) and red mullets (*Mullus barbatus* and *M. surmuletus*) in an exploited marine ecosystem of the Northwestern Mediterranean Sea. We explored the complementarity of both approaches, comparing results of food-web models previously informed with species distribution modelling results, aside from their applicability as independent techniques. The study shows that both modelling results are positively and significantly correlated with observational data. Predicted spatial patterns of biomasses show positive and significant correlations between modelling approaches and are more similar when using both methodologies in a complementary way: when using the E-HFC model previously informed with the environmental envelopes obtained from the B-HSD model outputs, or directly using niche calculations from B-HSD models to drive the niche priors of E-HFC. We discuss advantages, limitations and future developments of both modelling techniques.

## 1. **Introduction**

Marine resources and ecosystem services change in response to human stressors, such as fishing activities, habitat modification, and pollution (Halpern et al., 2015), in addition to environmental variability and change (Cury et al., 2008). The need to consider changes in the environment as well as human activities when analysing and managing marine ecosystems highlights the necessity to perform integrated analyses (Link, 2011). The productivity of marine resources depends on many factors: the state of communities, their structural and functional properties, the state of the ecosystems as a whole; external climatological factors; and human exploitation and the dynamics of target species in conjunction with the dynamics of non-target organisms. As such, environmental drivers and human impacts have to be included into the consideration to manage marine resources soundly (Christensen and Maclean, 2011).

To address the need for more holistic assessments, a wide variety of statistical and machine-learning methods predict, often in conjunction with geographic information systems and remote-sensing, spatial species distributions, abundance, and biomass from survey data. Frequently, the purpose of the species distribution statistical models (SDMs thereafter) is to use the information about where a species occurs and the relationship with associated environmental factors to predict how likely the species is to occur in unsampled locations and future environmental conditions. SDMs use the range of sampled environments and the same general time frame of the sampling to predict quantities of interest at un-sampled locations based on measured values at nearby sampled locations or future environmental factors. Spatial predictions of species distributions are thus directly related to the concept of the environmental niche, a specification of a species' response to a suite of environmental factors. Different techniques of SDMs have been applied to the marine environment (e.g., Jones et al., 2012; Kaschner et al., 2006), including the Mediterranean Sea (e.g., Morfin et al., 2012; Saraux et al., 2014). Guisan and Zimmermann (Guisan and Zimmermann, 2000) provide an extensive review of SDM statistical approaches.

75 Despite SDMs popularity, it is known that environmental drivers alone may not be sufficient to
76 account for species distributions (Navarro et al., 2015; Pollock et al., 2014). Other ecological
77 processes, including trophic interactions (such as competition, predation and facilitation), behavioural
78 parameters, and population dynamics may affect the spatial arrangement of a species, in addition to
79 human activities. Novel techniques take these processes into account when estimating species
80 distributions, such as Bayesian hierarchical species distribution models and mechanistic food-web
81 models using the new *Ecospace* habitat capacity model.

82 Bayesian hierarchical species distribution models (B-HSD models thereafter) estimate and predict
83 the abundance and distribution of marine species, using commonly environmental factors, but can
84 also indirectly include biological relationships and human activities. They are particularly appropriate
85 to identify and predict the distribution of the species as they allow both the observed data and model
86 parameters to be considered as random variables, resulting in a more realistic and accurate estimation
87 of uncertainty (Banerjee et al., 2014). B-HSD models are also used in data analyses with
88 geographically uneven levels of survey (sampling) effort, as such bias can be incorporated within the
89 analysis, reducing its influence on estimates of the effects of environmental variables. B-HSD models
90 allow for the incorporation of the spatial component as a random-effect term that represents the spatial
91 intrinsic variability of the data, after the exclusion of the environmental variables (Gelfand et al.,
92 2006). B-HSD model's applications in the Mediterranean Sea have illustrated that they can be
93 implemented for different purposes; for example to identify nurseries areas (Paradinas et al., 2015),
94 high density discards hot-spots (Paradinas et al., 2016; Pennino et al., 2014), or essential fish habitats
95 for vulnerable species (Pennino et al., 2016; Pennino et al., 2013).

96 Ecological processes and human activities, in addition to environmental factors, can be explicitly
97 considered in process-based oriented modelling (Fulton, 2010; Guisan and Zimmermann, 2000), like
98 in food-web models such as *Ecopath with Ecosim* (*EwE* thereafter) (Christensen and Walters, 2004).
99 *EwE* approach allows building food-web models by describing the ecosystem by means of energy

100 flows between functional groups, each representing a species, a sub-group of a species (e.g. juveniles
101 and adults) or a group of species that have functional and ecological similarities. *Ecospace* is the
102 spatial-temporal dynamic module of *EwE* that allows representing temporal and spatial 2D dynamics
103 of trophic web components (Walters et al., 2010; Walters et al., 1999). *Ecospace* has been widely
104 applied to quantify the spatial impact of fisheries on marine species (e.g., Christensen et al., 2003),
105 to analyses the impact of management scenarios such as the establishment of marine protected areas
106 (Walters, 2000), to develop spatial optimization routines (e.g., Christensen et al., 2009) and to assess
107 the impact of climate change on marine ecosystems by linking *Ecospace* with low trophic level
108 models (Fulton, 2011) or external spatial-temporal data (Steenbeek et al., 2013). To overcome
109 limitations of the original configuration of *Ecospace*, the Habitat Foraging Capacity model (E-HFC
110 model thereafter) (Christensen et al., 2014) was recently added to the spatial-temporal modelling
111 capabilities of *EwE*. The E-HFC model offers the ability to spatially drive foraging capacity of species
112 from the cumulative effects of multiple physical, oceanographic, and environmental and topographic
113 conditions and runs in conjunction with the food web and fisheries dynamics. This development, in
114 combination with the spatial-temporal framework module of *EwE* (Steenbeek et al., 2013), bridged
115 the gap between envelope environmental models and food-web models (Christensen et al., 2015;
116 Christensen et al., 2014).

117 A previous study applied the E-HFC model to analyse the historical and current distribution of three
118 commercially important fish species in a marine ecosystem of the NW Mediterranean Sea: European
119 hake (*Merluccius merluccius*), sardine (*Sardina pilchardus*) and anchovy (*Engraulis encrasicolus*)
120 (Coll et al., 2016). It evaluated the combined effects of environmental and topographic drivers
121 (primary production, salinity, temperature, substrate and depth), in addition to fishing dynamics and
122 food-web structure in their population spatial dynamics. Results illustrated the role of fishing and
123 environmental conditions on the biomass and distributions of these three species. Fishing had the
124 highest impact on results, while the spatial distribution of primary producers and depth followed in

125 importance.

126 In the parameterization of the E-HFC model, for each species in the food web previous knowledge is
127 required about which environmental parameters are important, and how species respond to these
128 environmental parameters. This knowledge can be obtained from local field studies or global
129 databases used for species distribution modelling initiatives (Kaschner et al., 2016). However, this
130 vital information can also be obtained from more accurate species distribution modelling studies at
131 local and regional scales.

132 In the present study, we investigated the distributions of five demersal commercial fish species:
133 European hake, anglerfish (*Lophius piscatorius* and *L. budegassa*) and red mullets (*Mullus*
134 *surmuletus* and *M. barbatus*) building from the previous applications of the E-HFC model (Coll et
135 al., 2016) and of B-HSD model (Munoz et al., 2013; Pennino et al., 2014; Pennino et al., 2013) in the
136 NW Mediterranean Sea. We firstly applied the B-HSD and E-HFC models independently and
137 compared results between both methodologies. Afterwards, relevant environmental parameters
138 selected by the B-HSD model and resulting functional responses or niche calculations were used to
139 parametrize the E-HFC model. The aim of the study was to explore the complementarity of both
140 approaches aside from their applicability as independent techniques. We discuss the benefits and
141 limitations of using both methodologies in parallel or in combination, and how this can complement
142 current knowledge about species distributions in the marine environment.

143 2. **Material and Methods**

144 ***a) Study area***

145 Our study area was located in the Southern Catalan Sea (within the Balearic Sea, NW Mediterranean
146 Sea, Figure 1). The NW Mediterranean Sea is an area of relatively high productivity due to a
147 combined effect of the Northern current and the runoff of the Ebro and Rhone Rivers (Bosc et al.,
148 2004; Estrada, 1996). In the northern area, the continental shelf is narrower, with the northern current

149 flowing south-westwards along the continental slope, towards the wider continental shelf surrounding
150 the Ebro Delta River. The area contains an important fishing ground for small pelagic fish and
151 demersal mesopredators (Palomera et al., 2007). It is also important for marine predatory species at
152 risk, such as marine mammals and seabirds (Coll et al., 2015).

153 According to previous studies, important environmental variables drive the dynamics of commercial
154 species in the area such as temperature, salinity and nutrients from the river run off (.e.g., Lloret et
155 al., 2004; Martin et al., 2012; Ospina-Alvarez et al., 2015; Palomera et al., 2007). In addition,
156 commercial marine resources are highly exploited (e.g., Colloca et al., 2013; Fernandes et al., 2017;
157 Tsikliras et al., 2015). Previous studies looking at the temporal dynamics of marine resources
158 identified that environmental factors, human activities and the structure of the food web were
159 important elements to predict ecosystem dynamics (Coll et al., 2008; Coll et al., 2016).

160 ***b) Selected species and environmental data***

161 Five demersal species were selected: European hake (*Merluccius merluccius*), two species of mullets:
162 striped red mullet (*Mullus surmuletus*) and red mullet (*Mullus barbatus*), and two species of angler
163 fish: black-bellied angler (*Lophius budegassa*) and angler (*Lophius piscatorius*). Due to the
164 difficulties of splitting fisheries statistics of the two species of red mullets and of anglerfish, these
165 four species were treated as two groups in the analyses and models: red mullets and anglerfish,
166 respectively. Henceforth we will refer to red mullets as *Mullus* spp., and *Lophius* spp. for anglerfishes.
167 Seven environmental and topographic variables were considered as potential predictors of the species
168 biomass. These included five climatic variables: Sea Surface Temperature (SST) and Sea Bottom
169 Temperature (SBT), both expressed in degrees Celsius (ºC); Sea Surface Salinity (SSS) and Sea
170 Bottom Salinity (SBS), both expressed in Practical Salinity Units (PSU); and primary production (PP,
171 mg C $m^{-2}$ $day^{-1}$); and two topographic features: depth (m); and type of the seabed or substrate.
172 Environmental variables (SST, SBT, SSS, SBS), as well as primary production (PP) were derived

from a regional application of the ROMS model (Shchepetkin and McWilliams, 2005) coupled to a biogeochemical nitrogen-based plankton model (Fennel et al., 2006). This coupled model implementation was tested in the Western Mediterranean (Alboran Sea) in previous work (Macias et al., 2011; Solé et al., 2016) and had been already used to drive a previous version of the *Ecospace* HFC model (Coll et al., 2016). The ROMS implementation was adapted to the Catalan Sea with a grid of 0.05º x 0.05º degrees resolution and a vertical resolution of 40 levels. Both boundary and atmospheric forcing conditions were climatologies. Boundary conditions were obtained from the NEMO model, which is available from http://www.nemo-ocean.eu, and the simulations used in this work were reported in Adani and co-authors (Adani et al., 2011). NEMO output was interpolated to the ROMS grid and imposed to a sponge layer of 10 horizontal grid points with a nudging relaxation time of 30 days. For the biological variables, boundary conditions were set-up as in Fennel and co-authors (Fennel et al., 2006). The meteorological forcing climatology was calculated from the European Center of Medium Weather Forecast (hereafter ECMWF) data, derived from ERA-40 reanalysis for air temperature, short wave radiation, long wave radiation, precipitation, cloud cover and salt flux. For pressure at the surface, we used the ECMWF Era-Interim reanalysis. QuickScat blind data was used for wind forcing (both zonal and meridional). The ROMS model was ran using both boundary conditions and atmospheric forcing climatologies to obtain a stable initial state during eight years. After this spin-up period, we used the ninth year as the year of study with the same climatological conditions used for the spin-up period and results were used as averaged climatology conditions (Coll et al., 2016).

In addition, depth and seabed habitat types were considered for the analyses as they are some of the main factors controlling species distribution and have been identified as key predictors to determine spatial patterns of many species (Lauria et al., 2015; Roos et al., 2015). Data for the seabed habitat types were obtained from the European Marine Observation Data Network (EMODnet) Seabed Habitats project (http://www.emodnet-seabedhabitats.eu/, European Commission's Directorate-

198 General for Maritime Affairs and Fisheries).

199 All the environmental data was aggregated at 0.05°×0.05° spatial resolution. Following a proposed
200 protocol (Zuur et al., 2010), these variables were explored for correlation, co-linearity, outliers, and
201 missing data before their use in the analysis and modelling (Figure 1 in Supplementary Material).
202 SBS was highly correlated to SBT and SSS ($r > 0.70$) and was the only one with a Generalized
203 Variance Inflation Factor (VIF) > 3 (Figure 2 in Supplementary Material). For this reason, these
204 variables were used alternatively in the Bayesian models taking particular attention to the SBS
205 variable. Specifically, separate model runs were performed including only one of each of the highly
206 correlated variables (SBS, SBT or SSS) to determine which one would explain more of the variance.
207 Finally, after an exploratory analysis, in order to better interpret, both the direction (positive or
208 negative) and magnitudes (effect sizes) of parameter estimates in relation to the others, the
209 explanatory variables were standardized (difference from the mean divided by the corresponding
210 standard deviation) (Gelman, 2008; Hereford et al., 2004).

211 ***c) Bayesian Hierarchical Species Distribution model***

212 Many SDMs are applied to predict the spatial distribution of species. However, these algorithms do
213 not always provide accurate results if they run using traditional prediction methods because there is
214 often a large amount of variability in the measurements of both response and environmental variables.
215 Bayesian statistical methods are increasingly used in fisheries (Munoz et al., 2013; Paradinas et al.,
216 2015; Pennino et al., 2014) because they offer several advantages over traditional statistical methods.
217 Using both observed data the model parameters as random variables produces more realistic and
218 accurate uncertainty estimations (Banerjee et al., 2014). Furthermore, Bayesian statistics integrate all
219 types of uncertainties using probability as a metric. By combining uncertainty in input data
220 (*likelihood*) with extra-data information (*prior distributions*), posterior probability distributions for
221 all unknown quantities of interest (i.e., parameters) are built using Bayes' theorem (Kinas and
222 Andrade, 2017).

Bayesian methods also allow for the incorporation of the spatial component as a random-effect term, thereby reducing its influence on estimates of the effects of geographical variables (Gelfand et al., 2006). Species distributions in the surrounding locations may influence observed species distributions due to contagious biotic processes (e.g., predation, competition, etc.). Similarly, environmental variables used to describe locations are neither randomly nor uniformly spatially distributed, but structured by physical processes causing gradients and/or patchy structures. One consequence of this general property of ecological variables is that the assumption of independence of the observations is not respected (Legendre, 1993). Therefore, it is necessary to incorporate the spatial structure of the data within the modelling process. Specifically, by treating the spatial effect as a variable of interest, these models can identify additional covariates that may improve the model fit and prediction and highlight the existence of possible effects that may affect the quantity of species biomass, such as human activities or trophic interactions.

Hence, in this study we used a hierarchical Bayesian point-reference spatial distribution model (B-HSD model) to estimate the species occurrence and biomass dependency with respect to chosen environmental variables (Munoz et al., 2013).

In particular, $Y_i$ and $Z_i$ denote, respectively, the spatial distributed occurrence and the conditional-to-presence biomass, where $i = 1,\ldots..,\ n$ is the spatial location. Then, as usual with this kind of variables, we modelled the occurrence, $Y_i$, using a Bernoulli distribution. In the case of the biomass, $Z_i$, our selection to model it was a Lognormal distribution. The mean of both variables was then related via the usual link functions (logit and log, respectively) to the environmental effects:

$$Y_i \sim \text{Bernoulli}(\pi_i) \qquad (1)$$

$$Z_i \sim \text{Lognormal}\ (\mu_i\ ,\ \sigma^2{}_i) \qquad (2)$$

$$\text{logit}(\pi_i) = \alpha^{(Y)} + X_i\beta + W_i^{(Y)}$$

$$\log(\mu_i) = \alpha^{(Z)} + X_i\beta + W_i^{(Z)}$$

where $\pi_i$ represents the probability of occurrence at location *i* and $\mu_i$ and $\sigma^2_i$ are the mean and variance of the conditional-to-presence biomass. The linear predictors containing the effects to which these parameters $\pi_i$ and $\mu_i$ are linked are formed with: $\alpha^{(Y)}$ and $\alpha^{(Z)}$, the terms representing the intercepts for each variable; $\beta$ is the vector of regression parameters, $X_i$ is the matrix of the explanatory covariates at location *i*; and the final terms $W_i^{(Y)}$ and $W_i^{(Z)}$ refer to the spatial structure of the occurrence and conditional-to-presence biomass respectively.

For all models, Bayesian parameter estimates and predictions were obtained throughout the Integrated Nested Laplace Approximations (INLA) approach (Rue et al., 2009) and package (http:\\www.r-inla.org) that is implemented in the R software (R Core Team 2017).

For the spatial effects (W), INLA implements the Stochastic Partial Differential Equations (SPDE) approach (Lindgren et al., 2011), that involves the approximation of a continuously indexed Gaussian Field (GF) with a Matérn covariance function ($Q$) by a Gaussian Markov Random Field (GMRF). In particular, a prior Gaussian distribution with a zero mean and covariance matrix was assumed for the spatial component which depend on the hyperparameters *k* and $\tau$, which determined its variance and range, respectively (see Munoz et al., 2013, for more detailed information about spatial effects). As recommended by Held and co-authors (Held et al., 2010) vague zero-mean Gaussian prior distribution with a variance of 1e5 were assigned for all fixed-effect parameters, which are approximations of non-informative priors designed to have little influence on the posterior distributions.

The environmental variable selection with all possible interaction terms was mainly based on the Watanabe-Akaike information criterion (WAIC), that is an improvement of the Deviance Information Criterion (DIC), traditionally used for Bayesian models, and better suited than the Akaike Information Criterion (AIC), usually applied within frequentist modelling procedures (Spiegelhalter et al., 2002). Unlike DIC, which is conditioned on a point estimate and is not fully Bayesian, WAIC is a fully Bayesian measure and uses entire posterior distributions to make inference about parameters; hence, estimations are more precise (Watanabe, 2010). Thus, the best (and most parsimonious) model was

chosen based on low WAIC values, containing only relevant predictors; i.e., those predictors with 95% credibility intervals not including the zero. In addition, other two criteria were computed to assess the models performance: the Root Mean Square Error (RMSE) and the adjusted coefficient of determination ($R^2$). The RMSE was assessed to check if the selected model has a low standard deviation between observed and predicted values, while the $R^2$ was used to prove that the selected model had a reasonable level of variance explained by the variables.

To validate the occurrence of B-HSD predictions, the common cross-validation procedure was implemented, which consists in randomly splitting the original dataset into two main subsets: a training dataset including 75% of the total observations, and a validation dataset containing the remaining 25% of the data (Fielding and Bell, 1997). The relationship between observed data and environmental variables was modeled by using the training dataset (i.e. models were run again with the traning dataset). The quality of predictions was then assessed by using the validation dataset. We repeated the validation procedure 10 times for the best model of each species and results were averaged over the different random subsets. We performed a validation procedure to formally evaluate overall model prediction using the area under the receiver-operating characteristic curve (AUC) (Fielding and Bell, 1997) and the "True Skill Statistic" (TSS) (Allouche et al., 2006). Both values range from 0 to 1 where values closer to 1 are better. For the biomass B-HSD models the Spearman correlation and the RMSE measures were calculated between the predicted and observed values (see section f below for details).

The B-HSD models were applied to data on the five demersal fish species previously mentioned. They were collected during the EU-funded MEDIterranean Trawl Survey (MEDITS) project (Bertrand et al., 2002), carried out from spring to early summer (April to June) from 2002 to 2012. The MEDITS project used a stratified sampling design based on depth (five bathymetric strata: 10 – 50, 51 – 100, 101 – 200, 201 – 500 and 501 – 700 m) and Geographical Sub-Area (GSA). Sampling stations were randomly placed within each stratum at the beginning of the project (Figure 2a-c). In

301 all subsequent years, sampling was performed in similar locations. The MEDITS data used in this
302 study concerns the area in the vicinity of the Ebro river mouth (Figure 1), and includes only the first
303 three bathymetric strata from 0 to 200 m. On average, 20 hauls per year divided into the various
304 bathymetric strata were performed every year in this zone (Figure 2a-c). Species biomass was
305 estimated for each haul as the total weight of each species (kg) per $km^2$ of trawling (Figure 2d-f).
306 Biomass values of all species were log transformed to down weight extreme values and to ensure a
307 normal distribution and homoscedasticity.

308 ***d) Ecospace Habitat Foraging Capacity modelling***

309 The basic routine of *Ecopath* provides a snapshot of the structure and flows of a food web and
310 describes the balance between production of functional groups and all consumptions within an
311 ecosystem (Christensen and Pauly, 1992; Polovina, 1984). Each functional group can represent a
312 species, a sub-group of a species (e.g. juveniles and adults) or a group of species that have functional
313 and ecological similarities. The *Ecopath* model uses a system of linear equations to describe the
314 average flows of mass and energy between these groups during a period of time, normally a year. The
315 flow to and from each group is described by the following equation:

316 $$B_i \cdot (P/B)_i = \sum B_j \cdot (Q/B)_j \cdot DC_{ij} + Y_i + E_i + BA_i + B_i \cdot (P/B)_i \cdot (1 - EE_i) \quad (3)$$

317 where $B_i$ is the biomass of group $i$, $(P/B)_i$ is the production per unit of biomass, $Y_i$ is the total fishery
318 catch rate, $E_i$ is the net migration rate (emigration−immigration), $BA_i$ is the biomass accumulation
319 rate, $EE_i$ is the Ecotrophic Efficiency', or the proportion of the production that is utilized in the
320 system, $B_j$ *is* the biomass of consumers or predators $j$, $(Q/B)_j$ *is* the consumption per unit of biomass
321 of $j$, and $DC_{ij}$ is the fraction of $i$ in the diet of $j$. To parameterize an *Ecopath* model it is required a
322 series of inputs for each functional group $i$, mostly Biomass ($Bi$), Diet ($DC_{ij}$), consumption and
323 production per unit of biomass ($(Q/B)_i$ and $(P/B)_i$), and fishing yields and other exports ($Y_i$ and $E_i$).
324 Among B, P/B, Q/B, EE, one parameter can be estimated by the model and the others are mandatory

inputs (Christensen and Walters, 2004; Christensen et al., 2008).

*Ecosim* is the temporal dynamic model of *EwE* and is used to simulate ecosystem effects of (mainly fishing) mortality changes and environmental forcing over time (Christensen and Walters, 2004; Walters et al., 1997; Walters et al., 2000). The model uses a system of time-dependent differential equations from the baseline mass-balance model (equation 3 above), where the biomass growth rate is calculated as:

$$dB_i/dt = g_i \cdot \sum Q_{ji} - \sum Q_{ij} + I_i - (M_i + F_i + e_i) \cdot B_i \quad (4)$$

where $dB_i/dt$ represents the growth rate of group $i$ during the time interval $dt$ in terms of its biomass $B_i$, $g_i$ is the net growth efficiency (production/consumption ratio, $P/Q$), $M_i$ is the non-predation $((P/B)_i \cdot B_i(1-EE_i))$ natural mortality rate, $F_i$ is fishing mortality rate, $e_i$ is emigration rate, $I_i$ is immigration rate and $e_i \cdot B_i - I_i$ is the net migration rate. The two summations in equation 4 estimate consumption rates, the first expressing the total consumption by group *i,* and the second the predation by all predators on the same group *i*. The consumption rates, $Q$, are calculated based on the 'foraging arena' concept, where $B_i$'s are divided into vulnerable and invulnerable components (Ahrens et al., 2012).

The set of *Ecosim* equations are used in the spatial routine *Ecospace,* the spatial-temporal model of *EwE*, that predicts the biomass dynamics in a two-dimensional space (Walters et al., 1999). 'Water' cells in *Ecospace* can be assigned to contain one or more habitat types and species can be assigned to preferred habitats (Christensen et al., 2014). Fishing fleets can be limited to fish in specific habitats, and can be subjected to zonal fishing regulations (no taking zones, e.g. Walters et al., 2000). Moreover, spatial variations of primary productivity and of fishing costs (e.g. costs related to the distance from fishing ports) can be incorporated. The model further incorporates dispersal rates of organisms and other behavioural parameters (Christensen et al., 2005).

In the original *Ecospace*, habitat types associated to biomass distributions and trophic interactions

were represented by a binary habitat use value, with each spatial cell being either entirely suitable (1) or entirely unsuitable (0) for species/functional groups. Therefore, the original version of *Ecospace* assumed homogenous conditions within each spatial cell, and local, but possibly relevant variations within cells, could not be represented. To overcome these limitations the Habitat Foraging Capacity model was recently developed (Christensen et al., 2014) to add the ability to spatially drive the foraging capacity of species across the *Ecospace* map from the cumulative responses to multiple physical, oceanographic, and environmental drivers.

In the Habitat Foraging Capacity model (Christensen et al., 2014) (E-HFC), the habitat capacity is defined as the suitability of a cell for a species or functional group to forage. This foraging capacity $C_{rcj}$ for each group $j$ in each cell $r,c$ is a continuous value from 0 to 1, and is calculated as a function of a vector of environmental attributes $H_{rc} = (H1,H2,...Hv)_{rc}$ in cell $r,c$, where $H1$ to $Hv$ represent any number of environmental factors such as water depth, proportion of hard bottom, surface temperature or salinity.

The $C_{rcj}$ values have to be linked to trophic interaction dynamics to specify how $C_{rcj}$ impacts food consumption and predation rates. This is done through a functional response of group $j$ to each environmental condition and through the basic foraging arena equations used to predict trophic interaction (food-web biomass flow) rates in the time dynamic module *Ecosim* (see Ahrens et al., 2012; Christensen and Walters, 2004; Walters et al., 1997; Walters et al., 2000; Walters and Martell, 2004). *Ecosim* represents biomass dynamics, trophic interaction and fishery effects as:

$$dB_j/dt = \frac{g_j \cdot a_{ij} \cdot v_{ij} \cdot B_j \cdot B_i}{2 \cdot v_{ij} + a_{ij} \cdot B_j} - Z_j \cdot B_j \qquad (5)$$

where $B_j$ is predator biomass, $B_i$ is prey biomass, $Z_j$ is total instantaneous mortality rate of $j$, $g_j$ is growth efficiency (corresponding to the production/consumption ratio, which can vary as predators grow in size), $v_{ij}$ is prey vulnerability exchange rate, and $aij$ is the rate of effective search by the predator. The vulnerable prey density $V_{ij}$ is represented by the foraging arena equation, which can be

expressed as:

$$V_{ij} = \frac{v_{ij} \cdot B_j}{2 \cdot v_{ij} + a_{ij} \cdot B_j} \quad (6)$$

where predation pressure in a cell depends on the foraging arena area in that cell. If we assume that variation in relative habitat capacity for the predator means variation in the foraging arena area over which a species can forage successfully, we can include a variation in relative habitat capacity in the *Ecospace* model by dividing the denominator $a_{ij} \cdot B_j$ term by relative habitat size or capacity $C_{rcj}$:

$$V_{ij} = \frac{v_{ij} \cdot B_j}{2 \cdot v_{ij} + a_{ij} \cdot B_j / C_{rcj}} \quad (7)$$

This assumption concentrates predation activity into smaller relative areas when $C$ (foraging arena size) is small, so this drives down vulnerable prey densities $V_{ij}$ more rapidly as $B_j$ increases in cases with less foraging arena area.

Importantly, including $C_{rcj}$ as a modifier in the $aij \cdot Bj/C_{rcj}$ predation rate term results in the equilibrium predator biomass ($B_j$ for which $dBj/dt=0$) being proportional to $C_{rcj}$, i.e.,

$$\mathrm{B_j} = \left(\mathrm{g_j} \cdot \mathrm{v_{ij}} \cdot \frac{\mathrm{B_i}}{\mathrm{Z_j}} \text{-} 2 \cdot \frac{\mathrm{v_{ij}}}{\mathrm{a_{ij}}}\right) \cdot \mathrm{C_{rcj}}\, B_j = \left(g_j \cdot v_{ij} \cdot \frac{B_i}{Z_j} - 2 \cdot \frac{v_{ij}}{a_{ij}}\right) \cdot C_{rcj} \quad (8)$$

Using the $C_{rcj}$ as modifiers of the foraging arena consumption rate equation results in spatial patterns of biomass of consumers being proportional to $C_{rcj}$, other factors (prey biomasses $B_i$ and mortality rates $Z_j$) being equal over space. The variation in habitat capacity also affects the vulnerability exchange rates $v_{ij}$, search rates $a_{ij}$, and predation rates $Z_j$, but the default assumption is that the dominant cause of 'poor' or relatively small habitat capacity is the lack of usable foraging arena area (Christensen et al., 2014).

In this study, we used an available ecosystem model developed with *EwE* (Christensen and Walters,

2004) representing the South Catalan Sea (NW Mediterranean) in 1978 that had been previously calibrated to time series of data from 1978 to 2010 using the temporal dynamic model *Ecosim* and validated with external data such as trophic levels and mortalities (Coll et al., 2013; Coll et al., 2008; Coll et al., 2006; Navarro et al., 2011). The model included 40 functional groups and four fishing fleets (bottom trawling, purse seine, long lining and tuna fishing), and covered an area of 5000 $km^2$ with depths from 50 to 400 m (Coll et al., 2006). The functional groups included primary and secondary producers, from phytoplankton to large predatory species and the units were expressed in terms of biomass as $t \cdot km^2$, and production and catch as $t \cdot km^2 \cdot year^{-1}$. Four of the five studied species, two of anglerfish and the other two of red mullets, were modelled as functional groups merging the two species of the genus *Lophius* and the two species of the genus *Mullus* together due to data availability and quality.

A previous E-HFC model developed to evaluate the combined effects of environmental conditions and fishing in the ecosystem dynamics of the Southern Catalan Sea was used as a starting point (Coll et al., 2016). The environmental variables used to parameterize the E-HFC model were the same as the ones used for the B-HSD model. In addition, the spatial pattern of primary production was used to drive the dynamics of the phytoplankton group (through the variation of the initial value of P/B) of the food-web model (Coll et al., 2016). The environmental envelops needed to be parameterized the functional responses in the E-HFC model that link the environmental variables with the response of hake, anglerfish and mullets were firstly obtained from the literature (following the implementation of Coll et al., 2016) and alternatively from results of the B-HSD model (Figure 3, *Path 1*, see next section for a detailed explanation).

***e) Testing modelling integration***

Two different approaches were implemented to test the possible integration of B-HSD and E-HFC and are schematized in Figure 3. The first one used the functional responses obtained from the

418 biomass B-HSD to inform E-HFC (1) and the second one used the niche calculations of the occurrence
419 B-HSD to drive the niche information within E-HFC (2):

420 1) In order to define the species environmental envelop in the E-HFC model, the functional responses
421 that link the environmental variables with the response of hake, anglerfish and mullets need to be
422 parameterized. Commonly, this is obtained from the literature (Coll et al., 2016). Here we proposed
423 an alternative approach that integrates the functional responses obtained by the biomass B-HSD
424 model in the E-HFC one (Figure 3, *Path 1*). In particular, the de-standardized functional responses
425 obtained from the biomass B-HSD model outputs were integrated in the E-HFC model. We named
426 this model E-HFC FR.

427 All other functional groups received the habitat foraging settings and the original configuration of
428 habitats, following the original implementation of the E-HFC model in the study area (Coll et al.,
429 2016).

430 2) In the E-HFC model it is also possible to inform a priori the niche of the species. This process can
431 be done using bibliographic information, online databases or other information. Here we set the
432 species niche using the output of the probability of occurrence obtained from the occurrence B-HSD
433 model to force the niche priors of the E-HFC model (Figure 3 - *Path 2*)[1]. We named this model E-
434 HFC Niche.

435 Finally, in order to compare the results obtained with above approaches (Figure 3) we also run the E-
436 HFC model without any input of the B-HSD ones using the default parameterization of the published
437 study (Coll et al., 2016). We named this model E-HFC original.

438 The *Ecospace* model was executed from 1978 to 2012 and results of model simulations were
439 compared averaging the biomass predictions from 2002 to 2012. Results from all modelling outputs
440 were also compared with results from the MEDITS survey for 2002 to 2012 as a validation procedure.

[1] The niche priors are entered in the *Ecospace* user interface through the Habitat Capacity Input map layers.

### *f) Comparison and complementarity of modelling results*

Results from biomass spatial predictions for European hake, anglerfish and mullets from the biomass B-HSD and E-HFC models were compared using Spearman's spatial correlations using the "corLocal" function of the R software that allows computing this measure for two spatial objects using a focal neighborhood and thus taking into account distance. Values of Spearman's correlations range from -1 to 1, being 1 equal to a perfect positive correlation between the two datasets.

In particular, in order to test the prediction power of each model, we first interpolated the observed data of each studied species using the "interpolation TIN" function of the Quantum GIS (available in: http://www.qgis.org/en/site/) to obtain a spatial raster with a 0.05º x 0.05º degree spatial resolution. Afterwards, we compared each model spatial output with observed data using the Spearman's spatial correlation. Additionally, we computed the Root Mean Square Error (RMSE) to measure the cell-by-cell difference between the two grids (i.e. observed data and models outputs). Secondly, we compared results of the E-HFC model previously informed with the biomass B-HSD modelling outputs (Figure 3) to assess how similar or different were the outputs using B-HSD to inform the food-web model in comparison to not doing it.

### *g) Introducing uncertainties in E-HFC model*

A formal fitting and validation of the time dynamic model *Ecosim* was previously performed (Coll et al., 2013; Coll et al., 2008; Navarro et al., 2011) and it is implemented in the models we used in this study. For the *Ecospace* model a formal fitting and validation assessment is still missing, partially due to the lack of independent spatial data to perform these kinds of analyses. In this study, we present a first attempt to validate *Ecospace* results comparing results with an independent dataset of spatial observations. In addition, as the Bayesian approach uses probabilities distributions, we used INLA posterior distributions for each hyperparameter to extract n random samples and obtain an explicit quantification of the uncertainties. Therefore, we generated six random samples from the posterior

distribution of the Bayesian results of occurrence for European hake, anglerfishes and red mullets that were used to run alternative *Ecospace* configurations to complement the prediction map of the mean values. These alternative distributions are possible acceptable solutions from the Bayesian results that allow assessing the impact in *Ecospace* results. Results from *Ecospace* were used to calculate an average of the predicted biomass and the standard deviation per species. This procedure illustrates that the process can be formally implemented if B-HSD and E-HFC models are used in a complementary way (e.g., such as in Figure 3 - *Path 2*).

## 3. **Results**

### ***a) European hake***

Among the 355 hauls sampled in the studied area between 2002 and 2012, European hake was present in 245 (Figure 2a).

Regarding European hake occurrence, the best statistical fitted model (based on the lowest WAIC, RMSE and higher $R^2$) resulting from the B-HSD method showed sea bottom temperature (SBT) and the random spatial effect as relevant variables (Table 1 Supplementary Material). Particularly, the SBT presented a negative relationship with the European hake occurrence (posterior mean = -0.25; 95% CI = [-0.53, -0.12]) (Table 2 and Figure 3a in Supplementary Material). The European hake occurrence B-HSD model achieved a 0.67 of AUC and 0.62 of TSS indicating a good degree of discrimination between locations where the species is present and those where it is absent. The final

European hake B-HSD biomass model retained as relevant predictors the SBT, sea surface salinity SSS and the random spatial effect (Table 3 Supplementary Material). Specifically, a positive relationship with SBT (posterior mean = 0.39; 95% CI = [0.23, 0.54]) and a negative with SSS (posterior mean = -0.90; 95% CI = [-1.51, -0.28]) and the expected biomass was found, i.e. suggesting that there was more hake in warmer and less saltier waters (Table 1 and Figure 4a and 4b).

The predicted spatial biomass for this species resulting from the statistical fitted B-HSD model showed a main hotspot of biomass located in the coastal area and continental shelf around the Ebro Delta (Figure 5a). Considering the food-web modelling approach, results from the original configuration of the model, E-HFC, showed a larger area of high concentration of hake coinciding with the continental shelf and upper slope (Figure 5b). Results from the E-HFC FR and E-HFC Niche were similar to the E-HFC original but extended the area to the coastal region surrounding the Ebro Delta and the southern part of the study area (Figure 5c and 5d). Taking into account the random samples from the probability distributions of European hake occurrence from the Bayesian modelling (Figure 4 and 5 in Supplementary Material) illustrated the low to moderate impact that uncertainty in the niche extension can have on the final spatial prediction of biomass for European hake from *Ecospace* model, E-HFC Niche (Figure 6a and b).

***b) Anglerfishes***

Anglerfishes were present in 180 of hauls sampled between 2002 and 2012 (Figure 2b). Specifically, *Lophius budegassa* was present in 146 hauls, while *Lophius piscatorius* in 127.

The final B-HSD model of the *Lophius* spp. occurrence included as relevant predictors the SBT (Table 4 Supplementary Material), showing a negative relationship (posterior mean = -0.28; 95% CI = [-0.68, -0.19]), in addition to the spatial effect (Table 2 and Figure 3b in Supplementary Material). Occurrence predictions recorded good values of predictions validation measures (AUC = 0.67; TSS= 0.61), reflecting a high ability of the model to predict true negative and true positive correctly.

*Lophius* spp. biomass B-HSD models retained SBS and the spatial random effect as final predictors (Table 5 Supplementary Material). The biomass of the *Lophius* spp. was positively affected by SBS (posterior mean = 0.53; 95% CI = [0.18, 0.87]) and the spatial random effect (Table 1 and Figure 4c).

Regarding B-HSD predictions, higher biomass of the species was found at the end of the continental shelf and the upper slope of the central and southern study area (around 200-400 m) (Figure 7a).

512 Results from the E-HFC model showed the main biomass located surrounding the Ebro Delta area
513 and the upper part of the continental shelf of the central and southern study area (Figure 7b), while
514 results from the E-HFC FR and E-HFC Niche also showed a concentration of biomass in the Ebro
515 Delta area and southern part of the continental shelf (Figure 7c and 7d).

516 Results of the random samples from the probability distributions of anglerfishes occurrence from the
517 Bayesian modelling (Figure 6 and 7 in Supplementary Material) illustrated the low impact that
518 uncertainty in the niche extension can have on the final spatial prediction of biomass for anglerfishes
519 from *Ecospace* model, E-HFC Niche (Figure 6c and d).

520 ***c) Red mullets***

521 Red mullets were present in 249 hauls sampled between 2002 and 2012 (Figure 2c). Particularly,
522 *Mullus surmuletu*s was present in 158 hauls, while the *Mullus barbatus* was in 222.

523 The B-HSD occurrence fitted model of *Mullus* spp. showed a positive relationship with primary
524 production (posterior mean = 0.64; 95% CI = [0.24, 1.28]) (Tables 2, 6 and Figure 3c Supplementary
525 Material). As for the other species, predictions measures presented good values (AUC = 0.68; TSS=
526 0.63). Similarly to the occurrence, for *Mullus* spp. biomass, the B-HSD model showed a positive
527 relationship with primary production (posterior mean = 4.23; 95% CI = [2.25, 5.54]) (Table 1 and
528 Figure 4d), as well with the random spatial effect (Table 7 Supplementary Material).

529 The predicted biomass of these species considering the B-HSD model was higher around the coastal
530 areas of the Ebro Delta and northern areas coinciding with the most productive region (Figure 8a).
531 Results from the E-HFC, E-HFC FR and E-HFC Niche models showed the main biomass also located
532 surrounding the Ebro Delta coast and northern areas and the last two showed more similar patterns to
533 the B-HSD results (Figure 8b, 8c and 8d).

534 Results of the random samples from the probability distributions of red mullets occurrence from the
535 Bayesian modelling (Figure 8 and 9 in Supplementary Material) illustrated the moderate impact that

uncertainty in the niche extension can have on the final spatial prediction of biomass for these species from *Ecospace* model, E-HFC Niche (Figure 6e and f).

***d) Comparison of results***

Modelling results regarding biomass distributions showed positive and significant correlations with observational data from the MEDITS project (Figure 9a, 9b and 9c). Correlations between modelling results and observational data were similar between modelling methodologies: they were between 72-81% for European hake (with largest values for the E-HFC-FR model), 77%-97% for anglerfishes (with largest values for the E-HFC-FR model) and 80%-98% for red mullets (with largest values for the E-HFC-FR and E-HFC-Niche models).

The Spearman correlations between modelling results showed highly positive and significant correlations between results for all the species considered (Figure 9a, 9b and 9c). In all cases, the correlations were higher when B-HSD results drove the E-HFC model, than when both methods were used individually.

With respect to the RMSE measure, low values were achieved by models ranging from 0.23 to 1.4. With the exception of the mullets, the RMSE of the E-HFC models driven by B-HSD results obtained lower RMSE with respect to the original E-HFC model (Table 2).

In terms of absolute predictions, results from the modelling techniques showed larger divergences, even when compared to observational data (Table 3). When comparing the mean values, the E-HFC modelling results yielded closest values compared to observational data for European hake and anglerfish, and B-HSD models were closer to red mullets. The B-HSD means were the largest values regarding hake and anglerfishes, while E-HFC Niche mean values were the largest for red mullets. In general, results from the Bayesian models predicted larger values of biomass than the food-web models for European hake and anglerfish, and lower for red mullets.

559 4. **Discussion**

560 Studying the factors that affect the spatial distribution of marine biodiversity is a central issue to
561 ecology, essential for evaluating biodiversity patterns, for predicting the impact of environmental
562 change and anthropogenic activities, and for designing useful management programs (Navarro et al.,
563 2015). Advancing the knowledge of marine species distributions is also essential to contribute to an
564 ecosystem-based management approach to marine resources (Pennino et al., 2016) and an ecosystem-
565 based spatial planning approach (Moore et al., 2016). Although effective management of marine
566 ecosystems requires information on the spatial distribution of marine species, there is still marked
567 paucity of our understanding of species-environment relationships. Modelling techniques have been
568 identified as key tools to contribute to this knowledge (Guisan and Zimmermann, 2000).

569 This study illustrates the application of two modelling techniques, applied in isolation and in
570 conjunction, to predict the distribution of occurrence and biomass of commercial marine species in a
571 small area of the NW Mediterranean Sea: a statistical approach, using a Bayesian Hierarchical Species
572 Distribution Model, and a mechanistic approach, using a spatial-temporal food-web model
573 (*Ecospace*).

574 Our analyses show that the results of both modelling approaches positively and significantly
575 correlated with observational data, confirming that both techniques are valid tools to predict species
576 distributions in the study area. In addition, both techniques showed high correlation with
577 observations, with a maximum of 29% of data not described by the models. Small divergences
578 between observations and predictions may be because the biological observational data used in this
579 case study represents one season only (late spring to early summer), while the environmental data
580 and food-web parameterization are expressed as annual mean values. A recent study has shown that
581 exploited demersal communities exhibit strong seasonal changes, even in small areas such as the one
582 used in our study (Vilás-González, 2016; Vilas-González et al., Submitted). Divergences between
583 models predictions and observations may also be because while the statistical Bayesian modelling

584 technique provides information about the potential niche of the sampled population, the food-web
585 modelling technique provides a prediction about the realized niche of the entire population. It is
586 necessary to undertake a similar study using observational data that represents the annual state of the
587 ecosystem to explore this issue deeper. Despite this data is still not available in our study area, this
588 study provides an example of application than can be repeated in a data-richer area.

589 Our results show clear heterogenic distributions of the demersal fish species in the study area, with
590 main concentrations in the coastal and shelf area. The Ebro River delta is known to be a productive
591 region; it is a high run-off area that boosts primary production and consequently generates the
592 preferential habitat for many species, especially in the recruits phase (Lloret and Lleonart, 2002).
593 Here, also important demersal species are surmised to concentrate (Navarro et al., 2015). However,
594 little is known about the spatial distribution patterns for the three species considered in the study area,
595 and the Western Mediterranean in general. Spatial information is mainly derived from studies of
596 distributions of eggs and larvae, and from sampling in recruitment areas (e.g., Hidalgo et al., 2008;
597 Morfin et al., 2012). In the case of European hake, Olivar and co-authors (Olivar et al., 2003) showed
598 that in the northern part of our study area, distributions of eggs and larvae coincide with the principal
599 fishing grounds of the spawning population, the shelf edge (Recasens et al., 1998). Maynou et al.
600 (Maynou et al., 2003) found that hake nursery areas were mainly located in the central part of the
601 continental shelf between 68 and 168 m depth. In addition, the waters surrounding the Columbretes
602 Islands show high European hake biomass estimates; these islands have been a Marine Protected Area
603 since 1989, and thus provide a stable, high quality ecosystem to stocks that could be exporting hake
604 adults and recruits to adjacent areas (Paradinas et al., 2015; Stobart et al., 2009).

605 In the case of *Lophius* spp. Lopez and co-authors (López et al., 2016) showed that juvenile specimens
606 in July 2013 were mainly located in coastal and continental shelve areas, while higher presence of
607 adults was found closer to the upper slope areas. *Mullus* spp. has shown to be distributed in coastal
608 and shelf areas, where *M. barbatus* is more abundant over muddy bottoms with maximum abundance

609 in the 50 – 200 m depth stratus. *M. surmuletus* prefers rough substrates and ranges between 10 to 50
610 m, at least in springtime (Lombarte et al., 2000; Tserpes et al., 2002). Overall, our results match the
611 available distribution information, although future spatial-temporal modelling work should aim to
612 analyse these species separately and their seasonal dynamics in-depth.

613 In terms of which variables could be driving the spatial distributions and abundance of the studied
614 species, our results from B-HSD models highlighted a strong correlation with temperature. European
615 hake and anglerfishes biomass was higher in warmer and less salty waters consistently with their
616 depth ranges, while mullets were more abundant in high productivity habitats. These finding are in
617 line with other studied from different areas such as Maravelias and co-authors (Maravelias et al.,
618 2007), who found that hake and red mullet abundance were higher in waters > 19 ºC in the Aegean
619 Sea. Similarly, Massutí and co-authors (Massutí et al., 2008) found a negative significant relationship
620 between sea bottom salinity and European hake catch per unit of effort around the Balearic Islands
621 waters. This is because European hake depth range coincides with a decrease of salinity
622 approximately from 50-400 m depth.

623 In addition, our results highlighted that the spatial random effects estimated from the B-HSD were
624 relevant for all the species in all selected models. The spatial random component is often used to
625 capture the effect of important missing predictors (F Dormann et al., 2007) or to account for
626 ecological processes (e.g., dispersal or aggregative behaviour) or anthropogenic effect that are
627 spatially structured (Merow et al., 2014). In our case, the spatial effect probably reflected a
628 combination of effects, including seasonality and impact of anthropogenic activities and their
629 synergies (Coll et al., 2012; Navarro et al., 2015; Vilas-González et al., Submitted).

630 Our study also showed positive correlations between methods, illustrating that statistical species
631 distribution and food-web spatial-temporal modelling can yield corresponding predictions. However,
632 biomass predictions from the Bayesian models were larger than the food-web models for European
633 hake and anglerfish, and lower for red mullets. This may be due to the fact that food-web models

directly include the effect of fisheries removals, reducing the potential biomass predicted by the Bayesian model for predators. On the contrary, higher biomass of red mullets could reflect indirect trophic effects such as competition or predation release.

Moreover, our findings suggest that it is best to use the modelling tools in a complementary way: projected species distributions are more similar to observations when the food-web model technique is informed with results from the species distribution modelling technique. In *Ecospace*, this was accomplished by either using the response functions derived from the statistical analysis in the *Ecospace* calculations, or by forcing the niche priors of *Ecospace* foraging capacity directly with results from the statistical models (Figure 3, *Path 1* and *Path 2*, respectively). These results represent a promising venue to develop further case studies to test the methodology complementarity, especially in larger study areas and with larger and more completed spatial-temporal datasets. Hybrid approaches between correlative and mechanistic methods can be pragmatic solutions to improving ecological predictions by adding key mechanisms to simple models (Urban et al., 2016). These methodologies can be part of an essential toolbox for promoting more efficient management by providing more accurate spatial species predictions.

Both techniques used in this study have advantages and disadvantages. The use of Bayesian Hierarchical Species Distribution models can be a powerful approach given that it quantifies both the spatial magnitude and the different sources of uncertainty. It has been demonstrated that this type of models can produce reliable species habitat predictions even in data-poor situation working with small sample size of 50 observations (Fonseca et al., 2017). However, despite these models popularity, they only implicitly consider interactions between species, disregarding the potentially important influence of biotic interactions (such as competition, predation and facilitation) and can become unreliable when they are used to extrapolate to novel conditions (Urban et al., 2016). When using the food-web modelling we can directly incorporate biotic interactions and the dynamics and effects of exploitation and spatial management measures (Christensen and Walters, 2004). However, spatial

food-web modelling approaches require large amounts of spatialized data and its integration needs to be informed by previous knowledge about which key environmental elements drive key processes. By combining the capabilities of the food-web modelling approach with the statistical modelling, we can gain clear advantages from both techniques and make food-web modelling more robust. For example, when using *Ecospace* with the response functions obtained from the statistical models (Figure 3, *Path 1*) we can directly model the impact of environmental change on species dynamics and interactions by capturing key spatial-temporal changes of the environmental drivers (Steenbeek et al., 2013). This is an essential aspect for developing future scenarios of change of marine ecosystems and predict future marine food-web configurations. Alternatively, forcing the niche calculations of *Ecospace* with results from external SDM tools (Figure 3, *Path 2*) may provide a valid shortcut when spatial-temporal datasets are not fully available but we have validated knowledge about niche changes.

Our study illustrates that by combining both approaches, uncertainty analyses can be developed. This is essential to contribute to the future development needs of *Ecospace.* Although there are several ways that uncertainty can be incorporated in *EwE* (Coll and Steenbeek, 2017; Heymans et al., 2016; Steenbeek et al., 2018), current rudimentary means to formally validate its spatial-temporal predictions against observations need to improve and *Ecospace* requires facilities to first fitting its behaviour to time series of spatial-temporal data such as it is supported in *Ecosim* (Mackinson et al., 2009; Scott et al., 2016). In our study we present a first attempt to move towards this direction to formally develop a validation tool considering data uncertainty, as it is illustrated here using a random sample of posterior distribution results from Bayesian statistics into *Ecospace* (Figure 6). The new *Ecospace* with the HCR model and the spatial-temporal framework will substantially benefit from these future capabilities (Coll et al., 2016; Romagnoni et al., 2015; Steenbeek et al., 2018).

683 **Acknowledgements**

684 MC was partially funded by the European Commission through the Marie Curie Career Integration
685 Grant Fellowships – PCIG10-GA-2011-303534 - to the BIOWEB project. This study is a contribution
686 to the project ECOTRANS (CTM2011-26333, Ministerio de Economía y Competitividad, Spain) and
687 SafeNET (EU-DGMARE MARE/2014/41). MC and JS acknowledge financial support by the
688 European Union´s Horizon research program grant agreement No 689518 for the MERCES project.
689 The authors express their gratitude to all the people that work in the MEDITS surveys. MEDITS data
690 collection has been co-funded by the EU through the European Maritime and Fisheries Fund (EMFF)
691 within the National Program of collection, management and use of data in the fisheries sector and
692 support for scientific advice regarding the Common Fisheries Policy.

693

694 **Figure captions**

695 **Figure 1.** Study area located in the Northwestern Mediterranean Sea.

696 **Figure 2.** (a-c) Sampling stations of the MEDITS surveys used for European hake (*Merluccius*
697 *merluccius*), anglerfishes (*Lophius* spp.) and red mullets (*Mullus* spp.); Presence is indicated
698 with red dots, while absence with black dots. (d-f) Observed log-transformed Biomass (kg/km2)
699 averages during 2002-2012.

700 **Figure 3.** Working path representing two directions on how Bayesian-Hierarchical Species
701 Distribution modelling (B-HSD) results can be integrated into *Ecospace* food-web model: (*Path*
702 *1*) by using response functions that describe the link between environmental factors and species
703 responses (E-HFC FR), or (*Path 2*) by forcing the niche priors of the Foraging Capacity Model
704 with niche calculations from B-HSD results (E-HFC Niche).

705 **Figure 4.** Functional response of B-HSD biomass models (kg/km$^2$) for a-b) European hake (*M.*
706 *merluccius*), c) anglerfishes (*Lophius* spp.) and d) red mullets (*Mullets* spp.). Predictor acronyms
707 are SBT = Sea Bottom Temperature; SSS= Sea Surface Salinity; SBS= Sea Bottom Salinity; PP=
708 Primary production. The solid line is the smooth function estimate and shaded regions represent
709 95% credibility interval (CI).

710 **Figure 5.** Predicted biomass distributions (log(kg/km$^2$)) for European hake (*M. merluccius*) using a)
711 Bayesian Hierarchical Species Distribution (B-HSD) model, b) *Ecospace* Habitat Foraging
712 Capacity model in its original configuration (E-HFC), c) *Ecospace* Habitat Foraging Capacity
713 model informed with functional responses from B-HSD (Figure 3, *Path 1*) (E-HFC FR), and d)
714 *Ecospace* Habitat Foraging Capacity model with the niche calculations driven by results from
715 B-HSD (Figure 3, *Path 2*) (E-HFC Niche).

716 **Figure 6.** Average and standard deviation of predicted biomass distributions (log(kg/km$^2$)) for

European hake (*M. merluccius*) (a-b), anglerfishes (*Lophius* spp.) (c-d) and red mullets (*Mullus* spp.) (e-f) resulting from *Ecospace* Habitat Foraging Capacity model informed with the niche calculations of random samples from the B-HSD results (Figure 3, *Path 2*) (E-HFC Niche) (six random samples are plotted in Supplementary Material, Figures 4-9) .

**Figure 7.** Predicted biomass distributions (log(kg/km$^2$)) for anglerfishes (*Lophius* spp.) using a) Bayesian Hierarchical Species Distribution (B-HSD) model, b) *Ecospace* Habitat Foraging Capacity model in its original configuration (E-HFC), c) *Ecospace* Habitat Foraging Capacity model informed with functional responses from B-HSD (Figure 3, *Path 1*) (E-HFC FR), and d) *Ecospace* Habitat Foraging Capacity model with the niche calculations driven by results from B-HSD (Figure 3, *Path 2*) (E-HFC Niche).

**Figure 8.** Predicted biomass distributions (log(kg/km$^2$)) for red mullets (*Mullus* spp.) using a) Bayesian Hierarchical Species Distribution (B-HSD) model, b) *Ecospace* Habitat Foraging Capacity model in its original configuration (E-HFC), c) *Ecospace* Habitat Foraging Capacity model informed with functional responses from B-HSD (Figure 3, *Path 1*) (E-HFC FR), and d) *Ecospace* Habitat Foraging Capacity model with the niche calculations driven by results from B-HSD (Figure 3, *Path 2*) (E-HFC Niche).

**Figure 9.** Spearman correlation results between observed biomass data and the four different modelling configuration and estimates from the MEDITS project (2002-2012) for a) European hake (*M. merluccius*), b) anglerfishes (*Lophius* spp.) and c) red mullets (*Mullets* spp.). B-HSD: Bayesian Hierarchical Species Distribution model; E-HFC: *Ecospace* Habitat Foraging Capacity original model; E-HFC FR: *Ecospace* Habitat Foraging Capacity model informed with functional responses from B-HSD; and E-HFC Niche: *Ecospace* Habitat Foraging Capacity model with the niche calculations driven by results from B-HSD.

742 **Tables**

743 **Table 1.** Numerical summary of the posterior distribution of the fixed effects for the best biomass B-
744 HSD of the three species studied. This summary contains the mean, the standard deviation, the
745 median and a 95% credible interval, which is a central interval containing 95% of the probability
746 under the posterior distribution. Predictors' acronyms are: SBT = Sea Bottom Temperature; SSS
747 = Sea Surface Salinity; SBS = Sea Bottom Salinity; PP = primary production (PP).

748

| Species | Predictor | Mean | SD | $Q_{0.025}$ | $Q_{0.5}$ | $Q_{0.975}$ |
|---|---|---|---|---|---|---|
| European hake | Intercept | 1.08 | 0.05 | 0.56 | 0.98 | 1.96 |
| | SBT | 0.39 | 0.07 | 0.23 | 0.39 | 0.54 |
| | SSS | -0.90 | 0.31 | -1.51 | -0.90 | -0.28 |
| Anglerfishes | Intercept | 0.23 | 0.01 | 0.11 | 0.25 | 0.54 |
| | SBS | 0.53 | 0.04 | 0.18 | 0.54 | 0.87 |
| Red mullets | Intercept | 0.23 | 0.02 | 0.13 | 0.12 | 0.52 |
| | PP | 4.23 | 0.09 | 2.25 | 3.96 | 5.54 |

749
750
751
752
753
754
755

**Table 2:** Root Mean Square Error (RMSE) results between observed biomass data and the four different modelling configurations. Acronyms are: B-HSD: Bayesian Hierarchical Species Distribution model; E-HFC: *Ecospace* Habitat Foraging Capacity original model; E-HFC FR: *Ecospace* Habitat Foraging Capacity model informed with functional responses from B-HSD; and E-HFC Niche: *Ecospace* Habitat Foraging Capacity model with the niche calculations driven by results from B-HSD.

| RMSE | B-HSD | E-HFC | E-HFC FR | E-HFC Niche |
|---|---|---|---|---|
| European hake (*M. merluccius*) | 0.51 | 1.4 | 0.37 | 0.95 |
| Anglerfishes (*Lophius* spp.) | 0.23 | 1.11 | 0.07 | 0.12 |
| Red mullets (*Mullets* spp.) | 0.38 | 0.26 | 0.44 | 0.43 |

771 **Table 3.** Comparison of spatial model minimum, mean and maximum values of biomass estimates
772 (kg/km$^2$) of European hake (*Merluccius merluccius*), anglerfishes (*Lophius* spp.) and red mullets
773 (*Mullus* spp.) under the four different modelling configuration and estimates from the MEDITS
774 project (2002-2012). B-HSD: Bayesian Hierarchical Species Distribution model; E-HFC: *Ecospace*
775 Habitat Foraging Capacity original model; E-HFC FR: *Ecospace* Habitat Foraging Capacity model
776 informed with functional responses from B-HSD; and E-HFC Niche: *Ecospace* Habitat Foraging
777 Capacity model with the niche calculations driven by results from B-HSD.

778
779

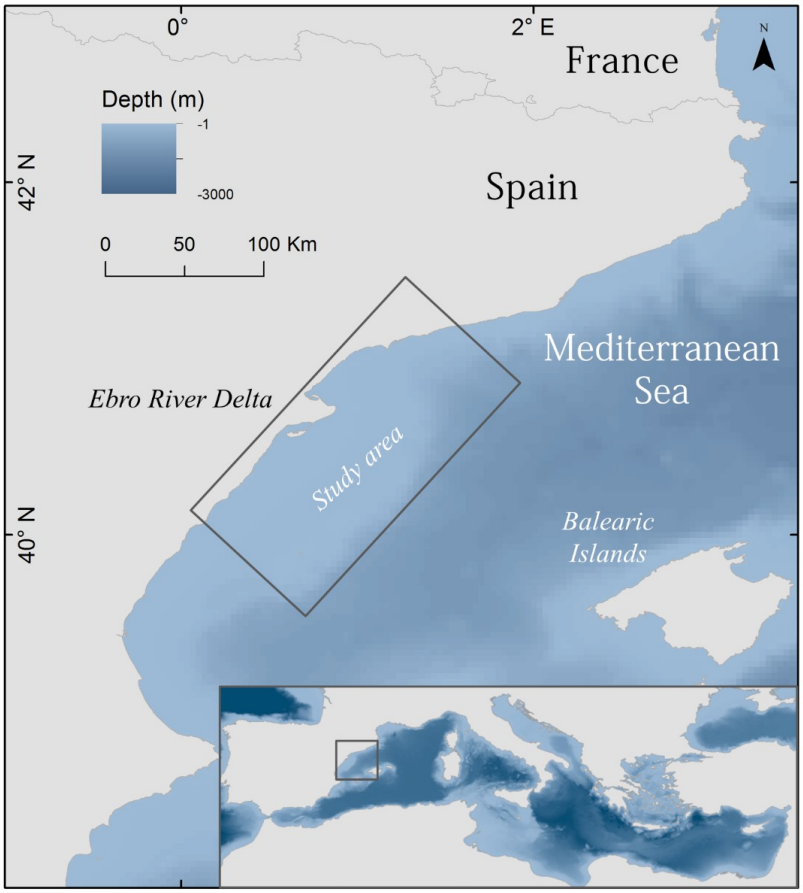
| Species | Model | Min | Mean | Max |
|---|---|---|---|---|
| European Hake | B-HSD | 0.00 | 475.60 | 1431.20 |
| | E-HFC | 0.00 | 17.41 | 435.41 |
| | E-HFC-FR | 0.00 | 106.29 | 3890.15 |
| | E-HFC-Niche | 0.00 | 259.14 | 20958.07 |
| | Observed biomass | 0.00 | 12.83 | 958.74 |
| Anglerfish | B-HSD | 0.00 | 210.70 | 951.40 |
| | E-HFC | 0.00 | 6.96 | 476.17 |
| | E-HFC-FR | 0.00 | 23.57 | 1744.93 |
| | E-HFC-Niche | 0.00 | 37.44 | 2356.01 |
| | Observed biomass | 0.00 | 3.14 | 389.76 |
| Red mullets | B-HSD | 0.00 | 32.2 | 722.15 |
| | E-HFC | 0.00 | 64.57 | 2300.44 |
| | E-HFC-FR | 0.00 | 162.05 | 5143.90 |
| | E-HFC-Niche | 0.00 | 164.25 | 5059.71 |
| | Observed biomass | 0.00 | 10.81 | 2404.08 |

780
781

782

783

784 **Figure 1.**
785




786
787

**Figure 2.**

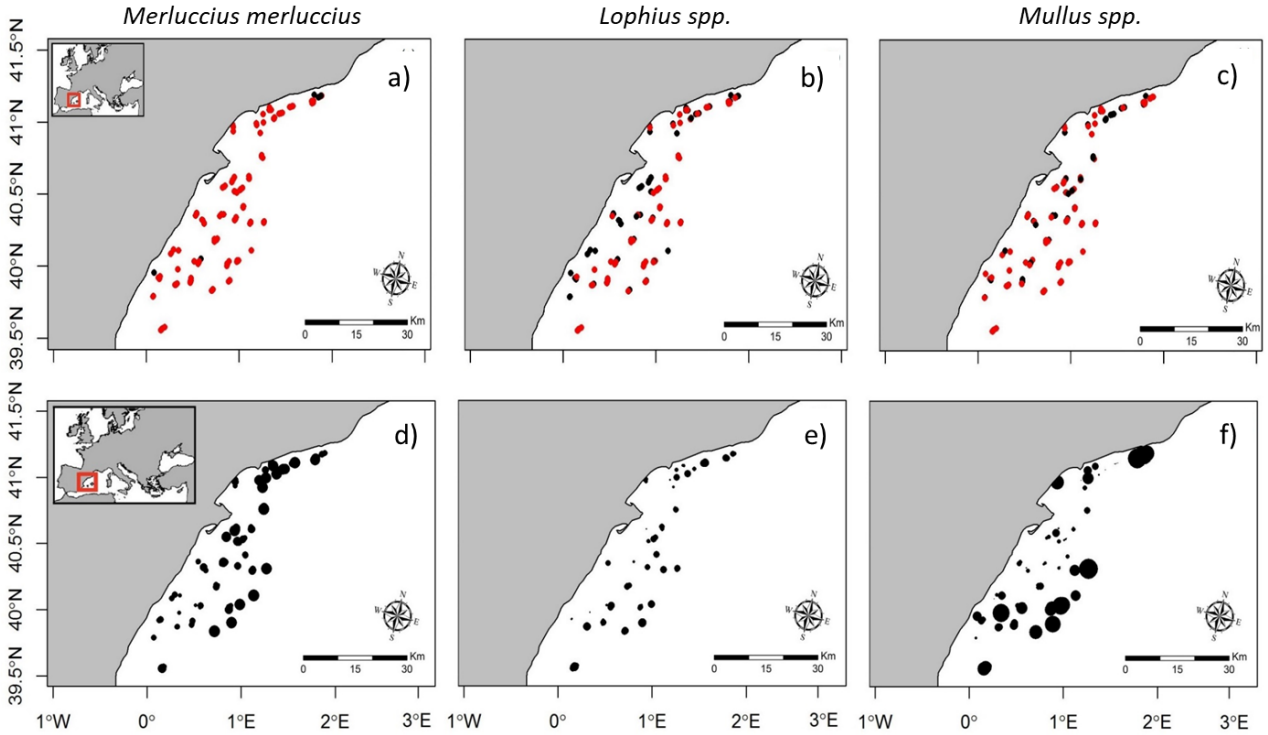

Merluccius merluccius
Lophius spp.
Mullus spp.
a)
b)
c)
d)
e)
f)
41.5°N
41°N
40.5°N
40°N
39.5°N
1°W
0°
1°E
2°E
3°E
0
15
30
Km

**Figure 3.**

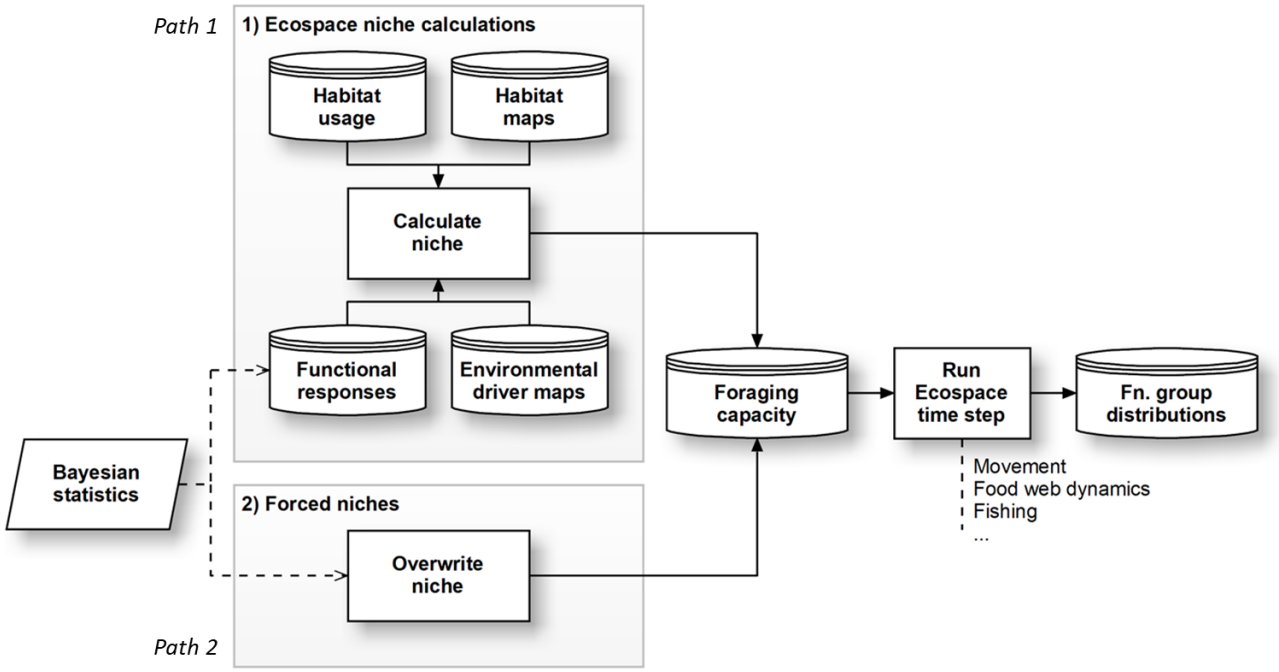

Path 1
1) Ecospace niche calculations
Habitat usage
Habitat maps
Calculate niche
Functional responses
Environmental driver maps
Foraging capacity
Run Ecospace time step
Fn. group distributions
Movement
Food web dynamics
Fishing
...
Bayesian statistics
2) Forced niches
Overwrite niche
Path 2

**Figure 4.**

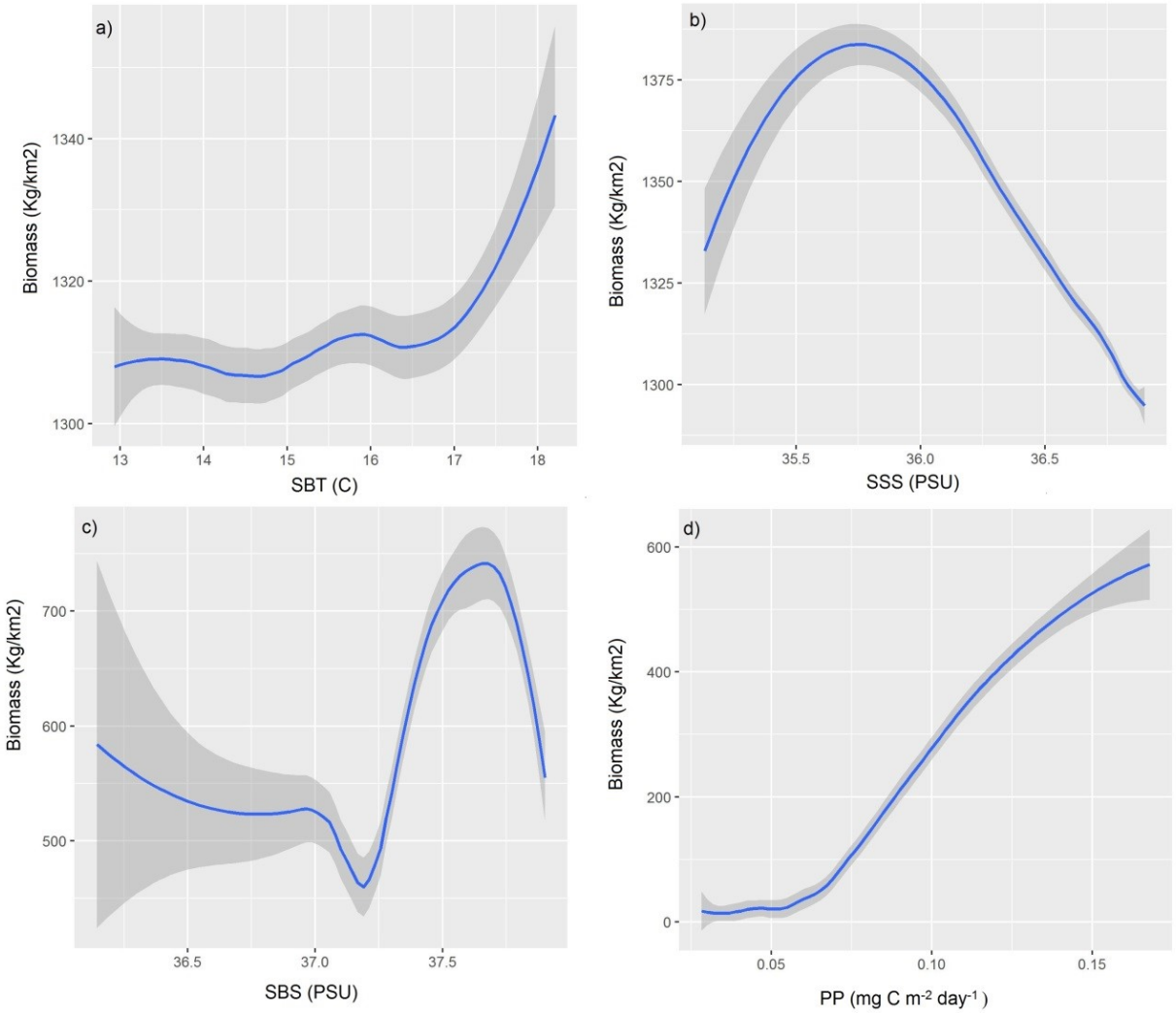
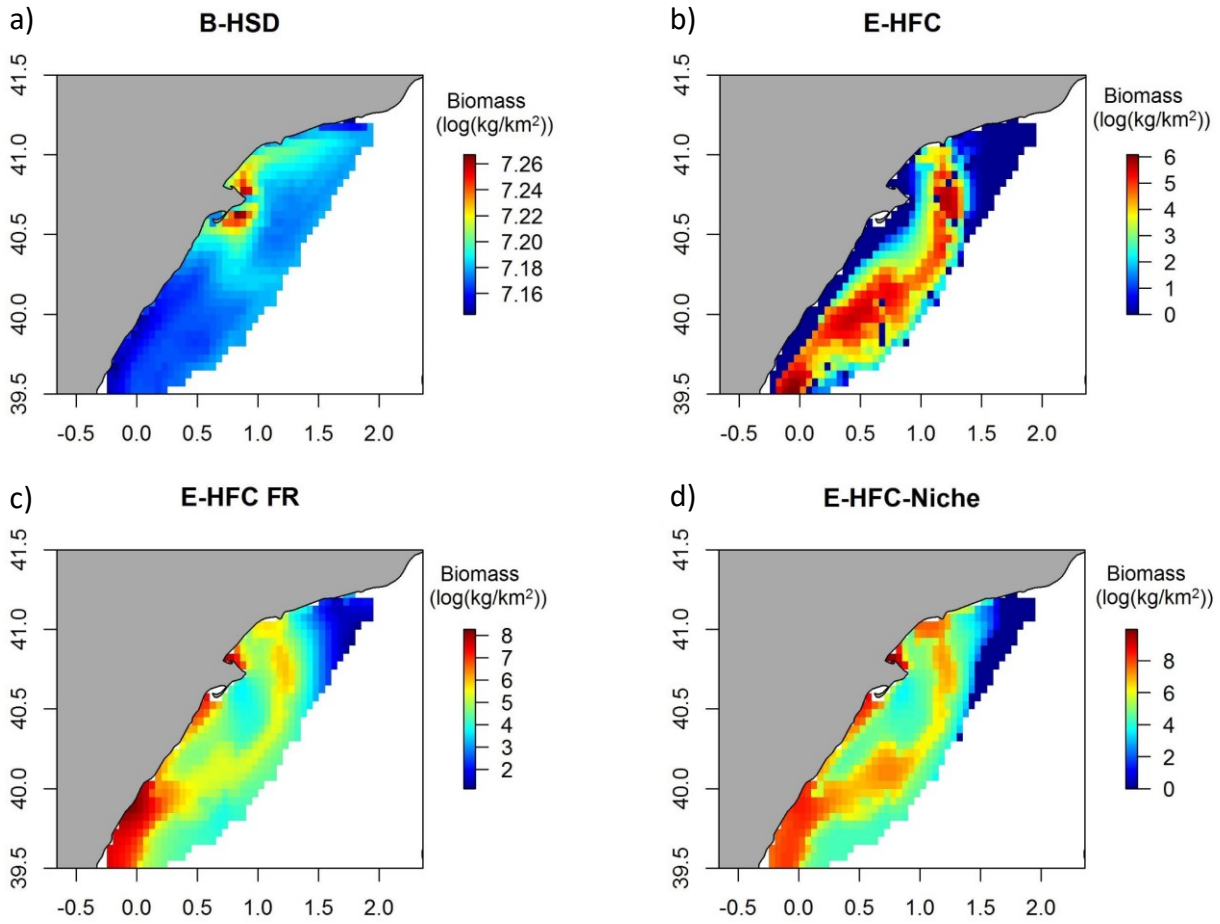
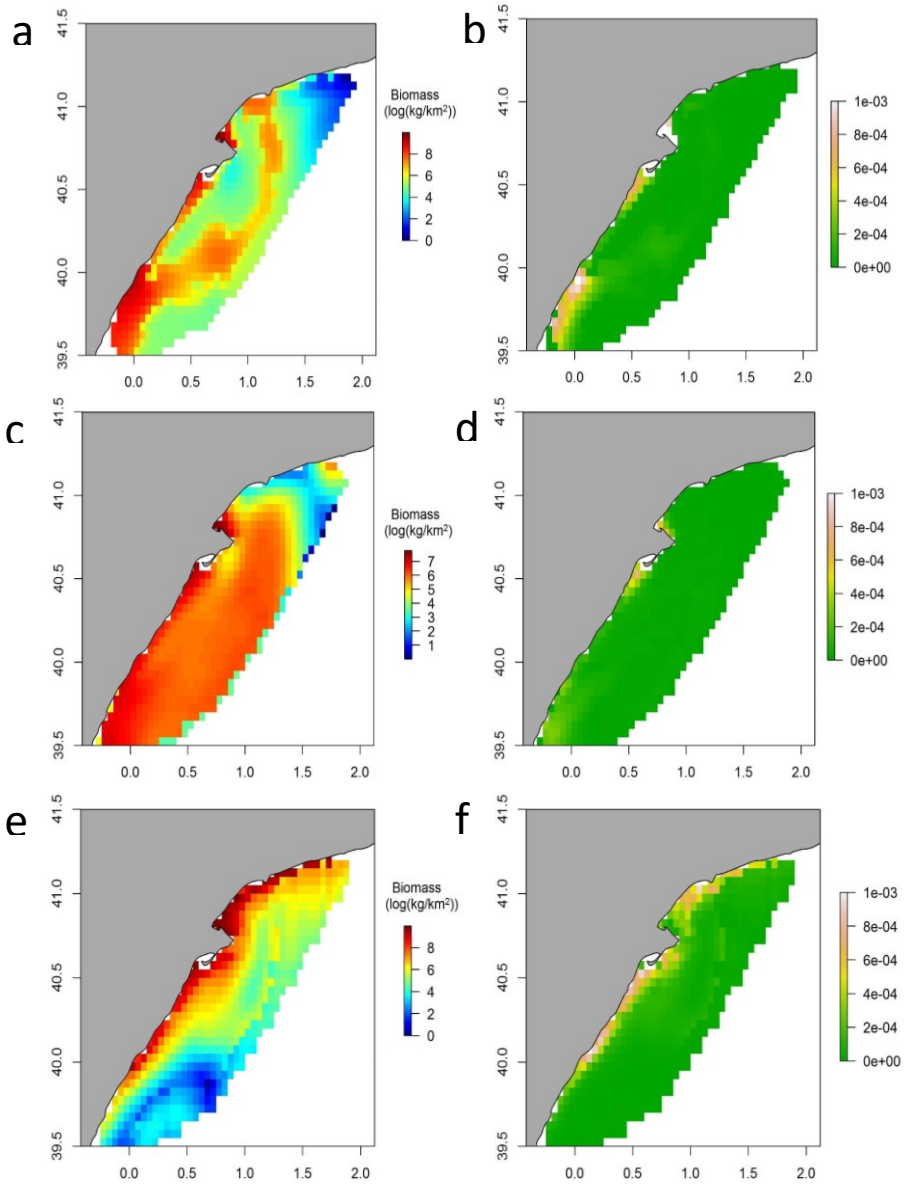

a)
Biomass (Kg/km2)
1340
1320
1300
13
14
15
16
17
18
SBT (C)
b)
Biomass (Kg/km2)
1375
1350
1325
1300
35.5
36.0
36.5
SSS (PSU)
c)
Biomass (Kg/km2)
700
600
500
36.5
37.0
37.5
SBS (PSU)
d)
Biomass (Kg/km2)
600
400
200
0
0.05
0.10
0.15
PP (mg C m-2 day-1 )

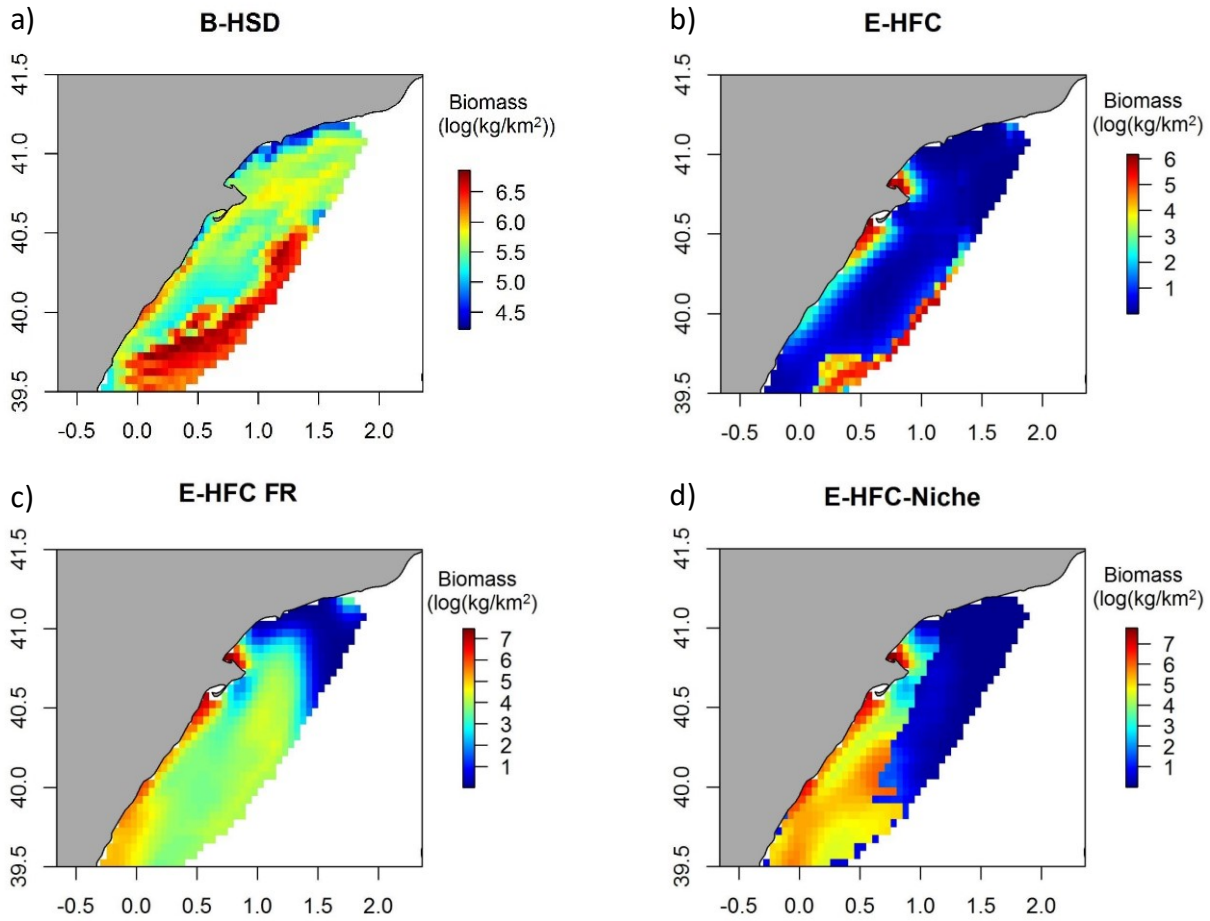
**Figure 5.**


a)
B-HSD
Biomass
(log(kg/km²))
7.26
7.24
7.22
7.20
7.18
7.16
b)
E-HFC
Biomass
(log(kg/km²))
6
5
4
3
2
1
0
c)
E-HFC FR
Biomass
(log(kg/km²))
8
7
6
5
4
3
2
d)
E-HFC-Niche
Biomass
(log(kg/km²))
8
6
4
2
0
41.5
41.0
40.5
40.0
39.5
-0.5
0.0
0.5
1.0
1.5
2.0

804 **Figure 6.**
805
806



807 **Figure 7.**
808




809
810

**Figure 8.**

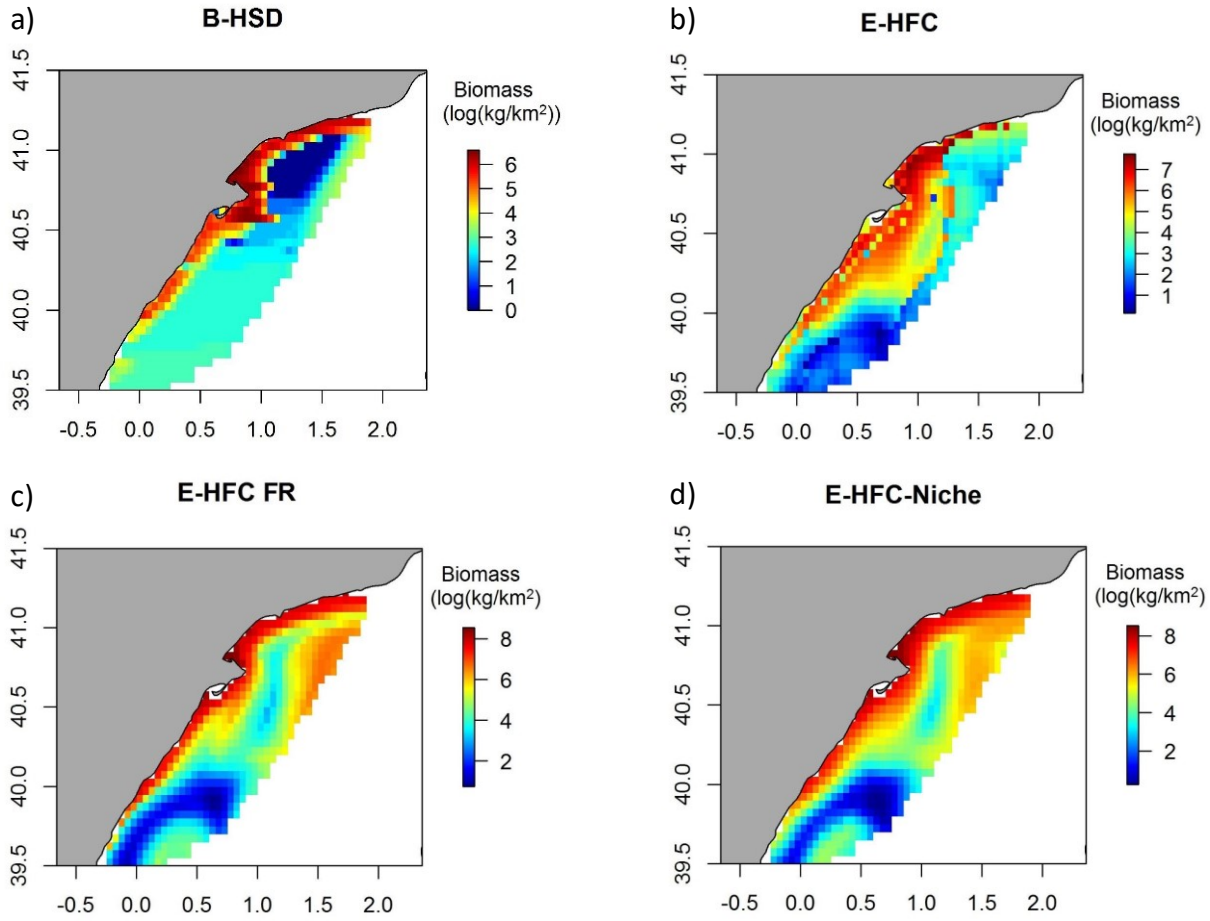
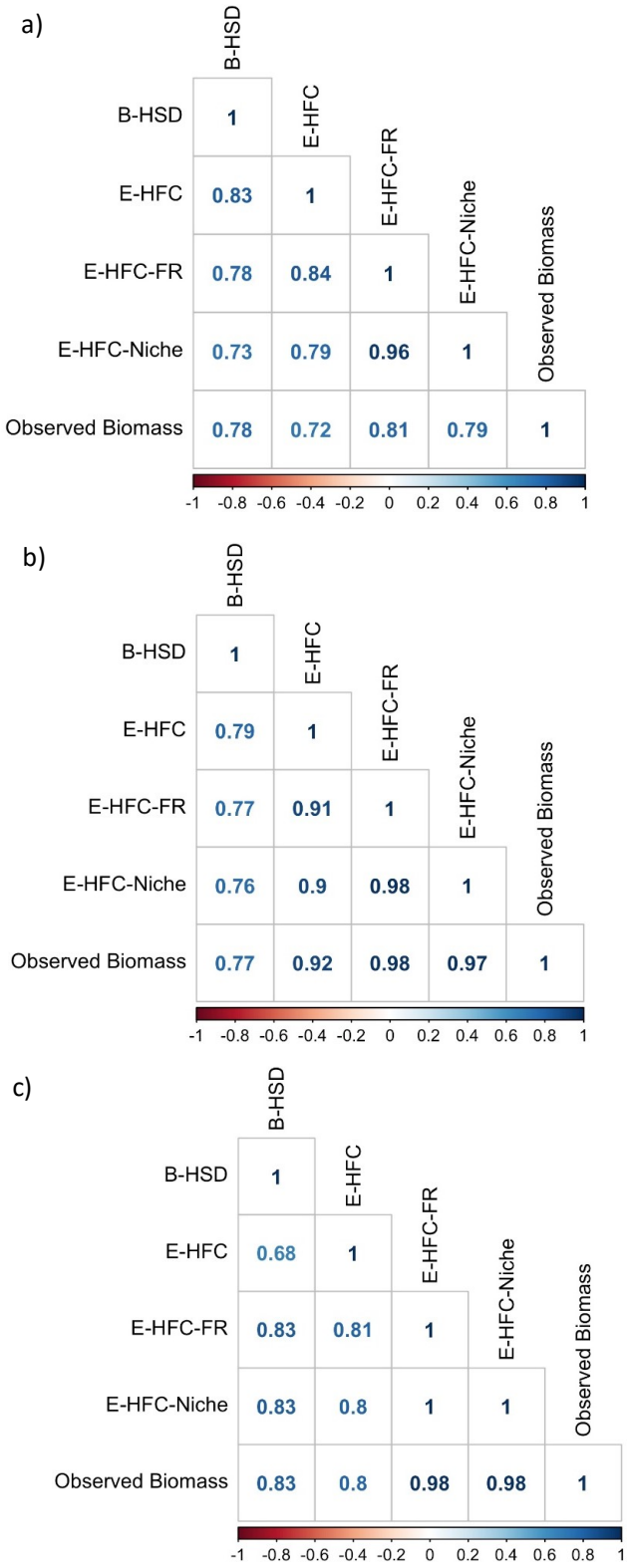

a)
B-HSD
Biomass
(log(kg/km²))
b)
E-HFC
Biomass
(log(kg/km²)
c)
E-HFC FR
Biomass
(log(kg/km²)
d)
E-HFC-Niche
Biomass
(log(kg/km²)
41.5
41.0
40.5
40.0
39.5
-0.5
0.0
0.5
1.0
1.5
2.0

**Figure 9.**


a)
B-HSD
E-HFC
E-HFC-FR
E-HFC-Niche
Observed Biomass
B-HSD 1
E-HFC 0.83 1
E-HFC-FR 0.78 0.84 1
E-HFC-Niche 0.73 0.79 0.96 1
Observed Biomass 0.78 0.72 0.81 0.79 1
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
b)
B-HSD
E-HFC
E-HFC-FR
E-HFC-Niche
Observed Biomass
B-HSD 1
E-HFC 0.79 1
E-HFC-FR 0.77 0.91 1
E-HFC-Niche 0.76 0.9 0.98 1
Observed Biomass 0.77 0.92 0.98 0.97 1
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
c)
B-HSD
E-HFC
E-HFC-FR
E-HFC-Niche
Observed Biomass
B-HSD 1
E-HFC 0.68 1
E-HFC-FR 0.83 0.81 1
E-HFC-Niche 0.83 0.8 1 1
Observed Biomass 0.83 0.8 0.98 0.98 1
-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

822 **References**


823 Adani, M., Dobricic, S., Pinardi, N., 2011. Quality assessment of a 1985-2007 mediterranean sea
824 reanalysis. Journal of Atmospheric and Oceanic Technology 28, 569-589.
825 Ahrens, R.N.M., Walters, C.J., Christensen, V., 2012. Foraging arena theory. Fish and Fisheries 13,
826 41-59.
827 Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models:
828 prevalence, kappa and the true skill statistic (TSS). Journal of applied ecology 43, 1223-1232.
829 Banerjee, S., Carlin, B.P., Gelfand, A.E., 2014. Hierarchical modeling and analysis for spatial data.
830 Crc Press.
831 Bertrand, J.A., De Sola, L.G., Papaconstantinou, C., Relini, G., Souplet, A., 2002. The general
832 specifications of the MEDITS surveys. Scientia Marina 66, 9-17.
833 Bosc, E., Bricaud, A., Antoine, D., 2004. Seasonal and interannual variability in algal biomass and
834 primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations.
835 Global Biogeochemical Cycles 18, doi:10.1029/2003GB002034.
836 Christensen, V., Coll, M., Buszowski, J., Cheung, W., Frölicher, T., Steenbeek, J., Stock, C.A.,
837 Watson, R., Walters, C., 2015. The global ocean is an ecosystem: Simulating marine life and
838 fisheries. Global Ecology and Biogeography 24 507-517.
839 Christensen, V., Coll, M., Steenbeek, J., Buszowski, J., Chagaris, D., Walters, C.J., 2014.
840 Representing variable habitat quality in a spatial food web model. Ecosystems 17, 1397-1412.
841 Christensen, V., Ferdaña, Z., Steenbeek, J., 2009. Spatial optimization of protected area placement
842 incorporating ecological, social and economical criteria. Ecological Modelling 220, 2583-2593.
843 Christensen, V., Guenette, S., Heymans, J.J., Walters, C., Watson, R., Zeller, D., Pauly, D., 2003.
844 Hundred-year decline of North Atlantic predatory fishes. Fish and Fisheries 4, 1-24.
845 Christensen, V., Maclean, J., 2011. Ecosystem Approaches to Fisheries: A Global Perspective.
846 Cambridge University Press, Cambridge 325 pp.
847 Christensen, V., Pauly, D., 1992. ECOPATH II - A software for balancing steady-state ecosystem
848 models and calculating network characteristics. Ecological Modelling 61, 169-185.
849 Christensen, V., Walters, C., 2004. Ecopath with Ecosim: methods, capabilities and limitations.
850 Ecological Modelling 72, 109-139.
851 Christensen, V., Walters, C., Pauly, D., Forrest, R., 2008. Ecopath with Ecosim version 6. User Guide
852 - November 2008. Lenfest Ocean Futures Project 2008, 235 pp.
853 Authro, 2005. Ecopath with Ecosim: A user's guide. Vancouver, B.C. (Canada).
854 Coll, M., Navarro, J., Palomera, I., 2013. Ecological role, fishing impact, and management options
855 for the recovery of a Mediterranean endemic skate by means of food web models. Biological
856 Conservation 157, 108-120.
857 Coll, M., Palomera, I., Tudela, S., Dowd, M., 2008. Food-web dynamics in the South Catalan Sea
858 ecosystem (NW Mediterranean) for 1978-2003. Ecological Modelling 217, 95-116.
859 Coll, M., Palomera, I., Tudela, S., Sardà, F., 2006. Trophic flows, ecosystem structure and fishing
860 impacts in the South Catalan Sea, Northwestern Mediterranean. Journal of Marine Systems 59,
861 63-96.
862 Coll, M., Piroddi, C., Albouy, C., Ben Rais Lasram, F., Cheung, W., Christensen, V., Karpouzi, V., Le
863 Loc, F., Mouillot, D., Paleczny, M., Palomares, M.L., Steenbeek, J., Trujillo, P., Watson, R., Pauly,
864 D., 2012. The Mediterranean Sea under siege: spatial overlap between marine biodiversity,
865 cumulative threats and marine reserves. Global Ecology and Biogeography 21, 465-480.
866 Coll, M., Steenbeek, J., 2017. Standardized ecological indicators to assess aquatic food webs: the
867 ECOIND software plug-in for Ecopath with Ecosim models. Environmental Modelling and
868 Software 89, 120-130.
869 Coll, M., Steenbeek, J., Ben Rais Lasram, F., Mouillot, D., Cury, P., 2015. “Low hanging fruits” for
870 conservation of marine vertebrate species at risk in the Mediterranean Sea. Global Ecology and
871 Biogeography 24, 226-239.

872 Coll, M., Steenbeek, J., Sole, J., Palomera , I., Christensen, V., 2016. Modelling the cumulative
873 spatial-temporal effects of environmental factors and fishing in a NW Mediterranean marine
874 ecosystem. Ecological Modelling 331, 100-114.
875 Colloca, F., Cardinale, M., Maynou, F., Giannoulaki, M., Scarcella, G., Jenko, K., Bellido, J.M.,
876 Fiorentino, F., 2013. Rebuilding Mediterranean fisheries: a new paradigm for ecological
877 sustainability. Fish and Fisheries 14, 89-109.
878 Cury, P.M., Shin, Y.J., Planque, B., Durant, J.M., Fromentin, J.M., Kramer-Schadt, S., Stenseth, N.C.,
879 Travers, M., Grimm, V., 2008. Ecosystem oceanography for global change in fisheries. Trends in
880 Ecology & Evolution 23, 338-346.
881 Estrada, M., 1996. Primary production in the northwestern Mediterranean. Scientia Marina 60, 55-
882 64.
883 F Dormann, C., M McPherson, J., B Araújo, M., Bivand, R., Bolliger, J., Carl, G., G Davies, R.,
884 Hirzel, A., Jetz, W., Daniel Kissling, W., 2007. Methods to account for spatial autocorrelation in
885 the analysis of species distributional data: a review. Ecography 30, 609-628.
886 Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J., Haidvogel, D., 2006. Nitrogen cycling in the
887 Middle Atlantic Bight: Results from a three-dimensional model and implications for the North
888 Atlantic nitrogen budget. Global Biogeochemical Cicles 20, GB3007.
889 Fernandes, P.G., Ralph, G.M., Nieto, A., Criado, M.G., Vasilakopoulos, P., Maravelias, C.D., Cook,
890 R.M., Pollom, R.A., Kovačić, M., Pollard, D., 2017. Coherent assessments of Europe's marine
891 fishes show regional divergence and megafauna loss. Nature Ecology & Evolution 1, 0170.
892 Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in
893 conservation presence/absence models. Environmental Conservation 24, 38-49.
894 Fonseca, V.P., Pennino, M.G., de Nóbrega, M.F., Oliveira, J.E.L., de Figueiredo Mendes, L., 2017.
895 Identifying fish diversity hot-spots in data-poor situations. Marine Environmental Research 129,
896 365-373.
897 Fulton, E., 2011. Interesting times: winners, losers, and system shifts under climate change around
898 Australia. ICES Journal of Marine Science 68, 1329–1342.
899 Fulton, E.A., 2010. Approaches to end-to-end ecosystem models. Journal of Marine Systems 81, 171-
900 183.
901 Gelfand, A.E., Silander, J.A., Wu, S., Latimer, A., Lewis, P.O., Rebelo, A.G., Holder, M., 2006.
902 Explaining species distribution patterns through hierarchical modeling. Bayesian Analysis 1, 41-
903 92.
904 Gelman, A., 2008. Scaling regression inputs by dividing by two standard deviations. Stat. Med. 27,
905 2865–2873.
906 Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological
907 Modelling 135, 147-186.
908 Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S.,
909 Rockwood, R.C., Selig, E.R., Selkoe, K.A., Walbridge, S., 2015. Spatial and temporal changes in
910 cumulative human impacts on the world's ocean. Nature Communications 7615.
911 Held, L., Schrödle, B., Rue, H., 2010. Posterior and cross-validatory predictive checks: a comparison
912 of MCMC and INLA. Statistical modelling and regression structures, 91-110.
913 Hereford, J., Hansen, T.F., Houle, D., 2004. Comparing strengths of directional selection: how strong
914 is strong? Evolution 58, 2133–2143.
915 Heymans, J.J., Coll, M., Link, J.S., Mackinson, S., Steenbeek, J., Christensen, V., 2016. Best practice
916 in Ecopath with Ecosim food-web models for ecosystem-based management. Ecological
917 Modelling 331, 173–184.
918 Hidalgo, M., Massutí, E., Moranta, J., Cartes, J., Lloret, J., Oliver, P., Morales-Nin, B., 2008. Seasonal
919 and short spatial patterns in European hake (Merluccius merluccius L.) recruitment process at the
920 Balearic Islands (western Mediterranean): the role of environment on distribution and condition.
921 Journal of Marine Systems 71, 367-384.

Jones, M.C., Dye, S.R., Pinnegar, J.K., Warren, R., Cheung, W.W., 2012. Modelling commercial fish distributions: Prediction and assessment using different approaches. Ecological Modelling 225, 133-145.
Kaschner, K., Kesner-Reyes, K., Garilao, C., Rius-Barile, J., Rees, T., Froese, R., 2016. AquaMaps: Predicted range maps for aquatic species. World wide web electronic publication, www.aquamaps.org, Version 08/2016.
Kaschner, K., Watson, R., Trites, A., Pauly, D., 2006. Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. Marine Ecology Progress Series 316.
Kinas, P.G., Andrade, H.A., 2017. Introdução à análise bayesiana (com R). Consultor Editorial.
Lauria, V., Gristina, M., Attrill, M., Fiorentino, F., Garofalo, G., 2015. Predictive habitat suitability models to aid conservation of elasmobranch diversity in the central Mediterranean Sea. scientific Reports 5, 13245.
Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74.
Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73, 423-498.
Link, J., 2011. Ecosystem-based fisheries Management: confronting tradeoffs. Cambridge University Press, Cambridge 207 pp.
Lloret, J., Lleonart, J., 2002. Recruitment dynamics of eight fishery species in the northwestern Mediterranean Sea. Scientia Marina 66, 77-82.
Lloret, J., Palomera, I., Salat, J., Sole, I., 2004. Impact of freshwater input and wind on landings of anchovy (Engraulis encrasicolus) and sardine (Sardina pilchardus) in shelf waters surrounding the Ebre (Ebro) River delta (north-western Mediterranean). Fisheries Oceanography 13, 102-110.
Lombarte, A., Recasens, L., González, M., de Sola, L.G., 2000. Spatial segregation of two species of Mullidae (Mullus surmuletus and M. barbatus) in relation to habitat. Marine Ecology Progress Series 206, 239-249.
López, N., Navarro, J., Barría, C., Albo-Puigserver, M., Coll, M., Palomera, I., 2016. Feeding ecology of two demersal opportunistic predators coexisting in the northwestern Mediterranean Sea. Estuarine, Coastal and Shelf Science 175, 15-23.
Macias, D., Catalan, I.A., Solé, J., Morales Nin, B., Ruiz, J., 2011. Atmospheric-induced variability of hydrological and biogeochemical signatures in the NW Alboran Sea. Consequences for the spawning and nursery habitats of European anchovy. Deep Sea Research I 58, 1175-1188.
Mackinson, S., Daskalov, G., Heymans, J.J., Neira, S., Arancibia, H., Zetina-Rejón, M., Jiang, H., Cheng, H.Q., Coll, M., Arreguin-Sanchez, F., 2009. Which forcing factors fit? Using ecosystem models to investigate the relative influence of fishing and changes in primary productivity on the dynamics of marine ecosystems. Ecological Modelling 220, 2972-2987.
Maravelias, C.D., Tsitsika, E.V., Papaconstantinou, C., 2007. Environmental influences on the spatial distribution of European hake (Merluccius merluccius) and red mullet (Mullus barbatus) in the Mediterranean. Ecol. Res. 22, 678-685.
Martin, P., Sabates, A., Lloret, J., Martin-Vide, J., 2012. Climate modulation of fish populations: the role of the Western Mediterranean Oscillation (WeMO) in sardine (Sardina pilchardus) and anchovy (Engraulis encrasicolus) production in the north-western Mediterranean. Climatic Change 110, 925-939.
Massutí, E., Monserrat, S., Oliver, P., Moranta, J., López-Jurado, J.L., Marcos, M., Hidalgo, M., Guijarro, B., Carbonell, A., Pereda, P., 2008. The influence of oceanographic scenarios on the population dynamics of demersal resources in the western Mediterranean: hypothesis for hake and red shrimp off Balearic Islands. Journal of Marine Systems 71, 421-438.
Maynou, F., Lleonart, J., Cartes, J.E., 2003. Seasonal and spatial variability of hake (Merluccius merluccius L.) recruitment in the NW Mediterranean. Fisheries Research 60, 65-78.

Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E., Elith, J., 2014. What do we gain from simplicity versus complexity in species distribution models? Ecography 37, 1267-1281.

Moore, C., Drazen, J.C., Radford, B.T., Kelley, C., Newman, S.J., 2016. Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. Marine Policy 69, 32-41.

Morfin, M., Fromentin, J.M., Jadaud, A., Bez, N., 2012. Spatio-Temporal Patterns of Key Exploited Marine Species in the Northwestern Mediterranean Sea. PLoS ONE 7, e37907.

Munoz, F., Pennino, M.G., Conesa, D., López-Quílez, A., Bellido, J.M., 2013. Estimation and prediction of the spatial occurrence of fish species using Bayesian latent Gaussian models. Stochastic Environmental Research and Risk Assessment 27, 1171-1180.

Navarro, J., Coll, M., Cardador, L., Fernández, A.M., Bellido, J.M., 2015. The relative roles of the environment, human activities and spatial factors in the spatial distribution of marine biodiversity in the Western Mediterranean Sea. Progress in Oceanography 131, 126-137.

Navarro, J., Coll, M., Louzao, M., Palomera, I., Delgado, A., Forero, M.G., 2011. Comparison of ecosystem modelling and isotopic approach as ecological tools to investigate food webs in the NW Mediterranean Sea. Journal of Experimental Marine Biology and Ecology 401, 97-104.

Olivar, M.P., Quílez, G., Emelianov, M., 2003. Spatial and temporal distribution and abundance of European hake, *Merluccius merluccius*, eggs and larvae in the Catalan coast (NW Mediterranean). Fisheries Research 60, 321-331.

Ospina-Alvarez, A., Catalan, I., Bernal, M., Roos, D., Palomera , I., 2015. From egg production to recruits: Connectivity and inter-annual variability in the recruitment patterns of European anchovy in the northwestern Mediterranean. Progress in Oceanography http://dx.doi.org/10.1016/j.pocean.2015.01.011.

Palomera, I., Olivar, M.P., Salat, J., Sabates, A., Coll, M., Garcia, A., Morales-Nin, B., 2007. Small pelagic fish in the NW Mediterranean Sea: An ecological review. Progress in Oceanography 74, 377-396.

Paradinas, I., Conesa, D., Pennino, M.G., Muñoz, F., Fernández, A.M., López-Quílez, A., Bellido, J.M., 2015. Bayesian spatio-temporal approach to identifying fish nurseries by validating persistence areas. Marine Ecology Progress Series 528, 245–255.

Paradinas, I., Marín, M., Pennino, M.G., López-Quílez, A., Conesa, D., Barreda, D., Gonzalez, M., Bellido, J.M., 2016. Identifying the best fishing-suitable areas under the new European discard ban. ICES Journal of Marine Science: Journal du Conseil, fsw114.

Pennino, M.G., Mérigot, B., Fonseca, V.P., Monni, V., Rotta, A., 2016. Habitat modeling for cetacean management: Spatial distribution in the southern Pelagos Sanctuary (Mediterranean Sea). Deep Sea Research Part II: Topical Studies in Oceanography.

Pennino, M.G., Muñoz, F., Conesa, D., López-Quílez, A., Bellido, J.M., 2014. Bayesian spatio-temporal discard model in a demersal trawl fishery. Journal of Sea Research 90, 44-53.

Pennino, M.G., Munoz, F., Conesa, D., López-Quílez, A., Bellido, J.M., 2013. Modeling sensitive elasmobranch habitats. Journal of Sea Research 83, 209-218.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A., McCarthy, M.A., 2014. Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). Methods in Ecology and Evolution 5, 397-406.

Polovina, J.J., 1984. Model of a Coral-Reef Ecosystem .1. the Ecopath Model and Its Application to French Frigate Shoals. Coral Reefs 3, 1-11.

Recasens, L., Lombarte, A., Morales-Nin, B., Torres, G., 1998. Spatio-temporal variation in the population structure of the European hake in the NW Mediterranean. Journal of Fish Biology 53, 387–401.

Romagnoni, G., Mackinson, S., Hong, J., Eikeset, A.M., 2015. The Ecospace model applied to the North Sea: Evaluating spatial predictions with fish biomass and fishing effort data. Ecological

1022 Modelling 300, 50-60.
1023 Roos, N.C., Carvalho, A.R., Lopes, P.F., Pennino, M.G., 2015. Modeling sensitive parrotfish
1024 (Labridae: Scarini) habitats along the Brazilian coast. Marine Environmental Research 110, 92-
1025 100.
1026 Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models
1027 by using integrated nested Laplace approximations. Journal of the royal statistical society: Series
1028 b (statistical methodology) 71, 319-392.
1029 Saraux, C., Fromentin, J.-M., Bigot, J.-L., Bourdeix, J.-H., Morfin, M., Roos, D., Van Beveren, E.,
1030 Bez, N., 2014. Spatial Structure and Distribution of Small Pelagic Fish in the Northwestern
1031 Mediterranean Sea. PLoS ONE 9, e111211.
1032 Scott, E., Serpetti, N., Steenbeek, J., Heymans, J.J., 2016. A Stepwise Fitting Procedure for automated
1033 fitting of Ecopath with Ecosim models. SoftwareX.
1034 Shchepetkin, A., McWilliams, J., 2005. The Regional Ocean Modeling System (ROMS): A split-
1035 explicit, free-surface, topography-following coordinates ocean model. Ocean Modelling 9, 347-
1036 404.
1037 Solé, J., Balabrera-Poy, J., Macías, D., Catalán, I.A., 2016. The role of ocean velocity in clorophyll
1038 variability. A modelling study in Alboran Sea. Scientia Marina 80S1, 249-256.
1039 Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model
1040 complexity and fit. J. R. Stat. Soc. Ser. B Stat. Methodol. 64, 583–616.
1041 Steenbeek, J., Coll, M., Gurney, L., Melin, F., Hoepffner, N., Buszowski, J., Christensen, V., 2013.
1042 Bridging the gap between ecosystem modeling tools and geographic information systems: Driving
1043 a food web model with external spatial–temporal data. Ecological Modelling 263, 139-151.
1044 Steenbeek, J., Corrales, X., Platts, M., Coll, M., 2018. Ecosampler: a new approach to assessing
1045 parameter uncertainty in Ecopath with Ecosim. SoftwareX 7, 198-204.
1046 Stobart, B., Warwick, R., González, C., Mallol, S., Díaz, D., Reñones, O., Goñi, R., 2009. Long-term
1047 and spillover effects of a marine protected area on an exploited fish community. Marine Ecolog
1048 Progress Series 384, 47−60.
1049 Tserpes, G., Fiorentino, F., Levi, D., Cau, A., Murenu, M., Zamboni, A., Papaconstantinou, C., 2002.
1050 Distribution of Mullus barbatus and M. surmuletus (Osteichthyes: Perciformes) in the
1051 Mediterranean continental shelf: implications for management. Scientia Marina 66, 39-54.
1052 Tsikliras, A.C., Dinouli, A., Tsiros, V.-Z., Tsalkou, E., 2015. The Mediterranean and Black Sea
1053 fisheries at risk from overexploitation. PLoS ONE 10, e0121188.
1054 Urban, M., Bocedi, G., Hendry, A., Mihoub, J.-B., Pe'er, G., Singer, A., Bridle, J., Crozier, L., De
1055 Meester, L., Godsoe, W., 2016. Improving the forecast for biodiversity under climate change.
1056 Science 353, aad8466.
1057 Vilás-González, D., 2016. Variabilidad espacial y estacional de la comunidad de peces, cefalópodos
1058 y crustáceos del Mar Mediterráneo Noroccidental y relación con los factores ambientales y
1059 humanos. Barcelona University.
1060 Vilas-González, D., Pennino, M.G., Bellido, J.M., Navarro, J., Palomera , I., Coll, M., Submitted.
1061 Seasonality of spatial patterns of abundance, biomass and biodiversity in a demersal community
1062 from the NW Mediterranean Sea. Progress in Oceanography.
1063 Walters, C., 2000. Impacts of dispersal, ecological interactions, and fishing effort dynamics on
1064 efficacy of marine protected areas: how large should protected areas be? Bulletin of Marine
1065 Science 66, 745-757.
1066 Walters, C., Christensen, V., Pauly, D., 1997. Structuring dynamic models of exploited ecosystems
1067 from trophic mass-balance assessments. Reviews in Fish Biology and Fisheries 7, 139-172.
1068 Walters, C., Christensen, V., Walters, W., Rose, K., 2010. Representation of multistanza life histories
1069 in Ecospace models for spatial organization of ecosystem trophic interaction patterns. Bulletin of
1070 Marine Science 86, 439-459.
1071 Walters, C., Pauly, D., Christensen, V., 1999. Ecospace: prediction of mesoscale spatial patterns in

trophic relationships of exploited ecosystems, with emphasis on the impacts of marine protected areas. Ecosystems 2, 539-554.
Walters, C., Pauly, D., Christensen, V., Kitchell, J.F., 2000. Representing Density Dependent Consequences of Life History Strategies in Aquatic Ecosystems: EcoSim II. Ecosystems 3, 70-83.
Walters, C.J., Martell, S.J.D., 2004. Foraging arena theory (II), in: Walters, C.J., Martell, S.J.D. (eds.), Fisheries Ecology and Management. Princeton University Press Princeton, NJ (USA), pp. 229-255.
Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research 11, 3571-3594.
Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution 1, 3-14.