Spatial standardization of Catch Per Unit Effort (CPUE) indices for the megrim (*Lepidorhombus whiffiagonis*) and the four-spot megrim (*L. boscii*) in North Atlantic Iberian waters (ICES divisions 8c and 9a).

Maria Grazia Pennino[1], Esther Abad[1] & Jaime Otero[1].

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

Background

Fishery-dependent data collected from fishery observers' on-board commercial vessels or logbooks can be used to construct standardized indices of relative biomass for stock assessment models.
Several standardization techniques have been used for fishery-dependent data of many species, and most of them highlight the inclusion of environmental variables and spatio-temporal effects (Thorson et al., 2016). Overall these methods have been proved to be a useful tool to address ecological and stock assessment issues.
Within this context here we present a spatial standardization of Catch Per Unit Effort (CPUE) indices for the megrim (*Lepidorhombus whiffiagonis*) and the four-spot megrim (*L. boscii*) using a time series (2003-2020) of observers on board data of the bottom otter trawl fleet that operate in the in North Atlantic Iberian waters (ICES divisions 8c and 9a).

Study Region

The region of interest for this study is the northern continental shelf of the Iberian Peninsula, a narrow area (10–60 km) of almost 18,000 km^2 that is characterized by important and marked hydro dynamism (Abad et al., 2019). Over the shelf, currents are driven by regional factors, such as tides and wind. In the winter, a warm and saline poleward current moves eastward along the Cantabrian coast and enters the Bay of Biscay (Izquierdo et al., 2021). In addition, the coastal upwelling off the Galician and Portuguese coasts appears during spring and summer which, combined with hydrographic mesoscale activities, has a strong influence on the primary production of the area (Sánchez and Olaso, 2004).

Input data

A dataset was compiled from the observers on board programme of the Northern Spanish coastal bottom otter trawl fleet developed by the Instituto Español de Oceanografía (IEO). These reference fleet includes trawlers that usually operate in waters from the continental shelf (from 100 to 350 m depth) with European hake (*Merluccius merluccius*) and anglerfish (*Lophius budegassa, L. piscatorius*), megrims (*Lepidorhombus boscii* and *L. whiffiagonis*), horse mackerel (*Trachurus sp.*), blue whiting (*Micromesistius poutassou*) and Norway lobster (*Nephrops norvegicus*) as target species. They make hauls of about 2–4 h, comprising about 4-5 fishing hauls per trip. At each haul, observers record all basic operational data (i.e., date and geographical position and time duration of the fishing haul), the number and weight of all retained and discarded taxa and environmental variables as the bathymetry of the seabed. Additionally, information about the fishing vessel are recorded as its total length (in meters). The time period considered in the present study extends from 2003 to 2020.

Methods

For the four-spot megrim the relationship between the catches and predictors was modelled using a Gamma distribution as no zeros were recorded. On the contrary for the megrim, a hurdle model was implemented as the percentage of zeros was higher than 30%. For this reason, two different response variables were analysed for the megrim: (1) a presence/absence variable to measure the occurrence probability of the species; (2) positive catches (in kg) as an indicator of the conditional-to-presence biomass of the species. The occurrence was modelled using a Binomial distribution with a logit link function and the catches with a Gamma distribution with a log link, to capture the overdispersion of the data.
In each model the response variable was modelled as a function of explanatory variables assumed to influence catchability including: fishing haul duration (in hours), total vessel length (in m), depth (in m) of the fishing haul, and two variables that assess the interannual (years, 2003-2010) and seasonal (quarter: 1,2,3,4) variability.
Prior to the analysis, the explanatory variables were standardized (i.e., difference from the mean divided by the corresponding standard deviation) (Gelman et al., 2014) to better interpret both the direction (positive or negative) and magnitude (effect sizes) of the parameter estimates.
As it is known that gear saturation can exert a significant nonlinear effect on catchability exploratory analysis were performed to verify the linear relationship between the response variables and the continuous predictors. For this reason, second order random walk (RW2) functions were applied to the haul duration and the bathymetry.
We further accounted for spatial autocorrelation by including a numeric vector with a mean of 0 and a Matern covariance function linking each observation to a spatial location (i.e., latitude and longitude). Thus, our model accounts for independent and region-specific noise not explained by the available covariates. For the parameters involved in the fixed effects, vague Gaussian priors with a mean of 0 and a variance of 100 were used. The random spatial effect only depends on two hyperparameters: the range and the variance of the spatial effect. Penalized complexity priors (Fuglstad et al., 2018) were used to describe prior knowledge on these hyperparameters. We set a prior range of 50 km with a probability of 0.001 for it to be lower and a prior variance of 2 with a probability of 0.001 for it to be higher. We performed a sensitivity analysis of the choice of priors for the spatial effect by testing different priors and verifying that the posterior distributions were consistent and concentrated well within the support of the priors.

Bayesian inference was performed using the Integrated Nested Laplace Approximations (INLA) approach (Rue et al., 2009) with its corresponding package. INLA uses the so-called Stochastic Partial Differential Equation approach to approximate the Gaussian field with the Matern covariance function by a Gaussian Markov random field (Rue et al., 2009).

We selected the most parsimonious model, starting with all, based on the goodness-of-fit using the deviance information criterion (DIC) (Spiegelhalter et al., 2002) and Watanabe–Akaike information criterion (WAIC) (Watanabe, 2010).

The final model was evaluated with the log-conditional predictive ordinate (log-CPO) (Roos and Held, 2011), which is a “leave-one-out” cross-validation index to assess the predictive power of the model (Pennino et al., 2019).

INLA has built-in functions allowing for a linear interpolation of the spatial effect within each triangle into a finer regular grid. The resulting high-resolution map of the spatial effect can be seen as a proxy for the species’ relative biomass.

All analyses and graphics were performed in R (R Core Team, 2020).

Results

Four-spot megrim (*L. boscii*)

Overall 3092 fishing hauls were analysed in which the *L. boscii* was caught, with values ranging from 1 and 370 kg by haul. The seasonal distribution of the catches seems overall homogeneous (Figure 1). However, it worth to be mentioned that in the 2020, due to the COVID pandemic the observers’ onboard program was performed only in the two last quarters of the year.

Overall catches showed a decreasing temporal trend along the time series 2003-2020, (Figure 1), recording the highest catches in the 2015, while the lowest value was recorded in the 2020 one (Figure 1).

The final Bayesian model, selected on the basis of the lowest DIC, WAIC and LCPO values, retained as predictors, the fishing haul time duration, depth, vessel total length, year, quarter and the spatial effect (Table 1). Finally, the depth variable required a smoothing spline, showing a negative relationship with the catches of the *L. boscii* (Figure 2).

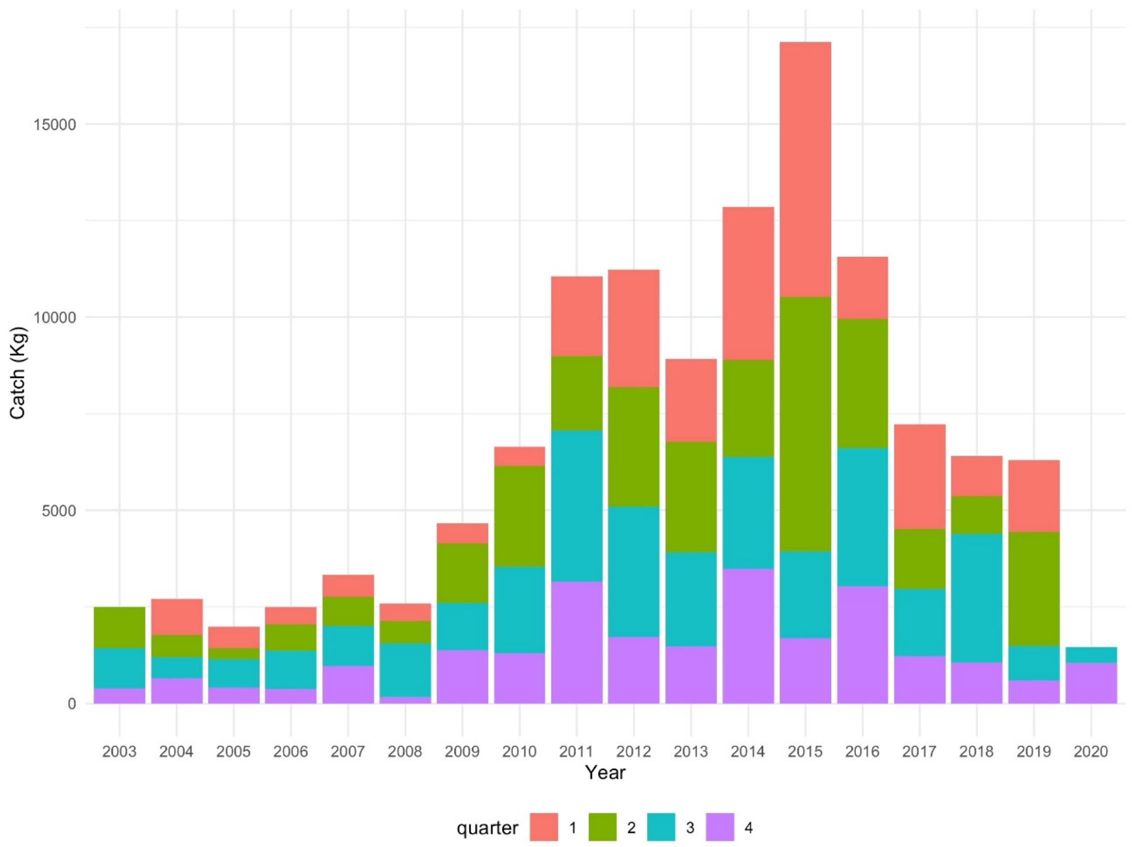

Figure 1: Temporal distribution (by quarter and year) of the megrim (*Lepidorhombus boscii*) catches (in Kg).

Fourth main hot-spots were identified in the area studied (Figure 3). Specifically, from south to north, the first was located off the Ria of Pontevedra and Vigo, the second one off the Costa de la Muerte, the third and largest of these covered most of the Artabrian gulf off La Coruña, and the fourth (a small one) was located off Gijon.

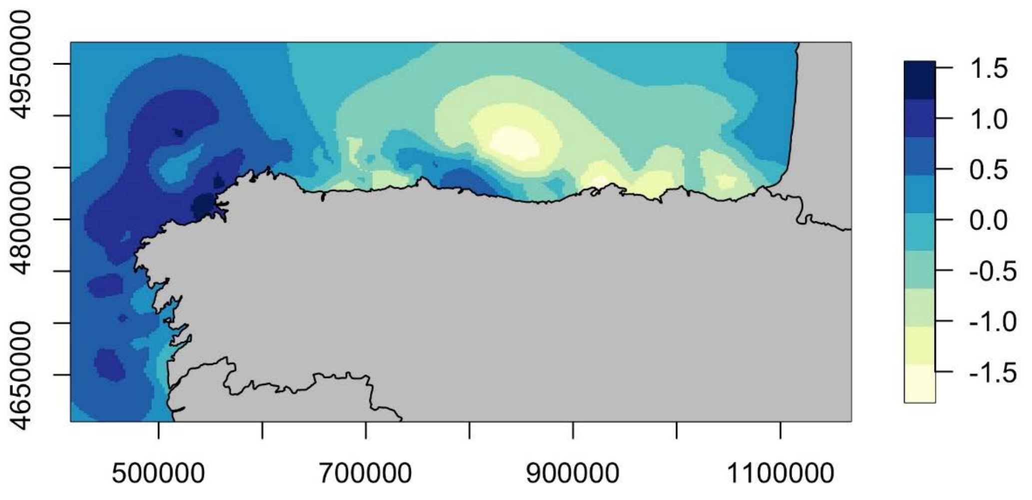

Figure 3: Posterior mean of the spatial effect for the four-spot megrim (*L. boscii*) model.

The fishing haul duration showed a positive relationship with the four-spot megrim catches, meaning that higher values of catches were in longer hauls. On the contrary, the vessel length showed a negative relationship with the catches (i.e., longer boats catch less quantity). Regarding the seasonal trend, there was a bit of heterogeneity between the different four-month periods, the first being the one with the highest catches (Figure 4). The annual trend showed a decreasing pattern in the last years of the time series, as also showed for the final standardized CPUE index (Figure 5).

Table 1. Model comparison for the occurrence and conditional-to-presence catch of the four-spot megrim (*L.boscii)*. Acronyms are deviance information criterion (DIC), Watanabe–Akaike information criterion (WAIC) and log-conditional predictive ordinate (log-CPO), Y = year, Q = quarter of the year, D = depth of the fishing haul, HD = haul duration, VL = vessel length, S = spatial effect, f = smoothing function. The final selected mode is highlighted in bold.

Model	DIC	WAIC	LCPO
1 + Y + Q + D + HD + VL	13755.97	13760.06	4.62
1 + Y + Q + D + HD + VL + S	13389.88	13428.34	4.45
1 + Y + Q + f(D) + HD + VL + S	13345.53	13378.56	4.41
1 + Y + Q +F(D) + f(HD) + VL + S	13348.21	13379.5	4.41

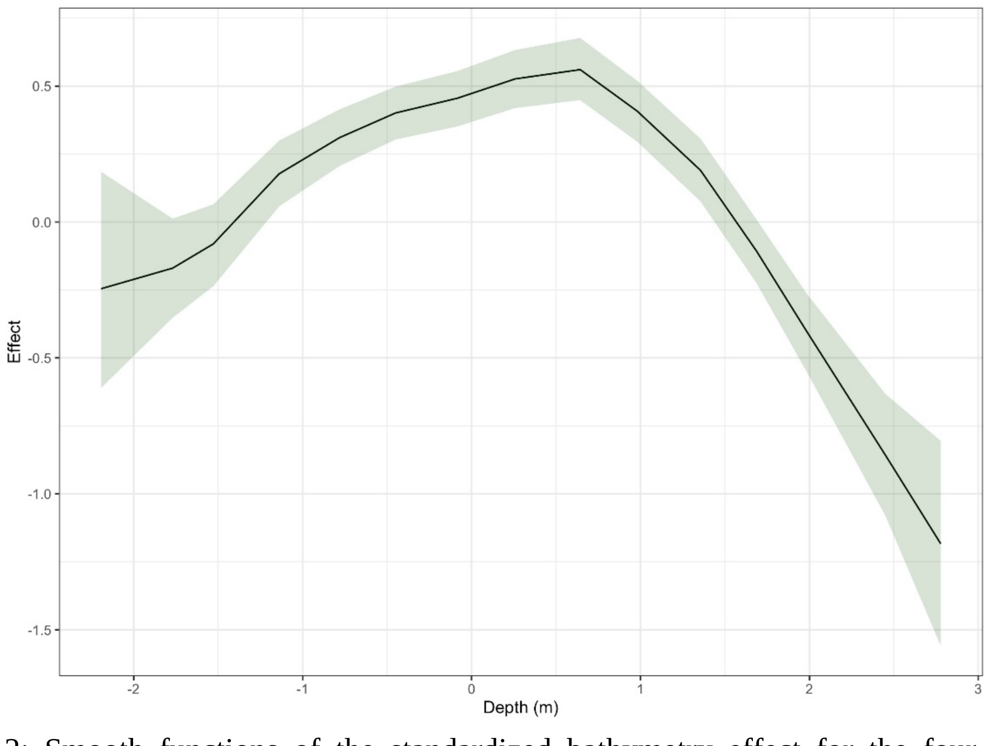

Figure 2: Smooth functions of the standardized bathymetry effect for the four-spot megrim (*L.boscii*) model. The solid line is the smooth function estimate, and shaded regions represent the approximate 95% credibility interval.

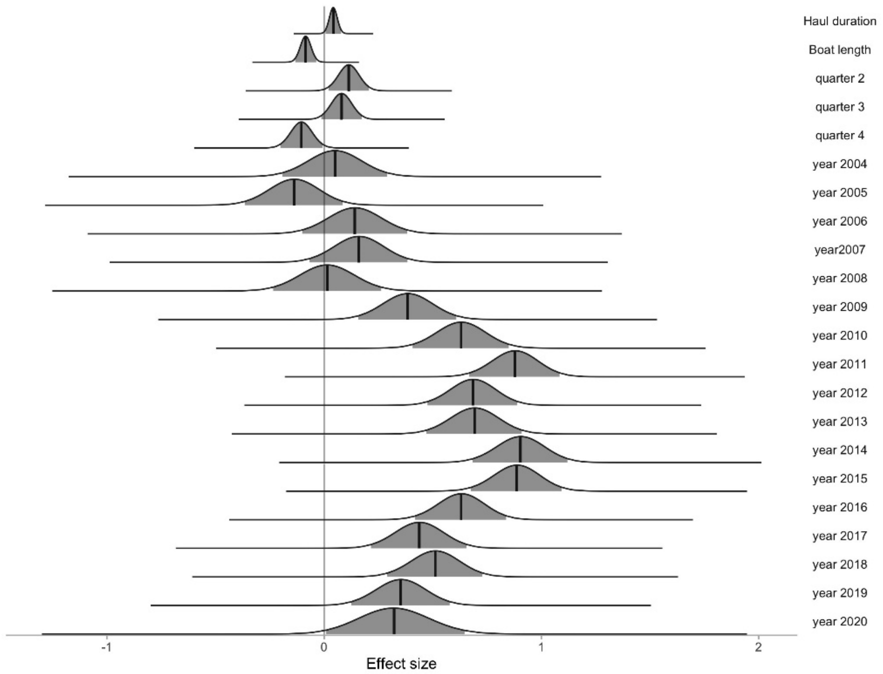

Figure 4: Posterior marginal distribution of the fixed effects of the catches four-spot megrim (*L. boscii)* model.

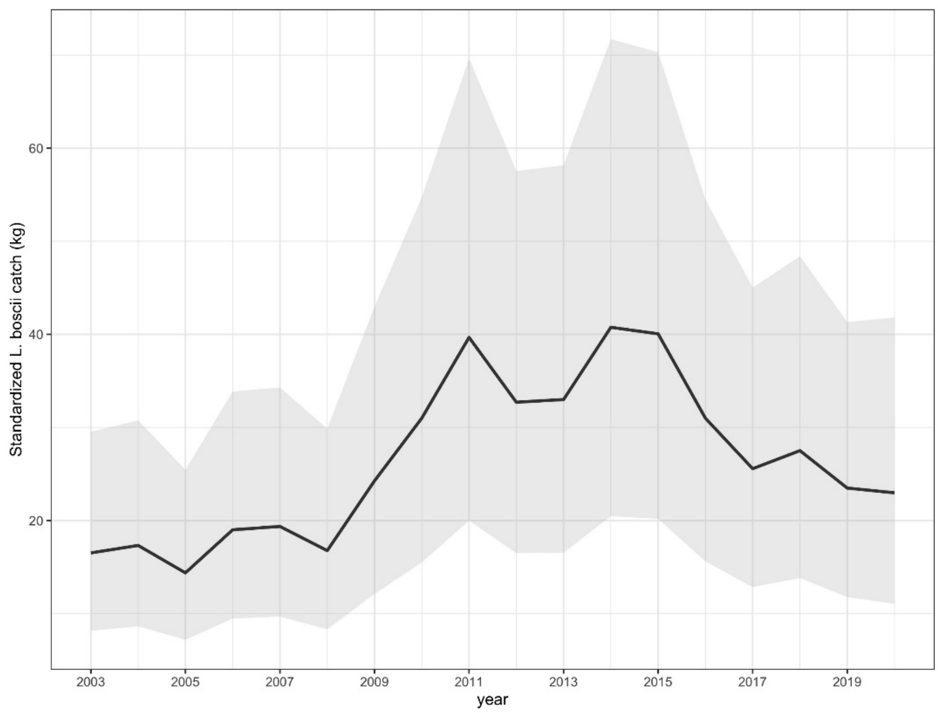

Figure 5: Standardized CPUE combined index for the four-spot megrim (*L. boscii*).

Megrim *(Lepidorhombus whiffiagonis*)

Among 3093 fishing hauls recorded, 1852 have positive catches ranging from 1 to 170 kg by haul. The seasonal distribution of the catches seems overall homogeneous (Figure 6).
No clear temporal trend can be observed in the catches along the time series 2003-2020, although in the last decade catches were higher than in the first one (Figure 6). The year with the highest catches was the 2011 while the lowest value was recorded in the 2008 one (Figure 6).

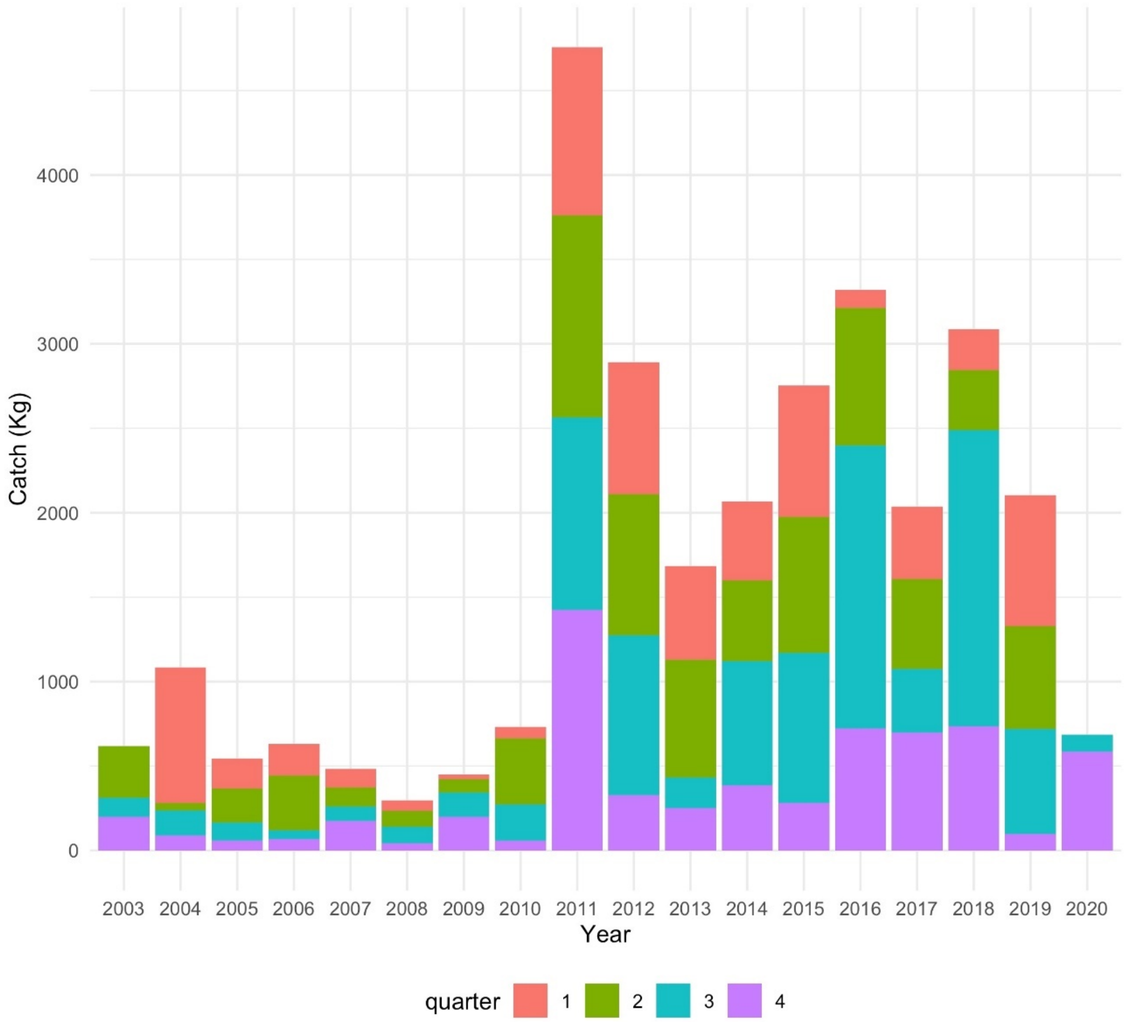

Figure 6: Temporal distribution (by quarter and year) of the megrim (*Lepidorhombus whiffiagonis*) catches (in Kg).

In the final Bayesian hurdle model for both occurrence and conditional-to-presence-catch retained as predictors, the fishing haul time duration, depth, vessel total length, year, quarter and the spatial effect (Table 2). In both models only, the depth variable required smoothing splines, showing a negative relationship with the probability of occurrence and the catches of the *L. whiffiagonis* (Figure 7).

Table 2. Model comparison for the occurrence and conditional-to-presence catch of the megrim (*L. whiffiagonis)*. Acronyms are deviance information criterion (DIC), Watanabe–Akaike information criterion (WAIC) and log-conditional predictive ordinate (log-CPO), Y = year, Q = quarter of the year, D = depth of the fishing haul, HD = haul duration, VL = vessel length, S = spatial effect, f = smoothing function. The final selected mode is highlighted in bold.

Model	DIC	WAIC	LCPO
Occurrence			
1 + Y + Q + D + HD + VL	3811.705	3812.281	0.62
1 + Y + Q + D + HD + VL + S	2902.212	2901.01	0.47
1 + Y + Q + f(D) + HD + VL + S	2876.866	2875.622	0.47
1 + Y + Q +F(D) + f(HD) + VL + S	2877.437	2876.036	0.47
Conditional-to-presence-catch			
1 + Y + Q + D + HD + VL	13755.97	13760.06	3.71
1 + Y + Q + D + HD + VL + S	13389.88	13428.34	3.67
1 + Y + Q + f(D) + HD + VL + S	13345.53	13378.56	3.62
1 + Y + Q +F(D) + f(HD) + VL + S	13348.21	13379.50	3.62

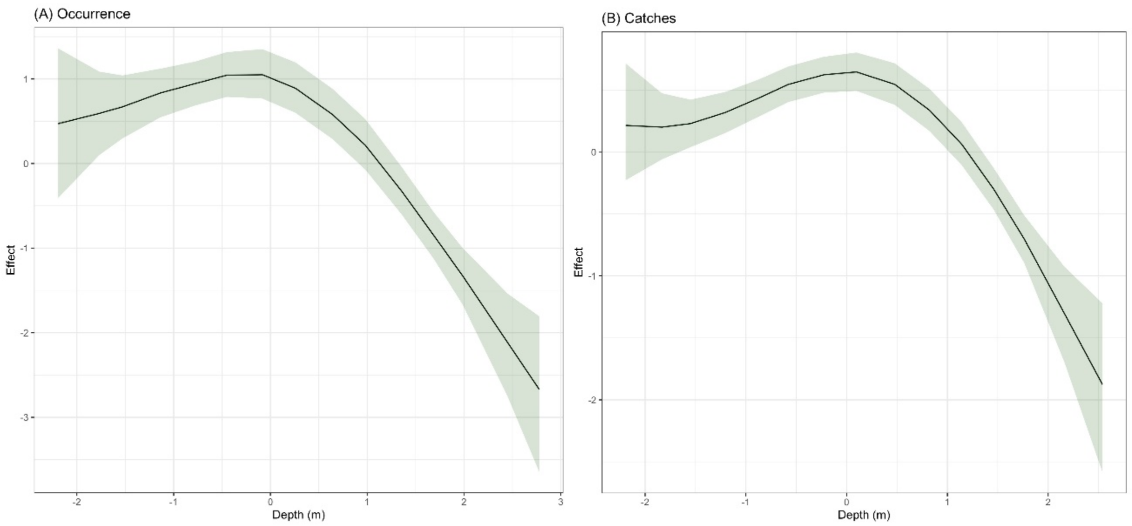

Figure 7: Smooth functions of the standardized bathymetry effects for the occurrence (A) and conditional-to-presence-catches (B) of the megrim (*L. whiffiagonis*) models. The solid line is the smooth function estimate, and shaded regions represent the approximate 95% credibility interval.

The spatial effect pattern was similar in both occurrence and conditional-to-presence-catches models (Figure 8), highlighting three main hotspots, the largest one covered most of the Artabrian gulf off La Coruña, the second one was located off Santander and the last one in more deeper waters in the Bay of Biscay.

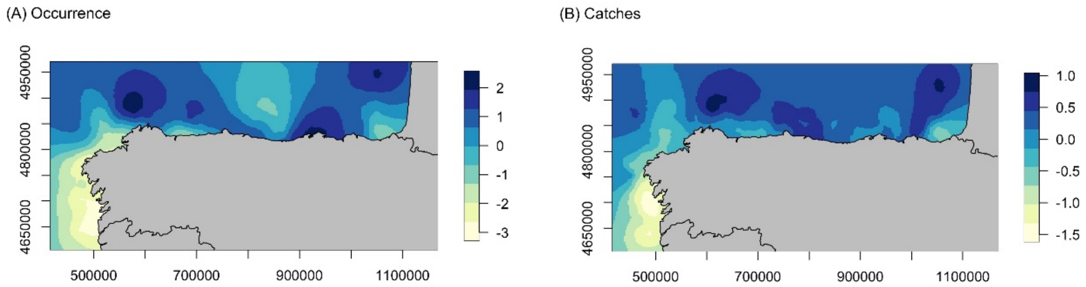

Figure 8: Posterior mean of the spatial effect for the (A) occurrence and (B) conditional-to-presence-catches for the megrim (*L. whiffiagonis*) models.

The fishing haul duration showed a positive relationship, i.e., higher values catches are recorded when the duration of the haul is longer. On the contrary the vessel length presented a negative relationship with the occurrence and the conditional-to-presence catches meaning the larger vessel caught lower quantity of megrim. However, looking at the magnitude of the estimated parameters for these two variables we can see that are not the most relevant in explaining the occurrence and catches variability (Figure 9). With respect the seasonal pattern, for the occurrence the quarter with higher probability of presence is the fourth of the year, while for the catches is the first one (Figure 9). For both models if we analysed the estimated parameters of the years we can see that overall there is an increasing trend in both occurrence and conditional-to presence-catches variables. This increasing trend is also reflected in the final standardized combined CPUE index (i.e., probability of occurrence * conditional-to presence-catches) (Figure 10).

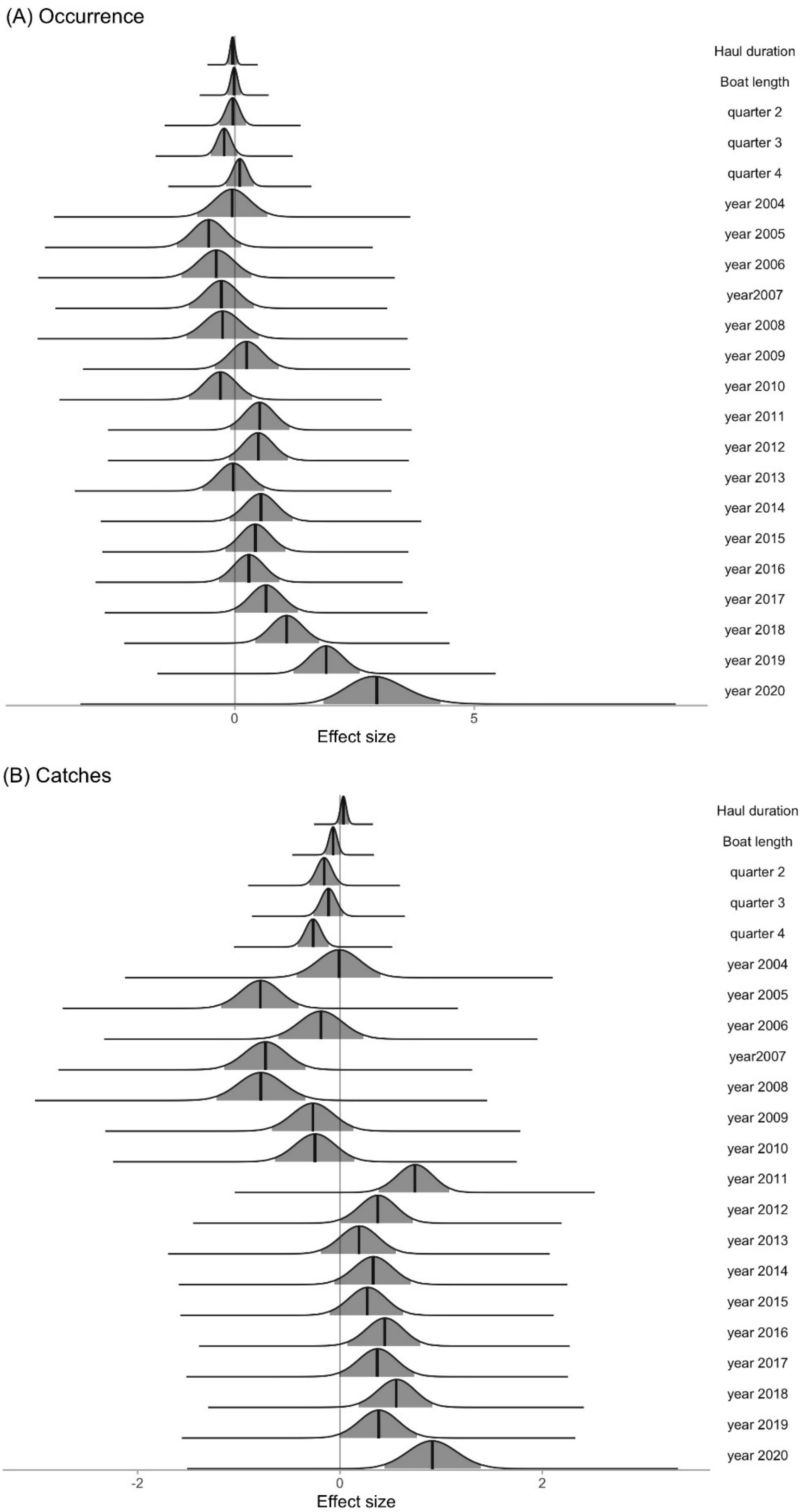

Figure 9: Posterior marginal distribution of the fixed effects of the occurrence (A) and conditional-to-presence-catches (B) megrim (*L. whiffiagonis)* models.

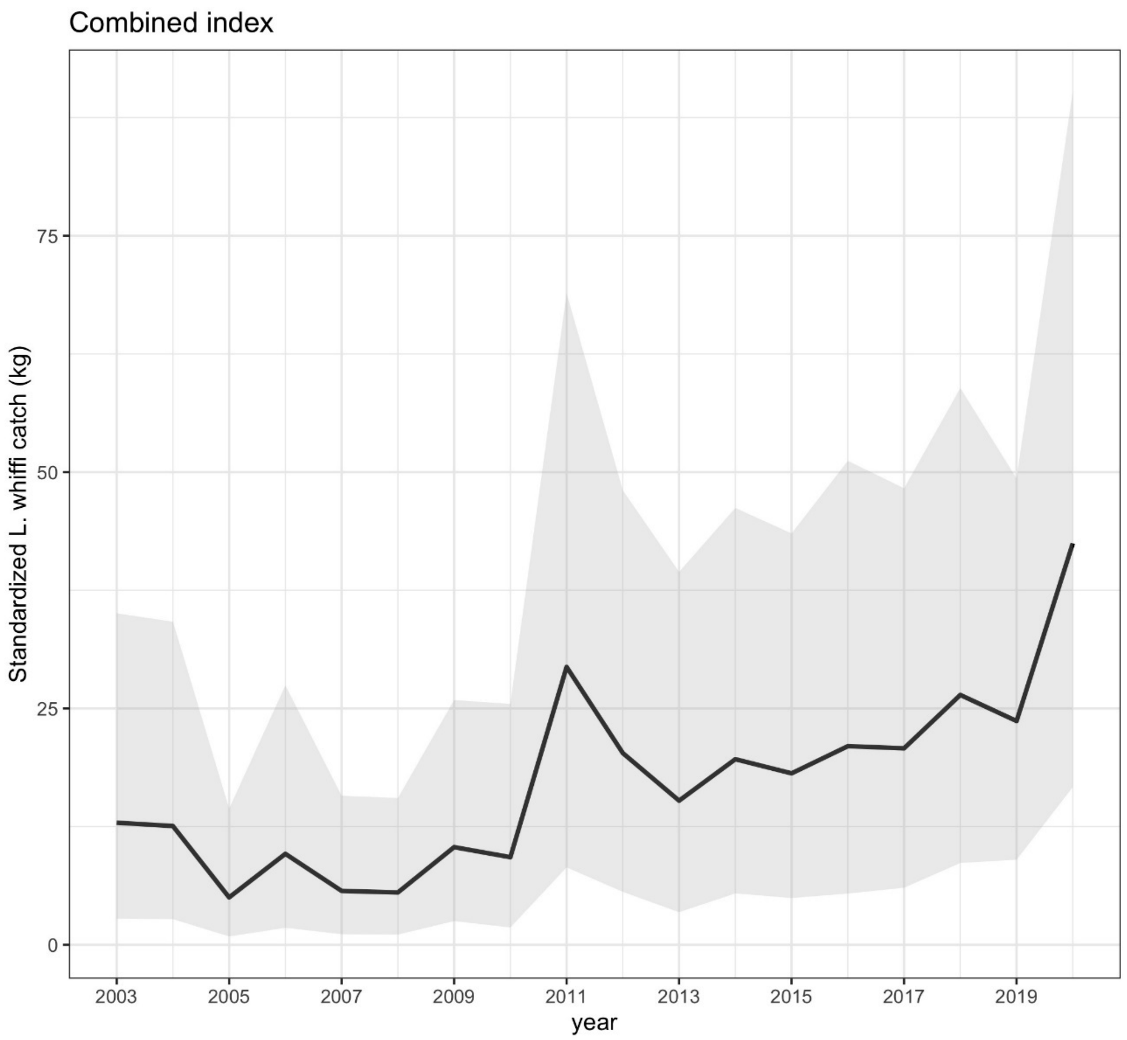

Figure 10: Standardized CPUE combined index for the megrim (*L. whiffiagonis*).

References

Abad, E., Pennino, M. G., Valeiras, J., Vilela, R., Bellido, J. M., Punzón, A., & Velasco, F. (2020). Integrating spatial management measures into fisheries: The Lepidorhombus spp. case study. *Marine Policy*, *116*, 103739.

Fuglstad, G. A., Simpson, D., Lindgren, F., and Rue, H. (2018). Constructing priors that penalize the complexity of Gaussian random fields. *J. Am. Stat. Assoc.* 114, 445–452. doi: 10.1080/01621459.2017.1415907

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2014). *Bayesian Data Analysis*, Vol. 2. Boca Raton, FL: Chapman & Hall.

Izquierdo, F., Paradinas, I., Cerviño, S., Conesa, D., Alonso-Fernández, A., Velasco, F., ... & Pennino, M. G. (2021). Spatio-temporal assessment of the European hake *(Merluccius merluccius*) recruits in the northern Iberian Peninsula. *Frontiers in Marine Science*, *8*, 1.

Pennino, M. G., Paradinas, I., Illian, J. B., Muñoz, F., Bellido, J. M., López-Quílez, A., et al. (2019). Accounting for preferential sampling in species distribution models. *Ecol. Evol.* 9, 653–663. doi: 10.1002/ece3.4789

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Roos, M., and Held, L. (2011). Sensitivity analysis in Bayesian generalized linear mixed models for binary data. *Bayesian Anal.* 6, 259–278. doi: 10.1214/11-BA609

Sánchez, F., & Olaso, I. (2004). Effects of fisheries on the Cantabrian Sea shelf ecosystem. *Ecological Modelling*, *172*(2-4), 151-174.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 64, 583–616. doi: 10.1111/1467-9868.00353

Thorson, J. T., Fonner, R., Haltuch, M. A., Ono, K., & Winker, H. (2016). Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(11), 1794-1807.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* 11, 3571–3594.