NRC Research Press

ARTICLE

Combining multiple data sets to unravel the spatiotemporal dynamics of a data-limited fish stock

Cecilia Pinto, Morgane Travers-Trolet, Jed I. Macdonald, Etienne Rivot, and Youen Vermard

Abstract: The biological status of many commercially exploited fishes remains unknown, mostly due to a lack of data necessary for their assessment. Investigating the spatiotemporal dynamics of such species can lead to new insights into population processes and foster a path towards improved spatial management decisions. Here, we focused on striped red mullet (*Mullus surmuletus*), a widespread yet data-limited species of high commercial importance. Aiming to quantify range dynamics in this data-poor scenario, we combined fishery-dependent and -independent data sets through a series of Bayesian mixed-effects models designed to capture monthly and seasonal occurrence patterns near the species' northern range limit across 20 years. Combining multiple data sets allowed us to cover the entire distribution of the northern population of *M. surmuletus*, exploring dynamics at different spatiotemporal scales and identifying key environmental drivers (i.e., sea surface temperature, salinity) that shape occurrence patterns. Our results demonstrate that even when process and (or) observation uncertainty is high, or when data are sparse, if we combine multiple data sets within a hierarchical modelling framework, accurate and useful spatial predictions can still be made.

Résumé : L'état biologique de nombreuses espèces commerciales n'est pas connu, principalement par manque de données nécessaires à l'évaluation de celle-ci. Explorer les dynamiques spatio-temporelles des espèces peut aboutir à une meilleure compréhension des drivers de ces dynamiques et potentiellement apporter les informations nécessaires à la mise en place de mesures de gestion spatiale. Ce travail se concentre sur le rouget barbet de roche (*Mullus surmuletus*), une espèce commune à forte valeur commerciale mais sur laquelle les données disponibles sont limitées. L'objectif de cette étude était de quantifier les dynamiques spatiales et saisonnières de ces 20 dernières années d'une espèce a données limitées en intégrant données de pêches scientifiques et commerciales dans un cadre de modèles mixtes Bayésiens. Cette méthode de modélisation intégrée nous a permis d'intégrer plusieurs sources de données et de couvrir toute la distribution de la population nord de *M. surmuletus* ainsi que de comprendre les distributions à différentes échelles spatio-temporelles et d'identifier les principales covariables environnementales définissant la présence de l'espèce. Nos résultats montrent que même s'il y a un haut niveau d'incertitude dans le processus et (ou) l'observation dans les zones où les données sont le plus limitées, l'intégration de données multiples dans un cadre de modélisation hiérarchique permet d'obtenir des prédictions spatiales utiles et précises.

Introduction

Long-term time series are a valuable resource for testing hypotheses on how temporal variability in recruitment or abundance or patterns of range expansion or distributional shift may relate to climatic and anthropogenic events (Doney et al. 2012; Hawkins et al. 2013). This is a prerequisite to forecast the response of populations under future scenarios of environmental change and additional anthropogenic pressures, such as fishing pressure (Szuwalski and Punt 2015).

Many fish stocks targeted by fisheries are not subjected to standardized assessment methods (Costello et al. 2012), meaning that both their exploitation level and their resilience to exploitation are uncertain. Nonassessed stocks not only comprise species of low commercial importance, some highly exploited species also fall outside the assessment process. This situation is often due to data scarcity, driven either by a lack of government investment in the fisheries management process or through the history of the data collection itself (Hilborn and Ovando 2014).

Stock assessment methods for so-called data-limited stocks have received considerable interest in recent years, with the development of new methods based on life history traits (e.g., body-size frequencies) or trends in abundance and fleets (ICES 2017; Kokkalis et al. 2017). However, data are still missing for many species or have been monitored only over short time scales. This critically hampers any evaluation process and potentially reduces viability of fish populations and associated fisheries (Costello et al. 2012).

When data collection on a species is sparse or limited, combining multiple data sources within a single analysis can help to overcome the limitation of single data sets considered separately. According to their respective spatiotemporal coverage, combining data sets allows for extending the time series, widening the

Received 12 April 2018. Accepted 23 September 2018.

C. Pinto.* IFREMER – Département Halieutique de Manche – Mer du Nord, 150 Quai Gambetta, BP 699, 62321 Boulogne s/mer, France; Joint Research Centre – Directorate D, Sustainable Resources – Unit D.0, Water and Marine Resources – TP 051, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy.
M. Travers-Trolet. IFREMER – Département Halieutique de Manche – Mer du Nord, 150 Quai Gambetta, BP 699, 62321 Boulogne s/mer, France.
J.I. Macdonald. Faculty of Life and Environmental Sciences, University of Iceland, 101 Reykjavík, Iceland.
E. Rivot. UMR 985 ESE Ecology and Ecosystem Health, Agrocampus Ouest, INRA, 35042 Rennes, France.
Y. Vermard. IFREMER – Département Ecologie et Modèles pour l'Halieutique, Rue de l'Ile d'Yeu, BP 21105, 44311 Nantes cedex 03, France.

Corresponding author: Cecilia Pinto (emails: cecilia.pinto@ec.europa.eu, pntccl@gmail.com).

*Present address: Joint Research Centre – Directorate D, Sustainable Resources – Unit D.0, Water and Marine Resources – TP 051, Via Enrico Fermi 2749, 21027 Ispra (VA), Italy.

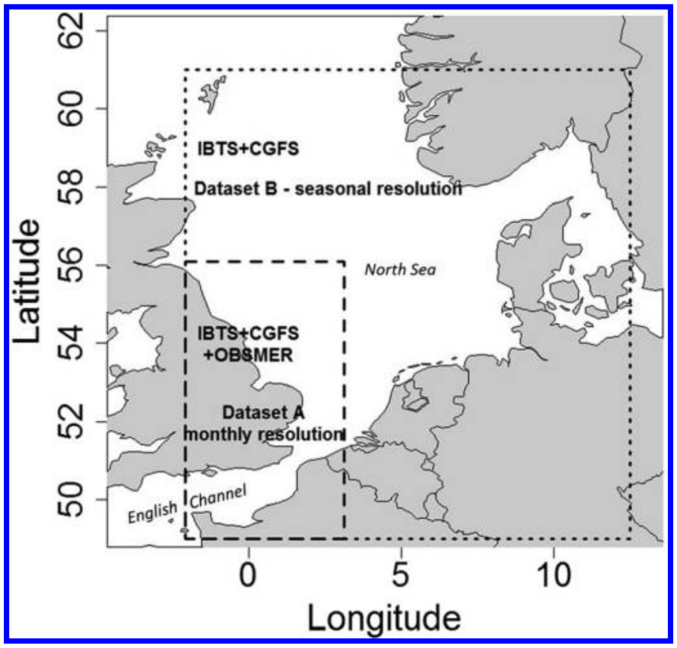
Can. J. Fish. Aquat. Sci. **76**: 1338–1349 (2019) dx.doi.org/10.1139/cjfas-2018-0149

area covered, and ultimately improving the power of the analysis and our understanding of population dynamics. The development of integrated analysis (as defined in Maunder and Punt 2013) as a tool to combine different data sources arising from different sampling methods (with their own spatial and temporal heterogeneity) within a single framework has received attention in the statistical ecology literature (McGeoch and Gaston 2002) and in fisheries sciences (Maunder and Punt 2013 and references therein). Hierarchical modelling approaches explicitly separate out process models from observation models and therefore offer an efficient framework for combining multiple data sets. The process equations allow for modelling multiple dependencies and stochasticity in a hierarchy of scales suitable to depict the spatial and temporal variability present within the data through latent parameters, while the set of observation equations define how the data relate to the state variables of the model (Gelfand 2012; Parent and Rivot 2012; Kéry and Royle 2016). This class of models is also particularly well suited to capturing residual correlation patterns through inclusion of spatial (or temporal) correlation structure in the latent variables (Legendre 1993; Elith and Leathwick 2009; Thorson and Minto 2015). Bayesian inferences on hierarchical models offer additional technical convenience and provide inferences in a probabilistic rationale that fully propagates uncertainty (Punt and Hilborn 1997; Harwood and Stokes 2003).

In this paper, we combine four fishery-dependent and fishery-independent data sets, spanning a 20-year period, within a single hierarchical model to explore monthly and seasonal occurrence patterns of striped red mullet (*Mullus surmuletus*), a demersal mullid of high commercial importance. We focus on the "northern subpopulation" that resides in the North Sea and eastern English Channel and shows little mixing with the "southern subpopulation" (Bay of Biscay; Mahe et al. 2014) and the "mixing zone subpopulation" (Celtic Sea and the western English Channel; Benzinou et al. 2013). Despite being commercially targeted across much of its range, information on the sensitivity of this species to changing environmental conditions is scarce. The hypothesized role of dynamic gradients (e.g., sea surface temperature) in shaping the migration and distribution patterns of the northern subpopulation (Beare et al. 2005; Engelhard et al. 2011) needs further enquiry using data covering the full geographic range of the subpopulation over several years. This northern subpopulation is also characterized by strong oscillations in abundance between consecutive years (Mahé et al. 2005). During the last 5 years, fluctuations have increased in magnitude, concurrently with the loss of the oldest and most efficient spawners from the population (ICES 2017). These indices suggest an effect of overexploitation (Iglésias et al. 2010) and are an alarm bell for future (and perhaps prolonged) depletion. Implementation of restrictive management options are currently debated, such as implementation of quota sharing within the total allowable catch for the subpopulation, as already established in a multilateral context for other species in the North Sea (Hannesson 2013).

Indications of a depleted population state, high abundance variability, and high uncertainty regarding spatial distribution drivers constitute strong motivations to fill in the gaps in biological and ecological knowledge for this species and eventually provide more reliable scientific advice for fisheries management. More specifically, the objectives of our study are twofold: (*i*) to clarify the role of environmental factors on shaping occurrence patterns across the full distributional range of the northern subpopulation of striped red mullet and (*ii*) to gain insight into the mechanisms underpinning the marked interannual fluctuations and seasonal migrations that characterize its spatiotemporal dynamics.

Fig. 1. Spatial coverage of data sets A and B.



Materials and methods

Presence–absence data

Our data are derived from three scientific bottom-trawl surveys and one set of commercial fishery catch records. The scientific surveys were the winter and summer International Bottom Trawl Survey (IBTS) (ICES 2017) and the Channel Ground Fish Survey (CGFS) (Coppin and Travers-Trolet 1989). The IBTS surveys take place over 1 month across January and February (winter survey, IBTS-Q1) and 1 month across August and September (summer survey, IBTS-Q3) and cover the whole of the North Sea. Since 2007, the winter survey has been expanded into the eastern English Channel. The CGFS takes place over 1 month in October and has covered the eastern English Channel since 1990. As the North Sea was not systematically sampled twice a year prior to 1995, only survey data from 1995 to 2015 are considered here. The commercial data come from the OBSMER French program (Cornou et al. 2016), which aims to collect data on landings and discards through onboard observers at sea. Catch data were collected throughout the year (for every fishing operation on each sampled trip) from 2003 to 2015.

The four initial data sets were first reclassified into two new data sets based on their spatial and temporal coverage. Data set A (n = 8391) comprises observations from IBTS-Q1, IBTS-Q3, CGFS, and OBSMER covering the eastern English Channel and the southern North Sea (Fig. 1) and spanning 1995 to 2015 at a monthly resolution. Data set B (n = 13 853) has the same temporal coverage (1995–2015) and covers a larger spatial area than data set A, as it includes the whole of the North Sea, but at the cost of making use of fishery-independent records only (i.e., IBTS-Q1, IBTS-Q3, and CGFS) and with a seasonal (i.e., winter, summer, and autumn) resolution (Fig. 1). The number of records available from each data source is presented in the online Supplementary material, Table S1[1].

For both data sets A and B, georeferenced point records describing the catches of striped red mullet captured at a particular location s and time t were transformed to presence–absence records. This is a critical simplification to limit the effect of heterogeneity in fishing effort and catchability among the various data sets and allows us to consider that all sampling methods are equiv-

---

[1] Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2018-0149.

alently informative relative to the presence–absence of the species. To further limit heterogeneity in the catchability and avoid "false zeros" due to low catchability (Martin et al. 2005) when using OBSMER data, only records from bottom trawlers using a mesh size between 70 and 90 mm were extracted for the analysis, as larger mesh sizes are typically used by boats targeting other species (e.g., *Pollachius virens*).

Environmental covariates

Presence–absence records were correlated with a set of environmental covariates thought to influence the occurrence of striped red mullet: depth at seabed, sediment type, sea surface temperature (SST), and sea surface salinity (SSS). Depth at seabed was extracted from the NORWegian ECOlogical Model system (NORWECOM) database (http://www.imr.no/~morten/wgoofe/). SSS was extracted at a monthly resolution from the NORWECOM website for the time interval from 1990 to 2008, while data for 2009–2015 were obtained by contacting the author of the model system directly. The SST data were obtained from satellite observations at a daily resolution, but for the purposes of this study a monthly mean was computed. Data for 1990–2008 was extracted from the AVHRR Pathfinder version 5.2 (PFV5.2) data set, provided by the US National Oceanographic Data Center and Group for High Resolution Sea Surface Temperature (http://pathfinder.nodc.noaa.gov) (Casey et al. 2010), while the 2009–2015 SST data was extracted from the ODYSSEA processing chain operated within the ESA/MEDSPIRATION project (Gohin et al. 2010). Seabed sediment types were adapted from Larsonneur et al. (1982) and Schluter and Jerosch (2008) and reclassified into five broad categories: mud, fine sand, coarse sand, gravel, and pebbles. To test for collinearity among covariates, we used the "vif.mer" function (variance inflation factor threshold set to 10) on a model object fitted using the "lme4" package (Bates et al. 2015) in R version 3.3.0 (R Core Team 2016) to calculate variance inflation factors for each predictor (R code available from https://github.com/aufrank/R-hacks/blob/master/mer-utils.R). As no collinearity among variables was detected, variance inflation factor values are not shown in the Results.

Modelling striped red mullet occurrence

Data sets A and B were analysed independently but using the same modelling approach.

Models were built in a hierarchical Bayesian framework using the SPDE (stochastic partial differential equations) approach in the "R-INLA" package (http://www.r-inla.org; Rue et al. 2009; Lindgren et al. 2011; Lindgren and Rue 2015) in R. This approach provides direct inference on the spatial and temporal dependencies in the data. The process equation models the probability, $p_t(s)$, of striped red mullet presence at time step t (i.e., either month or season) and location s as a random field on the logit scale:

$$\text{logit}(p_t(s)) = \boldsymbol{X}_t(s) \times \boldsymbol{\beta} + \theta_t(s) \tag{1}$$

where $\boldsymbol{X}_t(s)$ represents a vector of covariates (depth at seabed, sediment, SST, SSS) at time step t and location s, $\boldsymbol{\beta}$ represents a vector of coefficients (fixed effects) to be estimated, and $\theta_t(s)$ is a spatiotemporal random effect to account for variation not explicitly explained by covariates. Random effects are defined by a Gaussian random field that is spatially autoregressive (depending on the distance between locations) and temporally uncorrelated (for details see Cameletti et al. 2013). To avoid computational costs that rapidly arise in continuous space (the so-called "big n problem"; Lasinio et al. 2013), we modelled the spatial covariation within a Gaussian Markov random field on a discrete mesh that defines the area of interest (Krainski et al. 2016; see Fig. S2[1]). This way the influence of spatial covariance at any point s is reduced to a set of neighbours (Cameletti et al. 2013).

Given the latent field of presence probability $p_t(s)$ at any time t and location s, presence–absence data $y_t(s)$ are modelled as mutually independent and identically distributed Bernoulli variables

$$y_t(s) \sim \text{Bernoulli}(p_t(s)) \tag{2}$$

The full likelihood equation for the model then arises from the product of Bernoulli for all raw data (eq. 2). Because all data sources are considered as presence–absence, the strength of the hierarchical structure is that different data sources are integrated within a single analysis to infer a unique random field model for the probability of presence that captures the spatiotemporal covariations as defined in eq. 1.

Within the SPDE approach, eq. 1 can be rewritten as

$$\text{logit}(p_t(s)) = \mathbf{X}_t(s)\boldsymbol{\beta} + \mathbf{A}_t(s)\theta_t \tag{3}$$

where observation matrix $\mathbf{A}_t(s)$ is directly related to the space-discretizing mesh (Fig. S2[1]), as it extracts the values of the spatiotemporal random field at each location s and at each time step t. The realization of the random field can be represented through its mean density distribution and standard deviation, which in turn can be translated as the level of uncertainty at a certain location depending on the availability of data points (Cameletti et al. 2013). The quantification of such uncertainty, through the realization of the random field, allowed us to account for the heterogeneity across time and space of the sampling design, originating from the integration of different data sets.

Different mesh designs were compared visually, and the sensitivity of parameter estimation to the different designs were assessed (Cosandey-Godin et al. 2015). The best mesh designs for each data set (see Fig. S1[1]) include an outer extension to avoid a "boundary effect" (Lindgren and Rue 2015) and regularly shaped triangles, both in the inner and outer extensions and at the border between the two extensions (Krainski et al. 2016). Once the best mesh was selected, parameters values defining it were kept constant across models (i.e., at the same spatial resolution).

The simplified Laplace method was used to approximate the posterior marginal distributions (for details see Martins et al. 2013). We built and compared models of increasing complexity, from the null model including no covariates to the full model including all covariates and random effects. Models were compared through the deviance information criterion (DIC), the log marginal likelihood, and by estimating the variance contribution of random effects against that of fixed effects. To evaluate out-of-sample predictive capacity for each fitted model, we derived the conditional predictive ordinate, defined as the cross-validated predictive density at observation $y_t(s)$ with that observation removed (Roos and Held 2011). We used the conditional predictive ordinate values to compute the cross-validated logarithmic score (Gneiting and Raftery 2007), a measure of predictive quality, and the cross-validated Brier score (i.e., mean prediction error) for each model. This latter score evaluates the correspondence between fitted probabilities and observed binary outcomes (Schmid and Griffith 2005; Roos and Held 2011). Lower values on both scores reflect better predictions, with the Brier score interpreted relative to a reference value equal to sampling prevalence. The probability of presence was predicted across the whole area covered by each data set, but here we limit our spatial predictions to the areas where the standard deviation of the response was smaller than its mean value, also corresponding to the end of the asymptotic phase of its distribution (Fig. S3[1]). Following Ward et al. (2015), we also estimated the predictive accuracy of the best model through the area under the receiver operating characteristic curve using the "ROCR" package (Sing et al. 2005).

Table 1. Summary of models' deviance information criterion (DIC), log marginal likelihood, estimated spatial autocorrelation range (ρ), variance contribution of the spatial effect to the total variance (σ^2), cross-validated logarithmic score (CVLS), and Brier score.

	DIC	Log marginal likelihood	ρ spatial effect	σ^2 spatial effect	CVLS	Brier score
Data set A						
SED+DEP+SSS+SST	**7529**	**−4204**	**2.656599**	**4.604428**	**0.4468516**	**0.07572407**
SED+DEP+SSS	7556	−4222	2.768245	4.856019	0.4482658	0.14610750
SED+DEP	7558	−4219	2.726399	4.836133	0.4483667	0.14609570
SED	7557	−4210	2.770354	4.879648	0.4483378	0.14607490
SSS+SST	7541	−4221	2.431589	5.182673	0.4471932	0.14578490
NO COVS	7569	−4237	2.553919	5.495750	0.4486286	0.14622800
Data set B						
SED+DEP+SSS+SST	**6881**	**−3841**	**8.509531**	**6.554891**	**0.2532677**	**0.07857798**
SED+DEP+SSS	6894	−3854	8.017820	6.658798	0.2536428	0.07861613
SED+DEP	6895	−3848	8.000895	6.626146	0.2536606	0.07861056
SED	6900	−3839	7.889699	6.483401	0.2538337	0.07867360
SSS+SST	7672	−4257	7.474648	6.465650	0.2467375	0.05887704
NO COVS	7021	−3907	7.213805	6.803004	0.2583787	0.08070601

Note: SED, sediment type; DEP, depth at seabed; SSS, sea surface salinity; SST, sea surface temperature; NO COVS, no covariates. Best models for each data set are highlighted in bold.

Table 2. Estimated coefficients for the best models of data sets A and B.

	Data set A	Data set B
Mud	−4.4276 (−30.0362, +21.0294)	−0.7849 (−26.1550, +24.4351)
Fine sand	−4.5374 (−30.1481, +20.9218)	−0.8095 (−26.1797, +24.4106)
Gravels	−4.0908 (−29.7055, +21.3721)	−0.6337 (−26.0048, +24.5872)
Pebbles	−5.0858 (−30.6998, +20.3766)	−1.7490 (−27.1200, +23.4718)
Coarse sand	−3.8629 (−29.4751, +21.5977)	0.3191 (−25.0513, +25.5395)
SST	**+0.2079 (+0.1506, +0.2656)**	**+0.2680 (+0.1805, +0.3560)**
SSS	**+0.6516 (+0.0340, +1.2665)**	−0.0629 (−0.1820, +0.0558)
Depth	−0.0054 (−0.0136, +0.0027)	0.0022 (−0.0015, +0.0058)

Note: Values are posterior means and intervals are 95% Bayesian credibility intervals. Intervals not containing 0 are highlighted in bold.

Priors

We used the default priors for the fixed effects and hyperparameters as implemented in R-INLA (described in Lindgren and Rue 2015). Hyperparameters currently constitute an active area of research for the R-INLA team (see R-INLA documentation available from http://www.r-inla.org/). The latent field parameters θ_1 and θ_2 were defined by a multivariate normal distribution, which is a combination of θ_1 = N(0, 10), θ_2 = N(0, 10). All fixed parameter priors were defined by N(0, 1000), except the intercept that has a prior distribution: N(0, ∞).

Results

Model selection

Models with a month within year structure (for data set A) and season within year structure (for data set B) for the random effect were always preferred based on DIC. Models including all environmental covariates were selected as the best models on the balance of the DIC, the log marginal likelihood estimates, the reduced variance contribution of the spatial effect, and predictive quality (cross-validated logarithmic score and Brier score; Table 1). The spatial correlation range (nominal range) of the best model for data set A was 2.66 decimal degrees and 8.51 for data set B (Table 1). The area under the curve estimated for the best model for data set A was 0.61 and 0.69 for data set B.

Environmental parameters

SST and SSS were both found to be positively correlated with the presence of striped red mullet for data set A while for data set B only SST was significant (Table 2), suggesting that this species has a preference for areas where waters are warmer and more saline. Sediment types were not correlated with the presence of striped red mullet at the monthly time scale for data set A or at a seasonal scale for data set B (Table 2). Finally, depth at seabed had no effect on the distribution of striped red mullet at both resolutions.

Spatial latent field

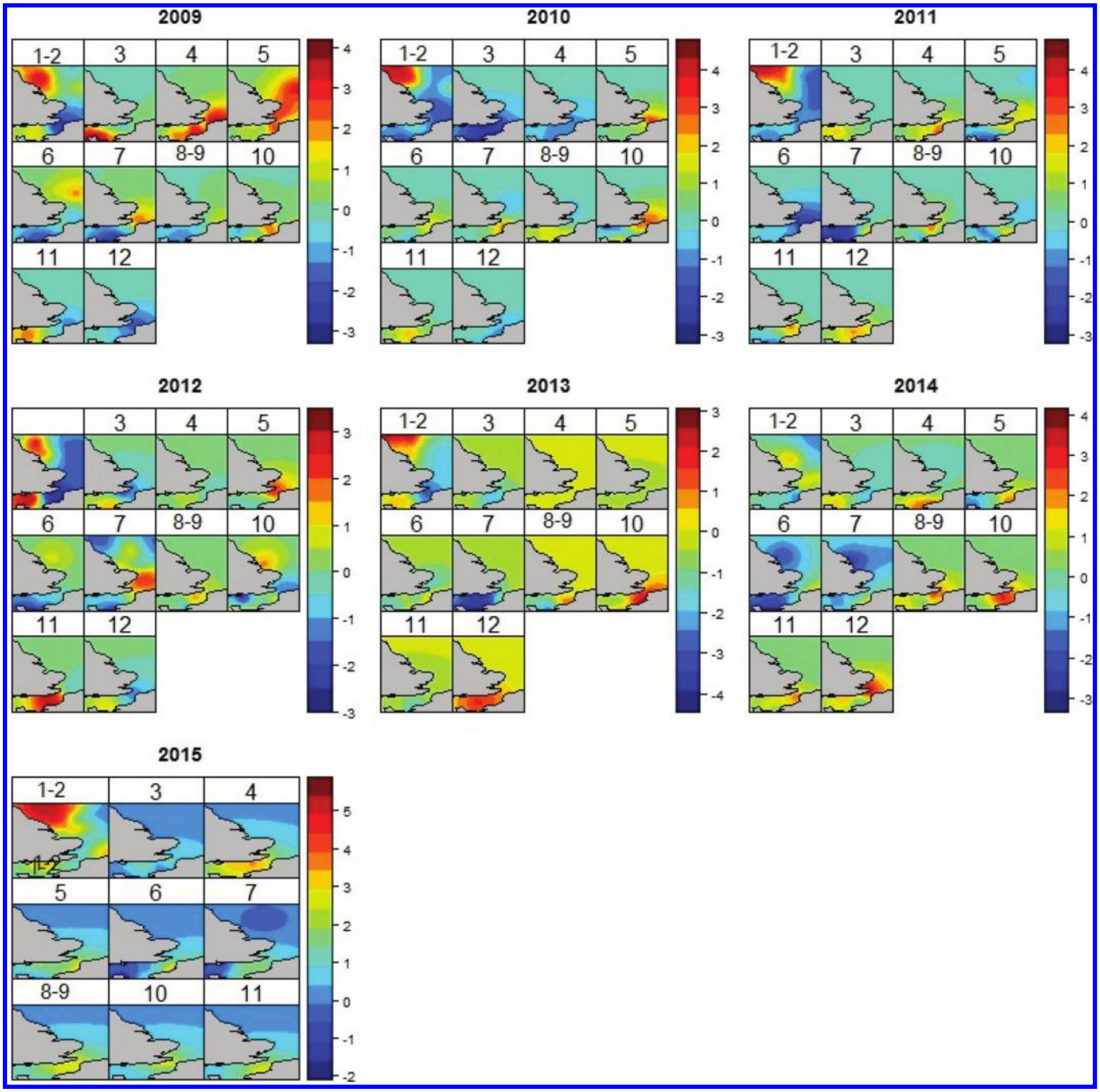
Posterior estimates of the spatial random effects inform about the spatiotemporal variability that is not captured by the effects of covariates in the model. Results are presented from 2009 onwards to allow direct comparison between data sets, as data for all months and seasons are available since 2009 only.

For data set A, posterior estimates of the spatial random effect revealed that the northern subpopulation of striped red mullet changed its distribution month by month, moving into and out of the eastern English Channel (Fig. 2). Owing to the high variation of the monthly distributional pattern among years, it is difficult to identify a consistent monthly movement trend across areas. For data set B in which sampling is consistent between winter and summer, results revealed a seasonal migration of the population moving from the northeast in the winter to the south in the summer, but with higher uncertainty during autumn, as data are available only for the eastern English Channel (Fig. 3). Random effects are estimated with higher uncertainty for data set A than for data set B (Figs. 3, 4), as a consequence of the higher spatial variability inherent in commercial sampling and fewer observations per month, as compared with the more spatially consistent survey observations in data set B.

Predicted probability of presence

Our modelling framework also allows for predicting the probability of presence at any point in the area provided that covariates are available. Figure 5 reports the predicted probability of

Fig. 2. Posterior mean of the spatial random effect for data set A — positive values indicate a high density of presence data, while negative values indicate a high density of absence data. The months of January–February, and August–September were grouped together to combine the parts of the IBTS survey that straddled months. [Colour online.]



presence of striped red mullet in the area covered by data set A, only in areas where the results are more reliable. These reliable areas were defined as the ones displaying a low standard deviation (i.e., low uncertainty) using a cutoff of 13.23 (the mean of the standard deviation also corresponding to the upper limit of the distribution where the standard deviation is more or less constant; see also Fig. S3[1]). Beyond the interannual and seasonal variability in the probability of striped red mullet presence, recurrent patterns can be detected. Results highlight a strong seasonal difference, with high predicted probability of presence (>70%) from July to October when the surface waters are warmer and there is a lower probability of presence (<50%) predicted for colder months (late winter) (Fig. 5 and Fig. S4A[1]). We also detected changes across years linked to SST; during the coldest springs of the series (2010 and 2013; Fig. S4B[1]), the probability of presence was lower for these seasons (Fig. 5 and Fig. S4A[1]). Predictions also suggest large-scale seasonal movements of the striped red mullet across the study area. The striped red mullet seems to spend the winter in the English Channel, before leaving this area in March and reaching the Dover Strait by April, although this pattern varies across years. We note also that in 2015, the probability of presence remains high in the English Channel throughout the year, but on average,

Fig. 3. Mean (left-hand side) and standard deviation (right-hand side) of the spatial random effect at a seasonal resolution for data set B — positive values of the mean indicate a high density of presence data, while negative values indicate a high density of absence data. The standard deviation increases with distance from the data points. [Colour online.]

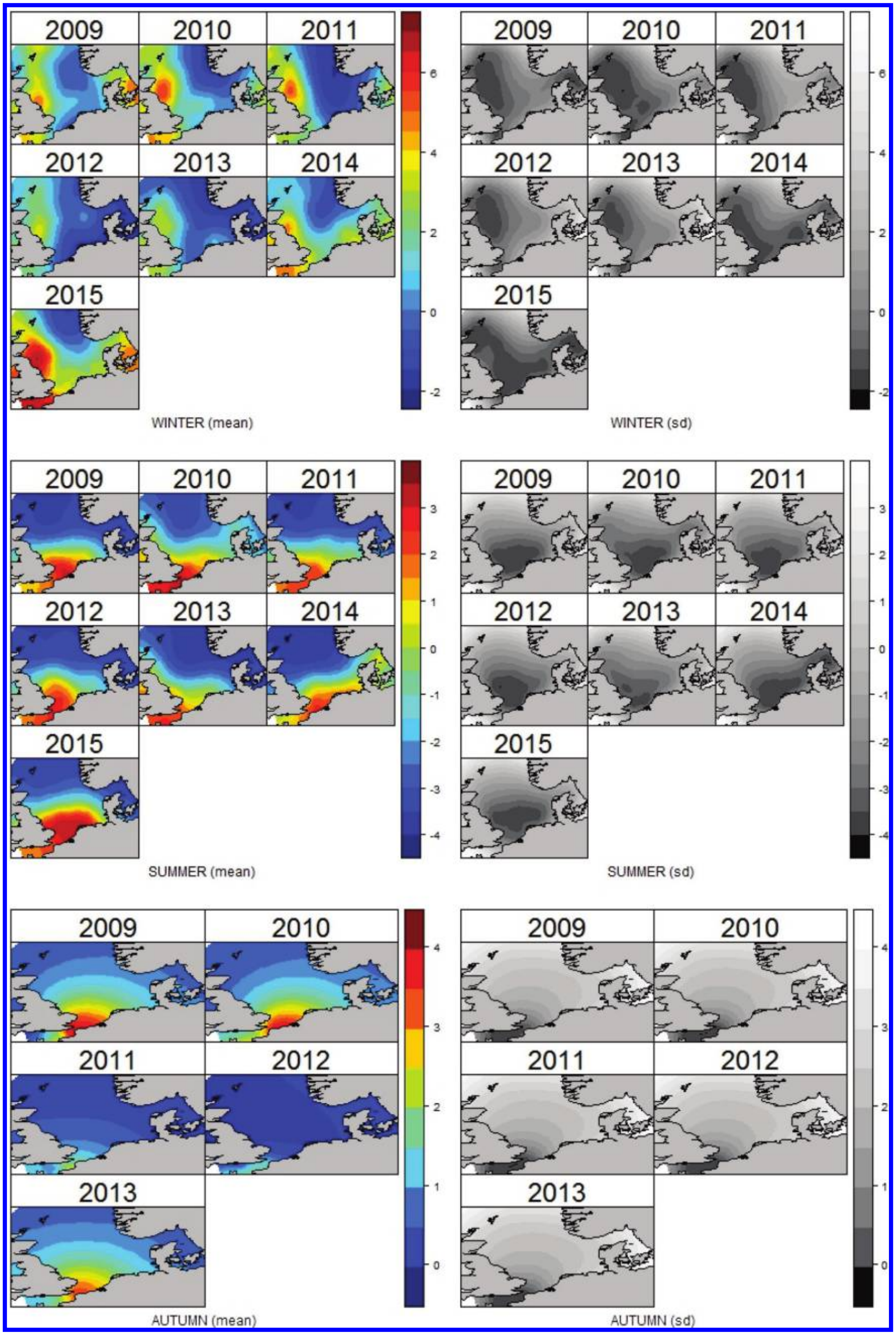
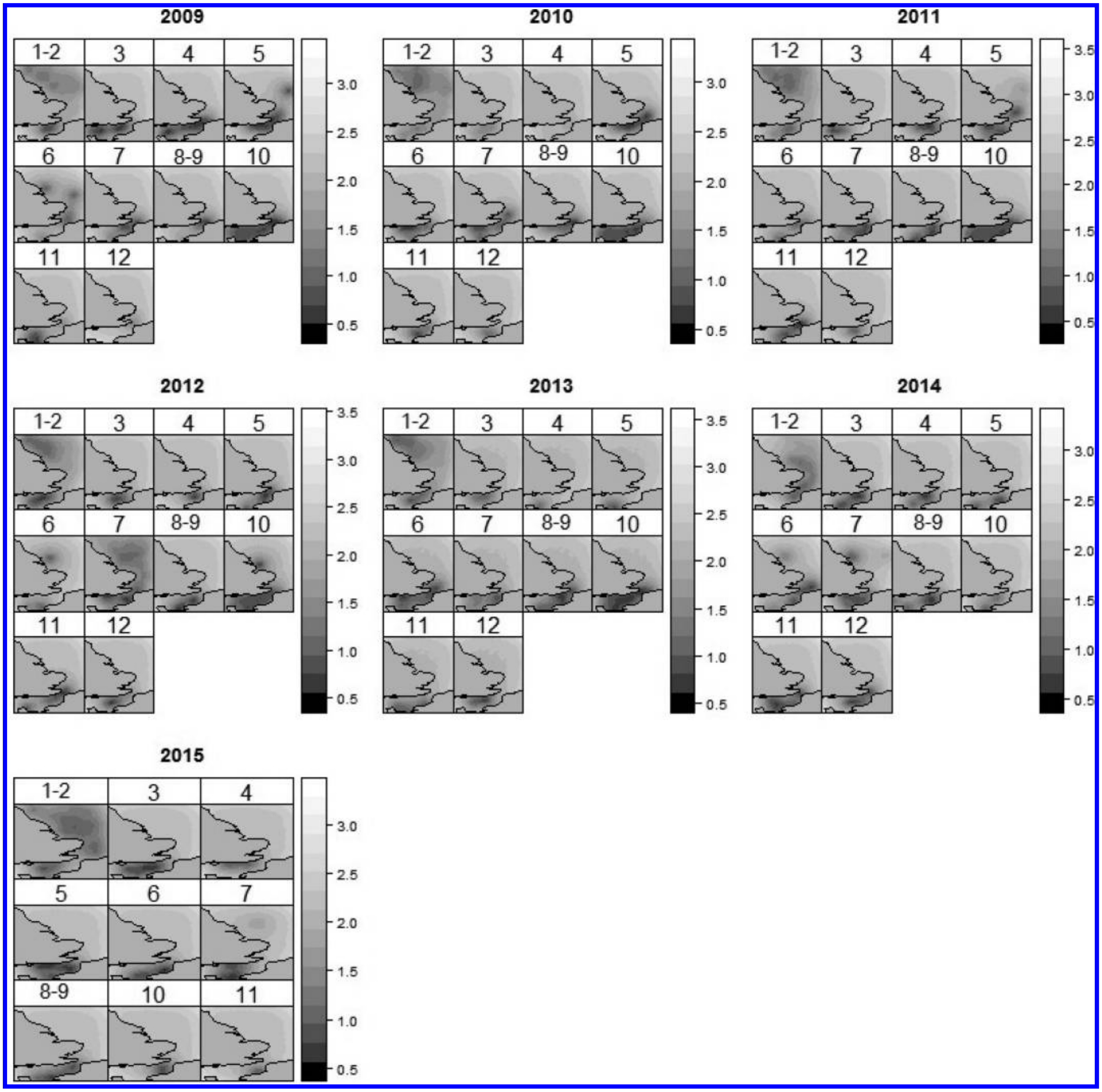


Fig. 4. Standard deviation of the spatial random effect for data set A. The months of January–February and August–September were grouped together to combine the parts of the IBTS survey that straddled months.



there are no strong variations in the registered SSTs compared with the previous years (Fig. S4A[1]).

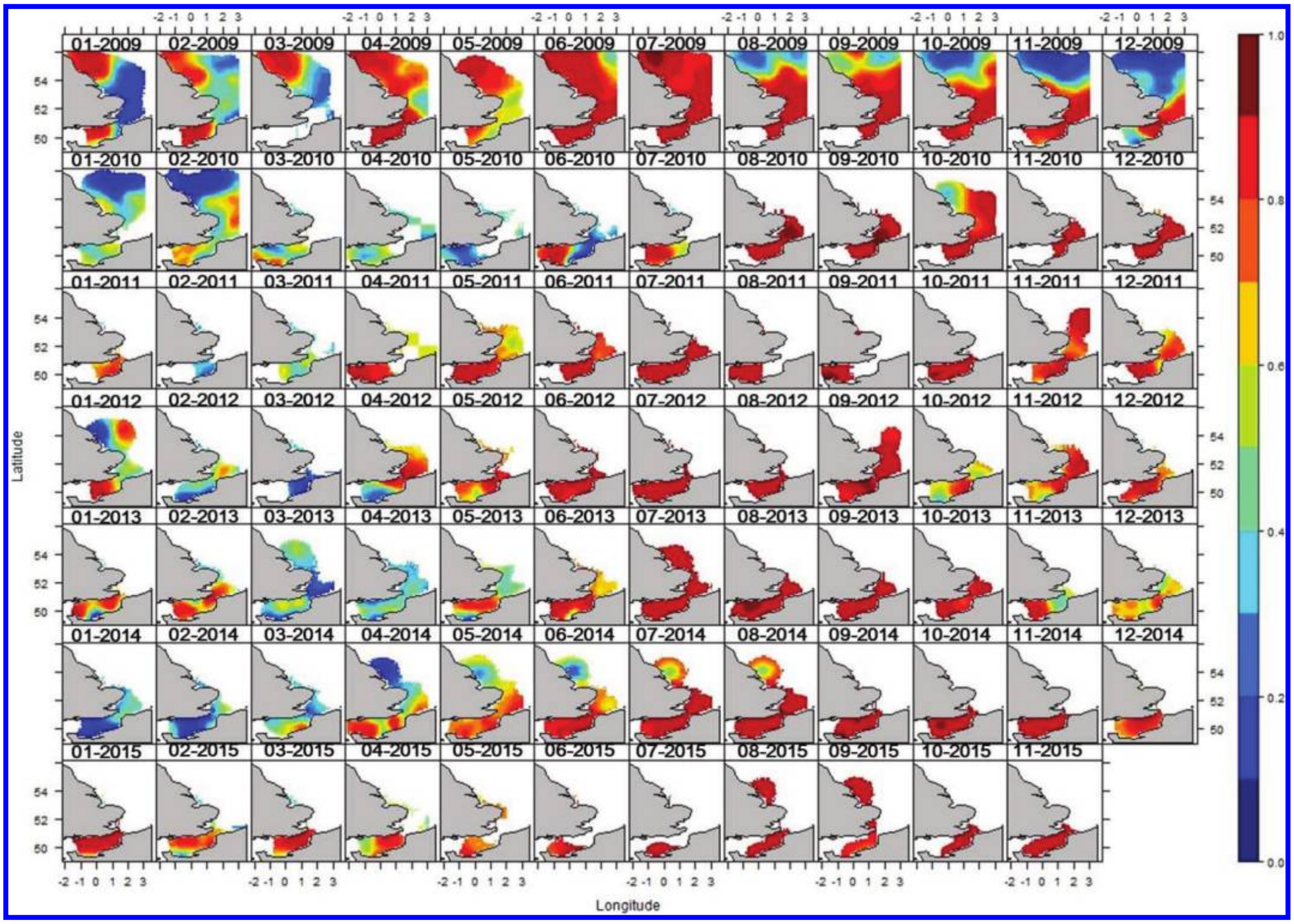
Predictions obtained from data set B (Fig. 6) seem contradictory, as they show a much higher probability of presence of striped red mullet during the winter, specifically in the northwest of the North Sea and the eastern English Channel. In the summer and the autumn, however, the probability of presence increased only in the eastern English Channel and, across all grid cells, was 20% lower than the winter period.

Discussion

This study provides the first spatially explicit analysis of how environmental parameters may shape the distribution of striped red mullet near its northern range boundary. All available information on this data-limited species was integrated into a single analysis that directly accounts for correlation structures in the data and the sources of uncertainty in data and process. The results provide a substantive contribution to our understanding of the spatiotemporal dynamics of this data-limited stock.

Our findings suggest that the occurrence of the northern subpopulation is positively correlated with water salinity and temperature. Results for the latter covariate match suggestions by Beare et al. (2005), who hypothesized that the presence of striped red mullet in northern waters in winter was related to increasing surface water temperatures. Moreover, our predictions show that certain years are characterized by larger occupied areas (e.g., 2011

Fig. 5. Spatial predictions of the probability of presence of striped red mullet in the eastern English Channel and southern North Sea at a monthly resolution from 2009 (top) to 2015 (bottom), as output from the best model for data set A. White areas represent grid cells in which the standard deviation was higher than the mean standard deviation (on a logit scale; see Fig. S2[1]). [Colour online.]

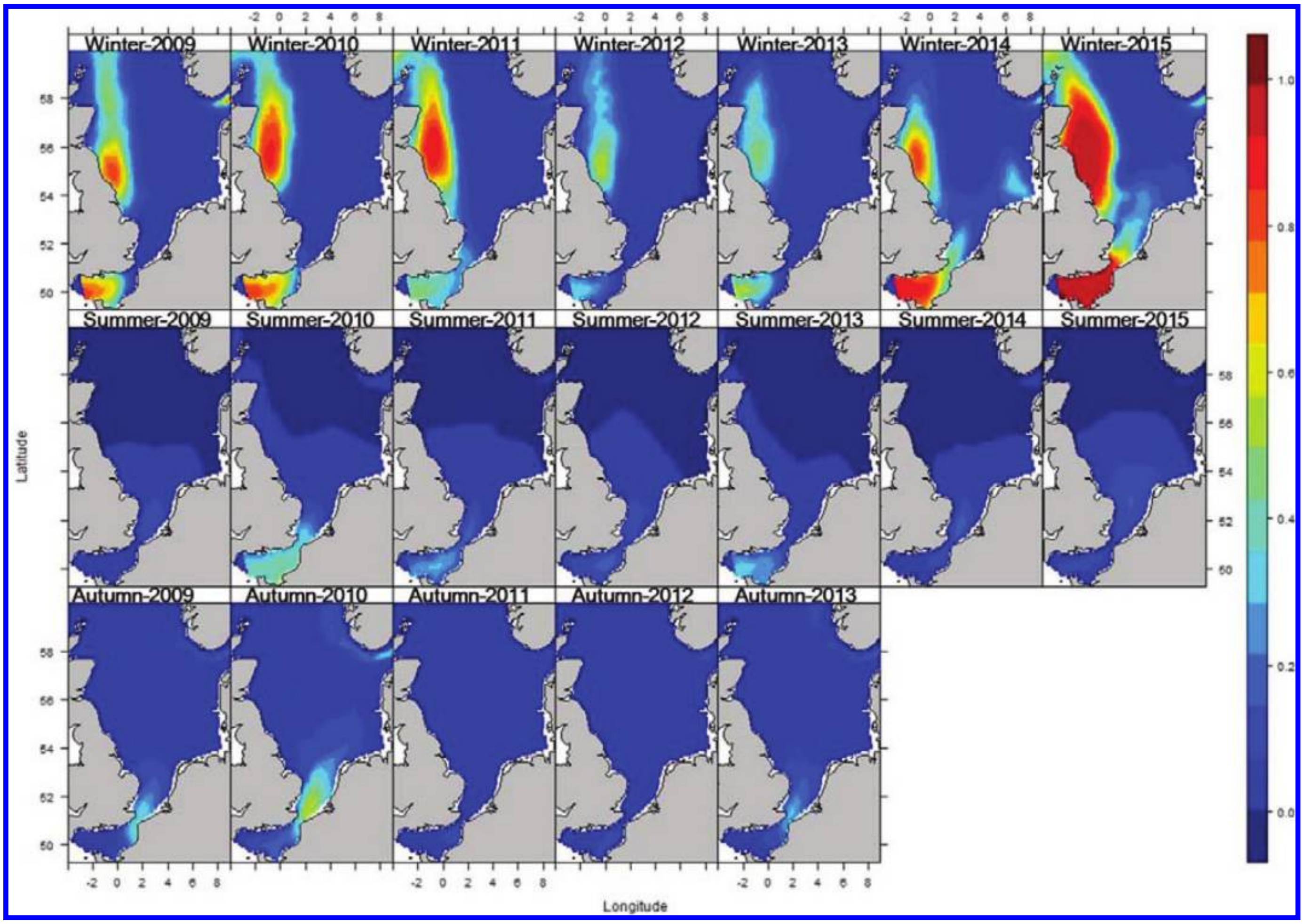


and 2015) interspersed with years of very low and (or) scattered concentrations (e.g., 2013; see Figs. 5 and 6). This complements previous descriptions of the strong interannual fluctuations in abundance within this subpopulation (Mahé et al. 2005; Carpentier et al. 2009). Whether range expansion is linked to population size in this species (see Fisher and Frank 2004) remains an open question. Yet, despite marked interannual variability in the distributional range, we detected patterns of seasonal migrations starting in both the northwest of the North Sea and the English Channel in winter, moving to the south of the North Sea in spring–summer, and entering the eastern English Channel in autumn. Previous work has often focused on the effect of water temperature (and other environmental parameters) on population abundances. For instance, striped red mullet has increased in abundance by 30% over the last two decades in the English Channel, concomitantly with a shift towards a warmer phase of the Atlantic multidecadal oscillation index (Auber et al. 2015). Cheung et al. (2013) predicted that species preferring warmer waters will increase in abundance and dominate fisheries catches in northern latitudes, as appears to be occurring in species such as Atlantic mackerel (*Scomber scombrus*) and hake (*Merluccius merluccius*) (Jansen 2014; Baudron and Fernandes 2015; Hughes et al. 2015).

However, abundances are not necessarily correlated to distribution extensions. Therefore, when developing spatial management frameworks to improve fisheries management, coupling the dynamics of both abundance and spatial distribution will likely prove productive to move forward. That said, presence–absence data are often more easily obtained and more widely available than abundance data, and modelling presence–absence can make the integration of data obtained from heterogeneous surveys simpler. Indeed, provided that detectability of the survey method(s) is considered to be 100%, meaning that at least one individual will be captured if the species is in fact present, presence–absence models allow us to largely ignore variation in catchability among different survey methods and sampling gears.

Although results from both data sets identify a positive effect of SST on the presence of striped red mullet, when looking at predictions we find that the highest probability of presence (both in terms of area and absolute values) is predicted in the summer months for data set A but not for data set B. Such discrepancy between the models built from the two data sets may result from several (nonmutually exclusive) hypotheses. (*i*) The effect of temperature on the population could be stronger in the winter than in the summer, causing the SST coefficients to vary throughout the seasons. (*ii*) Gradients in environmental factors could be steeper in data set B than in data set A, as a direct function of the larger area covered by data set B. This would be consistent with the fact that the relative variation in DIC when including covariates in the model is sharper in data set B than in data set A, suggesting that these have a stronger explanatory power than in data set A. This is confirmed also by the consistent contribution of the spatial variance and by the increased autocorrelation range,

Fig. 6. Spatial predictions of the probability of presence of striped red mullet in the eastern English Channel and the North Sea at a seasonal resolution from 2009 (left) to 2015 (right), as output from the best model for data set B. [Colour online.]



which are not affected by the addition of covariates in data set A. (*iii*) Last, the difference observed in model predictions between data sets could also result from an effect of sampling bias. Data set B is derived mainly from IBTS data that are consistently sampled every winter and every summer. Data set A also contains data from the IBTS surveys, but is complemented by the OBSMER data. Though incorporating true absences, this commercial data set is still potentially biased by variation in nominal and spatial commercial fishing effort that shifts not only between months but also between years (Fig. S6[1]). Although we cannot completely rule out seasonally variable fishing effort as contributing to our spatial predictions for data set A, we suggest that any effects are relatively minor given our use of presence–absence data as previously discussed. Building two different models based on the two different data sets allows us to glean the maximum possible information from both and improve our understanding of the species' dynamics at different spatial and temporal scales. Data set A provides insight into the spatiotemporal dynamics of the northern population of striped red mullet at a monthly level that could be missed using only data set B, which instead exposes the seasonal dynamics at a larger spatial scale, using spatially consistent survey information. Importantly, the lower level of sampling heterogeneity in data set B suggests where, spatially, the predictions from data set A may be less reliable due to the high uncertainty given by nonconsistent sampling.

Integrating multiple surveys in a single data set (either within A or B) allowed us to increase both the number of observations and our capacity to detect statistical flukes (Maunder and Punt 2013). Moreover, comparing the results from the separate analyses of data sets A and B allowed us to expand the geographical area (eastern English Channel and southern North Sea for data set A, whole North Sea for data set B) and explore the consistency of our inferences across two different spatial scales and at two different temporal resolutions (monthly for A and seasonal for B). As noted by Maunder and Punt (2013), when integrating multiple data sources, a trade-off should be found to maximize the scientific reward of integrated modelling. Integrating various sources of data in the same analysis does not necessarily give rise to improved understanding of the target system, as it may lead to conflicts in what the data sets tell us, in addition to increasing statistical complexity and computational costs. Separating the data set built by integrating the IBTS-Q1, IBTS-Q3, CGFS, and OBSMER data into two subsets that differed in spatial and temporal resolution was our trade-off. Analysing data set A alone allowed us to obtain inferences at a monthly level instead of just at a seasonal level. Additionally, a separate analysis of data sets A and B provided insights on the effects of environmental parameters at different spatial scales.

The major source of uncertainty in the data came from the lack of commercial data for single months in years prior to 2003 (the time series is complete for each month only from 2009 onwards). Confidence surrounding the estimates on this subpopulation during this time period is therefore relatively low, and further efforts are needed to improve data quality. A substantial impediment to progress on this front relates to the difficulties in accessing commercial catch data coming from observer programs that operate in countries bordering the North Sea. The advantages of having observer data from foreign fisheries targeting local stocks was demonstrated in the Alaskan fisheries (French et al. 1982) and stands in stark contrast with the situation in the eastern English Channel and North Sea area, where multiple countries similarly share the quota on several harvested stocks. Hannesson (2013) showed that cooperation always brings in more advantages than competition when stock harvesting is shared among parties. Hence, strong incentive exists to integrate all the available data — both fishery-dependent and -independent (e.g., national onboard observer programs) — to maximize coverage of spatiotemporal information in commercial stocks.

Species that are commercially exploited though not formally managed are particularly vulnerable to overexploitation, as their population dynamics are often not monitored, with no limits set on landings or minimum sizes. Using striped red mullet for illustration, our results have demonstrated some advantages of data integration and explicitly accounting for uncertainty under data limitation; however, it is important to note that these steps alone are not the silver bullet for successful fisheries management. Instead, we hope this work inspires future sampling designs, data collection, and multilateral data-sharing programs that in conjunction with appropriate modelling approaches can lead to better adaptive management decisions for data-limited populations (Walters 2007; Maunder and Punt 2013).

Acknowledgements

CP's postdoc was funded by Ifremer and France Filière Peche. The authors thank Bruno Ernande for suggestions and comments that improved the work during the analysis. The authors also thank two anonymous reviewers for their comments, which helped to improve the manuscript.

References

Auber, A., Travers-Trolet, M., Villanueva, M.C., and Ernande, B. 2015. Regime shift in an exploited fish community related to natural climate oscillations. PLoS ONE, **10**(7): e0129883. doi:10.1371/journal.pone.0129883. PMID:26132268.

Bates, D., Mächler, M., Bolker, B.M., and Walker, S.C. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. **67**: 1–48. doi:10.18637/jss.v067.i01.

Baudron, A.R., and Fernandes, P.G. 2015. Adverse consequences of stock recovery: European hake, a new "choke" species under a discard ban? Fish Fish. **16**(4): 563–575. doi:10.1111/faf.12079.

Beare, D., Burns, F., Jones, E., Peach, K., and Reid, D. 2005. Red mullet migration into the northern North Sea during late winter. J. Sea Res. **53**(3): 205–212. doi:10.1016/j.seares.2004.06.003.

Benzinou, A., Carbini, S., Nasreddine, K., Elleboode, R., and Mahé, K. 2013. Discriminating stocks of striped red mullet (*Mullus surmuletus*) in the Northwest European seas using three automatic shape classification methods. Fish. Res. **143**: 153–160. doi:10.1016/j.fishres.2013.01.015.

Cameletti, M., Lindgren, F., Simpson, D., and Rue, H. 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Adv. Stat. Anal. **97**: 109–131. doi:10.1007/s10182-012-0196-3.

Carpentier, A., Martin, C.S., and Vaz, S. 2009. Channel Habitat Atlas for marine Resource Management (CHARM phase II), INTERREG 3a Programme, IFREMER, Boulogne-sur-mer, France.

Casey, K.S., Brandon, T.B., Cornillon, P., and Evans, R. 2010. The Past, Present, and Future of the AVHRR Pathfinder SST Program. *In* Oceanography from Space, revisited. *Edited by* V. Barale, J.F.R. Gower, and L. Alberotanza. pp. 273–287. doi:10.1007/978-90-481-8681-5.

Cheung, W.W.L., Watson, R., and Pauly, D. 2013. Signature of ocean warming in global fisheries catch. Nature, **497**(7449): 365–368. doi:10.1038/nature12156.

Coppin, F., and Travers-Trolet, M. 1989. CGFS: Channel Ground Fish Survey [online]. Available from http://dx.doi.org/10.18142/11.

Cornou, A.-S., Diméet, J., Goascoz, N., Quinio-Scavinner, M., and Rochet, M.-J. 2016. Captures et rejets des metiers de peche francais Resultats des observations a bord des navires de peche professionnelle en 2015. *In* OBSMER.

Cosandey-Godin, A., Krainski, E.T., Worm, B., and Flemming, J.M. 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Can. J. Fish. Aquat. Sci. **72**(2): 186–197. doi:10.1139/cjfas-2014-0159.

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., and Lester, S.E. 2012. Status and solutions for the world's unassessed fisheries. Science, **338**: 517–520. doi:10.1126/science.1223389.

Doney, S.C., Ruckelshaus, M., Emmett Duffy, J., Barry, J.P., Chan, F., English, C.A., Galindo, H.M., Grebmeier, J.M., Hollowed, A.B., Knowlton, N., Polovina, J., Rabalais, N.N., Sydeman, W.J., and Talley, L.D. 2012. Climate change impacts on marine ecosystems. Annu. Rev. Mar. Sci. **4**(1): 11–37. doi:10.1146/annurev-marine-041911-111611.

Elith, J., and Leathwick, J.R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. **40**(1): 677–697. doi:10.1146/annurev.ecolsys.110308.120159.

Engelhard, G.H., Pinnegar, J.K., Kell, L.T., and Rijnsdorp, A.D. 2011. Nine decades of North Sea sole and plaice distribution. ICES J. Mar. Sci. **68**(6): 1090–1104. doi:10.1093/icesjms/fsr031.

Fisher, J., and Frank, K. 2004. Abundance-distribution relationships and conservation of exploited marine fishes. Mar. Ecol. Prog. Ser. **279**: 201–213. doi:10.3354/meps279201.

French, R., Nelson, R., Jr., and Wall, J. 1982. Role of the United States Observer Program in management of foreign fisheries in the Northeast Pacific Ocean and Eastern Bering Sea. N. Am. J. Fish. Manage. **2**(August): 122–131. doi:10.1577/1548-8659(1982)2<122:ROTUSO>2.0.CO;2.

Gelfand, A.E. 2012. Hierarchical modeling for spatial data problems. Spat. Stat. **1**: 30–39. doi:10.1016/j.spasta.2012.02.005.

Gneiting, T., and Raftery, A.E. 2007. Strictly proper scoring rules, prediction, and estimation. J. Am. Stat. Assoc. **102**: 359–378. doi:10.1198/016214506000001437.

Gohin, F., Saulquin, B., and Bryere, P. 2010. Atlas de la Température, de la concentration en Chlorophylle et de la Turbidité de surface du plateau continental français et de ses abords de l'Ouest européen. Ifremer.

Hannesson, R. 2013. Sharing a migrating fish stock. Mar. Resour. Econ. **28**(1): 1–17. doi:10.5950/0738-1360-28.1.1.

Harwood, J., and Stokes, K. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends Ecol. Evol. **18**(12): 617–622. doi:10.1016/j.tree.2003.08.001.

Hawkins, S.J., Firth, L.B., McHugh, M., Poloczanska, E.S., Herbert, R.J.H., Burrows, M.T., Kendall, M.A., Moore, P.J., Thompson, R.C., Jenkins, S.R., Sims, D.W., Genner, M.J., and Mieszkowska, N. 2013. Data rescue and re-use: recycling old information to address new policy concerns. Mar. Pol. **42**: 91–98. doi:10.1016/j.marpol.2013.02.001.

Hilborn, R., and Ovando, D. 2014. Reflections on the success of traditional fisheries management. ICES J. Mar. Sci. **71**(5): 1040–1046. doi:10.1093/icesjms/fsu034.

Hughes, K.M., Dransfeld, L., and Johnson, M.P. 2015. Climate and stock influences on the spread and locations of catches in the northeast Atlantic mackerel fishery. Fish. Oceanogr. **24**(6): 540–552. doi:10.1111/fog.12128.

ICES. 2017. Report of the ICES Workshop on the Development of Quantitative Assessment Methodologies based on Life-History Traits, Exploitation Characteristics, and Other Relevant Parameters for Stocks in Categories 3–6 (WKLIFEVI). 3–7 October 2016, Lisbon, Portugal.

Iglésias, S.P., Toulhoat, L., and Sellos, D.Y. 2010. Taxonomic confusion and market mislabelling of threatened skates: important consequences for their conservation status. Aquat. Conserv. Mar. Freshw. Ecosyst. **20**(3): 319–333. doi:10.1002/aqc.1083.

Jansen, T. 2014. Pseudocollapse and rebuilding of North Sea mackerel (*Scomber scombrus*). ICES J. Mar. Sci. **71**(2): 299–307. doi:10.1093/icesjms/fst148.

Kéry, M., and Royle, J.A. 2016. Chapter 2 — What are hierarchical models and how do we analyze them? *In* Applied hierarchical modeling in ecology. doi:10.1016/B978-0-12-801378-6.00002-3.

Kokkalis, A., Eikeset, A.M., Thygesen, U.H., Steingrund, P., and Andersen, K.H. 2017. Estimating uncertainty of data limited stock assessments. ICES J. Mar. Sci. **74**: 69–77. doi:10.1093/icesjms/fsw145.

Krainski, E.T., Lindgren, F., Simpson, D., and Rue, H. 2016. The R-INLA tutorial: SPDE models [online]. Available from http://www.r-inla.org/examples/tutorials/spde-tutorial

Larsonneur, C., Bouysse, P., and Auffret, J.P. 1982. The superficial sediments of the English Channel and its western approaches. Sedimentology, **29**: 851–864. doi:10.1111/j.1365-3091.1982.tb00088.x.

Lasinio, G.J., Mastrantonio, G., and Pollice, A. 2013. Discussing the "big *n* problem". Stat. Meth. Appl. **22**: 97–112. doi:10.1007/s10260-012-0207-2.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology, **74**(6): 1659–1673. doi:10.2307/1939924.

Lindgren, F., and Rue, H. 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. **63**(19): 1–26. doi:10.18637/jss.v063.i19.

Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. **73**: 423–498. doi:10.1111/j.1467-868.2011.00777.x.

Mahé, K., Destombes, A., Coppin, F., Koubbi, P., Vaz, S., Roy, D.L., and Carpentier, A. 2005. Le rouget barbet de roche *Mullus surmuletus* (L. 1758) en Manche orientale et mer du Nord. Ifremer/Crpmem, Vol. 39, pp. 1–199.

Mahé, K., Villanueva, M.C., Vaz, S., Coppin, F., Koubbi, P., and Carpentier, A. 2014. Morphological variability of the shape of striped red mullet *Mullus surmuletus* in relation to stock discrimination between the Bay of Biscay and the eastern English Channel. J. Fish Biol. **84**(4): 1063–1073. doi:10.1111/jfb.12345.

Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., and Possingham, H.P. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol. Lett. **8**(11): 1235–1246. doi:10.1111/j.1461-0248.2005.00826.x.

Martins, T.G., Simpson, D., Lindgren, F., and Rue, H. 2013. Bayesian computing with INLA: new features. Comput. Stat. Data Anal. **67**: 68–83. doi:10.1016/j.csda.2013.04.014.

Maunder, M.N., and Punt, A.E. 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. **142**: 61–74. doi:10.1016/j.fishres.2012.07.025.

McGeoch, M.A., and Gaston, K.J. 2002. Occupancy frequency distributions: patterns, artefacts and mechanisms. Biol. Rev. **77**(3): 311–331. doi:10.1017/s1464793101005887.

Parent, É., and Rivot, É. 2012. Introduction to hierarchical Bayesian modeling for ecological data. Chapman & Hall/CRC Press, Boca Raton, Fla.

Punt, A., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Rev. Fish Biol. Fish. **6**: 35–63.

R Core Team. 2016. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing. Vienna, Austria. Available from 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.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. **71**: 319–392. doi:10.1111/j.1467-9868.2008.00700.x.

Schluter, M., and Jerosch, K. 2008. Digital atlas of the North Sea (DANS). Geo-information regarding geology, geochemistry, oceanography and biology [online]. Available from https://www.researchgate.net/publication/274890019_Digital_Atlas_of_the_North_Sea_DANS_Geo-Information_regarding_geology_geochemistry_oceanography_and_biology.

Schmid, C.H., and Griffith, J.L. 2005. Multivariate classification rules: calibration and discrimination. *In* Encyclopedia of biostatistics. 2nd ed. *Edited by* P. Armitage and T. Colton. John Wiley & Sons, London, UK. pp. 3491–3497.

Sing, T., Sander, O., Beerenwinkel, N., and Lengauer, T. 2005. ROCR: visualizing classifier performance in R. Bioinformatics, **21**(20): 3940–3941. doi:10.1093/bioinformatics/bti623.

Szuwalski, C.S., and Punt, A.E. 2015. Can an aggregate assessment reflect the dynamics of a spatially structured stock? Snow crab in the eastern Bering Sea as a case study. Fish. Res. **164**: 135–142. doi:10.1016/j.fishres.2014.10.020.

Thorson, J.T., and Minto, C. 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J. Mar. Sci. **72**(5): 1245–1256. doi:10.1093/icesjms/fsu213.

Walters, C.J. 2007. Is adaptive management helping to solve fisheries problems? AMBIO J. Hum. Environ. **36**(4): 304–307. doi:10.1579/0044-7447(2007)36[304:IAMHTS]2.0.CO;2.

Ward, E.J., Jannot, J.E., Lee, Y., Ono, K., Shelton, A.O., and Thorson, J.T. 2015. Using spatiotemporal species distribution models to identify temporally evolving hotspots of species co-occurrence. Ecol. Appl. **25**: 2198–2209. doi:10.1890/15-0051.1.

Appendix A. Notes on the modelling approach

The mixed effects models we fit in this paper fall broadly within the class of "empirical" statistical models as defined by Levins (1966). These types of models are in essence correlative, although they may have mechanistic underpinnings related to the fundaments of Grinnellian and Eltonian niches (Hutchinson 1957; Soberón 2007; Beale et al. 2014). In lieu of the oft-lacking, detailed physiological knowledge needed for parameterization of an exciting new family of process-based models (e.g., Freitas et al. 2010; Jørgensen et al. 2012; Teal et al. 2012; see Peck et al. 2018 for a review), correlative models, which tend to compromise generality for realism and precision (Levins 1966; Dickey-Collas et al. 2014), remain widely used in ecology to explore the nature of relationships between species' distributions and biotic and abiotic factors, to build hypotheses, and to guide management decisions (Guisan and Thuiller 2005; Elith and Leathwick 2009; Robinson et al. 2011).

Our models were fitted in a Bayesian framework in R-INLA, using the SPDE approach to capture spatial and temporal dependence in the data (Rue et al. 2009; Lindgren et al. 2011). The merits of the Bayesian approach for this type of hierarchical model are many (Gelfand et al. 2006; Gelman and Hill 2007; Royle et al. 2007). Without reviewing these exhaustively here (see Elderd and Miller 2016 for a comprehensive appraisal), we highlight the inherent way in which random effects are handled as parameters of interest, resulting in fully specified probability distributions from which information on the intensity and uncertainty of the effects can be drawn; the option to incorporate prior knowledge based on empirical data or theory; and the ability to robustly quantify and propagate uncertainty through all modelling stages. Model fitting using INLA is computationally efficient and provides accurate approximations of the posterior marginal distributions of model parameters that show high concordance with MCMC simulations (Rue and Martino 2007; Rue et al. 2009; Held et al. 2010). Since Lindgren and colleagues proved that a continuously indexed Gaussian field described by a Matérn covariance function can be represented as a discretely indexed Gaussian Markov random field (Rue and Held 2005; Lindgren et al. 2011), rapid development of the SPDE approach within R-INLA has facilitated fitting of an expanding suite of hierarchical spatial and spatiotemporal models to spatial point patterns (Krainski et al. 2016). This approach has recently proven useful in analyses of georeferenced fisheries data sets, which are often data-rich and where inference at the scale of point locations, rather than grids, is required (e.g., Cosandey-Godin et al. 2015; Ono et al. 2016; Ward et al. 2015).

One of the well-noted criticisms of correlative species distribution models (see Elith and Leathwick 2009 for a review of different methods) has been their inability to adequately account for residual autocorrelation in space and (or) time. This situation can violate independence assumptions in regression models, leading to inference errors and (or) misrepresentation of covariate importance (Legendre 1993; Dormann 2007; Beale et al. 2010). The SPDE approach considers these correlation structures directly and allows great flexibility in their specification (e.g., Cosandey-Godin et al. 2015). We specified temporally independent realizations of the spatially structured error terms, but temporal dependence can easily be coded (e.g., Macdonald et al. 2018).

Our models were specific to striped red mullet in the North Sea and English Channel. However, the approach used is easily adaptable to other stocks and species for which questions on the drivers of distribution shifts remain open. The 20-year data set we analyzed represents a substantial compilation of georeferenced records on the environmental conditions experienced by *M. surmuletus* across a substantial part of its range. The model outputs therefore provide a basis for identifying physiological thresholds that can be used to develop more informative priors in future regression models (Simpson et al. 2015) or to guide parameterization of mechanistic models (Teal et al. 2018). We agree with Rochette et al. (2013), who advocate a hierarchical Bayesian framework as an appealing platform upon which to meld different types of data and models together, making it possible to assimilate the processes acting on different life history phases within the one "full life cycle" model. We see potential for the types of models developed here to contribute to the development of such a model for *M. surmuletus*.

References

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J., and Elston, D.A. 2010. Regression analysis of spatial data. Ecol. Lett. **13**: 246–264. doi:10.1111/j.1461-0248.2009.01422.x. PMID:20102373.

Beale, C.M., Brewer, M.J., and Lennon, J.J. 2014. A new statistical framework for the quantification of covariate associations with species distributions. Methods Ecol. Evol. **5**: 421–432. doi:10.1111/2041-210X.12174.

Cosandey-Godin, A., Krainski, E.T., Worm, B., and Flemming, J.M. 2015. Applying Bayesian spatiotemporal models to fisheries bycatch in the Canadian Arctic. Can. J. Fish. Aquat. Sci. **72**(2): 186–197. doi:10.1139/cjfas-2014-0159.

Dickey-Collas, M., Payne, M.R., Trenkel, V.M., and Nash, R.D.M. 2014. Hazard warning: model misuse ahead. ICES J. Mar. Sci. **71**: 2300–2306. doi:10.1093/icesjms/fst215.

Dormann, C.F. 2007. Effects of incorporating spatial autocorrelation into the analysis of species distribution data. Global Ecol. Biogeogr. **16**: 129–138. doi:10.1111/j.1466-8238.2006.00279.x.

Elderd, B.D., and Miller, T.E.X. 2016. Quantifying demographic uncertainty: Bayesian methods for integral projection models. Ecol. Monogr. **86**: 125–144. doi:10.1890/15-1526.1.

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

Freitas, V., Cardoso, J.F.M.F., Lika, K., Peck, M.A., Campos, J., Kooijman, S.A.L.M., and van der Veer, H.W. 2010. Temperature tolerance and energetics: a dynamic energy budget-based comparison of North Atlantic marine species. Philos. Trans. R. Soc. B Biol. Sci. **365**: 3553–3565. doi:10.1098/rstb.2010.0049.

Gelfand, A.E., Silander, J.A., Jr., Wu, S., Latimer, A., Lewis, P.O., Rebelo, A.G., and

Holder, M. 2006. Explaining species distribution patterns through hierarchical modeling. Bayesian Analysis, **1**: 41–92. doi:10.1214/06-BA102.

Gelman, A., and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, New York, USA.

Guisan, A., and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. **8**: 993–1009. doi:10.1111/j.1461-0248.2005.00792.x.

Held, L., Schrödle, B., and Rue, H. 2010. Posterior and cross-validatory predictive checks: a comparison of MCMC and INLA. *In* Statistical modelling and regression structures. *Edited by* T. Kneib and G. Tutz. Physica-Verlag, Berlin, Germany. pp. 91–110.

Hutchinson, G.E. 1957. Concluding remarks. Cold Spring Harb. Symp. Quant. Biol. **22**: 415–427. doi:10.1101/SQB.1957.022.01.039.

Jørgensen, C., Peck, M.A., Antognarelli, F., Azzurro, E., Burrows, M.T., Cheung, W.W.L., Cucco, A., Holt, R.E., Huebert, K.B., Marras, S., McKenzie, D., Metcalfe, J., Perez-Ruzafa, A., Sinerchia, M., Fleng Steffensen, J., Teal, L.R., and Domenici, P. 2012. Conservation physiology of marine fishes: advancing the predictive capacity of models. Biol. Lett. **8**: 900–903. doi:10.1098/rsbl.2012.0609. PMID:22859560.

Krainski, E.T., Lindgren, F., and Simpson, D. 2016. The R-INLA tutorial on SPDE models. Technical report to the Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology, **74**: 1659–1673. doi:10.2307/1939924.

Levins, R. 1966. The strategy of model building in population biology. Am. Sci. **54**: 421–431.

Lindgren, F., Rue, H., and Lindström, J. 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. **73**: 423–498. doi:10.1111/j.1467-9868.2011.00777.x.

Macdonald, J.I., Logemann, K., Krainski, E.T., Sigurðsson, Þ., Beale, C.M., Huse, G., Hjøllo, S.S., and Marteinsdóttir, G. 2018. Can collective memories shape fish distributions? A test, linking space-time occurrence models and population demographics. Ecography, **41**: 938–957. doi:10.1111/ecog.03098.

Ono, K., Shelton, A.O., Ward, E.J., Thorson, J.T., Feist, B.E., and Hilborn, R. 2016. Space-time investigation of the effects of fishing on fish populations. Ecol. Appl. **26**: 392–406. doi:10.1890/14-1874. PMID:27209782.

Peck, M.A., Arvanitidis, C., Butenschön, M., Canu, D.M., Chatzinikolaou, E., Cucco, A., Domenici, P., Fernandes, J.A., Gasche, L., Huebert, K.B., Hufnagl, M., Jones, M.C., Kempf, A., Keyl, F., Maar, M., Mahévas, S., Marchal, P., Nicolas, D., Pinnegar, J.K., Rivot, E., Rochette, S., Sell, A.F., Sinerchia, M., Solidoro, C., Somerfield, P.J., Teal, L.R., Travers-Trolet, M., and van de Wolfshaar, K.E. 2018. Projecting changes in the distribution and productivity of living marine resources: a critical review of the suite of modelling approaches used in the large European project VECTORS. Estuar. Coast. Shelf Sci. **201**: 40–55. doi:10.1016/j.ecss.2016.05.019.

Robinson, L.M., Elith, J., Hobday, A.J., Pearson, R.G., Kendall, B.E., Possingham, H.P., and Richardson, A.J. 2011. Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. Global Ecol. Biogeogr. **20**: 789–802. doi:10.1111/j.1466-8238.2010.00636.x.

Rochette, S., Le Pape, O., Vigneau, J., and Rivot, E. 2013. A hierarchical Bayesian model for embedding larval drift and habitat models in integrated life cycles for exploited fish. Ecol. Appl. **23**: 1659–1676. doi:10.1890/12-0336.1. PMID:24261047.

Royle, J., Kéry, M., Gautier, R., and Schmid, H. 2007. Hierarchical spatial models of abundance and occurrence from imperfect survey data. Ecol. Monogr. **77**: 465–481. doi:10.1890/06-0912.1.

Rue, H., and Held, L. 2005. Gaussian Markov random fields: theory and applications. Monographs on statistics and applied probability. Vol. 104. Chapman & Hall, London, UK.

Rue, H., and Martino, S. 2007. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. J. Stat. Plan. Infer. **137**: 3177–3192. doi:10.1016/j.jspi.2006.07.016.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J. R. Stat. Soc. Ser. B Stat. Methodol. **71**: 319–392. doi:10.1111/j.1467-9868.2008.00700.x.

Simpson, D.P., Rue, H., Martins, T.G., Riebler, A., and Sørbye, S.H. 2015. Penalising model component complexity: a principled, practical approach to constructing priors. arXiv:1403.4630.

Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecol. Lett. **10**: 1115–1123. doi:10.1111/j.1461-0248.2007.01107.x. PMID:17850335.

Teal, L.R., van Hal, R., van Kooten, T., Ruardij, P., and Rijnsdorp, A.D. 2012. Bio-energetics underpins the spatial response of North Sea plaice (*Pleuronectes platessa* L.) and sole (*Solea solea* L.) to climate change. Glob. Change Biol. **18**: 3291–3305. doi:10.1111/j.1365-2486.2012.02795.x.

Teal, L.R., Marras, S., Peck, M.A., and Domenici, P. 2018. Physiology-based modelling approaches to characterize fish habitat suitability: their usefulness and limitations. Estuar. Coast. Shelf Sci. **201**: 56–63. doi:10.1016/j.ecss.2015.11.014.

Ward, E.J., Jannot, J.E., Lee, Y., Ono, K., Shelton, A.O., and Thorson, J.T. 2015. Using spatiotemporal species distribution models to identify temporally evolving hotspots of species co-occurrence. Ecol. Appl. **25**: 2198–2209. doi:10.1890/15-0051.1. PMID:26910949.

This article has been cited by:

1. Ute Daewel, Corinna Schrum, Jed I. Macdonald. 2019. Towards end-to-end (E2E) modelling in a consistent NPZD-F modelling framework (ECOSMO E2E_v1.0): application to the North Sea and Baltic Sea. *Geoscientific Model Development* **12**:5, 1765-1789. [Crossref]