

L'Institut Agro Rennes-Angers
 Site d'Angers Site de Rennes

Année universitaire : 2022-2023

Spécialité :

Ingénieur Agronome

Spécialisation (et option éventuelle) :

« Sciences Halieutiques et Aquacoles »
(Ressources et Ecosystèmes Aquatiques)

Mémoire de fin d'études

d'ingénieur de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)

de master de l'Institut Agro Rennes-Angers (Institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement)

de l'Institut Agro Montpellier (étudiant arrivé en M2)

d'un autre établissement (étudiant arrivé en M2)

A data-integration approach with hierarchical models to reveal spatio-temporal variations of small pelagic fish distribution in the Bay of Biscay

Par : Aurel HEBERT--BURGGRAEVE



Soutenu à Rennes le 15/09/2023

Devant le jury composé de :

Président : Pablo Brosset

Maître de stage : Mathieu Doray,

Mathieu Authier et Maxime Olmos

Enseignant référent : Pablo Brosset

Autres membres du jury (Nom, Qualité) :

Etienne Rivot, Enseignant-Chercheur en écologie
halieutique, Institut Agro Rennes

Martin Huret, Chercheur en écologie marine, Ifremer
Brest

Ce document est soumis aux conditions d'utilisation «Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France» disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle de l'Institut Agro Rennes-Angers



Fiche de confidentialité et de diffusion du mémoire

Confidentialité

Non Oui si oui : 1 an 5 ans 10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible ⁽¹⁾.

Date et signature du maître de stage ⁽²⁾ : 15/09/23

(ou de l'étudiant-entrepreneur)

A la fin de la période de confidentialité, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant à renseigner).

Droits d'auteur

L'auteur ⁽³⁾ **HEBERT--BURGGRAEVE Aurel**

autorise la diffusion de son travail (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si oui, il autorise

la diffusion papier du mémoire uniquement⁽⁴⁾

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

(Facultatif) accepte de placer son mémoire sous licence Creative commons CC-By-Nc-Nd (voir Guide du mémoire Chap 1.4 page 6)

Date et signature de l'auteur : 15/09/23

Autorisation de diffusion par le responsable de spécialisation ou son représentant

L'enseignant juge le mémoire de qualité suffisante pour être diffusé (immédiatement ou à la fin de la période de confidentialité)

Oui Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.

Si oui, il autorise

la diffusion papier du mémoire uniquement⁽⁴⁾

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire

Date et signature de l'enseignant : 15/09/2023

(1) L'administration, les enseignants et les différents services de documentation de l'Institut Agré Rennes-Angers s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme

(3) Auteur = étudiant qui réalise son mémoire de fin d'études

(4) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option) sera signalée dans les bases de données documentaires sans le résumé

Acknowledgments

First of all, I would like to express my gratitude to my tutors Mathieu Doray, Matthieu Authier and Maxime Olmos. It was a great adventure not always easy because of the schedule of each.

I would like to thank Baptiste and Florian who made themselves available and answered my questions on several occasions.

I would like to thank Guillermo Boyra in charge of the JUVENA survey for the access to the spanish data used in this work.

I would also like to thank the Pelgas team for their warm welcome on board. Jean-Hervé accompanied me well in the footsteps of Mathieu and the acoustic hunt for small pelagic fish. Alejandro was a discreet ninja room-mate who introduced me to the comic strip "Il faut flinguer Ramirez".

I would also like to thank my many colleagues in Nantes. Victor's madeleines, Maud's kindness, Maël's nerf gun, Aïmen's slyness, this evening role-playing game with Corentin, Klervi, Ian, Hugo, Baptiste and Mathilde. I'd also like to say a special thank you to my office mates Paula and Ian, with whom I shared the same trials and tribulations of a final year work placement. And finally, Laetitia was a great babysitter during this very deserted summer.

I'd also like to mention the very warm welcome I received from the Porterie Handball Club and the good company of Delys, the navibus driver.

Thank you to all my REA friends who shared this year with me.

Thank you to all my family and my girlfriend for their unfailing support in my professional project.

Contents

Notation

1. Introduction.....	1
Context: common dolphin by-catch.....	1
Prey/predator relationship: small pelagic fishes and dolphins in the BoB.....	1
Sardine, anchovy and their environment.....	2
Species distribution modelling and data integration.....	2
2. Materials and Methods.....	4
Study area.....	4
Survey data.....	5
Commercial fishing data.....	7
Sardine and anchovy fisheries, and VMS data.....	7
Aggregating the data per season and matching the surveys timing.....	7
Data filtering and Preliminary analysis.....	8
Hierarchical Modelling.....	9
Spatio-temporal resolution of the model.....	9
Observation Process and likelihoods.....	11
Link functions and conditional expectations.....	12
Seasonal effect and environmental covariates.....	13
Priors.....	14
Inference tool INLA.....	14
Convergence assessment.....	15
Model selection.....	15
Modelling strategy.....	15
Derived quantities.....	16
3. Results.....	16
CPUEs and survey biomass comparison.....	16
Multi-season modelling.....	18
Convergence of models.....	18
Shared Gaussian Random Field.....	19
Predictions of probability of sardine presence.....	23
Predictions of sardine biomass / intensity.....	24

Prediction of occurrence of the sardine fisheries.....	25
4. Discussion.....	27
An integrated modelling framework to infer spatio-temporal distribution of fish abundance ..	27
A spatio-temporal model to account for seasonal and inter-annual variations in small pelagic fish.....	27
An integrated hierarchical model including survey and fishery data.....	30
Accounting for covariates.....	31
Computational limits to build an integrated modelling framework.....	31
A hierarchical model to allocate the different data information into shared terms (representing the distribution of seasonal and inter-annual variations) and terms specific to the different data sources.....	31
A shared random field to infer seasonal and inter-annual variations distribution from different data sources.....	31
A specific random field associated to each process (presence only, presence/absence and intensity of catches).....	32
5. Conclusion.....	33
Bibliography.....	34
Appendices.....	40

List of Figures

Figure 1: Bay of Biscay and study area

Figure 2: Spatial distribution of presence/absence per survey (JUVENA/PELGAS) and species (anchovy-ENGR-ENC / sardine-SARD-PIL) of all the year 2009 – 2022

Figure 3: Spatial distribution of presence commercial data per fishing fleet and per season over the period 2009 – 2022

Figure 4: Extrapolation grid of the study area (1 cell = 2,5 x 2,5 km)

Figure 5: Triangulated mesh of the study area

Figure 6: CPUE of September (t/h) by biomass of Juvena (t) for the sardine

Figure 7: CPUE of May of PS (t/h) by the biomass of Pelgas (t) for anchovy

Figure 8: CPUE of September (t/h) by the biomass of Juvena (t) for anchovy

Figure 9: Evolution of the intercepts of the models with covariate and models without covariate over the series (spring in green, summer in blue and fall in red)

Figure 10: Average posterior mean, standard deviation and average error of the common spatial random effect for all the years in logarithm scale

Figure 11: Posterior mean and average error of the common spatial random effect for 2015, 2017, 2022

Figure 12: Posterior mean of the common spatial random effect for all the series (2009-2022)

Figure 13: Average probability, standard deviation and average error of sardine presence for all the year (avg_pred_pres correspond to the average prediction of probability of presence) (Appendix 8)

Figure 14: Prediction, standard deviation and average error of sardine intensity for all the year in logarithm scale (avg_pred_int correspond to the average prediction of biomass intensity) (Appendix 9)

Figure 15: Prediction, standard deviation and average error of sardine fishery intensity for all the year in logarithm scale (avg_pred_fish_int correspond to the average prediction of fishery intensity) (Appendix 10)

List of Appendices

Appendix 1: Sardine model input data

Appendix 2: Prior: explanation and visualisation

Appendix 3: Fit indices provided by INLA to check the convergence

Appendix 4: Grids maps of the deviance residual associated to each process from the model in 2017

Appendix 5: Maps of the bathymetry (right) and the distance from the coast (left) of the BoB

Appendix 6: Summary of the convergence assessment indicators

Appendix 7: Focus on 2017

Appendix 8: Prediction of probability of sardine presence for all the year

Appendix 9: Prediction of sardine intensity for all the year

Appendix 10: Prediction of sardine fishery intensity for all the year

Notation

BoB: Bay of Biscay

Chl-a: Chlorophyll-a

CNRS: Centre National de la Recherche Scientifique

CPUE: Catch Per Unit Effort

EDSU: Elementary Distance Sampling Unit

Ifremer: Institut français de recherche pour l'exploitation de la mer

INLA: Integrated Nested Laplace Approximation

OFB: Office Français de la Biodiversité

PTM: Pelagic Pair Trawl Small Pelagic Fish

PS: Purse Seine

SDM: Species Distribution Model

SA: Sardine and Anchovy

SGF: Shared Gaussian random Field

SPF: Small Pelagic Fish

SSS: Sea Surface Salinity

SST: Sea Surface Temperature

VMS: Vessel Monitoring Systems

1. Introduction

Context: common dolphin by-catch

In the North-East Atlantic ocean, common dolphin (*Delphinus delphis* L. 1758) is a small oceanic dolphin accidentally caught in various fisheries (Taylor et al., 2022). Since the 1990's, an increasing number of common dolphins have been found stranded on the Atlantic seaboard every year. In the European countries, France has the highest record of common dolphins stranded on beaches (Rouby, 2022). This last 10 years, the mean of stranded dolphins per year is about 1250 (Dars et al., 2021). The stranded dolphins that could be necropsied by the French Stranding Network usually had their stomach full, indicating that they were feeding at the time of death (ICES, 2021), and bore fish net marks. The cause of death is by-catch, the incidental and non-intentional capture of non-target species in fishing gears.

From 2016 to 2018, the number of common dolphin by-catch was estimated between 5,000 and 10,000 per year (Peltier et al., 2020; Peltier et al., 2022). In 2017, during 482 days in the Bay of Biscay (BoB), observers on fishing fleets revealed 19 by-catch events involving a total of 65 common dolphins. Raised to the total effort of all fleets reporting common dolphins by-catches, it led to an estimate of 8,904 common dolphins bycaught (ICES, 2021). Deterministic projections from demographic model of population dynamics have suggested the population would be reduced to 20% of its current size in 30 years and be extinct in 100 years (Mannocci et al., 2014). The main part of the common dolphins stranding from by-catch come from the continental shelf of the BoB and correlations have been demonstrated between fishing activities (French gillnetters, especially those targeting high predators, Spanish bottom trawlers, Danish sennes, etc.) and by-catch (Peltier et al., 2019; Dars et al., 2021; Peltier et al., 2021).

In July 2020, the European Commission started an infringement procedure against three member states, including France, for the non-implementation of a strict protection system for common dolphins, as legally required under the provisions of the 'Habitats' Directive (European Commission, 2020). On the 20th march 2023, the state council of France issued an opinion asking the French Government to implement measures within 6 months to reduce the number of small cetaceans that are by-caught by fishing vessels flying the French flag in the BoB.

At the request of the Ministry for the Sea and the Ministry for Ecological Transition, the Office Français de la Biodiversité (OFB), Ifremer, La Rochelle University and the CNRS signed a declaration of intent to form a partnership in December 2020 to increase knowledge and propose remedial solutions. The Delmoges project follows on from this declaration and proposes to: i) produce new ecological and fisheries knowledge to improve understanding of the determinants of by-catch, ii) develop a range of scenarios to reduce by-catch and iii) assess these scenarios with their socio-economic and territorial consequences (Delmoges, 2021). The first purpose of this project focuses on dolphins distribution modelling to compare with fisheries effort and the production of risk maps to improve management. Our study focuses on modelling the dolphins' prey distribution to bring information better understand the impact of trophic interactions on the dolphin distribution.

Prey/predator relationship: small pelagic fishes and dolphins in the BoB

Prey and predator distributions influence each other. Quantifying the scale and valence of spatial prey-predator correlations is crucial in trophic ecology (Lambert et al., 2019). Common dolphins are opportunistic feeders that feed primarily on small epipelagic fish under 20 cm (Lahaye et al., 2005; Pusineri et al., 2007; Meynier et al., 2008; Spitz et al., 2010; Lambert et al., 2019). It was suggested that there is a correlation between energy density of preys and dolphins diet (Murphy et al., 2013). In the BoB, common dolphins consume mostly fat, i.e energetic, fish (Lahaye et al., 2005). Sardine seems to be the dominant prey in weight in the common dolphin diet in summer, fall and winter while horse mackerel is absent in summer and the quantity found during spring and fall was significantly higher from winter (Meynier et al., 2008). In contrast, anchovy proportion in diet did not show significant temporal variations (Meynier et al., 2008). The mean proportion of sardine, anchovy and horse mackerel from the dolphin daily food intake was respectively 45%, 21% and 5% (Meynier et al., 2008).

This haven't been done yet in the BoB but a model was developed for common dolphins in Greek waters including the effect of the probability of the presence of sardine (Giannoulaki et al., 2017). In this model, the sardine distribution was correlated positively with the common dolphin estimation. So modelling the distribution of the main prey of common dolphin, such as sardine, anchovies and horse mackerel (Lahaye et al., 2005; Meynier et al., 2008; Marçalo et al., 2018) can inform on the dolphin distribution.

Sardine, anchovy and their environment

In the Bay of Biscay, European sardine (*Sardina pilchardus*) and anchovy (*Engraulis encrasicolus*) (Whitehead, 1985; Coombs et al., 2006) (SA is the abbreviation for Sardine and Anchovy here after) are hypothesised to be one of the main prey for common dolphin (Pusineri et al., 2007; Meynier et al., 2008). SA live primarily in continental shelf waters (Fréon and Misund, 1999; Fréon et al., 2005; Petitgas et al., 2010). In the BoB, European sardine, or pilchard, lives during the day in dense schools near the seabed close to the coast, and sometimes near sea surface offshore (Doray et al., 2018a). Common anchovy is found in less dense schools near the seabed mainly in coastal areas and sometimes near sea surface in the BoB during daytime (Petitgas et al., 2010; Doray et al., 2018a). SA feed mainly on zooplankton (Chapuis et al., 2021; Modrak et al., 2022). They play a key ecological role in the coastal ecosystem, transferring plankton energy to high trophic levels (Fréon et al., 2005).

SA population dynamic is very dependent of the environment (Coombs et al., 2006; Doray et al., 2018a; Erauskin-Extramiana et al., 2019; Schickele et al., 2020; Fernández-Corredor et al., 2021). SA seem to be influenced by the local hydrological conditions of the Bay (Doray et al., 2018a). Studies showed the importance of river flows, bottom temperature, chlorophyll-a and mesozooplankton biomass in the dynamics of the Bay of Biscay pelagic ecosystem (Doray et al.,

2018b; Grandremy et al., 2022). The distribution of mesozooplankton has been highlighted as a factor in the spatial organisation of SA (Petitgas et al., 2006; Grandremy et al., 2022).

Species distribution modelling and data integration

Nowadays, species distribution models (SDM) are widely used in ecology for management and conservation purposes in terrestrial, marine and freshwater areas (Guisan and Thuiller, 2005; Phillips et al., 2006; Elith and Leathwick, 2009; Pennino et al., 2019; Martin Gonzalez et al., 2021). The growing interest in SDM stems from their ability to predict the distribution of species over large areas based on habitat description data (e.g. environmental and prey data) and species occurrence records (Giannoulaki et al., 2017; Pennino et al., 2019). Understanding the relative contribution of spatial and temporal components in the distribution of fish has direct implications for the implementation of spatially explicit management objectives, in particular by-catch mitigation (Martin Gonzalez et al., 2021). Model-based approaches to analyse ecosystemic survey data (Doray et al., 2018b) have become popular as they allow the construction of detailed maps of species distribution and density (Conn et al., 2017; Isaac et al., 2020). These techniques have already proven their effectiveness for several marine species as predicting cetacean distributions (Giannoulaki et al., 2017) or SA distribution modelling (Andrews et al., 2020; Schickele et al., 2020).

Data from scientific surveys are commonly used in SDM to understand the spatio-temporal distribution of species. Data from ecosystemic surveys have the advantage of being derived from controlled and repeated sampling schemes (Rufener, 2020). However, survey data have limitations: they are expensive and collected over a limited spatial and temporal scales (Rufener, 2020). Commercial fishing data are therefore increasingly used to complement scientific survey data and provide fine-scale information over the year to map the distribution of fish (Moriarty et al., 2020; Rufener, 2020; Martin Gonzalez et al., 2021; Alglave et al., 2023). But fishery-dependent data also present some limits (Maunder et al., 2006) fishing declarations can be erroneous, Catch Per Unit Effort (CPUEs) are not necessarily proportional to fish biomass, due to catchability or fishing behavior issues. The combination of the different data sources presents several advantages and disadvantages for modelling fish distribution (Martin Gonzalez et al., 2021). The first problem is the diversity of methods used to collect data (Moriarty et al., 2020). With the expansion of the types and amount of biodiversity data collected, there is a need to find ways to combine these different sources to provide consistent summaries of potential and realised distributions of species in space and time (Isaac et al., 2020). Pinto et al. (2018) proposed a spatial model based on occurrence (presence/absence) by combining surveys and commercial fishing data. This is a simplification to limit the effect of heterogeneity in fishing effort and catchability between different data sources and allows all sampling methods to be considered equally informative about the presence/absence of the species (Pinto et al., 2018; Martin Gonzalez et al., 2021).

Another issue are data derived from opportunistic sampling (e.g. tourist whale or bird watching, fishing data), where observers tend to look for a specific specie in areas where they expect to find it (Hefley and Hooten, 2016; Post van der Burg et al., 2020). These data are usually subject to preferential sampling (Diggle et al., 2010; Pennino et al., 2019; Post van der Burg et al., 2020;

Alglave et al., 2023). Conventional geostatistical methods assume that sampling is non-preferential (Diggle et al., 2010; Pennino et al., 2019). Diggle et al (2010) demonstrate that ignoring preferential sampling can lead to biased inferences. In the context of multiple data integration, given the preferential sampling of fishermen, hierarchical models are a particularly suitable modelling tool (Conn et al., 2017; Archambault et al., 2018; Post van der Burg et al., 2020; Rufener, 2020; Alglave et al., 2023).

The objective of this work was to develop a Species Distribution Model (SDM) for small pelagic species, such as anchovy or sardine in the Bay of Biscay. The SDM integrates biomass data from pelagic scientific surveys PELGAS (springtime, Doray et al., 2018b) and JUVENA (autumn, Boyra et al., 2020) informing on fish density on one hand, and French fishing data (SACROIS data combining VMS and logbooks) informing on the presence of fish all year round, on the other hand. In the framework of the Delmoges project, this SDM aims at modelling the spatio-temporal distribution of the common dolphin's main prey, the SPF in the Bay of Biscay, at seasonal and annual scales. The idea is to produce new ecological and fisheries knowledge to better understand the determinants of incidental catches (Delmoges, 2021). The SDM outputs should ultimately be used as covariate to model the dynamics of common dolphin distributions, taking into account physical and trophic conditions (Delmoges, 2021).

Quemper, (2021) developed an integrated modelling framework described by Alglave et al., (2022) map the distribution of sardine in the BoB, and investigate their preferential sampling defined by Diggle et al., (2010) by French commercial fishing fleets.

In this master thesis, we rely on and significantly extend the framework developed by Alglave et al., (2022) and Quemper et al (2021).

Quemper (2021) showed that taking into account preferential sampling in the inference had little influence on the inferred distribution of sardine presence (Quemper, 2021). We then analyzed commercial fisheries data as presence-only data with a Poisson point process and included a shared common Gaussian random field with survey data. This approach avoided biased inference when data were collected under preferential sampling (Diggle et al., 2010).

Additionally, the inclusion of environmental covariates in Quemper et al (2021)'s model did not reveal any significant effects of the covariates on the probability of sardine presence in the Bay of Biscay. Moreover, the SACROIS fishing data were subject to measurement error, due to the homogeneous reallocation of catches along the fishing path, which smooths and attenuates the signal, blurring correlations with environmental covariates (Quemper, 2021). Then our model only included the distance to the coast as environmental covariate, to explain and/or predict the distribution of SA in the BoB. The distance of commercial fishing operations to the coast was included in our model to take into account factors that explain the distribution of fishing operations, e.g. fuel cost.

Finally, in Quemper et al (2021) both fishery and survey data were integrated as presence absence processes, which did not properly account for the fishing process and reporting (only positive catch of marketable fish are reported), and did not allow for the use of accurate fish biomass (density)

estimates provided by scientific surveys. Our model extended Quemper's (2021) approach by modelling three processes relying on three types of data within a single hierarchical framework: presence-absence from the survey, intensity (biomass) from the survey and presence-only from the commercial data.

2. Materials and Methods

Study area

The study zone is the Bay of Biscay (BoB), an open oceanic bay located in the Northeast of the Atlantic Ocean between 48°5 and 43°5'N and 8 and 3'W (Lassalle et al., 2013). It covers an area of approximately 225,000 km² (Persohn, 2009). The continental shelf along the Spanish coast is narrow (~ 30 km) and it widens northward along the French coast, reaching 180 km off Brittany (Costoya et al., 2015). The limit of the plateau corresponds to the isobath of 200 m (Persohn, 2009).

The water masses in the upper layers (from 100 to 600 m depth) have temperature varying between 10.5 and 12 °C and salinity from 35.45 to 35.75 p.s.u (Koutsikopoulos and Cann, 1996), with sea surface temperature and salinity higher in the south of the Bay of Biscay (Persohn, 2009). Along the French continental shelf, freshwater input induces density gradients that favour a poleward circulation and carries essential minerals for the ecosystem chain. The Loire and Gironde rivers account for 75% of freshwater inputs into the Bay of Biscay (Persohn, 2009; Costoya et al., 2015).

This study is limited to the French continental shelf of the BoB (Figure 1). This area corresponds to the area sampled by the PELGAS scientific survey. It includes the continental shelf located between the Pointe du Raz and the Gouf de Capbreton.

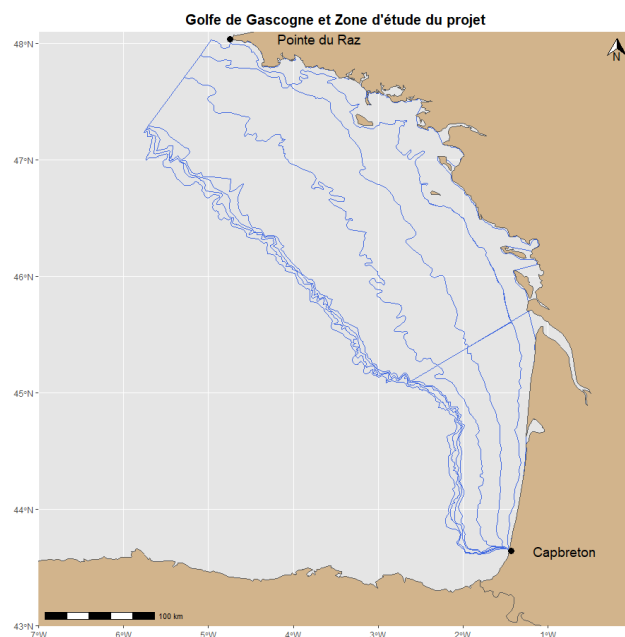


Figure 1: Bay of Biscay and study area

Survey data

Data from two integrated scientific surveys of the Bay of Biscay (BoB) pelagic ecosystem were used in this study: the PELGAS survey in May and JUVENA in September. One of the goal of these surveys is to assess the abundance, biomass, and age and size structure, of the BoB small pelagic fish community using an acoustic-trawl methodology (Doray et al., 2021). The PELGAS survey has been conducted by the Institut français de recherche pour l'exploitation de la mer (Ifremer), in collaboration with La Rochelle University and the Centre national de la recherche scientifique (CNRS) since 2000 over the French shelf of the BoB (Doray et al., 2018b). The JUVENA survey has been taking place since 2003. It has a focus on anchovy sub population and covers the Spanish and French coasts (Boyra et al., 2020). Only samples located in the study area were used in the analysis (i.e. samples located in the BoB continental shelf) (Figure 2).

PELGAS and JUVENA are acoustic surveys that use an echo sounder to emit short electrical pulses, transmitted in seawater as ultrasonic pulses, from the vessel hull to the seabed. In the presence of fish, a part of the acoustic signal is backscattered and recorded by the echosounder. The majority of pelagic fish being often out of reach of the sounder (in a layer of water between the surface and 10 meters of immersion) at night, acoustic surveys are carried out during daytime. Echo-integration makes it possible to evaluate the biomass present in an area by integrating the acoustic energy backscattered by all the fish targets insonified (Simmonds and MacLennan, 2007). These echo-integrations are carried out on the scale of an EDSU (Elementary Distance Sampling Unit), which corresponds to one nautical mile of acoustic linear sampling (Doray et al., 2018b). The acoustic densities, resulting from echo-integration, were backscattered by an assemblage of species. To determine the species and size composition of each EDSU, pelagic trawls are carried out regularly (2 to 3 trawls per day) and the species and size composition of the trawls are reallocated to the echo-integrals by EDSU (Doray et al., 2021).

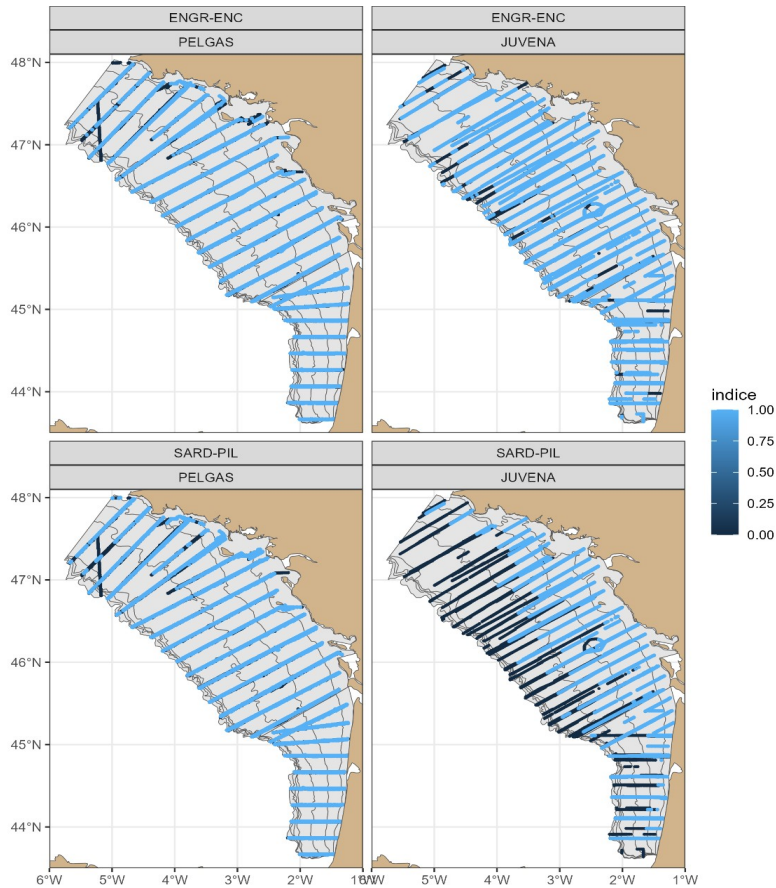


Figure 2: Spatial distribution of presence/absence per survey (JUVENA/PELGAS) and species (anchovy-ENGR-ENC / sardine-SARD-PIL) of all the year 2009 – 2022

Because PELGAS and JUVENA are scientific survey, they benefit from a standardized sampling plan. We then considered they provided an unbiased information on the spatio-temporal distribution of the species biomass. Then for PELGAS and JUVENA data on absence/presence and biomass (in tons) per EDSU were used in the model from 2009 to 2022 (Appendix 1).

Commercial fishing data

Sardine and anchovy fisheries, and VMS data

The fishing fleet targeting small pelagics in the Bay of Biscay brings together vessels with a variety of profiles, both in terms of technical characteristics (size, gear used, etc.) and their portfolio of targeted species or fishing grounds (Lahellec, 2020). This fleet is classically segmented into two groups: pelagic purse seiners and pelagic trawlers. They target the shoals they want to fish using sonar (Lahellec, 2020). If there is no fishing, it does not mean that there are no fish, but rather that there is a problem with the gear, or that the fish were not big enough and were thrown back. Additionally, the fisheries is driven by the market, because fisheries often have contracts with canneries that specify the size of fish and the tonnage they must catch. So we decided to not use the

absence data, but only presence data. Fishery catches reflects thus only a small part of the total population. And the fisheries don't have to declare discards in the case of sardines because there is no TAC for the stock.

Those issues specific to the BoB pelagic fisheries adds to classical CPUE biases. Due to the aggregation behaviour of pelagic fish, pelagic CPUEs remain relatively stable whatever the underlying biomass level (this phenomenon is also known as CPUE hyper-stability) (Pitcher, 1995). Moreover, fishery-dependent data can confound changes in fishing behavior with changes in fish abundance. This is because fish behave in more or less gregarious ways, during spawning or feeding phases for example, and some species travel in large, dense schools. It is also because of the preferential sampling (Diggle et al., 2010; Pennino et al., 2019; Post van der Burg et al., 2020; Alglave et al., 2023). Fishing operations are hence only performed in areas where fishermen expect to find the species of interest. Fishing absence data being not reliable, and CPUEs being not proportional to the biomass of small pelagic fish, we decided to use only presence data from the fishery.

The Vessel Monitoring System (VMS) monitors and registers with high accuracy (but low frequency) the geographical locations at sea of equipped fishing vessels (Phillip and Robert, 1998). VMS monitoring is mandatory for professional fishing vessels over 12 meters, flying the flag of Member States of the European Union, since January 1, 2012. VMS data allow the monitoring of fishing activity at a very fine spatio-temporal resolution (Alglave et al., 2022). These spatialized informations are coupled with landing data (logbooks) according to the methodology described by Hintzen et al. (2012). It estimates georeferenced catches per species by reassigning landings to the vessel positions meeting a set of conditions. The VMS positions are hence filtered to keep only the one more likely corresponding to fishing activity (mean speed inferior to 4.5 knots- algorithm AlgoPesca developed by Ifremer). The reassignment is uniform, this method cannot provide information on irregular catches. This study used the output of the algorithm SACROIS developed by the Ifremer based on this methodology.

Aggregating the data per season and matching the surveys timing

Only three seasons (spring, summer, fall) were defined in the model. To choose which month will be put in each season, we used the surveys as reference. PELGAS was mostly conducted in May, but started sometimes in April. In the study, spring hence corresponded to the months of May and April. The main Month for JUVENA is September but it can finish in October. We decided to put in fall the months of September and October. For summer, we just took the 3 months between the spring and fall months defined before: June, July and August, which also represented the months with highest commercial fisheries data.

Data filtering and Preliminary analysis

Commercial data include all sardine and anchovy landings declared by fishermen. An exploratory analysis of catches, effort and CPUE was carried out over the period 2009 – 2022. We produced

mean maps per year or month and gear, effort and Catch Per Unit Effort (CPUE) and compared with known information. A few suspicious data points have comforted the choice made in the work of Quemper, (2021) to keep only gears known to target sardines. In the study of Quemper, (2021), only the fleets that have more than one per cent of their captures constituted by sardines were kept. In 2021, about 86.7% of French sardine catches were made by purse-seiners while the remaining 13% was reported by pelagic pair-trawlers (ICES, 2022). Pelagic pair trawls (PTM) and purse seines (PS) have hence been retained for analysis (Figure 3, Appendix 1).

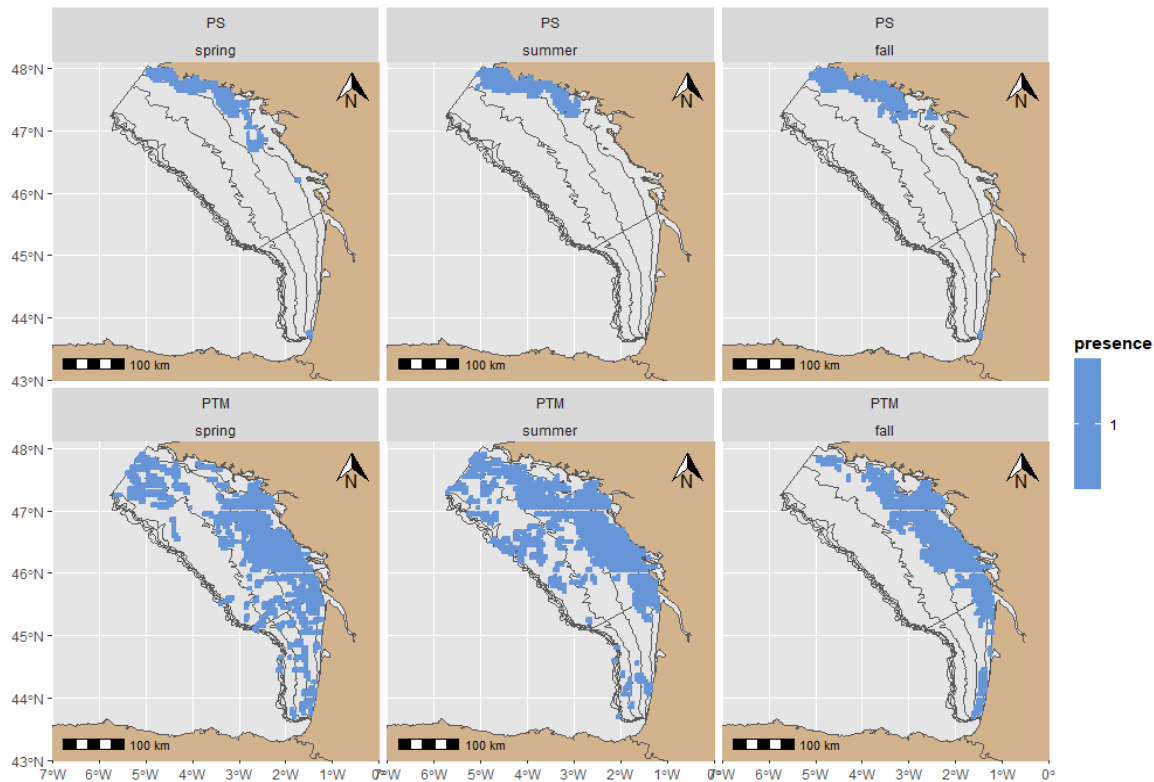


Figure 3: Spatial distribution of presence commercial data per fishing fleet and per season over the period 2009 – 2022

To confirm the absence of correlation between CPUE from commercial fleets and indices of abundance as biomass (Rufener, 2020), the correlation between CPUE and survey biomass was first explored. The commercial data of May and September were extracted and compared respectively with PELGAS and JUVENA data. All data were averaged on the same spatial grid to allow for their comparison. Fishing and survey data were merged by grid cell and year, and biomass vs CPUE scatter plots were produced.

Hierarchical Modelling

Spatio-temporal resolution of the model

Concerning the temporal resolution of the model, years and seasons were defined as integers: {2009, 2010, ..., 2022}, $k = 1$ for spring, $k = 2$ for summer, $k = 3$ for fall.

To define the spatial resolution of the model, we adopted the SPDE (Stochastic Partial Differential Equation) spatial framework which represents continuous Gaussian fields as a discrete Gaussian Markov random field (Lindgren, 2012). The number of knots determines the spatial resolution of the model (i.e resolution of the inference of the random fields). We used a k-means algorithm applied on samples locations to identify the location of knots (Figure 5). The SPDE approximation involves generating a triangulated mesh that has a vertex of a triangle at each knot using *R-INLA* (Lindgren, 2012)). Different mesh designs were compared visually and in terms of computing time. A single mesh has been kept for all the models. The mesh design includes an outer extension to avoid a “boundary effect” and regularly shaped triangles, both in the inner and outer extensions and at the border between the two extensions (Pinto et al., 2018). Then spatial variables at location were interpolated from knots to extrapolation grid using this triangulated mesh. (Figure 4, Figure 5)

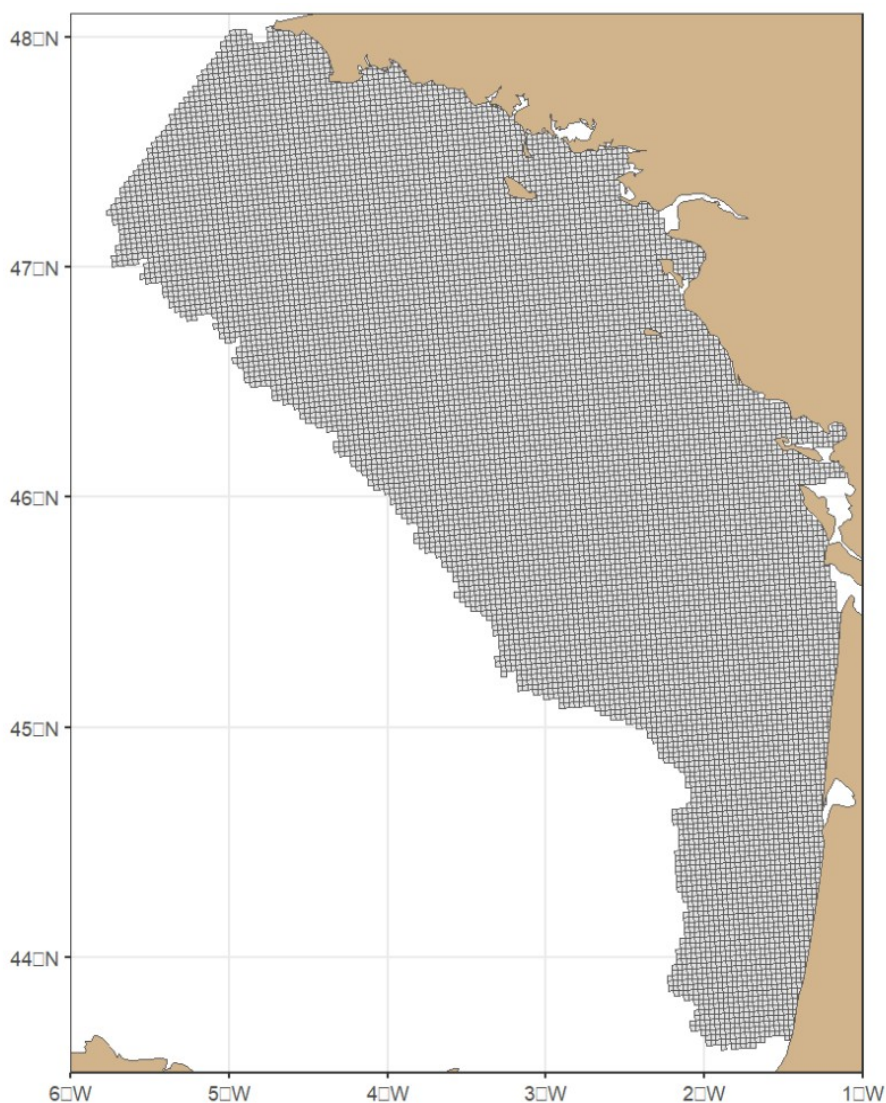


Figure 4: Extrapolation grid of the study area (1 cell = 2,5 x 2,5 km)

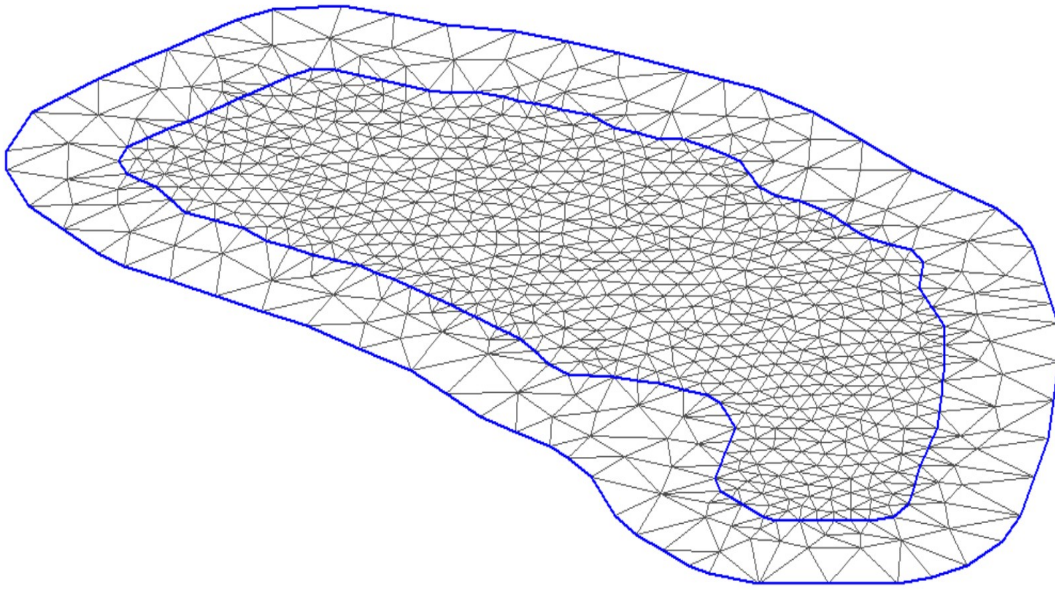


Figure 5: Triangulated mesh of the study area

The SPDE approach allows a Gaussian field to be approximated by a Gauss-Markov field (when the Gaussian field admits a Matérn function as its correlation function). This is achieved by reducing the spatial resolution at which the spatial structure of the latent field is estimated by modelling it on the scale of a sparse triangular mesh (a triangle has three nodes). At node level, the latent field is modelled explicitly. The finer the mesh resolution, the more accurate the approximation but the larger the number of nodes to be estimated and the longer the calculation time. Between the nodes, the value of the latent field will be a linear interpolation of the value of the latent field at each vertex of the triangle in which the observation is located. So the value of the latent field at an observation point depends not just on the nearest node, but on the 3 vertices of the triangle in which the observation lies, weighted by the distance between the observation and the corresponding node. More details are available in Rue et al. (2009), Moraga (2020) or Krainski et al. (2021).

Observation Process and likelihoods

Sardines and anchovy data were analysed independently using the same statistical approach. The hierarchical model consists in 3 parts: observations, latent processes and hyper-parameters. The observation models assume different likelihoods (Poisson Point Process / LogNormal / Binomial) depending on the nature of data (presence only fisheries data / survey biomass / survey presence-absence respectively). These observations are linked in the hierarchical model through latent processes (using Gaussian fields), which account for spatial and temporal dependencies of ecological interest. Finally, both observation and latent processes depend on higher-level parameters called hyper-parameters: variances and spatial ranges in the present case.

The model integrates different data sources in a way that retains the strengths of each (Isaac et al., 2020). There are 3 sources of data and each type of observation have a different likelihood that accounts for their idiosyncrasies.

Scientific surveys provide information on presence/absence and intensity (biomass). Initially, a positive SA biomass is coded as 1 and its absence as 0. Intensity only concerns strictly positive data so is conditional on presence: the biomass takes on the associated value in tons/EDSU. For commercial data, only the coordinates where SA have been fished are available. A value of 1 is associated to all the points corresponding to positive catch and produce presence-only maps.

The notion of detectability in this model was not addressed because the data did not permit it. The biases of acoustic sampling cannot be taken into account (fish in the echosounder blind zone near sea surface), so we considered that the survey data have a 100% detectability. For commercial fishing data, we considered that all declarations were accurate and faithful.

Let s denotes spatial position (a pair of longitude and latitude) in the BoB; t denotes the year (between 2009 and 2022) and k denotes the season (1 = Spring, 2 = Summer and 3 = Autumn). Let $z(s_i, t_i, k_i)$ represents the i th observed presence of sardines in the surveys for location s_i , year t_i and season k_i . z follows a Binomial distribution with parameters π , n . π corresponds to the probability of presence of fish and n equals 1.

$$(1) \quad z(s_i, t_i, k_i) \sim \text{Binomial}(\pi_{s_i, t_i, k_i}, n_{s_i, t_i, k_i} = 1)$$

Let $y(s_i, t_i, k_i)$ represents the i th observed intensity of presence when z is 1. y follows a LogNormal distribution with parameter μ and σ . These are related to the mathematical mean expectation and standard deviation of the variable's natural logarithm.

$$(2) \quad y(s_i, t_i, k_i) \sim \text{Log-N}(\mu_{s_i, t_i, k_i}, \sigma^2) \text{ with } \sigma > 0$$

Let S_1, \dots, S_n represent the occurrence of SA as seen by commercial fishing. S_1, \dots, S_n follows a Poisson point process (PPP) with parameter λ (Hefley and Hooten, 2016; Moreira et al., 2023). λ is the intensity of the counting process, which is derived from a (stochastic) Gaussian Process. Since the PPP is a statistical distribution over a spatial domain, it allows to model points in geographical space. In a homogenous PPP, it is assumed that points locations are independent from each other. This assumption is usually not adequate for biological species and can be relaxed in inhomogeneous PPP. If the intensity parameter λ is modelled with a Gaussian Process (on a log scale), the resulting process is called a Cox process. The fishing data were therefore modelled by a Cox process (Opitz, 2016) written as:

$$(3) \quad \{S_1, S_2, \dots, S_n\} \sim \text{PPP}(\lambda_{s_i, t_i, k_i})$$

Where λ_{s_i, t_i, k_i} is the intensity of the counting process for location s_i , year t_i and season k_i .

Link functions and conditional expectations

For observation processes, the parameters defining the likelihood follow stochastic processes.

For the presence/absence process, the linear predictor is modelled as the *logit* of the probability π , specified as

$$(4) \quad \text{logit}(\pi_{s_i, t_i, k_i}) = \underbrace{\alpha^\pi}_{\text{Intercept}} + \underbrace{W_\pi(s_i, k_i)}_{\text{specific Gaussian field}} + \underbrace{\xi_0(s_i, t_i, k_i)}_{\text{shared Gaussian field}} + \underbrace{V(k_i)}_{\text{season effect}} + \underbrace{\sum f_n(C_n(s_i, t_i, k_i))}_{\text{environmental covariables}}$$

For the intensity process, the LogNormal distribution parametrized with parameters, $E(\ln(y_{(s_i, t_i, k_i)})) = \mu_{(s_i, t_i, k_i)} + 0.5\sigma^2$ and $\text{Var}(\ln(y_{(s_i, t_i, k_i)})) = \sqrt{\exp(\sigma^2 - 1) * \exp(2\mu_{(s_i, t_i, k_i)} + \sigma^2)}$. The linear predictor is equal to:

$$(5) \quad \log(\mu_{s_i, t_i, k_i}) = \underbrace{\alpha^\mu}_{\text{Intercept}} + \underbrace{W_\mu(s_i, k_i)}_{\text{specific Gaussian field}} + \underbrace{\xi_0(s_i, t_i, k_i)}_{\text{shared Gaussian field}} + \underbrace{V(k_i)}_{\text{season effect}} + \underbrace{\sum f_n(C_n(s_i, t_i, k_i))}_{\text{environmental covariables}}$$

And for the presence-only process, the intensity of the Cox process is defined as

$$(6) \quad \log(\lambda_{s_i, t_i, k_i}) = \underbrace{\alpha^\lambda}_{\text{Intercept}} + \underbrace{W_\lambda(s_i, k_i)}_{\text{specific Gaussian field}} + \underbrace{\xi_0(s_i, t_i, k_i)}_{\text{shared Gaussian field}} + \underbrace{V(k_i)}_{\text{season effect}} + \underbrace{\sum f_n(C_n(s_i, t_i, k_i))}_{\text{environmental covariables}}$$

α^π , α^μ , α^λ are intercepts that represent the average effect in space and time (season and year) for the presence absence, the intensity and the presence only processes, respectively. $V(k)$ is a seasonal specific fixed effect. $W_\pi(s_i, k_i)$, $W_\mu(s_i, k_i)$, $W_\lambda(s_i, k_i)$ are spatial-seasonal terms that represent the unmeasured seasonal spatial variation in the presence/absence, the intensity and the presence only processes respectively. They are latent processes and are modelled using the SPDE approach (Lindgren et al., 2011). The effect is a zero-mean Gaussian random field whose covariance matrix follows a Matérn correlation function characterised by a range parameter r - the distance at which the correlation between two points is equal to 0.1, and a variance parameter σ^2 . This field reflects the SA distribution pattern for a season. It would give the preferential seasonal distribution given by this observation process.

ξ_0 is the spatio-temporal random effect that integrates seasons and years. It is a Gaussian random field shared between the 3 processes. In other words, each observation, whether absence/presence, positive biomass, or commercial fishing positive operation, is the realization of a stochastic process governed in part by this common latent process. The associated temporal correlation is modelled using an autoregressive process of order 1 parametrised by a parameter a . Finally C represents the environmental co-variables and γ is a fixed effect parameter which captures the species-habitat relationship with the function f_n modelling the species-habitat relationship. Here bathymetry, distance to coast, SST and Chl-a were selected as covariates to model species habitats.

Seasonal effect and environmental covariates

The season effect Vk_i is modelled as an intercept. In other words, it is an average pattern associated to the season which is added to α^π , α^μ , α^λ intercepts. It represents the difference between the spring,

the summer and the autumn intercepts. This seasonal effect may be due to the environment, but also differences between sampling from one season to the next. So a catchability effect that is not represented anywhere else in the model would be confounded with the season effect. Thus the season effect should not necessarily be interpreted as a biological effect.

The species-habitat relationship f_n in our model depends on space, year and season, as we postulate non-stationarity of the environmental effects. The intensity of the relationship, included in f_n , may change over the seasons. In other words, the process linking the distribution of fish and their habitat is the same, but the parameters may change over the seasons, particularly in relation to the species' biological cycle. Namely, requirements for reproduction are not the same as those for growth and feeding, and fish move to find optimal habitats along their life cycle.

Priors

We used the default priors for the hyper-parameters as implemented in R-INLA. Defining adequate default priors for hyper-parameters currently constitute an active area of research for the R-INLA team (Pinto et al., 2018). The shared latent field parameters specify that across time, the process evolves according to an auto-regressive AR(1) process. The prior for the autocorrelation parameter a is defined by a Penalised Complexity or PC prior for the autocorrelation parameter a where $a=1$ in the base model. Here we assume $P(a>0)=0.9$ (Krainski et al., 2021). For the hyper-parameters governing the Matern covariance function, PC priors were also used: the prior for the sill variance was set so that values greater than a factor of 5 had a probability of 0.01. The range indicates the minimum value at which we have spatial information. As the PELGAS data maximum spatial resolution was 1 nautical mile (1.852 km) the range was set to: value of 1.8/3, i.e. 0.6 km.

For the fixed effect, we split in 2 parts. For slope parameters (corresponding to the covariates) we put a prior such that the effect is between 0.2 (1/5) and 5 times the value of the corresponding intercept. We are working on a log scale so the mean is 0 and the precision (inverse of the variance) is $1/\left(\frac{\log(5)}{2}\right)^2$. For the season intercepts, we put a prior for LogNormal/PPP intercept parameters such that the effect is between 0.1 (1/10) and 10 for Binomial intercept parameters, insuring that the mean presence probability is between 0.1 and 0.9 with the possibility to have more extreme values. So the prior have a mean of 0 and a precision of $1/\left(\frac{\log(10)}{2}\right)^2$. These priors were weakly-informative in that they are putting some constraints on the parameter space and favoring plausible regions a priori. The shape of the chosen prior are depicted in the Appendix 2.

Inference tool INLA

The inference was performed in a Bayesian framework using the R-INLA tool. Rue et al., (2009) have developed the integrated nested Laplace approximation (INLA) to approximate the joint posterior distribution as an alternative to traditional Markov chain Monte Carlo methods (Moraga, 2020). INLA allows to perform approximate Bayesian inference in latent Gaussian models such as generalized linear mixed models and spatial and spatio-temporal models (Moraga, 2020). The aim

of the INLA methodology is to approximate the posterior marginals of the model effects and hyper-parameters. This is achieved by exploiting the computational properties of Gaussian Markov random fields and the Laplace approximation for multidimensional integration (Krainski et al., 2021).

R-INLA combines the inference by the integrated nested Laplace approximation with the SPDE (Stochastic Partial Differential Equations) approach.

Convergence assessment

For each model in this study, the Kullback-Leibler divergence and the posterior marginal distributions of the intercepts and the Gaussian field parameters were checked to assess the goodness of fit. The Kullback-Leibler divergence (kld) describes the difference between the Gaussian approximation and the simplified Laplace approximation for each model. If the model is well fitted, the kld vector is composed entirely of zeroes for the fixed effect.

The posterior marginal distribution of a parameter should be similar to a Gaussian distribution if the model fit is good.

The two goodness of fit indices provided by the INLA were also checked (see details in Appendix 3).

Upon convergence, the model goodness of fit was finally checked by mapping spatialized deviance residuals. Deviance residuals can be interpreted as the difference between your model's fit and the fit of an ideal model (Rundel, 2017). Deviance is a measure of goodness of fit in a similar way to the residual sum of squares (Rundel, 2017). Deviance residuals maps were checked for non-random patterns that may betray the presence of an effect that has not been taken into account, and that has significantly influenced the model's deviance residuals (Appendix 4).

Model selection

The criterion use to compare models was the Watanabe-Akaike information criterion (WAIC). WAIC is a fully Bayesian approach for estimating the out-of-sample expectation, starting with the computed log point wise posterior predictive density and then adding a correction for effective number of parameters to adjust for overfitting (Gelman et al., 2014).

Modelling strategy

The first model created was the Hurdle model on PELGAS data. A hurdle model is a class of statistical models where a random variable is modelled using two parts: the first is the probability of attaining value 0, and the second part models the intensity of the non-zero values. This model was performed using a Binomial likelihood for the presence/absence and a LogNormal for the intensity of presence. This model was a first approach to an integrated model for one season, fitted on several years.

We then fitted a model per season. k was set to 1 in the equation (4,5,6) and one model was fitted for spring using PELGAS and fishing data, and another for autumn, using JUVENA and fishing data. These two seasonal models were run for sardine and anchovy. The shared latent field parameters was modelled in each seasonal model as an auto-regressive AR(1) process representing year-to-year autocorrelation. The seasonal models allowed for testing the integration of survey and fishing data, as well as a first calibration of the PPP that demands a specific syntax in R-INLA.

The complete model included all 3 processes, all seasons and all years.

The inclusion of environmental covariates (bathymetry and satellite Sea Surface Temperature and surface Chlorophyll-a) in Quemper (2021)'s model did not improve the predictions. Due to time constraints, we then chose to just test the effect of distance from coast in our complete models, as a proxy for fuel consumption. Because of the independence of each year, the effect of the covariate can varies between year.

The complete model should have been fitted for all 14 years with three season at once. The shared latent field parameters should have been modelled as an auto-regressive AR(1) process representing season-to-season autocorrelation.

Unexpected incompatibility with R-INLA prevented the use of super computers to fit the complete model over 14 years. Personal computers only allowed to fit the complete model for a single year. A total of 14 annual models (with t set to 1) with 3 seasons were then fitted for sardine. The specific latent fields were hence independent between seasons and years for the 14 different models. The shared latent field parameters were modelled as an auto-regressive AR(1) process between seasons. All together, the outputs of the 14 annual models approximated the results of a single model fitted over 14 years at once. Fitting annual models however implied to assume that the sardine distribution was not stationary between years.

Due to time constraint, models could not be fitted on anchovy data.

Derived quantities

To summarise the information provided by the model, we have calculated the average shared Gaussian random field (SGF) and its inter-annual variability. For each grid cell and season, average and standard deviation (SD) of SGF values and prediction errors were calculated.

Our model produces 3 types of predictions: the probability of presence π , the biomass intensity μ and the catches fisheries densities λ . Predictions were obtained in each grid cell, season and year as the sum of an intercept and its variability V_{k_i} , a covariate effect, a process specific Gaussian random field and a shared Gaussian random field. Then, average and SD maps of predictions and predictions errors have been derived as for the SGF.

Average maps were analysed to identify mean patterns in SGF, predictions, and errors, SD maps were analysed to assess the inter-annual variability of each quantity in each grid cell.

3. Results

CPUEs and survey biomass comparison

All gear together, there was no significant correlation (Pearson correlation coefficient: 0,01 in spring, 0.14 in autumn) between sardine CPUEs and survey biomass estimates in May and September over the series. Most data were zeroes, with few large survey biomass values associated with low CPUEs, and high CPUEs associated with low survey biomass (Figure 6). We found the same pattern for each gear separately.

These small correlations confirmed that, for sardines, CPUE can not be used as an index of abundance.

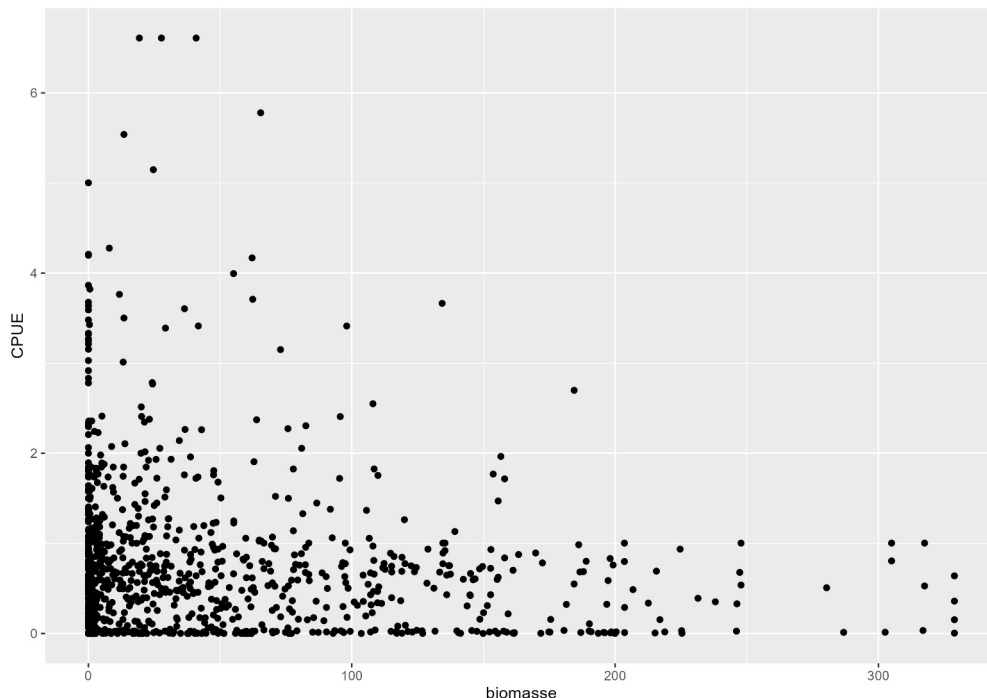


Figure 6: CPUE of September (t/h) by biomass of Juvena (t) for the sardine

All gear together, there was no significant correlation between anchovy CPUE and the PELGAS data (Pearson correlation coefficient = 0,24). However a significant positive correlation (Pearson correlation coefficient = 0,71) was found between purse seiners CPUEs and PELGAS survey biomass in May for of anchovy (Figure 7). The Pearson correlation coefficient associated was 0,71. There was also no significant correlation between anchovy CPUE and the JUVENA data (Pearson correlation coefficient = - 0,24) (Figure 8). As for sardine, most data were zeroes, with few large survey biomass values associated with low CPUEs, and high CPUEs associated with low survey biomass.

For the anchovy, the commercial fishing data were used as presence-only data like for the sardine.

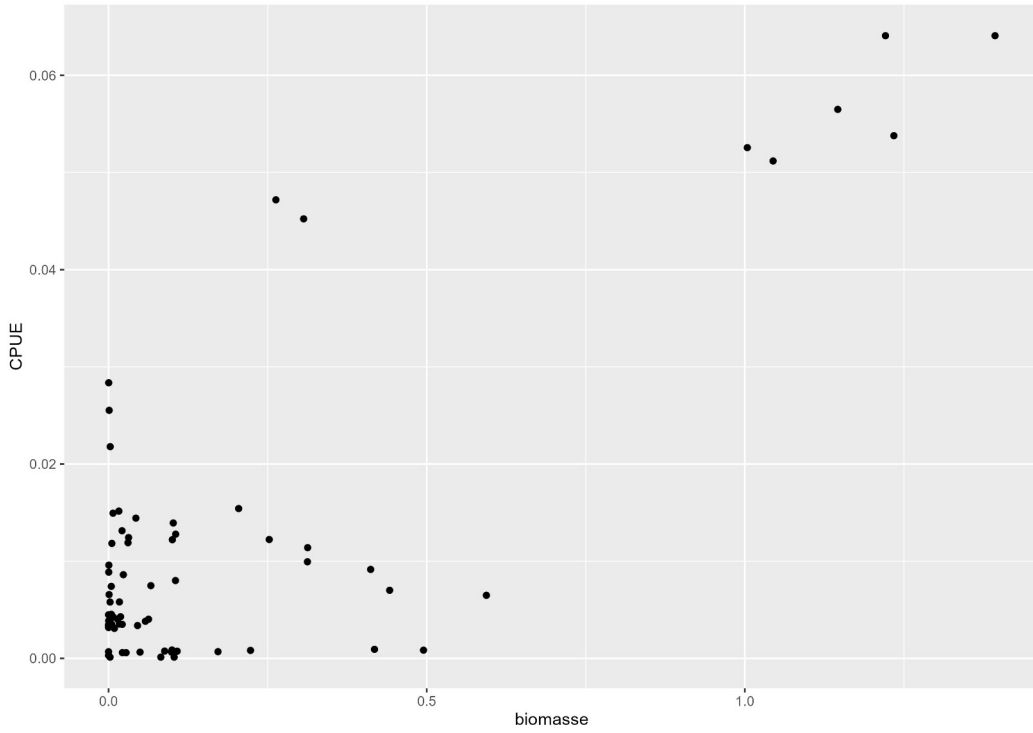


Figure 7: CPUE of May of PS (t/h) by the biomass of Pelgas (t) for anchovy

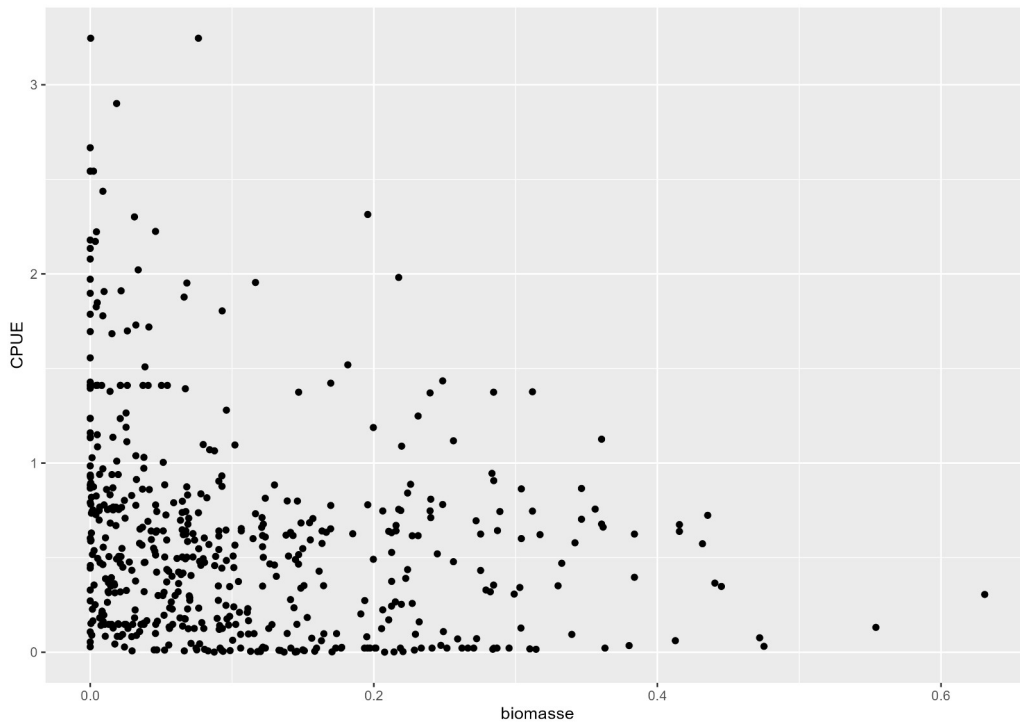


Figure 8: CPUE of September (t/h) by the biomass of Juvena (t) for anchovy

Multi-season modelling

Convergence of models

The indicator of convergence was true for all the models. The number of checks not passed during the posterior mode optimization for the hyper-parameters was small and did not exceed 0,5 % of the total effective parameters (Appendix 6). For all the fixed effect in the models, the kld was close to 0, suggesting no convergence issue (Appendix 6). The models were run with and without the covariate (distance from the coast). There were some differences in the intercept (Figure 9). The difference in WAIC between models with and without covariate varied a lot across years (appendix). For 7 years, the best model was the one with covariate and for 7 years the best model was the one without covariate. Finally, we kept the covariate in the models and all the results in the followed section were obtained with models including the distance to coast as covariate.

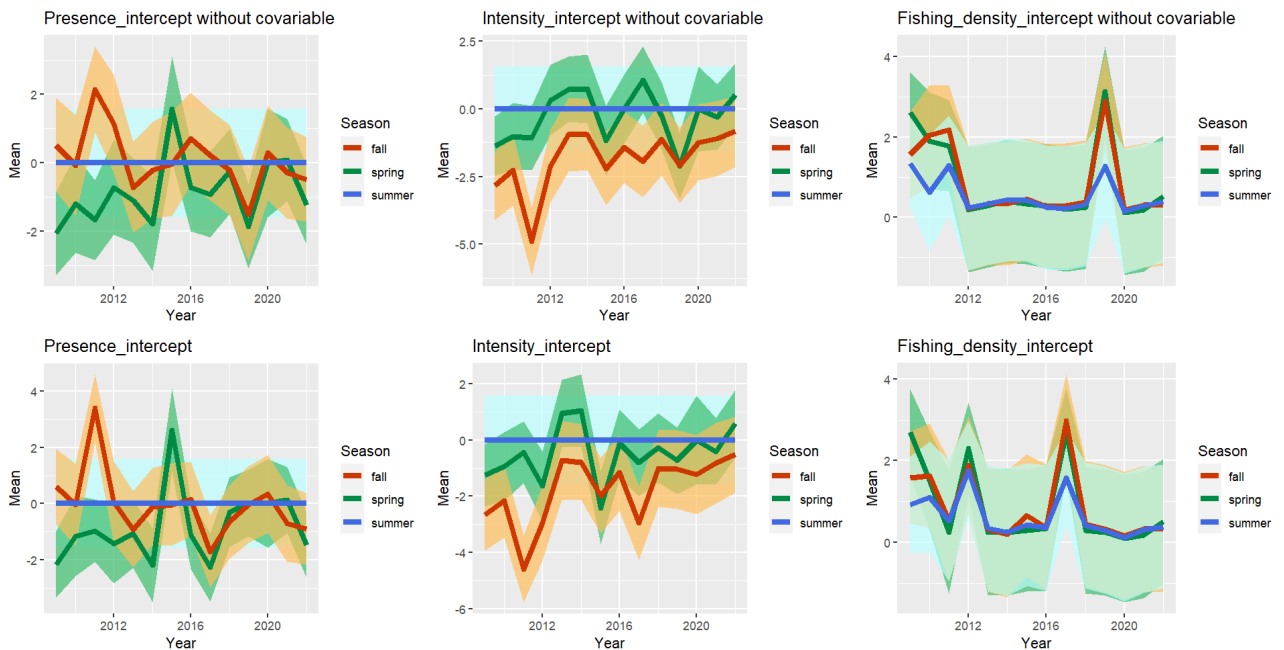


Figure 9: Evolution of the intercepts of the models with covariate and models without covariate over the series (spring in green, summer in blue and fall in red)

Shared Gaussian Random Field

The shared Gaussian random field (SGF) integrated information from all three likelihood processes. It represents the common spatial-temporal variability of the distribution of the sardine in the BoB that was not explain by the other effects. The SGF compiles all the information provided by all the different types of data available in a single map per season.

The posterior estimates of the common spatial random effect revealed a similar, globally coastal distribution for sardine in all seasons (Figure 10). Fish concentrations were found between the Gironde and the Loire estuaries, and in South West Brittany. Higher fish densities were observed in summer than spring and fall. The inter-annual variability followed a positive South-East to North-West gradient in all seasons. The average estimation error displayed an offshore-inshore gradient with higher error values in summer and fall offshore (Figure 10).

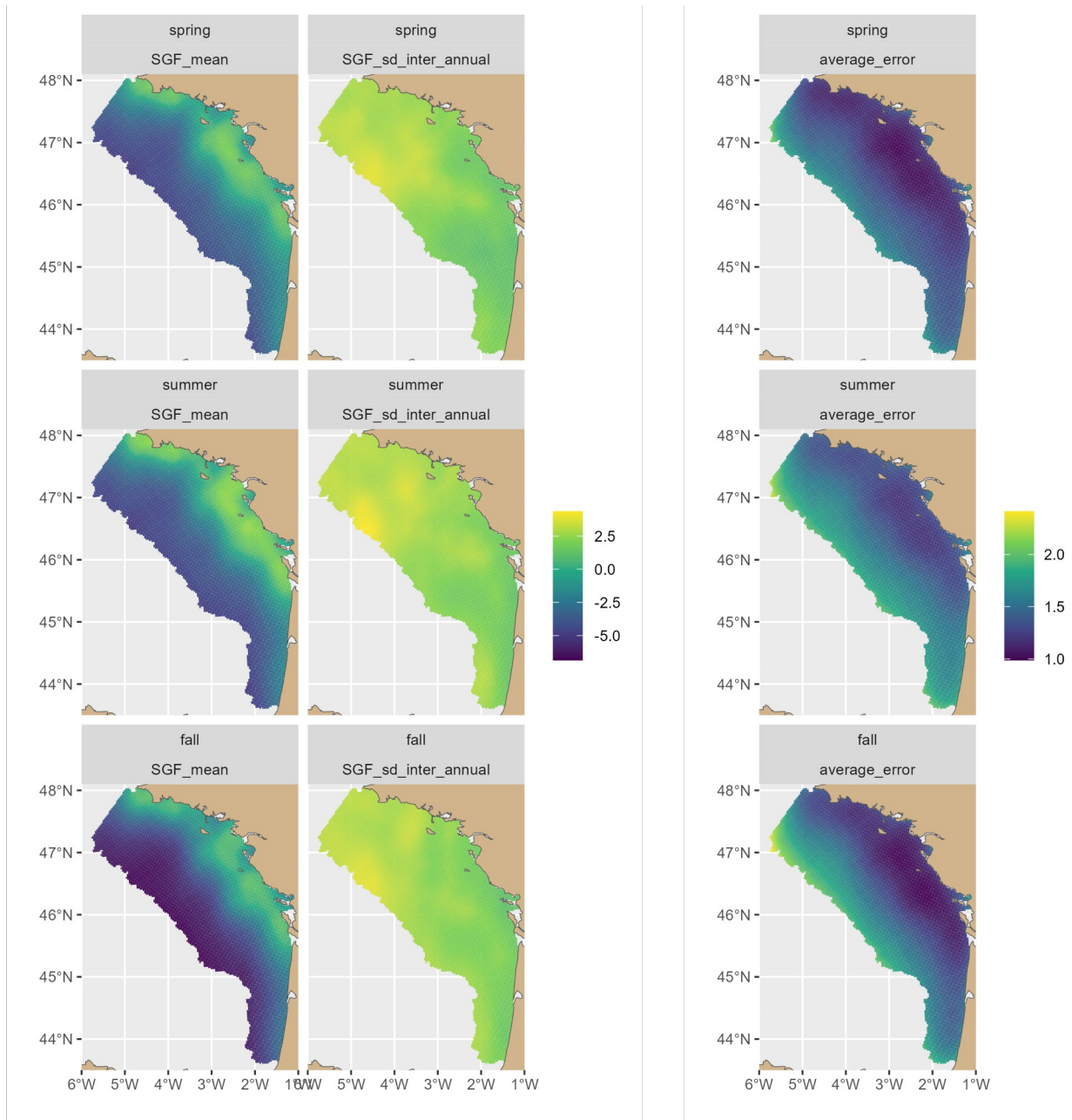


Figure 10: Average posterior mean, standard deviation and average error of the common spatial random effect for all the years in logarithm scale

Even if the average SGF did not vary much across seasons, we have identified 3 years with inter-annual and seasonal variability: 2015, 2017 and 2022 (Figure 11). Year 2015 was characterised by the presence of a sardine hotspot near the shelf break in NW BoB in spring and summer (Figure 11).

Fish density globally decreased from spring to fall. The disappearance of a fish hotspot located in Southern BoB from spring to fall suggested a northward migration. Conversely, a similar southern fish concentration appeared in fall in 2017 suggesting a southward migration. Beside this, year 2017 GRF map was similar to the average map. Contrary to other years, fish seemed to concentrate from spring to fall in 2022. The estimation error was lower in spring (2015, 2022) and higher in summer, with a negative offshore-inshore gradient. This gradient was particularly marked in 2017, with the lowest estimation errors near the coast, and highest offshore.

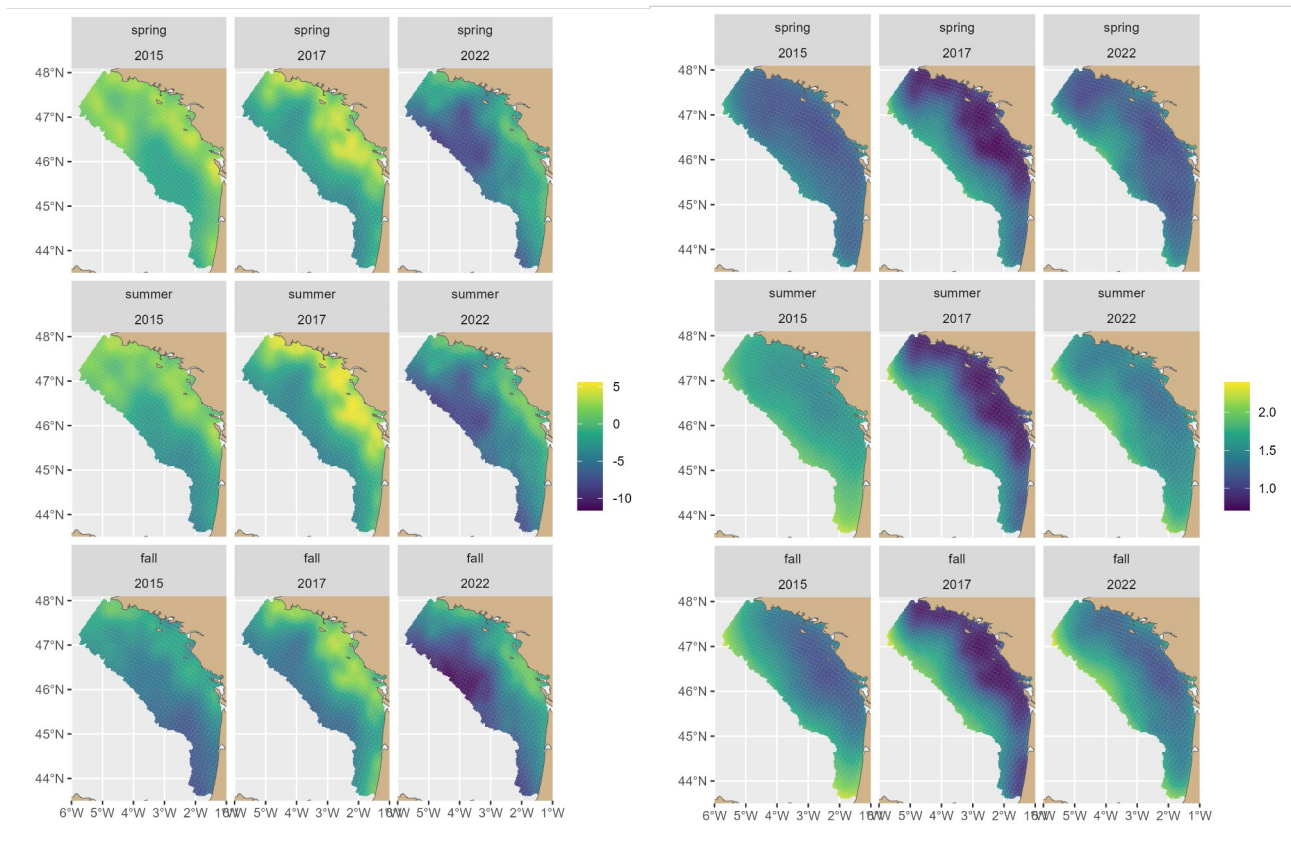


Figure 11: Posterior mean and average error of the common spatial random effect for 2015, 2017, 2022

In addition to inter-season variability, SGF displayed large inter-annual variability within each season (Figure 12) throughout the BoB. Despite a common coastal pattern, the intensity of presence varied from year to year (Figure 12). No clear trend was visible, except a slight overall decrease of SGF intensity in autumn since 2018 (Figure 12).

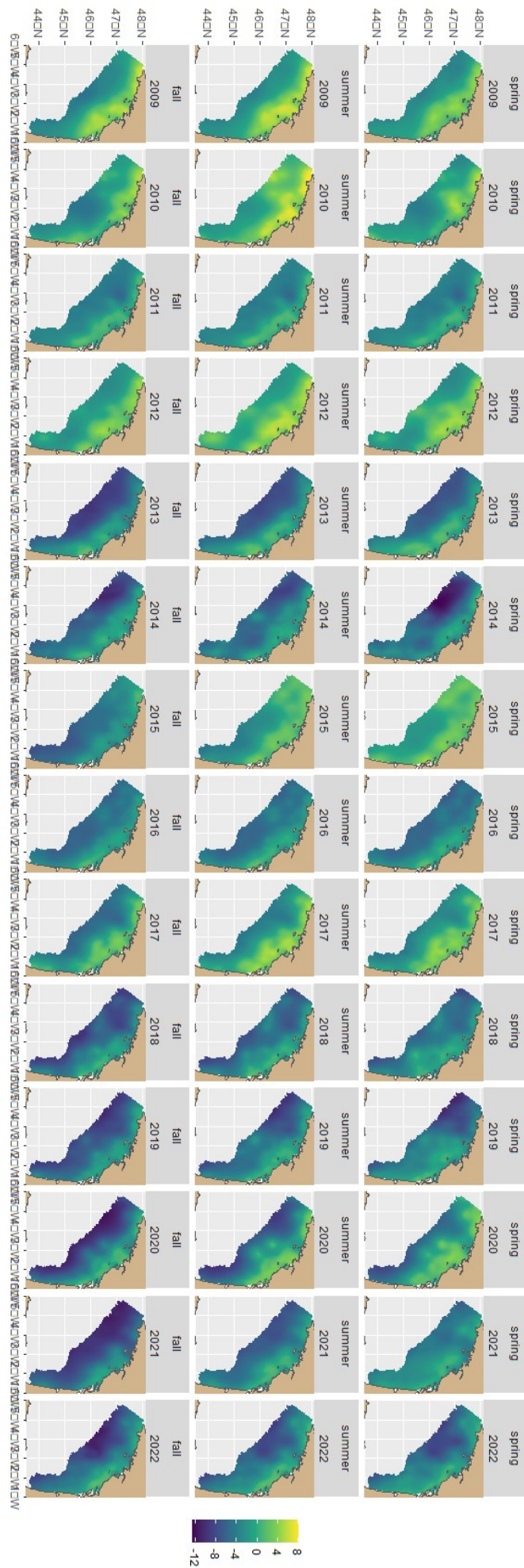


Figure 12: Posterior mean of the common spatial random effect for all the series (2009-2022)

Predictions of probability of sardine presence

The predictions associated to the Binomial likelihood represent the probability of presence of sardines. They are obtained by summing the specific Gaussian random field, shared Gaussian random field, season intercept and covariate intercept. The specific Gaussian random field of the presence is only informed by the survey so it is equal to zero in summer (no survey).

In spring, sardines were found on average in the coastal strip from the south to Belle Ile. The distribution was patchy, with small clusters of dense sardine schools (Figure 13). Small patches of high temporal variability were spread over the whole area. Average prediction error was minimal in areas of high probability of presence and along the survey acoustic transects.

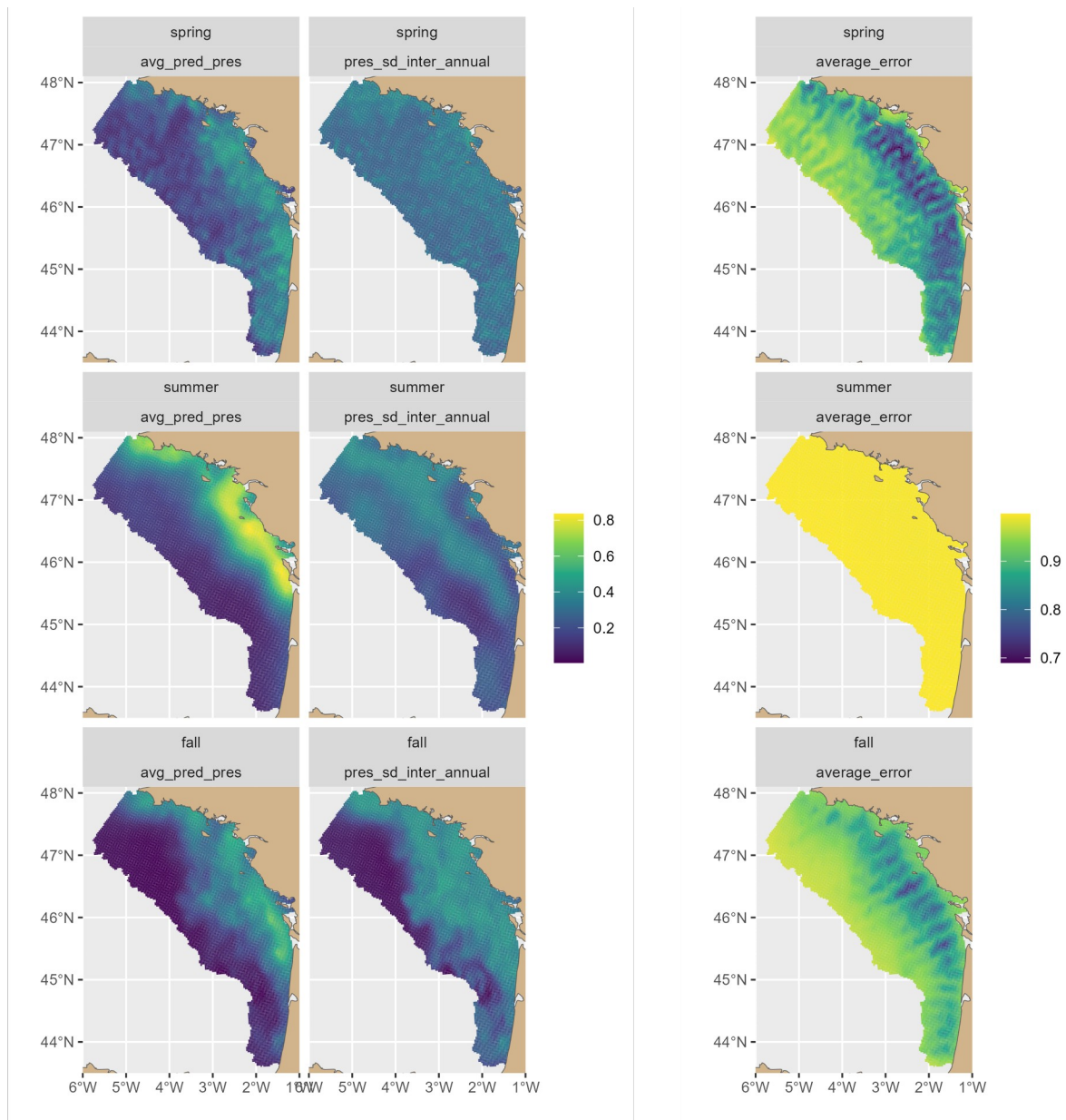


Figure 13: Average probability, standard deviation and average error of sardine presence for all the year (avg_pred_pres correspond to the average prediction of probability of presence) (Appendix 8)

In summer, sardines seemed on average to move North and concentrate near the coast: from the Gironde mouth to Belle Ile and also in south-western Brittany (Figure 13). The distribution was more continuous and concentrated than in spring, with a high probability of presence in the areas mentioned above. Inter-annual variability was the highest at the periphery of high probability of presence areas, probably denoting inter-annual expansion/contraction of those core distribution areas. Average prediction error was high throughout the zone in summer, as the prediction is only model-dependent.

In fall, sardine seemed on average to spread again offshore and towards central BoB. The distribution was more patchy, but less than in spring. Temporal variability was high in and at the periphery of areas of high probability of presence, suggesting that the sardine distribution was most variable in fall. Average prediction error was lower in areas of high probability of presence and along acoustic transects.

Predictions of sardine biomass / intensity

The predictions associated to the LogNormal likelihood represent the intensity / biomass of sardines, estimated in areas where the probability of presence equals one. Predictions are obtained by summing the specific Gaussian random field, shared Gaussian random field, season intercept and covariate intercept.

The overall patterns were the same as for the probability of presence, but the inter-annual variability was more patchy in spring and autumn (Figure 14). The sardine distribution was however more coastal in autumn, suggesting that denser shoals were found on average near the coast, and less dense school further offshore. Average predictions errors were lower in areas of high density and along acoustic transects.

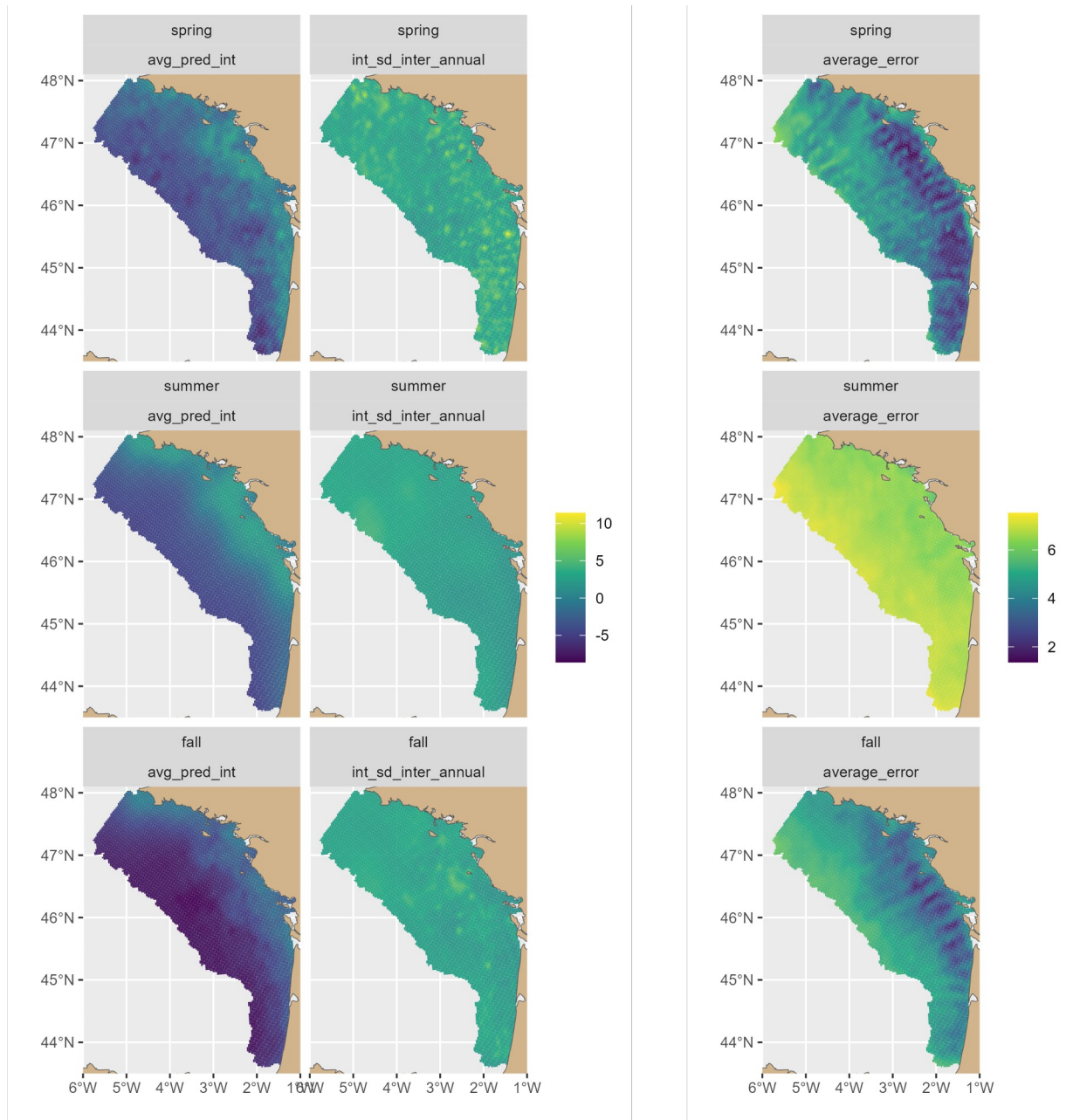


Figure 14: Prediction, standard deviation and average error of sardine intensity for all the year in logarithm scale (*avg_pred_int* correspond to the average prediction of biomass intensity) (Appendix 9)

Prediction of occurrence of the sardine fisheries

The prediction associated to the Poisson point process likelihood represents the density of positive catches by commercial fishermen. Predictions are obtained by summing the specific Gaussian random field, shared Gaussian random field, season intercept and covariate intercept. The specific Gaussian random fields were estimated for the 3 seasons, as fisheries occurred during the 3 periods.

The same average spatial pattern was seen for fishing activity between the 3 seasons (Figure 15). Sardine fishery was more active in a coastal area from the Gironde to Belle Ile and in South-West Brittany. Fishing intensity was the lowest in spring (early season), highest in summer and lowest in autumn. Average prediction errors showed a negative offshore-inshore gradient and were higher summer and lower in autumn.

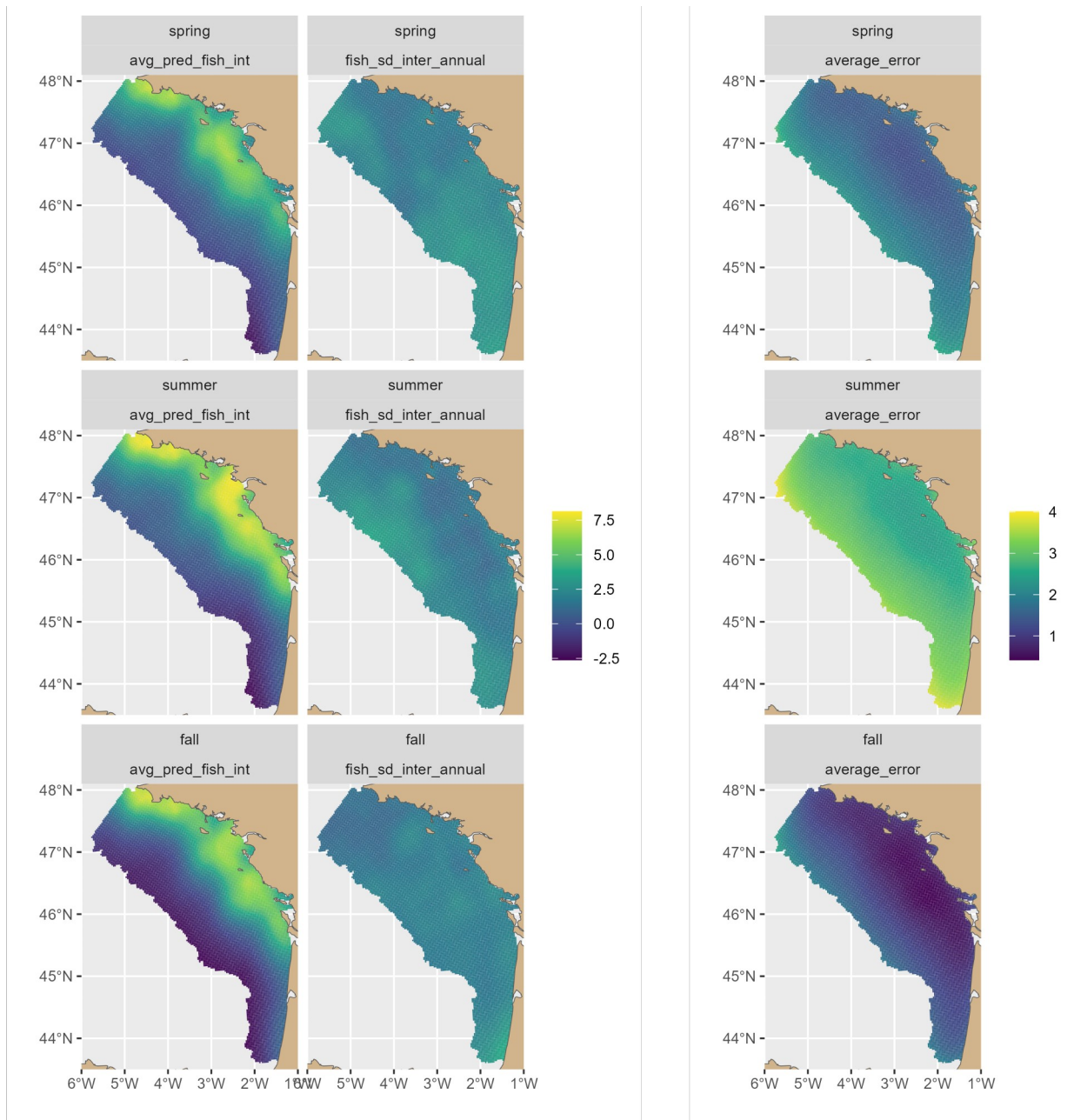


Figure 15: Prediction, standard deviation and average error of sardine fishery intensity for all the year in logarithm scale (avg_pred_fish_int correspond to the average prediction of fishery intensity) (Appendix 10)

4. Discussion

In this work, we have developed a modelling framework for combining data from scientific surveys and commercial fishing, to map the presence and density of small pelagic fishes, as well as fishery activity in the Bay of Biscay at seasonal and annual scales. The model was based on the work of Quemper (2021), which aimed at mapping the probability of sardine presence at monthly and annual scales, and assessing fishery preferential sampling. As Quemper (2021) did not find any preferential sampling, and showed that fishing data were too scarce to make reliable predictions in winter, we focused on: i) mapping fish density in addition to presence/absence, for seasons where enough data were available (spring, summer and autumn), and ii) developing a more realistic representation of the pelagic fishing process (presence only instead of presence/absence data). Our model integrated three processes and their associated likelihood. We considered the presence-absence and intensity data from scientific surveys and the fishing occurrence to inform Gaussian random fields. We defined a random field specific to each process and a common random field that integrated the variability from all data sources. The model has been run for 3 seasons from 2009 to 2022, each year independently. Model predictions provided for the first time a quantitative description of the seasonal spatial dynamics of the sardine and its associated fisheries in the BoB. Sardine core distribution areas were inshore, from the Gironde river mouth to South-West Brittany. Sardine distribution was on average more widespread and patchy in spring. Fish appeared to concentrate in their coastal core distribution areas in summer, and spread towards offshore areas in autumn, though being less dispersed than in spring. The inter-annual variability of the sardine distribution was the highest in spring and autumn. Summer predictions were however less reliable, as no survey data were available for this season. Model outputs confirmed that sardine fisheries essentially operated in fish coastal core distribution areas, with a peak activity in summer, medium and low activity in autumn and spring, respectively.

An integrated modelling framework to infer spatio-temporal distribution of small pelagic fish abundance

A spatio-temporal model to account for seasonal and inter-annual variations in small pelagic fish

Our new modelling approach estimates annual, seasonal variations in small pelagic spatial distribution by including both survey and fisheries data. It implied that the distribution of small pelagic fish varied over seasons and years, which is reasonable from a biological point of view, in the context of migratory, short-lived pelagic species exploiting the dynamic biotope of the Bay of Biscay. In fact, habitats occupied by sardines and anchovies for feeding are different from those for reproduction in the BoB (Petitgas, 2010; Politikos et al., 2015; Massé et al., 2018). The anchovies spawning grounds are located in the southeast corner of Biscay, and the feeding grounds are in the northern French shelf (Petitgas et al., 2010). The feeding migration occurs after spawning, mainly in July and the comeback south occurs in December. The migration is assumed to follow a north–south gradient in food and temperature (Petitgas, 2010; Politikos et al., 2015). There is also this type of migration for sardines in the Bay of Biscay (Petitgas et al., 2010). In the literature, the main environmental factors significantly related to the distribution of anchovy or sardine were

temperature, chlorophyll, salinity and depth (Coombs et al., 2006; Doray et al., 2018a; Erauskin-Extramiana et al., 2019; Fernández-Corredor et al., 2021)). It is expected inter-annual variability depending on annual environmental conditions (Olmos et al., 2023). Petitgas et al (2020) suggested that the distribution of sardine eggs shifted towards coastal areas from 2009 onwards. Since 1900, there has also been a change in the distribution of anchovy fisheries, which are moving northwards (Petitgas et al., 2010). Finally, according to the various IPCC scenarios, a turnover in species in the Bay of Biscay is expected to happen (Le Marchand et al., 2020). Our model could be used to follow these distributional changes at the seasonal and annual scales in the BoB.

The model was built on a seasonal time step instead of monthly step because fisheries data are less reliable than survey ones and do not cover the whole area and all seasons. It was therefore essential to base the model predictions on survey data. There is only one survey in the Bay of Biscay in May and one in September for small pelagic fish. Contrary to the fishing season that starts in April/May and end in September/October with a maximum in summer.

With respect to spatio-temporal models, this model provides a significant contribution in how to model seasonal and inter-annual variations in spatiotemporal model. Previous authors have included seasonal variation in isolation (Thorson et al., 2016; Grieve et al., 2017) or included both changes in spatial distribution among years and among seasons (Bourdaud et al., 2017; Kai et al., 2017; Kanamori et al., 2019; Akia et al., 2021). In our modelling approach we took the decision to first select the nature of the data, i.e only presence for fishery data, presence/absence for survey data and intensity of catch for survey data, to be able to use the most pertinent information from each data source.

Finally, because most ontogenetic migration between spawning and feeding occurs between summer and end of fall/beginning of winter, it could be interesting to extend our approach by integrating a winter season in the model using fishery data during winter and EVOHE (a winter bottom trawl survey which can inform on SA presence)

An integrated hierarchical model including survey and fishery data

The model combined both survey and fishery data within a single statistical framework. The exploitation of SA by commercial fleets is characterised by a very strong link between fishing activity and market demand, resulting in strong seasonality, high size-based fish selectivity, and fishing activity that is highly localised to the coast (Quemper, 2021; ICES, 2022). Catches of SA by commercial fishers provide information on the presence of marketable fish in geographically and temporally restricted areas. In our study we showed that the absence of correlation between CPUEs and survey biomass makes it impossible to use fishing data other than in terms of presence alone. This is because fishermen target sardines and anchovies of a certain weight and size and in recent years, high abundances have been concentrated at ages 0 and 1 (ICES, 2022). High biomasses during the surveys generally correspond to young sardines that are of no interest to fishermen. The correlation found between CPUEs of purse seiners and anchovy survey biomass in spring needs to be put into perspective, as it represents very few data in a restricted geographic area on the French Brittany coast (Duhamel et al., 2004). Previous studies have combined survey and fishery CPUE. Alglave et al. (2022; 2023) developed an integrated modelling framework to infer spatio-temporal distribution of fish abundance by combining commercial fishing data and scientific survey while accounting for preferential in the distribution of fishing effort. However as already discussed by

Quemper (2021), applying such a framework to SA might not improve significantly the inference of SA distribution due to the nature of the fishery. Future work could focus on understanding the factors driving the market (distance to canneries, fishing cost for example) and fisherman behavior to be able to better represent the preferential sampling of this fishery.

Accounting for covariates

Only one covariate, the distance to the coast, was considered in our model. Numerous studies showed that the distribution of Small Pelagic Fish (SPF) depends on environmental parameters such as SST, SSS, Chl-a and bathymetry. However, in the work of Quemper (2021), the effects of these covariates did not improve the models predictive capacities. We however tested the effect of the distance to shore, a covariate that had not been yet tested, and could act as a proxy for riverine influence from the fish point of view, and for oil consumption, from the fisherman point of view. Moreover, in the Bay of Biscay, bathymetry and distance from the coast are correlated along a latitudinal gradient and one can be used as a proxy for the other. The results showed that the effect of the distance to shore varied over the years. For 7 years on 14, the model with covariate was better than without. In the first try without covariates, the deviance residual map showed a coast-offshore gradient. Finally, adding the distance from the coast did not change the deviance residual maps. It is difficult to conclude about the effect of this covariate. Investigating the effects of more biologically structuring covariates (zooplankton preys ...) was beyond the scope of this study and is left for future ones: our model can incorporate covariates easily once these have been extracted and prepared.

Computational limits to build an integrated modelling framework

An important limit of our modelling framework is the absence of a single model for the entire series. Due to unexpected computing power limitations, it was not possible to run a model including all 14 years with 3 seasons on a personal computer. Fitting one model per year implied that the years were independent (although within year variations were modelled). We had to make the strong hypothesis that the intercepts and covariate effects associated with each process were non-stationarity over time, which would not have been the case in a single model fitted over all years. In the same way, specific random fields had to be treated as non stationary over time. They represented the seasonal variability of each process for one year, instead of the whole series.

A hierarchical model to allocate the different data information into shared terms (representing the distribution of seasonal and inter-annual variations) and terms specific to the different data sources

A shared random field to infer seasonal and inter-annual variations distribution from different data sources

In this study, we wanted to predict the distribution of SPF in summer, despite the absence of data from scientific surveys. The information provided by fisheries occurrences could not be used alone, because sampling does not cover the whole Bay of Biscay. This is why we inferred a random shared Gaussian field using the 3 data sources, which would vary over the seasons according to an auto-

regressive model of order 1 (AR1). The use of an auto-regressive term allows to use the more exhaustive and precise spring and fall data to predict the summer distribution. The common random field captures all available information and provides a summer distribution estimate. This makes it possible to identify hot spots (coast from the Gironde to the Loire estuaries, and in South West Brittany, (Figure 10) where sardines were abundant and the fishery active. As our model procedure makes it possible to integrate different processes, it could be interesting in the Delmoges project to create a model including dolphin occurrence data, to determine hot spots where SPFs, fishermen and dolphins co-occur.

The shared random field and predictions showed that sardines were found along the coast, essentially from the south of Brittany to the Gironde estuary. Quemper (2021) found a similar pattern and identified two areas described by Bellier et al. (2007) as being areas under riverine influence: (1) south-west Brittany, north of the Gironde estuary (2) south of the Gironde estuary. These areas have been identified in other studies as hotspots of adult sardine hotspots (Doray et al., 2018a) and sardine eggs (Huret et al., 2018; Petitgas et al., 2020).

However, some hotspots were only present for a few years. In 2015 (Figure 11), an area of positive values was clearly visible on the edge of the continental shelf between 46°N – 47°30'N. This area was also described as a spawning zone (Bellier et al., 2007). There is also a zone in Cape Breton in certain years (2012, 2014, 2015, 2017, Figure 12), documented by (Petitgas et al., 2020) as an area of high egg concentration in 2010. Sardine seemed to migrate North from spring to fall in certain years (2015, 2019), but no clear consistent migration pattern could be found in the SGF results.

A specific random field associated to each process (presence only, presence/absence and intensity of catches)

Even if the shared random fields provide very interesting information, it does not explain all the total variance. Specific random fields hence explain a larger part of the variability of spatio-temporal distribution of SPF. Also, if a random Gaussian field represents variability, it does not take into account the fixed effects of the model (intercepts and covariates). Predictions can only be made at the scale of a process. The probability of presence prediction maps were hence very different from the common random field distribution pattern. The predictions were driven by the specific random field, which has a larger range than the shared field. A comparison between the raw data and the prediction maps showed the same distribution pattern. Our model therefore takes good account of the quantity and quality of the survey data when making its estimates. The same is true for biomass predictions in the study area. As there is no scientific survey in summer, predicted distribution pattern corresponds exactly to that of the shared random field, modulo intercepts.

The analysis of average prediction maps derived from our model provided for the first time a quantitative description of sardine seasonal spatial dynamics in the BoB. If sardine core distribution areas were consistently coastal areas from Gironde to Brittany, their distribution appeared to be on average more widespread and patchy in spring. Our results suggest that sardine tend to concentrate in their coastal core distribution areas in summer, and spread towards offshore areas in autumn, though being less dispersed than in spring. Inter-annual variability was high at small scale over the whole BoB in spring, at the periphery of core distribution areas in summer and inside and at the

periphery of sardine concentrations in autumn. Inter-annual variability in sardine distribution hence seemed to be the highest in autumn.

Summer predictions were however less reliable than spring and autumn ones, as no survey data were available at this season. It is also worth noting that the fall JUVENA survey has a coarser resolution compared to the spring PELGAS survey (10 times less samples), as it must cover the off the shelf areas of the BoB in addition to shelf areas considered in this study (Doray et al., 2021). Differences in sampling intensities between seasons may hence contribute to explain some inter-seasonal differences in spatial patterns, at least at smaller scale (e.g. sardine distributions more patchy in spring compared to other seasons).

The specific field has only a positive effect on the predictions of fishing activity (Figure 18). It provides new information about hotspots of fishing activity, but the common field prevails elsewhere. In some cases, predictions of fishing occurrences are entirely driven by the shared field (2009, 2010, 2012, 2017).

Finally, the integration of commercial fishing data allows to supplement the information provided by scientific data in very localised coastal areas (Quemper, 2021). It also makes it possible to provide information for periods when there are no surveys. Nevertheless, our results highlight predictions are largely driven by survey data in the model, as surveys remain the most reliable and spatially extended and homogeneous data. Acoustic surveys are therefore necessary for a full understanding of the distribution of these species.

Further work would include to fit the model for other species, namely anchovy, and to integrate other data sources informing on the winter situation.

The Delmoges project has hence conducted a pilot acoustic survey using the uncrewed survey vehicle DriX in February 2023 that provided new quantitative information on small pelagic fish distribution in the BoB in winter. The EVHOE bottom trawl scientific survey also takes place in the BoB in November. Data from this other scientific survey could provide useful information on sardine distribution in early winter. The flexible structure of our model should allow for the inclusion of those new data source to provide a better understanding of the distribution of small pelagic fish in winter, at the peak dolphin by-catch season (Dars et al., 2021). This would be essential to map areas where SPF and dolphins co-occur, and assess if their co-occurrence increases the risk of dolphin by-catch.

5. Conclusion

This work demonstrated the interest of integrating different data sources to map the presence and density of a small pelagic fish, as well as fishing activity in the BoB at seasonal and annual scales. The specificities of small pelagic fish fisheries (CPUE over-stability, seasonality, size-based selectivity, market influence) required to develop a pelagic-specific SDM. Its use could provide information about the distribution of SPF during data poor period, within coastal areas where

fishing activity take place. Further integrating winter data and common dolphin distribution into a pelagic-SDM may allow to improve our understanding of the interactions between small pelagic fish, fisheries and dolphin in space and time. Using the outputs from our model and dolphins species distribution models, future works may use overlap metrics (Carroll et al., 2019) to investigate the spatio-temporal interactions between dolphins, small pelagic fish and fisheries and assess their impact on common dolphin by-catch in the BoB.

Bibliography

- Akia S, Amandé M, Pascual P, Gaertner D** (2021) Seasonal and inter-annual variability in abundance of the main tropical tunas in the EEZ of Côte d'Ivoire (2000-2019). *Fish Res* **243**: 106053
- Alglave B, Rivot E, Etienne M-P, Woillez M, Thorson JT, Vermard Y** (2022) Combining scientific survey and commercial catch data to map fish distribution. *ICES J Mar Sci* **79**: 1133–1149
- Alglave B, Vermard Y, Rivot E, Etienne M-P, Woillez M** (2023) Identifying mature fish aggregation areas during spawning season by combining catch declarations and scientific survey data. *Can J Fish Aquat Sci*. doi: 10.1139/cjfas-2022-0110
- Andrews S, Leroux S, Fortin M-J** (2020) Modelling the spatial-temporal distributions and associated determining factors of a keystone pelagic fish. doi: 10.1101/2020.04.16.044156
- Archambault B, Rivot E, Savina M, Le Pape O** (2018) Using a spatially structured life cycle model to assess the influence of multiple stressors on an exploited coastal-nursery-dependent population. *Estuar Coast Shelf Sci* **201**: 95–104
- Bellier E, Planque B, Petitgas P** (2007) Historical fluctuations in spawning location of anchovy (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*) in the Bay of Biscay during 1967-73 and 2000-2004. *Fish Oceanogr* **16**: 1–15
- Bourdaud P, Travers-Trolet M, Vermard Y, Cormon X, Marchal P** (2017) Inferring the annual, seasonal, and spatial distributions of marine species from complementary research and commercial vessels' catch rates. *ICES J Mar Sci* **74**: 2415–2426
- Boyra G, Rico I, Martínez U** (2020) Acoustic surveying of anchovy Juveniles in the Bay of Biscay: JUVENA 2020 Survey Report. doi: 10.13140/RG.2.2.36115.09768
- Carroll G, Holsman KK, Brodie S, Thorson JT, Hazen EL, Bograd SJ, Haltuch MA, Kotwicki S, Samhuri J, Spencer P, et al** (2019) A review of methods for quantifying spatial predator–prey overlap. *Glob Ecol Biogeogr* **28**: 1561–1577
- Chapuis A, Barrabes M, Don J** (2021) *Sardina pilchardus* (Walbaum, 1792).
- Conn PB, Thorson JT, Johnson DS** (2017) Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. *Methods Ecol Evol* **8**: 1535–1546
- Coombs SH, Smyth T, Conway D, Halliday NC, Bernal M, Stratoudakis Y, Alvarez P** (2006) Spawning season and temperature relationships for sardine (*Sardina pilchardus*) in the eastern North Atlantic. *J Mar Biol Assoc U K*. doi: 10.1017/S0025315406014251
- Costoya X, Decastro M, Gesteira M, Santos F** (2015) Changes in sea surface temperature seasonality in the Bay of Biscay over the last decades (1982 - 2014). *J Mar Syst* **150**: 91–101
- Dars C, Meheust E, Genu M, Méndez-Fernandez P, Peltier H, Wund S, Caurant F, Dabin W, Demaret F, Spitz J, et al** (2021) Les échouages de mammifères marins sur le littoral français en 2021. 43

- Delmoges** (2021) Présentation du projet DELMOGES (DELphinus MOuvement GESTion). 39
- Diggle PJ, Menezes R, Su T** (2010) Geostatistical inference under preferential sampling. *J R Stat Soc Ser C Appl Stat* **59**: 191–232
- Doray M, Boyra G, Kooij J van der** (2021) ICES Survey Protocols – Manual for acoustic surveys coordinated under ICES Working Group on Acoustic and Egg Surveys for Small Pelagic Fish (WGACEGG). doi: 10.17895/ices.pub.7462
- Doray M, Hervy C, Huret M, Petitgas P** (2018a) Spring habitats of small pelagic fish communities in the Bay of Biscay. *Prog Oceanogr* **166**: 88–108
- Doray M, Petitgas P, Huret M, Duhamel E, Romagnan JB, Authier M, Dupuy C, Spitz J** (2018b) Monitoring small pelagic fish in the Bay of Biscay ecosystem, using indicators from an integrated survey. *Prog Oceanogr* **166**: 168–188
- Duhamel E, Biseau A, Masse J, Villalobos H** (2004) The French anchovy fishery.
- Elith J, Leathwick JR** (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu Rev Ecol Evol Syst* **40**: 677–697
- Erauskin-Extramiana M, Alvarez P, Arrizabalaga H, Ibaibarriaga L, Uriarte A, Cotano U, Santos M, Ferrer L, Cabré A, Irigoien X, et al** (2019) Historical trends and future distribution of anchovy spawning in the Bay of Biscay. *Deep Sea Res Part II Top Stud Oceanogr* **159**: 169–182
- European Commission** (2020) Procédures d’infraction du mois de juillet: principales décisions.
- Fernández-Corredor E, Albo-Puigserver M, Pennino MG, Bellido JM, Coll M** (2021) Influence of environmental factors on different life stages of European anchovy (*Engraulis encrasicolus*) and European sardine (*Sardina pilchardus*) from the Mediterranean Sea: A literature review. *Reg Stud Mar Sci* **41**: 101606
- Fréon P, Cury P, Shannon L, Roy C** (2005) Sustainable exploitation of small pelagic fish stocks challenged by environmental and ecosystem changes: A review. *Bull Mar Sci* **76**: 385–462
- Fréon P, Misund O** (1999) Dynamics of pelagic fish distribution and behaviour: Effects on fisheries and stock assessment.
- Gelman A, Hwang J, Vehtari A** (2014) Understanding predictive information criteria for Bayesian models. *Stat Comput* **24**: 997–1016
- Gelman A, Jakulin A, Pittau MG, Su Y-S** (2008) A weakly informative default prior distribution for logistic and other regression models. *Ann Appl Stat* **2**: 1360–1383
- Gelman A, Shalizi CR** (2013) Philosophy and the practice of Bayesian statistics. *Br J Math Stat Psychol* **66**: 8–38

- Giannoulaki M, Markoglou E, Valavanis VD, Alexiadou P, Cucknell A, Frantzis A** (2017) Linking small pelagic fish and cetacean distribution to model suitable habitat for coastal dolphin species, *Delphinus delphis* and *Tursiops truncatus*, in the Greek Seas (Eastern Mediterranean). *Aquat Conserv Mar Freshw Ecosyst* **27**: 436–451
- Grandremy N, Romagnan J-B, Dupuy C, DORAY M, Huret M, Petitgas P** (2022) Hydrology and small pelagic fish drive the spatio-temporal dynamics of springtime zooplankton assemblages over the Bay of Biscay continental shelf. *Prog Oceanogr* **210**: 102949
- Grieve BD, Hare JA, Saba VS** (2017) Projecting the effects of climate change on *Calanus finmarchicus* distribution within the U.S. Northeast Continental Shelf. *Sci Rep* **7**: 6264
- Guisan A, Thuiller W** (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* **8**: 993–1009
- Hefley TJ, Hooten MB** (2016) Hierarchical Species Distribution Models. *Curr Landsc Ecol Rep* **1**: 87–97
- Hintzen NT, Bastardie F, Beare D, Piet GJ, Ulrich C, Deporte N, Egekvist J, Degel H** (2012) VMStools: Open-source software for the processing, analysis and visualisation of fisheries logbook and VMS data. *Fish Res* **115–116**: 31–43
- Huret M, Bourriau P, Doray M, Gohin F, Petitgas P** (2018) Survey timing vs. ecosystem scheduling: Degree-days to underpin observed interannual variability in marine ecosystems. *Prog Oceanogr* **166**: 30–40
- ICES** (2021) Workshop on fisheries Emergency Measures to minimize BYCatch of short-beaked common dolphins in the Bay of Biscay and harbor porpoise in the Baltic Sea. doi: 10.17895/ices.pub.7472
- ICES** (2022) Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA). doi: 10.17895/ices.pub.19982720.v2
- Isaac NJB, Jarzyna MA, Keil P, Dambly LI, Boersch-Supan PH, Browning E, Freeman SN, Golding N, Guillera-Arroita G, Henrys PA, et al** (2020) Data Integration for Large-Scale Models of Species Distributions. *Trends Ecol Evol* **35**: 56–67
- Kai M, Thorson JT, Piner KR, Maunder MN** (2017) Predicting the spatio-temporal distributions of pelagic sharks in the western and central North Pacific. *Fish Oceanogr* **26**: 569–582
- Kanamori Y, Takasuka A, Nishijima S, Okamura H** (2019) Climate change shifts the spawning ground northward and extends the spawning period of chub mackerel in the western North Pacific. *Mar Ecol Prog Ser* **624**: 155–166
- Koutsikopoulos C, Cann BL** (1996) Physical processes and hydrological structures related to the Bay of Biscay anchovy. *Sci. Mar.*
- Krainski ET, Gómez-Rubio V, Bakka H, Lenzi A, Castro-Camilo D, Simpson D, Lindgren F, Rue H** (2021) Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA.

- Lahaye V, Spitz J, Dabin W, Das K, Pierce G, Caurant F** (2005) Long-term dietary segregation of common dolphins *Delphinus delphis* in the Bay of Biscay, determined using cadmium as an ecological tracer. *Mar Ecol Prog Ser* **305**: 275–285
- Lahellec G** (2020) Typologie de la flotte de pêche exploitant les petits pélagiques dans le Golfe de Gascogne et identification des facteurs influençant leur stratégie d’exploitation. Institut Agro Rennes-Angers, Rennes
- Lambert C, Authier M, Doray M, Dorémus G, Spitz J, Ridoux V** (2019) Hide and seek in the Bay of Biscay—a functional investigation of marine megafauna and small pelagic fish interactions. *ICES J Mar Sci* **76**: 113–123
- Lassalle G, Pasqual J-S, Boët P, Rochet M-J, Trenkel V, Niquil N** (2013) Combining quantitative and qualitative models to identify functional groups for monitoring changes in the Bay of Biscay continental shelf exploited foodweb. *ICES J Mar Sci*. doi: 10.1093/icesjms/fst107
- Le Marchand M, Hattab T, Niquil N, Albouy C, Le Loc’h F, Lasram F** (2020) Climate change in the Bay of Biscay: Changes in spatial biodiversity patterns could be driven by the arrivals of southern species. *Mar Ecol Prog Ser*. doi: 10.3354/meps13401
- Lindgren F** (2012) Continuous domain spatial models in R-INLA. *ISBA Bull* **19**: 14–20
- Lindgren F, Rue H, Lindst J** (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach.
- Mannocci L, Catalogna M, Dorémus G, Laran S, Lehodey P, Massart W, Monestiez P, Van Canneyt O, Watremez P, Ridoux V** (2014) Predicting cetacean and seabird habitats across a productivity gradient in the South Pacific gyre. *Prog Oceanogr* **120**: 383–398
- Marçalo A, Nicolau L, Giménez J, Ferreira M, Santos JM, Araújo H, Silva A, Vingada J, Pierce G** (2018) Feeding ecology of the common dolphin (*Delphinus delphis*) in Western Iberian waters: has the decline in sardine (*Sardina pilchardus*) affected dolphin diet? *Mar Biol* **165**: 44
- Martin Gonzalez G, Wiff R, Marshall CT, Cornulier T** (2021) Estimating spatio-temporal distribution of fish and gear selectivity functions from pooled scientific survey and commercial fishing data. *Fish Res* **243**: 106054
- Massé J, Uriarte A, Angelico MM, Carrera P** (2018) Pelagic survey series for sardine and anchovy in ICES subareas 8 and 9 (WGACEGG) – Towards an ecosystem approach. *ICES Coop Res Rep* 332:
- Maunder MN, Sibert JR, Fonteneau A, Hampton J, Kleiber P, Harley SJ** (2006) Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES J Mar Sci* **63**: 1373–1385
- Meynier L, Pusineri C, Spitz J, Santos M, Pierce G, Ridoux V** (2008) Intraspecific dietary variation in the short-beaked common Dolphin *Delphinus delphis* in the Bay of Biscay: Importance of fat fish. *Mar Ecol Prog Ser* **354**: 277–287

- Modrak G, Caro V, Menut T** (2022) *Engraulis encrasicolus* (Linnaeus, 1758).
- Moraga P** (2020) Welcome | Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny.
- Moreira GA, Menezes R, Wise L** (2023) Presence-Only for Marked Point Process Under Preferential Sampling. *J Agric Biol Environ Stat.* doi: 10.1007/s13253-023-00558-x
- Moriarty M, Sethi SA, Pedreschi D, Smeltz TS, McGonigle C, Harris BP, Wolf N, Greenstreet SPR** (2020) Combining fisheries surveys to inform marine species distribution modelling. *ICES J Mar Sci* **77**: 539–552
- Murphy S, Pinn E, Jepson P** (2013) The short-beaked common dolphin (*Delphinus delphis*) in the North-eastern Atlantic: distribution, ecology, management and conservation status. *Oceanogr Mar Biol* **51**: 193–280
- Olmos M, Ianelli J, Ciannelli L, Spies I, McGilliard CR, Thorson JT** (2023) Estimating climate-driven phenology shifts and survey availability using fishery-dependent data. *Prog Oceanogr* **215**: 103035
- Opitz T** (2016) Modèles à processus gaussiens latents, inférence INLA*, R-INLA * Integrated Nested Laplace Approximation.
- Peltier H, Authier M, Caurant F, Dabin W, Daniel P, Dars C, Demaret F, Laran S, Meheust E, Ridoux V, et al** (2022) Bilan des échouages et des captures accidentelles de dauphins communs dans le golfe de Gascogne - Hiver 2021. 16
- Peltier H, Authier M, Caurant F, Dabin W, Daniel P, Dars C, Demaret F, Meheust E, Ridoux V, Van Canneyt O, et al** (2020) Bilan 2020 des évènements d'échouages de l'hiver et de l'été, cartographie des mortalités et corrélation spatiale avec les pêcheries. 12
- Peltier H, Authier M, Caurant F, Dabin W, Daniel P, Dars C, Demaret F, Meheust E, Van Canneyt O, Spitz J, et al** (2021) In the Wrong Place at the Wrong Time: Identifying Spatiotemporal Co-occurrence of Bycaught Common Dolphins and Fisheries in the Bay of Biscay (NE Atlantic) From 2010 to 2019. *Front. Mar. Sci.* **8**:
- Peltier H, Authier M, Caurant F, Dabin W, Dars C, Demaret F, Meheust E, Ridoux V, Van Canneyt O, Spitz J** (2019) Etat des connaissances sur les captures accidentelles de dauphins communs dans le golfe de Gascogne – Synthèse 2019. 23
- Pennino MG, Paradinas I, Illian JB, Muñoz F, Bellido JM, López-Quílez A, Conesa D** (2019) Accounting for preferential sampling in species distribution models. *Ecol Evol* **9**: 653–663
- Persohn C** (2009) Sensibilité des populations de poissons aux changements globaux en fonction de leurs traits d'histoire de vie étudiée par une approche empirique. Université de Bretagne Occidentale
- Petitgas P** (2010) Life cycle spatial patterns of small pelagic fish in the Northeast Atlantic. *ICES Coop Res Rep* **93**

- Petitgas P, Alheit J, Beare D, Bernal M, Casini M, Clarke M, Cotano U, Dickey-Collas M, Dransfeld L, Harma C, et al** (2010) Life-cycle spatial patterns of small pelagic fish in the Northeast Atlantic. doi: 10.17895/ices.pub.5408
- Petitgas P, Jacques M, Paul B, Pierre B, Bergeron J-P, Daniel D, Alain H, N K, Froidefond J-M, Santos M** (2006) Hydro-plankton characteristics and their relationship with sardine and anchovy distributions on the French shelf of the Bay of Biscay. *Sci. Mar.* 0214-8358 *Inst. Ciènc. Mar Barc.* CSIC 2006-06 Vol 70 N 1 P 161-172 70:
- Petitgas P, Renard D, Desassis N, Huret M, Romagnan J-B, Doray M, Woillez M, Rivoirard J** (2020) Analysing Temporal Variability in Spatial Distributions Using Min–Max Autocorrelation Factors: Sardine Eggs in the Bay of Biscay. *Math Geosci.* doi: 10.1007/s11004-019-09845-1
- Phillip M, Robert T** (1998) Fishing operations. 1. Vessel monitoring systems. FAO, Rome, Italy
- Phillips SJ, Anderson RP, Schapire RE** (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* **190**: 231–259
- Pinto C, Travers-Trolet M, Macdonald J, Rivot E, Vermard Y** (2018) Combining multiple data sets to unravel the spatiotemporal dynamics of a data-limited fish stock. *Can J Fish Aquat Sci.* doi: 10.1139/cjfas-2018-0149
- Pitcher TJ** (1995) The impact of pelagic fish behaviour on fisheries. *Sci Mar* **59**: 295–306
- Politikos DV, Huret M, Petitgas P** (2015) A coupled movement and bioenergetics model to explore the spawning migration of anchovy in the Bay of Biscay. *Ecol Model* **313**: 212–222
- Post van der Burg M, Austin JE, Wiltermuth MT, Newton W, MacDonald G** (2020) Capturing Spatiotemporal Patterns in Presence-Absence Data to Inform Monitoring and Sampling Designs for the Threatened Dakota Skipper (Lepidoptera: Hesperiiidae) in the Great Plains of the United States. *Environ Entomol* **49**: 1252–1261
- Pusineri C, Magnin V, Meynier L, Spitz J, Hassani S, Ridoux V** (2007) Food and feeding ecology of the common dolphin (*Delphinus delphis*) in the oceanic Northeast Atlantic and comparison with its diet in neritic areas. *Mar Mammal Sci* **23**: 30–47
- Quemper F** (2021) Modélisation de la distribution spatiale de la sardine du Golfe de Gascogne (*Sardina pilchardus*) par intégration de données commerciales et scientifiques : enjeux et limites. 90
- Rouby E** (2022) Population dynamics of elusive species : The case of the common dolphin in the North-East Atlantic Ocean. These de doctorat. La Rochelle
- Rue H, Martino S, Chopin N** (2009) Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *J R Stat Soc Ser B* **71**: 319–392
- Rufener M-C** (2020) Integrating commercial fisheries and scientific survey data: Advances, new tools and applications to model the fish and fishery dynamics. doi: 10.13140/RG.2.2.31991.34727
- Rundel C** (2017) Lecture 3 : Residual Analysis + Generalized Linear Models.

Schickele A, Leroy B, Beaugrand G, Goberville E, Hattab T, Francour P, Raybaud V (2020) Modelling European small pelagic fish distribution: Methodological insights. *Ecol Model* **416**: 108902

Simmonds J, Maclellan DN (2007) Fisheries acoustics: Theory and practice: Second edition. *Fish Acoust Theory Pract Second Ed* 1–252

Spitz J, Mourocq E, Schoen V, Ridoux V (2010) Proximate composition and energy content of forage species from the Bay of Biscay: High- or low-quality food? *ICES J Mar Sci* **67**: 909–915

Taylor N, Authier M, Banga R, Genu M, Macleod K, Gilles A (2022) Marine Mammal By-catch.

Thorson JT, Pinsky ML, Ward EJ (2016) Model-based inference for estimating shifts in species distribution, area occupied and centre of gravity. *Methods Ecol Evol* **7**: 990–1002

Whitehead J (1985) Clupeoid fishes of the world. *FAO species catalogue Vol 7*. An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, anchovies and wolf herrings. Part 1 : Chirocentridae, Clupeidae and Pristigasteridae.

Appendices

Appendix 1: Sardine model input data

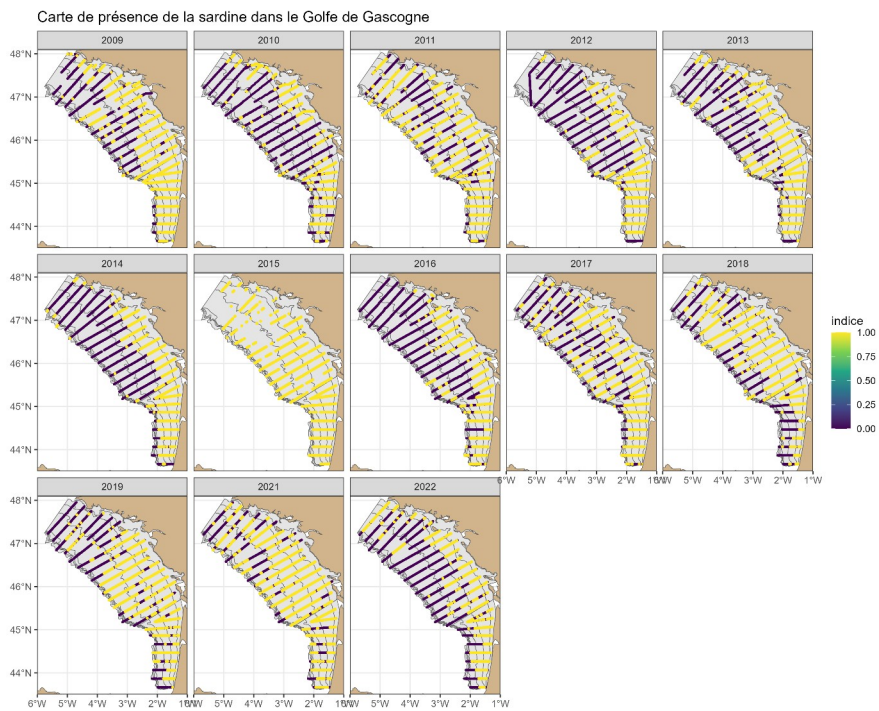


Figure 16: Spatial distribution of sardine presence/absence from PELGAS survey

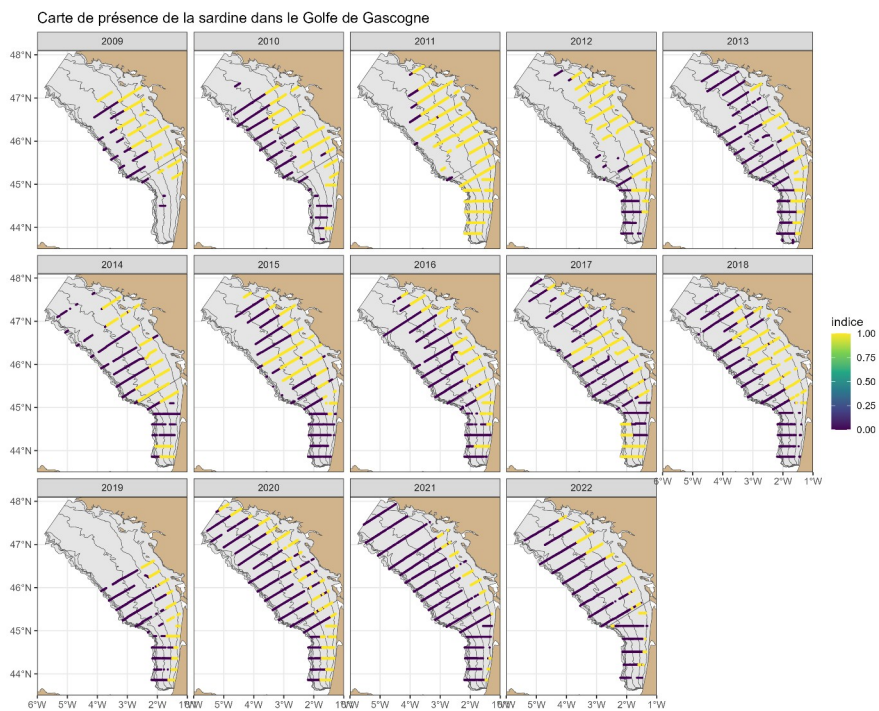


Figure 17: Spatial distribution of sardine presence/absence from JUVENA survey

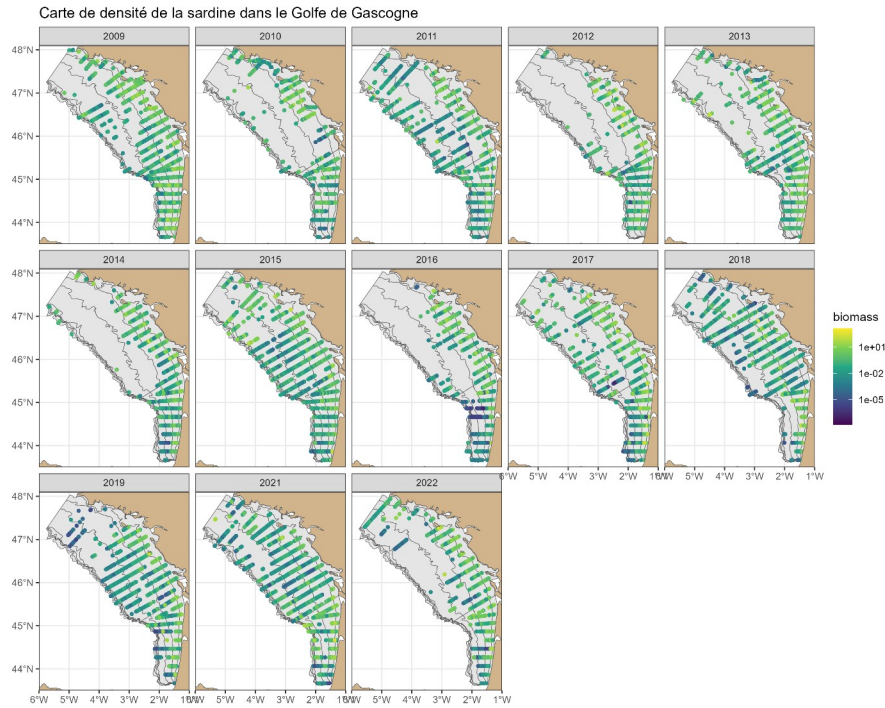


Figure 18: Spatial distribution of sardine biomass intensity from PELGAS survey (in log scale)

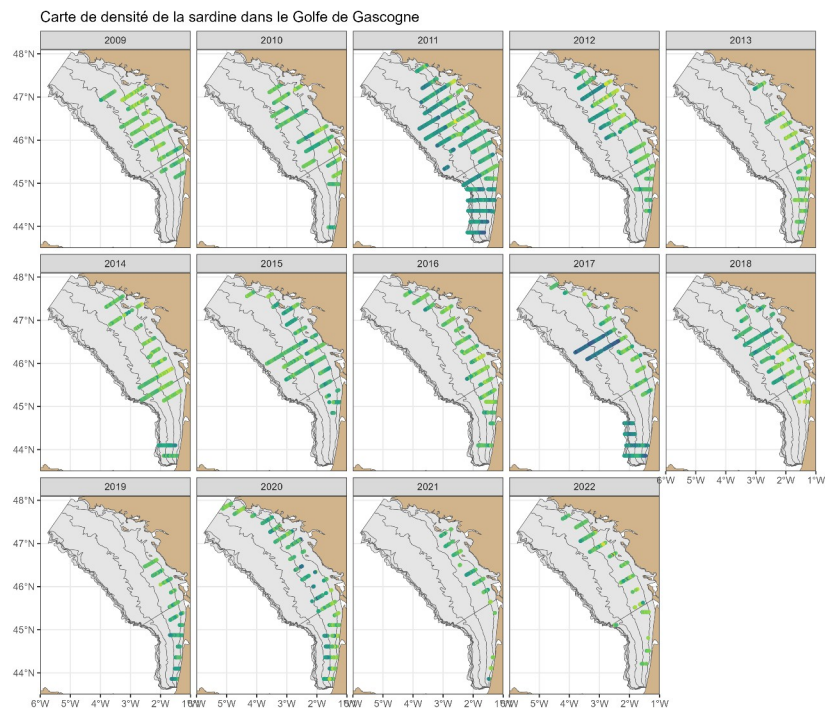


Figure 19: Spatial distribution of sardine biomass intensity from JUVENA survey (in log scale)

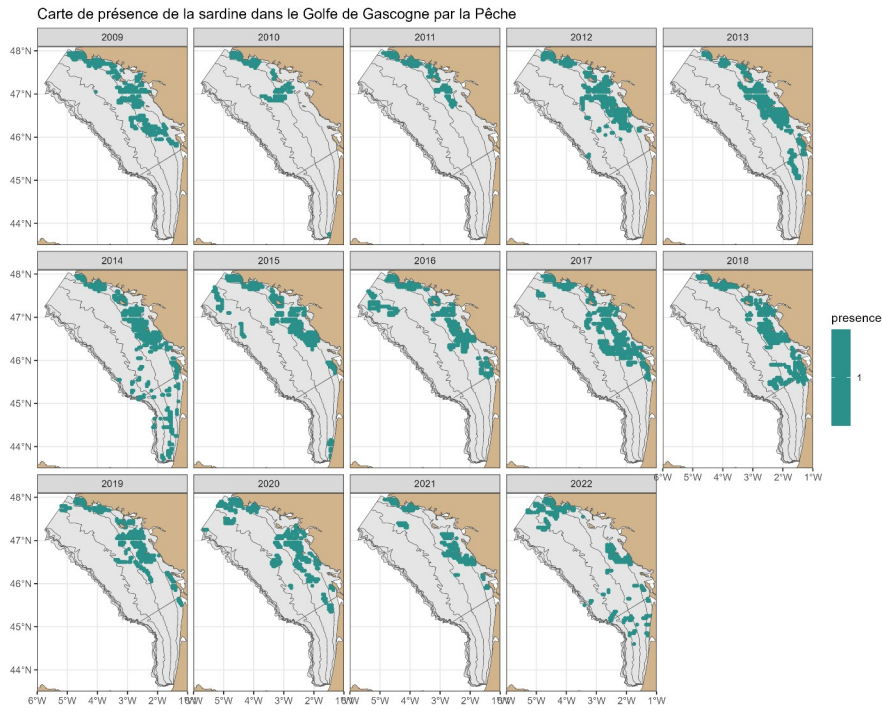


Figure 20: Spatial distribution of sardine presence commercial data (PS and PTM) in spring

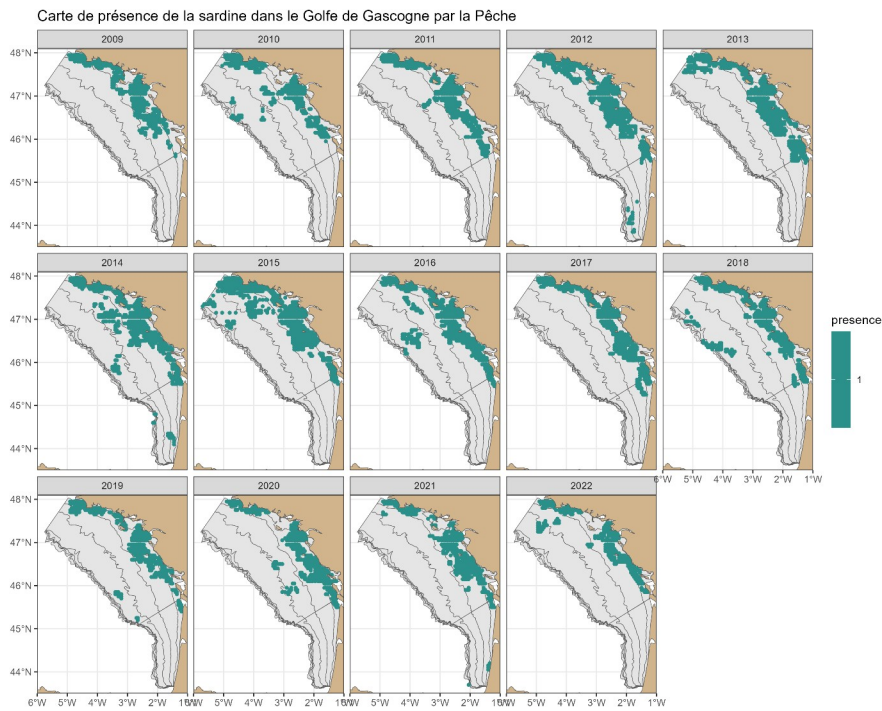


Figure 21: Spatial distribution of sardine presence commercial data (PS and PTM) in summer

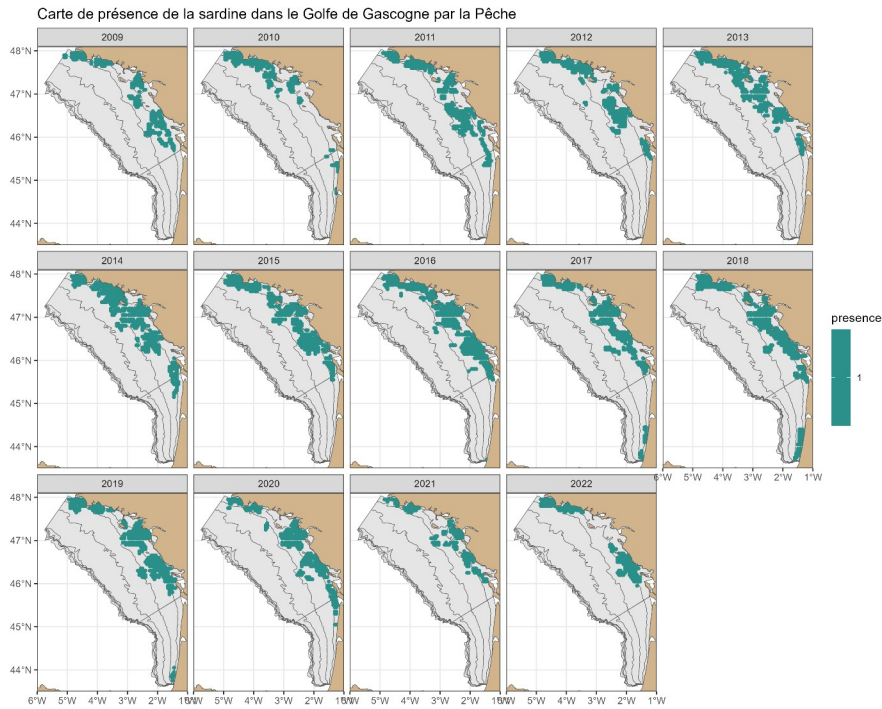


Figure 22: Spatial distribution of sardine presence commercial data (PS and PTM) in fall

Appendix 2: Prior: explanation and visualisation

Priors need to be specified in a Bayesian modelling framework. Priors can be constructively viewed as regularization devices (Gelman and Shalizi, 2013): they put soft constraints on the parameter space, and can help in obtaining better estimations (in a variance-bias tradeoff sense) esp. with sparse data.

We aimed at using so-called weakly-informative priors (Gelman et al., 2008), acknowledging that there is no precise definition of ‘weakly’. One important insight is that a prior that ‘looks’ uninformative (e.g. a uniform prior over a finite support) on a scale may not be ‘uninformative’ on another scale. The Figure 23 below illustrate this point with a prior on a logit scale commonly used in logistic regression. A large scale parameter is used (100) to translate naively an idea of vagueness (the corresponding variance is 100^2). However, this vagueness is actually very informative on the natural scale of a probability once the inverse logit transform is applied (Figure 23, right panel).

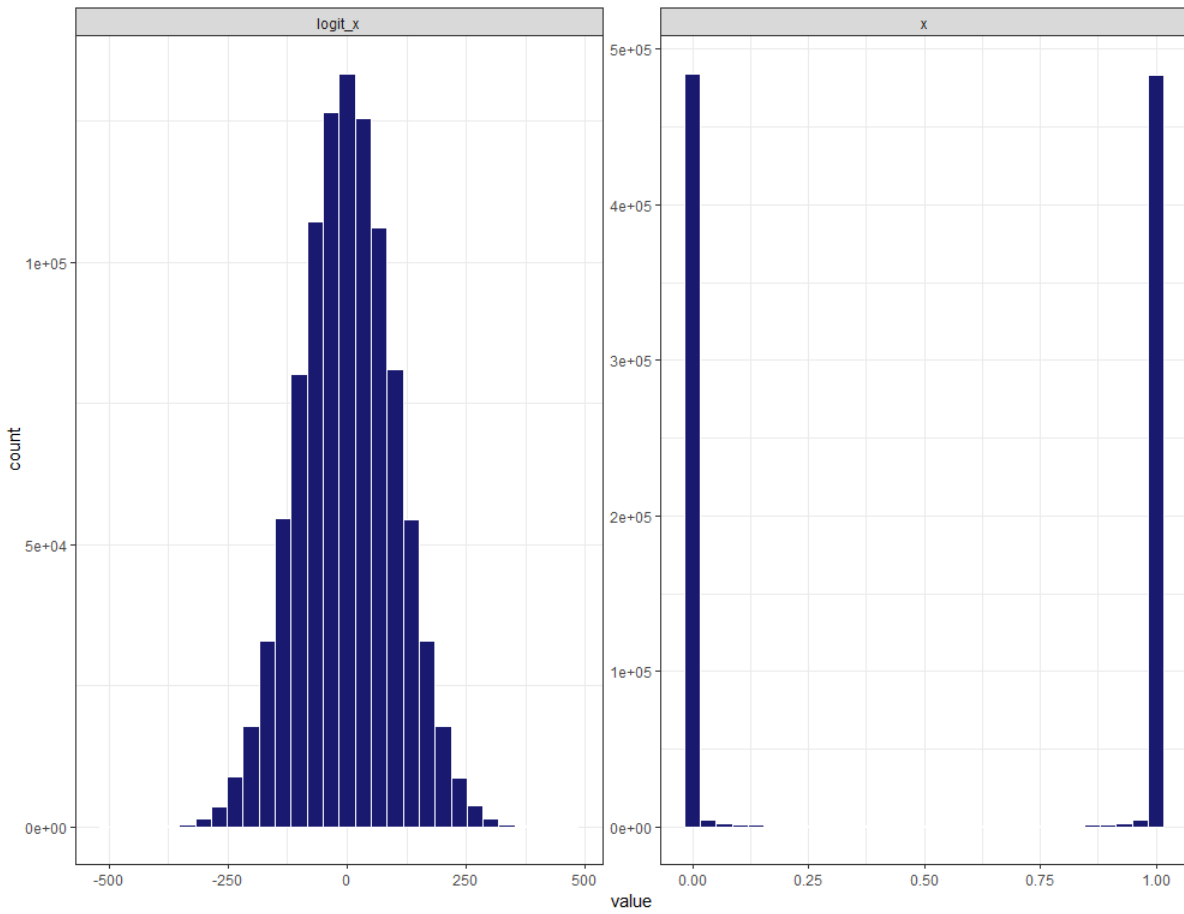


Figure 23: One million realization from a normal random variable $N(0, 100)$ represented as histograms with 30 bins. Left: on a logit scale. Right: on a probability scale.

Changing the scale parameter to be informative and a small value of 1.5 results in a weakly-informative prior on the probability scale (Figure 24).

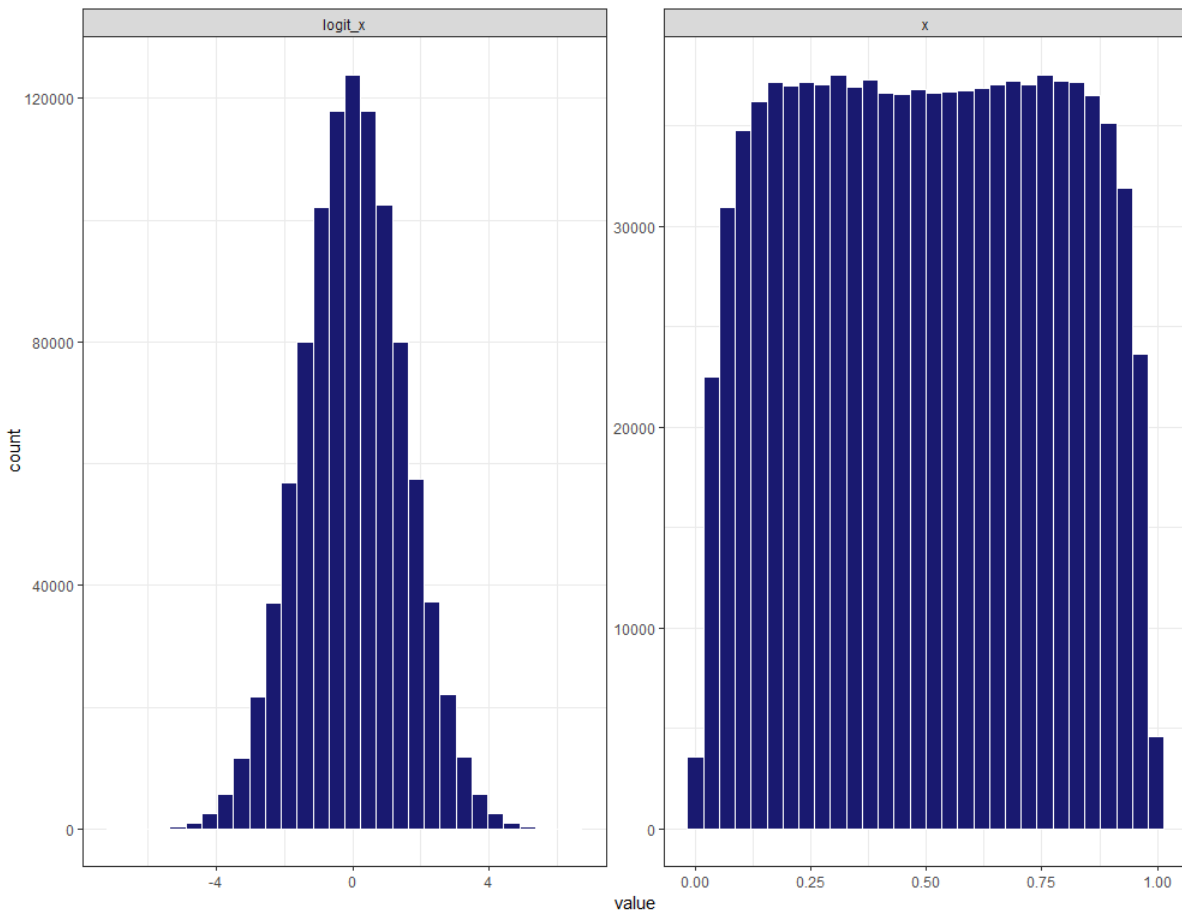


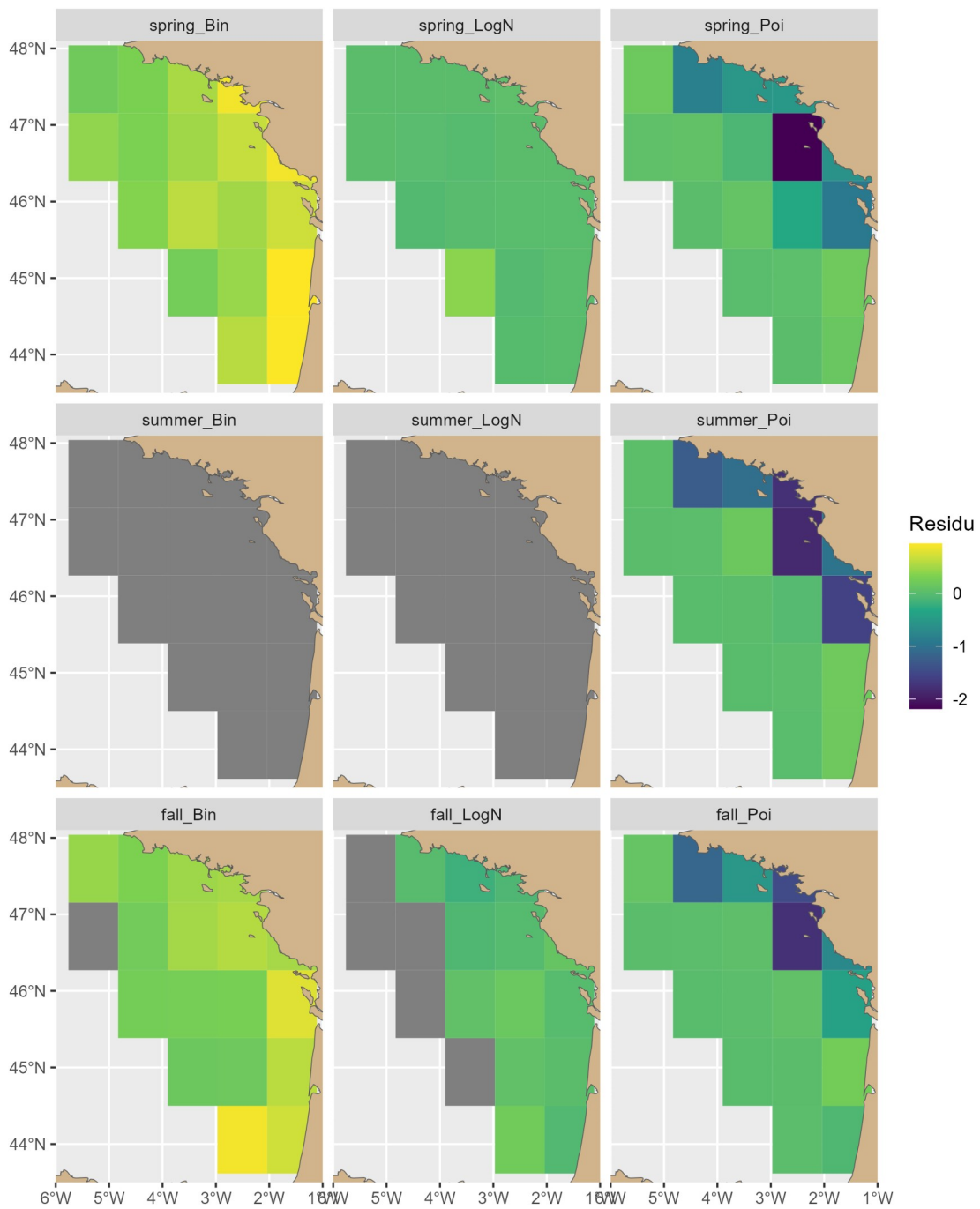
Figure 24: One million realization from a normal random variable $N\left(0, \frac{3}{2}\right)$ represented as histograms with 30 bins. Left: on a logit scale. Right: on a probability scale. Note the restricted range for the variable values on the logit scale compared to Figure 23.

Similar reasoning was applied for priors to be specified on a logarithmic scale. Prior specifications and choice were visualized to ensure that our choice were ‘reasonable’ in the sense of not putting too much weight on implausible regions on the parameter space.

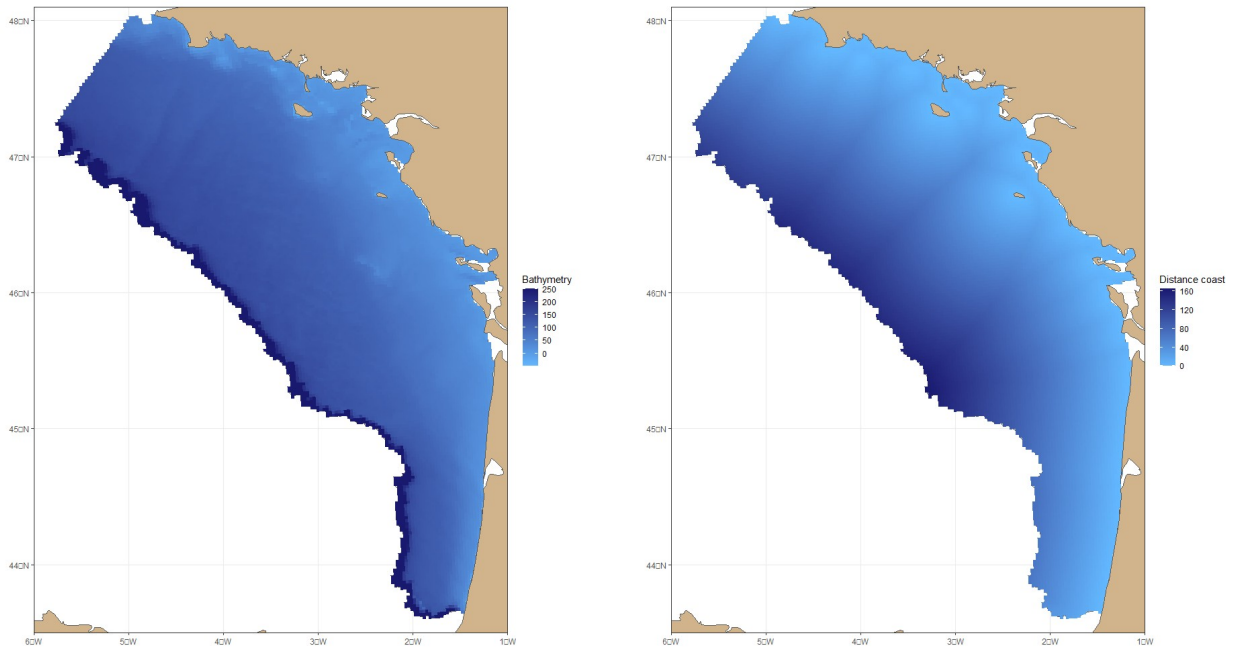
Appendix 3: Fit indices provided by INLA to check the convergence

The first is a logical indicator indicating whether the model fit was successful or not. The second component provides more specific information about the convergence of the mode-finding algorithm. It indicates whether the optimization algorithm used to find the mode of the joint posterior distribution was successful in converging to a valid solution. This number indicates the number of checks not passed during the posterior mode optimization for the hyper-parameters. A value of 0 means that the posterior mode of the hyper-parameters passed all the internal checks. Small values are usually acceptable, but not always, and higher values indicate bad convergence during the mode-finding process. If the Hessian matrix is not positive-definite, 1000 is added to the indicator as this is a significant issue that would require a rerun to locate the posterior mode.

Appendix 4: Grids maps of the deviance residual associated to each process from the model in 2017



Appendix 5: Maps of the bathymetry (right) and the distance from the coast (left) of the BoB



Appendix 6: Summary of the convergence assessment indicators

year	WAIC.null	Ok.null	mode.status.null	p.eff.null	WAIC.dist	Ok.dist	mode.status.dist	p.eff.dist
2009	39078,832	TRUE	3	1603,764	39057,405	TRUE	2	1599,052
2010	22002,965	TRUE	1	9088,384	21816,501	TRUE	0	8996,449
2011	-40347,575	TRUE	2	2017,966	-40381,505	TRUE	0	1990,798
2012	-3520,825	TRUE	0	2190,135	-3403,358	TRUE	1	2264,65
2013	18253,477	TRUE	0	2105,836	18239,526	TRUE	1	2081,456
2014	-3273,046	TRUE	1	2188,804	-1950,941	TRUE	0	2636,468
2015	-41460,404	TRUE	0	2242,735	-41453,114	TRUE	4	2214,165
2016	-12110,161	TRUE	0	2023,743	-12164,16	TRUE	9	2044,462
2017	-1369,609	TRUE	1	1829,151	-1249,915	TRUE	2	1908,232
2018	-13328,064	TRUE	0	2037,281	-13263,764	TRUE	0	2048,674
2019	-10690,468	TRUE	1	1705,003	-10756,276	TRUE	0	1659,562
2020	-21237,579	TRUE	0	1159,677	-21245,986	TRUE	0	1161,939
2021	2082,668	TRUE	0	1465,846	2090,3	TRUE	0	1468,263
2022	8169,698	TRUE	0	1392,419	8258,576	TRUE	0	1409,161

Appendix 7: Focus on 2017

The shared random field integrated information from all three likelihood processes. It represents the common spatial-temporal variability of the distribution of the sardine in the BoB that was not explain by the other effects.

The posterior estimates of the common spatial random effect revealed a globally coastal distribution for all the season (Figure 25). There was positive area between the Gironde estuary to south of Brittany. The pattern was the same over the season but the intensity was higher in summer.

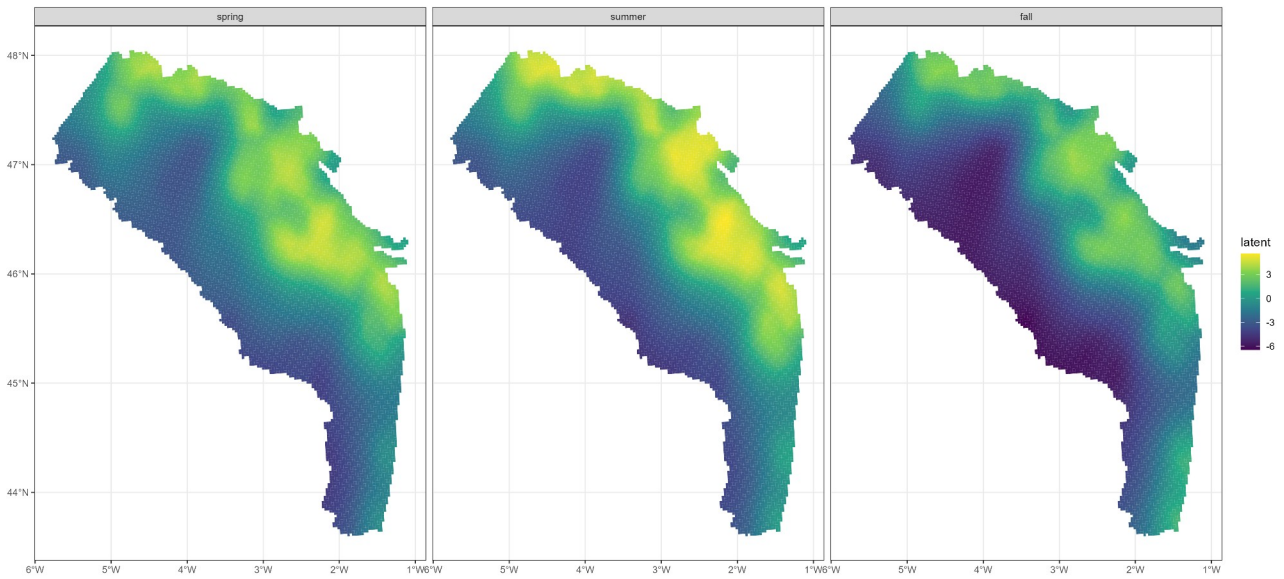


Figure 25: Posterior mean of the common spatial random effect in logarithm scale for 2017

Process specific random fields

The specific random fields represents the specific information brought by a single process, i.e. the spatial-temporal variability of the sardine distribution induced by the type of data and likelihood used. Those random effects explain the part of the distribution that was not captured by intercepts, the distance from the coast and the shared random field. We chose to focus on 2017 which correspond to a low value of presence intercept and a high value of fishery intercept.

The random field associated to the Binomial likelihood being informed by survey data only, the effect was null in summer. In spring, there was a high variability but the model predicted a positive effect in the offshore and southern parts of the BoB (Figure 26). This effect was not displayed by the shared random field. The addition of the specific random field allows to reconstruct the sardine distribution observed during the PELGAS 2017 survey, where sardines were observed near the coast, near the shelf break and in South Biscay in the BoB. In fall, the specific random field highlighted density hotspots near the coast, in central Biscay ($\sim 3^{\circ}\text{W} / 46.5^{\circ}\text{N}$) and in the southern tip of the Bay, that were present in survey data, but not in the shared random field (Figure 25).

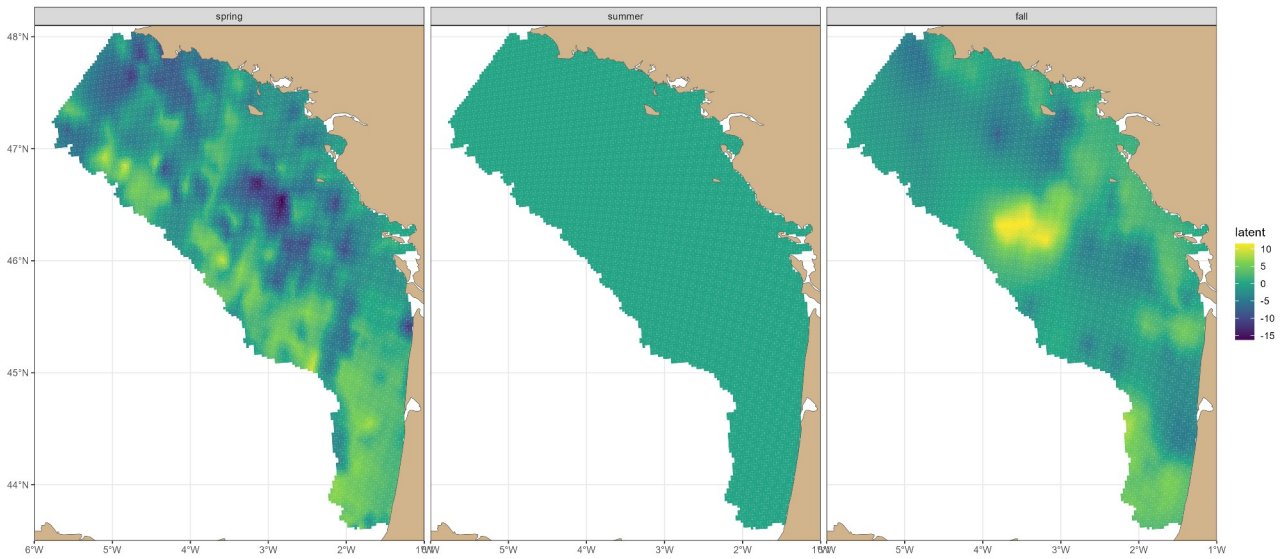


Figure 26: Posterior mean of the spatial random effect on the presence in 2017

The random field associated to the LogNormal likelihood was informed by survey data only, therefore the effect was null in summer (Figure 27). In spring, the presence and intensity random fields did not show the same patterns (Figure 27). The intensity random fields highlighted finer scale patches in southern coastal BoB and secondary near the shelf break, where sardine density was higher. Similar density patches can be observed in survey data.

In fall, a large part of the specific random field was null (zero biomass). It revealed 2 areas (in the middle and in the south of the study area) of negative values (Figure 27), corresponding to the highly positive areas as in the specific presence random field. Low biomass values were jence associated to high density of presence in fall 2017.

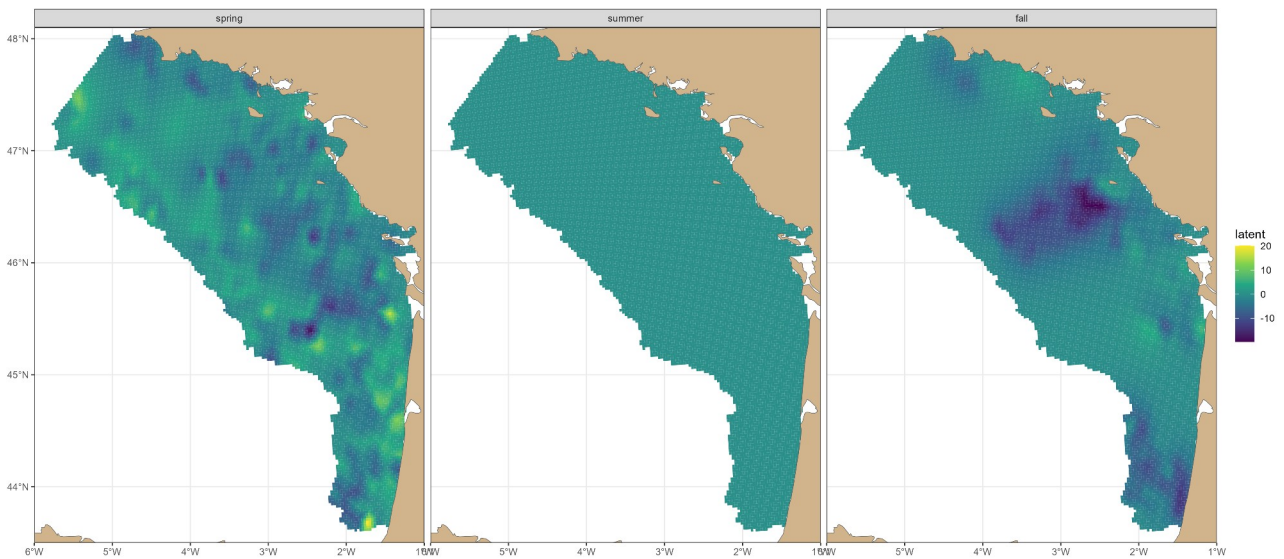


Figure 27: Posterior mean of the spatial random effect on the intensity in 2017

The random field associated to the Poisson Process likelihood was informed by the commercial fisheries data. In spring, the northern half of the BoB showed positive values and the south negative ones (Figure 28). There was an important area of positive in the middle of the study area. In summer, the positive values were coastal with a hotspot in front of the Gironde estuary (Figure 28). The offshore and southern part of the BoB showed lower probability of fishing presence. In fall, fishing hotspots were located along the Landes coast, South of Arcachon, in central BoB, South to Yeu island, and in Southern Brittany (Figure 28). The fishing activity was more diffuse, with the same Southern positive area as in the shared random field (Figure 25).

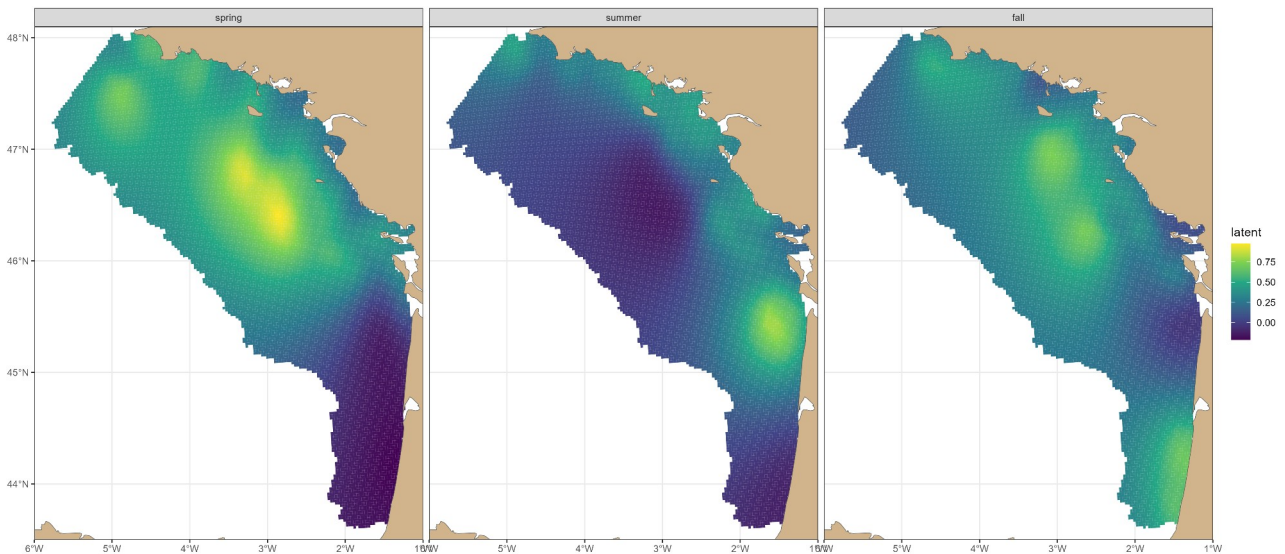


Figure 28: Posterior mean of the spatial random effect on the fishery presence in 2017

The value of the random effect fell between 1 and -0.25. The global fishing presence intercept in 2017 was very high (1,5) (Figure 9). The spatio-temporal variability represented by the random field was then lower than the global fishing intensity represented by the global intercept.

Model predictions

The predictions associated to the Binomial likelihood represent the probability of presence of sardines. In spring 2017, the predicted sardine presence map globally resembled the presence specific random field (Figure 26), but was more patchy (Figure 29). In fact, the variability of the presence random field was -15 to 15 (in log scale) (Figure 26) and only -5 to 5 in the case of the common random field (Figure 25). In fall, the sardine presence predictions map was less continuous, with well defined patches of sardine presence in centre BoB near and extreme North and South coastal areas and absence else where (Figure 29). The map resembled the presence specific random field (Figure 26, with sharper contours. The specific random field hence drove the

probability of presence prediction map for spring and fall. In summer, the only information was the shared random field which drove the predicted distribution of sardine presence (Figure 29).

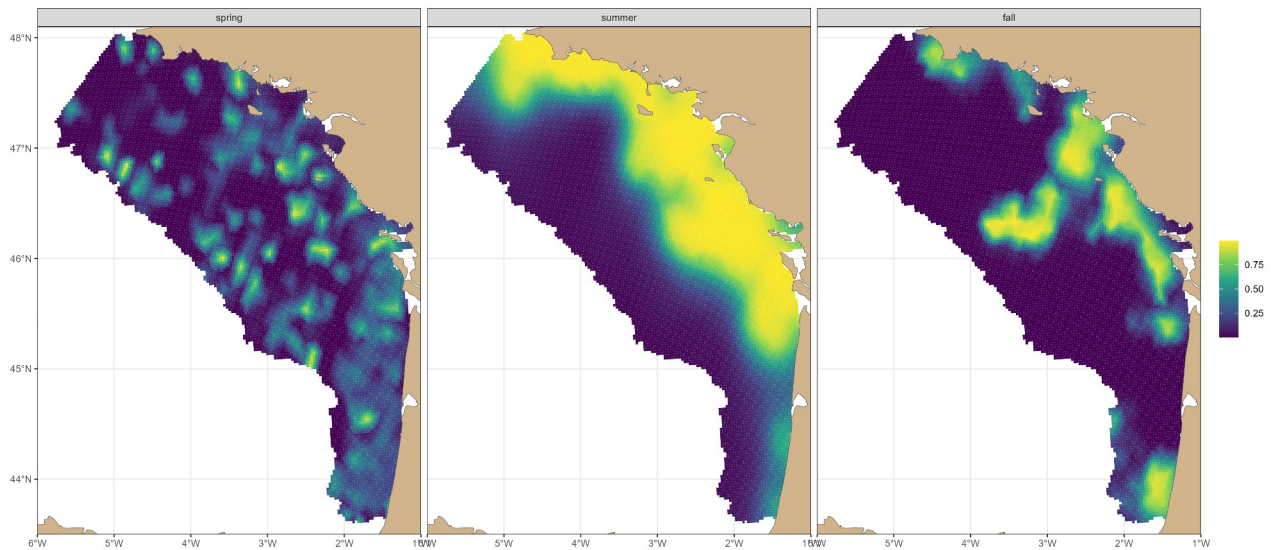


Figure 29: Probability of presence of sardines in 2017

The prediction associated to the LogNormal likelihood represents the intensity, or density of sardine in areas where they are present (Figure 30). Patterns of predicted intensity maps were similar to intensity specific random fields ones in spring and fall. Predicted intensities were however globally shifted relatively to specific random fields values, due to the addition of the seasonal intercept. In summer, the predicted intensity map was equal to the common random field, as no scientific survey data were available in this season (Figure 30).

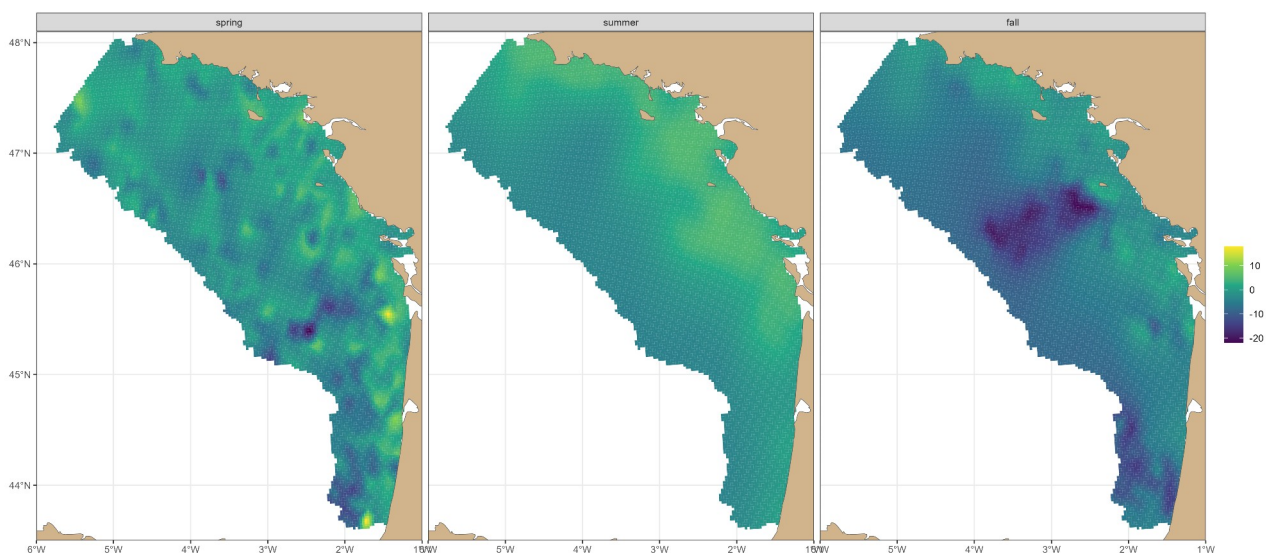


Figure 30: Intensity of sardines in 2017 in logarithm scale

The prediction associated to the Poisson point process likelihood represents the density of positive catches by commercial fishermen. This density of presence distribution was very close to the distribution given by the common random field (Figure 31). The fishing random field had values close to 0 (Figure 28). Its influence was very low compared to the common random field (Figure 25). So positive fishing operations were globally predicted to take place near the coast, with more activity in spring, except in the southern BoB in fall where a hotspot of fishing operations was predicted.

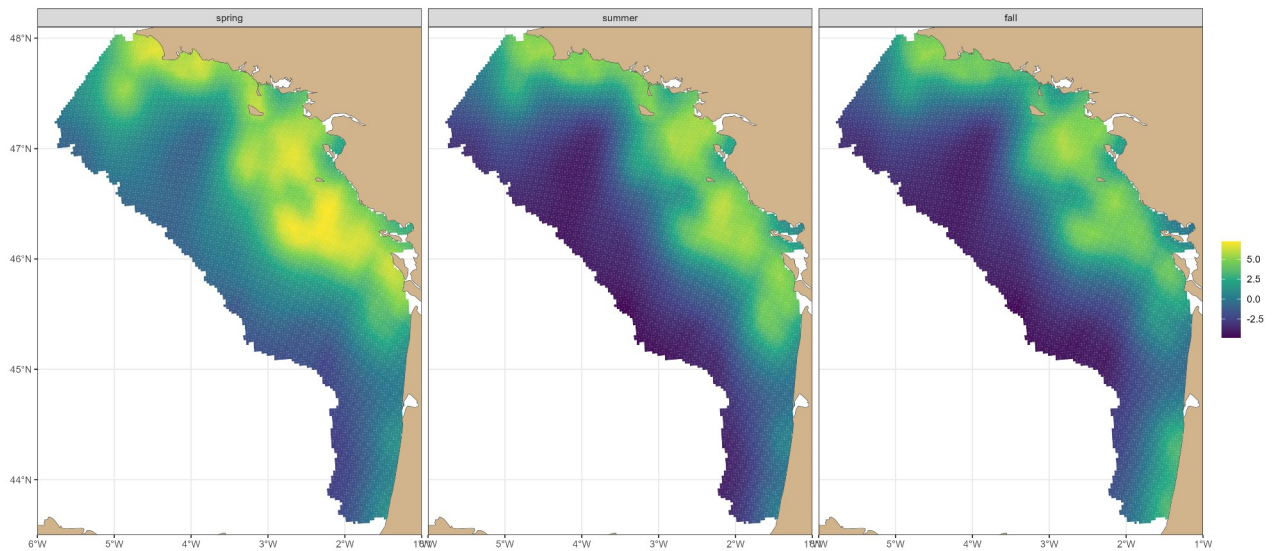
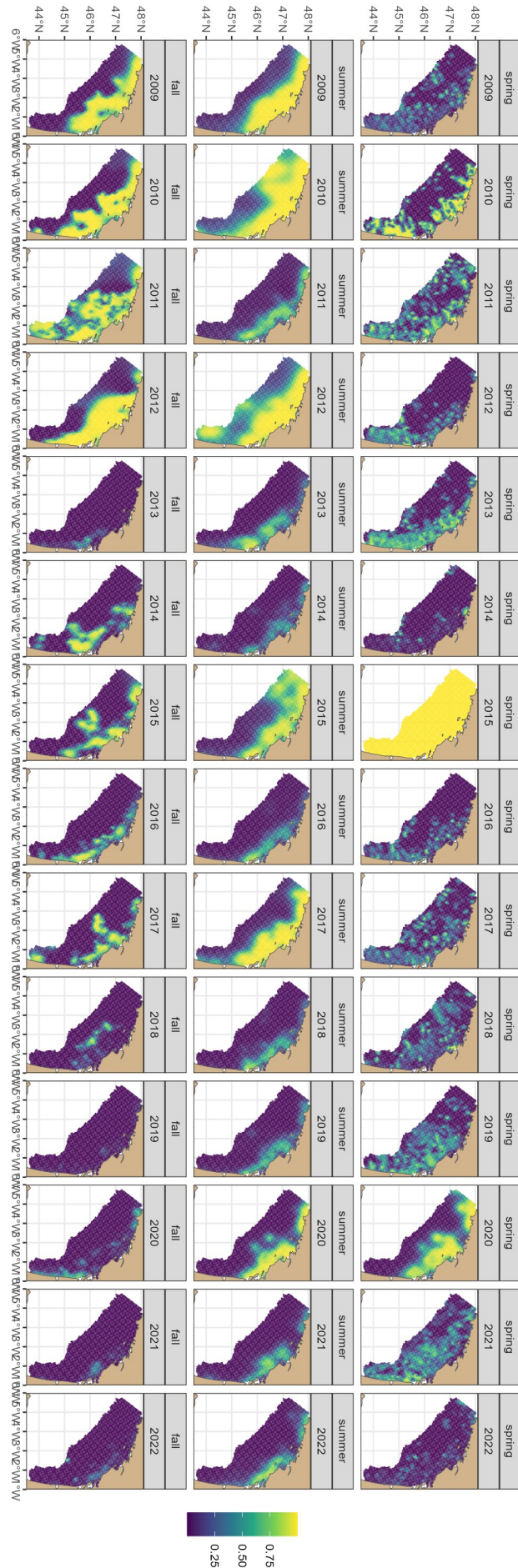
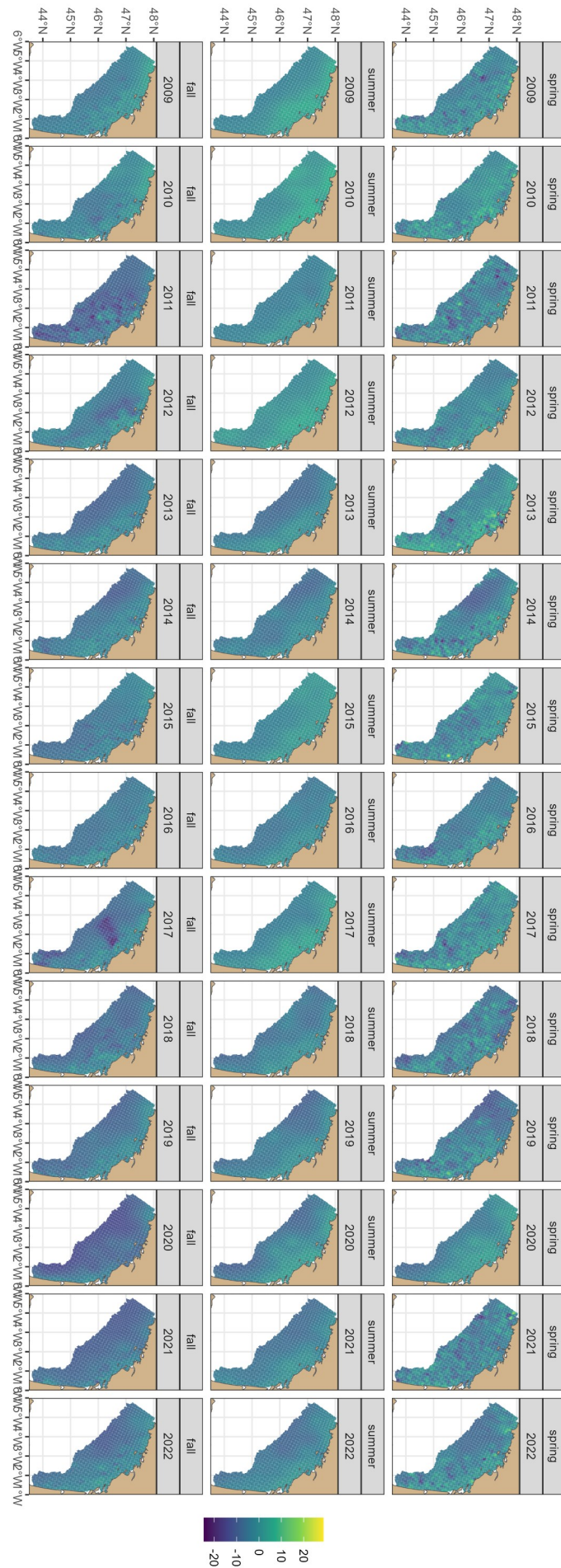


Figure 31: Density of presence of fished sardines in 2017 in logarithm scale

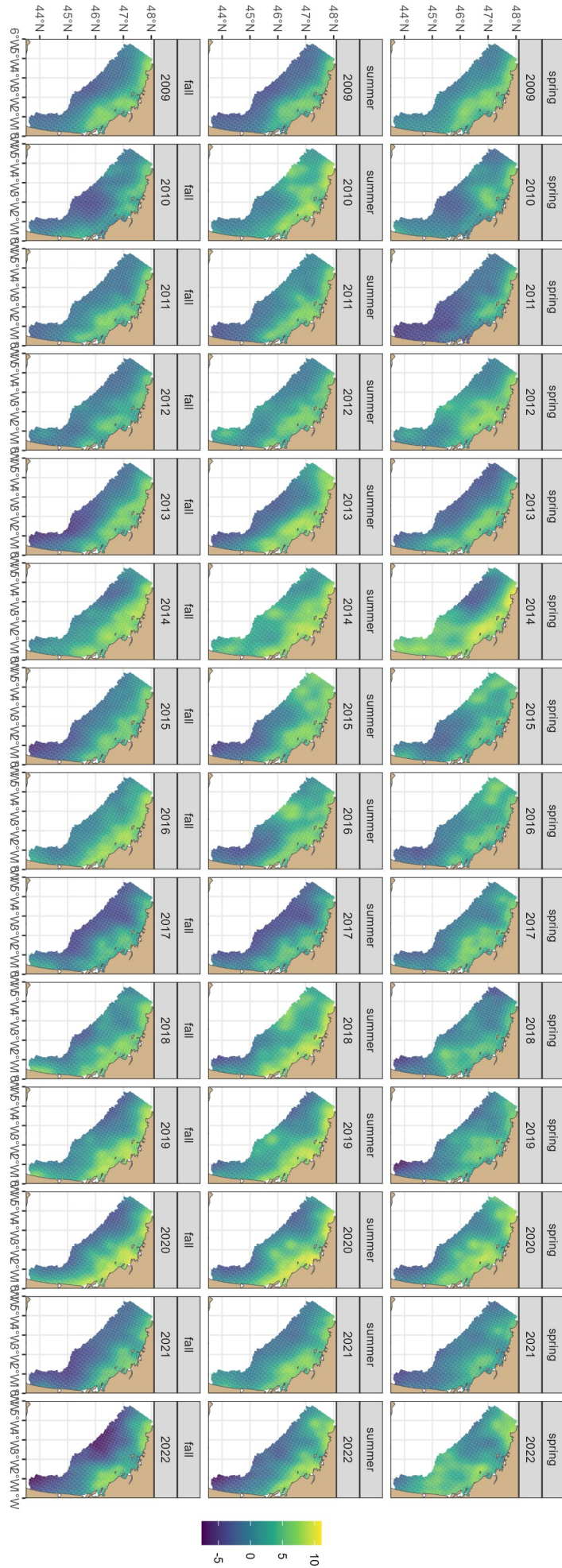
Appendix 8: Prediction of probability of sardine presence for all the year



Appendix 9: Prediction of sardine intensity for all the year



Appendix 10: Prediction of sardine fishery intensity for all the year



	Diplôme : Ingénieur Spécialité : Ingénieur Agronome Spécialisation / option : « Sciences Halieutiques et Aquacoles » (Ressources et Ecosystèmes Aquatiques) Enseignant référent : Pablo BROSSET	
Auteur(s) : Aurel HEBERT--BURGGRAEVE Date de naissance* : 13/04/2000	Organisme d'accueil : Ifremer – Centre Atlantique Adresse : 44311 Cedex 03, Rue de l'Île d'Yeu, 44980 Nantes	
Nb pages : 64 Annexe(s) : 10	Maître de stage : Mathieu DORAY, Matthieu Authier, Maxime Olmos	
Année de soutenance : 2023		
Titre français : Une approche d'intégration de données avec des modèles hiérarchiques pour caractériser les variations spatio-temporelles de la distribution des petits poissons pélagiques dans le Golfe de Gascogne		
Titre anglais : A data-integration approach with hierarchical models to reveal spatio-temporal variations of small pelagic fish distribution in the Bay of Biscay		
Résumé (1600 caractères maximum) : Le projet Delmoges vise à produire de nouvelles connaissances écologiques et halieutiques, pour réduire les captures accidentelles de dauphins dans le golfe de Gascogne (GdG). Notre étude vise à modéliser la variabilité saisonnière et inter-annuelle de la distribution des principales proies des dauphins, les petits poissons pélagiques, au moyen d'un modèle hiérarchique. Le modèle intègre 3 types de données : présence-absence et biomasse de poissons des campagnes scientifiques et présences de poissons par les données de pêche. Le modèle a fourni pour la première fois une description quantitative de la dynamique spatiale saisonnière de la sardine et des pêcheries associées. Les principales zones de distribution de la sardine étaient côtières, de l'embouchure de la Gironde au sud-ouest de la Bretagne. La sardine était en moyenne distribuée sur une plus large zone sous forme de petits agrégats au printemps. Elle semblait se concentrer dans ses zones de distribution principales en été et se disperser vers le large en automne, mais dans une moindre mesure qu'au printemps. Les pêcheries étaient concentrées dans les zones de distribution principale de la sardine à la côte, et étaient plus intenses en été, puis à l'automne et moins intenses au printemps. Après avoir intégré les données hivernales, l'application de notre modèle à l'anchois, et son couplage avec les distributions de dauphins communs devrait à terme permettre d'améliorer notre compréhension des interactions entre les petits poissons pélagiques et les dauphins et d'évaluer leur influence sur les captures accidentelles.		
Abstract (1600 caractères maximum) : The Delmoges project aims to generate new ecological and fisheries knowledge, in order to reduce dolphin bycatch in the Bay of Biscay (BoB). Within this framework, our study aims to model seasonal and inter-annual variability in the distribution of dolphins' main prey, small pelagic fish, using a hierarchical model. The model integrates 3 types of data: presence-absence and biomass of fish from scientific campaigns and presence of fish from fishing data. The model applied to GoG sardines has provided, for the first time, a quantitative description of the seasonal spatial dynamics of this species and associated fisheries. The sardine core distribution areas were coastal, from the mouth of the Gironde to south-west Brittany. On average, sardine distribution was most extended and patchy in spring. They seemed to concentrate in their core distribution areas inshore in summer, and to disperse offshore in autumn, but to a lesser extent than in spring. Fishing was concentrated in the sardine's main coastal distribution areas, and was more intense in summer, then autumn, and less intense in spring. After integrating winter data, applying our model to anchovy, and coupling it with common dolphin distributions, should ultimately improve our understanding of interactions between small pelagic fish and dolphins and their influence on dolphin bycatch.		
Mots-clés : Analyse bayésienne, statistique spatio-temporelle, modèle de distribution d'espèces, intégration de données, inférence INLA, petits poissons pélagiques, Golfe de Gascogne		
Key Words: Bayesian analysis, spatio-temporal statistic, species distribution model, data-integration, INLA inference, small pelagic fish, Bay of Biscay		