Investigating fish phenology and essential habitats through Empirical Orthogonal Functions

Baptiste Alglave[1], Maxime Olmos[2], Juliette Casemajor[2], Marie-Pierre Etienne[3], Etienne Rivot[4], Mathieu Woillez[2], and Youen Vermard[5]

[1]Université de Bretagne-Sud IUT de Vannes
[2]Ifremer
[3]IRMAR
[4]Agrocampus Ouest
[5]Ifremer Departement Ressources Biologiques et Environnement

March 11, 2024

Abstract

Fish spawning phenology is a major concern for conservation and fisheries management. New intensive data sources such as GPS-based tracking data or high resolution catch declaration data are progressively becoming available in the field of marine ecology. These benefit from high spatio-temporal resolution and open new research avenues to investigate inter-annual and seasonal variability of phenology. In this paper, we illustrate how catch declarations modeling coupled with spatio-temporal dimension reduction methods known as Empirical Orthogonal Functions (EOF) can be used to synthetize spatio-temporal signals in fish distribution; Specifically, we address the following questions; (1) can we identify spatio-seasonal patterns that can be interpreted in terms of seasonal migration between essential habitats? (2) can we identify changes in the phenology? (3) are those changes related to environmental drivers? The analysis is illustrated through the analysis of the reproduction phenology on three key commercial species in the Bay of Biscay (Hake, Sole and Sea Bass). The EOF analysis on these species emphasizes strong seasonal spatio-temporal patterns that correspond to migration patterns between feeding areas and reproduction areas. Based on this methodology, we identify seasonal variations in the timing of the reproduction and we relate these to Sea Surface Temperature, a key driver of fish reproduction.

Introduction

To complete their life cycle, fish require different habitats specific to different life stages (Harden, 1969). Those habitats, also known as Essential Fish habitats (Magnuson-Stevens Fishery Act, 2007) are associated with key demographic processes in the fish life cycle such as spawning, feeding and migrations and are characterized by a strong concentration of individuals within a spatially restricted area. However, rapid environmental changes may force fish to adapt, by tracking their essential habitat in space and time, and by changing the seasonal timing of their demographic processes (termed "phenology").

Understanding changes in phenology of demographic processes is critical for the management of fish population. Seasonal habitat utilization, timing of migrations and location of spawning areas are key knowledge to preserve fish essential habitats and ensure the renewal of marine resources (Delage and Le Pape, 2016; Lieth, 2013). For instance, areas where fish aggregate for spawning may require specific attention in terms of fisheries management (Biggs et al., 2021; Grüss et al., 2019). Also,marine Spatial Planning requires a good knowledge of fish essential habitats to implement offshore wind farms or limit the impact of marine aggregate extraction (Bastardie et al., 2015, 2014; Campbell et al., 2014).

Still the available data to investigate fish spatio-temporal demographic processes generally have sparse spatio-temporal coverage (*e.g.*scientific survey data, mark-recapture tagging data). Typically, scientific surveys usually occur once a year and provide samples only on the time span of the survey (Bastardie et al., 2015). Onboard observer data provide additional data on the whole year by recording fishing catches on a small portion of the commercial fleets (Rufener et al., 2021). With these data, it is possible to infer fish distribution at a seasonal or at a quarterly level at best (Kai et al., 2017; Olmos et al., 2023). However, this temporal resolution is generally not enough to investigate precisely the phenology of demographic processes that occur at a shorter temporal scale *e.g.* month, week (Biggs et al., 2021).

In the last decade, methods to combine fishermen declarations (logbook) with Vessel Monitoring System (VMS) data ('VMS x logbook' hereafter) have been developed to provide a fine scale information on fishing activity and fishing landings (Bastardie et al., 2010; Hintzen et al., 2012). In the last decade, 'VMS x logbook' data sources have been used to infer fish spatio-temporal distribution at a fine scale (Alglave et al., 2022; Azevedo and Silva, 2020; Dambrine et al., 2021; Murray et al., 2013). These data benefit from a high spatio-temporal resolution and consequently they open huge research avenues to investigate inter- and intra-annual variability of fish spatial distributions.

Recently, a modeling framework has been built (1) to integrate 'VMS x logbook' data from distinct fishing fleets to infer fish spatial distributions and (2) to handle preferential sampling of fisheries data (Alglave et al., 2022). The framework has been extended in time at a monthly time step. It has been applied to map fish aggregation areas to identify spawning grounds for a few key species of the Bay of Biscay (Alglave et al., 2023). Still, these approaches only investigate a small part of the time series: single year of data in Azevedo and Silva (2020) or a specific period over several years in Alglave et al. (2023). Consequently, they left apart the huge amount of information that can be extracted from the analysis of a long-term time series. One reason for this is the difficulty of simultaneously interpreting inter-, intra-annual and spatial variations in fish distribution.

Dimension-reductions techniques such as Empirical Orthogonal Functions (EOF - Hannachi et al., 2007; Lorenz, 1956) can provide insights into the spatio-temporal variability of fish population processes. EOF have been mostly used to characterize physical oceanography conditions. Some recent studies have investigated fish processes using EOF (Grüss et al., 2021; Petitgas et al., 2014; Thorson et al., 2020b, 2020a). However, to the best of our knowledge, previous studies on EOFs for biological processes have only aimed to synthesize the inter-annual variability of these processes and have never studied the intra-annual variability (*i.e.* phenology).

In this paper we aimed at demonstrating the potential of integrated spatio-temporal hierarchical models (ISTHM - Alglave et al., 2023, 2022) combined with EOF to:

- (i) identify spatio-seasonal patterns that can be interpreted in terms of essential habitats and migration between these habitats;
- (ii) infer temporal changes in phenology over long term time series;
- (iii) to explore the role of environmental drivers in controlling the phenology.

Taking sole, sea bass and hake in the Bay of Biscay as case studies, inferences derived from EOF analyses are compared to the literature. This allows to highlight the added value of our results with regards to the available knowledge of the location of spawning grounds, the intra-annual variability of spawning, and the environmental drivers of reproduction. Finally, we also expect that the EOF analysis will help to identify lesser-known or unknown essential habitats, such as feeding grounds.

Material and methods

Outline of the approach

Our approach includes different steps that are detailed hereafter:

- **Case studies and synthesis of the available knowledge on their phenology.** Sole, hake and sea bass are important fisheries of the Bay of Biscay. Based on a literature review, we provide expectations on the demographic processes, the essential habitat and the associated seasons to be compared with our results.
- **Inferring species distribution based on the ISTHM introduced by Alglave et al. (2022, 2023).** We rely on the framework developed by Alglave et al. (2023, 2022) to map the biomass of the mature fraction of the population for each species (hake, sole and sea bass) at a monthly time step over 2008 – 2018. The statistical approach integrates data from distinct trawler fleets that cover the whole Bay of Biscay.
- **EOF and clustering analysis of the model outputs.** To identify and visualize essential habitats and related seasons, we synthetize the temporal variation of the maps of abundance of mature fish through an EOF analysis (realized independently for each species) followed by a clustering analysis.
- **Investigating intra-annual variability of the demographic processes and relating phenological processes to environmental drivers.** Finally, we interpret the main modes of variability of the EOF with regards to adult reproduction phenology. We investigate intra-annual variability of reproduction and the drivers influencing reproduction timing.
- *Case studies description*

We selected three case studies that are important species in the Bay of Biscay and for which some knowledge on essential habitats is available but incomplete: **sole** , **hake** and **sea bass**(ICES 2020, 2022). Most of the literature on these species focus on spawning phenology (summarized in Figure 1). For **sole** , Arbault et al. (1986), Petitgas (1997) and Alglave et al. (2022) identified spawning grounds along the Bay of Biscay from January to March. For **hake** , Alvarez et al. (2004) provided similar analysis based on surveys conducted in the 90's and Poulard (2001) have investigated rough scale spatio-temporal distribution of hake based on logbook data. For **seabass** , recent analyses have investigated the spawning area and timing based on 'VMS x logbooks' data and provide information on phenology (Dambrine et al., 2021). Additional information on the adults feeding grounds is available for Sole (Figure 1).

Model structure and data to fit the model

Data and commercial fleets

We analyze catch per unit of effort (CPUE) of trawlers between 2008 – 2018, a relatively long period that allows to evidence intra- and inter-annual changes in species distribution and phenology.

As we only want to interpret the spatio-temporal dynamics of adult individuals, we filtered the mature fraction of the declarations by crossing catch declarations with the size distribution in each commercial category (see Alglave et al. (2023) for further details).

3

We selected the data of several trawler fleets as they benefit from a relatively opportunistic behavior, and they usually cover a wide area (Figure 2). Furthermore, their CPUE provides a good indicator of fish relative biomass (Hovgêrd et al. 2008). The selected fleets for each species are presented in Table 1.

Model structure and spatio-temporal resolution

To map the spatio-temporal distribution of the biomass of these different species, we used the framework developed in Alglave et al. (2023, 2022). The framework is a hierarchical integrated statistical model that combines multiple data sources to infer spatial distribution of fish density. The model is fitted to the data between 2008 and 2018 at a monthly time step on a 0.05° grid. It is structured in 3 layers:

(1) the latent field of relative biomass spatial distribution (the field we want to infer); (2) the observations layer; this layer can handle CPUE data from different fleets including the distinct catchability of the fleets; CPUE data are related to the same unique spatio-temporal field of relative abundance (3) unknown parameters, including the ones that control the shape of the biomass latent field;

We simplified the framework developed by Alglave et al. (2022) by ignoring the preferential sampling of fishermen. Indeed, previous results by Alglave et al. (2022) have shown that preferential sampling of trawlers is low. Considering it will therefore only slightly affect spatial predictions while strongly increasing the computation burden (Alglave et al., 2022).

EOF to identify essential habitats and to highlight changes in phenology

EOF Basics: a gentle overview

EOF was initially developed by Lorenz (1956) for weather forecasting. The broad idea is to generalize the classical dimension reduction techniques like Principal Component Analysis to spatio-temporal dimensions. EOF seeks to summarize the information brought by a set of spatio-temporal maps into a smaller set of maps that best describe and summarize the spatio-temporal patterns.

Let's defined $S(x,t)$ a biomass field defined at a time step t ($t = \{1, \ldots, T\}$) and spatial cell x, and the centered field of biomass $S^*(x,t) = S(x,t) - S(x,\cdot)$ (with $S(x,\cdot)$ the spatial average of $S(x,t)$). $S^*(x,t)$ is expressed as a linear combination of spatial patterns p_m (or maps, named EOF) related to temporal indices (or loading factors) $\alpha_m(t)$.

$$S^*(x,t) = \sum_{m=1}^{M} \alpha_m(t) \cdot p_m(x) \ ; \quad x \in \{1, ..., n\}, \ t \in \{1, ..., T\}, \ M \leq T$$

The loading factors $\alpha_m(t)$ and the spatial patterns $p_m(x)$ are defined to maximize the variation captured by the spatial patterns $p_m(x)$ and to ensure the spatial patterns and the loading factors are orthogonal between each other. The first spatial map $p_1(x)$ captures the biggest amount of spatial variation; the second spatial pattern $p_2(x)$ is orthogonal to the first one and captures the second biggest amount of spatial variation. In matrix terms, this falls back to a diagonalization problem and is equivalent to make a PCA analysis on a data frame where individuals are time steps and variables are locations (Lorenz, 1956). Classical PCA representation can be used to represent EOF results. Typically, the first two loading factors can be projected on the first two spatial patterns to get a visual representation of the spatio-temporal decomposition of the signal on the first plan of variability.

In practice, the diagonalization is performed through Singular Value Decomposition (Banerjee and Roy (2014). It is available in R through the function *svd* (R Core Team, 2023). Spatial patterns are normalized to 1 and loading factors are standardized by the square root of their eigenvalue.

Filtering EOF dimensions and locations of the spatial pattern

For each species, we filter the number of dimensions based on the graph of the variance explained by each dimension. As a commonly used empirical rule of thumb, we cut the graph at the dimension where there is

a drop in the variance explained. When plotting the spatial patterns, all the locations that contribute less to 1 / (number of grid cells over the spatial domain) are shaded to highlight the locations that contribute most to the variation. In standard PCA, this is equivalent to keeping only the variables (*i.e.* locations in our case) that explain or contribute more than a single variable (or location).

Identifying EOF results to phenological phases

Clustering analysis of EOF to identify seasons and essential habitats

To identify distinct essential habitats and to relate these with ecological season, a clustering analysis was performed on the loading factors and the EOF maps.

Through EOF or PCA, individuals (here time steps) and variables (here locations) are projected into two distinct spaces: the space of the individuals (time steps) and the space of the variables (locations). Often in standard PCA, clustering is realized in the space of the individuals only, but the same can be done in the space of the variables. While the first clustering allows to regroup individuals that have the same variable values, the second clustering allows to differentiate individuals.

This way, the clustering will regroup locations that have similar temporal patterns and time steps that have similar spatial patterns. Clusters of locations will be interpreted as distinct essential habitats and clusters of time steps will be interpreted as ecological seasons.

We performed clustering based on a Hierarchical Clustering on Principal Components (HCPC) through the package FactoMineR (Lê et al., 2008).

Relating EOF with literature knowledge

Most of the available knowledge available for our species focuses on spawning habitats (Table 2; Figure 1). By crossing the spatio-temporal patterns in the EOF with this knowledge it is possible to identify EOF principal components that spatially match with the spawning grounds and verify if the temporal dynamics of the associated loading factor match with the expected seasonality of spawning. It is then possible to investigate the intra-annual variability and the environmental drivers of spawning.

Investigating inter-annual variability

Reproduction is known to face intra-annual variability and is partly driven by the temperature (Fincham et al., 2013; Huret et al., 2018). Specifically, for sole, hake and sea bass, some studies have investigated the relationship between reproduction timing and SST and have evidenced an optimal range of temperature for reproduction (references are in Table 2).

We hypothesized that the peak of the loading factor associated with the spawning season represents the peak of the reproduction season, and we investigate the inter-annual variability of the peak.

We check if the spawning peak identified from EOF matches with the period of temperature optimal range (*e.g.* see Figure 4). SST data were extracted from the Marine Copernicus platform (*https://marine.copernicus.eu/*).

Results

Dimension filtrations and extracting average spatial patterns: Pre-analysis of the EOF

For **sole** , we select the six first EOF dimensions that capture 50% of the variance (Figure 3). For **hake** , we filter the two first dimensions that capture 30% of variance. For **seabass** , we select the first dimension only. It captures 30% of the variance. Other dimensions are not considered in this analysis as they are considered as noise.

The averaged spatial distribution (denoted $S(x, \cdot)$ in EOF equations) reveal specific average patterns for each species (Figure S1). EOF results presented in figure 4 have to be analyzed relatively with their average spatial pattern. For **sole** , average distributions are relatively coastal with high biomass offshore the Gironde Estuary

(2°W - 45°N). For **hake** , average spatial distribution is more offshore and corresponds to the slope area. For sea bass, the mean pattern is very coastal. Biomass is high along the Vendée coast (2°W-46°N to 3°W-47°N) with a hotspot near Belle Île (3°W - 47°N), and along the Landes coast (1.5°W-44°N to 1.5°W-45.5°N).

Identifying the essential habitats and associated seasons

Taking sole as illustration, we perform a clustering analysis on the loading factors (time steps) and eigen vectors (locations). This allows to identify several areas and seasons that characterize sole spatio-temporal dynamics (Figure 5, center and right). Six spatial clusters and three seasons can be identified (Figure 5 - Note that the clustering trees are available in Figure S3 and S4): (i) an area constituted by clusters 1-3 that is mainly correlated to winter months (November to February – Figure 5, left); (ii) a coastal area constituted by clusters 5 and 6 that mainly correlates to summer, autumn and early winter months (July to November, Figure 5, left); (iii) and an area constituted by cluster 4 that mainly corresponds to the average distribution for spring and early summer (March to June, Figure 5, left).

A deeper insight on the temporal dynamics of each cluster is given in the appendix (Figure S5). The clustering was also performed on the other species and are presented in the supplementary material (all figures after Figure S6).

Crossing the available knowledge with EOF to infer spawning phenology

All species present a strong seasonal pattern (Figure 4).

For **sole** , a periodic signal is revealed in the loading factors. Dimension 1 and 2 highlight high biomass in offshore areas in winter (December to April) and relatively coastal distribution in summer. EOF1 mainly captures the coastal and offshore seasonal migrations without highlighting spawning areas per se. Whereas in EOF2, offshore areas correspond to reproduction grounds highlighted in Figure 1. Also, PC2 maximums fall within the period where SST are favorable for reproduction. Then, for sole, dimension 2 seems to be the best descriptor for reproduction phenology: orange areas in EOF2 are spawning areas and PC2 maximums are spawning peaks. For **hake** , similar seasonal patterns can be evidenced in dimension 1 and 2. There are (1) shelf areas that are occupied during summer and (2) offshore areas on the edge of the shelf that are occupied during winter which coincides with spawning grounds from Figure 1. The maximum of PC1 falls within the period when SST is favorable for reproduction. EOF2 represents offshore and coastal seasonal migrations with less emphasis on reproduction. Hence, we consider EOF1 as the dimension that best corresponds to reproduction and we retain this dimension and the maximum values of PC1 to investigate the phenology of reproduction of hake.

For **sea bass** , the EOF 1 captures the variability off the central shelf of the Bay of Biscay *i.e.* off the Gironde estuary. The corresponding time amplitude showed a very strong seasonal pattern with high positive peaks occurring in January/February. These peaks match the period where SST is favorable for reproduction and the spawning areas from Figure 1.

Inter-annual variability of reproduction and relationship with SST

Our results also highlight inter-annual variability in reproduction phenology for the three species (Figure 6).

For **sole** , the months of reproduction identified through PC2 falls between January and March. In 2012/2013, reproduction seems to be a bit earlier; Maximum of PC2 is in December and falls outside the period where temperature is favorable for reproduction. When looking at the PC time series for 2012/2013 (Figure 4, PC2 for sole), the PC is flatter than for the other years and reproduction could also occur later on (there is another peak in March).

For **hake** , months of reproduction are a bit earlier and fall between December and February.

For **sole** and **hake** , both the reproduction period and the time range where temperature is favorable look to be relatively stable.

6

For **sea bass** , reproduction months emphasize more variability. The maximum of the PC1 time series falls between February and November specifically at the beginning of the time series. By contrast, the period where SST is favorable for reproduction is steady. This suggests that other covariates than temperature may strongly affect reproduction timing.

Discussion

'VMS x logook' data opens new gates to realize ecological analysis at a much finer spatio-temporal resolution than ever before (Azevedo and Silva, 2020; Gerritsen and Lordan, 2011; Murray et al., 2013), but still only few applications have evidenced this potential through concrete analysis on large time periods with massive amount of 'VMS x logbook' data.

In this paper, we combined an existing spatio-temporal model with a dimension-reduction approach (EOF) to investigate the phenology of three species of the Bay of Biscay (sole, hake and sea bass) based on 'VMS x logbook' data.

Here, 'VMS x logbook' data give access to monthly distributions. Combined with our modeling framework, it provides a way to analyze spawning phenology at a much finer temporal scale as other data sources (e.g., scientific survey) that would give access to quarterly distribution at best. Such modeling approach synthesizing inter and intra annual variability in spatial variation is a major result that should support a broader access to VMS data for science (Hintzen et al., 2012).

The need to combine with ancillary data to interpret the results

Our approach is not self-sufficient and relies partly on the availability of expert and/or literature knowledge to interpret the spatial patterns identified through EOF as ecological processes. EOF capture the main modes of variation of the spatio-temporal fields. Combining those results with ancillary data and knowledge about the timing of the main phenological events is needed to interpret the results.

Typically, for sole, some studies investigating the timing and the spatial distribution of egg and larvae distribution have been used to corroborate our interpretation of EOF (Arbault et al., 1986; Petitgas, 1997). For hake, survey data have recorded mature individuals during the spawning season (Alvarez et al., 2004) and catch declarations data have been used to evidence reproduction migration at the level of rough statistical rectangles (Poulard, 2001). They were proven to be consistent with our analysis too.

Still, those data are limited. In both cases, the data were restricted to few samples from very old spring surveys and they have not been updated since this period. Furthermore, in many other cases there might not be any reference data to corroborate the EOF results. One should encourage to increase the quantity of available data by conducting new surveys to obtain direct observations of spawning areas (Fox et al., 2008). However, this requires important investments and depends on the ability of institutions and government to raise funds for these surveys.

Alternatively, expert knowledge of fishermen could be a valuable complement to interpret the main modes of variation (Yochum et al., 2011) when other data are missing. Bezerra et al. (2021) and Silvano et al. (2006) proved the usefulness of fishermen knowledge to determine the temporality of fish spawning and to identify some spawning grounds by crossing the information of aggregation areas provided by several fishermen.

Enhanced EOF method to better disentangle ecological processes

An extensive literature has focused on EOF and has outlined both its usefulness and efficiency for dimensionality reduction but also its limits. Several authors have flagged that extra care needs to be taken when interpreting statistical structure as being physical or ecological processes (Monahan et al., 2009). Typically, a single process can be reflected through several modes of the EOF. In our application, we decided to restrict the interpretation to the dimensions that best highlight the signal of reproduction. Still, some reproduction signals can be evidenced in the other dimensions of the EOF. Other methods (in general linear transformations of EOF) have been developed to better disentangle the underlying ecological process (Hannachi et al., 2007). For instance, rotated EOF allows to obtain more distinct patterns by relaxing orthogonality

constraints. However, the choice of the rotation criterion and the number of dimensions used for rotation have to be set arbitrarily which implies to make additional choices. Developing alternative methods could be of great interest to help disentangling ecological processes.

Limits of using the commercial catch data

The use of commercial catch declaration limits the possibility to analyze phenology relative to smaller size classes. In our approach, we filtered the mature component of the populations based on the size distribution of each commercial category as described in Alglave et al. (2023). A similar approach was used in Azevedo and Silva (2020) (though their approach was more refined as proportion by length was considered to vary by size category and by zone also) and allowed to map different age classes of horse mackerel. But a strong limit of our approach is that it is hardly applicable to smaller size classes. Indeed, commercial catch declaration data likely provide a biased picture of the spatio-temporal distribution of smaller fish size class. The minimum landing size is often pretty high and individuals below this size are either rejected or not declared (Lehuta and Vermard, 2023). Furthermore, VMS data are not available for vessels below 12 m that may represent an important part of coastal vessels fishing near juvenile habitats.

As a consequence, mapping juveniles' habitat is not possible with catch declaration data only. In this case, nursery surveys are still the only available data to investigate juvenile spatio-temporal dynamics and should be considered as reference (*e.g.* Nurse, Nourseine et Noursom - Delaunay and Brind'Amour, 2018).

Also, another drawback when using commercial data is the targeting behavior of fishermen. This can lead to biased spatial predictions and to overestimated estimates of biomass. Here, we neglected it as trawlers don't have a strong targeting behavior. Still, the framework developed in Alglave et al. (2022) could be useful in the case of stronger preferential sampling (see Quemper, 2021).

Investigating the effect of environmental drivers on phenology

Temperature is an important factor of fish reproduction (Huret et al., 2018). Other covariates may also strongly affect the timing of reproduction. Especially for species such as sea bass that have part of their life cycle in the pelagic realm. Other factors such as the salinity or concentration in chlorophyll A could strongly affect reproduction timing. Literature does not explicitly mention any clear threshold for these covariates and in that case more extensive field sampling would be required to identify the determinant of reproduction and include these in our approach - see for instance Planque et al., (2011).

Furthermore, it is expected that in a changing environment, the time span of reproduction will shift following temperatures. Some results already illustrate such phenomenon for sole in the North Sea, the English Sea and the North East Channel (Fincham et al., 2013). We consider our approach as an interesting tooI to assess the effect of climate change on fish phenology in order to preserve fish essential habitats and ensure the renewal of marine resources in a context of rapid environmental changes.

Supplementary material

All the supplementary material documents are available at the online version of the manuscript.

Acknowledgment

The authors are grateful to the Direction des pêches maritimes et de l'aquaculture (DPMA) and Ifremer (Système d'Informations Halieutiques - SIH) who provided the aggregated VMS data. The findings and conclusions of the present paper are those of the authors.

Author contributions

All authors contributed to the conceptualization and methodology of the study. All authors contributed to analysis of findings as well as drafting and revising the manuscript.

8

Data availability statement

Logbooks and VMS data are confidential data and they are available on specific request to DPMA. Model outputs for the different species are available at *https://zenodo.org/records/10517096*. Codes are available at the link:*https://github.com/balglave/C4-EOF.git*

Conflict of interest statement

The authors have no conflicts of interest to declare.

References

Alglave, B., Rivot, E., Etienne, M.-P., Woillez, M., Thorson, J.T., Vermard, Y., 2022. Combining scientific survey and commercial catch data to map fish distribution. ICES Journal of Marine Science fsac032. https://doi.org/10.1093/icesjms/fsac032

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. Canadian Journal of Fisheries and Aquatic Sciences.

Alvarez, P., Fives, J., Motos, L., Santos, M., 2004. Distribution and abundance of European hake Merluccius merluccius (L.), eggs and larvae in the North East Atlantic waters in 1995 and 1998 in relation to hydrographic conditions. Journal of Plankton Research 26, 811–826.

Arbault, P.S., Camus, P., le Bec, C., 1986. Estimation du stock de sole (Solea vulgaris, Quensel 1806) dans le Golfe de Gascogne à partir de la production d'œufs. Journal of Applied Ichthyology 2, 145–156. https://doi.org/10.1111/j.1439-0426.1986.tb00656.x

Azevedo, M., Silva, C., 2020. A framework to investigate fishery dynamics and species size and age spatio-temporal distribution patterns based on daily resolution data: a case study using Northeast Atlantic horse mackerel. ICES Journal of Marine Science 77, 2933–2944.*https://doi.org/10.1093/icesjms/fsaa170*

Banerjee S, Roy A (2014) Linear Algebra and Matrix Analysis for Statistics, 1st edition. Chapman and Hall/CRC, Boca Raton Fla.

Bastardie, F., Nielsen, J.R., Eigaard, O.R., Fock, H.O., Jonsson, P., Bartolino, V., 2015. Competition for marine space: modelling the Baltic Sea fisheries and effort displacement under spatial restrictions. ICES Journal of Marine Science 72, 824–840. https://doi.org/10.1093/icesjms/fsu215

Bastardie, F., Nielsen, J.R., Miethe, T., 2014. DISPLACE: a dynamic, individual-based model for spatial fishing planning and effort displacement—integrating underlying fish population models. Canadian Journal of Fisheries and Aquatic Sciences 71, 366–386.

Bastardie, F., Nielsen, J.R., Ulrich, C., Egekvist, J., Degel, H., 2010. Detailed mapping of fishing effort and landings by coupling fishing logbooks with satellite-recorded vessel geo-location. Fisheries Research 106, 41–53.

Bezerra, I.M., Hostim-Silva, M., Teixeira, J.L.S., Hackradt, C.W., Félix-Hackradt, F.C., Schiavetti, A., 2021. Spatial and temporal patterns of spawning aggregations of fish from the Epinephelidae and Lutjanidae families: An analysis by the local ecological knowledge of fishermen in the Tropical Southwestern Atlantic. Fisheries Research 239, 105937.

Biggs, C.R., Heyman, W.D., Farmer, N.A., Kobara, S., Bolser, D.G., Robinson, J., Lowerre-Barbieri, S.K., Erisman, B.E., 2021. The importance of spawning behavior in understanding the vulnerability of exploited marine fishes in the US Gulf of Mexico. PeerJ 9, e11814.

Campbell, M.S., Stehfest, K.M., Votier, S.C., Hall-Spencer, J.M., 2014. Mapping fisheries for marine spatial planning: Gear-specific vessel monitoring system (VMS), marine conservation and offshore renewable energy. Marine Policy 45, 293–300. https://doi.org/10.1016/j.marpol.2013.09.015

Dambrine, C., Woillez, M., Huret, M., de Pontual, H., 2021. Characterising Essential Fish Habitat using spatio-temporal analysis of fishery data: A case study of the European seabass spawning areas. Fisheries oceanography 30, 413–428.

Delage, N., Le Pape, O., 2016. Inventaire des zones fonctionnelles pour les ressources halieutiques dans les eaux sous souveraineté française. Première partie: Définitions, critères d'importance et méthode pour déterminer des zones d'importance à protéger en priorité (Rapport de recherche). Pôle halieutique AGRO-CAMPUS OUEST, Rennes.

Delaunay, D., Brind'Amour, A., 2018. Campagnes « Nourriceries ».

Fincham, J.I., Rijnsdorp, A.D., Engelhard, G.H., 2013. Shifts in the timing of spawning in sole linked to warming sea temperatures. Journal of Sea Research 75, 69–76.

Fox, C.J., Taylor, M., Dickey-Collas, M., Fossum, P., Kraus, G., Rohlf, N., Munk, P., van Damme, C.J., Bolle, L.J., Maxwell, D.L., 2008. Mapping the spawning grounds of North Sea cod (Gadus morhua) by direct and indirect means. Proceedings of the Royal Society B: Biological Sciences 275, 1543–1548.

Gerritsen, H., Lordan, C., 2011. Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. ICES Journal of Marine Science 68, 245–252.

Gruss, A., Biggs, C.R., Heyman, W.D., Erisman, B., 2019. Protecting juveniles, spawners or both: A practical statistical modelling approach for the design of marine protected areas. Journal of Applied Ecology 56, 2328–2339. https://doi.org/10.1111/1365-2664.13468

Gruss, A., Thorson, J.T., Stawitz, C.C., Reum, J.C.P., Rohan, S.K., Barnes, C.L., 2021. Synthesis of interannual variability in spatial demographic processes supports the strong influence of cold-pool extent on eastern Bering Sea walleye pollock (Gadus chalcogrammus). Progress in Oceanography 194, 102569. https://doi.org/10.1016/j.pocean.2021.102569

Hannachi, A., Jolliffe, I.T., Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: A review. Int. J. Climatol. 27, 1119–1152. https://doi.org/10.1002/joc.1499

Harden, J., 1969. Fish Migration. Copeia 409–411.

Hintzen, N.T., Bastardie, F., Beare, D., Piet, G.J., 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. Fisheries Research 115–116, 31–43. https://doi.org/10.1016/j.fishres.2011.11.007

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. Progress in Oceanography 166, 30–40.

Kai, M., Thorson, J.T., Piner, K.R., Maunder, M.N., 2017. Spatiotemporal variation in size-structured populations using fishery data: An application to shortfin mako (Isurus oxyrinchus) in the Pacific Ocean. Canadian Journal of Fisheries and Aquatic Sciences 74, 1765–1780.

Le, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. Journal of statistical software 25, 1–18.

Lehuta, S., Vermard, Y., 2023. Contrasting impacts of the landing obligation at fleet scale: impact assessment of mitigation scenarios in the Eastern English Channel. ICES Journal of Marine Science 80, 518–531.

Lieth, H., 2013. Phenology and Seasonality Modeling. Springer Science & Business Media.

Lorenz, E.N., 1956. Empirical orthogonal functions and statistical weather prediction. Massachusetts Institute of Technology, Department of Meteorology Cambridge.

Monahan, A.H., Fyfe, J.C., Ambaum, M.H.P., Stephenson, D.B., North, G.R., 2009. Empirical Orthogonal Functions: The Medium is the Message. Journal of Climate 22, 6501–6514. https://doi.org/10.1175/2009JCLI3062.1

Murray, L.G., Hinz, H., Hold, N., Kaiser, M.J., 2013. The effectiveness of using CPUE data derived from Vessel Monitoring Systems and fisheries logbooks to estimate scallop biomass. ICES Journal of Marine Science 70, 1330–1340.

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.

Petitgas, P., 1997. Sole egg distributions in space and time characterised by a geostatistical model and its estimation variance. ICES Journal of Marine Science 54, 213–225.

Petitgas, P., Doray, M., Huret, M., Masse, J., Woillez, M., 2014. Modelling the variability in fish spatial distributions over time with empirical orthogonal functions: anchovy in the Bay of Biscay. ICES Journal of Marine Science 71, 2379–2389. https://doi.org/10.1093/icesjms/fsu111

Planque, B., Loots, C., Petitgas, P., LINDSTRoM, U.L.F., Vaz, S., 2011. Understanding what controls the spatial distribution of fish populations using a multi-model approach. Fisheries Oceanography 20, 1–17. https://doi.org/10.1111/j.1365-2419.2010.00546.x

Poulard, J.-C., 2001. Distribution of hake (Merluccius merluccius, Linnaeus, 1758) in the Bay of Biscay and the Celtic sea from the analysis of French commercial data. Fisheries Research 50, 173–187.

Quemper, F., 2021. Modelisation de la distribution spatiale de la sardine du Golfe de Gascogne (Sardina pilchardus) par integration de donnees commerciales et scientifiques: enjeux et limites. Institut Agro, Rennes.

Rufener, M.-C., Kristensen, K., Nielsen, J.R., Bastardie, F., 2021. Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. Ecological Applications e02453.

Silvano, R.A., MacCord, P.F., Lima, R.V., Begossi, A., 2006. When does this fish spawn? Fishermen's local knowledge of migration and reproduction of Brazilian coastal fishes. Environmental Biology of fishes 76, 371–386.

Thorson, J.T., Cheng, W., Hermann, A.J., Ianelli, J.N., Litzow, M.A., O'Leary, C.A., Thompson, G.G., 2020a. Empirical orthogonal function regression: Linking population biology to spatial varying environmental conditions using climate projections. Global Change Biology 26, 4638–4649. https://doi.org/10.1111/gcb.15149

Thorson, J.T., Ciannelli, L., Litzow, M.A., 2020b. Defining indices of ecosystem variability using biological samples of fish communities: A generalization of empirical orthogonal functions. Progress in Oceanography 181, 102244. https://doi.org/10.1016/j.pocean.2019.102244

Yochum, N., Starr, R.M., Wendt, D.E., 2011. Utilizing fishermen knowledge and expertise: keys to success for collaborative fisheries research. Fisheries 36, 593–605.

Tables

Table 1: Trawler fleets selected to infer species distribution for each species.

Species	Sole	Hake	Seabass

Fleets	OTB_DEF_>=70_0 bottom trawl targeting demersal fish	OTB_DEF_>=70_0 bottom trawl targeting demersal fish	OTB_DEF_>=70_0 bottom trawl targeting demersal fish
	OTB_CEP_>=70_0 bottom trawl targeting cephalopods	OTB_CEP_>=70_0 bottom trawl targeting cephalopods	OTB_CEP_>=70_0 bottom trawl targeting cephalopods
	OTT_DEF_>=70_0 otter trawl fleet targeting demersal	OTT_DEF_>=70_0 otter trawl fleet targeting demersal	PTM_DEF_>=70_0 pelagic trawl fleet targeting demersal fish

Note that the name of the fleets (*e.g.*OTB_CEP_>=70_0) presents first the gear of the fleet (OTB for bottom trawl), then the species caught (CEP for cephalopods, DEF for demersal species) and the mesh size (_>=70_0 for 70 mm and above).

Table 2 : Optimal range of temperature for each three species

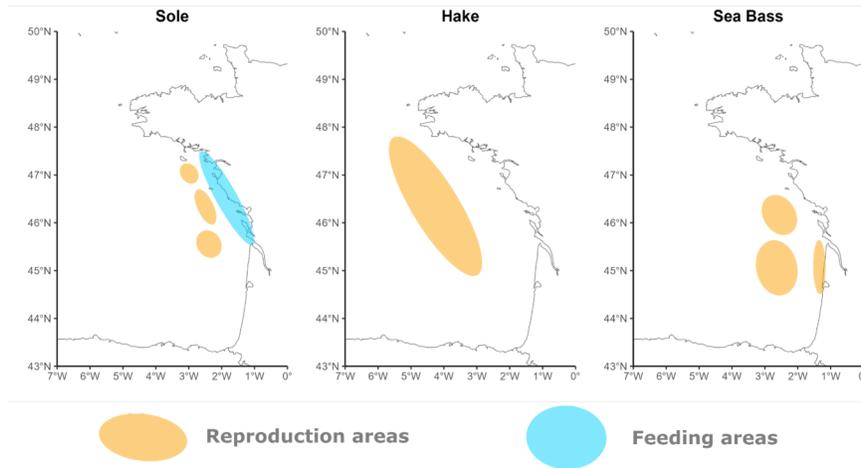| Species	*Sole*	*Hake*	*Seabass*
Optimal range temperature for reproduction	[10 ; 12.5°C]	[10 ; 12.5°C]	[11 ; 16°C]
Reference	Devauchelle (1986)	Murua (2010)	Devauchelle (1986)

Figures

Figure 1: Graphical synthesis of the available knowledge on essential habitats for the three species.
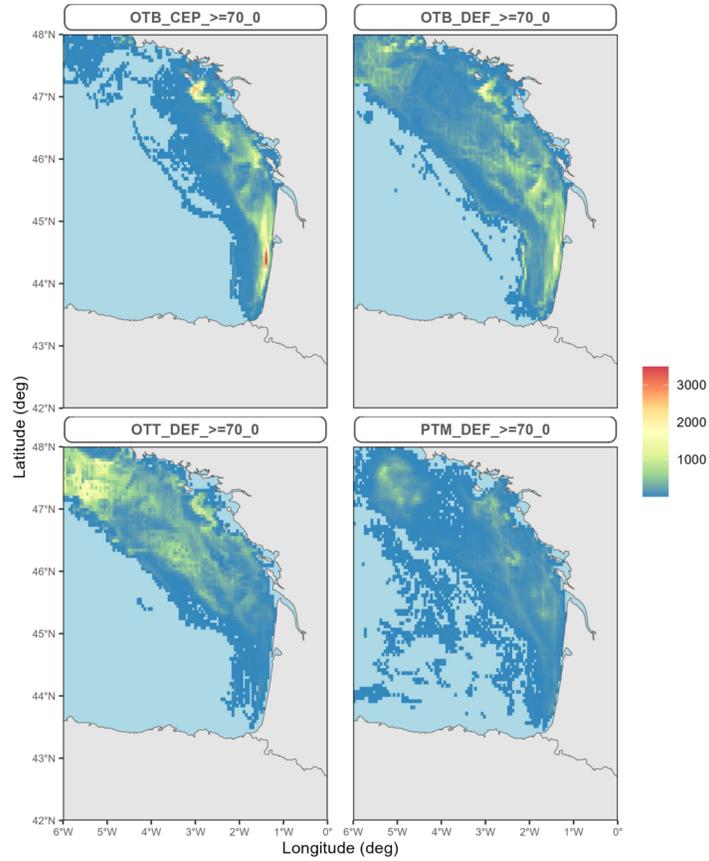
Figure 2: Spatial distribution of sampling effort (in hour) for each fleet aggregated over the period 2008 - 2018.
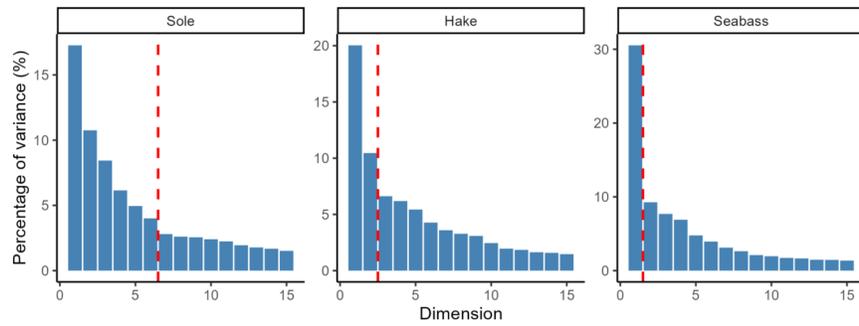

Figure 3: Proportion of the total variance explained by each dimension of the EOF for each species. Dashed line: threshold used to filter the interpretable dimensions of the EOF.
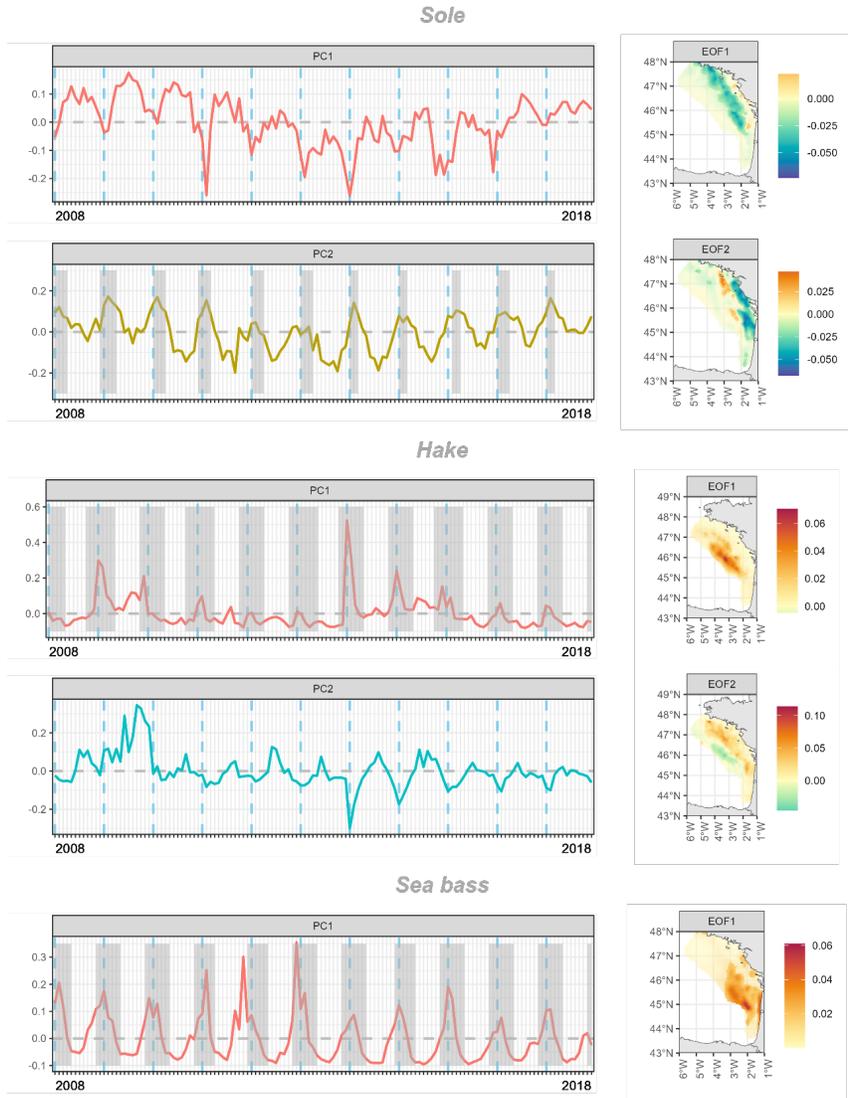
Figure 4: Loading factors (left) and EOF maps (right) for each species. For sole, only the two first dimensions are presented, the other dimensions are presented in Figure S2. The blue vertical line is the month of January for each year. The gray horizontal line is the 0 value. The gray bands are the period when temperature is within the optimal range of temperature for reproduction.
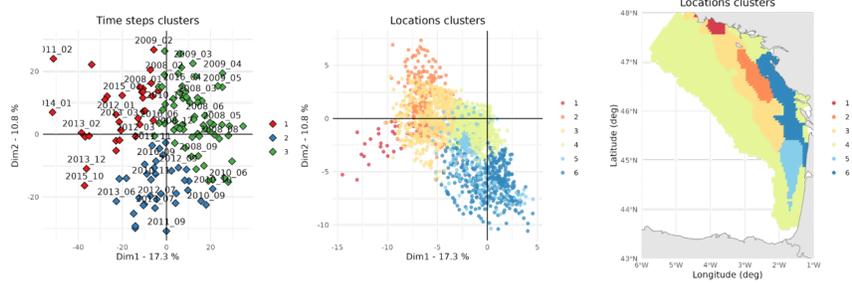
Figure 5: Sole case study. (Left) Projection of the loading factors on the two first dimensions of the EOF. Color: cluster identified through HAC analysis. Points on the left panel are identified by their time step (year/month). (Center) Projection of the eigen-vectors (EOF maps) on the two first dimensions. Color: cluster identified through HAC analysis. (Right) Spatial representation of the clusters.
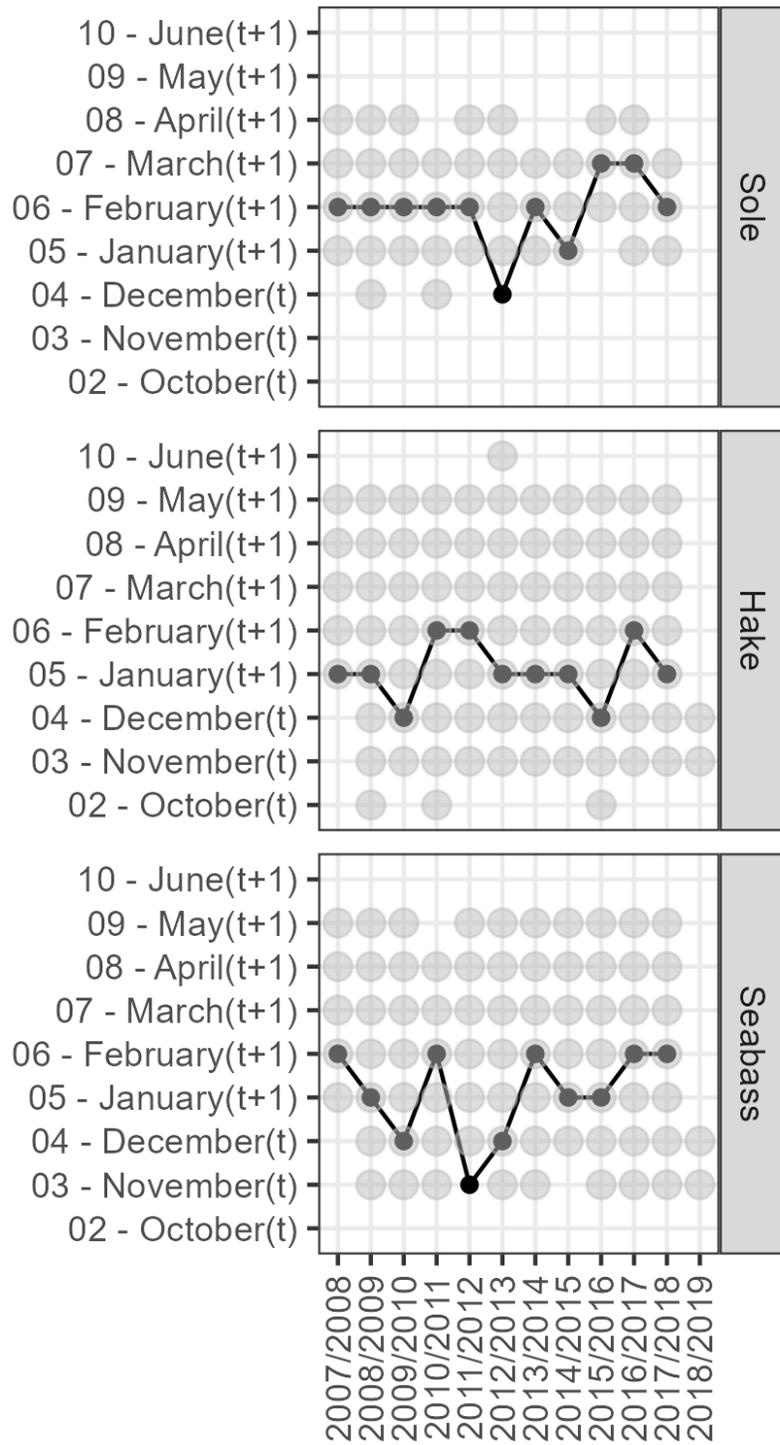
Figure 6: Variability of the period of reproduction identified based on the peak of the loading factor (Figure 4). For sole, we consider the maximum of PC2 (second dimension of the EOF). For hake, we consider the minimum of PC1 (first dimension of the EOF). For sea bass, we consider the maximum of PC1 (first dimension of the EOF). The gray dots are the months for which the average temperature is within the optimal range of temperature for reproduction.