Multi-species hotspots detection using self-organizing maps: Simulation and application to purse seine tuna fisheries management

Pauline Stephan, Daniel Gaertner, Ilan Perez, Loreleï Guéry

Received: 11 January 2022 | Accepted: 29 September 2022

DOI: 10.1111/2041-210X.14008

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Multi-species hotspots detection using self-organizing maps: Simulation and application to purse seine tuna fisheries management

Pauline Stephan[1,2,3] | Daniel Gaertner[2,3] | Ilan Perez[2,3] | Loreleï Guéry[4,5]

[1]ETH Zurich, Environmental Systems Science, Zurich, Switzerland

[2]MARBEC, University of Montpellier, CNRS, Ifremer, IRD, Sète, France

[3]Institut de Recherche pour le Développement (IRD), Sète, Cedex, France

[4]CIRAD, UMR PHIM, Montpellier, France

[5]PHIM, University of Montpellier, CIRAD, INRAE, Institut Agro, IRD, Montpellier, France

Correspondence
Pauline Stephan
Email: pauline.stephan14@gmail.com

Funding information
Fondation pour la Recherche sur la Biodiversite, Grant/Award Number: Gratification stages de master 2

Handling Editor: Aline Magdalena Lee

Abstract

1. To fulfil fisheries management objectives that often include implementing the precautionary and ecosystem-based approaches, multispecies fisheries data need to be analysed. Amongst the different methods dealing with these multidimensional data, self-organizing maps (SOMs) remain rarely used, although they are highly flexible in data input and offer visualization possibilities particularly suited to summarize complex datasets.
2. Here, we propose to combine SOMs with a clustering approach to break down data complexity and produce simple geographic maps showing catch hotspots, which can indicate sensitive zones in terms of fishery management. To promote this approach, we tested it first on simulated datasets and then on the open-access ICCAT commercial catch database of the tropical tuna fisheries of the Atlantic Ocean. We aimed to detect drifting fish aggregating devices (dFADs) catch hotspots of juveniles of two tuna species, bigeye and yellowfin tunas and of the silky shark, a commonly bycaught vulnerable shark species, in tropical tuna purse seine fisheries. Simulations on datasets increasing in complexity (in number, geographic and duration extent of the hotspots and number of species in the analysis) informed us about the method's sensitivity and limits.
3. Our findings showed that, in the context of multi-specific fisheries, the detection of the hotspot is dependent on a certain level of catch within the hotspot and that adding species to the analysis tended to mask small and short-duration hotspots. Applied to tropical tuna fisheries' data, the method confirmed the empirical knowledge on which first time-area closures were based and provided scientifical support.
4. All in all, the visual support provided by the method, its interpretability and its potential transferability to other fisheries' systems constitute its main strengths and imply a possible implementation in management decisions; specifically, as a tool to reach agreement between stakeholders in the definition of regulated areas for protecting juveniles of tunas and vulnerable associated species to dFAD practices.

KEYWORDS
bycatch, ecosystem-based approach, FAD, multispecies fisheries management, self-organizing maps, time-area closures, tropical tunas

1 | INTRODUCTION

To maintain fish stocks' exploitation rates at sustainable levels, a good knowledge of their ecology, and particularly of their spatiotemporal dynamics, is crucial. In multi-specific large-scale fisheries, species' distributions and their associative behaviours may change in space and time; hence disentangling seasonal or yearly dynamics and, in case of communities, isolating the key life stages of the target species and vulnerable species poses a great challenge to managers (Hilborn et al., 2021). To overcome the difficulties posed by the spatiotemporal dimension and the multi-specific characteristic of many fisheries, scientists have used different methods, including generalized linear mixed models (Kai et al., 2017), Bayesian modelling (Paradinas et al., 2015; Quiroz & Prates, 2018), joint dynamic species distribution models implemented in the vector autoregressive SPATIOTEMPORAL package (Guan et al., 2020), Poisson-link delta models (Grüss et al., 2019) or spatially explicit age-structured ecosystem models (Sibert et al., 2012). Such models are highly complex and require technical skills and consequential decisions from the users. They consequently prove to be poorly suited as decision tools for fishery managers and other stakeholders.

Aspiring to develop a method for hotspot identification easier to implement and to transfer to different systems and fishery managers, we here propose self-organizing maps (SOMs), an unsupervised artificial neural network (Kohonen, 1982). Their use is rising in ecological studies as they allow to explore patterns in large and complex datasets without a priori assumptions (Kalteh et al., 2008; Kangur et al., 2007; Park et al., 2006; Peña et al., 2015), yet they have been applied to few fisheries' studies until now (Conti et al., 2012; Hyun et al., 2005; Mendoza-Carranza et al., 2018; Russo et al., 2016; Simić et al., 2014). Amongst their main characteristics, SOM smooth data by filtering out noise. Furthermore, they sustain nonlinearity, high interaction between variables, outliers and zero values (Giraudel & Lek, 2001), amongst others typical attributes of commercial fishing data. Finally, they allow the user to visualize the multi-dimensional information on a two-dimensional plane, the 'map', easier to interpret by stakeholders and fishery administrators (Carlucci et al., 2018; Vesanto & Alhoniemi, 2000). In the case of spatiotemporal analysis, Andrienko et al. (2010) proposed to visualize the results either as space-in-time or time-in-space graphics. For species' abundance or catch hotspots within a defined geographic zone, these two possibilities translate into visualizing fixed spatial situations at a specific point in time or into visualizing changes over time in a specific place, respectively. The visual dimension is, beyond the flexibility in data input, undoubtedly the main appeal of SOMs.

In this multi-specific context, tropical tuna purse seiners in the Atlantic Ocean started deploying drifting fish aggregating devices (dFADs) in the early 1990s as it was known that tuna schools were aggregating under natural logs coming from river mouths (Arriz et al., 1999). Although setting the purse seine net (i.e. a fishing set) on a dFAD is less valuable in commercial size of tunas than on a free school, as it is composed mainly by juveniles of yellowfin (*Th. albacares*), bigeye (*Thunnus obesus*) tunas and all size classes of skipjack tunas (*Katsuwonus pelamis*), the rate of successful sets is higher (90% and 50%, respectively; Fonteneau et al., 2000). Further, the searching time to detect dFAD schools is reduced, particularly since the introduction of satellite-tracked buoys attached to dFADs in the late 1990s and then echo-sounders in 2010 (Baidai et al., 2020; Fonteneau et al., 2013; Torres-Irineo, Gaertner, et al., 2014). Their growing use represented 74% of total purse seiners catches in 2019 (ICCAT, 2021). Besides, the bycatch ratio, which includes associated non-tuna species, from other bony fishes to several vulnerable species as sharks, rays and turtles, is significantly higher in those sets (Torres-Irineo, Amandè, et al., 2014a). In the Atlantic Ocean, this fishery is regulated by the International Commission for the Conservation of Atlantic Tunas (ICCAT) and has been subject to varying static time-area closures of dFAD-fishing and other measures aiming at protecting juveniles of the main tropical tuna stocks since the late 1990s (ICCAT, 2020). However, these closures are not based on scientific advice and ongoing studies are evaluating their relevance.

In this study, we investigated the efficiency of SOMs combined with a clustering approach in detecting tuna juveniles' and vulnerable associated species' hotspots, candidate zones for time-area closures on dFAD activities (i.e. moratoria on dFAD). In a first step, we tested the SOM method's effectiveness with simulated mono- and multi-specific hotspots increasing in complexity both in number of species and level of abundances. In a second step, we applied it to real tropical tuna catch data and an index for juvenile silky shark (*Carcharhinus falciformis*) abundance derived from the literature (Lopez et al., 2020). More specifically, we asked (a) how large catches must be within a hotspot to be detected, particularly in relationship to other and to the same species in the surrounding areas and in the same temporal strata, (b) if both the number of species and the number of hotspots in the dataset play a role in the detection sensitivity, (c) if hotspots overlapping in time are correctly detected and finally (d) how to integrate a different type of data, for example, the case of bycatch of vulnerable species registered in general as presence/absence. It should be noted that the study was conducted in a practice-oriented mindset, exploring and refining the potential application of an already existing approach to a new purpose, rather than pursuing pure methodological objectives.

2 | MATERIALS AND METHODS

2.1 | Study context

Since the creation of the ICCAT in 1969, the contracting parties commit to cooperating in maintaining the populations of tunas and tuna-like species in the Atlantic Ocean and its adjacent seas

at sustainable levels for exploitation (ICCAT, 2019). Since 2015, they additionally commit to apply both ecosystem-based and precautionary approaches to their management practices, that is, to consider the ecosystem as a whole (including species either dependent on the stocks or negatively impacted by the fishing methods) and to base their decisions on best available scientific advice (ICCAT, 2015). To achieve those objectives, the commission started introducing and adapting restrictive measures in 1998, consisting progressively over time in a combination of catch limits, capacity limitations, limitations in number of dFADs by vessel, static time-area closures for dFAD-fishing, etc. (Figure 1). Only few studies until now have explored in a multispecies context the effectiveness of the implemented measures in protecting both juveniles' bigeye and yellowfin tunas and in reducing bycatches of vulnerable associated species (Escalle et al., 2017; Swimmer et al., 2020; Watson et al., 2009). Fonteneau et al. (2016) quantified the impact of the dFAD moratorium in place between January 2012 and February 2016 (Rec11, Figure 1) and found that, overall, this moratorium had limited results without significant reduction of European dFAD catches in 2013 and 2014. Furthermore, the authors did not observe major catches within the closure area after its reopening in March, and a 'fishing the line effect' emerged, that is, the catches around the borders of the closure were higher than usual. However, recent analyses on the Atlantic Ocean tuna tagging program's data suggest that this same moratorium was efficient for skipjack and yellowfin tunas, at least during closure months (Perez et al., 2022).

2.2 | Data

2.2.1 | Simulations

As an initial step, we simulated fisheries' data to test the robustness and sensitivity of the SOM method for the detection of dFADs catch hotspots in R (R Core Team, 2013). We generated five random datasets, that is, one per species (three targeted tuna species and two bycaught species), using the 'rtruncnorm' function from the TRUNCNORM package (Mersmann et al., 2018) that generates random values for the truncated normal distribution. For each species, we set the mean and standard deviation to the values from the ICCAT's tropical dFADs catch dataset to ensure that the simulations reflected a realistic case. In the ICCAT dataset by fishing mode, bigeye and yellowfin tuna made up between 10% and 20% of the total purse seiner dFADs catch, while skipjack made up for the rest. We were thus particularly interested in datasets where one species dominates the catch by far, and where that species is not the one considered overexploited, or at least not to be protected. For the generation of tuna hotspots, we manually increased the species' abundances in a selected geographic zone using the 'mutate' function in combination with 'case_when' function, both from 'dplyr' (Mailund, 2019). To identify the threshold of detectability, we did this gradually and separately for the juveniles of the two supposed overexploited, or close to overexploitation, tuna species (bigeye and yellowfin, respectively). Within the hotspot, we added between half and three times the mean catch by spatiotemporal unit (1° square × month, see

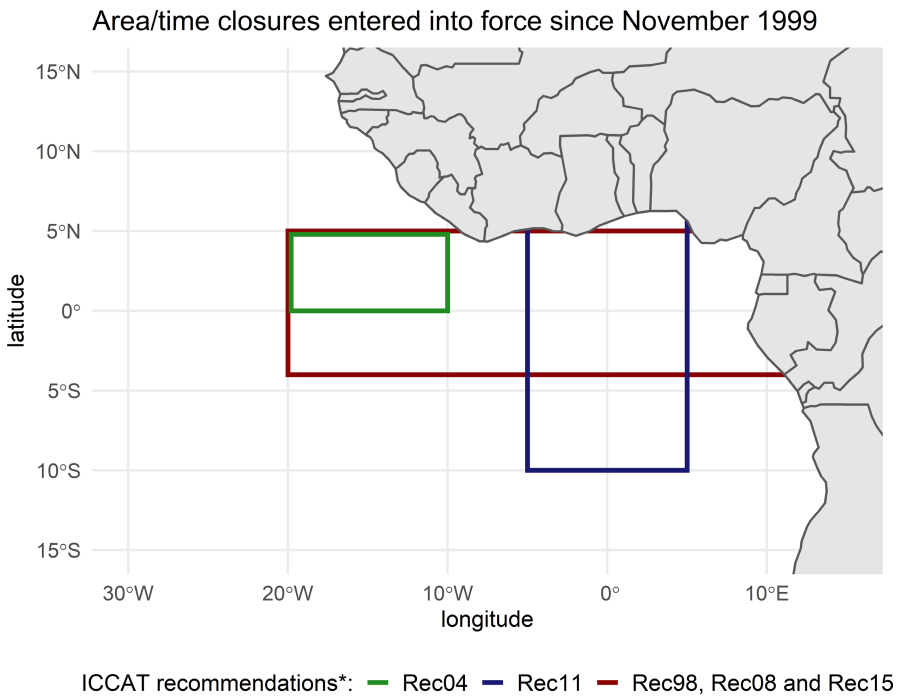


FIGURE 1 ICCAT recommendations since their introduction in 1999. FAD stands for fish aggregating device.

Table 3 and see the script 'DataSimulation_Over12months' on the stephanpau/SOM_TunaFisheries GitHub repository).

Moreover, with the aim to explore the robustness of SOM with regards to different data types, we simulated two associated bycatch species in dFAD operations whose estimated abundances were represented with indices ranging from 1 to 6, as catch of associated non-tuna species are not reported in the ICCAT database. We used 'rtruncnorm' again, yet with no reference dataset. We set the lower and upper limits to 0 and 6, and the means to 0.5 and 1.5, to have two different cases: a low occurrence species, and a species more likely to have high-abundance hotspots.

2.2.2 | Case study

With the aim to assess the performance of our approach on an open-access commercial catch dataset (available from: https://www.iccat.int/Data/t2ce_PS91-19_bySchool.7z), we applied the selected method to the tropical tuna fishery targeting tunas associated with dFADs in the Eastern Atlantic Ocean, as described above. To evaluate whether the dFAD moratorium's spatiotemporal strata may alter the detection of catch hotspots we ran our analyses on the period 1994–1998, before any time-area closure entered into force, and on the period 2012–2016, during an overall respected moratorium except for its first implementation year (Rec[11-01], Figure 2, ICCAT, 2020).

With regards to the juvenile silky shark, we created an estimated abundance index from the quarterly heatmaps that summarize the European observer programme's data (Spain and France) collected between 2003 and 2015 (Lopez et al., 2020). We recreated the paper's sampling cells using the 'SpatialPolygons' function from the SP package (Pebesma & Bivand, 2012). We then overlapped these large sampling zones with the ICCAT1 × 1° grid to assign a shark abundance index value to each of the ICCAT cells. We used the 'overGeomGeomDF' function from the RGEOS package (Bivand & Rundel, 2013). The shark data contained NA values, which eventually leads to rows being dropped during the SOM training. However, we kept them as NAs instead of setting them to 0 or to their neighbouring cells' values to illustrate the method's application to incomplete data. To check the robustness of this assumption, we ran the analysis once with NAs set to 0 and obtained overall the same results.

2.3 | Methods

2.3.1 | Self-organizing maps

SOMs are an unsupervised artificial neural network (Kohonen, 1982) used in ecology to detect patterns in complex datasets (Roigé et al., 2017). It consists of a two-dimensional map of a given number of hexagonal or rectangular nodes ordered in a grid. The network is trained on the data input matrix, constituted of input vectors

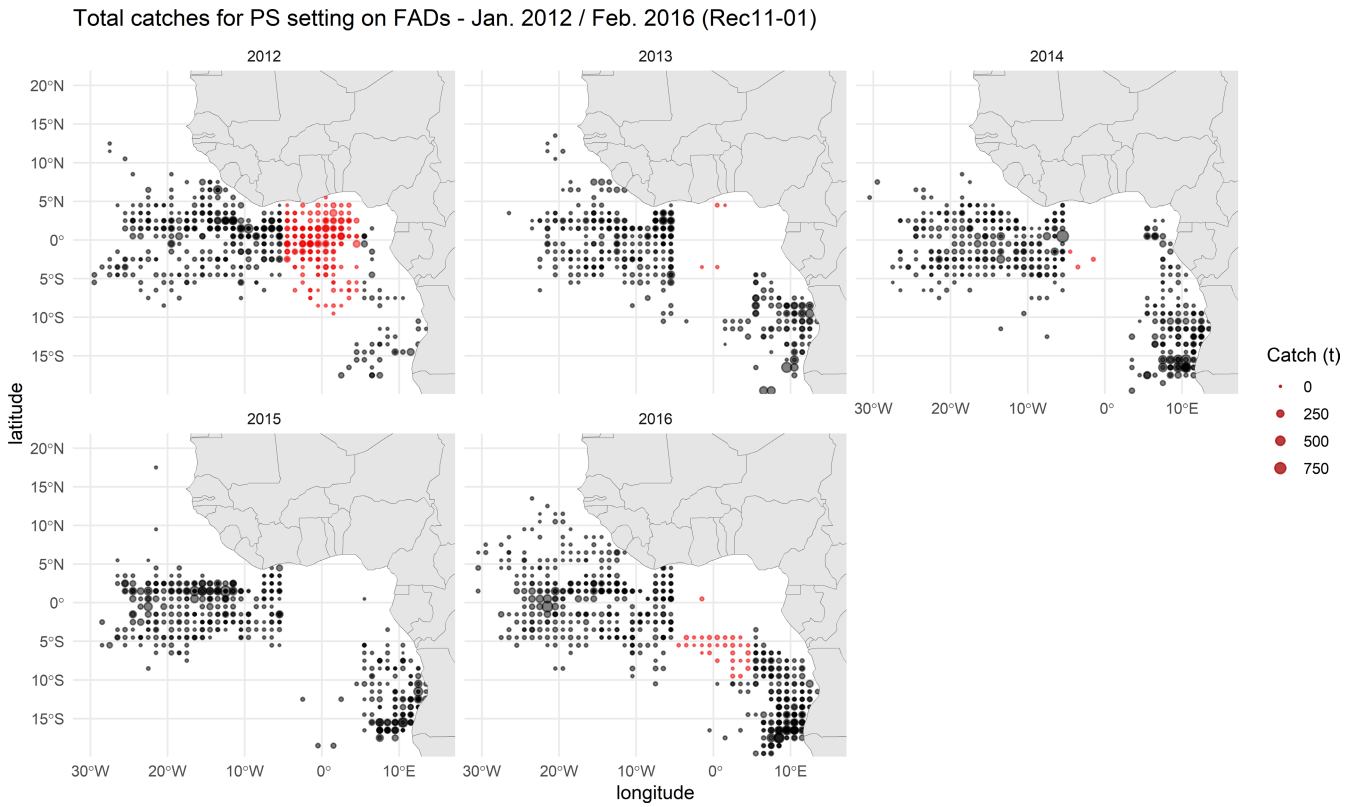


FIGURE 2 Purse seine tuna catch on dFADs during the moratorium period (rec[11-01], within (red) and outside (black) the time-area closure, source: ICCAT). The closure was less respected in the first and last years of its implementation.

of a given dimension. Each of the map's nodes is initialized with a 'codebook' vector of random values and of the same dimension as the input ones, a vector that will characterize the node throughout the analysis. The training then starts with a randomly selected input vector from the data input matrix. It is fed into the layer and a 'best-matching unit', that is, the node with the closest vector values, is selected. This node's codebook vector and its neighbourhoods are adjusted according to the neighbourhood radius to near the input vector. Vector by vector, the input matrix is fed in the network until convergence is reached (Chon, 2011). At the end of the training, each node is thus characterized by a codebook vector that now represents a few or more input vectors, that is, the codebook vector is very close to several ones from the input matrix. This repartition of observations within the two-dimensional network constitutes a first clustering step and its visualization already provides information on the ecosystem and especially on species' associations (Figure 3a–c).

With the aim to explore the robustness of SOMs to detect catch hotspots in presence of different data types, we performed two analyses: (1) by using an input matrix containing only the catch of each of the three species of tuna and (2) using an alternate input matrix to which we added the abundance indices of one (case study) or two (tests on simulated data, Table 3) associated pelagic species (Figure 3b). We explored different data pre-processing options that are explained in the 'Clustering' section below as those try-outs were combined with tests on the clustering method. In short, the results showed that centring and scaling the input matrix was the way to reduce the effect of the most caught species (otherwise skipjack catch drove the distribution of input vectors in the nodes; see Table 1; Figure S1 in the Supplementary Material for more details). We excluded log-transformation, as it emphasized low-abundance clusters, that is, 'cold spots' (Hyun et al., 2005).

We implemented the network training using the unsupervised 'som' function of the KOHONEN package (Wehrens & Buydens, 2007) in R (R Core Team, 2013). SOMs are trained in several steps and, for this function, require the users to define parameters and convert their data to a numeric matrix. The most influential parameters (Table 1) to define are the map size and its side ratios and consequently several heuristic rules have been proposed (Céréghino & Park, 2009; de Bodt et al., 2002; Park et al., 2006; Vesanto & Alhoniemi, 2000). The optimal grid size, or number of nodes, revolves around $5\sqrt{n}$ for n input vectors, and the optimal sides ratio approaches the ratio of the two highest eigenvalues of the input matrix. The learning rate *alpha* and the number of iterations *rlen* can be set for reasonable computation time and until convergence is reached (Andrienko et al., 2010).

2.3.2 | Clustering

Hierarchical clustering reduces the information to an amount easier to interpret and allows the geographical mapping of high-abundance clusters (Carlucci et al., 2018). The clustering algorithm is applied to the SOM's codebook vectors and can be visualized on the network (Figure 3e). To select the most reliable clustering approach, we first ran tests on simulated data with defined hotspots of one or two tuna species. Table 2; Figure S1 in the Supplementary Material summarize the conclusions of these tests. Ward's criterion, used for generating a dendrogram with the 'hclust' (R Core Team, 2016) function and already known to be more discriminatory than other criteria such as UPGMA (Mérigot et al., 2010), in combination with the 'cuTreeHybrid' function from the DYNAMICTREECUT package (Langfelder et al., 2008) performed best. This package was originally designed for genetic analyses and detects clusters based on the shapes of dendrogram's branches. The 'nbclust' function (Charrad et al., 2014), combining different methods to estimate the ideal number of clusters, was more dependent on data pre-processing and clustering criterion (see Figure S2). This has repercussions on the spatial distribution of the input data's assigned clusters (see script 'DataSimulation_Preprocessing_Clustering' on the stephanpau/SOM_TunaFisheries GitHub repository). We thus selected the approach combining the 'hclust' and 'cuTreeHybrid' functions.

After the clustering algorithm was applied to the codebook vectors and as each observation (i.e. each vector of tropical tuna catches for a 1° square × month cell) is assigned to a node and each node to a cluster, a cluster is assigned to each observation (Figure 3f). As each observation is identified in space and time, we can geographically map the clusters accordingly. We propose two options for visualization.

The first one consists in extracting only the clusters of interest, in our case, the high catch clusters, to facilitate the reading of the geographic maps and calculating the difference between medians of the clusters, to cut off above the highest distance to map only the ones above this limit (Figure 3g). We used the 'ave' function from the STATS package (R Core Team, 2016) to compute distance between medians of the clusters and filtered the clusters above the maximal distance. However, this method has limitations: it can happen that clusters that would have been included by a visual selection from the boxplots are excluded. Therefore, a careful attention to the boxplots is needed to avoid missing clusters.

The second option takes the form of a vulnerability categorized index from 1 to 10 for each cluster, easy to plot and to interpret (Figure 3h). This value between 1 and 10 is produced with the R package SCALES (Wickham, 2016), that is used to rescale the cluster's weighted mean catch or estimated abundance indices. Doing this by cluster rather than by observation ensured that the information from the SOM was not lost. When calculating the weighted mean catch, the users can set weight factors for each species according to their relevance in terms of fishery management, for example, depending on their change of stock status over time as observed in the Kobe plot (Kell, 2011; Walter et al., 2019). They thus pick the species to bring in the forefront. In case of two datatypes with different units, such as catch data and estimated abundance indices, we recommend calculating vulnerability indices separately. Fisheries' stakeholders can then compare those maps for decision-making: on one side juveniles of tuna and on the other vulnerable associated species they seek to protect.

TYPICAL WORKFLOW FOR A 5-SPECIES SYSTEM DURING ONE YEAR

TUNA CATCH DATA

Month	xLon	yLat	BET	YFT	SKJ
.	.	.	.	.	.

ACCESSORY SPECIES

Month	xLon	yLat	FAL	RAY
.	.	.	.	.

(a)

JOINT DATASET

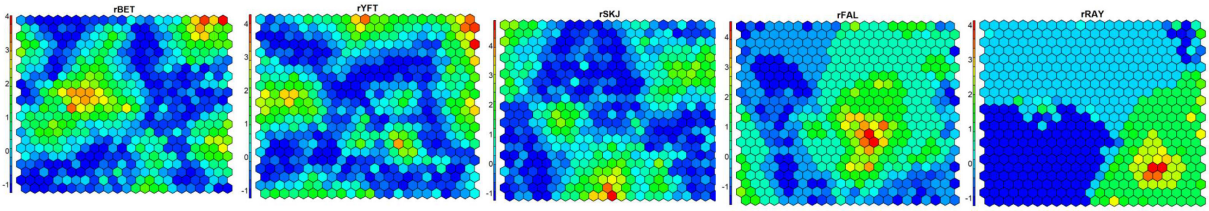
Year	Month	xLon	yLat	BET	YFT	SKJ	FAL	RAY	obsID
.	.	.	.	.	.	.	.	.	.

(b)

SOM INPUT MATRIX

BET	YFT	SKJ	FAL	RAY
.	.	.	.	.

(c)

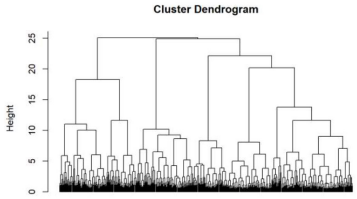


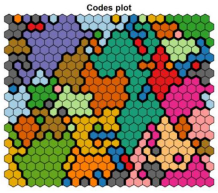
(d)

CODEBOOK VECTORS

BET	YFT	SKJ	FAL	RAY	Node
.	.	.	.	.	.

(e)





CODEBOOK VECTORS

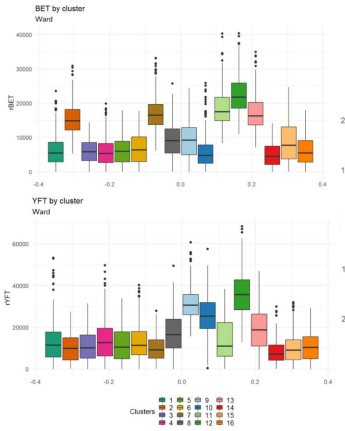
BET	YFT	SKJ	FAL	RAY	Node	Cluster
.	.	.	.	.	.	.

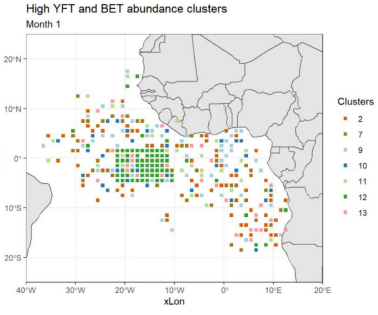
(f)

FINAL DATASET

Month	xLon	yLat	BET	YFT	SKJ	FAL	RAY	obsID	Node	Cluster
.	.	.	.	.	.	.		.	.	.

(g)





(h)

VULNERABILITY

Cluster	Weighted mean	Vulnerability
.	.	.

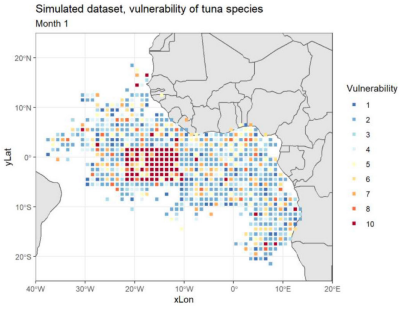


FIGURE 3 Typical workflow for a 5-species system during 1 year. (a) Joining of fisheries' catch data and bycatch data, in this case in the form of abundance indices between 0 and 5. (b) Conversion of species' columns to matrix format, then centring and scaling of the matrix. (c) SOM training and visualization. (d) Extraction of the codebook vectors. (e) Clustering of the nodes using the 'hclust' and 'cuTreeHybrid' functions and projection of the clusters on the SOM. (f) Assignment of the cluster to the original observations that belong to a node. (g) Extraction of BET and YFT high abundance clusters using the clusters' medians. (h) Creation of a vulnerability index by cluster for the tuna species and for the associated species independently of each other (the latter is not illustrated, see Figures 5 and 6).

TABLE 1 Main SOM parameters of the 'som' function from the 'kohonen' package. 'Data' refers to the SOM's data input matrix, centered and scaled.

Parameter	Set to
Grid size	Conventional rule: $(ceiling(\sqrt{nrow(data)}) \times 5$
Sides ratio (x/y)	Conventional rule: svd.data = svd(data) x/y = svd.data$d[1]/svd.data$d[2]
Topography (*topo*)	'hexagonal' (vs. 'rectangular')
Nb. of iteration (*rlen*)	In a first round, plot 'changes' and check when convergence is reached. Choose value optimizing computation time and convergence. In our case, it appeared than 15,000 iterations was an efficient choice
Learning rate (*alpha*)	Default: decline linearly from 0.05 to 0.01 over rlen updates
Neighbourhood radius (*radius*)	Default: start value covers 2/3 of all unit-to-unit distances

TABLE 2 Mono- and multi-species hotspot detection results of SOM's analyses on simulated data (one temporal unit) under different data pre-processing and clustering approaches. 'X' marks the correct identification of the hotspot. See script 'DataSimulation_Preprocessing_Clustering' on the stephanpau/SOM_TunaFisheries GitHub repository for more details.

Simulated data		hclust + cuTreeHybrid		nbclust	
Hotspots	Preprocessing	Ward	UPGMA	Ward	UPGMA
BET	None	—	—	—	—
	Center-scale	X	—	X	—
	log1p center-scale	—	—	—	—
	log1p	—	—	—	—
BET/YFT	None	—	—	—	—
	Center-scale	X	X	X	—
	log1p center-scale	X	—	—	—
	log1p	X	—	—	—

3 | RESULTS

3.1 | Simulation results

Analyses on simulated data to identify detection thresholds of juveniles' catch hotspots revealed that our method allowed to easily spot them on the maps every time the local catch within the hotspot is equal to or larger than 1.5 time the average catch of the tested species in the same month (see Table S1).

Furthermore, the method correctly identified most of the hotspots when increasing complexity in a 12-month simulation. Table 3 displays the performance of the method starting from a baseline case with no added hotspot. In case n°1, we added a bispecific (bigeye and yellowfin) hotspot that lasted 3 months, from January to March. The last column indicates that three hotspots out of three were detected, in other words that the bispecific hotspot was detected each month it was present. In case n°2, we kept the previous hotspot and added a single species' hotspot in November, which was also perfectly detected. From case n°3 onwards, we introduced associated species. In contrast to the previous scenarios, in case n°4, a bigeye hotspot over 1 month was not detected. As a sensitivity analysis, we then varied its added catch and its presence month. Increasing the catches within that non-detected hotspot (case n°5) resulted in detection, while conversely moving it only to another month to avoid temporal overlapping with the ray hotspot (case n°6) did not. Replacing the ray hotspot by a bigeye hotspot, that is, substituting by bigeye catches for the same zone and same months (case n°7), led again to detection. These results suggest that a 1-month and mono-specific hotspot does not weigh enough in a five species system.

TABLE 3 Results of 12 months, that is, 12 temporal units, simulations with increasing hotspots complexity. Hotspots belonging to the same species' column and case's row are in one geographical location over the same months. The names of the associated species correspond to usually bycaught elasmobranchs in dFAD fishing (FAL for silky sharks and RAY for any ray species). The number of detected spatiotemporal hotspots tells of the performance of the method. See script 'DataSimulation_Over12months' on the stephanpau/SOM_TunaFisheries GitHub repository for more details.

	Simulated hotspots											
	Juveniles of tuna species						Vulnerable associated species					
	BET			BET and YFT			FAL			RAY		
Case	Months	Added catch	Detected	Months	Added catch	Detected	Months	Added catch	Detected	Months	Added catch	Detected	No. of detected spatiotemp. hotspots
Baseline	–	–	–	–	–	–	–	–	–	–	–	–	0/0
n°1	–	–	–	1–3	Mean×2	Yes	–	–	–	–	–	–	3/3
n°2	11	Mean×2	Yes	1–3	Mean×2	Yes	–	–	–	–	–	–	4/4
n°3	11	Mean×2	Yes	1–3	Mean×2	Yes	9–12	+3	Yes	–	–	–	8/8
n°4	11	Mean×2	Partially	1–3	Mean×2	Yes	9–12	+3	Yes	3–6	+3	Yes	11/12
n°5	11	Mean×3	Yes	1–3	Mean×2	Yes	9–12	+3	Yes	3–6	+3	Yes	12/12
n°6	7	Mean×2	Partially	1–3	Mean×2	Yes	9–12	+3	Yes	3–6	+3	Yes	11/12
n°7	3–6	Mean×2	Yes	1–3	Mean×2	Yes	9–12	+3	Yes	Random only			11/11

3.2 | Case study: Tropical tuna fisheries of the Atlantic Ocean

3.2.1 | General trends

The SOMs were trained by individual year and by month over the 5-years periods (1994–1998 and 2012–2016), and for the sum of catch over all years (i.e. by summing up purse seiners tuna catch on dFADs by spatiotemporal cell and associating those with the juvenile silky shark abundance index). From the analysis of the outcomes, it can be observed that no information was lost when summarizing the dataset as a typical year. As this is the most efficient analysis for later interpretation and management decision-making, we present the results for the summed catch of the two time periods selected, before and after the implementation of the moratoria on dFADs.

The same and already known trends in tuna species' associations emerged from both periods of years and are reflected in the clusters, defined by the black borders and numbered (Figure 4). Clusters rallied the same type of observations in both periods, for example, clusters n°3 and n°4, for 1994–1998 and 2012–2016 periods respectively. Skipjack (SKJ) and bigeye (BET) tunas were more closely associated, as can be seen with their overlapping high-abundance zones: the upper left corner for the 1994–1998 period and in the bottom left corner for the 2012–2016 period. In both time periods as well, the silky shark juveniles appeared to reside in similar zones as yellowfin tuna, and less so with the two other species (Figure 4a upper part for 1994–1998 and Figure 4b cluster 5 for 2012–2016).

3.2.2 | Tuna juveniles' and vulnerable associated species' hotspots

During the first period (1994–1998) one main zone between November and March was isolated: a 10° wide horizontal hotspot along the Equator, approximately between 25°W and 5°W (Figure 5). This area corresponds to the core zone of the different dFAD seasonal closures implemented between 1999 and 2019 and defined following expert knowledge (dark red zone in Figure 1). It can be seen however, that juvenile silky sharks are not particularly associated with these hotspots (panels c and d of Figure 5).

For the 2012–2016 period, a period subject to an ICCAT dFAD time-area closure in January and February, several tuna hotspots are detected. In their turn the juvenile silky sharks' hotspots extracted from Lopez et al.'s (2020) distribution maps are correctly identified. These results give hints towards better fitted, smaller and shorter time-area closures. For instance, (i) between July and September, two small hotspots for both tunas and sharks emerged along the Mauritanian and the Gabonese coasts, (ii) in November and December, the same 10° wide horizontal hotspot appeared along the Equator, approximately between 25°W and 5°W and (iii) in February and March, during and right after the closure, a very small zone in the South of the Angolan coast emerges (Figure 6).

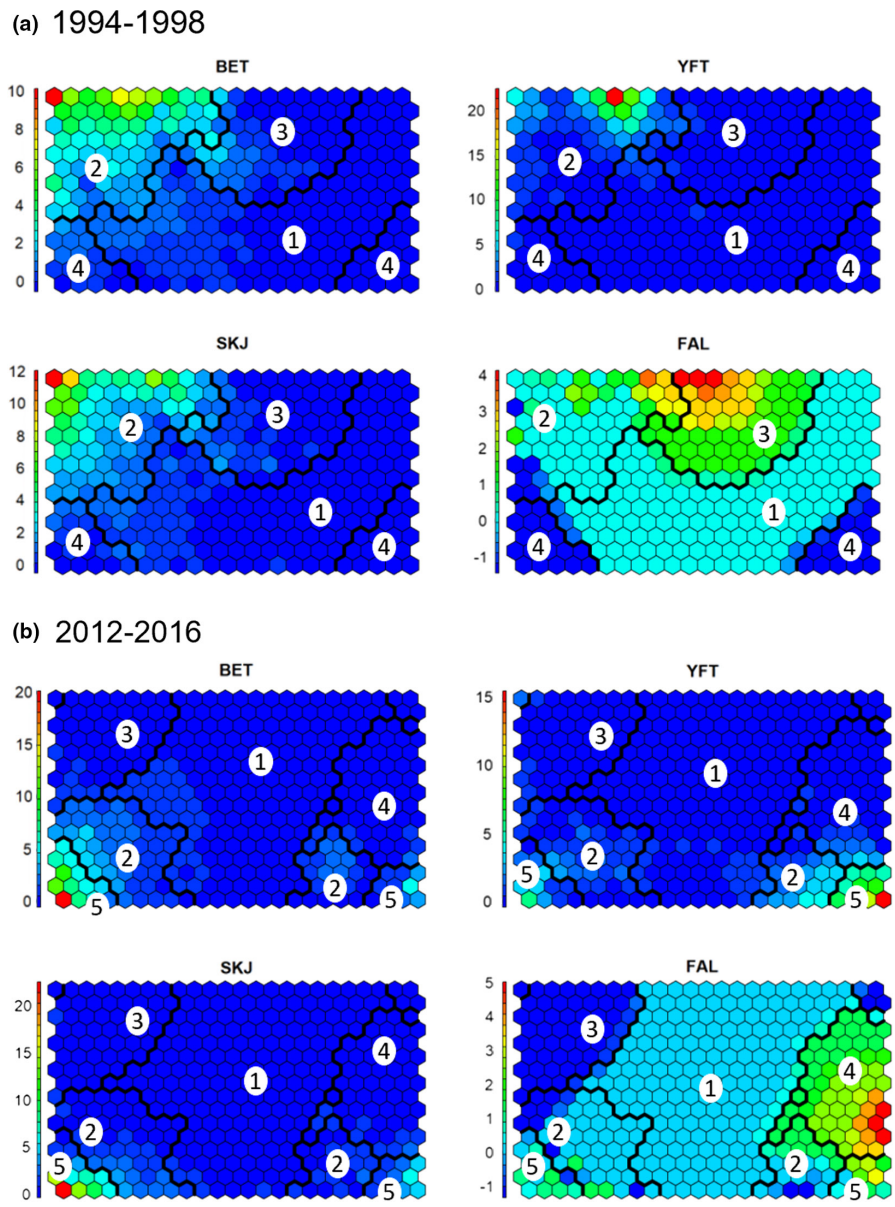


FIGURE 4 Heat maps for the four species by period, that is, 1994–1998 (a) and 2012–2016 (b). The numbers refer to the clusters delimited by the black borders. Warm colours correspond to high catch values or abundance index, while cold colours indicate low catch. See scripts 'FinalSOM_94_98' and 'FinalSOM_12_16' on the stephanpau/SOM_TunaFisheries GitHub repository for more details.

The discrepancy between the two periods of years in the duration of the central hotspot (November–December versus November–March) suggested that the decrease in dFADs catch due to the moratorium masked its location between 2012 and 2016. Besides, as can be seen in the difference in the geographical extent of the catch data, the fishery has greatly expanded and spread towards the south. This explain why the small hotspot in front of Angola could not have been detected in the 1990s, as well as the silky shark hotspots below the 10°S latitude.

4 | DISCUSSION

4.1 | Potential for fisheries management

The method proposed in this study was developed from a practice-oriented perspective. We tested it on randomly generated datasets whose characteristics (location, overall mean catch values and standard deviations) were based on the ICCAT's tropical purse seiner dFADs catch by 1°× month cells statistics. Those simulations and the following case study on tropical tuna fisheries of the Atlantic Ocean demonstrated the approach's effectiveness in identifying candidate zones for dFAD time-area closures. These zones, illustrated in patent geographic maps, have the potential to underpin management at the fishery scale, that is, to be brought around the table and to serve as a discussion basis. It should also be noted that the method would detect classical 'fishing the line' and effort reallocation effects (Torres-Irineo et al., 2011), which should both be considered before taking management decisions.

Besides, the method's R scripts can be run with minimal skills and require only minor changes from the users, who would adjust the factors to compute the vulnerability indices so as to better reflect the changes over time of the exploitation status of tuna species and the apparent abundance of associated species. Moreover, transferring the analysis workflow to a different dataset or ecosystem necessitates few simple adjustments, which in our opinion constitutes the main appeal of this approach in comparison to the ones mentioned

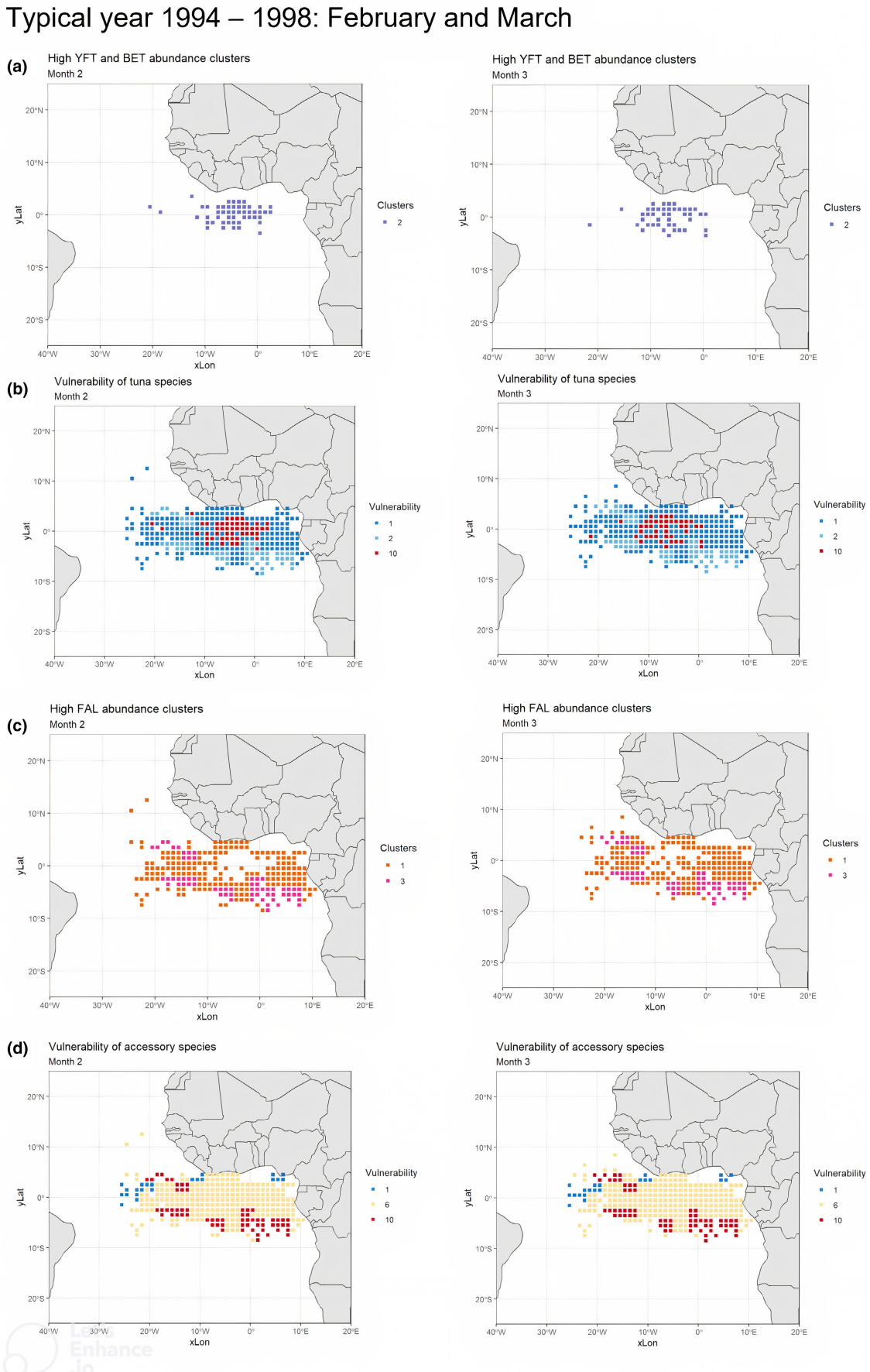


FIGURE 5 Illustration of parts of the results from the period 1994–1998, months 2 and 3. The central tuna hotspots that appears is present from November to March. This zone was empirically selected for a large time-area closure during three periods between 1998 and 2019 (see Figure 1).

in the introduction section (Grüss et al., 2019; Guan et al., 2020; Kai et al., 2017; Paradinas et al., 2015; Quiroz & Prates, 2018; Sibert et al., 2012). A further strength is undoubtedly the flexibility in data input, even if it implies cautious interpretation from the users.

4.2 | Limitations

The main limitation faced in the case study and most likely to be faced by anyone working on fisheries management was data quality and availability. We performed the analysis on an open-access database as it was, bearing in mind that we wanted to reflect a real-life situation with limited data access. We, thus, had to take account of the following points:

1. The ICCAT task 2 catch/effort data per fishing mode consists in logbook data corrected in terms of tuna species and size compositions based on samplings at landing. However, until now the correction process is done at large grid unit for the European purse seiners and with inequal quality with regards to data logbook collection and sampling effort for several other purse seiner fleets.
2. Reliable juvenile proportions of bigeye and yellowfin tunas estimated by sampling for every 1°cell with dFAD catch would have been needed to identify accurately juvenile catch hotspots. Still, as bigeye and yellowfin catch on dFADs are typically made up of about 80% of juveniles, we can reasonably assume that hotspots identified on this dataset are predominantly constituted of juveniles.
3. Bycatch data are collected by national or European observer programs (Escalle et al., 2016; Torres-Irineo, Amandè, et al., 2014b) and not freely accessible. Besides, to estimate bycatch in tuna purse-seine fisheries, data comes sometimes from low-coverage observer programs and is raised to the whole fisheries resulting in high uncertainty (Amandè et al., 2012). Nevertheless, in this study, we could show that spatially, temporally and quantitatively coarse abundance indices could efficiently be combined with precise catch data. This approach still requires minimal bycatch information, which is not available in the literature or online for most vulnerable species. We thus considered using species' distribution models (Lezama-Ochoa et al., 2020; Pikesley et al., 2013; Santos & Coelho, 2019) to build such an index, or even to directly use the modelled probability of presence (typically between 0 and 1) or the forecasted presence/absence as a variable in the input matrix. The analysis would have then detected probability of presence hotspots for the included species. These models also require input data to be built and are thus subject to similar limitations as we are. Still, the combination of catch data and probability of presence remains to be tested.
4. Finally, we cannot account for non-declared tuna catch, for instance within dFAD time-area closures or by non-contracting parties of the ICCAT (IUU fisheries). This lacking information could impact hotspot detection.

4.3 | Perspectives

To further validate the method and explore its potential in ecosystem description, next steps should include experimenting further types of fisheries' data, comparisons with different spatiotemporal models for external validation as well as inclusion of oceanographic variables. Following these, experimenting with supervised SOMs for near-future predictions of catch hotspots and incorporating them into the shift towards dynamic ocean management in large-scale international fisheries could be pursued.

To go beyond the current results, catch-per-unit-effort (CPUE) data could be an alternative to commercial tuna catch data, a shift that would imply seeking ecological hotspots rather than catch hotspots. Indeed, CPUE data, accounting for fishing effort, is assumed to reflect the abundance and consequently the ecology of each species. Unlike the catch, which is directly linked to the past fishing mortality, the apparent abundance expressed by the CPUE informs us about strata that are non-systematically explored but where regulating dFAD-fishing operations would benefit juveniles of tuna species. The calculation of CPUE needs however the definition of a common unit of effort, which remains problematic for dFADs sets as part of them are detected randomly (non-owned dFADs) and part of them are continuously tracked at-sea by the vessel owner.

Along with that, adding environmental, for example, sea surface temperature, chlorophyll-a concentration, etc., and fishing-related variables like dFAD density, to the input matrix could give valuable insights into recurring conditions associated with catch hotspots. This objective has already been pursued by Lopez et al. (2017) who explored environmental preferences of dFAD-associated species, yet using general additive mixed models and data from echosounder buoys. The authors found several significant relationships, amongst others between tuna hotspots and phytoplankton bloom. Similarly, Wang et al. (2018) sought to detect skipjack tuna environmental habitat preferences in the west-central Pacific Ocean with the help of classical feedforward neural networks. A novel SOM analysis would provide clues for environmental conditions of candidate conservation areas and could draw upon previous findings for validation (see Table 1 in Lopez et al., 2017). More importantly it would constitute a preliminary analysis before making the step to supervised models for near anticipation of effective conservation measures.

Forecasting near-future hotspots with oceanographic models, possible with SOMs using functions for supervised networks (Melssen et al., 2006), would allow the short-term planning of restrictive measures. With climate change, tuna habitat distribution limits are already spreading poleward at an overall estimated rate of about 6 kilometres per decade, varying by hemisphere and species (Erauskin-Extramiana et al., 2019). Amongst tropical tunas, bigeye tunas will decline in tropical zones and follow the general trend, while yellowfin and skipjack tunas are expected to thrive along tropical coastlines and in tropical areas. Along with that, technological improvements and changes in fishing strategies need to be accounted

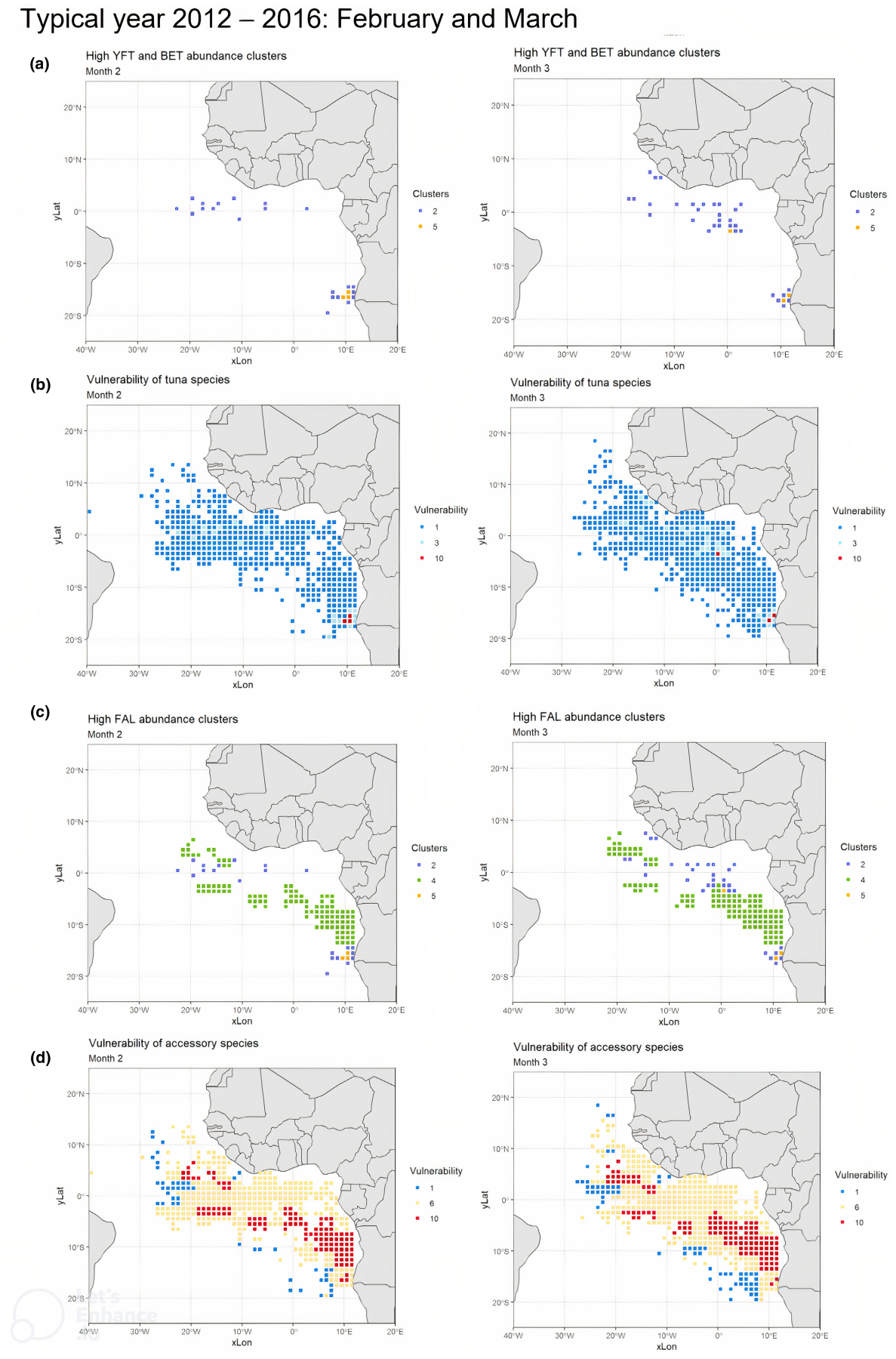


FIGURE 6 Illustration of parts of the results from the period 2012–2016, months 2 and 3. February is subject to an fish aggregating device closure (African coast, 10°S, 5°E and 5°W), which does not appear in panels (b and d). This is due to catches within the zone in 2012 and 2016 and to the construction of the typical year, summing up catches by cell. However, the effect of the moratorium appears in panel (a), as the high-abundance cluster usual for this period of time does not appear.

for (Eide, 2017; Torres-Irineo, Gaertner, et al., 2014), which is implicitly done when training SOMs on recent catch or CPUE data.

Apart from environmental variables, spatial coordinates can be introduced to the analysis using the 'Geo-SOM' algorithm developed by Bacao et al. (2004), yet not as a common variable but as a constraint. Concretely, the user can define a parameter *k* called 'geographical tolerance', that will define how close geographically a best-matching unit can be. In other words, during the SOM training, the best-matching unit will only be chosen from a subset of units that are located within a geographical radius defined by *k*. In a context of spatial fisheries management, this approach could prove very useful to design smaller time-area closures within very large fishing zones.

It follows from both all these considerations that spatiotemporal closures' management require dynamic approaches to best serve both ecological and economic targets. (Maxwell et al., 2015) advocate for the generalized used of dynamic ocean management, whose core principle is to integrate near real-time data to quickly adapt to new situations. The authors illustrate its effectiveness with two coastal fisheries, the American sea scallop fishery (O'Keefe & DeCelles, 2013) and the Australian southern bluefin tuna fishery (Hobday et al., 2010), where it led to a decrease in closure periods and areas while still resulting in encouraging improvements in stock conservation and bycatch issues. In the case of a large-scale and international fishery such as the purse-seiners' of the tropical Atlantic Ocean, measures' implementation needs more time and consensus than in the two examples mentioned above (Kaplan et al., 2014). Still, adapting measures on the basis of yearly analyses offers great potential for all stakeholders and for sustainability in international fisheries.

AUTHOR CONTRIBUTIONS

Pauline Stephan: Formal analysis, Methodology, Writing – Original Draft, Writing – Review & Editing. Daniel Gaertner: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition. Ilan Perez: Conceptualization. Loreleï Guéry: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition. All authors contributed and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank Pr. Jake Alexander from the ETH Zurich for formal supervision and useful comments on the manuscript. This work was funded by the Fondation pour la recherche sur la biodiversité (https://www.fondationbiodiversite.fr/la-frb-en-action/soutien-aux-jeunes-chercheurs/financement-stage-master-ii-recherche/).

CONFLICT OF INTEREST

The authors have no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.14008.

DATA AVAILABILITY STATEMENT

All R scripts and datasets are available on the GitHub repository 'stephanpau/SOM_TunaFisheries' (https://github.com/stephanpau/SOM_TunaFisheries) and on Zenodo (Stephan et al., 2022).

ORCID

Pauline Stephan https://orcid.org/0000-0002-5218-3852
Daniel Gaertner https://orcid.org/0000-0002-1194-4432
Loreleï Guéry https://orcid.org/0000-0003-3004-8462

REFERENCES

Amandè, M. J., Chassot, E., Chavance, P., Murua, H., Molina, A. D. D., & Bez, N. (2012). Precision in bycatch estimates: The case of tuna purse-seine fisheries in the Indian Ocean. *ICES Journal of Marine Science*, *69*(2012), 1501–1510. https://doi.org/10.1093/icesjms/fss106

Andrienko, G., Andrienko, N., Bremm, S., Schreck, T., Von Landesberger, T., Bak, P., & Keim, D. (2010). Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. *Computer Graphics Forum*, *29*(3), 913–922. https://doi.org/10.1111/j.1467-8659.2009.01664.x

Arriz, J., Delgado, A., Fonteneau, A., Gonzales Costa, F., & Pallares, P. (1999). Logs and tunas in the eastern tropical Atlantic. A review of present knowledge and uncertainties. Proceedings of the international workshop on fishing for tunas associated with floating objects. *ICCAT Special Report*, *11*, 21–65.

Bacao, F., Lobo, V., & Painho, M. (2004). Geo-self-organizing map (geo-SOM) for building and exploring homogeneous regions. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, *3234*, 22–37. https://doi.org/10.1007/978-3-540-30231-5_2

Baidai, Y., Dagorn, L., Amande, M. J., Gaertner, D., & Capello, M. (2020). Machine learning for characterizing tropical tuna aggregations under drifting fish aggregating devices (DFADs) from commercial echosounder buoys data. *Fisheries Research*, *229*(July 2019), 105613. https://doi.org/10.1016/j.fishres.2020.105613

Bivand, R., & Rundel, C. (2013). Package 'rgeos'. The Comprehensive R Archive Network, 61.

Carlucci, R., Bandelj, V., Ricci, P., Capezzuto, F., Sion, L., Maiorano, P., Tursi, A., Solidoro, C., & Libralato, S. (2018). Exploring spatio-temporal changes in the demersal and benthopelagic assemblages of the northwestern Ionian Sea (Central Mediterranean Sea). *Marine Ecology Progress Series*, *598*, 1–19. https://doi.org/10.3354/meps12613

Céréghino, R., & Park, Y. S. (2009). Review of the self-organizing map (SOM) approach in water resources: Commentary. *Environmental Modelling and Software*, *24*(8), 945–947. https://doi.org/10.1016/j.envsoft.2009.01.008

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). Nbclust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*(6), 1–36. https://doi.org/10.18637/jss.v061.i06

Chon, T. S. (2011). Self-organizing maps applied to ecological sciences. *Ecological Informatics*, *6*(1), 50–61. https://doi.org/10.1016/j.ecoinf.2010.11.002

Conti, L., Grenouillet, G., Lek, S., & Scardi, M. (2012). Long-term changes and recurrent patterns in fisheries landings from large marine ecosystems (1950-2004). *Fisheries Research*, *119–120*, 1–12. https://doi.org/10.1016/j.fishres.2011.12.002

de Bodt, E., Cottrell, M., & Verleysen, M. (2002). Statistical tools to assess the reliability of self-organizing maps. *Neural Networks*, *15*(8–9), 967–978. https://doi.org/10.1016/S0893-6080(02)00071-0

Eide, A. (2017). Climate change, fisheries management and fishing aptitude affecting spatial and temporal distributions of the Barents Sea cod fishery. *Ambio*, *46*(s3), 387–399. https://doi.org/10.1007/s13280-017-0955-1

Erauskin-Extramiana, M., Arrizabalaga, H., Hobday, A. J., Cabré, A., Ibaibarriaga, L., Arregui, I., Murua, H., & Chust, G. (2019). Large-scale distribution of tuna species in a warming ocean. *Global Change Biology*, *25*(6), 2043–2060. https://doi.org/10.1111/gcb.14630

Escalle, L., Gaertner, D., Chavance, P., Delgado De Molina, A., Ariz, J., & Mrigot, B. (2017). Forecasted consequences of simulated FAD moratoria in the Atlantic and Indian oceans on catches and bycatches. *ICES Journal of Marine Science*, *74*(3), 780–792. https://doi.org/10.1093/icesjms/fsw187

Escalle, L., Pennino, M. G., Gaertner, D., Chavance, P., Delgado de Molina, A., Demarcq, H., Romanov, E., & Merigot, B. (2016). Environmental factors and megafauna spatio-temporal co-occurrence with purse-seine fisheries. *Fisheries Oceanography*, *25*(4), 433–447. https://doi.org/10.1111/fog.12163

Fonteneau, A., Chassot, E., & Bodin, N. (2013). Global spatio-temporal patterns in tropical tuna purse seine fisheries on drifting fish aggregating devices (DFADs): Taking a historical perspective to inform current challenges. *Aquatic Living Resources*, *26*(1), 37–48. https://doi.org/10.1051/alr/2013046

Fonteneau, A., Gaertner, D., Maufroy, A., & Amandè, M. J. (2016). Effects of the ICCAT FAD moratorium on the tuna fisheries and tuna stocks. *Collected Volume of Scientific Papers ICCAT*, *72*(2), 520–533.

Fonteneau, A., Pallarés, P., & Pianet, R. (2000). A worldwide review of purse seine fisheries on FADs. *Pêche Thonière et Dispositifs de Concentration de Poissons*, *Caribbean-Martinique*, *1*, 15–35.

Giraudel, J. L., & Lek, S. (2001). A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling*, *146*(1–3), 329–339. https://doi.org/10.1016/S0304-3800(01)00324-6

Grüss, A., Thorson, J. T., & Jardim, E. (2019). Developing spatio-temporal models using multiple data types for evaluating population trends and habitat usage. *ICES Journal of Marine Science*, *76*(6), 1748–1761. https://doi.org/10.1093/icesjms/fsz075

Guan, L., Shan, X., Jin, X., Gorfine, H., Yang, T., & Li, Z. (2020). Evaluating spatio-temporal dynamics of multiple fisheries-targeted populations simultaneously: A case study of the Bohai Sea ecosystem in China. *Ecological Modelling*, *422*(November 2019), 1–10. https://doi.org/10.1016/j.ecolmodel.2020.108987

Hilborn, R., Akselrud, C. A., Peterson, H., & Whitehouse, G. A. (2021). The trade-off between biodiversity and sustainable fish harvest with area-based management. *ICES Journal of Marine Science*, *78*(6), 2271–2279. https://doi.org/10.1093/icesjms/fsaa139

Hobday, A. J., Hartog, J. R., Timmiss, T., & Fielding, J. (2010). Dynamic spatial zoning to manage southern bluefin tuna (Thunnus maccoyii) capture in a multi-species longline fishery. *Fisheries Oceanography*, *19*(3), 243–253. https://doi.org/10.1111/j.1365-2419.2010.00540.x

Hyun, K., Song, M. Y., Kim, S., & Chon, T. S. (2005). Using an artificial neural network to patternize long-term fisheries data from South Korea. *Aquatic Sciences*, *67*(3), 382–389. https://doi.org/10.1007/s00027-005-0771-8

ICCAT. (2015). *Resolution (15–11) by ICCAT Concerning the application of an ecosystem approach to fisheries management*. https://www.iccat.int/Documents/Recs/compendiopdf-e/2015-11-e.pdf

ICCAT. (2019). *ICCAT basic texts* (7th Rev. ed., pp. 1–120). https://www.iccat.int/documents/commission/basictexts.pdf

ICCAT. (2020). *Compendium management recommendations and resolutions adopted by ICCAT for the conservation of Atlantic tunas and tuna-like species* (pp. 1–447). https://www.iccat.int/com2020/ENG/PLE_102_ENG.pdf

ICCAT. (2021). *Statistical databases*. https://iccat.int/en/accesingdb.html

Kai, M., Thorson, J. T., Piner, K. R., & Maunder, M. N. (2017). Predicting the spatio-temporal distributions of pelagic sharks in the western and central North Pacific. *Fisheries Oceanography*, *26*(5), 569–582. https://doi.org/10.1111/fog.12217

Kalteh, A. M., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling and Software*, *23*(7), 835–845. https://doi.org/10.1016/j.envsoft.2007.10.001

Kangur, K., Park, Y. S., Kangur, A., Kangur, P., & Lek, S. (2007). Patterning long-term changes of fish community in large shallow Lake Peipsi. *Ecological Modelling*, *203*(1–2), 34–44. https://doi.org/10.1016/j.ecolmodel.2006.03.039

Kaplan, D. M., Chassot, E., Amande, J. M., Dueri, S., Demarcq, H., Dagorn, L., & Fonteneau, A. (2014). Spatial management of Indian Ocean tropical tuna fisheries: Potential and perspectives. *ICES Journal of Marine Science*, *71*(November), 1728–1749.

Kell, L. T. (2011). A standardised way of presenting species groups executive sumarises. *Collective Volume of Scientific Papers ICCAT*, *66*(5), 2213–2228.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69. https://doi.org/10.1007/BF00337288

Langfelder, P., Zhang, B., & Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: The dynamic tree cut package for R. *Bioinformatics*, *24*(5), 719–720. https://doi.org/10.1093/bioinformatics/btm563

Lezama-Ochoa, N., Lopez, J., Hall, M., Bach, P., Abascal, F., & Murua, H. (2020). Spatio-temporal distribution of the Spinetail devil ray Mobula Mobular in the eastern tropical Atlantic Ocean. *Endangered Species Research*, *43*, 447–460. https://doi.org/10.3354/ESR01082

Lopez, J., Alvarez-Berastegui, D., Soto, M., & Murua, H. (2020). Using fisheries data to model the oceanic habitats of juvenile silky shark (*Carcharhinus falciformis*) in the tropical eastern Atlantic Ocean. *Biodiversity and Conservation*, *29*(7), 2377–2397. https://doi.org/10.1007/s10531-020-01979-7

Lopez, J., Moreno, G., Lennert-Cody, C., Maunder, M., Sancristobal, I., Caballero, A., & Dagorn, L. (2017). Environmental preferences of tuna and non-tuna species associated with drifting fish aggregating devices (DFADs) in the Atlantic Ocean, ascertained through fishers' echo-sounder buoys. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *140*(March), 127–138. https://doi.org/10.1016/j.dsr2.2017.02.007

Mailund, T. (2019). Manipulating data frames: dplyr. In *R data science quick reference* (pp. 109–160). Apress. https://doi.org/10.1007/978-1-4842-4894-2_7

Maxwell, S. M., Hazen, E. L., Lewison, R. L., Dunn, D. C., Bailey, H., Bograd, S. J., Briscoe, D. K., Fossette, S., Hobday, A. J., Bennett, M., Benson, S., Caldwell, M. R., Costa, D. P., Dewar, H., Eguchi, T., Hazen, L., Kohin, S., Sippel, T., & Crowder, L. B. (2015). Dynamic Ocean management: Defining and conceptualizing real-time management of the ocean. *Marine Policy*, *58*, 42–50. https://doi.org/10.1016/j.marpol.2015.03.014

Melssen, W., Wehrens, R., & Buydens, L. (2006). Supervised Kohonen networks for classification problems. *Chemometrics and Intelligent Laboratory Systems*, *83*(2), 99–113. https://doi.org/10.1016/j.chemolab.2006.02.003

Mendoza-Carranza, M., Ejarque, E., & Nagelkerke, L. A. J. (2018). Disentangling the complexity of tropical small-scale fisheries dynamics using supervised self-organizing maps. *PLoS ONE*, *13*(5), 1–28. https://doi.org/10.1371/journal.pone.0196991

Mérigot, B., Urbec, J.-P., & Gaertner, J.-C. (2010). On goodness-of-fit measure for dendrogram-based analyses. *Ecology*, *91*(6), 1850–1859.

Mersmann, O., Trautmann, H., Steuer, D., & Bornkamp, B. (2018). *Package 'truncnorm'* (pp. 21–24). The Comprehensive R Archive Network.

O'Keefe, C. E., & DeCelles, G. R. (2013). Forming a partnership to avoid bycatch. *Fisheries*, *38*(10), 434–444. https://doi.org/10.1080/03632415.2013.838122

Paradinas, I., Conesa, D., Pennino, M. G., Muñoz, F., Fernández, A. M., López-Quílez, A., & Bellido, J. M. (2015). Bayesian spatio-temporal approach to identifying fish nurseries by validating persistence areas. *Marine Ecology Progress Series*, *528*, 245–255. https://doi.org/10.3354/meps11281

Park, Y. S., Tison, J., Lek, S., Giraudel, J. L., Coste, M., & Delmas, F. (2006). Application of a self-organizing map to select representative species in multivariate analysis: A case study determining diatom distribution patterns across France. *Ecological Informatics*, *1*(3 Spec. Iss), 247–257. https://doi.org/10.1016/j.ecoinf.2006.03.005

Pebesma, E., & Bivand, R. (2012). *Package 'sp'*. The Comprehensive R Archive Network. https://r-forge.r-project.org/projects/rspatial/

Peña, M., Carbonell, A., Tor, A., Alvarez-Berastegui, D., Balbín, R., dos Santos, A., & Alemany, F. (2015). Nonlinear ecological processes driving the distribution of marine decapod larvae. *Deep-Sea Research Part I: Oceanographic Research Papers*, *97*, 92–106. https://doi.org/10.1016/j.dsr.2014.11.017

Perez, I., Guéry, L., Authier, M., & Gaertner, D. (2022). Assessing the effectiveness of dFADs fishing moratorium in the eastern Atlantic Ocean for conservation of juvenile tunas from AOTTP data. *Fisheries Research*, *253*, 106360. https://doi.org/10.1016/j.fishres.2022.106360

Pikesley, S. K., Maxwell, S. M., Pendoley, K., Costa, D. P., Coyne, M. S., Formia, A., Godley, B. J., Klein, W., Makanga-Bahouna, J., Maruca, S., Ngouessono, S., Parnell, R. J., Pemo-Makaya, E., & Witt, M. J. (2013). On the front line: Integrated habitat mapping for olive ridley sea turtles in the Southeast Atlantic. *Diversity and Distributions*, *19*(12), 1518–1530. https://doi.org/10.1111/ddi.12118

Quiroz, Z. C., & Prates, M. O. (2018). Bayesian spatio-temporal modelling of anchovy abundance through the SPDE approach. *Spatial Statistics*, *28*, 236–256. https://doi.org/10.1016/j.spasta.2018.08.005

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

R Core Team. (2016). *Package 'stats'*. The Comprehensive R Archive Network.

Roigé, M., McGeoch, M. A., Hui, C., & Worner, S. P. (2017). Cluster validity and uncertainty assessment for self-organizing map pest profile analysis. *Methods in Ecology and Evolution*, *8*(3), 349–357. https://doi.org/10.1111/2041-210X.12669

Russo, T., Carpentieri, P., Fiorentino, F., Arneri, E., Scardi, M., Cioffi, A., & Cataudella, S. (2016). Modeling landings profiles of fishing vessels: An application of self-organizing maps to VMS and logbook data. *Fisheries Research*, *181*, 34–47. https://doi.org/10.1016/j.fishres.2016.04.005

Santos, C. C., & Coelho, R. (2019). Distribution patterns and indicators of the smooth hammerhead shark (*Sphyrna zygaena*) in the Atlantic Ocean. *Fisheries Research*, *212*(December 2018), 107–113. https://doi.org/10.1016/j.fishres.2018.12.015

Sibert, J., Senina, I., Lehodey, P., & Hampton, J. (2012). Shifting from marine reserves to maritime zoning for conservation of Pacific bigeye tuna (*Thunnus obesus*). *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), 18221–18225. https://doi.org/10.1073/pnas.1209468109

Simić, V. M., Simić, S. B., Stojković Piperac, M., Petrović, A., & Milošević, D. (2014). Commercial fish species of inland waters: A model for sustainability assessment and management. *Science of the Total Environment*, *497–498*, 642–650. https://doi.org/10.1016/j.scitotenv.2014.07.092

Stephan, P., Gaertner, D., Perez, I., & Guéry, L. (2022). Multi-species hotspots detection using self-organizing maps: Simulation and application to purse seine tuna fisheries management. *Zenodo* https://doi.org/10.5281/ZENODO.7128648

Swimmer, Y., Zollett, E. A., & Gutierrez, A. (2020). Bycatch mitigation of protected and threatened species in tuna purse seine and longline fisheries. *Endangered Species Research*, *43*, 517–542. https://doi.org/10.3354/ESR01069

Torres-Irineo, E., Amandè, M. J., Gaertner, D., de Molina, A. D., Murua, H., Chavance, P., Ariz, J., Ruiz, J., & Lezama-Ochoa, N. (2014a). Bycatch species composition over time by tuna purse-seine fishery in the eastern tropical Atlantic Ocean. *Biodiversity and Conservation*, *23*(5), 1157–1173. https://doi.org/10.1007/s10531-014-0655-0

Torres-Irineo, E., Gaertner, D., Chassot, E., & Dreyfus-León, M. (2014). Changes in fishing power and fishing strategies driven by new technologies: The case of tropical tuna purse seiners in the eastern Atlantic Ocean. *Fisheries Research*, *155*, 10–19. https://doi.org/10.1016/j.fishres.2014.02.017

Torres-Irineo, E., Gaertner, D., De Molina, A. D., & Ariz, J. (2011). Effects of time-area closure on tropical tuna purse-seine fleet dynamics through some fishery indicators. *Aquatic Living Resources*, *24*(4), 337–350. https://doi.org/10.1051/alr/2011143

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, *11*(3), 586–600. https://doi.org/10.1109/72.846731

Walter, J., Hiroki, Y., Satoh, K., Matsumoto, T., Winker, H., Ijurco, A. U., & Schirripa, M. (2019). Atlantic bigeye tuna stock synthesis projections and Kobe 2 matrices. *Collective Volumes of Scientific Papers*, *75*(7), 2283–2300.

Wang, J., Chen, X., Staples, K. W., & Chen, Y. (2018). The skipjack tuna fishery in the west-Central Pacific Ocean: Applying neural networks to detect habitat preferences. *Fisheries Science*, *84*(2), 309–321. https://doi.org/10.1007/s12562-017-1161-6

Watson, J. T., Essington, T. E., Lennert-Cody, C. E., & Hall, M. A. (2009). Trade-offs in the design of fishery closures: Management of silky shark bycatch in the eastern Pacific Ocean tuna fishery. *Conservation Biology*, *23*(3), 626–635. https://doi.org/10.1111/j.1523-1739.2008.01121.x

Wehrens, R., & Buydens, L. M. C. (2007). Self- and super-organizing maps in R: The kohonen package. *Journal of Statistical Software*, *21*(5), 1–19. https://doi.org/10.18637/jss.v021.i05

Wickham, H. (2016). *Scales: Scale functions for visualization* (pp. 1–50). The Comprehensive R Archive Network.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stephan, P., Gaertner, D., Perez, I., & Guéry, L. (2022). Multi-species hotspots detection using self-organizing maps: Simulation and application to purse seine tuna fisheries management. *Methods in Ecology and Evolution*, *13*, 2850–2864. https://doi.org/10.1111/2041-210X.14008