ICES Journal of Marine Science, 2023, **0**, 1–12
DOI: 10.1093/icesjms/fsad092
Original Article

An approach to map and quantify the fishing effort of polyvalent passive gear fishing fleets using geospatial data

Nuno Sales Henriques [1,2,*], Tommaso Russo [3,4], Luis Bentes [1], Pedro Monteiro [1], Antonio Parisi [3], Ramiro Magno [5], Frederico Oliveira [1], Karim Erzini [1,2], and Jorge M. S. Gonçalves [1]

[1]Centro de Ciencias do Mar, 8005-139, Faro, Portugal
[2]Universidade do Algarve, 8005-139, Faro, Portugal
[3]Università degli Studi di Roma Tor Vergata, 00133, Roma, Lazio, Italy
[4]CoNISMa, Roma, Lazio, Italy
[5]Pattern Institute, 8005-222, Faro, Portugal
∗*Corresponding author: tel: +351 289800051,* e-mail: nsaleshenriques@gmail.com.

The use of tracking devices, such as vessel monitoring systems or automatic identification system, enabled us to expand our knowledge on the distribution and quantification of fishing activities. However, methods and models based on vessel tracking data are mostly devised to be applied to towed gears, whereas applications to multi-gear and passive fisheries have been underrepresented. Here, we propose a methodology to deal with geospatial data to map and quantify the fishing effort, as soak time, of passive fishing gears used by a multi-gear fishing fleet. This approach can be adapted to other passive multi-or single-gear fisheries, since it requires only three variables that can be extracted from a pre-classified dataset, to identify the beginning (gear deployment) and the end (hauling) of passive fishing events. As far as we are aware, this is the first time a methodology that allows quantifying the soak time of static passive fishing events, within a polyvalent fishery context, is presented. We argue that the information that can be extracted from such approaches could contribute to improved management of multi-gear and static-gear fisheries and the ecosystem-based approach.

Keywords: AIS, fisheries mapping, fishing effort, passive fishing gears, polyvalent fishing fleet, soak time, vessel tracking data.

Introduction

To ensure appropriate management and conservation of the marine environment, it is important to understand the distribution and impact of human activities at sea. Fisheries, as one of the most important sources of food for human consumption, are also one of the most impactful extractive activities occurring in this environment (Pauly *et al.*, 1998; Swartz *et al.*, 2010). To ensure the proper management of the marine environment, by establishing comprehensive marine spatial plans, design effective marine protected areas, and protect vulnerable habitats and species it is vital to understand how this activity is performed and where it occurs (Halpern *et al.*, 2008; Campbell *et al.*, 2014; McCauley *et al.*, 2016; Vespe *et al.*, 2016). Despite the efforts to increase knowledge on fisheries, the distribution and quantification of fishing effort are far from being fully understood (Kroodsma *et al.*, 2018; Leblond *et al.*, 2019).

Accurate estimates of fishing effort are critical for ensuring sustainable fisheries, being essential for improving stock assessment, tracking market trends, and studying fishers' profitability (McCluskey and Lewison, 2008; Peterson *et al.*, 2017). Direct control of fishing effort could be a possible alternative to other traditional forms of fisheries management, such as the use of total allowable catches and quotas. But, due to the difficulty of precisely estimating and comparing the fishing effort of different vessels, gears, and locations, this management approach has seldomly been implemented (Shepherd, 2003).

With the introduction of tracking devices in fishing vessels, such as Vessel monitoring system (VMS) and automatic identification system (AIS), the possibility to study the spatial and temporal distribution of fishing activities and the quantification of fishing effort has improved dramatically (Witt and Godley, 2007). A growing number of studies have used data from tracking devices to identify fishing grounds (Gerritsen and Lordan, 2011; Jennings and Lee, 2012; Le Guyader *et al.*, 2017) and to estimate and map fishing effort (Natale *et al.*, 2015; Russo *et al.*, 2019) with an accuracy that was not previously possible. The underlying approaches when dealing with this data are generally to identify and classify different states of fishing trips (e.g. steaming, resting, and fishing) and to determine whether a vessel is fishing or not (Vermard *et al.*, 2010; Poos *et al.*, 2013). These approaches often rely on a simple statistical speed filter to detect fishing activities, under the assumption that fishing occurs at a speed much lower than steaming. In other cases, they use more complex models that rely on trajectory metrics, such as step length, turning angle, and correlation with environmental cues to infer the different phases of fishing trips (Lee *et al.*, 2010; Russo *et al.*, 2014; Leblond *et al.*, 2019; Mendo *et al.*, 2019a). Hidden Markov models (HMM) have been widely used to classify vessel tracking data using these metrics (Vermard *et al.*, 2010;

Received: 13 January 2023; **Revised:** 16 May 2023; **Accepted:** 23 May 2023

de Souza *et al.*, 2016; Whoriskey *et al.*, 2017; Mendo *et al.*, 2019a).

The majority of studies using geospatial data derived from AIS or VMS have focused on active gears such as trawls (Lee *et al.*, 2010; Natale *et al.*, 2015), with few studies on static or passive gears that represent an important proportion of fishing effort and impact (Kelleher *et al.*, 2012). One of the main reasons is that passive gear fishing events are more complex to model than active ones. Fishing with active gears generally involves continuous movement, with the vessel being connected to the gear throughout the entire fishing event, which usually lasts up to some hours. Meanwhile, a passive fishing event is composed of two phases: the first is the start of the fishing event, which is the gear deployment, and then it ends with the second phase of the fishing event, which is when the gear is retrieved onboard, i.e. the hauling of the gear. The period between these two phases is called soak time, and it can take hours to days or even weeks. During this period, the vessel is detached and away from the gear.

The existing studies on static gears have mainly focused on mapping fishing activities by identifying the moments when vessels hauled their gears. Hauling events are characterized by low vessel speed, usually in an approximately linear direction. These events are quite distinctive from other states of a fishing trip, like navigation or gear deployment (Marzuki *et al.*, 2018; Mendo *et al.*, 2019a). Yet, to estimate the fishing effort of passive gears in terms of soak time, the beginning of the fishing event, i.e. the deployment of the gear, needs to be identified. It is critical to identify deployment, soaking of gears, and hauling in the correct sequence. In fact, let's imagine the case in which one of the deployment or hauling phase occurring during a fishing day is absent within the spatial data: The corresponding estimation of soak time would be dramatically biased.

When the behaviour of vessels deploying a static gear is different from steaming, the identification of the deployments was shown to be possible (Charles *et al.*, 2014). But, in cases where the behaviour of the vessel when deploying is too similar to when it is steaming, this poses an additional challenge to identify gear deployments using only the behaviour variables of the vessel (Mendo *et al.*, 2019a) and therefore to calculate the soak time of that particular gear.

The quantification of soak time allows identification of illegal fishing behaviours, as some gears are not allowed to fish for more than a certain period of time (EU regulation No. 227/2013). It is also important in a stock assessment context, not only because captures vary, in quantity and variety, depending on the time a gear is fishing, but also because soak time affects catch rates of fishing gears (Boutillier and Sloan, 1987; Erzini *et al.*, 1997; Ward *et al.*, 2004; Morgan and Carlson, 2010; Li *et al.*, 2011).

In this study, we propose an approach to dealing with high-resolution geospatial data from a polyvalent (multi-gear) passive gear fishing fleet. We used AIS data from the Portuguese mainland coastal fishing fleet operating two main fishing gears: bottom nets (gillnets and trammel nets) and pots and traps. The methodology is primarily devised to be applied to single vessels to identify and classify the different behaviours within a trip: steaming, gear deployment, hauling of gears, and slow navigation. With the presented methodology, we aim to assess the distribution of this type of fishery and to provide insights on its effort, as soak time, by identifying the start (deployment) and the end (hauling) of a passive gear fishing event.

As far as the authors are aware, this is the first study to present a methodology to address the distribution and quantification of fishing effort, in terms of soak time, for a polyvalent passive gear fishing fleet using geospatial data.

Methods

Data

Fishery-dependent and spatial data from the Portuguese coastal polyvalent fishing fleet (length overall—LOA >12 m) using nets, pots, and traps, from 2014 to 2020, were used. This fleet operated within the continental Portuguese economic exclusive zone (EEZ; Figure 1).

The Directorate-General for Natural Resources, Safety and Maritime Services (DGRM) provided fishery-dependent data in the form of daily landings and logbook data. For spatial data, we used land-based AIS data collected from the AIS repository AIShub (www.aishub.net) and from MarineTraffic (www.marinetraffic.com).

To train and validate the classification models, two sources of data were used: (1) onboard GPS-collected data and (2) manually labelled AIS data. The onboard GPS data was collected from 32 fishing trips, from 9 different fishing vessels, of which 5 used nets (gill nets and trammel nets) and the remaining 4 vessels operated pots and traps. Fishing trips were classified into four different phases: steaming, deployment, hauling, and slow navigation. The latter corresponds to when a vessel is either drifting or slowly navigating, usually waiting for a gear to fish before starting the hauling process. The classified GPS and AIS data were then split into training and validation data.

For a summary on the raw AIS data, details of the process on selecting the AIS data from the desired vessels, and a description of the data used to train and validate the classification models, check the Supplementary material, Section 1.

A schematic representation of the entire workflow is represented in Figure 2.

Data pre-processing

Pre-processing of the data included: cleaning erroneous data points; split the AIS data into singular fishing trips; remove inadequate/incomplete fishing trips and interpolating the track data to a consistent frequency.

To reconstruct fishing tracks to have a timely consistence of datapoints, the tracks were interpolated using the algorithm developed by Russo *et al.* (2011, 2014) that relies on the Catmull–Rom approach, a modification of the hermit cubic spline algorithm (Tremblay *et al.*, 2006; Hintzen *et al.*, 2010). The interpolation was set to generate data points at 1-minute intervals, as this frequency has shown to be best suited to identify passive fishing events (Mendo *et al.*, 2019b). For a more detailed description of these steps, check the Supplementary material, Section 2.

Set up the classification variables

The approach intends to classify the AIS data into four different phases of a fishing trip: steaming, deploying, hauling,

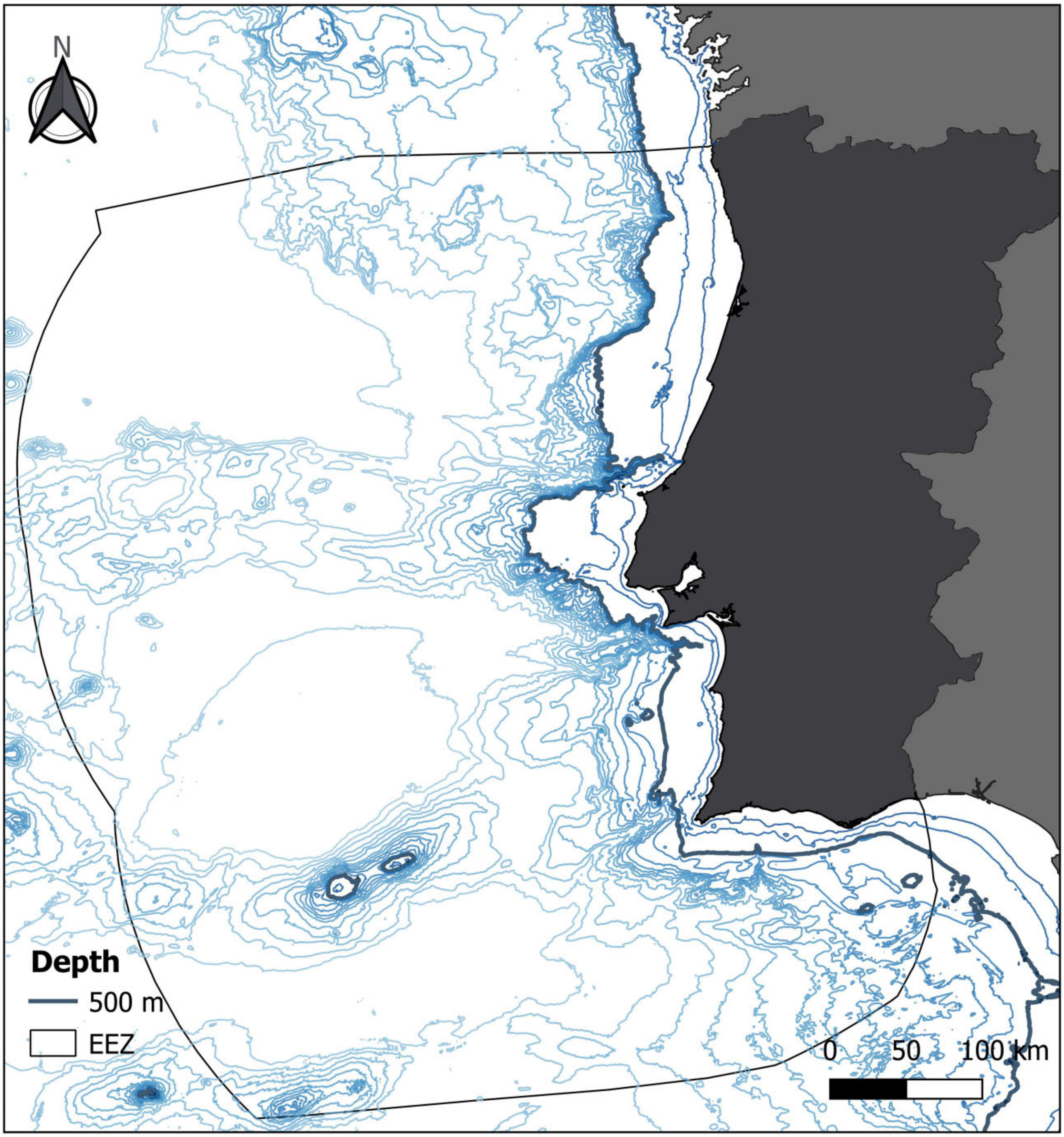

Figure 1. The current approach focuses on the Portuguese polyvalent fishing fleet operating within the Portuguese EEZ.

and slow navigation. To do so, we defined three classification variables: (1) vessel speed; (2) future overlap (FO); and (3) past overlap (PO). Vessel speed aided classification in states by observing that steaming and deployment were carried out at high navigation speeds, while slow navigation and hauling of gears were carried out at lower speeds (Figure 3 and Supplementary Figure S1).

The FO and PO aided in distinguishing (1) deployment events from steaming and (2) hauling events from slow navigation or drifting, respectively. We relied on the fact that if a gear is deployed, it must be hauled with vessels running the same path when deploying and hauling a gear. To identify a gear deployment, there must be overlap with a hauling track (carried out at a slow speed), happening in the future in relation to the deployment track. Likewise, to identify the hauling of a gear, there must be an overlap with a deployment track (carried at a fast speed), happening in the past in relation to the hauling track (Figure 3).

Given the previous rationale, we determined a distance to establish if two datapoints overlapped. To do so, we calculated the distance of 3620 deployment datapoints to the closest hauling datapoint, within the same fishing event and conversely, 14690 distances from the hauling datapoints to their closest deployment datapoints of the same fishing event (Supplementary Figure S2). With the calculated distances, we defined two threshold distances: (1) future overlap distance (FOD), being the threshold distance from a deployment datapoint to a hauling datapoint occurring in the future in

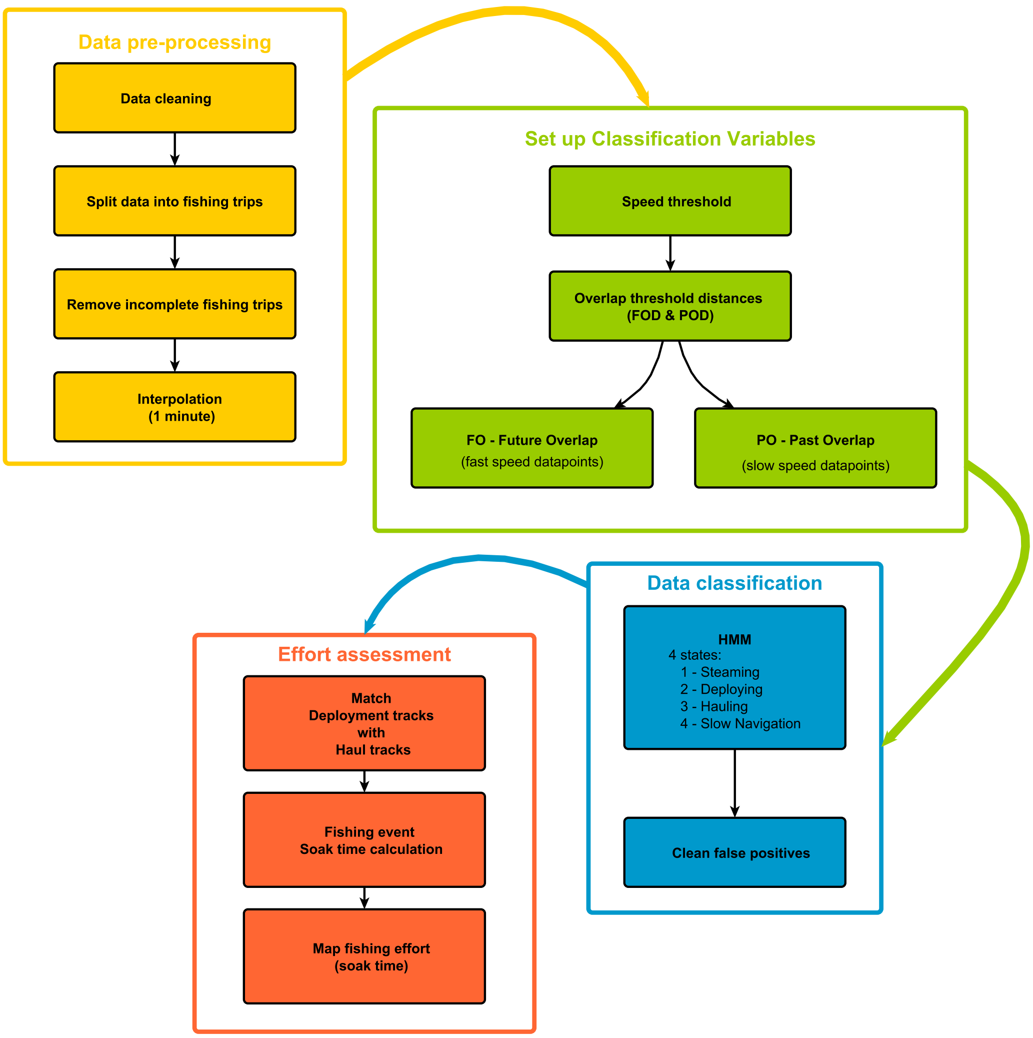

Figure 2. Schematic representation of the major steps carried out throughout our procedure. For a more detailed description of the data and the data selection procedure, check the Supplementary material.

relation to the deployment datapoint; and (2) past overlap distance (POD), being the threshold distance from a hauling datapoint to deployment datapoint that occurred in the past in relation to the hauling datapoint. These threshold distances were then defined using the normal distribution to model the distance value at which $P\ (X \leq x) = 0.99$, using the mean and standard deviation obtained from the calculated distances. The established distance values for FOD and POD were 150 m and 235 m, respectively (Supplementary Figure S3).

When calculating the minimum distances between datapoints, an important aspect to consider was the maximum soak time. The range of the soak time of the two gear types is very wide, and to identify the overlapping deployment and hauling datapoints, we precautionarily selected a time window that spanned from 90 minutes to 20 days. This window assumed that after a deployment, the gear is fishing for at least 90 minutes (in the case of nets) and that it can be soaking/fishing for up to 20 days (pots/traps).

Considering the parameters and assumptions previously described, for each datapoint with speed values corresponding to possible deployment or steaming events ($\geq$3.6 knots), we assigned a value equal to one to the binary variable FO. The assigned value of one means that for that given datapoint, there is an overlapping datapoint within the established threshold distance, at a low speed (<3.6 knots), 90 minutes to 20 days ahead in time to the referred datapoint. For high-speed datapoints without overlapping datapoints, and all low-speed datapoints were assigned zero for the FO variable. Conversely, for low speed datapoints (possible hauling or slow navigation/drifting datapoints), we assigned the PO binary variable, where value 1 means the existence of an overlapping datapoint, with speed of a deployment datapoint ($\geq$3.6 knots) that was generated 90 minutes to 20 days in the past, in relation to the low speed datapoint. Low-speed datapoints without overlapping datapoint and all datapoints with speed values $\geq$3.6 knots were assigned zero for the PO variable.

We are aware that in the case of pots and traps, fishers do not always haul the entire gear on deck and then deploy it elsewhere. Instead, fishers may run slowly along the line checking if there is any catch in each trap/pot whilst leaving the entire set in the water, and this behaviour can be repeated indefinitely for a given gear. This means that the overlapping tracks are all performed at a low speed. At the present, the identification of this particular fishing behaviour is beyond the scope of our approach.

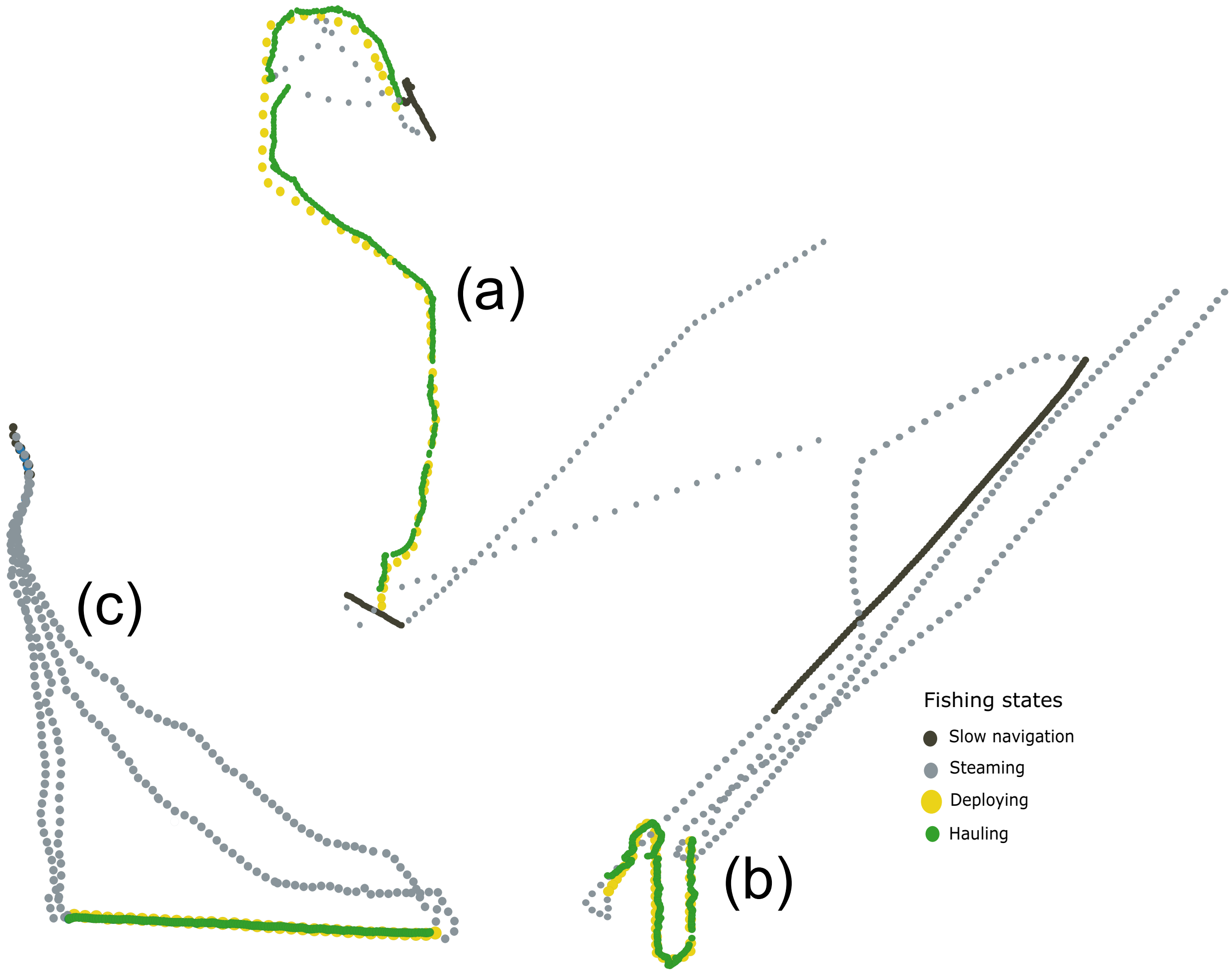

Figure 3. AIS tracks with three different fishing events. Fishing events are represented by the deployment datapoints (in yellow) that are displayed underneath the hauling datapoints (in green) when they overlap in space, meaning that a fishing event, as expected, starts with the deployment of a gear, at a fast navigation speed and terminates with its hauling, which is carried at low speed. The tracking data was interpolated at a 1-minute rate, meaning that the larger the distance between consecutive datapoints, the faster the vessel navigates, and *vice versa*. The behavioural diversity and complexity of this fishery are evident from this figure: (a) vessels can deploy two sets of fishing gears in a single event and haul them within the same trip; (b) deploy just one gear and after the deployment the vessel waits some hours, while slowly navigating/drifting, and then collect the gear within that same trip; or (c) a vessel can go out to sea to deploy a gear and then return on a following day to haul the gear. The soak time will depend on the type of gear and target specie.

For a complete description on the process of setting up the classification variables, check the Supplementary material, Section 3.

Data classification

The R package MomentuHMM (McClintock and Michelot, 2018) was used to classify each of the data points into one of the four different underlying hidden states of a fishing trip: steaming, deploying, hauling, and slow navigation.

We used three different variables in the HMM analysis: (1) speed, (2) future overlap (FO), and (3) past overlap (PO). To model speed values, we used the positive gamma distribution, and given that datapoints were interpolated into regular time intervals of 1 minute, we used the standard approach of MomentuHMM and converted speed values into distance values between datapoints (step length) as a proxy for speed. The distribution parameters of each fishing state, calculated from the training data, were used as the initial distribution parameters of the HMM model. Since the FO and PO were established as binary, we used the Bernoulli distribution to model these two variables. After assigning the overlap variables to the training data, we assessed the probability of FO = 1 and PO = 1 for each of the four states. Just like for the step parameters, these probability values, directly calculated from the training data were used as initial parameters for the HMM model. For a further description of the data classification with the HMM, check the Supplementary material, Section 4.1.

Because the HMM classification alone yielded an undesirable number of false positives for both states of the fishing event (deployment and hauling), the next step was to clean the miss labelled fishing datapoints. A description on the cleaning of false positives is described in Supplementary material, Section 4.2.

After establishing the distribution parameter for the three variables considered in the HMM analysis, and developing the procedure to clean false positives, we ran the entire classification analysis on the validation dataset. The rational of the assessment of the performance of the classification procedure is based on the comparison of the accuracy of the procedure's

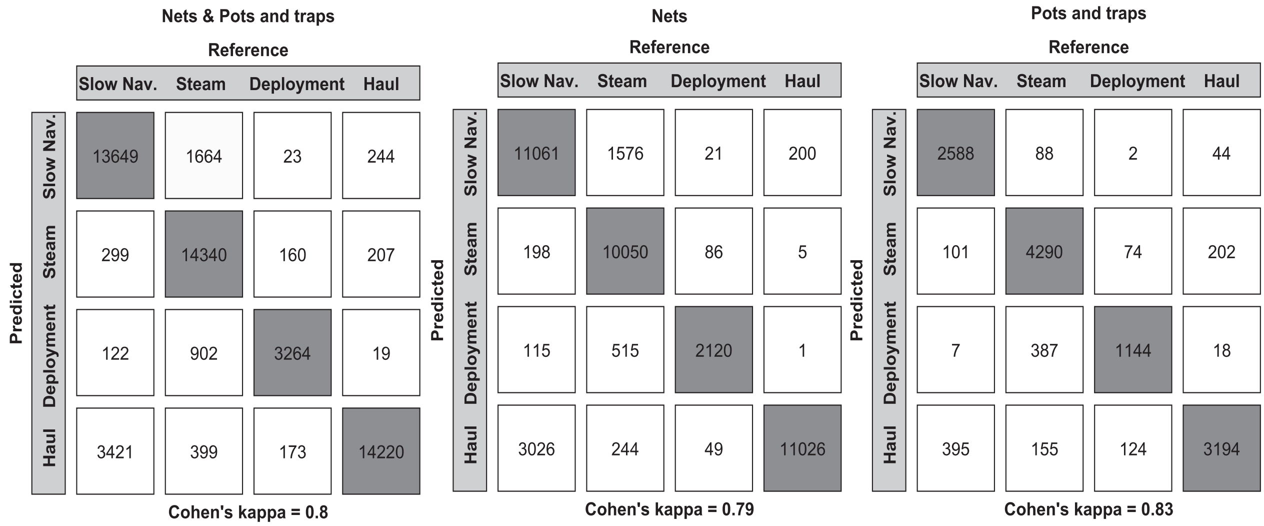

Figure 4. Confusion matrices and Cohen's coefficient values from the validation data classified by the procedure. Confusion matrices plot the number of datapoints, of each state, classified by the model (Predicted) against the actual number, of each state, of the validated data (Reference). The validation procedure was carried for the overall validation dataset, i.e. for all fishing trips regardless of the gear used (nets, pots, and traps) and for fishing trips using the same type of gear: nets or pots and traps.

output with the onboard ground-truthed and the manually labelled data, used as test/validation dataset to assess the model performance. For a more detailed description on the model validation, see Supplementary material, Section 4.3.

Footprint and effort assessment

After the classification of datapoints into the different states of a fishing trip, the last step was to identify the fishing events by matching the corresponding start (deployment) and end (hauling) of the fishing events. This is a critical step, as both moments of a passive fishing event need to be correctly matched to calculate the soak time of a gear. The approach to identify the fishing events was carried for each vessel, by matching each classified deployment datapoint with the most recent hauling datapoint that falls within the pre-established time window (90 minutes to 20 days) ahead of the deployment datapoint timestamp and within the FOD in relation to the deployment datapoint. For the full explanation on the process of identification of the fishing events and calculation of the soak time, check the Supplementary material, Section 5.

To map the fishing footprint of this fleet, we divided the study area into a 1×1 km grid. The footprint was then assessed through the presence/absence of the identified fishing events in each grid cell.

To map and calculate the total effort, as total soak time, we used the same 1×1 km grid. The effort was then calculated as the sum of the averages of soak time of each fishing event's datapoints present within each grid unit. The Effort formula for each grid unit can be represented as such:

$$\text{Effort} = \sum_{f=1}^{n} \frac{\sum_{j=1}^{N} S_{fj}}{N},$$

where j is the fishing vessel, f is the fishing event, S is the soak time, i.e. the difference between timestamps of the matching deployment and haul datapoints, and N is the number of pairs of matched datapoints, of each fishing event, within the grid square unit.

As already stated, AIS data have some limitations in terms of inconsistent spatial and temporal coverage (Metcalfe *et al.*, 2018; Shepperson *et al.*, 2018). For that reason, we studied the proportion of total effort that we were able to quantify and map. To do so, we compared the number of AIS fishing trips that passed all the steps of our approach, with the number of fishing trips of the same vessels, inferred from the landings dataset, under the assumption that each landing event corresponds to a fishing trip.

Results

Model validation

After setting the parameters that yielded the best accuracy values for the training data, which is the highest value of the average between the rate of true positives (Sensitivity) and the rate of true negatives (Specificity), we ran the classification procedure on the validation data and assessed its accuracy. The accuracy was assessed through confusion matrices for all the fishing trips, regardless of the gear used. To evaluate the classification performance for each type of gear used, we also carried out the validation of the model for trips using only nets and for trips operating only pots and traps (Figure 4 and Table 1). For a further description on the model validation, check the Supplementary material, Section 4.3.

Overall, the model performed well in identifying the start and end of fishing events, with an accuracy of more than 90% for both phases of the fishing events. Comparing the two types of gears, the model, as it was set up, seems to be better suited to identify fishing operations with nets than with pots and traps. The classification accuracy of deployments of pots and traps was the lowest, with only 85% of the deployment datapoints classified as such by the algorithm and because of that, the Balanced Accuracy for the deployment of pots and traps was the lowest. On the other hand, the model returned a higher rate of false positives for hauling events when classifying fishing events using nets (Specificity = 89%).

Through the analysis of Cohen's coefficient, the results of the classification demonstrate an overall strong agreement

Table 1. Results of the accuracy assessment for the deployment and hauling states for both types of gear.

Fishing state	All gears		Nets		Pots and Traps	
	Deployment	Haul	Deployment	Haul	Deployment	Haul
Sensitivity (recall)	0.90	0.97	0.93	0.98	0.85	0.92
Specificity	0.98	0.90	0.98	0.89	0.96	0.93
Balanced Accuracy	**0.94**	**0.93**	**0.96**	**0.93**	**0.91**	**0.93**
Number of vessels	11		7		4	
Number of trips	34		22		12	

The performance of the model was assessed through the Sensitivity (true positive rate), the Specificity (true negative rate), and the Balanced Accuracy, which is the arithmetic average of the Sensitivity and the Specificity of the model.

between the reference and the predicted data: kappa ≥0.8 (McHugh, 2012), especially when classifying fishing trips with pots and traps.

Soak time

The soak time of each fishing event was assessed through the difference between the timestamp of corresponding deployment and hauling datapoints. From the analysis of the gear used in each fishing event recorded in the logbook data, the analysed vessels commonly use both types of gear. Of the 146 vessels, 7 vessels reported using only one type of gear, with 6 vessels using only nets and one using just pots and traps during the period of 2014–2020.

To visually assess the distribution of the soak time of fishing events, we calculated the average soak time between matching deployment and haul datapoints of each fishing event (Average soak time). As expected, the distributions of soak time of these two types of gears are different. Pots and traps are known for being left fishing for a long period, ranging from some days to several weeks, while fishing with nets usually takes less time, from hours to a maximum of a few days (see example in Figure 5a and b). For most of the vessels that used both gears, the distribution of soak time, depending on the proportion of use of each type of gear, is expected to resemble, to some extent, the combination of the distributions of both types of gears (Figure 5c).

Footprint and effort

Applying the selection criteria for this study, i.e. fishing vessels of LOA > 12 m, operating in the waters of mainland Portugal and using only nets and/or pots and traps, from 2014 to 2020, resulted in a list of 301 fishing vessels. When querying the AIS databases from the list of these fishing vessels, and after cleaning the AIS dataset of erroneous datapoints, we started our analysis with ∼85.4 million AIS datapoints, from 151 vessels.

After the procedure of identification of the fishing trips and removing fishing tracks with gaps, 21950 fishing trips from 146 vessels remained to be interpolated and classified by the HMM procedure. Overall, the data availability, combined with the data requirement of the procedure, enabled us to classify 21.7% of the fishing trips carried out by these 146 vessels, during the period of 2014–2020 (Table 2). An example of the classified data is shown on Supplementary material section, Supplementary Figure S5.

To map the footprint (Supplementary Figure S6) and the effort (Figure 6) of this fishery, we selected the classified fishing trips containing matching states of the fishing events (deployments and hauls). This selection criteria, along with the existence and quality of the AIS data, allowed us to keep 13224 out of the 21950 classified fishing trips, from 84 of the 146 vessels from the classification dataset, corresponding to 12.1% of the total number of trips carried out by the number of vessels within the initial AIS dataset. From the remaining fishing trips, we were able to quantify 24353 complete fishing events, carried out by 84 vessels (Table 2).

Discussion

We introduce a method to *a posteriori* map the fishing effort, in terms of soak time, of passive fishing gears used by a polyvalent fishing fleet, using data from high-resolution tracking devices. The method relies on the ability to classify vessel tracking data into a set of *a priori* defined states, and in particular the beginning (deployment) and end (haul) of fishing events. By doing this for all vessels, it allows the footprint of the fishery to be studied and to quantify and explore, with high resolution, the spatial and temporal dynamics of the fishing effort of these polyvalent fishing vessels at an aggregated (e.g. fleet segment) scale.

Previous studies have also addressed the spatial dynamics of passive fishing gears through vessel tracking data. Mendo *et al.* (2019a) and Jennings and Lee (2012), for example, mapped the footprint of a passive fishery and studied the distribution of the fishing effort. The approach in Mendo *et al.* (2020) allows assessing the distribution of the footprint of a passive fishery, but it does not calculate the fishing effort. Meanwhile, the method used by Jenning and Lee (2012) defines the fishing effort as the time vessels spent hauling their gears, which is different from the time gears fished (soak time). Campos *et al.* (2023) took a different approach: The fishing effort was mapped and calculated as the distance carried by the vessels during the period of the fishing event as recorded on the logbook. This approach may overestimate the location and total length of fishing events, as the vessel after deploying a fishing gear can navigate while waiting for the gear to fish, or it can even deploy other gears before finishing (hauling) the initial fishing gear.

The proposed methodology brings the possibility of identifying the full duration of passive fishing events, which, until now, was only possible by means of onboard observers or from logbook data. Monitoring with onboard observers is very costly and generally only covers a small fraction of the fleet, while logbooks are compiled by fishers and their quality depends on the willingness of fishers to precisely log every important detail of the fishing activity, which is not always the case (Bastardie *et al.*, 2010; Sampson, 2011; Russo *et al.*, 2016a).

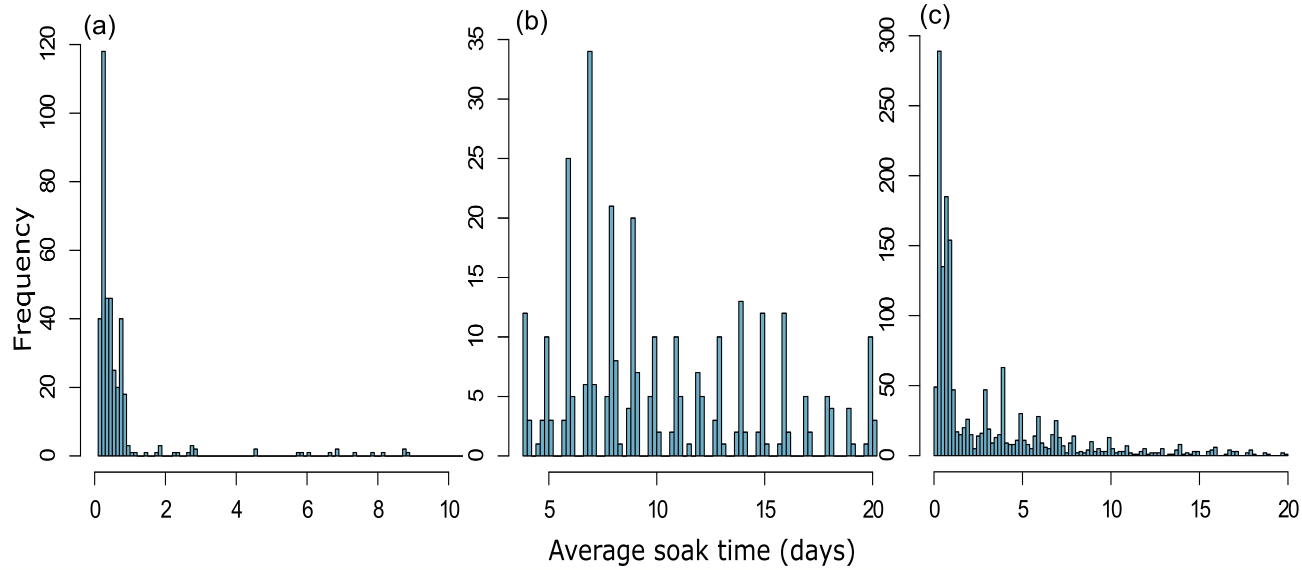

Figure 5. Distribution of the average soak time, in days, of each fishing event of one vessel using only nets (a), one vessel using only pots and traps (b), and of one vessel both types of gears (c). The average soak time of each fishing event was calculated as the average difference of timestamp between pairing deployment and hauling datapoints of a given fishing event. Vessels using nets tend to fish during short periods of time, usually up to one or two days, while vessels using pots and traps tend to leave the gears fishing for longer periods. The distribution of soak time of vessels using both types of gear is expected to resemble the combination of the distributions of nets (high number of fishing events with short periods of soak time) with the distribution of soak time values of pots and traps (a wider distribution of fishing events that can range for several days). The average values of soak time for each fishing event were calculated to the resolution of one decimal place. Meaning that a fishing event that lasts 0.5 days (12 hours) it is represented on a different bar that of a 24 hours (1 day) fishing event, for example.

Table 2. Data composition from the application of the vessel selection criteria until the final steps of the procedure.

	Initial list of vessels with AIS data	Classification dataset	Footprint and effort dataset
Number of vessels	151	146	84
Number of trips (from landing data)	109 006	101 152	49 876
Number of trips (from AIS)	NA	21 950	13 224
Proportion inferred from AIS	NA	21.7%	26.5%
Number of fishing events (from AIS)	NA	NA	24 353

Overall number of vessels, trips, and fishing events that were kept and identified from the AIS data, during the period 2014–2020, throughout the three main steps of our procedure.

The proposed method was developed and set up to be applied to a polyvalent passive fishery, regardless of which gears, with overlapping deployment and hauling tracks. Despite the good performance of the classification algorithm (>90% for deployment and hauling events), when analysing the accuracy for both gears separately, there is indeed a difference depending on gears and events. As the model was set up, the accuracy detecting fishing events using nets outperformed the ones using pots and traps. The deployments of pots and traps had a lower rate of true positives compared to the deployment of nets. The reason for this is probably related to the fact that in some cases, the deployment of pots and traps happens at a lower speed than the defined overall "fast speed threshold" (≥3.6 knots). On the other hand, when assessing the specificity (rate of true negatives) of hauling datapoints, fishing events with nets returned a higher number of false positives. This has to do with the behaviour of the vessel after deploying a net. The vessel commonly stays in the vicinity of where the net was deployed, usually slowly navigating or drifting, waiting for the net to fish. This proximity to the deployment or steaming tracks increases the chances of slow navigation datapoints being assigned the PO variable and consequently being classified as hauling datapoints. Contrary to most fishing events with nets, when fishing with pots and traps, the vessel leaves the area after deploying the gear, as these gears are almost never hauled within the same deployment trip. Therefore, the rate of false positives on hauling events using pots and traps was lower.

As anticipated, fishing events of nets, pots, and traps vary distinctly in terms of soak time. Fishing with nets usually means leaving the net fishing for some hours up to a few days, as caught fish tend to die and risk being eaten by scavengers. Also, the legal maximum soak time is 24 hours, except for nets of mesh size larger than 100 mm operating in zones of depth >300 m (Portuguese Ordinance No. 1102-H/2000). This exception is associated with particular fishing metiers targeting, for instance, monkfish (*Lophius* spp.; Szynaka *et al.*, 2021, 2022). In contrast, sheltering or trapping gears like pots and traps, the catches remain alive, allowing these gears to fish for much longer periods. Furthermore, unlike gillnets and

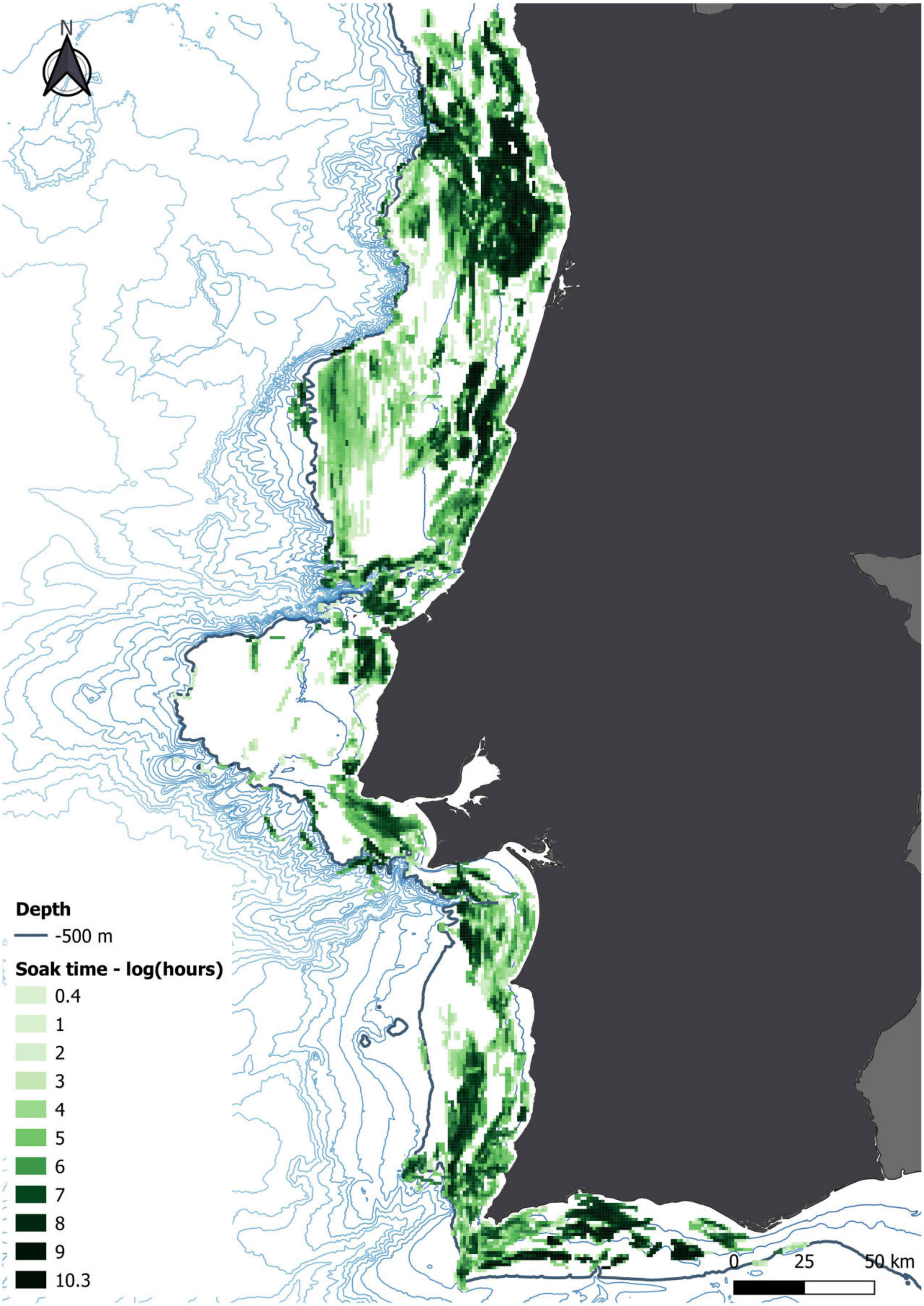

Figure 6. Map of fishing effort (soak time) of fishing events using nets, pots, and traps, during the period from 2014 to 2020. The resulting effort map represents 24353 fishing events, from 13224 fishing trips carried by 84 fishing vessels. The majority of the fishing activity is distributed within the 500 m depth, and the fishing effort presents a patchy distribution appearing to be more relevant in the northern part of the study area.

trammel nets, maximum catch rates of traps are often a week or more after deployment (Erzini *et al.*, 2008). Still, according to logbook data, one fishing vessel using nets had fishing events reaching up to 9 days of soak time. This either mean that this vessel used a gear such as pots or traps, or that it left nets fishing for longer periods than the maximum soak time permitted by law. In any case, neither the information about the use of pots and traps nor net soak times of up to 9 days of fishing were reported in the logbook, making this approach also useful to monitor the compliance of fishers with existing regulations, such as maximum soak time.

Within the analysed fleet, there were only seven vessels that reported using only one type of gear. For this reason, the distribution of soak time of vessels using both gears was studied. As expected, the distribution of the soak times of these vessels resembles the combination of the distribution of vessels that only use one type of gear, with several fishing events lasting around one and two days (nets) and several fishing events lasting up to 14 days, with a few soak times of up to 20 days (pots and traps).

The fishing footprint of this fleet is mainly located within the continental shelf, extending to the upper part of the continental slope, as also shown by Leitão *et al.* (2022). Fishing effort was patchily distributed and seems to be more relevant in the northern part of the study area and along two submarine canyons. It is important to stress that these maps do not represent the whole reality of this fleet. In fact, they represent the footprint and the fishing effort of 12.1% of the fishing trips carried by the 151 vessels included in the initial AIS dataset.

There are several issues with AIS systems, one of which is potential poor fishing fleet coverage (Russo *et al.*, 2016b, 2019; Shepperson *et al.*, 2018). Also, land-based AIS systems, even though more accurate than satellite-based AIS, have the disadvantage of having a limited range and are dependent of the existence of AIS antennas on land. Moreover, the signal transmission can be affected by external factors, such as weather conditions, and it can be deliberately switched off in case the skipper does not want to disclose his vessel's whereabouts (Russo *et al.*, 2016b; Emmens *et al.*, 2021). So, the footprint and distribution of fishing effort presented here mostly depend on the AIS fleet coverage, the signal reception infrastructure (existence of AIS antennas), and on the willingness of the skippers to leave the AIS transponder ON throughout the entire trip. In fact, when sorting the classified data from the 146 vessels to map and quantify the fishing effort, 62 vessels and 8726 fishing trips were discarded, as these trips did not include fishing events. From a GIS-based assessment of these incomplete trips, we saw that most of these vessels had the AIS transmitting while leaving the port, probably as a safety measure to avoid collisions with other vessels, but once at sea, the AIS would stop transmitting.

Another specificity of this approach that requires consistent data has to do with the fact that the identification of the start and end of a fishing event is dependent on each other. This means that in case the tracking data of a deployment is missing, the hauling data, even though existent, will not be classified as hauling because of the inexistence of the Past Overlap variable, but instead will be classified as slow navigation. A similar situation happens in case the data for the hauling event is missing and the deployment track is present: The deployment track will be classified as steaming, since the FO variable is missing. It is also important to stress that the frequency of datapoints generated by the tracking devices is of the upmost importance. This approach is able to identify the full duration of passive fishing events because the average frequency of datapoints generated by the AIS system is of 3 minutes (Supplementary Table S1). If the data frequency would be lower, the identification of the deployment events would probably not be possible, as these events can take 15 minutes to be completed.

The last process of grouping deployment datapoints into a single fishing event requires prior knowledge about the fishery being studied. By observing the training data, where deployments always took more than 17 minutes to be carried out, and given that these vessels deploy gears with several kilometres in length, we set a conservative threshold of 15 minutes, as the minimum time a deployment would need to be carried out. This value was set based on the specificity of this type of fishery, and indeed, if a vessel deploys a smaller gear, or takes <15 minutes to deploy it, then the fishing event would be discarded.

The development and tailoring of this methodology cannot ignore the specificities of the different fisheries it is intended to be applied to. The fact that this procedure allows the study of more than one type of gear or fishery is one of its major advantages. But, at the same time, because nets, pots, and traps are not handled exactly the same way, nor have the same distribution of soak time, there is the compromise of setting up parameters suitable to model different gears, which may be less suitable from those to model single-gear fisheries. For example, choosing a time window from 90 minutes to 20 days to calculate the overlapping distances between datapoints, does not make much sense from a pots and traps standpoint, as the soak time of these gears is longer than 1 or 2 days, and net presumably do not fish for up to 20 days. But because we are dealing with two types of gears that do differ considerably in terms of soak time, we needed to choose a time window that could accommodate the soak times of both gears. This is a trade-off that must be considered when dealing with a polyvalent fishery. But in case this method is used in fisheries with just one type of gear, or a particular fishery, then the parameters, such as speed and time window, can be set according to the peculiarities of the studied fishery or gear.

The comprehensive knowledge about fisheries distribution and the quantification of fishing effort is fundamental to improve ocean governance as well as to improve fisheries management and marine conservation (Halpern *et al.*, 2008; Campbell *et al.*, 2014; McCauley *et al.*, 2016; Vespe *et al.*, 2016). Yet, the precise and complete understanding of this sort of information is still lacking (Kroodsma *et al.*, 2018; Leblond *et al.*, 2019), especially regarding polyvalent fisheries, which comprise the majority of the fishing fleets worldwide (Kelleher *et al.*, 2012). Improving the resolution of fishing effort through the knowledge of where and for how long a gear is fishing will allow us not only to map the distribution of the fishing effort, but will also provide us with better estimates for stock assessment, better monitor fishing activities, and assess the relationships between soak time and landing composition or bycatch. By increasing the coverage of the different fishing fleets with high-resolution and high-frequency datapoint generation tracking devices, it will be possible to apply approaches such as the one described in this paper to a whole fleet segment, contributing to ecosystem-based approaches and improve management and conservation of the marine realm.

Funding

This work was funded in part by Portuguese national funds from the FCT—Foundation for Science and Technology through projects UIDB/04326/2020, UIDP/04326/2020, and LA/P/0101/2020 to CCMAR; NSH was funded by the FCT (Fundação para a Ciência e Tecnologia) PhD scholarship ID nr: 2020.05583.BD.

Acknowledgements

The authors would like to thank Magda Frade and Flávia Carvalho for helping in the collection of the onboard validation and training data. We would like to express our gratitude to the several skippers and crew members for sharing their valuable knowledge and for allowing us to join their fishing trips to collect the validation and training data. We would like to thank MarSensing Lda and Friedrich Zabel for providing and helping dealing with the AIS data. We are grateful to Jan Jaap Poos and the two anonymous reviewers who's comments and suggestions helped to improve the manuscript. We are thankful to the DGRM—Directorate-General for Natural Resources for providing the fishery-dependent data necessary to carry out this study.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Conflict of Interest

The authors have no conflict of interest to declare.

Data availability

The data underlying this article cannot be shared publicly since it provides the exact location of where fishing activities occur and contains confidential information. Therefore, we do not find ethically feasible to share this data. Moreover, under the confidentiality agreement with the Portuguese Directorate-General for Natural Resources (DGRM), we are not allowed to share this data as it contains private and confidential information about the fishing vessels. Yet, the R code used in the presented methodology is available at: https://github.com/NunoSH/Polyfish.git

Authors contribution

Conceptualization: NSH, TR, LB, PM; Data collection: NSH, LB, PM; Methodology: NSH, TR; R Script development: NSH, TR, AP, RM; Formal analysis: NSH; Supervision: JMSG, TR, KE; Funding acquisition: JMSG, KE; Writing—original draft: NSH; Writing—review & editing: NSH, TR, LB, PM, FO, KE, JMSG. The authors have no conflict of interest to declare.

References

Bastardie, F., Nielsen, J. R., Ulrich, C., Egekvist, J., and 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.

Boutillier, J. A., and Sloan, N. A. 1987. Effect of trap design and soak time on catches of the British Columbia prawn (*Pandalus platyceros*). Fisheries Research, 6: 69–79.

Campbell, M. S., Stehfest, K. M., Votier, S. C., and 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.

Campos, A., Leitão, P., Sousa, L., Gaspar, P., and Henriques, V. 2023. Spatial patterns of fishing activity inside the Gorringe bank MPA based on VMS, AIS and e-logbooks data. Marine Policy, 147: 105356.

Charles, C., Gillis, D., and Wade, E. 2014. Using hidden Markov models to infer vessel activities in the snow crab (*Chionoecetes opilio*) fixed gear fishery and their application to catch standardization. Canadian Journal of Fisheries and Aquatic Sciences, 71: 1817–1829.

de Souza, E. N., Boerder, K., Matwin, S., and Worm, B. 2016. Improving fishing pattern detection from satellite AIS using data mining and machine learning. PLoS One, 11: e0158248.

Emmens, T., Amrit, C., Abdi, A., and Ghosh, M. 2021. The promises and perils of automatic identification system data. Expert Systems with Applications, 178: 114975.

Erzini, K., Bentes, L., Coelho, R., Lino, P. G., Monteiro, P., Ribeiro, J., and Gonçalves, J. M. S. 2008. Catches in ghost-fishing octopus and fish traps in the northeastern Atlantic Ocean (Algarve, Portugal). Fishery Bulletin, 106: 321–327.

Erzini, K., Monteiro, C. C., Ribeiro, J., Santos, M. N., Gaspar, M., Monteiro, P., and Borges, T. C. 1997. An experimental study of gill net and trammel net 'ghost fishing' off the Algarve (southern Portugal). Marine Ecology Progress Series, 158: 257–265.

Gerritsen, H., and 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.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., Bruno, J. F. *et al.* 2008. A global map of human impact on Marine Ecosystems. Science, 319: 948–952.

Hintzen, N. T., Piet, G. J., and Brunel, T. 2010. Improved estimation of trawling tracks using cubic Hermite spline interpolation of position registration data. Fisheries Research, 101: 108–115.

Jennings, S., and Lee, J. 2012. Defining fishing grounds with vessel monitoring system data. ICES Journal of Marine Science, 69: 51–63.

Kelleher, K., Westlund, L., Hoshino, E., Mills, D., Willmann, R., de Graaf, G., and Brummett, R. 2012. Hidden Harvest: The Global Contribution of Capture Fisheries. 66469-GLB. The World Bank ; FAO.

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A. *et al.* 2018. Tracking the global footprint of fisheries. Science, 359: 904–908.

Le Guyader, D., Ray, C., Gourmelon, F., and Brosset, D. 2017. Defining high-resolution dredge fishing grounds with automatic identification system (AIS) data. Aquatic Living Resources, 30: 39.

Leblond, E., Woerther, P., Woillez, M., and Quemener, L. 2019. Sensor Systems for an Ecosystem Approach to Fisheries. In Challenges and Innovations in Ocean In Situ Sensors, pp. 173–288. Eric Delory. and Jay Pearlman (eds). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128098868000053(Accessed 11 May 2022).

Lee, J., South, A. B., and Jennings, S. 2010. Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. ICES Journal of Marine Science, 67: 1260–1271.

Leitão, P., Sousa, L., Castro, M., and Campos, A. 2022. Time and spatial trends in landing per unit of effort as support to fisheries management in a multi-gear coastal fishery. PLoS One, 17: e0258630.

Li, Y., Jiao, Y., and Reid, K. 2011. Gill-net saturation in Lake Erie: effects of soak time and fish accumulation on catch per unit effort of walleye and yellow perch. North American Journal of Fisheries Management, 31: 280–290.

Marzuki, M. I., Gaspar, P., Garello, R., Kerbaol, V., and Fablet, R. 2018. Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories. IEEE Journal of Oceanic Engineering, 43: 689–699.
McCauley, D. J., Woods, P., Sullivan, B., Bergman, B., Jablonicky, C., Roan, A., Hirshfield, M. *et al.* 2016. Ending hide and seek at sea. Science, 351: 1148–1150.
McClintock, B. T., and Michelot, T. 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods in Ecology and Evolution, 9: 1518–1530.
McCluskey, S. M., and Lewison, R. L. 2008. Quantifying fishing effort: a synthesis of current methods and their applications. Fish and Fisheries, 9: 188–200.
McHugh, M. L. 2012. Interrater reliability: the kappa statistic. Biochemia Medica, 22: 276–282.
Mendo, T., Smout, S., Photopoulou, T., and James, M. 2019a. Identifying fishing grounds from vessel tracks: model-based inference for small scale fisheries. Royal Society Open Science, 6: 191161.
Mendo, T., Smout, S., Russo, T., D'Andrea, L., and James, M. 2019b. Effect of temporal and spatial resolution on identification of fishing activities in small-scale fisheries using pots and traps. ICES Journal of Marine Science, 76: 1601–1609.
Metcalfe, K., Bréheret, N., Chauvet, E., Collins, T., Curran, B. K., Parnell, R. J., Turner, R. A. *et al.* 2018. Using satellite AIS to improve our understanding of shipping and fill gaps in ocean observation data to support marine spatial planning. Journal of Applied Ecology, 55: 1834–1845.
Morgan, A., and Carlson, J. K. 2010. Capture time, size and hooking mortality of bottom longline-caught sharks. Fisheries Research, 101: 32–37.
Natale, F., Gibin, M., Alessandrini, A., Vespe, M., and Paulrud, A. 2015. Mapping fishing effort through AIS data. PLoS One, 10: e0130746.
Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., and Torres, F. 1998. Fishing down marine food webs. Science, 279: 860–863.
Peterson, C. D., Gartland, J., and Latour, R. J. 2017. Novel use of hook timers to quantify changing catchability over soak time in longline surveys. Fisheries Research, 194: 99–111.
Poos, J. J., Turenhout, M. N. J., van Oostenbrugge, H.A. E., and Rijnsdorp, A. D. 2013. Adaptive response of beam trawl fishers to rising fuel cost. ICES Journal of Marine Science, 70: 675–684.
Russo, T.,Parisi, A., and Cataudella, S. 2011. New insights in interpolating fishing tracks from VMS data for different métiers. Fisheries Research, 108: 184–194.
Russo, T., Carpentieri, P., Fiorentino, F., Arneri, E., Scardi, M., Cioffi, A., and Cataudella, S. 2016a. Modeling landings profiles of fishing vessels: an application of self-organizing maps to VMS and logbook data. Fisheries Research, 181: 34–47.
Russo, T., D'Andrea, L., Parisi, A., and Cataudella, S. 2014. VMSbase: an R-Package for VMS and logbook data management and analysis in fisheries ecology. PLoS One, 9: e100195.
Russo, T., D'Andrea, L., Parisi, A., Martinelli, M., Belardinelli, A., Boccoli, F., Cignini, I. *et al.* 2016b. Assessing the fishing footprint using data integrated from different tracking devices: issues and opportunities. Ecological Indicators, 69: 818–827.
Russo, T., Franceschini, S., D'Andrea, L., Scardi, M., Parisi, A., and Cataudella, S. 2019. Predicting fishing footprint of trawlers from environmental and fleet data: an application of artificial neural networks. Frontiers in Marine Science, 6: 670.
Sampson, D. B. 2011. The accuracy of self-reported fisheries data: Oregon trawl logbook fishing locations and retained catches. Fisheries Research, 112: 59–76.
Shepherd, J. G. 2003. Fishing effort control: could it work under the common fisheries policy? Fisheries Research, 63: 149–153.
Shepperson, J. L., Hintzen, N. T., Szostek, C. L., Bell, E., Murray, L. G., and Kaiser, M. J. 2018. A comparison of VMS and AIS data: the effect of data coverage and vessel position recording frequency on estimates of fishing footprints. ICES Journal of Marine Science, 75: 988–998.
Swartz, W., Sala, E., Tracey, S., Watson, R., and Pauly, D. 2010. The spatial expansion and ecological footprint of fisheries (1950 to present). PLoS One, 5: e15143.
Szynaka, M. J., Erzini, K., Gonçalves, J. M. S., and Campos, A. 2021. Identifying métiers using landings profiles: an octopus-driven multigear coastal fleet. Journal of Marine Science and Engineering, 9: 1022.
Szynaka, M. J., Fernandes, M., Anjos, M., Erzini, K., Gonçalves, J. M. S., and Campos, A. 2022. Fishers, let us talk: validating Métiers in a multi-fear coastal fishing fleet. Fishes, 7: 174.
Tremblay, Y., Shaffer, S. A., Fowler, S. L., Kuhn, C. E., McDonald, B. I., Weise, M. J., Bost, C.-A. *et al.* 2006. Interpolation of animal tracking data in a fluid environment. Journal of Experimental Biology, 209: 128–140.
Vermard, Y., Rivot, E., Mahévas, S., Marchal, P., and Gascuel, D. 2010. Identifying fishing trip behaviour and estimating fishing effort from VMS data using Bayesian hidden Markov models. Ecological Modelling, 221: 1757–1769.
Vespe, M., Gibin, M., Alessandrini, A., Natale, F., Mazzarella, F., and Osio, G. C. 2016. Mapping EU fishing activities using ship tracking data. Journal of Maps, 12: 520–525.
Ward, P., Myers, R. A., and Blanchard, W. 2004. Fish lost at sea: the effect of soak time on pelagic longline catches. Fishery Bulletin, 102: 179–195.
Whoriskey, K., Auger-Méthé, M., Albertsen, C. M., Whoriskey, F. G., Binder, T. R., Krueger, C. C., and Flemming, J. M. 2017. A hidden Markov movement model for rapidly identifying behavioral states from animal tracks. Ecology and Evolution, 7: 2112–2121.
Witt, M. J., and Godley, B. J. 2007. A step towards seascape scale conservation: using vessel monitoring systems (VMS) to map fishing activity. PLoS One, 2: e1111.

Handling editor: Jan Jaap Poos

Received: 13 January 2023; **Revised:** 16 May 2023; **Accepted:** 23 May 2023