WORKSHOP ON GEO-SPATIAL DATA FOR SMALL-SCALE FISHERIES (WKSSFGEO)

VOLUME 4 | ISSUE 10

ICES SCIENTIFIC REPORTS

RAPPORTS
SCIENTIFIQUES DU CIEM

ICES INTERNATIONAL COUNCIL FOR THE EXPLORATION OF THE SEA
CIEM CONSEIL INTERNATIONAL POUR L'EXPLORATION DE LA MER

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H.C. Andersens Boulevard 44-46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

ISSN number: 2618-1371

ICES Scientific Reports

Volume 4 | Issue 10

WORKSHOP ON GEO-SPATIAL DATA FOR SMALL-SCALE FISHERIES (WKSS-FGEO)

Recommended format for purpose of citation:

ICES. 2022. Workshop on Geo-Spatial Data for Small-Scale Fisheries (WKSSFGEO).
ICES Scientific Reports. 4:10. 60 pp. http://doi.org/10.17895/ices.pub.10032

Editors

Josefine Egekvist • Marta Mega Rufino

Authors

Luís Manuel Correia Ferreira Bentes • Patricia Breen • Neil Campbell • Daniel Cano • Andre Carvalho • François Danhiez • Dan Edwards • Josefine Egekvist • Hannah Fennell • Tereza Fonseca • Gianluca Franceschini • Alessandro Galdelli • Maria Mar Gil • Gildas Glemarec • Emily Hague • Cherry Harper Jones • Nuno Henrique Henriques • Hilmar Hinz • Einar Hjörleifsson • Helen Holah • Patrik Jonsson • Isidora Katara • Stefanos Kavadas • Michael Kriegl • Pedro Leitão • Irida Maina • Adriano Mancini • Guillermo Martin • Roi Martinez • Maria Mateo • Anna Mujal-Colilles • Jeppe Olsen • Sara Palma Pedraza • Jacopo Pulcinella • Antonio Punzón • Jose Rodriguez • Julien Rodriguez • Marta Mega Rufino • Lara Salvany • Torsten Schulze • Daniela Silva • Jonathan Stounberg • Anna Nora Tassetti • Oliver Tully • ClaudioVasapollo • Christian von Dorrien

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

Contents

i Executive summary

The aim of the workshop on Geo-Spatial Data for Small-Scale Fisheries (WKSSFGEO) was to discuss methods for working with high-resolution geo-spatial data in small-scale fisheries (SSF) as well as large-scale fisheries (LSF) with low duration of fishing events.

Several case studies using this kind of data, obtained from national data collection programmes or research projects were presented during the workshop. There was a large variety of approaches to estimate fishing effort using high resolution data being implemented by the different countries, most still under development, which led to important and fruitful discussions to start building a common framework. During the workshop, it became clear that a common terminology was essential and should be established, and therefore the group built a terminology table based on definitions found in legislation and reports. The main steps needed to estimate fishing effort were summarised, and for each step, the different approaches were listed, namely for: (1) data sources; (2) data pre-processing; (3) identifying fishing trips; (4) methods to infer fishing activity; (5) model validation and procedures and (6) fishing effort indicators. Both methods using statistical procedures and machine learning have been addressed. Finally, r scripts to do the procedures were assembled and stored in a GitHub repository. It was concluded that in the case of SSF and static gears, higher resolution tracking data (seconds-minutes) is required to properly estimate fishing effort, and that more work is needed to further develop and explore the methods to classify these type of data into fishing activities for different types of gears. The current workshop should be seen as a starting point. One of the objectives of the WKSSFGEO was to provide information, show the gaps and highlight the important aspects for the EU-legislation on the collection of this type of data (SSF tracking) which is currently being developed, which was fully addressed.

ii Expert group information

Expert group name	Workshop on Geo-Spatial Data for Small-Scale Fisheries (WKSSFGEO)
Expert group cycle	Workshop
Chairs	Marta Mega Rufino, Portugal
	Josefine Egekvist, Denmark
Meeting venue and dates	29 November – 3 December 2021 , Lisbon, Portugal, online meeting (48 participants)

1 Terms of References

ToR a: Discuss and apply methods for identifying trips/hauls in small-scale fisheries, including passive gears, using high-resolution geo-spatial data. Participants need to bring their own data for case-studies to develop best practices and common methodologies

ToR b: Based on the best practices identified, develop an R-script that can be used as a template for analysis of geo-spatial for small-scale fisheries

ToR c: Evaluate how the use of high-resolution geo-spatial data improve effort estimates and can help quantify the extent of small-scale fisheries.

2 Introduction

WKSSFGEO was conducted as a hybrid meeting with participants physically present at IPMA, Lisbon, and participants present online via Teams. During the first day, several groups presented different case studies across the EU on implementation of high-resolution geo-spatial data. During the second day, four subgroups discussed the different pre-processing steps when working with this kind of data, namely how to define trips and infer fishing activities. In the third day of the workshop, four groups of participants worked on different subjects, namely one subgroup merged the information from the second day, one discussed validation of methods, another discussed effort variables (fishing indicators) and the last one discussed R-code for pre-processing of data and trip definition. It was clear that definitions of e.g. fishing days, fishing activity etc. were ambiguous and existing definitions not always applicable in the context of small-scale fisheries, high-resolution geo-spatial data and passive gears, so an additional subgroup collected and discussed the existing definitions from legislation and previous reports. Each day the meeting ended with a round table where each subgroup presented its conclusions, which were then discussed with all participants. On the last day of the workshop, a final round table of discussion was done, on the conclusions and recommendations for future work, where it was clear that another workshop was essential to expand on the findings reached during the present workshop.

2.1 Social media coverage

ICES
Twitter: *https://twitter.com/ICES_ASC/status/1466701270996168706?s=20*
Facebook: *https://www.facebook.com/ICES.Marine/posts/F4731435250234522*

The hybrid Workshop on Geo-Spatial Data for Small-Scale Fisheries (WKSSFGEO) took place at IPMA in Lisbon, Portugal, and remotely from 29 November to 3 December 2021.

Forty nine experts from 13 countries, representing 28 institutions travelled to IPMA in Lisbon, virtually and physically, with their suitcases full of fishing vessels tracks and data science knowledge to discuss and jointly develop a common framework to estimate fishing effort using high resolution data. This is essential for small-scale fisheries that represent 80% of the EU fleet and have a fundamental cultural heritage and human importance. The fishing effort information obtainable using this ***framework*** *is essential for biodiversity and conservation, marine spatial planning and fisheries management.*
Find more about WKSSFGEO: https://www.ices.dk/community/groups/Pages/WKSSFGEO.aspx

2.2 Policy context of small-scale fisheries

Small-scale fisheries (SSF) have different forms and modes of operation. They include the catching of fish, post-harvest treatment and marketing of the catches, as well as ancillary trades. These fisheries are found in coastal marine areas, brackish water lagoons, and along freshwater lakes, rivers, and reservoirs and are not homogenous within and across countries and regions. In the EU, SSF play a crucial role as they represent nearly 75% of all fishing vessels registered in the EU (EMFAF EU 2021/1139, EC (2022) Database - Eurostat. Available at: https://ec.europa.eu/eurostat/data/database). Small-scale fisheries are especially important in the Mediterranean and southern European countries, where over half of the sector is concentrated and where it has been playing a vital role in the local economy of coastal communities for centuries.

The Common Fisheries Policy (CFP) contemplates the specificities of small-scale coastal fisheries in a number of provisions. For example, the small-scale fleet is **exempted from certain obligations** that apply to larger vessels, such as fishing authorisations, landing declarations, sales notes, and separate stowage.

To develop appropriate management measures for small-scale fisheries, it is important to have accurate and sufficient information on their scope, stakeholders, operations and impacts.

As part of the revision of the regulation on fisheries control, the Commission proposed to modernise the EU rules governing fisheries data and to monitor small-scale vessels and provide an opportunity for small-scale fishers to become fully involved in the long-term management of the fish stocks (EU 2018). This includes an obligation for all vessels, namely those below 12 m overall length to report their catches electronically and have a tracking system, but with flexibility with regard to the specification of the vessels tracking system.

In addition, the Regulation (EU) 2021/1139 that establishes the European Maritime, Fisheries and Aquaculture Fund (EMFAF, EU 2021) lays down the priorities of the EMFAF, which includes enabling a sustainable blue economy in coastal, island, and inland areas.

2.3 Background of the workshop

The background of arranging this workshop is the absence of a common framework to analyse the data on the small-scale fisheries and passive gears spatio-temporal dynamics, namely in terms of fishing effort. According to EU 1224/2009, VMS data are mandatory for vessels larger than 12 metres and with a maximum interval between positions of two hours, which is often not adequate to map the activity of SSF, for which a fishing trip can often be shorter than 2 hours. ICES has an annual data call for VMS/Logbook data, requesting the spatial information for vessels with VMS for the period 2009–2020, however the smaller vessels without VMS are missing. These data are used for ICES advice requests to EU, OSPAR, and HELCOM on e.g. fishing abrasion (Swept Area Ratio) where the small-scale fishery is missing. The ICES WGSFD (Working Group on Spatial Fisheries Data) has a ToR to "Develop spatial effort indicators for static gears", but a need for a focused workshop on the use of high-resolution data in small-scale fisheries was identified during the WG meeting in 2021. In relation to ICES WGBYC (Working Group on Bycatch of Protected Species) data calls, fishing effort indicators relevant to bycatches of Protected, Endangered, and Threatened Species (PETS) are often not available at a fine-scale for passive gears, e.g. net length (for gillnets), number of hooks (for longlines), soak time, or the exact location of the gears. This might be improved if high-resolution spatial data and analysis methods are more widely available. The ICES WGCATCH (Working Group on Commercial Catches) has a ToR on small-scale fisheries to evaluate the use of geo-spatial data (e.g. GPS, AIS) to improve effort estimates and produce guidelines on how to calculate effort in the SSF.

As there is currently no EU legislation that makes it mandatory to track the positions of fishing vessels smaller than 12 m, there is no mandate to request these data in a data call. Yet, some national initiatives have been implemented as case studies in scientific projects and following local legislations on specific parts of the fleets. For instance, the entire Portuguese bivalve dredge fleet, which is mostly SSF, has been tracked since 2016. In the EU-MAP (EU 2021/1167), section 3.1, it is mentioned that effort variables should be reported for the whole fleet, and where there is no obligation under the control regulation (EU 1224/2009), alternative sampling methods shall be applied. Therefore, it was found highly relevant to discuss and share the methods and scripts for analysis of high-resolution geo-spatial data, the development of methods to calculate the fishing effort based on these data and begin to produce a common framework to be applied in high resolution data for taking into account the recent EU legislation (see the policy section on this document).

2.4 Recommendations for hybrid meetings

Hybrid meetings are starting to take place as a result of the COVID pandemic to allow the advantages of both meeting physically and also not restricting participation of the colleagues that are unable to travel to the meeting. Here we provide a list of recommendations for future hybrid meetings:

- A good internet connection is essential, preferably both Wi-Fi and cable network should be available (one network connection for 2–3 persons).
- All participants should identify themselves before speaking and should be equipped with headsets with microphones, or alternatively one or more hand-held wireless microphone could be used for the entire meeting room.
- Participants present physically in the room should log on to Teams (or the alternative platform used for the meeting) and follow the same rules as the online participants, e.g. keep their personal microphone off when not speaking, request to speak using the raise hand button and speak only when given the floor, and have their cameras turned on at all times.
- One person present physically should be assigned the responsibility to manage the online meeting, chats, speaking requests from participants, etc…
- Discussions between online and physically present participants worked better in subgroups than in plenary. Subgroups should be composed of mixed 5–6 participants (physical and remote) and each subgroup should be located in a different room with internet connection (i.e. 4 groups = 4 subgroup rooms).
- Clear communication of a schedule from the chair to both physically present and online participants is required, e.g. on time to meet for plenary.
- Limitations were found in the use of Sharepoint where it was not possible to work simultaneously in documents (as in google docs or in Teams). It is desirable to have a system where it is possible to work simultaneously/collaboratively, where chats are permitted, where review/changes are clearly identifiable, with the facility to make mind maps and with MS office compatibility. MS Teams could work both as a platform for online meetings, to document chat, and enable working simultaneously in documents, however some participants had problems with accessing the full functionality of Teams (i.e. chat was not visible, or it was not possible to share screen), while it worked well for other participants.
- The system (whatever is used) should be tested and introduced, preferably the week before the meeting to all participants that are interested to be sure it is fully functional when the WK starts.

2.5 Compilation of participants data: cloud, map, summary table

Below is an overview of the participants and the data they had available for the workshop. The workshop included participants from nine countries (Figure 2.1), bringing data principally from pots, gillnets, longline, and dredge fisheries (Figure 2.2).

- 49 (15 present; 33 online)
- 28 institutions
- 20 research projects
- 16 gear types
- SSF 33 / Large Scale Fisheries (LSF) 14
- Temporal resolution:
    - 1 min / 1 sec (mostly)
    - Maximum of 20 minutes

Source of data	Number of data sets
AIS	17
GPS	13
VMS	13

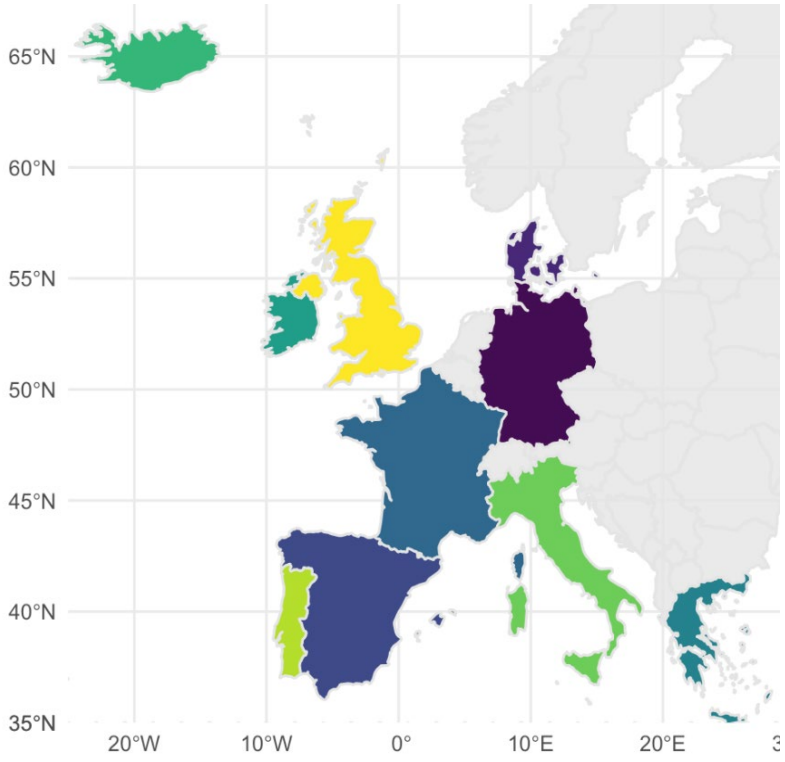


Figure 2.1. Map of WKSSFGEO participants countries.

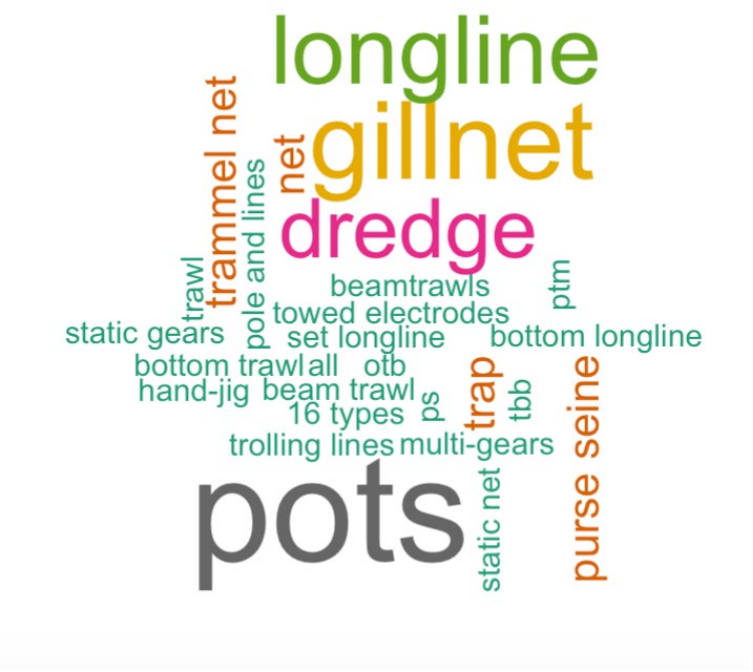


Figure 2.2. Word cloud of WKSSFGEO participants fishing gears corresponding of the data.

2.6 Presentations abstracts

Different teams from several countries presented different projects, issues with geospatial data and the approaches they use to analyse these data. These presentations are available on the WKSSFGEO SharePoint site.

Julien Rodriguez (Ifremer, France) and Francois Danhiez (Capgemini, France). RECOPESCA program and fishing gears identification from artificial intelligence applied to geospatial data.

Faced with a lack of data to accurately assess the spatial distribution of catches and fishing effort, and for the environmental characterization of the fishing area, Ifremer has implemented since 2005 a project called « Recopesca ». It is based on a sample of voluntary fishing vessels equipped with sensors coupled with onboard GPS monitors recording positions at least every fifteen minutes. The data gathered are used as a benchmark dataset for various research programs regarding the use of geolocation information to monitor fisheries.

As an example, a study investigated the use of machine learning to identify fishing gears from geospatial data. The data set gathered is composed of 25 612 trips recorded on 107 vessels from 2006 to 2020. It includes 16 gear types aggregated in 10 gear classes including steaming “SAIL” to be able to identify trips without fishing events. Five different algorithms (KNN, C5.0, SVM, RF and XGBoost) were tested on the basis of 96 motion, speed and time related variables computed on every trip. After the training of these algorithms on 70% of the data set, the algorithms precision was validated on the remaining 30 % of the data set. The model comparison shows that XGBoost achieved the classification of the 10 fishing gear classes with the highest accuracy (94.89 %). After choosing XGBoost as the algorithm to predict the gear classes, we conducted an opti-

mization leading to the building of a final algorithm exploiting only 34 of the 96 features to classify the fishing gears with the same performance. Thus, this work shows our ability to retrieve the fishing gear type or steaming using machine learning on GPS data collected every fifteen minutes and its applicability on VMS and AIS data need to be investigated in further works.

Oliver Tully, Guillermo Martin, Patricia Breen and Sara Palma Pedraza (Marine Institute, Ireland). iVMS and effort estimation for the Inshort Fleet in Ireland

Vessel monitoring systems (iVMS) have been deployed on some vessels under 12m in Ireland since 2014. It is currently mandatory for all vessels fishing hydraulic dredges for razor clams to have a functioning iVMS system onboard and to report GPS position at frequencies of 1-10min. There are about 70 such vessels. In the period 2017-2020 an additional 50 potters had iVMS installed voluntarily. From 2021 over 20 netters/potters have iVMS installed. iVMS systems have been procured from 5 different suppliers over the period 2014–2020 with an additional two companies supplying test units. There is a trade-off between initial capital cost, depreciation and swap out rates. Due to various power supply issues on vessels under 12m iVMS units that are autonomous of vessel power are now preferred and are a condition of procurement of new systems. Some autonomous units generating power from integrated solar panels do not maintain internal battery power and stop reporting. Two other units (one used in 2017–2020 on 25 potters, 1 on trial) have shown they can maintain reporting over winter on the west coast of Ireland. Conservation of battery power on solar powered units can be significantly improved by adjusting the ping frequency when vessels are in port using geofences (buffers around port). Various gear sensors have been trialled to obtain data on fishing effort. This has been discontinued. Any future use of gear sensors will be for validation of models designed to identify the location of fishing events. This will also be done by observers. There is an additional power demand if the iVMS unit needs to listen for the presence or absence of Bluetooth gear sensors that needs to be considered when using autonomous units.

Data are hosted on suppliers databases and accessed by the Marine Institute (fleet) and Skippers (own vessel) through interfaces provided by the supplier. Data are downloaded or otherwise accessed from the suppliers and maintained in a SQL database in the Marine Institute. Modelling of high frequency ping data drawn from the database, using Hidden Markov Models (HMM), has shown the potential to identify fishing events, transit between fishing events and steaming using date/time, speed and bearing data. More validation data are needed to quantify the proportion of pings correctly classified using such methods. An R Shiny application has been developed to show how high frequency iVMS data can be used to monitor fleet effort. The application uses the HMM classified pings and additional information on gear spacing to report metrics such as pots hauled, total length of fishing gear hauled or total length of the fishing track grouped by vessel, fleet, location and time.

Miguel B. Gaspar, André N. Carvalho and Marta Rufino (IPMA, Portugal). Real time tracking of Portuguese SSF fisheries

Small-scale fisheries (SSF) have a strong representation in Portugal, comprising 86% of the Portuguese fishing fleet (around 2500 active vessels). Most SSF vessels present an overall length < 9 m (71% of all vessels in 2019), equipped with outboard engines, and operating mainly within 3 nautical miles from the coastline. SSF have a high economic importance representing about 36% of landings in Euros (corresponding to 22% of tons landed). These fisheries use a broad combination of fishing gears and techniques (15 different main fishing gears in 2019), target a 274 different species (2019) with high commercial value and are responsible for most of the supply of fresh seafood to local markets (SSF landed on 61 fishing ports during 2019, whereas LSF in 33

ports). In addition, SSF promotes job creation (fishermen, traders, etc.), livelihoods and population settlement, and is part of an important heritage of culture and traditions. Despite their economic, environmental, social and cultural importance, it is broadly recognized that there is a lack of basic and high-quality information on SSF. The implementation of regular, broad coverage, efficient and cost-effective monitoring programs is hampered by the limited official data available, the high number of fishing vessels involved, the multiple fishing gears used, the multitude of caught fishing resources and the numerous landing sites. One of the major obstacles for long-term effective and responsive monitoring in SSF is the lack of spatial and temporal data.

In this context and since 2016, IPMA started the research project MONTEREAL (MAR2020), aiming to collect high spatio-temporal resolution data of the Portuguese bivalve dredge fleet, through real-time GPS tracking devices on installed on fishing vessels. This work permits to identify the main bivalve fishing grounds, assess the spatio-temporal distribution of the fishing effort, and contribute to the sustainable management of the fishery and maritime spatial planning. Currently 80 bivalve dredge vessels are equipped with GPS trackers to record and transmit the spatio-temporal position every 30 secs. The use of this equipment is already mandatory for this fleet. On a voluntary and experimental basis, the use of these GPS trackers is being extended to other SSF fleets, and they were installed in 10 trammel net vessels since 2020 and in 37 octopus pots and traps vessels since 2021. Additionally, in 2021 IPMA launched a mobile phone app "PescApanha" aiming to collect geolocated data and complement the information on the main species caught (e.g. weight). This app is based on the voluntary and responsible participation of fishermen and shellfish harvesters.

At present, within MONTEREAL, IPMA is developing a software that allows storing and processing all the daily received data and performing a dynamic analysis of the fishing effort. This is the most challenging phase of the project due to the need of automatically analyse millions of data. This involves developing algorithms to identify the fishing trips and fishing events, calculate the fishing effort for each fishing gear, pair the fishing days recorded with the GPS trackers with the official information on landings and auction sales, and finally plot the maps. The future challenge will be to identify the fishing trips with multi-gear events and assess how to use this data for the fisheries management.

Tania Mendo and Anna Mujal (St. Andrews, Scotland). Developing a nation-wide monitoring system for Small-Scale Fisheries: Experiences from Scotland

Increased positional data of fishing vessels have allowed great progress in identification of fishing grounds, effort, and impact to habitats. Experiences with VMS and AIS systems fitted on large scale fishing vessels have shown that in order to identify when and where fishing activities are occurring, fisheries-specific approaches need to be developed based on the fishing practices of each fleet. This presentation will guide you through our experiences in developing a nation-wide monitoring system for SSF in Scotland. We first explored the use of AIS for Small-Scale fisheries and then explored other cheaper tracking devices. We assess different methods and approaches to infer when hauling is occurring and estimate the number of creels deployed from geopositional data. The outputs of this system are currently being trialled in the Outer Hebrides, and a user friendly portal has been set-up to allow access to fishers, managers and researchers.

Josefine Egekvist, Jeppe Olsen, Gildas Glemarec, Jonathan Stounberg (DTU Aqua, Denmark). Integrating AI/VMS/BlackBox data and validation with EM and observer data

In Denmark, positional data are collected via different sources: VMS is mandatory for vessels larger than 12 m, AIS, which is mandatory for fishing vessels larger than 15 m but increasingly used by smaller vessels for security purposes, and an electronic monitoring (EM) system called Blackbox that is mandatory for mussel dredgers in certain areas and that equip a sample of the gillnet fleet.

With the combined data sources, VMS, AIS and BlackBox, the data coverage in commercial fisheries in Denmark varies between gears. In the period 2015–2020, the coverage for beam trawlers, bottom trawlers and Danish seines have been above 91%, for mussel dredgers, it has increased during the period from 76% to 95%. For nets the coverage is between 56 and 62%, for lines it has increased from 9% to 38% and for pots and traps, the coverage is low (3–20%). The coverage is calculated using the value of landings for those fishing trips with known positions versus the fishing trips where no positions have been found.

The AIS data are downloaded from the Danish Maritime Authority, but the vessel-id that is reported in logbooks are not directly available in the data. Therefore, the vessel-id is derived from MMSI/Call sign using databases like the fleet register. Alternatively, the vessel-id can sometimes be extracted from the vessel name that is available in the AIS data.

The VMS, AIS and BlackBox data are combined into one dataset (time zones can sometimes be tricky), and where there are gaps, an interpolation is applied (spline using position and heading), so that the resulting dataset has one position per minute. Fishing trips are defined based on harbour polygons. The data are then merged with logbooks/sales notes. Where logbooks are not available, sales notes are reported, and an algorithm developed with the RCG ISSG on Métier issues is applied to estimate the métier, including the gear code. A speed filter is applied to classify fishing or other activity, data are cleaned (points on land, erratic positions), and positions classified as fishing are turned into line features. For mobile bottom contacting gears, polygons are made by applying the width of the gear to estimate the swept area.

Issues with the data include false positives (e.g., points classified as fishing that are not) and false negatives (e.g., points that should be classified as fishing, but are not). This is caused by the vessel slowing down when not fishing, or fishing at unusual speeds.

An electronic monitoring program with video started up in Denmark in 2010, with the aim of collecting fine-scale effort and PETS bycatch data in commercial gillnet fisheries. The overarching goals are to develop methods to estimate PETS bycatch rates and extrapolate total PETS bycatch accurately, and ultimately to map PETS bycatch high-risk areas at species-level.

The frequency of the EM data is 10 seconds (location, speed, and course), and about 90% of the EM data have been analysed, so positions of fishing gears and bycatch events are known. The resolution of the data is high, but from a limited number of vessels (18 vessels since 2010 – 9 vessels in 2021). Extrapolation to the entire fleet is difficult, as logbook data have low resolution, but AIS/VMS data could be used to obtain better estimates of spatiotemporal distribution and intensity of fishing effort. The electronic monitoring can be used to validate fishing effort predictive models from AIS data.

In addition, data from the Data Collection Framework (DCF) at-sea observer programme can be used for validation of fishing activities, as observers register time and position of setting and hauling the gear. The majority of the observer trips are from bottom trawl fisheries, meaning that this type of data can't be used for validation of Danish gillnet fisheries.

After the workshop, the full VMS/AIS/BlackBox data set will be analysed for the Western Baltic testing the methods discussed during the workshop to analyse if there are any changes in fisheries behaviour that can be related to variations in oxygen levels, that have been mapped spatially by quarter back to 1990 in the EMFF project HypCatch.

Hilmar Hinz and Maria del Mar Gil (Imedea, Spain). Status of vessel monitoring systems of the artisanal fleet in the Balearic Island, Spain

Thus far the activities of artisanal fisheries in the Balearic Island have not been monitored despite its large fleet size (approx. 300 boats) and local economic importance. The artisanal fleet principally uses static fishing gears targeting different species in a seasonal pattern. In 2019, the local government equipped 15 boats with so called "green boxes" that are based on mobile phone technology for the geo-location of vessels. The pilot project served to test the feasibility of equipping the whole artisanal fleet with these systems. After the successful trial period green boxes are now being installed on all artisanal boats registered in the Balearic Islands. The completion of this process is being anticipated for 2023. Researchers from the IMEDEA (UIB-CSIC) are currently tasked to advise the government on the most useful configurations of these systems to aid later identification of fishing events, to be able to enforce fisheries regulations, to estimate fishing effort, as well as to monitor stocks and environmental impacts. The analysis of the fleet movement data will be explored within a dedicated project in the near future and standards developed within the WKSSFGEO workshop will be considered.

Anna Nora Tassetti, Alessandro Galdelli, Jacopo Pulcinella, Adriano Mancini, Luca Bolognini(Italy). ARGOS project. A low cost tracking solution for Italian small-scale fishery

During the last decade accurate spatial and quantitative information of industrial fisheries have been increasingly obtained using tracking technologies and machine learning analytical algorithms. However, in most small-scale fisheries, lack of spatial data has been a recurrent bottleneck as VMS and AIS, developed for vessels longer than 12 and 15 m in length respectively, have little applicability in these contexts. It follows that small-scale vessels (< 12 m in length and often without dedicated electrical systems) remain untracked and largely unregulated, even though they account for most of the fishing fleet in operation in the Mediterranean Sea. As such, the tracking of small-scale fleets tends to require the use of novel and low cost solutions.

Within the ARGOS project (Interreg V-A Italy-Croatia CBC Programme 2014–2020, Strategic calls for proposal), a scalable architecture is proposed, making use of a low-cost LoRaWAN/cellular network to acquire and process positioning data from small-scale vessels. The architecture relies on Traccar, while an efficient monitoring device is proposed using the high-tech and cost-efficient Teltonika FMM640. A hall-effect speed sensor attached to the hauler is used in tandem to record when and where the SSF vessel is in operation.

Preliminary results of a first installation of the prototype were presented, as well as the data collected and the algorithm developed to define their individual fishing trips. In the framework of the ARGOS project, we expect to monitor the movements of a sample of around 25/30 boats for 3 years, exerting their activity as widely as possible along the Marche Region (Italy).

The emergence of such a low-cost and open source technology coupled with artificial intelligence could open new opportunities for equipping small-scale vessels, collecting their trajectory data and estimating their fishing effort (information which has historically not been present).

Stefanos Kavadas, Irida Maina (HCMR, Greece). Greek VMS/AIS data and estimating fishing pressure in SSF

Small-scale fisheries (SSF) comprise 95% of the entire Greek fishing fleet (in total 13,241 vessels), while only 3.5% of these vessels are equipped with Vessel Monitoring System. An overview of the available datasets and methods used for analyzing VMS data for Greek SSF was presented to the group. The methods include: i) data quality control to remove common errors in VMS data, ii) identification of vessels with the same fisheries strategy (i.e. gillnets, trammel nets, bottom long lines), iii) characterization of vessels activity (such as fishing, steaming or mooring) based on speed thresholds and integration of fisheries legislation and iv) estimation of fishing effort. Additionally, given that data from monitoring devices (e.g. VMS/AIS) are not available (or complete) for vessels with length overall <12m, a Multi-Criteria Decision Analysis (MCDA, Kavadas *et al.*, 2015) used to estimate a fishing pressure index for SSF in data limited cases, was presented. The MCDA combines environmental/anthropogenic data such as bathymetry, distance from coast, meteorological conditions, fishing fleet characteristics and expert knowledge, with information from Data Collection Framework for estimating fishing effort for SSF in several spatiotemporal scales. Given that certain techniques (e.g. bottom longlines, static nets) might have impacts on the seabed (maerl beds, coralligenous formations etc.), assessing the actual pressure of SSF in Greece is quite important and needs to be further investigated. The above methods are planned to be used in the projects Med&BS RDBFIS ("Development of the regional database for the Mediterranean and Black Seas") and SEAwise ("Shaping ecosystem based fisheries management") for assessing the extent of the SSF fleet.

3 Terminology used in the report

During the workshop, it became clear that many of the discussions revolved around definitions and the need for a common terminology. Therefore, Table 3.1 below lists the terminology used in this report, and where possible, referring to definitions in legislation. A future task could be to develop this glossary further with infographics for a quick idea about the concepts through illustrations.

Table 3.1. Terminology used in this report.

Term	Definition	Discussion
Small Scale Fisheries (SSF)	The EU defines SSF as follows: "Small-scale coastal fishing is carried out by marine and inland fishing vessels of an overall length of less than 12 metres and not using towed fishing gear, and by fishers on foot, including shellfish gatherers." (EMFAF, EU 2021)	Stobberup *et al.*(2018) stated in a report to the PECH Committee of the European Parliament that this definition is too restrictive, as it does not take into account the specificities of fishing in the various Member States (MS).
High-resolution geo-spatial data	In this report it is defined as data with positions recorded with a frequency of maximum 15 minutes (i.e. secs to mins intervals). Examples of tracking devices are: AIS, VMS, iVMS (inshore VMS), Electronic Monitoring (such as BlackBox) and Greenbox	
Active gears	Same as mobile gears. In active gears, the target species are captured based on an aimed chase (e.g. trawls, dredges) (Bjordal 2009).	
Passive gears	Same as static gears. In passive gears, the capture of the target species is based on movement of the species towards the gear (Bjordal 2009) (e.g traps, gillnets, hook and line).	
Days at sea	The calculations of Days at Sea are based on the definition given in the EU-MAP Commission Decision (2021/1167/EU): "Any continuous period of 24 hours (or part thereof) during which a vessel is present within a defined fishing area and absent from port".	In the case of geo-spatial data, the time is known, and the days at sea can be calculated if the trip is defined. It was discussed if the Days at sea should only be counted as fishing effort if fishing activity is found within the trip. Further, SSF may take only 2 h and count as a day at sea, unlike LSF where they could be 24h fishing.
Fishing Days	The definition of fishing days has been discussed in several reports and is defined in the legislation. Fishing Days are based on dates recorded in logbooks. Another way to say this is considering they are based on 'calendar days'. The number of fishing dates registered	If fishing activity from a vessel is found in the geo-spatial data, the date should be recorded as a fishing day. However, the above comment also applies in this case to SSF. Further, many SSF that use vessels smaller than 10

	in logbooks can exceed the number of 24 hour periods in a trip. EU-MAP Commission Decision (2021/1167/EU): Fishing day: any calendar day at sea in which a fishing activity takes place. One fishing trip can contribute to both the sum of the fishing days for passive gears and the sum of the fishing days for active gears used on that trip. The 2016 Report on the 2nd Workshop on Transversal Variables workshop reviewed the definition of "fishing day" proposed by STECF-13-12 to be included in the DCF: *"Any day at sea with fishing operation. In case of passive gears, each day of a remained operational gear counts as fishing day and is associated to the fishing trip during which the gear was deployed."* *Since the Fishing time (which is equal to soaking time for passive gears) is currently not a mandatory field in logbooks, there may not be any information on whether gears remain at sea or not; this means that the definition of a "Fishing day" referring to passive gears cannot be followed in practice. The WK suggests that, the definition of "Fishing day" should be changed to:* ***Fishing day** - "Any day at sea with a fishing operation".*	metres length are not obliged to fill out logbooks.
Fishing trip	**(EU) 404/2011**: "any voyage of a fishing vessel during which fishing activities are conducted that starts at the moment when the fishing vessel leaves a port and ends on arrival in port."	If a vessel does several trips within a day, according to recommendations from the 2016 Report on the 2nd Workshop on Transversal Variables workshop, each trip is counted, and the fishing days are counted for each trip.
Vessel activity	The vessel is active, away from port/mooring.	This would include Days at Sea which are not Fishing Days.
Fishing activity	According to Control Regulation (EU 1224/2009): 'fishing activity' means searching for fish, shooting, setting, towing, hauling of a fishing gear, taking catch on board, transhipping, retaining on board, processing on board, transferring, caging, fattening and landing of fish and fisheries products.	

Fishing event	**The 2016 Report on the 2nd Workshop on Transversal Variables workshop** mentions fishing event as haul or passive gear use.	In the 2016 Report on the 2nd Workshop on Transversal Variables workshop, the fishing events are related to what is recorded in the logbooks. There might be fishing related events, e.g. retrieval of damaged gear with no catch or retrieval of stored/inactive passive gears that are not considered fishing events in logbook terms, but would be classified as fishing activity using high-resolution geo-spatial data only.
Fishing operation	All activities in connection with searching for fish, the shooting, towing and hauling of active gears, setting, soaking, removing or resetting of passive gears and the removal of any catch from the gear, keep nets, or from a transport cage to fattening and farming cages.	A fishing operation may include no fishing events!
Haul	Refers to the retrieval of a fishing gear, end of a fishing event	
Set/deploy	Refers to the placement of a fishing gear, beginning of a fishing event. The term "set" also refers to a collection of individual passive gears joined into a single string or fleet.	
Fishing time (hours)	(EU) 404/2011: Total time spent searching (e.g. using sonar) or fishing and equals the number of hours spent at sea minus the time spent in transit to, between and returning from the fishing grounds, dodging, inactive or waiting for repair. Use of haul by haul data can have advantages in calculating Fishing Days compared to a daily logbook entry.	The inclusion of searching time makes this definition problematic. It is similar to fishing operation, but apparently minus transit/inactive time. The effort variable is requested in the EU-MAP **2021/1167,** but regarding static gears, it is unclear if it refers to the vessel fishing time or the gear soaking time.
Fishing gear	International Convention for the Prevention of Pollution from Ships (MARPOL), Annex V: A fishing gear is any physical device or part thereof, or combination of items that may be placed on or in the water or on the seabed with the intended purpose of capturing or controlling for subsequent capture or harvesting marine or freshwater organisms.	

	Fishing gears are defined in the EU Master Data register. The link between fishing gears and métiers are defined in **EU-MAP (EU 2021/1167)** table 5, where the fishing gear is level 4.	
Soaking time	EU 2010/93: time calculated from the point where each individual unit of a passive gear has been set, to the time when the same unit starts to be removed. Average soak time varies among fisheries and is dependent on factors such as the target species.	When does soak time start or end? In some cases, it can take a long time to haul the gear, so does it start when any part of the gear is under the water and finish when the entire gear is over the deck?
Fishing effort	EU 1380/2013: the product of the capacity and the activity of a fishing vessel; for a group of fishing vessels it is the sum of the fishing effort of all vessels in the group. The amount of fishing gear of a specific type used on the fishing grounds over a given unit of time for example hours trawled per day, number of hooks set per day or number of hauls of a beach seine per day. When two or more kinds of gear are used, the respective efforts must be adjusted to some standard type before being added (FAO, 1997). FAO term portal (https://www.fao.org/faoterm/en/?defaultCollId=21): The effort may be nominal, reflecting the simple total of effort units exerted on a stock in a given time period). It may also be standard or effective when corrected to take account of differences in fishing power and efficiency and ensure direct proportionality with fishing mortality). Relates usually to a specific fishery and gear. If more than one gear is considered, standardization in relation to one of them is necessary. For biologists, a good measure of fishing effort should be proportional to fishing mortality. For economists it should be proportional to the cost of fishing.	
Fishing vessel	EU 1224/2009: any vessel equipped for commercial exploitation of living aquatic resources.	
Capacity	EU-MAP (EU 2021/1167) table 6: Number of vessels and vessel characteristics: GT (Gross Tonnage), kW (Kilowatt), Vessel age.	

Home port	Refers to the base port described below. Boat and gear activities are sampled from home ports or base ports, in contrast to catches and species composition, prices, etc. that are sampled at landing sites.	Note that some SSF boats may not leave from the port, but directly from a beach (e.g. in Scotland and Portugal).
Base port	The port from which fishing units operate, irrespective of where they are registered (home port). The differentiation between base ports and home ports occurs when fishing units migrate from the locations indicated by the frame survey to other sites, usually on a seasonal basis.	See comment above.
Landing sites	Locations at which boats land their catch. A landing site may be the same as the home port or base port but it can also be different. Boat and gear activities are sampled from home ports or base ports, in contrast to catches and species composition, prices, etc. that are sampled at landing sites.	
Point in harbour / boat parking areas	Area from which a vessel is berthed, moored, or launched.	
Observer data	Fisheries information collected onboard fishing vessels by independent observers.	
Electronic Monitoring (EM)	Van Helmond (2021): A typical electronic monitoring (EM) system consists of various activity sensors, GPS recording device and computer hardware which allow for detailed fishing effort registration without requiring additional onboard personnel, and (optionally) cameras for video monitoring of catches.	

4 Framework for working with high-resolution geo-spatial data

4.1 Data sources

Several data sources were identified and briefly discussed in the workshop (i.e. GPS trackers, AIS, iVMS). Further work is required towards the analysis and comparison of the different data sources, although see Quincoces (2021).

4.2 Preprocessing

Table 4.1 lists the recommended steps for the pre-processing of High Resolution Geo-Spatial Data stemming from various sources, but not accompanied by logbook data. Sources of these data include: iVMS, AIS, GPS, Electronic Monitoring, VMS, regardless of the frequency of the pings (although these would all be secs-mins). The steps are to be considered as suggestions and there might be exceptions, depending on the "metier". In some cases the order of the steps could have to be switched, while some steps may simply have to be ignored.

The Universal Transverse Mercator (UTM) map projection system can be used as long as the longitude/latitude (long/lat) distance between the points is less than 5º. Otherwise, haversine trigonometric functions should be used to compute distances between long/lat points. Generally speaking, use haversine for long/lat data and Euclidean for projected (UTM) data.

WARNING: When analysing data with R, try not to use the deprecated sp package, but use one of its more recent alternatives, like the sf package.

WARNING2: High resolution monitoring systems can produce large files that can be problematic when it comes to being treated "raw" with R. In these cases it is advisable, whenever possible, to carry out a prior filtering eliminating unnecessary variables, extracting the data from the area of interest or whatever is deemed appropriate by using other more suitable software, such as the awk, sed, cat, tr or similar tools, present in most UNIX and GNU/Linux operating systems. Note that on Windows, a Linux bash shell - the Windows Subsystem for Linux (WSL) - is available natively for Windows 10 version 2004 and higher (Build 19041 and higher) or Windows 11; older Windows versions can also use a Linux shell via an emulator or a virtual machine. In this way the size of the file can be significantly reduced, with subsequent processing with R being smoother and less CPU/memory demanding.

Table 4.1. Data pre-processing steps.

Issue	Method	Link to R function in GitHub	Discussion
Identify, flag, filter, faulty data			
UTF-encoding	Specifying working with UTF-8 format	Argument "encoding="UTF-8"" in read.table	Other languages besides EN (i.e., á,ú); note issues on importing from excel and system language/r-system language
Coordinates	• Careful with projections: lat-lon is usually WGS84 which is 4326, and UTM has to be in the zone of study (link to find the correct coding number for UTM zones) • C-squares method (*link*) • Use bathymetry (*link* for GEBCO) but if not available see NOAA, Copernicus) • Area of interest	Libraries: • sf (st_transform) • tidyverse • raster (stars/terra) Bathymetry c-squares. Inside vmsTools *https://github.com/ices-eg/WKSS-FGEO/blob/main/R-dev/ToICESGitHub/CustomizedProjectedCRS.R* creates a customized planar CRS that can be used for some calculations (e.g. distances) uses the deprecated PROJ4 format but could be adapted to PROJ6	• Wrong data • Outliers: impossible points, points out of your area of study. • Duplicates (see row below)
Time	• Time in R starts in 1970 (careful when importing time from Excel) • Check time-zone • Points in the future (check time range)	Need to create a template function • Library "lutz", miscellaneous functions for handling time (for example to convert GMT to local time)	
Speed	• Threshold • Percentile • Moving averages	• Be careful and compare transmitted speed (by the vessel) *vs.* calculated speed (e.g. $\Delta x/\Delta t$, using two subsequent positions & delta time) and use the most realistic one for this part • Speed threshold for AIS data • VMS data	
Duplicates	• Vessel, time_stamp, longs, lats • Buffer in time	Code for filtering Using vessel_id and time_stamp – 2 second buffer	To reduce vasts amounts of data from AIS namely, but also present in other systems
Points on land/in harbour			
Points on land	• Using shapefiles (*link*1, *link*2, *link3*, to a world public database of shapefiles)	• Points_on_land_polygon • sf::st_intersects / st_join /st_buffer	Allow for the inclusion of land polygons (e.g. Issues in vmstools with the position and

	defining land and filter out points on land • Use user-made polygon	• VMSTools	resolution of land polygons)
Points in harbour (boat parking areas)	• If vessels are harbour based, use polygon methods as in previous row • If vessels are not harbour based: a. Estimated 'harbours' like locations using lat/long frequency (hotspots) b. use a buffer c. use bathymetry	• VMS tools function: points_in_harbour has already a buffer in it (~1 NM) • *add_harbours function* • Bathymetry (fine -scale bathymetry) remove points in a water depth < D (D depends on the area) • RANN:nn2 for finding areas with high clustering of data points	-Issue: prevent false positives - Include beaches/launch sites? Problem: in some areas vessels are fishing very close to shore (10 metres). Expert knowledge needed -Further data /technical: Bathymetry sensors, any other device on collecting data/increase information (activity at position, N nets/pots, size of nets) -Area estimation approach: size of harbour area (buffer zone) estimated from the number of points (vessels) in a port
Downsampling and Interpolation (to a regular time series)			
Interpolation	• Linear interpolation between points • Use COG and speed with non-linear interpolation (VMS Tools) • *Kalman* filter • Other methods (AI)	Libraries: • stats::approx • VMS Tools : interpolateTacsat • AIS interpolation *interpolate_ais.R*	-Be careful on the amount of interpolated points (e.g. *Hintzen et al. (2010)*) -Interpolation may not be uniform (e.g. low frequency in steaming events and high frequency in fishing events). Careful this does not affect the model - not needed for all models. **Points_on_land_polygon**
Downsampling	• Mean • Nearest	• Downsample using timestep (nearest) *downsample_ais_timestep.R*	Too high resolution data could be a problem in terms of pro-

			cessing and won't provide additional information. Sensitivity assessment (develop a protocol in the future?)
Post-process	• Points in land • Outliers	• Points_on_land_polygon • sf::st_intersects / st_join /st_buffer • VMSTools • RANN:nn2 (must be the fastest method implemented in R to calculate distance between points)	• Recalculate variables • Check unrealistic values if using non linear interpolation and downsampling

4.3 Identifying trips

Table 4.2. Methods for identifying trips.

Issue	Method	Link to R function in GitHub	Discussion
How to define start and end of a trip	• If vessels are harbour based, check when the vessel leaves the polygon and when it enters the polygon • If there are hotspots not defined as harbours, use the same as in harbours. • If the vessels are not harbour based, use a buffer from the coastline • Use a time threshold (on points outside the harbours/coast), when the boat is stopped for more than a certain amount of time (for example, 3h but it varies with the fishery), a new trip would start	• *define_trips_pol function* • *https://github.com/MAPSirbim/AIS_data_processing/blob/main/R/global_functions.R*	Sometimes vessels will leave from the beach, or not turn AIS on until it is outside the harbour. Definition of harbor polygon Examples from participants include time interval thresholds from consecutive vessel/pings to identify fishing trip (this would work for daily trips)

4.4 Methods to infer fishing activity

It is generally possible to identify the fishing activity of a fishing vessel using high-resolution geospatial data, but vessels using active gears are typically easier to process than passive gear vessels. For active gears (trawl, seine, etc.), a fishing event is normally a continuous sequence with a clearly defined range of speeds within a certain amount of time and a typical geometrical pattern corresponding to the characteristic of the *métier*. On the other hand, *métiers* using passive gears (gillnet, longline, pots, etc.) are typically characterised by a 2-fold sequence corresponding to the setting (deployment) of the gear, and its hauling (retrieval) after a certain soak duration.

In some circumstances, e.g. when no logbook data are available to identify individual fishing trips, it may be useful to define fishing events first in the analysis framework, so that a collection of fishing events can then be aggregated as a fishing trip.

Table 4.3. Methods to infer fishing events.

Issue	Method	Link to R function in GitHub	Discussion
Identify fishing activity	• Using speed filter and turning angle • Hidden Markov Models • Expectation Maximization Algorithm • Random Forest • XGBoost • other ML approaches • Image identification techniques (CNN) • ML c 5.0 algorithm (*link*)	Libraries: • randomForest/ranger (really faster) • caret (large choice of machine-learning methods implemented) • kerasR - Interface to Python (large choice of machine-learning methods implemented, including CNN and other deep learning methods). • FactoMineR (principal components analysis and similar methods, useful for summarizing collinear variables and reducing dimension to avoid overfitting) Github: • *https://github.com/MAPSirbim/AIS_data_processing/blob/main/R/global_functions.R* • *https://github.com/ices-eg/WKSS-FGEO/tree/main/R-dev/ToICESGitHub*	Uncertainty with speed filter. Sometimes vessels slow down without fishing e.g. when approaching harbour, or when cleaning the gear from debris after a fishing events (typical for gillnetters) Increase information on activity if available (not dummy variables) using tools and technology such as; • Temperature logger (gear in water or not) • Bathymetry (sonic altimeter) • Sensor data (Electronic Monitoring): sensors measuring information on hydraulics e.g winches. Radio-frequency identification (RFID) transponder tag systems for tagged gear • Video-Based Electronic Monitoring • Mobile phone app, input of fisher • Fishing grounds Try to control overfitting. Be careful on what you want to model depending on the data you have and métier.

Additionally, a small example on how to apply machine-learning has been built during the workshop (link). The folder contains a master script and 10 associated functions. As an example, it uses a subset from the Danish small-scale fleet including validated fishing operations (courtesy of DTU Aqua). It shows how to apply randomForest with a home-made function for optimizing hyper-parameters, but also LDA and QDA models. An example is provided to implement a customised cross-validation process, which will be more suited than automated methods for estimating a realistic accuracy of the model. Moreover, alternative machine-learning methods included in the "caret" package have been used to test other kinds of algorithms (C5.0, SVM, XGBoost, TreeBagging). The accuracy achieved in that example is not satisfying at this stage and requires more investigation. In the future, and with a more relevant choice of covariates, the models could be greatly improved. This kind of approach could be used in combination with unsupervised methods like HMM (Hidden Markov Chains Models) or EM (Expectation Maximization), for an automated interpretation of different states. In the example described, it has been applied on fishing events, but it could be used for identifying different fishing operations or fishing gears. A preliminary reflection has to be made on how to apply this kind of method depending on the purpose (the data may need to be aggregated).

4.5 Model validation data and procedures

Subgroup discussion on model validation data and procedures has led to several recommendations on this topic. Key points are summarized in the items below:

- Type of validation data and collecting methods
- Validation measurements
- Geospatial variables to use for models to classify fishing gears
- Important aspects and AI algorithm requirements

Table summary on the discussion of the previous sections regarding the types and ways of collecting validation data to assess the quality of the models applied into geo-spatial data to classify fishing trips and events.

Table 4.4. Model validation data and procedures.

Issue	Validation data	Rational of usage	Comments
Types of validation data	Expert labelled geospatial data	Researchers and fishermen by looking at tracks on a GIS platform can identify the different phases of a fishing trip. This way it is possible to create data not only to validate the models, but also to train them	It is one of the least costly ways to get validation data and it can be used to validate all phases of a fishing trip. Very important to be very precisely quality controlled. Otherwise the model's performance can be compromised.
	Logbook data	Can be used to validate fishing events and types of gears, when present	Known to be useful and a standard way of validating model predictions/classification performance. Yet it is very

			common that the log-book data is missing and lacks information.
	Observer data	Onboard observer collecting important geo spatial information regarding the different phases of fishing trip	The most standard way of collecting data, yet it is very costly and time consuming.
	Sensors Gear sensors (e. g: temperature, accelerometers with magnetometers) Vessel sensors (e.g: hydraulic pressure, hauler activation)	Temperature sensors allow users to infer the soaktime and start and end of the fishing event for a particular set of gear and can be later merged with the positional data. Accelerometers with magnetometers can improve the calculation of variables related with vessel behaviour (acceleration, bearing, etc.) to identify vessel activity associated with fishing effort. Information of activation of haulers can be cross referenced with geospatial data to identify fishing activities.	Disadvantages of using sensors: Most sensors have short memory space and they have to be frequently collected in order to obtain data. Solution is to have a Hub on the vessel Example of sensor providers: Accelerometers > (https://rbr-global.com/products/sensors , *https://mbientlab.com/?gclid=Cj0KCQiA15yN-BhDTARIsAGnwe0Uxr1CgGsJGYoqAQ-MrWB9a-HDCOb9inZdnsI-btMiDyrdh9QklhiS-GkaApiZEALw_wcB*, not waterproof needs to be put into an underwater housing i.e. deep fishing lights Accelerometers > *https://www.technosmart.eu/axy-5-s/* (€390 per unit, or €290 for bulk buys. It's an extra €100 to integrate the pressure sensor. Accelerometer with magnetometer and option to integrate pressure sensor. They work to 300m, have a ~2 month battery life (recording continuously at 25Hz) which is rechargeable, and weight around 7g. They

			can be programmed to record at 1, 10, 25, 50 or 100Hz, and have a sensitivity range which can be set between +/- 2 and 16G). Temperature and depth -> Star Oddi *Archival tags for fish tagging (star-oddi.com)* Temperature > Tyni tag; *Data Loggers | Tinytag temperature, humidity, CO2 and energy recorders (geminidataloggers.com)*
Validation measurements	Confusion matrices	From confusion matrices it is possible to calculate different variables and statistics that allow analysts to assess the performance of the models, such as accuracy, kappa statistics and F-score. Various statistical parameters can be computed to compare methods and their performance when trying to classify fishing events, gear type, etc.	
	Error measurements	RMSD; MRAD, Bias, percentage of false positives/ false negatives of the classified ping	Validation and accuracy of the model is very dependent on the frequency of pings. It is very important for the detection of specific fishing parameters (soaktime, etc.). Low frequency data may miss fast fishing events e.g. shooting of trammel netting may be over in less than 20 minutes. These events may be missed, but could be picked up by additional sensors as discussed above. The usage of additional sensors may therefore compensate

			for the lower frequency of data to infer particular fishing events within tracks. Both geographical information and sensor data need to be merged via time tags.
	Point density distribution	Comparison between observed and expected distribution. Skewness and kurtosis measurements.	
	Regulations, cultural, and operational aspects of a fishery	Some fishing gears are not allowed to be used in certain areas or are not physically possible to operate in certain areas (e.g. trawlers used near the shore or over a rocky bottom) or at certain times of the year. On the other hand, seasonal and cultural factors can influence the targeting of certain species.	
	Awareness of the procedures of a fishing operation	Knowing the average time taken to haul a static gear, or how long a trawling fishing event is, can provide various insights on assessing if a particular track can or cannot be a fishing event. For example, if the model classifies a very short segment of a track as "fishing- hauling" and if it is known that the fishing gear has at least a certain length, then the most probable case is that the model misclassified this track segment.	To get validation data, it is better to concentrate the effort on fewer vessels but in a comprehensive timespan of trips/operations than otherwise. Important information can be extracted from past and future fishing trips that can improve the identification of particular fishing events, like site fidelity, time and day a gear was deployed and then hauled

Table 4.5. Geo-spatial variables to use in the model to classify the different fishing gears.

Type of variable	Rational/comments/examples
Time of the fishing event	Some fishing events only happen during a certain time of the day, depending on the target species and operational aspects of that metier.
Vessel speed	Even within passive gears some are more challenging and take longer to be deployed or hauled.
Turning angle/direction of the vessel when in a fishing event	Some gears are deployed/hauled along the coastline, for example or within the same bathymetry strata. On the other hand some gears can be operated regardless of changes in depth.
Acceleration	Some gear need to be operated at a constant and slow speed (e.g: trawls and passive gears), whilst others require the vessel to deploy the gear at a fast pace (e.g. seiners).
Straightness of path, Heading, Delta heading (rate of turn)	Gears like purse seiners have a circular path while fishing, while passive gear, like gillnets tend to maintain a more straight path. Trawlers also tend to keep a steady heading while fishing.
Bathymetry	Depending on the gear and target species, gears are operated at different depth strata.
Habitat/ substrate	Some gears have a strong association with a particular substrate.
Known wrecks/features/obstructions	Obstacles /obstructions can damage gears, specially active ones
Time of the year / seasons	Depending on the target species, some specific gears can only be used during certain time periods when that species is caught/available/desired.
Tidal cycles	Some gears can be used during certain tidal periods
Prior knowledge about the gear	Regulations, how the gear is used, where it is used, where it is not allowed, etc.
Vessel features	LOA (length overall), vessel type, engine power etc (e.g. information available in the EU fleet register).

Some important aspect and requirements of AI algorithms:

- Training data have to be quality checked before using.
- Artificial Intelligence (AI) algorithms are able to solve one problem at a time. If you use only one algorithm, you have to create a class for every combination of parameters you want to predict (e.g. gear, fishing/steaming, gear size or length…). It should be better to use an algorithm chain to identify gear, then fishing periods and finally fishing gear characteristics.
- What if some variables such as depth, vessel length, etc. are missing? Some algorithms are not able to deal with 'NA' (missing values). Some other algorithms just skip the lines where at least one variable = NA. Missing variables could be estimated (e.g. Depth can be estimated using functions such as the getNOAA.bathy() function).
- Do we need to know how the decision tree is built?
- Make sure that the training dataset is representative of the reality (different countries, metiers, …)
- The computation time could be very different from one model to another one. Some models are more willing to deal with large datasets.

- Avoid unbalanced dataset. This could be an issue and can lead to using oversampling methods or to only select algorithms allowing weighting between the classes.

4.6 Fishing effort indicators

First, the subgroup evaluated how geospatial data can inform and keep consistency with reporting requirements for the DCF (Effort variables defined in EU Map 2021/1167, Table 6). Special consideration was given to documenting the differences between mobile and static gears when assessing these reporting requirements.

The group assessed all variables in the EU Map table 6 and evaluated how these variables could be derived from geospatial data for SSF. Then, the subgroup discussed how high resolution geospatial information can add valuable information that is currently not implemented. A series of recommendations were also added to the summary table.

Subsequently, the subgroup evaluated which potential spatial fishing descriptors could be derived from highly resolved spatial data that would improve management of SSF. Case studies are used to illustrate the utility of these descriptors.

Table 4.6. Effort variables from EU-MAP table 6, and how they can be calculated using high-resolution geo-spatial data.

Effort variable, see definitions in glossary section 3	**How it can be calculated**	**SSF aspects**	**Recommendations**
Days at sea	To calculate days at sea we need to know the duration of a fishing trip, this requires identifying a fishing trip – see tools in section 4.3 Needed: high resolution spatial data only	With geo-spatial data we can calculate more precisely (hours at sea).	Hours at sea (h)
Hours fished (optional).	Mobile gears: Infer fishing Static gears: Soak time needs inferring deployment (time t) and hauling (time t+y) – might or might not be one trip. Needed: high resolution spatial data +models to infer setting and hauling activities	Identifying deployment might prove difficult from geospatial data if deployment spatial and speed patterns are not distinguishable from other activities.	Redundant for static gears, see soak time
Fishing days	Mobile gears: Infer fishing	>10-<12 m paper logbooks 10 and under – country dependent	If each day with "remained" operational gear is counted then the number of fishing days associated with each gear unit would be part of different fishing trips.

	Static gears: Soak time needs inferring deployment (time t) and hauling (time t+y) – might or might not be one trip Needed: high resolution spatial data +models to infer setting and hauling activities	Weekly reports, monthly, sale notes	
kW*Days at sea	Days at sea*kW Needed: high resolution spatial data Vessel information from fleet register		Information on vessel id is on EU vessel CFR code (to avoid issues with language)
GT*Days at sea	Days at sea*GT Needed: high resolution spatial data Vessel information from fleet register		Information on vessel id is on EU vessel CFR code (to avoid issues with language)
kW*Fishing days	Fishing days*kW Needed: high resolution spatial data +models to infer setting and hauling activities Vessel information from fleet register		Information on vessel id is on EU vessel CFR code (to avoid issues with language)
GT*Fishing days	Fishing days*GT	Vessel information out of date/not reliable in some SSF	Information on vessel id is on EU vessel CFR code (to avoid issues with language)
Number of trips	Needed: high resolution spatial data +models to infer setting and hauling activities + method to define start and end of trip - see table 4.3 for methods and tools		Port should be changed to include other starting point locations such as beaches. Possibly “Port and other boat parking areas”
Number of fishing operations	If fishing/steaming is identified in position data, count the number of fishing operations	Depending on specific static gears, the distinction between setting and hauling might be complicated.	In some cases, the setting of the gears are not easy to identify (due to high speeds), and therefore it makes more sense to only include hauling. It makes sense that for every haul, there should be an associated deployment event. It is only possible to consider hauling events where every event marked as hauling is definitely hauling and not setting.

			An alternative suggestion is that the 'number of fishing operations' is a more general term for when some type of fishing activity has happened. Another effort variable could be defined as 'number of hauling operations' for those cases where the hauling activity can be defined. Confirm there is a mention of the characteristics of fishing operations. Ideally (best case scenario), have information on set AND haul operation, including beginning and ending of spatio-temporal data for both events (set and haul). Possibly consider a three category verification indicator, 1-identified set and haul activity (ideal situation), 2- has identified only haul activity 3- identified only set activity (less reliable information, may be useful to identify fishing ground).
Length of nets (m)*soak time (days)	An estimation of net length can be provided based on positional data, i.e, accumulated distance between positional observations inferred as hauling. Soak time can in principle be calculated after identifying deployment and hauling events by overlapping these two operations and calculating the time between both events.	Issues defining deployment events and maybe hauling events using geospatial data only.	The total length of the nets can be validated from conversations with fishers which in some cases mention the number of panels/sets and might not be accurate. Need to work on an appropriate definition of soak time.
Number of nets/length	Length of net - as above: An estimation of net length can be provided based on positional data, i.e, accumulated distance between positional observations inferred as hauling.		

Number of hooks, number of lines	An estimation of number of hooks can be provided based on positional data, provided we have some validation data (from for example onboard observers, logbooks) i.e, accumulated distance between positional observations inferred as hauling. Soak time can in principle be calculated after identifying deployment and hauling events by overlapping these two operations and calculating the time between both events.		The total number of hooks need to be validated from conversations with fishers Need to work on an appropriate definition of soak time in relation to longlines
Number of pots, traps	Information from observers on the number of pots and distance between pots. From that the number of pots can be estimated. Vessel specific. Issue with underestimation/overestimation of number of pots, probably due to lack of accurate data from fishers on pot spacing. Hard to identify setting and hauling of the pots.		The total number of pots/traps need to be validated from conversations with fishers Need to work on an appropriate definition of soak time in relation to pots/traps. The length of the sets of pots/traps might also be a potentially useful effort indicator.

Some of the effort variables are more relevant to large-scale fisheries (LSF), whereas applying them to the SSF would overestimate the fishing effort in the small-scale fishery, e.g. fishing days instead of fishing hours. However, the effort variables that take the size of the vessel/engine into account (e.g. kW*Fishing days, GT*Fishing days) are partially accounting for this problem. Regarding the variable hours fished, it is not clearly defined in the EU-MAP for vessels fishing with passive gears if it is fishing hours for the vessel or fishing hours for the gear (soaking time). Also some standardization is needed for calculating soaking time, i.e. is it from the start, midpoint or end of setting/hauling the net, as for some gears, this can take a long time.

For some of the effort variables, a relationship can be found between the high-resolution data and effort (number of pots/traps, number of hooks) through case studies. This could be done through a representative survey for each main fishery with onboard observers. There should be awareness of new fishing gears appearing in the market, e.g. Australian octopus traps. Further, the importance of each effort variable should be studied for each fishery. For example in an octopus fishery with traps the relevant measure might be soaking time, while for fisheries with pots, the number of pots might be relevant together with the soaking time. In general, measures of soaking time for pots, traps and longlines are missing in the above effort metrics table.

It is useful to map fishing effort for the SSF, based on the high-resolution data, on a finer scale than what is often done in the LSF, as the activities are often more concentrated in time and space.

4.6.1 Case study of the North Irish Sea

The management of the north Irish Sea razor clam fishing fleet requires all vessels to record their position through an inshore VMS system. All vessels in the fleet fish for razor clams using hydraulic dredges (HMD) and are all less than 12m in length. The iVMS system records positional information, speed and bearing every 10 minutes. Hours fished (Figure 4.1) is derived using a speed based rule to define fishing activity from steaming activity. Additionally, a spatial buffer is used to exclude fishing activity unlikely to be HMD gear (as razor clam fishery in the North Irish Sea is well defined in space, and limited to depths less than about 12 m) or for which razor clams are not the target species.

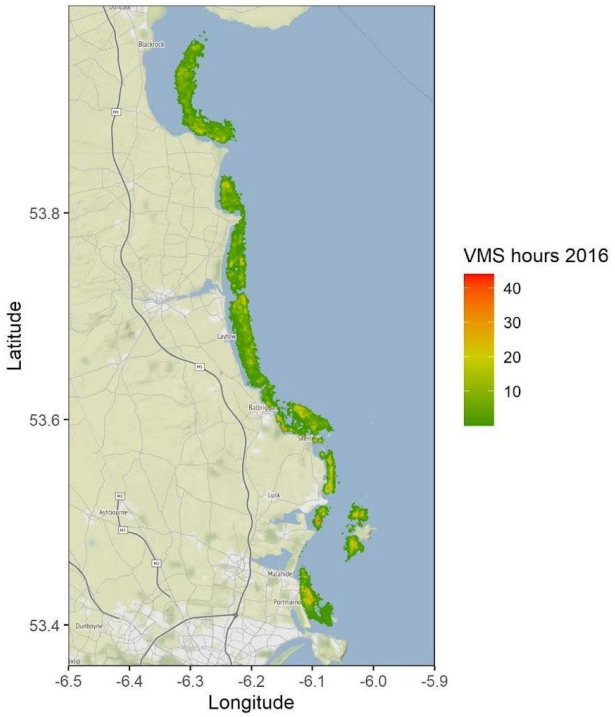


Figure 4.1. iVMS hours fished for the razor clam fleet in the North Irish Sea in 2016. iVMS is mandatory in this fleet, all vessels are under 12m in length and positional information is recorded every 10 minutes.

4.6.2 Case study of the Scottish trap fisheries

Increased use of marine areas in Scotland may lead to conflicts over space and resources. It is therefore important to accurately map fishing activities to inform local, regional and national fisheries management as well as marine planning and related policy commitments. The use of appropriate vessel tracking systems has been articulated by the Scottish government in the Scottish Inshore Fisheries Strategy and a national discussion paper on the future of fishers management in Scotland.

The Scottish Inshore Fisheries Integrated Data System (SIFIDS, https://masts.ac.uk/research_projects/scottish-inshore-fisheries-integrated-data-system-sifids-project/) project developed systems and processes for analysis of geospatial data from static gears. Using these processes, tracks are analyzed automatically to infer hauling activities (Figure 4.2). Using the distance covered

during a hauling event, an estimate of numbers of creels deployed in that trip is included. Main fishing grounds can therefore be evaluated at fine spatial scales, relevant to SSF management (see example in Figure 4.3).

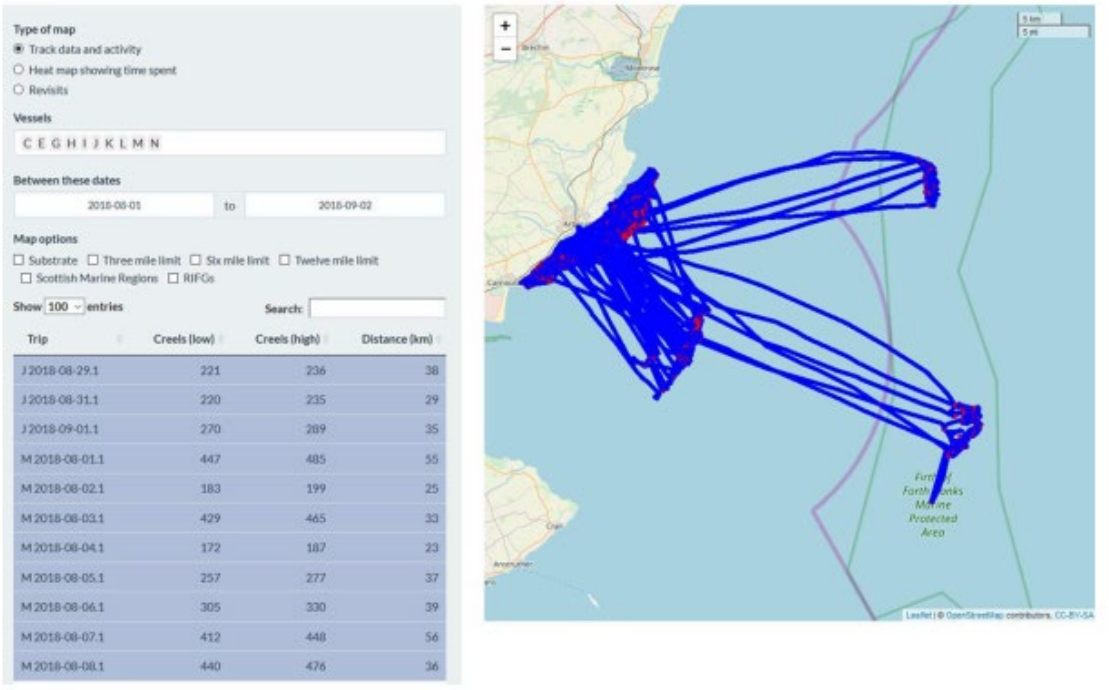


Figure 4.2. Map showing track data (blue) and hauling activities (red) identified from models.

Confidence intervals (95%) (creels (low), creels (high)) are shown in the data table of figure 4.2 for the estimated number of creels and the distance covered during each trip.

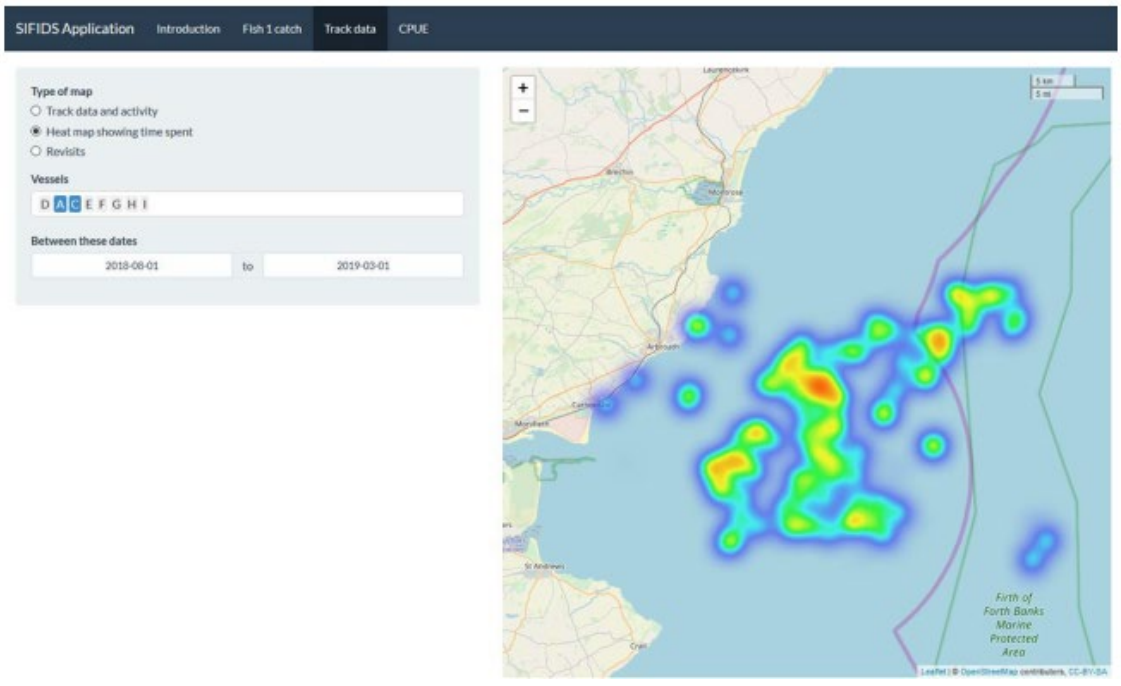


Figure 4.3. Heat map showing most important fishing grounds for 6 SSF vessels operating in the East Coast of Scotland (taken from Mendo *et al.*, 2019).

4.6.3 Case study of the South Portugal Octopus traps and pots fisheries

Octopus trap and pot fishery in the south of Portugal is the most important small-scale fishery of the region. However, use of space and spatially resolved fishing effort of this fishery remain unclear, reinforcing the need of accurate maps for fisheries management but also to contribute to better spatial management of the activities at sea.

In the national MAR2020 research project ParticiPESCA, almost 60 GPS trackers set in volunteer fishing boats are used to describe their spatial activity, while simultaneously, onboard observers track and describe fishing operations obtaining training data to allow automatic identification of setting and hauling activities.

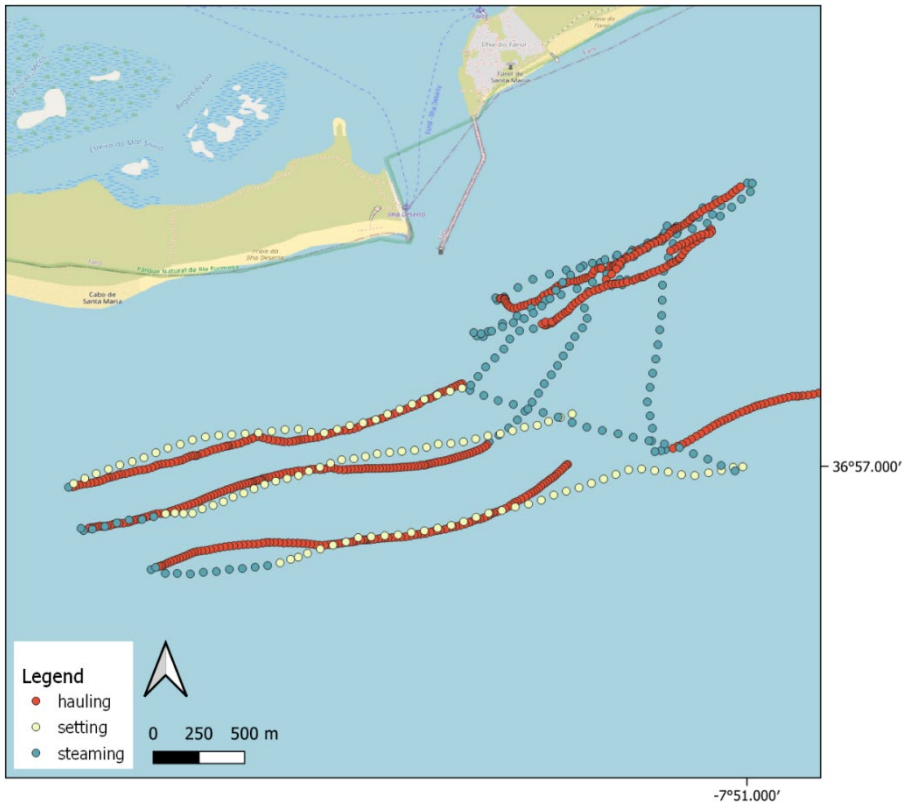


Figure 4.4. Map showing different fishing operations, hauling in red, setting in yellow and steaming in green for one vessel trip belonging to the octopus small-scale fisheries in the Algarve, Portugal.

4.6.4 Case study of the Portuguese bivalve dredge and octopus pots and traps fisheries

An important fishery targeting thirteen species of coastal bivalves takes place in Portugal, in three main fishing grounds located in the northwest, southwest and south mainland coast, using bivalve dredges. The fishery involves 80 vessels, with LOA between 5.45 and 15.82 m, that have captured 2100–2700 tons officially between 2018–2020. In the south of the country, the fishery operates from 3–15 m depth, whereas in the northwest it can reach 34 m depth.

Bivalve fisheries using dredges in Portugal have different types of tracks.

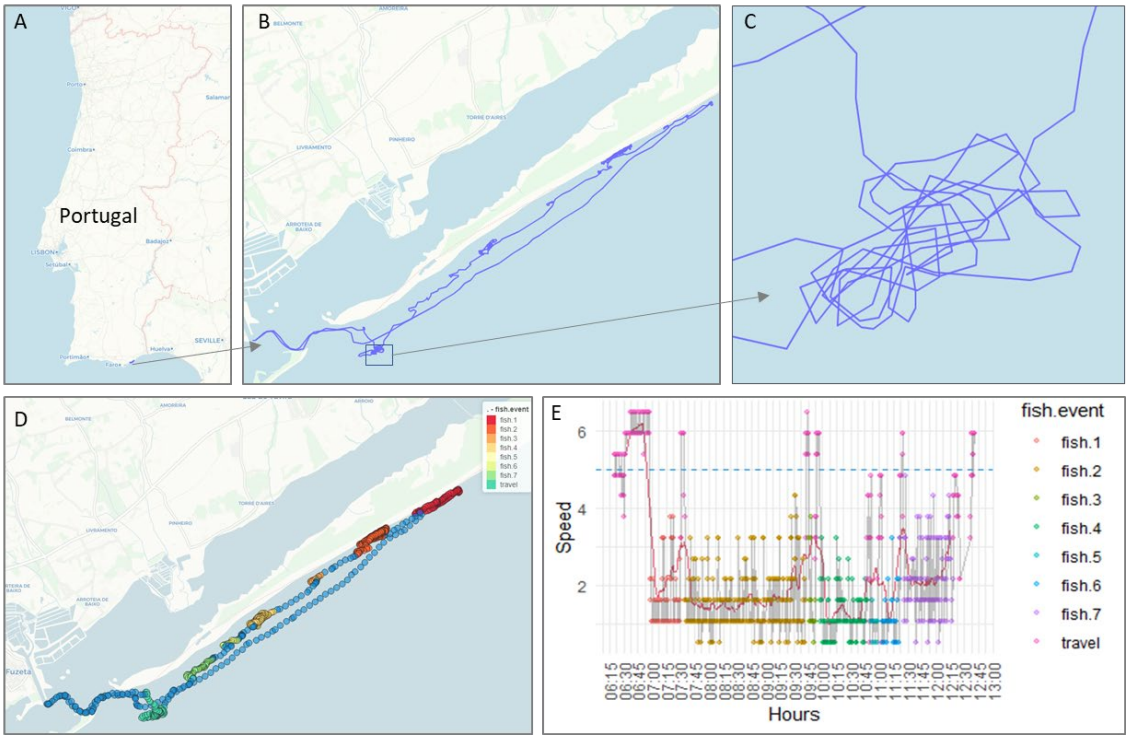


Figure 4.5. Map showing a track of one typical bivalve small-scale fishery trip in the Algarve, South of Portugal. Panel A, location in the country, B: complete fishing trip; C: zoom that shows the 'loops' typical of this fishery; D: map of fishing events; E: speed plot with signalled fishing events and moving speed average.

Since 2016 to present, all bivalve dredges vessels have been equipped with real-time GPS trackers that record their position and operation status every ~30 seconds, under the framework of research project MONTEREAL (MAR2020). This dataset is currently being processed to estimate and map high resolution fishing effort, which along with the annual independent bivalve fishing surveys carried out by IPMA (Portuguese Institute for the Sea and the Atmosphere), the official landings and auction sales recorded by the government and in collaboration with the stakeholders (fisherman and associations) will be used to manage the fishery and to better understand the functioning of this dynamic and vulnerable coastal ecosystem.

Within this research project, an additional five vessels targeting octopus using pots and traps and ten vessels using trammel nets have also been equipped with the same GPS trackers for exploratory purposes.

5 Methods/Scripts

A subgroup worked on scripts for processing the high-resolution data. A GitHub is available for testing and storing scripts: https://github.com/ices-eg/WKSSFGEO

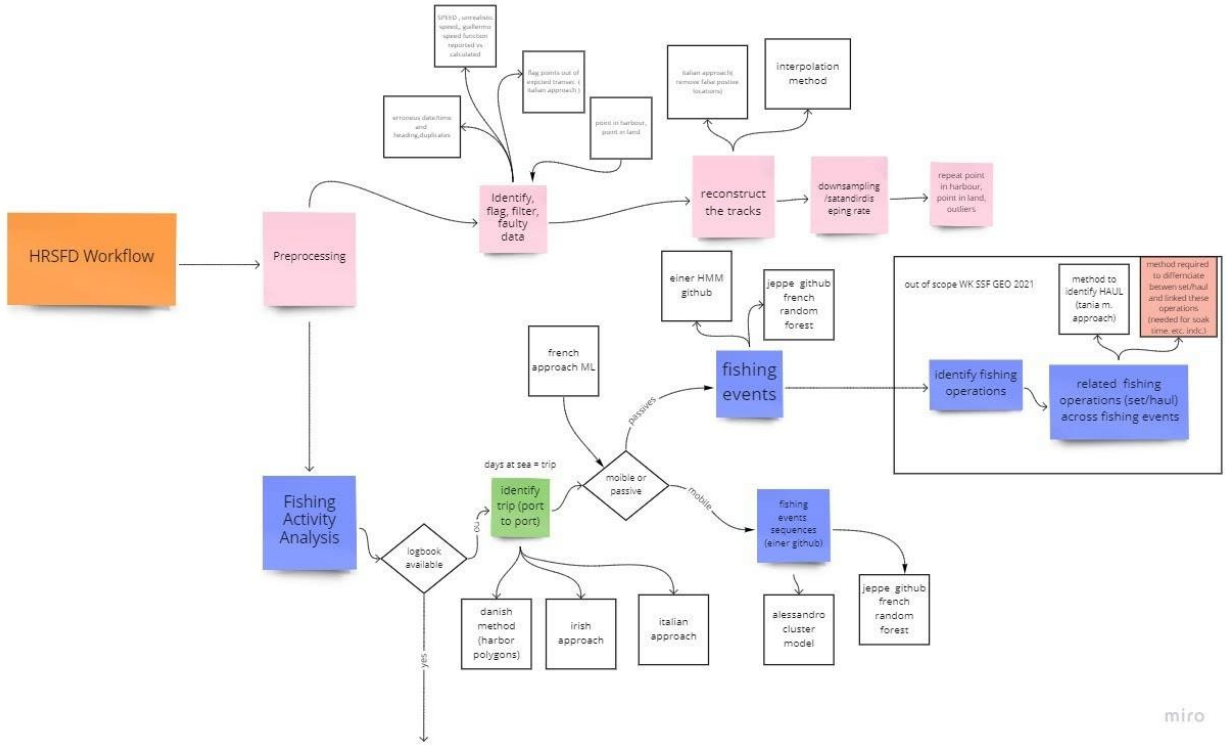


Figure 5.1. Workflow for high-resolution spatial fisheries data. Pre-processing steps are shown in pink, fishing activity analysis in blue and identifying trips in green.

5.1 Scripts for pre-processing steps

Pre-processing steps are also discussed in section 4.2. In table 5.1 scripts for identifying and filtering faulty data are listed.

Table 5.1. Pre-processing: identify and filter faulty data.

Function/ code	Methods description	author/ email/ link	type of data	gear/ metiers	Note
Italian approach		Alessandro Galdelli a.galdelli@univpm.it	AIS data 5 min ping rate	All type	The developed code is written in Matlab, so it is currently undisclosed. It is being converted to R language and will be available soon.

vmstools functions			VMS data, but pre-processing functions applicable to other types georeferenced spatial fisheries data	All	

In some cases, the intervals between positions vary. Therefore, it can be useful to reconstruct the tracks of the vessels using interpolation methods, see table 5.2.

Table 5.2. Reconstruct tracks/interpolation.

Function/code	Methods description	author/email/link	type of data	gear/ metiers	Note
altered vmstools approach	interpolation method	Jeppe/ Niels/ vmstools			

After interpolation, to reduce the amounts of raw data to process, it can be downsampled using a time-step, e.g. every 5 minutes to the nearest positions.

Table 5.3. Downsampling after interpolation or reducing raw data.

Function/code	Methods description	author/email/link	type of data	gear/ metiers	Note
Downsampling approach	Downsample using timestep (nearest)	Jeppe			

5.2 Fishing activity analysis

Some primary steps valid for all gear types are to make some additional variables based on the data, e.g. calculation of the speed from distance and time between position, mean speed over a time-period, changes in direction of the vessel.

Table 5.4. Make new variables from speed, course position etc.

Function/ code	Methods description	author/email/link	type of data	gear/ metiers	Note
French approach	Make new variables from AIS data	Francois	Jerk/ Directional change/ Acceleration/mean speed/sinuosity in time		

Table 5.5. Identify trip (port to port).

Function/ code	Methods description	author/ email/link	type of data	gear/metiers	Note
Danish method (harbour polygons)	Identify the time from leaving till returning to harbour. In case of large gaps in the data, the trip can be split into several trips	Jeppe Olsen (jepol@aqua.dtu.dk)	10 seconds EM (video validated data)	Tested on GN/OT, but applicable for all gears	
Irish method	Single trips for each vessel assumed as consecutive pings out of harbour	Guillermo Martin (guillermo.martin@marine.ie)	Inshore VMS data. 5 min ping rate	HMD but applicable to other gears	pointsinharbour() and sortTacsat() functions from vmstools package required
Italian approach	The create_fishing_trip function identifies vessel-specific fishing trips for each vessel, as sequences of points broadcasted by a vessel, from the time it leaves the port until it returns.	Alessandro Galdelli (*a.galdelli@univpm.it*)	AIS data 5 min ping rate. Test dataset is available *here*	All type	*create_fishing_trip(data = dat, ports = ports, ports_buffer = port_buf, coastal_ban_zone = coastal_ban_zone)* To run the function, different datasets are required: the sequence of AIS positions of a vessel, the coastal_ban_zone layer, and a layer related to the harbour. A recovery function was internally applied to join consecutive trips where the departure/arrival port was too far to be assigned.

Table 5.6. Identifying fishing event for passive gears.

Function/ code/	Methods description	author/ email/link	type of data	gear/ metiers	Note
Icelandic approach	Hidden Markov Models (HMM)	Einar			
French approach	Random Forest, toy example. Variables are not that well fitted and chosen to the example data.				Described in section 4.4

Table 5.7. Identifying fishing events sequences.

Function/ code	Methods description	author/ email/link	type of data	gear/metiers	Note
Italian approach	Cluster model	Alessandro Galdelli *a.galdelli@univpm.it*	AIS data 5 min ping rate. Test dataset is available *here*	OTB-PTM-TBB - PS (not tested)	classification_wrapper(vessel_data=dat_with_trip, pars=pars, write.output=T, output.name = "test1") Applies a cascade of classification algorithms on each fishing trip, using as input the AIS positions with their assigned trip (dat_with_trip) and the specific parameters needed for each classification algorithm (pars).
Ireland potter classification function	Hidden Markov Model (HMM)	Guillermo Martin guillermo.martin@marine.ie	VMS data at pings rates <= 5min	FPO	The function is constructed as having the georeferenced spatial fisheries data as a data input. Cleaning and fishing trip definition are executed within the function. More information is available in the *FPO_classifcation_user manual*

6 Conclusions and recommendations

Conclusions

Using data with high spatial and temporal resolution, is essential to estimate fishing effort of small-scale fisheries (vessels <12m) that use static and mobile gears, as well as large-scale fisheries with low duration of fishing events. Additional information coming from for example sensors placed on the fishing gear might help to refine the process of the fishing effort identification. The control regulation (EU 1224/2009) requires Vessel Monitoring Systems (VMS) on vessels larger than 12 m and with position intervals in large-scale fisheries up to 2 hours, which is often longer than many small-scale fisheries fishing trips, and therefore inadequate for estimating fishing events and fishing effort in SSF.

Geospatial data on fisheries are essential for marine spatial planning, fisheries management, seafood traceability, and for assessing and monitoring fishing impacts on the marine and coastal ecosystems including spatial control of fishing, as the SSF represent 85% (70 000 vessels) of the EU fleet. These data are also important for assessment of the economic and social value of small-scale fisheries and how small-scale fisheries link to coastal communities.

In the EU, VMS is mandatory for fishing vessels from 12 metres in length overall and more, and Automatic Identification Systems (AIS) from 15 metres. Logbooks are mandatory only from 10 metres in most Member States, and a large proportion of the SSF fleets are not monitored or under-monitored. The EU Control Regulation and the following Commission Implementing Regulations (404/2011 and 2015/1962) have accelerated the development of electronic technologies (ET) to monitor the activity of fishing vessels of all length classes in the last decade. In many cases however, it may be difficult to derive an estimation of fleet-wide fishing effort from devices originally conceived for other purposes. For instance, VMS was designed with fisheries control in mind, while AIS was intended to limit the risk of collisions and can be turned off manually by the user. Since the mid-2000s, electronic monitoring (EM) programmes that can collect a census of the fishing activity have started to emerge and are able to gather fine-scale fishery-dependent data in SSF reliably. Nevertheless, the cost of EM remains high and often prohibits the equipment of a large number of vessels. Still, the implementation and running costs of data collection with EM can be overcome using appropriate incentives and adequate financial compensations. Likewise, SSF fishers could benefit from specific financial support to implement ET for monitoring and reporting fishing activity data. Moreover, from a fisher's viewpoint, a constant tracking can be perceived as a threat to his/her own activity, and privacy and confidentiality concerns often arise from conversations with professionals (Plet-Hansen *et al.*, 2017, Barz *et al.*, 2021). To overcome trust concerns and to limit data misuse, it is essential that the end-users (e.g., fisheries scientists) are fully transparent in the way these sensitive data are utilised and that they are treated with the highest possible level of confidentiality (Dalskov *et al.*, 2021).

It is important to define different aspects of fishing activity e.g. fishing trips, fishing events, fishing effort prior to analysis of geospatial data. Clear objectives and metrics should help to guide the development of the conceptual framework for the reporting and analysis of geo-spatial data. The temporal intervals required to estimate fishing effort in these fisheries greatly depend on the fishery, but should always be within a range of seconds to minutes. Overall, a precautionary approach would be to acquire data with the highest resolution possible and to downsample where feasible afterwards. Sensitivity assessment for the minimum temporal resolution can be undertaken. Interpolation between reported pings can also enable 'upsampling' to increase frequency if this is required for successful data analysis. There is a big difference in requirements with respect to temporal resolution of the data between just locating fishing activity (which is

valuable in itself) and being able to make precise estimates of fishing effort and its footprint. More data to validate the capacity of geospatial data to estimate fishing effort are needed e.g., onboard observers could track the fishing operations, or a sample of the fleet (e.g., a reference fleet) could be monitored with EM for a period.

The workshop participants have suggested a common framework for the pre-processing and analysis of geospatial data in order to identify fishing trips and fishing events. A GitHub was set up for the workshop to organize the R-scripts for working with the high-resolution geo-spatial data, including workflow examples, scripts for data pre-processing and different methods for assessing trips and fishing events.

Many different national initiatives are underway in the EU to collect different types of fine-scale data in SSF (AIS, iVMS, EM, BlackBox, GreenBox, GPS, etc.), each with different advantages and challenges (Quincoces, 2021; van Helmond, 2021; Dalskov *et al.*, 2021). A more widespread use of high-resolution geo-spatial data associated with reliable methods for analysing these data would considerably improve the understanding of fishing activities of EU SSF fleets, and reduce the information needed to report in e.g., logbooks or surveys for the small-scale fisheries. Further work is required on the effort indicators for small-scale fisheries and homogenisation with the EU-MAP variables (table 6) which were developed mainly for large-scale fisheries. It is clear that these are relevant to small-scale fisheries, but need to be adapted, requiring further work gathering experts on the subject.

The maps of small-scale fisheries effort should always have a much finer gridding resolution, between 100 and 1000 m to be meaningful, taking into account the spatial scale of the fishery.

Recommendations for future work

As many terms and definitions relating to fishing activities are developed with trawlers filling in logbooks in mind, it was found that the definitions need to be refined for working with the high-resolution geo-spatial data. The workshop started the work and found that definitions can be ambiguous. Further discussions and agreements on definitions are needed. Further development of the glossary started in the workshop, with corresponding infographics and illustrations would be very useful.

More work is needed to develop and test methods to classify the positions and infer fishing operations, namely setting and hauling. Sensors can help inform on this for development of the methods, which can then be applied to larger datasets, e.g., AIS data. For details on how the different fishing gears operate, a collaboration with the Joint ICES/FAO Working Group on Fishing Technology and Fish Behaviour (WGFTFB), and specifically SSF technology experts on passive gears, might be useful.

Further work is required on the effort indicators for small-scale fisheries and harmonisation with the EU-MAP variables (table 6) which were developed mainly for large-scale fisheries (LSF). If the positions of fishing vessels are classified correctly into e.g., fishing trips, steaming, gear deployment, hauling and other activities the main information for effort indicators is obtained. For gillnets, the total length of nets might be calculated if positions are classified correctly, and the ping rate is sufficient to ensure an accurate approximation of the total length of the net fleets. For static gears, further development of methods to calculate soaking time is needed, but ongoing studies using machine learning processes are encouraging. In addition, parameters like number of pots/traps (for pots and traps fisheries), number of lines and hooks (for longline fisheries) could be estimated if relationships with known parameters are established.

If the use of high-resolution geo-spatial data becomes more wide-spread and standardised, it could fill the major data gaps which exist for the small-scale fishery segment. It would in turn be

very beneficial to integrate the SSF data into assessment frameworks for Marine Strategy Framework Directive Descriptors (D3 Commercial seafood and shellfish and D6 Seafloor Integrity), the spatial mapping of fishing effort done by ICES WGSFD, the assessment of cumulative effects, and Article 6 of the Habitat Directive.

WKSSFGEO recommends that an additional workshop is arranged to follow up on the work that was started in this Workshop, and to further develop and explore methods to classify the high-resolution geo-spatial data into fishing activities for different types of gears. If positions are recorded on a high resolution for a larger part of the fleet, this can result in better effort estimates for the small-scale fisheries and for the static gears. Future work within the ICES community on methods and workflows for processing the high-resolution geo-spatial data should be done with a strong link to ICES WGSFD (Working Group on Spatial Fisheries Data), that has a similar workflow for processing VMS data for the ICES VMS/Logbook data call. A possible collaboration with the Working Group on Technology Integration for Fishery-Dependent Data (WGTIFD) should also be considered as they also have ToRs relating to the development of standardized formats for data collected and analysed from EM systems, and on publishing recommendations for interoperability of EM systems.

7 Acknowledgements

We would like to acknowledge IPMA for hosting the meeting and all groups who provided test data and scripts for the workshop.

Marta M. Rufino was funded by the Mar2020 project MONTREAL.

Nuno Sales Henriques participation was funded by the PhD scholarship reference number 2020.05583.BD, by FCT - Fundação para a Ciência e Tecnologia.

8 References

Barz, F. *et al.* (2020) ‘`Boats don’t fish, people do’- how fishers’ agency can inform fisheries-management on bycatch mitigation of marine mammals and sea birds’, Marine Policy, p. 104268. doi:10.1016/j.marpol.2020.104268

Bjordal, A. (2009) A Fishery Manager’s Guidebook. Second Edition. Cochrane, K. and Garcia, S. (eds), The Food and Agriculture Organization of the United Nations and Wiley-Blackwell, Rome.

Dalskov, J., Glemarec, G., Kindt-Larsen, L., Kroner, A.M. and Nielsen, P., 2021. Workshop on electronic technologies for fisheries Part III: Systems adapted for small scale vessels, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels

FAO. 1997. Fisheries management. FAO Technical Guidelines for Responsible Fisheries No. 4. Rome, FAO. 82p.

EU 404/2011. Commission Implementing Regulation (EU) No 404/2011 of 8 April 2011 laying down detailed rules for the implementation of Council Regulation (EC) No 1224/2009 establishing a Community control system for ensuring compliance with the rules of the Common Fisheries Policy. OJ L 112, 30.4.2011, p. 1–153. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32011R0404

EU 1224/2009: Council regulation (EC) No 1224/2009 of 20 November 2009 establishing a Community control system for ensuring compliance with the rules of the common fisheries policy, amending Regulations (EC) No 847/96, (EC) No 2371/2002, (EC) No 811/2004, (EC) No 768/2005, (EC) No 2115/2005, (EC)

No 2166/2005, (EC) No 388/2006, (EC) No 509/2007, (EC) No 676/2007, (EC) No 1098/2007, (EC) No 1300/2008, (EC) No 1342/2008 and repealing Regulations (EEC) No 2847/93, (EC) No 1627/94 and (EC) No 1966/2006. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32009R1224

EU 1380/2013 of the European Parliament and of the Council of 11 December 2013 on the Common Fisheries Policy, amending Council Regulations (EC) No 1954/2003 and (EC) No 1224/2009 and repealing Council Regulations (EC) No 2371/2002 and (EC) No 639/2004 and Council Decision 2004/585/EC. OJ L 354, 28.12.2013, p. 22–61. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013R1380

EU 2010/93/EU. Commission Decision of 18 December 2009 adopting a multiannual Community programme for the collection, management and use of data in the fisheries sector for the period 2011-2013 (notified under document C(2009) 10121). OJ L 41, 16.2.2010, p. 8–71. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010D0093

EU 2021/1139. Regulation (EU) 2021/1139 of the European Parliament and of the council of 7 July 2021 establishing the European Maritime, Fisheries and Aquaculture Found and amending Regulation (EU) 2017/1004.

EU 2021/1167. COMMISSION DELEGATED DECISION (EU) 2021/1167 of 27 April 2021 establishing the multiannual Union programme for the collection and management of biological, environmental, technical and socioeconomic data in the fisheries and aquaculture sectors from 2022.

EU 2015/1962. Commission Implementing Regulation (EU) 2015/1962 of 28 October 2015 amending Implementing Regulation (EU) No 404/2011 laying down detailed rules for the implementation of Council Regulation (EC) No 1224/2009 establishing a Community control system for ensuring compliance with the rules of the common fisheries policy. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32015R1962

EU 2018: Proposal for a regulation of the European Parliament and of the Council mending Council Regulation (EC) No 1224/2009, and amending Council Regulations (EC) No 768/2005, (EC) No 1967/2006, (EC) No 1005/2008, and Regulation (EU) No 2016/1139 of the European Parliament and of the Council as regards fisheries control. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018PC0368&from=EN

Kavadas, S., Maina, I., Damalas, D., Dokos, I., Pantazi, M., Vassilopoulou, V., 2015, Multi-Criteria Decision Analysis as a tool to extract fishing footprints and estimate fishing pressure: application to small scale coastal fisheries and implications for management in the context of the Maritime Spatial Planning Directive. Mediterranean Marine Science, 16/2, 294-304.

Mendo, T., Smout, S. Ransijn, J., Durbach, I., McCann, P., Crowe, S. Carulla Fabrega, A., de Prado. I., James, M. 2019. Scottish Inshore Fisheries Integrated Data System (SIFIDS): Identifying fishing activities and their associated drivers. Report to Marine Scotland in fulfilment of EMFF: SCO1413 68pp.

Quincoces, 2021. Research for PECH Committee – Workshop on electronic technologies for fisheries - Part I: Transmitted positional data systems, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels

Schreiber Plet-Hansen, K., Qvist Eliasen, S., Mortensen, L. O., Bergsson, H., Olesen, H. J., & Ulrich, C. (2017). Remote electronic monitoring and the landing obligation – some insights into fishers' and fishery inspectors' opinions. Marine Policy, 76, 98-106. https://doi.org/10.1016/j.marpol.2016.11.028

Kim Stobberup, María Dolores Garza Gil, Aude Stirnemann-Relot, Arthur Rigaud,Nicolò Franceschelli, Roland Blomeyer, S 2017, Research for PECH Committee – Small-scaleFisheries and "Blue Growth" in the EU, European Parliament, Policy Department for Structuraland Cohesion Policies, Brussels. https://www.europarl.europa.eu/RegData/etudes/STUD/2017/573450/IPOL_STU(2017)573450_EN.pdf

van Helmond, A.T.M., 2021, Research for PECH Committee – Workshop on electronic technologies for fisheries − Part II: Electronic monitoring systems, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.

Annex 1: List of participants

Name	Institute	Country	Email
Adriano Mancini	Università Politecnica delle Marche	Italy	a.mancini@staff.univpm.it
Alessandro Galdelli	Università Politecnica delle Marche	Italy	a.galdelli@univpm.it
Andre Carvalho	IPMA	Portugal	andre.carvalho@ipma.pt
Anna Mujal-Colilles	University of St. Andrews	UK	amc53@st-andrews.ac.uk
Anna Nora Tassetti	Consiglio Nazionale delle Ricerche	Italy	annanora.tassetti@cnr.it
Antonio Punzón	INSTITUTO ESPAÑOL DE OCEANO-GRAFÍA	Spain	antonio.punzon@ieo.es
Cherry Harper Jones	CEFAS	UK	cherry.harperjones@cefas.co.uk
Christian von Dorrien	Thünen Institute of Sea Fisheries	Germany	christian.dorrien@thuenen.de
Claudio Vasapollo	ISPRA-Ambiente	Italy	claudio.vasapollo@isprambiente.it
Dan Edwards	JNCC	UK	dan.edwards@jncc.gov.uk
Daniel Cano	INSTITUTO ESPAÑOL DE OCEANO-GRAFÍA	Spain	daniel.cano@ieo.es
Daniela Silva	Applied Mathematics at the University of Minho	Portugal	danyelasylva2@gmail.com
Einar Hjörleifsson	Marine and Freshwater Research Institiute	Iceland	einar.hjorleifsson@hafogvatn.is
Emily Hague	Heriot-Watt University	UK	elh2001@hw.ac.uk
François Danhiez	Capgemini Engineering France	France	francois.danhiez@capgemini.com
Gianluca Franceschini	ISPRA-Ambiente	Italy	gianluca.franceschini@isprambiente.it
Gildas Glemarec	DTU Aqua	Denmark	ggle@aqua.dtu.dk
Guillermo Martin	Marine Institute, Ireland	Ireland	guillermo.martin@marine.ie
Hannah Fennell	Heriot-Watt University	UK	hf4@hw.ac.uk
Helen Holah	Marine Lab Scotland	UK	helen.holah@gov.scot
Hilmar Hinz	CSIC-IMEDEA	Spain	hhinz@imedea.uib-csic.es
Irida Maina	HCMR	Greece	imaina@hcmr.gr
Isidora Katara	CEFAS	UK	isidora.katara@cefas.co.uk
Jacopo Pulcinella	Consiglio Nazionale delle Ricerche	Italy	jacopo.pulcinella@irbim.cnr.it
Jeppe Olsen	DTU Aqua	Denmark	jepol@aqua.dtu.dk
Jonathan Stounberg	DTU Aqua	Denmark	jostou@aqua.dtu.dk
Jose Rodriguez	INSTITUTO ESPAÑOL DE OCEANO-GRAFÍA	Spain	jose.rodriguez@ieo.es
Josefine Egekvist (co-chair)	DTU Aqua	Denmark	jsv@aqua.dtu.dk
Julien Rodriguez	Ifremer	France	Julien.Rodriguez@ifremer.fr
Lara Salvany	ICES	Denmark	lara.salvany@ices.dk
Luís Manuel Correia Ferreira Bentes	CCMAR - Centre of Marine Sciences, University of Algarve	Portugal	lbentes@ualg.pt
Maria Mar Gil	CSIC-IMEDEA	Spain	mmgil@imedea.uib-csic.es
Maria Mateo	AZTI	Spain	mmateo@azti.es

Marta Mega Rufino (co-chair)	IPMA	Portugal	marta.rufino@ipma.pt
Michael Kriegl	Leibniz Centre for Tropical Marine Research	Germany	michael.kriegl@leibniz-zmt.de
Neil Campbell	Marine Lab Scotland	UK	neil.campbell@gov.scot
Nuno Henrique Henriques	CCMAR - Centre of Marine Sciences, University of Algarve	Portugal	nhhenriques@ualg.pt
Oliver Tully	Marine Institute	Ireland	oliver.tully@marine.ie
Patricia Breen	Marine Institute	Ireland	patricia.breen@marine.ie
Patrik Jonsson	SLU – Department of Aquatic research (SLU Aqua)	Sweden	Patrik.jonsson@slu.se
Pedro Leitão	IPMA	Portugal	pedro.leitao@ipma.pt
Roi Martinez	CEFAS	UK	roi.martinez@cefas.co.uk
Sara Palma Pedraza	Marine Institute	Ireland	Sara.PalmaPedraza@Marine.ie
Stefanos Kavadas	HCMR	Greece	stefanos@hcmr.gr
Tania Mendo	University of St. Andrews	UK	Tania.Mendo@st-andrews.ac.uk
Tereza Fonseca	Enviromental Defense Fund (EDF)	Portugal	terezacp@yahoo.com
Torsten Schulze	Thünen Institute of Sea Fisheries	Germany	Torsten.Schulze@thuenen.de

Annex 2: WKSSFGEO resolution

Workshop on Geo-Spatial Data for Small-Scale Fisheries (WKSSFGEO), chaired by Marta Rufino, Portugal; and Josefine Egekvist, Denmark, will be established and will meet in Lisbon, Portugal, 29 November – 3 December 2021 to:

a) Discuss and apply methods for identifying trips/hauls in small-scale fisheries, including passive gears, using high resolution geo-spatial data. Participants need to bring their own data for case-studies to develop best practices and common methodologies;
b) Based on the best practices identified, develop an R-script that can be used as a template for analysis of geo-spatial for small-scale fisheries;
c) Evaluate how the use of high resolution geo-spatial data improve effort estimates and can help quantify the extent of small-scale fisheries.

WKSSFGEO will report by 15 January 2022 (via HAPISG) for the attention of the ACOM and SCICOM.

Supporting information

Scientific justification	In relation to spatial data within the EU, VMS are available for vessels larger than or equal to 12 m since 2012, with a maximum ping rate of 2 hours. The ICES VMS/logbook data call requests VMS-based spatial data, but is missing information on fishery from vessels that are not carrying VMS. It is identified as a caveat in relation to the data outputs used for ICES Advice (e.g. ADGTRADE) that the small-scale fishery is missing, resulting in an underestimation of the fishing pressure, especially in coastal areas. Some national initiatives have been implemented to obtain spatio-temporal data from vessels < 12 m (e.g. AIS, GPRS trackers), but the methods to deal with this highly temporally resolved data are not harmonized/standardized. Several ICES members, such as the UK, are proposing the use of appropiate vessel tracking systems for the whole inshore fleet (DEFRA< 2018; Marine Scotland, 2019). Additionally, at the EU level current negotiations between the EU Commision, Parliament and Council are underway for the tracking on small scale fishing vessels by all Member States (P9_TA(2021)0076). Therefore, it is necessary to produce standardised protocols to identify fishing trips and infer fishing activities in SSF. With regards to passive gears, no matter the type of vessel, measures of fishing effort are often missing. Two types of effort is requested in the ICES RDBES Effort statistics: number of hours the vessel is conducting fishing and handling related activity and the soaking time. The workshop will test the use of highly resolved spatio-temporal data to identify setting and hauling events during fishing trips to infer other measures of effort (such as number of pots/traps, length of the net and/or gear soak time). The workshop will aim to discuss and develop standard procedures for identifying trips/hauls in SSF using geo-spatial data that can be compatible with VMS derived outputs. Participants will bring their own data for the case-studies. Namely, the workshop participants will explore the possibility of identifying the setting/hauling of passive gears. It will also be explored how different criteria applied affect the identification of fishing trips/hauls (e.g. through sensitivity analysis). The output will be an R-script for working with geo-spatial data for SSF.
Resource requirements	The research programmes which provide the main input to this group are already underway, and resources are already committed. The additional resources required to undertake additional activities in the framework of this group are negligible.
Participants	The group will be attended by members of WGSFD, WGCATCH and other invited experts.

Secretariat facilities	Standard EG support.
Financial	Funding will be requested for on site review.
Linkages to advisory committees	There are no obvious direct linkages with the advisory committees.
Linkages to other committees or groups	WGSFD, WGCATCH, WGBYC, WGTIFD, SCICOM, HAPISG
Linkages to other organizations	EU Regional Coordination Groups Intersessional Subgroups on Small-scale fisheries and Metier and transversal variable issues.

Annex 3: 'Create_fishing_trip' function

Create_fishing_trip_galdelli

Jacopo Pulcinella, Alessandro Galdelli, Anna Nora Tassetti
10/12/2021

Aims and description of the *create_fishing_trip* function

The results presented below report on the application of the function *create_fishing_trip*. It is part of the Italian workflow that was originally developed for t-AIS data and large scale fishery, and mainly for trawl fisheries. The function defines trips as sequences of points broadcasted by a vessel, from the time it leaves the port until it returns, and stores related information on the port of departure and of arrival (i.e., harbour name, country and statistical area). According to the discussion held during the WGSSFGEO such definition of fishing trip works better for mobile gears, while it might be incorrect for passive gears since it does not require at least one fishing event during the trip.

Since gaps in t-AIS data (i.e., loss of signal of at least 30 minutes) can hamper the identification of the departure and arrival ports, a recovery function was internally applied to join consecutive trips where the departure/arrival port was too far to be assigned. In order to join consecutive trips the function overlays the ending/starting points with the coastal_ban_zone, compares ids between consecutive ending/starting points, compares timestamps and forces a starting and ending port for each trip. In particular, fishing trips are joined and the nearest port is assigned if ending and starting points are consecutive, have a temporal distance shorter than 24 h and at least one is outside the coastal_ban_zone. At the end of the recovery process, for trips that still miss departure and/or arrival ports, the internal function closest_port_recovery is used to force port assignment under other conditions.

The original method used in the analysis was released by Galdelli at al., 2019 at https://doi.org/10.5281/zenodo.4761890. To run the workflow with different AIS or GPS data, a modified version was released at https://github.com/ices-eg/WKSSFGEO/tree/dev_branch/R-dev/galdelli_pulcinella_tassetti, together with the following and required additional layers:

- ICES and Mediterranean ports. It was obtained merging the Mediterranean ports (https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tassetti/data/med_harb_gsa.rData) with those available in the WGSSFGEO repository (https://github.com/ices-eg/WKSSFGEO/raw/main/data/harbours.rds).
- 3 nm buffer of the line coast of the Northern Europe (https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tassetti/data/coastal_ban_zone_nord_eu.rData)
- Trawling ban for the Mediterranean Sea, covering the minimum distances and depths for the use of towed gears, as defined by Article 13 of EU Council Regulation 1967/2006. It is bounded by the 3 nautical-mile line or by the 50 m isobath where that depth is reached at a shorter distance from the coast (https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tassetti/data/coastal_ban_zone_med.rData).

We applied the *create_fishing_trip function* and the *Worflow_jepol.R* (available at https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/jepol/Workflow_jepol.R) to the same

sample of AIS data available at https://github.com/ices-eg/WKSSFGEO/tree/main/data-examples. Resulting trips were compared in terms of numbers and duration.

Data analysis

Application of the function *create_fishing_trip* from Galdelli *et al.*, 2019

We loaded all the functions required from the source file and general settings for the workflow.

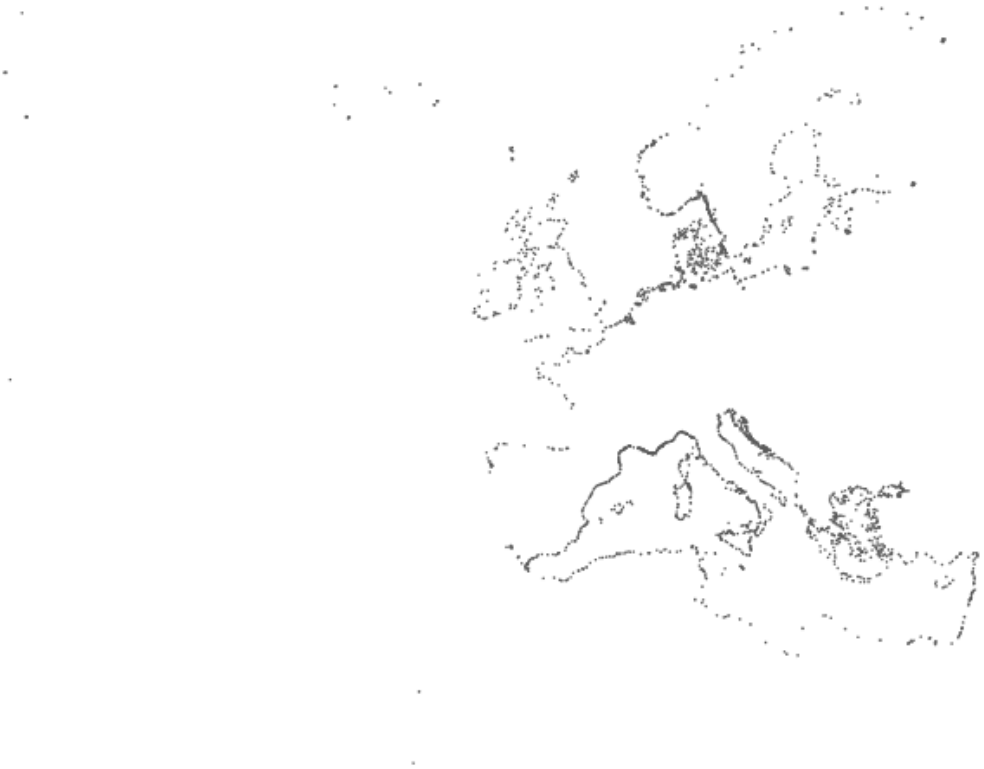
```
# general settings
install.missing.packages = F
wgs="+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0"
file_centroids<-"data/centroids.csv" # for classification
file_parameters<-"data/parameters.csv" # generic for different func-
tions
# source functions
devtools::source_url("https://raw.githubusercontent.com/ices-eg/WKSS-
FGEO/dev_branch/R-dev/galdelli_pulcinella_tassetti/R/global_functions.R")
# load parameters
centroids=inport_parameters(file.path(file_parameters), file.path(file_cen-
troids))[[2]]
pars=inport_parameters(file.path(file_parameters), file.path(file_cen-
troids))[[1]]
```

From the WGSSFGEO repository we selected a data sample from a single vessel (EX_1) and renamed the columns according to the Galdelli method.

```
# Download example file
dat <- read.csv(url("https://raw.githubusercontent.com/ices-eg/WKSS-
FGEO/main/data-examples/example_data_AIS.csv"))

head(dat)
## vessel_id time_stamp lon lat speed course gear behaviour
## 1 EX_1 2018-10-06 06:37:45 12.53560 55.95343 2.315  35.45 Other
## 2 EX_1 2018-10-06 06:37:55 12.53575 55.95356 3.788  34.93 Other
## 3 EX_1 2018-10-06 06:38:05 12.53599 55.95370 4.403  58.42 Other
## 4 EX_1 2018-10-06 06:38:15 12.53634 55.95371 4.298  99.60 Other
## 5 EX_1 2018-10-06 06:38:25 12.53668 55.95366 4.411 104.94 Other
## 6 EX_1 2018-10-06 06:38:35 12.53704 55.95363 4.458  89.72 Other
vessels <-"EX_1"
dat_1v = dat %>%
  filter(vessel_id == vessels)
all_dat_ita_format<-as.data.frame(dat_1v) %>%
  mutate(datetime = time_stamp, MMSI = vessel_id, longitude = lon, latitude
= lat)
all_dat_ita_format<-all_dat_ita_format[,c("MMSI", "datetime", "longitude",
"latitude", "speed")] # select fields of interest
```

- We merged the Mediterranean and Northern European ports in a single layer as required by the algorithm. As shown in the figure below, Portuguese harbours are missing.

```
# Mediterranean harbours

med_harb = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/med_harb_gsa.rData?raw=true")) %>%
  st_set_crs(4326) %>%
  mutate(SI_HARB = 1)

# Northern Harbours
nord_harb = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/main/data/harbours.rds?raw=true")) %>%
  st_set_crs(4326)
nord_harb_edit = nord_harb %>%
  mutate(Country = 1:nrow(nord_harb), harbour = 1:nrow(nord_harb), area =
1:nrow(nord_harb))

# combine harbours
all_harb = rbind(med_harb, nord_harb_edit)

# the file was exported in the link reported below
# import list of ports
ports = readRDS(url("https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-
dev/galdelli_pulcinella_tassetti/data/ices_med_harbours.rData?raw=true"))
ports <- ports %>%
  dplyr:::rename("GSA" = "area")
st_crs(ports) <- 4326
port_buf<-st_buffer(ports, 0.001) # create a buffer
st_crs(port_buf) <- 4326   # set crs

ggplot() +
  geom_sf(data = ports) +
  theme_void()
```



- The coastal ban polygon. Two coastal ban layers are available: the Mediterranean, created by the Italian teams and the northern European ban, created ad hoc for the WGSSFGEO using Qgis.

```
# Mediterranean ban
med_ban = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/coastal_ban_zone_med.rData?raw=true")) %>%
  st_set_crs(4326)
# Northern EU ban
nord_ban = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/coastal_ban_zone_nord_eu.rData?raw=true")) %>%
  st_set_crs(4326)
```

We considered the Northern European ban zone, according to the spatial extension of the vessel position data to process).

```
all_dat_ita_format_1v =all_dat_ita_format %>%
  filter(MMSI == vessels)
head(all_dat_ita_format_1v)
##   MMSI            datetime longitude latitude speed
## 1 EX_1 2018-10-06 06:37:45  12.53560 55.95343 2.315
## 2 EX_1 2018-10-06 06:37:55  12.53575 55.95356 3.788
## 3 EX_1 2018-10-06 06:38:05  12.53599 55.95370 4.403
## 4 EX_1 2018-10-06 06:38:15  12.53634 55.95371 4.298
## 5 EX_1 2018-10-06 06:38:25  12.53668 55.95366 4.411
## 6 EX_1 2018-10-06 06:38:35  12.53704 55.95363 4.458
# Download example file
coastal_ban_zone = nord_ban

# Fishing trip
dat_trip=create_fishing_trip(data=all_dat_ita_format_1v,
                             ports=ports,
                             ports_buffer=port_buf,
                             coastal_ban_zone=nord_ban)
##
================================================================================
====
##  trip create complete!!!
##  Recovery trip....complete!!!
dat_trip
##   MMSI trip     start_timestamp       end_timestamp departure arrival
## 1 EX_1    1 2018-10-06 06:39:35 2018-10-06 09:45:01        86      86
## 2 EX_1    2 2020-10-10 09:05:11 2020-10-10 10:23:49        86      86
## 3 EX_1    3 2020-10-11 06:42:26 2020-10-11 12:20:38        86      86
```

Application jepol workflow

An R script was available to reproduce the jepol's workflow with its own orginal data at *https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/jepol/Workflow_jepol.R*. We run the code available at *https://github.com/ices-eg/WKSSFGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tassetti/R/Workflow_jepol_edit_jp.R* and saved the result in the folder *https://github.com/ices-eg/WKSSFGEO/tree/dev_branch/R-dev/galdelli_pulcinella_tassetti/results*.

We applied the function *define_trips_pol* that resulted in a dataset of pings with the trip label *(jepol_pp_trip)* and in a trip table reporting the scheduling information of each trip (*schedule*).

```
# ping with trip label
jepol_pp_trip = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/jepol_pp_trip.rData?raw=true"))

# trip table
schedule = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/jepol_trips.rData?raw=true"))
```

```
head(schedule %>%
       filter(vessel_id == vessels))
##    vessel_id trip_id              depart              return duration_hrs
## 1:      EX_1  EX_1_1 2018-10-06 06:39:46 2018-10-06 09:45:01     3.087500
## 2:      EX_1  EX_1_2 2020-10-10 09:09:12 2020-10-10 10:23:29     1.238056
## 3:      EX_1  EX_1_3 2020-10-11 06:42:26 2020-10-11 12:20:17     5.630833
```

Comparison

As reported from the examples above, the functions *create_fishing_trips and define_trip_pol* gave consistent resulting trip tables for vessel EX_1.

The *create_fishing_trip* function was then applied locally to the whole dataset. Since the function requires about 5 minutes, we ran the analysis locally and we loaded the results in the link used in the code.

```
# results of create_fishing_trips, arrange columns to be consistent with the
preaviuos dataset
# load the result of the local run on the whole dataset
# Not run - it take ~5 mins, load files from git
# dat = as.data.frame(dat) %>%
#   mutate(datetime = time_stamp, MMSI = vessel_id, longitude = lon, lati-
tude = lat)
# dat<-dat[,c("MMSI", "datetime", "longitude", "latitude", "speed")]
# dat_trip=create_fishing_trip(data=dat,
#                              ports=ports,
#                              ports_buffer=port_buf,
#                              coastal_ban_zone=coastal_ban_zone)
# dat_with_trip=assign_trip(data=dat,
#                           trip_table=dat_trip)
# pings with trip labels
dat_pp_trip = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/galdelli_pp_trips.rData?raw=true"))
# trips table
dat_trip = readRDS(url("https://github.com/ices-eg/WKSS-
FGEO/blob/dev_branch/R-dev/galdelli_pulcinella_tas-
setti/data/galdelli_trips.rData?raw=true")) %>%
  mutate(duration_hrs = as.numeric(difftime(end_timestamp, start_timestamp,
units = "hours")))
```

For each vessel, we calculated the number of trips identified by each method (ntrip), their cumulative and average duration (duration_cum and duration_mean, respectively), and evaluated the difference between these metrics (see *trip_stat* table). We summarized this information in a new table that reports for each vessel and variable the percentage difference between the 2 methods.

```
# results of jepol workflow
trip_stat_jepol = schedule %>%
```

```
  group_by(vessel_id) %>%
  dplyr:::summarise(ntrip = length(trip_id),
                    duration_mean = round(mean(duration_hrs),2),
                    duration_cum = round(sum(duration_hrs),2)) %>%
  mutate(source = "jepol") %>%
  melt(id.var = c("vessel_id", "source"))

# results of galdelli workflow
trip_stat_galdelli = dat_trip %>%
  group_by(MMSI) %>%
  dplyr:::summarise(ntrip = length(trip),
                    duration_mean = round(mean(duration_hrs),2),
                    duration_cum = round(sum(duration_hrs),2)) %>%
  mutate(source = "galdelli") %>%
  mutate(vessel_id = MMSI) %>% dplyr:::select(-MMSI) %>%
  melt(id.var = c("vessel_id", "source"))

# combine and arrange datasets
trip_stat = rbind(trip_stat_jepol, trip_stat_galdelli) %>%
  dcast(vessel_id + variable ~ source, value.var = "value", fun.aggregate =
sum) %>%
  mutate(difference = round(galdelli - jepol, 2),
         difference_perc = round((1 - jepol/galdelli)*100, 2))
trip_stat
##    vessel_id      variable galdelli   jepol difference difference_perc
## 1       EX_1         ntrip     3.00    3.00       0.00            0.00
## 2       EX_1 duration_mean     3.36    3.32       0.04            1.19
## 3       EX_1  duration_cum    10.07    9.96       0.11            1.09
## 4      EX_10         ntrip    14.00   12.00       2.00           14.29
## 5      EX_10 duration_mean     8.09   10.09      -2.00          -24.72
## 6      EX_10  duration_cum   113.24  121.11      -7.87           -6.95
## 7      EX_11         ntrip    26.00   26.00       0.00            0.00
## 8      EX_11 duration_mean   102.54   12.15      90.39           88.15
## 9      EX_11  duration_cum  2666.08  316.00    2350.08           88.15
## 10     EX_12         ntrip     2.00    2.00       0.00            0.00
## 11     EX_12 duration_mean    32.22   32.11       0.11            0.34
## 12     EX_12  duration_cum    64.44   64.21       0.23            0.36
## 13     EX_13         ntrip     2.00    2.00       0.00            0.00
## 14     EX_13 duration_mean    64.44   64.29       0.15            0.23
## 15     EX_13  duration_cum   128.88  128.58       0.30            0.23
## 16     EX_14         ntrip     6.00    6.00       0.00            0.00
## 17     EX_14 duration_mean    16.53   16.94      -0.41           -2.48
## 18     EX_14  duration_cum    99.17  101.64      -2.47           -2.49
## 19      EX_2         ntrip     4.00    4.00       0.00            0.00
## 20      EX_2 duration_mean     4.97    4.97       0.00            0.00
## 21      EX_2  duration_cum    19.89   19.87       0.02            0.10
## 22      EX_3         ntrip    18.00   17.00       1.00            5.56
## 23      EX_3 duration_mean     6.42   47.31     -40.89         -636.92
## 24      EX_3  duration_cum   115.56  804.25    -688.69         -595.96
## 25      EX_4         ntrip     6.00    6.00       0.00            0.00
## 26      EX_4 duration_mean     3.06    3.35      -0.29           -9.48
## 27      EX_4  duration_cum    18.33   20.13      -1.80           -9.82
## 28      EX_6         ntrip     3.00    3.00       0.00            0.00
```

```
## 29      EX_6 duration_mean    25.68   16.24       9.44          36.76
## 30      EX_6  duration_cum    77.04   48.72      28.32          36.76
## 31      EX_7         ntrip     8.00    6.00       2.00          25.00
## 32      EX_7 duration_mean   105.91 1053.88    -947.97        -895.07
## 33      EX_7  duration_cum   847.25 6323.30   -5476.05        -646.33
## 34      EX_9         ntrip     2.00    3.00      -1.00         -50.00
## 35      EX_9 duration_mean    28.82   14.06      14.76          51.21
## 36      EX_9  duration_cum    57.64   42.17      15.47          26.84
```

There are 8 out of 12 vessels with a perfect correspondence in terms of number of trips identified by both methods.

```
# total vessel
length(unique(trip_stat$vessel_id))
## [1] 12
#correct
length(which(trip_stat$difference == 0 & trip_stat$variable == "ntrip"))
## [1] 8
# correct vessel
trip_stat$vessel_id[which(trip_stat$difference == 0 & trip_stat$variable ==
"ntrip")]
## [1] "EX_1"  "EX_11" "EX_12" "EX_13" "EX_14" "EX_2"  "EX_4"  "EX_6"
```

We applied a trip to trip comparison evaluating the overlap between the temporal intervals of the trips identified by both methods. We used the trip table of the jepol workflow as reference and evaluated the correspondence between trips also using the total duration of the identified intervals.

```
xvessels = unique(schedule$vessel_id)
out = NULL
no_trip = NULL
for(i in 1:length(xvessels)){
  #jepol track
  xdat_jepol = schedule %>%
    filter(vessel_id == xvessels[i]) %>%
    mutate(jepol_int = interval(depart, return, tzone = "UTC"))

  # formatting galdelli data
  xdat_galdelli = dat_trip %>%
    filter(MMSI == xvessels[i]) %>%
    mutate(galdelli_int = interval(start_timestamp, end_timestamp, tzone =
"UTC"))

  # extract only galdelli trips with temporal overlap with jepol trip
  for(j in 1:nrow(xdat_jepol)){
    xtrip = as.data.frame(xdat_jepol)[j,] %>%
      dplyr:::select(vessel_id, trip_id, jeopl_dur = duration_hrs,
jepol_int) %>%
      ungroup()
    xtrip_merge =  merge(xtrip,
                         xdat_galdelli %>%
```

```
                            ungroup() %>%
                            dplyr:::rename("vessel_id" = "MMSI") %>%
                            dplyr:::select(vessel_id, galdelli_int,
galdelli_dur = duration_hrs, trip),
                          by = c("vessel_id"))
    xtrip_intersect = xtrip_merge[which(as.logical(inter-
sect(xtrip_merge$galdelli_int, xtrip_merge$jepol_int))),]
    if(nrow(xtrip_intersect) == 0){
      xxtrip = cbind(xtrip, galdelli_dur = NA, galdelli_int = NA, trip = NA)
      no_trip = rbind(no_trip, xxtrip)
    }else{
      if(nrow(xtrip_intersect) == 1 & (xtrip_intersect$galdelli_dur >=
xtrip_intersect$jeopl_dur - xtrip_intersect$jeopl_dur*0.1 &
                                       xtrip_intersect$galdelli_dur <=
xtrip_intersect$jeopl_dur + xtrip_intersect$jeopl_dur*0.1)){
        next()
      }else{
        out = rbind(out, xtrip_intersect)
      }
    }
  }
}
out = out %>% dplyr:::select(-vessel_id, jepol_int, galdelli_int, trip_id,
trip)
out = out[,c("jepol_int", "galdelli_int", "trip_id", "trip")]
```

The following table reports the trips defined with jepol workflow with no single correspondences or difference in the duration with respect to the trips identified using the galdelli workflow. In some cases, the *create_fishing_trips* identified more trips than those defined by the jepol workflow. This happened for example for vessel EX_10 for which 3 trips were identified by the *define_trips_pol* function (EX_10_4, EX_10_5, EX_10_6) while only 2 trips were defined by the *create_fishing_trips* function (trip 13, 15). In other cases, the *create_fishing_trip*s was not able to split trips correctly, as happened for vessels EX_11 and EX_7 whose estimated trips lasted more than 2000 and 800 hours respectively. On the contrary, the durations of the trips EX_3_15 and EX_7_6 identified by *define_trips_pol* function, respectively 688 and 6301 hours, reveal some errors in their definition. According to these issues, the number of trips identified for vessels EX_3, EX_7, EX_9, and EX_10 by the *define_trips_pol* did not correspond with those identified with the *create_fishing_trips* function.

```
out
##                                          jepol_int
## 3  2013-10-05 05:18:12 UTC--2013-10-05 11:44:25 UTC
## 31 2013-10-05 11:45:26 UTC--2013-10-07 18:02:37 UTC
## 4  2013-10-05 11:45:26 UTC--2013-10-07 18:02:37 UTC
## 41 2013-10-08 02:28:55 UTC--2013-10-08 10:55:36 UTC
## 5  2013-05-29 04:06:44 UTC--2013-05-30 16:50:01 UTC
## 6  2013-09-07 03:20:02 UTC--2013-09-07 16:51:02 UTC
## 14 2020-09-20 14:33:46 UTC--2020-09-20 19:45:12 UTC
## 15 2020-09-20 23:09:10 UTC--2020-10-19 15:22:51 UTC
## 61 2016-12-04 12:21:01 UTC--2016-12-04 17:00:48 UTC
## 1  2015-09-09 02:04:55 UTC--2015-09-09 09:04:30 UTC
## 62 2015-07-26 09:01:26 UTC--2016-04-13 22:27:46 UTC
```

```
## 8  2016-05-03 22:08:30 UTC--2016-05-03 23:59:50 UTC
## 11 2018-12-30 02:40:41 UTC--2018-12-30 18:37:37 UTC
##                                       galdelli_int trip_id trip jepol_dur
## 3  2013-10-05 05:18:12 UTC--2013-10-06 12:30:15 UTC EX_10_4   13       6.4
## 31 2013-10-05 05:18:12 UTC--2013-10-06 12:30:15 UTC EX_10_5   13      54.3
## 4  2013-10-07 05:30:08 UTC--2013-10-08 10:55:57 UTC EX_10_5   15      54.3
## 41 2013-10-07 05:30:08 UTC--2013-10-08 10:55:57 UTC EX_10_6   15       8.4
## 5  2013-05-29 04:06:35 UTC--2013-05-29 04:16:06 UTC EX_11_5    7      36.7
## 6  2013-05-30 16:50:01 UTC--2013-09-07 16:51:22 UTC EX_11_6    8      13.5
## 14 2020-09-20 16:21:48 UTC--2020-09-20 19:45:32 UTC EX_3_14   91       5.2
## 15 2020-09-20 23:09:10 UTC--2020-09-20 23:59:59 UTC EX_3_15   92     688.2
## 61 2016-12-04 14:12:36 UTC--2016-12-04 17:01:28 UTC  EX_4_6   69       4.7
## 1  2015-09-07 19:58:03 UTC--2015-09-09 09:07:41 UTC  EX_6_1   56       7.0
## 62 2015-07-26 09:01:16 UTC--2015-07-26 23:59:59 UTC  EX_7_6   55    6301.4
## 8  2016-05-03 22:08:30 UTC--2016-06-06 17:33:10 UTC  EX_7_8   62       1.9
## 11 2018-12-28 20:57:59 UTC--2018-12-30 18:39:28 UTC  EX_9_2   75      15.9
##    galdelli_dur
## 3          31.2
## 31         31.2
## 4          29.4
## 41         29.4
## 5           0.2
## 6        2400.0
## 14          3.4
## 15          0.8
## 61          2.8
## 1          37.2
## 62         15.0
## 8         811.4
## 11         45.7
```

Conclusions

The application of the *create_fishing_trips* requires data to be formatted as indicated in the Galdelli 2019 workflow. (i.e., changing the column names according to the name of the dataset used in the original workflow). Further, for data that spatially extend outside the Mediterranean sea, it is required also an update of some input spatial layers.

The comparison between the results of the *define_trip_pol* with those of the *create_fishing_trips* revealed errors in both the procedures. At this stage an in-depth analysis of the errors was not carried out. However, the list of vessels with erroneous data was identified, promoting future comparison.

The *create_fishing_trip* function does not require interpolation and it works on preprocessed (i.e., cleaned) pings. However, using the spatial layer of the harbours and the polygon of the ban, the function is able to reconstruct trips also when data gaps are present and when final positions are not in harbours.

Since *a priori* knowledge of the duration of a trip is not available, the *create_fishing_trip* function does not require a specific definition of the minimum and maximum duration. The function reads each ping of a vessel using a sliding window of a certain size (from 3 to 6 pings) and han-

dles several features that occur (i.e. data gap, harbour assignment, or trip joining). On the contrary, the *define_trip_pol* uses a minimum and maximum duration to split the trips. However, applying this function with default parameters, it estimated trips with duration greater than the maximum allowed.

Further, it is relevant to note that the *create_fishing_trip* function requires about 5 minutes to process the whole dataset, while the *define_trip_po*l function used in the jepol workflow is faster (i.e., less than a minute). This probably due to the nature of the data for which the methods were developed. In t-AIS data, as those for which the *create_fishing_trips* function was originally designed, the presence of numerous gaps requires a ping to ping analysis to define the end of a trip and the starting of the subsequent one.

Finally, the results presented in this work refer to a single application of the create_fishing_trip function. We tried also to apply the method to the data available in at *https://github.com/ices-eg/WKSSFGEO/blob/main/data-examples/example_data_GPS.csv*, but in this area (i.e., Portugal) the *create_fishing_trips* function does not apply because harbours are not available.

Annex 4: On-site expert review of WKSSFGEO

Reviewer: Tania Mendo (St. Andrews University)
Working Group: WGSFD

1. Overview

During the ICES WGSFD meeting 7–11 June 2021, several aspects of geo-spatial data analysis were discussed. It was noted that while there are clear methods to infer fishing trips and fishing activities from Vessel Monitoring Systems (VMS), which generally transmit positional information every 2 hours, these are not generally suited for data with higher temporal resolution (scale of seconds to minutes) collected by other devices such as Automated Identification Systems or other trackers. These temporal resolutions are important especially for Small-Scale Fisheries (SSF) where fishing activities are of shorter durations that in large scale fisheries and therefore need higher frequencies to identify for example hauling events (e.g. Mendo *et al.*, 2019). Having data at higher temporal resolutions results in specific issues and challenges that were shared by several participants. The need for a dedicated workshop on geo-spatial data analysis for highly resolved spatial data was identified and considered relevant not only for WGSFD but also to WGCATCH, WGBYC and WGTIFD. As a result, a Workshop on Geo-spatial data for Small-Scale Fisheries (WKSSFGEO) was organized by ICES and chaired by Marta Rufino (IPMA) and Josefine Egekvist (DTU). This review document is structured according to some general remarks, evaluation of achievements in each Term of Reference and recommendations for future work.

2. General remarks

The workshop was set to discuss and develop standard procedures for identifying trips/hauls using highly resolved geo-spatial data in a collaborative manner. It was envisioned that participants would bring their own data and test several approaches to identify trips and fishing events. This would improve our understanding of advantages, disadvantages, and limitations of different methods to estimate fishing effort and establish best practices.

The workshop's terms of reference were as follows:

ToR a: Discuss and apply methods for identifying trips/hauls in small-scale fisheries, including passive gears, using high resolution geo-spatial data.

ToR b: Based on the best practices identified, develop an R-script that can be used as a template for analysis of geo-spatial for small-scale fisheries

ToR c: Evaluate how the use of high resolution geo-spatial data improve effort estimates and can help quantify the extent of small-scale fisheries.

WKSSFGEO was conducted in a hybrid format (physical and remote) and hosted by IPMA, Lisbon, Portugal.

3. Addressing the ToRs

ToR a: Discuss and apply methods for identifying trips/hauls in small-scale fisheries, including passive gears, using high resolution geo-spatial data.

Eight teams from 8 different countries provided presentations summarising current techniques and methods used, main issues encountered and best practices to identify trips and fishing events. Several common issues were identified and discussed during these presentations. Participants were then divided into subgroups (2 in person and 3 remote) to discuss methods, issues,

best practices for identifying trips and hauls in SSF, including passive gears using high resolution geospatial data.

The vast number of issues and approaches identified prompted a compilation of all different methods and issues encountered when analysing this type of data had to be conducted before being able to produce a standardised R script. A subgroup was established to collate all elements of this discussion into 5 sections: pre-processing, identification of trips, inferring fishing activities, model validation and effort variables. Another subgroup was established to collate different code available and used by different participants into the GitHub repository (see below).

Further steps: Due to the vast amount of specific issues and techniques used by participants, more emphasis was given to the discussion and compilation of methods than on applying these techniques on datasets brought by participants. The different methods now compiled should be compared using different case studies from participants in a future workshop.

ToR b: Based on the best practices identified, develop an R-script that can be used as a template for analysis of geo-spatial for small-scale fisheries

A subgroup focused on assembling the different approaches used by participants to identify fishing trips and fishing events and R code was developed and compiled into the Github repository https://github.com/ices-eg/WKSSFGEO

This repository currently hosts a series of codes to apply different techniques used currently by workshop participants, and ranges from methods to identify the start and end of a fishing trip using a polygons approach, to machine learning techniques to infer fishing events.

Further steps: A finalised R script with a tentative workflow needs to be developed and tested to develop best practices and a common methodology.

ToR c: Evaluate how the use of high resolution geo-spatial data improve effort estimates and can help quantify the extent of small-scale fisheries.

The effort subgroup focused on first evaluating if and how using geospatial data can inform reporting requirements for the Data Collection Framework (Effort variables defined in EU Map 2021/1167). The subgroup focused on evaluating the definitions for both mobile and static gear, the latter more widely used in SSF. It was revealed that geospatial data can inform several effort variables for the DCF and improve the accuracy of estimates of fishing time at a scale of minutes. Several definitions, however, are more appropriate for mobile gears, and need to be clearly defined for static gears, where gear is usually deployed during a trip and then recovered in a different trip. Under the current EU Map definitions, some variables become redundant for static gears, for example hours fished and soak time are the same. Several innovative ways to describe fishing effort are presented as case studies in the report. The difficulty of identifying when gear is being deployed and hauled was discussed as a major constraint to identify soak time in static gears.

Further steps: More work is needed to standardise effort variables specifically for static gears. Methods to identify deployment events in static gears need to be improved to allow for estimation of gear soak time.

4. General recommendations:

Further workshop is recommended to finalise the objectives set out in this first workshop, including:

a) Compile available R code to identify fishing trips and fishing events into R script with a tentative workflow.

b) Tentative workflow and methods to be tested using a range of case studies, with data brought by participants.

c) Further refinement of definitions of fishing effort variables,especially for SSF using static gears.

d) Further development of methods to estimate deployment and hauling events from geospatial data.

e) Further development of methods to estimate soak time need to be developed.