Harmonization and analysis of the Small-Scale Fisheries (SSF) fishing activity in the Atlantic Area (AA)

Authored By: Estanis Mugerza, Almudena Álvarez, Eneko Bachiller, André Carvalho, Arturo Colina, Sebastien Demanèche, Maria del Pino Fernández, Laura García, Lucía García Flórez, Miguel Gaspar, JMS Goncalves, Nuno S. Henriques, Mark James, Tania Mendo, Angela Muench, Paloma León, Antonio Punzón, Ana Ribeiro, Ignacio Sobrino, Inês Sousa, Declan Tobin

[Doc Date goes here]

Harmonization and analysis of the Small-Scale Fisheries (SSF) fishing activity data in the Atlantic Area (AA)

Background

The harmonization and standardization of estimation techniques of the variables related to fishing activity is a challenge across European countries. Under these activity variables, effort calculation is one of the most relevant variables. Moreover, this is even more complicated when it comes to the Small-Scale Fleet (SSF) due to the limitation in the data sources (e.g. official transversal data coming from the EU Control Regulation (EU, 2009)) is used to get these effort estimates. In order to try to harmonise and standardise these methodologies, the STECF has been organising different Workshops (WK) over the years (STECF 2013, STECF 2015, STECF 2017). Experts from different countries were involved in these WKs, trying to adopt some basic principles for fishing effort calculation and also developed a specific R package (fecR) for estimating the fishing effort in a standardized way.

One of the main objectives of CABFishMAN is to obtain a complete dataset on SSF fishing activity in the whole Atlantic Area (AA), using a standardised methodology among all the countries belonging to this AA. With this objective in mind, the same approach as the WKs mentioned above has been used, following the same basic principles and also the same R package developed. However, certain adaptations and assumptions have had to be made to each of the data sets coming from each Member State (MS) in order to use the aforementioned R package.

This document describes the methodology used, the data sources and also the adaptations and assumptions that have had to be made in order to obtain the effort estimates at the ICES rectangle spatial resolution level. This allows to see in a transparent way how this complete dataset on SSF fishing activity in the AA has been obtained. On the other hand, it can also be useful and helpful for other experts working with this type of dataset to see how certain variables (e.g. trip metiers allocation, ICES statistical rectangle by trip etc.) have been estimated in the absence of this information in the transversal data used.

Material and Methods

The methodology used to calculate the fishing effort was the fecR R package developed with this aim. In Annex 1 the document developed by Finlay Scott, Nuno Prista and Thomas Reilly, provides the details about how this package calculate the fishing effort.

Regarding the data sources used, transversal data collected under the EU Control Regulation has been used. This means, logbooks and First Sale Notes (FSN). Logbooks information was used for vessels equal or above to 10m LOA, and for vessels below 10m, FSN.

The fecR R package, requires adapting the variables collected under these transversal data to the requested formats (Annex 2). In addition, some of the variables required to calculate the effort estimates are not available in the transversal data (e.g. ICES rectangle in the case of the FSN).

Because of this, certain assumptions had to be made with the data. In the following section, the main gaps that were found in each MS datasets, and how these deficiencies were addressed, is described.

I. Portugal

To analyse the distribution of the SSF along the Portuguese mainland waters we relied on two different sources of data: 1, the EU fleet registry database, and 2, from daily sale notes of the SSF vessels. From the EU fleet registry we collected information about the characteristics and licenses of the fishing vessels, such as LOA, engine power in kW, gross tonnage, and fishing gears licences. Sales notes provided information on the ID of the vessel, the date of the sale of the catch, the location where the sale was done, the captured species, the weight of the captured species and the monetary value for each of the captured species. From the raw information provided by these two data sources, we were able to immediately extract 5 of the 15 variables required to run the effort calculation procedure by the fecR package. These were: vessel ID, LOA, gross tonnage, engine power and the economic zone.

Due to the lack of further information within the two data sources, to obtain the missing 10 variables it was required to treat the data based on some assumptions that are described below:

For the Trip ID, since dealing with SSF, it was assumed that each fishing vessel performs no more than one fishing trip per day. Therefore, each sale/landing event is a result of one fishing trip. To create the variable “Trip ID” we simply merged the vessel ID with its corresponding landing date. Example: PRT012345678_04_01_2015.

Since the Portuguese SSF is not required to have logbooks (vessels below 10m), it had to be assumed some parameters regarding the characteristics of the fishing trips. So, for the variables “departure date”, “return date” and “fishing date” were assigned the same date as the one on the sale note. This follows the assumption that a SSF vessel departs from the harbour, fishes and returns on the same day it sells its catch. With the absence of data regarding the departure and arrival time of the fishing vessels the values of “00:00” for the departure time and “12:00” for the arrival time were set. These two variables are quite unprecise, since the duration of a fishing trip can either last less or more than 12 hours, depending on many factors such as the type of fishery, the size of the vessel, or even the skipper. But without logbook information or vessel tracking data, there is no way to assign an accurate fishing trip duration with the provided data.

Due to the absence of logbooks, no information on the used gear was available within the two data sources. To assign the variable “gear” to each fishing trip, we had to proceed to cluster analysis on the sales notes to group them based on the species composition of each sale/landing event. For that, first the sales notes were split into the four seasons of the year: Spring, Summer, Autumn and Winter. Then a CLARA algorithm was used to group the daily sales into 40 clusters. This process was separately carried for each year between 2015 to 2019. This process returned a total of 800 clusters which were assigned a fishing gear based on expert judgement approach. This assignment was based on the species composition, the proportion between the weights of the captured species and on the fishing gears each vessel is licensed to operate. Some clusters were assigned the label “Not Known” for the cases where landing compositions generated too much doubt on the type(s) of gear(s) responsible for such landings.

The process of assigning fishing gears to groups of landings/sale events was a very difficult and subjective process. The magnitude of the data, i.e. the number of landing events, associated with the fact that some gears and metiers have a very heterogeneous composition of captured species, namely trammel nets, gillnets and longlines, made the process of assigning a fishing gear to a landing composition very difficult and without much certainty. Additionally, the Portuguese SSF fleet is mainly composed by polyvalent fisheries, meaning that one fishing vessel can operate more than one type of gear within the same fishing trip. This translates not only into more heterogeneity in the patterns of landing compositions, but it also means that more than one type of gear can be responsible for the catches landed/sold by these vessels.

fecR allows to assign value = 0 to the variable “gear mesh size”. Since the data sources did not provide any sort of information about the used gears, we set this value to zero given the lack of information.

Assigning fishing location to the data, at a resolution of ICES rectangles, was another fecR variable that required some assumptions, especially because Portuguese SSF vessels do not use vessel tracking data. The assignment of fishing effort location (ICES rectangle) derived from the landing location within the sales notes data. SSF vessels are only allowed to operate within the port jurisdiction which they are registered and in the contiguous jurisdictions, and within the 12nm from the coastal line. To assign the fishing trips to a given ICES rectangle, first the number of fishing trips of each vessel, were split by the regions corresponding to the port jurisdictions these vessels are authorized to operate in. Then, the number of fishing trips to the ICES rectangles that overlapped these jurisdictions and the 12nm inshore region were assigned.

Such as the approach to assign the fishing gear to the fishing trips, this one is also quite subjective. Fishers tend to have preferred fishing grounds that are influenced by many factors such as the types of habitats, depth or distance from port, for example. Therefore, splitting equally, the number of fishing trips by the jurisdictions a vessel is allowed to operate is a very unprecise way of allocating fishing effort, but this is probably the most suitable approach according to the available data sources for the Portuguese SSF fleet.

II. Spain

The raw information available for the SSF fleet along the Spanish mainland waters was obtained both for the northern fleet, corresponding to the Cantabrian and Galician coasts (i.e., fleet operating within ICES areas 8 and the norther region of 9a) and the southern fleet, for the coast of Cadiz and Huelva (i.e., southern region of ices 9a).

Missing data for either trips and/or vessels was filled based on some assumptions that are described in the following section.

In Spain, vessels under 10 meters are not obliged to fill diary logbooks but only supply the First Sales Notes (FSN). Therefore, data for trips carried out by such vessels were often with missing information on fishing gears and effort. Further, there is no obligation to register the catches (landings) if they are below 50 kg of live-weight equivalent for given species (EU, 2009), which supposed that trips with lower landings might not have been recorded in our datasets.

Under the EU Control Regulation, there are also exceptions for vessels between 12 and 15 meters concerning the installation of VMS devices. If their fishing activity is carried out in the national waters and the trip duration is less than 24 hours, these vessels will be exempt from using Vessel Monitoring Systems (VMS) devices. This way, for all vessels not obliged to use VMS, the only geospatial information recorded was the closest rectangle in the ICES divisions to their corresponding base port.

In some cases, the trip ID was empty, and that was filled by merging the vessel ID with the corresponding landing date.

For those vessels for which the gross tonnage (GT), LOA (m), engine power (kw) or the (main) metier was not defined, a specific review was made in the census (i.e., National Fleet register); all missing information available in the register was incorporated into our dataset filling the corresponding variables. For such vessels with known metier but undetermined mesh sizes, the most common values used by similar vessels by the same fleet were assigned.

For all vessels not providing information on the duration of their trips (i.e., vessels without logbooks), we assigned a duration of the fishing trip from 4a.m. to 4p.m. in all the cases with reported landings. This way, we assumed a 1-day fishing trip for all the smallest fleet segment, which resulted in 1 fishing day and 1 day-at-sea per fishing trip when running the fecR script. For some trips the metier and/or foCatEU5 variables was not available or was erroneous (e.g. GNS_LLS), and the fleet register was not accurate to determine a certain fishing gear. In such cases, only the vessels known by our institute(s) were manually determined, based on the vessel ID and the fishing period (for which a certain fishing activity was assigned, with its corresponding fishing gear; e.g. most of the GNS_LLS were known as gillnetters so that they were re-assigned as 'GNS' accordingly).

All ICES areas and rectangles were depurated and corrected when suggesting typographical errors. The cases with strange values were removed from the dataset (e.g., trips operating in areas out from the operating regions).

For the southern part of the Spanish SSF fleet, the length-range segment assigned to each vessel was corrected, given that vessels were initially defined based on different length-ranges in comparison with that required by WP4 – data formatting requirements. This way, while southern Spanish vessels were first ranged as LOA 0-10 m / 10-12 m / 12-18 m, length-range re-assignment was made to re-allocate all vessels and their corresponding trips into groups by LOA 0-6 m / 6-8 m / 8-10m / 10-12 m / 12-15 m. This re-assignment was made based on the census of the southern Spanish SSF. A re-assignment of length-classes was made to each trip*vessel*landed species, based on the percentages of active vessels registered for each range in the southern harbours within NUTS ES61 (i.e., harbours from Cádiz to Huelva).

III. France

The Northeast Atlantic French SSF fishing activity data used to answer the CABFishMAN WP4 data need (in the context of the Action N°2) are based on the following sources of information: **French fishing fleet register** (vessel characteristic (length overall, kilowatt, gross tonnage, age of the vessel))

Scientific census survey of annual fishing activity calendars1 (by vessels and month: active/inactive vessels, fishing area, métier and supra-region)

European logbooks (over 10m'vessels) and national monthly declarative fishing forms2 (coastal logbooks, less 10m' vessels) (by fishing trip / fishing sequence: total weight of landings by species, dates of the fishing trip (departure and return dates) and of the fishing sequences (day), declared 'vessel' fishing time by fishing sequence, fishing area, gear/gear dimension and mesh size)

Sales note data (landings statistics from auction markets) (total weight and value of landings by species, dates and vessels)

Geolocation data (esp. issued from VMS devices for over 12m' vessels) (calculated/estimated fishing trips and fishing sequences associated with: dates of the

---

[1] **Annual fishing activity calendar survey covers the whole of the reference population** in all the supra-regions where French vessels operated *(French fishing fleet register' vessels (FPC) including overseas fisheries, small-scale coastal fleets also vessels not covered by available control regulation declarative data)*. The survey is conducted by fishing observers *(observers' network of the Ifremer Fisheries Information System)* yearly in France on the basis of preliminary documentation provided by available control regulation declarative data *(fleet register, logbooks, monthly declarative forms, sales note data, geolocalisation data)* and take place every year in the first month of the year on the previous year. It is particularly instructive for the small-scale coastal fisheries, where catches and effort data are often incomplete.
It aims at characterizing each year the inactivity or activity of all the vessels each month of the year and, in the latter case, the metiers practiced *(metier is defined as the use of a gear to target one or several species)* and the main fishing areas with the corresponding range of operation *(distance to the coast of the fishing operation)*. In addition, fishing activity calendar identified each month the main port of exploitation, the number of fishermen on board and the number of days at sea and fishing days. The aim of collecting data about the activity of each vessel is to have a minimum but exhaustive information on the vessels, to have a complete picture of the whole fleet in terms of gears used and fishing activity, at least at a monthly scale.
Such surveys provide information on the part of fishing activity not included in available declarative data *(completeness check)* and also the basis, if necessary, to re-evaluate available fishing activity data estimates (*in case of incomplete data)*. They constitute also an input each year for the typological classifications of vessels by fleet and a description of their metiers which in turn makes also possible the definition of sampling plans to structure the routine data collection actions. They are also used to allocate metiers to each fishing trip and constitute the exhaustive basis for doing estimation based on the complementary on-site sampling data. Finally, some passive gears characteristics information are also collected during the survey with a minimum of 5% of the French fleet surveyed.
Detailed information about the survey could be found in Annex II presenting the fishing activity calendar' questionnaire and in the following document: *ICES CM 2008/K:12 "From fleet census to sampling schemes: an original collection of data on fishing activity for the assessment of the French fisheries." - Patrick BERTHOU, Olivier GUYADER, Emilie LEBLOND, Sébastien DEMANECHE, Fabienne DAURES, Claude MERRIEN, Patrick LESPAGNOL - https://www.ices.dk/sites/pub/CM%20Doccuments/CM-2008/K/K1208.pdf*.

[2] Declarative forms adapted to the special features of the small-scale coastal fisheries. Cf. Annex I. Arrêté du 18 mars 2015 relatif aux obligations déclaratives en matière de pêche maritime (https://www.legifrance.gouv.fr/loda/id/LEGITEXT000030441176/2019-07-16/). See also Annex I for more information.

fishing trip (departure and return dates) and of the fishing sequences (day), spatial (by fishing area) 'vessel' fishing time by fishing sequence).

The definition of the reference fleet population follows the definition of Commission decision 2016/12513 (any vessel registered on 31 December, or which has fished at least one day in the year up to 31 December) in order to have a comprehensive view of the fishing activity applied during the year.

Then for the Northeast Atlantic French SSF, the definition of all the fishing trips with their associated features (dates, fishing area, métier, gear and mesh size, total weight, and value of landings by species) is based on a **cross-validation tool of the different available declarative data: SACROIS.**

SACROIS4 is a cross-validation tool for the fisheries statistics, aiming at providing the best possible fishing statistics data by cross-checking available data from the different declarative control regulation data flows (multiple declarative sources complementary and sometimes inconsistent). The application is crossing information, at the most disaggregated level, from the fishing fleet register, logbooks and coastal logbooks, sales notes data, geolocation data and the scientific census of annual fishing activity calendars, in order to build a dataset compiling the most accurate and complete information for each individual French fleet' fishing trip with its associated features (dates, fishing area, metier, gear and mesh size, total weight and value of landings by species). The application verifies and controls the different sources of data, linking and comparing them, with the aim of displaying validated, adjusted and qualified spatial landings per species and fishing effort data series. The application compiles them into a single, verified and consistency, controlled data flow. The application provides also several quality indicators and evaluates the completeness of the data flows.

---

[3] Commission Implementing Decision (EU) 2016/1251 of 12 July 2016 adopting a multiannual Union programme for the collection, management and use of data in the fisheries and aquaculture sectors for the period 2017-2019. ELI: http://data.europa.eu/eli/dec_impl/2016/1251/oj

[4] https://sih.ifremer.fr/Debarquements-effort-de-peche/Sacrois

(2022) **Sacrois. Un algorithme de croisement de données**. https://archimer.ifremer.fr/doc/00774/88631/

IFREMER SIH (2022). **SACROIS - Algorithme de consolidation des données déclaratives.** IFREMER SIH(IFREMER)

https://doi.org/10.12770/6510e8e0-788d-45ba-9792-3d0585fe1009

SACROIS tools fit then with the needs identified: 1) to have available a single unique fishing activity data flow validated and qualified to answer all the end-user's requirement (asset to produce consistent answer for all the fishing data needs) and 2) compulsory EU regulations (e.g. EC EU Control Reg. 404/20115 (art. 145)).

SACROIS functionalities have been developed in order to estimate the best possible fishing statistics features (fishing effort, gear, gear dimension and mesh size, fishing area etc.) by fishing trip considering all the information available in the declarative data cross validated. The main functionalities are:

- Linking input data flows and infer estimated SACROIS fishing trips,
- Consolidation, validation and adjustment of the estimated spatial *(by fishing area)* 'vessel' fishing time and dates of the SACROIS fishing trips *(departure, return dates and fishing sequences' days),*
- Consolidation, validation and adjustment of the SACROIS fishing trips' fishing area *(incl. EEZ and regulatory boundaries – e.g. UK 6–12nm)*,
- Consolidation, validation and adjustment of the SACROIS fishing trips' landings by species and faunal composition associated,
- Estimation of the value of the SACROIS fishing trips' landings by species,
- Allocation of a single "metier" to each SACROIS fishing trip/fishing sequence.

Each of these SACROIS functionalities are well described in the related documentation detailing the algorithm applied and hypothesis associated (SACROIS has been developed as a modular tool, each of them being described by an algorithm and the hypothesis associated). For example, fishing effort data available in logbooks (over 10m' vessels) or monthly declarative forms (less 10m' vessels) are cross validated, controlled and adjusted considering the existing geolocation data (e.g. issued from the VMS devices) and the scientific census survey of annual fishing activity calendars.

Same type of algorithm is applied to consolidate, validate and eventually adjust the fishing areas associated to each SACROIS fishing trip. The objectives are:

[5] 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.
ELI: http://data.europa.eu/eli/reg_impl/2011/404/oj

- to better spatialize the declarative spatial fishing activity data taking into consideration the existing geolocation data *(esp. issued from VMS devices)*
- to spatialize SACROIS fishing trip for which there are imprecise or no existing spatialized information.

To consolidate the spatialized information available, the main principles retained are the following:

1) **to prioritize the geolocation data** *(which allow calculating high quality and accurate fishing area information deducted from an algorithm*[6] *based on the basic positions (latitude and longitude) of the vessels)* when they are available.
2) **to take into consideration the scientific census survey of annual fishing activity calendars** *(incl. at least monthly spatial information by vessel at ICES rectangle level and sometimes more precise as sub-rectangle level and/or with the corresponding range*[7] *of operation)* **when geolocation data are not available and spatial data from declarative information are missing** *(e.g. SACROIS fishing trips issued only from sales note data with no other declarative data available to complete the information)* **or imprecise** *(e.g. SACROIS fishing trips issued only from logbooks data or monthly declarative forms with either missing fishing area or fishing area declared at a less precise level than ICES rectangle although this is the minimum level asked in logbooks and monthly declarative forms)*

The scientific census survey of annual fishing activity calendars is especially useful to consolidate the SACROIS fishing trips' spatialization (at least at ICES rectangle level) for non-geolocated vessels (e.g. less than 12m' vessels with no VMS devices).

In the same way, a specific **algorithm** is included into SACROIS tool **to estimate the value of landings by species** based on existing sales note data (sometimes directly deducted from them) or estimation of an average price. For some fleet segment, estimated price based on expert knowledge is also used.

A specific **algorithm** is also included into SACROIS **to allocate a single metier to each SACROIS fishing trip/fishing sequence**, based on the dominant landed species (or group of species) in value, the vessel' activity calendar survey and eventually the declared gear (see

[6] Ifremer. Système d'Informations Halieutiques (2021). Algorithme de traitement de données de géolocalisation ALGOPESCA. Note synthétique. https://archimer.ifremer.fr/doc/00682/79405/

[7] Distance to the coast, the following range of operations could be informed depending of the area where the "vessel*month" is operating: "Fluvial, Estuarien" (in inland water), "3 milles" (inside the 3 nautical miles), "3-12 milles" (inside the 3-12 nautical miles), "Côtier" (inside the 12 nautical miles), "Mixte" (inside and outside the 12 nautical miles), "Large" (outside the 12 nautical miles) and "Etranger" (exclusively in foreign area).

detailed methodology explained in 'Anonymous, DCF metier workshop report, 2018'8, Annex 5). The methodology to determine the dominant landed species (or group of species) is based on the raw ordination of the landed species.

In the end, the definition of all the fishing trips of the Northeast Atlantic French SSF fleet with their associated features (dates, fishing area incl. EEZ and regulatory boundaries, gear, gear dimension and mesh size, total weight and value of landings by species) **result from the application of the SACROIS algorithms**. Completeness (evaluated against the exhaustive Ifremer activity survey) and reliability of the declarative data calculated via the SACROIS tools are qualified as good quality and sufficient to produce the reference fishing activity' estimates (capacity, fishing effort and landings) for these fleets.

SACROIS data detail, at the best possible level regarding the existing declarative data, the fishing activity of each vessel' trip (e.g. the quantity and value of the landed species per catch day, landing location/fishing area incl. EEZ and regulatory boundaries and type of gear/mesh size/dimension used …) and constitute then the principal source of information on fishing activities.

The CABFishMAN WP4 data need (in the context of the Action N°2) were then calculated and formatted on the basis of the SACROIS data which are conformed to the requested aggregation (by year, quarter, month, vessel length classes, fishing technique, landing location, fishing área, gear and métier) and covered all the French SSF fishing vessels operating in the Northeast Atlantic.

Quarter and month information have been deducted from the SACROIS fishing trip' return date when landing country have been deducted from the estimated landing location. Spatial information (at ICES rectangle level), gear/métier or landings by species in weight and in value were also directly deducted from the SACROIS data (see above).

[8] See detailed methodology explained in 'Anonymous, DCF metier workshop report, 2018', Annex4, https://datacollection.jrc.ec.europa.eu/documents/10213/891027/2018_Workshop_DCF+Metiers.pdf/6b928c8a-c2ac-4507-840c-98155e0f07d9

Vessel length ranges have been deducted from the fishing fleet register for each vessel.

Fishing effort estimates (number of trips, days at sea and fishing days) have not been directly calculated by using the generic R script (fecR package) as this is not suitable for the SACROIS data format. Therefore, a specific adapted R script has been developed adapted to the SACROIS data format to calculate fishing effort metrics following the different principles and the common joint methodology developed during the 2nd transversal variables workshop9. Consequently, they are completely comparable with the metrics which should have been calculated from the FecR package and therefore with the data calculated by other CABFishMAN members for other countries.

Finally, the **vessel Fleet Segment DCF** also asked in the data format requested, has been calculated vessel by vessel **from the scientific census survey of annual fishing activity calendars** considering their exploitation strategy and annual fishing activity detailed in their fishing activity calendars. In this calculation, it was considered than one vessel could use several gears during the year but belong to only one DCF fleet segment for a given year. To define the principal fishing technique performed by the vessels, the number of "vessel*month" performed by fishing technique during the given year was considered. In case there is no principal (all less than 50% of the total vessel fishing activity) then the vessel was allocated to a polyvalent (active gears, passive gears or combination of the two) DCF fleet segment. More information about the methodology could be found in the working document10 "Previous experiences, tests for application in the French context and recommendations" drafted as an input for the 2nd workshop on a fisheries-based fleet segmentation approach11.

---

[9] Castro Ribeiro, C., Holmes, S., Scott, F., Berkenhagen, J., Demaneche, S., Prista, N., Reis, D., Reilly, T., Andriukaitiene, J., Aquilina, M., Avdič Mravlje, E., Calvo Santos, A., Charilaou, C., Dalskov, J., Davidiuk, I., Diamant, A., Egekvist, J., Elliot, M., Ioannou, M., Jakovleva, I. Kuzebski, E., Ozernaja, O., Pinnelo, D., Thasitis, I., Verlé, K., Vitarnen, J., Wójcik, I..Report of the 2nd Workshop on Transversal Variables. Nicosia, Cyprus. 22-26 February 2016. A DCF ad-hoc workshop. 109pp.EUR 27897; doi 10.2788/042271. 614eeaf9-6b37-4e59-9738-7e96d47e3ec2 (europa.eu)

[10] Demaneche Sebastien, Guyader Olivier, Le Grand Christelle, Merzereaud Mathieu, Vigneau Joel, Quentin Laurent (2022). Alternative approaches to the segmentation of the EU fishing fleets. Workshop II - 28-30th March 2022. Previous experiences, tests for application in the French context and recommendations. PDG-RBE-HISSEO, PDGRBE- EM, PDG-RBE-HMMN-RHPB. https://doi.org/10.13155/89336

[11] Sulanke, Erik & Berkenhagen, Jörg & Sykkö, Anti & Valve, Joonas & Mantziaris, Stamatis & Daniel, Grigoras & Catalin, Paun & Demaneche, Sebastien & Guyader, Olivier & Grand, Christelle & Merzereaud, Mathieu & Vigneau, Joel & Quentin, Laurent & Fernandes, Ana & Cano, Suzana. (2022). Report of the second workshop on an alternative approach to the segmentation of fishing fleets. 10.13140/RG.2.2.15792.23043. July 2022. DOI:10.13140/RG.2.2.15792.23043

Following that, **Northeast Atlantic French SSF fishing activity data for the years 2015 to 2019** have been provided to feed the **CABFishMAN WP4 data need** (in the context of Action N°2) and thereafter the CABFishMAN geotool.

IV. Ireland

Detailed data describing fishing activity of the small-scale fleets of Ireland has been received by BIM from the Marine Institute. The data provided includes fishing activity for all vessels under 15m in length from the logbook, landings declarations and sales notes databases. The variables provided in each of the three databases are listed in table 1. A common vessel identifier is used across the three databases which is not linked to publicly available vessel registers.

In the logbook database describing daily operations variables describing the vessel (ID, provenance, overall length), the date, the location of fishing (ICES division and rectangle), the gear (gear code, mesh size), the species caught and the estimated weight in KG were included. In the landings declarations database describing the fishing trips variables describing the vessel (ID, provenance, overall length), the date, the location of fishing port (port name, ICES division), the gear (gear code, mesh size), the species caught and the landed and estimated weight in KG were included.

In the sales notes database describing the sale of fish variables describing the vessel (ID, provenance, overall length), the landings and sale dates, the location of the fishing port (port name, ICES division), the forms of fish species sold (species name, size category, presentation, freshness) with the quantities sold and prices and value received.

As part of action 2 of WP 4 the following variables were required to be added to each of the three databases: ICES statistical rectangle and metier (level 5 or 6).

(PDF) Report of the second workshop on an alternative approach to the segmentation of fishing fleets (researchgate.net)

Addition of ICES statistical rectangles

ICES statistical rectangles are included in the logbook database which describes daily activity. Given the nature of the landings declarations and sales notes databases this variable is not included as these databases describe volumes of fish landed and sold that can be from multiple rectangles depending on the length of the fishing trip.

The landings declarations database includes the variable Division, describing the ICES division while the sales notes databases includes the variable zone_desc, which describes the FAO fishing area. In order to incorporate statistical rectangle, the fishing port will be used as a proxy of spatial activity, with the rectangle where that port is located in used. In many cases this should reflect the reality of local small-scale vessels fishing in proximity to their home port. This method has been applied to both the landings declarations and the sales notes databases.

Addition of metier

None of the three databases included a variable describing the metier utilised. Metier combines a number of variables that describe the fishing trip such as the gear used, the location of activity and the target species. Utilising the variables already contained within the databases the metier could be estimated. Here, metiers must be described at level 5 or 6. An example of a metier level 5 is OTB_CRU_27.7.e-k where the gear is described first (OTB), followed by the target assemblage of species (CRU representing crustacean fisheries) and finally the location (27.7.e-k representing the Celtic Sea excluding Galway Bay and Porcupine). Metier level 6 includes the gear mesh size additionally. Given the incomplete data here for gear mesh size the metier level 5 is added to the three databases (however it is metier level 4 that is required in this exercise so this is also included).

Logbooks

The logbook database contains the variable Gear which includes 40 different gear codes. These are summarised into 11 main gears.

Each species is assigned a target species code classifying the species as demersal, pelagic, crustacean, and so on, with eight categories defined. To define the target species for each trip a code must be made to identify the unique trip. This is composed of the vessel ID combined with the date. The landings for each trip are assessed and if over 50% of the trip landings are under one of the target species categories, then the trip is assigned that category. The majority of trips did contain over 50% of the landings under one category however a significant number of trips had landings under more than one trip and below the 50% threshold. A number of mixed categories exist describing landings of crustaceans and demersals (MCD), cephalopod and demersal (MCF), pelagic and demersal (MPD) among others.
Location is defined by the ICES rectangle variable included in the database.
The calculated variables for gear and target species are combined in excel to form the metier.

Landings declarations

The landings declarations database contains the variable GearType which includes 40 different gear codes. These are summarised into 11 main gears.

Each species is assigned a target species code classifying the species as demersal, pelagic, crustacean, and so on, with eight categories defined. To define the target species for each trip a code must be made to identify the unique trip. This is composed of the vessel ID combined with the landing date. The landings for each trip are assessed and if over 50% of the trip landings are under one of the target species categories, then the trip is assigned that category. The majority of trips did contain over 50% of the landings under one category however a significant number of trips had landings under more than one trip and below the 50% threshold. A number of mixed categories exist describing landings of crustaceans and demersals (MCD), cephalopod and demersal (MCF), pelagic and demersal (MPD) among others.

Location is defined by the ICES division variable included in the database.
The calculated variables for gear and target species are combined in excel to form the metier.

Sales notes

The sales notes database does not contain any variable that describes the fishing gear utilised as this database is filled in by the buyer of the fish, not the seller. It does include information on who sold the fish but more detailed information on the fishing vessel is not included. To estimate the gear used a code is created combing the vessel ID and the landing date which is then cross-referenced with the Landings Declarations database. The gear code shown in that database is then used in the sales notes, where a match occurs. Where a match does not occur an estimation of the gear is made by analysing the target species category by trip and matching a likely gear to that assemblage. The gear codes are summarised into 11 main gears.

Each species is assigned a target species code classifying the species as demersal, pelagic, crustacean, and so on, with eight categories defined. To define the target species for each trip a code must be made to identify the unique trip. This is composed of the vessel ID combined with the landing date. The landings for each trip are assessed and if over 50% of the trip landings are under one of the target species categories, then the trip is assigned that category. Most trips did contain over 50% of the landings under one category however a significant number of trips had landings under more than one trip and below the 50% threshold. Several mixed categories exist describing landings of crustaceans and demersals (MCD), cephalopod and demersal (MCF), pelagic and demersal (MPD) among others.

Location is adapted from the landing port variable included in the database with the ICES rectangle of the port utilised. It is assumed that small-scale fishing vessels sell their landings at their home port (landing port variable) and that they do not travel far out to sea and remain in the ICES rectangle of the port.
The calculated variables for gear and target species are combined in excel to form the metier.

Anonymisation

Currently there is a high degree of anonymity in each of the three databases. Analysis of metiers identifies those containing 3 vessels or more and 5 vessels or more. Results can be

seen in table 1. In this exercise only those metiers containing at least five vessels were included.

Table 1: assessment of metiers estimated within each database

Database	Total Metiers	Metiers +3 vessels	Metiers +5 vessels
Logbooks	209	106	80
Landings Declarations	351	128	93
Sales Notes	107	54	42

Database merge

The three individual datasets were assessed in terms of vessel numbers and fishing trips. In the logbook dataset there were 485 vessels of which 483 were present in the landings dataset and 424 were present in the sales notes dataset. Of the landings dataset's 483 vessels, all appear in the logbooks and 423 in the sales notes. In the sales there are 2000 vessels, of which 424 appear in the logbooks and 423 in the landing declarations.

The logbooks contain 247,254 trips of which 22,899 appear in the landings and 5,621 in the sales notes. In the landings dataset there are 196,346 trips with 22,899 in logbooks and 49,127 in sales. In the sales notes there are 344,701 trips, of which 5,621 are in the logbooks and 49,127 in the landings.

In terms of vessel numbers there are very similar amounts present in the logbooks and landings declarations (unsurprisingly) however there are significantly more trips present in the logbooks (again unsurprisingly). Therefore, the logbook dataset was chosen to merge with the sales notes dataset to incorporate the most vessels and trips across the distinct datasets. Duplicates across the logbook and sales notes datasets were removed from the sales notes and the files were merged in MS Access.

Assumptions

Variables that were missing from the files and required for the data processing included departure date, departure time, arrival date and arrival time. An important assumption here is that effort is a maximum of one day at sea. This is based on the fact that the assessment here focuses on small scale fishing vessels that typically do not fish for more than one day at a

time. However, there certainly are vessels that have fishing trips that are over 24 hours. To reduce the impact of this assumption on these vessels a range of hours per day were assigned depending on the metier with those metiers more likely to have multi-day fishing trips given higher hours per day. Additionally, as logbooks contain data for each day of each fishing trip, the effect of this assumption will be ameliorated.

Three different daily hours at sea were utilised; 8 hours, 12 hours and 16 hours. All FPO gears(pots) were assigned with 8 hours per day. The gears of dredges, gillnetters, pair trawlers, seiners and beam trawls were assigned 12 hours per day. Longline, trawls, purse seines and pelagic trawls were assigned 16 hours per day.

Using these assumptions, the variables departure date, departure time, arrival date and arrival time are filled in.

Vessel statistics were added into the file detailing the engine power (in kWs) and gross tonnage using official vessel registers.

Data analysis in R using fec-R and checking data

In order to run the fec-R package, which estimates effort, data must be arranged in a data frame in R and formatted correctly for the script to run. Variable formats include character, double, integer, numeric among others. Missing data points and duplicate entries were other common errors appearing in the data checking exercise. Missing data points were filled in using the median value for that variable.

Once the data frame passed the checks as specified in the document 'fecR: Checking the fishing trip input data', the fec-R script can be run on the data frame with effort then estimated. In the final data file to be inputted into the cabfishman geotool the following variables are collated; Year, Quarter, Month, vslFlgCtry, vslFtechCod, vslLenCls, vslLandLocp, landCtry, landLoc, area, rect, eez, foCatEu6, foCatEu5, gear, NbFishTrips, SeaDays, HoursAtSea, FishDays, sppCode, sppName, LandWG, LandVL. All variables here are extracted from the original database collated except for the four effort variables which are sourced from the fec-R script run previously.

V. United Kingdom

Data for UK was taken from the fisheries activity database provided by the Marine Management Organization (MMO). This database reports vessel specific information (e.g., length, engine size) as well as trip level information (departure and landing date and port, fishing location, gear and mesh size, catch (species, live weight and value) for all vessels landing in the UK or fishing the UK EEZ. MMO processes the data for UK following the procedure outlined in: **UK Sea Fisheries Statistics 2019 Methodology (publishing.service.gov.uk)**. As such the data provided by MMO included all the information needed to provide the data input for fecR. Although the project focus on the North Atlantic area, data was provided for all vessels recording their landings in the fisheries activity database from MMO. This was done to be aligned with the data collected for the other workstreams in this project. Preparing the data for fecR returned some mismatch within the MMO fisheries activity data gear classification and the DCF gear classification. These mismatches were corrected manually.

Final data was provided as anonymous trip-level landings and fishing effort data from UK vessels up to 18 meters length registered in the cooperation area of Interreg Atlantic area. Data are provided for the years 2015-2018 both years included and contain all fishing trips with return dates (landings date) from 2015-01-01 to 2018-31-12, i.e. some departure dates may be before 2015-01-01. Raw data on individual fishing trips are provided.

For the methodology to calculate and partition the fishing effort variables (days at sea, hours at sea and fishing days), we have used the one developed and admitted in the DCF context (see "https://op.europa.eu/en/publication-detail/-/publication/8c5583fa-c360-11e6-a6db-01aa75ed71a1/language-en").

Conclusions

The standardisation of the effort variable is fundamental when assessing the activity of different fisheries. In this standardisation work, the fecR package has been used to standardise the effort of SSF in the AA. However, due to the limitations of the official

transversal data (FSN, Logbooks) collected through the EU Control Regulation, several assumptions have had to be made in order to comply with the formats required by fecR, and to be able to run this code developed in R.

The absence of some variables has occurred mainly in the case of vessels with a length of less than 10m, and which only have to report FSN. Two main assumptions had to be made due to the absence of these variables in the data sources:

The first assumption was made for the allocation of spatial information at the trip level (e.g. ICES Division and ICES rectangle). This information is missing in the FSN and is not always complete in the case of SSF logbooks. For this purpose and knowing that these vessels usually fish close to their home ports, the ICES rectangle corresponding to the home port of each of the vessels has been assigned. This information is available in the European Fleet Register for most fleets (home port is not available for the Irish fleet, only town of registration). This has allowed all fishing trips to be assigned a corresponding ICES rectangle per trip. The other major difficulty is the allocation of the fishing gear corresponding to each of the trips. SSFs are very polyvalent fisheries, using a wide variety of fishing gears. It is also the case in some Member States that even on the same fishing trip, the same vessel uses different fishing gears. This is why skippers, when filling in their logbooks or FSNs, should indicate the main fishing gear used. In addition, in some cases, they often use groupings of different fishing gears, as may be the case for net fishing gears. When assigning the fishing gear, within CABFishMAN, different alternatives have been used, such as the use of statistical tools (e.g. cluster analysis), information collected through sampling of SSFs etc.

Effort information by gear is essential for SSF management. However, especially in the case of smaller vessels, this information is not always reported, or the resolution of the information reported is not always the best. Furthermore, over the last few years, the importance of having this information at a spatial level has become essential, due to the need to be able to carry out spatial management of the marine area.

CABFishMAN has been able to develop a complete data set for the AA, where all SSF fishing trips have been identified by gear and at a spatial resolution of ICES statistical rectangle. While this is an important step forward, there is still room for improvement in terms of collecting

this type of information, even for SSF. The obligatory reporting of information on the fishing gear or gears used by fishing trip by the skkipers, together with other variables such as number of nets, hooks, etc., including spatial information, would considerably improve the knowledge of the SSF activity. Furthermore, as the demand for high resolution spatial data is becoming more and more urgent and necessary, the use of devices adapted to these fleets would allow basic information to be obtained that would greatly improve maritime spatial management.

References

Anon 2013. Report of the Working group on Common understanding and statistical methodologies to estimate/re-evaluate transversal data in small-scale fisheries. 21-23 May 2013, Nantes, France. 76pp.

Almendra Castro Ribeiro C, Holmes S, Scott F, Calvo Santos A, Verlé K, Vukov I, Dalskov J, Egekvist J, Baikov A, Ahvonen A, Korhonen P, Leonardi S, Demanèche S, Berkenhagen J, Tzouramani I, Papoutsis A, Labanchi L, Davidjuka I, Ozernaja O, Jakovleva I, Balnis A, Aquilina M, Muscat E, Hamon K, Avdič E, Paulrud A, Ringdahl K, Elliott M, Reilly T. Report on the Workshop on Transversal Variables. (Linking economic and biological effort data (call) design). 19th -23rd January 2015. EUR 27153. Luxembourg (Luxembourg): Publications Office of the European Union; 2015. JRC95206.

Castro Ribeiro, C., Holmes, S., Scott, F., Berkenhagen, J., Demaneche, S., Prista, N., Reis, D., Reilly, T., Andriukaitiene, J., Aquilina, M., Avdič Mravlje, E., Calvo Santos, A., Charilaou, C., Dalskov, J., Davidiuk, I., Diamant, A., Egekvist, J., Elliot, M., Ioannou, M., Jakovleva, I. Kuzebski, E., Ozernaja, O., Pinnelo, D., Thasitis, I., Verlé, K., Vitarnen, J., Wójcik, I..Report of the 2nd Workshop on Transversal Variables. Nicosia, Cyprus. 22-26 February 2016. A DCF ad-hoc workshop. 109pp.EUR 27897; doi 10.2788/042271.

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.

Annex I. fecR R package

fecR: Calculating fishing effort

Finlay Scott[1], Nuno Prista[2] and Thomas Reilly[3]

1. European Commission, DG Joint Research Centre, Directorate D - Sustainable Resources, Unit D.02 Water and Marine Resources, via Enrico Fermi 2749, 21027 Ispra (VA), Italy

2. SLU, Swedish University of Agricultural Sciences, Havsfiskelaboratoriet, Turistgatan 5, 453 30 LYSEKIL

3. Marine Scotland Science, PO Box 101, 375 Victoria Road, Aberdeen, AB11 9DB, Scotland, UK

2017-09-08

Contents

1 Introduction

This vignette introduces the *fecR* package and how to use it. The package implements fishing effort calculations that were developed at the 2nd Workshop on Transversal Variables held in Nicosia, Cyprus on 22-26 February 2016 (Castro Ribeiro et al., 2016).

fecR can be used to calculate two types of fishing effort: *Days at Sea* and *Fishing Days*. These effort types measure different things. Days at Sea is concerned with the economics of fishing whereas Fishing Days is concerned with the fishing pressure on the stock. These measures are explained in more detail below.

2 Workflow

Using *fecR* is straightforward. It can be broken into three stages:

- Prepare the data;
- Check the data;
- Calculate efort.

2.1 Preparing the data

For the effort calculation function to work the input data must be prepared in the correct format. Details of this format can be found in the package vignette *checking_data* and also in the Annexes of the Nicosia report (Castro Ribeiro et al., 2016).

The data must be stored as a *data.frame* in R. Each column of the data.frame has a specfic name and a specific format. For example, the departure data column must be called *depdate* and the entries of that column must be a character string of numeric characters of the format "YYYYMMDD", e.g. "20161027".

It is probably easier to prepare your data using a spreadsheet and export it as a CSV file. The CSV file can then be read into R as a *data.frame* using the *read.csv()* function.

2.2 Checking the data

It is important to check that the data is in the correct format before trying calculate the fishing effort. The function to do this is *check_format()*. Details and examples of how to use this function can be found in the package vignette *checking_data*. It is possible to for the *check_format()* to perform some simple corrections to the data.

The data should be checked and revised until calling *check_format()* produces a happy success message and no warnings.

2.3 Calculating the effort

The final step is to call the *calc_fishing_effort()* function with the data. This function calculates both effort measures and returns a list of two *data.frame*s. One holds the Days at Sea, the other Fishing Days.

3 Fishing trips

Fishing effort is calculated for each fishing trip. It is therefore important to explore what we mean by trip and what happens during a trip. A fishing trip has the same vessel identifier and departure and return dates and times. Each fishing trip has a unique trip identifier.

An example data set is created here. It consists of two fishing trips. As mentioned above, it is essential that the input data is formatted correctly. More information can be seen in the *checking_data* vignette.

```
# Make fake data of two trips
trip1 <- data.frame(
    eunr_id = "my_boat", loa = 2000, gt = 70, kw = 400, trip_id = "trip1",
    # 4 day trip
    depdate = "20140718", deptime = "0615", retdate = "20140721", rettime = "1615",
    # Only fish on 2 of those
    fishdate = c("20140719", "20140719", "20140719", "20140719", "20140720", "20140720", "20140720"),
    gear = c("OTB","OTB","OTB ","GN","OTB","GN","FPO"),
    gear_mesh_size = c(80,80,80,50,80,50,0),
    fishing_area = "27.4.B", economic_zone = "EU",
    rectangle = c("39F0","40F0","41F0","41F0","41F0","41F0","41F0"),
    stringsAsFactors = FALSE
)
trip2 <- data.frame(
    eunr_id = "my_boat", loa = 2000, gt = 70, kw = 400, trip_id = "trip2",
    # 2 day trip
```

```
    depdate = "20140722", deptime = "0615", retdate = "20140723", rettime = "0600",
    # Only fish on 2 of those
    fishdate = c("20140722", "20140723", "20140723", "20140723"),
    gear = c("OTB","OTB","GN","FPO"),
    gear_mesh_size = c(80,80,50,0),
    fishing_area = "27.4.B",
    economic_zone = "EU",
    rectangle = c("39F0","39F0"),
    stringsAsFactors = FALSE
)
# Stich them together to make a data set of two fishing trips
dat <- rbind(trip1, trip2)
dat
```

```
##    eunr_id  loa gt  kw trip_id  depdate deptime  retdate rettime fishdate
## 1  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140719
## 2  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140719
## 3  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140719
## 4  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140719
## 5  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140720
## 6  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140720
## 7  my_boat 2000 70 400   trip1 20140718    0615 20140721    1615 20140720
## 8  my_boat 2000 70 400   trip2 20140722    0615 20140723    0600 20140722
## 9  my_boat 2000 70 400   trip2 20140722    0615 20140723    0600 20140723
## 10 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600 20140723
## 11 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600 20140723
##    gear gear_mesh_size fishing_area economic_zone rectangle
## 1   OTB             80       27.4.B            EU      39F0
## 2   OTB             80       27.4.B            EU      40F0
## 3   OTB             80       27.4.B            EU      41F0
## 4    GN             50       27.4.B            EU      41F0
## 5   OTB             80       27.4.B            EU      41F0
## 6    GN             50       27.4.B            EU      41F0
## 7   FPO              0       27.4.B            EU      41F0
## 8   OTB             80       27.4.B            EU      39F0
## 9   OTB             80       27.4.B            EU      39F0
## 10   GN             50       27.4.B            EU      39F0
## 11  FPO              0       27.4.B            EU      39F0
```

The data above consists of two fishing trips. The trips are identified by the *trip_id* column which must be unique for each trip. Within each trip the vessel identifier, departure and return dates and times must be the same (the *eunr_id*, *depdate*, *deptime*, *retdate* and *rettime* columns).

Each fishing trip is made up of different fishing activites. A fishing activity is the use of a particular gear in a particular area on a particular date. Gear is a combination of the gear type and the gear mesh size. Gears of the same type but of different mesh size are considered to be different gears. Area is a combination of the economic zone, the fishing area and the fishing rectangle.

Specifically, fishing activity is defined as the use of a type of gear (*gear*) with a particular mesh size (*gear_mesh_size*) in a particular economic zone (*economic_zone*) in a particular fishing area (*fishing_area*) in a particular rectangle (*rectangle*) on a particular date (*fishdate*). Each row of the data table represents an instance of fishing activity. Within each trip each fishing activity is unique.

For example, *trip1* in the above data departed port on the 18th July 2014 at 06:15 and returned on the 21st July at 1615. During the trip seven unique fishing activities were performed over two days. Four activities happened on the 19th July and three on the 20th. On the 19th July an OTB gear with a mesh size of 80 was

used in three different rectangles and a gill net (GN) was also used. The three OTB entries on the 19th July count as three different activities because they happened in three different rectangles and therefore different areas. The use of the GN gear type happened in the same area as one of the OTB activities. Although it is the same area, because this is a different gear type it is considered to be another activity.

Days at sea is reported at the gear (type and mesh size), fishing area and economic zone level. Fishing days is reported at the gear (type and mesh size), fishing area, economic zone and rectangle level, i.e. a lower geographical level.

4 Calculating Days at Sea

The total Days at Sea of a trip is calculated as the number of commenced 24 hour periods of the trip. Only the total duration of trip is considered, i.e. the difference between departing and returning.

For example, *trip1* above starts on 18th July 2014 at 06:15 and returns on the 21st July 2014 at 16:15. The total duration of the trip is therefore 3*24 + 10 = 85 hours. This takes the trip into the 4th commencing 24 hour period. The total Days at Sea is therefore 4 days.

We want to know the Days at Sea attributed to each activity on a trip. The total Days at Sea are split equally across each day on the trip on which fishing occurs, i.e. the number of unique fishing dates on the trip. In *trip1* the only days on which fishing occurs are the 19th and 20th of July, i.e. the number of fishing dates is 2. The total Days at Sea is split equally among them so each fishing date gets 2 Days at Sea each.

Within each fishing date, the Days at Sea attributed to that day is split equally across the fishing activities on that day. In *trip1* on the 19th July there are four different fishing activities (given by the different gear and area combinations). Each fishing activity on the 19th July therefore gets 2 / 4 = 0.5 fishing day each. On the 20th July there are three fishing activities so each one gets 2 / 3 = 0.67 Fishing Days each.

The Days at Sea for each trip are summed over the fishing dates and rectangles so that the Days at Sea of each trip is reported at the gear (type and mesh size), fishing area and economic zone level.

5 Calculating Fishing Days

Calculating the Fishing Days of a trip requires splitting the gears types used during the trip into active and passive. The distinction is held in the *gear_codes* data set included in the package. In essence, fishing with passive gears happens in parallel while fishing with active gears happens in series.

Considering the active gears, each date which has a fishing activity using an active gear is allocated 1 fishing day. This is split equally among the active gear activities on that date. For example, the 19th July of *trip1* has four fishing activities, of which three are active gears (OTB is an active gear, GN is a passive gear). Each of the active gear activities is therefore allocated 1 / 3 = 0.33 Fishing Days. The second day, the 20th July, has three activities of which only one uses an active gear, the OTB gear (FPO is a passive gear). The OTB activity is therefore allocated 1 fishing day.

Considering the passive gears, each fishing activity using a passive gear is allocated 1 fishing day. For example, the 19th July of *trip1* has one passive gear activity (the GN activity). This activity is allocated 1 fishing day. If there are multiple fishing activities with passive gears on the same fishing date, they each get allocated 1 fishing day. For example, the 20th July of *trip1* has three fishing activities two of which are with pasive gears (GN and FPO). Both of the passive fishing activities get 1 fishing day each.

The Fishing Days for each trip are summed over the fishing dates so that the Fishing Days of each trip is reported at the gear (type and mesh size), fishing area, economic zone and rectangle level.

There may be more Fishing Days on a trip than Days at Sea.

6 Simple demonstration

First we load the *fecR* package:

```
library(fecR)
```

6.1 Checking the data with *check_format()*

Before calculating effort we want to check the data is in the correct format. We do this with the *check_format()* function.

```
dat <- check_format(dat)
```

```
## [1] "============================"
## [1] " STECF Transversal2 checks on formats"
## [1] "============================"
## [1] "Checking column names..."
## [1] "Checking eunr_id..."
## [1] "Checking loa..."
## [1] "Checking gt..."
## [1] "Checking kw..."
## [1] "Checking depdate..."
## [1] "Checking retdate..."
## [1] "Checking fishdate..."
## [1] "Checking deptime..."
## [1] "Checking rettime..."
## [1] "Checking gear..."
## [1] "ATT: Whitespace found in gear column"
## [1] "Problem rows: 3"
## [1] "ATT: Unknown code in gear column"
## [1] "Problem rows: 3"
```

```
## Warning in check_format(dat): ATT: Unknown code in gear column
```

```
## [1] "Checking gear_mesh_size..."
## [1] "Checking fishing_area..."
## [1] "Checking economic_zone..."
## [1] "Checking rectangle..."
## [1] "Checking trip_id"
## [1] "Checking uniqueness of trip_id..."
## [1] "Checking duplicates..."
## [1] "============================"
## [1] "Attention: There are problems with this data set."
## [1] "============================"
```

A deliberate error has been included in the data and check shows a warning. The problem is that whitespace has been found in the *gear* column on row 3. We can take a look and see that there is whitespace at the end of the entry.

```
dat[3,"gear"]
```

```
## [1] "OTB "
```

We could fix this ourselves but this simple error can be corrected by the autocorrection option. We call *check_format()* again, this time wih the *correct* option.

```
dat <- check_format(dat, correct=TRUE)
```

```
## [1] "============================"
## [1] " STECF Transversal2 checks on formats"
## [1] "============================"
## [1] "================================"
## [1] "WARNING WARNING WARNING WARNING"
## [1] "correct=TRUE selected => changes other than simple formatting will be made to the dataset"
## [1] "The changes may not be what you want and implemented only for testing purposes"
## [1] "The corrected data is returned as the output"
## [1] "Correct input data should run with correct=FALSE"
## [1] "================================"
## [1] "Checking column names..."
## [1] "Checking eunr_id..."
## [1] "Checking loa..."
## [1] "Checking gt..."
## [1] "Checking kw..."
## [1] "Checking depdate..."
## [1] "Checking retdate..."
## [1] "Checking fishdate..."
## [1] "Checking deptime..."
## [1] "Checking rettime..."
## [1] "Checking gear..."
## [1] "ATT: Whitespace found in gear column"
## [1] "Problem rows: 3"
## [1] "Attempting to correct by removing whitespace from gears"
## [1] "Checking gear_mesh_size..."
## [1] "Checking fishing_area..."
## [1] "Checking economic_zone..."
## [1] "Checking rectangle..."
## [1] "Checking trip_id"
## [1] "Checking uniqueness of trip_id..."
## [1] "Checking duplicates..."
## [1] "============================"
## [1] "The returned data passes the check."
## [1] "============================"
```

The output shows that the returned data has been corrected and should now pass check. We pass it in again just to make sure:

```
# Problem has been fixed
dat[3,"gear"]
```

```
## [1] "OTB"
```

```
# Check again
dat <- check_format(dat)
```

```
## [1] "============================"
## [1] " STECF Transversal2 checks on formats"
## [1] "============================"
## [1] "Checking column names..."
## [1] "Checking eunr_id..."
## [1] "Checking loa..."
## [1] "Checking gt..."
## [1] "Checking kw..."
```

```
## [1] "Checking depdate..."
## [1] "Checking retdate..."
## [1] "Checking fishdate..."
## [1] "Checking deptime..."
## [1] "Checking rettime..."
## [1] "Checking gear..."
## [1] "Checking gear_mesh_size..."
## [1] "Checking fishing_area..."
## [1] "Checking economic_zone..."
## [1] "Checking rectangle..."
## [1] "Checking trip_id"
## [1] "Checking uniqueness of trip_id..."
## [1] "Checking duplicates..."
## [1] "============================"
## [1] "The returned data passes the check."
## [1] "============================"
```

The check now shows that there are no problems with the data.

6.2 Calculating fishing effort with *calc_fishing_effort()*

We can now calculate the fishing effort for our data using the *calc_fishing_effort()*: The check function is automatically run again unless we turn it off.

```
effort <- calc_fishing_effort(dat, check=FALSE)
```

This returns a list with two elements: *days_at_sea* and *fishing_days*. We can inspect these to look at the different effort types.

Days at Sea is given by trip, gear, fishing area and economic zone.

```
effort$days_at_sea
```

```
##   eunr_id  loa gt  kw trip_id  depdate deptime  retdate rettime gear
## 1 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  FPO
## 2 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615   GN
## 3 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 4 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600  FPO
## 5 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600   GN
## 6 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600  OTB
##   gear_mesh_size fishing_area economic_zone days_at_sea
## 1              0       27.4.B            EU   0.6666667
## 2             50       27.4.B            EU   1.1666667
## 3             80       27.4.B            EU   2.1666667
## 4              0       27.4.B            EU   0.1666667
## 5             50       27.4.B            EU   0.1666667
## 6             80       27.4.B            EU   0.6666667
```

The sum of the Days at Sea for a trip is the number of commenced 24 hour periods of that trip.

Fishing Days is given by trip, gear (type and mesh size), fishing area, economic zone and rectangle.

```
effort$fishing_days
```

```
##   eunr_id  loa gt  kw trip_id  depdate deptime  retdate rettime gear
## 1 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  FPO
## 2 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615   GN
```

```
## 3 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 4 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 5 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 6 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600  FPO
## 7 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600   GN
## 8 my_boat 2000 70 400   trip2 20140722    0615 20140723    0600  OTB
##   gear_mesh_size rectangle fishing_area economic_zone fishing_days
## 1              0      41F0       27.4.B            EU    1.0000000
## 2             50      41F0       27.4.B            EU    2.0000000
## 3             80      39F0       27.4.B            EU    0.3333333
## 4             80      40F0       27.4.B            EU    0.3333333
## 5             80      41F0       27.4.B            EU    1.3333333
## 6              0      39F0       27.4.B            EU    1.0000000
## 7             50      39F0       27.4.B            EU    1.0000000
## 8             80      39F0       27.4.B            EU    2.0000000
```

We can inspect each trip individually:

```
subset(effort$fishing_days, trip_id=="trip1")
```

```
##   eunr_id  loa gt  kw trip_id  depdate deptime  retdate rettime gear
## 1 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  FPO
## 2 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615   GN
## 3 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 4 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
## 5 my_boat 2000 70 400   trip1 20140718    0615 20140721    1615  OTB
##   gear_mesh_size rectangle fishing_area economic_zone fishing_days
## 1              0      41F0       27.4.B            EU    1.0000000
## 2             50      41F0       27.4.B            EU    2.0000000
## 3             80      39F0       27.4.B            EU    0.3333333
## 4             80      40F0       27.4.B            EU    0.3333333
## 5             80      41F0       27.4.B            EU    1.3333333
```

We can see that the GN gear type has 2 Fishing Days allocated to it. This is because it is used in two activities during the trip.

7 References

Castro Ribeiro, C., Holmes, S., Scott, F., Berkenhagen, J., Demaneche, S., Prista, N., Reis, D., Reilly, T., Andriukaitiene, J., Aquilina, M., Avdič Mravlje, E., Calvo Santos, A., Charilaou, C., Dalskov, J., Davidiuk, I., Diamant, A., Egekvist, J., Elliot, M., Ioannou, M., Jakovleva, I. Kuzebski, E., Ozernaja, O., Pinnelo, D., Thasitis, I., Verlé, K., Vitarnen, J., Wójcik, I..Report of the 2nd Workshop on Transversal Variables. Nicosia, Cyprus. 22-26 February 2016. A DCF ad-hoc workshop EUR; doi

Annex II. fecR format

CABFishMAN WP4 Format proposal Output Data

Columns 4 to 7 are conditional to the vessel performing the fishing trip considered. Columns 1 to 3 and 8 to 15 are conditional to the characteristics *(time, gear, landings location, spatialization …)* of the fishing trip considered. Columns 20-21 are conditional to the species landed from the fishing trip considered. Columns 16-19 present the cumulative fishing effort metrics calculated by the 15 first previous fields ***(!Do not cumulated them by species!)***. Columns 22-23 present the cumulative landings metrics *(in value and weight)* calculated by the 15 first fields and the following species fields 20-21.

Field	Variable name	Description	Format	Code list or example
1	Year	Year	Character string of length 4.	**2015–2018** Reference date = return date of the vessels' fishing trips.
2	Quarter	Quarter	Character string of length 1.	**1-4** Reference date = return date of the vessels' fishing trips.
3	Month	Month	Character string of length 2.	**01-12** Reference date = return date of the vessels' fishing trips.
4	vslFlgCtry	Vessel Flag Country The flag country of the vessel.	Character string of length 3.	Finite code list- ISO 3166 country alpha-3 codes: **PRT, ESP, FRA, GBR, IRL** *https://www.iso.org/obp/ui/#search/code/*
5	vslFtechCod	Vessel Fleet Segment DCF / EU-MAP. Principal fishing technique *(or combination of)* performed by the vessel.	Character string of length 3.	Finite code list- defined by the EU-MAP Commission Delegated Decision (EU) 2021/1167**: DFN, DRB, DTS, FPO, HOK, MGO, MGP, PGO, PGP, PMP, PS_, OTM, TBB** *https://datacollection.jrc.ec.europa.eu/wordef/fleet-segment-dcf*
6	vslLenCls	Vessel Length Class Vessel length class overall (m)	Character string of length 6.	Finite code list: **VL0006, VL0608, VL0810, VL1012, VL1215, VL1518, VL18XX** *e.g. VL0006 = vessels less than 6 meters in length (<6m) VL0608 = vessels between 6 meters and 8 meters in length ([6-8[m).*
7	vslLandLocp	Principal landing location of the vessels.	Character string of length 5; the first 2 letters correspond to the country, the remaining 3 to the	Finite code list - UNECE/LOCODE codification, e.g.: **FRBES, ESVGO, GBABD, PTLIS, IEKBS …** *https://unece.org/trade/cefact/unlocode-code-list-country-and-territory*

			unique location code.	
8	landCtry	Landing country associated with the landing location The country where the fishes have been landed	Character string of length 3.	Finite code list- ISO 3166 country alpha-3 codes: **PRT, ESP, FRA, GBR, IRL** *https://www.iso.org/obp/ui/#search/code/* Depending of the vessels' fishing trips landing location
9	landLoc	Landing location of the vessels' fishing trips. The location where the fishes have been landed	Character string of length 5 the first 2 letters correspond to the country, the remaining 3 to the unique location code.	Finite code list - UNECE/LOCODE codification, e.g.: **FRBES, ESVGO, GBABD, PTLIS, IEKBS …** *https://unece.org/trade/cefact/unlocode-code-list-country-and-territory*
10	area	FAO area codes *(ICES division)* recorded for the fishing effort variables and the species landings	Character string of length 6	Finite code list– FAO codification, e.g.: **27.8.a, 27.7.e …** Detailed FAO division: *https://www.fao.org/fishery/area/Area27/en*
11	rect	ICES statistical rectangles recorded for the fishing effort variables and the species landing	Character string of length 4	Finite code list- ICES codification, e.g.: **31E2, 24E6 …** Use "**-9**" if unknown Detailed ICES Statistical rectangles: *https://www.ices.dk/data/maps/Pages/ICES-statistical-rectangles.aspx*
12	eez	Exclusive Economic Zone recorded for the fishing effort variables and the species landings	Character string of length 3	Finite code list consisting of: **XIN** for **International Waters**. *https://circabc.europa.eu/sd/a/b137e37d-8243-4f76-a7ba-7bb2928a5cf1/MDRMatrixIndex.html* & ISO 3166 country alpha-3 codes for **national waters**: **PRT, ESP, FRA, GBR, IRL** *https://www.iso.org/obp/ui/#search/code/*
13	foCatEu6	Metier DCF level 6 https://datacollection.jrc.ec.europa.eu/wordef/fishing-activity-metier	Character string with gear type, target assemblage of species, mesh size ranges and other selective device components, underscore separated.	Finite code list; DCF codification, e.g.: **DRB_MOL_0_0_0 , OTB_DEF_>=120_0_0 …** See attached Excel file.
14	foCatEu5	Metier DCF level 5 https://datacollection.jrc.ec.europa.eu/	Character string with gear type and target assemblage of species,	Finite code list; DCF codification, e.g.: **DRB_MOL, OTB_DEF …** See attached Excel file.

		wordef/fishing-activity-metier	underscore separated.	
15	gear	Gear type *(Metier DCF level 4)* https://datacollection.jrc.ec.europa.eu/wordef/fishing-activity-metier	Character string with gear type.	Finite code list; DCF codification, e.g.: **DRB, OTB …** See attached Excel file. Derived from the International Standard Statistical Classification of Fishing Gear (ISSCFG) *https://www.fao.org/cwp-on-fishery-statistics/handbook/capture-fisheries-statistics/fishing-gear-classification/en/*
16	NbFishTrips	Number of fishing trips Calculated and partitioned following the DCF agreed methodology. *https://op.europa.eu/en/publication-detail/-/publication/8c5583fa-c360-11e6-a6db-01aa75ed71a1/language-en*	Numeric Unique value calculated by the 15 first previous fieds. **Do not cumulated them by species.**	
17	SeaDays	Days at sea Calculated and partitioned following the DCF agreed methodology. *https://op.europa.eu/en/publication-detail/-/publication/8c5583fa-c360-11e6-a6db-01aa75ed71a1/language-en*	Numeric Unique value calculated by the 15 first previous fields. **Do not cumulated them by species.**	
18	HoursAtSea	Hours at sea Calculated and partitioned following the DCF agreed methodology. *https://op.europa.eu/en/publication-detail/-/publication/8c5583fa-c360-11e6-a6db-01aa75ed71a1/language-en*	Numeric Unique value calculated by the 15 first previous fields. **Do not cumulated them by species.**	
19	FishDays	Fishing days Calculated and partitioned following the DCF agreed methodology. *https://op.europa.eu/en/publication-detail/-/publication/8c5583fa-c360-11e6-a6db-01aa75ed71a1/language-en*	Numeric Unique value calculated by the 15 first previous fields. **Do not cumulated them by species**	
20	sppCode	Species code (3A_code) of the fishes landed The species codes of all the recorded fishes landings from the fishing trip considered.	Character string of length 3.	Finite code list– FAO codification, e.g.: **SCE, BSS, SOL …** ASFIS List of Species for Fishery Statistics Purposes *http://www.fao.org /fishery/collection/ asfis/en*

21	sppName	Species scientific name of the fishes landed	Character string.	Finite code list- FAO codification, e.g.: **Pecten maximus, Dicentrarchus labrax, Solea solea …** ASFIS List of Species for Fishery Statistics Purposes *http://www.fao.org /fishery/collection/ asfis/en*
22	LandWG	Landed Weight The live weight equivalent in KG for each recorded species (fields 20-21) and the 15 first fields.	Numeric Unique value calculated by the 15 first previous fields and the species fields 20-21.	
23	LandVL	Landed Value estimated The landed value estimated in € for each recorded species (fields 20-21) and the 15 first fields	Numeric Unique value calculated by the 15 first previous fields and the species fields 20-21.	

Other fields that could be discussed.

Further information conditional to the vessel performing the fishing trip considered:

X	vslAgeCls	Vessel Age Class	Character string of length 6.	Finite code list consisting of: **VA0010, VA1020, VA2030, VA3040, VA40XX**
X	fshAgeCls	Fishermen Age Class	Character string of length 6	Finite code list consisting of: **FA0020, FA2030, FA3040, FA4050, FA5060, FA60XX**
X	vslNfhCls	Vessel average crew on-board (nb of fishermen) Class	Character string of length 7.	Finite code list consisting of: **VNf0001, VNf0102, VNf0203, VNf0304, VNf0405, VNf0510, VNf1020, VNf20XX**
X	vslPwrCls	Vessel Power (kW) Class	Character string of maximum length 9	Finite code list consisting of: **VP0050, VP50100, VP100150, VP150200, VP200500, VP5001000, VP1000XX**
X	vslTonCls	Vessel gross tonnage (GT) Class	Character string of maximum length 12	Fixed code list consisting of: **VGT00100, VGT100300, VGT300500, VGT5001000, VGT10005000, VGT500010000, VGT10000XX**

Further information conditional to the characteristics of the fishing trip considered:

X	*foCatEu7*	*Metier DCF level 7 (**national level**)*	*Character string with gear and target group of species (FAO code)*	*Fixed code list; Examples:* *DRBSCE LTLBSE*

Further cumulative fishing effort metrics: KWSeaDays, KWFishingDays, GTSeaDays, GTFishingDays

Methodology to calculate the vessels DCF Fleet Segment.

The vessels DCF Fleet Segment should be calculated vessel by vessel on the basis of the annual fishing activity of the vessels. Take care than one vessel could use several gears during a year but belong to only one DCF fleet segment for a given year. The DCF fleet segment of a vessel expressed the principal fishing technique *(or combination of in case there is no principal (all less than 50% of the fishing activity of the vessel))* performed by the vessel. The DCF fleet segment are detailed on the STECF website: https://datacollection.jrc.ec.europa.eu/web/dcf/wordef/fleet-segment-dcf and in the table8 of the DCF Commission Delegated Decision (EU) 2021/1167: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32021D1167&from=EN

The reference table linking metier and fishing technique associated is available in attached Excel file calculated from the reference list of metier developed by the ISSG on metier issue *(see for the reference list of metier by RCG - https://github.com/ices-eg/RCGs/tree/master/Metiers)*.

The different DCF fleet segment retained are the following:

Vessels using 'passive' gears:

DFN = Drift and/or fixed netters *(DFN = principal (more than 50%) fishing technique performed by the vessel)*
FPO = Vessels using pots and/or traps *(FPO = principal (more than 50%) fishing technique performed by the vessel)*
HOK = Vessels using hooks *(HOK = principal (more than 50%) fishing technique performed by the vessel)*
PGO = Vessels using other passive gears *(PGO = principal (more than 50%) fishing technique performed by the vessel)*
PGP = Vessels using polyvalent passive gears only *(vessels using a combination of the following fishing techniques: DFN – FPO – HOK – PGO with no principal (all less than 50%) and no other fishing techniques (i.e. active gears) used)*

Vessels using 'active' gears:

DRB = Dredgers *(DRB = principal (more than 50%) fishing technique performed by the vessel)*
DTS = Demersal trawlers and/or demersal seiners *(DTS = principal (more than 50%) fishing technique performed by the vessel)*
TBB = Beam trawlers *(TBB = principal (more than 50%) fishing technique performed by the vessel)*
OTM = Pelagic trawlers *(OTM = principal (more than 50%) fishing technique performed by the vessel)*
PS = Purse seiners *(PS = principal (more than 50%) fishing technique performed by the vessel)*
MGO = Vessels using other active gears *(MGO = principal (more than 50%) fishing technique performed by the vessel)*
MGP = Vessels using polyvalent active gears only *(vessels using a combination of the following fishing techniques: DRB-DTS-TBB-OTM-PS-MGO with no principal (all less than 50%) and no other fishing techniques (i.e. passive gears) used)*

Vessels using 'active' and 'passive' gears:

PMP = Vessels using active and passive gears *(vessels using a combination of the fishing techniques: DRB-DTS-TBB-OTM-PS-MGO-DFN-FPO-HOK-PGO with no principal (all less than 50%))*

The notion of 'principal' fishing technique *(calculation of the % of fishing activity performed for a given fishing technique for a vessel during a year)* should be based on fishing effort metric *(e.g. number of fishing trips or days at sea performed by the vessel considered with a specific fishing technique during the given year / total number of fishing trips or days at sea performed by the vessel considered during the given year)*.

At the end, each fishing vessel should be allocated to one and only one DCF fleet segment DCF for a given year although he could performed more than one fishing technique during the year.