Construction of socio-economic indicators on commercial fishing fleets in mainland France: part 1 - sampling plan

RBE/EM[1], RBE/STH/LBH[2]
Merzéréaud M.[1], Daurès F.[1], Guyader O.[1], Le Grand C.[1], Leonardi S.[1], Macher C.[1], Spagnol C.[2]
Date: October 2021

Fiche documentaire

<table>
<tr><td colspan="2">Titre du rapport/ Report title : Construction of socio-economic indicators on commercial fishing fleets in mainland France : part 1 - sampling plan</td></tr>
<tr><td>Référence interne : R.DEP/UNIT/LABO AN-NUM

Diffusion :
☒ libre (internet)

☐ restreinte (intranet) – date de levée d'embargo : AAA/MM/JJ

☐ interdite (confidentielle) – date de levée de confidentialité : AAA/MM/JJ</td><td>Date de publication : AAA/MM/JJ
Version : 1.0.0

Référence de l'illustration de couverture
Crédit photo/titre/date

Langue(s) :</td></tr>
<tr><td colspan="2">Résumé/ Abstract :
Under the EU Data Collection Framework (DCF) Regulation, Member States are obliged to collect and provide socio-economic data on fisheries as a basis for advice on fisheries management or for other purposes.
In metropolitan (mainland) France, the data producer is the *Service Statistiques de la Prospective* (SSP; Statistics and Prospective Service) of the Ministry of Agriculture and Food. For data collection, the SSP relies on its two partners: IFREMER (*Institut Français de Recherche pour l'Exploitation de la Mer*; French Research Institute for Exploitation of the Sea), which carries out field surveys, and *Laboratoire d'Économie et de Management de Nantes-Atlantique* (LEMNA; Nantes-Atlantique Economics and Management Laboratory), which collects accounting data from management centres. This document presents the stages in the construction of the annual sampling plan that will subsequently be used for the collection of economic data. This is the first stage in the construction of socio-economic indicators, before data collection and validation.</td></tr>
<tr><td colspan="2">Mots-clés/ Keywords :
Commercial fisheries, economic surveys, socio-economics, economic indicators, sampling plan, collection, validation</td></tr>
<tr><td colspan="2">Comment citer ce document/ How to cite this document :
Merzéréaud M., Daurès F., Guyader O., Le Grand C., Leonardi S., Macher C., Spagnol C. 2021. Construction of socio-economic indicators on commercial fishing fleets in mainland France : part 1 - sampling plan, Report IFREMER-RBE-EM-SIH</td></tr>
<tr><td colspan="2">Disponibilité des données de la recherche :</td></tr>
<tr><td colspan="2">DOI :</td></tr>
</table>

Commanditaire du rapport :	
Nom / référence du contrat : ☐ Rapport intermédiaire (réf. bibliographique : XXX) ☐ Rapport définitif (réf. interne **du rapport intermédiaire** : R.DEP/UNIT/LABO AN-NUM/ID ARCHIMER)	
Projets dans lesquels ce rapport s'inscrit (programme européen, campagne, etc.) :	
Auteur(s) / adresse mail	**Affiliation / Direction / Service, laboratoire**
Merzéréaud Mathieu	RBE/EM
Daurès Fabienne	RBE/EM
Guyader Olivier	RBE/EM
Le Grand Christelle	RBE/EM
Leonardi Sophie	RBE/EM
Macher Claire	RBE/EM
Spagnol Charlène	RBE/STH/LBH
Contributeur(s) / adresse mail	**Affiliation / Direction / Service, laboratoire**
Minne Marie-Dominique	SSP, Ministère de l'Agriculture et de l'Alimentation
Encadrement(s) :	
Destinataire :	
Validé par :	

Table of contents

1 Introduction

Under European Regulation (EU) 2017/1004, known as the DCF (Data Collection Framework)[1] regulation, member states are required to collect and provide socio-economic data on fisheries, as specified in the Commission Delegated Decision (EU) 2021/1167 of 27 April 2021[2], to provide a source of advice for fisheries management and other purposes.

In metropolitan (mainland) France, data production is the responsibility of the *Service Statistiques de la Prospective* (SSP; Statistics and Prospective Service) within the Ministry for Agriculture and Food. The SSP delegates data collection to its two partners: IFREMER (*Institut Français de Recherche pour l'Exploitation de la Mer* ; French Research Institute for Exploitation of the Sea), which conducts field surveys, and the *Laboratoire d'Économie et de Management de Nantes-Atlantique* (LEMNA; Nantes-Atlantique Economics and Management Laboratory), which collects accounting data from management centres.

To produce these data, fishing companies take part in annual surveys; the purpose being to calculate economic indicators per fleet segment, and, more generally, to evaluate the economic performance and development of different sub-fleets. The main topics covered in these surveys are income, costs, physical capital, physical investments and the financial situation, and jobs. These surveys have been endorsed by the *Conseil National de l'Information Statistique* (CNIS; National Council for Statistics) since 2012[3] under the category 'public interest and statistical quality'. This accreditation is a guarantee of rigorous methodology and full compliance with data and statistical confidentiality requirements.

This document sets out the steps involved in developing the annual sampling plan, which is then used by the SSP's two partners – IFREMER and LEMNA – to collect economic data. It, therefore, presents the first stage of establishing socio-economic indicators and should be considered in conjunction with the following documents:

Spagnol C., Le Grand C., Guyader O., 2021. Construction of socio-economic indicators on commercial fishing fleets in mainland France: part 2 - data collection

Le Grand C., Daurès F., Guyader O., Macher C., Leonardi S., Merzereaud M. 2021. Construction of socio-economic indicators on commercial fishing fleets in mainland France: part 3 - validation method

[1] Regulation (EU) 2017/1004 of the European Parliament and of the Council of 17 May 2017: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R1004&from=EN

[2] See Table 7 in the Commission Delegated Decision (EU) 2021/1167 of 27 April 2021: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32021D1167&from=EN

[3] Production de données économiques dans le secteur des pêches maritimes (Enquête pour la) (2021A704AG) - https://www.cnis.fr/enquetes/enquete-pour-la-production-de-donnees-economiques-dans-le-secteur-des-peches-maritimes-2021a704ag/

2 Developing the sampling plan

The Statistics and Prospective Service (SSP) is responsible for developing the annual sampling plan used to collect economic data on the fisheries sector, under the national programme in metropolitan France. This task involves its partners: IFREMER and LEMNA. IFREMER is responsible for the first two steps involved in developing the sampling plan, namely, establishing the survey frame and calculation of adjustment parameters for annual individual income using the FILEMO composite variable. This section of the document also includes a simplified description of the consecutive phases that the SSP implements (calculations for the allocation of surveys, random sampling, and dividing data collection tasks between the two partners), to give readers an overall view of the process.

2.1 Establishing the survey frame

All sampling plans need a **survey frame** that describes the population studied and from which the units surveyed will be selected. This frame also includes sets of individual information that can be drawn on in different planned, operational phases, from random sampling to estimating final indicators (and including stratification, allocation calculations based on predetermined precision objectives, and statistical inference procedures such as calibration).

This section describes the steps of establishing the survey frame used within the sampling plan to collect economic data in metropolitan France. **Two distinct versions** of this frame will be created: one for the **English Channel-North Sea-Atlantic** coastline (identified by the abbreviation AT in the rest of the document and covering vessels operating in the FAO 27 fishing zone) and one for vessels based in the **Mediterranean** (identified by the abbreviation ME in the rest of the document and corresponding to the FAO 37 zone). These two frames are established using a **common methodology** and differ simply in terms of population considered and the specific characteristics of the typological classification employed. In this description, **year N** is taken as **the temporal reference for implementation of the collection exercise**, on the understanding that the survey will collect data about the year N-1.

2.1.1 Step 1. Deployment of the two data sources used: CFR fleet and fishing vessel activity data files

Both survey frames are produced from the same exhaustive sources of individual data. They draw primarily on the **Community Fishing Fleet register (CFR) national data file**, which compiles administrative information and technical details for every French fishing vessel. This file will then be supplemented by an advanced typological classification, which is based on **exhaustive data on fishing vessel activity** collected by reconstructing activity journals[4]. To make sure that there is coherence between the information in the survey frame (established at the end of the year N-1) and the time period focused on in the survey (also the year N-1), an updated version of this file is used. However, a second version of the file, dated 31 December, year N-2, complements this information in two specific ways. Firstly, it includes **the most up-to-date, validated and exhaustive typological classification of fishing sub-fleets** available at the point in time when the base is established. Moreover, it provides a picture going further back, which, when considered alongside the picture acquired from the updated file, can be used to **isolate the vessels present**

[4] https://sih.ifremer.fr/Activite-socio-economie/Activite-des-navires

throughout year N-1 and vessels that experienced notable changes in situation during this year (it will be shown later that this information provides exclusion criteria for the survey frame).

2.1.2 Step 2. Defining segmentations suited to different indicator reporting needs

Segmentation is a way of dividing a population of vessels into 'segments', with **each segment combining the sub-fleet the vessel belongs to[5] and its size class.** Segmentations are created using classifications produced from activity journals at 31/12 of year N-2. They provide the basis for the stratifications that will structure the future sampling plan. As the purpose of the latter is to establish socio-economic indicators suitable both for DCF needs and the research and expert advice-giving needs of IFREMER (conducted at different scales), it is important that the development process incorporates **stratifications** appropriate for the **different levels of reporting anticipated.** Thus, for each of the two coastlines, vessel size classes will be divided up according to specific criteria and using two distinct sub-fleet typologies (the **'DCF' typology**, in keeping with the nomenclatures required by European bodies, and the **'SIH' typology**, which corresponds more to scientific needs) to eventually establish two segmentations per coastline [6].

2.1.3 Step 3. Producing individual indicators for use as inclusion criteria in the survey frame

As mentioned previously, by considering two pictures of the same information base (taken at N-2 and N-1) in parallel, it will be possible to define multiple **individual indicators on changes in status,** and thus to show the **stability** of fishing activity and the vessel **management approach** followed during the period of interest to the survey. As this survey focuses on activity conducted over a full year, applying this notion of stability to various significant factors will naturally identify the vessels it would be pertinent to study, and which will make up the survey frame. Inclusion in the base will, therefore, be conditioned by the combination of five distinct binary indicators:

- **Presence in the fleet on 31/12 of year N-2** (equivalent to the presence of a value entered for sub-fleet typology): if 0, the vessel will not have covered the N-1 'period' fully (exclusion).
- **Exit from the fleet** during year N-1 (equivalent to the presence of a reference date in the CFR file earlier than the majority reference date denoting the last time the file was updated): if 1, the vessel left the fleet before the update (exclusion).
- **Change of owner** during year N-1 (equivalent to a change of value in the Owner field between versions N-2 and N-1 of the data sources): if 1, exclusion from the survey frame.
- **Change of coastline** during year N-1 (equivalent to a change of value in the Coastline field between versions N-2 and N-1 of the data sources): if 1, exclusion from the survey frame.

[5] The purpose of a sub-fleet typology is to group together vessels that have similar behaviours or fishing strategies (the same fishing activities or combinations of activity), or according to criteria such as technical characteristics or sector of origin or activity.

[6] The DCF sub-fleet typology is based on the dominant gear type used in that year whereas the IFREMER SIH typology is based on the vessels' versatility in terms of gear used (IFREMER SIH (2022). Activité des navires. IFREMER https://doi.org/10.12770/d31aac89-5d6b-4bee-9e92-611579a846d2)(Berthou Patrick, Daures Fabienne, Guyader Olivier, Leblond Emilie, Merrien Claude, Demaneche Sebastien, Jezequel Michele (2003). Typologies des flottes de pêche : Méthodes Ifremer-SIH. DRV/SIH/N°4/082003. https://archimer.ifremer.fr/doc/00705/81686/). For example, a vessel used to dredge scallops for five months and to trawl fish for seven months will be allocated to the DCF sub-fleet for trawlers and seiners and to the IFREMER fleet/sub-fleet for non-exclusive trawlers/dredger-trawlers.

- **Inactive** vessel on 31/12 of year N-2 (equivalent to an entry in the 'inactive for fishing' category in the relevant fields in the sub-fleet typologies[7]): if 1, exclusion from the survey frame.

Regarding this last indicator of inactivity, it should be noted that a more up-to-date situational analysis would, of course, be desirable, to preserve the survey frame as much as possible for vessels that were possibly inactive during the year of interest, N-1. However, as mentioned previously, this updated information is not available either at the time when the sampling plan is drawn up, or even as the data is being collected. This gap in time may represent a slight bias due to potential exclusion from the survey frame of vessels that were indeed active during year N-1 and inclusion of vessels that became inactive in that year. Analysis has been conducted, therefore, to assess the likelihood of a vessel becoming active again having been inactive the previous year. This shows a **strong tendency for the status of inactive to continue for some time**, making it reasonable to assume a **minimal bias.**

A vessel's inclusion, or not, in the survey frame is deduced from these five indicators. It can be assumed, therefore, that a vessel can be included if it was present in the fleet on 31/12 of year N-2 and that, according to the most recent information available, it was still operating on the same coastline, with the same owner, and was still active.

2.1.4 Composition of the resulting survey frame

The **fields** that make up the **survey frame** as supplied to the SSP are presented in **summary** in Table1. This survey frame will serve as a reference at different steps of developing the collection exercise.

[7] Vessels practicing marine worm fishing (dredging or diving) are also categorised as 'inactive for fishing'.

Table1. Constituent fields of the survey frame

FIELD	DESCRIPTION
NAVS_COD	Vessel registration code
AN_REF	Reference year
NAVP_LONGUEUR_HT	Length overall (in cm)
LONGUEUR_CLASS_LIB	Length class
QAM_LIB	Name of registration district
QAM_COD	Code for registration district
QAM_RG	Geographical range of district
SEG_SIHECO_LIB	Name of vessel's specific sector
SEG_SIHECO_COD	Code for vessel's specific sector
SEG_SIHECO_LIB	Name of vessel's specific sub-fleet
SEG_SIHECO_COD	Code for vessel's specific sub-fleet
SEG_DCR_LIB	Name of DCR segment
SEG_DCR_COD	Code for DCR segment
SEG_DCR_LIB	Name of DCR sub-fleet
SEG_DCR_COD	Code for DCR sub-fleet
SEG_DCR_Standard	Standard for DCR segment
CARN_EFFECTIF	Average number of crew on board
NB_MOIS_ACT	Number of months of activity
FILEMO	Product of the values of the 3 components NAVP_LONGUEUR_HT, CARN_EFFECTIF and NB_MOIS_ACT (see section 2.2)
ind_Sortie_N_1	Binary indicator of exit from the fleet during year N-1
ind_Present3112_N_2	Binary indicator of presence in the fleet on 31 December, year N-2
ind_ChgtArm_N_1	Binary indicator of a change of owner during year N-1
ind_ChgtFac_N_1	Binary indicator of a change of coastline during year N-1
ind_Inactifs	Binary indicator of inactivity during year N-1
ind_BaseSondage	Binary indicator of the vessel's inclusion in the survey frame (deduced from the 5 previous indicators)

2.1.5 In summary

An annual survey frame for collecting economic data in metropolitan France (with information on activity in year N-1 collected during year N) is established for each of the two coastlines, taking vessels listed in the CFR file in the reference year N-1 as the basis. Individual information provided by the CFR is supplemented by a typological classification of vessels provided by IFREMER and based on activity journals for the year N-2 (the most up-to-date source available at this point in time). This individual information is used to define the criteria for including vessels in the survey frame, to help decide the make-up of the stratifications used, and to provide a basis for calculating how to allocate vessels (FILEMO). The gap in time between when typologies are available and when the sampling plan is formed can lead to a moderate bias in two ways: the manner in which vessels are allocated in a segmentation, and, by extension, their potential for inclusion in the survey. This is because the validity of the inactive vessel indicator, whether positive or negative, carries a degree of uncertainty.

2.2 Calculating adjustment parameters for annual individual income using the FILEMO composite variable

2.2.1 Introduction to the method

The allocation method for stratified collection effort is based on **objectives of precision** applied to the **stratification** under consideration. This involves adjusting the numbers surveyed to a controlled coefficient of variation (CV) of estimators. The adjustment must be applied to an **individual variable that is exhaustive at target population level** and sufficiently correlated to the key indicators to be evaluated from the sampled data that their precision is as close as possible to the initial value. The exhaustive variable chosen and applied in this context will be a **proxy of the annual income,** similar to a power function applied to a specific composite variable (the function therefore resulting from a log-log linear regression). This proxy will be known as an '**auxiliary variable**' or an '**optimisation variable**'.

This composite variable, known as **FILEMO,** is the **product of three individual variables** generated from CFR data and activity journal data: NAVP_LONGUEUR_HTNAVP_LONGUEUR_HT (an individual vessel's length in cm), CARN_EFFECTIF (average number of crew on board the vessel) and NB_MOIS_ACT (number of months of vessel activity[8]) (Van Iseghem, S. *et al.,* 2011). It is incorporated into the survey frame every year and is used to form a **predictive model of individual income** that will produce the **exhaustive proxy**. Every year, this model is adjusted to data on income collected from surveys conducted during years N-2, N-3 and N-4 (remember that data for year N-1 is collected in year N). Two versions of this model are implemented, one for each coastline.

To this end, a **log-log linear adjustment** (linear adjustment of the logarithm for income using the FILEMO logarithm) is applied to data available for the three years considered, generating a pair (a, b) with slope and intercept, such that a+b.log(FILEMO) is a proxy of log (Income). By applying the logarithm, it is possible to linearise the initial 'power' type link which connects the Income/IN variable to FILEMO, thus simplifying the problem posed. Using this linear regression also has the advantage of facilitating the validation process for the adjustment obtained. Every year, an adjustment report is sent to the SSP, describing the two pairs of estimated values (one for coastline AT and one for coastline ME) so that the support variable in the allocation calculation is

[8]A vessel is considered to be active in a given month if it is engaged in fishing activity for at least one day during that month.

evaluated at an individual level across the whole population, and the algorithm that determines allocations for the collection exercise can be started.

2.2.2 Summary analysis of results obtained: applied to AT

The FILEMO adjustment procedure applied to the English Channel-North Sea-Atlantic coastline is described here. The steps are the same when applied to the Mediterranean coastline.

The procedure starts by exploring the link between log(FILEMO) and log(IN), distinguishing the year considered using a simple scatter plot. By **testing the 'year' effect,** it is possible, first and foremost, to validate the adjustment choice over a period of three years by merit of the **stability of the link**, which must be established between the two variables.

Despite a more pronounced distribution of incomes corresponding to the smaller FILEMO values, denoting a weakened correlation for these extreme values, **a clear trend towards the linear link across the points in their entirety** can be distinguished. **The 'year' effect does not appear to stand out, at least not in graph form.** The adjustment is applied to all vessels surveyed during the period year N-5 to year N-3 and results in a **regression line**.

The pair of values thus determined will be used to calculate the proxy of IN from individual FILEMO values for vessels operating on the English Channel-North Sea-Atlantic coastline, based on the following formula:

$IN = e^{a} \times FILEMO^{b}$

where a is the intercept and b is the slope.

It is important to note that the (a,b) pair obtained defines the optimal linear link between log(FILEMO) and log(IN), but once this equality is converted to exponential form (non-linear function) to establish the link between FILEMO and IN, this optimality is only approximate. The value of b arrived at is thus checked to make sure it does not diverge too much from the value of b such that:

- the **linear correlation coefficient** between IN and $FILEMO^{b}$ is **maximal** (remember that this coefficient measures the *linear* link that exists between two quantitative variables);
- the $FILEMO^{b}$ **coefficient of variation** is **similar** to that for IN (remember that the CV of a numeric variable is a relative measure of the dispersal of data around the mean).

The two graphs below (Figure 1) make it possible to verify the validity of the value of b chosen

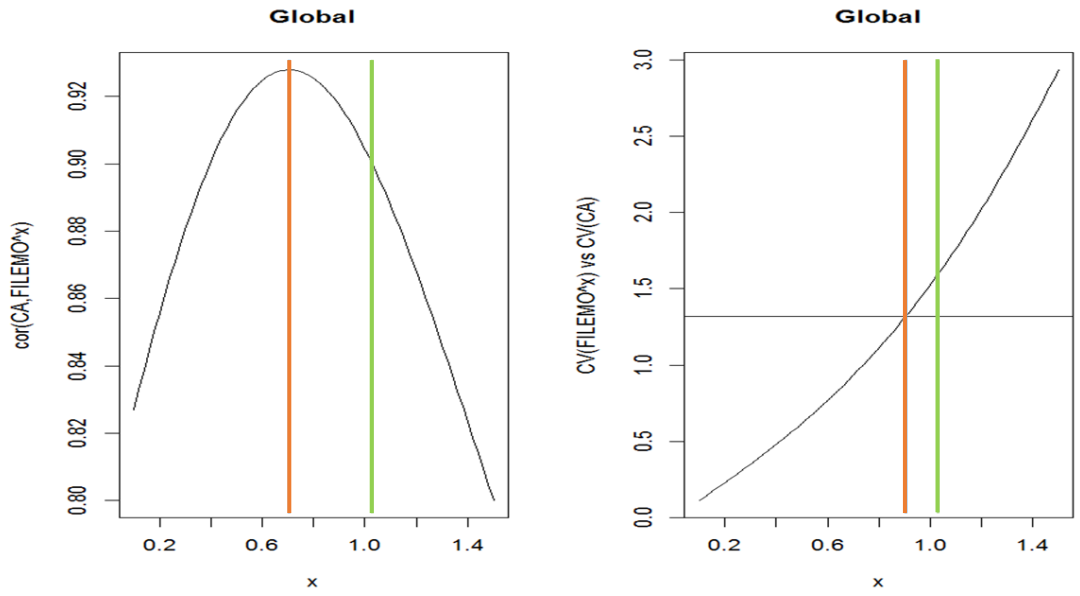

Figure 1. Examples of graphs to validate the chosen parameters (left - in terms of linear correlation cor(IN,FILEMO^X); right - in terms of variation CV(FILEMO^X) vs CV(IN). On both graphs, the 'optimal power' parameter is shown in orange, and the adjusted parameter b in green. (Sources: Administrative data from the 'Community Fishing Fleet Register' (CFR) file on vessel activity, IFREMER fisheries information system (*Système d'Informations Halieutiques*: SIH), and the SSP survey on the production of economic data in the maritime fishing sector).

2.3 Allocation methodology applied to the sampling plan: a problem of Neyman optimisation with constraints

The allocation method used in developing the sampling plan for economic data collection is based on **Neyman optimisation** (minimisation of the variance of the overall mean estimator) **under various constraints**. In terms of the **statistical quality** of indicators produced from the collection, it **determines** the **numbers to be surveyed**, according to different stratifications and depending on specific needs, also known as constraints.

In this case, the priority constraint is one of **local precision**, which can be applied to the initial segmentation characterising the stratified sampling plan (indexed *h* in what follows), or in an aggregated way to a group of these segments (indexed *p)*. This precision relates, most often, to a **coefficient of variation** (CV_p), but can also be expressed as a maximum relative error with a given probability.

Another constraint concerns the **restriction of numbers** allocated per segment/stratum, these being set between a fixed minimum number and a fixed maximum number which is naturally the population of the segment under consideration. An additional constraint, albeit secondary in the context of the sampling plan as it does not impede the set objectives, is a **total fixed allocation**. Most often this concerns a **budgetary**-type constraint. This problem can, therefore, be formalised in the following way:

$$\begin{cases} Min_{n_1,\dots,n_H} V\left(\hat{\bar{Y}}\right) = \sum_h \left(\frac{N_h}{N}\right)^2 . \left(1 - \frac{n_h}{N_h}\right) . \frac{S_h^2}{n_h} & (1) \\ n_{h\,min} \leq n_h \leq N_h, \forall h = 1, \dots, H & (2) \\ \sum_{h=1,\dots,H} n_h = n & (3) \end{cases}$$

where *h* represents the segment, n_h the corresponding allocation, N_h the corresponding total population, *Y* the exhaustive variable considered (so here this is the optimisation variable, the proxy of the income produced using FILEMO; see previous section), and S_h^2 the variance of the Y variable within the segment. The $n_{h\,min}$ value is the allocation per segment *h* used to achieve the set precision objective CV_p, with *p* representing the group of segments to which the precision targeted applies (if the precision objectives concern the initial stratification, this can be written as *p = h*). Here, therefore, it represents the minimum allocation according to the precision constraints required, and conditions the constraint concerning restrictions on the allocation (2). It should be noted that in these conditions, an allocation higher than $n_{h\,min}$ is equivalent to a resulting CV value lower than CV_p, thus improving precision within the group (Koubi et al., 2009). The value *n* (total number of observations) conditions the total allocation constraint (3). Constraint (1) represents the optimisation objective in the context of (2) and (3).

2.4 Drawing and allocating the sample between collection partners

Once the exhaustive variable produced from FILEMO has been established as an approximation of vessels' individual income, and this approximation has been used as an optimisation variable in the process for calculating a stratified allocation, the next step is to draw the sample according to the units thus distributed. This is done through **systematic stratified random sampling per allocation segment** (defined as the combination 'DCF segment' and 'collection operator'). For each stratum, this involves (Figure 2):

- Isolating vessels belonging to the stratum in the survey frame (for example, all trawlers and seiners under 10 m targeted by the IFREMER survey).
- Comparing the total number of vessels in the population in stratum N with the corresponding allocation n.
- **Placing** vessels **in order**, **according to their specific maritime district** (coastal distribution, from north to south for the AT coastline and from west to east for the ME coastline) **and then according to their length** (in centimetres).
- Separating all vessels **into n subsets of similar numbers** ($\left\lfloor \frac{N}{n} \right\rfloor$individuals per subset with an additional individual for $N - n \times \left\lfloor \frac{N}{n} \right\rfloor$subsets divided in a homogenous way over the whole set). For example, if N = 46 and n = 7, numbers for the subsets will be E1 = 7, E2 = 6, E3 = 7, E4 = 6, E5 = 7, E6 = 6 and E7 = 7 (see Figure 2).
- **In an equiprobable way, randomly drawing** an integer i between 1 and$\left\lfloor \frac{N}{n} \right\rfloor$, and selecting the i^{th} item from each subset. The individuals thus selected will form the sample for the stratum considered (another possibility is to draw an integer i between

1 and $\left\lfloor \frac{N}{n} \right\rfloor$+1 and take the ith item from the subset when that is possible, and the (i-1)th if not).

Navire	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24	N25	N26	N27	N28	N29	N30	N31	N32	N33
Quartier	Q1	Q1	Q1	Q1	Q1	Q1	Q1	Q2	Q2	Q2	Q2	Q2	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q3	Q4	Q4	Q4	Q5	Q5	Q5	Q5	Q5
Taille	573	607	692	706	788	919	960	625	772	773	779	795	519	543	549	609	615	622	643	726	765	785	804	935	955	508	686	713	516	628	655	745	784

Figure 2. Method for segmenting the survey frame into subsets according to vessel and quarter before systematic sampling

After random sampling has been applied to each of the strata, a final arbitration step validates (or not) the assignment of vessels to each collection operator. The result is the definitive distribution between the two partners of the vessels to be surveyed.

LEMNA and IFREMER conduct the collection exercise following two distinct and complementary procedures. LEMNA collects data from companies' accounts. IFREMER collects data directly from fishing vessel owners/company bosses by means of surveys conducted by interviewers working along the coastline. Once the sample has been drawn, LEMNA indicates the vessels it needs to collect information from through its network[9]. The sample is then divided between LEMNA and IFREMER.

This collection approach results in a highly differentiated distribution according to vessel size. In effect, the largest vessels all have accounting systems and more chance of being approached by LEMNA (Figure 3). In the Mediterranean, where there are higher numbers of smaller vessels, the share surveyed by IFREMER is noticeably larger than it is on the Atlantic coast.

[9]Their network of partners on the French coastline includes approved management centres, chartered accountancy firms and fishing companies.

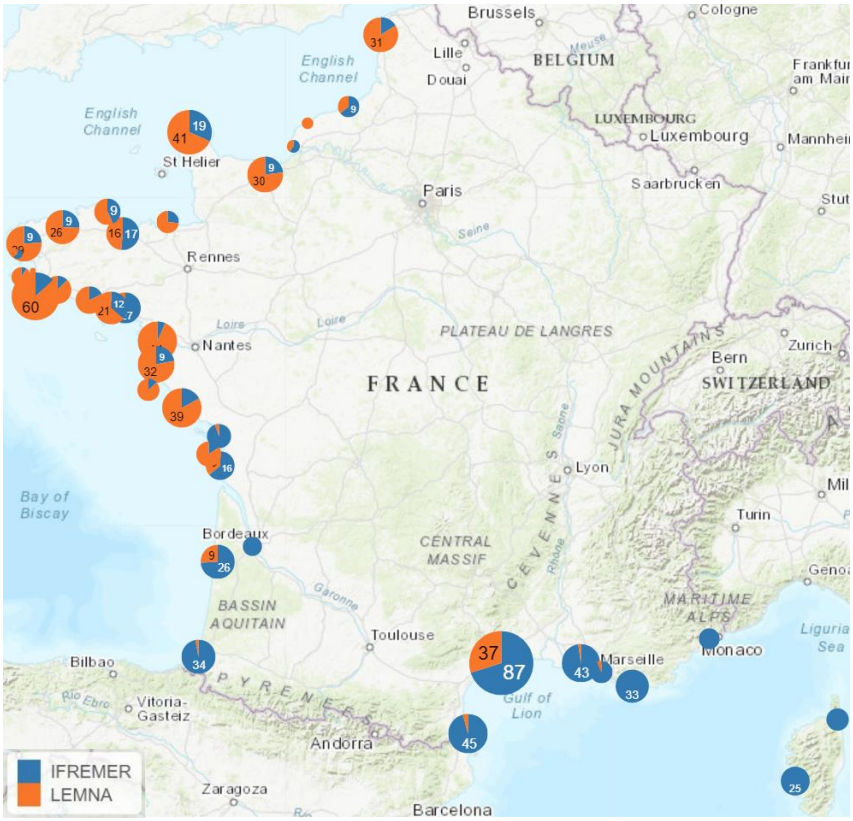

Figure 3. Distribution of vessels 'collected' according to maritime district and research organisation (LEMNA in Orange, IFREMER in blue) in 2017. (Source: Gitton et al., 2020)

Once vessels drawn from the national sampling plan have been distributed between the two partners, each partner can proceed with organising the budgetary and logistical aspects of the data collection exercise, according to the vessels they have been allocated.

3 Bibliography

Gitton François-Pierre, Minne Marie-Dominique, Baranger Laurent, Souffez Arnaud, Guyader Olivier, Le Grand Christelle, Merzereaud Mathieu (2020). Enquête pour la production de données économiques dans le secteur des pêches maritimes 2021-2025. Dossier de présentation au Comité du label de la statistique publique. Séance du 21 octobre 2020. 288p. https://archimer.ifremer.fr/doc/00694/80622/

Koubi, Malik; Mathern, Sandrine Résolution d'une des limites de l'allocation de Neyman, Actes des Xèmes Journées de Méthodologie Statistique, Volume 2009 (2009). http://www.jms-insee.fr/2009/S22_3_ACTE_KOUBI_JMS2009.PDF

Van Iseghem, S., Quillérou, E., Brigaudeau, C., Macher, C., Guyader, O., and Daurès, F. (2011). Ensuring representative economic data: survey data-collection methods in France for implementing the Common Fisheries Policy. – ICES Journal of Marine Science, 68: 1792–17