

# Construction des indicateurs socio-économiques sur la pêche professionnelle : méthodologies appliquées à la métropole de 2009 à 2019

De la collecte des données à la construction des indicateurs





## Fiche documentaire

Titre du rapport : Construction des indicateurs socio-économiques sur la pêche professionnelle : méthodologies appliquées à la métropole de 2009 à 2019

Référence interne : RBE/SIH-UEM/2021-1

Date de publication : 2021/09/30

Version : 1.0.0

Diffusion :

libre (internet)

Référence de l'illustration de couverture

restreinte (intranet) - date de levée d'embargo : RAS

Langue(s) : français

interdite (confidentielle) - date de levée de confidentialité : AAA/MM/JJ

---

### Résumé/ Abstract :

La production d'indicateurs économiques et sociaux sur la pêche professionnelle en France est un enjeu important pour suivre l'évolution de la situation des pêches pour ces flottilles et territoires français et évaluer les conséquences de mesures de gestion sur la durabilité des pêcheries.

Pour la France et au sein du ministère de l'Agriculture et de l'Alimentation (MAA), la Direction des pêches Maritimes et de l'Aquaculture (DPMA) est chargée du programme national de collecte de données (Data Collection Framework) et le Service Statistique et de la Prospective (SSP) coordonne la collecte et la production d'indicateurs économiques et sociaux sur les flottilles de pêches. L'Ifremer est un partenaire et contributeur au programme national de la DCF. Sur le volet économique et social, l'action socio-économie du Système d'Informations Halieutiques de l'Ifremer contribue à la collecte de données économiques à la fois en métropole et dans certaines régions d'outre-mer. Cette collecte sur le terrain est menée chaque année par l'Ifremer depuis l'année 2000 sur les façades Manche - Mer du Nord - Atlantique et depuis l'année 2002 sur la façade Méditerranée.

Ce document vise à décrire et expliciter les méthodologies mises en œuvre par l'Ifremer en vue d'aboutir à l'élaboration des indicateurs socioéconomiques sur la flotte de pêche métropolitaine, depuis la construction du plan d'échantillonnage, en passant par la collecte des données et leur validation et jusqu'au calcul des variables DCF.

Ce document est en accès libre et constitue une référence méthodologique pour l'ensemble des travaux utilisateurs des données de l'enquête nationale. Il fonctionne en binôme avec un autre rapport méthodologique plus détaillé, car à visée opérationnelle, et donc en diffusion interne.

The production of economic and social indicators on professional fishing fleets in France is an important issue to monitor the evolution of the fisheries situation for these fleets and French territories and to evaluate the consequences of management measures on the sustainability of the fisheries.

For France and within the Ministry of Agriculture and Food (MAA), the Directorate of Maritime Fisheries and Aquaculture (DPMA) is in charge of the national data collection program (Data Collection Framework) and the Statistical and Prospective Service (SSP) coordinates the collection and production of economic and social indicators on the fishing fleets. Ifremer is a partner and contributor to the national DCF program.

On the economic and social component, the socio-economic action of the Ifremer Fisheries Information System contributes to the collection of economic data both in Metropolitan France and in some overseas regions. This field collection has been carried out every year by Ifremer since 2000 on the Channel-North Sea-Atlantic coast and since 2002 on the Mediterranean coast.

The aim of this document is to describe and explain the methodologies implemented by Ifremer in order to elaborate the socio-economic indicators on the fishing fleet in Metropolitan France, from the construction of the sampling plan, through the collection and validation of the data, to the calculation of the DCF variables.

This document is freely available and constitutes a methodological reference for all work using the national survey data. It works in tandem with another more detailed methodological report, which is intended to be operational, and is therefore distributed internally.

**Mots-clés/ Key words :**

Pêche professionnelle, France métropolitaine, méthodologie, socio-économie, plan d'échantillonnage, collecte, validation, indicateurs

Professional fishing, Metropolitan France, methodology, socio-economics, sampling plan, collection, validation, indicators

**Comment citer ce document :**

Leonardi Sophie, Le Grand Christelle, Merzereaud Mathieu, Daures Fabienne, Macher Claire, Guyader Olivier (2021). Construction des indicateurs socio-économiques sur la pêche professionnelle : méthodologies appliquées à la métropole de 2009 à 2019. RBE/SIH-UEM/2021-1. <https://doi.org/10.13155/83613>

**Disponibilité des données de la recherche :**

[Enquête](#) sur la production des données économiques dans le secteur des pêches maritimes. Enquête labellisée de qualité statistique et d'intérêt général. Les données sont protégées par le secret statistique. Le comité du secret du conseil national de l'information statistique (CNIS) peut accorder des dérogations d'accès aux données confidentielles dans la mesure où la demande est effectuée à des fins de statistique publique ou de recherche scientifique ou historique

---

DOI : 10.13155/83613

---

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.) : Data Collection Framework (DCF), Système d'Informations Halieutiques, [Enquête](#) sur la production des données économiques dans le secteur des pêches maritimes.

Auteur(s) / adresse mail

Affiliation / Direction /  
Service, laboratoire

Leonardi Sophie / [sophie.leonardi@ifremer.fr](mailto:sophie.leonardi@ifremer.fr)

PDG-RBE-EM

Christelle Le Grand /  
[christelle.le.grand@ifremer.fr](mailto:christelle.le.grand@ifremer.fr)

PDG-RBE-EM

PDG-RBE-EM

Merzéréaud Mathieu /  
[mathieu.merzereaud@ifremer.fr](mailto:mathieu.merzereaud@ifremer.fr)

Fabienne Daurès / [fabienne.daures@ifremer.fr](mailto:fabienne.daures@ifremer.fr)

PDG-RBE-EM

Macher Claire / [claire.macher@ifremer.fr](mailto:claire.macher@ifremer.fr)

PDG-RBE-EM

Guyader Olivier / [olivier.guyader@ifremer.fr](mailto:olivier.guyader@ifremer.fr)

PDG-RBE-EM

Encadrement(s) :

Destinataire :

**Validé par** : Olivier Guyader

### Remerciements

Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence nationale de la Recherche au titre du programme Investissements d'Avenir portant la référence ANR-10-EQPX-17 (Centre d'accès sécurisé aux données – CASD).

## Sommaire

### Table des matières

<b>1</b>	<b>Liste des tableaux</b> .....	<b>7</b>
<b>2</b>	<b>Liste des figures</b> .....	<b>9</b>
<b>3</b>	<b>Liste des acronymes</b> .....	<b>10</b>
<b>4</b>	<b>Contexte et objectifs</b> .....	<b>11</b>
<b>5</b>	<b>Construction du plan d'échantillonnage</b> .....	<b>15</b>
5.1	Constitution de la base de sondage.....	15
5.2	Calculs des paramètres d'ajustement du chiffre d'affaires annuel individuel par la variable composite FILEMO.....	19
5.3	Méthodologie d'allocation appliquée au plan d'échantillonnage : problème d'optimisation de Neyman sous contraintes.....	22
5.4	Tirage de l'échantillon et répartition entre partenaires de collecte.....	23
<b>6</b>	<b>Collecte et validation des données sur le terrain</b> .....	<b>27</b>
6.1	Résumé.....	27
6.2	Un questionnaire unique et son guide d'utilisation.....	28
6.3	Données économiques locales.....	29
6.4	Procédure de collecte.....	30
6.5	Consolidation terrain par les enquêteurs.....	30
<b>7</b>	<b>Méthode de validation</b> .....	<b>33</b>
7.1	Les indicateurs : méthodes de calcul et variables utilisées.....	33
7.2	Les données de l'enquête économique.....	36
7.3	Les étapes de la validation.....	38
<b>8</b>	<b>Conclusion</b> .....	<b>41</b>
<b>9</b>	<b>Références bibliographiques</b> .....	<b>42</b>
<b>10</b>	<b>ANNEXE</b> .....	<b>43</b>
10.1	Annexe 12 : Calcul du coût du personnel.....	43

## 1 Liste des tableaux

TABLEAU 1 : CHAMPS CONSTITUTIFS DE LA BASE DE SONDRAGE .....	18
TABLEAU 2 : AJUSTEMENT GLOBAL DE FILEMO SUR LA FAÇADE MANCHE-MER DU NORD-ATLANTIQUE SUR LA PERIODE 2014-2016 : INDICATEURS RELATIFS A LA REGRESSION LINEAIRE $\text{LOG(CA)}=F(\text{LOG(FILEMO)})$ SOUS R. SOURCES: DONNEES ADMINISTRATIVES DU FICHIER "FLOTTE DE PECHE COMMUNAUTAIRE" (FPC), DONNEES ACTIVITE DES NAVIRES -IFREMER - SYSTEME D'INFORMATIONS HALIEUTIQUES, SSP (SERVICE DE LA STATISTIQUE ET DE LA PROSPECTIVE) -ENQUETE SUR LA PRODUCTION DE DONNEES ECONOMIQUES DANS LE SECTEUR DES PECHEES MARITIMES. ....	21
TABLEAU 3 : VARIABLES COLLECTEES DANS L'ENQUETE ET UTILISEES POUR CALCULER LES VARIABLES SOCIO-ECONOMIQUES REQUISES PAR LA DCF.....	34
TABLEAU 4 : QUESTIONS DE L'ENQUETE UTILISEES POUR PRODUIRE LES VARIABLES DCF : CALCUL + VALIDATION.....	37
TABLEAU 5 : TRAÇABILITE ET POINTEURS DE QUALIFICATION DES DONNEES.....	39
TABLEAU 6 : TRAÇABILITE ET POINTEURS DE CORRECTION DES DONNEES .....	40
TABLEAU 7 : L'ASSIETTE ENIM UTILISEE POUR LE CALCUL DES COTISATIONS .....	43
TABLEAU 8 : COMPOSITION DES COTISATIONS SOCIALES PATRONALES ET SALARIALES.....	43



## 2 Liste des figures

FIGURE 1 : PROCESSUS DE CONSTRUCTION DES VARIABLES DCF A L'ECHELLE NATIONALE ET IMPLICATION DE L'IFREMER .....	12
FIGURE 2 : AJUSTEMENT DE FILEMO SUR LA FAÇADE MANCHE-MER DU NORD-ATLANTIQUE : 2.A. LOG(CA) EN FONCTION DE LOG(FILEMO) PAR ANNEE, 2.B. EN GLOBAL : DROITE DE REGRESSION LINEAIRE $\text{LOG(CA)}=F(\text{LOG(FILEMO)})$ ET INTERVALLE DE PREDICTION A 95% SOURCES : DONNEES ADMINISTRATIVES DU FICHIER "FLOTTE DE PECHE COMMUNAUTAIRE" (FPC), DONNEES ACTIVITE DES NAVIRES -IFREMER - SYSTEME D'INFORMATIONS HALIEUTIQUES, SSP (SERVICE DE LA STATISTIQUE ET DE LA PROSPECTIVE) -ENQUETE SUR LA PRODUCTION DE DONNEES ECONOMIQUES DANS LE SECTEUR DES PECHEES MARITIMES. ....	20
FIGURE 3 : GRAPHES DE VALIDATION DES PARAMETRES RETENUS (A GAUCHE, EN TERMES DE CORRELATION LINEAIRE, A DROITE, EN TERMES DE VARIATION). POUR CHACUN, EN BLEU, LE PARAMETRE « PUISSANCE OPTIMAL », EN VERT LE PARAMETRE B AJUSTE. SOURCES : DONNEES ADMINISTRATIVES DU FICHIER "FLOTTE DE PECHE COMMUNAUTAIRE" (FPC), DONNEES ACTIVITE DES NAVIRES -IFREMER - SYSTEME D'INFORMATIONS HALIEUTIQUES, SSP (SERVICE DE LA STATISTIQUE ET DE LA PROSPECTIVE) -ENQUETE SUR LA PRODUCTION DE DONNEES ECONOMIQUES DANS LE SECTEUR DES PECHEES MARITIMES..	22
FIGURE 4 : MODE DE DECOUPAGE DE LA BASE DE SONDAGE EN SOUS-ENSEMBLES AVANT LE TIRAGE SYSTEMATIQUE.....	24
FIGURE 5 : REPARTITION DES NAVIRES 'COLLECTES' SELON LE QUARTIER MARITIME ET L'ORGANISME ENQUETEUR (LEMNA EN ORANGE, IFREMER EN BLEU) EN 2017. SOURCE : GITTON ET AL., 2020. ....	25
FIGURE 6 : THEMES ABORDES DANS LE QUESTIONNAIRE DES DONNEES SOCIO-ECONOMIQUES. SOURCE : IFREMER (2020).....	28
FIGURE 7 : ETAPES DE LA VALIDATION ET SOURCES DE DONNEES MOBILISEES .....	38

### 3 Liste des acronymes

ASSEDIC : association pour l'emploi dans l'industrie et le commerce

AT : Manche – Mer du Nord – Atlantique

CA : chiffre d'affaires

CAF : caisse d'allocations familiales

CPO : cotisation professionnelle obligatoire

CSTEP : comité scientifique, technique et économique de la pêche

DCF: data collection framework

DPMA : direction des pêches maritimes et de l'aquaculture

EBE : excédent brut d'exploitation

ENIM : établissement national des invalides de la marine

ETP : équivalent temps plein

FPC : flotte de pêche communautaire

LEMNA : laboratoire d'économie et de management de Nantes-Atlantique

ME : Méditerranée

SIH : système d'informations halieutiques

SSM : service social maritime

SSP : service statistique et de la prospective

## 4 Contexte et objectifs

La production d'indicateurs économiques et sociaux sur la pêche professionnelle en France est un enjeu important pour suivre l'évolution de la situation des pêches pour ces flottilles et territoires français et évaluer les conséquences de mesures de gestion sur la durabilité des pêcheries. Ces indicateurs permettent non seulement de décrire la situation socio-économique de ces activités et leur évolution pour une meilleure compréhension des enjeux, mais aussi de modéliser les socio-écosystèmes concernés afin de simuler les impacts de nouvelles réglementations ou d'évènements naturels et anthropiques sur l'ensemble du système. Les pêcheries françaises, au même titre que celles des autres Etats membres, sont concernées par la Politique Commune de la Pêche et soumises au cadre communautaire pour la collecte de données halieutiques dénommé Data Collection Framework<sup>1</sup> définissant le programme pluriannuel de collecte de données devant être mis en œuvre par chaque Etat membre dans un plan de travail national.

Pour la France et au sein du ministère de l'Agriculture et de l'Alimentation (MAA), la Direction des pêches Maritimes et de l'Aquaculture (DPMA) est chargée du programme national de collecte de données (DCF) et le Service Statistique et de la Prospective (SSP)<sup>2</sup> coordonne la collecte et la production d'indicateurs économiques et sociaux sur les flottilles de pêches. L'Ifremer est un partenaire et contributeur au programme national de la DCF<sup>3</sup>. Sur le volet économique et social, l'action socio-économie<sup>4</sup> du Système d'Informations Halieutiques de l'Ifremer<sup>5</sup> contribue à la collecte de données économiques à la fois en métropole et dans certaines régions d'outre-mer. Dans le cadre de la coordination organisée par le SSP, l'Ifremer met en œuvre des enquêtes directes sur le terrain en face à face avec les patrons armateurs sur la base d'un questionnaire unique (Ifremer, AMURE, 2020)<sup>6</sup>. Le Laboratoire d'Economie et de Management de Nantes-Atlantique (LEMNA) collecte des données comptables auprès d'un réseau de centres de gestion. Les données collectées sont notamment restituées sous forme d'indicateurs par segments de pêche dans le rapport économique annuel du Comité scientifique, technique et économique de la pêche<sup>7</sup> (CSTEP).

Cette collecte sur le terrain est menée chaque année par l'Ifremer depuis l'année 2000 sur les façades Manche – Mer du Nord – Atlantique et depuis l'année 2002 sur la façade Méditerranée, et ce, sur un échantillon national représentatif de l'ensemble des flottilles de pêche métropolitaines jusqu'en 2011. L'obtention du label de qualité statistique et d'intérêt général pour l'enquête nationale<sup>8</sup> portée par le SSP en 2012, a permis de répartir l'échantillon national entre les deux partenaires de collecte (Ifremer et Lemna) et d'assurer un respect strict de la

---

<sup>1</sup> <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A32013R1380>

<sup>2</sup> <https://agreste.agriculture.gouv.fr/agreste-web/statistiques/ORGA-SSP/listeTypeStatistique/>

<sup>3</sup> <https://sih.ifremer.fr/Le-SIH/Appui-aux-politiques-publiques/Application-en-France>

<sup>4</sup> <https://sih.ifremer.fr/Activite-socio-economie>

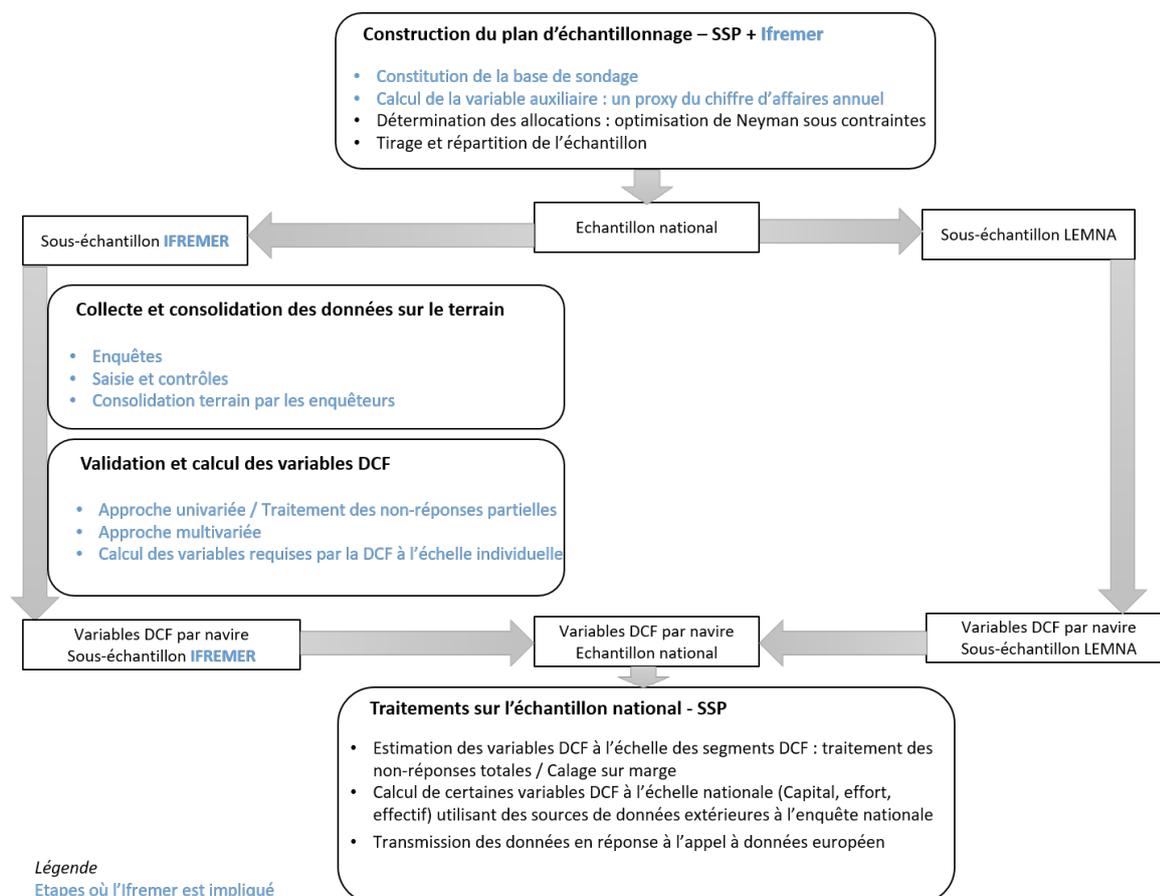
<sup>5</sup> <https://sih.ifremer.fr/>

<sup>6</sup> <https://archimer.ifremer.fr/doc/00646/75841/>

<sup>7</sup> <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/2018-annual-economic-report-eu-fishing-fleet-stecf-18-07>

<sup>8</sup> <https://www.cnis.fr/enquetes/enquete-pour-la-production-de-donnees-economiques-dans-le-secteur-des-peches-maritimes-2020a705ag/?producer=475>

confidentialité des données, protégées par le secret statistique<sup>9</sup>. Ce label a été reconduit en 2015 et tout récemment en 2020 (Gitton et al., 2020)<sup>10</sup> sur la période 2021 - 2025.



**Figure 1 : Processus de construction des variables DCF à l'échelle nationale et implication de l'Ifremer**

Ce document vise à décrire et expliciter les procédures mises en œuvre en vue d'aboutir à l'élaboration des indicateurs socioéconomiques sur la flotte de pêche métropolitaine, depuis la construction du plan d'échantillonnage, en passant par la collecte des données et leur validation et jusqu'au calcul des variables DCF. Ce rapport se concentre sur les travaux qui sont du ressort de l'Ifremer (Figure 1), à savoir la contribution à la construction du plan d'échantillonnage national, en particulier la constitution de la base de sondage et l'ajustement de la variable auxiliaire, ainsi que la collecte, la validation, le traitement des données et le calcul des variables DCF pour les navires du sous-échantillon qui lui est confié chaque année et qu'il restitue à l'échelle individuelle au SSP.

Les travaux menés par le SSP ou le Lemna peuvent être appréhendés à travers le dossier de présentation de l'enquête nationale au Comité du label de la statistique publique (Gitton et al., 2020).

<sup>9</sup> <https://www.cnis.fr/comite-du-secret-statistique/>

<sup>10</sup> <https://archimer.ifremer.fr/doc/00694/80622/>

La construction des indicateurs socio-économiques sur la pêche professionnelle en Guyane et en Guadeloupe est décrite dans deux autres rapports (Leonardi et *al.*, 2020 et Le Grand et *al.*, 2020)<sup>11</sup>.

Ce document est en accès libre et constitue une référence méthodologique pour l'ensemble des travaux utilisateurs des données de l'enquête nationale. Il fonctionne en binôme avec un autre rapport méthodologique plus détaillé, car à visée opérationnelle, et donc en diffusion interne.

---

<sup>11</sup> <https://archimer.ifremer.fr/doc/00649/76105/> et <https://archimer.ifremer.fr/doc/00649/76107/>



## 5 Construction du plan d'échantillonnage

Avant de rentrer dans les aspects méthodologiques, il est utile de rappeler que la construction du plan d'échantillonnage annuel pour la collecte de données économiques sur le secteur de la pêche dans le cadre du programme national en métropole est sous la responsabilité du SSP et implique les partenaires de collecte, à savoir l'Ifremer et le Lemna. Bien que l'Ifremer ne soit en charge que des deux premières étapes de la construction du plan d'échantillonnage, à savoir la construction de la base de sondage ainsi que le calcul des paramètres d'ajustement du chiffre d'affaires annuel individuel par la variable composite FILEMO (Figure 1), il est également proposé dans cette partie une description simplifiée des étapes 3 et 4 mises en œuvre par le SSP permettant ainsi au lecteur de disposer d'une vue d'ensemble du processus.

### 5.1 Constitution de la base de sondage

Tout plan d'échantillonnage nécessite une **base de sondage** décrivant la population étudiée et au sein de laquelle seront tirées les unités à enquêter. Cette base a également vocation à intégrer un ensemble d'informations individuelles qui permettront de répondre aux besoins inhérents aux différentes phases opérationnelles planifiées, allant du tirage aléatoire jusqu'à l'estimation des indicateurs finaux (en passant par la constitution de la stratification, les calculs d'allocation basés sur des objectifs de précision prédéterminés, des procédures d'inférence statistique telles que le calage sur marges, ...).

Dans cette partie seront décrites les étapes de construction de la base de sondage utilisée dans le cadre du plan d'échantillonnage pour la collecte des données économiques en France métropolitaine. **Deux versions distinctes** de cette base seront créées, une pour la **façade Manche-Mer du Nord-Atlantique** (qu'on distinguera par l'abréviation « AT » dans la suite du document, et couvrant les navires opérant en zone FAO 27), une autre décrivant les navires basés en **Méditerranée** (affectés à la zone FAO 37). La construction de ces 2 bases se fonde sur une **méthodologie commune**, leur différenciation reposant simplement sur la population considérée ainsi que sur une caractérisation spécifique de la typologie employée. Dans cette description, on prendra pour **référence temporelle l'année N de mise en œuvre du processus de collecte**, sachant que l'enquête s'attachera à collecter les données portant sur l'année N-1.

Etape n°1 : Mobilisation des deux sources de données utilisées : le fichier flotte FPC et les données d'activité des navires de pêche

Les deux bases de sondage sont issues des mêmes sources de données individuelles exhaustives. Elles prennent en effet majoritairement appui sur le **fichier national de la Flotte de Pêche Communautaire (FPC)**, fichier compilant des informations administratives et des caractéristiques techniques relatives à chaque navire de la flotte de pêche française. Ce fichier sera enrichi d'une typologie avancée mise en place grâce au **recensement exhaustif de l'activité des navires** par le biais d'un travail de reconstitution des [calendriers d'activité](#). Afin de s'assurer de la cohérence des informations contenues dans la base de sondage (constituée en fin d'année N-1) avec la période ciblée par l'enquête (pour rappel, elle aussi l'année N-1), une version actualisée de ce fichier est utilisée. Toutefois, une seconde version de ce fichier, datée au 31 décembre de l'année N-2, est mise à contribution de façon complémentaire sur deux aspects spécifiques. D'une part, elle intègre une **typologie des flottilles validée et exhaustive, la plus actualisée** qui puisse être mise à disposition au moment de la construction de cette base. D'autre part, elle propose une photographie plus ancienne qui, mise en parallèle avec l'image procurée

par le fichier actualisé, permet **d'isoler les navires présents tout au long de l'année N-1**, tout comme ceux qui auront connu durant cette année des changements de situation notables (on verra plus loin que cela constituera des critères d'exclusion de la base de sondage).

### Étape n°2 : Construction de segmentations adaptées à différents besoins de restitution d'indicateurs

Une segmentation est une manière de compartimenter une population de navires en « segments », **un segment étant la combinaison d'une flottille d'appartenance<sup>12</sup> et d'une classe de taille** de navire. On construit, à partir des typologies issues des calendriers d'activité au 31/12 de l'année N-2, des segmentations sur lesquelles reposeront les stratifications structurant le futur plan d'échantillonnage. La finalité de ce dernier étant de construire des indicateurs socioéconomiques en réponse aux besoins de la DCF, mais également des travaux de recherche et d'expertise de l'Ifremer menés à différentes échelles, il est nécessaire de prendre en compte dans le processus de construction des **stratifications** répondant à ces **différents niveaux de restitution attendus**. Ainsi, à chacune des deux façades correspondra un découpage spécifique des classes de taille des navires, auquel on associera deux typologies de flottilles distinctes (une **typologie « DCF »** conforme aux nomenclatures requises par les instances européennes, et une **typologie « SIH »** correspondant davantage au besoin scientifique) pour aboutir finalement à deux segmentations par façade<sup>13</sup>.

### Étape n°3 : Production d'indicateurs individuels servant de critères d'inclusion dans la base de sondage

Comme énoncé précédemment, la mise en parallèle de deux photographies d'une même base d'information prises aux instants N-2 et N-1 va permettre la construction de multiples **indicateurs individuels de changements d'état**, témoignant de la **stabilité** de l'activité de pêche et du **mode de gestion** des navires considérés durant la période ciblée par l'enquête économique. Cette enquête portant sur l'activité opérée au cours d'une année entière, cette notion de stabilité appliquée à différents facteurs significatifs va naturellement discriminer les navires que l'on considèrera comme pertinents à enquêter, et qui composeront ainsi notre base de sondage. Ainsi, l'intégration dans la base sera conditionnée par la combinaison de 5 critères binaires distincts :

- indicateur de **présence en flotte au 31/12 de l'année N-2** (équivalent à la présence d'une valeur de typologie de flottilles renseignée) : si 0, le navire n'aura pas couvert entièrement la « période » N-1 (exclusion) ;

---

12 Une typologie de flottille a pour but de regrouper des navires ayant des comportements ou des stratégies de pêche similaires (mêmes métiers ou combinaisons de métiers), ou selon des critères tels que les caractéristiques techniques des navires ou les secteurs de provenance ou d'activité.

13 Pour rappel, la typologie des flottilles DCF est basée sur l'utilisation d'un engin dominant à l'année tandis que la typologie Ifremer SIH est basée sur la polyvalence des navires du point de vue des engins utilisés (<https://sih.ifremer.fr/Description-des-donnees/Module-Activite-de-peche-et-socio-economie/Activite-des-navires/Typologie-des-flottilles> , [https://sih.ifremer.fr/content/download/8065/54741/file/2003-04\\_Methodo\\_typologie\\_flotte\\_de\\_peche.pdf](https://sih.ifremer.fr/content/download/8065/54741/file/2003-04_Methodo_typologie_flotte_de_peche.pdf)).

A titre d'exemple, un navire pratiquant 5 mois la drague à coquille-St-Jacques et 7 mois le chalut de fond à poissons sera affecté à la flottille DCF des senneurs et chalutiers de fond et aux flottille / sous-flottille Ifremer des chalutiers non exclusifs / chalutiers dragueurs.

- indicateur de **sortie de flotte** en cours d'année N-1 (équivalent à la présence dans le fichier FPC d'une date de référence antérieure à la date de référence majoritaire correspondant à la dernière actualisation de la base) : si 1, le navire est sorti de flotte avant l'actualisation (exclusion) ;
- indicateur de **changement d'armateur** au cours de l'année N-1 (équivalent à un changement de valeur du champ Armateur entre les versions N-2 et N-1 des sources de données) : si 1, exclusion de la base de sondage ;
- indicateur de **changement de façade** au cours de l'année N-1 (équivalent à un changement de valeur du champ Façade entre les versions N-2 et N-1 des sources de données) : si 1, exclusion de la base de sondage ;
- indicateur de navires **inactifs** au 31/12 de l'année N-2 (équivalent à la présence de la modalité « Inactifs à la pêche » dans les champs relevant des typologies de flottille<sup>14</sup>) : si 1, exclusion de la base de sondage.

Il faut noter, concernant ce dernier indicateur d'inactivité, qu'un état des lieux plus actualisé serait bien entendu souhaitable afin de préserver au maximum la base de sondage de navires ayant été éventuellement inactifs durant l'année d'intérêt N-1. Mais comme dit précédemment, cette information actualisée ne peut être obtenue au moment de la constitution du plan d'échantillonnage, ni même au moment de la campagne de collecte. Ce décalage peut représenter un léger biais dans le fait de potentiellement exclure de la base de sondage des navires ayant pourtant été bel et bien actifs durant l'année N-1, tout comme dans le fait d'y avoir peut-être inscrit des navires étant devenus inactifs cette année-là. Une analyse a toutefois été menée à propos de la probabilité d'un navire de redevenir actif après avoir été inactif l'année précédente. Celle-ci montre une **forte tendance du statut d'inactif à perdurer dans le temps**, ce qui peut laisser supposer **un biais minimisé**.

De ces 5 indicateurs, on déduit finalement l'appartenance ou non d'un navire à la base de sondage : on considèrera ainsi qu'un navire peut y être intégré s'il était présent en flotte au 31/12 de l'année N-2 et qu'il l'était toujours selon les dernières informations disponibles, dans la même façade et avec le même armateur, et qu'il était en activité selon les dernières sources d'information disponibles.

#### Composition de la base de sondage résultante

Nous présentons ici le Tableau 1 **synthétique** décrivant les **champs** constituant la **base de sondage** telle que fournie au SSP. Cette base de sondage servira de support de référence aux différentes étapes de mise au point de la campagne de collecte.

---

<sup>14</sup> Sont également compris dans les « inactifs à la pêche » les navires pratiquant la pêche des vers (drague ou plongée).

Tableau 1 : Champs constitutifs de la base de sondage

CHAMPS	DESCRIPTION FR
NAVS_COD	Code d'immatriculation du navire
AN_REF	Année de référence
NAVP_LONGUEUR_HT	Longueur Hors Tout (en cm)
LONGUEUR_CLASS_LIB	Classe de longueur
QAM_LIB	Libellé du quartier d'immatriculation
QAM_COD	Code du quartier d'immatriculation
QAM_RG	Rang géographique du quartier
SEG_SIHECO_LIB	Libellé du segment spécifique d'appartenance
SEG_SIHECO_COD	Code du segment spécifique d'appartenance
F_SIHECO_LIB	Libellé de la flottille spécifique d'appartenance
F_SIHECO_COD	Code de la flottille spécifique d'appartenance
SEG_DCR_LIB	Libellé du segment DCR
SEG_DCR_COD	Code du segment DCR
F_DCR_LIB	Libellé de la flottille DCR
F_DCR_COD	Code de la flottille DCR
SEG_DCR_Standard	Segment DCR standard
CARN_EFFECTIF	Effectif moyen à bord
NB_MOIS_ACT	Nombre de mois d'activité
FILEMO	Produit des valeurs des 3 composantes NAVP_LONGUEUR_HT, CARN_EFFECTIF et NB_MOIS_ACT (voir partie 2.2)
ind_Sortie_N_1	Indicateur binaire de sortie de flotte au cours de l'année N-1
ind_Present3112_N_2	Indicateur binaire de présence en flotte au 31 décembre de l'année N-2
ind_ChgtArm_N_1	Indicateur binaire de changement d'armateur au cours de l'année N-1
ind_ChgtFac_N_1	Indicateur binaire de changement de façade au cours de l'année N-1
ind_Inactifs	Indicateur binaire d'inactivité au cours de l'année N-2
ind_BaseSondage	Indicateur binaire d'inclusion du navire dans la base de sondage (déduit des 5 indicateurs précédents)

## En résumé

Une base de sondage annuelle pour la collecte des données économiques en métropole (collecte opérée au cours de l'année N sur l'activité de l'année N-1) est constituée pour chacune des deux façades maritimes sur la base des navires présents dans le FPC de l'année de référence N-1. Les informations individuelles délivrées par le FPC sont enrichies par une typologie des navires fournie par l'Ifremer et mise en place à partir des calendriers d'activité portant sur l'année N-2 (source la plus actualisée disponible à cet instant). Cet ensemble d'informations individuelles permet ainsi de caractériser les critères d'inclusion des navires dans la base de sondage, de proposer les éléments qui composeront les stratifications utilisées ainsi que les supports pour les calculs d'allocation (FILEMO). Le décalage temporel des typologies disponibles constituera toutefois un biais modéré sur deux plans : la manière dont les navires seront affectés à une segmentation, et par extension, leur potentialité à être enquêtés puisque la validité de l'indice d'inactivité, qu'il soit positif ou négatif, sera empreint d'une relative incertitude.

### 5.2 Calculs des paramètres d'ajustement du chiffre d'affaires annuel individuel par la variable composite FILEMO

#### Présentation de la méthode

La méthode d'allocation d'effort de collecte stratifiée mise en place s'appuie sur des **objectifs de précision visés sur la stratification** considérée. Cet ajustement d'effectifs enquêtés vers un coefficient de variation (CV) d'estimateurs maîtrisé doit s'appliquer sur une **variable individuelle exhaustive au niveau de la population ciblée**, et de plus suffisamment corrélée aux indicateurs clés qui seront évalués à partir des données échantillonnées, afin que leur précision se rapproche autant que possible de la valeur ciblée initialement. La variable exhaustive choisie, et mise à contribution dans ce cadre d'application, sera un **proxy du chiffre d'affaires annuel**, image d'une fonction puissance appliquée à une variable composite particulière (fonction résultant donc d'une régression linéaire log-log). On appellera ce proxy « **variable auxiliaire** », ou encore « **variable d'optimisation** ».

Cette variable composite, appelée **FILEMO**, est le **produit de 3 variables individuelles** issues des données FPC et des données des calendriers d'activité : NAVP\_LONGUEUR\_HT (longueur individuelle du navire, en cm), CARN\_EFFECTIF (effectif moyen annuel embarqué à bord du navire) et NB\_MOIS\_ACT (nombre de mois d'activité du navire<sup>15</sup>) (Van Iseghem, S. et al., 2011). Elle est intégrée chaque année au sein de la base de sondage, et sert à la constitution d'un **modèle prédictif du chiffre d'affaires individuel** qui fournira le **proxy exhaustif**. Ce modèle est ajusté tous les ans sur les données de chiffre d'affaires issues des enquêtes opérées durant les années N-2, N-3 et N-4 (on le rappelle, pour une collecte de données sur N-1 opérée durant l'année N). Deux versions de ce modèle sont implémentées, une pour chaque façade.

Pour cela, sur les données disponibles pour les trois années considérées, on procède à un **ajustement linéaire log-log** (ajustement linéaire du logarithme du CA par le logarithme de FILEMO), générant un couple (a,b) d'intercept et de pente, tel que  $a+b \cdot \log(\text{FILEMO})$  soit un proxy de  $\log(\text{Chiffre d'Affaires})$ . Le passage au logarithme permet de linéariser le lien initial de type « puissance » qui relie la variable Chiffre d'Affaires à FILEMO, simplifiant ainsi la problématique posée. L'utilisation de la régression linéaire présente également l'avantage de faciliter le

---

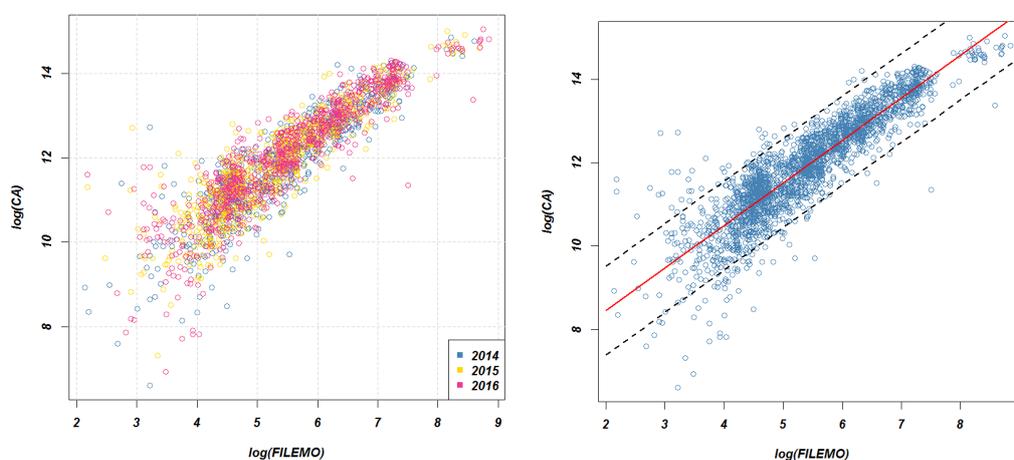
<sup>15</sup> On considère que le navire a été actif à la pêche un mois donné à partir de moment où il a été actif à la pêche au moins 1 jour au cours de ce mois.

processus de validation de l'ajustement obtenu. Un rapport d'ajustement est envoyé chaque année au SSP, décrivant en particulier les deux couples de valeurs estimées (un pour la façade AT, l'autre pour la façade ME), afin que la variable support du calcul d'allocation soit évaluée au niveau individuel sur l'ensemble de la population, et que l'algorithme de détermination des allocations d'effort de collecte puisse ainsi être lancé.

### Synthèse de l'analyse des résultats obtenus au cours de l'analyse : application AT en N=2018

On décrit ici la procédure d'ajustement de FILEMO appliquée à la façade Manche-Mer du Nord-Atlantique. Les étapes demeurent les mêmes pour l'application sur la façade Méditerranée.

On commence par explorer le lien entre  $\log(\text{FILEMO})$  et  $\log(\text{CA})$ , en distinguant l'année considérée, au moyen d'un simple nuage de points (voir Figure 2). **Tester l'effet « Année »** permet en premier lieu de valider le choix d'ajustement sur une période de 3 années par la **stabilité de la relation** qui doit être établie entre les deux variables.



**Figure 2 : Ajustement de FILEMO sur la façade Manche-Mer du Nord-Atlantique : 2.a.  $\log(\text{CA})$  en fonction de  $\log(\text{FILEMO})$  par année, 2.b. en global : droite de régression linéaire  $\log(\text{CA})=f(\log(\text{FILEMO}))$  et intervalle de prédiction à 95%**

Sources : Données administratives du fichier "Flotte de Pêche communautaire" (FPC), Données Activité des navires - Ifremer - Système d'Informations Halieutiques, SSP (Service de la Statistique et de la Prospective) -Enquête sur la production de données économiques dans le secteur des pêches maritimes.

Malgré une dissémination plus prononcée des chiffres d'affaires (CA) correspondant aux plus petites valeurs de FILEMO (partie gauche du graphe), dénotant une corrélation fragilisée pour ces valeurs extrêmes, on distingue déjà une **nette tendance au lien linéaire sur la globalité des points**. **L'effet « année » ne semble pas ressortir, du moins graphiquement**. L'ajustement s'exerce sur la globalité des navires enquêtés sur la période 2014-2016, et aboutit à la **droite de régression** illustrée et décrite dans la Figure 2 et le Tableau 2 (traitements réalisés sous le logiciel R).

**Tableau 2 : Ajustement global de FILEMO sur la façade Manche-Mer du Nord-Atlantique sur la période 2014-2016 : Indicateurs relatifs à la régression linéaire  $\log(CA)=f(\log(FILEMO))$  sous R.**

Sources: Données administratives du fichier "Flotte de Pêche communautaire" (FPC), Données Activité des navires - Ifremer - Système d'Informations Halieutiques, SSP (Service de la Statistique et de la Prospective) -Enquête sur la production de données économiques dans le secteur des pêches maritimes.

```

Call:
lm(formula = log(CA) ~ log(FILEMO_sih), data = DF)

Residuals:
Min          1Q          Median          3Q          Max
-3.0937    0.3191    -0.2769    0.0221    3.3073

Coefficients:
              Estimate          Std. Error    t value    Pr(>|t|)
(Intercept)  6.41408          0.05680    112.92    <2e-16 ***
log(FILEMO_sih)  1.01967          0.01028    99.19    <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5438 on 2271 degrees of freedom
Multiple R-squared:  0.8125,    Adjusted R-squared:  0.8124
F-statistic: 9839 on 1 and 2271 DF, p-value: < 2.2e-16

```

Le couple de valeurs ainsi déterminé sera utilisé pour calculer le proxy du CA à partir des valeurs individuelles FILEMO pour les navires appartenant à la façade Manche-Mer du Nord-Atlantique, à partir de la formule :

$$CA = e^a \times FILEMO^b$$

où a est l'intercept et b la pente.

Il ne faut pas négliger que le couple obtenu (a,b) est bien le couple définissant le lien linéaire optimal entre  $\log(FILEMO)$  et  $\log(CA)$ , mais qu'une fois l'égalité passée à l'exponentielle (fonction non linéaire) pour établir le lien entre FILEMO et CA, cette optimalité n'est alors plus qu'approximative. Nous vérifions alors que la valeur de b trouvée ne s'écarte pas trop de la valeur de b telle que :

- le **coefficient de corrélation linéaire** entre CA et  $FILEMO^b$  est **maximal** (on rappelle que ce coefficient permet de mesurer la liaison *linéaire* qui existe entre deux variables quantitatives) ;
- le **coefficient de variation** de  $FILEMO^b$  est **similaire** à celui de CA (on rappelle que le CV d'une variable numérique est une mesure relative de la dispersion des données autour de la moyenne).

Les deux graphiques ci-dessous (cf. Figure 3) permettent de vérifier la validité de la valeur de b retenue (pour rappel,  $b=1.01967$ ).

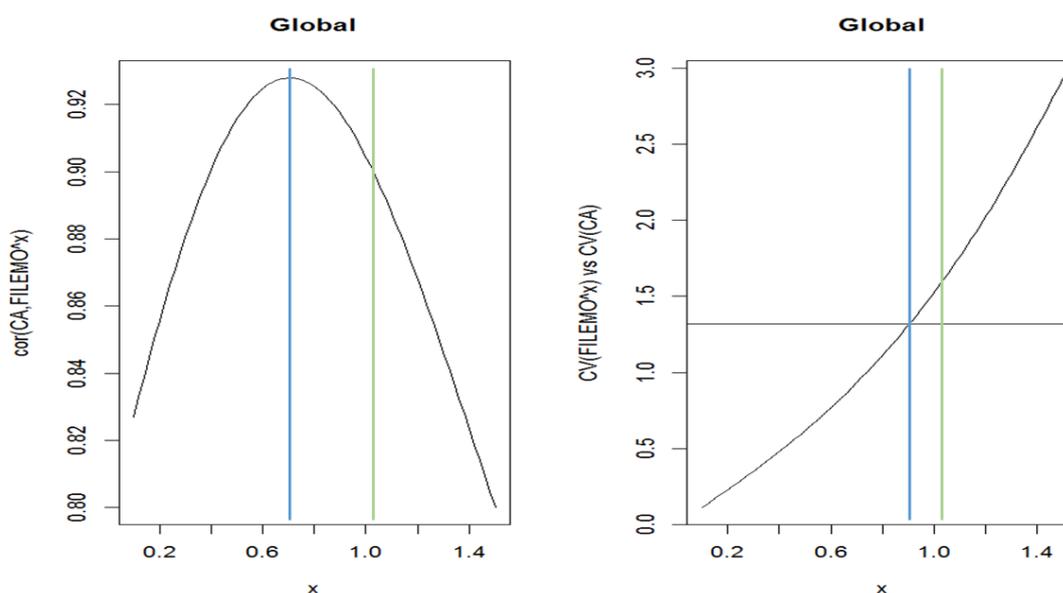


Figure 3 : Graphes de validation des paramètres retenus (à gauche, en termes de corrélation linéaire, à droite, en termes de variation). Pour chacun, en bleu, le paramètre « puissance optimal », en vert le paramètre  $b$  ajusté.  
Sources : Données administratives du fichier "Flotte de Pêche communautaire" (FPC), Données Activité des navires - Ifremer - Système d'Informations Halieutiques, SSP (Service de la Statistique et de la Prospective) -Enquête sur la production de données économiques dans le secteur des pêches maritimes.

### 5.3 Méthodologie d'allocation appliquée au plan d'échantillonnage : problème d'optimisation de Neyman sous contraintes

La **méthode d'allocation** utilisée dans le cadre de la construction du plan d'échantillonnage pour la collecte des données économiques est basée sur **une optimisation de Neyman** (minimisation de la variance de l'estimateur moyen global) **sous diverses contraintes**. Elle vise à **déterminer** selon une stratification donnée **les effectifs à enquêter** en fonction de besoins particuliers, également appelées contraintes, quant à la **qualité statistique** des indicateurs issus de la collecte.

La contrainte prioritaire dans notre cas sera une contrainte de **précision locale**, qui pourra s'appliquer au sein de la segmentation initiale caractérisant le plan d'échantillonnage stratifié (indexée  $h$  dans ce qui suit), ou bien encore au niveau agrégé d'un regroupement de ces segments (indexé  $p$ ). Cette précision se rapporte le plus souvent à un **coefficient de variation** ( $CV_p$ ), mais peut tout aussi bien s'exprimer en tant qu'erreur relative maximale avec une probabilité donnée.

Une autre contrainte repose sur **l'encadrement des effectifs** alloués par segment/strate, ceux-ci étant compris entre un effectif minimal fixé, et un effectif maximal qui est naturellement la population du segment considéré. Enfin, une troisième contrainte, secondaire dans le cadre de notre plan d'échantillonnage car non limitante sur les objectifs fixés, est une **allocation totale fixe**. Cette dernière découle le plus souvent d'une contrainte de type **budgétaire**. On peut donc formaliser ce problème de la façon suivante :

$$\left\{ \begin{array}{l} \text{Min}_{n_1, \dots, n_H} V(\hat{Y}) = \sum_h \left(\frac{N_h}{N}\right)^2 \cdot \left(1 - \frac{n_h}{N_h}\right) \cdot \frac{S_h^2}{n_h} \quad (1) \\ n_{h \min} \leq n_h \leq N_h, \forall h = 1, \dots, H \quad (2) \\ \sum_{h=1, \dots, H} n_h = n \quad (3) \end{array} \right.$$

où  $h$  désigne le segment,  $n_h$  l'allocation correspondante,  $N_h$  la population totale correspondante,  $Y$  la variable exhaustive considérée (donc ici notre variable d'optimisation, le proxy du chiffre d'affaires issu de FILEMO ; voir chapitre précédent), et  $S_h^2$  la variance de la variable  $Y$  au sein du segment. La valeur  $n_{h \min}$  est l'allocation par segment  $h$  permettant d'atteindre l'objectif de précision fixé  $CV_p$ , avec  $p$  désignant le regroupement de segments sur lequel porte la précision visée (si nos objectifs de précision portent sur la stratification initiale, on peut poser  $p=h$ ). Elle incarne donc bien ici l'allocation minimale selon les contraintes de précision à atteindre, et conditionne la contrainte (2) d'encadrement de l'allocation. Il faut noter dans ces conditions qu'une allocation supérieure à  $n_{h \min}$  équivaut à une valeur de CV résultant inférieure à  $CV_p$ , c'est-à-dire à une précision améliorée au sein du regroupement (Koubi et al, 2009). La valeur  $n$  (nombre total d'observations) conditionne quant à elle la contrainte (3) d'allocation totale. Enfin, la contrainte (1) représente l'objectif d'optimisation dans le cadre de (2) et de (3).

#### 5.4 Tirage de l'échantillon et répartition entre partenaires de collecte

Une fois la variable exhaustive issue de FILEMO construite comme approximation du chiffre d'affaires individuel des navires, et une fois celle-ci utilisée comme variable d'optimisation dans la procédure de calcul des allocations stratifiées, l'étape suivante va consister à tirer l'échantillon selon les effectifs ainsi distribués. On procède pour cela à un **tirage aléatoire systématique stratifié par segment d'allocation** (défini comme la combinaison « segment DCF » et « opérateur de collecte »). Celui-ci consiste, pour chaque strate d'appartenance à (cf. Figure 4):

- isoler dans la base de sondage les navires appartenant à la strate (par exemple, tous les chalutiers et senneurs de fond de moins de 10m visés par l'enquête Ifremer) ;
- rapprocher le nombre total de navires de la population dans cette strate  $N$  à l'allocation correspondante  $n$  ;
- **ordonner** les navires **selon le quartier maritime d'appartenance** (répartition côtière, du nord au sud pour la façade AT, d'ouest en est pour la façade ME) **puis selon la longueur** quantitative du navire (en cm) ;
- **séparer** l'ensemble des navires **en  $n$  sous-ensembles d'effectifs similaires**  $\left(\frac{N}{n}\right)$  individus par sous-ensemble, avec un individu supplémentaire pour  $N - n \times \left\lfloor \frac{N}{n} \right\rfloor$  sous-ensembles répartis de manière homogène sur l'ensemble). Par exemple, si  $N=46$  et  $n=7$ , alors les effectifs des sous-ensembles seront  $E_1=7, E_2=6, E_3=7, E_4=6, E_5=7, E_6=6$  et  $E_7=7$  (voir schéma ci-dessous) ;
- **tirer aléatoirement de manière équiprobable** un entier  $i$  entre 1 et  $\left\lfloor \frac{N}{n} \right\rfloor$ , et sélectionner le  $i$ -ème élément de chaque sous-ensemble. Les individus ainsi sélectionnés constitueront l'échantillon pour la strate considérée (une autre

possibilité consiste à tirer un entier  $i$  entre 1 et  $\left\lfloor \frac{N}{n} \right\rfloor + 1$ , et de prendre le  $i$ -ème élément du sous-ensemble lorsque cela est possible, le  $(i-1)$ -ème sinon).

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	Q2	Q2	Q2	Q2	Q2	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 4 : Mode de découpage de la base de sondage en sous-ensembles avant le tirage systématique

Une fois le tirage aléatoire effectué sur chacune des strates, un dernier arbitrage est réalisé afin de valider (ou non) l’assignation des navires à chaque opérateur de collecte, aboutissant finalement à la répartition définitive des navires à enquêter entre les deux partenaires.

La collecte est réalisée par le Lemna et l’Ifremer selon deux procédures distinctes et complémentaires. Le Lemna collecte les données à partir des comptabilités d’entreprise. L’Ifremer collecte directement les données auprès des patrons/armateurs des navires de pêche par voie d’enquêtes réalisées par des enquêteurs répartis sur le littoral. Quand l’échantillon est tiré, le Lemna indique pour quels navires il doit pouvoir collecter l’information grâce à la mobilisation de son réseau<sup>16</sup>. L’échantillon est ensuite réparti entre le Lemna et l’Ifremer.

Il découle de ce mode de collecte une répartition très différenciée selon la taille. En effet, les plus gros navires tiennent tous une comptabilité et ont beaucoup plus de chance d’être approchés par le Lemna (cf. Figure 5). En Méditerranée, où les navires de petite taille sont beaucoup plus fréquemment, la part de navires enquêtée par l’Ifremer est sensiblement plus importante que sur la façade Atlantique.

<sup>16</sup> Leur réseau de partenaires situés sur tout le littoral français, constitués de Centres de gestion agréés, de Cabinets d’expertise comptable ou d’Armements à la pêche.

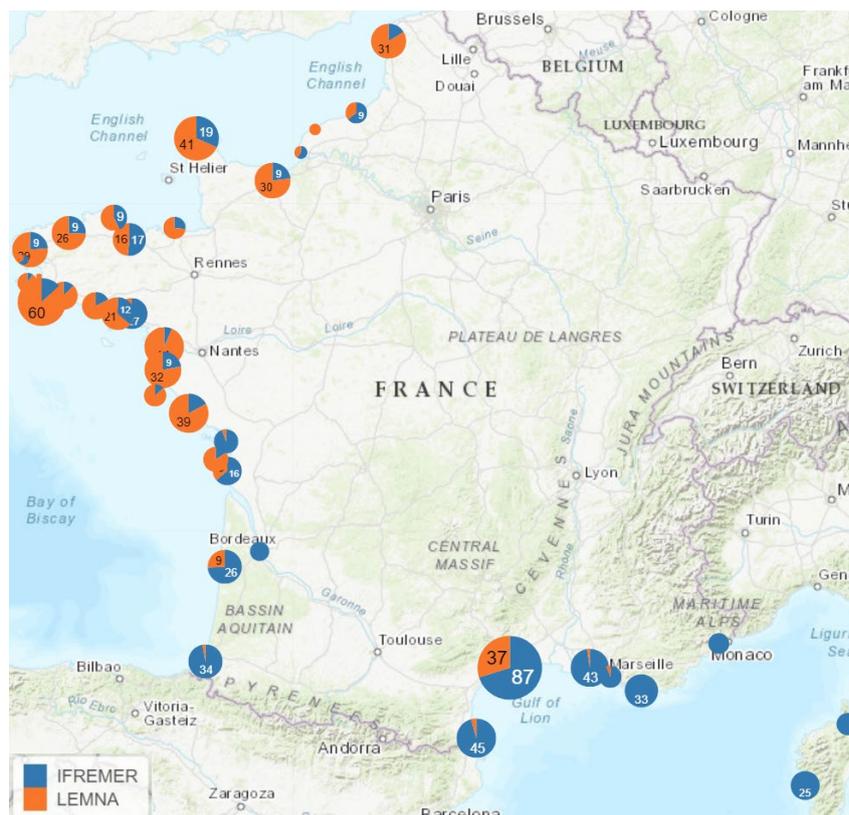


Figure 5 : Répartition des navires ‘collectés’ selon le quartier maritime et l’organisme enquêteur (Lemna en orange, Ifremer en bleu) en 2017.  
 Source : Gitton et al., 2020.

Une fois la répartition des navires tirés dans le plan d’échantillonnage national entre les deux partenaires de collecte, peut alors commencer pour chacun des partenaires de collecte le travail d’organisation budgétaire et logistique de la campagne de collecte en fonction des navires leur ayant été alloués.



## 6 Collecte et validation des données sur le terrain

### 6.1 Résumé

Un questionnaire unique est **administré en face à face** par des enquêteurs sur le terrain auprès des patrons et/ou armateurs de pêche. Il est accompagné d'un **guide** dont l'objectif est de préciser ce qui est attendu derrière chaque question et d'apporter ainsi une aide au remplissage du questionnaire.

Au-delà du questionnaire permettant de collecter des données à l'échelle du navire de pêche, un formulaire (Ifremer 2019) de « **données économiques locales** » est également à compléter par les enquêteurs dès le début de la campagne de collecte. Il a vocation à collecter des données, à **l'échelle des ports d'immatriculation**, relatives aux **prix unitaires des divers avitaillements et matériels**. Ces données sont utiles pour **compléter le questionnaire** lorsque les enquêtés ne disposent pas du détail des prix unitaires ainsi que pour mieux **appréhender le contexte socio-économique local** en rencontrant les acteurs clefs (coopératives, fournisseurs, comités des pêches...).

Sur le terrain, la collecte est opérée par des observateurs salariés de l'Ifremer ou d'un prestataire. Sur les façades Manche – Mer du Nord – Atlantique et Méditerranée, entre 700 et 800 navires sont tirés chaque année par façade pour un taux de réponse moyen réel de 35% à 40%. L'expérience nous fait comptabiliser **un jour de travail par enquête** (démarchage téléphonique, déplacement, réalisation de l'enquête, relecture, saisie et validation inclus).

Afin d'assurer la qualité des données collectées, d'une part, **une procédure de collecte** est remise aux enquêteurs détaillant les étapes de la mise en œuvre des enquêtes socio-économiques depuis la prise de contact avec les patrons pêcheurs jusqu'à la restitution des questionnaires papier à l'Ifremer. D'autre part, **les enquêteurs** sont également en charge d'une **phase de consolidation terrain** qui s'opère à travers quatre étapes.

En outre, des enquêteurs bien formés favorise également une collecte de qualité. La formation a pour objectif **l'appropriation** et la bonne compréhension du questionnaire et des différents outils et documents techniques par ces derniers. Elle inclut une **formation théorique**, une **formation technique à l'halieutique** ainsi qu'un **accompagnement sur le terrain** des nouveaux enquêteurs.

Par ailleurs, un **suivi régulier** de la réalisation des enquêtes depuis le lancement jusqu'à la clôture de la campagne est piloté par le coordinateur de l'action socio-économie. Il permet de suivre la **bonne réalisation des enquêtes** en adéquation avec le plan d'échantillonnage, de **répondre aux questions** des enquêteurs et d'apporter des solutions aux **difficultés** rencontrées.

Afin de favoriser la participation des pêcheurs professionnels de manière pérenne à l'enquête socio-économique, plusieurs actions de communication sont mises en œuvre chaque année.

L'ensemble des documents techniques sont mis à disposition des enquêteurs sur le site web du SIH. Ils incluent le logiciel FESTIF qui permet **la saisie des données d'enquête** mais également **le suivi de la réalisation du plan d'échantillonnage**.

## 6.2 Un questionnaire unique et son guide d'utilisation

La version 2020 du questionnaire unique (Ifremer, AMURE, 2020) utilisé en métropole sur les façades Manche-Mer du Nord-Atlantique et Méditerranée est issue de nombreuses étapes de construction à savoir :

- la création **d'une première version du questionnaire en 2000** par l'équipe du Service d'Economie Maritime (SEM) répondant au cadre communautaire pour la collecte de données halieutiques initiée en 2000<sup>17</sup> mais aussi aux travaux de recherche et d'expertise du service sur cette thématique ;
- **l'actualisation chaque année** du questionnaire en fonction des mises à jour du règlement européen relatif à la collecte de données halieutiques, des besoins des chercheurs et des retours des enquêteurs ;
- la prise en compte de la **labellisation en 2012, en 2015 et en 2020** de l'enquête nationale devenue 'Enquête d'intérêt général et de qualité statistique sur la production des données économiques dans le secteur des pêches maritimes' et impliquant une mise en forme spécifique (encart de la première page, logos...).



Figure 6 : Thèmes abordés dans le questionnaire des données socio-économiques.

Source : Ifremer (2020).

Les questions présentes dans ce questionnaire de 20 pages répondent à différentes finalités (cf. Figure 6) :

- **renseigner les variables requises par le cadre communautaire** en vigueur pour la collecte de données halieutiques<sup>18</sup>, la plupart des variables étant construites à partir de plusieurs questions (par exemple, les 'autres coûts variables' sont obtenus en sommant les frais de débarquement, les coûts de glace, d'appâts, de vivre et d'huiles) ;

<sup>17</sup>[https://datacollection.jrc.ec.europa.eu/c/document\\_library/get\\_file?uuid=f01aff47-e0e7-4c18-9364-bbc53beea54a&groupId=10213](https://datacollection.jrc.ec.europa.eu/c/document_library/get_file?uuid=f01aff47-e0e7-4c18-9364-bbc53beea54a&groupId=10213)

<sup>18</sup> <https://datacollection.jrc.ec.europa.eu/legislation/current/obligations>

- **valider ces variables** (par exemple, la part du chiffre d'affaires pêche par espèce, circuit de commercialisation ou criée/hors criée, permettent de valider et consolider le chiffre d'affaires pêche annuel) ;
- récupérer des **informations qualitatives complémentaires** utiles pour les travaux de recherche : par exemple, les signes de différenciation utilisés du type labels, pin's, marque collective qui permet d'en savoir plus sur les démarches de valorisation des produits mises en œuvre par les pêcheurs professionnels.

Ce questionnaire est **administré en face à face** par des enquêteurs sur le terrain auprès des patrons et/ou armateurs de pêche.

Ce questionnaire s'accompagne d'un **guide** (Leonardi et al., 2019) à destination des enquêteurs en charge de l'administrer. Il a pour objectif de préciser ce qui est attendu derrière chaque question et apporte ainsi une aide au remplissage du questionnaire. Ce document de 34 pages à vocation à être alimenté au gré des enquêtes et passe en revue l'ensemble des questions dans l'ordre d'apparition du questionnaire. Il contient également des annexes qui précisent d'une part comment **désagréger des revenus et des coûts** lorsque ceux-ci sont agrégés sur plusieurs navires et/ou activités, et d'autre part, comment renseigner les différentes sections du questionnaire dans **les cas où la pêche professionnelle embarquée n'est pas l'unique activité pratiquée par le patron sur le navire tiré**, et notamment :

- CAS 1: Autre activité que la pêche en mer embarquée : **pêche à pied** ;
- CAS 2: Autre activité que la pêche en mer : **conchyliculture** ;
  - CAS 2 - a: Pratique de la pêche en mer et de la conchyliculture **avec le même navire** ;
  - CAS 2 - b: Pratique de la pêche en mer et de la conchyliculture **avec deux navires distincts** ;
- CAS 3: le navire de pêche est utilisé pour **d'autres activités en mer**. Ce cas concerne les navires utilisés pour la pêche-promenade, location, remorquage, collecte de déchets...dans le cas d'activités conchylicoles, se reporter au cas 2-a ;
- CAS 4: Activité principale: pêche en mer / Activité secondaire: **commercialisation de produits de la mer** autres que ceux débarqués par le navire enquêté. Ce cas concerne les pêcheurs qui achètent des produits de la mer: moules, crevettes de Madagascar, tourteaux... (en plus de leur propre pêche) pour achalander leur étal ;
- CAS 5: le propriétaire du navire exploite/possède **plusieurs navires de pêche** à l'année. Le cas d'un propriétaire qui possède plusieurs navires de pêche et qui peut être patron embarqué sur tout ou partie de ses navires ;
- CAS 6: le propriétaire du navire est **matelot embarqué sur un autre navire de pêche**. Ce cas concerne les armateurs embarqués à bord de leur navire comme patron de pêche mais également à bord d'autres navires dont ils ne sont pas propriétaires en tant que matelot.

### 6.3 Données économiques locales

Au-delà du questionnaire permettant de collecter des données à l'échelle du navire de pêche, un formulaire (Ifremer, 2019) de « données économiques locales » est également à compléter par les enquêteurs dès le début de la campagne de collecte. Il a vocation à collecter des données, à **l'échelle des ports d'immatriculation** ou regroupement de ces ports d'immatriculation, relatives

aux **prix unitaires** des **divers avitaillements et matériels** : carburant, glace et huile, licences de pêche, taxes et frais de débarquements et engins de pêche.

Ces données sont utiles :

- pour **compléter le questionnaire** lorsque les enquêtés ne disposent pas du détail des prix unitaires ;
- pour mieux **appréhender le contexte socio-économique local** en rencontrant les acteurs clefs (coopératives, fournisseurs, comités des pêches...).

#### 6.4 Procédure de collecte

La mise en œuvre des enquêtes socio-économiques consiste en une succession d'étapes depuis la prise de contact avec les patrons pêcheurs jusqu'à la restitution des questionnaires papier à l'Ifremer, telles que :

- **compléter** le fichier des **données économiques locales** (prix du carburant, de l'huile, de la glace, ...) par quartier maritime ;
- **contacter l'ensemble des patrons pêcheurs** tirés dans le plan d'échantillonnage et les solliciter pour la réalisation d'une enquête sur la base du volontariat ;
- **réaliser chaque année un maximum d'enquêtes économiques** auprès des patrons pêcheurs, réparties sur les quartiers maritimes de la façade Méditerranée, selon le plan d'échantillonnage fourni par l'Ifremer et les méthodologies définies ;
- **saisir les données** collectées dans le logiciel "FESTIF" mis à disposition par l'Ifremer ;
- **pointer les enquêtes réalisées, les refus, les enquêtes en cours** dans le logiciel "FESTIF" qui est doté d'un outil de suivi de la réalisation des enquêtes en conformité avec le plan d'échantillonnage. Ce fichier de pointage doit être envoyé tous les 15 jours au coordinateur opérationnel de l'Ifremer à partir du démarrage du programme d'enquêtes et ce jusqu'à la fin du programme ;
- valider la cohérence des variables métiers, nombre de jours de mer annuel et effectif moyen annuel renseignés dans l'enquête économique avec ceux de l'enquête activité ;
- **valider la cohérence** de la variable **chiffre d'affaires** renseignée dans l'enquête économique avec les autres sources disponibles dans la base de données Harmonie (Ventes en criées, SACROIS) ;
- **restituer** à l'Ifremer **les questionnaires papier** sous une forme lisible et exploitable, en veillant à ce que les données saisies dans "FESTIF" correspondent strictement aux informations inscrites dans les formulaires ;
- **compléter le fichier « retour enquêteur »** afin d'assurer un retour du déroulement de la saison d'enquête auprès du coordinateur opérationnel.

#### 6.5 Consolidation terrain par les enquêteurs

La validation des données d'enquête collectées et saisies permet d'en assurer la qualité. En amont de la phase de validation opérée par l'ingénieur statisticien de l'équipe (cf. partie 4), une phase de consolidation terrain par l'enquêteur est également opérée à travers les quatre étapes décrites ci-après.

## Prédocumentation et approche experte de l'enquêteur

La prédocumentation consiste à **mettre à disposition des enquêteurs des indicateurs à l'échelle individuelle** construits à partir de **sources de données extérieures à l'enquête** et déjà disponibles pour renseigner, parfois de manière partielle ou sur l'année antérieure, un certain nombre de variables collectées à travers l'enquête. Ainsi les sources de données disponibles au sein du SIH permettent de renseigner en amont des enquêtes les variables suivantes :

- les caractéristiques techniques du navire ;
- des informations sur l'armateur et notamment s'il exploite plusieurs navires et/ou s'il pratique d'autres activités que la pêche avec le navire tiré ;
- l'effectif moyen annuel (source Calendriers d'activité<sup>19</sup>) ;
- le nombre de jours de mer annuel (sources SACROIS<sup>20</sup> et Ventas) ;
- la durée moyenne des marées (source SACROIS) ;
- le volume et la valeur des débarquements annuels (sources SACROIS et Ventas) ;
- les cinq principales espèces débarquées à l'année en valeur (sources SACROIS et Ventas) ;
- une vue historique de la participation du navire aux campagnes d'enquêtes socio-économiques (tiré, enquêté, motif du refus) ;
- des indicateurs économiques moyens pour la façade\*flottille\*classe de longueur de navire ;
- les informations qualitatives non labellisées extraites de la dernière enquête économique sur le navire.

La mise à disposition de ces données à l'enquêteur lui permet de **se forger une certaine expertise** à l'échelle de chaque navire mais aussi par segment de flotte. Il est ensuite à même de prendre du recul sur les données transmises par l'enquêté et/ou sur la qualité des sources extérieures pour chaque navire enquêté. Il est invité à **transmettre un maximum de commentaires relatifs à la qualité des données** via la saisie du commentaire final sous le logiciel FESTIF.

### Contrôles à la saisie

Au cours de la saisie sous le logiciel FESTIF, un certain nombre de contrôles existent permettant de :

- vérifier la cohérence des données saisies pour un navire ;
- identifier les valeurs nulles ou manquantes ;
- identifier des valeurs saisies au-dessus ou en deçà de valeurs de référence seuils.

Ces contrôles, étant bloquants pour la saisie ou de simples messages informatifs, permettent **d'alerter l'enquêteur** en charge de la saisie de **réels ou éventuels erreurs et oublis de saisie**.

### Relecture post-saisie

A l'issue de la saisie, le logiciel FESTIF met à disposition de l'enquêteur un récapitulatif de l'ensemble des informations saisies pour **relecture immédiate** et **correction des éventuels erreurs et oublis**.

---

<sup>19</sup> <https://sih.ifremer.fr/Activite-socio-economie/Activite-des-navires>

<sup>20</sup> <https://sih.ifremer.fr/Debarquements-effort-de-peche/Sacrois>

## Validation croisée par l'enquêteur

La dernière étape consiste à **mobiliser l'expertise de l'enquêteur terrain** pour nous faire un retour sur un certain nombre de données qui ressortiraient comme possiblement erronées.

Pour ce faire, l'ingénieur statisticien en charge de la validation des données fait tourner un certain nombre de vérifications sur les fichiers de saisie des enquêteurs et **pointe des potentielles incohérences** (entre variables, entre plusieurs sources de données, par rapport à des seuils ou par rapport à l'historique), en particulier sur les variables chiffre d'affaires, métiers, nombre de jours de mer et effectif embarqué. Ces dernières sont recensées par navire et transmises à l'enquêteur via un fichier Excel pour :

- Validation et explicitation des différences ou écarts ;
- Correction.

Cette étape est prévue *a minima* 15 jours avant la fin des enquêtes terrain. A la demande des enquêteurs, plusieurs sessions de ce type peuvent être organisées permettant une validation au fil de l'eau des enquêtes.

Cette **approche individuelle** qui **croise la connaissance des données**, de leurs traitements et algorithmes d'une part avec **l'expertise terrain** d'autre part est très enrichissante et nous permet de mieux prendre en compte la **réalité du terrain** mais aussi de pointer **les limites des données et algorithmes** dont on dispose avec parfois la capacité de les faire évoluer.

## 7 Méthode de validation Méthode de validation

La présente section détaille les étapes de validation mises en œuvre par l’Ifremer jusqu’à la restitution des variables DCF à l’échelle du navire pour le sous-échantillon affecté à l’Ifremer.

Ce travail s’inscrit dans un processus de validation plus global (Figure 1) qui s’étend de la collecte et des contrôles réalisés par les enquêteurs<sup>21</sup> jusqu’à la restitution des indicateurs agrégés en réponse à l’appel à données de la Commission européenne par le SSP (Gitton et *al.*, 2020).

### 7.1 Les indicateurs : méthodes de calcul et variables utilisées

Les indicateurs requis dans le cadre de la DCF sont précisés dans le Tableau 3 suivant.

Ils sont calculés de manière individuelle pour chaque navire à l’échelle annuelle, à partir des données de l’enquête économique et de données complémentaires. Les données individuelles sont restituées au Service de la statistique et de la prospective (SSP) du ministère de l’agriculture qui opère la suite des traitements (y compris le calcul à l’échelle des segments).

---

<sup>21</sup> Cf. partie 3.6 ‘Consolidation terrain’.

Tableau 3 : Variables collectées dans l'enquête et utilisées pour calculer les variables socio-économiques requises par la DCF

Catégorie	Variable DCF	Unité	Restitution individuelle Ifremer	Variables de l'enquête utilisées
<b>Effort</b>	Jours en mer	Jours	x	Nombre de jours de mer
	Consommation d'énergie	Litres	x	volume de carburant
<b>Emploi</b>	Effectif embarqué	Nombre	x	Effectif moyen
	ETP	Nombre	x	Effectif moyen, nombre de jours de mer
	Nombre total d'heures travaillées par an	Nombre		
	Main-d'œuvre non rémunérée	Nombre		
<b>Revenus</b>	Valeur brute des débarquements	Euro	x	Chiffre d'affaires pêche
	Revenus provenant de la location de quotas ou d'autres droits de pêche	Euro		
	Subventions d'exploitation	Euro		
	Autres revenus	Euro		
	Subventions à l'investissement	Euro		
<b>Dépenses</b>	Consommation de capital fixe	Euro		
	Coûts de l'énergie	Euro	x	coût du carburant
	Autres coûts non variables	Euro	x	coût des engins, autres dépenses d'armement, prime d'assurance, licences, cotisation centre de gestion
	Autres coûts variables	Euro	x	huile, appâts, vivres, glace, taxes de débarquement
	Coût du personnel	Euro	x	Coût du personnel
	Coûts d'entretien et de réparation	Euro	x	Coût d'entretien et réparations
	Paievements au titre de la location de quotas ou de droits de pêche	Euro		
<b>Capital</b>	Valeur de la main-d'œuvre non rémunérée	Euro		
	Valeur des quotas et des autres droits de pêche	Euro		
	Investissements	Euro		
	Taux d'endettement à court/long terme	Euro	x	part de l'emprunt dans le financement du navire
	Valeur de remplacement du capital	Euro		
	Total des actifs	Euro		

Il est précisé ci-après, le cas échéant, les méthodes de calcul des indicateurs restitués par l'Ifremer à l'échelle du navire pour le sous-échantillon qui lui a été confié par le SSP.

- **Valeur des débarquements**
- **Nombre de jours de mer**
- **Effectif embarqué**

L'Ifremer restitue cette variable, mais elle est également calculée par le SSP, à l'échelle des segments, sur l'ensemble de la population nationale.

- **Equivalent Temps Plein (ETP)**

$$ETP = \min(\text{effectif moyen}; \frac{\text{effectif moyen} * \text{nombre de jours de mer}}{250})$$

L'équivalent temps plein (ETP) par navire est calculé à partir de l'effectif moyen, du nombre de jours de mer et du nombre de jours de mer référent. Un arrêté du code du travail maritime précise que le temps plein d'un marin ne peut dépasser 250 jours, aussi, si le nombre de jours de mer du navire est supérieur à 250 jours, la valeur retenue pour l'ETP est l'équipage moyen.

Le nombre de jours de mer référent est le standard européen : 250 jours. L'Ifremer restitue cette variable, mais elle est également calculée par le SSP, à l'échelle des segments, sur l'ensemble de la population nationale.

- **Volume et coût du carburant**
- **Coût de la réparation**
- **Autres coûts non variables**

*Autres coûts non variables*

$$= \text{Coût engin} + \text{cotisation centre gestion} + \text{prime d'assurance} \\ + \text{autres dépenses d'armement} + \text{coût des licences} \\ + \text{cotisation professionnelle obligatoire}$$

L'indicateur « Autres coûts non variables » est la somme des différents coûts fixes collectés dans l'enquête économique : le coût des engins, la cotisation au centre de gestion, la prime d'assurance du navire, les autres dépenses d'armement, le coût des licences et la cotisation professionnelle obligatoire (non collectée dans l'enquête mais estimée en appliquant le taux de la cotisation à la base de l'Établissement national des invalides de la marine (ENIM)).

- **Autres coûts variables**

*Autres coûts variables = huile + appâts + vivres + glace + taxes de débarquement*

L'indicateur « Autres coûts variables » est constitué des coûts de l'huile, des appâts, des vivres, de la glace et des taxes de débarquement.

- **Coût du personnel**

Le coût du personnel inclus les salaires, les cotisations sociales patronales (ENIM, Caisse d'allocations familiales (CAF), Association pour l'emploi dans l'industrie et le commerce (ASSEDIC), taxe d'apprentissage, Service Social Maritime (SSM)), les congés payés et les primes. Il provient soit de l'enquête, lorsqu'il est disponible, soit d'un calcul réalisé à partir des données de l'enquête (effectif, nombre de jours de mer, catégorie ENIM de l'équipage, part équipage) et des taux de cotisations (source ENIM).

Ces données permettent de calculer l'assiette ENIM, base des calculs des cotisations sociales. Le détail de ces calculs est en annexe 10.1.

Le coût du personnel est alors calculé de la manière suivante :

$$\text{Coût du personnel} = CA - (\text{Taxes Deb.} + \text{Carbu} + \text{Huile} + \text{Appâts} + \text{Vivres} + \text{Glace}) \\ \times \text{Part}_{\text{Equipage}} \\ + \text{Cotisations Sociales}_{\text{patronales}} + \text{congés payés}$$

- **Valeur de remplacement du capital**

L'Ifremer ne restitue pas les variables de capital, elles sont calculées par le SSP.

Pour paramétrer le modèle d'estimation du capital PIM (Perpetual inventory method), le SSP s'appuie notamment sur les valeurs d'achat du navire et les valeurs d'assurance, collectées dans l'enquête économiques (Daurès et Le Grand, 2015).

- **Taux d'endettement**

$$\text{Financement du navire} = x\% \text{ autofinancement} + y\% \text{ subventions} + z\% \text{ emprunt}$$

Le taux d'endettement est estimé par la part de l'emprunt ( $z$ ) dans le financement du navire, lors de l'achat.

## 7.2 Les données de l'enquête économique

Le Tableau 4 présente les questions de l'enquête qui sont utilisées pour produire les indicateurs socio-économiques requis par la DCF. Il contient les variables utilisées pour le calcul mathématique des indicateurs et les questions dont les réponses permettent de valider (comparaison, cross-validation) la valeur calculée de l'indicateur.

Tableau 4 : Questions de l'enquête utilisées pour produire les variables DCF : calcul + validation

Variable DCF	Questions de l'enquête utilisées pour produire les variables DCF	Thème du questionnaire
Valeur des débarquements	<b>5.1 Chiffre d'affaires</b> 5.2 Part criée 5.3 Répartition du CA hors criée 5.5 Détail du CA par espèce 5.6 Cotisation caisse intempérie 3.1 Effort de pêche 3.2 Principaux métiers 5.13 Diversification sans utiliser le navire 5.14 Diversification en utilisant le navire 5.16 Retraite 1.2 Travail en association de navires 1.3 Exploitation autres navires 1.5 Activité professionnelle du conjoint	5. Recettes et coûts 5. Recettes et coûts 5. Recettes et coûts 5. Recettes et coûts 5. Recettes et coûts 3. Activité 3. Activité 5. Recettes et coûts 5. Recettes et coûts 5. Recettes et coûts 1. Information sur l'enquêté et son entourage 1. Information sur l'enquêté et son entourage 1. Information sur l'enquêté et son entourage
Jours de mer	<b>3.1 Effort de pêche</b> 5.13 Diversification sans utiliser le navire 5.14 Diversification en utilisant le navire 3.2 principaux métiers 3.3 durée moyenne d'une marée	3. Activité 5. Recettes et coûts 5. Recettes et coûts 3. Activité 3. Activité
Effectif embarqué	<b>6.2 Effectif moyen annuel</b> 6.1 Patron seul à bord 6.3 Rotation d'équipage 6.4 Nombre de marins étrangers	6. Equipage et mode de rémunération 6. Equipage et mode de rémunération 6. Equipage et mode de rémunération 6. Equipage et mode de rémunération
ETP	<i>Calculé à partir de Effectif embarqué et Nombre de jours de mer</i>	
Consommation et coût de l'énergie	<b>5.8 consommations intermédiaires</b> 3.1 Effort de pêche	5. Recettes et coûts 3. Activité
Coûts d'entretien et de réparation	<b>5.17 entretien et réparation du navire</b> 5.18 Changement de moteur	5. Recettes et coûts 5. Recettes et coûts
Autres coûts non variables	<b>5.19 autres dépenses d'armement</b> <b>2.8 prime et valeur d'assurance</b> <b>3.4 licences</b> <b>4.1 engins</b> <b>4.2 gréements</b> <b>5.10 cotisation centre de gestion</b> <b>5.11 CPO (cotisation professionnelle obligatoire)</b>	5. Recettes et coûts 2. Le Navire 3. Activité 4. Engin de pêche et gréements 4. Engin de pêche et gréements 5. Recettes et coûts 5. Recettes et coûts
Autres coûts variables	<b>5.8 consommations intermédiaires</b> <b>5.7 taxes de débarquement</b> 5.9 adhésion OP (organisation de producteurs)	5. Recettes et coûts 5. Recettes et coûts 5. Recettes et coûts
Coût du personnel	<b>6.7 coût total annuel de l'équipage</b> 6.4 Nombre de marins étrangers 6.5 rémunération à la part 6.6 Primes pour l'équipage 6.8 Détail des charges de personnel 6.9 Salaire d'un matelot 6.10 Catégorie ENIM de l'équipage 1.1 Statut du patron	6. Equipage et mode de rémunération 6. Equipage et mode de rémunération 1. Information sur l'enquêté et son entourage
Valeur de remplacement du capital	2.1 année d'acquisition du navire 2.2 construction neuve 2.4 degré de possession du navire 2.5 prix d'achat 2.7 remboursement de prêts 2.8 prime et valeur d'assurance 2.9 valeur du navire sur le marché de l'occasion	2. Le Navire 2. Le Navire 2. Le Navire 2. Le Navire 2. Le Navire 2. Le Navire 2. Le Navire
Taux d'endettement	<b>2.6 part de l'emprunt dans le financement du navire</b>	2. Le Navire

### 7.3 Les étapes de la validation

Les différentes étapes de la validation sont détaillées ci-après, incluant une partie relative à la traçabilité des données : des pointeurs de qualification et de correction sont définis pour chaque variable validée afin de tracer les éventuelles corrections réalisées.

Pour ce faire, des routines de validation sont déployées mobilisant les données d'enquêtes saisies et consolidées par les enquêteurs, d'autres sources de données du SIH<sup>22</sup> ainsi que des données économiques validées historiques (Figure 10) :

- une approche univariée qui permet de qualifier et corriger les valeurs pour chaque variable : traitement des non-réponses partielles ;
- une approche multivariée qui apporte une lecture globale experte de l'ensemble des navires de l'échantillon ;
- le calcul des variables requises par la DCF à l'échelle individuelle.

Ce travail de validation est réalisé sur la façade Atlantique et sur la Méditerranée.

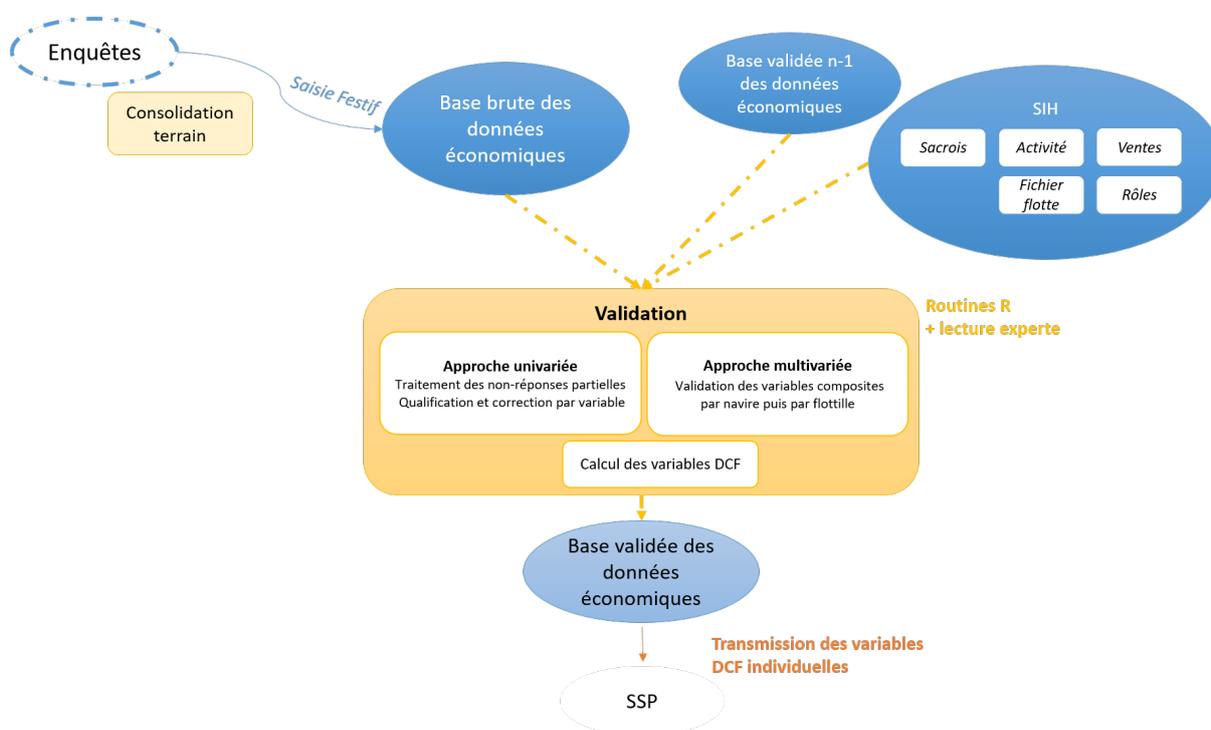


Figure 7 : Etapes de la validation et sources de données mobilisées

#### Approche univariée – Traitement des non-réponses partielles

##### 1.1.1.1 Validation et qualification des données individuelles

Les variables de l'enquête sont qualifiées : les valeurs manquantes, extrêmes ou aberrantes sont repérées par comparaison de la valeur du navire avec :

- les valeurs de référence d'un groupe de navires homogène (segment ou autre niveau de restitution) du navire pour l'année considérée ;

<sup>22</sup> <https://sih.ifremer.fr/Donnees/Catalogue-de-donnees>

- des référentiels historiques constitués pour des groupes de navires homogènes ;
- des indicateurs pouvant provenir d'autres sources de données disponibles à l'échelle de chaque navire (par exemple données d'effort de pêche et de production déclarées pour le navire considéré, regroupées dans la base SACROIS).

Des routines de validation permettent notamment d'identifier des valeurs potentiellement aberrantes comme par exemple :

- à l'échelle d'un segment  $Valeur_n < 0.85 * P5 Valeur_{Segt}$  ou  $Valeur_n > 1.15 * P95 Valeur_{Segt}$  ;
- en procédant à des calculs de cohérence de coûts ( $Valeur_n / Chiffre\ d'affaires_n > x\%$ ) ;
- ou encore en vérifiant certains ratios (par exemple Chiffre d'affaires / homme / par jour de mer, coût du carburant / Chiffre d'affaires, ...).

Les valeurs à vérifier sont tracées et codées de la manière suivante :

Tableau 5 : Traçabilité et pointeurs de qualification des données

Pointeur	Signification	Commentaire
0	Valeur non qualifiée	Valeur par défaut au début de la période de travail
1	Valeur bonne	Valeur qui, a priori, a pu sembler douteuse, mais qui se révèle, après vérification, bonne
2	Valeur statistique hors	Peu utilisé
3	Valeur douteuse	Utilisé pour les valeurs déclarées par les patrons qui semblent douteuses, aberrantes
4	Valeur fausse	Utilisé uniquement pour les erreurs de l'enquêteur : erreurs de saisie, erreurs de calculs
9	Valeur manquante	Absence de données pour la variable et le navire concerné.

Lorsqu'un navire présente de nombreuses incohérences, il est identifié et son enquête est vérifiée afin de trouver des éléments qui expliquent d'éventuelles valeurs aberrantes. Le navire peut alors être écarté du calcul des référentiels par groupe de navires utilisés pour le traitement de la non réponse partielle pour certains indicateurs ou complètement supprimé de l'échantillon s'il est jugé globalement peu fiable.

#### 1.1.1.2 Traitement des non-réponses partielles et des valeurs qualifiées de fausses

Les non-réponses partielles (valeurs manquantes) et les valeurs qualifiées de douteuses (pointeur 3) ou fausses (pointeur 4) à l'étape précédente sont traitées par des méthodes d'imputation :

- soit en utilisant les données de l'enquête (erreur de saisie ou utilisation d'autres données de l'enquête) qui permettent de déduire la valeur ;
- soit par déduction de la valeur en utilisant d'autres sources de données disponibles à l'échelle de chaque navire (activité, effort et production de chaque navire, ...) ;
- soit par imputation d'une valeur de référence calculée au niveau de la flottille ;
- soit en utilisant une valeur historique supposée fixe dans le temps (par exemple valeur d'assurance), provenant d'une enquête antérieure du navire concerné.

Un système de pointeurs trace les corrections réalisées (Tableau 6).

Tableau 6 : Traçabilité et pointeurs de correction des données

Pointeur	Signification	Commentaire
0	Valeur non corrigée	Valeur par défaut au début de la période de travail
5	Valeur corrigée par imputation	Valeur imputée
6	Valeur mise en cohérence	Valeur qui existe dans l'enquête (erreur de saisie ou de calcul) ou qui peut être directement déduite des informations de l'enquête
7	Valeur corrigée manuellement	Valeur déduite d'une source externe de correction ou experte
8	Valeur panel	Utilisation de la valeur déclarée l'année précédente ou récente (3 ans maximum) si le navire a déjà été enquêté

#### Approche multivariée / Calcul des variables DCF

La validation variable par variable est terminée. La dernière étape de la validation est la lecture globale, experte, de l'ensemble des navires de l'échantillon.

Les variables définies par la DCF ainsi que des indicateurs économiques composés de plusieurs variables validées sont calculés (ex. : coûts du personnel, salaire net par homme, CA/homme/Jmer, part du carburant sur le CA, excédent brut d'exploitation (EBE), ...). Les valeurs aberrantes sont repérées :

- CA/Homme/Jour de mer < 100 ou CA/Homme/Jour de mer > 1100 ;
- CA < 8000€ ;
- coût du personnel / CA < 25% ou Coût du personnel / CA > 75% ;
- EBE / CA < - 25% ou EBE / CA > 50% ;
- salaire < 0.

Les enquêtes qualifiées de « mauvaise » par l'enquêteur et les navires pour lesquels de nombreuses imputations ont été réalisées sont également vérifiés. Les tendances des variables par flottille sont calculées en utilisant les présentes données, cumulées à la base de données validées historique.

Les navires « douteux » sont observés sur l'ensemble de leurs indicateurs, par rapport aux autres navires de leur flottille. Ils sont soit conservés dans l'échantillon (cohérence des diverses informations de l'enquête), soit jugés peu fiable (il est alors préférable de les sortir de l'échantillon).

L'échantillon final restitué au SSP est composé des variables requises par la DCF sur cet échantillon.

## 8 Conclusion

Ce rapport décrit à un instant t les méthodologies mises en œuvre par l’Ifremer pour les étapes qui lui incombent dans le processus global depuis la collecte des données jusqu’au transfert des variables requises par la DCF à l’échelle individuelle au SSP. Ces méthodologies ont vocation à évoluer en fonction du règlement DCF et des groupes de travail qui y sont rattachés ainsi que des travaux de recherche utilisateurs de ces données.

## 9 Références bibliographiques

Daurès Fabienne, Le Grand Christelle (2015). Contribution à l'estimation du capital à la pêche à partir de séries de prix à la construction collectées par échantillon - Façade Mer du Nord Manche Atlantique.

Dentes De Carvalho Gaspar, N., Keatinge, M. et Guillen Garcia, J. rédacteur(s) en chef. Le rapport économique annuel 2018 sur la flotte de pêche de l'UE. (CSTEP 18-07) EUR 28359 EN, Office des publications de l'Union européenne, Luxembourg, 2018, ISBN 978-92-79-79390-5, doi:10.2760/56158, JRC112940.

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/>

Ifremer (2019). Enquête économique Secteur Pêche – France. Données économiques locales.

Ifremer (2020). L'économie de la pêche - Participez à l'enquête ! <https://archimer.ifremer.fr/doc/00349/46054/>

Ifremer, AMURE (2020). Enquête d'intérêt général et de qualité statistique. Enquête sur la production des données économiques dans le secteur des pêches maritimes. Année de référence 2019. Période de collecte : mars à juin 2020. Questionnaire Économique - N°2020-01. <https://archimer.ifremer.fr/doc/00646/75841/>

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](http://www.jms-insee.fr/2009/S22_3_ACTE_KOUBI_JMS2009.PDF)

Le Grand Christelle, Merzereaud Mathieu, Leonardi Sophie, Guyader Olivier (2020). Indicateurs socio-économiques sur la pêche professionnelle. Guyane et Guadeloupe. Guide méthodologique. Rapport Ifremer-RBE-SIH-EM, 24 p. <https://archimer.ifremer.fr/doc/00649/76107/>

Leonardi Sophie, Daurès Fabienne, Guyader Olivier, Macher Claire, Le Grand Christelle, Merzereaud Mathieu (2019). Guide questionnaire socio-économique SIH. Aide au remplissage du questionnaire « Enquête sur la production des données économiques dans le secteur des pêches maritimes » Programme de collecte SIH 2019. SIH/N°2019-01

Leonardi Sophie, Le Grand Christelle, Merzereaud Mathieu, Bettali Tiphaine, Blanchard Fabian, Mansuy Emmanuel, Cisse Abdoul, Guyader Olivier (2020). Méthodologie de collecte des données socio-économiques sur la pêche professionnelle. Guyane et Guadeloupe. Année 2020. Rapport Ifremer-RBE-SIH-EM-BIODIVHAL. 18p. <https://archimer.ifremer.fr/doc/00649/76105/>

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

## 10 ANNEXE

### 10.1 Annexe 12 : Calcul du coût du personnel

Il s'agit dans un premier temps de calculer l'assiette ENIM sur laquelle vont s'appliquer les taux des différentes cotisations sociales. L'assiette ENIM se calcule à partir des catégories ENIM des matelots. On distingue la catégorie ENIM du patron (lorsqu'il est propriétaire embarqué) et celle du reste de l'équipage.

$$\text{Assiette ENIM}_{\text{patron}} = \text{cat. ENIM}_{\text{patron}} \times \text{Base ENIM}_{\text{jour}} \times \text{Nb jours de mer}$$

$$\text{Assiette ENIM}_{\text{Equip}} = \text{cat. ENIM}_{\text{Reste Equip}} \times \text{Base ENIM}_{\text{jour}} \times \text{Nb jours de mer} \times (\text{effectif} - 1)$$

$$\text{Assiette ENIM}_{\text{Totale}} = \text{Assiette ENIM}_{\text{patron}} + \text{Assiette ENIM}_{\text{Equip}}$$

Différentes taxes vont ensuite être calculées sur la base de l'assiette ENIM équipage ou Totale. Les cotisations salariales et patronales sont composées des taxes précisées dans le Tableau 24 et le Tableau 25 suivants.

**Tableau 7 : L'assiette ENIM utilisée pour le calcul des cotisations**

Cotisation	Assiette ENIM		
	Equip	Patro	Tota
	.	n	l
ENIM salariales			x
ENIM patronales propriétaire embarqué		x	
ENIM patronales reste de l'équipage	x		
CAF			x
CSG RDS			x
Assedic patronales			x
Assedic salariales			x
CPO (taxes comité)			x
Taxes d'apprentissage			x

**Tableau 8 : Composition des cotisations sociales patronales et salariales**

	Cotisations sociales	
	Salariale	Patronale
	s	s
ENIM salariales	x	
ENIM patronales propriétaire embarqué		x
ENIM patronales reste de l'équipage		x
CAF		x
CSG RDS	x	
Assedic patronales		x
Assedic salariales	x	
CPO (taxes comité)		
Taxes d'apprentissage		x
SSM		x

Le salaire forfaitaire et les congés payés se déduisent facilement à partir de ces informations :

$$\text{Salaire Forfaitaire Net} = \text{Assiette ENIM}_{\text{totale}} - \text{Cotisations Sociales}_{\text{salariales}}$$

$$\text{Congés Payés} = \text{Salaire Forfaitaire Net} / 10$$

Avec ces informations, il est possible de calculer le coût du personnel et les salaires moyens.

$$\begin{aligned} \text{Coût du personnel} &= CA - (\text{Taxes Deb.} + \text{Carbu} + \text{Huile} + \text{Appâts} + \text{Vivres} + \text{Glace}) \\ &\quad \times \text{Part}_{\text{Equipage}} \\ &\quad + \text{Cotisations Sociales}_{\text{patronales}} + \text{congés payés} \end{aligned}$$

$$\begin{aligned} \text{Salaire moyen brut}_{\text{homme}} &= \text{Coût du personnel} - (\text{Cotisations Sociales}_{\text{patronales}} + \text{congés payés}) \\ &\quad \div \text{effectif} \end{aligned}$$

$$\begin{aligned} \text{Salaire moyen net}_{\text{homme}} &= \text{Coût du personnel} - (\text{Cotisations Sociales}_{\text{patronales}} + \text{congés payés}) \\ &\quad - \text{Cotisations Sociales}_{\text{salariales}} \\ &\quad \div \text{effectif} \end{aligned}$$