

COMPARAISON DE PLANS D'ÉCHANTILLONNAGE STRATIFIÉS ADAPTÉS AUX DONNÉES D'ACTIVITÉ DE PÊCHE PROFESSIONNELLE

Angelina El Ghaziri¹, Lise Bellanger², Nicolas Rollo³ & Brice Trouillet³

¹*Université de Nantes, Maison des Sciences de l'Homme-USR3491, 5 allée Jacques Berque, BP 12015, 44021 Nantes Cedex 1.
angelina.el-ghaziri@univ-nantes.fr*

²*Université de Nantes, Laboratoire de Mathématiques Jean Leray-UMR CNRS 6629, bât.10 bureau 116, UFR Sciences et Techniques, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex3
lise.bellanger@univ-nantes.fr*

³*Université de Nantes, LETG-Nantes, Géolittomer-UMR CNRS 6554, BP 81227, 44312 Nantes Cedex 3.
nicolas.rollo@univ-nantes.fr; brice.trouillet@univ-nantes.fr*

Résumé. Dans un contexte de partage croissant d'espace maritime, la spatialisation de l'activité de pêche est un enjeu scientifique et d'aménagement important, surtout pour l'activité des petits navires qui ne sont pas équipés d'un système de suivi par satellite. Le GIS VALPENA (éVALuation des Pratiques de Pêches au regard des Nouvelles Activités) a été créé en 2014 pour encadrer le développement d'une méthode générique (basée sur des enquêtes) permettant la production en routine de données de spatialisation homogènes, à une échelle spatiale et temporelle adaptée aux besoins. Il rassemble des scientifiques de plusieurs laboratoires de l'université de Nantes et des représentants des marins-pêcheurs (6 comités des pêches sur les 9 en métropole). De cette collaboration interdisciplinaire, pluridisciplinaire (géographie, géomatique, statistique) et opérationnelle émergent plusieurs problématiques de recherche, parmi lesquelles l'élaboration d'un plan d'échantillonnage stratifié adapté permettant de surmonter entre autres les risques liés à la sur-enquête des pêcheurs. Différentes stratégies d'échantillonnage ont été testées basées sur la théorie des sondages et le savoir-faire des comités des pêches. Les résultats des simulations réalisées à partir de données exhaustives ont permis d'identifier la meilleure stratification pour l'estimation des indicateurs spatio-temporels utilisés par les comités. Le plan d'échantillonnage actuellement retenu est un plan stratifié aléatoire simple à allocation proportionnelle.

Mots-clés. Théorie des sondages, plan d'échantillonnage, stratification, estimation spatialisée, pêche maritime, indicateurs spatio-temporelles, aménagement de l'espace maritime.

1. Introduction

Les activités en mer ne cessent de s'accroître avec les développements industriels (*e.g.* les énergies marines renouvelables, les sites d'extractions de granulats etc.) et écologiques (*e.g.* les aires marines protégées pour le respect de la biodiversité) contraignant les marins pêcheurs à modifier leurs pratiques et/ou leurs espaces de travail. De plus, les activités de pêche des navires qui ne sont pas équipés d'un système de suivi par satellite (les moins de 12 m sauf exceptions) et qui représentent

près de 80% des flottilles, ne sont pas représentées, créant ainsi une source de conflit lors d'implémentation d'un nouveau usage en mer. Le projet VALPENA¹ (éVALuation des activités de PÊche au regard des Nouvelles Activités) a été initié en Pays de Loire en 2010 et est devenu un Groupement d'Intérêt Scientifique (GIS) en 2014. Le but est de développer des outils et des méthodes permettant aux pêcheurs d'être eux-mêmes en capacité de spatialiser leurs activités pour les besoins de l'aménagement de l'espace maritime.

Les données VALPENA sont basées sur des enquêtes annuelles des patrons de pêches sur leurs activités de l'année précédente. Pour éviter la sur-enquête des pêcheurs et pour réduire les coûts financiers d'enquêtes, il a été décidé de mettre en place un plan d'échantillonnage.

Nous présenterons les différents plans d'échantillonnage élaborer en utilisant la théorie des sondages et les résultats des simulations de Monte-Carlo permettant d'évaluer le meilleur plan d'échantillon en terme de représentativité des activités de pêche. L'estimation de l'activité de pêche est réalisée en utilisant les indicateurs spatio-temporels développés au sein du GIS VALPENA et la comparaison des plans d'échantillonnage sera illustrée sur un des indicateurs le plus utilisé par les comités.

Matériel et Méthodes

2.1 Matériel

Les données d'enquêtes VALPENA collectées concernent : (i) la zone travaillée (réparti sur des mailles d'environ 3 milles nautiques de côtés), (ii) la date (année et mois), (iii) l'engin déployé et (iv) l'espèce ciblée (dans la limite de trois). Le référentiel géographique et les nomenclatures des engins et des espèces sont interopérables (*e.g.* le maillage spatial utilisé est une subdivision des rectangles statistiques du Conseil International pour l'Exploration de la Mer, CIEM). Les comités disposent également d'une base de données technique et administrative sur les navires: le port d'attache, la longueur du navire et des informations personnels sur l'armateur, etc... Par ailleurs, un logiciel (VALPEMAP) a été construit permettant de saisir directement les enquêtes sur une plateforme en ligne. Cette plateforme permet d'éviter les erreurs de saisies lors des enquêtes (*e.g.* fautes de frappe ou couple engin/espèce non existant). Pour plus d'information, le site valpena.fr, décrit exhaustivement la méthodologie de collecte et de validation des données recueillies ainsi que le mode de fonctionnement du projet.

Les résultats présentés seront basés sur les données exhaustives de l'activité de l'année 2014 de la flotte de Normandie. Les plans d'échantillonnage testés sont de type stratifié aléatoire simple avec allocation proportionnelle mais avec des stratifications différentes (cf. Tillé (2001), p134). Nous comparerons différentes stratifications possibles à l'aide d'une méthode de Monte-Carlo appliquée aux données exhaustives.

2.2 Méthodes

En concertation avec les différents comités membres du GIS, deux indicateurs ont été choisis pour construire, comparer et valider les différents plans d'échantillonnage: l'Indicateur de Densité (ID) et l'Indicateur d'Intensité (II). Nous présenterons les résultats sur l'Indicateur d'Intensité (II).

2.2.1 Variable d'intérêt

Indicateur d'Intensité II

Pour une période donnée, l'Indicateur d'Intensité, II, mesure l'intensité de l'activité pêche sur une maille. Plus précisément, l'indicateur II correspond au nombre de mois déclarés travaillés sur la

¹ Pour plus d'informations, voir le site internet : valpena.fr

maille, cumulés pour tous les navires de la flotte étudiée durant une période déterminée. L'indicateur d'intensité, II , est la variable d'intérêt.

En notant N le nombre de navires de la flottille étudiée et MO le nombre de mois de la période étudiée, pour une maille k , l'indicateur d'intensité, II , est défini par :

$$II(\text{maille}_k) = \sum_{i=1}^N \sum_{m=1}^{MO} \mathbb{1}_{\text{nav}_i \times \text{mois}_m \times \text{maille}_k}$$

Où $\mathbb{1}_{\text{nav}_i \times \text{mois}_m \times \text{maille}_k}$ est égale à 1 si le navire i a eu une activité le mois m sur la maille k pendant la période étudiée.

2.2.2 Fonction d'intérêt

Indicateur d'Intensité en proportion (II_p)

La fonction d'intérêt qu'on cherche à estimer est l'indicateur II en proportion, noté II_p . Elle représente la proportion de l'intensité de pêche par rapport au nombre de navires et de mois étudiés. Elle est défini par :

$$II_p(\text{maille}_k) = \frac{II(\text{maille}_k)}{MO \times N}$$

Une proportion est un cas particulier d'une moyenne et se classe parmi les fonctions d'intérêt linéaires. Ainsi, nous utiliserons l'estimateur de Horvitz et Thompson qui est un estimateur linéaire et sans biais et qui peut s'utiliser dans le cadre de tout plan d'échantillonnage (Tillé (2001)).

2.2.3 Les différents PE étudiés

Plusieurs plans d'échantillonnage ont été testés afin de déterminer : la meilleure stratification et la taille de l'échantillon à enquêter (Ardilly (2006) et Tillé (2001)). La composition des strates a été établie grâce aux connaissances des comités de pêches sur l'activité des navires. Nous présenterons les résultats entre deux choix de stratification et pour une taille d'échantillon optimisée en utilisant une précision fixée par chaque comité variant entre 5 et 10%. Les deux plans d'échantillonnage présentés utilisent l'allocation proportionnelle et se différencient par la stratification mise œuvre :

PE1 : stratification basée sur le port d'attache du navire, la catégorie du premier engin et son type (actif ou passif) et la longueur du navire.

En effet, un navire peut être équipé de plusieurs engins. Le premier engin est défini comme étant l'engin utilisé le plus longtemps (cela a été normalisé pour tous les comités).

PE2 : stratification basée sur le port d'attache du navire, sous-flottille (*i.e.* par métier particulier) et la longueur du navire.

La définition des typologies des sous-flottilles est propre à chaque comité et une sous-flottille peut regrouper plusieurs catégories d'engins utilisés par le navire.

2.2.4 Méthode de comparaison des PE

Les données issues d'une année exhaustive sont utilisées pour comparer les deux plans d'échantillonnage. Nous calculons d'abord l'indicateur d'intensité en proportion par maille k (pour $k=1, \dots, K$) pour la population totale (*i.e.* enquête exhaustive sur l'année d'étude) ; la moyenne, l'écart type et quelques quantiles des II_p sur les K mailles sont également calculés. Ensuite, il faut déterminer dans le cadre d'un sondage, la taille de l'échantillon de navires à enquêter, n , pour

pouvoir estimer II_p avec la précision du plan d'échantillonnage défini par chaque comité. Une fois n fixé, nous allons mettre en œuvre les deux plans PE1 et PE2 en répartissant les navires en strate avec un minimum de 5 navires par strate h ($h=1...H$). Comme nous disposons initialement des données exhaustives, pour étudier les propriétés des deux plans d'échantillonnage, nous simulons des échantillons de taille n associées à chaque plan d'échantillonnage en suivant les trois étapes suivantes:

1- Pour chaque plan d'échantillonnage (PE1 et PE2), on simule 500 échantillons de navires sans remise (Monte Carlo), $c=1 \dots 500$:

1.1- Tirage d'un échantillon $S(c)$ de taille n en tirant n_h navires dans chaque strate h de taille N_h selon l'allocation proportionnelle :

$$\frac{n_h}{N_h} = \frac{n}{N}$$

1.2- Calcul de II_p estimés pour chaque échantillon $S(c)$ et chaque maille $k=1, \dots, K$.

2- Comparaison à l'échelle de chaque maille k des deux plans d'échantillonnage : calcul pour chaque maille k , la moyenne des II_p estimés, le biais, la variance et l'Erreur Quadratique Moyenne relatifs (EQMr) calculés à partir des 500 échantillons. Des cartographies de ces valeurs sont également réalisées.

3- Comparaison globale sur l'ensemble des K mailles des deux plans d'échantillonnage : calcul de la moyenne, l'écart type et les quantiles des II_p estimés sur la zone d'étude (ie les K mailles) et les 500 échantillons.

Le meilleur plan d'échantillonnage en termes de représentativité de l'activité de pêche sera choisi en comparant : à l'échelle de chaque maille (i) les cartographies des EQMr, (ii) les cartographies des II_p estimés par rapport aux II_p de l'année exhaustive et, à l'échelle de la zone d'étude (iii) les valeurs de la moyenne, de l'écart type et des quantiles de II_p sur toutes les mailles.

2- Résultats et perspectives

Nous montrerons que le meilleur plan d'échantillonnage adapté aux données VALPENA est un plan stratifié aléatoire simple avec allocation proportionnelle. Les résultats seront présentés sur l'activité 2014 de la flotte de Normandie. La construction du plan d'échantillonnage repose sur trois principes:

- La stratification de la flotte basée sur trois critères : la situation géographique (port/quartier maritime/groupement de port), l'activité des navires (catégorie et type du premier engin) et la longueur des navires (variables selon le comité).
- Le tirage de l'échantillon des navires de chaque strate à l'aide d'un sondage aléatoire simple sans remise : chaque navire a la même probabilité d'être enquêté ou non et ne peut être enquêté qu'une seule fois.
- L'allocation proportionnelle déterminant le nombre de navires à enquêter dans chaque strate. Le nombre de navires à enquêter, sera ainsi proportionnel à la taille de la strate.

D'autres travaux de simulation sont en cours pour comparer l'allocation proportionnelle à d'autres types d'allocations. Par ailleurs, ce travail ouvre la voie à plusieurs réflexions, allant du traitement pluriannuel des données issues d'enquêtes exhaustives et de PE, au traitement des données manquantes pour les années de plan d'échantillonnage.

Bibliographie

Ardilly, P. (2006), *Les techniques de sondages*, Technip, Paris

Tillé, Y. (2001), *Théorie des sondages*, Dunod, Paris.