

Plans d'échantillonnage adaptés aux données spatialisées d'activité de pêche professionnelle issues de l'observatoire VALPENA

*Elodie PLISSONNEAU ⁽¹⁾, Lise BELLANGER ⁽²⁾, Guillaume CHAUVET ⁽³⁾, Ion TILLIER ⁽⁴⁾,
Brice TROUILLET ⁽⁵⁾*

⁽¹⁾ Maison des Sciences de l'Homme – USR3491 , Nantes. elodie.plissonneau@univ-nantes.fr

⁽²⁾ Laboratoire de Mathématiques Jean Leray – UMR CNRS 6629, Nantes.

⁽³⁾ ENSAI, Rennes

⁽⁴⁾ CRPMEM Pays de la Loire – Les Sables d'Olonne.

⁽⁵⁾ LETG-Nantes, Géolittomer – UMR 6554 CNRS, Nantes.

Sommaire

I) Introduction

II) Matériel et Méthodes

III) Résultats

IV) Conclusions et Perspectives

I) Introduction

➤ **VALPENA** : é**VAL**uation des activités de **P**êche au regard des **N**ouvelles **A**ctivités

(Voir <http://www.gis-valpena.fr/>)

- Projet issu d'une collaboration entre Université de Nantes et le COREPEM en 2010. Depuis 2014, c'est un **GIS** hébergé à la Maison des Sciences de l'Homme (Nantes).
- **Participants** : 7 des 9 CRPMEM à l'échelle nationale métropolitaine (soit 2/3 des navires de la flotte métropolitaine):
Nord-Pas-de-Calais/Picardie, Haute et Basse Normandie, Bretagne, Pays de la Loire, PACA et Poitou-Charentes.

- **But du dispositif** : donner des arguments aux pêcheurs dans les problématiques de partage de l'espace marin:
 - Création et exploitation de données par les structures professionnelles des pêches.
- **Principe** : Enquête déclarative auprès des pêcheurs, chaque année, pour inventorier leurs activités de pêche
- **Limites** : Enquêtes exhaustives impossibles à réaliser chaque année
⇒ Mise en place de Plans d'Échantillonnage (PE) sur les navires.

➤ Travaux réalisés depuis 2014 :

- Comparaison par simulation de différents PE envisageables pour échantillonner les navires :
 - ✓ À partir d'une variable d'intérêt particulière : Indicateur d'Intensité (II).
 - ✓ Pour les données recueillies en Bretagne (BZH), Nord-Pas-de-Calais/Picardie (NPCP), Pays de la Loire (PDL) et Basse Normandie.
- Mise en œuvre opérationnelle du PE sélectionné :
 - ✓ en 2015 : liste de navires à enquêter pour NPCP.
 - ✓ en 2016 : liste de navires à enquêter pour NPCP, PDL, BN et BZH.

➤ Objectifs de ce travail :

- Comparer les performances (ER) des PE pour l'Indicateur d'Intensité pour différentes allocations : proportionnelle et optimale
- Aboutir à des recommandations de mise en œuvre sur le terrain et dans le traitement des données

II) Matériel et Méthodes

2.1) Matériel

➤ **Base de sondage** : Base Navires des Pays de la Loire

- Croisement du Fichier Flotte Communautaire (EU) et du savoir des membres des comités.
- Données techniques et administratives:
 - Port d'attache et d'exploitation
 - Longueur
 - Puissance
 - Engins utilisés
 - Armateur
 - Licences
- Au départ, base plus ou moins fiable et complète en fonction des comités.

2.1) Matériel

➤ Informations recueillies :

couverture des deux-tiers de la flotte française métropolitaine

- Population U à enquêter de taille N : Navires (patrons pêcheurs)
- Information à récolter pour chaque maille du domaine spatial étudié
 - ✓ Utilisation du logiciel VALPEMAP pour la saisie des enquêtes
- Données stockées sous la forme:

registration_nbr	id_maille	année	mois	engin	espèce
xxxxxxx	26E7C5	2010	1	OTB	SCE
xxxxxxx	26E7D3	2010	2	OTB	SCR
xxxxxxx	26E7D4	2010	3	DRB	SCR
xxxxxxx	26E7D4	2010	3	OTB	SCR
...

Accueil >> test >> 2005 >> DEF (001)

Nouvelle saisie de travail pour le mois de janvier

Annuler

Enregistrer

Engins

Ajouter

Code	Nom
aucun engin sélectionné	

Espèces

Ajouter

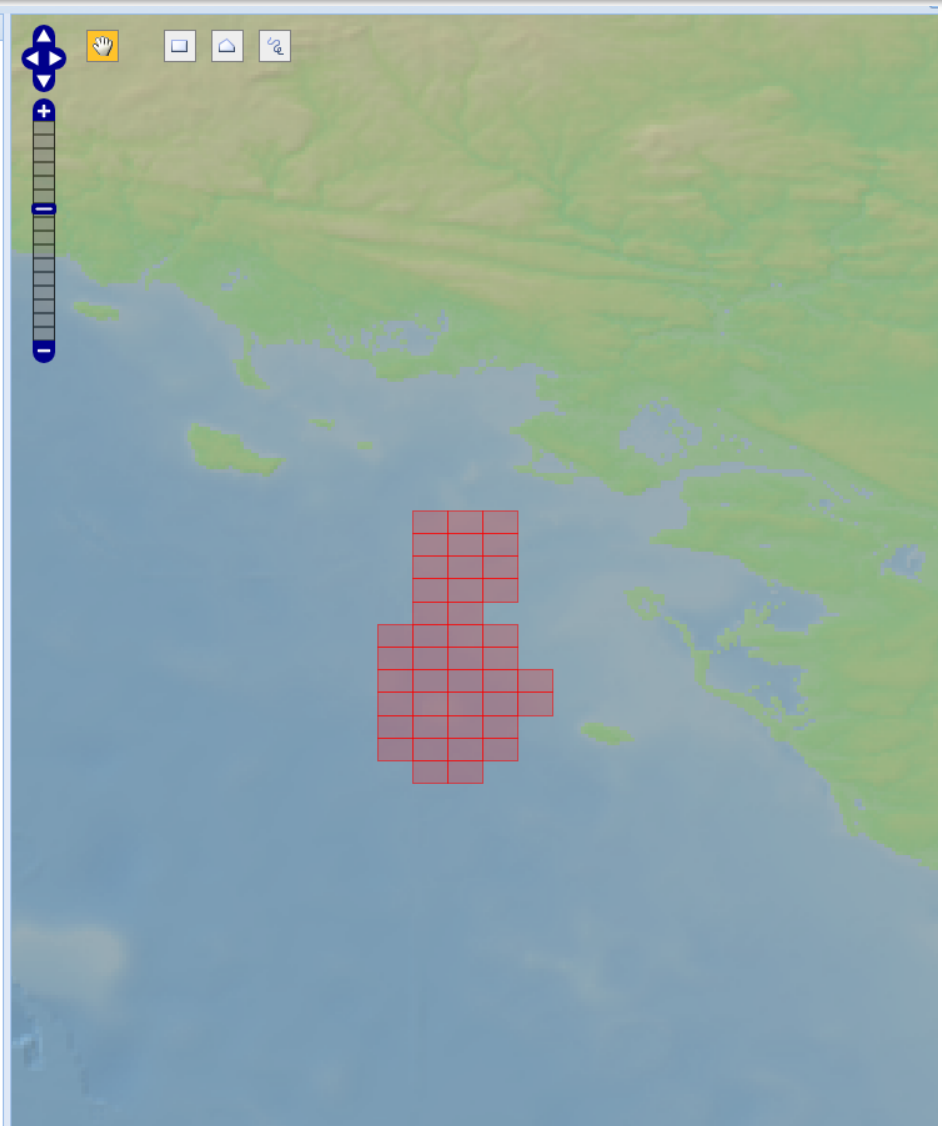
Code	Nom
aucune espèce sélectionnée	

Nombre de mailles sélectionnées : 42

Commentaires :

Couches cartographiques affichées

Bathymétrie	<div>100%</div>	<div>▼</div>
Maillage CIEM	<div>40%</div>	<div>▲</div>



2.1) Matériel

➤ Informations recueillies :

couverture des deux-tiers de la flotte française métropolitaine

- Population à enquêter de taille N : Navires (patrons pêcheurs)
- Information à récolter pour chaque maille du domaine spatial étudié
 - ✓ Utilisation du logiciel VALPEMAP pour la saisie des enquêtes
- Données stockées sous la forme:

registration_nbr	id_maille	année	mois	engin	espèce
xxxxxx	26E7C5	2010	1	OTB	SCE
xxxxxx	26E7D3	2010	2	OTB	SCR
xxxxxx	26E7D4	2010	3	DRB	SCR
xxxxxx	26E7D4	2010	3	OTB	SCR
...

2.2) Méthodes

2.2.1) Variable d'intérêt : l'Indicateur d'Intensité

- *Indicateur d'Intensité*: Nombre de mois travaillés, pour le navire i , maille k fixée :

$$II_{nav_i}^k = \sum_{m=1}^{MO} 1_{mois_m \times maille_k \times nav_i} ; k \text{ fixé}, i = 1, \dots, N$$

- *Variable d'intérêt* : II pour le navire i moyennée sur toutes les mailles :

$$Y_i = \frac{\sum_{k=1}^K II_{nav_i}^k}{K}$$

- *Paramètre d'intérêt à optimiser* : Moyenne de la variable d'intérêt :

$$\bar{Y}_U = \frac{1}{N} \sum_{i=1}^N Y_i$$

Où : - N est la taille de la population U

- MO le nombre de mois de la période étudiée, souvent l'année

2.2) Méthodes

2.2.2) Plans d'Echantillonnage

➤ But du travail :

Comparaison par simulations

des performances de différents plans de sondage
stratifiés aléatoires simples,

à allocation proportionnelle ou optimale,

pour estimer un paramètre d'intérêt //.

➤ Définition de l'allocation:

Procédure d'affectation des n unités de l'échantillon à chacune des strates.

2.2) Méthodes

2.2.2) Plans d'Echantillonnage

i. Types d'allocations

➤ Allocation proportionnelle :

- Le principe est que la proportion d'individus à enquêter dans chaque strate est identique à la proportion d'individus à enquêter dans la population totale.
- La formule est la suivante : $\frac{n_h}{N_h} = \frac{n}{N}$
Où : n_h est la taille de l'échantillon de la strate h ;
 N_h est la taille de la strate h ;
 n est la taille de l'échantillon à prélever sur la population totale ;
 N est la taille de la population totale.
- Le paramètre d'intérêt estimé est alors : $\hat{Y} = \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_h} \sum_{i \in S_h} Y_i = \frac{1}{n} \sum_{i \in S} Y_i$
Où : S_h est l'échantillon tiré dans la strate h .

2.2) Méthodes

2.2.2) Plans d'Echantillonnage

i. Types d'allocations

➤ Allocation optimale de Neyman :

- Le principe est de tirer une proportion d'individu à enquêter d'autant plus grande que la variance au sein de la strate est grande.

- La formule est la suivante : $n_h = n \frac{N_h S_{yh}}{\sum_{l=1}^H N_l S_{yl}}, h = 1, \dots, H$

Où : $S_{yh} = \sqrt{S_{yh}^2}$;

$S_{yh}^2 = \frac{1}{N_h - 1} \sum_{k \in U_h} (Y_k - \bar{Y}_h)^2$ est la variance corrigée de la strate h ;

U_h est la population de la strate h ;

$\bar{Y}_h = \frac{\sum_{k \in U_h} Y_k}{N_h}$ est la moyenne des Y_k pour $k \in U_h$.

- Le paramètre d'intérêt estimé est alors : $\hat{Y} = \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_h} \sum_{i \in S_h} Y_i$

2.2) Méthodes

2.2.2) Plans d'Echantillonnage

ii. Les différents PE étudiés

Nom du PE	Stratification	Taille de l'échantillon	Type d'allocation
<i>PEactuel</i>	QM/1 ^{er} engin Actif/Passif	1/3 de la flotte	Proportionnelle
<i>PEstrat</i>	Port/1 ^{er} Engin/Longueur	1/3 de la flotte	Proportionnelle
<i>PEstratopt</i>	Quartier Maritime	1/3 de la flotte	Optimale
<i>PEstratn</i>	Port/1 ^{er} Engin/Longueur	$n \geq \frac{N \left(b^2 + \widehat{p}_y (1 - \widehat{p}_y) z_{1-\frac{\alpha}{2}}^2 \right)}{b^2 N + \widehat{p}_y (1 - \widehat{p}_y) z_{1-\frac{\alpha}{2}}^2} \quad (A)$	Proportionnelle
<i>PEstratnopt</i>	Quartier Maritime	$n \geq \frac{N \left(b^2 + \widehat{p}_y (1 - \widehat{p}_y) z_{1-\frac{\alpha}{2}}^2 \right)}{b^2 N + \widehat{p}_y (1 - \widehat{p}_y) z_{1-\frac{\alpha}{2}}^2} \quad (A)$	Optimale

- Avec:
- N nombre total de navires et n la taille de l'échantillon de navires
 - b la précision souhaitée pour le paramètre d'intérêt étudié (ici fixée à 10%)
 - $z_{1-\frac{\alpha}{2}}^2$ quantile d'ordre $1-\alpha/2$ d'une loi $N(0,1)$
 - \widehat{p}_y estimation de \overline{I}^k pour la maille k (prise ici dans la pire situation i.e. égale à 0.5)

2.2) Méthodes

2.2.3) Comparaison des PE : étude de simulation

➤ La méthode est la suivante :

1. Pour l'ensemble de la base navires PDL enquêtée en 2011 (activité 2010) :
 - Calcul de l'indicateur $\bar{II}^k = \frac{1}{N} \sum_{i=1}^N II_{nav_i}^k$ pour chaque maille k ;

2.2) Méthodes

2.2.3) Comparaison des PE : étude de simulation

➤ La méthode est la suivante :

Pour chaque PE (PE_{actuel} , PE_{strat} , $PE_{stratopt}$, PE_{stratn} , $PE_{stratnopt}$) à comparer :

2. Ré-échantillonnage sans remise (méthode de Monte Carlo)

- Pour $s=1,...,500$:
 - tirage aléatoire d'un échantillon s , de taille fixée (1/3 de la flotte ou n selon les PE), suivant une allocation proportionnelle ou optimale;
 - calcul des estimations de Horvitz-Tompson de l'indicateur II à partir de l'échantillon s , pour chaque maille:

$$\bar{II}_{(s)}^k = \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_h} \sum_{i \in S_h^{(s)}} II_{navi}^k$$

2.2) Méthodes

2.2.3) Comparaison des PE : étude de simulation

➤ La méthode est la suivante :

Pour chaque PE (PE_{actuel} , PE_{strat} , PE_{stratn}) à comparer :

3. Estimation de l'indicateur II et de sa précision par maille k sur l'ensemble des échantillons:

- Estimation de : $\hat{II}^k = \sum_{s=1}^{500} \frac{\bar{II}_{(s)}^k}{500}$
- Erreur relative : $ER(\hat{II}^k) = \frac{\sqrt{EQM(\hat{II}^k)}}{\bar{II}^k}$

2.2) Méthodes

2.2.3) Comparaison des PE : étude de simulation

➤ Comparaison globale des résultats, sur l'ensemble des mailles du domaine d'étude :

- Estimation du paramètre d'intérêt : $\bar{\hat{I}}$;
- Etendue et IC 95% par percentiles des valeurs des mailles ;
- Erreur relative moyennée sur l'ensemble des mailles.

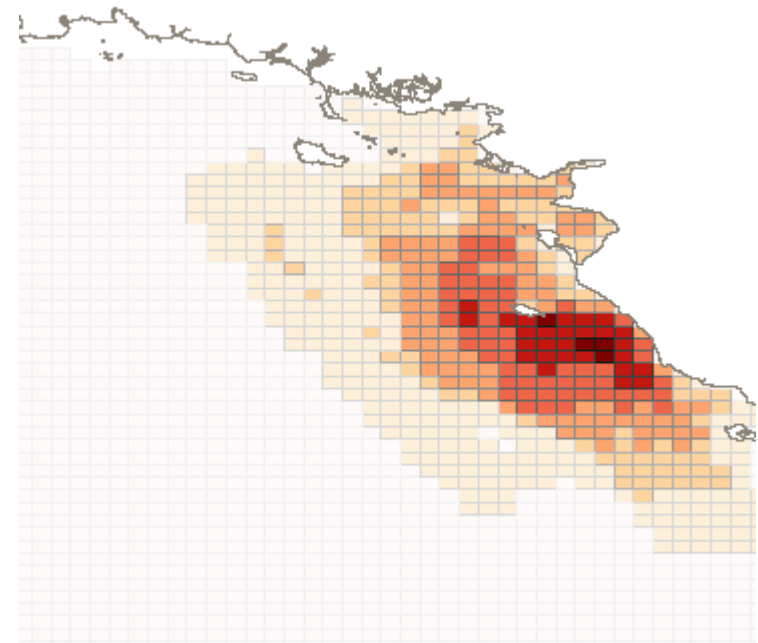
➤ Cartographie du domaine d'étude :

- Indicateur d'Intensité par maille : \hat{I}^k ;
- Erreur relative moyennée sur l'ensemble des échantillons, par maille.

III) Résultats

3.1) Description des données PDL

- Année d'étude : activité 2010
- $K = 15000$ mailles
- $N = 345$ navires
- Stratification :
9 strates (opt) ou 28 strates (prop)
- n : nombre de navires échantillonnés
77 (n tds) ou 115 (1/3 population)



III) Résultats

3.1) Description des données PDL

- Stratification des 345 navires et Allocation optimale pour $n = 77$ navires

Strates	Nh	nh (proportionnel)	nh_rectif	Probabilité d'inclusion	Poids
AI	30	0 (10)	4	0.13	7.50
ES	7	0 (2)	4	0.57	1.75
LS	44	6 (15)	5	0.11	8.80
NO	39	4 (13)	4	0.10	9.75
PI	49	4 (16)	4	0.08	12.25
RB	48	2 (16)	4	0.08	12.00
SG	42	11 (14)	10	0.24	4.20
SN	56	27 (9)	23	0.41	2.43
YE	30	22 (10)	19	0.63	1.58

3.2) Cartographie : résultats sur *II* pour *PEstratn* et *PEstratnopt*

	Flotte totale	PEstratn	PEstratnopt
Taille	n=345	n=77	n=77
Amplitude (%)	[0.00 ; 12.36]	[0.01 ; 12.34]	[0.00 ; 12.91]
IC Percentiles (%)	[0.05 ; 3.53]	[0.01 ; 3.29]	[0.05 ; 3.56]
Moyenne globale (%)	0.43	0.43	0.45
Erreur Relative globale	-	0.05	4.49

3.2) Cartographie : résultats pour *PEstratn* et *PEstratnopt*

$\hat{\Pi}_h$ PEstratn PDL10

II_h PDL10

$\hat{\Pi}_h$ PEstratnopt PDL10

Cartes non diffusables

3.2) Cartographie : résultats pour *PEstratn* et *PEstratnopt*

ER $\hat{\Pi}^k$ *PEstratn* PDL10

ER $\hat{\Pi}^k$ *PEstratnopt* PDL10

Cartes non diffusables

ER moyenne : 0.05

ER moyenne : 4.49

- *PEstratnopt* semble moins efficace que *PEstratn* pour estimer $\bar{\Pi}^k$
- *PEstratnopt* semble quand même correct dans les zones les plus fréquentées

IV) Conclusions et Perspectives

4.1) Conclusions

- Tests sur l'allocation optimale **non concluants**
- **Causes possibles** de ce résultats :
 - *Stratification trop imprécise*
 - *Choix du paramètre d'intérêt à revoir :*
 - Allocation proportionnelle : permet estimation spatialisée
pas d'hypothèse non spatialisée
 - Allocation optimale: hypothèse sur la moyenne sur toutes les mailles
hypothèse non spatialisée pour estimation spatialisée

IV) Conclusions et Perspectives

4.2) Perspectives

➤ *Amélioration du PE :*

- Amélioration du PE avec l'allocation optimale:
 - Autres paramètres d'intérêt
 - Stratification plus fine
- Améliorations de la stratification : inclusion du critère « Métier » possible ;
- Travail sur le choix du n ;

➤ *Réflexion sur une utilisation possible d'une **approche modèle**.*

➤ *Traitements des données en aval :*

- Etudes longitudinales pour les données recueillies avec PE et sans PE ;
- Post-stratification pour calculer d'autres fonctions d'intérêt ;
- Simulation de scenarii (projet éolien, ...) impactant l'espace maritime et donc le calcul des variables d'intérêt.

Plans d'échantillonnage adaptés aux données spécialisées d'activité de pêche professionnelle issues de l'observatoire VALPENA

Merci pour votre attention !

Elodie PLISSONNEAU, GIS VALPENA - Maison des Sciences de l'Homme, Nantes. elodie.plissonneau@univ-nantes.fr

Références

- Ardilly, P. (2006). *Les techniques de sondage (nouvelle édition)*, Paris : Editions Technip.
- Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R*. Hoboken, NJ, USA: Wiley.
- Plissonneau, E. (2014). Normalisation de l'exploitation des données VALPENA par la mise en place d'indicateurs géostatistiques et validation de plans d'échantillonnage avec la théorie des sondages, Rapport de stage de Master 2, 66p.
- Tillé, Y. (2001). *Théorie des sondages – Échantillonnage et estimation en populations finies*. Paris : Dunod.
- Tillier, I., Cautain, B., Brivoal, F. and N. Gace-Rimaud (2014). Etude d'impact du raccordement électrique du projet éolien du banc de Guérande : Analyse de l'activité de pêche professionnelle, Rapport produit dans le cadre de l'étude d'impact du projet, Réseau de Transport d'Electricité, 99p.
- Trouillet, B. (2013). Connaissance des pratiques spatiales dans le domaine des pêches maritimes : l'expérience VALPENA. Colloque « Exploitation et conservation des écosystèmes aquatiques : une question d'échelle ? », AFH, Bordeaux, 19-21 juin 2013.