

Inférer la distribution spatiale d'espèces halieutiques :

Effet de la réallocation uniforme des captures commerciales



Océane GUITTON, Chloé TELLIER, Juliette THEOLEYRE
Cursus M2 Agronome - Spécialisation Science des données

Collaboration avec l'IFREMER et le Pôle halieutique, mer et littoral d'AgroCampus Ouest



l'institut Agro
agriculture • alimentation • environnement



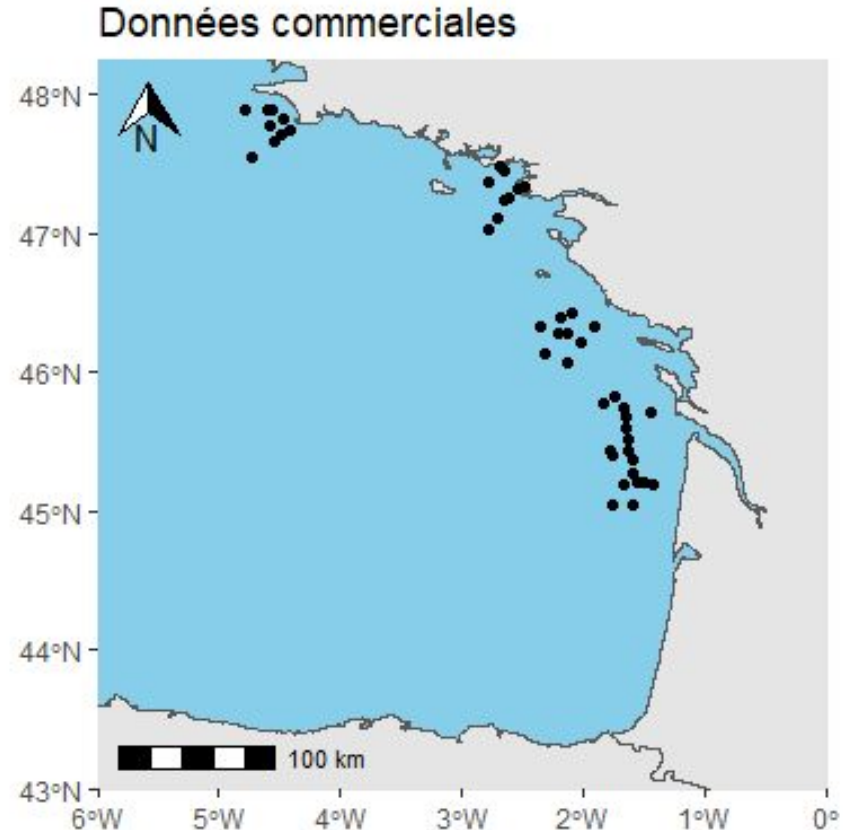
I - Estimation spatialisée de l'abondance





Des données valorisables : les déclarations de capture = **données commerciales**.

- **Localisation** des points de pêche et **quantité** de poissons pêchés.
- **Peu coûteuses** et disponibles en **grande quantité**.
- **Ciblage des zones de forte abondance** par les pêcheurs.





Le modèle considéré présente les *trois composantes* suivantes :

- La **densité de biomasse** : variable latente spatialisée.
- Un **processus d'échantillonnage** correspondant à la génération de la localisation des **points de pêche**.
- Un **modèle d'observation** qui génère la **quantité de poisson pêchée** en chaque point de pêche.

Composantes du modèle

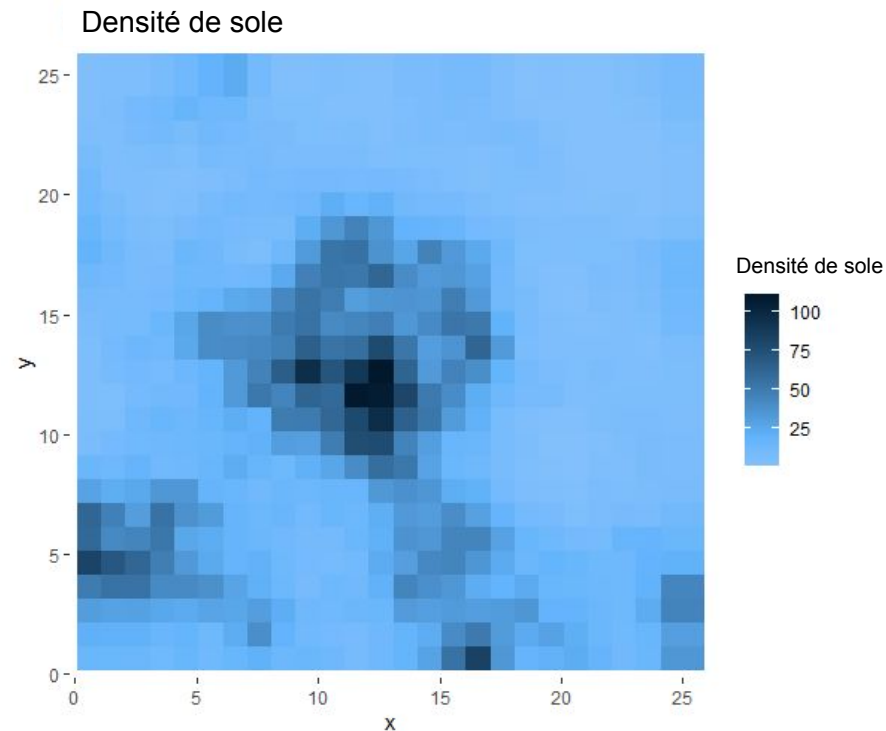
Densité de soles



La **densité de soles** est modélisée de la manière suivante :

$$S(x) = \exp(\alpha_s + \Gamma_s(x)^T \cdot \beta_s + \delta(x))$$

- α_s : **intercept**.
- $\Gamma_s(x)$: **covariables environnementales**.
- B_s : **paramètres** associés à ces **covariables**.
- $\delta(x)$: **effet spatial aléatoire** qui traduit la corrélation spatiale de la distribution des soles.



Composantes du modèle

Le processus d'échantillonnage

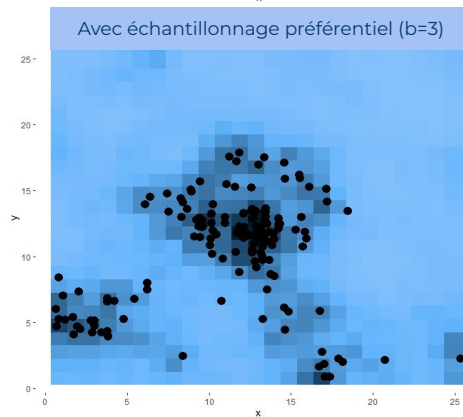
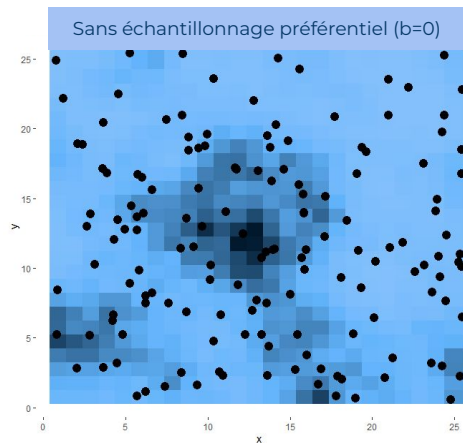
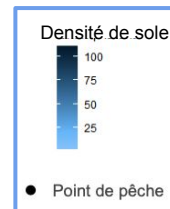


Génération des localisations des points de pêche :
processus de Poisson non homogène.

$$Xcom \sim IPP(\lambda(x))$$

$$\lambda(x) = \exp(\alpha_x + b \cdot \log(S(x)) + \Gamma_x(x)^T \cdot \beta_x + \eta(x))$$

- α_x : **intercept**.
- b : paramètre traduisant un **échantillonnage préférentiel**.
- $S(x)$: valeur de la **densité de sole** au point considéré.
- $\Gamma_s(x)$: **covariables**.
- β_x : **paramètres** associés aux **covariables**.
- $\eta(x)$: **effet spatial aléatoire**.



Composantes du modèle

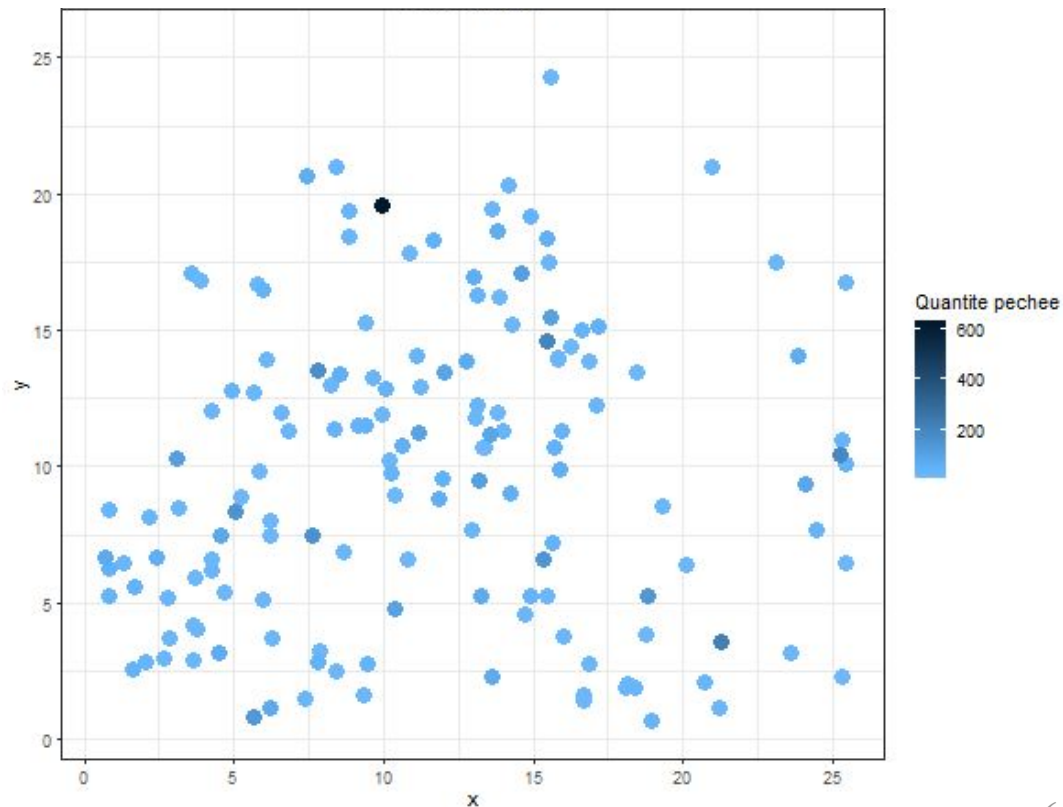
Modèle d'observation



Attribution de **la quantité exacte de poissons pêchée** en chaque point de pêche.

$$Y_i \mid x_i, S(x_i) \sim L (S(x_i), \sigma^2)$$

Avec **Yi la quantité pêchée**.

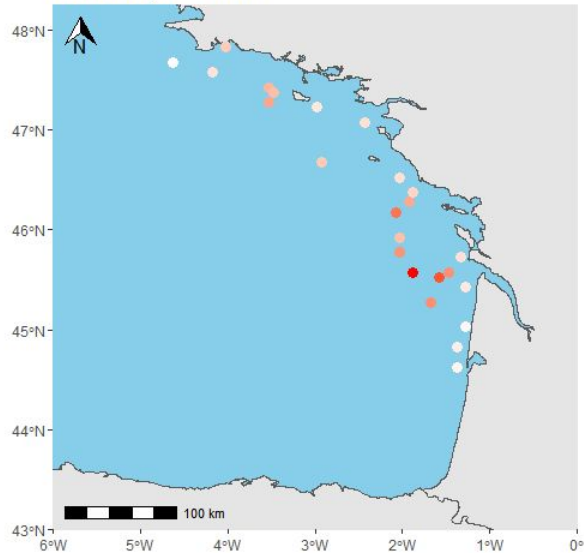


Problématisation

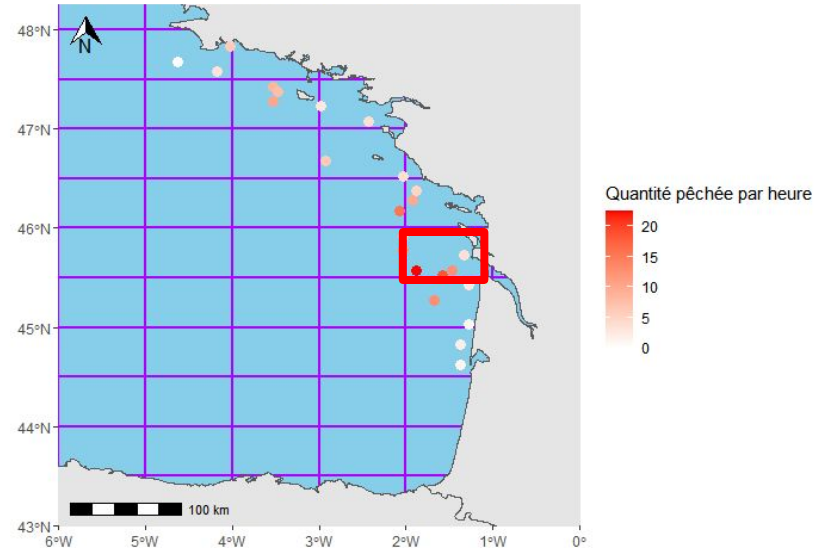
...



1) Données commerciales du modèle d'Alglave et al.



2) Ajout des carrés statistiques du golfe de Gascogne



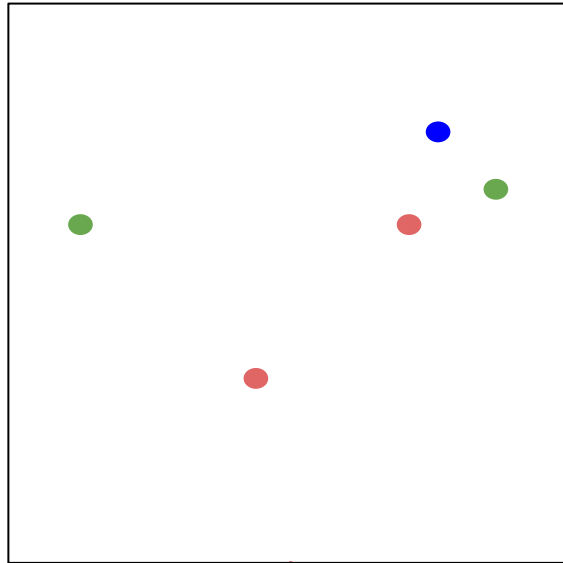
- En réalité, **on n'a pas la quantité pêchée par point de pêche, mais à l'échelle d'un carré statistique** pour un jour donné et un bateau donné.
- On se place à **l'échelle d'un carré statistique**.

Problématisation

La réallocation uniforme



3) Captures exactes

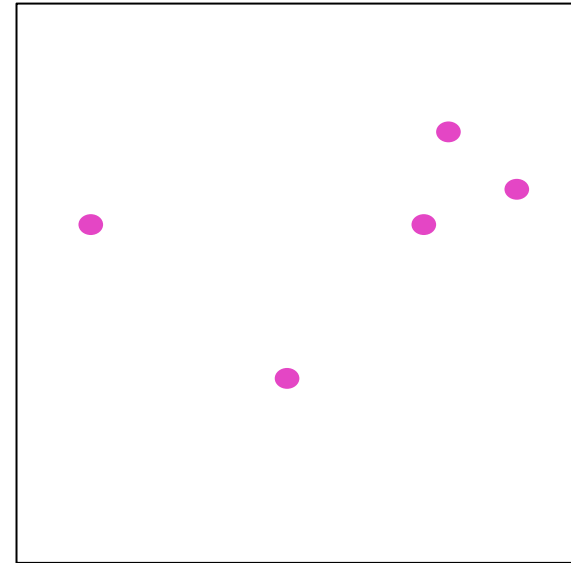


Quantité pêchée
par heure

- 20
- 10
- 5

Quantité totale pêchée (y_{total}) = **50**
Nombre de points de pêche (n) = **5**

4) Captures ré-allouées



Quantité pêchée
par heure

- 10

$$\frac{(y_{\text{total}})}{(n)} = \frac{50}{5} = 10 = y_{\text{réallouée}}$$

→ Quantités pêchées ré-allouées :
identiques pour tous les points de
pêche du bateau d'un carré statistique. 9

Problématique

Quelle est la conséquence de cette procédure de ré-allocation sur les sorties du modèle, sur l'estimation des paramètres, et notamment sur l'estimation de l'abondance des poissons ?

II - Etude de l'influence de la réallocation des captures commerciales sur la performance du modèle





➤ **Etude par simulation** sur différents scénarios.

➤ **Simplification** du modèle :

- Densité de soles :

$$S(x) = \exp(\alpha S + \Gamma S(x)T \cdot \beta S + \delta(x))$$

- Intensité du processus d'échantillonnage :

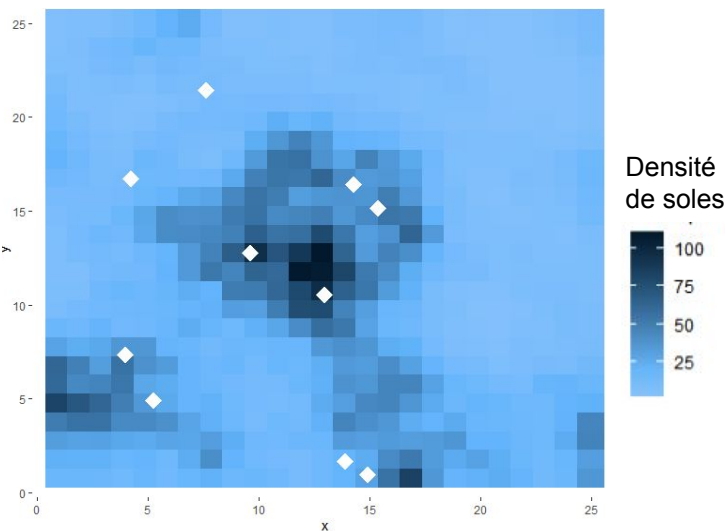
$$\lambda f(x) = \exp(\alpha Xf + bf.\log(S(x)) + \Gamma X(x)T \cdot \beta Xf + \eta f(x))$$



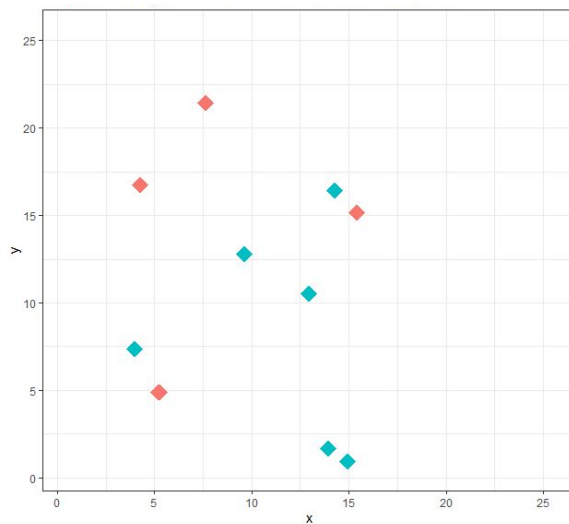
- **150** points de pêche.
- **Échantillonnage préférentiel** plus ou moins fort :
 - Échantillonnage non préférentiel ($b=0$).
 - Échantillonnage préférentiel modéré ($b=1$).
 - Échantillonnage préférentiel fort ($b=3$).



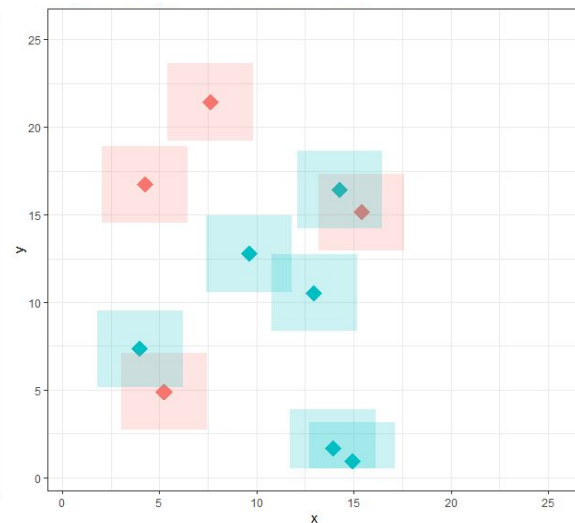
1) Génération des centres de pêche



2) Attribution des centres aux bateaux

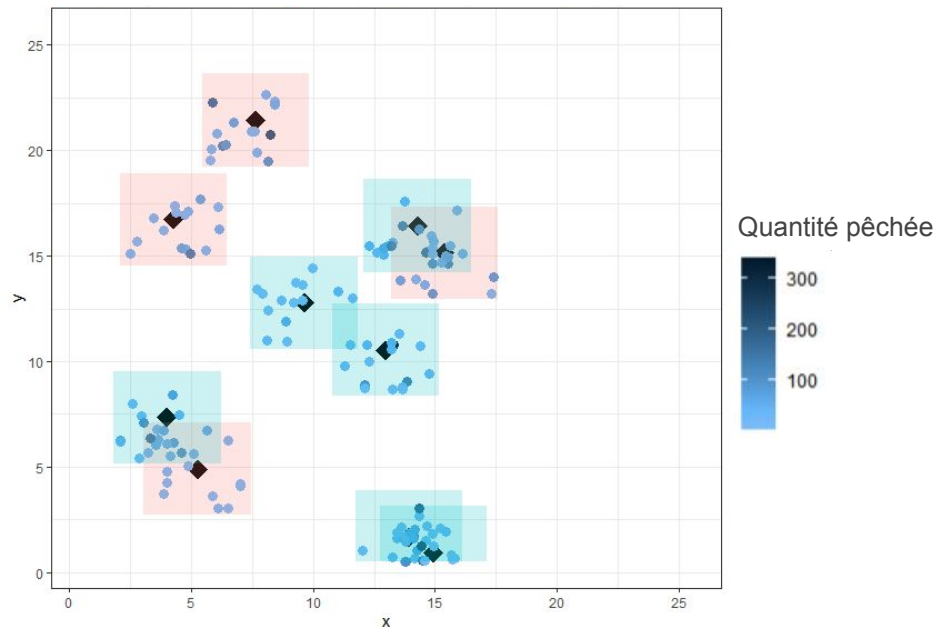


3) Délimitation des zones de pêche

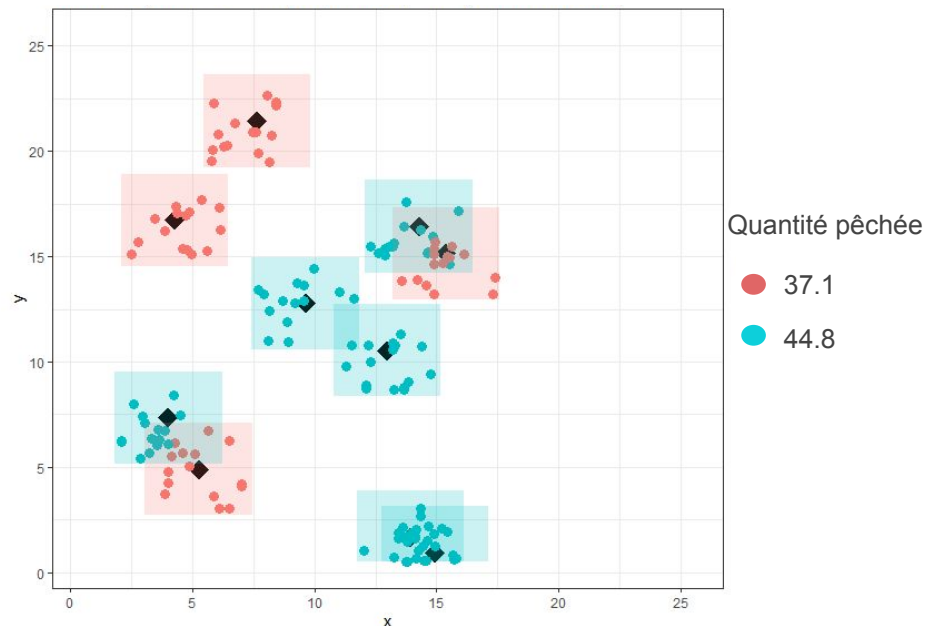




4) Génération des points de pêche et des quantités pêchées



5) Réallocation uniforme



◆ Centre de pêche

● Point de pêche

■ Délimitation d'une zone de pêche



- **Nombre de bateaux de pêche** considérés dans le protocole de simulation **(P)**.
- **Nombre moyen de zones de pêches** visitées par l'ensemble des bateaux **(Z)**.
- **Structuration spatiale** dans la localisation des points de pêche.

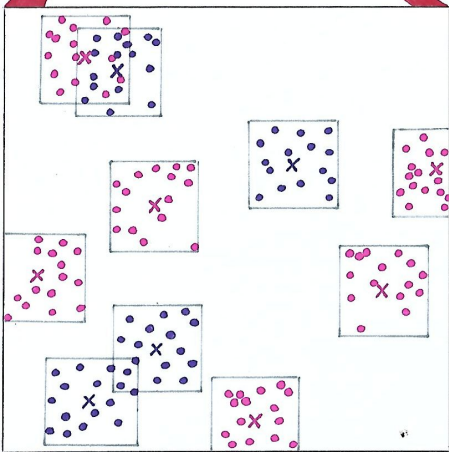


Scénarios de simulation

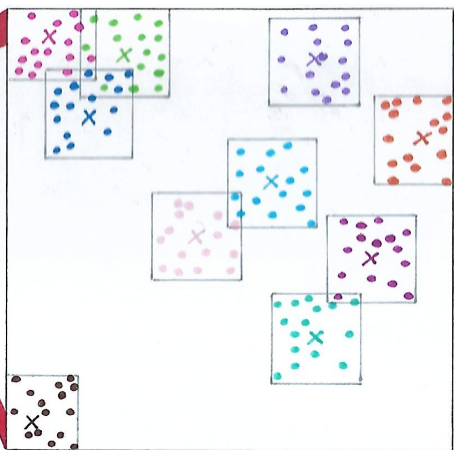
Les différentes valeurs des paramètres

- 3 valeurs de nombre de bateaux possibles (P).
- 3 valeurs de nombre moyen de zones de pêche possibles (Z).
- Exemple des couples (P = 2, Z = 5) et (P = 10, Z = 1).

z \ P	2		5		10	
	C	V	C	V	C	V
1	2	75	5	30	10	15
3	6	25	15	10	30	5
5	10	15	25	6	50	3



• P_1
 • P_2



• P_1 • P_3 • P_5 • P_7 • P_9
 • P_2 • P_4 • P_6 • P_8 • P_{10}

P = Nombre de séquences de pêche
 Z = Nombre moyen de zones de pêche par séquence de pêche
 C = Nombre de centres de pêche générés (= P*Z)
 V = Nombre de points de pêche dans chaque zone
 p_k = Séquence de pêche n^*k

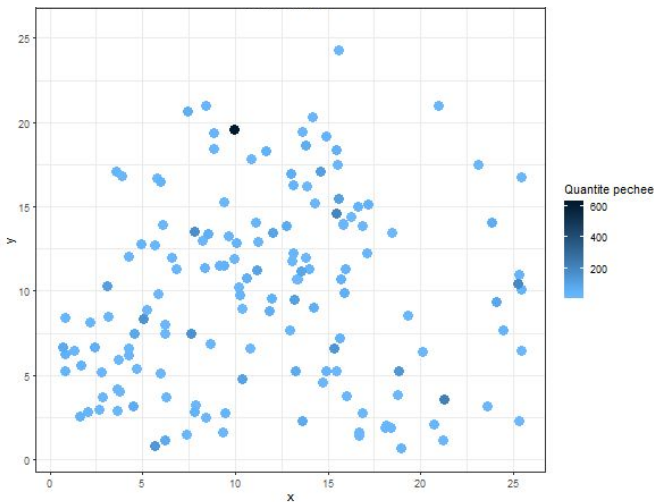
✕ Centre de pêche
 □ Délimitation d'une zone de pêche
 ● Point de pêche

Scénarios de simulation

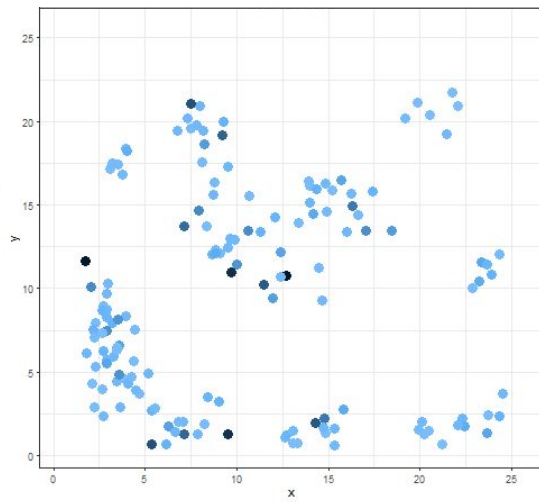
Comparaison de trois situations



Situation idéale,
sans structuration
ni réallocation

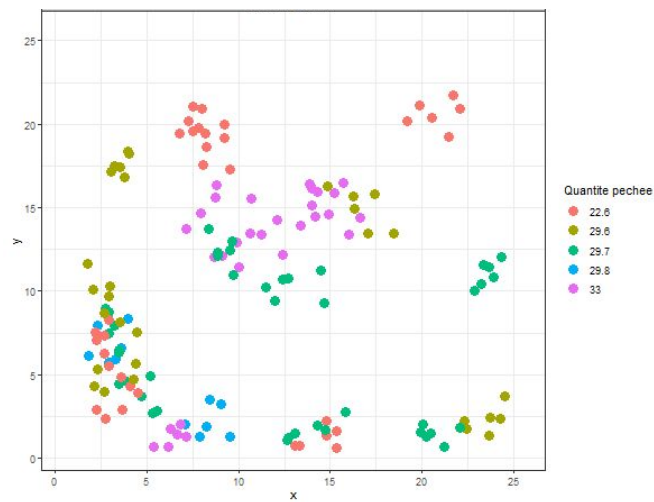


Situation avec
structuration et
sans réallocation



● Point de pêche

Situation avec
structuration et
avec réallocation





Utilisation du package TMB :

- **Estimation des paramètres** par méthode du maximum de vraisemblance.
- **TMB = Template Model Builder.**
 - Deux étapes :
 - **Approximation de Laplace** pour calculer la vraisemblance.
 - Recherche du **maximum de vraisemblance** grâce à un algorithme de descente de gradient.

Déroulé des scénarios



1) Situation sans structuration, sans réallocation

Echantillonnage préférentiel **fixé** (b)

Simulation sans structuration

2) Situation avec structuration, sans réallocation

Echantillonnage préférentiel **fixé** (b)
Nombre de bateaux (P)
Nombre moyen de zones par bateau (Z)

Simulation avec structuration

Biomasse totale **simulée** (N)
Densité de soles **simulée** (S)

Réallocation uniforme

Estimation des paramètres (TMB)

Nouveauté dans le modèle

Echantillonnage préférentiel **estimé** (b)
Densité de soles **estimée** (S)
Biomasse totale **estimée** (N)



Objectif : évaluer l'**efficacité** de notre modèle, sans ou avec **réallocation uniforme**.

Pour cela, nous allons nous intéresser à **trois métriques de performance** :

- Biais de la **biomasse totale N**.
$$= \frac{N_{\text{estimée}} - N_{\text{simulée}}}{N_{\text{simulée}}}$$
- **Root Mean Squared Prediction Error (RMSPE)**.
$$= \sqrt{\frac{1}{n} \cdot \sum_x (S(x)_{\text{simulé}} - S(x)_{\text{estimé}})^2}$$
- Biais de **l'échantillonnage préférentiel b**.
$$= \frac{b_{\text{estimé}} - b_{\text{fixé}}}{b_{\text{fixé}} \text{ (ou 1 si } b_{\text{fixé}} = 0)}$$

III - Résultats



Résultats

2 types de résultats :

- L'effet de la **structuration**.
- L'effet de la **réallocation uniforme**.

Rappels :

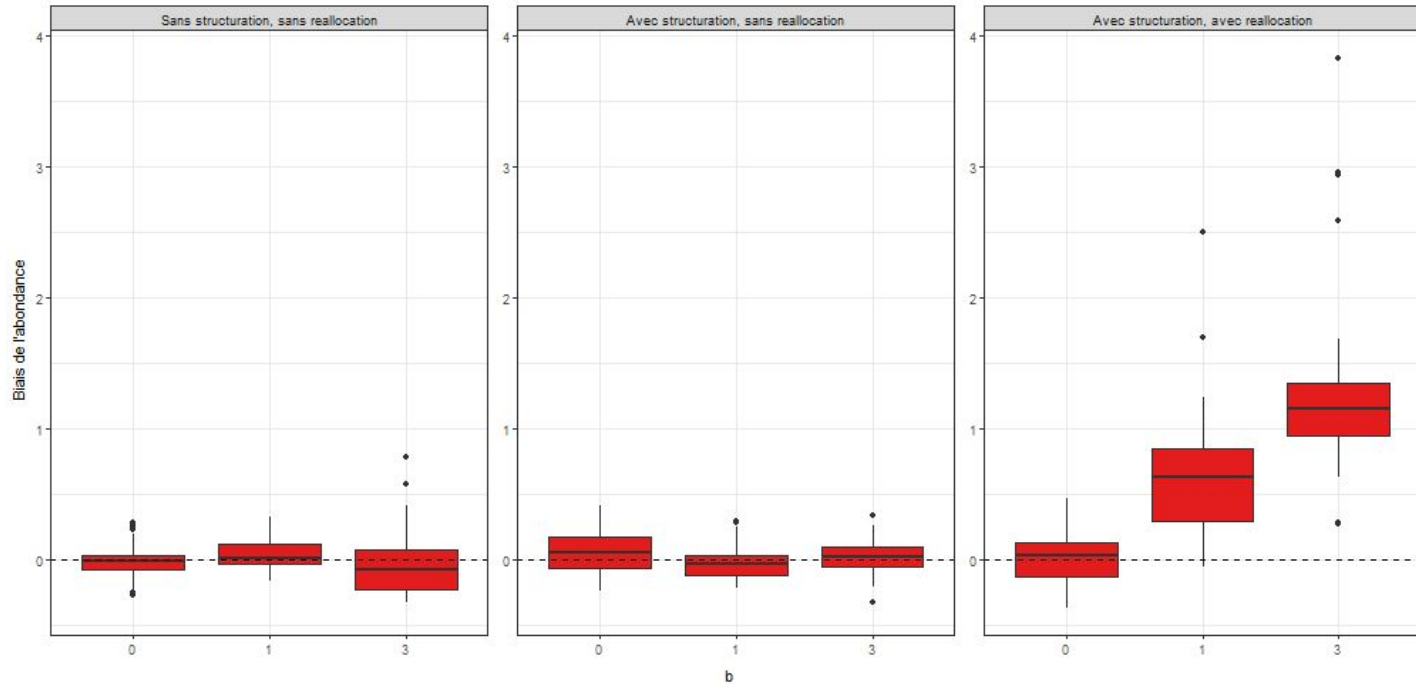
- **Un scénario** = une valeur d'échantillonnage préférentiel (b), un nombre de bateaux (P), un nombre moyen de zones de pêche par bateau (Z).
- **27 scénarios différents** : $3 (b) * 3 (P) * 3 (Z)$.
- Un total de **54 scénarios** : $27 * 2$ (sans ou avec réallocation).
- **100 simulations** de chaque scénario.



5 bateaux (P=5)
5 zones de pêche en
moyenne par bateau (Z=5)

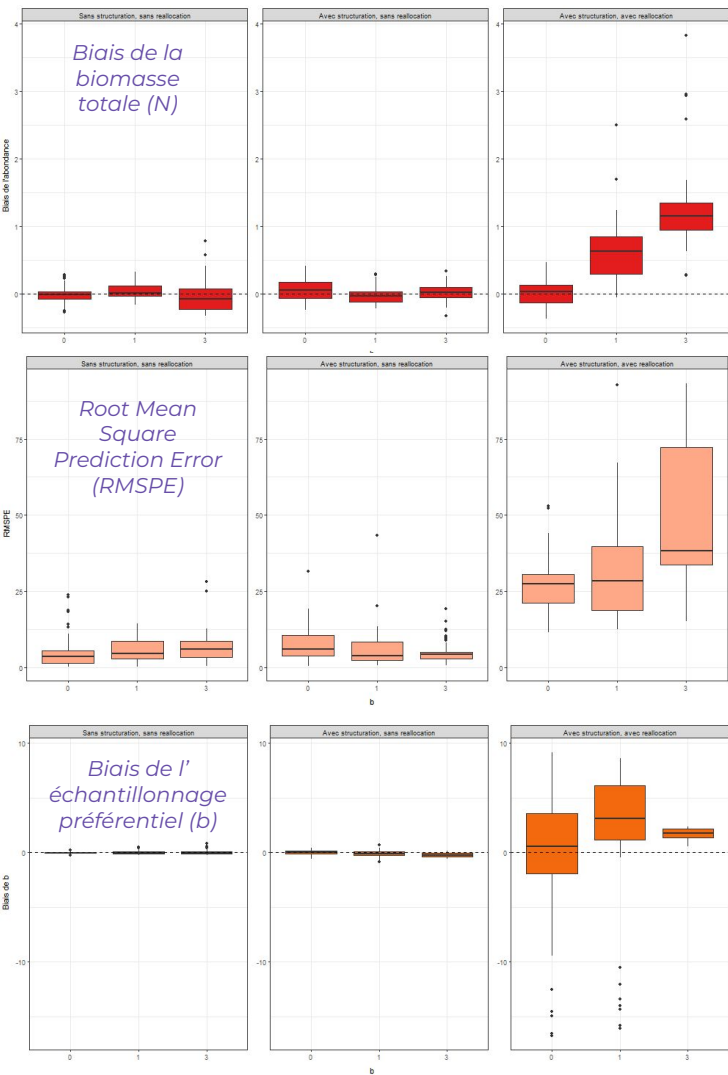
Biais de la biomasse totale (N)

$$= \frac{N_{\text{estimée}} - N_{\text{simulée}}}{N_{\text{simulée}}}$$

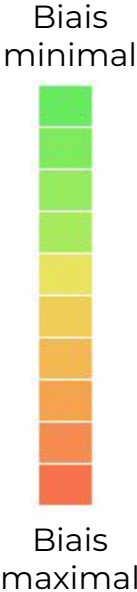


Résultats

Effet de la structuration et de la réallocation uniforme



- La **structuration** “en zones” de nos points de pêche commerciaux **n'introduit pas de biais** par rapport à des points de pêche non structurés.
- En revanche, la **réallocation uniforme** des captures commerciales :
 - **Augmente fortement le biais de l'échantillonnage préférentiel (b), le biais de la biomasse totale (N) et le RMSPE.**
 - **Augmente la variabilité** de ces métriques de performance.
 - Pour le biais de la biomasse totale (N) et le RMSPE, cet effet est d'autant plus important que **l'échantillonnage préférentiel est fort.**



Biais médian de N pour les différentes situations ZP, avec réallocation uniforme et $b=1$

Z/P	2	5	10
1	0.28	0.20	0.38
3	0.44	0.55	0.68
5	0.47	0.63	0.59

Rappels :

P = nombre de bateau de pêche
Z = nombre moyen de zones de pêche

Résultats



Effet de la structuration en présence de réallocation

Biais médian de N pour les différentes situations ZP, avec réallocation uniforme et $b=1$

Z/P	2	5	10
1	0.28	0.20	0.38
3	0.44	0.55	0.68
5	0.47	0.63	0.59

Médiane de RMSPE pour les différentes situations ZP, avec réallocation uniforme et $b=1$

Z/P	2	5	10
1	19.03	15.76	17.62
3	25.54	25.19	29.45
5	29.61	28.38	24.95

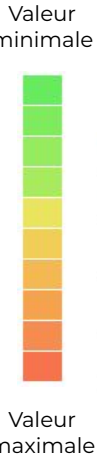
Biais médian de b pour les différentes situations ZP, avec réallocation uniforme et $b=1$

Z/P	2	5	10
1	1.39	0.87	1.13
3	0.35	3.41	3.95
5	1.12	3.13	4.33

Rappels :

P = nombre de bateau de pêche
Z = nombre moyen de zones de pêche

- La qualité d'estimation se dégrade avec l'augmentation du nombre moyen de zones visitées par bateau.
- Le nombre de bateaux ne semble pas influencer sur la qualité de nos résultats.



IV - Conclusion et discussion





- La structuration en zones de pêche n'introduit pas de biais dans l'estimation en comparaison avec une absence de structuration.
 - A quel point la réallocation détériore-t-elle les performances du modèle ?
 - **Introduction de biais** sur l'estimation de l'**échantillonnage préférentiel** et sur l'estimation de la **biomasse totale** ; **erreurs locales dans l'estimation de la densité de soles**.
 - **Plus le nombre de zones homogénéisées est élevé, plus la réallocation est grossière** → biais et erreurs plus importants.
- **Nécessité de modifier le modèle pour prendre en compte la réallocation.**



➤ **Perspectives :**

- Simuler des **captures 'sans succès'**.
- Prendre en compte les **effets spatiaux aléatoires**.

➤ **Peut on utiliser la donnée commerciale pour inférer la distribution spatiale des espèces halieutiques ?**

→ Introduire de la donnée scientifique qui pourrait être une **source d'information supplémentaire** non négligeable en présence de **réallocation uniforme**.



**Merci de
votre écoute !**

Un grand merci à **Marie-Pierre Etienne** et à **Baptiste Alglave** pour
leur accompagnement tout le long du projet.