

Cours / TD - Introduction à la modélisation spatiale pour l'écologie

Baptiste Alglave, Sophie Elliott

Janvier 2022

- **Introduction:** Rappel contexte / sujet de thèse
- **Partie 1:** Modélisation intégrée et échantillonnage préférentiel
“Integrated framework accounting for preferential sampling to estimate fish spatial distribution: a base framework, limits and perspectives”
- **Partie 2:** Modélisation spatio-temporelle
“Can we trust commercial landings data to identify essential habitats of harvested fish?”
- **Partie 3:** Questions de changements de support
“Going from coarse landings data to fine scale species distribution”
- **Partie 4:** Détails techniques / comment estimer la corrélation spatiale efficacement ?
Approche SPDE et INLA

- **Introduction:** Rappel contexte / sujet de thèse
- **Partie 1:** Modélisation intégrée et échantillonnage préférentiel
“Integrated framework accounting for preferential sampling to estimate fish spatial distribution: a base framework, limits and perspectives”
- **Partie 2:** Modélisation spatio-temporelle
“Can we trust commercial landings data to identify essential habitats of harvested fish?”
- **Partie 3:** Questions de changements de support
“Going from coarse landings data to fine scale species distribution”
- **Partie 4:** Détails techniques / comment estimer la corrélation spatiale efficacement ?
Approche SPDE et INLA

**Différence d'objectif
avec Sophie**

*Modéliser la relation à l'habitat
vs.
Modéliser les structures de
corrélation spatio-temporelles*

Introduction

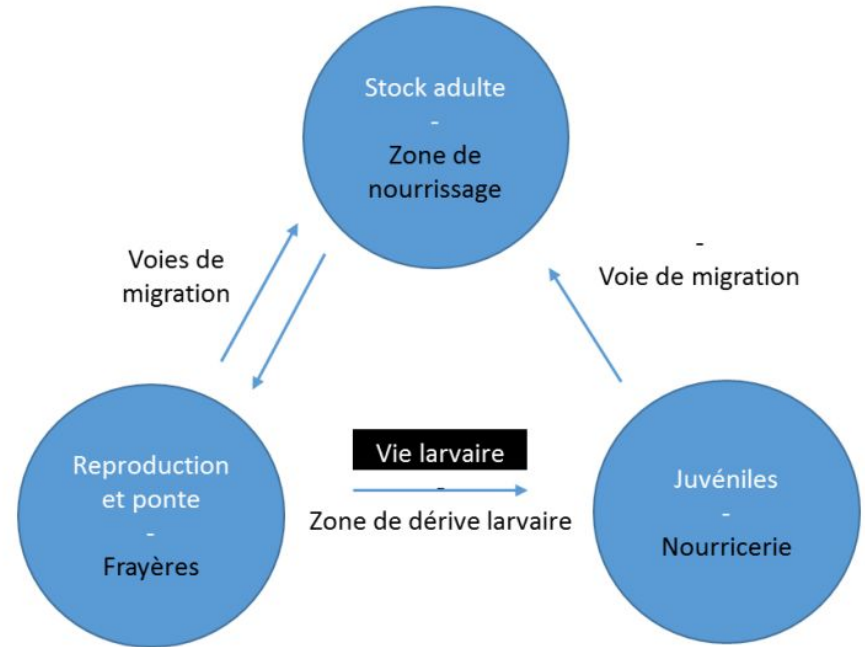
Contexte halieutique

Le renouvellement des ressources halieutiques

⇒ repose sur des zones clés de leur cycle de vie

=

Zones Fonctionnelles Halieutiques d'importance



Représentation schématique du cycle de vie d'une espèce d'intérêt halieutique

Introduction

Contexte halieutique

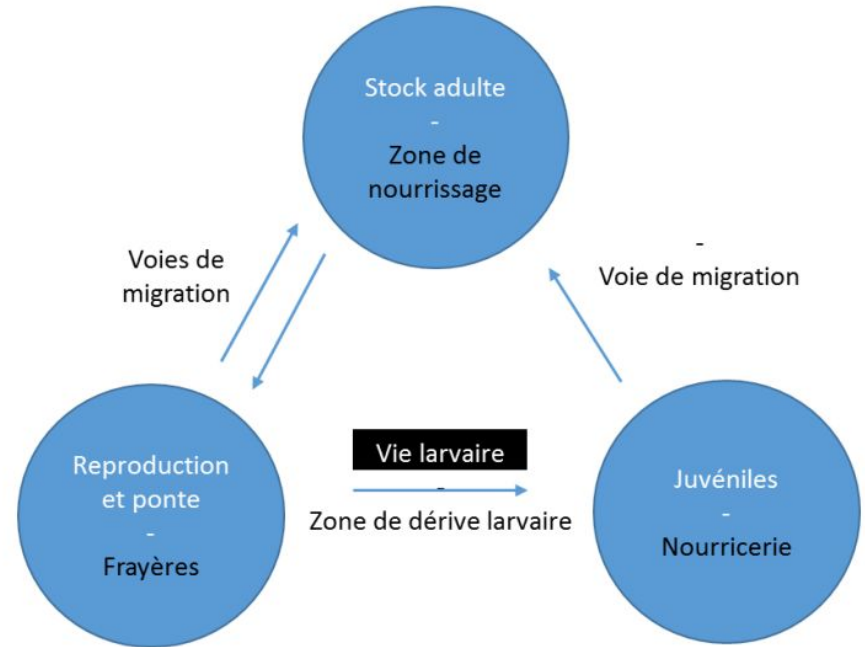
Le renouvellement des ressources halieutiques

⇒ repose sur des zones clés de leur cycle de vie

=

**Zones Fonctionnelles
Halieutiques d'importance**

Identification et caractérisation de ces zones dans le temps et l'espace ?



Représentation schématique du cycle de vie d'une espèce d'intérêt halieutique

Introduction

Les données disponibles

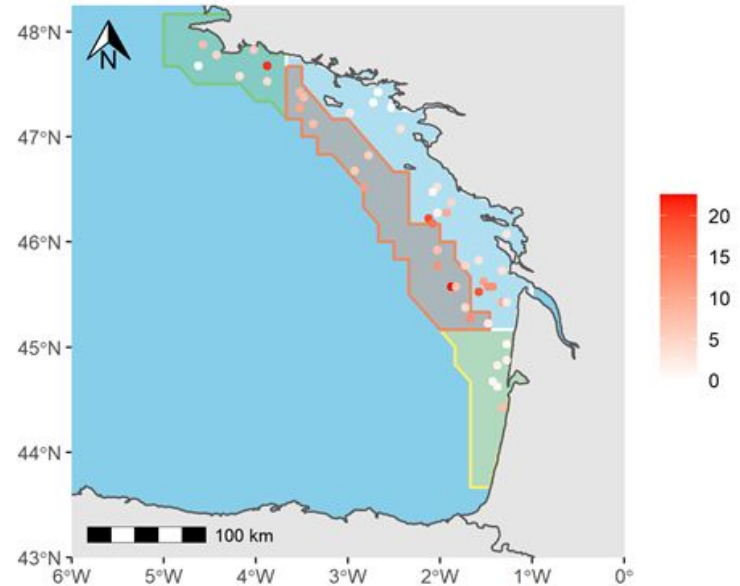
Campagnes
scientifiques



- Plan d'échantillonnage standardisé

- Données coûteuses

- Couverture spatio-temporelle limitée



Données scientifiques

A noter les différents types de campagnes scientifiques (démersales vs. pélagiques)

Observations captures au chalut \neq observations acoustiques

Nb. échantillons campagnes au chalut \ll Nb. échantillons campagnes acoustiques

Sole commune - campagne ORHAGO 2018.

Unité : kg.

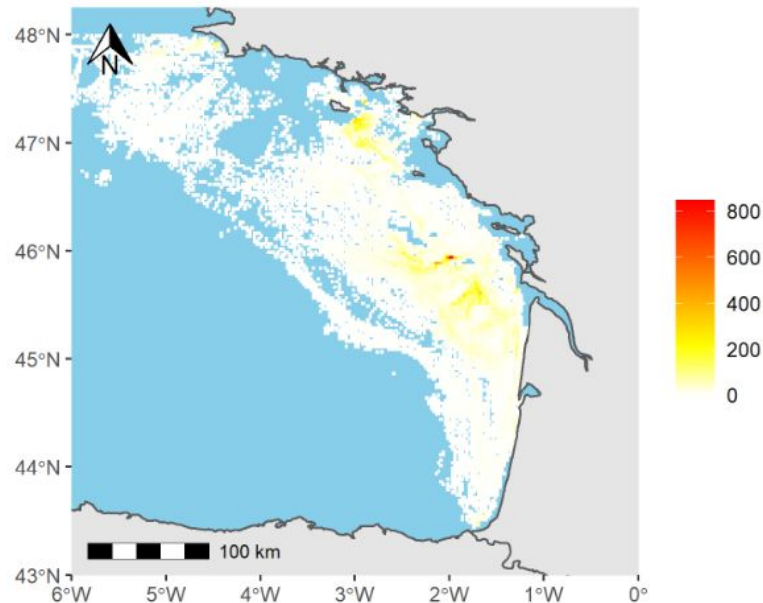
Introduction

Les données disponibles

Données logbooks
x
Données VMS



- Échantillonnage dense dans l'espace et le temps
- Peu coûteuse (pour les scientifiques)
- Échantillonnage non-standardisé et dépendant du comportement de pêche



Débarquements commerciaux

Sole commune - novembre 2018. Métier OTB_DEF :
chaluts de fond ciblant les espèces démersales.
Unité : kg.

***Vous connaissez ?
Bcp + de données que Sophie***

Introduction

Les données disponibles

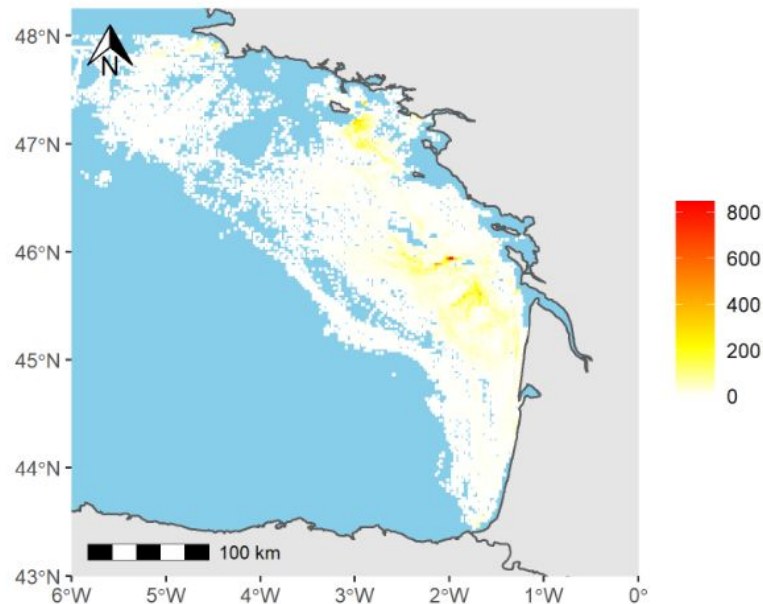
Données logbooks
x
Données VMS



- Échantillonnage dense dans l'espace et le temps
- Peu coûteuse (pour les scientifiques)
- Échantillonnage non-standardisé et dépendant du comportement de pêche

Proche des questions liées à l'échantillonnage des données en sciences participatives

Isaac et al., 2020



Débarquements commerciaux

Sole commune - novembre 2018. Métier OTB_DEF : chaluts de fond ciblant les espèces démersales.

Unité : kg.

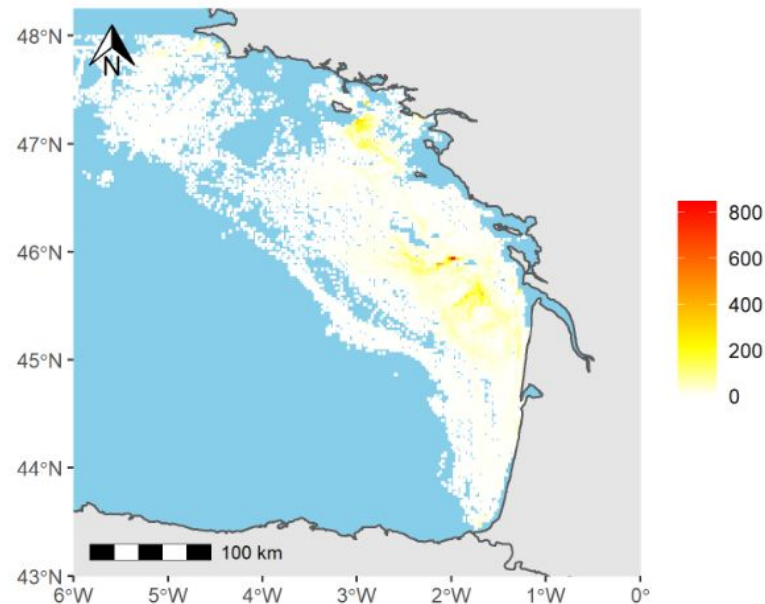
Introduction

Les données disponibles

Données logbooks
x
Données VMS



- Échantillonnage dense dans l'espace et le temps
- Peu coûteuse (pour les scientifiques)
- Échantillonnage non-standardisé et dépendant du comportement de pêche
- Flottes de pêche non homogènes
- Hypothèses pour répartir les captures sur les points VMS



Débarquements commerciaux

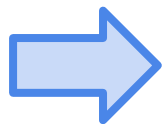
Sole commune - novembre 2018. Métier OTB_DEF : chaluts de fond ciblant les espèces démersales.

Unité : kg.

Introduction

Objectif de la thèse :

Construction d'une approche de **modélisation spatio-temporelle intégrée** combinant données scientifiques et données commerciales



Identification de zones fonctionnelles halieutiques d'importance

Une thèse à forte composante méthodologique :

Méthodologique

Écologique

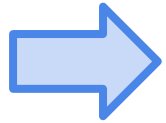
Halieutique

Gestion et
conservation

Introduction

Objectif de la thèse :

Construction d'une approche de **modélisation spatio-temporelle intégrée** combinant données scientifiques et données commerciales



Identification de zones fonctionnelles halieutiques d'importance

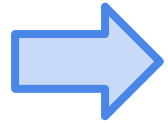
Une thèse à forte composante méthodologique :



Introduction

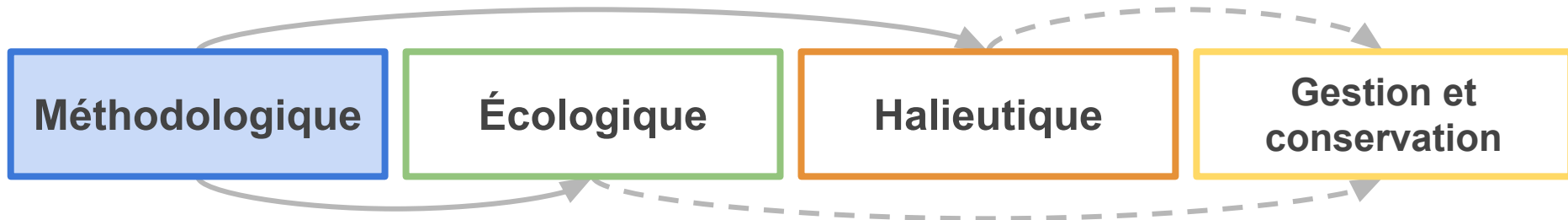
Objectif de la thèse :

Construction d'une approche de **modélisation spatio-temporelle intégrée** combinant données scientifiques et données commerciales



Identification de zones fonctionnelles halieutiques d'importance

Une thèse à forte composante méthodologique :



Partie 1 : Modèle intégré et échantillonnage préférentiel

Integrated framework accounting for preferential sampling to estimate fish spatial distribution: a base framework, limits and perspectives

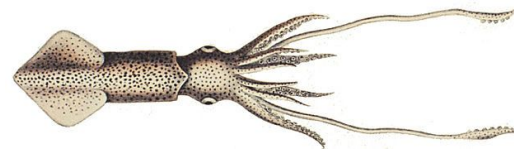
- Présentation du modèle
- Evaluation du modèle par scénarios de simulation-estimation
- Application à 3 espèces démersales du golfe de Gascogne



Merlu



Sole



Loliginidae / encornet

Partie 1 : Modèle intégré et échantillonnage préférentiel

Integrated framework accounting for preferential sampling to estimate fish spatial distribution: a base framework, limits and perspectives

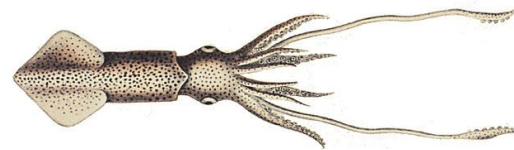
- Présentation du modèle ← **Focus**
- Evaluation du modèle par scénarios de simulation-estimation
- Application à 3 espèces démersales du golfe de Gascogne



Merlu



Sole



Loliginidae / encornet

Partie 1 : Modèle intégré et échantillonnage préférentiel

Integrated framework accounting for preferential sampling to estimate fish spatial distribution: a base framework, limits and perspectives

Rq: bcp moins de 0 que chez Sophie (30 - 40 %)

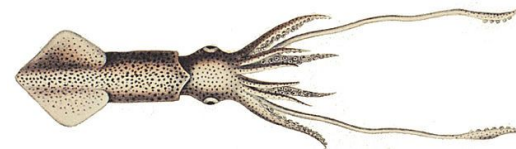
- Présentation du modèle ← **Focus**
- Evaluation du modèle par scénarios de simulation-estimation
- Application à 3 espèces démersales du golfe de Gascogne



Merlu



Sole



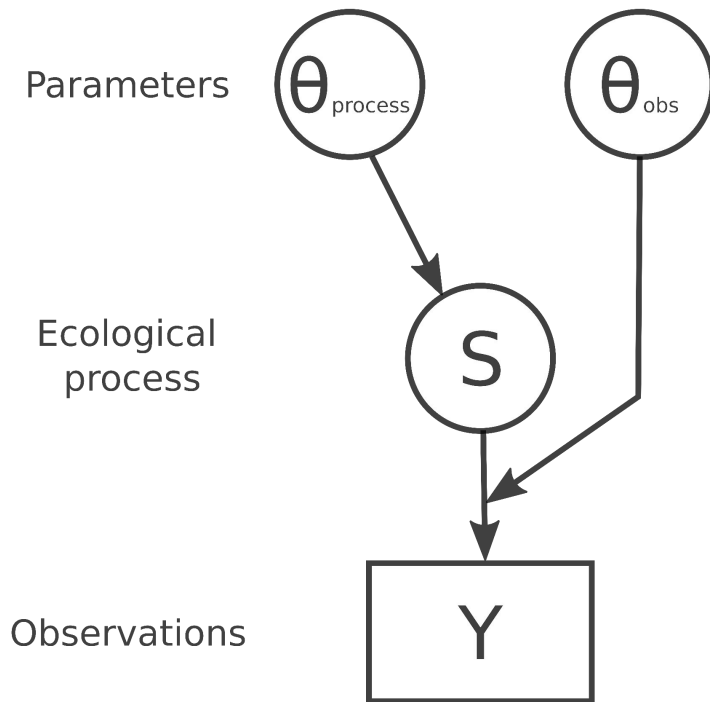
Loliginidae / encornet

Partie 1 : Modèle intégré et échantillonnage préférentiel

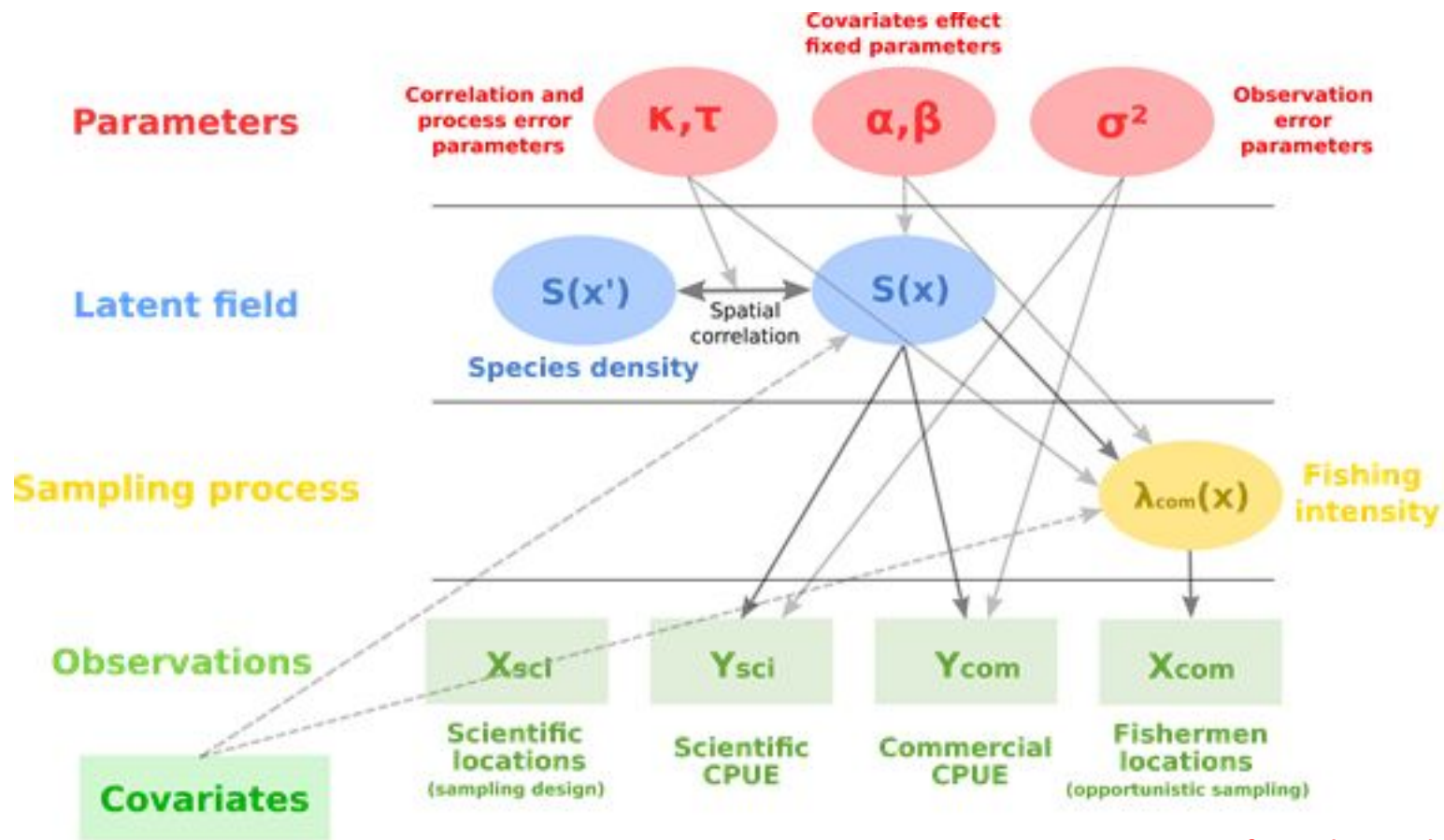
Petit rappel modèles hiérarchiques (ou modèles à espace d'état)

Modèles qui distinguent :

- **Observations**
- **Champ latent** (non-observé)
 - le processus écologique
(i.e. la variable d'intérêt)
- **Paramètres** contrôlant
 - la structure du champ latent
 - la variance/l'erreur d'observation

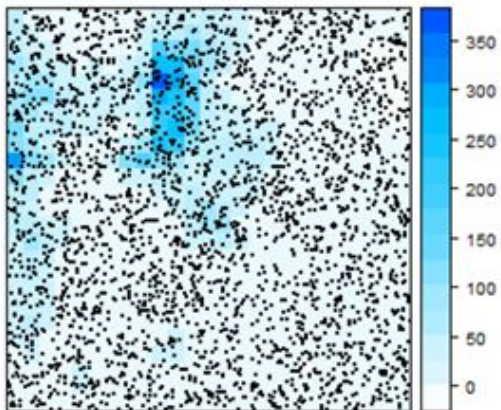


Partie 1 : Modèle intégré et échantillonnage préférentiel

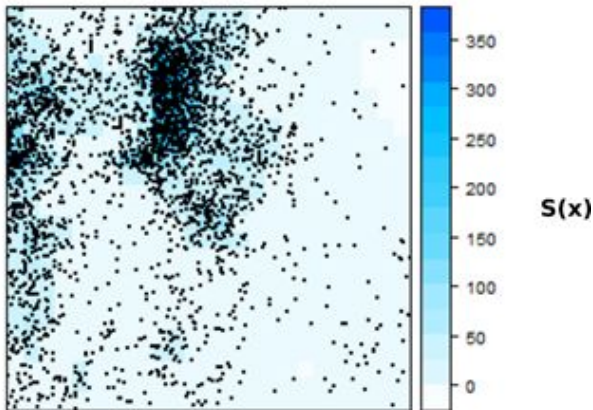


Partie 1 : Modèle intégré et échantillonnage préférentiel

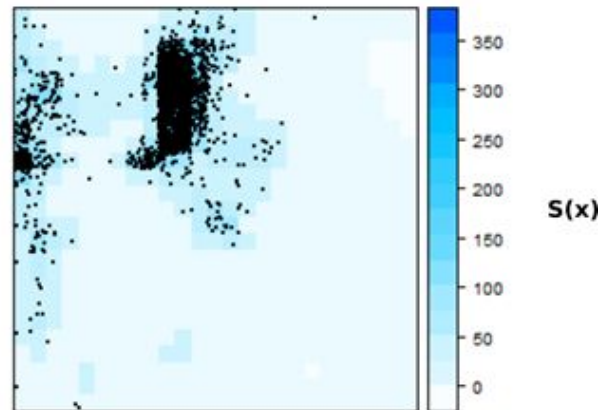
Echantillonnage
uniforme



Échantillonnage
préférentiel modéré



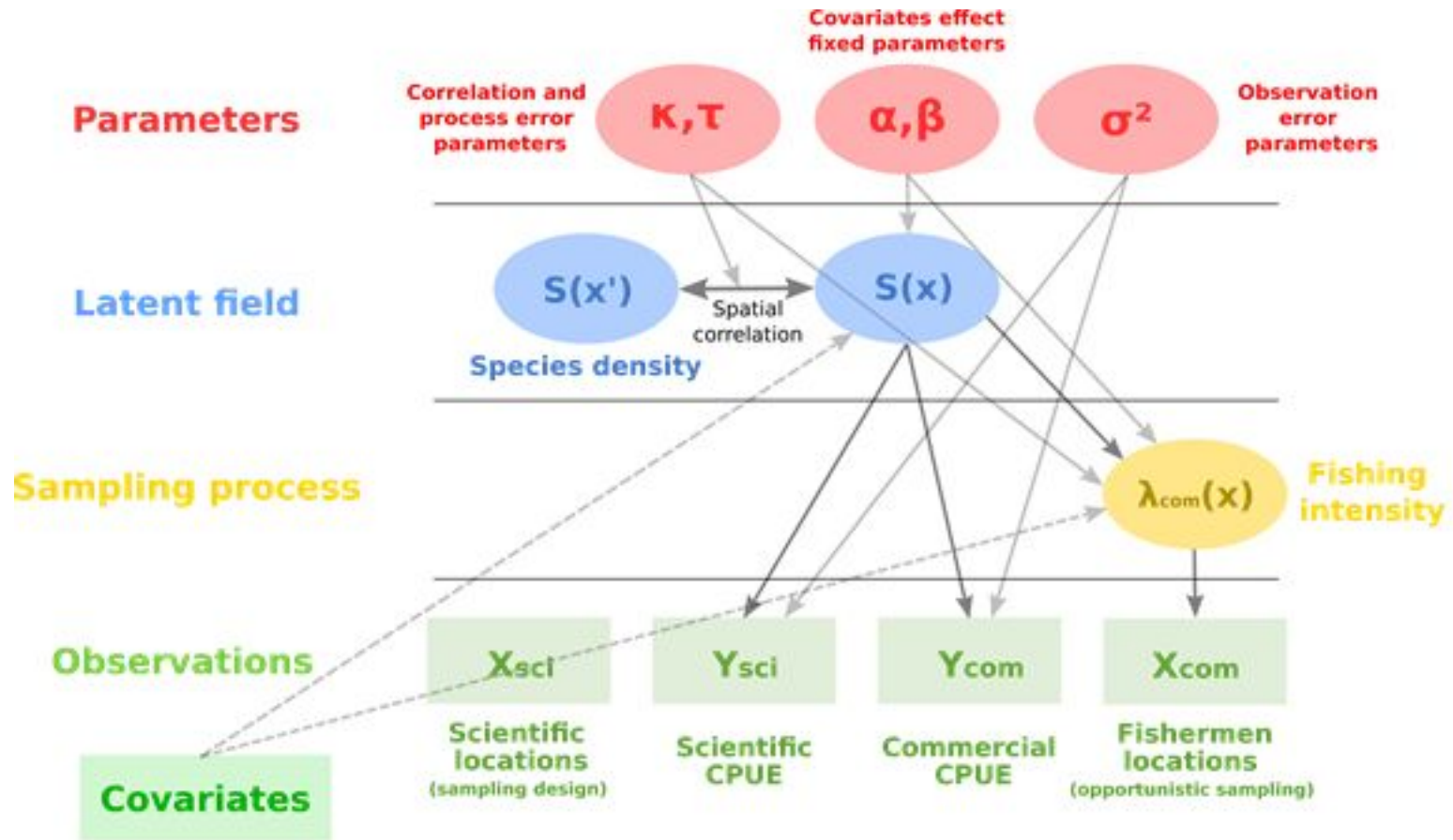
Échantillonnage
préférentiel fort



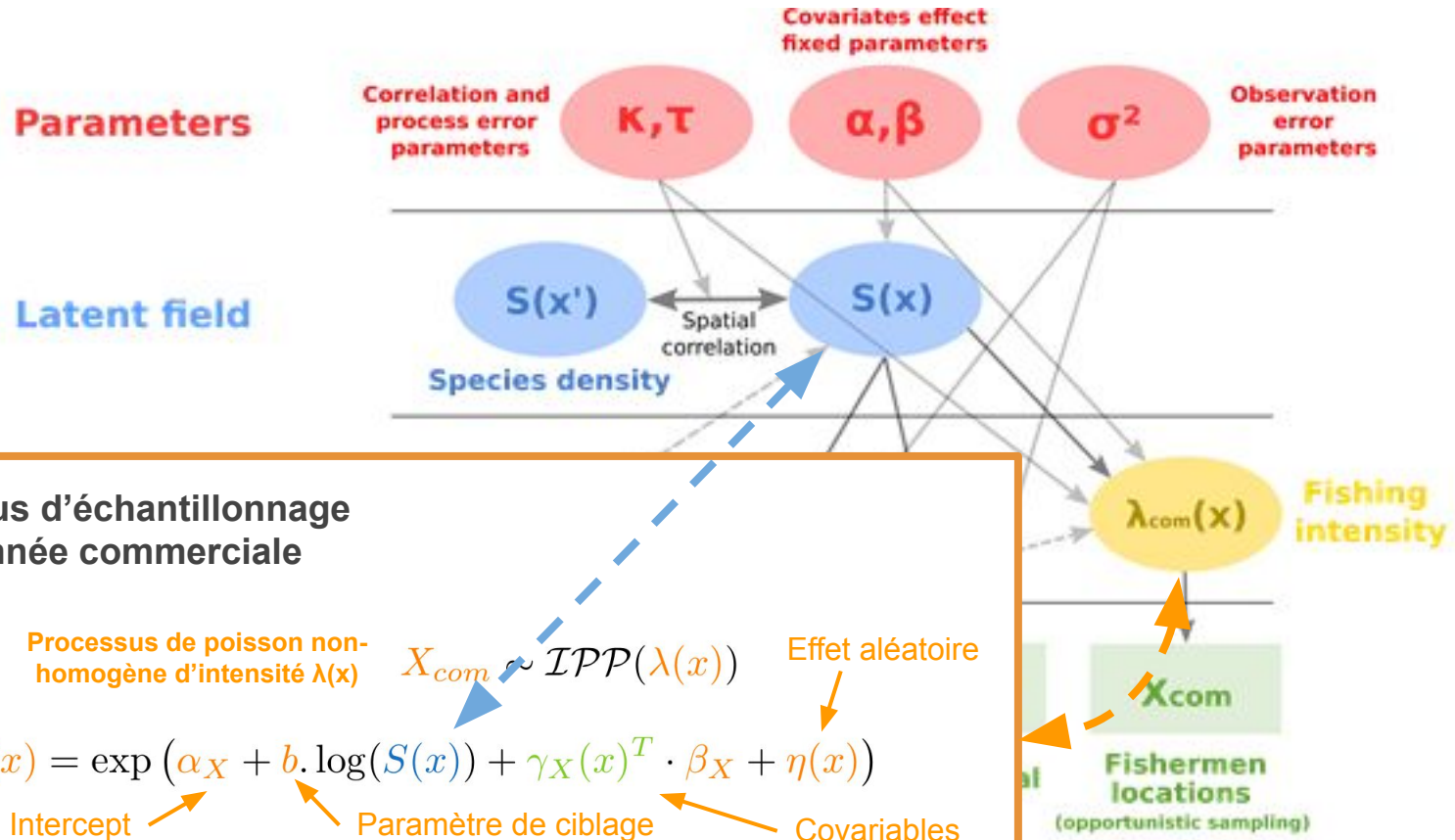
Simulation du processus ponctuel pour 3 niveaux d'échantillonnage préférentiel

Bleu : champ latent, points : échantillons commerciaux

Partie 1 : Modèle intégré et échantillonnage préférentiel



Partie 1 : Modèle intégré et échantillonnage préférentiel



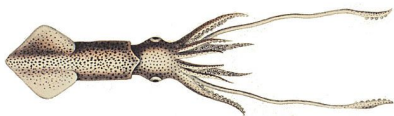
Partie 1 : Modèle intégré et échantillonnage préférentiel



$b = 1$

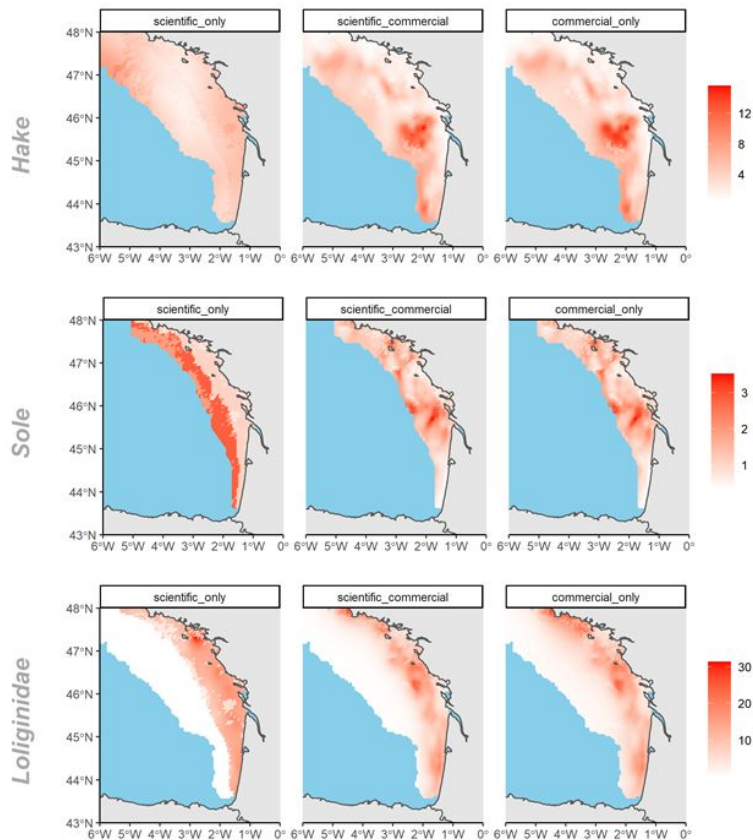


$b = 2$

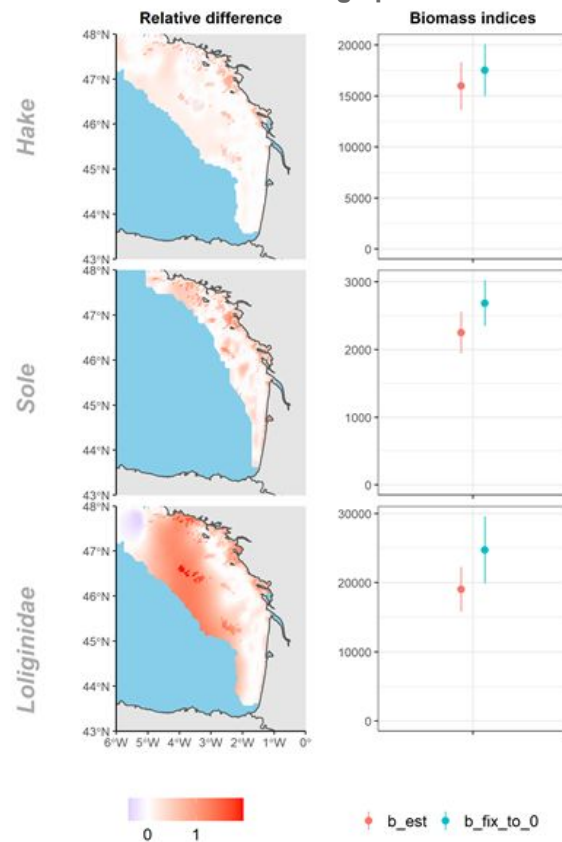


$b = 3.5$

Contribution des sources de données dans l'inférence



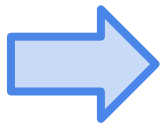
Pris en compte de l'échantillonnage préférentiel



Partie 1 : Modèle intégré et échantillonnage préférentiel

Plusieurs questions

- Données commerciales >> Données scientifiques
⇒ Comment faire en sorte que les deux sources de données informent l'estimation ?
(data-weighting / model misspecification)
- Ne pas modéliser l'échantillonnage préférentiel
⇒ Surestimation de la biomasse
- L'approche permet de prendre en compte de façon parcimonieuse l'échantillonnage préférentiel
⇒ ⚠ simplifie beaucoup l'ensemble des processus de ciblage



Passage à une approche
spatio-temporelle

Partie 2 : Approche spatio-temporelle

Can we trust commercial landings data to identify essential habitats of harvested fish ?

- Passage à un modèle spatio-temporel
- Analyse des patrons de distribution pour la sole commune
- Evolution de l'échantillonnage préférentiel

Modèle spatio-temporel



Sole

Modèle spatial - résultats préliminaires dans le cadre du projet MACCO



Bar



Merlu



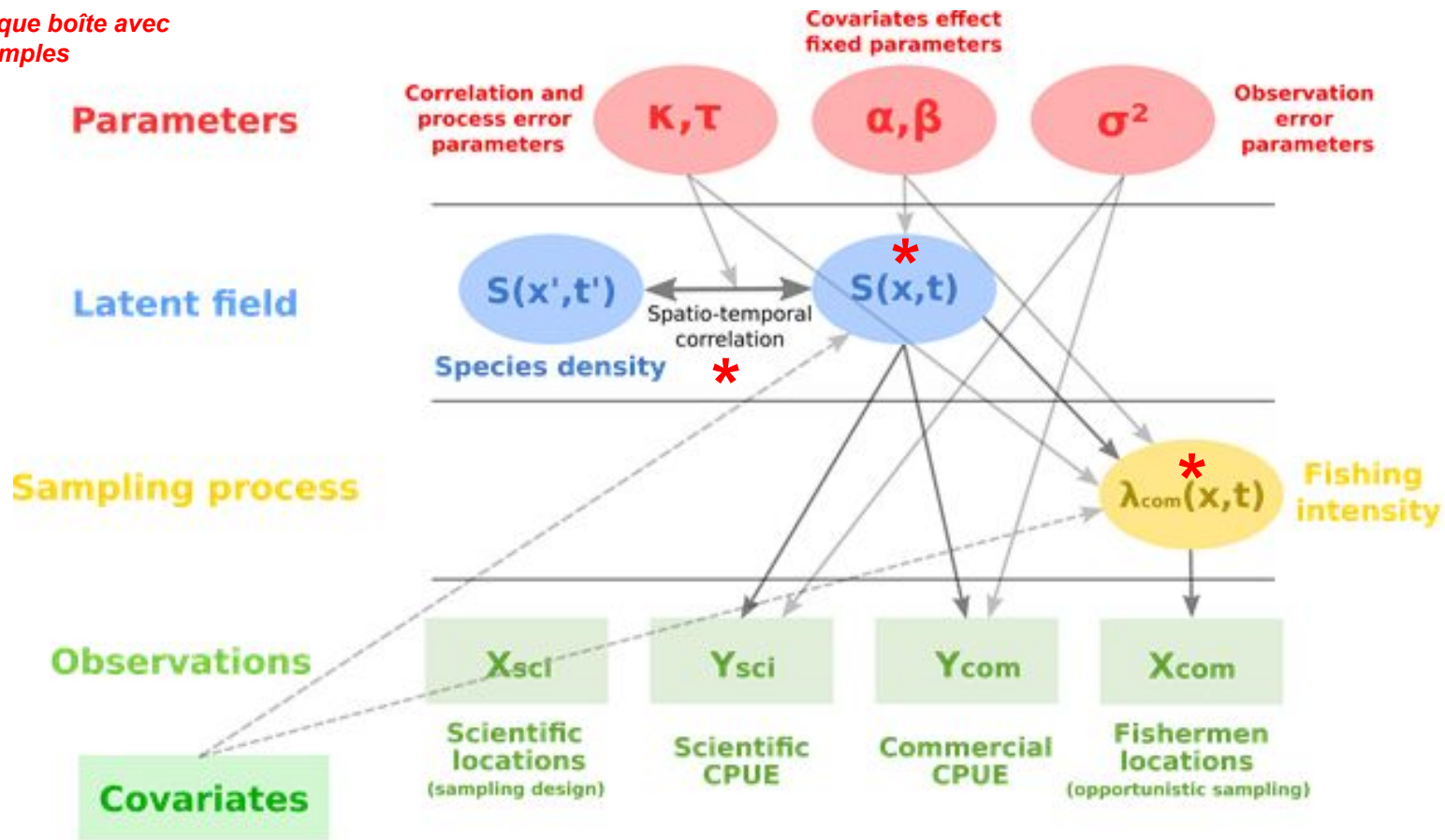
Merlan



Baudroie

Partie 2 : Approche spatio-temporelle

Décrire chaque boîte avec
des mots simples



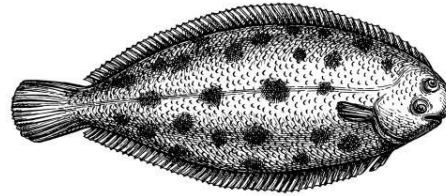
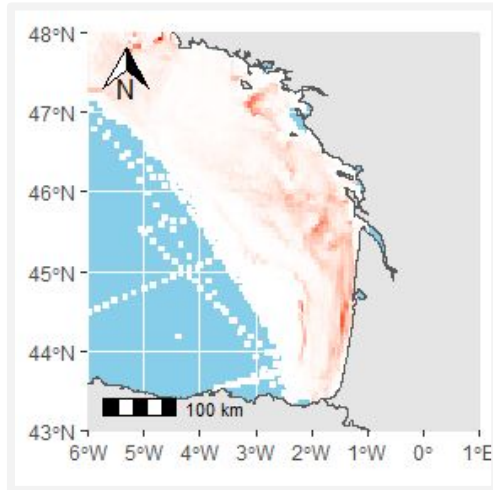
Partie 2 : Approche spatio-temporelle

Modèle spatial → Modèle spatio-temporel

- Auto-corrélation temporelle du **champ de biomasse** ($S(x,t)$)
- Niveau moyen de la **biomasse** : effet année / effet saison
- **Paramètre de ciblage** (b) :
 - Effet année / effet saison
 - Effet aléatoire gaussien
- Niveau moyen d'**intensité de pêche** : effet année / effet saison

Partie 2 : Approche spatio-temporelle

**Modèle
spatio-temporel**



SSB

Définition

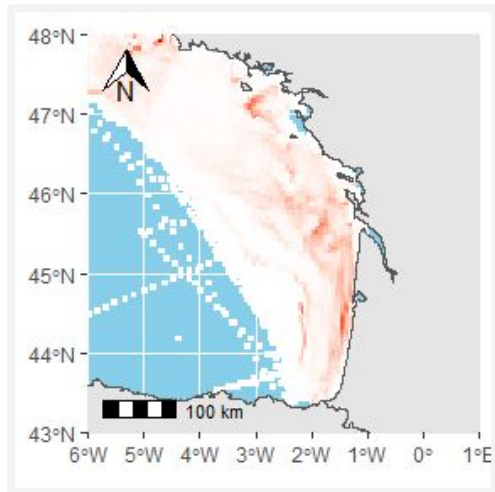
Effort - métier OTB_DEF

Partie 2 : Approche spatio-temporelle

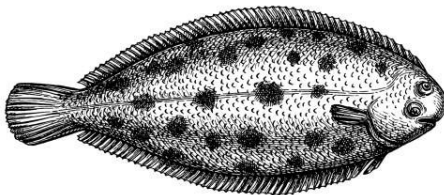
Modèle
spatio-temporel

Ajustement du modèle sur la
période 2008 - 2018 pour la
sole du golfe de Gascogne

Pas de tps mensuel = 132 cartes



Effort - métier OTB_DEF

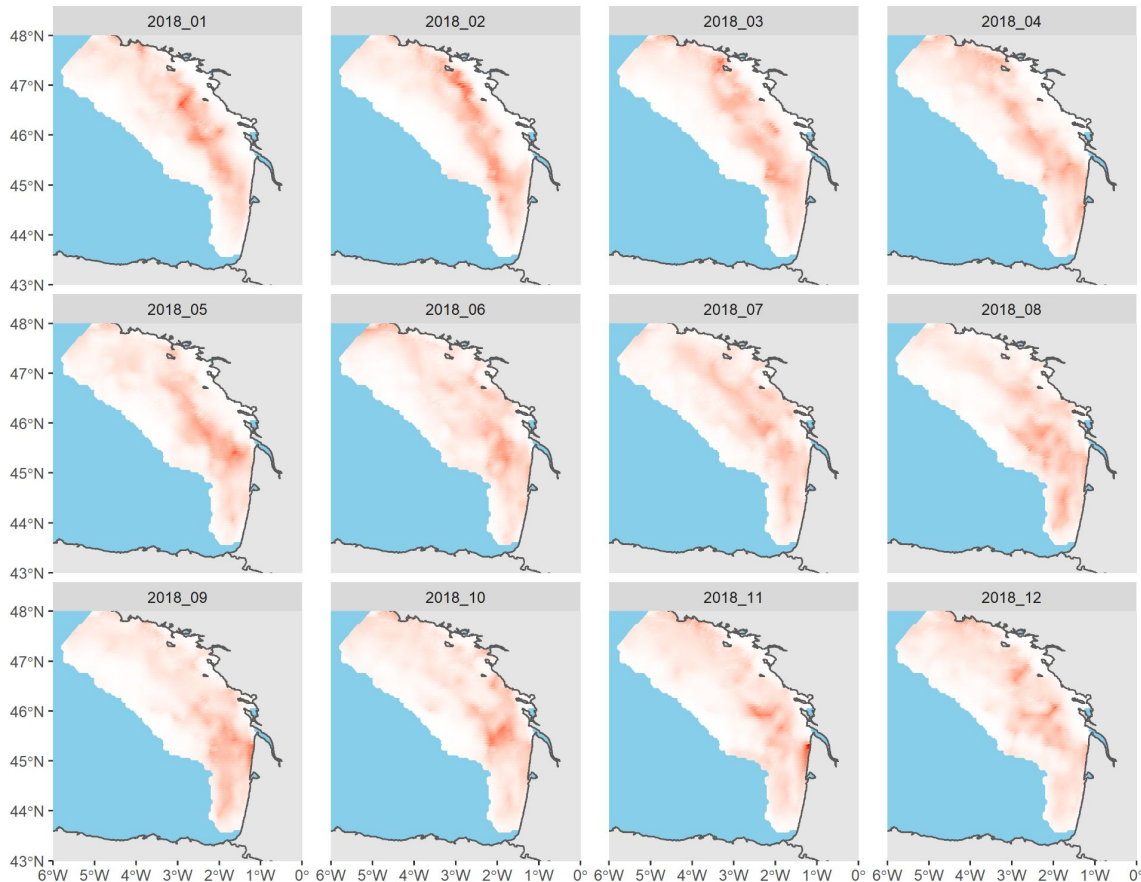


SSB

Définition

Partie 2 : Approche spatio-temporelle

**Prédictions du modèle
pour l'année 2018**

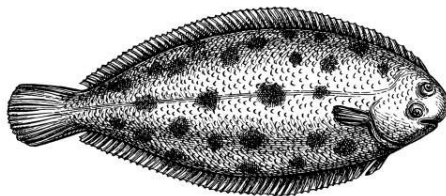
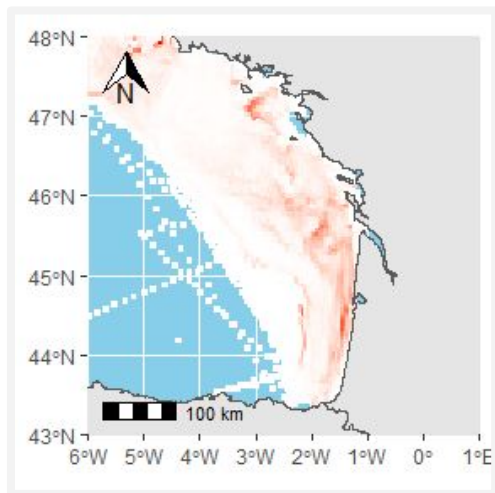


Partie 2 : Approche spatio-temporelle

Modèle
spatio-temporel

Ajustement du modèle sur la
période 2008 - 2018 pour la
sole du golfe de Gascogne

Pas de tps mensuel = 132 cartes



SSB

Effort - métier OTB_DEF

Partie 2 : Approche spatio-temporelle

Modèle
spatio-temporel

Ajustement du modèle sur la
période 2008 - 2018 pour la
sole du golfe de Gascogne

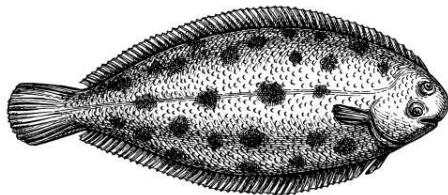
Pas de tps mensuel = 132 cartes

EOF

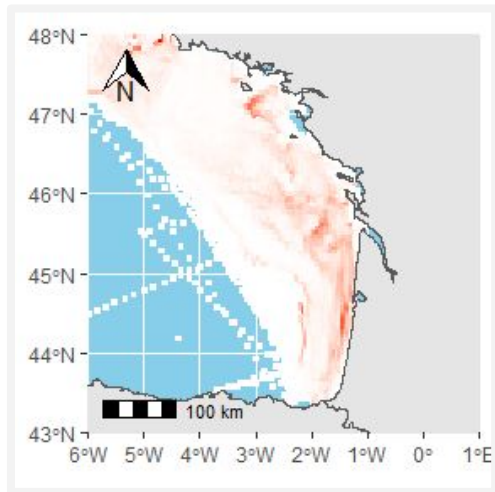
capture des principaux
patrons de distribution

=

Qql. cartes + time series



SSB



Effort - métier OTB_DEF

Partie 2 : Approche spatio-temporelle

Empirical Orthogonal Functions (EOF)

$$\vec{s}_t^* = \sum_{k=1}^K \alpha_k(t) \vec{p}^k + \vec{n}_t$$

Diagram illustrating the Empirical Orthogonal Functions (EOF) model. The equation shows the normalized biomass field \vec{s}_t^* as a sum of loading factors $\alpha_k(t)$ multiplied by spatial patterns \vec{p}^k , plus a residual variability term \vec{n}_t . Arrows point from the labels below to the corresponding terms in the equation:

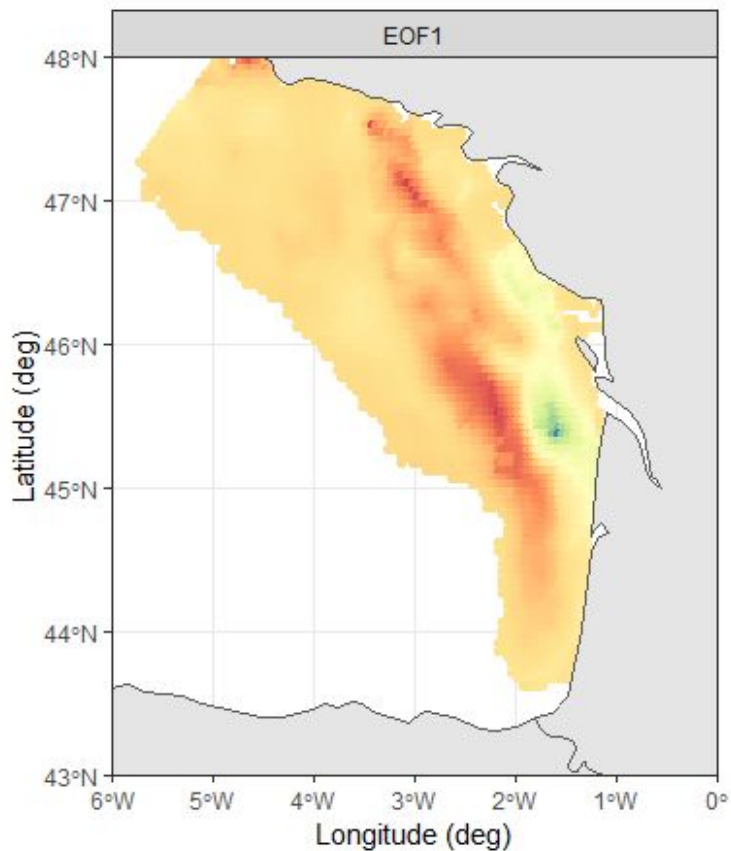
- Champ de biomasse normalisé (points to \vec{s}_t^*)
- Loading factors (points to $\alpha_k(t)$)
- Patrons spatiaux (points to \vec{p}^k)
- Variabilité résiduelle (points to \vec{n}_t)

Dans une **EOF**, les indices α_k et les patterns p^k sont construits tel que

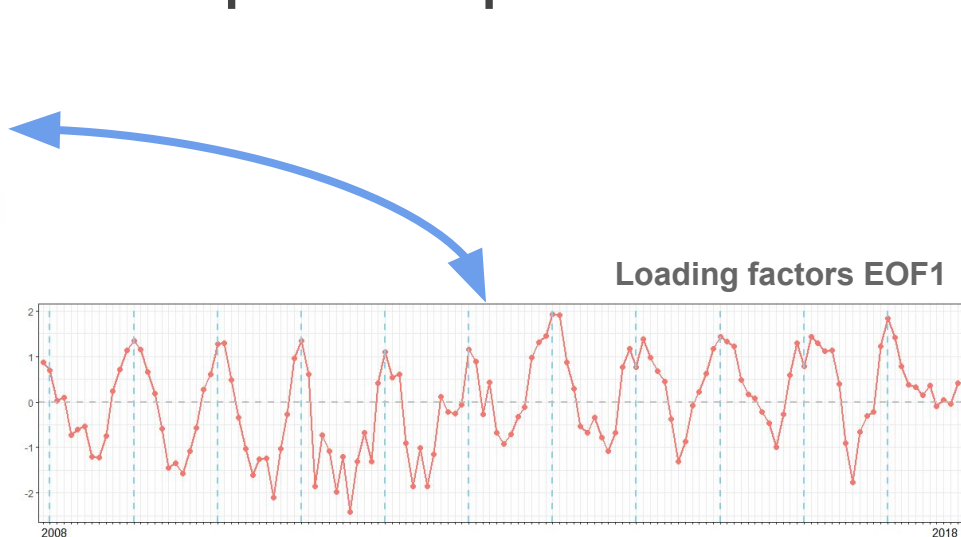
- la variance résiduelle n_t soit minimale
- les patrons p^k soient orthogonaux

\equiv ACP

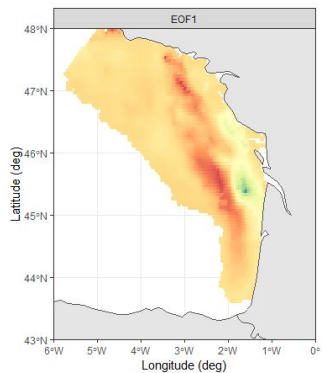
Partie 2 : Approche spatio-temporelle



= Patron de distribution
capturant le plus de variabilité

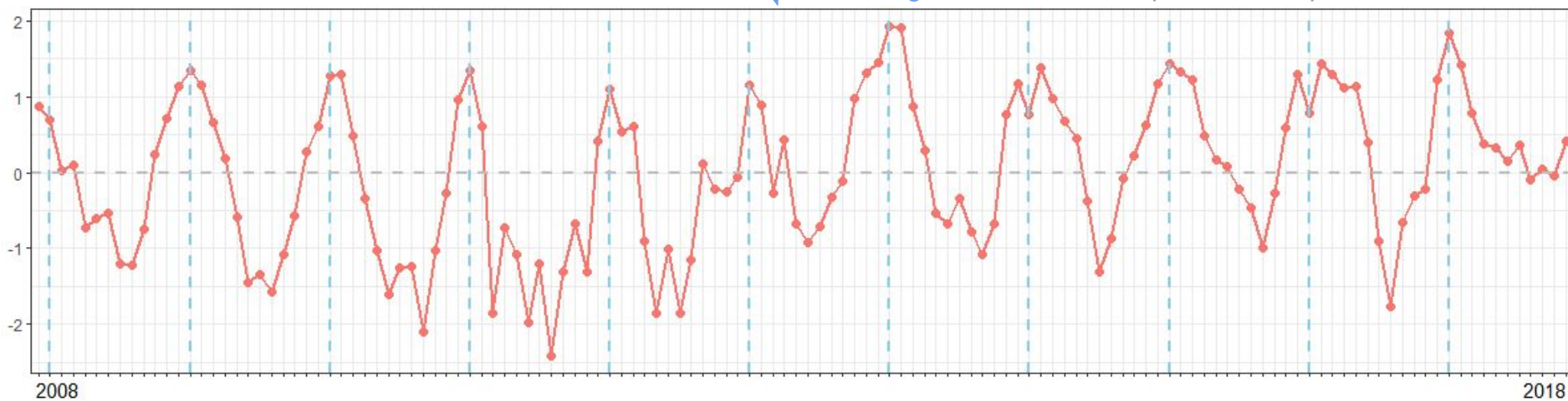


Partie 2 : Approche spatio-temporelle



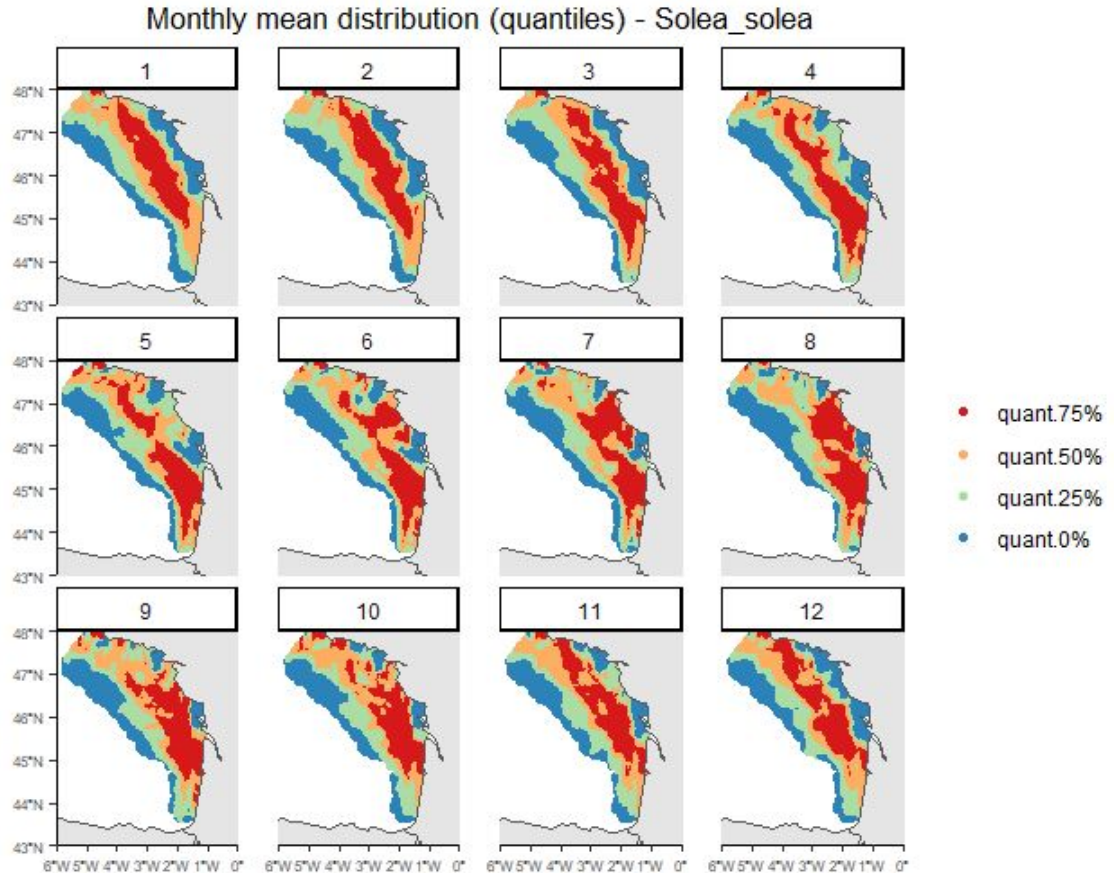
Loading factors EOF1

Ligne bleue : février ~ période de reproduction de la sole

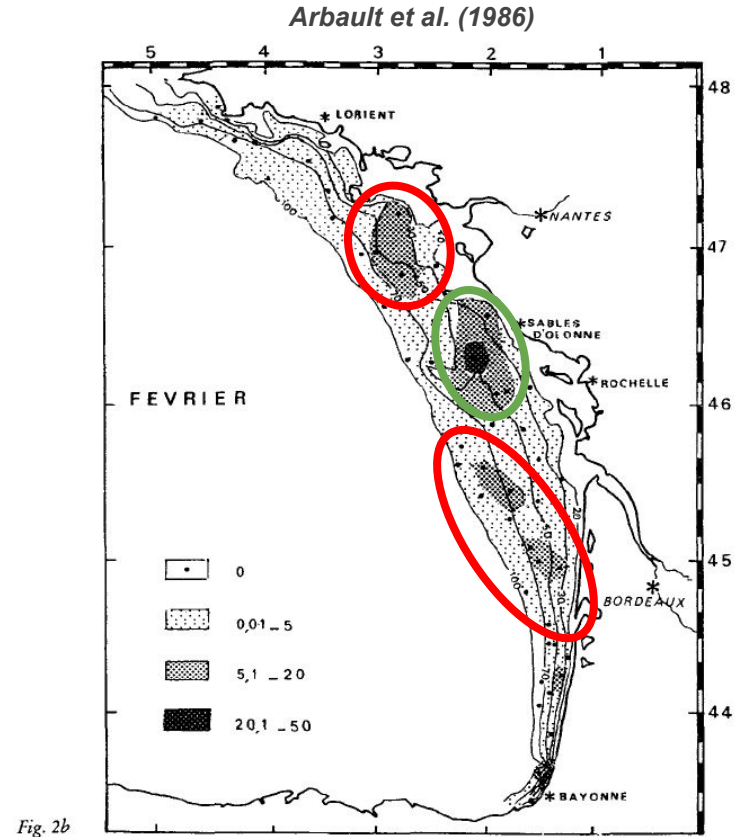
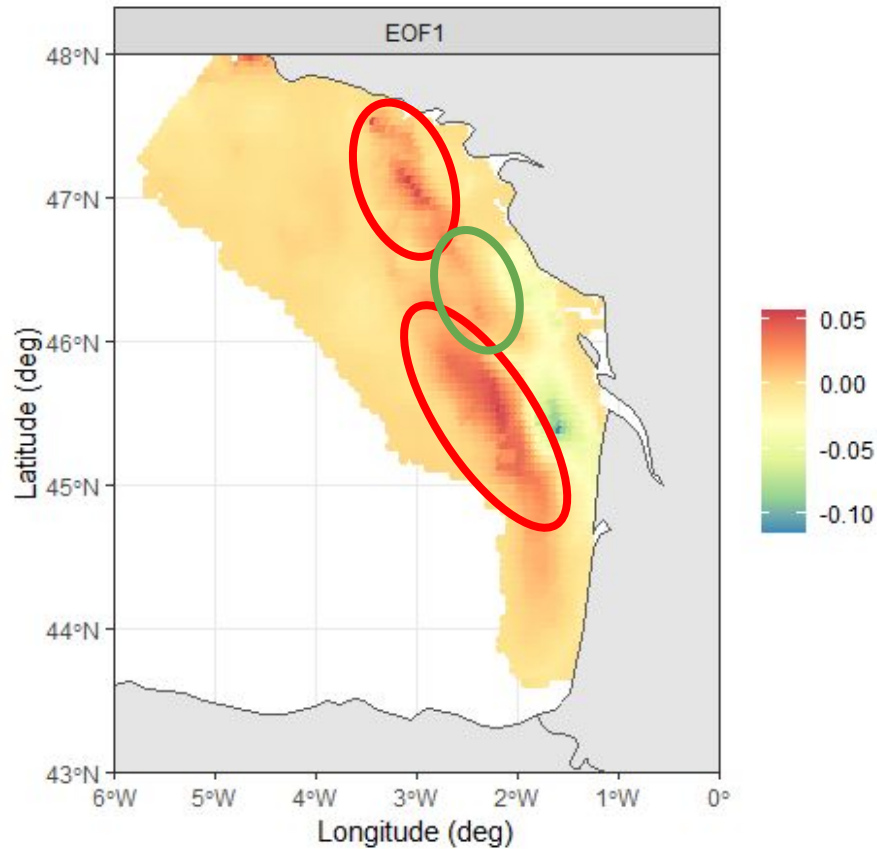


Partie 2 : Approche spatio-temporelle

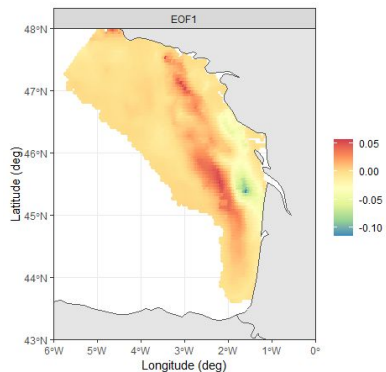
Migration 'côte-large' ⇒



Partie 2 : Approche spatio-temporelle



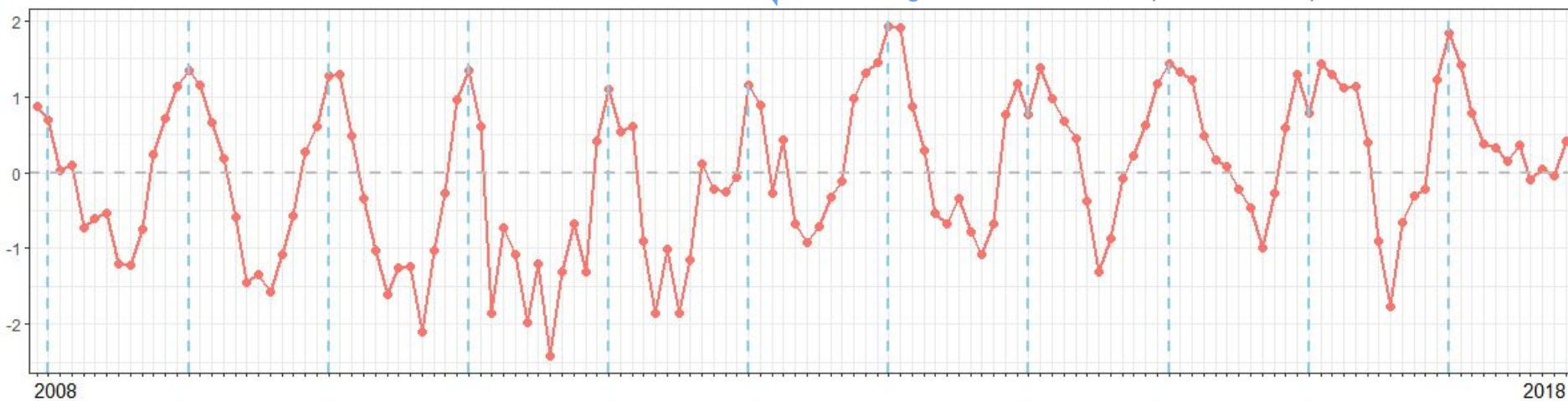
Partie 2 : Approche spatio-temporelle



⇒ Mise en évidence d'un patron de distribution saisonnier ⇔ zone et période de reproduction

Loading factors EOF1

Ligne bleue : février ~ période de reproduction de la sole

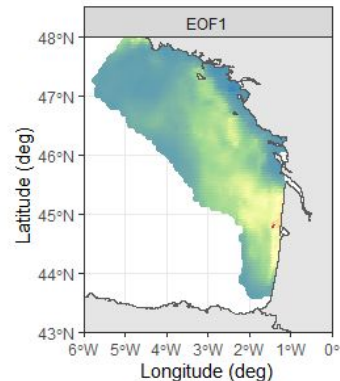
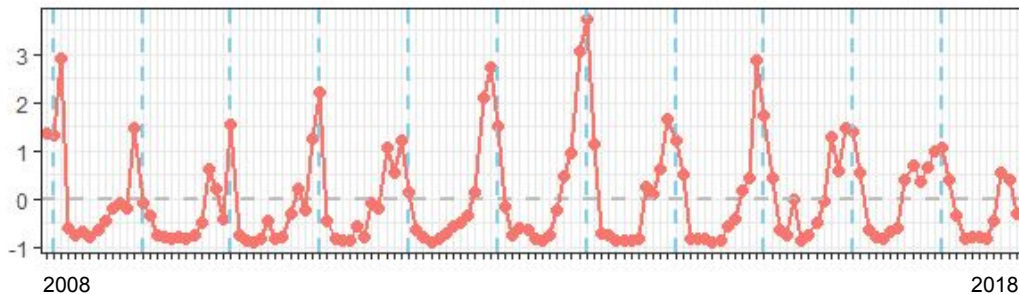


Partie 2 : Approche spatio-temporelle

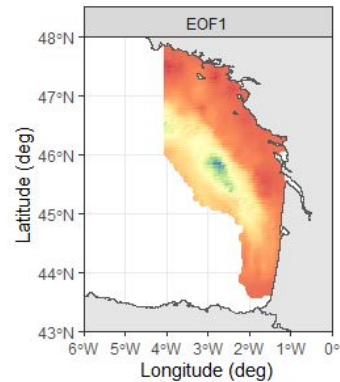
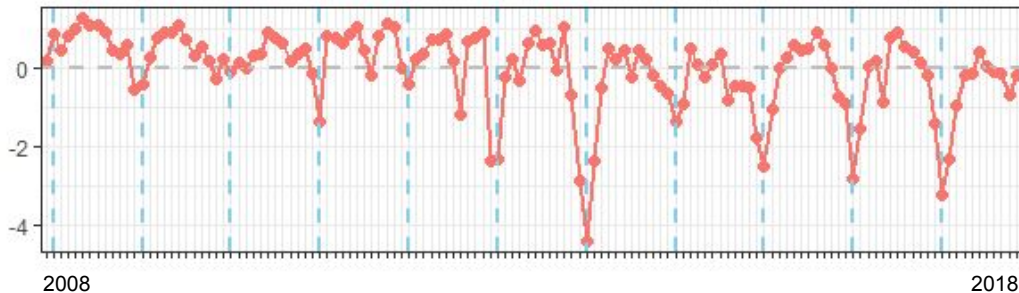
Et pour les autres espèces ?



Bar

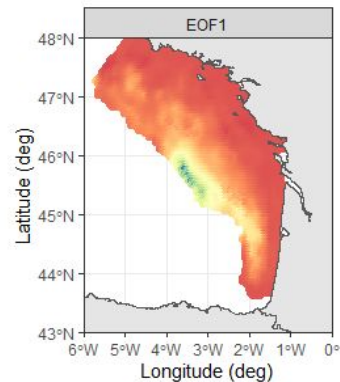
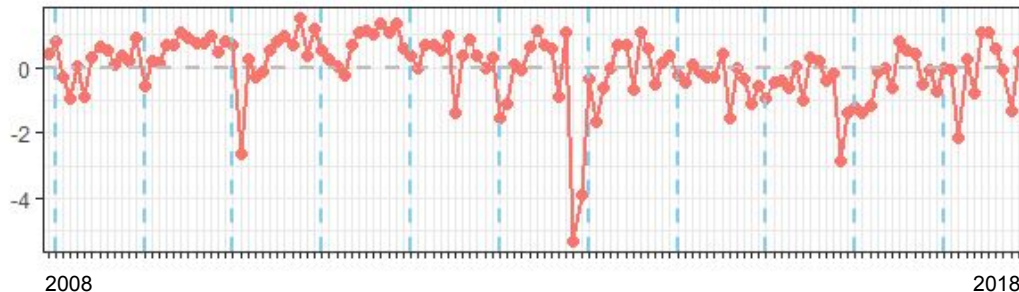
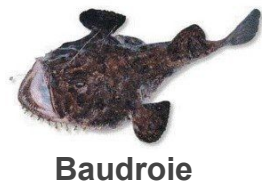
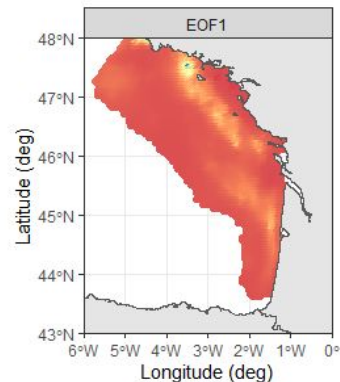
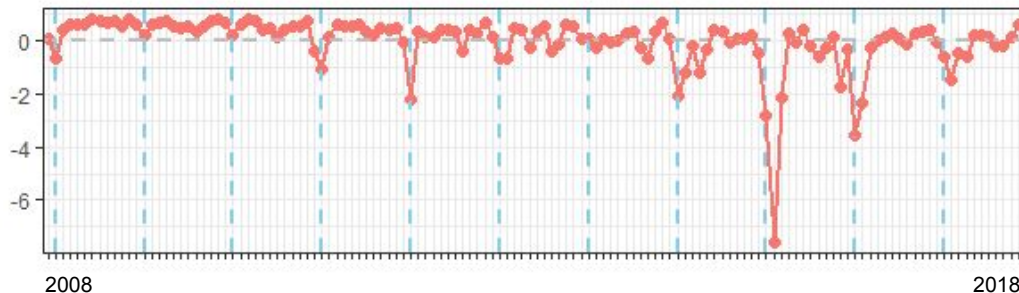
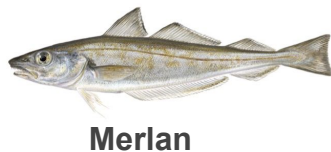


Merlu



Partie 2 : Approche spatio-temporelle

Et pour les autres espèces ?



Partie 2 : Approche spatio-temporelle

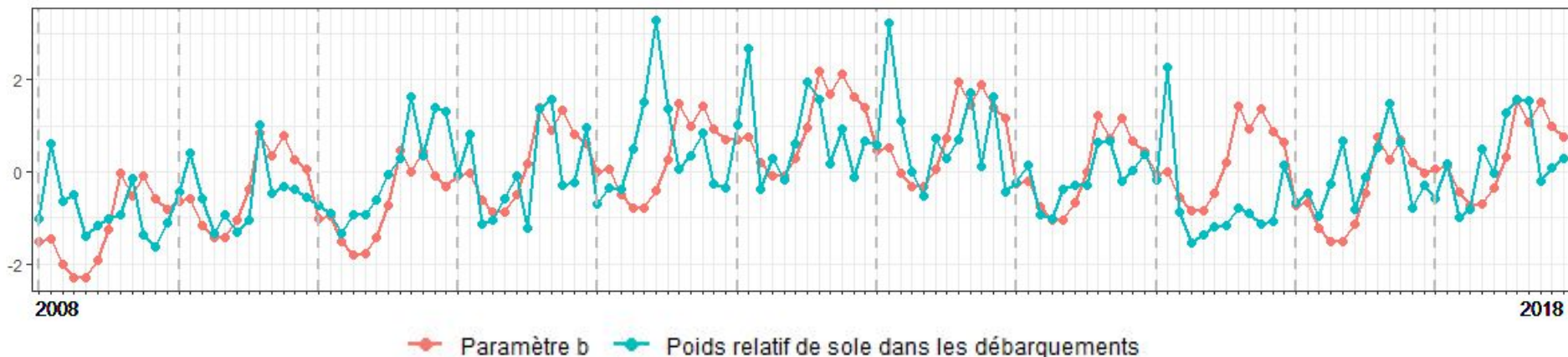
Evolution temporelle du paramètre de ciblage (b)

Rappel paramètre b

Paramètre b
[0.8 ; 1.2]

vs.

**Proportion relative de sole
dans les débarquements**
(→ indicateur de ciblage)
[2,5% ; 7.5%] des débarquements



Coefficient de corrélation = 0.5

Partie 2 : Approche spatio-temporelle

En résumé

Approche permettant :

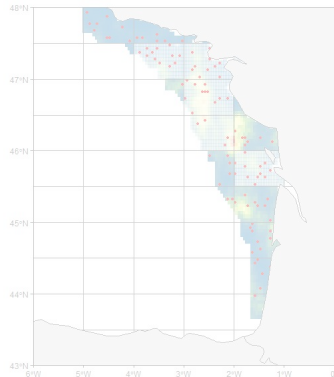
- de capturer les principaux patrons de distributions des espèces halieutiques à partir de la donnée commerciale ↔ reproduction
- d'estimer la force de l'échantillonnage préférentiel vis-à-vis de l'espèce d'intérêt et son évolution dans le temps

Partie 3 : Changement de support

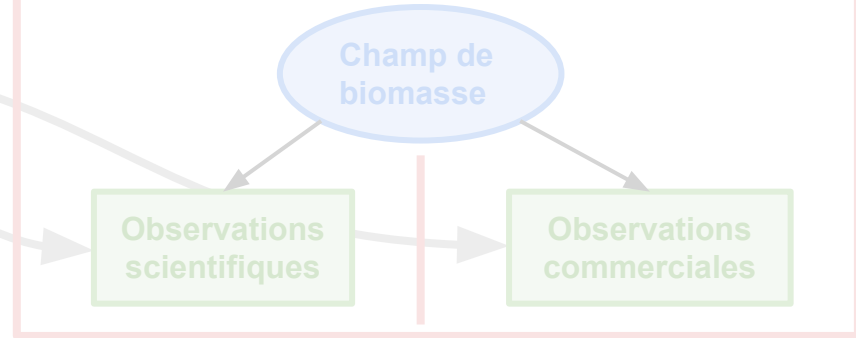
**Going from coarse landings data
to fine scale species distribution**

Modèle de simulation

- Champ de biomasse (cov + RE)
- Données commerciales
- Données scientifiques



Modèle d'estimation



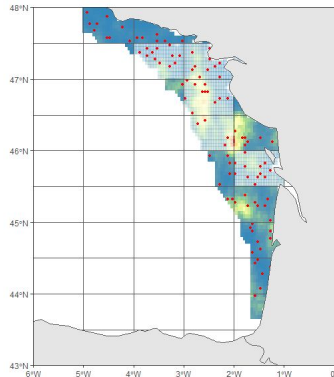
Rappel simulation-estimation

Partie 3 : Changement de support

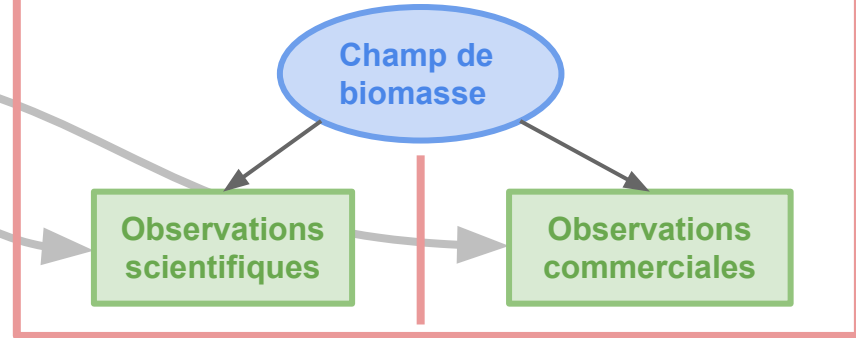
Going from coarse landings data
to fine scale species distribution

Modèle de simulation

- Champ de biomasse (cov + RE)
- Données commerciales
- Données scientifiques



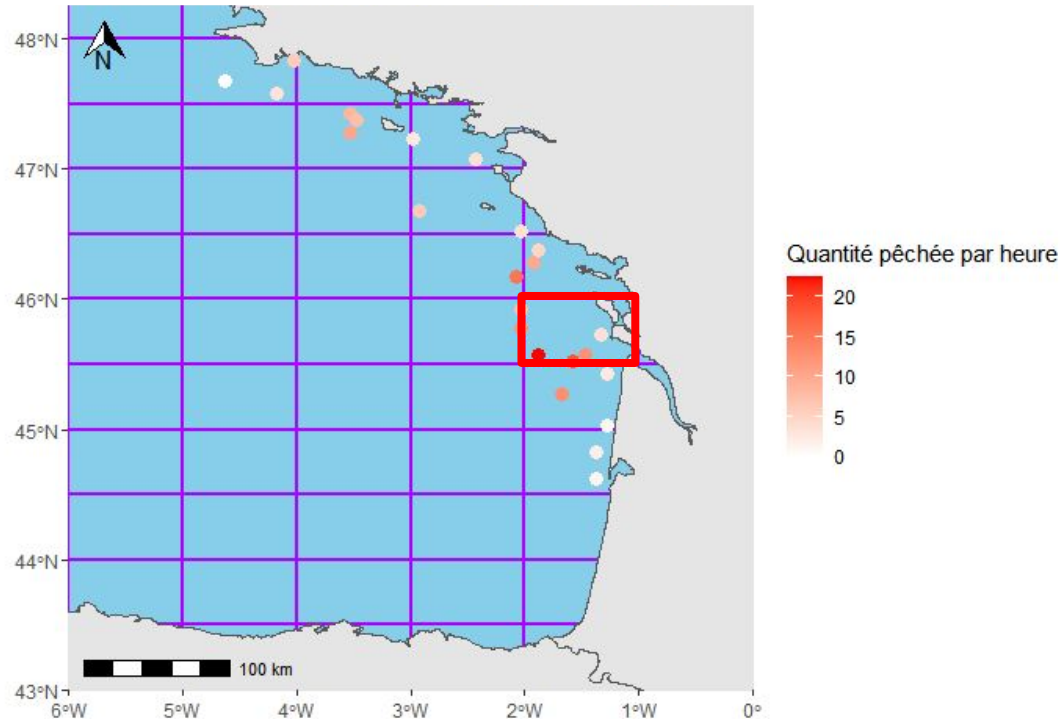
Modèle d'estimation



Rappel simulation-estimation

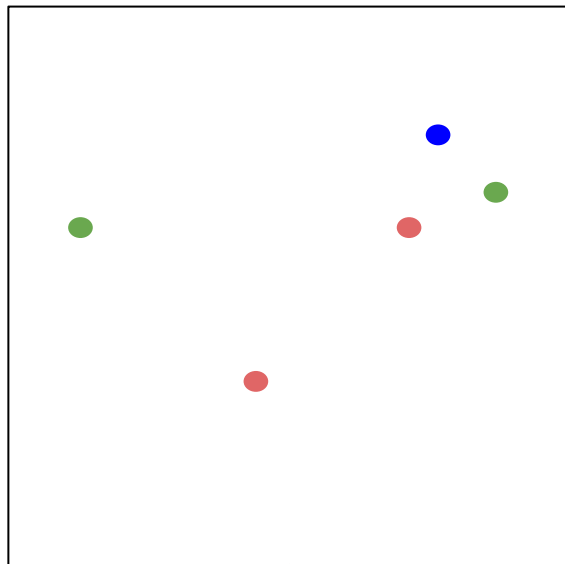
Partie 3 : Changement de support

Les données de déclaration de pêche sont déclarées à l'échelle de carrés statistiques



Partie 3 : Changement de support

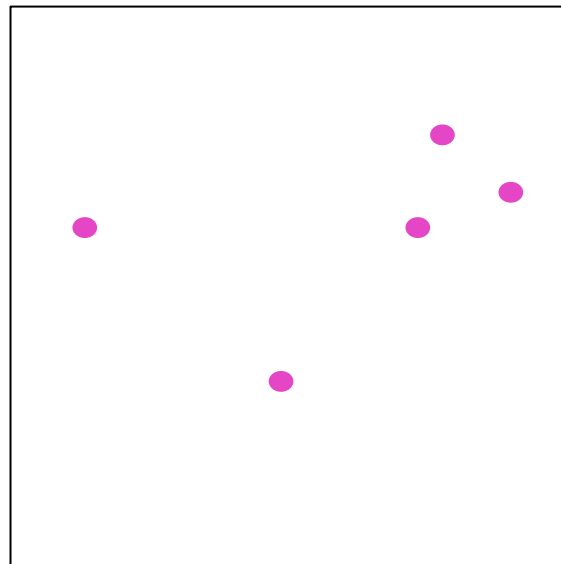
Captures exactes



Quantité pêchée
par heure

- 20
- 10
- 5

Captures ré-allouées



Quantité pêchée
par heure

- 10

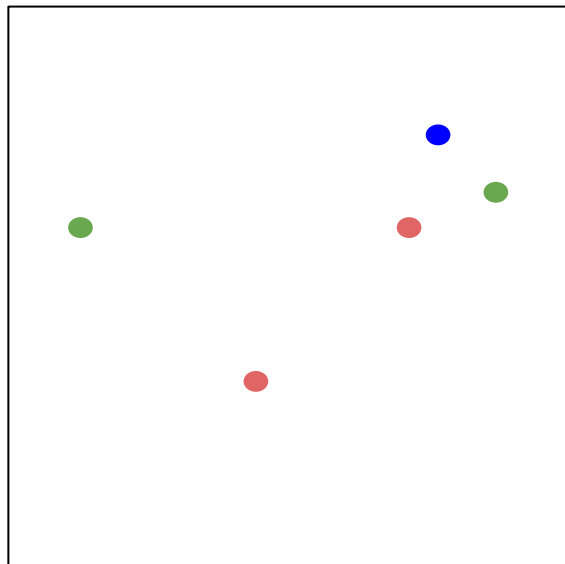
Quantité totale pêchée (y_{total}) = **50**
Nombre de points de pêche (n) = **5**

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

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

Partie 3 : Changement de support

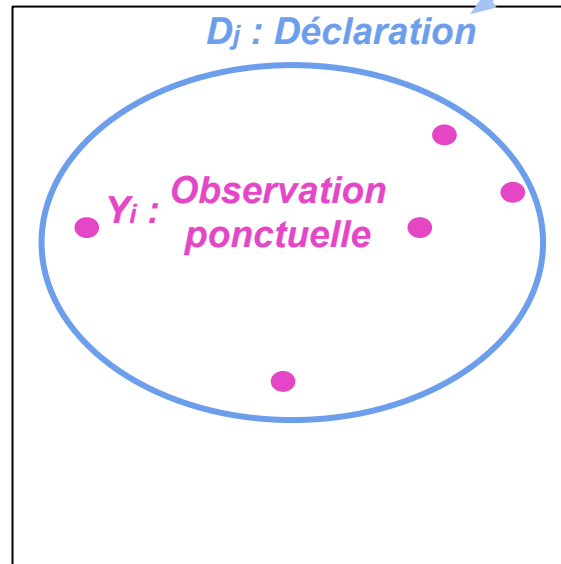
Captures exactes



Quantité pêchée
par heure

- 20
- 10
- 5

Captures ré-allouées



Résolution:
Rectangle stat. x jour
x marée x engin de pêche

Quantité pêchée
par heure

- 10

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

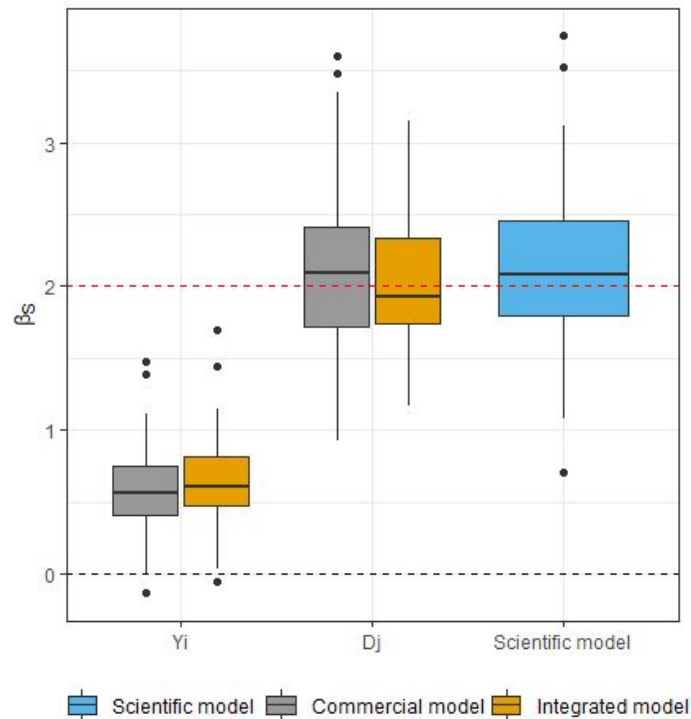
$$\frac{(y_{\text{total}})}{(n)} = \frac{50}{5} = 10 = y_{\text{réallouée}}$$

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

Partie 3 : Changement de support

Résultats

Paramètre de la relation espèce-habitat



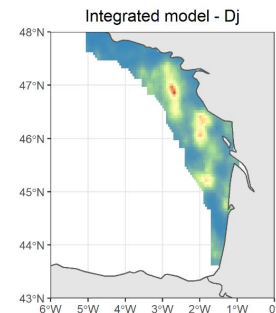
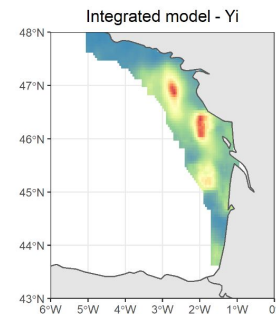
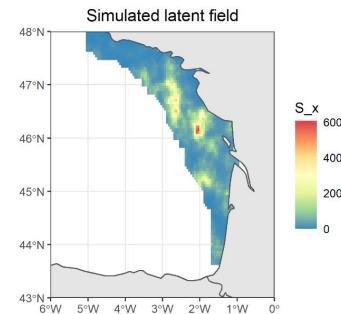
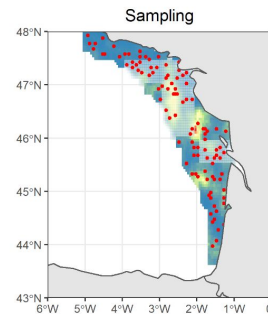
Y_i : Observation ponctuelle

$$Y_i \sim \text{Zinfl Lognorm}(\mu_i, \sigma^2)$$

D_j : Déclaration

$$D_j = \sum_{i \in \mathcal{P}_j} Y_i$$

Vecteur de tous les points de pêche associés à la donnée de déclaration j



Annexes

Champ latent

Champ de biomasse

$$S(x, t) = \exp(\alpha_S(t) + \Gamma_S(x)^T \cdot \beta_S + \epsilon(x, t))$$

Effet aléatoire
spatio-temporel
(AR1)

$$\epsilon(x, t) = \rho \cdot \epsilon(x, t - 1) + \omega(x, t)$$

Covariables
(propres au
champ latent)

$$\omega(x, t) \sim N(0, Q_S^{-1})$$

Effet aléatoire
spatio-temporel (AR1)

$$\epsilon(x, 0) \sim N(0, \sigma_0^2 \cdot Id)$$

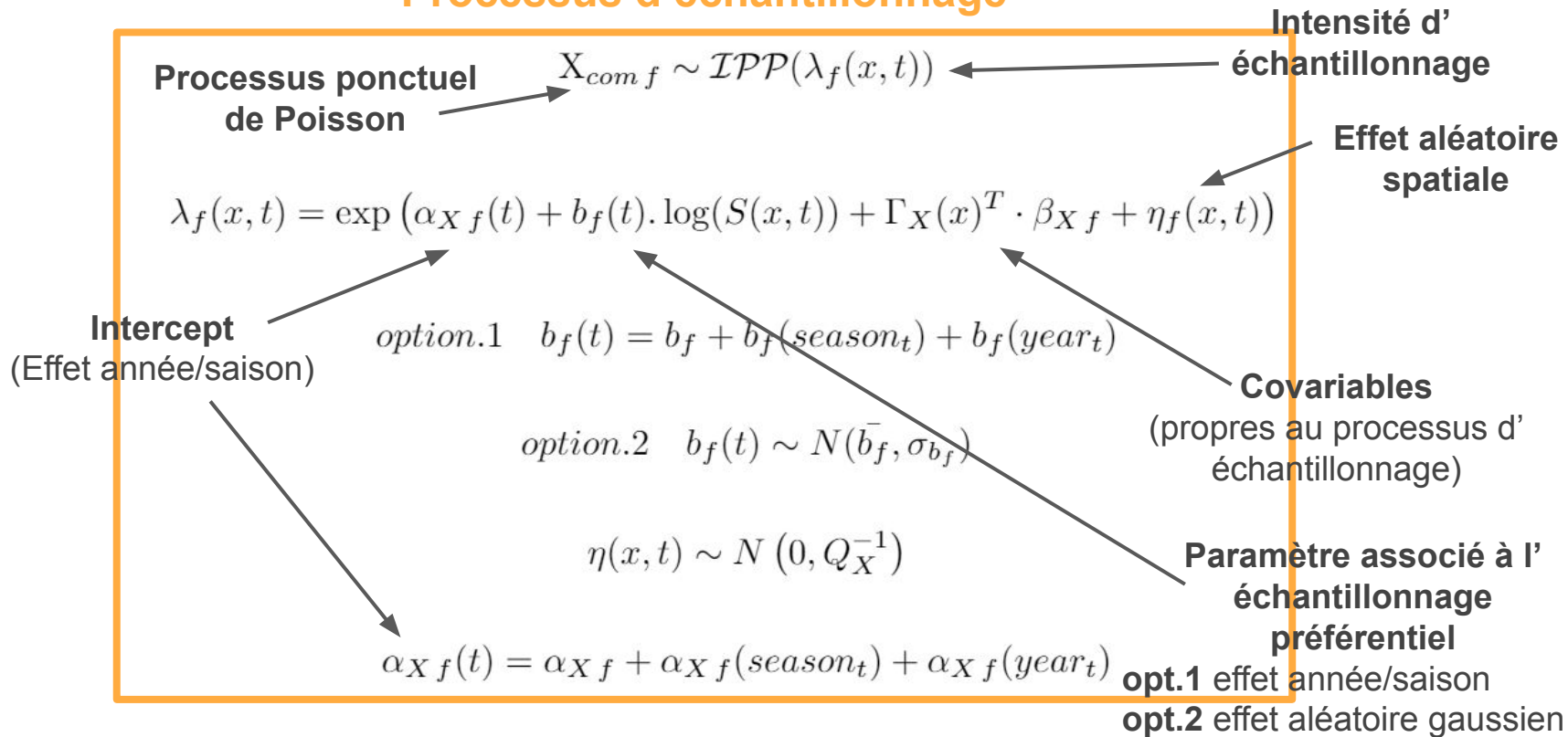
Initialisation de
l'AR1

Intercept

(Effet année/saison)

$$\alpha_S(t) = \alpha_S + \alpha_S(\text{season}_t) + \alpha_S(\text{year}_t)$$

Processus d'échantillonnage



Processus d'observation

→ homogène dans le temps et l'espace

Paramètre de
zero-inflation

$$P(Y_i = y_i | x_i, t_i, S(x_i, t_i)) = \begin{cases} \exp(-e^{\xi_f} \cdot \mu_f(x_i, t_i)) & \text{if } y_i = 0 \\ (1 - \exp(-e^{\xi_f} \cdot \mu_f(x_i, t_i))) \cdot L\left(y_i, \frac{\mu_f(x_i, t_i)}{(1 - \exp(-e^{\xi_f} \cdot \mu_f(x_i, t_i)))}, \sigma_f^2\right) & \text{if } y_i > 0 \end{cases}$$

Espérance des
observations

$$\mu_f(x_i) = q_f \cdot S(x_i, t_i)$$

Champ de biomasse

$$q_{fest} = k_f * q_{ref}$$

Capturabilité

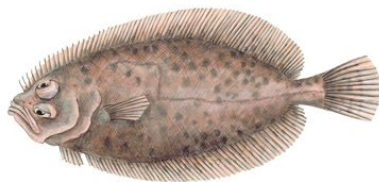
Scaling factor

Perspectives

→ Application à d'autres cas d'études / d'autres aires géographiques



Merlu



Cardine



Baudroie



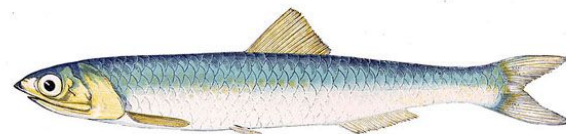
Merlan



Bar

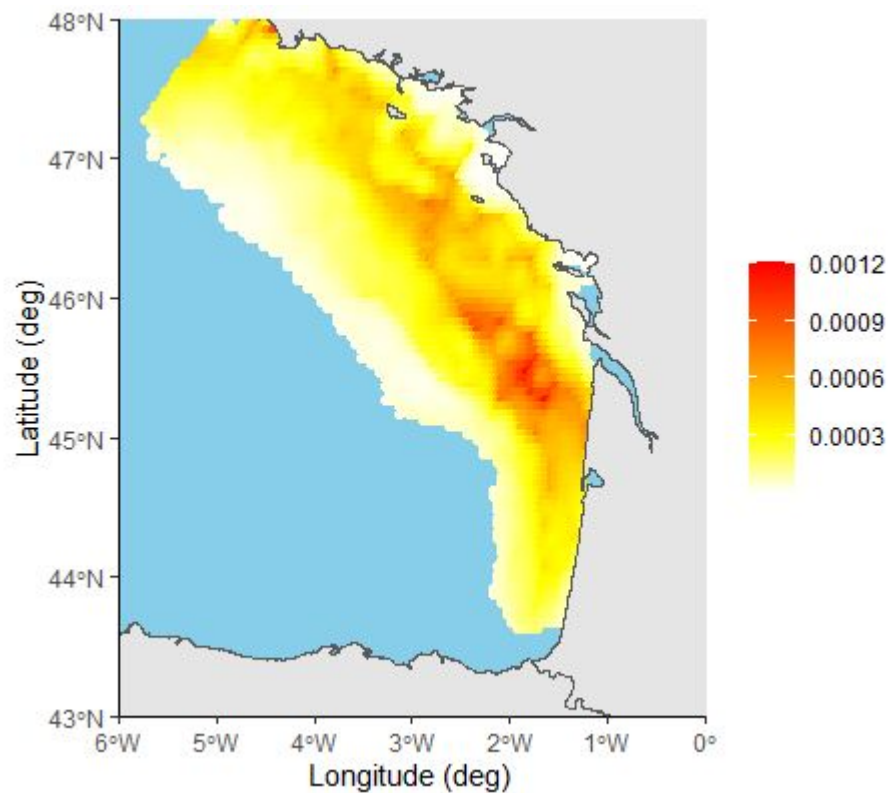


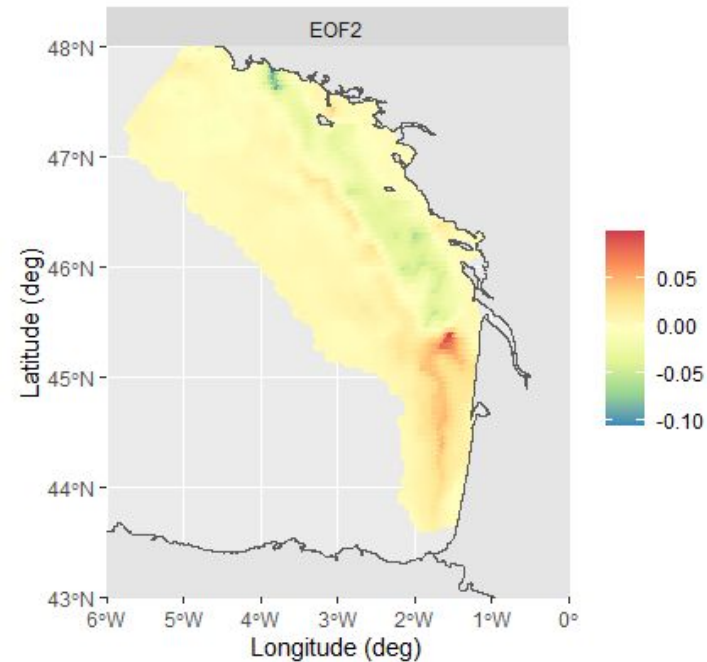
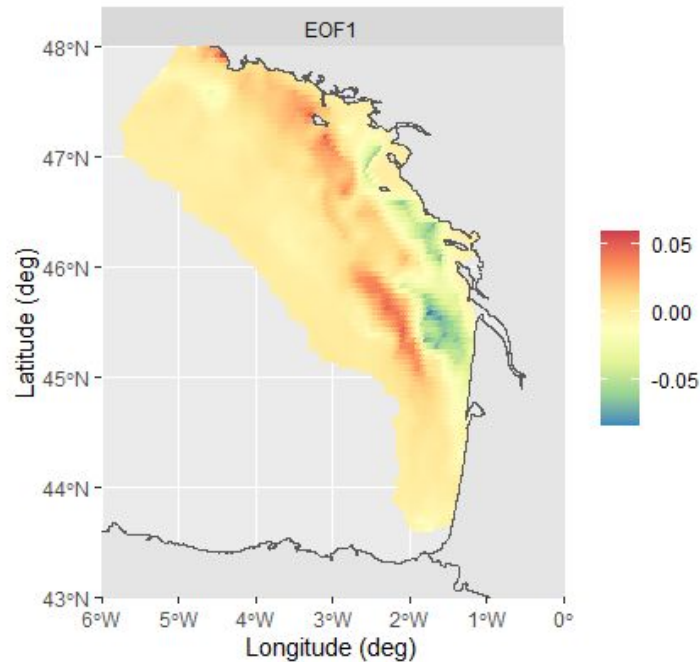
Sardine

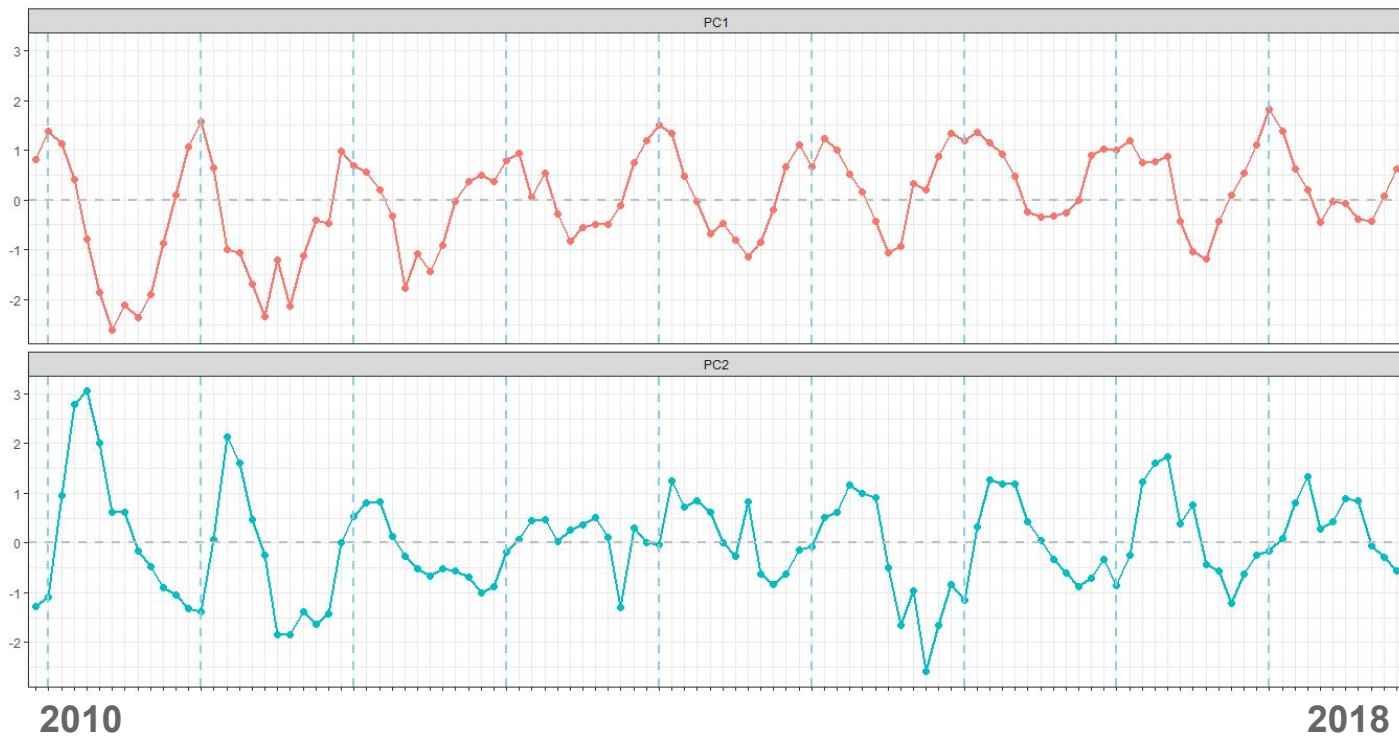
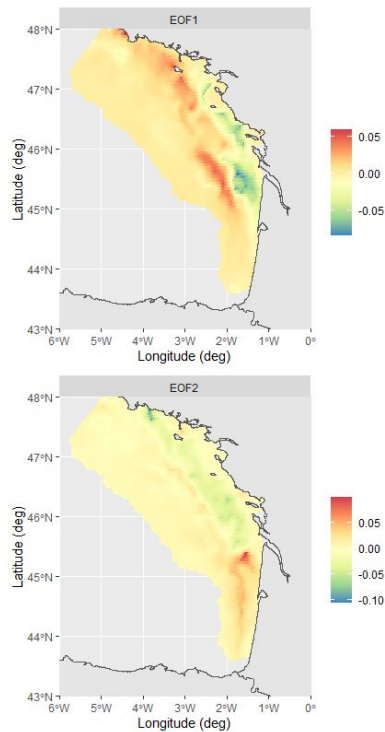


Anchois

Patron moyen







Introduction

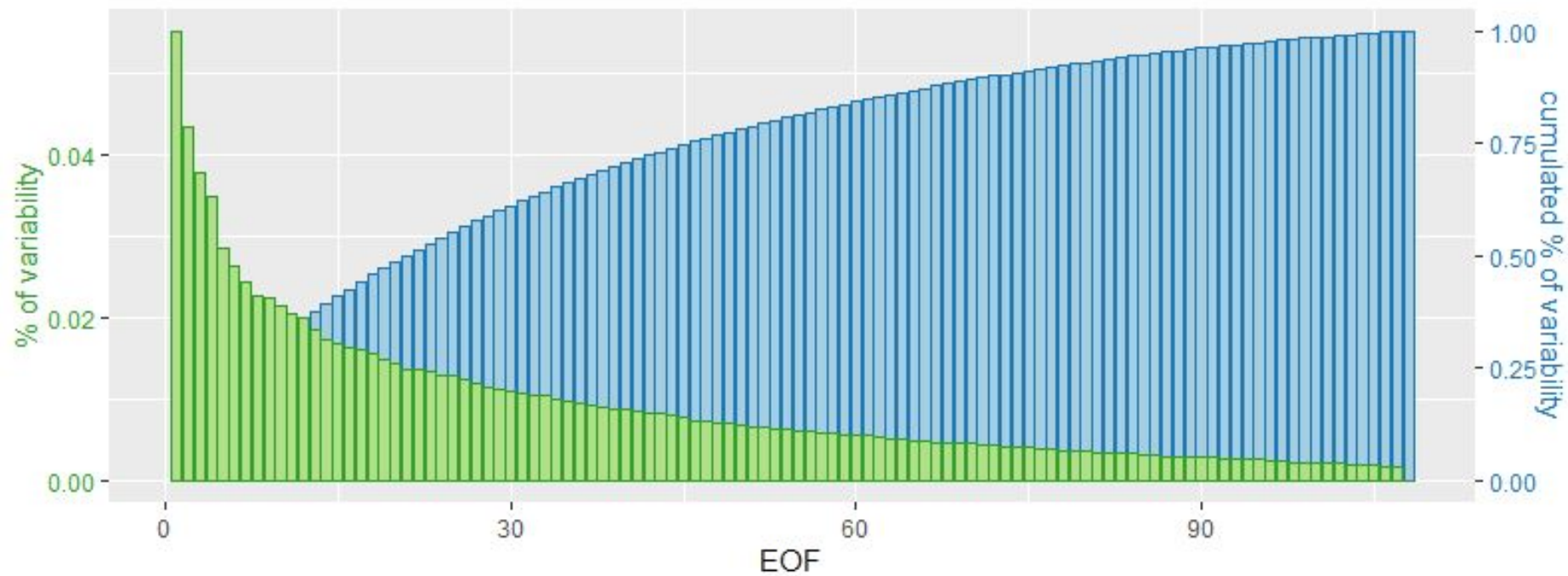
Chapitre 1
Modèle intégré

Chapitre 2
Approche ST

Autres axes de
recherches

Planning

Discussions



Stocks d'intérêt

Quelles cas d'étude choisir ?

Des stocks où la donnée est disponible :

- ⇒ stock évaluées,
- ⇒ où les captures sont majoritairement françaises,
- ⇒ où les données VMS sont disponibles (bateaux > 12 m)

Manche-Est/Mer du Nord	Golfe de Gascogne
Sole (Solea solea - Division 27.7.d (Eastern English Channel) Fr (50%) - UK (25%) - Belg. (25%)	Bay of Biscay Sole (Solea solea - sole commune - 8a, 8b) Fr
Cephalopods ???	Nephrops (Divisions 8.a,b, FU 23-24) Fr
Striped red mullet (Mullus surmuletus) - rouget barbet de roche - Subarea 4 (North Sea), divisions 7.d (Eastern English Channel) and 3.a (Skagerrak, Kattegat) Neth (½)- Fr (½) - UK - Belg	European seabass (Dicentrarchus labrax - Bar - Divisions 8.a-b (Bay of Biscay North and Central) Fr
Plaice in Division 7.d Belg. (½) - Fr (½) - UK	Sardine in divisions 8a, b, d Fr (¾)- Esp
Whiting (Merlangius merlangus - merlan - Subarea 4 (North Sea), Division 7.d (Eastern English Channel) and 3.a (Skagerrak, Kattegat) Vild : Fr (¾) Neth UK Belg.	Hake (Merluccius merluccius - merlu - Division 3.a, Subareas 4, 6 and 7 and Divisions 8.a,b,d (Northern stock) Fr (¼) - Sp - UK - Dan. - Ir.
	Anglerfish (Lophius piscatorius and Lophius budegassa - Baudroie - Sub-area 7 and Divisions and 8.a,b,d) Fr (60%) - UK - Ir - Sp
	Pollack (Pollachius pollachius - lieu jaune - Subarea 8 and Division 9.a) Fr
	Whiting (Merlangius merlangus - merlan - Subarea 8 and Division 9.a) Fr
	Herring - Subarea 8 (Bay of Biscay) - Data poor Fr
	Plaice (Pleuronectes platessa - Plie - Subarea 8 and Division 9.a) Fr (2/3) - Port.

Constrained refined Delaunay triangulation

