

# Projet requins peau bleue

Rindra, Emma, Léa

8 janvier 2021

## Contents

<b>1</b>	<b>Importations et transformation des données</b>	<b>1</b>
<b>2</b>	<b>Fonction de détection</b>	<b>1</b>
<b>3</b>	<b>Fonction de densité</b>	<b>2</b>
3.1	Choix des covariables de la fonction de densité . . . . .	2
3.2	Ajustement des modèles par session . . . . .	3
<b>4</b>	<b>Abondance : résultats et cartes</b>	<b>6</b>
4.1	Résultats numériques . . . . .	6
4.2	Carte d'abondance pour la session 2 et availability = 1 . . . . .	7

## 1 Importations et transformation des données

On récupère les données suivantes :

- `obsdata`, `segdateet predata`
- `distdata`, une jointure entre `predata` et `obsdata` sur le segment

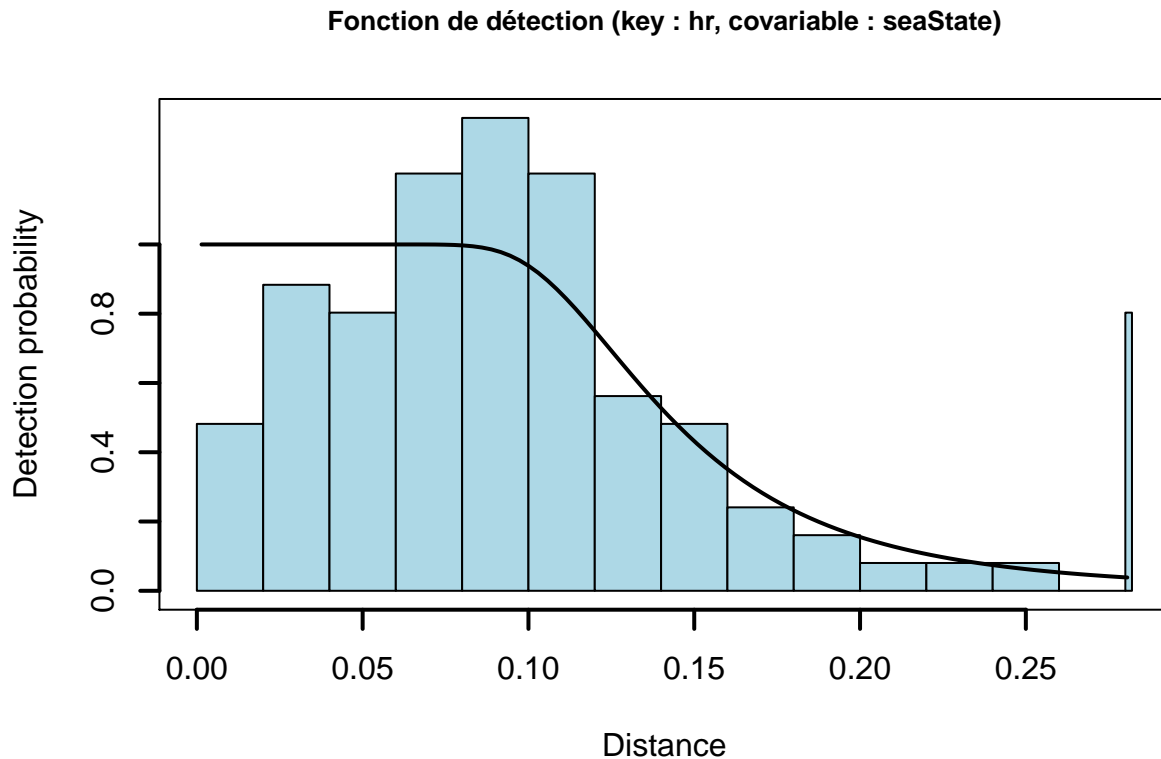
On centre et réduit les covariables (présentes dans `distdata`, `segdata` et `predata`).

## 2 Fonction de détection

Un algorithme a été écrit pour déterminer la meilleure fonction de détection. Cet algorithme teste, pour les deux distributions possibles (half-normal et hazard-rate), les différentes combinaisons de covariables possibles : aucune, le beaufort `seaState`, l'observateur `observerId` ou les deux. Pour chaque fonction de détection ajustée grâce aux données, on obtient l'AIC.

formula <chr>	key <chr>	AIC <dbl>
~1	hn	-293.4346
~seaState	hn	-290.1428
~observerId	hn	-288.9999
~observerId+seaState	hn	-288.8973
~1	hr	-297.3398
~seaState	hr	-297.3362
~observerId	hr	-295.5575
~observerId+seaState	hr	-295.8082

Pour la fonction avec le plus petit AIC (**hr** hazard-rate ; sans covariable), il y avait une erreur d'ajustement. On a donc choisi la fonction de détection avec le 2ème plus petit AIC (**hr** hazard-rate ; avec la covariable **seaState**).



### 3 Fonction de densité

#### 3.1 Choix des covariables de la fonction de densité

On crée une fonction `selec_dsm_aic_fwd` qui a permis de sélectionner de manière forward les covariables à inclure dans le modèle `dsm` sur un critère d'AIC. Cette fonction récursive ajoute à chaque étape la covariable permettant d'obtenir l'AIC le plus faible. Elle s'arrête lorsque l'ajout d'une covariable ne diminue plus l'AIC.

iter <dbl>	formule <chr>	aic <dbl>
1	count ~ s(SST_4w_mea)	1840.352
2	count ~ s(SST_4w_mea) + s(X, Y)	1761.374
3	count ~ s(SST_4w_mea) + s(X, Y) + s(CHL_4w_mea)	1756.541

C'est le modèle 3 que nous avons retenu ici, avec les covariables suivantes :

- température de la mer en moyenne les 4 semaines avant la session,
- la localisation (interaction entre X et Y),
- la teneur en chlorophylle en moyenne les 4 semaines avant la session.

Les covariables ont été choisies en ajustant les modèles sur l'ensemble des données, sans distinction de session, pour avoir le même modèle pour les 2 sessions par la suite.

### 3.2 Ajustement des modèles par session

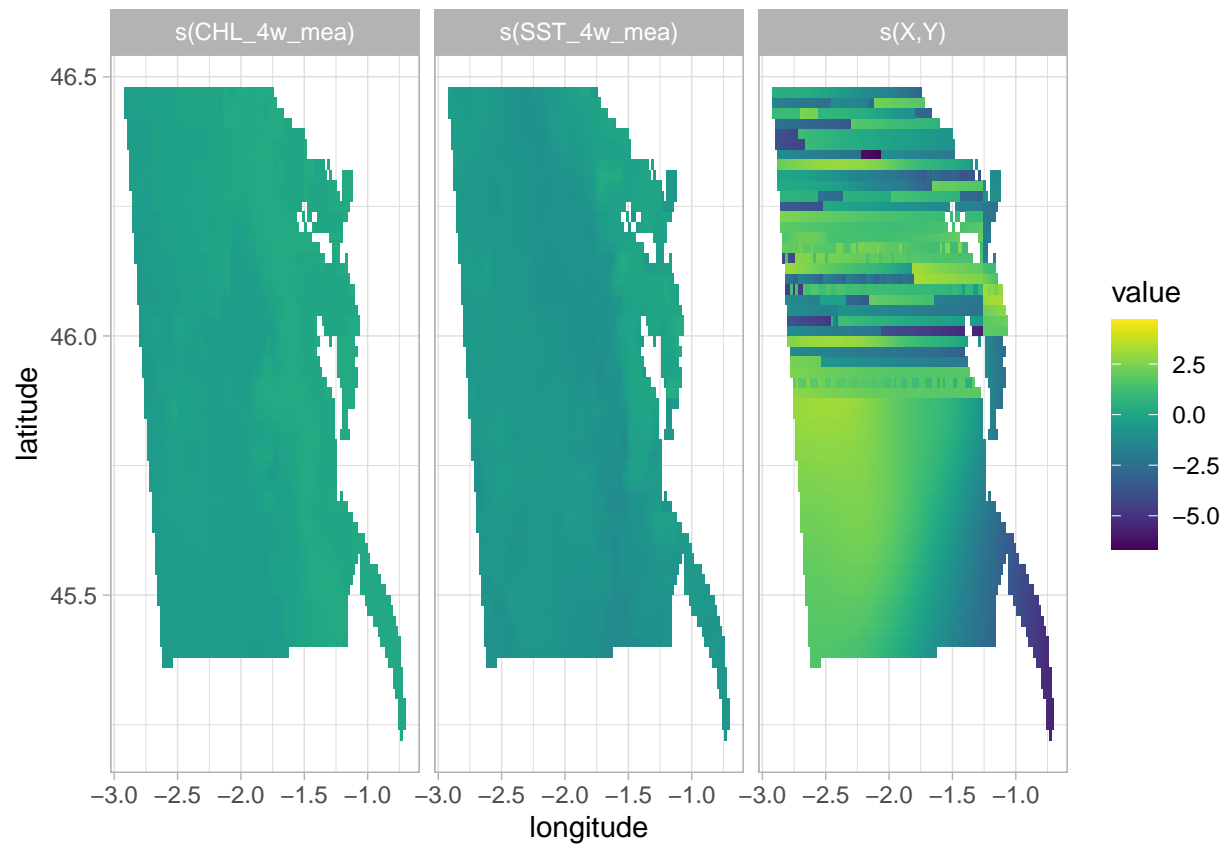
6 modèles de densité ont été ajustés grâce à la fonction `dsm` :

- Avec un biais de disponibilité constant à 1 : un modèle pour la session 2 et un pour la session 3
- Avec un biais de disponibilité constant à 0.41 : un modèle pour la session 2 et un pour la session 3
- Avec un biais de disponibilité dépendant de la profondeur : un modèle pour la session 2 et un pour la session 3

On note “on-shelf” quand la profondeur est inférieure à 150m, et “off-shelf” si la profondeur est supérieure à 150m.  $availability_{offshelf} = 0,1357617$  et  $availability_{onshelf} = 0,6332016$

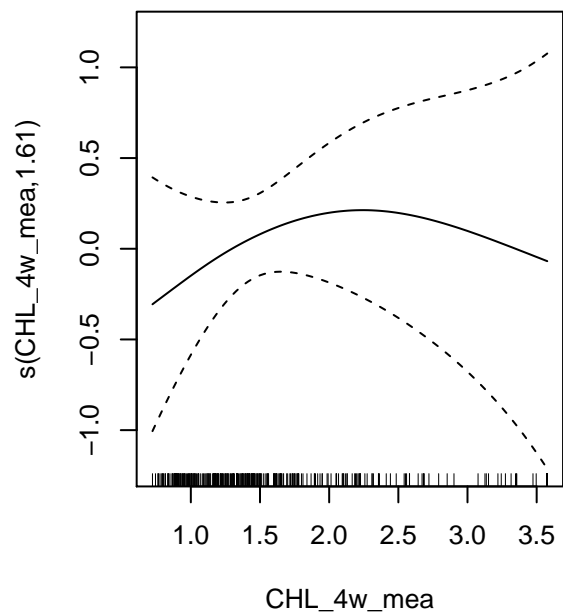
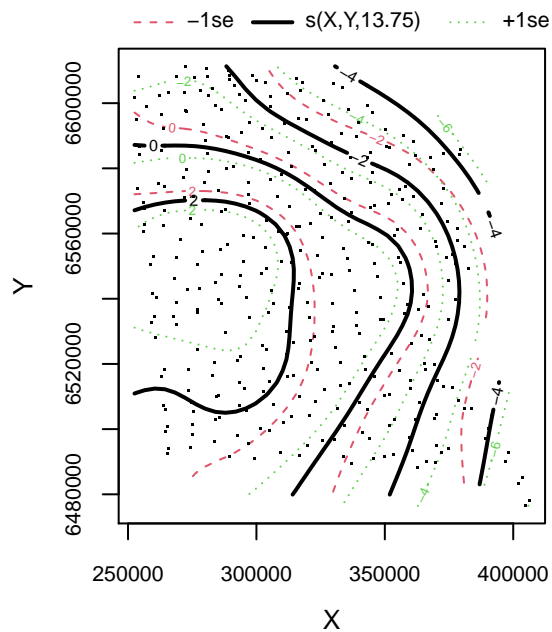
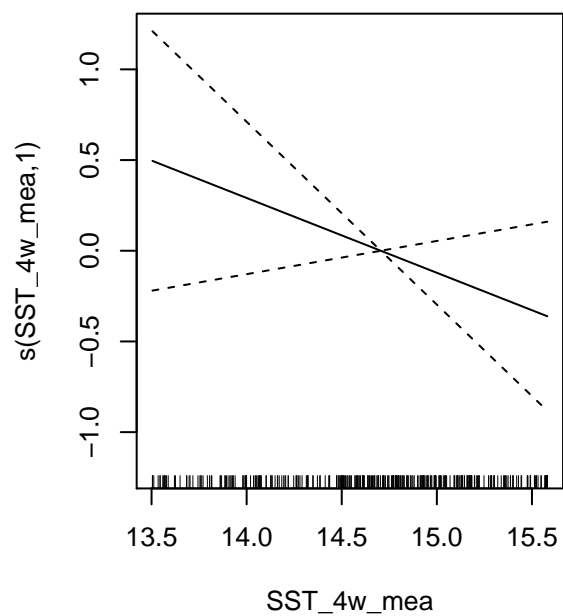
On peut visualiser le modèle GAM ajusté. Par exemple, nous allons regarder le modèle pour la session 2 avec un biais de disponibilité constant fixé à 1.

```
plot_pred_by_term(dsm_s2_av1,
                  predata_output$predata,
                  c("longitude", "latitude")) + scale_fill_viridis()
```



On remarque un manque de régularité surprenant dans la carte pour  $s(X, Y)$ .

```
par(mfrow = c(2,2))
plot.gam(dsm_s2_av1)
```



*Lignes pointillées : intervalle de confiance à 95% pour la forme moyenne de l'effet.*

## 4 Abondance : résultats et cartes

L'abondance est calculée avec la fonction `predict` (`predict.dsm` du package `dsm`).

### 4.1 Résultats numériques

```
## [1] "----- Session 2 -----"

## [1] "Avec availaibility = 1, on estime qu'il y a 51414 requins au total dans la zone."

## [1] "Avec availaibility = 0.41, on estime qu'il y a 125398 requins au total dans la zone."

## [1] "Avec availaibility selon on/off shelf, on estime qu'il y a 81194 requins au total dans la zone."

## [1] "----- Session 3 -----"

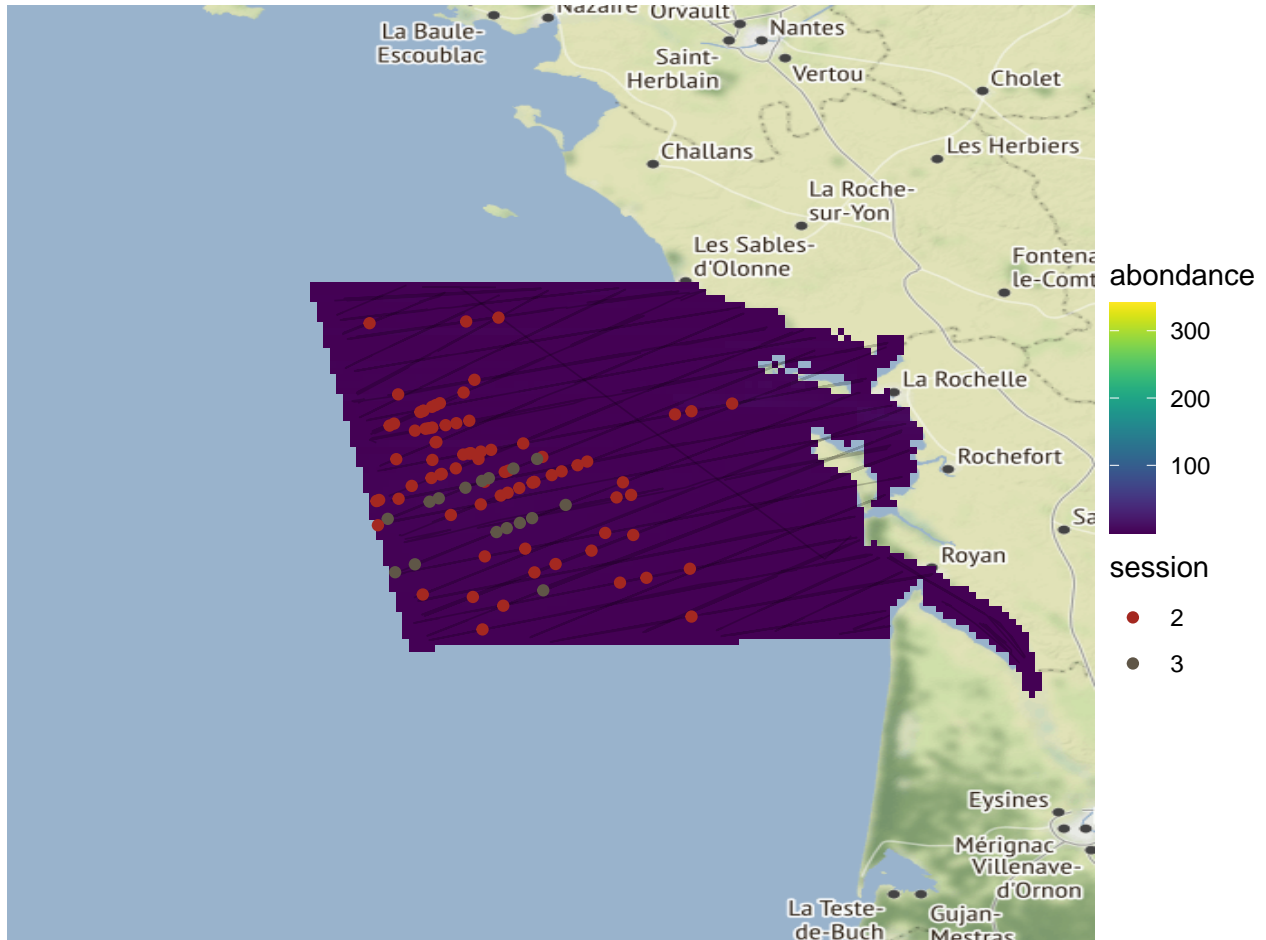
## [1] "Avec availaibility = 1, on estime qu'il y a 1609 requins au total dans la zone."

## [1] "Avec availaibility = 0.41, on estime qu'il y a 3925 requins au total dans la zone."

## [1] "Avec availaibility selon on/off shelf, on estime qu'il y a 2541 requins au total dans la zone."
```

Les résultats nous surprennent, étant donné le peu de requins détectés, surtout pour la session 2. Il semble étrange d'estimer 125000 requins (avec un biais de disponibilité de 0,41) dans 14947 km<sup>2</sup>.

#### 4.2 Carte d'abondance pour la session 2 et availability = 1



La carte est illisible : on a quelques points de la grille, d'environ 4 km<sup>2</sup>, qui contiennent jusqu'à 300 requins d'après les estimations faites ici. Le reste des points contiennent très peu de requins.

On peut regarder le nombre requins estimés dans les 25 ‘cases’ les plus peuplées :

```
## [1] 341.14018 256.39783 248.33158 203.20937 151.07222 144.57324 97.80301
## [8] 95.35479 94.38298 92.81302 92.38439 92.20136 91.07586 87.92166
## [15] 87.48004 86.82170 85.96427 84.95655 80.99995 80.47083 79.94930
## [22] 77.63538 74.00506 70.54337 70.28423
```

Un résumé des données prédites par ‘case’ :

```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
##  0.0001   0.0632   0.3980   1.9781   1.7913  341.1402
```