

Année universitaire : 2019 - 2020

Diplôme :

« Ingénieur agronome »

Spécialité :

« Sciences halieutiques et aquacoles »,  
préparée à Agrocampus Ouest

Option :

« Ressources et Ecosystèmes  
Aquatiques »

### Mémoire de fin d'études

d'ingénieur de l'École nationale supérieure des sciences agronomiques, agroalimentaires, horticoles et du paysage (AGROCAMPUS OUEST), école interne de l'institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement

de master de l'École nationale supérieure des sciences agronomiques, agroalimentaires, horticoles et du paysage (AGROCAMPUS OUEST), école interne de l'institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement

d'ingénieur de l'Institut national supérieur des sciences agronomiques, de l'alimentation et de l'environnement (AgroSup Dijon)

## Modélisation statistique de la distribution spatiale des espèces *Zostera marina* et *Zostera noltei* à l'échelle de la façade Manche-Atlantique

Par : Romain MOUILLARD



Parc naturel marin d'Iroise ©

Soutenu à Rennes, le 17/09/2020

### Devant le jury composé de :

Président : Olivier Le Pape

Maîtres de stage : Mickaël Vasquez, Martin Marzloff

Enseignant référent : Olivier Le Pape

Autres membres du jury :

Auriane Jones - Enseignant-chercheur Agrocampus Ouest

Nicolas Desroy - Chercheur à Ifremer Dinard

Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST

## Fiche de confidentialité et de diffusion du mémoire

### Confidentialité

Non  Oui si oui :  1 an  5 ans  10 ans

Pendant toute la durée de confidentialité, aucune diffusion du mémoire n'est possible <sup>(1)</sup>.

Date et signature du **maître de stage** <sup>(2)</sup> :

**A la fin de la période de confidentialité**, sa diffusion est soumise aux règles ci-dessous (droits d'auteur et autorisation de diffusion par l'enseignant à renseigner).

### Droits d'auteur

L'auteur<sup>(3)</sup>

autorise la diffusion de son travail (immédiatement ou à la fin de la période de confidentialité)

Oui  Non

Si oui, il autorise

la diffusion papier du mémoire uniquement<sup>(4)</sup>

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire (joindre dans ce cas la fiche de conformité du mémoire numérique et le contrat de diffusion)

(Facultatif)  accepte de placer son mémoire sous licence Creative commons CC-By-Nc-Nd (voir Guide du mémoire Chap 1.4 page 6)

Date et signature de l'**auteur** :

### Autorisation de diffusion par le responsable de spécialisation ou son représentant

L'enseignant juge le mémoire de qualité suffisante pour être diffusé (immédiatement ou à la fin de la période de confidentialité)

Oui  Non

Si non, seul le titre du mémoire apparaîtra dans les bases de données.

Si oui, il autorise

la diffusion papier du mémoire uniquement<sup>(4)</sup>

la diffusion papier du mémoire et la diffusion électronique du résumé

la diffusion papier et électronique du mémoire

Date et signature de l'**enseignant** :

(1) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme

(3) Auteur = étudiant qui réalise son mémoire de fin d'études

(4) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé

## **Remerciements**

Je tiens à remercier tout particulièrement Martin et Mickael pour m'avoir encadré durant ce stage de 6 mois, pour avoir partagé leur expérience et pour leur aide dans la rédaction de ce rapport. Malgré le contexte particulier et les 5 mois de télétravail, je n'ai jamais été abandonné seul dans la nature et pour ça je vous remercie énormément.

Je n'oublie évidemment pas Touria, qui par son aide dans le projet, mais également sa bonne humeur et sa gentillesse, m'a permis de réaliser mon stage dans les meilleures conditions.

Enfin, je tiens également à remercier toute l'équipe du LEBCO (Aurélien, Benjamin, Pierre-Olivier etc.) qui ont participé de près ou de loin à ce projet en discutant avec moi autour d'un café ou d'un repas. Je vous ai malheureusement peu vu à cause de la crise sanitaire, mais j'ai réellement apprécié la très bonne ambiance qui règne au sein de l'équipe.

# Table des matières

<b>I – Introduction</b> .....	1
<b>II - Matériel &amp; Méthodes</b> .....	4
1. Données de l'étude.....	4
a. Cas d'étude - les herbiers de zostère de la façade française Manche-Atlantique.....	4
b. Données environnementales.....	6
i. Présentation des variables environnementales.....	6
ii. Préparation des variables environnementales.....	8
iii. Constitution des jeux de données pour la modélisation.....	9
2. Approches de modélisation.....	11
a. Modèles prédictifs.....	11
b. Construction des modèles.....	12
c. Comparaison des performances des différentes approches de modélisation.....	13
3. Prédiction spatiales des modèles.....	14
4. Caractérisation de la niche environnementale.....	15
5. Prédiction des conséquences potentielles du changement climatique.....	16
<b>III – Résultats</b> .....	17
1. Comparaison des performances des modèles.....	17
2. Prédiction spatiales des modèles Random Forest.....	18
3. Caractérisation de la niche environnementale.....	21
4. Prédiction des conséquences potentielles du changement climatique.....	27
<b>Discussion</b> .....	30
<b>Conclusions</b> .....	35
<b>Bibliographie</b> .....	36
<b>Annexes</b> .....	42
<b>Fiche Résumé</b> .....	61

## Liste des figures

**Figure 1.** Photographies d'herbiers de *Zostera noltei* en zone médiolittorale (à gauche) et de *Zostera marina* en zone infralittorale (à droite)

**Figure 2.** Illustrations de pressions locales sur les herbiers de zostères : pêche à pied sur l'herbier de *Z. noltei* (à gauche) et mouillage des bateaux de plaisance en saison estivale sur les herbiers (à droite)

**Figure 3.** Données disponibles depuis 2000 dans le bassin d'Arcachon (a) et dans le golfe du Morbihan (b).

**Figure 4.** Corplot sur les données environnementales, avec la variable *Température annuelle de l'air* (à gauche) et sans (à droite)

**Figure 5.** Variogrammes multivariés sur les données de *Z. marina* (à gauche) et *Z. noltei* (à droite)

**Figure 6.** Performances des modèles : AUC sur données de validation (en haut) et taux d'erreur par modèle (faux positifs / faux négatifs) (en bas) pour *Z. marina* (à gauche) et *Z. noltei* (à droite)

**Figure 7.** Cartographie de l'indice d'habitat potentiel de *Z. marina* en Bretagne

**Figure 8.** Cartographie de l'écart-type de l'indice d'habitat potentiel de *Z. marina* en Bretagne

**Figure 9.** Cartographie des présences et absences de *Z. marina* en Bretagne

**Figure 10.** Cartographie des présences et absences de *Z. marina* en Bretagne avec le masque Substrat

**Figure 11.** Représentation de l'importance relative des variables sur la base du critère *MeanDecreaseAccuracy* dans les modèles Random Forest pour *Zostera marina*

**Figure 12.** Représentation de l'effet marginal des variables les plus structurantes dans les modèles de distribution sur la présence de *Zostera marina* à l'aide de partial plot

**Figure 13.** Représentation de l'interaction entre les 3 variables les plus structurantes (*Lumière*, *Fetch*, *Chlorophylle\_a*) dans les prédictions du modèle Random Forest pour *Z. marina*

**Figure 14.** Représentation de l'importance relative des variables sur la base du critère *MeanDecreaseAccuracy* dans les modèles Random Forest pour *Zostera noltei*

**Figure 15.** Représentation de l'effet marginal des variables les plus structurantes dans les modèles de distribution sur la présence de *Zostera noltei* à l'aide de partial plot

**Figure 16.** Représentation de l'interaction entre les 3 variables les plus structurantes (*Fetch*, *Lumière*, *Température estivale de l'air*) dans les prédictions du modèle Random Forest pour *Z. noltei*

**Figure 17.** Cartes de prédictions obtenues avec les modèles RF, à 11km de résolution, de la distribution potentielle actuelle de *Z. marina* (à gauche) et de *Z. noltei* (à droite)

**Figure 18.** Conséquences prédites du changement climatique (RCP4.5 et RCP 8.5) sur la distribution spatiale de l'habitat potentiel de *Z. marina* sur la façade Manche-Atlantique aux horizons 2050 et 2100.

**Figure 19.** Conséquences prédites du changement climatique (RCP4.5 et RCP 8.5) sur la distribution spatiale de l'habitat potentiel de *Z. noltei* sur la façade Manche-Atlantique aux horizons 2050 et 2100.

## Liste des tableaux

**Tableau 1.** Descriptions et sources des différents paramètres utilisés pour la modélisation.

**Tableau 2.** Description d'une matrice de confusion.

**Tableau 3.** Présentation des conditions environnementales moyennes des écorégions

**Tableau 4.** Prédictions spatiales dans les écorégions

## Liste des annexes

**Annexe 1 :** Carte de distribution du jeu de données de validation (20%) à 500 m

**Annexe 2 :** Représentation de l'importance relative des variables dans les modèles BRT

**Annexe 3 :** Représentation de l'importance relative des variables dans les modèles GLM

**Annexe 4 :** Méthodes de calcul du nombre de clusters optimal dans un partitionnement de type k-means : Indice de Calinski Harabasz (à gauche) ; méthode du coude (à droite)

**Annexe 5 :** ACP sur les clusters issus du partitionnement

**Annexe 6 :** Présentation géographique des écorégions et de leur profil environnemental exprimé en valeur relative au travers de radarplot

**Annexe 7 :** Tableaux bilan des prédictions sur le changement climatique

**Annexe 8 :** Carte de prédiction de l'indice d'habitat potentiel de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 9 :** Carte de l'écart-type de l'indice d'habitat potentiel de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 10 :** Carte de prédiction de présence/absence de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 11 :** Carte de prédiction de l'indice d'habitat potentiel de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 12 :** Carte de l'écart-type de l'indice d'habitat potentiel de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 13 :** Carte de prédiction de présence/absence de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m

**Annexe 14 :** Carte de prédiction de présence/absence de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m avec un masque *Substrat*

## I – Introduction

Les herbiers marins sont des prairies sous-marines qui poussent dans la plupart des mers du globe. Ils sont composés de plantes à fleurs (phanérogames marines) et non d'algues. Ces plantes appartiennent aux familles *Zosteraceae*, *Posidoniaceae*, *Cymodoceaceae* et *Hydrocharitaceae*. A l'instar des mangroves et récifs coralliens, ces espèces jouent le rôle d'ingénieurs écologiques (Jones *et al.*, 1997) car leur présence modifie l'environnement côtier, et notamment ses caractéristiques physico-chimiques. En effet, les herbiers marins remplissent de nombreuses fonctions écologiques qui permettent à une diversité d'espèces de se développer : importante production photosynthétique (Duarte & Chiscano, 1999), puit de carbone (Duarte *et al.*, 2004 ; Fourqurean *et al.*, 2012), modification des conditions hydro-sédimentaires (Fonseca *et al.*, 1992 ; Hansen *et al.*, 2012), provision d'un habitat complexe, zone de reproduction et nourricerie etc. (Orth *et al.*, 2006). Les herbiers sont par conséquent des espèces à fortes valeurs écologiques et patrimoniales.

En France métropolitaine, plusieurs espèces d'herbiers marins sont recensées avec la présence de posidonie (*Posidonia oceanica*) en Méditerranée et la présence de zostère marine (*Zostera marina*) et de zostère naine (*Zostera noltei*) le long des côtes Manche/Atlantique (Figure 1).



Les herbiers de zostères se développent sur les sédiments sableux et sablo vaseux intertidaux et infralittoraux des côtes. *Z. marina* se développe dans la zone infralittorale, c'est-à-dire dans la partie du littoral constamment immergée avec une luminosité compatible avec la vie des végétaux. Elle se trouve en moyenne entre 0 et 4 m de profondeurs (10 m exceptionnellement). *Z. noltei* se développe au milieu de la zone médiolittorale (zone de balancement des marées) à des taux d'émergence de 40 à 70 % en moyenne. Ainsi, les deux espèces ne se rencontrent ensemble sur l'estran qu'au niveau des limites basses de *Z. noltei* et limites hautes de *Z. marina* (Bajjouk & Hily, 2015).

**Figure 1.** Photographies d'herbiers de *Zostera noltei* en zone médiolittorale (à gauche) et de *Zostera marina* en zone infralittorale (à droite) – extrait du rapport de Bajjouk *et al.* (2019)

Cependant, les herbiers marins sont des habitats côtiers vulnérables, et ont connu un déclin important ces dernières décennies face aux pressions anthropiques croissantes qui s'exercent à la fois à l'échelle locale (e.g. dragage - Wu *et al.*, 2017) et globale (e.g. augmentation des événements extrêmes - Oliver *et al.*, 2018). En effet, se basant sur 215 sites répartis à travers le monde, Waycott *et al.* (2009) quantifient le déclin moyen des herbiers marins à hauteur de 110 km<sup>2</sup> par an depuis 1980, soit une perte totale de 29 % de la surface estimée en 1879. Le déclin s'accélère, passant de 0,9 % par an avant 1940 à 7 % par an depuis 1990. Les herbiers marins auraient donc des taux de déclin comparables à ceux des mangroves, des récifs coralliens ou encore des forêts pluviales tropicales. D'après cette étude, les principaux facteurs responsables de cette évolution seraient les maladies, le changement climatique, les espèces invasives, les aménagements côtiers, la dégradation de la qualité de l'eau, les pratiques de pêche ou encore la destruction par les ancres et les hélices des bateaux.



Le déclin des herbiers marins en France s'inscrit pleinement dans ce contexte mondiale de perte en quantité et en qualité de ces habitats marins côtiers. Le bilan mondial réalisé par Waycott *et al.* (2009) fait écho à de nombreux travaux en France (Figure 2) : impacts de maladies comme le phénomène appelé « Wasting disease » ayant décimé les herbiers dans les années 1930 (Burdick *et al.*, 1993) ; impacts de la pêche à pied dans le Golfe du Morbihan (Hily *et al.*, 2004) ; impacts des vagues de chaleurs dans le bassin d'Arcachon (Auby *et al.*, 2011) ; impacts des espèces invasives (Peirano *et al.*, 2005) ; ou encore impact de l'extraction des sédiments, des activités portuaires, de l'ostréiculture, de la mytiliculture et donc à plus vaste échelle de l'eutrophisation des eaux côtières (Bajjouk & Hily, 2015).



**Figure 2.** Illustrations de pressions locales sur les herbiers de zostères : pêche à pied sur l'herbier de *Z. noltei* (à gauche) et mouillage des bateaux de plaisance en saison estivale sur les herbiers (à droite)

Dans ce contexte de dégradation en quantité et en qualité des herbiers marins, il est critique de mieux connaître leur écologie afin d'appréhender leur vulnérabilité aux pressions anthropiques globales et locales et d'appuyer des mesures de gestion efficaces. Il est par ailleurs crucial de bien caractériser leur état écologique et leur répartition spatiale afin de détecter les changements (i.e. perte ou restauration).

Ainsi, dans la continuité des directives internationales mises en œuvre pour la protection et la conservation des habitats marins (« Directive Habitat » (92/43) ; « Directive Cadre sur l'Eau » (2000/60/CE) ; convention OSPAR ; liste rouge de l'UICN), l'Agence Française pour la Biodiversité (aujourd'hui Office Français de la Biodiversité) a lancé en 2017 le projet Life MarHa (Marine Habitats), dans lequel s'inscrit mon stage. D'une durée de 8 ans, ce projet poursuit l'objectif d'évaluer l'efficacité du réseau Natura2000 en mer et de rétablir ou maintenir un bon état de conservation des habitats marins en France, dont les herbiers de zostères et de posidonies font partie (Bajjouk *et al.*, (2019, 2020)).

Depuis leurs apparitions dans les années 1990 (notamment via la popularisation des GLMs Generalised Linear Model - McCullagh & Nelder, 1989), les modèles de distribution d'espèces sont désormais très largement utilisés pour caractériser la répartition spatiale des espèces et des habitats écologiques en fonctions de paramètres environnementaux (Bekkby *et al.*, 2008 ; Valle *et al.*, 2011 ; Downie *et al.*, 2013 ; Schubert *et al.*, 2015). Le premier avantage des modèles est qu'ils permettent de produire de la connaissance sur la distribution des habitats et espèces plus rapidement et à moindre coût comparé aux méthodes acoustiques ou de télédétection. De plus, ils contribuent à une appréhension de la niche environnementale et apportent des éléments importants dans la compréhension de la biologie et de l'écologie des espèces. Enfin, dans un contexte mondial où les habitats et espèces marines sont soumis à des multiples pressions, les modèles permettent de prédire les réponses écologiques à des scénarios futurs (pollution, changement climatique etc.).



Dans le cadre du projet Life MarHa, en collaboration avec l'équipe de recherche du laboratoire LEBCO du centre Ifremer de Brest, mon stage visait à appliquer des méthodes de modélisation statistique de distribution d'espèces (SDM dans la suite du rapport pour *Species Distribution Modelling* en anglais) pour caractériser la distribution spatiale des espèces *Zostera marina* et *Zostera noltei* à l'échelle de la façade Manche-Atlantique, mieux comprendre leur écologie et produire des cartographies de référence en appui à la gestion. En effet, les observations directes disponibles concernant la répartition des herbiers de zostères sont peu fréquentes, limitées dans le temps et restreintes géographiquement à certaines zones du littoral français donc ne permettent pas d'appréhender les changements passés et en cours à l'échelle de toute la façade côtière Manche-Atlantique.

Dans un premier temps, j'ai comparé différentes approches afin d'identifier une méthode robuste de modélisation de la distribution spatiale potentielle des deux espèces de zostères (i.e. naines et marines) à l'échelle de la façade Manche-Atlantique. Les modèles retenus m'ont ainsi permis de décrire les caractéristiques communes et spécifiques des niches environnementales de chacune des deux espèces. Dans un second temps, j'ai évalué via les prédictions des modèles les impacts potentiels de différents scénarios de changement climatique sur l'habitat potentiel de ces deux espèces.

## **II - Matériel & Méthodes**

Le travail de modélisation mis en place lors de mon stage afin d'appréhender la distribution spatiale potentielle des deux espèces *Z. marina* et *Z. noltei* à l'échelle de la façade Manche-Atlantique peut se résumer en 5 étapes majeures : (1) la constitution de jeux de données robustes, (2) l'évaluation de la performance d'approches alternatives (GLM, RF, BRT) pour modéliser l'habitat potentiel des herbiers, (3) la caractérisation de la répartition géographique des deux espèces, ainsi que (4) de leur niche environnementale. (5) La dernière étape visait à prédire les effets de scénarios de changement climatique sur l'habitat potentiel des herbiers.

### **1. Données de l'étude**

#### **a. Cas d'étude - les herbiers de zostère de la façade française Manche-Atlantique**

Afin de consolider un jeu de données concernant l'occurrence des deux espèces d'intérêt, j'ai tout d'abord recensé les données géoréférencées existantes concernant les herbiers de zostères.

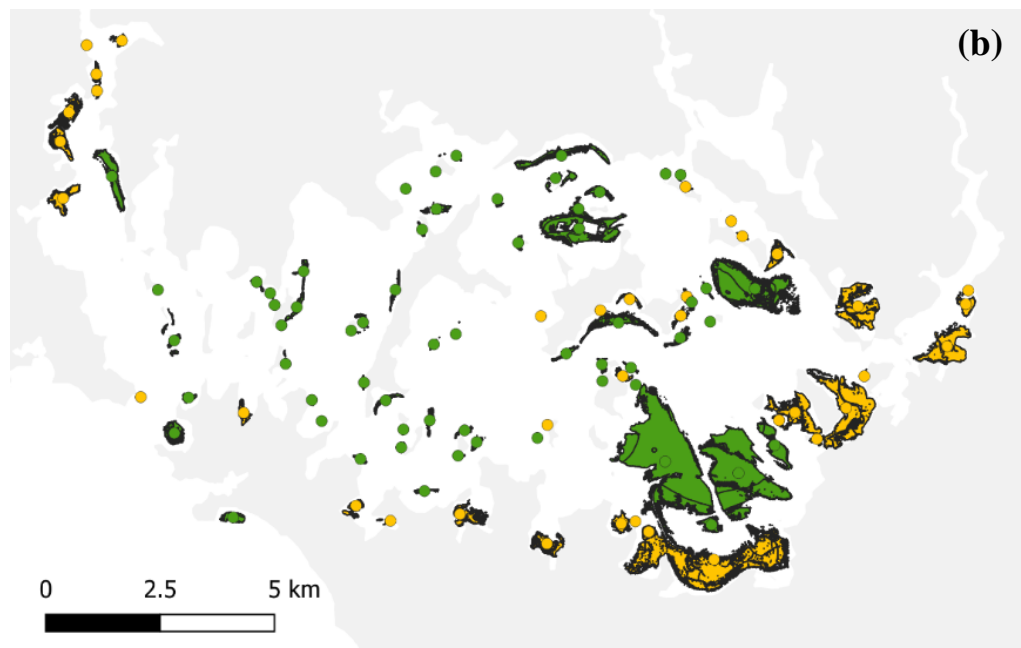
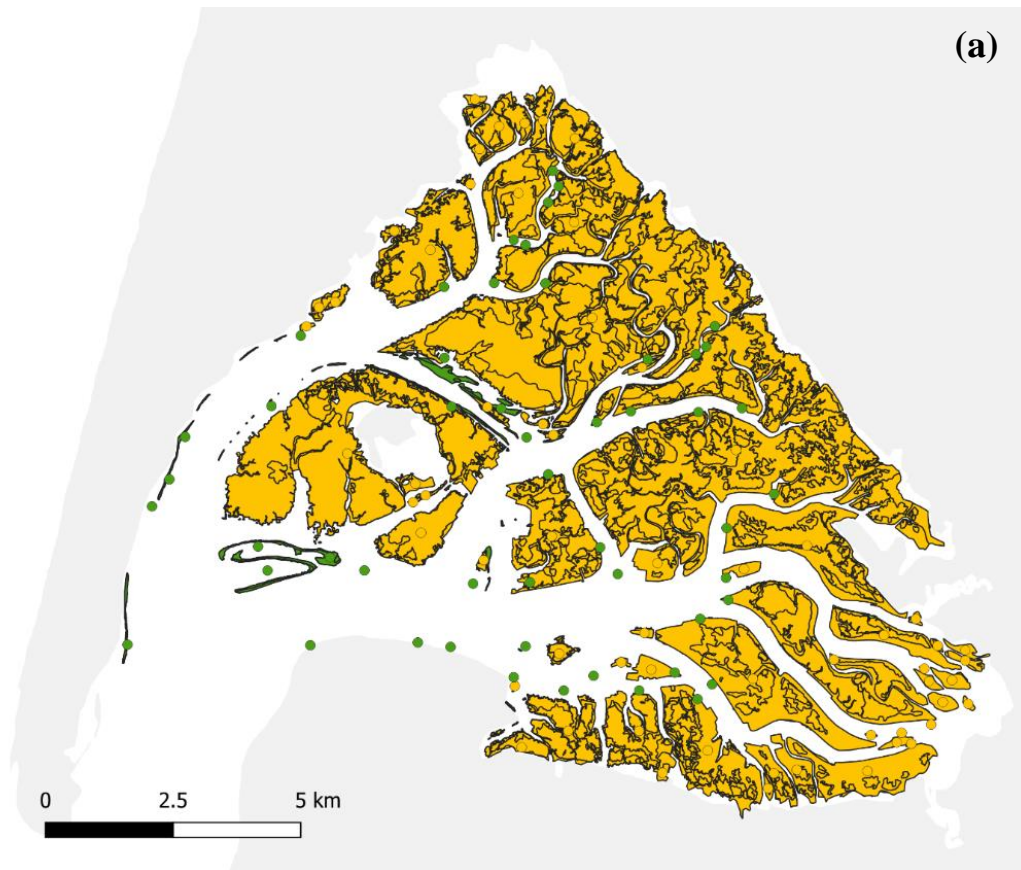
Les données recensées (Bajjouk *et al.*, 2020) concernant la répartition des herbiers proviennent d'études locales diverses réalisées par différents organismes (Ifremer, Office Français de la Biodiversité, Parcs naturels marin etc.) dans le cadre de divers suivis écologiques (Réseau Natura 2000, réseau OSPAR, réseau CARTHAM, réseau REBENT etc.). L'ensemble des relevés, à la fois surfaciques et ponctuels, couvrent une période allant de 1960 à 2016, et reposent sur des méthodes de prospection variées (par exemples : photos aériennes, plongées etc.). La figure 3 illustre par exemple les données acquises localement au niveau du bassin d'Arcachon et Golfe du Morbihan.

Afin de modéliser l'habitat potentiel actuel des herbiers de zostères, j'ai uniquement considéré les relevés récents (post-2000) pour lesquels les métadonnées étaient complètes, notamment en termes de différenciation des deux espèces d'intérêt *Zostera marina* et *Zostera noltei*.

Mon stage ne vise pas à étudier la biomasse ou l'étendue d'un herbier local, mais à caractériser l'habitat potentiel des deux espèces à l'échelle de la façade française Manche-Atlantique. Aussi, j'ai discrétisé les données surfaciques (i.e. polygones) disponibles afin de constituer un jeu de données d'occurrence correspondant à des présences ponctuelles à l'échelle de la façade Atlantique métropolitaine. J'ai réalisé ce travail de discrétisation des données d'occurrence à deux résolutions horizontales différentes, à 500 m ou à 1000m respectivement, afin de tester l'influence du nombre d'observations et de la résolution spatiale sur la performance des modèles SDM. Ainsi, j'ai obtenu un total de 2079 points de présence à 500m de résolution (ou 926 à 1000m) pour *Zostera marina* ; et, un total de 1336 points de présence à 500m de résolution (ou 521 à 1000m) pour *Zostera noltei*.

Si des connaissances expertes pourraient informer localement les zones d'absence des deux espèces (par exemple, Christian Hily à l'échelle de la Bretagne ou Isabelle Auby à l'échelle du bassin d'Arcachon), la localisation de vraies absences n'était pas envisageable de manière robuste et homogène à l'échelle de la façade Manche-Atlantique.

Ainsi, pour chaque espèce, j'ai constitué un jeu de données de pseudo-absences afin de modéliser leur habitat potentiel. Spécifiquement, j'ai tiré aléatoirement autant de points de pseudo-absence que de points de présence recensés pour chaque espèce à une résolution donnée. L'échantillonnage aléatoire des pseudo-absences ciblait les pixels non associés à des présences en zone infralittorale (lumière  $\geq 0.7$  mol.phot.m<sup>-2</sup>.j<sup>-1</sup> - Populus *et al.*, 2017).



- Points d'observations de *Z. noltei*
- Points d'observations de *Z. marina*
- Polygones de *Z. noltei*
- Polygones de *Z. marina*

**Figure 3.** Données disponibles depuis 2000 dans le bassin d'Arcachon (a) et dans le golfe du Morbihan (b).

## b. Données environnementales

### i. Présentation des variables environnementales

En complément des données d'occurrence, j'ai compilé les informations de différentes sources (e.g. satellites, modèles hydrodynamiques...) afin de caractériser l'environnement côtier où se développent les herbiers de zostères. Ces paramètres présentés dans le Tableau 1 peuvent se résumer selon 3 catégories : propriétés physico-chimiques des masses d'eaux, conditions hydrodynamiques (et atmosphériques pour l'espèce intertidale *Z. noltei*), ainsi que les caractéristiques sédimento-morphologiques des fonds marins.

**Tableau 1.** Descriptions et sources des différents paramètres utilisés pour la modélisation.

Variable	Résolution horizontale	Résolution temporelle	Couverture temporelle	Description / Sources
<i>Données utilisées pour les modèles de distribution de l'habitat potentiel actuel</i>				
Bathymétrie	100 m	--	--	<i>Exprimé en m</i> <i>Source : EMODnet Bathymetry (2018)</i>
Kd <sub>PAR</sub>	250 m	Journalière	2005 – 2009	<i>Kd<sub>PAR</sub> ("diffuse attenuation coefficient of downwelling photosynthetically available radiation") est un coefficient de pénétration de la lumière dans la colonne d'eau. Il est estimé à partir de la radiançe et mesuré par satellite. Il s'exprime en m<sup>-1</sup>.</i> <i>Source : EMODnet Seabed Habitats(Vasquez et al.,2020)</i>
PAR	4000 m	Journalière	2005 – 2009	<i>Le rayonnement photosynthétiquement actif (photosynthetically active radiation - PAR) est un rayonnement dont les longueurs d'onde s'étendent de 400 à 700 nm que les organismes photosynthétiques peuvent utiliser dans le processus de photosynthèse. Il s'agit du rayonnement reçu à la surface de l'eau. Il s'exprime en mol.phot.m<sup>-2</sup>.j<sup>-1</sup>.</i> <i>Source : EMODnet Seabed Habitats(Vasquez et al.,2020)</i>  <i>Dans notre étude, nous avons utilisé un proxy de la lumière disponible au fond (PAR<sub>fond</sub>) obtenu à partir de la bathymétrie, du Kd<sub>PAR</sub> et du PAR, et calculé avec la formule :</i> $PAR_{fond} = PAR \times e^{-bathymétrie \times KdPAR}$
Courant – Energie cinétique au fond	500 m	Horaire	2010-2015	<i>Exprimé en N.m<sup>-2</sup></i> <i>Source : MARS3D-AGRIF - Zone Manche-Gascogne (Caillaud et al., 2016)</i>
Vagues – Energie cinétique au fond	300 m	Horaire	2010-2015	<i>Exprimé en N.m<sup>-2</sup></i> <i>Source : Homère (Boudiere et al., 2013)</i>
Fetch	100 m	--	--	<i>Burrows (2012) considère le fetch comme un indice d'exposition aux vagues. La valeur de cet indice en un point est calculée comme l'exposition totale de ce point à des vagues pouvant arriver à 360° autour de ce point. A la différence de la variable Vagues – Energie cinétique au fond, la variable Fetch prend en compte l'orientation des vagues relativement à l'exposition du trait de côte.</i>  <i>Dans notre étude, nous avons considéré ce paramètre comme un proxy de la morphologie du littoral et de l'hydrodynamisme, en complément de l'information apportée par la variable Vague.</i> <i>Source : Non publié (Burrows, 2012)</i>

Substrat - Nature géologique des fonds	Multi-échelle	--	2020	<p><i>Classification de Folk en 7 classes : 1 - Roche, 2 - sédiments grossiers, 3 - sédiments mixtes, 4 - limon (limon <math>\geq 90\%</math>, sable <math>&lt; 10\%</math>, graviers <math>&lt; 5\%</math>), 5 - limon sablonneux (limon 50-90%, sable 10-50%, graviers <math>&lt; 5\%</math>), 6 - Sable (limon <math>&lt; 10\%</math>, sable <math>\geq 90\%</math>, graviers <math>&lt; 5\%</math>), 7 - Sable limoneux (limon 10-50%, sable 50-90%, graviers <math>&lt; 5\%</math>)</i></p> <p><i>Source : EMODnet Geology</i></p>
Température de l'eau	500 m	Horaire	2010-2015	<p><i>Exprimé en °C</i></p> <p><i>Source : MARS3D-AGRIF - Zone Manche-Gascogne (Caillaud et al., 2016)</i></p>
Salinité	500 m	Horaire	2010-2015	<p><i>Exprimé en PSU</i></p> <p><i>Source : MARS3D-AGRIF - Zone Manche-Gascogne (Caillaud et al., 2016)</i></p>
Température annuelle de l'air	1000 m	Horaire	2012-2018	<p><i>Exprimé en K</i></p> <p><i>Source : Arome – Météo France (Seity et al., 2011)</i></p>
Température estivale de l'air	1000 m	Horaire	2012-2018	<p><i>Exprimé en K</i></p> <p><i>Source : Arome – Météo France (Seity et al., 2011)</i></p>
Concentration en oxygène dissous	4000 m	Journalière	2012-2018	<p><i>Exprimé en mg.L<sup>-1</sup></i></p> <p><i>Source : MARC (Arduin, 2013)</i></p>
Concentration en chlorophylle-a	4000 m	Journalière	2012-2017	<p><i>Exprimé en µg.L<sup>-1</sup></i></p> <p><i>Source : MARC (Arduin, 2013)</i></p>
Concentration en phosphate	4000 m	Journalière	2012-2018	<p><i>Exprimé en µmol.L<sup>-1</sup></i></p> <p><i>Source : MARC (Arduin, 2013)</i></p>
Concentration en nitrate	4000 m	Journalière	2012-2018	<p><i>Exprimé en µmol.L<sup>-1</sup></i></p> <p><i>Source : MARC (Arduin, 2013)</i></p>
<i>Données utilisées pour l'évaluation de la résilience des espèces face au changement climatique</i>				
Vagues – hauteur significative	12000 m	Mensuelle	1970-2004 2040-2050 (RCP4.0 / 8.5) 2090-2100 (RCP4.0 / 8.5)	<p><i>Exprimé en m</i></p> <p><i>Source : NOC (Bricheno et al., 2018)</i></p>
Courant – Energie cinétique au fond	11000 m	Mensuelle	2010-2020 2040-2050 (RCP4.0 / 8.5) 2090-2100 (RCP4.0 / 8.5)	<p><i>Cette variable est la résultante de deux composantes : la vitesse de surface horizontale vers l'Est (u) et la vitesse de surface horizontale vers le Nord.</i></p> <p><i>Elle a été obtenue avec la formule :</i></p> <p><i>Courant = <math>\sqrt{u^2 + v^2}</math></i></p> <p><i>Exprimé en m.s<sup>-1</sup></i></p> <p><i>Source : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)</i></p>
Température de l'eau	11000 m	Mensuelle	2010-2020 2040-2050 (RCP4.0 / 8.5) 2090-2100 (RCP4.0 / 8.5)	<p><i>Exprimé en °C</i></p> <p><i>Source : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)</i></p>
Salinité	11000 m	Mensuelle	2010-2020 2040-2050 (RCP4.0 / 8.5) 2090-2100 (RCP4.0 / 8.5)	<p><i>Exprimé en PSU</i></p> <p><i>Source : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)</i></p>

Température estivale de l'air	5 minutes	Mensuelle	1970-2000 2041-2060 2081-2100 (RCP4.0 / 8.5)	Sur les 19 indicateurs, nous avons analysé l'indicateur « BIO10 = Mean Temperature of Warmest Quarter » Exprimé en °C <i>Source</i> : Eyring et al., (2016) ; Fick et al., (2017)
Concentration en oxygène dissous	11000 m	Mensuelle	2010-2020 2040-2050 2090-2100 (RCP4.0 / 8.5)	Exprimé en mol.m <sup>-3</sup> <i>Source</i> : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)
Concentration en chlorophylle-a	11000 m	Mensuelle	2010-2020 2040-2050 2090-2100 (RCP4.0 / 8.5)	Exprimé en kg.m <sup>-3</sup> <i>Source</i> : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)
Concentration en phosphate	11000 m	Mensuelle	2010-2020 2040-2050 2090-2100 (RCP4.0 / 8.5)	Exprimé en mol.m <sup>-3</sup> <i>Source</i> : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)
Concentration en nitrate et nitrite	11000 m	Mensuelle	2010-2020 2040-2050 2090-2100 (RCP4.0 / 8.5)	Exprimé en mol.m <sup>-3</sup> <i>Source</i> : POLCOMS-ERSEM (Copernicus Climate Change Service, 2020)

\* Dans cette étude, le nom des variables a été simplifié : PAR = Lumière ; Vagues – Energie cinétique au fond = Vagues ; Courant – Energie cinétique au fond = Courant ; Température de l'eau = Température ; Température estivale de l'air = Tair\_ete ; Concentration en Chlorophylle\_a = Chlo\_a ; Concentration en oxygène dissous = Oxygène ; Concentration en phosphate = Phosphate ; Concentration en nitrate = Nitrate.

Afin de caractériser l'habitat potentiel des herbiers de zostères, mon travail de modélisation repose sur des climatologies récentes des différentes variables dynamiques (e.g. température, courant etc...), et sur les données les plus récentes concernant les variables statiques (i.e. substrat), comme détaillée dans le tableau ci-dessus. J'ai utilisé des moyennes temporelles pour la plupart des variables. Cependant, afin d'appréhender le rôle des événements extrêmes d'un point de vue climatologique, j'ai parfois utilisé des percentiles (par exemple le percentile 90 pour la variable vague).

## ii. Préparation des variables environnementales

La préparation des données environnementales a été réalisée en 3 étapes :

- Etape 1 : Définition de la zone infralittorale.

L'étendue géographique de mon étude a été restreinte à la zone infralittorale en utilisant la variable *Lumière* comme proxy. En s'appuyant sur les travaux de Populus et al. (2017), on considère qu'un pixel est dans la zone infralittorale si la valeur de la variable *Lumière* dans ce pixel est supérieure à 0.7 mol.photons m<sup>-2</sup>.j<sup>-1</sup>.

- Etape 2 : Homogénéisation de la résolution horizontale des variables

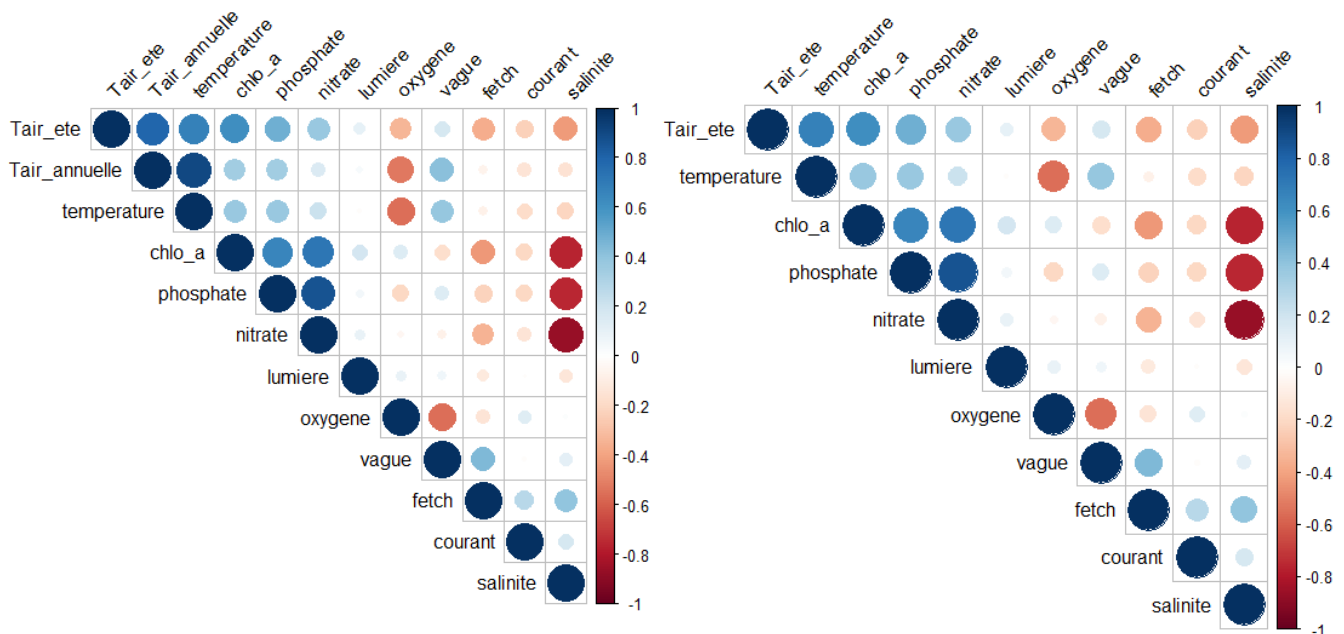
Comme il est possible de le constater dans le tableau 1, les variables environnementales n'ont pas été construites avec les mêmes résolutions. Nous avons eu recours à la fonction *resample* et à la méthode des plus proches voisins pour reprojeter toutes les variables environnementales à la même résolution (500 m ou 1000 m).



- Etape 3 : Correction de la multi-colinéarité dans le jeu de données environnementales

La colinéarité entre variables explicatives peut être source de biais importants dans la modélisation, et notamment mener à la surestimation de certains coefficients de régression, à inverser le signe des coefficients de régressions (et donc à des conclusions écologiques contre-intuitives ou erronées), ou encore à surestimer la variance des paramètres (Foucart, 2006). Afin de réduire la colinéarité des prédicteurs, j'ai utilisé le facteur d'inflation de la variance (VIF pour *Variance Inflation Factor*) pour quantifier la colinéarité. Le VIF propose une quantification de la multi-colinéarité entre variables explicatives en s'appuyant sur des régressions qui vise à expliquer la variables explicatives  $X_i$  en fonction de toutes les autres  $X_j$ .

Sur la base des conseils de Guisan *et al.* (2017), j'ai opéré à une sélection de variables afin de maintenir des scores VIF < 10. Seule la variable *Température annuelle de l'air* avait une valeur de VIF > 10 (VIF = 12.40) de part des corrélations fortes à trois autres variables : *Température de l'eau*, *Vagues* et *Concentration en oxygène dissous*. Au vu de l'importance écologique pour les herbiers des variables *Température de l'eau* (Van Katwijk *et al.*, 1997 ; Greve *et al.*, 2003) et *Vagues* (La Nafie *et al.*, 2012), j'ai éliminé la variable *Température annuelle de l'air* du jeu de données final, et conservé les 12 autres variables explicatives.



**Figure 4.** Corrplot sur les données environnementales, avec la variable *Température annuelle de l'air* (à gauche) et sans (à droite)

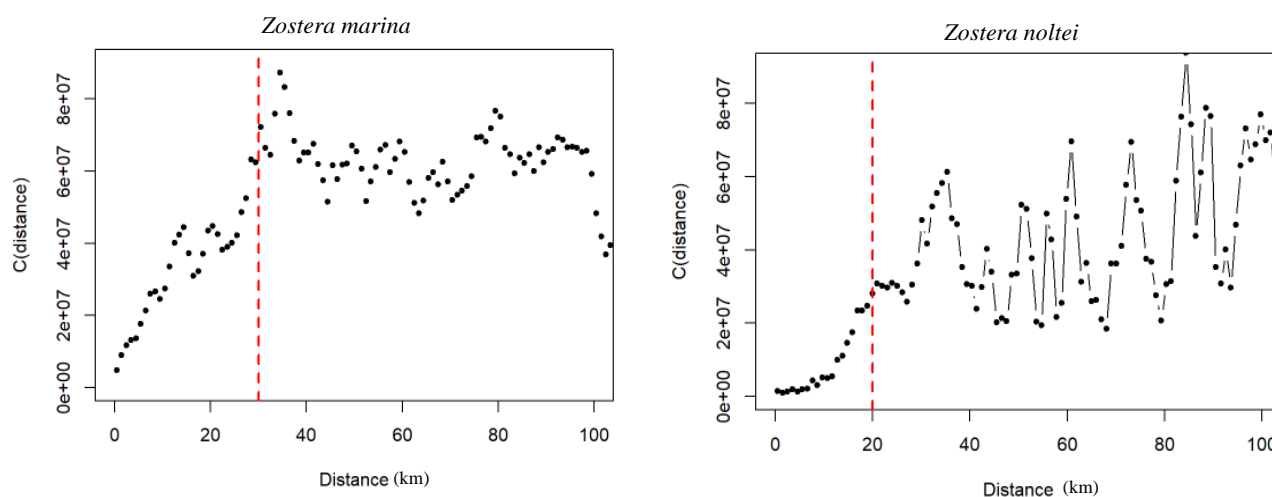
### iii. Constitution des jeux de données pour la modélisation

J'ai ensuite procédé à la compilation des données d'occurrence avec les données environnementales, en associant à chaque point de présence/absence une combinaison de variables environnementales. Certaines variables mal résolues à la côte et en particulier pour la zone intertidale, n'ont pas de valeurs définies à la localisation précise de certains points de présence/absence. Ainsi, afin de surmonter ce problème, la valeur de ces variables correspond parfois à celle du pixel le plus proche.

Le cas de *Z. noltei* est particulier puisque les variables océanographiques sont majoritairement définies en dessous du zéro hydrographique et non sur la zone d'estran. Afin d'appréhender la distribution biogéographique de l'habitat potentiel de *Z. noltei*, j'ai donc utilisé la valeur des variables environnementales dans le pixel situé en zone infralittorale le plus proche.

Avant d'entreprendre le travail de modélisation, une dernière étape préliminaire visait à quantifier l'autocorrélation spatiale au sein de chaque jeu de données environnementales (à différentes résolutions et pour chaque espèce). Anselin et Bera (1998) ont proposé une définition de l'autocorrélation spatiale comme l'absence d'indépendance entre des observations géographiquement proches. Ainsi, la présence d'autocorrélation spatiale dans les données équivaut à la non-indépendance des observations, ce qui peut biaiser les résultats de modèles statistiques. Le test de l'autocorrélation spatiale est donc une précaution indispensable à toute étude empirique mobilisant des données géographiques. Pour étudier cela, j'ai utilisé la fonction *variogmultiv* du package *adespatial* pour réaliser un variogramme multivarié prenant en compte l'ensemble des variables environnementales (Figure 5).

Un moyen de réduire l'autocorrélation spatiale en amont de la modélisation consiste à réduire le jeu de données en éliminant les points géographiquement proches. Par exemple, on peut, étant donné la position du plateau sur la Figure 5 (panel de gauche pour *Z. marina*, et droite pour *Z. noltei*), fortement réduire l'autocorrélation spatiale des données en ne sélectionnant qu'un point de présence de *Z. marina* tous les 30 km (resp. 20 km pour *Z. noltei*). Un tel filtrage spatial réduit de manière importante le nombre d'observations à 54 points pour *Z. marina* à 1000 m de résolution (resp. 44 points pour *Z. noltei*), et 88 points pour *Z. marina* à 500 m de résolution (resp. 70 points pour *Z. noltei*).



**Figure 5.** Variogrammes multivariés sur les données de *Z. marina* (à gauche) et *Z. noltei* (à droite)

Étant donné le compromis entre réduction de l'autocorrélation spatiale et diminution de la taille du jeu de données, j'ai également construit un jeu de données intermédiaire entre le jeu de données original et le jeu de données très réduit corrigeant l'autocorrélation spatiale. Ainsi, afin de garder un nombre d'occurrences suffisantes pour permettre le sous-échantillonnage, et notamment le recours à de la validation croisée, j'ai appliqué un filtre spatial à 5 km qui a abouti à un ensemble de 207 points d'observations à 1000 m de résolution pour *Z. marina* (resp. 111 points pour *Z. noltei*), et 393 points pour *Z. marina* à 500 m de résolution (resp. 204 points pour *Z. noltei*).

Afin d'évaluer la sensibilité des résultats aux choix réalisés dans la préparation des données, j'ai ainsi construit des jeux de données alternatifs selon les critères suivants : avec ou sans certaines variables (e.g. avec ou sans les variables *Concentration en oxygène dissous*, *Concentration en nitrate*, *Concentration en phosphate* et *Concentration en chlorophylle-a* qui sont les variables les moins bien définies spatialement, et avec la plus grande résolution), à deux résolutions horizontales (500 ou 1000 m) et selon les deux stratégies de correction de l'autocorrélation spatiale présentées ci-dessus (filtre à 30 km pour *Zostera marina* (resp. 20 km pour *Zostera noltei*) ou filtre à 5 km).

La variable *Substrat* est une variable structurante incontestable de l'habitat des herbiers de zostères, puisque les herbiers de zostères ne se développent pas sur la roche mais uniquement sur les substrats meubles. Des études montrent également l'importance de la nature du substrat meuble en opposant notamment les fonds sableux et les fonds limoneux (Kenworthy & Fonseca, 1976 ; Van Katwijk *et al.*, 1997). Nous aurions souhaité regarder si la distribution des herbiers dépend de la nature du substrat meuble, et nous avons pour cela envisagé d'utiliser la source de données d'EMODnet Geology sur les substrats, qui est une compilation des meilleures sources de données en Europe, et qui discrimine 7 types de substrats (sable, sable vaseux, vase sableuse, vase, sédiment mixte, substrat grossier et roche). Cependant, nous n'avons finalement pas pris en compte cette variable dans nos modèles car l'information est de qualité inégale spatialement. Il est par exemple connu qu'elle est mauvaise en Bretagne Nord. Utiliser cette variable aurait donc biaisé nos modèles.

## **2. Approches de modélisation**

### **a. Modèles prédictifs**

Dans le cadre du volet 'habitat côtier' du projet Life MarHa, le Work Package « Modélisation » vise à combiner différentes approches de modélisation, et notamment des approches déterministes (par exemple en couplant des modèles de croissance des herbiers à des modèles hydro-sédimentaires) et statistiques. Dans mon étude, j'ai uniquement mis en œuvre des modèles statistiques de type SDM.

Les méthodes de modélisation statistique sont aujourd'hui très nombreuses. On peut citer par exemple les modèles linéaires gaussien classique (régression linéaire, analyse de variance, analyse de covariance), les modèles linéaires généralisés (GLM), les modèles non linéaires (régression non paramétrique, modèles additifs généralisés GAM, les réseaux de neurones etc.), les modèles mixtes, les modèles par arbre binaire de régression et de classification etc. Dans cette étude, trois méthodes de modélisation statistique ont été utilisées : les modèles linéaires généralisés (GLM), les Random Forest (RF) et les modèles Boosted Regression Trees (BRT).

Le modèle linéaire généralisé (GLM) (Nelder & Wedderburn, 1972) est une généralisation de la régression linéaire. Le GLM a trois caractéristiques principales : (1) la distribution des erreurs de la variable réponse n'est plus nécessairement gaussienne (comme dans le cas des modèles linéaires) mais rattachée à une loi de la famille exponentielle (Normale, Poisson, Binomiale, Gamma) ; (2) la fonction de lien permet de modéliser des relations non-linéaires entre variable réponse et variables explicatives ; (3) enfin, l'estimation du maximum de vraisemblance du modèle repose sur la méthode des moindres carrés. Dans cette étude, la probabilité d'occurrence suit une loi binomiale dans un contexte de présence-absence, et la fonction de lien est la fonction *logit*.

Le modèle Random Forest (RF) (Breiman, 2001) est un cas particulier de bagging (*bootstrap aggregating*) appliqué aux arbres de décision de type CART (*Classification And Regression Trees*). L'idée de cette méthode est de construire une forêt de N arbres de classification (500 dans cette étude). Chaque arbre est construit sur un jeu d'apprentissage qui sous-échantillonne les individus via un tirage avec remise (i.e. méthode de *bootstrapping*) et à chaque nœud de l'arbre les prédicteurs potentiels sont issus d'un tirage aléatoire parmi l'ensemble des variables explicatives. La force des forêts aléatoires (ou *Random Forest* en anglais) est d'agrèger les prévisions des N modèles entraînés sur des sous-échantillons des données à prédire et des variables explicatives (comme décrit ci-dessus). Dans le contexte d'une classification binaire de type présence/absence, deux types de prédictions peuvent être exploitée, soit la proportion de présence prédite sur les N arbres (i.e. une estimation de la probabilité de présence), soit le mode dominant (i.e. une estimation de la réponse binaire – présence ou absence).

Le modèle Boosted Regression Trees (BRT) (Friedman, 2002) combine la méthode des arbres de régression et une technique d'amplification (*boosting*) pour améliorer la performance prédictive du modèle d'ensemble en dédiant chaque nouveau modèle à l'explication des résidus (i.e. de la variance non-expliquée). Ainsi, le modèle BRT est une technique d'apprentissage supervisé dont le but est de créer à partir d'arbres de décision un modèle d'ensemble prédisant la valeur d'une variable-cible à partir de la valeur de plusieurs variables d'entrée.

Notre intérêt s'est porté sur ces trois méthodes pour deux raisons. La première est la simplicité d'exécution, de compréhension et d'interprétation de ces modèles. La deuxième raison est la diversité de ces méthodes. Un des objectifs initiaux de mon étude est de construire une approche robuste de modélisation statistique, donc il était important d'évaluer l'influence du modèle utilisé sur les résultats, et de comparer forces et limites de chaque approche. J'ai donc réalisé mon analyse à l'aide de ces trois méthodes afin de sélectionner par la suite la plus performante.

## **b. Construction des modèles**

Les observations disponibles correspondent uniquement à des présences donc l'entraînement des modèles d'habitat potentiel repose sur l'échantillonnage aléatoire de pseudo-absence dans la zone infralittorale (Chefaoui *et al.*, 2008 ; Phillips *et al.*, 2009 ; VanDerWal *et al.*, 2009).

Afin de limiter les biais liés (1) à l'échantillonnage des pseudo-absences et (2) au jeu de données de présence, j'ai eu recours à des techniques de validation croisée, et à l'utilisation d'un modèle d'ensemble pour la prédiction.

La méthode de l'ensemble consiste à créer plusieurs jeux de données, à entraîner un modèle par jeu de données puis de faire la moyenne des prévisions des modèles indépendants générés. Dans mon cas, j'ai réalisé 10 échantillonnages aléatoires de pseudo-absences, et donc généré 10 jeux de données. Cette méthode permet de quantifier la stabilité des résultats d'un modèle à un autre, et ainsi de caractériser la robustesse des prédictions (i.e. la concordance entre les modèles) pour différentes zones géographiques.

Par ailleurs, la validation croisée permet via le sous-échantillonnage des données en jeux indépendants d'entraînement et de validation d'obtenir une estimation robuste de la performance du modèle. Dans cette étude, on peut distinguer deux niveaux de validation indépendante (1) de la capacité prédictive des modèles et (2) de leur capacité explicative, respectivement :

(1) Le premier niveau de validation indépendante a consisté à utiliser 80% des données pour l'entraînement et 20% pour la validation des prédictions pour les trois types de modèles (GLM, RF et BRT). Voir Annexe 1 concernant les détails de cette partition des données.

(2) La phase d'entraînement elle-même sur les 80% des données repose entièrement sur de la validation croisée. Pour la méthode GLM, il s'agit d'une validation croisée basée sur 10 sous-échantillonnage avec remise 70% / 30% pour chacun des 10 jeux de données. Cette validation croisée est inhérente aux approches Random Forest et BRT, de par leur construction et l'entraînement de chaque arbre sur des sous-échantillons (cf. description des modèles).

Pour résumer, trois méthodes statistiques (GLM, RF, BRT) sont utilisées pour prédire la répartition spatiale de chaque espèce. Les modèles d'ensemble RF et BRT moyennent les résultats obtenus à partir de 10 jeux de données. Dans le cas du modèle d'ensemble GLM, chacun des 10 jeux de données étant sous-échantillonné 10 fois aléatoirement avec remise, l'ensemble moyenne les résultats obtenus à partir de 100 jeux de données. Notez que l'ensemble des modèles GLM générés a été obtenu par une procédure *StepAIC*, et que les modèles RF et BRT ont été obtenus à l'aide de la librairie CARET (*Classification And REgression Training*) qui permet d'obtenir les paramètres optimaux de calibration des modèles.

### c. Comparaison des performances des différentes approches de modélisation

Le premier critère de comparaison est la courbe ROC (*receiver operating curve characteristic*) des modèles générés, et plus précisément la valeur AUC. Une courbe ROC est un graphique représentant les performances d'un modèle de classification. Cette courbe trace le taux de vrais positifs en fonction du taux de faux positifs. On associe à cette courbe une valeur AUC (« Area Under Curve » ou "aire sous la courbe ROC"). Cette valeur fournit une mesure agrégée des erreurs de classification. Les valeurs d'AUC sont comprises entre 0 à 1. Si toutes ses prédictions sont correctes, l'AUC est de 1, alors qu'un modèle nul (qui classe les observations au hasard) aura un AUC voisin de 0.5. Il y a plusieurs avantages à utiliser cette métrique : (1) il n'est pas seuil-dépendant et est donc une mesure de la performance du modèle seule, et 2) peut être utilisé pour toutes les méthodes, donc permet une comparaison inter-méthode des résultats.

Le deuxième critère de comparaison est la matrice de confusion associée à chaque modèle. La matrice de confusion détaille la qualité de la classification pour chacune des classes, à savoir ici les présences et les pseudo-absences. Elle permet donc d'aller au-delà de l'évaluation globale proposée par la métrique AUC en quantifiant les taux d'erreurs commises, et leur type. Dans cette étude, les matrices de confusion ont été obtenues à partir de seuils calculés avec les données d'entraînement selon le critère *Sensitivity = Specificity*. Ce critère permet d'obtenir de classer les données en faisant autant d'erreurs sur les présences que sur les absences.

Dans une matrice de confusion, pour un problème binaire de présence/absence, on a ainsi accès à 4 indicateurs : les vrais positifs quand le modèle prédit correctement la classe *Présence* ; les vrais négatifs quand le modèle prédit correctement la classe *Absence* ; les faux positifs quand le modèle prédit incorrectement la classe *Présence* ; et enfin, les faux négatifs quand le modèle prédit incorrectement la classe *Absence*. A partir de cette matrice, on peut savoir si le modèle fait beaucoup d'erreur en regardant le taux d'erreur moyen du modèle (faux négatif + faux positif), ainsi que la nature de ces erreurs en regardant les taux d'erreurs spécifiques.

**Tableau 2.** Description d'une matrice de confusion.

Matrice de confusion	Réalité	
	Présence réelle d'un herbier	Absence réelle d'un herbier
Prédiction du modèle		
Présence prédite	Vrai positif. Classification juste	Faux négatif. On prédit présence alors qu'il y a une absence réelle.
Absence prédite	Faux positif. On prédit absence alors qu'il y a une présence réelle.	Vrai négatif. Classification juste

Ces deux critères d'évaluation obtenus pour chacun des modèles générés, ont également été évalué pour les modèles d'ensemble.

Quel que soit le modèle utilisé (RF, GLM ou BRT), les performances de chacun des modèles sont meilleures avec le jeu de données constitué de toutes les variables, avec un filtre de 5 km pour corriger l'autocorrélation spatiale et à une résolution de 500m.

### 3. Prédictions spatiales des modèles

Un premier objectif de mon stage était de réaliser des prédictions spatiales de la distribution potentielle de l'habitat des deux espèces *Z. marina* et *Z. noltei*. A l'aide de modèles d'ensemble, j'ai réalisé pour chaque espèce trois types de prédictions spatiales : (1) une cartographie de l'indice d'habitat potentiel, (2) la cartographie de l'écart-type associé, ainsi (3) qu'une cartographie de présence/absence.

L'indice d'habitat potentiel est un indice quantifiant la probabilité que les conditions environnementales soient favorables à la persistance de l'espèce modélisée. Cet indice prend des valeurs entre 0 (pour un environnement statistiquement défavorable) et 1 (pour un environnement très fortement favorable). La cartographie à l'échelle de la façade Manche-Atlantique de cet indice d'habitat potentiel revient à quantifier la présence potentielle des deux espèces le long du littoral, et ainsi à définir une cartographie de référence des zones favorables aux herbiers de zostères.

La carte d'indice d'habitat potentiel de chaque espèce a été obtenue avec un modèle d'ensemble en calculant la moyenne des prédictions des modèles pondérées par le poids des modèles. Le poids d'un modèle est égale à la valeur AUC de ce modèle divisée par la somme des valeurs AUC de tous les modèles composant l'ensemble. Ainsi, les meilleurs modèles, ayant une plus grande valeur AUC, ont un poids plus important dans la construction de la carte de prédiction de l'indice d'habitat potentiel.

La cartographie de l'écart-type associé à la carte d'indice d'habitat potentiel permet de quantifier la stabilité des résultats (i.e. la variabilité des prédictions) entre les différents modèles, et propose donc une appréciation de la robustesse des prédictions le long de la façade.

A partir d'un modèle, il est possible d'avoir accès à deux types de prédictions : une prédiction entre 0 et 1 (indice d'habitat potentiel) ; ou une prédiction valant 0 ou 1 (présence = 1 ; absence = 0). L'accès à la prédiction de présence/absence n'est pas direct puisqu'il est nécessaire de calculer un seuil pour discrétiser la probabilité (indice d'habitat potentiel) en un indice binaire (0/1). Dans cette étude, j'ai fixé la condition de calcul du seuil  $Sensitivity = Specificity$ . En effet, je souhaitais obtenir un seuil de discrétisation faisant autant d'erreur sur les faux positifs que sur les faux négatifs. Ce seuil est spécifique à un modèle donné. Notez que les valeurs du seuil optimal pour discrétiser les prédictions d'indice d'habitat potentiel en Présence/Absence varient entre les modèles puisque ce seuil est fonction du jeu de données. Pour *Z. marina*, les seuils varient de 0.43 à 0.54 avec les GLM, de 0.31 à 0.705 avec les RF et de 0.48 à 0.61 avec les BRT. Pour *Z. noltei*, les seuils varient de 0.43 à 0.595 avec les GLM, de 0.25 à 0.69 avec les RF et de 0.3 à 0.66 avec les BRT.

Au sein du modèle d'ensemble, il y a plusieurs modèles, chacun peut produire une carte d'indice d'habitat potentiel, et à partir du seuil spécifique à chacun de ces modèles, il est possible d'obtenir la carte de présence/absence prédite par chacun de ces modèles. La carte de présence/absence finale a été obtenue en superposant l'ensemble des cartes de présence/absence générées par tous les modèles constituant l'ensemble, et en considérant que le modèle d'ensemble prédit une présence dans un pixel uniquement si plus de 50% des modèles prédisent une présence dans ce pixel. J'ai souhaité réaliser une cartographie de présence/absence afin de pouvoir rendre plus lisible l'information dans les cartes, et les rendre pratique pour les gestionnaires.

Notez que pour *Z. noltei*, qui est une espèce vivant en zone médiolittorale, seules les prédictions pour le premier pixel à partir du zéro hydrographique sont cartographiées.



#### 4. Caractérisation de la niche environnementale

Un deuxième objectif de mon stage était de quantifier et décrire l'influence des variables environnementales dans les prédictions des modèles afin de caractériser la niche écologique de chacune des deux espèces étudiées. Cette analyse peut se résumer selon 3 étapes.

La première étape vise à quantifier l'importance relative des variables explicatives aux prédictions du modèle. Pour les modèles GLM, l'importance relative des variables correspond à la déviance moyenne expliquée. Pour les modèles Random Forest, le critère *Mean Decrease in Accuracy* donne une estimation approximative de la perte de performance prédictive due à l'omission d'une variable. Enfin, pour les modèles BRT, la fonction *ggInfluence* (Jouffray *et al.*, 2019), qui repose également sur la perte associée à l'omission d'une variable, a été utilisée pour déterminer l'influence relative de chaque prédicteur dans les arbres de décisions.

La deuxième étape concerne la description de l'effet marginal des variables les plus influentes sur l'occurrence prédite des herbiers (effet linéaire ou non, effet positif ou négatif, présence de seuils etc.). Cette étape repose sur l'analyse de graphiques d'influence partielle des variables les plus importantes tels que présentés en figures 12 et 15. Afin d'analyser l'influence combinée des prédicteurs et de qualifier les interactions entre variables explicatives sur l'indice d'habitat potentiel, les figures 13 et 16 présentent des représentations en deux dimensions de l'influence croisée des variables les plus influentes sur les prédictions.

Enfin la dernière étape vise à dépasser l'analyse variable par variable de la niche écologique des deux espèces (Etape 1 et 2). Il s'agit ici de décrire l'occurrence potentielle de chaque espèce avec une approche biogéographique reposant sur des ensembles environnementaux homogènes identifiés par clustering selon la méthode des k-means le long de la façade Manche-Atlantique. Les clusters créés – appelés ici « écorégions » - représentent des environnements côtiers particuliers de la façade Manche-Atlantique française, et permettent de résumer de manière complémentaire la niche écologique des deux espèces d'herbiers en prenant en compte les combinaisons locales des 11 variables environnementales.

Le partitionnement par la méthode des k-means est basé sur des données centrées-réduites. Le nombre optimal de clusters estimé selon (1) l'indice de Calinski Harabasz, et (2) la méthode du coude est de 2 (Annexe 4). Cependant, afin de dépasser cette définition grossière selon 2 écorégions évidentes de part et d'autre de la pointe Finistère (i.e. la Manche et le Golfe de Gascogne), j'ai décidé de discriminer l'environnement littoral selon 7 écorégions. Ce choix se justifie autant empiriquement (de par l'interprétabilité écologique de chacune de ces 7 écorégions), que statistiquement. En effet, la visualisation par ACP (Annexe 5), et le test formel via une PERMANOVA ( $p$ -value < 0.05) confirment la pertinence du partitionnement selon ces 7 clusters. Les profils environnementaux des 7 écorégions construites correspondent à des environnements côtiers distincts et sont retranscrit dans le tableau 3 ci-dessous et en Annexe 6.

L'écorégion 1 (« Golfe Normand-Breton ») est caractérisée par des températures de l'air et de l'eau plus froide, des concentrations en phosphate, nitrate et chlorophylle a plus faibles et des valeurs de lumière plus faible que la moyenne de la façade.

L'écorégion 2 (« Atlantique abritée ») est caractérisée par des températures de l'air et de l'eau beaucoup plus chaude, des concentrations en phosphate, nitrate et chlorophylle a plus fortes, une salinité plus faible, une faible exposition aux vagues (fetch ; énergie cinétique des vagues) et au courant, et par une luminosité plus grande que la moyenne de la façade.

L'écorégion 3 (« Bretagne-Sud ») est caractérisée par des valeurs proches des moyennes de la façade mais se démarque par des températures de l'eau plus chaude que la moyenne.

L'écorégion 4 (« Manche-Est ») est caractérisée par des températures de l'eau beaucoup plus froide, des concentrations en phosphate, nitrate, chlorophylle a et oxygène dissous plus fortes et une salinité plus faible que la moyenne de la façade.

L'écorégion 5 (« Pointe de la Bretagne ») est caractérisée par des températures de l'air plus faibles, des concentrations en chlorophylle a plus faibles et une exposition aux vagues (fetch ; énergie cinétique des vagues) et au courant plus forte que la moyenne de la façade.

L'écorégion 6 (« Atlantique-Sud ») est caractérisée par des valeurs de températures de l'air et de l'eau beaucoup plus chaudes, une concentration en phosphate plus forte et une très forte exposition aux vagues (fetch ; énergie cinétique des vagues) comparées à la moyenne.

L'écorégion 7 (« Zones lumineuses ») est caractérisée par des valeurs proches de celles de la façade mais avec des valeurs de lumière bien plus forte que la moyenne.

**Tableau 3.** Présentation des conditions environnementales moyennes des écorégions

Variables \ Ecorégion N°	1	2	3	4	5	6	7	Moyenne de la façade
Surface (km <sup>2</sup> )	2805	978	2153	841	1094	736	1473	10080
Courant (N.m <sup>-2</sup> )	0.14	0.06	0.02	0.07	0.20	0.04	0.05	0.09
Vagues (N.m <sup>-2</sup> )	30.41	28.13	82.64	41.37	162.75	379.49	86.71	90.34
Fetch (m)	18384.97	6352.58	16200.73	16869.74	29516.71	23647.37	14874.54	17704.83
Temp.air_ete (°C)	16.99	19.11	17.53	17.21	16.27	19.65	17.23	17.48
Température (°C)	12.53	13.81	13.69	11.93	12.94	14.54	12.84	13.09
Lumière (mol.pho.m <sup>-2</sup> .j <sup>-1</sup> )	3.99	8.40	3.33	4.68	5.25	5.57	18.70	6.74
Salinité (PSU)	35.04	32.20	34.02	32.52	35.58	34.20	34.12	34.18
Oxygène (mg.L <sup>-1</sup> )	8.53	8.42	8.21	8.74	8.33	7.15	8.47	8.34
Chlo_a (µg.L <sup>-1</sup> )	1.72	4.85	2.80	3.58	1.03	2.36	2.47	2.49
Phosphate (µmol.L <sup>-1</sup> )	0.34	0.81	0.61	0.87	0.43	0.88	0.51	0.56
Nitrate (µmol.L <sup>-1</sup> )	6.28	28.54	16.09	31.74	6.62	15.89	13.26	14.42

## **5. Prédiction des conséquences potentielles du changement climatique**

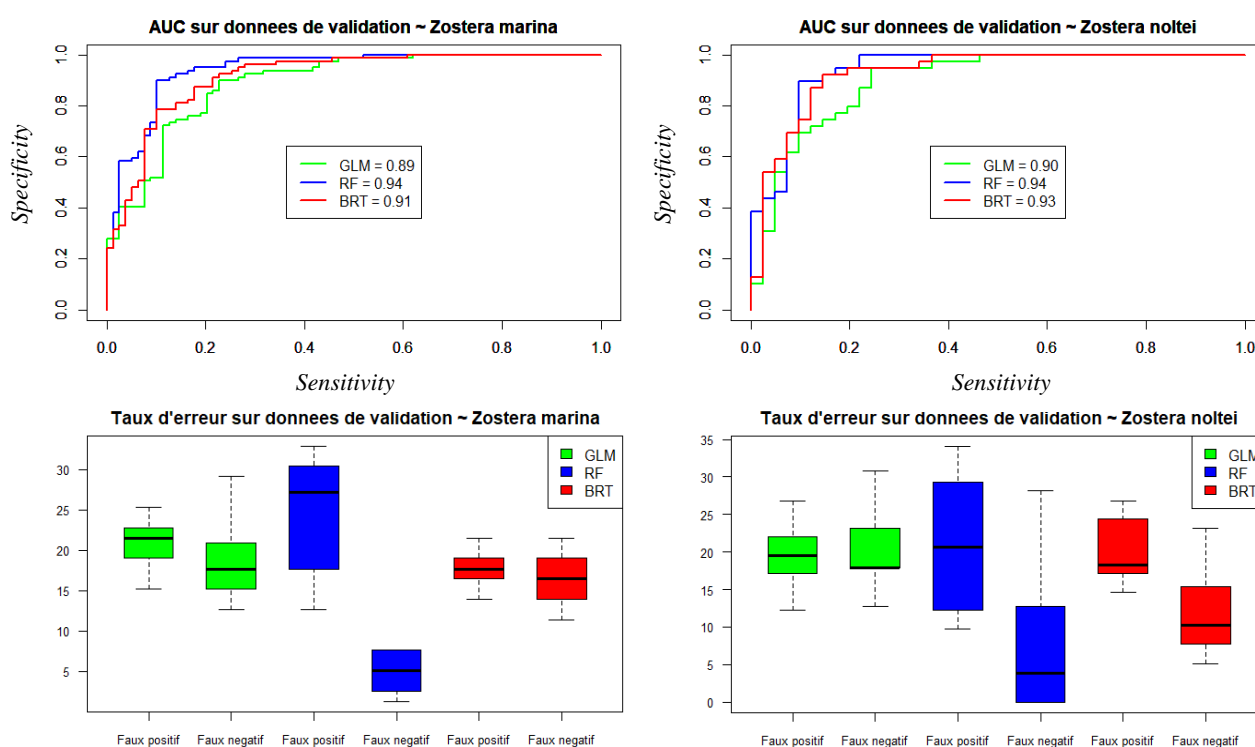
Enfin, le dernier objectif de cette étude était d'appréhender les conséquences potentielles du changement climatique sur la distribution de l'habitat potentiel des deux espèces *Z. marina* et *Z. noltei*. Pour cela, les modèles ont été construits en utilisant les 10 variables explicatives disponibles pour caractériser de manière consistante les climatologies récentes (2010-2020) et futures à horizon 2040-2050 et 2090-2100 (Tableau 1). Les prédictions spatiales ont été faites sur les décennies 2040-2050 et 2090-2100 selon les RCP4.5 (scénario socio-économique dit 'intermédiaire') et RCP8.5 (scénario socio-économique pessimiste dit 'Business-as-usual') en utilisant les prédictions régionalisées de changement climatique disponibles via le programme Copernicus (Copernicus Climate Change Service, 2020).

### III - Résultats

#### 1. Comparaison des performances des modèles

Les performances de chacun des modèles (GLM, RF, BRT) sont meilleures avec le jeu de données constitué de toutes les variables, avec un filtre de 5 km pour corriger l'autocorrélation spatiale et à une résolution de 500m. Il s'agit donc du jeu de données final pour cette étude.

Les 3 approches testées (GLM, Random Forest, Boosted Regression Trees) ont des excellentes performances prédictives sur les 20% de données de validation pour les deux espèces. En effet, les AUC varient entre 0.89 (GLM pour *Z. marina*) et 0.94 (Random Forest pour les deux espèces) (Figure 6). Par ailleurs, les taux moyens d'erreur de prédiction obtenus par validation croisée (faux positifs et faux négatifs confondus) varient entre 20,3% et 14,7%. Notez que les meilleures performances prédictives sont obtenues avec les modèles Random Forest. En effet, quelle que soit l'espèce modélisée, les modèles Random Forest sont les modèles ayant la meilleure valeur AUC sur les données de validation (*Z. marina* = 0.94 ; *Z. noltei* = 0.94), devant les BRT (*Z. marina* = 0.91 ; *Z. noltei* = 0.93) puis les GLM (*Z. marina* = 0.89 ; *Z. noltei* = 0.90). Si on s'intéresse aux taux d'erreurs de ces modèles, le taux d'erreur moyen des Random Forest est également le plus faible (*Z. marina* = 16.7% ; *Z. noltei* = 14.7%), encore une fois devant les BRT (*Z. marina* = 17.2% ; *Z. noltei* = 15.9%) et les GLM (*Z. marina* = 19.8% ; *Z. noltei* = 20.3%). Ainsi, il semblerait dans notre cas que les modèles du type forêt d'arbre décisionnel (RF, BRT) performent mieux que les modèles de type GLM. Les modèles Random Forest se démarquent particulièrement par leur faible taux de faux négatifs (*Z. marina* = 8.7% ; *Z. noltei* = 8.2%). Ces modèles ont donc une meilleure capacité que les BRT et GLM pour prédire la présence d'herbiers. Cependant, comme avec les autres modèles, le taux de faux positifs estimé par validation croisée est autour de 20-25%. Ainsi, les taux de faux positifs modérément élevés relativement aux faibles taux de faux négatifs révèlent que les prédictions des modèles ont tendance à surestimer la présence potentielle des deux espèces.



**Figure 6.** Performances des modèles : AUC sur données de validation (en haut) et taux d'erreur par modèle (faux positifs / faux négatifs) (en bas) pour *Z. marina* (à gauche) et *Z. noltei* (à droite)

En raison de leurs meilleures performances prédictives relativement aux deux autres approches envisagées, je ne présente dans les sections suivantes que les résultats des modèles Random Forest notamment en termes de prédictions spatiales, de caractérisation de la niche environnementales des deux espèces et de prédiction des effets du changement climatique sur la distribution des deux espèces.

## **2. Prédictions spatiales des modèles Random Forest**

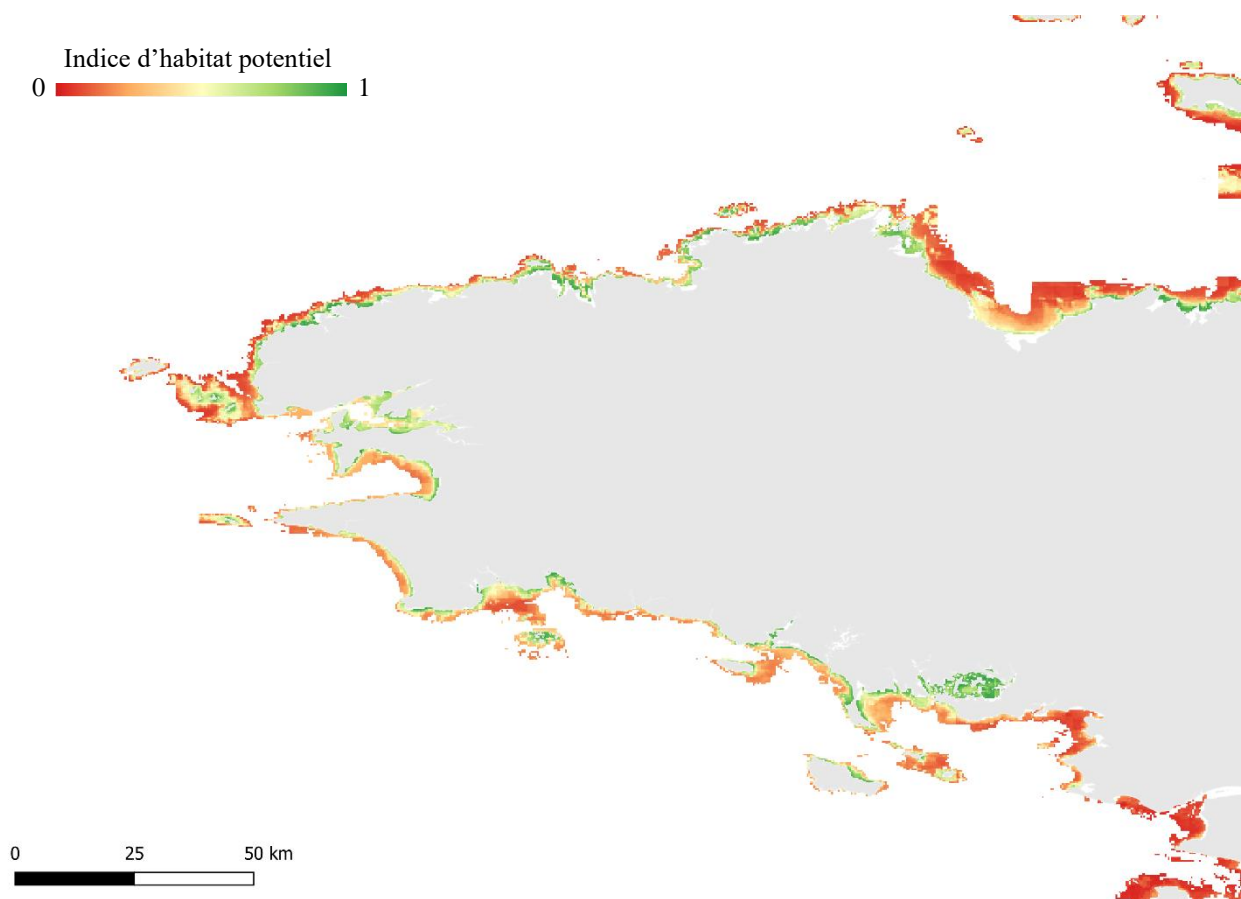
A l'aide du modèle d'ensemble Random Forest, j'ai réalisé trois types de prédictions spatiales : (1) une cartographie de l'indice d'habitat potentiel, (2) la cartographie de l'écart-type associé, ainsi (3) qu'une cartographie de présence/absence. L'ensemble de ces prédictions spatiales à l'échelle de la façade Manche-Atlantique française est disponible en Annexe (Annexes 8 à 14). Les figures 7 à 10 présentent les prédictions obtenues sur la distribution potentielle de l'habitat de *Zostera marina* à l'échelle de la Bretagne.

La variabilité des prédictions d'un modèle à l'autre (au sein de l'ensemble) reflète une certaine stabilité des résultats en termes d'indice d'habitat avec un écart type moyen de 0.073 pour *Zostera marina* et 0.067 pour *Zostera noltei*. Il existe toutefois des zones, comme la rade de Brest pour *Z. noltei* ou le golfe Normand-Breton pour *Z. marina*, où le degré de précision des prédictions est plus faible, avec des valeurs d'écart-type maximale de 0.26 et 0.29 pour *Z. marina* et *Z. noltei*.

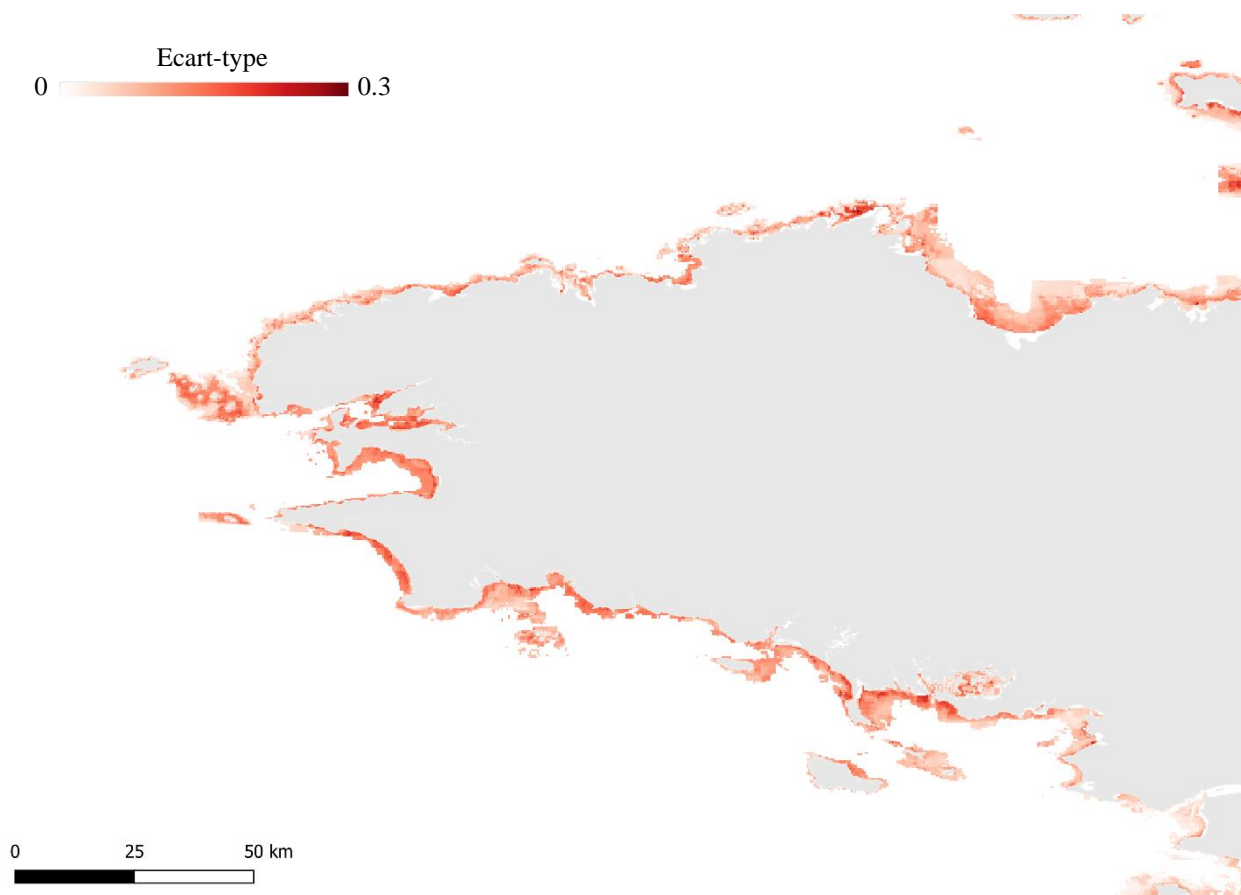
Au-delà de la validation quantitative des performances des modèles (via les AUC et taux d'erreur), la bonne concordance entre connaissances empiriques ou expertes des zones d'occurrence des herbiers et la distribution potentielle prédite pour chacune des espèces à l'échelle de la façade apporte une forme de validation qualitative de la performance prédictive des modèles. Par exemple, les prédictions des modèles sont écologiquement cohérentes pour plusieurs sites le long de la façade avec des valeurs élevées d'indice d'habitat potentiel dans le Golfe du Morbihan et sur littoral de Bretagne Nord et Bretagne Sud pour *Z. marina* ; et dans le bassin d'Arcachon, au niveau des pertuis Breton et pertuis d'Antioche, ou encore en Bretagne Nord pour *Z. noltei*.

Cependant, les cartes mettent également en valeur des zones où la présence des espèces est peu vraisemblable. Par exemple, les prédictions d'un habitat potentiel favorable à *Z. noltei* aux niveaux des estuaires de la Seine, de la Loire, de la Vilaine, en Rade de Brest ou encore en Baie de Douarnenez sont peu réalistes. De même, on observe une présence potentielle prédite de *Z. marina* sur l'intégralité du littoral breton, ou encore présence prédite importante dans l'ensemble de bassin d'Arcachon. Comme on le peut l'observer dans la figure 3, la principale espèce présente dans le bassin d'Arcachon est *Z. noltei*. Ainsi, ces résultats suggèrent comme énoncé plus haut une tendance du modèle à surestimer la présence potentielle des deux espèces.

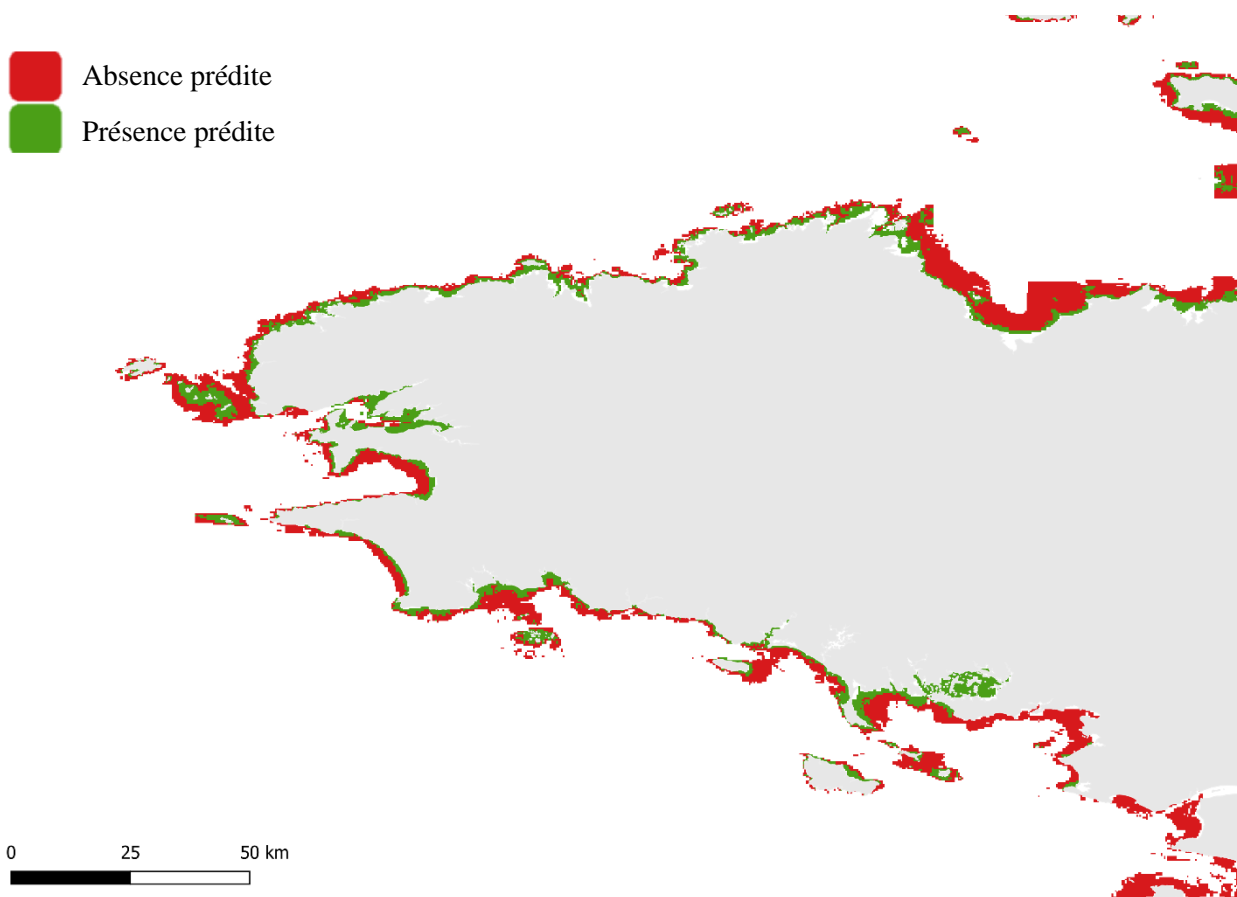
Les cartes de prédictions obtenues avec les modèles prennent uniquement en compte les conditions climatiques et biogéochimiques du milieu, et non la nature géologique des fonds marins. Toutefois, une carte a été produite en utilisant la carte de prédiction des présences/absences et en utilisant la variable *Substrat* comme un masque. Cela a permis d'obtenir une carte dans laquelle il n'y a pas de présence prédite à des endroits où il est connu qu'il y a de la roche. Si on compare les figures 9 (sans le masque *Substrat*) et 10 (avec le masque *Substrat*), on peut observer que le modèle a surestimé la présence de *Z. marina* en Bretagne et l'utilisation de ce masque permettrait d'obtenir des résultats plus réalistes à l'échelle de la façade Manche-Atlantique. Nous reviendrons sur la qualité des cartes en discussion.



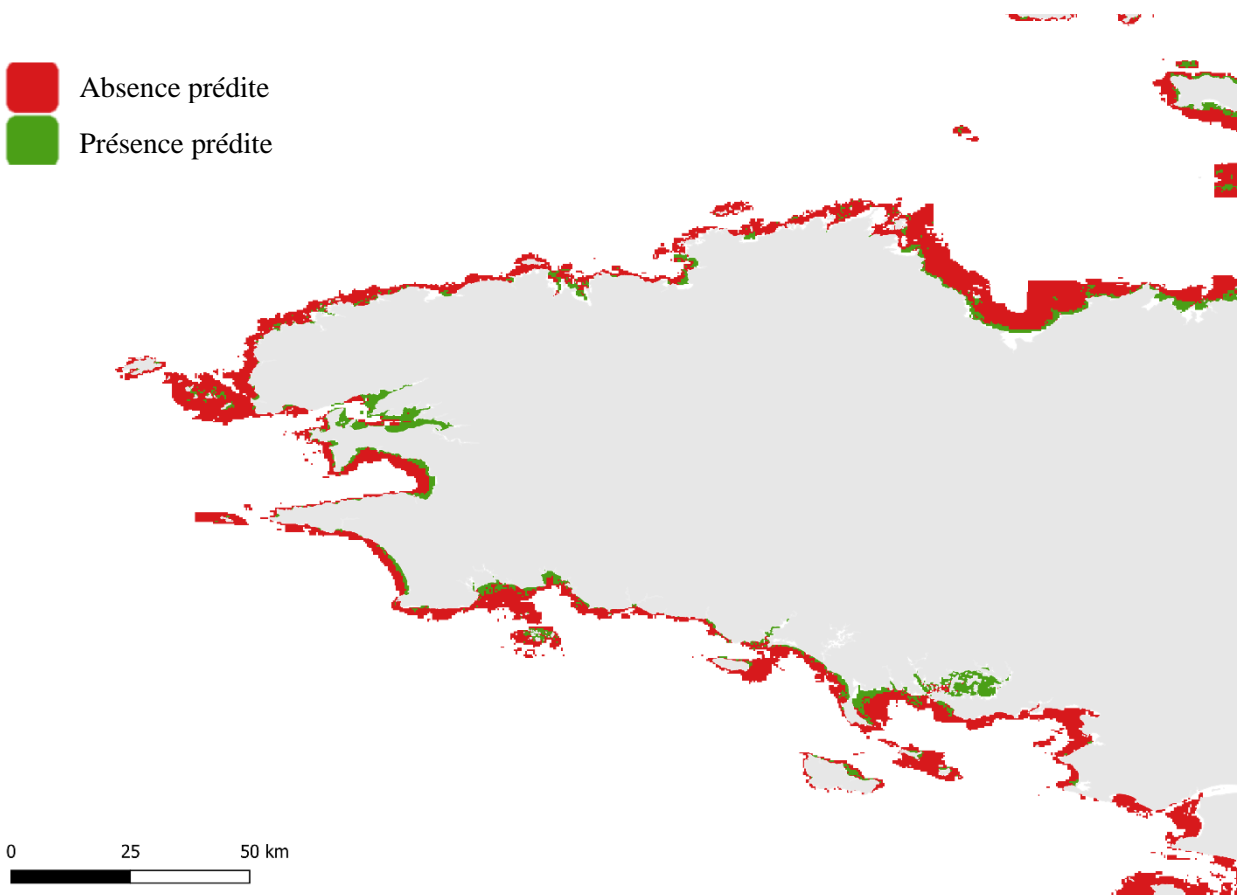
**Figure 7.** Cartographie de l'indice d'habitat potentiel de *Z. marina* en Bretagne



**Figure 8.** Cartographie de l'écart-type de l'indice d'habitat potentiel de *Z. marina* en Bretagne



**Figure 9.** Cartographie des présences et absences de *Z. marina* en Bretagne



**Figure 10.** Cartographie des présences et absences de *Z. marina* en Bretagne avec le masque Substrat



### 3. Caractérisation de la niche environnementale

Cette partie quantifie et décrit l'influence des variables environnementales dans les prédictions des modèles afin de caractériser la niche écologique de chacune des deux espèces étudiées.

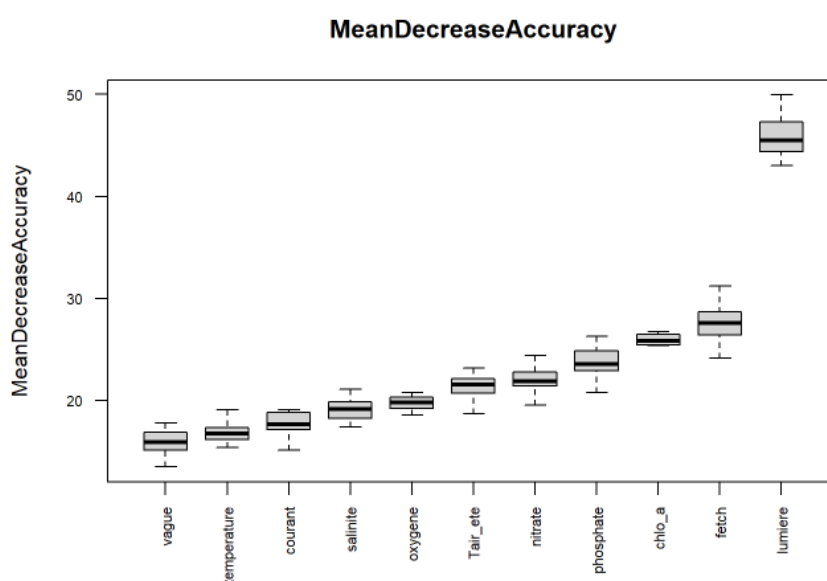
Je rappelle ici que la première étape de notre raisonnement vise à quantifier l'importance relative des variables explicatives aux prédictions du modèle. La deuxième étape concerne la description de l'effet marginal des variables les plus influentes sur l'occurrence prédite des herbiers. Enfin la dernière étape vise à décrire l'occurrence potentielle de chaque espèce avec une approche biogéographique.

- *Zostera marina*

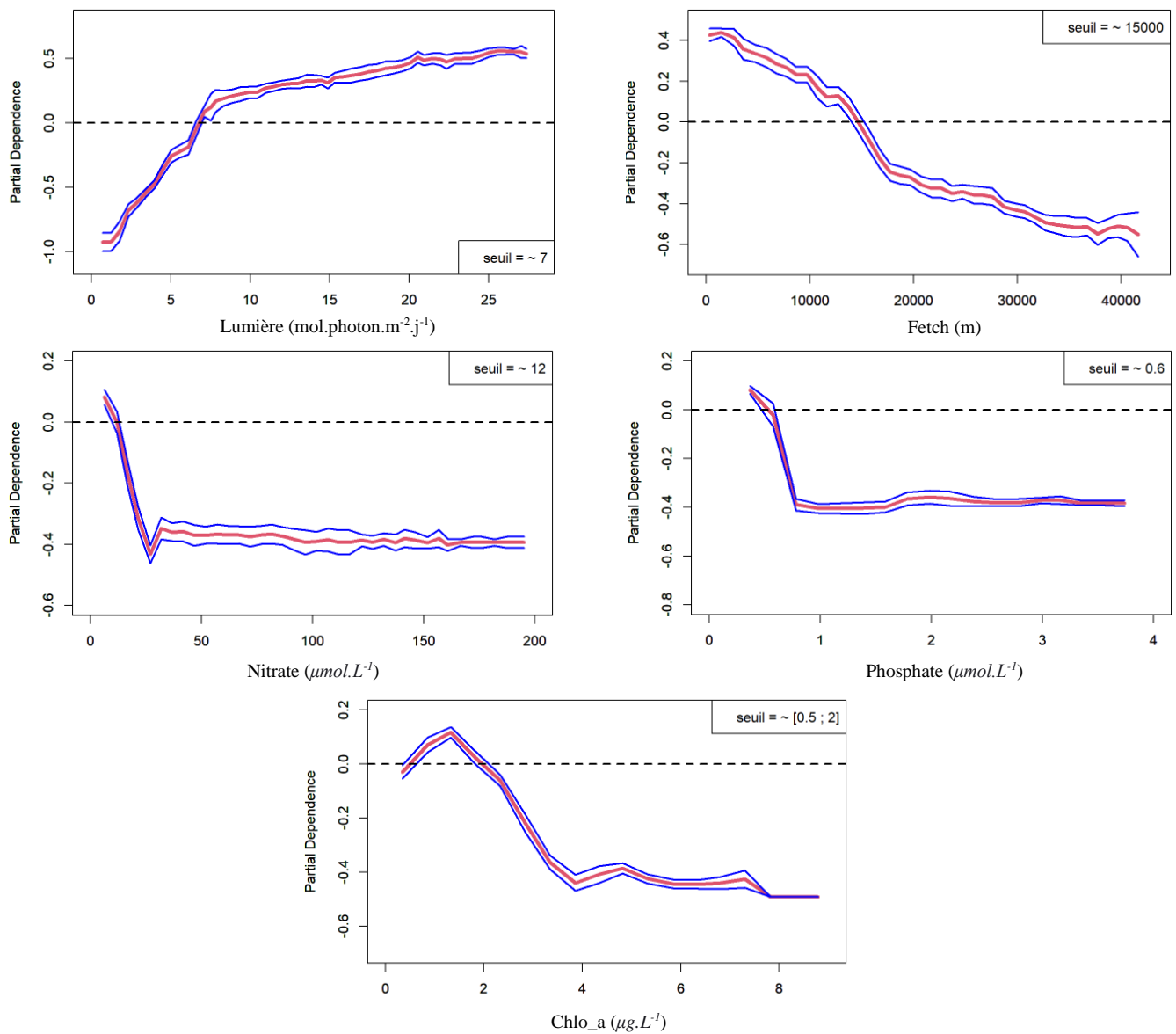
Les prédictions de l'habitat potentiel de *Z. marina* dépendent par ordre décroissant d'influence de la *Lumière* avec une valeur moyenne de *MeanDecreaseAccuracy* de 45.5%, du *Fetch* (27.6%), suivi de la *Chlorophylle\_a* (25.6%), du *Phosphate* (23.8%) et du *Nitrate* (21.9%) (Figure 11).

Les conditions favorables à la présence de *Z. marina* en zone infralittorale correspondent à de forts niveaux de lumière ( $> 7 \text{ mol.photons.m}^{-2}.\text{j}^{-1}$ ), une faible exposition aux vagues (valeurs de *Fetch*  $< 15000 \text{ m}$ ), des concentrations modérées à faibles en nutriments (nitrate  $< 12 \mu\text{mol.L}^{-1}$ ; phosphate  $< 0.6 \mu\text{mol.L}^{-1}$ ), et des niveaux faibles à modérés de chlorophylle a (entre 0.5 et  $2 \mu\text{g.L}^{-1}$ ) (Figure 12). Cette combinaison de conditions environnementales suggère un habitat potentiellement favorable à *Z. marina* correspondant à des zones côtières abritées, peu turbides et peu eutrophisée (nitrate, phosphate, chlorophylle a). (Figure 12).

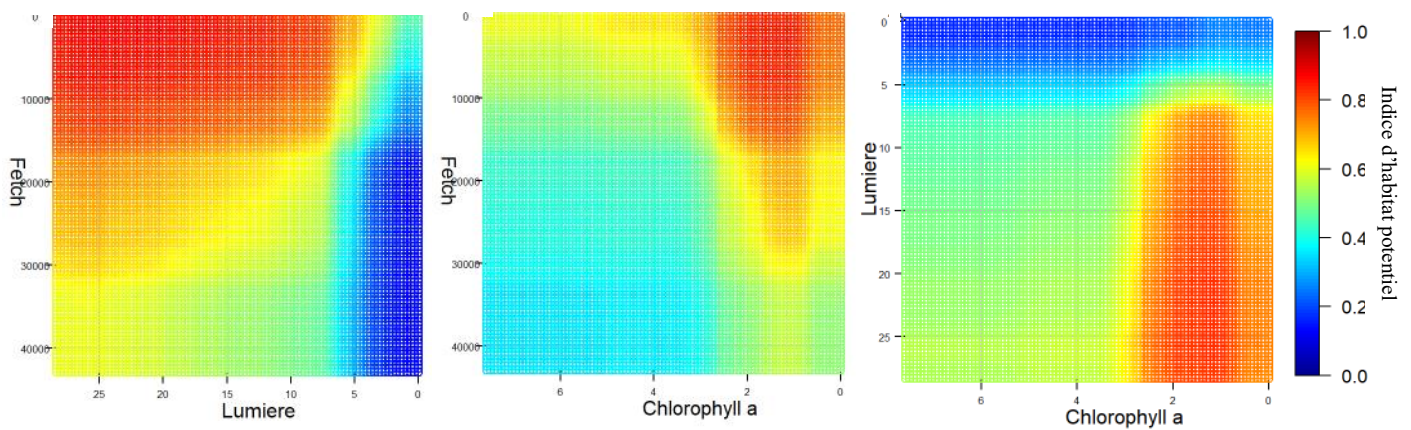
Les graphes des interactions 2 à 2 entre les 3 variables les plus structurantes sur les prédictions du modèle Random Forest (Figure 13) suggèrent que les herbiers de *Z. marina* requièrent une combinaison de conditions favorables en termes de *Lumière*, de *Fetch*, et de *Chlorophylle\_a*. Aussi, les partial plots (Figure 12) et plots d'interactions (Figure 13) permettent de caractériser les seuils environnementaux qui délimitent les conditions favorables à la présence de *Z. marina* ( $Lumière > 7 \text{ mol.photons.m}^{-2}.\text{j}^{-1}$ ;  $Fetch < 15000 \text{ m}$ ; *Chlorophylle\_a* : entre 0.5 et  $2 \mu\text{g.L}^{-1}$ ).



**Figure 11.** Représentation de l'importance relative des variables sur la base du critère *MeanDecreaseAccuracy* dans les modèles Random Forest pour *Zostera marina*



**Figure 12.** Représentation de l'effet marginal des variables les plus structurantes dans les modèles de distribution sur la présence de *Zostera marina* à l'aide de partial plot



**Figure 13.** Représentation de l'interaction entre les 3 variables les plus structurantes (*Lumière*, *Fetch*, *Chlorophylle\_a*) dans les prédictions du modèle Random Forest pour *Z. marina*

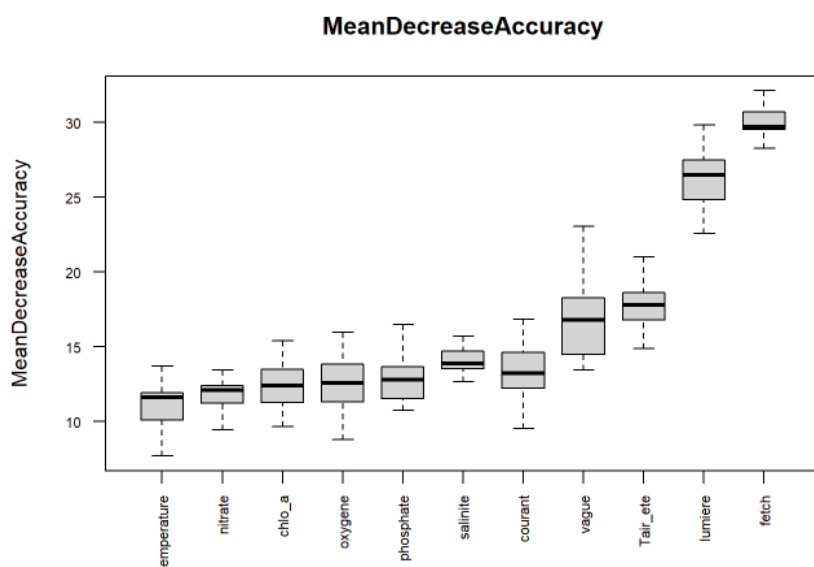
- *Zostera noltei*

Les variables *Fetch* et *Lumière* sont les deux variables les plus structurantes du modèle, avec des valeurs moyennes de *MeanDecreaseAccuracy* de 32.1% pour le *Fetch* et 27.6% pour la *Lumière*. Suivent ensuite les variables *Température estivale de l'air* (17.4%), *Vague* (16.8%), *Courant* (14.1%) et *Salinité* (13.9%) (Figure 14).

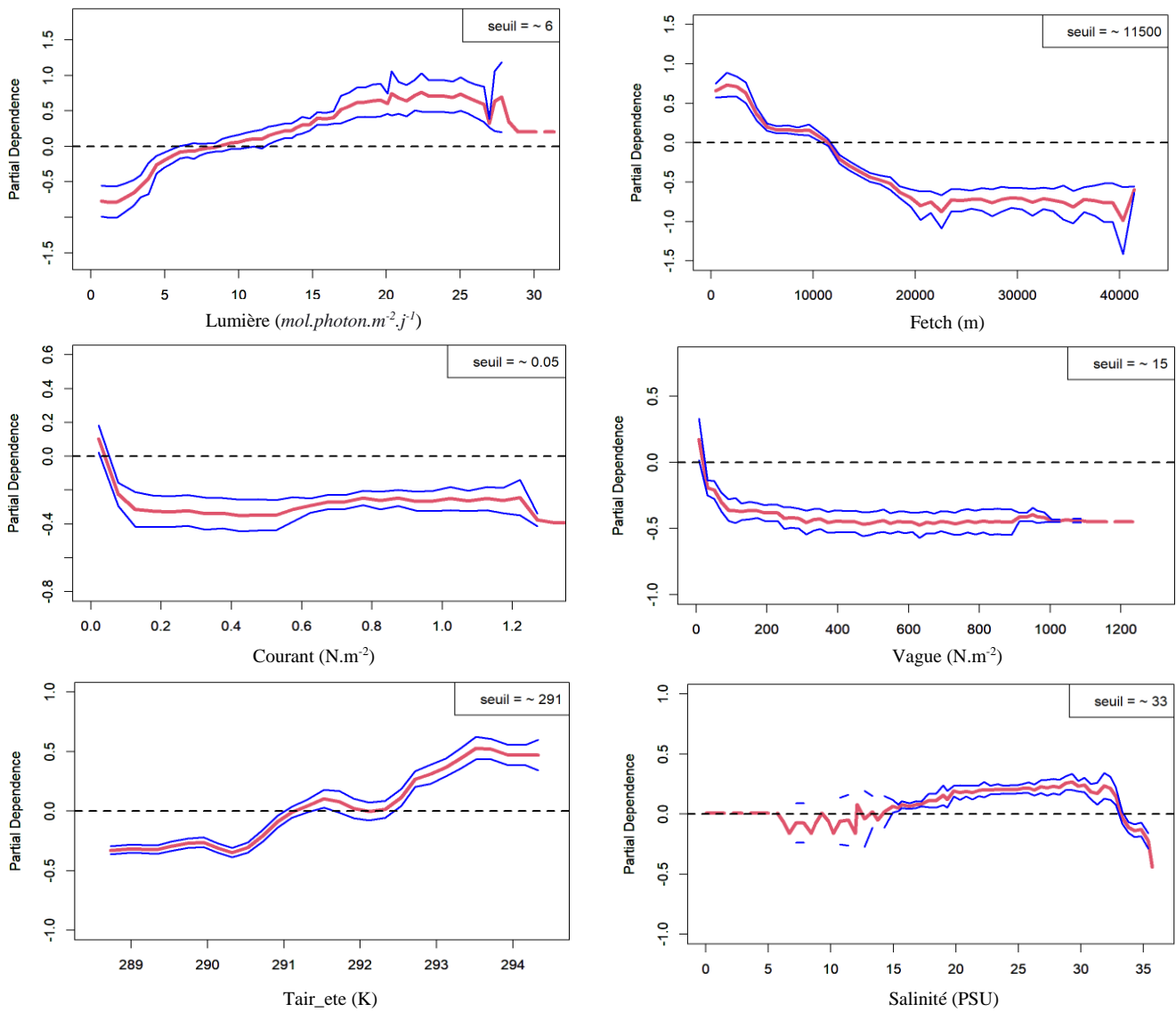
Il est intéressant d'observer dans un premier temps l'importance de la température de l'air dans la structuration de notre modèle. En effet, nous avons fait l'hypothèse en amont de cette étude que prendre cette variable pouvait être une variable structurante pour *Z. noltei* dans la mesure où cette espèce vit en zone médiolittorale, et est par conséquent exposée à l'air à marée basse.

Les modèles Random Forest prédisent que les zones favorables à la présence de *Z. noltei* sont peu exposées aux vagues (valeurs de fetch < 11500m ; et énergie cinétique moyenne induite par les vagues < 15 N.m<sup>-2</sup>) ou au forts courants (< 0.05 N.m<sup>-2</sup>). Ces zones sont également caractérisées par une turbidité réduite (fortes valeurs de lumière > 6 mol.photons.m<sup>-2</sup>.j<sup>-1</sup>), des valeurs de salinité entre 5 et 33 PSU et des températures estivales moyennes de l'air > 291K ~ 17°C. La variable *Température estivale de l'air* permet d'appréhender l'habitat potentiel sur un gradient Nord/Sud en opposant notamment la façade Manche et la façade Atlantique. L'influence de la *Salinité* révèle un optimal environnemental associé à des eaux saumâtres, notamment aux niveaux des estuaires ou de baies semi-ouvertes (par exemple Arcachon ou golfe du Morbihan), zones d'échange entre l'eau douce et l'eau de mer. Ainsi, l'habitat potentiel de *Z. noltei* correspondrait à des zones intertidales abritées du vent, des vagues et du courant, peu turbides et saumâtres (Figure 15).

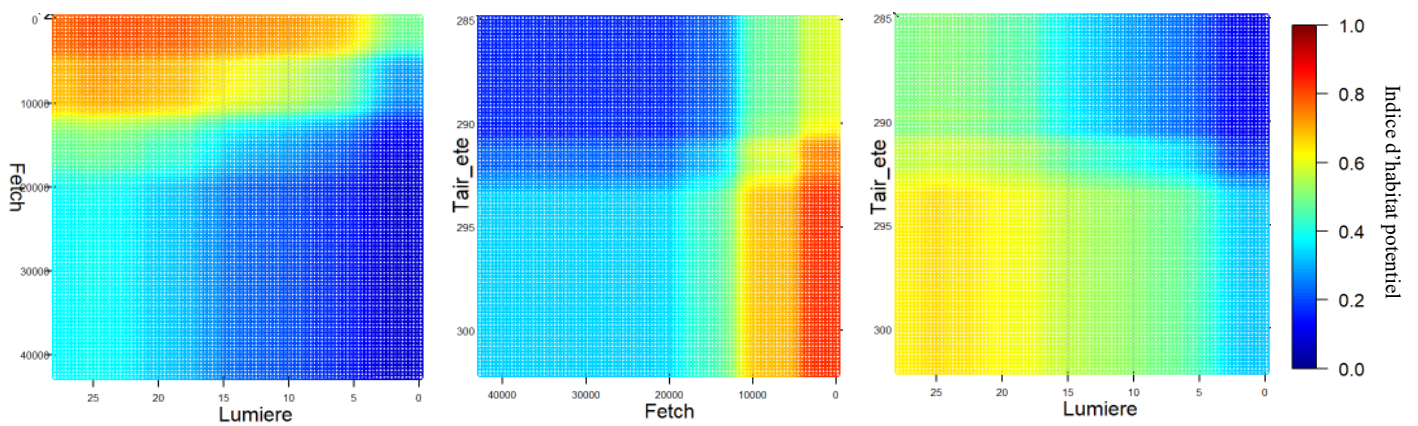
La figure 16 représente l'effet des interactions 2 à 2 entre les 3 variables les plus influentes sur les prédictions d'habitat potentiel de *Z. noltei*. La plage environnementale la plus favorable à *Z. noltei* correspond à des valeurs de lumière élevées (> 15 mol.photons.m<sup>-2</sup>.j<sup>-1</sup>), des valeurs de températures estivales de l'air > 291K et des valeurs de fetch faibles (> 5000 m).



**Figure 14.** Représentation de l'importance relative des variables sur la base du critère *MeanDecreaseAccuracy* dans les modèles Random Forest pour *Zostera noltei*



**Figure 15.** Représentation de l'effet marginal des variables les plus structurantes dans les modèles de distribution sur la présence de *Zostera noltei* à l'aide de partial plot



**Figure 16.** Représentation de l'interaction entre les 3 variables les plus structurantes (*Fetch*, *Lumiere*, *Température estivale de l'air*) dans les prédictions du modèle Random Forest pour *Z. noltei*

- Bilan par écorégions.

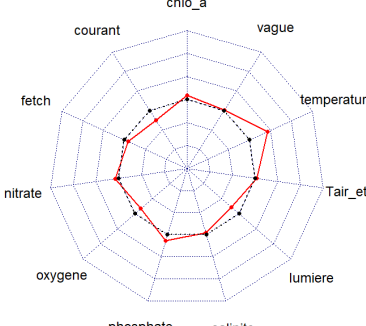
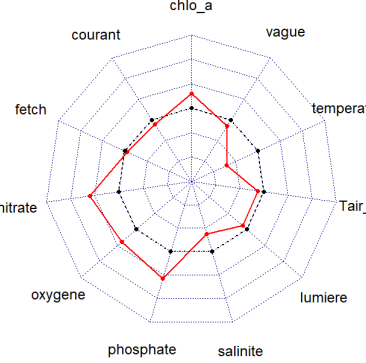
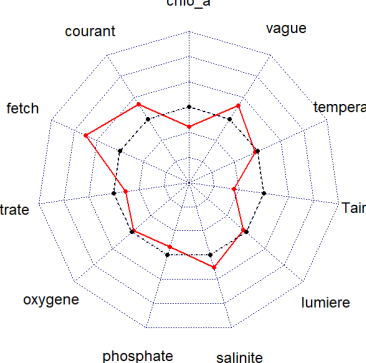
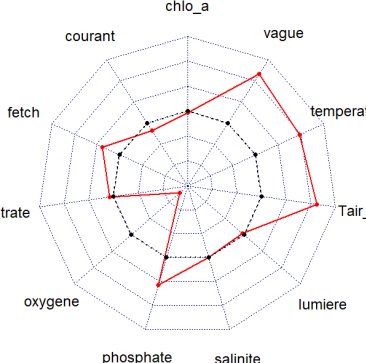
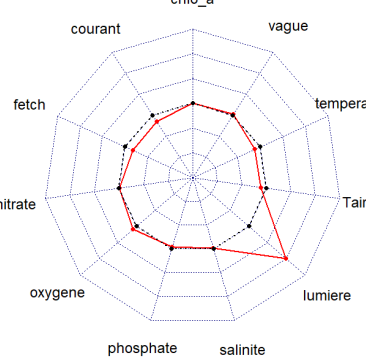
Afin de compléter la définition de la niche écologique de chacun des deux espèces, je présente ici une synthèse des prédictions d’habitat potentiel au sein de chacune des sept écorégions, i.e. des 7 ensembles environnementaux homogènes identifiés par clustering selon la méthode des k-means le long de la façade Manche-Atlantique. Les profils environnementaux des 7 écorégions caractérisent des environnements côtiers distincts (voir Tableau 3 et les radars plots dans le Tableau 4).

Le tableau 4 résume les prédictions obtenues dans les différentes écorégions. Tout d’abord, l’écorégion 7 (« Zone lumineuse ») correspond à l’habitat potentiel le plus favorable aux deux espèces. En effet, 42.6 % de l’habitat potentiel de *Z. marina* et 42.9% de celui de *Z. noltei* se situeraient cette zone. Ainsi, les zones avant tout très lumineuses (donc peu turbides et peu agitées) seraient les plus favorable aux deux espèces. Par ailleurs, un second résultat majeur de ce bilan par écorégions révèle que les deux espèces ne possèdent pas la même distribution potentielle à l’échelle de la façade Manche-Atlantique française. En effet, 80% de l’habitat potentiel de *Z. marina* se répartit dans les écorégions 1, 3 et 7 alors que 80% de l’habitat potentiel de *Z. noltei* se répartit dans les écorégions 2 et 7. Enfin, cette synthèse par écorégion permet également d’identifier les zones défavorables à la présence de chaque espèce, à savoir les écorégions 4 et 6 pour *Zostera marina*, et les écorégions 4, 5 et 6 pour *Zostera noltei*.

**Tableau 4.** Prédications spatiales dans les écorégions

Ecorégion	Descriptions de l’écorégion	<i>Zostera marina</i>		<i>Zostera noltei</i>	
		%Ecorégion occupée	Proportion de l’espèce	%Ecorégion occupée	Proportion de l’espèce
N°1 <i>Golfe Normand – Breton</i>		15.5	17.9 %	3.0	5.4 %
N°2 <i>Atlantique abritée</i>		19.1	7.7 %	62.9	39.9 %



<p>N°3 <i>Bretagne Sud</i></p>		21.2	18.8 %	6.9	9.6 %
<p>N°4 <i>Manche Est</i></p>		1.0	0.3 %	1.6	0.9 %
<p>N°5 <i>Pointe de la Bretagne</i></p>		27.7	12.5 %	0.2	0.2 %
<p>N°6 <i>Atlantique Sud</i></p>		0.0	0.0 %	2.2	1.1 %
<p>N°7 <i>Zones lumineuses</i></p>		69.9	42.6 %	44.9	42.9 %



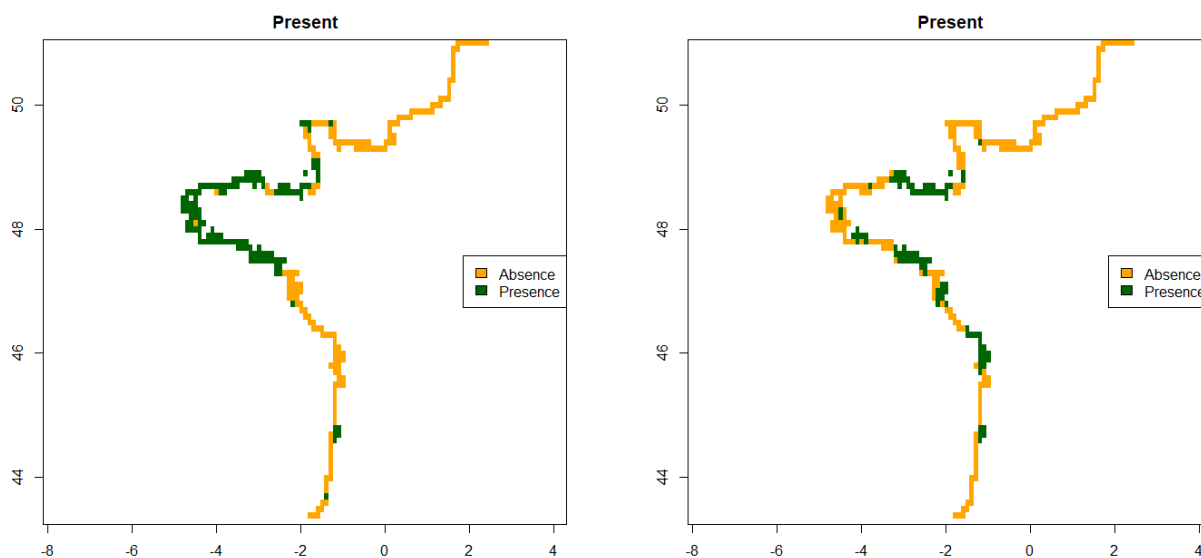
#### 4. Prédiction des conséquences potentielles du changement climatique

Dans cette dernière partie concernant les effets du changement climatique, les modèles Random Forest de distribution ont été entraînés avec 10 variables explicatives pour lesquelles des projections futures étaient disponibles mais à une résolution spatiale de 11 km (donc dégradée par rapport aux résultats présentés ci-dessus). Les performances de ces modèles restent néanmoins excellentes et globalement comparables à celles des modèles présentés précédemment : pour *Zostera marina*, la valeur AUC sur les données de validation est de 0.92, et le taux d'erreur moyen est de 18%, et le taux de Faux négatifs de 12.6% (*plus bas que pour les modèles à fine résolution*). Pour *Zostera noltei*, les performances sont légèrement moins bonnes avec une valeur AUC sur les données de validation de 0.85, et un taux d'erreur moyen de 27.3%.

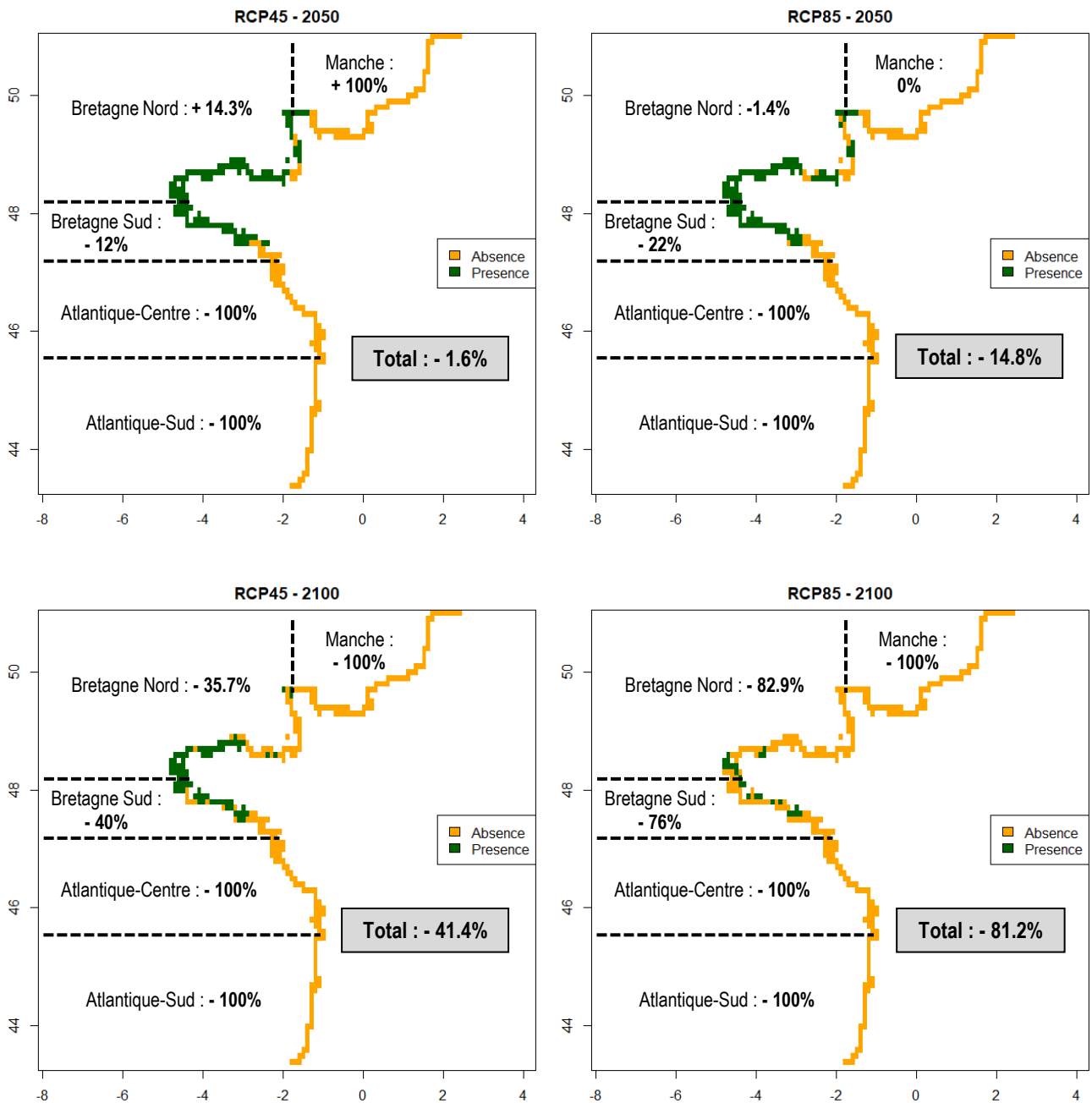
Les prédictions suggèrent quel-que-soit le scénario considéré un déclin important des herbiers de *Z. marina* sur le long terme (horizon 2100). Bien qu'une expansion locale de leur répartition (+14.3% en Bretagne Nord) soit possible à l'horizon 2050 (RCP4.5), la tendance à l'échelle de la façade pour les herbiers de *Z. marina* est au déclin à l'horizon 2100. Le scénario RCP8.5 serait le plus critique pour cette espèce avec un déclin prédit de 81.2% de son habitat potentiel le long de la façade (Figure 18).

L'habitat potentiel de *Z. noltei* risque de décliner en surface et devenir plus défavorable à l'échelle de la façade Atlantique, alors que les conditions environnementales pourraient devenir plus favorables en Manche. Quel que soit le scénario RCP, le changement climatique pourrait conduire à une expansion de la répartition des herbiers de *Z. noltei* sur la façade Manche-Atlantique de l'ordre de 39% avec le RCP4.5 et de 29% avec le RCP8.5 (Figure 19).

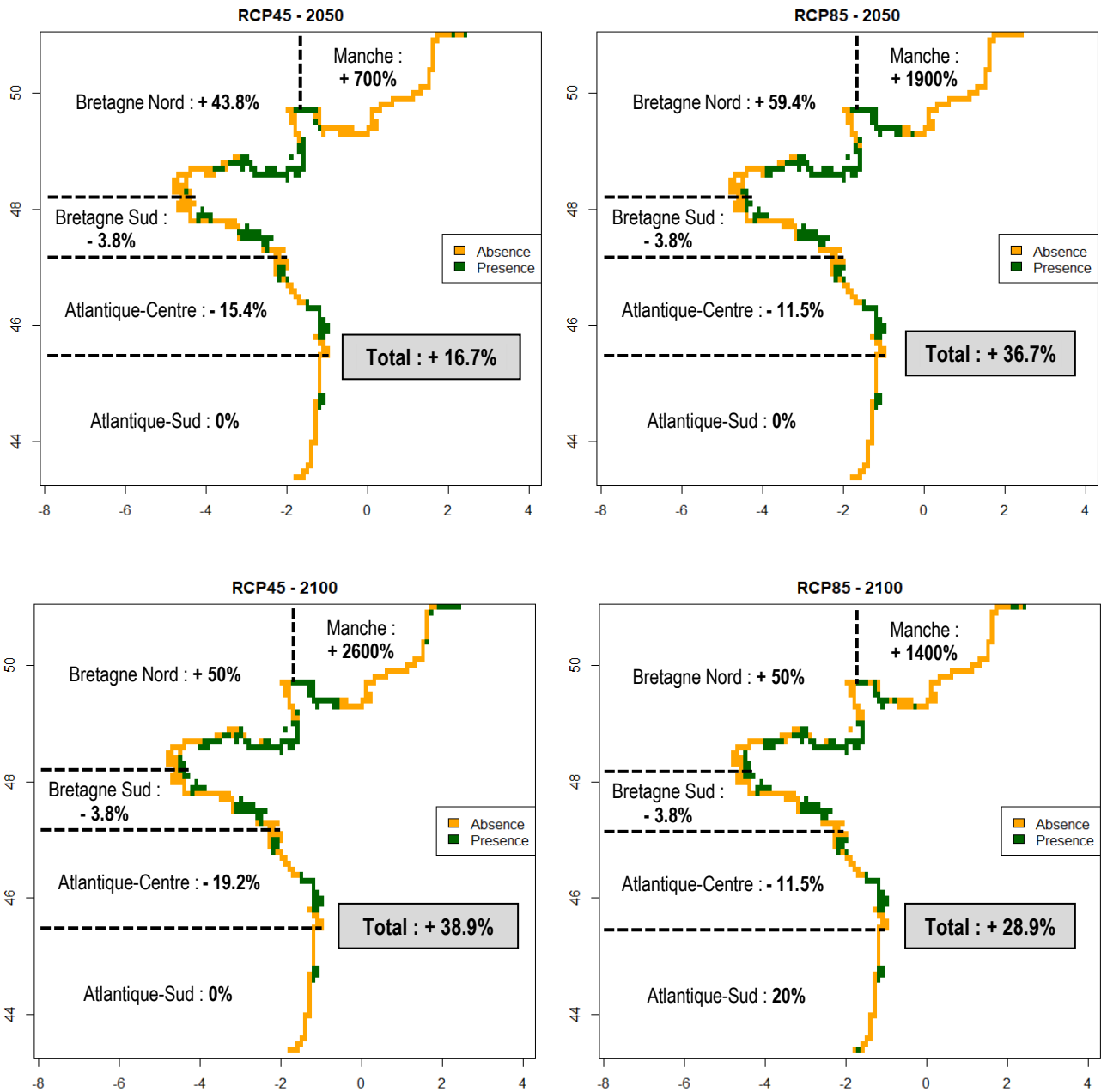
Enfin, les prédictions suggèrent des réponses contrastées de l'habitat potentiel de *Z. noltei* selon le scénario RCP envisagé. Avec le RCP4.5, on observe une tendance long-terme à l'expansion, de l'ordre de +~17% à horizon 2050, puis de +~22% entre 2050 et 2100. En revanche, dans le cas du scénario plus pessimiste RCP8.5 'business-as-usual', on observe une forte expansion de +~37% à l'horizon 2050, suivi d'une phase de déclin modérée de 6% entre 2050 et 2100. Ainsi, avec le RCP8.5, il semblerait que les effets du changement climatique bénéficient aux herbiers de *Z. noltei* dans un premier temps, avant les conditions environnementales ne leur deviennent finalement défavorables à l'horizon 2100 (Figure 19).



**Figure 17.** Cartes de prédictions obtenues avec les modèles RF, à 11km de résolution, de la distribution potentielle actuelle de *Z. marina* (à gauche) et de *Z. noltei* (à droite)



**Figure 18.** Conséquences prédites du changement climatique (RCP4.5 et RCP 8.5) sur la distribution spatiale de l'habitat potentiel de *Z. marina* sur la façade Manche-Atlantique aux horizons 2050 et 2100.



**Figure 19.** Conséquences prédites du changement climatique (RCP4.5 et RCP 8.5) sur la distribution spatiale de l'habitat potentiel de *Z. noltei* sur la façade Manche-Atlantique aux horizons 2050 et 2100.

## Discussion

- **Impacts potentiels du changement climatique et nécessité de mettre en place des mesures pour la gestion des espèces et habitats**

Dans un contexte mondial où habitats et espèces marines sont soumis à de multiples pressions en particulier en zone littoral (Turner *et al.*, 1996 ; Lotze *et al.*, 2006 ; Halpern *et al.*, 2008), la résilience des écosystèmes côtiers est mise à rude épreuve. Les résultats de notre étude montrent que le changement climatique pourrait avoir des conséquences néfastes sur l'habitat potentiel des herbiers de zostères, et entraîner leur déclin à l'échelle de la façade Manche-Atlantique à horizon 2100. Si les sorties numériques des modèles sont à interpréter avec précaution, le déclin prédit des herbiers de *Zostera marina* de 81% à l'échelle de la façade Manche-Atlantique à l'horizon 2100 dans le cas du RCP8.5 suggère néanmoins une tendance généralisée de perte en quantité et en qualité de cet habitat côtier.

Ehlers *et al.* (2008) mentionnent dans leur étude que la diversité génotypique est l'élément clé pour la résilience et le maintien du fonctionnement de l'écosystème des herbiers marins dans un contexte de changement climatique. Toutefois, en s'appuyant sur notre étude, on pourrait penser que la diversité génétique puisse être affectée par le déclin, ne subsistant que les génotypes adaptés au nouvel environnement. Le déclin des herbiers pourrait donc avoir des conséquences importantes sur la biodiversité fonctionnelle des écosystèmes côtiers français, et donc sur les services écosystémiques qu'ils sont susceptibles de fournir (filtration de l'eau, modification des conditions hydro-sédimentaires (Fonseca *et al.*, 1992 ; Hansen *et al.*, 2012), zone de reproduction et nourricerie (Orth *et al.*, 2006) etc.). Ainsi, le déclin prédit des herbiers de zostères pourrait avoir des impacts importants sur la biodiversité et/ou les services écosystémiques qu'ils fournissent.

On est donc en droit de se questionner sur les conséquences potentielles d'une telle perturbation de l'habitat que représente les herbiers de zostères (Dobson *et al.*, 2006 ; Thompson *et al.*, 2017). Est-ce que les herbiers de zostères seront remplacés s'ils disparaissent ? Par quelle(s) espèce(s) ? Les fonctions des herbiers dans l'écosystème seront-elles réalisées par le(s) espèce(s) remplaçante(s) ? (Marzloff *et al.*, (2016, 2018)) En effet, si on se réfère au forêt de kelp, cet habitat a décliné en quantité et en qualité dans de nombreuses régions du globe ces dernières décennies. Les forêts laminaires productives et structurellement complexes ont parfois décliné pour laisser place à de véritables déserts (Ling *et al.*, (2009, 2014) ; Marzloff *et al.*, 2013), ou parfois à des tapis d'algues filamenteuses (*turf* en anglais) dont la valeur écologique est quasiment nulle en termes de productivité, de complexité d'habitat et donc de biodiversité associées (Filbee-Dexter & Wernberg, 2018 ; Feehan *et al.*, 2019 ; Layton *et al.*, 2019). Ainsi, le déclin de cet habitat clé en zone tempérée affecte de manière irréversible la biodiversité fonctionnelle des zones anciennement colonisées par le kelp (Marzloff *et al.*, 2015 ; Johnson *et al.*, 2017 pour l'irréversibilité ; Ling, 2008 pour la perte de biodiversité). A l'instar des forêts de kelp, les herbiers seront-ils remplacés par des espèces similaires ou laisseront-ils place à des turfs ou des déserts subaquatiques ?

Dans un contexte global où les habitats sont affectés par de multiples pressions (pollutions herbicides, antifouling, eutrophisation, espèces invasives, maladies etc.), il faut rappeler que l'impact potentiel du changement climatique n'est qu'une goutte de plus dans l'océan. Il est dans notre intérêt de protéger les écosystèmes des perturbations et de prendre les mesures nécessaires pour supprimer ou réduire au maximum les impacts des activités humaines sur ces écosystèmes afin de maintenir ou atteindre un bon état écologique du milieu marin, et ainsi bénéficier de façon optimale des biens et services qu'il nous fournit. Toutefois, même si les pressions étaient réduites, la question de la réversibilité des impacts sur les herbiers est importante. Par exemple, à Arcachon, les herbiers ont du mal à revenir depuis les grosses canicules entre 2003 et 2006 (Cognat, 2020), alors que la pression a disparue.

Il faut rappeler qu'il est difficile d'appréhender les impacts cumulés des pressions puisque de la plupart sont difficiles à prédire et donc rarement prise en compte dans les modèles (pollution etc.). Il serait donc intéressant d'initier des travaux dans cette voie pour améliorer la compréhension des impacts et mettre en place des mesures de gestion plus efficaces.

Afin d'adopter des mesures de gestion efficaces pour préserver les herbiers à échelle locale, comme à l'échelle de la façade métropolitaine, il est également critique de mieux connaître l'écologie des herbiers, ainsi que de mieux caractériser leurs distributions spatiales dans le but de détecter les changements. Un autre aspect de mon étude rentre parfaitement dans cette démarche de mieux connaître les espèces pour mieux les gérer et mieux les protéger.

- **Description de la niche environnementale et la biogéographie des espèces**

- *Conditions hydrodynamiques*

Nos résultats mettent en avant l'importance des variables *Vagues*, *Courant* et *Fetch* dans la niche environnementale des herbiers de zostères, et plus particulièrement la gamme de valeurs de ces variables. Une forte exposition aux vagues et au courant pourrait exposer les herbiers à un risque d'arrachement des pieds non négligeable, mais également permettrait la mise en suspension des sédiments et limiterait donc l'accès à la lumière. Bien que les herbiers aient un effet tampon sur le courant et les vagues en les atténuant au sein du lit (Widdows *et al.*, 2008 ; Hansen et Reidenbach, 2012 ; Ganthly *et al.*, 2015), il semblerait que la colonisation par les herbiers dans des zones exposées soient difficiles, ce qui irait dans le sens des hypothèses avancées pour expliquer les difficultés des herbiers dans le bassin d'Arcachon à retrouver l'étendue qu'ils occupaient il y a plusieurs décennies. En effet, à la suite de la canicule de 2003, les herbiers ont décliné dans le bassin, le courant et les vagues se sont renforcés puisque les herbiers ne les atténuaient plus autant, et désormais, cette boucle rétro négative limiterait le retour des herbiers (Cognat, 2020).

- *Activité photosynthétique des herbiers*

Nos résultats mettent en avant l'importance de la variable *Lumière*. La présence des herbiers serait favorisée dans des zones où la turbidité est faible (peu de chlorophylle a, peu de nutriments, faible exposition aux vagues et courant responsable de la mise en suspension des sédiments) (Moore *et al.*, (1996, 1997)), avec peu de compétition pour la lumière (Short *et al.* 1995 ; Hauxwell *et al.*, 2001) et dans de faibles profondeurs (Dennison *et al.*, 1985).

Les herbiers ont besoin de lumière mais également de nutriments organiques et inorganiques pour leur développement. En milieu marin, les nutriments sont rarement limitants avec les apports des bassins versants, des estuaires etc. Cependant, la tolérance des herbiers aux nutriments semble limitée comme l'indiquent nos résultats, avec la mise en évidence de seuils au-delà desquels la tolérance de *Z. marina* à de fortes valeurs en nutriments est très limitée. Ces résultats sont confirmés par de nombreuses études qui montrent les effets négatifs de charges importantes en nutriments sur le développement et/ou la survie des herbiers (Burkholder *et al.*, 1992 ; Van Katwijk *et al.*, 1997 ; Brun *et al.*, 2008 ; La Nafie *et al.*, 2012 ; Govers *et al.*, 2014). Cela pourrait s'expliquer par des conditions redox réduites, par des concentrations potentiellement toxiques d'ammonium, par la stimulation de diverses formes d'algues (phytoplancton, algues épiphytes et macro-algues) qui concurrenceraient dans un premier temps les herbiers pour l'accès aux nutriments puis à la suite de leur développement, concurrenceraient efficacement les herbiers pour la lumière par ombrage (Short *et al.*, 1995 ; Hauxwell *et al.*, 2001). La concentration en chlorophylle-a peut être considérée comme un proxy de la concentration en nutriments inorganiques dissous favorables à la production primaire et donc au développement des herbiers, mais également comme un proxy de l'état d'eutrophisation des zones côtières. Il est donc également très intéressant d'observer que nos modèles captures cela en montrant un effet positif de faibles concentrations en chlorophylle-a ( $< 2 \mu\text{g.L}^{-1}$ ), et un effet négatif en cas d'excès (zone eutrophisée, panaches de fleuves très chargé en nutriments).

- *Propriétés physico-chimiques de l'eau*

D'après l'étude de Greve *et al.* (2003), les températures de l'eau élevées ont un effet important sur l'équilibre en oxygène interne des zostères. L'augmentation de la température de l'eau stimulerait davantage la respiration des plantes que la photosynthèse, favorisant des conditions anoxiques dans le méristème, même à la lumière, à des températures supérieures à 30 °C. Si on s'intéresse aux résultats de l'étude de Van Katwijk *et al.* (1997), de fortes températures de l'eau auraient un effet négatif sur la longueur des feuilles et la densité de pieds des herbiers, et favoriseraient leur nécrose. Cependant, ayant choisi de travailler avec la température moyenne annuelle de l'eau, et non avec les valeurs maximales saisonnières, nos résultats ne reflètent pas de tels effets de cette variable sur les herbiers.

Nos résultats indiquent également que les deux espèces ont une tolérance différente à la salinité. *Z. noltei* préférerait des conditions saumâtres (< 33 PSU), alors que *Z. marina* préférerait des conditions marines de salinité (> 34 PSU). Cela pourrait s'expliquer de plusieurs façons. D'après Hootsmans *et al.* (1987), les taux de germination seraient optimaux pour *Z. noltei* à des teneurs faibles de salinité, pour un optimal entre 10 et 20 ‰ pour *Z. marina*. Nejrup *et al.* (2008) indiquent que de faibles valeurs de salinité affectent la morphologie des pieds, la biomasse, la croissance et la capacité photosynthétique de *Z. marina* mais que les effets s'estompent au-delà de 10 ‰. Enfin, d'après Charpentier *et al.* (2005), le taux de sédimentation serait plus faible à faible salinité, ce qui limiterait la pénétration de la lumière. Ainsi, il semblerait que la salinité puisse avoir des effets sur les herbiers et leur distribution en ayant une influence sur l'accès à la lumière, le taux de germination ou encore la morphologie et la survie des plantes.

Certaines variables comme la température de l'eau ou la concentration en oxygène dissous se démarquent peu, et il pourrait être intéressant de refaire cette étude avec des valeurs extrêmes (min / max / P90 etc.) pour affiner nos résultats et la caractérisation des niches écologiques.

Notez qu'il est nécessaire d'être critique sur les variables utilisées, leur qualité et leur résolution. En effet, certaines ne sont pas bien résolues à la côte, et de nombreuses pressions locales potentiellement structurantes ne sont pas considérées dans cette étude (e.g. pollutions).

- *Bilan sur la niche écologique modélisée des deux espèces*

La niche écologique modélisée est donc en bonne adéquation avec les connaissances disponibles sur l'écologie des deux espèces de zostères. Ainsi, l'habitat potentiel modélisé de *Z. noltei* correspond à des zones intertidales abritées du vent, des vagues et du courant, peu turbides et saumâtres, correspondant aux baies lumineuses et abritées de la façade Atlantique et de Bretagne Nord. En revanche, l'habitat potentiel modélisé de *Z. marina* correspond à des zones côtières abritées, peu turbides et peu eutrophisée (nitrate, phosphate, chlorophylle a) principalement réparties le long du littoral en Bretagne Nord et Bretagne Sud.

Toutefois, les valeurs d'indice d'habitat potentiel prédites sont localement élevées dans des zones où la présence des herbiers est moins bien connue (perspectives de prospections), mais également dans des zones peu vraisemblables. On atteint donc ici les limites de cette étude. Une validation des cartes seraient donc nécessaire à la suite de cette étude.

- **Limites de l'étude**

Cette étude apporte ainsi des éléments de compréhension de l'écologie des herbiers de zostères, ainsi que des informations concernant l'aire de distribution potentielle des herbiers le long de la façade Manche-Atlantique. Les résultats numériques présentés ici sont toutefois considérés comme indicatif et les prédictions détaillées restent à interpréter avec précaution. Par exemple, la présence ou l'absence locale des herbiers ne peuvent être affirmées sur la seule base des prédictions des modèles. En effet les résultats des modèles sont par essence associés à divers biais et de l'incertitude (en lien avec la méthode choisie ou les données utilisées).

- *Biais liés à la méthodologie employée*

D'un point de vue méthodologique, deux sources de biais sont selon moi importantes à considérer. Tout d'abord, un biais relatif à l'autocorrélation spatiale dans nos données (Legendre, 1993 ; Jayet, 2001 ; Oliveau, 2010). En effet, j'ai choisi d'utiliser le jeu de données sur lequel j'avais corrigé partiellement l'autocorrélation spatiale. Cette correction partielle à l'aide d'un filtre à 5 km se justifiait de deux façons : la première est qu'on a considéré qu'au-delà 5 km, il y avait peu de chance que deux points d'observations proviennent du même herbier ; la deuxième était la volonté de constituer un jeu de données d'occurrences plus important pour modéliser l'habitat potentiel à l'échelle de la façade et ainsi avoir assez de points d'occurrences pour faire de la validation croisée. Cependant, il est certain que ce choix est discutable puisque la correction de l'autocorrélation spatiale est ici partielle et que l'autocorrélation spatiale résiduelle dans un modèle peut biaiser l'entraînement et donc les prédictions des modèles statistiques. Ainsi, si la collecte de nouvelles données permet d'augmenter la taille du jeu de données d'occurrences sur les herbiers de zostères, utiliser les distances recommandées par les variogrammes multivariés permettrait d'avoir une approche plus robuste puisque cela permettrait de gommer ces interactions spatiales.

Le deuxième point méthodologique sur lequel j'aimerais revenir est le calcul des seuils de discrétisation pour obtenir les prédictions de présence/absence. En effet, j'ai fait le choix de calculer mes seuils sur les données d'entraînement selon le critère *Sensitivity = Specificity* (Liu *et al.*, 2005 ; Jiménez-Valverde & Lobo, 2007 ; Nenzen & Araujo, 2011). Cependant, on observe un déséquilibre important entre les faux positifs et les faux négatifs sur les données de validation : le modèle penche vers la conservation, en surestimant les présences. Ainsi, le modèle d'ensemble prédit moins bien sur les données de validation (20%) que sur les données d'entraînement, et le critère *Sensitivity = Specificity* n'est pas validé, ce qui est un signe de surapprentissage (*overfitting*). Les cartes de présence/absence sont biaisées par cela et il serait donc intéressant dans de futures études de régler ce problème : voir la différence dans les prédictions en calculant les seuils sur les données test et non plus sur les données d'entraînement pourrait être une piste.

- *Biais liés aux données*

En ce qui concerne les biais liés aux données, on peut diviser cela en trois parties : les données de présence/absence, les données environnementales et les données utilisées pour l'analyse des conséquences du changement climatique.

(1) Données d'occurrences : Tout d'abord, discutons de la qualité du jeu de données initial de présence/pseudo-absences. Il est important de rappeler que les données recensées proviennent d'études locales diverses réalisées par différents organismes et que l'ensemble des relevés, à la fois surfaciques et ponctuels, couvrent une période allant de 1960 à 2016, et reposent sur des méthodes de prospection variées. Ainsi, je vois dans cette synthèse plusieurs sources de biais. Le premier biais est la date des données utilisées. En effet, il est possible que nos modèles aient surestimé la présence d'herbiers en se basant sur des points de présence qui ont peut-être disparus depuis. Ensuite, la présence de nombreuses sources peut également amener un biais dans la construction de la donnée. Deuxièmement, j'ai choisi de construire un jeu de données de pseudo-absences par échantillonnage aléatoire dans la zone infralittorale, faute de vraies d'absences. Il existe donc un biais dans la mesure où il est possible qu'il y ait eu des pseudo-absences au niveau d'herbiers non recensés. A l'avenir, il pourrait être pertinent de constituer un vrai jeu de présence/absence par exemple via un échantillonnage aléatoire dans toute la zone considérée et d'une caractérisation de chaque point par interprétation visuelle des orthophotographies du littoral, ce qui permettrait de pouvoir affiner cette étude sur la distribution et la compréhension de la niche environnementale des deux espèces.



(2) Données environnementales : De nombreuses limites sont également liées aux données environnementales utilisées pour les modèles de distribution à 500 m de résolution. Tout d'abord, j'ai choisi dans cette étude de travailler avec des moyennes pluriannuelles pour approcher un habitat potentiel moyen des espèces. Cependant, un tel choix entraîne l'absence de prise en compte des extrêmes saisonniers qui pourraient limiter nos présences. Par exemple, les concentrations minimales en oxygène dissous, les concentrations maximales en phosphate, nitrate, chlorophylle a ou encore les valeurs maximales de températures de l'eau pourraient impacter la survie des herbiers et donc la cartographie de l'habitat potentiel de chacune des espèces. Ainsi, il est important de rappeler ici que j'ai lissé les effets saisonniers et les extrêmes annuels en prenant en compte les moyennes pluriannuelles pour réaliser cette étude. Refaire cette étude en prenant en compte les variations et extrêmes saisonniers pourrait être intéressant pour affiner la description de la tolérance des herbiers à chacune des variables et ainsi mieux appréhender la niche écologique des deux espèces. Ensuite, je tiens à noter que la qualité des données environnementales a pu être biaisée par l'utilisation de la fonction resample pour calibrer des données de résolutions différentes, et que la qualité des variables ayant une résolution < 500 m a par conséquent été dégradée. De plus, j'ai associé à certains points la valeur de certaines variables environnementales issue du pixel le plus proche quand les variables étaient mal définies localement. Bien que je n'aie pas eu le choix, fautes de données, je suis conscient que cela puisse biaiser les relations statistiques au sein du modèle. Enfin, je tiens à discuter des données environnementales utilisées pour la modélisation de l'habitat potentiel de *Z. noltei*. En effet, faute de données en zone médiolittorale, j'ai utilisé les données les plus proches de l'estran i.e. le premier pixel de la zone infralittorale. Il est vrai que cette hypothèse est contestable dans la mesure où je réalise une prédiction de la distribution de cette espèce à partir des conditions les plus proches des points de présence/absence, et non des conditions au point. Cependant, ne pas le faire aurait conduit à ne pas étudier cette espèce, ce qui aurait été regrettable quand on observe que la qualité des modèles est bonne avec un taux d'erreur moyen de 14.7%.

(3) Données sur le changement climatique : Enfin, les données utilisées pour la partie changement climatique méritent également d'être discutées. Tout d'abord, pour cette dernière partie, le nombre de points d'occurrences était très limité puisque nous avons travaillé à une résolution de 11km. Ensuite, il est important de rappeler que nous n'avons pas pu avoir accès à une variable *Lumière* dans cette partie, ce qui est dommage quand on connaît l'importance relative que cette variable a eu dans nos précédents modèles. Dans de telles conditions, il n'est pas surprenant d'obtenir des modèles de moins bonnes qualités, avec des taux d'erreurs moyens plus élevés (18% pour *Z. marina* ; 27% pour *Z. noltei*) et des incohérences importantes comme le prouve par exemple l'expansion de *Z. noltei* en Manche-Est, ce qui semble relativement improbable. Dans cette partie, on capte une tendance intéressante sur l'impact potentiel du changement climatique sur l'habitat potentiel des herbiers, mais les résultats numériques restent toutefois largement améliorables. Pour avoir une approche plus robuste, il serait intéressant de prendre en compte l'ensemble de l'aire de distribution de *Z. marina* et *Z. noltei* en Europe, du sud du Portugal au nord de la Norvège, en passant par les herbiers d'Angleterre, d'Ecosse et du Danemark. De cette manière, on appréhenderait davantage les tolérances minimales et maximales des espèces et on donnerait davantage d'importance aux variations environnementales liées au changement climatique. Je pense en effet que ne pas prendre l'ensemble de l'aire de distribution et donc de la niche environnementale des espèces dans une étude d'impact du changement climatique peut biaiser de manière importante les prédictions, et conduire à des valeurs surprenantes comme c'est le cas des prédictions sur *Z. noltei* pour qui les modèles prédisent une expansion à l'horizon 2100.

- **Perspectives intéressantes**

Dans le cadre du projet MarHa, à la suite de mon stage, plusieurs perspectives intéressantes mériteraient une attention particulière :

- 1 – Valider formellement les cartes de prédictions spatiales construites à 500 m à partir d’outils de télédétection, acoustiques et des connaissances d’experts.

- 2 – Consolider un jeu de données plus complet, de meilleure qualité (pour combler les limites citées ci-dessus).

- 3 – Construire un modèle plus général de la distribution biogéographique des deux espèces (prenant en compte la totalité de leurs zones de répartition en Europe) afin d’affiner la compréhension de leur écologie, et la prédiction des conséquences biogéographiques du changement climatique.

- 4 – Construire des modèles ciblés à haute résolution afin de décrire la distribution spatiale des espèces à l’échelle de sites spécifiques (et non à l’échelle de la façade côtière Manche-Atlantique).

## **Conclusions**

Les herbiers de zostères ont un rôle écologique majeur dans les écosystèmes côtiers français. Les aspects de nurserie pour de nombreuses espèces marines, écologiquement importantes ou commercialement exploitées, montrent par exemple l’importance de cet habitat. Cependant, ces habitats sont depuis plusieurs décennies soumis à de multiples pressions, dont il est crucial de comprendre, prédire et gérer les impacts sur les herbiers. Ainsi, ce travail apporte des connaissances importantes sur l’écologie et la répartition spatiale des herbiers de zostères le long de la façade Manche-Atlantique et contribue ainsi à une meilleure compréhension de ces écosystèmes, de leur état écologique et de leurs réponses potentielles au changement climatique.

Les performances prédictives de nos modèles sont très bonnes, et ce pour les deux espèces *Z. noltei* et *Z. marina*. La niche environnementale modélisée permet de caractériser les conditions et zones côtières favorables à chacune des deux espèces : l’habitat potentiel de *Z. noltei* correspondrait à des zones intertidales abritées du vent, des vagues et du courant, peu turbides et saumâtres ; et, celui de *Z. marina* correspond à des zones côtières abritées, peu turbides et peu eutrophisée (nitrate, phosphate, chlorophylle a).

Cette étude ouvre de nombreuses perspectives de recherches et il est de notre ressort de toutes les explorer afin d’améliorer nos connaissances sur ces espèces et les protéger au mieux.

## Bibliographie

### • **Données utilisées**

#### ○ **Données sur les herbiers**

Bajjouk T. et al., 2020. Synthèse cartographique des herbiers de zostères. Produit multi-sources Ifremer-Life Marha (en préparation)

#### ○ **Données environnementales**

Ardhuin Fabrice (2013). Données issues du modèle ECO-MARS3D. <https://sextant.ifremer.fr/record/7ec6aa6e-2406-42d1-a798-5aa2984e6311/>

Boudiere Edwige, Maisondieu Christophe, Ardhuin Fabrice, Accensi Mickael, Pineau-Guillou Lucia, Lepesqueur Jeremy (2013). A suitable metocean hindcast database for the design of Marine energy converters. *International Journal of Marine Energy*, 3-4, e40-e52. Publisher's official version: <http://dx.doi.org/10.1016/j.ijome.2013.11.010>, Open Access version: <http://archimer.ifremer.fr/doc/00164/27524/>

Bricheno, L.M., Wolf, J., 2018. Future Wave Conditions of Europe, in Response to High-End Climate Change Scenarios. *Journal of Geophysical Research: Oceans* 123, 8762–8791. <https://doi.org/10.1029/2018JC013866>

Burrows, M., 2012. Influences of wave fetch, tidal flow, and ocean colour on subtidal rocky communities. *Mar. Ecol. Prog. Ser.* 445, 193–207. <https://doi.org/10.3354/meps09422>

Caillaud Matthieu, Petton Sebastien, Dumas Franck, Rochette Sebastien, Vasquez Mickael (2016). Rejeu hydrodynamique à 500 m de résolution avec le modèle MARS3D-AGRIF - Zone Manche-Gascogne. Ifremer. <https://doi.org/10.12770/3edee80f-5a3e-42f4-9427-9684073c87f5>

Copernicus Climate Change Service. « Marine biogeochemistry data for the Northwestern European shelf and Mediterranean Sea from 2006 up to 2100 derived from climate projections ». ECMWF, 2020. <https://doi.org/10.24381/CDS.DCC9295C>.

EMODnet Bathymetry Consortium, 2018. EMODnet Digital Bathymetry (DTM 2018). EMODnet Bathymetry Consortium. <https://doi.org/10.12770/18ff0d48-b203-4a65-94a9-5fd8b0ec35f6>

EMODnet Geology (2020). <https://www.emodnet-geology.eu/data-products/seabed-substrates/>

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, *Geosci. Model Dev.*, 9, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016. Données consultées le 3 juin 2020 sur. [https://www.worldclim.org/data/cmip6/cmip6\\_clim5m.html](https://www.worldclim.org/data/cmip6/cmip6_clim5m.html).

Fick, Stephen E., et Robert J. Hijmans. « WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas ». *International Journal of Climatology* 37, n° 12 (octobre 2017): 4302-15. <https://doi.org/10.1002/joc.5086>. Données consultées le 3 juin 2020 sur <https://www.worldclim.org/data/worldclim21.html>

Seity Y, Brousseau P, Malardel S, Hello G, Bénard P, Bouttier F. 2011. The AROME-France convective-scale operational model. *Mon. Weather Rev.* 139: 976–991.

Vasquez Mickael, Manca Eleonora, Inghilesi Roberto, Martin Simon, Agnesi Sabrina, Al Hamdani Ziad, Annunziatellis Aldo, Bekkby Trine, Pesch Roland, Askew Natalie, Bentes Luis, Castle Lewis, Doncheva Valentina, Drakopoulou Vivi, Gonçalves Jorge, Laamanen Leena, Lillis Helen, Loukaidi Valia, McGrath Fergal, Mo Giulia, Monteiro Pedro, Muresan Mihaela, O'Keeffe Eimear, Populus Jacques, Pinder Jordan, Ridgeway Amy, Sakellariou Dimitris, Simboura Mika, Teaca Adrian, Tempera Fernando, Todorova Valentina, Tunesi Leonardo, Virtanen Elina (2020). EUSeaMap 2019, A European broad-scale seabed habitat map, technical report. EASME/EMFF/2018/1.3.1.8/Lot2/SI2.810241–EMODnet Thematic Lot n° 2 – Seabed Habitats. <https://doi.org/10.13155/74782>

### • **Littérature**

Anselin L. et Bera A.K. (1998). “Spatial Dependence in Linear Regression Models with an Application to Spatial Econometrics”, in Ullah A. et Giles D.E.A. (eds.), *Handbook of Applied Economics Statistics*, Springer-Verlag, pp. 21-74.

Auby Isabelle, Bost Charles-André, Budzinski Helene, Dalloyau Sebastien, Desternes Amandine, Belles Angel, Trut Gilles, Plus Martin, Pere Caroline, Couzi Laurent, Feigne Claude, Steinmetz Julien (2011). Régression des herbiers de zostères dans le Bassin d'Arcachon : état des lieux et recherche des causes. RST/ODE/LER/AR/11.007. <https://archimer.ifremer.fr/doc/00054/16507/>

- Bajjouk Touria, Duchene Julie, Guillaumont Brigitte, Bernard Maud, Blanchard Michel, Derrien-Courtell Sandrine, Dubois Stanislas, Grall Jacques, Hamon Dominique, Hily Christian, Le Gal Aodren, Rigolet Carine, Rossi Nadège, Ledard Michel (2015). *Les fonds marins de Bretagne, un patrimoine remarquable : Connaître pour mieux agir*. Edition Ifremer-DREAL Bretagne, 152 p. <http://dx.doi.org/10.13155/42243>
- Bajjouk T., Blanchet A., Cordier C., De Casamajor M.N., Foveau A., Liabot P.O., Marzloff M. & Ouisse V., 2019. *Life Marha, Référentiel projet et plan prévisionnel pour les actions de l'Ifremer du volet côtier*. Version 2019. 37p.
- Bajjouk T., Blanchet A., Cordier C., De Casmajor M. N., Liabot P.O., Marzloff M., Ouisse V., Derrien L., Droual G., Le Bris A., Tempera F., 2020. *Projet Life Intégré Marha. Bilan d'activités du volet côtier. Rapport Ifremer de la phase 1- Réf. DYNECO/LEBCO/202003TB*. 94p.
- Bekkby, Trine, Eli Rinde, Lars Erikstad, Vegar Bakkestuen, Oddvar Longva, Ole Christensen, Martin Isæus, et Pål Erik Isachsen. « *Spatial Probability Modelling of Eelgrass (Zostera Marina) Distribution on the West Coast of Norway* ». *ICES Journal of Marine Science* 65, n° 7 (1 octobre 2008): 1093-1101. <https://doi.org/10.1093/icesjms/fsn095>.
- Breiman, L. *Random Forests*. *Machine Learning* 45, 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
- Brun FG, Olivé I, Malta E, Vergara JJ, Hernández I, Pérez-Lloréns JL (2008) *Increased vulnerability of Zostera noltii to stress caused by low light and elevated ammonium levels under phosphate deficiency*. *Mar Ecol Prog Ser* 365:67-75. <https://doi.org/10.3354/meps07512>
- Burdick, D.M., F.T. Short, and J. Wolf. 1993. *An index to assess and monitor the progression of wasting disease in eelgrass Zostera marina*. *Marine Ecology Progress Series* 94:83-90. <http://www.int-res.com/articles/meps/94/m094p083.pdf>
- Burkholder, JoAnn M, Katherine M Mason, et Howard B Glasgow. « *Water-Column Nitrate Enrichment Promotes Decline of Eelgrass Zostera Marina: Evidence from Seasonal Mesocosm Experiments* ». *Mar. Ecol. Prog. Ser.*, 1992, 16.
- Charpentier, A., P. Grillas, F. Lescuyer, E. Coulet, et I. Auby. « *Spatio-Temporal Dynamics of a Zostera Noltii Dominated Community over a Period of Fluctuating Salinity in a Shallow Lagoon, Southern France* ». *Estuarine, Coastal and Shelf Science* 64, n° 2 (1 août 2005): 307-15. <https://doi.org/10.1016/j.ecss.2005.02.024>.
- Chefaoui, Rosa M., et Jorge M. Lobo. « *Assessing the Effects of Pseudo-Absences on Predictive Distribution Model Performance* ». *Ecological Modelling* 210, n° 4 (10 février 2008): 478-86. <https://doi.org/10.1016/j.ecolmodel.2007.08.010>.
- Cognat, M. (2020). *Rôles des facteurs environnementaux et des interactions biomorphodynamiques sur l'évolution spatio-temporelle des herbiers de zostères dans une lagune mésotidale*. Ecole doctorale des sciences et environnements
- Dennison, William & Alberte, RS. (1985). *Role of daily light period in depth distribution of Zostera marina (eelgrass)*. *Marine Ecology-progress Series - MAR ECOL-PROGR SER*. 25. 51-61. 10.3354/meps025051.
- Dobson, Andrew, David Lodge, Jackie Alder, Graeme S. Cumming, Juan Keymer, Jacquie McGlade, Hal Mooney, et al. « *Habitat Loss, Trophic Collapse, and the Decline of Ecosystem Services* ». *Ecology* 87, n° 8 (2006): 1915-24. [https://doi.org/10.1890/0012-9658\(2006\)87\[1915:HLTCAT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[1915:HLTCAT]2.0.CO;2).
- Downie, Anna-Leena, Mikael von Numers, et Christoffer Boström. « *Influence of Model Selection on the Predicted Distribution of the Seagrass Zostera Marina* ». *Estuarine, Coastal and Shelf Science* 121-122 (avril 2013): 8-19. <https://doi.org/10.1016/j.ecss.2012.12.020>.
- Duarte, Carlos M, et Carina L Chiscano. « *Seagrass Biomass and Production: A Reassessment* ». *Aquatic Botany* 65, n° 1 (1 novembre 1999): 159-74. [https://doi.org/10.1016/S0304-3770\(99\)00038-8](https://doi.org/10.1016/S0304-3770(99)00038-8).
- Duarte C.M, Middelburg J.J., Caraco N.. *Major role of marine vegetation on the oceanic carbon cycle*. *Biogeosciences Discussions, European Geosciences Union*, 2004, 1 (1), pp.659-679. hal-00297772
- Ehlers, A, B Worm, et Tbh Reusch. « *Importance of Genetic Diversity in Eelgrass Zostera Marina for Its Resilience to Global Warming* ». *Marine Ecology Progress Series* 355 (26 février 2008): 1-7. <https://doi.org/10.3354/meps07369>.
- Feehan, Colette J., Sean P. Grace, et Carla A. Narvaez. « *Ecological Feedbacks Stabilize a Turf-Dominated Ecosystem at the Southern Extent of Kelp Forests in the Northwest Atlantic* ». *Scientific Reports* 9, n° 1 (8 mai 2019): 7078. <https://doi.org/10.1038/s41598-019-43536-5>.
- Filbee-Dexter, Karen, et Thomas Wernberg. « *Rise of Turfs: A New Battlefield for Globally Declining Kelp Forests* ». *BioScience* 68, n° 2 (1 février 2018): 64-76. <https://doi.org/10.1093/biosci/bix147>.

- Fonseca, Mark S., et Jennifer A. Cahalan. « A Preliminary Evaluation of Wave Attenuation by Four Species of Seagrass ». *Estuarine, Coastal and Shelf Science* 35, n° 6 (1 décembre 1992): 565-76. [https://doi.org/10.1016/S0272-7714\(05\)80039-3](https://doi.org/10.1016/S0272-7714(05)80039-3).
- Foucart, Thierry. « Colinéarité et régression linéaire ». *Mathématiques et sciences humaines*, n° 173 (1 mars 2006). <https://doi.org/10.4000/msh.2963>.
- Fourqurean, James W., Carlos M. Duarte, Hilary Kennedy, Núria Marbà, Marianne Holmer, Miguel Angel Mateo, Eugenia T. Apostolaki, et al. « Seagrass Ecosystems as a Globally Significant Carbon Stock ». *Nature Geoscience* 5, n° 7 (juillet 2012): 505-9. <https://doi.org/10.1038/ngeo1477>.
- Friedman, Jerome H. « Stochastic Gradient Boosting ». *Computational Statistics & Data Analysis* 38, n° 4 (février 2002): 367-78. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2).
- Ganthy, Florian, Laura Soissons, Pierre-Guy Sauriau, Romaric Verney, et Aldo Sottolichio. « Effects of Short Flexible Seagrass *Zostera Noltei* on Flow, Erosion and Deposition Processes Determined Using Flume Experiments ». *Sedimentology* 62, n° 4 (2015): 997-1023. <https://doi.org/10.1111/sed.12170>.
- Govers, Laura L., Jan H.F. de Brouwer, Wouter Suykerbuyk, Tjeerd J. Bouma, Leon P.M. Lamers, Alfons J.P. Smolders, et Marieke M. van Katwijk. « Toxic Effects of Increased Sediment Nutrient and Organic Matter Loading on the Seagrass *Zostera Noltii* ». *Aquatic Toxicology* 155 (octobre 2014): 253-60. <https://doi.org/10.1016/j.aquatox.2014.07.005>.
- Greve, Tina Maria, Jens Borum, et Ole Pedersen. « Meristematic Oxygen Variability in Eelgrass (*Zostera Marina*) ». *Limnology and Oceanography* 48, n° 1 (2003): 210-16. <https://doi.org/10.4319/lo.2003.48.1.0210>.
- Guisan, Antoine, Wilfried Thuiller, et Niklaus E. Zimmermann. *Habitat Suitability and Distribution Models: With Applications in R*. Cambridge University Press, 2017.
- Halpern, Benjamin S., Shaun Walbridge, Kimberly A. Selkoe, Carrie V. Kappel, Fiorenza Micheli, Caterina D'Agrosa, John F. Bruno, et al. « A Global Map of Human Impact on Marine Ecosystems ». *Science* 319, n° 5865 (15 février 2008): 948-52. <https://doi.org/10.1126/science.1149345>.
- Hansen, Jer, et Ma Reidenbach. « Wave and Tidally Driven Flows in Eelgrass Beds and Their Effect on Sediment Suspension ». *Marine Ecology Progress Series* 448 (23 février 2012): 271-87. <https://doi.org/10.3354/meps09225>.
- Hansen, Jennifer C. R., et Matthew A. Reidenbach. « Seasonal Growth and Senescence of a *Zostera Marina* Seagrass Meadow Alters Wave-Dominated Flow and Sediment Suspension Within a Coastal Bay ». *Estuaries and Coasts* 36, n° 6 (novembre 2013): 1099-1114. <https://doi.org/10.1007/s12237-013-9620-5>.
- Hauxwell, Jennifer, Just Cebrián, Christopher Furlong, et Ivan Valiela. « Macroalgal Canopies Contribute to Eelgrass (*Zostera Marina*) Decline in Temperate Estuarine Ecosystems ». *Ecology* 82, n° 4 (2001): 1007-22. [https://doi.org/10.1890/0012-9658\(2001\)082\[1007:MCCTEZ\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2001)082[1007:MCCTEZ]2.0.CO;2).
- Hily C., Gacé N. (2004). *Impact de la pêche à pied sur les peuplements et les habitats de l'estran : cas des herbiers de zostères marines*. In *Contribution à la gestion et à la conservation des espaces marins insulaires protégés (Manche - Atlantique) : les activités de pêche à pied et de plongée ; impacts sur la biodiversité et mise au point d'outils d'évaluation* (coord. C. Hily). Programme de recherche Espaces protégés (Minist. Env.). Rapport Univ. Bret. Occ. Brest : Chap 3 , 34pp.
- Hootsmans, M. J. M., J. E. Vermaat, et W. Van Vierssen. « Seed-Bank Development, Germination and Early Seedling Survival of Two Seagrass Species from The Netherlands: *Zostera Marina* L. and *Zostera Noltii* Hornem ». *Aquatic Botany* 28, n° 3 (1 août 1987): 275-85. [https://doi.org/10.1016/0304-3770\(87\)90005-2](https://doi.org/10.1016/0304-3770(87)90005-2).
- Hubert Jayet. *Econométrie et données spatiales - Une introduction à la pratique*. Cahiers d'Economie et de Sociologie Rurales, INRA Editions, 2001, 58-59, pp.105-129. [ffhal-01200924](https://hal-01200924).
- Jiménez-Valverde, Alberto, et Jorge M. Lobo. « Threshold Criteria for Conversion of Probability of Species Presence to Either-or Presence-Absence ». *Acta Oecologica* 31, n° 3 (1 mai 2007): 361-69. <https://doi.org/10.1016/j.actao.2007.02.001>.
- Johnson, C.R., Chabot, R.H., Marzloff, M.P. & Wotherspoon, S. (2017). *Knowing when (not) to attempt ecological restoration*. *Restoration Ecology*, 25, 140-147.
- Jones, Clive G., John H. Lawton, et Moshe Shachak. « Positive and Negative Effects of Organisms as Physical Ecosystem Engineers ». *Ecology* 78, n° 7 (1997): 1946-57. [https://doi.org/10.1890/0012-9658\(1997\)078\[1946:PANEEO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[1946:PANEEO]2.0.CO;2).
- Jouffray JB et al. 2019 *Parsing human and biophysical drivers of coral reef regimes*. *Proc. R. Soc. B* 286: 20182544.

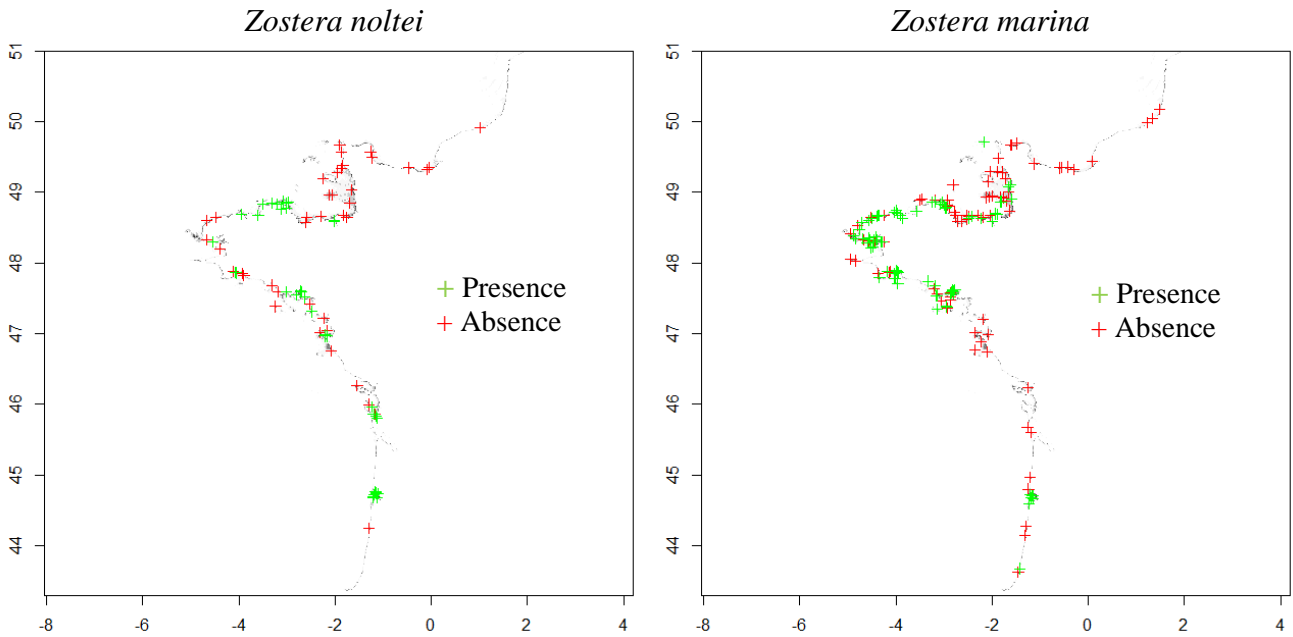
- Katwijk, Mm van, Lht Vergeer, Ghw Schmitz, et Jgm Roelofs. « Ammonium Toxicity in Eelgrass *Zostera Marina* ». *Marine Ecology Progress Series* 157 (1997): 159-73. <https://doi.org/10.3354/meps157159>.
- Kenworthy, W. Judson, et Mark Fonseca. « Reciprocal Transplant of the Seagrass *Zostera Marina* L. Effect of Substrate on Growth ». *Aquaculture*, Volume 12, Issue 3, November 1977, Pages 197-213.
- Layton, Cayne, Matthew J. Cameron, Victor Shelamoff, Pamela A. Fernández, Damon Britton, Catriona L. Hurd, Jeffrey T. Wright, et Craig R. Johnson. « Chemical Microenvironments within Macroalgal Assemblages: Implications for the Inhibition of Kelp Recruitment by Turf Algae ». *Limnology and Oceanography* 64, n° 4 (2019): 1600-1613. <https://doi.org/10.1002/lno.11138>.
- Legendre, P. (1993). *Spatial Autocorrelation: Trouble or New Paradigm?* *Ecology*, 74, 1659-1673.
- Ling, S.D. (2008). *Range expansion of a habitat-modifying species leads to loss of taxonomic diversity: a new and impoverished reef state.* *Oecologia*, 156, 883-894.
- Ling, S.D., Johnson, C.R., Frusher, S.D. & Ridgway, K.R. (2009). *Overfishing reduces resilience of kelp beds to climate-driven catastrophic phase shift.* *Proc. Natl. Acad. Sci. U. S. A.*, 106, 22341-22345.
- Ling, S.D., Scheibling, R.E., Rassweiler, A., Johnson, C.R., Shears, N., Connell, S.D. et al. (2014). *Global regime shift dynamics of catastrophic sea urchin overgrazing.* *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370.
- Liu, Canran, Pam M. Berry, Terence P. Dawson, et Richard G. Pearson. « Selecting Thresholds of Occurrence in the Prediction of Species Distributions ». *Ecography* 28, n° 3 (2005): 385-93. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>.
- Lotze, Heike K., Hunter S. Lenihan, Bruce J. Bourque, Roger H. Bradbury, Richard G. Cooke, Matthew C. Kay, Susan M. Kidwell, Michael X. Kirby, Charles H. Peterson, et Jeremy B. C. Jackson. « Depletion, Degradation, and Recovery Potential of Estuaries and Coastal Seas ». *Science* 312, n° 5781 (23 juin 2006): 1806-9. <https://doi.org/10.1126/science.1128035>.
- Marzloff, M.P., Johnson, C.R., Little, L.R., Soulié, J.-C., Ling, S.D. & Frusher, S.D. (2013). *Sensitivity analysis and pattern-oriented validation of TRITON, a model with alternative community states: Insights on temperate rocky reefs dynamics.* *Ecological Modelling*, 258, 16-32.
- Marzloff, M.P., Little, L.R. & Johnson, C.R. (2015). *Building Resilience Against Climate-Driven Shifts in a Temperate Reef System: Staying Away from Context-Dependent Ecological Thresholds.* *Ecosystems*, 19, 1-15.
- Marzloff, M.P., Melbourne-Thomas, J., Hamon, K.G., Hoshino, E., Jennings, S., van Putten, I.E. et al. (2016). *Modelling marine community responses to climate-driven species redistribution to guide monitoring and adaptive ecosystem-based management.* *Global Change Biology*, 22, 2462-2474.
- Marzloff, M.P., Oliver, E.C.J., Barrett, N.S., Holbrook, N.J., James, L., Wotherspoon, S.J. et al. (2018). *Differential vulnerability to climate change yields novel deep-reef communities.* *Nat. Clim. Chang.*, 8, 873-878.
- McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*. 2nd Edition, Chapman and Hall/CRC, Washington, DC. <http://dx.doi.org/10.1007/978-1-4899-3242-6>
- Moore, Ka, Ha Neckles, et Rj Orth. « *Zostera Marina* (Eelgrass) Growth and Survival along a Gradient of Nutrients and Turbidity in the Lower Chesapeake Bay ». *Marine Ecology Progress Series* 142 (1996): 247-59. <https://doi.org/10.3354/meps142247>.
- Moore, Kenneth A, Richard L Wetzel, et Robert J Orth. « *Seasonal Pulses of Turbidity and Their Relations to Eelgrass (Zostera Marina L.) Survival in an Estuary* ». *Journal of Experimental Marine Biology and Ecology* 215, n° 1 (15 juillet 1997): 115-34. [https://doi.org/10.1016/S0022-0981\(96\)02774-8](https://doi.org/10.1016/S0022-0981(96)02774-8).
- Nafie, Yayu A. La, Carmen B. de los Santos, Fernando G. Brun, Marieke M. van Katwijk, et Tjeerd J. Bouma. « *Waves and High Nutrient Loads Jointly Decrease Survival and Separately Affect Morphological and Biomechanical Properties in the Seagrass Zostera Noltii* ». *Limnology and Oceanography* 57, n° 6 (2012): 1664-72. <https://doi.org/10.4319/lo.2012.57.6.1664>.
- Nejrup, Lars & Pedersen, Morten. (2008). *Effects of salinity and water temperature on the ecological performance of Zostera marina.* *Aquatic Botany*. 88. 239-246. [10.1016/j.aquabot.2007.10.006](https://doi.org/10.1016/j.aquabot.2007.10.006).
- Nelder, John; Wedderburn, Robert (1972). "Generalized Linear Models". *Journal of the Royal Statistical Society. Series A (General)*. Blackwell Publishing. 135 (3): 370–384.
- Nenzén, H. K., et M. B. Araújo. « *Choice of Threshold Alters Projections of Species Range Shifts under Climate Change* ». *Ecological Modelling* 222, n° 18 (24 septembre 2011): 3346-54. <https://doi.org/10.1016/j.ecolmodel.2011.07.011>.
- Oliveau, Sébastien. « *Autocorrélation spatiale : leçons du changement d'échelle* ». *LEspace géographique* Vol. 39, n° 1 (17 mars 2010): 51-64.
- Oliver, E.C.J., Donat, M.G., Burrows, M.T., Moore, P.J., Smale, D.A., Alexander, L.V. et al. (2018). *Longer and more frequent marine heatwaves over the past century.* *Nature Communications*, 9, 1324.

- Orth, Robert J., Tim J. B. Carruthers, William C. Dennison, Carlos M. Duarte, James W. Fourqurean, Kenneth L. Heck, A. Randall Hughes, et al. « A Global Crisis for Seagrass Ecosystems ». *BioScience* 56, n° 12 (1 décembre 2006): 987-96. [https://doi.org/10.1641/0006-3568\(2006\)56\[987:AGCFSE\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)56[987:AGCFSE]2.0.CO;2).
- Peirano, Andrea, Valentina Damasso, Monica Montefalcone, Carla Morri, et Carlo Nike Bianchi. « Effects of Climate, Invasive Species and Anthropogenic Impacts on the Growth of the Seagrass *Posidonia Oceanica* (L.) Delile in Liguria (NW Mediterranean Sea) ». *Marine Pollution Bulletin* 50, n° 8 (1 août 2005): 817-22. <https://doi.org/10.1016/j.marpolbul.2005.02.011>.
- Phillips, Steven J., Miroslav Dudík, Jane Elith, Catherine H. Graham, Anthony Lehmann, John Leathwick, et Simon Ferrier. « Sample Selection Bias and Presence-Only Distribution Models: Implications for Background and Pseudo-Absence Data ». *Ecological Applications* 19, n° 1 (2009): 181-97. <https://doi.org/10.1890/07-2153.1>.
- Populus J, Vasquez M, Albrecht J, Manca E, Agnesi S, et al., 2017. EUSeaMap, a European broad-scale seabed habitat map. 174p. DOI: 10.13155/49975
- Schubert, Pr, W Hukriede, R Karez, et Tbh Reusch. « Mapping and Modeling Eelgrass *Zostera Marina* Distribution in the Western Baltic Sea ». *Marine Ecology Progress Series* 522 (2 mars 2015): 79-95. <https://doi.org/10.3354/meps11133>.
- Short, Frederick T., David M. Burdick, et James E. Kaldy. « Mesocosm Experiments Quantify the Effects of Eutrophication on Eelgrass, *Zostera Marina* ». *Limnology and Oceanography* 40, n° 4 (1995): 740-49. <https://doi.org/10.4319/lo.1995.40.4.0740>.
- Thompson, Patrick L., Bronwyn Rayfield, et Andrew Gonzalez. « Loss of Habitat and Connectivity Erodes Species Diversity, Ecosystem Functioning, and Stability in Metacommunity Networks ». *Ecography* 40, n° 1 (2017): 98-108. <https://doi.org/10.1111/ecog.02558>.
- Turner, R. K., S. Subak, et W. N. Adger. « Pressures, Trends, and Impacts in Coastal Zones: Interactions between Socioeconomic and Natural Systems ». *Environmental Management* 20, n° 2 (mars 1996): 159-73. <https://doi.org/10.1007/BF01204001>.
- Valle, Mireia, Guillem Chust, Andrea del Campo, Mary S. Wisz, Steffen M. Olsen, Joxe Mikel Garmendia, et Ángel Borja. « Projecting Future Distribution of the Seagrass *Zostera Noltii* under Global Warming and Sea Level Rise ». *Biological Conservation* 170 (février 2014): 74-85. <https://doi.org/10.1016/j.biocon.2013.12.017>.
- VanDerWal, Jeremy, Luke P. Shoo, Catherine Graham, et Stephen E. Williams. « Selecting Pseudo-Absence Data for Presence-Only Distribution Modeling: How Far Should You Stray from What You Know? » *Ecological Modelling* 220, n° 4 (24 février 2009): 589-94. <https://doi.org/10.1016/j.ecolmodel.2008.11.010>.
- Villeneuve, Valérie, Stéphane Légaré, Jean Painchaud, et Warwick Vincent. « Dynamique et modélisation de l'oxygène dissous en rivière ». *Revue des sciences de l'eau* 19, n° 4 (17 janvier 2007): 259-74. <https://doi.org/10.7202/014414ar>.
- Waycott, Michelle, Carlos M. Duarte, Tim J. B. Carruthers, Robert J. Orth, William C. Dennison, Suzanne Olyarnik, Ainsley Calladine, et al. « Accelerating Loss of Seagrasses across the Globe Threatens Coastal Ecosystems ». *Proceedings of the National Academy of Sciences* 106, n° 30 (28 juillet 2009): 12377-81. <https://doi.org/10.1073/pnas.0905620106>.
- Widdows, J, Nd Pope, Md Brinsley, H Asmus, et Rm Asmus. « Effects of Seagrass Beds (*Zostera Noltii* and *Z. Marina*) on near-Bed Hydrodynamics and Sediment Resuspension ». *Marine Ecology Progress Series* 358 (21 avril 2008): 125-36. <https://doi.org/10.3354/meps07338>.
- Wu, P.P.-Y., McMahon, K., Rasheed, M.A., Kendrick, G.A., York, P.H., Chartrand, K. et al. (2017). Managing seagrass resilience under cumulative dredging affecting light: Predicting risk using dynamic Bayesian networks. *J. Appl. Ecol.*, 55, 1339-1350.

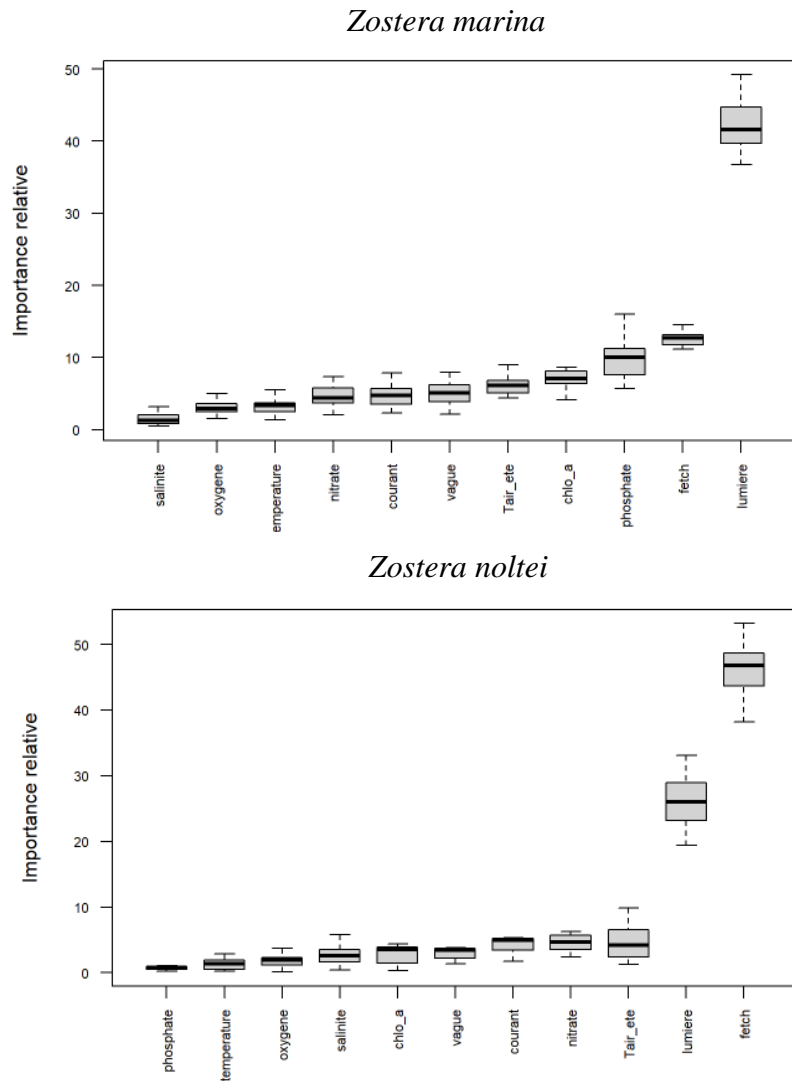


# Annexes

**Annexe 1 :** Carte de distribution du jeu de données de validation (20%) à 500 m

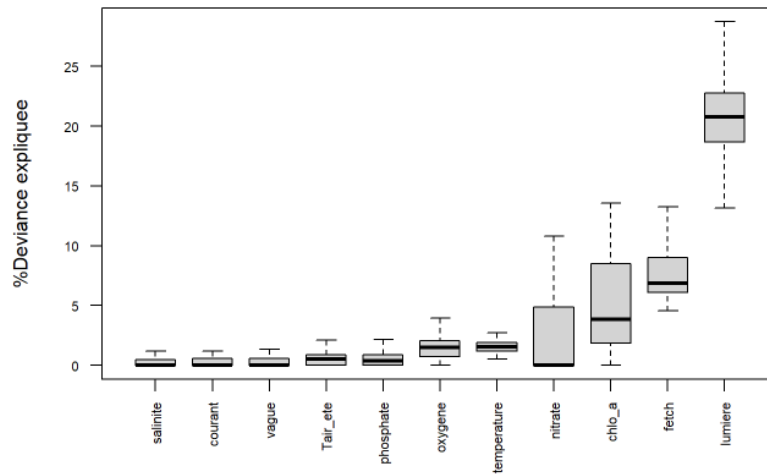


**Annexe 2 :** Représentation de l'importance relative des variables dans les modèles BRT

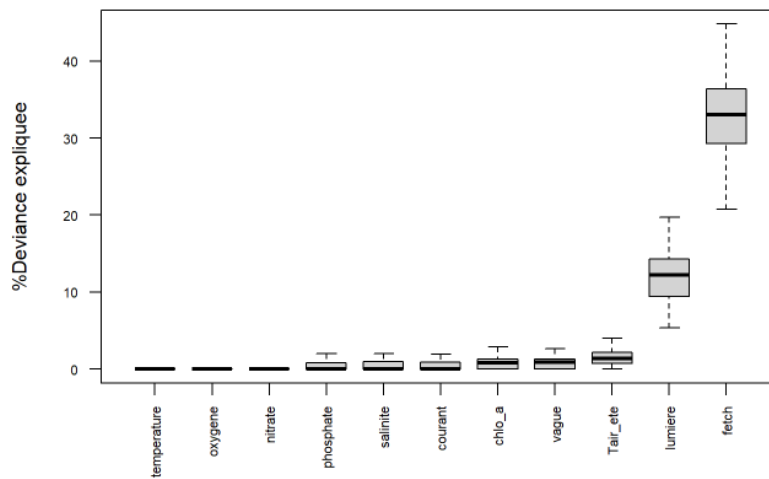


**Annexe 3 :** Représentation de l'importance relative des variables dans les modèles GLM

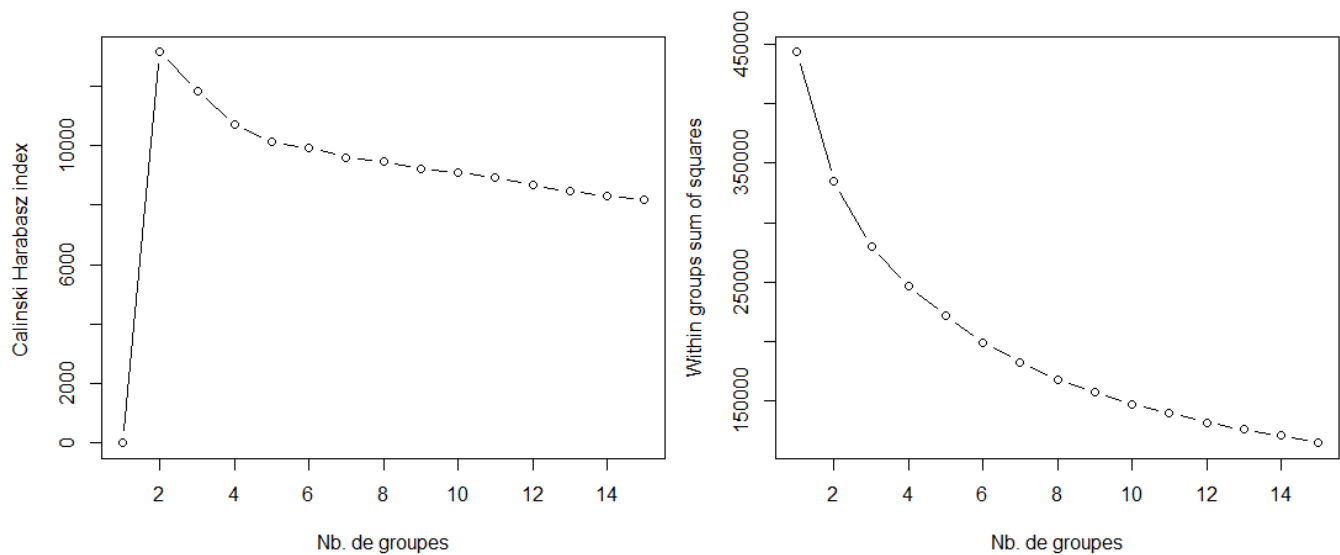
*Zostera marina*



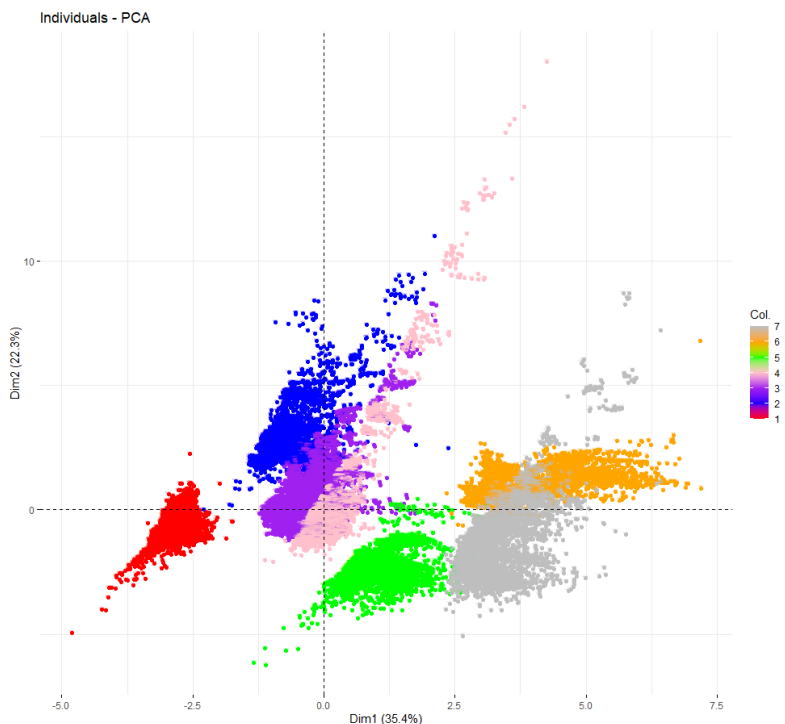
*Zostera noltei*



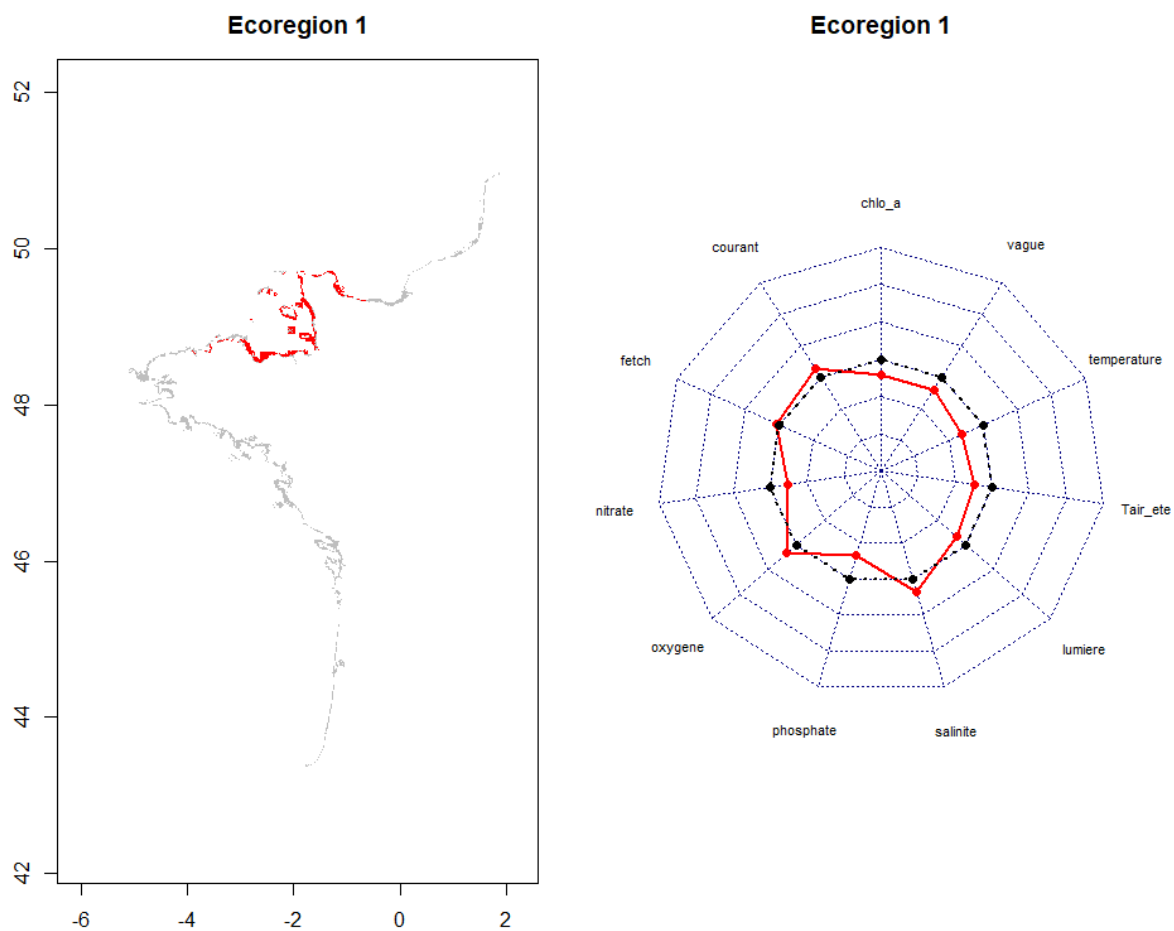
**Annexe 4 :** Méthodes de calcul du nombre de clusters optimal dans un partitionnement de type k-means : Indice de Calinski Harabasz (à gauche) ; méthode du coude (à droite)



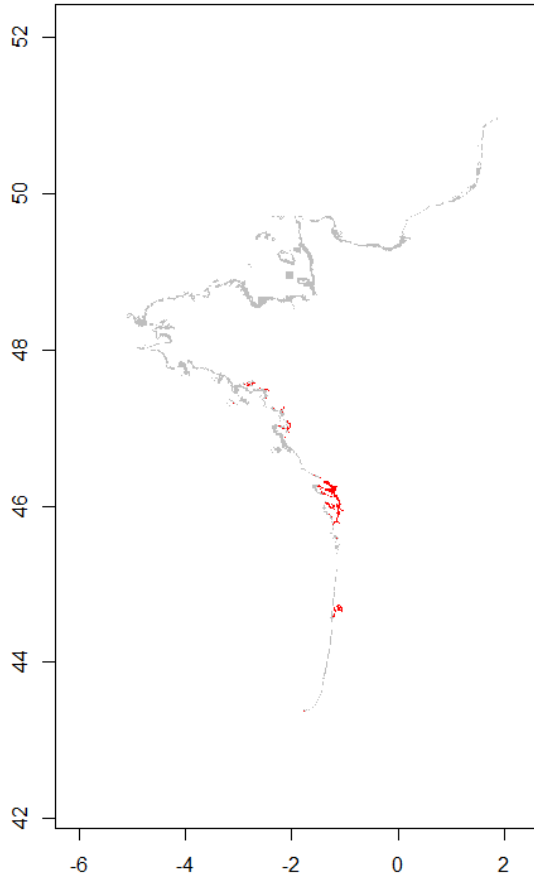
**Annexe 5 :** ACP sur les clusters issus du partitionnement



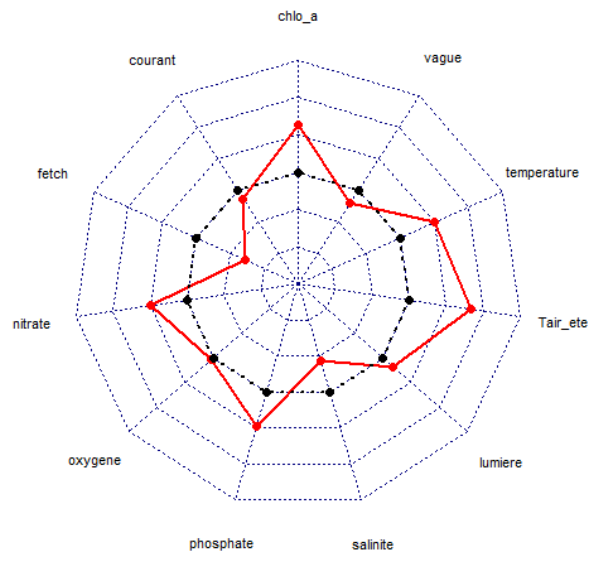
**Annexe 6 :** Présentation géographique des écorégions et de leur profil environnemental exprimé en valeur relative au travers de radarplot



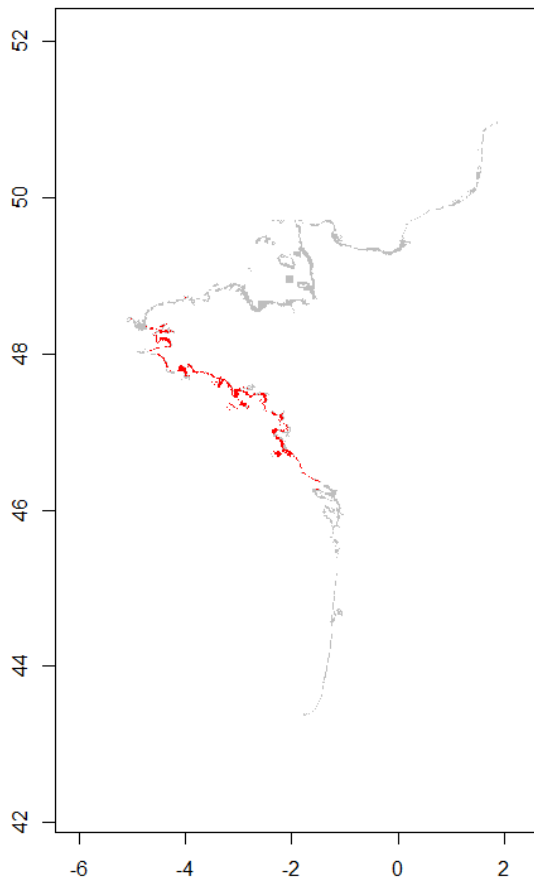
**Ecoregion 2**



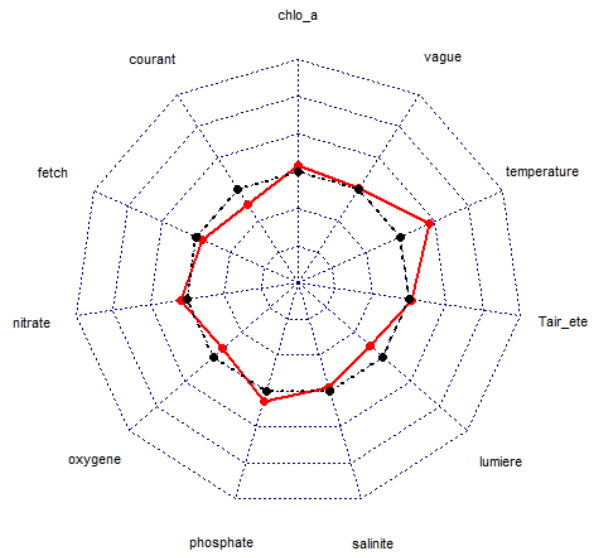
**Ecoregion 2**



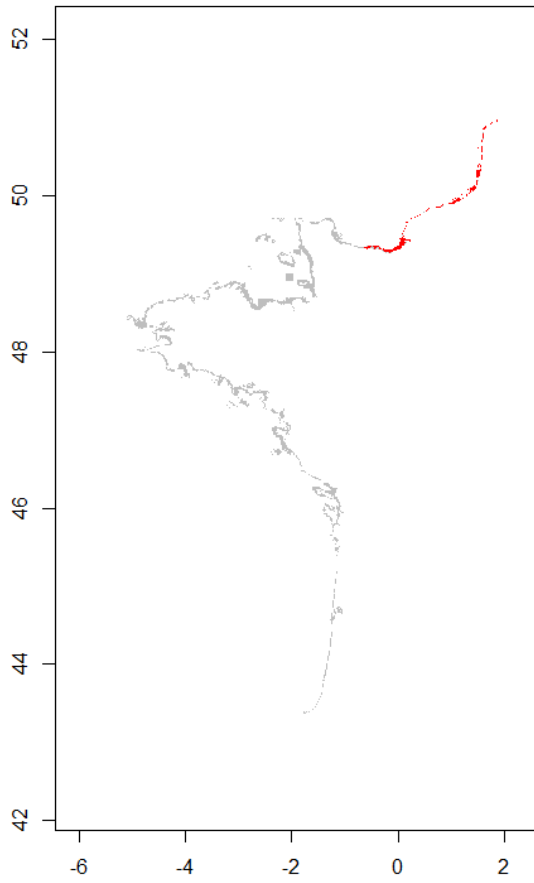
**Ecoregion 3**



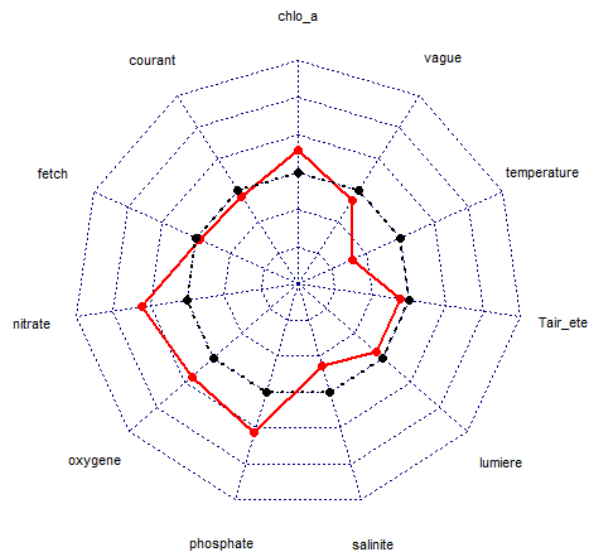
**Ecoregion 3**



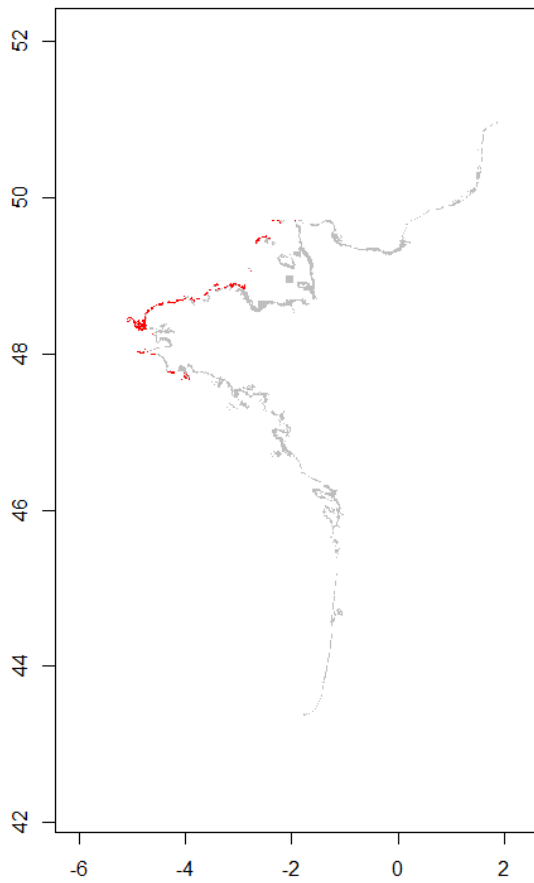
**Ecoregion 4**



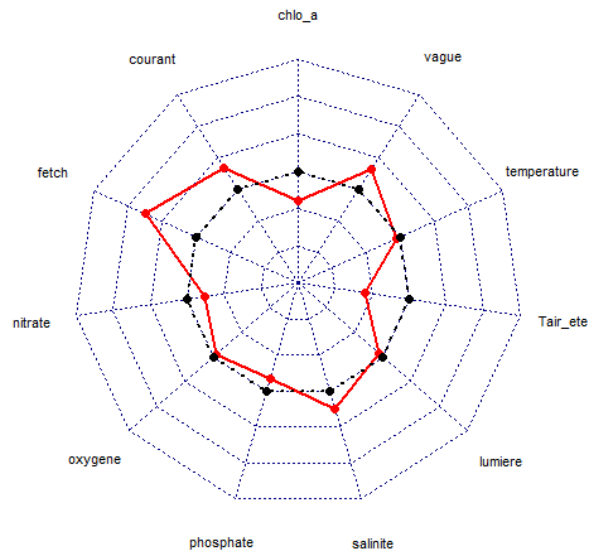
**Ecoregions 4**



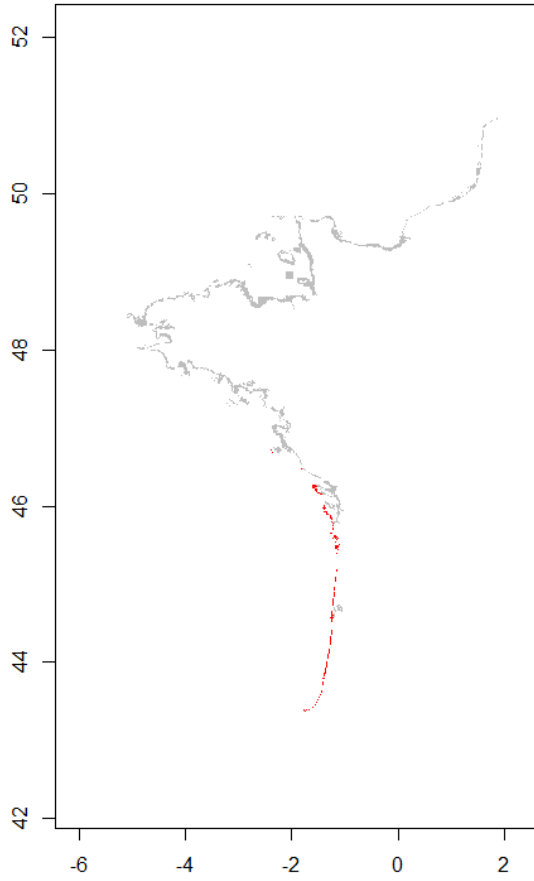
**Ecoregion 5**



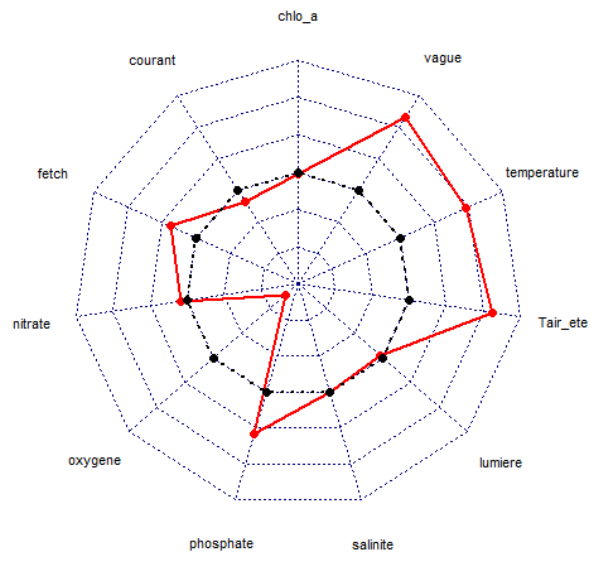
**Ecoregions 5**



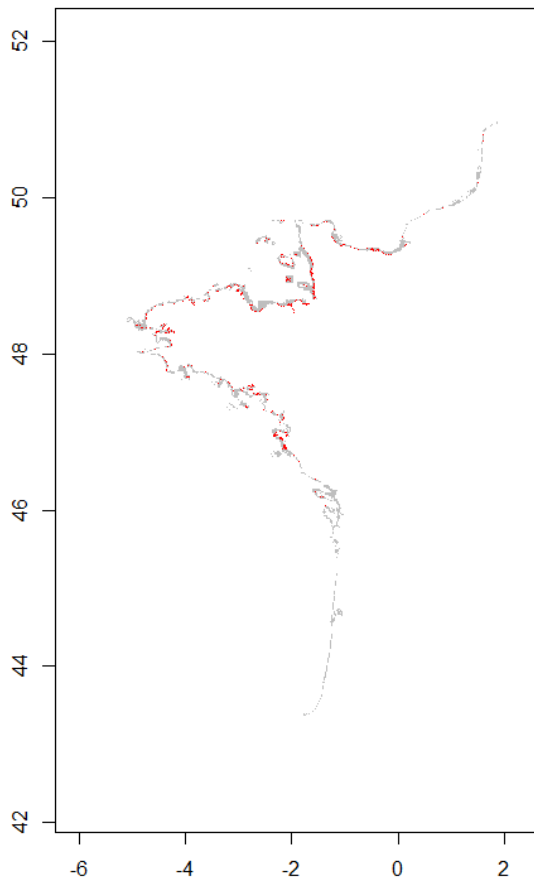
**Ecoregion 6**



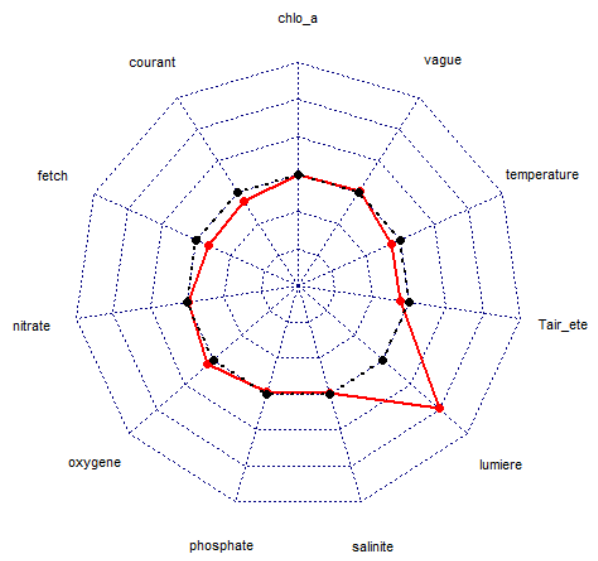
**Ecoregion 6**



**Ecoregion 7**



**Ecoregion 7**



**Annexe 7** : Tableaux bilan des prédictions sur le changement climatique

	Changement de superficie de <i>Z. marina</i> (%) avec le RCP4.5			Changement de superficie de <i>Z. marina</i> (%) avec le RCP8.5		
	Present / 2040-2050	2040-2050 / 2090-2100	Present / 2090-2100	Present / 2040-2050	2040-2050 / 2090-2100	Present / 2090-2100
Atlantique Sud	-100%	0%	<b>-100%</b>	-100%	0%	<b>-100%</b>
Atlantique Centre	-100%	0%	<b>-100%</b>	-100%	0%	<b>-100%</b>
Bretagne Sud	-12%	-31.2%	<b>-40%</b>	-22%	-69.2%	<b>-76%</b>
Bretagne Nord	+14.3%	-43.8%	<b>-35.7%</b>	-1.4%	-82.6%	<b>-82.9%</b>
Manche	+100%	-100%	<b>-100%</b>	0%	-100%	<b>-100%</b>
<b>Façade</b>	<b>-1.6%</b>	<b>-39.8%</b>	<b>-41.4%</b>	<b>-14.8%</b>	<b>-66.4%</b>	<b>-81.2%</b>

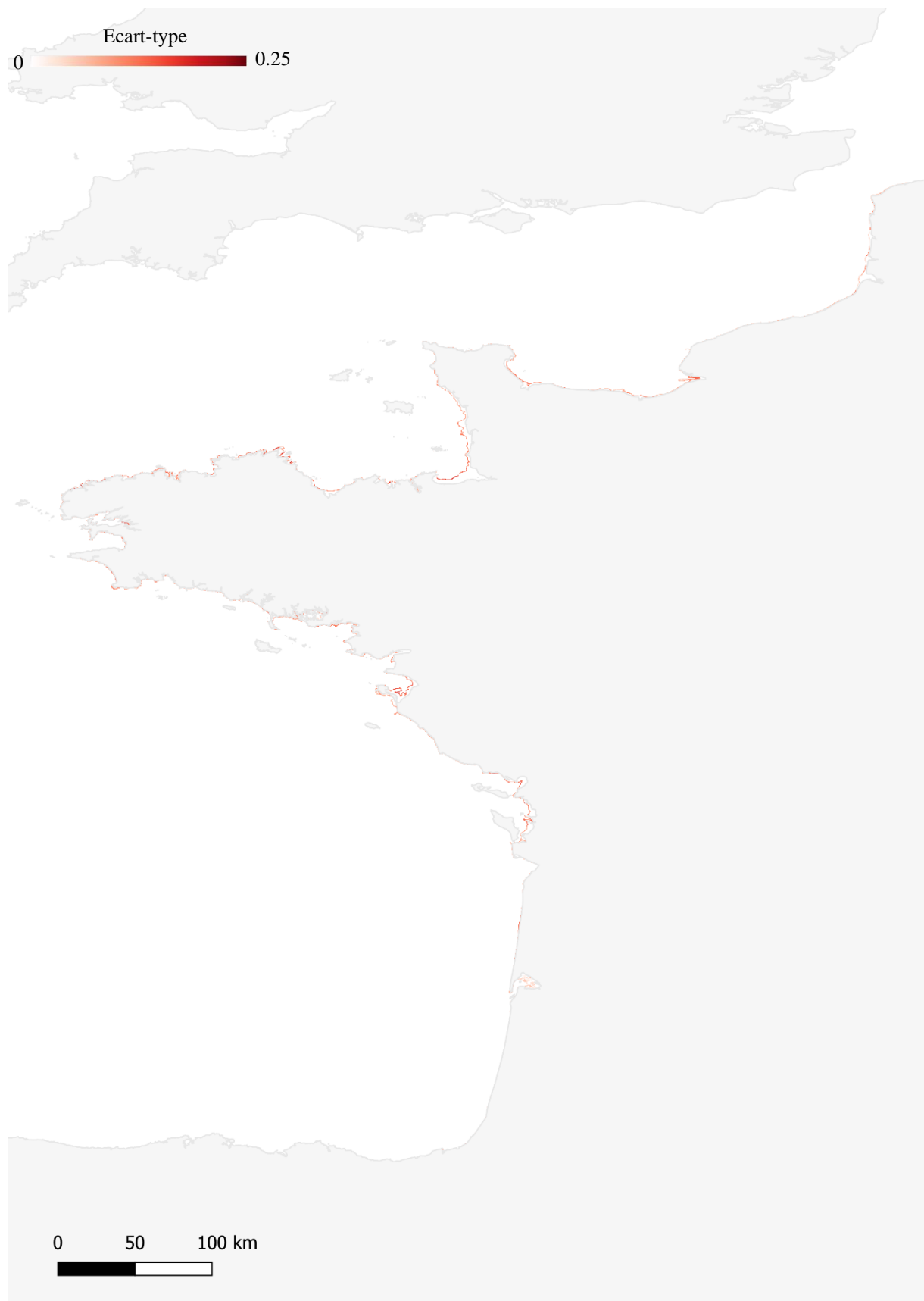
	Changement de superficie de <i>Z. noltei</i> (%) avec le RCP4.5			Changement de superficie de <i>Z. noltei</i> (%) avec le RCP8.5		
	Present / 2040-2050	2040-2050 / 2090-2100	Present / 2090-2100	Present / 2040-2050	2040-2050 / 2090-2100	Present / 2090-2100
Atlantique Sud	0%	0%	<b>0%</b>	0%	+20%	<b>20%</b>
Atlantique Centre	-15.4%	-4.5%	<b>-19.2%</b>	-11.5%	0%	<b>-11.5%</b>
Bretagne Sud	-3.8%	0%	<b>-3.8%</b>	-3.8%	0%	<b>-3.8%</b>
Bretagne Nord	43.8%	+4.3%	<b>50%</b>	+59.4%	-5.9%	<b>+50%</b>
Manche	+700%	+371.4%	<b>+2600%</b>	+1900%	-26.3%	<b>+1400%</b>
<b>Façade</b>	<b>+16.7%</b>	<b>+19%</b>	<b>+38.9%</b>	<b>36.7%</b>	<b>-5.7%</b>	<b>28.9%</b>



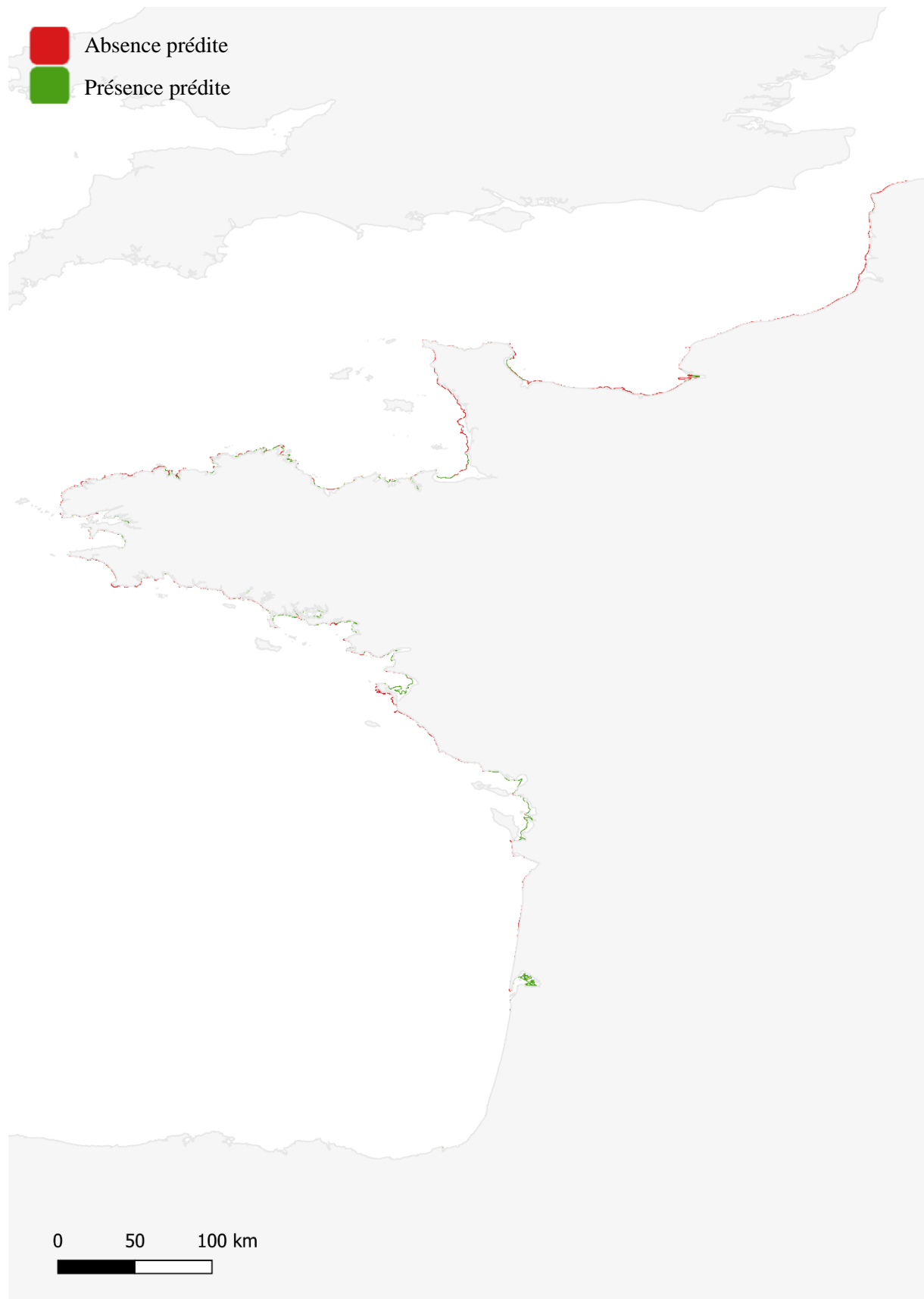
**Annexe 8 :** Carte de prédiction de l'indice d'habitat potentiel de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



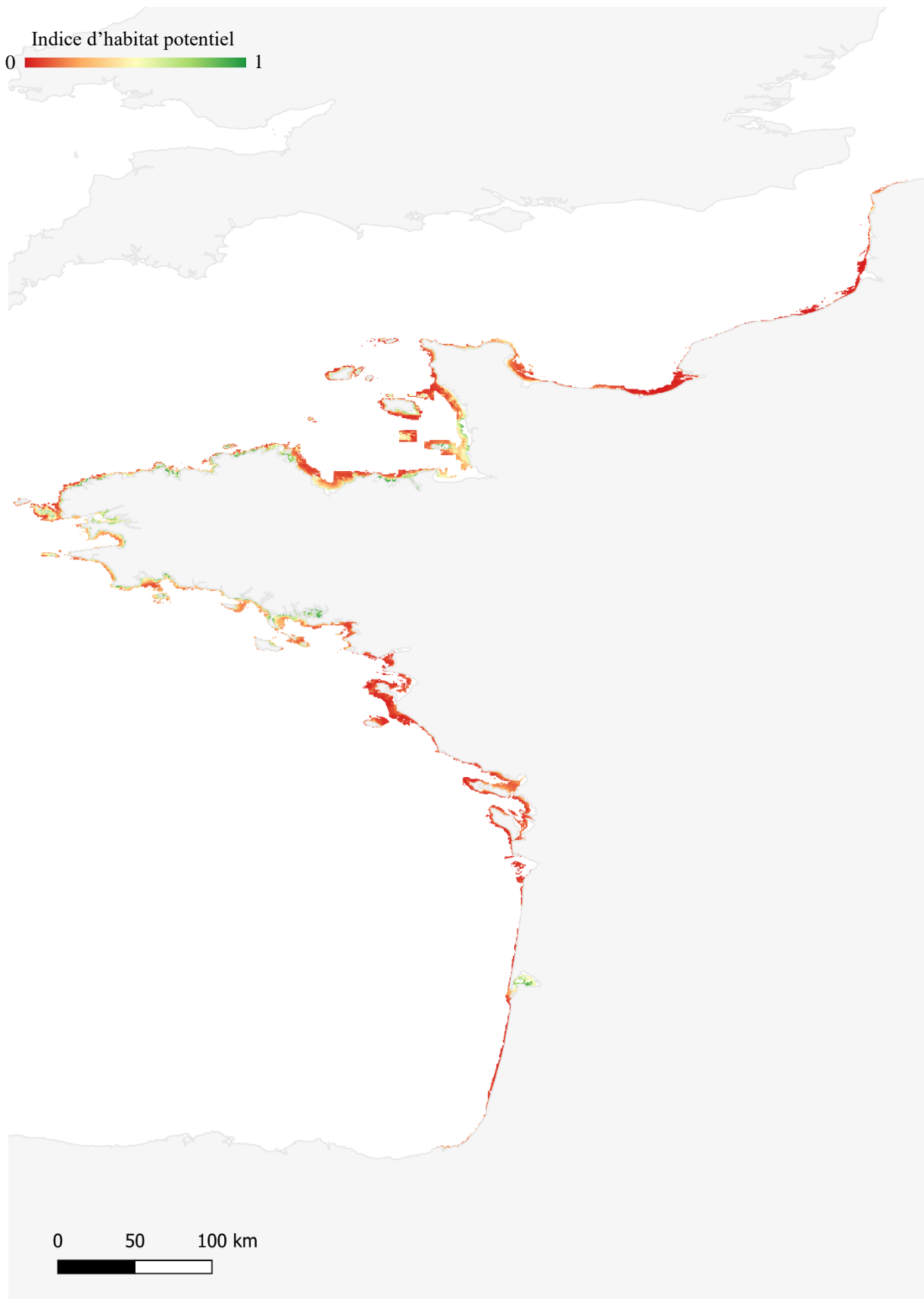
**Annexe 9 :** Carte de l'écart-type de l'indice d'habitat potentiel de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



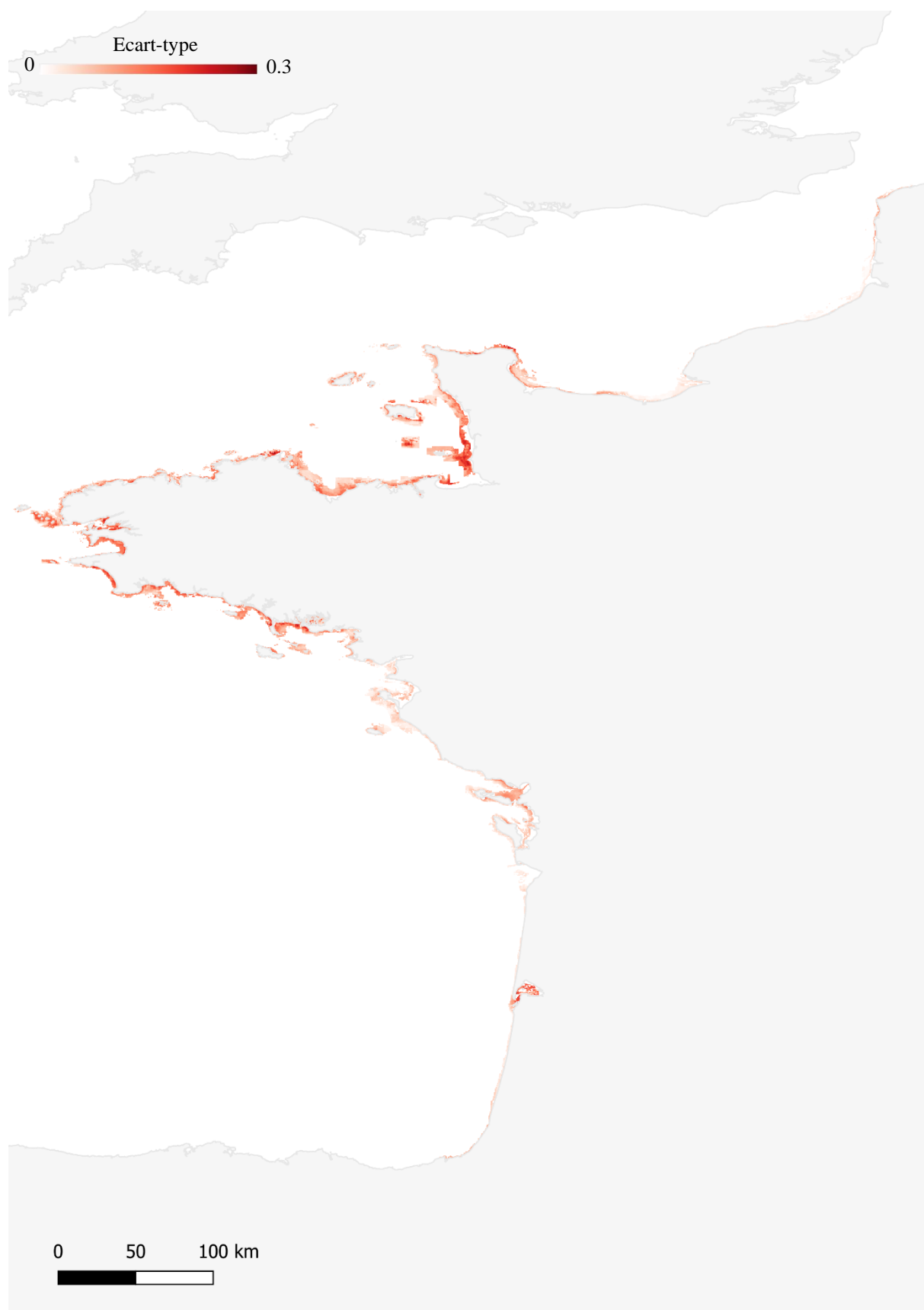
**Annexe 10 :** Carte de prédiction de présence/absence de *Z. noltei* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



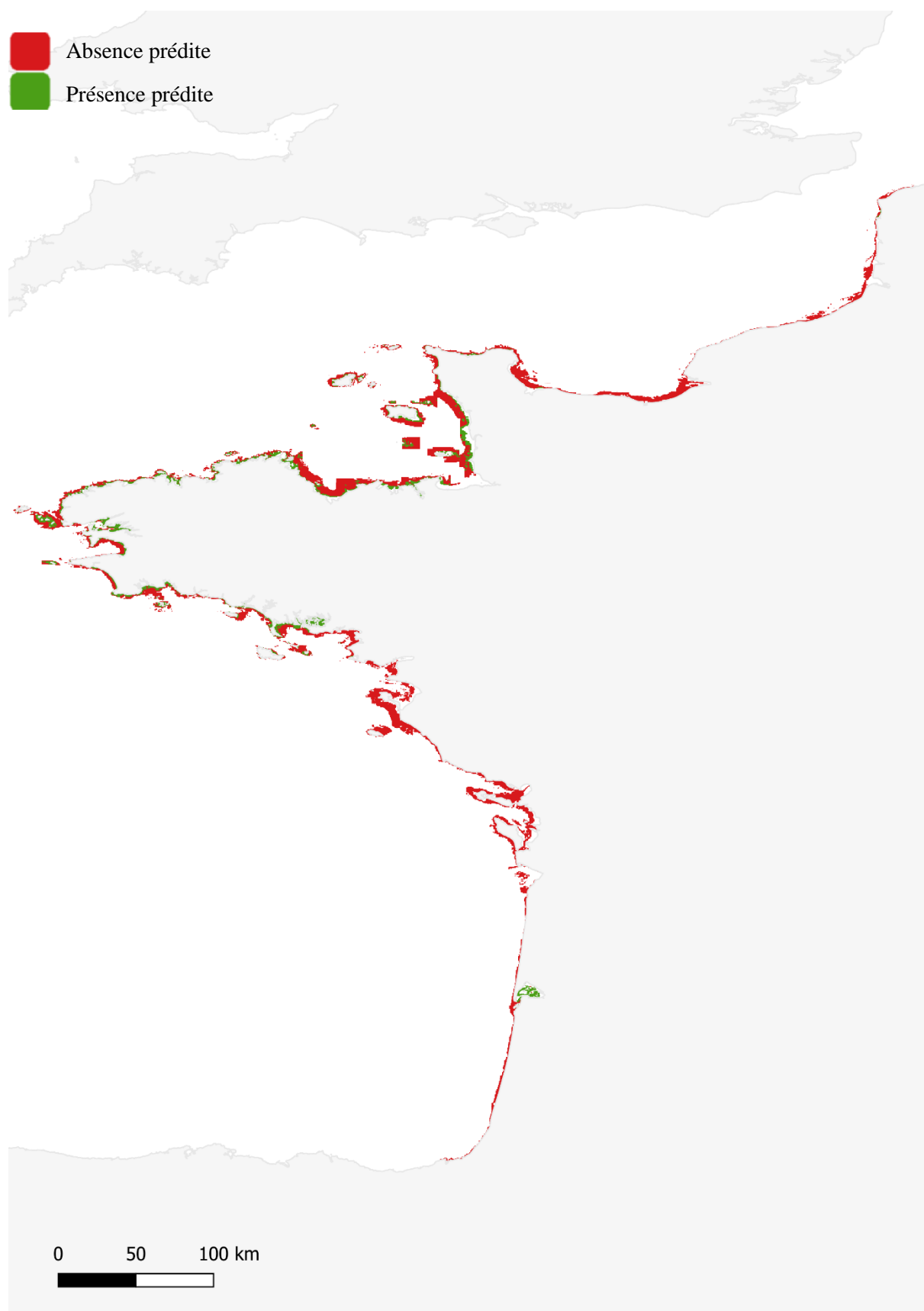
**Annexe 11 :** Carte de prédiction de l'indice d'habitat potentiel de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



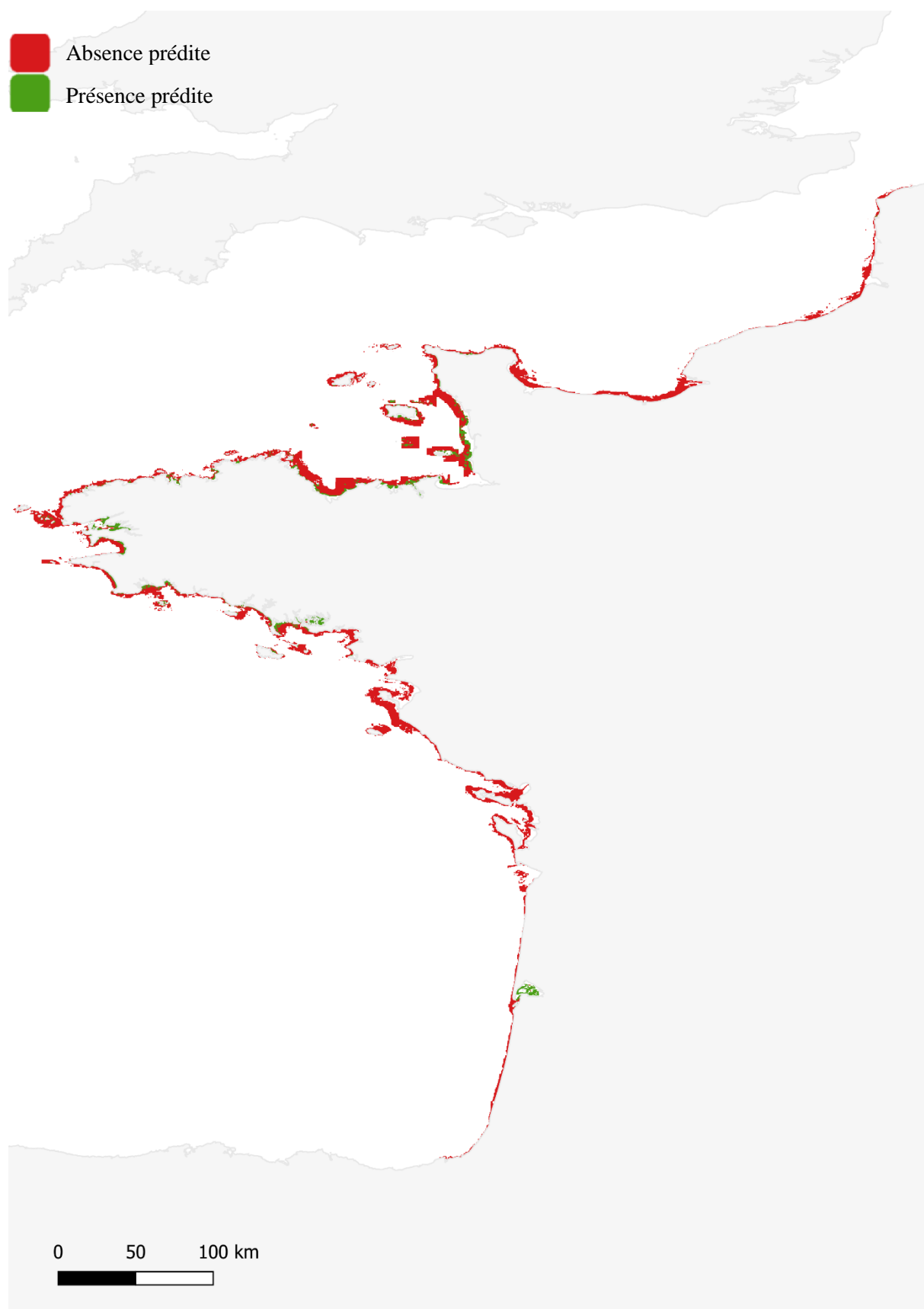
**Annexe 12 :** Carte de l'écart-type de l'indice d'habitat potentiel de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



**Annexe 13 :** Carte de prédiction de présence/absence de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m



**Annexe 14 :** Carte de prédiction de présence/absence de *Z. marina* à l'échelle de la façade Manche-Atlantique obtenue avec un modèle d'Ensemble Random Forest à 500 m avec un masque *Substrat*





Diplôme : Ingénieur agronome

Spécialité : « *Sciences halieutiques et aquacoles* » (SHA)

Option : « *Ressources et Ecosystèmes Aquatiques* » (REA)

Auteur : Romain MOUILLARD

Date de naissance : 09/05/1996

Nombre de pages : 35

Année de soutenance : 2020

Organisme d'accueil : IFREMER Brest

Unité de recherche DYNECO, Laboratoire LEBCO

Adresse : 1625 route de Sainte Anne (29280 – Plouzané)

Maitres de stage : Mickael Vasquez & Martin Marzloff

Titre en français : Modélisation statistique de la distribution spatiale des espèces *Zostera marina* et *Zostera noltei* à l'échelle de la façade Manche-Atlantique

Titre en anglais : Statistical modelling of the spatial distribution of *Zostera marina* and *Zostera noltei* along the French Atlantic and English Channel coastline

Résumé : Le projet Life MarHa coordonné par l'Office Français de la Biodiversité poursuit l'objectif de rétablir et/ou maintenir le bon état écologique des habitats naturels marins en France. Dans le cadre de ce projet, ma mission a été de réaliser la modélisation statistique de la distribution spatiale des herbiers de zostères (*Zostera marina* et *Zostera noltei*) à l'échelle de la façade Manche-Atlantique. Un premier aspect visait à analyser la sensibilité des résultats à l'approche adoptée et notamment au type de modèle utilisé (GLM, RF, BRT) afin de caractériser la robustesse de la démarche, et d'identifier les modèles les plus performants. Ensuite, l'objectif a été de construire à partir des modèles retenus des cartes de prédiction de la distribution spatiale potentielle de ces deux espèces à l'échelle de la façade Manche-Atlantique, puis de décrire la niche environnementale des deux espèces. Enfin, les prédictions des modèles ont permis de décrire les impacts potentiels des différents scénarios du changement climatique sur l'habitat potentiel de ces deux espèces. A l'aide d'un ensemble de modèles Random Forest, cette étude renforce les connaissances actuelles concernant la niche environnementale des espèces et leur distribution le long du littoral métropolitain atlantique. Cette étude met également en avant la vulnérabilité de ces espèces au changement climatique et la nécessité de mettre en place des politiques de gestion et des mesures efficaces.

Abstract: The MarHa Life project coordinated by the Office Français de la Biodiversité aims at restoring and/or maintaining the good ecological status of natural marine habitats in France. As part of this project, my mission was to statistically model the spatial distribution of zostera seagrass beds (*Zostera marina* and *Zostera noltei*) along the French Atlantic and English Channel coastline. A first objective was to assess results sensitivity to the modelling approach and in particular to the type of model used (GLM, RF, BRT) so as to characterize the performance of the method and to identify the best-suited models. Then, using the selected models, a second objective was to predict the potential spatial distribution of these two species along the French Atlantic-Channel coast, and to describe their environmental niche. Finally, the last objective was to model the potential impacts of different climate change scenarios on the potential habitat of these two species. Using an ensemble of Random Forest models, this study consolidated our current knowledge of *Zostera spp.* environmental niches and spatial distribution along the Western French metropolitan coastline. This study also highlights the vulnerability of these species to climate change and the need for effective management policies and measures.

Mots clés : *Zostera marina* ; *Zostera noltei* ; Modèle d'habitat ; Habitat potentiel ; Niche environnementale  
Key words : *Zostera marina* ; *Zostera noltei* ; Habitat modelling ; Potential habitat ; Environmental niche