

Modélisation de la niche écologique du Lézard ocellé en France hexagonale

Jérôme Prunier, ADENEKO

25/10/2024



© Loïc Brepson

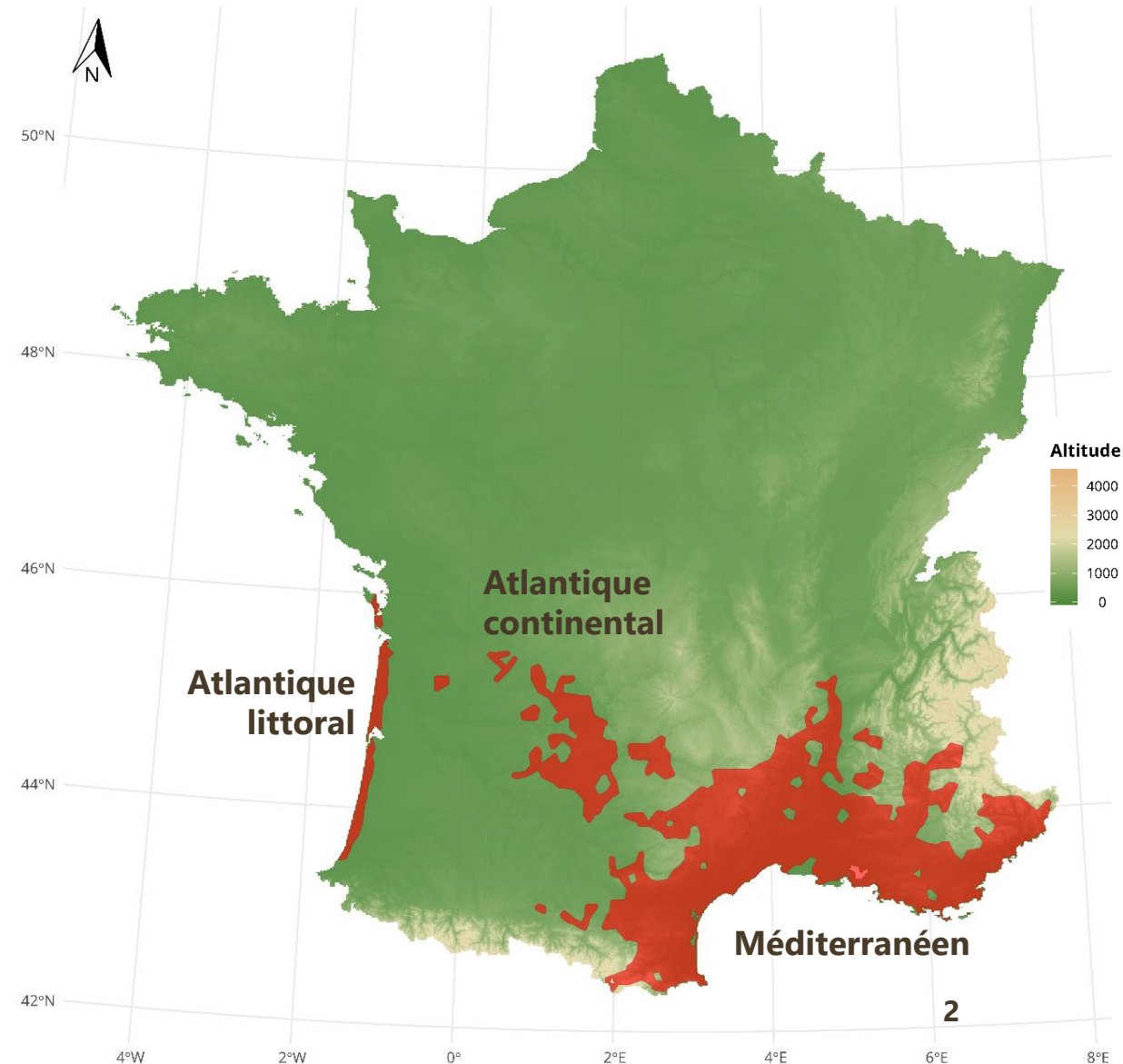


© adeneko



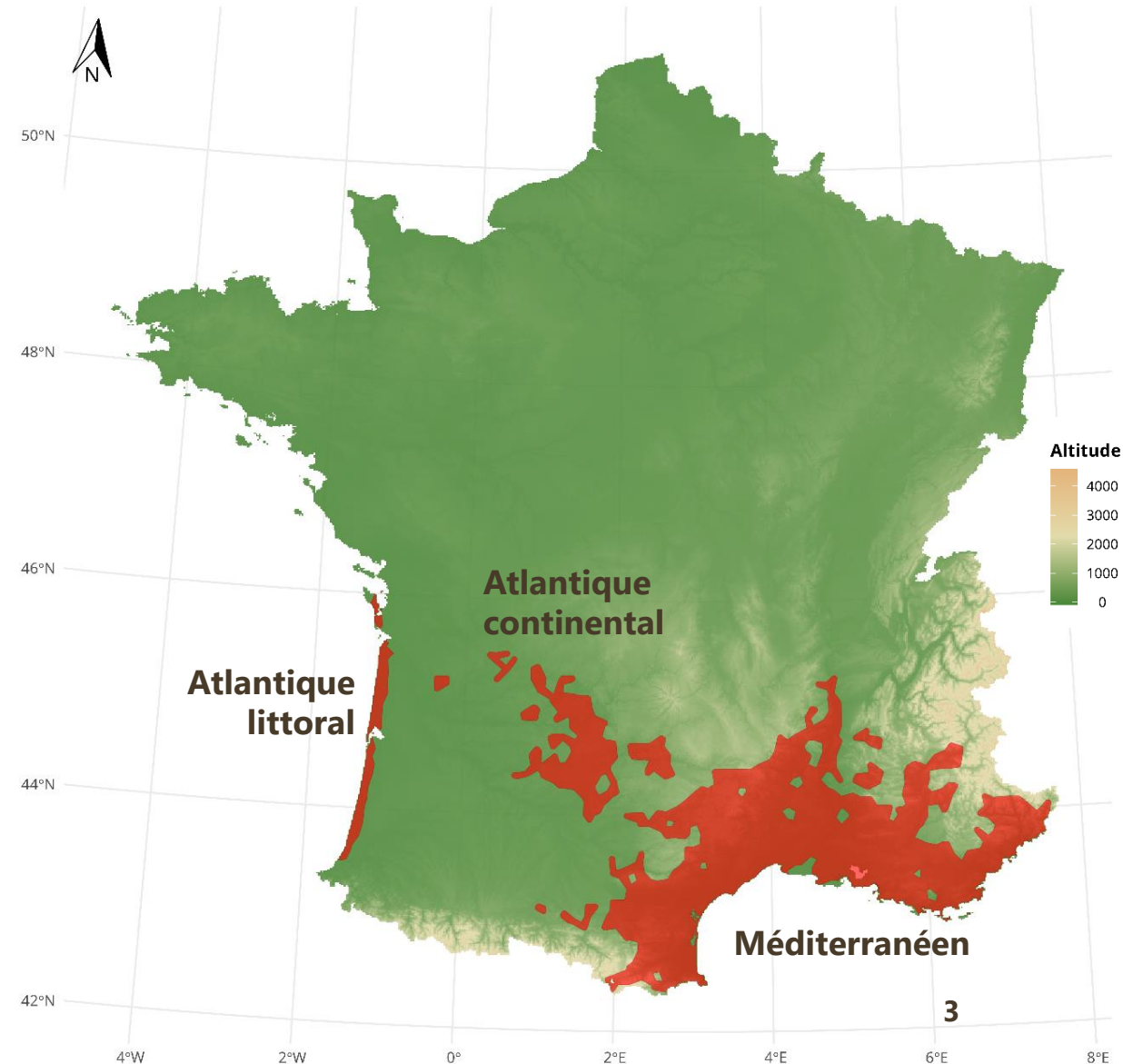
Modélisation de la niche écologique du **Lézard ocellé**

- Trois grandes zones de peuplement en France hexagonale (absent de Corse)



Modélisation de la niche écologique du **Lézard ocellé**

- ▶ Trois grandes zones de peuplement en France hexagonale (absent de Corse)
- ▶ Espèce en déclin continu depuis plusieurs décennies



Modélisation de la niche écologique du **Lézard ocellé**

- ▶ Trois grandes zones de peuplement en France hexagonale (absent de Corse)
- ▶ Espèce en déclin continu depuis plusieurs décennies
- ▶ Second PNA 2020-2029
 - Volet renforcement des connaissances sur la répartition
 - Guider les efforts de prospection
 - Orienter les actions de conservation



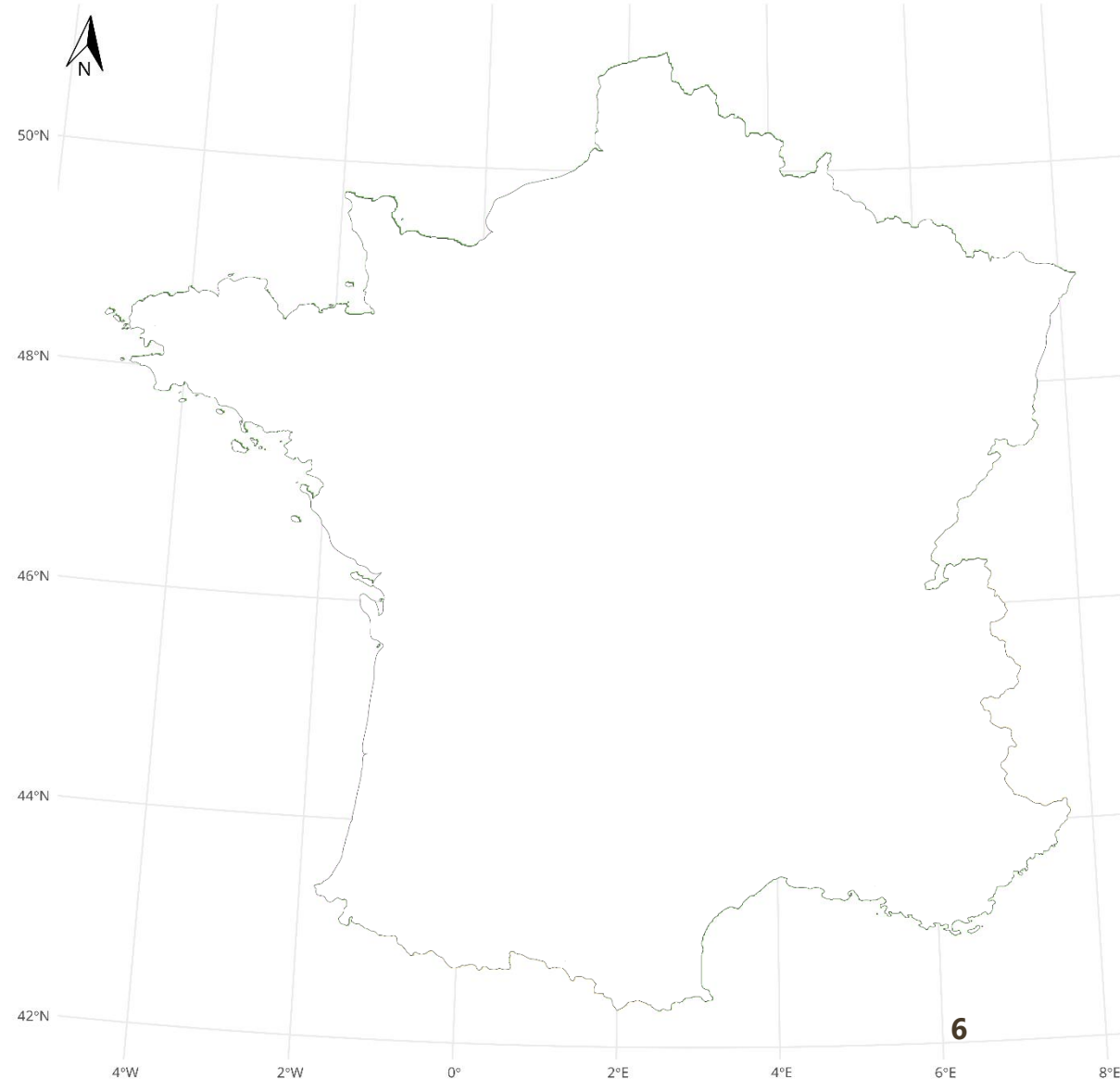
Modélisation de la niche écologique du Lézard ocellé

- ▶ Trois grandes zones de peuplement en France hexagonale (absent de Corse)
- ▶ Espèce en déclin continu depuis plusieurs décennies
- ▶ Second PNA 2020-2029
 - Volet renforcement des connaissances sur la répartition
 - Guider les efforts de prospection
 - Orienter les actions de conservation
 - Modélisation de la niche écologique à l'échelle hexagonale



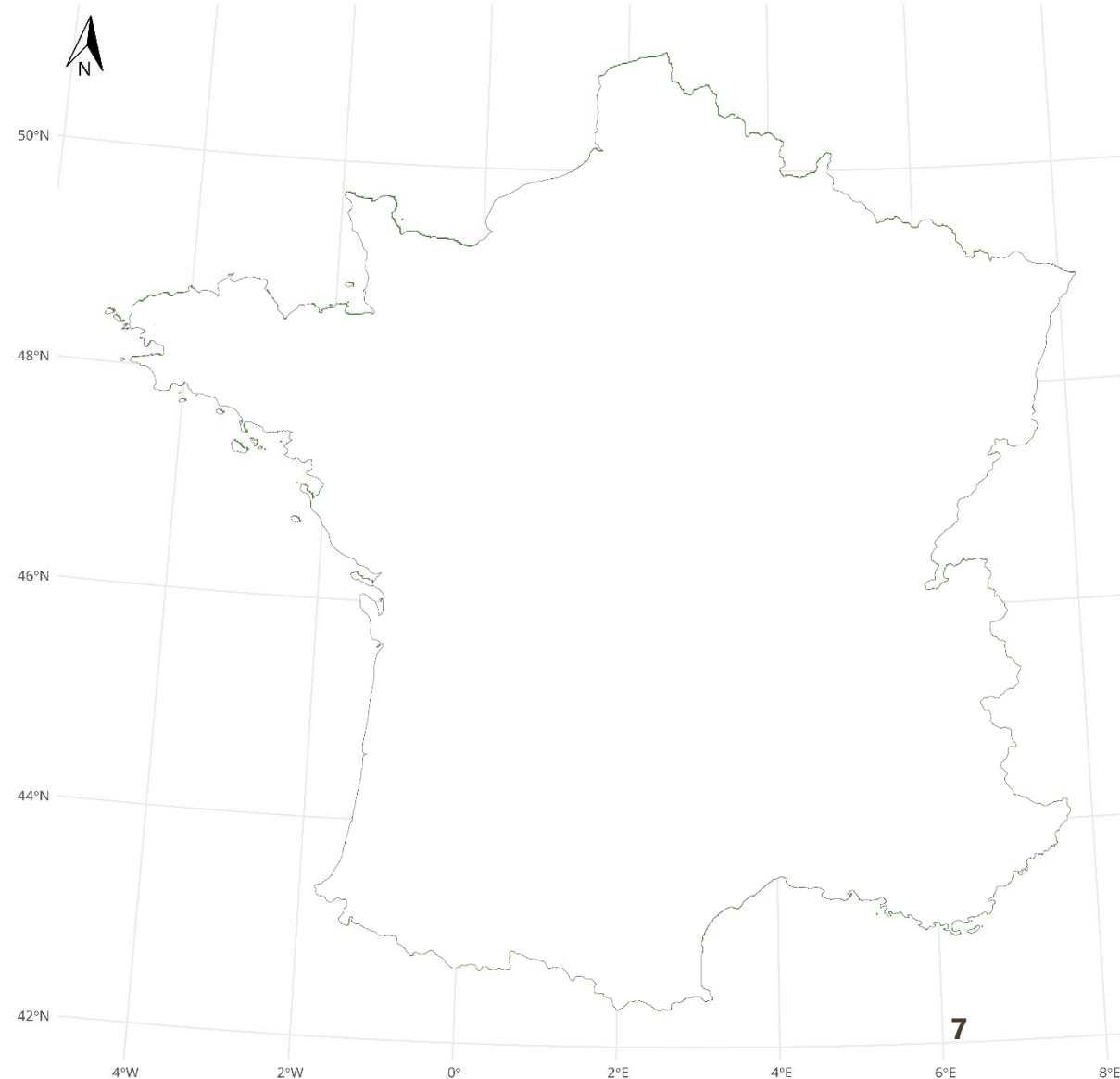
Modélisation de la niche écologique du Lézard ocellé

- Ensemble des caractéristiques environnementales permettant à un organisme de boucler son cycle vital



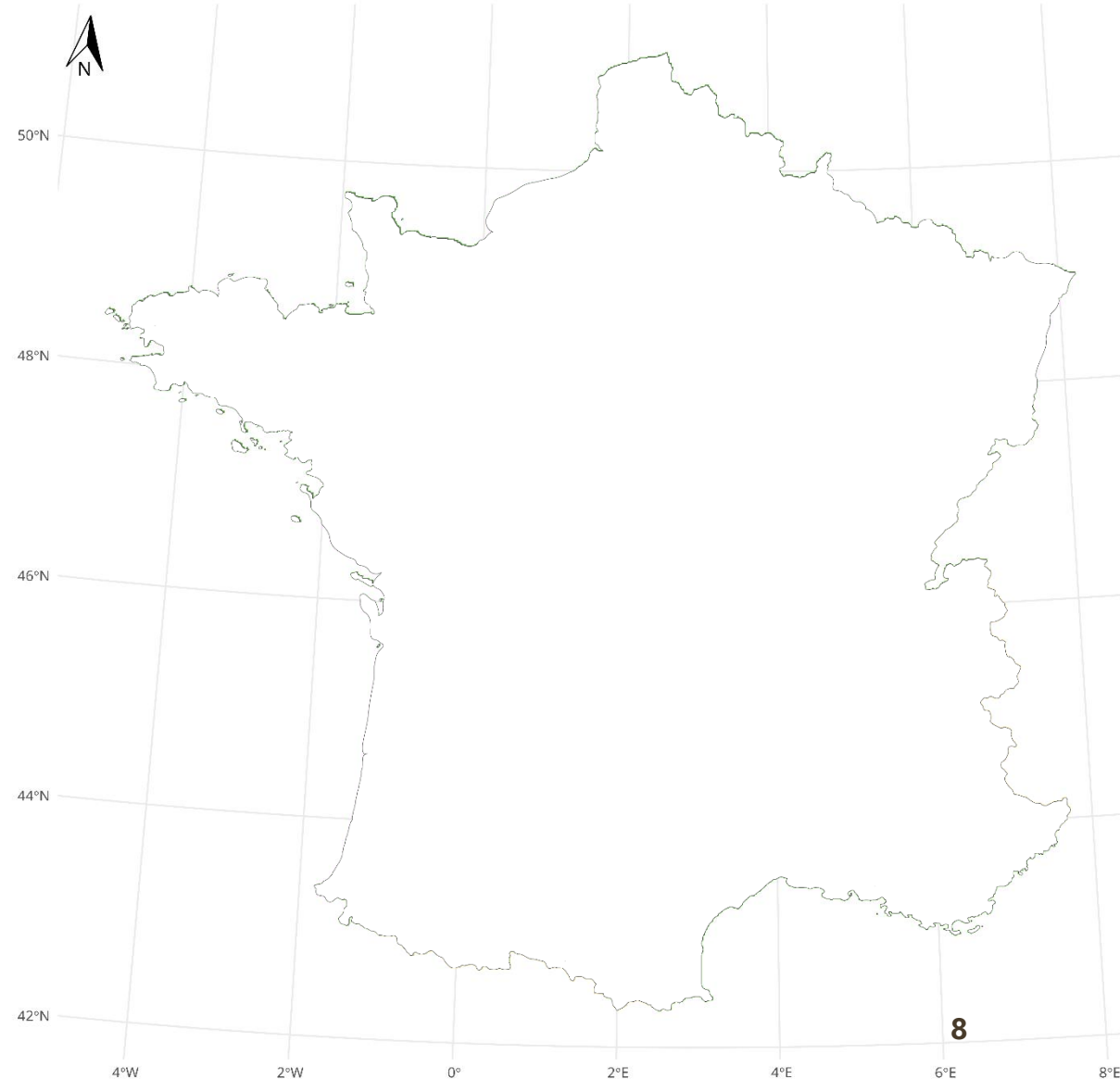
Modélisation de la niche écologique du Lézard ocellé

- Prédire la répartition géographique des besoins écologiques de l'espèce



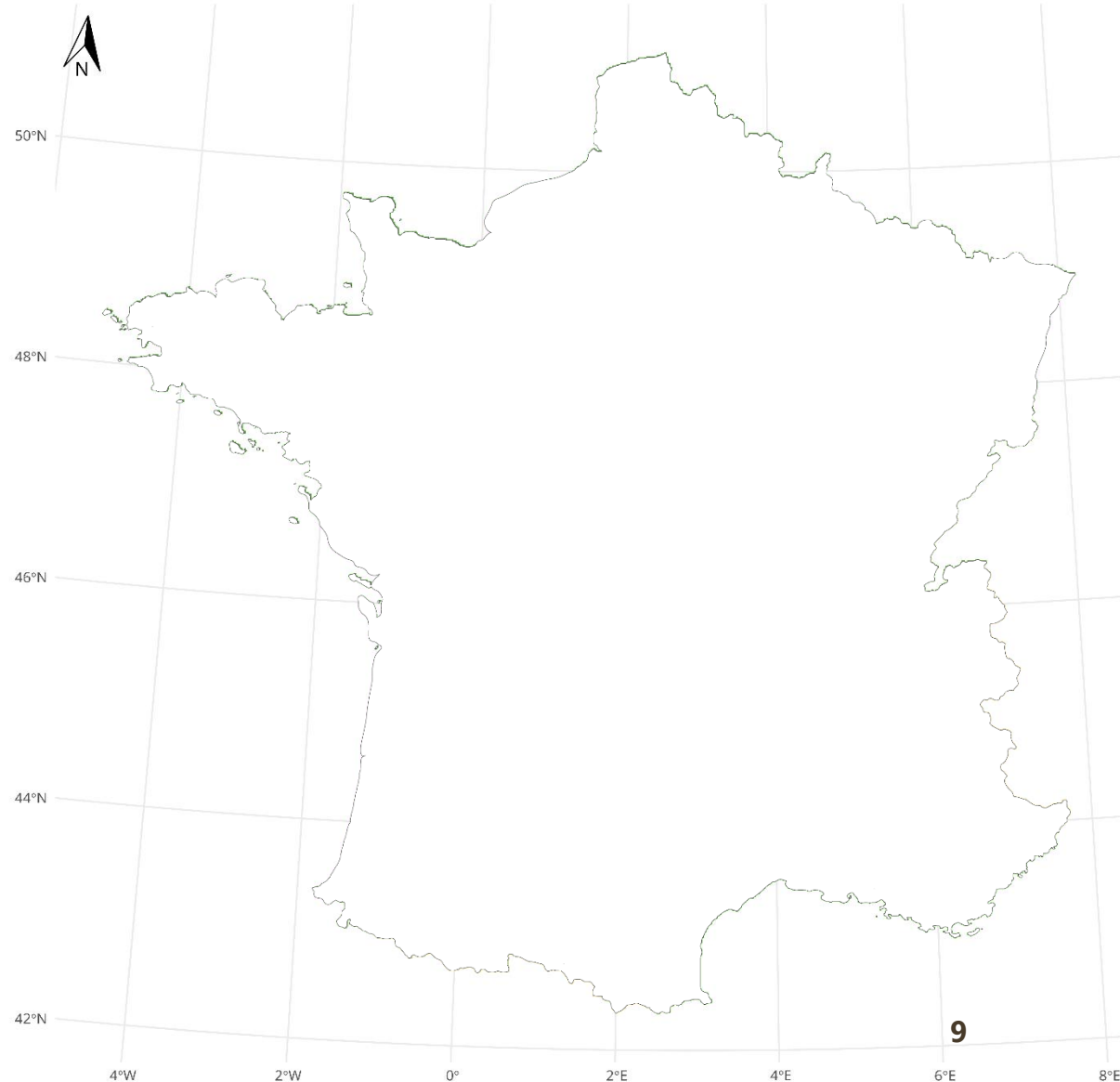
Modélisation de la niche écologique du Lézard ocellé

- ▶ Prédire la répartition géographique des besoins écologiques de l'espèce
- ↕
- ▶ Prédire la répartition de la qualité de l'habitat



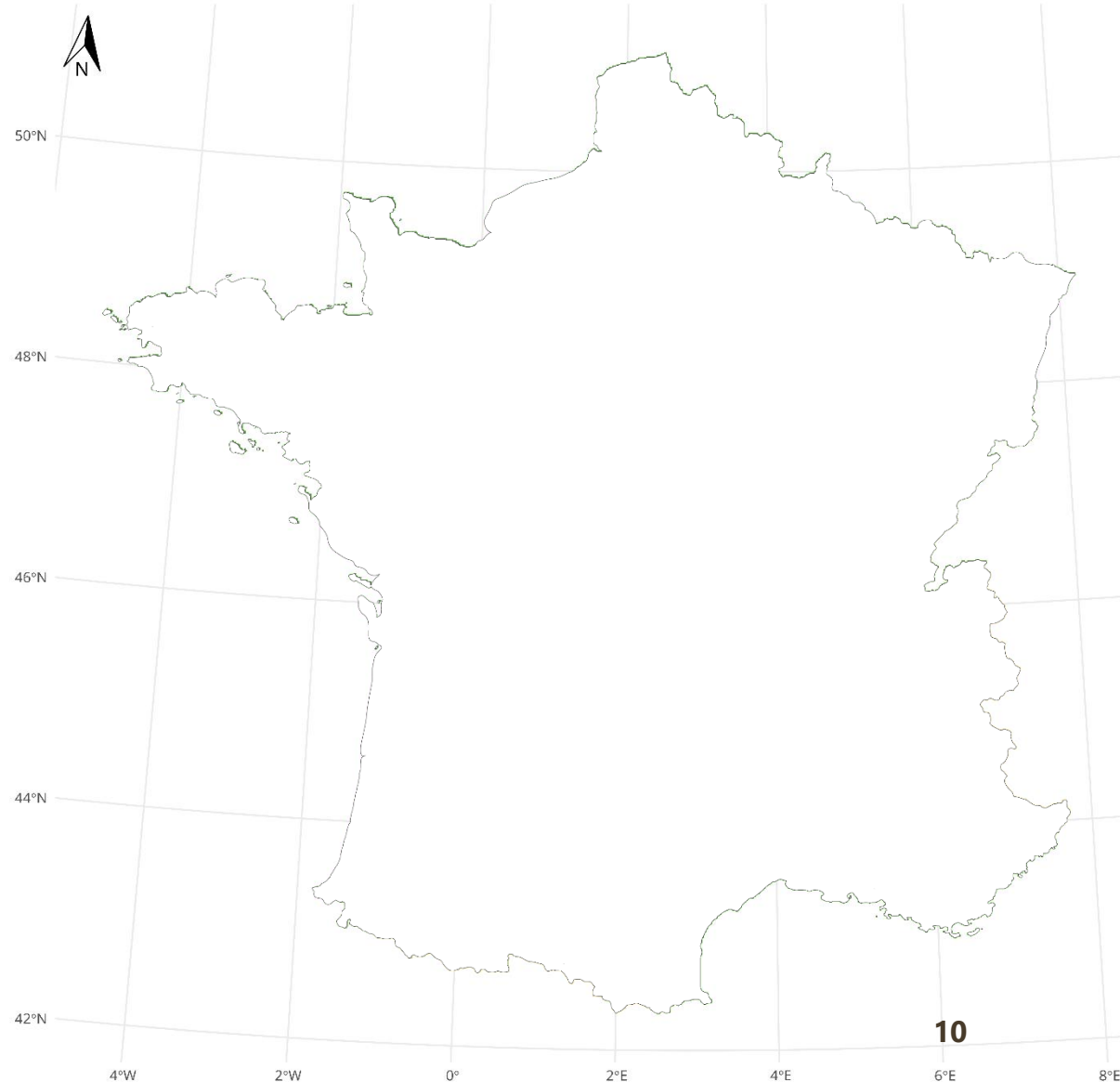
Modélisation de la niche écologique du Lézard ocellé

- ▶ Prédire la répartition géographique des besoins écologiques de l'espèce
- ↕
- ▶ Prédire la répartition de la qualité de l'habitat
- ↕
- ▶ Prédire la répartition potentielle de l'espèce



Modélisation de la niche écologique du Lézard ocellé

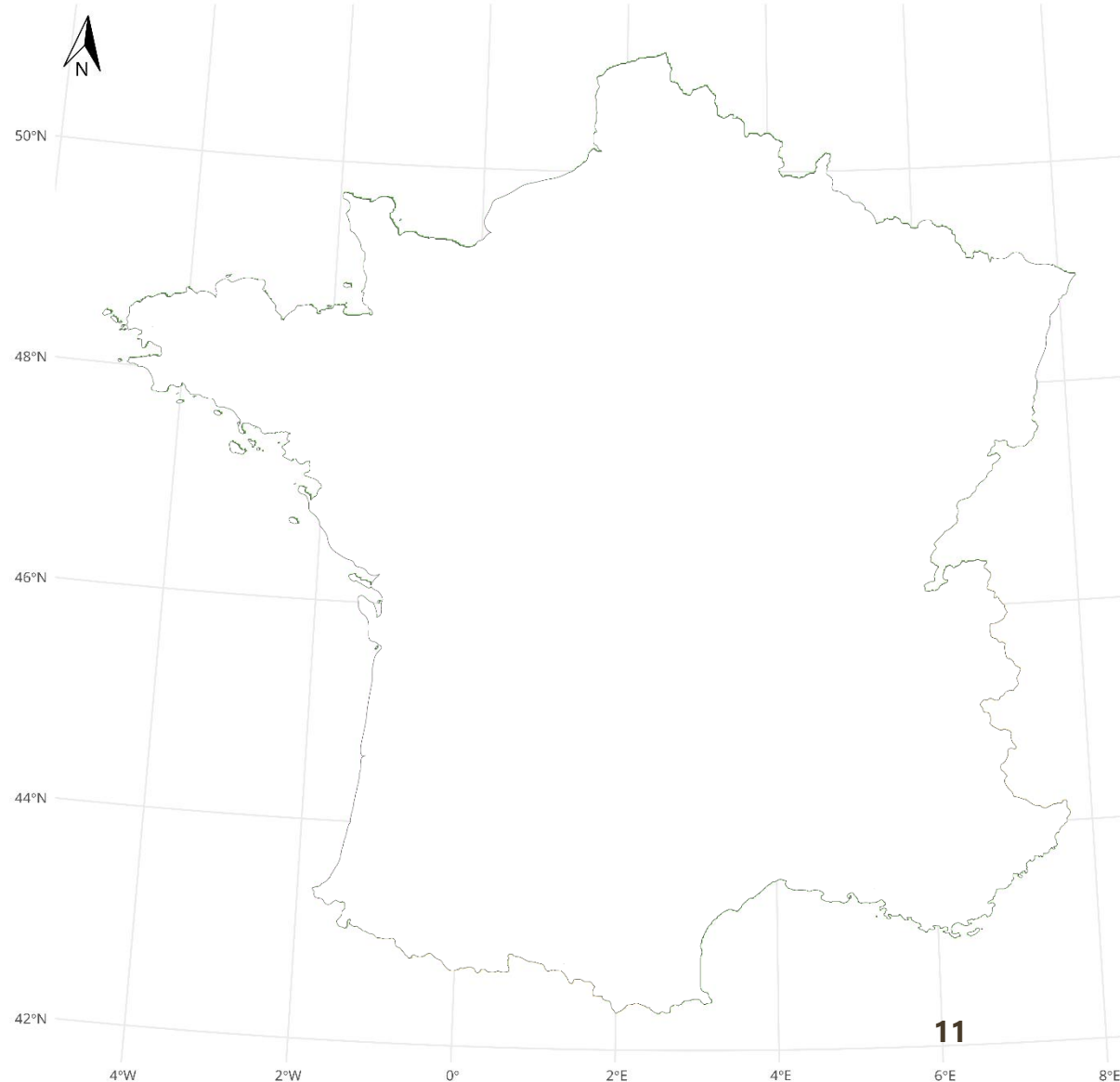
- ▶ Prédire la répartition géographique des besoins écologiques de l'espèce
- ↕
- ▶ Prédire la répartition de la qualité de l'habitat
- ↕
- ▶ Prédire la répartition potentielle de l'espèce
- ↕
- ▶ Prédire la probabilité de présence de l'espèce



Modélisation de la niche écologique du Lézard ocellé

- ▶ Prédire la répartition géographique des besoins écologiques de l'espèce
- ▶ Prédire la répartition de la qualité de l'habitat
- ▶ Prédire la répartition potentielle de l'espèce
- ▶ Prédire la probabilité de présence de l'espèce

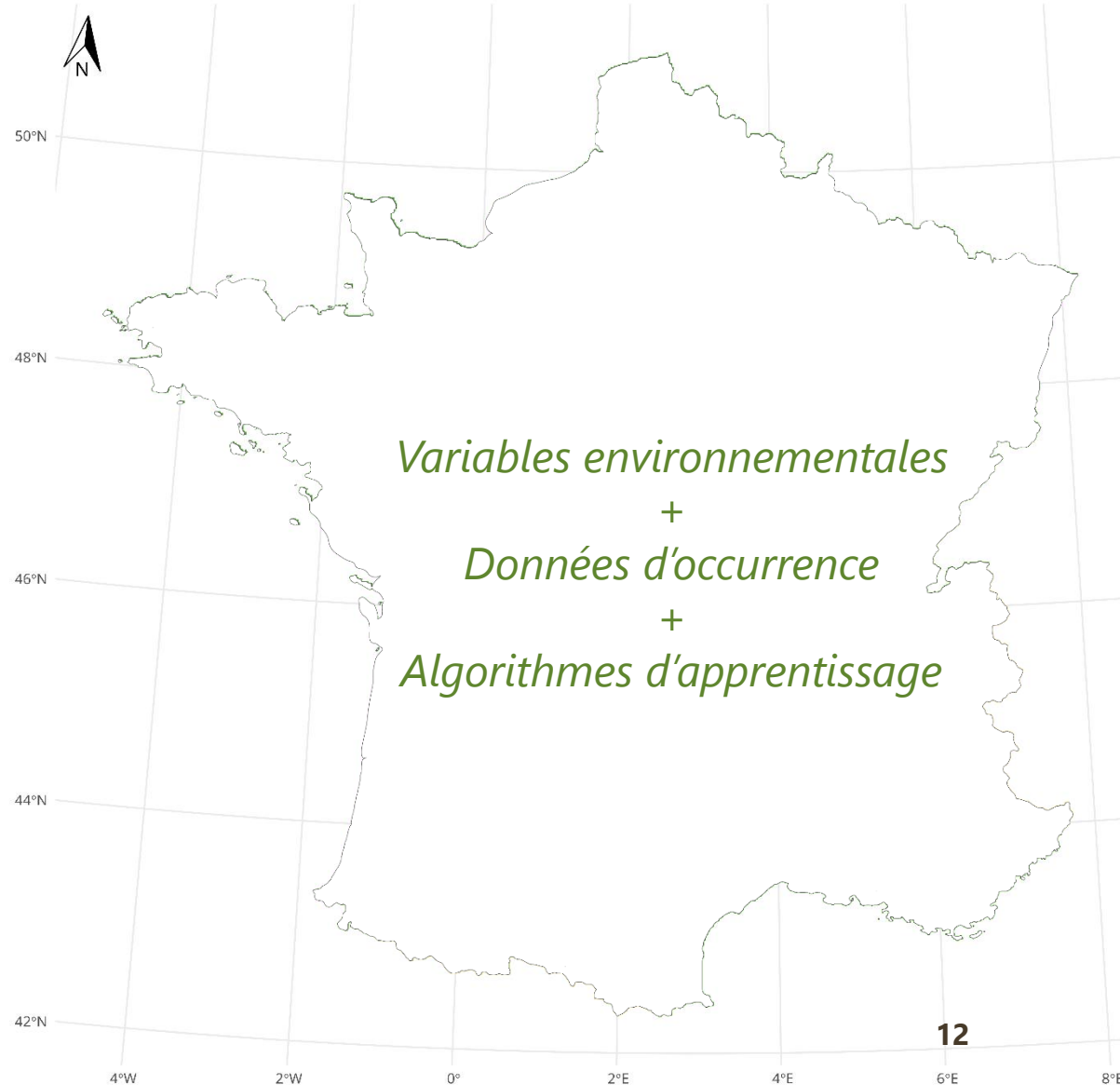
Intelligence artificielle



Modélisation de la niche écologique du Lézard ocellé

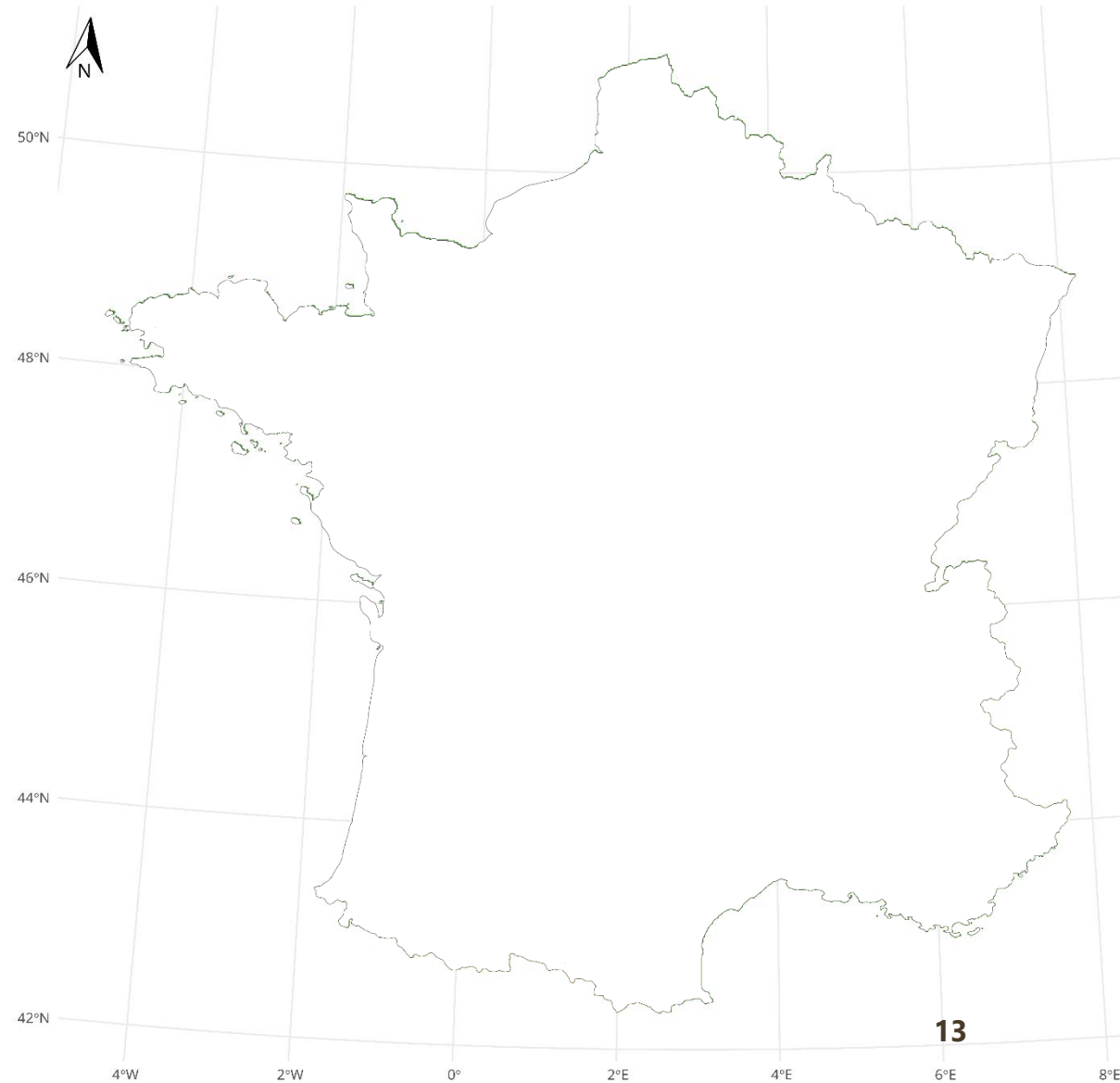
- ▶ Prédire la répartition géographique des besoins écologiques de l'espèce
- ▶ Prédire la répartition de la qualité de l'habitat
- ▶ Prédire la répartition potentielle de l'espèce
- ▶ Prédire la probabilité de présence de l'espèce

Intelligence artificielle



Données environnementales

► Résolution spatiale : 200m

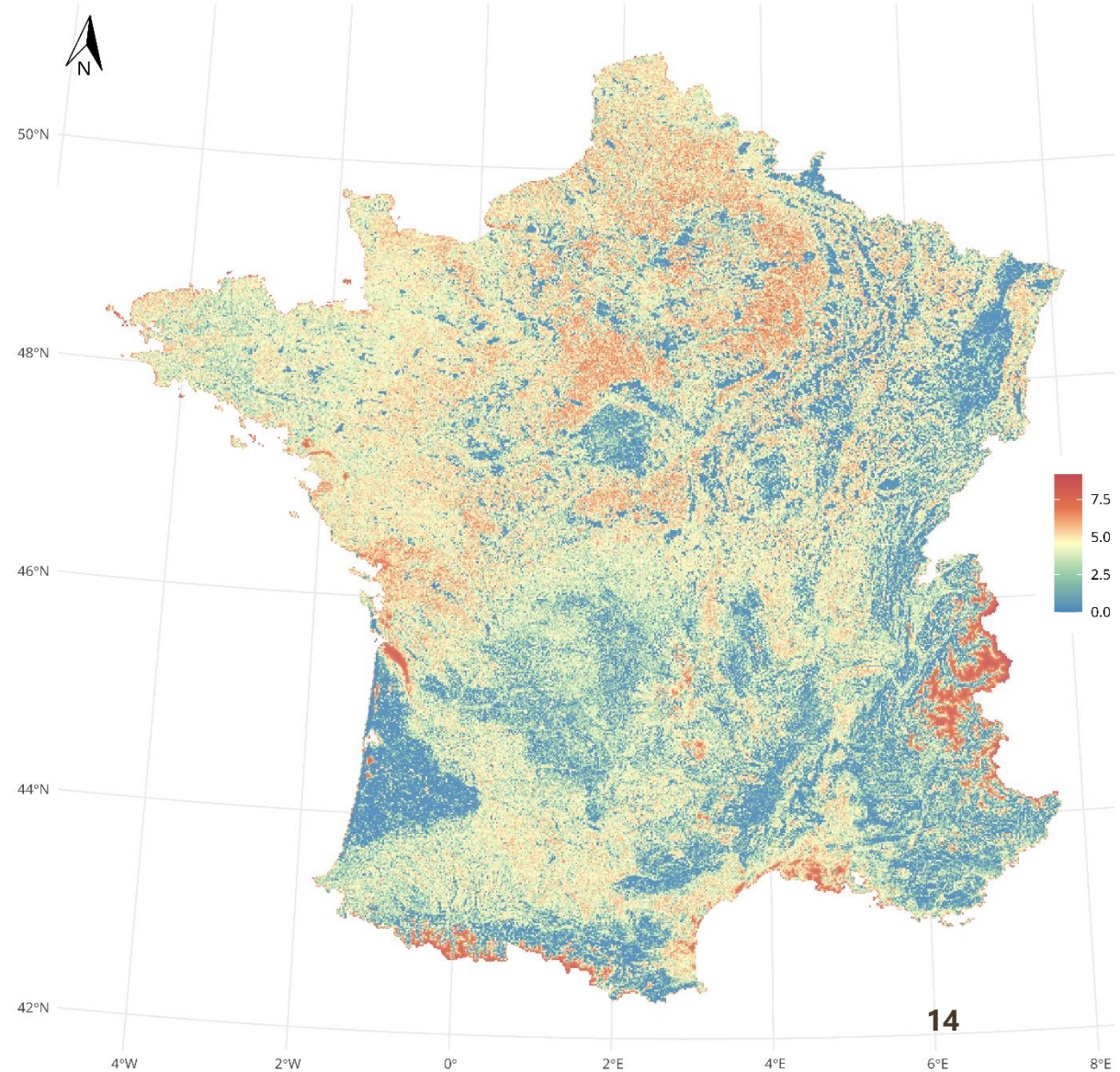


Données environnementales

- ▶ Résolution spatiale : 200m
- ▶ 9 variables d'occupation du sol
 - zones xériques ouvertes
 - zones urbaines, routes, forêts, zones agricoles, etc.

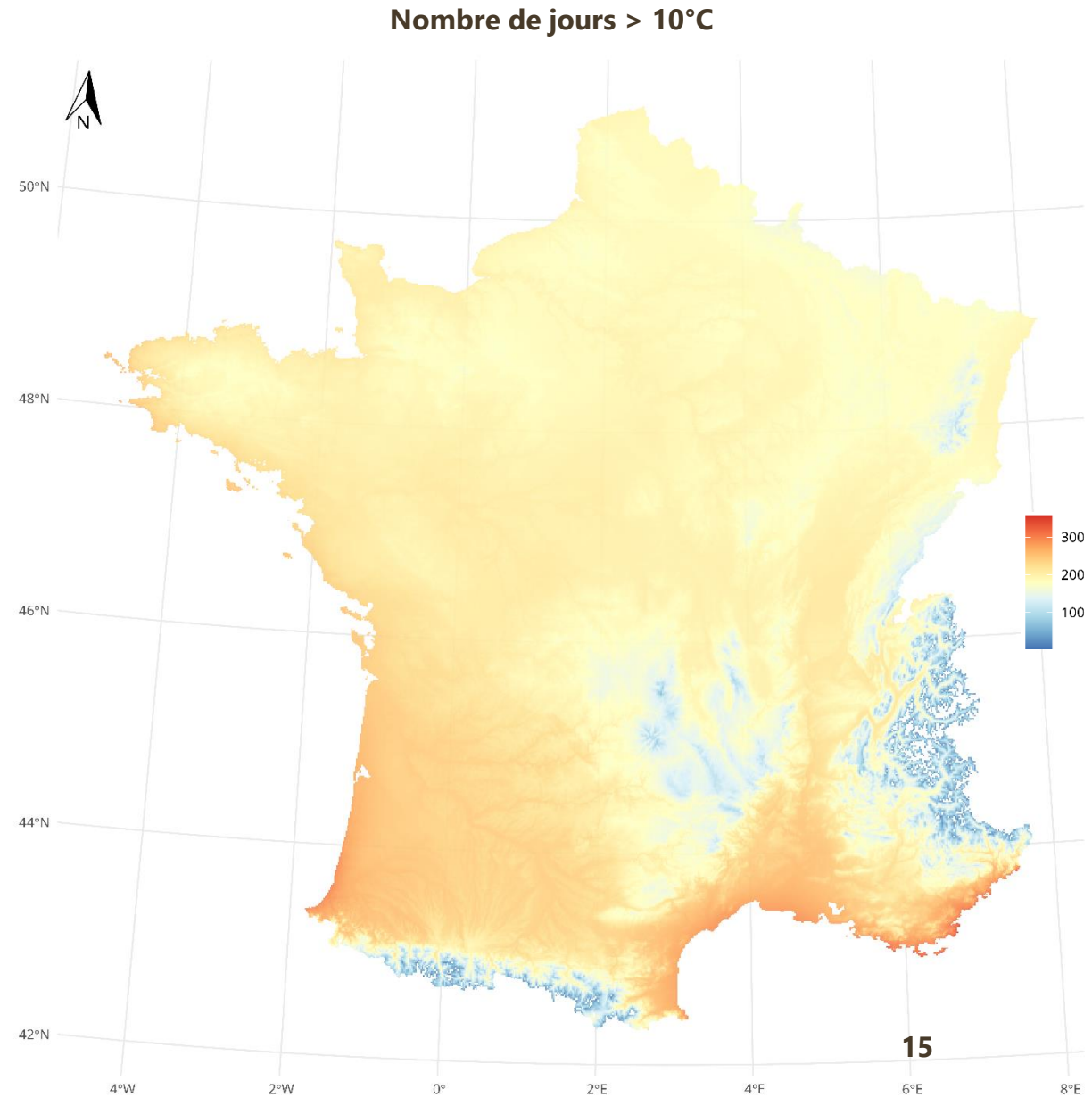


Distance au boisement le plus proche en km (log)



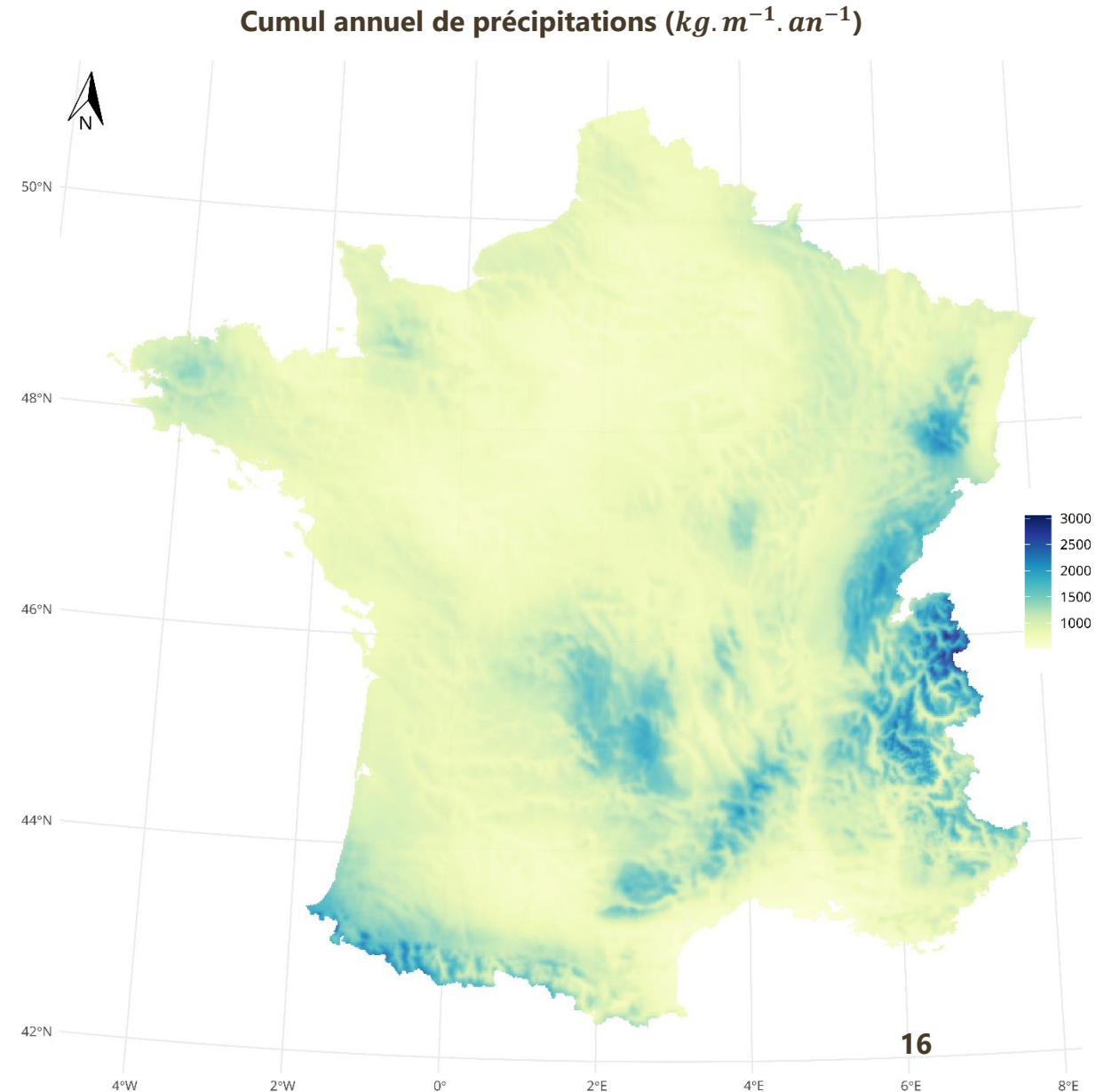
Données environnementales

- ▶ Résolution spatiale : 200m
- ▶ 9 variables d'occupation du sol
- ▶ 16 variables climatiques
 - 8 variables thermiques
 - températures, cumul de chaleur, énergie solaire



Données environnementales

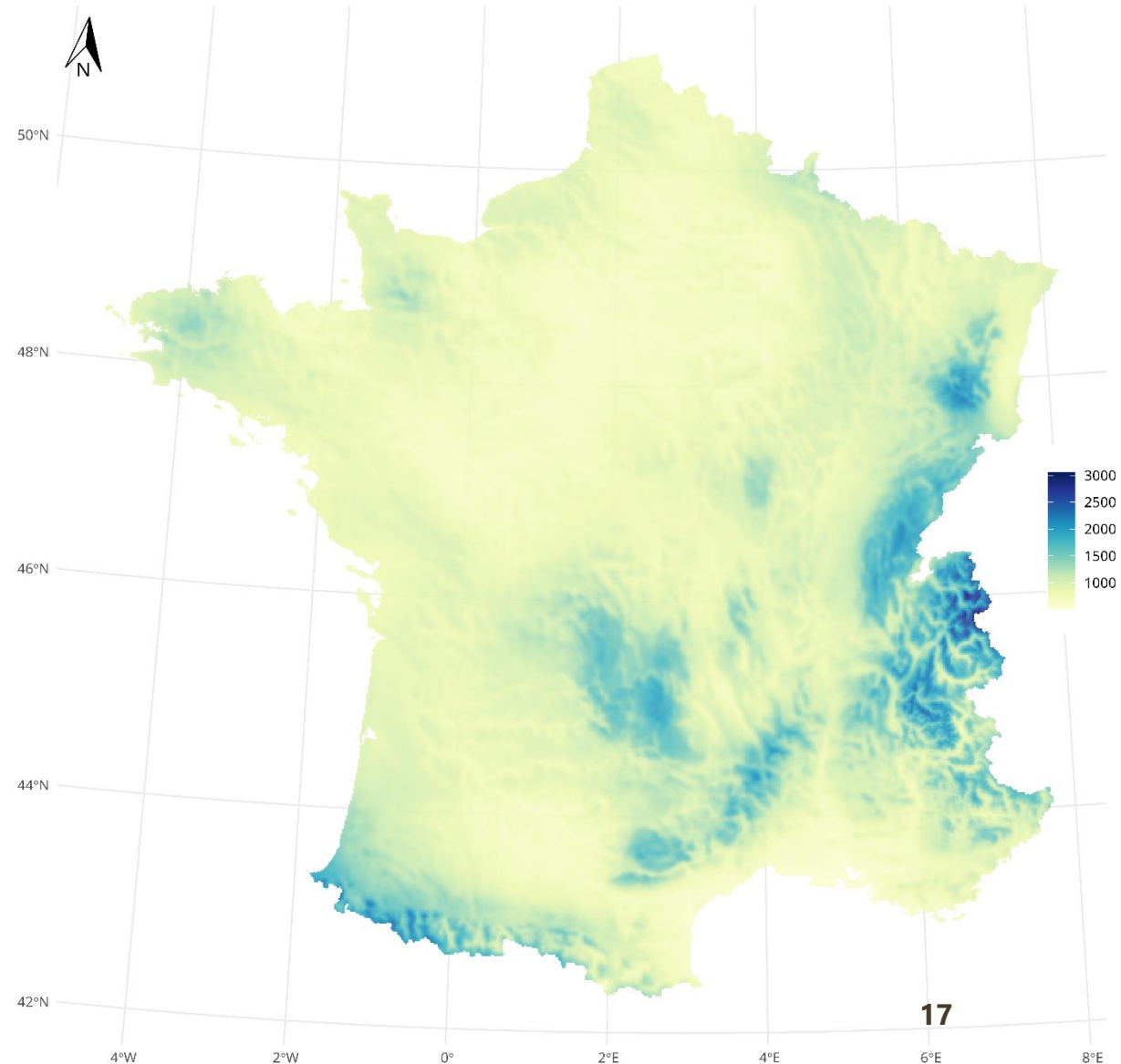
- ▶ Résolution spatiale : 200m
- ▶ 9 variables d'occupation du sol
- ▶ 16 variables climatiques
 - 8 variables thermiques
 - températures, cumul de chaleur, énergie solaire
 - 8 variables hydriques
 - précipitations, stress hydrique, saturation en eau, évaporation



Données environnementales

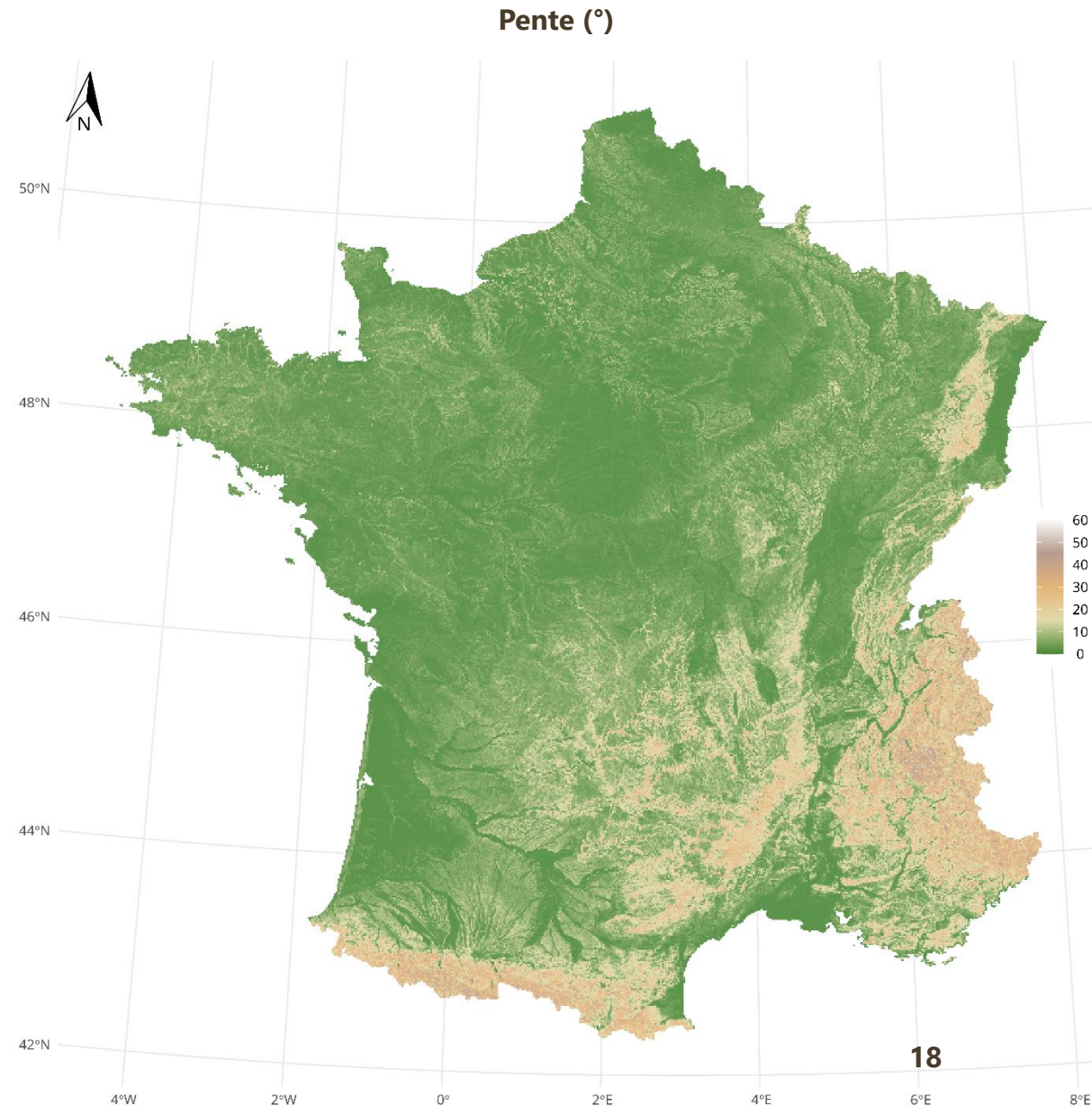
- ▶ Résolution spatiale : 200m
- ▶ 9 variables d'occupation du sol
- ▶ 16 variables climatiques
 - 8 variables thermiques
 - températures, cumul de chaleur, énergie solaire
 - 8 variables hydriques
 - précipitations, stress hydrique, saturation en eau, évaporation
 - Période 1981-2010
 - Projections climatiques à l'horizon 2100

Cumul annuel de précipitations ($kg \cdot m^{-1} \cdot an^{-1}$)



Données environnementales

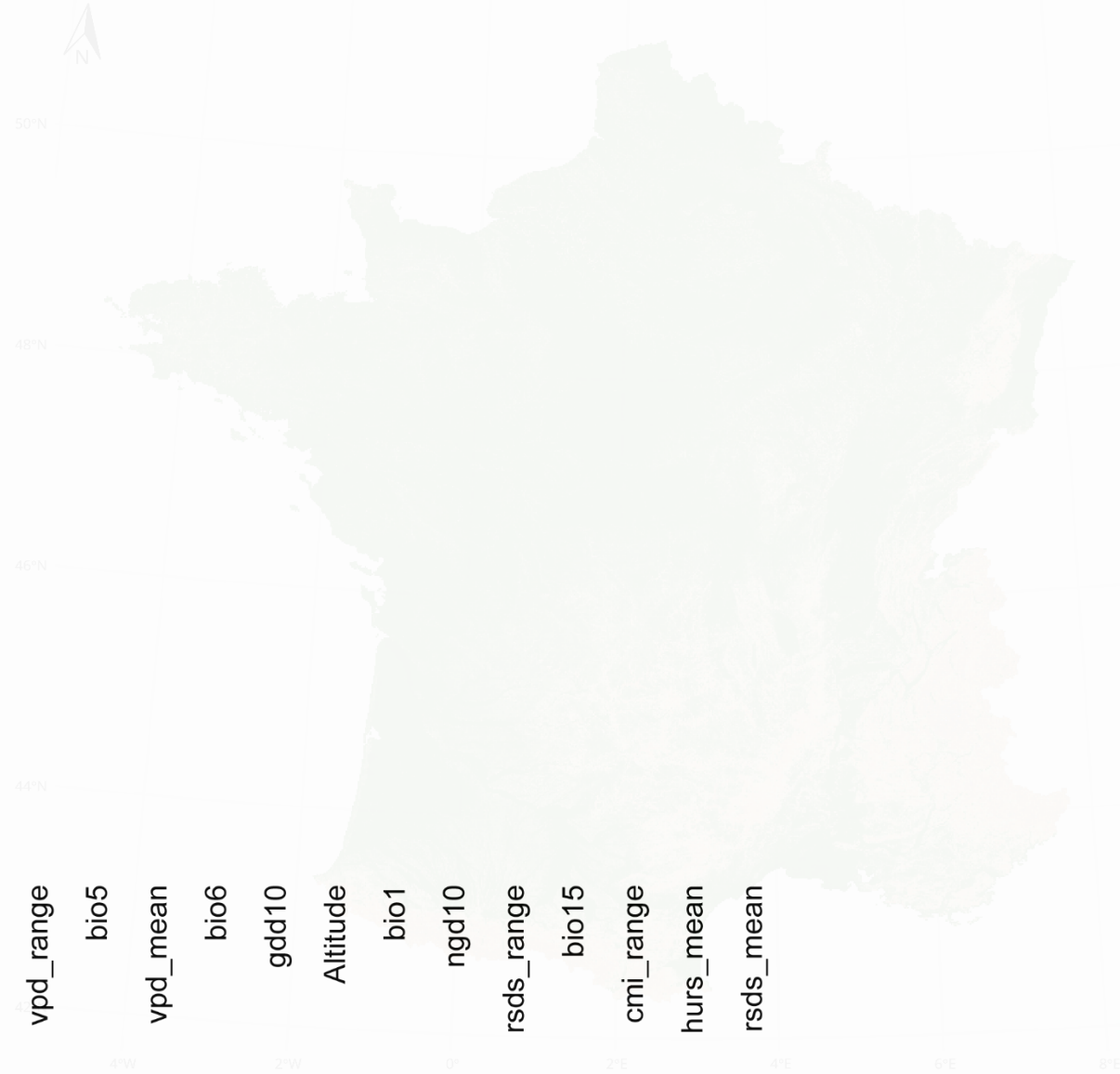
- ▶ Résolution spatiale : 200m
- ▶ 9 variables d'occupation du sol
- ▶ 16 variables climatiques
- ▶ 4 variables topographiques
 - altitude, pente, rugosité, orientation



Données environnementales

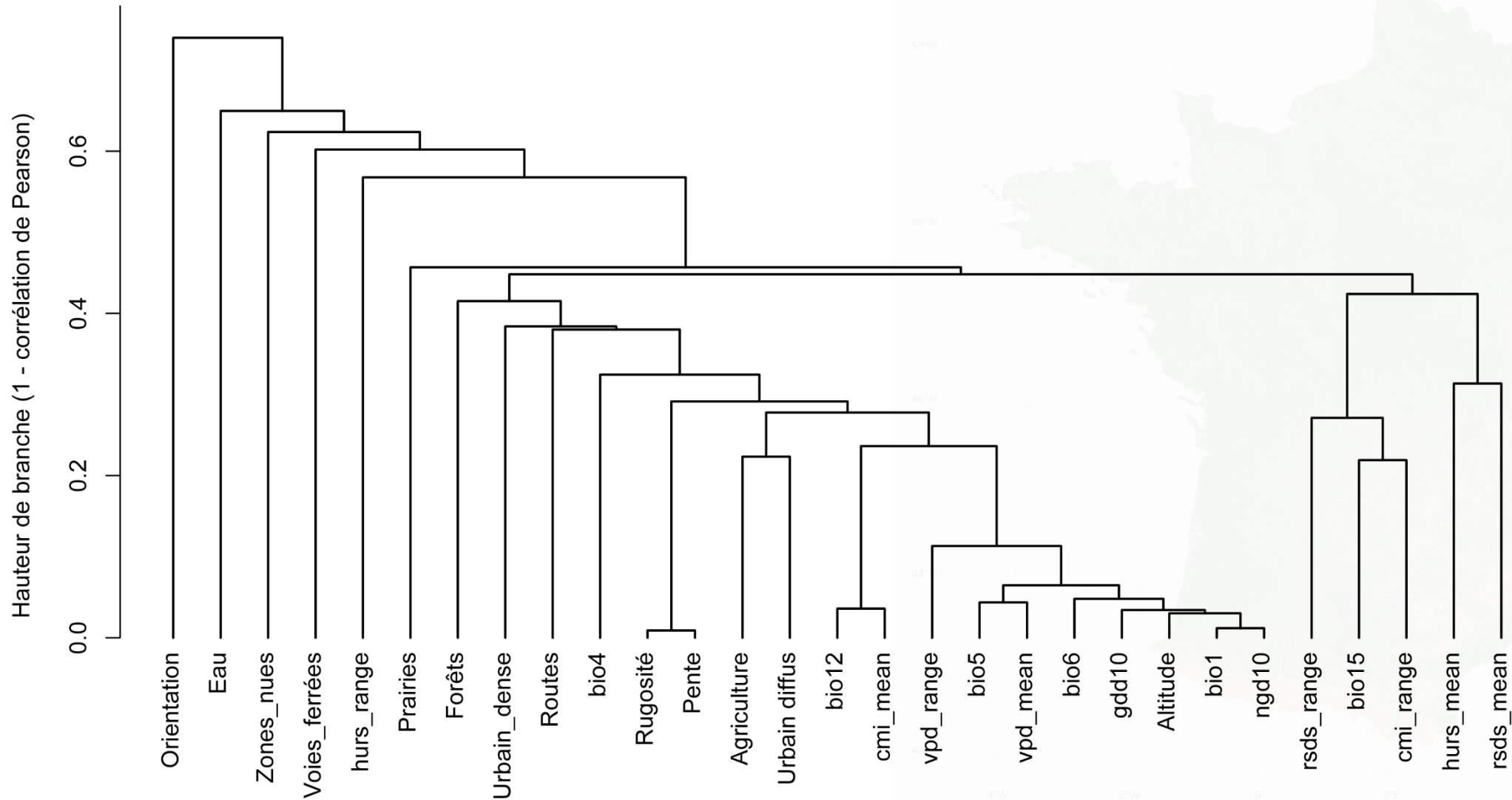
► 29 variables

Orientation
Eau
Zones_nues
Voies_ferrées
hurs_range
Prairies
Forêts
Urban_dense
Routes
bio4
Rugosité
Pente
Agriculture
Urban diffus
bio12
cmi_mean
vpd_range
bio5
vpd_mean
bio6
gdd10
Altitude
bio1
ngd10
rsds_range
bio15
cmi_range
hurs_mean
rsds_mean



Données environnementales

► 29 variables partiellement redondantes...

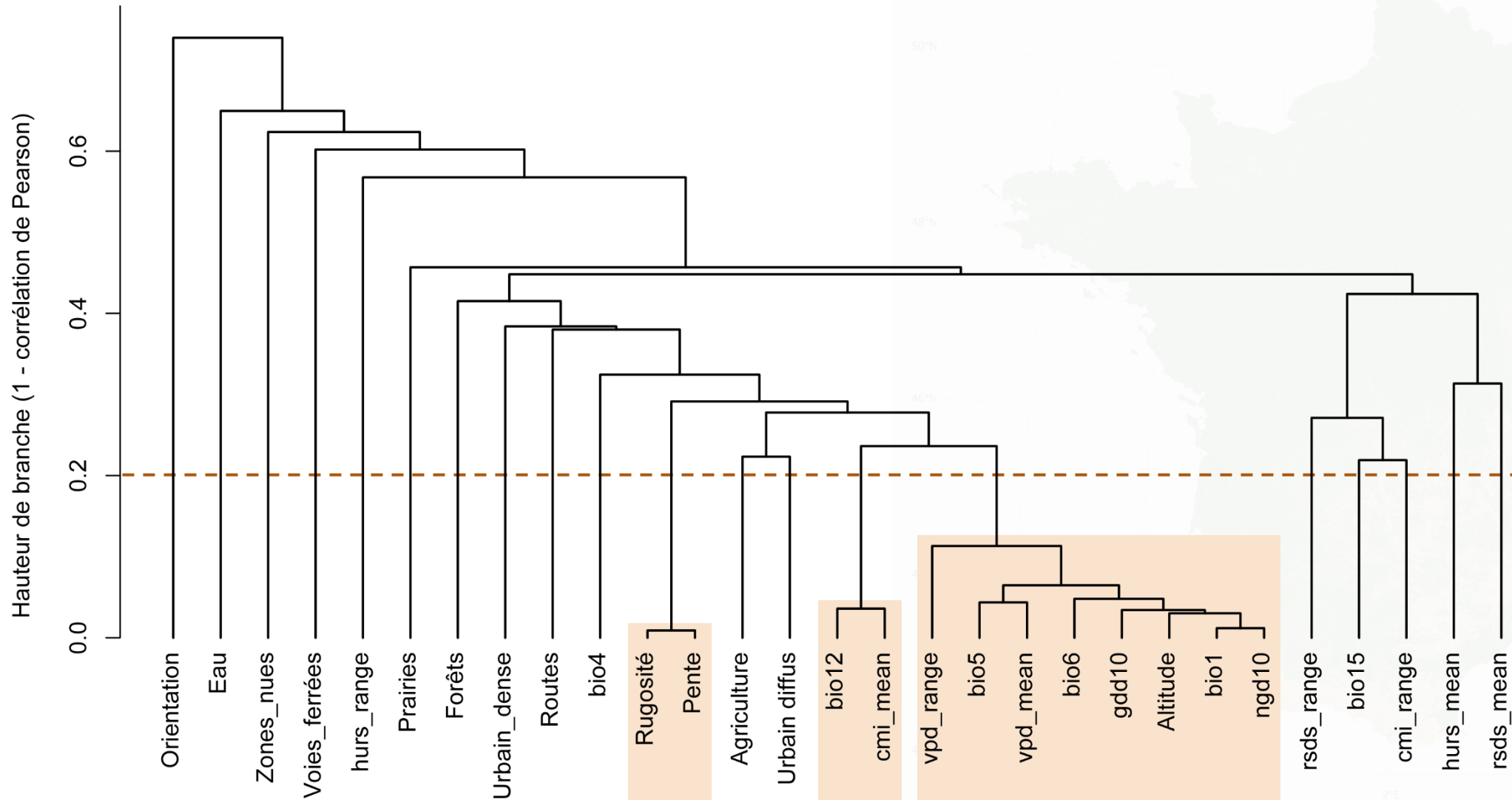


$VIF_{max} = 1147$ 🤯

Données environnementales

▶ 29 variables partiellement redondantes...

- 3 clusters (correlations > 80%)



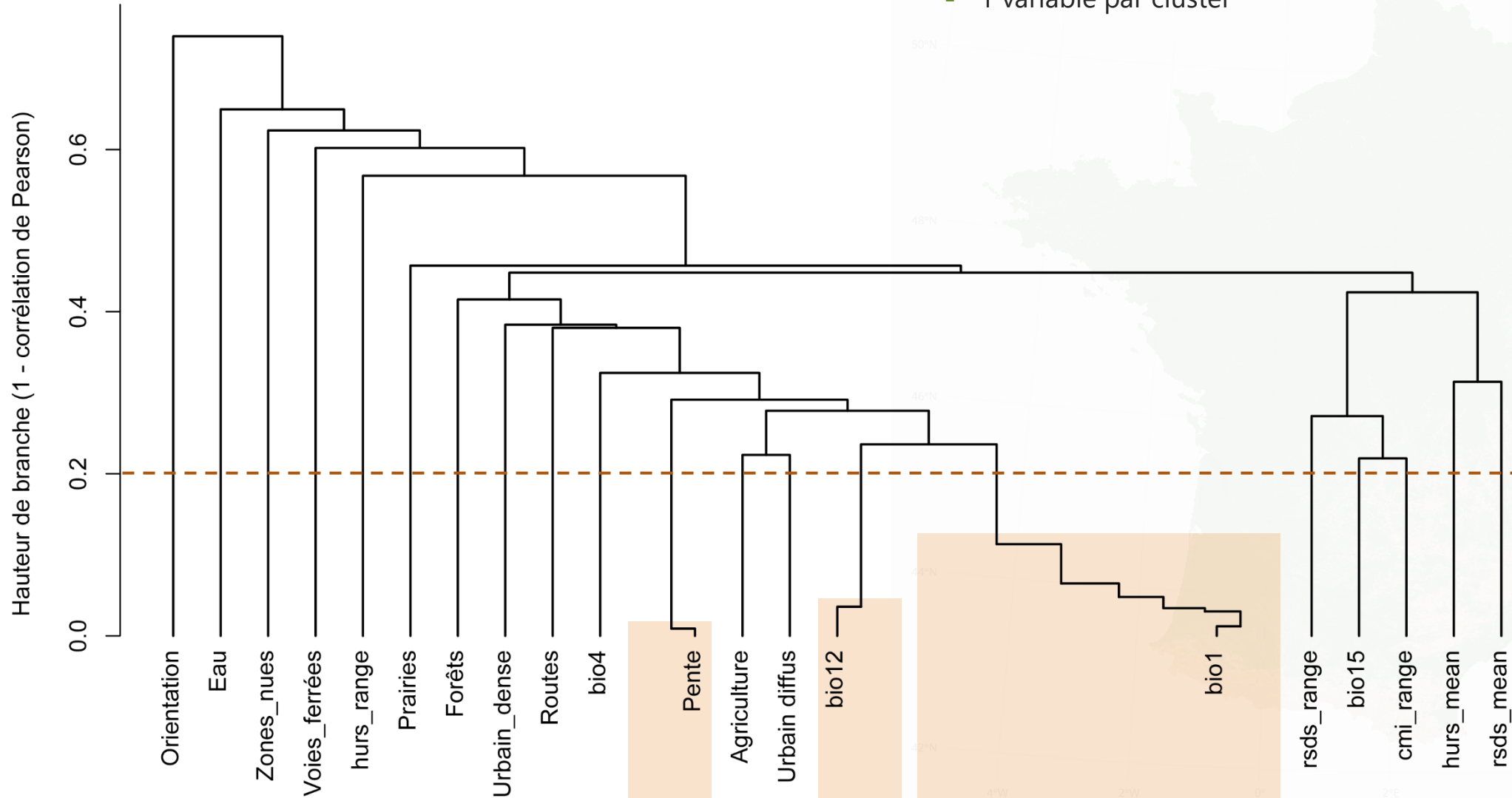
$VIF_{max} = 1147$



Données environnementales

▶ 29 variables partiellement redondantes...

- 3 clusters (correlations > 80%)
- 1 variable par cluster



$VIF_{max} = 1147$ 😱

$VIF_{max} = 7.5$ 😊

Données environnementales

▶ 29 variables partiellement redondantes...

- 3 clusters (correlations > 80%)
- 1 variable par cluster

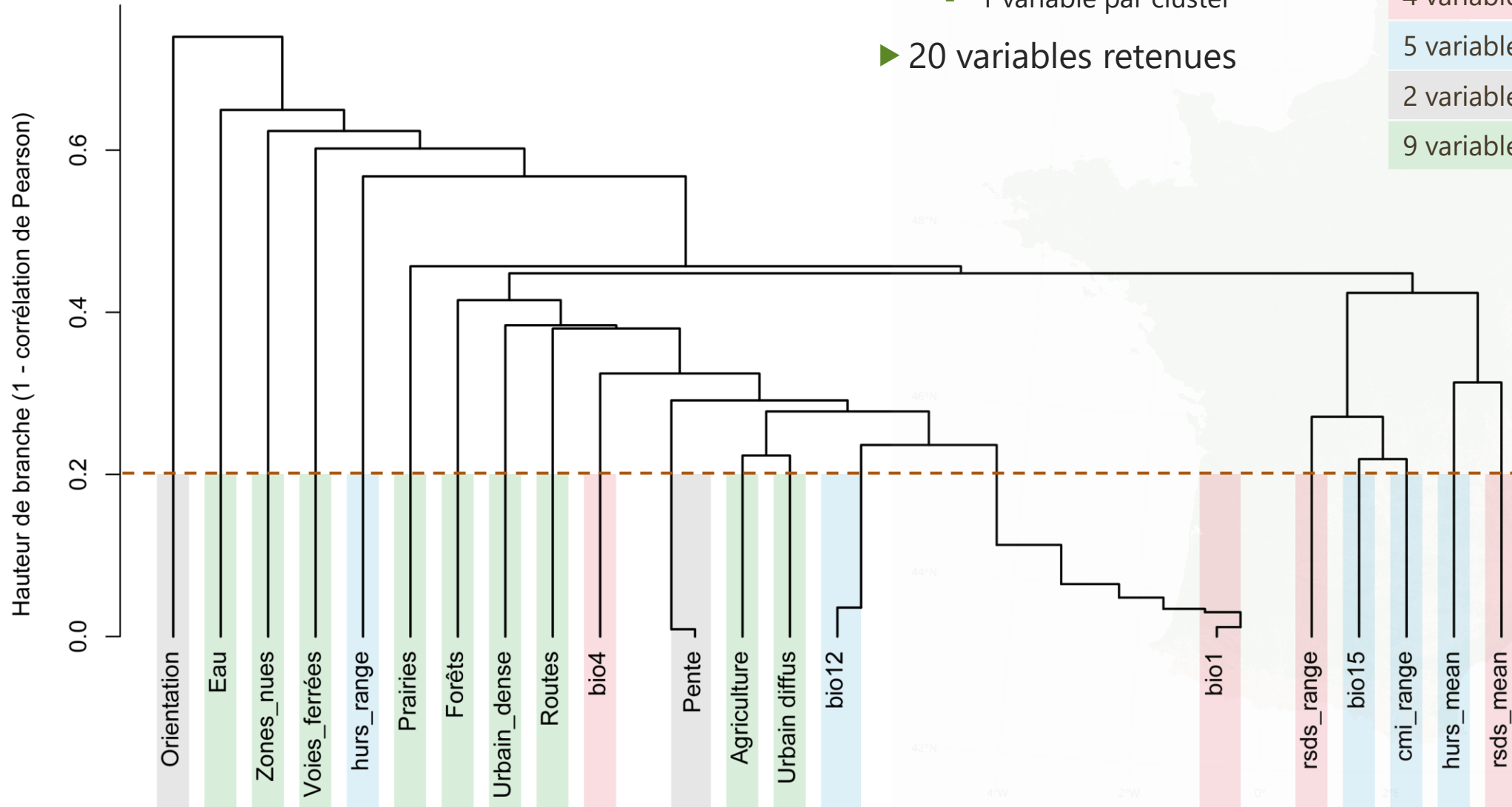
▶ 20 variables retenues

4 variables thermiques

5 variables hydriques

2 variables topographiques

9 variables d'occupation du sol



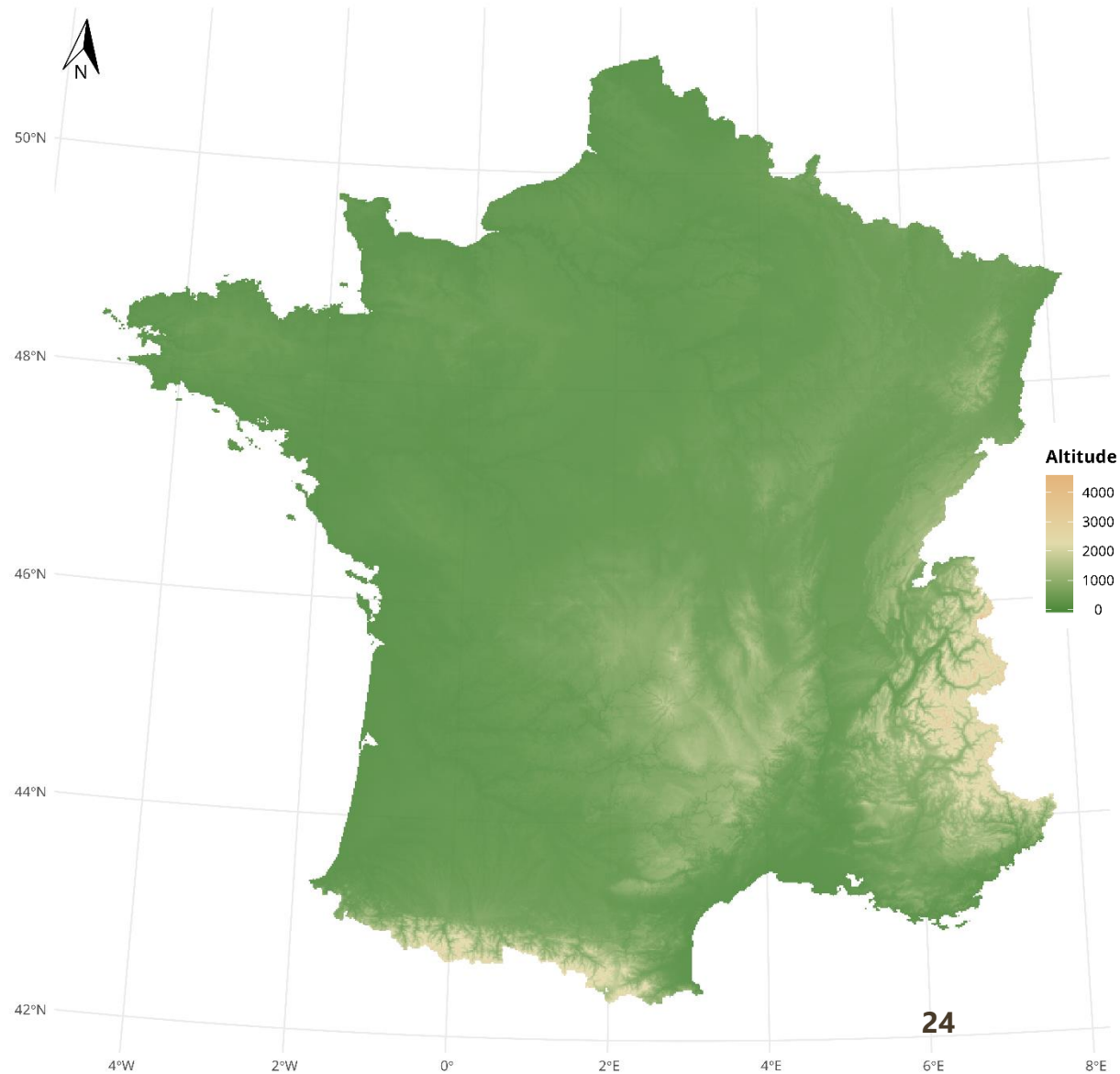
$$VIF_{max} = 1147$$



$$VIF_{max} = 7.5$$



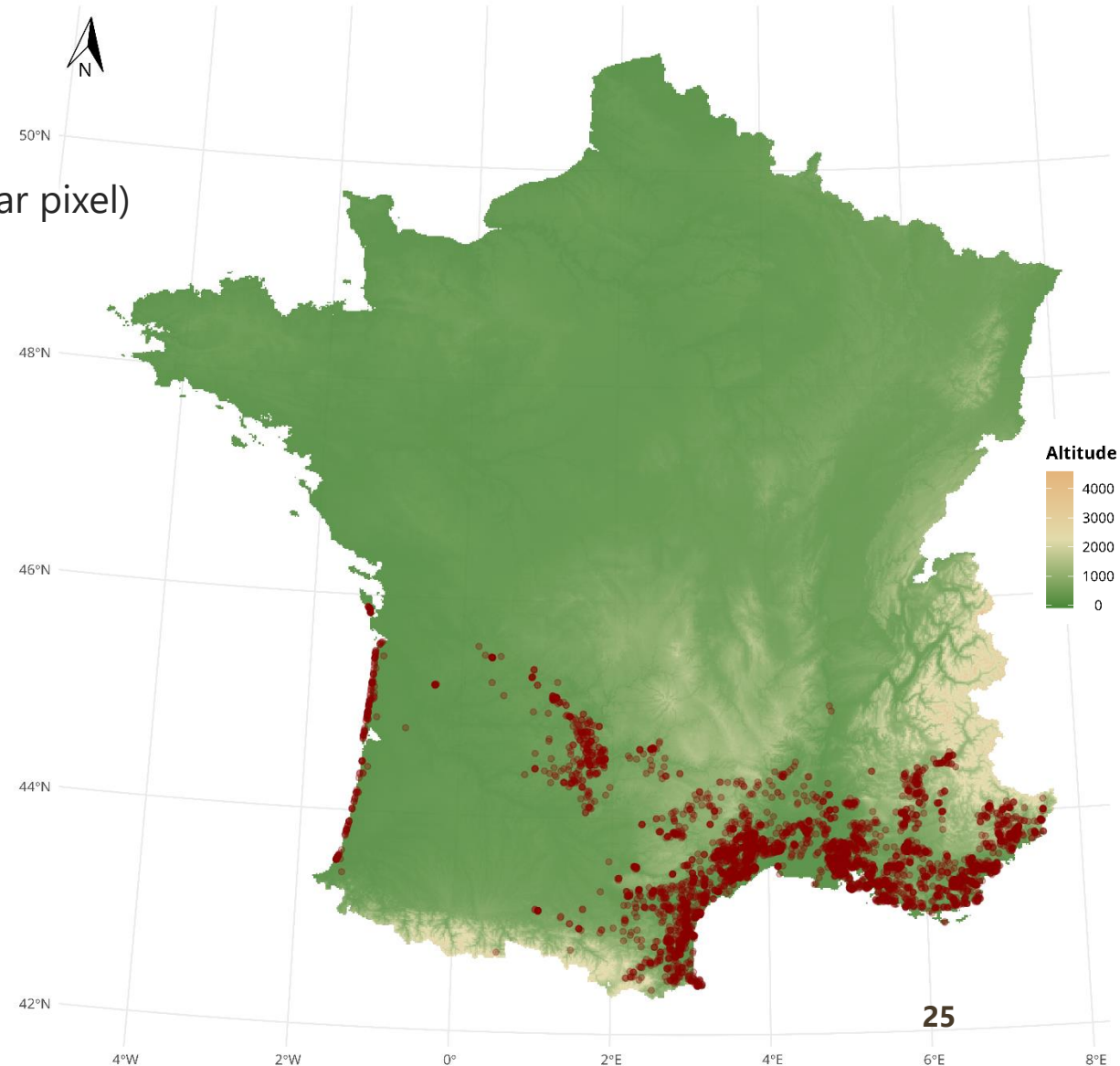
Données d'occurrence



Données d'occurrence

► Présence

- > 12000 données (SINP 2000-2024)
- 5893 données retenues (une donnée de présence par pixel)



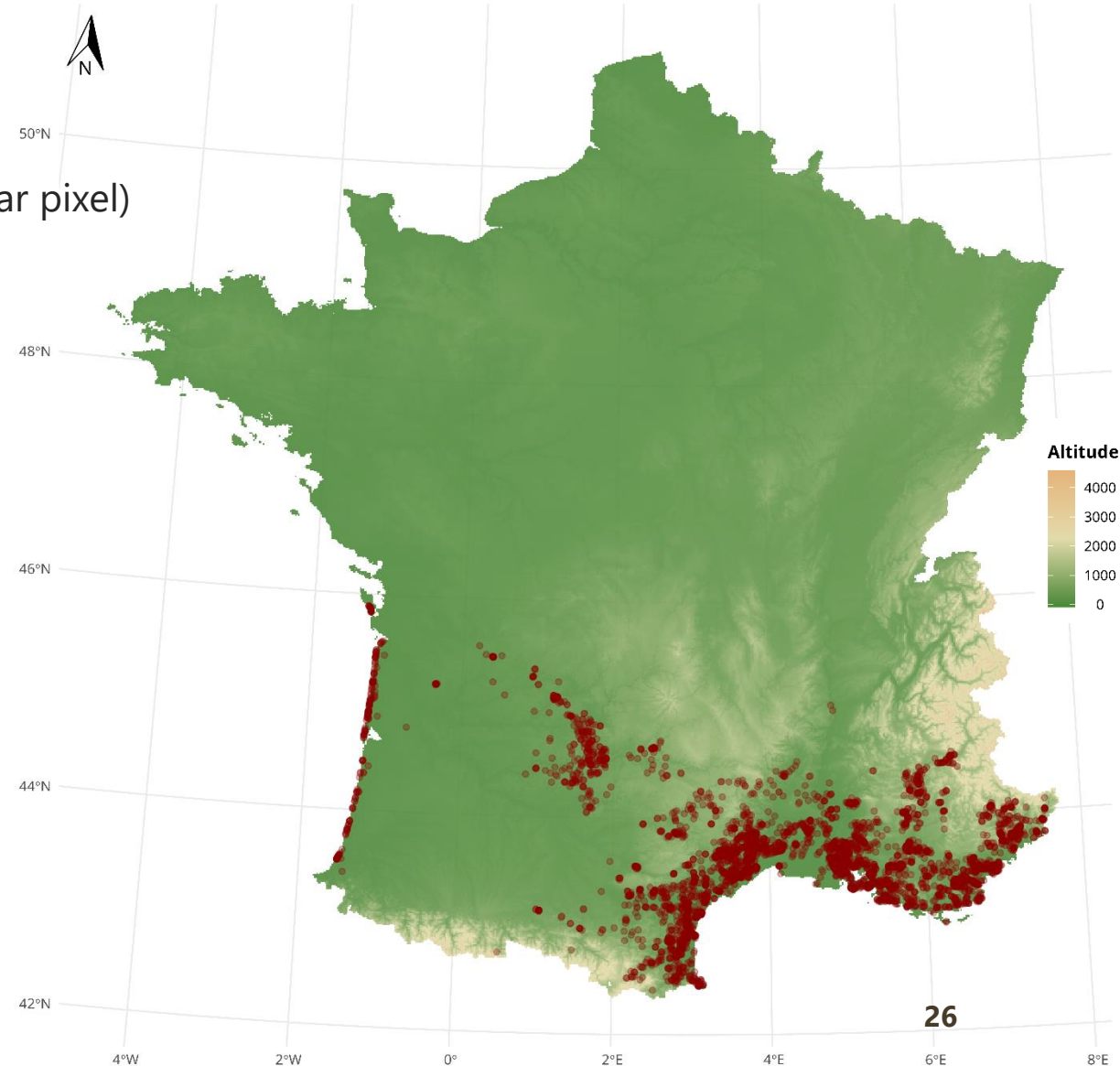
Données d'occurrence

► Présence

- > 12000 données (SINP 2000-2024)
- 5893 données retenues (une donnée de présence par pixel)

► Pas de réelles données d'absence

- Inventaires non-protocolés



Données d'occurrence

► Présence

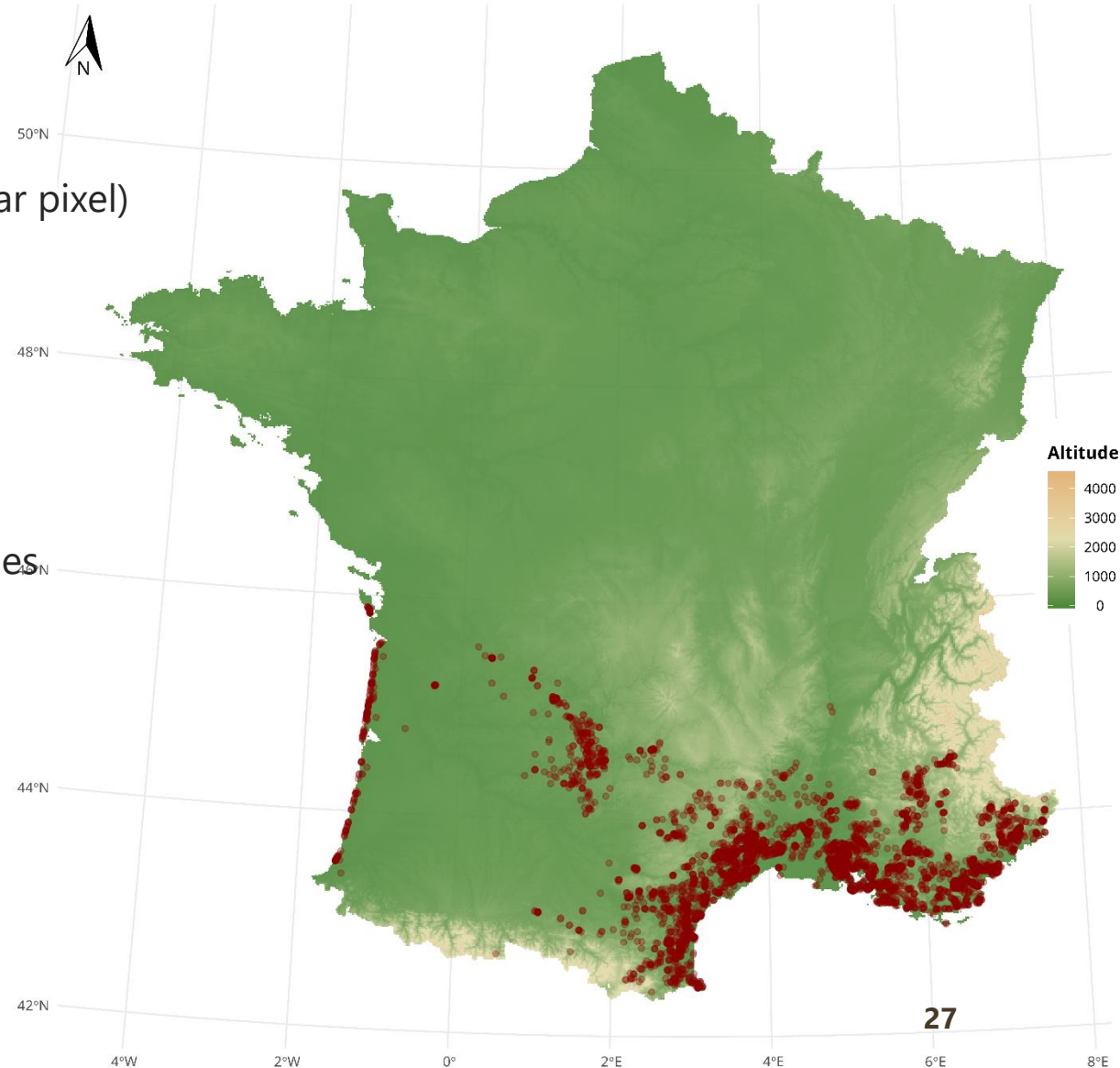
- > 12000 données (SINP 2000-2024)
- 5893 données retenues (une donnée de présence par pixel)

► Pas de réelles données d'absence

- Inventaires non-protocoles

► Données de « fond » (« pseudo-absences »)

- Ensemble des conditions environnementales auxquelles l'espèce est susceptible d'être confrontée



Données d'occurrence

► Présence

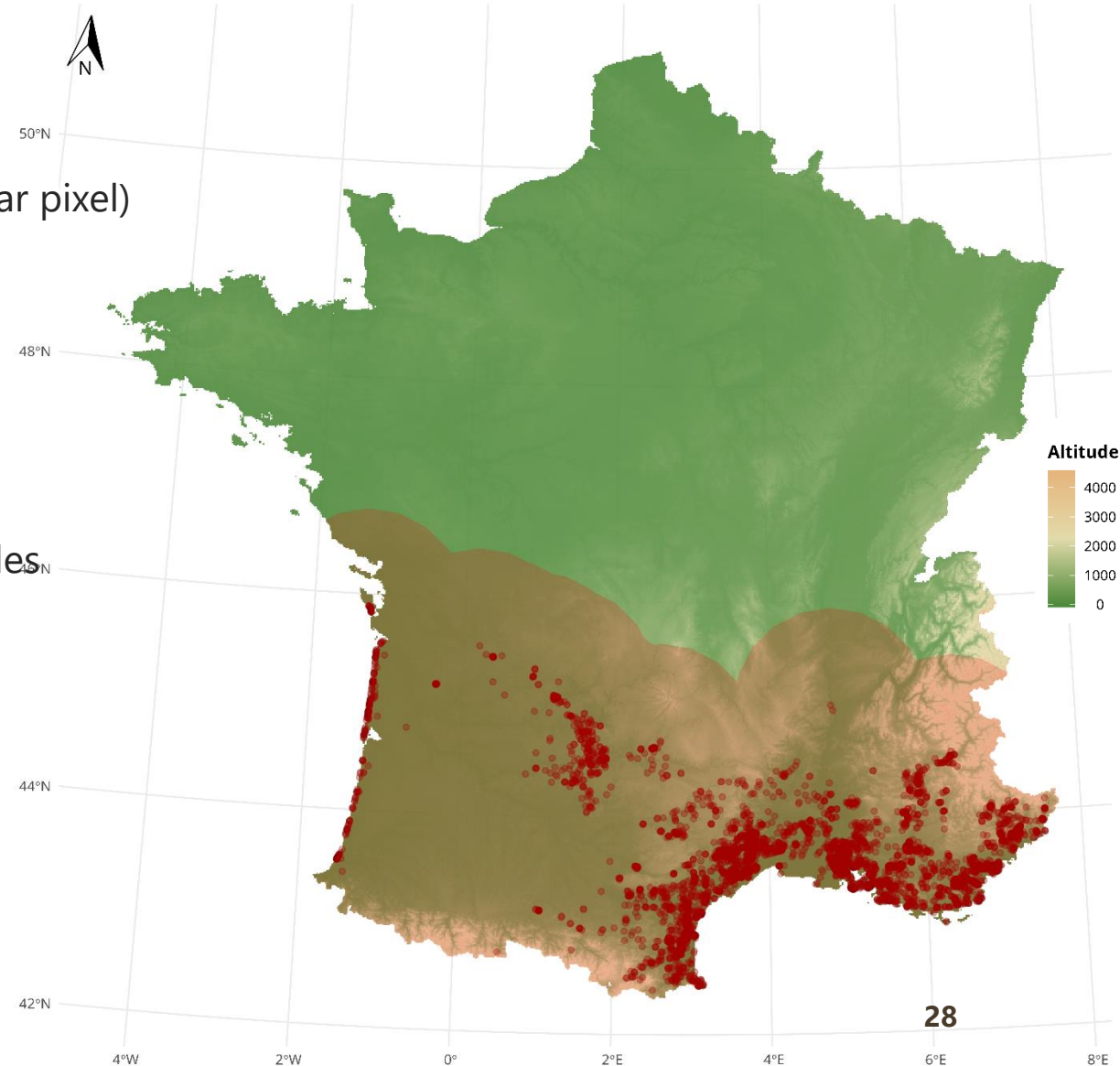
- > 12000 données (SINP 2000-2024)
- 5893 données retenues (une donnée de présence par pixel)

► Pas de réelles données d'absence

- Inventaires non-protocoles

► Données de « fond » (« pseudo-absences »)

- Ensemble des conditions environnementales auxquelles l'espèce est susceptible d'être confrontée



Données d'occurrence

► Présence

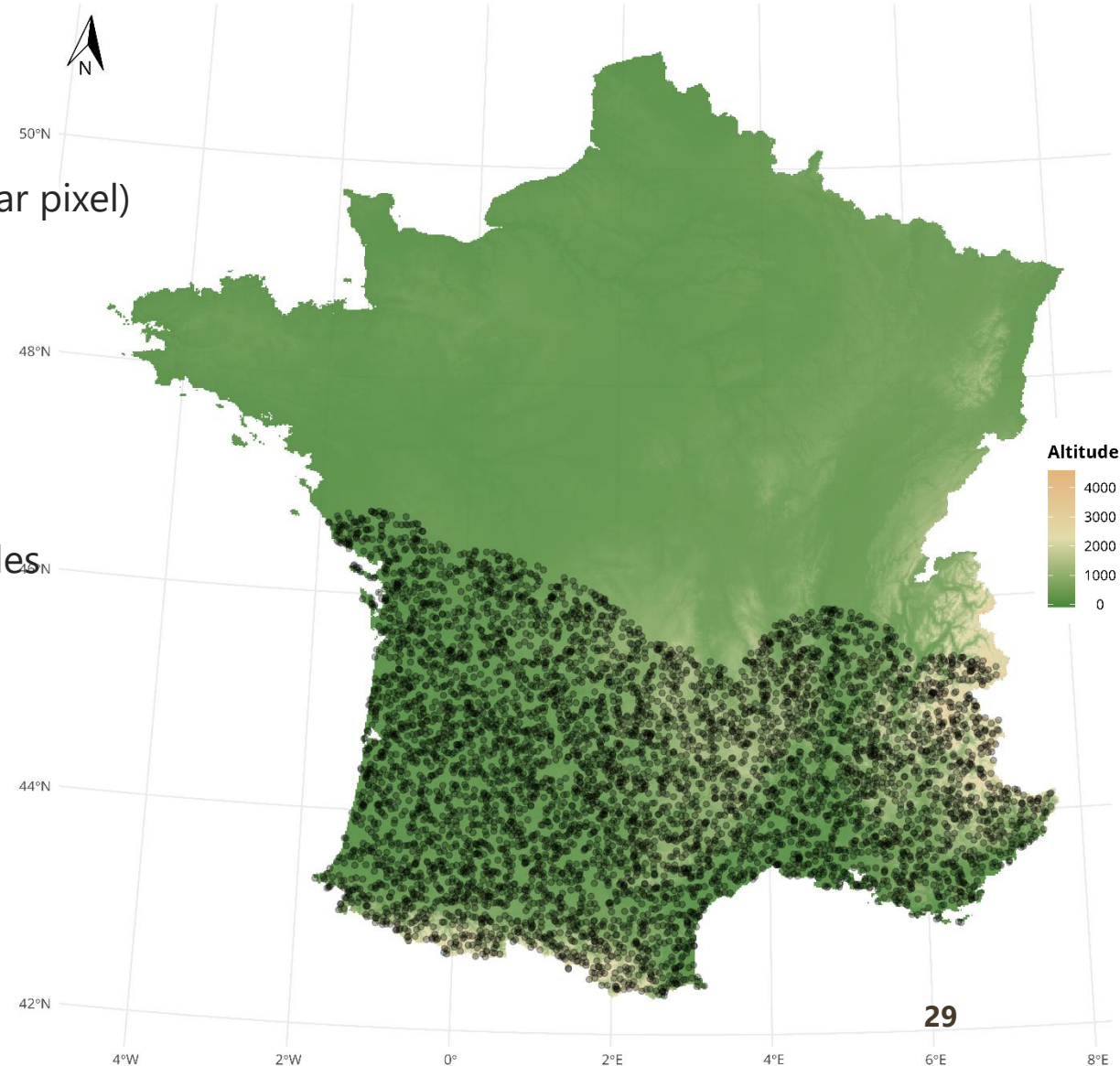
- > 12000 données (SINP 2000-2024)
- 5893 données retenues (une donnée de présence par pixel)

► Pas de réelles données d'absence

- Inventaires non-protocoles

► Données de « fond » (« pseudo-absences »)

- Ensemble des conditions environnementales auxquelles l'espèce est susceptible d'être confrontée
- 5893 données de fond tirées au hasard



Outils de modélisation

Variables
environnementales

Données d'occurrence
Données de fond
Données de présence

Outils de modélisation

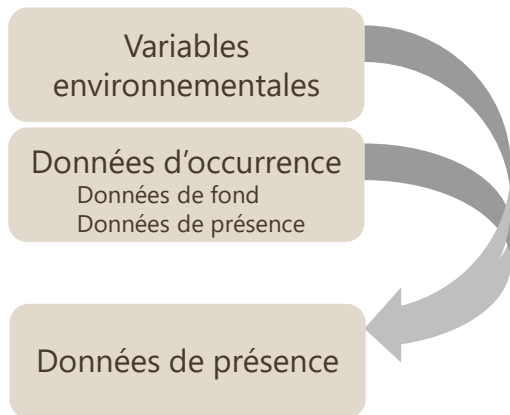
- ▶ Apprentissage automatique supervisé (IA)

Variables
environnementales

Données d'occurrence
Données de fond
Données de présence

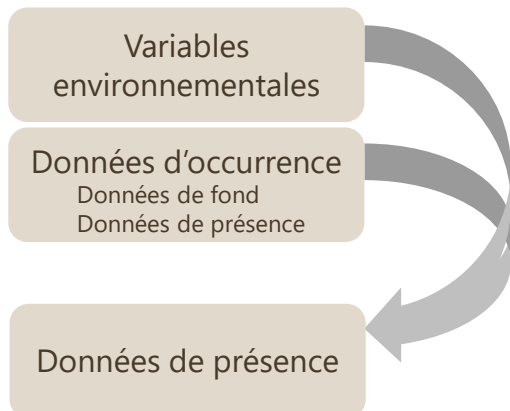
Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence



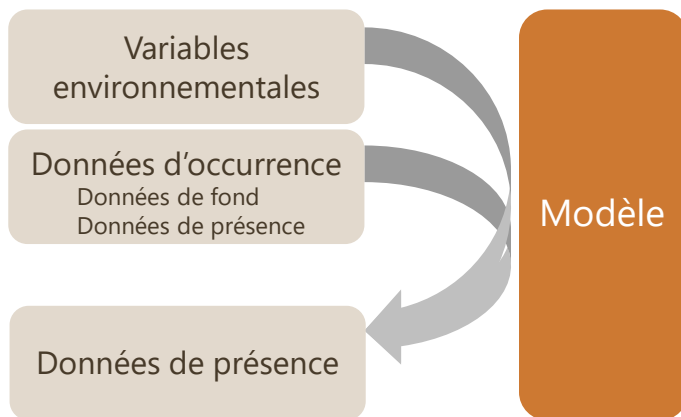
Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...



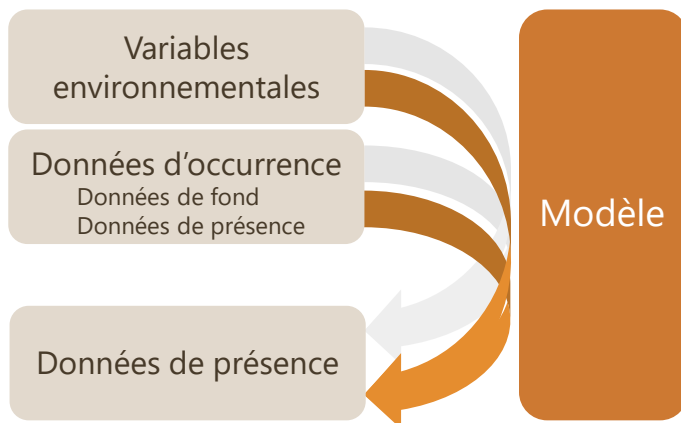
Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence



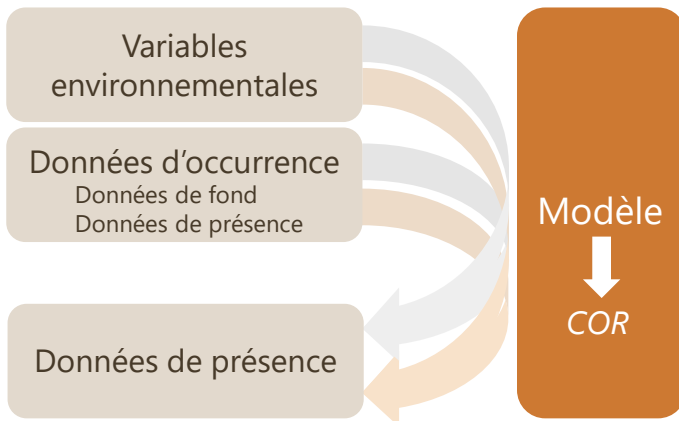
Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence
 - 2) Valider le modèle sur 30% des données d'occurrence



Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence
 - 2) Valider le modèle sur 30% des données d'occurrence
 - 3) Evaluer la performance du modèle (COR)

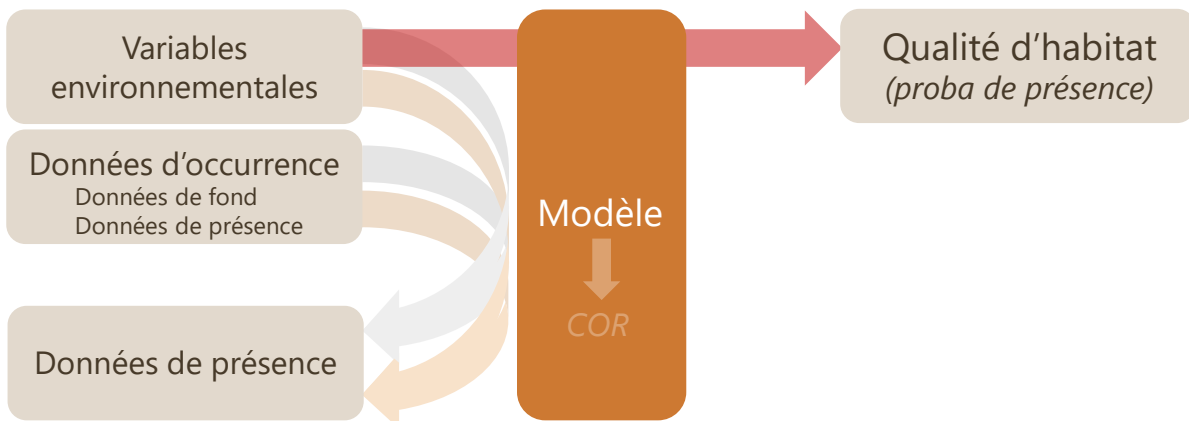


COR

- Probabilité qu'une observation soit correctement prédite par l'algorithme
- Varie entre 0 et 1
- Objectif : > 0.9

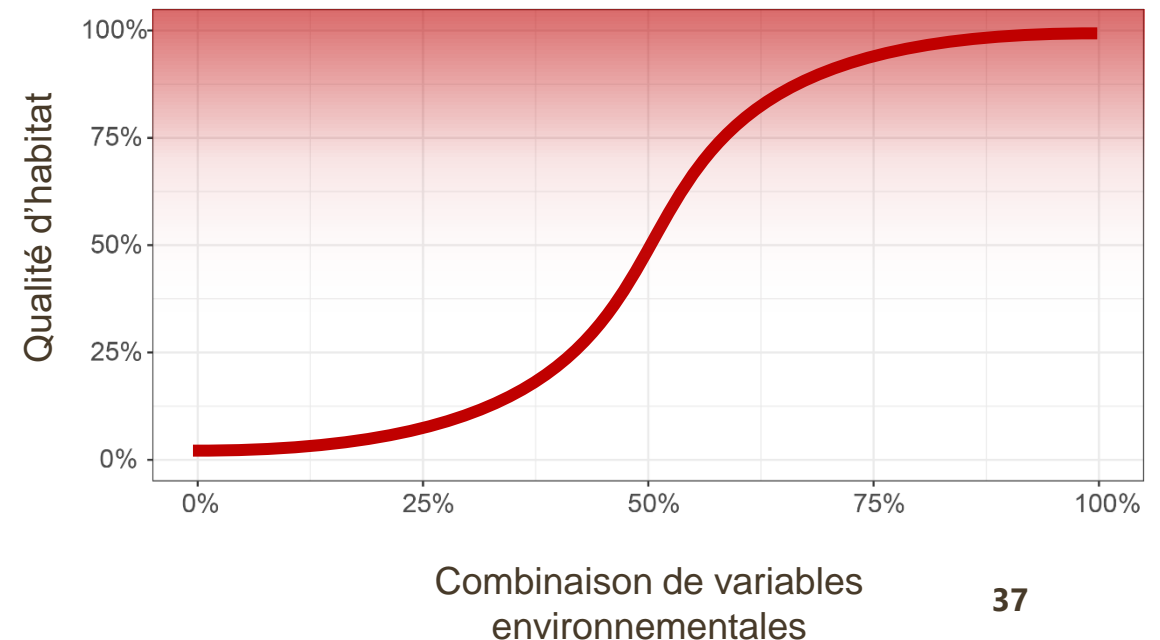
Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence
 - 2) Valider le modèle sur 30% des données d'occurrence
 - 3) Evaluer la performance du modèle (COR)
 - 4) Prédire la qualité d'habitat (extrapolation)



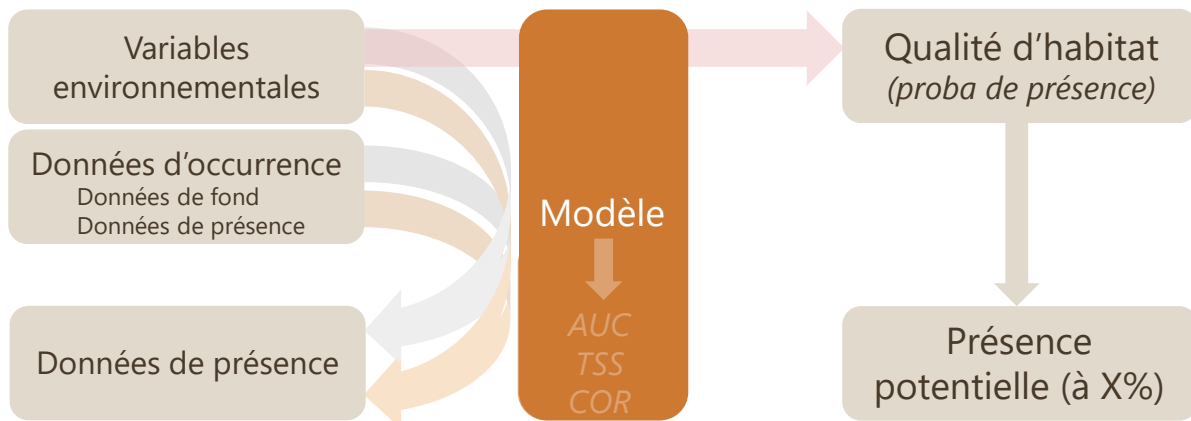
COR

- Probabilité qu'une observation soit correctement prédite par l'algorithme
- Varie entre 0 et 1
- Objectif : > 0.9



Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence
 - 2) Valider le modèle sur 30% des données d'occurrence
 - 3) Evaluer la performance du modèle (COR)
 - 4) Prédire la qualité d'habitat (extrapolation)
 - 5) Définir un seuil optimal de prédiction (OPT)

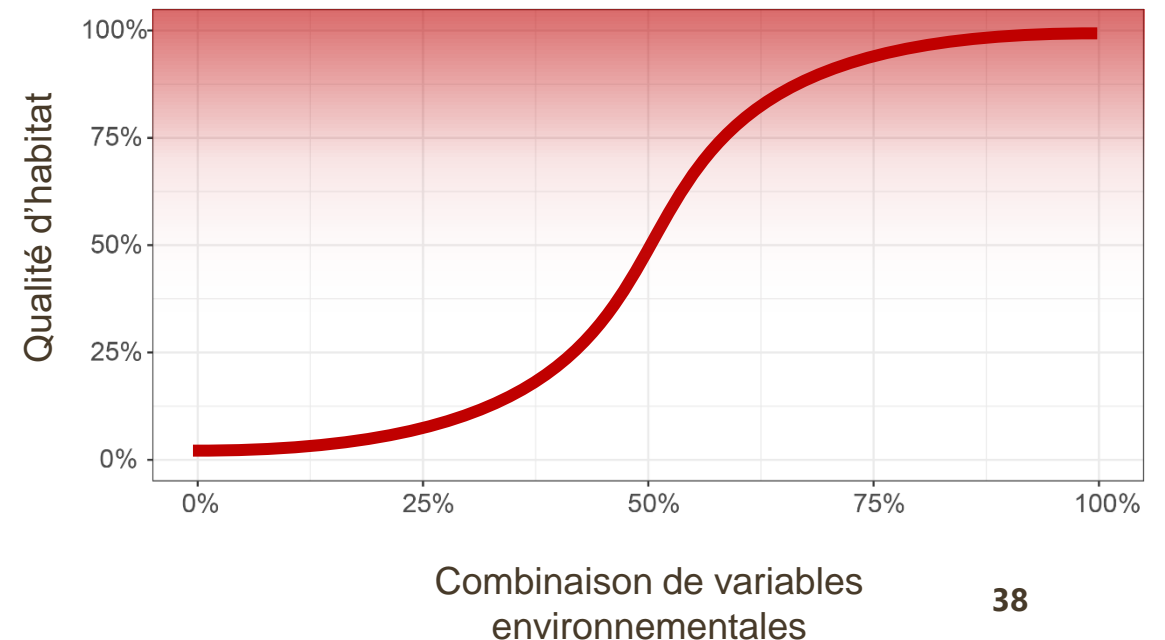


COR

- Probabilité qu'une observation soit correctement prédite par l'algorithme
- Varie entre 0 et 1
- Objectif : > 0.9

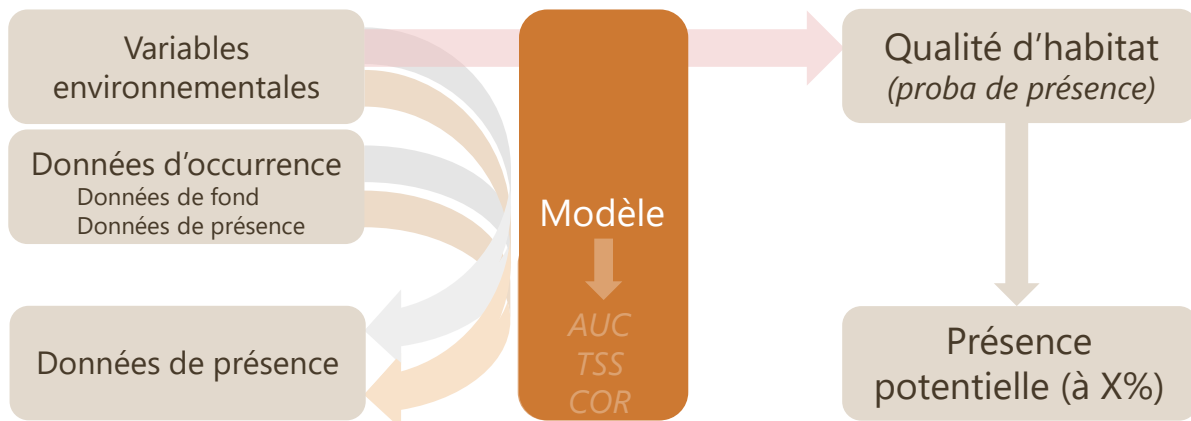
OPT

- Seuil de qualité d'habitat à partir duquel l'espèce sera considérée comme présente sur un pixel
- Estimé automatiquement par l'algorithme



Outils de modélisation

- ▶ Apprentissage automatique supervisé (IA)
 - Entraîner un modèle à prédire avec précision les données de présence
- ▶ 9 algorithmes
 - MAXENT, GLM, RandomForest, ...
- ▶ 5 étapes
 - 1) Entraîner le modèle sur 70% des données d'occurrence
 - 2) Valider le modèle sur 30% des données d'occurrence
 - 3) Evaluer la performance du modèle (COR)
 - 4) Prédire la qualité d'habitat (extrapolation)
 - 5) Définir un seuil optimal de prédiction (OPT)

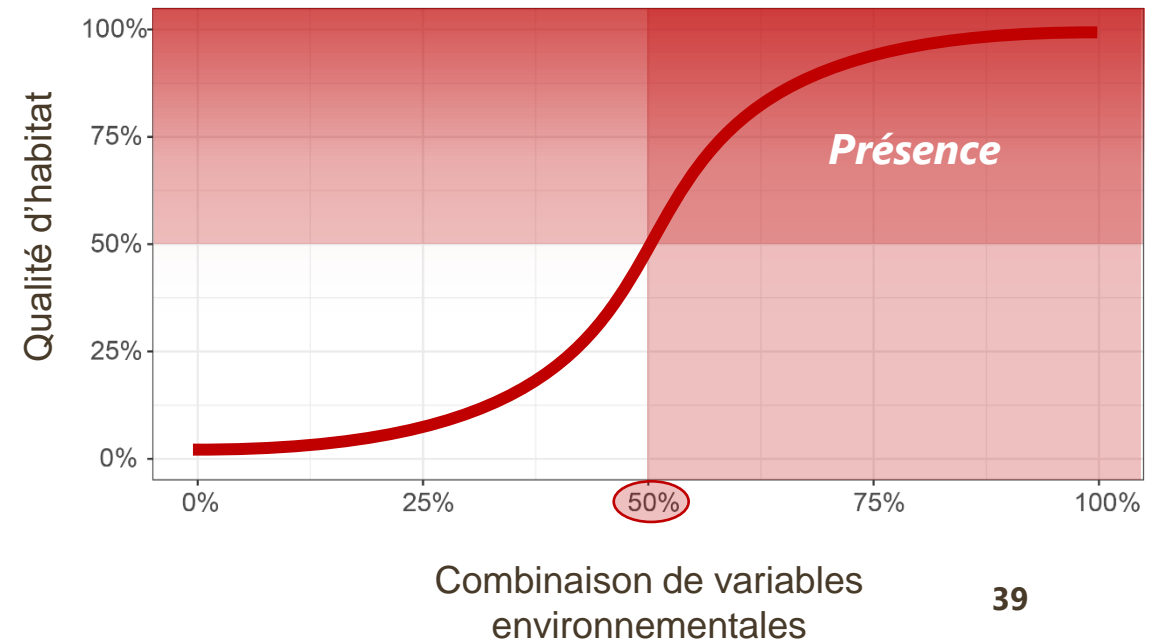


COR

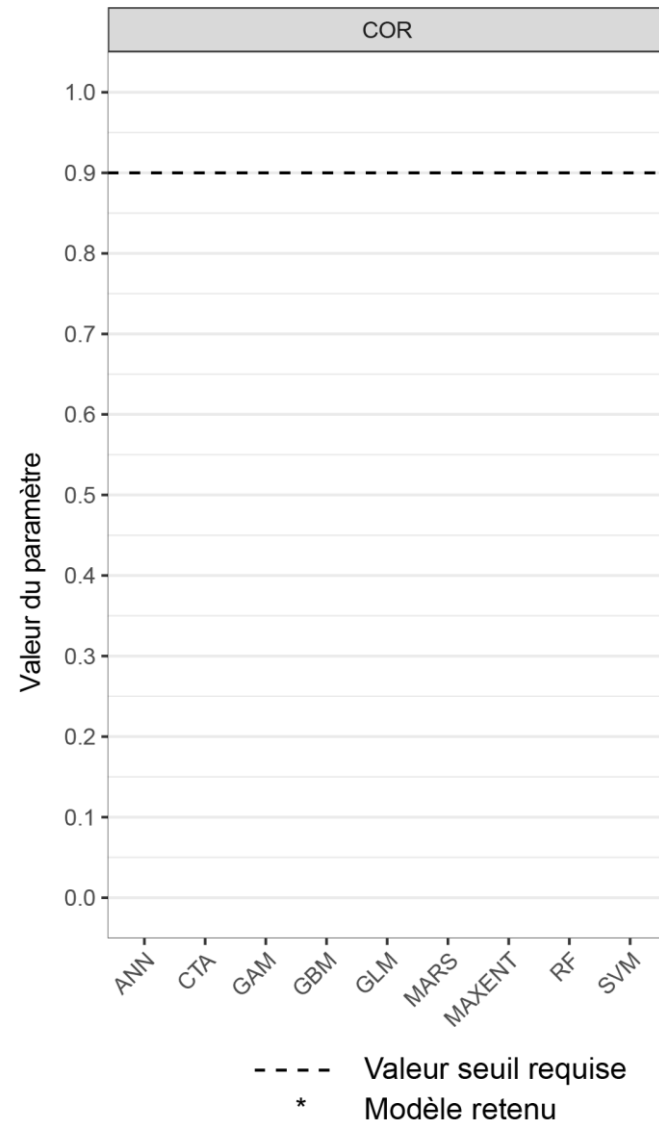
- Probabilité qu'une observation soit correctement prédite par l'algorithme
- Varie entre 0 et 1
- Objectif : > 0.9

OPT

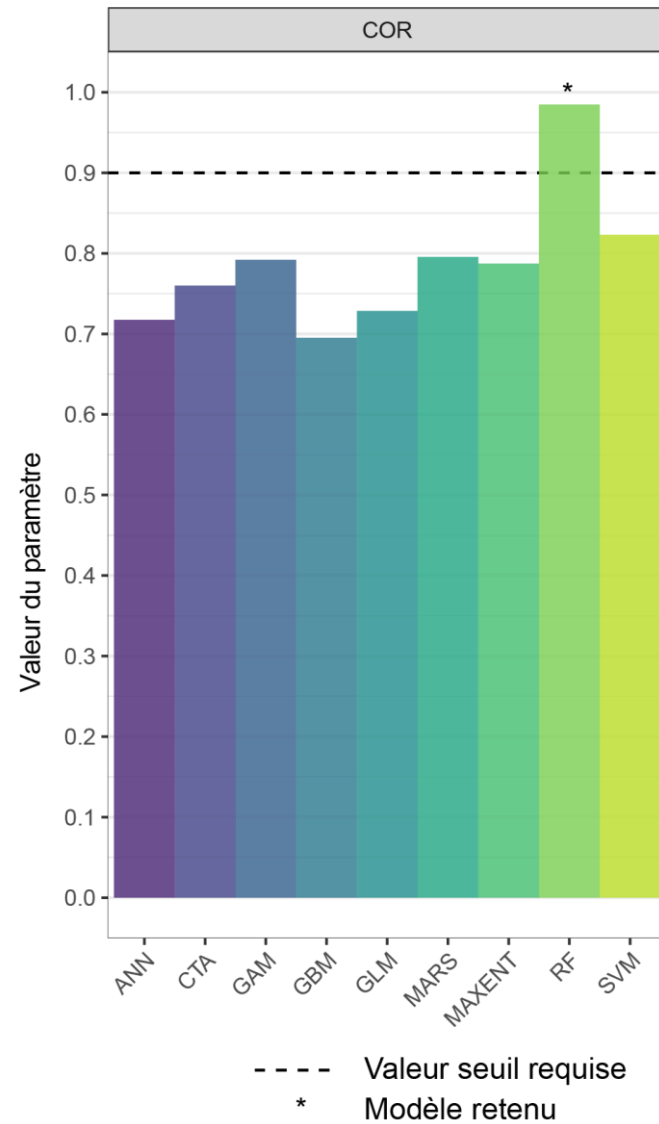
- Seuil de qualité d'habitat à partir duquel l'espèce sera considérée comme présente sur un pixel
- Estimé automatiquement par l'algorithme



Performance des modèles

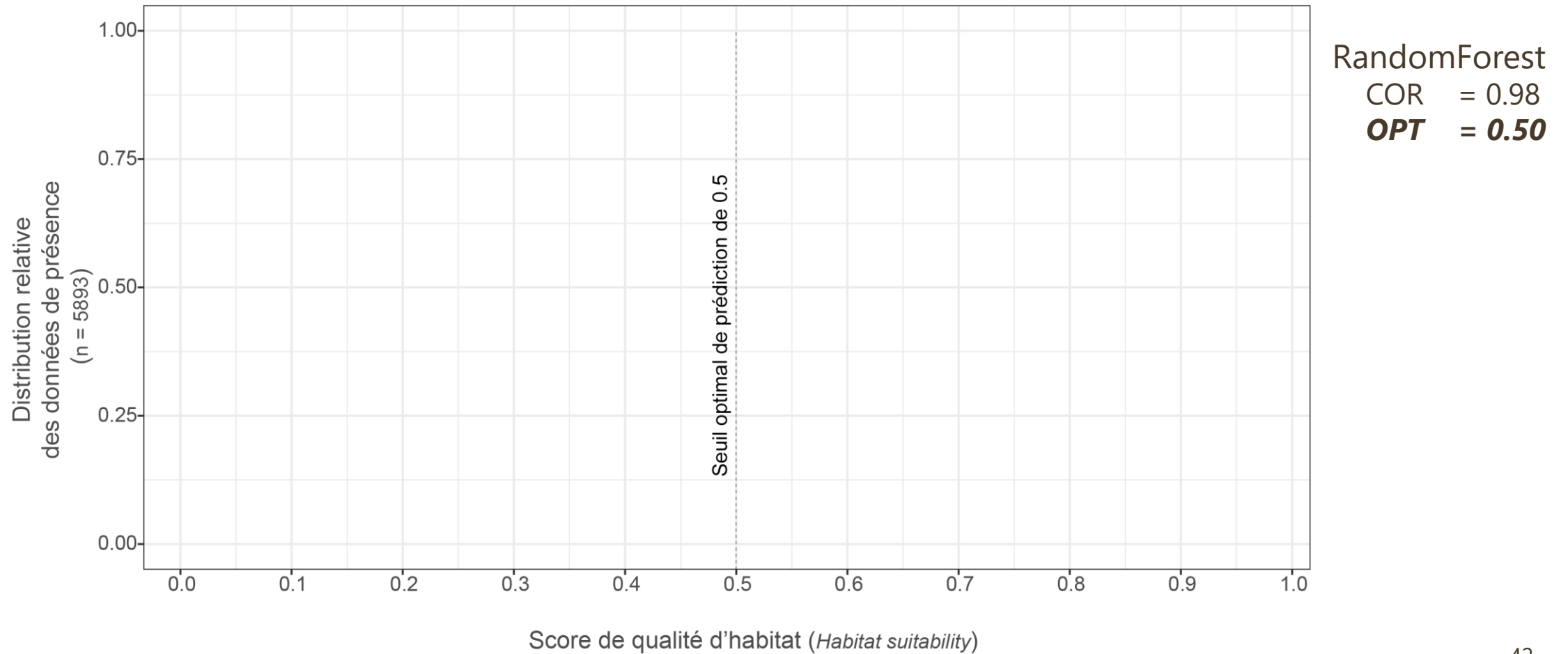


Performance des modèles

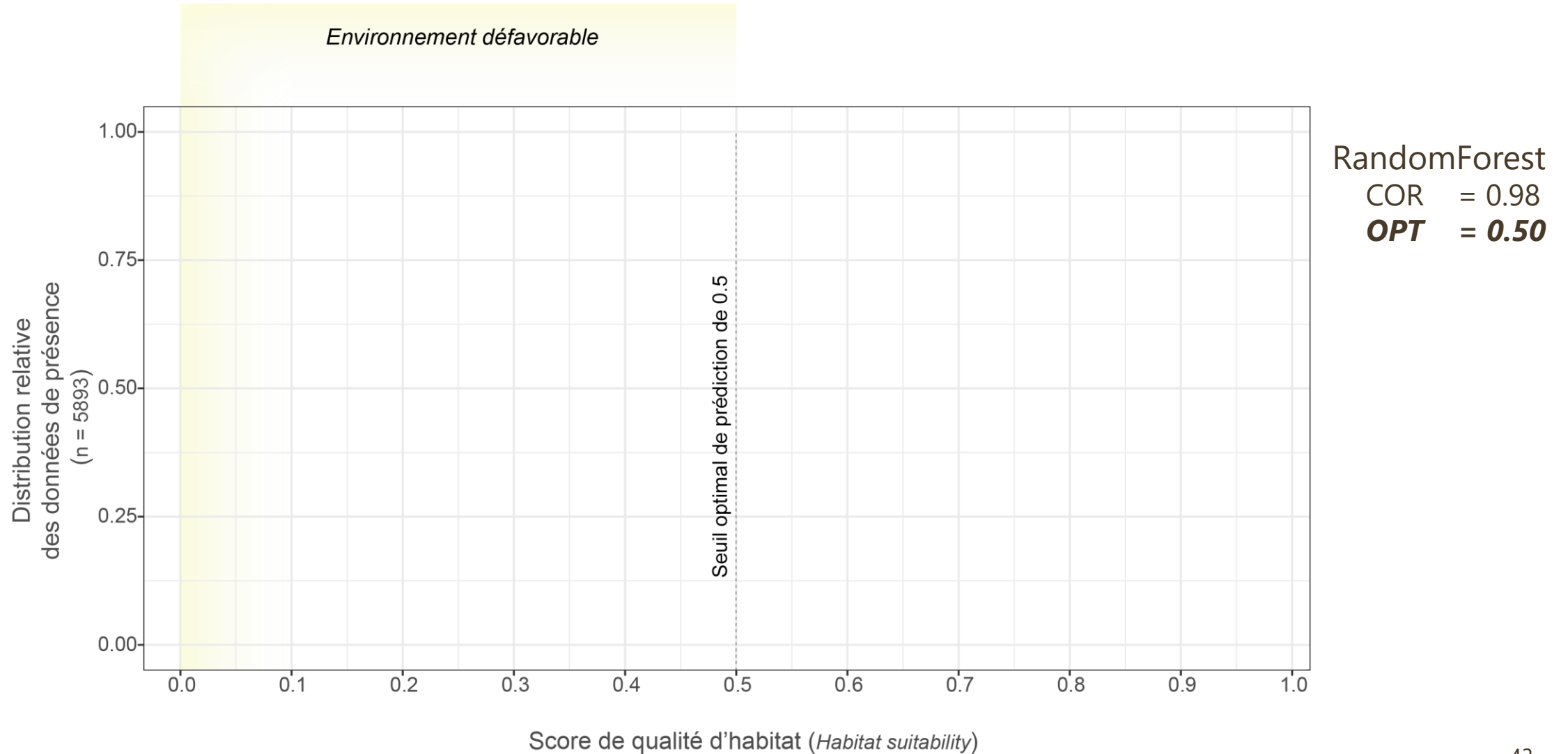


RandomForest
COR = 0.98

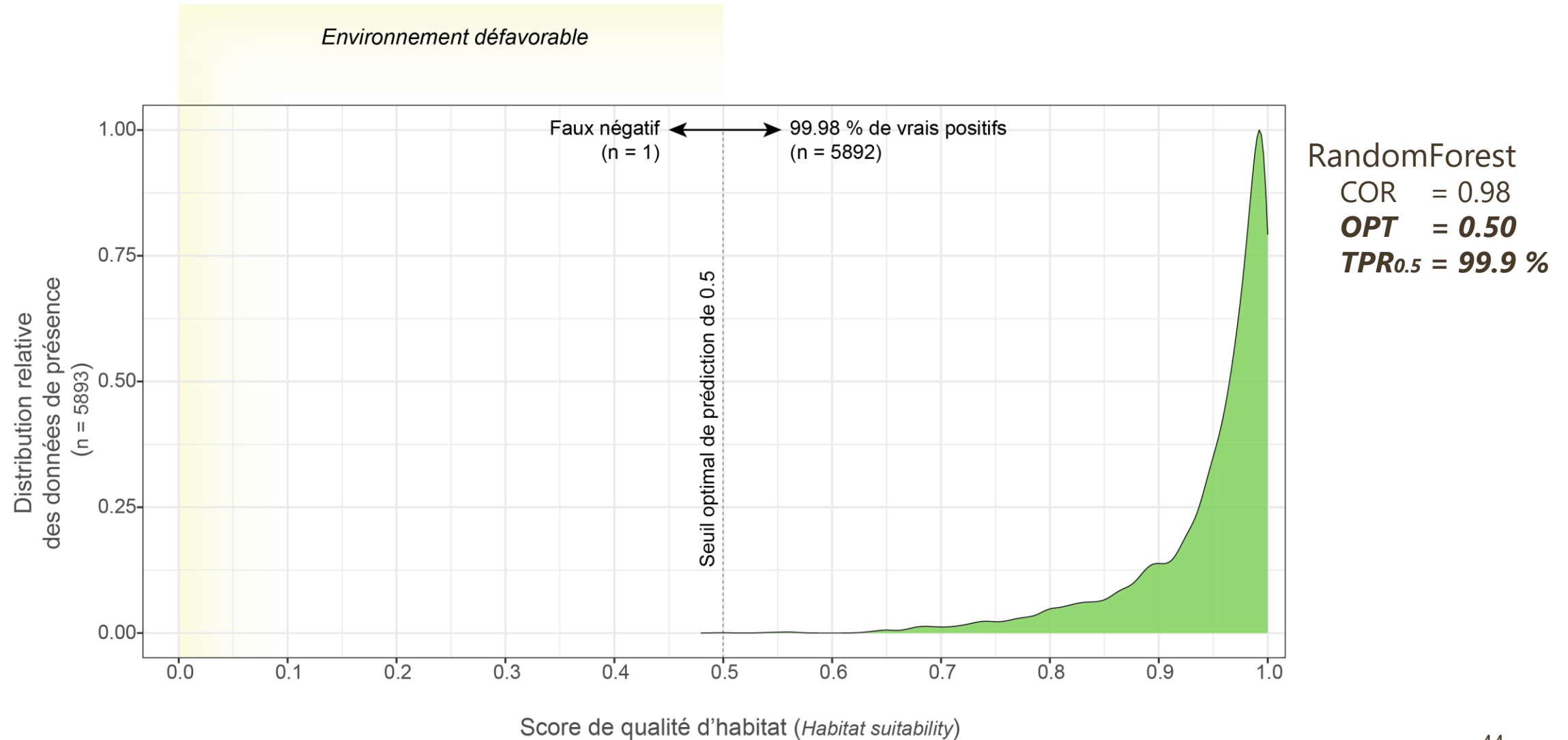
Sensibilité et seuillage du modèle RandomForest



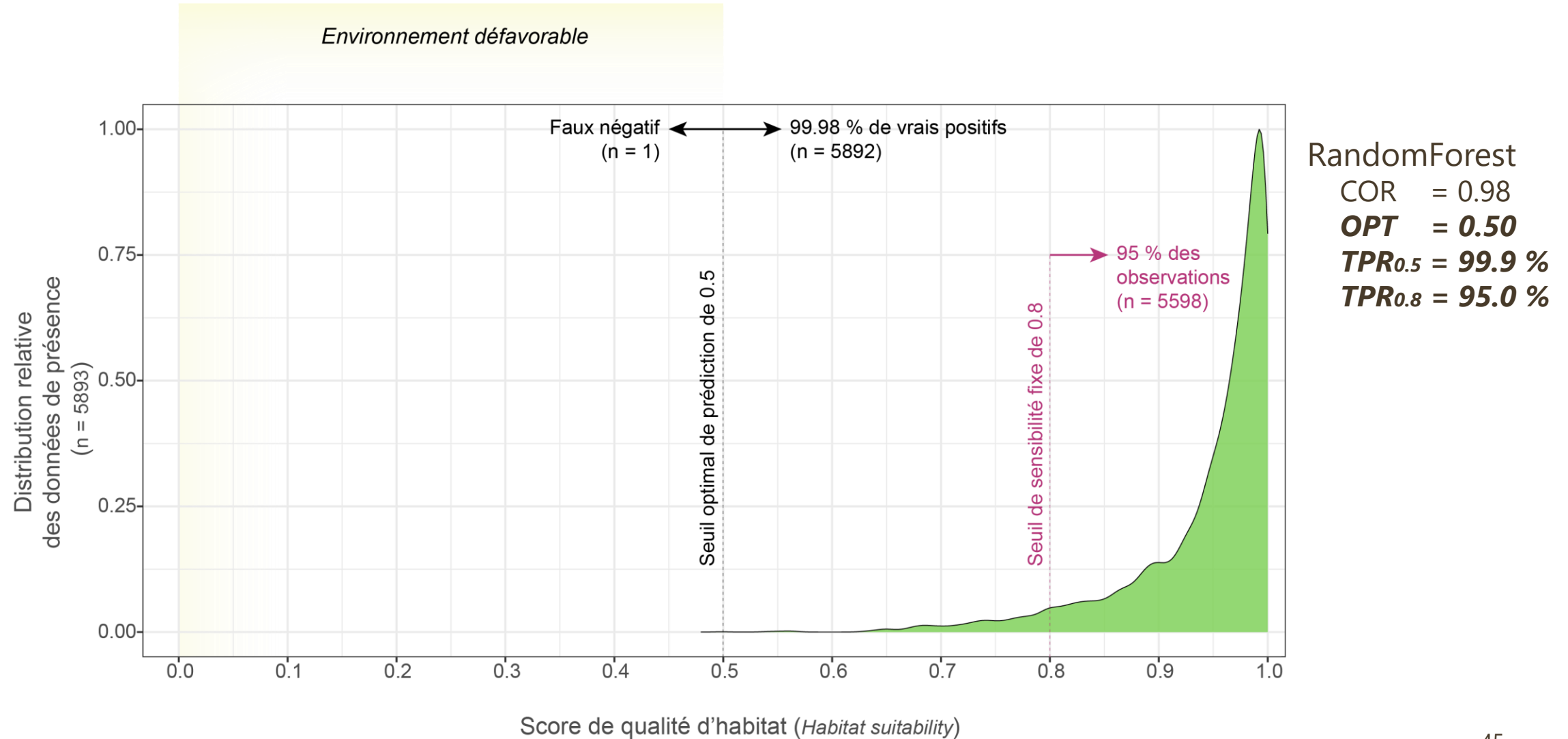
Sensibilité et seuillage du modèle RandomForest



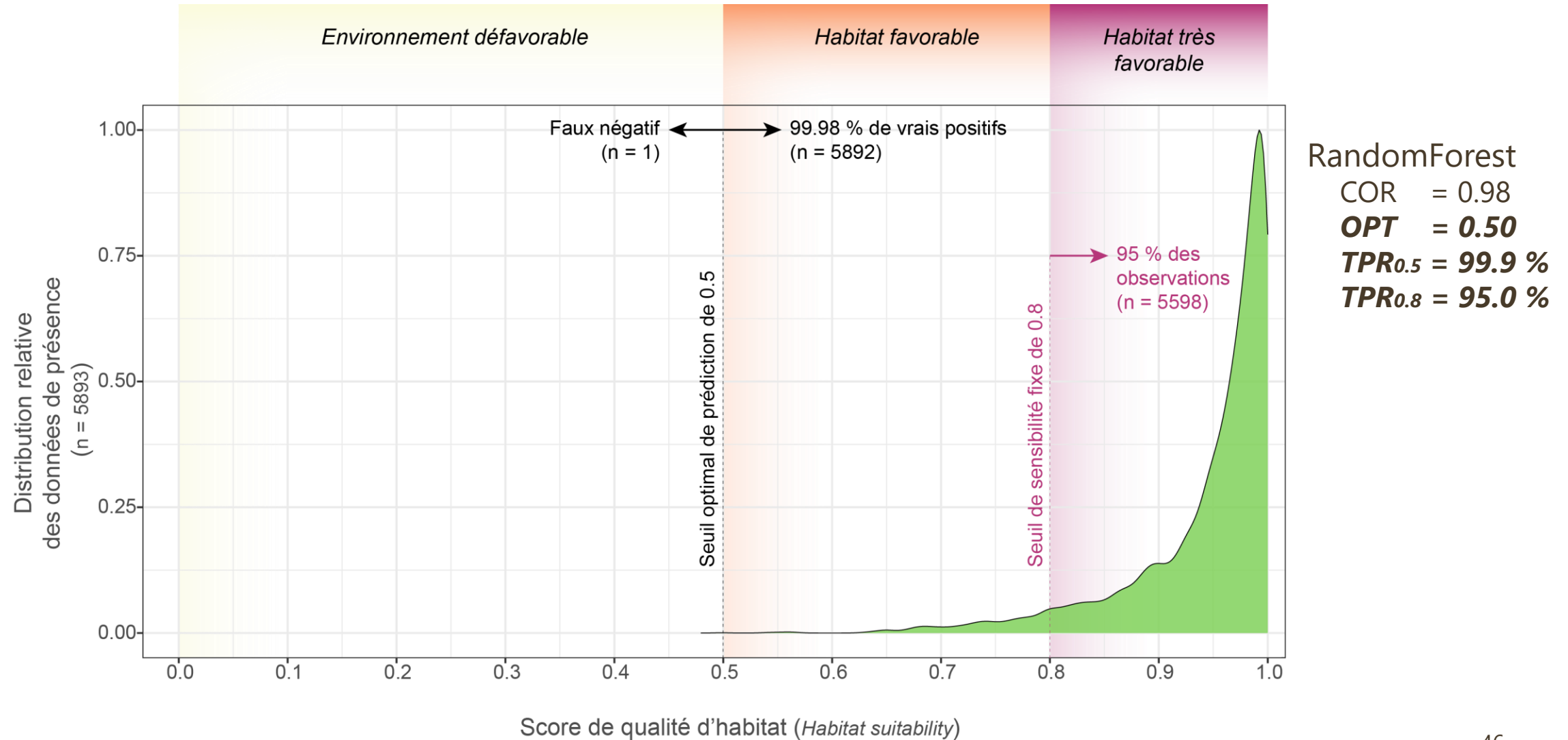
Sensibilité et seuillage du modèle RandomForest



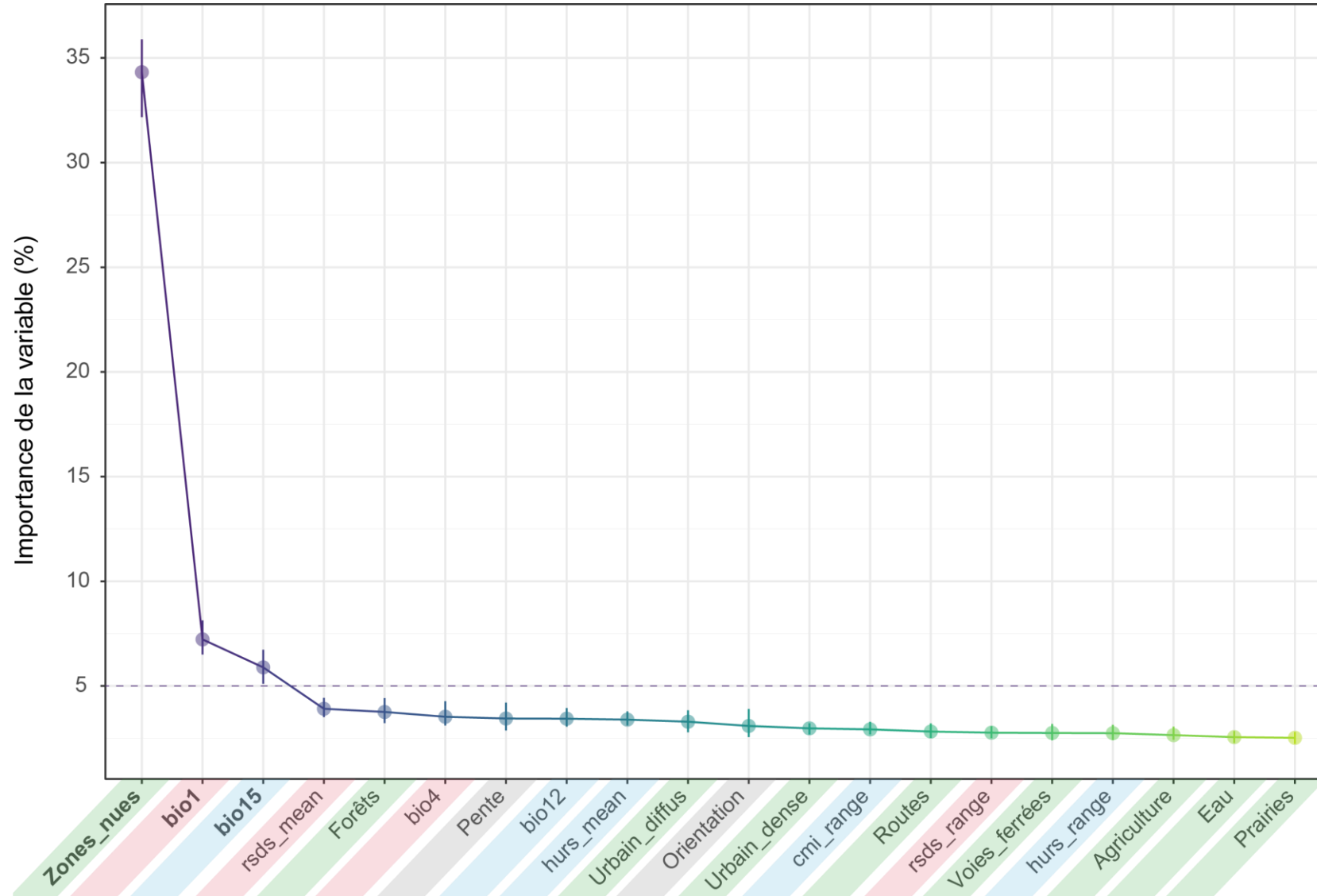
Sensibilité et seuillage du modèle RandomForest



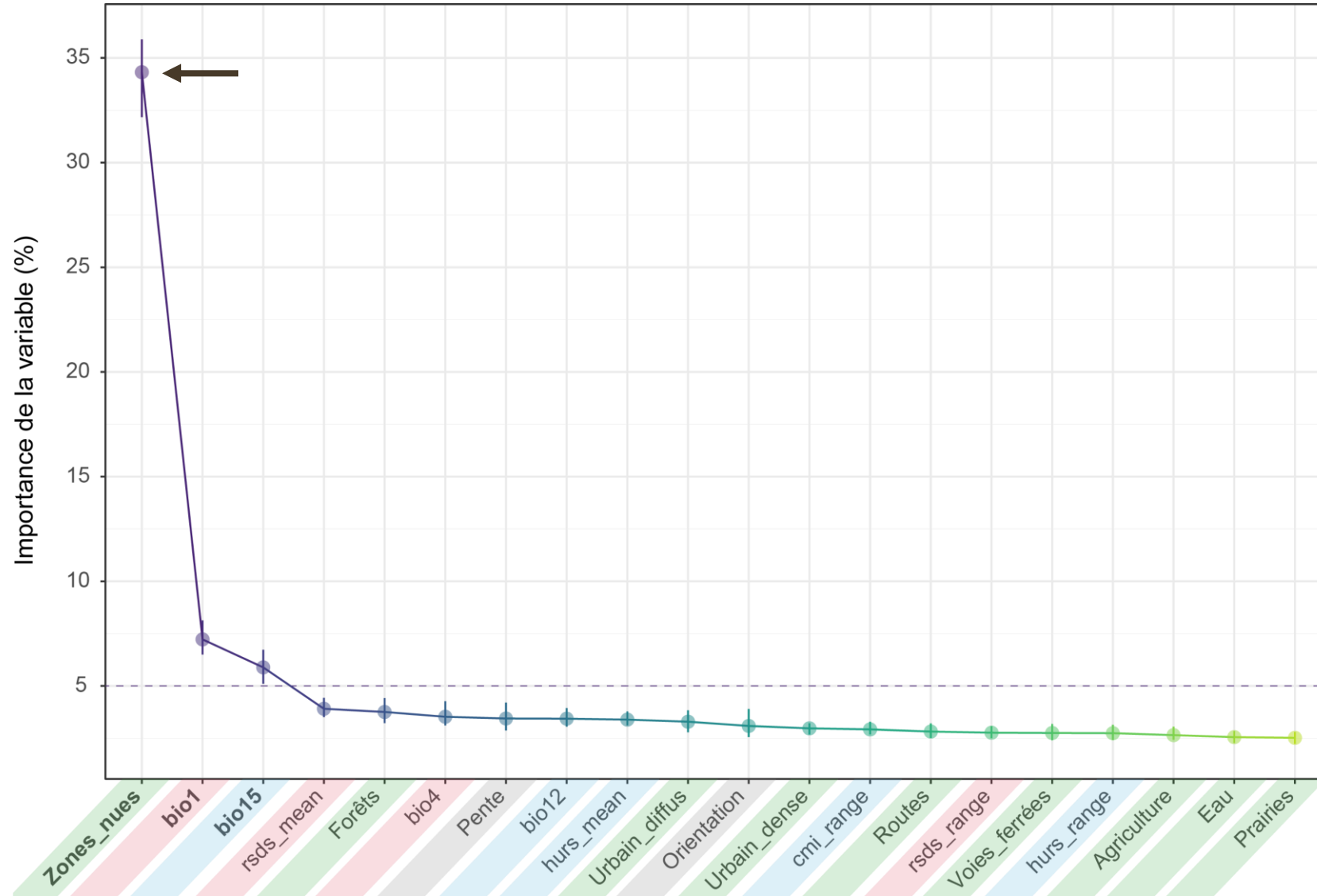
Sensibilité et seuillage du modèle RandomForest



Contribution des variables au score de qualité d'habitat



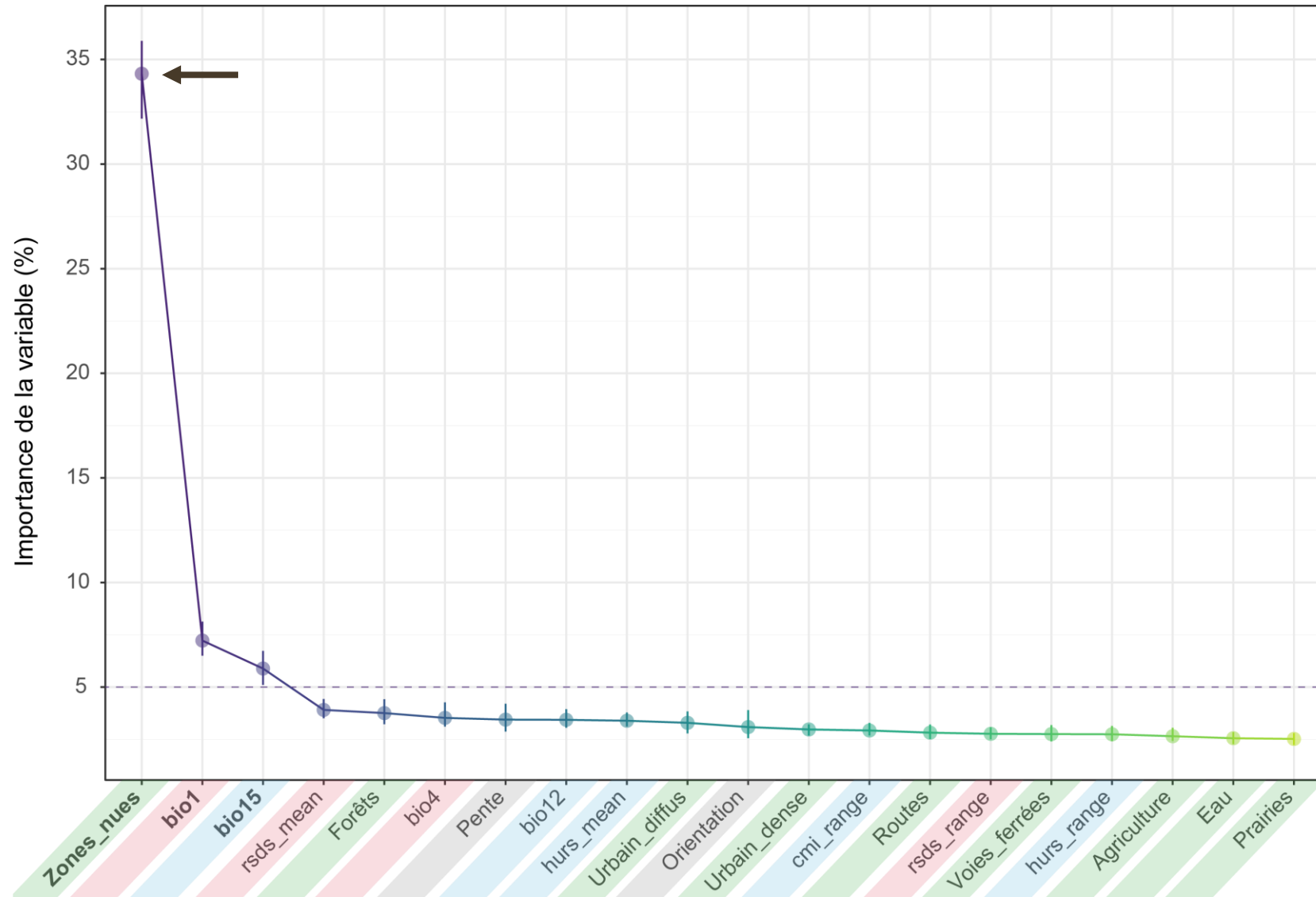
Contribution des variables au score de qualité d'habitat



► Distance aux Zones xériques ouvertes

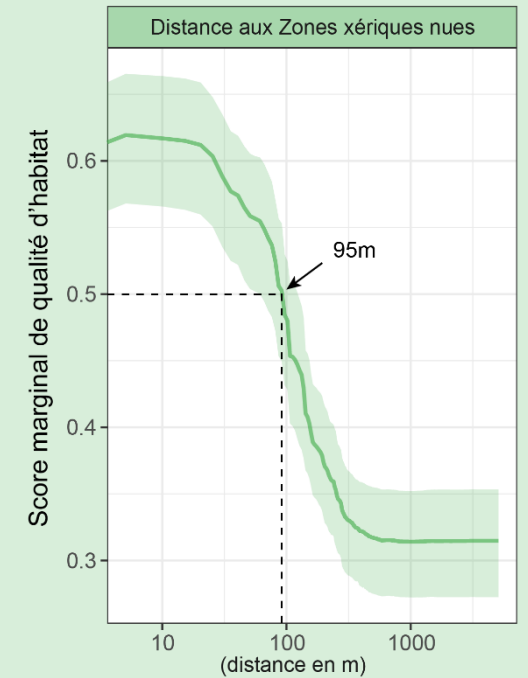
- variable *Zones_nues*
- Contribution : 34.3%

Contribution des variables au score de qualité d'habitat

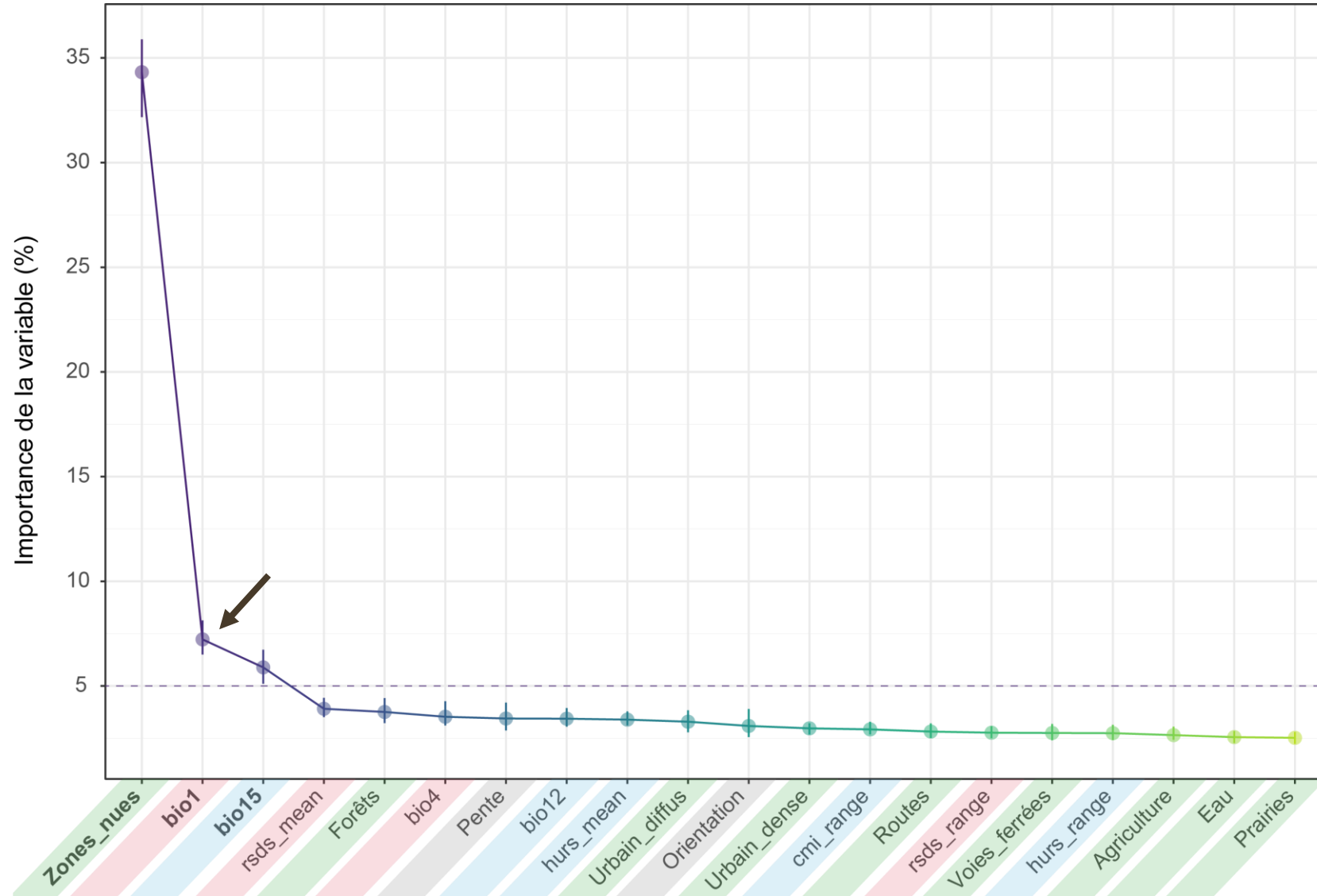


► Distance aux Zones xériques ouvertes

- variable *Zones_nues*
- Contribution : 34.3%

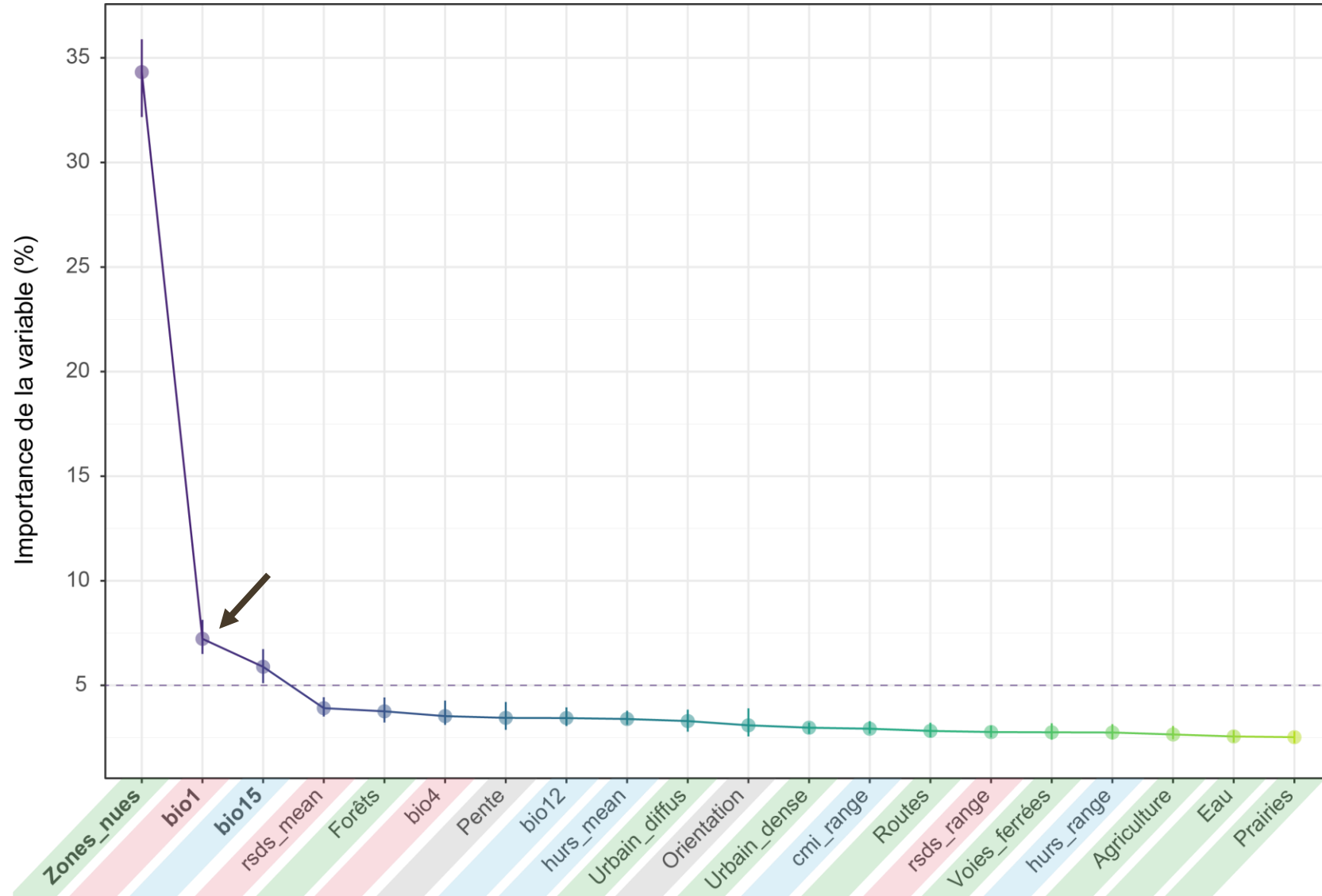


Contribution des variables au score de qualité d'habitat

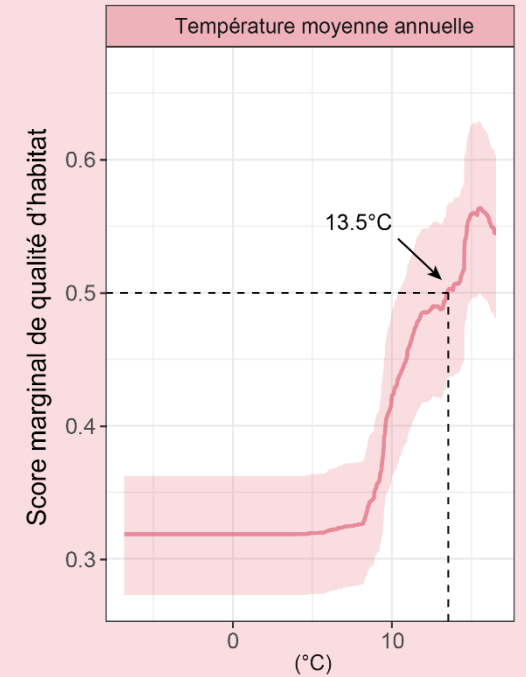


- Température moyenne annuelle
 - variable *bio1*
 - Contribution : 7.2%

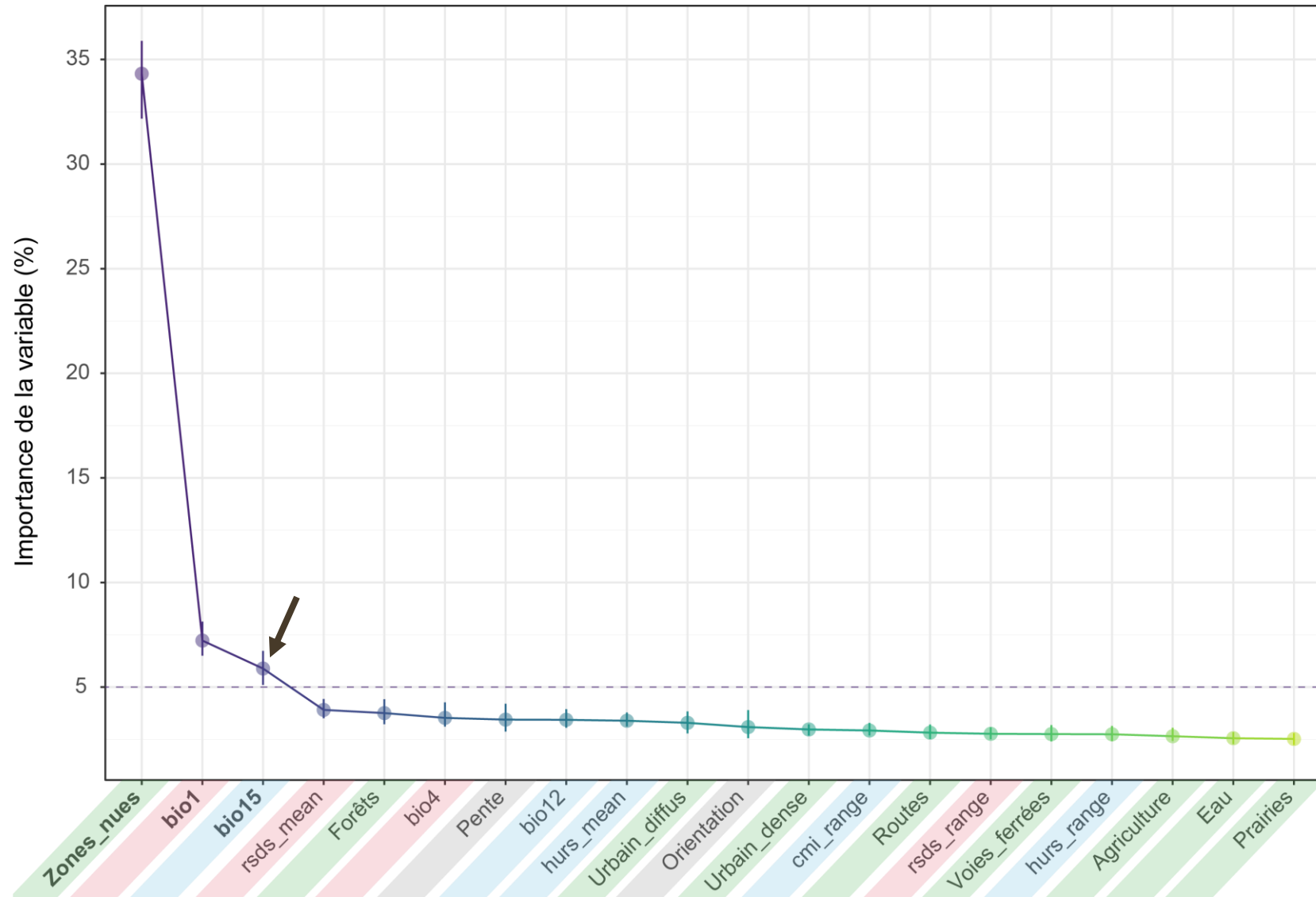
Contribution des variables au score de qualité d'habitat



- Température moyenne annuelle
 - variable *bio1*
 - Contribution : 7.2%

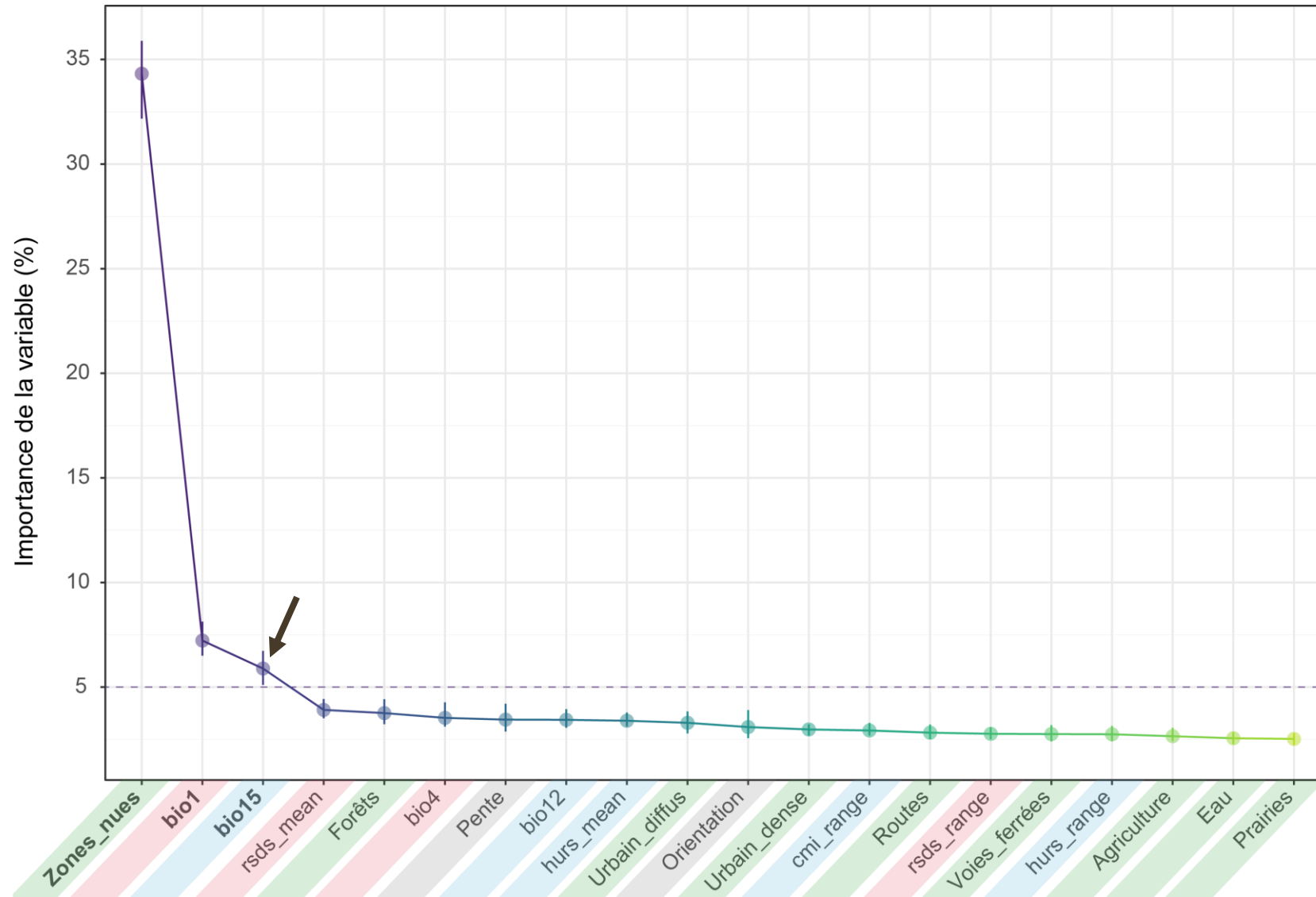


Contribution des variables au score de qualité d'habitat



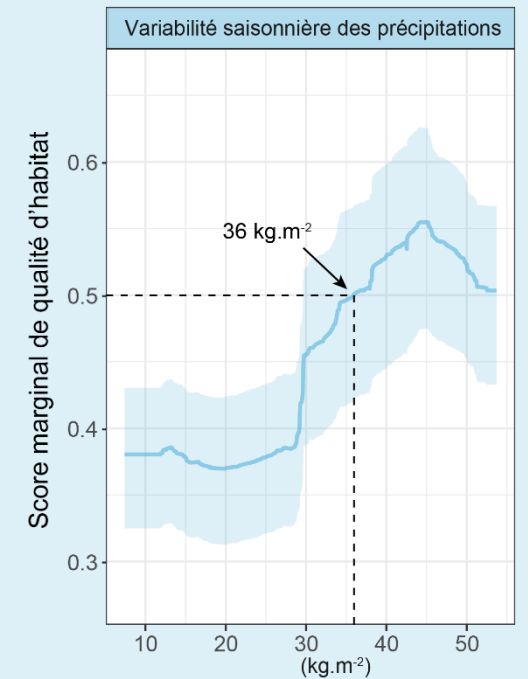
- ▶ Variabilité saisonnière des précipitations
 - variable *bio15*
 - Contribution : 5.9%

Contribution des variables au score de qualité d'habitat

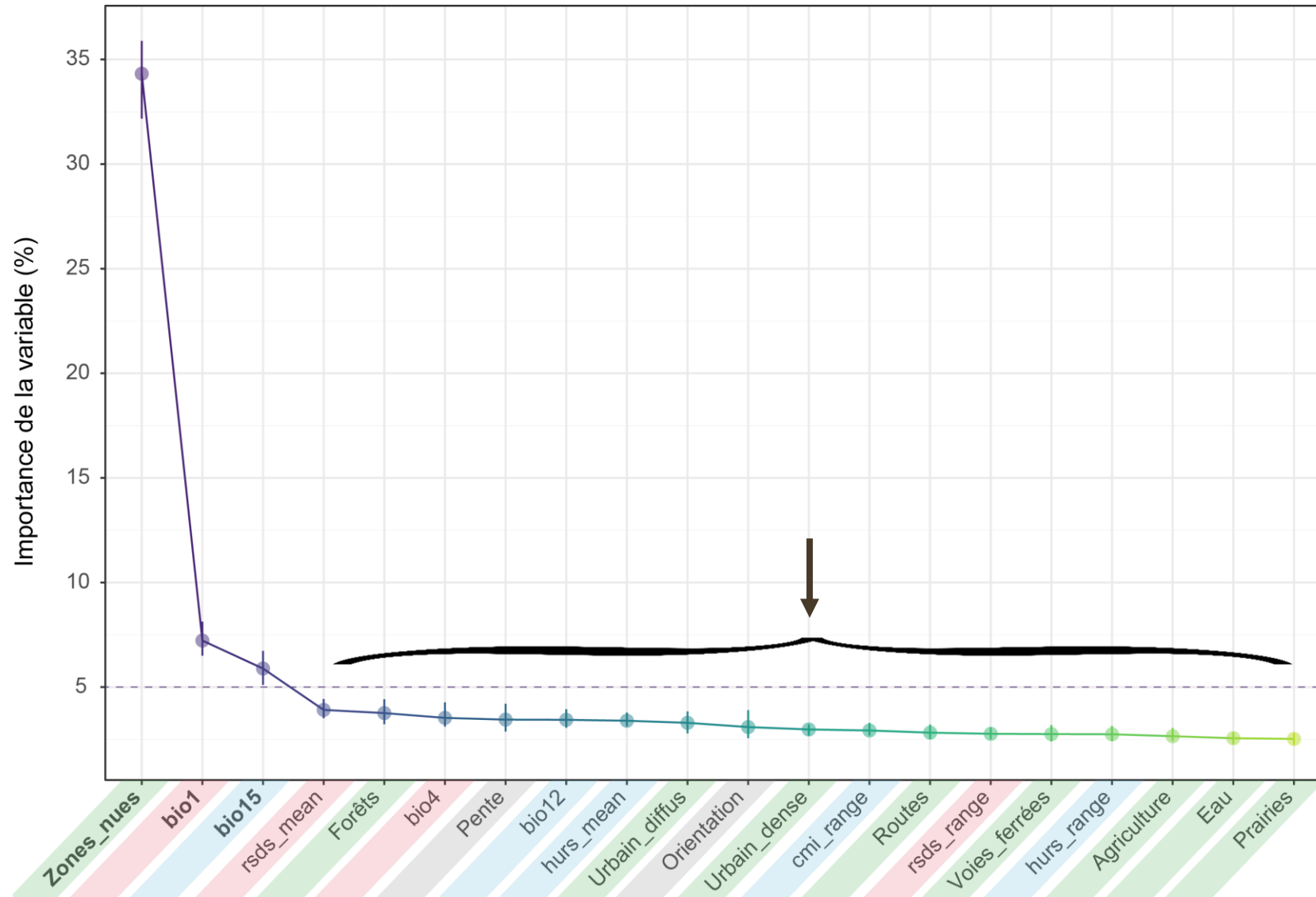


► Variabilité saisonnière des précipitations

- variable *bio15*
- Contribution : 5.9%



Contribution des variables au score de qualité d'habitat



- ▶ Autres variables
 - Contributions <5% mais >2.5%

Limites du modèle

- ▶ Données de présence non exhaustives



Limites du modèle

- ▶ Données de présence non exhaustives
- ▶ Données environnementales non exhaustives (micro-habitat)



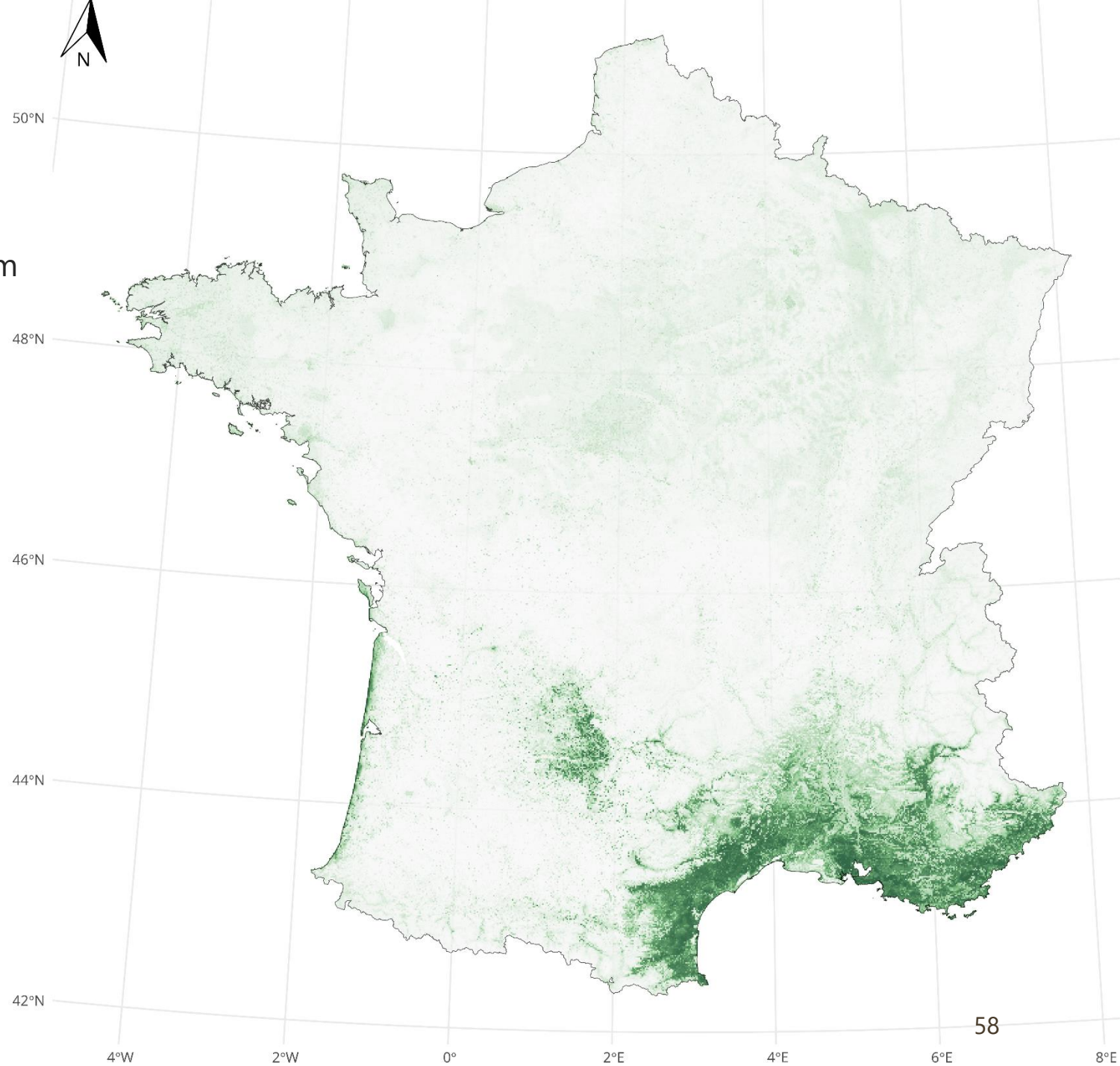
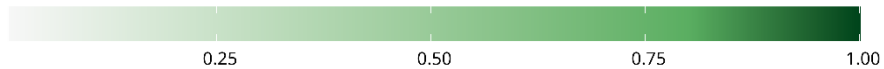
©Xavier Rozec

Limites du modèle

- ▶ Données de présence non exhaustives
- ▶ Données environnementales non exhaustives (micro-habitat)
- ▶ Le modèle n'intègre **aucune information sur la connectivité fonctionnelle** du paysage
 - Certains pixels peuvent être prédits comme hautement favorables mais rester totalement inaccessibles...

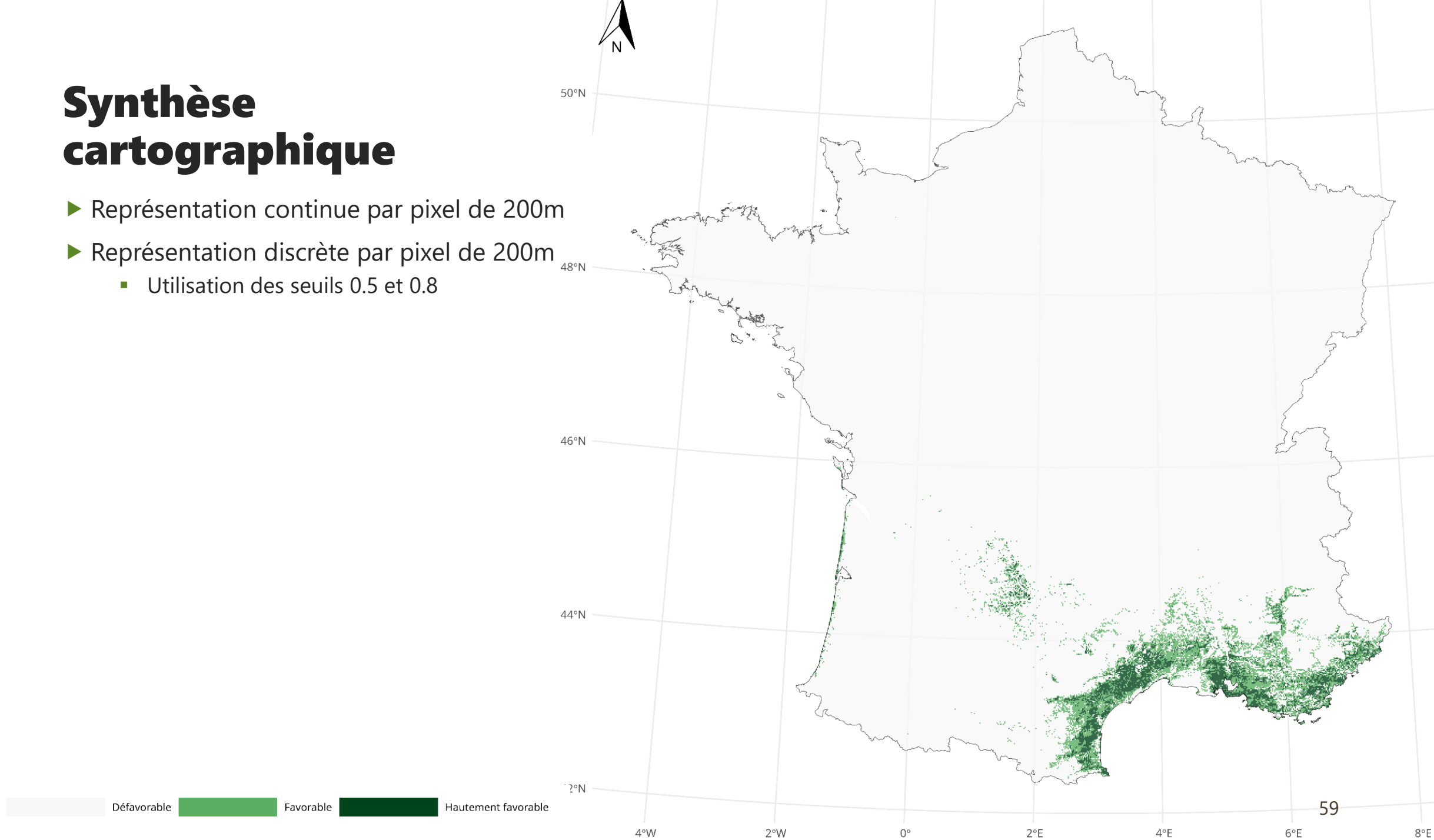
Synthèse cartographique

► Représentation continue par pixel de 200m



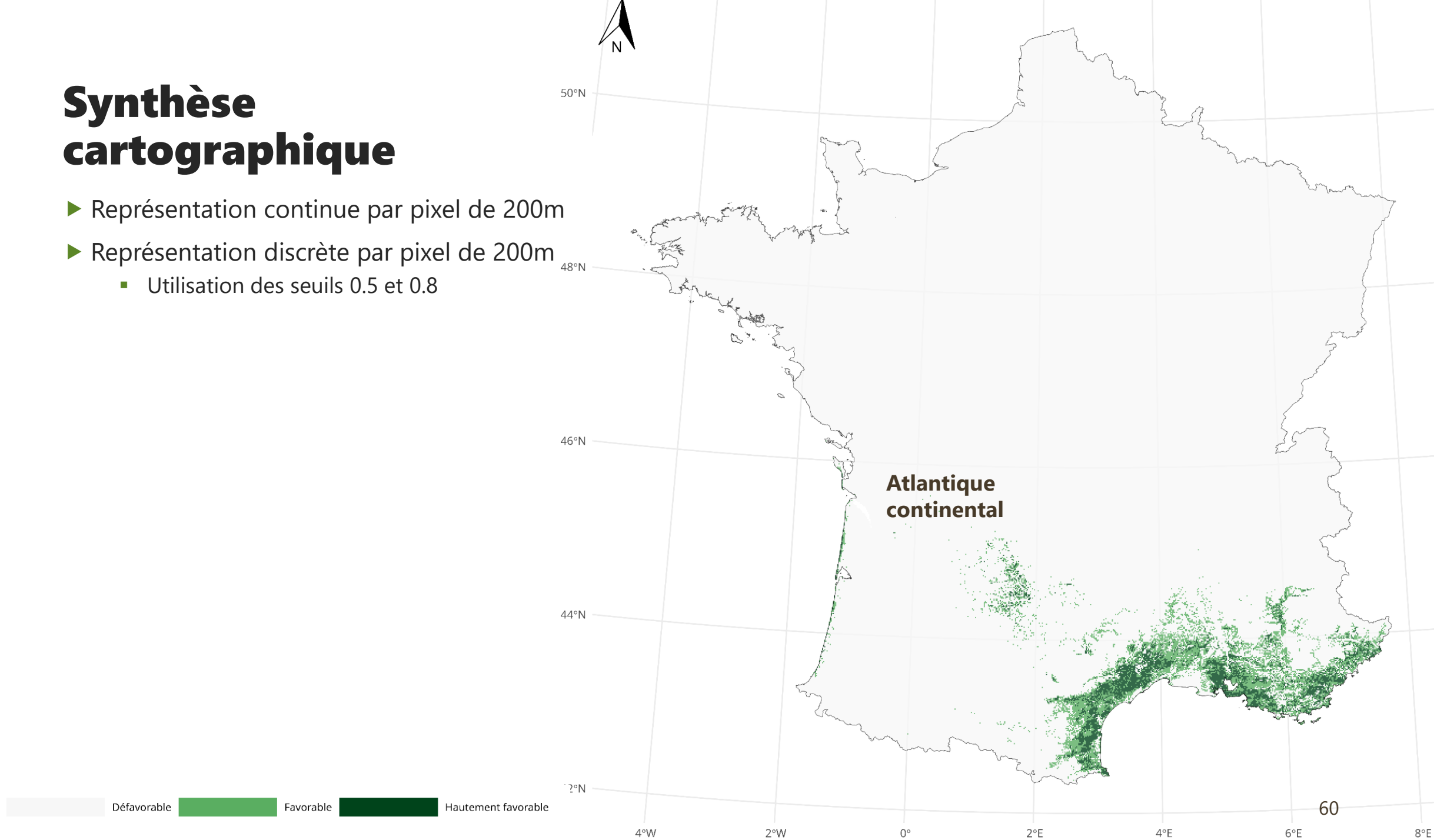
Synthèse cartographique

- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8



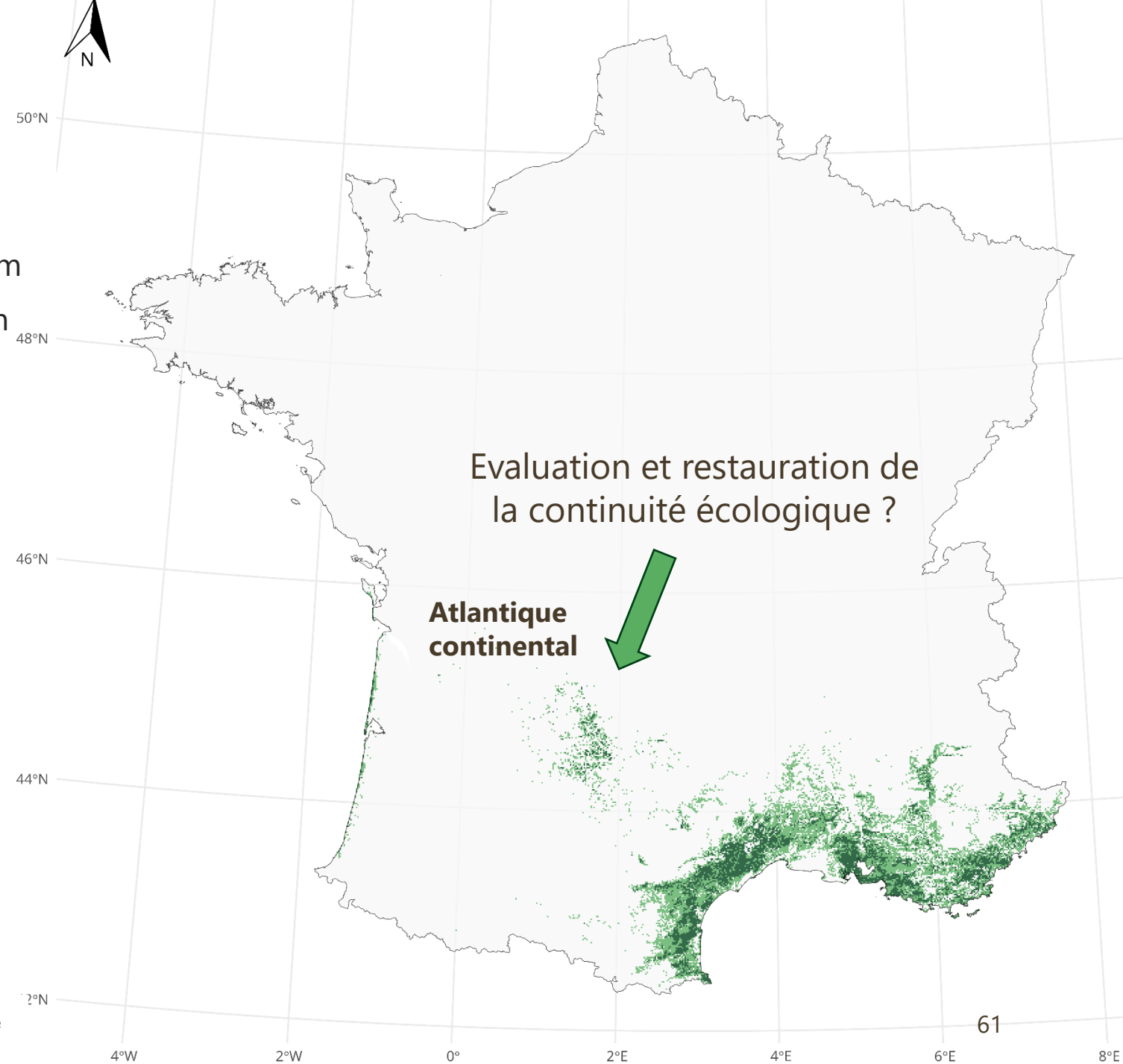
Synthèse cartographique

- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8



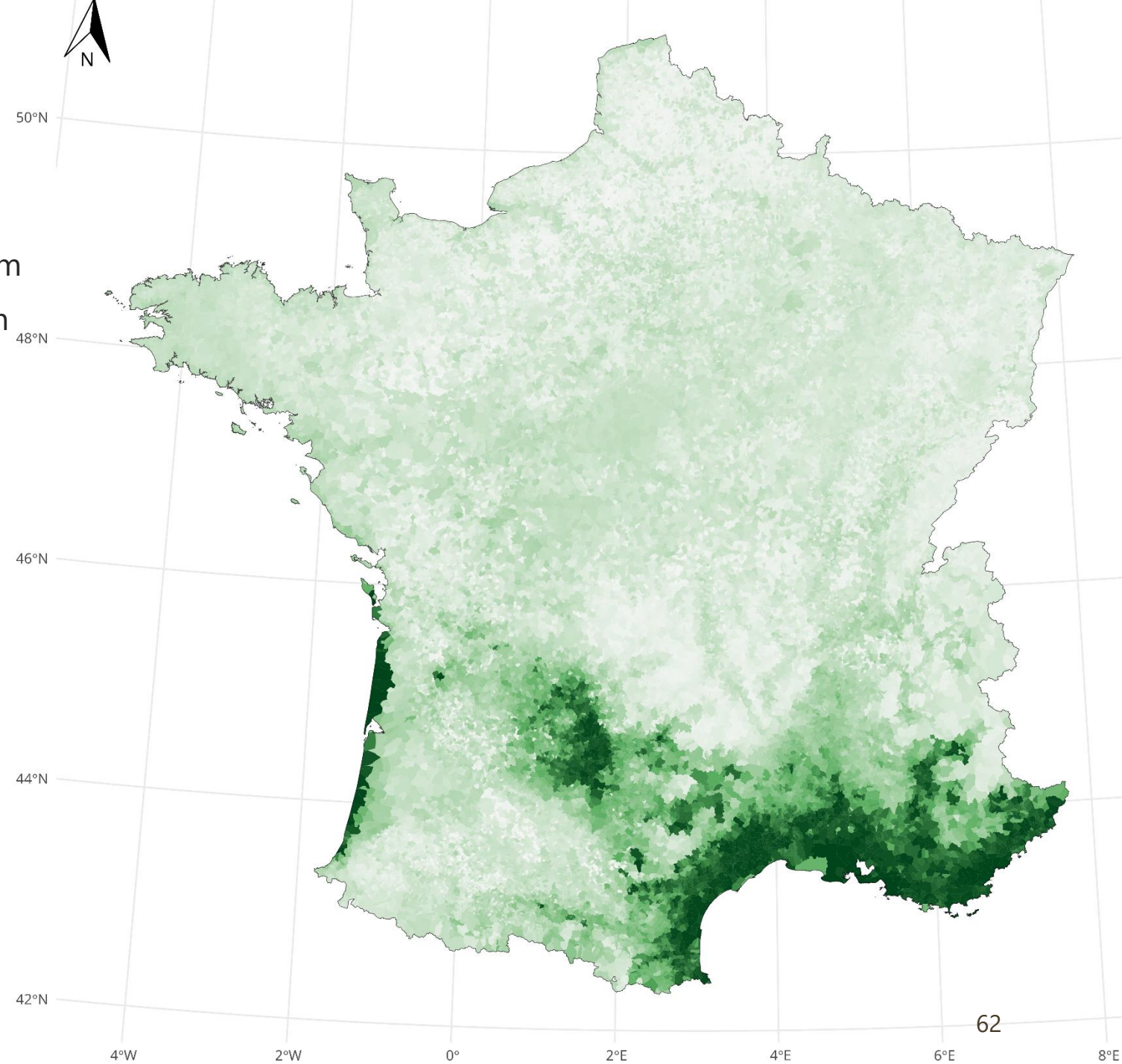
Synthèse cartographique

- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8



Synthèse cartographique

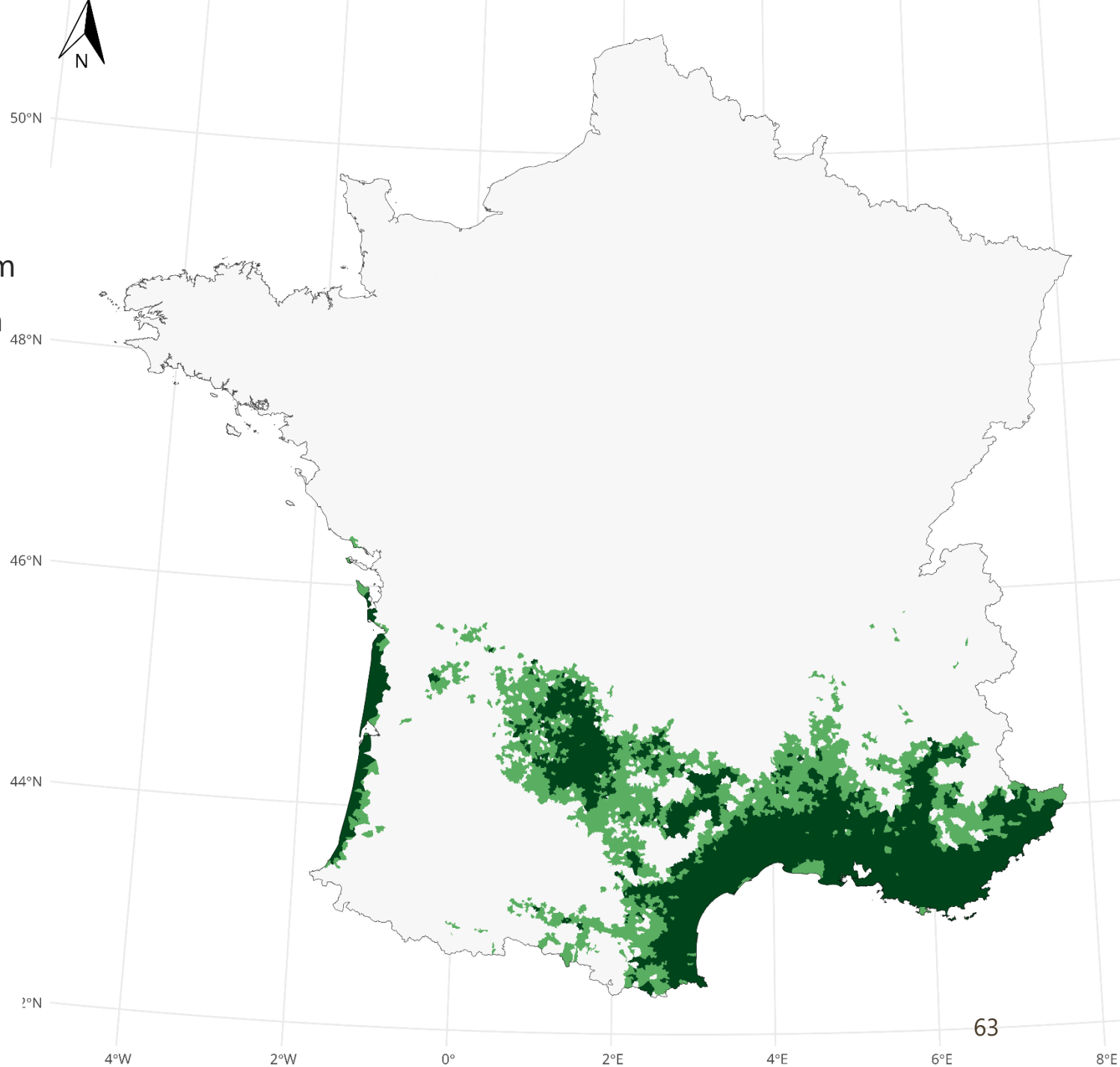
- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8
- ▶ Représentation continue par commune
 - Valeur maximale par commune



Synthèse cartographique

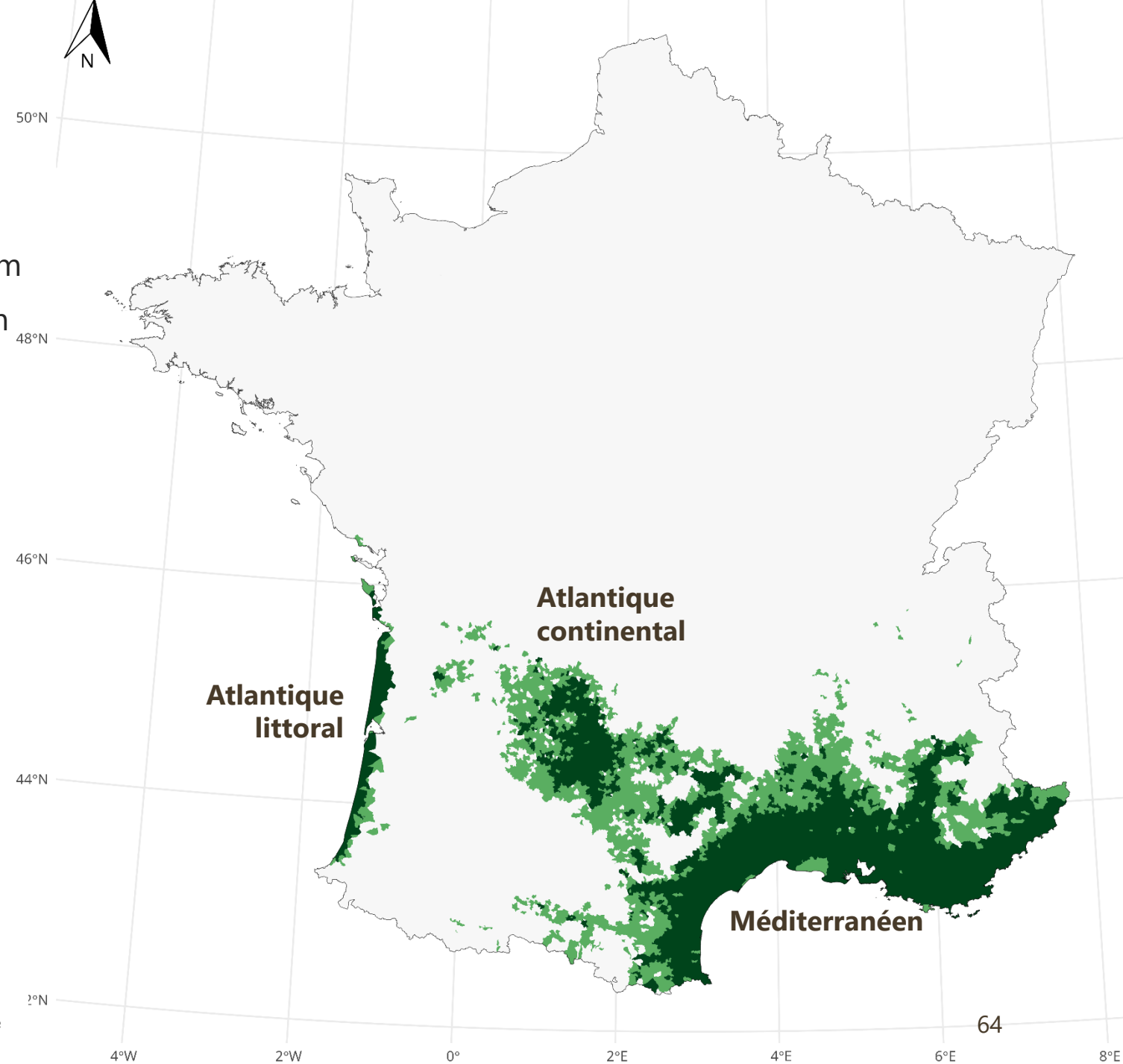
- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8
- ▶ Représentation continue par commune
 - Valeur maximale par commune
- ▶ Représentation discrète par commune
 - Valeur maximale par commune

Défavorable Favorable Hautement favorable



Synthèse cartographique

- ▶ Représentation continue par pixel de 200m
- ▶ Représentation discrète par pixel de 200m
 - Utilisation des seuils 0.5 et 0.8
- ▶ Représentation continue par commune
 - Valeur maximale par commune
- ▶ Représentation discrète par commune
 - Valeur maximale par commune



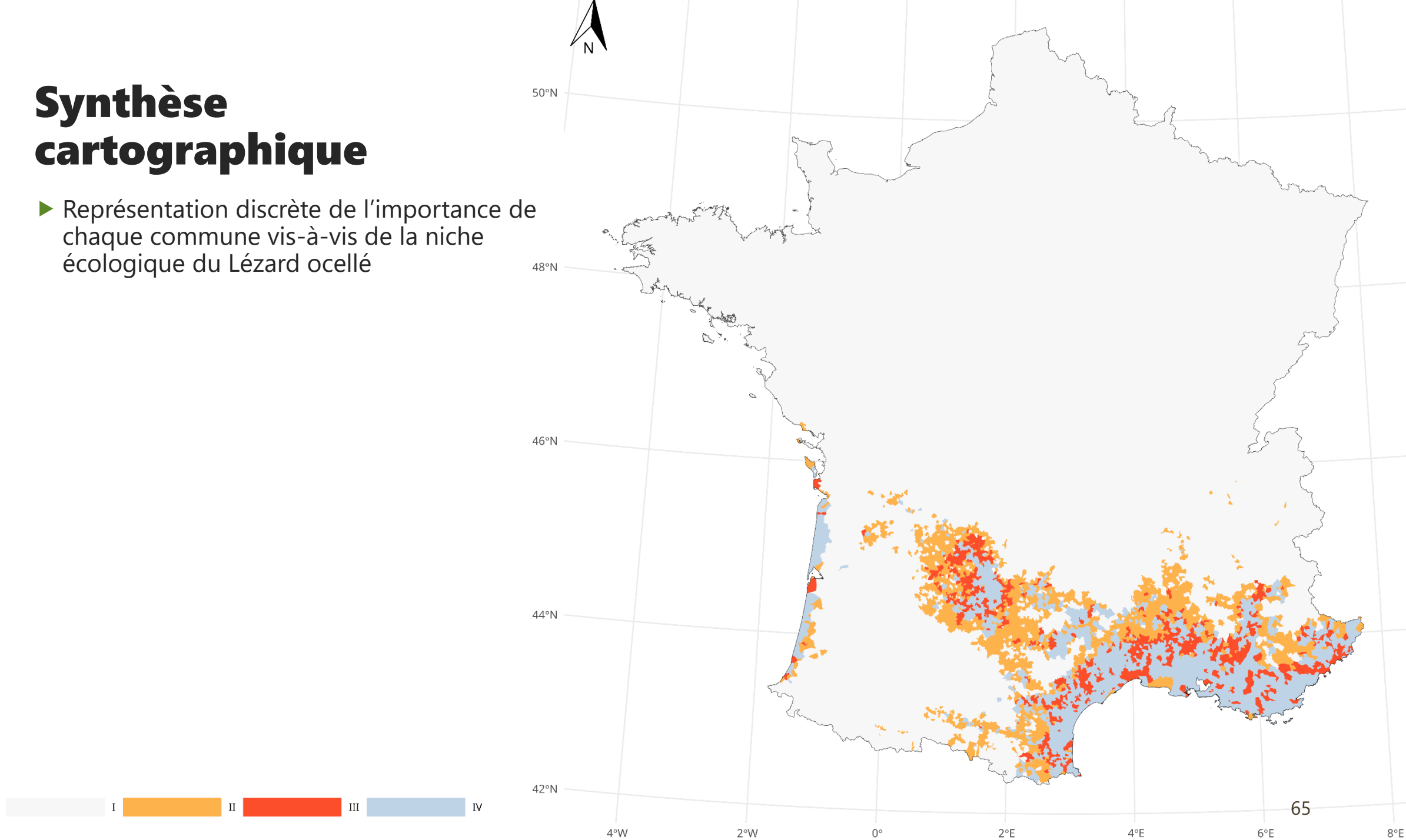
Défavorable

Favorable

Hautelement favorable

Synthèse cartographique

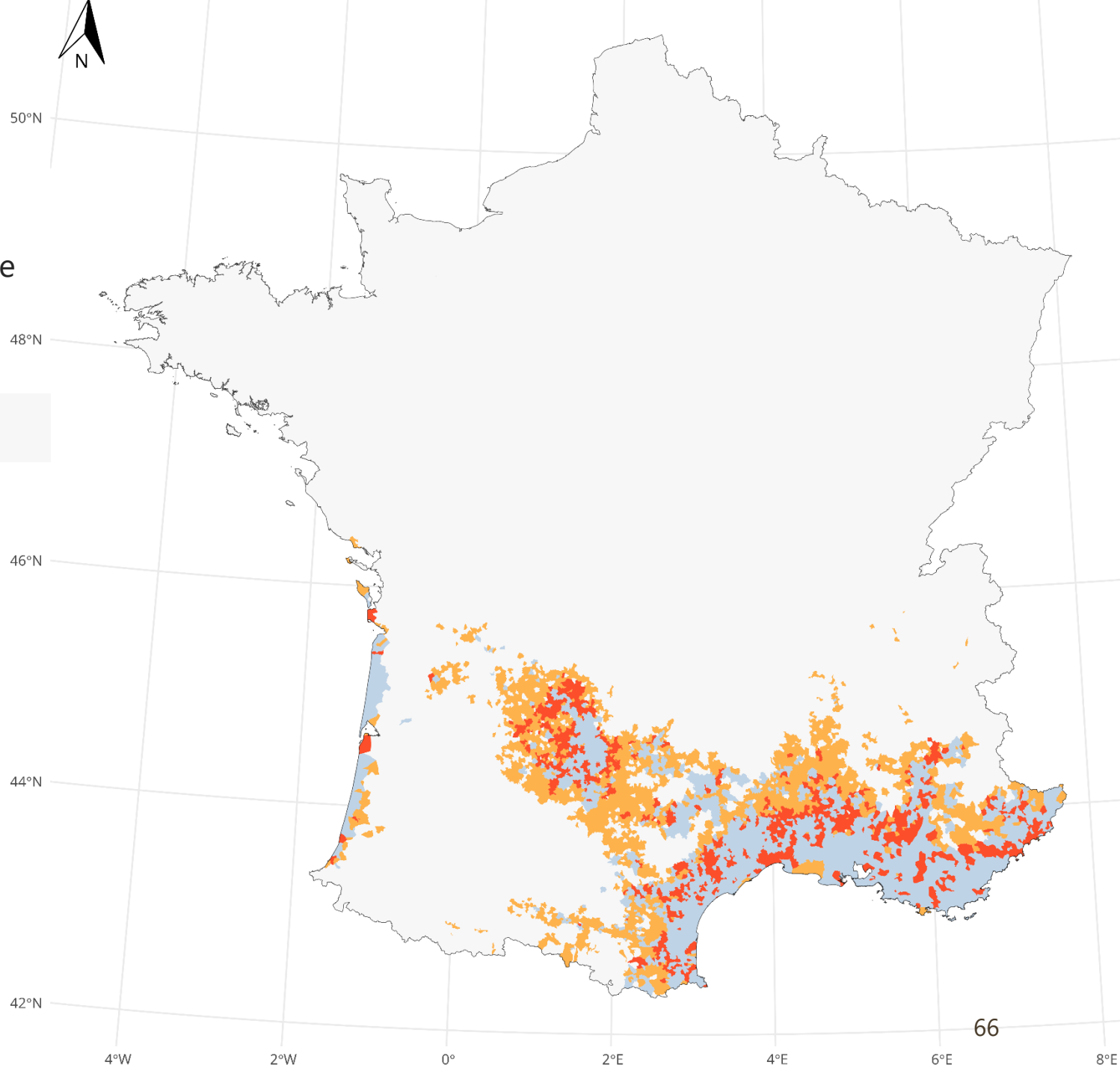
- Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé



Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

- Niveau I : Espèce absente et milieu défavorable

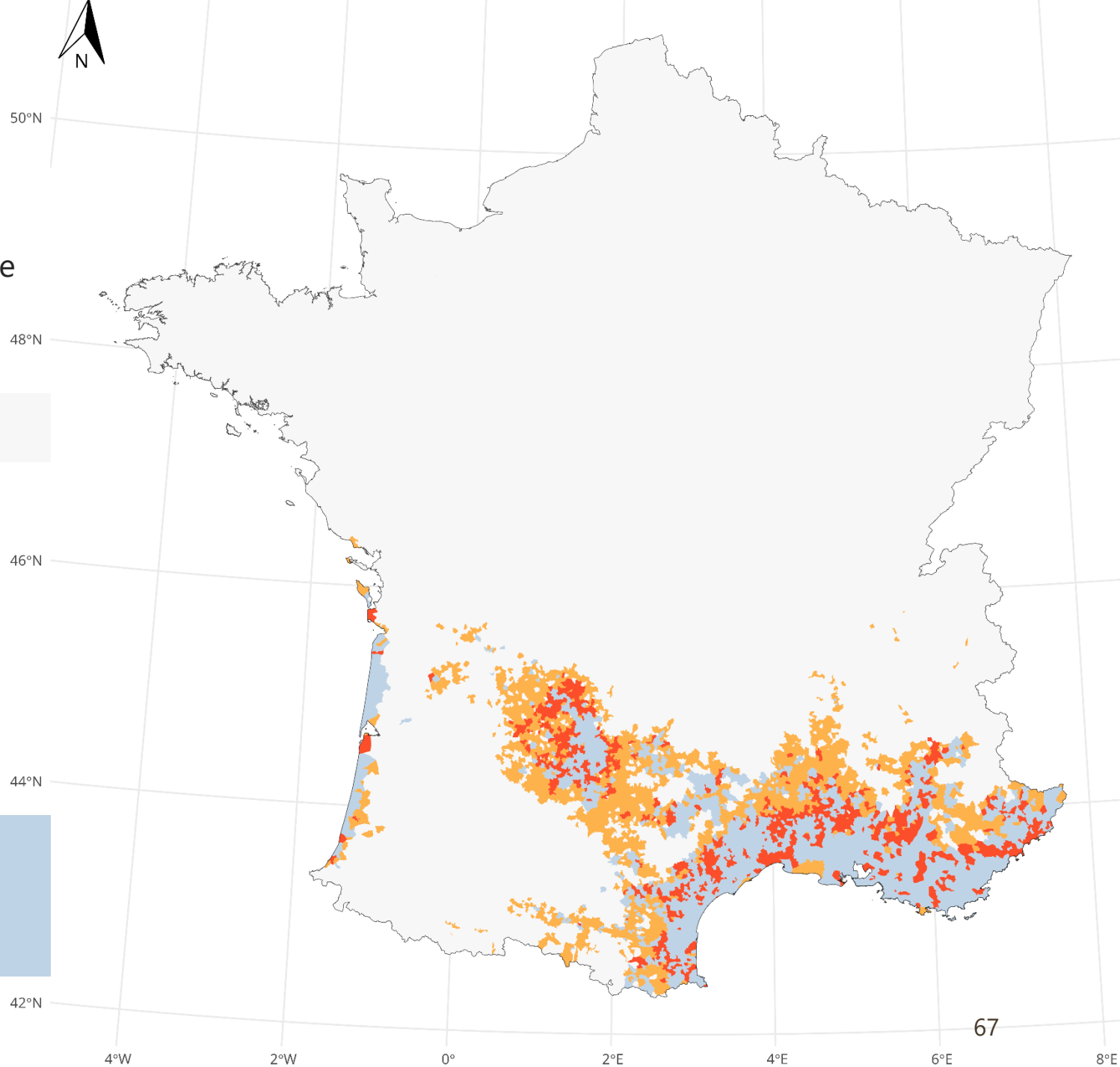


Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

■ Niveau I : Espèce absente et milieu défavorable

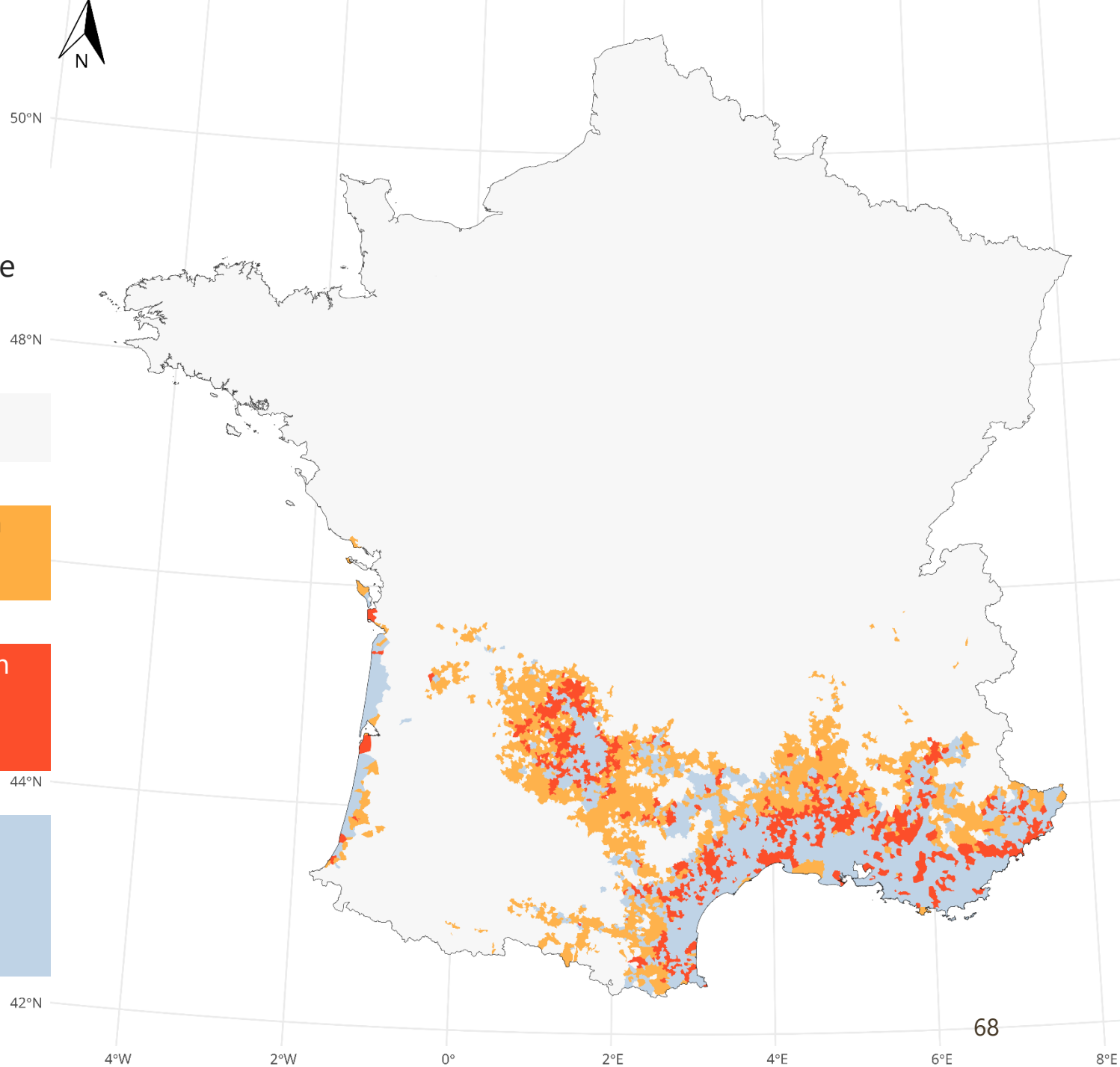
■ Niveau IV : Espèce d'ores et déjà détectée et habitat hautement favorable, impliquant un effort de conservation des populations locales



Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

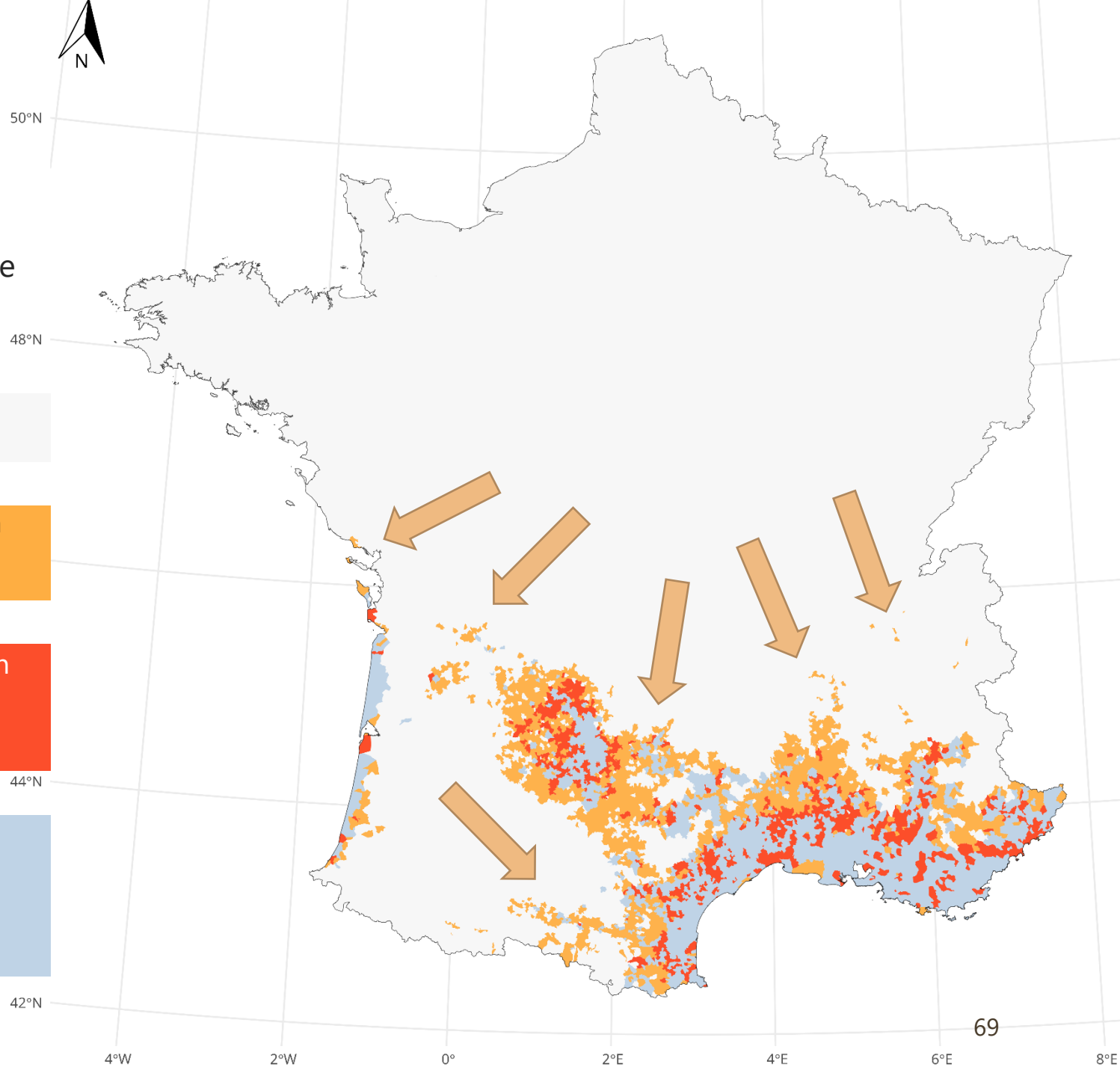
- Niveau I : Espèce absente et milieu défavorable
- Niveau II : Espèce non détectée malgré un habitat favorable, invitant à un effort de prospection
- Niveau III : Espèce non détectée malgré un habitat hautement favorable, invitant à un effort soutenu de prospection
- Niveau IV : Espèce d'ores et déjà détectée et habitat hautement favorable, impliquant un effort de conservation des populations locales



Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

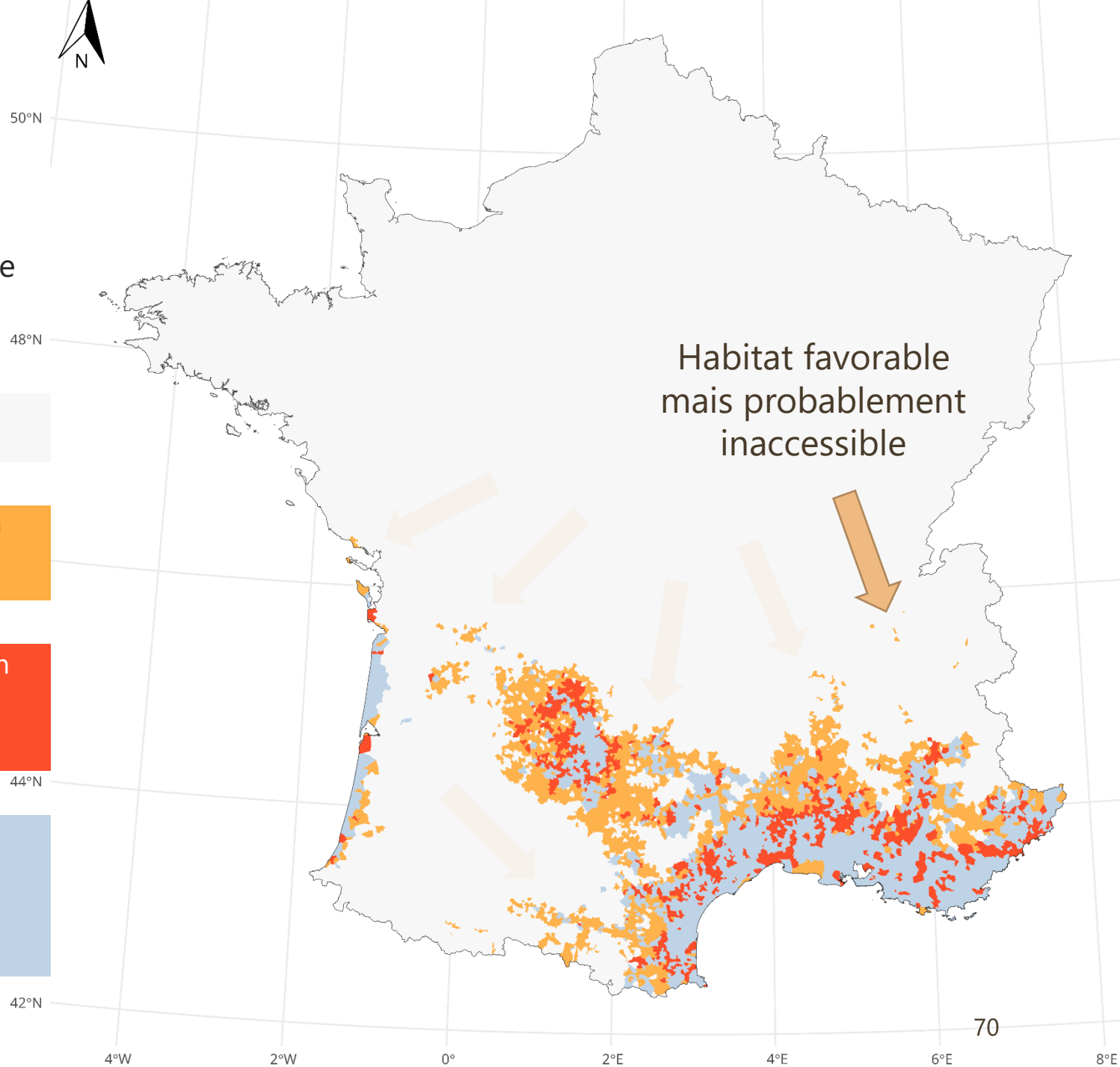
- Niveau I : Espèce absente et milieu défavorable
- Niveau II : Espèce non détectée malgré un habitat favorable, invitant à un effort de prospection
- Niveau III : Espèce non détectée malgré un habitat hautement favorable, invitant à un effort soutenu de prospection
- Niveau IV : Espèce d'ores et déjà détectée et habitat hautement favorable, impliquant un effort de conservation des populations locales



Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

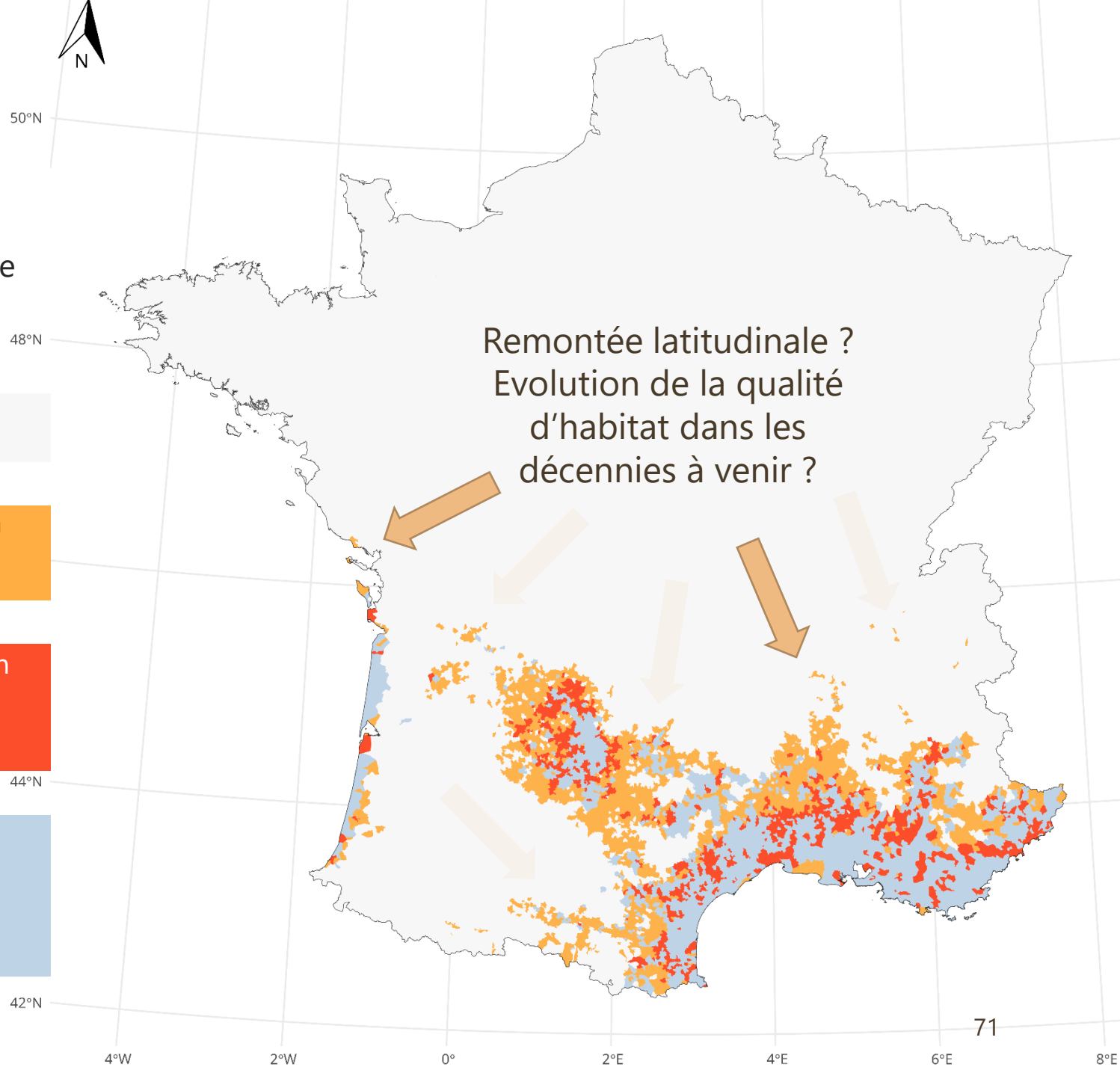
- Niveau I : Espèce absente et milieu défavorable
- Niveau II : Espèce non détectée malgré un habitat favorable, invitant à un effort de prospection
- Niveau III : Espèce non détectée malgré un habitat hautement favorable, invitant à un effort soutenu de prospection
- Niveau IV : Espèce d'ores et déjà détectée et habitat hautement favorable, impliquant un effort de conservation des populations locales



Synthèse cartographique

► Représentation discrète de l'importance de chaque commune vis-à-vis de la niche écologique du Lézard ocellé

- Niveau I : Espèce absente et milieu défavorable
- Niveau II : Espèce non détectée malgré un habitat favorable, invitant à un effort de prospection
- Niveau III : Espèce non détectée malgré un habitat hautement favorable, invitant à un effort soutenu de prospection
- Niveau IV : Espèce d'ores et déjà détectée et habitat hautement favorable, impliquant un effort de conservation des populations locales





Merci pour votre attention

Contact

Jérôme PRUNIER

ADENKO

jerome.prunier@adeneko.fr

Cartes et rapport d'étude sont accessibles ici :

http://www.adeneko.fr/archives/ADENKO_Ocelle_NENM.zip

Événement organisé par :



Avec le soutien technique et financier de :

