

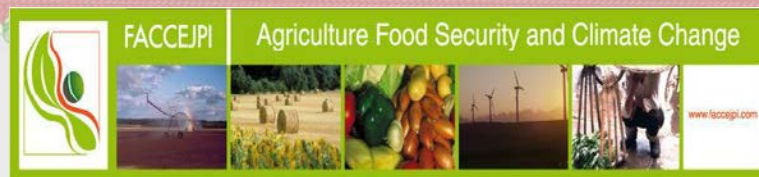
Leyli Borner

Post-Doc

Projets Genomite & NiCoTools

Maria Navajas, Alain Migeon, Gaël Kergoat, Bruno Le Ru, Christine Meynard

Modélisation de la répartition de ravageurs de cultures: défis méthodologiques et projections climatiques



UMR CBGP
Centre de Biologie pour
la Gestion des Populations



Double objectif:

1. 15 espèces ravageurs de culture (Genomite) => connaître la répartition potentielle => cartographie des enjeux

WP1 : Species distribution models and tritrophic interactions under CC scenarios in Europe

interactions entre organismes– conséquences sur la dynamique des chaînes trophiques – aires de distribution

En réponse à l'enjeu:

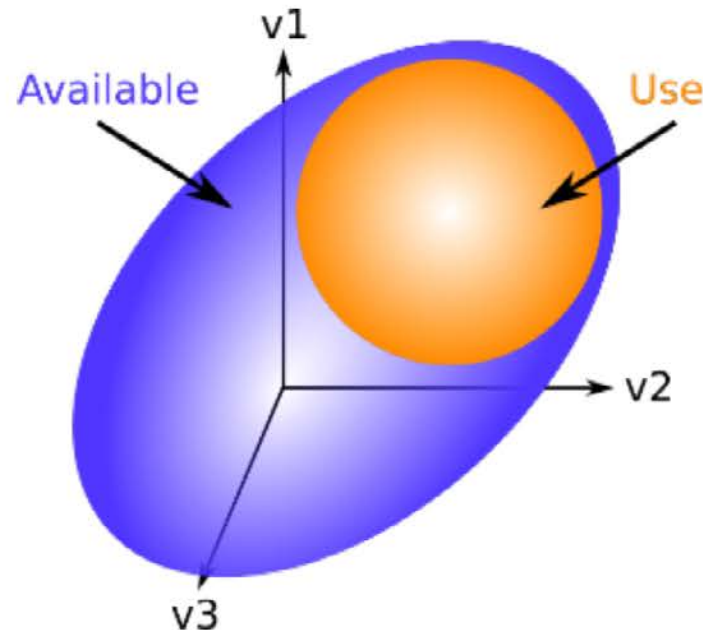
- ❖ **Estimer et prédire les risques agronomiques. Modélisation écologique (modification des aires de répartition, bioinvasions) : enjeu 2**

2. Problématique méthodologique de modéliser avec seulement des données de présence => trouver quelle est la meilleure méthode

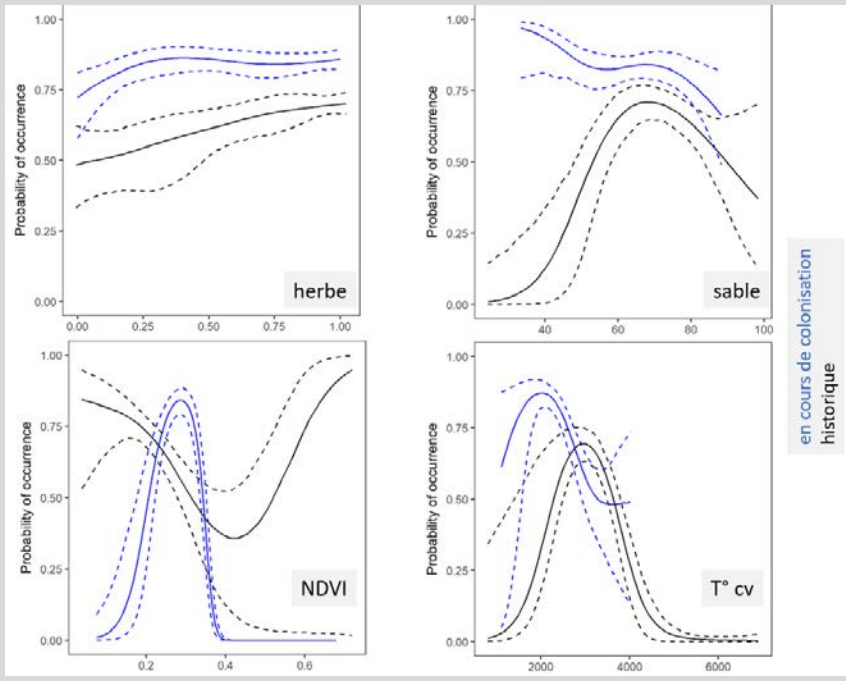
La modélisation de niche écologique: qu'est-ce que c'est?

Niche écologique d'une espèce: ensemble des conditions environnementales nécessaires à la survie et à la persistance de l'espèce

Principe méthodologique: comparaison des caractéristiques des sites utilisés par l'espèce aux caractéristiques des sites disponibles

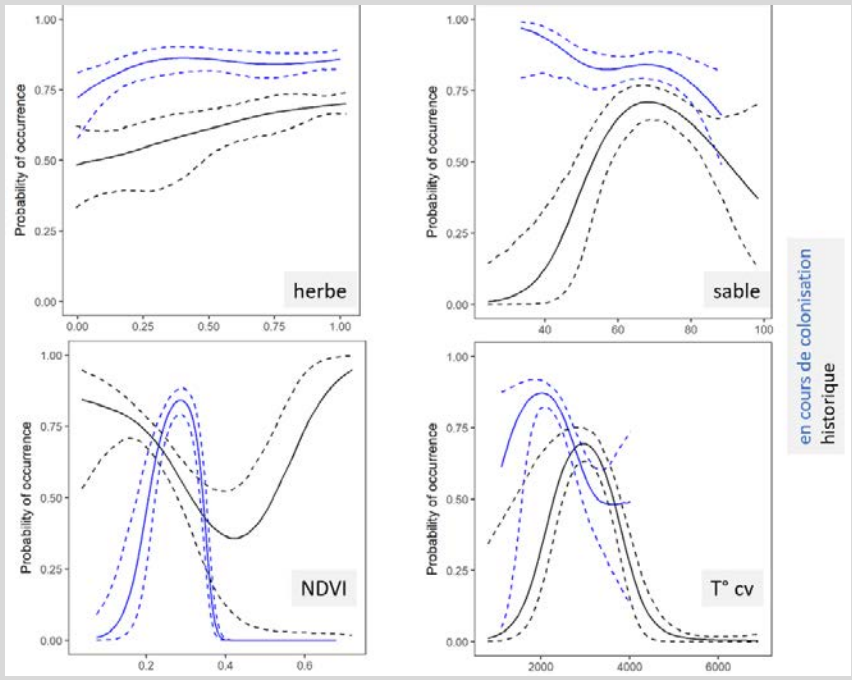


1. Niche environnementale

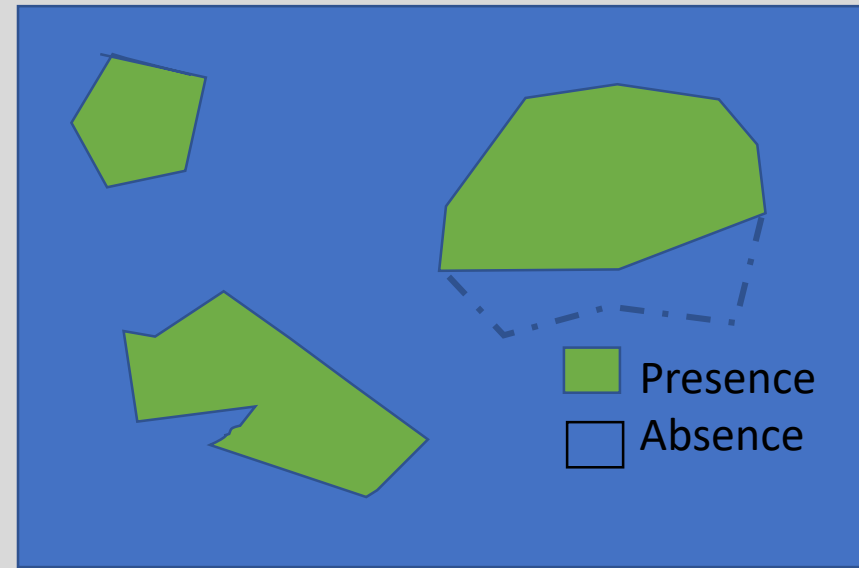


Borner et al. (in prep) CERISE

1. Niche environnementale

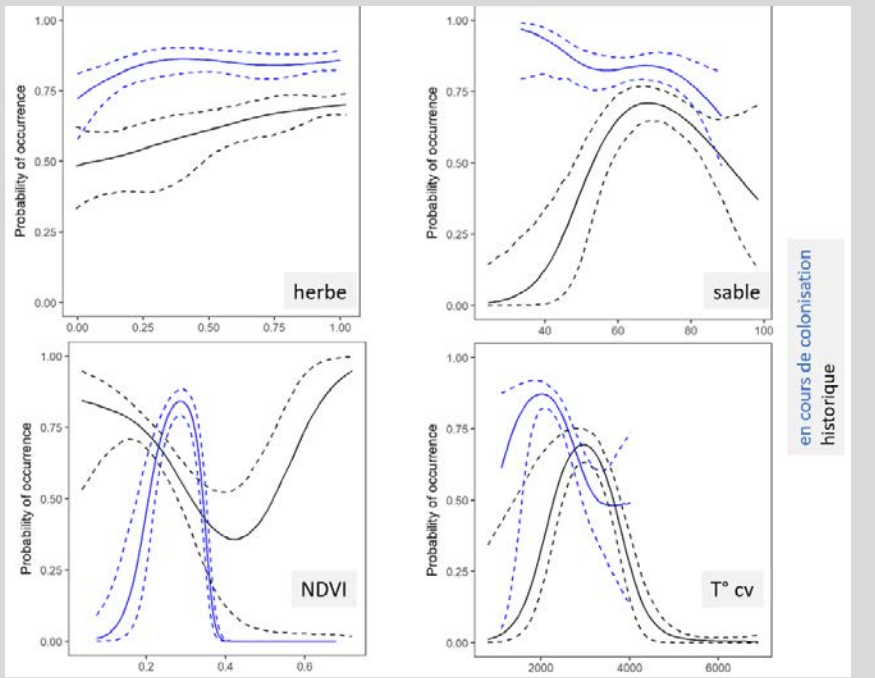


2. Présence/Absence – biais histo/géo

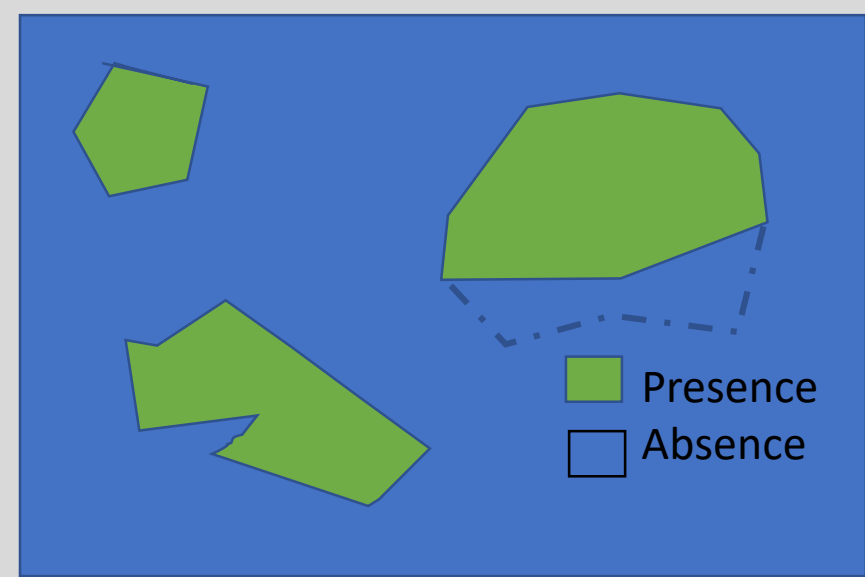


Borner et al. (in prep) CERISE

1. Niche environnementale

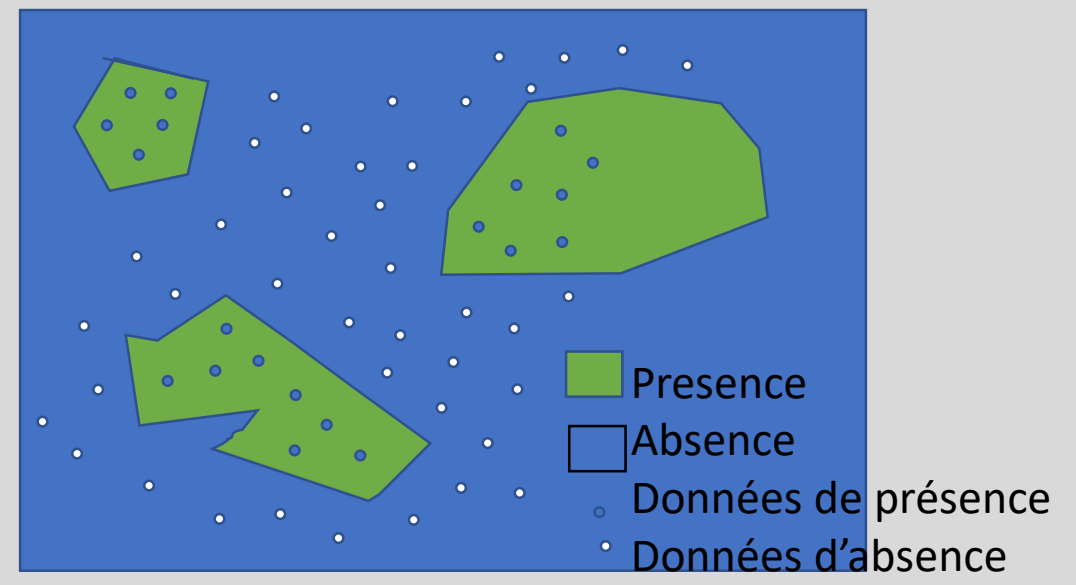


2. Présence/Absence – biais histo/géo



Borner et al. (in prep) CERISE

3. Observation – collecte des données

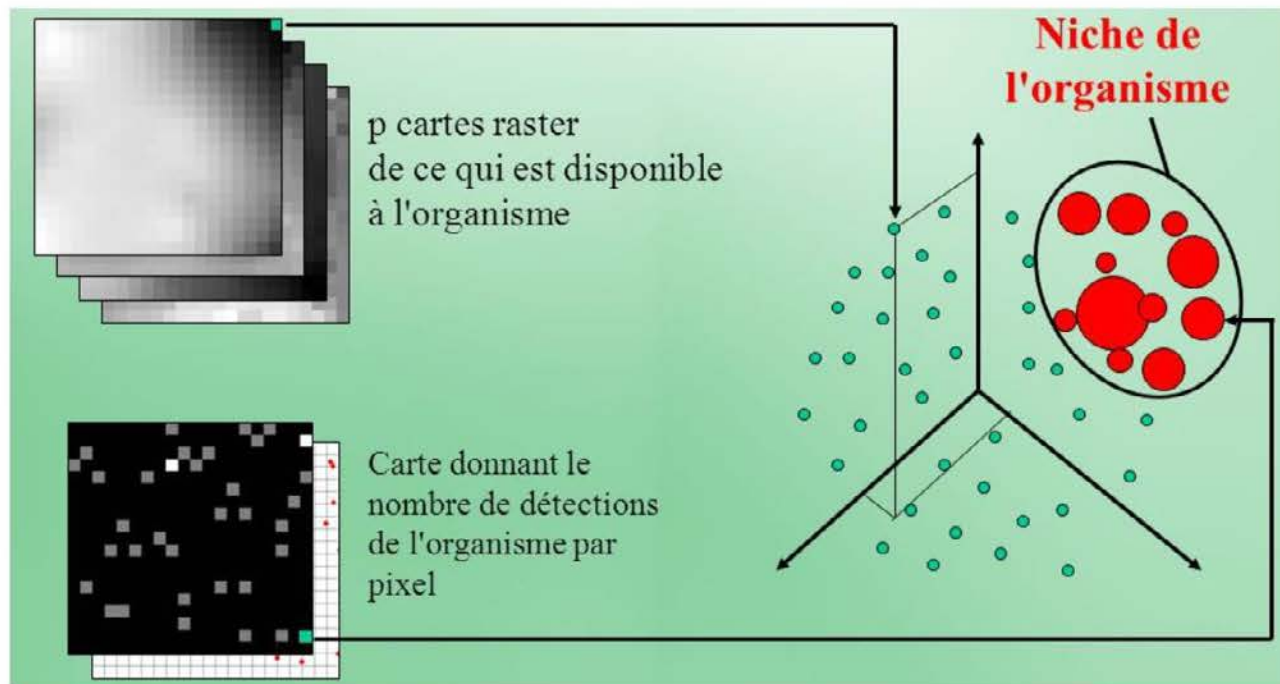


Informations à fournir en entrée:

- Données géoréférencées de présence des espèces
- Description spatialisée de variables environnementales (par ex: habitats, altitude, pente, orientation, bâti et routes)

Informations fournies en sortie:

- Caractérisation de la niche écologique utilisée par l'espèce
- Cartographie de la distribution potentielle de l'espèce (*suitability*)



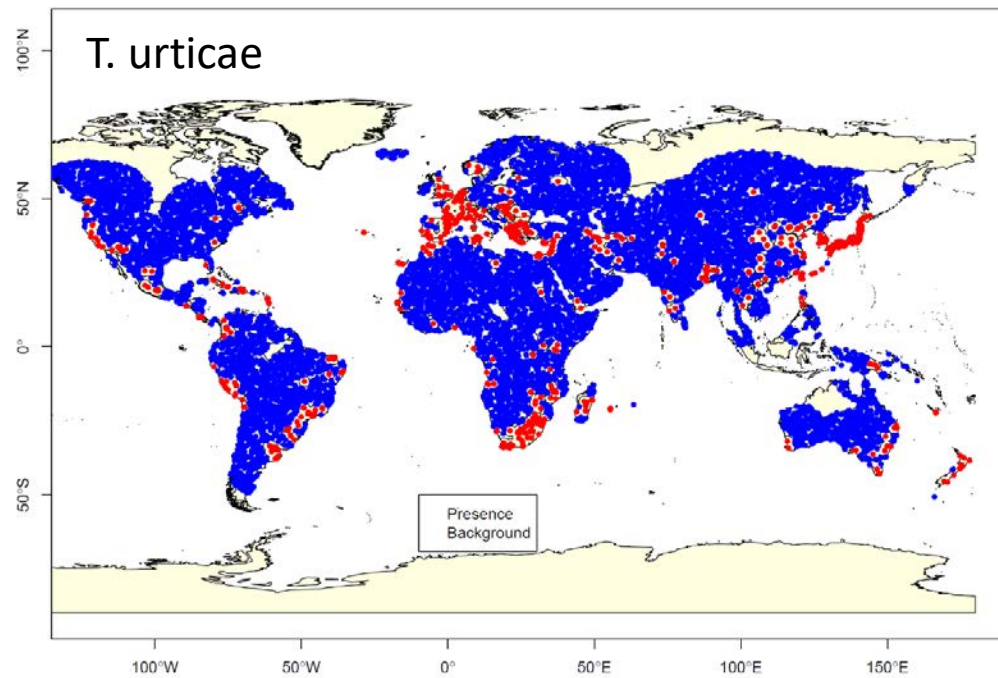
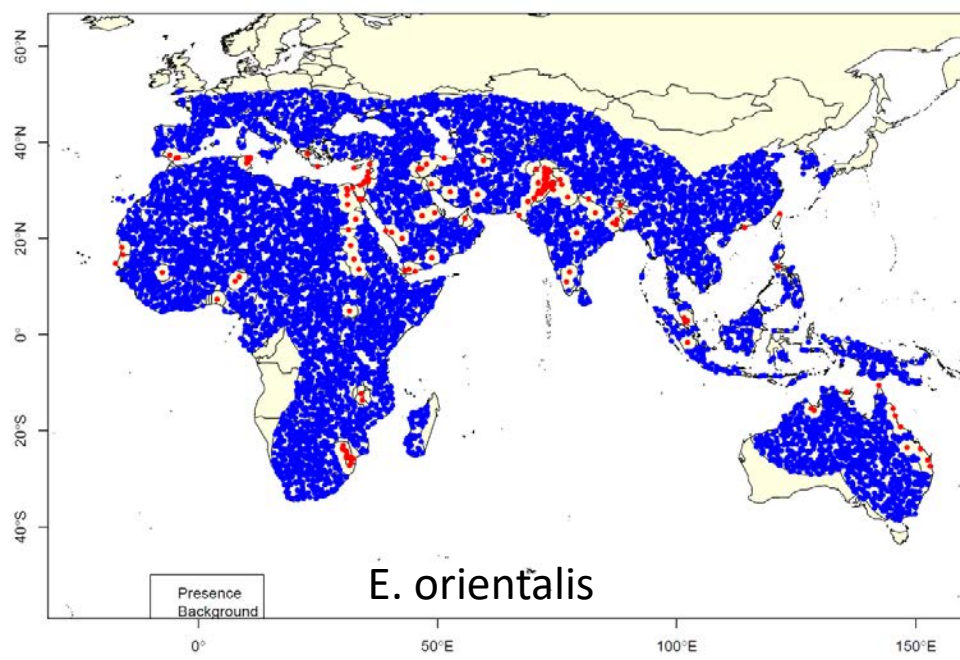
Applications à des ravageurs des cultures – Genomite project – Spider Mites - Migeon & Dorkeld (2006-2017)



| Genre | Espec | Code SPMWeb | Hote |
|---------------|----------------|-------------|---|
| Tetranychus | urticae | 872 | polyphage |
| Tetranychus | neocaledonicus | 852 | polyphage |
| Tetranychus | kanzawai | 1023 | polyphage |
| Tetranychus | evansi | 882 | Solanaceae |
| Panonychus | citri | 624 | Rutaceae |
| Tetranychus | ludeni | 884 | polyphage |
| Eutetranychus | orientalis | 321 | Rutaceae, Euphorbiaceae |
| Bryobia | rubrioculus | 57 | Arbres fruitiers (Rosaceae) |
| Bryobia | praetiosa | 54 | Polyphage herbacées (Poaceae, Asteraceae, Fabaceae) |
| Panonychus | ulmi | 623 | Arbres fruitiers (Rosaceae) |
| Tetranychus | turkestani | 893 | polyphage |
| Bryobia | kissophila | 89 | Lierre |
| Oligonychus | coffea | 327 | Thé, café, arbres tropicaux persistants |
| Eutetranychus | banksi | 567 | Rutaceae, Euphorbiaceae |
| Eutetranychus | africanus | 322 | polyphage |

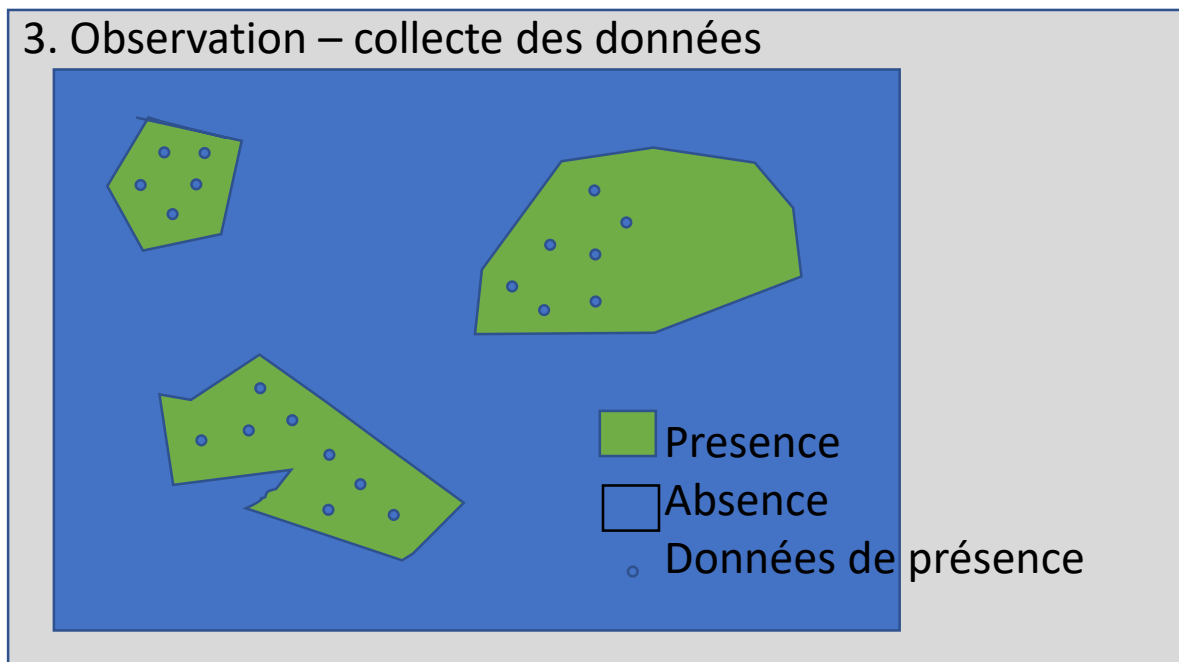
Espèces limitées par le climat et non par les cultures

Grande diversité des distributions des espèces



Que faire quand il n'y a que des données de présence à disposition?
Comment générer des absences sur des sites qui ne soient pas associés
à des conditions environnementales favorables pour l'espèce?

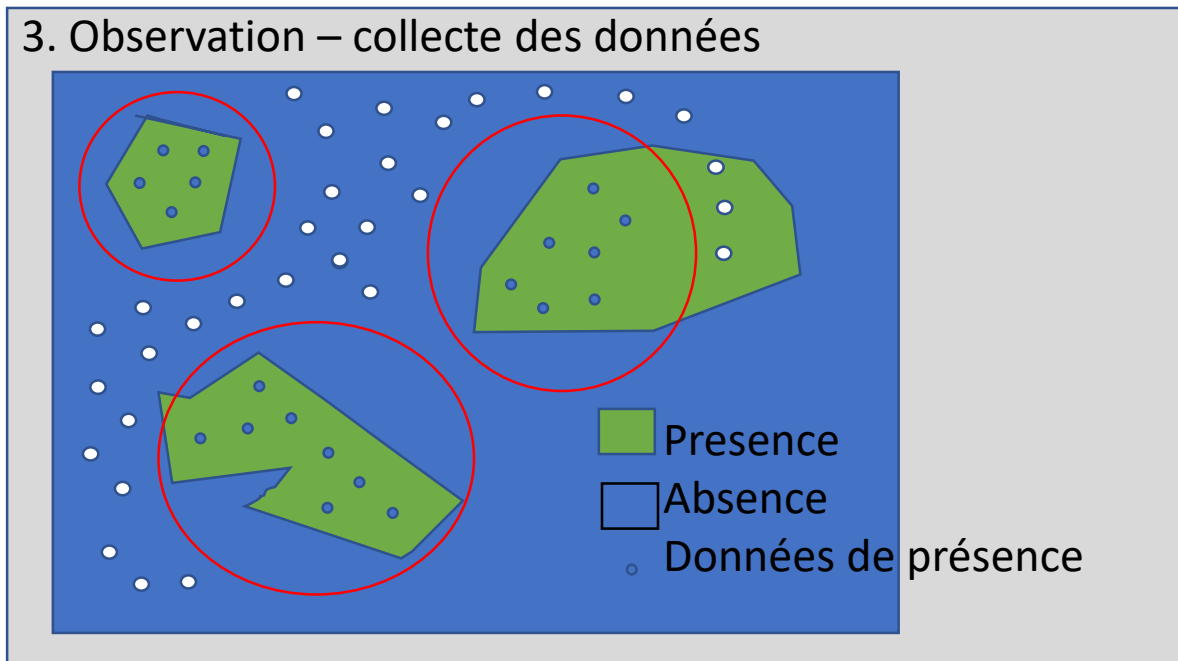
Quelques modèles ne se basent que sur des données de présence (**Bioclim, Domain**)
la majorité des modèles requièrent des données de présence et d'absence (**GAM, GLM, BRT, Random Forest, Maxent**) => création de **pseudo-absences**



Que faire quand il n'y a que des données de présence à disposition?
Comment générer des absences sur des sites qui ne soient pas associés
à des conditions environnementales favorables pour l'espèce?

Comparaison et évaluation de 3 méthodes:

1. À une distance fixe des données de présence (dist min)

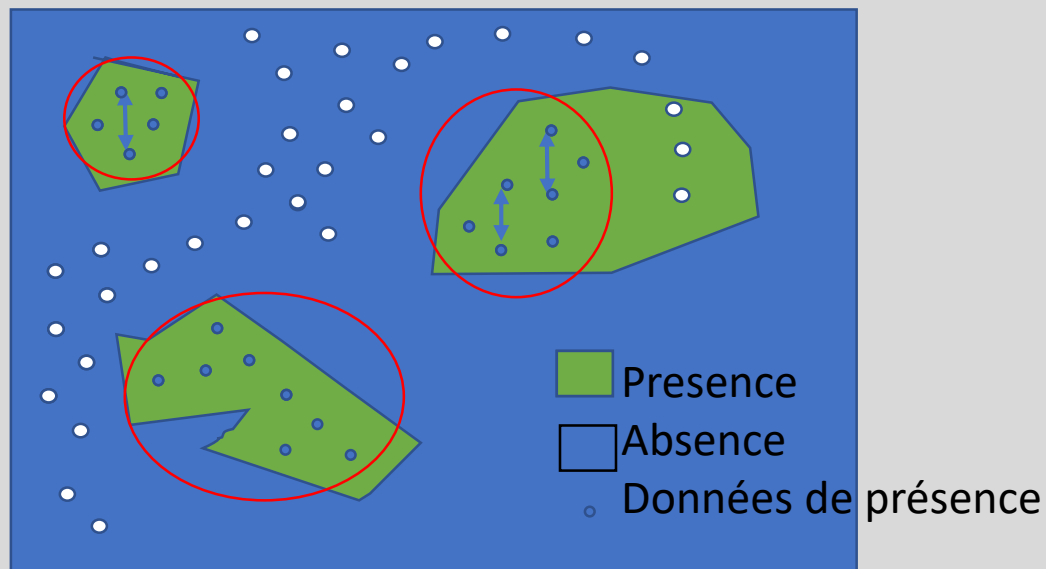


Que faire quand il n'y a que des données de présence à disposition?
Comment générer des absences sur des sites qui ne soient pas associés
à des conditions environnementales favorables pour l'espèce?

Comparaison et évaluation de 3 méthodes:

1. À une distance fixe des données de présence
2. A une distance basée sur les distances entre données de présence (dist clust)

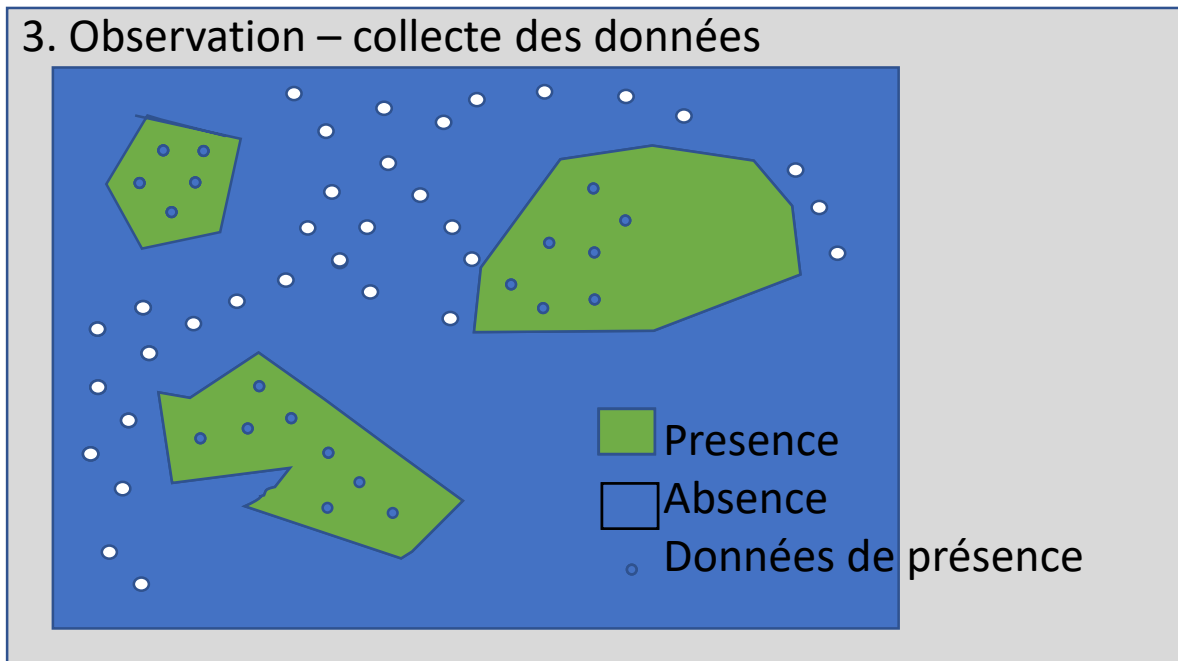
3. Observation – collecte des données



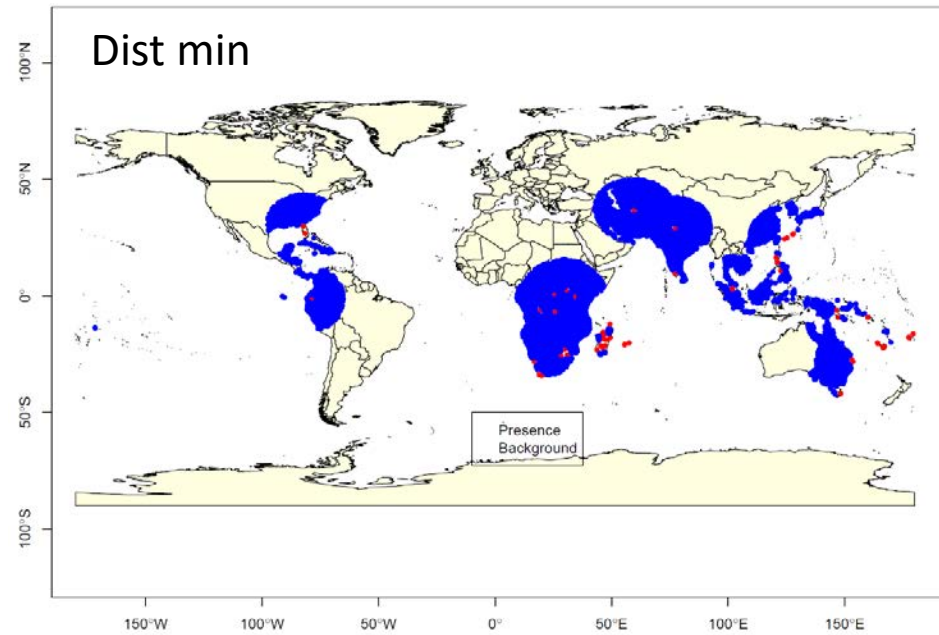
Que faire quand il n'y a que des données de présence à disposition?
Comment générer des absences sur des sites qui ne soient pas associés
à des conditions environnementales favorables pour l'espèce?

Comparaison et évaluation de 3 méthodes:

1. À une distance fixe des données de présence (dist min)
2. A une distance basée sur les distances entre données de présence (dist clust)
3. **A distance des conditions environnementales associées aux données de présence (dist env)**

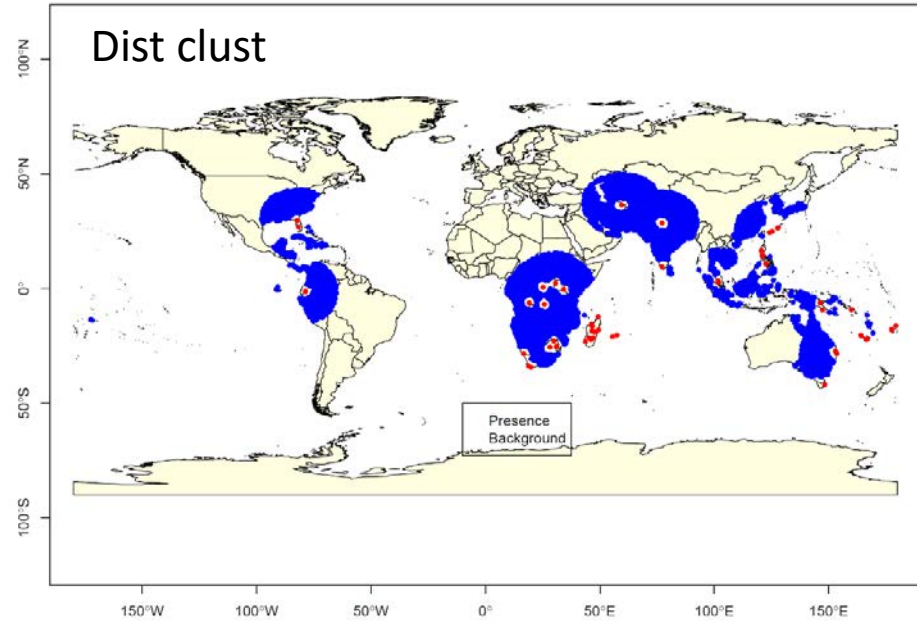


Dist min

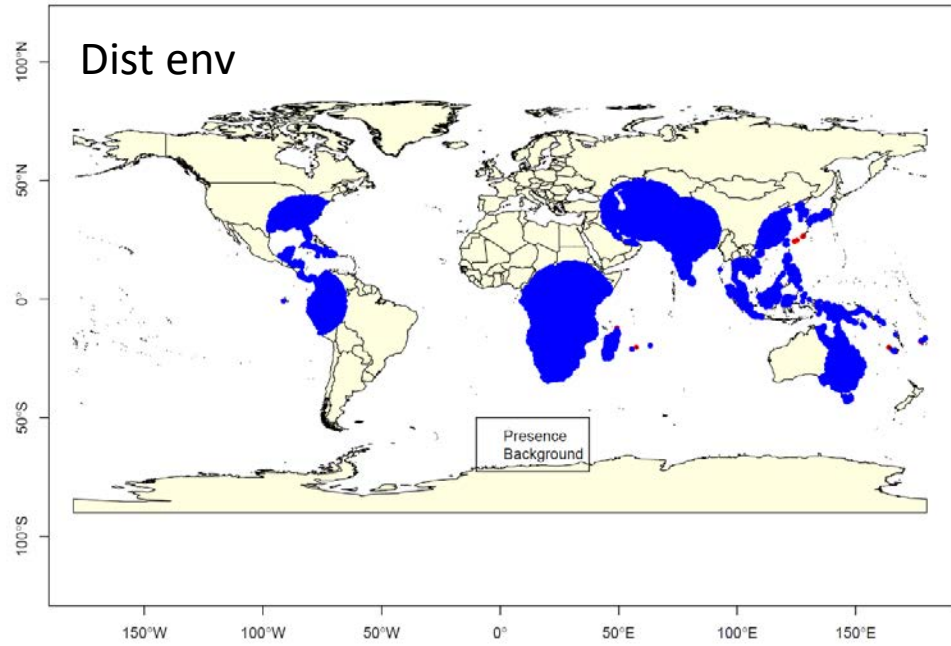


Exemple de *O. coffeae*

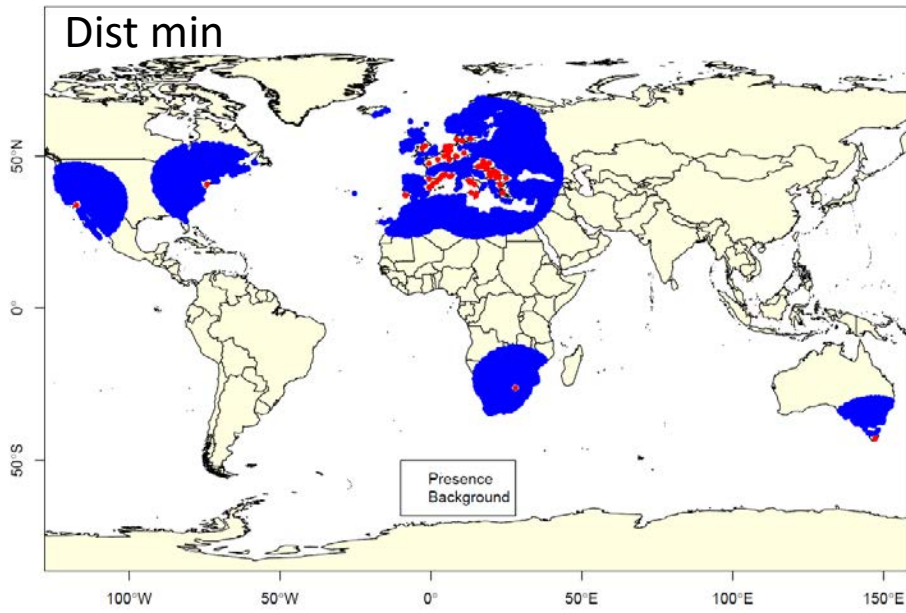
Dist clust



Dist env

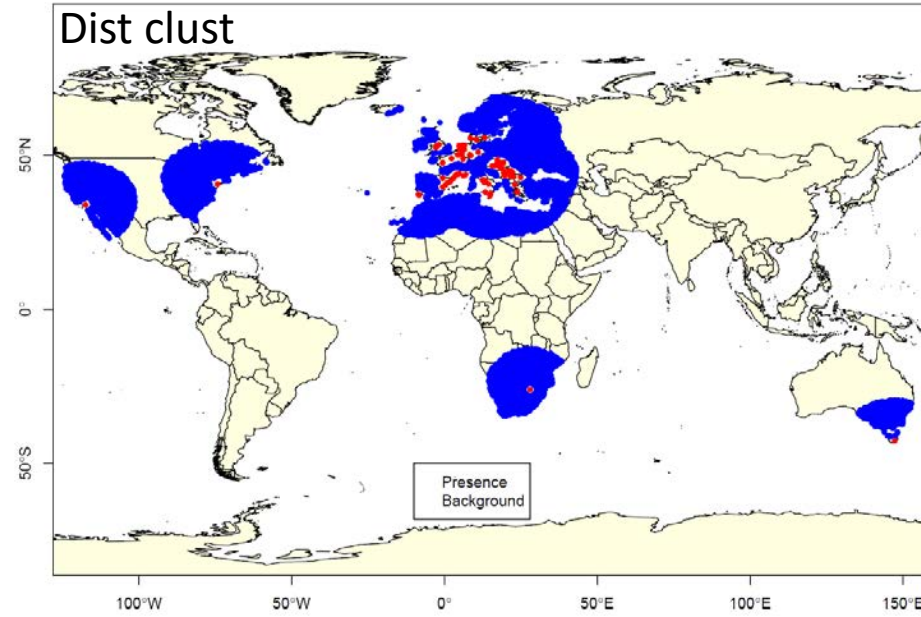


Dist min

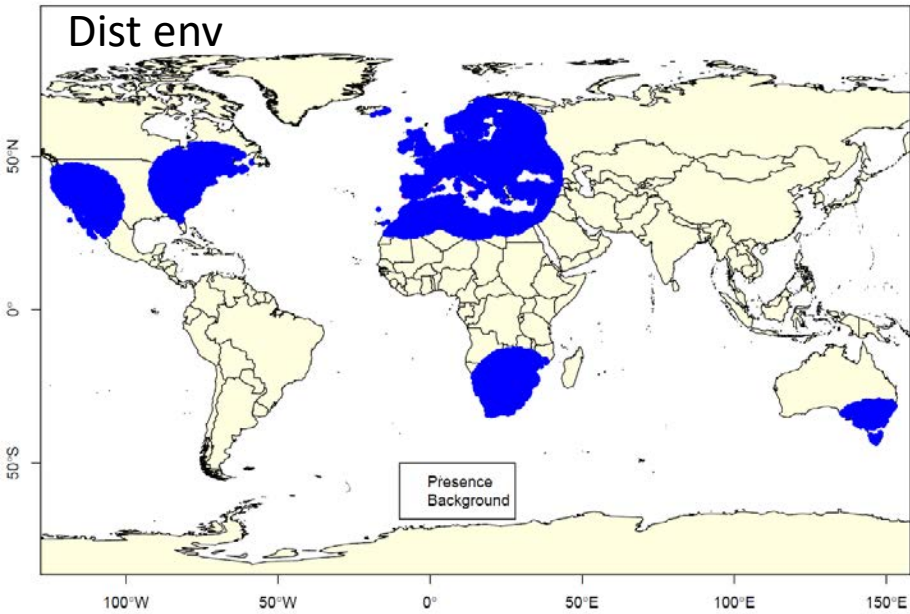


Exemple de *B. kissophila*

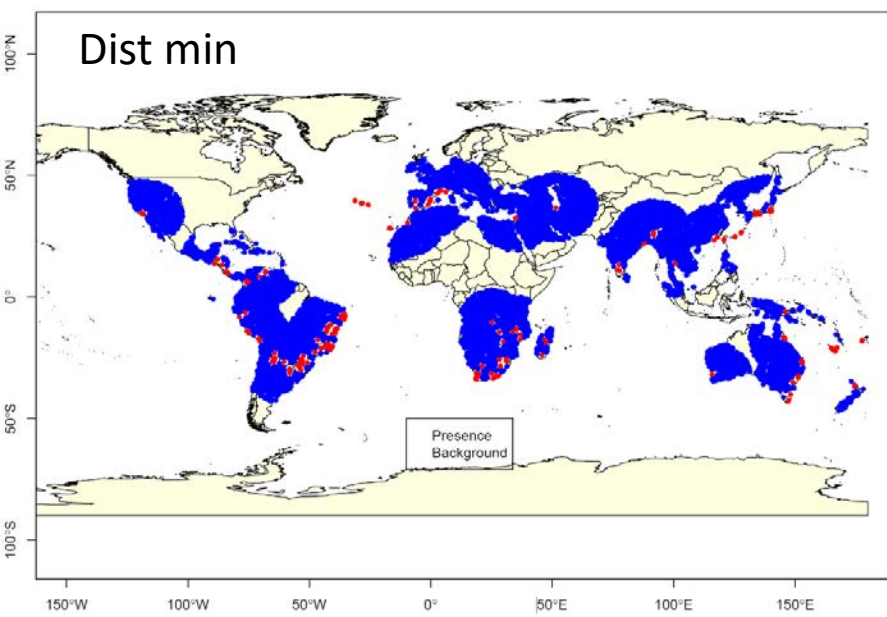
Dist clust



Dist env

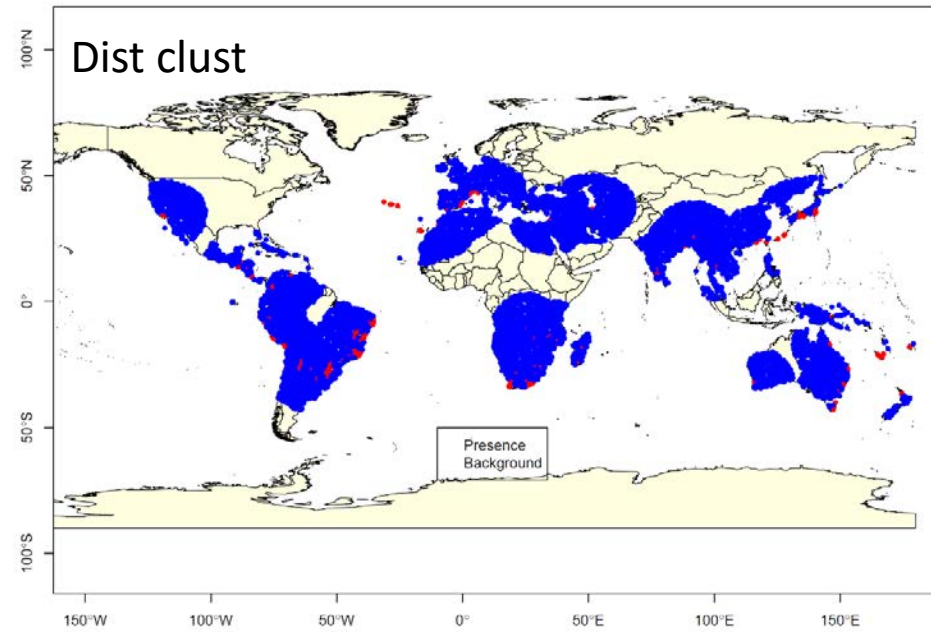


Dist min

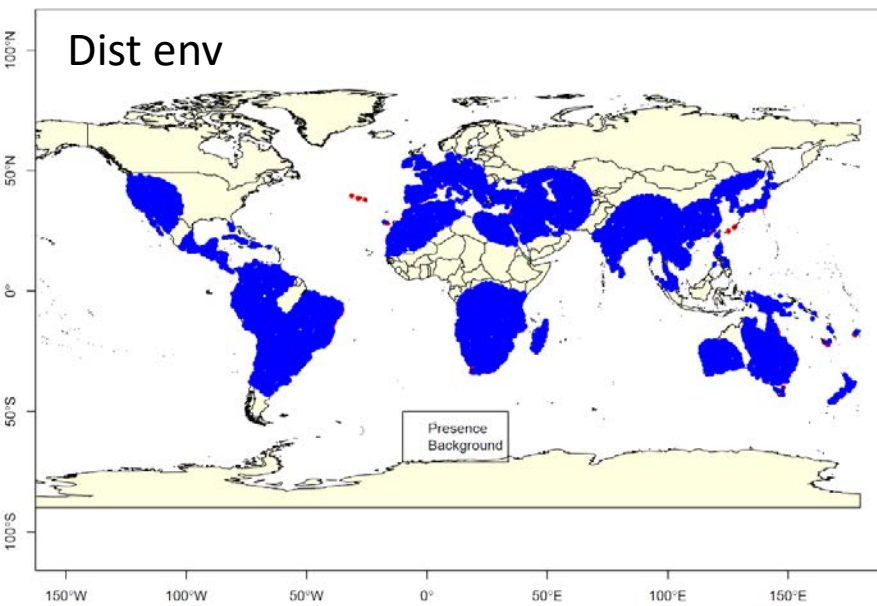


Exemple de *T. ludeni*

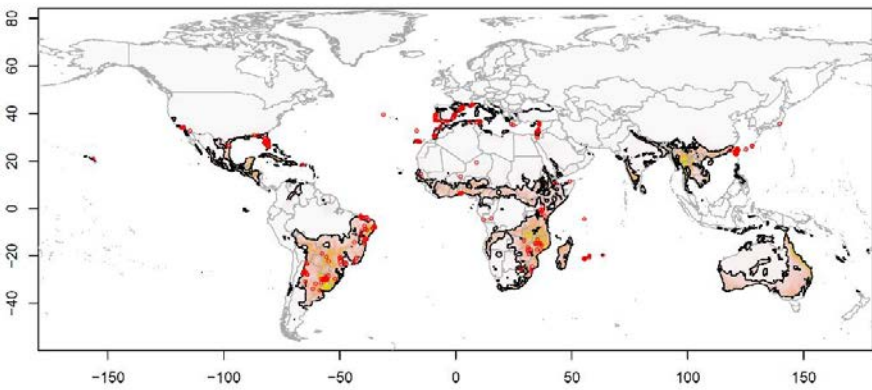
Dist clust



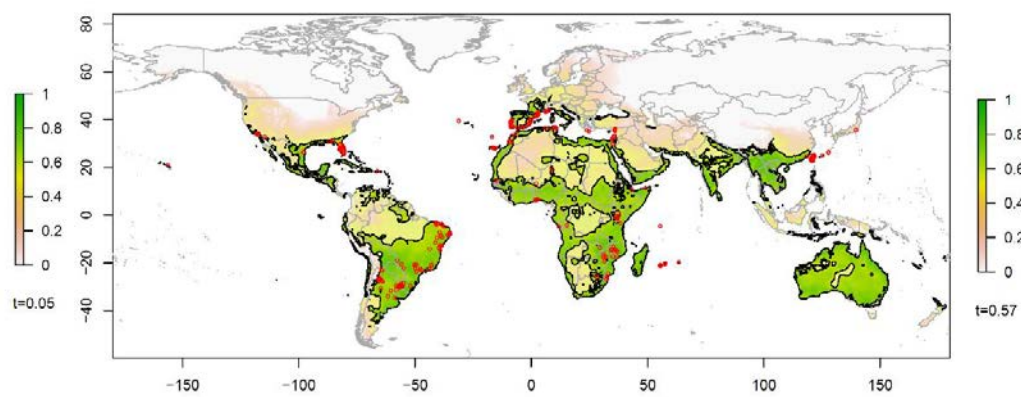
Dist env



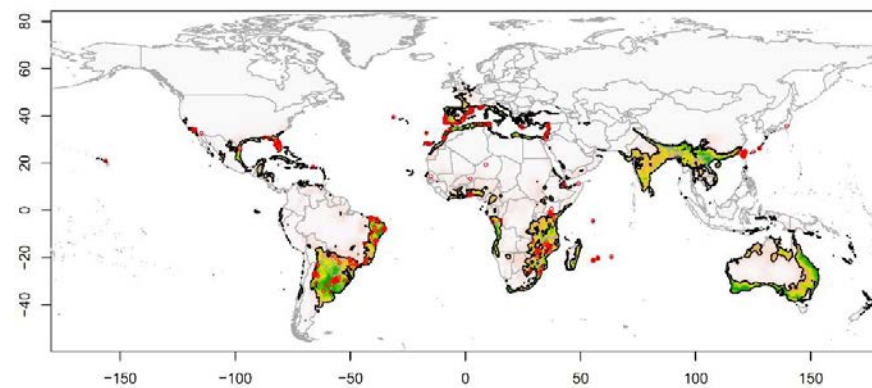
Bioclim, raw values projection outside of range



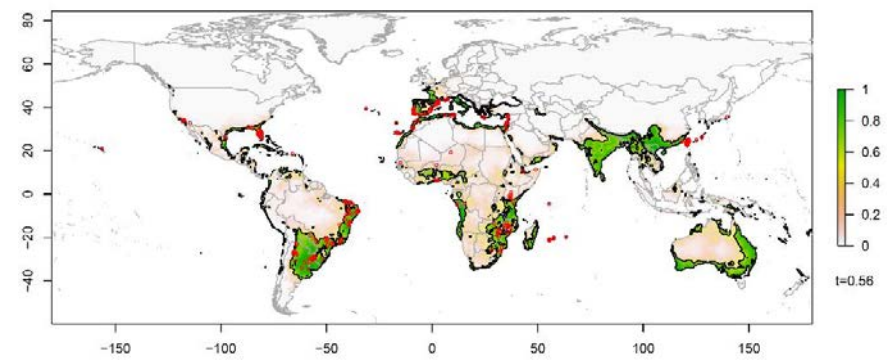
Domain, raw values projection outside of range



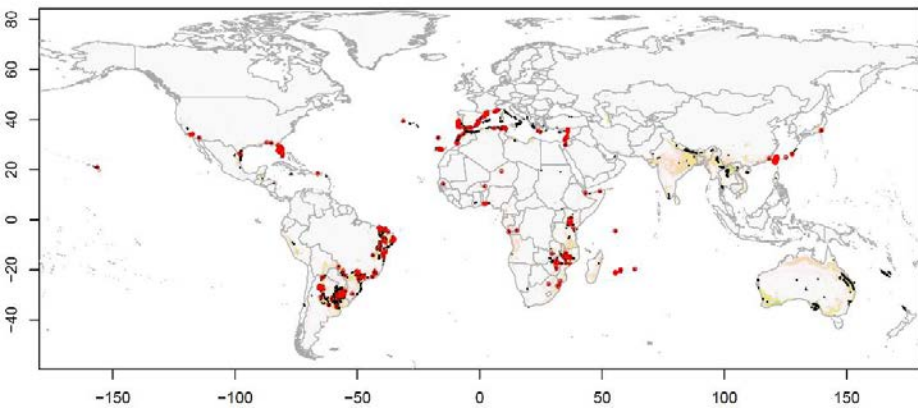
Maxent, raw values projection outside of range



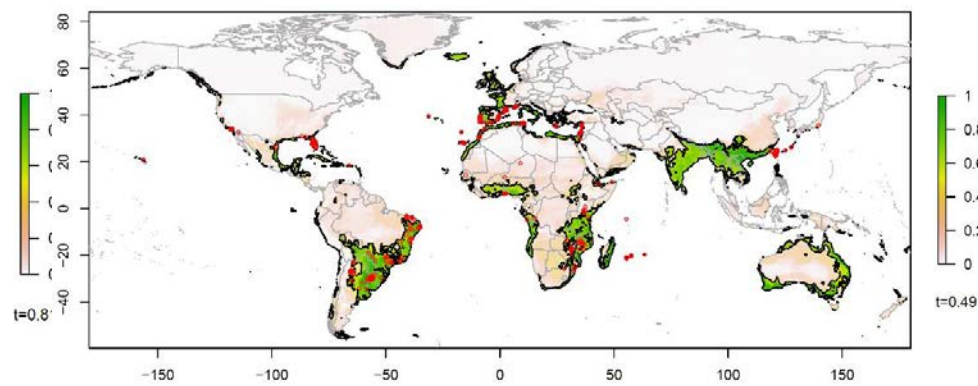
GAM, raw values projection outside of range



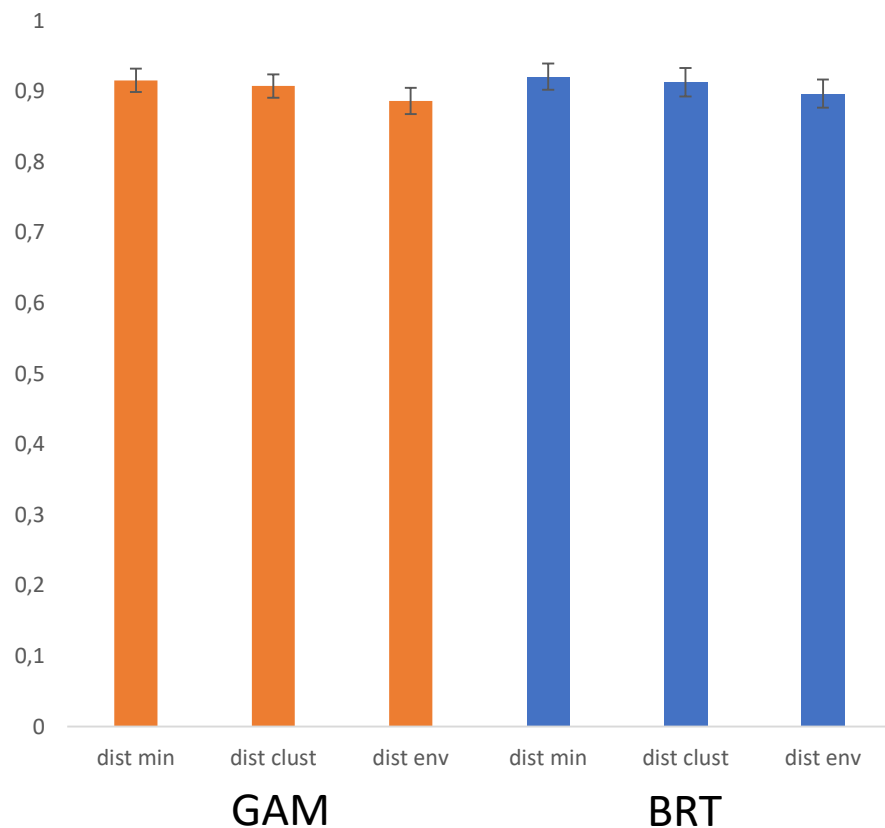
Random Forest, raw values projection outside of range



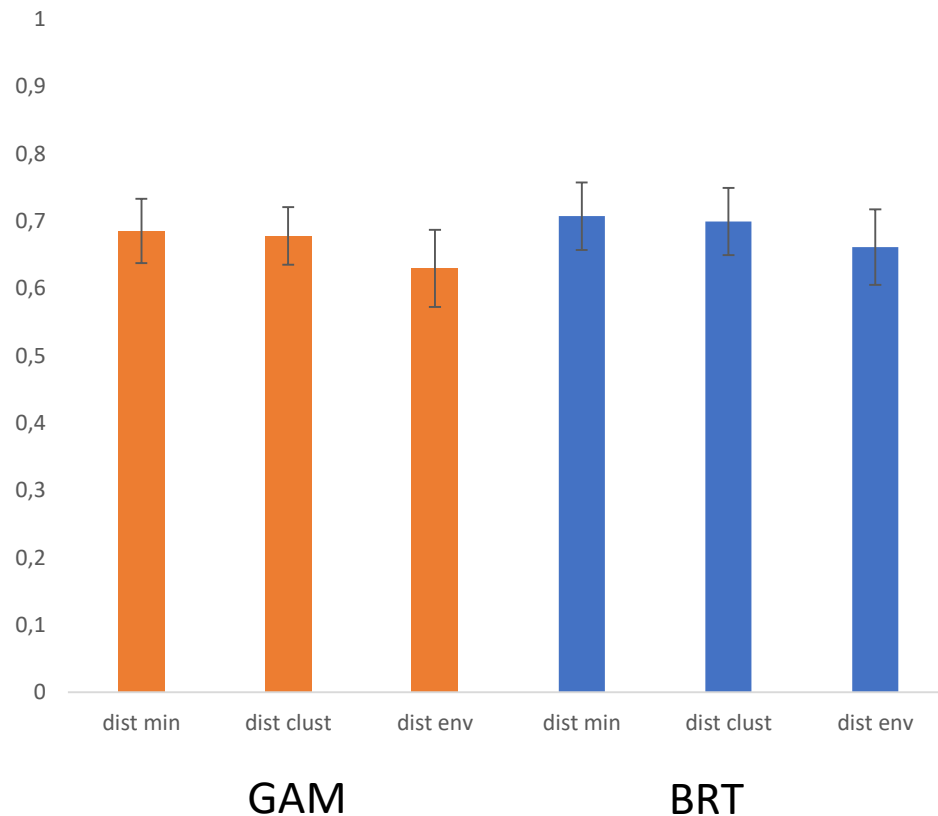
BRT, raw values projection outside of range



AUC



TSS



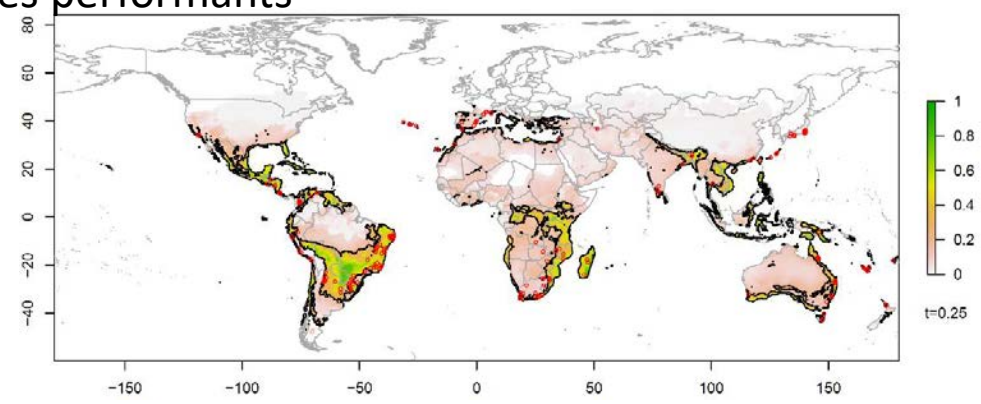
En termes méthodologiques:

- Les trois méthodes sont performantes ($AUC \approx 0.9$ et $TSS > 0.6$)
- Aucune méthode ne ressort comme étant plus performante

Consensus des modèles performants

Consensus mean, projection inside of range

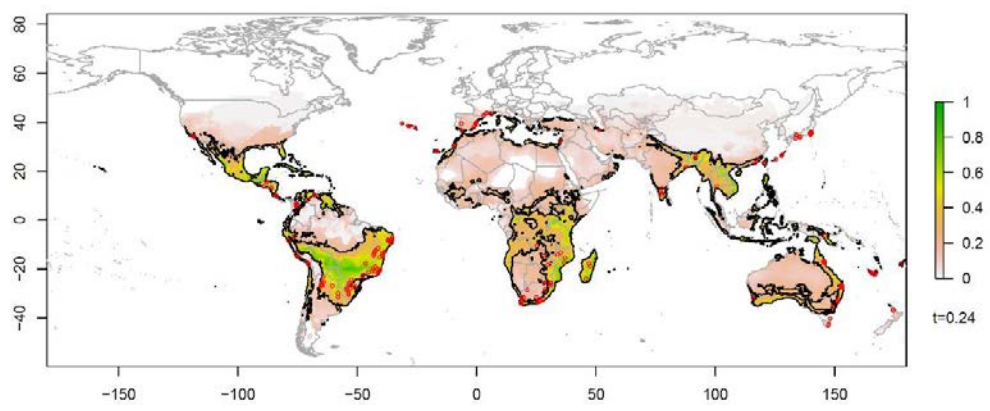
Dist min



T. ludeni

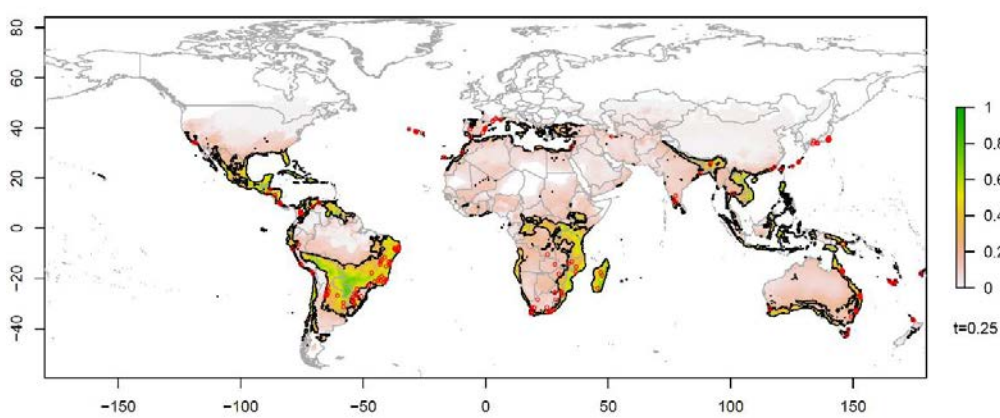
Dist clust

Consensus mean, projection inside of range



Dist env

Consensus mean, projection inside of range



| | threshold | auc | sensitivity | specificity | TSS |
|----------------|-----------|------|-------------|-------------|------|
| <i>bioclim</i> | 0.05 | 0.84 | 0.75 | 0.77 | 0.52 |
| <i>domain</i> | 0.61 | 0.79 | 0.71 | 0.74 | 0.45 |
| <i>brt</i> | 0.53 | 0.94 | 0.88 | 0.88 | 0.76 |
| <i>gam</i> | 0.56 | 0.93 | 0.87 | 0.88 | 0.75 |
| <i>maxent</i> | 0.21 | 0.94 | 0.9 | 0.86 | 0.76 |

D'un point de vue cartographie des enjeux:

- Les trois méthodes sont performantes (AUC ≥ 0.8)
 - Convergence des cartographies consensus entre les trois méthodes
- ⇒ validation des cartes
- ⇒ Projection pour 2050 et 2100 avec scénario A2 CSIRO et A1B MIROC-H

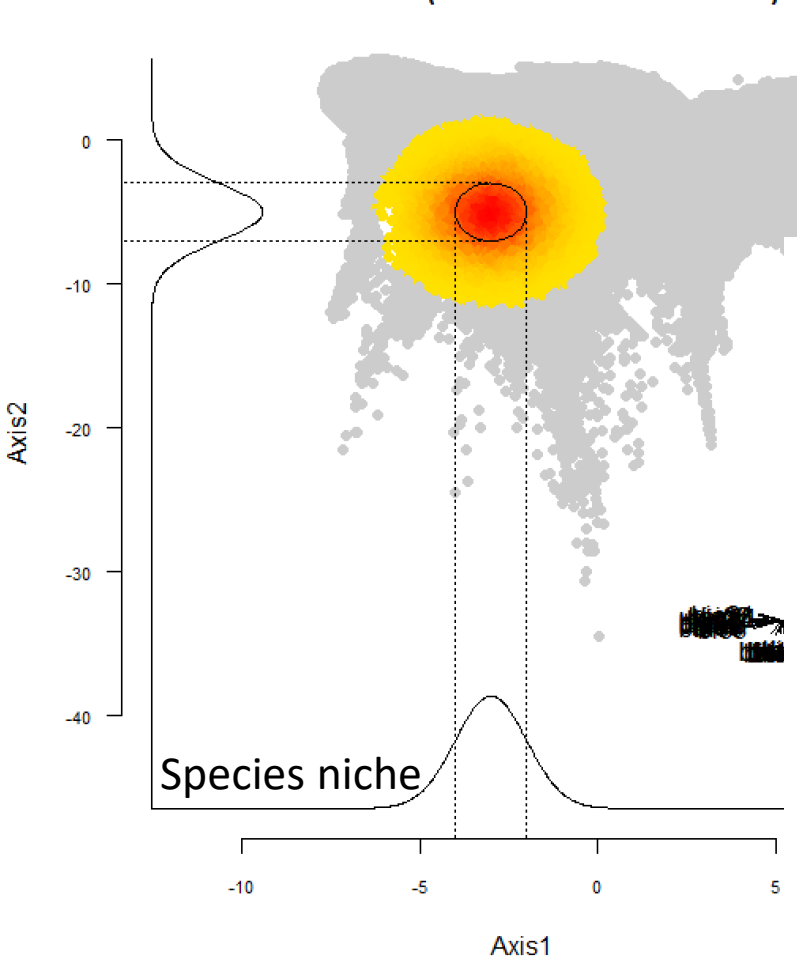
D'un point de vue méthodologique:

- Aucune méthode ne ressort comme étant plus performante
- Problème posé par l'évaluation de données qu'on a nous même créées – comment être sûr?

| | threshold | auc | sensitivity | specificity | TSS |
|---------------------|-----------|------|-------------|-------------|------|
| <i>bioclim</i> | 0.06 | 0.81 | 0.74 | 0.74 | 0.48 |
| <i>domain</i> | 0.61 | 0.76 | 0.72 | 0.69 | 0.41 |
| <i>brt</i> | 0.52 | 0.92 | 0.88 | 0.85 | 0.73 |
| <i>gam</i> | 0.52 | 0.91 | 0.87 | 0.83 | 0.69 |
| <i>maxent</i> | 0.26 | 0.92 | 0.88 | 0.83 | 0.71 |
| <i>randomforest</i> | 0.8 | 0.97 | 0.74 | 1 | 0.73 |

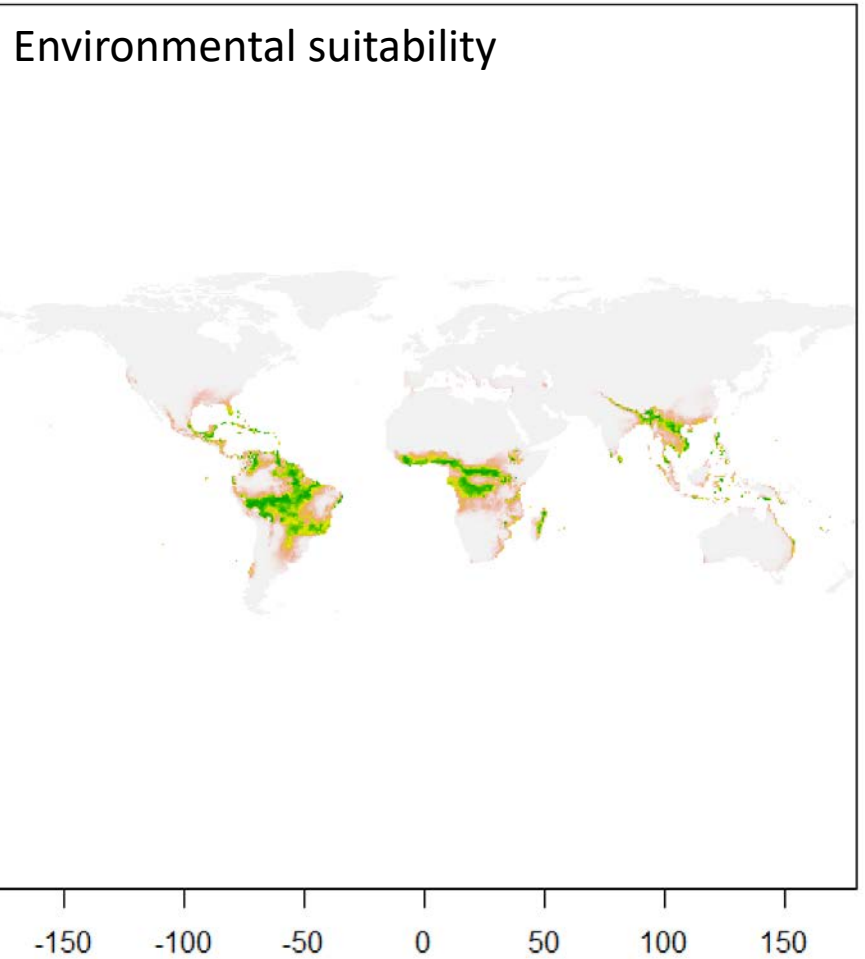
Génération d'espèces virtuelles - espèce spécialiste

PCA of environmental conditions
Axes 1 & 2 (2 axes included in total)

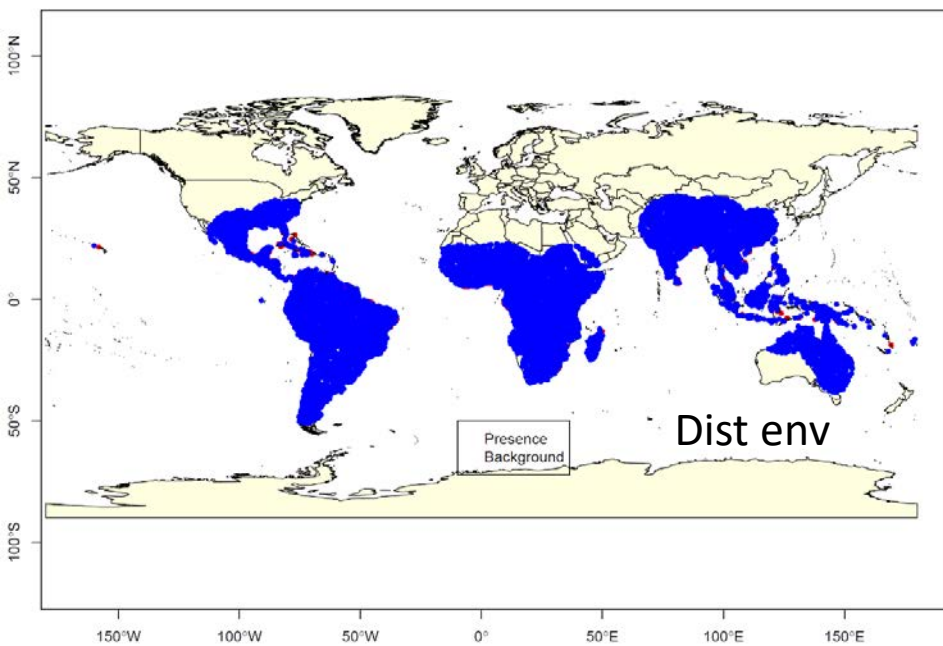
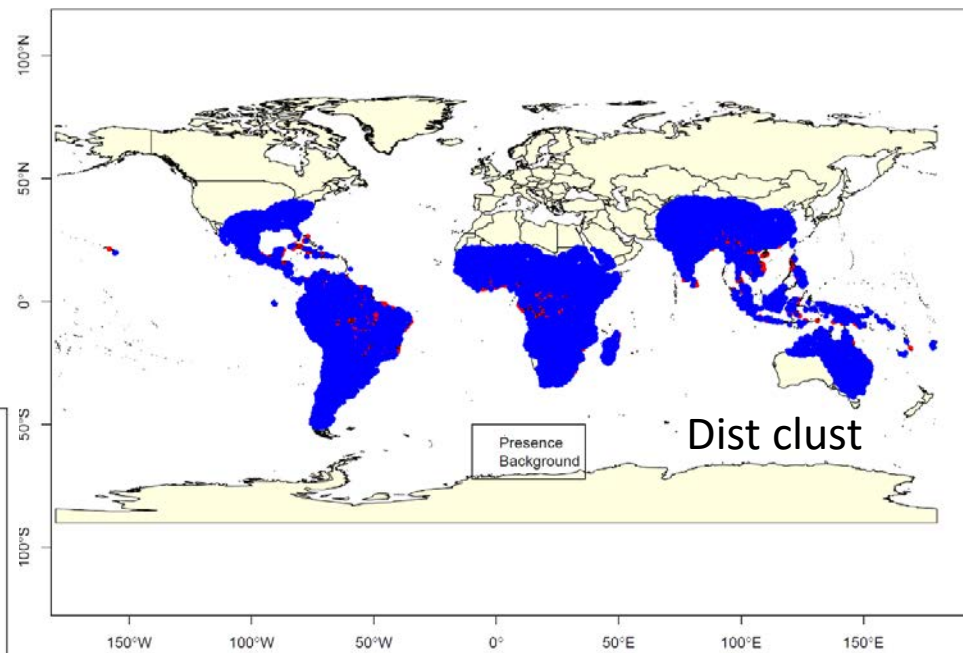
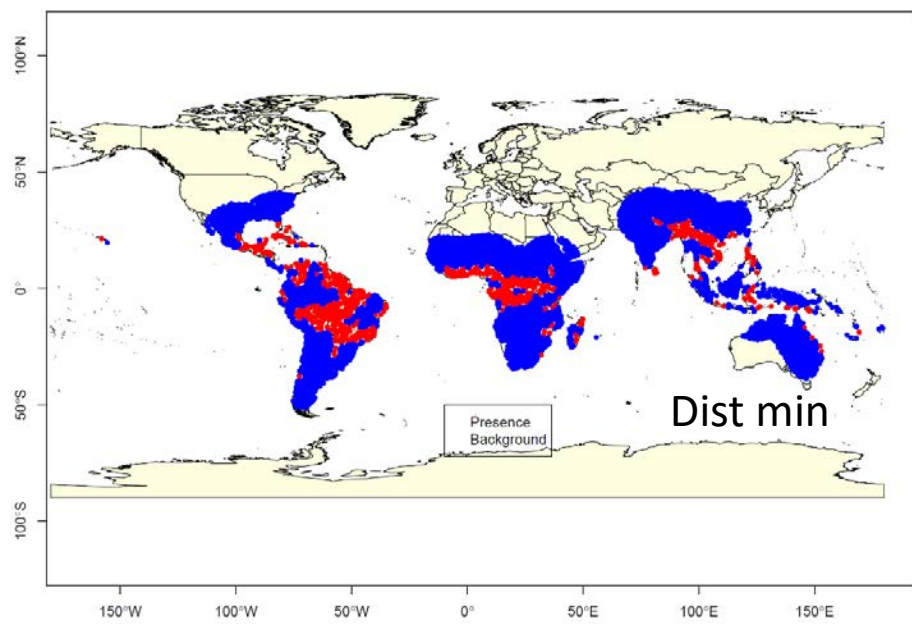


Pixel suitability

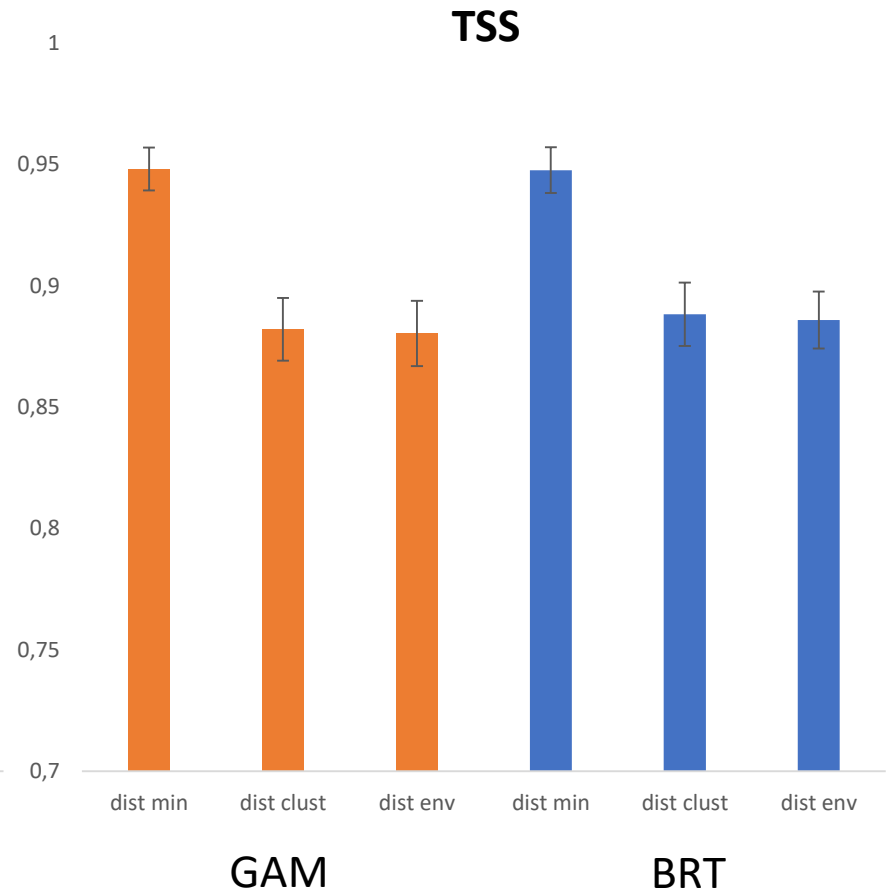
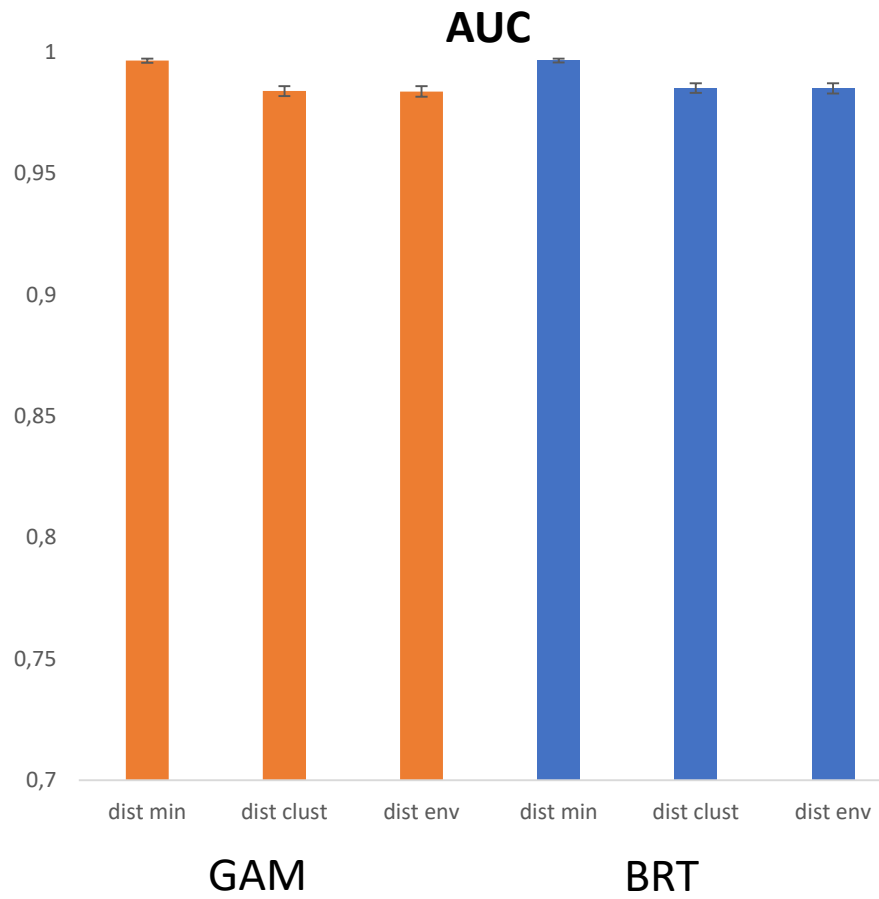
- 1
- 0.8
- 0.6
- 0.4
- 0.2
- 0



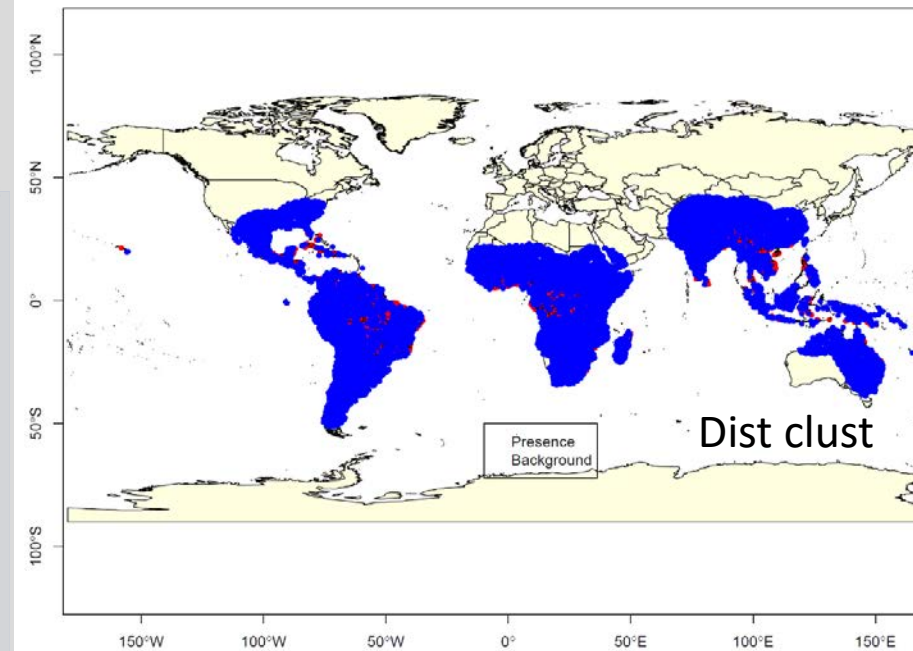
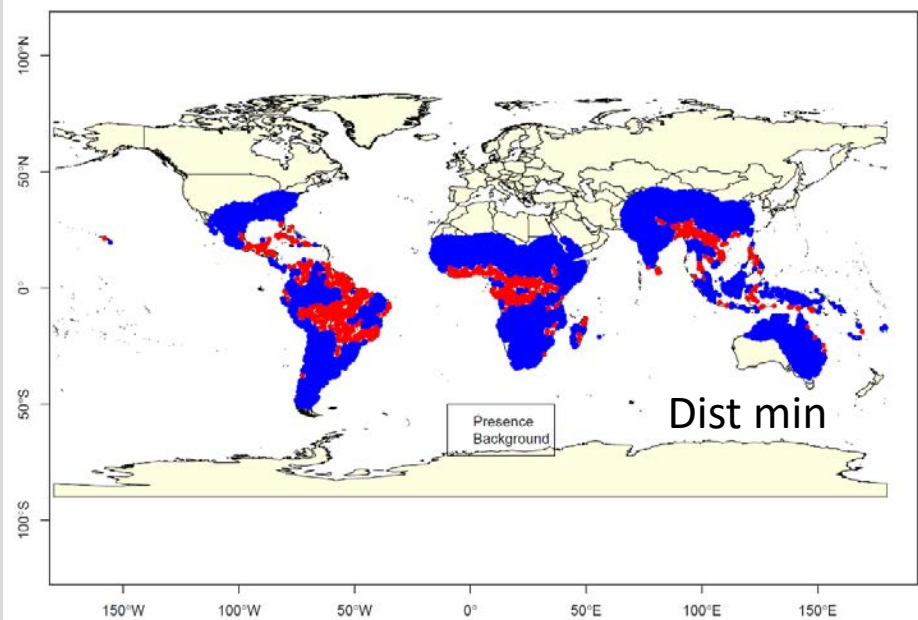
0.8
0.6
0.4
0.2
0.0



1000 données

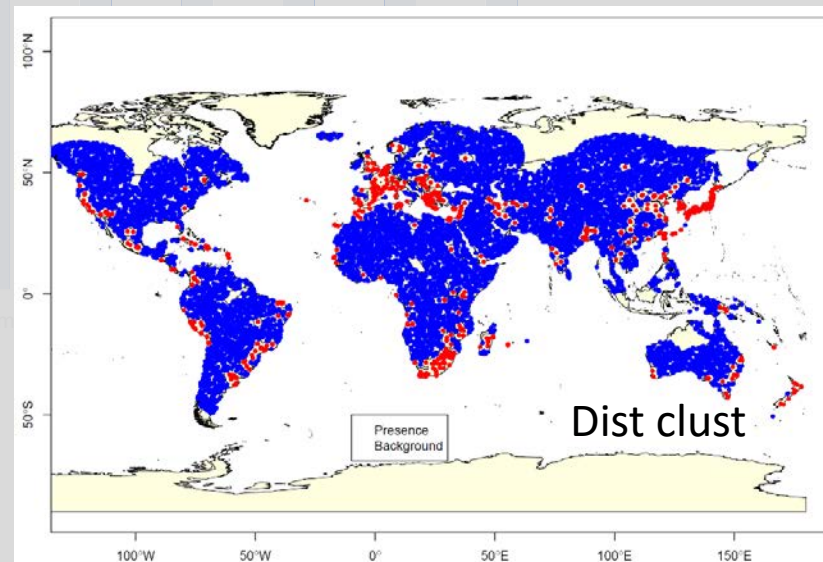
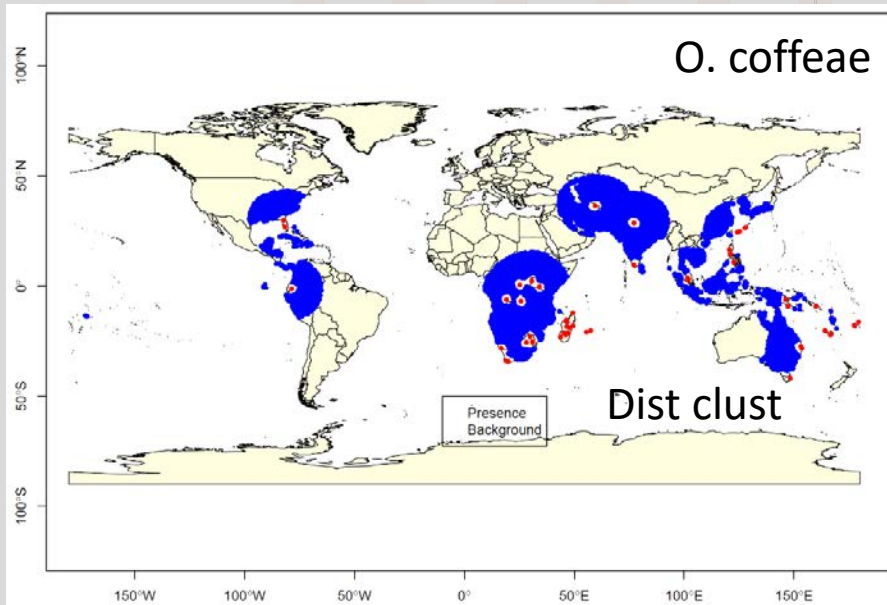


Espèce virtuelle



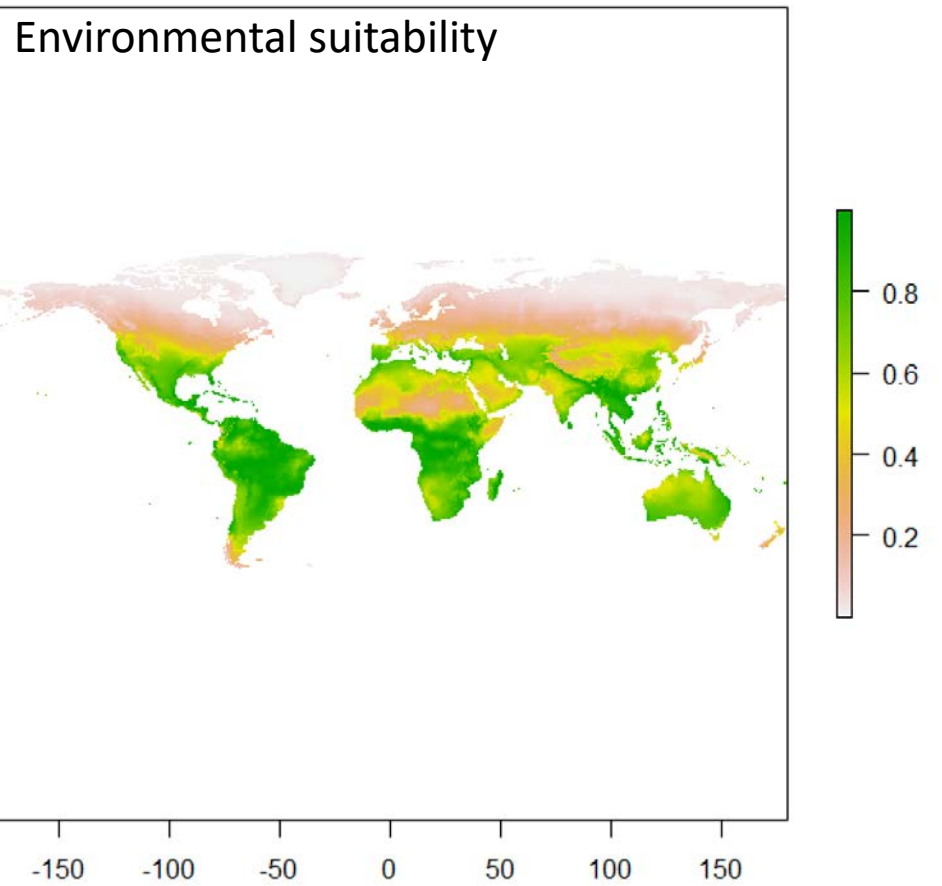
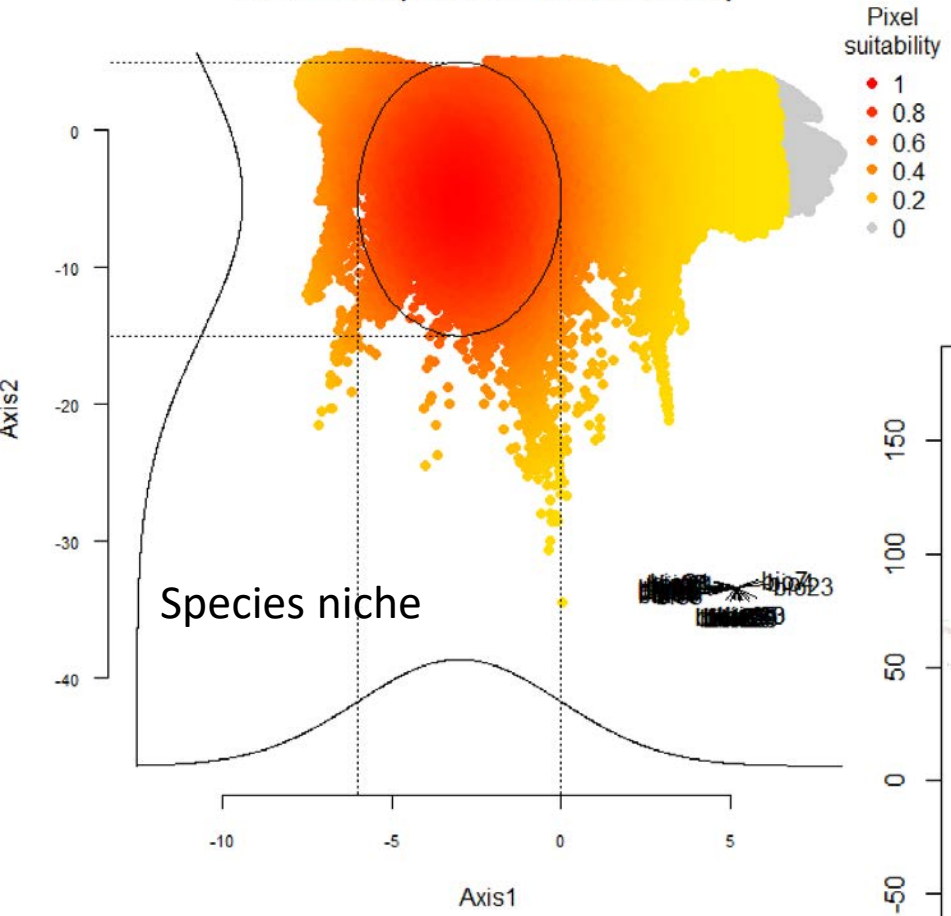
Espèces réelles

Absence de structure spatiale

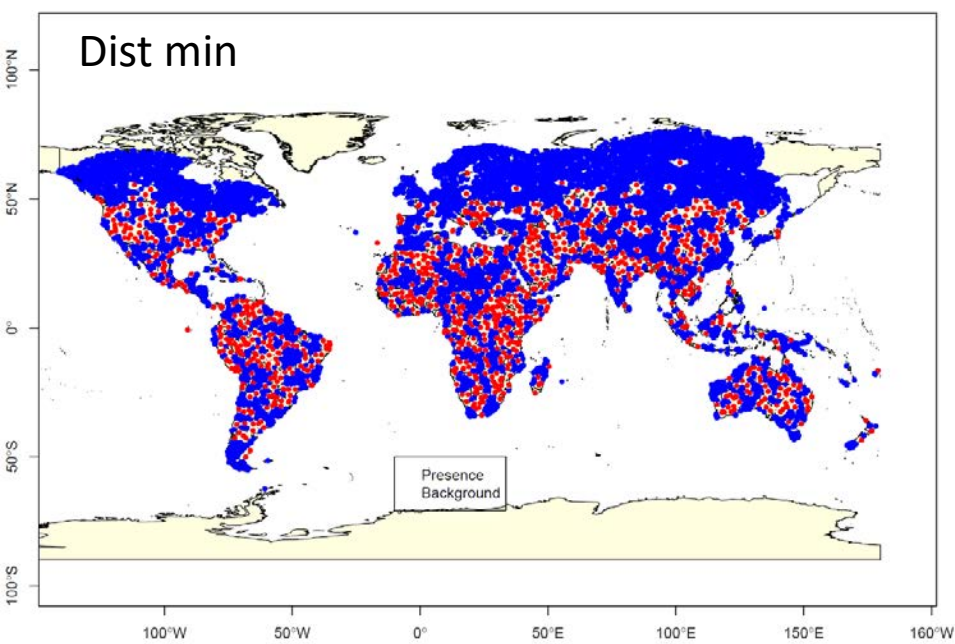


Génération d'espèces virtuelles - espèce généraliste

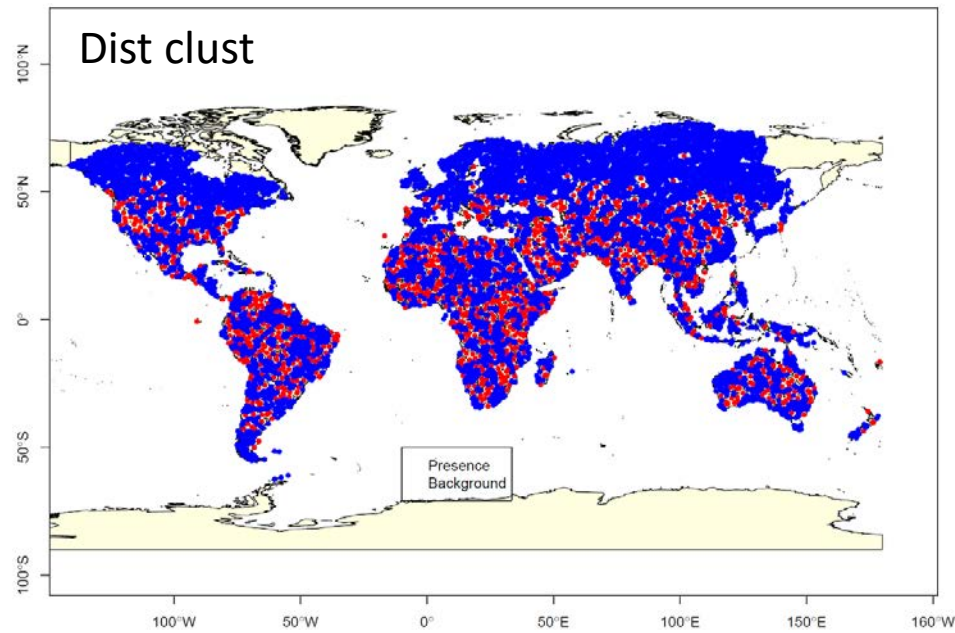
PCA of environmental conditions
Axes 1 & 2 (2 axes included in total)



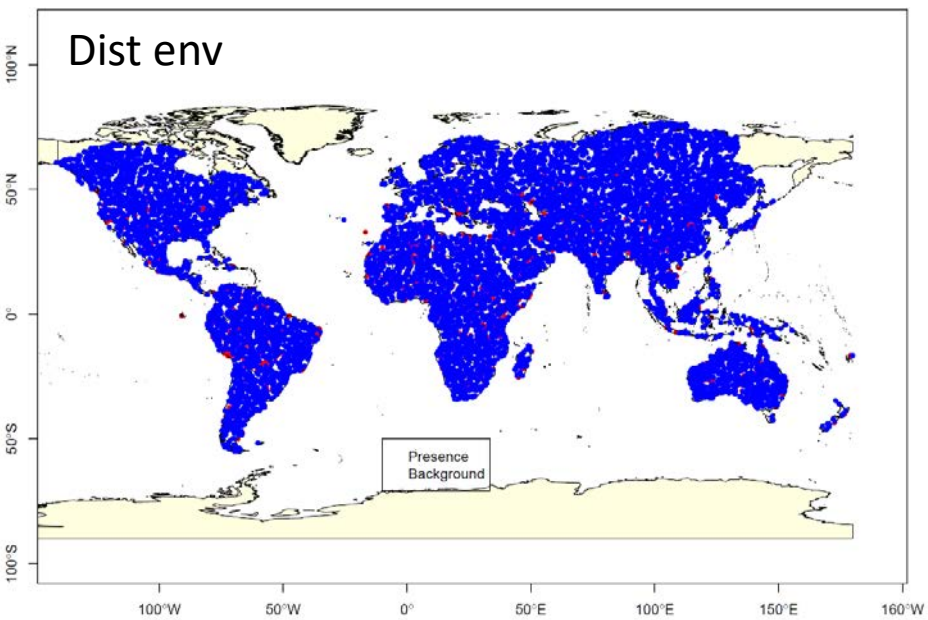
Dist min



Dist clust

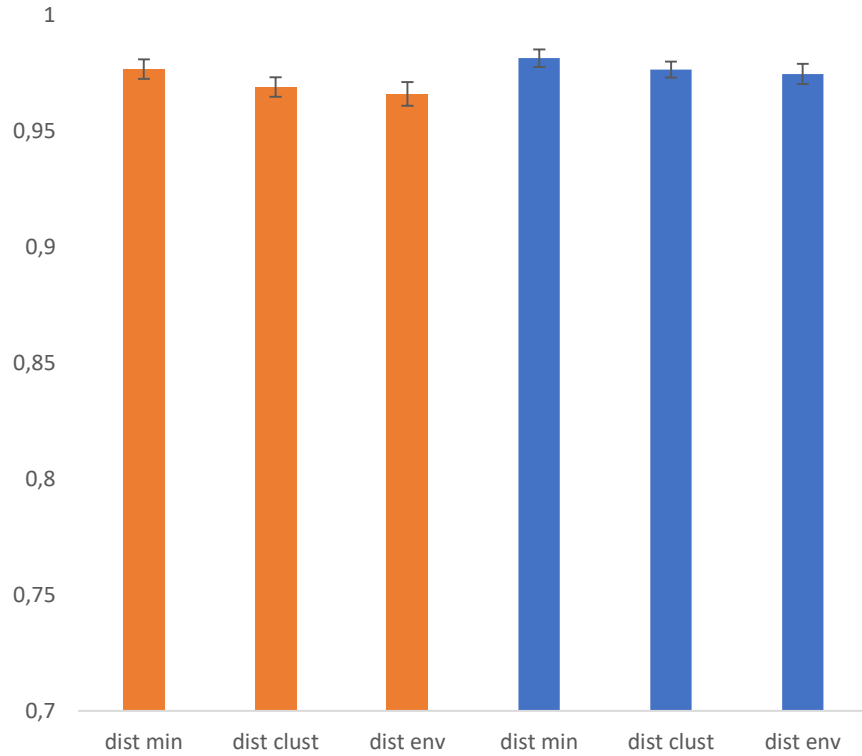


Dist env



1000 données

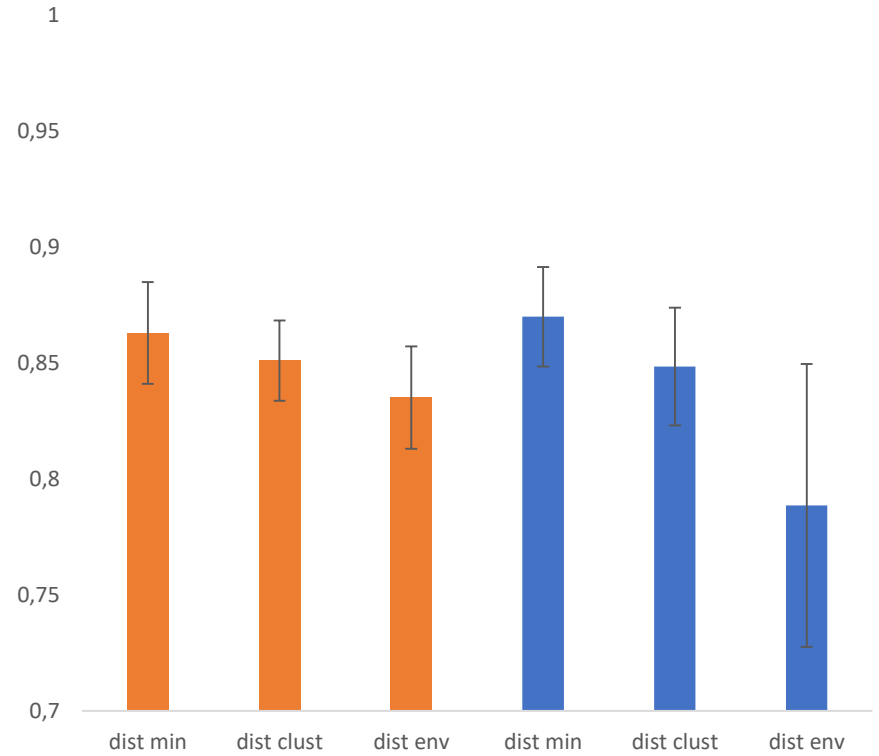
AUC



GAM

BRT

TSS

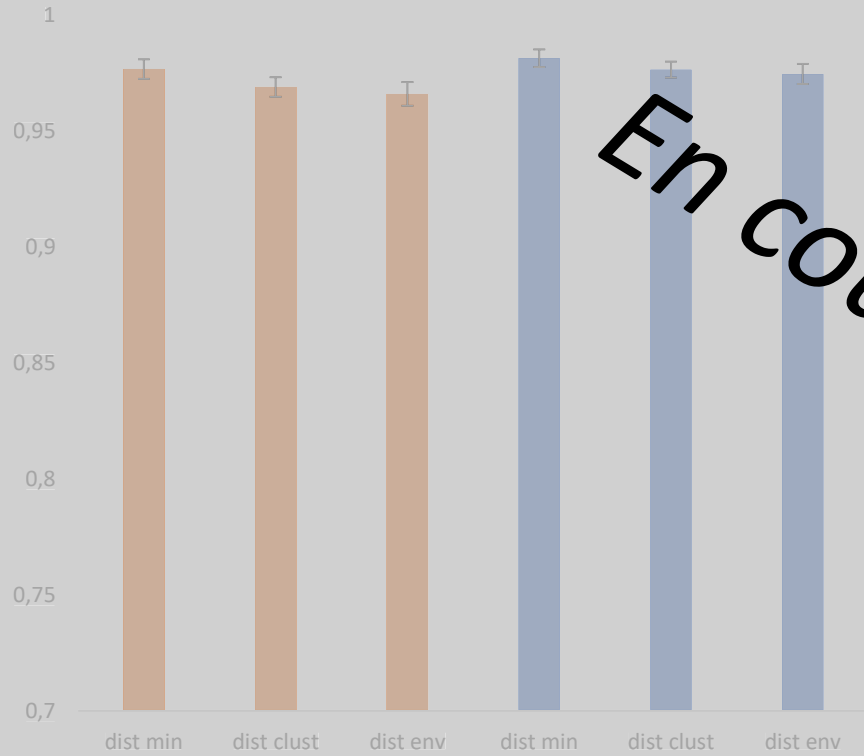


GAM

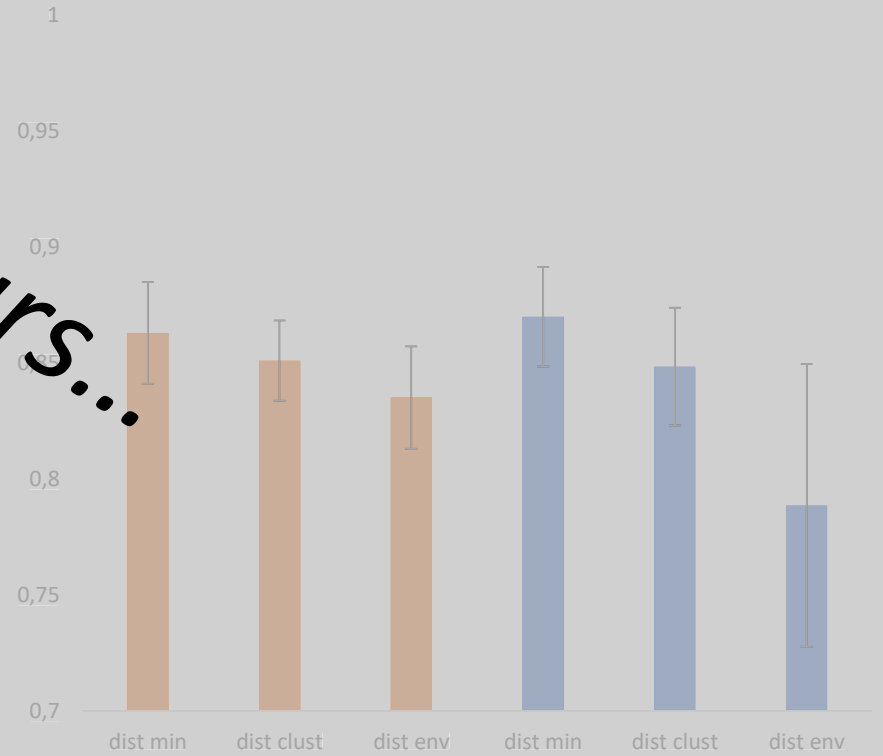
BRT

1000 données

AUC



TSS



En cours...

GAM

BRT

GAM

BRT

Conclusion – espèces virtuelles

- Quelques ajustements encore à faire sur espèces virtuelles
- Ajout de structuration spatiale dans les données pour s'approcher plus de la réalité
- Utilisation des espèces virtuelles pour choisir une méthode adaptée pour la génération des pseudo-absences (différence entre spécialiste et généraliste)



Merci de votre attention!

