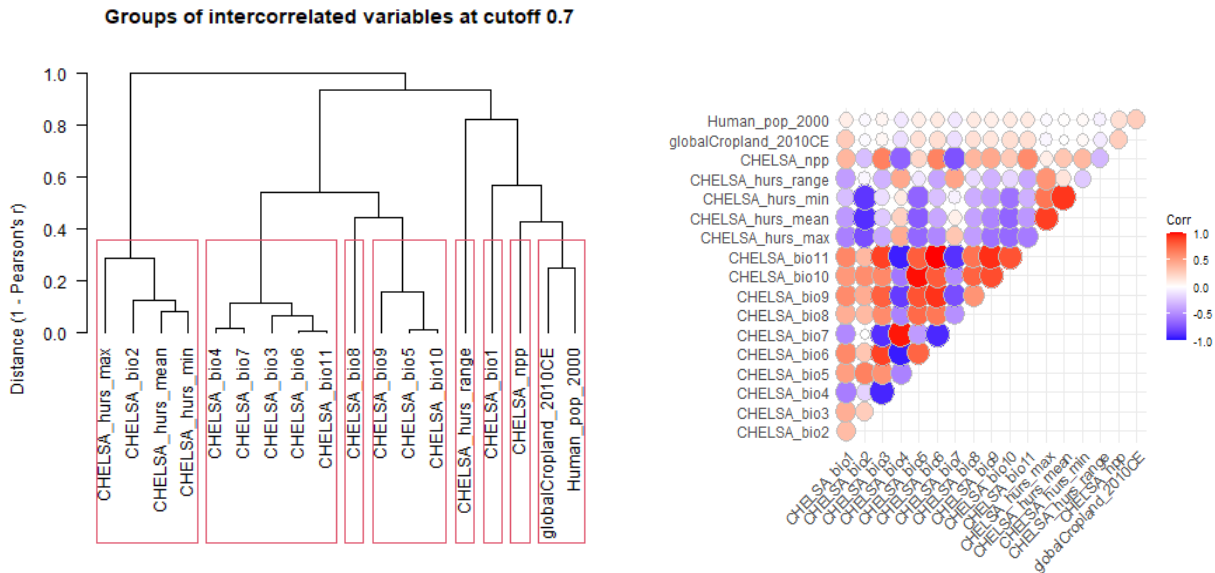


Test-0

Eléna

2025-01-20

J'ai appliqué un facteur 15 à tous mes rasters ($0.5\text{arcminutes} \times 15 = 7.5\text{arcminutes}$) pour que le code tourne plus rapidement.



Nous avons fait le choix de ne garder deux variables de températures (bio5 et bio7) et deux variables d'humidité (Hurs min et range). En effet, humidité minimale joue un rôle très important chez les insectes et l'intervalle d'humidité est plus important que l'humidité maximale.

Pour ce test, nous utilisons donc 6 variables :

- Bio5 (température maximale)
- Bio7 (range température)
- Hurs_min (humidité minimale)
- Hurs_range (range d'humidité)
- NPP (productivité primaire nette)
- globalCropland_2010CE (terres cultivées 2010)

Valeurs au sein ds intervalles = 80 :

```
intervals <- list( bio5 = seq(min(combinations[, 1]), max(combinations[, 1]), length.out = 80),  
hurs_min = seq(min(combinations[, 2]), max(combinations[, 2]), length.out = 80),
```

```

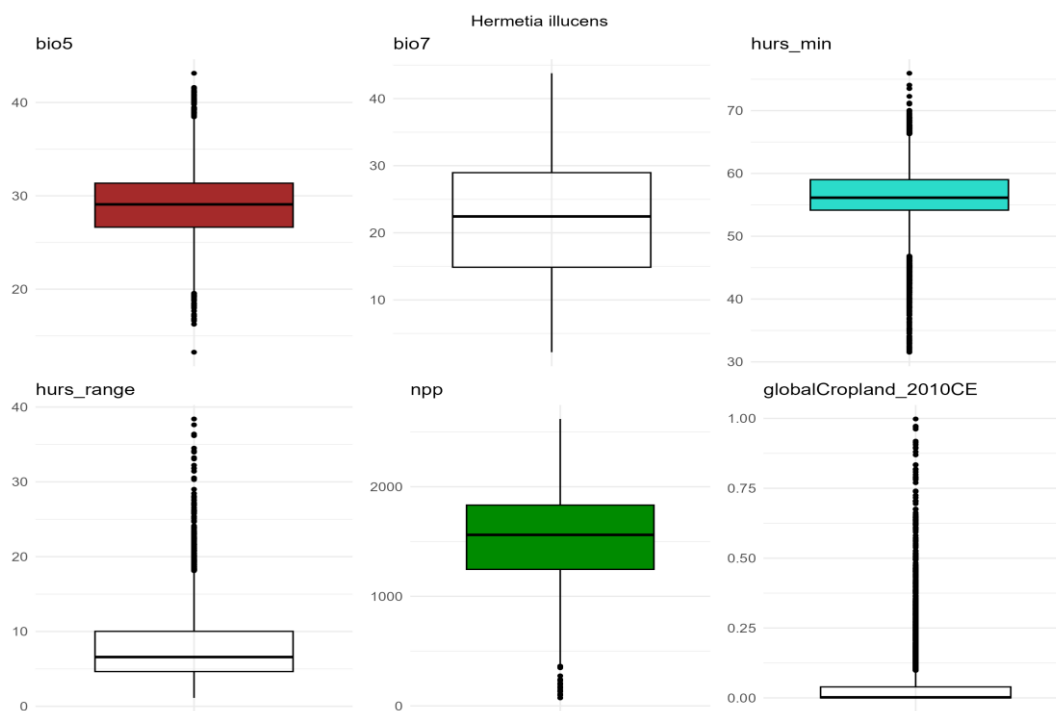
hurs_mean = seq(min(combinations[, 3]), max(combinations[, 3]), length.out = 80),
hurs_range = seq(min(combinations[, 4]), max(combinations[, 4]), length.out = 80), npp =
seq(min(combinations[, 3]), max(combinations[, 3]), length.out = 80), croplands =
seq(min(combinations[, 6]), max(combinations[, 6]), length.out = 80) )

```

Espace environnemental :

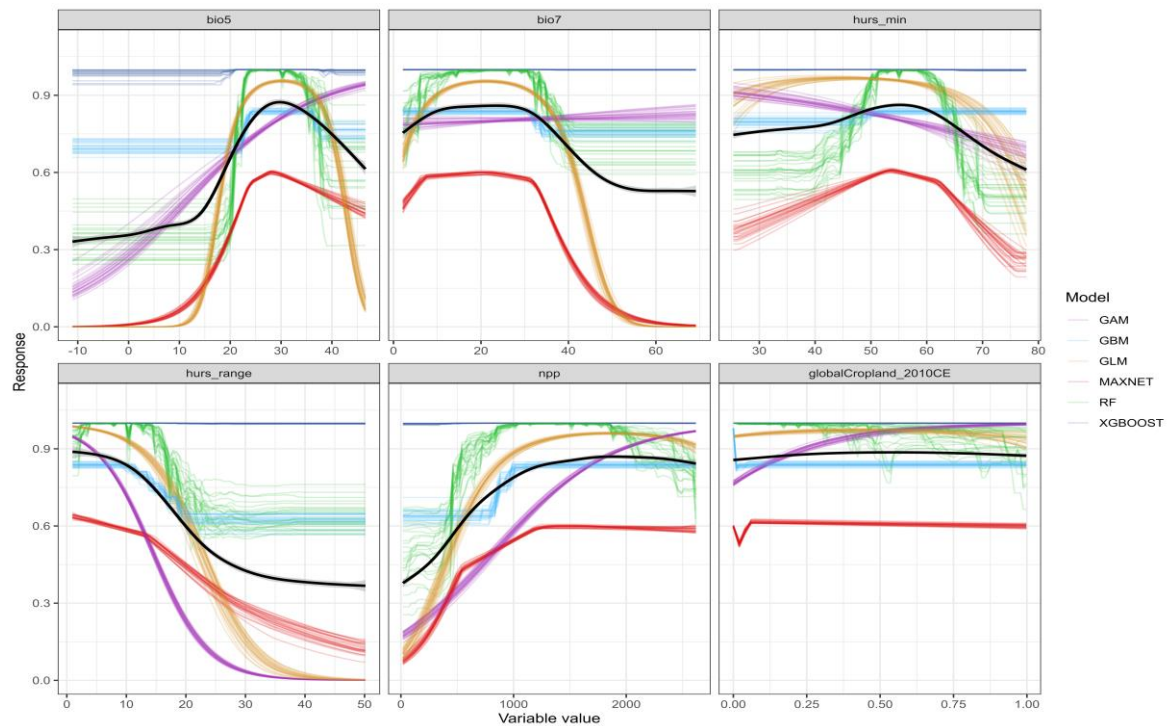
- Total number of cells with environmental conditions in the geographical space: 979220
- Number of duplicated conditions: 374902
- Number of unique cells (environmental space): 604318

Espèce pour le test : mouche soldat noire (*Hermetia illucens*) qui compte 17 216 occurrences avant les étapes de filtrations des occurrences puis 5690 après.



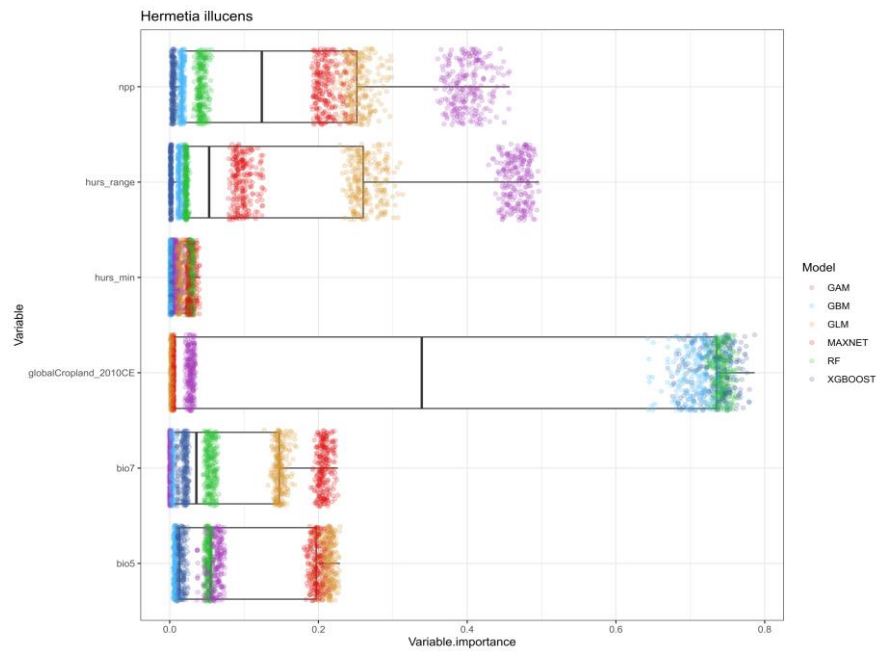
Repartition des occurrences le long des variables

- Majorité des occurrences dans des gammes environnementales assez restreintes.
- Quelques outliers surtout pour l'Humidité min, hurs_range et les terres cultivées.



Courbes de réponses des variables utilisées pour le test 0

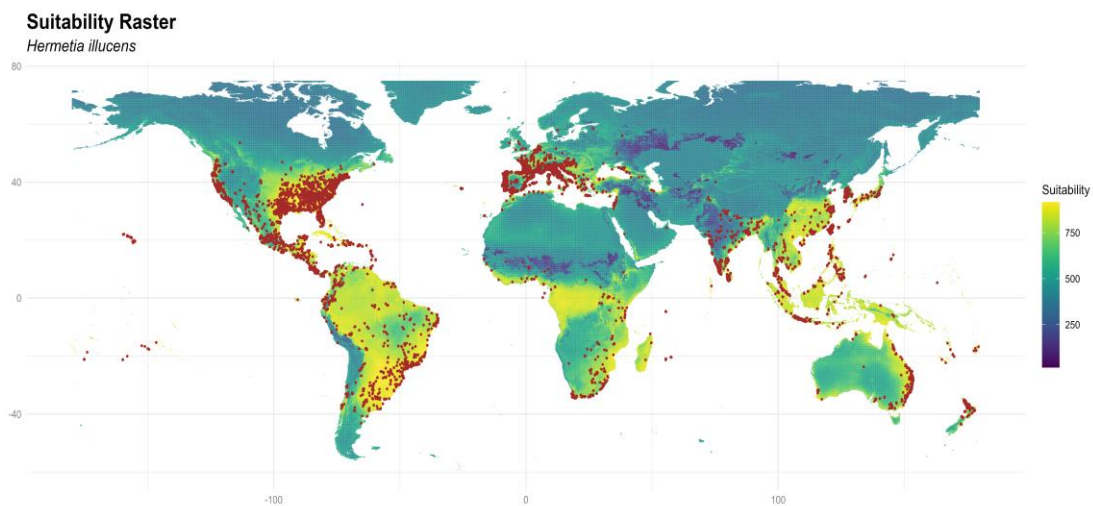
- Les algorithmes présentent une réponse globalement cohérente aux variables en fonction de la présence de l'espèce (en comparaison avec la figure ci-dessus).
- Une réponse inattendue est observée pour la productivité primaire nette, où l'on aurait anticipé une courbe en cloche, compte tenu de la distribution des occurrences le long du gradient.
- L'algorithme XGBOOST ne montre aucune réponse aux variables.
- Les algorithmes ne réagissent pas à la variable des terres cultivées.



Importance des variables dans la prédiction de l'établissement de l'espèce par les modèles

- Les résultats des algorithmes étant très différents, on n'observe pas vraiment de différence d'importance entre les variables.

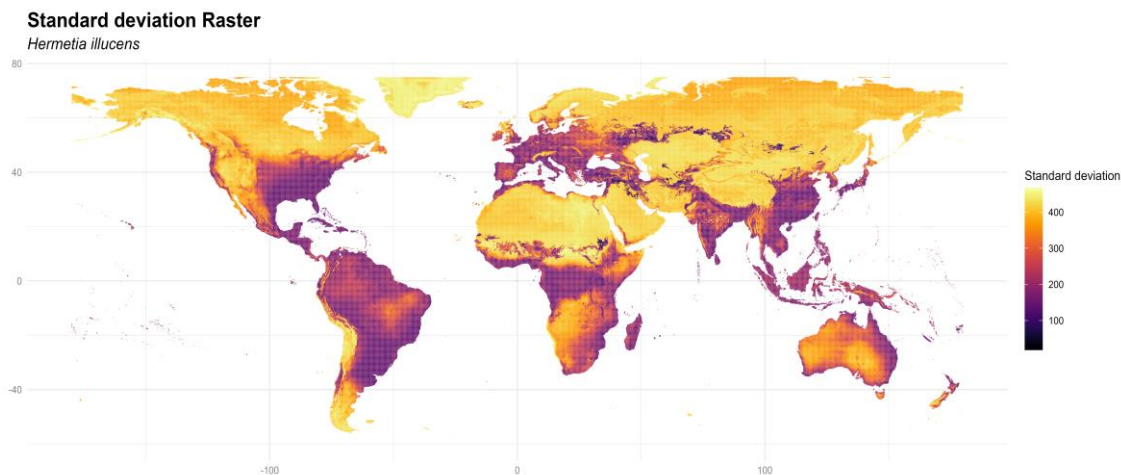
Carte d'indice de favorabilité



Indice de favorabilité final

- Le potentiel d'établissement de l'espèce est très fort aux endroits d'occurrences de l'espèce et apparaît globalement moyen (autour de 500) à travers le monde, ce qui suggère que les modèles peinent à fournir une évaluation claire et cohérente. Cela met en évidence un manque de robustesse dans les prédictions, probablement lié à des étapes préalables insuffisamment précises ou adaptées.

Carte incertitude



Incertitude des modèles

Incertitude forte au niveau des régions où la prédiction de l'établissement semble moyen. Les algorithmes doivent produire des réponses différentes ce qui donne cet écart-type très importants.