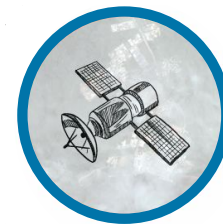


Intégration des métriques paysagères et de données entomologiques multi-sources pour la modélisation de la distribution des moustiques en Asie du Sud-Est

Claire Teillet, Sokeang Hoeun, Sébastien Boyer, Trang Thi Thuy Huynh,
Sébastien Marcombe, Diep Nguyen Tran Bich, Dung Nguyen Van, Vincent Herbreteau

claire.teillet@ird.fr

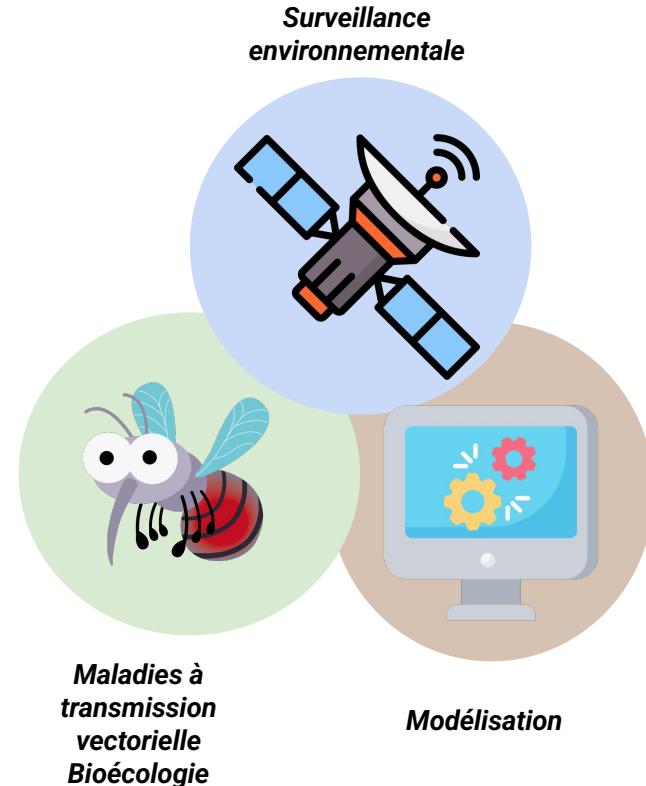
Study supported by ECOMORE Project (AFD), and SEA-ROADS Project (ANRS)



Introduction

Les maladies à transmission vectorielle constituent un enjeu majeur de santé publique en Asie du Sud-Est

- **Climat et environnement** : facteurs influençant la survie, le développement et la distribution des moustiques à différentes échelles
- **Manque de connaissances** :
 - **Distribution des espèces** de moustiques et **relations avec l'environnement** à l'échelle régionale en Asie du Sud-Est
 - **Variables paysagères peu étudiées** dans les modèles de distribution d'espèces (SDM) à l'échelle **régionale**, avec une **résolution spatiale adaptée à la biologie des vecteurs**



Introduction

Les maladies à transmission vectorielle constituent un enjeu majeur de santé publique en Asie du Sud-Est :

- Climat et facteurs bioclimatiques influençant la survie, la reproduction et la distribution des vecteurs (moustiques)
- Manque de données et de modèles pour la surveillance et la gestion des maladies à transmission vectorielle
 - Rôle des variables paysagères dans la modélisation de la répartition des espèces (à l'échelle régionale, avec une résolution spatiale adaptée à la biologie des vecteurs)

Comment les facteurs bioclimatiques et paysagers influencent-ils les niches écologiques d'*Ae. aegypti* et d'*Ae. albopictus* à l'échelle régionale (Asie du Sud-Est continentale) ?

Surveillance

Maladies à
transmission
vectorielle
Bioécologie

Machine et
Deep learning

Données entomologiques

Occurrences
1965-2025

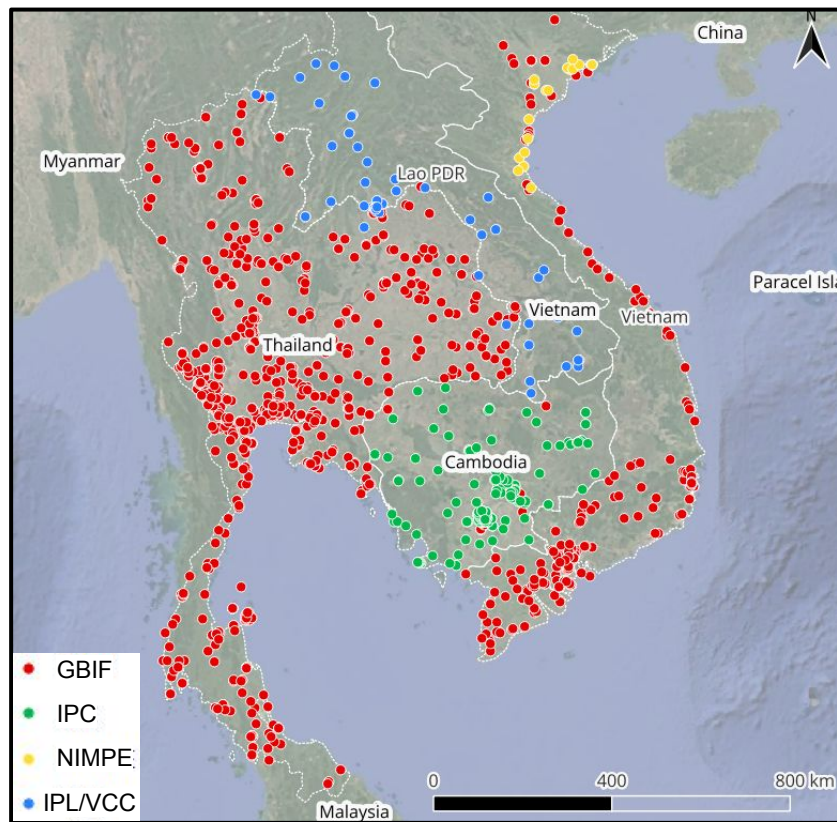


Plateforme d'accès
libre aux données de
biodiversité

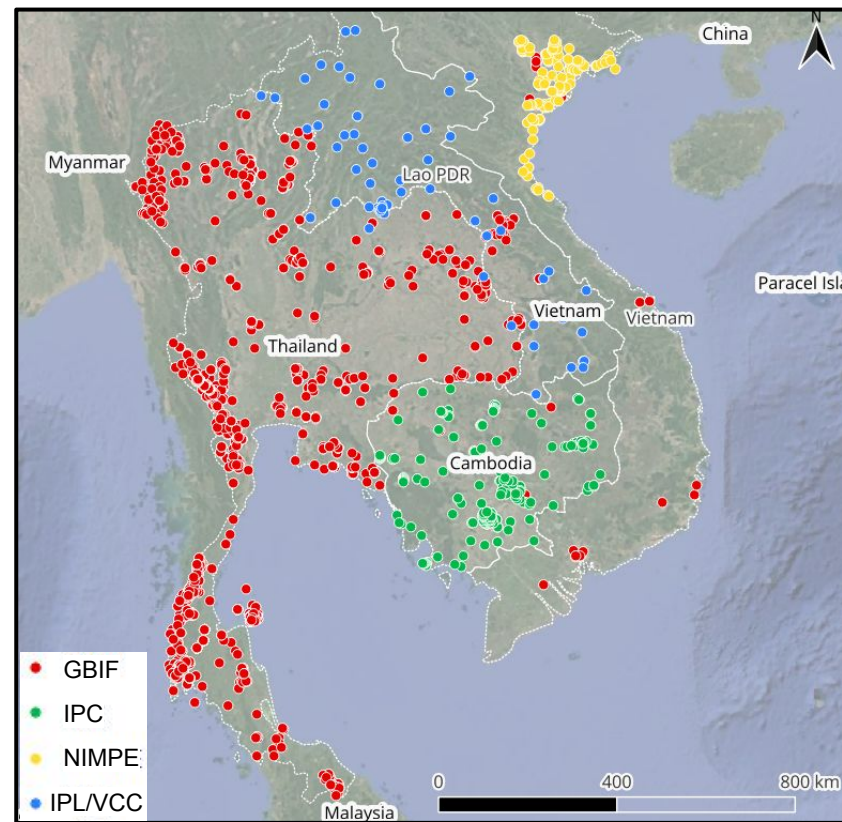


partenaires
locaux

Aedes aegypti [1230]

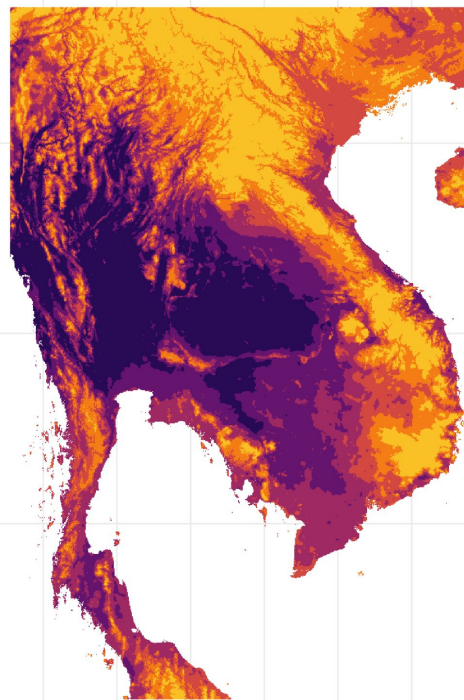


Aedes albopictus [1296]

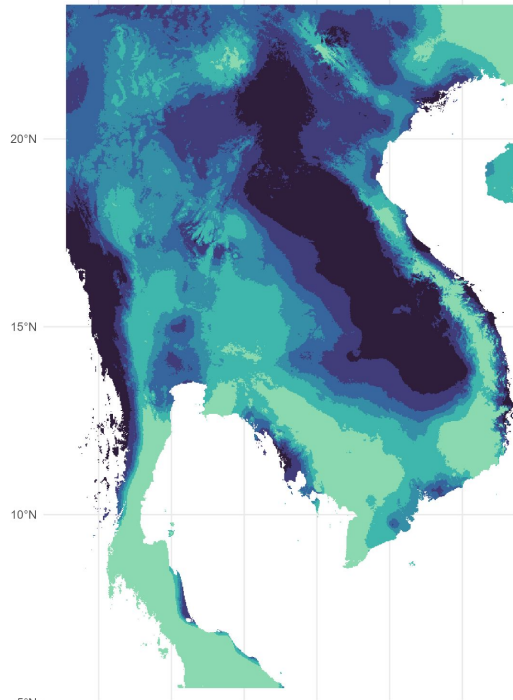


Données bioclimatiques

20 variables bioclimatiques à **1km de résolution spatiale**
WorldClim (1970-2000)



bio_5
(quantiles)
39
35
34
33
32
30
18



bio_15
(quantiles)
120
95
88
85
82
76
29

Température

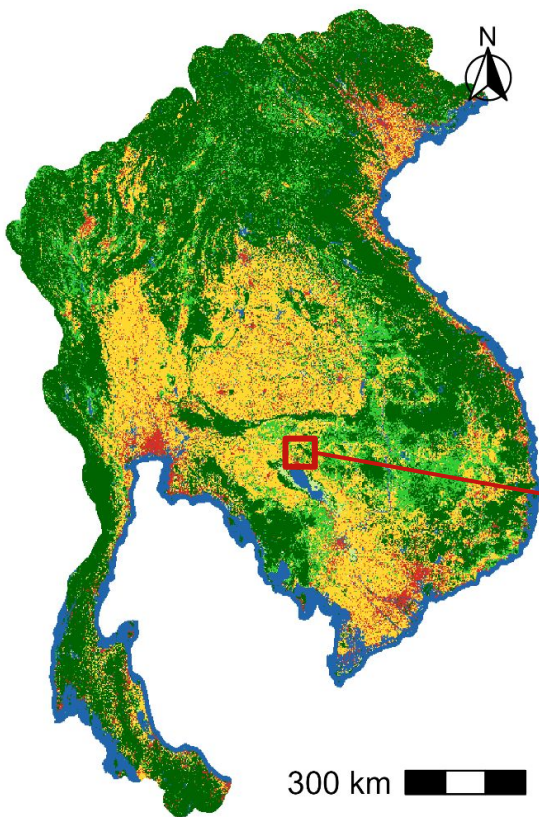
- BIO1 = Annual Mean Temperature
- BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
- BIO3 = Isothermality ($BIO2/BIO7$) ($\times 100$)
- BIO4 = Temperature Seasonality (standard deviation $\times 100$)
- BIO5 = Max Temperature of Warmest Month
- BIO6 = Min Temperature of Coldest Month
- BIO7 = Temperature Annual Range ($BIO5 - BIO6$)
- BIO8 = Mean Temperature of Wettest Quarter
- BIO9 = Mean Temperature of Driest Quarter
- BIO10 = Mean Temperature of Warmest Quarter
- BIO11 = Mean Temperature of Coldest Quarter

- BIO12 = Annual Precipitation
- BIO13 = Precipitation of Wettest Month
- BIO14 = Precipitation of Driest Month
- BIO15 = Precipitation Seasonality (Coefficient of Variation)
- BIO16 = Precipitation of Wettest Quarter
- BIO17 = Precipitation of Driest Quarter
- BIO18 = Precipitation of Warmest Quarter
- BIO19 = Precipitation of Coldest Quarter

Précipitations

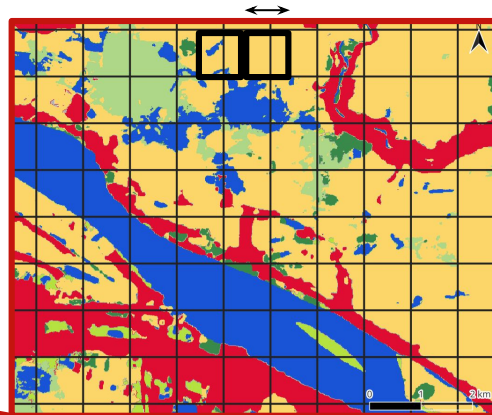
Métriques paysagères

[ESRI Land Cover Classification](#)
(2025), à **10 m de résolution spatiale**
(Sentinel-2 images)

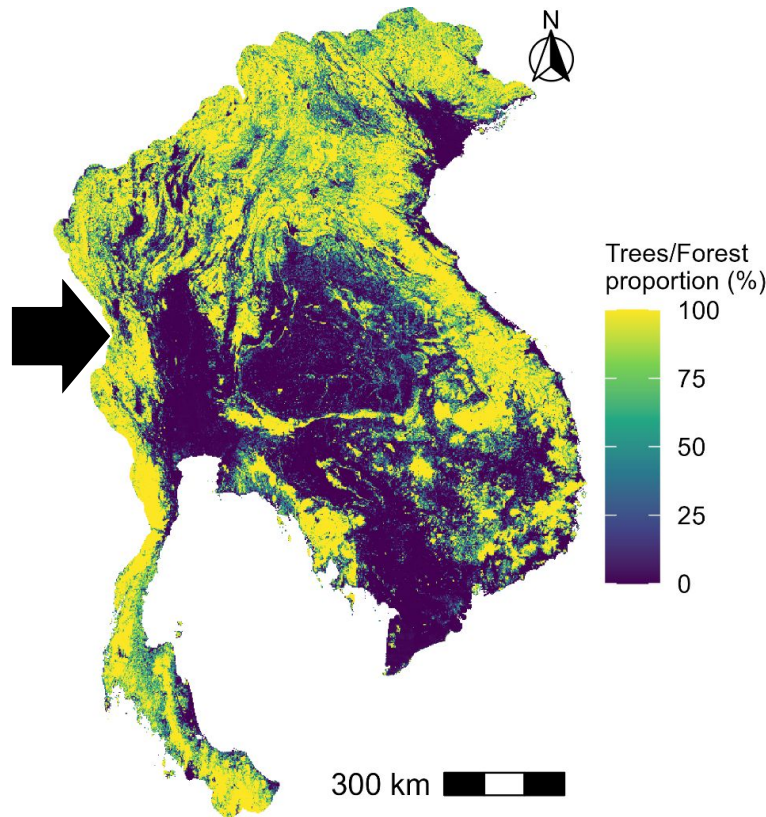


Métriques paysagères
calculées pour chaque classe
Bosch M. 2019

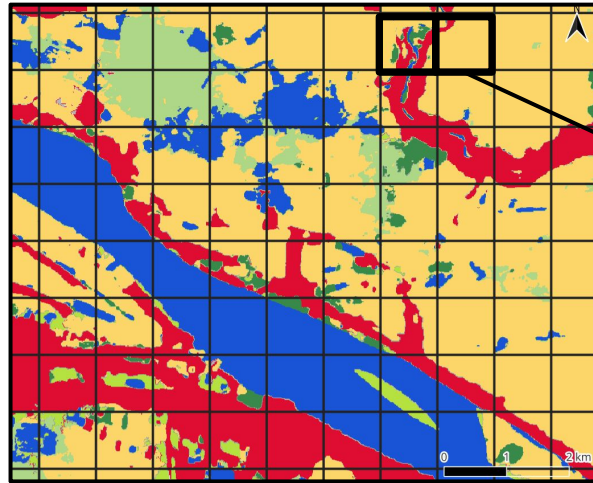
Cellule de la grille à **1 km**
(variables bioclimatiques)



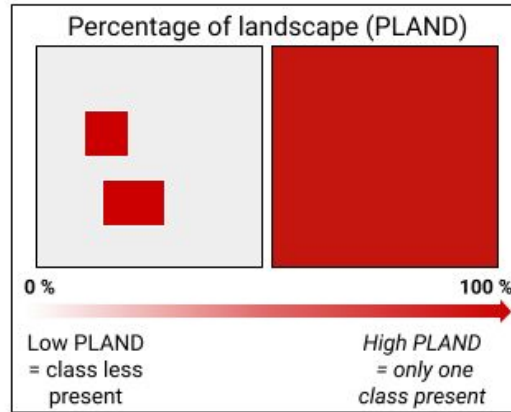
Exemple d'une métrique paysagère
proportion de forêt (1 km)



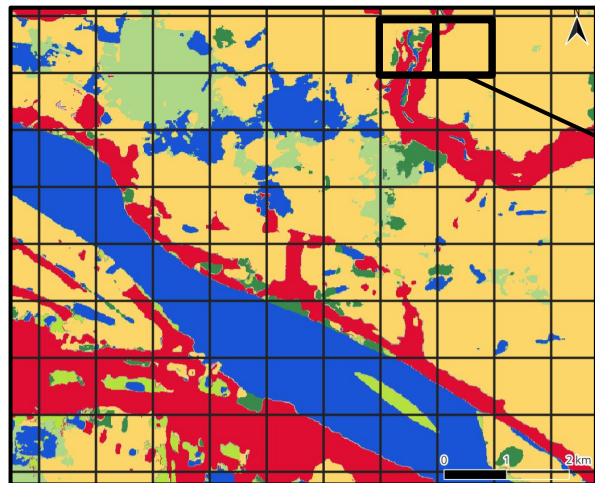
Métriques paysagères



COMPOSITION



Métriques paysagères



COMPOSITION

Percentage of landscape (PLAND)



0 %

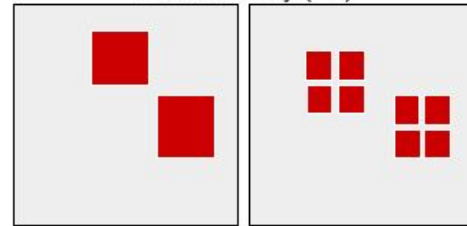
100 %

Low PLAND
= class less
present

High PLAND
= only one
class present

FRAGMENTATION

Patch density (PD)



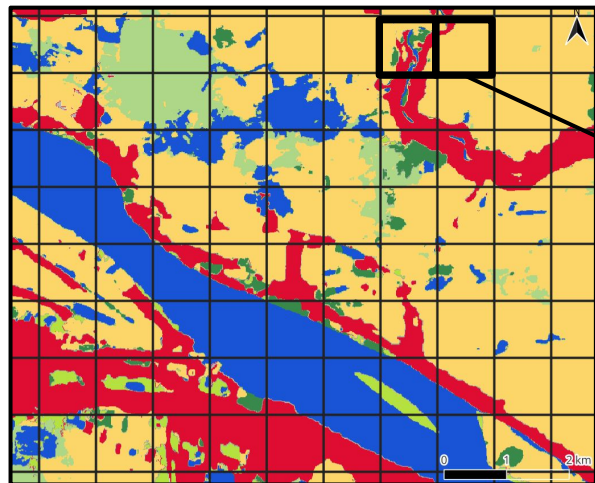
0

% / ha

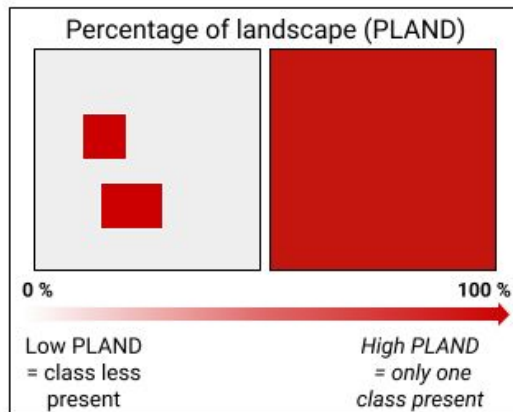
Low PD =
less
fragmented

High PD
= more
fragmented

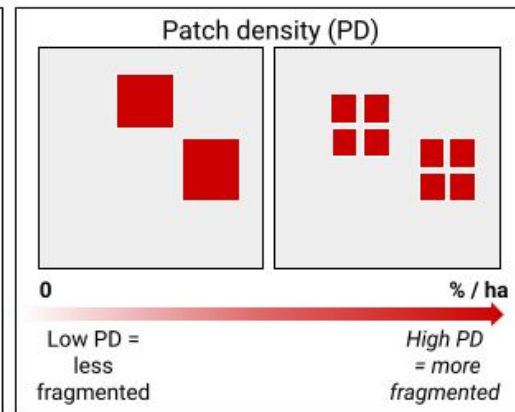
Métriques paysagères



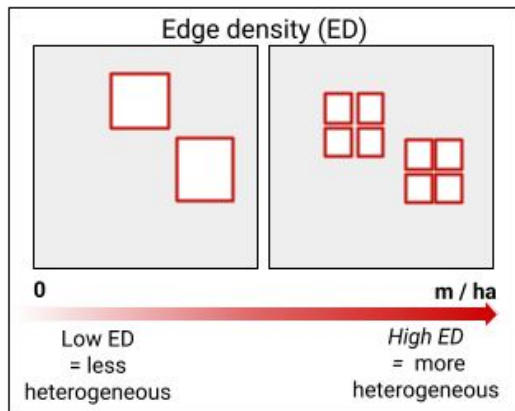
COMPOSITION



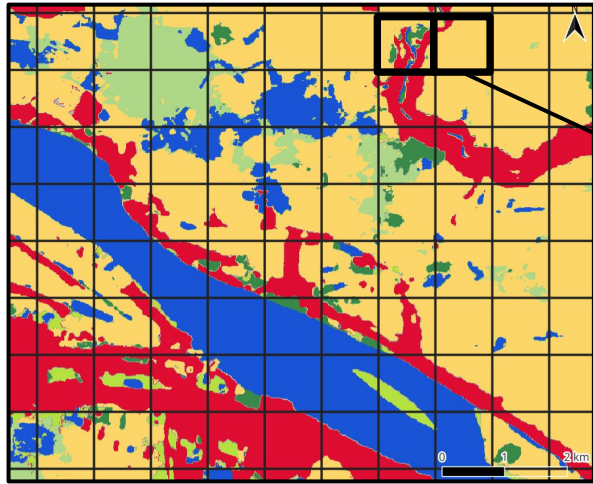
FRAGMENTATION



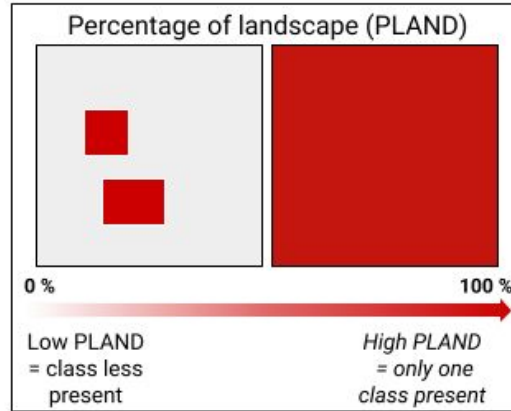
HÉTÉROGÉNÉITÉ



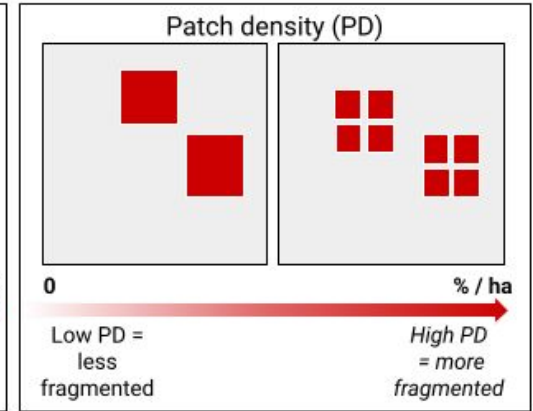
Métriques paysagères



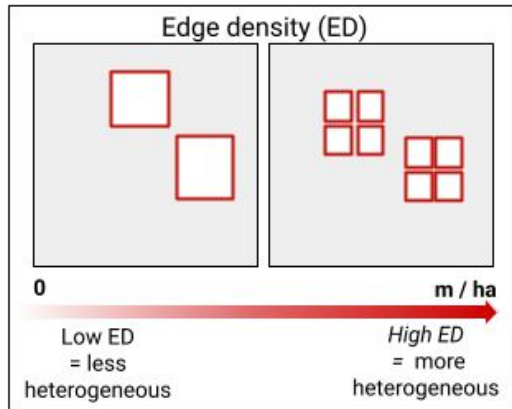
COMPOSITION



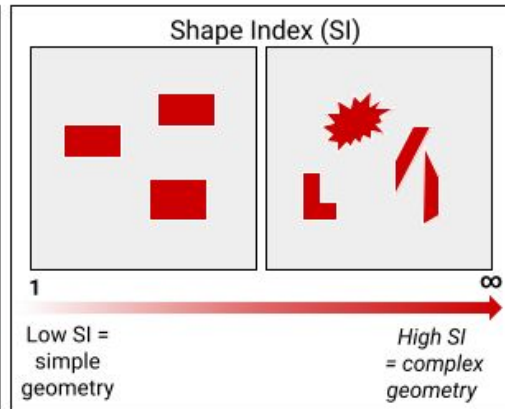
FRAGMENTATION



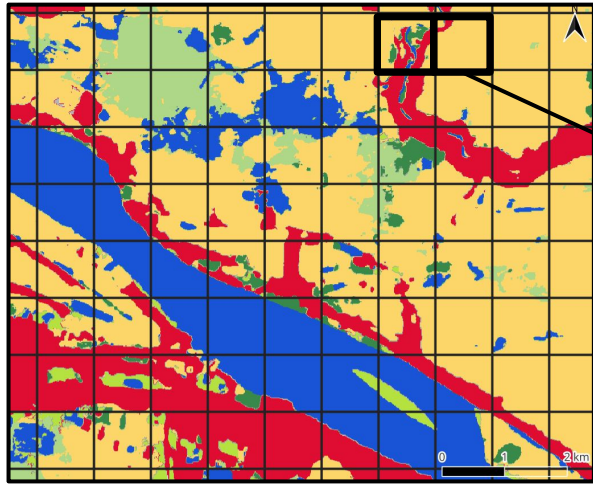
HÉTÉROGÉNÉITÉ



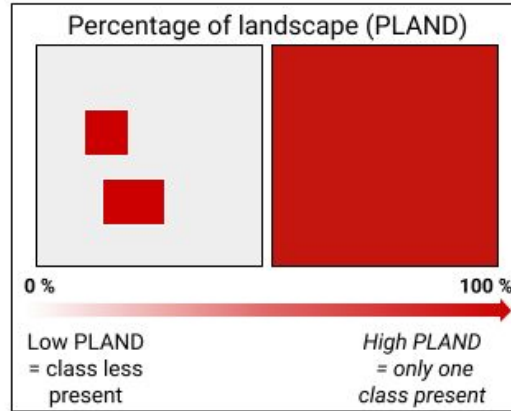
FORME



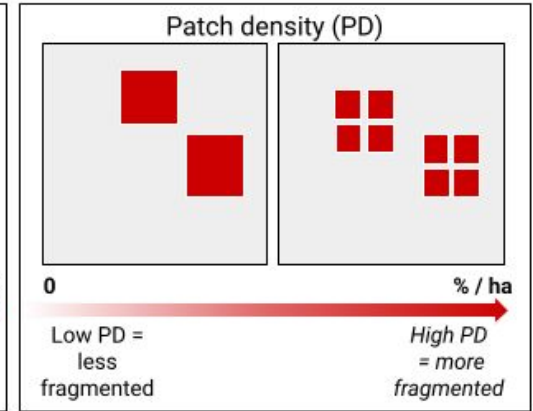
Métriques paysagères



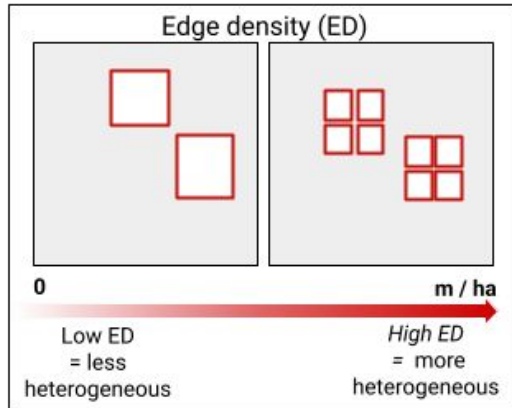
COMPOSITION



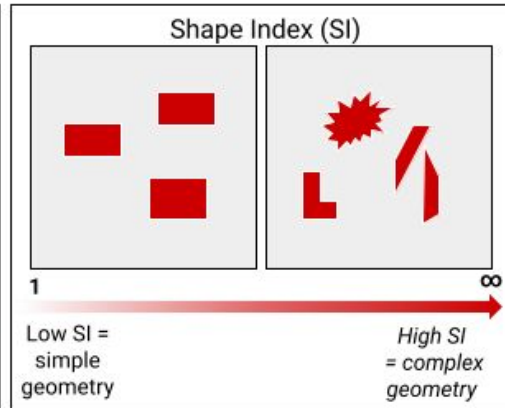
FRAGMENTATION



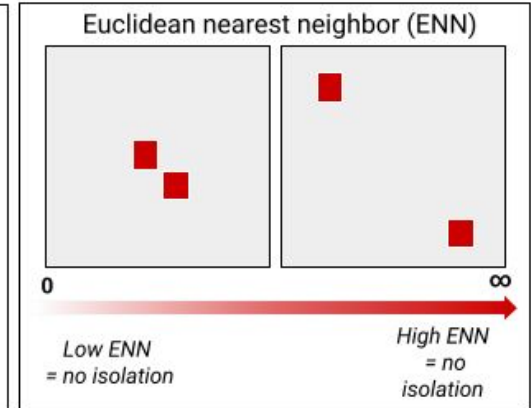
HÉTÉROGÉNÉITÉ



FORME



ISOLEMENT



Optimisation du code pour calculer à l'échelle régionale et à 1 km de résolution spatiale

package Pylandstats, Bosh M. 2019

Création d'un lanceur en ligne applicable sur n'importe quelle unité spatiale et classification

<<

Data Source & Export

Work & Export Directory

Landscape Raster (.tif)

Spatial Analysis Unit (.gpkg, .shp)

Target Classes (e.g., 1, 2, 3)

1, 2, 3, 4, 5, 6, 7

Chunk Size (Cells per batch)

5000

RUN TEST (500 ROWS)

RUN FULL ANALYSIS (CHUNK MODE)



Landscape Analysis

- Use a projected CRS (meters), not geographic coordinates (degrees)

Documentation:

- PyLandStats (Bosch et al., 2019) : <https://pylandstats.readthedocs.io/en/latest/>
- Metrics reference (FRAGSTATS-inspired): <https://www.fragstats.org/index.php/documentation>

Metrics Selection

Area & Edge

- ☐ Total Area
- ☒ Proportion of Landscape
- ☐ Number of Patches
- ☒ Patch Density
- ☐ Largest Patch Index
- ☐ Total Edge
- ☐ Edge Density
- ☐ Mean Patch Area
- ☐ Mean Perimeter

Shape

- ☒ Mean Shape Index
- ☐ Mean Perimeter-Area Ratio
- ☐ Mean Fractal Dimension

Core

- ☐ Total Core Area
- ☐ Core Area Prop. of Landscape
- ☐ Number of Disjunct Core Areas
- ☐ Disjunct Core Area Density
- ☐ Mean Core Area

Aggregation

- ☐ Landscape Shape Index
- ☐ Effective Mesh Size
- ☐ Mean Euclidean Nearest Neighbor

Landscape-level metrics

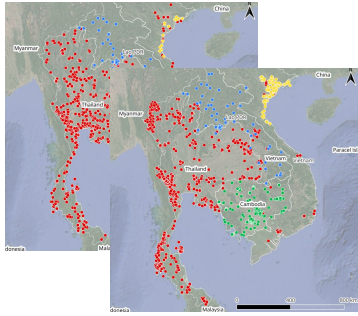
- ☒ Shannon Diversity Index
- ☐ Entropy
- ☐ Contagion
- ☐ Joint Entropy
- ☐ Conditional Entropy
- ☐ Mutual Information
- ☐ Relative Mutual Information



Live Processing Console

Stage de
Matthew Martines

Méthodologie

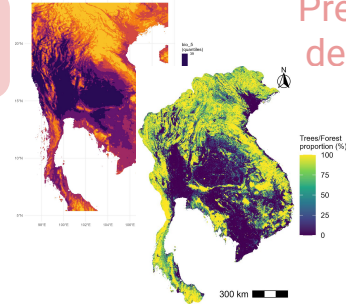


Données de présence
(*Aedes aegypti* or *Aedes albopictus*)

**62 variables climatiques
et paysagères**

+ *Pseudo-absences*

Sélection des variables
 $|r| < 0,8$; $VIF < 10$



Prétraitement
des données

Méthodologie

Prétraitement
des données

Données de présence
(*Aedes aegypti* or *Aedes albopictus*)

62 variables climatiques
et paysagères

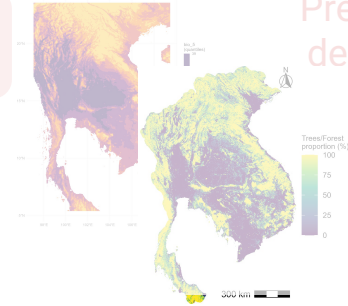
Sélection des variables
 $|r| < 0,8$; $VIF < 10$

Variables
bioclimatiques

Variables bioclimatiques
et paysagères

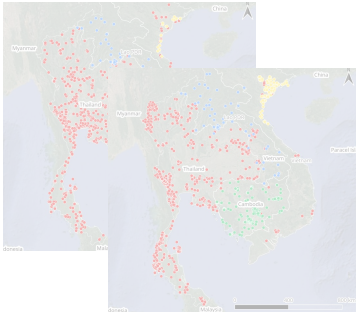
Entraînement : GLM, GBM, MAXNET, RF...

Évaluation : validation croisée spatiale
(AUCroc, TSS)



Modèle de
distribution
d'espèces
(SDM)

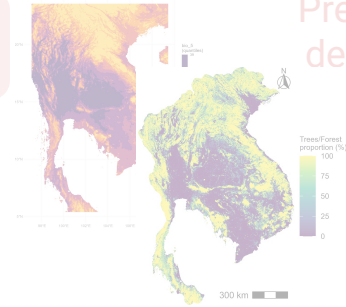
Méthodologie



Données de présence
(*Aedes aegypti* or *Aedes albopictus*)

62 variables climatiques
et paysagères

Sélection des variables
 $r^2 < 0,8$; VIF < 10



Prétraitement
des données

Variables
bioclimatiques

Variables bioclimatiques
et paysagères

Entraînement : GLM, GBM, MAXNET, RF...

Évaluation : validation croisée spatiale
(AUCroc, TSS)

Sélection : AUC > 0.70

Modèle
d'ensemble

Modèle
d'ensemble

Prédiction de la
probabilité de présence

Importance des
variables

Courbes de réponse

Incertitude des
prédictions (EMcv)



Modèle de
distribution
d'espèces
(SDM)

Performance

Ae. aegypti

AUCroc = 0.89

Ae. albopictus

AUCroc = 0.88

Performance

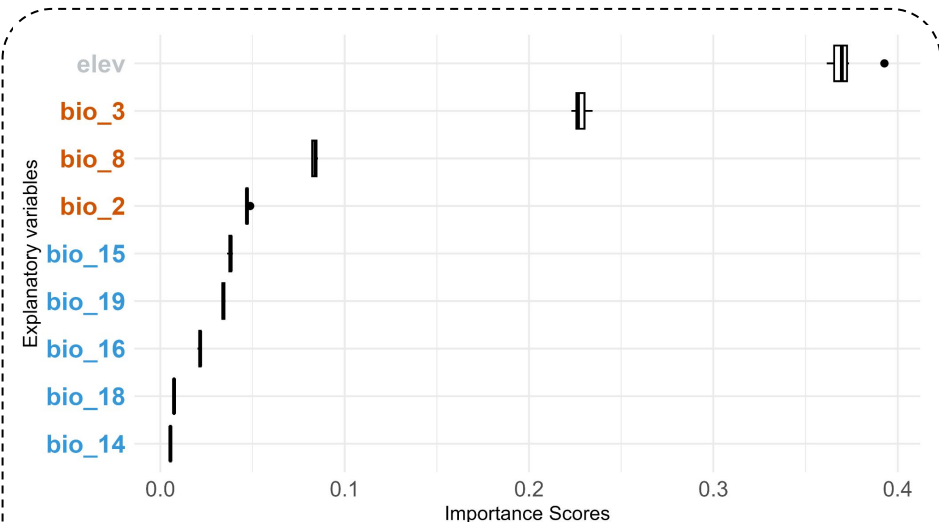
Ae. aegypti

AUCroc = 0.89

Ae. albopictus

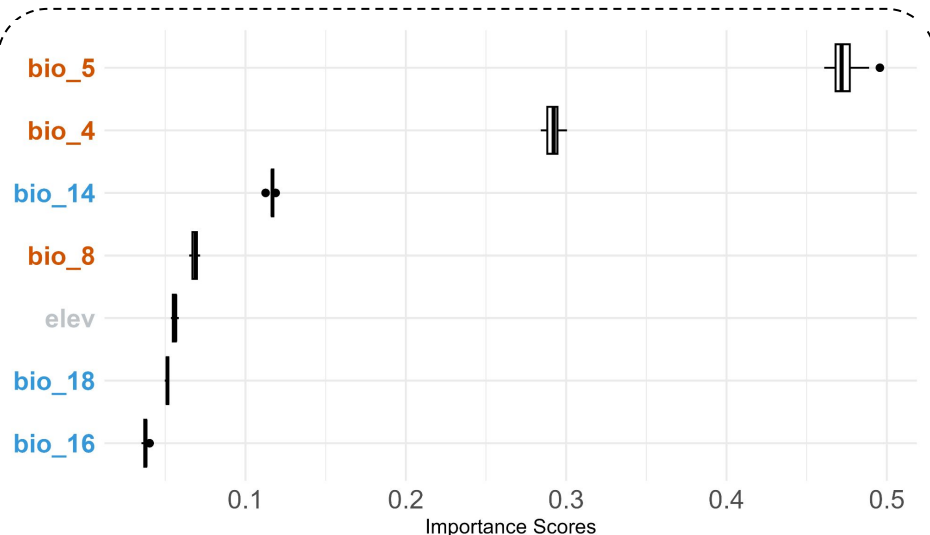
AUCroc = 0.88

Importance des variables : quelle est la contribution de chaque variable dans le modèle de prédiction ?



elev = altitude

bio 3 = isothermality (rapport entre l'amplitude thermique journalière moyenne et l'amplitude annuelle)



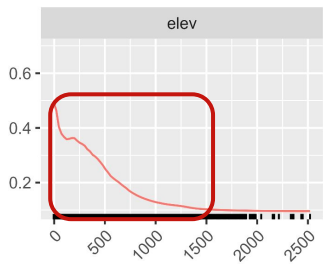
bio 5 = température maximale du mois le plus chaud

bio 4 = variation saisonnière des températures

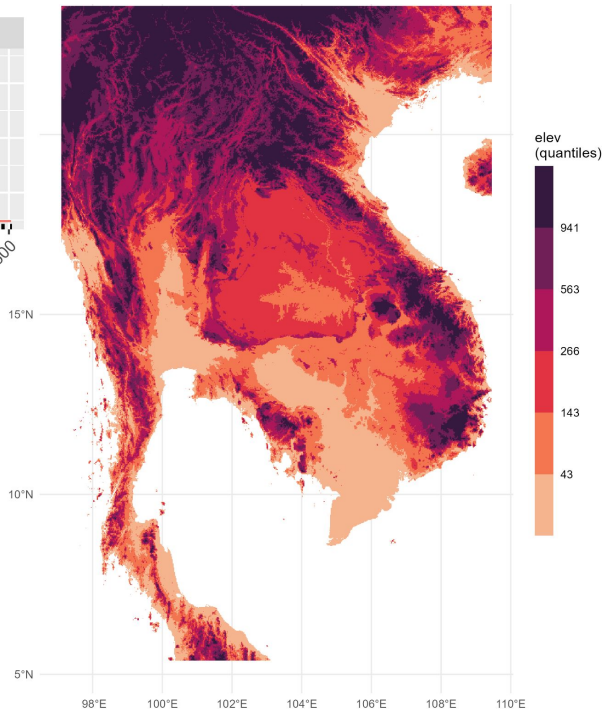
Courbes de réponse : Comment la probabilité de présence varie-t-elle en fonction de chaque variable ?

Ae. aegypti

Elevation

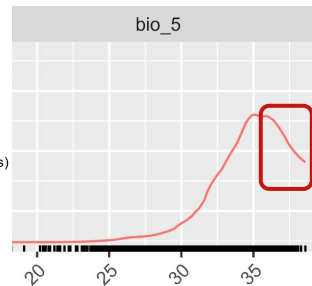


→ Diminution
de la probabilité
de présence en
fonction de
l'altitude

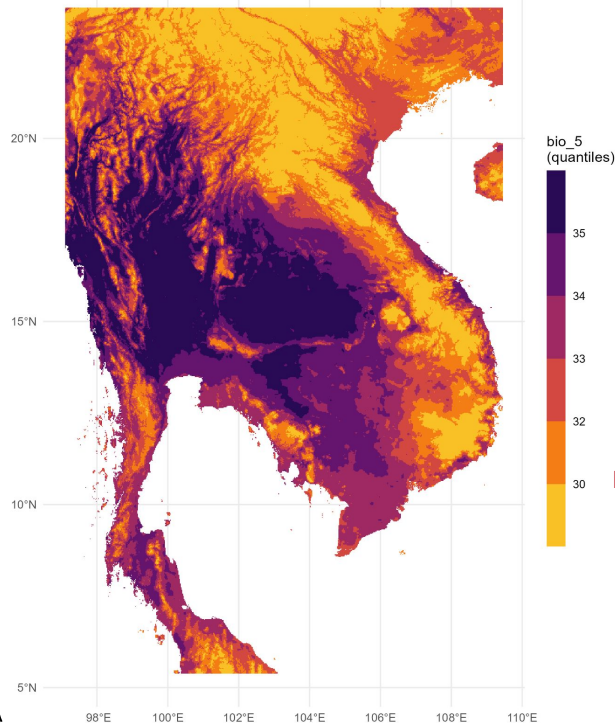


Ae. albopictus

BIO5 = Max Temperature of Warmest Month



→ probabilité de
présence augmente
avec la température
maximale du mois le
plus chaud jusqu'à
environ 35 °C, puis
diminue



Performance

Ae. aegypti

AUCroc = 0.93 (0.89)

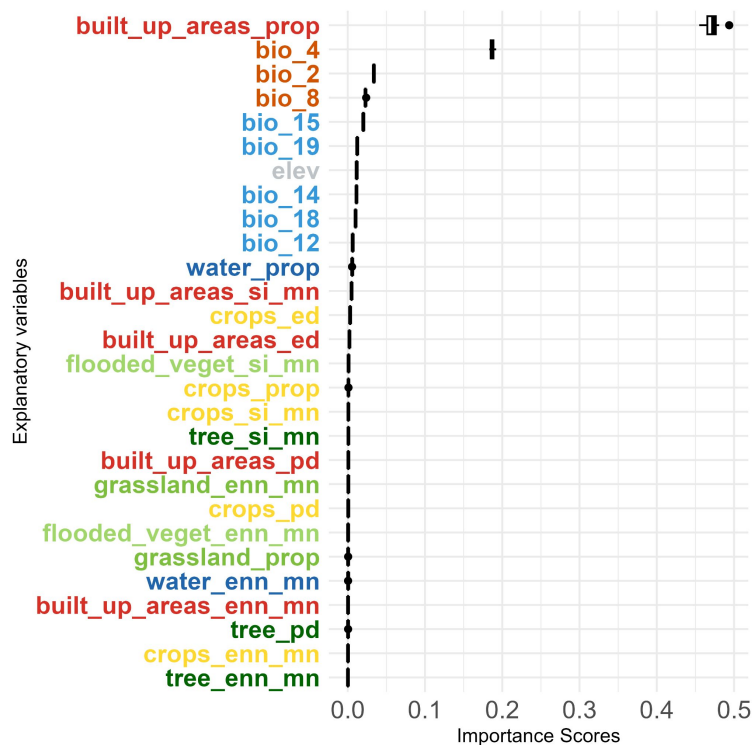
Ae. albopictus

AUCroc = 0.89 (0.88)

Importance des variables

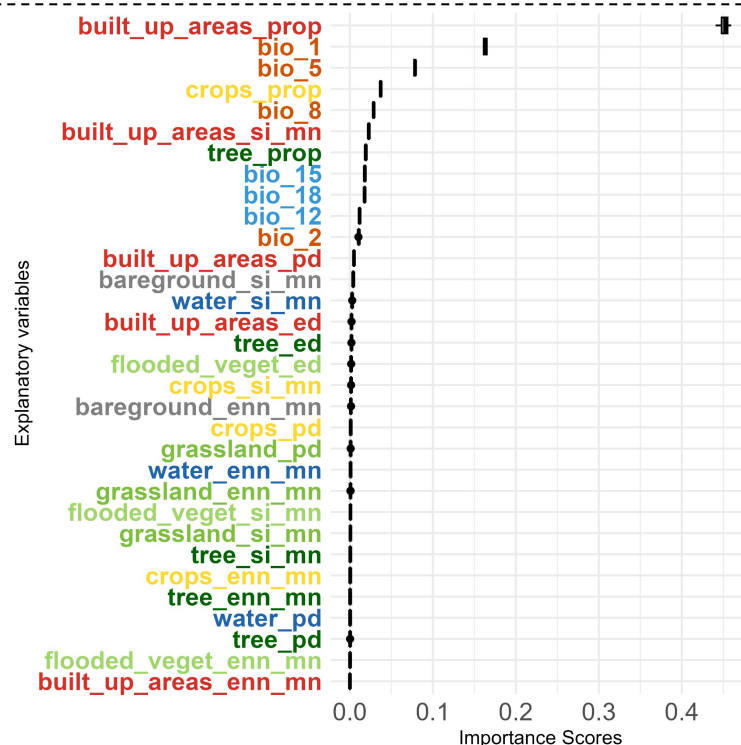
Ae. aegypti

AUCroc = 0.93 (0.89)



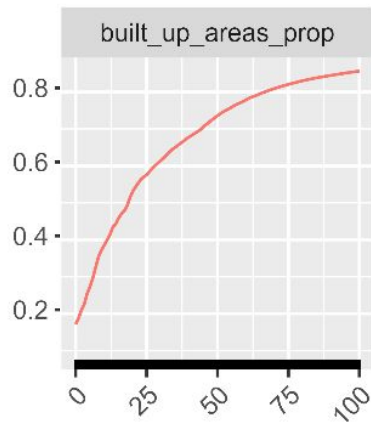
Ae. albopictus

AUCroc = 0.89 (0.88)



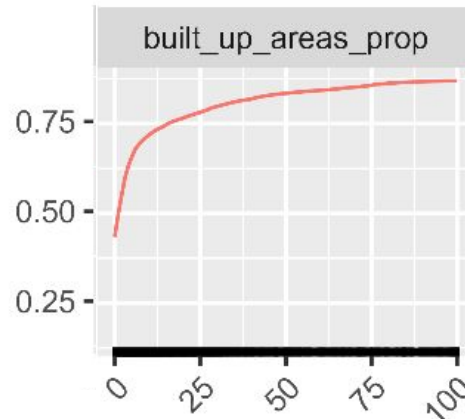
Courbes de réponse

Ae. aegypti



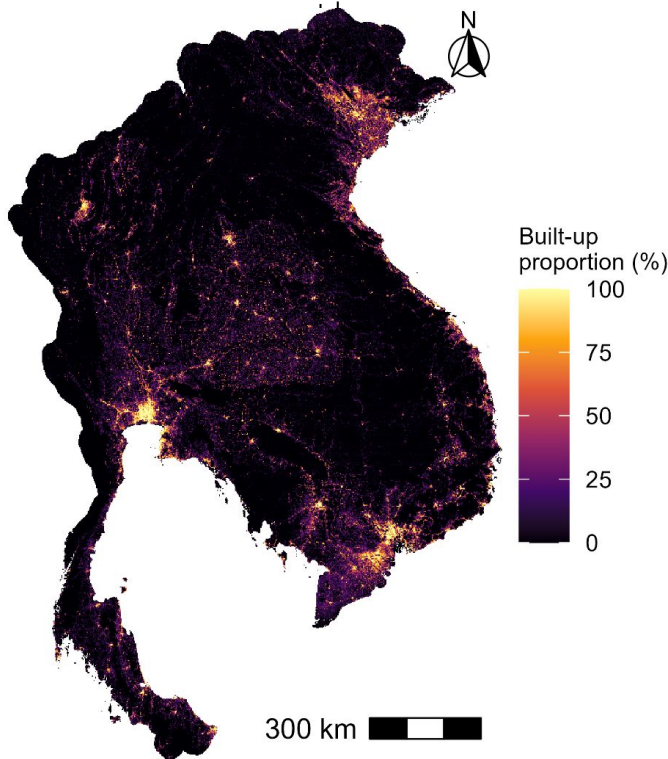
→ augmentation progressive avec la proportion de zones urbanisées

Ae. albopictus



→ augmentation rapide dès les faibles proportions de zones urbanisées, puis plateau

Proportion des zones urbanisées



Probabilité de présence

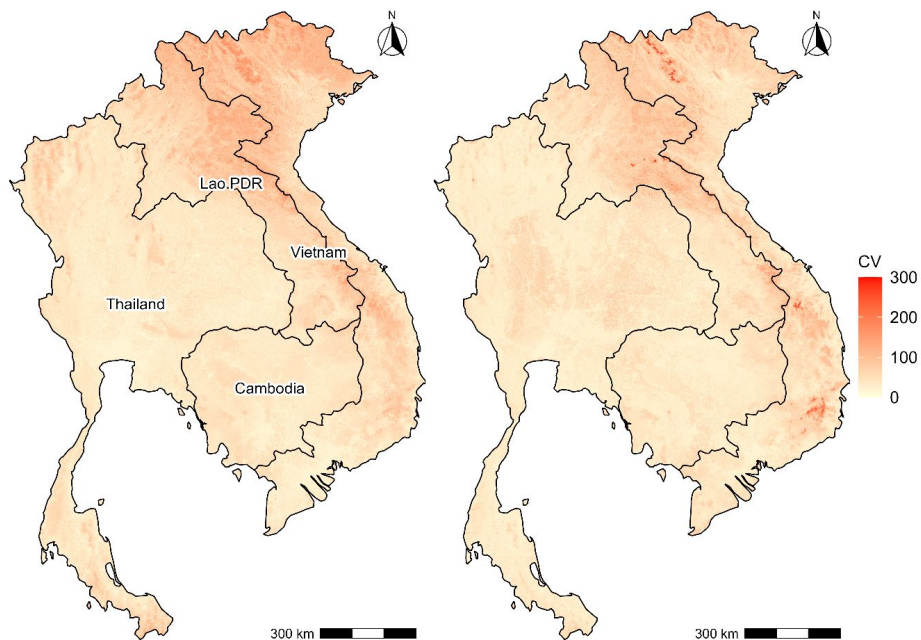
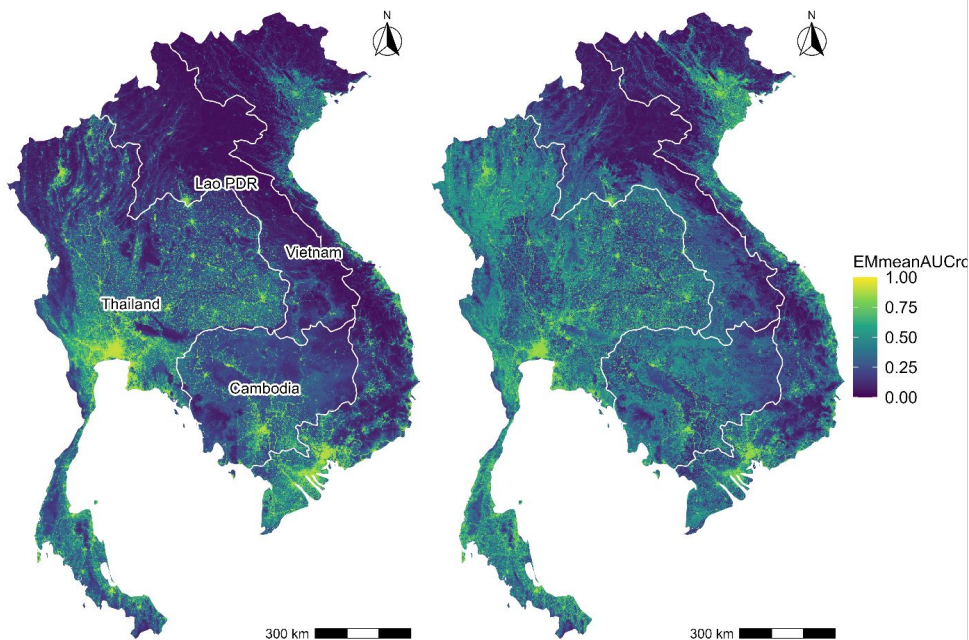
Incertitude des prédictions

Ae. aegypti

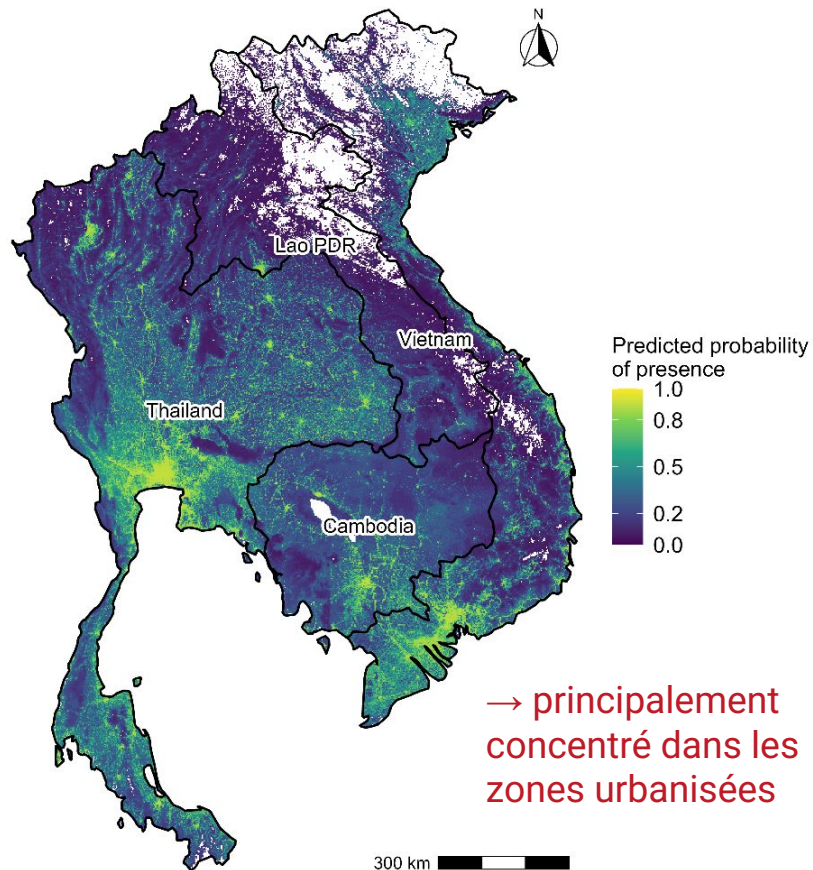
Ae. albopictus

Ae. aegypti

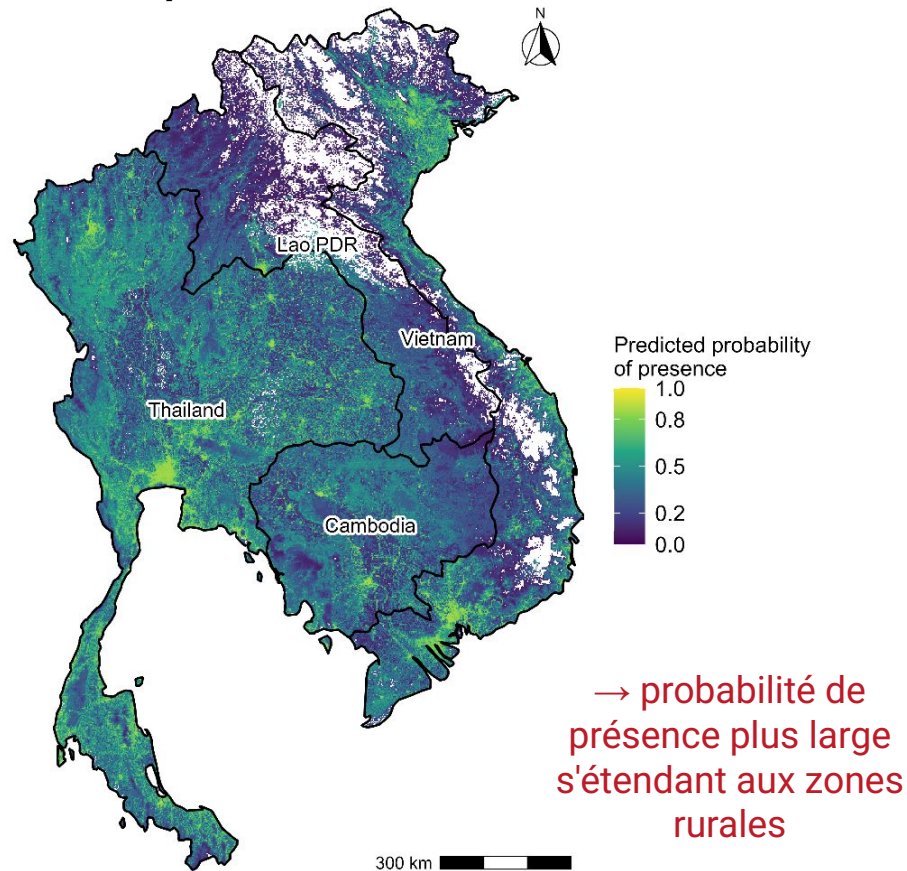
Ae. albopictus



(a) *Ae. aegypti*



(b) *Ae. albopictus*



Comment les facteurs bioclimatiques et paysagers influencent-ils les niches écologiques d'*Ae. aegypti* et d'*Ae. albopictus* à l'échelle régionale ?

Climat

Variables bioclimatiques délimite l'enveloppe climatique des deux espèces

Urbanisation

Forte influence de la proportion des zones urbanisées pour ces deux espèces avec deux réponses différentes

Configuration du paysage

Composition prime sur la configuration pour ces deux espèces
Gain de performance (*Ae. aegypti*)

Probabilité de

présence avec incertitude, sur la région, à 1 km de résolution spatiale

Conclusion / Perspectives



Modélisation

Cadre de modélisation complet et reproductible

- sélection des variables
- analyse des performances
- validation croisée spatiale
- incertitude des modèles
- biais d'échantillonnage environnemental



Données environnementales

Indicateurs paysagers avec une résolution de 1 km accessibles pour l'Asie du Sud-Est



Limites : Variables bioclimatiques (1970-2000), Classification (2025), Occurrences (1965-2025)



Données entomologiques

Application à d'autres espèces de moustiques :

Culex + *Armigeres* + *Anopheles*

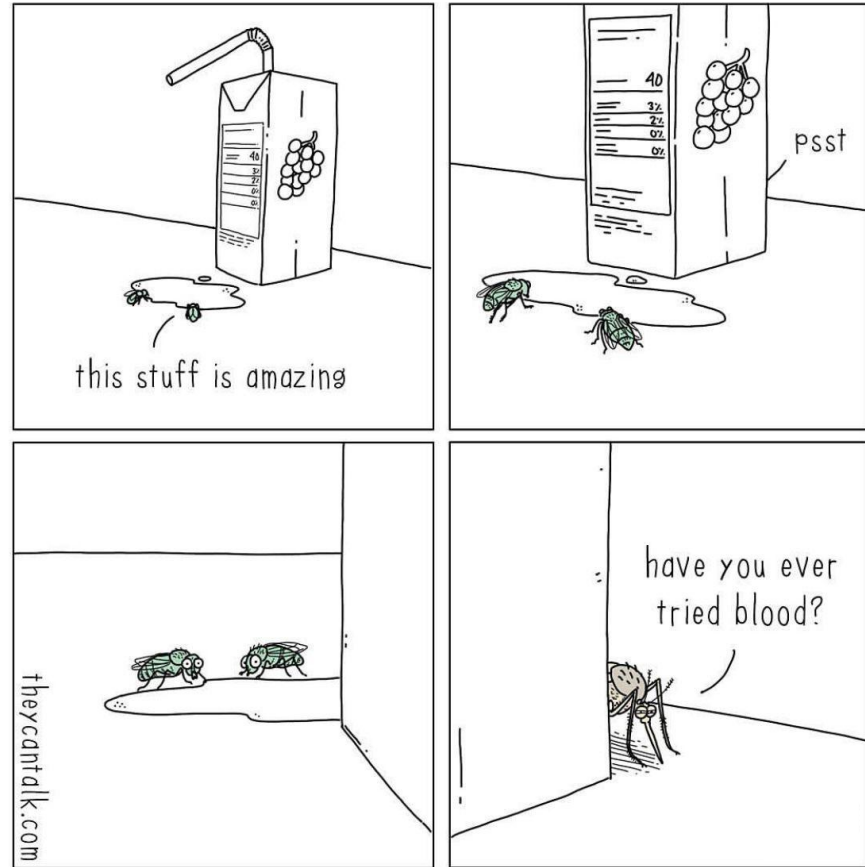
1. ~~*Aedes aegypti*~~
2. ~~*Aedes albopictus*~~
3. *Culex quinquefasciatus*
4. *Culex brevipalpis*
5. *Aedes vittatus*
6. *Armigeres subalbatus*
7. *Armigeres kesseli*
8. *Culex vishnui*
9. *Culex nigropunctatus*
10. *Culex gelidus*
11. *Culex fuscocephala*



Une seule santé

Interdisciplinarité + Interopérabilité
comprendre les interactions complexes
au sein des écosystèmes à différentes
échelles imbriquées

**Merci de votre
attention !**

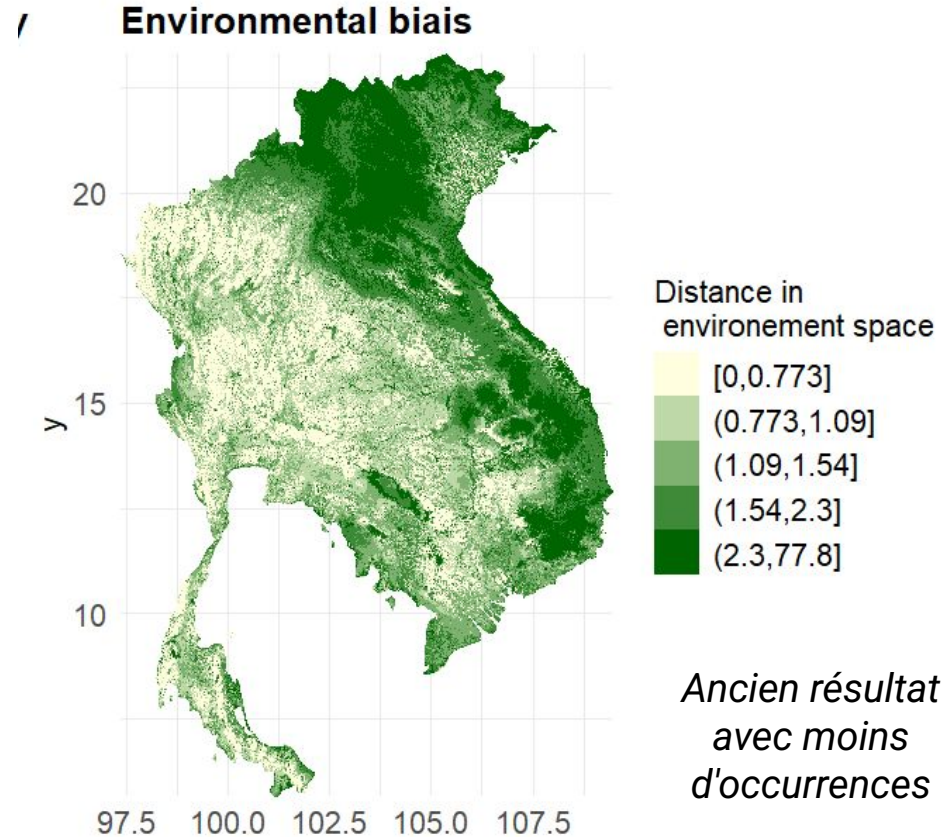


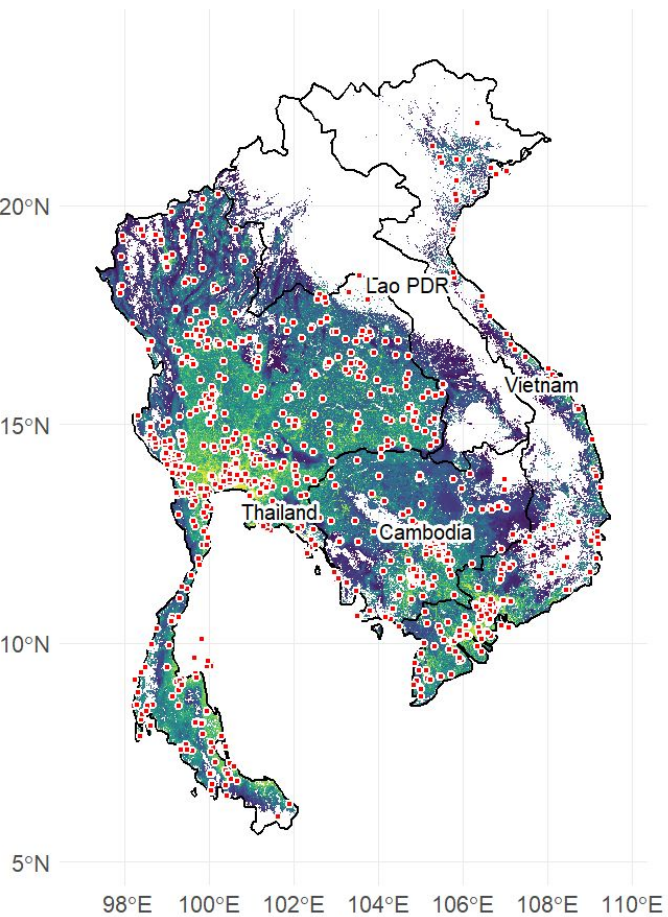
Evaluation du biais d'échantillonnage environnemental

Dans quelle mesure les occurrences utilisées ont-elles couvert l'ensemble des conditions environnementales de la zone d'étude ?

Valeurs les plus élevées = distances les plus importantes entre chaque cellule et la cellule échantillonnée la plus proche

Visualisation des conditions environnementales **qui n'ont pas été échantillonnées**

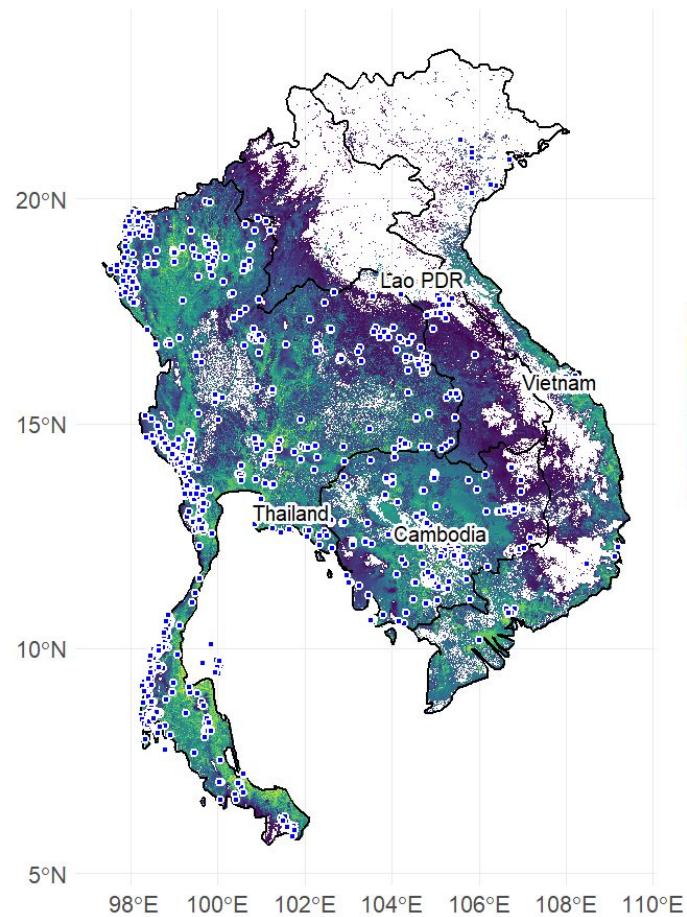




EMmeanAUCroc masked

0.8
0.6
0.4
0.2

• *Ae. aegypti* occurrences



EMmeanAUCroc masked

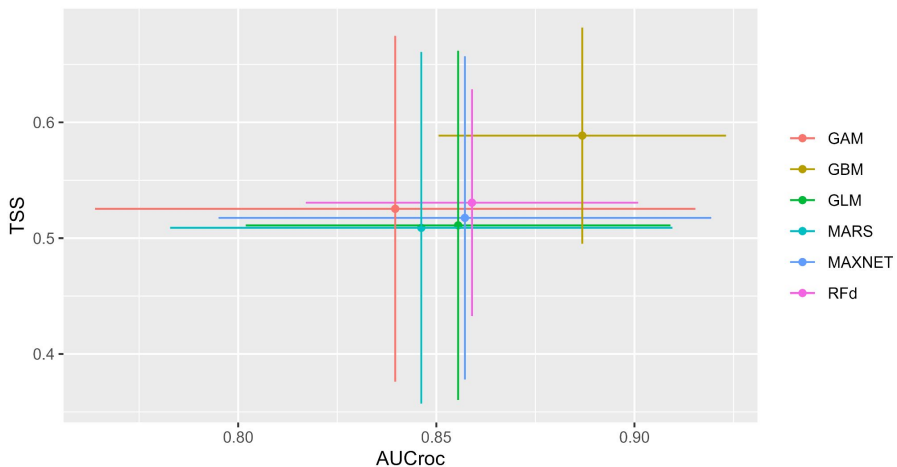
0.75
0.50
0.25

• *Ae. albopictus* occurrences

Single model for climatic and landscape variables

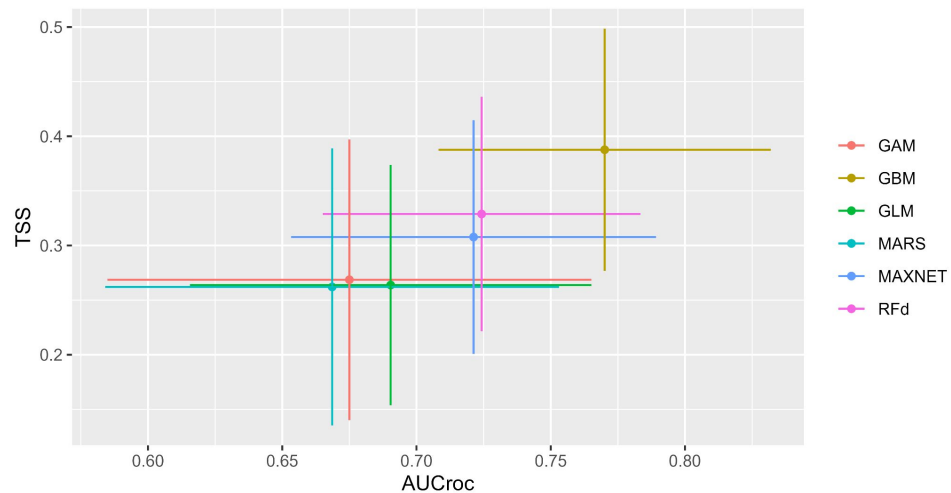
Ae. aegypti

Validation dataset

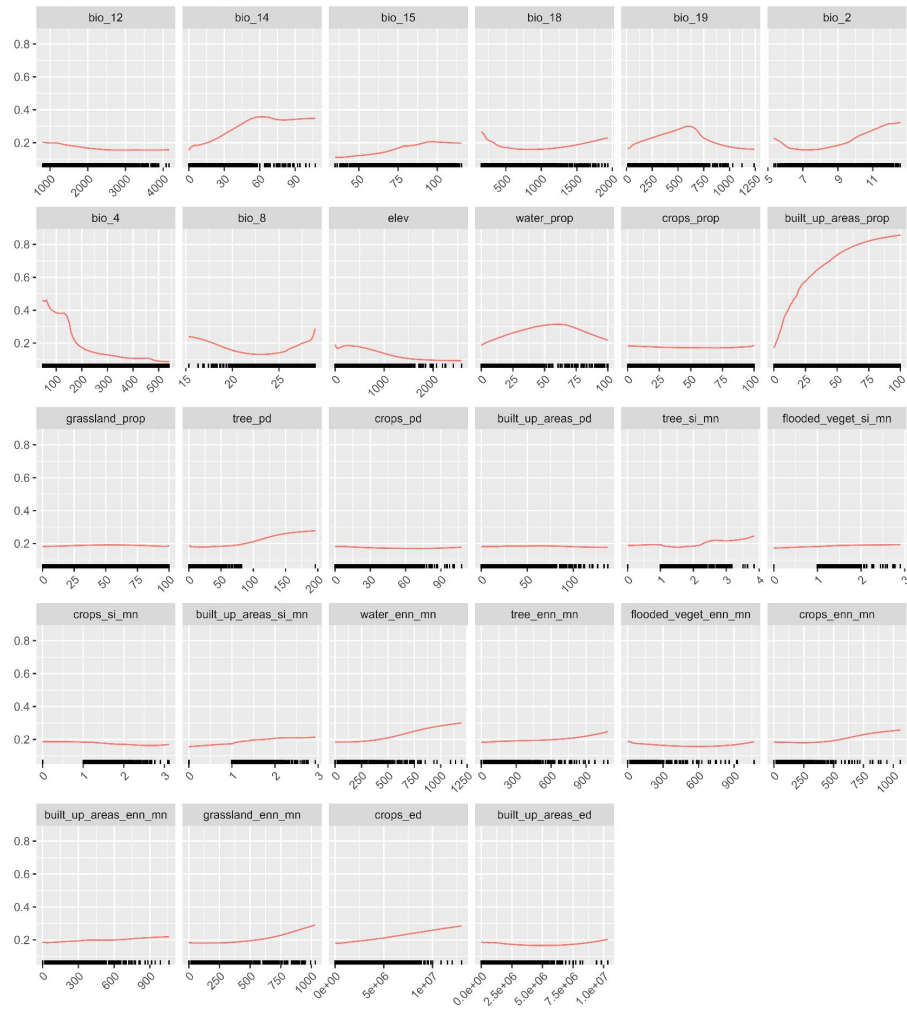


Ae. albopictus

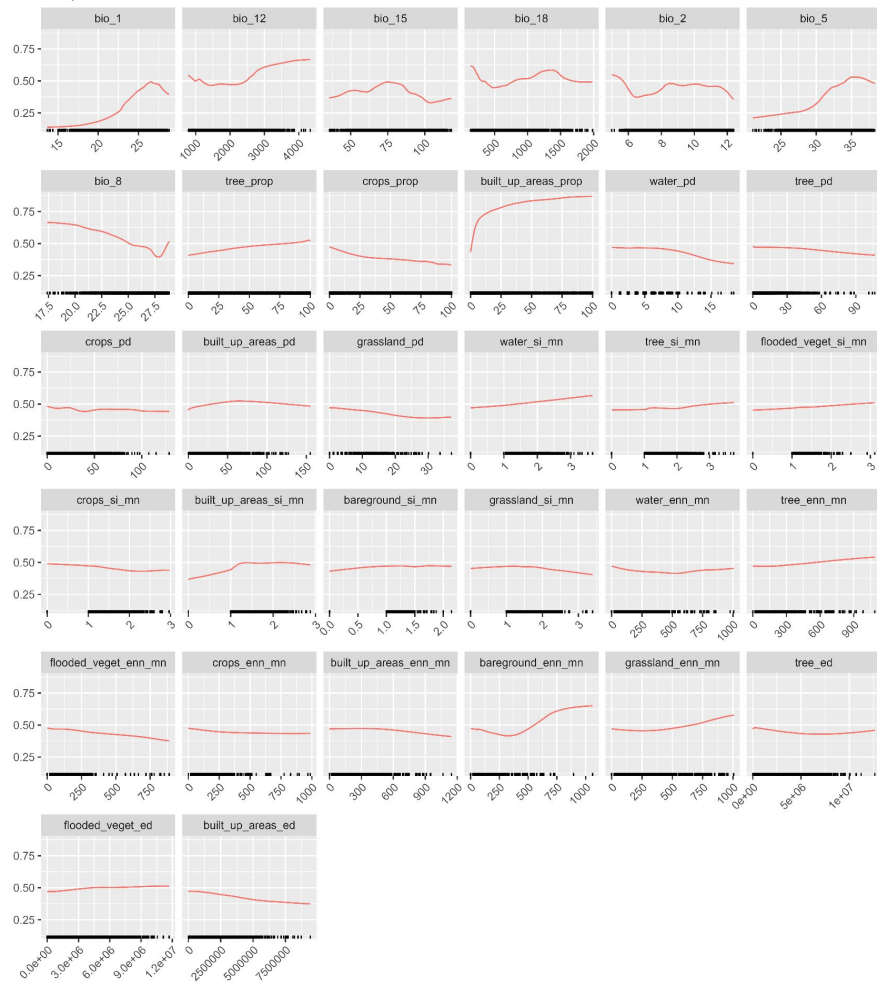
Validation dataset

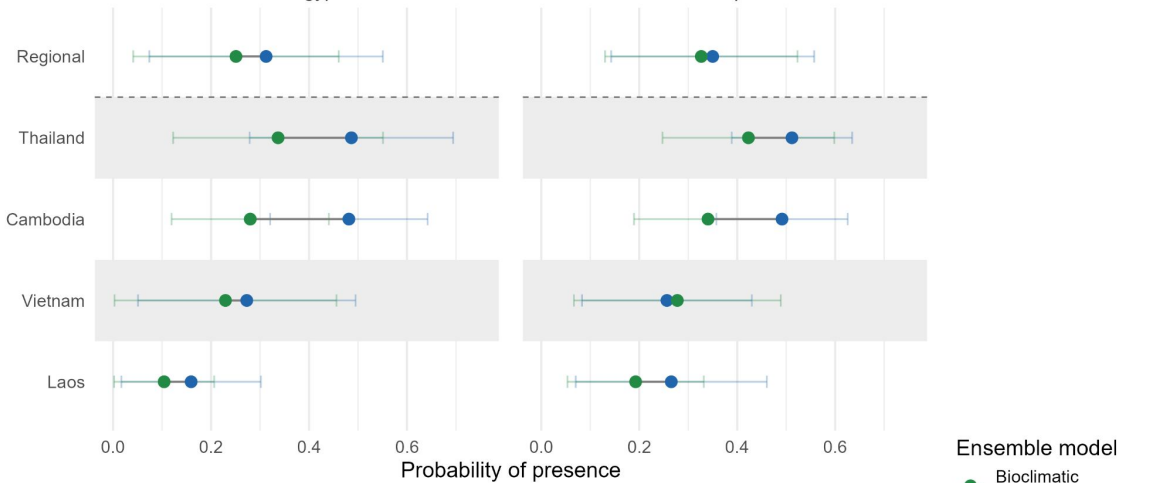


Response curves for aegypti's models



Response curves for albo's models



*Ae. aegypti**Ae. albopictus*

Ensemble model

- Bioclimatic + Landscape metrics
- Bioclimatic

*Ae. aegypti**Ae. albopictus*