

Occurrence des parasitoses et paysages urbains : les rongeurs porteurs de Trypanosomes dans la ville de Niamey (Niger)

J.-P. Rossi, I. Kadaouré, M. Garba, M. Godefroid, P. Gauthier, C. Tatard, G. Dobigny

27 septembre 2016

I-Cadre conceptuel

Relations entre l'hétérogénéité environnementale et les patrons de distribution des populations (espace et temps)

- 1 - Expliquer les patrons
- 2 - Inférer les processus sous jacents

2-Objectif : relier la structure du paysage urbain aux occurrences d'une maladie

L'épidémiologie cherche à identifier les facteurs environnementaux corrélés à la présence de la maladie.

Facteurs environnementaux en milieu urbain :

- Facteurs abiotiques/biotiques (écologiques) ex. humidité
- Structure de l'habitat urbain (facteurs sociaux-économiques, historiques, ...)

Ces facteurs sont liés et covariant (écologie urbaine).

3-Structure de l'habitat urbain

Le travail présenté aujourd'hui concerne une petite partie du problème et traite des **liens entre habitat urbain et occurrences d'une parasitose**.

Pour tester un lien parasitose-paysage il faut :

- 1 Traduire le paysage (vu comme un patchwork d'objets) en variables aléatoires quantitatives (VA)
- 2 Trouver un moyen efficace de rechercher et tester les **relations** entre ces VA et les **données épidémiologiques**

Ces questions sont communes à l'écologie du paysage et à l'agroécologie (par exemple).

4-Choix méthodologiques

4a Caractérisation du paysage urbain

On utilise les **métriques paysagères** pour quantifier la **composition** et la **structure** du paysage urbain.

Ces indices mathématiques sont nombreux et diversifiés (Gustafson 1998).

Il faut faire un choix car les VA sont fortement corrélées et souvent colinéaires

→ problèmes possibles lors des analyses statistiques ou la modélisation.

Métriques retenues dans cette étude préliminaire :

Métriques de classes

- PLAND (% landscape cover) mesure de composition
- ED (edge density) mesure de fragmentation

Métriques paysage

- Diversité (SIDI, SIEI), contagion, fragmentation, forme

Gustafson 1998. Ecosystems 1: 143–56.

4b Lien paysage - occurrences de la maladie

Il existe différentes stratégies pour lier un cortège de variables explicatives à une variable cible.

Il faut trouver une approche qui n'est pas trop sensible à la colinéarité et peut travailler sur un nombre d'observations relativement restreint.

On a testé l'approche régression PLS et différents algorithmes utilisés en modélisation de niche écologique.

Les résultats présentés aujourd'hui concernent l'algorithme **maxent** (maximum entropy: Elith et al 2011)

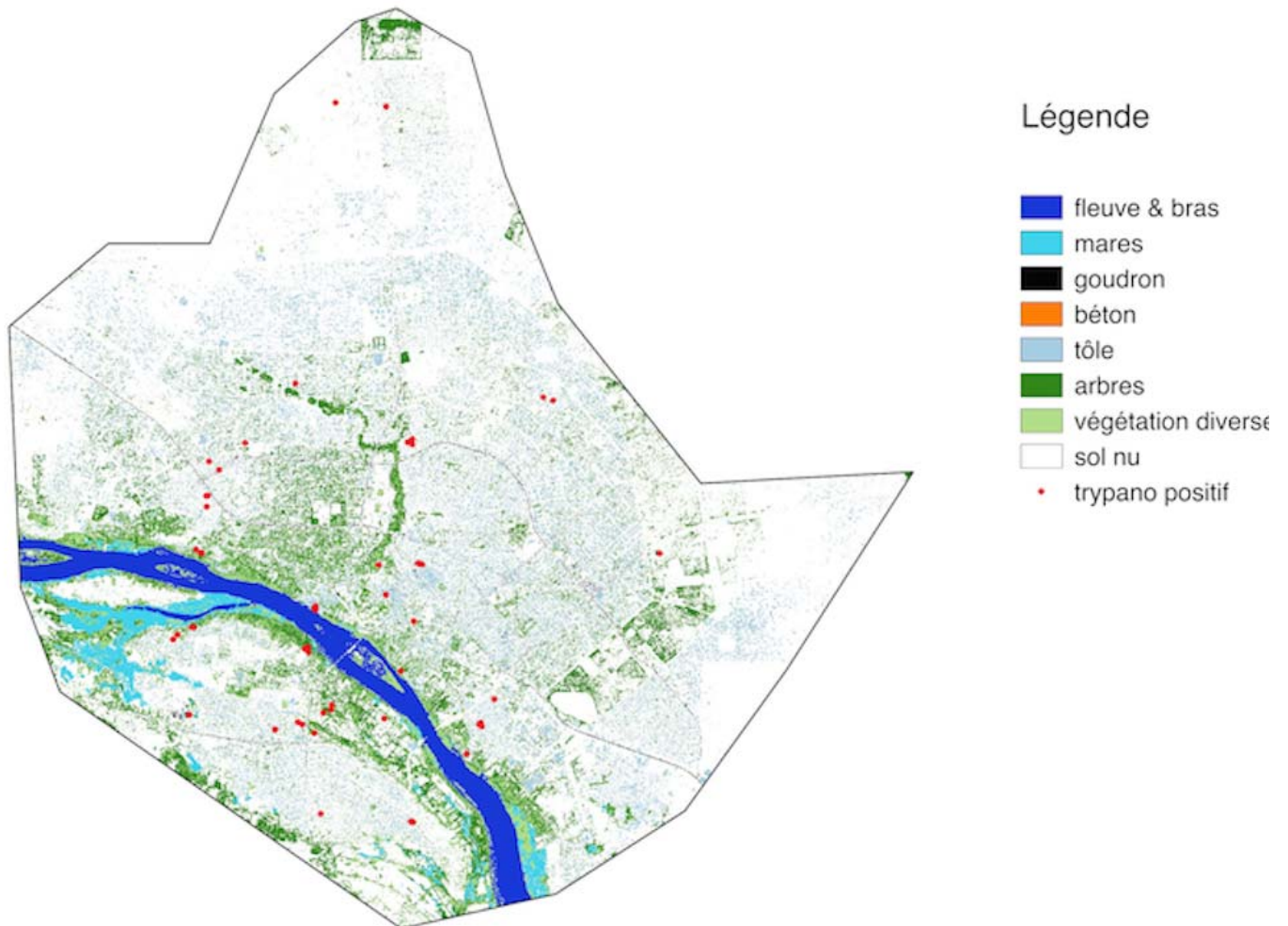
Elith et al. 2011. Diversity and Distributions 17: 43–57.

5-Données utilisées

Nous avons travaillé sur une série de données d'occurrence de *Trypanosoma lewisi* chez les rongeurs capturés dans la ville de Niamey.

Les données sur paysage urbain sont constituées d'un raster à 8 classes d'occupation du sol:

Sol nu, fleuve, mares, tôles, goudron, béton, arbres, végétation autre.

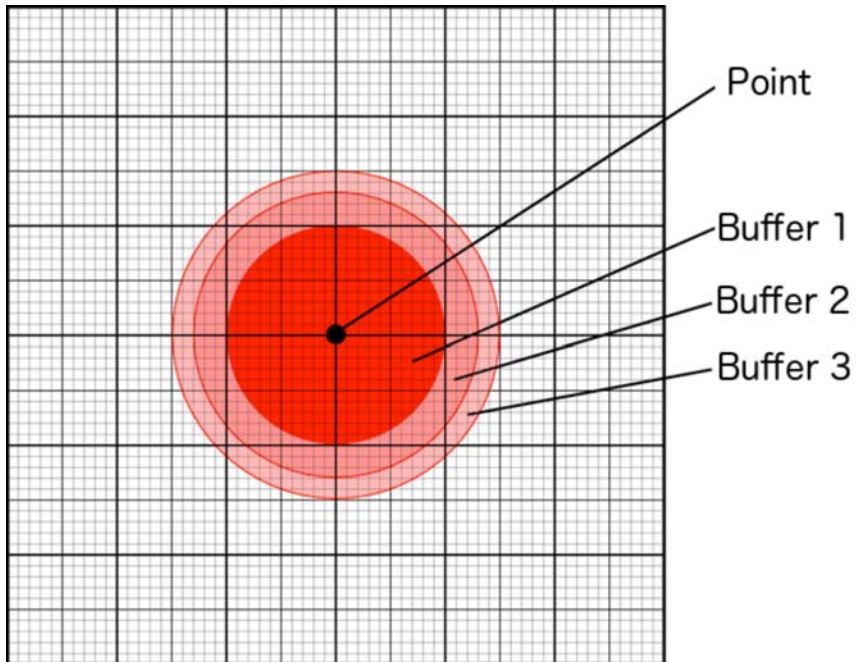


Carte de Niamey

6-Analyses statistiques / modélisation

Extraction de buffers autour des points d'occurrence

On utilise différents diamètres == différents grains/résolution



Analyse par fenêtres glissantes sur l'ensemble du domaine d'étude (5 928 529 fenêtres...)



SIDI (Indice de diversité de Simpson)

Modélisation de la distribution potentielle des cas positifs avec maxent

- Modèle “presence only” avec bonnes capacités prédictives
- Gère bien les situations avec peu d’observations et beaucoup de variables explicatives

Principe de l’analyse

- occurrences + paysage local → ajustement d’un modèle maxent
- modèle maxent + paysage complet (ville) → calcul de l’indice de “suitability” et cartographie sur toute la ville

Attention ! On a peu de points d’occurrence donc il faut s’assurer de la robustesse de l’approche.

Réalisation de 500 ajustements indépendants sur des sous-ensembles de données d’occurrence puis analyse statistique des AUC (= mesure de la qualité de l’ajustement)

7-Résultats

56 points avec cas positifs

127 points sans cas positifs

On travaille avec 5 grains différents → tailles de fenêtres (10, 20, 30, 40 et 50 m)

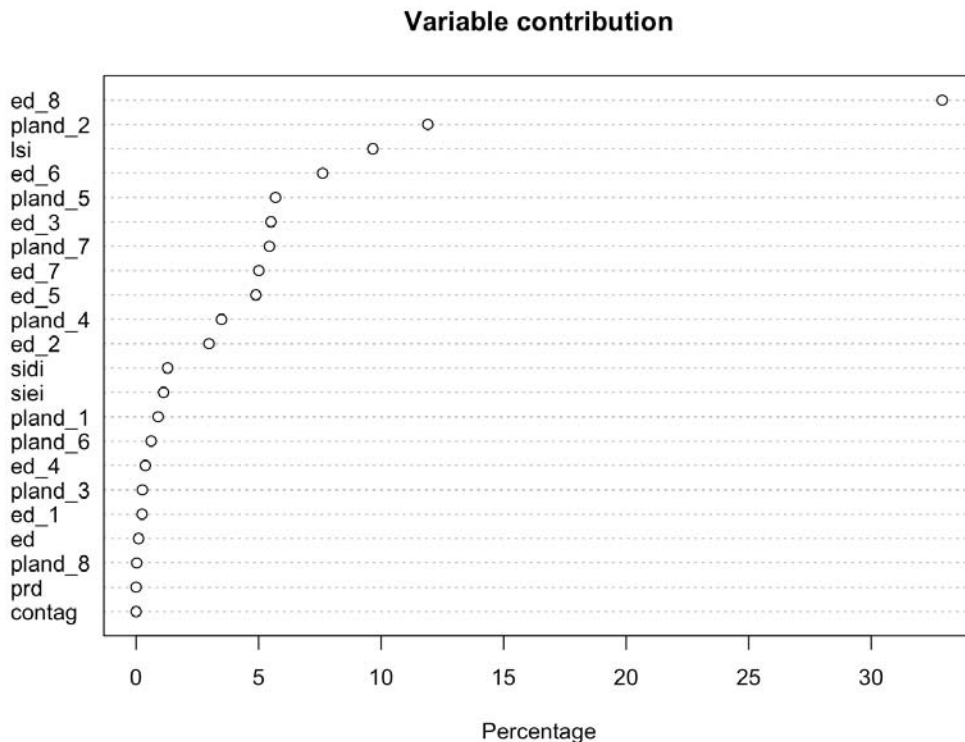
7a Un exemple avec des fenêtres circulaires de rayon = 50 m.

Mesure de la “qualité” du modèle

Le AUC est 0.7439086

Variation d'importance

1 = fleuve ; 2 = mares ; 3 = goudron ; 4 = béton 5 = tôles ; 6 = arbres ; 7 = autre végétation 8 = sol nu



7b Cartes de prédiction

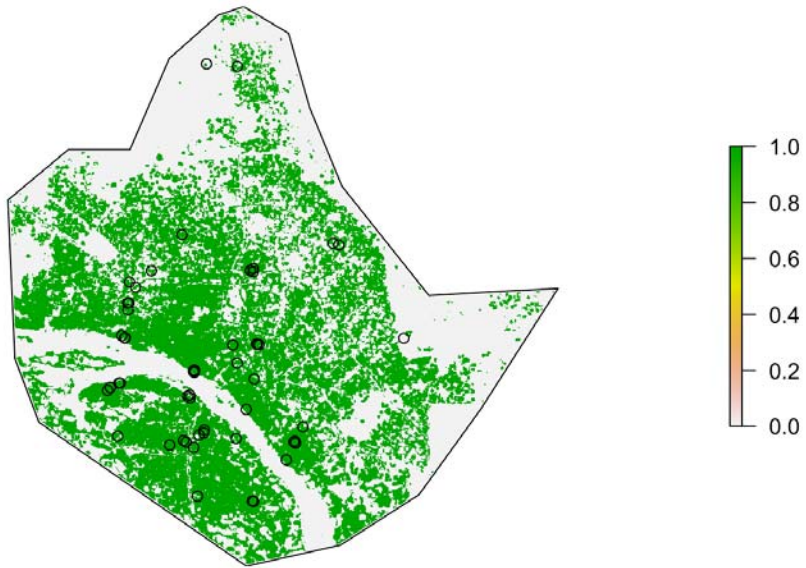
Le modèle permet d'estimer la qualité de l'habitat vis-à-vis de la variable à expliquer pour tous les points de l'espace. On obtient une carte.

“habitat suitability”



En appliquant un “seuil” sur les valeurs de l’“habitat suitability” on peut distinguer les pixels associés à des valeurs suffisamment élevées pour que le pixel soit considéré comme favorable. Les autres sont considérés comme défavorables.

“présence/absence” : seuil = maximise la sensibilité et la spécificité (taux de vrais positifs et vrais négatifs)



“no omission” : seuil = valeur pour laquelle aucune présence n’est omise.

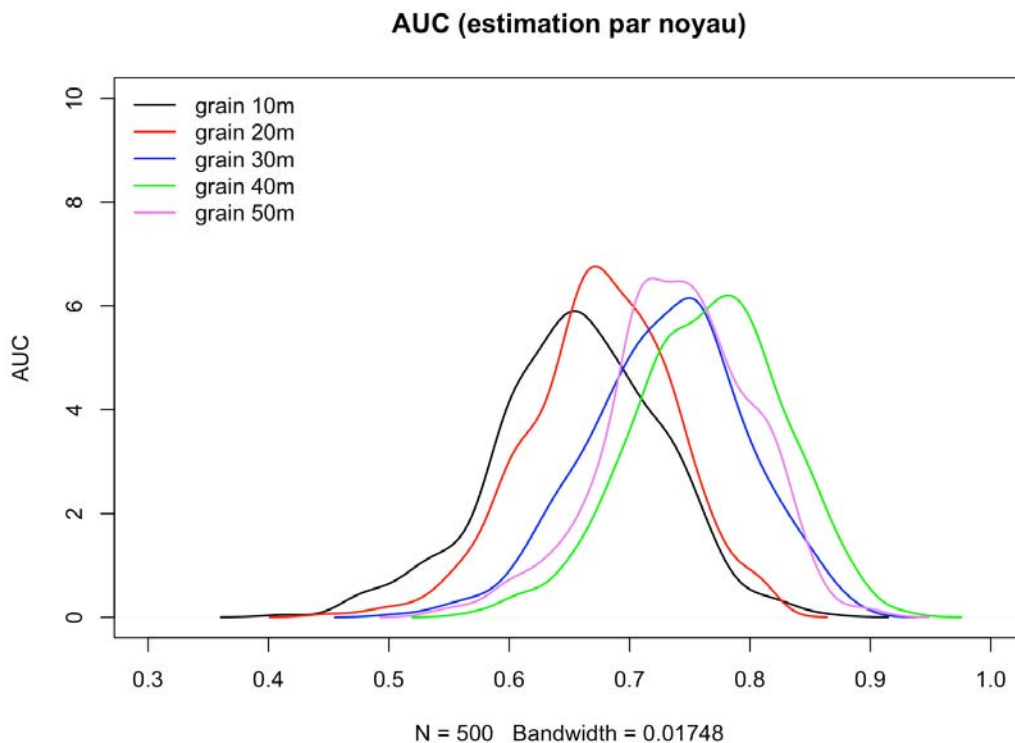


7c Approche statistique

Comme on a peu d'occurrences, les résultats du modèle dépendent fortement du choix aléatoire initial du lot de points utilisés pour l'ajustement (même effet lors de l'évaluation du modèle).

Pour avoir une idée claire des performances du modèle, on fait $n=500$ ajustements indépendants suivis de $n=500$ évaluations.

L'opération est répétée pour chaque grain (10 m, 20 m, 30 m, 40 m, 50 m).



Il y a un effet d'échelle clair. La qualité du modèle augmente pour des grains situés entre 10 et 40 m puis chute ensuite (50 m).

Pour grain > 50 calculs en cours.

7d Incertitude spatiale

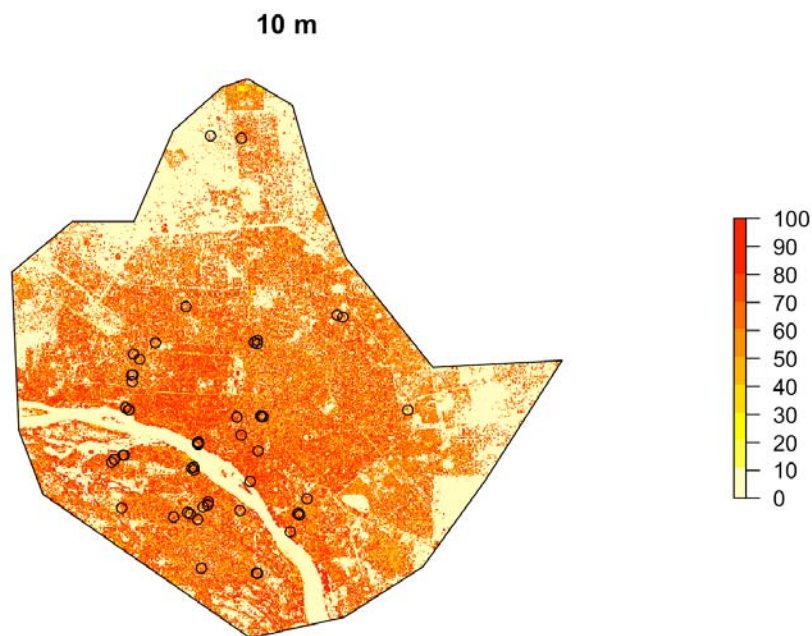
On ne peut se baser sur une seule réalisation, cf. forme des histogrammes.

On travaille alors sur les N modèles utilisés pour calculer les distributions de fréquence précédentes pour faire une sorte d'ensemble forecasting façon-à-à.

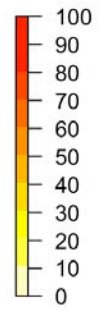
Chaque pixel est associé à une valeur qui est égale au nombre de fois où il a été "prédit" comme étant favorable (spec_sens).

La variabilité spatiale de la carte qui en résulte traduit l'incertitude spatiale sur la prédiction.

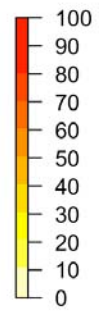
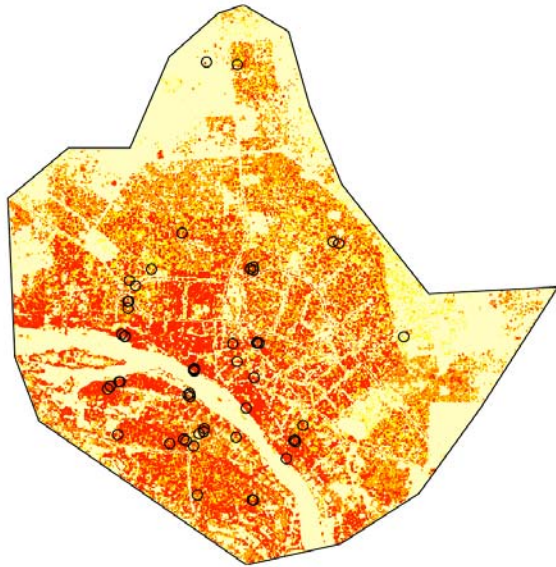
Voici des exemples basés sur des séries de $n=100$ répliquats.



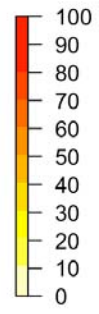
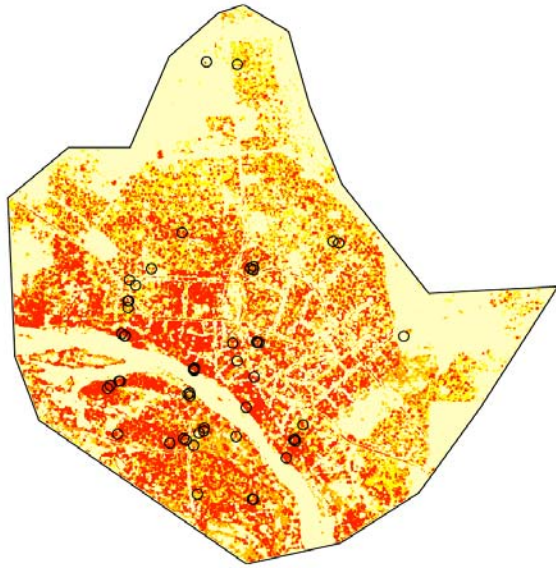
20 m



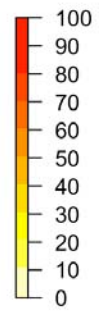
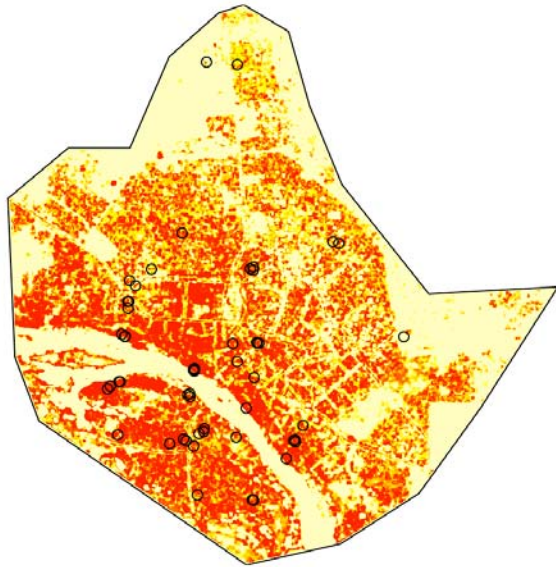
30 m



40 m



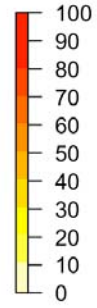
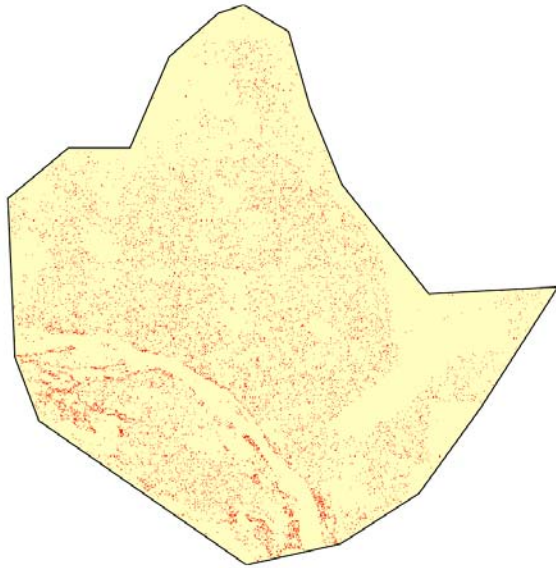
50 m



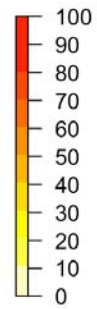
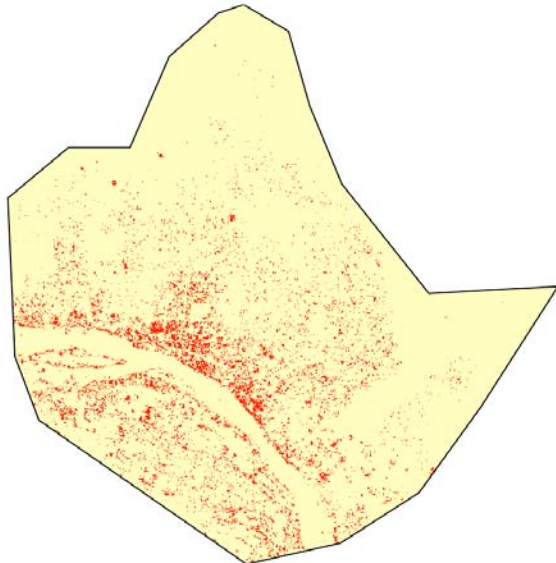
7e Vers des cartes de risque...

Proportion de simulations > 0.95

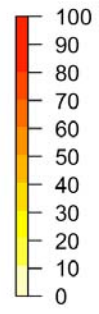
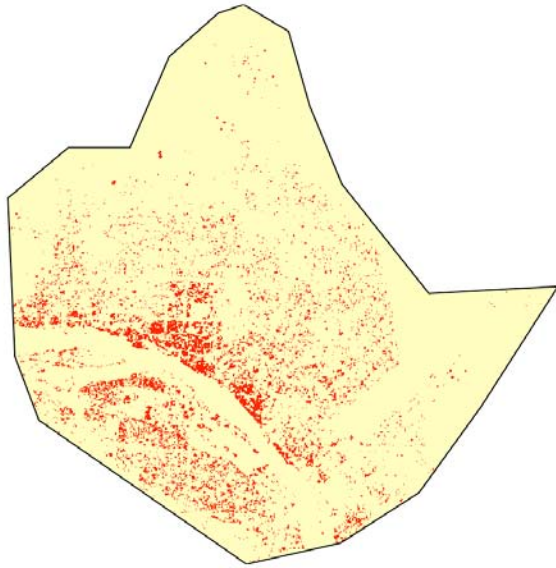
10 m



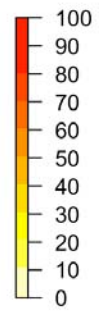
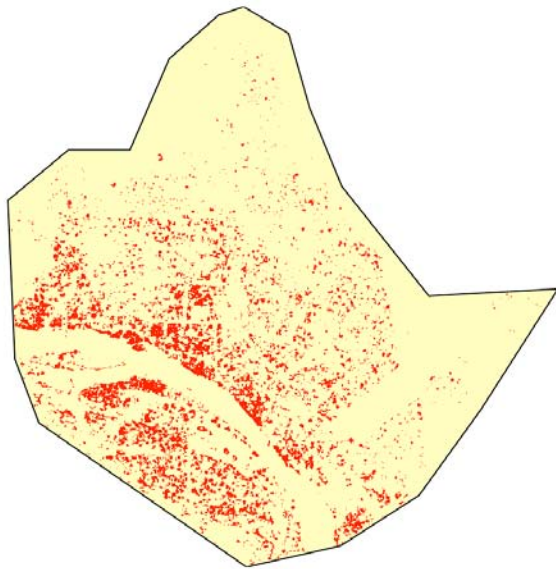
20 m



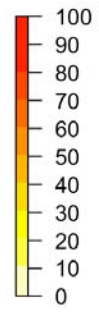
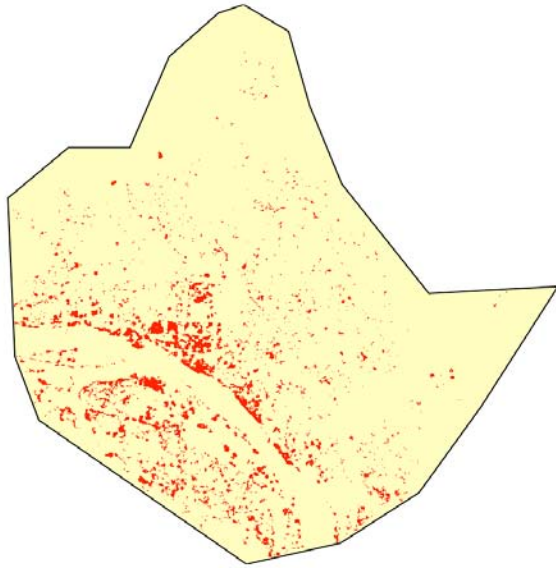
30 m



40 m



50 m

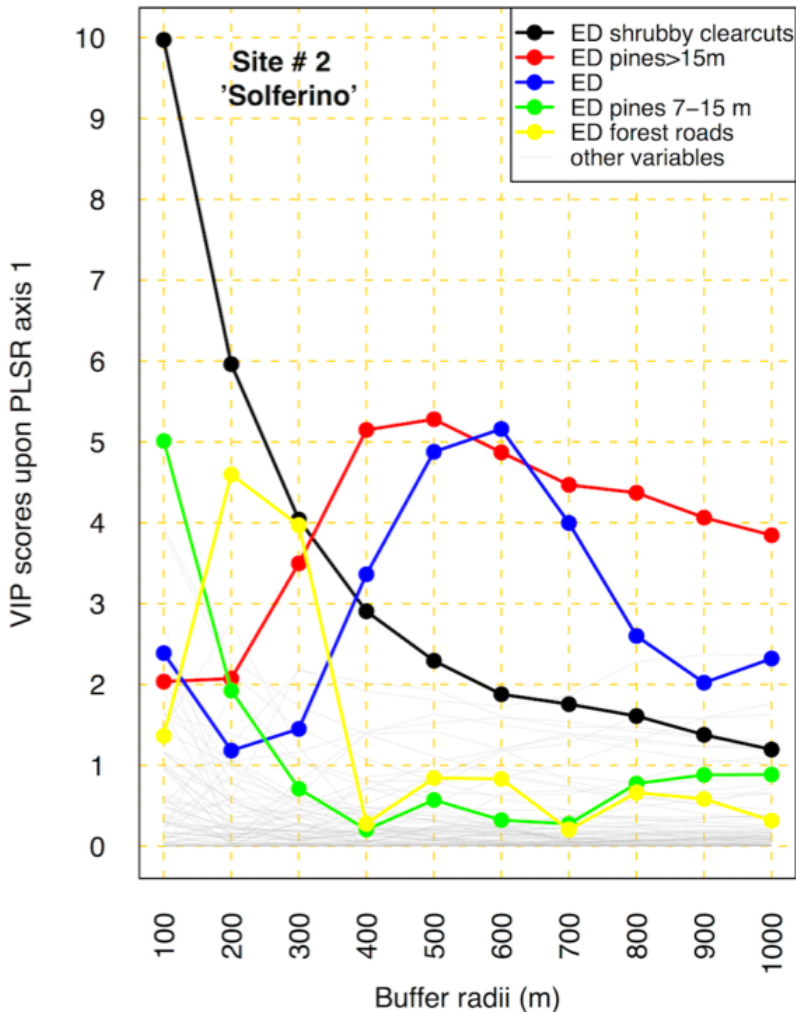


8-Conclusions *provisoires*

- La distribution des occurrences de la maladie n'est pas indépendante de la structure et de la composition du paysage urbain à Niamey.
- L'algorithme maxent permet de trouver les liens paysage-occurrences et de les formaliser dans un modèle stat.
- Le modèle permet d'estimer la distribution potentielle de la maladie à l'échelle de la ville.
- Une large proportion de la ville offre des conditions favorables.
- L'approche présentée pourraient se révéler utile pour informer le public et les autorités sur les zones potentiellement infectées et aider les chercheurs à mettre en place des suivis, des manip ou à optimiser les plans d'échantillonnage.
- La question des effets d'échelle est fascinante et mérite d'être considérée pour chaque modèle biologique. Si dans le cas présent, 40 m semble la solution optimale (reste à confirmer), rien ne prouve que ce serait la même chose pour la toxoplasmose ou la leptospirose.

9-Perspectives

- Examiner des grains de tailles > 50
- Regarder si les effets d'échelles jouent sur la contribution des variables (une variable ayant une bonne contribution à une échelle la conserve-t-elle à une autre ?)



Ex lien richesse spécifique des papillons et métriques de paysage (Rossi & Van Halder, 2010 Ecol. Indicat. 10, 452-458).

- Construction de modèle “multi-échelles”
- Tester d'autres type de modèle (BRT, ...)
- Réaliser un grand nombre de predictions et analyser l'incertitude spatiale

- Explorer d'autres jeux de données comparables (toxco, leptos et toutes les maladies que Gauthier attrapera à Cotonou ou ailleurs...)

Merci de votre attention !

