

# Calculer des tendances régionales à partir d'une enquête auprès des ménages : cas de la prévalence du VIH dans les Enquêtes Démographiques et de Santé (EDS)

J. Larmarange<sup>a</sup>, R. Vallo<sup>b</sup>, S. Yaro<sup>c</sup>

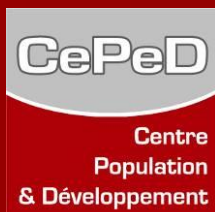
P. Msellati<sup>d</sup>, N. Méda<sup>c</sup>, B. Ferry<sup>a</sup>

<sup>a</sup> IRD / CEPED UMR 196 • <sup>b</sup> Université Montpellier EA 4205

<sup>c</sup> Centre Muraz • <sup>d</sup> UMR 145 IRD / Université de Montpellier I

Conférence francophone sur les Sondages • Tanger • 23 mars 2010

[joseph.larmarange@ceped.org](mailto:joseph.larmarange@ceped.org)

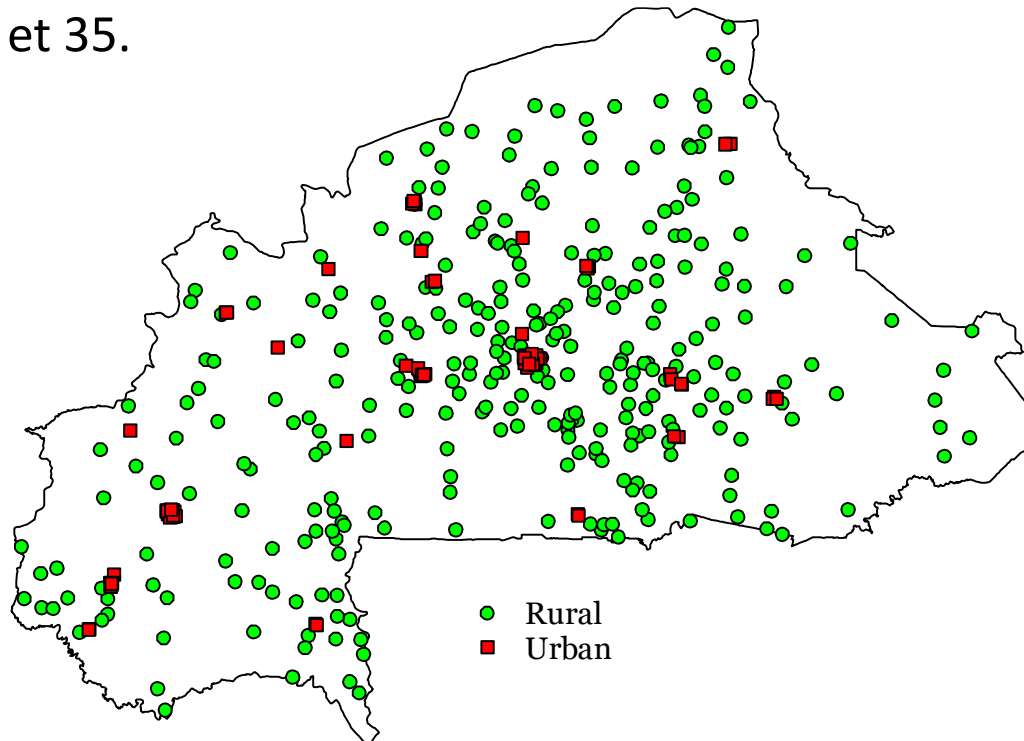


# Enquêtes Démographiques et de Santé (EDS)

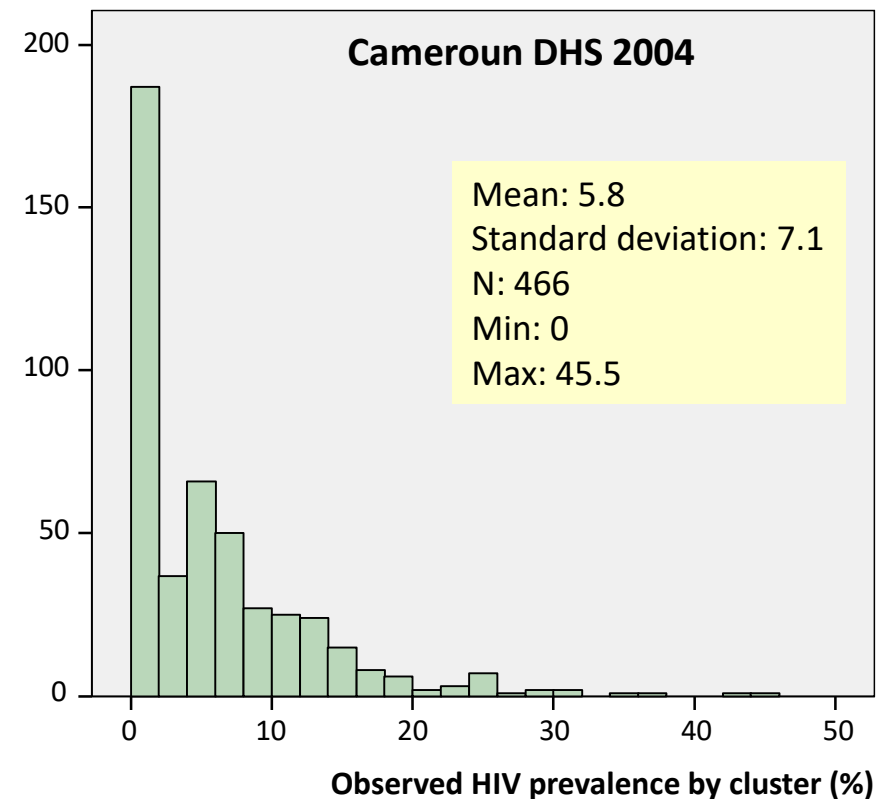
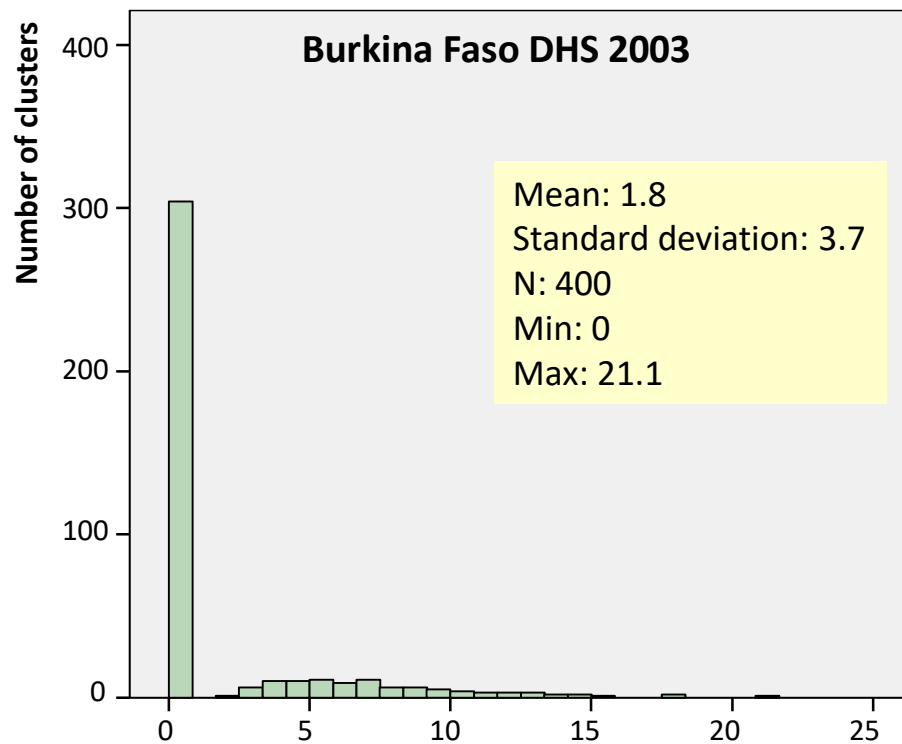
- Les EDS sont bien connues des démographes.
- Depuis 1984, plus de 200 EDS ont été réalisées dans plusieurs pays à faible et moyen revenu, notamment en Afrique.
- Sur certains sujets, comme la prévalence du VIH, les EDS constituent la seule source de données nationale en population générale.
- Certaines EDS collectent la longitude et la latitude des zones d'enquêtes, mais cette information reste peu exploitée.
- Les EDS sont échantillonnées afin d'être représentative au niveau national et régional. Mais les cartes par région sont rarement adaptés pour explorer les variations spatiales d'un phénomène.

# Échantillonnage d'une EDS : exemple du Burkina Faso 2003

- Burkina Faso 2003 : échantillonnage stratifié en grappes à deux degrés
  - 400 grappes (zones d'enquêtes)
  - 3179 ménages • 7151 personnes testées pour le VIH
  - ➔ 17,9 personnes testées par grappe en moyenne
- Selon l'EDS, le nombre moyen de personnes testées par grappe varie entre 10 et 35.



# Prévalence du VIH observée par grappe



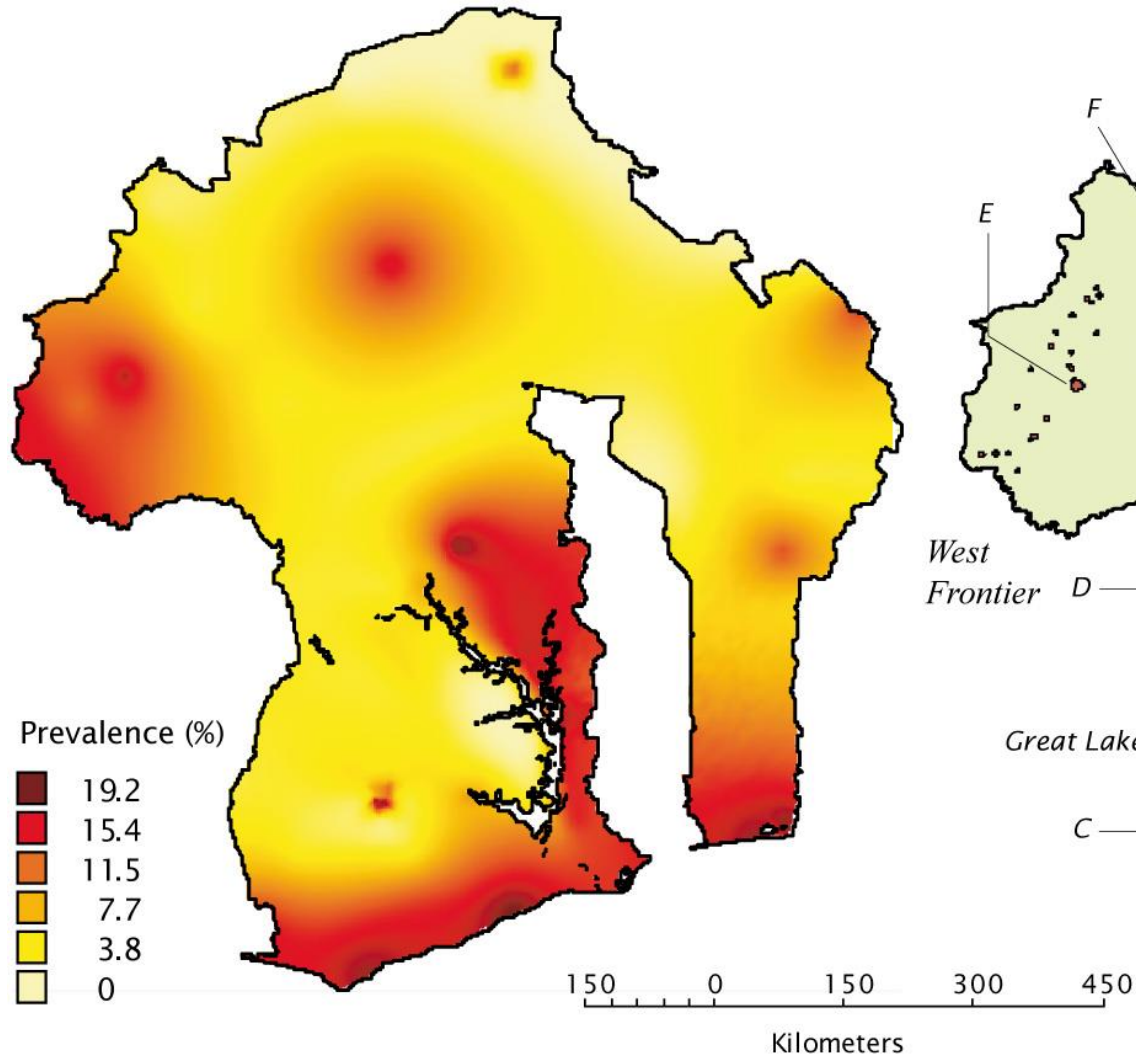
*Prévalence observée = Prévalence réelle + Erreur aléatoire (EA)*

# Approche développée

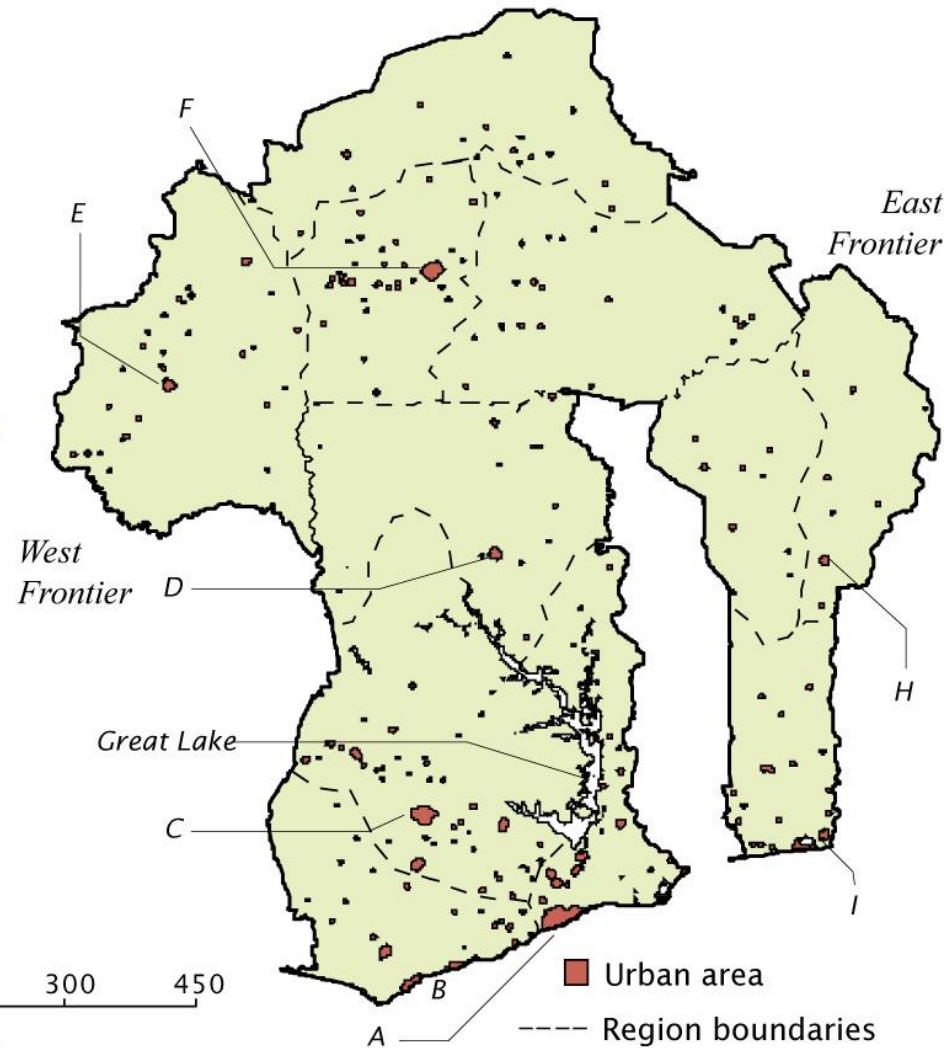
- Inspirée des analyses en composantes d'échelle
- Prév. obs. = tendance régionale + composante locale + EA
- Peut-on estimer une tendance régionale pour chaque grappe ? (compensation du second tirage)
- Puis procéder à une interpolation spatiale de ces tendances régionales ? (compensation du premier tirage)
- Pour tester notre approche, élaboration d'un pays fictif avec une épidémie fictive.
- Objectif :
  - pas une estimation précise de la prévalence en chaque point de la carte,
  - mais reproduire les principales variations spatiales de l'épidémie.

# Pays fictif

a. Prevalence in the model (national level of 10%)



b. Regions and urban areas

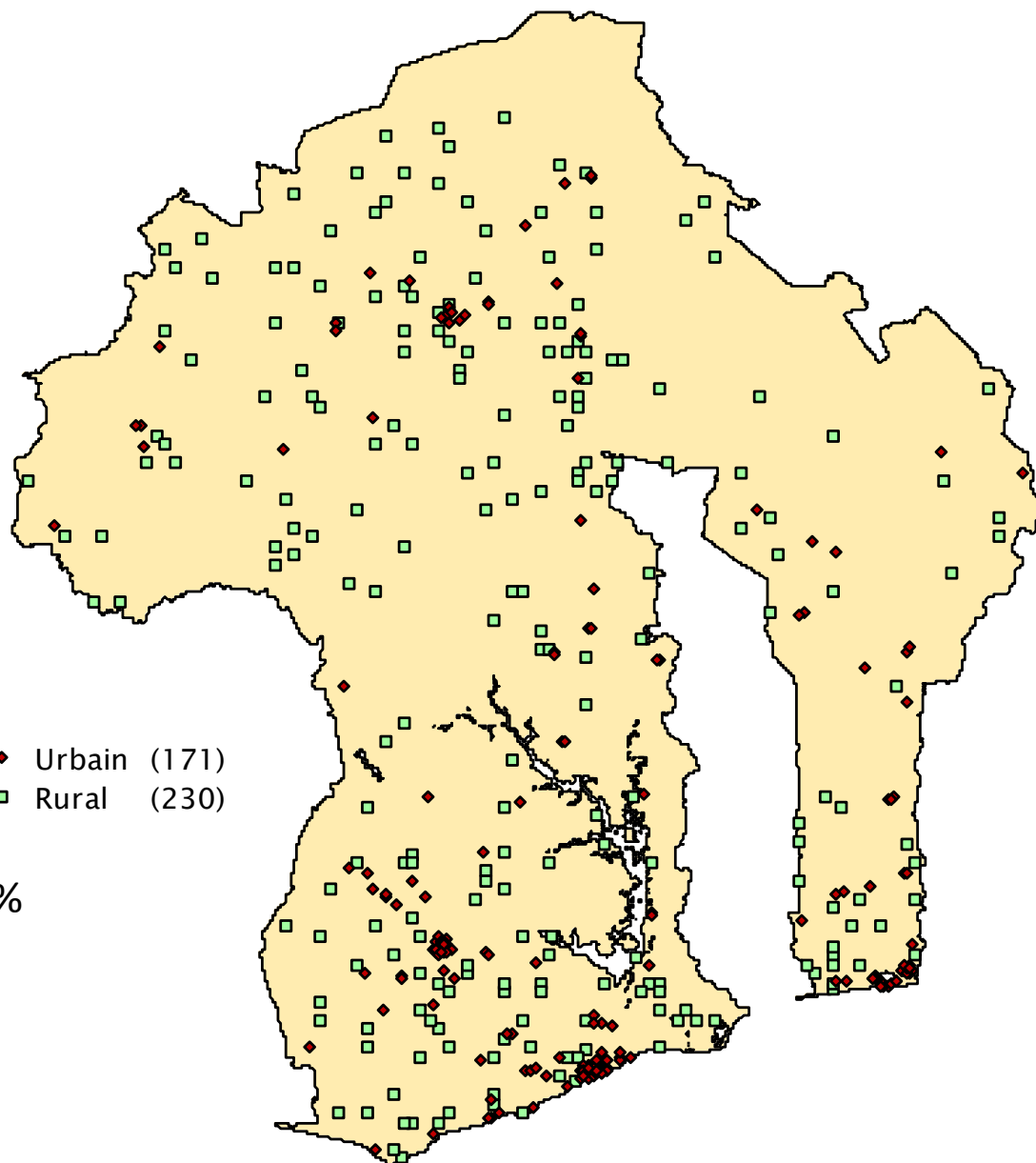


# Simulation d'une EDS

- Exemple avec :

- Prévalence nationale de 10%
- 8000 personnes testées
- 400 grappes

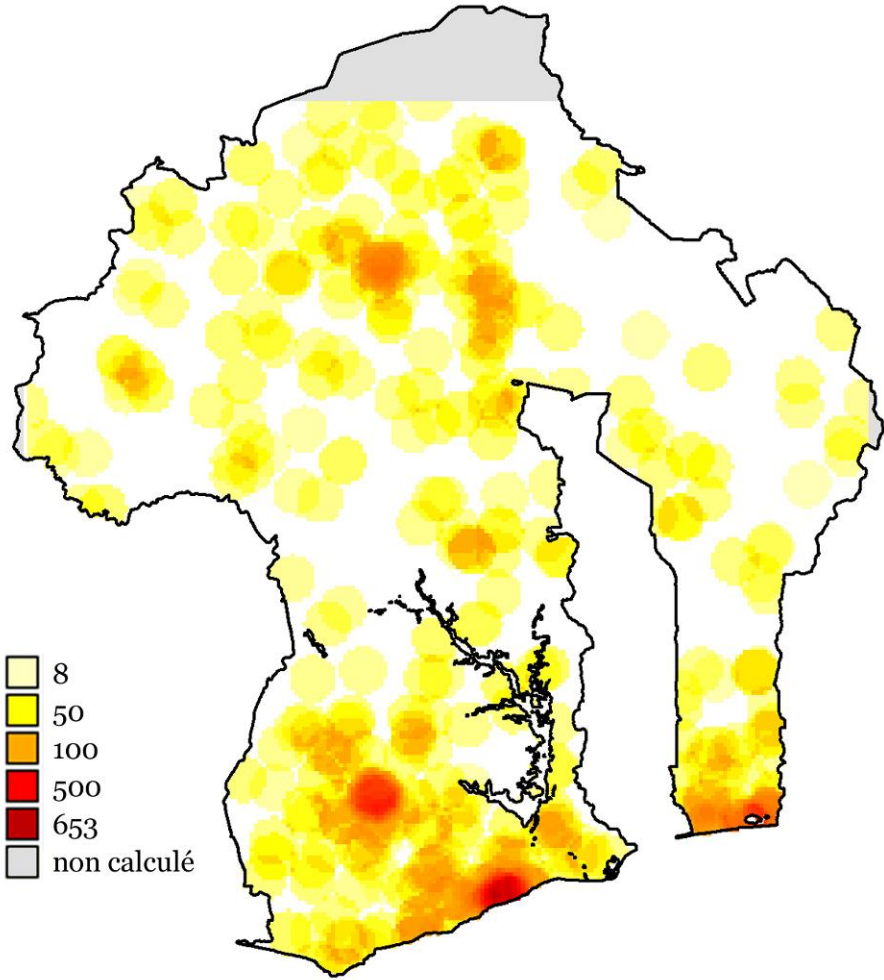
◆ Urbain (171)  
■ Rural (230)



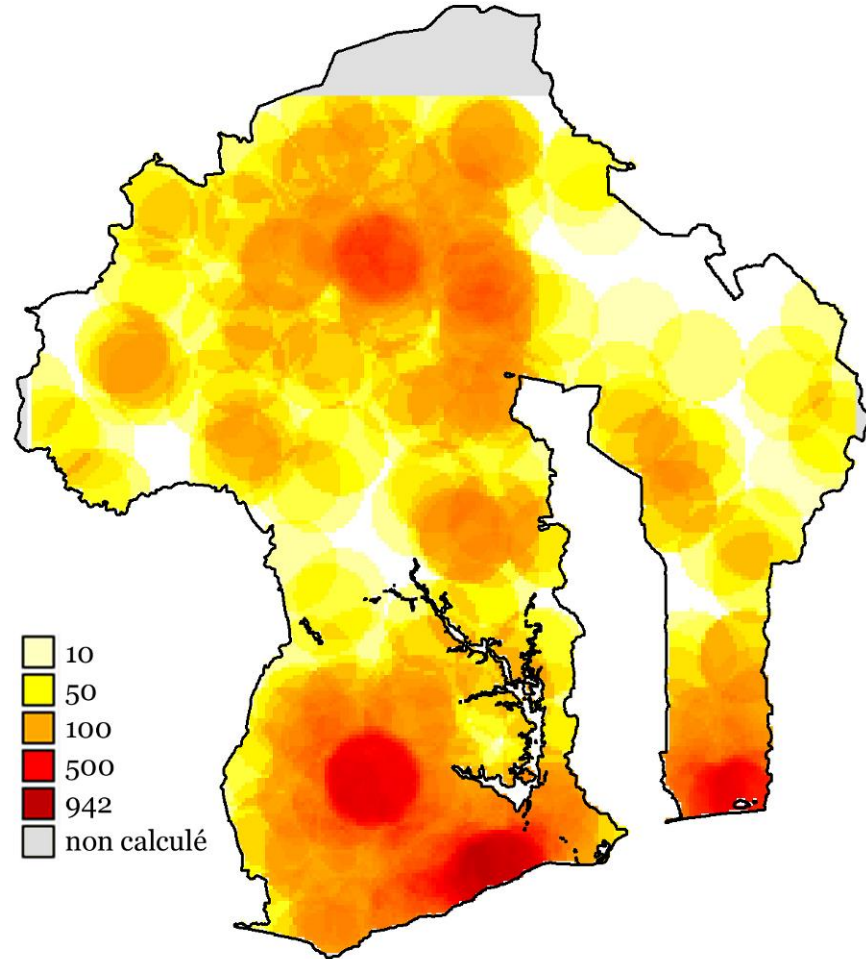


# Personnes testées dans un rayon de 50 et 100 kilomètres

50 km



100 km





# Méthode N : cercles de même effectif

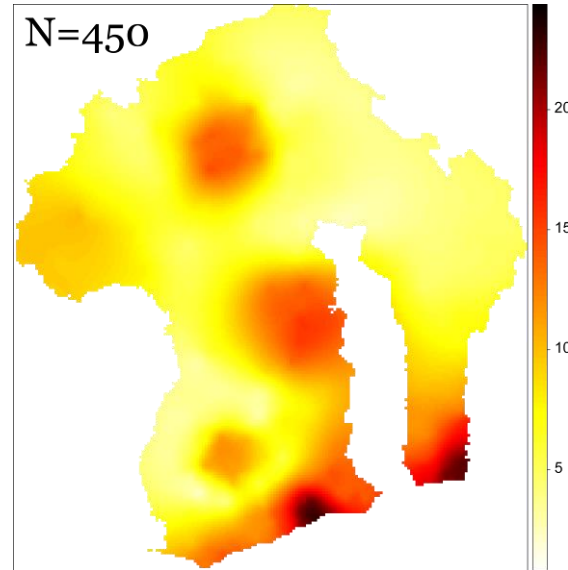
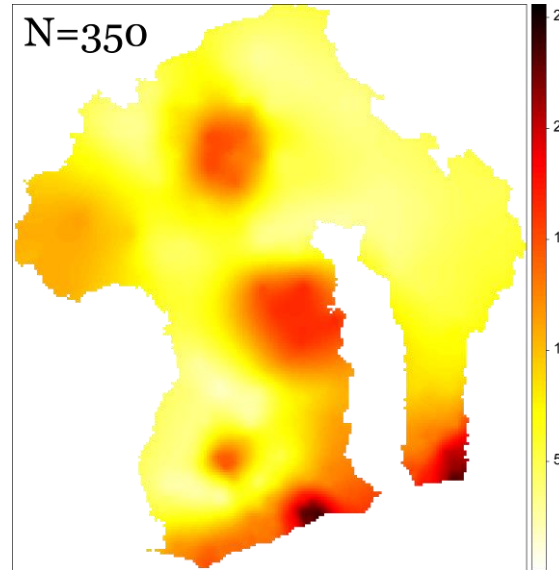
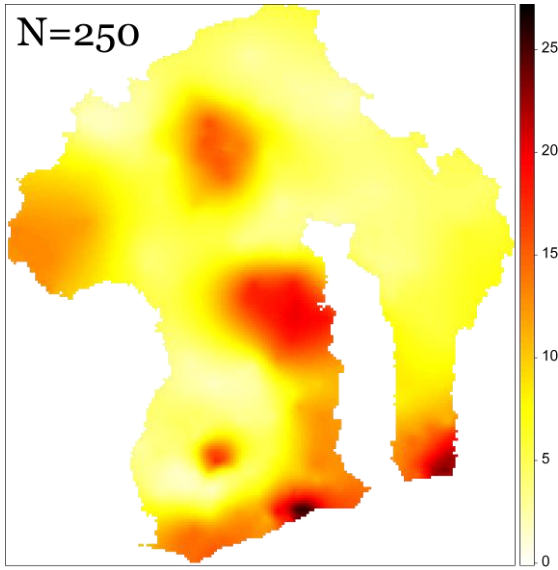
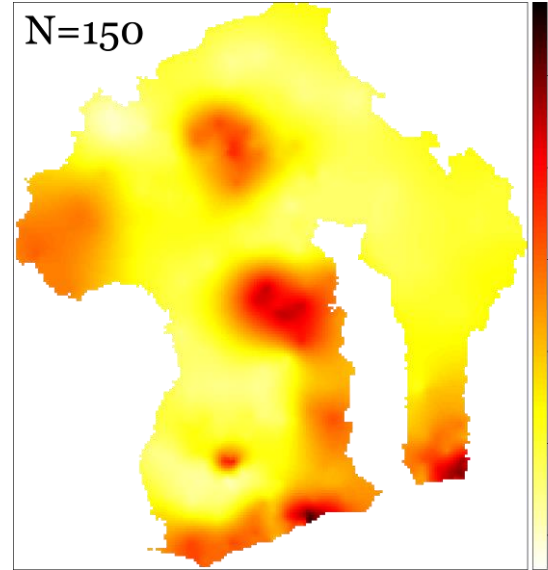
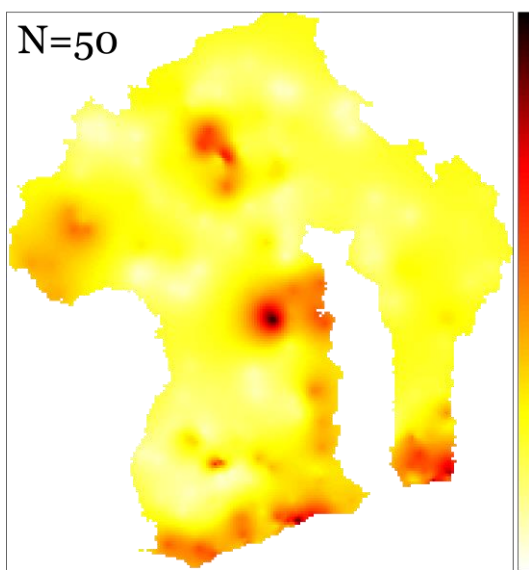
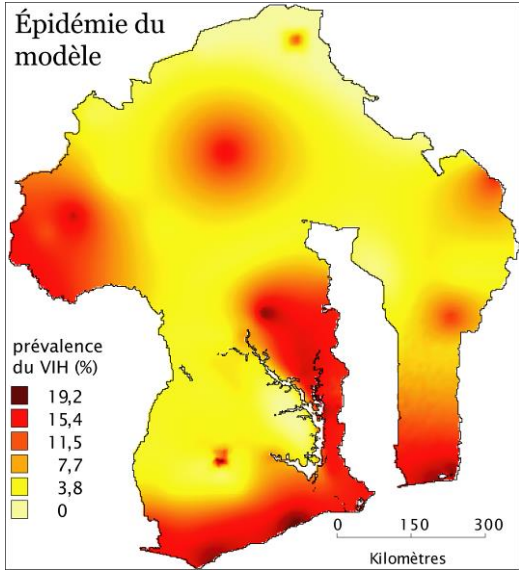
Dans cet exemple,  
N est fixé à 250

123 km

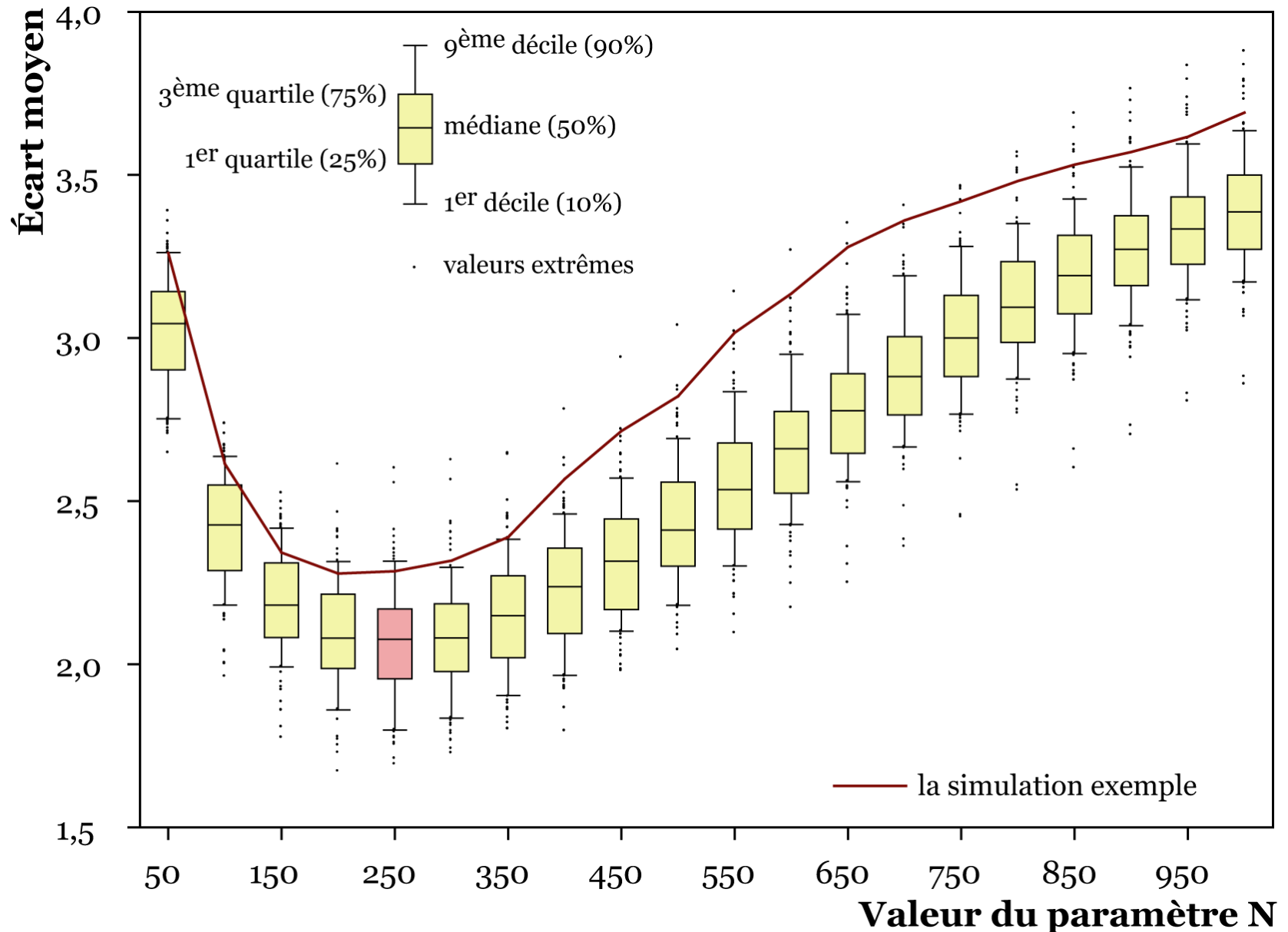
$n = 2157$

$p = 2,6\%$

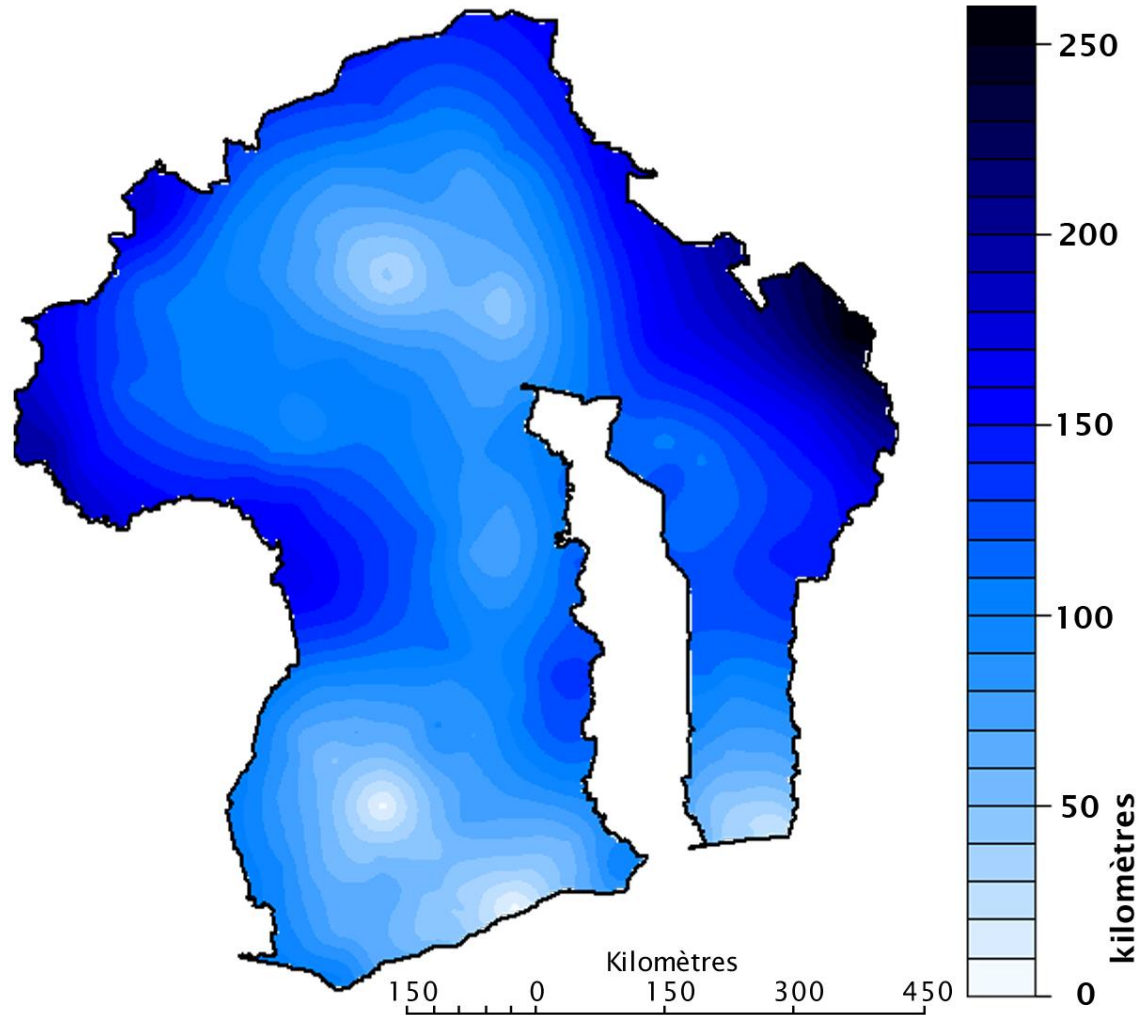
# Méthode N et interpolation spatiale par krigeage



# Écart moyen pour 100 simulations identiques



# Rayon du cercle de lissage pour N=250



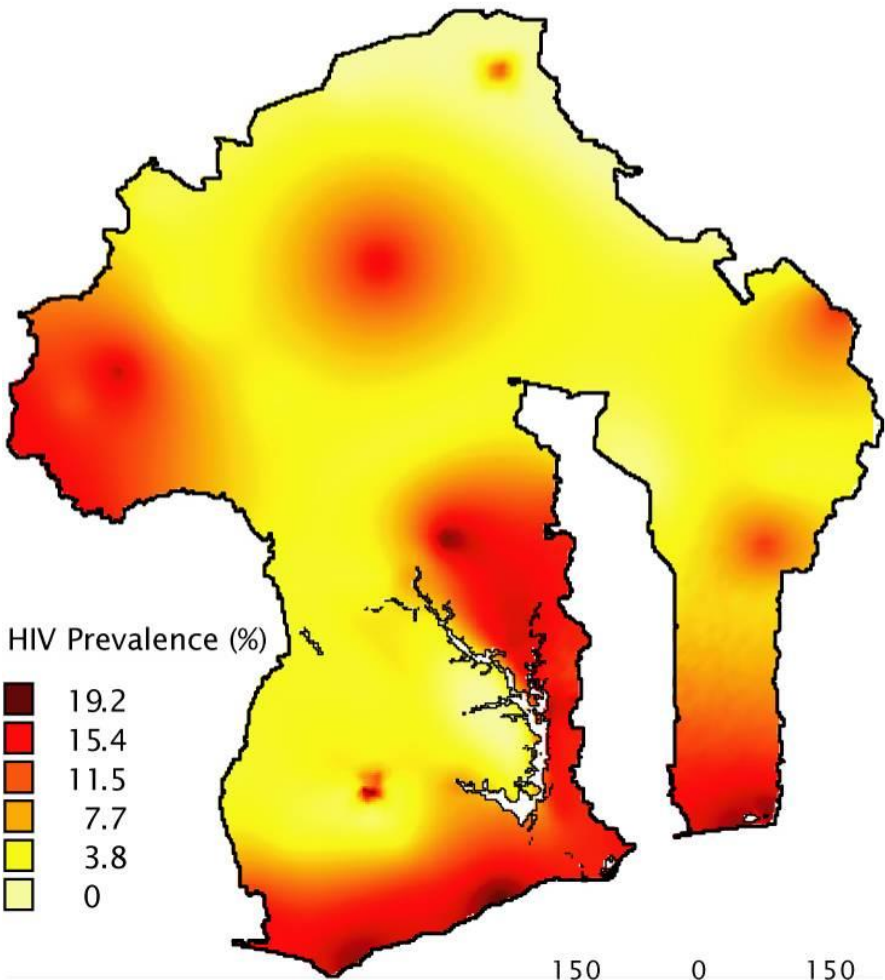
- Dans les zones peu enquêtées
  - Rayon très élevé
  - Proposition : ajout d'un second paramètre R, correspondant à un rayon maximal.
  - Nous avons fait le choix du 9<sup>e</sup> décile.
  - Dans cet exemple, R=128 km

# Prendre en compte les agglomérations urbaines

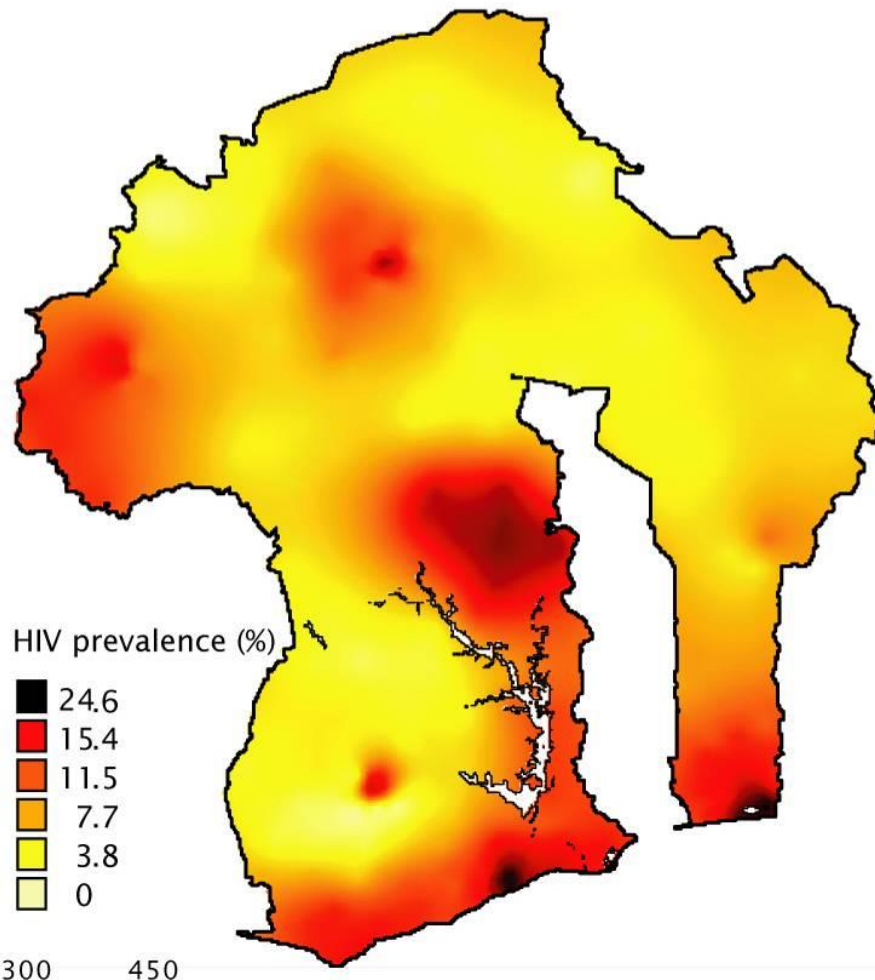
Agglomération	Grappes	Observations	Prévalence observée
A	24	489	21.2 %
I	18	352	23.3 %
C	17	346	15.2 %
F	7	151	18.0 %
E	3	48	16.3 %
H	2	32	10.3 %

# Épidémie du modèle et tendances régionales estimées

a. Prevalence of the model (national level of 10%)



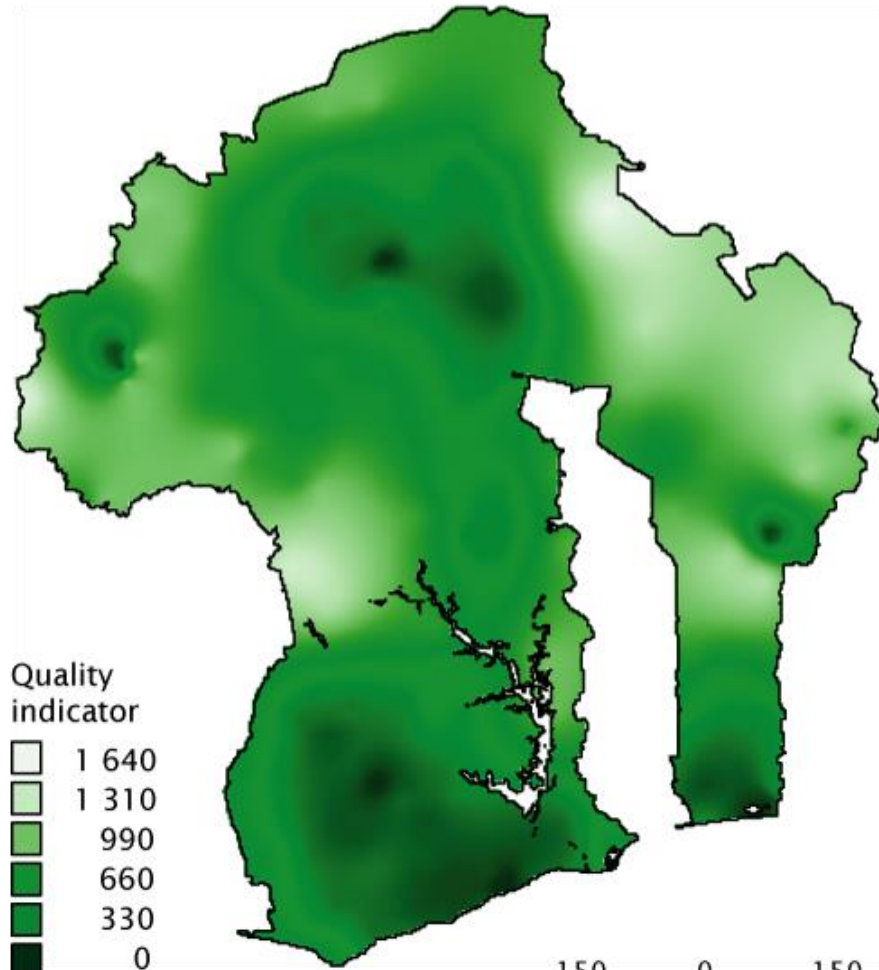
b. Estimated prevalence from a DHS simulation



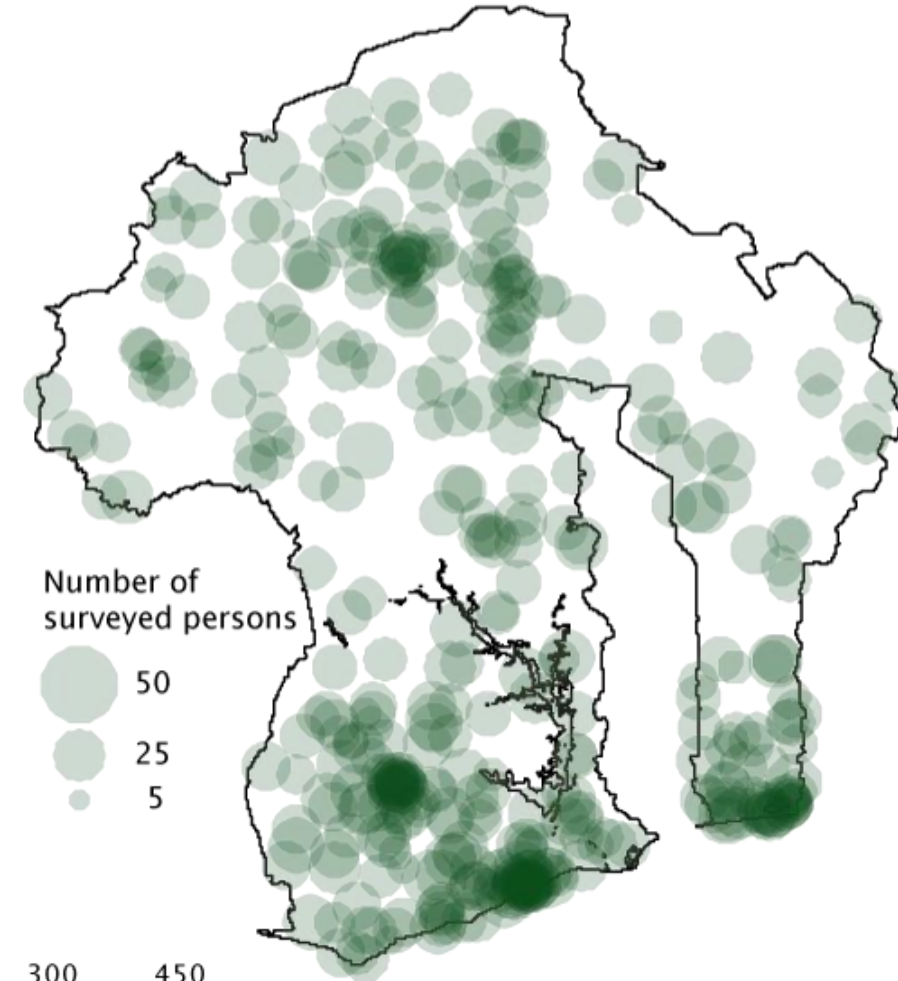


# Cartes complémentaires

c. Quality Indicator

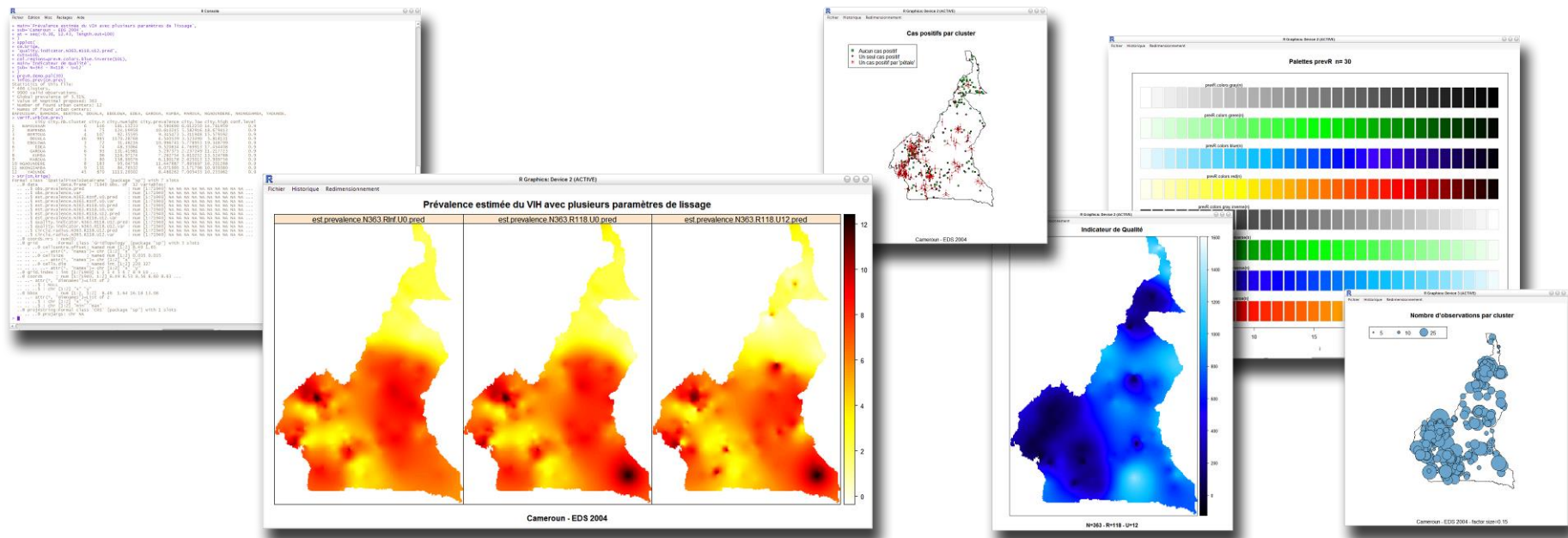


c. Surveyed persons by cluster



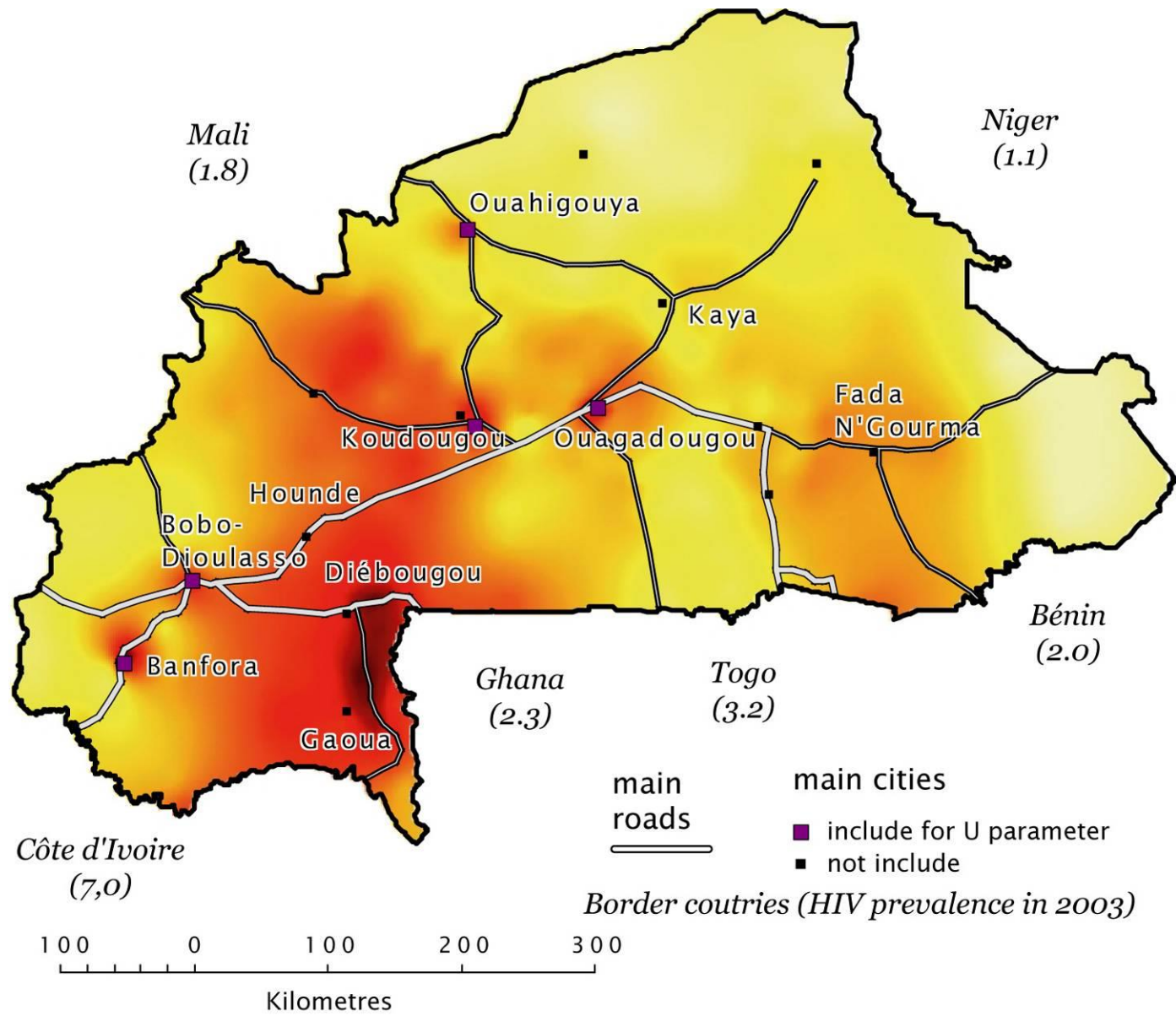


# prevR : réalisation de ces analyses sous R



- Disponible gratuitement sur <http://www.ceped.org/prevR>
- Bilingue : anglais / français
- Import des données, analyses, export dans des formats standards
- En cours de réécriture pour offrir une interface graphique utilisateur

# Application au Burkina Faso: prévalence du VIH 2003



# Comment choisir le bon paramètre de lissage ?

- Comme pour les estimateurs à noyau, toute la difficulté consiste à choisir le bon paramètre de lissage, le paramètre  $N$  dans le cas présent.
- Pour le pays fictif, un paramètre optimal peut être calculé puisque la surface à estimer est connue, ce qui n'est pas le cas avec des données réelles d'enquêtes.
- Nous avons fait le choix d'une épidémie présentant des variations spatiales correspondant aux tendances régionales que nous espérons estimer.
  - Une autre épidémie fictive peut amener à d'autres valeurs optimales.
- L'application à des données réelles des valeurs optimales calculées sur notre pays fictif nous a permis néanmoins d'obtenir des résultats cohérents.

- Les cercles de même effectif permettent de prendre en compte l'hétérogénéité de la distribution des observations.
- La méthode est simple à mettre en œuvre avec prevR.
- Les résultats sont limités :
  - estimations indirectes à partir des grappes voisines
  - approche descriptive et non explicative
  - la difficulté posée par le choix du paramètre N
- Mais cohérents :
  - avec les connaissances actuelles sur les épidémies de VIH.
- Cette méthode n'est qu'une première étape dans une recherche :
  - Disposer à partir d'une EDS d'une première image des variations spatiales d'un phénomène avant la réalisation d'une enquête spécifique permettant de comprendre et d'expliquer ces variations.

# Merci de votre attention

<http://www.ceped.org/prevR>

