



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

BS, DMV, PhD,  
DACVPM  
Professeur  
Département  
d'études sur les  
animaux de  
production  
Université de  
Prétoira  
Co-rédacteur en  
chef de *Médecine  
Vétérinaire  
Préventive*



**Geoffrey T. Fosgate**

**GF-TAD**

**Fièvre aphteuse**

**Atelier de formation sur l'évaluation des risques  
Johannesburg, Afrique du Sud 19-21 septembre  
2023**

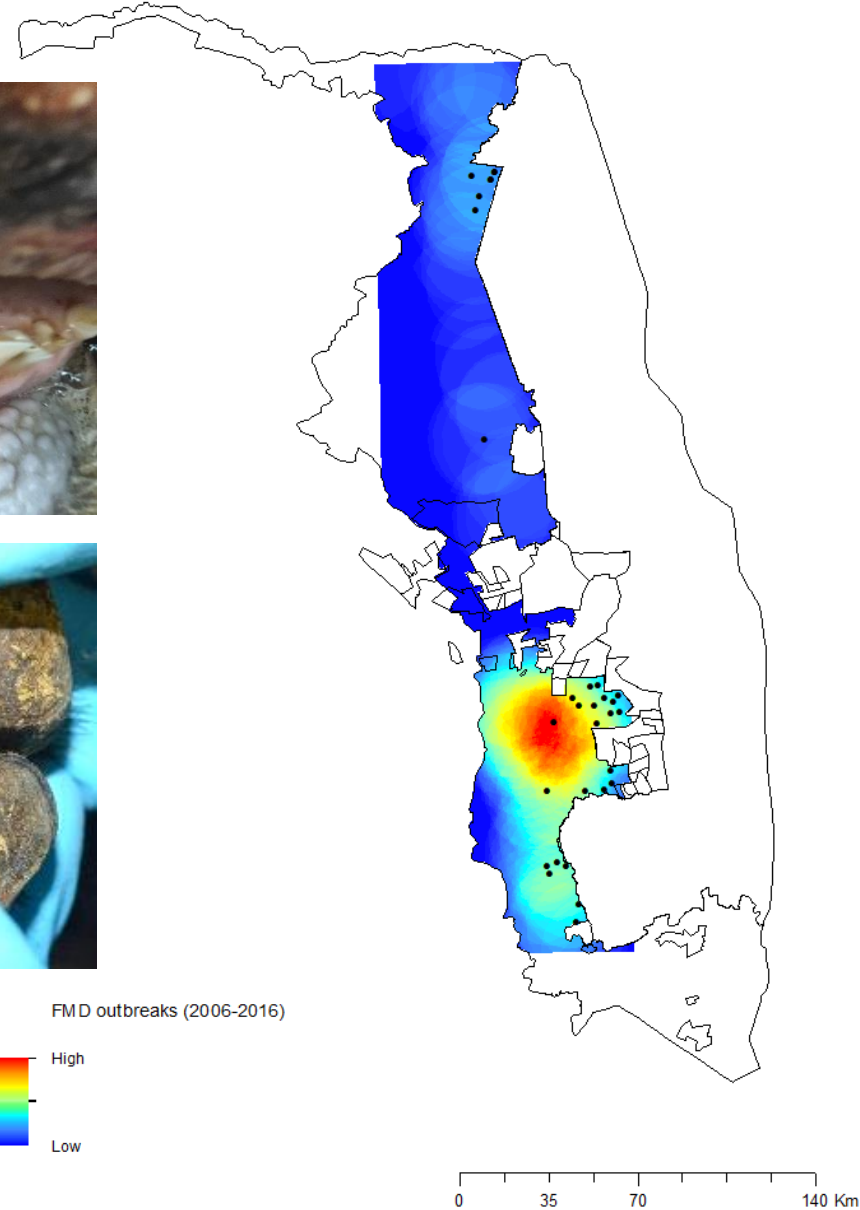
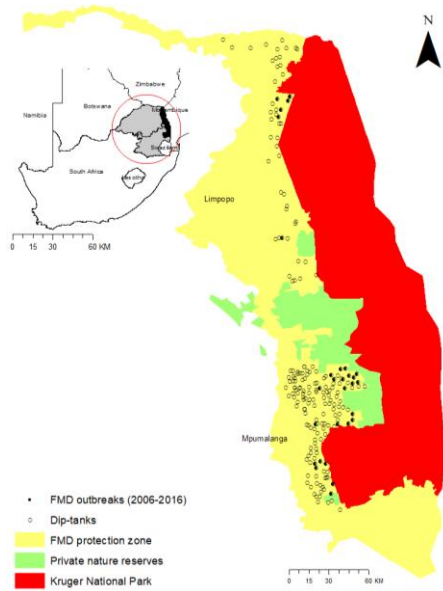


# Épidémiologie spatiale

- Introduction à l'épidémiologie spatiale
- Définition du cluster
- Présentation descriptive
  - Cartes de points
  - Cartes choroplèthes
- Cartographie des risques
  - Pondération inverse de la distance
  - Krigeage
- Détection de clusters
  - Temporel
  - Spatial
  - Temporospatal



# Épidémiologie spatiale



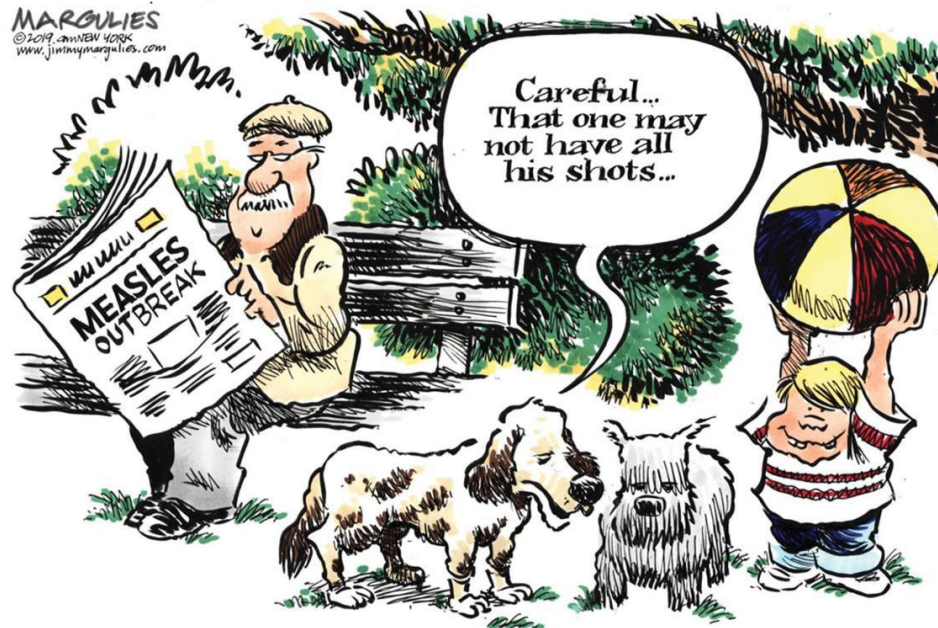
# Épidémiologie spatiale

- **L'étude de la répartition spatiale des états liés à la santé et des déterminants de la santé dans les populations**
- **L'épidémiologie spatiale fournit un cadre pour examiner les influences de l'espace et du lieu sur la santé**
  - **Décrire et analyser les types de maladies**
  - **Explorer et analyser les modèles spatiaux**
  - **Faire des hypothèses sur d'éventuelles relations causales**
- **Méthodes**
  - **Cartographie descriptive des maladies**
  - **Cartographie des risques**
  - **Détection de clusters**
- **Le lieu peut être utilisé comme substitut pour les influences sur la maladie**
  - **Exposition aux risques environnementaux**
  - **Réseaux de déplacements d'animaux**
- **Facteurs de gestion**



# Épidémiologie spatiale

- Une épidémie est une augmentation, souvent soudaine, du nombre de cas d'une maladie au-delà de ce qui est normalement attendu dans la population d'une zone donnée.
- Une épidémie est également une augmentation soudaine de l'incidence d'une maladie, mais elle est souvent utilisée pour des répartitions géographiques limitées.
- Un cluster est une agrégation de cas dans le lieu et dans le temps qui sont plus



[https://madison.com/ct/news/opinion/column/cartoons-of-the-week/collection\\_d62e61e8-61ff-11e9-8ac9-27446eda148d.html/](https://madison.com/ct/news/opinion/column/cartoons-of-the-week/collection_d62e61e8-61ff-11e9-8ac9-27446eda148d.html/)

# Présentation descriptive

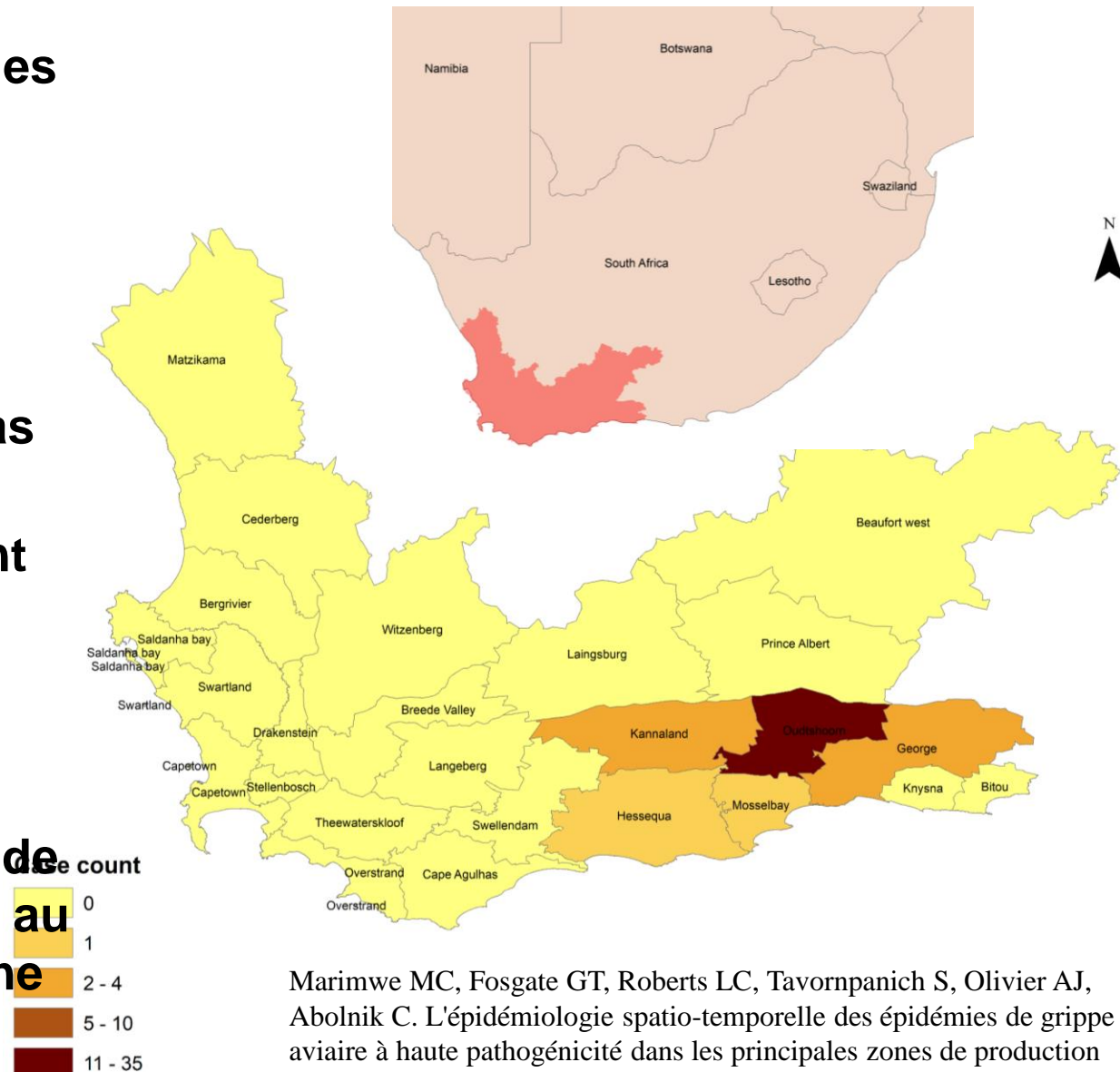


- Les cartes ponctuelles permettent une distribution simple des rapports de cas
- Ne tient pas compte de la population sous-jacente à risque
- Peut donner une indication des zones à haut risque et des facteurs de risque possibles

Chimera ET, Fosgate GT, Etter EMC, Jemberu WT, Kamwendo G, Njoka P. Modèles spatio-temporels et facteurs de risque de la fièvre aphteuse au Malawi entre 1957 et 2019. *Prev Vet Med* 2022 ;204 : 105639 .

# Présentation descriptive

- Les cartes choroplèthes sont regroupées pour les régions géopolitiques
- Lorsque des emplacements individuels ne sont pas disponibles
- Les démarcations sont arbitraires et sans rapport avec des facteurs épidémiologiques
- Pourrait tenir compte de la population à risque au sein des unités de zone

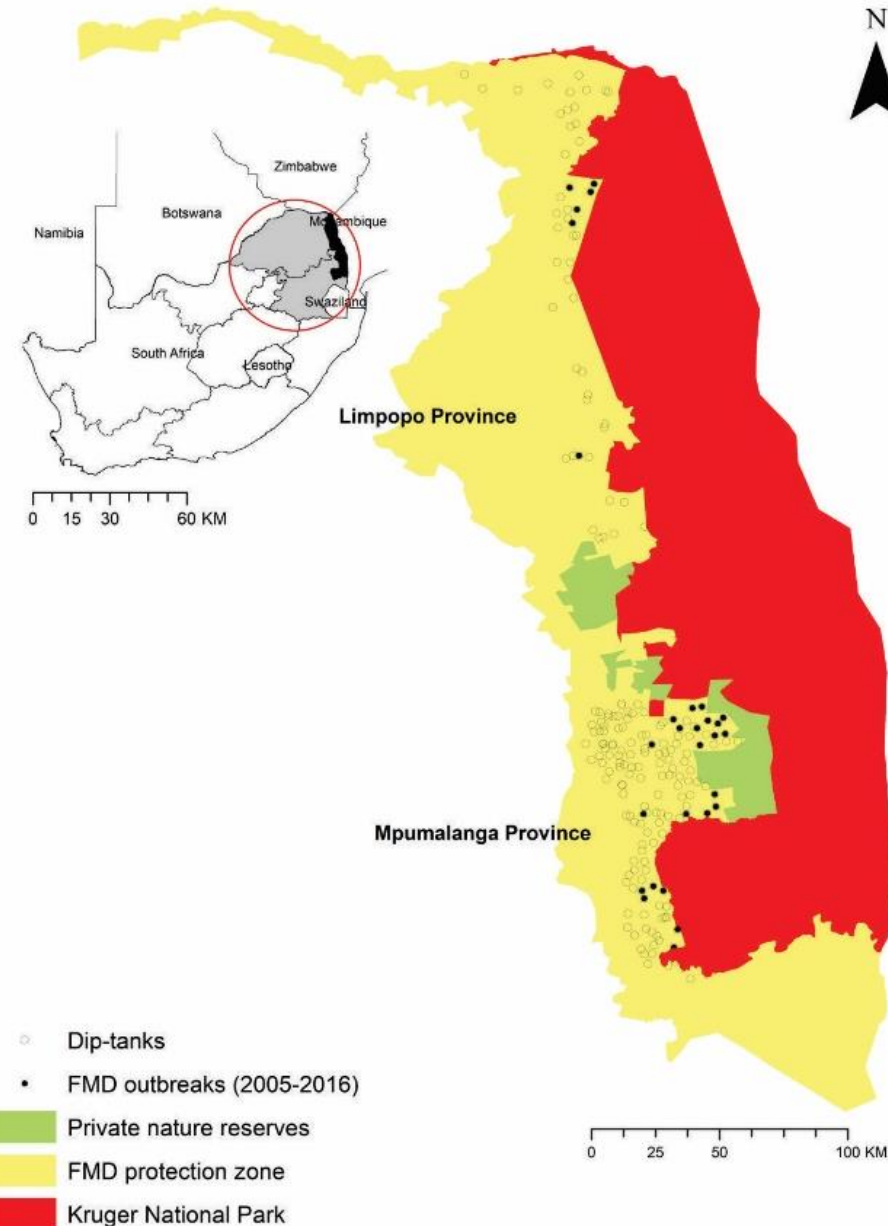




# Présentation descriptive

- La proximité de fonctionnalités spécifiques peut être évaluée
- Est la forme la plus simple d'évaluation de l'exposition spatiale
- Suppose que tous les individus se trouvant à une distance spécifique d'une source ont la même exposition
- Couramment utilisé pour mesurer l'accès aux ressources et l'exposition aux risques environnementaux
- Distance à la clôture anti-maladie

Sirdar MM, Fosgate GT, Bignault D, Mampane RL, Rikhotso O, Du Plessis B, Gumbo B. Répartition spatiale des épidémies de fièvre aphteuse (FMD) en Afrique du Sud (2005-2016). *Trop Anim Santé Prod* 2021;53:376.

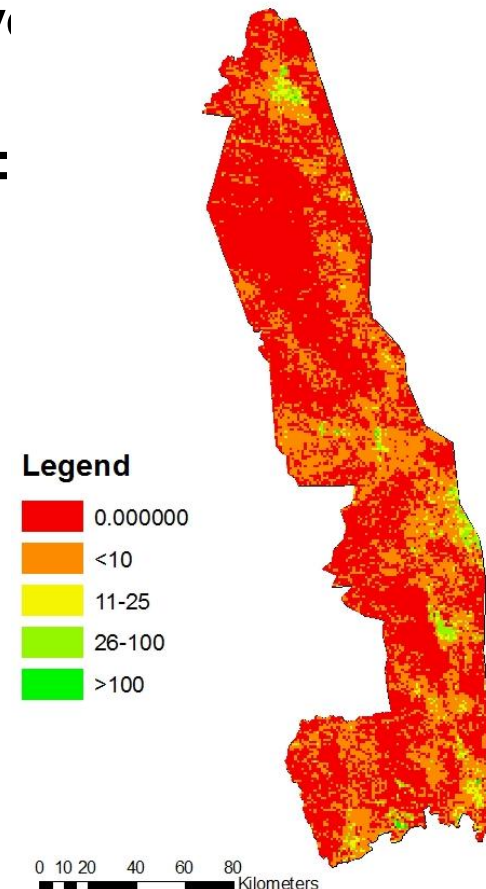




# Cartographie des risques

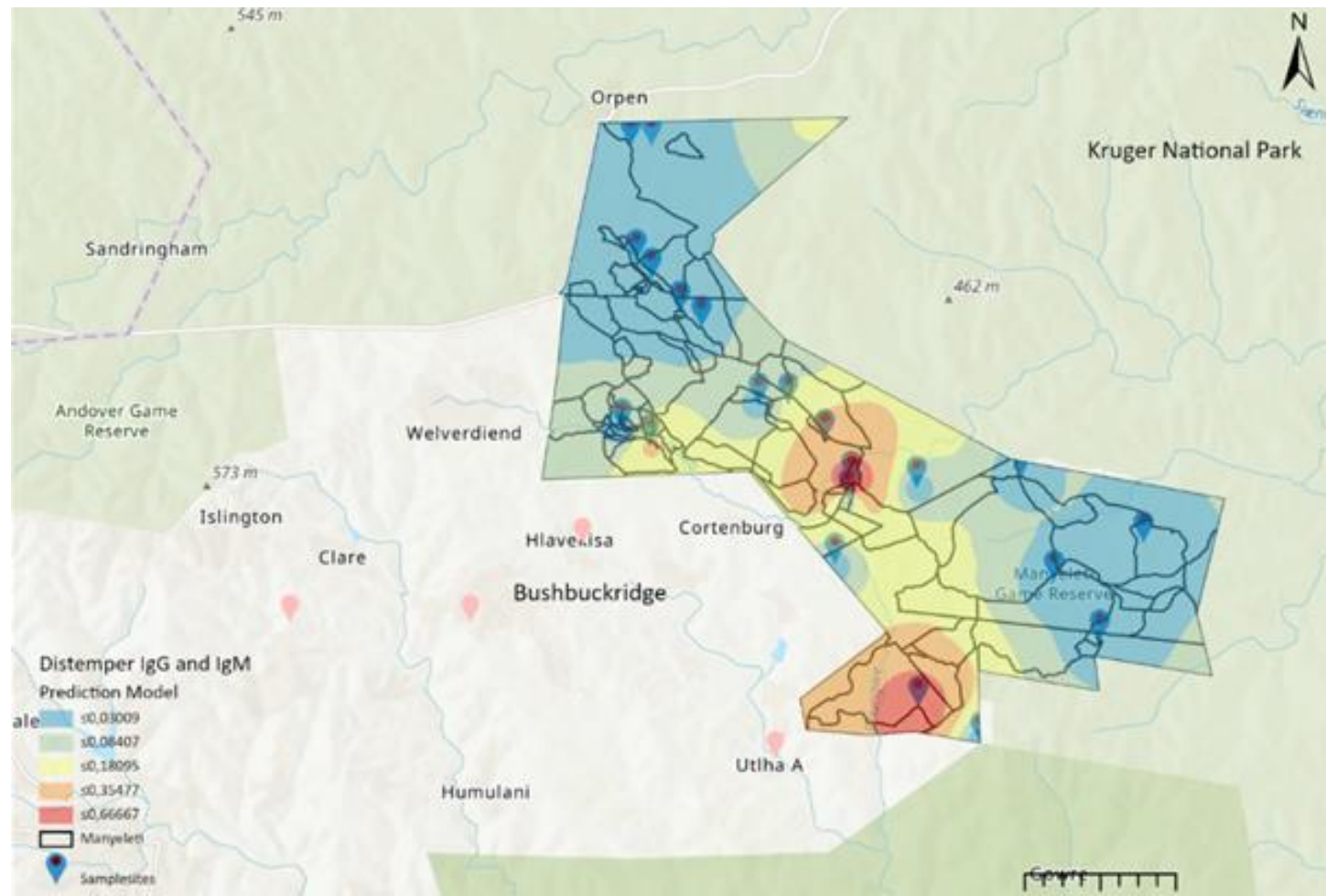
- L'interpolation spatiale est utilisée pour estimer la valeur d'une variable à un emplacement non échantillonné à partir de mesures effectuées sur d'autres sites .
- L'interpolation spatiale est basée sur l'idée que les points proches les uns des autres dans l'espace ont tendance à avoir des valeurs similaires.
- De nombreuses méthodes différentes disponibles :
  - Exact ou approximatif
  - Déterministe ou géostatistique
  - Local ou mondial
  - Graduel ou brusque

Hughes K, Fosgate GT, Budke CM, Ward MP, Kerry R, Ingram B. Modélisation de la répartition spatiale du buffle d'Afrique ( *Syncerus caffer* ) dans le parc national Kruger, Afrique du Sud. *PLOS ONE* 2015; 10(12): e0142000.



# Cartographie des risques

- Techniques déterministes
  - Interpolation polynomiale
  - Pondération inverse de la distance



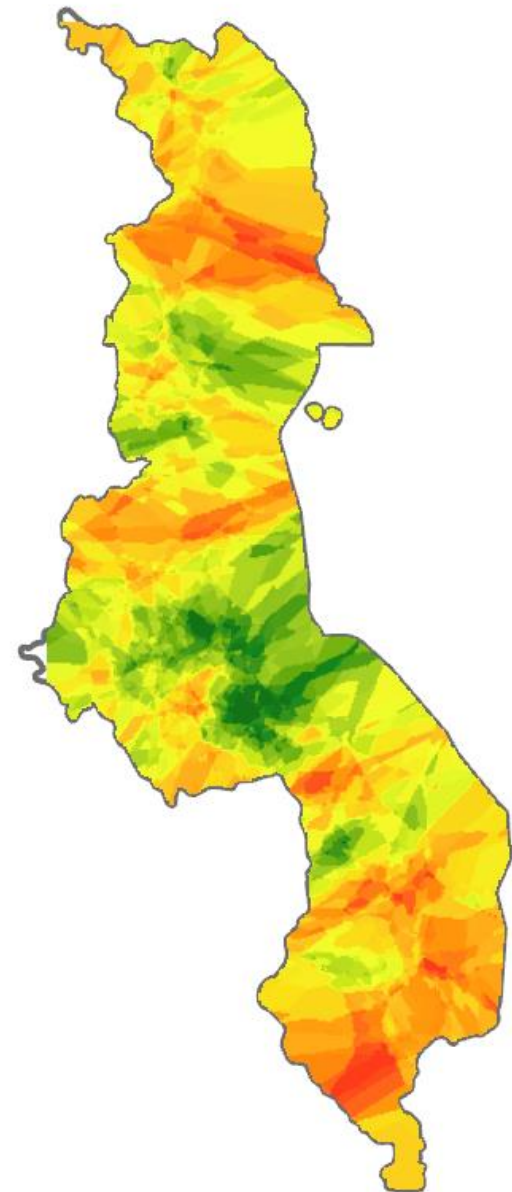
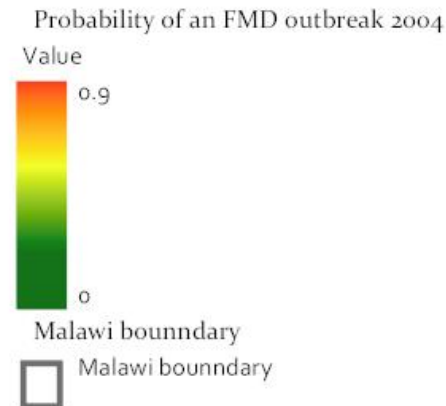
# Cartographie des risques

- **Approches géostatistiques (Krigage)**
  - **Krigage ordinaire**
  - **Krigage simple**
  - **Krigage universel**
  - **Krigage bayésien empirique**

$$d(x, y) = \sum_{i=1}^n w_i d_i$$

$$\gamma(h) = \frac{\sum (z(x) - z(x+h))^2}{2n}$$

Chimera ET, Fosgate GT, Etter EMC, Jemberu WT, Kamwendo G, Njoka P. Modèles spatio-temporels et facteurs de risque de la fièvre aphteuse au Malawi entre 1957 et 2019. *Prev Vet Med* 2022 ;204 : 105639 .





# Détection de clusters

- **Les clusters sont des groupes d'occurrences géographiquement et/ou temporellement délimités, d'une taille et d'une concentration suffisantes, peu susceptibles de se produire par hasard .**
- **Les clusters sont soit liés les uns aux autres par un mécanisme social ou biologique, soit ils ont une relation commune avec un autre événement ou circonstance .**
- **Les animaux présentant des caractéristiques similaires ont tendance à se regrouper et leurs caractéristiques communes expliquent en partie l' association entre la maladie et le lieu.**
- **Les attributs environnementaux influencent des groupes entiers et affectent la maladie au-delà des caractéristiques individuelles globales**

# Détection de clusters

- J'identifie les emplacements, les formes et les tailles des régions spatiales potentiellement anormales
- Déterminer si chacun de ces clusters potentiels est plus susceptible d'être un « vrai » cluster ou une occurrence fortuite
- Est- ce qu'il se passe quelque chose d'inattendu, et si oui, où ?

Existe-t-il des zones présentant un nombre élevé de maladies suggérant une épidémie ou des zones à haut

risque?

Quelle est l'ampleur des maladies attendues dans la

région ?

a- t-il des zones où la maladie est beaucoup plus répandue que prévu ?

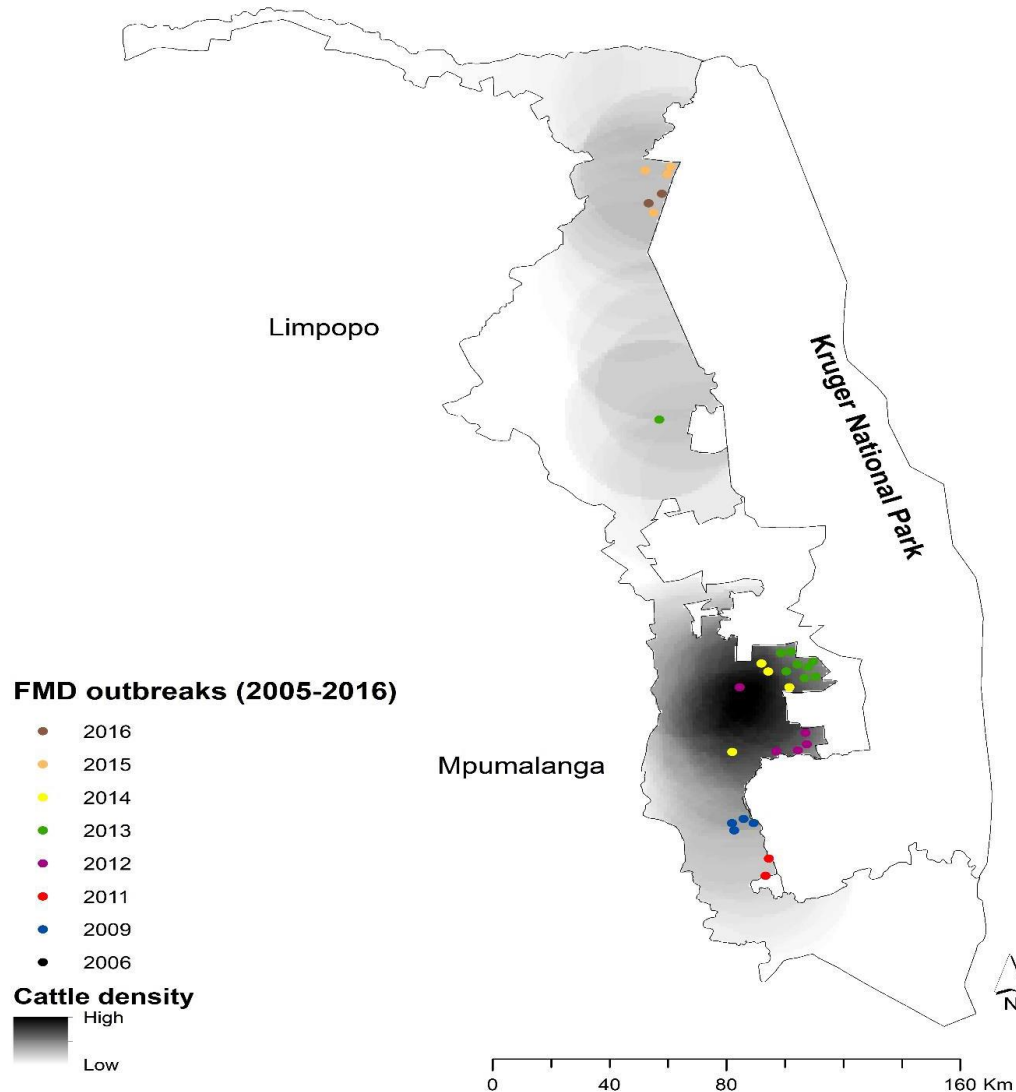
# Détection de clusters

- Les tests de clusters globaux recherchent des clusters spatiaux n'importe où dans une zone d'étude, mais n'identifient pas nécessairement l'endroit où les clusters se produisent, et sont utilisés pour identifier les écarts par rapport au caractère aléatoire spatial lorsque le modèle spatial global est pris en compte .
- Les tests de grappes locaux identifient les endroits où il existe un certain excès/déficit (un point chaud/froid) n'importe où dans une zone d'étude.
- Données temporelles uniquement
- Données spatiales uniquement
- Données de cas uniquement
- Données cas-témoins (ou transversales)
- Prédicteurs continus
- spatiale et temporelle combinée



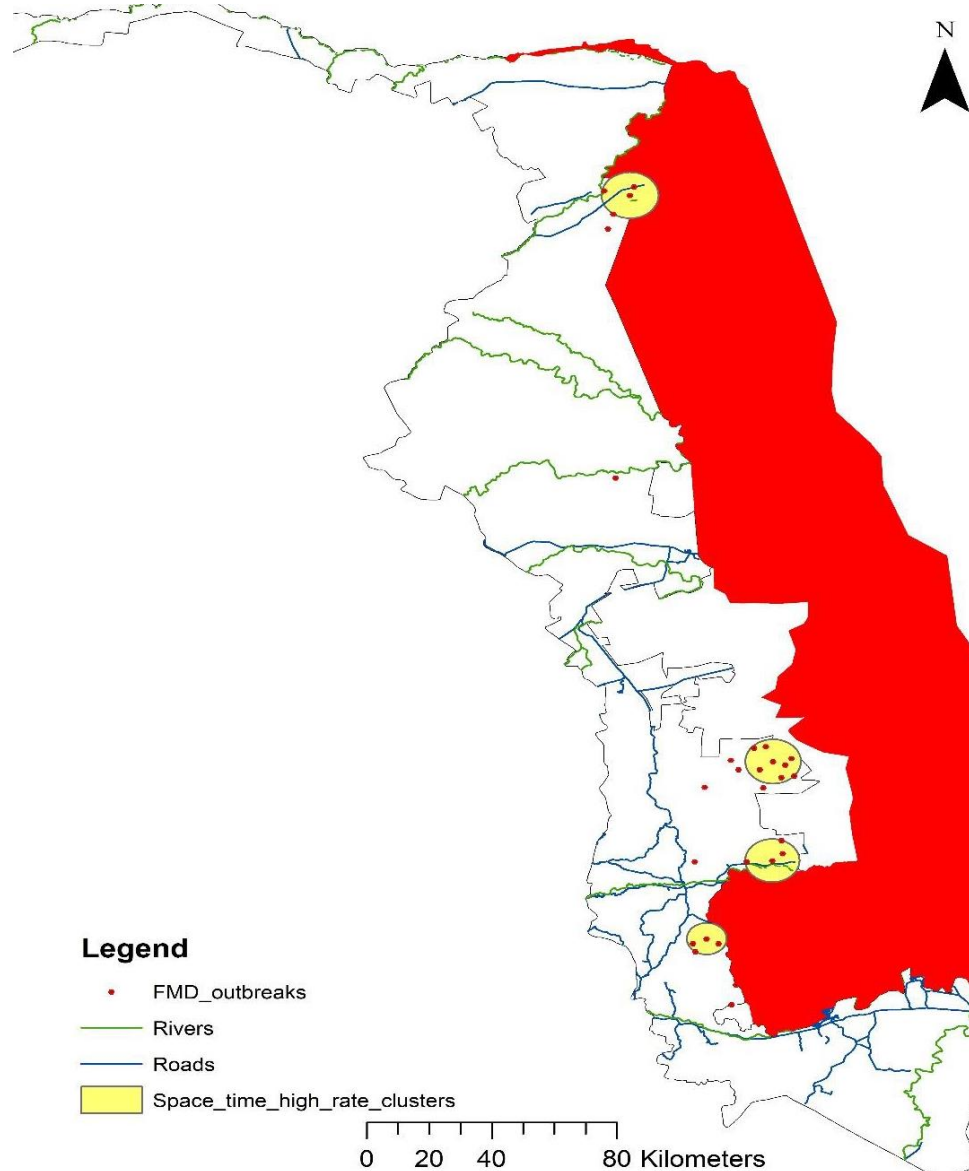
# Est-ce que FMD se regroupe ?

Sirdar MM, Fosgate GT, Blignaut B, Mampane RL, Rikhotso O, Du Plessis B, Gummow B. Répartition spatiale des épidémies de fièvre aphteuse (FMD) en Afrique du Sud (2005-2016). *Trop Anim Santé Prod* 2021;53:376.

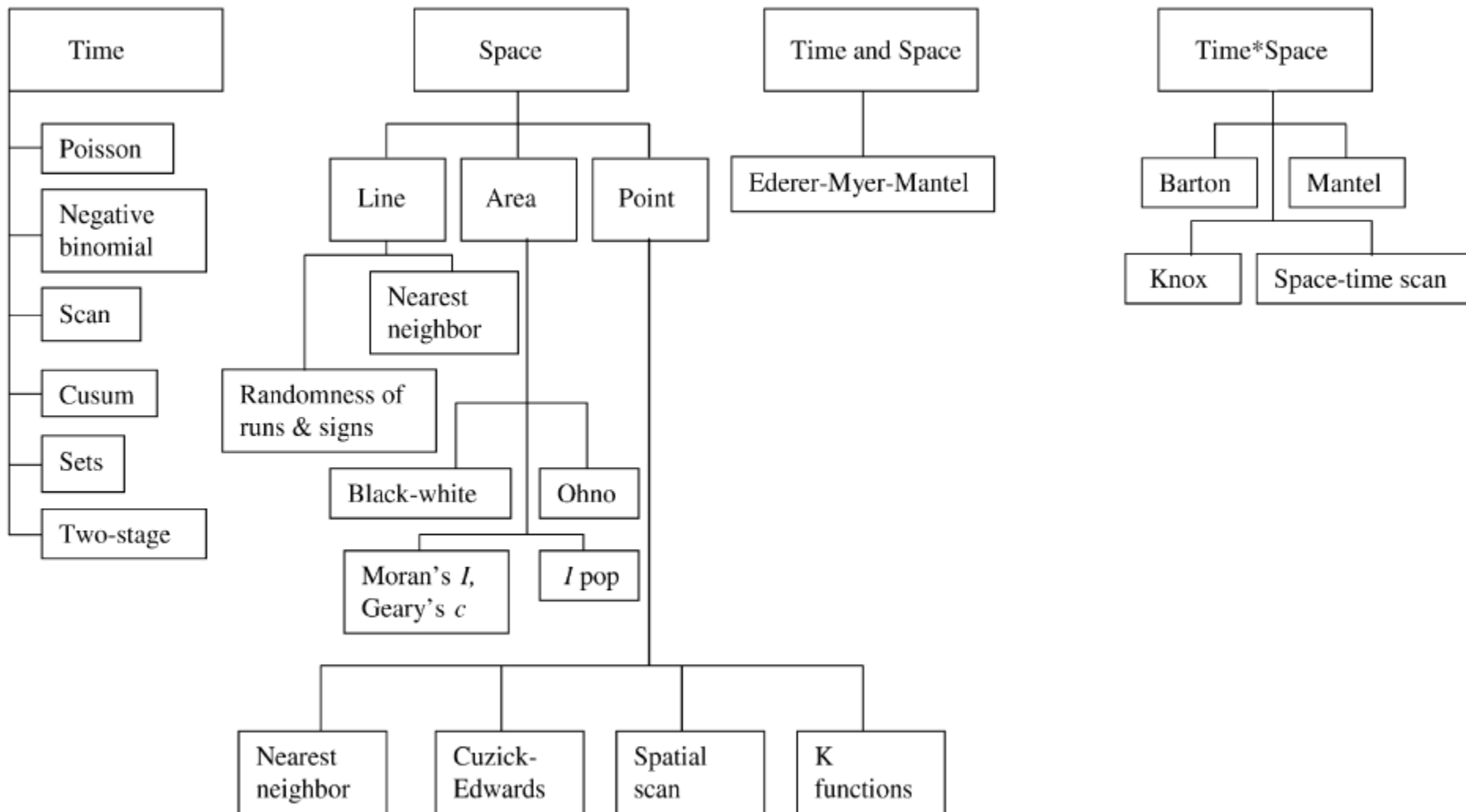


# Est-ce que FMD se regroupe ?

Sirdar MM, Fosgate GT, Blignaut B, Mampane RL, Rikhotso O, Du Plessis B, Gummow B. Répartition spatiale des épidémies de fièvre aphteuse (FMD) en Afrique du Sud (2005-2016). *Trop Anim Santé Prod* 2021;53:376.



# Détection de clusters





# Détection de clusters

Spatial		
Mondial	Locale	Temporospatial
Fonction K de Ripley	Statistiques d'analyse de Kulldurff	Test d'Ederer, Myers et Mantel (EMM)
Test de Cuzick-Edwards	LISA	Statistiques d'analyse de Kulldurff
Moi de Moran		
Ipop		

# Regroupement spatial mondial

- **Fonction k de Ripley**
  - Analyse les données ponctuelles liées aux distances entre les emplacements affectés et non affectés
  - Évalue les distributions groupées et dispersées
- **Test de Cuzick-Edwards**
  - Analyse les données ponctuelles liées aux cas et aux contrôles
  - Identifie le voisin le plus proche plutôt que les distances réelles
  - Peut évaluer le regroupement et les distributions dispersées
- **Moi de Moran**
  - Analyse les données ponctuelles ou surfaciques pour obtenir des résultats quantitatifs
  - Peut évaluer les distributions groupées ou dispersées
- **Ipop**
  - Modification de Moran's I pour tenir compte de la population à risque

# Fonction K de Ripley

- Utilisé pour analyser la configuration spatiale des données de points d'incident
- Résume la dépendance spatiale (regroupement ou dispersion) sur une plage de distances
- La fonction K de Ripley peut être utilisée pour évaluer comment le regroupement ou la dispersion spatiale change lorsque la taille du

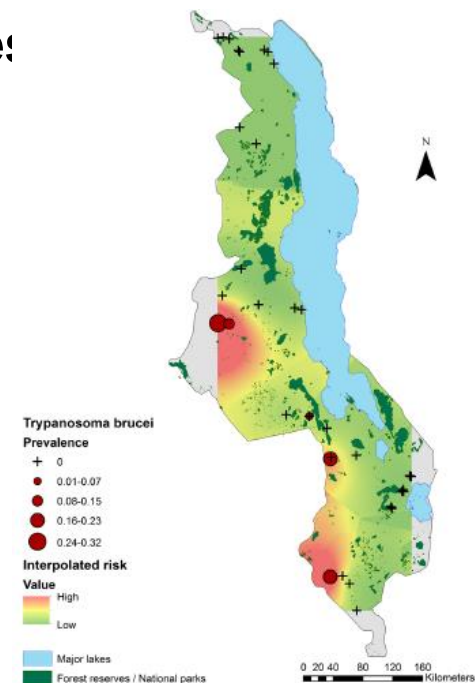
$$\widehat{K}(t) = \lambda^{-1} \sum_{i \neq j} \frac{I(d_{ij} < t)}{n}$$

- Où  $d_{ij}$  est la distance euclidienne entre les  $i$ ème et  $j$ ème points dans un ensemble de données de  $n$  points,  $t$  est le rayon de recherche,  $\lambda$  est la densité moyenne de points (généralement estimée comme  $n/A$ , où  $A$  est l'aire de la région contenant tous les points) et  $I$  est la fonction indicatrice (1 si son opérande est vrai, 0 sinon)



# Test de Cuzick-Edward

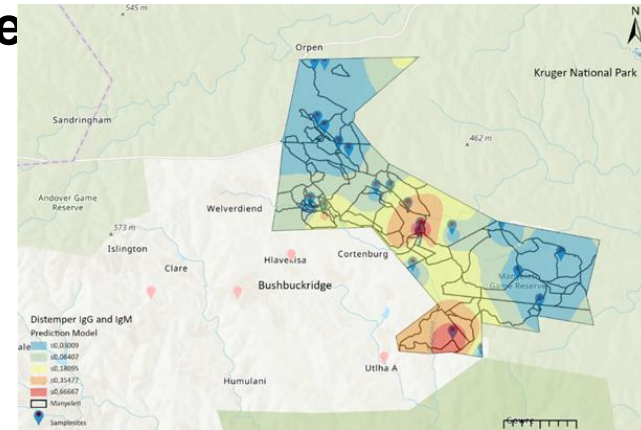
- Le nombre total de paires cas-cas est additionné et comparé au nombre attendu sur la base d'une distribution hypergéométrique
- $E[m] = np = n_1 (n_1 - 1) / n(n - 1)$
- $p$  est la probabilité qu'une paire cas/cas se produise (voisins les plus proches)
- $n$  est le nombre total d'observations (cas et contrôles)
- $n_1$  est le nombre total de cas
- La statistique du test est un test Z typique
- $Z = (m + 0,5 - E[m]) / \sqrt{np(1-p)}$



# Moi de Moran

- Utilisé pour l'évaluation de données continue

Coefficient de Moran (I) = 
$$\frac{n \sum (x_i - \bar{x})(x_j - \bar{x})}{J \sum (x - \bar{x})^2}$$



**n** = nombre de zones étudiées

**J** = nombre total de contigüités

$x_i$  &  $x_j$  = valeurs de zone adjacente (de chaque côté du lien)

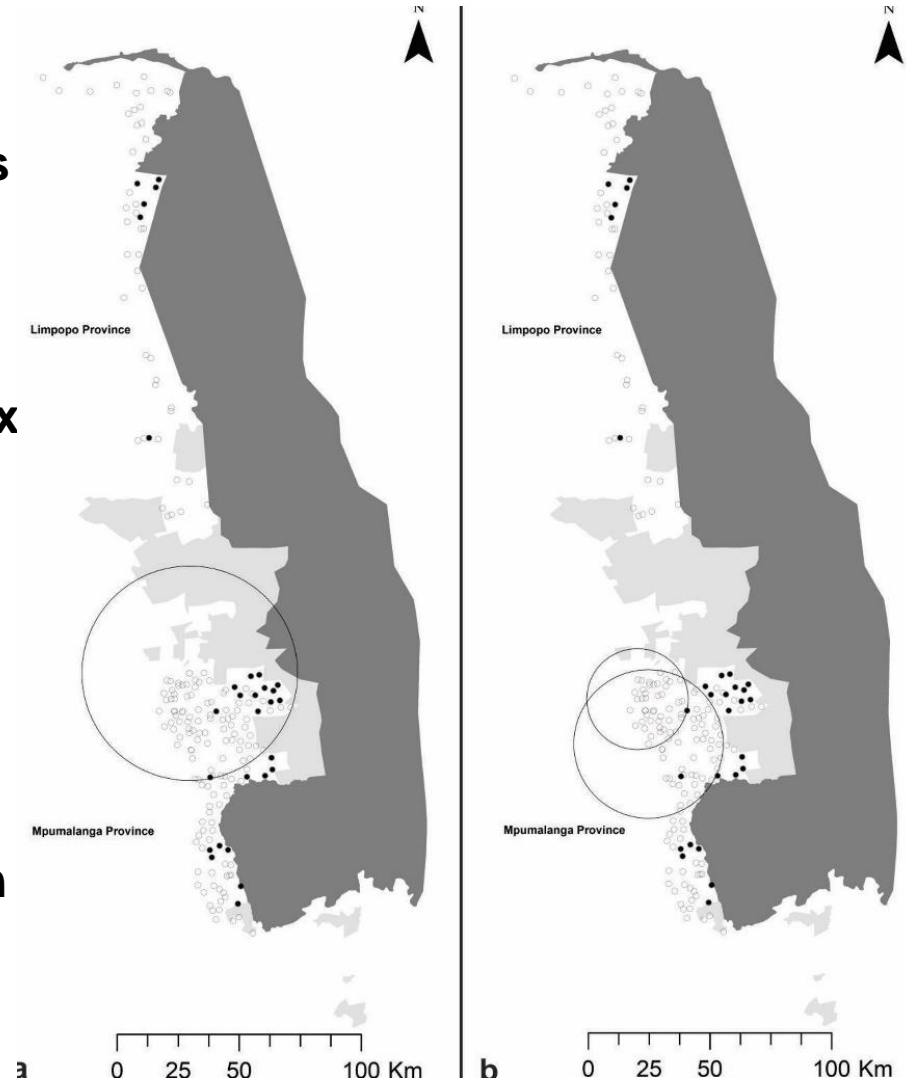
$x$  = valeur de zone &  $\bar{x}$  = moyenne de toutes les valeurs (zones)

$$E_{je} = -1 / n - 1$$

$$\sigma_I = \sqrt{\frac{n[J(n^2 + 3 - 3n) + 3J^2 - n\sum L^2] - k[J(n^2 - n) + 6J^2 - 2n\sum L^2]}{J^2(n-1)(n-2)(n-3)}}$$

# Regroupement spatial local

- **Statistiques d'analyse spatiale**
  - L'approche peut utiliser plusieurs types de données
  - Bernoulli ou données de cas/témoins
  - Poisson pour les données de taux d'incidence
- **Indicateurs locaux d'autocorrélation spatiale (LISA)**
  - Utilise des données quantitatives
  - Modification du I de Moran
  - Identifie les emplacements responsables de l'autocorrélation (clustering ou dispersion)





# Analyse spatiale

- **Pourquoi utiliser une statistique d'analyse**
  - **Nous ne savons pas où les maladies surviendront**
  - **Nous ne connaissons pas leur étendue géographique**
- 1. Obtenir des données pour un ensemble d'emplacements spatiaux  $s_i$**
- 2. Choisissez un ensemble de régions spatiales  $S$  à rechercher**
- 3. Choisissez des modèles de données sous l'hypothèse nulle  $H_0$  (pas de clusters) et des hypothèses alternatives  $H_1(S)$  (cluster dans la région  $S$ ).**
- 4. Dériver une fonction de score  $F(S)$  basée sur  $H_1(S)$  et  $H_0$**
- 5. Trouver les régions les plus anormales (c'est-à-dire les régions  $S$  avec le  $F(S)$  le plus élevé)**
- 6. Déterminer si chacun de ces clusters potentiels est réellement un cluster anormal**

# Analyse spatiale

- Créer une grille régulière ou irrégulière de centroïdes pour la zone d'étude
- Créez un nombre infini de cercles autour de chaque centroïde, avec un rayon compris entre zéro et un maximum de 50 % de la population.
- Pour chaque cercle :
  - Obtenir le nombre réel et attendu de cas à l'intérieur et à l'extérieur du cercle
  - Calculer la fonction de vraisemblance
- Comparez les cercles :
  - Choisissez le cercle avec la fonction la plus probable comme cluster le plus probable
- Inférence:
  - Générer des répliques aléatoires de l'ensemble de données sous l'hypothèse nulle d'absence de clusters (échantillonnage de Monte Carlo )



# Analyse spatiale

- une densité de population inhomogène
- Teste simultanément les clusters de n'importe quelle taille et n'importe quel emplacement à l'aide de fenêtres circulaires à rayon variable
- L'approche prend en compte plusieurs tests
- Il est possible d'inclure des covariables qui pourraient être une source de confusion
- L'approche peut être utilisée pour des données ponctuelles ou agrégées
- Peut analyser des données basées sur plusieurs distributions
  - Bernoulli – cas/témoins ou données transversales
  - Poisson – taux d'incidence
  - Normal – similaire à une analyse LISA pour les données continues (diapositive suivante)
  - Exponentiel – analyse de survie
  - Ordinal – peu courant en médecine vétérinaire mais pourrait être une stratification du cancer

- Indicateurs locaux d'autocorrélation spatiale
- Modification du I global de Moran
- les clusters HH et LL
- Le clustering conventionnel ne parvient pas à préserver la contiguïté

$$MC = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n c_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{1}{\sum_{i=1}^n \sum_{j=1}^n c_{ij}} \frac{\sum_{i=1}^n Z_{Y,i} \sum_{j=1}^n c_{ij} Z_{Y,j}}{(n-1)/n}$$



# Détection de clusters spatio-temporels

- **Test d'Ederer, Myers et Mantel (EMM)**
  - Test unilatéral pour le clustering
  - Sensible aux changements dans la population à risque et donc déconseillé pendant de nombreuses (> 5) périodes
  - Peut être calculé relativement facilement à la main (feuille de calcul)
- **Statistiques d'analyse**
  - Peut être utilisé pour identifier les clusters à haut débit et à faible débit (et non unilatéraux)
  - Peut être utilisé pendant plusieurs périodes
  - Peut être coûteux en calcul et impossible à calculer manuellement
  - Couramment employé en épidémiologie vétérinaire

# Test spatio-temporel EMM

- Regroupement oui/non – impossible de faire la distinction entre les distributions aléatoires et uniformes (dispersées) si aucun regroupement n'est détecté
- Un test de regroupement temporel dans plusieurs séries temporelles simultanément
- Pas utilisé aussi couramment depuis la disponibilité des tests d'analyse spatiale

$$\chi^2 = \frac{[\sum m_i - E(\sum m_i) - 0.5]^2}{\sum V(m_i)}$$

$m_i$  = maximum number of cases in single time period for location  $i$

$\sum m_i$  = sum of  $m_i$  over all areas;  $E(\sum m_i)$  = expected value for sum

$V(m_i)$  = variance of areal maxima; 0.5 is the  $\chi^2$  continuity correction factor

# Analyse spatio-temporelle

- **Utilisez une fenêtre cylindrique, avec la base circulaire représentant l'espace et la hauteur représentant le temps**
- **Pour chaque cylindre**
  - **Obtenez le nombre réel et prévu de caisses à l'intérieur et à l'extérieur du cylindre.**
  - **Calculer la fonction de vraisemblance**
- **Comparer les cylindres**
  - **Sélectionnez le cylindre avec la fonction la plus probable comme cluster le plus probable**
- **Inférence:**
  - **Générer des répliques aléatoires de l'ensemble de données sous l'hypothèse nulle d'absence de clusters (échantillonnage de Monte Carlo )**
  - **Comparez les clusters les plus probables dans des ensembles de données réels et aléatoires (test du rapport de vraisemblance )**

# Permutation espace-temps

- Données de cas uniquement
- Pour chaque cylindre, calculer le nombre attendu de cas conditionné aux totaux marginaux

$$\mu_{st} = \sum_s c_{st} \times \sum_t c_{st} / C$$

- où  $c_{st}$  = nombre de cas au temps  $t$  dans l'emplacement  $s$
- et  $C$  = nombre total de cas
- Calculez ensuite la statistique du test

$$T_{st} = [c_{st} / \mu_{st}]^{c_{st}} \times [(C c_{st}) / (C - \mu_{st})]^{C c_{st}}$$

si  $c_{st} > \mu_{st} = 1$ , sinon

de test  $T = \max_{st} T_{st}$



# Permutation espace-temps

- Générer des répliques aléatoires de l'ensemble de données conditionnées sur les totaux marginaux en permutant les paires d'emplacements spatiaux et de temps
- Comparez les statistiques de test dans des ensembles de données réelles et aléatoires à l'aide des tests d'hypothèse de Monte Carlo :
- $p = \text{rang}(T_{\text{réel}}) / (1 + \text{nombre de répétitions})$
- S'ajuste aux clusters purement géographiques et purement temporels
- Teste simultanément les épidémies de toute taille à n'importe quel endroit à l'aide d'une fenêtre cylindrique à rayon (espace) et hauteur (temps) variables.
- Compte pour plusieurs tests
- Données agrégées ou non agrégées

# Résumé

- **Problèmes d'épidémiologie spatiale décrivant l'apparition de la maladie en termes de localisation géographique**
- **Il existe de nombreuses procédures disponibles pour créer des cartes de risques et évaluer la présence de regroupements.**
- **Certaines méthodes peuvent être réalisées « à la main » tandis que d'autres nécessitent un logiciel spécialisé**
- **Les cartes de risques (interpolation) sont généralement réalisées dans un logiciel SIG**
- **SaTScan est un logiciel gratuit ( [www.satscan.org](http://www.satscan.org) ) couramment utilisé pour étudier le clustering spatial.**



Merci



World Organisation  
for Animal Health  
Founded as OIE



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA  
Faculty of Veterinary Science