Thèse

En vue de l'obtention du grade de

Docteur de l'université de Guyane

Spécialité : *Sciences de la vie et de la santé*

Présentée et soutenue publiquement le *13 octobre 2020* par :

SOUSHIETA JAGADESH

Biogeography of Emerging Infectious Diseases, v.1.10

JURY

Pierre COUPPIÉ	HDR	Président du Jury
Antoine ADENIS	PU, DR, MCF ou CR	Examinateur
Maylis DOUINE	PU, DR, MCF ou CR	Examinateur
Benjamin ROCHE	HDR	Rapporteur
Tom FLETCHER	HDR	Rapporteur
Rodolphe GOZLAN	HDR	Directeur de Thèse
Mathieu NACHER	HDR	Co-Directeur de Thèse

Ecole Doctorale n°587 : *Diversités, Santé et développement en Amazonie*

Biogeography of Emerging Infectious Diseases

Contents

List of Figures

Acknowledgements

The Université de Guyane funded this thesis. I thank the members of the Jury: Pierre Couppié , Benjamin Roche, Tom Fletcher, Antoine Adenis, Maylis Douine, Marine Combe and my PhD supervisors.

I am incredibly thankful to the various people influential in the smooth progression and completion of this thesis. A few, however, deserve special mention:

My supervisors – Rudy Gozlan and Mathieu Nacher for their brilliant and rapid feedback, instilling confidence whilst facing publication demons, and being timekeepers over a three-year marathon.

My co-supervisor- Marine Combe for being a cheery inspiration for women in STEM, friend and confidant.

My thesis collaborators – Pierre Couppié and Loïc Epelboin for being great sports in sharing their curated databases and their constructive criticism.

Albert – for being the best neighbor, the ABBA, and tea time conversations.

Pierre – for the Renaud, the rigorous French lessons, and sharing the love of peanut butter.

Richard – simply for being Richard.

The ORYs – for being the most incredible and accepting family

A special shout to Benjamin and Marie-Sylvie – for helping me translate portions of my thesis.

My dad, mum & Gaju – for the support and love at all spatial, emotional, and financial scales.

Benjamin – partner in geekiness for extensive otolith quests, adventures, and the future.

Abbreviations and Definitions

Abbreviations

A	
AIDS	Acquired Immunodeficiency Syndrome
AM	Adaptive Metropolis
ART	Anti-Retroviral Therapy
AUC	Area Under Curve
B	
BAME	black, Asian, and minority ethnic communities
BPD	Blueprint list of Priority Diseases
BU	Buruli Ulcer
C	
CCHF	Crimean Congo Hemorrhagic Fever
CDC	Center for Disease Control
CI	Confidence Interval
CNIL	Commission Nationale Informatique et Libertés
COREVIH	Regional Coordination for the fight against HIV
COVID-19	Coronavirus disease 2019
CRS	Coordinate Reference System
D	
DEM	Digital Elevation Model
DIC	Deviance Information Criterion
E	
EID	Emerging Infectious Disease
EMD	Emergence des Maladies Infectieuses
EVD	Ebola virus disease
F	

FAO	Food and Agriculture Organisation
FG	French Guiana
FHDH	French Hospital Database for HIV
G	
GLM	Generalized Liner Model
H	
HAART	Highly Active Anti-Retroviral Therapy
HAC	Hierarchical Agglomerative Clustering
HC	Histoplasma capsulatum var. capsulatum
HIV	Human Immunodeficiency Virus
I	
IDW	Inverse Distance Weighted
INSERM	Institut National de la Recherche Médicale
IUCN	International Union for Conservation of Nature
L	
LF	Lassa Fever
M	
MAT	Microscopic Agglutination Test
MAXENT	Maximum entropy
MCMC	Markov Chain Monte Carlo
MERS-COV	Middle East Respiratory Syndrome Coronavirus
MVD	Marburg Viral Disease
P	
PHLA	People Living with HIV/AIDS
PPRI	Plan de Prevention du Risque Inondation
Q	
QPPV	le Quartier Prioritaire de la Politique de la Ville

R	
ROC	Receiver Operating Characteristic
RVF	Rift Valley Fever
S	
SARS	Severe Acute Respiratory Syndrome
SARS-COV-1	Severe Acute Respiratory Syndrome Coronavirus 1
SARS-COV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SDM	Species Distribution Modelling
SOM	Self-organizing maps
SRTM	Shuttle Radar Topography Mission
SSDM	Stacked Species Distribution Modelling
SVM	Support Vector Machine
T	
TWI	Topological Wetness Index
U	
USAID	United States Agency for International Development
USGS	United States Geological Survey
W	
WHO	World Health Organization

Definitions and concepts

Biogeography

Biogeography, the study of the geographic distribution of plants, animals, and other forms of life. It is concerned not only with habitation patterns but also with the factors responsible for variations in distribution.

Disease biogeography is the study of the geographic distribution of infectious diseases.

Reference: Escobar LE, Craft ME. Advances and limitations of disease biogeography using ecological niche modeling. Front Microbiol (2016) 7:1174. doi:10.3389/fmicb.2016.01174

Dilution effect

The 'dilution effect' hypothesis suggests that the net effects of biodiversity (including host and non-host species) reduce the risk of certain diseases in ecological communities.

Reference: Keesing, F., Holt, R. D. and Ostfeld, R. S. (2006). Effects of species diversity on disease risk. Ecol. Lett. 9, 485-498.

Emerging Infectious Diseases

'Emerging infectious diseases' are defined as 'those whose incidence in humans has increased within the past two decades or threatens to increase in the near future. Emergence may be due to the spread of a new agent, to the recognition of an infection that has been present in the population but has gone undetected, or to the realization that an established disease has an infectious origin. Emergence may also be used to describe the reappearance (or re-emergence) of a known infection after a decline in incidence'

Reference: Lederberg J., Shope R.E., Oaks S.C. National Academy Press; Washington, D.C.: 1992. Institute of Medicine (U.S.). Committee on emerging microbial threats to health. Emerging infections: microbial threats to health in the United States

Neighborhood effect

Neighborhood effects refer to (a) the processes by which various neighborhood conditions influence the well-being of residents collectively or individually or (b) outcomes associated with negative neighborhood conditions.

Reference: Roosa, M. W. and White, R. M. B. (2014) 'Neighborhood Effects', in Encyclopedia of Quality of Life and Well-Being Research. Springer Netherlands, pp. 4328–4331.

Remote sensing

Remote sensing is the science and practice of acquiring information about an object without actually coming into contact with it. In terms more appropriate for this thesis, remote sensing is a technology for sampling reflected and emitted electromagnetic (EM) radiation from the Earth's terrestrial and aquatic ecosystems and atmosphere.

Refernce: Horning N. Remote Sensing. In: Encyclopedia of Ecology, Five-Volume Set. Elsevier Inc.; 2008. p. 2986–94.

Spatial Bayesian model

A key concept behind spatial models is based on following objectives: scale of the ecological process, estimation and inference of parameter estimates, model specification and comparison, and prediction.

The Bayesian approach to spatial modelling fulfills the above objectives and has two main advantages over the conventional frequentist methods. First, it is not based on approximation (like the penalized quasilikelihood) and thus provides exact results even for binary responses. Second, it correctly propagates the uncertainties linked to the estimation of the variogram parameters.

Reference: Korner-Nievergelt F, Roth T, von Felten S, Guélat J, Almasi B, Korner-Nievergelt P. Chapter 6 - Assessing Model Assumptions: Residual Analysis. Bayesian Data Anal Ecol Using Linear Model with R, BUGS, STAN. 2015;75–94. Available from: http://www.sciencedirect.com/science/article/pii/B978012801370000006X

Banerjee, S., B.P. Carlin, and A.E. Gelfand. 2003. Hierarchical Modeling and Analysis for Spatial Data. Boca Raton: Chapman & Hall/CRC.

Spatial dependence

Spatial dependence is defined as "the property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations"

Reference: Legendre, P., Legendre, L. and Legendre, L. (1998) Numerical ecology. Elsevier.

Spillover transmission

Spillover transmission is defined as the processes that enable a pathogen from a vertebrate animal to establish infection in a human.

Reference: Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., & Lloyd-Smith, J. O. (2017). Pathways to zoonotic spillover. Nature reviews. Microbiology, 15(8), 502–510. https://doi.org/10.1038/nrmicro.2017.45

SOM

Self-organizing maps (SOM) are feed-forward networks that use an unsupervised learning approach through a process called self-organization. When an input pattern is fed to the network, the units in the output layer compete with each other, and the winning output unit is typically the one whose incoming connection weights are closest to the input pattern (Euclidean distance). Thus, the input is presented, and each output unit computes its closeness or match score to the input pattern.

Reference: Sajja PS, Akerkar R. Bio-Inspired Models for Semantic Web. In: Swarm Intelligence and Bio-Inspired Computation. Elsevier Inc.; 2013. p. 273–94.

Zoonosis

A zoonosis is any disease or infection that is naturally transmissible from vertebrate animals to humans. Animals thus play an essential role in maintaining zoonotic infections in nature. Zoonoses may be bacterial, viral, or parasitic, or may involve unconventional agents.

Reference: WHO | Zoonoses. WHO. 2017

Résumé

La récente pandémie de Covid19 nous rappelle, si cela était encore nécessaire, que la propagation des maladies infectieuses ignore les frontières géographiques. Les changements combinés de biodiversité locale et l'utilisation des terres, l'augmentation de la connectivité internationale par le transport et le commerce ainsi que la menace imminente du changement climatique a accru le risque d'émergence et de réémergence des maladies infectieuses (EMI). Jusqu'à présent la réponse des politiques de santé publique a été la surveillance passive sans toutefois s'avérer réellement efficace dans la prévention et le contrôle des épidémies. Le choix qui a été fait ici est celui d'une nouvelle approche anticipative, par identification des zones à haut risques d'EMI en se basant sur la détection des facteurs environnementaux les plus favorisant. Parmi ces facteurs on trouve la conversion des terres, la diminution drastique de la biodiversité ou encore le changement climatique. Ainsi la méthode biogéographique a permis d'étudier et d'analyser les EMI à travers différents groupes de taxons de pathogènes comme les bactéries, les virus, les protozoaires et les champignons. L'étude a été portée globalement, ainsi que localement, en Guyane Française, un territoire français d'outre-mer situé en Amérique du Sud. Dans les deux cas, à travers les différents groupes de pathogènes, les risques d'inondation, les récentes conversions de parcelles de forêts en terres agro-minières et l'augmentation du minimum de température due au changement climatique se sont avérés être des facteurs significatifs dans l'émergence globale et locale des maladies infectieuses étudiées. Les principaux résultats de cette thèse sont les suivantes :

1. Une approche biogéographique de modélisation de la distribution des EMI en utilisant les bases de données existantes sur les cas cliniques, l'imagerie satellite et un modèle statistique non conventionnel est efficace pour détecter précocement les régions à risque, permettre d'améliorer la prévention, et contrôler leur diffusion.
2. Il est possible d'anticiper les EMI en identifiant et en gérant précocement les facteurs favorisant ayant un lien direct avec l'anthropisation de l'environnement.

Abstract

The COVID-19 pandemic highlights that the spread of infectious diseases goes beyond geographical boundaries. Simultaneous changes in local biodiversity and land use, the increasing international connectivity through human transport and trade and the imminent threat of climate change have increased the risk of the emergence and reemergence of infectious diseases. The current public health response to emerging infectious diseases (EID) by passive surveillance has proven largely ineffective in preventing and controlling disease outbreaks. The way toward is to "get ahead of the curve" by identifying potential hotspots of disease emergence and detecting the environmental triggers such as land transformation, biodiversity loss and climate change. I used a biogeographic approach to study and analyze disease emergence across different taxonomic pathogen groups such as bacterial, viral, protozoal and fungal, globally and in French Guiana, a French Overseas territory located in South America. I found that regions at risk of floods, recent conversion of forest to agricultural lands and increasing minimum temperature (i.e. temperature at night) caused by climate change were drivers for disease emergence locally and globally across the different pathogen groups. The main findings of the PhD thesis are the following:

1. Biogeographic approach to mapping the distribution of EIDs with using existing human cases data, remote sensing imagery and unconventional statistical models is effective to "get ahead of the curve" in the detection of regions at risk and the management of EIDs.
2. EIDs are not unprecedented but predictable by identifying and managing the triggers of disease emergence, which have a direct link with the anthropization of the environment.

1.Introduction

Global Biogeography of Viral Zoonoses

1.1. Introduction (Français)

La propagation des maladies infectieuses n'a pas de limites spatiales. L'émergence du COVID19 a démontré que l'émergence d'une maladie dans une zone du monde pouvaits en quelques mois seulement évoluer en pandémie du fait du caractère international et toujours plus rapide de la mobilité humaine et commerciale. La préoccupation grandissante dont fait l'objet la rapide propagation de l'émergence des maladies infectieuses (EMI) et l'augmentation du risque d'émergence des pathogènes (Jones *et al*,. 2008, Fisher *et al*., 2012) a mené à un regain d'intérêt au sein de la communauté scientifique, la population, et les autorités nationales et internationales, afin d'investir dans le contrôle et la prévention des maladies émergentes à toutes les échelles.

Jusqu'à maintenant la réponse des pouvoirs de santé publique à l'émergence des maladies a été approximative et réactionnaire. Les facteurs locaux déclencheurs tel que les marchés de produits frais sont surveillés de près, mais les autres facteurs environnementaux comme le changement climatique et la fragmentation des terres découlant d'un accroissement global de la demande de produits issus de l'agriculture sont souvent négligés et ce alors même que les études récentes ont montré que la pression anthropogénique sur les terres utilisées pour l'expansion agricole et le bétail favorisait les EMI.

Dans les îles de Sumatra par exemple, la migration de chauves-souris frugivores suite aux feux de forêts déclenchés en vue de créer des surfaces de culture pour l'huile de palme a mené à l'émergence du Nipah (NiV). Ce virus qui s'est propagé à travers des cochons infectés provoque une encéphalite infectieuse aiguë avec un taux élevé de mortalité et a touché principalement des fermiers et des ouvriers d'abattoirs en Malaysie. (Lo and Rota, 2008; Daszak et al., 2013). Les mesures de prévention censées être prises sont finalement plutôt réactive et de contrôle, faisant suite à une émergence déjà établie, et sont souvent accompagnées d'une surveillance passive des EMI qui reste inefficace. Il devient donc nécessaire d'anticiper autant que possible, en identifiant et contrôlant les facteurs de maladies infectieuses, et en détectant les régions à risque de futures émergences..

Lors des quatre dernières décennies plus de 70% des émergences de maladies infectieuses étaient des zoonoses (Taylor, Latham and Woolhouse, 2001; Jones et al., 2008). Puisque la plupart des pathogènes émergents utilisent des réservoirs animaux non humains, des hôtes intermédiaires, et/ou des espèces vectrices dans leur cycle de transmission, les interactions locales humain-animal-écosystème jouent un rôle majeur dans la transmission des pathogènes de l'environnement à l'homme et par conséquent dans l'émergence de maladies infectieuse au sein de la population humaine. L'augmentation rapide de la population humaine, les 'activités économiques, et le changement climatique pèsent indéniablement sur ces interactions.

Les quatre principaux facteurs d'émergence des EMI

L'impact de l'anthropisation sur l'environnement est sans précédent (McMichael and Butler, 2011). Tous les jours nous voyons disparaitre environ 32 000 hectares de forêt tropicale, et environ 32 000 autres sont lourdement dégradés (Measuring the Daily Destruction of the World's Rainforests - Scientific American, no date). Parallèlement à ces pertes et dégradations, 135 espèces de macro-organismes (animaux et végétaux) disparaissent quotidiennement. Le changement climatique, la transformation des terres, la perte de biodiversité et le commerce d'animaux peuvent être considérés comme des facteurs prépondérants à l'origine de l'émergence des maladies infectieuses.

Les données actuelles confirment que le changement climatique a contribué de manière significative à l'émergence des maladies infectieuses, notamment dans des régions jusqu'à maintenant présent préservées, en impactant des un spectre de latitudes et des 'altitudes toujours plus largealtitudes toujours plus grandes. Par ailleurs dans les régions qui étaient déjà touchées l'effet de la saisonnalité qui pouvait jusqu'à maintenant encadrer et donc limiter les plages périodes d'émergence tend à s'estomper (McMichael, Woodruff and Hales, 2006). On reporte aujourd'hui des maladies associées aux crues et inondations comme la leptospirose dans des environnements aussi divers que l'Inde, l'Argentine, l'Australie, le Brésil ou encore l'Italie (Lau et al., 2010).

Déforestation, urbanisation, agriculture et autres formes de transformation des terres prennent part également dans l'augmentation du risque d'EMI. L'augmentation de la population dans les grandes zones urbaines n'encourage pas à une production de nourriture durable et accentue les inégalités dans l'accès et la qualité des soins favorisant ainsi les conditions de propagation. (Rodwin and Gusmano, 2002). La déforestation, en augmentant la probabilité de contacts directs et indirects des humains avec les petits mammifères et les oiseaux accroit les risques d'émergence de zoonose. On constate aussi que la déforestation en vue de l'exploitation agricole, l'élevage de bétail ou encore la culture de riz sont en association directe avec l'émergence des maladies (Daszak *et al.*, 2013; Jagadesh, Combe, Couppié, Le Turnier, *et al.*, 2019).

Localement, la diminution de biodiversité des prédateurs due à leur déclin ou encore à leur migration impacte directement les populations de petits mammifères et d'insectes vecteurs dans des chaines d'interactions complexes. En Amérique du Nord, l'extinction du pigeon voyageur a rendu disponible une grande quantité de glands jusqu'alors consommé en grande partie par ces oiseaux. La conséquence directe a été une augmentation rapide des rongeurs qui les consomment alors même qu'ils ont la particularité d'être le principal réservoir de la maladie de Lyme via les tiques qu'ils transportent. (Blockstein, 1998). La consommation de viande sauvage et le commerce d'animaux découlant de la demande croissante en protéine animale provoque des changements significatifs dans l'interface homme-animal réservoir (Wolfe et al., 2005). L'épidémie de SRAS de 2002 s'est propagée à l'homme à travers la consommation de civette palmiste à masque et de chien martre, tous deux intensivement élevés en chine à des fins alimentaires. On pense savoir que le patient zéro de cette épidémie était un cuisinier. La chauve-souris fer à cheval s'est avéré être l'hôte originel du SARS-CoV 1 avec la civette palmiste à masque et le chien martre comme hôtes intermédiaires du virus. (Li et al., 2005; Wang and Eaton, 2007).

Les régions tropicales comme zones à haut risque d'EID

Les régions qui risquent le plus de subir les conséquences de ces changements sont les tropiques et plus particulièrement les pays en voie de développement qui ont un bagage déjà existant en matière de maladie infectieuse. Aujourd'hui, presque 70% des évènements météorologiques extrêmes frappent les régions d'Asie, du Pacifique, d'Afrique ou du Moyen Orient (Costello et al., 2009). Dans les régions tempérées l'augmentation des températures sur l'ensemble de l'année a aussi mené à une extension significative des maladies émergentes, particulièrement celle ayant comme vecteur de transmission des invertébrés, comme la dengue, le chikungunya et la CCHF (Vescio et al., 2012; Tesla et al., 2018). Cependant, l'inéquitable disponibilité des tests, vaccins et médicaments qui permettraient de détecter, prévenir et contrôler les maladies infectieuses, associés à un système de services de santé publique réduit et de faibles revenus conduit inexorablement à une charge disproportionnée de souffrance humaine et de décès dus aux EMI dans les pays en voie de développement.

Plus d'un milliard de personnes, considérés comme population pauvre, vit dans des régions à haut risque environnemental et social où la prévalence des maladies infectieuses est haute, alors même que ces régions ne sont à l'origine que de 3% de l'empreinte carbone planétaire. Les perturbations environnementales dues aux activités humaines, nous obligent à considérer l'importance des facteurs socio-économique dans l'épidémiologie des EMI (Schneider and Machado, 2018). L'exode rural a résulté en une pauvreté structurelle dans des communautés où peuvent se cumuler une hygiène inadéquate, un accès limité aux services de soins et de la sous-nutrition, favorisant ainsi la transmission et la persistance d'un large spectre de maladies infectieuses. (Farmer et al., 2006). Les facteurs socio-économiques tel que l'inégalité des soins, la pauvreté, l'urbanisation, la stigmatisation sociale, les handicaps et la migration tissent des liens complexes dans l'épidémiologie des EMI. (Anthony J McMichael and Ulisses Confalonieri, no date). Les changements climatiques influencent de nombreux facteurs socio-économiques dans les populations les plus vulnérables,

comme la migration, la pauvreté et l'augmentation des pénuries, augmentant ainsi les inégalités déjà existantes.

Prendre les devants

La pandémie de COVID-19, actuellement en cours au moment d'écrire cette thèse, amende l'idée aujourd'hui incontournable que l'on ne peut plus et que l'on ne doit plus se permettre d'attendre l'heureuse découverte d'un vaccin ou d'un médicament miracle lors d'une pandémie. A l'ère de l'augmentation des EMI, nous avons plus que jamais besoin de changement ce qui implique un nouveau paradigme qui intégrerait la centralité fondamentale de l'environnement naturel et son rôle crucial de tampon dans l'émergence des maladies infectieuses émergentes. La prédiction des émergences est essentielle afin de prendre les devants. Les modèles et les prévisions mathématiques rendent possible une estimation rapide permettant d'assister le contrôle et les efforts de prévention potentielle lorsque le temps pour les études épidémiologiques est réduit (Siettos and Russo, 2013). La cartographie des EMI et la prédiction des futures émergences, rendus possible grâce à une résolution de l'imagerie satellitaire toujours plus pointue, permet de déplacer le nouveau paradigme vers une identification précoce des facteurs environnementaux potentiels plutôt qu'une surveillance conventionnelle des maladies infectieuses.

Le but de cette thèse était de : (1) comprendre les dépendances spatiales et la distribution des EMI en utilisant une approche biogéographique ; (2) utiliser les bases de données existantes ainsi que les données d'image satellitaire libre d'accès et valider leur efficacité et leur précision dans la prédiction des EMI; (3) identifier les facteurs environnementaux et socio-économiques des EMI à travers différents groupes de pathogènes, globalement et plus localement en Guyane Française (FG). Les objectifs de recherche de ma thèse étaient les suivants :

1. Cartographier la distribution des EMI de différents groupes taxonomiques de pathogènes comme les bactéries, les virus, les champignons et les protozoaires en utilisant les bases de données existantes que ce soit globalement ou pour une région tropicale d'intérêt.
2. Identifier les facteurs environnementaux et socio-économiques des maladies émergentes en utilisant les bases de données en libre accès des images satellitaires, les cartographies climatiques et les autres trames disponibles (raster).
3. Evaluer les modèles prédictifs de risque pour les maladies émergentes et identifier les zones à risque des EMI étudiées.
4. Comparer la distribution, les facteurs influençant l'émergence et la diffusion des EMI à travers différents groupes de pathogènes.

L'approche biogéographique des maladies infectieuses est une nouvelle méthode basée sur la cartographie des évènements épidémiologiques, la distribution des facteurs de maladies, les hôtes réservoirs ou intermédiaires et l'identification des facteurs influençant leur répartition spatiale et temporelle. Cette approche ne procure pas seulement une analyse descriptive de la dynamique de transmission mais aussi des mesures quantitatives des associations entre les maladies et leurs facteurs ainsi que des prédictions en vue des futures émergences d'épidémie. (Escobar and Craft, 2016).

Pertinence de la localisation de l'étude

La Guyane Française est un territoire français d'outre-mer (3.9339° N, 53.1258° W) en Amérique du Sud (*Figure 1.1*). Son choix se justifie pour notre étude car c'est une région tropicale comprenant une population concentrée sur le littoral et plus de 95% de sa superficie est classée comme forêt primaire, formant ainsi une portion majeure de l'importante biodiversité du plateau des Guyanes, menacée par l'anthropisation due à l'augmentation rapide de la population. La couverture terrestre inclue aussi différents milieux bien distincts comme les savanes, les zones humides ou encore les mangroves côtières.

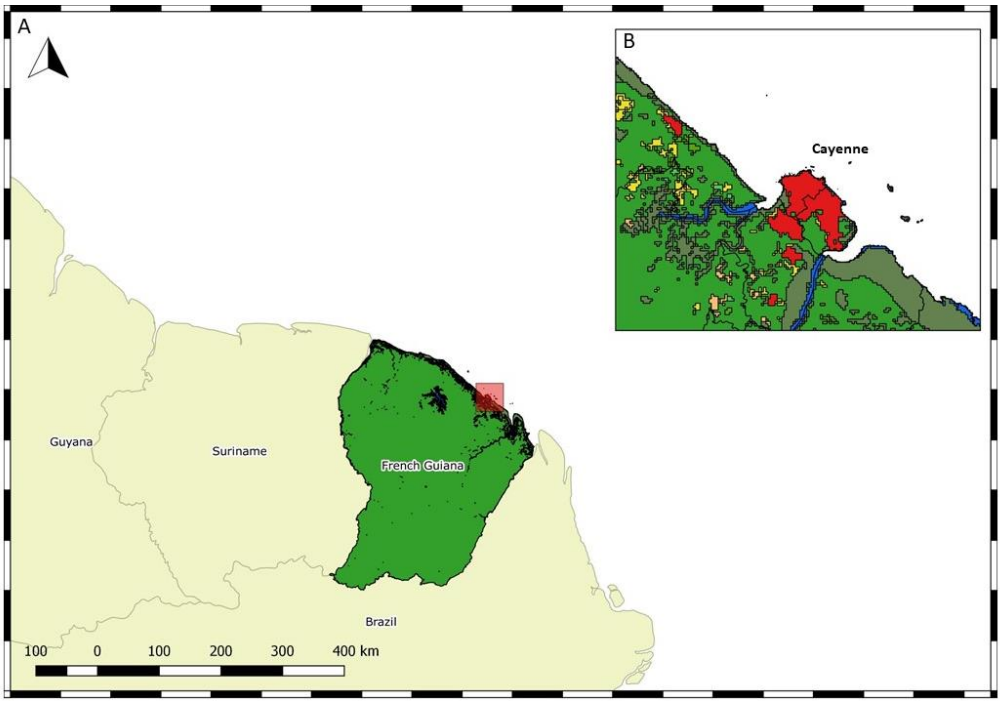

Figure 1.1 *Une carte de la Guyane française montrant les différentes couvertures terrestres : forêts primaires en vert, mangroves couleur olive, eau en bleu, zone urbaine en rouge, zone arbustive en orange et les terres cultivées en jaune. Encart B. Carte de couverture terrestre illustrant la proximité des zones de forêt primaire avec la région urbanisée de Cayenne, chef-lieu de Guyane Française.*

Mais le critère prévalent fut le centre hospitalier de Cayenne, chef-lieu de Guyane Française, où ont été méticuleusement enregistrées les EMI dans la région depuis 1960. Le territoire est divisé en 22 zones administratives nommées « communes ». Malgré une surface de 83 534 km², le territoire a une faible densité de population de 3.11 hab./km² avec 72.78% (95 CI 0.726–0.728) de la population totale vivant le long du littoral.

Cadre théorique

Cette thèse se divise en 5 chapitres, chacun des chapitres analyse l'émergence des maladies infectieuses en utilisant une approche biogéographique d'un groupe spécifique de pathogène. Le chapitre final résume et discute les résultats des précédents chapitres (*Figure 1.2*). Toutes les méthodes utilisées dans les chapitres sont différentes et innovantes dans le dom domaine de l'épidémiologie des maladies infectieuses.

Chacun des chapitres a déjà fait l'objet d'une publication, ou est en cours d'examen dans une revue à comité de lecture internationale.

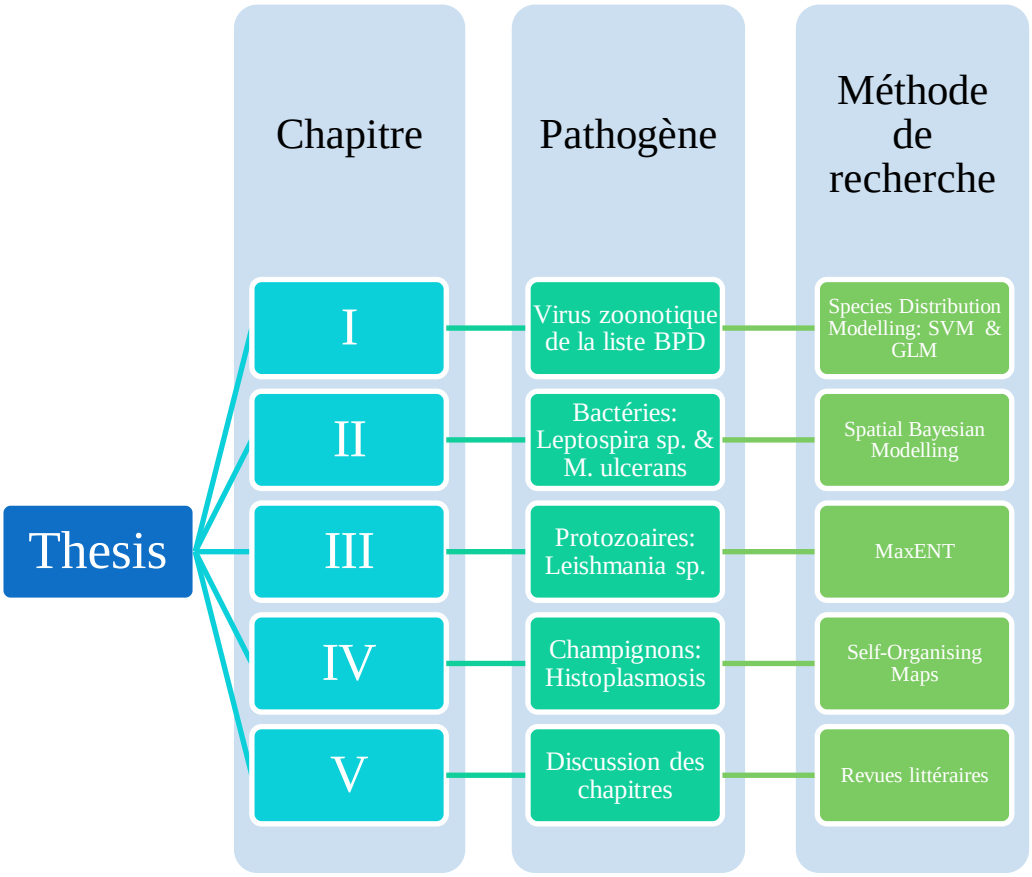

Figure 1.2 *Cadre théorique de la présente thèse*

1.2. Introduction (English)

The spread of infectious diseases has no spatial limitations. The COVID19 outbreak has demonstrated that a disease emerging in one corner of the world has the potential to turn into a widespread pandemic in the span of a couple of months due to the speed and international inter-connectivity in human mobility and trade. The concern over the rapid spread of Emerging Infectious disease (EIDs) and the increasing risk of emerging pathogens (Jones *et al*,. 2008, Fisher *et al*., 2012), has led to a resurgence in interest by the scientific community, the general public and the national and international authorities in the control and prevention of disease emergence at all spatial scales.

The current public health approach to disease outbreaks is proximate and reactionary. The local drivers triggering outbreaks such as for example wet markets are scrutinized. However, other environmental drivers like climate change and land fragmentation arising from a growing global agricultural demand, are often overlooked. Recent studies have shown that anthropogenic pressures on land use for agricultural expansion and livestock trigger EIDs. For example, the fruit-bat migration driven by the deforestation for palm tree cultivation through forest fires in the islands of Sumatra led to the emergence of Nipah disease through infected pigs, a viral encephalitis with a high fatality rate, in farmers and abattoir workers in Malaysia (Lo and Rota, 2008; Daszak *et al.*, 2013). The reactive measures of outbreak prevention and control following disease emergence, along with passive surveillance of EIDs are not effective. We need to "get ahead of the curve" i.e. identify and manage the drivers of disease emergence and detect regions at risk for future disease outbreaks.

In the past four decades, over 70% of emerging human infections were zoonoses (Taylor, Latham and Woolhouse, 2001; Jones *et al.*, 2008). Since most emerging pathogens have non-human animal reservoir, intermediate host, and/or vector species in their transmission cycle, the local human-animal-ecosystem inter-

actions in the environment play a major role in driving pathogen transmission from the environment to humans and thus disease emergence in human populations. Increasing human population, economic activity and climate change weigh upon these interactions.

The four drivers of EIDs

The impact of anthropization on the environment is unparalleled (McMichael and Butler, 2011). We are losing over 80,000 acres of tropical rainforest daily, and significantly degrading another 80,000 acres every day (*Measuring the Daily Destruction of the World's Rainforests - Scientific American*, no date). Along with this loss and degradation, we lose around 135 species per day as the forests fall. Climate change, land transformation, biodiversity loss, and animal trading could be considered as major drivers of disease emergence. Current evidence suggests that climate change is likely to be contributing to the diseases emergences in regions previously unaffected at higher latitudes and altitudes and to their persistence and changed seasonality in endemic areas (McMichael, Woodruff and Hales, 2006). In addition, flood-associated outbreaks of diseases such as leptospirosis have been reported in environments as diverse as India, Argentina, Australia, Brazil and Italy (Lau *et al.*, 2010).

Deforestation, urbanization, agriculture and other forms of landscape transformation have also increased the risk of EIDs. Increased urbanization leads to unsustainable agricultural and protein demand, increasing informal settlements and rising urban poverty (Rodwin and Gusmano, 2002). The urban poor are often homeless or survive in slums and have poor access to healthcare, which indirectly increases the risk of EIDs. Deforestation allows direct and indirect human contact with small mammal and birds, thereby increasing the risk of zoonoses emergence. Agricultural activities such as land clearing by deforestation, livestock rearing and rice cultivation have been found to have direct association to disease emergence (Daszak *et al.*, 2013; Jagadesh, Combe, Couppié, Le Turnier, *et al.*, 2019).

Biodiversity loss such as loss or displacement of predators can alter the population density of small mammal reservoirs and insect vectors in complex ways. For example, the extinction of the passenger pigeon

in North America, has led to an increase in the rodent population because acorns, formerly consumed by these birds, became more available for rodents to thrive. In turn, this increased the abundance of the animal reservoir of Lyme disease transmitting ticks (Blockstein, 1998). Bushmeat consumption and animal trading, arising from the growing demand of animal protein, causes significant changes at the human-animal reservoir interface (Wolfe *et al.*, 2005). The SARS epidemic of 2002 spread to humans through the consumption of Himalayan palm civets and raccoon dogs, both of which are intensively farmed in China. A chef is considered as the potential patient zero of the SARS epidemic. Chinese horseshoe bats were the original host for SARS-CoV 1,with civets and raccoons being the secondary, intermediate viral hosts (Li *et al.*, 2005; Wang and Eaton, 2007).

Tropics regions at risk of disease emergence

The regions at higher risk for bearing the environmental burnt are at the Tropics and are often developing countries with an existing burden of infectious diseases. Today, nearly 70% of extreme weather events occur in regions of Asia, the Pacific, Africa and the Middle East (Costello *et al.*, 2009). However, the increasing temperatures all-year around in the temperate regions have led a significant latitudinal shift in disease emergence, particularly vector-borne diseases such as dengue, chikungunya and CCHF (Vescio *et al.*, 2012; Tesla *et al.*, 2018). Furthermore, inequitable distribution of diagnostic tests, vaccines, and medicines to detect, prevent and control infectious diseases, along with weak public health delivery systems in developing and low-income countries, a disproportionate burden of human suffering and death from EIDs continue to occur in developing countries.

The world's poorest billion people live in environmentally and socially high-risk regions, which are also regions, where the prevalence of infectious disease is highest and are yet responsible for only 3% of the global carbon footprint. With environmental perturbation and disruption due to human activities, we need to acknowledge the significance of socioeconomic factors in the epidemiology of EIDs (Schneider and Machado, 2018). Rural-urban migration has resulted in structural poverty in communities where inadequate

sanitation, limited access to health care and under-nutrition are widespread, and promotes the transmission and persistence of a wide range of infectious diseases (Farmer *et al.*, 2006). Socioeconomic factors such as health inequality, poverty, urbanization, social stigma, disability, and migration have complex links to EID epidemiology (Anthony J McMichael and Ulisses Confalonieri, no date). Climate change imposed socio-economic factors on vulnerable populations such as migration, poverty and increasing demand due to food and water shortages increase the existing health inequality.

Get ahead of the curve

The ongoing COVID-19 pandemic has thought us that we cannot wait for novels discoveries of vaccines and 'miracle' drugs. We have moved passed the golden age of magic bullets and serendipitous discoveries. In era of increasing EIDs, we are desperately in need for a paradigm change. A new paradigm that integrates the fundamental centrality of the natural environment and its major role in buffering disease emergence. Predicting disease emergence is essential to "get ahead of the curve". Mathematical modelling and prediction provides quick assessment for control and potential preventive efforts when time for epidemiological studies is scarce (Siettos and Russo, 2013). Mapping of EIDs and prediction of emergence coupled with the growing resolution of remote sensing satellite imagery has shifted the new paradigm towards identifying potential environmental drivers ahead of disease emergence rather than the conventional surveillance of EIDs themselves.

The aim of this PhD thesis was to: (1) understand the spatial dependency and distribution of EIDs using a biogeographic approach; (2) use existing disease databases, freely available remote sensing data and validate their effectiveness and accuracy in predicting EIDs; (3) identify the environmental and socioeconomic drivers of disease emergence across different pathogen groups, globally and in French Guiana (FG). We sought to fulfill the following research objectives:

1. Mapping of the distribution of EIDs of different taxonomic pathogen groups such as bacterial, viral, fungal and protozoal using existing disease databases globally or in a tropical region of interest.
2. Identify the environmental and socioeconomic drivers of disease emergence using freely available remote sensing data like landscape imagery, climate maps and other rasters.
3. Measure the predictive risk of disease emergence and identify the potential hotspots of EIDs under study.
4. Compare the distribution and the factors influencing emergence and spread across the different pathogen groups.

A biogeographic approach for infectious diseases is a novel method based on mapping disease events, the distribution of disease vectors, reservoirs and/or intermediate hosts, and identifying the factors influencing their spatial patterns and trends. This approach not only provides a descriptive analysis on the transmission dynamics, but also quantitative measures of associations between disease and its influencing factors and in the prediction of future outbreaks (Escobar and Craft, 2016).

Significance of the study location

FG is a French Overseas territory located at 3.9339° N, 53.1258° W in South America (*Figure 1.3*). We chose FG as our study location as it is tropical region comprising of a densely crowded coastline with over 95% of the total land area is classified as primary rainforest forming a major portion of the highly biodiverse Guiana shield, which is in danger of anthropization to meet the demand of the increasing population. The land cover also includes distinct areas of savannahs, wetlands, and coastal mangroves.

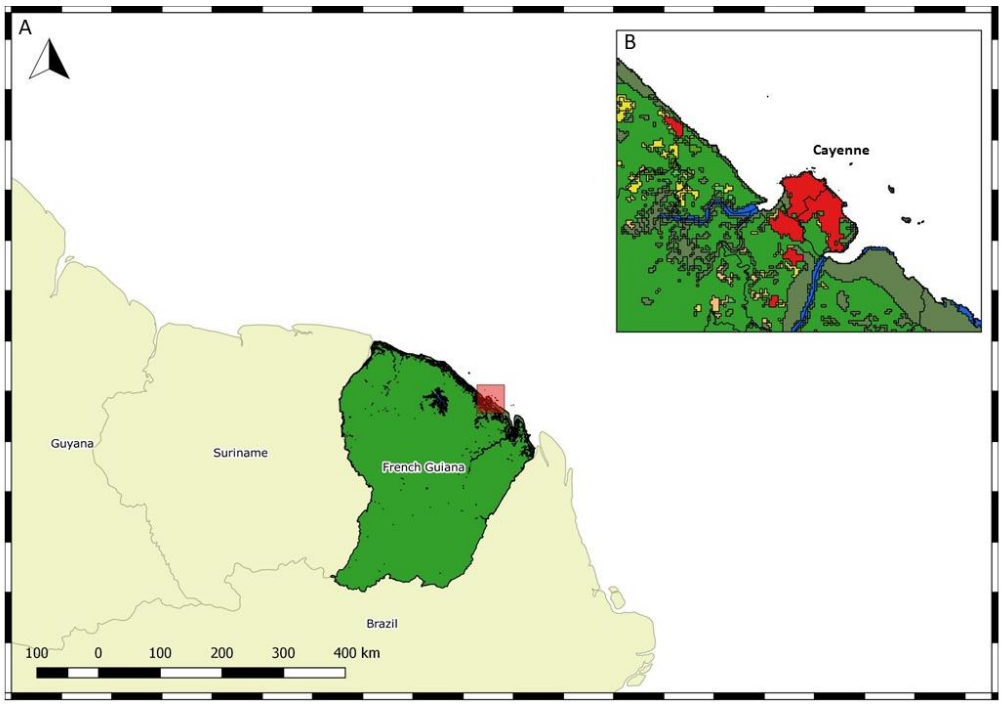

Figure 1.3. *A. Map of French Guiana showing the diverse land cover: primary forests in green, mangroves in olive, water in blue, urban area in red, shrubland in orange and cropland in yellow. Inset B. Land cover map illustrating the proximity of primary forests to the urban regions of Cayenne, capital of French Guiana.*

Most importantly, the central hospital at Cayenne, capital of FG, has meticulous records EIDs in the region since 1960. The territory is divided in 22 administrative units termed as communes. Despite its land area of 83,534 km2, the territory has a low population density of 3.11/km2 with 72.78% (95 CI 0.726–0.728) of the total population living along the littoral region.

Theoretical Framework

This thesis is divided into five chapters, with each chapter analyzing disease emergence using a biogeographic approach in a specific taxonomic pathogen group and the final chapter, summarizing and discussing the results of the preceding chapters (*Figure 1.4.*). All statistical methods used in the chapters are different and are novel in the study of the epidemiology of infectious diseases. The chapters have already been published or are under consideration in international peer-reviewed journals.

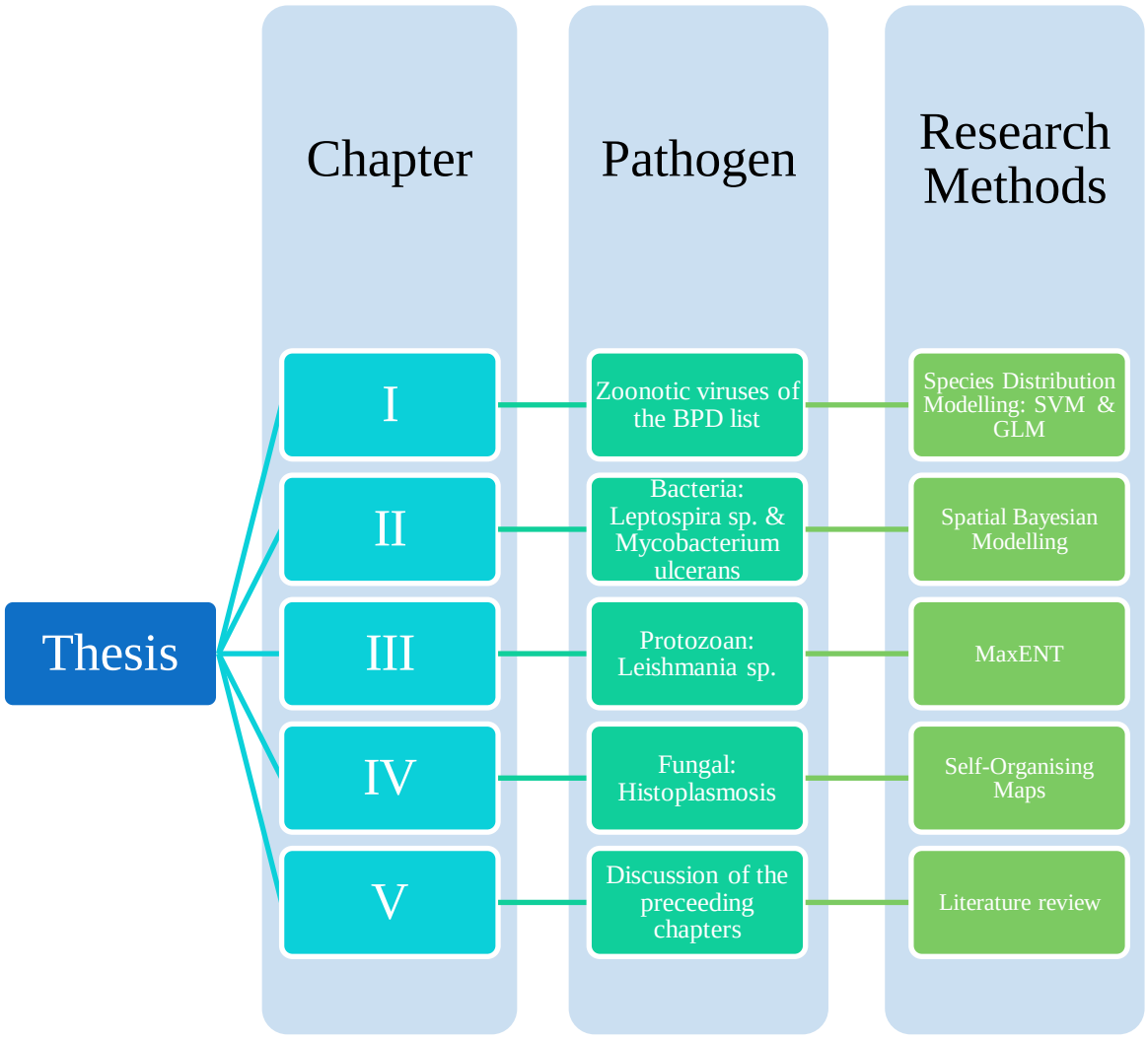

Figure 1.4. *Theoretical framework of the PhD thesis*

2.Chapter I

Global Biogeography of Viral Zoonoses

Title of journal article

In search for the hotspots of Disease X: A biogeographic approach to mapping the predictive risk of WHO's Blueprint Priority Diseases

Authors: Soushieta Jagadesh, Marine Combe, Mathieu Nacher & Rodolphe Elie Gozlan

Journal: Submitted to International Journal of Epidemiology (13th May 2020)

Conference: Accepted for the 19th International Congress on Infectious Diseases as an oral presentation and the abstract was awared the **ProMed Award**.

DOI preprint: 10.1101/2020.03.27.20044156

2.1. In the search of the hotspots of Disease X: A biogeographic approach

Abstract

Anthropization of natural habitats including climate change along with overpopulation and global travel have been contributing to emerging infectious diseases outbreaks. The recent COVID-19 outbreak in Wuhan, highlights such threats to human health, social stability and global trade and economy. We used species distribution modelling and environmental data from satellite imagery to model Blueprint Priority Diseases occurrences. We constructed classical regression and Support Vector Machine models based on environmental predictor variables such as landscape, tree cover loss, climatic covariates. Models were evaluated and a weighed mean was used to map the predictive risk of disease emergence. We mapped the predictive risk for filovirus, Nipah, Rift Valley Fever and coronavirus diseases. Elevation, tree cover loss and climatic covariates were found to significant factors influencing disease emergence. We also showed the relevance of disease biogeography and in the identification potential hotspots for Disease X in regions in Uganda and China.

Introduction

Changes to local land use and biodiversity, the increasing international connectivity through human transport and trade and climate change provide optimal conditions for the emergence of zoonotic infectious diseases. The displacement of the geographical footprint of pathogens and/or infected hosts leads to such EIDs (Ogden, AbdelMalik and Pulliam, 2017) with the COVID-19 as an ongoing example at the center of international scrutiny. The World Health Organization (WHO) has developed identify list of zoonotic diseases posing of a large-scale public health risk due to 1) their epidemic potential and 2) current absence or limited number of treatment and control measures available ('WHO | List of Blueprint priority diseases', 2018). The current list, established in 2018, includes Crimean-Congo hemorrhagic fever (CCHF), Ebola virus and Marburg virus disease (EVD & MVD), Lassa fever (LF), Middle East respiratory syndrome coronavirus (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS), Nipah and henipaviral diseases, Rift Valley fever (RVF), Zika and Disease X. The term "Disease X" refers to a potential international epidemic caused by a currently unknown pathogen, for which WHO calls for preparedness in the event of a novel disease emergence. This list of diseases in need of urgent research was termed as Blueprint listof Priority Diseases (BPDs).

All the BPDs in the current list can be classified as zoonotic viral infections. In the past four decades, over 70% of the emerging infections are or were zoonoses (Taylor, Latham and Woolhouse, 2001; Jones *et al.*, 2008). The increasing unpredictability in the global climate and the decreasing distance between the local human- animal- ecosystem interactions play a major role in driving the emergence of infection in human populations. The rising mean temperatures have been reported to have a significant effect in the tick-borne CCHF incidence and mosquito-borne Zika sustainability in subtropical and temperate regions (Vescio *et al.*, 2012; Tesla *et al.*, 2018). Bushmeat consumption and animal trading, arising from

the growing demand of animal protein, causes significant changes at the human-animal reservoir interface (Wolfe *et al.*, 2005). Among the BPDs, studies demonstrate that the SARS and EVD outbreaks were directly linked to the consumption of infected bushmeat (Smith *et al.*, 2012; 'An update on the risk of transmission of Ebola virus (EBOV) via the food chain', 2014). LF, MVD and EVD flourish in West and central Africa where the consumption of bushmeat is four times greater than the Amazon, which is richer in biodiversity (FA *et al.*, 1995). Moreover, EIDs are also triggered by anthropogenic pressures on land use for agricultural expansion and livestock farming to meet the demand of a growing human population. The fruit-bat migration driven by the deforestation through forest fires in the islands of Sumatra lead to the emergence of Nipah disease in farmers and abattoir workers in Malaysia (Lo and Rota, 2008; Daszak *et al.*, 2013). Once sufficient infection cycles between human-animal without the sustainable transmission among humans termed as viral chatter is established, the emergence of human-to- human transmission is inevitable (Wolfe *et al.*, 2005). Mathematical modelling and prediction provides quick assessment for control and potential preventive efforts when time for epidemiological studies is scarce (Siettos and Russo, 2013).

Modelling of infection dynamics analyses diseases outbreaks in animal population and estimates the rate of transmission as well as the potential chance of spillover. Yet, the increasing trends of EIDs risks surpass our capacity in the surveillance and detection of spillovers and outbreaks. Thus, the current public health response to EIDs is to "get ahead of the curve" and the growing resolution of satellite imagery has shifted the paradigm towards identifying potential environmental drivers such as deforestation, land fragmentation, biodiversity loss and climate change rather than the surveillance of the EIDs themselves.

Here we proposed mapping the predictive risk of the BPDs using Species Distribution Models (SDM) based on potential environmental drivers such as deforestation, landscape and climatic covariates derived from satellite imagery following the year 2000. The aim is to provide a global perspective, measure predictive risks and evaluate the use of biogeography on predicting diseases outbreaks. We

also used our approach to EIDs using disease biogeography as a tool to identify the potential hotspots for an unknown Disease X listed in the BPDs.

Methods

To construct robust models and generate predictive risk data of the BPDs' distribution, we adapted the following steps from species distribution modelling: 1) compilation of the geographical coordinates of the BPDs emergence; 2) construction of spatial buffer around the spatial points; 3) generation of pseudo absence points; 4) extraction of environmental predictor covariates; 5) fitting of the model and 6) calculation of the predictive risk of each BPD.

Location of BPD emergence

We extracted the distributional data on the global occurrence of the BPDs in humans across the years 2000 to 2019 from WHO archives, Promed mail and published studies (Appendix Table 2.1). In cases where the origin of the BPDs were unclear, we narrow down to the general region or district of origin. In other words, the localization of BPDs emergence correspond to patient zero of each outbreak. We did not include the diseases endemic to countries such as CCHF and Lassa fever. The calculation of predictive risk for endemic diseases and pandemics such as Zika may not contribute to the existing evidence. Laboratory outbreaks and outbreak in domestic and wildlife were also excluded. We georeferenced the sites of origin or the centroids of the region of occurrence to the nearest 0.0001°. Spatial buffers were constructed around the geographical coordinates depending upon the mobility range of the respective pathogens reservoirs and intermediate buffers (Appendix table 2.2). The construction of buffers also mitigates the relatively less precise geographical logical coordinates in cases where the exact point of origin was unclear.

Presence and Pseudo-absence points

We constructed spatial buffers to mitigate any potential inaccuracies in the localization of the occurrence points. For the presence points, we generated 100 random points at a 100km spatial buffer around the point of origin of each BPDs outbreak. Then, a spatial bufferzone of 150km was used as a mask and we generated 1000 pseudo-absence points within the geographical extent of the reservoirs of each of the BPDs. We chose the buffer of radius 150km to account to the average flying range of the order Chiroptera, the reservoirs of most BPD.

Environmental predictor datasets

We extracted the climatic covariates such as monthly maximum and minimum temperatures and precipitation from 'Terra Climate' high-resolution global data from 2000- 2017 (Abatzoglou *et al.*, 2018). Land cover data used was a composite of the 2017 MODIS satellite imagery and the deforestations mdata was derived from the Hansan's deforestation tree loss-year maps from 2000-2018. The topological data including altitude and hydrological models was extracted from one arc-second digital elevation model (DEM) of 30m resolution, derived from NASA (Version 3.0) SRTM imagery available in the United States Geological Survey USGS. We obtained the geographical distribution of the primaryhosts and reservoir mammals from the IUCN red list (Isberg, Balaguera-Reina and Ross, 2017). The shapefiles were rasterized and resampled into rasters of the 5km resolution raster model. We used livestock distribution rasters (2010) from the Food and Agriculture Organisation (FAO) for diseases such as RVF and Nipah including cattle, goats and pigs (Robinson *et al.*, 2014). The raster layers were resampled to a fixed resolution of 5km and stacked to a raster brick.

Fitting of model and Model prediction

The disease distribution models were fitted using the r package "discmo". In our study, we used the main two methods for species distribution modelling (SDM) i) the classical generalized linear models (glm) using gaussian regression methods and ii) the machine learning method, support vector machine

(SVM). We choose the glm models to analyze the influence of the environmental factors on the emergence of BPDs. SVM are popular in SDM using presence and pseudo-absence data. The models were evaluated using ROC curves and the area under curves (AUC) for the produced thresholds were calculated. Studies have critized the use of AUCs in the evaluation of SDMs, especially when the study involves large extent. We removed the "spatial sorting bias" through "point-wise distance sampling" as explained by Hijmanns, 2012. Model prediction was made using the "predict" function to map the predictive risk of the diseases based on the values of the independent variables extracted from the environmental predictor rasters. A weighed average of the glm and SVM models were calculated using the AUCs for each BPD model and a final composite prediction was made.

Deforestation analysis

To analyze the impact of deforestation on the emergence of the BPDs, we conducted a detailed spatiotemporal analysis using the R package "gfcanalytics" on the Hansan's tree cover loss maps. We used the presence points at a spatial buffer of 100km, used for the BPD models, for this analysis.

Results

We modelled five out of nine BPDs. We used glm and SVM methods of SDM for each of the disease. The AUC was maintained at an average of 0.9323 for glm models and 0.9597 for the SVM models. We have tabulated the AUC for each model in *Table 2.1*.

Table 2.1 *Generalized linear models (GLM) and support vector machine (SVM) regression coefficients for the AUC of the blue print diseases list analysis (BPD).*

Disease/Model	**Gaussian/GLM regression**	**Support Vector Machine**
Ebola virus (EVD)	0.9115	0.8680
Marburg virus disease (MVD)	0.9885	0.9913
Coronavirus (CoV)	0.8625	0.9831

Nipah and henipaviral diseases	0.9811	0.9911
Rift Valley fever (RVF)	0.9190	0.9653

Filovirus diseases

The distribution of the filovirus diseases, EVD and MVD, was restricted to the African subcontinent (*Figure 2.1* & *Figure 2.2*). The climatic covariates such as average total monthly precipitation [95%CI 0.0091 to 0.0305; P 0.0003], monthly maximum [95%CI -10.0125 to -2.07884; P 0.0034] and minimum temperature [95%CI -6.77 to -1.838; P 0.0008] were found to have a significant relationship with distribution of EVD emergence. From 2001 to 2018, the area of EVD emergence lost 2.25Mha of tree cover, equivalent to a 7.2% decrease in tree cover since 2000. MVD distribution was also observed to have an inverse relationship with minimum temperature [95%CI -6.7728 to -1.8380; P 0.0008] and a direct association with maximum temperature [95%CI 1.8275 to 8.0317; P 0.0018] and total precipitation [95%CI 0.015921 to 0.06010; P 0.001]. From 2001 to 2018, the regions contributing to the emergence of MVD lost 593kha of tree cover, equivalent to a 9.6% decrease in tree cover since 2000.

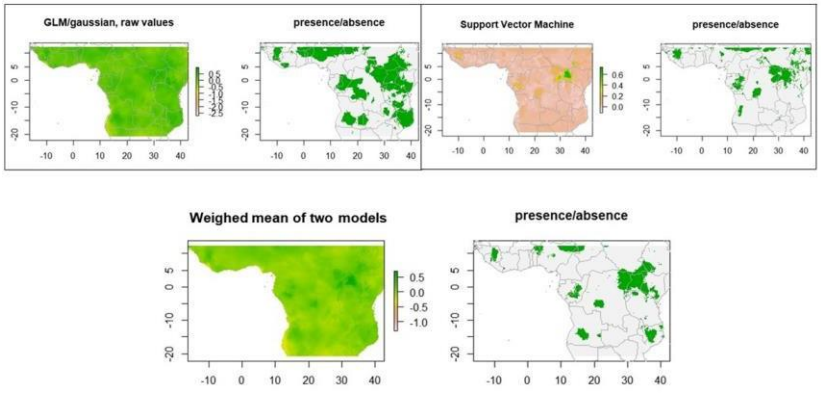

Figure 2.1 *Map illustrating the predictive risk of Ebola viral disease using glm, SVM and the weighed composite models.*

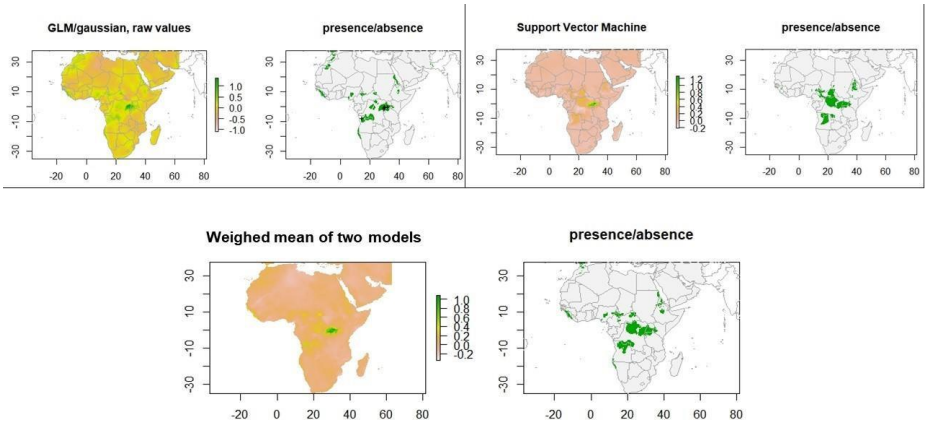

Figure 2.2 *Map demonstrating the predictive risk of Marburg viral disease using glm, SVM and the weighed composite models.*

Nipah virus disease

Elevation [95%CI -6.0114 to -1.4460; P 0.0014] along with the average maximum temperature [95%CI -8.5197 to -3.9108; P <0.0001] were found to have a negative correlation while the average minimum temperature [95%CI 2.6843 to 8.5860; P 0.0002] had a positive association with the distribution of Nipah disease (*Figure 2.3*). The occurrence sites in Bangladesh are comprised of the Lower Gangetic Plains Sundarbans mangroves vegetation and no intact forest. However, between the years 2016 and 2018, Kerala lost 23.2kha of tree cover, equivalent to a 0.89% decrease in tree cover since 2000.

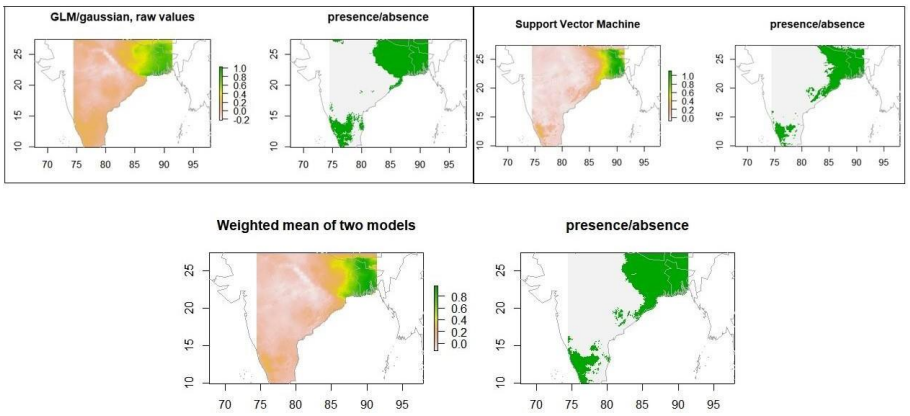

Figure 2.3 *Map illustrating the predictive risk of Nipah using glm, SVM and the weighed composite models.*

Rift Valley Fever

We found no significant relationship between the distribution of the RVF emergence (*Figure 2.4*) and the regions of cattle and goat grazing. However, we observed significant correlations with the average minimum temperature [95%CI 0.05881 to 0.14845; P <0.0001]. The tree cover loss in the regions with RVF occurrence is negligible.

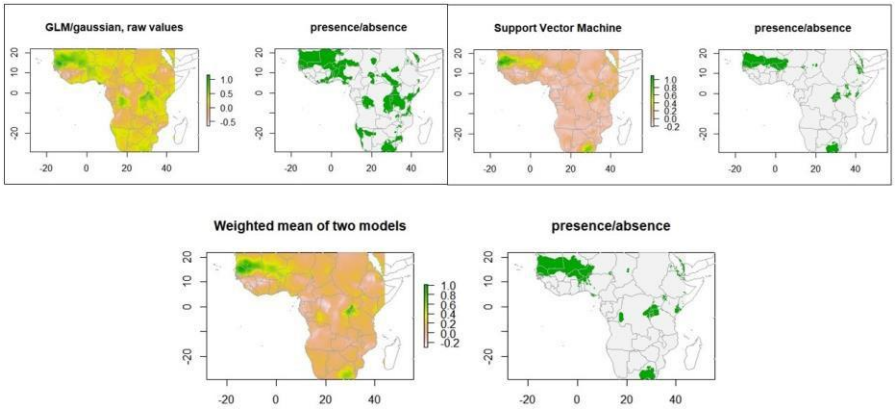

Figure 2.4 *Map demonstrating the predictive risk of Rift Valley Fever using glm, SVM and the weighed composite models.*

Coronavirus diseases

Elevation [95%CI -0.0014 to -0.0007; P <0.0001] and the average minimum temperature [95%CI 0.0067 to 0.0218; P 0.0003] had a positive association with the distribution of coronavirus diseases (*Figure 2.5*). From 2001 to 2003, Guangdong, the province where SARS emerged, lost 83.3kha of tree cover, equivalent to a 0.96% decrease in tree cover since 2000. The regions involved MERS had a negligible loss in tree cover.

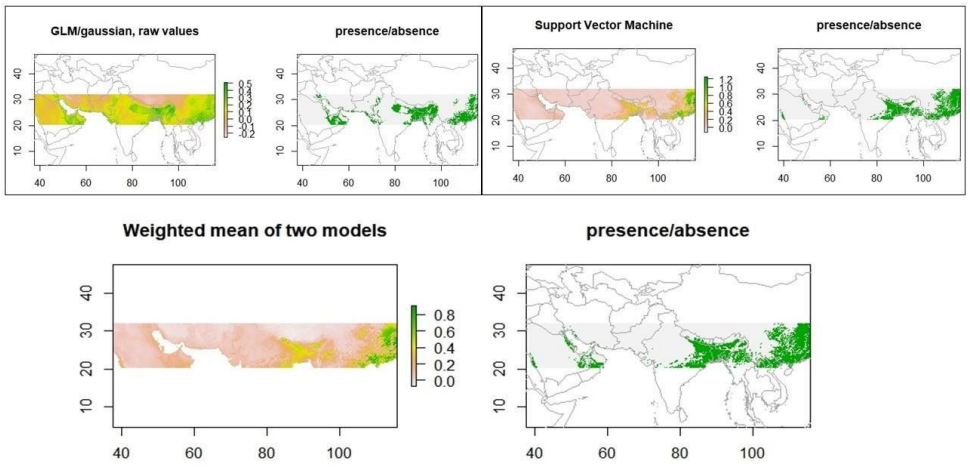

Figure 2.5 *Map illustrating the predictive risk of Coronavirus diseases – SARS and MERS using glm, SVM and the weighed composite models.*

Disease X

The weighed means of the AUC from the BPD models were used to create a threshold and the potential hotspots for the emergence of disease X was mapped (*Figure 2.6*).

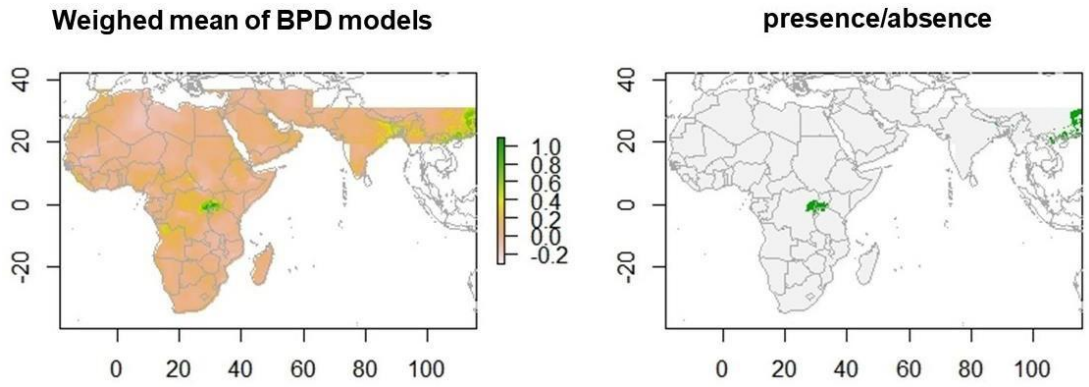

Figure 2.6 *Map illustrating the potential hotspots of Disease X based on the weighed composite models of the BPDs.*

Discussion

Our study was the first to map the predictive risk of the disease emergence using species distribution models and to establish the impact of environmental factors on their emergence. We also identified the potential hotspots for future emerging infectious diseases based on the predictive models of BPDs. The results of our study has shown that the disease emergence is spatially dependent on the bioclimatic factors such as elevation, tree cover loss and climatic covariates. These factors could thus be utilized to identify and predict the hotspots of Disease X.

We found elevation to have a significant influence on the distribution of Nipah and coronavirus diseases. Studies have hypothesized that nipah emergence in lower Gangetic plains and at the low-lying backwaters could be attributed to flooding, which causes destruction of mammal habitats (Ambat *et al.*, 2019). Rapid changes to habitats lead to starvation and migration of the known reservoirs of the Nipah virus, the bats, with contamination of fruit trees near human dwelling and increased exposure to the pathogen. Our results confirm this hypothesis associated with low- lying plains, flooding and emergence of Nipah disease. Coronaviruses are also linked with bat reservoirs. We hypothesize that similar events leading to destruction of bat habitats due to flooding at the low-lying regions such the wadis near Jeddah and Zhujiang Basin at the Province of Guangdong triggered the emergence of MERS and SARS respectively.

A recent study found that the increase in surface temperature and the unpredictability in the seasonal rainfall due to climate change, has an indirect effect on disease emergence through the sudden changes to the reservoirs habitats, biodiversity loss and small mammal migration (García *et al.*, 2018). Our results demonstrates that increase in the residual temperature also known as the minimum temperature has a direct influence on the emergence and distribution of the BPDs. This direct spatial dependence of disease emergence on minimum temperatures is worrying. With climate change, the increasing night minimum temperatures lengthens the freeze-free season in most mid- and high latitude regions (Lead *et al.*, no date). These conditions with worsening climate change could increase the potential latitudinal extent of disease emergence. This hypothesis has been established from the vector-borne BPDs emergence predictions like CCHF and Zika (Gale *et al.*, 2012; Tesla *et al.*, 2018).

Our study observed that deforestation is another important factor driving the emergence of BPDs. Studies have established the impact of deforestation and migration of bats on the occurrence of EVD, Nipah and SARS diseases (Wilcox and Ellis, no date; Lo and Rota, 2008; Olivero *et al.*, 2017). Most of the models showed tree cover loss at 100km around the occurrence of the BPDs (see for example EVD, Figure 7).

Biodiversity loss, due to anthropogenic changes to their habitat such as deforestation, land fragmentation for agriculture, increase in demand for protein and climate change lead to decrease in predators and migration of the small mammals towards human habitations. We hypothesize that a decrease in predators could cause an imbalance in the natural predator-prey equilibrium leading to an increase in small mammal reservoirs and thus to viruses transmission via disease vectors such as ticks. Lesser species diversity and inter-species interactions makes spillover easier towards the accidental human hosts. This has been termed as the “dilution effect” (Guégan, 2008).

As with all mathematical models, our study has its limitations. The size of the environmental predictor raster layers limited our spatial extent of the models. We chose the quality of the satellite imagery and resolution over a global perspective with poor or outdated data. Our study does not have a temporal

component in the form of times series, which would be interesting especially with the climatic covariates. We mitigated this by choosing recent raster data corresponding to the period of the study and linking the spatiotemporal presence and pseudo-absence points to corresponding climatic monthly covariates. Despite these limitations, our study is the first to confirm the validity and effectiveness of using SDMs and other mathematical models to predict and identify the potential hotspots for BPDs. The use of a biogeographic approach in disease modelling offers a wider perspective on the environmental drivers and highlights the importance of climate change in the context of disease emergence.

Most of all, our study results observed that the potential hotspots for an unknown disease X is located in Uganda and China (*Figure 2.6*). It is interesting to note that the associated predictive risk map includes the region of Wuhan, the epicenter of the ongoing COVID-19 outbreak.

Conclusion and recommendations

In our study with the use of a biogeographic approach, satellite imagery and SDMs, we were able to identify Wuhan as a potential hotspot of disease emergence in the absence of COVID-19 data. We confirm that distribution of disease emergence in humans is spatially dependent on environmental factors such as landscape, tree cover loss and climatic covariates. The following are our recommendations to get ahead of the curve in the prediction and prevention of disease emergence:

1. Evaluate regions at high risk of flooding and identify them as hotspots for disease emergence at the tropics. With the unpredictability in rainfall and rising sea levels, these regions are in need of active disease surveillance.
2. The direct relationship of the disease emergence with rapid changes in surface temperatures pose the threat of spatial latitudinal extension of the disease. Urgent need of global efforts to communicate the impact of climate change on future emergence of a disease like COVID-19 and thus to include EIDs when evaluation the economic costs of climate change.
3. Alternate solutions to deforestation, land fragmentation for livestock farming and bush-meat consumption to meet the growing protein demand.

4. The ongoing COVID-19 is not unprecedented but rather predictable using simple mathematical modelling techniques and freely available satellite imagery. We recommend the using a biogeographic approach to predictive risk mapping to identify potential hotspots of disease emergence.

3.Chapter III

Biogeography of Bacterial diseases of aquatic origin

Title of journal article

Emerging human infectious diseases of aquatic origin: a comparative biogeographic approach using Bayesian spatial modelling

Authors: Soushieta Jagadesh, Marine Combe, Pierre Couppié, Paul Le Turnier, Loïc Epelboin, Mathieu Nacher & Rodolphe Elie Gozlan

Citation: Jagadesh, S. et al. (2019) 'Emerging human infectious diseases of aquatic origin: A comparative biogeographic approach using Bayesian spatial modelling', International Journal of Health Geographics. BioMed Central Ltd., 18(1), p. 23.

DOI: https://doi.org/10.1186/s12942-019-0188-6

Title of journal article

Global emergence of Buruli Ulcer

Authors: Soushieta Jagadesh, Marine Combe, Pierre Couppié, Mathieu Nacher & Rodolphe Elie Gozlan

Citation: Jagadesh, S. et al. (2019) 'Global Emergence of Buruli Ulcer', EcoHealth. 16, pages 591–593

DOI: https://doi.org/10.1007/s10393-019-01445-z

3.1. Emerging human infectious diseases of aquatic origin: A comparative biogeographic approach using Bayesian spatial modelling

Abstract

Background: With the increase in unprecedented and unpredictable disease outbreaks due to human-driven environmental changes in recent years, we need new analytical tools to map and predict the spatial distribution of emerging infectious diseases and identify the biogeographic drivers underpinning their emergence. The aim of the study was to identify and compare the local and global biogeographic predictors such as landscape and climate that determine the spatial structure of leptospirosis and BU.

Methods: We obtained 232 hospital-confirmed leptospirosis (2007–2017) cases and 236 BU cases (1969–2017) in French Guiana. We performed non-spatial and spatial Bayesian regression modelling with landscape and climate predictor variables to characterize the spatial structure and the environmental drivers influencing the distribution of the two diseases.

Results: Our results show that the distribution of both diseases is spatially dependent on environmental predictors such as elevation, topological wetness index, proximity to cropland and increasing minimum temperature at the month of potential infection. However, the spatial structure of the two diseases caused by bacterial pathogens occupying similar aquatic niche was different. Leptospirosis was widely distributed across the territory while BU was restricted to the coastal riverbeds.

Conclusions: Our study shows that a biogeographic approach is an effective tool to identify, compare and predict the geographic distribution of emerging diseases at an ecological scale, which are spatially dependent to environmental factors such as topography, land cover and climate.

Introduction

In recent years, rapid global and local environmental changes (e.g. deforestation, pollution, climate change) in the tropics have led to significant modifications in biodiversity (Pereira, Navarro and Martins, 2012). Such rapid changes have underpinned alteration of host–pathogen patterns leading to more frequent and random EID outbreaks (Roche and Guégan, 2011; Murray *et al.*, 2015; Morris *et al.*, 2016). The direct and indirect consequences of these human-driven environmental changes result in the alteration of the geographic distribution of aquatic hosts and/or reservoir with a direct effect on the distribution of their pathogens (i.e. bacteria, viruses, parasites, fungi) (Murray *et al.*, 2015). Thus, the past and current geographical distribution of suitable habitats constrain pathogens survival and growth (Roche and Guégan, 2011), leading to the redistribution of EIDs' risk in local human populations. For example, deforestation has recently been linked to an increased risk of BU, a bacterial skin disease due to *Mycobacterium ulcerans* with high tropism for skin and causing severe ulcerations in humans (Morris *et al.*, 2016). Globally, over 20 years, the need for arable lands in the tropics in response to a rapid growing human population has led to a 28% decrease of primary forest (Gibbs *et al.*, 2010). Studies have demonstrated that small physical changes influenced by deforestation and climate in freshwater systems lead to significant restriction in the distribution of biota, altering the dynamics of these ecosystems (Harvell *et al.*, 2002; Perkin *et al.*, 2011). These freshwater habitats also act as a carrier of certain pathogens capable of causing human disease. Thus, rapid alterations in freshwater habitats can influence the exposure of the pathogen to humans, resulting in emergence of disease in a region over time. The global significance of EIDs caused by pathogens found freshwater habitats has been less studied in comparison to terrestrial ones (Harvell *et al.*, 2002; Lafferty, Porter and Ford, 2004; Ostfeld, Keesing and Eviner, 2008) at an ecological scale. Outside malaria and cholera, there is little evidence linking spatial–temporal patterns of EIDs in human population and biogeographical freshwater drivers.

To construct a better understanding of biogeographical drivers influencing the local distribution of EIDs, the dynamics of human cases has to be analyzed in space and time in light of significant factors

such as topography, land cover and climate. All three factors have direct implications in the availability of freshwater, the distribution in time and space of this water and on the type of aquatic habitats (i.e. flowing vs stagnant). It remains unclear if the emergence, transmission and distribution of EIDs in a region can be determined based on the spatiotemporal relationship between the disease incidence and these environmental factors. We hypothesize that the spatial structures of two EIDs caused by pathogenic bacteria occupying similar freshwater niche may overlap if a spatial pattern exists.

Here we used a couple of pathogenic aquatic bacteria, Leptospira spp., responsible for leptospirosis and *M. ulcerans* responsible of BU, as biological models to test our biogeographical predictions. These two infections are severe re-emerging diseases of epidemiological concern in humans in intertropical regions like FG, the area of study (Douine *et al.*, 2017b; L. Epelboin *et al.*, 2017; Loïc Epelboin *et al.*, 2017). Although the transmission dynamics of the two bacteria are widely different, the mode of transmission to humans involves direct contact with aquatic habitats contaminated by the bacteria (*Figure 3.1*). Leptospirosis is one of the most widespread zoonotic diseases as it occurs in temperate and tropical regions, and in urban and rural settings, dependent on the spatial distribution of its mammal reservoir, especially rodents such as *Rattus rattus* and *Rattus norvegicus*. Studies demonstrated that culturable pathogenic Leptospira were detected in soil for at least 16 days and in spring water for 28 days (Casanovas-Massana *et al.*, 2018). This suggests that the environment is not a multiplication reservoir but rather a temporary carrier for pathogenic Leptospira. While BU, on the other hand, is a generalist pathogen, globally more restricted in its spatial distribution to regions near wetlands and slow-moving rivers, notably areas prone to flooding in humid tropical and subtropical areas. *M. ulcerans* DNA has been detected in sediments, mud, detritus, biofilms, and aquatic invertebrates in still lentic and flowing lotic systems in the environment (Combe *et al.*, 2017b). The transmission dynamics of BU still remains unclear, but is believed to be related to exposure to freshwater systems that contains *M. ulcerans* through abraded skin (Combe *et al.*, 2017b). The public health response to the presence of pathogenic bacteria in the envi-

ronment at present is reactionary. However, systematic surveillance of the pathogenic bacteria in the environment would aid in the prediction and control of outbreaks.

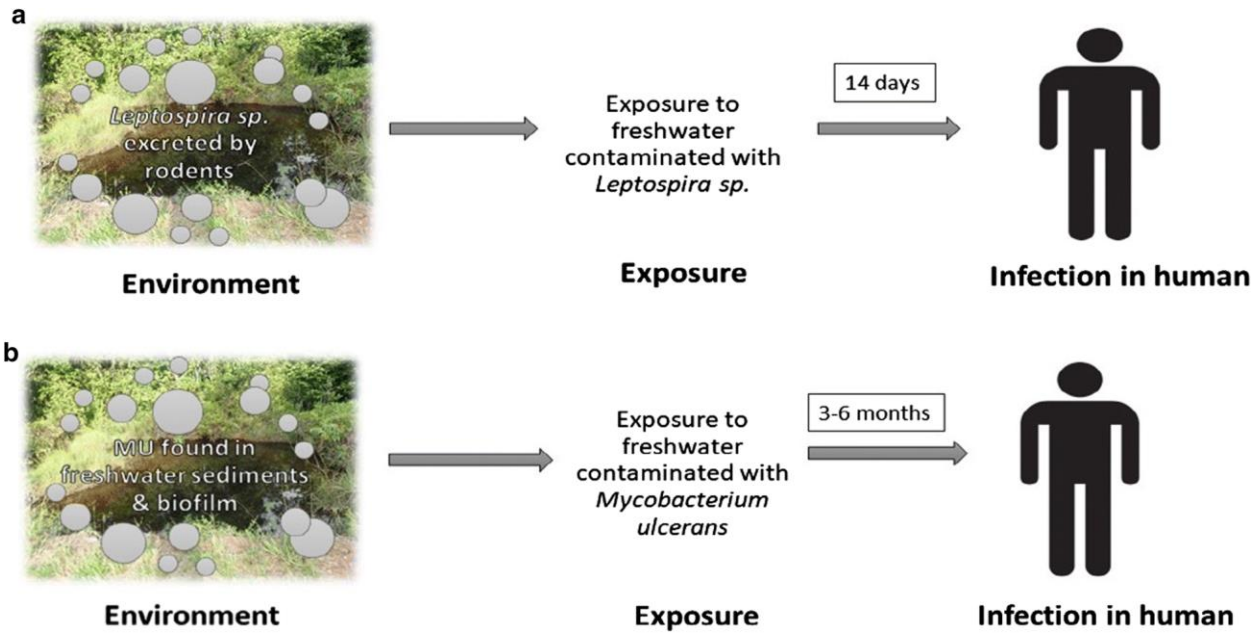

Figure 3.1 ***a)*** *Transmission dynamics of leptospirosis illustrating the host–pathogen–environment interface. b) Transmission dynamics of BU; MU represents Mycobacterium ulcerans*

The overall aim of our study was to identify the patterns of leptospirosis and BU cases distribution and to quantify the local and regional biogeographic drivers underpinning such distribution. We hypothesized (1) that the local topography, land cover and climate spatially influenced the distribution of both diseases, and (2) that a distinct spatial pattern exists for the environmental drivers of both diseases, although spatial patterns may overlap as both infectious agents occupy similar freshwater niches. Such understanding is important to local Health Agencies in order to optimize local developments and habitat management as well as to incorporate EIDs risk in the decisions of local planers.

Materials and methods

Ethics statement

The study protocol was approved by Cayenne General Hospital authorities according to French ethical rules. The leptospirosis and BU database were declared to the CNIL NO 2068308 and CNIL NO 3X#02254258 respectively following the requirements imposed by the French law. The database was anonymized and excluded from variables that facilitates identification of the patients. The leptospirosis and BU cases received appropriate treatment as per the French laws in public health.

Study area

Our study was conducted in French Guiana (FG), an overseas territory of France (*Figure 3.2*). The region is characterized by cyclic wet and dry seasons.

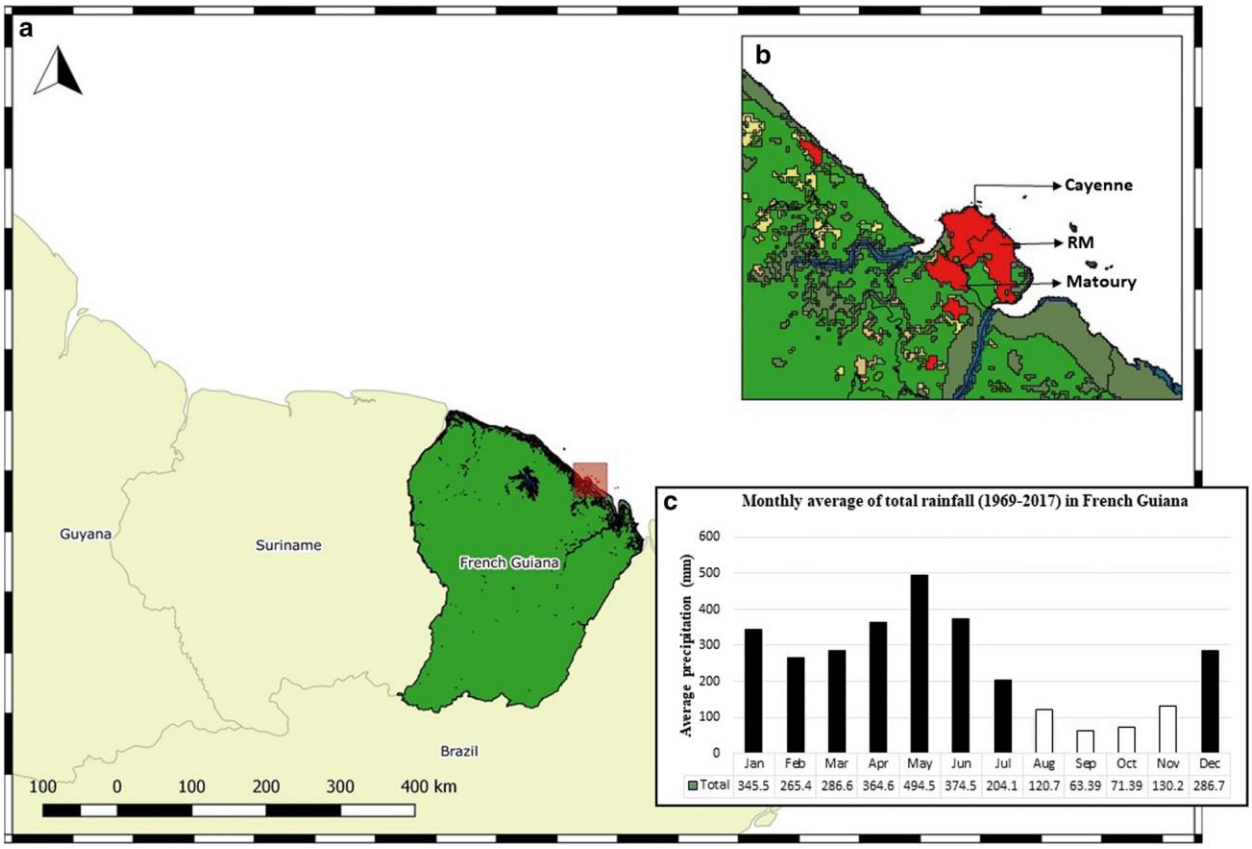

Figure 3.2 *a) Map of French Guiana showing the diverse land cover: primary forests in green, mangroves in olive, water in blue, urban area in red, shrubland in orange and cropland in yellow. Inset b) Land cover map illustrating the proximity of primary forests to the urban regions of Cayenne, Rémire-Montjoly and Matoury. Insert c) is a graphical representation of the monthly average of total rainfall in millimeters (mm) from six meteorological station across FG; black bars represent the wet season and white bars dry season*

Diagnostic criteria and patient data

Cases of leptospirosis were defined as a case with ongoing symptoms that are compatible with the clinical description of the disease along with a positive PCR result or a fourfold rise in titer in 2 weeks measured by MAT or a positive IgM ELISA. Cases from 2007 to 2014 (from January 1, 2007 to September 30, 2014) were validated by checking one by one all the medical charts. These cases were confirmed by having a positive PCR from blood or urine samples, cerebrospinal fluid and/or a MAT seroconversion with MAT titers ≥ 100 and/or a fourfold increase in MAT titers on two consecutive sera samples, and/or MAT titers ≥ 200 and/or a positive MAT titer with IgM seroconversion or IgM elevated titer. Cases from 2007 to 2014 were used in a previously published paper where the methodology is described in details (Cropet *et al.*, 2018). Cases from 2014 to 2017 (from October 1, 2014 to December 31, 2017) were added to this study. No medical charts were checked and the microbiological diagnosis relied on positive PCR on any fluid, and/or positive MAT and/or positive IgM (Niloofa *et al.*, 2015).

A confirmed case of BU was defined as a probable case with a clinically compatible cutaneous or bone lesion meeting the WHO clinical definition of BU (World Health Organization, 2012) and the detection of *M. ulcerans* in smear or by histological examination using Ziehl–Neelsen microscopy or/and IS2404 PCR.

In both cases, we excluded duplicate entries based on the address of the patient; to exclude patients with recur- rent infection or reinfection. The age, sex, date of diagnosis and the spatial coordinates (Cartesian coordinates) of the patient's residence were extracted from the central hospital database managed by the specialist in-charge. The cases were projected at WGS84/UTM21 onto a shapefile (.shp) and the same coordinate reference system (CRS) was maintained for all spatial and statistical analysis done.

Non-spatial vs spatial modelling

To test the objectives of our study, we developed non- spatial and spatial models using Bayesian regression to identify if there was a spatial structure in the distribution of the cases or if the cases occurred in random. The model achieving a better fit was chosen. We also analyzed the influence of spatial drivers such as landscape and climate on distribution in the models. Finally, we compared the significant models

of both the diseases under study to delineate the similarities in the biogeography of two bacterial diseases of similar freshwater origins.

Topography

A one arc-second DEM of 30 m resolution, derived from NASA Version 3.0 SRTM imagery available in the USGS website, was constructed using QGIS (version: 2.8, Las Palmas). As the SRTM Global 1 arc-second product from 2016 is void filled, no further processing was done prior to hydrological modelling. The mean, minimum, maximum measures of the elevation and TWI were derived for each spatial point from a 30 m resolution DEM. TWI is an index measure that illustrates the capacity of a region to accumulate water in presence of rainfall (Beven and Kirkby, 1979). TWI is a function of local upslope draining through a certain point per unit contour length (a) and slope (tanβ). We calculated TWI, ln(a/tanβ), from DEM models using TOPMODEL, a runoff method created by Beven and Kirby using to detect flood prone regions. The index detects potential ponding areas, regions of increased soil moisture, and rainfall runoff. The mean elevation was categorized into nine levels: 0–5 m, 5–10 m, 10–50 m, 50–100 m, 100–200 m, 200–300 m, 300–400 m, 400–500 m, and 500–600 m above sea level to explore the environmental attitude threshold for both infections.

Land cover data

To assess the relationship between land cover and dis- ease incidence, we used the MODIS-based Global Land Cover Climatology data developed by the USGS (Broxton *et al.*, 2014). The data is derived from the Collection 5.1 MCD12Q1 land cover type data and is based across a time period of 10 years (2000–2010). It provides the land cover classification with the highest confidence as validated by Broxton et al. The composite raster of French Guiana was cut from the global map, matched in scale and resolution to the topological raster and its land cover type was broadly classified into eight classes (*Table 3.1*) using QGIS. Due the persistent cloud cover over the trans-equatorial zone in the imagery prior to 2000, disease cases exclusively from 2000 to 2017 were used for the land cover model generation.

Table 3.1*Proportion of each land cover in French Guiana*

Land cover class	Proportion of total land covered (%)
Primary forest	96.38
Mangroves	2.14
Water	0.77
Shrub land (flooded and non-flooded)	0.28
Crop land	0.13
Mosaic forest	0.13
Urban	0.12
Grassland	0.04

To analyze the land cover surrounding the disease cases, spatial buffers of radius 2 km, 5 km and 10 km were constructed. The radii of the spatial buffers were chosen to represent the land use and the environment in proximity to the sample population. The proportion of the land cover class contained in the three buffers for each spatial point was extracted and the resulting land cover variables was used for the regression models. One of the limitations in our study is the lack of land cover data at 30 m resolution for the years prior to 2000. As a result, the BU data prior to 2000 could not be used in the construction of land cover models and restricted the construction of the model combinations for BU. We chose to restrict the time period of land cover BU model, to focus on the spatial drivers of disease emergence to test our hypothesis. Using topological, land cover and climate grids at 30 m resolution eliminates the spatial discrepancy between the models and ensures the quality of the spatial data used.

Meteorological covariates

The meteorological data including the monthly maximum temperature (Tmax), minimum temperature (Tmin) and monthly total precipitation for the year 1968 to 2017 were obtained from six meteorological stations distributed across FG. Météo France provided the data. The monthly variables matched

the temporal resolution for the disease cases used for Bayesian modelling. Climate covariates were spatially interpolated from points to climate grids using an IDW approach in QGIS. IDW uses the nearest neighbor interpolation method, which takes on the value of the closest sample (Attorre *et al.*, 2007). However, using local interpolation might not show micro influences where neighboring data is not local enough. In our study, we used meteorological data from six weather stations located in the most habited regions and across the country to mitigate the underestimation of micro influences. The spatial resolution of the climate grid models matched the DEM models to preserve the resolution through the multilevel modelling. The value of the pixel that fell under the points representing the spatial points in the shape file was extracted, resulting in monthly meteorological covariates for each spatial point. In addition, in the case of BU, for each spatial point, the meteorological data 1 to 6 months prior to the reported date were also included to account for the unknown time of exposure, incubation period, appearance of symptoms and delay in health seeking behavior. To illustrate, the observed climatic covariates for a case reported in mm/yy (m) at a specific spatial point (xy) is noted and the climatic observations 6 months (m−1, m−2, m−3, m−4, m−5 and m−6) prior to "m" are also included. The incubation period of leptospirosis is distinctly shorter i.e. around 14 days, and so lag time was not included for leptospirosis.

Regression modelling

To identify the significant relationships between the two diseases and the various topographical, landscape, meteorological and demographic variables across the country, regression modelling was used. We used mixed effects MCMC approach for developing non-spatial and spatial models. To achieve optimal power in the regression models, we generated random spatial points across a spatial polygon file of FG (water bodies were excluded) stratified by the proportion of population in each commune. The number of background points was optimized by power calculation using G power (version 3.1) and an a priori number of 500 spatial points (approximate 1:2 = presence: absence ratio) were generated as controls. The same set of background points were used in both leptospirosis and BU modelling. To address the temporal nature of the controls, dates in dd/mm/yyyy format were randomly generated for the spatial

point across the time periods; from 2007 to 2017 for the leptospirosis dataset and from 1968 to 2017 for the BU dataset.

Non-spatial models

A non-spatial GLM was tested on the 232 and 236 cases of leptospirosis and BU respectively, and the strength of their association with elevation, landscape and climate covariates. The predictor covariates for each case was indexed by location, K = {ki,…,k}, where each k is a vector recording of the longitude and latitude at UTM 21 N projection. The response variable y(k) was the presence or absence of disease at generic location k. The covariates for each response variable at k were recorded. Simple logistic regression was done to check for the presence of association between the presence of case and each covariate. We then used a MCMC sampler for Multivariate Generalized Linear Mixed Models to establish the relationship between the dependent variable and the covariates introduced as fixed effects. The "MCMCglmm" package was used for analysis (Hadfield, 2018). The DIC was extracted from the models for comparison with the spatial models.

Spatial models

All models were generated using the "binomial" family of spGLM function from spBayes R package. The model parameters were estimated using MCMC methods utilizing an AM algorithm with a 43% acceptance rate (Gelman and Rubin, 1992). The starting coefficient values and the beta tuning were obtained from the non-spatial logistic regression models. The predictor variables not significant in the logistic and MCMC regression non-spatial models were not analyzed in spatial modelling. Both the spatial and non-spatial models for each disease were tested individually for elevation and TWI under topographically variables, each of the land cover predictors, and climatic covariates such as maximum, minimum temperatures and total precipitation.

Model comparison and verification

Models were compared using the DIC for the Bayesian models. The models with lower DIC values indicate better model performance during model comparison similar to the AIC values (Spiegelhalter *et al.*, 2002). Prior and posterior predictive checks were conducted to ensure the robustness of the models. All statistical tests were set at the conventional 5% significance level. All statistical models are based on assumptions and the spatial models in our study are no exception. In the Bayesian spatial models used, data was assumed to have a spatial structure. This assumption was mitigated by conducting a preliminary spatial cluster analysis to confirm the presence of a spatial structure of the leptospirosis and BU cases. The cluster analysis was done using "satscan" R package. The area-level random effects are not assumed constant but is under the assumption that the outcome between two neighboring spatial points is more similar than that between two distant spatial points in Bayesian models. Statistical R packages and datasets used for each model are detailed in Appendix Table 3.1.

Results

Our results show that leptospirosis was widely distributed across FG, occurring in 18 of the 22 communes (Appendix 2: Table 2). The incidence of leptospirosis in FG was found to be 0.96 [95% confidence interval 0.8–1.1] per 1000 people during the period 2007–2017. BU was found to be restricted to the cities and towns along the coast, occurring in nine communes. During 1969–2017, the incidence of BU was found to be 1.9 [95% CI 1.7–2.2] per 1000 people in FG. The incidences of both diseases for each commune has been provided in Appendix 2: Table 2.

Spatial vs non-spatial models

Our results demonstrate that most spatial models (9 out of 12; 75%) produced lower DIC values in comparison to the non-spatial models (*Table 3.2*). This illustrates a spatial dependence of leptospirosis and BU cases distribution in FG towards environmental drivers such as elevation, TWI, land cover and climate. On comparison of the regression coefficients between the spatial and non-spatial models, we observed that the non-spatial models overestimated the significance of the environmental variables likely attributed to the violation of the basic model assumptions (*Table 3.3*).

Table 3.2 *DIC value of the leptospirosis and BU non-spatial vs spatial models*

Model	Disease dataset	Non-spatial model	Spatial model
Mean elevation	Leptospirosis	683.01	413.67
	Buruli Ulcer	626.81	542.97
Mean TWI	Leptospirosis	842.56	523.21
	Buruli Ulcer	875.98	676.46
Land cover at 2 km	Leptospirosis	665.18	574.21
	Buruli Ulcer	6.61	304.14
Land cover at 5 km	Leptospirosis	591.65	534.07
	Buruli Ulcer	16.73	301.28
Land cover at 10 km	Leptospirosis	626.14	512.82
	Buruli Ulcer	6.05	279.83
Minimum temperature	Leptospirosis	748.95	518.39
	Buruli Ulcer	928.31	674.07

Topological models

The spatial elevation models of leptospirosis and BU had lower DIC values than the non-spatial elevation models (Table 3). Elevation over 10 m (10–50 m, 50–100 m and 100–200 m) were negatively correlated to the presence of leptospirosis in the geographical area [95% CI− 0.1759 to − 0.0505, − 0.1690 to − 0.0831, − 0.1178 to − 0.0443 respectively]. Similarly, elevation at 10–50 m was inversely associated with the presence of BU [95% CI − 0.0946 to − 0.0301]. TWI was found to be a spatially dependent environmental driver with a significant positive correlation with the disease positivity ([95% CI 0.1353 to 0.2461 and 0.0651 to 0.1245] for leptospirosis and BU respectively).

Table 3.3 *Comparison of predictor variables of the different statistical models: logistic, Bayesian non-spatial and spatial models*

Model	Disease dataset	Logistic regression	Non-spatial MCMC model	Spatial MCMC model

Mean elevation	Leptospirosis	(−)	(−)	(−)
	Buruli Ulcer	(−)	(−)	(−)
Mean TWI	Leptospirosis	(+)	(+)	(+)
	Buruli Ulcer	(+)	(+)	(+)
Land cover at	Leptospirosis	(−) primary forest	(−) primary forest	(−) primary forest
2 km	Buruli Ulcer	(+) urban	(+) urban	(+) urban
		(+) cropland	(+) cropland	(+) cropland
Land cover at	Leptospirosis	(−) primary forest	(−) primary forest	(−) primary forest
5 km		(−) mangroves	(−) mangroves	(−) mangroves
		(−) urban	(−) urban	X
	Buruli Ulcer	(+) urban	(+) urban	X
		(+) cropland	(+) cropland	X
Land cover at	Leptospirosis	(−) mangroves	(−) mangroves	X
10 km		(−) primary forest	(−) primary forest	(-) primary forest
		(+) cropland	(+) cropland	(+) cropland
	Buruli Ulcer	(+) cropland	(+) cropland	(+) cropland
		(−) urban	(−) urban	X
		(−) mangroves	(−) mangroves	(−) mangroves
		(−) primary forest	(−) primary forest	(−) primary forest
Maximum tem-	Leptospirosis	X	X	–
perature	Buruli Ulcer	X	X	–
Minimum tem-	Leptospirosis	(+) 0 month	(+) 0 month	(+) 0 month
perature	Buruli Ulcer	(+) − 4 months	(+) − 4 months	X
Total precipita-	Leptospirosis	X	X	–
tion	Buruli Ulcer	X	X	–

*The (+) and (−) indicate the positive or negative correlation of the significant coefficients and "X" denotes non-significant coefficients

Land cover

The spatial models of leptospirosis achieved a better fit than those of BU in comparison to the non-spatial models for predictor variables, land cover at 2, 5 and 10 km. At a spatial buffer of 2 km, the presence of leptospirosis was inversely related to the presence of primary rainforest [95% CI − 5.4550 to − 2.926]. While at buffer of 5 km, the presence of mangroves [95% CI − 7.5813 to − 2.277] along with primary rainforest [95% CI − 8.5772 to − 4.834] negatively influenced the distribution of leptospirosis. At a buffer of 10 km, primary forest [95% CI − 9.1840 to − 5.3070] remained a negative predictor of leptospirosis while presence of cropland [95% CI 5.5302 to 18.2710] was found to have a positive influence. The distribution of BU on the other hand was found to be spatially independent based on the DIC values. However, the confidence intervals of the non-spatial models were wide and so the spatial models are reported instead given their methodological robustness. At spatial buffers of 2 km, urban land cover [95% CI 1.6300 to 2.934] along with cropland [95% CI 4.1979 to 6.888] were found to be a positive predictor of BU incidence. No significant variables were observed at buffer of 5 km. The presence of mangroves [95% CI − 7.7912 to − 2.3001] and primary rainforest [95% CI − 6.7331 to − 3.4672] were inversely associated with the presence of BU disease.

Meteorological covariates

Increase in the minimum monthly temperature (Tmin) was found to have a positive influence in the distribution of leptospirosis and BU. The leptospirosis spatial model demonstrated a significant association with the predictor variable; Tmin [95% CI 1.128 to 1.339] at month zero while in the BU spatial model, the regression coefficient was not significant. The other meteorological variables, maximum temperature and total precipitation, were found not significant in the non-spatial models.

Model combinations

For leptospirosis, the best fitting model combination was mean elevation, cropland at 10 km and Tmin at month 0, i.e. the month of potential infection. BU cases from the year 2000 was used for the model combinations due to the lack of landscape covariates for the earlier years. The best model fit was found to be mean elevation, cropland at 10 km, and primary forest at 10 km.

Discussion

To our knowledge, this is the first study to identify and compare the effects of biogeographic factors on the spatiotemporal distribution of two emerging bacterial diseases of aquatic origin at an ecological scale. Our main finding is that the statistical models demonstrate the significance of spatial structure in the distribution of the two diseases. The environmental covariates were found to significantly influence the spatial distribution of both diseases. The robustness of the spatial Bayesian models along with the narrow confidence intervals of the predictor variables and lower DIC values support our hypothesis that biogeographic factors influence the spatial distribution of the two diseases. The spatial structure of the two diseases correspond with the geographical distribution of Leptospira spp. and M. ulcerans in the environment, as described previously in FG (M. Combe *et al.*, 2019).

The top ranked model combination, ranked based on DIC, for both diseases included low elevation, high TWI and cropland at 10 km. This can be partly attributed to a higher proportion of population living in river basins of low elevation, which attracts human settlements providing easy irrigation for agriculture. However, the inverse association with increased elevation and the positive relationship with the TWI, demonstrates that both diseases show spatial predilection towards low-lying regions that are also prone to flooding. A prospective cohort study in Brazil showed that households at low elevation had a high leptospirosis infection risk (Hagan *et al.*, 2016). WHO reports the increased propensity of leptospirosis outbreaks following floods as flooding facilitates the spread of the pathogen through proliferation of rodents, which shed large amounts of pathogenic leptospires in their urine and thus increase the exposure to a susceptible population (‘WHO | Flooding and communicable diseases fact sheet’, 2012). The positive

relationship between cropland and leptospirosis incidence is congruent with previous studies on leptospirosis that demonstrated that croplands and associated farming practices also enhanced the rodent population leading to increased exposure with surface water and soil contaminated by rodent excreta (Martin *et al.*, no date; Kuriakose *et al.*, 2008; Dhewantara *et al.*, 2018). Other studies showed that freshwater bodies near rain- forests were hotspots for leptospirosis (Baker, 1965; Lau and DePasquale, 2012), that is also congruent with our observation of leptospirosis in villages found on the banks of the river Maroni that is surrounded with primary rainforest. The river basin and croplands were found to serve as a network that spatially attracts and concentrate small mammals that are potential reservoirs to leptospirosis, thus maintaining the bacteria in the region.

In our study, the site of exposure to the leptospira was unknown and so we assumed that the patients were infected in the vicinity of their residence as demonstrated by various studies (Barcellos and Sabroza, 2001; Rood *et al.*, 2017). This is however in contradiction to le Turnier et al. 2018, who hypothesized that occupation such as gold mining in proximity to the primary forest was the likely source of infection in French Guiana. If that were the case, our spatial models would not demonstrate a significant spatial structure but rather a random occurrence of cases. Our results are further supported by a recent study from French Guiana (M. Combe *et al.*, 2019), which carried out environmental microbiological sampling for Leptospira spp. in urban and rural areas in proximity to forests. The study observed that the bacteria were detected in modified urban ecosystems rather than in areas near forests.

Low elevation, river basins and agricultural activities were found to be also significant risk factors in the spatial distribution of BU. The BU cases were found occupy regions known to be prone to flooding in accordance to the cartographic regulatory document assessing the flood-prone areas in the commune known as the PPRI (Figure. 3). Studies have demonstrated the occurrence of BU disease outbreaks in West Africa and Australia associated with unprecedented flooding of rivers and lakes, damming of rivers and modification of wetlands into agricultural lands or recreational facilities (Merritt, Benbow and Small, 2005; Duker, Portaels and Hale, 2006). Flooding has been proposed to facilitate the transfer of

M. ulcerans among aquatic reservoirs by providing a potential route for inter-water body dispersion (Merritt, Benbow and Small, 2005; Combe *et al.*, 2017a). Previous studies in FG detected positive samples of *M. ulcerans* from freshwater bodies in the floodplains (Morris *et al.*, 2014). We observed a sharp decline in the incidence of cases in the region of Sinnamary following 1994, which corresponds to the construction of the Petit-Saut dam (***Figure 3.3***) on the Sinnamary River. The building of the reservoir has been shown to influence flooding, in this case with a potential reduction of exposure to the pathogen in a susceptible population resulting in 11.38% annual decrease in incidence (p-value: 0.001) as supported by previous work from FG (Douine *et al.*, 2017b). In West Africa, studies have demonstrated a positive association between BU incidence and agriculture (Brou *et al.*, no date; Marston *et al.*, 1995; Wagner, M Eric Benbow, *et al.*, 2008; Wagner *et al.*, 2008), and a case–control study in Benin showed that farmers were associated with an increased risk of BU (Wagner *et al.*, 2008). Overall, agricultural regions were found to have higher prevalence of BU due to an increase in nutrients favorable to biofilm growth and also decrease in dissolved oxygen content in surrounding freshwater bodies, which provides an ideal environment for *M. ulcerans's* persistence (Merritt *et al.*, 2010). Both bacteria occupying similar aquatic niche, despite of differing modes of transmission (***Figure 3.1***), show similar relationships with topological risk factors, suggesting that low elevation, flood-prone regions near croplands are common risk factors for these specific bacterial diseases.

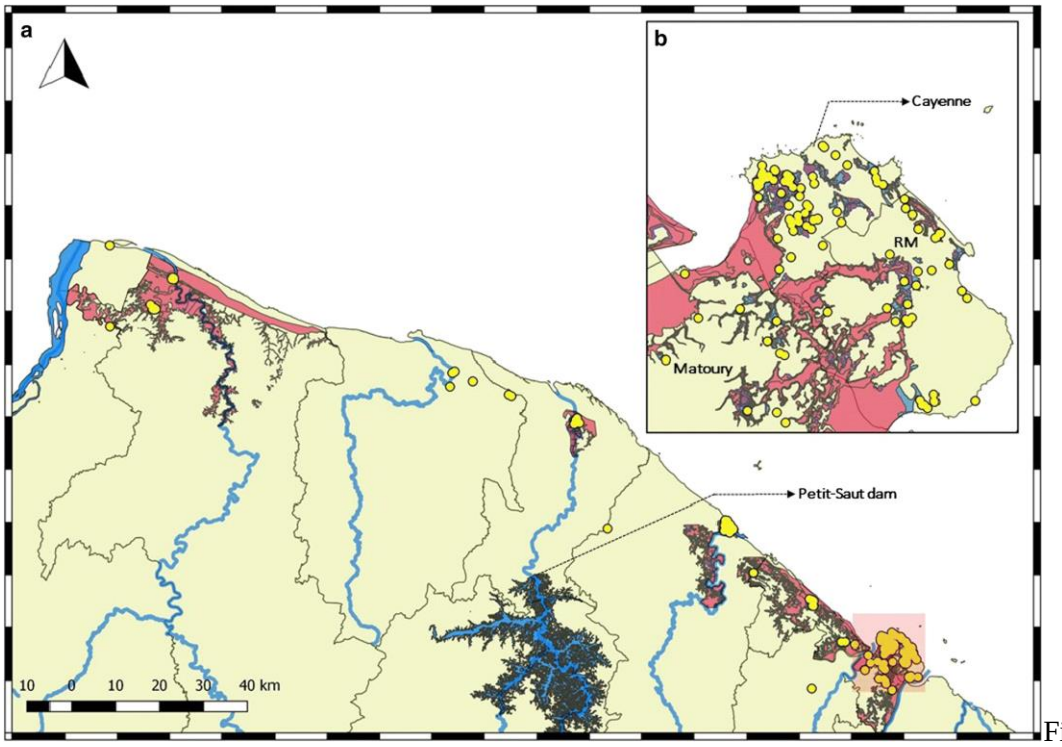

Figure.

Figure 3.3 *a) BU cases (in yellow) occupying flood-prone regions along the banks of the major rivers in French Guiana. b) BU cases in the flood-prone urban regions of Cayenne, Remire-Montjoly (RM) and Matoury with (1) zone red: regions in high risk of flooding, (2) zone blue: regions in average risk and (3) zone purple: regions of low risk.*

Interestingly our results demonstrate a positive relationship between urban land cover and BU incidence, which is in contradiction with other land cover studies conducted in West Africa on BU prevalence (Wagner *et al.*, 2008; Smithwick and Wu, 2016). The results from our datasets demonstrate that in the years following 2000, 79% of the new BU cases occurred in urban settings in comparison to 52% in the earlier years (79% vs 52%; p-value < 0.0001, 95% CI 14.96–37.94). This is supported by a recent study conducted in FG that also showed that modified urban ecosystems might favor BU emergence (M. Combe *et al.*, 2019). Also, the increasing number of cases in the urban regions of Victoria, Australia, provides a new perspective to BU incidence, which was previously thought to be restricted to rural areas (O'Brien *et al.*, 2018).

The increased urban incidence of BU was demonstrated by the sharp rise of the disease in the sub-urban populations, namely Rémire-Montjoly and Matoury (*Figure 3.3*) between the years 2002 and 2004. These regions were subjected to deforestation, modification of marshy lowlands into habitable areas during the early 2000s, as evidenced by the increase in annual rate of population growth by 3.3 and 7.8 between the years 1990–1999, 1.9 and 4.6 between 1999–2010 in Rémire-Montjoly and Matoury respectively. It is worth noting that these regions are more habited due to increased urbanization along the coastal regions resulting in large proportion of susceptible population. However, in contrast, the regions near the coast in Benin were found to have lower than expected BU prevalence, which was attributed to an access of pumped water sources in urban settings (Wagner, M. Eric Benbow, *et al.*, 2008). We propose that flood-prone regions associated with an increasing naïve population is at risk to develop BU due to the persistent maintenance of the bacteria in the environment as seen in Australia (Veitch *et al.*, 1997; van Ravensway *et al.*, 2012).

Whilst previous studies from FG, including a time- series analysis, report the influence of rainfall on both disease incidences (Morris., *et al.*, 2014; Cropet *et al.*, 2018), our study being a spatial biogeographical analysis did not found a significant correlation between disease incidences and rainfall. We observed that increase in minimum temperature during the time of potential infection could influence the prevalence of the diseases. Our results establish a positive relationship between increase in Tmin at − 4 months (m−4) prior BU diagnosis, which corresponds approximately to the time of potential infection (i.e. the incubation period of the disease) (Loftus *et al.*, 2018). A study from Australia also report higher BU disease incidence with Tmin conditions, with BU occurrence associated with Tmin at − 18 months. However, such a lag phase does not correspond to the incubation period of the bacteria. By comparison, leptospirosis was found to have a similar association to Tmin during the month of diagnosis. Increase in Tmin (1 °C) at a lag of 11 weeks was significantly associated with the increase in leptospirosis cases in the Republic of Korea (Joshi, Kim and Cheong, 2017). In our study, the association between increase in Tmin during the potential exposure period and disease incidence demonstrates that increase in minimum

temperature plays a significant role during infection. It is interesting to note that studies on climate change report warming trends in minimum Tmin over time due to greater heat accumulation with consistent Tmin increasing more than Tmax (Trenberth *et al.*, 2007). This establishes an indirect connection between climate change and increasing incidence of both diseases. However, a time-series analysis will be needed to analyze this relationship further and to forecast potential implications in the future emergence of both diseases under scenarios of climate change in FG. This research has demonstrated that two aquatic diseases of bacterial origin are spatially dependent at an ecological scale and a biogeographic approach is important in identifying the factors influencing the disease emergence and maintenance in a region. This approach is especially useful when information on the host and pathogen distribution are unavailable.

Conclusions

On comparing the environmental factors influencing the spatial distribution of two aquatic bacterial diseases of freshwater origin, we conclude that low-lying regions prone to flooding with nearby agricultural land and increased minimum temperature during the time of infection were found to be at risk for the increased incidence of both diseases. The trends of positive population growth rate in the urban regions of FG predict that deforestation and habitat fragmentation will continue to accommodate the needs of the growing population. Such human-driven regional environmental modifications along with global climate change affects vulnerable freshwater systems resulting in increased host–pathogen contact and ensure the maintenance of the aquatic bacteria at an ecological scale. Based on our results, we recommend the following to reduce the incidence of the two disease in developing tropical regions:

i. Better urban planning by construction in regions of low flood risk or low TWI, calculated from global satellite imagery. Regions of high flood risk need better drainage systems that would decrease human-pathogen contact.

ii. Croplands to be developed further away (over 10 km) from the population would also reduce human contact with rodents and with aquatic systems favorable for *M. ulcerans and Leptospira sp.*

iii. Increased minimum temperature during the time of infection signals the play of global environmental factors i.e. climate change. Climate change to be tackled at global scale to reduce the risk of disease emergence in tropical regions.

iv. Finally, conducting passive disease surveillance and measuring disease risk using biogeographic approach in regions where data on pathogen and reservoirs is scarce, is useful in the prevention and control of diseases.

3.2. Global emergence of Buruli Ulcer

Rapid ecological changes, underpinned by human activities or climate change, influence the geographic distribution of emerging pathogens(Jones *et al.*, 2008).Recent works on BU emergence, a neglected tropical disease of the skin and soft tissue, linked the infectious agent *Mycobacterium ulcerans* to deforestation of primary forests for agriculture and mining activities(Morris *et al.*, 2016). A clustering of BU cases in the alluvial gold mining towns of Kakerifu and Kasongo was observed in Democratic Republic of Congo (Janssens *et al.*, 2005). In addition, the increase of artisanal gold mining in the Birimian Greenstone belt during the 1980s correlated with the emergence of BU in West Africa(van der Werf *et al.*, 1989). In Ghana the Amansie West district, where the Ashanti gold mines are located, was the most BU-endemic region with 150.8 cases per 100000 people reported in 1999(Amofah *et al.*, 2002). The Ghanaian districts exhibiting Birimian meta-sedimentary rock, an Archean greenstone rich in arsenopyrites, were found more susceptible to BU, because arsenic contained in this rocks could lead to local human populations more susceptible to *M. ulcerans*(Mantey *et al.*, 2012).

Greenstone belts contribute to most of the gold mined in BU-endemic regions, notably the Precambrian Guinean/Birimian shield of West Africa and Guianas, and Lachlan Fold Belt of Victoria in Australia. Open-pit quarrying, the preferred method of gold mining from the Greenstone belts, involves deforestation and destruction of surface rocks to explore the greenstone layers, resulting thus in massive biodiversity loss. Gold is then extracted by cyanide heap leaching or amalgamation with mercury using freshwater. The resulting run-off containing high levels of heavy metals, including arsenic and sulfides, are released into the surrounding water bodies leading to accumulation of heavy metals in sediments, biofilm and aquatic communities followed with a rapid change of physicochemical parameters in the downstream aquatic ecosystems. The resulting low *pH* and anaerobic conditions favors the growth of *M. ulcerans* (Hagarty *et al.*, 2015; Combe *et al.*, 2017a), in addition to the indirect effect of deforestation and urbanization caused by mining.

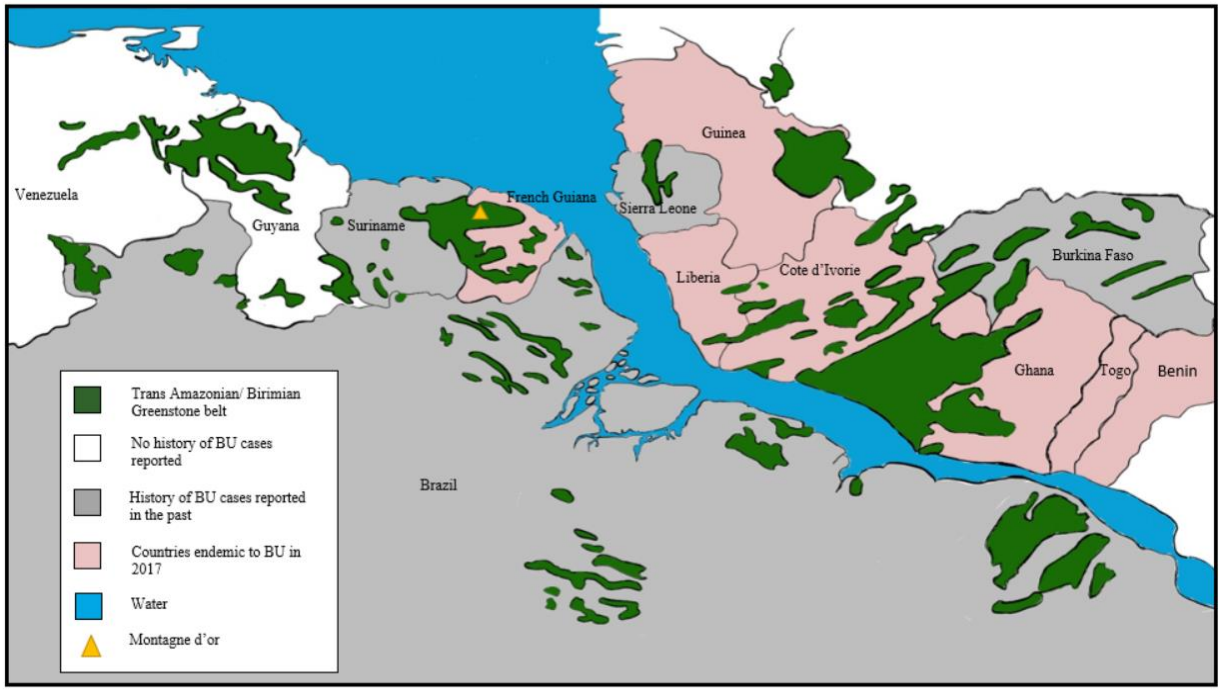

Figure 3.4 *Relationship between the spatial distribution of the Tranamazonian/ Birimian greenstone belt and the countries endemic to BU*

Montagne d'or is the largest gold mining project taken up in French Guiana (South America), which is presently highly debated due to its ecological impacts in one of the most biodiverse regions of the world, and its economic benefits. It involves an Archean greenstone deposit (*Figure 3.4*), located in between two natural reserves of the Amazonian rainforest and irrigated by the largest inland river, the Mana. Such project will result in severe biodiversity loss due to primary forest destruction and the increase in gold mining activity is expected to result in an increase of BU incidence in the drainage basin of river Mana, a region with previously reported cases of BU. In 2008, the French Government introduced "Operation Harpie" a military operation with the aim of reducing illegal gold-mining, notably those involving significant environmental damages. The number of BU cases noticeably decreased by 50% in the following years (i.e. 36 cases, 2009-2017*vs* 79 cases, 2000-2008 (Douine *et al.*, 2017a)). Similarly, an unprecedented rise by 72% of BU cases (182 new cases, the highest ever reported) was observed in 2016 in the Murray Basin, Victoria,

Australia (O'Brien *et al.*, 2018)and coincided with an increase in gold production rates in Victoria. Victoria being the state with highest number of cases reported in Australia.

The global demand for gold had a seven folds increase in the last decade. Millions of people in the developing countries depend on artisanal and small-scale gold mining for their livelihoods (World Health Organization, 2016). In 2010, the motion for a resolution urging the European Commission for a complete ban on cyanide leach-mining was initiated. The Commission rejected the proposal to implement the ban, arguing that closing of gold mines using cyanides would be "detrimental to employment"('Answer given by Mr Potočnik on behalf of the Commission', 2010)but ignoring the indirect implications of open pit mining on emerging infectious diseases in tropical countries. This lack of awareness and regulation may however drive new type of disease emergence in local communities. Gold mining is an understudied risk factor for BU emergence in endemic areas exhibiting suitable environmental conditions prone to *M. ulcerans* growth. Despite obvious economic interests, gold mining and in particular greenstone belts like in Africa, French Guiana or in Australia could drive the re-emergence of a debilitating disease at epidemiological levels as recently seen in Australia.

4.CHAPTER III

Spatial variations in the distribution of protozoal pathogens

Title of journal article

Spatial variations between Leishmania species: A biogeographic approach to mapping the distribution of Leishmania species in French Guiana

Authors: Soushieta Jagadesh, Marine Combe, Marine Ginoves, Pierre Couppié, Mathieu Nacher & Rodolphe Elie Gozlan

Journal: To be submitted to Journal of Biogeography (June 2020)

4.1. Spatial variations between Leishmania species: A biogeographic approach to mapping the distribution of Leishmania species in French Guiana

Abstract

Cutaneous Leishmaniosis (CL) is the most prevalent form of Leishmaniosis and is widely endemic in the Americas. Several species *Leishmania sp*. are responsible for CL, a severely neglected tropical disease. Treatments vary according to the different species of *Leishmania* responsible of the disease. We proposed to map the distribution of the *Leishmania sp.* reported in French Guiana using a biogeographic approach using environmental predictors. We also measured species endemism i.e. the uniqueness of the species to a defined geographic location. Our results show that the distribution patterns varied between *Leishmania sp.* The species distribution modelling of the eco-epidemiological spatial patterns of the *Leishmania sp.* is the first to measure endemism based on bioclimatic factors in FG. Precipitation seasonality and altitude were found to contribute the least in the distribution of the *Leishmania sp*. The study also emphasizes the impact of tree cover loss on the increasing distribution of *L. (V.) braziliensis* in the mort antropized regions.

Introduction

With the worldwide attention on novel emerging infectious diseases, research on 'older' diseases of high burden in the tropics remain limited. Due to a lack of global priority setting, these tropical diseases are often left neglected. In 2014, the WHO ranked Leishmaniosis as a group 1 emerging and uncontrolled disease Neglected Tropical Disease (NTD) (Mackey *et al.,* 2014). Cutaneous Leishmaniosis (CL), the most prevalent form of the disease, is widely distributed from the Indian subcontinent, across West Asia and the Middle East, through the Mediterranean and Northern Africa to Central and South America. An estimated 600,000 to 1,200,000 cases of CL occur annually (*Disease Background – Cutaneous Leishmaniasis – DNDi*, no date). CL is a dermal infection caused by protozoan parasites belonging to the genus *Leishmania* (family *Trypanosomatidae*) and is transmitted by infected female sandflies.

CL can be etiologically divided into Anthroponotic (old world) CL endemic to Asia and the Mediterranean caused by *L. tropica* and Zoonotic or new world CL limited to the Americas. It is caused by two subgenera of *Leishmania*: *Leishmania leishmania* complex and *Leishmania viannia* complex. The species of the *L. viannia* complex systematically disseminate in the skin and lead to clinical complications such as mucosal leishmaniasis (ML). This is particularly associated with *L. (V.) braziliensis,* where ML can develop in 2-5 % of patients with CL, days to years after recovery (Strazzulla *et al.*, 2013). ML destroys the mucous membranes of the nose, mouth and throat and can lead to long-term structural scaring and secondary bacterial infections (Meissner WHO, 2010). Systemic antileishmanial drugs are often used to treat CL caused by *L. (V.) braziliensis* species, to promote healing of the primary lesion and reduce the risk of developing ML (Lainson, no date; Herwaldt, 1999).

In FG, CL was described for the first time in 1954 (Floch, 1954). With over 95% of the country attributed to primary rainforests, FG hosts a high diversity of mammals that can act as reservoirs for *Leishmania sp.*. Five human *Leishmania* species have been described in FG: 86.9% *L. (V.) guyanensis*, 9.7% *L. (V.) braziliensis*, 2.8% *L. (L.) amazonensis*, 1.3% *L. (V.) lainsoni* and sporadic cases *L. (V.) naiffi*

(Simon *et al.*, 2017). Apart from the species-dependent clinical manifestations, a main challenge for public health care is that the various *Leishmania sp.* differ in terms of treatment sensitivities. Therefore, characterizing the distribution of the five *Leishmania sp.* across FG could prove to be significant to the management of clinical cases. Such environmental mapping is currently lacking.

Here, we mapped the distribution of the five *Leishmania sp*. in FG using a biogeographic approach using only environmental predictors. In addition to mapping leishmania species distribution and richness, we also measured species endemism i.e. the uniqueness of the species to a defined geographic location. To measure the impact of deforestation on the different *Leishmania sp*., we analyzed compared the tree cover loss around the occurrence sites with pseudo-absence sites.

Methods

Ethics statement

All patients were informed using written documents that case records and biological data might be further used in research and that they had the right to refuse. The retrospective anonymized case database was approved by the National Commission for Informatics and Liberties (CNIL; number 1805118v0).

Study area

FG is is split into the littoral, that consists of a thin coastal strip at the north of the territory and is more urbanized and the primary rainforests, where there are sparse settlements and sporadic illegal gold mining camps.

Patient data

We used for our analysis, the records of patients in consultation for a suspicion of leishmaniosis at the Cayenne General Hospital and associated administratively health centers between January 1994 and January 2015. New cases were defined as cases without a history of leishmaniosis in the previous 12 months. Diagnosis was confirmed by the presence of *Leishmania* parasites on microscopic examination of May-

Grünwald Giemsa-colored skin smears, and/or by a positive culture on RPMI culture medium, and/or by the detection of *Leishmania* DNA by PCR. Species identification was done by PCR-RFLP retrospectively for the cases before 2006 (Simon *et al.*, 2017). Not all cases could be identified due to insufficient *Leishmania* DNA quantity.

Location of exposure

The patient database contained the details on the location of potential exposure including the distance from the road or river in cases of forest chalets or gold mining camps. The sites of exposure were georeferenced using IGN (l'Institut National de l'Information Géographique et Forestière/ National Institute of Geography and Forestry) maps and projected at WGS84 and this projection was maintained throughout the study.

Environmental variables

We extracted the climatic covariates at 2.5 minutes spatial resolution from Worldclim database (Booth *et al.*, no date). All bioclim variables from one to nineteen was used in the analysis (*Table 4.1*). The topological data was extracted from one arc-second digital elevation model (DEM) of 30 m resolution, derived from NASA (Version 3.0) Shuttle Radar Topography Mission (SRTM) imagery available in the United States Geological Survey (USGS). The raster layers were then resampled to a fixed resolution of approximately 4km^2 and stacked to a raster stack.

Table 4.1 *Relative variable importance of the individual SDM and SSDM models*

Bioclimatic variables	*AMAZONENSIS SDM*	*BRAZILIENSIS SDM*	*GUYANENSIS SDM*	*LAINSONI SDM*	*SSDM mean [95% CI]*
Altitude	**4.72**	**4.61**	**4.72**	4.79	**4.71 [4.7- 4.72]**
Annual Mean Temperature	4.77	4.80	4.77	4.79	4.78 [4.78 -4.78]

Mean Diurnal Range	**4.72**	4.78	4.77	4.79	4.76 [4.75 -4.77]
Isothermality	4.77	4.78	4.74	4.79	4.77 [4.77-4.77]
Temperature Seasonality	4.77	4.80	4.75	**4.25**	4.64 [4.59- 4.69]
Max Temperature of Warmest Month	4.77	4.79	4.73	4.79	4.77 [4.76- 4.78]
Min Temperature of Coldest Month	4.77	4.80	4.77	4.79	4.78 [4.78- 4.78]
Temperature Annual Range	4.76	4.77	4.77	4.79	4.78 [4.78- 4.78]
Mean Temperature of Wettest Quarter	4.77	4.80	4.78	4.79	4.78 [4.78- 4.78]
Mean Temperature of Driest Quarter	4.76	4.80	4.77	4.79	4.78 [4.78- 4.78]
Mean Temperature of Warmest Quarter	4.77	4.80	4.78	4.79	4.78 [4.78- 4.78]
Mean Temperature of Coldest Quarter	4.77	4.80	4.78	4.79	4.78 [4.78- 4.78]
Annual Precipitation	4.77	4.76	4.77	4.79	4.77 [4.77- 4.77]
Precipitation of Wettest Month	4.76	4.79	4.75	4.79	4.77 [4.77- 4.77]
Precipitation of Driest Month	4.77	4.77	**4.72**	4.79	4.76 [4.75- 4.77]
Precipitation Seasonality	4.77	**4.49**	4.77	4.79	**4.70 [4.68- 4.73]**

Precipitation of Wettest Quarter	4.77	4.79	4.76	4.79	4.78 [4.78- 4.78]
Precipitation of Driest Quarter	4.77	4.78	4.77	**4.71**	4.76 [4.75- 4.77]
Precipitation of Warmest Quarter	4.77	4.74	4.77	4.79	4.77 [4.77- 4.77]
Precipitation of Coldest Quarter	4.77	4.75	4.77	4.79	4.77 [4.77- 4.77]

Fitting of model and Model prediction

The individual species distribution models (SDMs) and stacked species distribution models (SSDMs) were fitted using the r package "SSDM". In our study, we used Maximum entropy model (Maxent) for the construction of the SDM. Maxent is a popular choice for SDM models using presence-only data. Finally, we built SSDMs using the same algorithm and the five *Leishmania* species. The outputs of the different species are aggregated in SSDM maps of local species richness and a weighted endemism index (WEI) was used for the endemism map. All models were evaluated using ROC curves and the area under curves (AUC) for the produced thresholds were calculated.

Tree-cover loss analysis

To analyze the impact of deforestation on the distribution of the *Leishmania* species, we conducted a detailed spatiotemporal analysis using the R package "gfcanalytics" on the Hansan's tree cover loss maps. We used the presence and generated pseudo-absence points from 2001-2016 at a spatial buffer of 15km, for this analysis. Regression analysis was done for each of the *Leishmania* sp. to establish the relationship between deforestation and the presence of the pathogen.

Results

We georeferenced 470 exposures of *Leishmania sp*., which resulted in 107 geographically resampled, presence-only occurrences. We modelled four of the five *Leishmania sp.* observed in FG. The occurrences of L. *(V.) naiffi* were scarce (1/470; 0.93%) and could not be resampled to be used in the training process. The composition of the training data included *L. (L.) amazonensis* (8/107; 7.48%), *L. (V.) braziliensis* (18/107; 16.82%), *L. (V.) guyanensis* (76/107; 71.03%), and *L. (V.) lainsoni* (5/107; 4.67%). All models used MaxENT as the base algorithm.

AMAZONENSIS SDM

The SDM model of *L. (L.) amazonensis* (*Figure 4.1*) was evaluated and measured to have an AUC of 0.982 and a kappa of 0.034. The model sensitivity and specificity of 100% and 96% respectively. All bioclimatic variables excepting altitude and mean diurnal range contributed equally to the model construction (*Table 4.1*).

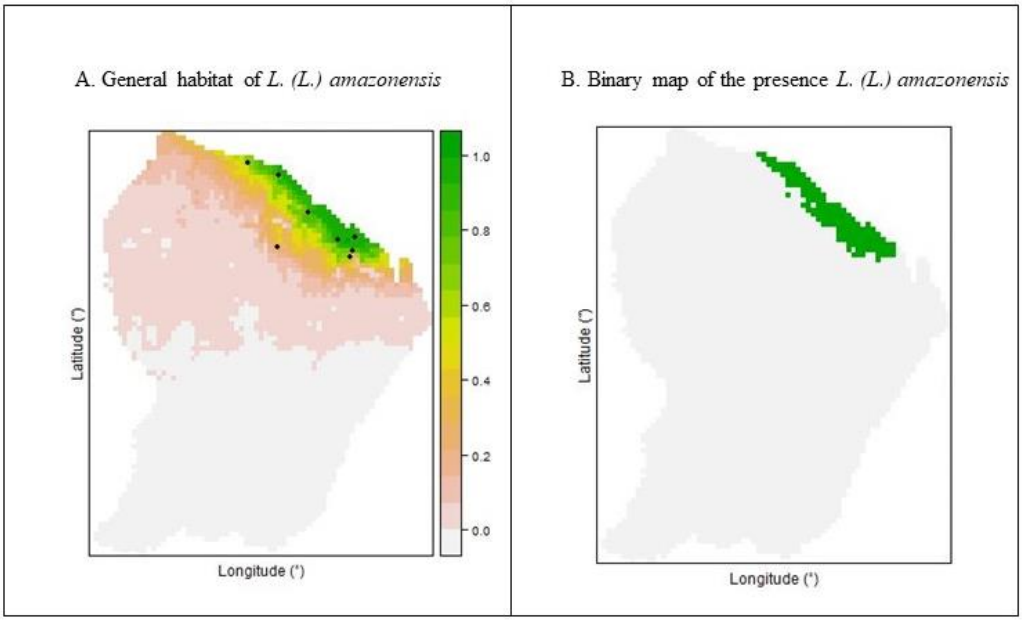

Figure 4.1 *Distribution map of L. (L.) amazonensis in French Guiana, South America. Panel A – demonstrating the general habitat suitability of L. (L.) amazonensis, with black dots indicating L. (L.) amazonensis occurrences ;Panel B: representing the presence of L. (L.) amazonensis*

BRAZILIENSIS SDM

The predictive risk of *L. (V.) braziliensis* is mapped in *Figure 4.2* The evaluation of SDM model observed an AUC of 0.632, kappa 0.003 and sensitivity and specificity of 60% and 66% respectively. Precipitation seasonality contribute to a lesser extent in comparison to the other bioclimatic variables (*Table 4.1*).

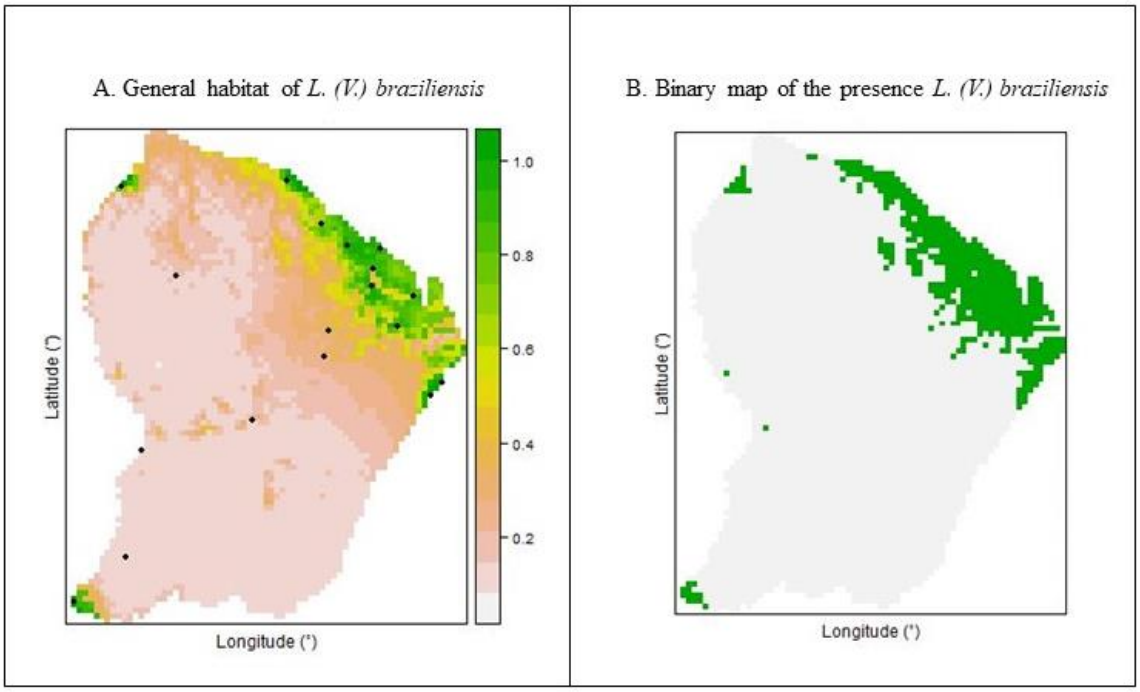

Figure 4.2 *Distribution map of L. (V.) braziliensis in French Guiana, South America. Panel A – demonstrating the general habitat suitability L. (V.) braziliensis, with black dots indicating L. (V.) braziliensis occurrences; Panel B: representing the presence of L. (V.) braziliensis*

GUYANENSIS SDM

SDM model of *L. (V.) guyanensis* mapped the predictive risks of the potential sites of exposure (*Figure 4.3*) The evaluation of the model observed an AUC, kappa, sensitivity, and specificity

of 0.778, 0.036, 78% and 78% respectively. All bioclimatic variables excepting altitude and precipitation at the driest quarter contributed equally to the model construction (*Table 4.1*).

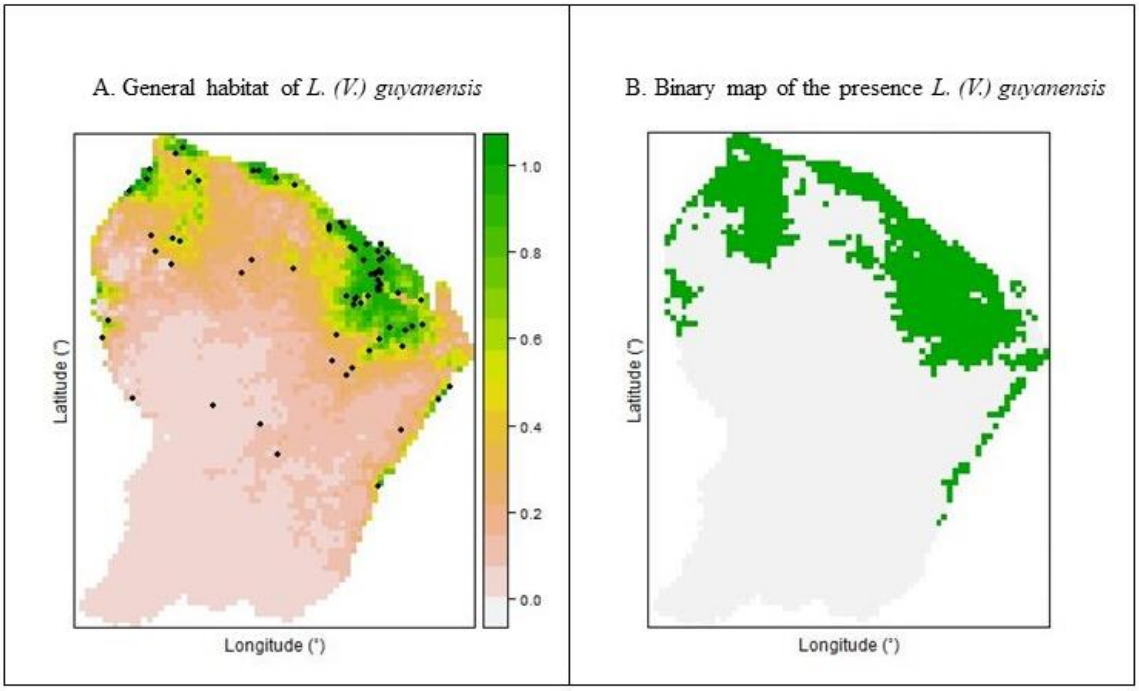

Figure 4.3 *Distribution map of L. (V.) guyanensis in French Guiana, South America. Panel A – demonstrating the general habitat suitability L. (V.) guyanensis, with black dots indicating L. (V.) guyanensis occurrences; Panel B: representing the presence of L. (V.) guyanensis*

LAINSONI SDM

The predictive risk of *L. (V.) lainsoni* illustrated in *Figure 4.4*.The evaluation of SDM model observed an AUC, kappa, sensitivity, and specificity of 0.976, 0.013, 100% and 95% respectively. All climatic variables contributed equally to the model construction. Temperature seasonality and precipitation at the driest quarter contributed to a lesser extent in comparison to the other bioclimatic variables (*Table 4.1*).

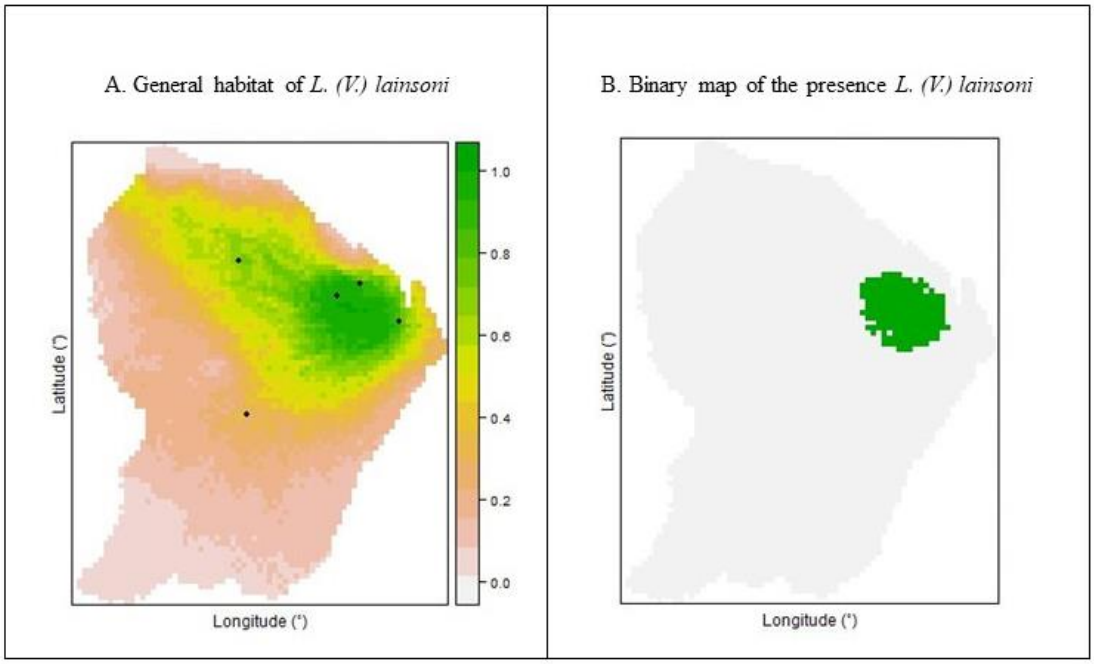

Figure 4.4 *Distribution map of L. (V.) lainsoni in French Guiana, South America. Panel A – demonstrating the general habitat suitability L. (V.) lainsoni, with black dots indicating L. (V.) lainsoni occurrences; Panel B: representing the presence of L. (V.) lainsoni*

SSDM of *Leishmania sp*.

The SDM models were stacked and outputs of the four *Leishmania sp.* were aggregated in SSDM maps of species richness and composition using the summing continuous habitat suitability maps stacking method (pSSDM). The local *Leishmania* species richness and endemism maps are illustrated in *Figure 4.5*. We observed a mean species richness error of 0.654 and prediction success, sensitivity and specificity of 70%, 80% and 67% respectively. The community similarity calculated by the Jaccard index was 4.8%. Precipitation and temperature seasonality along with altitude contributed to a lesser extent than other bioclimatic variables (*Table 4.1*).

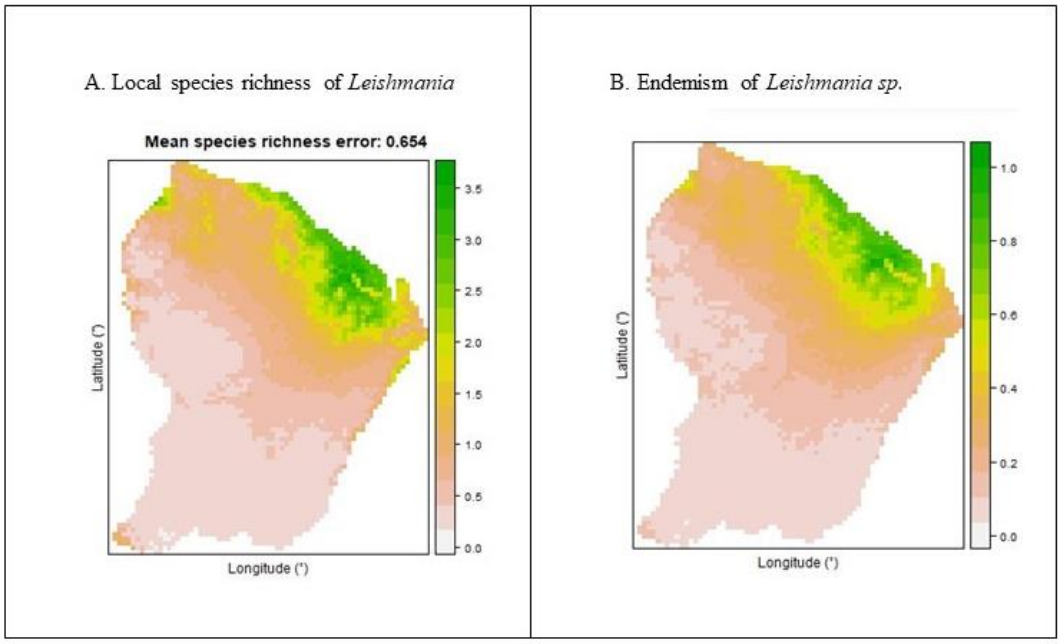

Figure 4.5 *Distribution of the four Leishmania sp. in French Guiana. Panel A: illustrating the local species richness of Leishmania sp.; Panel B: demonstrating the Endemism map of Leishmania.*

Deforestation analysis

Occurrences of *L. (V.) guyanensis* (3050.21; 95% CI [2251.44, 3848.98]) and *L. (V.) braziliensis* (1261.04, 95% CI [781.22, 1740.86]) were found to have higher tree cover loss at the radius of 15km than the pseudo-absence points. While the tree cover loss around the regions of *L. (L.) amazonensis* 368.66, 95% CI [-96.22, 833.54].and *L. (V.) lainsoni* 111.26, 95% CI [-158.81, 381.33] occurrence were not found to be significant in comparison with the pseudo-absence points.

Discussion

The major finding of this study is the presence of variations in the spatial patterns and clusters, suggesting that each *Leishmania sp.* has a specific climate niche and potentially different or exclusive vectors and reservoirs. The northeastern region of FG was found to have the highest species richness (*Figure 4.6*). Cases with reported exposure from this region must be subject to molecular

diagnosis for species identification prior to treatment. Especially for travelers visiting this region. This can limit treatment failure especially with *L. (V.) braziliensis* usually fail to respond to pentamidine isethionate, the first-line treatment of *L. (V.) guyanensis* CL in French Guiana and instead, relies on parenteral meglumine antimoniate or liposomal amphotericin B (Schwartz, Hatz and Blum, 2006). Evidence on the seasonal distribution of the *Leishmania sp.* is scarce but might provide an insight if there are temporal variations between the different species. Our species distribution modelling of the eco-epidemiological spatial patterns of the *Leishmania sp.* is the first to measure endemism based on bioclimatic factors in FG. Precipitation seasonality and altitude were found to contribute the least in the distribution of the *Leishmania sp.* This can be explained by the influence of environmental predictors on the disease vector, sandflies, rather than the pathogen itself. The study also highlights the impact of deforestation on the increasing distribution of *L. (V.) braziliensis* in FG.

Studies prior to 1980 report no evidence of *L. (V.) braziliensis* in FG (Dedet, Pradinaud and Gay, 1989; Desjeux and Dedet, 1989). In the recent decade, there has been an increase in *L. (V.) braziliensis* cases in FG. Our results demonstrate that *L. (V.) braziliensis* was the second most common *Leishmania sp.* in FG with its distribution along the coast and clusters along the Oyapock, Maroni Rivers and deep in the primary rainforests at the tri-country border (shared between FG, Suriname and Brazil). Studies hypothesized that the increase in *L. (V.) braziliensis* cases was due to the increasing presence of humans in the deeply forested areas where transmission of *L. (V.) braziliensis* usually occurs (Loiseau *et al.*, 2019) as cases occurred mostly amongst illegal gold miners and military personal. This explains the cluster of cases at the tri-country border. However, our study findings note that the spatial distribution of *L. (V.) braziliensis* along the coast and more anthropized areas. This can be explained by the high tree cover loss in the proximity of *L. (V.) braziliensis*, which lead to increase contact of the vector and the species reservoir with the humans.

Deforestation is a controversial risk factor incidence of Leishmaniosis. Some studies suggest that continued deforestation lengthens the distance between habitations and forest, thus decreasing the contact between host and the vector (Rodrigues *et al.*, 2019). While other argue that deforestation increase the exposure of the vector to human habitations (Desjeux, 2001; Rosales, Yang and Avila Blas, 2014). Our study found that *L. (V.) braziliensis* along with *L. (V.) guyanensis* were found to occur in regions with high tree cover loss. Increasing cases of *L. (V.) braziliensis* in the anthropized regions along with high tree cover loss suggests increased exposure of the sandfly vector with potential domestic reservoirs such as dogs leading to the establishment of a peridomestic cycle. The potential presence of a peridomestic cycle was hypothesized in a previous study (Martin-Blondel *et al.*, 2015). This is supported by the frequent finding of dogs with cutaneous lesions due to L. (V.) braziliensis leading to the conclusion that severe deforestation led to the migration of infected wild animals (particularly rodents) to residential areas in Brazil (Rangel and Lainson, 2009). Domestic cats have also been identified as a reservoir of *L. (V.) braziliensis* in FG (Rougeron *et al.*, 2011).

The impact of deforestation on leishmania distribution is further emphasized by the restricted spatial spread of *L. (V.) lainsoni*. Tree cover loss was not found associated with *L. (V.) lainsoni*, which was found in a deeply forested region as opposed to three other widespread *Leishmania sp. L. (L.) amazonensis*, found distributed along the coastal strip in this study, was also found to be independent of deforestation. This may be explained by two distinct cycles of leishmaniases occurring at different levels of the rainforest canopy were described in FG (Dedet, Pradinaud and Gay, 1989). The *L. (V.) guyanensis* cycle, located in the canopy with the arboreal sand fly *Lutzomyia umbratilis* as the vector and mammal reservoir of the canopy, the two-toed sloth. Tree cover loss forces the vector and reservoir to migrate, increasing the exposure to humans. While the other cycle with *L. (L.) amazonensis* occurs at ground level with *Lu. flaviscutellata* as the vector and *Proechymys cuvieri*, a rodent, as the main reservoir host (Rotureau, 2006). This results in the continued transmission at the ground level despite tree cover or canopy loss.

Our results show that the environmental variables found to contribute to the spatial models of Leishmania. Precipitation seasonality, the tendency for a place to have more rainfall in certain months, was found to have lesser influence in the distribution of certain species and the stacked model. A decrease in rainfall is linked to increased cases 2 months later in FG (Roger *et al.*, 2013). Precipitation seasonality was also found to contribute the least to the distribution of the Phlebotomine sandflies in central Europe along with precipitation at the warmest and coldest quarter (Koch *et al.*, 2017). Altitude was found to contribute unanimously less across the models. This concurs with studies that demonstrate the higher risk of other *Leishmania sp.* in low-lying regions (Ostyn *et al.*, 2015; Galgamuwa, Dharmaratne and Iddawela, 2018). This demonstrates that the bioclimatic predictors influence the sandfly vectors rather than the pathogen. New world sandflies of the genus *Lutzomyia* have been reported in a wide range of altitudes (Gomez *et al.*, 2014). The decrease in Leishmania exposure in high altitudes in FG is due to the poor access and absence of human settlements at the high-altitude regions.

This study shows the use of SSDMs in analyzing the spatial patterns of different species and provides a future scope for studies exploring spatial variations in malarial strains or between the different clades of HIV in a region. Studies looking into the variations in the seasonal distribution of *Leishmania sp.* would help us understand the transmission dynamics of the different species better.

Conclusion

In our study with the use of a biogeographic approach, we were able to characterize the spatial variations in the distribution of *Leishmania sp* and map the predictive risk based on bioclimatic factors in FG. The presence of a peridomestic cycle of *L. (V.) braziliensis* was confirmed. We also analyzed the role of tree cover loss in the distribution in the different *Leishmania sp.* SSDMs were found to be effective in detecting regions in FG with high species richness, which is particularly useful in the clinical management of the cases.

5.CHAPTER IV

Socioeconomic factors influencing the distribution of a fungal disease in a vulnerable population, French Guiana.

Title of journal article

Mapping priority neighborhoods: A novel approach to cluster identification in HIV/AIDS population

Authors: Soushieta Jagadesh, Marine Combe, Pierre Couppié, Rodolphe Elie Gozlan & Mathieu Nacher

Journal: Submitted to BMC Public Health (6th February 2020)

DOI preprint: 10.21203/rs.3.rs-17843/v1

5.1. Mapping priority neighborhoods: A novel approach to cluster identification

Abstract

Background: Urban disadvantaged neighborhoods have higher HIV risk behavior and higher levels of AIDS-related mortality. Studies demonstrate that interventions at the community level focusing on risk groups have increased success rates than individual patient-based management in the context of HIV/AIDS. We tested a novel approach to identify population groups in need of greater public health efforts to achieve UNAIDS 90-90-90.

Methods: We extracted retrospective data on 2141 HIV/AIDS patients, recruited from 1997-2017 in the regional hospitals in French Guiana. Self-organizing maps were constructed and clusters were identified based on demographic and socioeconomic variables such as age, sex, CD4 counts at Nadir, type of neighborhood, unemployment rate, and presence of opportunistic illness such as Histoplasmosis and Hepatitis B in the sample population.

Results: Neighborhood unemployment rates were identified to have a large impact in the distribution of HIV/AIDS. Also, the risk of disseminated histoplasmosis, the most common AIDS-defining illness in French Guiana, was not associated to any particular neighborhood suggesting that urban socioeconomic features are not the primary drivers of exposure risk.

Conclusion : Socioeconomically disadvantaged neighborhoods remain hotspots for HIV/AIDS. We conclude that SOM is an effective tool in the identification of risk clusters that may guide public health efforts to optimize HIV prevention and testing in French Guiana and other developing country

Introduction

The residential environment, particularly disadvantaged neighborhoods, has been shown to affect individual behavior and health through direct or indirect socioeconomic processes (Mayer and Jencks, no date). Neighborhoods have emerged as a determinant of public health providing both objective physical infrastructure and social characteristics that influence individual and community health (Koh *et al.*, 2010). Exploring neighborhood effects is thus important as it looks at community-based interventions to promote population health and health equity in regions with racial/ethnic minorities.

Characteristics of disadvantaged neighborhoods often correlate with individual HIV risk behavior (Latkin *et al.*, 2013). High school dropout rates, low employment rate, substance and alcohol abuse, increased street violence usually correlate with greater HIV risk (Singer *et al.*, 2006). Furthermore, neighborhoods with such features clustering exhibited higher HIV mortality and delayed anti-retroviral therapy (ART) initiation (Latkin *et al.*, 2013).

Histoplasmosis is one of the most frequent, often overlooked endemic disseminated mycosis in HIV patients in the Americas (Nacher *et al.*, 2016, 2019). The disease caused by the inhalation of the microconidia and mycelial forms of a fungus, Histoplasma capsulatum var. capsulatum (HC), found in guano-enriched soils. In French Guiana (FG), histoplasmosis is the most common AIDS-defining event and the leading cause of AIDs related deaths (Nacher *et al.*, 2014), with 75% of HIV-infected persons being foreign citizens. The fight against HIV has struggled to reach the undiagnosed reservoir with 30% of patients in Cayenne and 50% in Saint Laurent du Maroni, the two major cities of FG, diagnosed with advanced HIV-disease, with certain population groups particularly at risk of late testing. Although individual epidemiologic risk factors are important, spatialized community approaches to HIV programs seem to have an operational advantage. We thus tested the hypotheses that the distribution of HIV/AIDS and that of AIDS-related histoplasmosis varied between neighborhoods using self-organized maps, a two-dimensional data visualization tool that is trained using an unsupervised process.

Methods

Study Settings

In January 2015, France rolled out "le Quartier Prioritaire de la Politique de la Ville (QPPV)" or priority district of the city's policy to identify and improve the socially disadvantaged areas in France. Neighborhoods were identified by the poverty rate defined by INSEE (Institut National de la Statistique et des Études Économiques) as the proportion of the population living in regions that are under 60% of the median metropolitan living standards. In FG, the QPPV identified 32 priority neighborhoods in the urban, suburban, and peri-urban regions. In our study, we refer to the QPPV neighborhoods as urban disadvantaged neighborhoods.

In FG, all patients diagnosed with HIV receive free ART regardless of their socioeconomic status or country of origin. HIV care is accessible and is comparable to mainland France.

Patient Data

HIV-positive patients in follow-up at the hospitals of major cities in FG between April 1995 and May 2017 were enrolled in the French Hospital Database for HIV (FHDH).

Self-organizing maps (SOMs)

We used self-organizing maps (SOMs) to identify clusters who would benefit most from public health measures. SOM are a type of artificial neural network that consists of an array of units, called nodes, arranged in a fixed position on a grid. The key feature of the SOM is that the topology of the original input data is conserved, i.e. similar variables are grouped together on the map. We trained the SOM model of 2120 nodes and variables (n=9) with 2000 iterations at a learning rate between 0.05 – 0.01 to reach a minimum plateau. The nodes were modeled on a 20x20 grid with a hexagonal topology. Each node represented an HIV patient and his/her neighborhood with its position fixed in the grid. The variables included age, sex, CD4 counts at Nadir, neighborhood type, CDC classification of HIV, histoplasmosis/hepatitis

positivity, unemployment rate and neighborhood security. All SOM statistics were performed using the 'kohonen' package in R version 3.6.1.

Ethics

All patients enrolled in the FHDH gave written informed consent to the use of data for research. The data is anonymized and encrypted before transfer to the Ministry of health and the Institut National de la Recherche Médicale (INSERM), which centralize data from Regional Coordination for the fight against HIV (COREVIH) across France. This cohort has been approved by the Commission Nationale Informatique et Libertés (CNIL) since 1992 and has been used for numerous publications.

Results

Sociodemographic characteristics

We recruited a sample population of 2141 patients diagnosed with HIV between 1997 and 2017 in FG. The sample consisted of 1101 (51.4%) male and 1040 (48.6) female patients with an average age of 37.4 ± 0.26 years at the time of diagnosis. The patients were classified based on CDC classification system of HIV infection; A (n=1256 [58.7%]), B (n=249 [11.6%]), and C (n=636 [29.7%]) and had an average CD4 count of 205.7 cells/mm3. Our dataset included cases from rural regions (n=356 [16.6%]) and socioeconomically disadvantaged (n=1455 [68%]) and favorable (n=330 [15.4%]) neighborhoods. The HIV patients from disadvantaged neighborhoods had 1.33 times [95% CI: 1.01 – 1.75; P 0.04] higher odds of progressing to AIDS in comparison to the favorable ones. We found no significant risk of developing histoplasmosis [0.66 – 1.82; 0.74] in HIV patients inhabiting poorer neighborhoods.

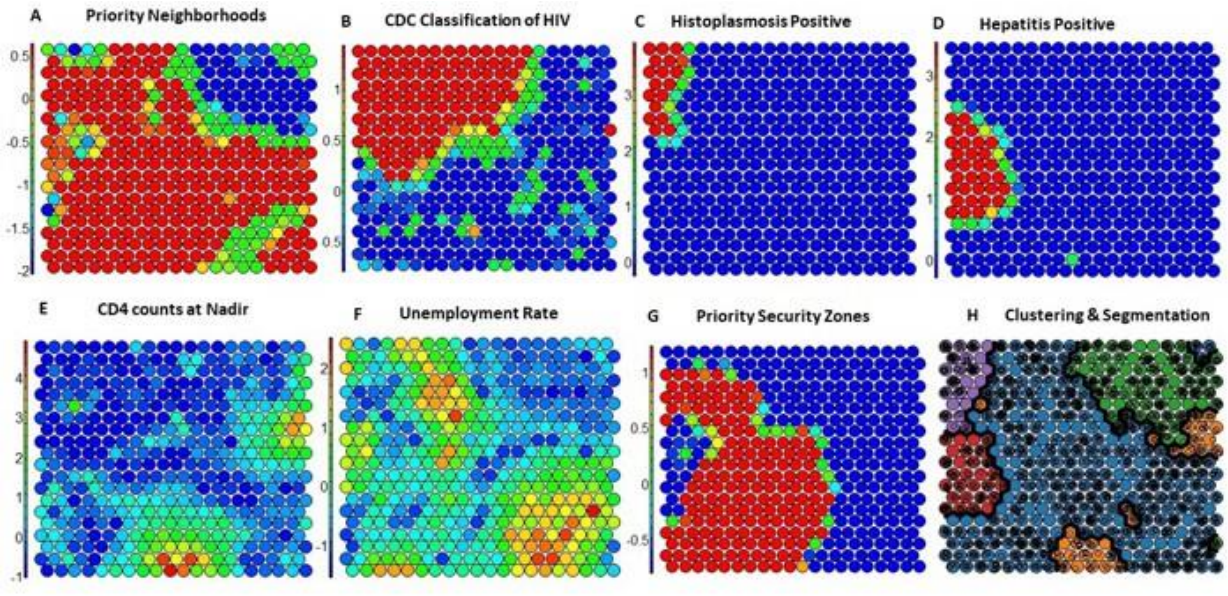

Figure 5.1 *Self-organizing maps (SOM) model of 2120 nodes on a 20x20 grid. The spectrum of colors represented in the y-axis represents the scale of the map. A) demonstrates the structure of HIV/AIDS population where the patient living in urban disadvantaged neighborhoods are illustrated (in red), in rural (yellow-green spectrum) and favorable urban neighborhoods (blue); B-E) illustrate the spectrum of severity of HIV/AIDS; F & G) illustrate the neighborhood unemployment rate and insecurity; H) represents the hierarchical agglomerative clustering and segmentation on the SOM model.*

SOM model and clusters

The heatmaps illustrate the distribution of seven variables in the sample population (*Figure 5.1 A-G*). We used hierarchical agglomerative clustering (HAC) to determine the optimal number of clusters in the sample population (*Figure 5.2*). The demographic and socio-economic clusters in the HIV-related histoplasmosis population in FG is visualized in Fig. 1H. We observed that the largest cluster (*Figure 5.1 H*; blue; n=1365) included people living with HIV/AIDS (PHLA) with a low employment rate living in urban disadvantaged neighborhoods and rural areas. The cluster also included regions where crime and insecurity are higher than in richer neighborhoods. The significantly smaller second cluster (*Figure 5.1 H*; green,) was represented by cases from favorable suburbs with a higher employment rate and relatively higher proportion of cases belonging to CDC class A. The histoplasmosis (*Figure 5.1 H*; purple; n=113) and hepatitis (red; n=141) cases were grouped in separate clusters. The hepatitis and histoplasmosis clusters had significantly higher males ($P < 0.0001$), increased mortality ($P = 0.009$, $P = 0.002$), and lower CD4

counts at nadir (P = 0.008, P <0.0001), respectively. The descriptive statistics are detailed in Appendix Table 5.1. The smallest cluster (*Figure 5.1*; orange; n=109) were cases with higher CD4 counts at nadir.

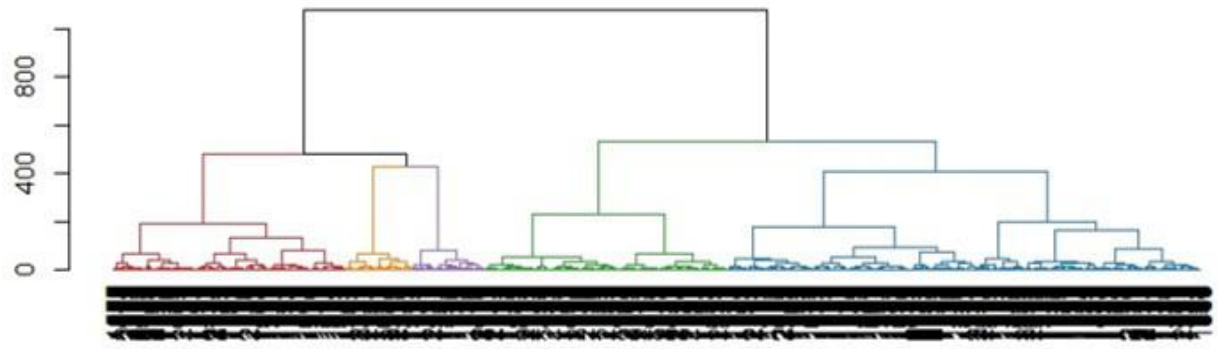

Figure 5.2 *Cluster dendrogram to determine the optimal number of clusters in the SOM model. The colors correspond to the colors of the SOM model.*

Discussion

Here, for the first time, we mapped the neighborhood effect on HIV/AIDS using self-organizing maps and showed that the spatial boundaries of HIV/AIDS distribution follow those of neighborhoods. Our second hypothesis that HIV-related histoplasmosis risk differed by neighborhood was rejected as the risk of developing histoplasmosis in HIV/AIDS was similar in both the socioeconomically privileged and underprivileged neighborhoods.

We found unemployment to have a higher structural impact than the other tested variables on the identified clusters. Studies demonstrate that unemployed HIV patients had an increased risk of HIV mortality and disease progression than those with stable employment in the HAART era (Maruthappu *et al.*, 2017). This coincides with our results, which showed that HIV-infected persons in economically poorer neighborhoods had higher odds of progressing to AIDS. Unemployed individuals have lower access to -and/or- underutilize preventive healthcare services than their employed counterparts (Tefft and Kageleiry, 2014), which potentially delays HIV diagnosis and treatment. Furthermore, economic insecurity amongst youth,

particularly women, has been shown to increase the HIV burden in developing countries including FG (Nacher *et al.*, 2010; Austin, Choi and Berndt, 2017).

Our results showed that neighborhood insecurity formed a sizable portion of the largest cluster. Research has shown that neighborhood crime was associated with HIV risk taking behaviors such as unprotected sex and multiple sexual partnerships (Nacher *et al.*, 2010; Ojikutu *et al.*, 2018). Furthermore, a recent study observed that HIV-related mortality reached over 8% of released prison inmates in FG (Huber *et al.*, 2017).

UNAIDS aims for 90% of HIV-infected persons aware of their diagnosis, 90% of those diagnosed on antiretroviral treatment and 90% of these virologically suppressed by 2020. However, the first 90% seems hard to reach in most countries and notably in FG where the proportion of patients diagnosed each year with advanced HIV disease remains stable despite efforts to scale up and diversify HIV-testing. The present approach may bring strategic and operational insights to improve the capacity to reach the hidden reservoir of undiagnosed infections (Nacher *et al.*, 2018).

We observed that the risk of histoplasmosis in PLHA was similar in urban neighborhoods and rural regions. Thus, it seemed that the incidence of histoplasmosis in FG was ubiquitous, influenced by the environmental distribution and endemicity of spores rather than socioeconomic factors or microenvironmental nuances. In our study, histoplasmosis, the first AIDS-defining infection in FG, was a significant cluster in PLHA (Nacher *et al.*, 2011). Our results on the histoplasmosis cluster is consistent with previous studies from FG (Nacher *et al.*, 2014), where the risk factors include male sex and <50 cells/mm3 CD4 at nadir. We also observed higher mortality in histoplasmosis and hepatitis clusters, which demonstrates the significance of cluster identification for community-level management of HIV/AIDS.

We did not analyze the environmental factors influencing the distribution of histoplasmosis, providing thus scope for future studies to analyze the impact of soil acidity and rainfall on disease incidence. In addition, the lack of socioeconomic variables, such as built environment, is another limitation of our study.

Conclusions

We conclude that SOMs applied to HIV/AIDS cases are an effective tool to identify and prioritize clusters in a large dataset. Whilst unemployment and neighborhood insecurity have a significant impact on HIV risk and socioeconomically disadvantaged neighborhoods remain hotspots for HIV/AIDS, histoplasmosis was found dependent on the environmental distribution of the pathogen's spores. Public health efforts should prioritize disadvantaged neighborhoods for HIV/STI awareness, and screening and management of hepatitis and histoplasmosis.

6.CHAPTER V

Discussion

"Make no mistake; they are connected, these disease outbreaks coming one after another. And, they are not simply happening to us; they represent the un-intended results of things we are doing. They reflect the convergence of two forms of crisis on our planet.

The first crisis is ecological, the second is medical."

- David Quammen

6.1. Discussion

My PhD thesis aimed to understand the spatial dependency and distribution of EIDs using a biogeographic approach. This chapter compares and discusses the major findings of the previous chapters on the use of existing data to identify the drivers of EIDs and to predict potential hotspots of disease emergence. Also included is a discussion on the use of anticipatory 'get ahead of the curve' approach in the management and prevention of disease outbreaks. The discussion is divided into three themes: spatial dependence, statistics in biogeography and drivers of disease emergence. The chapter concludes with the discussion of the limitations, the future scope of research in disease emergence, and a graphical summary.

Significance of spatial dependence

This thesis set out with a null hypothesis that the distribution of the EIDs were spatially independent to the environment. The major findings of the thesis reject the null hypothesis. The results unanimously demonstrate that the distribution of the EIDs are spatially dependent and the local and global environment has a direct impact on it. Although disease emergence is a complex and multifactorial process, for most EIDs, the pathogens are already present in the local environment and are given a selected advantage due to perturbations at the host-pathogen-environment interface, leading to their exposure to humans (Morse, 1991).

In epidemiology, spatial dependence is defined by global or local clustering of cases. Local or first order clustering statistics measure the tendency of events (i.e. cases or outbreaks) to occur around a particular point in space while global or second order clustering measure the tendency of events to cluster in space in general (Diggle, Besag and Gleaves, 1976). This thesis focuses on local clustering statistics by developing risk maps (Chapters 1, 3) or identifying disease 'hotspots' (Chapters 1, 2 & 3). Chapter 4, however is based on second order clustering.

Grids and clusters

Grid is a 2-dimentional series of contiguous cells used for spatial indexing. Grids are important for measuring spatial dependence of EIDs. A cluster is an aggregation of cases of a disease or another health-related condition. In the figure below (*Figure 6.1*), grids A, B and C represent different patterns of disease distribution (disease represented by the red, yellow and green circles). The third grid demonstrates clustering of disease green and red at A1 and D2-3 respectively. This, grids are essential for spatial identification of clusters and analyze the drivers influencing the location of the clusters.

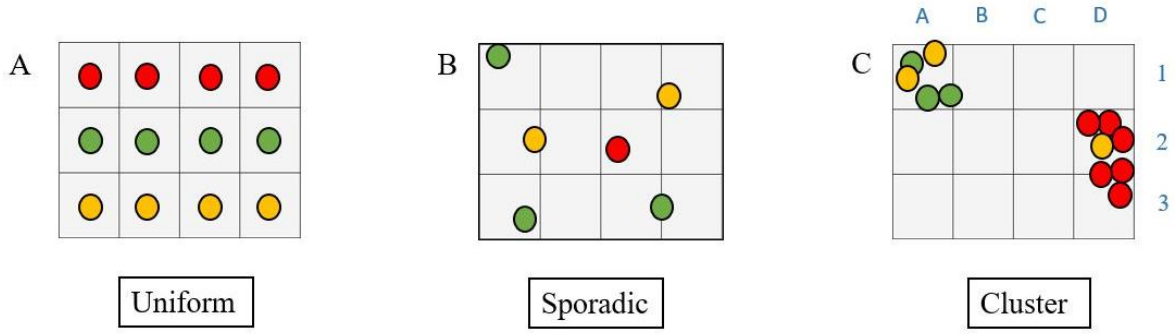

Figure 6.1 *Explaining spatial clustering.*

As over 70% of the EIDs are of zoonotic origin, the distribution of the pathogen is often dependent on the geographical distribution of the reservoir or intermediate hosts. The chapter 1 and 3 discussing viral zoonoses and leishmaniosis illustrate this spatial relationship. Studies show that zoonotic risk spatially depends on the distribution of mammal groups of high zoonotic potential (Han, Kramer and Drake, 2016a). Thus, regions with overlapping or high species richness of chiroptera (bats), primates, Rodentia (rodents), Carnivora (carnivores) and the hoofed ungulates were at a higher risk for EIDs. Among mammals, rodents contribute to the greater number of zoonotic hosts than any other order: approximately 10.7% of rodents are zoonotic hosts (Han *et al.*, 2015). Rodents, Rattus rattus and Rattus norvegicus, the principal reservoirs in the transmission of Leptospirosis and their role in the distribution of Leptospirosis in FG is discussed in chapter 2. The chapter demonstrates the influence of flooding and proximity to

croplands linked to rodent distribution on Leptospirosis incidence. Bats are also associated with numerous EID events in humans including SARS, MERS, EVD and 9.8% of bat species are known to be zoonotic hosts (Luis *et al.*, 2013). Around 49% of all species of order Carnivora were found to carry zoonotic pathogens (Han, Kramer and Drake, 2016b). Carnivores tend to accumulate the infectious agents of their prey. This plays a large role in transmission of pathogens through wet markets, where exotic game is butchered and sold. The SARS epidemic spread to humans through the consumption of carnivores, civets and raccoon dogs, both of which are intensively farmed in China. Leishmaniosis, discussed in chapter 3, can also be transmitted to humans by domestic carnivores, such as dogs and cats (Dantas-Torres, 2007; Kent *et al.*, 2013).

In the case of generalist pathogens, such as Mycobacterium *ulcerens,* discussed in detail in chapter 2, the distribution was found directly dependent to the surrounding aquatic environment. Low elevation, river basins and agricultural activities were found to be also significant risk factors in the spatial distribution of BU. We also observed interesting spatial clusters of BU in regions where gold mining on greenstone belts are prevalent (Jagadesh, Combe, Couppié, Nacher, *et al.*, 2019). This has led to changes in global distribution with new epidemics in Australia and Nigeria, where gold mining is also on the rise (O'Brien *et al.*, 2019).

In diseases transmitted by direct transmission such as HIV/AIDS, discussed in chapter 4, the scale of human movement along with behavioral and socio-economic factors play a larger role in spatial dependence than environmental factors. In such cases, spatial dependence by underlying risk factors or behaviors can lead to disease clustering. Unemployment and neighborhood insecurity have a significant impact on HIV risk and socioeconomically disadvantaged neighborhoods were found hotspots for HIV/AIDS. Another example for behavior-based clustering is the geographical clustering of pertussis in the USA due to vaccine refusal in communities (Omer *et al.*, 2006).

The surrounding environment thus plays a major role in the distribution of EIDs and intrinsically links the pathogen to its geographical location. The thesis confirms spatial dependency through various mathematical models in the preceding chapters.

Statistics in Biogeography

A biogeographic approach, defined by Evelyn C. Pielou, is observing, recording and explaining the geographic ranges of all living things (Pielou, 1979). Biogeographers over the course of time have introduced a variety of mathematical models to explain the observed distributions and events (Heikkinen and Ogmander, no date). With most models being stochastic by nature, using statistical methods is inevitable. The statistical software R version 3.0 to 4.0 was used for all statistical analysis (*Figure 6.2*). I chose R over STATA for epidemiological model. Spatial statistics in R is extensively developed, offering a wide choice for model construction and development of minor tweaks to the formulae to adapt models used in other life sciences for epidemiological purposes.

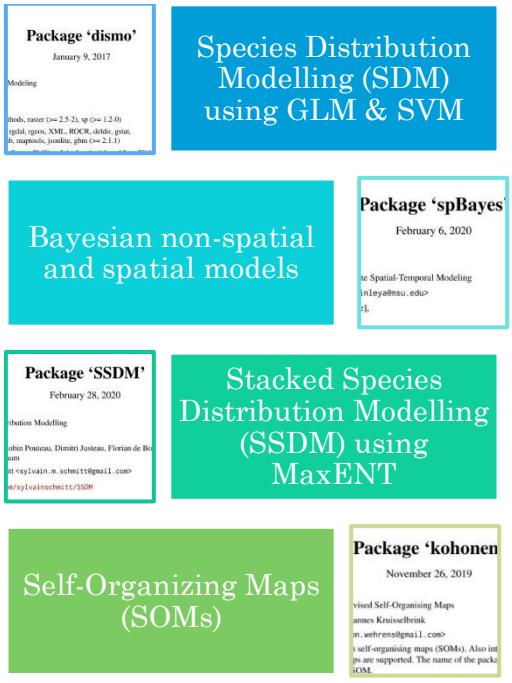

Figure 6.2 *Statistical models and main R packages used in the thesis*

Untapped data mines

With the rising issue of climate change, the geography of pathogen distribution is dependent on multiple environmental processes and interactions. Most models used in mapping the biogeography of EIDs include statistical, process-based and landscape-based models using a range of environmental variables ('Environmental change, climate and health: issues and research methods', 2003). For the construction of mathematical models throughout the thesis, I extracted the environmental variables from remote-sensing freely accessible data. The advancement of high-definition technology has made high-resolution satellite imagery freely available and a durable source of environmental data. Bioclimatic variables including topology, land cover, deforestation and climate data. Data on global distribution of terrestrial mammals and global biodiversity is also freely accessible from IUCN and GBIF databases. Also, existing datasets such as the Global Health Atlas, WHO archives and from other disease surveillance sites like Promed mail provide information on the global EID outbreaks. Chapter 1 of this thesis uses occurrence data from the above databases. Such technology is particularly useful in regions where environmental data is scarce. Use and expansion of this data offers various advantages:

1. Improvement of prediction and diagnostic capacity in disease endemic countries.
2. Regular and standardized inclusion of bioclimatic variables in epidemiological studies.
3. Advocates the inclusion of EIDs in the political climate agenda.

Mathematical models used

SDMs are a family of statistical learning methods that predicts distribution of species from a set of georeferenced observations and environmental predictors, often-used in macroecological studies. It relies on ecological theory of processes that mediate species distributions and abundance (Austin, 2002). In the first chapter of the thesis, SDM was used for the first time to map the predictive risk of the disease emergence and to establish the impact of environmental factors on their emergence. The species occurrences were replaced by outbreak events and pseudo-absence data within the geographical range of the reservoir. Furthermore, SDM proved to be a valuable tool in prediction of disease emergence by forecasting the regions in China (including Wuhan) to be hotspots for future diseases in the absence of COVID-19 data.

While biogeographical models are often based on stochastic processes, spatial Bayesian models are found to perform better when faced with limited, clumped or short-term occurrence data (Redding *et al.*, 2017). Chapter 2 confirms the robustness of spatial Bayesian models. In this case, the leptospirosis occurrences were of relatively shorter time period when compared to the BU cases and so, the choice to use bayesian models were made. The study also compared the use of non-spatial and spatial models in epidemiology and demonstrated that spatial Bayesian models had the narrow confidence intervals of the predictor variables and lower DIC values. The significance of spatial structure on the distribution of the two diseases was highlighted. In addition, the environmental covariates were found to significantly influence the spatial distribution of both diseases.

SSDMs is a useful statistical method to map the species distribution of multiple species. Chapter 3 explores the biogeography of five different species of Leishmania in FG. SSDMs aggregates outputs of the different species into maps of local species richness. MaxENT was the algorithm used in the SSDMs as the study used presence-only data. Along with the environmental predictor contributions, the endemism map illustrates the geographical range of the endemic species.

Although used in ecological and economic studies, SOMs are underutilized in epidemiology. SOMs proved to be effective in global cluster identification in HIV/AIDS, a disease transmitted by human-human direct transmission. Identifying clusters in a vulnerable population of an illness influenced largely by socioeconomic factors is a complex process. SOMs were able to identify clusters in need of targeted public health efforts. Histoplasmosis, however, was found dependent on the environmental distribution of the pathogen's spores.

Despite their unconventionality in epidemiology, all statistical models proved satisfactory. Expanding, using and validating existing databases ensures standardized, durable and reproducible research.

Drivers of EIDs

Across the preceding chapters, various environmental and socioeconomic drivers have been found to influence the emergence and the distribution of EIDs. The bioclimatic variables used in the thesis include elevation, TWI, land cover, climate, deforestation and domestic and terrestrial mammal distribution. Here, we compare and discuss the major drivers across the various taxonomic groups and modes of transmission (*Figure 6.3*).

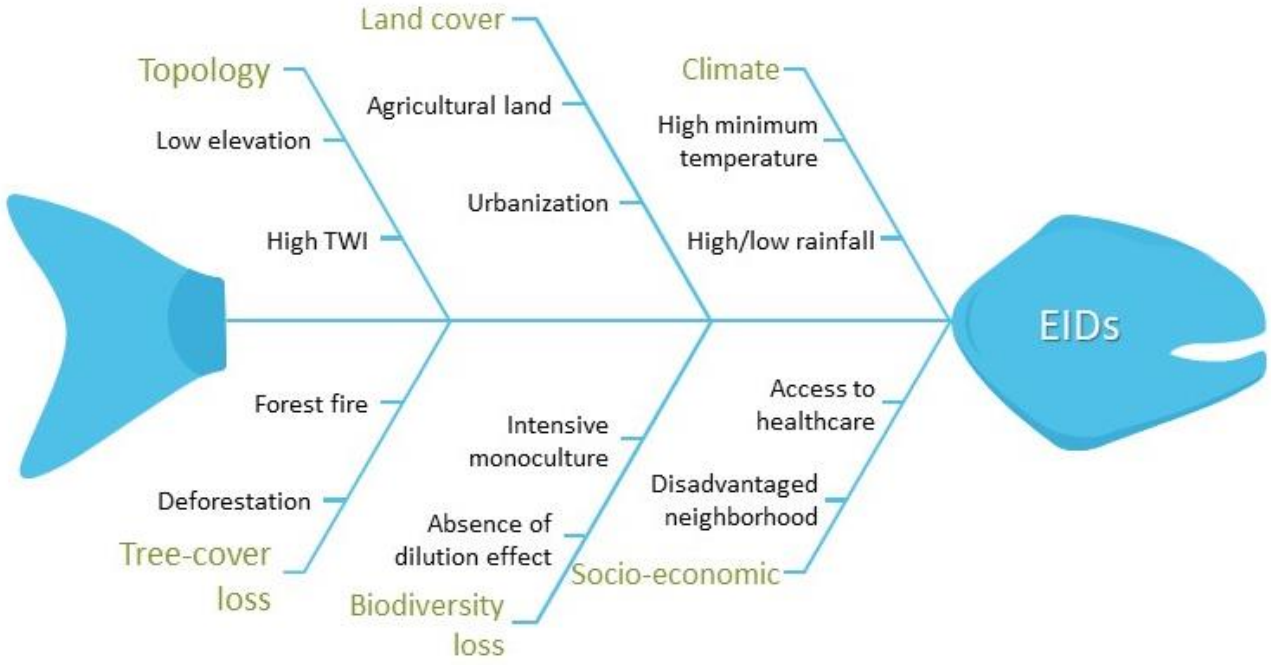

Figure 6.3 *The fishbone diagram illustrates the causal relationship between EIDs and its drivers, inferred from the thesis. The categories of the causes are highlighted in green and the causes written in black.*

River basins - Cradle of life and disease emergence

Low-lying floodplains are found to be at high-risk for EIDs as demonstrated with viral diseases (Nipah, SARS and MERS; Chapter 1), bacterial diseases of aquatic origin (BU and Leptospirosis; Chapter 2) and protozoal zoonotic disease (Leishmaniosis; Chapter 3). EIDs are often observed in river basins and plains as the fertile lands attracts human settlements for agriculture. The high population density alone does not account for the propensity of infections in these areas. Flooding associated with low-lying has been established to cause emergence of disease outbreaks. Flooding is also associated with the effects of El Niño/ La Niña and the increasing sea levels due to climate change. Chapter 1 hypothesized that flooding led to destruction of bat habitats, triggering disease emergence. While chapter 2 emphasizes that

flooding facilitates the spread of the pathogen through proliferation of rodents, which shed large amounts of pathogenic leptospires in their urine and thus increase the exposure to a susceptible population. BU, caused by an environmental generalist pathogen, is pendent on flooding onto agricultural or recreational area for increased contact with the susceptible population.

Studies demonstrate that flooding alters the balance of the host-pathogen-environment interface and often creates a conducive environment for the development of pathogens and vectors (Okaka and Odhiambo, 2018). Numerous studies demonstrates effects of flooding on vector-borne, rodent-borne and water-borne diseases as they require a vehicular transfer from host to host (in the case of water-borne) or a host/vector as part of its life cycle (Patz *et al.*, 2008; Brown and Murray, 2013). With water-borne diseases, intense rainfall mobilizes pathogens in the environment and displaces them into the aquatic environment and vice versa (Marcheggiani *et al.*, 2010; Cann *et al.*, 2013). The increasing trend of water-borne epidemics worldwide from 1980–2006 coincides with the rising number of flooding events (Adikari and Yoshitani, 2009). Rodent-borne diseases increase during flooding events because of altered patterns of human-pathogen-rodent contact (Ahern *et al.*, 2005) as discussed with leptospirosis in chapter 2. Vector-borne diseases have been extensively studied in relation to heavy rain and flooding (Hubálek and Halouzka, 1999; Medlock *et al.*, 2012; Weiler *et al.*, 2017).

The results of this thesis identifies river basins and low-lying coastal plains to be at high-risk for EIDs. Improved urban planning and avoiding contraction of habitations at river or sea front in expanding cities could prevent potential EIDs.

Agriculture - Sowing seeds for EIDs

Intensive agriculture, often a byproduct of overpopulation, causes rapid changes to the host-pathogen-environment nexus. Chapter 1 discusses the emergence of Nipah disease as a consequence of deforestation for extensive farming of oil palms (*Elaeis guineensis*) in the islands of Sumatra. Chapter 2 demonstrated that croplands and associated farming practices increased the rodent population leading to

increased exposure with surface water and soil contaminated by rodent excreta. The risk of BU also found to be high in proximity to croplands and rice farming.

Large-scale palm oil plantations have fueled the onset of a variety of neglected and emerging infectious diseases. In Colombia and Venezuela, the palm-oil plantations have provided optimal habitats for Chagas disease vectors (Guhl, Pinto and Aguilera, 2009). A systematic review on the palm oil industry demonstrates the increased risk of malaria, leptospirosis, Melioidosis and Onchocerciasis associated with frequent flooding due to the massive deforestation (Myzabella *et al.*, 2019). Furthermore, the Nipah virus outbreaks in Bangladesh and Kerala have been associated with drinking raw data palm sap contaminated with bat excreta (Islam *et al.*, 2016).

Rapid land changes caused by conversion of primary forests to arable land for large-scale monoculture is a recipe for disease emergence. This thesis confirms the less desirable effects of intensive agriculture on its role in increasing the risk of EIDs.

Climate change - Tempest amidst us

The increasing effects of climate change have led to rising sea levels, erratic rainfall and increasing minimum temperature (temperature at night; T_{min}). One of the major findings of this thesis is the direct impact of climate change on disease emergence and distribution of EIDs. In chapter 1, increasing average T_{min} was associated with the emergence of Nipah, RVF and coronavirus diseases, while it was inversely associated with filovirus diseases. Similarly, chapter 2 establishes a direct relationship between increased T_{min} and aquatic bacterial disease, leptospirosis and BU.

Studies observe that rising sea levels can act synergistically with climate change and then interact in a complex manner with other environmental drivers to generate an increased transmission of vector-borne infectious diseases (Ramasamy and Surendran, 2011). The evidence linking temperatures to EIDs is scarce, particularly for non vector-borne diseases. Warmer weather was found to permit vector survival at

higher elevations, spreading malaria beyond its historical geographic range in highland of eastern Kenya (Omumbo *et al.*, 2011). In addition, higher temperatures could increase the potential latitudinal extent of disease emergence. This hypothesis has been established from the vector-borne disease emergence predictions like CCHF and Zika (Gale et al., 2012; Tesla et al., 2018). The El Niño oscillations causing abnormally heavy rainfall have been associated with increased RVF in Africa due to increase habitat suitability for vector populations (Pittiglio *et al.*, no date).

Our results demonstrates that increase in the residual temperature, T_{min}, has a direct influence on the emergence and distribution of the EIDs. Rising sea levels leading to flooding along with unpredictable rainfall increase the risk of EIDs in tropical and temperate regions.

Biodiversity- the final frontier

Deforestation and biodiversity loss are tightly wound in a cause and effect relationship. Deforestation, as a consequence of urbanization and need of arable land, leads to biodiversity loss. Deforestation has been linked to numerous outbreaks such as EVD, Nipah and SARS as discussed in Chapter 1. Intensive monoculture of mammals for domestic, recreational and exotic game purposes has also been linked to disease emergence. For instance, camel breeding for races with MERS, extensive rearing of goat and cattle with RVF and captive breeding of wild animals for bushmeat with SARS.

Biodiversity loss, due to deforestation and increase in protein demand along with climate change has led to migration of the small mammals towards human habitations and decrease in predators. Furthermore, the global trade in exotic game parallels that of domesticated animals. Nearly 500 million kilograms of bushmeat are consumed annually in the tropics alone, more than six times the sustainable rate (Bradshaw, Sodhi and Brook, 2009). Decreasing species diversity and inter-species interactions makes the jump of pathogens from mammals easier towards the accidental human hosts. This has been termed as the "dilution effect" (Schmidt and Ostfeld, 2001; Khalil *et al.*, 2016)

This thesis demonstrates the impact of deforestation and the resulting biodiversity loss on EIDs. Monoculture of mammals and breeding wild animals especially predators for bushmeat directly influence and promote disease spillover.

Inequality in healthcare

Chapter 4 discusses in detail the effects of disadvantaged neighborhoods in the clustering diseases transmitted by direct transmission such as HIV/AIDS. Unemployment and neighborhood insecurity were also found to be key factors in these clusters. Vulnerable urban subpopulations, particularly immigrants, struggle to access health services (Rodwin and Gusmano, 2002). This inequality is even higher in asylum seekers and refugees. Outbreak control in these settings requires tackling poverty, malnutrition and over-crowding.

The ongoing COVID-19 outbreak portrays this disparity in health justice. The pandemic was observed to disproportionately affect black, Asian, and minority ethnic (BAME) communities (Kirby, 2020) in the USA. BAME communities often live in neglected neighborhoods (Quillian and Pager, 2001). They have poorer access to healthcare due underinsurance, which is defined as cost sharing that represents significant financial barriers to care or risk of catastrophic medical expenditures (Collins *et al.*, 2014). Death rates, from New York City, among black or African American people and Hispanic people were significantly higher than that of white or Asian people. The IFS report showed that in the UK, the death rate for people of black African descent was 3.5 times higher than for white British people, while for those of black Caribbean and Pakistani descent, death rates were 1.7 times and 2.7 times higher, respectively (*Are some ethnic groups more vulnerable to COVID-19 than others? | Inequality: the IFS Deaton Review*, no date). This difference in death rates can be explained by the increased comorbidities in the minority populations resulting from years of poor access to healthcare.

Socioeconomic factors play a pivotal role in the prevention and controls of EIDs. Sustainable and accessible healthcare is a human right regardless of nationality and borders. The pathogens respect neither imaginary boundaries nor ethnicities; why should healthcare be any different.

The bigger picture

Human driven environmental changes in the tropics pose ongoing, escalating threat on future disease emergence. There is an urgent need to integrate Nature at the center of disease epidemiology. Our current response to emerging infectious diseases is based on therapeutic control and prevention by immunization. With pathogens emerging, evolving and remerging at alarming rates, despite our advance in science, generating vaccines or novel drugs for each pathogen is beyond our current capacity. In the 21th century, the political agenda needs to move on from 'miracle cures' and focus as well as invest on prevention prior to actuation of outbreaks.

Predict to prevent

Epidemiology needs a change of paradigm towards a more predictive stance, switching from prevent and control to forecast and prevent. Monitoring disease emergence in wildlife and humans using a One Health approach is a part of several National and WHO led surveillance programs. PREDICT-2, the last-standing USAID Emerging Pandemic Threats funding program, which supported virology, ecology and epidemiology around the world since 2009 ended just a couple of weeks prior to the COVID19 outbreak (Carlson, 2020). The results of the program, however, led to one conclusive discovery of a zoonosis, the Bas-Congo virus (Grard *et al.*, 2012). Although PREDICT discovered hundreds of potential zoonoses (Kelly *et al.*, 2020), without human infection the potential spillovers of the pathogens remains unclear. Over the years, active surveillance programs have been found to be less attractive to stakeholders as their results are projected over a long-term period, which leads to pressure around accountability, lack of immediate action and uncertainty in general impact.

The results of this thesis recommends using a biographic approach to assess the predictive risk of EIDs using environmental and socioeconomic predictors. Mapping regions at high-risk and identifying hotspots using mathematical models have proved to be an effective way to prioritize mitigation and prevention of future pandemics in key regions. This approach bypasses the need of extensive environmental sampling.

Mathematical models used in disease risk predictions are based on assumptions and as with all statistical models, they are not without errors. Nevertheless, they provide enough evidence to be considered a 'reasonable basis for action'(Woolhouse, 2011). Weather forecasting depends on similar methods of predictive modelling. If weather forecasts provide enough basis for preventive action, why not consider predictive modelling in making decisions in the prevention of EIDs. Unfortunately, the COVID-19 pandemic is months or years past the stage where predictive risk modelling could have made the most difference.

Environmental sustainability

At an estimated annual worldwide loss of 0.8% of forests, studies predict that between 0.1 and 0.3% of tropical forest species (14 000 to 40 000 species) may be disappearing each year (Randall Hughes *et al.*, 2007). Destruction of natural habitats is one of the major causes for decline in global biodiversity (Rands *et al.*, 2010). Such destruction results in land fragmentation (i.e. the division of habitat into smaller and more isolated fragments separated by human-transformed land cover (Haddad *et al.*, 2015)) with disrupted host-pathogen-environment interface at the front of and human-transformed area increasing the risk of spillover.

Today, the "bushmeat" crisis is also one of the most severe but preventable threats to host-pathogen- environment interfaces. Bushmeat hunting is facilitated by deforestation and habitat fragmentation, which provides access to the deeper parts of the forest. For instance in Cameroon, the large and growing urban demand for bushmeat coincided with the opening up of logging concessions in the East Province

(Wolfe *et al.*, 2005). The hunters, particularly the butchers handling the animals are at high risk of infection due to exposure of the contaminated body fluids (Kurpiers *et al.*, 2015).

However, in many countries in sub-Saharan Africa, bushmeat is the only source of protein available to communities. In remote parts of Cameroon, bushmeat comprised 80–98 % of animal protein consumption in communities with very few opportunities to buy alternative sources of protein (Muchaal and Ngandjui, 1999). Similarly, in rural Equatorial Guinea, bushmeat consumption contributed to 43 % of total protein consumption in remote villages, but only 18 % in well-connected villages (Allebone-Webb, no date). On the contrary, in southern China, bushmeat also known as "yewei" or wild flavor and is consumed for its exotic flavor as it is believed that consuming "yewei" taps into the strength of the animals and is a sign of wealth (Karl Taro Greenfeld, 2006).

Solutions lie in banning wildlife trade and promoting sustainable and durable alternative protein sources in remote communities. In addition, stringent disincentives for logging and deforestation leading to fragmentation deters hunters by cutting off their access into the forests. Detecting deforestation through remote-sensing data could help in identifying hotspots for potential spillovers. The massive scale and intensity of human activity and its environmental effects, forces us to look at the bigger picture. Eventually, we need to adapt to a rapidly changing environment but we also need to rapidly mitigate the damage. Invest in sustainable food production and livestock farming; manage climate change, better urban planning, and construct human settlements without encroaching on the natural environment to sustain the growing population.

'Get ahead of the curve'

By acknowledging the influence of the environmental predictors on EIDs and moving towards a predictive approach to disease emergence, we could nip outbreaks in the bud.

Assumptions, limitations and delimitations

As with all mathematical models, the spatial models used in the thesis are based on following assumptions *(Table 6.1)*.

Table 6.1 *Assumptions to be acknowledged in this thesis*

Assumptions	**Mitigation**
Pathogens are in equilibrium with the geographical range of the model	Chapter 1: The geographical range was set by the geographical distribution of the mammal reservoirs and host. Chapter 2: Previous evidence from environmental sampling (Marine Combe *et al.*, 2019) Chapter 3: FG is endemic to Leishmaniosis Chapter 4: Not applicable
Environmental stability during the timeframe of the analysis	Chapter 1: The timeframe of the study was shortened between 2000-2018 Chapter 2: Time-period of BU case longer than Leptospirosis due to poor case density but less than a 50-year period. Chapter 3: Short time period Chapter 4: Not applicable
Training samples are representative of the occurrences	Chapter 1: Pseudo-absence pointes generated Chapter 2: Not applicable Chapter 3: MaxENT used Chapter 4: Not applicable

The major limitation of this thesis is the absence of temporal component in the form of times series, which would be interesting especially with the climatic covariates. We mitigated this by choosing recent raster data corresponding to the period of the study and linking the spatiotemporal points to corresponding climatic monthly covariates. In chapter 4, the environmental drivers influencing Histoplasmosis distribution is poorly analyzed. As the Histoplasmosis studied was dependent on HIV/AIDS there were also many confounding factors.

The delimitation of the study is the its relatively small spatial scale. The rasters used in the thesis were of high resolution and thus the analysis was time-consuming. With the time constraints of the thesis, limiting the spatial scale (especially in chapter 1) was a good compromise for better quality and robust models. As vector-borne diseases has been extensively studied in the context of climate change, I chose diseases with different transmission dynamics.

Future scope of research

Throughout the thesis, novel statistical methods using a biogeographic approach were in infectious disease epidemiology. We chose models that can be used with limited case data, in the absence of environmental pathogen data and easily accessible bioclimatic variables. Expanding and validating these models in infectious diseases with different modes of transmission is required. For instance, modelling the poleward shift of vector-borne diseases using SSDMs. Socioeconomic factors of disease emergence remains a relatively underexplored area.

Spatial Bayesian analysis to compare different diseases of the same origin could help in underpinning common factors. Policy-ready recommendations from these studies would be of interest to policy-makers.

Explore the competition in the niches of Leishmania species. Chapter 3 observed the difference in the geographical distribution between species of Leishmania and discussed in detail the influence of environment in the spatial patterns. The population demographics and the potential reservoirs of the different

species need further study to explain the spatial trends. Addition of the temporal component should be considered.

SOMs have tremendous potential in prioritizing clusters in the epidemiology, grouping cases with similar characteristics, especially in large datasets. Testing SOMs in identifying clusters based on the bio-climatic variables of EIDs in the region from existing databases can be less time consuming than large raster based spatial models.

A systematic methodological review on prediction modelling is needed to evaluate the effectiveness of prediction modelling on disease emergence.

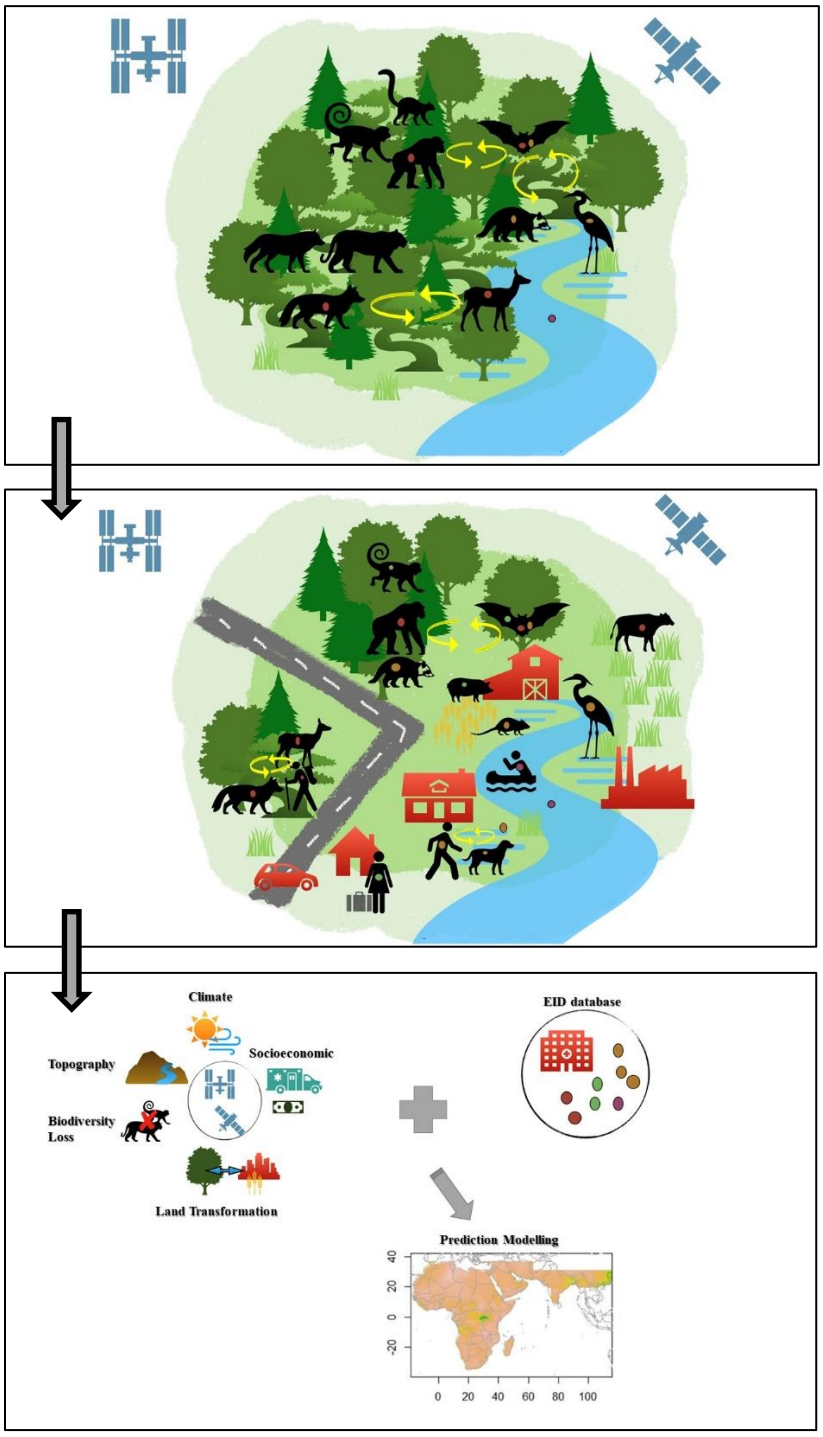

Figure 6.4 *Summary of the thesis*

6.2. Recommendations

The major findings of this thesis infer the following policy-relevant recommendations on emergence and distribution of EIDs:

1. Expand, use and validate existing databases and remote-sensing data to ensure durable, reproducible and standardization in EID research.
2. Better urban planning to avoid construction of residential settlements at flood-prone, low-lying regions.
3. Environmental sustainability must be included in the global EID prevention and control agenda.
4. Deforestation and habitat fragmentation to be mitigated by land corianders to promote the 'dilution effect' and avoidance of livestock farming in fragmented land.
5. Ban bushmeat trading and the logging practices that promote it.
6. Disease emergence to be a part of the international policy of climate change (IPCC).
7. Sustainable and accessible healthcare equally available across population groups with special public health focus on vulnerable communities.

7. CONCLUSIONS

7.1. Conclusions (Français)

Le choix d'une approche biogéographique permettant de cartographiée la distribution des EID en combinant les données sur les cas clinique, l'imagerie satellitaire et des modèles statistiques non conventionnels s'est montré particulièrement efficace dans l'analyse prédictive des risques et la détermination des zones à haut risque émergence de maladies infectieuses. Cette thèse a permis d'établir que l'émergence des maladies infectieuses peut être anticipée en identifiant et en contrôlant les différents facteurs ayant un lien direct avec l'anthropisation de l'environnement. Ainsi, les zoonoses et les maladies véhiculées par l'eau se sont avérées spatialement dépendantes de facteurs environnementaux tels que les basses altitudes, la déforestation, la transformation des terres et l'augmentation du minimum de température alors que ce sont les facteurs socio-économiques qui régissent essentiellement la transmission des maladies d'humain à humain.

.

Nouveautés

- Il existe une dépendance spatiale dans la distribution des EMI.
- La distribution des émergences des BPDs est influencée de manière directe par la déforestation et la fragmentation des habitats.
- Les plaines inondables sont des zones à haut risque pour les maladies bactériennes d'origine aquatique.
- L'exploitation aurifère est un facteur de risque sous-estimé de l'émergence de BU et de son endémicité.
- Les publications issues de cette thèse sont les premières à établir le rôle significatif de l'augmentation du minimum de température, conséquence majeure du changement climatique, dans l'émergence des maladies infectieuses et de leur distribution.

- Les différentes espèces du genre Leishmania en Guyane Française ont une répartition spatiale territoriale.
- Des modèles statistiques de type SDM dans les régions à haut risque d'EMI sont efficaces et ont permis notamment d'identifier la région de Wuhan (Chine) comme zone à haut risque d'émergence, sans même avoir de données cliniques/épidémiologiques sur le COVID-19.

Résumé

L'environnement immédiat joue un rôle central dans la distribution des EMI et permet d'associer le pathogène à une localisation géographique. Les modèles prédictifs utilisant une approche biogéographique permettent de combiner les données cliniques existantes ainsi que les indicateurs environnementaux accessibles afin de cartographier la distribution des EMI et identifier précocement les zones à risque. Cette approche est particulièrement idoine lorsque les données sont rares, voir absentes dans le cas d'un nouveau pathogène notamment. Par son caractère anticipatif, cette approche de la prévention et le contrôle des EMI pourrait améliorer considérablement l'efficacité de gestion des épidémies à venir.

7.2. Conclusions (English)

Biogeographic approach to mapping the distribution of EIDs with using existing human cases data, remote sensing imagery and unconventional statistical models was found effective in analyzing the predictive risk, and detection of hotspots. The thesis established that EIDs are not unprecedented but predictable by identifying and managing the triggers of disease emergence, which have found to have a direct link with the anthropization of the environment. Zoonoses, and water-borne diseases were found to be spatially dependent on environmental factors such as low elevation, deforestation, transformed land and minimum temperature while direct human to human transmission diseases where influenced by socio-economic factors.

Gaps filled

- **Spatial dependence** in the distribution of EIDs was confirmed in this thesis.
- The distribution of the BPD outbreaks was influenced directly by **deforestation** and habitat fragmentation.
- **Low-lying floodplains** are at high risk of bacterial diseases of aquatic origin
- **Gold mining** was observed to be an understudied risk factor of BU emergence and endemicity.
- The studies included in the thesis were the first to establish the role of **rising minimum temperature**, a major consequence of climate change, in disease emergence and distribution.
- The **spatial patterns** of the different Leishmania species were observed to be territorial in FG.
- SDMs were used to predict the regions at high risk of **future disease emergence** and were effective in identifying the region of Wuhan (Hubei, China) as a hotspot for EIDs in the absence of COVID-19 data.

Summary

The surrounding environment plays a major role in the distribution of EIDs and intrinsically links the pathogen to its geographical location. Prediction modelling using a biogeographic approach utilizes existing clinical data, and accessible environmental predictors, in situations where data on the pathogen is scarce or in cases of a novel pathogen, to map the distribution of EIDs and identify disease hotspots. An anticipatory approach for EID prevention and control can be the missing link needed to prevent future pandemics.

References

Abatzoglou, J. T. *et al.* (2018) 'TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958-2015', *Scientific Data*. Nature Publishing Groups, 5(1), pp. 1–12. doi: 10.1038/sdata.2017.191.

Adikari, Y. and Yoshitani, J. (2009) 'Global Trends in Water-Related Disasters: an insight for policymakers Water in a Changing World Global Trends in Water- Related Disasters an insight for policymakers', *International Centre for Water Hazard and Risk Management (ICHARM)*. Available at: www.grundini.com (Accessed: 31 May 2020).

Ahern, M. *et al.* (2005) 'Global Health Impacts of Floods: Epidemiologic Evidence', *Epidemiologic Reviews*, 27(1), pp. 36–46. doi: 10.1093/epirev/mxi004.

Allebone-Webb, S. (no date) *Evaluating dependence on wildlife products in rural Euatorial Guinea*.

Ambat, A. S. *et al.* (2019) 'Nipah virus: A review on epidemiological characteristics and outbreaks to inform public health decision making', *Journal of Infection and Public Health*. Elsevier Ltd, 12(5), pp. 634–639. doi: 10.1016/j.jiph.2019.02.013.

Amofah, G. *et al.* (2002) 'Buruli ulcer in Ghana: results of a national case search.', *Emerging infectious diseases*, pp. 167–170. Available at: https://www.ncbi.nlm.nih.gov/pubmed/11897068.

'An update on the risk of transmission of Ebola virus (EBOV) via the food chain' (2014) *EFSA Journal*, 12(11). doi: 10.2903/j.efsa.2014.3884.

'Answer given by Mr Potočnik on behalf of the Commission' (2010). European Parliament. Available at: http://www.europarl.europa.eu/sides/getAllAnswers.do?reference=P-2010-3589&language=EN.

Anthony J McMichael and Ulisses Confalonieri (no date) 'Environment, climate change, social factors and the implications for controlling infectious diseases of poverty', in. Available at:

https://www.who.int/tdr/capacity/global_report/2012/chapitre2_web.pdf (Accessed: 4 May 2020).

Are some ethnic groups more vulnerable to COVID-19 than others? | Inequality: the IFS Deaton Review (no date). Available at: https://www.ifs.org.uk/inequality/chapter/are-some-ethnic-groups-more-vulnerable-to-COVID-19-than-others/ (Accessed: 31 May 2020).

Attorre, F. *et al.* (2007) 'Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale', *International Journal of Climatology*. John Wiley & Sons, Ltd, 27(13), pp. 1825–1843. doi: 10.1002/joc.1495.

Austin, K. F., Choi, M. M. and Berndt, V. (2017) 'Trading sex for security: Unemployment and the unequal HIV burden among young women in developing nations', *International Sociology*, 32(3), pp. 343–368. doi: 10.1177/0268580917693172.

Austin, M. P. (2002) 'Spatial prediction of species distribution: An interface between ecological theory and statistical modelling', *Ecological Modelling*. Elsevier, 157(2–3), pp. 101–118. doi: 10.1016/S0304-3800(02)00205-3.

Baker, H. J. (1965) 'Leptospirosis in Malaysia', *Military Medicine*. Oxford University Press, 130(11), pp. 1101–1102. doi: 10.1093/milmed/130.11.1101.

Barcellos, C. and Sabroza, P. C. (2001) 'The place behind the case: leptospirosis risks and associated environmental conditions in a flood-related outbreak in Rio de Janeiro.', *Cadernos de saúde pública / Ministério da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública*, 17 Suppl, pp. 59–67. doi: 10.1590/s0102-311x2001000700014.

Beven, K. J. and Kirkby, M. J. (1979) 'A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant', *Hydrological Sciences Bulletin*. Taylor & Francis Group, 24(1), pp. 43–69. doi: 10.1080/02626667909491834.

Blockstein, D. E. (1998) 'Lyme Disease and the Passenger Pigeon?', *Science*. American Association for

the Advancement of Science (AAAS), 279(5358), pp. 1831c – 1831. doi: 10.1126/science.279.5358.1831c.

Booth, T. H. *et al.* (no date) 'BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies'. doi: 10.1111/ddi.12144.

Bradshaw, C. J. A., Sodhi, N. S. and Brook, B. W. (2009) 'Tropical turmoil: A biodiversity tragedy in progress', *Frontiers in Ecology and the Environment*. John Wiley & Sons, Ltd, pp. 79–87. doi: 10.1890/070193.

Brown, L. and Murray, V. (2013) 'Examining the relationship between infectious diseases and flooding in Europe', *Disaster Health*. Informa UK Limited, 1(2), pp. 117–127. doi: 10.4161/dish.25216.

Broxton, P. D. *et al.* (2014) 'A Global Land Cover Climatology Using MODIS Data', *Journal of Applied Meteorology and Climatology*, 53(6), pp. 1593–1605. doi: 10.1175/JAMC-D-13-0270.1.

Cann, K. F. *et al.* (2013) 'Extreme water-related weather events and waterborne disease', *Epidemiology and Infection*. Cambridge University Press, pp. 671–686. doi: 10.1017/S0950268812001653.

Carlson, C. J. (2020) 'From PREDICT to prevention, one pandemic later', *The Lancet Microbe*. Elsevier BV, 1(1), pp. e6–e7. doi: 10.1016/s2666-5247(20)30002-1.

Casanovas-Massana, A. *et al.* (2018) 'Quantification of Leptospira interrogans Survival in Soil and Water Microcosms', *Applied and Environmental Microbiology*. Edited by S.-J. Liu, 84(13). Available at: http://aem.asm.org/content/84/13/e00507-18.abstract.

Collins, S. R. *et al.* (2014) 'Too high a price: out-of-pocket health care costs in the United States. Findings from the Commonwealth Fund Health Care Affordability Tracking Survey. September-October 2014', *Issue brief (Commonwealth Fund)*, 29, pp. 1–11.

Combe, M. *et al.* (2017a) 'Global and local environmental changes as drivers of Buruli ulcer emergence.',

Emerging microbes & infections, p. e22. Available at: https://www.ncbi.nlm.nih.gov/pubmed/28442755.

Combe, M. *et al.* (2017b) 'Global and local environmental changes as drivers of Buruli ulcer emergence', *Emerging Microbes & Infections*. doi: 10.1038/emi.2017.7.

Combe, M. *et al.* (2019) 'Comparison of Mycobacterium ulcerans (Buruli ulcer) and Leptospira sp. (Leptospirosis) dynamics in urban and rural settings', *PLoS Neglected Tropical Diseases*, 13(1). doi: 10.1371/journal.pntd.0007074.

Combe, Marine *et al.* (2019) 'Comparison of Mycobacterium ulcerans (Buruli ulcer) and Leptospira sp. (Leptospirosis) dynamics in urban and rural settings', *PLOS Neglected Tropical Diseases*. Edited by B. Althouse. Public Library of Science, 13(1), p. e0007074. doi: 10.1371/journal.pntd.0007074.

Costello, A. *et al.* (2009) 'Managing the health effects of climate change. Lancet and University College London Institute for Global Health Commission', *The Lancet*. Elsevier Limited, pp. 1693–1733. doi: 10.1016/S0140-6736(09)60935-1.

Cropet, C. *et al.* (2018) 'Epidemiology of Human Leptospirosis in French Guiana (2007–2014): A Retrospective Study', *The American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 99(3), pp. 590–596. doi: 10.4269/ajtmh.17-0734.

Dantas-Torres, F. (2007) 'The role of dogs as reservoirs of Leishmania parasites, with emphasis on Leishmania (Leishmania) infantum and Leishmania (Viannia) braziliensis', *Veterinary Parasitology*, 149(3–4), pp. 139–146. doi: 10.1016/j.vetpar.2007.07.007.

Daszak, P. *et al.* (2013) 'Interdisciplinary approaches to understanding disease emergence: The past, present, and future drivers of Nipah virus emergence', *Proceedings of the National Academy of Sciences of the United States of America*, 110(SUPPL. 1), pp. 3681–3688. doi: 10.1073/pnas.1201243109.

Dedet, J. P., Pradinaud, R. and Gay, F. (1989) 'Epidemiological aspects of human cutaneous leishmaniasis in french guiana', *Transactions of the Royal Society of Tropical Medicine and Hygiene*. Trans R Soc Trop Med Hyg, 83(5), pp. 616–620. doi: 10.1016/0035-9203(89)90375-1.

Desjeux, P. (2001) ‘The increase in risk factors for leishmaniasis worldwide’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*. Royal Society of Tropical Medicine and Hygiene, pp. 239–243. doi: 10.1016/S0035-9203(01)90223-8.

Desjeux, P. and Dedet, J. P. (1989) ‘Isoenzyme characterization of 112 leishmania isolates from french guiana’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*. Trans R Soc Trop Med Hyg, 83(5), pp. 610–612. doi: 10.1016/0035-9203(89)90373-8.

Dhewantara, P. W. *et al.* (2018) ‘Geographical and temporal distribution of the residual clusters of human leptospirosis in China, 2005–2016’, *Scientific Reports*. Nature Publishing Group, 8(1), p. 16650. doi: 10.1038/s41598-018-35074-3.

Diggle, P. J., Besag, J. and Gleaves, J. T. (1976) ‘Statistical Analysis of Spatial Point Patterns by Means of Distance Methods’, *Biometrics*. JSTOR, 32(3), p. 659. doi: 10.2307/2529754.

Disease Background – Cutaneous Leishmaniasis – DNDi (no date). Available at: https://www.dndi.org/diseases-projects/leishmaniasis/disease-background-cl/ (Accessed: 5 June 2020).

Douine, M. *et al.* (2017a) ‘Mycobacterium ulcerans infection (Buruli ulcer) in French Guiana, South America, 1969–2013: an epidemiological study’, *The Lancet Planetary Health*. Elsevier, 1(2), pp. e65–e73. doi: 10.1016/S2542-5196(17)30009-8.

Douine, M. *et al.* (2017b) ‘Mycobacterium ulcerans infection (Buruli ulcer) in French Guiana, South America, 1969–2013: an epidemiological study’, *The Lancet Planetary Health*. Elsevier, 1(2), pp. e65–e73. doi: 10.1016/S2542-5196(17)30009-8.

Duker, A. A., Portaels, F. and Hale, M. (2006) ‘Pathways of Mycobacterium ulcerans infection: a review.’, *Environment international*, pp. 567–573. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16492390.

‘Environmental change, climate and health: issues and research methods’ (2003) *Bulletin of the World Health Organization*, 81(3), pp. 226–227. doi: 10.1590/S0042-96862003000300016.

Epelboin, L. *et al.* (2017) ‘La leptospirose en Guyane française et sur le bouclier des Guyanes. État des connaissances en 2016’, *Bulletin de la Société de pathologie exotique*. Lavoisier, 110(3), pp. 165–179. doi: 10.1007/s13149-017-0559-9.

Epelboin, Loïc *et al.* (2017) ‘La leptospirose humaine en Guyane : état des connaissances et perspectives’, *Bulletin d’épidémiologie hebdomadaire*, 8–9, pp. 168–175. Available at: https://jeannicod.ccsd.cnrs.fr/SANTE_PUB_INSERM/inserm-01503825v1 (Accessed: 26 March 2019).

Escobar, L. E. and Craft, M. E. (2016) ‘Advances and limitations of disease biogeography using ecological niche modeling’, *Frontiers in Microbiology*. doi: 10.3389/fmicb.2016.01174.

FA, J. E. *et al.* (1995) ‘Impact of Market Hunting on Mammal Species in Equatorial Guinea’, *Conservation Biology*, 9(5), pp. 1107–1115. doi: 10.1046/j.1523-1739.1995.951107.x.

Farmer, P. E. *et al.* (2006) ‘Structural Violence and Clinical Medicine’, *PLoS Medicine*. Public Library of Science, 3(10), p. e449. doi: 10.1371/journal.pmed.0030449.

Floch, H. (1954) ‘[Notes and Discussions on Forest Leishmaniasis in French Guiana] - PubMed’, *Publ Inst Pasteur Guyane Fr Inini* . Available at: https://pubmed.ncbi.nlm.nih.gov/13229406/ (Accessed: 5 June 2020).

Gale, P. *et al.* (2012) ‘Impact of climate change on risk of incursion of Crimean-Congo haemorrhagic fever virus in livestock in Europe through migratory birds’, *Journal of Applied Microbiology*, 112(2), pp. 246–257. doi: 10.1111/j.1365-2672.2011.05203.x.

Galgamuwa, L. S., Dharmaratne, S. D. and Iddawela, D. (2018) ‘Leishmaniasis in Sri Lanka: Spatial distribution and seasonal variations from 2009 to 2016’, *Parasites and Vectors*. BioMed Central Ltd., p. 60. doi: 10.1186/s13071-018-2647-5.

García, F. C. *et al.* (2018) ‘Changes in temperature alter the relationship between biodiversity and ecosystem functioning’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 115(43), pp. 10989–10994. doi: 10.1073/pnas.1805518115.

Gelman, A. and Rubin, D. B. (1992) ‘Inference from Iterative Simulation Using Multiple Sequences’, *Statist. Sci.* The Institute of Mathematical Statistics, 7(4), pp. 457–472. doi: 10.1214/ss/1177011136.

Gibbs, H. K. *et al.* (2010) ‘Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s’, *Proceedings of the National Academy of Sciences*, 107(38), pp. 16732 LP – 16737. Available at: http://www.pnas.org/content/107/38/16732.abstract.

Gomez, E. A. *et al.* (2014) ‘Distribution of Lutzomyia ayacuchensis, the vector of Andean-type cutaneous leishmaniasis, at different altitudes on the Andean slope of Ecuador’, *Acta Tropica*. Elsevier, 137, pp. 118–122. doi: 10.1016/j.actatropica.2014.05.006.

Grard, G. *et al.* (2012) ‘A Novel Rhabdovirus Associated with Acute Hemorrhagic Fever in Central Africa’, *PLoS Pathogens*. Edited by D. Wang. Public Library of Science, 8(9), p. e1002924. doi: 10.1371/journal.ppat.1002924.

Guégan, J.-F. (2008) *Ostfeld RS Keesing F, Eviner VT: Infectious Disease Ecology. Effects of Ecosystems on Disease and of Disease on Ecosystems, Parasites & Vectors - Parasites Vectors*. doi: 10.1186/1756-3305-1-28.

Guhl, F., Pinto, N. and Aguilera, G. (2009) ‘Sylvatic triatominae: A new challenge in vector control transmission’, *Memorias do Instituto Oswaldo Cruz*. Fundacao Oswaldo Cruz, 104(SUPPL. 1), pp. 71–75. doi: 10.1590/S0074-02762009000900012.

Haddad, N. M. *et al.* (2015) ‘Habitat fragmentation and its lasting impact on Earth’s ecosystems’, *Science Advances*. American Association for the Advancement of Science, 1(2), p. e1500052. doi: 10.1126/sciadv.1500052.

Hadfield, J. (2018) ‘Package “ MCMCglmm ”’.

Hagan, J. E. *et al.* (2016) ‘Spatiotemporal Determinants of Urban Leptospirosis Transmission: Four-Year Prospective Cohort Study of Slum Residents in Brazil’, *PLoS neglected tropical diseases*. Public Library of Science, 10(1), pp. e0004275–e0004275. doi: 10.1371/journal.pntd.0004275.

Hagarty, J. *et al.* (2015) ‘Chemistry of natural waters and its relation to Buruli ulcer in Ghana’, *Journal of Hydrology: Regional Studies*. doi: 10.1016/j.ejrh.2015.03.006.

Han, B. A. *et al.* (2015) ‘Rodent reservoirs of future zoonotic diseases’, *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1501598112.

Han, B. A., Kramer, A. M. and Drake, J. M. (2016a) ‘Global Patterns of Zoonotic Disease in Mammals’, *Trends in Parasitology*. Elsevier Ltd, pp. 565–577. doi: 10.1016/j.pt.2016.04.007.

Han, B. A., Kramer, A. M. and Drake, J. M. (2016b) ‘Global Patterns of Zoonotic Disease in Mammals’, *Trends in Parasitology*. doi: 10.1016/j.pt.2016.04.007.

Harvell, C. D. *et al.* (2002) ‘Climate warming and disease risks for terrestrial and marine biota.’, *Science (New York, N.Y.)*. United States, 296(5576), pp. 2158–2162. doi: 10.1126/science.1063699.

Heikkinen, J. and Ogmander, H. H. (no date) *STATISTICS IN BIOGEOGRAPHY*.

Herwaldt, B. L. (1999) ‘Leishmaniasis’, *Lancet*. Elsevier Limited, pp. 1191–1199. doi: 10.1016/S0140-6736(98)10178-2.

Hubálek, Z. and Halouzka, J. (1999) ‘West Nile fever - A reemerging mosquito-borne viral disease in Europe’, *Emerging Infectious Diseases*. Centers for Disease Control and Prevention, 5(5), pp. 643–650. doi: 10.3201/eid0505.990505.

Huber, F. *et al.* (2017) ‘High mortality among male HIV-infected patients after prison release: ART is not enough after incarceration with HIV’. doi: 10.1371/journal.pone.0175740.

Isberg, S., Balaguera-Reina, S. A. and Ross, J. P. (2017) ‘The IUCN Red List of Threatened Species’. doi: 10.2305/IUCN.UK.2017.

Islam, M. S. *et al.* (2016) ‘Nipah virus transmission from bats to humans associated with drinking traditional liquor made from date palm sap, Bangladesh, 2011-2014’, *Emerging Infectious Diseases*. Centers for Disease Control and Prevention (CDC), 22(4), pp. 664–670. doi: 10.3201/eid2204.151747.

Jagadesh, S., Combe, M., Couppié, P., Le Turnier, P., *et al.* (2019) 'Emerging human infectious diseases of aquatic origin: A comparative biogeographic approach using Bayesian spatial modelling', *International Journal of Health Geographics*. BioMed Central Ltd., 18(1), p. 23. doi: 10.1186/s12942-019-0188-6.

Jagadesh, S., Combe, M., Couppié, P., Nacher, M., *et al.* (2019) 'Global Emergence of Buruli Ulcer', *EcoHealth*. doi: 10.1007/s10393-019-01445-z.

Janssens, P. *et al.* (2005) 'Buruli ulcer: An historical overview with updating to 2005'. Available at: https://www.researchgate.net/publication/292417711.

Jones, K. E. *et al.* (2008) 'Global trends in emerging infectious diseases', *Nature*. Nature Publishing Group, 451, p. 990. Available at: http://dx.doi.org/10.1038/nature06536.

Joshi, Y. P., Kim, E.-H. and Cheong, H.-K. (2017) 'The influence of climatic factors on the development of hemorrhagic fever with renal syndrome and leptospirosis during the peak season in Korea: an ecologic study', *BMC Infectious Diseases*. BioMed Central, 17(1), p. 406. doi: 10.1186/s12879-017-2506-6.

Karl Taro Greenfeld (2006) *China Syndrome: The True Story of the 21st Century's First Great Epidemic*. Available at: https://books.google.fr/books/about/China_Syndrome.html?id=HRz3_mgUnqYC&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false (Accessed: 2 June 2020).

Kelly, T. R. *et al.* (2020) 'Implementing One Health approaches to confront emerging and re-emerging zoonotic disease threats: lessons from PREDICT', *One Health Outlook*. Springer Science and Business Media LLC, 2(1), pp. 1–7. doi: 10.1186/s42522-019-0007-9.

Kent, A. *et al.* (2013) 'Is the Dog a Possible Reservoir for Cutaneous Leishmaniasis in Suriname?', *Journal of Tropical Medicine*. Edited by A. P. Dash. Hindawi Publishing Corporation, 2013, p. 324140. doi: 10.1155/2013/324140.

Khalil, H. *et al.* (2016) 'Declining ecosystem health and the dilution effect', *Scientific Reports*. Nature Publishing Group, 6(1), pp. 1–11. doi: 10.1038/srep31314.

Kirby, T. (2020) ‘Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities’, *The Lancet Respiratory Medicine*. Elsevier BV, 0(0). doi: 10.1016/s2213-2600(20)30228-9.

Koch, L. K. *et al.* (2017) ‘Modeling the climatic suitability of leishmaniasis vector species in Europe’, *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–10. doi: 10.1038/s41598-017-13822-1.

Koh, H. K. *et al.* (2010) ‘Translating research evidence into practice to reduce health disparities’, *American journal of public health*, 100 Suppl, pp. S72-80. doi: 10.2105/AJPH.2009.167353.

Kuriakose, M. *et al.* (2008) ‘Leptospirosis in a midland rural area of Kerala State’, *Indian Journal of Medical Research*. doi: 10.1063/1.2165194.

Kurpiers, L. A. *et al.* (2015) ‘Bushmeat and emerging infectious diseases: Lessons from Africa’, in *Problematic Wildlife: A Cross-Disciplinary Approach*. Springer International Publishing, pp. 507–551. doi: 10.1007/978-3-319-22246-2_24.

Lafferty, K. D., Porter, J. W. and Ford, S. E. (2004) ‘Are Diseases Increasing in the Ocean?’, *Annual Review of Ecology, Evolution, and Systematics*. Annual Reviews, 35(1), pp. 31–54. doi: 10.1146/annurev.ecolsys.35.021103.105704.

Lainson, R. (no date) ‘The Neotropical Leishmania species: a brief historical review of their discovery, ecology and taxonomy’. doi: 10.5123/S2176-62232010000200002.

Latkin, C. A. *et al.* (2013) ‘Neighborhoods and HIV: A social ecological approach to prevention and care’, *American Psychologist*. American Psychological Association Inc., 68(4), pp. 210–224. doi: 10.1037/a0032704.

Lau, C. L. *et al.* (2010) ‘Climate change, flooding, urbanisation and leptospirosis: Fuelling the fire?’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, pp. 631–638. doi: 10.1016/j.trstmh.2010.07.002.

Lau, C. L. and DePasquale, J. M. (2012) ‘Leptospirosis, diagnostic challenges, American Samoa.’,

Emerging infectious diseases. Centers for Disease Control and Prevention, 18(12), pp. 2079–81. doi: 10.3201/eid1812.120429.

Lead, C.-O. *et al.* (no date) *Observed Climate Variability and Change 2 Contents*.

lesphore Brou, T. *et al.* (no date) 'Landscape Diversity Related to Buruli Ulcer Disease in Cô te d'Ivoire'. doi: 10.1371/journal.pntd.0000271.

Li, W. *et al.* (2005) 'Bats are natural reservoirs of SARS-like coronaviruses', *Science*, 310(5748), pp. 676–679. doi: 10.1126/science.1118391.

Lo, M. K. and Rota, P. A. (2008) 'The emergence of Nipah virus, a highly pathogenic paramyxovirus', *Journal of Clinical Virology*, 43(4), pp. 396–400. doi: 10.1016/j.jcv.2008.08.007.

Loftus, M. J. *et al.* (2018) 'The incubation period of Buruli ulcer (Mycobacterium ulcerans infection) in Victoria, Australia - Remains similar despite changing geographic distribution of disease.', *PLoS neglected tropical diseases*. Public Library of Science, 12(3), p. e0006323. doi: 10.1371/journal.pntd.0006323.

Loiseau, R. *et al.* (2019) 'American cutaneous leishmaniasis in French Guiana: an epidemiological update and study of environmental risk factors', *International Journal of Dermatology*. Blackwell Publishing Ltd, 58(11), pp. 1323–1328. doi: 10.1111/ijd.14625.

Luis, A. D. *et al.* (2013) 'A comparison of bats and rodents as reservoirs of zoonotic viruses: Are bats special?', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 280(1756). doi: 10.1098/rspb.2012.2753.

Mackey, T. K. *et al.* (2014) 'Emerging and reemerging neglected tropical diseases: A review of key characteristics, Risk factors, And the policy and innovation environment', *Clinical Microbiology Reviews*. American Society for Microbiology, 27(4), pp. 949–979. doi: 10.1128/CMR.00045-14.

Mantey, S. *et al.* (2012) 'Spatial Dependency of Buruli Ulcer Disease on Geological Settings in Ghana',

International Journal of Science and Research (IJSR) ISSN (Online Impact Factor, 3(9), pp. 2319–7064.

Marcheggiani, S. *et al.* (2010) 'Risks of water-borne disease outbreaks after extreme events', *Toxicological and Environmental Chemistry*. Taylor & Francis Group , 92(3), pp. 593–599. doi: 10.1080/02772240903252140.

Marston, B. J. *et al.* (1995) 'Emergence of Buruli Ulcer Disease in the Daloa Region of Cote D'Ivoire', *The American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 52(3), pp. 219–224. doi: 10.4269/ajtmh.1995.52.219.

Martin-Blondel, G. *et al.* (2015) 'Outbreak of Leishmania braziliensis cutaneous Leishmaniasis, Saül, French Guiana', *Emerging Infectious Diseases*. Centers for Disease Control and Prevention (CDC), pp. 892–894. doi: 10.3201/eid2105.141181.

Martin, R. J. *et al.* (no date) *Leptospiral Interspecies Infections on an Illinois Farm*. Available at: https://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC1919866&blobtype=pdf (Accessed: 24 January 2019).

Maruthappu, M. *et al.* (2017) 'Unemployment and HIV mortality in the countries of the Organisation for Economic Co-operation and Development: 1981–2009', *JRSM Open*. SAGE Publications, 8(7), p. 205427041668520. doi: 10.1177/2054270416685206.

Mayer, S. E. and Jencks, C. (no date) 'Growing up in Poor Neighborhoods: How Much Does it Matter?', *Science*. American Association for the Advancement of Science, pp. 1441–1445. doi: 10.2307/1703125.

McMichael, A. J. and Butler, C. D. (2011) 'Promoting Global Population Health While Constraining the Environmental Footprint', *Annual Review of Public Health*. Annual Reviews, 32(1), pp. 179–197. doi: 10.1146/annurev-publhealth-031210-101203.

McMichael, A. J., Woodruff, R. E. and Hales, S. (2006) 'Climate change and human health: Present and future risks', *Lancet*. Elsevier Limited, pp. 859–869. doi: 10.1016/S0140-6736(06)68079-3.

Measuring the Daily Destruction of the World's Rainforests - Scientific American (no date). Available at: https://www.scientificamerican.com/article/earth-talks-daily-destruction/ (Accessed: 4 May 2020).

Medlock, J. M. *et al.* (2012) 'A review of the invasive mosquitoes in Europe: Ecology, public health risks, and control options', *Vector-Borne and Zoonotic Diseases*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA , pp. 435–447. doi: 10.1089/vbz.2011.0814.

Meissner WHO, D. (2010) *Working to Overcome the Global Impact of Neglected Tropical Diseases Annexe I First WHO report on neglected tropical diseases*.

Merritt, R. W. *et al.* (2010) 'Ecology and Transmission of Buruli Ulcer Disease: A Systematic Review', *PLoS Neglected Tropical Diseases*. Edited by R. O. Phillips. Public Library of Science, 4(12), p. e911. doi: 10.1371/journal.pntd.0000911.

Merritt, R. W., Benbow, M. E. and Small, P. L. C. (2005) 'Unraveling an emerging disease associated with disturbed aquatic environments: the case of Buruli ulcer', *Frontiers in Ecology and the Environment*. John Wiley & Sons, Ltd, 3(6), pp. 323–331. doi: 10.1890/1540-9295(2005)003[0323:UAEDAW]2.0.CO;2.

Morris, A., Gozlan, R. E., *et al.* (2014) 'Complex temporal climate signals drive the emergence of human water-borne disease.', *Emerging microbes & infections*, 3(8), p. e56. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26038751.

Morris, A., Gozlan, R., *et al.* (2014) 'First Detection of Mycobacterium ulcerans DNA in Environmental Samples from South America', *PLoS Neglected Tropical Diseases*. doi: 10.1371/journal.pntd.0002660.

Morris, A. L. *et al.* (2016) 'Deforestation-driven food-web collapse linked to emerging tropical infectious disease, Mycobacterium ulcerans.', *Science advances*, 2(12), p. e1600387. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27957534.

Morse, S. S. (1991) 'Emerging viruses: Defining the rules for viral traffic', *Perspectives in Biology and Medicine*. Perspect Biol Med, 34(3), pp. 387–409. doi: 10.1353/pbm.1991.0038.

Muchaal, P. K. and Ngandjui, G. (1999) ‘Impact of Village Hunting on Wildlife Populations in the Western Dja Reserve, Cameroon’, *Conservation Biology*. John Wiley & Sons, Ltd, 13(2), pp. 385–396. doi: 10.1046/j.1523-1739.1999.013002385.x.

Murray, K. A. *et al.* (2015) ‘Global biogeography of human infectious diseases’, *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1507442112.

Myzabella, N. *et al.* (2019) ‘Occupational health and safety in the palm oil industry: A systematic review’, *International Journal of Occupational and Environmental Medicine*. NIOC Health Organization, pp. 159–173. doi: 10.15171/ijoem.2019.1576.

Nacher, M. *et al.* (2010) ‘What is driving the HIV epidemic in French Guiana?’, *International Journal of STD and AIDS*, 21(5), pp. 359–361. doi: 10.1258/ijsa.2010.009570.

Nacher, M. *et al.* (2011) ‘Short report: What is AIDS in the Amazon and the Guianas? Establishing the burden of disseminated histoplasmosis’, *American Journal of Tropical Medicine and Hygiene*, 84(2), pp. 239–240. doi: 10.4269/ajtmh.2011.10-0251.

Nacher, M. *et al.* (2014) ‘Risk Factors for Disseminated Histoplasmosis in a Cohort of HIV-Infected Patients in French Guiana’, *PLoS Neglected Tropical Diseases*. Public Library of Science, 8(1), p. 7. doi: 10.1371/journal.pntd.0002638.

Nacher, M. *et al.* (2016) ‘Disseminated histoplasmosis in Central and South America, the invisible elephant: The lethal blind spot of international health organizations’, *AIDS*. Lippincott Williams and Wilkins, pp. 167–170. doi: 10.1097/QAD.0000000000000961.

Nacher, M. *et al.* (2018) ‘Temporal trend of the proportion of patients presenting with advanced HIV in French Guiana: Stuck on the asymptote?’, *BMC Research Notes*. BioMed Central Ltd., 11(1). doi: 10.1186/s13104-018-3944-y.

Nacher, M. *et al.* (2019) ‘The Fight against HIV-Associated Disseminated Histoplasmosis in the Americas: Unfolding the Different Stories of Four Centers’, *Journal of Fungi*, 5(2), p. 51. doi:

10.3390/jof5020051.

Niloofa, R. *et al.* (2015) ‘Diagnosis of leptospirosis: Comparison between microscopic agglutination test, IgM-ELISA and IgM rapid immunochromatography test’, *PLoS ONE*. Public Library of Science, 10(6). doi: 10.1371/journal.pone.0129236.

O’Brien, D. P. *et al.* (2018) ‘Tackling the worsening epidemic of Buruli ulcer in Australia in an information void: time for an urgent scientific response.’, *The Medical journal of Australia*. Australia, 208(7), pp. 287–289.

O’Brien, D. P. *et al.* (2019) ‘The changing epidemiology worldwide of Mycobacterium ulcerans’, *Epidemiology and Infection*. Cambridge University Press. doi: 10.1017/S0950268818002662.

Ogden, N., AbdelMalik, P. and Pulliam, J. (2017) ‘Emerging infectious diseases: prediction and detection’, *Canada Communicable Disease Report*. Infectious Disease and Control Branch (IDPCB) - Public Health Agency of Canada, 43(10), pp. 206–211. doi: 10.14745/ccdr.v43i10a03.

Ojikutu, B. O. *et al.* (2018) ‘Neighborhood crime and sexual transmission risk behavior among black men living with HIV’, *Journal of Health Care for the Poor and Underserved*. Johns Hopkins University Press, 29(1), pp. 383–399. doi: 10.1353/hpu.2018.0026.

Okaka, F. O. and Odhiambo, B. D. O. (2018) ‘Relationship between Flooding and Out Break of Infectious Diseasesin Kenya: A Review of the Literature’, *Journal of Environmental and Public Health*. Hindawi Limited, 2018. doi: 10.1155/2018/5452938.

Olivero, J. *et al.* (2017) ‘Recent loss of closed forests is associated with Ebola virus disease outbreaks’, *Scientific Reports*. Nature Publishing Group, 7(1), p. 14291. doi: 10.1038/s41598-017-14727-9.

Omer, S. B. *et al.* (2006) ‘Nonmedical exemptions to school immunization requirements: secular trends and association of state policies with pertussis incidence.’, *JAMA*, 296(14), pp. 1757–63. doi: 10.1001/jama.296.14.1757.

Omumbo, J. A. *et al.* (2011) 'Raised temperatures over the Kericho tea estates: Revisiting the climate in the East African highlands malaria debate', *Malaria Journal*. BioMed Central, 10(1), p. 12. doi: 10.1186/1475-2875-10-12.

Ostfeld, R. S., Keesing, F. and Eviner, V. T. (eds) (2008) *Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems*. Princeton University Press, Princeton, New Jersey. Available at: http://press.princeton.edu/titles/8692.html.

Ostyn, B. *et al.* (2015) 'Transmission of Leishmania donovani in the Hills of Eastern Nepal, an Outbreak Investigation in Okhaldhunga and Bhojpur Districts', *PLOS Neglected Tropical Diseases*. Edited by A. Warburg. Public Library of Science, 9(8), p. e0003966. doi: 10.1371/journal.pntd.0003966.

Patz, J. A. *et al.* (2008) 'Disease Emergence from Global Climate and Land Use Change', *Medical Clinics of North America*. Med Clin North Am, pp. 1473–1491. doi: 10.1016/j.mcna.2008.07.007.

Pereira, H. M., Navarro, L. M. and Martins, I. S. (2012) 'Global Biodiversity Change: The Bad, the Good, and the Unknown', *Annual Review of Environment and Resources*. Annual Reviews , 37(1), pp. 25–50. doi: 10.1146/annurev-environ-042911-093511.

Perkin, E. *et al.* (2011) *The influence of artificial light on stream and riparian ecosystems: Questions, challenges, and perspectives*, *Ecosphere*. doi: 10.1890/ES11-00241.1.

Pielou, E. C. (1979) *Biogeography*. Wiley.

Pittiglio, C. *et al.* (no date) *Africa - El Niño and increased risk of Rift Valley fever - Warning to countries*.

Quillian, L. and Pager, D. (2001) *Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime 1*, *AJS*.

Ramasamy, R. and Surendran, S. N. (2011) 'Possible impact of rising sea levels on vector-borne infectious diseases', *BMC Infectious Diseases*. BioMed Central, 11(1), p. 18. doi: 10.1186/1471-2334-11-18.

Randall Hughes, A. *et al.* (2007) 'Reciprocal relationships and potential feedbacks between biodiversity and disturbance', *Ecology Letters*. Ecol Lett, pp. 849–864. doi: 10.1111/j.1461-0248.2007.01075.x.

Rands, M. R. W. *et al.* (2010) 'Biodiversity conservation: Challenges beyond 2010', *Science*. American Association for the Advancement of Science, pp. 1298–1303. doi: 10.1126/science.1189138.

Rangel, E. F. and Lainson, R. (2009) 'Proven and putative vectors of American cutaneous leishmaniasis in Brazil: Aspects of their biology and vectorial competence', *Memorias do Instituto Oswaldo Cruz*. Fundacao Oswaldo Cruz, 104(7), pp. 937–954. doi: 10.1590/S0074-02762009000700001.

van Ravensway, J. *et al.* (2012) 'Climate and landscape factors associated with Buruli ulcer incidence in Victoria, Australia.', *PloS one*. Public Library of Science, 7(12), p. e51074. doi: 10.1371/journal.pone.0051074.

Redding, D. W. *et al.* (2017) 'Evaluating Bayesian spatial methods for modelling species distributions with clumped and restricted occurrence data', *PLOS ONE*. Edited by L. Bosso. Public Library of Science, 12(11), p. e0187602. doi: 10.1371/journal.pone.0187602.

Robinson, T. P. *et al.* (2014) 'Mapping the global distribution of livestock', *PLoS ONE*. Public Library of Science, 9(5). doi: 10.1371/journal.pone.0096084.

Roche, B. and Guégan, J.-F. (2011) 'Ecosystem dynamics, biological diversity and emerging infectious diseases', *Comptes Rendus Biologies*, 334(5), pp. 385–392. doi: https://doi.org/10.1016/j.crvi.2011.02.008.

Rodrigues, M. G. de A. *et al.* (2019) 'The role of deforestation on American cutaneous leishmaniasis incidence: spatial-temporal distribution, environmental and socioeconomic factors associated in the Brazilian Amazon', *Tropical Medicine & International Health*. Blackwell Publishing Ltd, 24(3), pp. 348–355. doi: 10.1111/tmi.13196.

Rodwin, V. and Gusmano, M. (2002) 'The World Cities Project: Rationale, Organization, and Design for Comparison of Megacity Health Systems', *Journal of urban health : bulletin of the New York Academy of*

Medicine. J Urban Health, 79(4). doi: 10.1093/JURBAN/79.4.445.

Roger, A. *et al.* (2013) 'Climate and leishmaniasis in french guiana', *American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 89(3), pp. 564–569. doi: 10.4269/ajtmh.12-0771.

Rood, E. J. J. *et al.* (2017) 'Environmental risk of leptospirosis infections in the Netherlands: Spatial modelling of environmental risk factors of leptospirosis in the Netherlands', *PLoS ONE*. Public Library of Science, 12(10). doi: 10.1371/journal.pone.0186987.

Rosales, J. C., Yang, H. M. and Avila Blas, O. J. (2014) 'Variability Modeling of Rainfall, Deforestation, and Incidence of American Tegumentary Leishmaniasis in Orán, Argentina, 1985–2007', *Interdisciplinary Perspectives on Infectious Diseases*. Edited by A. Grinberg. Hindawi Publishing Corporation, 2014, p. 461013. doi: 10.1155/2014/461013.

Rotureau, B. (2006) 'Ecology of the Leishmania Species in the Guianan Ecoregion Complex - PubMed', *The American journal of tropical medicine and hygiene*. Available at: https://pubmed.ncbi.nlm.nih.gov/16407350/ (Accessed: 10 June 2020).

Rougeron, V. *et al.* (2011) 'First clinical case of cutaneous leishmaniasis due to Leishmania (Viannia) braziliensis in a domestic cat from French Guiana.', *Veterinary parasitology*, 181(2–4), pp. 325–8. doi: 10.1016/j.vetpar.2011.04.028.

Schmidt, K. A. and Ostfeld, R. S. (2001) 'BIODIVERSITY AND THE DILUTION EFFECT IN DISEASE ECOLOGY', *Ecology*. John Wiley & Sons, Ltd, 82(3), pp. 609–619. doi: 10.1890/0012-9658(2001)082[0609:BATDEI]2.0.CO;2.

Schneider, M. C. and Machado, G. (2018) 'Environmental and socioeconomic drivers in infectious disease', *The Lancet Planetary Health*. Elsevier B.V., pp. e198–e199. doi: 10.1016/S2542-5196(18)30069-X.

Schwartz, E., Hatz, C. and Blum, J. (2006) 'New world cutaneous leishmaniasis in travellers', *Lancet*

Infectious Diseases. Elsevier, pp. 342–349. doi: 10.1016/S1473-3099(06)70492-3.

Siettos, C. I. and Russo, L. (2013) 'Mathematical modeling of infectious disease dynamics', *Virulence*. Taylor and Francis Inc., pp. 295–306. doi: 10.4161/viru.24041.

Simon, S. *et al.* (2017) 'Cutaneous leishmaniasis in French Guiana: Revising epidemiology with PCR-RFLP', *Tropical Medicine and Health*. BioMed Central Ltd., 45(1). doi: 10.1186/s41182-017-0045-x.

Singer, M. C. *et al.* (2006) 'Syndemics, sex and the city: Understanding sexually transmitted diseases in social and cultural context', *Social Science and Medicine*, 63(8), pp. 2010–2021. doi: 10.1016/j.socscimed.2006.05.012.

Smith, K. M. *et al.* (2012) 'Zoonotic viruses associated with illegally imported wildlife products', *PLoS ONE*, 7(1). doi: 10.1371/journal.pone.0029505.

Smithwick, E. A. H. and Wu, J. (2016) 'Landscape Fragmentation as a Risk Factor for Buruli Ulcer Disease in Ghana', *The American Journal of Tropical Medicine and Hygiene*. doi: 10.4269/ajtmh.15-0647.

Spiegelhalter, D. J. *et al.* (2002) *Bayesian measures of model complexity and fit (with discussion)*, *Journal of the Royal Statistical Society, Series B*. doi: 10.1111/1467-9868.02022.

Strazzulla, A. *et al.* (2013) 'Mucosal Leishmaniasis: An Underestimated Presentation of a Neglected Disease', *BioMed Research International*, 2013, pp. 1–7. doi: 10.1155/2013/805108.

Taylor, L. H., Latham, S. M. and Woolhouse, M. E. J. (2001) 'Risk factors for human disease emergence', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 356(1411), pp. 983–989. doi: 10.1098/rstb.2001.0888.

Tefft, N. and Kageleiry, A. (2014) 'State-level unemployment and the utilization of preventive medical services', *Health Services Research*, 49(1), pp. 186–205. doi: 10.1111/1475-6773.12091.

Tesla, B. *et al.* (2018) 'Temperature drives Zika virus transmission: Evidence from empirical and

mathematical models', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society Publishing, 285(1884). doi: 10.1098/rspb.2018.0795.

Trenberth, K. E. *et al.* (2007) 'Observations: Surface and Atmospheric Climate Change'. Cambridge University Press, pp. 235–336. Available at: https://archive-ouverte.unige.ch/unige:18698 (Accessed: 11 September 2019).

Veitch, M. G. K. *et al.* (1997) *A large localized outbreak of Mycobacterium ulcerans infection on a temperate southern Australian island, Epidemiol. Infect*. Cambridge University Press. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2809003/pdf/9440434.pdf (Accessed: 21 February 2019).

Vescio, F. M. *et al.* (2012) 'Environmental correlates of crimean-congo haemorrhagic fever incidence in Bulgaria', *BMC Public Health*, 12(1), p. 1116. doi: 10.1186/1471-2458-12-1116.

Wagner, T., Benbow, M Eric, *et al.* (2008) 'A Landscape-based model for predicting Mycobacterium ulcerans infection (Buruli Ulcer disease) presence in Benin, West Africa.', *EcoHealth*, pp. 69–79. Available at: https://www.ncbi.nlm.nih.gov/pubmed/18648799.

Wagner, T., Benbow, M. Eric, *et al.* (2008) 'Buruli ulcer disease prevalence in Benin, West Africa: Associations with land use/cover and the identification of disease clusters', *International Journal of Health Geographics*. doi: 10.1186/1476-072X-7-25.

Wang, L. F. and Eaton, B. T. (2007) 'Bats, civets and the emergence of SARS', *Current Topics in Microbiology and Immunology*, pp. 325–344. doi: 10.1007/978-3-540-70962-6_13.

Weiler, M. *et al.* (2017) 'Tick abundance: a one year study on the impact of flood events along the banks of the river Danube, Austria', *Experimental and Applied Acarology*. Springer International Publishing, 71(2), pp. 151–157. doi: 10.1007/s10493-017-0114-1.

van der Werf, T. S. *et al.* (1989) 'Mycobacterium ulcerans infection in Ashanti region, Ghana.', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 83(3), pp. 410–413. Available at:

https://www.ncbi.nlm.nih.gov/pubmed/2617591.

‘WHO | Flooding and communicable diseases fact sheet’ (2012) *WHO*. World Health Organization. Available at: https://www.who.int/hac/techguidance/ems/flood_cds/en/ (Accessed: 22 January 2019).

‘WHO | List of Blueprint priority diseases’ (2018) *WHO*. World Health Organization.

Wilcox, B. and Ellis, B. (no date) ‘Forests and emerging infectious diseases of humans’.

Wolfe, N. D. *et al.* (2005) ‘Bushmeat hunting, deforestation, and prediction of zoonotic disease emergence’, *Emerging Infectious Diseases*. Centers for Disease Control and Prevention (CDC), pp. 1822–1827. doi: 10.3201/eid1112.040789.

Woolhouse, M. (2011) ‘How to make predictions about future infectious disease risks’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, pp. 2045–2054. doi: 10.1098/rstb.2010.0387.

World Health Organization (2012) ‘Treatment of mycobacterium ulcerans disease (buruli ulcer): guidance for health workers. World Health Organization.’ Available at: http://www.who.int/iris/handle/10665/77771.

World Health Organization (2016) ‘Artisanal and small-scale gold mining and health - Environmental and Occupational Health Hazards Associated with Artisinal and Small-scale Gold Mining’, p. 36. Available at: http://apps.who.int/iris/bitstream/10665/247195/1/9789241510271-eng.pdf?ua=1.

Appendices

Appendix Table 2.1: Blue print diseases (BDPs) geographic data point of outbreak emergence

Virus	Country	Site of origin	Month/Year	Lat (y)	Long (x)
Ebola	Uganda	Gulu municipality	10/2000	2.8599	32.2801
Ebola	Gabon	Medemba village, La Zadie district	11/2001	0.8439	14.0704
Ebola	Republic of Congo	Mbomo, Mbomo district	12/2002	-2.353	13.621
Ebola	South Sudan	Yambio	05/2004	4.5721	28.3955
Ebola	Republic of Congo	Etoumbi	04/2005	0.01667	14.8999
Ebola	Democratic Republic of Congo	Mweka	09/2007	-4.85187	21.5595
		Luebo		-5.35218	21.42192
Ebola	Uganda	Bundibugyo	08/2007	0.71117	30.06469
Ebola	Uganda	Nakisamata village, Luwero District	06/2011	0.6444	32.7286
Ebola	Uganda	Kibaale	07/2012	0.9833	31.0833
Ebola	DRC	Isiro	08/2012	2.8666	27.6667
Ebola	Uganda	Luwero	11/2012	0.8492	32.4731
Ebola	DRC	Inkanamongo village, near Boende	07/2014	-0.2816	20.8805
Ebola	Guinea	Meliandou, Gueckedou District	12/2013	8.6167	-10.0611
Ebola	DRC	Likati	05/2017	2.8833	24.0500
Ebola	DRC	Bikoro	04/2018	-0.75	18.116667
Ebola	DRC	Mangina	05/2018	0.604421	29.308032
Marburg	Uganda	Kabale	10/2012	-1.24857	29.98993
Marburg	Angola	Uige	10/2004	-7.5749	15.0678
Marburg	Uganda	Kamwenge	06/2007	0.1866	30.4539

Mar-burg	Uganda	Cave in Queen Elizabeth Forest	12/2007	-0.4011	30.0477
Mar-burg	Uganda	Cave in Queen Elizabeth Forest	06/2008	-0.4011	30.0477
Mar-burg	Uganda	Kampala-Mpigi	09/2014	0.2476	32.3493
RVF	Saudi Arabia	Jizan region	08/2000	17.1864	42.6943
RVF	Kenya	Near Garissa	12/2006	-0.4563	39.6985
RVF	South Africa	Free State	03/2010	-28	27
RVF	Mauritania	Atar	09/2010	20.5108	-13.0248
RVF	Mauritania	Assaba	09/2015	16.5833	-11.5833
RVF	Mauritania	Assaba	09/2012	16.5833	-11.5833
RVF	Uganda	Kabale	03/2016	-1.24857	29.98993
RVF	Niger	Tchintabaraden	08/2016	15.8999	5.8048
RVF	Senegal	Linguere	09/2013	15.3972	-15.1154
MERS	Saudi Arabia	Jeddah	06/2012	21.54238	39.19797
Nipah	Bangladesh	Faridpur district	04/2004	23.5705	89.8618
Nipah	India	Kozhikode	07/2018	11.25	75.77
Nipah	Bangladesh	Meherpur district	04/2000	23.7685	88.6608
Nipah	Bangladesh	Rajbari	02/2008	23.7649	89.6505
Nipah	Bangladesh	Manikganj	02/2008	23.8667	89.95
Nipah	Bangladesh	Faridpur district	12/2009	23.5705	89.8618
Nipah	Bangladesh	Khulna	01/2020	22.8768	89.5114
Nipah	Bangladesh	Naogaon	1/2003	24.8025	8.9490
Nipah	Bangladesh	Tangali	12/2004	24.2482	89.9196
Nipah	Bangladesh	Thakurgaon	02/2007	26.0266	88.4558
Nipah	Bangladesh	Kushtia	03/2007	23.9010	89.1324

Nipah	Bangladesh	Lalmonirhat	02/2011	25.9182	89.4581
SARS	China	Shunde district	11/2002	22.7685	113.1951

Appendix Table 2.2: Reservoirs of Blueprint priority diseases (BPDs) studied in this study

Virus	Reservoir	Order	Latin Name
MERS	Serotine bat	Chiroptera	Neoromicia_zuluensis
MERS	Egyptian tomb bat	Chiroptera	Taphozous_perforatus
MERS	Greater mouse-tailed bat	Chiroptera	Rhinopoma_microphyllum
MERS	Kuhl's pipistrelle	Chiroptera	Pipistrellus_kuhlii
MERS	Botta's serotine	Chiroptera	Eptesicus_bottae
MERS	Straw-colored fruit bat	Chiroptera	Eidolon_helvum
MERS	Egyptian fruit bat	Chiroptera	Rousettus_aegyptiacus
Nipah	Fruit bats/ Flying foxes		Pteropus_anetianus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_aruensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_caniceps
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_capistratus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_chrysoproctus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_cognatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_conspicillatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_dasymallus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_faunulus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_fundatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_giganteus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_gilliardorum
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_griseus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_howensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_hypomelanus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_intermedius
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_keyensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_leucopterus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_livingstonii
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_lombocensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_loochoensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_lylei
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_macrotis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_mahaganus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_mariannus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_melanopogon
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_melanotus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_molossinus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_neohibernicus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_niger
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_nitendiensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_ocularis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_ornatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_pelewensis

Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_personatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_pohlei
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_poliocephalus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_pselaphon
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_pumilus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_rayneri
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_rennelli
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_rodricensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_rufus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_samoensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_scapulatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_seychellensis
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_speciosus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_temminckii
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_tonganus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_tuberculatus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_ualanus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_vampyrus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_vetulus
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_voeltzkowi
Nipah	Fruit bats/ Flying foxes	Chiroptera	Pteropus_woodfordi
Ebola	Greater Long-fingered bat	Chiroptera	Miniopterus_inflatus
Ebola	Franquet's epauletted fruit bat	Chiroptera	Epomops_franqueti
Ebola	Hammer-headed fruit bat	Chiroptera	Hypsignathus_monstrosus
Ebola	Little collared fruit bats	Chiroptera	Myonycteris_torquata
Marburg	Egyptian fruit bat	Chiroptera	Rousettus_aegyptiacus
SARS	Greater horseshoe bat	Chiroptera	Rhinolophus_ferrumequi-num
SARS	Civet	Carnivora	Paguma_larvata

Appendix Table 3.1: Datasets and statistical models used in the study

Model	**Objective**	**Variables used**	**Time Period**	**Model**	**R package used**
Preliminary Cluster analysis	Detection of significant spatial structure	Leptospirosis and BU case as spatial points	Leptospirosis: 2007 - 2017 BU: 1969 - 2017	Poisson regression	satscan
Topographic model	To correlate elevation and flooding to incidence of Leptospirosis and BU in non-spatial and spatial models.	Mean, minimum, maximum measures of the elevation and TWI calculated from DEM of 30m. Leptospirosis and BU case as spatial points along with the randomly generated spatial points.	Leptospirosis: 2007 - 2017 BU: 1969 - 2017	Non-spatial vs Spatial Logistic regression	Stats MCMCglmm spBayes
Land cover model	Identification of land cover type that flavors the incidence of Leptospirosis and BU in non-spatial and spatial models.	Proportion of land cover in spatial buffers of 2km, 5km and 10km radii around Leptospirosis, BU cases and random spatiotemporal points.	Leptospirosis: 2007 - 2017 BU: 2000 - 2017	Non-spatial vs Spatial Logistic regression	Stats MCMCglmm spBayes
Meteorological model	To correlate climatic covariates to the incidence of Leptospirosis and BU in non-spatial and spatial models.	Interpolated from points to climate grids of 30m resolution using an Inverse Distance Weighted (IDW) approach Leptospirosis (months 0&-1) and BU case data (months 0 upto -6) as spatial points along with the randomly generated spatiotemporal points.	Leptospirosis: 2007 - 2017 BU: 2000 - 2017	Non-spatial vs Spatial Logistic regression	Stats MCMCglmm spBayes

Appendix Table 5.1: Significant descriptive statistics of the Hepatitis and Histoplasmosis clusters.

Variables measured	Hepatitis			Histoplasmosis		
	Positive n=158	Negative n=1966	Difference ± SE [95% CI ; P value]	Positive n=130	Negative n=2011	Difference ± SE [95% CI ; P value]
Male (%)	100 (63.3)	904 (46.0)	17.3 [9.3-24.8;<0.0001]	83 (63.8)	928 (46.2)	15.3 [8.9-25.7;<0.0001]
Mortality (%)	23 (14.6)	165 (8.39)	6.2 [1.3-12.6; 0.009]	21 (16.5)	167 (8.3)	7.9 [2.4-15.2; 0.002]
Average CD4 counts at nadir (SD)	169.3 (195.2)	210.0 (185.2)	46.1±15.4 [10.5-70.8; 0.008]	55.6 (75.8)	215.4 (195.9)	159.9±17.3 [126.0-193.7; <0.0001]