HAL
open science

Modélisation multi-échelle de la dynamique spatiale de la Dengue : application à la Nouvelle-Calédonie et à la région Pacifique

Magali Teurlai

To cite this version:

Magali Teurlai. Modélisation multi-échelle de la dynamique spatiale de la Dengue : application à la Nouvelle-Calédonie et à la région Pacifique. Modélisation et simulation. Université Montpellier II - Sciences et Techniques du Languedoc, 2014. Français. NNT : 2014MON20167 . tel-01875272

HAL Id: tel-01875272

https://theses.hal.science/tel-01875272

Submitted on 17 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Collège Doctoral du Languedoc-Roussillon

THÈSE
Pour obtenir le grade de
Docteur

Délivré par **l'Université Montpellier 2**

Préparée au sein de l'école doctorale Information Structures Systèmes
Et des unités de recherche UMR 182 LOCEAN et UMR 228 ESPACE-DEV

Spécialité : **Informatique**

Présentée par **Magali Teurlai**

Modélisation multi-échelle de la dynamique spatiale de la dengue

Application à la Nouvelle-Calédonie et à la région Pacifique

Soutenue le 18 décembre 2014 devant le jury composé de

Mr Bernard CAZELLES, Professeur, UPMC Paris 6	Examinateur
Mr Pierre GANÇARSKI, Professeur, Université de Strasbourg	Rapporteur
Mr Jean GAUDART, Maître de conférences, HDR, Université Aix-Marseille	Rapporteur
Mme Thérèse LIBOUREL, Professeur, Université Montpellier 2	Directeur de thèse
Mr Paulo Sergio LUCIO, Professeur, Universidade Federal do Rio Grande do Norte	Examinateur
Mr Morgan MANGEAS, CR1, IRD	Co-encadrant
Mr Christophe MENKES, CR1, IRD	Co-encadrant
Mr Alain VIARI, DR, INRIA	Président

UNIVERSITÉ MONTPELLIER 2
SCIENCES ET TECHNIQUES

« Les grandes personnes aiment les chiffres. Quand vous leur parlez d'un nouvel ami, elles ne vous questionnent jamais sur l'essentiel. Elles ne vous disent jamais « Quel est le son de sa voix ? Quels sont les jeux qu'il préfère ? Est-ce qu'il collectionne les papillons ? » Elles vous demandent : « Quel âge a-t-il ? Combien a-t-il de frères ? Combien pèse-t-il ? Combien gagne son père ? » Alors seulement elles croient le connaître… »

Antoine de Saint Exupéry

Le Petit Prince

REMERCIEMENTS

A Monsieur Pierre Gançarski, Professeur à l'Université de Strasbourg, et à Monsieur Jean Gaudart, Maître de conférence (HDR) à l'Université d'Aix-Marseille et praticien hospitalier à la faculté de Médecine de Marseille, qui nous ont fait l'honneur d'accepter de rapporter cette thèse,

Sincère reconnaissance

A Messieurs Paulo Sergio Lucio, Professeur à l'Universidade Federal do Rio Grande do Norte, et Alain Viari, Directeur de recherche à l'INRIA, qui nous ont fait l'honneur d'accepter de participer à ce jury thèse,

Hommages respectueux

A Monsieur Bernard Cazelles, Professeur à l'Université Pierre et Marie Curie, Paris 6, qui nous a fait l'honneur de participer à ce jury de thèse, et en remerciement des échanges scientifiques durant mon stage de Master 2,

Hommages respecteux

A la mémoire de Frédérique Messud-Petit, Maître de conférences en épidémiologie à l'Ecole Nationale Vétérinaire de Toulouse…

Cette thèse a été financée en partie par un prix d'encouragement à la recherche de la Province Sud de la Nouvelle-Calédonie, que je tiens à remercier ici. Je remercie également le centre IRD de Nouméa, pour l'accueil et l'hébergement.

Je tiens à remercier mes encadrants en Nouvelle-Calédonie. Merci pour votre disponibilité sans limite, je sais combien elle est précieuse, merci également pour ces échanges scientifiques incessants, extrêmement enrichissants du fait de nos différentes disciplines. Morgan, merci pour ces années d'échanges, ta vision claire et synthétique, et ton ouverture d'esprit face à mon esprit parfois un peu rebelle ;-). Christophe, après quatre années dans le bureau de tes stagiaires, les run WRF de la SPCZ sur JFM au format *ncdf* sont maintenant beaucoup moins mystérieux ! Merci pour la liberté totale que tu m'as accordée, tes remarques constructives, et ton attention aux détails, j'ai appris à aller jusqu'au bout du fond du micro-détail scientifique grâce à toi, et à ne jamais rien affirmer sans l'avoir vérifié 1 645 378 768 fois (oui, j'ai compté). Enfin, je vous remercie tous les deux infiniment pour votre compréhension et votre soutien lors des six premiers mois de thèse.

Merci également à Thérèse Libourel, qui a accepté de diriger cette thèse, malgré sa charge de travail et les 9 h de décalage horaire, pour son soutien, son efficacité, et son extrême réactivité.

Avec plus de 17 000 km entre l'Ecole Doctorale et moi, je ne serais jamais venue à bout des démarches administratives sans l'aide de Nicolas Serrurier, secrétaire de l'Ecole Doctorale I2S. Je pense que tu es en réalité un être créé par une puissance supérieure, venu sur Terre pour aider les thésards, qu'il vente, qu'il pleuve ou qu'il soit 3h du matin. MERCI ! (Pour la 365 789 657ème fois, oui j'ai compté aussi).

Merci à Nicolas Dégallier, qui avait initialement accepté la direction de cette thèse, pour sa perception fine des êtres, et ses idées toujours intéressantes sur le travail à réaliser.

Merci aux personnes de la DASS (Jean-Paul Grangeon, Patrice Nicot, Anne Pfannstiel, Maguy Daurès, Ludo, Michèle, Martine) pour leur collaboration et leur accueil. Merci aux personnes de la Mairie de Nouméa et du SIPRES (Kévin Lucien, Henrick Baraduc, et Odile Qaeze) pour leur collaboration. Merci aux personnes de l'IPNC (Ann-Claire Gourinat, Myrielle Dupont-Rouzeyrol, Laurent Guillaumot, et Cyrille Goarant) pour leurs informations précieuses en matière de diagnostic, de virologie moléculaire ou d'entomologie. Laurent, merci pour ces soirées passées au son d'une certaine flûte andine, la quena…

Je tiens à remercier tous les collègues qui ont également contribué au travail réalisé dans cette thèse, Jérôme Lefèvre, Pascal Douillet et Cyril Dutheil, de l'IRD, pour votre aide sur les données climatiques (et les run WRF de la SPCZ sur JFM !). Merci à Alizé Mercier, et à Lucie Lap, les « stagiaires-dengue », pour le travail impeccable réalisé sur Nouméa et le Pacifique. Merci à Annelise Tran, ainsi que Françoise Mathieu-Daudé, pour leur participation aux comités de thèse, et pour les différents échanges tout au long de la thèse. Merci à Emmanuel Roux, et Paulo pour leurs conseils sur les ACP. Merci à Bernard Cazelles, pour son aide sur les ondelettes durant mon stage M2. Merci à Elodie Descloux, avec qui j'ai

partagé la condition de doctorante « dengue » au sein de l'IRD une année durant, pour les « apéros scientifiques ». J'espère collaborer avec toi dans le futur, et faire plein d'autres apéros ! Merci également à Jean-Sébastien Dehecq, et François Taglioni, les collègues de la Réunion. Jean-Séb, ce fût un réel plaisir de te rencontrer, en espérant pouvoir organiser encore des déménagements de Bernard l'Hermite.

Merci aux membres de l'UMR-ESPACE-DEV, Céline, Catherine Sabinot, Flo (merci pour ta patience lors de la préparation des cours, et merci de m'avoir fait comprendre les arcanes matrcielles des ACP), Marc Despinoy, aux membres du GIE Océanide (Jean-Brice, Sven, Tonio, et Estelle), et aux informaticiens de l'IRD (Michel, François et Julien).

Merci au Docteur Virlon, sans les compétences et le professionnalisme duquel cette thèse n'aurait pu débuter.

Enfin, cette thèse est pour moi l'aboutissement d'une longue démarche scientifique, et je tiens à remercier ici l'ensemble de mes professeurs et maître de stage durant mon parcours. En particulier, mes professeurs de mathématiques au collège et au lycée (Mr Galzy et Mr Biscara), qui m'ont donné les bases nécessaires (quoique minces) au travail réalisé dans cette thèse. A Mme Hélène Cordoliani, professeur de biologie au lycée Saint-Louis, dont le visage m'est apparu plusieurs fois au cours de la rédaction du manuscrit, prononçant des phrases ancrées dans ma mémoire depuis 14 ans comme un baobab dans un astéroïde (« définissez les termes en introduction !», « Mon Dieu mes pauvres petits, vous avez tous des têtes d'enterrements. Allez vous aérer un peu »). Merci à Christophe Baehr, pour m'avoir fait retrouver la petite étincelle du gamin ébahit par les sciences, et ma passion scientifique. Merci à mes professeurs de l'Ecole Nationale Vétérinaire de Toulouse : à la mémoire de Madame Frédérique Messud-Petit, qui m'a appris les bases de l'épidémiologie, à Stéphane Bertagnoli, pour ses cours passionnants sur l'émergence des arboviroses en France, l'évolution des virus et le lien entre les phénomène d'émergence et la théorie du chaos, à Philippe Jacquiet qui aura su nous intéresser à la vie des vers, et à Monsieur Bodin, qui nous a présenté le corps des êtres vivants comme une véritable forteresse, résistant vaillamment, à coup d'anticorps et de cytokines aux assauts de cavaliers armés de toxines et de morceaux d'ADN. Ce sont ces cours de l'ENVT qui m'ont fait choisir la voie de l'épidémiologie, et je vous en remercie. Merci enfin à mes maîtres de stage, Flavie Goutard, Pascal Hendrikx, et Sirenda Vong, pour tout ce que vous m'avez apporté. Pour finir, mon parcours a aussi été l'occasion d'une rencontre humaine exceptionnelle : gracias, Ventura, el Angito. Te deseo lo mejor.

Un grand merci à l'ensemble des personnes du couloir d'océanographie (les deux Sophie, Fréd, Anne, les deux Jérôme, JB, Solange, Alex et ceux que j'oublie), pour vos petits mots de soutien, barres chocolatées et autres biscuits déposés sur mon bureau pendant la rédaction. Ça m'a réellement aidée !

Merci à la « thésard team », et autres (nombreux) stagiaires et post-doctorants de l'IRD : Christelle (130 trèfle !), Haizea, Séverin (plein de bonnes choses à vous trois !), Séb, Armelle, Laury (allez courage, fini l'harmonica à 5 h du matin :-)), Cyril, Margot (and the amazing color palette eyes), Lucie (alias « la fille

qui avait des ondes magiques face aux ordinateurs »), Simon, Adrien (l'homme qui comprend les stats plus vite que son ombre !), Laure, Laurent (plein de bonnes choses à vous trois !), Audrey (plein de bonnes choses à vous trois !). Merci pour toutes ces soirées partagées. Mention spéciale à Houssem, le Tunisien, et les oggias de retour de soirée à 3h du matin ;-). Merci pour ces soirées passées à refaire le monde, à parler de sujets aussi variés que Dieu, les mojitos, la théorie de la relativité, la finesse du sable sur les plages tunisiennes, ou ce qu'est le temps, au final…. Nos discussions me manquent.

Merci aux « filles », Fréd (Simon-Garel), Marine, Karine, Christelle et Laury. Puissent les soirées hammam être encore nombreuses.

Merci à Stéph, alias « Santiago », pour les soirées à philosopher, à Karine, pour son amitié et son soutien précieux au cours des derniers mois, et aux autres dauphins avec qui j'ai pu explorer le lagon (Stéph « Presidente », Chloé, Sosé, Nico, Manu) Le rhum arrangé nous attend ! Merci aussi à Aquanature, et aux dauphins toulousains, qui m'ont enseigné l'apnée, et permis de vivre des moments exceptionnels auprès des raies mantas : Jean, Gauhtier, Xav, Fabien et Jeff.

Merci à mes amis vétérinaires toulousains. L'éloignement aurait été difficile à tenir si je ne savais pas que vous serez là, toujours. Maud et Simon, Romu, Adrien, Ben, Guérric et Caro… Merci aussi à la clique des Saint-Louisards, Yann (merci pour les hébergements à Montpellier !), Gus et Pauline, Marie-Anne et Aurélie. J'ai hâte de renter passer un peu de temps avec vous, 3 ans, c'est trop long…

Merci à ma famille, qui a toujours été là pour me soutenir.

Enfin, merci Stéph, pour ces moments de bonheur évident, de liberté et d'humanité partagés à tes côtés, ces soirées karaokés, et ces moments de franche rigolade, à l'îlot Maître ou ailleurs. Ta présence a rendu ces 4 années exceptionnelles. Je ne peux exprimer ici tout ce que je ressens, mais le cœur y est.

Christophe, peut-être publieras-tu un jour tes écrits (je ne parle bien évidemment pas d'une certaine H…) ? En attendant, en hommage à ton âme poétique, et en remerciement des livres et des quelques soirées apaisantes passées, à ton insu, à t'écouter jouer de la flûte traversière, je vois dans cette thèse l'opportunité d'en publier au moins un :

OK Asteroideb639 de Ganymède.
des lunes et des planètes,
des cieux et des soleils,
des clouxes et des enxlumes,
des enfants et des femmes,
du pain et du vin
déride et des sourires.

Signé: Astéroid b612

Résumé français

Depuis les années 1970, les pays du Pacifique sont de plus en plus fréquemment touchés par des maladies vectorielles telles que la Dengue, le Chikungunya ou le Zika. Le contrôle de ces maladies nécessite la connaissance de leur distribution spatio-temporelle au sein de la population ainsi que la compréhension des facteurs et mécanismes, souvent multiples, régissant cette distribution. Dans cette thèse, nous nous intéressons à la modélisation spatio-temporelle des déterminants de la dynamique spatiale de la dengue à l'échelle régionale du Pacifique, l'échelle territoriale de la Nouvelle-Calédonie, et l'échelle d'une ville, Nouméa, capitale de la Nouvelle-Calédonie. Dans le Pacifique, la dengue survient sous la forme de vagues épidémiques successives dues à l'introduction et à la diffusion régionale d'un nouveau sérotype tous les cinq à sept ans. En Nouvelle-Calédonie, la dengue présente une dynamique épidémique saisonnière, le sérotype dominant étant celui circulant dans la région. L'émergence d'une épidémie nécessite des conditions climatiques précises, et un indicateur annuel prédictif du risque d'émergence est maintenant utilisé de manière opérationnelle par les autorités de santé. Sur le plan spatial, au cours d'une épidémie, en moyenne, la circulation du virus est plus intense dans les zones où la température moyenne ainsi que les densités locales de population sont élevées. Que ce soit sur le territoire entier ou dans la ville de Nouméa, lors de la ré-émergence d'un même sérotype, la diffusion spatiale du virus paraît limitée par l'immunité de population créée par les épidémies précédentes. Cette thèse permet de mettre en évidence la nature complexe et multi-factorielle des maladies vectorielles, et de souligner l'intérêt d'analyses multi-échelles pour l'étude de leur épidémiologie. Au-delà des résultats obtenus sur la dengue dans la région Pacifique, notre volonté était de développer un cadre méthodologique pour l'analyse spatio-temporelle des données de surveillance épidémiologique applicable à d'autres contextes géographiques ou épidémiologiques.

Mots-clés : Epidémiologie ; Modélisation spatio-temporelle ; Climat et environnement ; Maladies vectorielles ; Dengue ; Pacifique

Summary

Since the 1970's, the frequency of vector-borne diseases such as Dengue, Chikungunya or Zika has significantly increased in the Pacific region. Understanding the factors and mechanisms underlying the spatio-temporal distribution of these diseases provides useful information regarding their control and prevention. In this thesis, we identified dengue spatio-temporal patterns and used modeling tools to identify the factors associated with an increased epidemiological risk at a regional scale (Pacific), a territorial scale (New-Caledonia), and a city scale (Noumea, the capital of New-Caledonia). Every five to seven years, dengue spreads over the entire Pacific as large epidemics caused by the introduction and regional diffusion of one of the four dengue virus serotypes. In New Caledonia, dengue has a seasonal epidemic pattern. The emergence of an epidemic requires specific climatic conditions. The identification of these conditions led to the implementation of an operational early warning system to predict dengue annual epidemic risk. Spatially, at the territorial scale, during epidemic years, high levels of viral circulation are found in areas with higher mean temperature and higher local population densities. Whether at the territorial scale or at the city scale, the spatial diffusion of the virus during epidemics caused by the re-emergence of the same serotype seems limited by the population immunity created by past epidemics. This thesis highlights the complexity and the multi-factorial aspect of vector-borne diseases, and discusses the usefulness of a multi-scale approach in modelling their epidemiology. Besides enhancing our understanding of dengue epidemiology over the Pacific area, we also developed a methodological framework that can be used in other geographical or epidemiological settings for the spatio-temporal analysis and modeling of epidemiological surveillance data.

Key words : Epidemiology ; Spatio-temporal modelling ; Climate and environment ; Vector-borne diseases ; Dengue ; Pacific

REMERCIEMENTS *5*

RESUMES *11*

SOMMAIRE *13*

LISTE DES FIGURES *17*

LISTE DES TABLEAUX *19*

LISTE DES ABREVIATIONS *21*

LISTE DES PUBLICATIONS ASSOCIEES A LA THESE *23*

INTRODUCTION *25*

PARTIE I : CONTEXTE ET DONNEES 35

CHAPITRE 1 : Revue bibliographique sur la dengue et les facteurs influençant son épidémiologie 37

1.1. La dengue ---------- 39
1.1.1. L'agent pathogène 39
1.1.2. Les vecteurs 41
1.1.3. Formes cliniques et épidémiologie des formes cliniques 46
1.1.4. Diagnostic 48
1.1.5. Traitement 49
1.1.6. Prophylaxie médicale et sanitaire 49
1.1.7. Epidémiologie 50
1.2. Modélisation de la dynamique spatiale de la dengue : revue bibliographique ---------- 55
1.2.1. Objectifs de la modélisation 55
1.2.2. Les approches possibles en modélisation 56
1.2.3. Caractéristiques des études de modélisation 56
1.2.4. Description de la distribution spatio-temporelle des cas de dengue 57
1.2.5. Facteurs influençant la distribution spatiale de la dengue 58
1.2.6. Facteurs influençant la distribution temporelle de la dengue 61
1.2.7. Contraintes des études de modélisation 63
1.3. Conclusion ---------- 65

CHAPITRE 2 : Contexte Calédonien, données et modèles utilisés 67

2.1. Contexte calédonien ---------- 71
2.1.1. Contexte géographique et historique 71
2.1.2. Contexte épidémiologique 76
2.2. Données analysées ---------- 83
2.2.1. Echelle d'étude et résolution 83
2.2.2. Nature des données 84
2.2.3. Données épidémiologiques 86

2.2.4. Données entomologiques 94
2.2.5. Données climatiques 97
2.2.6. Données démographiques et socio-économiques 106
2.2.7. Données de télédétection 107
2.2.8. Données topologiques 108
2.3. Outils d'analyse et de modélisation 110
2.3.1. Les approches possibles en modélisation 110
2.3.1. Définition des cibles de modélisation 111
2.3.2. Les études possibles lors de l'analyse statistique de champs spatio-temporels 114
2.3.3. Modélisation statistique au cours des études descriptives 115
2.3.4. Modélisation statistique au cours des études analytiques 120
2.3.5. Modèles mécanistiques : modélisation des processus de transmission au niveau individuel 122
2.3.6. Décomposition orthogonale aux valeurs propres (« Empirical Orthogonal Functions ») 124

PARTIE II : ANALYSE ET MODELISATION MULTI-ECHELLE 137

CHAPITRE 3 : Analyse de la dengue à l'échelle du Pacifique 139

3.1. Analyse spatio-temporelle des données épidémiologiques 141
3.1.1. Transformation des données 141
3.1.2. Caractérisation globale de la distribution spatio-temporelle des cas de dengue 144
3.1.3. Analyse temporelle par pays 149
3.1.4. Analyse par vague pandémique 152
3.1.5. Discussion et perspectives 157

CHAPITRE 4 : Analyse de la dengue à l'échelle nationale ou territoriale 167

4.1. Analyse spatio-temporelle 169
4.1.1. Dynamique de la dengue en Nouvelle-Calédonie 169
4.1.2. Structure spatiale de la dengue en Nouvelle-Calédonie 171
4.1.3. Etude des schémas de propagation de la dengue par la méthode des ondelettes 173
4.1.4. Etude de la variabilité des schémas spatiaux de la dengue par la méthode des EOF : application à la Nouvelle-Calédonie 175
4.2. Analyse temporelle 183
4.2.1. Modélisation statistique non linéaire du risque épidémique annuel sur le territoire entier à partir de variables climatiques 183
4.3. Analyse spatiale : 206
4.3.1. Modélisation non linéaire de l'intensité de la circulation virale et projections futures 206

CHAPITRE 5 : Analyse de la dengue à l'échelle locale 247

5.1. Analyse spatio-temporelle 249
5.1.1. Dynamique de la dengue à Nouméa 249
5.1.1. Structure spatiale de la dengue à Nouméa 250

5.1.2. Etude de l'évolution temporelle des schémas spatiaux de la dengue par la méthode des EOF : application à Nouméa 253
5.2. Analyse temporelle ---------- 256
5.2.1. Modélisation déterministe du risque épidémique 256
5.3. Analyse spatiale ---------- 271
5.3.1. Objectifs et études réalisées 271
5.3.2. Enquête spatiale des facteurs de risque socio-économiques de la dengue à Nouméa... 271
5.3.3. Modélisation non linéaire de l'intensité de la circulation virale..... 274

CHAPITRE 6 : Discussion et perspectives 291

6.1. Synthèse des résultats ---------- 293
6.1.1. Tableau récapitulatif 293
6.1.1. Structure temporelle de la dengue et facteurs influençants..... 294
6.1.2. Structure spatiale de la dengue et facteurs influençants 298
6.2. Implications opérationnelles ---------- 301
6.3. Perspectives ---------- 303
6.3.1. Facteurs associés à la dynamique spatiale des épidémies dans le Pacifique 303
6.3.2. Facteurs associés à la dynamique spatiale des épidémies en Nouvelle-Calédonie 304
6.3.3. Mécanismes impliqués dans la dynamique spatiale 305
6.3.4. Estimation des paramètres de transmission en Nouvelle-Calédonie 306
6.3.5. Intégration des échelles..... 309
6.3.6. Autres arboviroses 309

CONCLUSION 311

BIBLIOGRAPHIE 313

ANNEXES..... 337

Annexe 1 : Glossaire ---------- 337
Annexe 2 : Formulaires de déclaration des cas de dengue en Nouvelle-Calédonie ---------- 341
Annexe 3 : Liste des champs disponibles dans les bases de données mises à disposition par la DASS-NC ---------- 343
Annexe 4 : Liste des codes quartier utilisés dans les différents fichiers ---------- 347
Annexe 5 : Formulaire utilisé pour la collecte des données entomologiques à Nouméa ---------- 349
Annexe 6 : Article 2 : étude des processus de propagation de la dengue au Cambodge par la technique d'analyse en ondelette ---------- 351
Annexe 7 : Equations utilisées pour l'analyse en ondelette ---------- 361
Annexe 8 : Questionnaire utilisé dans l'enquête spatiale des facteurs de risque socio-économiques de la dengue à Nouméa---------- 365

LISTE DES FIGURES

Figure 1 : Schéma des interactions (symbolisées par des flèches) entre l'hôte susceptible, le vecteur et l'agent pathogène menant à la dynamique observée des maladies vectorielles 30
Figure 1.2 : Le virus de la dengue..... 40
Figure 1.3 : Répartition des vecteurs de la dengue dans la Pacifique..... 41
Figure 1.4 : Cycle de vie des moustiques 42
Figure 1.5 : Cinétique des paramètres virologiques et sérologiques 48
Figure 2.1 : Carte du Pacifique et population par pays 71
Figure 2.2 : Positions des zones de convergence de la région Indo-Pacifique..... 72
Figure 2.3 : Densité de population dans les communes de Nouvelle-Calédonie 74
Figure 2.4 : Villes, tribus et stations météorologiques en Nouvelle-Calédonie 75
Figure 2.5 : Epandages d'insecticides à Nouméa. 83
Figure 2.6 : Représentation schématique des données analysées au cours de la thèse 85
Figure 2.7 : Carte des secteurs de la surveillance entomologique à Nouméa 96
Figure 2.8 : Carte des précipitations quotidiennes moyennes de Nouvelle-Calédonie obtenue par krigeage des moyennes quotidiennes de chaque station de 1991 à 2000 99
Figure 2.9 : Validation spatiale des variables atmosphériques simulées par le modèle WRF.....101
Figure 2.10 : Validation temporelle des variables atmosphériques simulées par le modèle WRF.....102
Figure 2.11 : Données historiques et évolution de la température moyenne et maximale à Nouméa à l'horizon 2100 105
Figure 2.12 : Image satellite RapiEye de Nouméa.....108
Figure 2.13 : Etudes possibles lors de l'analyse de champs spatio-temporels.....114
Figure 2.14 : Principe de l'analyse en ondelette.....117
Figure 2.15 : Exemple de l'utilisation de la méthode des ondelettes pour la détection de processus de propagation.....118
Figure 2.16 : Représentation géométrique des taux d'incidence spatio-temporels.....130
Figure 2.17 : Axe principal de la variabilité spatio-temporelle des taux d'incidence.....131
Figure 2.18 : Représentation de deux champs spatio-temporels de taux d'incidence.....132
Figure 2.19 : Comparaison des valeurs propres d'une analyse EOF effectuées sur deux jeux de données.....133
Figure 3.1 : Comparaison de deux méthodes pour définir les années épidémiques à l'échelle du Pacifique.....143
Figure 3.2 : Distribution spatio-temporelle des taux d'incidence annuels standardisés dans vingt PTIP (1971-2009) 144
Figure 3.3 : Carte du nombre d'épidémies entre 1971 et 2009 dans le Pacifique 146
Figure 3.4 : Evolution cyclique du nombre de pays touchés par une épidémie de dengue dans le Pacifique 147
Figure 3.5 : Evolution de la fréquence des épidémies au cours du temps dans le Pacifique 148
Figure 3.6 : Analyse EOF des taux d'incidence standardisés dans le Pacifique 149
Figure 3.7 : Evolution de la fréquence des épidémies dans le Pacifique 150
Figure 3.8 : Différents profils épidémiques dans les îles du Pacifique.....152
Figure 3.9 : Année de début de circulation d'un nouveau sérotype dans le Pacifique.....154
Figure 3.10 : Carte des pays les plus touchés par chacune des vagues pandémiques.....156
Figure 3.11 : Dynamique des quatre sérotypes de dengue dans le Pacifique entre 1971 et 2009 157
Figure 3.12 : Abondance relative des différents sérotypes en Thaïlande, 1973-2002 160
Figure 4.1 : Dynamique de la dengue en Nouvelle-Calédonie, 1971-2013.....169
Figure 4.2 : Structure spatiale des épidémies de dengue en Nouvelle-Calédonie, 1995-2012.....172

Figure 4.3 : Structure spatiale des anomalies des taux d'incidence annuels de la dengue en Nouvelle-Calédonie au cours des années épidémiques de la période 1995-2012 176
Figure 4.4 : Analyse EOF des taux d'incidence de la dengue en Nouvelle-Calédonie, période 1995-2012 177
Figure 4.5 : Analyse EOF des précipitations quotidiennes moyennes sur 24 stations météorologiques en Nouvelle-Calédonie, période 1995-2011 179
Figure 4.6 : Analyse EOF des températures moyennes sur 24 stations météorologiques en Nouvelle-Calédonie, période 1995-2011 180
Figure 4.7 : general map of New-Caledonia 232
Figure 4.8: maps of the 5 explanatory variables selected for modeling dengue incidence rates 233
Figure 4.9: Spatial heterogeneity of dengue annual incidence rates in New Caledonia 234
Figure 4.10: Principal component analysis over the set of climatic variables and socio-economic variables.....235
Figure 4.11: Results of the best multi-variate model of dengue incidence rates 236
Figure 4.12: Maps of observed and predicted annual incidence rates.....237
Figure 4.13: trends of dengue spatial distribution under global warming.....238
Supplementary Figure 4.14: Mean dengue annual incidence rates (in number of declared cases per 10,000 people per year) during epidemic years between 1995 and 2012 according to age classes.....239
Supplementary Figure 4.15: Increase of mean temperature in Noumea as projected by six coupled atmosphere-ocean models at the 2100 horizon.....239
Supplementary Figure 4.16: Comparison of six coupled ocean-atmosphere model projections 240
Supplementary Figure 4.17: Relationship between dengue annual incidence rates and the 5 selected explanatory variables 241
Supplementary Figure 4.18: Spatial influence between the selected explanatory variables and dengue incidence rates 242
Figure 5.1 : Dynamique de la dengue à Nouméa 249
Figure 5.2 : Population (A), densité de population (B) et nom des quartiers (C) du Grand Nouméa 251
Figure 5.3 : Structure spatiale des épidémies de dengue à Nouméa, 1995-2012 252
Figure 5.4 : Structure spatiale des anomalies des taux d'incidence annuels de la dengue à Nouméa au cours des années épidémiques de la période 1995-2012 254
Figure 5.5 : Analyse EOF des taux d'incidence annuels de la dengue à Nouméa, période 1995-2012 255
Figure 5.6 : Indices entomologiques à Nouméa : comparaison des années épidémiques et non épidémiques (2000-2011).....256
Figure 5.7 : Description des étapes de la construction du modèle entomologique du risque épidémique à Nouméa.....258
Figure 5.8 : Résultats du modèle entomologique du risque d'émergence épidémique à Nouméa 262
Figure 5.9 : Seuils entomologiques nécessaires au développement d'une épidémie à Nouméa en fonction de la température 263
Figure 5.10 : Relation entre le R0 calculé et les variables d'entrée du modèle 264
Figure 5.11 : Influence de la température sur les cycles de vie d'Aedes aegypti. 265
Figure 5.12 : Analyse de sensibilité du modèle aux différentes valeurs des paramètres.....266
Figure 5.13 : Lien entre température, mortalité et âge des moustiques Aedes aegypti adultes.....270
Figure 5.14 : Taux d'incidence annuels de la dengue en 2009 à Nouméa par îlot 278
Figure 5.15 : Semi-variogramme des taux d'incidence annuels par îlot en 2009.....279
Figure 5.16 : Analyse en composantes principales des variables explicatives de la dengue à l'échelle de Nouméa.....282
Figure 5.17 : Semi-variogrammes des résidus de trois modèles SVM bivariés prédisant les taux d'incidence par îlot à Nouméa en 2009 285

LISTE DES TABLEAUX

Table 2.1 : Tests de laboratoire utilisés pour le diagnostic de la dengue au cours du temps 81
Table 2.2: Incidence annuelle de la dengue dans vingt Pays et Territoires Insulaires du Pacifique..... 87
Table 2.3: Données manquantes et statistiques de chacun des champs de la base finale 89
Table 2.4: Pourcentage mensuel de cas de dengue confirmés parmi les cas enregistrés en Nouvelle-Calédonie..... **Erreur ! Signet non défini.**
Table 2.5 : résultats du geo-référencement des données épidémiologiques dans les communes du Grand-Nouméa..... 92
Table 2.6 : Indices entomologiques collectés par l'IPNC et le SIPRES (mars 2000 – septembre 2011) 96
Table 2.7 : Données manquantes des variables météorologiques..... 98
Table 2.8 : Indicateurs de télédétection pour l'étude spatiale à Nouméa en 2009.....109
Table 2.9 : Indicateurs géographiques pour l'étude spatiale à Nouméa en 2009.....109
Table 2.10 : Champ spatio-temporel de taux d'incidence X(t,s), avec n = 6 et p = 2 130
Table 3.1 : Comparaison de la précocité d'apparition d'un nouveau sérotype dans les vingt PTIP 155
Table 4.1 : Chronologie des épidémies annuelles de dengue dans les communes de Nouvelle-Calédonie.....174
Table 4.2 : Correlation between dengue incidence rates and socio-economic or climate variables.....243
Table 4.3 : Bivariate and multivariate modelling of dengue incidence rates: variable selection according to the RMSE of the SVM models.....246
Table 4.4 : Projections of temperature increase and predicted incidence rates for three time periods in the future.....246
Table 5.1 : Paramètres et variables du modèle 259
Table 5.2 : Etapes du calcul du nombre de moustiques infectieux produits par jour et par hôte infecté (noté A) 260
Table 5.3 : Etapes du calcul du nombre de personnes infectées « produites » par moustique infectieux qui s'est infecté au cycle n (noté B) 260
Table 5.4 : Résultats de l'enquête spatiale des facteurs de risque sociologiques de la dengue à l'échelle de Nouméa.....273
Table 5.5 : Coefficients de corrélation (Rho) entre les variables explicatives et les taux d'incidence de la dengue à l'échelle de Nouméa 280
Table 5.6 : Modélisation bivariée et multi-variée des taux d'incidence de la dengue : sélection des variables par comparaison de l'erreur de prédiction des modèles SVM 284
Table 6.1 : Synthèse des résultats par échelle spatio-temporelle et par nature des études abordées.....295

LISTE DES ABREVIATIONS

AOGCM	Atmosphere/Ocean Global Circulation Model
AR	Assessment Report
AUC	Area Under the (ROC) Curve
BI	Brigthness Index
CHT	Centre Hospitalier Territorial
CMIP	Coupled Model Intercomparison Project
CPS	Secrétariat général de la Communauté du Pacifique
DASS-NC	Direction des Affaires Sanitaires et Sociales - Nouvelle-Calédonie
DENV-x	Virus de la dengue, sérotype x
DITTT	Direction des Infrastructures, de la Topographie et des Transports Terrestres
ELISA	Enzyme-Linked ImmunoSorbent Assay
ENSO	El Niño Southern Oscillation
EOF	Empirical Orthogonal Function
FF	Force des vents
FSM	Federated States of Micronesia
GIEC	Groupe d'experts Intergouvernemental sur l'Evolution du Climat
GLM	Generalised Linear Model
IFI	ImmunoFluorescence Indirecte
IH	Inhibition de l'hémagglutination
IPCC	Intergovernmental Panel on Climate Change
IPNC	Institut Pasteur de Nouvelle-Calédonie
ISEE	Institut de la Statistique et des Etudes Economiques
LAV	Lutte Anti-Vectorielle
NC	Nouvelle-Calédonie
NDVI	Normalised Difference Vegetation Index
NDVI(RE)	Normalised Difference Vegetation Index (Red edge)
NDWI	Normalised Difference Water Index
OMS	Organisation Mondiale de la santé
PNG	Papouasie Nouvelle-Guinée
PTIP	Pays et Territoires Insulaires du Pacifique
PVI	Perpendicular Vegetation Index
ROC (courbe)	Receiver Operating Characteristic (curve)
ROSSP	Réseau Océanien de Surveillance de la Santé Publique
RR	Précipitations
RT-PRC	Reverse-Transcriptase Polymerase Chain Reaction
SIPRES	Service d'Inspection et de Prévention des Risques Environnementaux et Sanitaires
SPCZ	South-Pacific Convergence Zone
SVM	Support Vector Machine
TN	Température miNimale
TNDVI	Transformed Normalised Difference Vegetation Index
TX	Température maXimale
UN	Humidité relative miNimale
URE-EnM	Unité de Recherche et d'Expertise en Entomologie Médicale
UX	Humidité relative maXimale
WRF	Weather Reasearch and Forecast Model
ZEE	Zone économique exclusive

LISTE DES PUBLICATIONS ASSOCIEES A LA THESE

Publications soumises ou acceptées

- Article 3 : **Teurlai M**, Mangeas M, Menkes C, Lefevre J, Cavarero V, Pfannstiel A, Grangeon JP, Guillaumot L, Gourinat AC. Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: a Worked Example in New Caledonia (En révision à PLoS NTD).

- Article 2 : **Teurlai M**, Huy R, Cazelles B, Duboz R, Baehr C, Vong S. (2012) Can human movements explain heterogeneous propagation of dengue fever in Cambodia ? *PLoS Negl Trop Dis*. doi: 10.1371/journal.pntd.0001957

- Article 1 : Descloux E, Mangeas M, Menkes C, Lengaine M, Leroy A, Tehei T, Guillaumot L, **Teurlai M**, Gourinat AC, Benzler J, Pfannstiel A, Grangeon JP, Degallier N, De Lamballerie X. (2012) Climate-based models for understanding and forecasting dengue epidemics. *PLoS Negl Trop Dis*. doi: 10.1371/journal.pntd.0001470

Publications en cours de préparation

- **Teurlai** *et al.* On the use of Empirical Orthogonal Functions to explore the spatio–temporal distribution of infectious diseases : application to dengue in New Caledonia.

- **Teurlai** et al. Multi-factorial spatial modeling of dengue epidemics at a city scale: the case of Noumea in 2009, New-Caledonia

- Mercier *et al.* The insular dynamics of Dengue fever in the South Pacific over the 40 last years (1971-2011).

- Mangeas *et al.* Towards a new method for modeling and forecasting intra-annual variability of dengue outbreak risk: application within a New-Caledonian Context (The Lancet Infectious Diseases)

Communications orales et poster

- **Teurlai M**, Mangeas M, Menkes C. Analyse multi-factorielle des épidémies de dengue dans les communes de Nouvelle-Calédonie : impact du climat et des facteurs socio-environnementaux. Journée médicale Calédonienne – 3 octobre 2014 – Nouméa – Nouvelle-Calédonie.

- **Teurlai M**, Mangeas M, Menkes C. Study of the spatial distribution of dengue cases in New Caledonia: the impact of present and future climate. Journée scientifique de l'Institut Pasteur de Nouvelle-Calédonie – Institut Pasteur de Nouvelle-Calédonie – 21 novembre 2013 – Nouméa – Nouvelle-Calédonie.

- **Teurlai M**, Mangeas M, Menkes C. The role of climate in the epidemiology of dengue fever: understanding and predicting the epidemiologic risk. IRD-USP collaborative workshop on climate Change – University of the South Pacific 26-28 august 2013 – Suva – Fiji.

- **Teurlai M**, Mangeas M, Menkès C. Recherches sur la dengue en Nouvelle-Calédonie. Séminaire Province Nord de Nouvelle-Calédonie; Quelles recherches scientifiques en Province Nord ? – 19-21 juin 2013 – Poindimié – Nouvelle-Calédonie.

- **Teurlai M**, (avec Descloux E, Mangeas M, Menkes C, Lengaine M, Leroy A, Tehei T, Guillaumot L, Gourinat AC, Benzler J, Pfannstiel A, Grangeon JP, Degallier N, De Lamballerie X). Climate-based models for understanding and forecasting dengue epidemics and perspectives for the South Pacific area. University of Fiji – 13-15 September 2012 – Nandi – Fiji.

- **Teurlai M** (with Huy R, Cazelles B, Duboz R, Baehr C, Vong S). Human movements can explain heterogeneous propagation of dengue fever in Cambodia. Poster. Young Researchers Scientific Meeting. Pasteur Institute International Network – November 8-10, 2011 – Paris - France.

Histoire de la lutte contre les maladies infectieuses

Les maladies infectieuses ont été pendant longtemps, et sont encore, une cause importante de mortalité des hommes. Certaines épidémies ont eu des effets dévastateurs sur les populations humaines et ont si profondément marqué les esprits qu'il en reste encore des traces dans notre langage courant. Les expressions « fuir comme la peste » ou « avoir une peur bleue » sont par exemple apparues suite aux vastes épidémies de peste et de choléra ayant frappé l'Europe respectivement au Moyen-Age et au 19ème siècle. De tous temps, les hommes ont cherché à comprendre l'origine des maladies infectieuses ainsi que leur mode de transmission afin de pouvoir mieux s'en prémunir. Les travaux de Pasteur puis ceux de Koch au 19ème siècle ont marqué un réel tournant dans la lutte contre ces maladies. Le développement de la théorie de l'origine microbienne des maladies infectieuses a permis l'identification de nombreux agents pathogènes, une meilleure compréhension de leurs mécanismes de transmission ainsi que le développement de vaccins efficaces. La mise en place de mesures de lutte intégrant les connaissances sur l'épidémiologie des maladies, l'utilisation massive de la vaccination, puis, dans la deuxième moitié du 20ème siècle celle des antibiotiques, ont permis de réduire de manière significative le nombre de morts liés aux maladies infectieuses, du moins dans les pays développés. Portée par l'efficacité croissante des moyens thérapeutiques préventifs ou curatifs, la compréhension de plus en plus fine des processus de transmission et les succès subséquents de la lutte contre les maladies infectieuses, l'Organisation Mondiale de la Santé (OMS) lance en 1966 une grande campagne mondiale de lutte contre la variole et annonce son éradication officielle de la surface du globe en 1980 (WHO, 1980). Les esprits sont à l'optimisme : il est possible d'éradiquer de manière définitive les maladies infectieuses. C'était sans compter sur les processus évolutifs, qui font émerger sans cesse de nouvelles formes de vie, et permettent aux agents infectieux existant d'échapper aux défenses que leur imposent les hôtes qu'ils parasitent. A la fin des années 1970, un syndrome caractérisé par la déficience des cellules du système immunitaire et la mort de tous les patients fait son apparition et se propage à l'ensemble du globe : le SIDA (Syndrome d'Immuno-Déficience Acquise). De nombreuses bactéries deviennent résistantes aux antibiotiques, des agents pathogènes inconnus émergent partout dans le monde (coronavirus du Syndrome Respiratoire Aigu Sévère), d'autres, confinés jusqu'alors aux animaux, passent la barrière d'espèce et provoquent des épidémies chez l'homme (virus de la grippe H5N1, virus Ebola). Les espoirs retombent. L'émergence permanente de nouveaux agents infectieux exclut la possibilité d'une élimination totale.

Les agents pathogènes émergents

Une étude récente a recensé 1407 agents pathogènes pour l'homme dont 177 sont considérés comme émergents ou ré-émergents (Woolhouse and Gowtage-Sequeria, 2005). Le concept d'émergence est un

concept apparu récemment, et il n'existe pas encore de définition universellement admise. En France, la Cellule permanente des maladies infectieuses émergentes a défini une maladie infectieuse émergente comme « un phénomène infectieux – ou présumé infectieux – inattendu (en référence à ses propriétés intrinsèques ou à la connaissance de sa biologie), touchant l'homme, l'animal ou les deux ». Il peut s'agir :

- D'une entité clinique d'origine infectieuse nouvellement apparue ou identifiée ;
- D'une maladie infectieuse connue, mais dont l'incidence augmente ou dont les caractéristiques (cliniques, évolutives) se modifient dans un espace ou dans un groupe de population donné (Leport et al., 2011).

Les maladies émergentes sont caractérisées par l'incertitude quant à leur impact sanitaire ou économique futur. C'est cette incertitude, associée aux conséquences parfois dramatiques des émergences passées (comme la grippe espagnole de 1918, la grippe asiatique de 1957 ou le SIDA dans les années 1980) qui conduisent les autorités sanitaires à leur accorder une grande importance, malgré une incidence parfois faible. Un exemple d'actualité frappant est l'émergence du virus Ebola au Libéria en mars 2014. C'est également cette incertitude et la difficulté de prévoir l'évolution future des maladies émergentes qui pousse les autorités sanitaires à privilégier la mise en place de mesures préventives et de campagnes de communication massives qui peuvent parfois sembler excessives *a posteriori*, comme ce fût le cas pour la grippe H1N1 en 2009.

Parmi les agents pathogènes ayant récemment émergé, les virus, et particulièrement les virus à ARN, ont une place prépondérante par rapport aux bactéries, protozoaires, helminthes, ou infections fongiques. En effet, sur les 208 virus connus, 77 (37 %) ont émergé dans les dernières décennies, contre 54 des 538 espèces connues de bactéries (10 %), 22 des 317 espèces de champignons pathogènes (7 %), 14 des 57 espèces de protozoaires (25 %) et 10 des 287 espèces d'helminthes (3 %) (Woolhouse and Gowtage-Sequeria, 2005). Une autre étude sur les agents pathogènes émergents a recensé 335 « évènements infectieux émergents » entre 1940 et 2004, parmi lesquels 25 % seulement sont dus à des virus (Jones et al., 2008). Cependant cette étude ne donne pas la proportion des différents pathogènes responsables de l'ensemble des maladies, et ne permet donc pas de calculer un risque relatif d'émergence entre les différentes espèces de pathogènes.

Sur les 20 familles de virus comprenant des espèces pathogènes pour l'homme, seules 4 familles regroupent plus de la moitié des virus pathogènes recensés ou émergents (Woolhouse and Gowtage-Sequeria, 2005): les *Bunyaviridae* (famille qui comprend le virus Hantaa ainsi que celui de la fièvre de la vallée du Rift), les *Flaviviridae* (virus de la dengue, de la fièvre jaune, de la fièvre Zika ou de l'encéphalite japonaise), les *Togaviridae* (virus responsable du Chikungunya, de la fièvre du Nil Occidental – ou West Nile – ou de la maladie de Ross River) et les *Reoviridae* (virus de la bluetongue chez les animaux, diarrhées à rotavirus chez les hommes).

Une grande partie (26/77 soit près d'un tiers, Liam Brierley, communication personnelle) des virus ayant émergé récemment sont des arbovirus (Jones et al., 2008), contraction de l'anglais « **ar**thropod-**bo**rne **virus** », c'est-à-dire des « virus entretenus dans la nature par transmission biologique de vertébré à vertébré par l'intermédiaire d'arthropodes hématophages » (Toma et al., 2004), le plus souvent un moustique. L'installation récente et pérenne du moustique *Aedes albopictus,* vecteur de nombreuses maladies, dans le Sud de l'Europe, y compris la France, fait craindre une transmission accrue des arboviroses dans le futur (Talbalaghi et al., 2010; Vazeille et al., 2010). La forte proportion d'arbovirus parmi les pathogènes émergents (46/177, Liam Brierley, communication personnelle) ainsi que la sensibilité de leur cycle de transmission aux conditions climatiques (Focks et al., 1993a) font également craindre une augmentation de leur circulation dans le futur du fait du changement climatique annoncé (Colwell et al., 1998; Githeko et al., 2000; Gubler et al., 2001; Patz et al., 1998; Reiter, 2001; Smith et al., 2014b).

La dengue

La dengue est actuellement l'arbovirose la plus importante au monde, touchant chaque année plus de 390 millions de personnes, dont près de 100 millions d'infections se manifestant de manière apparente (Bhatt et al., 2013). Elle est la première cause de décès chez les enfants en Asie, et pose problème par le nombre d'hospitalisations nécessaires à la bonne prise en charge des patients (WHO, 2009). L'OMS estime à 2,5 milliards le nombre de personnes exposées au risque de contracter la dengue dans le monde actuellement (WHO, 2009).

La dengue est transmise par les moustiques du genre *Aedes* et provoque un syndrome fébrile le plus souvent bénin, mais pouvant se compliquer d'hémorragies et aboutir au décès des personnes affectées. La dengue est apparue voilà près de 1000 ans (Gubler, 1998; Holmes and Twiddy, 2003). Elle est cependant considérée comme émergente à cause de l'augmentation drastique de son incidence et de sa forte expansion géographique depuis les années 1970, ainsi qu'à cause de l'apparition récente de formes sévères de dengue hémorragique (Gubler, 2004) (voir Encadré 1 : Histoire naturelle de la dengue dans le monde).

L'impact de la dengue dépend en grande partie du niveau de développement des infrastructures de santé des pays dans lesquels elle sévit et de la capacité de ces infrastructures à prendre en charge les formes sévères. Dans les pays en développement où la proportion de formes hémorragiques est importante, l'impact sanitaire est considérable. Dans le Pacifique, où la dengue sévit sous la forme de vagues épidémiques et où la proportion des formes hémorragiques est moins élevée qu'en Asie (Kiedrzynski et al., 1998; Singh et al., 2005), l'impact diffère selon les pays.

En Nouvelle-Calédonie, le taux de morbidité (voir glossaire en Annexe 1, page 337), est élevé au cours des années épidémiques et la dengue représente un enjeu de santé publique majeur (DASS-NC, 2010). Cependant le taux de létalité est beaucoup plus faible (0,3 % en 2003 (DASS-NC, 2010)) que dans

d'autres pays grâce à l'existence d'infrastructures de santé solides. La dengue a en revanche un impact économique considérable, du fait de la gratuité des soins pour les patients, et de la prise en charge des arrêts maladie par les systèmes d'assurance sociale. De plus, l'évolution future de la dengue en Nouvelle-Calédonie inquiète : en effet, les épidémies semblent de plus en plus fréquentes, et d'une ampleur croissante, faisant craindre la saturation de la capacité des hôpitaux dans un futur proche.

Encadré 1 : Histoire naturelle de la dengue dans le monde

L'origine du virus de la dengue n'est pas connue de manière certaine, cependant des études phylogénétiques laissent penser qu'il serait apparu il y a environ mille ans (Twiddy, Holmes, et Rambaut 2003) au sein d'un cycle sylvatique singe-moustique-singe, soit en Asie, soit en Afrique (Duane J. Gubler 1998; Holmes et Twiddy 2003). Le virus se serait alors différencié en 4 sérotypes, et chacun des sérotypes aurait traversé la barrière d'espèce entre singes et humains il y a entre ~320 (DENV-2) et ~125 ans (DENV-1) (Holmes et Twiddy 2003). Des descriptions cliniques d'une maladie ressemblant à la dengue datent de l'an 992 en Chine. Des épidémies apparues en 1635 aux Antilles Françaises et en 1699 au Panama sont également potentiellement attribuées à la dengue (Duane J. Gubler 1998). La première épidémie documentée a eu lieu à Philadelphie, en 1779 (voir figure 1, (Duane J. Gubler 1998)). C'est sûrement au cours du 18ème et du 19ème siècle qu'est apparu un cycle urbain de transmission impliquant le moustique *Aedes aegypti* comme espèce vectrice, et l'on pense que la dengue est alors devenue endémique dans de nombreuses villes tropicales. Au début du 20ème siècle, c'est une maladie majeure. Elle provoque des pandémies régulières dans les Caraïbes. Elle est également présente aux Etats-Unis dans les années 30, mais la lutte intensive contre la fièvre jaune à cette époque provoque sa disparition. La circulation sur le continent américain décroit ensuite jusque dans les années 1970 (WHO 2009). Entre 1900 et 1950, les 4 sérotypes du virus sont également très présents en Asie du Sud-Est, où de vastes épidémies se développent régulièrement.

Durant la seconde guerre mondiale, la circulation virale s'intensifie en Asie du Sud-Est, probablement à cause de mouvements importants de troupes militaires permettant un renouvellement rapide de la population d'hôtes susceptibles. Ainsi, dans les années 50 la dengue devient hyper-endémique, c'est-à-dire avec co-circulation simultanée des 4 sérotypes (Gubler, 2004). Cette hyper-endémicité est supposée être à l'origine de l'apparition des premières formes sévères de dengue. L'année 1953 marque un tournant dans l'histoire de la dengue avec l'apparition de la première épidémie de dengue hémorragique à Manille, aux Philippines. La dengue hémorragique se développe ensuite à Bangkok en 1958, à Singapour, au Vietnam et en Malaisie au début des années 60.

Alors qu'en 1970 la dengue hémorragique n'est présente qu'en Asie, l'accroissement de la population et l'urbanisation intensive qui l'accompagne, l'augmentation des échanges internationaux, l'absence d'un contrôle effectif des espèces de moustique vectrices ainsi que la fragilité des infrastructures de santé publique de certains pays conduisent à la dissémination rapide de la maladie à travers le monde (Gubler,

1998). Après plusieurs décennies d'absence, les 4 sérotypes de la dengue sont ainsi réintroduits en Amérique du Sud en 1977 (DENV-1), 1981 (DENV-2 et DENV-4) et 1994 (DENV-3), où ils causent des vagues épidémiques régulières depuis. La dengue fait également sa ré-apparition dans le Pacifique en 1971 (DENV-1) après près de 30 ans d'absence, la dernière épidémie rapportée dans cette région du monde datant de 1941. Cette expansion géographique sans précédent a conduit à qualifier la dengue de maladie ré-émergente.

Elle est devenue, depuis 4 décennies, l'arbovirose la plus importante au monde, affectant chaque année plus de 390 millions de personnes réparties dans plus 100 pays, et causant plusieurs centaines de milliers de décès (Bhatt et al., 2013; Gubler, 1998).

En l'absence de vaccin ou de traitement spécifique, la lutte contre la dengue dans le monde repose essentiellement sur la mise en place d'une surveillance épidémiologique afin d'identifier les zones et les périodes à risque, la sensibilisation des populations et le contrôle du moustique vecteur. Dans ce contexte, il est fondamental de mieux comprendre les facteurs et mécanismes impliqués dans le processus de transmission virale. La compréhension des mécanismes menant à la distribution spatio-temporelle observée de la dengue doit permettre :

- A l'échelle de la planète, de prévoir au mieux l'évolution future de la maladie, en particulier dans le contexte d'expansion géographique rapide, et de la possible augmentation future de l'aire de répartition des moustiques vecteurs sous l'hypothèse d'un réchauffement global ;
- A l'échelle d'un pays ou d'une ville: de fournir une information pertinente aux autorités sanitaires quant aux zones à cibler en priorité et aux facteurs de risque, afin de rationaliser les choix en matière de lutte ;
- A l'échelle d'un pays ou d'une ville : de fournir éventuellement un système d'alerte précoce, permettant d'anticiper le risque et d'améliorer la prise en charge des épidémies.

Modélisation de la dynamique spatiale de la dengue

Comme toute maladie vectorielle, la dynamique spatiale de la dengue (*i.e.* la distribution spatio-temporelle des cas de dengue) résulte d'interactions complexes entre l'agent pathogène, l'hôte et le moustique vecteur, interactions qui peuvent être modifiées par l'environnement dans lequel ces trois composantes évoluent (Figure 1). Nous entendons ici le terme « environnement » au sens large, c'est-à-dire englobant l'ensemble des facteurs écologiques, climatiques, mais également économiques, sociologiques ou comportementaux pouvant influencer le cycle de transmission de la dengue. Nous entendons par « dynamique » l'évolution du nombre de cas dans le temps, et par « dynamique spatiale » l'évolution du nombre de cas dans le temps et dans l'espace.

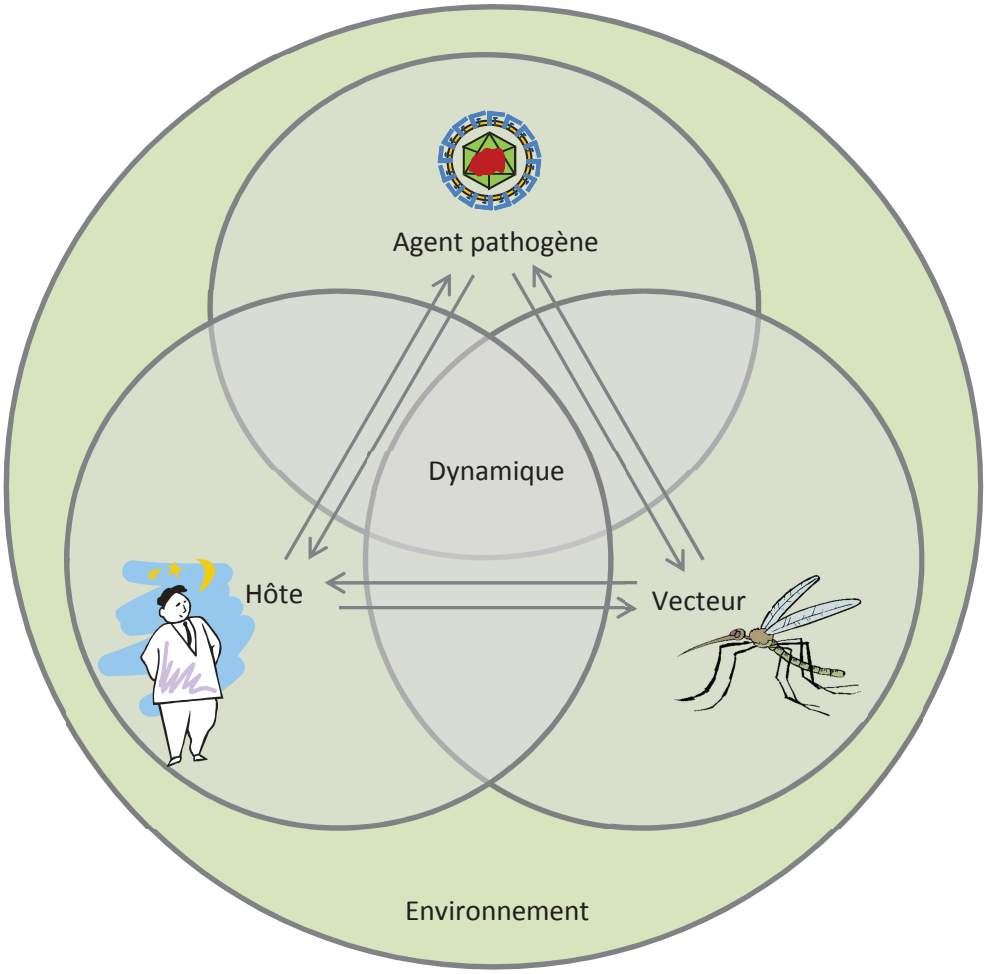


Figure 1 : Schéma des interactions (symbolisées par des flèches) entre l'hôte susceptible, le vecteur et l'agent pathogène menant à la dynamique observée des maladies vectorielles. *Hôtes, vecteurs et agents pathogènes évoluent dans un environnement commun pouvant influencer ces interactions.*

La nature multi-factorielle des déterminants de la dynamique spatiale de la dengue et les nombreuses interactions entre les différents facteurs rendent complexes la compréhension des mécanismes impliqués dans le cycle de transmission et altèrent le caractère prévisible des épidémies. Malgré les nombreuses études réalisées sur le sujet, études qui seront abordées dans le premier chapitre de cette thèse, certaines incohérences dans les résultats limitent toujours notre compréhension des déterminants de la dynamique spatiale de la dengue.

Objectif global de la thèse

L'objectif premier de cette thèse, outre les potentielles retombées opérationnelles, est de contribuer à éclaircir les mécanismes impliqués dans le cycle de transmission de la dengue, en étudiant la distribution spatio-temporelle des cas de dengue et les facteurs associés à la distribution observée dans une région du

monde où les études étaient peu nombreuses au début de cette thèse : le Pacifique, et plus particulièrement la Nouvelle-Calédonie.

Données épidémiologiques, hypothèses de travail, méthodes et objectifs spécifiques

La grande quantité d'information épidémiologique relativement fiable dans le Pacifique et la qualité de cette information en Nouvelle-Calédonie nous donne l'opportunité assez unique de pouvoir étudier ces mécanismes à différentes échelles spatiales et temporelles. A l'exception d'une courte incursion dans le domaine des modèles déterministes, nous nous sommes limités dans le cadre de cette thèse à une approche statistique.

Dans le Pacifique, depuis la réapparition de la dengue en 1971, afin de suivre l'évolution de la maladie et d'orienter les mesures de contrôle, les pays ont mis en place des réseaux de surveillance épidémiologique permettant la création de bases de données possédant une profondeur temporelle importante (Kiedrzynski et al., 1998). Dans la cadre de cette thèse, nous avons eu accès à trois bases de données épidémiologiques : une base régionale recensant le nombre de cas annuels dans chacun des pays du Pacifique, une base territoriale recensant le nombre de cas mensuels enregistrés en Nouvelle-Calédonie depuis 1971, et une base de données rassemblant des informations individuelles anonymes (dont la localisation du cas et la date) sur chacun des cas de dengue déclarés aux autorités sanitaires de Nouvelle-Calédonie depuis 1995, et pouvant être agrégés dans le temps ou dans l'espace à différentes échelles et résolutions.

Un des premiers objectifs de la thèse était d'identifier d'éventuels schémas spatio-temporels récurrents, et d'identifier les facteurs environnementaux déterminant la distribution spatio-temporelle des cas à chacune des échelles. Nous avons étudié trois échelles spatiales : l'échelle du Pacifique, l'échelle du territoire de Nouvelle-Calédonie et l'échelle locale de Nouméa, capitale de la Nouvelle-Calédonie. A cette fin, pour chacune des trois échelles, nous avons réalisé une analyse approfondie de la variabilité spatio-temporelle des données épidémiologiques. Pour les deux dernières échelles, nous avons récolté des données environnementales sur les facteurs pouvant potentiellement influencer le cycle de transmission de la dengue et construit des modèles statistiques permettant d'expliquer ou de prédire le risque épidémiologique à partir de ces facteurs. L'identification de facteurs communs à toutes les échelles conduirait à l'hypothèse forte de l'existence d'un mécanisme commun sous-jacent à tout risque épidémiologique et contribuerait à améliorer notre capacité à anticiper ces risques quelle que soit l'échelle ou la zone d'étude.

Un deuxième objectif de cette thèse était de proposer une méthode générique de l'approche de ce type de problème en épidémiologie, et d'identifier, pour chacune des échelles étudiées, les outils les mieux adaptés pour répondre à un objectif donné.

Enfin, cette thèse nous a donné l'opportunité assez unique d'étudier une même maladie à différentes échelles spatiales. C'était donc l'occasion de comprendre comment les différents processus épidémiologiques intervenant à différentes échelles interfèrent les uns avec les autres. En particulier, l'objectif final de cette thèse était de développer un indicateur de risque épidémiologique pour la Nouvelle-Calédonie intégrant les informations acquises à chacune des échelles.

Plan de la thèse

Dans une première partie, nous exposerons les éléments de contexte indispensables à la bonne compréhension et à l'interprétation des études réalisées. Le premier chapitre fait l'état des connaissances sur la dengue, ainsi qu'un état de l'art synthétique sur les études déjà réalisées quant aux facteurs de risque de la maladie. Dans le deuxième chapitre, nous décrirons l'ensemble des données analysées au cours de cette thèse, les éléments nécessaires à l'évaluation de leur fiabilité ainsi que les outils statistiques utilisés au cours du travail. Le regroupement de l'ensemble des données dans un même chapitre permet de souligner le travail important de nettoyage des données qui a dû être réalisé, et permet d'épurer les chapitres consacrés à l'analyse des données, dans une volonté de clarifier les résultats obtenus. En ce qui concerne les outils statistiques, nous passerons rapidement en revue les différents outils à disposition, et détaillerons plus particulièrement trois outils innovants utilisés régulièrement au cours de la thèse : i) l'analyse en ondelette (Torrence and Compo, 1998) pour l'étude des processus de propagation spatio-temporels ; ii) un outil d'analyse de la variabilité temporelle des schémas spatiaux issu du domaine de la météorologie et très rarement utilisé en épidémiologie, la « décomposition orthogonale aux valeurs propres » ou « Empirical Orthogonal Function » (EOF) en anglais (Hannachi et al., 2007), et iii) un outil d'apprentissage automatique récent permettant l'identification de liens non-linéaires entre facteurs environnementaux et risque épidémiologique, les machines à vecteur de support (ou « Support Vector Machine » (SVM) en anglais) (Vapnik, 1998). L'utilisation de l'analyse en ondelette et des SVM pour les études épidémiologiques était déjà bien développée au début de cette thèse. Ces deux techniques ne sont donc exposées ici que dans un souci de compréhension de l'exposé par le lecteur. En revanche, l'utilisation des EOF a nécessité d'adapter l'outil au cadre épidémiologique, en particulier en ce qui concerne l'interprétation des résultats.

Dans une deuxième partie nous exposerons le travail d'analyse et de modélisation réalisé à trois échelles spatiales différentes. Chaque chapitre correspond à une des trois échelles spatiales d'étude : l'échelle régionale du Pacifique (chapitre 3), l'échelle territoriale de Nouvelle-Calédonie (chapitre 4) et l'échelle locale de Nouméa (chapitre 5). Chacun de ces chapitres est organisé selon l'approche méthodologique utilisée, à savoir une première partie exploratoire de la dynamique spatio-temporelle des données épidémiologiques, puis une analyse des facteurs de risque déterminant la dynamique des cas de dengue, et pour terminer une analyse des facteurs de risque déterminant la distribution spatiale des cas. Enfin, dans un dernier chapitre (chapitre 6), nous discutons de l'apport d'une étude multi-échelle pour éclairer les connaissances sur l'épidémiologie de la dengue. Les facteurs identifiés dans le risque de transmission

de la dengue diffèrent selon que l'on étudie le risque d'émergence d'une épidémie ou sa sévérité, selon qu'on étudie la dynamique de la maladie ou sa distribution spatiale, et enfin, pour un même risque, les facteurs déterminant du risque différent en fonction de l'échelle d'étude. Dans ce dernier chapitre, nous dégageons les implications des résultats en termes opérationnels et ébauchons des pistes de recherche future.

Pour conclure, nous ferons une synthèse des résultats et des apports de la thèse.

Outre les résultats qu'elle a permis de mettre en évidence, la méthodologie développée dans le cadre de cette thèse dépasse le cadre d'étude de la dengue et peut servir à quiconque voudrait aborder un problème d'analyse des données spatio-temporelles dans d'autres contextes épidémiologiques.

Les lecteurs trouveront un glossaire des termes épidémiologiques, entomologiques et statistiques utilisés en Annexe 1, page 337.

PARTIE I : CONTEXTE ET DONNEES

CHAPITRE 1 : REVUE BIBLIOGRAPHIQUE SUR LA DENGUE ET LES FACTEURS INFLUENÇANT SON EPIDEMIOLOGIE

1.1. La dengue 39

1.1.1. L'agent pathogène 39
Structure de l'agent pathogène 39
Diversité génétique, pouvoir antigène et pouvoir immunogène 39
1.1.2. Les vecteurs 41
1.1.2.a. Espèces vectrices 41
1.1.2.b. Cycle de vie et facteurs influençant 42
Cycle de vie 42
Particularités des *Aedes* 43
Influence des facteurs environnementaux sur les traits de vie des *Aedes* 44
1.1.3. Formes cliniques et épidémiologie des formes cliniques 46
Formes cliniques 46
Proportion de formes inapparentes 47
Facteurs de risque de la DHF 47
1.1.4. Diagnostic 48
1.1.4.a. Diagnostic expérimental 48
1.1.4.b. Diagnostic différentiel 49
1.1.5. Traitement 49
1.1.6. Prophylaxie médicale et sanitaire 49
Prophylaxie médicale 49
Prophylaxie sanitaire 50
1.1.7. Epidémiologie 50
1.1.7.a. Les hôtes 50
1.1.7.b. Les modes de transmission 50
1.1.7.c. Le cycle de transmission viral 51
1.1.7.d. Facteurs liés à l'hôte influençant le cycle de transmission 51
Statut immunitaire de l'hôte 51
Comportement de l'hôte 51
Phase virémique 52
1.1.7.e. Facteurs liés au vecteur influençant le cycle de transmission 52
Traits de vie du moustique 53
Compétence vectorielle et période d'incubation extrinsèque 53
Capacité vectorielle 54
1.1.7.f. Facteurs liés au virus influençant le cycle de transmission 54
Influence sur la compétence et la capacité vectorielle 54
Influence sur l'expression clinique de la maladie 55

1.2. Modélisation de la dynamique spatiale de la dengue : revue bibliographique 55

1.2.1. Objectifs de la modélisation ---------- 55
1.2.2. Les approches possibles en modélisation ---------- 56
1.2.3. Caractéristiques des études de modélisation ---------- 56
1.2.4. Description de la distribution spatio-temporelle des cas de dengue ---------- 57
1.2.5. Facteurs influençant la distribution spatiale de la dengue ---------- 58
1.2.5.a. Echelle régionale 58
Facteurs influençant la distribution spatiale des vecteurs 58
Facteurs influençant la distribution spatiale de la dengue 59
1.2.5.b. Echelle Nationale 59
1.2.5.c. Echelle locale 60
Facteurs influençant la distribution spatiale du vecteur 60
Lien entre densité vectorielle et distribution spatiale de la dengue 60
Facteurs influençant la distribution spatiale de la dengue 60
1.2.6. Facteurs influençant la distribution temporelle de la dengue ---------- 61
1.2.6.a. Facteurs liés à la variabilité saisonnière de la dengue 61
1.2.6.b. Facteurs liés à la variabilité inter-annuelle de la dengue 61
Facteurs intrinsèques : immunité de population 62
Facteurs intrinsèques : variabilité génétique du virus 62
Facteurs extrinsèques 62
1.2.7. Contraintes des études de modélisation ---------- 63
Contraintes sur les modèles utilisés 63
Contraintes sur les facteurs étudiés 64
Impact sur notre compréhension de la maladie 64

1.3. Conclusion 65

1.1. La dengue

1.1.1. L'agent pathogène

1.1.1.a. Structure de l'agent pathogène

La dengue est due à un arbovirus de la famille des *Flaviviridae,* genre *Flavivirus*. Comme tous les *Flavivirus*, c'est un virus enveloppé à ARN. Il mesure environ 50 nm de diamètre. Son matériel génétique est protégé par une nucléocapside icosaédrique constituée de la protéine C, entourée d'une membrane lipidique dans laquelle sont enchâssées deux protéines (E et prM/M) formant une enveloppe protéique externe (Figure 1.2). Le génome est constitué d'ARN monocaténaire positif d'une longueur d'environ 10 700 nucléotides, codant pour trois protéines structurales (les protéines C, prM/M et E), et pour sept protéines non structurales (NS1, NS2A, NS2B, NS3, NS4A, NS4B, et NS5) (Kuhn et al., 2002). Les protéines structurales E et prM/M sont impliquées dans les mécanismes de fusion cellulaire avec la membrane des cellules de l'hôte (Leitmeyer et al., 1999), tandis que les protéines non structurales sont impliquées dans les mécanismes de transcription/réplication du matériel génétique ainsi que dans les mécanismes d'adaptation aux défenses immunitaires de l'hôte (Muñoz-Jordán et al., 2003).

1.1.1.b. Diversité génétique, pouvoir antigène et pouvoir immunogène

Génétiquement, le virus de la dengue est étroitement lié aux autres *Flavivirus* dont la transmission se fait par l'intermédiaire d'un moustique du genre *Aedes*, dont le tropisme est non-neurotrope (*i.e.* provoquant des syndromes fébriles hémorragiques plutôt que des encéphalites), et qui infectent essentiellement des primates comme le virus Zika, ayant émergé en 2013 en Nouvelle-Calédonie, ou le virus Spondweni (Gaunt et al., 2001).

Sur la base de leurs propriétés antigéniques et de la réponse immunitaire induite, les virus de la dengue sont classés en quatre sérotypes (DENV-1 à DENV-4). L'infection par l'un des sérotypes provoque chez l'hôte une immunité croisée transitoire contre l'ensemble des sérotypes durant probablement entre 2 et 9 mois (Sabin, 1952) et une immunité probablement à vie contre le sérotype en cause. Une première infection avec l'un des sérotypes favoriserait l'apparition de formes cliniques graves lors d'infection secondaire par un autre sérotype par l'intermédiaire d' « anticorps facilitants » (Burke et al., 1988; Halstead, 1970). L'infection par deux des sérotypes protègerait d'une infection par les deux sérotypes restants. En effet, la plupart des personnes adultes habitant en Asie sont immunisées contre les quatre sérotypes, tandis que très peu d'infections tertiaires ou quaternaires apparentes sont déclarées.

Comme tous les virus à ARN, le virus de la dengue est un virus dont l'évolution est rapide par rapport aux virus à ADN du fait d'un taux de réplication élevé, et de l'absence de mécanisme de correction des

erreurs de l'enzyme effectuant la réplication du matériel génétique, l'ARN-polymérase (Holmes and Burch, 2000; Steinhauer and Holland, 1987). Une autre source de variabilité génétique des virus de la dengue est la possibilité de recombinaison entre différents virus présents au sein d'un même organisme, au moins pour le sérotype 1 (Aaskov et al., 2007; Holmes et al., 1999; Worobey et al., 1999).

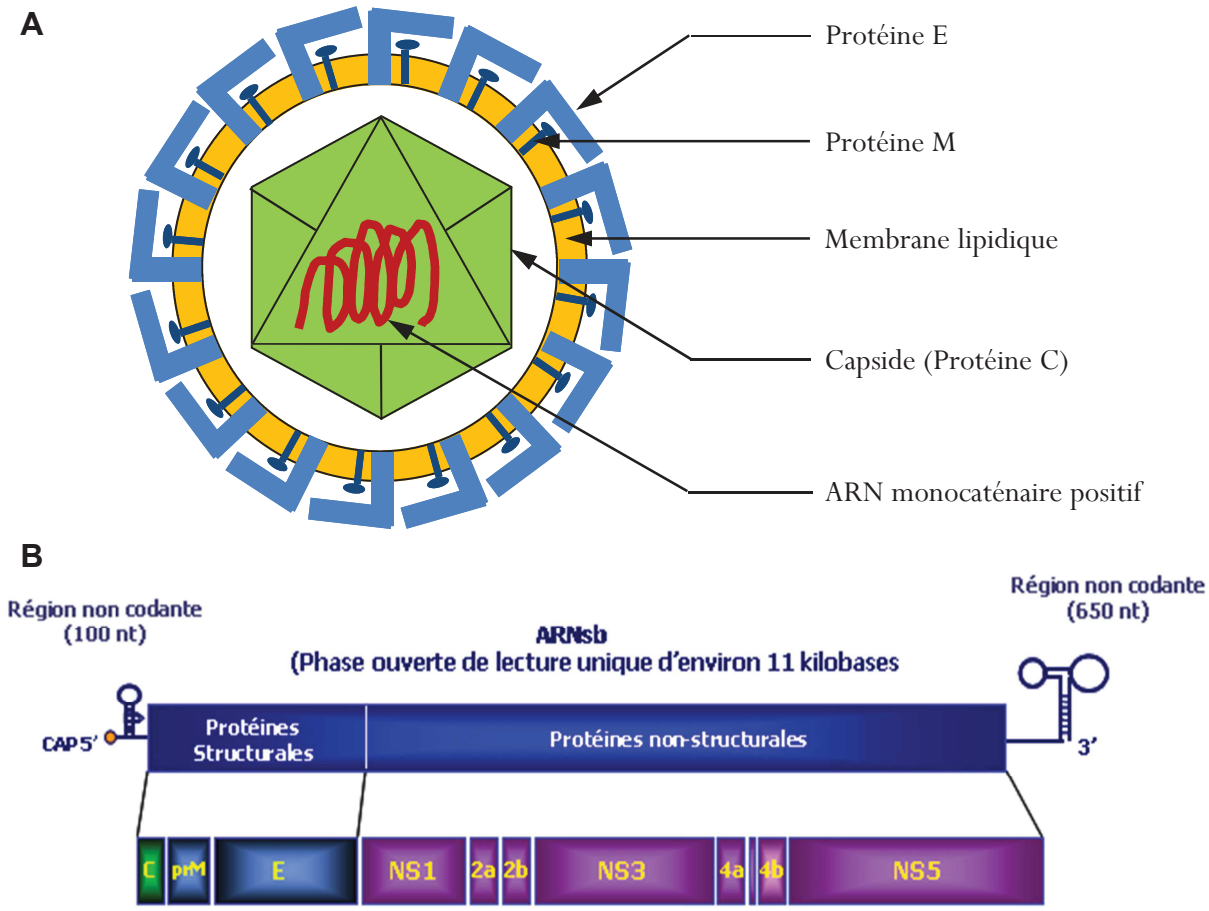


Figure 1.2 : Le virus de la dengue. A. *Structure du virus.* ***B.*** *Structure du génome (source B : Institut Louis Malardé http://www.ilm.pf/Dengueetvecteurs-virus). L'ARN présente une phase ouverte de lecture unique, c'est-à-dire qu'il est transcrit en une seule polyprotéine qui est clivée en différentes protéines par la suite.*

Chaque sérotype comprend différents génotypes, *i.e.* des groupes de virus présentant des similarités génétiques. La variabilité génétique intra-sérotypique des virus de la dengue a été mise en évidence pour la première fois par Rico-Hesse et Trent (Rico-Hesse, 1990; Trent et al., 1990). Rico-Hesse a défini les génotypes de manière arbitraire comme des groupes de virus présentant moins de 6 % de divergence dans leur séquence d'ARN, la portion d'ARN analysée dans cette étude faisant 240 nucléotides de long. Depuis, de nombreuses autres études génétiques, le plus souvent basées sur la séquence complète du gène E, ont été réalisées (Chungue et al., 1993; Holmes et al., 1999; Twiddy et al., 2002). La définition d'un génotype dépend de la portion d'ARN sélectionnée pour effectuer les études phylogénétiques et n'est pas encore bien unifiée (Rico-Hesse, 2003). Le sérotype 2 est celui qui a été le plus étudié, et la plupart des études montrent qu'il existe 6 génotypes, nommés en fonction de leur localisation

géographique : Américain, Américain/Asiatique, Asiatique 1, Asiatique 2, cosmopolite et sylvatique (ou Malaysien) (Holmes and Twiddy, 2003)

Le plus souvent, les virus appartenant à un même génotype sont regroupés dans l'espace ou dans le temps. Les analyses phylogénétiques permettent ainsi de reconstruire de manière assez fine l'histoire naturelle des virus (A-Nuegoonpipat et al., 2004; Cao-Lormeau et al., 2014a; Dupont-Rouzeyrol et al., 2014).

1.1.2. Les vecteurs

1.1.2.a. Espèces vectrices

La plupart des vecteurs de la dengue sont des moustiques appartenant au genre *Aedes*, reconnaissables à l'aspect rayé de leurs pattes. Dans le monde, les deux principales espèces vectrices sont *Aedes albopictus* et *Aedes aegypti*, mais un grand nombre d'autres espèces ont été reconnues comme vectrices de la dengue, incluant *Ae. mediovittatus*, *Ae. africanus*, *Ae. luteocephalus* ou *Ae. niveus*. Les espèces impliquées dans les cycles sylvatiques de transmission singe-moustique-singe sont *Ae. taylori-furcifer*, *Ae. luteocephalus* et *Ae. opok* en Afrique (Holmes and Twiddy, 2003) et des moustiques du genre *Ochlerotatus* en Malaysie (Gubler, 1998).

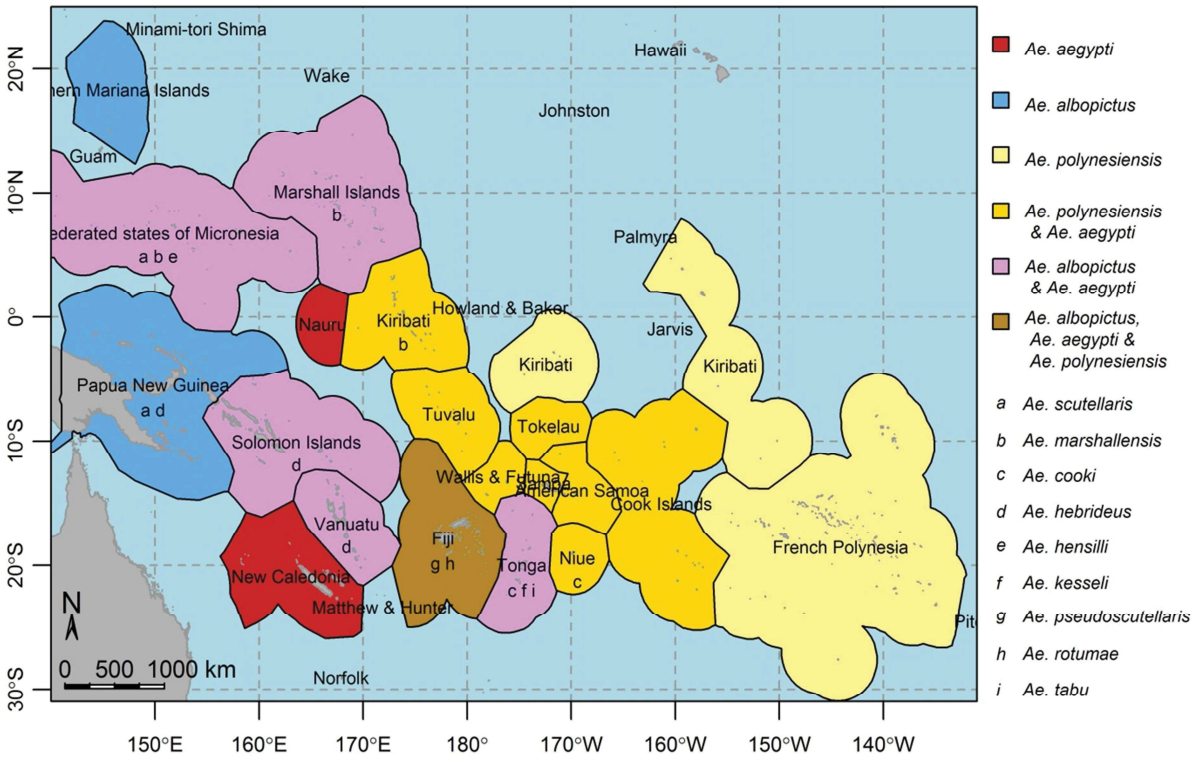


Figure 1.3 : Répartition des vecteurs de la dengue dans la Pacifique. *Les couleurs correspondent aux trois principales espèces vectrices de la dengue, Ae. aegypti, Ae. albopictus et Ae. polynesiensis. La présence d'autres espèces est indiquée par des lettres sous le nom de chacun des pays.*

Dans le Pacifique, outre *Ae. aegypti* et *Ae. albopictus*, les espèces identifiées comme capables de transmettre la dengue sont : *Ae. polynesiensis*, *Ae. scutellaris*, *Ae. marshallensis*, *Ae. cooki*, *Ae.hebrideus*, *Ae. hensilli*, *Ae. kesseli*, *Ae. pseudoscutellaris*, *Ae. rotumae*, et *Ae. tabu* (Guillaumot, 2005). La Figure 1.3 montre une carte de la répartition des différentes espèces vectrices dans chacun des pays. La seule espèce vectrice de la dengue en Nouvelle-Calédonie est *Ae. aegypti*. Les moustiques de l'espèce *Ae. vigilax*, très présents en Nouvelle-Calédonie à proximité des mangroves, ne transmettent pas la dengue.

1.1.2.b. Cycle de vie et facteurs influençants

Cycle de vie

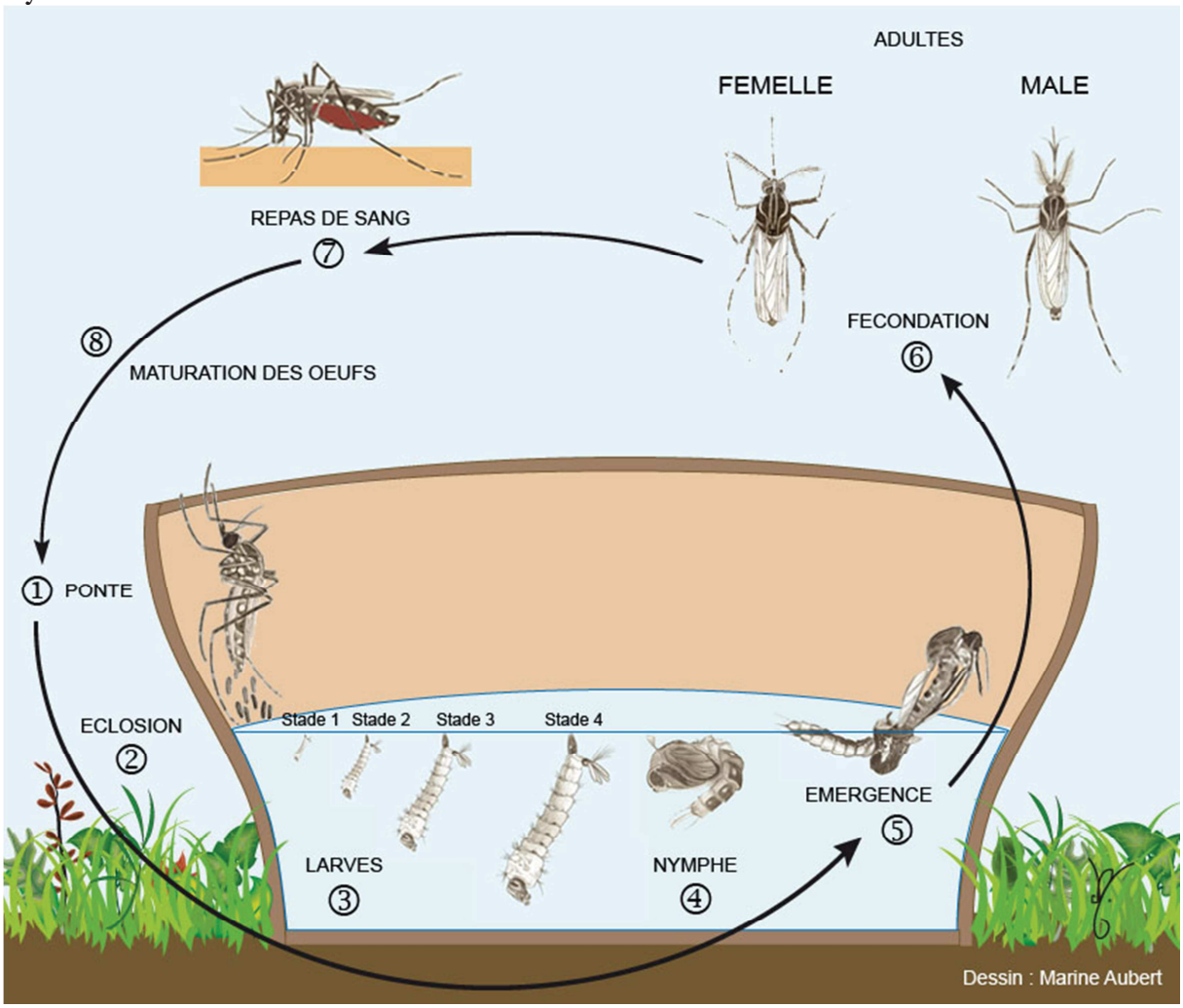


Figure 1.4 : Cycle de vie des moustiques *(source : Institut Pasteur de Nouvelle-Calédonie, IPNC).*

Les moustiques ont une phase de vie adulte aérienne et une phase de vie larvaire aquatique (Figure 1.4). Les femelles pondent leurs œufs à proximité d'une zone aquatique. Les gîtes et les modalités de ponte varient en fonction des espèces. Après un temps qui dépend de chaque espèce de moustique et des conditions environnementales, chaque œuf éclot pour donner une larve. Il existe, chez tous les

moustiques, quatre stades larvaires (notés L1 à L4). Les larves ont un mode de vie aquatique. Elles se nourrissent de matières organiques présentes dans l'eau. Elles respirent soit directement à la surface, soit, comme les larves d'*Aedes*, au moyen d'un siphon situé au bout de l'abdomen. Au cours de la pupation, les larves de stade 4 se transforment en nymphe, un stade éphémère de métamorphose permettant au moustique d'acquérir ses caractéristiques d'adulte, suite à quoi il y a émergence de l'adulte. Après reproduction, les femelles partent en quête d'un repas sanguin, nécessaire au développement des œufs : seules les femelles piquent des hôtes. Après une phase de maturation des œufs, et en fonction des conditions environnementales, la femelle pond ses œufs et recommence un cycle de reproduction. L'ensemble des phases de reproduction (fécondation, repas sanguin et ponte) des adultes femelles est appelé « cycle gonotrophique », chaque femelle pouvant effectuer plusieurs cycles au cours de sa vie. La durée d'un cycle gonotrophique correspond à la durée séparant deux étapes similaires du cycle, par exemple, deux repas sanguins.

Particularités des *Aedes*

Les femelles *Aedes* ne pondent pas directement dans l'eau, mais sur des substrats humides, le plus souvent la paroi d'un récipient contenant de l'eau. En conséquence, les œufs d'*Aedes* restent émergés à l'air libre, et n'éclosent qu'au premier contact avec l'eau, ce qui nécessite que le niveau d'eau du récipient augmente et ce qui rend l'éclosion partiellement dépendante des précipitations. Les *Aedes* préfèrent généralement pondre dans des gîtes ombragés contenant de l'eau claire et pauvre en matière organique (Delatte et al., 2008). Les œufs sont capables de survivre à l'air libre pendant plusieurs années. Les larves d'*Aedes* présentent la particularité de pouvoir se développer dans de très petites quantités d'eau (1 mL), ce qui complique la recherche des gîtes larvaires et rend le contrôle des populations difficile. Les gîtes de ponte peuvent être très variés et incluent les citernes non couvertes servant au stockage d'eau dans les pays endémiques (Barbazan et al., 2008; Espinoza Gómez et al., 2001; Stewart Ibarra et al., 2013) les pneus, les boîtes de conserve abandonnées, les sous-pots de plantes, les récipients plastiques (Favier et al., 2006a; Stewart Ibarra et al., 2013) ou encore les creux d'arbre ou de rocher (Delatte et al., 2008).

Les femelles *Aedes* fractionnent leur repas sanguin, c'est-à-dire qu'elles piquent plusieurs hôtes au cours d'un même cycle gonotrophique (Scott et al., 2000), augmentant le potentiel de transmission des maladies. Les femelles *Aedes* piquent durant la journée, à différents moments selon l'espèce. En conséquence, contrairement au paludisme, l'utilisation nocturne de moustiquaires ne permet pas de réduire le risque de contracter la dengue.

Les hôtes diffèrent d'une espèce d'*Aedes* à l'autre. Le caractère fortement anthropophile d'*Aedes aegypti* le rend inféodé au milieu urbain. *Aedes aegypti* est un moustique généralement endophile (vivant à l'intérieur des maisons), ce qui lui permet parfois de survivre malgré des conditions climatiques *a priori* défavorables (Jansen and Beebe, 2010). Les *Aedes albopictus* ont un spectre d'hôte plus large. Ils sont donc capables de vivre en zone semi-urbaine ou rurale. On les trouve également en lisière de forêt.

La distance de vol des *Aedes* varie en fonction de la disponibilité en hôtes. Dans les zones urbaines, pour *Aedes aegypti*, la majorité des adultes restent dans la maison dans laquelle ils ont émergé ou les maisons adjacentes, se déplaçant en moyenne de 40 à 80 m (Getis et al., 2003; Harrington et al., 2005; Maciel-de-Freitas, et al., 2007; Muir and Kay, 1998), mais des distances plus élevées pouvant aller jusqu'à 500 mètres ont été enregistrées (Harrington et al., 2005; Liew and Curtis, 2004; Muir and Kay, 1998).

Influence des facteurs environnementaux sur les traits de vie des *Aedes*

Les traits de vie pouvant influencer la transmission de la dengue sont tous ceux qui ont trait à la densité vectorielle des femelles adultes (survie des femelles adultes, fécondité des femelles en nombre d'œufs par cycle gonotrophique, mortalité des différents stades immatures, ratio du nombre de femelles et de males à l'émergence, rapidité de la succession des générations, i.e. durée des stades pré-imaginaux), ainsi que ceux qui ont trait au nombre d'hôtes qu'une femelle peut piquer au cours de sa vie (durée du cycle gonotrophique des femelles adultes et nombre d'hôtes par repas sanguin). Bien que la taille des adultes influence la capacité de certaines espèces de moustique à transmettre des arbovirus (Baqar et al., 1980; Grimstad and Haramis, 1984; Grimstad and Walker, 1991; Nasci and Mitchell, 1994; Takahashi, 1976), son influence sur la transmission de la dengue n'est pas connue, certaines études montrant un lien effectif (Sumanochitrapon et al., 1998), et d'autres infirmant ce lien (Bosio et al., 1998; Schneider et al., 2007).

Des études utilisant principalement des techniques de capture – marquage – recapture ont permis d'estimer les **taux de survie journalier des femelles adultes** d'*Aedes agypti* : ces taux sont compris entre 0,80 et 0,95 (moyenne 0,913) en Guyane Française (Fouque et al., 2006), entre 0,86 et 0,91 en Australie (Muir and Kay, 1998), et estimés à 0,9 dans le centre du Vietnam (Hugo et al., 2014). Certaines études montrent que le taux de survie varie avec l'âge (Harrington et al., 2008; Muir and Kay, 1998; Styer et al., 2007) tandis que d'autres suggèrent que la variation des taux de survie avec l'âge n'intervient que si les conditions environnementales sont suffisamment favorables pour permettre un vieillissement des populations de moustique (Brady et al., 2013; Hugo et al., 2014). La durée de vie des femelles adultes varie de 2 à 29 jours (moyenne 14 jours) en conditions naturelles en Guyane (Fouque et al., 2006). Une étude de modélisation des données issues de 410 études ayant estimé la survie des femelles adultes de manière expérimentale ou sur le terrain a permis d'estimer à 38 jours l'âge médian des populations d'*Aedes aegypti* et à 60 jours celui des *Aedes albopictus* dans des conditions environnementales optimales, la plus grande longévité enregistrée sur le terrain étant respectivement de 100 et 120 jours pour chacune des deux espèces (Brady et al., 2013).

Les facteurs influençant la mortalité des adultes sont la prédation, les maladies, l'âge (Harrington et al., 2008; Muir and Kay, 1998), et pour les conditions environnementales, la température et l'humidité. La mortalité des femelles adultes augmente lorsque l'humidité relative de l'air diminue (Canyon et al., 1999), et elle est minimale pour une plage de températures comprises entre 20 °C et 30 °C, l'influence

de la température sur la mortalité dépendant de l'âge et de l'espèce de moustique (Brady et al., 2013; Degallier et al., 2012; Rueda et al., 1990). *Aedes albopictus* présente des taux de survie plus élevés qu'*Aedes aegypti*, mais les jeunes adultes *Aedes aegypti* sont capables de supporter une gamme de température plus large. Lorsque les populations vieillissent (> 40 jours), *Aedes aegypti* a une survie optimale uniquement pour des températures autour de 21° C, tandis que *Aedes albopictus* peut survivre à une gamme de température variant de ~15 à 35 °C (Brady et al., 2013).

La durée du **cycle gonotrophique** diminue lorsque la température augmente (Focks et al., 1993a; Pant and Yasuno, 1973).

Une vaste étude de terrain a montré qu'une grande majorité des *Aedes aegypti* se nourrit sur au moins deux hôtes, et a évalué le nombre d'hôtes piqué par jour par femelle à 0,76 en Thaïlande, et à 0,63 à Puerto Rico. L'effet de la température sur la **fréquence des repas sanguin** varie en fonction des zones d'étude (Scott et al., 2000).

Des études en laboratoire ont montré que la **fécondité des femelles** est contrôlée par l'humidité (la fécondité augmente avec l'humidité (Canyon et al., 1999)), ainsi que par la variation diurne de température (de larges variations diurnes induisent une baisse de fécondité de 25 % (Carrington et al., 2013)).

Le temps de développement embryonnaire et la survie des **œufs** dépendent de la température et de l'humidité (Focks et al., 1993a; Mohammed and Chadee, 2011). Les œufs d'*Aedes albopictus* sont capables d'entrer en diapause, et de survivre plusieurs années malgré des conditions climatiques adverses (Pumpuni et al., 1992) tandis que ceux d'*Aedes aegypti* présentent des taux de mortalité plus élevés pour des températures tempérées (Fischer et al., 2011). Les précipitations influencent la densité vectorielle d'adultes en mettant en eau les gîtes sur les parois desquels les œufs sont situés, l'immersion étant nécessaire à l'éclosion des œufs (Focks et al., 1993a).

La **durée de développement des stades immatures** d'*Aedes aegypti* est contrôlée par la température moyenne (Couret et al., 2014), la variation diurne de température (Carrington et al., 2013), la quantité de nourriture disponible pour les larves (Arrivillaga and Barrera, 2004; Couret et al., 2014), la densité de larves par gîte (Couret et al., 2014). Les valeurs seuils de température au-delà et en-deçà desquelles les larves d'*Aedes aegypti* ne peuvent se développer en adulte varient en fonction des études (14 et 36 °C (Bar-Zeev, 1958), 8 °C et 42 °C (Focks et al., 1993a)). La durée totale de l'éclosion des œufs à l'émergence des adultes peut varier de 16 jours à 21°C, à 10 jours à 30 °C (Couret et al., 2014). Le cycle de développement des larves peut varier de 18 jours si les larves n'ont accès qu'à une quantité faible de nourriture, à 8 jours si la quantité de nourriture disponible est suffisante (Arrivillaga and Barrera, 2004)).

La **survie des larves** a été moins étudiée que celle des adultes. Elle dépend de la prédation, de la quantité de nourriture disponible (27 % de survie à 0,1 mg de nourriture/larve/jour, 88% de survie à 1

mg/larve/jour, puis la survie diminue si la quantité de nourriture dépasse 1 mg/larve/jour (Arrivillaga and Barrera, 2004)), de la présence de larves d'autres espèces, de la température (survie plus élevée pour des températures élevées (Teng and Apperson, 2000)) ou de la variation journalière de la température (une baisse de 6% de survie des larves est enregistrée pour des variations diurnes de 18 °C (Carrington et al., 2013)). Les précipitations peuvent jouer de manière défavorable sur la mortalité des larves, puisque de fortes précipitations peuvent faire déborder les gîtes et les « lessiver ».

Le **ratio** du nombre de **femelles** par rapport au nombre de **mâles** à l'émergence augmente avec la température moyenne lorsqu'il existe des variations diurnes importantes de température, passant de 0,9 à une température moyenne de 26/27 °C à 1,2 à une température moyenne de 34/35 °C (Mohammed and Chadee, 2011).

1.1.3. Formes cliniques et épidémiologie des formes cliniques

Formes cliniques

La dengue se présente sous des formes cliniques variées, allant dans la majorité des cas de l'infection inapparente à une forme fébrile bégnine dite dengue classique (DF pour « Dengue Fever »). Elle peut aussi survenir sous sa forme sévère, la dengue hémorragique (DHF pour « Dengue Haemorrhagic Fever »), caractérisée par une anomalie de l'hémostase, et une perméabilité capillaire accrue pouvant entraîner des hémorragies, se compliquer d'un état de choc (DSS pour « Dengue Shock Syndrome ») et aboutir au décès (Gubler, 1998; Guzman et al., 2010). La précocité de la prise en charge des formes sévères est un facteur déterminant de l'issue de la DHF. La période d'incubation est de 4 à 7 jours (extremum de 2 à 14 jours).

La dengue classique est caractérisée par une fièvre d'apparition brutale et durant entre 5 et 7 jours. La fièvre est souvent accompagnée d'un ou plusieurs des symptômes suivants : fatigue intense, rougeur de la face, érythème, douleurs généralisées, myalgie, arthralgie, maux de tête, nausées, vomissements. Les examens sanguins révèlent souvent une leucopénie ou une légère thrombocytopénie (Halstead, 1970; Rigau-Pérez et al., 1998). Des signes d'hémorragie peuvent apparaître, sous la forme de pétéchies, d'épistaxis, de saignements gingivaux, d'hématurie, ou du signe du tourniquet positif (apparition de pétéchies après pose d'un brassard comprimant les vaisseaux sanguins) (Rigau-Pérez et al., 1998). Certains patients peuvent présenter des formes avec lésions hépatiques, cardiomyopathie ou encéphalopathie et complications neurologiques dont les conséquences peuvent être graves (Rigau-Pérez et al., 1998). La sévérité des formes de dengue classique varie avec l'âge, en général bégnines chez les enfants, plus sévères chez les adultes.

La dengue hémorragique débute le plus souvent comme une dengue classique. Elle est caractérisée par l'apparition, entre le 3ème et le 7ème jour, de saignements pouvant devenir importants, d'une

thrombocytopénie (concentration de plaquettes inférieure à 100 000 / µL), et de signes de fuite plasmatique (hématocrite élevé, hypo-protéinémie, hypo-albuminémie, effusions pleurales ou abdominales) (Rigau-Pérez et al., 1998).

Les fuites plasmatiques importantes des patients avec DHF peuvent conduire à un choc circulatoire, dont les signes d'alerte sont hypotension, douleur abdominale prolongée, vomissements fréquents, apathie ou excitation, et hypothermie avec suées, et prostration (Rigau-Pérez et al., 1998).

La proportion des différentes formes varie en fonction de la zone géographique, du contexte épidémiologique, du statut immunitaire des patients et du sérotype circulant.

Proportion de formes inapparentes

La proportion de formes inapparentes est variable d'un endroit à l'autre. Elle est cependant toujours supérieure à la proportion de formes cliniques. Différentes enquêtes épidémiologiques prospectives ont montré que la proportion de personnes présentant des formes inapparentes parmi l'ensemble des personnes enrôlées et ayant présenté une séroconversion est de 52-53 % en Thaïlande (Endy et al., 2002; Mammen Jr et al., 2008), de 64 % aux Pays-Bas (cas importés uniquement) (Baaten et al., 2011), de 75 % dans différents pays d'Asie et d'Amérique Latine (Dussart et al., 2012), de 78 % à Singapour (Yap et al., 2013), 87 % dans une autre Province de Thaïlande (Burke et al., 1988) et jusqu'à 93 % au Vietnam (Thai et al., 2011).

Dans une même région, la proportion de formes inapparentes varie en fonction des années, de l'incidence de l'année en cours, et, pour une même année, en fonction de facteurs géographiques locaux (proportions très différentes parmi les enfants de plusieurs écoles distantes de moins de 15 kilomètres) (Endy et al., 2002, 2011). Certaines études ont trouvé différentes proportions de formes inapparentes en fonction de l'âge, les résultats étant contradictoires entre les différentes études (Thai et al., 2011; Yap et al., 2013), tandis que d'autres n'ont trouvé aucune influence de l'âge, du sexe, ou encore de l'historique de vaccination contre l'encéphalite japonaise (Endy et al., 2002, 2011). Une revue récente note que le facteur le plus constamment associé à une forte proportion de formes asymptomatiques est un taux d'incidence élevé l'année précédente, suggérant un rôle important de l'immunité croisée transitoire entre différents sérotypes dans l'expression clinique de la maladie (Grange et al., 2014).

Facteurs de risque de la DHF

Les deux principaux facteurs de risque identifiés dans le développement d'une DHF sont une infection secondaire hétérologue (Burke et al., 1988; Halstead, 1970), et la virulence de l'agent pathogène en cause (Rosen, 1996).

Halstead a développé la théorie des « anticorps facilitants » (ADE pour « Antibody-Dependant Enhancement » en anglais), selon laquelle les anticorps non-neutralisants acquis lors d'une première

infection faciliteraient l'infection des cellules mononuclées et la libération de substances vasoactives responsables d'une perméabilité capillaire accrue en cas d'infection secondaire par un sérotype hétérologue (Brandt et al., 1982). Cette théorie a été renforcée par de nombreuses études épidémiologiques montrant que le risque de développer une DHF était plus important lors d'infections secondaires par des sérotypes hétérologues que lors d'infections primaires (Burke et al., 1988; Guzman and Kouri, 2003). Cependant, des formes de DHF peuvent aussi survenir en cas d'infection primaire.

Le sérotype 2 est plus souvent associé à la DHF que les autres sérotypes. En particulier, le génotype Asiatique semble plus virulent que le génotype d'Amérique Latine, et est responsable de l'émergence de la DHF en Amérique Latine dans les années 1970 (Guzmán et al., 2002; Rico-Hesse et al., 1997).

Les autres facteurs de risque de DHF incluent une charge virale élevée (Murgue et al., 2000; Vaughn et al., 2000), l'âge en Asie (les enfants sont plus touchés) (Guzman and Kouri, 2003; Vaughn et al., 2000), les co-morbidités chroniques (Guzman and Kouri, 2003; Leo et al., 2011) telles que l'asthme, la thalassémie et le diabète de type II, être né d'une mère infectée par la dengue (Kliks et al., 1988), l'origine ethnique (à Cuba, les blancs ont plus de chance de développer une dengue hémorragique que les populations d'origine africaine) (Kouri et al., 1989) ou des variations génétiques intra-génotypiques (Steel et al., 2010).

1.1.4. Diagnostic

1.1.4.a. Diagnostic expérimental

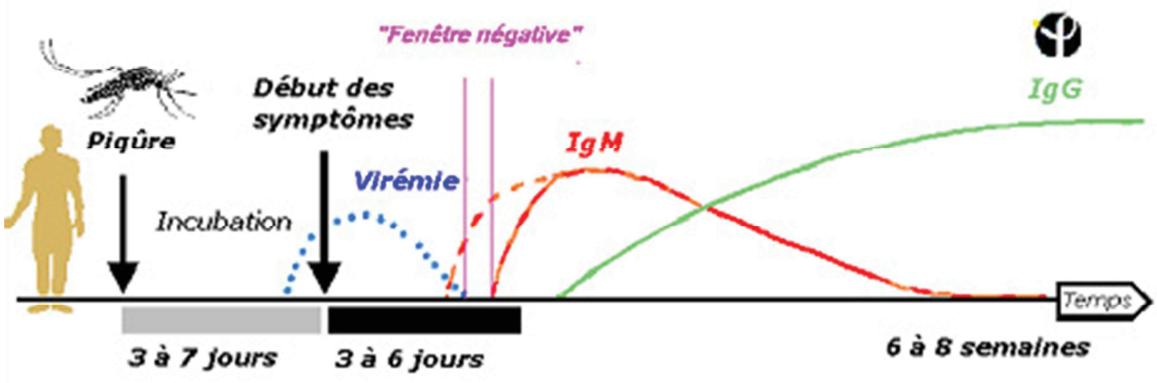


Figure 1.5 : Cinétique des paramètres virologiques et sérologiques. *La « fenêtre négative » correspond à une période pendant laquelle la dengue est difficile à détecter de manière expérimentale (baisse des titres viraux, et immunité humorale non encore en place) (source : IPNC, http://www.ipnc.nc/spip.php?article69#top).*

L'infection d'un patient par le virus de la dengue peut être confirmée en laboratoire soit par l'identification directe du virus durant la phase virémique (J1 à J5) soit par la détection indirecte des anticorps dirigés contre le virus à partir de J5 (Figure 1.5). Les tests virologiques comprennent

l'isolement du virus sur culture cellulaire, la détection de l'ARN viral par des techniques d'amplification comme la RT-PCR (pour « Reverse-Transciptase Polymerase Chain Reaction »), ou la détection d'antigènes viraux (le plus souvent l'antigène NS1) par ELISA (pour « Enzyme-Linked ImmunoSorbent Assay »). Les tests sérologiques identifient les anticorps (immunoglobulines précoces, *i.e.* IgM, ou immunoglobulines tardives, *i.e.* IgG) dirigés contre le virus de la dengue par ELISA ou inhibition de l'hémagglutination (IH). Seuls les tests ELISA NS1 ne permettent pas d'identifier le sérotype en cause.

1.1.4.b. Diagnostic différentiel

La dengue classique, dont les symptômes sont non spécifiques, peut être confondue avec de nombreuses autres maladies provoquant un syndrome fébrile (Rigau-Pérez et al., 1998; WHO, 2009) comme le chikungunya, la leptospirose, la fièvre Zika (toutes présentes en Nouvelle-Calédonie (DASS-NC, 2011)), la rougeole, la fièvre typhoïde, l'encéphalite japonaise ou le palu (présentes de manière anecdotique ou absentes du territoire). La non spécificité des symptômes de la dengue classique entraîne un biais dans les données de surveillance des pays où plusieurs virus provoquant les mêmes symptômes circulent et qui n'ont pas les moyens financiers de confirmer le diagnostic de manière expérimentale.

1.1.5. Traitement

Il n'existe à l'heure actuelle aucun traitement spécifique de la dengue, le traitement est uniquement symptomatique, et seule une prise en charge précoce des formes hémorragiques (hospitalisation et transfusion de plasma) permet de réduire le risque de survenue d'un choc circulatoire.

1.1.6. Prophylaxie médicale et sanitaire

Prophylaxie médicale

Il n'existe, à l'heure actuelle, aucun vaccin contre la dengue. En effet, le phénomène des anticorps facilitants, qui augmente le risque de DHF lors d'infections secondaires, oblige à créer des vaccins tétravalents, *i.e.* protégeant contre les quatre sérotypes à la fois, ce qui est problématique (Guzman et al., 2010). Le manque de modèle animal pour l'étude de la dengue est également un frein important à l'élaboration d'un vaccin.

Le candidat-vaccin le plus avancé dans les phases d'essais cliniques est un vaccin créé par Sanofi – Pasteur. C'est un vaccin recombinant vivant atténué composé de chimères spécifiques à chaque sérotype. Lors des essais cliniques de phase II-b, ce candidat vaccin a montré une bonne protection contre les virus DENV-1, DENV-3 et DENV-4, mais une protection insuffisante contre le virus DENV-2 (Sabchareon et al., 2012). Les essais cliniques de phase III montrent une efficacité de 50,0 % contre le

sérotype 1, de 35,0 % contre le sérotype 2, de 78,4 % contre le sérotype 3 et de 66,7 % contre le sérotype 4 (Capeding et al., 2014).

Prophylaxie sanitaire

Le seul moyen de lutte à l'heure actuelle consiste à diminuer la fréquence des contacts hôte/vecteur. Les moyens de lutte peuvent être individuels (utilisation d'anti-moustique), ou collectifs. La lutte collective consiste pour l'instant, dans la majorité des pays, en une lutte chimique dirigée contre les adultes par vaporisation dans l'air d'insecticides, une lutte chimique dirigée contre les larves dans les pays où certains types de containers faciles à atteindre ont été identifiés comme responsables de la majorité de production des larves, et en des campagnes de formation et de sensibilisation de la population pour l'élimination des gîtes larvaires. Plusieurs moyens innovants de lutte sont à l'étude, notamment l'utilisation d'une bactérie parasite des vecteurs, *Wolbachia*, pour remplacer les populations de vecteur existantes par des populations infectées moins compétentes. Ces populations de vecteurs infectés ont montré de bons taux d'invasion des populations naturelles, et seraient plus sûrs d'utilisation que des populations génétiquement modifiées, la bactérie étant déjà présente à l'état naturel (Hoffmann et al., 2011). A ce jour, ces nouvelles techniques de lutte n'ont pas encore été mises en place de manière opérationnelle.

1.1.7. Epidémiologie

1.1.7.a. Les hôtes

L'hôte principal du virus de la dengue est l'homme. Il peut également infecter les singes du genre *Macaca* et *Presbytis* en Asie, et les singes de l'espèce *Erythocebus patas* en Afrique (Gubler, 1998; Holmes and Twiddy, 2003).

1.1.7.b. Les modes de transmission

Le mode principal de transmission entre hôtes est la transmission vectorielle. Des cas de transmission par transfusion sanguine à partir d'individus porteurs sains ont également été rapportés (Tambyah et al., 2008). En Nouvelle-Calédonie, un cas de transmission potentielle de la mère à l'enfant par le lait a été rapporté (Barthel et al., 2013).

Chez le moustique, il existe une transmission verticale transovarienne : les œufs pondus peuvent être infectés, et donner naissance à des adultes infectés (Rosen and Gubler, 1974; Rosen et al., 1983; Shroyer, 1990). Les taux de transmission varient en fonction du sérotype et de l'espèce vectrice en cause, mais sont en général relativement faibles (~1 % des femelles *Ae. albopictus* sont capables de transmettre le virus à leur descendance (Shroyer, 1990)).

1.1.7.c. Le cycle de transmission viral

Un hôte infectieux possède des particules virales dans son sang, particules qu'il transmet à un moustique sain au cours d'un repas sanguin. Un moustique infecté ne peut devenir infectieux (*i.e.* capable de transmettre le virus à un hôte sain) qu'après une phase de migration et de réplication virale au sein du moustique appelée **incubation extrinsèque**, d'une durée de 1 à 3 semaines.

Il existe trois cycles de transmission virale : un cycle sylvatique singe – moustique – singe en Afrique et en Asie, un cycle de transmission en zone semi-urbaine ou rurale, dont le vecteur principal est *Aedes albopictus*, capable, grâce à son large spectre d'hôtes, de survivre dans des zones peu habitées, et un cycle urbain, dont le vecteur principal est *Aedes aegypti*, que l'on ne trouve qu'en milieu urbain (Gubler, 1998).

Les épidémies liées à *Aedes aegypti* sont souvent plus « explosives » que celle liées à *Aedes albopictus*. En effet, le large spectre d'hôte d'*Aedes albopictus* augmente l'intervalle de temps nécessaire au passage du virus d'un hôte à un autre. Le cycle de transmission s'étale donc dans le temps (Lambrechts et al., 2010).

1.1.7.d. Facteurs liés à l'hôte influençant le cycle de transmission

Les facteurs liés à l'hôte influençant le cycle de transmission sont tous ceux modifiant la fréquence des contacts hôte/vecteur, ceux modifiant la probabilité qu'un moustique sain s'infecte lorsqu'il pique un hôte infectieux, et ceux modifiant la probabilité qu'un hôte sain s'infecte lorsqu'il se fait piquer par un moustique infectieux.

Statut immunitaire de l'hôte

Un hôte immunisé contre l'un des sérotypes ne peut être infecté de nouveau par le même sérotype.

Comportement de l'hôte

L'hôte peut réduire la fréquence à laquelle il se fait piquer en mettant en place des mesures de protection contre les piqures de moustique (pose de moustiquaires dans les maisons, utilisation d'anti-moustique, lutte anti-vectorielle à l'échelle collective pour réduire la densité vectorielle, port de vêtements particuliers…). Souvent négligé dans les études de modélisation du fait de l'absence de données, cet aspect semble pourtant primordial, puisqu'il permet d'expliquer l'absence de cas de dengue dans certains pays développés aux conditions climatiques favorables (Reiter et al., 2003).

Phase virémique

La phase virémique est la phase au cours de laquelle le sang d'un hôte contient des particules virales et correspond donc à la période au cours de laquelle l'hôte peut transmettre le virus à des vecteurs sains (ou à d'autres humains via transfusion sanguine (Tambyah et al., 2008)). Les paramètres importants dans la transmission des virus au moustique sont la **durée de la virémie** et la **charge virale** des hôtes au cours de la virémie.

La virémie des patients asymptomatiques n'a été que très peu étudiée, et très peu d'information existe quant au potentiel infectieux de ces patients. La possibilité de transmission par transfusion sanguine et, en Nouvelle-Calédonie, la présence certaines années de cas index autochtones sans qu'aucun cas importé n'ait été détecté dans le mois précédent laissent penser que les patients asymptomatiques virémiques représentent une source de transmission. Une étude à Jakarta a montré que 2 patients sur 8 asymptomatiques présentaient une virémie détectable (Beckett et al., 2005). Une autre étude au Cambodge a montré que 17 patients asymptomatiques présentaient des titres viraux non significativement différents des titres viraux des 243 patients symptomatiques inclus dans l'étude (Duong et al., 2011). La durée de la virémie des patients asymptomatiques est inconnue, du fait de la difficulté de détecter ces patients au bon moment.

Chez les patients symptomatiques, la virémie dure en moyenne 4 à 5 jours (Gubler et al., 1981; Murgue et al., 2000; Rigau-Pérez et al., 1998; Vaughn et al., 2000), avec un pic virémique le plus souvent au 3ème jour (Vaughn et al., 2000) et une décroissance des titres viraux concomitante avec la défervescence (Brown et al., 1996; Murgue et al., 2000) et l'augmentation des titres d'anticorps. La virémie peut débuter (au moins) un jour avant l'apparition des symptômes (Tambyah et al., 2008), et varier de 2 à 12 jours (Gubler et al., 1981). Le lien entre la charge virale des hôtes, la sévérité des symptômes et la parité de l'infection (primaire ou secondaire) diffère d'une étude à l'autre (Gubler et al., 1981; Murgue et al., 2000; Vaughn et al., 2000).

Il serait intéressant de connaître les facteurs influençant la durée de la virémie, ou la charge virale des patients. Cependant, c'est un domaine très peu étudié. D'autre part, la dose infectieuse nécessaire à l'infection des moustiques semble elle-même très variable (Pongsiri et al., 2014).

1.1.7.e. Facteurs liés au vecteur influençant le cycle de transmission

Les facteurs liés au vecteur influençant le cycle de transmission sont tous ceux modifiant la fréquence des contacts hôte/vecteur (*i.e.* les traits de vie du moustique), ceux modifiant la capacité d'un moustique infecté à devenir infectieux (appelée « **compétence vectorielle** »), ainsi que ceux modifiant la vitesse à laquelle un moustique infecté devient infectieux (*i.e.* la durée de la période d'incubation extrinsèque).

Traits de vie du moustique

Voir 1.1.2.b page 42.

Compétence vectorielle et période d'incubation extrinsèque

La compétence vectorielle est l'aptitude intrinsèque d'un arthropode vecteur à assurer la transmission biologique et active d'un agent infectieux. Elle est mesurée en infectant oralement des moustiques et en mesurant la proportion de femelles présentant l'agent pathogène dans les glandes salivaires (ou parfois dans la tête) au cours du temps, tout en contrôlant les conditions environnementales (Talbalaghi et al., 2010).

Après ingestion par le moustique, les virus infectent les cellules épithéliales de l'intestin dans lesquelles il y a multiplication. Cette étape nécessite une liaison physique entre les virus et les cellules épithéliales de l'intestin du moustique via des récepteurs. Elle peut constituer une première barrière à la transmission virale. Les virions matures sortent ensuite de ces cellules pour aller dans l'hémolymphe, ce qui constitue une deuxième barrière à la transmission (Black et al., 2002). Ils infectent ensuite des organes secondaires, dont les glandes salivaires. La nécessité de pénétrer dans les cellules des glandes salivaires peut constituer une troisième barrière à la transmission (Cox et al., 2011). L'existence de cette troisième barrière implique que les études de compétence vectorielle mesurées sur la tête entière des moustiques (étude de la dissémination du virus dans le corps de l'insecte) sont difficilement comparables à celles étudiant la présence du virus dans la salive (études de compétence vectorielle au sens propre).

La compétence vectorielle dépend :

- des titres viraux (et donc de la virémie des hôtes infectés) (Black et al., 2002; Girod et al., 2011; Watts et al., 1987) ;
- de l'espèce vectrice en cause (*Aedes aegypti* est plus compétent qu'*Aedes albopictus* pour le virus de la dengue, tandis que pour le virus du Chikungunya, c'est l'inverse (Lambrechts et al., 2010; Talbalaghi et al., 2010; Vazeille et al., 2008)) ;
- peut-être de la taille des moustiques adultes (Bosio et al., 1998; Schneider et al., 2007; Sumanochitrapon et al., 1998) ;
- de facteurs environnementaux comme la température moyenne (Watts et al., 1987) ou les variations diurnes de température (Liu-Helmersson et al., 2014; Lambrechts et al., 2011) ;
- de facteurs génétiques liés au vecteur (Bosio et al., 2000; Fansiri et al., 2013). Ainsi, différentes sous-populations d'une même espèce de moustique peuvent présenter des compétences vectorielles différentes au sein d'un même pays (Ye et al., 2014) ;
- de facteurs génétiques liés au virus (Fansiri et al., 2013).

Un paramètre important lié à la compétence vectorielle est la vitesse à laquelle un moustique infecté devient infectieux, *i.e.* la durée de la période d'incubation extrinsèque. La durée d'incubation extrinsèque est sous le contrôle principal de la température moyenne (Liu-Helmersson et al., 2014; Lambrechts et al., 2011; Watts et al., 1987), mais également sous le contrôle des titres viraux (Girod et al., 2011; Watts et al., 1987) et de facteurs génétiques liés au virus (Anderson and Rico-Hesse, 2006). Globalement, la durée d'incubation extrinsèque diminue (et donc la transmission virale augmente) lorsque les températures augmentent, le temps nécessaire à ce que 50 % de femelles infectées deviennent infectieuses passant par exemple de 12 jours à 30 °C à 7 jours à 32 °C ou 35 °C (Watts et al., 1987).

Capacité vectorielle

L'ensemble des paramètres liés au vecteur et influençant la transmission de la dengue (traits de vie, compétence vectorielle et durée d'incubation extrinsèque) sont regroupés sous le terme de capacité vectorielle. La capacité vectorielle est un concept développé pour l'étude de la Malaria. Elle est définie comme le nombre moyen d'infections potentielles créés par une population de moustiques par personne infectée par unité de temps (Garrett-Jones, 1964; MacDonald, 1957). Ainsi, *Aedes albopictus* a une moins bonne compétence vectorielle qu'*Aedes aegypti* pour la dengue, mais sa capacité vectorielle est augmentée du fait de sa plus grande longévité (Talbalaghi et al., 2010). Par contre elle est diminuée du fait de son spectre d'hôtes plus large (Lambrechts et al., 2010).

1.1.7.f. Facteurs liés au virus influençant le cycle de transmission

Influence sur la compétence et la capacité vectorielle

L'influence des facteurs génétiques liés au virus sur la capacité vectorielle semble particulièrement importante dans les phénomènes d'émergence. En effet, les espèces vectrices d'Amérique Latine présentent une capacité vectorielle nettement plus élevée pour le génotype asiatique que pour le génotype américain du virus DENV-2, du fait d'une période d'incubation extrinsèque plus courte. Cette capacité accrue pour le génotype asiatique serait à l'origine du remplacement du génotype américain par le génotype asiatique et l'émergence de la DHF associée à ce génotype en Amérique Latine (Anderson and Rico-Hesse, 2006). Un autre exemple est l'émergence du virus Chikungunya à la Réunion suite à une mutation spécifique du virus qui a rendu le vecteur *Aedes albopictus* très compétent et a permis une large diffusion des virus mutants (Arias-Goeta et al., 2013).

Influence sur l'expression clinique de la maladie

La variabilité génétique des virus induit des variations dans l'expression clinique de la maladie. L'impact des variations cliniques sur le cycle de transmission est peu connu, du fait du manque de données sur les formes asymptomatiques, mais elle modifie de manière significative la perception que nous avons de ce cycle au travers des données collectées par les réseaux de surveillance.

1.2. Modélisation de la dynamique spatiale de la dengue : revue bibliographique

1.2.1. Objectifs de la modélisation

Les études citées précédemment, réalisées en laboratoire ou sur le terrain, permettent d'estimer certains paramètres précis du cycle de transmission, et d'étudier l'impact éventuel de facteurs environnementaux sur ces paramètres. Elles ne permettent cependant pas de comprendre comment tous ces paramètres interagissent entre eux, ni de quantifier leur influence sur la distribution spatio-temporelle des cas de dengue. Pour comprendre ces interactions et identifier les facteurs clés de la transmission, des études de modélisation sont nécessaires. Les objectifs de la modélisation sont :

- Décrire la distribution spatio-temporelle des cas de dengue afin d'identifier les zones et les périodes à risque, ainsi que d'éventuels schémas spatio-temporels récurrents (Cummings et al., 2004; Teurlai et al., 2012; Thai et al., 2010) ;
- Comprendre par quels mécanismes hôtes, virus et vecteurs interagissent entre eux pour aboutir à la distribution spatio-temporelle observée (Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006; Favier et al., 2006b);
- Quantifier et hiérarchiser l'impact des différents facteurs sur la distribution spatio-temporelle des cas, et ainsi identifier les facteurs clés (Descloux et al., 2012; Johansson et al., 2009a; Wu et al., 2009) ;
- En fonction des résultats, fournir une information pertinente aux autorités de santé sur les mesures de contrôle les plus à même de faire diminuer l'incidence de la maladie, ou, si les facteurs identifiés ne sont pas contrôlables, comme la température, fournir des indicateurs de prédiction du risque afin d'améliorer la planification des actions (Degallier et al., 2010; Descloux et al., 2012; Lowe et al., 2011, 2013; Racloz et al., 2012).

1.2.2. Les approches possibles en modélisation

Quelle que soit la maladie étudiée, il existe deux grands types d'approches complémentaires dans les études de modélisation en épidémiologie. La première approche est statistique. Elle part des données observées et utilise des outils statistiques pour explorer la variabilité spatio-temporelle des cas d'une maladie et découvrir des liens statistiques globaux avec des facteurs explicatifs ou prédictifs (climatiques, socio-économiques…) (voir par exemple (Cazelles et al., 2005; Cummings et al., 2004; Descloux et al., 2012; Hales et al., 1999; Johansson et al., 2009a; Wu et al., 2009)).

La deuxième approche est mécanistique. Le but est de modéliser le cycle de transmission au niveau des individus afin de simuler des variables épidémiologiques (ou entomologiques pour les maladies vectorielles) qui sont ensuite comparées aux données observées afin de tester la validité des hypothèses sur lesquelles sont construits les modèles ou afin d'estimer certains paramètres de transmission (voir par exemple (Ferguson et al., 1999; Hay et al., 2000; Hopp and Foley, 2003; Wearing and Rohani, 2006; Smith et al., 2014a)).

1.2.3. Caractéristiques des études de modélisation

Les études de modélisation diffèrent par plusieurs caractéristiques :

- **La nature de l'étude** : les études peuvent être spatio-temporelles (Cummings et al., 2004; Thai et al., 2010; Teurlai et al., 2012; Cuong et al., 2013), temporelles (Cazelles et al., 2005; Descloux et al., 2012; Gharbi et al., 2011; Johansson et al., 2009b) ou spatiales (Carbajo et al., 2012; Hales et al., 2002; Honório et al., 2009; Hopp and Foley, 2003; Salje et al., 2012; Wu et al., 2009). Du fait de la difficulté à modéliser à la fois la dimension spatiale, la dimension temporelle et la nature multifactorielle de la dynamique spatiale des maladies, la plupart des études de la variabilité spatio-temporelle des données épidémiologiques n'incorporent aucun facteur de risque. Dans la plupart des cas, les études des déterminants de la dynamique spatiale des maladies distinguent la variabilité spatiale de la variabilité temporelle, et les études intégrant à la fois les dimensions spatiales, temporelles et différents facteurs de risque sont rares (Lowe et al., 2011) ;

- **L'échelle d'étude** : les études spatiales peuvent être locales (Honório et al., 2009; Salje et al., 2012), nationales (Carbajo et al., 2012; Wu et al., 2009) ou régionales (Hales et al., 2002; Hopp and Foley, 2003) tandis que les études temporelles peuvent analyser soit la variabilité inter-annuelle (Descloux et al., 2012; Johansson et al., 2009b) soit la variabilité saisonnière (Descloux et al., 2012; Gharbi et al., 2011) ;

- **La cible de la modélisation** : certaines analyses ont pour objet l'étude des facteurs liés à l'émergence d'une épidémie (Descloux et al., 2012) tandis que d'autres étudient les facteurs liés à l'intensité des épidémies (Johansson et al., 2009b) ;

Nous pourrions rajouter une quatrième caractéristique : **les objectifs de l'étude**, et distinguer les résultats des modèles dont le but est purement prédictif et visant à construire des systèmes d'alerte précoce (Descloux et al., 2012; Yu et al., 2011) de ceux des modèles à visée explicative (Descloux et al., 2012; Hopp and Foley, 2001; Wearing and Rohani, 2006), mais cette distinction est parfois difficile à effectuer.

Contrairement à ce qui est couramment fait dans les nombreuses revues bibliographiques qui trient les études sur la dynamique spatiale de la dengue par facteur de risque (Halstead, 2008; Kyle and Harris, 2008; Morin et al., 2013), nous avons décidé de trier les différentes études en fonction de la nature et de l'échelle de chacune des études.

1.2.4. Description de la distribution spatio-temporelle des cas de dengue

A l'échelle du globe, la dengue survient essentiellement dans les zones de distribution des vecteurs, *i.e.* en zone inter-tropicale (Bhatt et al., 2013; WHO, 2009). L'Asie est le continent le plus touché, avec près de 70 % du total des infections (Bhatt et al., 2013).

Sur le plan temporel, la dynamique de la dengue dépend des régions du Monde étudiées. En Asie (Cazelles et al., 2005; Cummings et al., 2004; Huy et al., 2010; Vong et al., 2010; Zhang et al., 2005), la dengue se présente sous une forme hyper-endémique (*i.e.* avec co-circulation des quatre sérotypes). Des épidémies surviennent tous les ans, et il n'y a jamais d'arrêt de la circulation virale. Tous les 3 à 5 ans, des épidémies de plus forte ampleur sont enregistrées, souvent concomitantes avec un changement de sérotype (Cummings et al., 2004; Duong et al., 2013a; Vong et al., 2010). En Thaïlande, un autre cycle de 10 ans a été enregistré (Zhang et al., 2005).

En Amérique Latine, entre 1970 et 2000, la dengue survenait sous la forme de vagues épidémiques régionales dues à l'émergence et à la diffusion d'un sérotype dominant sur l'ensemble du continent (Gubler, 2004; Guzman and Kouri, 2003). A partir des années 2000, la dynamique de la dengue s'est rapprochée de ce qui est observé en Asie du Sud-Est, avec des épidémies annuelles, et une co-circulation de plusieurs sérotypes (Chowell et al., 2008; Gubler, 2004; Guzman and Kouri, 2003).

Dans le Pacifique, depuis, 1971, la dengue survient sous la forme de vagues épidémiques dues à l'émergence et à la diffusion d'un sérotype dominant (Kiedrzynski et al., 1998; Singh et al., 2005). La co-circulation de différents sérotypes au sein d'un même pays est rare.

Quelle que soit la région du monde étudiée, les épidémies annuelles de dengue présentent souvent une saisonnalité marquée.

A l'échelle d'un pays, la dengue a longtemps été considérée comme une maladie essentiellement urbaine (Cummings et al., 2004), mais récemment, des études ont montré une charge plus importante de la

maladie en zone rurale par rapport aux zones urbaines au Vietnam (Cuong et al., 2013) et au Cambodge (Vong et al., 2010)

A des échelles très locales, comme celle des villes, les foyers infectieux de dengue sont très localisés dans le temps (entre 5 et 17 jours d'intervalle) et dans l'espace (taille des foyers infectieux entre 200 m et 800 m) (Aldstadt et al., 2012; Honório et al., 2009; Liebman et al., 2012; Mammen Jr et al., 2008; Morrison et al., 1998; Rabaa et al., 2013; Salje et al., 2012; Tran et al., 2004; Vazquez-Prokopec et al., 2010; Yoon et al., 2012).

Sur le plan spatio-temporel, des ondes progressives (onde temporelle se déplaçant dans l'espace) d'incidence ont été mises en évidence. En Thaïlande, la propagation de la maladie se fait vers les provinces à partir de la capitale Bangkok selon un cycle de 2-3 ans (Cummings et al., 2004). Au Cambodge (Teurlai et al., 2012) et dans le Sud du Vietnam (Cuong et al., 2013; Thai et al., 2010) une propagation annuelle à partir de quelques zones rurales vers les villes majeures a été mise en évidence. Les zones de démarrage de l'épidémie annuelle sont systématiquement les mêmes au Cambodge, tandis qu'elles diffèrent d'une année à l'autre au Vietnam.

1.2.5. Facteurs influençant la distribution spatiale de la dengue

1.2.5.a. Echelle régionale

Qu'elles soient basées sur des modèles statistiques (Hales et al., 2002) ou sur des modèles biologiques de transmission (Degallier et al., 2010; Hopp and Foley, 2003; Jetten and Focks, 1997), il n'existe que très peu d'études des déterminants de la distribution spatiale de la dengue dans le monde, probablement du fait de l'absence de bases de données épidémiologiques régionales. Ces études n'ont considéré que les déterminants environnementaux (climat, occupation des sols), probablement également du fait de l'absence de bases de données régionales sur les autres facteurs possibles (type de virus, facteurs socio-économiques…).

Facteurs influençant la distribution spatiale des vecteurs

Une étude utilisant un lien statistique entre la distribution observée de sites de présence d'*Aedes albopictus* et des variables concernant le climat (température et précipitations) et l'occupation des sols (topographie et couverture en eau) a permis d'établir une carte mondiale des niches écologiques favorables à l'implantation d'*Aedes albopictus* (Benedict et al., 2007). Cette étude prévoit une implantation d'*Aedes albopictus* dans une zone plus étendue que la zone inter-tropicale, couvrant par exemple une majorité de l'Europe. L'utilisation de modèles biologiques permettant de modéliser les

populations d'*Aedes aegypti* à partir de variables uniquement climatiques (Focks et al., 1993a) ont également permis d'établir une carte mondiale de la distribution d'*Aedes aegypti* (Hopp and Foley, 2001). Cette étude ne prévoit pas l'implantation d'*Aedes aegypti* en Europe.

Facteurs influençant la distribution spatiale de la dengue

Hales a réussi à créer un modèle statistique d'absence/présence de la dengue dans le monde basé uniquement sur la pression de vapeur d'eau (lié à l'humidité) (Hales et al., 2002). Les autres études régionales utilisent un modèle biologique de la capacité vectorielle pour lier les variables climatiques (température et précipitations), la biologie du moustique et l'occurrence de la dengue dans le monde (Degallier et al., 2010; Hopp and Foley, 2003; Jetten and Focks, 1997; Patz et al., 1998). Ces études sont cependant basées uniquement sur des variables climatiques. Elles permettent d'établir les limites de répartition de la dengue dans le monde, mais l'absence d'intégration d'autres facteurs rend les prédictions incertaines.

L'influence des variables climatiques sur la distribution régionale des cas de dengue fait craindre une augmentation du nombre de pays touchés par la maladie dans le futur sous l'effet du changement climatique (Banu et al., 2011; Degallier et al., 2010; Githeko et al., 2000; Hopp and Foley, 2003; Jetten and Focks, 1997; McMichael et al., 2006; Patz et al., 1998)

Le facteur prépondérant de la dissémination de la dengue de pays infectés vers des pays indemnes semble être les mouvements humains de population (Kuno, 1995).

1.2.5.b. Echelle Nationale

Il n'y a, à notre connaissance, que cinq études des déterminants de la distribution spatiale de la dengue à l'échelle d'un pays entier, utilisant toutes une approche statistique (Carbajo et al., 2012; Hu et al., 2011; Lowe et al., 2011; Tipayamongkholgul and Lisakulruk, 2011; Wu et al., 2009). Les facteurs étudiés comprennent des facteurs climatiques, socio-économiques, géographiques ou démographiques. L'influence de la température et de facteurs démographiques sur la transmission de la dengue est cohérente entre toutes les études : globalement, à l'échelle d'un pays, la transmission de la dengue est accrue dans les zones où la température est plus élevée, où le degré d'urbanisation est fort, et où la densité locale de population est importante. Les deux études ayant incorporé des facteurs socio-économiques ne mettent en évidence aucun lien entre ces facteurs et le degré de transmission de la dengue (Hu et al., 2011; Wu et al., 2009). Le rôle des précipitations est plus controversé : une seule étude montre une association statistique positive forte entre les précipitations et les taux d'incidence de la dengue (Lowe et al., 2011), les autres études ne montrent qu'une association statistique faible (Hu et al., 2011) ou inexistante (Carbajo et al., 2012; Wu et al., 2009).

1.2.5.c. Echelle locale

Facteurs influençant la distribution spatiale du vecteur

A des échelles très locales, les facteurs de risque de la présence du vecteur de la dengue ou de densités vectorielles accrues sont :

- la présence de gîtes larvaires comme des containers d'eau, des sous-pots de plantes, des boîtes abandonnées ou encore des pneus (Estallo et al., 2013; Favier et al., 2006a; Seidahmed et al., 2012; Thammapalo et al., 2005) ;
- le comportement humain, comme certaines pratiques religieuses qui utilisent des vases, ou l'absence de mesures préventives, l'absence de moustiquaires aux fenêtres (Spiegel et al., 2007; Thammapalo et al., 2005) ;
- certains facteurs socio-économiques, comme être au chômage, ou appartenir à une classe socio-économique basse (Seidahmed et al., 2012; Spiegel et al., 2007) ;
- certains facteurs environnementaux, comme la proximité de plantations, ou de forêts (Sarfraz et al., 2012; Thammapalo et al., 2005)

Lien entre densité vectorielle et distribution spatiale de la dengue

Cependant, les études de l'influence de la densité vectorielle sur la distribution spatiale des cas de dengue à des échelles très locales sont contradictoires : certaines études montrent une corrélation positive (Mammen Jr et al., 2008; Pessanha et al., 2012; Yoon et al., 2012), d'autres ne permettent de faire aucun lien entre les facteurs entomologiques et la maladie (Honório et al., 2009), et d'autres montrent une corrélation négative (Seidahmed et al., 2012).

Facteurs influençant la distribution spatiale de la dengue

L'influence des facteurs socio-économiques sur la distribution spatiale de la dengue à des échelles très locales semble également varier en fonction des études (Costa et al., 2013; Johansson et al., 2009a; Penna and Lucia, 2004).

Les facteurs explicatifs les plus robustes de la distribution spatiale des foyers infectieux de dengue au sein d'une ville semblent être les schémas de mouvement humain (Adams and Kapan, 2009; Barmak et al., 2011; Honório et al., 2009; Liebman et al., 2012; Penna and Lucia, 2004; Rabaa et al., 2013; Reiner et al., 2014; Stoddard et al., 2013; Vazquez-Prokopec et al., 2010). Les lieux fréquentés de manière régulière par un grand nombre de personnes, comme les centres commerciaux, paraissent associés à un risque élevé de transmission. Un autre facteur robuste semble être le comportement humain

(notamment l'entretien des jardins) et la mise en place de mesures individuelles de protection (Bohra and Andrianasolo, 2001; Favier et al., 2006a).

Récemment, une étude statistique a mis en évidence l'influence de l'immunité de population sur la distribution spatiale des cas de dengue à des échelles très fines (Salje et al., 2012).

1.2.6. Facteurs influençant la distribution temporelle de la dengue

1.2.6.a. Facteurs liés à la variabilité saisonnière de la dengue

De nombreuses études empiriques ont montré que les variations saisonnières des taux d'incidence de la dengue sont liées aux variations saisonnières des températures, des précipitations et de l'humidité relative (Banu et al., 2014; Cuong et al., 2013; Gharbi et al., 2011; Huang et al., 2013a), cependant, de nombreuses exceptions ont également été notées

Johansson (Johansson et al., 2009a), après avoir pris en compte cet effet saisonnier, a montré un lien consistant entre élévation mensuelle de température ou de précipitations et élévation mensuelle des taux d'incidence de la dengue à Puerto Rico.

Le même type de lien saisonnier peut être observé sur le terrain entre les indices servant à suivre les populations de vecteurs dans le temps et les taux d'incidence de la dengue (Dibo et al., 2008; Fouque et al., 2006; Pham et al., 2011).

1.2.6.b. Facteurs liés à la variabilité inter-annuelle de la dengue

En plus du cycle saisonnier, la dengue présente souvent des cycles multi-annuels dus, en Asie, à l'émergence, certaines années, d'épidémies de plus forte ampleur submergeant les systèmes de santé (Cazelles et al., 2005; Cummings et al., 2004; Zhang et al., 2005). Les causes de la survenue de ces épidémies de grande ampleur ne sont pas encore comprises. D'autre part en Asie où les 4 sérotypes circulent, les épidémies dues à chacun des sérotypes se succèdent dans le temps et ne sont pas en phase les unes avec les autres. L'enjeu de la modélisation de la variabilité inter-annuelle de la dengue est de démêler l'effet des facteurs intrinsèques liés à l'hôte, au virus ou au vecteur, de l'effet des facteurs extrinsèques liés à l'environnement (Koelle and Pascual, 2004; Rohani, 2009).

Facteurs intrinsèques : immunité de population

Plusieurs études utilisant des modèles de dynamique des populations pour simuler l'évolution des taux d'incidence dans le temps (Anderson and May, 1999) ont permis d'explorer l'effet des phénomènes immunitaires de l'hôte sur la dynamique inter-annuelle de la dengue. Cependant ces études sont contradictoires : certaines montrent que le phénomène immunitaire en jeu est l'existence d'anticorps facilitants (Ferguson et al., 1999; Cummings et al., 2005), tandis que d'autres montrent l'importance de l'existence d'une immunité croisée transitoire entre sérotypes hétérologues (i.e. différents) (Adams et al., 2006; Hay et al., 2000; Reich et al., 2013; Wearing and Rohani, 2006), et l'absence d'influence du phénomène d'anticorps facilitants (Wearing and Rohani, 2006).

Facteurs intrinsèques : variabilité génétique du virus

Plusieurs études en Asie du Sud-Est ont montré que la variabilité génétique intra-sérotypique des virus de la dengue suivait une dynamique interannuelle complexe, avec des successions d'émergence et d'extinction de nouvelles lignées génétiques au cours du temps. Ces études ont montré que la diversité génétique d'un sérotype donnée augmente avec l'intensité de circulation de ce dernier (Duong et al., 2013a, 2013b; Zhang et al., 2005). Certaines études ont montré que l'émergence d'un sérotype dominant dans les pays hyper-endémiques pouvait faire suite à un changement de lignée au sein de ce sérotype (Lee et al., 2010). Une étude en laboratoire a montré qu'un changement de génotype survenu au Vietnam au sein du sérotype 2 n'étais pas lié à une augmentation de la compétence vectorielle, mais à une augmentation des titres viraux chez les patients virémiques, donnant un avantage évolutif au génotype remplaçant (Ty Hang et al., 2010).

Facteurs extrinsèques

Le lien entre variables météorologiques locales et taux d'incidence de la dengue à l'échelle inter-annuelle est difficile à évaluer, car peu d'études utilisent une méthodologie permettant de distinguer l'effet saisonnier de l'effet inter-annuel. Lorsqu'il est montré, ce lien semble dépendre de la localisation géographique de l'étude. Au Vietnam, l'analyse des taux d'incidence annuels de la dengue a montré que l'augmentation régulière de la température et des précipitations permet d'expliquer une tendance à l'augmentation de la sévérité des épidémies annuelles. Cependant les variables climatiques ne permettent pas d'expliquer la survenue de deux épidémies de grande ampleur au cours de la période d'étude (Cuong et al., 2011). A Puerto Rico, les précipitations ont été identifiées comme un facteur de variabilité inter-annuelle des taux d'incidence de la dengue, mais pas la température, tandis qu'en Thaïlande, c'est l'inverse. Au Mexique, aucune variable climatique locale n'est statistiquement associée aux variations inter-annuelles de la dengue (Johansson et al., 2009b).

Le phénomène El Niño est un phénomène climatique survenant dans le Pacifique équatorial et ayant une influence considérable sur les variations inter-annuelles du climat dans plusieurs régions du Monde.

Plusieurs études ont étudié l'influence d'El Niño sur la survenue d'épidémies annuelles de dengue (Gagnon et al., 2001) ou sur les variations inter-annuelles des taux d'incidence de la dengue (Cazelles et al., 2005; Hales et al., 1999; Johansson et al., 2009b). L'intérêt de ces études réside dans le fait que le phénomène El Niño est prévisible plusieurs mois à l'avance. Ainsi la mise en évidence d'un lien entre ce phénomène et un risque accru de circulation virale permettrait d'anticiper le risque épidémique (Rohani, 2009). L'influence d'El Niño sur la dengue semble varier d'un pays à l'autre, notamment parce que les conséquences d'El Niño sur les variables climatiques locales ne sont pas les mêmes partout. Hales a montré que dans le Pacifique, sur un plan régional, le nombre de pays touchés par une épidémie une année donnée est partiellement dépendant de ce phénomène (Hales et al., 1996). En désagrégeant les données par pays, les coefficients de corrélation entre les indices temporels de suivi d'El Niño et les incidences annuelles de la dengue varient de -0,18 à 0,57 en fonction des pays, et El Niño n'est positivement corrélé aux incidences annuelles de la dengue que dans 5 pays (Tonga, Fiji, Tokelau, Samoa et aux Samoa américaines) sur les 14 étudiés (Hales et al., 1999). D'autres études ont montré une association transitoire entre El Niño et les taux d'incidence de la dengue sur des cycles de 2 à 3 ans en Thaïlande (Cazelles et al., 2005; Johansson et al., 2009b), à Puerto Rico, mais pas au Mexique (Johansson et al., 2009b). Enfin, en Indonésie, au Surinam et en Guyane Française, la survenue d'épidémies une année donnée est plus fréquente lors des évènements El Niño.

Il est supposé que le lien statistique observé entre El Niño et la dengue s'explique par l'influence de ce phénomène sur les conditions climatiques locales, et par voie de conséquence sur le cycle de transmission (Hales et al., 1996), cependant les résultats de Johansson (Johansson et al., 2009b), qui a cherché à tester cette hypothèse, laissent perplexes : en effet, par exemple à Puerto Rico, El Niño influence uniquement les températures, et non les précipitations, or les variations inter-annuelles d'incidence de la dengue sont corrélées uniquement avec les précipitations et pas avec les températures.

1.2.7. Contraintes des études de modélisation

Les caractéristiques des différentes études de modélisation (nature, échelle et cible) imposent des contraintes à la fois en termes de modèles à utiliser, et en termes de facteur étudiés.

Contraintes sur les modèles utilisés

La cible de modélisation (émergence *vs* sévérité) et la nature de l'étude (spatiale *vs* temporelle) imposent des contraintes sur **les modèles utilisés**. En particulier, en ce qui concerne la cible de la modélisation, l'étude de l'intensité de l'épidémie s'appuie souvent sur le calcul des taux d'incidence, ou de valeurs similaires, comme les taux standard de morbidité (Elliott et al., 2001), qui sont toutes des variables quantitatives continues. Les modèles pour étudier les facteurs liés à l'intensité d'une épidémie sont donc des modèles de régression. L'étude de l'émergence des épidémies nécessite quant à elle d'utiliser un critère de définition d'une épidémie ou d'un foyer épidémique, ce qui peut devenir problématique,

puisqu'il n'existe pas de définition unique de ce qu'est une épidémie (voir Costagliola et al., 1991; Coutinhoa et al., 2006; Cullen et al., 1984; Farrington et al., 1996; Heisterkamp et al., 2006; Held et al., 2005; Höhle and Paul, 2008; Höhle et al., 2009; Rossi et al., 1999; Serfling, 1963; Stroup et al., 1989; Turbelin et al., 2009 pour quelques exemples de définition de seuils épidémiques). Dans l'étude des phénomènes d'émergence épidémique, la variable réponse est dichotomique (une entité spatiale ou une entité temporelle, *i.e.* un individu statistique peut être soit épidémique soit non épidémique), et les modèles utilisés sont des modèles de classification.

Contraintes sur les facteurs étudiés

La nature et l'échelle d'étude quant à elles, contraignent fortement le type de facteur étudié.

Par exemple, la plupart des études temporelles statistiques n'incorporent que des variables climatiques (Cazelles et al., 2005; Descloux et al., 2012; Johansson et al., 2009a; Pinto et al., 2011) ou des indices de suivi entomologique (Descloux et al., 2012), probablement du fait du temps nécessaire à la compilation ou à la création de séries temporelles longues concernant d'autre facteurs (socio-économiques ou démographiques par exemple). Toujours sur le plan temporel, du fait de l'absence de données de séroprévalence sous forme de séries temporelles régulières, l'influence de l'immunité de population sur la dynamique de la dengue est essentiellement étudiée via des modèles mécanistes (Hay et al., 2000; Wearing and Rohani, 2006) dans lesquels il est difficile d'incorporer d'autres variables.

Sur le plan spatial, comme il n'existe en général qu'un nombre restreint de stations météorologiques par ville, le climat n'est jamais étudié à des échelles locales, alors qu'il régule fortement les cycles de vie des moustiques. De même, toujours sur le plan spatial, il est extrêmement difficile d'évaluer le rôle des facteurs liés à l'entomologie. En effet, il n'existe pas de données entomologiques à des échelles autres que locales et les études nationales ou régionales tentant d'évaluer l'impact des facteurs entomologiques sur la dengue passent par des modèles mécanistes de simulation des populations de moustique (Hopp and Foley, 2003). A l'échelle locale, il est possible d'avoir des séries temporelles d'indices de suivi entomologique, par contre, l'effort d'échantillonnage à mener pour obtenir des données entomologiques pouvant être utilisées spatialement est considérable, et de fait n'est souvent réalisé que sur des périodes de temps courtes, rarement plus de deux ou trois ans (Honório et al., 2009; Mammen Jr et al., 2008; Seidahmed et al., 2012; Yoon et al., 2012).

Impact sur notre compréhension de la maladie

Ces contraintes liées à la nature ou à l'échelle des études, en limitant le nombre de facteurs pouvant être incorporés dans chacune des études, rendent nos connaissances de l'épidémiologie de la dengue parcellaires, et, du fait de la nature multi-factorielle de la dynamique spatiale de la dengue, compliquent notre compréhension de l'épidémiologie de cette maladie.

Trier les différentes études existantes en fonction de la nature et de l'échelle d'étude nous a permis de mettre en évidence que les résultats apparemment contradictoires entre études de même nature n'apparaissent que dans certains types d'études.

Ainsi, les liens entre facteurs explicatifs et risque infectieux qui semblent constants et robustes d'une étude à l'autre sont :

- L'influence de la température sur la dynamique saisonnière des épidémies de dengue (Descloux et al., 2012; Hii et al., 2009; Huang et al., 2013b; Lu et al., 2009; Pinto et al., 2011; Shang et al., 2010), et sur la structure spatiale de la sévérité des épidémies à des échelles au moins nationales (Carbajo et al., 2012; Hopp and Foley, 2003; Hu et al., 2011; Lowe et al., 2011; Wu et al., 2009);
- L'influence des mouvements humains et de la structuration des contacts sociaux sur la structure spatiale de la dengue à des échelles locales (Adams and Kapan, 2009; Barmak et al., 2011; Honório et al., 2009; Liebman et al., 2012; Penna and Lucia, 2004; Rabaa et al., 2013; Reiner et al., 2014; Stoddard et al., 2013; Vazquez-Prokopec et al., 2010) ;
- L'influence de facteurs immunitaires sur la dynamique inter-annuelle et l'alternance des 4 sérotypes en zone hyper-endémique (Adams et al., 2006; Cummings et al., 2005; Ferguson et al., 1999; Hay et al., 2000; Reich et al., 2013; Wearing and Rohani, 2006).

Les liens qui semblent inconstants sont :

- L'influence des précipitations sur la dynamique (inter-annuelle ou saisonnière) et sur la structure spatiale de la dengue quelle que soit l'échelle d'étude ;
- L'influence des températures sur la dynamique inter-annuelle de la dengue (Descloux et al., 2012; Hales et al., 1999; Johansson et al., 2009b), ou sur la dynamique à court terme après avoir retiré l'effet saisonnier (Johansson et al., 2009a) ;

L'influence des facteurs socio-économiques (Bohra and Andrianasolo, 2001; Costa et al., 2013; Penna and Lucia, 2004; Reiter et al., 2003 et paragraphe 5.2, page 267) et entomologiques (Honório et al., 2009; Mammen Jr et al., 2008; Pessanha et al., 2012) sur la structure spatiale de la dengue à des échelles très locales.

1.3. Conclusion

Malgré les nombreuses études réalisées sur le sujet, notre compréhension de l'épidémiologie de la dengue et des facteurs influençant le cycle de transmission reste limitée. Si certains facteurs prépondérants peuvent être dégagés, comme l'importance de l'immunité des hôtes dans la dynamique inter-annuelle observée en Asie, l'influence des facteurs climatiques sur la dynamique saisonnière ou

encore l'importance des mouvements humains dans la structuration spatiale des épidémies à des échelles locales, des questions persistent, comme celle de l'influence des facteurs climatiques sur la dynamique inter-annuelle des épidémies de dengue, ou celle de l'absence de lien apparent entre densités vectorielles et structuration spatiale de la dengue à des échelles locales, ou encore la question du mécanisme permettant de relier les études statistiques réalisées sur des populations d'individus aux interactions moléculaires entre hôtes, vecteurs et virus. Nous espérons, au cours de cette thèse, par l'analyse et la modélisation de données épidémiologiques et environnementales à différentes échelles spatiales et temporelles, aider à éclaircir certaines de ces questions. En effet, il est rare de disposer de jeux de données aussi conséquents, comprenant des données épidémiologiques fiables, ainsi que des données climatiques, géographiques et entomologiques. Il est tout aussi rare de pouvoir réaliser ces études à des échelles à la fois régionales, nationales et locales.

CHAPITRE 2 : CONTEXTE CALEDONIEN, DONNEES ET MODELES UTILISES

2.1. Contexte calédonien 71

2.1.1. Contexte géographique et historique 71
2.1.1.a. Le Pacifique 71
Situation générale 71
Climat 71
2.1.1.b. La Nouvelle-Calédonie 73
Climat 73
Situation générale 73
Histoire, peuplement et conséquences épidémiologiques 76
2.1.2. Contexte épidémiologique 76
2.1.2.a. Histoire naturelle de la dengue en Nouvelle-Calédonie et dans le Pacifique 76
2.1.2.b. Organisation des systèmes de surveillance de la santé dans le Pacifique 77
2.1.2.c. Organisation du système de santé en Nouvelle-Calédonie 78
2.1.2.d. Organisation du système de surveillance de la dengue en Nouvelle-Calédonie 78
Modalités de collecte et de transmission des données 78
Facteurs pouvant influencer le nombre de cas enregistrés 79
Définition des cas 79
Standardisation spatiale et temporelle du réseau 80
2.1.2.e. Organisation de la lutte contre les maladies vectorielles en Nouvelle-Calédonie 81
Lutte péri-focale 82
Actions systématiques de prévention et de communication 82
LAV en dehors du Grand Nouméa 83

2.2. Les données analysées 83

2.2.1. Echelle d'étude et résolution 83
2.2.2. Nature des données 84
2.2.3. Les données épidémiologiques 86
2.2.3.a. Données dans le Pacifique (1971- 2009) 86
Résolution spatio-temporelle des données épidémiologiques « Pacifique » 86
2.2.3.b. Données en Nouvelle-Calédonie 88
Données agrégées : résolution spatiale territoire et résolution temporelle mensuelle (janvier 1971 – janvier 2014) 88
Données brutes géo-localisées : (01/01/1995 – 30/09/2012) 88
Résolution spatio-temporelle minimale des données brutes 89
Géo-référencement des cas en coordonnées lon/lat sur Nouméa 90
Encadré 2 : algorithme de géo-référencement des cas 93
2.2.4. Les données entomologiques 94
2.2.4.a. Données entomologiques dans le Pacifique 94

2.2.4.b. Données entomologiques à Nouméa (mars 2000 – septembre 2011) 94
Protocole d'échantillonnage..... 94
2.2.4.c. Résolution spatiale et temporelle des données entomologiques de Nouméa 95
Résolution spatiale secteurs – Résolution temporelle mensuelle..... 95
Résolution spatiale grappes – Résolution temporelle annuelle 95
2.2.5. Les données climatiques ---------- 97
2.2.5.a. Données observées en Nouvelle-Calédonie..... 97
Données de la station de Nouméa (1er janvier 1971- 31 décembre 2009) 97
Données sur l'ensemble du territoire (1er janvier 1995 – 31 décembre 2010). 97
Données de précipitations issues d'un krigeage des données observées (résolution spatiale 1 km – carte moyenne de 2000 à 2011) 98
2.2.5.b. Données issues de simulations numériques autour de la Nouvelle-Calédonie 99
Résolution spatio-temporelle des données climatiques simulées 100
Validation des données simulées..... 100
2.2.5.c. Données de projections climatiques en Nouvelle-Calédonie à l'horizon 2100 (1971-2100) 103
Séries temporelles non corrigées (Nouméa)..... 103
Séries temporelles corrigées (Nouméa, Touho, Bourail, Lifou et Koné). 104
Etude des augmentations moyennes de température en Nouvelle-Calédonie 104
2.2.5.d. Données ENSO 105
2.2.6. Les données démographiques et socio-économiques ---------- 106
2.2.6.a. Données des îles du Pacifique 106
2.2.6.b. Données de Nouvelle-Calédonie 106
Données de population (1971 – 2009) 106
Données d'âge par commune 106
Données socio-économiques 106
2.2.6.c. Données de Nouméa..... 107
2.2.7. Les données de télédétection---------- 107
2.2.8. Les données topologiques ---------- 108

2.3. Les outils d'analyse et de modélisation 110

2.3.1. Les approches possibles en modélisation---------- 110
2.3.1. Définition des cibles de modélisation ---------- 111
Indicateurs de sévérité des épidémies..... 111
Indicateurs de présence d'épidémie ou de foyer épidémique 112
2.3.2. Les études possibles lors de l'analyse statistique de champs spatio-temporels --------- 114
2.3.2.a. Etudes possibles..... 114
2.3.2.b. Etudes non abordées au cours de la thèse 115
2.3.3. Modélisation statistique au cours des études descriptives ---------- 115
2.3.3.a. Objectifs des études descriptives 115
2.3.3.b. Détermination des processus de propagation..... 115
2.3.3.c. Extraction de schémas spatiaux ou temporels récurrents 118
Périodicité et auto-corrélation des signaux temporels..... 118
Détection de cluster et auto-corrélation spatiale 119

2.3.4. Modélisation statistique au cours des études analytiques ---------- 120
2.3.4.a. Objectifs de l'étude 120
2.3.4.b. Etude des facteurs liés à la sévérité des épidémies 120
2.3.4.c. Etude des facteurs liés à l'émergence des épidémies 122
2.3.5. Modèles mécanistiques : modélisation des processus de transmission au niveau individuel 122
2.3.5.a. Modèles de dynamique des populations de type « SIR »..... 123
2.3.5.b. Modèles de capacité vectorielle 123
2.3.6. Décomposition orthogonale aux valeurs propres (« Empirical Orthogonal Functions ») 124
Notations et transformation des données 125
Calcul des EOF 127
Représentation géométrique pour p = 2..... 129
Interprétation d'une analyse EOF : 131
Applications et intérêts d'une analyse EOF..... 135
Limites de l'analyse EOF classique et extensions possibles (Hannachi et al., 2007) 136

2.1. Contexte calédonien

2.1.1. Contexte géographique et historique

2.1.1.a. Le Pacifique

Situation générale

L'Océan Pacifique est le plus vaste océan du monde, couvrant un tiers de la surface du globe. Il s'étend entre les latitudes 60° nord et 60° sud. Il comprend une quarantaine de pays et territoires insulaires, situés majoritairement entre 10° nord et 30° sud (Figure 2.1).

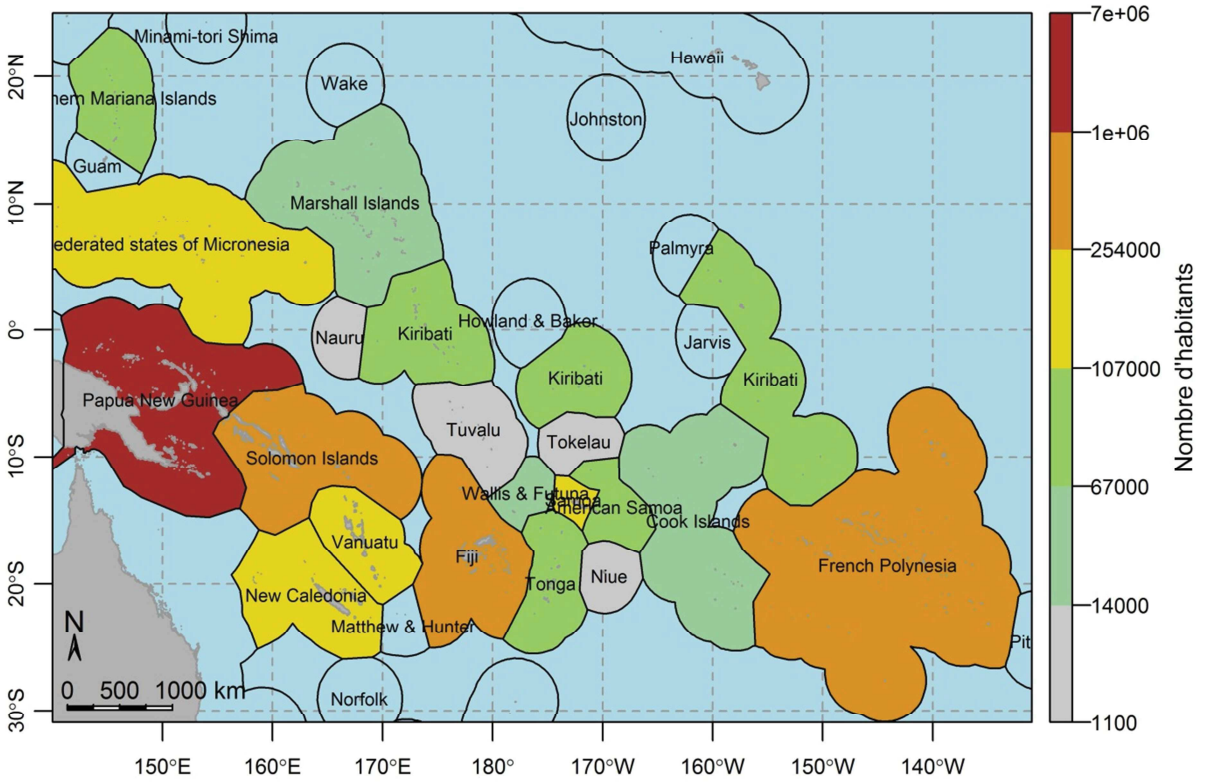


Figure 2.1 : Carte du Pacifique et population par pays. *Les pays sont entourés de leur zone économique exclusive (ZEE), dont la couleur correspond à la population du pays. Les pays dont la population est donnée sont les pays pour lesquels des données épidémiologiques sont disponibles.*

Climat

Le climat sur l'ensemble des îles du Pacifique est un climat tropical à sub-tropical. Il est régi par la présence d'anticyclones stables (Sinclair, 1996), et de deux zones de convergence : la Zone de Convergence Inter-Tropicale, et la Zone de Convergence du Pacifique Sud (SPCZ) (Lefèvre et al., 2010; Vincent, 1994; Vincent et al., 2011). La position géographique ainsi que l'intensité des anticyclones et des zones de convergence varient de manière saisonnière suivant la migration du soleil, et

varient également à d'autres échelles temporelles sous l'influence de divers phénomènes climatiques à large échelle spatiale comme le phénomène El Niño (ENSO, variations inter-annuelles) (Rasmusson and Carpenter, 1982), ou encore l'Oscillation de Madden-Julian (variations sur des périodes de 40 à 50 jours) (Madden and Julian, 1971). Le climat observé ainsi que l'impact des différents phénomènes climatiques peuvent être très variables d'un pays à l'autre, en raison de l'étendue spatiale couverte par l'ensemble des territoires.

La SPCZ est une zone de convergence des alizés d'orientation nord-ouest/sud-est s'étendant de la Papouasie Nouvelle-Guinée à approximativement 30°S/160°W (Figure 2.2). Elle est caractérisée par la présence quasi permanente d'une couverture nuageuse et de fortes précipitations (Vincent, 1994; Vincent et al., 2011).

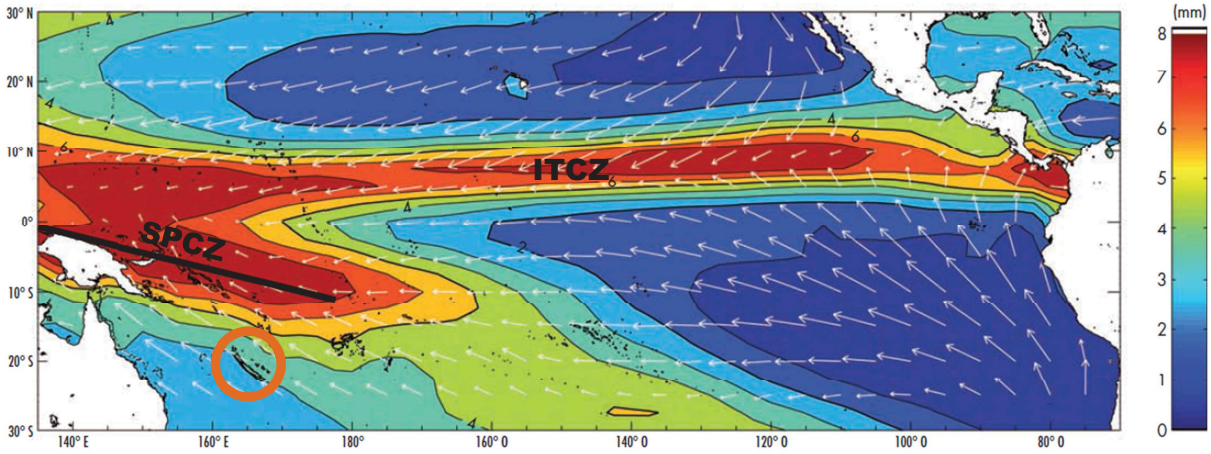


Figure 2.2 : Positions des zones de convergence de la région Indo-Pacifique. *La figure présente une carte des précipitations quotidiennes moyennes de 1979 à 2000 dans le Pacifique (source : Christophe Menkès et Atlas de Nouvelle-Calédonie, IRD (Bonvallot et al., 2013)). SPCZ : Zone de Convergence du Pacifique Sud ; ITCZ : Zone de Convergence Inter-Tropicale. La position de la Nouvelle-Calédonie est signalée par le cercle orange. Les flèches représentent les vents dominants.*

La position de la SPCZ varie de manière saisonnière : durant l'été austral, saison pendant laquelle elle est la plus marquée, elle se déplace vers le Sud. En juillet, durant l'hiver austral, la SPCZ s'atténue, et sa position remonte vers le Nord, laissant la place à l'anticyclone situé au-dessus de l'Australie et au Nord de la Nouvelle-Zélande (Vincent, 1994).

ENSO est le nom donné à un phénomène climatique survenant de manière inter-annuelle à des échelles spatiales de bassin. Au cours de ce phénomène, on observe des variations de pression de surface sur l'ensemble du Pacifique équatorial entraînant soit un affaiblissement des alizés par rapport à la moyenne (phénomène « El Niño »), soit un renforcement des alizés (phénomène « La Niña »). Sous l'effet des alizés, l'eau chaude de surface du Pacifique Sud est normalement entraînée vers l'ouest, créant la « warm pool » du Pacifique. Durant El Niño, l'affaiblissement des alizés entraîne un déplacement des

masses d'eau chaude vers l'Est tandis que les eaux froides situées le long de la côte Péruvienne sont repoussées en profondeur. Ces déplacements ont des conséquences sur l'ensemble du climat de la planète : globalement, il fait plus chaud et humide dans le Pacifique Est, tandis qu'il fait plus froid et sec dans le Pacifique Ouest (comme en Nouvelle-Calédonie). Durant La Niña, le phénomène inverse se produit. ENSO a une période (dans le sens fréquentiel) privilégiée de 2 à 8 ans, et est en grande partie responsable des variations inter-annuelles du climat dans les îles équatoriales du Pacifique (Rasmusson and Carpenter, 1982), et dans une moindre mesure, dans les îles tropicales et sub-tropicales comme la Nouvelle-Calédonie.

2.1.1.b. La Nouvelle-Calédonie

Climat

Comme la plupart des îles du Pacifique Sud, le climat de la Nouvelle-Calédonie dépend essentiellement de la position de la SPCZ ainsi que de la position d'anticyclones et de dépressions mobiles.

Les variations saisonnières de la position de la SPCZ sont à l'origine des deux saisons observées : la saison chaude et humide s'étend de novembre à avril tandis que la saison fraîche et sèche s'étend de mai à octobre. Les variations inter-annuelles sont partiellement influencées par ENSO : la Nouvelle-Calédonie étant située à l'Ouest du Pacifique, il fait plus chaud et humide durant La Niña, et plus froid et sec durant El Niño.

La Nouvelle-Calédonie est une île allongée orientée dans le sens Nord-Ouest/Sud-Est. Sur le plan spatial, la présence d'une chaîne montagneuse culminant à 1628 m et séparant la côte Est de la côte Ouest entraîne une disparité de climat entre les deux côtes. En effet, les alizés chauds et humides venant du Sud-Est se heurtent à la paroi formée par les montagnes sur la côte Est, influençant le niveau et la distribution spatiale des précipitations sur le territoire (Lefèvre et al., 2010). Les précipitations varient ainsi de 800 mm/an pour certaines stations de la côte Ouest à 3200 mm/an sur la côte Est. Les températures peuvent descendre jusqu'à 10 °C durant la saison fraîche et s'élever jusqu'à plus de 35 °C durant l'été austral.

Situation générale

La Nouvelle Calédonie est un territoire français insulaire situé à ~1500 kilomètres à l'Est de l'Australie. Il comprend une île principale, la Grande Terre, d'environ 400 km de long sur 60 km de large, ainsi que cinq îles secondaires (Bélep au nord, Ouvéa, Maré et Lifou à l'est, et l'île des Pins au sud). La côte Ouest de la Grande Terre est séparée de la côte Est par une chaîne montagneuse culminant à 1 628 m d'altitude. Sur la grande terre, le réseau routier est constitué de deux routes territoriales longeant chacune des côtes, ainsi que de 4 routes transversales. Le reste du réseau routier est constitué de pistes, dont certaines ne sont praticables qu'en véhicules tout terrain.

Le territoire est divisé en 3 provinces (Nord, Sud et Province des Iles) et 33 communes (Figure 2.3), dont l'équivalent administratif métropolitain serait le département. Les communes de Nouméa, la capitale, du Mont-Dore et de Dumbéa sont regroupées sous le nom de « Grand Nouméa », et sont elle-mêmes divisées en 67 quartiers, 78 iris, et 1563 îlots (dont 41 quartiers, 46 iris et 868 îlots à Nouméa). Sur les 245 580 habitants recensés en 2009 (Institut de la statistique et des études économiques (ISEE), 2009), 147 365 habitent dans le grand Nouméa. Le reste de la population est réparti dans des villes d'environ 2000 habitants, le plus souvent côtières, ou dans des tribus (Figure 2.4).

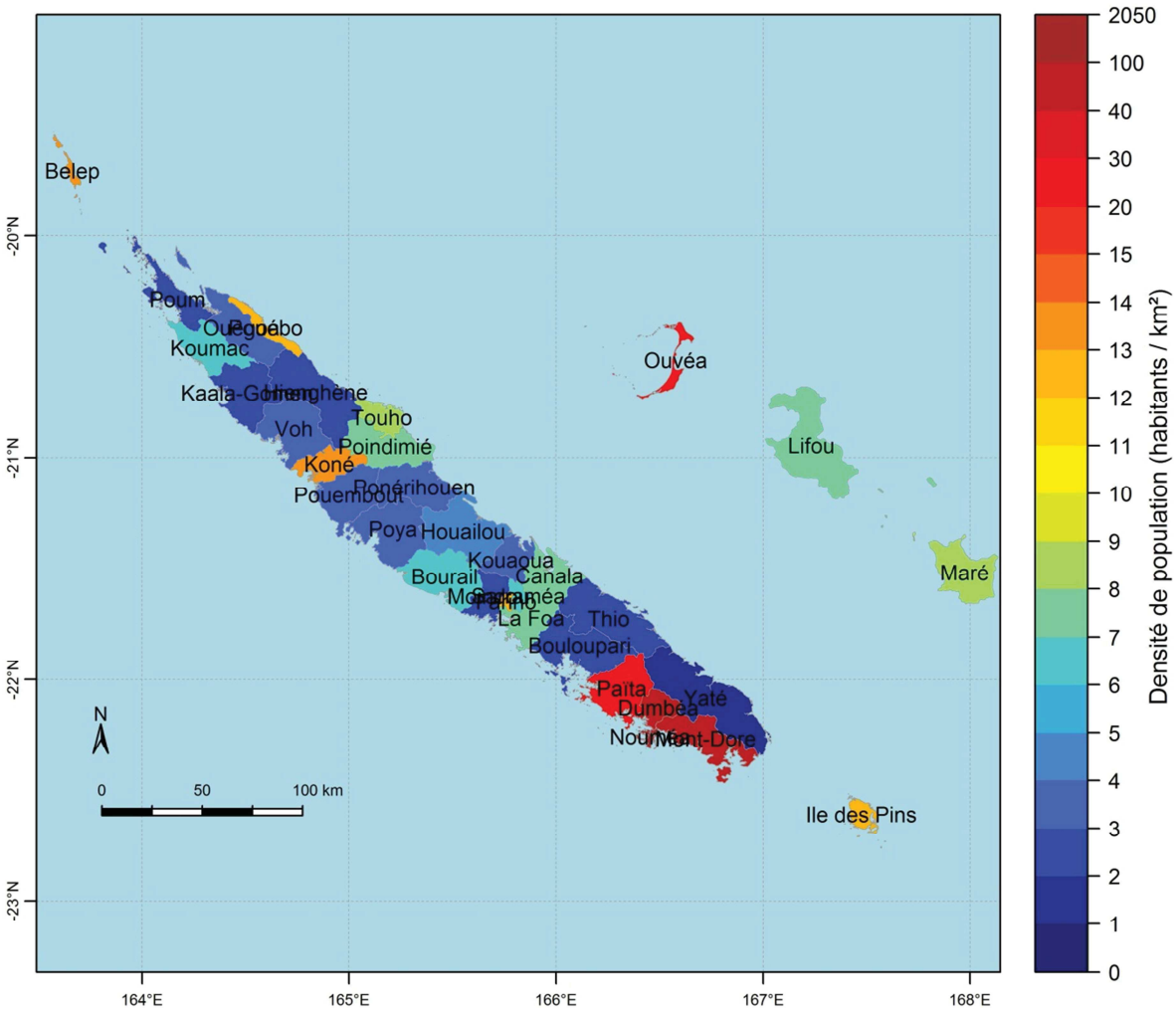


Figure 2.3 : Densité de population dans les communes de Nouvelle-Calédonie. *L'échelle de densité n'est pas linéaire.*

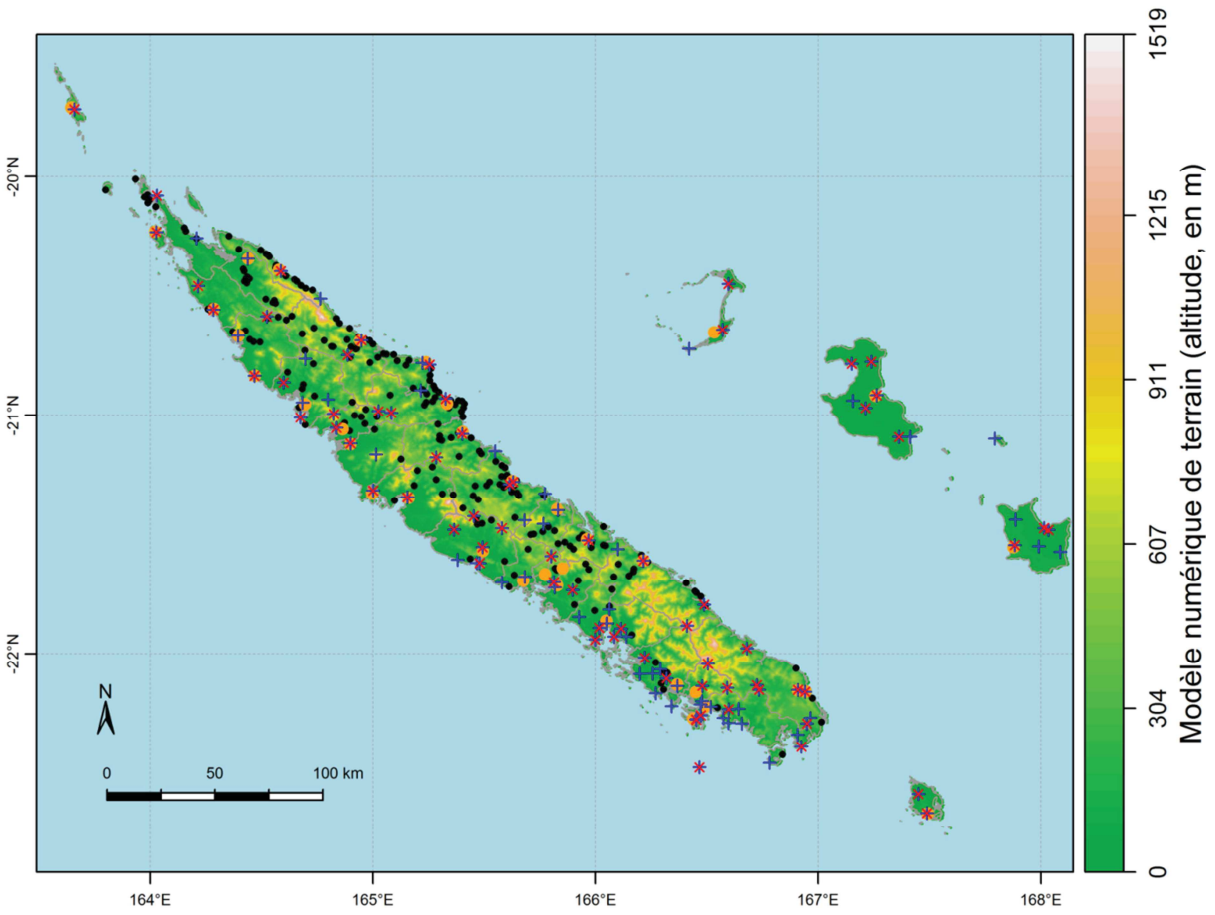


Figure 2.4 : Villes, tribus et stations météorologiques en Nouvelle-Calédonie. *Les villes sont représentées par des points orange, les tribus par des points noirs, les stations météorologiques par des croix rouges orientées en diagonale (67 stations mesurant la température) ou bleues orientées de manière droite (113 stations mesurant les précipitations), les « étoiles » correspondent donc à des stations où sont mesurées à la fois température et précipitations.*

Il y a 35 millions d'années, un phénomène géologique d'obduction survient : une couche de roches ultrabasiques située dans les profondeurs de l'enveloppe terrestre remonte en surface et recouvre le Sud de l'île (Collot et al., 2008). Ces roches sont riches en minéraux. Ainsi, la Nouvelle-Calédonie possède maintenant entre 20 et 25 % des réserves mondiales de nickel. Cette abondance de nickel a influencé et influence toujours actuellement les mouvements de population entre la Calédonie et le reste du monde, avec un impact potentiel sur les risques épidémiologiques. Suite à ce phénomène d'obduction ainsi qu'à d'autres phénomènes géologiques (Cluzel et al., 2001), les sols Calédoniens ont des propriétés physico-chimiques particulières, ce qui entraîne leur colonisation par des espèces végétales et animales endémiques, conduisant à la présence de plusieurs écosystèmes très particuliers dans l'ensemble du territoire (Heads, 2008). Une question récurrente est celle de savoir s'il existe un lien entre la particularité des écosystèmes calédoniens et l'absence remarquable de certaines espèces vectrices de maladie en Nouvelle-Calédonie, comme *Aedes albopictus*, ou les moustiques du genre *Anophele*, espèces qui sont pourtant présentes dans la plupart des îles du Pacifique (Figure 1.3).

Histoire, peuplement et conséquences épidémiologiques

La Nouvelle-Calédonie est originellement peuplée par le peuple kanak, dont la civilisation apparaît vers l'an 1000. En 1774, lors des grandes expéditions maritimes lancées par les Etats européens, James Cook découvre la Grande Terre qu'il surnomme « Nouvelle-Calédonie ». La France prend possession du pays en 1853. La Nouvelle-Calédonie reste alors une colonie française jusqu'en 1946, année au cours de laquelle elle devient un territoire d'Outre-Mer (Bonvallot et al., 2013). Suite à une guerre civile qui prendra fin en 1988 avec la signature des accords de Matignon, puis en 1998 la signature des accords de Nouméa (JORF n°121 du 27 mai 1998 page 8039), la Nouvelle-Calédonie est maintenue dans la France, mais avec un statut dont l'autonomie est augmentée.

Au cours de son histoire, la Nouvelle-Calédonie a connu de nombreux mouvements de population. Durant la seconde guerre mondiale, elle est une importante base militaire accueillant de nombreux soldats américains. Les mouvements de troupes à cette époque sont supposés être à l'origine, en 1942-1943, de la première épidémie de dengue sur le territoire dont le virus est identifié (DENV-1) (Fauran, 1996). A partir des années 1960, l'exploitation du nickel prend une place importante dans l'économie du pays, et de nombreuses vagues d'immigration vont se succéder, au grès des cours du nickel et des besoins de main d'œuvre pour les mines. Ainsi, la Nouvelle-Calédonie est maintenant caractérisée par une population mixte, composée de kanaks, de calédoniens non kanaks présents depuis plusieurs générations, de personnes d'origine asiatique, océanienne ou métropolitaine. Ces populations ont toutes un mode de vie différent ainsi que des activités propres les rendant plus ou moins à risque vis-à-vis de la dengue ou d'autre maladies. De plus, en particulier chez les kanaks, le rapport à la santé et à la maladie n'est pas le même que dans le monde occidental.

Sur le plan spatial, la séparation géographique entre les différentes populations est relativement nette : la côte Est est presque entièrement peuplée par des kanaks, vivant en tribu selon un mode de vie ancestral, tandis que la côte Ouest est peuplée essentiellement par les descendants des colons, qui ont monté des grandes propriétés agricoles ou fermières. La plupart des métropolitains arrivés récemment sont concentrés sur Nouméa. Le Nord et le Sud de la Grande Terre affichent également des disparités en termes de densité de population et de mode de vie.

2.1.2. Contexte épidémiologique

2.1.2.a. Histoire naturelle de la dengue en Nouvelle-Calédonie et dans le Pacifique

Dans le Pacifique, les premières épidémies de Dengue ont été rapportées à Hawaï en 1844, en Polynésie Française en 1852, aux îles Fiji et en Nouvelle-Calédonie en 1883 (Descloux, 2011; Perry, 1950). Durant la seconde guerre mondiale, les mouvements de troupe dans le Pacifique sont à l'origine d'une vague épidémique sur de nombreuses îles, et pour la première fois, le virus est isolé et le sérotype est identifié (DENV-1) (Steel et al., 2010). Entre les années 1950 et les années 1970, alors que la dengue

circule de manière permanente en Asie, seules deux épidémies ont été rapportées dans le Pacifique, toutes deux en Polynésie Française, en 1964 et 1969. En 1971, après plus de 30 ans d'absence (sauf les deux épidémies de Polynésie Française), la dengue émerge dans le Pacifique, sous la forme de vagues épidémiques régulières touchant un grand nombre d'îles et de territoires de la région, et au cours desquelles les premiers cas de DHF du Pacifique sont enregistrés (Kiedrzynski et al., 1998; Singh et al., 2005).

2.1.2.b. Organisation des systèmes de surveillance de la santé dans le Pacifique

Dans le Pacifique, chaque pays possède son propre système de surveillance, plus ou moins fonctionnel en fonction des ressources allouées au système. Les données de chacun des pays sont centralisées par l'Organisation Mondiale de la Santé dans le Pacifique. Le Secrétariat général de la Communauté du Pacifique (CPS), une organisation inter-gouvernementale au service du développement dans le Pacifique, anime également un réseau de surveillance régional, le Réseau Océanien de Surveillance de la Santé Publique (ROSSP), auquel est intégré un système de surveillance syndromique (Kool et al., 2012) dont les données sont accessibles sur internet via le serveur PacNet (http://www.spc.int/phs/PPHSN/Services/PacNet/intro.htm). Régulièrement, la CPS compile les données récoltées par les différents réseaux et systèmes de surveillance, pour produire des rapports régionaux, tous les 4 à 5 ans pour la dengue par exemple (Kiedrzynski et al., 1998; Singh et al., 2005).

Les Pays et Territoires Insulaires du Pacifique (PTIP) sont en général de petits pays, dont la population est restreinte (la moitié des PTIP a moins de 90 000 habitants), et dont les ressources financières et techniques sont limitées, sauf pour les PTIP rattachés à la France (Nouvelle-Calédonie, Polynésie Française et Wallis et Futuna). Dans les pays en développement, le fonctionnement des réseaux de surveillance, et notamment leur stabilité dans le temps, peut être altéré par un certain nombre de contraintes financières, politiques et techniques. Les ressources parfois faibles allouées aux agents en charge de la surveillance, l'absence de laboratoire de diagnostic, les difficultés techniques auxquelles doivent faire face les laboratoires existant (budget insuffisant, manque de personnel qualifié, difficulté d'approvisionnement en réactifs) sont autant de facteurs qui rendent difficiles la collecte et la transmission des données épidémiologiques sur le terrain. Dans le Pacifique, à ces contraintes il faut également ajouter l'isolement géographique, et l'absence parfois de réseaux de communication efficaces (Kool et al., 2012). Ces contraintes financières et techniques sont parfois aggravées par le contexte politique du pays qui conditionne les moyens et ressources allouées à la gestion de la santé (Bendali, 2006; Roger et al., 2004). Dans ce contexte, il est important de garder à l'esprit que les données de surveillance récoltées à l'échelle du Pacifique peuvent présenter des biais spatiaux et temporels.

2.1.2.c. Organisation du système de santé en Nouvelle-Calédonie

Il y a, en Nouvelle-Calédonie, 7 centres hospitaliers publics (2 en Province Nord, 5 à Nouméa), 3 cliniques privées (à Nouméa), 30 centres médicaux sociaux répartis sur l'ensemble du territoire et gérés au niveau provincial (5 dans la Province des îles, 15 dans la Province Nord, et 10 dans la Province Sud), pour une capacité totale de 711 lits d'hospitalisation installés en 2011. Toujours en 2011, il y avait sur le territoire 626 médecins libéraux (densité de 22,34 médecins pour 10 000 habitants) et 275 équivalents temps plein médecins salariés des structures publiques (DASS-NC, 2011). Les centres de soin sont donc accessibles à l'ensemble de la population.

Malgré les fortes disparités économiques, le système de couverture sociale de la Nouvelle-Calédonie permet aux plus démunis de bénéficier d'une aide médicale pouvant aller jusqu'à la gratuité des soins.

2.1.2.d. Organisation du système de surveillance de la dengue en Nouvelle-Calédonie

En Nouvelle-Calédonie, la dengue est une maladie à déclaration obligatoire depuis 1991. C'est-à-dire qu'un médecin suspectant un patient d'avoir contracté la dengue est dans l'obligation de déclarer le cas à la Direction des Affaires Sanitaires et Sociales de Nouvelle-Calédonie (DASS-NC).

Modalités de collecte et de transmission des données

Actuellement, le réseau de surveillance de la dengue en Nouvelle-Calédonie fonctionne sur un mode passif. Les données sont collectées par l'ensemble des médecins du territoire. Si un patient présente un tableau évocateur de la dengue, le médecin remplit une fiche de déclaration (voir Annexe 2) et effectue ou non un prélèvement. Si aucun prélèvement n'a été effectué, la fiche de déclaration est envoyée directement à la DASS-NC, qui enregistre les informations relatives au cas dans une base de données territoriale. Si un prélèvement est effectué, il est envoyé à l'Institut Pasteur de Nouvelle-Calédonie, qui effectue les tests diagnostiques nécessaires, et transmet la fiche de prélèvement et le résultat des tests à la DASS-NC si le résultat du test est positif. La base de données de la DASS-NC analysée dans cette thèse enregistre donc des cas cliniques et des cas confirmés par un test diagnostic.

Il existe, en parallèle, un réseau sentinelle, composé en 2007 de 8 médecins à Nouméa, 9 médecins en Province Sud ailleurs qu'à Nouméa, 3 médecins en Province Nord, et 3 médecins en Province des îles. Depuis 2013, ces médecins, en plus de la déclaration des cas dans le cadre de la surveillance passive, envoient un rapport hebdomadaire du nombre de consultations effectuées par jour ainsi que des informations sur les consultations concernant les arboviroses.

Facteurs pouvant influencer le nombre de cas enregistrés

Il est possible qu'en 1991 le changement de statut de la dengue en maladie à déclaration obligatoire ait entraîné une augmentation du nombre de cas déclarés ne correspondant pas nécessairement à une augmentation du nombre de cas effectifs sur le terrain. Les niveaux d'incidence d'avant 1991 et d'après 1991 paraissent donc difficilement comparables.

Un autre facteur pouvant influencer le nombre de cas enregistrés par le réseau de surveillance est la prise en charge financière des consultations et analyses liées à la maladie. Cette prise en charge varie au cours du temps en Nouvelle-Calédonie. En période inter-épidémique, ces frais sont pris en charge par le système classique de sécurité sociale, et il est possible que des personnes atteintes par la dengue mais dont la couverture sociale est insuffisante ne consultent pas leur médecin. En période épidémique, ces frais sont pris en charge à 100 % directement par la DASS-NC, quelle que soit la couverture sociale du patient. Depuis 2008, la prise en charge totale par la DASS-NC des frais de consultation et d'analyse liés à la dengue est effective quel que soit le médecin déclarant. Avant 2008, cette prise en charge n'était assurée que pour un nombre restreint de médecins déclarants appartenant au réseau sentinelle. L'impact de ce changement en 2008 sur les données de surveillance doit être évalué.

Hors de Nouméa, où la plupart des personnes bénéficient d'une aide médicale totale, en 2008, l'augmentation du nombre de médecins pour lesquelles la prise en charge est assurée à 100 % n'a probablement pas eu d'impact sur le nombre de cas déclarés et confirmés à la DASS-NC. A Nouméa, par contre, où les couvertures sociales diffèrent d'un patient à l'autre en fonction des mutuelles d'assurance, il est possible que la proportion de cas pour lesquels un prélèvement a été effectué ait augmenté en 2008. La dengue étant une maladie à déclaration obligatoire (c'est-à-dire que la fiche de prélèvement doit être remplie par le médecin, que ce médecin effectue ou non un prélèvement, et que ce médecin appartienne ou non au réseau sentinelle), le nombre de cas totaux déclarés (cliniques et confirmés) n'a probablement pas été fortement impacté par ce changement de mesure.

Définition des cas

En 2014, un **cas suspect** (cas clinique) est défini comme un cas sans examen sanguin réalisé ou en attente des résultats d'un examen et présentant une fièvre ou une éruption papulo-maculeuse érythémateuse d'apparition brutale associée à au moins deux des signes suivants : hyperhémie conjonctivale, œdèmes des mains et/ou des pieds, céphalées, douleurs rétro-orbitaires, douleurs musculaires ou articulaires, asthénie, éruption cutanée maculo-papulaire, manifestations hémorragiques (pétéchies, purpura, saignements des gencives, du nez ou digestifs).

Un **cas possible** est défini comme un cas suspect survenant en période inter et pré-épidémique, avec une sérologie IgM positive pour la dengue sur un seul examen sanguin réalisé en phase tardive, sans argument épidémiologique et sans autre diagnostic. Les arguments épidémiologiques sont l'entourage

proche d'un cas confirmé, la notion de cas regroupés dans l'espace, ou le retour récent depuis une zone de circulation endémique du virus de la dengue.

Un **cas probable** est défini comme un cas suspect survenant en période inter et pré-épidémique, avec une sérologie IgM positive pour la dengue sur un seul examen sanguin réalisé en phase tardive, avec argument épidémiologique, ou bien un cas suspect survenant en période épidémique (c'est-à-dire lorsque la DASS-NC a déclenché une alerte épidémique suite à l'augmentation du nombre de cas déclarés), avec une sérologie IgM positive pour la dengue sur un seul examen sanguin réalisé en phase tardive.

Un **cas confirmé** est un cas suspect pour lequel l'un des trois tests suivant est positif : i) identification de l'antigène NS1 ; ii) identification d'ARN viral par RT-PCR ; iii) séroconversion IgM (test négatif à positif sur une paire de sérums prélevés sur le même patient à 7 jours d'intervalle) (DASS-NC, 2011).

La définition du cas a évolué au cours du temps en Nouvelle-Calédonie, notamment du fait de l'évolution des tests diagnostiques à disposition (voir Table 2.1). En particulier, jusqu'en 1991, le test diagnostic de référence était l'inhibition de l'hémagglutination. Les cultures cellulaires et la RT-PCR n'étaient effectuées alors que de manière ponctuelle sur quelques cas pour sérotypage. Les différences de sensibilité et de spécificité des différents tests utilisés au cours du temps ont pu entraîner des variations temporelles dans les niveaux d'incidence enregistrés.

Standardisation spatiale et temporelle du réseau

En épidémiologie, la standardisation est l'ensemble des « actions consistant à rendre comparables plusieurs séries d'observations sur des individus qui diffèrent par un ou plusieurs paramètres » (Toma et al., 2001). En matière de surveillance épidémiologique, afin de pouvoir suivre l'évolution d'une maladie dans le temps et dans l'espace, il est indispensable de rendre comparables des données récoltées dans des zones géographiques différentes, par des personnes variées, ou encore récoltées au même endroit mais à des moments différents. En général, cette standardisation est assurée par la définition d'un protocole de surveillance.

En Nouvelle-Calédonie, quelle que soit la période d'étude, les modalités de collecte (définition du cas, densité de médecins déclarant, modalités de transmission et d'enregistrement des cas) ont été homogènes spatialement sur l'ensemble du territoire. La standardisation spatiale est donc bonne, et les données de surveillance peuvent être comparées d'une zone géographique à l'autre.

En revanche, sur le plan temporel, les modalités de surveillance ont évolué au cours du temps, rendant difficile la comparaison des niveaux d'incidence entre périodes de temps successives (par exemple années 1970 et années 2000).

Table 2.1 : Tests de laboratoire utilisés pour le diagnostic de la dengue au cours du temps (source : IPNC, Ann-Claire Gourinat)

Détecté	Technique	Sensibilité[1]	1971 - 1989	1989 - 1990	1990 - 1991	1991 - 1995	1995 - 2001	2001 - 2007	2007 - 2009	2009 – présent
IgG	IH[2] (Institut Pasteur Paris)		X	X	X	X				
	IH[2]						X	X	X	X
IgM	IFI[3]			X	X					
	MAC ELISA[4]					X	X	X		
	Capture ELISA[4] Kit® (PANBIO)	95%							X	X
	Anti-Dengue Virus IgM Detection PA Kit ® PENTAX								X	
Virus	Culture cellulaire			X	X	X	X			
NS1	STRIP ® BIO-RAD	87 %[5]							X	X
	ELISA[4] Platelia® dengue NS1 BIO-RAD	90 %							X	X
ARN	RT-PRC[6]	> 99 %			X	X	X	X	X	X

1 pourcentage de tests avec un résultat positif parmi les patients infectés (en règle générale, patients dont la culture cellulaire est positive)
2 Inhibition de l'Hémagglutination
3 Immuno-Fluorescence Indirecte
4 Enzyme-linked ImmunoSorbent Assay
5 pourcentage de tests avec un résultat positif parmi les patients ayant un résultat positif en ELISA Platelia® dengue NS1 BIO-RAD
6 Reverse-Transcriptase Polymerase Chain Reaction

2.1.2.e. Organisation de la lutte contre les maladies vectorielles en Nouvelle-Calédonie

Nous décrirons ici uniquement la lutte anti-vectorielle (LAV) organisée par les pouvoirs publics, bien qu'un certain nombre de compagnies privées réalise également des prestations d'épandage insecticide. La DASS-NC émet des recommandations sur les mesures à mettre en œuvre, mais la mise en place des actions sur le terrain est de la compétence des communes (LAV gérée par la Mairie principale de chacune des communes).

A Nouméa, de 1980 à 2002, la LAV a consisté en des épandages systématiques tous les 15 jours en saison chaude, indépendamment du statut épidémique ou non dans le territoire. A partir de 2003, l'épandage systématique a été arrêté, au profit d'actions de lutte peri-focale (*i.e.* uniquement autour des foyers infectieux) et d'actions systématiques de prévention et de communication.

Lutte péri-focale à Nouméa

Dès qu'un cas d'arbovirose (dengue, chikungunya, fièvre Zika) est déclaré à la DASS-NC, les informations concernant le patient sont transmises au Service d'Inspection et de Prévention des Risques Environnementaux et Sanitaires (SIPRES). Les agents du SIPRES se déplacent alors au domicile du patient, et dans un périmètre de 100 m autour de ce domicile :

- Visitent toutes les maisons pour éliminer les gîtes larvaires ;
- Réalisent un épandage d'insecticide au moyen d'un Swing Fog® (Figure 2.5, photo de gauche), ou d'un LECO® (ou équivalent) monté sur un véhicule (Figure 2.5, photo de droite).

Certaines années, le nombre de cas déclarés devient trop important, et les visites des domiciles particuliers sont irréalisables en pratique sur le terrain. L'épandage est alors réalisé sous la forme de frappes de zones. La fréquence de passage est alors d'une fois par semaine dans chacun des quartiers de Nouméa.

Les années pour lesquelles des frappes de zone ont été réalisée sont les années 2003, 2009, 2013 (larges épidémies de dengue) et 2014 (épidémie de fièvre Zika). Les insecticides utilisés sont la deltaméthrine et le malathion (dont l'usage est autorisé en Nouvelle-Calédonie), des adulticides agissant sur le système nerveux des insectes adultes (Narahashi, 1982, 1996). L'usage de l'une ou l'autre des deux molécules est déterminé en fonction des résultats de tests sur la résistance des moustiques à ces molécules.

Actions systématiques de prévention et de communication

Les actions de prévention sont réalisées de manière permanente, indépendamment de la survenue d'une épidémie ou pas. Leur fréquence varie en fonction de la disponibilité des agents de terrain. Elles consistent en :

- Des programmes d'éducation sanitaire mis en place dans les écoles ;
- Des campagnes de communication via la presse locale ;
- L'emploi d'agents « PPIC » (Programme Provincial d'Insertion Citoyenne) qui passent régulièrement dans les maisons, visitent le jardin avec les propriétaires, et apprennent aux personnes à reconnaître et éliminer les gîtes larvaires.

Figure 2.5 : Epandages d'insecticides à Nouméa (Photos : Laurent Guillaumot, IPNC).

LAV en dehors du Grand Nouméa

Les actions de lutte décrites précédemment sont implémentées de manière homogène sur le Grand Nouméa. Dans les autres communes de Nouvelle-Calédonie, la mise en place de ces actions sur le terrain dépend de la volonté politique du Maire de chaque commune. Il n'existe pour l'instant pas de données sur ce qui est réalisé effectivement dans chacune des communes ailleurs que sur le Grand Nouméa.

2.2. Données analysées

2.2.1. Echelle d'étude et résolution

Dans cette thèse, sur le plan spatial, nous distinguerons plusieurs échelles spatiales en fonction de l'étendue de la zone d'étude. Bien que le terme « échelle » soit souvent employé dans le langage courant pour désigner également la résolution des données, il convient ici de bien distinguer les deux notions. Ainsi, pour une même échelle, les données peuvent être disponibles à basse ou haute résolution. Les données épidémiologiques n'étant pas disponibles sous forme de grilles régulières, nous désignerons par « résolution spatiale » l'unité géographique utilisée pour l'agrégation des données (correspondant le plus souvent à une division administrative). Nous distinguerons trois échelles spatiales :

- Nous désignerons par « échelle régionale » une échelle d'étude englobant plusieurs pays (exemple : le Pacifique). La résolution des bases de données régionales est souvent le pays, mais il serait théoriquement possible d'obtenir des données régionales à plus haute résolution. Cette échelle est caractérisée par l'hétérogénéité des infrastructures de santé et des moyens de lutte mis en œuvre entre les différentes unités spatiales ainsi que la grande distance pouvant séparer deux foyers infectieux ;

- Nous désignerons par « échelle nationale » (ou « échelle territoriale » pour la Nouvelle-Calédonie) une échelle d'étude comprenant un pays entier ou une partie relativement étendue (disons plus de 50

km) d'un même pays. La limite de 50 km est arbitraire, et est donnée ici uniquement à titre indicatif. Cette échelle est caractérisée par une homogénéité des infrastructures de santé, mais une répartition hétérogène de la population, ainsi qu'une hétérogénéité géographique (climat, altitude, occupation des sols…) ;

– Nous désignerons par « échelle locale » une échelle d'étude pour laquelle la zone étudiée est une ville (incluant les villes-pays comme Singapour), ou une portion de ville (exemple : Nouméa). Cette échelle est caractérisée par la densité importante de la population d'hôtes et la fréquence élevée des contacts entre hôtes en résultant, ainsi que la faible distance séparant différents foyers infectieux.

Sur le plan temporel, nous distinguerons différentes échelles en fonction de la périodicité des phénomènes étudiés. Pour chacune des échelles, les données peuvent être disponibles à différentes résolutions. Nous appellerons « résolution temporelle » des données le pas de temps utilisé pour enregistrer ou agréger les données (quotidien, hebdomadaire, mensuel, annuel…).

– Nous désignerons par « échelle inter-annuelle » l'échelle des études visant à étudier les variations d'un phénomène quelconque d'une année à une autre ;

– Nous désignerons par « échelle intra-annuelle » l'échelle des études visant à étudier les variations d'un phénomène quelconque au sein d'une même année. Cette échelle comprend donc l'étude des cycles saisonniers (cycles ayant une période de un an) et des cycles infra-saisonniers (cycles possédant une période inférieure à l'année).

2.2.2. Nature des données

Il existe deux types de données permettant d'effectuer la modélisation : les données épidémiologiques, décrivant la distribution des cas dans le temps et dans l'espace, et les données concernant les potentiels facteurs explicatifs (variables climatiques, entomologiques, socio-économiques, environnementales, etc.), encore appelées co-variables.

Les données épidémiologiques analysées sont en général des données de surveillance épidémiologique récoltées par les autorités sanitaires des différents pays. Elles se présentent soit sous la forme brute d'informations individuelles sur chacun des patients, soit sous des formes directement agrégées dans le temps et dans l'espace. Quelle que soit leur forme, ces données ont un caractère spatio-temporel, nécessitant l'utilisation d'outils de modélisation particuliers.

Au cours de cette thèse, quelle que soit le format initial des données, nous les avons transformées sous la forme de tableaux de données spatio-temporels, dans lesquels les données épidémiologiques individuelles sont agrégées dans le temps et dans l'espace (Figure 2.6).

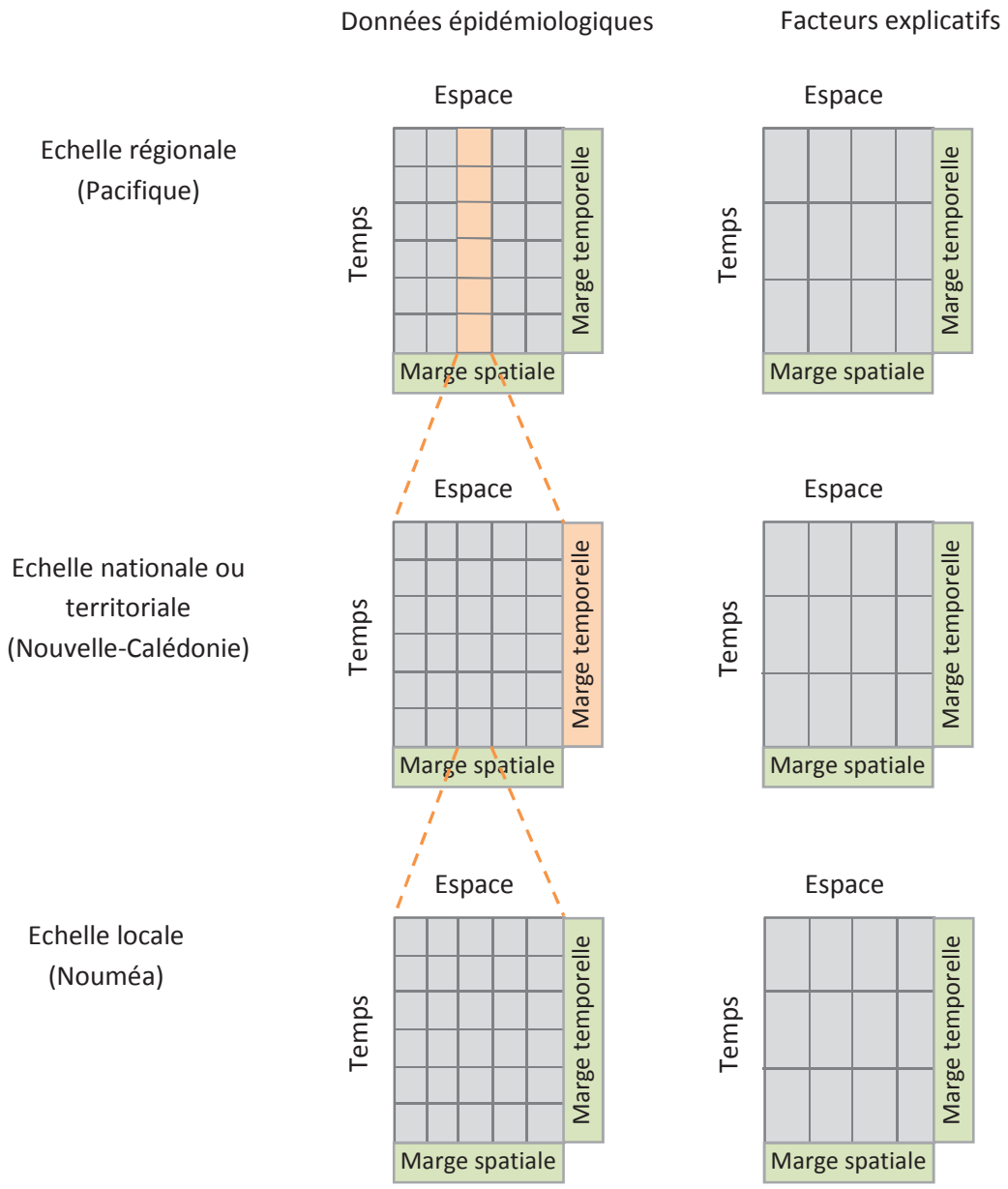


Figure 2.6 : Représentation schématique des données analysées au cours de la thèse. *Chacun des tableaux est un champ spatio-temporel, où chacune des colonnes représente une série temporelle à un endroit donné, et chacune des lignes représente une carte spatiale à un pas de temps donné. Les marges des tableaux correspondent à des cartes spatiales des moyennes sur l'ensemble des pas de temps (marge spatiale) ou à des séries temporelles moyennes obtenues en agrégeant les données sur l'ensemble de la zone d'étude (marge temporelle). La résolution des données épidémiologiques peut être différente de celle des facteurs explicatifs, ce qui peut représenter une contrainte importante en termes de modélisation. Un tableau de données d'une échelle inférieure correspond à la désagrégation dans l'espace de l'une des séries temporelles de l'échelle supérieure. La marge temporelle d'un tableau de données correspond à l'une des séries temporelles du tableau de données de l'échelle supérieure (par exemple, les deux colonnes rouges peuvent représenter la série temporelle des cas de dengue sur l'ensemble du territoire calédonien).*

2.2.3. Données épidémiologiques

2.2.3.a. Données dans le Pacifique (1971- 2009)

Les données épidémiologiques concernant les vingt PTIP membres de la CPS (Nouvelle Calédonie, Polynésie Française, Wallis et Futuna, Iles Cook, Fiji, Palau, Tonga, Vanuatu, Western Samoa, American Samoa, Etats Fédérés de Micronésie (FSM), Kiribati, îles Marshall, Nauru, Niue, Mariana du Nord, Papouasie Nouvelle Guinée (PNG), îles Salomon, Tokelau, et Tuvalu, voir Figure 2.1) ont été compilées par Elodie Descloux durant sa thèse (Descloux, 2011), en collaboration avec la CPS (Tom Kiedrzynski, Justus Benzler et Christelle Lepers). Afin de collecter le nombre de cas rapportés chaque année entre 1971 et 2009 dans les vingt PTIP, une grande enquête a été menée auprès des autorités de santé. Ces données ont été complétées avec les données disponibles dans les bases de données de la CPS et de l'OMS. Les chiffres pouvant différer d'une source à l'autre en fonction de la définition du cas utilisée (cas clinique, cas probable, cas confirmé biologiquement, ou nombre de cas global estimé), plusieurs réunions d'experts préalables à ma thèse ont été nécessaires afin d'établir un nombre « consensus » de cas annuels de dengue dans chacun des pays, correspondant à des cas cliniques plus ou moins confirmés. Ces chiffres sont présentés dans la Table 2.2. Comme déjà expliqué précédemment, les chiffres sont difficilement comparables d'un pays à l'autre puisque que les systèmes de surveillance ne sont pas homogènes. Dans le Pacifique, contrairement aux régions hyper-endémiques d'Asie et d'Amérique Latine, la co-circulation de plusieurs sérotypes au cours d'une même année épidémique dans un même pays est rare. Ainsi, pour la plupart des épidémies, le sérotype en cause est également connu.

Résolution spatio-temporelle des données épidémiologiques « Pacifique »

A l'échelle du Pacifique, les données sont disponibles uniquement à la résolution spatiale du pays, et à la résolution temporelle annuelle.

Table 2.2: Incidence annuelle de la dengue dans vingt Pays et Territoires Insulaires du Pacifique. La couleur correspond au sérotype responsable de l'épidémie (bleu pour le DENV-1 ; vert pour le DENV-2 ; rose pour le DENV-3 et orange pour le DENV-4). Les années au cours desquelles le sérotype dominant n'était pas précisé dans les données à disposition sont surlignées en gris. Les années pour lesquelles aucune donnée épidémiologique n'est disponible sont notées « NA » (pour « Non Available »).

Year	PF	Kiribati	Cook	Niue	Am Samoa	Tokelau	W Samoa	Tonga	Wallis	Tuvalu	Fiji	Marshall	Nauru	Vanuatu	NC	Solomon	FSM	PNG	NMariana	Palau
1971	12943	NA	NA	NA	NA	NA	NA	NA	790	NA	2960	NA	NA	NA	0	NA	NA	NA	NA	NA
1972	NA	NA	NA	790	NA	NA	NA	NA	0	NA	446	NA	3000	NA	54	NA	NA	1000	NA	NA
1973	56	0	0	9	0	0	2	NA	0	0	132	0	0	3	288	0	0	19	0	0
1974	48	0	0	0	0	108	0	15000	0	NA	25	0	3500	0	0	0	26	2	0	0
1975	1542	0	0	26	239	67	2547	8384	0	0	16203	0	0	1412	11	0	12	9	0	0
1976	757	0	333	0	0	10	157	2	520	0	10	0	0	93	282	0	0	564	0	0
1977	472	0	249	0	0	0	15	0	NA	0	1	0	0	17	260	0	0	60	0	0
1978	334	0	2	8	0	18	18	18	0	0	3	0	0	56	68	0	0	0	0	0
1979	1832	49	0	0	0	0	248	0	270	956	6914	0	52	77	118	0	0	0	0	0
1980	362	808	357	749	0	0	120	3552	21	15	127	0	538	16	796	0	0	0	0	0
1981	401	116	0	88	0	0	19	432	4	0	18	0	0	7	1	0	0	217	0	0
1982	246	11	0	2	0	0	25	1035	0	0	676	0	0	0	0	NA	0	0	0	0
1983	328	0	0	0	0	0	0	575	0	0	238	0	0	3	2	0	0	0	0	0
1984	419	5	0	0	0	0	0	180	0	0	190	0	0	0	6	0	0	0	0	0
1985	671	4	0	48	0	0	0	13	0	0	31	0	0	0	142	0	0	0	0	0
1986	391	0	0	269	0	7	0	8	0	0	269	0	0	0	426	0	0	0	0	0
1987	407	0	0	6	0	0	0	0	0	0	432	0	0	0	31	0	0	0	0	0
1988	545	45	0	4	0	0	0	17	0	0	22	0	0	2	194	0	0	0	0	1254
1989	8183	62	0	0	4	141	450	15	2854	58	2225	10	0	5606	2554	0	0	217	0	0
1990	2979	9	5	0	0	43	49	298	77	10	1473	0	0	989	326	0	13	0	0	0
1991	1196	0	1776	2	0	3	22	122	11	19	65	56	0	74	26	0	0	475	0	0
1992	1456	0	5	0	0	0	3	35	0	811	358	0	0	989	10	0	0	0	0	0
1993	1050	0	0	0	0	0	2	9	0	0	39	0	0	52	4	0	0	0	0	0
1994	1071	0	1	0	246	0	0	1	0	0	17	0	0	138	11	0	0	0	0	0
1995	1003	0	786	0	0	0	388	0	5	0	206	0	0	100	2211	1	66	202	0	857
1996	2175	0	2	1	49	0	1013	3	0	0	2653	0	0	19	2121	3	67	0	0	0
1997	2521	0	1075	0	NA	NA	163	0	0	0	24780	0	0	37	250	0	NA	NA	0	0
1998	525	170	0	0	205	NA	49	460	395	2	430	0	0	131	2618	0	700	NA	0	14
1999	522	0	0	0	1	NA	0	0	0	0	14	0	0	110	359	0	0	NA	0	1
2000	258	0	0	1	0	NA	200	0	0	0	3	0	0	0	15	0	14	NA	0	872
2001	4164	0	21	0	500	70	300	0	0	0	0	0	0	0	34	0	0	0	1418	486
2002	853	0	2256	0	1185	0	100	0	2000	0	0	0	0	0	105	30	0	22	0	65
2003	361	0	1	0	0	0	0	5192	938	0	0	0	233	0	5673	0	0	NA	0	16
2004	416	0	8	0	0	0	1	3	41	0	2	0	0	26	792	0	658	NA	0	72
2005	385	0	0	0	0	0	2	0	0	0	0	0	0	2	46	0	0	NA	0	10
2006	2833	0	479	0	1	0	29	0	0	0	25	0	0	168	53	0	0	NA	0	29
2007	4253	0	1224	0	126	0	23	62	0	0	100	0	0	82	49	0	6	NA	0	170
2008	450	831	91	3	649	0	677	195	0	0	1895	0	2	96	1179	3	10	NA	0	204
2009	2479	7	1404	0	373	0	NA	4	15	0	374	0	0	4	8456	0	4	NA	374	8

2.2.3.b. Données en Nouvelle-Calédonie

Données agrégées : résolution spatiale territoire et résolution temporelle mensuelle (janvier 1971 – janvier 2014)

Dans le cadre de la thèse de Descloux, la DASS-NC a mis à disposition le nombre de cas total rapportés chaque mois sur l'ensemble du territoire pour la période 1971-2014, sous la forme d'une seule série temporelle mensuelle agrégée.

Données brutes géo-localisées : (01/01/1995 – 30/09/2012)

La DASS-NC a mis à disposition les données détaillées et anonymes sur chacun des cas de dengue rapportés en Nouvelle-Calédonie du 1er janvier 1995 au 30 septembre 2012. A partir de 1995, les informations spatiales suivantes ont été enregistrées : la commune de résidence pour chacun des cas, et le quartier de résidence ainsi que l'adresse pour chacun des cas survenus dans les communes du Grand Nouméa (Nouméa, Dumbéa et Mont-Dore). A partir de 2008, la tribu de résidence également été enregistrée.

Ces données ont été fournies sous la forme de trois tableaux (1995-2005, 2006-2007 et 2008-2012) où chacune des lignes représente un cas déclaré de dengue, et chacune des colonnes représente une information relative au cas (voir Annexe 3 pour la liste des champs disponibles dans chacune des bases de données, ainsi que le codage utilisé).

Un premier travail a consisté à homogénéiser les codes utilisés sur l'ensemble des trois bases afin de créer une base de données unique couvrant l'ensemble de la période 1995 – 2012 (voir Annexe 3, base 1995-2012 pour le codage utilisé dans la base finale homogénéisée).

Ces données sont d'une excellente qualité, au vu du faible pourcentage de données manquantes pour les champs permettant une agrégation spatio-temporelle (Table 2.3).

Dans cette base, les cas enregistrés correspondent soit à des cas cliniques, soit à des cas confirmés. En période inter-épidémique (mai à octobre), entre 85 et 90 % des cas enregistrés dans la base sont des cas confirmés. Durant les pics épidémique (mars à juin), lorsque le nombre de cas devient trop important, seuls 50 à 70 % des cas enregistrés sont confirmés (voir **Erreur ! Source du renvoi introuvable.**). Cependant, étant donné que la majorité des cas non confirmés surviennent en période de forte prévalence de la maladie, il y a de grandes chances qu'ils correspondent réellement à des cas de dengue (la valeur prédictive positive d'un tableau clinique évocateur est élevée en période épidémique). Nous avons donc décidé d'analyser la totalité des cas, qu'ils soient confirmés ou non, car restreindre les études aux seuls cas confirmés aurait amputé les données de près de la moitié des cas en période épidémique.

Table 2.3: Données manquantes et statistiques de chacun des champs de la base finale

Nom du champ	Nombre de données manquantes	Pourcentage de données manquantes	Moyenne	Ecart-type	Autre
Quartier de résidence*	313*	2.5*	-	-	-
Commune de résidence	103	0.4	-	-	-
Sexe	512	2.1	-	-	Homme : 50.6 % Femme : 47.2 %
Age	1746	7.2	31.3	17.7	-
Date de consultation	0	0	-	-	-
Type de cas	0	0	-	-	Cas clinique : 70.5 % Cas confirmé : 29.5 %
Serotype	17367	71.5	-	-	-
Tribu (code)	1938**	43.3**			
Tribu (nom)	4045**	90.4**			

* Calculé uniquement sur les 12 388 cas géo-référencées à la résolution du quartier et situés dans les trois communes du Grand Nouméa

** Calculé uniquement sur les 4 475 cas situés en dehors du grand Nouméa et étant survenus après 2008

Table 2.4 : Pourcentage mensuel de cas de dengue confirmés parmi les cas enregistrés en Nouvelle-Calédonie

Mois	Janv	Fév	Mars	Avril	Mai	Juin	Juillet	Août	Sept	Oct	Nov	Déc
Pourcentage de cas confirmés	85	87	74	69	58	55	70	70	85	76	87	90

Résolution spatio-temporelle minimale des données brutes

Pour les analyses temporelles, l'information disponible est la date de consultation, au jour près. Nous pourrions donc théoriquement travailler sur des données quotidiennes. Mais l'information intéressante sur le plan épidémiologique est la date d'infection du patient. Or dans la base de données, seule la date de consultation est disponible de manière complète sur l'ensemble de la période. Au vu de l'incertitude de 1 à 7 jours sur la durée d'incubation de la maladie, ainsi que l'incertitude de 1 à 5 jours sur le temps entre l'apparition des premiers symptômes et la date de consultation, l'unité d'agrégation la plus fine possible sur le plan temporel est la semaine.

Pour les analyses spatiales, l'unité d'agrégation la plus fine possible est la commune si l'on travaille sur l'étendue du territoire, et le quartier, si l'on travaille sur l'étendue du Grand Nouméa. Pour les analyses à la résolution de la commune, les données d'incidence ne sont exploitables à Maré qu'à partir de 1996 (IPNC, 1996), à Lifou à partir de 1999 (IPNC, 1999) et à l'île des Pins à partir de 2003 (IPNC, 2003), dates avant lesquelles *Aedes aegypti*, l'unique vecteur de la dengue en Nouvelle-Calédonie, était absent.

Géo-référencement des cas en coordonnées lon/lat sur Nouméa

Au cours des analyses, il s'est avéré nécessaire d'augmenter la résolution spatiale des données épidémiologiques sur Nouméa. L'information sur l'adresse du patient étant disponible, nous avons entrepris de géo-référencer les 12 388 cas déclarés sur le Grand Nouméa. Cependant, les adresses ayant été enregistrées à la main de manière non standardisée, ce travail a pris 3 mois et ne nous a laissé que peu de temps pour la modélisation elle-même.

L'adresse de résidence du patient était manquante pour l'ensemble de l'épidémie de 2008/2009. Or cette épidémie était particulièrement intéressante à étudier, car c'est l'épidémie de plus grande ampleur durant la période d'étude. Pour le compte de la DASS-NC, nous avons complété leur base de données à partir de la base de données concernant la lutte anti-vectorielle péri-focale tenue par le SIPRES (mairie de Nouméa). La DASS-NC nous a ensuite fourni les adresses manquantes. La mairie de Nouméa vérifiant les adresses avant d'aller effectuer les épandages péri-focaux, les adresses de l'épidémie de 2008/2009 de la base finale sont donc plus propres que celles des autres années.

Grâce à une carte du cadastre des communes du Grand Nouméa pour laquelle chacune des adresses est liée à une coordonnée géographique, nous avons pu géo-référencer de manière certaine un total de 6 979 cas de dengue sur le 12 388 cas déclarés sur le Grand-Nouméa (~56,3 %). Les résultats finaux du travail de géo-référencement sont présentés en Table 2.5. Nous avons commencé par construire un algorithme de reconnaissance des caractères afin de géo-référencer les cas de manière automatique (voir Encadré 2). Le résultat de cet algorithme a ensuite été vérifié cas par cas, à la main, et une classification a été construite en fonction de la raison de l'erreur commise par l'algorithme. Cet algorithme a relativement bien fonctionné sur la commune de Nouméa à partir de 2008 grâce au travail de vérification des adresses effectué par la mairie pour l'épidémie de 2009. Le faible pourcentage de géo-référencement sur les communes du Mont-Dore et de Dumbéa est lié au fait que la plupart des patients habitant ces communes donnent leur adresse sous forme de lotissements (groupements résidentiels d'une dizaine de rues), pour lesquels nous n'avions à disposition aucune coordonnée géographique. La vérification des résultats de l'algorithme adresse par adresse a également permis de vérifier qu'un nombre important de cas (2 123, 17,1 %) pour lesquels l'algorithme n'avait pas fonctionné pouvaient potentiellement être géo-référencés adresse par adresse de manière semi-automatique car ils possédaient un nom de rue clairement identifiable. De plus ces cas n'étaient pas répartis aléatoirement dans l'espace (l'algorithme faisait des erreurs de géo-référencement sur certains noms de rue uniquement). L'exploitation de ces données pour l'analyse spatiale ne pouvait donc pas se passer d'un géo-référencement fiable de ces 2 123 cas. A cette fin, nous avons créé un deuxième algorithme permettant d'effectuer des requêtes semi-automatique dans la base cadastrale. Par manque de temps, nous avons limité ce géo-référencement semi-automatique aux années 1995 -1996 et aux années supérieures à 2005 (voir colonne « p_main » dans la Table 2.5).

Au vu des résultats du géo-référencement, nous avons décidé de limiter les analyses spatiales à haute résolution sur l'agglomération de Nouméa à l'épidémie majeure, constituée de 2 323 cas déclarés de

septembre 2008 à août 2009 dans la commune de Nouméa, excluant les communes du Mont-Dore et de Dumbéa. Au final, sur les 2 323 cas déclarés sur la commune de Nouméa durant l'année épidémique 2009, nous avons géo-référencé de manière certaine 1 762 cas (75,9 %). Les 561 cas non géo-référencés sont répartis comme suit :

- 35 cas (1,5 %) pour lesquels l'adresse du patient a été répertoriée comme un nom de bâtiment dont nous n'avions pas l'adresse (exemple : « auberge de jeunesse ») ;
- 225 cas (9,7 %), pour lesquels l'adresse du patient possédait un nom de rue identifiable, mais pas de numéro permettant de le localiser dans la rue ;
- 90 cas (3,9 %) pour lesquels l'adresse du patient a été répertoriée sous forme de lotissement (exemple : « tours de Magenta ») ;
- 98 cas (4,2 %,), pour lesquels il n'y avait aucune information spatiale exploitable ;
- 113 cas (4,9 %), pour lesquels le patient habitait dans un « squat », dont il n'existe aucun inventaire officiel, et encore moins de données géo-référencées les concernant ;

Si l'on peut émettre l'hypothèse que pour les 4 premières possibilités les cas sont répartis de manière aléatoire dans l'espace, l'absence de géo-référencement des squats introduit un facteur de confusion certain pour les analyses spatiales. Ce biais est à garder en tête lors de l'interprétation des résultats.

Table 2.5 : résultats du geo-référencement des données épidémiologiques dans les communes du Grand-Nouméa

Commune	Année	nb_cas*	nb_geo*	p_geo*	p_manquant**	p_num**	p_lotiss**	p_bat**	p_squat*	p_main*
Nouméa	1995	613	364	59,4	8,6	12,6	9,8	8,6	1,0	0
	1996	1002	671	67,0	9,2	9,6	8,5	5,6	0,2	0
	1997	88	29	33,0	6,8	2,3	8,0	10,2	5,7	34,1
	1998	814	316	38,8	3,1	4,8	7,2	7,9	3,3	34,9
	1999	149	57	38,3	0	4,7	6,7	6,0	1,3	43
	2000	6	3	50,0	0	16,7	0	0	0	33,3
	2001	17	4	23,5	11,8	0	5,9	0	0	58,8
	2002	44	21	47,7	0	4,5	6,8	9,1	0	31,8
	2003	1688	737	43,7	2,8	11,9	8,5	3,8	2,4	26,9
	2004	397	219	55,2	2,0	5,0	6,8	2,8	0	28,2
	2005	21	16	76,2	9,5	0	14,3	0	0	0
	2006	19	16	84,2	0	0	0	10,5	5,3	0
	2007	26	18	69,2	19,2	7,7	0	0	3,8	0
	2008	333	266	79,9	4,5	7,2	1,8	3	3,6	0
	2009	2323	1762	75,9	4,2	9,7	3,9	1,5	4,9	0
	2010	63	49	77,8	3,2	9,5	6,3	0	3,2	0
	2011	12	11	91,7	0	0	0	8,3	0	0
	2012	145	117	80,7	0,7	11,7	0,7	0,7	5,5	0
	Total	7760	4676	60,3	3,4	9,2	6,4	4,1	2,8	2,9
Mont-Dore	1995	217	3	1,4	7,4	17,1	59,4	7,8	6,9	0
	1996	329	1	0,3	18,2	17,0	49,8	4	10,6	0
	1997	21	0	0	19,0	19,0	42,9	9,5	9,5	0
	1998	235	0	0	11,5	11,1	46,8	4,7	17,4	8,5
	1999	73	0	0	11,0	15,1	58,9	2,7	6,8	5,5
	2000	6	0	0	16,7	0	50,0	0	16,7	16,7
	2001	8	0	0	0	12,5	75,0	0	0	12,5
	2002	10	0	0	0	20,0	70,0	10,0	0	0
	2003	535	1	0,2	3,9	12,5	61,9	4,7	8,6	8,2
	2004	116	0	0	1,7	20,7	60,3	6,9	4,3	6
	2005	11	0	0	0	18,2	63,6	18,2	0	0
	2006	5	0	0	0	40	20	40	0	0
	2007	7	0	0	0	28,6	57,1	0	14,3	0
	2008	70	26	37,1	8,6	18,6	28,6	4,3	2,9	0
	2009	1056	646	61,2	8,5	8,9	16,5	0,9	4,0	0
	2010	23	13	56,5	0	8,7	26,1	4,3	4,3	0
	2011	3	0	0	0	33,3	33,3	0	33,3	0
	2012	13	9	69,2	7,7	23,1	0	0	0	0
	Total	2739	699	25.5	6,2	12,7	39,6	3,5	7,2	7,3
Dumbéa	1995	109	8	7,3	12,8	20,2	55	2,8	1,8	0
	1996	103	11	10,7	21,4	28,2	38,8	1	0	0
	1997	7	1	14,3	42,9	28,6	14,3	0	0	0
	1998	83	11	13,3	34,9	21,7	18,1	2,4	3,6	6
	1999	54	6	11,1	42,6	16,7	16,7	0	0	13
	2000	0	0	0	0	0	0	0	0	0
	2001	1	0	0	100	0	0	0	0	0
	2002	14	2	14,3	7,1	50,0	21,4	0	0	7,1
	2003	570	21	3,7	13,2	29,3	30,4	0,7	9,3	13,5
	2004	61	1	1,6	8,2	37,7	32,8	0	0	19,7
	2005	8	1	12,5	0	37,5	50	0	0	0
	2006	2	0	0	50	0	0	0	50	0
	2007	2	0	0	50	0	50	0	0	0
	2008	142	48	33,8	2,1	26,8	26,8	4,2	6,3	0
	2009	690	319	46,2	10	14,5	14,1	1	14,2	0
	2010	14	9	64,3	0	28,6	7,1	0	0	0
	2011	3	2	66,7	0	0	0	0	33,3	0
	2012	21	11	52,4	4,8	19,0	4,8	4,8	14,3	0
	Total	1889	455	24,1	4,4	22,6	24,5	1,3	9,0	9,1

* nb_cas : nombre total de cas de dengue déclarés
nb_geo : nombre de cas ayant été géoréférencés
p_geo : pourcentage de cas ayant été géoréférencés
** p_manquant : % de cas pour lesquels aucune information spatiale exploitable n'était disponible
p_num : % de cas pour lesquels la rue de résidence du patient a été identifiée, mais dont le numéro est manquant
p_lotiss : % de cas pour lesquels le patient a donné une adresse sous forme de lotissement
p_bat : % de cas pour lesquels le patient a donné une adresse sous forme d'un bâtiment
p_squat : % de cas pour lesquels le patient réside dans un squat
p_main : % de cas potentiellement référençables au cas par cas (adresse avec avec un nom de rue et un numéro)

Encadré 2 : algorithme de géo-référencement des cas

Les adresses enregistrées dans la base de données épidémiologique ainsi que celles du cadastre ont d'abord été épurées, afin d'enlever tous les accents, les caractères spéciaux, la dénomination de la rue (« rue », « allée », « impasse », « de », « du », « des »…) ainsi que les informations ne correspondant pas à un nom de rue (numéro de BP, nom du quartier…). Les numéros de rue ont été isolés et placés dans une base à part. N'ont été conservés, dans les deux bases, que le nom de rue (exemple : l'adresse du cas « BP 99999 vallée du tir 83 R. A. St Euxpéry » a été transformé en « A St Euxpery », et le numéro « 83 » enregistré dans une base à part, et l'adresse de la base cadastre « 83 rue Antoine de Saint-Exupéry » a été transformée en « Antoine Saint Exupery »).

Puis trois algorithmes ont été utilisés afin d'associer, à chacune des adresses correspondant à un cas épidémiologique, une adresse de la base cadastre. Le premier algorithme a consisté pour chacune des adresses de la base cadastre (exemple « Rue Antoine de Saint Exupéry »), à chercher les correspondances entre la chaîne de caractères épurée (exemple « Antoine Saint Exupery ») et l'ensemble des chaînes de caractères épurées de la base épidémiologique, en laissant à la fonction de reconnaissance une marge d'erreur de 10 % pour prendre en compte les inversions manuelles de lettres, ou les lettres éventuellement manquantes. Cet algorithme a bien fonctionné, sauf pour les adresses composées d'un nom et d'un prénom, et pour lequel le prénom n'avait pas été noté dans la base épidémiologique (exemple : « A St Euxpery »). Afin de mieux géo-référencer ces adresses composées d'un nom et d'un prénom, nous avons donc utilisé un deuxième algorithme basé sur la même démarche, mais n'utilisant que le dernier mot de l'adresse cadastre (exemple « Exupery »).

A chaque nouvelle adresse de la base cadastre vérifiée, la proposition correspondant à des adresses déjà vérifiées était écrasée. Ainsi, si certaines adresses ayant un nom de rue très court (exemple : rue « Ry »), étaient vérifiées en dernier dans la base cadastre, tous les noms de rue de la base épidémiologique contenant la chaîne de caractère « ry » (comme « A St Euxpery ») étaient associées à cette rue de la base cadastre (exemple : même si la rue « Exupery » avait déjà été reconnue, cette rue était écrasée et remplacée par la rue « Ry »). Pour éviter ce problème, nous avons, au préalable, trié les noms de rue de la base cadastre par nombre croissant de caractères, afin de ne pas écraser les noms de rue détectés au fur et à mesure. Enfin, nous avons utilisé un dernier algorithme, basé sur une démarche inverse. C'est-à-dire qu'au lieu de chercher, pour chacune des adresse de la base cadastre, les adresse correspondantes dans la base épidémiologique, nous avons cherché, pour chacune des adresses de la base épidémiologique (exemple : « A St Euxpery), l'adresse correspondante dans la base cadastre. L'ensemble de ces trois algorithmes nous a permis, pour chacune des adresses de la base épidémiologique, d'obtenir trois propositions d'adresses de la base cadastre.

Nous avons ensuite validé, à la main, pour chacune des adresses épidémiologiques, laquelle des trois propositions était correcte tout en enregistrant également la raison pour laquelle aucun des trois propositions n'était correcte le cas échéant. Dans l'exemple, c'est le deuxième algorithme qui a

fonctionné et reconnu comme identiques les chaînes de caractère « Exupery » et « Euxpery », et a ainsi permis d'associer l'adresse initiale du cas « BP 99999 vallée du tir 83 R. A. St Euxpéry » à l'adresse «83 Rue Antoine de Saint-Exupéry » de la base cadastre ce qui a permis d'associer également les coordonnées géographiques de longitude et de latitude correspondantes.

2.2.4. Données entomologiques

2.2.4.a. Données entomologiques dans le Pacifique

A l'échelle du Pacifique, l'Unité de Recherche et d'Expertise d'Entomologie Médicale (URE-EnM) de l'IPNC nous a fourni la localisation actuelle (2014) des espèces vectrices de la dengue (voir Figure 1.3 page 41). L'information historique existe (Belkin, 1962), mais n'est pas encore compilée sous forme de données exploitables.

2.2.4.b. Données entomologiques à Nouméa (mars 2000 – septembre 2011)

Dans la commune de Nouméa, des données entomologiques mensuelles sont collectées conjointement par l'IPNC et le SIPRES (mairie de Nouméa) depuis août 1997. Entre 1997 et février 2000, les données ont été collectées par des agents temporaires, dont la formation était hétérogène. A partir de mars 2000, le nombre d'agents en charge de la collecte a été réduit, et seul un nombre restreint de personnes qualifiées et formées sont impliquées dans le réseau de surveillance entomologique, assurant ainsi la fiabilité des données. Un système de surveillance entomologique similaire à celui de Nouméa a été mis en place dans les communes du Mont-Dore à partir de 1997, et de Dumbéa à partir de 2006, mais le manque d'agents permanents en charge de la collecte rend les données trop parcellaires pour pouvoir être exploitées en modélisation. Nous présentons donc ici uniquement les informations relatives à la commune de Nouméa

Protocole d'échantillonnage

Afin de suivre dans le temps quatre indices entomologiques (voir liste et définition dans la Table 2.6), la commune de Nouméa a été découpée en trois grands secteurs (Figure 2.7) : le secteur Sud (1 545 maisons recensées), le secteur Ouest (2 600 maisons recensées) et le secteur Est (2 463 maisons recensées). Dans chacun des secteurs, environ cent maisons sont visitées chaque mois afin de collecter un certain nombre d'informations entomologiques (voir le formulaire utilisé en Annexe 5) et de construire les indices entomologiques présentés en Table 2.6. Dans chacun des secteurs, l'échantillonnage des maisons est un échantillonnage aléatoire en grappe. Chaque grappe comporte entre 10 et 27 maisons (moyenne 18,6 maisons par grappe). L'ensemble des maisons d'une grappe tirée au

sort est visité. Les immeubles et locaux commerciaux sont exclus de la surveillance pour des raisons d'homogénéité.

A Nouméa, seules trois personnes sont en charge de la collecte des données afin de réduire la subjectivité dans l'appréciation de certaines données entomologiques non précisément définies par une formule numérique.

2.2.4.c. Résolution spatiale et temporelle des données entomologiques de Nouméa

Résolution spatiale secteurs – Résolution temporelle mensuelle

Le protocole d'échantillonnage a été construit afin d'obtenir des données mensuelles pour chacun des trois secteurs (Sud, Est et Ouest). L'agrégation est faite par le l'URE-EnM de l'IPNC, qui produit des tableaux synthétiques annuels. Nous avons réalisé l'homogénéisation de l'ensemble des tableaux annuels jusqu'en septembre 2011 afin d'obtenir des séries temporelles couvrant l'ensemble de la période d'étude.

Ces données peuvent être utilisées pour effectuer des analyses temporelles à la résolution temporelle mensuelle et la résolution spatiale des secteurs. Avec seulement trois secteurs géographiques distincts, ces données ne permettent pas de faire des analyses spatiales fiables.

Résolution spatiale grappes – Résolution temporelle annuelle

Les données à la résolution spatiale des grappes ont été informatisées par le SIPRES à partir de 2012. Mais à cette résolution spatiale, l'échantillonnage temporel mensuel est perdu, le tirage au sort d'une grappe donnée pour un mois donné étant aléatoire. Pour que ces données soient exploitables pour des analyses spatiales, ou des analyses temporelles annuelles, il faudrait :

- Vérifier que chacune des grappes a été tirée au sort au moins une fois par an ;
- Créer un fichier « *shapefile* » de l'ensemble des grappes**,** afin de pouvoir agréger les données épidémiologiques à la même résolution spatiale ;
- Standardiser les données entomologiques, afin d'enlever l'effet de la saisonnalité sur les niveaux observés des indices entomologiques. En effet, au cours d'une même année, deux grappes peuvent être échantillonnées à six mois d'intervalle, conduisant nécessairement à des niveaux d'indice entomologiques plus élevés dans la grappe ayant été échantillonnée en saison chaude et humide. De même, une grappe donnée a pu être échantillonnée en saison froide une année, et en saison chaude une autre année, conduisant à des disparités interannuelles uniquement dus au biais d'échantillonnage à cette résolution spatio-temporelle.

Ayant reçu les données à la résolution des grappes tardivement au cours de la thèse, nous n'avons malheureusement pas eu le temps de les traiter. Ainsi, les indices entomologiques n'ont pu être exploités que pour des analyses temporelles mensuelles.

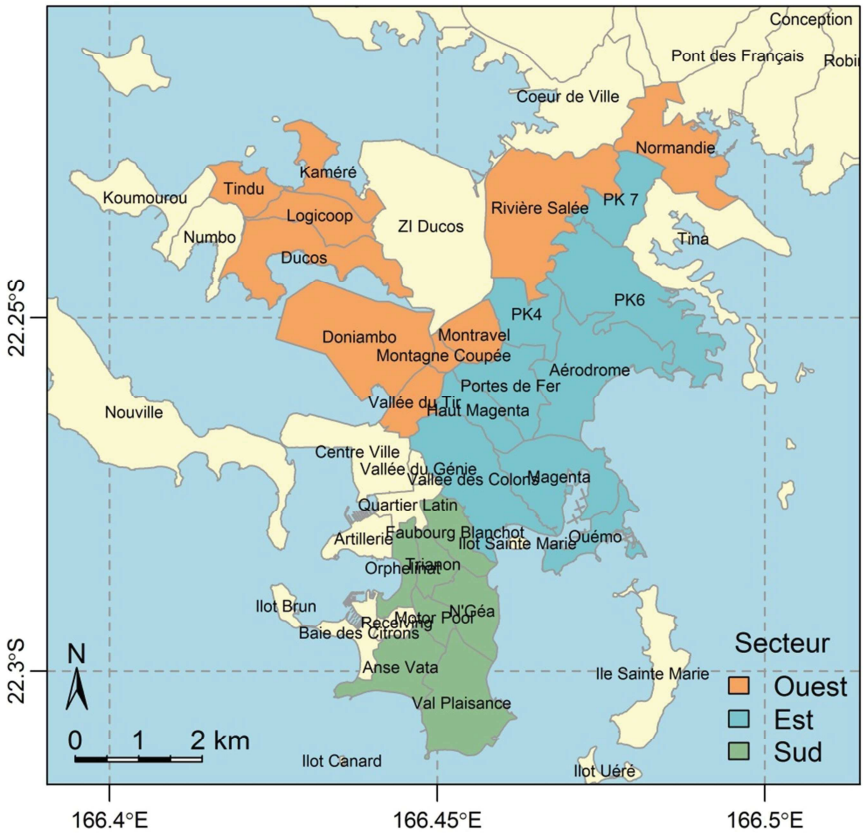


Figure 2.7 : Carte des secteurs de la surveillance entomologique à Nouméa

Table 2.6 : Indices entomologiques collectés par le SIPRES et l'IPNC (mars 2000 – septembre 2011)

Abréviation	Nom	Définition
IM	Indice Maison	Nombre de maisons dans lesquelles il y a au moins un gîte larvaire positif pour *Ae. aegypti* / nombre de maisons visitées
IBP	Indice de Breteau Pondéré	Nombre de gîtes positifs pour *Ae. aegypti* pondéré par la productivité des gîtes / nombre de maisons visitées
IPA	Indice de Productivité d'Adultes	Nombre de stades productifs (L4 + nymphes) d'*Ae. aegypti* / nombre de maisons visitées
IPPC	Indice Pondoirs Pièges Collants	Nombre de femelles adultes (données disponibles pour la thèse : 2007-2011)

2.2.5. Données climatiques

2.2.5.a. Données observées en Nouvelle-Calédonie

Les données observées de variables climatiques proviennent de mesures effectuées par Météo France au moyen de stations géo-localisées. Elles consistent, pour chaque station météorologique, en des séries temporelles souvent longues. Les séries temporelles directement mesurées présentent des ruptures qui peuvent par exemple avoir été introduites par une erreur de mesure, un changement de capteur, un défaut de calibration de capteur, ou encore un changement de l'environnement de la station. Météo France réalise donc régulièrement un travail d'homogénéisation des séries temporelles, qui consiste à détecter et corriger ces ruptures d'homogénéité, ainsi qu'à interpoler les données manquantes lorsque cela est possible. L'ensemble des séries analysées dans le cadre de ce travail de thèse sont des séries homogénéisées. Les séries mesurées par Météo France ont en général une résolution temporelle horaire, et peuvent être obtenues à n'importe quelle résolution temporelle inférieure à l'heure (http://www.meteo.nc/donnees-publiques-enseignement).

Données de la station de Nouméa (1er janvier 1971- 31 décembre 2009)

Dans le cadre d'un accord de partage des données, Météo France a mis à disposition les séries temporelles quotidiennes de la station de Nouméa (code station 98818001), allant du 1er janvier 1971 au 31 décembre 2009, pour les variables suivantes : température minimale (TN), température maximale (TX), précipitations (RR), humidité relative minimale (UN), humidité relative maximale (UX) et force des vents (FF). La longueur de cette série (près de 40 ans) permet de l'utiliser pour réaliser des analyses temporelles à n'importe quelle échelle (infra-saisonnière, saisonnière ou interannuelle).

Données sur l'ensemble du territoire (1er janvier 1995 – 31 décembre 2010)

Afin de réaliser les analyses spatiales, Météo France a mis à disposition les séries temporelles quotidiennes de RR (118 stations météorologiques), TN et TX (70 stations météorologiques), et l'humidité relative moyenne (moyenne de UN et UX, 42 stations) de stations réparties sur l'ensemble du territoire (voir Figure 2.4 page 75, et Figure 4.7 page 232). Les séries les plus longues couvrent la période allant du 1er janvier 1995 au 31 décembre 2010. La date du 1er janvier 1995 a été choisie car les données épidémiologiques ne sont géo-référencées par commune que depuis cette date. Pour certaines stations il existe des données manquantes sur plusieurs années en raison d'une installation de la station ultérieure à 1995, ou de pannes prolongées de capteurs (Table 2.7).

Après vérification de la qualité des séries temporelles, il s'est avéré que les capteurs d'humidité étaient peu fiables, et qu'une majorité des données d'humidité relative des stations autres que celle de Nouméa étaient dégradées. En conséquence, les analyses temporelles utilisant l'humidité relative ne sont pas réalisables en dehors de la ville de Nouméa. La couverture spatiale des stations dans lesquelles il y a au

moins une année complète de données n'est pas non plus suffisante pour réaliser une analyse spatiale utilisant l'humidité à la résolution des communes.

Table 2.7 : Données manquantes des variables météorologiques. Les chiffres dans la table correspondent au nombre de stations présentant 1, 2, … ou 16 années manquantes.

Nombre d'années de données manquantes	1 an	2 ans	3 ans	4 ans	5 ans	6 ans	7 ans	8 ans	9 ans	10 ans	11 ans	12 ans	13 ans	14 ans	15 ans	16 ans
RR	50	23	3	5	6	2	3	3	4	2	2	2	2	2	4	5
TN et TX	33	4	7	2	2	3	4	0	2	2	0	2	1	2	3	3
Humidité relative	16	6	2	1	0	1	4	0	2	0	1	0	0	0	2	7

Dans le cadre de l'évaluation du lien entre la circulation virale de la dengue et les variables climatiques, ces données peuvent être utilisées pour :

- Des analyses spatiales uniquement, à la résolution des communes, en calculant des moyennes temporelles sur l'ensemble de la période d'étude ;
- Des analyses temporelles dans chacune des 33 communes ;

Dans ces deux cas, les données épidémiologiques étant analysables à la résolution des communes uniquement, il est nécessaire de trouver un moyen de calculer un climat moyen par commune à partir des données de quelques stations réparties dans la commune.

Ces données ne peuvent pas être utilisées pour :

- Des analyses spatio-temporelles conjointes dengue /climat, du fait du nombre important de données manquantes et de la couverture spatiale insuffisante des séries temporelles complètes.

Données de précipitations issues d'un krigeage des données observées (résolution spatiale 1 km – carte moyenne de 1991 à 2000)

Météo France a réalisé un krigeage (interpolation sur une grille spatiale de données mesurées en certains points de l'espace, basée sur un modèle statistique de l'autocorrélation spatiale des données observées) des données quotidiennes de précipitations moyennées sur la période 1991 – 2000 (Figure 2.8). Ce krigeage a été fait à une résolution de 1 km. En plus des krigeages classiques basé uniquement sur les données observées, Météo France a intégré les perturbations potentielles de précipitations dues à l'orographie dans AURHELY, le modèle statistique ayant servi à kriger les données (Bénichou and Le Breton, 1987).

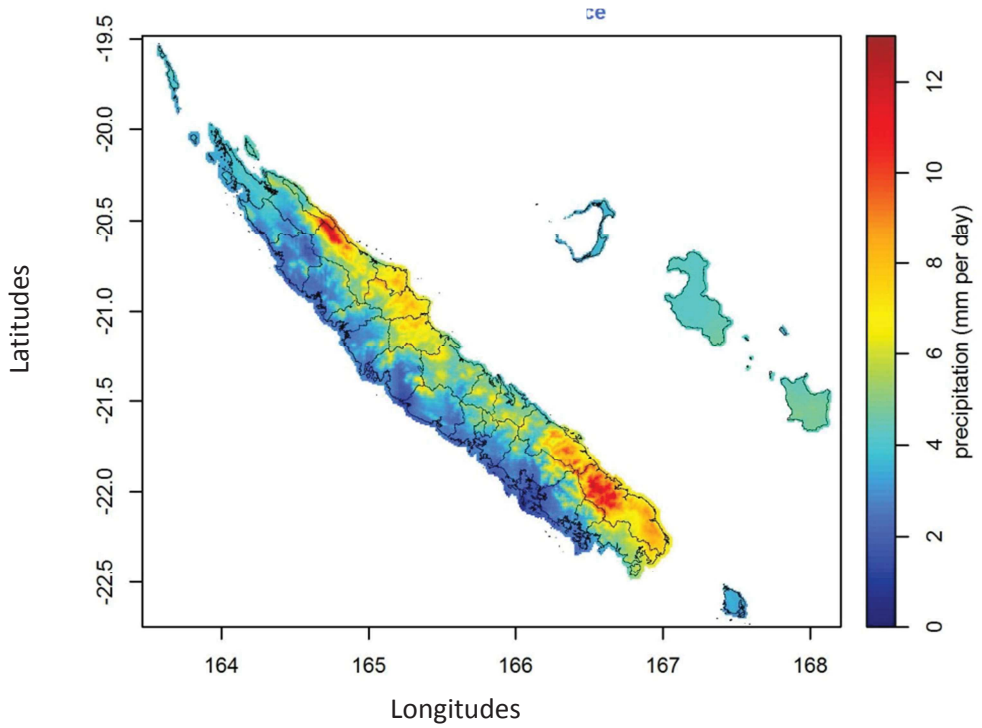


Figure 2.8 : Carte des précipitations quotidiennes moyennes de Nouvelle-Calédonie *obtenue par krigeage des moyennes quotidiennes de chaque station de 1991 à 2000.*

2.2.5.b. Données issues de simulations numériques autour de la Nouvelle-Calédonie

Afin d'étudier les régimes de temps et la circulation orographique autour de la Nouvelle-Calédonie, une équipe d'océanographes et de climatologues de l'IRD et de Météo France avait, préalablement à mon arrivée, réalisé des simulations numériques de variables atmosphériques (Lefèvre et al., 2010) en utilisant le modèle régional WRF (Weather Research and Forecast model, (Skamarock et al., 2005)). Les simulations consistent à :

- Résoudre numériquement les équations décrivant la circulation des fluides atmosphériques ;
- Réaliser une descente d'échelle dynamique : les données d'environnement atmosphérique sont généralement simulées à large échelle spatiale. La descente d'échelle consiste à redéfinir une grille de calcul des simulations dont la résolution est plus fine, et à se servir des simulations basse résolution pour forcer les simulations haute résolution. En plus d'augmenter la résolution spatiale des données climatiques, le modèle WRF possède des éléments permettant de prendre en compte les perturbations de la circulation atmosphérique créées par l'orographie, et permet donc de reproduire le climat de Nouvelle-Calédonie de manière plus fidèle que les données à large échelle.

Résolution spatio-temporelle des données climatiques simulées

Les variables simulées comprenaient, entre autres, la température moyenne de surface, les précipitations totales, et l'humidité relative. Les sorties du modèle sont enregistrées au format « *ncdf* », un format qui permet de stocker de manière compacte un grand nombre de variables sur des grilles multidimensionnelles non nécessairement régulières. Les variables ont été simulées à deux résolutions spatio-temporelles différentes :

- Une résolution spatiale de 4 km et temporelle de 6 h pour les étés austraux (1er novembre au 30 avril) 2005/2006, 2006/2007 et 2007/2008 ;
- Une résolution spatiale de 20 km et temporelle de 6 h du 1er mars 2000 au 1er janvier 2011.

Les données à 4 km n'étant disponibles que pour trois étés, elles ne pouvaient servir qu'à effectuer des analyses spatiales. Les données à 20 km pouvaient éventuellement servir à tout type d'analyse spatio-temporelle, sauf les analyses temporelles interannuelles, puisque la période couverte (2000-2011) n'était pas suffisamment longue.

Validation des données simulées

Avant de pouvoir utiliser ces données simulées dans des modèles explicatifs de la variabilité spatio-temporelle de la dengue, il était nécessaire de vérifier leur validité. La Figure 2.9 montre les cartes spatiales moyennes (marge spatiale) de chacun des champs de données simulées. Nous voyons que les données à haute résolution spatiale peuvent être utilisées pour des analyses spatiales, tandis que celle simulées sur des grilles à 20 km de résolution ne peuvent être utilisées spatialement.

Nous avons effectué la validation temporelle des données simulées en comparant les séries temporelles simulées aux séries temporelles observées enregistrées par 24 stations météorologiques réparties sur le territoire (voir les points sur la Figure 2.9.C). Pour chacune des paires de série temporelle (simulée vs observée), nous avons calculé la racine de l'erreur quadratique moyenne du modèle (racine de la somme, sur chacun des pas de temps, des différences entre observé et simulé au carré), les différences de moyenne, les différences d'écart-type et les corrélations entre séries temporelles observées et séries temporelles simulées. La Figure 2.10 montre les résultats des coefficients de corrélation. Les données simulées de température sont corrélées à plus de 0,8 avec les données observées quelle que soit la résolution ou le jeu de données simulé utilisé. Les données simulées de température peuvent donc être utilisées pour des analyses et modélisations temporelles avec les données épidémiologiques. Les données d'humidité relative, en revanche, sont moins bien corrélées (coefficient de corrélation moyen de 0,5 quelle que soit la résolution temporelle ou le jeu de données utilisé), et ne peuvent être utilisées pour des modélisations futures. Les données simulées de précipitations peuvent être utilisées à une résolution hebdomadaire ou mensuelle pour le jeu de données à 4 km de résolution, et à une résolution temporelle mensuelle pour le jeu de données à 20 km de résolution (coefficients de corrélation > 0,5).

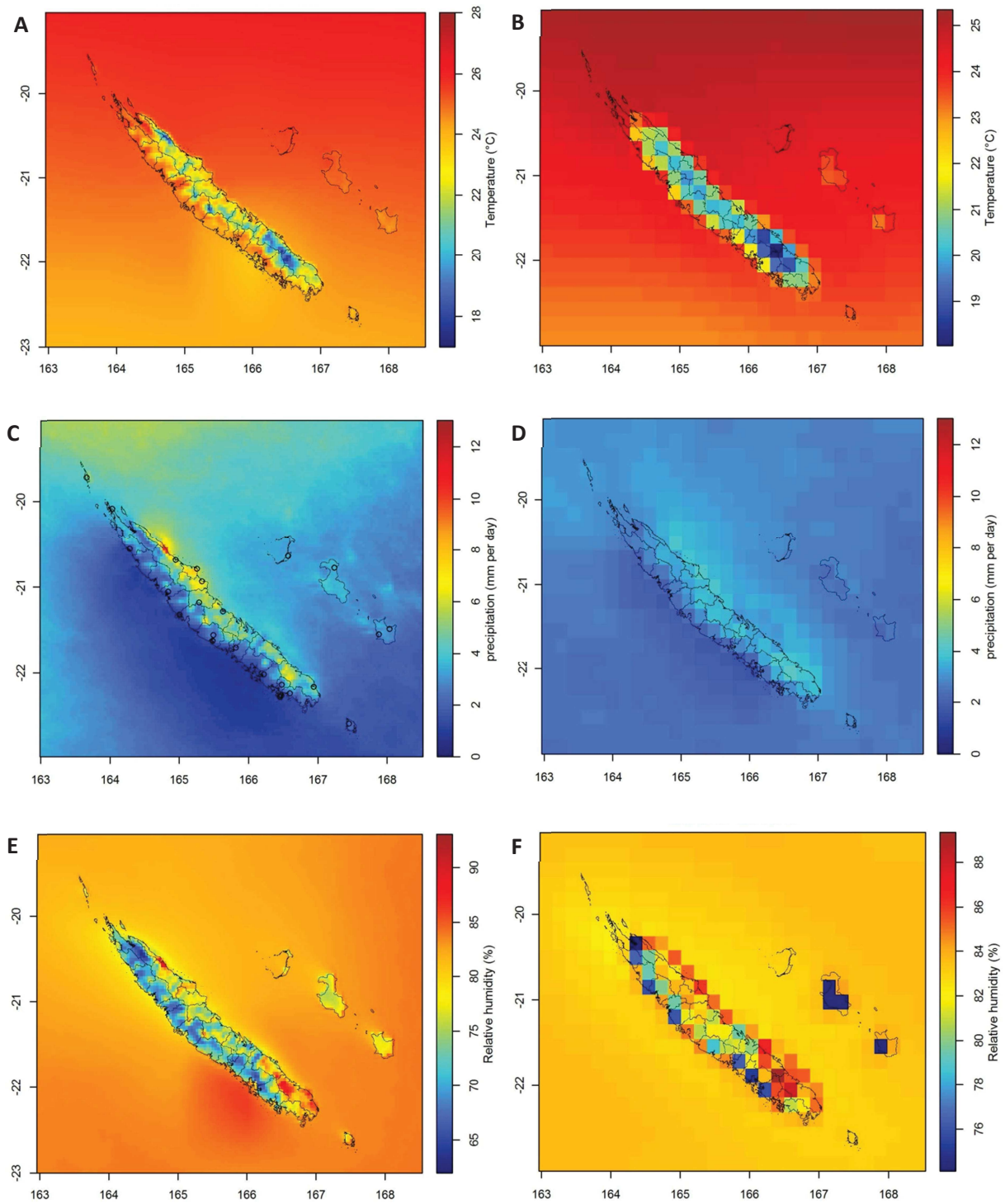


Figure 2.9 : Validation spatiale des variables atmosphériques simulées par le modèle WRF. *Chaque carte représente la carte spatiale moyenne des données de température (**A** et **B**), précipitations (**C** et **D**) et humidité relative (**E** et **F**) simulées par le modèle WRF utilisant une grille à 4 km de résolution (A,C et E) ou à 20 km de résolution (B,D et F). Chaque pixel correspond à la moyenne temporelle de la série temporelle correspondante (étés austraux de 2005 à 2008 pour les cartes A,C et E, et période 2000-2011 pour les cartes B, D et F). Nous pouvons voir sur la carte C que la correspondance spatiale avec les données obtenues grâce au krigeage des données observées par Météo France est bonne. En revanche, les simulations effectuées à une résolution de 20 km ne permettent pas d'obtenir une bonne représentation spatiale des variables climatiques (voir carte D).*

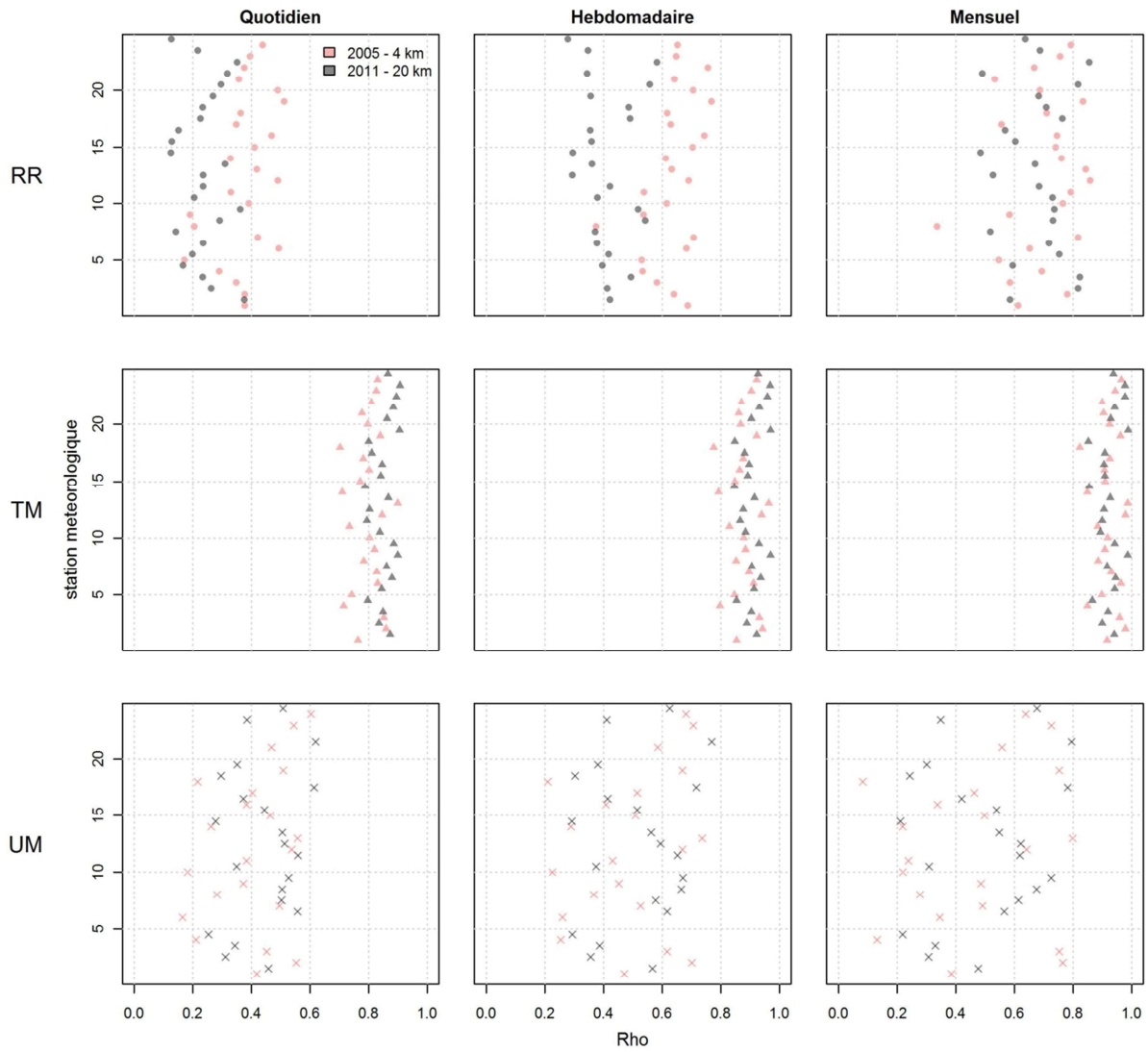


Figure 2.10 : Validation temporelle des variables atmosphériques simulées par le modèle WRF. *Sur chaque panneau, l'abscisse des points représente le coefficient de corrélation entre la série simulée et la série observé, l'ordonnée représente le numéro de la station météorologique utilisée pour la validation (24 stations en tout). Chaque ligne représente les résultats de la validation pour une variable climatique donnée (Précipitations en ligne 1, température moyenne en ligne 2 et humidité relative en ligne 3), et chaque colonne correspond à la résolution temporelle des séries. La couleur des points correspond au jeu de données simulé utilisé pour la comparaison avec les séries observées : en rouge pour les données simulées sur les été austraux de 2005 à 2008 à une résolution spatiale de 4 km, et en noir pour les données simulées sur la période 2000-2011 à une résolution spatiale de 20 km.*

Les données haute résolution (4 km) ont donc une excellente validité à la fois spatiale et temporelle pour les températures et précipitations. Malheureusement, elles ne sont disponibles que sur trois étés austraux. Au vu du temps nécessaire au paramétrage du modèle WRF et au calcul numérique des simulations atmosphériques (plusieurs mois), et suite à une panne et un remplacement du cluster de calcul numérique de l'IRD, il n'a pas été possible d'effectuer de nouvelles simulations à 4 km afin d'obtenir des séries temporelles de même longueur que celles des données épidémiologiques.

2.2.5.c. Données de projections climatiques en Nouvelle-Calédonie à l'horizon 2100 (1971-2100)

Sous l'égide du Programme Mondial de Recherche sur le Climat, financé en grande partie par les Nations Unies, un groupe de travail sur les modèles climatiques couplés Océan/Atmosphère a lancé en 1995 le projet CMIP (Coupled Models Intercomparison Project) visant à coordonner la recherche sur le changement climatique en fournissant un protocole standardisé pour l'étude et l'évaluation des modèles de circulation globale couplés Océan/Atmosphère (AOGCM). Les sorties des modèles sont ainsi évaluées selon le protocole fourni et sur la base du volontariat par l'ensemble de la communauté de recherche sur le climat (chercheurs, institutions internationales…). L'ensemble des résultats de l'évaluation des modèles est ensuite synthétisé par le Groupe d'Experts Intergouvernemental sur l'Evolution du Climat (GIEC, an anglais « Intergovernmental Panel on Climate Change » ou IPCC), qui produit périodiquement des rapports d'évaluation (AR) du changement climatique. Ces rapports permettent d'identifier les points faibles des modèles, et de donner des orientations pour leur amélioration. Les séries climatiques utilisées dans le cadre de mon travail ont été créées et évaluées suivant le protocole CMIP5-AR4, correspondant à la 5ème phase du projet CMIP (Taylor et al., 2012), pour lesquels les modèles climatiques ont été modifiés pour prendre en compte les recommandations de l'AR4. Ces données sont disponibles gratuitement en ligne (http://www-pcmdi.llnl.gov).

Parmi l'ensemble des AOGCM, des experts de l'IRD (Christophe Menkès) et de Météo France (Anne leroy, Temaui Tehei, Virgil Cavarero) en ont sélectionné six ("CanESM2", "CNRM-CM5", "IPSL-CM5A-MR", "IPSL-CM5B-LR", "MPI-ESM-LR", et "NorESM1-M"), sur la base de leur capacité à reproduire de manière fiable la position de la SPCZ et ses variations, et donc à reproduire de manière relativement fidèle le climat du Pacifique Sud. Différents *scenarii* d'émission de gaz à effet de serre sont utilisés dans ces modèles : le scénario RCP 4.5 et le scénario RCP 8.5 qui diffèrent par les hypothèses de mise en œuvre de mesures visant à réduire les émissions dans le futur. Le scénario RCP 8.5 est plus pessimiste que le scénario RCP 4.5. Plusieurs variables sont disponibles, mais nous n'avons utilisé que les températures.

Pour chacun des modèles, les séries temporelles sont simulées sur la période 1971-2005 (données « historiques ») et sur la période 2006-2100 (« projections »). Ces données sont disponibles sur des grilles spatiales basse résolution qui diffèrent d'un modèle à l'autre. Les points de grille peuvent être situés en mer et relativement éloignés de la Nouvelle-Calédonie. D'autre part, ces simulations sont réalisées à basse résolution, sans nécessairement tenir compte de la géographie locale. Les séries temporelles historiques peuvent donc s'écarter significativement des séries temporelles réellement observées en Nouvelle-Calédonie, et les données doivent être corrigées. Météo France a réalisé cette correction au moyen d'une descente d'échelle statistique (Cavarero et al., 2012).

Séries temporelles non corrigées (Nouméa).

Les données brutes non corrigées correspondent à des séries temporelles quotidiennes de température moyenne. Pour chacune des stations météorologiques de Nouméa, Touho, Bourail, Lifou et Koné,

Virgil Cavarero (Météo France) a téléchargé les séries temporelles correspondant au point de grille le plus proche de chacune des stations. Dans cette thèse, nous n'avons eu besoin, pour chacun des modèles, que de la série de température moyenne du point de grille le plus proche de Nouméa (voir Figure 2.11 pour les données historiques et projections à l'horizon 2100 de la température moyenne à Nouméa).

Séries temporelles corrigées (Nouméa, Touho, Bourail, Lifou et Koné).

Pour chacune des stations météorologiques, Météo France a réalisé une descente d'échelle des données simulées au moyen d'une méthode quantile-quantile (Cavarero et al., 2012) qui consiste en quatre étapes : i) calcul des percentiles des données observées à la station météorologique ; ii) calcul des percentiles des données simulées historiques ; iii) pour chaque percentile, calcul du facteur à appliquer aux données historiques pour les ramener au niveau des données observées ; iv) les données simulées corrigées sont obtenues en appliquant les facteurs de correction calculés à chacun des percentiles des données simulées brutes.

Ces facteurs de correction, calculés sur les données historiques, peuvent ensuite être appliqués aux percentiles des données de projections futures pour obtenir des projections corrigées. Les séries observées utilisées pour calculer les facteurs de correction étaient des séries de température maximale et de température minimale. Les séries simulées non corrigées de température moyenne ont donc été transformées en séries corrigées de température maximale ou minimale.

Nous avons, au final, pour les stations météorologiques de Nouméa, Touho, Bourail, Lifou et Koné, et pour chacun des six modèles sélectionnés, eu accès à des séries temporelles quotidiennes corrigées de température maximale et de température minimale sur la période 1971-2100 (voir Figure 2.11 pour un exemple à Nouméa).

Etude des augmentations moyennes de température en Nouvelle-Calédonie

Dans le cadre de cette thèse, nous avons utilisé ces séries temporelles uniquement pour calculer des augmentations moyennes de température dans le futur. Afin d'évaluer la possibilité d'utiliser les projections de manière spatiale, nous avons comparé les augmentations de température maximale prévues pour les différentes stations. A moins d'un dixième de degré Celsius près, les augmentations de température maximale et minimale prévues sont les mêmes à Bourail, Koné, Nouméa, Lifou et Touho. Nous avons donc considéré l'augmentation moyenne de température comme homogène sur l'ensemble du territoire. Afin de simplifier les analyses, nous avons comparé, à Nouméa, l'augmentation prévue de température en utilisant soit les données corrigées (température maximale et minimale), soit les données non corrigées (température moyenne). L'augmentation moyenne de température étant homogène sur le territoire, et étant identique que l'on utilise les séries corrigées ou non corrigée pour la

calculer (Figure 2.11), nous avons décidé, au final de n'utiliser que les séries temporelles non corrigées correspondant à la station de Nouméa.

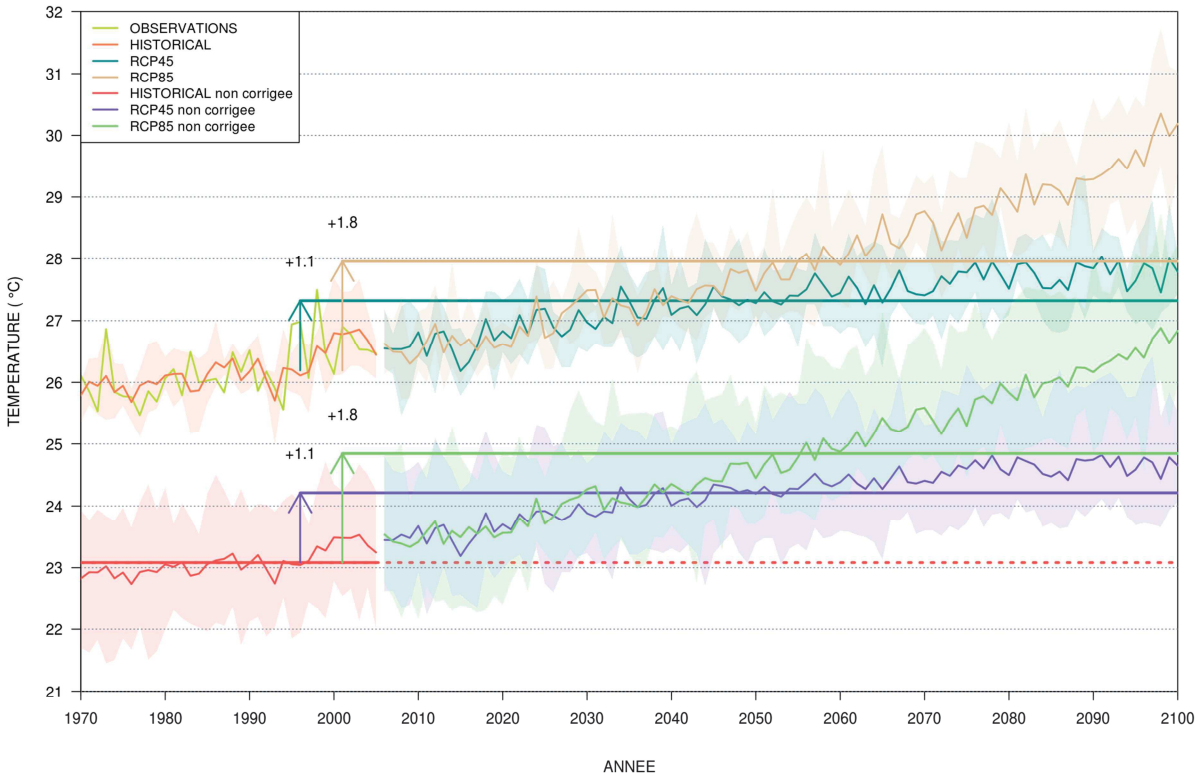


Figure 2.11 : Données historiques et évolution de la température moyenne et maximale à Nouméa à l'horizon 2100. *La courbe vert kaki 1971-2005 correspond à la moyenne annuelle des températures maximales observées à la station de Nouméa. Les autres courbes correspondent aux données simulées (les courbes sont la moyenne, pour chaque pas de temps, des données des six modèles, et les ombres représentent le minimum et le maximum sur l'ensemble des six modèles). Les 3 courbes du bas correspondent aux simulations brutes non corrigées (température moyenne), les trois courbes du haut correspondent aux simulations corrigées (température maximale). Les flèches verticales et les chiffres au-dessus représentent l'augmentation de température, en °C, entre la moyenne des données historiques et la moyenne des données de projection. Que l'on utilise les données corrigées ou non corrigées, l'augmentation de température prévue est la même. (Figure amicalement créée par Cyril Dutheil, IRD, sur la base des scripts de V. Cavarero, Météo France).*

2.2.5.d. Données ENSO

L'intensité d'ENSO est mesurée par différents indices dont le Multivariate ENSO Index (MEI) et la Southern Ocsillation Index (SOI). La SOI est souvent la plus utilisée. Elle mesure la différence de pression de surface entre Tahiti (Polynésie Française) et Darwin (Australie). Elle est négative durant « El Niño » et positive durant « La Niña ». Les indices ENSO peuvent être obtenus par exemple via le Centre de Prévisions Climatiques de la National Oceanic and Atmospheric Administration (NOAA, http://www.cpc.ncep.noaa.gov/data/indices/)

2.2.6. Données démographiques et socio-économiques

2.2.6.a. Données des îles du Pacifique

Dans le cadre de sa thèse, Descloux a collecté, en parallèle des données épidémiologiques et toujours en collaboration avec la CPS, les données de recensement de population dans chacun des 20 PTIP, et a interpolé les données de manière linéaire pour obtenir les données pour les années entre deux recensements.

2.2.6.b. Données de Nouvelle-Calédonie

Les données démographiques de Nouvelle-Calédonie proviennent des recensements généraux de population de Nouvelle-Calédonie effectués par l'Institut de la Statistique et des Etudes Economiques (ISEE).

Données de population (1971 – 2009)

Les séries temporelles d'effectif de population de même résolution spatio-temporelle que les données épidémiologiques (résolution temporelle hebdomadaire, mensuelle ou annuelle, résolution spatiale du territoire ou de la commune) ont été obtenues par interpolation linéaire des données de recensement 1956, 1963, 1969, 1976, 1983, 1989, 1996, 2004 et 2009.

Données d'âge par commune

Afin de réaliser l'étude spatiale sur le territoire à la résolution des communes, nous avons utilisé les données de pyramide des âges des recensements de 1989, 1996, 2004 et 2009 (Institut de la statistique et des études économiques (ISEE), 1989, 1996, 2004, 2009). Pour les années 1989, 1996 et 2004, ces données n'étaient disponibles que sous format papier dans les archives même de l'ISEE, et nous avons dû les informatiser.

Données socio-économiques

Pour l'étude spatiale à la résolution des communes intégrant les facteurs socio-économiques comme variables explicatives du risque épidémiologique de dengue, nous avons utilisé les données du recensement de 2009 disponibles gratuitement en ligne (Institut de la statistique et des études économiques (ISEE), 2009). Deux types de variables sont disponibles : des variables individuelles, et des variables caractérisant les logements. Afin de pouvoir comparer les chiffres d'une commune à l'autre, nous avons transformé chacune des variables en pourcentage, en utilisant le nombre d'individus total et

le nombre de logements totaux comme dénominateur pour le calcul des pourcentages sur les variables individuelles ou « logement ».

Sur la base de l'épidémiologie de la dengue, nous avons sélectionné uniquement les variables pouvant avoir un lien potentiel avec la transmission de la dengue (activité ou type de logement pouvant favoriser un contact vecteur – hôte, voir article 3, Table 4.2 page 243 pour une liste complète).

2.2.6.c. Données de Nouméa

Pour l'étude spatiale à l'échelle de Nouméa en 2009, les données socio-économiques de la ville de Nouméa proviennent du recensement de population 2009. Ces données sont accessibles gratuitement en ligne à la résolution des quartiers. Nous avons effectué une demande motivée par courrier pour les obtenir à la résolution de l'îlot.

Le travail à l'échelle de Nouméa ayant été réalisé dans le cadre d'un projet de recherche en collaboration avec des collègues réunionnais, les variables pouvant influencer la transmission de la dengue ont été sélectionnées de manière collégiale par l'ensemble des chercheurs impliqués dans le projet.

Afin de réduire le nombre de variables étudiées, et d'augmenter la pertinence de chacune d'entre elles, les modalités de certaines variables ont été agrégées (exemple : la variable « lieu de résidence avant 2004 » a été sélectionnée parce qu'elle reflète potentiellement le statut susceptible ou immun des hôtes lors de l'épidémie de 2009 ; ainsi les modalités correspondant à des pays indemnes de dengue (métropole, étranger autre) ont été regroupées entre elles, et les modalités correspondant à des pays touchés par la dengue ont été regroupées entre elles).

2.2.7. Données de télédétection

Afin d'étudier les facteurs associés à la distribution spatiale des cas de dengue sur Nouméa en 2009, nous avons commandé une image satellite RapidEye datant du 16/10/2010 (Figure 2.12). Cette image a une résolution de 4,5 m. A partir de cette image, des experts en télédétection de l'IRD (Marc Despinoy, Alizé Mercier) ont calculé une série d'indicateurs présentés en Table 2.8. Les données de télédétection permettent d'avoir accès à des informations sur l'occupation des sols, telles que la présence de végétation ou la présence de zones humides. L'occupation des sols peut avoir une influence sur le cycle de vie du moustique.

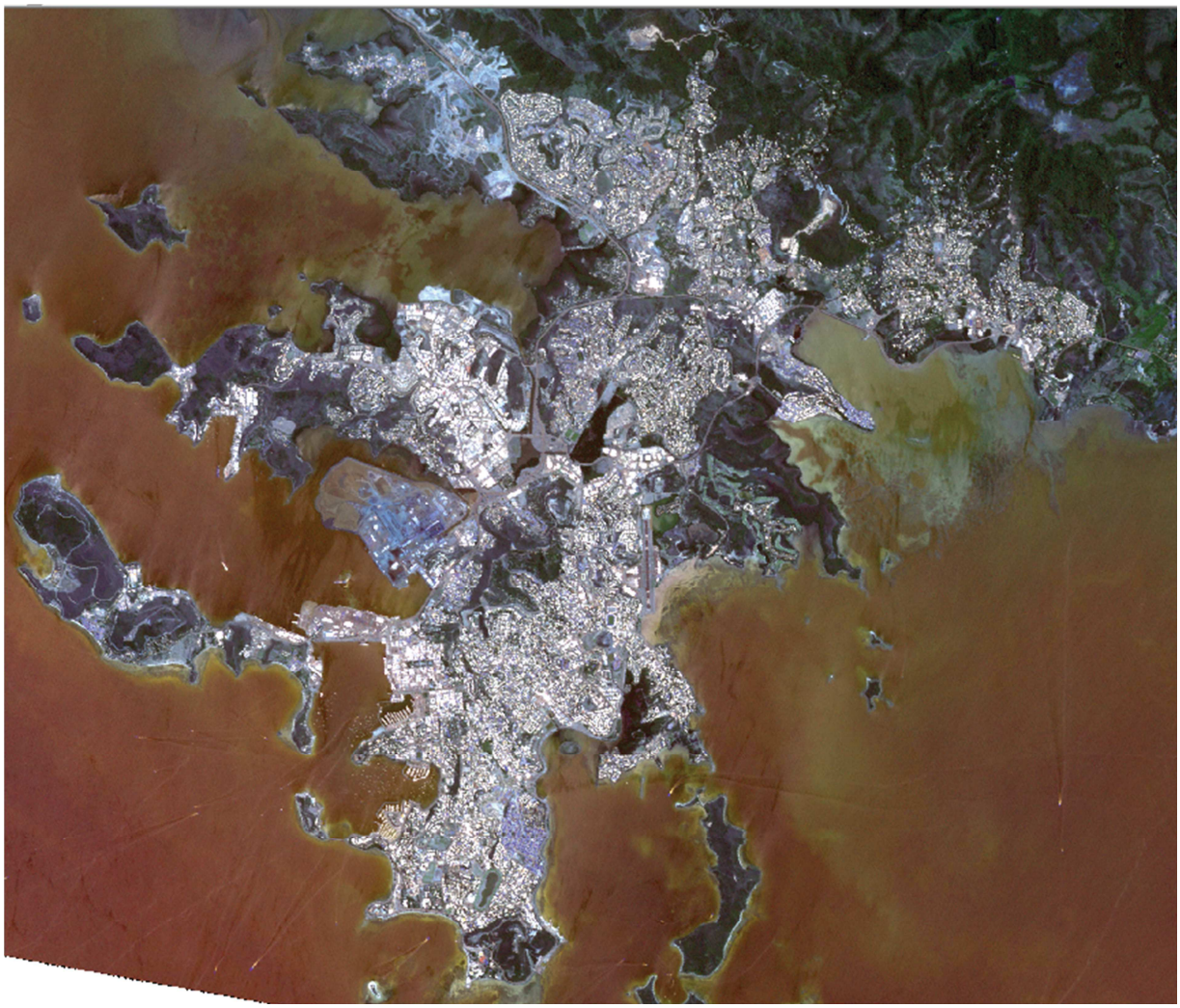


Figure 2.12 : Image satellite RapiEye de Nouméa *(image du 16/10/2010, résolution 4,5m).*

2.2.8. Données topologiques

Les données topologiques proviennent de la base de données topologique de la Direction des Infrastructures, de la Topographie et des Transports Terrestres (DITTT) de la Nouvelle-Calédonie. Elles sont disponibles sous la forme de couches d'informations cartographiques au format « *shapefile* ». Elles comprennent de nombreuses informations comme l'emplacement et le type des bâtiments, la présence de points d'eau, l'emplacement de ruines et de cimetières, les infrastructures routières, ou l'occupation des sols.

Une réflexion collégiale, dans laquelle nous avons listé l'ensemble des facteurs pouvant influencer la transmission de la dengue à une échelle locale nous a conduits à construire, à partir de cette base, cinq indicateurs potentiels du risque de transmission, listés dans la Table 2.9. L'hétérogénéité des indicateurs au sein d'un même quartier nous a conduits à réaliser une étude à une résolution plus fine : celle de l'îlot.

Table 2.8 : Indicateurs de télédétection pour l'étude spatiale à Nouméa en 2009

Indicateur	Calcul	Commentaire
NDVI	$\frac{(PIR - IR)}{(PIR + IR)}$	PIR = réflectance dans le proche infrarouge IR = réflectance dans l'infrarouge Caractérisation de la couverture végétale -1≤ NDVI < 0 : eau ou nuages NDVI = 0 : sol nu 1 ≥ NDVI > 0 : Biomasse végétale
NDVI(RE)	$\frac{(RE - R)}{(RE + R)}$	RE = réflectance dans le « red edge » R = réflectance dans le rouge Permet une classification de la végétation plus fine qu'avec le NDVI
TNDVI	$= \sqrt{NDVI + L}$	L = 0,5 Valeurs comprise entre 0 et 1 Sa valeur augmente avec la densité du couvert végétal
PVI	$= \sqrt{(SR - R)^2 + (SPIR - PIR)^2}$	SR = réflectance du sol nu dérivé de la droite des sols nus dans le rouge R = réflectance dans le rouge SPIR = réflectance du sol nu dérivé de la droite des sols nus dans le proche infrarouge PIR réflectance dans le proche infrarouge Il caractérise l'humidité des sols. PVI < 0 : sols hydromorphes PVI = 0 : sols nus PVI > 0 : végétation (augmente en fonction de la surface foliaire)
BI	$= \sqrt{PIR^2 + R^2}$	PIR = réflectance dans le proche infrarouge R = réflectance dans le rouge Indice de contraste
NDWI	$= \frac{(G - PIR)}{(G + PIR)}$	G = réflectance dans le vert PIR = réflectance dans le proche infrarouge Il caractérise les zones humides. -1≤ NDVI < 0 : Végétation ou sol nu 1 ≥ NDVI > 0 : Eau

Table 2.9 : Indicateurs géographiques pour l'étude spatiale à Nouméa en 2009

Indicateurs	Description
Densité d'adresse	Nombre d'adresses répertoriées dans un rayon de 100m autour de n'importe quel point de l'espace. Des moyennes par îlot ont ensuite été calculées
Densité de population	Nombre d'habitants par km², calculée par îlot
Pourcentage de surface bâtie	Pourcentage de surface bâtie dans chacune des parcelles cadastrales. Des moyennes par îlot ont ensuite été calculées.
Distance aux points d'eau	Pour chaque point de l'espace, distance au point d'eau le plus proche. Des moyennes par îlot ont ensuite été calculées
Densité de route	Pour chaque point de l'espace, densité de route dans un rayon de 100m. Des moyennes par îlot ont ensuite été calculées. Cet indicateur a été créé pour indiquer un éventuel morcellement de l'espace qui freinerait la diffusion spatiale des moustiques vecteurs.

2.3. Outils d'analyse et de modélisation

Le but de ce paragraphe n'est pas d'effectuer une liste exhaustive des outils de modélisation utilisables, mais d'établir une classification des outils disponibles en fonction du type et de l'objectif de l'étude. La plupart des outils de modélisation utilisés dans cette thèse étant déjà largement décrits par ailleurs (modèles de régression linéaire, modèles de classification…), nous décrirons, pour les modèles utilisés dans cette thèse, le minimum nécessaire à la bonne compréhension des résultats. Nous ne détaillerons qu'un seul des outils utilisés : la décomposition orthogonale aux valeurs propres (EOF pour « Empirical Orthogonal Function »), un outil qui permet l'extraction de schémas spatio-temporels. Nous avons décidé de détailler cet outil car son utilisation dans le domaine de l'épidémiologie est encore actuellement limitée, et il n'existe, à notre connaissance, que deux études ayant déjà employé cette technique en épidémiologie (Ortíz Bultó et al., 2006; Shi et al., 2014).

2.3.1. Les approches possibles en modélisation

Nous rappelons qu'il existe deux types d'approche pour améliorer la compréhension des facteurs influençant le cycle de transmission d'une maladie. La première approche part des données observées et utilise des outils statistiques pour explorer la variabilité spatio-temporelle des cas d'une maladie et découvrir des liens avec des facteurs explicatifs ou prédictifs (climatiques, socio-économiques…). La deuxième approche consiste à modéliser le cycle de transmission au niveau des individus afin de simuler des variables épidémiologiques (ou entomologiques pour les maladies vectorielles) qui sont ensuite comparées aux données observées afin de tester la validité des hypothèses sur lesquelles sont construits les modèles ou afin d'estimer certains paramètres de transmission. En ce qui concerne la dengue, voir (Degallier et al., 2010; Favier et al., 2005; Ferguson et al., 1999; Focks et al., 1995; Hay et al., 2000; Hopp and Foley, 2003; Smith et al., 2014a; Wearing and Rohani, 2006).

La première approche permet une exploration rapide des données et des liens existant entre variables explicatives et distribution spatio-temporelle des cas. Elle ne nécessite pas de formuler d'hypothèse sur le mécanisme d'action des variables sur le cycle de transmission : si le lien existe, les modèles statistiques le trouvent. Cette approche est utile dans un premier temps pour s'approprier les données, et mieux connaître la maladie étudiée. En revanche elle ne permet pas la compréhension fine des mécanismes en jeu, et le signal entre une variable donnée et la variable épidémiologique peut être difficile à observer en présence d'interactions entre les différentes variables explicatives. Cette approche est utile à la construction de modèles mécanistiques car elle permet l'identification des facteurs à intégrer à ces modèles.

La deuxième approche permet de construire des modèles qui ont l'avantage :

- d'expliciter les mécanismes de transmission ;

- de permettre de tester de nombreuses hypothèses sur les mécanismes de transmission en faisant varier la structure ou les paramètres des modèles : ces modèles sont extrêmement utiles pour distinguer l'effet de variables dites « intrinsèques », *i.e.* directement liées aux interactions hôte/vecteur/virus (immunité de population, période d'incubation extrinsèque) de l'effet des variables dites « extrinsèques », *i.e.* l'environnement dans lequel évoluent les hôtes, les vecteurs et les agents pathogènes (Koelle and Pascual, 2004). Ces modèles sont particulièrement utiles pour l'étude de l'immunité de population, car il existe peu de jeux de données spatio-temporels observés de manière régulière dans le temps et permettant une analyse empirique (Salje et al., 2012) ;
- de pouvoir intégrer des informations complexes comme les mouvements humains ou les contacts sociaux (Eubank et al., 2004; Smith et al., 2014a);
- d'avoir une bonne capacité d'extrapolation et de pouvoir être utilisés pour effectuer des prédictions dans le futur ou à de large échelles spatiales (Degallier et al., 2010; Hopp and Foley, 2003) ;
- de permettre de simuler l'effet de différentes stratégies de contrôle (Hethcote, 1989).

Les limites à l'utilisation des modèles mécanistiques sont la nécessité d'effectuer un calibrage des paramètres, une difficulté à intégrer l'une des deux dimensions spatiale ou temporelle, et des simulations parfois chronophages et nécessitant une grande puissance de calcul, en particulier pour les modèles basés sur l'individu (IBM ou ABM pour « Individual/Agent-Based Models »), ou les modèles basés sur des graphes (Eubank et al., 2004).

2.3.1. Définition des cibles de modélisation

En épidémiologie, il est possible d'étudier soit les phénomènes d'émergence d'une épidémie, soit la sévérité des épidémies. Les données épidémiologiques peuvent donc être agrégées de différentes manières en fonction de l'objectif.

Indicateurs de sévérité des épidémies

Le premier indicateur est l'incidence de la maladie. L'incidence est le nombre de nouveaux cas apparaissant par unité de temps dans une zone géographique donnée. Les incidences peuvent être hebdomadaires, mensuelles, annuelles, en fonction de la résolution temporelle d'agrégation des données. Cet indicateur est utile dans la planification opérationnelle des actions de lutte, car il permet d'estimer l'effort de lutte dans chacune des zones géographiques.

Le nombre de cas d'une maladie apparaissant dans une zone géographique donnée est proportionnel au nombre d'habitants de la zone étudiée. Ainsi, afin de pouvoir comparer les niveaux d'incidence entre différentes zones géographiques, ou entre différents pas de temps, il est nécessaire de rapporter ce

nombre à la population à risque, c'est le taux d'incidence. Le taux d'incidence est l'indicateur le plus utilisé dans les études épidémiologiques, car il reflète l'intensité de la circulation virale, et donc le cycle de transmission de la maladie. Les séries temporelles de taux d'incidence sont appelées des courbes épidémiques.

Lorsque la maladie touche préférentiellement certaines populations (par exemple maladies touchant plus les femmes, ou maladies touchant préférentiellement les personnes âgées), les taux d'incidence peuvent être influencés par la structure démographique de la population sous-jacente (âge, sexe). Ainsi, si deux régions ont des structures démographiques différentes, les taux d'incidence peuvent être plus élevés dans une des deux régions alors que l'intensité de la transmission de l'agent pathogène est la même. Par exemple, si une maladie touche préférentiellement les personnes âgées, à transmission égale, les régions possédant une proportion plus importante de personnes âgées présenteront des taux d'incidence naturellement plus élevées que les régions dont la population est plus jeune. L'étude des cycles de transmission nécessite alors de standardiser les taux d'incidence, de manière à prendre en compte l'effet de la structure démographique sous-jacente. Différentes méthodes de standardisation existent. Le principe est d'utiliser une population de référence de structure démographique donnée arbitraire (méthode de standardisation directe) ou existante (méthode de standardisation indirecte), et de calculer les taux d'incidence attendus dans chacune des régions si les populations possédaient toutes cette même structure à chacun des pas de temps (Ahmad et al., 2001).

L'ensemble des indicateurs de sévérité des épidémies sont des variables quantitatives, et nécessitent donc des outils de modélisation adaptés à l'étude de ces variables.

Indicateurs de présence d'épidémie ou de foyer épidémique

Certains épidémiologistes définissent une épidémie (chez l'Humain), ou épizootie (chez l'animal) « lorsque l'incidence d'une maladie subit de **fortes variations**, au cours de périodes successives, partant d'un niveau faible, voire nul, pour augmenter considérablement (*accélération*) avant de se stabiliser et de décroitre pour parfois s'annuler » (Toma et al., 2001). D'autres auteurs utilisent le terme « épidémie » pour désigner « une maladie dont l'incidence est simplement nettement supérieure à celle attendue dans une région pendant une période donnée » (Toma et al., 2001). Cette définition est volontairement subjective et non quantitative, puisque chaque pays définit lui-même « l'incidence attendue », ainsi que les seuils épidémiques en fonction de ses objectifs, de ses possibilités d'action, de la contagiosité ainsi que de la gravité de la maladie. Par exemple, pour une maladie extrêmement létale comme Ebola, un nombre restreint de cas, une dizaine, suffit à déclarer un statut épidémique. Pour une maladie moins grave comme la grippe, le même nombre restreint de cas sur une zone géographique limitée peut constituer un foyer épidémique, mais n'est pas suffisant pour déclencher des alertes épidémiques nationales.

Les indicateurs de seuil épidémique peuvent être définis de deux manières : soit de manière empirique, à partir des données observées, en établissant des seuils quantitatifs sur taux d'incidence, soit de manière théorique, en fonction du potentiel de transmission de la maladie.

Différents seuils empiriques sont utilisés à travers le monde (voir Costagliola et al., 1991; Coutinhoa et al., 2006; Cullen et al., 1984; Farrington et al., 1996; Heisterkamp et al., 2006; Held et al., 2005; Höhle and Paul, 2008; Höhle et al., 2009; Rossi et al., 1999; Serfling, 1963; Stroup et al., 1989; Turbelin et al., 2009 pour quelques exemples de définition de seuils épidémiques). Il n'existe pas de définition universelle, et chaque situation nécessite une réflexion sur le seuil le mieux adapté.

La définition théorique du potentiel de transmission d'une maladie passe par la notion de nombre de reproduction de base, noté « *R0* ». Le nombre de reproduction de base est défini comme le nombre de cas secondaires apparaissant suite à l'introduction d'une personne infectieuse dans une population entièrement susceptible au cours de la période pendant laquelle cette personne reste infectieuse (Anderson and May, 1999). Si ce nombre est inférieur à 1, le nombre de cas ne peut que décroître au cours du temps, et si la maladie est absente de la zone géographique étudiée, elle ne peut émerger. Si ce nombre est supérieur à 1, le nombre de cas croît au cours du temps : une épidémie peut potentiellement se développer dans une zone d'où elle était absente.

Le *R0* peut être déterminé de différentes manières. Il peut être calculé de manière empirique, à partir des courbes épidémiques observées uniquement (Vynnycky et al., 2007; Ward et al., 2009), ou à partir d'études de séroprévalence (Farrington et al., 2001). Le plus souvent, il est estimé en construisant d'abord un modèle de transmission, en dérivant une formule permettant de le calculer à partir des paramètres du modèle, et en estimant ces paramètres à partir de courbes épidémiques observées (Britton, 2001; Degallier et al., 2005; Farrington et al., 2001; Favier et al., 2006b; Ferrari et al., 2005; Massad et al., 2009). Voir Dietz *et al.* (Dietz, 1993) pour une revue des différentes méthodes d'estimation.

En pratique, il est plus aisé de travailler avec le nombre de reproduction effectif, noté « *R* », qui tient compte de l'évolution de la proportion d'hôtes susceptibles au cours du temps. Il représente le nombre de cas secondaires apparaisant suite à l'introduction d'une personne infectieuse dans une population partiellement immune. Ce nombre peut être estimé en temps réel à partir de courbes épidémiques, si le temps de génération est connu, le temps de génération étant l'intervalle de temps entre le moment de l'infection du cas primaire et les infections des cas secondaires dues à ce cas primaire (Cauchemez et al., 2006; Fine, 2003; Wallinga and Teunis, 2004).

Ces nombres sont des variables quantitatives. Des indicateurs du potentiel d'émergence d'épidémies ou de foyers épidémiques peuvent être construits à partir des nombres de reproduction en utilisant la valeur seuil de 1. La valeur de ces nombres est liée à la sévérité des épidémies, et ils pourraient éventuellement

être utilisés pour l'étude de la sévérité des épidémies. Cependant le lien entre ces nombres et la sévérité des épidémies n'est pas linéaire (Woolhouse and Gowtage-Sequeria, 2005). Il est donc préférable, lors des études sur la sévérité des épidémies, d'utiliser directement les indicateurs liés à l'incidence.

2.3.2. Les études possibles lors de l'analyse statistique de champs spatio-temporels

2.3.2.a. Etudes possibles

A partir des données décrites précédemment, un grand nombre d'études peuvent être effectuées. Ces études peuvent être séparées en deux catégories : i) des études descriptives de la variabilité spatio-temporelle d'un jeu de donnée, en l'occurrence les données épidémiologiques ; ii) des études analytiques permettant de détecter les associations entre une variable réponse (encore appelée variable dépendante, ou cible de modélisation, *i.e.* les données épidémiologiques) et une ou plusieurs variables explicatives (indifféremment appelées co-variables, facteurs, facteurs explicatifs, variables indépendantes). Les études analytiques peuvent être soit spatiales, soit temporelles, soit spatio-temporelles (Figure 2.13).

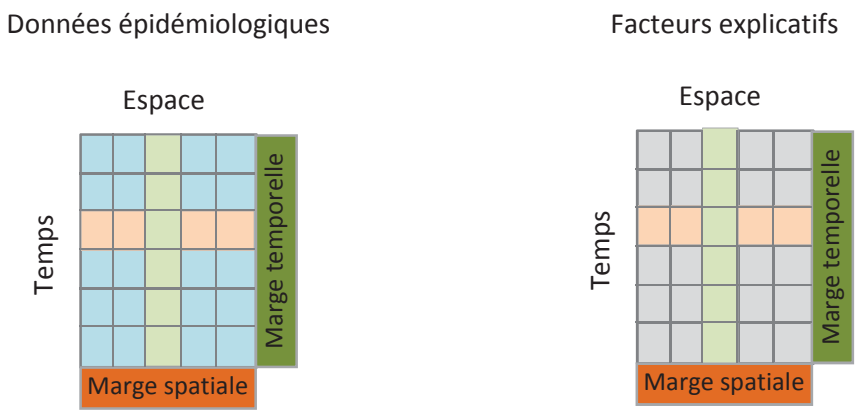


Figure 2.13 : Etudes possibles lors de l'analyse de champs spatio-temporels. *Le schéma représente les champs de données spatio-temporelles analysables quelle que soit l'échelle spatiale d'étude (régionale, nationale ou locale). Les couleurs correspondent aux types d'études possibles : les études descriptives de la variabilité spatio-temporelle d'un champ de données (en bleu), et les études analytiques entre une variable réponse (ici une variable épidémiologique) et une ou plusieurs variables explicatives. Les études analytiques peuvent porter soit sur des associations spatiales (en rouge), soit sur des associations temporelles (en vert). Les études temporelles peuvent être réalisées soit dans chacune des entités spatiales ayant servi à l'agrégation des données (en vert clair), soit sur les données moyennées sur l'ensemble des entités spatiales (analyse des marges moyennes, en vert foncé). De même les études spatiales peuvent porter sur les moyennes sur l'ensemble des pas de temps (rouge foncé), ou sur un pas de temps donné, le plus souvent une année (en rouge clair)*

2.3.2.b. Etudes non abordées au cours de la thèse

Pour les études analytiques, nous nous sommes limités, dans cette thèse aux analyses présentées dans la Figure 2.13 à chacune des échelles spatiales, et nous n'avons donc pas abordé :

- L'analyse conjointe de plusieurs champs spatio-temporels de données, i.e. les études analytiques intégrant à la fois la dimension spatiale et la dimension temporelle des données ;
- La prise en compte des phénomènes de diffusion inhérents aux maladies contagieuses. Par exemple, l'étude analytique temporelle des facteurs de risque d'émergence d'un foyer épidémique dans une commune donnée sur le territoire nécessite la prise en compte des variables explicatives de la commune étudiée, mais également la prise en compte de la présence de foyers épidémiques dans les communes voisines.

Pour les études à l'échelle spatiale de Nouméa, une fois les cas de dengue géo-référencés par des coordonnées géographiques, ces données auraient pu être traitées comme des processus ponctuels aléatoires, qui nécessitent l'utilisation d'outils particuliers (Gatrell et al., 1996). Cependant, nous avons préféré les agréger à une résolution spatiale très fine afin d'homogénéiser les analyses, mais également parce que les outils d'analyse et de modélisation des données spatiales agrégées sont plus nombreux et plus flexibles, et parce que les analyses en processus ponctuels permettent difficilement de distinguer les effets des variables explicatives de l'effet de la densité de population sur le risque d'apparition d'un cas à un endroit donné (Elliott et al., 2001).

2.3.3. Modélisation statistique au cours des études descriptives

2.3.3.a. Objectifs des études descriptives

L'objectif des études descriptives est l'identification de schémas spatio-temporels récurrents : présence de cycles dans les données, de processus de propagation, de structures spatiales ou temporelles répétitives. La présence de schémas récurrents renforce l'idée de l'existence d'un déterminisme dans la distribution spatio-temporelle des données. L'objectif des études analytiques est d'identifier les facteurs statistiquement associés à la variabilité spatiale ou temporelle de la variable réponse. Cependant la mise en évidence d'un lien statistique ne suffit pas à déterminer un lien de causalité.

2.3.3.b. Détermination des processus de propagation

Peu de méthodes sont disponibles pour détecter les processus de propagation. Elles sont toutes basées sur le calcul de décalages temporels existant entre les courbes épidémiques à différents endroits. Elles sont basées soit sur le calcul d'une fonction non paramétrique de covariance spatiale (Bjørnstad and

Falck, 2001; Bjørnstad et al., 1999; Grenfell et al., 2001), soit sur des méthodes permettant d'extraire la phase de phénomènes cycliques, comme les analyses spectrales. Dans cette thèse, nous avons décidé d'utiliser une méthode spectrale, l'analyse en ondelettes (Torrence and Compo, 1998).

La Figure 2.14 illustre le principe de l'analyse en ondelettes, et la Figure 2.15 donne une exemple du résultat obtenu en utilisant la technique sur des données épidémiologiques de surveillance de la dengue au Cambodge de 2002 à 2008 (voir Annexe 6 (Teurlai et al., 2012)). Les équations utilisées sont décrites en Annexe 7, ainsi que dans plusieurs publications (Cazelles et al., 2005, 2007; Torrence and Compo, 1998).

Le principe de l'analyse en ondelettes d'une série temporelle $x(t)$ est le suivant :

- un signal cyclique d'une fréquence donnée f_0 est choisi : l'ondelette mère. Il existe différentes formes d'ondelettes, choisies en fonction de la forme du signal à analyser et des objectifs de l'analyse ;
- Pour chacun des pas de temps t du signal $x(t)$ à analyser, un coefficient $W(f_0, t)$ correspondant à la convolution entre le signal et la forme conjuguée de l'ondelette-mère est calculé. Ce coefficient est appelé coefficient d'ondelette et mesure la « ressemblance » entre l'ondelette et le signal à l'instant t. Le carré du coefficient d'ondelette représente la puissance du signal $x(t)$ dans la fréquence f_0 au temps t. On obtient ainsi une série temporelle des coefficients d'ondelette, de même longueur que le signal $x(t)$, correspondant à l'évolution temporelle de la puissance du signal $x(t)$ dans la fréquence f_0 ;
- Puis l'ondelette-mère est étirée, sa fréquence correspondant alors à une fréquence quelconque f, et le processus est alors répété, permettant d'obtenir une autre série temporelle de l'évolution au cours du temps de la puissance du signal $x(t)$ pour la fréquence f ;
- L'ensemble du carré des coefficients d'ondelette pour chacun des pas de temps et chacune des fréquences analysées est appelé « spectre d'ondelette ».

L'étude du spectre d'ondelette permet de déterminer les fréquences présentes dans le signal ainsi que leur évolution au cours du temps. A partir de ce spectre, il est également possible de filtrer le signal dans une fréquence f quelconque, en reconstruisant un signal filtré uniquement à partir des coefficients d'ondelette correspondant à cette fréquence. Par exemple, filtrer les séries temporelles de taux d'incidence dans la bande de fréquence annuelle permet d'extraire la composante épidémique saisonnière.

Lorsque l'ondelette mère choisie est de forme complexe, les coefficients d'ondelette peuvent s'écrire en fonction de $e^{i\theta(f,t)}$ où $\theta(f,t)$ représente la phase de la composante filtrée du signal $x(t)$ dans la fréquence f au temps t. Dans l'exemple donné, la phase permet de localiser les épidémies dans le temps. Par exemple, les instants t pour lesquels $\theta(f,t) = 0, f=1/an$ sont les pas de temps auxquels surviennent les pics épidémiques annuels.

En calculant la phase de plusieurs signaux localisés à des endroits différents, il est possible de calculer le décalage de phase entre les deux séries temporelles, ce décalage étant directement proportionnel au décalage temporel existant entre les deux signaux. En épidémiologie, la connaissance du décalage temporel entre les épidémies dans différents endroits permet ainsi de reconstituer la propagation de l'épidémie.

Les avantages de cette technique sont qu'elle permet d'analyser des signaux non stationnaires (i.e. dont la moyenne, la variance ou la fréquence varie au cours du temps), et que le calcul des phases est indépendant des niveaux d'incidence observés, et donc robuste aux biais de surveillance pouvant exister.

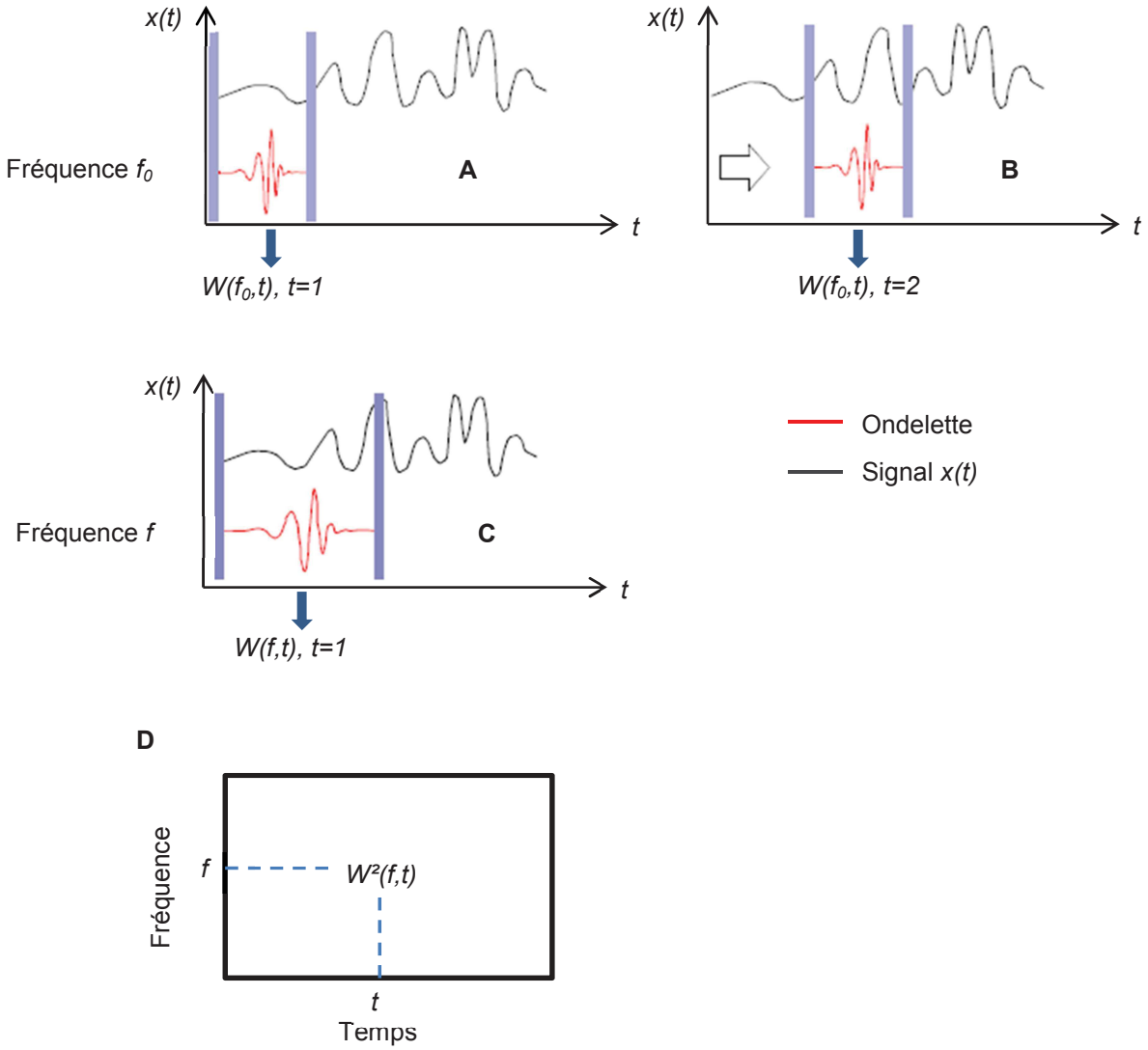


Figure 2.14 : Principe de l'analyse en ondelette. A. *Le coefficient d'ondelette $W(f_0,1)$ (convolution entre le signal et l'ondelette) est calculé pour la fréquence f_0 de l'ondelette mère au premier pas de temps.* ***B.*** *Puis l'ondelette est décalée dans le temps, et le coefficient $W(f_0,2)$ est calculé pour le pas de temps suivant, et ainsi de suite pour l'ensemble des pas de temps.* ***C.*** *L'ondelette est ensuite étirée (ou contractée) afin d'analyser une autre fréquence f, et le processus répété pour l'ensemble des fréquences d'analyse.* ***D.*** *L'ensemble du carré des coefficients d'ondelette est appelé spectre d'ondelette.*

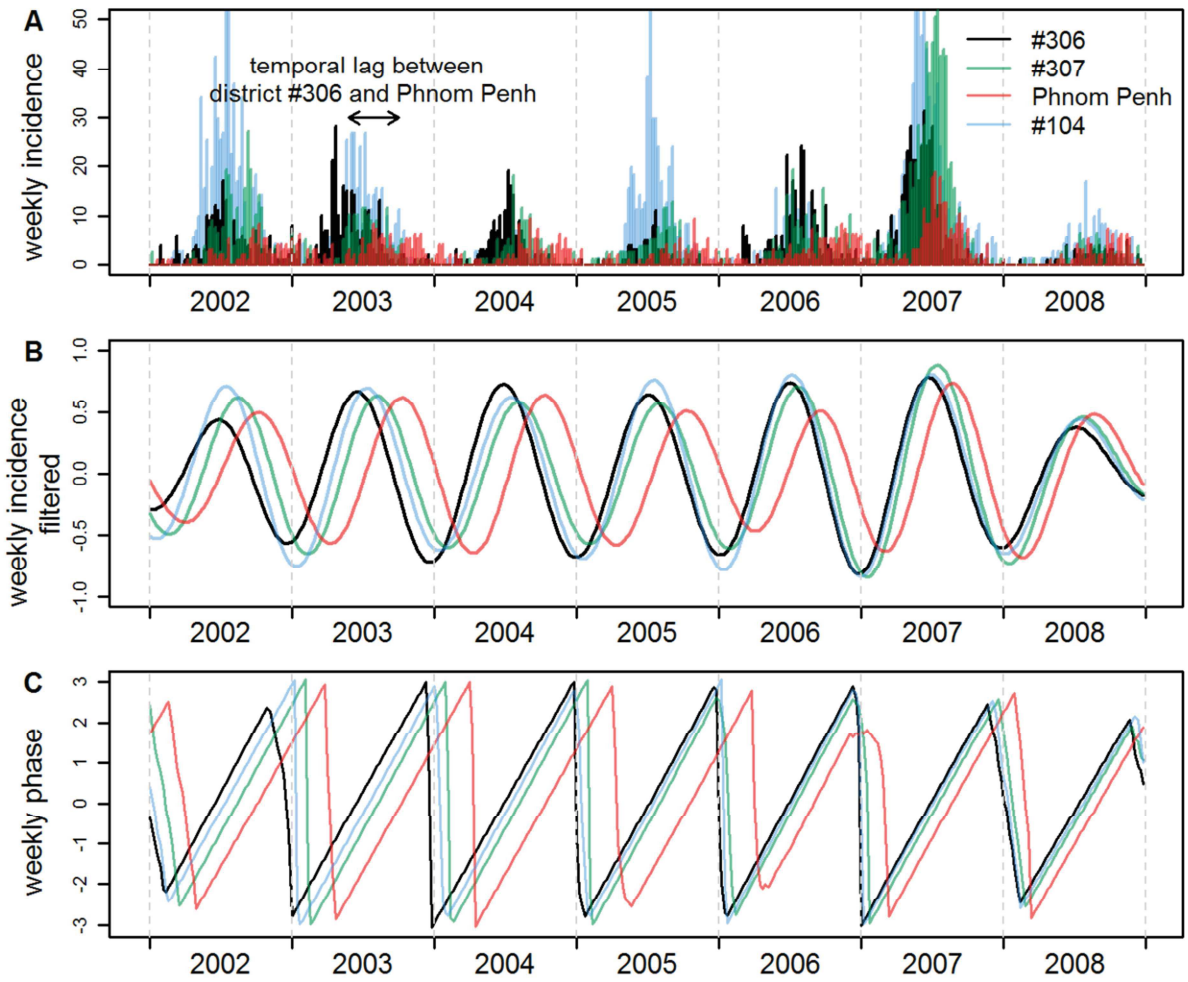


Figure 2.15 : Exemple de l'utilisation de la méthode des ondelettes pour la détection de processus de propagation. *Les données analysées sont les taux d'incidence hebdomadaires de la dengue dans quatre districts du Cambodge (Phnom Penh (rouge), et districts n°306 (noir), 307 (vert) et 104 (bleu).* ***A.*** *Séries temporelles analysées.* ***B.*** *Composante annuelle des séries temporelles de taux d'incidence, obtenues en filtrant les séries dans la bande périodique 0,8 – 1,2 an par la méthode des ondelettes.* ***C.*** *Phase de chacune des composantes épidémiques annuelles de la dengue. La différence de phase entre deux composantes annuelles données permet de mesurer le décalage temporel existant entre les deux épidémies correspondantes. Dans l'exemple, nous avons estimé que l'épidémie annuelle survient à Phnom Penh avec 4 mois de retard par rapport à l'épidémie du district n°306.*

2.3.3.c. Extraction de schémas spatiaux ou temporels récurrents

Périodicité et auto-corrélation des signaux temporels

Sur le plan temporel, les techniques spectrales d'analyse des signaux dans le domaine fréquentiel permettent d'étudier les cycles des séries temporelles. Ces techniques comportent des techniques adaptées à l'analyse de signaux stationnaires, comme la décomposition de Fourrier, ou des techniques adaptées à l'analyse de signaux non stationnaires, comme l'analyse en ondelette (Torrence and Compo,

1998) ou la décomposition empirique des modes (EMD pour « Empirical Mode Decomposition », (Cummings et al., 2005; Huang et al., 1998)), qui permet également l'analyse de signaux non-linéaires.

Sur le plan temporel toujours, il existe également des outils permettant de réaliser des analyses dans le domaine temporel, comme la fonction d'auto-corrélation, ou la fonction d'auto-corrélation croisée, qui calculent la corrélation entre un signal et lui-même (autocorrélation) ou entre un signal et un autre signal (auto-corrélation croisée), en décalant les deux séries d'un ou plusieurs pas de temps.

Détection de cluster et auto-corrélation spatiale

En épidémiologie, la contagiosité des maladies introduit souvent une agrégation spatiale des cas, et une auto-corrélation spatiale des taux d'incidence, puisque l'apparition d'un cas à un endroit donné favorise l'apparition de cas dans les endroits voisins.

Nous n'avons pas, dans cette thèse, utilisé de méthodes de détection de cluster, mais plusieurs méthodes sont disponibles, utilisant en général des algorithmes qui scannent la zone spatiale d'étude, évaluent la probabilité d'apparition de cas à différentes endroits, et donnent une mesure de si cette probabilité est significativement plus élevée qu'une valeur de référence (Elliott et al., 2001).

La présence d'auto-corrélation spatiale peut être détectée au moyen de divers indices (indice de Moran, indice de Geary, indice Tango) (Elliott et al., 2001). Ces indices donnent une indication de la présence/absence d'auto-corrélation, mais ne donnent pas d'information sur son éventuelle étendue spatiale. Pour connaître l'étendue spatiale de l'auto-corrélation des données, il est nécessaire d'utiliser soit des variogrammes, soit des corrélogrammes de Moran. Les corrélogrammes de Moran sont basés sur la définition d'un graphe de voisinage entre les entités spatiales, et le corrélogramme donne l'étendue de l'auto-corrélation en nombre de lags dans le graphe (Bivand et al., 2013). Si une auto-corrélation est présente, les coefficients de corrélations diminuent avec la distance. Nous avons, dans cette thèse, préféré utiliser des semi-variogrammes, qui donnent l'étendue de l'auto-corrélation directement en unité de distance de la carte utilisée. Le semi-variogramme empirique est défini comme l'évolution de la semi-variance d'une variable spatiale en fonction de la distance. La distance étant continue, la semi-variance est en fait calculée en discrétisant les distances en différentes classe h. Le principe de calcul du semi-variogramme empirique d'un champ spatial $x(s)$, $s = 1, 2, \ldots, p$, avec p le nombre d'entités spatiales des données, est le suivant : pour chaque paire de points $(i), x(j), i, j \in 1,2, \ldots p, i \neq j$, deux valeurs sont calculées : d_{ij} la distance entre les points $x(i)$ et $x(j)$, et $\left(x(i) - x(j)\right)^2$. La valeur $\gamma(h)$ du semi-variogramme du champ saptial $x(s)$, $s = 1, 2, \ldots, p$, pour la classe de distance h est alors définie comme :

$$\gamma(h) = \frac{1}{2N} \sum_{d_{ij} \in h} \left(x(i) - x(j)\right)^2$$

où N représente le nombre de paires de points séparés par une distance appartenant à la classe de distance *h*. S'il n'y a pas d'auto-corrélation, les valeurs de *x(i)* et *x(j)* sont indépendantes l'une de l'autre, et les semi-variances sont stables quelle que soit la classe *h* de distance considérée. Si une auto-corrélation est présente, les valeurs de *x(i)* et *x(j)* sont d'autant plus proches (et donc la semi-variance faible) que la distance d_{ij} séparant les deux points est faible.

2.3.4. Modélisation statistique au cours des études analytiques

2.3.4.a. Objectifs de l'étude

L'objectif des études analytiques est l'identification des facteurs de variabilité spatiale ou temporelle des données épidémiologiques. Ces études reposent sur l'utilisation de modèles de régression (si la variable réponse est quantitative) ou de classification (si la variable réponse est qualitative), qui permettent de quantifier les associations statistiques existant une variable réponse (un des indicateurs épidémiologiques) et une ou plusieurs variables explicatives.

2.3.4.b. Etude des facteurs liés à la sévérité des épidémies

L'étude de la sévérité des épidémies nécessite la modélisation d'indicateurs quantitatifs (incidence, taux d'incidence,…), et utilisent donc des modèles de régression.

L'étude de la qualité du modèle passe par le calcul de l'écart moyen entre les valeurs de la variable réponse prédite par le modèle et les valeurs observées. Différents indicateurs d'erreur sont utilisables, en particulier l'erreur quadratique moyenne, l'erreur relative quadratique moyenne ou encore le coefficient de détermination (Scherrer et al., 2009).

Deux grandes catégories d'outils de régression existent : les modèles paramétriques de régression, et les modèles non paramétriques d'apprentissage automatique.

Les modèles paramétriques de régression reposent sur certaines hypothèses, en particulier, les résidus des modèles doivent avoir une distribution normale et une variance constante (hypothèse d'homoscédasticité), et les données utilisées doivent être indépendantes. En épidémiologie, les données d'incidence suivent une distribution de Poisson, et des modèles linéaires généralisés doivent être utilisés afin de prendre en compte la particularité de cette distribution. La présence d'auto-corrélation spatiale peut entraîner une sur-dispersion des données, c'est-à-dire que la variance des données, au lieu d'être égale à la moyenne comme c'est normalement le cas pour des données suivant la distribution de Poisson, est supérieure à la moyenne.

Différents moyens sont possibles pour prendre en compte l'auto-corrélation et gérer la sur-dispersion des données : utilisation de modèles à effets mixtes (Bolker et al., 2009; Elston et al., 2001), utilisation de modèles explicitant clairement la dépendance spatiale ou temporelle entre les données (Dormann et al., 2007), ou encore utilisation de modèles linéaires généralisés basés sur la distribution négative binomiale (mais ces modèles ne permettent de prendre en compte que la sur-dispersion des données, et non l'auto-corrélation) (Elston et al., 2001).

De plus l'estimation des paramètres des modèles paramétriques peut devenir instable lorsque les variables explicatives sont corrélées entre elles. Voir Dormann *et al.*(Dormann et al., 2013) pour une revue des moyens disponibles pour gérer la multi-colinéarité entre variables explicatives.

Les modèles non paramétriques d'apprentissage automatique supervisé sont basés sur l'utilisation d'algorithmes permettant de construire un lien de manière automatique entre une variable réponse et une ou plusieurs variables explicatives. Ces modèles ont l'avantage de ne reposer sur aucune hypothèse concernant la distribution de la variable réponse. L'absence de paramètres rend ces modèles extrêmement flexibles, et permet le plus souvent d'établir des liens non linéaires sans avoir à déterminer au préalable la forme de ce lien. Le corollaire est que les seules sorties de ces modèles sont les valeurs prédites de la variable réponse : il n'y pas de paramètre à interpréter, et l'exploitation des résultats nécessite donc une étape supplémentaire d'exploration de la nature du lien entre variables explicatives et variables réponses.

Ces modèles peuvent présenter des problèmes de sur-apprentissage, c'est-à-dire que le lien construit peut être trop spécifique du jeu de données analysé, et non extrapolable à d'autres contextes. Pour éviter le sur-apprentissage, il est nécessaire d'effectuer une validation de ces modèles, en découpant le jeu de données en un jeu d'apprentissage et un jeu de validation dans lequel les individus statistiques (c'est-à-dire les entités spatiales ou les entités temporelles dans notre cas) sont répartis de manière aléatoire. Le lien entre variable réponse et variables explicatives est alors construit sur le jeu d'apprentissage, puis le modèle est appliqué aux variables explicatives du jeu de validation afin de prédire les valeurs de la variable réponse. Les erreurs du modèle sont alors calculées en comparant les valeurs prédites aux valeurs observées du jeu de validation, jeu qui n'a pas servi à la construction du modèle. Plusieurs méthodes de validation existent : une méthode simple pour laquelle un seul jeu de validation est construit, et une méthode de validation croisée pour laquelle le processus de découpage du jeu de données en jeu d'apprentissage et jeu de validation est répété plusieurs fois (méthode « k-fold », où le jeu de données total est coupé en k partitions, chaque partition servant à tour de rôle de jeu de validation pendant que les partitions restantes servent de jeu d'apprentissage, ou bien méthode « leave-one-out », qui est un cas particulier de la méthode « k-fold » pour lequel le nombre k de partitions est égal au nombre d'individus statistiques du jeu de donné total). Ces méthodes diffèrent par le biais et la précision de l'estimation des erreurs du modèle (Kohavi, 1995).

Dans cette thèse, nous avons utilisé les modèles SVM (« Support Vector Machine ») (Cortes and Vapnik, 1995; Vapnik, 1998) dans les problèmes de régression, pour leur flexibilité, leur absence de sensibilité à la présence de colinéarité entre les variables explicatives et leur capacité à modéliser des liens non linéaires.

2.3.4.c. Etude des facteurs liés à l'émergence des épidémies

L'étude de l'émergence des épidémies nécessite la modélisation d'indicateurs qualitatifs souvent binaires (présence vs absence d'un foyer épidémique ou d'une épidémie), et utilisent donc des modèles de classification.

Il existe de nombreux critères de comparaison de la performance des modèles de classification (Hand, 2001). Le critère le plus utilisé est le calcul des erreurs de classification, qui passe par le calcul de tables de contingence appelées tables de confusion, dans laquelle les prédictions du modèle sont comptabilisées en tant qu'éléments bien classés ou éléments mal classés par comparaison aux valeurs observées. Différents indicateurs d'erreur sont utilisables : la sensibilité, la spécificité, la valeur prédictive positive, la valeur prédictive négative (voir glossaire, Annexe 1). L'indicateur global de performance le plus utilisé dans les problèmes de classification supervisée est l'aire sous la courbe ROC (« Receiver Operating Characteristic »), encore appelée AUC pour « Area Under the Curve » (Hanley and McNeil, 1982).

Comme pour les modèles de régression, il est possible d'utiliser soit des modèles paramétriques, soit des modèles non paramétriques d'apprentissage automatique, les éléments décrits pour chacune de ces classes de modèle au paragraphe précédent restant valides.

Dans cette thèse, nous avons utilisé les modèles SVM (Cortes and Vapnik, 1995; Vapnik, 1998) dans les problèmes de classification, pour leur flexibilité, leur absence de sensibilité à la présence de colinéarité entre les variables explicatives et leur capacité à modéliser des liens non linéaires.

2.3.5. Modèles mécanistiques : modélisation des processus de transmission au niveau individuel

Les outils permettant de modéliser les processus de transmission au niveau individuel incluent les modèles de dynamique de population (Anderson and May, 1999), les modèles basés sur des graphes (Eubank et al., 2004), les modèles de méta-population, les automates cellulaires ou encore les modèles basés sur l'individu (Smith et al., 2014a). Nous ne décrirons ici que les modèles de dynamique des populations.

2.3.5.a. Modèles de dynamique des populations de type « SIR »

Les modèles de dynamique de population (Anderson and May, 1999; Hethcote, 1989) découpent la population d'hôtes en différents compartiments en fonction de l'état des individus composant la population (sain, en période d'incubation, infectieux, immun…). Des équations différentielles intégrant différentes hypothèses sur la maladie (immunité, durée, mortalité induite, mode de transmission) permettent de décrire l'évolution temporelle de la proportion d'individus au sein de chaque compartiment. Les paramètres à intégrer sont en général des paramètres démographiques (mortalité, naissance), ainsi que des paramètres régissant les contacts entre hôtes sains et infectieux.

Ces modèles peuvent être utilisés dans une approche qualitative, afin de tester des hypothèses sur les mécanismes de transmission (par exemple pour la dengue, la durée de l'immunité transitoire (Hay et al., 2000; Wearing and Rohani, 2006), ou les modalités de contact entre hôtes (Smith et al., 2014a)). Dans ce cas, plusieurs modèles sont construits, basés chacun sur une hypothèse différente concernant les mécanismes de transmission, et les variables de sortie (taux d'incidence par exemple) sont comparées aux données observées pour savoir quelle hypothèse (*i.e.* quel mécanisme) est la plus plausible.

Une fois un modèle construit, sa structure validée et ses paramètres ajustés, il peut être utilisé de manière quantitative pour dériver un grand nombre d'informations épidémiologiques importantes, comme les taux d'incidence, le nombre de reproduction de base, le nombre de reproduction ou encore la capacité vectorielle. Ces modèles décrivent naturellement les phénomènes d'immunité de population et sont utiles pour étudier l'impact de ces phénomènes sur la transmission d'une maladie, et simuler de manière quantitative l'impact de différentes mesures de contrôle.

Ces modèles sont uniquement dynamiques, et s'il est possible d'y intégrer une dimension spatiale simple (Grenfell et al., 2001), ils peuvent difficilement être utilisés pour produire des cartes de risque. De plus, ils sont sensibles aux valeurs des paramètres utilisés.

2.3.5.b. Modèles de capacité vectorielle

Les modèles de capacité vectorielle ont été créés pour l'étude de la malaria (MacDonald, 1957), et sont adaptés à l'étude de l'émergence épidémique. Ces modèles décrivent la capacité du vecteur à répandre une maladie vectorielle donnée dans une population d'hôtes, en décrivant les interactions de transmission entre l'hôte, le virus et le vecteur, supposition faite de la présence de ces trois éléments sur la zone géographique considérée. La capacité vectorielle C est définie comme (Garrett-Jones, 1964; MacDonald, 1957) :

$$C = \frac{a^2 p_a p_b \exp(-\mu T_{eip})}{\mu}$$

où a est le nombre de piqures par jour par vecteur, p_a et p_b sont respectivement la probabilité qu'un hôte sain s'infecte lorsqu'il est piqué par une moustique infectieux, et la probabilité qu'un moustique sain s'infecte lorsqu'il pique un hôte infectieux, μ est le taux de mortalité journalière du vecteur, et T_{eip} est la durée d'incubation extrinsèque de la maladie. L'ensemble de ces paramètres dépend de facteurs environnementaux comme la température (Focks et al., 1993a; Liu-Helmersson et al., 2014).

La capacité vectorielle est étroitement liée au nombre de reproduction de base (à la durée de la virémie près (Patz et al., 1998)). Ce modèle est donc adapté à l'étude de l'émergence épidémique, et permet d'expliciter les mécanismes par lesquels les variables environnementales influencent le cycle de transmission des maladies vectorielles. Il est très facile à spatialiser. Un inconvénient majeur est qu'il ne permet pas d'intégrer les facteurs liés à l'hôte comme l'immunité de population.

2.3.6. Décomposition orthogonale aux valeurs propres (« Empirical Orthogonal Functions »)

Une question courante dans l'étude de l'évolution spatio-temporelle d'une maladie est celle de confirmer ou d'invalider l'existence de schémas spatiaux récurrents au cours d'épidémies successives. En effet, l'existence de tels schémas permet :

- D'identifier les zones à risque à cibler en priorité par les autorités de santé publique ;
- D'émettre des hypothèses sur les facteurs sous-tendant ces schémas spatiaux et les mécanismes en cause ;
- De mieux anticiper le risque épidémiologique et d'améliorer les modèles de prédiction.

A l'heure actuelle, l'identification de ces schémas spatiaux récurrents dans le domaine de l'épidémiologie se fait principalement par comparaison visuelle de cartes de taux d'incidence moyennes sur un certain nombre de périodes de temps successives. La tâche de comparaison peut s'avérer fastidieuse lorsque le nombre de cartes à comparer augmente (longue période d'étude).

Dans cette thèse, nous proposons d'adapter aux données épidémiologiques la technique de décomposition orthogonale aux valeurs propres (appelée « EOF » en anglais, pour Empirical Orthogonal Functions), couramment utilisée en météorologie pour l'extraction de schémas spatio-temporels (Hannachi et al., 2007). Cette technique a été formalisée dans le domaine de la météorologie en 1956 par Lorenz (Lorenz, 1956). Elle a permis depuis l'identification d'oscillations importantes du climat, comme l'oscillation de Madden-Julian (Madden and Julian, 1971). Une grande partie des informations fournies dans ce paragraphe proviennent de la revue de Hannachi (Hannachi et al., 2007), mais il existe évidemment de nombreux articles sur le sujet dans le domaine de la météorologie. Une certaine confusion existe dans la littérature du fait de la multitude de noms donnés aux différents éléments de l'analyse. Nous utiliserons ici les appellations de Hannachi (Hannachi et al., 2007).

L'analyse EOF est une méthode d'analyse exploratoire qui permet de décomposer un champ spatio-temporel d'anomalies $X^c(t, \mathbf{s})$ en une somme de schémas spatiaux *f(s)* (également appelés « EOF ») et de schémas temporels *f(t)* (composantes principales, « PC » pour « Principal Components ») :

$$X^c(t, \mathbf{s}) = \sum_{k=1}^{M} f(t) f(\boldsymbol{s})$$

où M représente le nombre de modes de variabilité des données.

Les schémas sont triés par ordre décroissant de la part de variance des données initiales qu'ils expliquent. Si des structures spatiales temporellement récurrentes sont présentes dans les données, la visualisation d'un nombre restreint d'EOF et de composantes principales permet de synthétiser une grande partie de l'information spatio-temporelle contenue initialement dans les données. En ce sens, ils sont très utiles pour réduire la dimension d'étude d'un champ spatio-temporel présentant un grand nombre de pas de temps ou un grand nombre d'entités spatiales. Dans une analyse EOF classique, les schémas spatiaux obtenus sont orthogonaux (*i.e.* la corrélation entre deux des schémas spatiaux obtenus est nulle, *i.e.* la matrice de covariance des schémas spatiaux est diagonale).

L'analyse EOF classique est strictement équivalente à une analyse en composante principale (ACP) non normée réalisée sur un champ spatio-temporel dont les entités temporelles représentent les individus (en ligne), et les entités spatiales représentent les variables (en colonne). Si les équations permettant l'extraction des schémas spatio-temporels sont établies depuis longtemps, la difficulté de l'utilisation de cette technique réside dans l'interprétation des schémas obtenus. Nous nous efforcerons donc dans cette thèse de donner une interprétation épidémiologique des différents éléments du résultat de l'analyse. A cette fin, l'ensemble des équations sera décrit en termes épidémiologiques bien que ces équations soient en réalité applicables à n'importe quelle matrice de données quantitatives.

En épidémiologie, à notre connaissance, cette technique n'a été utilisée que deux fois (Ortíz Bultó et al., 2006; Shi et al., 2014). Ortíz Bultó *et al.* l'ont utilisée comme une ACP classique pour filtrer les modes de variabilité des données épidémiologiques et climatiques avant d'effectuer des analyses statistiques de régression sur les modes principaux de variabilité. Shi *et al.* l'ont utilisée comme nous proposons de le faire, pour extraire les schémas spatiaux temporellement récurrents. Cependant, ils n'ont pas explicité l'interprétation des différents éléments sur un plan épidémiologique.

Notations et transformation des données

Nous noterons les matrices par des lettres majuscules, et les vecteurs par des lettres minuscules en gras. La transposée du vecteur ligne $\mathbf{x} = (x_1, x_2, \dots x_n), n$ quelconque sera notée $\mathbf{x}^T$, et la transposée de la matrice X sera notée X^T

Soit $X(t, \mathbf{s})$ un champ spatio-temporel de taux d'incidence où t représente les différents pas de temps et $\mathbf{s}$ un vecteur dénotant la position spatiale (exemple : longitude et latitude du centroïde d'une division administrative). Le taux d'incidence au temps $t_i,\ i = 1, 2, \dots, n$ dans la région spatiale $\mathbf{s}_j, j = 1, 2, \dots, p$ est noté x_{ij}. Le champ spatio-temporel des taux d'incidence est alors représenté par la matrice de données de dimension $n \times p$:

$$\begin{array}{cc} & \xleftrightarrow{\quad p \text{ régions spatiales} \quad} \\ & \mathbf{y}_j \\ X = & \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & \cdots & \cdots & x_{2p} \\ \vdots & \vdots & x_{ij} & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \begin{array}{c} \\ \mathbf{x}_i \\ \\ \\ \end{array} \begin{array}{c} \\ \updownarrow \ n \text{ pas de temps} \\ \\ \\ \end{array} \end{array}$$

où $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip}),\ i = 1, \dots, n$ représente la carte des taux d'incidence au temps $t = i$, et où $\mathbf{y}_j = (x_{1s}, x_{2s}, \dots, x_{ns})^T,\ s = 1, \dots, p$ représente la série temporelle des taux d'incidence dans la région spatiale s.

Soit $\overline{x}_{.j}$ la moyenne temporelle des taux d'incidence dans la $j^{ème}$ région spatiale calculée de la manière suivante :

$$\overline{x}_{.j} = \frac{1}{n} \sum_{i=1}^{n} x_{ij}$$

La carte de la moyenne temporelle des taux d'incidence est donnée par :

$$\overline{\mathbf{x}} = (\overline{x}_{.1}, \dots, \overline{x}_{.p}) = \frac{1}{n}\, \mathbf{1}_n\, X$$

où $\mathbf{1}_n = (1, \dots, 1)$ est le vecteur ligne de longueur n contenant uniquement des « 1 ».

Si à chacune des cartes spatiales de taux d'incidence $\mathbf{x}_i, i = 1, \dots, n$ nous retirons la carte spatiale moyenne $\overline{\mathbf{x}}$ (i.e. si nous centrons la matrice X), nous obtenons le champ spatio-temporel $X^C(t, \boldsymbol{s})$ des anomalies temporelles des taux d'incidence par rapport à la carte spatiale moyenne :

$$X^C = X - \mathbf{1}_n^T\, \overline{\mathbf{x}} = \begin{pmatrix} x_{11}^c & x_{12}^c & \cdots & x_{1p} \\ x_{21}^c & \cdots & \cdots & x_{2p}^c \\ \vdots & \vdots & x_{ij}^c & \vdots \\ x_{n1}^c & x_{n2}^c & \cdots & x_{np}^c \end{pmatrix}$$

où $x_{ij}^c = x_{ij} - \overline{x}_{.j},\ i = 1, \dots, n\,; j = 1, \dots, p$ représente l'écart entre le taux d'incidence observé dans la région j au temps i et le taux d'incidence moyen dans la région j.

Calcul des EOF

Le principe d'une analyse EOF ainsi que celui de l'ACP est de déterminer les m axes principaux du nuage de points, i.e de déterminer les axes pour lesquels la variance du nuage projeté sur ces axes est maximale, avec la contrainte de l'unité de la norme des vecteurs directeurs.

Soit U la matrice formée de m vecteurs spatiaux $\mathbf{u}_k = (u_{1k}, \ldots, u_{pk})^T$, $k = 1, 2, \ldots, m$, représentant les coordonnées des vecteurs $\mathbf{u}_k$ dans l'espace de même base que le champ $X^C(t, \boldsymbol{s})$:

$$U = \begin{array}{c} \mathbf{u}_k \\ \begin{pmatrix} u_{11} & u_{12} & \boxed{\cdots} & u_{1m} \\ u_{21} & \ddots & \boxed{\cdots} & u_{2m} \\ \vdots & \ddots & \boxed{u_{jk}} & \vdots \\ u_{p1} & u_{p2} & \boxed{\cdots} & u_{pm} \end{pmatrix} \end{array}, \qquad t.q. \begin{cases} \mathbf{u}_k^T \cdot \mathbf{u}_k = \mathbf{1} \; \forall \, k \\ \mathbf{u}_i^T \cdot \mathbf{u}_j = \mathbf{0} \; \forall \, i \neq j \end{cases}$$

Dans ce cas, les coordonnées de la projection du champ spatio-temporel $X^C(t, \boldsymbol{s})$ dans l'espace de base U sont définies par :

$$A = X^C U$$

Notons

$$A = \begin{array}{c} \mathbf{a}_k \\ \begin{pmatrix} a_{11} & \boxed{a_{12}} & \cdots & a_{1m} \\ a_{21} & \boxed{\ddots} & \cdots & a_{2m} \\ \vdots & \boxed{a_{ik}} & \ddots & \vdots \\ a_{n1} & \boxed{a_{n2}} & \cdots & a_{nm} \end{pmatrix} \end{array}$$

Chaque vecteur $\mathbf{a}_k = (a_{1k}, a_{ik}, \ldots, a_{nk})^T, i = 1, 2, \ldots, n$; $k = 1, 2, \ldots, m$ est appelé composante principale. C'est une série temporelle des coordonnées du champ spatio-temporel $X^C(t, \boldsymbol{s})$ projeté sur l'axe de vecteur directeur $\mathbf{u}_k$, ou sous forme matricielle :

$$\mathbf{a}_k = X^C \mathbf{u}_k$$

On a donc

$$a_{ik} = \sum_{j=1}^{p} x_{ij} \, u_{jk}$$

La variance de $\mathbf{a}_k = X^c\mathbf{u}_k$ peut s'écrire :

$$var(X^T\mathbf{u}_k) = \mathbf{u}_k^T X^{CT} X^C \mathbf{u}_k = \mathbf{u}_k^T S \mathbf{u}_k$$

où S est la matrice de covariance (temporelle) du champ spatio-temporel $X(u, \mathbf{s})$:

$$S = \frac{1}{n} X^{CT} X^C = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & \ddots & \dots & s \\ \vdots & s_{ij} & \ddots & \vdots \\ s_{p1} & s_{n2} & \dots & s_{pp} \end{pmatrix}$$

avec s_{ij} la covariance entre les séries temporelles des taux d'incidence de n'importe quelle paire de points spatiaux$(\mathbf{s}_i, \mathbf{s}_j)$, $i, j = 1, \dots, p$, soit :

$$s_{ij} = \frac{1}{n} \sum_{t=1}^{n} x_{ti}^c \, x_{tj}^c = \frac{1}{n} \mathbf{x}_i^c \, \mathbf{x}_j^c$$

Le problème de recherche des axes principaux revient donc à chercher les vecteurs $\mathbf{u}_k$, $k = 1, \dots m$ de manière à ce que $\mathbf{u}_k^T S \mathbf{u}_k$ soit maximal, avec la contrainte de norme unitaire des vecteurs, soit $\mathbf{u}_k^T \cdot \mathbf{u}_k = \mathbf{1} \; \forall \, k$. Ce problème peut se résoudre par la méthode des multiplicateurs de Lagrange. On cherche alors l'*extremum* de l'équation suivante :

$$L(\mathbf{u}_k) = \mathbf{u}_k^T S \mathbf{u}_k - \lambda_k(\mathbf{u}_k^T \cdot \mathbf{u}_k - \mathbf{1})$$

En utilisant la dérivée matricielle, on obtient :

$$\frac{\partial L(\mathbf{u}_k)}{\partial \mathbf{u}_k} = 2S\mathbf{u}_k - 2\lambda_k \mathbf{u}_k = 0$$

Les vecteurs $\mathbf{u}_k$, $k = 1, \dots m$ sont les vecteurs directeurs des axes principaux s'ils sont les solutions de l'équation :

$$S\mathbf{u} = \lambda \mathbf{u}$$

Les vecteurs unitaires $\mathbf{u}_k$, $k = 1, \dots m$ des axes principaux sont donc les vecteurs propres de la matrice de covariance des taux d'incidence, associés aux valeurs propres λ_k.

En multipliant à gauche les deux termes de l'équation $S\mathbf{u}_k = \lambda \mathbf{u}_k$ par $\mathbf{u}_k^T$, on remarque que $\mathbf{u}_k^T S \mathbf{u}_k = \lambda_k \mathbf{u}_k^t \mathbf{u}_k$, i.e.

$$var(X^c \mathbf{u}_k) = \lambda_k$$

La valeur propre λ_k est donc égale à la variance de la $k^{ième}$ composante principale et les valeurs propres sont donc positives.

Dans la littérature des sciences atmosphériques, les vecteurs $\mathbf{u}_k,\ k = 1,\ldots m$ sont appelés EOF, et les projections $\mathbf{a}_k = (a_{1k},\ldots,a_{nk})^T,\ k = 1,\ldots m$ du champ d'anomalie des taux d'incidence sur les axes principaux sont appelées composantes principales.

Il est courant de déterminer la part de variance expliquée par chacun des axes sous forme de pourcentage :

$$\%\ de\ variance\ expliquée\ par\ l'axe\ k = 100 \cdot \frac{\lambda_k}{\sum_{k=1}^{m} \lambda_k}\ \%$$

et de représenter ces pourcentages sous la forme d'un diagramme en bâtons. Contrairement à une analyse ACP classique où la présentation des résultats s'arrête à une représentation géométrique, dans l'analyse EOF, les vecteurs propres peuvent être représentés sous forme de cartes, et les composantes principales sous forme de séries temporelles.

Le champ spatio-temporel d'incidence de départ $X(t, \mathbf{s})$ peut être reconstruit intégralement uniquement à partir des vecteurs propres et des composantes principales :

$$X = \mathbf{1}_n{}^T \overline{\mathbf{x}} + AU^T$$

Le taux d'incidence de la région spatiale j au temps t peut être retrouvé par :

$$x_{tj} = \overline{x}_{.j} + \sum_{k=1}^{m} a_{tk}\, u_{jk}$$

Représentation géométrique pour p = 2

L'objectif de ce paragraphe est d'utiliser une représentation géométrique des données d'incidence pour faciliter l'interprétation des résultats de l'analyse EOF, et faire le lien avec l'interprétation classique des résultats d'une ACP normée. Le champ spatio-temporel $X(t, \mathbf{s})$ peut être représenté sous la forme d'un nuage de points des entités temporelles dans un repère formé par les entités spatiales. Afin de permettre des représentations dans un espace plan, nous prenons ici un exemple fictif où le champ $X(t, \mathbf{s})$ est composé de taux d'incidence pour $n = 6$ années et $p = 2$ régions spatiales (voir Table 2.10). Dans cet exemple, l'unité utilisée pour les taux d'incidence est arbitraire (mettons en nombre de cas / 100 habitants / an).

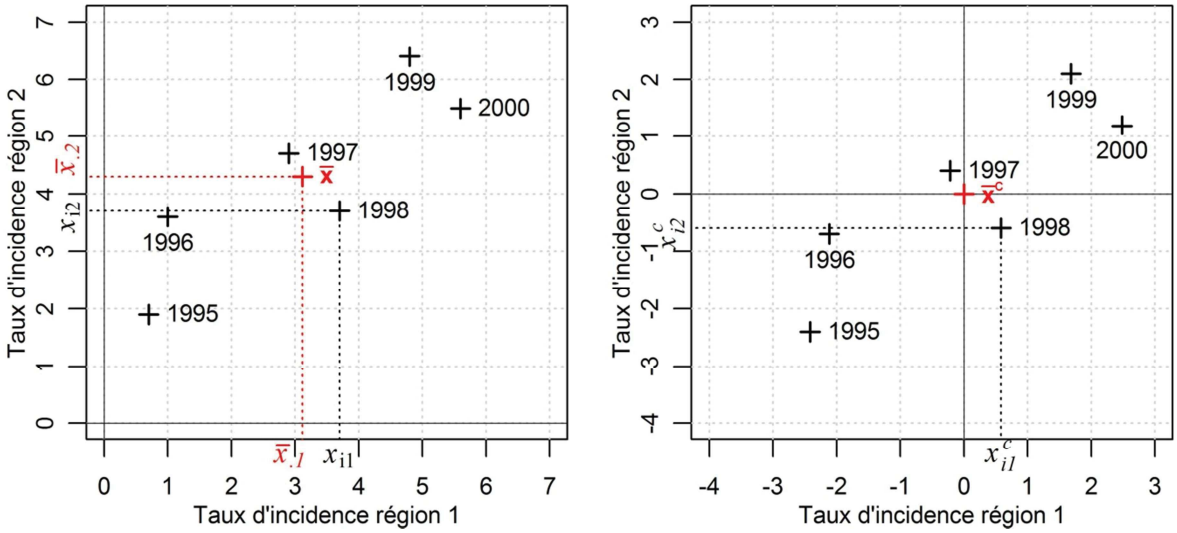


Figure 2.16 : Représentation géométrique des taux d'incidence spatio-temporels. A : *Champ X(t,s) : chaque point représente la carte des taux d'incidence au temps t. Le centre de gravité du nuage de points représente la carte moyenne des taux d'incidence.* ***B :*** *champ $X^c(t,s)$. Chaque point représente une carte d'anomalies d'incidence par rapport à la carte moyenne (qui devient donc le centre du repère).*

Table 2.10 : Exemple de champ spatio-temporel de taux d'incidence
*X(t,***s***)*, avec $n = 6$ et $p = 2$

	Région 1	Région 2
1995	0,7	1,9
1996	1	3,6
1997	2,9	4,7
1998	3,7	3,7
1999	4,8	6,4
2000	5,6	5,5

Les vecteurs propres de la matrice de covariance du champ présenté en Table 2.10 sont donc les vecteurs directeurs unitaires des axes principaux. Les éléments $u_{1k}, \ldots, u_{pk}$ du vecteur $\mathbf{u}_k$ représentent les coordonnées du vecteur $\mathbf{u}_k$ dans le repère formé par les entités spatiales. Les éléments a_{1k}, $a_{ik}, \ldots, a_{nk}$ du vecteur $\mathbf{a}_k$ représentent les coordonnées du champ spatio-temporel projeté sur l'axe k (Figure 2.17).

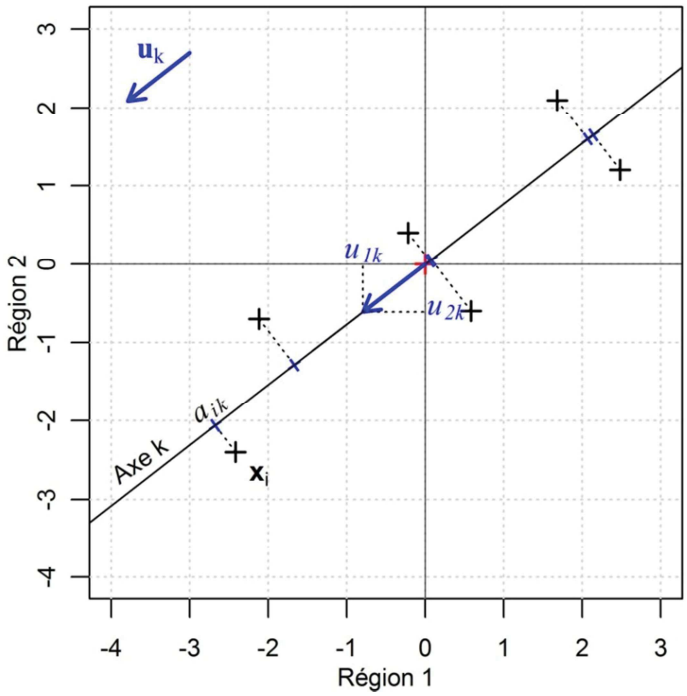


Figure 2.17 : Axe principal de la variabilité spatio-temporelle des taux d'incidence. *La valeur propre* λ_k *associée à l'axe k est égale à la variance de* $(a_{1k}, \ldots, a_{nk})$

Interprétation d'une analyse EOF :

Les éléments à interpréter dans une analyse EOF sont :

- La variance totale du nuage de points ;
- La part de variance expliquée par chacun des axes ;
- Les composantes principales ;
- Les cartes d'EOF.

1. Variance totale du nuage de points (somme des valeurs propres)

La variance totale du nuage de points, ou inertie totale, est égale à la somme des valeurs propres et rend compte de la dispersion du nuage de points autour de la moyenne, *i.e.* rend compte de la variabilité temporelle des schémas spatiaux de taux d'incidence observés par rapport à la carte spatiale moyenne. Une grande inertie totale signifie qu'il existe une grande variabilité temporelle des schémas spatiaux : certaines des cartes de taux d'incidence observées aux pas de temps t_i, $i = 1, \ldots, n$ s'écartent très fortement de la moyenne. Inversement, une faible inertie totale montre une faible variabilité temporelle des schémas spatiaux : toutes les cartes de taux d'incidence se ressemblent et sont proches de la carte moyenne (Figure 2.18).

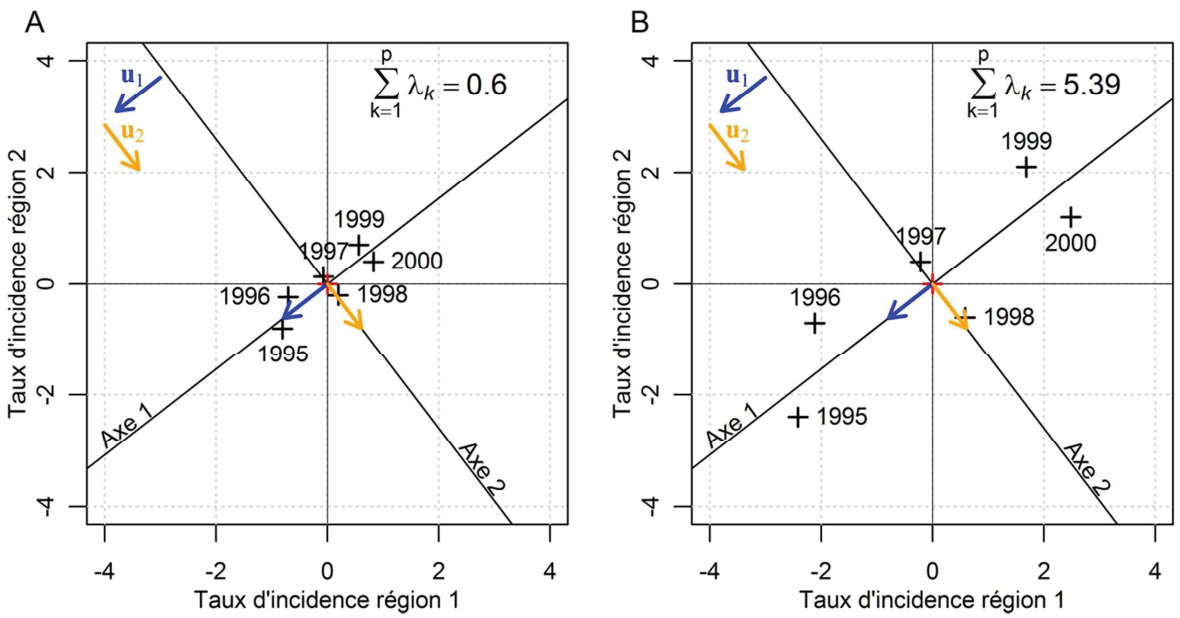


Figure 2.18 : Représentation de deux champs spatio-temporels de taux d'incidence *possédant les mêmes axes principaux, mais des variances totales différentes. A. Les cartes de taux d'incidence annuelle ressemblent à la carte moyenne. B. Les cartes de taux d'incidence annuelle s'écartent de la carte moyenne.*

2. Part de variance expliquée par chacun des axes (étude des valeurs propres)

Dans le cas où l'inertie totale du nuage de points est petite, il n'est pas utile d'aller plus loin dans l'étude la variabilité spatio-temporelle des incidences. Dans le cas où l'inertie totale du nuage est grande, les schémas spatiaux observés à certains pas de temps s'écartent fortement de la moyenne. Ces schémas peuvent soit présenter une certaine cohérence temporelle, se répétant d'un pas de temps à l'autre, soit être tous différents les uns des autres.

Si un groupe de régions spatiales ont des dynamiques qui co-varient, *i.e.* si les pas de temps où un fort taux d'incidence est observé dans une région sont également ceux où un fort taux d'incidence est observé dans d'autres régions, les structures spatiales observées sont cohérentes d'un pas de temps à l'autre, les points du nuage auront tendance à s'aligner selon une direction privilégiée. L'axe principal correspondant capturera alors une grande partie de la variance totale du nuage de points. La variance ou valeur propre associée à cet axe aura donc tendance à être forte. L'EOF correspondant montrera le groupe de régions spatiales dont les taux d'incidence s'écartent fortement de la moyenne (présentant une grande variance) et dont la dynamique co-varie (voir Figure 2.19.A). L'EOF sera représentatif d'un ensemble de pas de temps au cours desquels le schéma spatial identifié par l'EOF se répète. Les composantes principales identifieront cet ensemble de pas de temps présentant des schémas spatiaux similaires. Dans ce cas, l'étude d'un nombre réduit d'EOF et de composantes principales sera suffisante à l'étude de la variabilité spatio-temporelle des données.

Au contraire, si certains schémas spatiaux s'écartent fortement de la moyenne (grande inertie totale) mais que les dynamiques de chacune des régions spatiales ne sont pas co-variantes entre elles, les points seront dispersés autour de la moyenne de manière éparse, sans aucune direction privilégiée. La variance

du nuage de points se répartira alors sur plusieurs axes (voir Figure 2.19.B et D). Dans ce cas, l'interprétation des composantes principales et des EOF devient aussi fastidieuse que celle de l'étude de cartes spatiales des taux d'incidence bruts, puisqu'elle nécessitera l'examen d'un nombre important de modes de variabilité pour rendre compte de la totalité de l'information contenue initialement dans les données.

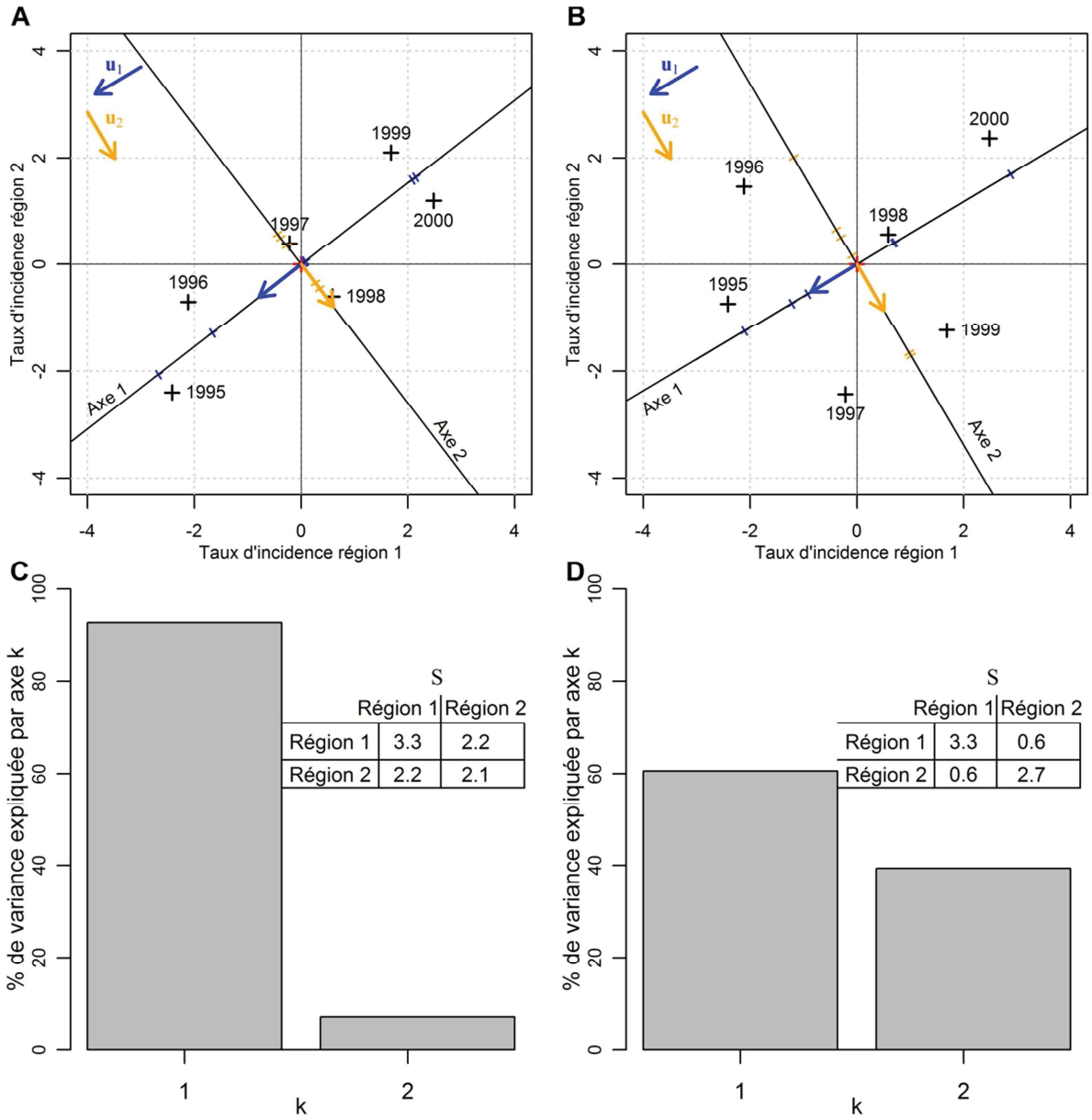


Figure 2.19 : Comparaison des valeurs propres d'une analyse EOF effectuées sur deux jeux de données. A et B. *Champ X(t,s) et axes principaux d'un jeu de données dont les deux régions présentent soit une forte covariance temporelle (A), soit une faible covariance temporelle (B).* ***C et D.*** *Valeurs propres (exprimées en pourcentage de variance expliquées par chacun des axes) de la matrice de covariance de chacun des jeux de données présentés respectivement en A et B. En A, la covariance entre les deux séries temporelles rend les structures spatiales presque identiques d'une année à l'autre, toutes proportionnelles au schéma spatial* $\boldsymbol{u}_1$ *à un facteur d'expansion près : la composante principale* $\boldsymbol{a}_1$.

Le nombre d'axes à interpréter pour rendre compte de l'information initiale est en général déterminé de manière à expliquer un certain pourcentage de la variance des données initiales. Le pourcentage de

variance spatio-temporelle des données initiales expliquées par un nombre i donné d'axes principaux est donnée par :

$$\%\ variance\ expliquée\ par\ les\ i\ premiers\ axes\ principaux = 100 \cdot \frac{\sum_{k=1}^{i} \lambda_k}{\sum_{k=1}^{m} \lambda_k}\ \%$$

Le spectre des valeurs propres donne également une idée de la dégénérescence des valeurs propres. La dégénérescence survient lorsque deux valeurs propres ont des valeurs proches. Les EOF correspondants doivent alors être interprétés avec prudence, car le phénomène épidémiologique sous-jacent est en fait réparti sur plusieurs EOF, et n'importe quelle combinaison linéaire de ces EOF est toute aussi significative que les EOF eux-mêmes (Hannachi et al., 2007). Voir North *et al* (North et al., 1982) pour le calcul de critères de dégénérescence.

3. Interprétation des EOF

Les EOF $\mathbf{u}_k = \left(u_{1k}, \ldots, u_{pk}\right)^T,\ k = 1, \ldots, m$ peuvent être représentés sous la forme de m cartes spatiales représentant les structures spatiales d'anomalies de taux d'incidence dominantes dans les données. La composante principale $\mathbf{a}_k$ associée à chaque EOF donne l'évolution de cette structure spatiale au cours du temps. Les EOF doivent être interprétés par ordre de valeur propre décroissante, chaque EOF représentant alors une part de la variance initiale des données de plus en plus réduite. Les EOF étant des vecteurs unitaires, les valeurs $u_{jk},\ j, k = 1, \ldots, m$ sont toujours comprises entre -1 et 1. Le signe d'un EOF donné est arbitraire, et doit être interprété conjointement au signe de la composante principale correspondante.

L'interprétation se limite uniquement aux EOF expliquant une part importante de la variance totale (λ_k élevée). Sur le $k^{ième}$ EOF, les régions spatiales ayant les plus fortes valeurs absolues (positives ou négatives) de u_{jk} sont celles présentant la plus forte anomalie d'incidence par rapport à leur moyenne respective $\overline{x}_{.j}$. L'examen des EOF permet donc d'identifier les zones géographiques où un agent pathogène donné a circulé de manière anormalement faible ou anormalement élevée au cours de certaines années par rapport à la moyenne de la circulation dans ces zones géographiques toutes années confondues. Si l'EOF, explique une part importante de la variance, deux régions spatiales $\mathbf{s}_i$ et $\mathbf{s}_j$ présentant des valeurs u_{ik} et u_{jk} proches (même couleur sur la carte du k^{ième} EOF) ont une dynamique des taux d'incidence qui co-varie fortement.

Les schémas spatiaux identifiés par chacun des EOF peuvent ensuite être analysés ou interprétés sur le plan épidémiologique comme n'importe quelle carte de taux d'incidence. A ce titre, les EOF constituent un filtre des structures spatiales des données, et les analyses des facteurs de risque peuvent être effectuées uniquement sur les structures spatiales principales d'anomalies (Ortíz Bultó et al., 2006).

4. Interprétation des composantes principales

De manière similaire aux EOF, les composante principales $\mathbf{a}_k = (a_{1k}, \ldots, a_{nk})^T,\ k = 1, \ldots m$ peuvent être représentées sous la forme de séries temporelles et doivent être interprétées par ordre décroissant de valeur propre, et conjointement à l'EOF associé. Seules les composantes principales de plus forte

variance doivent être interprétées. La variance de la $k^{ième}$ composante principale est égale à λ_k : au même titre que l'inertie totale, la valeur absolue des éléments a_{ik} de la $k^{ième}$ composante principale rend compte de la faible ou forte variabilité temporelle des schémas spatiaux par rapport à la carte moyenne. L'écart du taux d'incidence de la région j à la moyenne $\overline{x}_{.j}$ au temps t est d'autant plus fort que la valeur absolue de a_{tk} est élevée. Le signe de l'anomalie est donné par le signe du produit $a_{tk}\ u_{jk}$ (ceci n'est vrai que pour les axes dont la variance λ_k est élevée). Les composantes principales permettent donc d'identifier les pas de temps au cours desquels un agent pathogène donné a circulé de manière anormalement faible ou anormalement élevée dans certaines régions par rapport à la circulation moyenne de l'agent pathogène dans les régions considérées. Sur la $k^{ième}$ composante principale, deux pas de temps t_i et t_j possédant des valeurs proches de a_{ik} et a_{jk} ont une structure spatiale proche (exemple : sur le jeu de données représenté en Figure 2.19.A, les cartes spatiales de taux d'incidence des années 1999 et 2000 se ressemblent).

L'analyse dynamique des composantes principales permet d'identifier les fréquences avec lesquelles un phénomène spatial cohérent survient (exemple pour El Niño : tous les 2 à 8 ans, ou pour l'oscillation de Madden-Julian : tous les 50 à 80 jours (Madden and Julian, 1971)).

Applications et intérêts d'une analyse EOF

Nous venons de voir que l'analyse EOF est extrêmement riche de sens, et permet :

- De savoir si les schémas spatiaux de taux d'incidence observés à différents pas de temps sont stables dans le temps ou non (faible inertie totale vs forte inertie totale) ;
- De savoir, dans le cas où les schémas spatiaux sont variables dans le temps, si des schémas spatiaux particuliers (anomalies par rapport à la moyenne) sont récurrents ou s'ils sont tous différents ;
- D'identifier les pas de temps ainsi que les régions spatiales pour lesquels un agent pathogène a circulé de manière « anormalement » faible ou anormalement élevée dans certaines régions ;
- D'identifier les groupes de régions spatiales dont la dynamique temporelle co-varie ainsi que les groupes de pas de temps dont les schémas spatiaux d'anomalie d'incidence sont similaires.

Tout ceci pourrait également être fait à l'aide de l'inspection visuelle des n cartes de taux d'incidence, des p cartes de covariance ou de corrélation, avec l'aide de calculs d'indicateurs de moyennes, d'écart-types et de covariances. L'intérêt d'une analyse EOF, comme toute ACP, réside essentiellement :

- Dans la réduction du nombre de dimensions à analyser pour interpréter la variabilité spatio-temporelle des données d'incidence : en effet, s'il existe des structures spatiales dominantes, souvent, seuls les deux ou trois premiers axes principaux sont à interpréter, synthétisant une part importante de l'information initiale (inertie), et ce quelles que soient les dimensions n et p du champ spatio-temporel de départ ;

– Dans la simplicité et la rapidité des algorithmes de mise en œuvre : en effet, seules quelques lignes de programmation dans des logiciels permettant le calcul matriciel suffisent à obtenir les résultats.

Nous nous sommes limités dans cette thèse à l'application qui peut être faite d'une analyse EOF dans la description synthétique de la variabilité spatio-temporelle des taux d'incidence et l'identification des structures spatio-temporelles dominantes. Mais l'analyse EOF peut également être utilisée pour filtrer les données, en ne reconstruisant le signal d'origine qu'à partir d'un nombre restreint de modes de variabilités.

Limites de l'analyse EOF classique et extensions possibles (Hannachi et al., 2007)

En climatologie, la contrainte d'orthogonalité des schémas spatiaux obtenus peut rendre difficile l'interprétation physique des résultats. Il devrait en être de même en épidémiologie : si deux processus épidémiologiques corrélés entre eux sont à l'origine de la variabilité spatio-temporelle observée des données, les schémas spatiaux en résultant seront également corrélés entre eux et certains schémas spatiaux cohérents sur le plan épidémiologique se répartiront sur plusieurs EOF, rendant leur interprétation difficile. Pour surmonter ces problèmes d'interprétation, les chercheurs ont développé des méthodes spécifiques. Parmi elles, les REOF (pour « Rotated » EOF) (Horel, 1981; Richman, 1981).

L'analyse EOF classique ne tient pas compte de la nature spatiale ou temporelle des données, et ne permet donc pas d'incorporer les processus d'auto-corrélation temporelle ou d'auto-corrélation spatiale existant pourtant dans les données épidémiologiques, du fait de la nature infectieuse ou environnementale d'une maladie. Les chercheurs ont développé les « extended » EOF, qui permettent la prise en compte explicite des schémas d'auto-corrélation, ainsi que la détection de tendances, d'oscillation, ou de processus de propagation (Plaut and Vautard, 1994; Weare and Nasstrom, 1982).

L'analyse EOF classique rend compte de l'évolution temporelle des schémas spatiaux, mais ne permet pas, en tant que telle, l'identification de processus de propagation. Si le but de l'analyse est la détection de tels processus, les EOF dans le domaine fréquentiel (EOF complexes) peuvent être utilisés (Rasmusson et al., 1981; Wallace, 1972; Wallace and Dickinson, 1972).

L'analyse EOF classique ne permet d'analyser qu'un seul champ spatio-temporel à la fois. L'analyse simultanée de plusieurs champs spatio-temporels (par exemple un champ de taux d'incidence et un champ de variable climatique) possédant les mêmes pas de temps mais pouvant être mesurés sur des grilles spatiales différentes peut être réalisée par la méthode de décomposition des valeurs singulières (SVD) ou par des analyses de corrélation canonique (Bretherton et al., 1992).

PARTIE II : ANALYSE ET MODELISATION MULTI-ECHELLE

CHAPITRE 3 : ANALYSE DE LA DENGUE A L'ECHELLE DU PACIFIQUE

3.1. Analyse spatio-temporelle des données épidémiologiques 141

3.1.1. Transformation des données ---------- 141
3.1.1.a. Standardisation spatiale des données 141
3.1.1.b. Définition d'une année épidémique 142
3.1.2. Caractérisation globale de la distribution spatio-temporelle des cas de dengue -------- 144
3.1.2.a. Analyse spatio-temporelle 144
3.1.2.b. Analyse de la structure spatiale moyenne 145
3.1.2.c. Analyse de la dynamique globale des épidémies de dengue dans le Pacifique..... 146
3.1.2.d. Analyse de la variabilité spatio-temporelle de la dengue dans le Pacifique par la méthode des EOF 149
3.1.3. Analyse temporelle par pays ---------- 149
3.1.3.a. Analyse des fréquences épidémiques par pays 149
3.1.3.b. Caractérisation des profils dynamiques 151
3.1.4. Analyse par vague pandémique ---------- 152
3.1.4.a. Etude de la propagation d'un sérotype au sein du Pacifique 153
3.1.4.b. Etude de la structure spatiale des vagues 155
3.1.4.c. Dynamique des vagues pandémiques..... 157
3.1.5. Discussion et perspectives ---------- 157
3.1.5.a. Résumé des résultats à l'échelle régionale 157
3.1.5.b. Emergence d'un sérotype dominant 158
3.1.5.c. Dynamique cyclique des pandémies de dengue dans le Pacifique 161
3.1.5.d. Particularités épidémiologiques inexpliquées..... 163
Circulation prolongée du DENV-1 163
Co-circulation du virus DENV-3 avec d'autres sérotypes 164
Circulation endémique en Polynésie Française 165
3.1.5.e. Absence de schémas régionaux..... 165

Le travail présenté dans ce chapitre fait suite au travail initial effectué par Descloux durant sa thèse (Descloux, 2011), et a été effectué en collaboration avec Mlle Alizé Mercier, dont j'ai co-encadré le stage dans mon unité. A l'échelle du Pacifique, à l'exception du climat pour lequel des données issues de simulations numériques existent, nous n'avons eu accès qu'à des données parcellaires concernant les covariables que nous pouvions potentiellement étudier. Nous avons donc limité notre analyse à une étude de la dynamique spatio-temporelle des données épidémiologiques présentées au Chapitre 2 (page 86). Les analyses concernant le climat ont été effectuées par d'autres personnes de l'équipe que je n'ai pas co-encadrées. Elles ne sont donc pas présentées dans cette thèse.

3.1. Analyse spatio-temporelle des données épidémiologiques

3.1.1. Transformation des données

3.1.1.a. Standardisation spatiale des données

A une échelle supra-nationale comme celle du Pacifique, les réseaux de surveillance ne sont pas standardisés sur le plan spatial : chaque pays utilise ses propres modalités de collecte des données en fonction de ses objectifs sanitaires et de ses ressources financières ou humaines. Or le nombre de cas déclarés est fortement dépendant de ces modalités de collecte. Les taux d'incidence calculés à cette échelle ne peuvent donc en aucun cas être comparés tels quels d'un pays à l'autre. Ils nécessitent une standardisation spatiale.

Sur le plan temporel par contre, même si certaines modalités de collecte ont été modifiées ente 1971 et 2009, comme par exemple la définition du cas, ou les tests utilisés pour le diagnostic biologique, nous pouvons supposer une certaine homogénéité au sein d'un même pays, permettant une interprétation de la dynamique épidémique au sein d'un pays, et la comparaison des dynamiques des différents pays entre elles.

Nous avons donc opté pour une standardisation qui consiste, dans chacun des vingt PTIP étudiés, à diviser les taux d'incidence par le taux d'incidence maximal observé dans le pays considéré. Ainsi, la cible des analyses par la suite est un chiffre adimensionnel, qui conserve l'information relative à l'intensité de l'épidémie, laisse la dynamique au sein d'un même pays intacte, mais permet de ramener l'ensemble des taux d'incidence observés sur une même échelle allant de 0 à 1 quelles que soient les modalités de collecte des données utilisées. Les niveaux des taux d'incidence standardisés pour un pas de temps donné restent non comparables d'un pays à l'autre.

3.1.1.b. Définition d'une année épidémique

Très rapidement, il a fallu, pour chacun des vingt PTIP, distinguer les années épidémiques des années non épidémiques. Dans un objectif d'analyse et de modélisation de la dynamique spatio-temporelle, il a fallu trouver une définition quantitative qui soit robuste à des changements important des réseaux de surveillance d'une année à l'autre, et qui soit applicable à l'ensemble des pays.

Descloux avait proposé, dans sa thèse (Descloux, 2011), de déclarer une année comme étant épidémique si son taux d'incidence se situe dans le tercile supérieur des taux d'incidence calculés sur les trente-neuf années d'étude (voir Figure 3.1.A pour une illustration en Nouvelle-Calédonie). Cette définition est extrêmement robuste aux biais pouvant résulter de l'absence de standardisation des réseaux de surveillance et fonctionne bien pour les pays dans lesquels on observe régulièrement des épidémies, comme en Nouvelle-Calédonie.

Un des inconvénients de cette définition est qu'elle fixe à priori le nombre d'années épidémiques à un tiers de la période d'étude, c'est-à-dire treize années épidémiques. Cette définition, appelée « méthode des terciles », devient donc trop sensible et peu spécifique dans certains pays (*i.e.* classe comme épidémiques des années au cours desquels seuls un ou deux cas ont été déclarés dans un pays). En effet, lorsqu'il y a moins de treize années avec des cas déclarés, l'ensemble de ces années sont déclarées comme épidémiques, y compris les années où seuls un ou deux cas ont été déclarés (voir Figure 3.1.C pour une illustration aux Samoa Américaines). La définition basée sur les terciles est par contre trop peu sensible, et trop spécifique (i.e classe comme non épidémiques des années qui à priori présentent un nombre relativement important de cas) dans les pays où il y a plus que treize années épidémiques, comme en Polynésie Française (Figure 3.1.D).

Nous avons donc exploré d'autres définitions, et avons opté pour définir une année comme étant épidémique dans un pays donné si le taux d'incidence observé cette année est supérieure ou égal à 2,5% du taux d'incidence maximal observé dans le pays sur les 39 années d'étude, ce qui est équivalent à définir une année comme étant épidémique si le taux d'incidence standardisé est supérieur à 0,025 (voir Figure 3.1.B pour une illustration en Nouvelle-Calédonie). Comme illustré en Figure 3.1 C et D, cette méthode utilisant un seuil fixe est moins sensible (et donc plus spécifique) que la méthode des terciles dans les pays où les épidémies sont rares, et moins spécifique (et donc plus sensible) dans les pays où les épidémies sont régulières. Le seuil de 0,025 a été défini après inspection visuelle des vingt courbes épidémiques semblables à celles présentées en Figure 3.1 C et D.

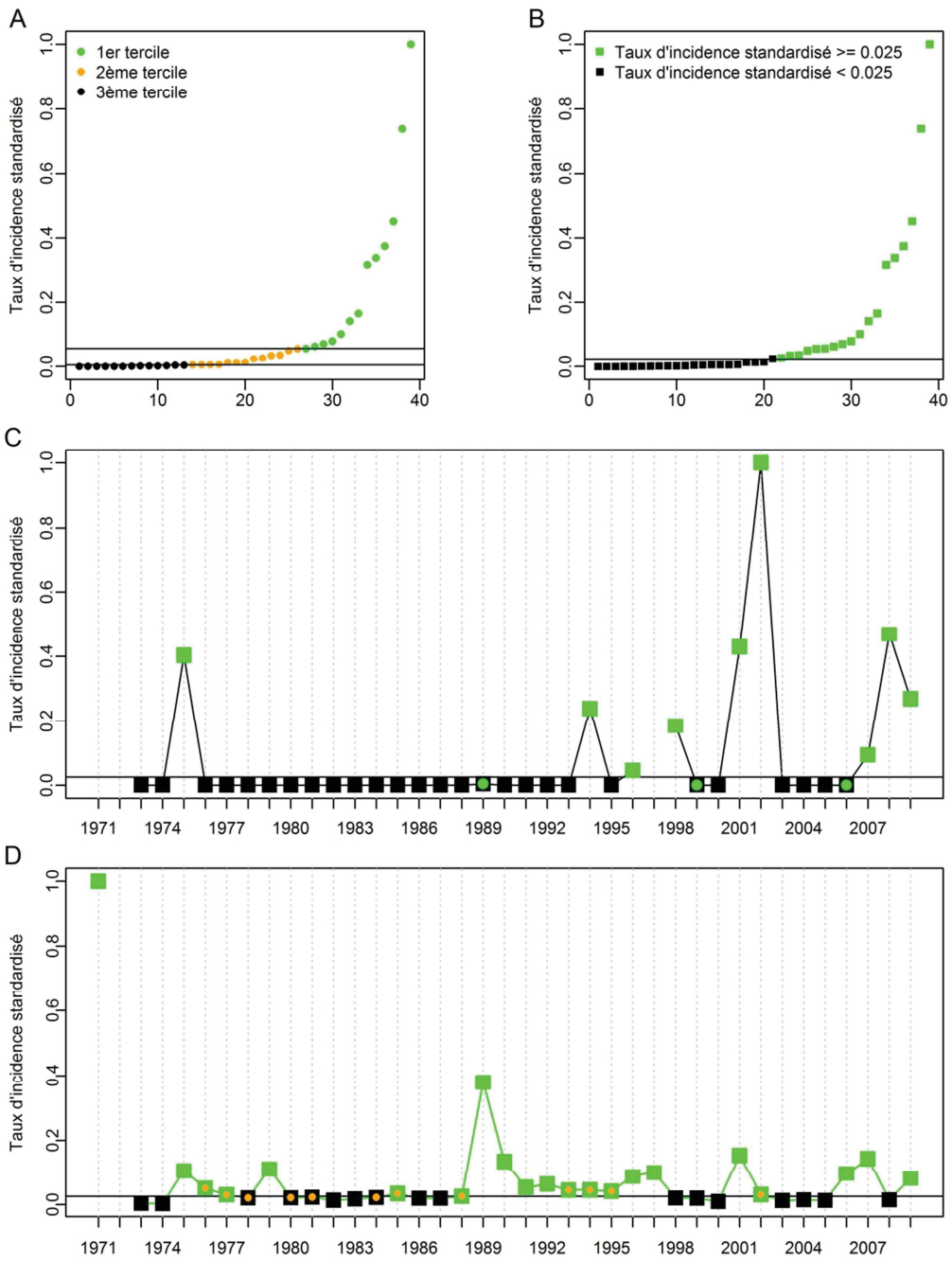


Figure 3.1 : Comparaison de deux méthodes pour définir les années épidémiques à l'échelle du Pacifique. ***A et B****. Méthodes des terciles (A) et du seuil (B) appliquées aux taux d'incidence standardisés de Nouvelle-Calédonie. Les taux d'incidence sont triés par ordre croissant. Les années dont le taux d'incidence est supérieur au deuxième tercile (A) ou supérieur ou égal à un seuil de 0.025 (B) sont classées comme épidémiques (en vert). Les années dont le taux d'incidence est inférieur au premier tercile (A) ou à 0.025 (B) sont classées comme non épidémiques. Les années dont le taux d'incidence est entre les deux terciles sont classées comme indéfinies. Afin d'améliorer la spécificité de la définition, elles ont ensuite été classées comme non épidémiques.* ***C et D*** *: comparaison de la sensibilité et de la spécificité des deux méthodes pour les American Samoa, pays où les épidémies sont rares (C) et la Polynésie Française, pays où des épidémies de faible ampleur surviennent presque tous les ans (D). Les légendes sont les même qu'en A et B. Les carrés d'une couleur uniforme indiquent que les deux méthodes donnent une classification identique. La ligne horizontale représente la valeur seuil de 0,025 du taux d'incidence maxial.*

3.1.2. Caractérisation globale de la distribution spatio-temporelle des cas de dengue

3.1.2.a. Analyse spatio-temporelle

La Figure 3.2 montre l'évolution spatio-temporelle des taux d'incidence standardisés entre 1971 et 2009 dans les vingt PTIP étudiés. La co-circulation de plusieurs sérotypes au sein d'un même pays a été rapportée, mais reste rare comparée aux pays hyperendémiques d'Asie du Sud-Est. La couleur sur la Figure 3.2 code donc pour le sérotype dominant. La couleur grise correspond aux années pour lesquelles le sérotype n'est pas connu, ou pour lesquelles deux sérotypes ont circulé en même temps, sans que le sérotype dominant ne soit précisé dans les données à notre disposition.

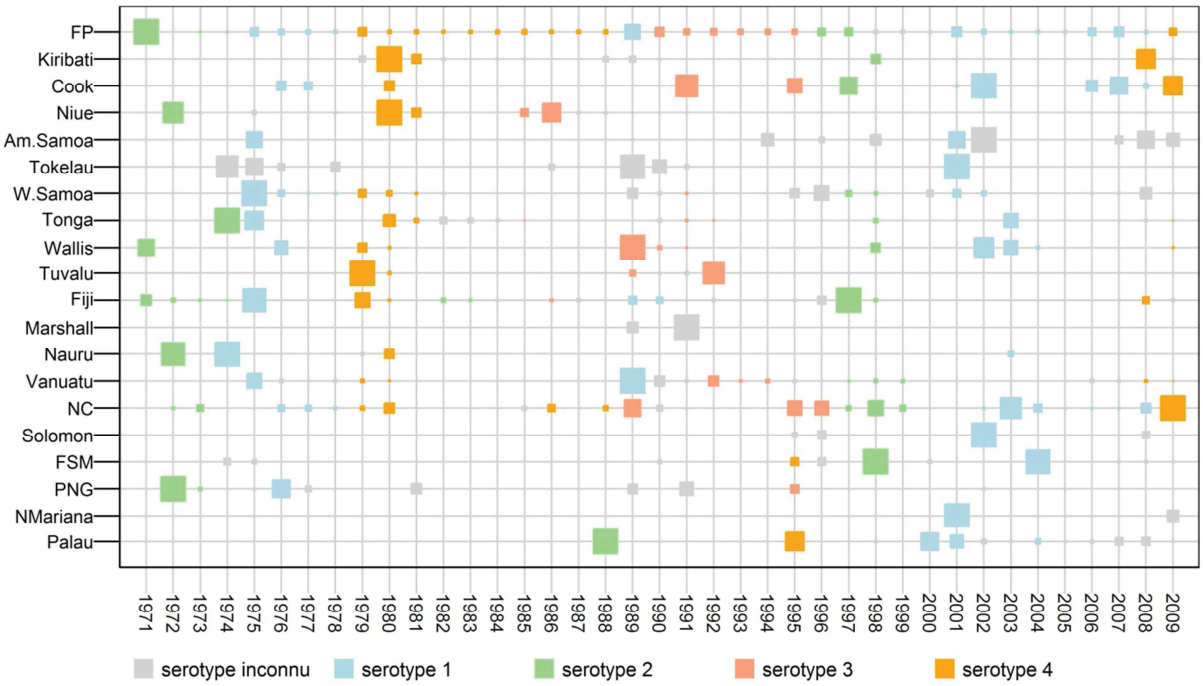


Figure 3.2 : Distribution spatio-temporelle des taux d'incidence annuels standardisés dans vingt PTIP (1971-2009). *Pour chaque pays, la surface du carré est proportionnelle au taux d'incidence standardisé, et le carré de taille maximale représente un taux d'incidence standardisé de 1 (taux d'incidence observé maximal pour le pays). La couleur code pour le sérotype dominant ayant circulé au cours de l'année.*

La dengue dans le Pacifique survient sous forme de vagues épidémiques suite à l'émergence et à la diffusion d'un nouveau sérotype dans la région. La diffusion sur une majorité des îles du Pacifique dure entre 3 et 6 ans, temps au bout duquel un nouveau sérotype émerge et circule de façon dominante. Les sérotypes semblent se succéder selon l'ordre DENV-2, DENV-1, DENV-4, puis DENV-3. Cependant, les séries ne sont pas suffisamment longues pour que ce résultat soit robuste. Les sérotypes 2, 1 et 4 ont été responsables de deux vagues pandémiques au cours des 39 années d'étude. Ces vagues dues à un même sérotype sont espacées de 20 à 24 ans (entre 1975 et 1995 pour le sérotype 2, entre 1978 et 2000 pour le sérotype 1, et entre 1984 et 2008 pour le sérotype 4), si l'on excepte quelques résurgences

touchant un nombre limité de pays (sérotype 2 touchant un pays en 1988, sérotype 1 touchant trois pays à la fin des années 1980 et sérotype 4 touchant trois pays en 1995).

A partir de ces résultats, plusieurs questions peuvent être posées :

- Quelle est la dynamique globale de la dengue dans le Pacifique, tous pays confondus ? La fréquence de survenue des épidémies augmente-t-elle au cours du temps ?
- Quelle est la structure spatiale globale de la dengue dans le Pacifique ? Certains pays sont-ils systématiquement plus touchés que d'autres ? Y a-t-il une évolution de la distribution géographique des pays touchés ?
- Pour chacun des pays, quelle est la dynamique de la dengue ? Peut-on repérer des groupes de pays présentant la même dynamique ? Si oui, ces groupes ont-ils une structure géographique particulière ?
- Pour chacune des vagues épidémiques, existe-t-il un schéma de propagation particulier ? Un sérotype nouveau émerge-t-il systématiquement dans un même pays ou groupe de pays ? Les pays sont-ils toujours touchés dans le même ordre ?
- La structure spatiale moyenne de la dengue est-elle la même d'une vague épidémique à l'autre ?

3.1.2.b. Analyse de la structure spatiale moyenne

Ce paragraphe vise à répondre à la question : y a-t-il des pays ou régions géographiques plus touchées que d'autres ?

Les taux d'incidence ne peuvent pas être comparés d'un pays à l'autre à cette échelle. Nous avons donc décidé de comparer le nombre d'épidémies déclarées dans chacun des pays entre 1971 et 2009 (Figure 3.3).

D'après la Figure 3.3, si l'on suppose que les systèmes de surveillance des îles du Pacifique sont à même de détecter des épidémies, les pays les plus régulièrement touchés par la dengue sont les pays situés dans le Sud du Pacifique Sud, entre 15° et 20° de latitude sud. Les deux pays présentant le plus grand nombre d'épidémies déclarées sont la Nouvelle-Calédonie et la Polynésie Française. Ces deux territoires français présentent la particularité d'avoir une partie de leur population adulte renouvelée chaque année (affectations professionnelles depuis la métropole), et dont la population a les moyens financiers de voyager par avion dans différentes régions du monde potentiellement infectées par la dengue. Cependant cette carte est à interpréter avec prudence. En effet, ces deux territoires français sont également ceux qui possèdent les meilleures infrastructures de santé publique dans la région, et les moyens de surveillance les plus importants. Cette carte contient potentiellement des informations intéressantes, mais ne peut être interprétée que conjointement à une carte de la fiabilité des systèmes de surveillance dans les pays considérés.

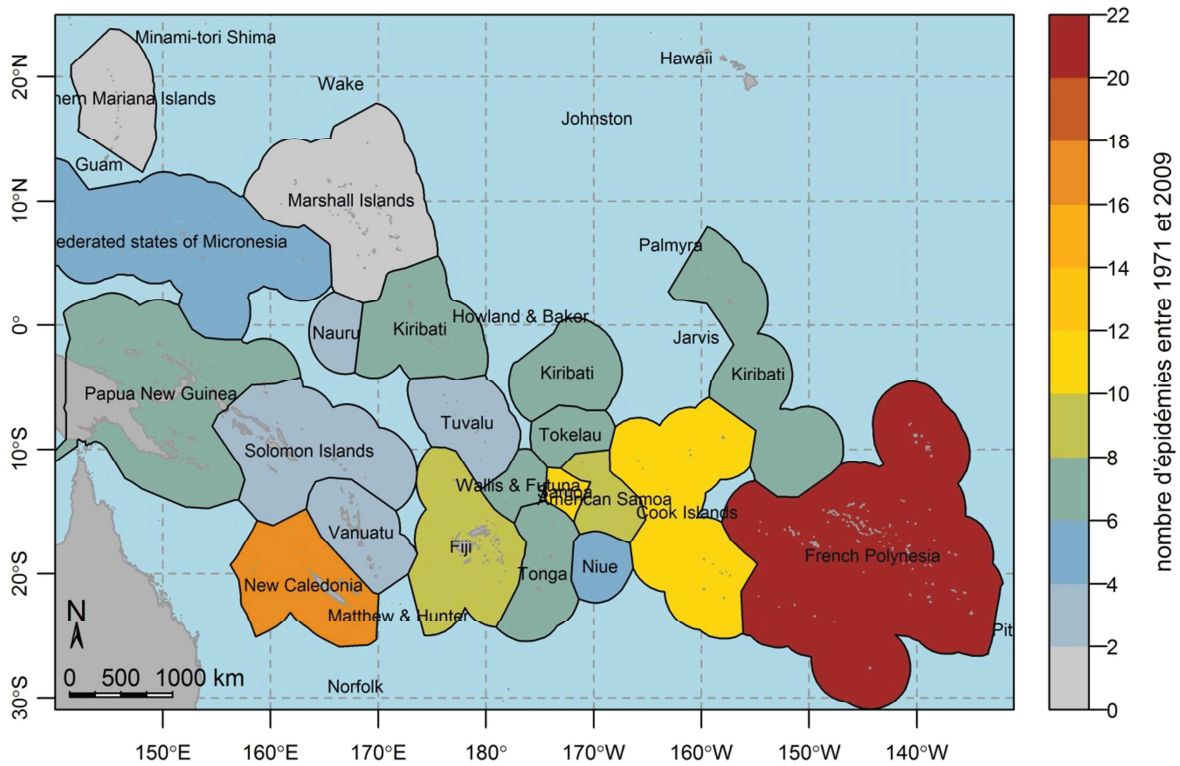


Figure 3.3 : Carte du nombre d'épidémies entre 1971 et 2009 dans le Pacifique.

3.1.2.c. Analyse de la dynamique globale des épidémies de dengue dans le Pacifique

La Figure 3.4.A montre l'évolution du nombre de pays touchés par une épidémie de dengue entre 1971 et 2009, tous sérotypes confondus. Sur l'ensemble de la période d'étude, il n'existe que trois années au cours desquelles aucun pays n'a été touché par une épidémie : 1984, 1987 et 2005.

Afin d'évaluer la durée des vagues épidémiques, nous avons réalisé une analyse en ondelettes de la série temporelle présentée en Figure 3.4.A (voir Figure 3.4 B et C). L'analyse en ondelettes permet de quantifier les fréquences présentes dans un signal temporel et d'en étudier l'évolution au cours du temps. Le spectre d'ondelette montre que les vagues épidémiques surviennent de manière cyclique, tous les 4 ans entre 1976 et 1980, et tous les 5 à 7 ans depuis la fin des années 1980.

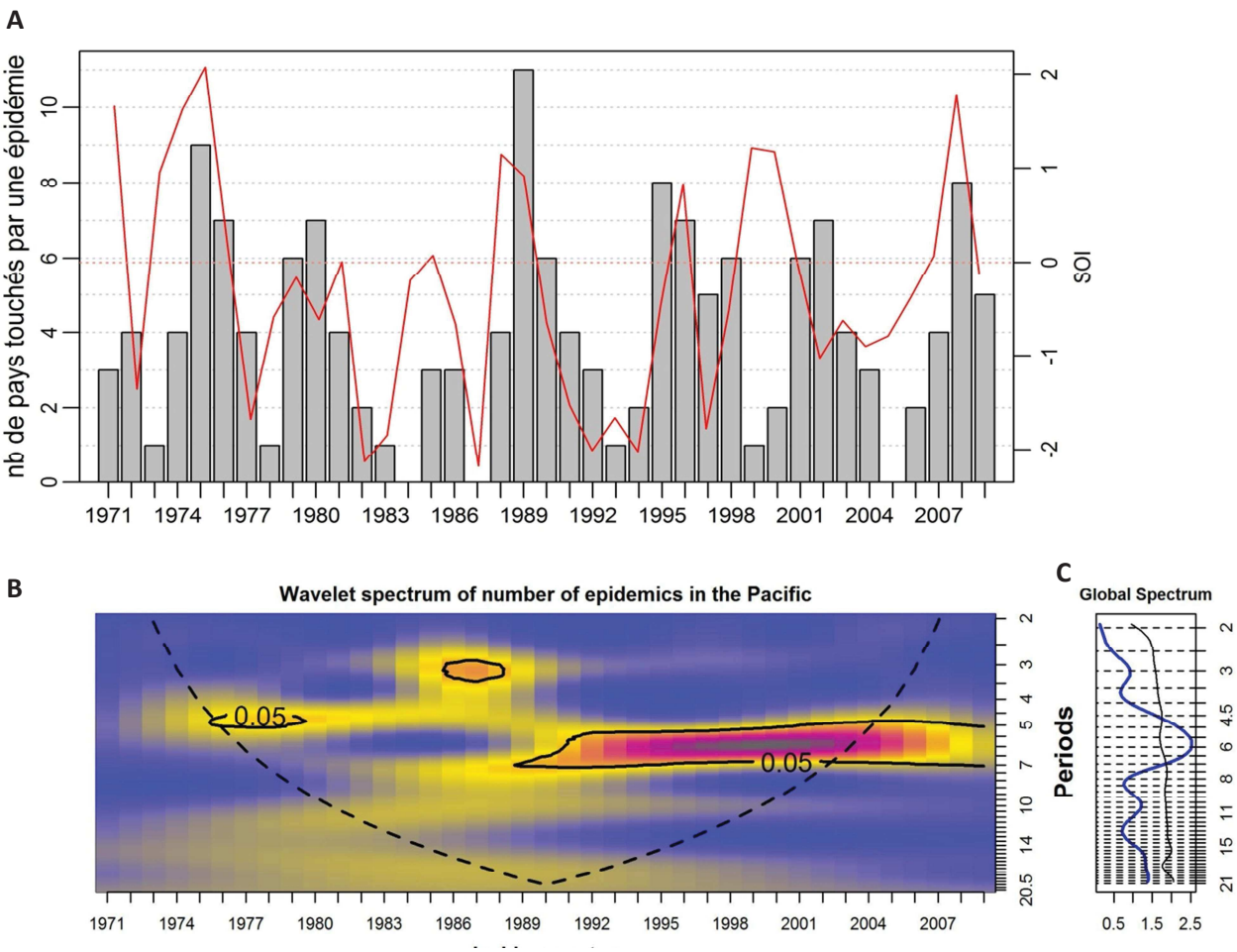


Figure 3.4 : Evolution cyclique du nombre de pays touchés par une épidémie de dengue dans le Pacifique. ***A*** *: Les barres grises représentent le nombre de pays touchés par une épidémie de dengue (taux d'incidence standardisé supérieur à 0.025) au cours de la période d'étude. La courbe rouge est la valeur du Southern Oscillation Index, un indicateur de la force du phénomène El Niño (valeurs négatives) ou La Niña (valeurs positives).* ***B*** *: Spectre d'ondelette de A. L'abscisse représente le temps, de 1971 à 2009. L'ordonnée représente les périodes étudiées, en années (de 2 à 20,5 ans), la couleur code pour les coefficients d'ondelette du signal A. Ces coefficients représentent la puissance du signal pour une année donnée et une période donnée, allant du bleu pour une faible puissance, jusqu'au violet pour une forte puissance. Les traits en pointillé représentent les limites du cône d'influence, zone au-delà de laquelle les effets de bord empêchent d'interpréter les coefficients d'ondelette (partie basse du spectre d'ondelette). Les traits noirs délimitent la zone de significativité des coefficients d'ondelette. L'axe du temps (abscisses) de A et B sont superposables.* ***C*** *: spectre global de A (assimilable à un spectre de Fourier). L'ordonnée représente la période (inverse de la fréquence) étudiée, et l'abscisse représente la puissance. La ligne bleue représente la moyenne temporelle des coefficients d'ondelette présentés en B (puissance moyenne du signal entre 1971 et 2009). La ligne noire représente le seuil de significativité de la puissance du signal pour la période considérée.*

L'augmentation de la période du signal depuis la fin des années 1980 traduit une augmentation de la durée de circulation virale au cours d'une vague épidémique, et l'augmentation de la durée avant laquelle un nouveau sérotype remplace le précédent dans le Pacifique. Au vu de la cible de l'analyse (nombre de pays touchés par une épidémie), cette augmentation de la durée de la circulation virale peut être interprétée comme :

- Une augmentation de la durée de la circulation virale au sein d'un pays touché, le nombre total de pays touchés au cours d'une vague restant constante (augmentation de la fréquence des épidémies au sein des pays touchés) ;
- Une augmentation du nombre de pays touchés au cours d'une vague épidémique (accroissement géographique), la durée de l'épidémie au sein d'un même pays restant constante.

La corrélation entre la SOI et le nombre d'épidémies annuelles est de 0,35 (p-valeur = 0,03) sur la période 1971-2009, et de 0,46 (p-valeur = 0,02) sur la période 1971-1996, tandis que Hales avait trouvé une corrélation de 0,58 sur la période 1971-1996 (Hales et al., 1996). Une partie de la variabilité inter-annuelle du nombre de pays touchés par la dengue peut donc être expliquée par des variations climatiques à grande échelle.

Afin d'évaluer la fréquence du nombre d'épidémies dans la région entière (qui combine l'information sur la durée des vagues pandémiques, telle qu'étudiée par l'analyse en ondelettes, avec l'intensité de chaque vague), nous avons calculé le nombre d'épidémies déclarées dans le Pacifique par période glissante de 7 ans, et effectué une régression linéaire de ce nombre en fonction du temps (Figure 3.5). La période de 7 ans a été choisie car elle correspond à la durée des vagues pandémiques depuis les années 1980, mais des fenêtres d'agrégation allant de 5 à 10 ans ont été testées.

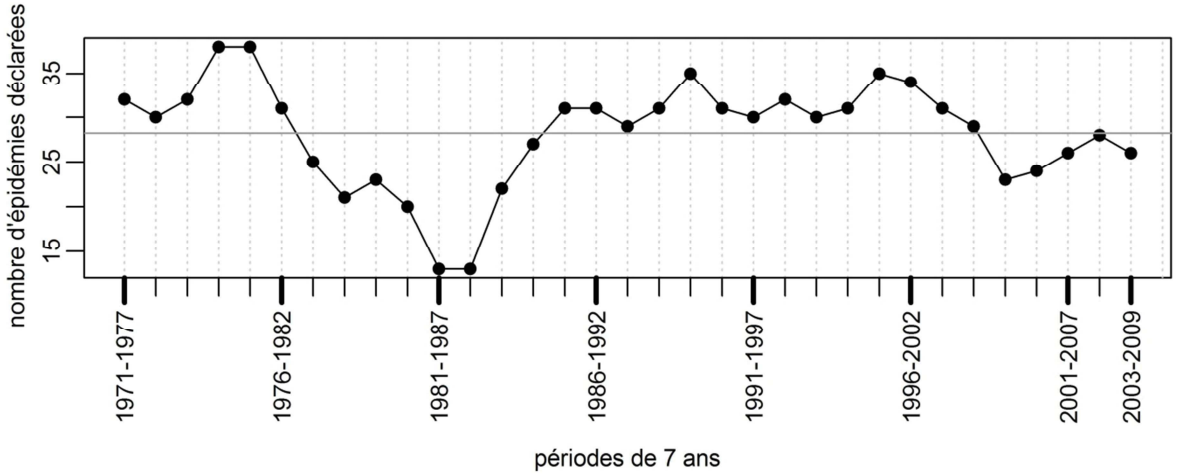


Figure 3.5 : Evolution de la fréquence des épidémies au cours du temps dans le Pacifique. *Chaque point représente le nombre d'épidémies déclarées au cours d'une période de 7 ans. La droite grise représente la régression du nombre d'épidémies par période glissante par rapport au temps. Sa pente représente l'augmentation annuelle du nombre d'épidémies par période de 7 ans. Elle est non significativement différente de zéro (p-value 0.99).*

De manière globale dans le Pacifique, il n'y a pas eu d'augmentation de la fréquence du nombre d'épidémies au cours des 39 dernières années (ce résultat reste valable pour les autres tailles de fenêtre glissante testées).

3.1.2.d. Analyse de la variabilité spatio-temporelle de la dengue dans le Pacifique par la méthode des EOF

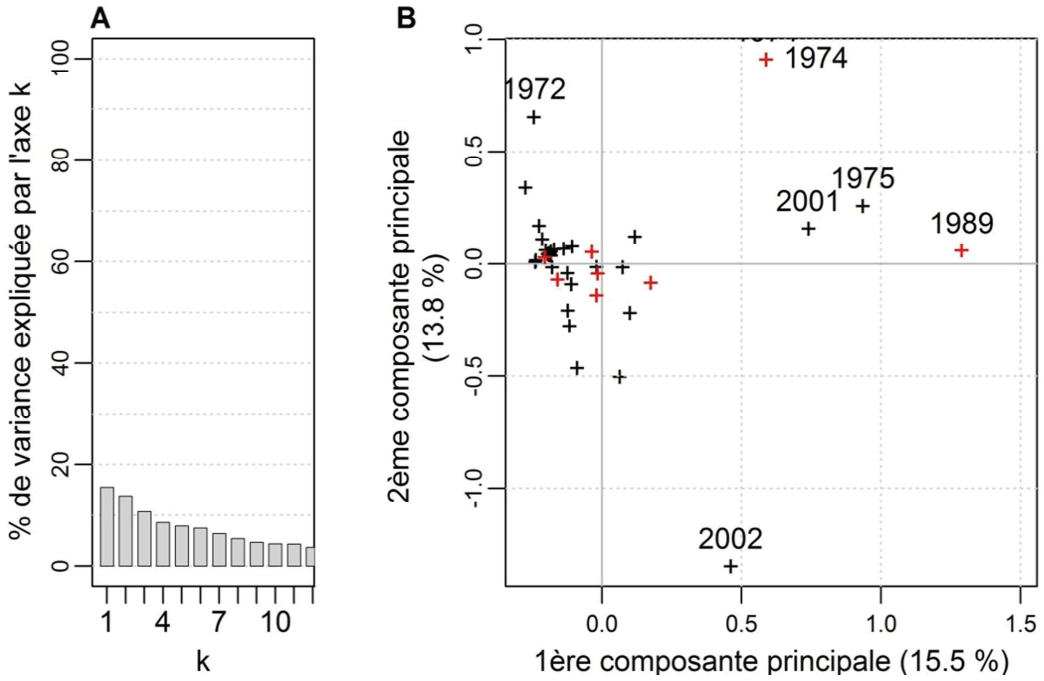


Figure 3.6 : Analyse EOF des taux d'incidence standardisés dans le Pacifique. A. *Valeurs propres (λ_k dans les équations des EOF).* ***B.*** *Représentation des cartes spatiales de taux d'incidence annuels standardisés ($\boldsymbol{x}_t$, $t = 1971, \ldots, 2009$) dans le plan formé par les deux premières composantes principales. Les points en rouge correspondent aux années au cours desquelles il y a eu apparition d'un nouveau sérotype dans le Pacifique.*

La Figure 3.6 montre les résultats de l'analyse EOF effectuée sur les taux d'incidence annuels standardisés dans vingt PTIP. L'étude des valeurs propres, avec une variance répartie équitablement sur un grand nombre d'axes, et la nécessité d'étudier au moins 9 axes pour synthétiser plus de 80 % de la variance spatio-temporelle initiale montre qu'il n'y a pas de direction privilégiée dans le nuage de points, *i.e.* il n'y a pas de structure spatiale cohérente se répétant d'une année à l'autre.

3.1.3. Analyse temporelle par pays

3.1.3.a. Analyse des fréquences épidémiques par pays

L'objectif de ce paragraphe est d'analyser l'évolution de la fréquence des épidémies au cours du temps, mais à une résolution nationale au sein du Pacifique. Nous avons donc analysé les fréquences épidémiques dans chacun des vingt PTIP en calculant la pente de la droite de régression, par rapport au temps, du nombre d'épidémies sur une fenêtre glissante de 10 ans (Figure 3.7.A pour une illustration en Nouvelle-Calédonie).

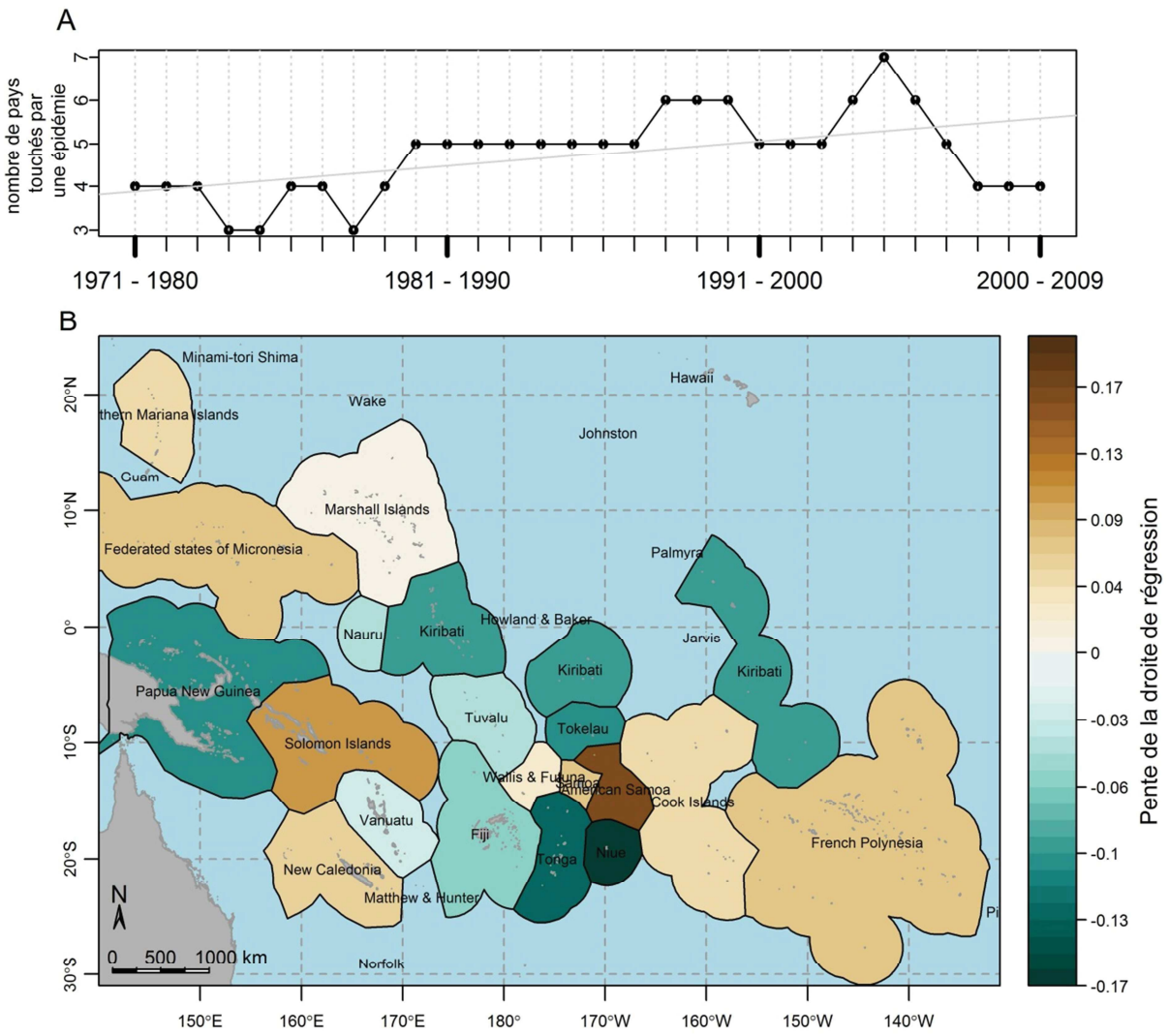


Figure 3.7 : Evolution de la fréquence des épidémies dans le Pacifique, *résolution spatiale des pays.* ***A.*** *Nombre d'épidémies par fenêtre de 10 ans en Nouvelle-Calédonie. La droite grise est la droite de régression, dont la pente représente l'augmentation annuelle du nombre d'épidémie par fenêtre de 10 ans.* ***B.*** *Carte des pentes des droites de régression pour chacun des pays, calculées comme illustré en A.*

Tous pays confondus, la fréquence des épidémies n'augmente pas dans le Pacifique (Figure 3.5). Mais ce résultat masque l'hétérogénéité existant d'un pays à l'autre lorsque l'on augmente la résolution spatiale d'analyse. En effet, les épidémies de dengue sont de plus en plus fréquentes dans les Etats Fédérées de Micronésie, en Nouvelle-Calédonie, en Polynésie Française, aux îles Cook ou aux Samoa Américaines, tandis qu'elles se font plus rares dans les autres pays. Aux îles Salomon ou aux îles Marshall, le nombre d'épidémies répertoriées est trop faible pour pouvoir interpréter les résultats de cette analyse.

Il ne semble pas exister de structure spatiale particulière, comme un gradient nord/sud, ou un gradient est/ouest. Les facteurs liés à l'augmentation ou à la diminution de la fréquence des épidémies au sein d'un pays semblent donc être des facteurs intrinsèques à chacun des pays, comme l'immunité de

population, les mouvements internationaux, ou encore les actions de lutte, plutôt que des facteurs environnementaux à large échelle comme le climat.

3.1.3.b. Caractérisation des profils dynamiques

Dans ce paragraphe, nous visons à regrouper les pays ayant éventuellement des profils dynamiques similaires, afin d'émettre des hypothèses sur les causes des variations spatiales de survenue d'épidémie. Dans le Pacifique, l'inspection visuelle des courbes épidémiques montre que la dynamique de la dengue diffère d'un pays à l'autre. En effet, comme montré sur la Figure 3.8, il semble exister trois profils dynamiques différents :

- Un profil épidémique sporadique : il n'y a aucun cas déclaré durant la majorité des années, et des épidémies de forte ampleur (pour le pays considéré) surviennent de manière sporadique, très espacées dans le temps. Exemple : îles Salomon (Figure 3.8. A et B) ;
- Un profil épidémique régulier : des épidémies majeures surviennent régulièrement, séparées par de courtes périodes au cours desquelles les seuls cas déclarés sont des cas importés ne donnant pas lieu à des épidémies. Exemple : Nouvelle-Calédonie (Figure 3.8. C et D) ;
- Un profil endémo-épidémique : des épidémies surviennent presque tous les ans, sans qu'il y a ait de période nette d'arrêt de la circulation virale entre les épidémies. Exemple : Polynésie Française (Figure 3.8. E et F).

Nous avons tenté de classer les vingt PTIP étudiés selon ces trois profils dynamiques de manière quantitative ou semi-quantitative par deux méthodes :

- Une méthode de classification ascendante hiérarchique sur les courbes de taux d'incidence standardisés triés par ordre croissant, utilisant la distance euclidienne et l'algorithme d'agrégation de Ward (Ward, 1963) ;
- Une méthode basée sur l'estimation empirique de trois paramètres des courbes de taux d'incidence standardisés : le nombre d'années non épidémiques (élevé pour le profil épidémique sporadique, plus faible pour les deux autres profils), la moyenne des taux incidences standardisés durant les années non épidémiques (plus élevé pour les profils endémo-épidémiques), et la pente de la courbe formée par les années épidémiques, lorsque les taux d'incidence standardisés sont triés par ordre croissant (plus élevée pour les profils sporadiques).

Malheureusement, ces deux méthodes n'aboutissent pas à une classification cohérente des dynamiques des différents pays, et nous ne pouvons donc présenter une carte de répartition géographique des différents profils.

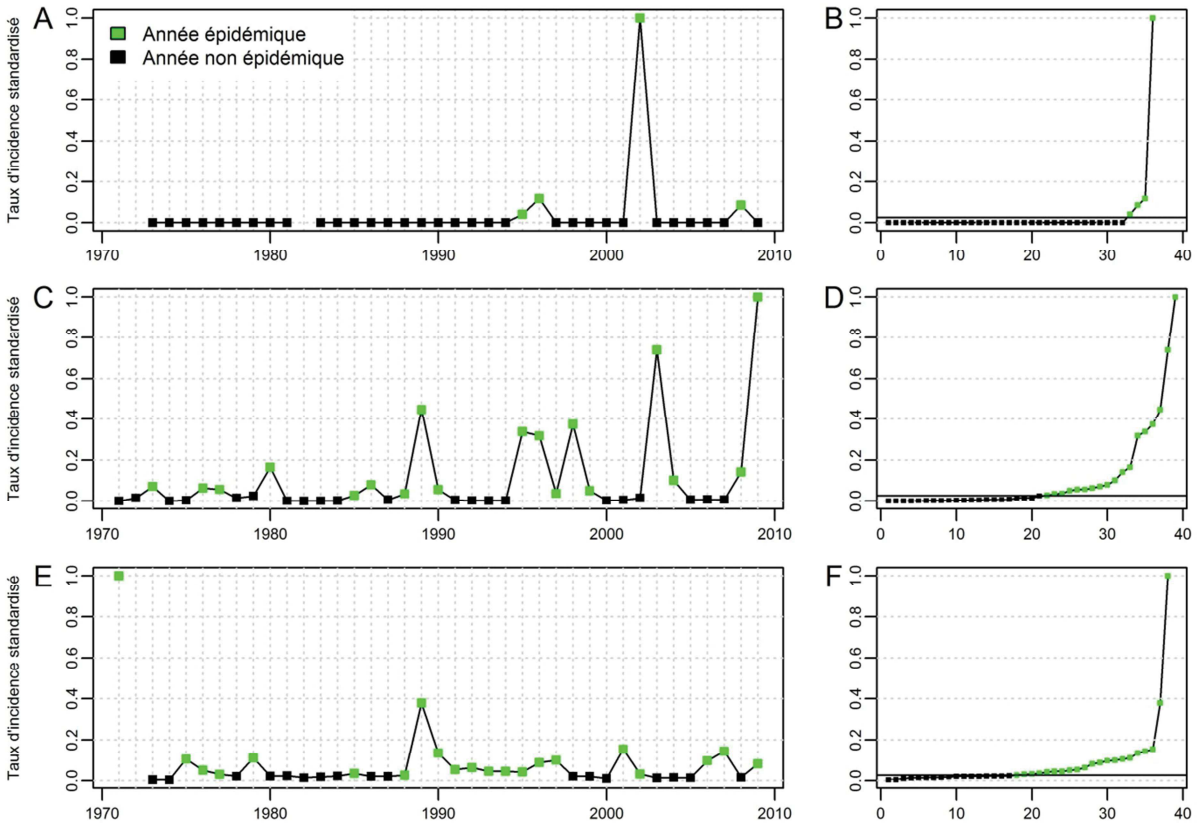


Figure 3.8 : Différents profils épidémiques dans les îles du Pacifique. *Les courbes représentent les taux d'incidence standardisés classés par ordre chronologique (A, C et E), ou par ordre croissant (B,D et F) pour les îles Salomon (A et B), la Nouvelle-Calédonie (C et D) et la Polynésie Française (E et F).*

3.1.4. Analyse par vague pandémique

Nous avons défini une vague pandémique comme la période de circulation d'un sérotype donné sans qu'il y ait interruption (au moins un pays déclare un cas du à ce sérotype chaque année composant la vague pandémique), et causant une épidémie dans au moins quatre PTIP. Nous avons ainsi distingué huit vagues pandémiques :

- Vague 1 : de 1971 à 1974 (sérotype 2, 8 pays touchés) ;
- Vague 2 : de 1974 à 1977 (sérotype 1, 11 pays touchés) ;
- Vague 3 : de 1979 à 1981 (sérotype 4, 11 pays touchés) ;
- Vague 4 : de 1989 à 1996 (sérotype 3, 7 pays touchés) ;
- Vague 5 : de 1996 à 1999 (sérotype 2, 9 pays touchés) ;
- Vague 6 : de 2000 à 2004 (sérotype 1, 13 pays touchés) ;
- Vague 7 : de 2006 à 2008 (sérotype 1, 5 pays touchés) ;
- Vague 8 : de 2008 à 2009 (sérotype 4, 5 pays touchés, toujours en cours en 2010).

3.1.4.a. Etude de la propagation d'un sérotype au sein du Pacifique

L'objectif de ce paragraphe est de répondre aux questions suivantes : un sérotype nouveau émerge-t-il systématiquement dans un même pays ou groupe de pays ? Une fois un nouveau sérotype détecté, existe-t-il des trajectoires de propagation systématiques ?

Afin d'étudier la localisation de l'initiation des vagues pandémiques, pour chacune des vagues, nous avons identifié, pour chacun des vingt PTIP, l'année au cours de laquelle le sérotype nouveau provoque la première épidémie (Figure 3.9). Nous pouvons remarquer que la Polynésie Française semble souvent touchée en début de vague, les îles Cook sont souvent touchées l'année suivant l'apparition d'une épidémie en Polynésie Française, tandis que la Nouvelle-Calédonie semble touchée plutôt en fin de vague.

Afin de vérifier et de synthétiser cette information, nous avons compté, pour chacun des pays, le nombre de fois où il était touché lors de la première ou de la dernière année d'une vague pandémique (Table 3.1). Cette analyse confirme la survenue précoce des épidémies dues à un nouveau sérotype en Polynésie Française. D'autre part la Nouvelle-Calédonie, qui a été touchée par l'ensemble des vagues pandémiques, est presque toujours (7 fois sur 8) touchée après d'autres pays. La surveillance régionale de la dengue dans le pacifique pourrait donc éventuellement améliorer la prédiction du risque épidémique dans ce pays.

L'étude des cartes d'initiation des vagues pandémiques (Figure 3.9) semble montrer qu'il n'existe pas de trajectoire de propagation systématique particulière. Cela montre que la part de la variation interannuelle de survenue d'épidémie non expliquée par la corrélation avec la SOI dans un pays donné est liée à des facteurs intrinsèques au pays, ce qui est cohérent avec le résultat trouvé à propos des fréquences épidémiques au sein de chacun des pays.

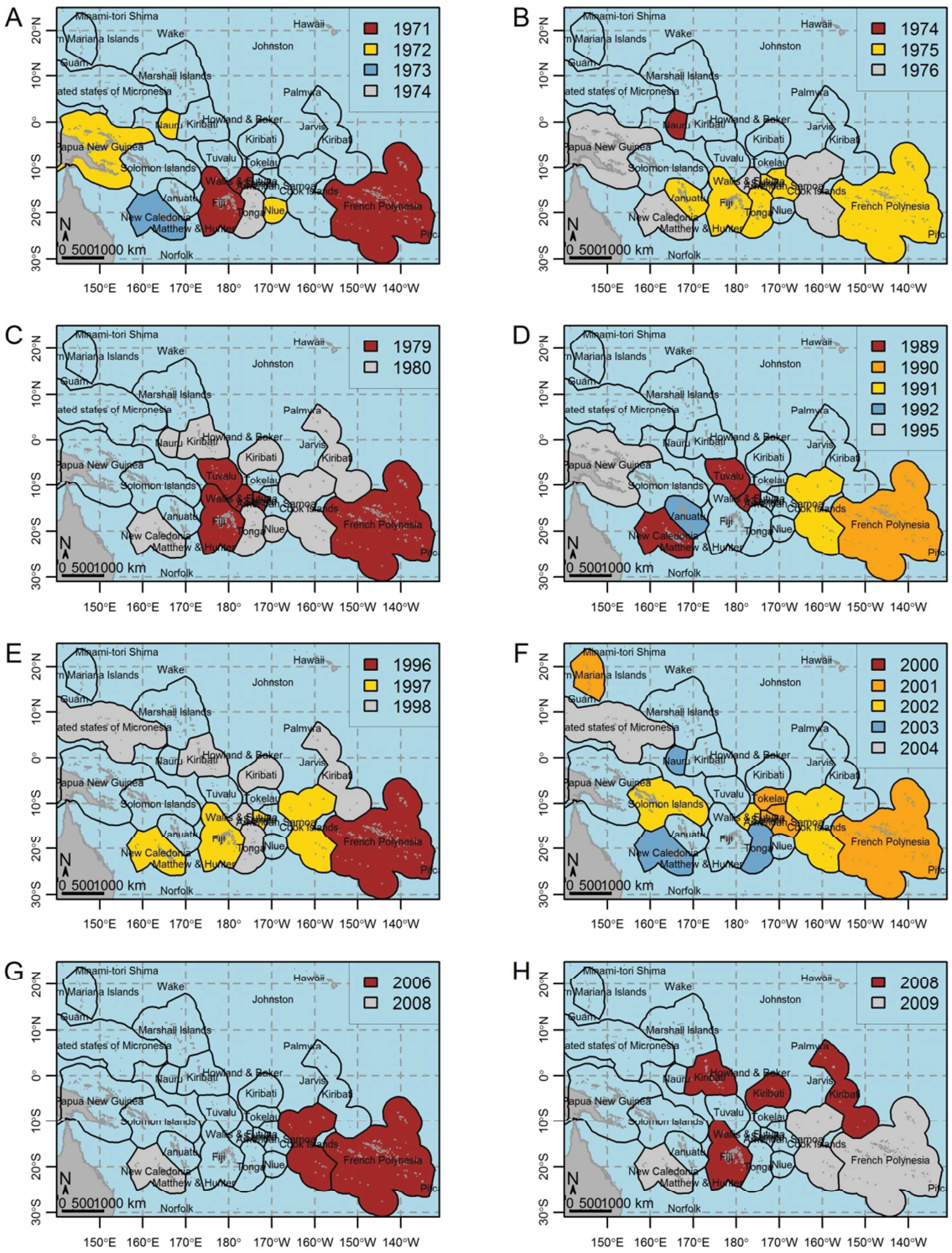


Figure 3.9 : Année de début de circulation d'un nouveau sérotype dans le Pacifique. A. *circulation du sérotype 2 entre 1971 et 1974 ;* ***B.*** *circulation du sérotype 1 entre 1974 et 1977 ;* ***C.*** *Circulation du sérotype 4 entre 1979 et 1981 ;* ***D.*** *Circulation du sérotype 3 entre 1989 et 1996 ;* ***E.*** *Circulation du sérotype 2 entre 1996 et 1999 ;* ***F.*** *Circulation du sérotype 1 entre 2000 et 2004 ;* ***G.*** *Circulation du sérotype 1entre 2006 et 2008 ;* ***.*** *Circulation du sérotype 4 entre 2008 et 2009.*

Table 3.1 : Comparaison de la précocité d'apparition d'un nouveau sérotype dans les vingt PTIP

Pays	Nb premier*	Nb dernier**	Nb vague***
FP	4	1	8
Fiji	3	0	5
Wallis	3	2	6
Tuvalu	2	0	2
Palau	1	0	1
NC	1	4	8
Nauru	1	1	4
Western Samoa	1	0	4
Cook	1	3	7
Kiribati	1	2	3
Northern Mariana	0	0	1
PNG	0	2	3
FSM	0	2	2
Solomon	0	0	1
Vanuatu	0	0	2
Marshall	0	0	0
Tonga	0	3	5
Tokelau	0	0	1
American Samoa	0	0	2
Niue	0	1	2

* Nombre de fois où le pays a été touché dès la première année d'une vague pandémique (trié par ordre croissant)
** Nombre de fois où le pays a été touché la dernière année d'une vague pandémique
*** Nombre total de fois où le pays a été touché par une vague pandémique

3.1.4.b. Etude de la structure spatiale des vagues

Afin de déterminer s'il existe une structure spatiale se répétant d'une vague à l'autre, nous avons réalisé une carte par vague pandémique montrant l'incidence maximale observée dans chacun des pays (Figure 3.10). Comme pour les analyses temporelles, il ne semble pas exister de schéma spatial systématique, laissant penser que les facteurs environnementaux à grande échelle n'ont que peu d'influence sur la variabilité spatiale de la dengue.

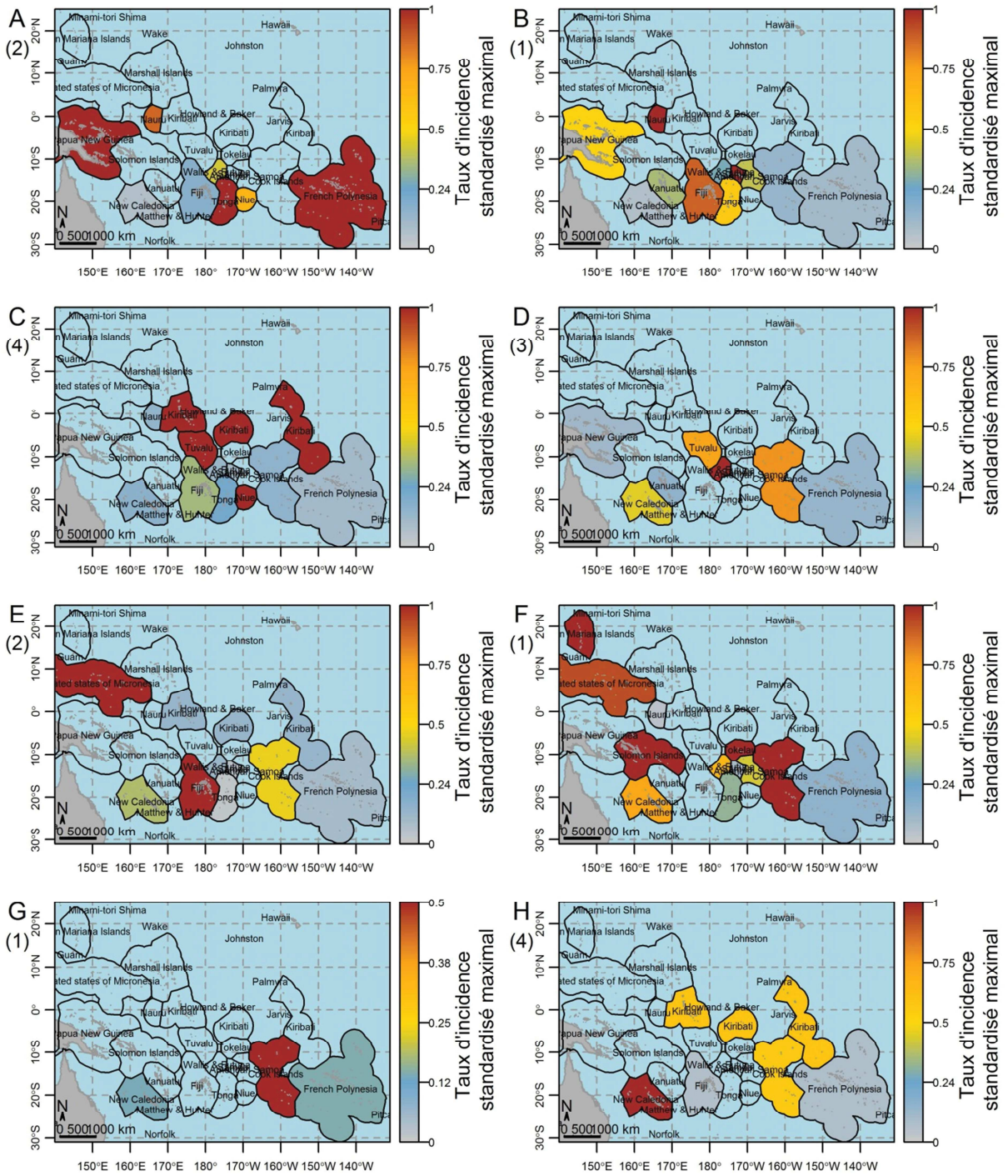


Figure 3.10 : Carte des pays les plus touchés par chacune des vagues épidémiques. *La couleur code pour le taux d'incidence standardisé maximal observé durant une vague. L'échelle de couleurs va du minimum au maximum observé pour chaque vague, afin de faciliter les comparaisons entre vagues. Le chiffre entre parenthèses sous la lettre correspondant à la figure indique le sérotype responsable de la vague épidémique.* ***A.*** *Circulation du sérotype 2 entre 1971 et 1974 ;* ***B.*** *Circulation du sérotype 1 entre 1974 et 1977 ;* ***C.*** *Circulation du sérotype 4 entre 1979 et 1981 ;* ***D.*** *Circulation du sérotype 3 entre 1989 et 1996 ;* ***E.*** *Circulation du sérotype 2 entre 1996 et 1999 ;* ***F.*** *Circulation du sérotype 1 entre 2000 et 2004 ;* ***G.*** *Circulation du sérotype 1entre 2006 et 2008 ; . Circulation du sérotype 4 entre 2008 et 2009.*

3.1.4.c. Dynamique des vagues épidémiques

La Figure 3.11 montre la dynamique de la dengue sérotype par sérotype, tous pays confondus.

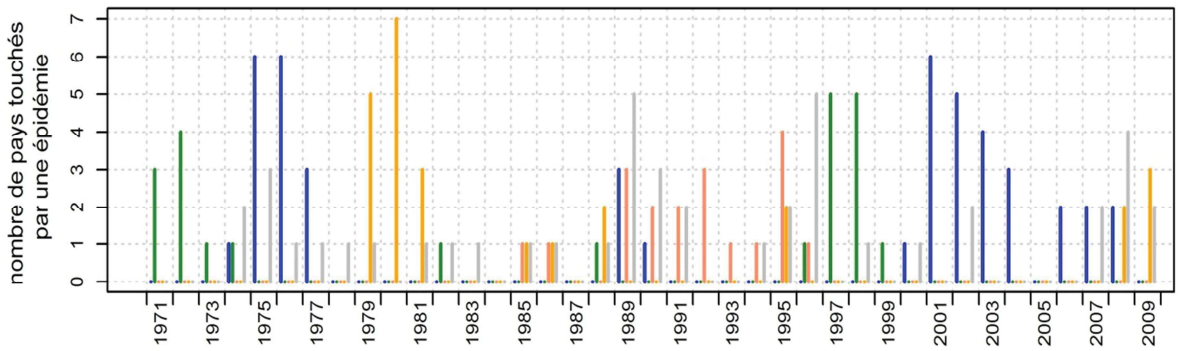


Figure 3.11 : Dynamique des quatre sérotypes de dengue dans le Pacifique entre 1971 et 2009. *La figure représente le nombre d'épidémies déclarées au cours du temps pour chacun des quatre sérotypes de dengue. La couleur correspond au sérotype dominant. Le sérotype inconnu correspond soit à une absence d'information concernant le sérotype en cause, soit à la co-circulation de plusieurs sérotypes sans que l'information sur le sérotype dominant soit disponible dans les données.*

Nous pouvons remarquer que l'ensemble des vagues épidémiques a débuté dans un nombre limité de pays (de 1 à 3) avant d'exploser dans l'ensemble de la région. Un système d'alerte régional devrait donc permettre à certains pays touchés peu précocement d'anticiper le risque épidémique. Les années 1985 à 1989 sont particulières du fait de la présence d'épidémies dues à différents sérotypes déclarées à différents endroits du Pacifique au cours de la même année. La faible circulation apparente du DENV-3 entre les années 1989 et 1995 est due à la présence de co-circulation de plusieurs sérotypes au sein d'un même pays dans de nombreux pays entre les années 1989 et 1995. La co-circulation enregistrée à cette période-là impliquait le DENV-3 et soit le DENV-4, soit le DENV-1, sans que l'information sur le sérotype dominant ne soit disponible dans les données à disposition.

3.1.5. Discussion et perspectives

3.1.5.a. Résumé des résultats à l'échelle régionale

Après plus de trente ans d'absence, la dengue se propage depuis 1971 dans le Pacifique sous la forme de vagues épidémiques survenant de manière cyclique tous les 4 ans (années 1970) ou tous les 5 à 7 ans (depuis la fin des années 1980). Contrairement aux autres régions du monde, la co-circulation de plusieurs sérotypes est rare, et chaque vague est due à l'émergence et à la circulation dominante d'un seul des quatre sérotypes. L'émergence d'un sérotype ayant déjà circulé semble survenir tous les 20 à 24 ans. Lors de l'arrivée d'un nouveau sérotype, la Polynésie Française est touchée de manière plus précoce que les autres pays. Par contre il ne semble pas exister de schéma de propagation particulier à partir du pays d'émergence, ou de structure spatiale particulière se répétant d'une vague à l'autre, ni de

cohérence entre les dynamiques des différents pays, ou de structure régionale de la fréquence des épidémies au sein d'un pays.

3.1.5.b. Emergence d'un sérotype dominant

L'émergence et la diffusion d'un sérotype dominant semble être un phénomène mondial. En effet, à partir des années 1970, de larges épidémies touchant les Caraïbes, l'Amérique Centrale et le Nord de l'Amérique du Sud ont succédé de près à celles du Pacifique, avec le même sérotype dominant (circulation du sérotype 1 de 1977 à 1980, du sérotype 4 de 1981 à 1988, et du sérotype 3 après plus de 15 ans d'absence, de 1994 au début des années 2000 (Gubler, 1998; Guzman and Kouri, 2003)).

L'émergence et la circulation d'un nouveau sérotype dans le Pacifique pourrait être due soit à l'introduction de ce sérotype à partir d'autres régions du monde, soit à la résurgence locale d'un sérotype circulant de manière silencieuse. L'absence complète de circulation virale apparente pendant près de 30 ans avant les années 1970 rend l'hypothèse d'introduction à partir de pays endémiques plus plausible. Les études phylogénétiques ont confirmé cette hypothèse et ont montré que les introductions ont lieu à partir de pays d'Asie ou d'Amérique (A-Nuegoonpipat et al., 2004; Dupont-Rouzeyrol et al., 2014; Steel et al., 2010).

Au vu de l'absence de schéma de propagation particulier, il est légitime de se demander si les différentes épidémies survenant au cours d'une vague épidémique sont le fait de l'introduction d'un sérotype à partir de pays endémiques suivie de la diffusion régionale de ce sérotype dans le Pacifique, ou bien si chacun des pays du Pacifique subit des introductions indépendantes les unes des autres à partir d'autres régions du monde. Les études phylogénétiques montrent que les vagues de sérotype 2 de 1971 à 1974 (Steel et al., 2010), de sérotype 3 de 1989 à 1995 (Chungue et al., 1993) et de sérotype 4 de 2008 à 2009 (Li et al., 2010) sont dues à l'introduction unique d'une souche virale suivie de la diffusion régionale de la souche. Par contre, d'autres études montrent des introductions multiples de plusieurs génotypes d'un même sérotype à partir de différents pays endémiques pour les vagues dues au sérotype 1 de 2000 à 2004 (A-Nuegoonpipat et al., 2004; Dupont-Rouzeyrol et al., 2014), et au sérotype 3 en 2013 (Cao-Lormeau et al., 2014a). Plus précisément, la circulation du DENV-2 de 1971 à 1974 serait consécutive à l'introduction d'une souche unique de DENV-2 en 1971 en Polynésie Française et à Fiji à partir d'Amérique (Steel et al., 2010). La circulation du DENV-3 entre 1989 et 1995 serait due à l'introduction d'une souche unique de DENV-3 en 1988 depuis l'Asie (Chungue et al., 1993). De 2000 à 2004 dans le Pacifique (A-Nuegoonpipat et al., 2004), ainsi qu'en 2002/2003 en Nouvelle-Calédonie (Dupont-Rouzeyrol et al., 2014), différents génotypes de DENV-1 originaires d'Asie ou du Pacifique ont circulé. De même, les épidémies de DENV-3 observées en Polynésie Française et aux Salomon en 2013 sont dues à l'introduction respective d'un génotype de Guyane Française, et d'un génotype d'Asie (Cao-Lormeau et al., 2014a). L'ensemble des études phylogénétiques montre l'existence d'une circulation régionale des souches introduites, qu'il y ait eu une ou plusieurs introductions (A-

Nuegoonpipat et al., 2004; Cáceres et al., 2008; Cao-Lormeau et al., 2011; Chungue et al., 1993; Dupont-Rouzeyrol et al., 2014; Effler et al., 2005; Li et al., 2010; Steel et al., 2010).

Les pays d'Asie du Sud-Est et les pays d'Amérique Latine étant hyperendémiques, il est logique de supposer alors que les quatre sérotypes sont introduits chaque année via des voyageurs infectés. Il serait donc intéressant de comprendre pourquoi un des sérotypes prend le dessus sur les autres. Les hypothèses pouvant expliquer ce phénomène sont nombreuses et complexes.

La première est une hypothèse d'immunité de population. Plusieurs études de modélisation ont montré que les phénomènes immunitaires et les interactions hôte/pathogène qu'ils engendrent pouvaient expliquer entièrement la dynamique observée en Asie du Sud-Est (Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006), notamment l'alternance de sérotypes dominant. On peut supposer que dans le Pacifique, la population est d'autant plus susceptible à un sérotype donné que ce sérotype n'a pas circulé depuis longtemps. Cependant, cette hypothèse n'explique pas la dominance d'un sérotype donné lors de sa première apparition dans les années 1970 ou 1980, après plus de trente ans d'absence de circulation apparente. D'autre part, les niveaux de circulation virale sont beaucoup moins élevés dans les pays du Pacifique qu'en Asie, et il n'est pas sûr que le seuil d'immunité de population suffisant à empêcher la survenue d'une nouvelle-épidémie soit atteint au bout de sept ans, comme l'a montré l'épidémie de DENV-1 survenue en Nouvelle-Calédonie en 2013, après plus 10 ans de circulation de ce sérotype dans le pays.

Une autre hypothèse pouvant expliquer la dominance d'un sérotype par rapport aux autres serait un lien épidémiologique avec les pays d'Amérique ou d'Asie. En effet, nous pouvons supposer que l'émergence d'un sérotype est facilitée si le sérotype est introduit plus fréquemment que les autres et que le nombre de virus d'un sérotype donné introduit à partir des régions endémiques est proportionnel à l'intensité de la circulation virale de chacun des sérotypes dans les pays d'origine (voir Figure 3.12 pour une illustration en Thaïlande, par exemple l'augmentation progressive de la proportion de DENV-1 circulant en Thaïlande pourrait peut-être entraîner une augmentation de l'introduction du sérotype 1 dans le Pacifique et ainsi favoriser son émergece). Si des données relatives à l'abondance des différents sérotypes existent pour quelques pays d'Asie (Huy et al., 2010; Zhang et al., 2005) il n'existe pas encore de revue régionale sur les sérotypes ayant circulé et dominé en Asie du Sud-Est permettant d'étudier cette hypothèse de manière plus précise.

Que ce soit en Asie, en Amérique ou dans le Pacifique une autre hypothèse pour expliquer l'émergence et la domination de la circulation d'un sérotype par rapport aux autres est celle de particularités génétiques intrinsèques au génotype introduit, et qui pourrait conférer à ce génotype un avantage évolutif. En effet, une étude en Thaïlande a montré qu'il existait une corrélation positive entre la diversité génétique au sein d'un sérotype et sa prévalence (Zhang et al., 2005). D'autre études au Cambodge (Duong et al., 2013a) et au Vietnam (Ty Hang et al., 2010) ont montré une dynamique

complexe des lignées génétiques au sein d'un même sérotype, avec des remplacements majeurs d'un génotype par un autre parfois concomitants avec des changements observés de sérotype dominant.

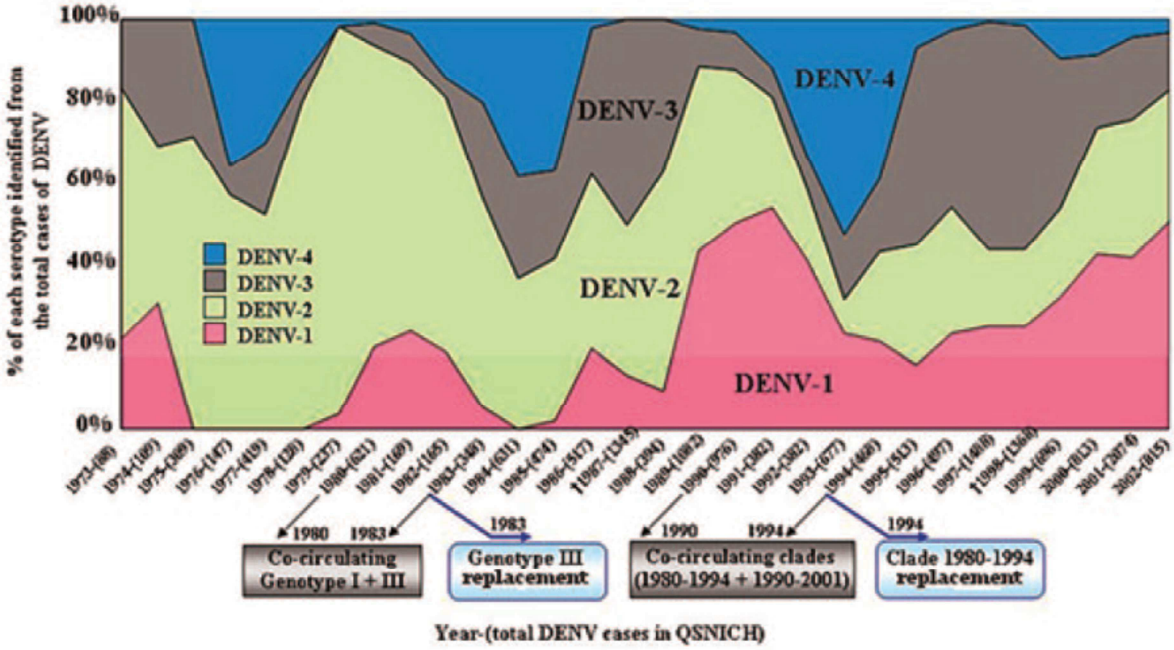


Figure 3.12 : Abondance relative des différents sérotypes en Thaïlande, 1973-2002. *Cette figure est reproduite à partir de Zhang et al. (Zhang et al., 2005), et a été construite à partir des cas hospitalisés dans un hôpital de Bangkok, Thaïlande. Le sérotype dominant en Thaïlande varie au cours du temps : DENV-2 de 1973 à 1986 ; DENV-3 en 1987 ; DENV-2 en 1988-1989 ; DENV-1 de 1990 à 1992 ; DENV-4 en 1993 et 1994 ; DENV-3 de 1995 à 1999 ; et DENV-1 de 1999 à 2002. Nous pouvons remarquer que, comme dans le Pacifique, le sérotype 4 ne circule que de manière épisodique et durant des périodes plus courtes que les autres sérotypes. Nous pouvons également remarquer l'augmentation progressive de la circulation du DENV-1 depuis la fin des années 1980 et l'opposition de phase existant entre le DENV-1 et le DENV-4.*

Il serait intéressant de comprendre les mécanismes impliqués dans l'émergence d'un sérotype donné plutôt qu'un autre. Les études pourraient inclure :

- Des études phylogénétiques similaires à celles de Duong *et al.* au Cambodge (Duong et al., 2013a, 2013b) afin d'étudier la dynamique des lignées génotypiques. Ces études phylogénétiques pourraient se concentrer sur la période 1988-1996, période au cours de laquelle le sérotype 3 n'a pas totalement dominé dans le Pacifique, et au cours de laquelle des épidémies dues à d'autres sérotypes ont persisté.

- La construction de modèles déterministes de type SEIR afin de tester les hypothèses relatives aux phénomènes immunitaires. De telles études seraient nettement facilitées par la réalisation préalable d'enquêtes de séroprévalence dans les pays du Pacifique.

L'augmentation du nombre d'introductions depuis d'autres régions endémiques ainsi que l'observation d'une co-circulation de plusieurs sérotypes au sein d'un même pays comme en Nouvelle-Calédonie en 2008/2009 (DENV-1 et DENV-4) ou en Polynésie Française en 2013 (DENV-1 et DENV-3 (Cao-Lormeau et al., 2014a)) montrent que l'épidémiologie de la dengue dans le Pacifique pourrait se transformer pour aboutir à une situation hyper-endémique comme cela a été le cas en Asie dans les années 1970 ou en Amérique dans les années 2000 (Guzman and Kouri, 2003).

3.1.5.c. Dynamique cyclique des pandémies de dengue dans le Pacifique

De par la résolution annuelle des données étudiées, nous n'avons pu étudier que les variations interannuelles (cycles dont la période est supérieure à 2 ans) à cette échelle régionale. Les pandémies de dengue dans le Pacifique présentent une variabilité interannuelle régulière : cycle de 4 ans dans les années 1970, et de 5 à 7 ans depuis les années 1980. Il serait intéressant de connaître les causes de cette synchronisation régionale des variations interannuelles de la dengue.

Dans les situations hyperendémiques comme en Asie, la dengue présente également une dynamique complexe mêlant des cycles saisonniers et des cycles interannuels avec une alternance des différents sérotypes. La période de ces cycles interannuels est légèrement plus courte que dans le Pacifique (périodicité de 3 ans des courbes épidémiques globales, et périodicité de chacun des sérotypes de 9 à 10 ans en Thaïlande (Cummings et al., 2004; Wearing and Rohani, 2006) et au Cambodge (Huy et al., 2010)). Si le cycle saisonnier est clairement le résultat de l'impact saisonnier du climat sur les populations vectorielles, les facteurs responsables de la dynamique interannuelle observée en Asie ne sont pas encore bien déterminés. Certaines études montrent une relation transitoire entre la dynamique interannuelle de la dengue et des facteurs extrinsèques climatiques tels ENSO (Cazelles et al., 2005), tandis que d'autres soulignent l'importance de facteurs intrinsèques comme l'immunité de population (Hay et al., 2000), la présence transitoire d'une immunité croisée entre les différents sérotypes (Wearing and Rohani, 2006) ou l'existence du phénomène d'« anticorps facilitants » (Ferguson et al., 1999). Cependant, les études utilisant des modèles déterministes explicitant clairement les mécanismes immunitaires en jeu (Ferguson et al., 1999; Wearing and Rohani, 2006) supposent la co-circulation de plusieurs sérotypes au même endroit. Ce qui n'est pas le cas du Pacifique.

Dans le Pacifique, en l'absence de co-circulation des différents sérotypes, et sachant que les différentes vagues sont toutes dues à des introductions à partir de pays endémiques, nous pouvons imaginer plusieurs scénarii conduisant à la périodicité observée : i) Des facteurs environnementaux à large échelle, comme le climat, forcent la dynamique de la dengue dans l'ensemble des PTIP. Le phénomène El Niño, qui a une périodicité de 4 à 8 ans, et une forte influence sur les variations interannuelles du climat dans le Pacifique est un bon candidat ; ii) Des phénomènes environnementaux cycliques forcent la dynamique de la dengue uniquement dans certains pays clés, comme la Polynésie Française, et la

dynamique des îles alentours est la même, du fait de la diffusion régionale d'un sérotype donné ; iii) La périodicité observée dans le Pacifique est forcée par celle observée en Asie ; iv) des facteurs intrinsèques comme l'immunité de population, ou la densité de population régulent le cycle interannuel de la dengue ; v) un mélange de i, ii, iii et iv.

Nous n'avons trouvé qu'un faible impact du phénomène El Niño comme élément synchronisateur des épidémies sur l'ensemble du Pacifique (corrélation de 0,35 entre SOI et nombre annuel d'épidémies). La différence entre nos résultats et ceux de Hales (Hales et al., 1996) provient certainement des données utilisées (la base de données de Hales a été complétée par Descloux), ainsi que de la définition utilisée pour déclarer une année comme étant épidémique. La faible corrélation globale peut provenir de l'impact différentiel d'El Niño dans les différentes régions du Pacifique. En effet, durant un phénomène El Niño, si les températures augmentent à l'est du Pacifique, elles diminuent à l'ouest, et il serait intéressant de mener des études temporelles entre ENSO et les variables climatiques locales, puis entre ENSO et les dynamiques épidémiques au sein de chacun des pays séparément, à l'image de l'étude Hales (Hales et al., 1999). La faiblesse de la corrélation montre l'importance d'étudier d'autres facteurs tels que les facteurs démographiques, immunitaires ou les mouvements de population dans la synchronisation régionale des épidémies.

Les différences entre le nombre de pays touchés par une épidémie chaque année tel que calculé ici et tel que calculé par Hales illsutre une limite de l'analyse des données régionales, et la sensibilité des résultats à la définition utilisée pour définr un évènement comme épidémique.

Le scénario ii) paraît plausible, et est étayé par les résultats d'une étude sur l'impact du phénomène El Niño dans le Pacifique (Hales et al., 1999). Cette étude a montré une corrélation positive entre les séries mensuelle de cas de dengue et le SOI dans 5 des 14 pays du Pacifique étudiés (Tokelau, Samoa, Samoa Américaines, Fiji, et Tonga). Elle a également montré des schémas de propagation régionale depuis les grandes îles vers les plus petites : depuis Fiji vers Tonga, Samoa, Samoa Américaines, et Tokelau d'une part, et depuis la Polynésie Française vers la Nouvelle-Calédonie, le Vanuatu et Wallis d'autre part.

Cependant, à part aux îles Fiji, les pays dans lesquels El Niño semble avoir une influence sur la dynamique interannuelle ne sont pas les pays dans lesquels les différentes vagues épidémiques sont initiées (Polynésie Française, Wallis et Fiji, voir Table 3.). El Niño peut donc favoriser l'apparition d'une épidémie dans ces pays-là une fois un nouveau sérotype installé, mais il n'explique pas la périodicité de l'émergence d'un nouveau sérotype.

D'autre part, nos résultats concernant l'absence de dynamique de propagation d'une vague à l'autre, et l'absence de groupes de pays ayant des dynamiques similaires sont incohérents avec l'étude de Hales. Ceci peut être dû au fait que Hales a travaillé sur des données mensuelles, et nous sur des données annuelles. De plus les donées analysées ici sont des données agrégées annuellement de janvier à janvier : pour les pays du Pacifique Sud, ce découpage ne correspond pas aux périodes épidémiques annuelles, et

peut répartir les cas d'une seule épidémie sur deux années, ce qui peut biaser les données. Il est également possible que les corrélations positives trouvées par Hales ne soient vérifiées que de manière transitoire sur la période d'étude, les séries épidémiologiques étant souvent non stationnaires.

Le scénario iii) est peu plausible, du fait de la différence existant entre la périodicité de la dengue en Asie et celle du Pacifique (3 ans, vs 5 à 7 ans). Enfin le scénario iv) nécessiterait une enquête de séroprévalence sérotype spécifique dans plusieurs pays ou bien la construction de modèles déterministes afin d'être testée.

Il semble important d'arriver à mieux comprendre les mécanismes impliqués dans les variations interannuelles de la dengue dans le Pacifique, et notamment d'arriver à distinguer l'influence de facteurs intrinsèques (immunité) de celle des facteurs extrinsèques (climat, environnement). Les pistes d'étude sont :

- une étude approfondie des phénomènes de propagation d'un sérotype donné au sein du Pacifique, afin de reconstruire l'histoire précise de la dengue pour chacune des vagues pandémiques, et ainsi infirmer ou confirmer les liens épidémiologiques existant entre les pays. Cette connaissance peut être acquise soit au moyen d'analyse et de modélisation des données épidémiologiques, soit au moyen d'études phylogénétiques fines. L'analyse et la modélisation devra utiliser des techniques permettant de gérer la non-stationnarité des données épidémiologiques (comme l'analyse en ondelettes), et ne donnera vraisemblablement des résultats intéressants que si elle est effectuée sur des données à résolution mensuelle.
- La construction d'un modèle déterministe multi-sérotype afin de tester les hypothèses liées à l'immunité de l'hôte ou à la virulence des souches virales introduites dans la périodicité observée.

L'absence de co-circulation prolongée dans cette région du monde pourrait simplifier la tâche et éclairer les connaissances sur la dynamique plus complexe existant dans les pays hyperendémiques.

3.1.5.d. Particularités épidémiologiques inexpliquées

Circulation prolongée du DENV-1

Le sérotype 1 présente une dynamique atypique par rapport à celle des autres sérotypes. En effet, les études phylogénétiques montrent que la vague pandémique de 2006 à 2008 est due à des virus génétiquement très proches de ceux isolés entre 2000 et 2004 (Dupont-Rouzeyrol et al., 2014). On peut donc considérer les vagues de 2000 à 2004 et de 2006 à 2008 comme une seule et même vague, impliquant une circulation du virus DENV-1 sur l'ensemble des îles du Pacifique pendant près de dix ans. D'autre part, alors qu'à partir de 2008 le sérotype 4 a progressivement remplacé le sérotype 1 dans la plupart des îles du Pacifique, en Nouvelle-Calédonie le virus DENV-4 n'a provoqué qu'une seule

épidémie en 2008/2009, et a ensuite complètement disparu, sans remplacer le DENV-1, qui a provoqué en 2012/2013 la plus grande épidémie jamais enregistrée dans le pays. La réapparition du DENV-1 en Nouvelle-Calédonie puis dans d'autres îles du Pacifique en 2013, après plus de 10 ans de circulation, et alors que le DENV-4 circulait dans l'ensemble du Pacifique est un phénomène remarquable. Une étude phylogénétique a montré que la résurgence du DENV-1 en 2012 en Nouvelle-Calédonie est due à l'introduction d'un nouveau génotype à partir d'Asie, différent de celui qui avait circulé de 2000 à 2008 (Dupont-Rouzeyrol et al., 2014). Cette circulation prolongée dans un pays relativement peuplé comme la Nouvelle-Calédonie pose une question intéressante sur l'immunité de population ; en effet, cet évènement semble indiquer que les épidémies dans les îles du Pacifique sont d'une ampleur suffisamment faible, ou bien que la population d'hôtes susceptibles est renouvelée de manière suffisamment rapide dans certains pays pour permettre l'émergence d'un même sérotype pendant plusieurs années de suite. Les pays du Pacifique ne sont donc pas à l'abri d'une augmentation de la fréquence des épidémies dans le futur (par exemple due au réchauffement climatique (Githeko et al., 2000; Patz et al., 1998)). D'autre part, ce phénomène est un phénomène intéressant à étudier pour répondre à la question des facteurs impliqués dans l'émergence d'un des quatre sérotypes par rapport aux autres.

Co-circulation du virus DENV-3 avec d'autres sérotypes

Le sérotype 3 est le seul sérotype pour lequel une co-circulation a été enregistrée de manière prolongée dans le Pacifique. Sa circulation n'a pas empêché la survenue de quelques épidémies dues à d'autres sérotypes dans certains pays entre 1985 et 1989. En 2103, ce sérotype a co-circulé avec le DENV-1 en Polynésie Française, sans que nous ne sachions encore lequel des deux sérotypes prendra le dessus.

Si le schéma épidémiologique observé à la fin des années 1980 se répète, nous pouvons nous attendre à une dominance de la circulation du virus DENV-3 dans le Pacifique dans les années qui viennent (comme déjà anticipé (Singh et al., 2005)), tout en enregistrant des épidémies dues à d'autres sérotypes dans un nombre restreint de pays, ou des cas de co-circulation au sein d'un même pays, comme ce qui est observé en Polynésie Française.

Il serait intéressant de comprendre les caractéristiques de ce virus, pour lequel, à priori, l'immunité de population en 1989 était plus faible que pour les autres sérotypes. La seule étude phylogénétique du DENV-3 ayant circulé entre 1989 et 1996 a porté sur 28 virus, dont 13 seulement étaient issus du pacifique et avaient tous été isolés en Nouvelle-Calédonie ou en Polynésie Française (Chungue et al., 1993). Il serait intéressant de reprendre une étude sur ce sérotype, en élargissant l'origine géographique des virus étudiés.

Circulation endémique en Polynésie Française

La dengue présente une dynamique particulière en Polynésie Française, différente de celle des autres pays. En effet, tandis que la dengue se présente dans la plupart des pays sous une forme épidémique, avec des épidémies d'une ampleur relativement importante séparées par des périodes plus ou moins longues d'absence de circulation virale, elle se présente en Polynésie Française plutôt sous une forme endémique avec des épidémies déclarées quasiment chaque année, sans période nette d'arrêt de la circulation virale. Les raisons de la circulation à bas bruit observée dans ce pays par rapport aux autres îles du Pacifique peuvent inclure des relations accrues avec les pays endémiques via les voyageurs internationaux, une faible mise en œuvre des mesures de lutte, un renouvellement plus rapide de la population d'hôtes susceptibles, ou des particularités géographiques, climatiques, démographiques et socio-économiques favorisant les contacts hôte/vecteur.

Dans tous les cas, le fait que la Polynésie Française a été touchée par l'ensemble des vagues épidémiques d'une manière relativement précoce, ainsi que la présence d'un réseau de surveillance fiable font de la Polynésie Française un pays clé en termes de surveillance épidémiologique pour le déclenchement d'une alerte régionale lors de la survenue d'un nouveau sérotype. Cette remarque est d'autant plus importante que les vagues pandémiques ont tendance à émerger dans un nombre limité de pays avant d'exploser dans le Pacifique (Figure 3.11)

3.1.5.e. Absence de schémas régionaux

L'absence de schémas stables de propagation régionale d'un sérotype donné reste à confirmer (ou plutôt infirmer) de manière statistiquement plus robuste, par exemple avec une analyse en ondelettes (Cuong et al., 2013; Grenfell et al., 2001; Teurlai et al., 2012; Thai et al., 2010). En effet les études phylogénétiques ont montré l'existence d'une circulation régionale des virus de la dengue (A-Nuegoonpipat et al., 2004; Cao-Lormeau et al., 2011; Chungue et al., 1993; Dupont-Rouzeyrol et al., 2014; Effler et al., 2005; Li et al., 2010; Steel et al., 2010). Or nous pouvons nous attendre à ce que cette circulation se fasse au moyen de transports internationaux, qui sont supposés être stables d'une année à l'autre, hormis l'influence ponctuelle d'évènements culturels internationaux (jeux du Pacifique…). L'absence de schémas de propagation est également incohérente avec l'étude déjà réalisée sur les données épidémiologiques dans le Pacifique (Hales et al., 1999), bien que cette étude ait été réalisée sur des données ayant une meilleure résolution temporelle (mensuelle au lieu de annuelle).

Malgré une synchronisation régionale apparente de la dynamique interannuelle de la dengue, il est également raisonnable de penser qu'une partie de la variabilité interannuelle de la dengue dépend de facteurs intrinsèques à chacun des pays. L'émergence d'une épidémie à un sérotype donné dans un pays donné une année donnée serait alors une combinaison de facteurs régionaux de circulation d'un sérotype avec des facteurs locaux favorisant ou limitant l'émergence d'une épidémie. Cette hypothèse

expliquerait l'absence de schémas de propagation stables ainsi que l'absence de schéma spatial dans l'augmentation ou la diminution de la fréquence des épidémies.

Il apparaît donc nécessaire d'étudier la dynamique de la dengue à l'échelle de chacun des pays afin de mieux comprendre l'échelle régionale. C'est l'objet des deux chapitres suivants, qui focalisent à la Nouvelle-Calédonie l'étude de la dynamique spatio-temporelle de la dengue en relation avec divers facteurs environnementaux, socio-économiques ou entomologiques.

CHAPITRE 4 : ANALYSE DE LA DENGUE A L'ECHELLE NATIONALE OU TERRITORIALE

4.1. Analyse spatio-temporelle 169

4.1.1. Dynamique de la dengue en Nouvelle-Calédonie 169
4.1.1.a. Dynamique globale 169
4.1.1.b. Définition d'une année épidémique 171
4.1.2. Structure spatiale de la dengue en Nouvelle-Calédonie 171
4.1.3. Etude des schémas de propagation de la dengue par la méthode des ondelettes 173
4.1.4. Etude de la variabilité des schémas spatiaux de la dengue par la méthode des EOF : application à la Nouvelle-Calédonie 175
4.1.4.a. Analyse EOF des données épidémiologiques 175
4.1.4.b. Analyse EOF du climat 178
4.1.4.c. Discussion 181
Résumé des principaux résultats 181
Inversion des schémas spatiaux au cours des années 1996, 2004 et 2008 181
Perspectives : interprétation des schémas spatiaux observés 183

4.2. Analyse temporelle 183

4.2.1. Modélisation statistique non linéaire du risque épidémique annuel sur le territoire entier à partir de variables climatiques 183
4.2.1.a. Contexte et méthodologie 183
4.2.1.b. Résultats et conclusions 184
4.2.1.c. Article 1 185

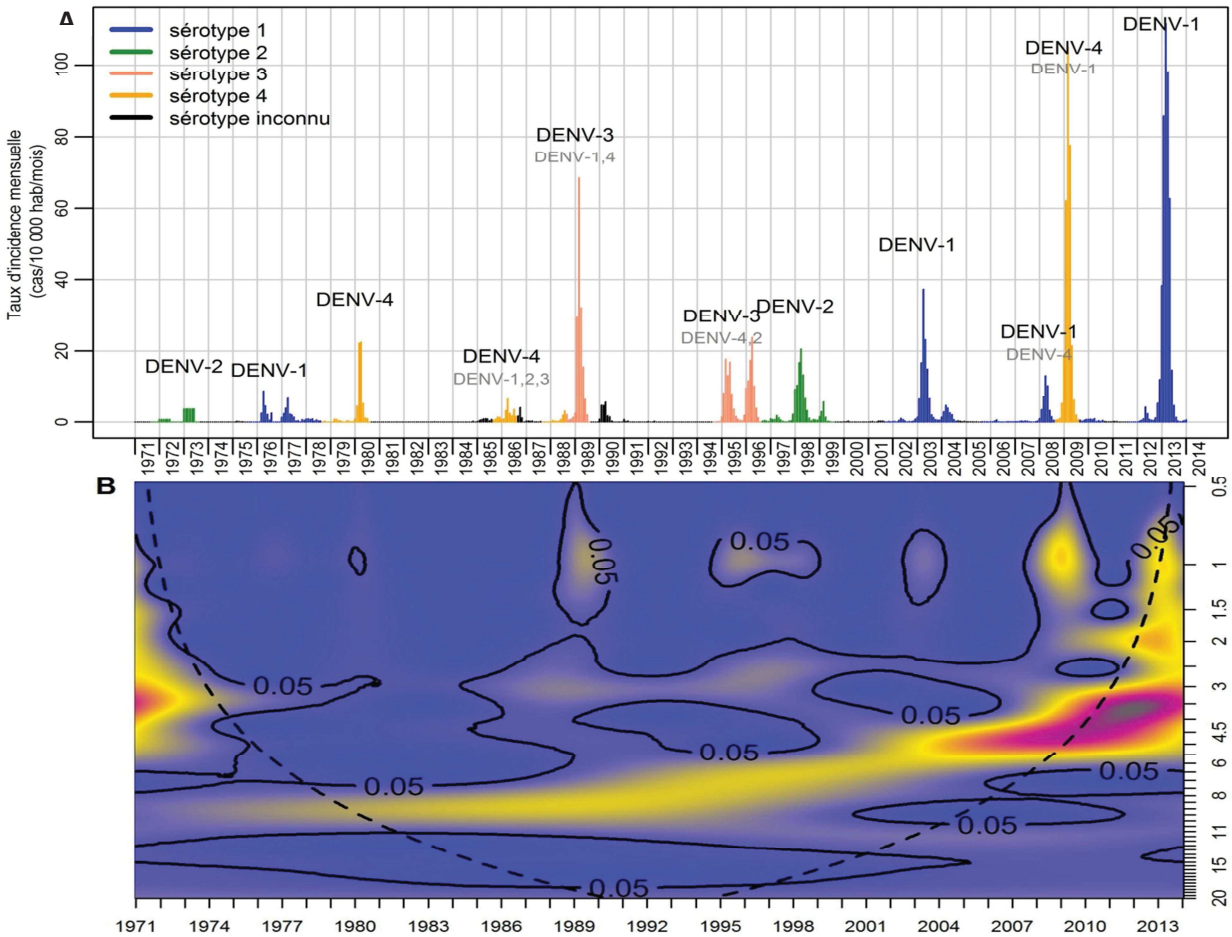
4.3. Analyse spatiale : 206

4.3.1. Modélisation non linéaire de l'intensité de la circulation virale et projections futures 206
4.3.1.a. Contexte et méthodologie 206
4.3.1.b. Résultats et discussion 206
Influence de la température moyenne sur la sévérité des épidémies 206
Influence des variables socio-économiques sur la sévérité des épidémies 207
4.3.1.c. Article 3 208

4.1. Analyse spatio-temporelle

4.1.1. Dynamique de la dengue en Nouvelle-Calédonie

4.1.1.a. Dynamique globale



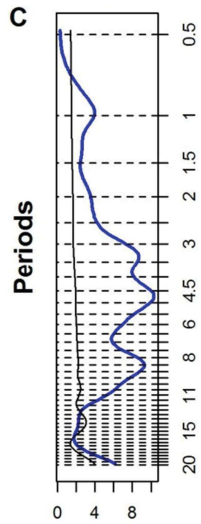


Figure 4.1 : Dynamique de la dengue en Nouvelle-Calédonie, 1971-2013. A. *Taux d'incidence mensuels sur le territoire. La couleur code pour le sérotype dominant. Le noir code pour les années au cours desquelles le sérotype est inconnu, ainsi que celles au cours desquelles plusieurs sérotypes ont été identifiés sans que l'information sur le sérotype dominant ne soit disponible. Avant 1974, nous n'avons eu accès qu'à des données d'incidence déjà agrégées annuellement. Nous avons alors réparti le nombre de cas sur les 6 premiers mois de l'année pour pouvoir les intégrer à cette courbe épidémique mensuelle.* ***B.*** *Spectre d'ondelette de A. L'abscisse représente le temps, de 1971 à 2013. L'ordonnée représente les périodes étudiées, en années (de 0.5 à 20 ans), la couleur code pour les coefficients d'ondelette du signal A. Ces coefficients représentent la puissance du signal pour un mois donné et une période donnée, allant du bleu pour une faible puissance, jusqu'au violet pour une forte puissance. Les traits en pointillé représentent les limites du cône d'influence, zone au-delà de laquelle les effets de bord empêchent d'interpréter les coefficients d'ondelette (partie basse du spectre d'ondelette). Les traits noirs délimitent la zone de significativité des coefficients d'ondelette. Les axes du temps (abscisses) de A et B sont superposables.* ***C.*** *Spectre global de A (assimilable à un spectre de Fourier). L'ordonnée représente la période (inverse de la fréquence) étudiée, et l'abscisse représente la puissance. La ligne bleue représente la moyenne temporelle des coefficients d'ondelette présentés en B (puissance moyenne du signal entre 1971 et 2013). La ligne noire représente le seuil de significativité de la puissance du signal.*

La Figure 4.1 montre les taux d'incidence mensuels de la dengue de 33 749 cas de dengue déclarés entre janvier 1971 et janvier 2014. En Nouvelle-Calédonie, la dengue se présente sous une forme épidémique (Figure 4.1.A). Comme c'est en général le cas pour les arboviroses, les courbes épidémiques sont caractérisées par une saisonnalité marquée, avec une concentration des cas entre janvier et mai, correspondant à l'été austral ainsi qu'à la saison chaude et pluvieuse en Nouvelle-Calédonie. Chaque année entre 10 et 30 cas de dengue sont importés à partir de pays d'Asie ou du Pacifique (Anne Pfannstiel, DASS-NC, communication personnelle). Certaines années, l'introduction de ces cas sur le territoire donne lieu à des épidémies d'ampleur majeure, débutant généralement entre novembre et décembre et finissant en juin. L'épidémie de 2013 représente une exception à cette dynamique, puisqu'il n'y a pas eu d'arrêt complet de la circulation virale pendant les mois hivernaux de 2012, conduisant à une flambée des cas dès le mois de septembre 2012 et à l'épidémie la plus importante jamais enregistrée sur le territoire en 2013. L'analyse des facteurs climatiques liés à la variabilité inter-annuelle du risque d'émergence de la dengue sur le territoire sera présentée dans l'article 1 (Descloux et al., 2012).

Les niveaux d'incidence sont difficiles à interpréter, du fait de changements majeurs du réseau de surveillance (test diagnostic utilisé, pourcentage de médecins déclarant) au cours de la période d'étude.

Les épidémies sont dues à l'émergence et à la diffusion sur le territoire d'un sérotype dominant. La co-circulation de deux sérotypes est rare. Comme à l'échelle du Pacifique, elle a été enregistrée lorsque le virus en cause était du type DENV-3, pendant les années 1980/1990. Le virus DENV-4 a également co-circulé avec le virus DENV-1 en 2008, a causé une épidémie majeure en 2009, et contrairement aux autres îles du Pacifique, a complètement disparu du territoire depuis. Les épidémies de 2012 et 2013 sont remarquables par la réémergence d'un sérotype (DENV-1) ayant déjà circulé pendant près de 10 ans. Une étude phylogénétique montre que le génotype en cause était un génotype différent de celui ayant circulé de 2003 à 2008 (Dupont-Rouzeyrol et al., 2014).

Les Figure 4.1.B et C montrent le spectre d'ondelette des taux d'incidence mensuels de la dengue sur le territoire. On observe des pics de puissance significatifs dans la bande annuelle (1980, 1990, 1995 à 1998, 2003, 2004 et 2009), correspondant à la dynamique saisonnière des épidémies. Le pic de puissance observé dans la bande périodique des 3 ans correspond à la circulation d'un même sérotype sur le territoire au cours de plusieurs années consécutives, avec des ralentissements de la circulation pendant la période hivernale. Enfin, le dernier pic de puissance, couvrant des périodes diminuant de la bande périodique des 8/10 ans à la bande périodique de 4/5 ans montre que la fréquence des épidémies augmente sur le territoire depuis les années 1990.

4.1.1.b. Définition d'une année épidémique

Comme à l'échelle du pacifique, il a été rapidement nécessaire de distinguer les années épidémiques des années non épidémiques. Descloux avait proposé dans sa thèse de définir une année comme étant épidémique si le taux d'incidence observé se trouvait dans le tercile supérieur des taux d'incidence observés entre 1971 et 2013. Nous avons conservé cette définition, qui est robuste aux changements éventuels du réseau de surveillance survenus au cours du temps. Les années épidémiques ainsi retenues en Nouvelle-Calédonie sont : 1973, 1976, 1977, 1980, 1986, 1989, 1995, 1996, 1998, 2003, 2004, 2008, 2009, 2013. Par rapport à la méthode du seuil utilisée au chapitre précédent, les années 1985, 1988, 1990, 1997 et 1999 ne sont plus considérées comme épidémiques. Une année épidémique va du 1er septembre de l'année *n-1* au 31 août de l'année *n*.

4.1.2. Structure spatiale de la dengue en Nouvelle-Calédonie

La structure spatiale de la dengue au cours d'une épidémie ne peut être étudiée que depuis 1995, année à partir de laquelle les cas de dengue ont été géo-référencés à la résolution de la commune. Pour la période 1995 à 2012, nous avons eu accès aux données relatives à l'âge des patients et avons pu ainsi réaliser la standardisation des taux d'incidence par rapport à l'âge.

La Figure 4.2 montre la structure spatiale de 24 272 cas de dengue déclarés dans les 33 communes composant la Nouvelle-Calédonie sur la période 1995-2012 ainsi qu'au cours de différentes années épidémiques. En moyenne (Figure 4.2.A), la dengue circule de manière hétérogène sur le territoire : les communes de la côte Est sont globalement plus touchées que celles de la côte Ouest. Nous pouvons remarquer les taux d'incidence particulièrement élevés dans les communes de Hienghène, Ouégoa et Pouébo, situées dans le Nord-Est du territoire. L'analyse des facteurs climatiques et socio-économiques liés à l'hétérogénéité spatiale moyenne sera présentée dans l'article 3 (Teurlai *et al.*, soumis à PLoS NTD).

Les taux d'incidence observés au cours des années épidémiques sont plus élevés lors de l'arrivée d'un nouveau sérotype sur le territoire (1995, 1998, 2003, 2009) que lors de sa réémergence les années suivantes (1996, 2004, 2008) (taux d'incidence moyens de 249 cas/10 000 habitants/an, dans le premier cas, vs 44 cas / 10 000 habitants / an dans le deuxième cas).

Les forts niveaux d'incidence dans les trois communes du Nord-Est du territoire par rapport aux autres communes du pays sont observés pour les années 1995, 1998, et dans une moindre mesure 2003 et 2004. Nous pouvons également noter les taux d'incidence particulièrement faibles tout au long des années dans les communes de Koumac (sauf en 2009) et Kaala-Gomen. Les épidémies dues à la réapparition du virus DENV-1 en 2004 et 2009 ont été de faible ampleur sur l'ensemble du territoire, et ont surtout touchés les îles qui avaient été épargnées jusque-là (Maré en 2004, Ile des Pins et Lifou en 2008).

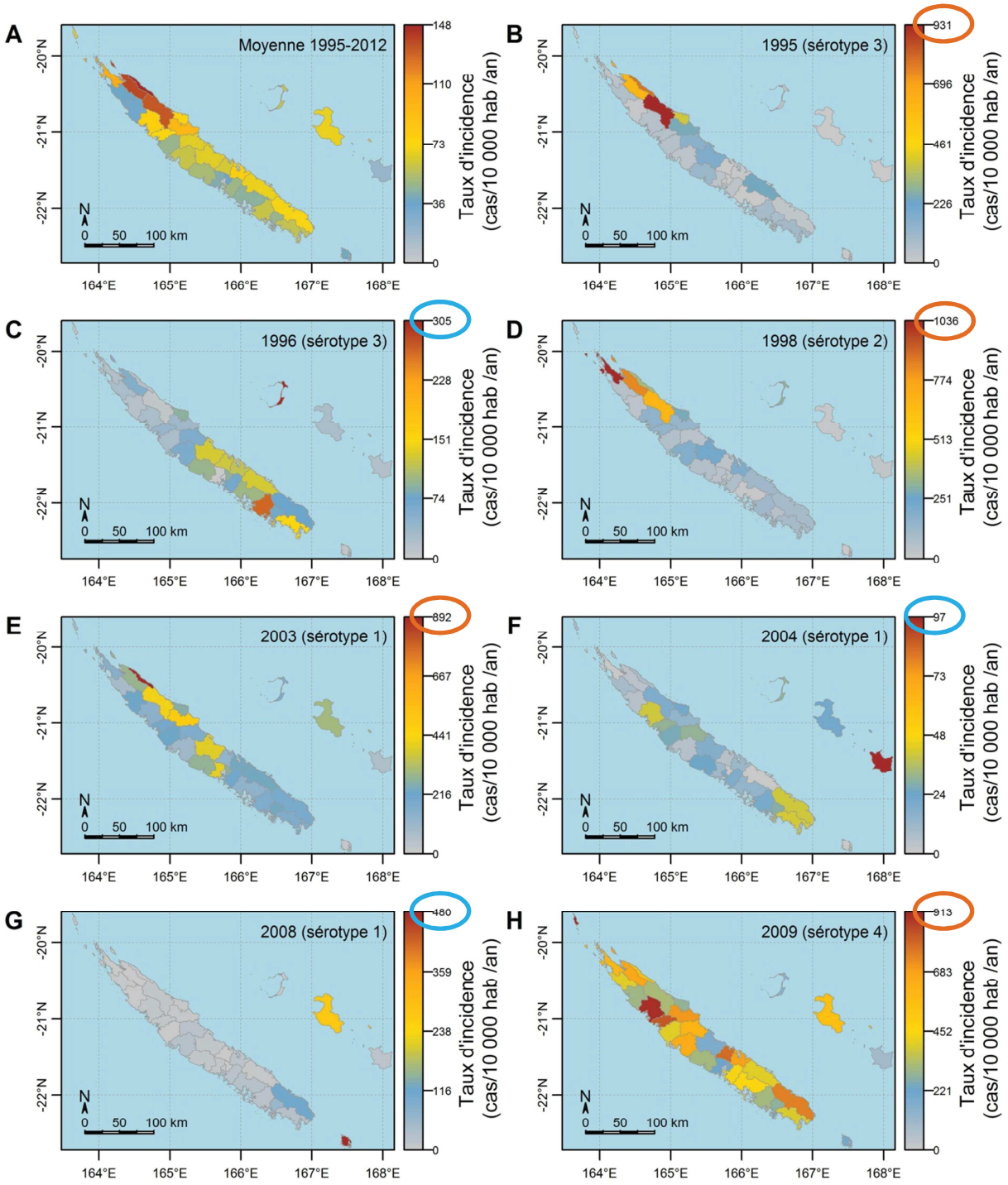


Figure 4.2 : Structure spatiale des épidémies de dengue en Nouvelle-Calédonie, 1995-2012. A. *Carte de la moyenne des taux d'incidence annuelle de la dengue entre 1995 et 2012 ($\overline{x}$ dans les équations des EOF).* ***B, C, D, E, F, G et H.*** *Taux d'incidence annuelle de la dengue au cours des années épidémiques (correspond à x_{1995}, x_{1996}, x_{1998}, x_{2003}, x_{2004}, x_{2008}, x_{2009} dans les équations des EOF). Les taux d'incidence annuels ont été calculés en agrégeant les cas de septembre de l'année n-1 à août de l'année n et en effectuant une standardisation par rapport à l'âge. Les cartes annuelles n'ont pas la même échelle de couleur, le maximum d'incidence sur chacune des cartes est encerclé en rouge pour les années de première apparition d'un sérotype, et en bleu lors de la ré-émergence d'un même sérotype.*

Les taux d'incidence annuels ont été calculés en agrégeant les cas du mois de septembre de l'année *n-1* au mois d'août de l'année *n* afin de tenir compte de la saisonnalité observée des épidémies et du climat en Nouvelle-Calédonie, avec des températures qui atteignent un maximum durant l'été austral.

4.1.3. Etude des schémas de propagation de la dengue par la méthode des ondelettes

L'objet de ce paragraphe est de déterminer s'il existe un schéma de propagation de la dengue se répétant d'une année à l'autre, et, le cas échéant, d'identifier les zones de démarrage des épidémies, ainsi que d'émettre des hypothèses sur les facteurs sous-tendant la diffusion spatio-temporelle de la maladie sur le territoire. L'approche méthodologique utilisée est la même que celle déjà utilisée pour étudier les schémas de propagation au Cambodge (Teurlai et al., 2012), et est présentée au chapitre 3.3.2 ainsi qu'à l'article 2, Annexe 6 et en Annexe 7.

La méthode est basée sur l'utilisation de l'analyse en ondelettes (Grenfell et al., 2001; Torrence and Compo, 1998), qui permet :

- 1) Le calcul du décalage temporel entre la dynamique de chacune des communes et celle d'une commune de référence, ici Nouméa. Cette méthode est une méthode d'analyse non stationnaire, c'est-à-dire qu'elle permet de calculer les décalages temporels pour chacun des pas de temps de la série temporelle, et donc d'obtenir un décalage moyen pour chacune des années ;
- 2) Sur la base des décalages temporels par rapport à Nouméa, il est alors possible, chaque année, de trier les communes par ordre chronologique de développement de l'épidémie annuelle dans chacune d'entre elles. Les résultats du tri sont présentés en Table 4.1.

Chaque épidémie n'a pas forcément affecté l'ensemble des communes. Ainsi, le nombre de cas peut être très faible dans une commune donnée lors de certaines années. Nous avons décidé de n'interpréter les phases et décalages temporels que lorsqu'au moins un foyer épidémique a été détecté durant l'année considérée dans une commune donnée. Un foyer épidémique au sein d'une commune a été défini comme l'apparition de plus 10 cas de dengue en moins de deux semaines. Ce seuil a été choisi car il correspond approximativement au seuil à partir duquel la DASS-NC déclenche des réunions afin d'examiner la situation épidémiologique (le côté approximatif provient du fait qu'il n'existe pas de seuil bien défini en Nouvelle-Calédonie, et le déclenchement des alertes se fait au cas par cas après examen de la situation épidémiologique).

Les tests de corrélation de Spearman, permettant de comparer l'ordre des communes au cours de deux années données, montre qu'il n'existe pas, en Nouvelle-Calédonie, de schéma de propagation particulier se répétant d'une année à l'autre. Les facteurs de propagation de l'épidémie semblent donc aléatoires, reflétant des mécanismes survenant certainement à des échelles spatiales faibles. Il est connu, par

exemple, que l'épidémie due au virus DENV-1 de 2003 est survenue suite à l'importation de plusieurs cas de dengue à partir du Vanuatu à l'occasion du voyage d'une classe de l'école de Poindimié.

Table 4.1 : Chronologie des épidémies annuelles de dengue dans les communes de Nouvelle-Calédonie. Les communes ont été classées par ordre croissant du décalage de phase entre l'épidémie de la commune considérée et l'épidémie de Nouméa, décalage de phase calculé par la méthode des ondelettes.

Rang de la commune	1995	1996	1998	2003	2004	2008	2009
1	Païta	Ouvéa	Ouvéa	Kaala-Gomen	Maré	Dumbéa	Ouvéa
2	Nouméa	Païta	Ouégoa	Poindimié	Mont-Dore	Nouméa	Poum
3	Mont-Dore	Mont-Dore	Poum	Hienghène	Nouméa	Mont-Dore	Poya
4	Bourail	Nouméa	Pouébo	Thio		Lifou	Koné
5	Hienghène	Dumbéa	Païta	Ouégoa		Ile des Pins	Voh
6	Ponérihouen		Koné	Touho		Dumbéa	Kouaoua
7	Thio		Poindimié	Mont-Dore		Nouméa	Kaala-Gomen
8	Pouébo		Poya	Ponérihouen		Mont-Dore	Hienghène
9	Touho		Canala	Nouméa		Lifou	Canala
10	Houailou		Mont-Dore	Dumbéa		Ile des Pins	Pouembout
11	Dumbéa		Nouméa	Maré			Ouégoa
12	Poindimié		Hienghène	Voh			Koumac
13	Ouégoa		Houailou	Païta			Yaté
14			Touho	Koné			Ponérihouen
15			Dumbéa	Bourail			Dumbéa
16				La Foa			Poindimié
17				Pouébo			Pouébo
18				Bouloupari			Païta
19				Canala			Nouméa
20				Houailou			Belep
21				Lifou			La Foa
22				Ouvéa			Bouloupari
23							Touho
24							Houailou
25							Bourail
26							Lifou
27							Mont-Dore
28							Ile des Pins
29							Thio

La précocité de l'épidémie à Ouvéa au cours des années 1996, 1998 et 2009 n'est pas due au fait que les épidémies débutent tôt, mais plutôt au fait que les foyers épidémiques annuels sont d'une durée limitée dans cette commune par rapport à l'épidémie de Nouméa. La dynamique épidémique globale calculée sur les six mois d'été apparaît donc en avance.

Bien que les premiers foyers de circulation autochtone soient toujours enregistrés très tôt dans la saison épidémique à Nouméa, le développement de l'épidémie n'y est pas plus précoce ou plus tardif que dans

d'autres villes beaucoup moins peuplées du territoire, contrairement au Cambodge (Teurlai et al., 2012) et au Vietnam (Thai et al., 2010) où les épidémies se développent de manière très tardive dans les grosses villes, et contrairement à la Thaïlande où les épidémies de dengue à Bangkok servent de « pace-maker » aux épidémies observées dans les autres provinces (Cummings et al., 2004).

4.1.4. Etude de la variabilité des schémas spatiaux de la dengue par la méthode des EOF : application à la Nouvelle-Calédonie

4.1.4.a. Analyse EOF des données épidémiologiques

La Figure 4.3 montre les cartes d'anomalies des taux d'incidence observés au cours des années épidémiques. Chacune des cartes présentées de B à F a été obtenue en retirant la carte d'incidence moyenne (A) à chacune des cartes de taux d'incidence présentées en Figure 4.2.B à F. Nous retrouvons de manière frappante le faible niveau d'incidence observé (anomalie d'incidence négative sur l'ensemble du territoire) lors des épidémies de 1996, 2004 et 2008).

En 1995 et 1998, le virus de la dengue a circulé de manière particulièrement intense dans les trois communes de Hienghène, Ouégoa et Pouébo, qui présentent des anomalies d'incidence positives par rapport à une moyenne déjà plus élevée que les autres communes du territoire. Par contre, ces trois communes semblent être protégées lors de la réémergence d'un même sérotype (anomalie d'incidence négative proche de la moyenne, et taux d'incidence presque égaux à 0 cas/10 000 habitants/an en 1996, 2004 et 2008). Ce phénomène de protection semble se produire également dans la commune de Poindimié (côte Est) entre 2003 et 2004.

Afin de quantifier la variabilité temporelle des cartes de taux d'incidence annuels autour de la carte moyenne, et d'identifier des schémas spatiaux récurrents d'une année à l'autre, nous avons effectué une analyse EOF sur le champ spatio-temporel $X(t, \mathbf{s}), t = 1995,\ 1996,\ \ldots, 2012,\ \mathbf{s} = 1, 2, \ldots\ 33,$ des taux d'incidence annuels de la dengue par commune (il y a 33 communes en tout sur le territoire).

La Figure 4.4 montre les résultats de l'analyse EOF. Le premier élément à interpréter est le niveau de variabilité des taux d'incidence annuels par rapport à la moyenne. En effet, si les anomalies d'incidence sont faibles par rapport à la moyenne, les cartes annuelles, bien que pouvant présenter des structures spatiales différentes sont toutes proches de la carte moyenne, et l'analyse EOF elle-même n'est pas nécessaire. Comme le premier axe représente près de 70 % de la variance totale des données, l'intensité des anomalies peut être approximée en effectuant le produit des valeurs de la première composante principale (CP1) avec les valeurs du premier EOF (EOF1) dans chacune des communes (*i.e.* le produit $a_{ik}u_{jk},\ k = 1,\ i = 1995,\ 1996, \ldots, 2012,\ j =\ 1, 2, \ldots 33$).

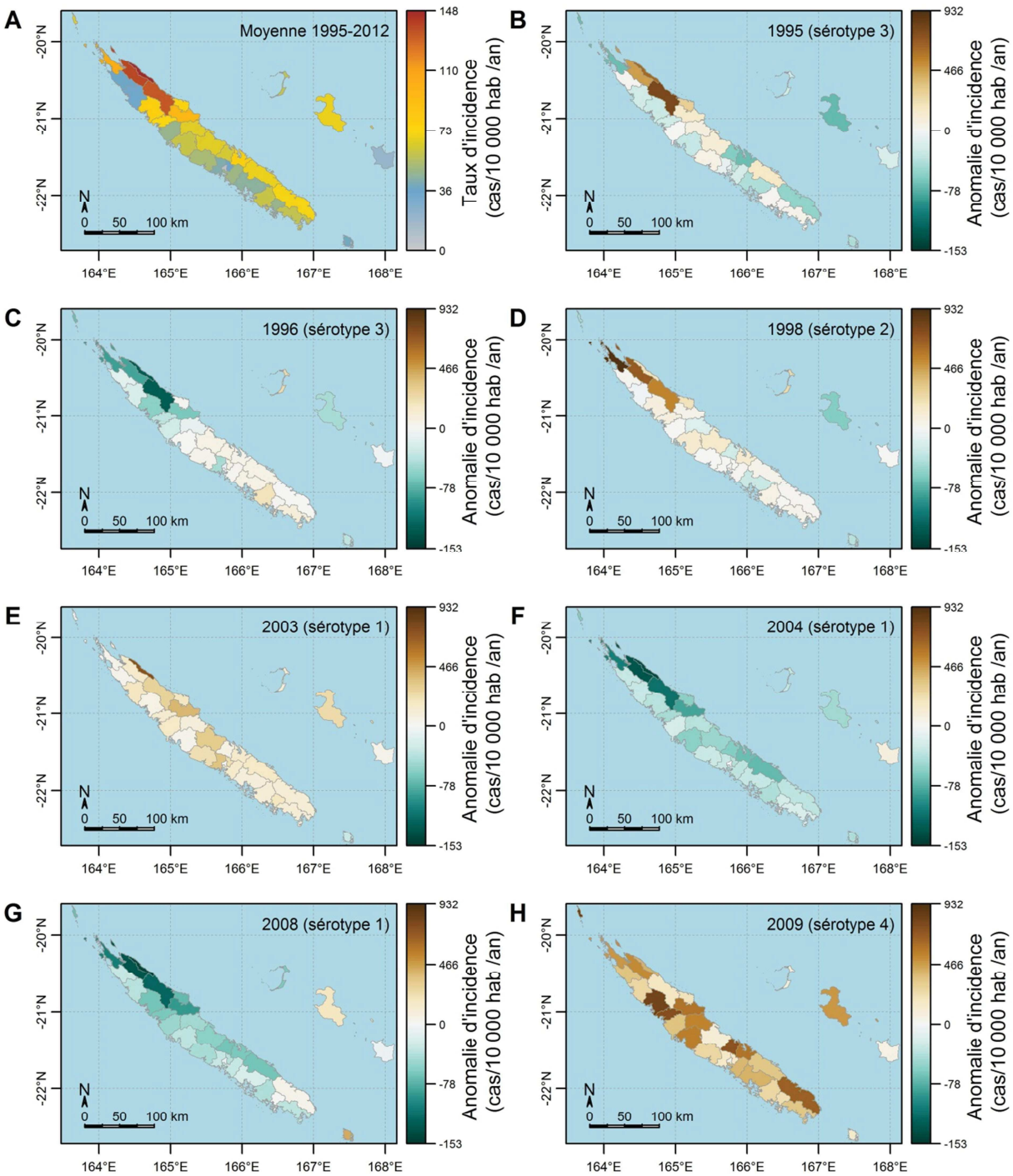


Figure 4.3 : Structure spatiale des anomalies des taux d'incidence annuels de la dengue en Nouvelle-Calédonie au cours des années épidémiques de la période 1995-2012. A. *Carte de la moyenne des taux d'incidence annuelle de la dengue entre 1995 et 2012 (*$\overline{\boldsymbol{x}} = \overline{x}_{\cdot j}, j = 1, \ldots, p$*, dans les équations).* ***B, C, D, E, F, G et H.*** *Anomalie du taux d'incidence annuelle de la dengue par rapport à la moyenne des taux d'incidence de chacune des communes sur la période d'étude, au cours des années épidémiques (correspond à* $\boldsymbol{x}_t^c = \boldsymbol{x}_t - \overline{x}_{\cdot j}$, $t \in \{1995, 1996, 1998, 2003, 2004, 2008, 2009\}$*).*

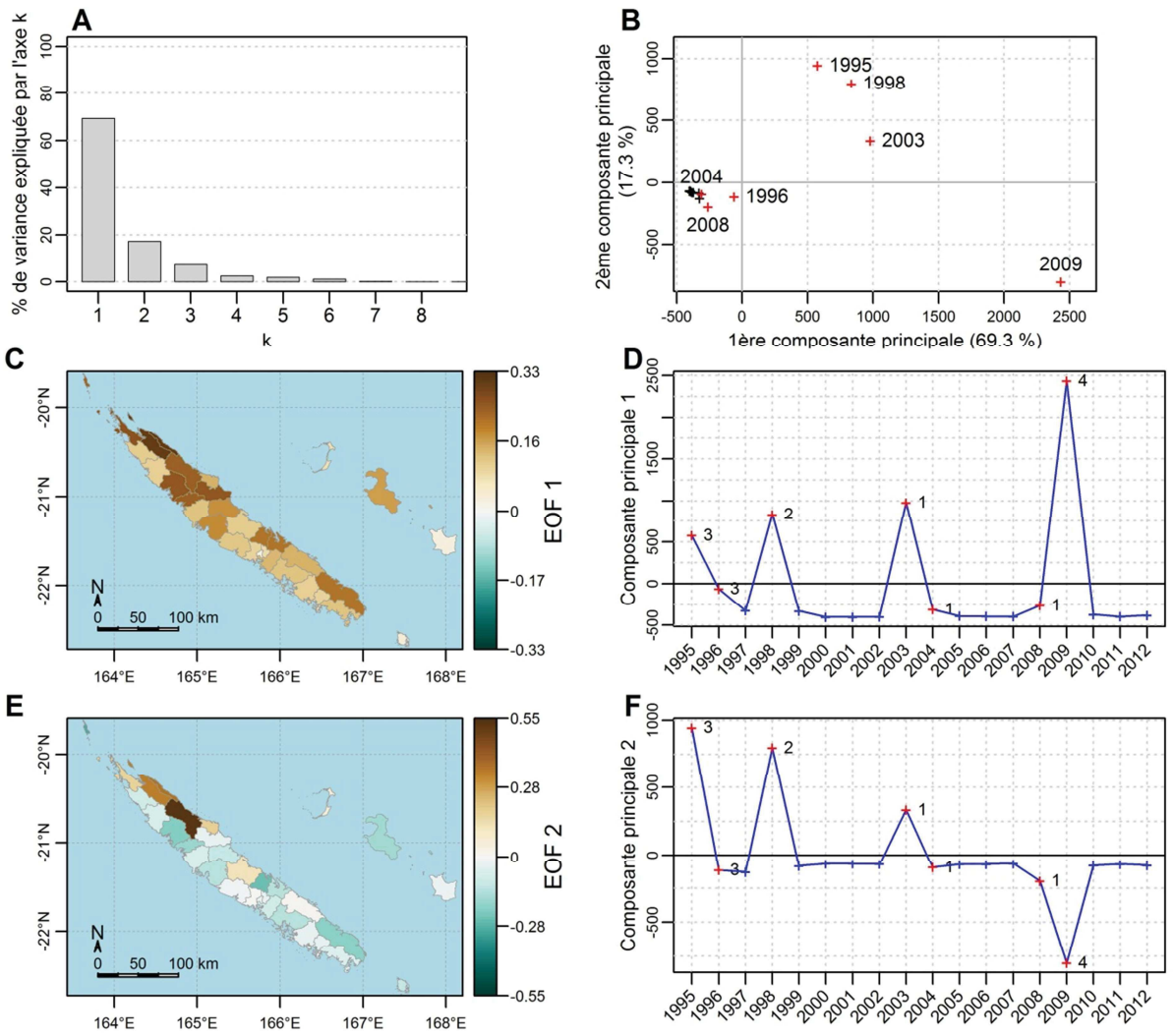


Figure 4.4 : Analyse EOF des taux d'incidence de la dengue en Nouvelle-Calédonie, période 1995-2012. ***A.*** *Valeurs propres (λ_k, $k = 1,\ldots,p$, dans les équations).* ***B.*** *Projection des cartes d'anomalies de taux d'incidence dans le plan formé par les deux premiers axes principaux, i.e les deux premières composantes principales ($\boldsymbol{a}_k = X^c\boldsymbol{u}_k = (a_{1k},\ldots,a_{nk}), k \in \{1,2\}$).* ***C à F.*** *Cartes des deux premiers EOF ($\boldsymbol{u}_k = (u_{1k},\ldots,u_{pk})$, k=1 en C et $k = 2$ en E) et composantes principales associées (D et F pour $k = 1$ et 2 respectivement). Sur les panneaux B, D et F, les croix rouges représentent les années épidémiques, et les croix noires ou bleues les années non épidémiques. Les numéros indiquent le sérotype dominant au cours de l'épidémie.*

Ce produit est beaucoup plus élevé que les taux d'incidence moyens (pour $k = 1$, l'anomalie d'incidence est d'environ 700 * 0.33 = 230 cas/10 000 habitants par an pour i = 1995, 1998 et 2003, et environ 2500*0.33 = 825 cas / 10 000 habitants /an pour i = 2009, à comparer à une moyenne comprise entre 0 et 150 cas/10 000 habitants/an, Figure 4.4. C, et D, et Figure 4.3.A). La variance totale du nuage du point est donc élevée par rapport à la moyenne : il existe une forte variabilité interannuelle de la structure spatiale des taux d'incidence de la dengue.

Le deuxième élément à interpréter est le spectre des valeurs propres. Les deux premiers axes représentent 86,6 % de la variance totale. Soit il existe des structures spatiales fortes dans les données d'incidence, se répétant d'une année à l'autre, soit le schéma observé une année donnée s'écarte particulièrement de la moyenne.

L'examen de la première composante principale montre que le premier axe de variabilité est fortement déterminé par la structure spatiale de 2009 (Figure 4.4.B et D). Ceci paraît normal, puisque 2009 est l'année où l'épidémie a été de plus grande ampleur sur la période d'étude. C'est donc également celle qui présente la plus grande variance. Les valeurs élevées de la première composante principale ($\mathbf{a}_1$) pour les années 1995, 1998 et 2003 indiquent que la première structure spatiale d'anomalie (1er EOF $\mathbf{u}_1$, Figure 4.4.C) observée en 2009 se retrouve également en 1995, 1998 et 2003 (Figure 4.4. B et C). Ce résultat indique que cette structure spatiale est une structure cohérente tout au long des années épidémiques. Les valeurs de $\mathbf{u}_1$ sont de même signe dans l'ensemble des communes (Figure 4.4.C) : la dynamique des taux d'incidence présentée par la première composante principale est donc homogène sur l'ensemble du territoire.

L'examen du signe du produit $a_{i1}u_{j1},\ i = 1995,\ldots,2012; j = 1,\ldots,33$ du premier EOF par la première composante principale permet de retrouver le résultat observé sur les cartes d'anomalies : il est positif pour les années $i = 1995$, 1998, 2003 et 2009, ce qui indique des taux d'incidence largement supérieurs à la moyenne de chacune des communes durant ces années-là. Il est négatif pour les années $i = 1996$, 2003 et 2004, indiquant des taux d'incidence inférieurs à la moyenne, et ce sur l'ensemble du territoire. Le fait que le premier axe explique près de 70% de la variance permet de dire, avant d'interpréter les autres axes, que ce résultat, directement visible sur les cartes d'anomalies pour les trois communes du Nord-Est, est également valable sur les autres communes du territoire, du moins pour celle qui présentent une valeur absolue élevée de u_{j1}.

L'examen de la deuxième composante principale montre que le deuxième axe de variabilité est déterminé par la présence d'une structure spatiale inverse au 2ème EOF en 2009 (composante principale $a_{ik},\ k = 2,$ fortement négative pour $i = 2009$). C'est-à-dire que l'anomalie d'incidence dans les trois communes du Nord-Est est beaucoup moins élevée en 2009 que pour les autres années (pour savoir si ces trois communes ont effectivement un taux d'incidence plus élevé ou plus faible que la moyenne en 2009, il faut également regarder le premier axe, ou bien regarder directement la carte d'anomalie de 2009).

4.1.4.b. Analyse EOF du climat

L'un des facteurs pouvant potentiellement expliquer la dynamique interannuelle des taux d'incidence observés dans les 33 communes est le climat. En effet, la dynamique climatique interannuelle devrait être homogène sur l'ensemble du territoire, mais avec sûrement des impacts différenciés spatialement. Nous avons donc réalisé une analyse EOF simple sur les champs spatio-temporels de température moyenne et de précipitations, utilisant les séries temporelles de 24 stations météorologiques pour lesquelles nous avions des séries complètes sur l'ensemble de la période d'étude (Figure 4.5 et Figure 4.6).

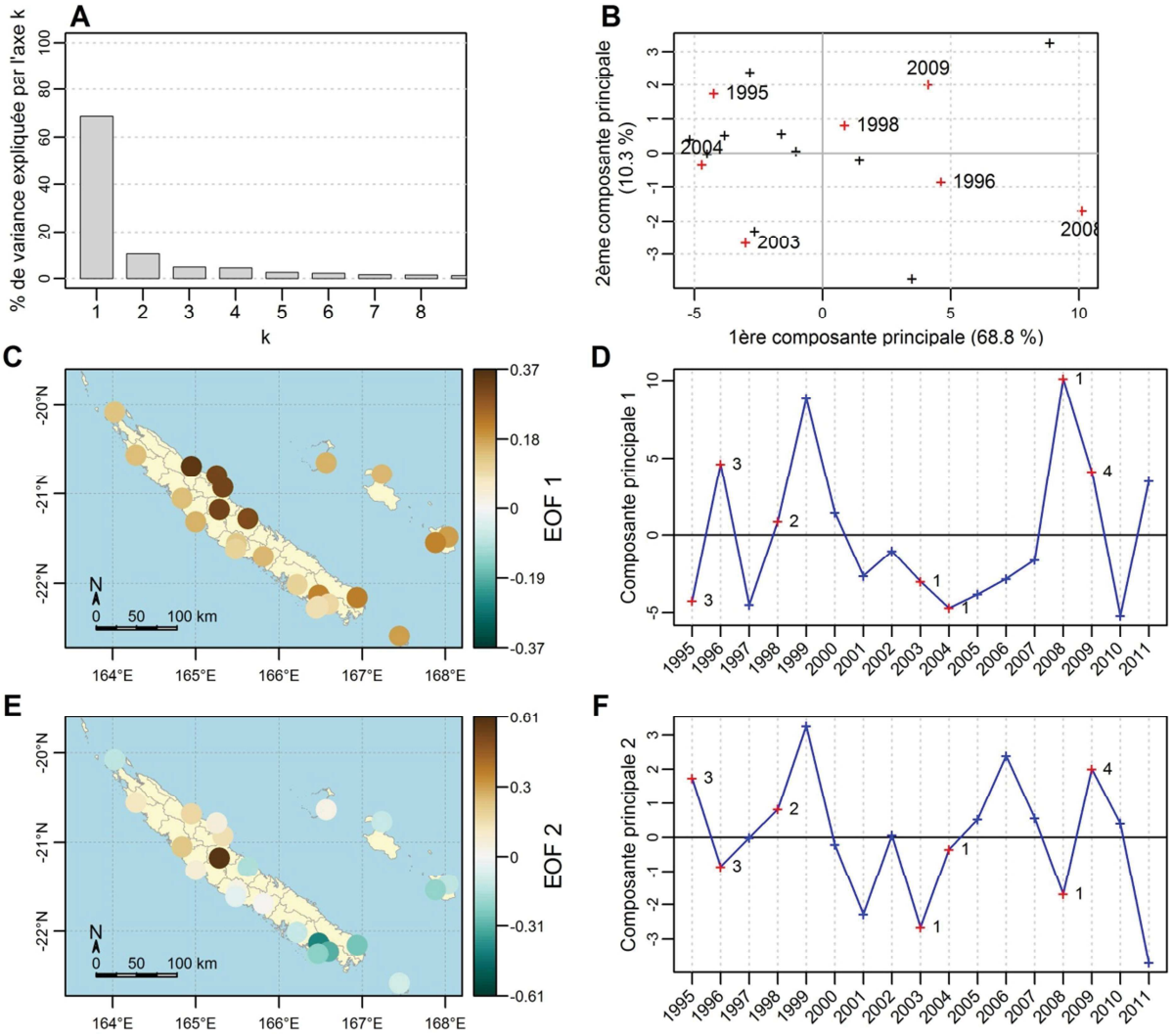


Figure 4.5 : Analyse EOF des précipitations quotidiennes moyennes sur 24 stations météorologiques en Nouvelle-Calédonie, période 1995-2011. A. *Valeurs propres.* ***B.*** *Scatter plot des deux premières composantes principales.* ***C à F.*** *Cartes des deux premiers EOF (EOF 1 en C et EOF2 en E) et composantes principales associées (D et F respectivement). Sur les panneaux B, D et F, les croix rouges représentent les années épidémiques, et les croix noires ou bleues les années non épidémiques. Les numéros indiquent le sérotype dominant au cours de l'épidémie.*

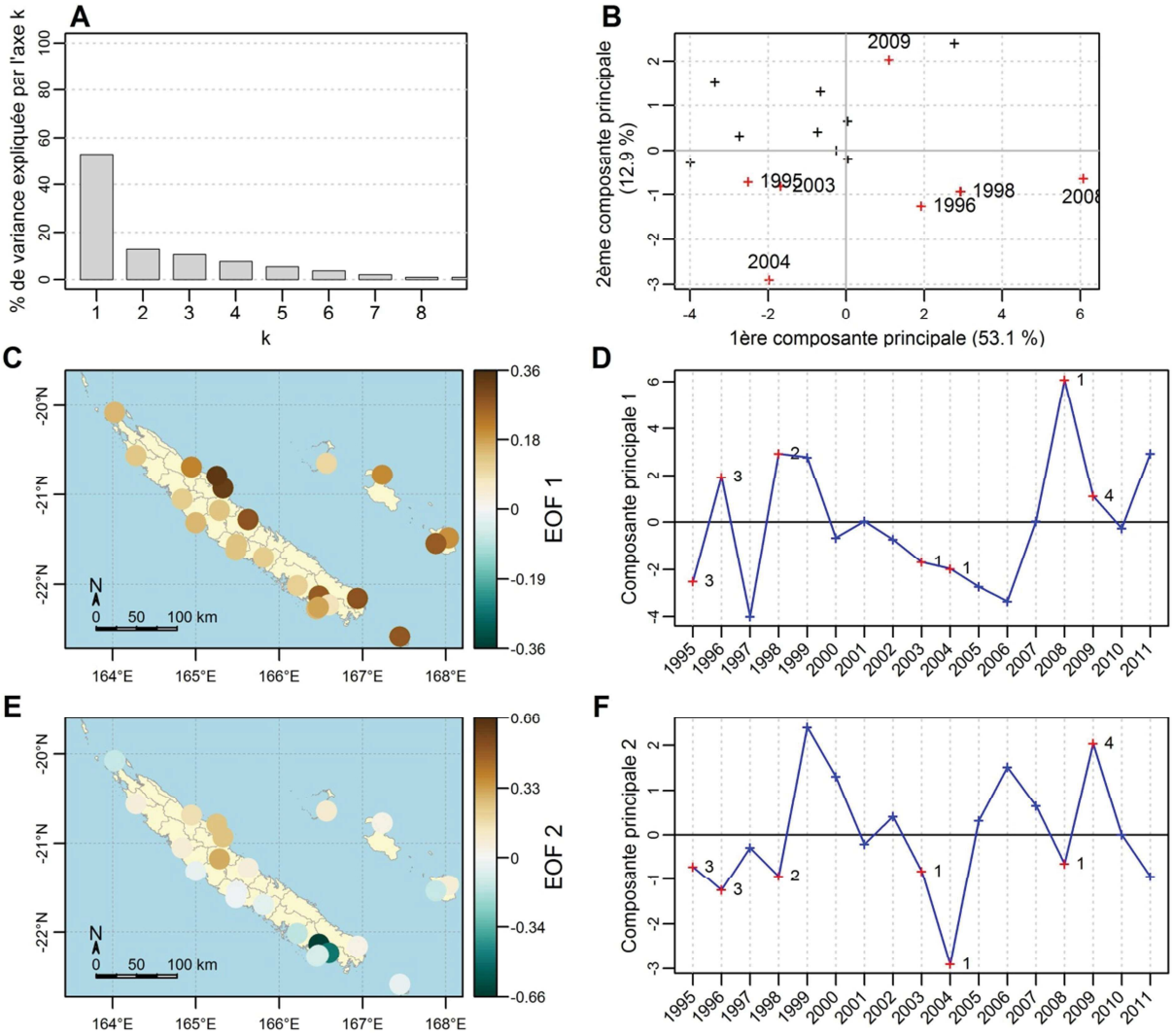


Figure 4.6 : Analyse EOF des températures moyennes sur 24 stations météorologiques en Nouvelle-Calédonie, période 1995-2011. A. *Valeurs propres.* ***B.*** *Scatter plot des deux premières composantes principales.* ***C à F.*** *Cartes des deux premiers EOF (EOF 1 en C et EOF2 en E) et composantes principales associées (D et F respectivement). Sur les panneaux B, D et F, les croix rouges représentent les années épidémiques, et les croix noires ou bleues les années non épidémiques. Les numéros indiquent le sérotype dominant au cours de l'épidémie.*

L'étude des valeurs propres montre qu'il existe une forte cohérence spatiale de la dynamique des deux variables étudiées (79,2 % de variance et 66 % de variance expliquée par les deux premiers axes pour les précipitations et les températures moyennes respectivement). Le fait que le premier EOF de chacune des variables climatiques a le même signe sur l'ensemble du territoire montre que la dynamique interannuelle est homogène sur l'ensemble des communes : comme nous nous y attendions, les années plus chaudes ou plus pluvieuses sont plus chaudes et plus pluvieuses partout. Par contre, la dynamique interannuelle principale observée est différente de celle de la dengue. Les coefficients de corrélation de Spearman entre la première composante de l'analyse des données épidémiologiques et la première composante de l'analyse des données climatiques est de 0,03 (*p*-valeur de 0.91) pour les précipitations, et de 0.09 (*p*-valeur de 0.72) pour les températures moyennes. Il ne semble donc y avoir aucune

corrélation entre les anomalies interannuelles principales de taux d'incidence et les anomalies climatiques interanuelles principales. De plus, les zones dans lesquelles les variations interannuelles du climat sont les plus intenses (Sud Est de l'île, et communes de Touho et Poindimié) ne correspondent pas aux zones de variations interannuelles des taux d'incidence (Hienghène, Ouégoa, Pouébo, Poum, Poindimié, Koné, Voh, Yaté et Canala). Il serait intéressant d'étudier d'autres variables (température minimale, température maximale, nombre de jours pour lesquels une variable dépasse un certain seuil).

4.1.4.c. Discussion

Résumé des principaux résultats

En résumé, en Nouvelle-Calédonie, la circulation virale au cours des années épidémiques de 1995 à 2012 est spatialement hétérogène (Figure 4.2.A). En moyenne, les communes de la côte Est sont plus touchées que celle de la côte Ouest. La dynamique interannuelle des taux d'incidence est homogène sur le territoire : globalement, l'ensemble des communes ont été plus touchées par la dengue en 1995, 1998, 2003 et 2009 qu'en 1996, 2004, et 2008. La dynamique interannuelle résulte de deux schémas spatiaux de variations :

- un premier schéma spatial (EOF1, 69,3 % de la variance, voir Figure 4.4.C), soulignant que la différence d'intensité de la circulation virale entre les années épidémiques 1995, 1998, 2003 et 2009 (circulation virale intense) d'une part et les années 1996, 2004 et 2008 d'autre part (circulation virale faible) est particulièrement importante dans les communes du nord (Hienghène, Ouégoa, Pouébo, Poum, Poindimié, Koné, et Voh) et certaines communes de la côte Est (Yaté, Canala). Ce premier schéma spatial est fortement influencé par les taux d'incidence observés en 2009 ;
- Un deuxième schéma (EOF2, 17,3 % de la variance, voir Figure 4.4.E) soulignant l'intensité accrue de la circulation virale dans les communes du Nord-Est lors des épidémies de 1995, 1998 et 2003 (mais pas durant celle de 2009) par rapport à la circulation moyenne dans ces communes.

Les précipitations et les températures ne semblent pas liées aux schémas épidémiologiques observés.

Inversion des schémas spatiaux au cours des années 1996, 2004 et 2008

L'étude du premier EOF montre que les communes les plus touchées par la dengue en 1995, 1998, 2003 et 2009 sont également celles les moins touchées en 1996, 2004 et 2008. L'hypothèse majeure qui pourrait expliquer le schéma observé est celle d'une immunité de population sérotype spécifique, qui limiterait la circulation virale lors de la ré-introduction d'un même sérotype dans une commune donnée. En effet, les années 1996, 2004 et 2008 sont dues à un sérotype ayant déjà circulé entre une et cinq années auparavant. Ce résultat n'est pas robuste sur le plan statistique puisque basé uniquement sur l'observation de trois changements de sérotype, et de trois ré-introductions d'un même sérotype.

Une grande inconnue dans l'étude de la dengue en Nouvelle-Calédonie est la durée nécessaire au renouvellement d'une population susceptible à la suite d'une épidémie. Cette durée dépend :

- De l'intensité de la circulation virale au cours de l'épidémie immunisant la population ;
- De la proportion de cas non détectés par le réseau de surveillance (malades ne se rendant pas dans un centre de soins, ou cas asymptomatiques) ;
- Du taux de croissance démographique ;
- Des mouvements migratoires entrants et sortants du territoire.

Si l'hypothèse d'une protection immunitaire est juste, et que l'on se base sur l'observation du taux d'incidence très faible au cours de l'épidémie de 2008 par rapport à celle de 2003 (DENV-1), et sur le fait que l'épidémie de 2008 n'a touché que les îles qui avaient été épargnées auparavant, il semble que l'effet protecteur de l'immunisation de la population dure au moins 4 ans. Par contre, le sérotype 1 a également été responsable en 2013 de l'épidémie de plus forte ampleur jamais enregistrée sur le territoire (Figure 4.1.A), ce qui prouve qu'avec les niveaux d'incidence observés en Nouvelle-Calédonie, dix années suffisent à ce qu'une population importante de personnes non immunes soit reconstruite et permette de nouveau le développement d'une épidémie majeure. Il serait intéressant de construire un modèle mathématique déterministe de type SIR différenciant au moins deux sérotypes, afin de tester cette hypothèse immunitaire. La compréhension du rôle joué par les facteurs immunitaires est d'autant plus importante qu'un vaccin devrait bientôt être disponible. L'effet de l'immunisation de la population sur la dynamique de la dengue doit donc être compris.

Si cette hypothèse de protection immunitaire de la population est vraie, les conséquences opérationnelles sont immédiates : lors de la ré-introduction d'un sérotype sur le territoire, les autorités sanitaires peuvent s'attendre à un nombre de cas plus élevé dans les communes ayant été peu touchées les années précédentes.

Les autres facteurs potentiellement associés aux variations interannuelles de l'intensité de la circulation virale sur l'ensemble du territoire incluent :

- Des facteurs climatiques, selon un mécanisme plus complexe que l'influence directe de la température annuelle moyenne et des précipitations annuelles moyennes sur le cycle de vie du moustique vecteur (e.g. variations climatiques infra-saisonnières comme les variations mensuelles du nombre de jours où une variable persiste au dessus d'un certain seuil) ;
- Des facteurs liés au virus : même si cela semble peu probable, des mutations génétiques conduisant à des modifications de la compétence vectorielle d'*Aedes aegypti* pour un sérotype donné au cours du temps pourraient expliquer l'atténuation de la circulation virale ;

- Des facteurs liés au vecteur, comme la lutte anti-vectorielle, qui modifie les densités de populations vectorielles au cours du temps. Cependant, si la lutte anti-vectorielle peut être hétérogène spatialement sur le territoire, elle est censée être relativement stable dans le temps, sauf au cours de certaines années particulières de changement drastique de stratégie de lutte, comme en 2003 ;
- Des facteurs liés à l'hôte, autres que les phénomènes immunitaires.

Perspectives : interprétation des schémas spatiaux observés

Sans autre donnée, ni observation de terrain, il est difficile d'interpréter les schémas spatiaux observés en termes épidémiologiques. En effet, l'hétérogénéité spatiale de la circulation virale au cours des années épidémiques, ainsi que les deux schémas d'anomalie principaux (constitués du 1er EOF et de la circulation intensément élevée ou intensément faible dans le Nord du territoire) peuvent être dus à une combinaison de multiples facteurs.

L'étude des facteurs expliquant l'hétérogénéité de la circulation virale moyenne sur la période 1995-2012 fait l'objet de l'article 3 présenté page 208.

4.2. Analyse temporelle

4.2.1. Modélisation statistique non linéaire du risque épidémique annuel sur le territoire entier à partir de variables climatiques

4.2.1.a. Contexte et méthodologie

Nous avons vu qu'en Nouvelle-Calédonie, la dengue présente une dynamique épidémique, avec des épidémies survenant tous les 4 à 5 ans, séparées par de longues périodes au cours desquelles seuls des cas importés sont détectés. Ces épidémies surviennent souvent une ou deux années après l'émergence d'épidémies dues au même sérotype dans d'autres îles du Pacifique (Figure 3.9, page 154). Or les autorités sanitaires calédoniennes estiment qu'il y a entre dix et trente cas de dengue importés chaque année sur le territoire. Quels sont alors les facteurs associés à l'émergence ou à l'absence d'émergence d'une épidémie sur le territoire ?

Parmi l'ensemble des facteurs pouvant influencer le cycle de transmission de la dengue, seules des variables climatiques et entomologiques étaient disponibles sous la forme de séries temporelles longues (2000 - 2009 pour les données entomologiques, et 1971 - 2010 pour les données climatiques).

Dans l'analyse qui suit, nous avons étudié les relations saisonnières entre les variables épidémiologiques, climatiques et entomologiques sur Nouméa pour la période 2000 - 2009. Nous avons ensuite utilisé des modèles SVM (Cortes and Vapnik, 1995; Vapnik, 1998) afin d'étudier l'association statistique entre les

facteurs climatiques et la variabilité interannuelle des épidémies de dengue sur l'ensemble du territoire de 1971 à 2009, toutes communes confondues. Pour cette étude interannuelle, nous avons transformé les données mensuelles d'incidence de la dengue (Figure 4.1.A) en une série temporelle annuelle binaire (0 pour l'absence d'épidémie et 1 pour la présence d'une épidémie). Les variables climatiques étudiées incluent la température, les précipitations et l'humidité relative quotidiennes sur Noumea, des variables construites à partir de ces trois variables brutes (variables mensuelles, trimestrielles, annuelles, et indicateurs de persistance de condition climatique sur des durées plus ou moins longues) ainsi que des indicateurs mesurant l'intensité annuelle du phénomène climatique ENSO (SOI et MEI). Les modèles SVM permettent ici de prédire la probabilité d'occurrence d'une épidémie l'année *n* à partir des valeurs observées des variables climatiques. Deux types de modèles SVM ont été construits :

- Des modèles explicatifs utilisant comme variables explicatives les variables climatiques de l'année *n* ;
- Des modèles prédictifs utilisant les variables climatiques de l'année *n-1*, dans le but de construire un système d'alerte précoce de la survenue des épidémies.

Afin d'identifier les variables climatiques ayant la plus forte association statistique avec le risque inter-annuel d'occurrence d'une épidémie, nous avons construit l'ensemble des modèles utilisant toutes les combinaisons possibles d'une, deux ou trois variables climatiques, et mesuré la qualité de chacun des modèles à l'aide d'un indicateur, l'AUC.

4.2.1.b. Résultats et conclusions

Les analyses ont montré que :

- La dynamique de la dengue est saisonnière. Le cycle épidémique annuel saisonnier survient un à deux mois après le pic de température, et est en phase avec le pic de précipitations et de densité de vecteurs ;
- Le risque épidémique annuel de la dengue en Nouvelle-Calédonie est fortement influencé par les conditions climatiques locales (température, humidité relative, précipitations) durant l'été austral ;
- Le phénomène ENSO ne semble pas avoir d'impact sur le risque de survenue d'une épidémie de dengue en Nouvelle-Calédonie ;
- Le lien entre les conditions climatiques et la probabilité d'occurrence est non linéaire, faisant apparaître des effets seuils des variables climatiques, difficilement détectables par l'utilisation de modèles statistiques linéaires paramétriques classiques tels que les GLM ;
- Les meilleures variables permettant d'expliquer la survenue ou non d'une épidémie sont le nombre de jours où la température dépasse 32 °C en janvier/février/mars (été austral) et le nombre de jours où l'humidité relative dépasse 95 % en janvier. Le risque d'émergence d'une épidémie est accru si

ces variables dépassent 11 jours et 10 jours respectivement, montrant l'importance de la persistance de conditions climatiques favorables dans le risque d'émergence par rapport aux variables météorologiques classiquement étudiées et ne reflétant que des valeurs moyennes ;

- Le modèle explicatif présente une AUC de 0,8, ce qui démontre l'impact majeur du climat sur le risque épidémique annuel en Nouvelle-Calédonie. Pour un seuil de probabilité à 0,65 pour différencier les années épidémiques des années non épidémiques, ce modèle a une sensibilité (pourcentage d'années épidémiques détectées) de 75 % et une spécificité (pourcentage d'années non épidémiques détectées) de 95 % ;

- Le modèle prédictif a une AUC de 0,69. Pour un seuil de probabilité à 0,65 pour différencier les années épidémiques des années non épidémiques sa valeur prédictive positive est de 79 %, et sa valeur prédictive négative est de 65 %, *i.e.* lorsque le modèle donne une probabilité de survenue d'une épidémie supérieur à 0,65, il y 4 chances sur 5 qu'une épidémie survienne. Lorsqu'il donne une probabilité de survenue de l'épidémie inférieure à 0,65, il y a 3 chances sur 5 qu'il n'y ait pas d'épidémie. Ce modèle prédictif est maintenant utilisé de manière opérationnelle par les autorités sanitaires.

L'influence des conditions climatiques locales peut s'expliquer par l'influence du climat sur les composante virales (durée d'incubation extrinsèque, vitesse de réplication du virus), vectorielles (influence du climat sur le cycle de vie d'*Aedes aegypti*) et humaines (comportements influençant les contacts hôtes/vecteurs) de la dynamique de la dengue.

L'absence d'influence du phénomène ENSO sur le risque épidémique peut s'expliquer par la position relativement éloignée de la Nouvelle-Calédonie par rapport au centre majeur d'influence d'El Niño, et également par l'influence d'autres facteurs de variation climatique comme l'oscillation de Madden-Julian pouvant brouiller le signal ENSO.

Le modèle prédictif est maintenant utilisé de manière opérationnelle par les autorités sanitaires de Nouvelle-Calédonie (DASS-NC) comme l'une des composantes d'un indicateur annuel du risque épidémique de dengue.

4.2.1.c. Article 1

L'analyse dont nous venons de donner un résumé a été menée par Descloux durant sa thèse et fait l'objet d'un article publié dans PLoS NTD, présenté ci-après. Ma contribution a porté sur l'analyse des données entomologiques climatiques et épidémiologiques sur Nouméa de 2000 à 2009, sur l'interprétation de l'ensemble des résultats ainsi que sur la relecture critique du manuscrit.

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Climate-Based Models for Understanding and Forecasting Dengue Epidemics

Elodie Descloux[1,2]*, Morgan Mangeas[3], Christophe Eugène Menkes[4], Matthieu Lengaigne[5], Anne Leroy[6], Temaui Tehei[6], Laurent Guillaumot[7], Magali Teurlai[3], Ann-Claire Gourinat[8], Justus Benzler[9], Anne Pfannstiel[10], Jean-Paul Grangeon[10], Nicolas Degallier[5], Xavier De Lamballerie[1]

1 UMR190, Emergence of Viral Pathologies, Institute of Research for the Development, Aix-Marseille University, Marseille, France, **2** Department of Internal Medicine, Territorial Hospital Centre of New Caledonia, Noumea, New Caledonia, **3** UMR ESPACE-DEV 228, Institute of Research for the Development, Noumea, New Caledonia, **4** UMR 7159/UR 182, LOCEAN, Institute of Research for the Development, Noumea, New Caledonia, **5** UMR 7159/UR 182, LOCEAN, Institute of Research for the Development, University Paris VI, Paris, France, **6** Météo-France, Noumea, New Caledonia, **7** Laboratory of Medical Entomology, Pasteur Institute, Noumea, New Caledonia, **8** Laboratory of Virology, Pasteur Institute, Noumea, New Caledonia, **9** Public Health Surveillance and Communicable Disease Control Section, Public Health Division, Secretariat of the Pacific Community, Noumea, New Caledonia, **10** Health Department, Direction of Health and Social Affairs of New Caledonia, Noumea, New Caledonia

Abstract

Background: Dengue dynamics are driven by complex interactions between human-hosts, mosquito-vectors and viruses that are influenced by environmental and climatic factors. The objectives of this study were to analyze and model the relationships between climate, *Aedes aegypti* vectors and dengue outbreaks in Noumea (New Caledonia), and to provide an early warning system.

Methodology/Principal Findings: Epidemiological and meteorological data were analyzed from 1971 to 2010 in Noumea. Entomological surveillance indices were available from March 2000 to December 2009. During epidemic years, the distribution of dengue cases was highly seasonal. The epidemic peak (March–April) lagged the warmest temperature by 1–2 months and was in phase with maximum precipitations, relative humidity and entomological indices. Significant inter-annual correlations were observed between the risk of outbreak and summertime temperature, precipitations or relative humidity but not ENSO. Climate-based multivariate non-linear models were developed to estimate the yearly risk of dengue outbreak in Noumea. The best explicative meteorological variables were the number of days with maximal temperature exceeding 32°C during January–February–March and the number of days with maximal relative humidity exceeding 95% during January. The best predictive variables were the maximal temperature in December and maximal relative humidity during October–November–December of the previous year. For a probability of dengue outbreak above 65% in leave-one-out cross validation, the explicative model predicted 94% of the epidemic years and 79% of the non epidemic years, and the predictive model 79% and 65%, respectively.

Conclusions/Significance: The epidemic dynamics of dengue in Noumea were essentially driven by climate during the last forty years. Specific conditions based on maximal temperature and relative humidity thresholds were determinant in outbreaks occurrence. Their persistence was also crucial. An operational model that will enable health authorities to anticipate the outbreak risk was successfully developed. Similar models may be developed to improve dengue management in other countries.

Citation: Descloux E, Mangeas M, Menkes CE, Lengaigne M, Leroy A, et al. (2012) Climate-Based Models for Understanding and Forecasting Dengue Epidemics. PLoS Negl Trop Dis 6(2): e1470. doi:10.1371/journal.pntd.0001470

Editor: Assaf Anyamba, NASA Goddard Space Flight Center, United States of America

Received May 18, 2011; **Accepted** November 21, 2011; **Published** February 14, 2012

Copyright: © 2012 Descloux et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by funds from the French Overseas Ministry (research program SEPDE/RC/n°508 titled "Prévention et prévision des épidémies de dengue en Nouvelle-Calédonie"). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: elodie.descloux@hotmail.com

Introduction

Dengue viruses are the most important arthropod-borne viruses affecting humans. During the past century, the four serotypes (DENV 1 - DENV 4) have spread to about a hundred countries in the tropical and subtropical world including Asia, Africa, the Americas and the Pacific. Each year, an estimated 50 million people contract dengue fever with at least 500,000 cases of dengue haemorrhagic fever or dengue shock syndrome leading to 25,000 deaths [1]. The spatial distribution of this emerging infectious disease largely reflects the distribution of its primary urban mosquito vector, *Aedes aegypti* [2]. As no effective vaccine and specific treatment exist, vector control currently represents the only resource to mitigate dengue outbreaks.

Epidemic dynamics of dengue, like those of other vector-borne diseases, are driven by complex interactions between hosts, vectors and viruses that are influenced by environmental and climatic factors. Several determinants in dengue fever emergence have been identified including human population growth, accelerated urbanization, increased international transport, weakened public

Author Summary

Dengue fever is a major public health problem in the tropics and subtropics. Since no vaccine exists, understanding and predicting outbreaks remain of crucial interest. Climate influences the mosquito-vector biology and the viral transmission cycle. Its impact on dengue dynamics is of growing interest. We analyzed the epidemiology of dengue in Noumea (New Caledonia) from 1971 to 2010 and its relationships with local and remote climate conditions using an original approach combining a comparison of epidemic and non epidemic years, bivariate and multivariate analyses. We found that the occurrence of outbreaks in Noumea was strongly influenced by climate during the last forty years. Efficient models were developed to estimate the yearly risk of outbreak as a function of two meteorological variables that were contemporaneous (explicative model) or prior (predictive model) to the outbreak onset. Local threshold values of maximal temperature and relative humidity were identified. Our results provide new insights to understand the link between climate and dengue outbreaks, and have a substantial impact on dengue management in New Caledonia since the health authorities have integrated these models into their decision making process and vector control policies. This raises the possibility to provide similar early warning systems in other countries.

health infrastructure as well as a lack of effective vector control and disease surveillance [3–6]. On the other hand, there is growing interest in the impact of climate change on the emergence or re-emergence of vector-borne infectious diseases such as dengue [7–10]. It has been shown that climate-induced variations in modelled *A. aegypti* populations were strongly correlated to reported historical dengue cases (1958–1995) at the global scale [11], and a potential increase in the latitudinal and altitudinal distribution of *A. aegypti* and dengue are expected under global warming [5,12].

In a specific ecosystem, the required conditions for the occurrence of a dengue outbreak include i) the presence of a dengue virus, ii) the presence and a sufficient density of competent vectors, iii) a sufficient number of susceptible humans that is serotype-specific, and iv) favorable environmental and climatic conditions for dengue transmission. Despite evidence that climate can influence dengue like other vector-borne diseases (i.e. vector population size and distribution, vector-pathogen-host interactions, and pathogen replication [7,10,13–14]), the relationships between climate, *Aedes* mosquitoes density and behaviour, human populations and dengue incidence are not well understood.

Previous studies have shown that temperature influences the lengths of the mosquito gonotrophic cycle and the extrinsic incubation period of the virus within the mosquito, the survival rate of adults, the mosquitoes population size and feeding behaviours and the speed of virus replication [7,13,15–19]. Water is necessary for eggs and larva development, mosquito breeding, and humidity affects adult mortality [16–17,20–22]. Temperatures and precipitations have been identified as influencing incidence rates of dengue in several endemic areas in the world (i.e. Thailand [23–24], Taiwan [25–27], Singapore [28], and Puerto Rico [24,29]). On a broader scale, it is plausible that El Niño-Southern Oscillation (ENSO) also influences patterns of dengue transmission [23–24,30–31]. This coupled ocean-atmosphere phenomena results in warm waters displacement and changes in sea surface temperatures (SST) across the Pacific Ocean, and has a strong influence on regional climates, particularly in the Pacific. ENSO can induce large temperature, humidity and precipitation changes for months (see the websites of the International Research Institute for Climate and Society (IRI, www.iri.org), and the National Oceanic and Atmospheric Administration (NOAA, www.noaa.gov) for more details). Importantly, previous studies revealed a positive correlation between ENSO, as measured by the Southern Oscillation Index (SOI), and dengue outbreaks in the South Pacific islands [30–31].

Our study was conducted in New Caledonia where dengue represents a major public health problem like in many Pacific Islands Countries and Territories [32]. The first dengue outbreak in New Caledonia occurred in 1884–1885 [33]. Disease transmission increased after World War II, and successive waves of epidemics involving all four serotypes were reported. Since 2000, serotype 1 has been predominant [34] causing more than 6,000 cases during the 2003–2004 epidemics [35] and about one thousand of cases in 2008. Although the serotype 4 [36] was involved in a major outbreak in 2009 (8,456 cases), the serotype 1 is still circulating. New Caledonia has had an effective surveillance system for dengue and access to high quality meteorological data for many years. Since 2000, regular entomological surveillance is performed. This provides an opportunity to study the influence of climate variations on dengue dynamics.

We analyzed the epidemiology of dengue fever in Noumea, the capital of New Caledonia, from 1971 to 2010 together with local and remote climate influences. The objectives of this study were i) to improve our knowledge of the relationships between meteorological variables, entomological surveillance indices and dengue fever dynamics at seasonal to inter-annual time scales, ii) to identify suitable conditions for an epidemic occurrence, and iii) to develop a predictive model for dengue outbreaks that can be integrated in an early warning system in New Caledonia.

Methods

Study area

New Caledonia is a French overseas territory located in the subregion of Melanesia in the southwest Pacific, about 1,200 kilometres east of Australia and 1,500 kilometres northwest of New Zealand. It lies astride the Tropic of Capricorn, between 19° and 23° south latitude. Its climate is tropical.

This archipelago of 18,575 square kilometres is made up of a main mountainous island elongated northwest-southeast 400 kilometres in length and 50–70 kilometres wide, the Loyalty Islands (Mare, Lifou, and Ouvea), and several smaller islands (e.g. Isle of Pines). The population was estimated in January 2009 to be 245,580 [37]. Approximately half of inhabitants are concentrated in the southeast region of the main island around Noumea, the capital.

A. aegypti is the only mosquito vector of dengue in New Caledonia. The two others vectors of dengue present in the Pacific region, *A. albopictus* and *A. polynesiensis*, have never been detected in this archipelago [38–40]. In Noumea, most of *A. aegypti* breeding sites are outdoors and therefore rainfall dependent.

Data collection

Epidemiological data. All cases of dengue fever and dengue haemorrhagic fever reported from January 1971 to December 2010 were collected from the Pasteur Institute, the Health Department of the Direction of Health and Social Affairs of New Caledonia, and the Communicable Disease Surveillance Division, Secretariat of the Pacific Community. A clinical case was defined as sustained fever and at least two of the following criteria: nausea or vomiting, myalgia or arthralgia, headache or retro-orbital pain, rash and/or spontaneous bleeding. A laboratory

positive case was defined as a single or paired serum sample positive in serological assays (hemagglutination inhibition, IgM detection by indirect immunofluorescence or ELISA), or direct detection of dengue virus by reverse-transcriptase polymerase chain reaction (RT-PCR using a pandengue technique), virus isolation, or NS1 antigen detection (ELISA or rapid immunochromatographic test). The serotyping of positive samples using RT-PCR with specific primers for DENV-1, 2, 3, and 4 was regularly performed at the Pasteur Institute over the 1971–2010 period.

Since 1995, a georeferencing of dengue cases per council has been performed. To identify the possible origin of dengue infection, travel history and first day of illness were determined by the Health Department of the Direction of Health and Social Affairs of New Caledonia. Imported dengue cases were defined as laboratory positive dengue cases with travel history to endemic countries within 14 days before the date of disease onset.

Incidence rates of dengue were calculated each year in New Caledonia (number of dengue cases per 10,000 inhabitants per year) using population data based on linear extrapolations of local census reports (1969, 1976, 1983, 1989, 1996, 2004, 2009 [37]). In Noumea, annual dengue incidence rates were computed for the 1995–2010 period from observed data. As georeferencing was not available before 1995, and as there is a strong linear relationship between incidence rates observed in Noumea and those observed in the entire territory over the 1995–2010 period, we used a linear fit to estimate dengue incidence rates in Noumea for the 1971–1994 period (Figure 1).

Two methods were used to separate the years during which dengue outbreaks occurred (denoted epidemic years) or did not occur (denoted non epidemic years) on the basis of annual incidence rates of dengue cases computed from January to December in Noumea. The first method denoted "tercile method" divided the years into three groups: epidemic years when the dengue incidence rate belonged to the upper tercile, non epidemic years when the dengue incidence rate belonged to the lower tercile, and unclassifiable years when the dengue incidence rate belonged to the central tercile. The second method denoted "median method", divided years into two groups: epidemic years when the annual incidence rate was greater than the median of the annual dengue incidence rates over the 1971–2010 period, and non epidemic years when the annual incidence rate was lower than the median. The first method allowed the problem of epidemic threshold to be minimised and to ensure a clear separation between epidemic and non epidemic years but with a 30% data loss while the second one allowed models to be built using the whole set of data.

Meteorological data. Two types of meteorological data were used: meteorological data measured at the reference weather station of Météo-France in central Noumea, and ENSO indices.

Data collected at the Noumea weather station for the period January 1971 to December 2010, the time period of the available dengue data, were analyzed. This station provides observations that are representative of the local climate around Noumea which contributes the most dengue cases in New Caledonia (Figure 1), and where dengue outbreaks usually begin. From these daily data, monthly, quarterly and annual means were calculated as well as monthly and quarterly number of days with a daily parameter greater than a given threshold. Quarterly data were generated with a sliding window each month. Monthly and quarterly parameters were named "parameter_month", and "parameter_first letter of each month of the quarter", respectively. The meteorological parameters of interest were daily minimum, mean, and maximum temperatures (min Temp, mean Temp, max Temp), daily minimum, mean, and maximum relative humidity (min RH, mean RH, max RH), and cumulative precipitations (Precip). Other parameters that may influence the productivity of larval breeding sites and mosquitoes populations were also

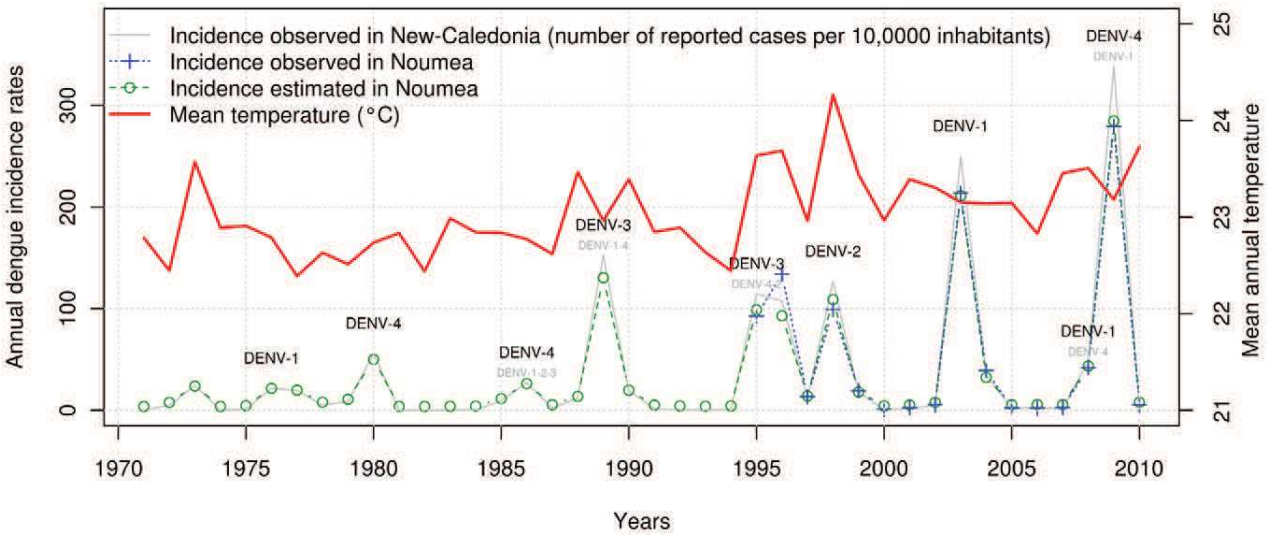


Figure 1. Epidemiology of dengue fever and evolution of annual mean temperature in Noumea-New Caledonia (1971–2010). The predominant circulating serotype (DENV-1, DENV-2, DENV-3 or DENV-4) is indicated in black characters. When other serotypes were detected, they are indicated in little grey characters. Annual dengue incidence rates observed in Noumea over the 1995–2010 period are highly correlated with dengue incidence rates observed in New Caledonia (Spearman coefficient *rho* = 0.99, *p*-value = $1*10^{-14}$). Annual dengue incidence rates in Noumea (1971–1994) were estimated (green dotted line with circles) on the basis of the relationship between incidence rates observed in New Caledonia (grey line) and those observed in Noumea (blue dotted line with crosses) using a linear model. During the 1971–2010 period, dengue incidence rates and annual mean temperatures (from January to December) were significantly correlated in Noumea (Spearman's coefficient *rho* = 0.426, *p*-value = 0.007). An increasing trend of dengue outbreaks amplitude and annual mean temperatures were observed during this 40-year study period.
doi:10.1371/journal.pntd.0001470.g001

considered such as mean daily wind force at 10 meters (WF), potential evapotranspiration by Penman-Monteith (ETP) and potential hydric balance sheets (HB = Precip-ETP) reflecting water resources. Numbers of days with a parameter over a threshold x were named NOD_parameter_threshold x. Several temperature thresholds were analyzed for min Temp, mean Temp and max Temp, ranging from 21 to 25°C, 24 to 28°C, and 25 to 35°C, respectively. Different thresholds were also analyzed for min RH (50%, 60%, 70%), mean RH (70%, 80%) and max RH (80%, 90%, 95%), for Precip (0.1, 1, 2, 3, 4, 5, 10, 25 mm/day), ETP (4, 5, 6, 7), HB (−5, 0, 5, 10 mm), and WF (3, 4, 5, 6, 7, 8 m/s).

Several ENSO indices were integrated in the analysis: Niño 3, Niño 3.4, Niño 4, Southern Oscillation Index (SOI), and Multivariate ENSO Index (MEI). Time series of these monthly ENSO indices were obtained from the NOAA Climate Prediction Center [41].

Altogether approximately 4000 meteorological data were generated for the 1971–2010 period (monthly, quarterly, and annual values). They were aggregated in seven families: temperature, relative humidity, precipitations, wind force, potential evapotranspiration, hydric balance sheet and ENSO.

Entomological surveillance data. An entomological surveillance network was established since 1997 at the initiative of the Pasteur Institute, the Health Department of the Direction of Health and Social Affairs of New Caledonia, and councils of Noumea and its neighbouring towns.

Since March 2000, about one hundred randomly selected houses in each of three districts of Noumea (East, West, and South) in a homogeneous and representative panel of 6,608 houses were visited each month to determine the number and type of larval developmental places, and the number of *A. aegypti* larvae, pupae and female adults. The following surveillance indices were computed monthly by the Laboratory of Medical Entomology of the Pasteur Institute:

i) House Index (HI) = number of houses with at least one larval breeding site positive for *A. aegypti* x 100)/number of inspected premises.

ii) Breteau Index (BI) = number of larval breeding sites positive for *A. aegypti*/100 inspected premises.

iii) Adult Productivity Index (API) = number of *A. aegypti* pupae and stage 4 larvae/number of inspected premises.

Statistical analysis and modelling

Bivariate and multivariate analyses were conducted using the R software package (R development Core Team version 2.9.1 [42]).

Time series analysis. Time series analysis of monthly, quarterly and annual data of dengue incidence rates, entomological indices and climatic variables were studied. Their temporal evolution was studied at inter-annual and seasonal scales. Global trends were computed for epidemiological and meteorological time series using linear regression (trend line).

Bivariate analysis. The relationships between epidemiological and meteorological data, entomological and meteorological data, and entomological and epidemiological data were studied in Noumea at different time-scales using a Spearman's method with p-values below 0.05 indicating statistical significance. At the annual scale, time series of annual dengue incidence rates and annual means of meteorological variables were analyzed from 1971 to 2010. At the monthly scale, time-lagged correlation analyses (lag being equal to 0, 1, 2 and 3 months) were performed on time series of monthly means of meteorological variables, entomological indices and dengue incidence rates from March 2000 to December 2009.

Comparative analysis of epidemic years and non epidemic years. To minimize the influence of changes in disease surveillance and diagnosis over the 1971–2010 period, we decided to use series of epidemic years (0 for non epidemic years, 1 for epidemic years, according to the tercile method described above) rather than dengue incidence rates.

Epidemic and non epidemic years were compared to identify suitable seasonal meteorological patterns for dengue outbreak occurrence. Monthly and quarterly meteorological data observed in Noumea during epidemic and non epidemic years were compared from August (year y-1) to July (year y) and means and 95% confidence interval (IC95%) were calculated. Categorical variables were compared using a two-sided t-test and correlation analyses were performed using a Spearman's rank correlation test. The p-values below 0.05 were considered to indicate statistical significance.

Multivariate modelling of dengue outbreak risk. The final objective of this study was to design two types of model to predict the risk of dengue outbreak in Noumea. The first model named hereafter "explicative model" was expected to identify suitable conditions for an epidemic occurrence using data from September (year y-1) to April (year y), i.e. four months before and after the outbreak onset (in January). The second model named hereafter "predictive model" was intended to help the health authorities of New Caledonia to anticipate the risk of a dengue outbreak. Only meteorological variables available prior to the outbreak onset, i.e. from September (year y-1) to December (year y-1) were used in this framework. On the basis of the bivariate analysis results, we decided to focus on the monthly and quarterly meteorological data. Poorly correlated variables such as wind force were excluded from the pool of potential input variables.

The type of classification method used for both explicative and predictive models was the Support Vector Machines (SVM) which is a supervised pattern recognition technique recently introduced in Statistical Learning Theory [43]. The main advantage of this method is that SVM are based on the principle of Structural Risk Minimization rather than on the error rates as do many other methods. SVM focus on generalizing well rather than correctly classifying the training dataset (i.e. minimizing the generalization error rather than the training error). The concept of SVM is to design a function which correctly classifies all of the objects of the training dataset. In the linearly separable case, SVM allow the identification of an hyperplane which is defined by the following equation: $\boldsymbol{w}.x+b=0$ where $\boldsymbol{w}$ is a vector normal to the hyperplane and b is the bias. In the non linear case, the separating surface is found by mapping the input points onto a higher dimensional space where the training dataset become linearly separable and by using an appropriate kernel (here a Gaussian kernel) in the optimization process [43].

In our study, the SVM took as input a set of meteorological data and predicted, for each given input, which one of the two possible classes the input is a member (epidemic year or non epidemic year). All the available data (40 years) were used for training the model and the median method, introduced above, was applied to separate the years. The results were then supplied as probability estimates of dengue outbreak occurrence using the method developed by Wu et al. [44].

The selection of the most relevant model was achieved using a forward stepwise selection method based on the corrected Akaike Information Criterion (AIC_c) [45–46]. This method not only rewards goodness of fit, but also includes a penalty that discourages overfitting.

The robustness of the explicative and predictive models was estimated using a leave-one-out cross validation method: a single observation (year *y*) from the original sample (1971–2010 years) was retained as a validation data for testing the model, and the remaining observations were used as training data. This process was repeated 40 times such that each yearly observation in the sample was used once as the validation data. The results from the folds then were averaged to produce a single estimation of dengue outbreak risk in Noumea each year. The performance of the models was estimated with the Receiver Operator Characteristics - Area Under the Curve (ROC-AUC). The sensitivity, specificity, positive predictive value and negative predictive value were calculated for each model.

Results

Time series analysis

Dengue data. During the 1971–2010 period, successive waves of dengue outbreaks involving the four serotypes were recorded in New Caledonia with an increasing magnitude, particularly in Noumea where dengue outbreaks usually begin (Figure 1). The annual dengue incidence rates revealed a global upward linear trend (mean increase of 65.4 dengue cases per 10,000 inhabitants over the studied period in Noumea). The most severe outbreaks were caused by DENV-1 and more recently DENV-4 in 2003 (5673 reported cases, 733 hospitalizations, 19 deaths), 2008 (1170 reported cases, ~100 hospitalizations, two deaths) and 2009 (8456 reported cases, 470 hospitalizations, three deaths). On four occasions, dengue outbreaks were repeated in two successive years: in 1976–1977 (DENV-1), 1995–1996 (DENV-3), 2003–2004 (DENV-1), and 2008–2009 (DENV-1 and DENV-4).

The analysis of monthly reported and laboratory positive cases revealed a strong seasonal distribution of dengue cases during epidemic years (Figure 2). The majority of outbreaks displayed a similar seasonal evolution: beginning in January, an epidemic peak between March and May, and ending in July. The temporal distribution of dengue cases during non epidemic years was different, with an occurrence of cases every month. Imported dengue cases from different locations in Asia and the Pacific (particularly Indonesia, the Philippines and French Polynesia) were recorded once or several times a year without a clear seasonal pattern.

Entomological data. Entomological surveillance data were available from March 2000 to December 2009 in Noumea and a decreasing trend of all entomological indices was observed (supporting Figure S1). Indices reflecting the distribution and the abundance of larval developmental places (HI and BI), and the vector density (API) were strongly correlated (HI versus BI: $rho = 0.98$, p-value<0.001; API versus HI: $rho = 0.82$, p-value<0.001; API versus BI: $rho = 0.84$, p-value<0.001).

Monthly means of HI, BI and API revealed a strong seasonal pattern with highest values between January and July (Figure 3).

Meteorological data. Over the 1971–2010 period, time series of annual means of daily mean Temp, Precip, and mean RH were characterized by a strong inter-annual variability. A number of ENSO events were observed including the strongest El Niño events of the century (i.e. 1982–1983 and 1997–1998). A global upward linear trend of annual mean Temp (mean increase of 0.75°C over the studied period, Figure 1) was observed in contrast with the Precip and mean RH time series that did not display any trend.

Rainfall is highly seasonal in New Caledonia. There are two main seasons: a warm and wet season (November–April), and a cooler and drier season (May–October). From November to April, max Temp in Noumea commonly reaches 30°C (on average during 42 days) and 6-month cumulative Precip 630 mm, whereas from May to October, max Temp rarely reaches 30°C (on average during only 2 days) and 6-month cumulative Precip are around 430 mm. The peak of mean Temp (February) precedes the peak of Precip and mean RH (March) with a lag of one month.

Bivariate analysis

During the 1971–2010 period, a significant correlation was found between dengue incidence rates and mean annual mean Temp in Noumea (Spearman's coefficient $rho = 0.426$, p-value $= 0.007$, Figure 1) but there was no significant correlation with annual mean RH and Precip. Similar results were obtained with conserved trends and detrended data. Anomalies of annual means of mean Temp, Precip and mean RH were significantly correlated with ENSO, as measured by Niño 3.4 ($rho = -0.365$, p-value $= 0.029$; $rho = -0.481$, p-value $= 0.003$; rho -0.486, p-value $= 0.003$, respectively). During El Niño (positive value of Niño 3.4), the weather was cooler and drier. During La Niña (negative value of Niño 3.4), the weather was warmer and wetter. However, no direct correlation was found between ENSO and dengue incidence rates at the inter-annual scale ($rho = -0.106$, p-value $= 0.539$). Dengue outbreaks occurred during either El Niño, La Niña or neutral phases of ENSO.

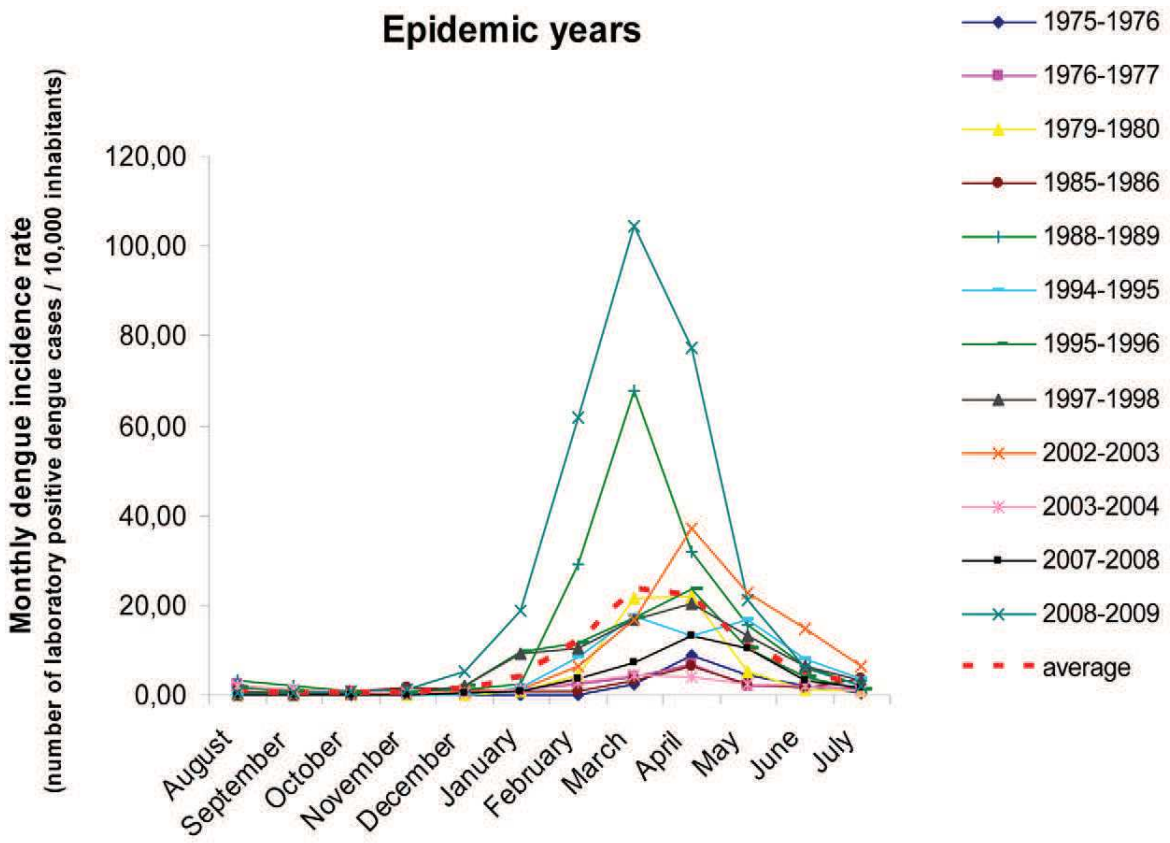
During the 2000–2009 period, dengue incidence rates, meteorological and entomological data were analyzed in Noumea at a monthly scale. A strong seasonal distribution of HI, BI and API was observed (Figure 3), and significant correlations were found between monthly entomological surveillance indices and climate variables (data not shown). Although the highest dengue incidence rates and the highest values of HI, BI and API were observed during the same period of the year (from January to July), no significant time-lagged correlation has been found between monthly entomological indices and dengue incidence rates reported in Noumea over the 2000–2009 period (supporting Figure S1). We did not find relevant entomological patterns during dengue outbreaks. Accordingly, entomological surveillance indices were not used for the modelling of dengue outbreak risk.

Comparative analysis of epidemic and non epidemic years

Based on the tercile method, there were 13 epidemic years (dengue incidence rate in the upper tercile, i.e. >19.48 cases/10 000 inhabitants) and 13 non epidemic years (dengue incidence rate in the lower tercile, i.e. <4.13 cases/10 000 inhabitants). A detailed analysis was performed based on monthly and quarterly meteorological data measured from September (year *y*-1) to April (year *y*), i.e. four months before and after the outbreak onset.

Temperatures (min Temp, mean Temp and max Temp) were higher during epidemic years than during non epidemic years. The peak of max Temp, observed usually in February, preceded the epidemic peak of dengue with a lag of 1–2 months (Figure 4a). Analysis of daily data allowed identifying important temperature thresholds. It revealed that the number of days with max Temp exceeding 32°C, mean Temp exceeding 27°C, and min Temp exceeding 22°C were significantly higher during epidemic years than during non epidemic years. The most important and significant differences were observed during the first quarter of the year, principally in February for max Temp (p-$value<0.01$ using a t-test, Figure 4b).

By contrast, the relationships between Precip, mean RH and dengue dynamics were not clear, as shown in supporting Figure S2. Highest Precip and mean RH were observed in February–March–April during the epidemic phase of dengue. Using a t-test,


Epidemic years
Monthly dengue incidence rate
(number of laboratory positive dengue cases / 10,000 inhabitants)
120,00
100,00
80,00
60,00
40,00
20,00
0,00
August
September
October
November
December
January
February
March
April
May
June
July
1975-1976
1976-1977
1979-1980
1985-1986
1988-1989
1994-1995
1995-1996
1997-1998
2002-2003
2003-2004
2007-2008
2008-2009
average

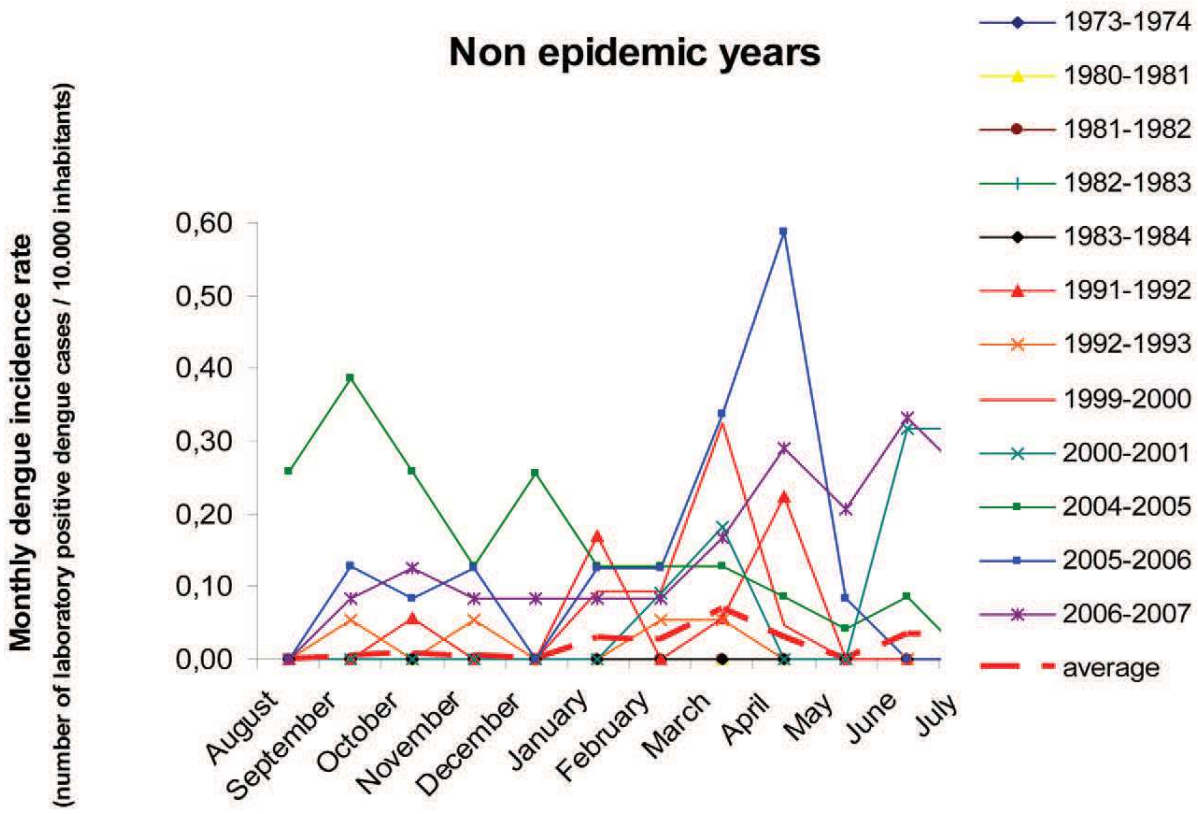

Non epidemic years
Monthly dengue incidence rate
(number of laboratory positive dengue cases / 10.000 inhabitants)
0,60
0,50
0,40
0,30
0,20
0,10
0,00
August
September
October
November
December
January
February
March
April
May
June
July
1973-1974
1980-1981
1981-1982
1982-1983
1983-1984
1991-1992
1992-1993
1999-2000
2000-2001
2004-2005
2005-2006
2006-2007
average

Figure 2. Monthly distribution of laboratory positive dengue cases during epidemic and non epidemic years. A strong seasonality in the dengue cases distribution was observed during epidemic years with outbreaks occurring usually between January and July. By contrast, dengue cases occurred almost every month without a clear seasonal pattern during non epidemic years.
doi:10.1371/journal.pntd.0001470.g002

Precip and mean RH were significantly lower in February during epidemic years than during non epidemic years (p-value<0.01 and $=0.04$, respectively). Inversely, the ETP was significantly higher in February (p-value $=0.02$). WF, HB, ENSO indices and entomological surveillance indices were not significantly different between epidemic and non epidemic years.

Meteorological variables showing strongest correlations with the epidemic years series, as defined in the Methods section, are presented for each family of variables in Table 1. Significant correlations were identified with several local meteorological variables (particularly Temp, Precip, RH, and ETP) but not with ENSO indices. No or poor correlation was found with WF and HB. In accordance with Figure 4 and supporting Figure S2, Temp were positively correlated with dengue outbreaks in Noumea, whereas Precip and RH measured in February were negatively correlated with dengue outbreaks. A positive correlation was found between the ETP measured in February and the occurrence of dengue outbreaks.

Multivariate modelling of dengue outbreak risk

First, in order to produce an explicative model of dengue outbreak, we selected meteorological variables observed within the period of dengue outbreak onset, i.e. from January to April (Figure 2). The best SVM model based on the minimum AIC_c (-79.21) was obtained using two meteorological variables, i.e. the number of days with maximal temperature exceeding 32°C during the first quarter of the year (NOD_max Temp_32_JFM), and the number of days with maximal relative humidity exceeding 95% during January (NOD_max RH_95_January). The addition of a third meteorological variable did not improve the performance of the model. Results obtained in leave-one-out cross validation (Figure 5) were close to those obtained with the complete dataset (Figure S3) and were characterized by a high ROC-AUC value reaching 0.80 and 0.85, respectively. As indicated by the ROC curves, most of epidemic years were predicted correctly with high probability and few false alarms. Importantly, with bivariate analysis, NOD_max Temp_32_JFM was positively correlated with the occurrence of dengue outbreak ($rho=0.57$, p-value $=0.002$) whereas NOD_max RH_95_January did not appear to be a discriminatory meteorological variable ($rho=-0.11$, p-value $=0.58$). With multivariate analysis, these two variables were highly informative and discriminatory. Scatter plots of epidemic and non epidemic years as a function of these two variables allowed the identification of three distinct groups (Figure 6): group A including years characterized by low NOD_max Temp_32_JFM (<12 days) and low NOD_max RH_95_January (<12 days), group B including years characterized by high NOD_max Temp_32_JFM (>12 days) and low NOD_max RH_95_January, and group C including years characterized by low NOD_max Temp_32_JFM and high NOD_max RH_95_January (>12 days). According to the tercile method of years classification, all non epidemic years belonged to group A

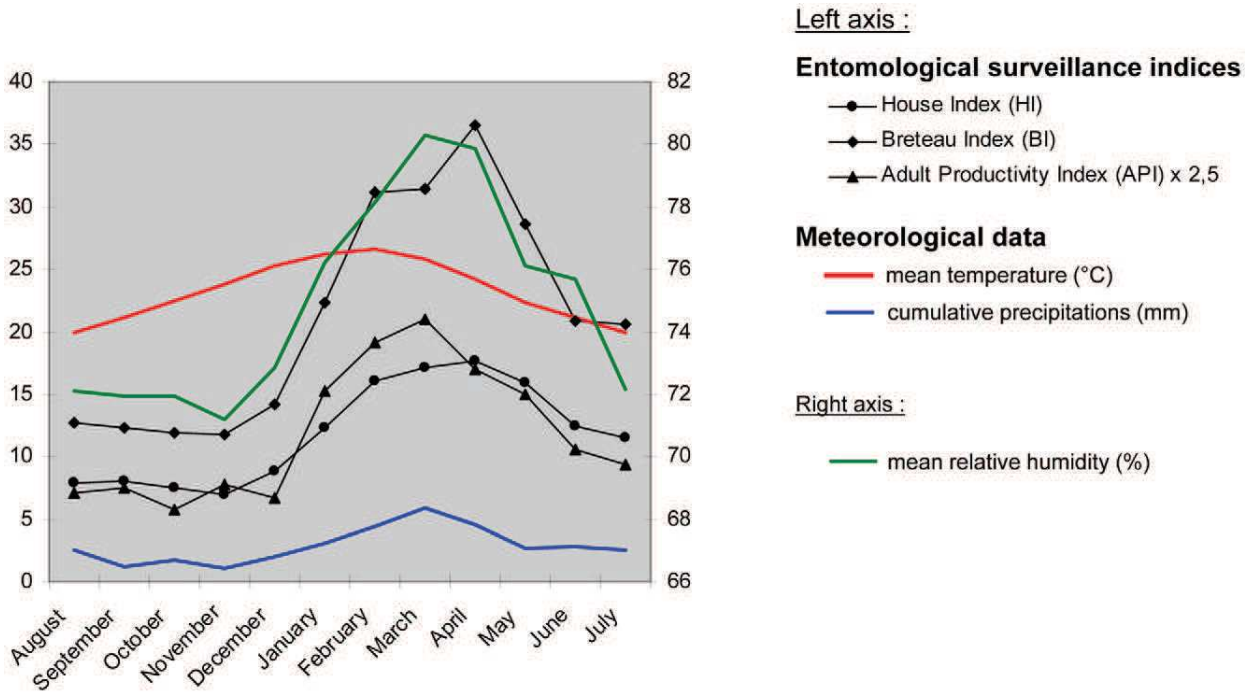
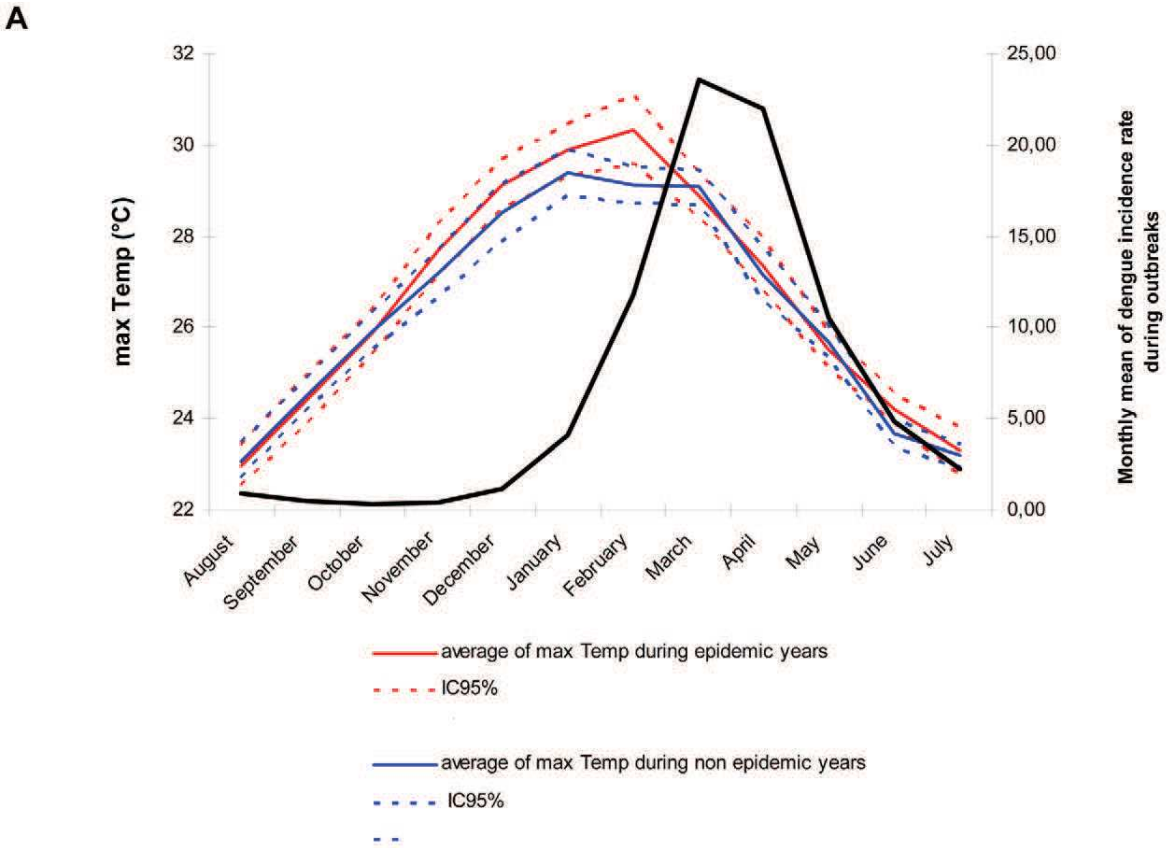


Figure 3. Seasonal evolution of monthly entomological surveillance indices and meteorological data in Noumea (August 2000–July 2009). HI, BI and API evolution display a strong seasonal cycle, with highest values between January and July. Entomological surveillance indices were significantly correlated with meteorological data at the seasonal scale. The peak of mean Temp preceded the peak of Precip, mean RH and API with a lag of one month, and the peak of HI and BI with a lag of two months.
doi:10.1371/journal.pntd.0001470.g003


A
max Temp (°C)
32
30
28
26
24
22
25,00
20,00
15,00
10,00
5,00
0,00
Monthly mean of dengue incidence rate during outbreaks
August
September
October
November
December
January
February
March
April
May
June
July
average of max Temp during epidemic years
IC95%
average of max Temp during non epidemic years
IC95%

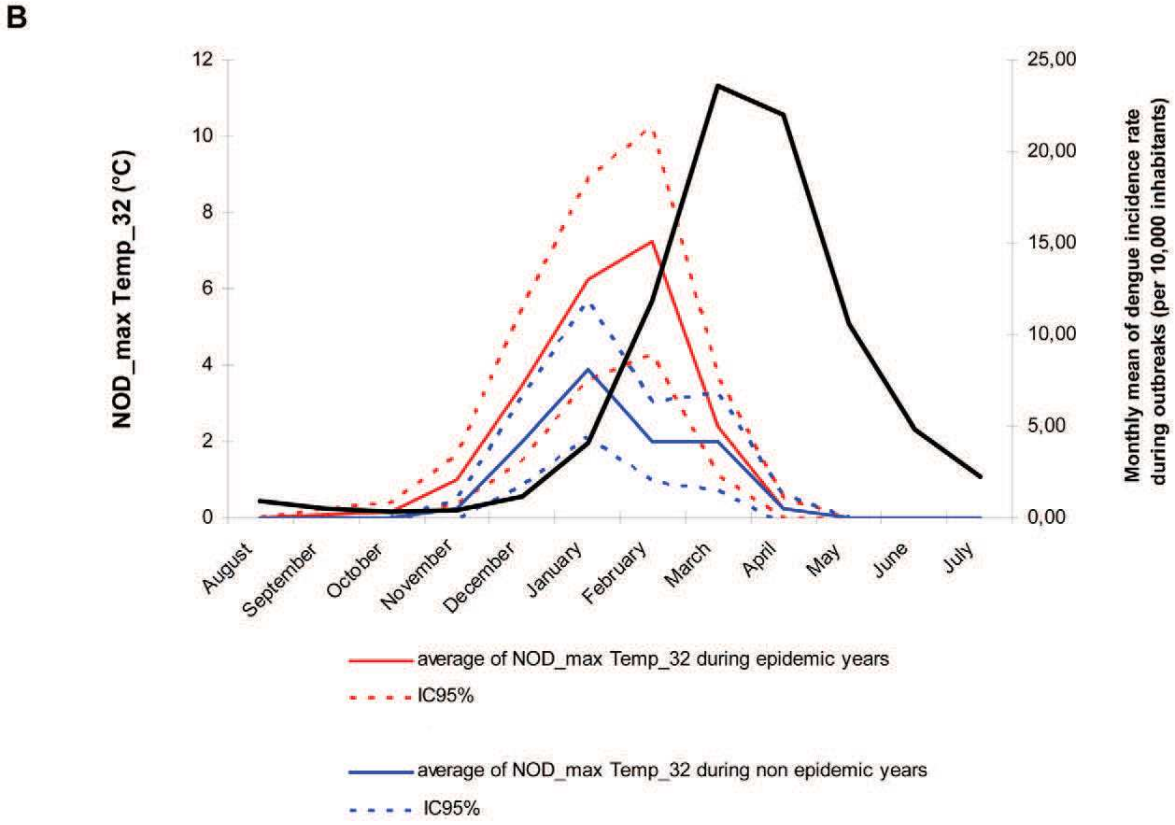

B
NOD_max Temp_32 (°C)
12
10
8
6
4
2
0
25,00
20,00
15,00
10,00
5,00
0,00
Monthly mean of dengue incidence rate during outbreaks (per 10,000 inhabitants)
August
September
October
November
December
January
February
March
April
May
June
July
average of NOD_max Temp_32 during epidemic years
IC95%
average of NOD_max Temp_32 during non epidemic years
IC95%

Figure 4. Relationship between maximal temperatures and dengue outbreaks in Noumea. Averages and 95% confidence intervals (IC95%) of max Temp (Figure 4a) and NOD_max Temp_32 (Figure 4b) calculated monthly during epidemic and non epidemic years were compared from August (year *y*-1) to July (year *y*). The peak of max Temp preceded the epidemic peak of dengue with a lag of 1–2 months. The number of days with max Temp exceeding 32°C during the first quarter of the year was significantly higher during epidemic years than during non epidemic years, especially in February (NOD_max Temp_32_February = 7.25 versus 2 days, respectively).
doi:10.1371/journal.pntd.0001470.g004

whereas all epidemic years, except 1973 and 2003, belonged to either group B or group C. Similar results were obtained using the median method ensuring the inclusion of all years, preferable for the development of SVM models. Only four years (1978, 1979, 1985, and 2002) belonging to the middle tercile (dengue incidence rate ranging from 4.13 to 19.48 cases/10 000 inhabitants/year) were incorrectly classified using the median method. In 2002, although favorable climatic conditions for dengue outbreak were observed, the incidence rate (5.24 dengue cases/10 000 inhabitants/year) was close to the median (7.65 dengue cases/10 000 inhabitants/year). In 1978, 1979 and 1985, the low values of NOD_max Temp_32_JFM and NOD_max RH_95_January were not favorable for dengue outbreak. However, incidence rates (7.74, 10.63, and 11.24 dengue cases/10 000 inhabitants/year, respectively) were close to the median. Two years (1973 and 2003) belonging to epidemic years using either a tercile or a median method of classification were characterized by low NOD_max RH_95_January and intermediate NOD_max Temp_32_JFM, as members of group A (non epidemic years). However, dengue outbreaks occurred with high incidence rates (23.64 and 213.58 dengue cases/10 000 inhabitants/year in 1973 and 2003, respectively). These mismatches indicate that i) the model fails for years that are difficult to classify as their dengue incidence rates were close to the median and in the middle tercile and, ii) NOD_max Temp_32_JFM and NOD_max RH_95_January alone cannot account for all dengue outbreaks (Figure 6). It is likely that other climate events and other factors influencing dengue dynamics contribute to the epidemic spread of dengue viruses during these peculiar years. We were thus able to build an efficient explicative model of dengue epidemics based on meteorological variables contemporaneous to the outbreak.

Table 1. Correlations between meteorological variables and dengue outbreaks in Noumea.

	Spearman's rank correlation test	
	rho coefficient	*p*-value
Temperature (°C)		
NOD_min Temp_22_JFM	0.58	<0.01
NOD_mean Temp_27_NDJ	0.59	<0.01
NOD_max Temp_32_JFM	0.51	<0.01
Relative humidity (%)		
NOD_min RH_70_February	−0.47	0.01
NOD_max RH_95_February	−0.47	0.01
NOD_max RH_80_SON	0.47	0.02
Precipitations (mm)		
Precip_February	−0.57	<0.01
NOD_Precip_0.1_December	−0.43	0.03
NOD_Precip_10_February	−0.41	0.04
Potential evapotranspiration (mm)		
ETP_February	0.44	0.02
NOD_ETP_4_February	0.50	0.01
NOD_ETP_6_FMA	0.47	0.01
Hydric balance sheet (mm)		
NOD_HB_10_February	−0.45	0.02
NOD_HB_0.5_FMA	−0.37	0.06
NOD_HB_5_ DJF	−0.32	0.11
Wind Force (m/s)		
NOD_WF_3_NDJ	0.41	0.04
NOD_WF_7_SON	0.37	0.07
WF_September	0.23	0.26
ENSO		
NINO.3.4_November	0.20	0.33
NINO.3.4_ASO	0.19	0.34
NINO.4_November	0.18	0.37

Monthly and quarterly meteorological data measured from September (year *y*-1) to April (year *y*), i.e. four months before or after the outbreak onset, were analyzed from 1971 to 2010 in Noumea. For each family of meteorological variables, the three variables most correlated with the occurrence of dengue outbreaks are presented, *p*-value<0.05 indicating statistical significance. Monthly and quarterly parameters were named "parameter_month", and "parameter_first letter of each month of the quarter", respectively. Number of days with a parameter over a threshold *x* were named NOD_parameter_threshold *x*.
doi:10.1371/journal.pntd.0001470.t001

Another challenge was to construct a predictive model for dengue epidemics using variables available prior to the outbreak onset, i.e. from September (year *y*-1) to December (year *y*-1). Accurate predictive skill (AICc = −66.64) was achieved with the SVM model built from the value of the two following variables: the quarterly mean of maximal relative humidity during October–November–December (max RH_OND), and the monthly mean of maximal temperature in December (max Temp_December) of the year *y*-1 with a ROC-AUC value of 0.83 (supporting Figure S4). Probabilities obtained in leave-one-out cross validation (Figure 7) and the corresponding ROC-AUC value reaching 0.69 illustrate the robustness of this predictive model. Importantly, max RH_OND and max Temp_December were not significantly correlated with the risk of dengue outbreak with bivariate analysis (*rho* = 0.24, *p*-value = 0.14; and *rho* = 0.25, *p*-value = 0.14, respectively).

Scatter plots of epidemic years and non epidemic years built from the combination of meteorological variables used for the SVM explicative model (Figure 8) and for the SVM predictive model development (Figure 9) show that dengue outbreaks occurred in distinct climatic conditions in Noumea. With the SVM predictive model, as noted with the SVM explicative model, epidemic years belonged to two different groups of data according to the value of max RH_OND and max Temp_December (see the two red kernels corresponding to high risk of dengue outbreak in Figure 9). Dengue outbreaks occurred following either years characterized by high max Temp_December and relatively low max RH_OND, or years characterized by high max RH_OND_December, and max Temp_December. To note, the high value of max Temp_December (31.2°C) and the relatively low value of max RH_OND (86.8%) measured in 2010 indicate a high risk (74%) of dengue outbreak for 2011.

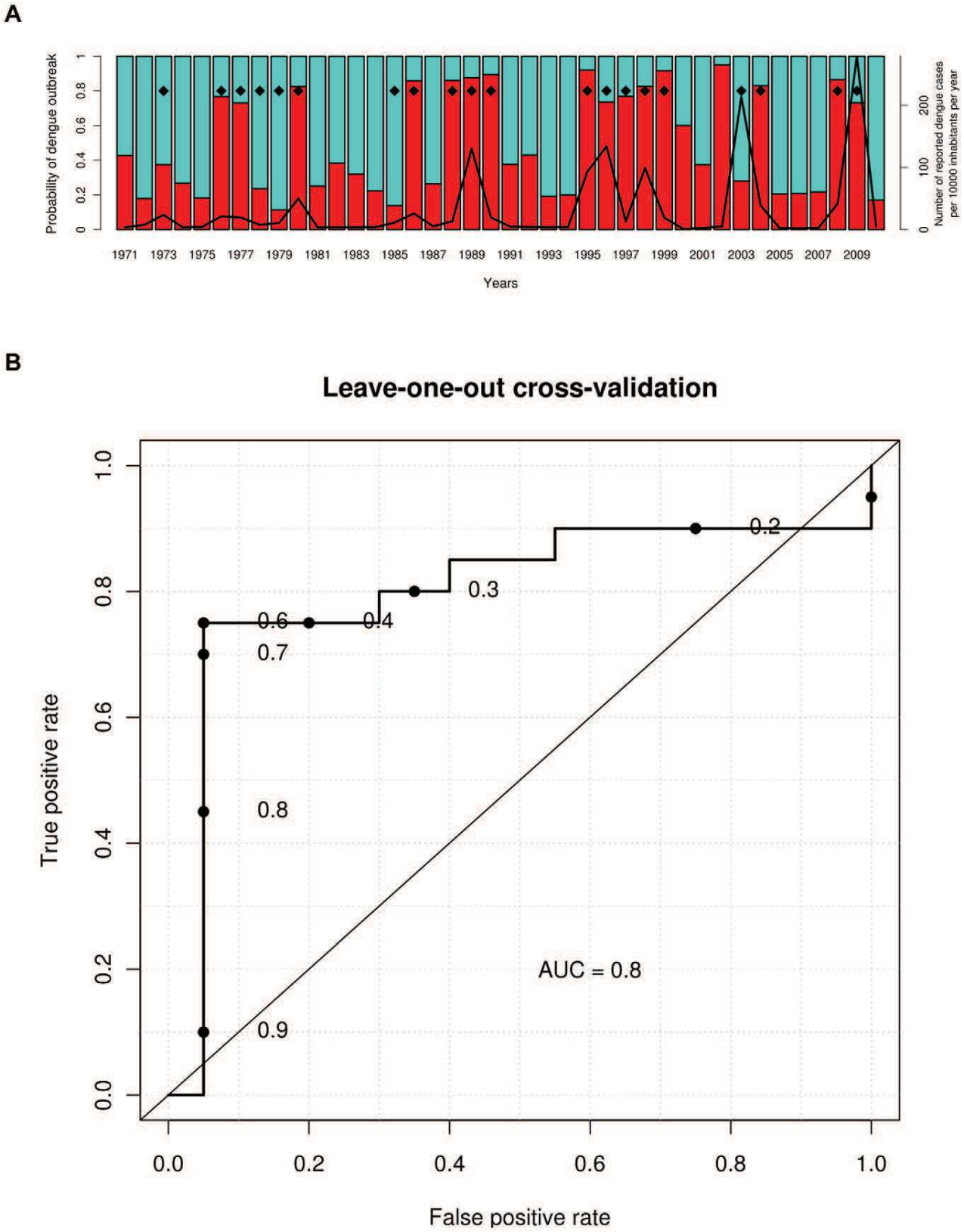


Figure 5. SVM explicative model of dengue outbreaks in Noumea (leave-one-out cross validation). The model estimates the probability of dengue outbreak occurrence (red bars) each year according to the number of days with maximal temperature exceeding 32°C during the first quarter of the year (NOD_max Temp_32_JFM), and the number of days with maximal relative humidity exceeding 95% during January (NOD_max

RH_95_January). Results obtained in leave-one-out cross validation are presented in Figure 5a. The black line indicates the annual dengue incidence rate, and black diamonds indicate epidemic years according to the median method. The ROC curve (Figure 5b) indicates the rates of true and false positives for different detection thresholds. For example, for a probability of dengue outbreak above 65% (0.65), 15 of 20 epidemic years are predicted correctly (true positive rate = 75%) with only one false alarm (false positive rate = 5%). The sensitivity of the model for this threshold is 75% (15 epidemic years predicted correctly/20 epidemic years), the specificity 95% (19 non epidemic years predicted correctly/20 non epidemic years), the positive predictive value 94% (15 epidemic years predicted correctly/16 epidemic years predicted by the model), and the negative predictive value 79% (19 non epidemic years predicted correctly/24 non epidemic years predicted by the model).
doi:10.1371/journal.pntd.0001470.g005

A comparison of the results obtained with the explicative model and the predictive model was performed together with a detailed analysis of the relationships between meteorological variables used to build the explicative model (NOD_max Temp_32_JFM and NOD_max RH_95_January) and those used to build the predictive model (max RH_OND and max Temp_December). As shown in Figure S5, strong relationships exist between the values of max Temp and max RH measured at the end of the year *y*-1, and those measured at the beginning of the year *y*. Low max RH_OND and max Temp_December (year *y*-1) were predictive of low NOD_max Temp_32_JFM and NOD_max RH_95_January (years *y*, group A). High max RH_OND and max Temp_December (year *y*-1) were predictive of either high NOD_max Temp_32_JFM and low NOD_max RH_95_January (years *y*, group B), or low NOD_max Temp_32_JFM and high NOD_max RH_95_January (years *y*, group C). Results obtained with the predictive model were highly consistent with those obtained with the explicative model with similar probabilities of dengue outbreak risk obtained for 30 of the 40 studied years. Failures of the predictive model can be explained by a lack of correlation between these meteorological variables on a few occasions (e.g. 1982, 1983, 1995). For example, although the predictive model estimated a risk of dengue outbreak close to 5% in 1995, the explicative model estimated a risk over 90%, and a major outbreak occurred. The value of max RH_OND and max Temp_December measured in 1994 (87% and 27.6°C, respectively) were relatively low and therefore not predictive of outbreak risk. However, climatic conditions were favorable for a dengue outbreak occurrence (NOD_max Temp_32_JFM = 20 days, NOD_max RH_95_January = 0 day, group B). This suggests that

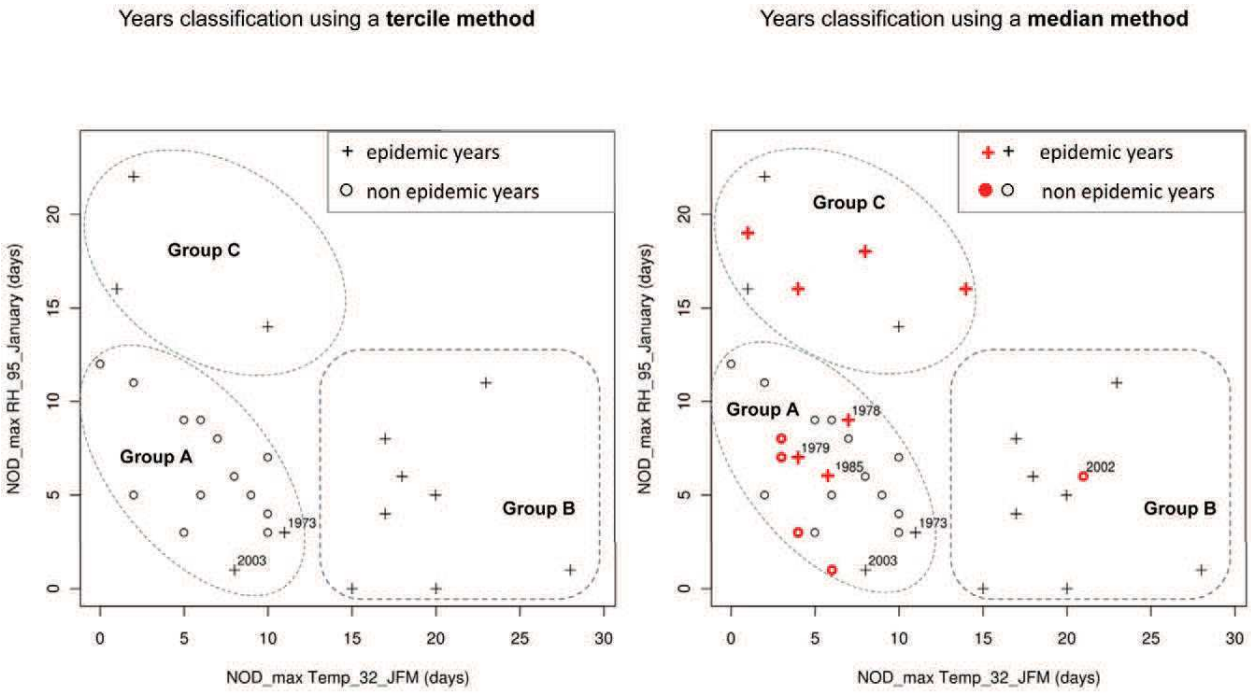


Figure 6. Scatter plots of epidemic and non epidemic years with regards to NOD_max Temp_32_JFM and NOD_max RH_95_January. Each year, the number of days with maximal temperature exceeding 32°C during January–February–March (NOD_max Temp_32_JFM) and the number of days with maximal relative humidity exceeding 95% during January (NOD_max RH_95_January) were calculated. Two methods denoted "tercile method" and "median method" were used to separate the years on the basis of annual dengue incidence rates in Noumea (see Methods). On the left panel, epidemic years (dengue incidence rate in the upper tercile, i.e. >19.48 cases/10,000 inhabitants/year) and non epidemic years (dengue incidence rate in the lower tercile, i.e. <4.13 cases/10,000 inhabitants/year) are presented. The distribution of crosses (epidemic years) and circles (non epidemic years) permits the identification of three groups (A, B, C). All non epidemic years belonged to group A whereas all epidemic years, except 1973 and 2003, belonged to either group B or group C suggesting that dengue outbreaks can occur in distinct climatic conditions. On the right panel, epidemic years (dengue incidence rate greater than the median, i.e. 7.65 cases/10,000 inhabitants/year) and non epidemic years (dengue incidence rate lower than the median) are presented with the advantage of a whole set of data being usable for modelling. Years that were not considered with the tercile method (dengue incidence rate in the middle tercile) are coloured in red. Further epidemic (red crosses) and non epidemic years (red circles) are considered with the median method, and similar groups (A, B, C) were identified. With the median method, three epidemic years (1978, 1979 and 1985) and one non epidemic year (2002) were incorrectly classified. These four years were characterized by annual dengue incidence rates closed to the median.
doi:10.1371/journal.pntd.0001470.g006

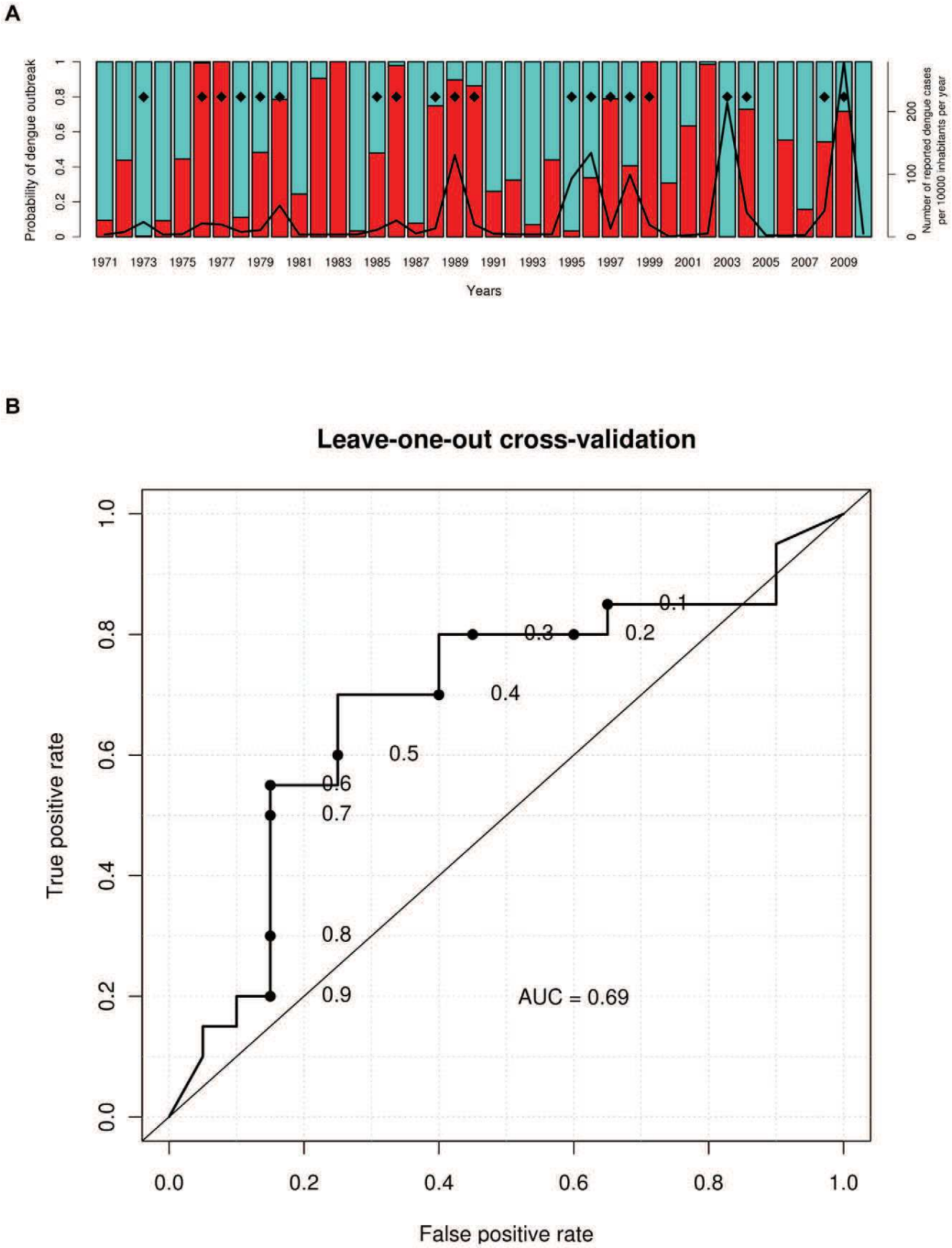


Figure 7. SVM predictive model of dengue outbreaks in Noumea (leave-one-out cross validation). The model estimates the probability of dengue outbreak occurrence (red bars) each year *y* according to the quarterly mean of maximal relative humidity during October–November–December (max RH_OND), and the monthly mean of maximal temperature in December (max Temp_December) year *y*-1. Results obtained in leave-

one-out cross validation are presented in Figure 7a. The black line indicates the annual dengue incidence rate, and black diamonds indicate epidemic years according to the median method. The ROC curve (Figure 7b) indicates the rates of true and false positives for different detection thresholds. For example, for a probability of dengue outbreak above 65% (0.65), 11 of 20 epidemic years were predicted correctly (true positive rate = 55%) with three false alarms (false positive rate = 15%). The sensitivity of this model for this threshold is 55% (11 epidemic years predicted correctly/10 epidemic years), the specificity 85% (17 non epidemic years predicted correctly/20 non epidemic years), the positive predictive value 79% (11 epidemic years predicted correctly/14 epidemic years predicted by the model), and the negative predictive value 65% (17 non epidemic years predicted correctly/26 non epidemic years predicted by the model).
doi:10.1371/journal.pntd.0001470.g007

other climate variables or meteorological processes may impact on the local value of NOD_max Temp_32_JFM and NOD_max RH_95_January.

Discussion

The influence of climate on dengue dynamics in Noumea, the capital of New Caledonia, over the 1971–2010 period has been analyzed at different time scales using high quality and high resolution meteorological observation data, along with epidemiological and entomological surveillance data. During epidemic years, dengue outbreaks peaked around March–April at the end of summer season. The epidemic peak lagged the warmest temperature by 1–2 months and was in phase with maximum precipitations and maximum relative humidity. The seasonal

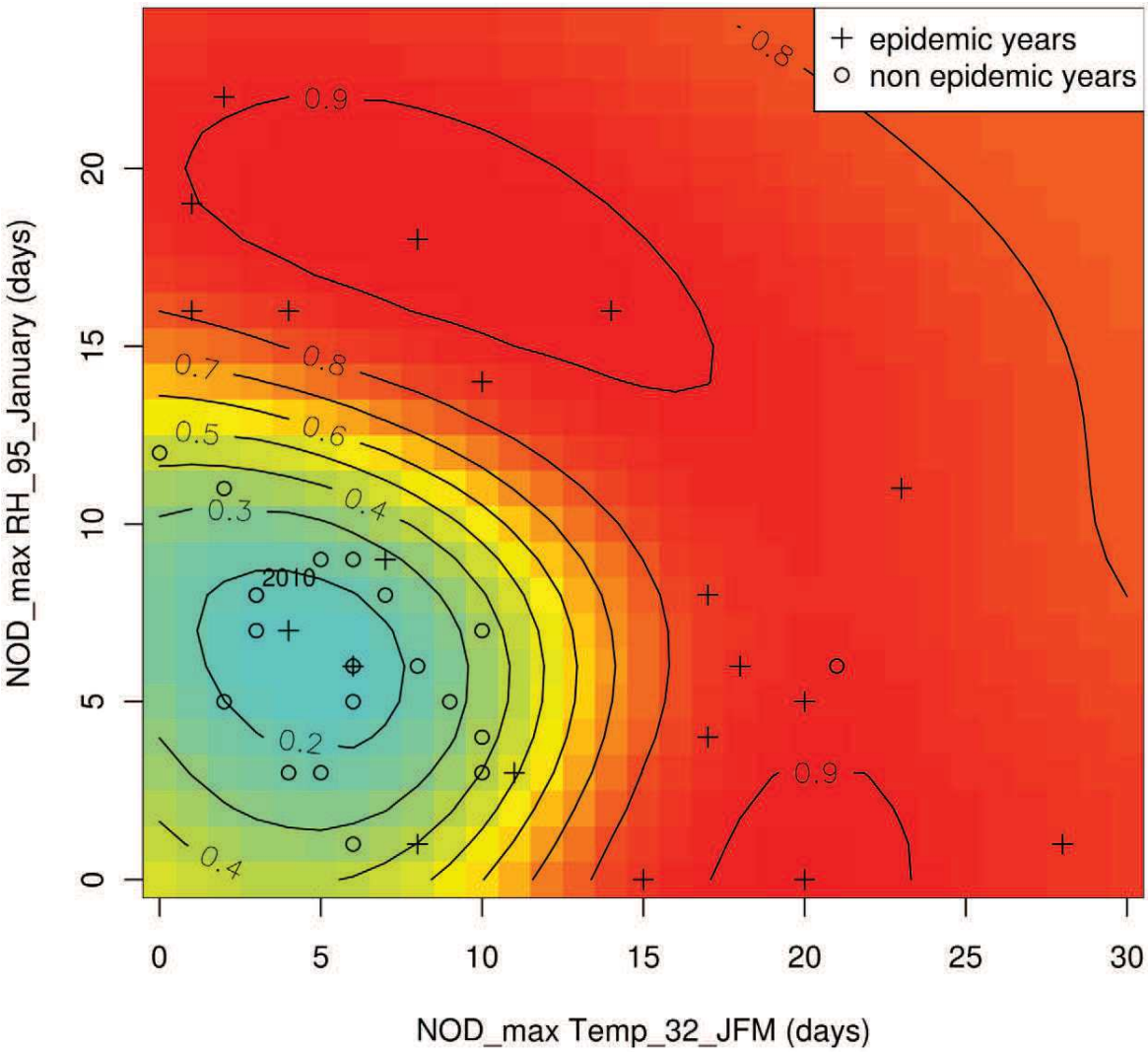


Figure 8. SVM explicative model probability contours superimposed with NOD_max Temp_32_JFM and NOD_max RH_95_January during epidemic/non epidemic years. Line-curves indicate the estimated probability of dengue outbreak occurrence given by the model. Blue colour indicates low risk, yellow colour indicates intermediate risk, and red colour indicates high risk of dengue outbreak. Meteorological parameters used to build the SVM models are shown for epidemic years (crosses) and non epidemic years (circles). The number of days with maximal temperature exceeding 32°C during January–February–March (NOD_max Temp_32_JFM) and the number of days with maximal relative humidity >95% during January (NOD_max RH_95_January) of the year *y* were used to build the SVM explicative model.
doi:10.1371/journal.pntd.0001470.g008

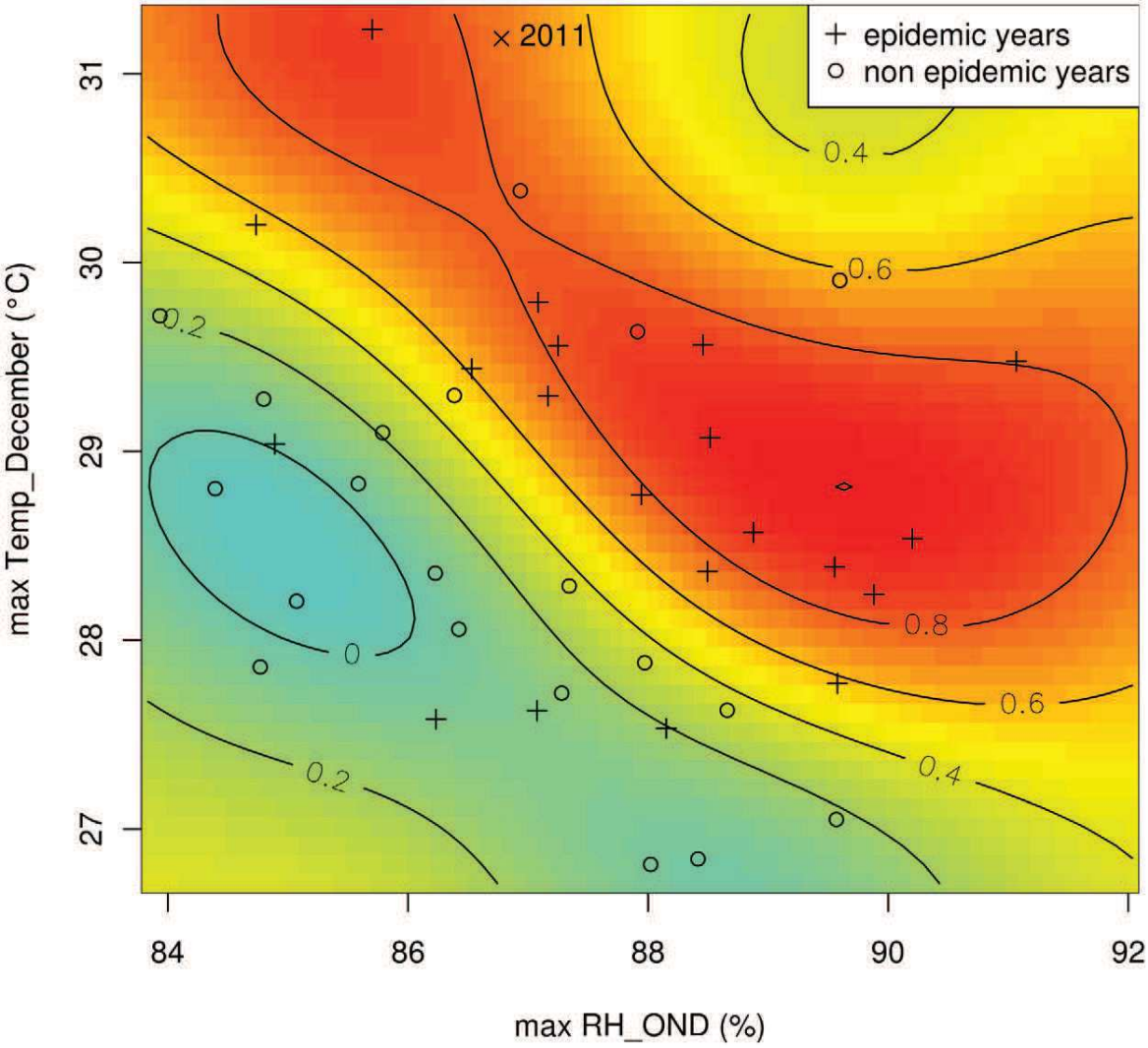


Figure 9. SVM predictive model probability contours superimposed with max RH_OND and max Temp_December during epidemic/non epidemic years. Similarly to the SVM explicative model (Figure 8), the quarterly mean of maximal relative humidity during October–November–December (max RH_OND), and the monthly mean maximal temperature in December (max Temp_December) of the year *y*-1 were used to build the SVM predictive model.
doi:10.1371/journal.pntd.0001470.g009

evolution of entomological indices (e.g, Breteau, House and Adult productivity indices) matched the seasonality of dengue outbreaks.

No relationship was found between the inter-annual variations of dengue incidence rates and those of the entomological data. On the other hand, a number of meteorological indices developed from summertime temperature, precipitation or relative humidity showed a significant correlation with dengue occurrence.

New explicative and operational predictive models of dengue outbreak were developed. We used a multivariate SVM model to identify the best set of meteorological variables explaining dengue epidemics. We found that a non linear combination of two meteorological variables strongly outperforms a model based on a single variable or a linear approach, as commonly employed in the literature. We found the best explicative variables to be the number of days with max Temp exceeding 32°C during January–February–March (NOD_max Temp_32_JFM) and the number of days with max RH exceeding 95% during January (NOD_max RH_95_January). When the model gives a probability of dengue outbreak above 65%, these two variables explain 94% of the epidemic years and 79% of the non epidemic years (Figure 5). Most dengue outbreaks occurred within two kinds of distinct climatic conditions: high NOD_max Temp_32_JFM and low NOD_max RH_95_January, or low NOD_max Temp_32_JFM and high NOD_max RH_95_January. We were also able to build another SVM model based on two variables to predict dengue outbreaks in advance: the maximal temperature in December (max Temp_December) and maximal relative humidity during October–November–December (max RH_OND) of the year prior to the epidemics. For a probability of dengue outbreak above 65%, this model can predict 79% of the epidemic years and 65% of the non epidemic years (Figure 7).

Influence of local meteorological conditions on dengue dynamics

Overall, the high performance of the climate-based models of dengue outbreak risk developed in our study suggest that dengue

dynamics were essentially driven by climate during this 1971–2010 period in Noumea. The explicative model provides important and new information. We have shown that maximal values of temperature and relative humidity were determinant in dengue outbreaks occurrence and precise thresholds of their value were identified. Importantly, we found that the most relevant meteorological variables explaining dengue outbreaks were built using the number of days for which the variable was greater than a threshold value introducing the importance of the persistence of suitable climatic conditions. Our findings are compatible with the mosquito biology and viral transmission cycle.

The length of *Aedes* gonotrophic cycle is shorter at temperatures above 32°C and feeding frequency is more than twofold at 32°C as compared to 24°C; pupae development period reduced from four days at 22°C to less than one day at 32–34°C [16–17,47]. Additionally, the experimental infection of *A. aegypti* with DENV-2 viruses showed that the extrinsic incubation period shortens from 12 days at 30°C to seven days at 32–35°C leading to an increasing risk of viral transmission from an infected mosquito to a susceptible host [15]. The influence of temperature on the rate of virus replication inside mosquitoes was also evidenced in the study of Watts *et al.* Temperatures may also influence the vector size and its biting rate [19,21]. Consequently, it is likely that the increased level of viral transmission characterizing dengue outbreaks in Noumea at temperatures exceeding 32°C may be a consequence of shortening of the *A. aegypti* gonotrophic cycle and extrinsic incubation period, and of increased vector feeding frequency.

Mortality rate of larvae, pupae and adult mosquitoes as a function of temperature between 10 and 40°C can be represented by a wide-base 'U' graphical shape with lower mortality rate at temperature ranging from 15 to 30°C [16–20,22]. Hence, *A. aegypti* mortality rate may be relatively constant at temperatures observed usually in Noumea, and the increasing mortality rate expected above 32°C is not likely to be an important limiting parameter in the spread of dengue viruses in this specific ecosystem.

Larval breeding places are mostly outdoors in Noumea and mosquito abundance increases during the rainy and humid season. Moreover, relative humidity may be determinant in *A. aegypti* egg development and adult population size that may itself be correlated with vectorial capacity [48]. High humidity shortens incubation and blood-feeding intervals; it favours adult mosquito longevity [20] and thus dengue transmission. This may explain why a sustained high RH during January is associated with a higher risk of dengue outbreak in Noumea.

Influence of remote climate conditions on dengue dynamics

On a broader scale, a growing number of studies have shown that ENSO may be associated with changes in the risk of mosquito borne diseases such as dengue [23–24]. By contrast, Hales *et al.* [31] further analyzed the relationships between the annual number of dengue cases in New Caledonia, ENSO, temperature and rainfall using global atmospheric reanalyses climate based data, and they did not find any significant correlation between SOI and dengue (Pearson's coefficient = 0.20). In accordance with this study, and with the advantage of observational and long term data, we found significant inter-annual correlations between ENSO and our local climate but not between ENSO and dengue (Table 1). Moreover, the selection process of multivariate models did not select any ENSO index neither in explicative mode nor in predictive mode. These findings suggest that, in New Caledonia, large-scale climate indices such as ENSO cannot account for the complexity of the local meteorological inter-annual situations. However, at a larger scale, Hales *et al.* showed that the number of dengue outbreaks in the South Pacific islands (aggregated data, 1970–1995) were positively correlated with the SOI [30], suggesting that La Niña may favour dengue outbreaks in this region of the world. The impact of ENSO on local weather in the South Pacific may strongly vary from one place to another. New Caledonia, located around 20° south latitude in the western Pacific is relatively far from the main centre of action of ENSO located in the equatorial central/eastern equatorial Pacific and its local weather is thus not only influenced by ENSO, but also by other climate modes such as the Madden-Julian Oscillation which strongly influences local meteorological parameters at intra-seasonal (30 to 90 days) time scales [49]. In contrast, ENSO influence may be stronger in islands located closer to the equator, the relationship between ENSO and dengue epidemics being therefore more straightforward [29].

Our long-term study also suggests an increasing risk of dengue outbreaks in New Caledonia in the context of global warming (Figure 1). Even though a global upward trend of dengue incidence rates was noted along the 1971–2010 period, and as surveillance methods and laboratory tests have evolved, it is difficult to know if the amplitude of dengue outbreaks is significantly growing.

Dengue dynamics driven by multiple factors

Even though climate influenced the disease epidemiology in Noumea during this forty-year period, the reasons of dengue emergence in New Caledonia are multiple, including population growth (119,710 inhabitants in 1973 to 245,580 in 2009), accelerated urbanization particularly around Noumea, tourism development and increasing international and inter-islands traffic [50]. The emergence of dengue fever in other parts of the world, particularly South East Asia where dengue is endemic with a co-circulation of the four serotypes, represents an increasing source of virus introduction into New Caledonia. Indeed, multiple and repeated introductions of dengue viruses have been detected from several countries in Asia [34]. Moreover, the geographical distribution of *A. aegypti* has expanded during recent decades in New Caledonia (Paupy and Guillaumot, unpublished data).

Well known factors may have contributed to the epidemic dynamics such as the size of susceptible human hosts and vectors populations. In the absence of seroprevalence data, and due to the lack of long term entomological data, these variables were not included in the input dataset of the models. Nevertheless, as dengue is known to confer a prolonged serotype-specific immunity in the long term, herd immunity represents an important factor in understanding dengue dynamics [51–54]. In New Caledonia, successive waves of dengue outbreaks involving the same serotype were reported in 1980 and 1986 (DENV-4), 1989 and 1995 (DENV-3), 2003 and 2008 (DENV-1). This constant interval time between two epidemics involving the same serotype has already been observed in other South Pacific Islands [55–57]. Recently, a large molecular characterization of DENV-1 viruses collected regularly in French Polynesia between the 2001 and 2006 outbreaks revealed that the virus responsible for the severe 2001 outbreak was introduced from South-East Asia, and evolved under an endemic mode until its re-emergence under an epidemic mode five years later [56]. These findings suggest that 5–6 years may be necessary for the renewal of the susceptible population in these islands. In New Caledonia, at four occasions, dengue outbreaks were detected between January and July during two successive years: in 1976–1977 (DENV-1), 1995–1996 (DENV-3), 2003–2004 (DENV-1), and 2008–2009 (DENV-1 and DENV-4). This

suggests that environmental conditions may be not favorable for dengue transmission all through the epidemic year, particularly during the second semester of the year characterized by lower values of entomological indices. It is likely that dengue re-emerged the following year when climatic conditions were favorable for dengue transmission (as suggested by the results of our explicative model in 1977, 1996, 2004 and 2009) and the size of the mosquito-vector and susceptible human populations were still sufficient for a large spread of dengue viruses. In these four examples of recurrent outbreaks during two consecutive years, it is more likely that the end of the epidemic was driven by limiting climatic factors and intricate entomological factors rather than by the depletion of the susceptible population.

The relationship between *Aedes* density and the intensity of dengue transmission remains unclear [47,58–60]. Although dengue viruses cannot circulate if mosquito vectors are not present, the vector density of adult female *A. aegypti* necessary for dengue viruses to become endemic or epidemic remains unknown. In Noumea, entomological indices (HI, BI and API) were not correlated with the incidence rate of dengue, they were sometimes lower during epidemic than during non epidemic periods and lowest values were measured during the largest outbreak in 2009. The fact that these usual entomological surveillance indices (particularly API) are good indicators of adult density in Noumea suggests that the mosquito density threshold under which dengue viruses cannot spread widely may be very low and has never been reached up to now. Moreover, mosquito populations are influenced by human behaviours and meteorological variables alone cannot account for their geographical distribution and abundance [14,61]. At the domestic level, *A. aegypti* populations are also influenced by global trends in urbanization, socioeconomic conditions, and vector control efforts. For instance, the outbreak predicted in 2002 with a probability close to 90% did not occur. A possible explanation is that strong vector control policies (e.g. increased efforts to reduce mosquito breeding sites and undertake human population education, development of perifocal spraying of insecticides) were undertaken in New Caledonia at the time of large dengue outbreaks in the other Pacific French overseas territories (French Polynesia in 2001, Wallis and Futuna in 2002). A relaxation in vector control efforts at the end of 2002 may have allowed the resurgence of dengue in the East coast and the spread of the virus through the archipelago during the next year.

Overall, our results suggest that the local climate had a major effect on dengue dynamics in Noumea during the last forty years. It is likely that other factors, not included in the input dataset of the models, had a lower influence on dengue epidemic dynamics. The introduction of dengue viruses may have been relatively constant, and the number of human hosts susceptible to a given serotype and of mosquito-vectors may have been always sufficient for an epidemic to occur when suitable climate conditions were met. It is likely that the susceptibility of human populations influenced the serotype involved in the outbreak and the epidemic magnitude. The variability of the length of the gonotrophic cycle, the extrinsic incubation period, and the life span of infected mosquitoes under climate change rather than the overall vector density may play a major role on the epidemic dynamics of dengue at the seasonal scale.

Epidemics forecasting model

Although the meteorological variables contemporaneous to the epidemic season provide crucial information on local dengue dynamics as discussed above, prediction models are needed to anticipate the risk before the dengue outbreak onset and to make the model useful for health authorities in New Caledonia. In this study, we were able to build such a predictive model relying on maximal temperature and relative humidity measured in Noumea at the end of the previous year.

Biological interpretations about statistical associations between specific climatic conditions and the yearly risk of dengue outbreak in Noumea can be made in the frame of the explicative model as it uses relevant climatic variables that occur within the period of outbreak onset. The meteorological variables selected in the frame of the predictive model are tightly connected with the explicative meteorological variables (Figure S5).

As Noumea concentrates the majority of inhabitants and of dengue cases, as this city has been affected by all dengue outbreaks that occurred in New Caledonia during the last 40 years, and as dengue epidemics usually begin in Noumea, our predictive model is useful to anticipate the risk of dengue outbreak in New Caledonia. However, climatic conditions in Noumea can not account for dengue epidemics in other localities in New Caledonia that would not involve Noumea, even if this situation has never been observed in 40 years.

Depending on the user's objectives, different detection thresholds corresponding to a probability of dengue outbreak can be used. In the case of dengue, it is likely that decision makers would prefer to choose a detection threshold with high true positive rate and low false positive rate, as obtained with a detection threshold of 65% (Figure 7b). The model initialized in December 2009 indicated no risk of dengue outbreak for 2010 that was in accordance with the current epidemiological situation. To note, a high risk of dengue outbreak is predicted for 2011 (74%, Figure 9). Up to now, only a few cases of dengue fever have been reported. Only one case imported from the Philippines was possible to type and belonged to the serotype 1. It is likely that a significant part of the human population is immunized against the serotypes 1 and 4 involved in the largest dengue outbreaks reported in New Caledonia in 2008 and 2009 but the introduction of a new serotype (DENV-2 or DENV-3) may lead to another epidemic. However, several important confusing factors may interfere with dengue dynamics this year such as the massive rainfalls brought by the tropical cyclone Vania in middle January 2011 with its unknown effects on vector populations, the introduction and worrying local diffusion of Chikungunya viruses transmitted by the same mosquito and the subsequent enhancement of vector control policies.

Conclusions and perspectives

In conclusion, the epidemic dynamics of dengue fever were strongly influenced by climate variability in Noumea during the 1971–2010 period. Local thresholds of maximal temperature and relative humidity have been identified with precision allowing the development of explicative and predictive climate-based models of dengue outbreak risk. The health authorities of New Caledonia have now integrated these models into their new decision making process in order to improve their management of dengue, in combination with clinical, laboratory (e.g. serotype determination), and entomological surveillance data. This work provides an example of the practical utility of research projects in operational public health fields and reinforces the need for a multidisciplinary approach in the understanding and management of vector-borne diseases. Our results provide also new insights for future experimental studies. It seems important now to study the impact of maximal temperatures exceeding 32°C and maximal relative humidity exceeding 95%, and the influence of their duration (more or less than 12 days) on the length of the extrinsic incubation period, feeding frequency and longevity of *A. aegypti* from New Caledonia.

The epidemic dynamics of dengue are driven by complex interactions between human-hosts, mosquito-vectors and viruses. These interactions are influenced by environmental and climatic factors that may have more or less burden according to the geographical localisation, the local climatic conditions, the vector characteristics (e.g. *Aedes* species and strains), the size and movements of human populations and the epidemiology of dengue. Consequently, our results can not be applied to other ecosystems. However, the methodology of analysis used in this study could be extended to other localities highly threatened by the emergence of dengue in the South Pacific, like in other tropical and subtropical countries. As global atmospheric reanalyses climate based data exist, there is hope for the development of local predictive models of dengue outbreak in countries where no reliable weather data are available.

Supporting Information

Figure S1 Evolution of House Index, Adult Productivity Index and dengue cases reported in Noumea (2000–2009). The monthly incidence rate of dengue cases (histograms) reported in Noumea from March 2000 to December 2009 was not significantly correlated (time-lag being equal to 0, 1, 2, or 3 months) with the value of HI (orange line) reflecting the abundance of larval resting places, and API (green line) reflecting the vector density. Although highest dengue incidence rates and highest values of entomological surveillance indices were observed during the same period of the year (from January to July), no relevant entomological patterns were identified during dengue outbreaks. A decreasing trend of entomological indices was observed that may reflect the impact of strengthened vector control policies. Sometimes, higher indices were measured during non epidemic than during epidemic years, and lowest indices were observed in 2009 whereas a major dengue outbreak occurred suggesting that the minimal vector density allowing the occurrence of dengue outbreaks may be very low.
(TIF)

Figure S2 Relationship between monthly cumulative precipitations, mean relative humidity and dengue outbreaks in Noumea. Averages and 95% confidence intervals (IC95%) of Precip (Figure S2a) and mean RH (Figure S2b) calculated monthly during epidemic and non epidemic years were compared from August (year *y*-1) to July (year *y*). Highest Precip and mean RH were observed during the epidemic phase of dengue.
(TIF)

Figure S3 SVM explicative model of dengue outbreaks in Noumea (complete dataset). The model estimates the probability of dengue outbreak occurrence (red bars) each year according to the number of days with maximal temperature exceeding 32°C during the first quarter of the year (NOD_max Temp_32_JFM), and the number of days with maximal relative humidity exceeding 95% during January (NOD_max RH_95_January). Results obtained with the complete dataset are presented in Figure S3a. The black line indicates the annual dengue incidence rate, and black diamonds indicate epidemic years according to the median method. The ROC curve (Figure S3b) indicates the rates of true and false positives for different detection thresholds.
(TIF)

Figure S4 SVM predictive model of dengue outbreaks in Noumea (complete dataset). The model estimates the probability of dengue outbreak occurrence (red bars) each year *y* according to the quarterly mean of maximal relative humidity during October–November–December (max RH_OND), and the monthly mean of maximal temperature in December (max Temp_December) of the year *y*-1. Results obtained with the complete dataset are presented in Figure S4a. The black line indicates the annual dengue incidence rate, and black diamonds indicate epidemic years according to the median method. The ROC curve (Figure S4b) indicates the rates of true and false positives for different detection thresholds.
(TIF)

Figure S5 Relationships between predictive climate variables (year *y*-1) and explicative climate variables (year *y*). Line-curves indicate the probability of dengue outbreak occurrence estimated by the SVM predictive model. Blue colour indicates low risk, yellow colour indicates intermediate risk, and red colour indicates high risk of dengue outbreak. The values of the quarterly mean of maximal relative humidity during October–November–December (max RH_OND), and the maximal temperature in December (max Temp_December) of the year *y*-1 used to build the SVM predictive model were calculated each year during the 1971–2010 period. The point coordinates were associated each year with the letter A, B, or C according to the value of the two climate variables used to build the SVM explicative model, i.e. the number of days with maximal temperature exceeding 32°C during January–February–March (NOD_max Temp_32_JFM) and the number of days with maximal relative humidity >95% during January (NOD_max RH_95_January). As in Figure 7, members of group A correspond to years *y* with a low NOD_max Temp_32_JFM and a low NOD_max RH_95_January. Members of group B correspond to years *y* with high NOD_max Temp_32_JFM and low NOD_max RH_95_January. Members of group C correspond to years *y* with high NOD_max Temp_32_JFM and high NOD_max RH_95_January. Most of members of the group A correspond to non epidemic years whereas most of members of the group B or C correspond to epidemic years. This figure illustrates the strong relationship existing between the predictive and the explicative climate variables used to build the models. Low max RH_OND and max Temp_December (year *y*-1) were predictive of low NOD_max Temp_32_JFM and NOD_max RH_95_January (years *y*, group A). High max RH_OND and max Temp_December (year *y*-1) were predictive of either high NOD_max Temp_32_JFM and low NOD_max RH_95_January (years *y*, group B), or low NOD_max Temp_32_JFM and high NOD_max RH_95_January (years *y*, group C).
(TIF)

Acknowledgments

We are grateful to Bérangère Arnould and Cédric Baillif, students at Météo-France for their help in climate data collection and analysis. We thank Suzanne Chanteau, director of the Pasteur Institute of New Caledonia, for helpful discussion and comments about the manuscript, Luc Monimeau from the Health Department of the Direction of Health and Social Affairs of New Caledonia, for maps construction. We thank Tony Kolbe and Simon Reid from the Secretariat of the Pacific Community, for their help with English translation and comments about the manuscript. We also thank all the sentinel physicians of New Caledonia for reporting dengue cases.

Author Contributions

Conceived and designed the experiments: MM CEM ED LG JB. Analyzed the data: ED MM CEM MT AL TT LG JB JPG ML ND XDL. Wrote the paper: ED MM CEM ML. Performed the epidemiological data collection and analysis: ACG AP JPG ED. Performed the entomological data collection and analysis: LG MT ED MM CEM. Performed the climate data collection and analysis: AL TT MM CM ED. Performed modelling: MM AL.

References

1. WHO Dengue guidelines for diagnosis, treatment, prevention and control (2009) Available: http://whqlibdoc.who.int/publications/2009/9789241547871_eng.pdf. Accessed 10 December 2010.
2. Gubler DJ (1998) Dengue and Dengue Hemorrhagic Fever. Clin Microbiol Rev 11: 480–496.
3. Rigau-Pérez JG, Clark GG, Gubler DJ, Reiter P, Sanders EJ, et al. (1998) Dengue and dengue haemorrhagic fever fever. Lancet 352: 971–977.
4. Gubler DJ (2002) Epidemic dengue/dengue hemorrhagic fever as a public health, social and economical problem in the 21th century. Trends Microbiol 10: 100–103.
5. Hales S, de Wet N, Maindonald J, Woodward A (2002) Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. Lancet 360: 830–834.
6. Mackenzie JS, Gubler DJ, Petersen LR (2004) Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. Nat Med 10: S98–S109.
7. Patz JA, Epstein PR, Burke TA, Balbus JM (1996) Global climate change and emerging infectious diseases. JAMA 275: 217–223.
8. Patz JA, Campbell-Lendrum D, Holloway T, Foley JA (2005) Impact of regional climate change on human health. Nature 438: 310–317.
9. Chan NY, Ebi KL, Smith F, Wilson TF, Smith AE (1999) An Integrated Assessment Framework for Climate Change and Infectious Diseases. Env Health Persp 107: 329–337.
10. Tabachnick WJ (2010) Challenges in predicting climate and environmental effects on vector-borne disease episystems in a changing world. J Exp Biol 213: 946–954.
11. Hopp MJ, Foley JA (2003) Worldwide fluctuations in dengue fever cases related to climate variability. Clim Res 25: 85–94.
12. Jetten TH, Focks DA (1997) Potential changes in the distribution of dengue transmission under climate warming. Am J Trop Med Hyg 57: 285–297.
13. Martens WJM, Jetten TH, Focks DA (1997) Sensitivity of malaria, schistosomiasis and dengue to global warming. Climatic Change 35: 145–156.
14. Jansen CC, Beebe NW (2010) The dengue vector *Aedes aegypti*: what comes next. Microbes Infect 12: 272–279.
15. Watts DM, Burke DS, Harrison BA, Whitmire RE, Nisalak A (1987) Effect of temperature on the vector efficiency of *Aedes aegypti* for dengue 2 virus. Am J Trop Med Hyg 36: 143–152.
16. Focks DA, Haile DG, Daniels E, Mount GA (1993) Dynamic life table model for *Aedes aegypti* (Diptera: *Culicidae*): analysis of the literature and model development. J Med Entomol 30: 1003–1017.
17. Focks DA, Haile DG, Daniels E, Mount GA (1993) Dynamic life table model for *Aedes aegypti* (Diptera: *Culicidae*): simulation results and validation. J Med Entomol 30: 1018–1028.
18. Focks DA, Daniels E, Haile DG, Keesling JE (1995) A simulation model of the epidemiology of urban dengue fever: litterature analysis, model development, preliminary validation, and samples of simulation results. Am J Trop Med Hyg 53: 489–506.
19. Yang HM, Macoris MLG, Galvani KC, Andrighetti MTM, Wanderley DMV (2009) Assessing the effects of temperature on the population of *Aedes aegypti*, the vector of dengue. Epidemiol Infect 137: 1188–1202.
20. Christophers SR (1960) *Aedes aegypti* (L.) The Yellow Fever Mosquito. Its Life History, Bionomics and Structure. University Press, Cambridge.
21. Rueda LM, Patel KJ, Axtell RC, Stinner RE (1990) Temperature-dependent development and survival rates of *Culex quinquefasciatus* and *Aedes aegypti* (Diptera: *Culicidae*). J Med Entomol 27: 892–898.
22. Hopp MJ, Foley JA (2001) Global-scale relationships between climate and the Dengue fever vector, *Aedes aegypti*. Clim Change 48: 441–463.
23. Cazelles B, Chavez M, McMichael AJ, Hales S (2005) Nonstationary influence of El Niño on the synchronous dengue epidemics in Thailand. PLoS Med 2(4): e106.
24. Johansson MA, Cummings DA, Glass GE (2009) Multiyear climate variability and dengue-El Niño southern oscillation, weather, and dengue incidence in Puerto Rico, Mexico, and Thailand: a longitudinal data analysis. PLoS Med 6(11): e1000168.
25. Hsieh YH, Chen CW (2009) Turning points, reproduction number, and impact of climatological events for multi-wave dengue outbreaks. Trop Med Int Health 14: 628–638.
26. Shang CS, Fang CT, Liu CM, Wen TH, Tsai KH, et al. (2010) The role of imported cases and favorable meteorological conditions in the onset of dengue epidemics. PLoS Negl Trop Dis 4(8): e775.
27. Chen SC, Liao CM, Chio CP, Chou HH, You SH, et al. (2010) Lagged temperature effect with mosquito transmission potential explains dengue variability in southern Taiwan: insights from a statistical analysis. Sci Total Environ 408: 4069–4075.
28. Hii YL, Rocklöv J, Ng N, Tang CS, Pang FY, et al. (2009) Climate variability and increase in intensity and magnitude of dengue incidence in Singapore. Glob Health Action;doi: 10.3402/gha.v2i0.2036.
29. Johansson MA, Dominici F, Glass GE (2009) Local and global effects of climate on dengue transmission in Puerto Rico. PLoS Negl Trop Dis 3(2): e382.
30. Hales S, Weinstein P, Woodward A (1996) Dengue fever epidemics in the South Pacific: driven by El Niño Southern Oscillation? Lancet 348: 1664–1665.
31. Hales S, Weinstein P, Souares Y, Woodward A (1999) El Niño and the dynamics of vectorborne disease transmission. Environ Health Perspect 107: 99–102.
32. Singh N, Kiedrzynski T, Lepers C, Kamisan Benyon E (2005) Dengue in the Pacific-an update of the current situation. Pac Health Dialog 12: 111–119.
33. Perry WJ (1950) The mosquitoes and mosquito-borne diseases on New-Caledonia, an historic account: 1885–1946. Am J Trop Med 30: 103–114.
34. A-Nuegoonpipat A, Berlioz-Arthaud A, Chow V, Endy T, Lowry K, et al. (2004) Sustained transmission of dengue virus type 1 in the Pacific due to repeated introductions of different Asian strains. Virology 329: 505–512.
35. Bouldouyre MA, Baumann F, Berlioz-Arthaud A, Chungue E, Lacassin F (2006) Factors of severity at admission during an epidemic of dengue 1 in New Caledonia (South Pacific) in 2003. Scand J Infect Dis 38: 675–681.
36. Li D-S, Liu W, Guigon A, Mostyn C, Grant R, Aaskov J (2010) Rapid displacement of Dengue virus type 1 by type 4, Pacific region, 2007–2009. Emerg Inf Dis 16: 123–125.
37. Historique du dénombrement de la population: population des communes de la Nouvelle-Calédonie de 1956 à 2009. Institut de la statistique et des études économiques de Nouvelle-Calédonie (ISEE). Available: http://www.isee.nc/population/population.html. Accessed 18 April 2011.
38. Rosen L, Rozeboom LE, Sweet BH, Sabin AB (1954) The transmission of dengue by *Aedes polynesiensis* Marks. Am J Trop Med Hyg 3: 878–882.
39. Failloux AB, Rodhain F (1999) Importance of mosquito population genetic studies in medical entomology. Examples from French Polynesia. Ann Soc Entomol Fr 35: 1–16.
40. Guillaumot L, Arboviruses and their vectors in the Pacific – Status report (2005) Pac Health Diag 12: 45–52.
41. Monthly atmospheric and SST Indices. Climate Prediction Center. Available: http//www.cpc.ncep.noaa.gov/data/indices. Accessed 10 January 2011.
42. R Development Core Team (2009) R: A Language and Environment for Statistical Computing. Available: http://www.r-project.org. Accessed 18 April 2011.
43. Vapnik V (1998) Statistical learning theory. New York: Wiley. 736 p.
44. Wu TF, Lin CJ, Weng RC (2004) Probability estimates for multi-class classification by pairwise coupling., J Mach Learn Res 5: 975–1005.
45. Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth edition. New York: Springer. 495 p.
46. Burnham KP, Anderson DR (2002) Model selection and multimodel interference: a practical information-theoretic approach. 2nd edition. Colorado State University: Springer-Verlag. 347 p.
47. Focks DA, Brenner RJ, Hayes J, Daniels E (2000) Transmission thresholds for dengue in terms of *Aedes aegypti* pupae per person with discussion of their utility in source reduction efforts. Am J Trop Med Hyg 62: 11–18.
48. Morales Vargas RE, Ya-Umphan P, Phumala-Morales N, Komalamisra N, Dujardin JP (2010) Climate associated size and shape changes in *Aedes aegypti* (Diptera: *Culicidae*) populations from Thailand. Infect Genet Evol 10: 580–585.
49. Lefèvre J, Marchesiello P, Jourdain N, Menkes C, Leroy A (2010) Weather regime and orographic circulation around New Caledonia. Marine Pollution Bulletin 61: 413–431.
50. Statistiques touristiques rapides, Nouvelle Calédonie 2009. Institut de la statistique et des études économiques de Nouvelle-Calédonie (ISEE). Available: http://www.isee.nc/tourisme/telechargement/statouristrapid2009.pdf. Accessed 18 April 2011.
51. Hay SI, Myers MF, Burke DS, Vaughn DW, Endy T, et al. (2000) Etiology of interepidemic periods of mosquito-borne disease. Proc Natl Acad Sci USA 97: 9335–9339.
52. Wearing HJ, Rohani P (2006) Ecological and immunological determinants of dengue epidemics. Proc Natl Acad Sci USA 103: 11802–11807.
53. Adams B, Holmes EC, Zhang C, Mammen MP, Jr., Nimmannitya S, et al. (2006) Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. Proc Natl Acad Sci USA 103: 14234–14239.
54. Lambrechts L, Knox TB, Wong J, Liebman KA, Albright RG, et al. (2009) Shifting priorities in vector biology to improve control of vector-borne disease. Trop Med Int Health 14: 1505–1514.
55. Chungue E, Deubel V, Cassar O, Laille M, Martin PMV (1993) Molecular epidemiology of dengue 3 viruses and genetic relatedness among dengue 3 strains isolated from patients with mild or severe form of dengue fever in French Polynesia. J Gen Virol 74: 2765–2770.
56. Descloux E, Cao-Lormeau VM, Roche C, De Lamballerie X (2009) Dengue 1 diversity and microevolution, French Polynesia 2001–2006: connection with epidemiology and clinics. PLoS Negl Trop Dis 3(8): e493.
57. Li DS, Liu W, Guigon A, Mostyn C, Grant R, et al. (2010) Rapid displacement of dengue virus type 1 by type 4, Pacific region, 2007–2009. Emerg Infect Dis 16: 123–125.
58. Morrison AC, Zielinski-Gutierrez E, Scott TW, Rosenberg R (2008) Defining challenges and proposing solutions for control of the virus vector *Aedes aegypti*. PLoS Med 5(3): e68.
59. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz Soldan V, Kochel TJ, et al. (2009) The role of human movement in the transmission of vector-borne pathogens. PLoS Negl Trop Dis 2009 3(7): e481.

60. Honório NA, Nogueira RM, Codeço CT, Carvalho MS, Cruz OG, et al. (2009) Spatial evaluation and modeling of Dengue seroprevalence and vector density in Rio de Janeiro, Brazil. PLoS Negl Trop Dis 3(11): e545.

61. Pontes RJ, Freeman J, Oliveira-Lima JW, Hodgson JC, Spielman A (2000) Vector densities that potentiate dengue outbreaks in a Brazilian city. Am J Trop Med Hyg 62: 378–383.

4.3. Analyse spatiale :

4.3.1. Modélisation non linéaire de l'intensité de la circulation virale et projections futures

4.3.1.a. Contexte et méthodologie

Nous avons vu qu'une fois une épidémie amorcée sur le territoire, la circulation virale moyenne diffère d'une commune à l'autre (Figure 4.2.A). Cette partie vise à étudier les facteurs déterminant l'hétérogénéité spatiale des niveaux d'incidence de la dengue au cours des années épidémiques enregistrées entre 1995 et 2012 (*i.e.* 1995, 1996, 1998, 2003, 2004, 2008 et 2009). Nous avons décidé dans un premier temps de réaliser une étude spatiale globale, c'est-à-dire en considérant uniquement les moyennes temporelles sur l'ensemble de la période d'étude.

L'approche méthodologique utilisée est la même que celle de l'étude des facteurs de variation temporelle du risque épidémiologique, *i.e.* utilisation de modèles non-linéaires pour lier les variables explicatives potentielles à la variable épidémiologique à étudier, et sélection des variables statistiquement associées au risque épidémiologique par la comparaison de la qualité de plusieurs modèles utilisant toutes les combinaisons possibles d'une, deux ou trois des variables explicatives.

La température jouant un rôle essentiel dans l'hétérogénéité observée, nous avons décidé d'étudier l'impact du changement climatique sur la distribution spatiale future des cas de dengue à l'horizon 2100.

4.3.1.b. Résultats et discussion

Parmi les 19 variables climatiques et les 51 variables socio-économiques étudiées, les meilleurs variables explicatives de l'hétérogénéité des taux d'incidence annuels moyens entre 1995 et 2012 en Nouvelle-Calédonie sont la température moyenne, le nombre de personnes par foyer (au sens de foyer familial), et, dans une moindre mesure, le pourcentage de personnes au chômage dans chacune des communes. Les précipitations ne semblent pas avoir d'influence particulière sur les différences de niveaux d'incidence entre les communes. Si les températures augmentent de 3 °C d'ici à 2100, comme prédit par les modèles de circulation atmosphérique, le niveau moyen d'incidence de la dengue pourrait doubler, passant d'une moyenne de 160 cas/100 000 habitants/an pour l'ensemble du territoire à 320 cas/100 000 habitants/an, doublant ainsi l'impact économique et sanitaire de la dengue, notamment au niveau de la prise en charge par les hôpitaux.

Influence de la température moyenne sur la sévérité des épidémies

L'influence de la température moyenne sur la transmission de la dengue peut être expliquée via son impact sur le moustique (allongement de la durée de vie, accélération du développement des larves…),

sur le virus (diminution de la durée d'incubation extrinsèque, augmentation de la vitesse de réplication virale) ou sur l'homme (comportement favorisant le contact hôte/vecteur).

Ce qu'il est intéressant de noter ici est que la variable climatique la plus associée à la sévérité des épidémies au sein des communes (température moyenne) est différente de celles identifiées comme favorisant l'émergence des épidémies (nombre de jours où la température maximale dépasse 32°C, et nombre de jours où l'humidité relative dépasse 95%. Ce résultat implique que les mécanismes par lesquels le climat influence l'émergence des épidémies sont à différencier des mécanismes favorisant la circulation virale une fois l'épidémie amorcée. Il montre également la difficulté à démêler les différentes composantes de l'impact du climat sur les cycles épidémiologiques des maladies infectieuses. Il serait intéressant de monter un modèle mathématique déterministe de type SIR afin de clarifier les mécanismes en cause.

Influence des variables socio-économiques sur la sévérité des épidémies

L'influence du nombre de personnes par foyer sur les niveaux d'incidence de la dengue est relativement facile à interpréter : elle peut être expliquée par la faible dispersion du moustique *Aedes aegypti* au cours de sa vie (quelques centaines de mètres au maximum autour d'une maison (Harrington et al., 2005)), et le caractère extrêmement focal des foyers épidémiques de dengue (souvent limité à quelques maisons voisines (Honório et al., 2009; Mammen Jr et al., 2008)).

Le pourcentage de personnes au chômage est une variable très corrélée aux variables socio-économiques représentatives des différents modes de vie en Nouvelle-Calédonie. Elle doit donc être interprétée comme telle. Les communes dans lesquelles un grand nombre de personnes vivent selon un mode de vie tribal, dans des habitations traditionnelles, avec un accès à l'eau potable et à l'électricité limité, et un fort taux de chômage sont en moyenne plus sévèrement touchées par la dengue que les communes où une forte proportion de personnes vivent selon un mode de vie occidental, dans des habitations en dur, avec l'air conditionné et des moyens de transport motorisés. Ceci peut-être expliqué non seulement par le fait que le mode de vie tribal favorise le contact entre l'homme et le moustique (vie dehors, etc.), mais également par le fait que les populations mélanésiennes ont une approche particulière du risque sanitaire. En effet, la maladie est considérée dans les populations kanaks comme une sorte de fatalité arrivant pour une raison précise (par exemple un différend ancien avec des personnes appartenant à d'autres tribus) et que les mesures de prévention telles que l'emploi de moustiquaires ou de répulsifs ne peuvent empêcher[1].

L'influence majeure de la composante sociologique sur les niveaux d'incidence en Nouvelle-Calédonie souligne l'importance des campagnes de prévention et de communication auprès des populations dans le contrôle de la maladie.

[1] Ces informations sont tirées d'une conférence sur la sociologie de la santé, organisée par la Société Calédonienne de Santé Publique. Conférencier : Patrice Godin.

4.3.1.c. Article 3

Les résultats de cette étude font l'objet d'un article soumis à PLoS NTD, et accepté avec révisions. Le manuscrit soumis est présenté ci-après. Les figures et tableaux se trouvent après le texte et les références.

PLOS Neglected Tropical Diseases

Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: a Worked Example in New Caledonia

--Manuscript Draft--

Manuscript Number:	PNTD-D-14-01289
Full Title:	Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: a Worked Example in New Caledonia
Short Title:	Multi-factorial Modelling of Dengue Epidemics
Article Type:	Research Article
Keywords:	Dengue fever; Spatial analysis; Territorial scale; Statistical modelling; Support Vector Machines; Climate; Climate change; Socio-economic; Health
Corresponding Author:	Magali Teurlai Institut de Recherche pour le Développement Nouméa Cédex, NEW CALEDONIA

Short title

Multi-factorial Modelling of Dengue Epidemics

Title

Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: a Worked Example in New Caledonia

Authors

Magali Teurlai[1,2,*], Christophe Eugène Menkès[2], Virgil Cavarero[3], Nicolas Degallier[4], Elodie Descloux[5], Jean-Paul Grangeon[6], Laurent Guillaumot[7], Paulo Sergio Lucio[8], Françoise Mathieu-Daudé[9], Morgan Mangeas[1]

Affiliations

[1] UMR 228, ESPACE-DEV, Institute for Research and Development (IRD), Noumea, New Caledonia

[2] UMR 182, LOCEAN, Institute for Research and Development (IRD), Noumea, New Caledonia

[3] Météo-France, Noumea, New Caledonia

[4] UMR 182, Laboratoire d'Océanographie et du Climat, Expérimentation et Approches Numériques (LOCEAN), Institute for Research and Development (IRD), Paris, France

[5] Department of Internal Medicine and Infectious Diseases, Territorial Hospital Centre, Noumea, New Caledonia

[6] Health Department, Direction of Health and Social Affairs of New Caledonia, Noumea, New Caledonia

[7] Laboratory of Medical Entomology, Pasteur Institute, Noumea, New Caledonia

[8] Centro de Ciências Exatas e da Terra (CCET), Universidade Federal do Rio Grande do Norte (UFRN), Campus Universitário - Lagoa Nova, 59078-970 - Natal – RN, Brazil

[9] UMR 224 MIVEGEC, Institute for Research and Development (IRD), Noumea, New Caledonia

*Correspondence to: magali.teurlai@ird.fr;

Abstract

Background/Objectives: Understanding the factors underlying the spatio-temporal distribution of infectious diseases provides useful information regarding their prevention and control. Dengue spatio-temporal patterns result from complex interactions between the virus, the host, and the vector. These interactions can be influenced by environmental conditions. Our objectives were to analyse dengue fever spatial distribution over New Caledonia during epidemic years, to identify some of the main underlying factors, and to predict the spatial evolution of dengue fever under changing climatic conditions, at the 2100 horizon.

Methods: We used principal component analysis and support vector machines to analyse and model the influence of climate and socio-economic variables on the mean spatial distribution of 24,272 dengue cases reported from 1995 to 2012 in thirty-three communes of New Caledonia. We then modelled and estimated the evolution of dengue incidence rates using a regional downscaling of future climate projections.

Results: The spatial distribution of dengue fever cases is highly heterogeneous. The variables most associated with this observed heterogeneity are the mean temperature, the mean number of people per premise, and the mean percentage of unemployed people, a variable highly correlated with people's way of life. Rainfall does not seem to play an important role in the spatial distribution of dengue cases during epidemics. By the end of the 21st century, if temperature increases by approximately 3 °C mean incidence rates could be multiplied by two.

Conclusion: In New Caledonia, a subtropical insular environment, both temperature and socio-economic conditions are influencing the spatial spread of dengue fever. Extension of this study to other countries worldwide should improve the knowledge about climate influence on dengue burden and about the complex interplay between different factors. This study presents a methodology that can be used as a step by step guide to model dengue spatial heterogeneity in other countries.

Author summary

Dengue is the most important viral arthropod-borne disease worldwide and its geographical expansion during the past decades has been of growing concern for scientists and public health authorities because of its heavy sanitary burden and economic impacts. In the absence of an effective vaccine, control is currently limited to vector-control measures. In this context, understanding the sociologic, entomologic

and environmental factors underlying dengue dynamics is essential and can provide public health authorities with sound information about control measures to be implemented. In this study, we analyse socio-economic, climatic and epidemiological data to understand the impact of the studied factors on the spatial distribution of dengue cases during epidemic years in New Caledonia, a French island located in the South Pacific. We identify at risk areas, and find that temperature and people's way of life are key factors determining the level of viral circulation in New Caledonia. Hence, communication campaigns fostering individual protection measures against mosquito bites could help reduce dengue burden in New Caledonia. Using projections of temperature under different scenarii of climate change, we find that dengue incidence rates could be multiplied by two by the end of the century, with areas at low risk of dengue fever now being highly affected in the future.

Introduction

Dengue fever is the most important mosquito-borne viral disease, with an estimated 50 million people being infected each year and 2.5 billion people living in areas at risk of dengue worldwide [1](WHO, 2009). The true burden of apparent dengue could be twice as high, and the total burden of dengue fever infections could reach 390 million people when including asymptomatic cases [2](Bhatt et al., 2013). Whereas only nine countries were affected by dengue epidemics in the 1970's, more than a hundred countries are now reporting dengue outbreaks on a regular basis, making dengue fever the most rapidly spreading mosquito-borne viral disease in the world [1,2](Bhatt et al., 2013; WHO, 2009). This rapid global spatial spread over the past 40 years results probably from recent socio-economic changes such as global population growth and uncontrolled urbanisation. Lack of effective mosquito control in endemic areas, increased international air traffic or decay in public health infra-structure in developing countries are also important factors that could explain the rapid regional spread of the disease[3–6](Gubler, 1998; Guzman et al., 2010; Kyle and Harris, 2008; Monath, 1994). However, in a given country where there are sufficient numbers of susceptible hosts, these factors need to be associated with suitable climate conditions before dengue fever can establish, since it is transmitted by mosquito species whose life cycle is influenced by temperature, humidity and rainfall [7–9](Focks et al., 1993a, 1995; Morin et al., 2013). Indeed, several studies have pointed out that the current geographic distribution of dengue fever or its vector worldwide could be predicted accurately based on climate variables using either statistical models [10](Hales et al., 2002) or deterministic models [11–13](Degallier et al., 2010; Hopp and Foley, 2003; Jetten and Focks, 1997). Other studies have pointed out that climate change could have profound consequences on the epidemiology of dengue fever, because increased temperature and rainfall could facilitate viral transmission and could lead to the geographic expansion of the mosquito species responsible for its transmission [11,14–17](Githeko et al., 2000; Jetten and Focks, 1997; McMichael et al., 2006; Naish et al., 2014; Patz et al., 1998).

The complex interplay and relative importance of each factor in the occurrence and spread of dengue fever epidemics might differ from one country to another, depending on the specific climate conditions, cultural and socio-economic environment the virus circulates in [18](Reiter, 2001). Identifying the factors limiting dengue fever spatial spread at a national level could help understand the worldwide pattern of dengue disease, could help predict its future spatial distribution, and could provide national decisions-makers with useful information regarding the appropriate control measures to be implemented.

Most studies trying to identify dengue risk factors spatially were performed at a city scale or a local scale (< 80 km) [19–32](Aziz et al., 2012; Bohra and Andrianasolo, 2001; Costa et al., 2013; Estallo et al., 2013; Honório et al., 2009; Khormi et al., 2011; Nakhapakorn and Tripathi, 2005; Penna and Lucia, 2004; Reiter et al., 2003; Sarfraz et al., 2012; Seidahmed et al., 2012; Spiegel et al., 2007; Stoddard et al., 2013; Thammapalo et al., 2005). Among these studies, some have identified risk factors for the presence of *Aedes* mosquito species, such as socio-economic factors [29,30](Seidahmed et al., 2012; Spiegel et al., 2007), proximity to specific plantations [28,32](Sarfraz et al., 2012; Thammapalo et al., 2005), proximity of potential breeding sites [22,28,29,32](Estallo et al., 2013; Sarfraz et al., 2012; Seidahmed et al., 2012; Thammapalo et al., 2005), or human behaviour [30,32](Spiegel et al., 2007; Thammapalo et al., 2005). Some studies highlighted the importance of human movement [23,26,31,33](Honório et al., 2009; Liebman et al., 2012; Penna and Lucia, 2004; Stoddard et al., 2013) or population immunity [34](Salje et al., 2012) in shaping the spatial transmission of dengue fever at small spatial scales. High dengue incidence rates have also been found in neighbourhoods with low social income [20,26](Costa et al., 2013; Penna and Lucia, 2004), difficult access to piped water [21,26](Bohra and Andrianasolo, 2001; Penna and Lucia, 2004), or no implementation of mosquito protection measures [21,27](Bohra and Andrianasolo, 2001; Reiter et al., 2003).

Spatial analyses at a country or territorial scale (> 200 km) are scarce. Some of these studies focused on the spatio-temporal dynamics of the disease only [35–37](Cummings et al., 2004; Teurlai et al., 2012; Cuong et al., 2013) and proposed hypothesis about the underlying processes, but did not include analyses of potential explicative factors. To our knowledge, there are only five studies to date identifying and quantifying spatial risk factors for dengue at a "national" scale > ~200 km and < ~1000 km. Four studies identified temperature as having a major influence on the spatial distribution of locally acquired dengue cases [38–41](Carbajo et al., 2012; Hu et al., 2011; Lowe et al., 2011; Wu et al., 2009), the last one did not assess the role of climate factors [42](Tipayamongkholgul and Lisakulruk, 2011). The role of other factors in the spatial distribution of dengue cases varied from one place to another. For example, in Australia (Queensland) [38](Hu et al., 2011) and Taiwan [40](Wu et al., 2009), rainfall seemed to play a minor role whereas in Brazil, rainfall played a major role [39](Lowe et al., 2011). In Taiwan [40](Wu et al., 2009) and Argentina [41](Carbajo et al., 2012), urbanization level was a key factor in dengue fever spatial distribution, and in one province of Thailand [42](Tipayamongkholgul and Lisakulruk, 2011), the main factor identified was the proximity to major urban centres. No association

was found with socio-economic covariates in Argentina [41](Carbajo et al., 2012) or Australia [38](Hu et al., 2011).

New Caledonia, where the present study takes place, has a unique situation: it is a developed insular territory located in the inter-tropical area of the South Pacific where the access to high quality data and the lack of terrestrial borders with other countries make it a natural laboratory to study dengue dynamics. A gross average of ten imported cases is detected each year by the Public Health authorities. However, large dengue epidemics develop only every three to five years [43](Descloux et al., 2012). A recent study analysing the temporal relationship between dengue epidemic occurrence and climate variables at an inter-annual scale showed that the development of an epidemic in New Caledonia needs precise climate conditions relying on both temperature and relative humidity [43](Descloux et al., 2012).

The objectives of the present study were i) to characterise the spatial distribution of dengue cases in New Caledonia once an epidemic spreads over the territory; ii) to determine which of possible covariates are shaping the observed distribution; iii) to explore the potential spatial distribution of dengue cases under future climate projections.

We present a complete methodology, from data collection, data transformation, variable selection, and application to future climate projections. We address a number of methodological issues such as spatial autocorrelation, correlation between explanatory variables, or potential non-linearity between epidemiological data and explanatory covariates.

Material and methods

All analyses and figures were performed using R software [44](R Core Team, 2014).

1. Study area

New Caledonia is a French territory, located in the Pacific Ocean 1,500 km East of Australia. It is divided into 33 communes covering 18,576 km². Out of the 245,344 inhabitants (2009), around 58% (147,365 people) live in Noumea, the main city, and its surroundings. The rest of the population is scattered in small towns of about 2,000 people, or live in rural areas, including traditional Melanesian settlements locally called “tribes” (Figure 4.7). Although the average population density outside Noumea is very low (5.3 inhabitants per km²), local densities can be high as people gather in small settlements.

New Caledonia is located at the limit of the tropical zone between latitudes 19° and 23° South. The East coast and the West coast are separated by a mountain range culminating at 1629 m. Climate is heterogeneous: the East coast and the southern tip of the main island get more rain than the West coast,

as mountains provide a vertical lift to the warm and humid air brought by trade winds. Average rainfalls range from 800 mm/year in some western weather stations to 3200 mm/year in the East. Besides, temperature can drop below 10 °C during cool season on clear nights and sometimes rise above 35 °C due to the influence of tropical air masses [45](Météo France, 2014).

From an oversimplified point of view, there are three population groups, having different cultural and social habits: Melanesian people, people of French ascent who migrated two hundred years ago, and people from various origins who migrated recently. Although the three groups are spatially partially mixed, Melanesian people live mostly on the East coast, whereas the second group live mostly on the West coast and the third group live mainly in Noumea.

In New Caledonia, dengue represents a major public health problem with large epidemics affecting the territory every three to five years and involving a succession of all four serotypes [43,46–48](A-Nuegoonpipat et al., 2004; Bouldouyre et al., 2006; Descloux et al., 2012; Li et al., 2010). Co-circulation of different dengue virus serotypes (DENV1-4) during major epidemics is rare, and has been observed only once (2009). Vector control measures include continuous large communication and prevention campaigns fostering source reduction aimed at all citizens, as well as chemical control of adult mosquitoes around declared cases. Public Health infrastructure is reliable, and the surveillance system for dengue fever has been efficient for many years. All people have access to medical care, even though people living in remote areas might have more difficulties to reach local health centres.

2. Data

2.1 Epidemiological data

We studied 24,272 cases of dengue fever reported to the Direction of Sanitary and Social Affairs (DASS-NC) from 1st January 1995 to 30th September 2012, a period over which the spatial location of cases was recorded at the commune level. Cases were recorded on a clinical basis, but more than 71% of declared cases have also been laboratory confirmed either by virus isolation, viral RNA detection (RT-PCR) and NS1 antigen detection (ELISA), or by detection of IgM (ELISA). During the study period, changes in the surveillance system have been homogeneous over the country, making spatial analysis of mean incidence rates possible.

For each of the 33 communes composing New Caledonia, we calculated mean yearly age-standardised incidence rates over epidemic years (as defined by Descloux et al. [43](Descloux et al., 2012)) during the study period, i.e. 1995, 1996, 1998, 2003, 2004, 2008 and 2009. As the majority of outbreaks display a similar seasonal evolution (beginning in January, epidemic peak between March and May, and end in July) and according to the climate seasonality in New Caledonia, we aggregated data annually from 1st September of year n-1 to 31st August of year n before calculating the mean. Average incidence

rates over epidemic years of the study period reflect mean viral circulation in a commune once an epidemic strikes the territory.

Age-standardisation was necessary as the age structure of the population varies from one commune to another, and age groups are affected differently by dengue epidemics (see Supplementary Figure 4.14). We performed age-standardisation using a direct method [49](Ahmad et al., 2001), which consists in calculating the expected incidence rate in each commune if all communes had the same (arbitrary) age-structure. We chose the World Health Organisation standardised population as the reference population for the arbitrary age-structure [49](Ahmad et al., 2001). Population numbers and age-structure of each commune used to calculate age-specific incidence rates were interpolated linearly from the 1989, 1996, 2004 and 2009 territorial censuses [50–53](Institut de la statistique et des études économiques (ISEE), 1989, 1996, 2004, 2009).

Average incidence rates over the study period were artificially low in 4 of the islands surrounding New Caledonia (Mare, Lifou, Ouvéa, Isle of Pines) as the only vector of dengue in New Caledonia, *Aedes aegypti*, was not present over the entire study period [54–56](IPNC, 1996, 1999, 2003). We thus decided to limit the analysis to New Caledonia's mainland, comprising 28 communes.

2.2 Socio-economic covariates

From the 2009 territorial census [53](Institut de la statistique et des études économiques (ISEE), 2009), recorded at the commune level, we selected a total of 51 variables that could potentially influence dengue transmission. These variables can broadly be categorised in sociologic variables reflecting people's activity and financial means, variables linked to human movement within or between communes, variables linked to people's housing, and variables linked to human local density (see Table 4.2 for a complete list of these variables). Because human density per commune is very low and do not reflect local densities, we decided not to include it in the analysis.

2.3 Climate covariates

2.3.1 Observed data and upscaling to the communal scale

Observed rainfall (RR), maximal temperature (maxT) and minimal temperature (minT) were provided by Meteo France on a daily basis from 1995 to 2012 in 118 weather stations scattered all over New Caledonia. We first selected data from the same time periods as the epidemiological data (September of epidemic year *n-1* to August of epidemic year *n*). To handle missing data, only time series covering at least a full continuous year were kept resulting in 67 locations for temperature, and 113 locations for rainfall (see Figure 4.7). For each variable, at each weather station, we then calculated a typical year

time-series, 365 days long, and calculated the mean value for the climate variable as the average of this time series, hereafter called "climate index".

To model the link between climate and incidence rates by means of regression models, we had to project the punctual station-based climate data to the same spatial level as the epidemiological data (i.e. the commune level). In order to take into account the fact that incidence rates are calculated from individual data, and the fact that the population is distributed heterogeneously within each commune, we assigned every tribe the climate index value of the closest station, assuming all people within a tribe to be exposed to the same climate conditions. We then averaged the climate index values over all tribes in the same commune, repeated the process over all towns of the same commune, and then calculated an average of the tribe and town values, weighted by the respective proportion of people living in either place, thus obtaining a mean climate index value per commune (see Figure 4.8). The distance between a tribe or a town and the closest station ranges from 0 km to 27 km, with a mean of 9 km (see Figure 4.7). This "projection" process was done for rainfall, minimal and maximal temperature, and for 16 other climate indices built from these three raw variables (see Table 4.2 for a complete list). The 16 climate indices we built included the key variables most explicative of the inter-annual variability of dengue epidemics in New Caledonia, i.e. the number of days when maximal temperature exceeds 32°C during January/February/March [43](Descloux et al., 2012), and the number of days when precipitations exceeds 2 mm in January [57](Leroy, 2011). They also included variables known to have an influence on biological processes (see all other climate indices in table 1) [58](Hijmans et al., 2005). All observed climate variables and those built from these observed variables are referred to as "observed data" in the article.

2.3.2 Assessing the trends of future mean temperature in New Caledonia

To obtain maps of mean temperature under different global warming scenarii, we first retrieved historical (1971-2005) and future (2006-2100) projections of mean temperature simulated by six coupled ocean-atmosphere models from the 5th Phase of the Coupled Model Intercomparison Project – Assessment Report 4 (CMIP5-AR4) experiments [59,60](Anonymous; Taylor et al., 2012). The six selected models are "CanESM2", "CNRM-CM5", "IPSL-CM5A-MR", "IPSL-CM5B-LR", "MPI-ESM-LR", and "NorESM1-M". They were selected based on their capacity to reproduce a reasonable climate in the South Pacific [61]. Two scenarii of emission (greenhouse gases and aerosols) - referred to as "Representative Concentration Pathways" (RCPs) – were chosen: RCP8.5 refers to a high emission scenario while RCP4.5 refers to a midrange mitigation emission scenario. In this article, these data are respectively denoted "historical" (1971-2005), "RCP4.5" and RCP8.5" (2006-2100).

Model simulations are available at different spatial locations depending on the spatial grid used by each model. For each model, we selected time-series from the spatial point closest to Noumea. We first assessed the necessity of performing a statistical downscaling to correct the retrieved time series so they

fit the observed temperature time series [62](Cavarero et al., 2012) but concluded that correction was not necessary for estimating average increases in mean temperature in New Caledonia. For the periods 2011-2030, 2041-2060 and 2081-2100, we calculated the increase of mean temperature averaged over the 20-year periods compared to the 1981-2000 historical data average (see Supplementary Figure 4.15). By comparing this mean increase to the mean increase in time series taken from other close grid points around New Caledonia, we concluded that this mean increase could be considered homogeneous over the whole territory. As all 6 selected models predictions were coherent (see Supplementary Figure 4.16), we averaged the mean increases over all models. By applying this mean increase on current temperature maps, we obtained temperature maps for the periods 2011-2030, 2041-2060 and 2081-2100.

3. Multivariate modelling of present dengue incidence rates

3.1 Spatial autocorrelation of the response variable

We assessed spatial autocorrelation (SAC) of mean yearly dengue incidence rates by comparing the semi-variogram of observed data to 200 semi-variograms obtained by random permutations (i.e. without replacement) of the spatial locations of each commune centroid [63](Bivand et al., 2013). Distances between pairs of commune centroids were calculated along the roads to take into account the potential influence of the central mountainous chain in the spatial spread of the disease.

3.2 Pre-selection of explanatory variables

To reduce the number of variables to study, the number of models to build and to facilitate interpretation, we first used principal component analysis (PCA) [64](Escoffier and Pagès, 2008) separately on climate and socio-economic variables to identify clusters of correlated explanatory variables. In each cluster, we selected the variable most correlated with dengue incidence rates as representative of all other variables in the cluster. We were then left with a set of three climate variables and two socio-economic variables as inputs for modelling the incidence rates.

3.3 Modelling the spatial association between explanatory variables and dengue incidence rates

To overcome several methodological issues such as multi-collinearity between explanatory variables [65,66](Dormann et al., 2013; Scherrer, 2008) or the possible existence of complex non-linear links between the explanatory and the response variables, we chose to use Support Vector Machines (SVM) to model the link between the five selected explanatory variables and mean yearly dengue incidence rates. SVM is a non-parametric supervised learning algorithm that can be used in high dimensional multivariate

classification or regression problems [67,68](Cortes and Vapnik, 1995; Vapnik, 1998). It can be used to model non-linear links without any a priori knowledge of the shape of the link [69](Ben-Hur and Weston, 2010).

To classify the selected explanatory variables according to their importance in shaping dengue spatial distribution, we compared the overall performance of models based on all possible combinations of one, two or three variables. For each possible combination of explanatory variables, we optimised the SVM parameters combination by building several models using a range of different values for the parameters and selecting the best parameters combination by minimising the models root mean squared errors (RMSE), calculated using a 10-fold cross validation procedure [70](Kohavi, 1995). The influence of each explanatory variable on the spatial distribution of dengue cases was then assessed by comparing the RMSE of all optimised models.

4. Specific model for exploring dengue spatial trends in the future under changing climatic conditions

As described below in the results section, temperature is a key factor determining dengue spatial variability over New Caledonia. We thus decided to explore the evolution of dengue incidence rates under changing climate conditions (considering all others variables as remaining constant), by applying the best explicative multivariate model with inputs from maps of temperature for the future (see methods/data/climate covariates). Because the use of kernels in non–linear SVM models impairs predictions outside the observed range of explanatory variables, we built a linear approximation of the best SVM model on present observed data. The linear approximation consists of a simple linear model linking the two best explanatory variables to observed dengue age-standardised incidence rates as the response variable. Normality and homoscedasticity of residuals were confirmed by the Shapiro-Wilks' test and the Bartlett's test respectively [66](Scherrer, 2008).

Results

Spatial distribution of dengue cases

Figure 4.9 shows that once an epidemic strikes the territory, dengue cases are distributed heterogeneously. Mean annual age-standardised incidence rates during epidemic years range from 22 to 375 cases per 10,000 people per year, with a mean of 168 cases per 10,000 people per year and a standard deviation of 83 cases per 10,000 people per year. On average the East coast is more affected than the West coast, as can be seen on Figure 4.9.B. When communes are sorted by increasing incidence rates (Figure 4.9.C), we see that dengue incidence rates increase regularly from one commune

to another, except from a sharp increase in three communes located in the North-East where dengue incidence rates are 2 to 3 times higher than the territorial mean.

Autocorrelation of the response variable

The semi-variogram of dengue incidence rates did not reveal any significant spatial autocorrelation. This suggests that the local spread of dengue viruses around a case imported in a commune do not exceed the mean radius of a commune in New Caledonia, e.g. approximately 13 kilometres.

Pre-selection of explanatory variables

Table 4.2 shows Pearson's correlation coefficients (rho) between each explanatory variable and dengue age-standardised incidence rates. Dengue is positively correlated with variables related to temperature and precipitation, but is negatively correlated with variables reflecting mean thermal range or extreme thermal conditions (see "isothermality", "Trange" or "nb_D_maxT_32_JFM" in Table 1). This suggests that, in a given commune, marked temporal variations of temperature is a factor limiting viral circulation. Based on the linear dependence measure of correlation, dengue is also more strongly associated with temperature than with precipitations. Socio-economic variables are highly spatially correlated to dengue incidence rates. Variables reflecting people's way of life (e.g. place of birth), local human density (e.g. mean number of people per household, percentage of premises under 40 m²), or human movement are more correlated with dengue incidence rates than variables related to the housing type (e.g. premises with inside toilets) (absolute value of rho up to 0.75 for the former and 0.58 for the latter). In particular, the fact that the place where people were born is significantly associated with dengue fever incidence rates (correlation coefficient around 0.5 for people born in New Caledonia and –0.5 for people born elsewhere) whereas the type of premise is not (absolute correlation coefficient lower than 0.3 for variables describing access to water or electricity) suggests that individual behaviours have a stronger influence on incidence rates than local housing conditions.

Figure 4.10 shows the PCA results. For clarity reasons, we only show the results of PCA performed on the variables most correlated with dengue incidence rates, with an absolute Pearson correlation coefficient over 0.6 for socio-economic variables, and over 0.4 for climate variables (these thresholds were selected after verifying that they did not modify the variable pre-selection results).

PCA of climate variables (Figure 4.10.A) shows that in New Caledonia, temperature is the factor accounting for most of the spatial climatic variability among communes. Temperature is highly correlated with the first PCA axis which represents 68% of the total climatic variance Temperature and rainfall are not spatially correlated at the commune level. In each group of temperature or rainfall variables, the variables most spatially correlated with dengue incidence rates were the average mean

temperature (meanT) and the mean daily rainfall during the wettest quarter of the year (wettest_quarter) (see Table 4.2). In addition to these two variables, we decided to keep a third variable, the average daily rainfall (RR), for further statistical modelling, as this variable is more easily available in other countries or model simulations.

Figure 4.10.B shows that the spatial variability of socio-economic factors mainly reflects the spatial distribution of people with different cultural habits. Communes where a high proportion of inhabitants live in a tribal way, in small premises, with few means of transportation and a high percentage of unemployment are opposed to communes where many people live a western way of life, in permanent buildings, using air conditioning and getting around using cars. Even though the number of people per premise seems to be slightly correlated with the proportion of people living in tribes, we kept this variable as it stands out of the cluster of variables representing the way of life. We thus decided to keep the percentage of unemployed people and the mean number of inhabitants per housing as representative of socio-economic factors for further statistical modelling.

Modelling the spatial association between explanatory variables and dengue incidence rates

Table 4.3 shows the RMSE of the optimised models built on all possible combinations of one, two or three of the five selected explanatory variables (meanT, wettest_quarter, RR, nb_per_household and act_unemployed). When looking at bivariate non-linear SVM models, the best variable explaining the spatial heterogeneity of dengue incidence rates is the percentage of unemployed people per commune. The second most important explanatory variable is the mean temperature. Rainfall is the least explanatory variable. Moreover the RMSE of models based on observed rainfall almost equal the initial standard deviation of dengue incidence rates, which means that rainfalls are poor predictors of dengue incidence rates. The relationship between dengue incidence rates and each of the explanatory variables is linear, except for the fraction of unemployed people (Supplementary Figure 4.17). When looking at the spatial structure of dengue incidence rates predicted by SVM models based only on one of the selected variables (Supplementary Figure 4.18), we see that temperature captures mainly the South to North gradient of increasing incidence rates (Supplementary Figure 4.18.B) whereas socio-economic variables captures the spatial heterogeneity between the West coast and the East coast (Supplementary Figure 4.18.E and Supplementary Figure 4.18.F). Temperature seems to have no influence in communes located below 21°S (Supplementary Figure 4.18.B).

All models based on two explanatory variables and including at least one variable related to rainfall (best RMSE of ~58 cases per 10,000 people per year) performed worse than the best bivariate model (RMSE of ~53 cases per 10,000 people per year). This suggests that in New Caledonia, rainfall has little influence on the spatial variability of dengue viral circulation at the commune level. Models combining two explanatory variables (excluding rainfalls) performed better than bivariate models. The addition of a

third explanatory variable did not improve significantly model performances. Hence we focused our attention on models combining two explanatory variables.

The best explicative model is a model predicting increasing incidence rates as the mean temperature and the mean number of people per premise increase (see Figure 4.11). The influence of these two variables on dengue incidence rates is close to linear as shown by almost parallel contour lines on Figure 4.11.A. This model accurately predicts the sharp mean increase in incidence in the three communes of the North East of New Caledonia (Hienghène, Ouégoa and Pouébo). The maximal error of the model is observed for Farino (West coast), which is the only commune where all inhabitants live at an altitude higher than 200 m above sea level.

Figure 4.12 shows the spatial structure of incidence rates as predicted by the best SVM model based on two explanatory variables (Figure 4.12.B). This model captures the observed spatial heterogeneity in incidence rates between the East coast and the West coast, as well as the sharp increase in the three communes of the North East.

Future trends in the spatial distribution of dengue cases under changing climatic conditions

Table 4.4 shows, for both climate change scenarii, the average increase of mean temperature for three 20-year periods compared to the 1981-2000 historical simulations. All models (see Supplementary Figure 4.16 for the mean increase predicted by each model) predict that the mean temperature will increase over time, with projections being more pessimistic for RCP8.5 simulations. Temperature could increase by more than 3 °C by the end of the next century.

Figure 4.12.C shows that the linear model based on the two best explanatory variables is suitable. This is confirmed by the low RMSE of the linear model (~45 cases per 10,000 people per year), which is almost equal to the one of the SVM model based on the same variables. The Shapiro-Wilks and the Bartlett's test confirmed the normality and homoscedasticity of residuals. According to RCP8.5 climate projections, the average dengue incidence rates over the territory could rise by 34 cases per 10,000 people per year in the near future, and by 164 cases per 10,000 people per year by the end of the century, multiplying global dengue burden by two in New Caledonia (Table 4.4).

Figure 4.13 shows the spatial distribution of potential future trends of dengue incidence rates. By the end of the century, dengue incidence rates during epidemic years could reach a maximum of 383 cases per 10,000 people per year under the RCP4.5 scenario (Figure 4.13.C), and 470 cases per 10,000 people per year under the RCP8.5 scenario (Figure 4.13.D). The increase in dengue burden per commune relative to the projected 1981-2100 incidence rates might be higher in communes that are currently at low risk of dengue fever compared to the communes already at high risk now (Figure 4.13.B). Under RCP8.5 scenario, communes at low risk now will have incidence rates during epidemic

years higher than 160 cases per 10,000 people per year, the current territorial mean annual incidence rates.

Discussion

Association between climate factors and dengue spatio-temporal dynamics in New Caledonia

The association found between temperature and dengue incidence rates in New Caledonia can be explained by the influence of temperature on the life cycle of the mosquito transmitting the virus in New Caledonia, *Aedes aegypti*. High temperatures increase the productivity of the breeding sites through an acceleration of the metabolism of the mosquito themselves, and a faster development of the micro-organisms the larvae feed on, resulting in a higher vector density even with the same number of breeding sites [7,71–73](Couret et al., 2014; Focks et al., 1993a, 2000; Mohammed and Chadee, 2011). High temperatures also speed up the extrinsic incubation period [7,74,75](Dégallier, 2011; Focks et al., 1993a; Watts et al., 1987), with the effect that an increased proportion of females *Ae. aegypti* can reach the infectious stage before dying. Finally, warmer temperatures accelerate the mosquito gonotrophic cycle, and make females *Ae. aegypti* more aggressive [7,72,76–78](Focks et al., 1993a, 2000; Hopp and Foley, 2001; Scott et al., 2000; Yang et al., 2009), increasing the biting rate and potential transmission of viral particles to susceptible hosts.

In Noumea, the main city, precise climate variables and important thresholds values have been identified as necessary conditions to trigger an epidemic (e.g. number of days when maximal temperature exceeds 32°C in January/February/March, and number of days with maximal relative humidity exceeds 95% during January [43](Descloux et al., 2012). At the scale of the entire territory, we found that the spatial distribution of dengue cases during epidemic years is strongly influenced by the average mean temperature. These results suggest that temperature has a major role in dengue dynamics in an insular territory characterised by climate seasonality. However, we did not find a strong association between the spatial distribution of dengue cases during epidemics and average rainfalls or with the number of days when maximal temperature exceeds 32 °C. The variables influencing either the triggering of an epidemic [43](Descloux et al., 2012) or its spatial distribution are not the same. This suggests that the underlying processes differ somehow. Our findings highlight the complexity of studying and understanding dengue dynamics, the importance of well separating the two epidemiological processes of epidemic triggering in a susceptible population, and its intensity once it has started by clearly defining the modelling target (incidence rates for epidemic intensity, or dummy variables for epidemic triggering), and the importance of well defining the scale of study (temporal evolution, or spatial distribution).

The positive association found between the mean temperature and dengue incidence rates is consistent with the one found in previous studies having analysed the spatial distribution of dengue cases at spatial scale > 200 km [38–42](Carbajo et al., 2012; Hu et al., 2011; Lowe et al., 2011; Tipayamongkholgul and Lisakulruk, 2011; Wu et al., 2009). In these studies as well as in ours, all regions were located between 10° and 25° of latitude, at the fringe of the tropical area, except Argentina, where the region studied extends to 35° South. In the 10° - 25° latitudinal band, annual mean temperature lies in a range of temperature where the life cycle of the mosquito is very sensitive to temperature changes [7](Focks et al., 1993a).

Some *Aedes* species, including *Ae. aegypti*, are able to breed in very small amount of water, e.g. snails' shells. Rainfalls can play a role in dengue transmission cycle by filling up potential breeding sites [7](Focks et al., 1993a), thus influencing the vector density. Rainfalls also increase the relative humidity, which extends the mosquitoes' lifespan and therefore the likelihood of those who had an infectious blood meal to reach the infectious stage. However, our study suggests that in New Caledonia, there is no strong association between rainfalls and the spatial distribution of cases during epidemics. A plausible explanation can be the multi-factorial nature of dengue fever, and the relative influence each factors plays on dengue dynamics: despite suitable rainfall conditions, dengue might not circulate well if other factors are limiting dengue viral circulation, such as some human behaviour influencing the contact between vector and host. This aspect has been highlighted very clearly in the United States [27](Reiter et al., 2003). Some studies have found that the effect of rainfall on vector density can be modulated by human activities such as water storage practices [79](Stewart Ibarra et al., 2013). However, in New Caledonia, we are not aware of specific practices to store water that could explain the lack of association between rainfall and virus circulation intensity. Another potential explanation could be that in dry areas, breeding sites are filled up by other non-climatic mechanisms, such as automatic irrigation or plant watering.

Worldwide, the association between rainfalls and the spatial distribution of dengue cases at a "national" scale (> 200 km) is not as clear as the one for temperature: one spatial study did not find any association between dengue incidence rates and rainfall [40](Wu et al., 2009), whereas two others did [38,39](Hu et al., 2011; Lowe et al., 2011). Other factors influencing dengue transmission (e.g. anthropic factors influencing the availability of filled breeding sites) and not included in the different studies might blur the rainfall signal.

It would be interesting to perform the same kind of multi-factorial spatial analysis in areas of epidemic or endemic transmission located closer to the equator, where the mean temperature is higher, to see what climatic factors impact the spatial distribution of cases. For example in French Polynesia, where temperature is warmer than in New Caledonia in average, temperature does not seem to play a role in dengue spatial variability (unpublished results). This kind of study could help understand better the complex interplay between the different factors (climate, socio-economic, immunologic, viral, entomologic…) associated with dengue fever transmission.

Association between socio-economic factors and the spatial distribution of dengue cases in New Caledonia

Our results are consistent with previous studies that have pointed out the importance of socio-economic factors on the spatial distribution of dengue cases, whatever the spatial scale studied: national (>200 km) [39–42](Carbajo et al., 2012; Lowe et al., 2011; Tipayamongkholgul and Lisakulruk, 2011; Wu et al., 2009) or local (<10 km) [20,26,29,80](Costa et al., 2013; Degallier et al., 2000; Penna and Lucia, 2004; Seidahmed et al., 2012). The association between the percentage of unemployed people and dengue in New Caledonia cannot be interpreted in terms of lack of economic activity only, as shown by the PCA on socio-economic factors. This variable we selected as input for the models is indeed highly correlated with other variables reflecting the way of life, socio-economic and cultural differences existing in New Caledonia, which are in turn highly correlated to housing type. Moreover, a part of Melanesian people categorized as unemployed by the territorial census actually have a family-sized cropping activity. Therefore, at this spatial scale in New Caledonia, it is difficult to statistically differentiate the role played by human behavior, human activity or housing type in dengue fever transmission. However those three factors influence the contact rate between viraemic patients or susceptible hosts on one hand, and mosquitoes on the other hand. This highlights the importance of limiting the contact between humans and vectors and should lead local authorities to strengthen communication campaigns about personal protection measures towards populations at risk.

Influence of global warming on future dengue incidence rates in New Caledonia

The results about climate change must be interpreted keeping in mind that they represent a climate risk only, and that the spatial association between dengue incidence rates and temperature might change over time depending on socio-demographic changes, or changes in dengue control strategy. Assuming all other factors remain constant in time, our results suggest that Public Health authorities can expect the dengue burden to raise significantly during the next century, and can expect the dengue spatial range to increase. The impact of climate change will be relatively higher in communes at low risk of dengue fever now, such as areas located in altitude. Indeed, with increasing temperature, *Ae. aegypti* is expected to be able to breed in areas that are currently free of the mosquito [11](Jetten and Focks, 1997).

Methodological issues

The analysis of the spatial pattern of infectious diseases, in relation with environmental or socio-economic factors raises a number of methodological issues, such as the presence of spatial auto-correlation, the spatial scale of aggregation of the data, the existence of possible non-linear links

between the response and the explanatory variables, or the presence of multi-collinearity between the response variables. Most issues have already been addressed in the past, and solutions already exist to handle them. For example, the review by Dormann et al. deals with multi-collinearity [65](Dormann et al., 2013) or spatial-autocorrelation [81](Dormann et al., 2007). The main issue we have been confronted with in our study was the spatial upscaling of meteorological data observed at precise locations to the same spatial level of aggregation as the epidemiological data. In existing spatial studies of dengue fever at a national scale, authors have geo-spatially interpolated climate variables on regular grids using kriging methods, and have averaged gridded values over a given administrative division [38–41](Carbajo et al., 2012; Hu et al., 2011; Lowe et al., 2011; Wu et al., 2009). This approach has two drawbacks in New Caledonia. Simple kriging models do not take into account the potential elevation between two given points, leading to biased estimates of temperature in mountainous regions. Moreover, the traditional approach used in climatology, which consist in aggregating temperature over grid points taken uniformly over the whole aggregative area makes the implicit assumptions that people at risk are distributed homogeneously over the aggregative area. This is particularly problematic in New Caledonia where large areas are not inhabited. In our approach, as epidemiological data are collected at the individual level, we tried to estimate the climate conditions for each individual (and therefore for the mosquitoes surrounding each individual). However, the algorithm used introduces some noise, due to the fact that the weather stations are sometimes kilometres away from some towns or tribes. High spatial resolution climate data obtained from high resolution modelling of atmospheric conditions could be used, but some noise will be introduced by the modelling error compared to the observed data. This issue needs further attention in the future to increase the quality of spatial epidemiological and environmental studies.

Factors not taken into account and perspectives

Some factors that could influence the spatial distribution of dengue cases have not been taken into account in this study: the location of the first cases introduced each year, the spatial variability in population immunity, viral factors such as the serotype circulating, or factors associated to the mosquito such as the spatial variability in vector competence, or dengue vector control measures. We decided not to include the serotype, as we performed the analysis on averages over several years, and as, except in 2009, there was no co-circulation of different serotypes. Therefore, spatial differences in the level of viral circulation cannot be associated with genetic differences between serotypes for each year a new serotype was introduced but past seroprevalence data are not available. A territorial seroprevalence survey to assess serotype specific population immunity has been implemented recently in New Caledonia, but data are not available yet. Spatial variations in *Ae. aegypti* vector competence could potentially influence the spatial variability of dengue virus circulation observed during epidemic years. However, such information is not known in New Caledonia and could be interesting to stutdy. It could be interesting to include environmental variables derived from GIS data or remote-sensing in this kind of

study. The amount of vector control effort implemented in each commune is heterogeneous on the territory, as this activity falls within the commune's authority, and each commune is free to implement or not the territorial guidelines. The vector control effort in each commune is thus difficult to quantify, and data were not available yet at the time of analysis. As soon as these data will be collected by local authorities, they could be incorporated in the modelling process to assess the efficiency of vector control measures.

Acknowledgments

We would like thank Cyril Dutheil for his help with the analysis of CMIP5 climate projections, Anne Pfannstiel (DASS-NC) for useful discussion about the dengue notification system and vector control measures, Ann-Claire Gourinat (Pasteur Institute) for information regarding the biological diagnosis of dengue fever performed at the Pasteur Institute, and last but not least, we would like to thank all the practitioners that have been reporting dengue cases for many years.

References

1. WHO (2009) Dengue guidelines for diagnosis, treatment, prevention, and control. Geneva: TDR : World Health Organization. Available: http://site.ebrary.com/id/10363988. Accessed 19 June 2014.

2. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, et al. (2013) The global distribution and burden of dengue. Nature 496: 504–507. doi:10.1038/nature12060.

3. Gubler DJ (1998) Dengue and dengue hemorrhagic fever. Clin Microbiol Rev 11: 480–496.

4. Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, et al. (2010) Dengue: a continuing global threat. Nat Rev Microbiol 8: S7–S16. doi:10.1038/nrmicro2460.

5. Kyle JL, Harris E (2008) Global Spread and Persistence of Dengue. Annu Rev Microbiol 62: 71–92. doi:10.1146/annurev.micro.62.081307.163005.

6. Monath TP (1994) Dengue: the risk to developed and developing countries. Proc Natl Acad Sci 91: 2395–2400.

7. Focks DA, Haile DG, Daniels E, Mount GA (1993) Dynamic life table model for Aedes aegypti (Diptera: Culicidae): analysis of the literature and model development. J Med Entomol 30(6): 1003–1017.

8. Focks DA, Daniels E, Haile DG, Keesling JE (1995) A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. Am J Trop Med Hyg 53 (5): 489–506.

9. Morin CW, Comrie AC, Ernst K (2013) Climate and Dengue Transmission: Evidence and Implications. Environ Health Perspect 121: 1264–1272. doi:10.1289/ehp.1306556.

10. Hales S, Wet N d., Maindonald J, Woodward A (2002) Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. The Lancet 360: 830–834.

11. Jetten TH, Focks DA (1997) Potential changes in the distribution of dengue transmission under climate warming. Am J Trop Med Hyg 57 (3): 285–297.

12. Degallier N, Favier C, Menkes CE, Lengaigne M, Ramalho WM, et al. (2010) Toward an early warning system for dengue prevention: modeling climate impact on dengue transmission. Clim Change 98: 581–592.

13. Hopp MJ, Foley JA (2003) Worldwide fluctuations in dengue fever cases related to climate variability. Clim Res 25: 85–94.

14. McMichael AJ, Woodruff RE, Hales S (2006) Climate change and human health: present and future risks. The Lancet 367: 859–869. doi:10.1016/S0140-6736(06)68079-3.

15. Githeko AK, Lindsay SW, Confalonieri UE, Patz JA (2000) Climate change and vector-borne diseases: a regional analysis. Bull World Health Organ 78: 1136–1147.

16. Naish S, Dale P, Mackenzie JS, McBride J, Mengersen K, et al. (2014) Climate change and dengue: a critical and systematic review of quantitative modelling approaches. BMC Infect Dis 14: 167. doi:10.1186/1471-2334-14-167.

17. Patz JA, Martens WJ, Focks DA, Jetten TH (1998) Dengue fever epidemic potential as projected by general circulation models of global climate change. Environ Health Perspect 106: 147–153.

18. Reiter P (2001) Climate change and mosquito-borne disease. Environ Health Perspect 109: 141–161.

19. Aziz S, Ngui R, Lim Y a. L, Sholehah I, Nur Farhana J, et al. (2012) Spatial pattern of 2009 dengue distribution in Kuala Lumpur using GIS application. Trop Biomed 29: 113–120.

20. Costa JV, Donalisio MR, Silveira LV de A (2013) Spatial distribution of dengue incidence and socio-environmental conditions in Campinas, São Paulo State, Brazil, 2007. Cad Saúde Pública 29: 1522–1532. doi:10.1590/0102-311X00110912.

21. Bohra A, Andrianasolo H (2001) Application of GIS in Modeling Dengue Risk Based on Sociocultural Data: Case of Jalore, Rajasthan, India. Available: http://repository.searo.who.int/handle/123456789/15847. Accessed 11 July 2014.

22. Estallo EL, Más G, Vergara-Cid C, Lanfri MA, Ludueña-Almeida F, et al. (2013) Spatial Patterns of High Aedes aegypti Oviposition Activity in Northwestern Argentina. PLoS ONE 8: e54167. doi:10.1371/journal.pone.0054167.

23. Honório NA, Nogueira RMR, Codeço CT, Carvalho MS, Cruz OG, et al. (2009) Spatial Evaluation and Modeling of Dengue Seroprevalence and Vector Density in Rio de Janeiro, Brazil. PLoS Negl Trop Dis 3: e545. doi:10.1371/journal.pntd.0000545.

24. Khormi HM, Kumar L, Elzahrany RA (2011) Modeling spatio-temporal risk changes in the incidence of dengue fever in Saudi Arabia: a geographical information system case study. Geospatial Health 6: 77–84.

25. Nakhapakorn K, Tripathi NK (2005) An information value based analysis of physical and climatic factors affecting dengue fever and dengue haemorrhagic fever incidence. Int J Health Geogr 4: 13. doi:10.1186/1476-072X-4-13.

26. Penna F, Lucia M (2004) Ecological Study of Rio de Janeiro City DEN-3 Epidemic, 2001-2002. Available: http://repository.searo.who.int/handle/123456789/15955. Accessed 11 July 2014.

27. Reiter P, Lathrop S, Bunning M, Biggerstaff BJ, Singer D, et al. (2003) Texas lifestyle limits transmission of dengue virus. Emerg Infect Dis 9: 86–89.

28. Sarfraz MS, Tripathi NK, Tipdecho T, Thongbu T, Kerdthong P, et al. (2012) Analyzing the spatio-temporal relationship between dengue vector larval density and land-use using factor analysis and spatial ring mapping. BMC Public Health 12: 853.

29. Seidahmed OME, Hassan SA, Soghaier MA, Siam HAM, Ahmed FTA, et al. (2012) Spatial and Temporal Patterns of Dengue Transmission along a Red Sea Coastline: A Longitudinal Entomological and Serological Survey in Port Sudan City. PLoS Negl Trop Dis 6: e1821. doi:10.1371/journal.pntd.0001821.

30. Spiegel JM, Bonet M, Ibarra A-M, Pagliccia N, Ouellette V, et al. (2007) Social and environmental determinants of Aedes aegypti infestation in Central Havana: results of a case-control study nested in an integrated dengue surveillance programme in Cuba: Social and environmental determinants of Aedes aegypti infestation. Trop Med Int Health 12: 503–510. doi:10.1111/j.1365-3156.2007.01818.x.

31. Stoddard ST, Forshey BM, Morrison AC, Paz-Soldan VA, Vazquez-Prokopec GM, et al. (2013) House-to-house human movement drives dengue virus transmission. Proc Natl Acad Sci 110: 994–999. doi:10.1073/pnas.1213349110.

32. Thammapalo S, Chongsuwiwatwong V, Geater A, Lim A, Choomalee K (2005) Socio-demographic and environmental factors associated with Aedes breeding places in Phuket, Thailand. Southeast Asian J Trop Med Public Health 36: 426–433.

33. Liebman KA, Stoddard ST, Morrison AC, Rocha C, Minnick S, et al. (2012) Spatial Dimensions of Dengue Virus Transmission across Interepidemic and Epidemic Periods in Iquitos, Peru (1999–2003). PLoS Negl Trop Dis 6: e1472. doi:10.1371/journal.pntd.0001472.

34. Salje H, Lessler J, Endy TP, Curriero FC, Gibbons RV, et al. (2012) Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. Proc Natl Acad Sci 109: 9535–9538. doi:10.1073/pnas.1120621109.

35. Cummings DAT, Irizarry RA, Huang NE, Endy TP, Nisalak A, et al. (2004) Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature 427: 344–347.

36. Teurlai M, Huy R, Cazelles B, Duboz R, Baehr C, et al. (2012) Can Human Movements Explain Heterogeneous Propagation of Dengue Fever in Cambodia? PLoS Negl Trop Dis 6: e1957. doi:10.1371/journal.pntd.0001957.

37. Cuong HQ, Vu NT, Cazelles B, Boni MF, Thai KTD, et al. (2013) Spatiotemporal Dynamics of Dengue Epidemics, Southern Vietnam. Emerg Infect Dis 19: 945–953. doi:10.3201/eid1906.121323.

38. Hu W, Clements A, Williams G, Tong S, Mengersen K (2011) Spatial Patterns and Socioecological Drivers of Dengue Fever Transmission in Queensland, Australia. Environ Health Perspect 120: 260–266. doi:10.1289/ehp.1003270.

39. Lowe R, Bailey TC, Stephenson DB, Graham RJ, Coelho CAS, et al. (2011) Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil. Comput Geosci 37: 371–381. doi:10.1016/j.cageo.2010.01.008.

40. Wu P-C, Lay J-G, Guo H-R, Lin C-Y, Lung S-C, et al. (2009) Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan. Sci Total Environ 407: 2224–2233. doi:10.1016/j.scitotenv.2008.11.034.

41. Carbajo AE, Cardo MV, Vezzani D, others (2012) Is temperature the main cause of dengue rise in non-endemic countries? The case of Argentina. Int J Health Geogr 11: 26.

42. Tipayamongkholgul M, Lisakulruk S (2011) Socio-geographical factors in vulnerability to dengue in Thai villages: a spatial regression analysis. Geospatial Health 5: 191–198.

43. Descloux E, Mangeas M, Menkes CE, Lengaigne M, Leroy A, et al. (2012) Climate-Based Models for Understanding and Forecasting Dengue Epidemics. PLoS Negl Trop Dis 6: e1470. doi:10.1371/journal.pntd.0001470.

44. R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available: http://www.R-project.org/.

45. Météo France (2014). Available: http://www.meteo.nc/climat/normales. Accessed 21 July 2014.

46. Li D, Liu W, Guigon A, Mostyn C, Grant R, et al. (2010) Rapid Displacement of Dengue Virus Type 1 by Type 4, Pacific Region, 2007–2009. Emerg Infect Dis 16: 123–125. doi:10.3201/eid1601.091275.

47. Bouldouyre MA, Baumann F, Berlioz-Arthaud A, Chungue E, Lacassin F (2006) Factors of severity at admission during an epidemic of dengue 1 in New Caledonia (South Pacific) in 2003. Scand J Infect Dis 38: 675–681. doi:10.1080/00365540600606432.

48. A-Nuegoonpipat A, Berlioz-Arthaud A, Chow V, Endy T, Lowry K, et al. (2004) Sustained transmission of dengue virus type 1 in the Pacific due to repeated introductions of different Asian strains. Virology 329: 505–512. doi:10.1016/j.virol.2004.08.029.

49. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, et al. (2001) Age standardisation of rates: a new WHO standard. GPE Discussion Paper Series. World Health Organization.

50. Institut de la statistique et des études économiques (ISEE) (1989) Recensement de la population - 1989.

51. Institut de la statistique et des études économiques (ISEE) (1996) Recensement de la population - 1996.

52. Institut de la statistique et des études économiques (ISEE) (2004) Recensement de la population - 2004.

53. Institut de la statistique et des études économiques (ISEE) (2009) Recensement de la population - 2009. Available: http://www.isee.nc/population/population.html. Accessed 21 November 2011.

54. IPNC (1996) Rapport technique - Année 1996. Nouméa: Institut Pasteur de Nouvelle-Calédonie.

55. IPNC (1999) Rapport technique - Année 1999. Nouméa: Institut Pasteur de Nouvelle-Calédonie.

56. IPNC (2003) Rapport Technique - Année 2003. Institut Pasteur de Nouvelle-Calédonie. Available: www.institutpasteur.nc/rapports-dactivites/. Accessed 19 July 2014.

57. Leroy A (2011) Prévisibilité des épidémies de dengue par Arpège à échéance saisonnière - Prévisibilité des fréquences d'occurrence de jours chauds et jours pluvieux. Météo France.

58. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965–1978. doi:10.1002/joc.1276.

59. Taylor KE, Stouffer RJ, Meehl GA (2012) An Overview of CMIP5 and the Experiment Design. Bull Am Meteorol Soc 93: 485–498. doi:10.1175/BAMS-D-11-00094.1.

60. Anonymous (n.d.) Program for Climate Model Diagnosis and Intercomparison. Available: http://www-pcmdi.llnl.gov/. Accessed 25 July 2014.

61. Bellenger H, Guilyardi E, Leloup J, Lengaigne M, Vialard J (2014) ENSO representation in climate models: from CMIP3 to CMIP5. Clim Dyn 42: 1999–2018. doi:10.1007/s00382-013-1783-z.

62. Cavarero V, Peltier A, Aubail X, Leroy A, Dubuisson B, et al. (2012) Les évolutions passées et futures du climat de la Nouvelle-Calédonie. La Météorologie 77: 13–21.

63. Bivand RS, Pebesma E, Gómez-Rubio V (2013) Applied Spatial Data Analysis with R. 2nd ed. Springer New York. 405 p.

64. Escoffier B, Pagès J (2008) Analyses factorielles simples et multiples. Objectifs, méthodes et interprétation. 4th ed. Paris: Dunod. 318 p.

65. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 27–46. doi:10.1111/j.1600-0587.2012.07348.x.

66. Scherrer B (2008) Biostatistique. 2nd ed. Montréal: Gaëtan Morin éditeur - Chènelière éducation. 816 p.

67. Vapnik V (1998) Statistical learning theory. New York: Wiley. 736 p.

68. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20: 273–297.

69. Ben-Hur A, Weston J (2010) A user's guide to support vector machines. In: Carugo O, Eisenhaber F, editors. Data Mining Techniques for the Life Sciences. Methods in Molecular Biology. Totowa, NJ: Humana Press, Vol. 609.

70. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI. Vol. 14. pp. 1137–1145.

71. Couret J, Dotson E, Benedict MQ (2014) Temperature, Larval Diet, and Density Effects on Development Rate and Survival of Aedes aegypti (Diptera: Culicidae). PLoS ONE 9: e87468. doi:10.1371/journal.pone.0087468.

72. Focks DA, Brenner RJ, Hayes J, Daniels E (2000) Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. Am J Trop Med Hyg 62: 11–18.

73. Mohammed A, Chadee DD (2011) Effects of different temperature regimens on the development of Aedes aegypti (L.) (Diptera: Culicidae) mosquitoes. Acta Trop 119: 38–43. doi:10.1016/j.actatropica.2011.04.004.

74. Dégallier N (2011) Capacidade vetorial : importancia na modelagem da transmissao de dengue. In: Marcondes CB, editor. Entomologia medica e veterinaria : segunda ediçao. Biblioteca Biomédica. Sao Paulo: Atheneu. pp. 37–44.

75. Watts DM, Burke DS, Harrison BA, Whitmire RE, Nisalak A (1987) Effect of temperature on the vector efficiency of Aedes aegypti for dengue 2 virus. Am J Trop Med Hyg 36(1): 143–152.

76. Hopp MJ, Foley JA (2001) Global-scale relationships between climate and the Dengue fever vector, Aedes aegypti. Clim Change 48: 441–463.

77. Yang HM, Macoris MLG, Galvani KC, Andrighetti MTM, Wanderley DMV (2009) Assessing the effects of temperature on the population of Aedes aegypti, the vector of dengue. Epidemiol Infect 137: 1188–1202. doi:10.1017/S0950268809002040.

78. Scott TW, Amerasinghe PH, Morrison AC, Lorenz LH, Clark GG, et al. (2000) Longitudinal studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: blood feeding frequency. J Med Ent 37: 89–101.

79. Stewart Ibarra AM, Ryan SJ, Beltrán E, Mejía R, Silva M, et al. (2013) Dengue Vector Dynamics (Aedes aegypti) Influenced by Climate and Social Factors in Ecuador: Implications for Targeted Control. PLoS ONE 8: e78263. doi:10.1371/journal.pone.0078263.

80. Degallier N, Vilarinhos P d. TR, Carvalho M d. SL d., Knox MB, Caetano J (2000) People's knowledge and practice about dengue, its vectors and control means in Brasilia (DF), Brazil: its relevance with entomological factors. J Amer Mosq Cont Assoc 16 (2): 114–123.

81. Dormann CF, M. McPherson J, B. Araújo M, Bivand R, Bolliger J, et al. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30: 609–628. doi:10.1111/j.2007.0906-7590.05171.x.

Figures

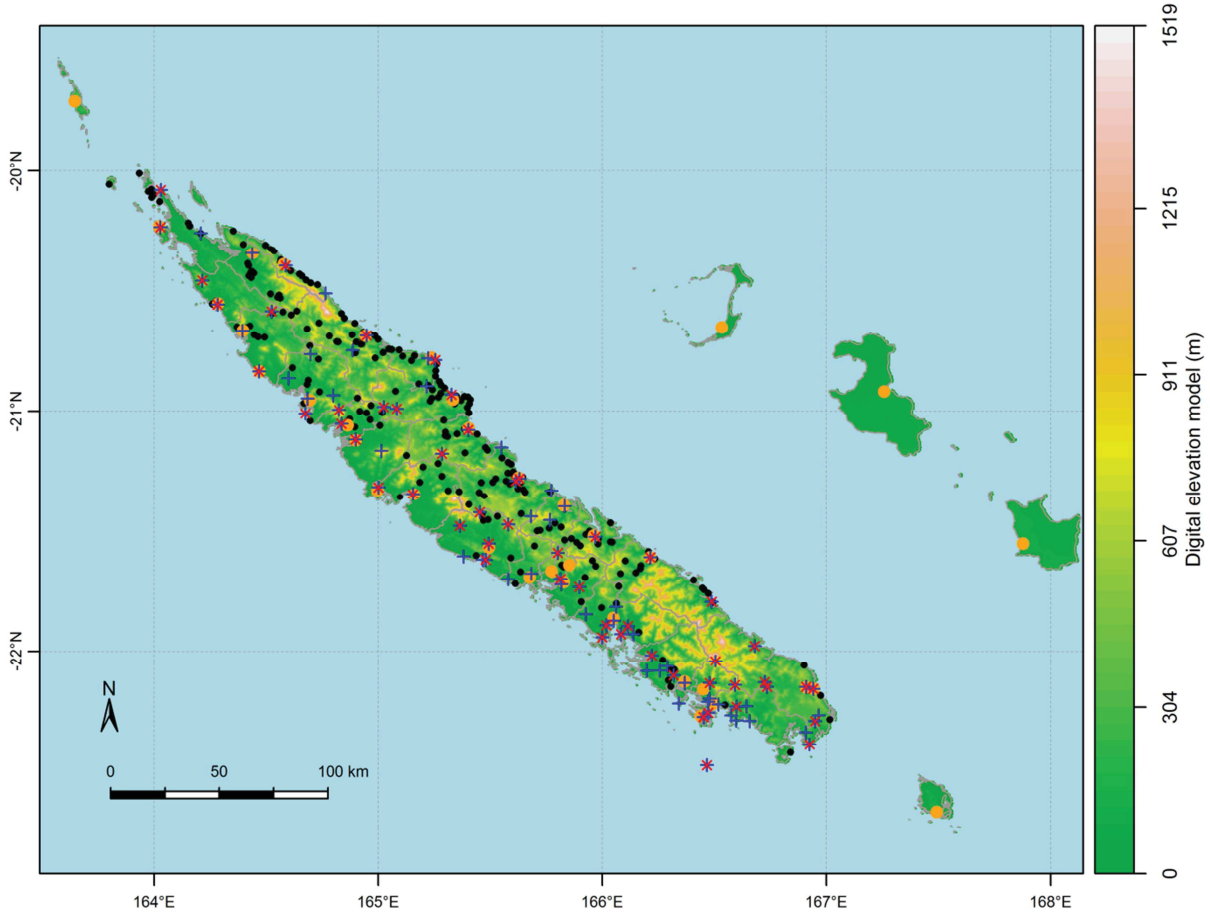


Figure 4.7 : general map of New-Caledonia. *The map shows the location of towns (orange dots), tribes (black dots), and weather stations registering temperature (red crosses) and rainfall (blue crosses) in New Caledonia. The background colour represents the digital elevation model (altitude).*

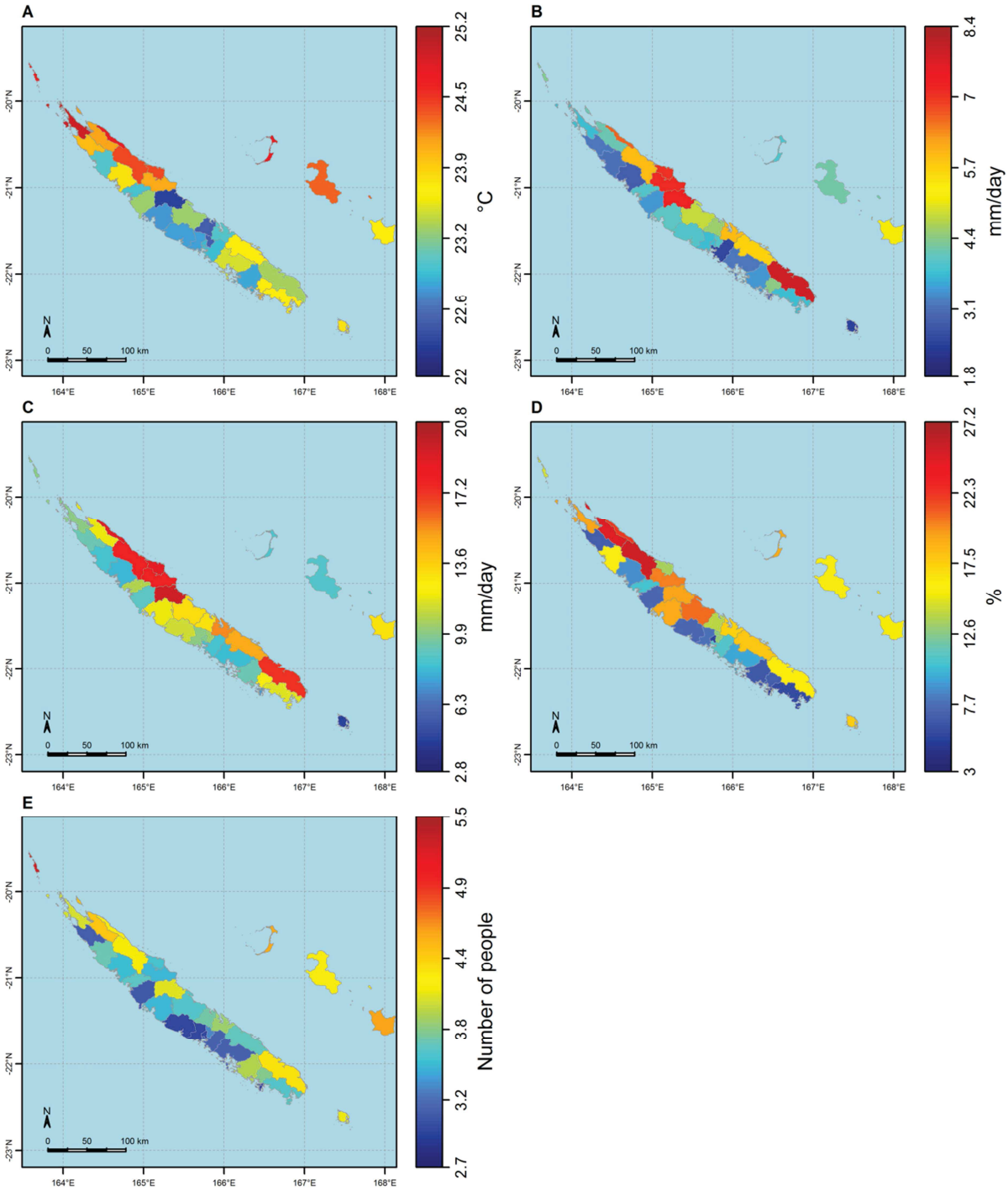


Figure 4.8: maps of the 5 explanatory variables selected for modeling dengue incidence rates. A: *average mean temperature;* ***B:*** *average daily rainfall;* ***C:*** *average daily rainfall during the wettest quarter;* ***D:*** *percentage of unemployed people;* ***E:*** *average number of people per premise.*

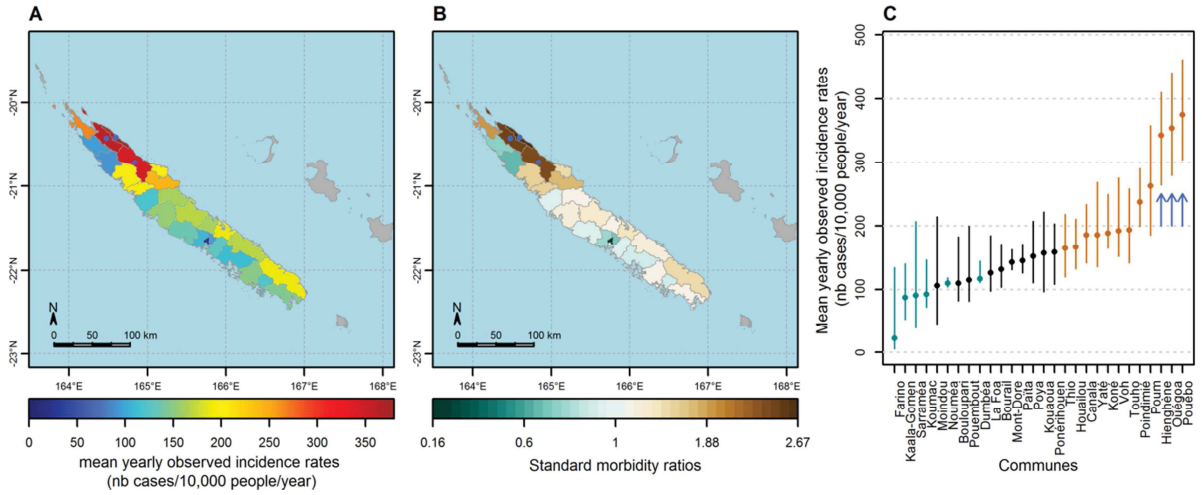


Figure 4.9: Spatial heterogeneity of dengue annual incidence rates in New Caledonia. A: *map of the average annual incidence rates per commune calculated over the period 1995-2012 during epidemic years;* ***B:*** *map of the ratio between the average incidence rate of a given commune and the territorial mean;* ***C:*** *average annual incidence rates (y-axis) per commune (x-axis) calculated over the period 1995-2012 during epidemic years. Communes are sorted by increasing annual incidence rate. Communes with unusually high dengue incidence rates are highlighted by blue dots (A and B) or blue arrows (C).*

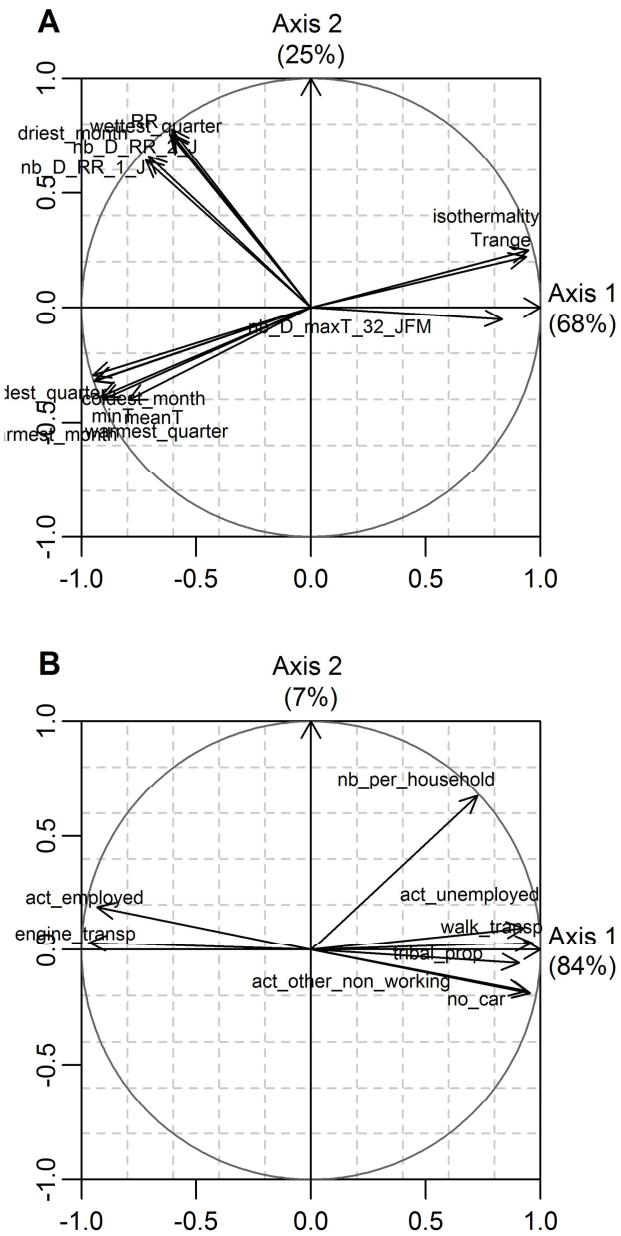


Figure 4.10: Principal component analysis over the set of climatic variables (A) and socio-economic variables (B). *The figure shows the correlation circles of PCA performed on the variable most correlated with dengue incidence rates (see methods). Pearson correlation coefficients between variables can be approximated by the angle between the corresponding arrows: 1 for a 0° angle, 0 for a 90° angle, and -1 for a 180° angle.*

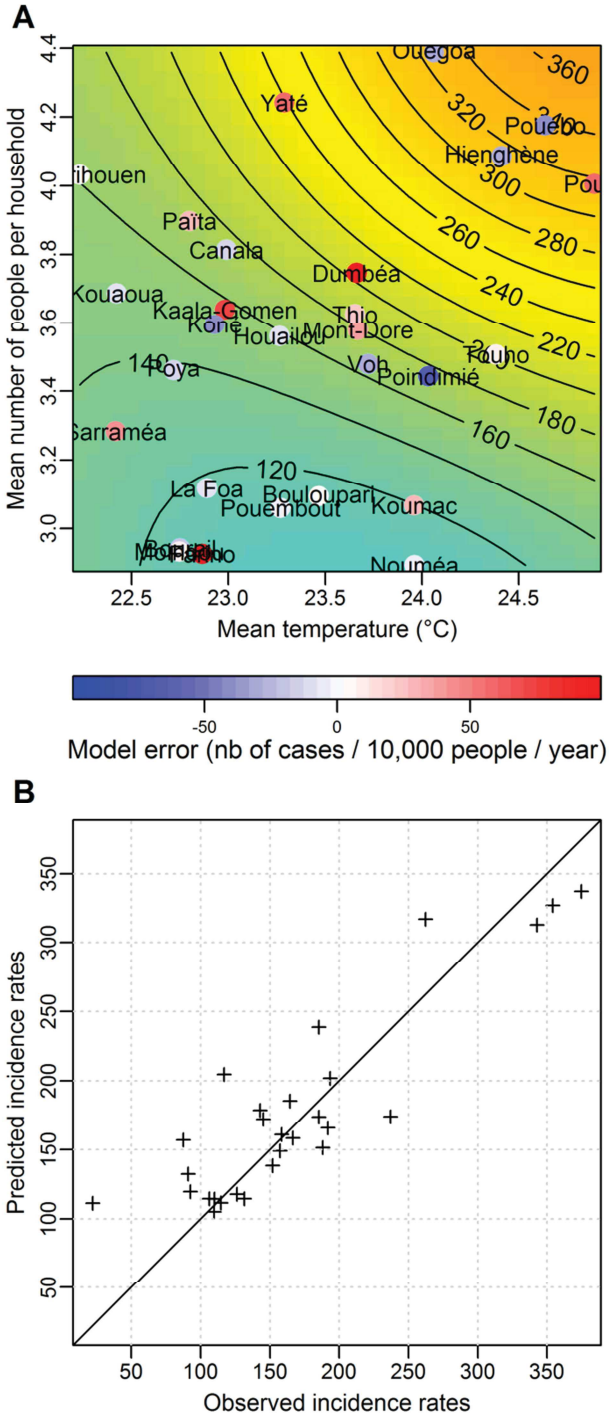


Figure 4.11: Results of the best multi-variate model of dengue incidence rates. A: *Predicted mean annual incidence rates as a function of the two best explanatory variables (mean temperature and mean number of people per premise). The axes represent the value of the two best explanatory variables. Predicted annual incidence rates are represented by the colour (blue for low incidence rates to orange for high incidence rates) and by the contour lines giving incidence rates in number of cases per 10,000 people per year. Each commune that has been used to build the model is placed on the graph according to the observed value of the two explanatory variables in the commune. Its position on the graph hence provides the incidence rate in the commune as predicted by the model. For each commune, the coloured dot represents the difference between the predicted and the observed incidence rate (model error).* ***B:*** *Scatter plot of the predicted and observed incidence rates for each of the 28 communes. The RMSE of this model is 45 cases/10,000/year*

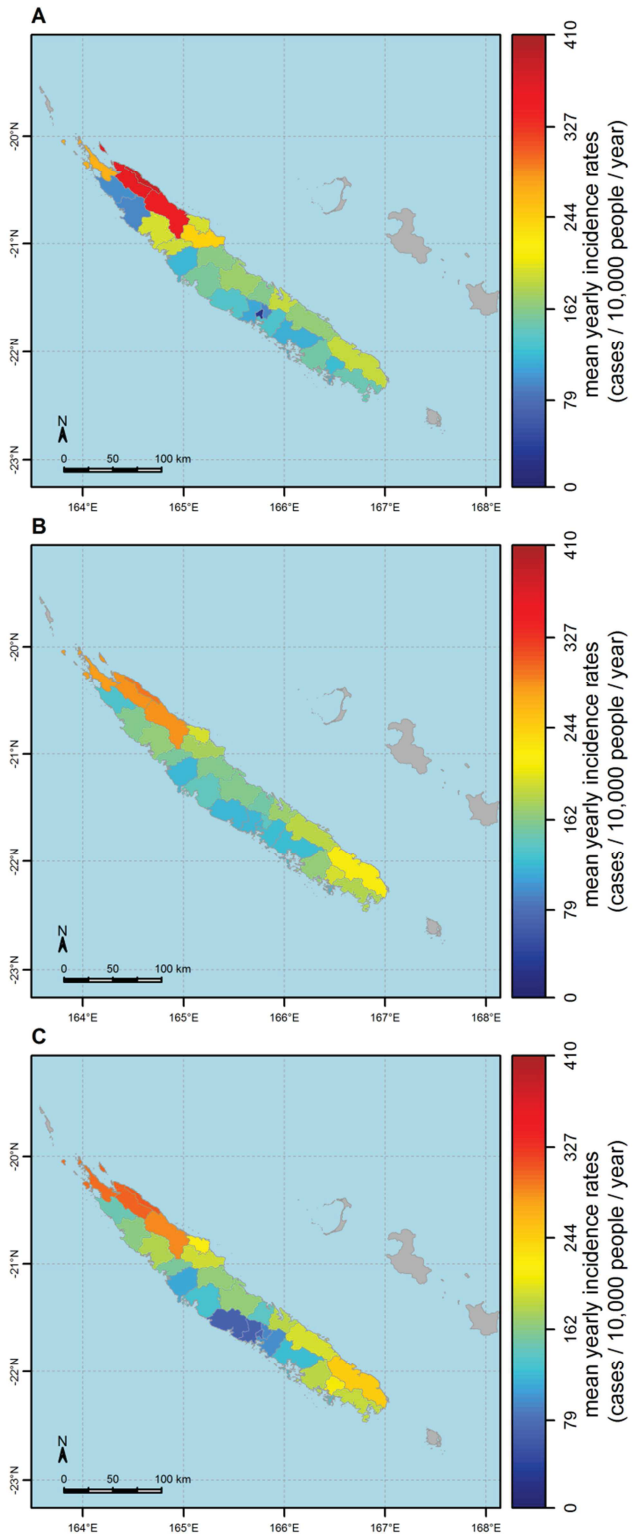


Figure 4.12: Maps of observed and predicted annual incidence rates. A: *map of observed dengue annual incidence rates.* ***B and C:*** *maps of dengue annual incidence rates predicted by the SVM model (B) and the linear model (C) based on the mean temperature and the mean number of people per premise.*

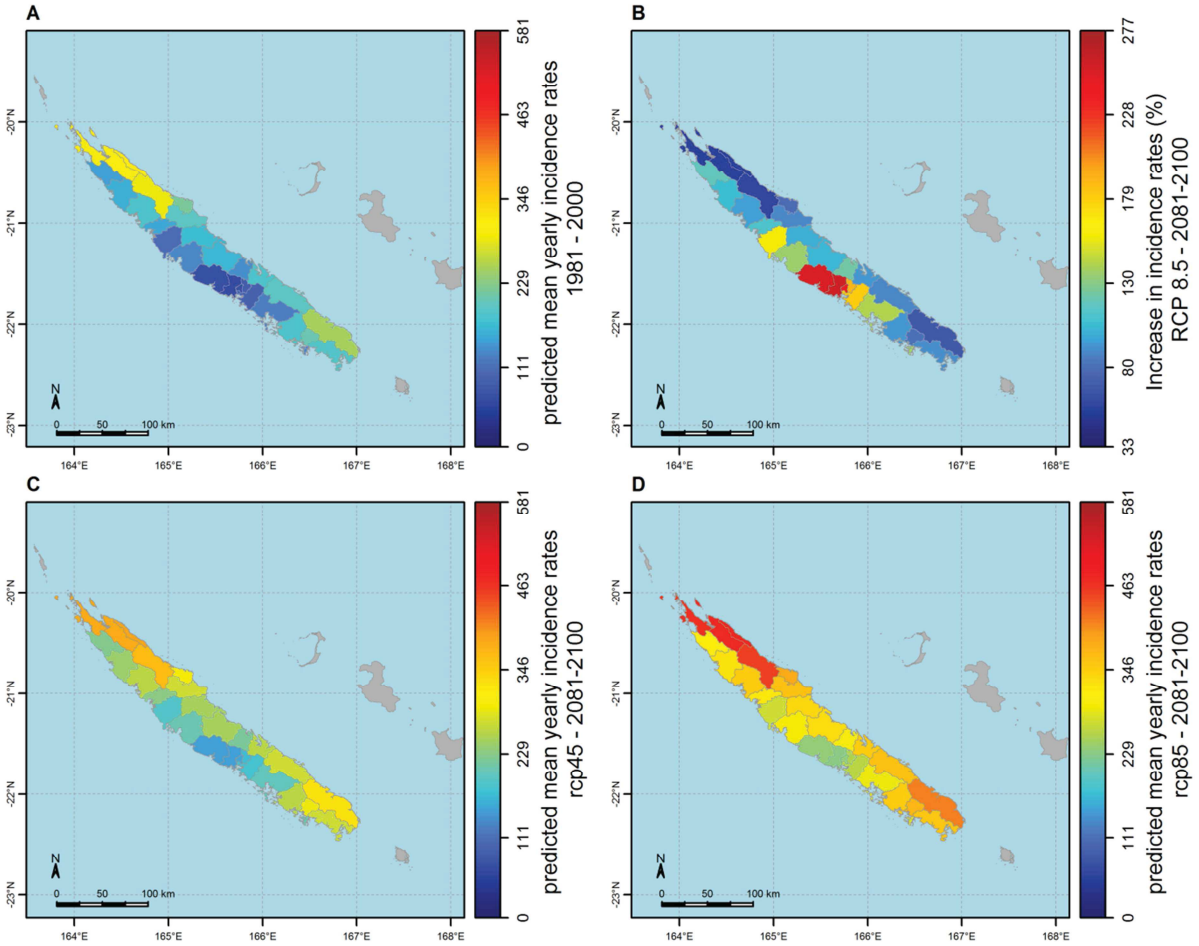
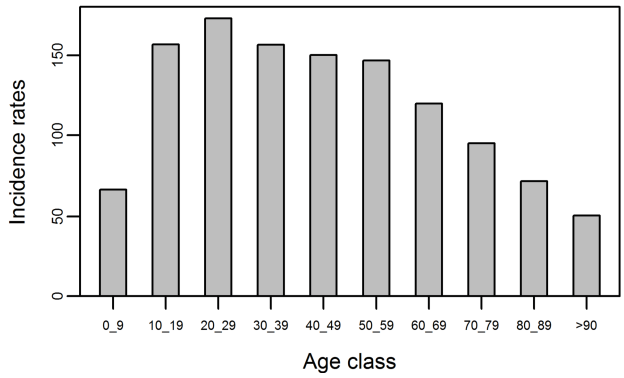
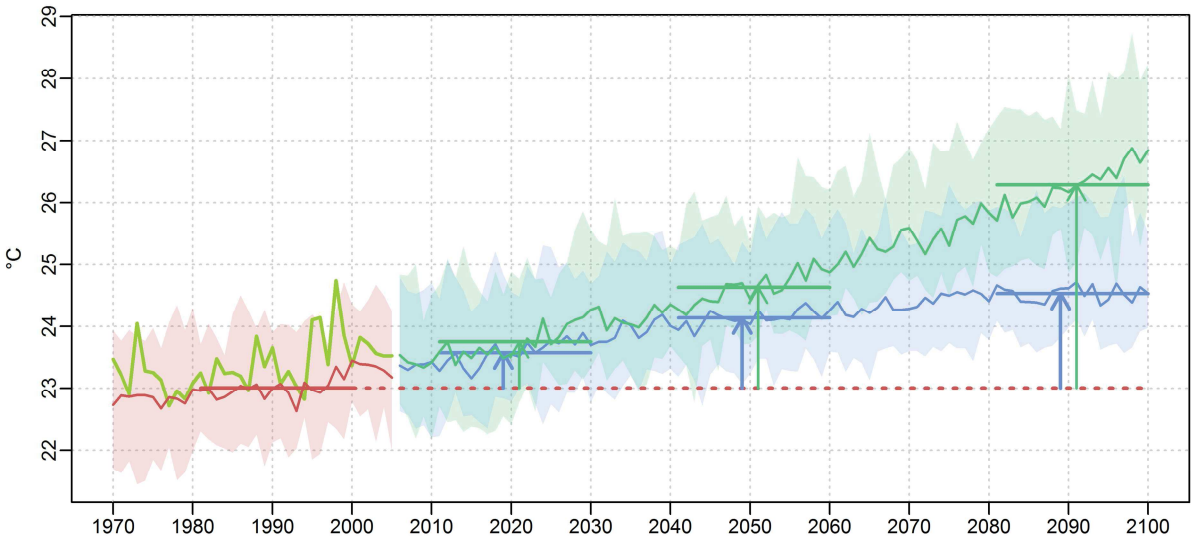


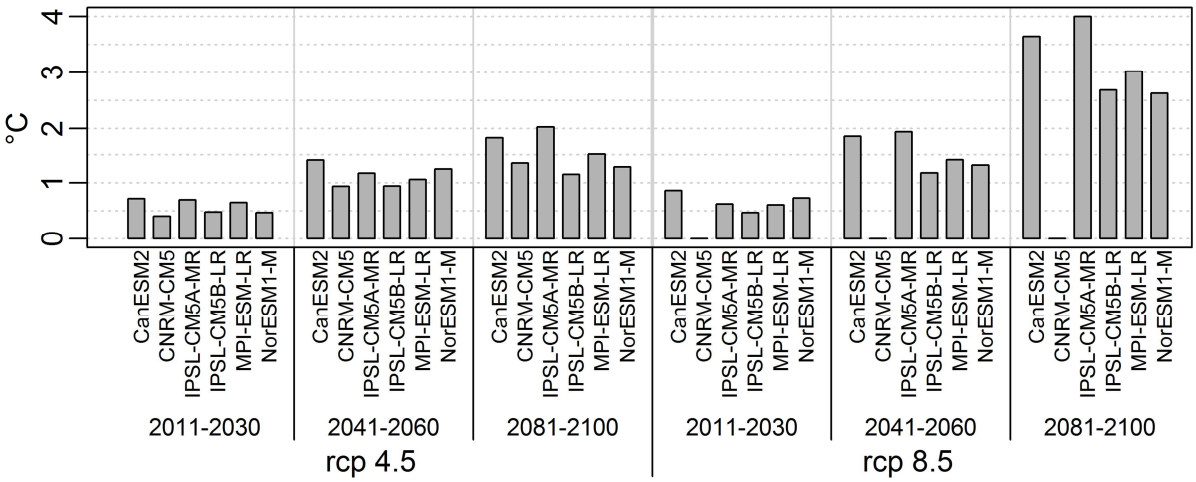
Figure 4.13: trends of dengue spatial distribution under global warming. A: *Dengue annual incidence rates as predicted by the linear model applied to historical simulations of climate (1981-2000).* ***B:*** *Relative increase of annual incidence rates (in % of the incidence rates shown in A) as projected in 2081-2100 under the RCP 8.5 scenario.* ***C and D:*** *Annual incidence rates as projected in 2081-2100 under the RCP 4.5 (C) and the RCP 8.5 (D) scenarii.*



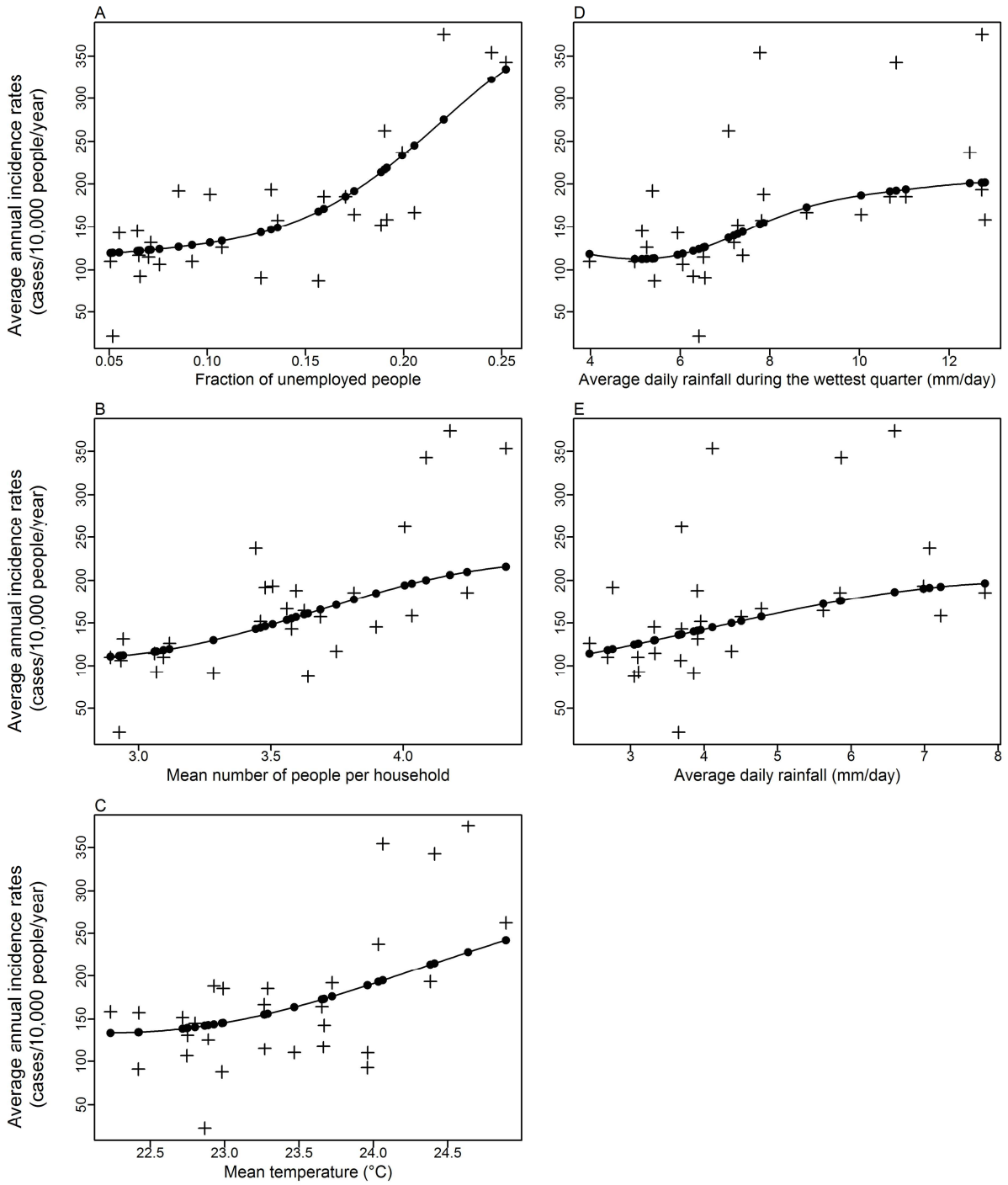
Supplementary Figure 4.14: Mean dengue annual incidence rates (in number of declared cases per 10,000 people per year) during epidemic years between 1995 and 2012 according to age classes.



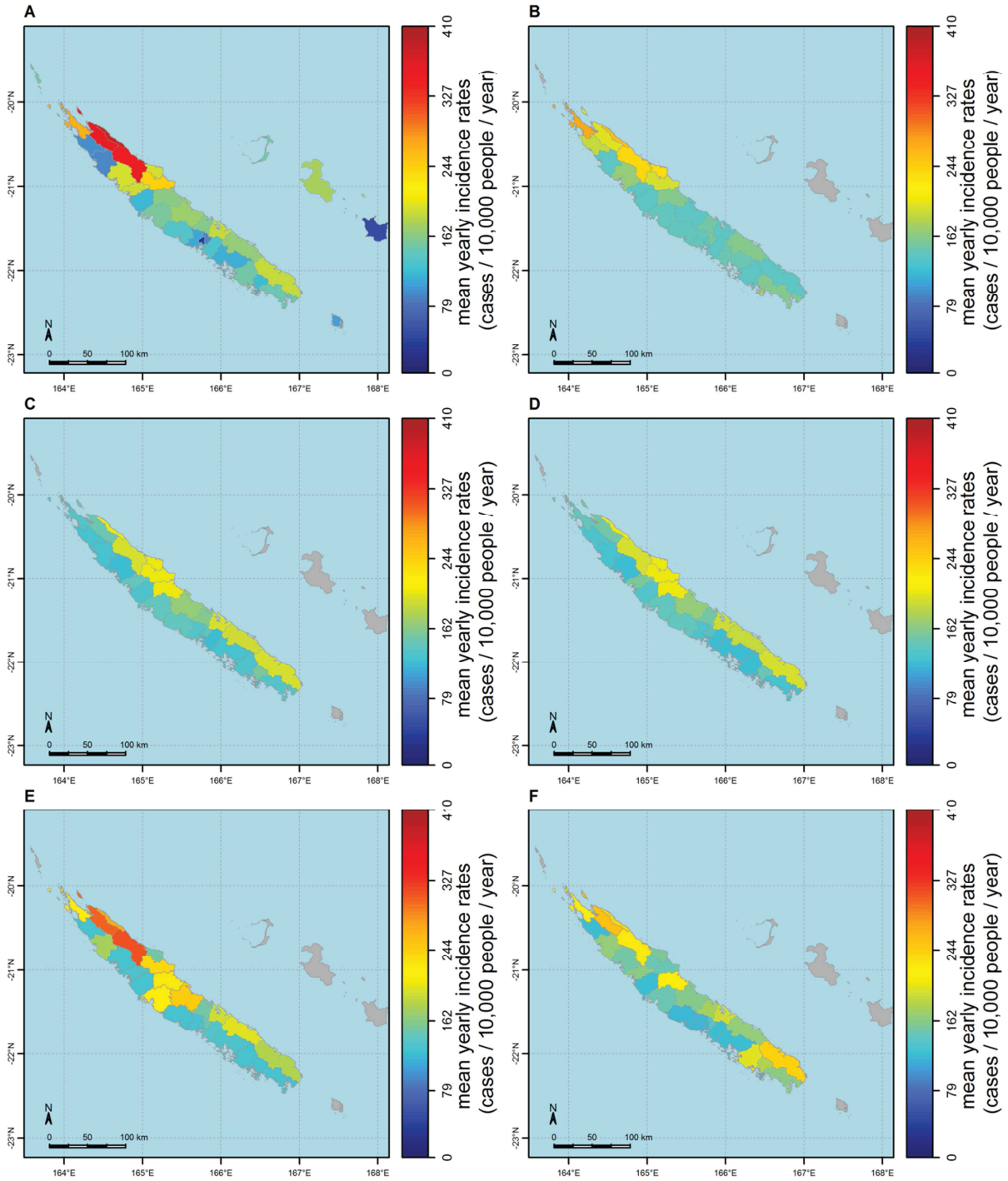
Supplementary Figure 4.15: Increase of mean temperature in Noumea as projected by six coupled atmosphere-ocean models at the 2100 horizon. *The yellow-green curve (1971-2005) represents the observed mean temperature in Noumea. The other curves represent the average of the annual time series simulated by the six models over the historical period (1971-2005, red curve), and the 2006-2100 period under RCP 4.5 (blue curve) and RCP 8.5 (green curve) scenarii. Horizontal segments represent the average of the time series over the given periods (1981-2000, 2011-2030, 2041-2060 and 2081-2100). Arrows represent the average increase in mean temperature for each time period (see table 3). Background shadings represent the annual range of mean temperature simulations over the six models.*



Supplementary Figure 4.16: Comparison of six coupled ocean-atmosphere model projections *for the increase in mean temperature in Noumea for two climate change scenarii (RCP4.5 and RCP 8.5) and three time periods relative to the historical series of temperature (1980-2000).*



Supplementary Figure 4.17: Relationship between dengue annual incidence rates and the 5 selected explanatory variables. *For each selected explanatory variable (**A to E**), mean annual incidence rates observed in 28 communes of new Caledonia (crosses), and mean annual incidence rates predicted by the bivariate SVM model based on corresponding explanatory variable (dots). The curve represents the bivariate SVM model predictions over the whole observed range of the explanatory variable (non-linear regression curve).*



Supplementary Figure 4.18: Spatial influence between the selected explanatory variables and dengue incidence rates. A: *map of the observed mean dengue annual incidence rates.* ***B to E****: maps of mean annual incidence rates predicted by bivariate SVM models based only on one of the 5 selected variables: mean temperature (B), average daily rainfall (C), average daily rainfall during the wettest quarter (D), percentage of unemployed people (E) and mean number of people per household (F).*

Tables

Table 4.2 : Correlation between dengue incidence rates and socio-economic or climate variables

Type	Variable	Rho*	p value	Category	Description
Climate	meanT	0,619	0.0004	temperature	Average mean temperature
	coldest_quarter	0,617	0.0005	temperature	Average minimum temperature during the coldest quarter
	coldest_month	0,608	0.0006	temperature	Average minimum temperature during the coldest month
	minT	0,586	0,0010	temperature	Average minimum temperature
	wettest_quarter	0,563	0,0018	rainfall	Average daily rainfall during the wettest quarter
	driest_month	0,559	0,0020	rainfall	Average daily rainfall during the driest month
	isothermality	-0,515	0,0051	temperature range	Average of the difference between mean and minimum temperature, divided by mean temperature
	warmest_quarter	0,508	0,0058	temperature	Average minimum temperature during the warmest quarter
	nb_D_maxT_32_JFM	-0,508	0,0058	temperature range	Average number of days with maximal temperature exceeding 32 °C in January, February and March
	Trange	-0,491	0,0080	temperature range	Average temperature range (difference between maximum and minimum temperature)
	warmest_month	0,483	0,0093	temperature	Average minimum temperature during the warmest month
	RR	0,473	0,0109	rainfall	Average daily rainfall
	nb_D_RR_1_J	0,419	0,0264	rainfall	Average number of days with daily rainfall exceeding 1mm during January
	nb_D_RR_2_J	0,408	0,0313	rainfall	Average number of days with daily rainfall exceeding 2mm during January
	wettest_month	0,296	0,1262	rainfall	Average daily rainfall during the wettest month
	nb_D_maxT_30_JFM	-0,202	0,3015	temperature range	Average number of days with maximal temperature exceeding 30 °C in January, February and March
	dem	-0,186	0,3440		Digital elevation model (altitude)
	driest_quarter	0,125	0,5265	rainfall	Average daily rainfall during the driest quarters
	maxT	-0,062	0,7548	temperature	Average maximum temperature
	nb_D_maxT_28_JFM	-0,007	0,9700	temperature range	Average number of days with maximal temperature exceeding 28 °C in January, February and March
Socio-economic	act_unemployed	0,759	< 0.0001	Activity	Percentage of unemployed people
	engine_transp	-0,744	< 0.0001	Human movement	Percentage of people using a motorised vehicle to get around
	nb_per_household	0,739	< 0.0001	Human density	Average number of people per premise
	walk_transp	0,722	< 0.0001	Human movement	Percentage of people getting around by foot
	act_employed	-0,708	< 0.0001	Activity	Percentage of people working (any occupational activity)
	no_car	0,704	< 0.0001	Human	Percentage of premises with no car

				movement	
	act_other_non_working	0,648	0.0002	Activity	Percentage of people that are unoccupied but are neither students nor retired
	tribal_prop	0,632	0.0003	Activity	Percentage of people living in a tribe
	spc_agri	0,597	0.0008	Activity	Percentage of people working as farmers
	WC_inside	-0,568	0,0016	Housing	Percentage of premises with toilets inside
	surf_40_m2	0,568	0,0016	Human density	Percentage of premises under 40 m²
	house	-0,540	0,0030	Housing	Percentage of premises that are houses
	live_same04	0,532	0,0036	Human movement	Percentage of people living in the commune before 2004
	air_con	-0,528	0,0039	Housing	Percentage of premises with at least one room with air conditioning
	live_other_com04	-0,525	0,0042	Human movement	Percentage of people living in another commune before 2004
	act_retired	-0,490	0,0081	Activity	Percentage of retired people
	electricity	-0,488	0,0084	Housing	Percentage of premises with access to public electricity
	spc_arti	-0,480	0,0097	Activity	Percentage of people working as artisans
	work_other_com	-0,477	0,0103	Human movement	Percentage of people employed in another commune
	birth_NC	0,475	0,0106	Activity	Percentage of people born in New Caledonia
	birth_mainland_France	-0,470	0,0115	Activity	Percentage of people born in France, mainland
	birth_FP	-0,466	0,0125	Activity	Percentage of people born in French Polynesia
	concrete_slab	-0,460	0,0139	Housing	Percentage of premises built on a concrete slab
	surf_40_120_m2	-0,453	0,0154	Human density	Percentage of premises between 40 m² and 120 m²
	surf_120_m2	-0,448	0,0169	Human density	Percentage of premises over 120 m²
	birth_abroad	-0,433	0,0215	Activity	Percentage of people born abroad
	nb_rooms	-0,428	0,0230	Human density	Average number of bedrooms per premise
	hut	0,387	0,0421	Housing	Percentage of premises that are huts
	no_bikes	0,360	0,0599	Human movement	Percentage of premises with no bicycle nor motorcycle
	live_main_france04	-0,349	0,0685	Human movement	Percentage of people living in the mainland of France before 2004
	indiv_water	0,345	0,0722	Housing	Percentage of premises using an individual watering place located outside
	shed	0,330	0,0866	Housing	Percentage of premises that are temporary sheds
	birth_WF	-0,305	0,1150	Activity	Percentage of people born in Wallis and Futuna
	tap_water	-0,302	0,1179	Housing	Percentage of premises with access to tap water
	sect1	0,262	0,1774	Activity	Percentage of people employed in the primary sector (agriculture…)
	main_home	0,258	0,1849	Human movement	Percentage of premises that are the main home of a household

	sect2	-0,244	0,2115	Activity	Percentage of people employed in the secondary sector (industry…)
	spc_exec	-0,233	0,2334	Activity	Percentage of people working as managers or executives
	spc_lab	-0,206	0,2926	Activity	Percentage of people working as labour workers
	live_abroad04	-0,170	0,3861	Human movement	Percentage of people living abroad before 2004
	public_transp	0,164	0,4037	Human movement	Percentage of people using public transportation to get around
	studying	0,148	0,4529	Activity	Percentage of people unoccupied and studying
	act_undergrad_student	0,148	0,4529	Activity	Percentage of people registered in an undergraduate course
	coll_water	0,132	0,5037	Housing	Percentage of premises using collective watering places
	sect3	0,024	0,9015	Activity	Percentage of people employed in the tertiary sector (sales, trade, services…)
	spc_other	0,024	0,9049	Activity	Percentage of people working but not as farmers nor managers nor workmen nor artisans nor employees
	spc_employee	0,001	0,9958	Activity	Percentage of people working as employees

* Pearson correlation coefficient (Rho) with dengue incidence rates and associated p-value.
Variables are sorted by category (socio-economic or climate) and by decreasing order of their absolute Pearson correlation. Variables selected for the multivariate modelling are in bold

Table 4.3 : Bivariate and multivariate modelling of dengue incidence rates: variable selection according to the RMSE of the SVM models

Variable 1*	Variable 2*	Variable 3*	RMSE**
act_unemployed	-	-	53
nb_per_household	-	-	68
meanT	-	-	69
wettest_quarter	-	-	72
RR	-	-	75
meanT	**nb_per_household**	**-**	**45**
meanT	**act_unemployed**	**-**	**47**
act_unemployed	**nb_per_household**	**-**	**49**
RR	nb_per_household	-	58
RR	meanT	-	65
wettest_quarter	act_unemployed	-	66
wettest_quarter	meanT	-	67
RR	act_unemployed	-	68
wettest_quarter	nb_per_household	-	69
wettest_quarter	RR	-	73
meanT	**nb_per_household**	**act_unemployed**	**42**
meanT	**nb_per_household**	**RR**	**47**
meanT	**act_unemployed**	**RR**	**48**
meanT	**act_unemployed**	**wettest_quarter**	**49**
nb_per_household	**act_unemployed**	**RR**	**50**
nb_per_household	**meanT**	**wettest_quarter**	**52**
nb_per_household	act_unemployed	wettest_quarter	53
nb_per_household	RR	wettest_quarter	63
meanT	RR	wettest_quarter	66
act_unemployed	RR	wettest_quarter	65

* Variables included as explanatory variables for modelling dengue incidence rates
** Root mean square error of each model, in number of cases /10,000 people / year. Models are classified first by the number of explanatory variables used, then by increasing RMSE.
Models highlighted in bold perform better than the best bivariate model

Table 4.4 : Projections of temperature increase and predicted incidence rates for three time periods in the future.

Scenario	Period	Projected increase of mean temperature * (°C)	Standard deviation of the projected mean temperature increase (°C)	Predicted incidence rates: territory average	Increase of territory average dengue incidence rates**
-	1981-2000	-	-	168	-
RCP4.5	2011-2030	0.57	0.14	197	29
	2041-2060	1.13	0.18	226	58
	2081-2100	1.53	0.34	247	78
RCP8.5	2011-2030	0.66	0.15	202	34
	2041-2060	1.54	0.34	247	79
	2081-2100	3.20	0.60	332	164

* Average of the mean temperature increase predicted by 6 coupled ocean-atmosphere models (see methods)
** Relative to the territory average predicted for the 1981-2000 period

CHAPITRE 5 : ANALYSE DE LA DENGUE A L'ECHELLE LOCALE

5.1. Analyse spatio-temporelle 249

5.1.1. Dynamique de la dengue à Nouméa 249
5.1.1. Structure spatiale de la dengue à Nouméa 250
5.1.2. Etude de l'évolution temporelle des schémas spatiaux de la dengue par la méthode des EOF : application à Nouméa 253

5.2. Analyse temporelle 256

5.2.1. Modélisation déterministe du risque épidémique 256
5.2.1.a. Objectifs du modèle 256
5.2.1.b. Construction du modèle 257
Description des étapes 257
Paramètres, variables et équations 259
5.2.1.c. Exploration du modèle 261
Durée des cycles 261
Nombre de cas secondaires suite à l'introduction d'un cas dans une population entièrement susceptible 262
Seuil entomologique nécessaire au développement d'une épidémie 263
5.2.1.d. Relation entre le *R0*, les indices entomologiques et la température, et analyse de sensibilité 264
Influence des variables d'entrées sur le R0 calculé 264
Effet du changement de résolution temporelle de la température sur la variabilité inter-annuelle du *R0* 265
Analyse de sensibilité du modèle aux différents paramètres 266
5.2.1.e. Discussion et perspectives 267
Hypothèses du modèle 267
Valeurs seuils d'indices entomologiques 267
Paramètres du cycle de vie des moustiques influençant la transmission 268
Variabilité inter-annuelle du risque d'émergence épidémique 268

5.3. Analyse spatiale 271

5.3.1. Objectifs et études réalisées 271
5.3.2. Enquête spatiale des facteurs de risque socio-économiques de la dengue à Nouméa 271
5.3.2.a. Matériel et méthode 271
Protocole d'enquête 271
Analyses statistiques 272
5.3.2.b. Résultats 272
5.3.3. Modélisation non linéaire de l'intensité de la circulation virale 274
5.3.3.a. Matériel et méthodes 274
Données épidémiologiques 274
Variables explicatives 275
Pré-sélection des variables explicatives 276

Modélisation multi-variée du risque de transmission 276
Etudes d'auto-corrélation..... 277
5.3.3.b. Résultats..... 277
Distribution spatiale moyenne des taux d'incidence 277
Auto-corrélation de la variable réponse 278
Pré-sélection des variables explicatives..... 279
Modélisation multivariée..... 284
Etudes d'auto-corrélation des résidus des modèles 285
5.3.3.c. Discussion..... 286

5.1. Analyse spatio-temporelle

5.1.1. Dynamique de la dengue à Nouméa

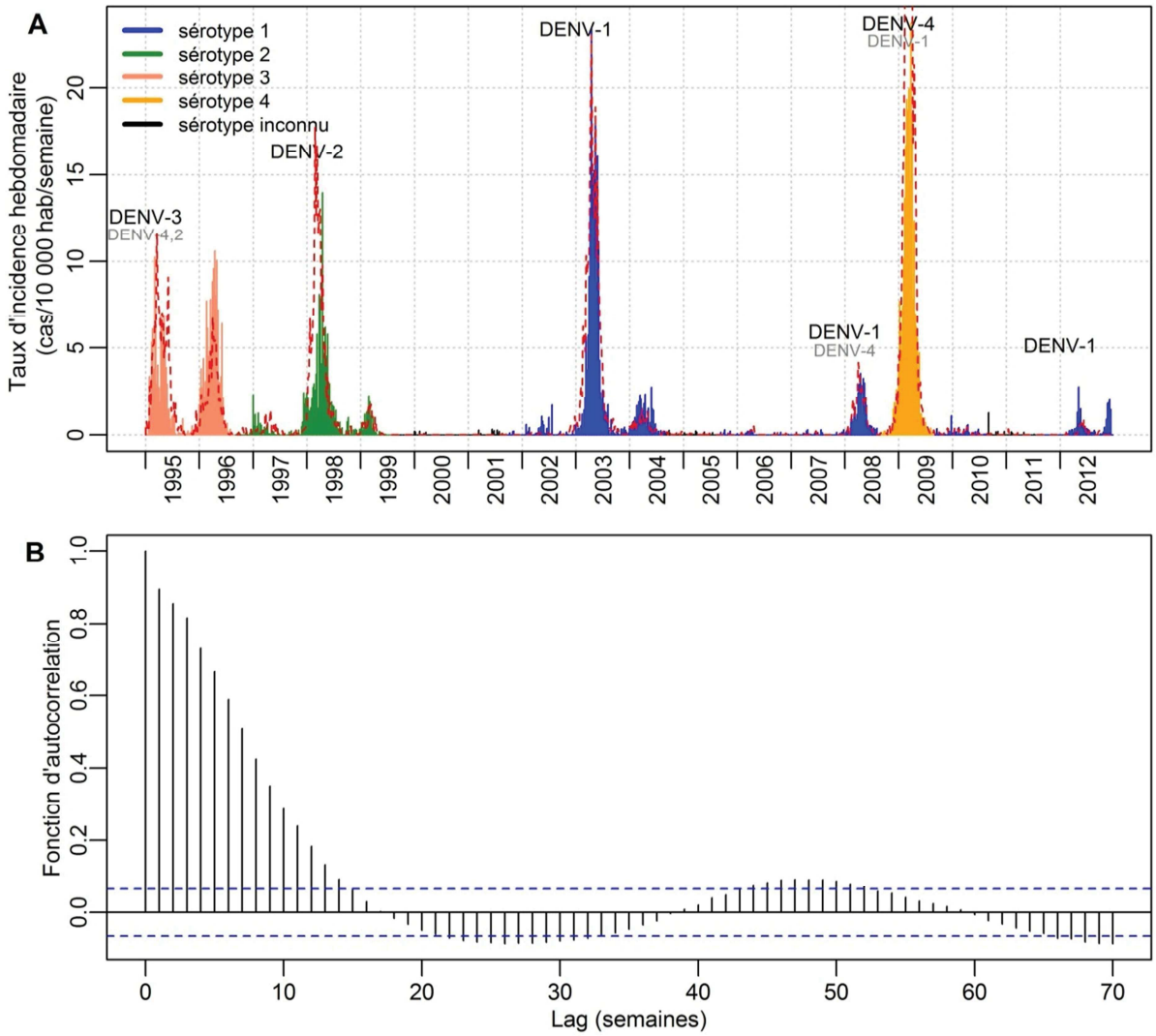


Figure 5.1 : Dynamique de la dengue à Nouméa. A. *Taux d'incidence hebdomadaire dans la commune de Nouméa. La couleur code pour le sérotype dominant. La ligne pointillée rouge représente les taux d'incidence hebdomadaires dans l'ensemble de la Nouvelle-Calédonie.* ***B.*** *Fonction d'auto-corrélation des taux d'incidence de la dengue à Nouméa (corrélation entre x(t) et x(t-k), k dénotant le nombre de semaines de décalage entre les deux séries). Les deux lignes horizontales marquent les limites de significativité des coefficients de corrélation.*

A Nouméa, qui comprend près des deux-tiers de la population Calédonienne (Institut de la statistique et des études économiques (ISEE), 2009), la dynamique de la dengue est similaire à celle de la Calédonie entière (corrélation de Pearson entre les taux d'incidence hebdomadaires à Nouméa et sur le territoire de 0,87, voir Figure 5.1). La dengue y présente une forme épidémique saisonnière, avec un démarrage des épidémies en décembre/janvier (sauf l'épidémie de 2013 qui a débuté en septembre 2012), un pic en avril, et une fin en juin/juillet (Figure 5.1.A). Quelle que soit la commune d'apparition des premiers cas autochtones de dengue sur le territoire, les premiers cas de dengue à Nouméa sont toujours détectés

dans les premières semaines de l'épidémie (résultats non montrés). Par contre le développement des épidémies n'y est pas plus rapide ou moins rapide qu'ailleurs (voir analyse en ondelettes du chapitre précédent). Les taux d'incidence n'y sont pas plus élevés que sur d'autres communes du territoire, ce qui pourrait étonner étant donnée la densité de population sur Nouméa par rapport aux autres communes.

Les taux d'incidence présentent une auto-corrélation jusqu'à environ 12 semaines (trois mois). L'auto corrélation à un an apparaît très faible, contrairement à ce qui a été observé en Asie (Cazelles et al., 2005; Johansson et al., 2009b). Un cycle annuel est pourtant présent dans les données, mais uniquement de manière transitoire : les données d'incidence ne sont pas stationnaires. Nous aurions pu ici utiliser un spectre d'ondelette pour détecter ce cycle annuel, cependant il aurait été quasiment identique à celui du territoire entier (Figure 4.1.B), et nous avons préféré monter ici la fonction d'auto-corrélation, pour illustrer comment la non-stationnarité des données épidémiologiques peut masquer certaines informations lorsque des techniques ne prenant pas en compte cette non-stationnarité sont utilisées (voir également Cazelles and Hales, 2006 sur la question de la non-stationnarité).

5.1.1. Structure spatiale de la dengue à Nouméa

La Figure 5.3 montre la structure spatiale des taux d'incidence annuels de la dengue sur le grand Nouméa. En moyenne, nous pouvons observer une hétérogénéité spatiale de la dengue, avec un gradient Sud/Nord : les quartiers du Sud de la ville sont moins touchés que ceux du Nord, et ce quelle que soit l'année considérée. Certaines années, nous pouvons également observer un gradient Est/Ouest : les quartiers Ouest de la ville (Ducos, Logicoop, Kaméré, Rivière Salée) sont souvent les plus touchés, sauf en 2004.

Le calcul des taux d'incidence (incidence rapportée à la population de chacun des quartiers) est une étape nécessaire avant de pouvoir comparer les niveaux de circulation virale. Cependant, lorsque les taux d'incidence sont calculés à haute résolution spatiale, il est nécessaire de faire attention à leur précision : l'estimation des taux d'incidence est d'autant moins précise que le quartier est peu peuplé. Ainsi, il faut être prudent dans l'interprétation des taux d'incidence des quartiers de Koumourou et Numbo (254 habitants), Doniambo (279 habitants), Montagne coupée (252 habitants), Receiving (314 habitants) et Vallée du Génie (264 habitants), où l'apparition d'un faible nombre de cas peut faire augmenter les taux d'incidence rapidement (augmentation d'environ 30 cas / 10 000 habitants pour chaque nouveau cas recensé).

La structure d'âge étant considérée homogène sur l'ensemble de la ville, nous n'avons pas standardisé les taux d'incidence par rapport à l'âge. La distribution spatiale de la dengue semble relativement stable d'une année à l'autre (Figure 5.3.B. à F), ce que nous allons vérifier au moyen d'une analyse EOF. Du fait de la faible précision de l'estimation des taux d'incidence dans les quartiers peu peuplés, nous avons limité l'analyse aux quartiers comprenant plus de 800 habitants.

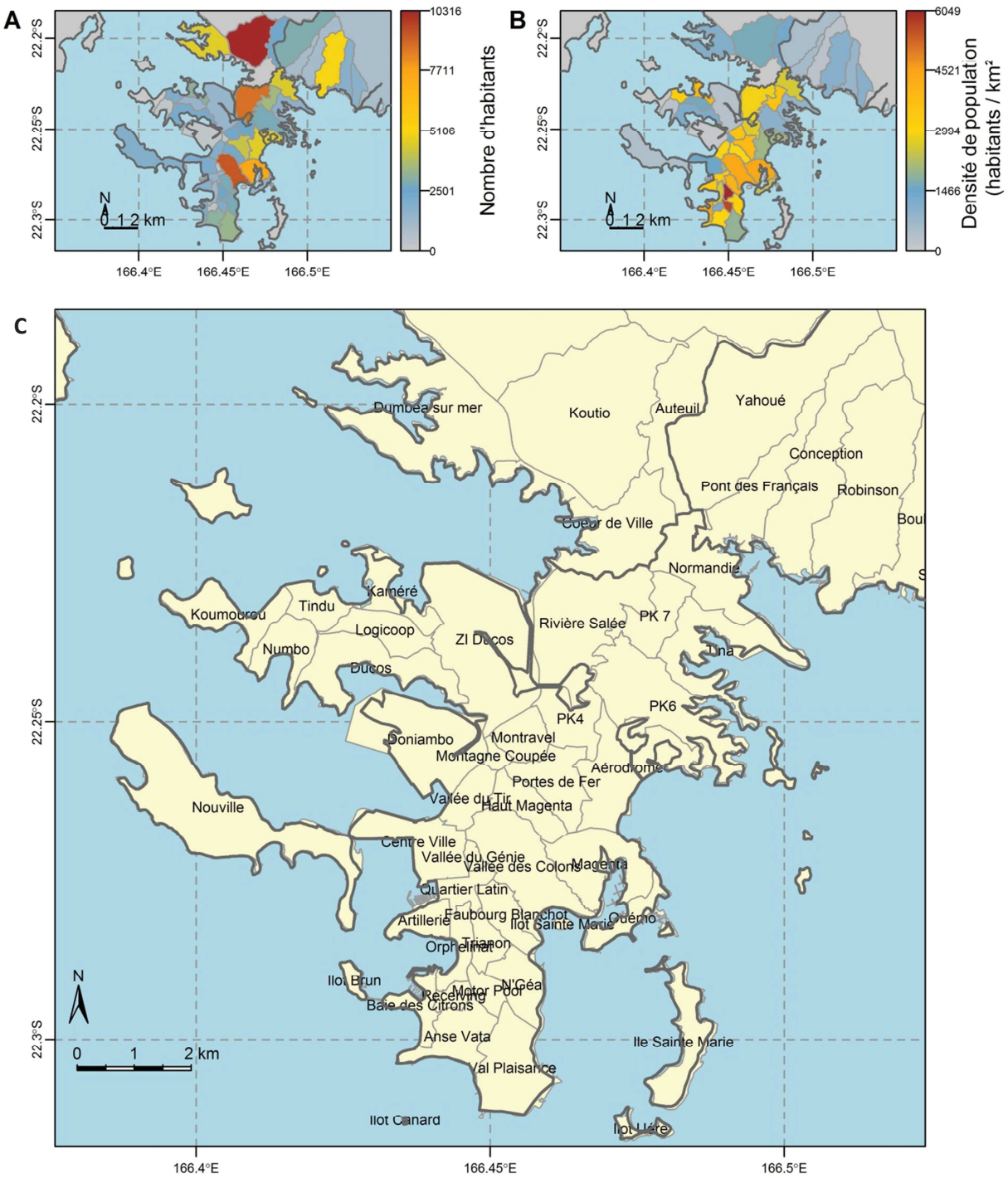


Figure 5.2 : Population (A), densité de population (B) et nom des quartiers (C) du Grand Nouméa. *Le grand Nouméa comprend les communes de Nouméa (Sud), Mont-Dore (Nord-Est) et Dumbéa (Nord-Ouest).*

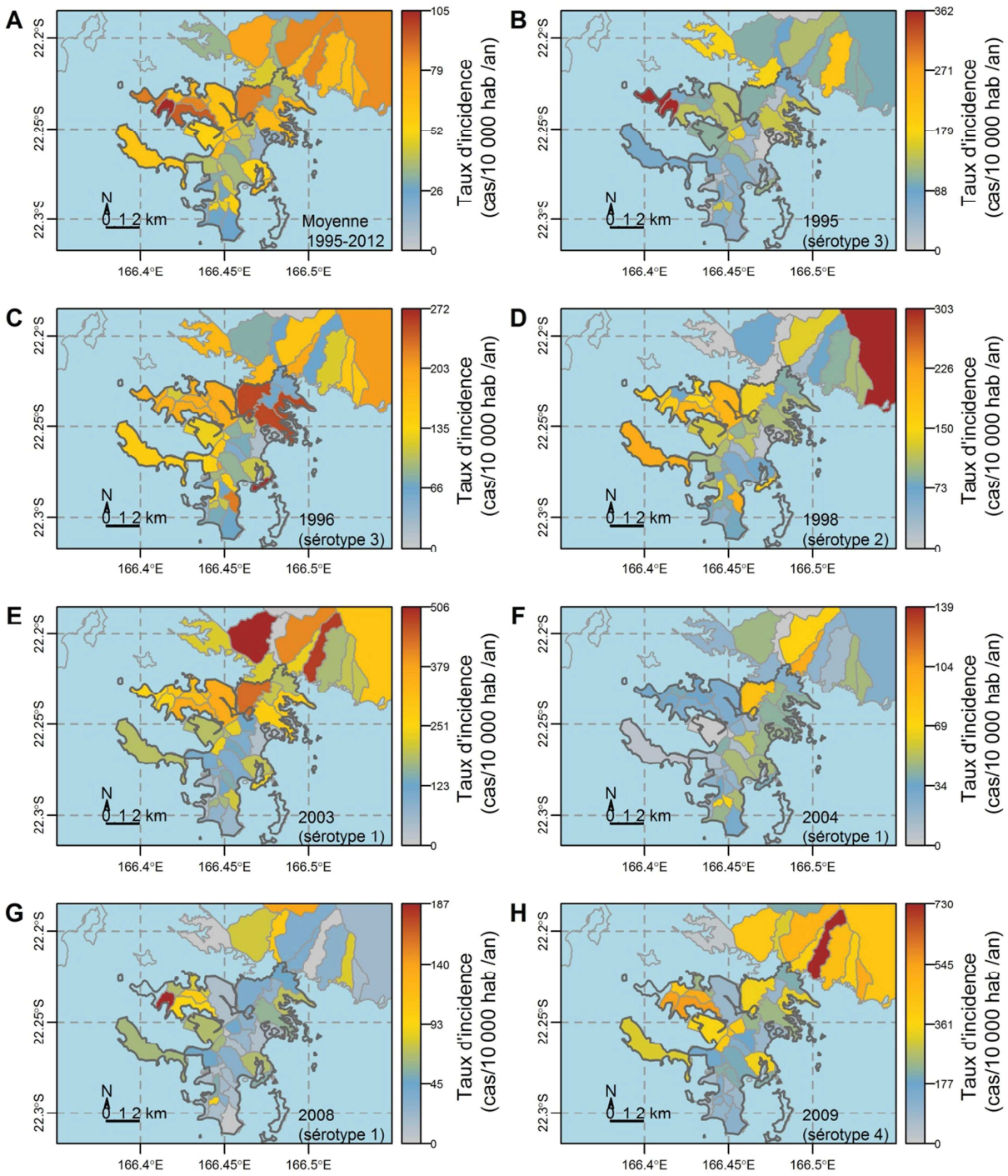


Figure 5.3 : Structure spatiale des épidémies de dengue à Nouméa, 1995-2012. ***A.*** *Carte de la moyenne des taux d'incidence annuels de la dengue entre 1995 et 2012 (* $\overline{\mathbf{x}}$ *dans les équations des EOF).* ***B, C, D, E, F, G et H.*** *Taux d'incidence annuels de la dengue au cours des années épidémiques (correspond à* $\mathbf{x}_{1995}$, $\mathbf{x}_{1996}$, $\mathbf{x}_{1998}$, $\mathbf{x}_{2003}$, $\mathbf{x}_{2004}$, $\mathbf{x}_{2008}$, $\mathbf{x}_{2009}$ *). Les taux d'incidence annuels ont été calculés en agrégeant les cas de septembre de l'année n-1 à août de l'année n. Le trait plus épais délimite la commune de Nouméa, par rapport à celles du grand Nouméa (Mont-Dore et Dumbéa) plus au Nord.*

5.1.2. Etude de l'évolution temporelle des schémas spatiaux de la dengue par la méthode des EOF : application à Nouméa

L'étude des cartes annuelles d'anomalie d'incidence (Figure 5.4) montre la faible amplitude des variations interannuelles (anomalies d'incidence relativement faibles par rapport à la moyenne), sauf en 2003 et 2009. Les quartiers présentant le plus souvent des taux d'incidence élevés sont Ducos et Rivière Salée (Nouméa), Koutio (Dumbéa) et Conception (Mont-Dore). En 2004 et 2008, la circulation virale a été d'autant plus faible que les quartiers avaient été très touchés par la dengue en 2003.

La Figure 5.5 montre les résultats de l'analyse EOF effectuée sur les quartiers de plus de 800 habitants. L'étude des valeurs propres, avec 87,2 % de la variance interannuelle expliquée par le premier axe montre que les variations interannuelles présentent une grande cohérence spatiale : les quartiers les plus touchés une année sont également ceux les moins touchés les autres années.

Le deuxième axe (6,85 % de variance expliquée) capture essentiellement la particularité spatiale de l'année 2009, à savoir des anomalies d'incidence plus faibles que les autres années pour les quartiers de Ducos, Rivière salée et Dumbéa sur Mer (produit $u_{jk}.a_{ik}$, négatif pour $k = 2$, $i = 2009$ et $j =$ Ducos, Rivière Salée, et Dumbéa Sur Mer, Figure 5.5.E et F).

Comme à l'échelle du territoire, l'essentiel de la variance spatio-temporelle est créée par une inversion du schéma spatial en 2004 et 2008, années de ré-émergence du sérotype 1, ce qui renforcerait l'hypothèse émise à l'échelle des communes que la variabilité inter-annuelle observée est essentiellement due à des phénomènes d'immunité de population. Par contre, nous n'observons pas cette inversion pour l'année 1996 par rapport à l'année 1995, alors que les deux épidémies sont causées par le même sérotype. Soit l'hypothèse émise est fausse, soit les taux d'incidence de 1995 ont été suffisamment faibles pour que l'immunité de population ne puisse pas se mettre en place. Comme pour l'échelle des communes, la création d'un modèle mathématique déterministe de type SIR pour tester cette hypothèse serait intéressante.

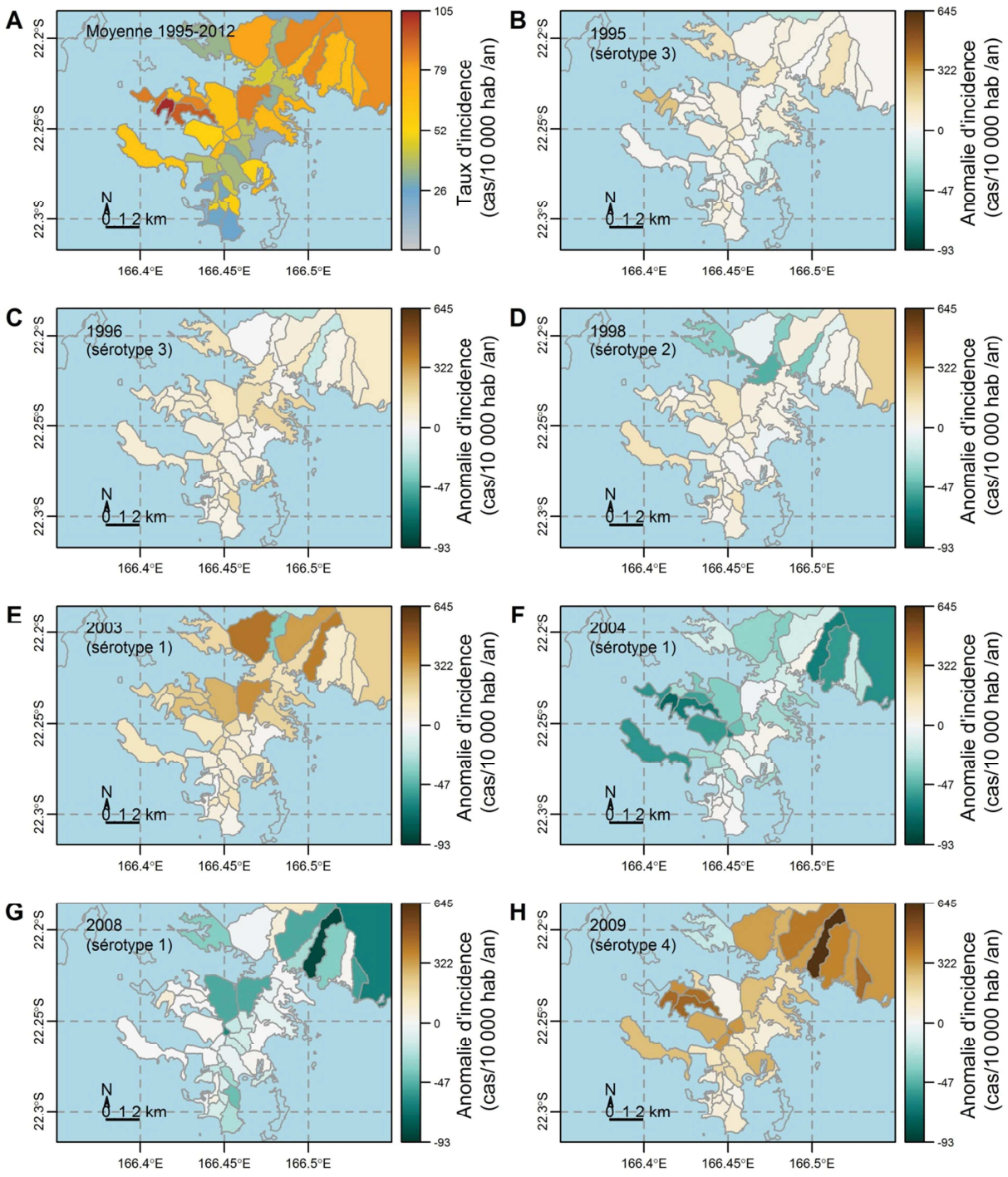


Figure 5.4 : Structure spatiale des anomalies des taux d'incidence annuels de la dengue à Nouméa au cours des années épidémiques de la période 1995-2012. A. *Carte de la moyenne des taux d'incidence annuels de la dengue entre 1995 et 2012 (* $\overline{\mathbf{x}} = \overline{x}_{.j}, j = 1, \ldots, p$, *dans les équations EOF).* ***B, C, D, E, F, G et H.*** *Anomalie du taux d'incidence annuel de la dengue par rapport à la moyenne des taux d'incidence de chacune des communes sur la période d'étude, au cours des années épidémiques (correspond à* $\mathbf{x}_t^c = \mathbf{x}_t - \overline{x}_{.j}$, $t \in \{1995, 1996, 1998, 2003, 2004, 2008, 2009\}$ *). Les taux d'incidence annuels ont été calculés en agrégeant les cas de septembre de l'année n-1 à août de l'année n.*

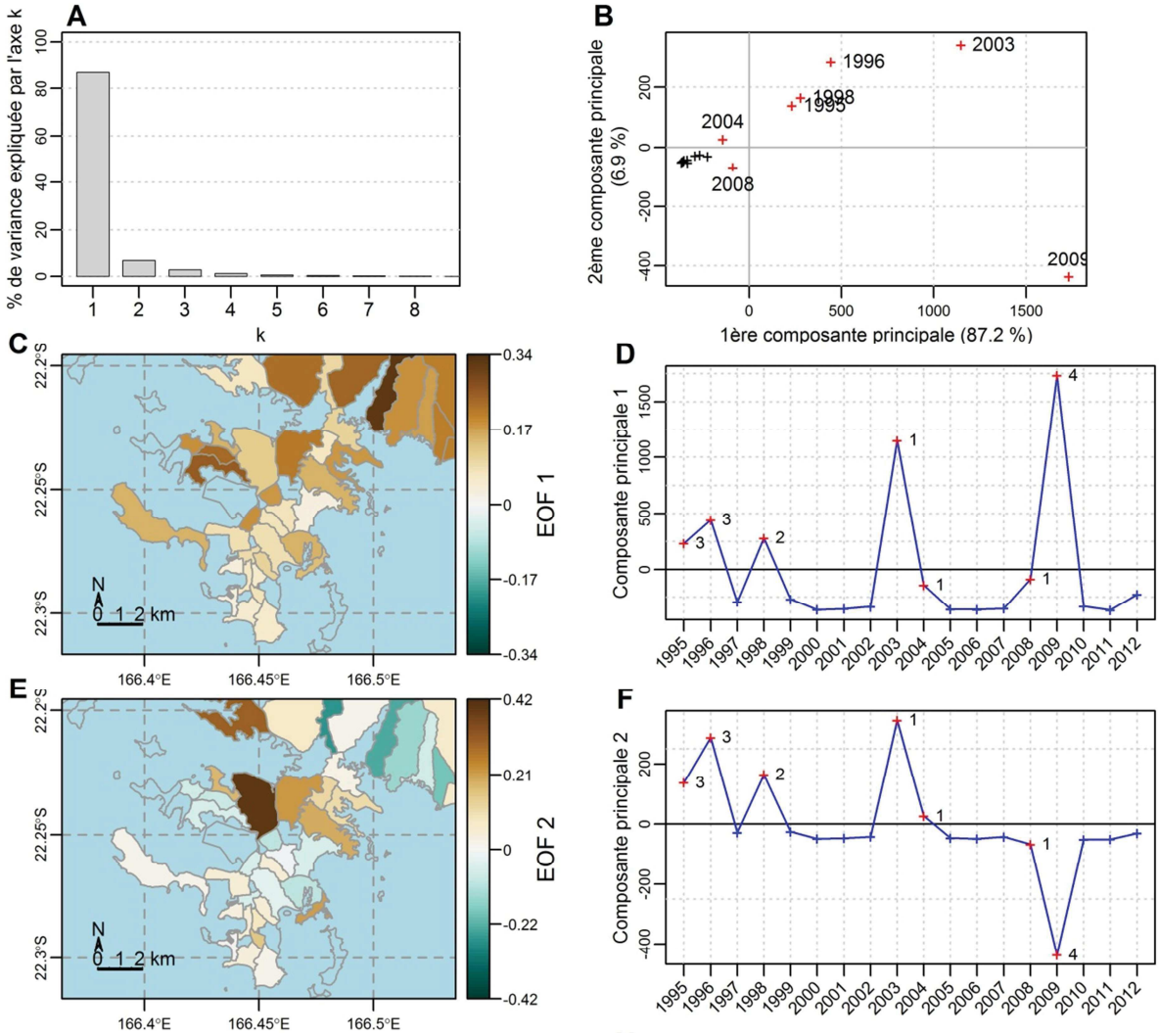


Figure 5.5 : Analyse EOF des taux d'incidence annuels de la dengue à Nouméa, période 1995-2012. A. *Valeurs propres (λ_k, $k = 1, \ldots, p$, dans les équations).* ***B.*** *Projection des cartes d'anomalies de taux d'incidence dans le plan formé par les deux premiers axes principaux, i.e les deux premières composantes principales ($\mathbf{a}_k = X^c\mathbf{u}_k = (a_{1k}, \ldots, a_{nk}), k \in \{1,2\}$).* ***C à F.*** *Cartes des deux premiers EOF ($\mathbf{u}_k = (u_{1k}, \ldots, u_{pk})$, k=1 en C et $k = 2$ en E) et composantes principales associées (D et F pour $k = 1$ et 2 respectivement). Sur les panneaux B, D et F, les croix rouges représentent les années épidémiques, et les croix noires ou bleues les années non épidémiques. Sur les paneaux D et F, les chiffres indiquent le sérotype dominant au cours de l'épidémie.*

5.2. Analyse temporelle

5.2.1. Modélisation déterministe du risque épidémique

5.2.1.a. Objectifs du modèle

L'analyse temporelle à l'échelle du territoire a montré que le risque annuel d'émergence d'une épidémie nécessitait des conditions climatiques précises, ce qui laisse penser que les facteurs entomologiques ont une importance prépondérante dans l'émergence des épidémies. Le risque inter-annuel d'émergence à l'échelle du territoire est le même qu'à Nouméa. Or à Nouméa, l'analyse des données entomologiques n'a montré aucune corrélation inter-annuelle entre la densité vectorielle et le risque d'émergence (voir article 1 (Descloux et al., 2012)). En effet, les indices entomologiques représentatifs des densités vectorielles ne sont pas plus élevés au cours des années épidémiques qu'au cours des années non épidémiques (Figure 5.6).

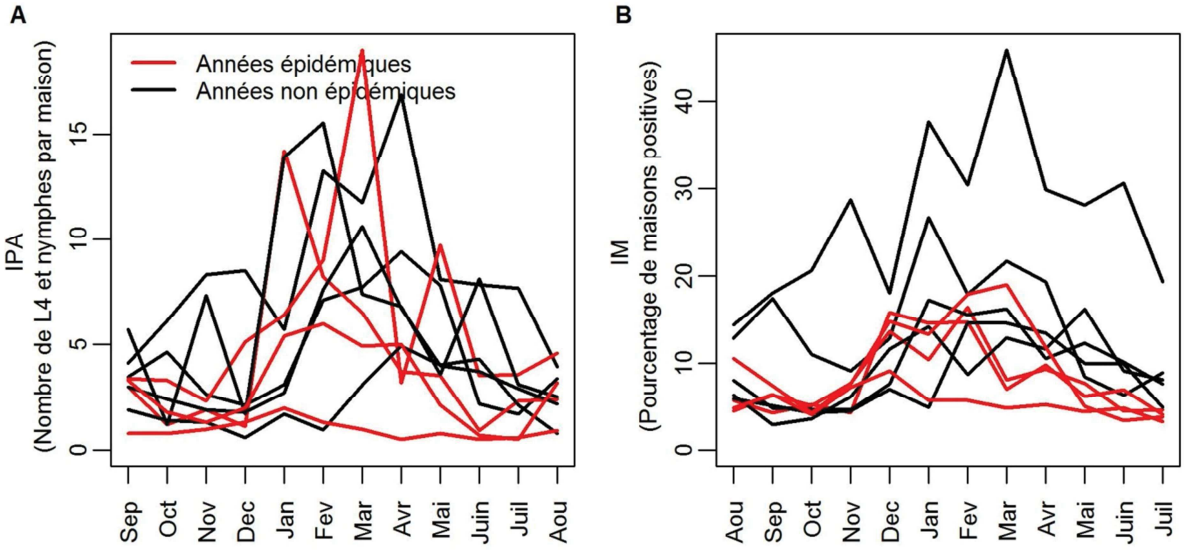


Figure 5.6 : Indices entomologiques à Nouméa : comparaison des années épidémiques et non épidémiques (2000-2011). ***A****. Indice de productivité d'adultes (IPA = nombre de Larves de stade 4 (L4) et de nymphes divisé par le nombre de maisons visitées).* ***B.*** *Indice maison (IM = nombre de maisons infestées divisé par le nombre de maisons visitées * 100). Les indices se rapportent uniquement à Aedes aegypti. Les années épidémiques (en rouge) sont les années 2003, 2004, 2008 et 2009. Elles vont de septembre de l'année n-1 à août de l'année n.*

Dans cette partie, nous avons cherché à comprendre par quels mécanismes le climat influence le risque épidémique annuel. Nous avons pour cela construit un modèle permettant de prédire le risque d'émergence épidémique en prenant en compte les paramètres du cycle de transmission autres que les densités vectorielles (*i.e.* paramètres du cycle de vie du moustique, et paramètres du cycle de transmission du virus). Ce modèle a pour cible le nombre de reproduction de base, dont la valeur seuil de « 1 » permet de distinguer l'existence et l'absence de risque d'émergence.

5.2.1.b. Construction du modèle

Les modèles existants du même type que celui construit lient la capacité vectorielle ou le nombre de reproduction de base aux paramètres et variables entomologiques, dont les densités vectorielles (Degallier et al., 2010; Githeko et al., 2000; Jetten and Focks, 1997). Ces modèles sont souvent utilisés pour établir des cartes mondiales du potentiel épidémique en fixant la capacité vectorielle ou le nombre de reproduction de base à 1 afin d'estimer des seuils critiques de densité vectorielle nécessaire au développement d'une épidémie. Plus la densité critique est faible, plus le potentiel épidémique est élevé. Dans notre étude, nous avions à disposition les données observées de l'indice de productivité d'adultes (IPA), collecté de manière mensuelle à Nouméa de mars 2000 à décembre 2009. Nous avons donc construit un modèle qui a pour cible le nombre de reproduction de base, et pour variables d'entrée le nombre de nymphes par habitant sur Nouméa, ainsi que les variables climatiques influençant le cycle de transmission (température et humidité relative). Dans ce modèle, à chaque pas de temps (mensuel ou annuel, au choix), nous considérons l'ensemble de la population d'hôtes de Nouméa comme susceptible, et essayons d'estimer l'effet de l'introduction d'une seule personne infectieuse dans cette population. Le but du modèle est alors de connaître le nombre de cas secondaires créés à partir du cas index (*i.e.* le nombre de reproduction de base *R0*), connaissant la densité vectorielle sur Nouméa, ainsi que les conditions météorologiques.

Description des étapes

La Figure 5.7 schématise les étapes du cycle de transmission prises en compte dans la construction du modèle : à partir de l'indice entomologique IPA (nombre de L4 et de nymphes / maison visitée), nous avons calculé un nombre de nymphes par habitant *N*. Ces nymphes mettent un certain nombre de jours à se développer (*Tn*) pour finalement émerger avec un taux de survie à l'émergence *Sn*. Parmi les nymphes qui émergent, une proportion *r* sont des femelles. Ces femelles adultes meurent selon un taux de mortalité μ, indépendant de l'âge de la femelle et mettent un certain temps à compléter leurs cycles gonotrophiques (*Tg*). En supposant qu'il n'y a qu'une seule personne infectieuse parmi la population d'hôtes susceptibles (*NH)*, et connaissant le nombre moyen de personnes piquées par repas sanguin par un moustique (*k*) ainsi que la probabilité qu'un moustique sain s'infecte s'il pique une personne infectée (*pa*), nous pouvons calculer, pour chaque femelle, la probabilité de s'infecter lors d'un repas sanguin. Dans les équations, nous avons cherché à prendre en compte le fait qu'une femelle peut s'infecter au premier cycle gonotrophique, ou à des cycles ultérieurs, une femelle infectée tardivement ayant la possibilité de transmettre le virus sur des périodes plus courtes qu'une femelle infectée précocement au cours de sa vie. Une fois la femelle infectée, il faut un certain temps avant qu'elle ne puisse de nouveau transmettre le virus (durée d'incubation extrinsèque, *Teip*). A partir du moment où une femelle a survécu à la durée d'incubation extrinsèque et devient infectieuse, elle peut potentiellement transmettre le virus à chaque repas sanguin (séparés dans le temps par la durée du cycle gonotrophique) avec une probabilité *pb*. Le nombre de personnes susceptibles qu'une femelle infectieuse va pouvoir infecter au cours de sa vie peut être calculé en connaissant le nombre de personnes piquées par repas sanguin (*k*),

ainsi que la probabilité qu'un moustique infectieux transmette le virus à un hôte sain lors d'une piqure (*pb*), tout en prenant en compte le nombre de cycles gonotrophiques que cette femelle va pouvoir compléter avant de mourir.

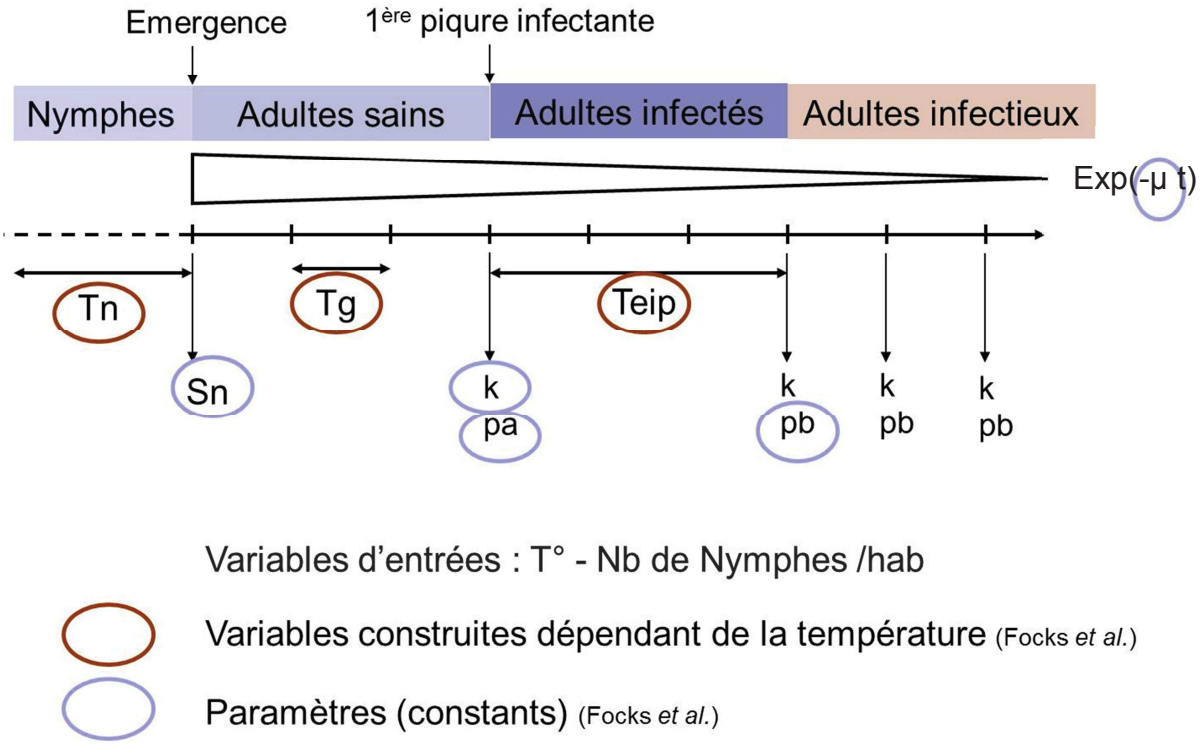


Figure 5.7 : Description des étapes de la construction du modèle entomologique du risque épidémique à Nouméa. *Le bandeau supérieur note le stade de développement des moustiques. La flèche décroissante juste en dessous représente la population de moustiques femelles adultes vivantes, décroissant au cours du temps selon une loi exponentielle (taux de mortalité supposé constant, exp(-μt) représente donc la proportion de femelles qui survivent jusqu'au temps t). La flèche graduée représente l'axe du temps, chaque graduation dénotant un repas sanguin (i.e. complétion d'un cycle gonotrophique). Les différents paramètres et variables sont décrits dans les tableaux ci-après.*

Les paramètres de survie à l'émergence (*Sn*), de mortalité (*μ*) ainsi que les paramètres liés à la probabilité de transmission virale lors d'un repas sanguin (*k*, *pa* et *pb*) ne sont pas connus en Nouvelle-Calédonie, où il manque un laboratoire entomologique possédant des chambres thermiques pour pouvoir les estimer. Nous avons donc pris les valeurs trouvées dans la littérature (Focks et al., 1993b).

L'influence de la température sur le cycle de transmission est prise en compte au travers de son influence sur la durée des cycles (développement des nymphes, cycle gonotrophique et durée d'incubation extrinsèque) ainsi que sur le paramètre de mortalité des adultes. A chaque pas de temps, le calcul de la durée des différents cycles est basé sur des équations de cinétique enzymatique décrites par ailleurs (Focks et al., 1993b). Le taux de mortalité μ, supposé constant et indépendant de l'âge des femelles, reste constant pour des températures allant de 2 °C à 38 °C, et augmente brutalement jusqu'à 100 % de mortalité pour des températures inférieures à 0 °C ou supérieures à 40 °C (Focks et al., 1993b). L'influence de l'humidité relative sur le cycle de transmission n'est prise en compte qu'au

travers de son influence sur la mortalité des adultes. Pour les valeurs seuils utilisées à partir desquelles température et humidité influencent la mortalité des adultes (Focks et al., 1993b), sur l'ensemble de la période d'étude (2000 – 2011), l'humidité relative a toujours été suffisamment élevée à Nouméa et la température toujours comprise entre 2 °C et 38 °C. Au final, dans le modèles construit, ces variables n'influencent donc pas les taux de mortalité en Nouvelle-Calédonie, et le modèle possède donc deux variables d'entrée : l'IPA et la température (via son influence sur la durée des cycles).

Paramètres, variables et équations

La Table 5.1 présente les paramètres et variables du modèle entomologique du risque épidémique. Le taux de mortalité est supposé indépendant de l'âge du moustique, mais il est ici classé dans les variables construites car nous y avons intégré l'effet de la température et de l'humidité relative sur la survie du moustique (Focks et al., 1993b).

Le nombre de reproduction de base de la dengue (*R0*) est calculé à chaque pas de temps en multipliant la durée de virémie de l'hôte infectieux introduit dans la population (fixée à 6 jours) par le nombre de moustiques infectieux que cet hôte va « produire » par jour (noté « *A* », voir Table 5.2), et par le nombre de personnes qu'un moustique infectieux peut infecter durant le reste de sa vie (noté « B », voir Table 5.3).

Table 5.1 : Paramètres et variables du modèle

VARIABLES OBSERVEES	IPA		Nombre L4 + Nymphes / maison visitée
	N		Nombre de nymphes /habitant = IPA/4 / nb d'habitant par maison visitée *
	TM		Température quotidienne moyenne de la station de Nouméa (= pour chaque pas de temps, moyenne entre la température minimale et la température maximale)
PARAMETRES	NH	100000	Nombre d'Hôtes dans la population (nb d'habitants sur Nouméa)
	Sn	0,83	% de nymphes qui survivent à durée du cycle nymphe + émergence
	r	0,5	% de femelles parmi les adultes qui émergent
	Tv	6	Durée de la virémie (en jours)
	k	1,8	Nb de personnes différentes piquées /moustique /repas complet
	pa	0.9	Probabilité qu'un moustique s'infecte sachant que le moustique a piqué une personne infectieuse
	pb	0.9	Probabilité qu'un hôte s'infecte sachant que l'hôte a été piqué par un moustique infectieux
VARIABLES CONSTRUITES	Tn		Durée du stade nymphal
	Tg		Durée du cycle gonotrophique
	Teip0		Durée d'incubation extrinsèque vraie
	Teip		Durée d'incubation extrinsèque effective = durée entre le moment où le moustique s'infecte et le 1er cycle gonotrophique suivant la période d'incubation extrinsèque (= *Teip0* + durée pour compléter le cycle gonotrophique en cours)
	μ_A	0,09	Taux de mortalité des adultes (le taux de mortalité de 0,09 /jour est ensuite modulé par la température et l'humidité relative)

* Le nombre *N* de nymphes par habitant a été estimé à partir de l'IPA. A partir de janvier 2009, le réseau de surveillance entomologique a également collecté un indice du nombre de nymphes par maison (INM). Une régression linéaire entre l'IPA et l'INM sur la période janvier 2009 – juillet 2010 a permis d'estimer que l'INM est égal à ¼ de l'IPA. D'autre part, le nombre d'habitants par maison est estimé à 3,37 (source : URE-EnM IPNC). D'où N = IPA / (4 * 3,37).

Table 5.2 : Etapes du calcul du nombre de moustiques infectieux produits par jour et par hôte infecté (noté *A*)

$N.\,Sn$	Nombre de moustiques adultes émergeant pendant la durée du stade nymphal
$N.\,\frac{Sn}{Tn}$	Nombre de moustiques adultes émergeant /jour
$N.\,\frac{Sn}{Tn}.\,r$	Nombre de femelles adultes émergeant / jour
$N.\,\frac{Sn}{Tn}.\,r.\sum_{n=1}^{\infty} p\,(X=n)$	Nombre de femelles adultes s'infectant / jour / hôte infecté ; Chaque femelle peut s'infecter au cycle *n, n=1…infini*.
$p\,(X=n)=\exp(-\mu_A\cdot Tg\cdot n)\cdot\left(1-\frac{k\cdot pa}{NH}\right)^{n-1}\cdot\frac{k\cdot pa}{NH}$	Proportion de femelles s'infectant (pour la première fois) au cycle *n* (avec taux de prévalence = 1/*NH*, fixé par la définition du *R0*) = proportion de femelles qui survivent jusqu'au cycle *n* * probabilité qu'une femelle ne s'infecte pas pendant les *n-1* premiers cycles * probabilité qu'elle s'infecte au cycle *n*
$A=N.\,\frac{Sn}{Tn}.\,r.\exp(-\mu_A\cdot Teip)\sum_{n=1}^{\infty} p\,(X=n)$	Nombre de femelles adultes infectieuses produites / jour /hôte infecté (*i.e.* les femelles infectées survivent à la durée d'incubation extrinsèque *Teip*, la proportion de femelles infectées survivant à *Teip* étant égale à *exp(-μTeip)*)
$N.\,\frac{Sn}{Tn}.\,r.\exp(-\mu_A\cdot Teip)\sum_{n=1}^{\infty}\left(p\,(X=n).\,B\right)$	Chaque femelle qui s'est infectée au cycle *n* va pouvoir infecter un nombre *B* d'hôtes durant le reste de sa vie. Ainsi cette équation représente le nombre d'hôtes infectés produits / jour /hôte infectieux.
$R0=Tv.\,N.\,\frac{Sn}{Tn}.\,r.\exp(-\mu_A\cdot Teip)\sum_{n=1}^{\infty}\left(p\,(X=n).\,B\right)$	En multipliant l'équation précédente par la durée de la virémie, on obtient le nombre d'hôtes infectés produits / hôte infectieux, c'est-à-dire le *R0*.

Table 5.3 : Etapes du calcul du nombre de personnes infectées « produites » par moustique infectieux qui s'est infecté au cycle *n* (noté *B*)

$\frac{\exp\left[-\mu_A\cdot(n\cdot Tg+Teip)\right]}{\mu_A}$	Espérance de vie d'un moustique infectieux qui s'est infecté au cycle *n* (*i.e.* qui a déjà survécu à *n* cycles gonotrophiques et à la durée d'incubation extrinsèque)
$\frac{\exp\left[-\mu_A\cdot(n\cdot Tg+Teip)\right]}{\mu_A\cdot Tg}+1$	Espérance du nombre de cycles gonotrophiques complétés /moustique infectieux qui s'est infecté au cycle *n*
$B=\left(\frac{\exp\left[-\mu_A\cdot(n\cdot Tg+Teip)\right]}{\mu_A\cdot Tg}+1\right)\cdot k\cdot pb$	Nombre moyen de personnes infectées par moustique infectieux qui s'est infecté au cycle *n*

Ce qui conduit à l'expression suivante du *R0*, à partir de deux variables (IPA et température, l'humidité étant toujours suffisamment élevée pour ne pas influencer la mortalité des vecteurs en Nouvelle-Calédonie) :

$$R0 = Tv \cdot \frac{N}{NH} \cdot \frac{Sn}{Tn} \cdot r \cdot \exp(-\mu_A \cdot Teip) \cdot k^2 \cdot p_a \cdot p_b \cdot$$
$$\sum_{n=1}^{\infty} \left[\exp(-\mu_A \cdot n \cdot Tg) \cdot \left(1 - \frac{k \cdot pa}{NH}\right)^{n-1} \cdot \left(\frac{\exp[-\mu_A \cdot (n \cdot Tg + Teip)]}{\mu_A \cdot Tg} + 1 \right) \right]$$

Ou encore :

$$R0 = Tv \cdot \frac{IPA}{4 \cdot 3{,}37 \cdot NH} \cdot \frac{Sn}{Tn} \cdot r \cdot \exp(-\mu_A \cdot Teip) \cdot k^2 \cdot p_a \cdot p_b \cdot$$
$$\sum_{n=1}^{\infty} \left[\exp(-\mu_A \cdot n \cdot Tg) \cdot \left(1 - \frac{k \cdot pa}{NH}\right)^{n-1} \cdot \left(\frac{\exp[-\mu_A \cdot (n \cdot Tg + Teip)]}{\mu_A \cdot Tg} + 1 \right) \right]$$

Avec *Tn*, *Tg* et *Teip* des fonctions de la température (Focks et al., 1993a).

5.2.1.c. Exploration du modèle

Durée des cycles

A Nouméa, la durée des cycles au cours de l'année, simulée à partir des températures est la suivante :

- Cycle de développement des larves : de 4 jours en été à 11 jours en hiver ;
- Cycle de développement des nymphes *Tn* : de 2 à 4 jours ;
- Cycle gonotrophique *Tg* : de 2 à 4,5 jours ;
- Durée d'incubation extrinsèque effective *Teip* : de 10 à 22 jours (toujours égale à 4 fois la durée du cycle gonotrophique).

La durée de ces cycles varie au cours du temps suivant les variations de température. Les extrêmes donnés correspondent aux valeurs durant l'été (cycles courts) ou l'hiver (cycles longs). Nous pouvons remarquer ici que le nombre seuil de moustiques nécessaires au développement d'une épidémie varie donc également au cours du temps de manière saisonnière : plus les cycles sont courts (été), plus ce seuil est bas.

Nombre de cas secondaires suite à l'introduction d'un cas dans une population entièrement susceptible

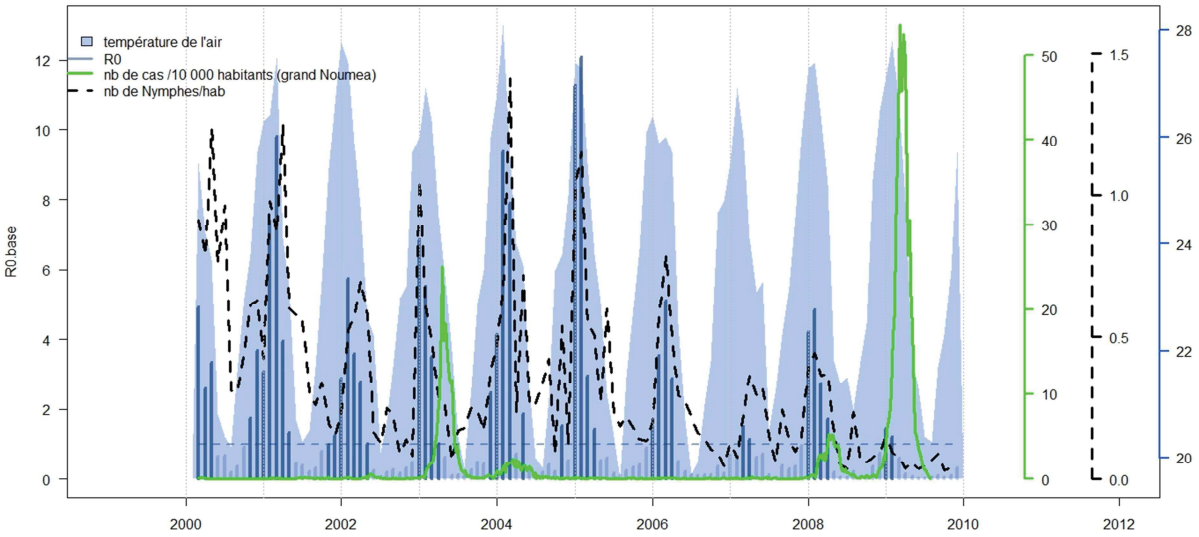


Figure 5.8 : Résultats du modèle entomologique du risque d'émergence épidémique à Nouméa. *L'abscisse est le temps, et l'ordonnée la valeur des différentes variables reproduites. Sont représentés : i et ii) les deux variables d'entrée du modèle, i.e. température mensuelle moyenne (station de Nouméa) en bleu plein, et le nombre de Nymphes par habitant (= IPA/(4*3,37)) en noir pointillé ; iii) la variable de sortie du modèle, .i.e. le R0, sous la forme de barres verticales pour rappeler que le modèle est réinitialisé à chaque pas de temps, sans tenir compte des pas de temps précédents, ni du développement éventuel d'une épidémie qui modifierait la proportion d'hôtes susceptibles dans la population. Les barres sont en bleu clair si R0 <1 et en bleu foncé si R0 >1. La limite R0 = 1 est matérialisée par la ligne pointillée horizontale ; iv) les taux d'incidence hebdomadaires de la dengue sur le Grand Nouméa, en vert.*

La Figure 5.8 présente les résultats de la modélisation. Lors des années épidémiques, le premier pas de temps au cours duquel le *R0* devient supérieur à 1 (favorable au développement d'une épidémie) correspond bien à la période de démarrage des épidémies observées. Sur le plan saisonnier, le modèle arrive à reproduire la période la plus à risque (décembre à avril), démontrant ainsi l'influence du climat et de la densité de vecteurs sur la saisonnalité des épidémies.

Le modèle indique également que d'après les indices entomologiques observés et les températures moyennes à Nouméa, le risque d'émergence épidémique est présent chaque année : ce modèle déterministe n'arrive pas à prévoir la variabilité inter-annuelle du risque épidémique. D'autre part le niveau du *R0* (lui-même lié au nombre total de cas à la fin d'une épidémie (Woolhouse and Gowtage-Sequeria, 2005)) ne correspond pas aux niveaux d'incidence observés durant les épidémies : 2009 est l'année où le risque d'émergence épidémique lié à l'entomologie est le plus faible, et où l'épidémie a été de plus forte ampleur.

Seuil entomologique nécessaire au développement d'une épidémie

De la même manière que ceux qui utilisent le modèle pour établir des cartes mondiales du risque d'émergence (Degallier et al., 2010; Jetten and Focks, 1997; Patz et al., 1998), il est possible de calculer les valeurs seuils de l'IPA nécessaires au développement d'une épidémie en fonction de la température (Figure 5.9). D'après ces seuils, une épidémie aurait du mal à se développer lorsque la température moyenne est inférieure à 19 °C. Pour une température de 28 °C comme régulièrement observé en été, le seuil d'IPA en dessous duquel une épidémie ne pourrait pas se développer est de 1. Cependant ce graphe est à interpréter avec prudence, du fait que le modèle décrit ne permet pas de prédire le risque épidémique réellement observé.

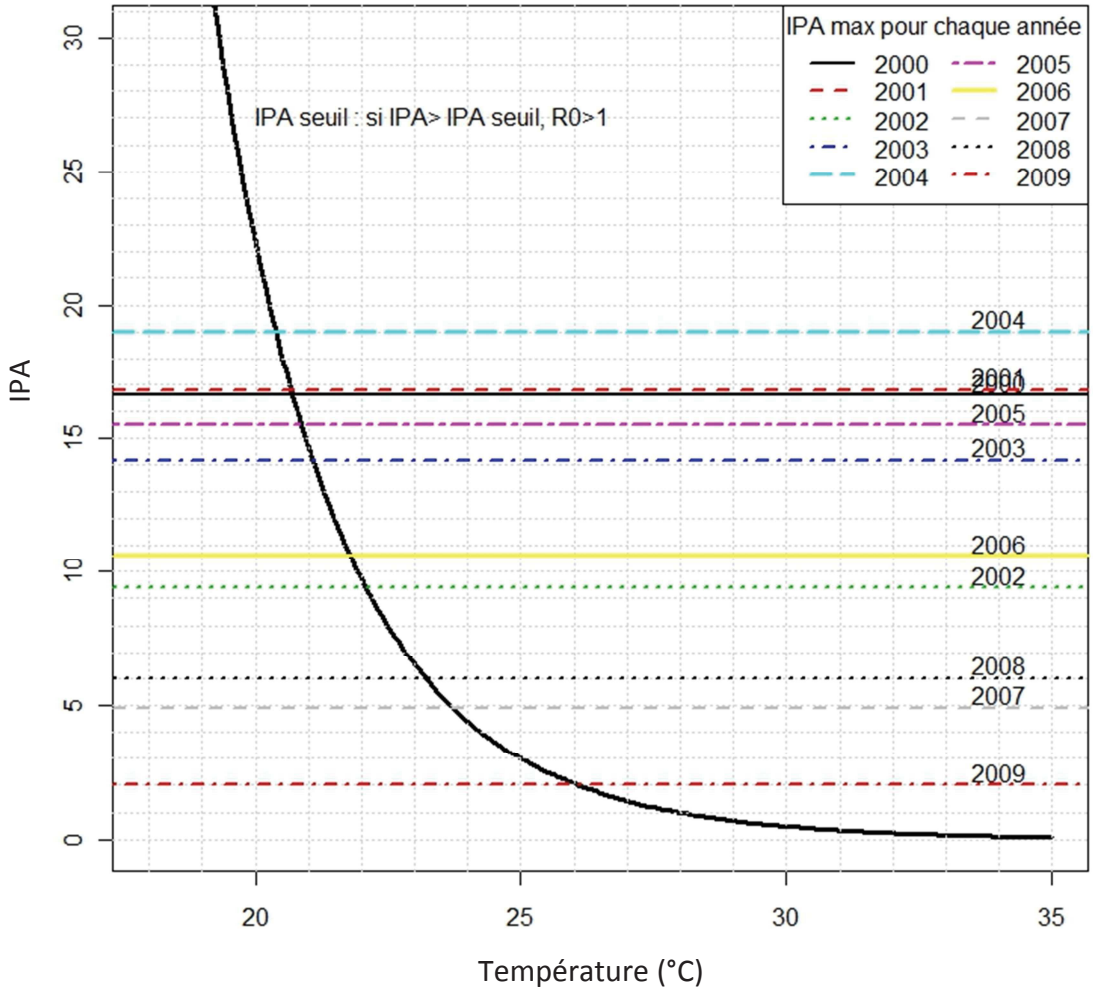


Figure 5.9 : Seuils entomologiques nécessaires au développement d'une épidémie à Nouméa en fonction de la température. *La ligne noire pleine représente la valeur de l'indice IPA nécessaire au développement d'une épidémie (i.e. R0>1). Les lignes horizontales représentent le maximum de l'indice IPA réellement enregistré chaque année. Nous voyons que même en 2009, où les indices entomologiques étaient très bas, le nombre de L4 et de nymphes était suffisant au développement d'une épidémie pour les températures habituellement observées en été (28 à 30 °C).*

5.2.1.d. Relation entre le *R0*, les indices entomologiques et la température, et analyse de sensibilité

Afin d'essayer de comprendre pourquoi le modèle n'arrive pas à reproduire la variabilité inter-annuelle du risque d'émergence des épidémies de dengue, nous avons cherché à étudier :

- La forme de la relation liant les variables d'entrée au *R0*. En effet, le modèle statistique a montré que le lien entre risque d'émergence et température n'est pas linéaire, mais qu'il y a des phénomènes d'effet seuil (voir article 1 (Descloux et al., 2012)) ;
- L'effet de l'augmentation de la résolution temporelle du modèle (utilisation de la température moyenne quotidienne) afin de prendre en compte la persistance de conditions climatiques favorables au développement des épidémies, puisque que l'une des variables statistiques identifiée comme prédictive du risque d'émergence est le nombre de jour où la température maximale dépasse 32 °C ;
- La sensibilité du modèle aux différentes valeurs des paramètres.

Influence des variables d'entrées sur le R0 calculé

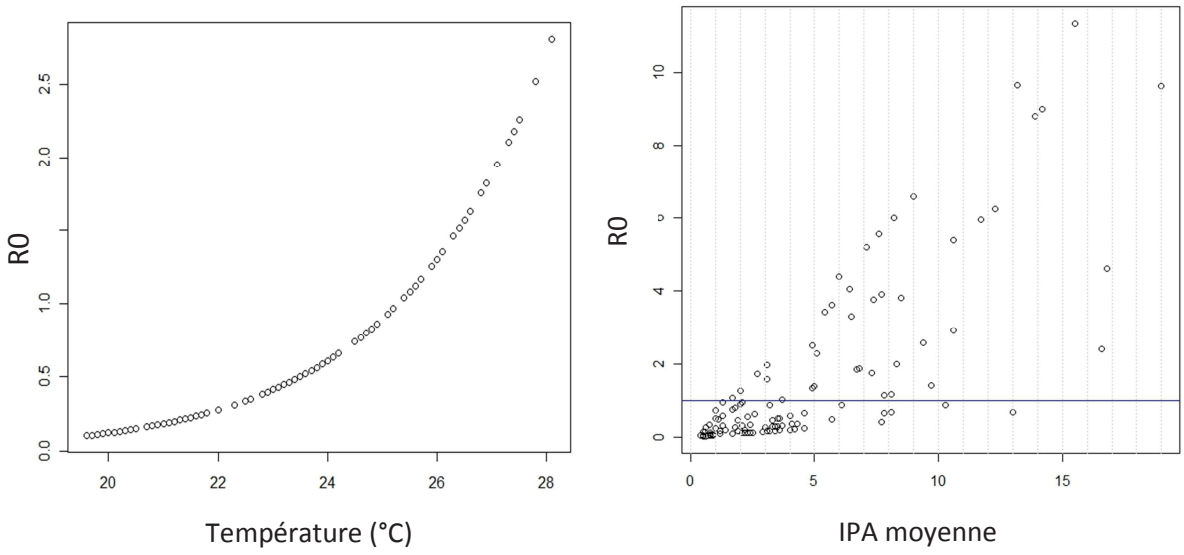


Figure 5.10 : Relation entre le R0 calculé et les variables d'entrée du modèle. A. *Valeur de R0 en fonction de diverses valeurs de température. La variable entomologique d'entrée N est ici fixée de manière arbitraire à 0,2 nymphes / habitants.* ***B.*** *Valeurs de R0 en fonction des valeurs observées de l'indice entomologique IPA, les variations inter-annuelles de la température ayant été enlevées en calculant une année-type de température moyenne. La variabilité des valeurs de R0 pour chacune des valeurs d'IPA provient de la variabilité saisonnière de l'année-type des températures.*

La Figure 5.10 montre le lien entre le *R0* calculé et chacune des deux variables d'entrée du modèle, l'autre variable ayant été fixée. Nous pouvons voir que les équations construites conduisent à des liens

dans lesquels il n'y a pas d'effet seuil. Le lien entre le *R0* et la température moyenne est exponentiel, le lien entre le *R0* et l'IPA est quasi linéaire.

La température moyenne influence le calcul du *R0* via son influence sur la durée des cycles. La Figure 5.11 montre que le cycle sur lequel la température a le plus d'effet est la durée d'incubation extrinsèque de la maladie. Les équations utilisées n'introduisent aucun effet seuil dans la durée des cycles.

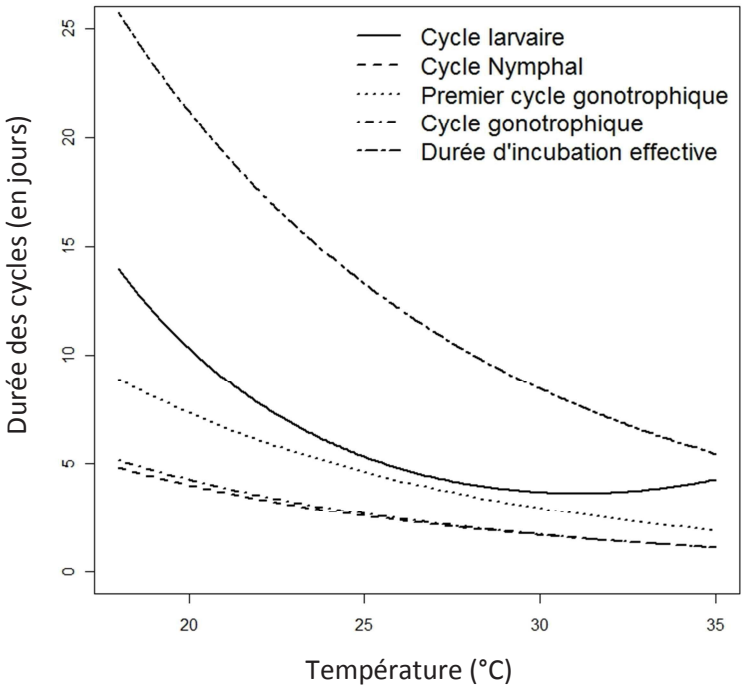


Figure 5.11 : Influence de la température sur les cycles de vie d'Aedes aegypti.

Effet du changement de résolution temporelle de la température sur la variabilité inter-annuelle du *R0*

L'utilisation de données quotidiennes, ainsi que l'introduction d'effets seuils de manière artificielle pour casser la continuité du lien entre température et *R0* (par exemple durée d'incubation extrinsèque réduite à 1 jour pour des températures dépassant 32 °C) n'ont pas permis de retrouver la variabilité inter-annuelle observée du risque d'émergence épidémique annuel.

Analyse de sensibilité du modèle aux différents paramètres

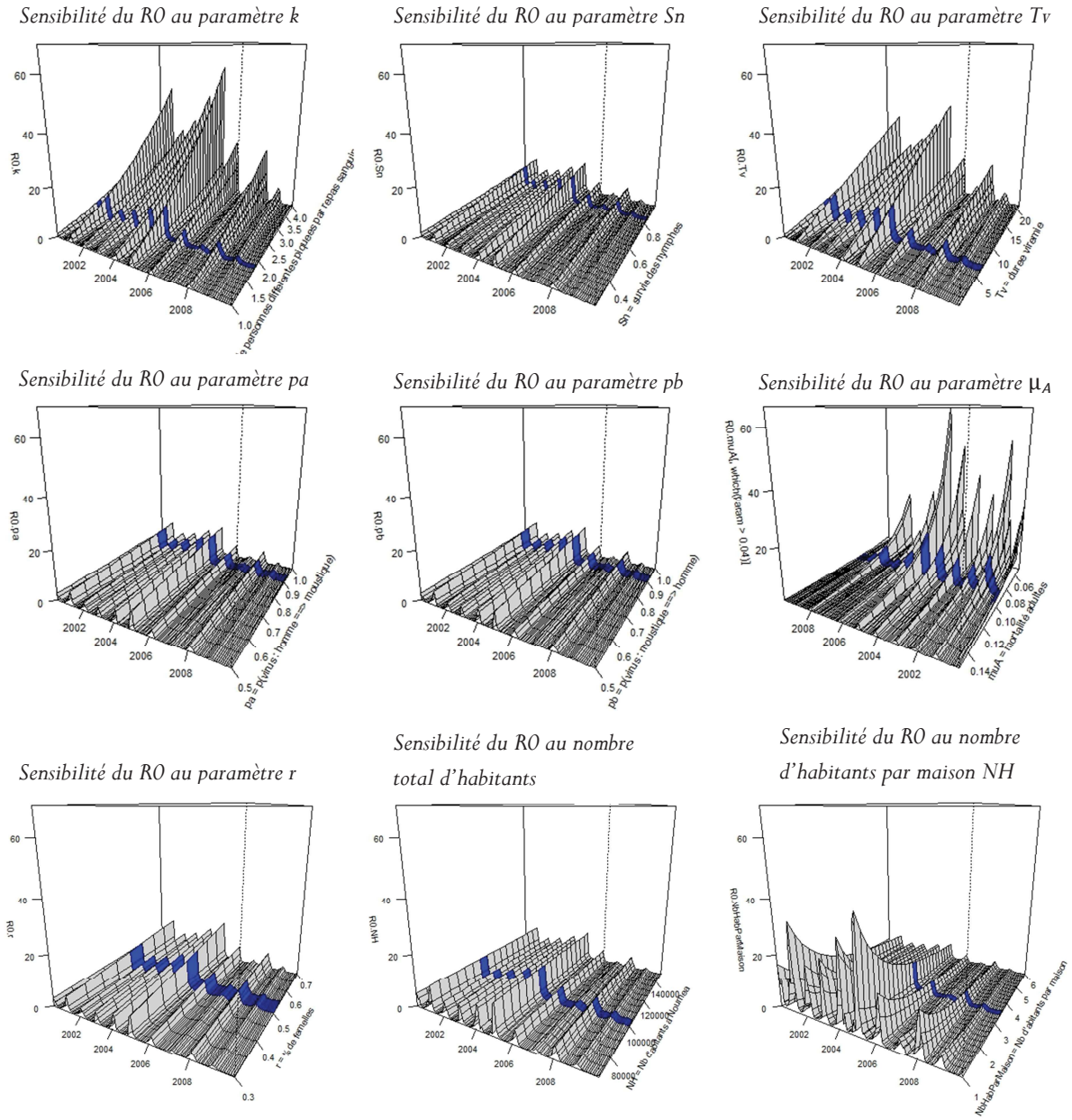


Figure 5.12 : Analyse de sensibilité du modèle aux différentes valeurs des paramètres. *Pour chacun des paramètres du modèle, l'axe vertical représente le R0 calculé. Tous les axes verticaux sont à la même échelle. L'axe des x représente le temps, de 2000 à gauche à 2009 à droite. L'axe des profondeurs représente les différentes valeurs de paramètre testées. La ligne bleue représente le calcul du R0 pour la valeur du paramètre utilisée jusqu'ici (voir Table 5.1).*

La Figure 5.12 montre les résultats de l'analyse de sensibilité du modèle aux paramètres. Globalement, le modèle est peu sensible aux paramètres influençant la densité vectorielle (pourcentage de moustiques femelles *r*, survie des nymphes *Sn*), au nombre total d'habitants sur Nouméa (*NH*), ainsi qu'aux

probabilités de transmission virale lors d'un repas sanguin entre un hôte infecté et un moustique sain (*pa*) ou l'inverse (*pb*). Le paramètre pour lequel le *R0* est le plus sensible est la mortalité. Pour une température moyenne de 26 °C, une baisse du taux de mortalité de 0,09 moustiques/jour à 0,08 moustiques/jour suffit à multiplier le *R0* calculé par 1,45, tandis qu'il faut multiplier par deux la survie des nymphes (c'est-à-dire passer d'une survie de 42 % à une survie de 83 %, *i.e.* multiplier par deux les densités vectorielles) pour obtenir une telle augmentation du *R0*. Ce résultat montre l'importance de la longévité des moustiques dans le cycle de transmission de la dengue. Le *R0* est également relativement sensible au nombre de personnes piquées par un moustique au cours d'un repas sanguin (*k*), au nombre d'habitants par maison, ainsi qu'à la durée de la virémie (*Tv*).

5.2.1.e. Discussion et perspectives

Hypothèses du modèle

Dans ce modèle, nous avons fait plusieurs hypothèses :

- A chaque pas de temps, la population d'hôtes est considérée comme entièrement susceptible, indépendamment du développement effectif d'une épidémie ;
- Le taux de mortalité des nymphes est égal à 0 ;
- Le taux de mortalité des adultes est supposé constant, indépendant de l'âge des femelles et indépendant de la température pour des valeurs de température comprises entre 2 °C et 38 °C.

Valeurs seuils d'indices entomologiques

Ce modèle montre que les valeurs seuils d'indices entomologiques à atteindre pour empêcher le développement d'une épidémie sont extrêmement bas (IPA < 1 nymphe ou larve de stade 4 par maison pour une température de 28 °C). Ce résultat pourrait expliquer le développement d'épidémies en Nouvelle-Calédonie malgré les efforts importants de lutte anti-vectorielle réalisés, et la difficulté à maintenir des densités vectorielles en dessous d'un seuil critique de manière opérationnelle (Ooi et al., 2006). Cependant ce seuil est difficilement utilisable de manière opérationnelle comme valeur seuil cible à atteindre tant qu'un modèle reproduisant de manière fiable la variabilité inter-annuelle du risque d'émergence épidémique observé en Nouvelle-Calédonie n'aura pas été construit.

Paramètres du cycle de vie des moustiques influençant la transmission

L'analyse de sensibilité du modèle pourrait être utile aux chercheurs d'autres domaines, car elle permet d'identifier les paramètres à estimer en priorité lors des études entomologiques expérimentales. Les paramètres à identifier en priorité sont :

- les taux de mortalité des moustiques adultes (cf (Brady et al., 2013) pour des exemples d'études), et l'influence des variables climatiques ;
- le nombre d'hôtes piqués par repas sanguin (cf (Scott et al., 1993, 2000) pour un exemple)
- la durée de la virémie.

Variabilité inter-annuelle du risque d'émergence épidémique

Alors qu'un lien statistique a été trouvé entre température et variabilité inter-annuelle du risque épidémique (Descloux et al., 2012), le modèle mécanistique n'arrive pas reproduire la variabilité inter-annuelle du risque. Les raisons qui pourraient expliquer cette contradiction dans les résultats incluent :

- Un processus du cycle de vie du moustique non pris en compte (prédation des larves et des adultes par exemple) ;
- Un processus lié à l'hôte non pris en compte (comme l'immunité de la population due aux épidémies précédentes) ;
- Un processus lié au virus non pris en compte (comme le nombre de virus introduits par an sur le territoire) ;
- Un processus lié à l'environnement non pris en compte (comme les précipitations, nécessaires au remplissage des gîtes larvaires et à l'éclosion des œufs, ou encore la température maximale, ou des variables démographiques) ;
- Des équations liant les variables d'entrée, en particulier la température, au risque épidémique annuel ne représentant pas la réalité (notamment, dans les données il est observé un effet seuil de la température et de l'humidité tandis que les équations utilisées n'introduisent aucun effet seuil) ;
- Une mauvaise adéquation entre les valeurs des paramètres du modèle tirés de la littérature et les valeurs réelles en Nouvelle-Calédonie ;
- Non prise en compte de la persistance de conditions climatiques favorables (dans ce modèle, le *R0* est calculé à chaque pas de temps à partir d'une seule valeur de température moyenne mensuelle).

Nous avons vu, grâce aux analyses EOF, que l'immunité de population pouvait potentiellement jouer un rôle important dans les niveaux d'incidence observés à Nouméa. La prise en compte de cette variable ne peut se faire par un modèle de capacité vectorielle uniquement, et nécessiterait la construction d'un

modèle SIR complet, intégrant les différents sérotypes. Ce type de modèle a notamment permis de reproduire la variabilité inter-annuelle de la dynamique de la dengue dans les régions hyper-endémiques d'Asie (Adams et al., 2006; Cummings et al., 2005; Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006).

D'après la DASS-NC, qui effectue des enquêtes épidémiologiques auprès des cas index chaque année, le nombre de cas introduits est relativement constant, compris entre 10 et 30 cas cliniques par an. Il est peu probable que le nombre de virus introduits par an puisse alors expliquer la variabilité inter-annuelle du risque d'émergence.

En ce qui concerne les précipitations, elles pourraient être prises en compte par des équations permettant de simuler le remplissage des gîtes larvaires. Ces équations ont déjà été développées dans le cadre de la construction d'un modèle de dynamique de population de moustique, le modèle CIMSIM (Focks et al., 1993a).

L'introduction d'effets seuils dans les équations peut se faire de deux manières : soit au travers des équations liant la température à la durée des cycles, soit au travers des équations liant la température et l'humidité relative à la mortalité des larves et des moustiques adultes. Les équations liant la température à la durée des cycles sont des équations de cinétique enzymatique. Or les enzymes ont un fonctionnement perturbé pour des températures autour de 39 °C/40 °C (*i.e.* la température de la fièvre). Il n'y a donc pas de raison d'introduire d'effet seuil pour des températures à 32 °C. En revanche, nous avons supposé la mortalité des adultes constante tout au long de la période d'étude (les températures étant comprises entre 2 °C et 38 °C). Or des études ont montré récemment que pour des températures moyennes supérieures à 18 °C, de larges variations diurnes de température impactent négativement la transmission de la dengue par leur influence négative sur la mortalité (Lambrechts et al., 2011), ainsi que sur le temps de développement des larves (Carrington et al., 2013). De même une étude récente de modélisation des données issues de 400 études de l'influence de la température sur la mortalité des *Aedes aegypti* adultes a montré un lien complexe et non linéaire entre température et mortalité (voir Figure 5.13 (Brady et al., 2013)), pour lequel les seuils de température au-delà et en-deça desquels la mortalité augmente fortement sont beaucoup plus étroits que ceux utilisés. L'utilisation des équations développées par *Brady et al.* afin de prendre en compte la complexité de ce lien permettrait certainement d'introduire dans le modèles des effets seuils plus réalistes.

Etant donnée l'analyse de sensibilité du modèle, l'imprécision sur les paramètres *r* (pourcentage de moustiques femelles), *Sn* (pourcentage de survie des nymphes), *NH* (nombre total d'habitants sur Nouméa, *pa* et *pb* (probabilités de transmission virale lors d'un repas sanguin entre un hôte infecté et un moustique sain ou l'inverse, respectivement) n'a qu'un influence très modérée sur les résultats du modèle. En revanche il conviendrait de mieux estimer les valeurs réelles du taux de mortalité des femelles adultes (cf précédemment), de la durée de la virémie ainsi que du nombre de personnes piquées par vecteur par repas sanguin.

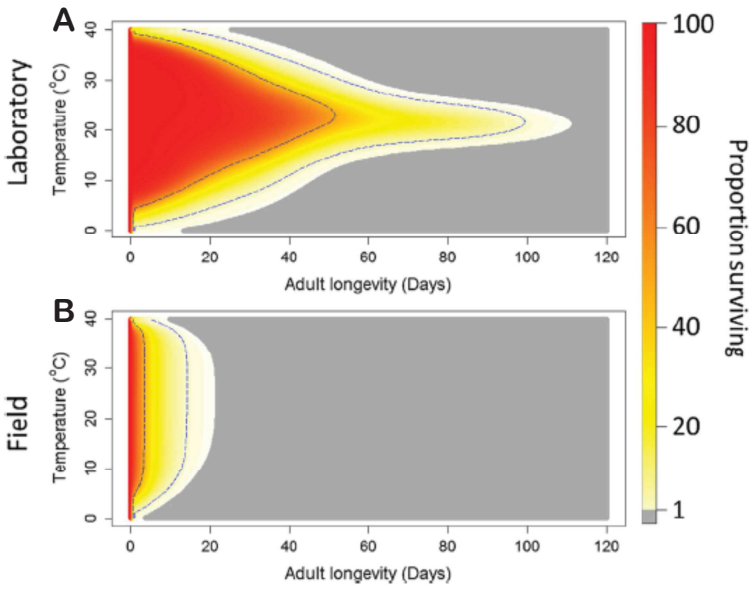


Figure 5.13 : Lien entre température, mortalité et âge des moustiques Aedes aegypti adultes. *La figure est reproduite à partir de Brady et al. (Brady et al., 2013). Les panneaux montrent les résultats des équations de modélisation des données issues d'études de laboratoire (**A**) ou de terrain (**B**). Ces équations lient le taux de mortalité des adultes (1 – proportions d'adultes survivant), représentée par une couleur, faible mortalité en rouge, forte mortalité en jaune, à l'âge des moustiques (abscisse) et à la température (ordonnée). En laboratoire, des moustiques âgés de 1 mois ont des taux de mortalité qui augmentent drastiquement pour des températures inférieures à environ 10 °C ou supérieures à 30 °C.*

Dans ce modèle, nous avons utilisé les températures moyennes mensuelles, comprises entre 24 °C et 28 °C toute l'année. Or la variable au-delà duquel le risque d'émergence augmente fortement en Nouvelle-Calédonie identifiée par les modèles statistiques concerne la persistance de températures maximales au-delà d'un seuil de 32 °C sur plusieurs jours. Nous avons donc construit également des simulations quotidiennes du *R0* utilisant les températures maximales observées à Nouméa, puis effectué des moyennes mensuelles ou annuelles des valeurs du *R0* quotidiennes simulées, ces valeurs mensuelles ou annuelles devant être plus importantes lors des mois ou années où le nombre de jours où la température maximale dépasse 32 °C est élevé. Les simulations ainsi obtenues continuaient de prédire un risque d'émergence pour l'ensemble des années sans exception, et les valeurs les plus faibles des *R0* simulés ne correspondaient pas aux années non-épidémiques. Donc même si les équations étaient modifiées pour introduire un effet seuil sur les températures, il n'est pas sûr que le modèle arrive à reproduire la variabilité inter-annuelle du risque d'émergence épidémique. Ceci renforce l'idée qu'une variable importante du cycle de transmission manque dans les équations et guide les perspectives futures vers la construction d'un modèle SIR prenant en compte l'immunité de population.

5.3. Analyse spatiale

5.3.1. Objectifs et études réalisées

Nous avons vu que les schémas spatiaux de la dengue à l'échelle de Nouméa, résolution « quartiers », sont relativement stables dans le temps (voir analyse spatio-temporelle). L'objectif des études présentées ici est l'identification des facteurs spatialement associés à des taux d'incidence élevés.

Pour répondre à cet objectif, nous avons mené, dans le cadre d'un projet financé par le Ministère des Affaires étrangères (projet MOM) réalisé en collaboration avec des partenaires de l'île de la Réunion, deux types d'études :

- Une enquête de terrain ;
- Une étude de modélisation similaire à celle effectuée à l'échelle des communes (voir article 3).

L'enquête de terrain a été réalisée par Lucie Laplaige, stagiaire de Master 2 Géographie au sein de l'unité mixte de recherche 228 ESPACE-DEV. J'ai participé à l'encadrement de ce stage, fournissant un appui méthodologique et technique dans le montage du protocole d'enquête et dans l'analyse des données, mais le remplissage des questionnaires, la construction de la base de données, les analyses elles-mêmes et les tableaux de résultats ont été réalisés par Mlle Laplaige. Les résultats sont présentés ici, quoique de manière succincte, car ils apportent un éclairage intéressant sur l'étude de modélisation.

L'étude de modélisation a été réalisée en collaboration avec une stagiaire bénéficiant d'une bourse « Leonardo » : Alizé Mercier, ainsi qu'avec l'aide de Marc Despinoy, expert en télédétection à l'UMR ESPACE-DEV. Mon rôle a été d'encadrer Mlle Mercier, de coordonner les réunions pour le tri des données et la construction des indicateurs, de construire la base de données épidémiologique en géo-référençant les cas et de construire les modèles statistiques. Le travail de construction de la base de données des indicateurs (variables explicatives) définis au cours des réunions, ainsi que les analyses exploratoires (ACP) ont été effectuées par Mlle Mercier, sous mon encadrement.

5.3.2. Enquête spatiale des facteurs de risque socio-économiques de la dengue à Nouméa

5.3.2.a. Matériel et méthode

Protocole d'enquête

Nous avons utilisé une méthode d'enquête de type « cas-témoin » dont le but est de valider l'hypothèse selon laquelle l'exposition d'un individu à un ou plusieurs facteurs augmente ou diminue le risque d'attraper la maladie. Nous avons cherché à collecter des données concernant deux types de facteurs d'exposition : des facteurs identiques aux variables explicatives utilisées dans l'étude de modélisation, afin de valider sur le terrain les résultats de la modélisation, et des facteurs accessibles uniquement par

des enquêtes de terrain, pour lesquels il n'existe pas de données analysables par modélisation (présence de gîtes dans les jardins, mesures de protection personnelles utilisées…). Certains facteurs concernaient les individus, d'autres l'habitat et l'environnement du moustique. L'annexe 8 présente le questionnaire utilisé, avec la liste des facteurs d'exposition analysés. Nous avons défini six zones géographiques sur Nouméa, trois zones « cas », et trois zones « témoin ». Les zones témoins ont été sélectionnées de manière à être adjacentes aux zones « cas » afin de minimiser l'effet des facteurs de confusion (*i.e.* facteurs pouvant potentiellement influencer le risque, mais non pris en compte dans l'étude et qui pourraient différer entre les zones « cas » et les zones « témoin », biaisant ainsi les analyses). Les zones « cas » ainsi que les zones « témoin » ont été sélectionnées sur la base d'analyses de cluster des données d'incidence par îlot en 2009. L'individu statistique est le chef de foyer (un questionnaire par maison visitée). Les enquêtes épidémiologiques de type cas-témoin se font en général sur la base du statut de cas ou de témoin de l'individu statistique. Ici, ne connaissant pas le statut réel du chef de foyer vis-à-vis de la dengue en 2009 (cas ou témoin), chaque personne interrogée a été désignée comme « cas » ou « témoin » sur la base de l'appartenance de son lieu d'habitation aux zones géographiques définies précédemment. Dans chaque zone géographique, les maisons ont été sélectionnées par un échantillonnage aléatoire simple. Un objectif de 400 questionnaires (67 questionnaires par zone) a été établi sur la base d'un calcul de puissance. Au total, 383 questionnaires exploitables ont pu être analysés (197 dans les zones « cas », et 186 dans les zones « témoin »).

Analyses statistiques

Le rôle des facteurs d'exposition binaire (l'individu est exposé au risque ou non) en tant que facteurs de risque de la dengue a été évalué au moyen d'odd-ratios, et les intervalles de confiance des odd-ratios calculés au moyen de régressions logistiques. Un odd-ratio significativement inférieur à 1 (borne supérieure de l'intervalle de confiance à 95 % inférieur à 1) pour un facteur d'exposition donné indique que le facteur en question est un facteur de protection : les personnes exposées à ce facteur ont significativement moins de chances d'attraper la dengue que les personnes non exposées. Inversement, un odd-ratio significativement supérieur à 1 (borne inférieure de l'intervalle de confiance à 95 % supérieur à 1) indique que le facteur est un facteur de risque : les personnes exposées à ce facteur ont significativement plus de chances d'attraper la dengue. Un odd-ratio non significativement différent de 1 indique que le facteur n'a pas d'influence sur le risque d'attraper la dengue.

5.3.2.b. Résultats

La Table 5.4 présente les résultats de l'analyse des facteurs d'exposition binaires. Les facteurs de risque prépondérants sont l'absence de l'entretien des jardins, ainsi qu'une densité locale d'habitation élevée. Ces résultats seront discutés en même temps que les résultats de l'étude de modélisation.

Table 5.4 : Résultats de l'enquête spatiale des facteurs de risque sociologiques de la dengue à l'échelle de Nouméa. Les facteurs sont triés par odd-ratio croissant (des facteurs protecteurs aux facteurs de risque). Les facteurs protecteurs sont surlignés en vert, les facteurs de risque en rouge.

Facteur	Odd-ratio	IC-*	IC+*
avec bateau	0.42	0.18	0.98
efficacité anti-moustique	0.45	0.24	0.84
habitat avec plusieurs étages	0.49	0.21	1.16
> 30 pots	0.52	0.14	1.87
rue large	0.53	0.3	0.92
habitat sur pilotis	0.56	0.28	1.13
habitat sans étage	0.59	0.39	0.89
jardin aéré	0.63	0.41	0.97
jardin entretenu	0.65	0.43	0.98
pas de gouttière	0.66	0.37	1.19
< 10 pots	0.67	0.36	1.26
climatisation	0.67	0.44	1.02
sans objets	0.71	0.46	1.09
avec des oiseaux/volailles	0.78	0.29	2.12
Piqure lieu de travail	0.87	0.49	1.55
Piqure matin	0.89	0.51	1.55
haies denses	0.92	0.59	1.43
habitat collectif	0.93	0.58	1.5
avec des chiens/chats	0.94	0.63	1.41
arbres fruitiers	0.97	0.65	1.46
piqure soir	0.98	0.65	1.47
haies aérée	0.99	0.66	1.48
sans animaux	1.03	0.69	1.55
piqure midi	1.05	0.23	4.76
gouttière en mauvais état	1.1	0.62	1.96
grands arbres	1.12	0.74	1.7
pas de haies	1.13	0.7	1.84
piqure lieu de sociab	1.13	0.7	1.82
moustiquaire	1.17	0.64	2.13
piqure intérieur	1.22	0.81	1.84
serre/ombrière	1.24	0.56	2.72
habitat indep	1.27	0.82	1.97
piqure nuit	1.27	0.83	1.95
gouttière en bon état	1.31	0.83	2.06
avec encombrants	1.34	0.86	2.08
avec objets divers	1.35	0.9	2.03
info qd épidémie	1.43	0.84	2.43
avec pneus	1.5	0.79	2.86
jardin peu/pas entretenu	1.52	1	2.32
se protège + qd informé	1.57	1.04	2.37
jardin dense	1.66	1.07	2.57
piqure tte journée	1.77	0.99	3.15
10-30 pots	1.99	0.95	4.15
habitat avec 1 étage	1.99	1.32	3
habitat mitoyen	2.28	1.18	4.05

* Bornes inférieures (IC-) et supérieurs (IC+) de l'intervalle de confiance à 95 % des odd-ratios

5.3.3. Modélisation non linéaire de l'intensité de la circulation virale

Pour l'étude de modélisation, nous avons commencé par réaliser une étude des facteurs pouvant expliquer l'hétérogénéité spatiale de la dengue à la résolution des quartiers de Nouméa. Nous avons obtenu un signal, relativement faible, entre les variables socio-économiques et la distribution spatiale des taux d'incidence. Dans le même temps, l'exploration des données géographiques de la base de données topographique a révélé une grande hétérogénéité spatiale intra-quartiers, dans les surfaces bâties, la taille des jardins ou encore la densité d'adresses autour d'une maison. La faiblesse du signal obtenu à la résolution des quartiers ainsi que l'hétérogénéité intra-quartier des variables explicatives nous ont conduits à augmenter la résolution de l'analyse pour essayer d'obtenir une information plus claire. L'unité spatiale de l'étude présentée ici est donc l'îlot, unité utilisée pour le recensement de population et comprenant une trentaine de maisons en moyenne. Cette étude a nécessité le géo-référencement précis des cas de dengue à partir de l'adresse des patients (voir 2.2.3.a, page 88). N'ayant pu géo-référencer l'ensemble des cas, l'étude est limitée à l'année 2009 et à la commune de Nouméa (excluant le Mont-Dore et Dumbéa). Les objectifs de cette étude sont : i) identifier les zones les plus à risque dans Nouméa et ii) identifier les variables responsables d'une augmentation du risque de transmission.

Une particularité de la ville de Nouméa par rapport aux villes de métropole est la présence de « squats », groupements d'habitations sauvages, le plus souvent constituées de tôle ondulée, dans laquelle habitent des personnes d'origine kanak venues de tribus de brousse pour chercher du travail, mais n'ayant pas les moyens financiers de payer un loyer, ou préférant vivre de façon communautaire. Le plus souvent, les squats ne sont pas reliés au réseau d'eau courante ni à l'électricité, ni à un système municipal d'évacuation des eaux usées. Les gens restent dehors la plupart de la journée. Ils sont donc plus exposés aux piqûres des moustiques *Aedes aegypti* à activité diurne que les personnes travaillant dans les bureaux climatisés par exemple, ce qui devrait entraîner un risque accru de contracter la dengue au cours d'une épidémie.

5.3.3.a. Matériel et méthodes

Données épidémiologiques

Sur les 2 323 cas de dengue déclarés entre le 1er septembre 2008 et le 31 août 2009 sur la commune de Nouméa, 1 762 cas (75,9 %) ont pu être géo-référencés précisément. Les cas non géo-référencés comprennent 98 cas (4,2 % du total des cas) pour lesquels il n'existe aucune information spatiale, 9,7 % de cas pour lesquels un nom de rue a été identifié, mais auquel le numéro dans la rue est manquant, 3,9 % de cas dont l'adresse est enregistrée sous la forme d'un lotissement, 1,5 % de cas dont l'adresse

est enregistrée sous forme d'un bâtiment, et 4,9 % de cas survenus sur des patients habitant dans des squats.

L'absence de géo-référencement des cas survenus dans les squats ainsi que dans les trois seuls bâtiments de Nouméa comprenant un grand nombre d'appartements (les tours de Magenta, la cité Pierre Lenquête, et la cité Saint Quentin) entraîne des taux d'incidence artificiellement bas dans les îlots correspondants, et donc un biais dans l'analyse. Ces cas représentent au maximum 6,4 % du total des cas déclarés à Nouméa en 2009. Le reste des cas non géo-référencés peuvent être considérés comme répartis aléatoirement dans l'espace et donc sans incidence sur la structure spatiale étudiée.

Nous avons ensuite agrégé les cas par îlot et construit des taux d'incidence annuels par îlot. Le nombre médian d'habitants par îlot est de 88. Comme les taux d'incidences sont calculés en divisant le nombre de cas au sein d'un îlot par le nombre d'habitants, l'apparition d'un cas dans un îlot de 100 habitants entraîne une augmentation du taux d'incidence de 1 %. A cette résolution spatiale, nous avons donc calculé les taux d'incidence en nombre de cas pour cent habitants par an. Le taux d'incidence annuel correspondant à l'estimation du pourcentage de personnes touchées par la dengue au cours de l'année, la précision de l'estimation de ce pourcentage diminue lorsque le nombre d'habitants dans un îlot est faible. Nous avons décidé de n'étudier que les îlots dans lesquels il y a plus de dix habitants, le nombre seuil de 10 habitants/îlot ayant été choisi après inspection de l'histogramme des taux d'incidence par îlot. En effet, certains îlots présentaient des valeurs extrêmement élevées de taux d'incidence, uniquement parce qu'un seul cas était apparu dans l'îlot au sein d'une population réduite.

Variables explicatives

Les variables explicatives ont été construites à partir des données du recensement de 2009 (Institut de la statistique et des études économiques (ISEE), 2009), des données topographiques de la DITTT, et des données de télédétection. Elles sont de plusieurs types :

- Des variables socio-économiques décrivant le niveau socio-économiques ou l'origine culturelle des individus (lieu de naissance, catégorie socio-professionnelle…) ;
- Des variables décrivant les logements et permettant de caractériser l'habitat (maison vs appartement, présence d'air conditionné, toilettes à l'intérieur ou à l'extérieur…);
- Des variables reflétant la densité de population (densité de population et densité d'adresse, pourcentage de surface bâtie…) ;
- Des variables reflétant l'environnement du moustique (indices de télédétection décrivant le couvert végétal, indice de morcellement de l'espace, distance au point d'eau…).

Elles ont été calculées à la résolution de l'îlot. L'ensemble des variables est listé dans la Table 5.5.

Pré-sélection des variables explicatives

Afin de réduire le nombre de variables à étudier, et de limiter la colinéarité entre variables explicatives, nous avons pré-sélectionné certaines variables en utilisant des ACP pour identifier les groupes de variables corrélées entre elles. Nous avons réalisé une ACP par type de variable (socio-économique, bâti, densité de population et environnement). Dans chacun des groupes de variables corrélées entre elles, nous en avons sélectionné une seule comme représentative de toutes les autres. La variable sélectionnée est celle présentant le plus grand coefficient de corrélation (en valeur absolue) avec les taux d'incidence de la dengue par îlot. Nous n'avons pas sélectionné les variables non significativement corrélées aux taux d'incidence de la dengue. L'ACP des variables socio-économiques n'a pas permis d'identifier de réels groupes de variables. Nous avons donc sélectionné les variables socio-économiques de manière itérative, en ne sélectionnant que les variables significativement corrélées à la dengue, puis en éliminant les variables corrélées à d'autres variables avec un coefficient de corrélation de plus 0,5 en valeur absolue.

Au total, nous avons sélectionné neuf variables explicatives pour la suite de la modélisation :

- Pourcentage de maisons ;
- Pourcentage de logements reliés à l'électricité ;
- Pourcentage de personnes habitant le même logement qu'en 2004 ;
- Pourcentage de personnes de catégorie socio-professionnelles « artisan » ;
- Pourcentage de personnes nées en Nouvelle-Calédonie ;
- Pourcentage de personnes n'ayant pas d'activité rémunérée ;
- Pourcentage de personnes travaillant dans le 1er secteur ;
- Pourcentage de personnes travaillant dans le deuxième secteur ;
- Densité d'adresses.

Modélisation multi-variée du risque de transmission

Afin de prendre en compte l'existence de liens possiblement non linéaires entre la variable réponse et les variables explicatives, nous avons décidé d'utiliser les modèles SVM (Cortes and Vapnik, 1995; Vapnik, 1998) pour quantifier le lien entre le risque de transmission et les variables explicatives. Afin d'identifier les variables les plus associées au risque de transmission de la dengue, nous avons construit plusieurs modèles utilisant toutes les combinaisons possibles de une, deux ou trois variables explicatives. Pour chacun des modèles créés, nous avons optimisé les paramètres du modèle en testant une série de valeurs possibles, et en sélectionnant les paramètres du modèle ayant une erreur minimale. L'erreur de chacun des modèles est estimée par le calcul de la RMSE (racine carrée de l'erreur quadratique moyenne), en utilisant une méthode de validation croisée découpant le jeu de données en 10 sous-échantillons utilisés à tour de rôle comme jeu de validation du modèle (« 10-fold cross validation ») (Kohavi, 1995). La

comparaison des RMSE de l'ensemble des modèles construits et de leur écart-type permet ensuite d'identifier les variables les plus associées au risque de transmission.

Etudes d'auto-corrélation

Nous avons étudié l'auto-corrélation de la variable réponse (taux d'incidence de la dengue par îlot) en comparant le semi-variogramme de la variable étudiée avec 200 semi-variogrammes obtenus par permutation aléatoire des coordonnées géographiques des centroïdes des îlots (Bivand et al., 2013).

Ayant décelé une auto-corrélation significative des taux d'incidence, nous avons cherché à comprendre si cette auto-corrélation était due au processus infectieux lui-même (diffusion locales des cas) ou s'il elle était due au forçage extérieur d'une variable elle-même auto-corrélée. Nous avons donc également quantifié l'auto-corrélation des résidus de chacun des modèles bivariés créés.

5.3.3.b. Résultats

Distribution spatiale moyenne des taux d'incidence

La Figure 5.14 montre les taux d'incidence par îlot à Nouméa. Le risque de transmission de la dengue est spatialement hétérogène à Nouméa. Les parties Nord et Ouest de Nouméa ont été touchées de manière plus sévère lors de l'épidémie de 2009. Cette carte fait également apparaître des zones très localisées de transmission plus intense dans des quartiers où la circulation virale est en moyenne relativement basse, comme le Nord de Faubourg-Blanchot, la baie de l'Orphelinat ou Magenta. Les zones de transmission les plus intenses en 2009 étaient les zones habitées de la zone industrielle de Ducos, une partie de Logicoop ainsi que la plupart des îlots de Rivière Salée. La moyenne des taux d'incidence annuels sur l'ensemble des îlots est de 2,0 cas pour 100 habitants/an, et l'écart-type de 2,5 cas pour 100 habitants/an. Nous pouvons voir sur cette carte l'hétérogénéité spatiale existant au sein d'un même quartier (voir la zone industrielle de Ducos où seuls les îlots proches de Logicoop sont touchés, ou le quartier P.K6, où seuls les îlots proches de Rivière Salée et de Normandie sont touchés). L'étude des taux d'incidence agrégés par quartiers peut donc masquer certains foyers de transmission très localisés.

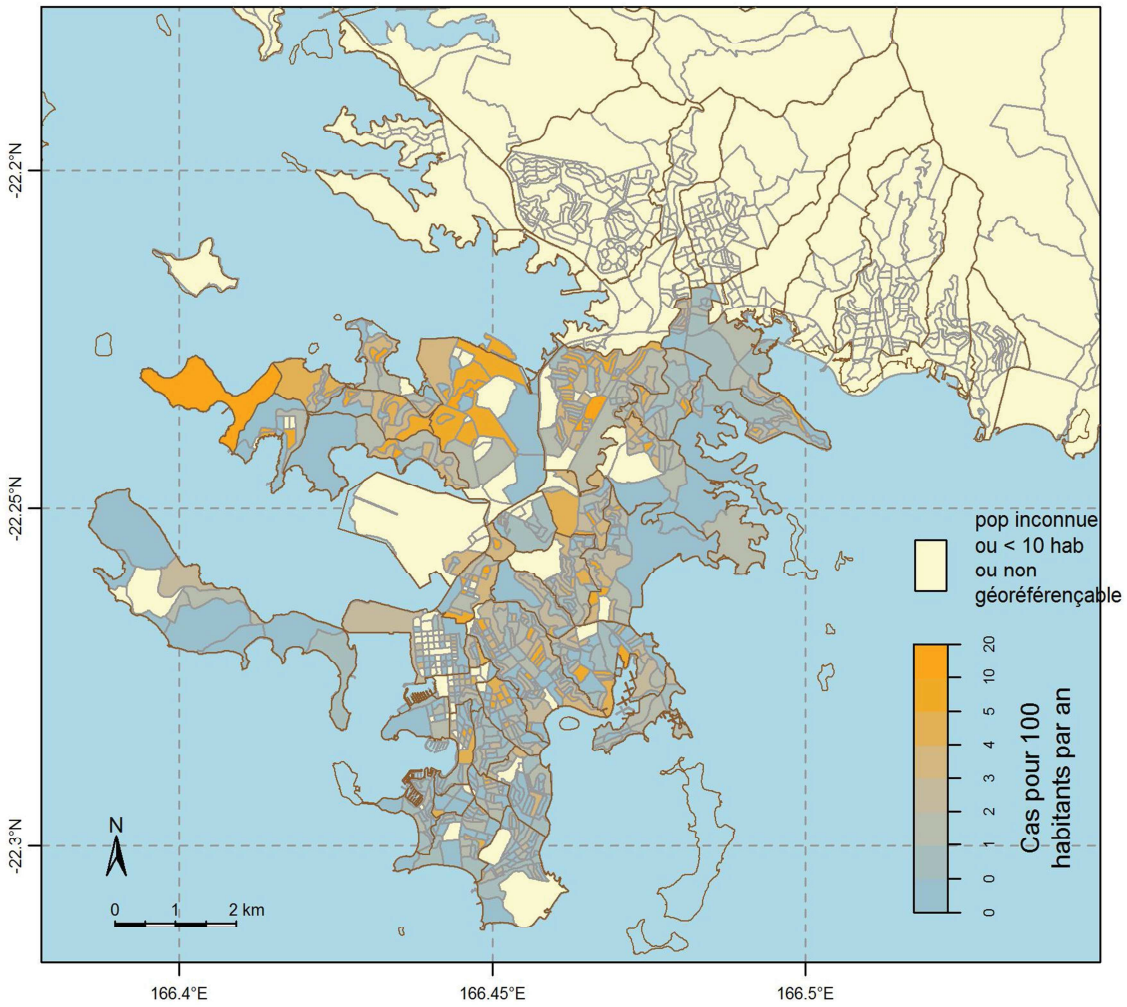


Figure 5.14 : Taux d'incidence annuels de la dengue en 2009 à Nouméa par îlot. *L'échelle de couleur des taux d'incidence est non linéaire. La carte avec les noms des différents quartiers peut être trouvée* ***Figure 2.7*** *page 96.*

Auto-corrélation de la variable réponse

La Figure 5.15 montre le semi-variogramme des taux d'incidence. Les taux d'incidence présentent une auto-corrélation significative jusqu'à près de trois kilomètres. Or dans la litérature, la taille des foyers infectieux de dengue à des échelles très locales est souvent comprise entre 200 et 800 m (Aldstadt et al., 2012; Honório et al., 2009; Liebman et al., 2012; Mammen Jr et al., 2008; Morrison et al., 1998; Rabaa et al., 2013; Salje et al., 2012; Tran et al., 2004; Vazquez-Prokopec et al., 2010; Yoon et al., 2012). La forte étendue de l'auto-corrélation des taux d'incidence de la dengue observés à Nouméa, ainsi que les valeurs du semi-variogramme empirique plus élevées que l'intervalle de confiance pour des distances élevées font penser que cette auto-corrélation provient d'une tendance spatiale de la distribution des cas de dengue (le terme de « tendance spatiale » décrit une variable dont la valeur augmente dans une direction spatiale donnée).

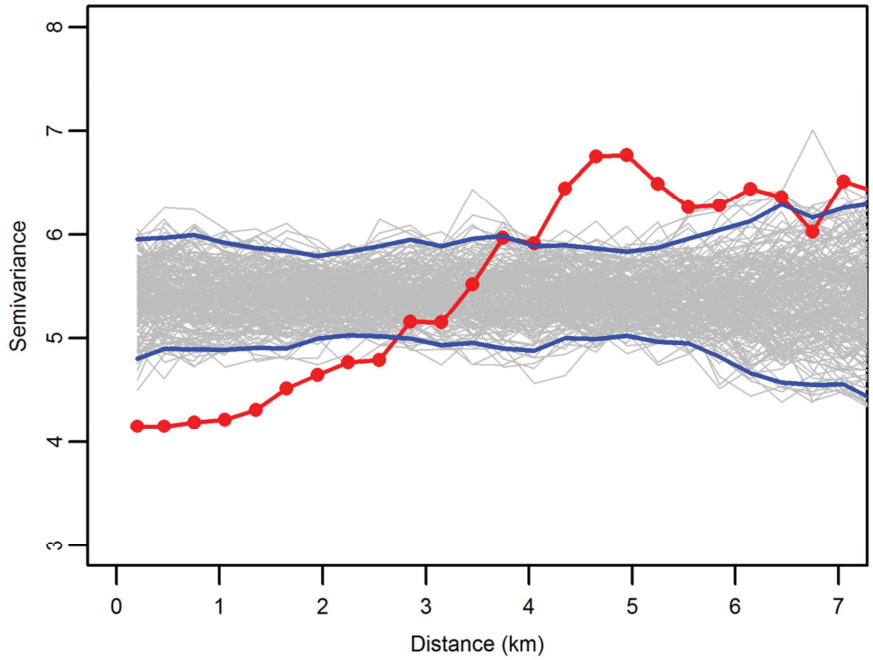


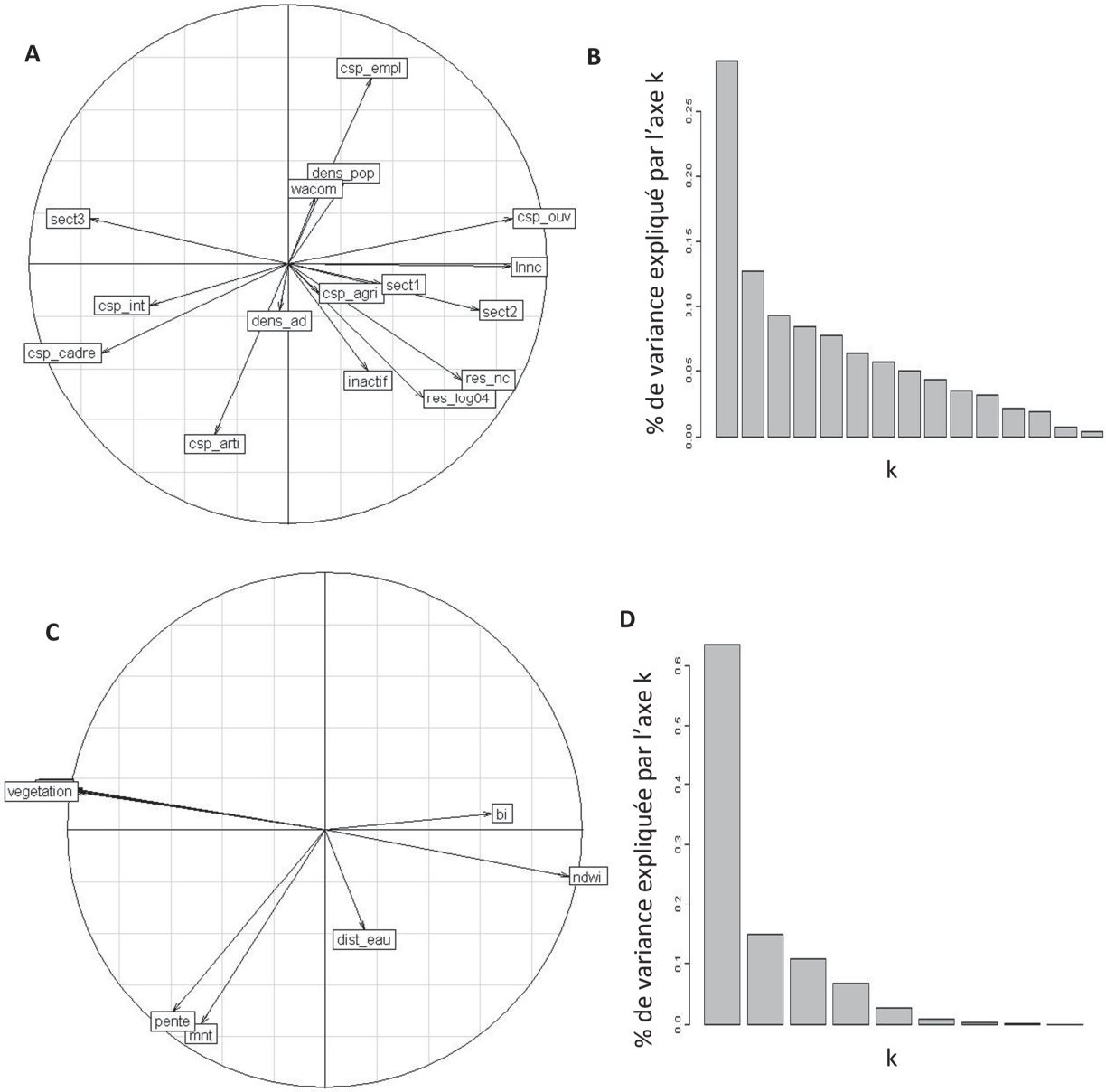
Figure 5.15 : Semi-variogramme des taux d'incidence annuels par îlot en 2009. *La courbe rouge est le semi-variogramme empirique des données d'incidence. Les courbes grises représentent deux cents semi-variogrammes obtenus par permutation aléatoire des coordonnées des centroïdes des îlots. Les courbes bleues représentent les 2,5 et 97,5 percentiles des semi-variogrammes aléatoires. Les données présentent une auto-corrélation spatiale significative lorsque les points proches dans l'espace présentent des valeurs proches de la variable étudiée (i.e semi-variance faible). La distance à partir de laquelle le semi-variogramme empirique n'est plus distinguable des semi-variogrammes théoriques construits sous l'hypothèse d'absence d'auto-corrélation donne l'étendue de l'auto-corrélation spatiale.*

Pré-sélection des variables explicatives

La Table 5.5 montre les coefficients de corrélation de Pearson entre chacune des variables explicatives et les taux d'incidence de la dengue. Globalement, les variables sont très peu corrélées aux taux d'incidence de la dengue (meilleur coefficient de corrélation de 0.248). Les variables socio-économiques ainsi que la variable « maison » semblent plus associées que les autres aux niveaux d'incidence observés. Parmi les variables socio-économiques, le lieu de naissance, qui est lié aux habitudes culturelles en Nouvelle-Calédonie, est la variable prépondérante associée à l'intensité de la transmission. Aucune des variables environnementales construites ne semble liée aux taux d'incidence de la dengue à Nouméa.

Table 5.5 : Coefficients de corrélation (Rho) entre les variables explicatives et les taux d'incidence de la dengue à l'échelle de Nouméa. Les variables peuvent être regroupées en 4 catégories : socio-économiques, bâti, densité de population et environnement. Les variables sont triées par ordre décroissant de la valeur absolue du coefficient de corrélation. Les variables sélectionnées pour la suite de la modélisation sont surlignées en gris. Les variables en italique sont celles non significativement corrélées à la dengue.

Variable	Category	Rho	p-value	Description
maison	Bâti	0.248	< 0.001	Percentage of housing that are houses, as opposed to appartments, sheds, boats or huts
Lnnc	Socio-économique	0.247	< 0.001	Percentage of people born in New-Caledonia
res_h04	Socio-économique	0.247	< 0.001	Percentage of people living in the same housing since before 2004
res_nc	Socio-économique	0.214	< 0.001	Percentage of people living in New-Caledonia since before 2004
Appart	Bâti	-0.208	< 0.001	Percentage of housing that are appartments as opposed to houses, boats, sheds or huts
sect2	Socio-économique	0.191	< 0.001	Percentage of people working in the secondary sector
sect3	Socio-économique	-0.187	< 0.001	Percentage of people working in the tertiary sector
csp_lab	Socio-économique	0.180	< 0.001	Percentage of people working as labor workers
csp_exec	Socio-économique	-0.167	< 0.001	Percentage of people working as executives
Inactif	Socio-économique	0.142	< 0.001	Percentage of people that are inactive (retired, working at home, unemployed and other inactive)
csp_arti	Socio-économique	-0.107	0.004	Percentage of people working as a artisans
dens_ad	Densité de population	0.100	0.006	Density of adresses in each block (number of adresses per m²)
elec1	Bâti	0.100	0.007	Percentage of housing connected to the general electricity network
sect1	Socio-économique	0.097	0.008	Percentage of people working in the primary sector
elec2	Bâti	-0.094	0.011	Percentage of housing with an electrical generator
Abri	Bâti	-0.090	0.015	Percetange of housing that are sheds as opposed to houses, appartments, boats or huts
eau1	Bâti	0.080	0.030	Percentage of housing with running water
elec3	*Bâti*	*-0.068*	*0.064*	*Percentage of housing with solar pannels or other*
eau2	*Bâti*	*-0.067*	*0.070*	*Percentage of housing with access to water through a public water hole*
csp_agri	*Socio-économique*	*0.062*	*0.091*	*Percentage of people working as farmers*
ndwi	*Environnemental*	*-0.056*	*0.125*	*Normalized Difference Water Index*
Bateau	*Bâti*	*-0.049*	*0.184*	*Percetange of housing that are boats as opposed to houses, appartments, sheds or huts*
eau3	*Bâti*	*-0.047*	*0.198*	*Percentage of housing that has access to water through a public water hole*
csp_empl	*Socio-économique*	*0.046*	*0.211*	*Percentage of people working as employers*
Mnt	*Environnemental*	*-0.043*	*0.247*	*Digital Elevation Model*
dens_pop	*Densité de population*	*-0.039*	*0.289*	*Population density (inhabitant/km²)*
tndvi	*Environnemental*	*0.038*	*0.301*	*Transformed Normalized Difference Vegetation Index*
dist_eau	*Environnemental*	*0.035*	*0.343*	*Distance to the nearest water source (lake, pond) (in meters)*
Cabane	*Bâti*	*-0.033*	*0.368*	*Percetange of housing that are huts as opposed to houses, appartments, sheds or boats*
surf_bat	*Densité de population*	*0.033*	*0.372*	*Percentage of built-area per cadastral parcel*
Ndvi	*Environnemental*	*0.027*	*0.458*	*Normalized Difference Vegetation Index*
csp_int	*Socio-économique*	*-0.023*	*0.535*	*Percentage of people working in an intermediate profession*
Pente	*Environnemental*	*-0.023*	*0.538*	*Slope (in degrees)*
Bi	*Environnemental*	*-0.017*	*0.638*	*Brightness Index*
Lwdm	*Socio-économique*	*-0.015*	*0.681*	*Percentage of people living and working in different municipalities*
ndvi_RE	*Environnemental*	*0.013*	*0.714*	*Normalized Difference Vegetation Index – Red Edge*
vegetation	*Environnemental*	*-0.008*	*0.830*	*Normalized Difference Vegetation Index in unbuilt-areas*


A
csp_empl
dens_pop
wacom
sect3
csp_ouv
lnnc
sect1
csp_agri
csp_int
sect2
dens_ad
csp_cadre
inactif
res_nc
res_log04
csp_arti
B
% de variance expliqué par l'axe k
0.00
0.05
0.10
0.15
0.20
0.25
k
C
vegetation
bi
ndwi
dist_eau
pente
mnt
D
% de variance expliquée par l'axe k
0.0
0.1
0.2
0.3
0.4
0.5
0.6
k

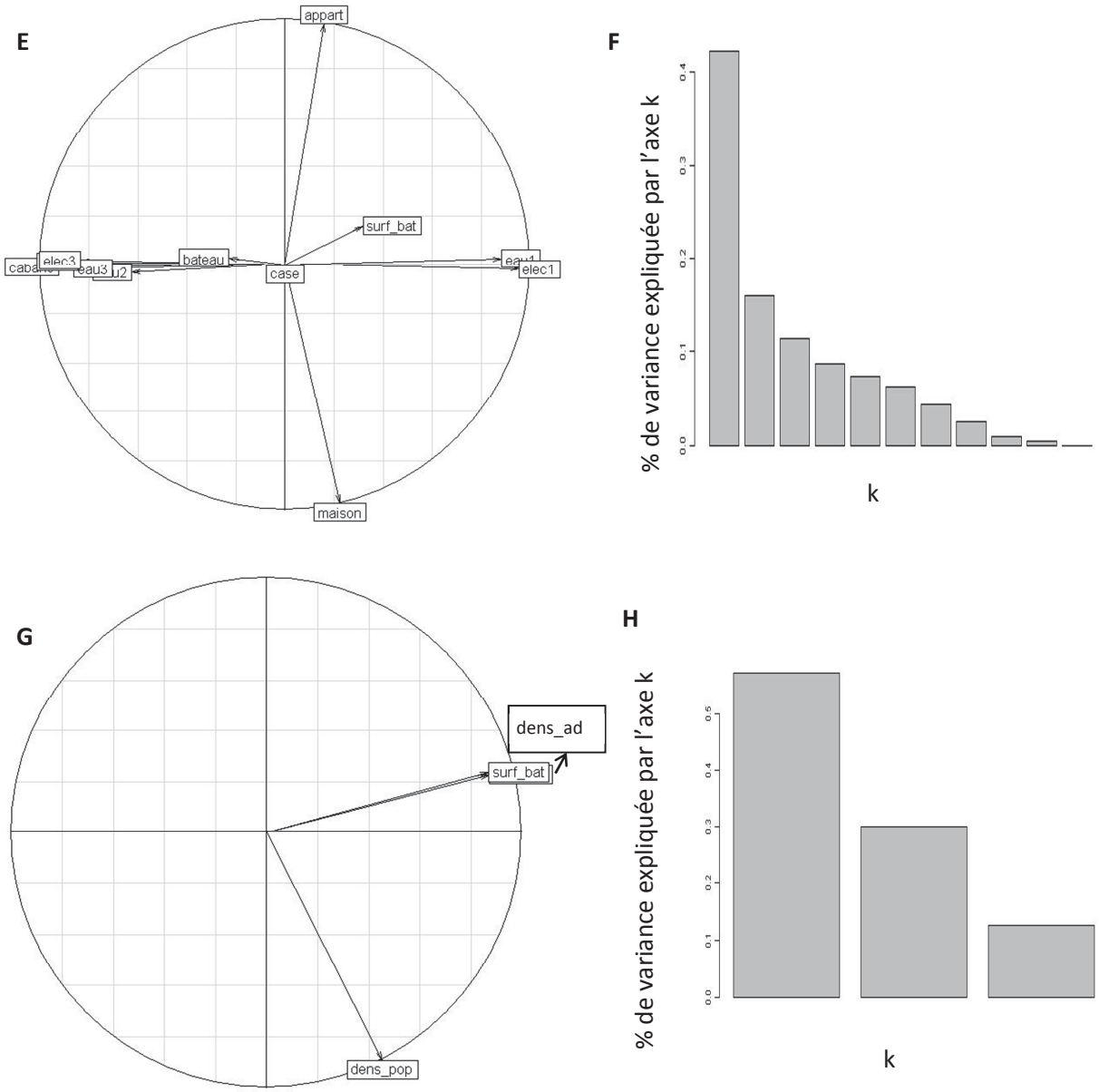


Figure 5.16 : Analyse en composantes principales des variables explicatives de la dengue à l'échelle de Nouméa. *Les figures montrent le cercle des corrélations (axe 1 en abscisse, axe 2 en ordonnée) et les valeurs propres (panneaux de droite) des ACP réalisées sur les variables socio-économiques (**A et B**), environnementales (**C et D**), type de bâti (**E et F**) et densité de population (**G et H**). Lorsque les variables sont proches du cercle (de rayon 1), le coefficient de corrélation de Pearson entre deux variables peut être estimé approximativement par l'angle formé entre les deux flèches : corrélation de 1 pour un angle de 0°, corrélation de 0 pour un angle de 90°, et corrélation de -1 pour un angle de 180°. Sur le panneau G, l'étiquette cachée correspond à la variable de densité d'adresse par îlot.*

La Figure 5.16 montre les résultats de l'ACP réalisée sur chacune des quatre catégories de variables explicatives. Globalement, il est difficile de différencier les îlots entre eux sur la base des variables socio-économiques (Figure 5.16.A et B., pourcentage de variance expliqué par les deux premiers axes d'à peine 42 %), ce qui montre une certaine mixité socio-économique au sein des îlots. Le premier axe oppose les îlots dans lesquels un grand pourcentage de personnes sont nées en Nouvelle-Calédonie et travaillent en tant qu'ouvriers ou dans le secteur secondaire, aux îlots dans lesquels la plupart des personnes sont cadres ou travaillent dans le commerce. Le deuxième axe oppose les îlots où il y a beaucoup d'employés à ceux où il y a beaucoup d'artisans.

Parmi les groupes de variables socio-économiques corrélées entre elles, nous avons sélectionné le pourcentage de personnes nées en Nouvelle-Calédonie (« lnnc »), le pourcentage de personnes travaillant comme artisans (« csp_arti»), le pourcentage de personnes habitant le même logement depuis 2004 (« res_h04 »), le pourcentage de personnes inactives (« inactif ») ainsi que le pourcentage de personnes travaillant dans le secteur primaire (« sect1 ») ou secondaire (« sect2 »).

Les Figure 5.16.C et D montrent que les variables de télédétection décrivant le couvert végétal ou l'humidité ont une distribution spatiale similaire (elles sont toutes corrélées entre elles, positivement ou négativement). La présence de végétation ou d'humidité n'est pas liée à l'altitude ou à la pente. Sur la base des coefficients de corrélation présentés en Table 5.5, nous aurions pu sélectionner la variable d'altitude (« mnt ») ainsi que l'indice télédétection « ndwi » pour la suite de la modélisation. Cependant, aucune de ces deux variables n'est significativement corrélée à la dengue. Nous ne les avons donc pas conservées pour la suite.

Les Figure 5.16.E et F montrent que le premier axe de variabilité entre îlots à Nouméa en termes de bâti est celui de la présence de squats. En effet, la présence de cabanes et d'habitation non reliées au réseau d'eau courante (eau2 et eau3) ou d'électricité (elec3) permet de distinguer les îlots d'autres îlots où un grand nombre d'habitations ont accès à l'eau (eau1) ou à l'électricité (elec1). Le deuxième axe de variabilité oppose les îlots où une forte proportion de personnes habite dans des maisons à ceux où les personnes habitent en appartement. Nous avons décidé de conserver le pourcentage de personnes habitant dans des maisons ainsi que le pourcentage de logements rattachés à l'électricité (elec1) pour la suite de la modélisation.

Les Figure 5.16.G et H montrent que le pourcentage de surface bâtie par îlot est très corrélé à la densité d'adresses (étiquette cachée sur la Figure 5.16.G), mais n'est pas du tout corrélée à la densité de population. Une hypothèse pour expliquer l'absence de corrélation entre densité d'adresse et densité de population est la présence de petits immeubles dans certains îlots, possédant un grand nombre d'appartements ayant tous la même adresse. Nous avons décidé de conserver la variable « densité d'adresses » pour la suite de la modélisation. Nous avons écarté la variable « densité de population » sur la base de sa faible corrélation avec les taux d'incidence de la dengue.

Modélisation multivariée

Table 5.6 : Modélisation bivariée et multi-variée des taux d'incidence de la dengue : sélection des variables par comparaison de l'erreur de prédiction des modèles SVM

Variable 1*	Variable 2*	Variable 3*	RMSE**
Lnnc	-	-	2,3
res_h04	-	-	2,3
Maison	-	-	2,3
sect1	-	-	2,4
sect2	-	-	2,4
dens_ad	-	-	2,4
Inactif	-	-	2,4
elec1	-	-	2,4
spc_arti	-	-	2,4
Maison	Lnnc	-	2,3
Maison	Inactif	-	2,3
Maison	Res_h04	-	2,3
Res_h04	Dens_ad	-	2,3
Maison	Sect2	-	2,3
Elec1	Sect2	-	2,3
Lnnc	Dens_ad	-	2,3
Res_h04	Lnnc	-	2,3
Elec1	Lnnc	-	2,3
Maison	Csp_arti	-	2,3
Lnnc	Inactif	-	2,3
Res_h04	Csp_arti	-	2,3
Elec1	Res_h04	-	2,3
Maison	Sect1	-	2,3
Csp_arti	Lnnc	-	2,3
Maison	Lnnc	Inactif	2,2
Maison	Lnnc	Res_h04	2,3
Maison	Res_h04	Csp_arti	2,3
Maison	Csp_arti	Inactif	2,3
Maison	Lnnc	Dens_ad	2,3
Maison	Csp_arti	Sect2	2,3
Maison	Csp_arti	Lnnc	2,3
Maison	Elec1	Lnnc	2,3
Maison	Csp_arti	Dens_ad	2,3
Maison	Res_h04	Dens_ad	2,3
Res_h04	Lnnc	Dens_ad	2,3
Res_h04	Csp_arti	Dens_ad	2,3
Maison	Res_h04	Sect2	2,3
Lnnc	Inactif	Dens_ad	2,3
Maison	Lnnc	Sect2	2,3

* Variables incluses comme variables explicatives pour la modélisation des taux d'incidence

** Racine carrée de l'erreur quadratique moyenne, en nombre de cas /100 habitants / an. Les modèles sont classés par nombre croissant de variables incluses, puis par ordre croissant d'erreur.

La Table 5.6 montre les résultats de la modélisation. Pour les modèles basés sur deux ou trois variables explicatives, seuls les quinze meilleurs modèles sont montrés. La RMSE peut servir à calculer l'intervalle de confiance de la prédiction des taux d'incidence. Or ces taux ne peuvent être calculés avec une précision supérieur à 1 cas/100 habitants /an. Nous pouvons donc considérer les RMSE comme toutes

équivalentes, aucun des modèles ne permettant d'améliorer la précision de la prédiction des taux d'incidence par îlot. Il apparaît donc difficile de classer les variables par ordre décroissant d'influence sur la transmission virale au sein d'un îlot. Nous pouvons cependant remarquer que les « meilleurs » modèles sont basés sur les variables les plus corrélées à la dengue, *i.e.* le pourcentage de personnes habitant en maison, ainsi que le pourcentage de personnes nées en Nouvelle-Calédonie. Malgré l'absence de colinéarité entre les variables explicatives (chacune des variables explicative a un coefficient de corrélation inférieur à 0,5 en valeur absolue avec l'ensemble des autres variables), l'association de plusieurs variables explicatives ne permet pas d'améliorer la performance globale de la prédiction des taux d'incidence.

Etudes d'auto-corrélation des résidus des modèles

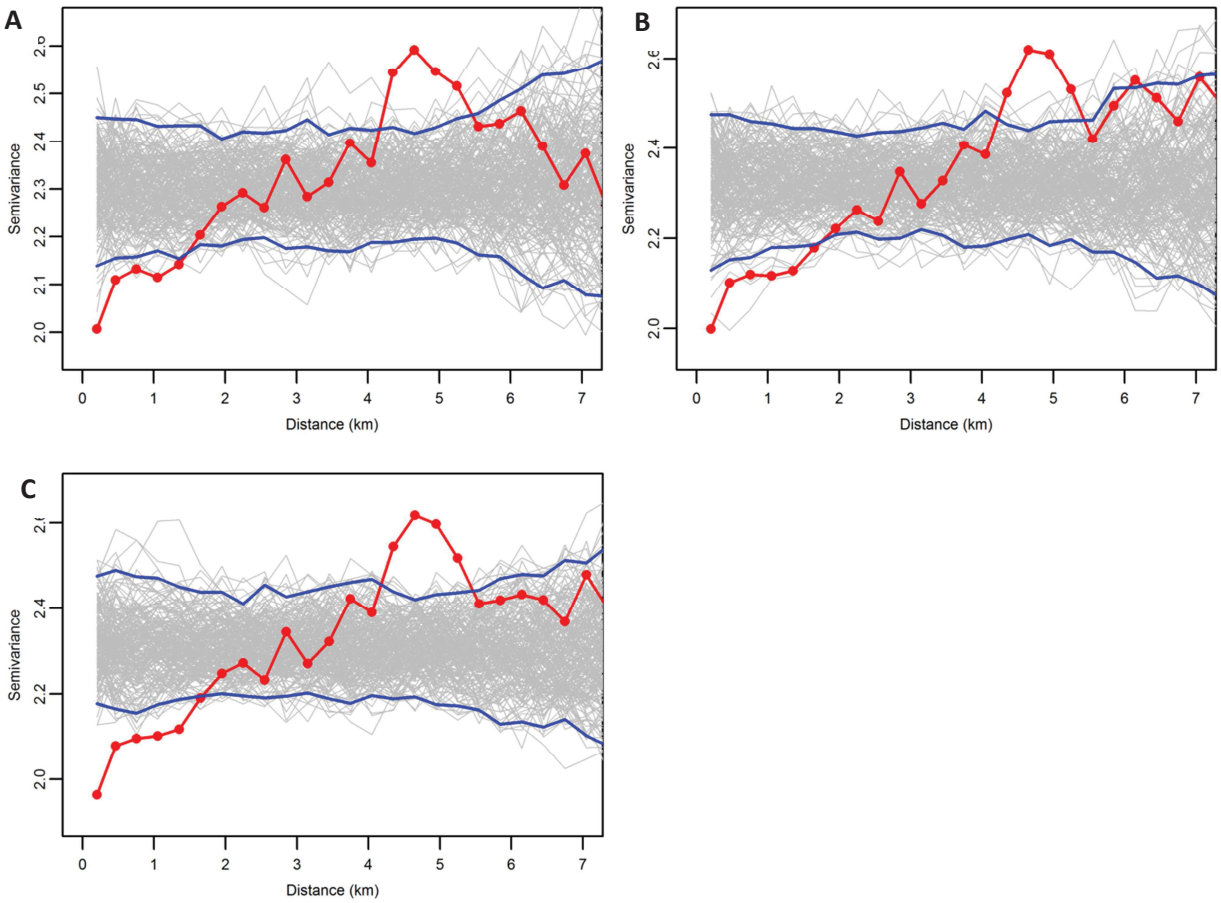


Figure 5.17 : Semi-variogrammes des résidus de trois modèles SVM bivariés prédisant les taux d'incidence par îlot à Nouméa en 2009 *à partir d'une des variables explicative sélectionnées. La variable explicative de chacun des modèles est le pourcentage de personnes nées en Nouvelle-Calédonie (A), le pourcentage de personnes habitant en maison (B) et le pourcentage de personnes n'ayant pas changé de logement depuis 2004 (C). Pour chacun des modèles, la courbe rouge est le semi-variogramme empirique des résidus du modèle. Les courbes grises représentent deux cents semi-variogrammes obtenus par permutation aléatoire des coordonnées des centroïdes des îlots. Les courbes bleues représentent les 2,5 et 97,5 percentiles des semi-variogrammes aléatoires.*

La Figure 5.17 montre les semi-variogrammes des résidus de trois modèles SVM bivariés prédisant les taux d'incidence à partir du pourcentage de personnes nées en Nouvelle-Calédonie, du pourcentage de personnes habitant en maison ou du pourcentage de personnes n'ayant pas changé de logement depuis 2004. Nous pouvons voir que les résidus présentent une auto-corrélation significative jusqu'à 1,5 km seulement (à comparer à la limite de 3 km de l'auto-corrélation des taux d'incidence). L'étude de l'auto-corrélation des résidus des autres modèles bivariés n'a pas mis en évidence de baisse significative de l'étendue de l'auto-corrélation. L'auto-corrélation jusqu'à 3 km observée sur les taux d'incidence initiaux semble donc en partie due au forçage extérieur de l'une des trois variables les plus corrélées aux taux d'incidence : ces variables présentent probablement une auto-corrélation spatiale, et leur association avec les taux d'incidence entraîne une auto-corrélation des taux d'incidence supérieure à celle induite uniquement par les processus infectieux. Mais cette hypothèse est encore à vérifier.

5.3.3.c. Discussion

Globalement, les indicateurs construits pour l'étude de modélisation ne permettent d'expliquer qu'une part très faible de l'hétérogénéité spatiale des taux d'incidence à la résolution des îlots à Nouméa en 2009. Il nous a donc manqué certaines variables clé. Parmi les variables pouvant potentiellement influencer la transmission virale et non prises en compte dans cette étude il y a la lutte anti-vectorielle et son corollaire, la résistance des populations de moustiques aux insecticides utilisés, l'immunité de population, des variables individuelles non accessibles via des données SIG comme les mesures individuelles de protection mises en place (climatisation, utilisation de moustiquaires, utilisation de répulsifs anti-moustique, etc.), le climat, les mouvements humains ou les variables liées à l'aspect entomologique de la transmission, comme la densité vectorielle ou la présence de gîtes larvaires.

L'immunité de population a vraisemblablement peu joué sur la distribution spatiale des cas de dengue en 2009. En effet, la dernière épidémie due au virus DENV-4 date de 1988, ce qui a laissé plus de 21 ans au renouvellement d'une population d'hôtes susceptibles.

Contrairement à l'échelle des communes, la lutte anti-vectorielle est homogène spatialement sur la ville de Nouméa. Elle est effectuée par des agents spécialisés et bien formés. De plus en 2009, l'ampleur de l'épidémie a conduit les agents du SIPRES à effectuer des frappes de zone au cours desquelles l'ensemble des quartiers de Nouméa a été pulvérisé une fois par semaine. Les différences d'implémentation de la lutte anti-vectorielle municipale ne peuvent donc expliquer l'hétérogénéité des taux d'incidence à l'échelle de Nouméa. Une étude est en cours qui étudie la résistance des différentes populations de moustiques de Nouméa. Les résultats préliminaires montrent une hétérogénéité spatiale des résistances en 2010 (soit à la fin de l'épidémie). Les épandages municipaux ayant été homogènes en 2009, il est possible que l'hétérogénéité spatiale de la résistance des populations de moustiques aux insecticides provienne de l'utilisation intensive d'insecticides domestiques dans certains quartiers (Dr. Mathieu-Daudé, communication personnelle). Cependant, l'échantillonnage spatial de cette étude de résistance

est insuffisant pour pouvoir conclure sur le lien potentiel entre les résistances observées et les taux d'incidence de la dengue. Il serait intéressant, lors de la prochaine épidémie, de conduire des études de résistance en augmentant le nombre de populations de moustiques échantillonnées pour pouvoir effectuer une cartographie.

Avec une seule station météorologique sur la commune de Nouméa, il n'est pas possible d'évaluer l'influence du climat sur la distribution spatiale des cas de dengue. A ces échelles, si le climat a une influence, ce serait plutôt via l'existence de microclimats locaux (*e.g.* humidité plus élevée dans les zones ombragées…) très difficiles à quantifier.

Lors de l'enquête, la comparaison des zones « cas » et des zones « témoin » a révélé trois principaux facteurs influençant la transmission de la dengue à l'échelle locale : la promiscuité entre les personnes (le fait d'habiter dans un habitat possédant un mur commun avec la maison voisine est un facteur de risque), l'environnement très local du moustique vecteur ainsi que l'implication de chaque individu dans le maintien d'un environnement « propre » (la présence d'un jardin aéré et/ou bien entretenu est un facteur protecteur tandis que la présence d'un jardin dense et/ou pas ou peu entretenu est un facteur de risque), et enfin les mesures individuelles de lutte (le fait de mettre en place des mesures de protection anti-moustique individuelles est un facteur de protection). Le type d'habitat, quant à lui, est difficile à classer, le fait d'habiter dans une maison à 1 étage représentant un facteur de risque, mais le fait d'habiter dans une maison sans étage étant un facteur protecteur. Les résultats de l'enquête sont globalement cohérents avec ceux de l'étude de modélisation.

Le lien entre densité de population et risque d'attraper la dengue est moins clair que celui détecté à l'échelle des communes. A l'échelle de Nouméa, ce lien est complexe et difficile à interpréter. L'enquête a montré que la proximité entre habitations est un facteur de risque (habiter dans une rue large est un facteur de protection, tandis qu'avoir un une habitation dont l'un des murs est mitoyen avec l'habitation voisine est un facteur de risque). Dans le même temps, l'étude de modélisation montre une corrélation, bien que faible, entre densité d'adresse et taux d'incidence élevé de dengue, tandis qu'il n'existe aucune corrélation significative entre densité de population et sévérité de l'épidémie de 2009. Il semble donc que la promiscuité entre les hôtes soit un facteur de risque à l'échelle de Nouméa (maisons proches). En revanche, l'absence de corrélation entre densité de population et taux d'incidence de la dengue ne peut alors s'expliquer que si le fait d'habiter dans une maison est un facteur de risque. En effet, les îlots dans lesquels la majeure partie de la population habitent en immeuble ont certainement une densité de population élevée. La corrélation négative entre la proportion de personnes habitant en appartement et les taux d'incidence de la dengue pourraient ainsi masquer l'effet de la densité de population lors de l'étude de modélisation. La promiscuité entre habitations augmente la probabilité de contact entre un moustique infectieux et un hôte susceptible, surtout dans un contexte comme la dengue où l'espèce vectrice a un rayon d'action dépassant rarement le kilomètre au cours de sa vie (Harrington et al., 2005). L'influence de fortes densités locales d'habitations sur les taux d'incidence pourrait expliquer le caractère extrêmement focal et explosif des foyers épidémiques de dengue lorsque

l'espèce vectrice en cause est *Aedes aegypti* (Lambrechts et al., 2010). Il est également dommage que nous n'ayons pas pu avoir accès à la variable « nombre d'habitants par foyer » à l'échelle des îlots, afin de pouvoir comparer son influence sur les taux d'incidence à deux échelles différentes.

Nous n'avons trouvé aucune corrélation entre les variables environnementales issues des indices de télédétection et la dengue, ce qui pourrait surprendre, étant donnée la très haute résolution de l'image satellite utilisée (4,5 m). Or l'enquête a mis en évidence des facteurs de risque liés à l'environnement proche du moustique. Les facteurs étudiés par l'enquête sont des facteurs non accessibles via la télédétection (entretien des jardins, type de végétation, présence de gîtes potentiels….). L'incohérence entre les deux études peut être expliquée par le mode de ponte des *Aedes aegypti*. En effet, contrairement aux moustiques du genre *Anophèle*, vecteurs du palu, qui pondent leurs œufs dans des environnements aquatiques caractérisables par télédétection (Tran et al., 2008), les moustique du genre *Aedes* sont capables de pondre dans des gîtes contenant à peine quelques millilitres d'eau (Delatte et al., 2008), et indétectables autrement que par l'œil humain. Ce résultat montre ici la limite de l'approche par modélisation dans l'étude des facteurs de risque liés au moustique vecteur, et les apports des enquêtes de terrain. La mise en évidence, par l'enquête, des facteurs environnementaux comme facteurs de risque est cohérente avec la relation temporelle saisonnière trouvée entre la densité de moustiques et les taux d'incidence de la dengue (Descloux et al., 2012).

L'impact du type d'habitat (maison vs appartement) sur les taux d'incidence locaux est relativement facile à expliquer. La présence d'un jardin favorise la présence de gîtes larvaires, et donc la présence de moustiques. D'autre part, les personnes habitant dans des maisons passent certainement plus de temps à l'extérieur que les personnes habitant en appartement (entretien du jardin, barbecue…). La présence d'un jardin, en plus d'augmenter la densité vectorielle pourrait ainsi favoriser les contacts hôte/vecteur.

Enfin, en ce qui concerne les facteurs individuels de protection, à l'échelle des communes, nous avions montré que les facteurs socio-économiques représentatifs d'un mode de vie tribal étaient prédictifs d'un risque de circulation virale accru au cours des épidémies. A l'échelle des quartiers, les facteurs socio-économiques ne sont que faiblement liés aux taux d'incidence, mais ils apparaissent comme prépondérants par rapport aux variables environnementales (voir Table 5.5). Enfin l'enquête a fait ressortir la mise en place de mesures de protection individuelle comme facteur protecteur. La forte cohérence des trois résultats montre l'importance de l'implication de chacun des individus dans la diminution du risque de contracter la dengue. Ces résultats montrent l'importance d'une approche préventive dans la diminution du risque infectieux, approche utilisant des campagnes de communication et de prévention destinées à responsabiliser chaque citoyen dans la mise en œuvre de mesures de protection.

L'autocorrélation résiduelle après avoir retiré l'effet de variables externes est compatible avec les valeurs déjà publiées de la littérature (Aldstadt et al., 2012; Honório et al., 2009; Liebman et al., 2012; Mammen Jr et al., 2008; Morrison et al., 1998; Rabaa et al., 2013; Salje et al., 2012; Tran et al., 2004;

Vazquez-Prokopec et al., 2010; Yoon et al., 2012). L'auto-corrélation résiduelle de 1,5 kilomètre peut être due au processus infectieux lui-même : l'apparition d'un cas à un endroit donné influence l'apparition de cas dans un rayon de 1,5 km. Le caractère très localisé de l'auto-corrélation provient du comportement des moustiques vecteurs, dont le rayon d'action dépasse rarement le kilomètre. L'étendue de l'auto-corrélation résiduelle observée à Nouméa ainsi que les résultats de plusieurs études ayant montré que les *Aedes* sont capables de voler plus de 100 m pour trouver des gîtes de ponte devraient conduire à réfléchir au périmètre défini pour la lutte anti-vectorielle péri-focale, qui est actuellement de 100 m à Nouméa.

Enfin, un facteur important non pris en compte est le mouvement humain. En effet, à l'échelle locale, c'est le seul facteur dont le rôle est stable entre les différentes études (Adams and Kapan, 2009; Barmak et al., 2011; Honório et al., 2009; Liebman et al., 2012; Penna and Lucia, 2004; Rabaa et al., 2013; Reiner et al., 2014; Stoddard et al., 2013; Vazquez-Prokopec et al., 2010). A Nouméa, la zone industrielle de Ducos est une zone potentiellement à haut risque pour la dengue. En effet, un grand nombre d'entreprise entreposent du matériel à l'air libre susceptible de permettre la création de nombreux gîtes (casses automobiles, entreprises de bâtiment…). La zone industrielle de Ducos est également un lieu de brassage important de population. Or nous pouvons remarquer sur la carte des taux d'incidence que l'un des endroits les plus touchés par la dengue est à proximité de cette zone. Il serait intéressant de monter des études permettant l'analyse des mouvements humains, et l'analyse du rôle de la zone de Ducos dans la diffusion des épidémies sur Nouméa.

Une limite de l'étude de modélisation est de n'avoir étudié la distribution spatiale des cas qu'en 2009, sans tenir compte de la variabilité inter-annuelle des schémas spatiaux. En effet, les facteurs étudiés (densité de population, mnt, facteurs socio-économiques) sont des variables relativement statiques, avec peu de variations inter-annuelles, ce qui pourrait expliquer qu'elles ne soient pas représentatives du risque si le schéma spatial de 2009 s'écarte de celui des autres années épidémiques. Cependant l'analyse EOF a montré que les schémas spatiaux étaient stables d'une année à l'autre (87,2 % de variance expliquée par le premier axe). Le schéma spatial de 2009 s'écarte de celui des autres années, notamment par des taux d'incidence légèrement inférieurs dans la zone industrielle de Ducos et de Rivière Salée par rapport aux années 1995, 1998 et 2003 (voir le 2ème EOF, Figure 4.4.E et F.). Cependant cette différence ne représente que 6,9 % de la variance totale des taux d'incidence. Le schéma spatial de 2009 est donc représentatif des autres années d'apparition d'un nouveau sérotype. Au vu des résultats de l'analyse EOF, il aurait peut-être été intéressant d'effectuer une analyse sur la moyenne des schémas spatiaux de 1995, 1998, 2003 et 2009. Nous pouvons noter ici l'importance d'effectuer une analyse exploratoire approfondie de la variabilité spatio-temporelle des données épidémiologiques avant d'effectuer des études analytiques des facteurs de risque. En effet, si les schémas spatiaux n'avaient pas été stables, il aurait fallu porter notre attention sur des facteurs de risque non statiques, présentant une variabilité temporelle inter-annuelle importante.

Une autre limite de l'étude est la non-exhaustivité du géo-référencement des cas. Le biais introduit par le non géo-référencement des cas survenus dans les squats est peut-être ce qui conduit à observer une corrélation positive entre les taux d'incidence de la dengue et le pourcentage de foyers reliés au réseau électrique général (« élec1 ») et une absence de corrélation entre les variables décrivant le bâti et les taux d'incidence. Il est possible que d'autres biais de non géo-référencement existent et que ces biais soient à l'origine de la faiblesse du signal statistique détecté au cours de cette étude.

Une dernière limite de l'étude provient de la nature des variables modélisées à ces échelles très locales. Il existe un problème bien connu en épidémiologie spatiale inhérent à l'étude de données agrégées par unité spatiale dont les limites sont définies de manière arbitraire. En effet, si les cas enregistrés sont distribués de manière très hétérogène au sein des zones d'agrégation, une modification, même légère, des limites des zones d'agrégation peut conduire à des cartes spatiales d'apparence très différentes. Ce phénomène est connu sous le nom de « MAUP » en anglais, pour « Modifiable Area Unit Problem » (Elliott and Wartenberg, 2004). L'augmentation de la résolution de l'analyse devrait limiter ce problème puisqu'elle permet d'améliorer l'homogénéité des variables étudiées au sein de chaque unité spatiale. Cependant, l'augmentation de la résolution d'analyse introduit également une composante aléatoire importante dans l'estimation des taux d'incidence. Une erreur minime dans le nombre de cas enregistrés peut alors conduire à l'apparition artificielle de zones de transmission ou d'absence de transmission. Une méthode possible pour limiter ce phénomène est l'utilisation de méthode bayesiennes pour lisser les données (Elliott and Wartenberg, 2004).

L'augmentation de la résolution de l'analyse (passage d'une résolution « quartiers » à une résolution « îlots ») a également introduit un grand nombre d'entités spatiales pour lesquelles les valeurs de la variable réponse ou des variables explicatives sont égales à zéro, ce qui réduit fortement la performance des modèles ne prenant pas en compte cette particularité. Peut-être les liens entre les variables étudiées et les taux d'incidence de la dengue pourraient être ici mieux analysés par l'emploi de modèles linéaires généralisés de type « zero-inflated Poisson model ». Un moyen possible de gérer cette difficulté est d'effectuer des analyses de régression logistique, pour identifier les facteurs associés à l'absence complète de cas dans un îlot (vs présence de cas, quel que soit le nombre de cas).

CHAPITRE 6 : DISCUSSION ET PERSPECTIVES

6.1. Synthèse des résultats 293

6.1.1. Tableau récapitulatif 293
Précisions sur les cibles de modélisation 293
6.1.1. Structure temporelle de la dengue et facteurs influençant 294
6.1.1.a. Echelle Pacifique 294
6.1.1.b. Liens entre échelle Pacifique et échelles territoriales ou nationales 296
6.1.1.c. Echelle territoriale 297
6.1.1.d. Lien entre échelle territoriale et locale 297
6.1.1.e. Echelle locale de Nouméa 298
6.1.2. Structure spatiale de la dengue et facteurs influençant 298
6.1.2.a. Echelle régionale et lien avec les dynamiques nationales ou territoriales 298
6.1.2.b. Echelles territoriales et locales, et lien avec la dynamique à l'échelle du Pacifique 299

6.2. Implications opérationnelles 301

6.3. Perspectives 303

6.3.1. Facteurs associés à la dynamique spatiale des épidémies dans le Pacifique 303
6.3.1.a. Dynamique inter-annuelle 303
6.3.1.b. Diffusion régionale 304
6.3.2. Facteurs associés à la dynamique spatiale des épidémies en Nouvelle-Calédonie 304
6.3.3. Mécanismes impliqués dans la dynamique spatiale 305
6.3.4. Estimation des paramètres de transmission en Nouvelle-Calédonie 306
6.3.4.a. Nombre de cas introduits par an 306
6.3.4.b. Proportion de formes asymptomatiques 307
6.3.4.c. Virémie des patients et dose infectieuse 307
6.3.4.d. Compétence vectorielle 308
6.3.4.e. Mortalité des *Aedes aegypti* 308
6.3.4.f. Renouvellement de la population 308
6.3.5. Intégration des échelles 309
6.3.6. Autres arboviroses 309

L'objectif premier de cette thèse était d'améliorer nos connaissances sur l'épidémiologie de la dengue, une maladie à transmission vectorielle d'une importance sanitaire croissante, ayant un fort impact économique et pour laquelle il n'existe à l'heure actuelle aucun vaccin ou traitement spécifique. Durant cette thèse, nous avons eu l'opportunité relativement unique de pouvoir répondre à des questions sur l'épidémiologie d'une maladie vectorielle à différentes échelles spatiales, allant d'une échelle régionale à une échelle très locale. Ce travail nous a permis d'obtenir une vision globale de l'épidémiologie de la dengue dans la Pacifique Sud. Nous nous efforcerons donc dans un premier temps d'effectuer une synthèse des résultats. Ces résultats ayant déjà été discutés séparément dans chacun des chapitres précédents, nous nous limiterons ici à en faire une synthèse rapide et à souligner les connexions existant entre les différentes études. Dans un deuxième temps nous dégagerons les implications opérationnelles. Enfin, nous synthétiserons les pistes futures de recherche.

6.1. Synthèse des résultats

6.1.1. Tableau récapitulatif

Les résultats apportés par chacune des études menées au cours de cette thèse sont synthétisés dans la Table 6.1. Chaque case correspond à une étude potentiellement réalisable. Chaque ligne du tableau correspond à une échelle spatio-temporelle : échelle régionale, territoriale ou locale pour les échelles spatiales, et études inter-annuelles ou intra-annuelles pour les études temporelles. Les colonnes correspondent à la nature des études. Les deux premières colonnes correspondent aux études de la variabilité spatio-temporelle des données épidémiologiques uniquement, n'intégrant aucun facteur de risque. Ces études ont eu deux objectifs distincts : i) la caractérisation de la distribution spatiale et de la distribution temporelle des cas de dengue à chacune des échelles, avec identification des structures principales de la distribution spatiale des cas de dengue au cours du temps au moyen d'analyses EOF ; ii) la détection d'éventuels processus de propagation, au moyen d'analyses descriptives ou d'analyses en ondelette. Les deux dernières colonnes du tableau correspondent à des études analytiques des facteurs environnementaux associés à la variabilité spatio-temporelle des cas de dengue. En fonction des échelles spatio-temporelles et des données disponibles, nous avons étudié soit les facteurs associés à un risque d'émergence épidémique, soit ceux associés à la sévérité des épidémies.

Précisions sur les cibles de modélisation

Pour l'ensemble des études des facteurs associés à la sévérité des épidémies, nous avons utilisé les taux d'incidence comme variable cible. Pour les études des facteurs associés à l'émergence épidémique, nous avons dû, au préalable, définir ce qu'était un « évènement épidémique », et plusieurs méthodes ont été utilisées : à l'échelle du Pacifique, une année a été définie comme épidémique si le taux d'incidence annuel dans le pays était supérieur à 0,025 du taux d'incidence annuel maximal dans le pays entre 1971

et 2009. En Nouvelle-Calédonie et à Nouméa, une année a été définie comme épidémique si le taux d'incidence du territoire appartenait au tercile supérieur des taux d'incidences de 1971 à 2011. Enfin, pour le modèle mécanistique d'émergence des épidémies, nous avons modélisé le nombre de reproduction de base, $R0$, dont la valeur seuil de 1 permet de distinguer les pas de temps au cours desquels le risque d'émergence existe ($R0 > 1$) des pas de temps au cours desquels le risque d'émergence est nul ($R0 < 1$). D'autres définitions peuvent éventuellement être utilisées en fonction du contexte épidémiologique (Farrington et al., 1996; Serfling, 1963; Stroup et al., 1989).

6.1.1. Structure temporelle de la dengue et facteurs influençants

6.1.1.a. Echelle Pacifique

A l'échelle du Pacifique, du fait de la disparité des réseaux de surveillance entre les différents pays, et du fait de l'évolution des modalités de surveillance dans le temps, nous n'avons pu étudier que le risque d'émergence épidémique, et non la sévérité des épidémies. Du fait de la résolution annuelle des données, seule la variabilité inter-annuelle du risque d'émergence des épidémies a pu être étudiée. Dans le Pacifique, la dengue se propage depuis les années 1970 sous la forme de vagues pandémiques survenant tous les 5 à 7 ans et touchant l'ensemble des îles. La fréquence de ces vagues est stable depuis les années 1990. Contrairement aux autres régions du monde (Vong et al., 2010; Zhang et al., 2005), la co-circulation de plusieurs sérotypes est limitée, à l'exception des périodes de circulation régionale du sérotype 3, concomitante avec des épidémies éparses dues à d'autres sérotypes. Chaque vague est due à la diffusion régionale d'un sérotype dominant introduit à partir des zones endémiques d'Asie et d'Amérique Latine (A-Nuegoonpipat et al., 2004; Cao-Lormeau et al., 2011; Chungue et al., 1993; Dupont-Rouzeyrol et al., 2014; Steel et al., 2010).

Les raisons de l'émergence régionale régulière d'un seul des quatre sérotypes sont inconnues. En Asie, le type de dynamique observé, avec alternance régulière des sérotypes, a été lié à des phénomènes immunitaires au sein des hôtes (Adams et al., 2006; Cummings et al., 2005; Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006), et à des facteurs de variabilité génétique intra-sérotypique (Zhang et al., 2005; Ty Hang et al., 2010; Duong et al., 2013b).

Dans le Pacifique, la synchronisation régionale des épidémies sous forme de vagues est due en faible partie au phénomène climatique El Niño, avec un plus grand nombre de pays touchés par une épidémie durant la phase « La Niña » de ce phénomène, lorsque l'Ouest du Pacifique Sud est plus chaud et plus humide que la normale (voir 3.1.2.c page 146 et (Hales et al., 1999)). Cependant, d'après nos résultats, ENSO n'explique que 9 % de la variabilité inter-annuelle du nombre de pays touchés par an. La synchronisation des épidémies dues à un même sérotype semble également due en partie à un

Table 6.1 : Synthèse des résultats par échelle spatio-temporelle et par nature des études abordées. Les cellules grisées correspondent à un travail propre à la thèse. Les cellules en blanc correspondent à la participation à un travail collectif, soit avec d'autres doctorants, soit par l'encadrement de stagiaire.

Echelle \ Nature		Etudes spatio-temporelles des données épidémiologiques		Etudes des facteurs de risque	
		Structures spatiales et temporelles	Processus de propagation	émergence épidémique	intensité des épidémies
Régionale	Inter-ann	- Vagues pandémiques tous les 5 à 7 ans	Pas de schéma de propagation	- Cible = nombre de pays touchés par une épdiémie par an - Facteur faiblement associé : ENSO	-
	Int	-		-	-
	Spatiale	- Pas de structure spatiale de de la fréquence des épidémies - Pas de structure spatiale du nombre d'épidémies par pays - Structure spatiale des incidences difficile à interpréter		-	-
Territoriale	Inter-annuelle	- Cible : Taux d'incidence spatio-temporels - Structure spatiale de la dengue cohérente d'une année à l'autre - Facteurs : immunité de population ?	Pas de schéma de propagation	- Cible = épidémie vs pas épidémie - Facteurs associés : NBJ_TX_32_JFM et NBJ_RH_95_J - Facteurs non associés : ENSO, densité vectorielle	-
	Intra-a	-		- Cible = Taux d'incidence mensuelle - Facteurs associés = température et précipitations	-
	Spatiale	-		-	- Cible = taux d'incidence - Facteurs associés : température moyenne, nombre de personnes par maison, comportement humain - Facteurs faiblement associés : précipitations, NBJ_TX_32_JFM, NBJ_RH_95_J
Locale	Inter-annuelle	- Cible : Taux d'incidence spatio-temporels - Structure spatiale de la dengue cohérente d'une année à l'autre - Facteurs : immunité de population ?	Pas de schéma de propagation	- Cible : R0 modélisé - Pas de mécanisme identifié	-
	Intra-annu	-		- Cible : R0 modélisé - Mécanisme = influence de la température sur les cycles de vie d'Aedes aegypti	-
	Spatiale	-		-	- Cible = taux d'incidence - Facteurs associés : ??? Environnement très fin (cf enquête) - Facteurs faiblement associés : ??

phénomène de diffusion régionale du virus, qui a été mis en évidence par des études statistiques (Hales et al., 1999) ou phylogénétiques (A-Nuegoonpipat et al., 2004; Cáceres et al., 2008; Cao-Lormeau et al., 2011, 2014a; Dupont-Rouzeyrol et al., 2014; Effler et al., 2005; Li et al., 2010; Steel et al., 2010). Ce phénomène de diffusion régional est encore à étudier. Nous n'avons pas, au cours de cette thèse, pu mettre en évidence de schéma de propagation stable dans le temps. Il conviendrait donc, pour chacune des vagues épidémiques, d'établir une chronologie précise de la diffusion virale, afin de pouvoir au mieux identifier les facteurs favorisant cette diffusion.

6.1.1.b. Liens entre échelle Pacifique et échelles territoriales ou nationales

La fréquence de survenue des épidémies n'a pas augmenté globalement dans le Pacifique. Cependant ce résultat masque l'hétérogénéité existant entre les différents pays, certains pays comme la Nouvelle-Calédonie voyant le nombre d'épidémies augmenter au cours du temps, tandis que d'autres pays voient le nombre d'épidémies diminuer, sans qu'aucune structure spatiale n'ait pu être mise en évidence. Ce résultat, ainsi que l'absence de schéma de propagation stable et de schéma spatial stable à l'échelle du Pacifique semblent indiquer que les facteurs de variabilité inter-annuelle de survenue des épidémies dans chacun des pays sont sous le contrôle de facteurs locaux propres à chacun des pays. Ceci est cohérent avec les disparités existant entre les études de la variabilité inter-annuelle de la dengue réalisées à des échelles nationales dans différents pays (Thaïlande, Puerto Rico, Mexique, et autres villes du Pacifique (Cazelles et al., 2005; Hales et al., 1999; Johansson et al., 2009b)). Nous pensons que les disparités spatiales du lien temporel inter-annuel entre dengue et facteurs de risque sont dues à la nature multi-factorielle de la dynamique de la dengue, et que l'émergence d'une épidémie est sous le contrôle de facteurs limitants qui peuvent varier d'un pays à l'autre, en fonction du contexte géographique, socio-économique, démographique ou épidémiologique (nombre d'introductions, renouvellement de population, taille critique de population nécessaire au développement d'une épidémie, présence de plusieurs espèces vectrices, etc.). En effet, par exemple, comme suggéré par Johansson (Johansson et al., 2009a), il est probable que la température ait plus d'influence sur la dynamique épidémique dans les pays où la température moyenne est suffisamment basse pour empêcher la circulation virale pendant une partie de l'année, comme la Nouvelle-Calédonie, que dans les pays où la température est favorable à la transmission de la dengue durant l'année entière. Nous pensons donc que la compréhension de la variabilité inter-annuelle de la dengue pourrait être améliorée en effectuant une cartographie mondiale des différentes études réalisées sur le sujet, et, dans le futur, en faisant l'effort d'intégrer des variables autres que climatiques dans les études dynamiques.

Du fait de la diffusion régionale des virus, la survenue de vagues pandémiques à l'échelle du Pacifique a une forte influence sur l'épidémiologie de la dengue dans chacun des pays. En effet, par exemple en Nouvelle-Calédonie, l'ensemble des épidémies observées ont été causées par le sérotype dominant circulant dans le Pacifique, à l'exception de l'épidémie de 2013 due au virus DENV-1 alors que le

sérotype circulant dans le Pacifique était le virus DENV-4 (Dupont-Rouzeyrol et al., 2014). La circulation régionale d'un sérotype augmente donc potentiellement le risque annuel de survenue d'une épidémie dans chacun des pays. Le risque de survenue d'une épidémie dans un pays donné serait donc une combinaison du risque de diffusion régionale du virus, et du risque d'émergence propre à chacun des pays. En Nouvelle-Calédonie, il conviendrait d'intégrer ce risque régional de manière quantitative dans le système national d'alerte précoce du risque de survenue d'une épidémie.

6.1.1.c. Echelle territoriale

En Nouvelle-Calédonie, la dengue survient sous une forme épidémique saisonnière. Tous les 5 à 10 ans, un nouveau sérotype est introduit sur le territoire, causant une épidémie majeure. Le sérotype introduit est le plus souvent le sérotype dominant dans le Pacifique. Une fois un sérotype introduit, il circule alors sur le territoire pendant plusieurs années, parfois séparées par des périodes silencieuses.

D'après les autorités sanitaires, entre 10 et 30 cas de dengue sont importés chaque année sur le territoire. Ce nombre est très certainement sous-estimé, du fait de la forte proportion de formes asymptomatiques de dengue (Baaten et al., 2011; Grange et al., 2014). Malgré l'introduction régulière de personnes infectieuses, les épidémies ne se développent que tous les 3 à 5 ans. Nous avons montré dans cette thèse que le risque d'émergence d'une épidémie en Nouvelle-Calédonie était accru lorsque des conditions climatiques précises à Nouméa sont réunies, *i.e.* lorsqu'il y plus de 11 jours an Janvier/Février/Mars où la température maximale dépasse 32 °C, ou lorsqu'il y a plus de 10 jours en Janvier où l'humidité relative dépasse 95 % (Descloux et al., 2012). Le phénomène ENSO, quant à lui, n'a pas d'influence sur la variabilité inter-annuelle du risque d'émergence d'une épidémie en Nouvelle-Calédonie, ce qui est en accord avec l'étude de Hales (Hales et al., 1999).

En Nouvelle-Calédonie, lors des années épidémiques, la dengue présente un caractère saisonnier, avec un démarrage en décembre/janvier, un pic en avril, et une fin des épidémies en juin. Ce caractère saisonnier est dû à l'influence des variables climatiques sur le cycle de vie des moustiques. Le pic épidémique est observé un à deux mois après le pic de température, et est en phase avec le pic de précipitations, ce qui est cohérent avec les mécanismes individuels de transmission du virus (Focks et al., 1993a).

6.1.1.d. Lien entre échelle territoriale et locale

En Nouvelle-Calédonie, à l'échelle du territoire, la sévérité des épidémies et le risque inter-annuel de survenue d'une épidémie est étroitement corrélé à la dynamique de la dengue à Nouméa. A tel point que le risque inter-annuel de survenue d'une épidémie sur le territoire entier peut être expliqué en utilisant uniquement les variables météorologiques enregistrées à Nouméa (Descloux et al., 2012). Nouméa, qui concentre près de la moitié de la population calédonienne semble donc avoir une

importance primordiale dans le risque d'émergence épidémique. Ce constat est renforcé par le fait que quelle que soit la commune dans laquelle sont détectés les premiers cas de dengue au cours d'une épidémie, le virus circule toujours de manière précoce à Nouméa, qui pourrait donc agir comme amplificateur du virus, et réservoir de virus pour les autres communes.

6.1.1.e. Echelle locale de Nouméa

A l'échelle locale de Nouméa, la dynamique inter-annuelle du risque d'émergence épidémique ainsi que la dynamique saisonnière des épidémies sont toutes deux sous le contrôle de facteurs climatiques (Descloux et al., 2012). Par contre, en ce qui concerne les facteurs entomologiques, si un lien a pu être mis en évidence entre l'évolution saisonnière des densités vectorielles et le caractère saisonnier des épidémies de dengue, aucun lien statistique n'a pu être mis en évidence entre la variabilité inter-annuelle des densités vectorielles et la variabilité inter-annuelle du risque d'émergence épidémique. Ce résultat montre qu'en Nouvelle-Calédonie, les densités vectorielles sont toujours suffisamment élevées pour permettre le développement d'une épidémie, *i.e.* l'indice maison est toujours supérieur à 1 % (Pontes et al., 2000) et l'indice de Breteau est toujours supérieur à 4 (Sanchez et al., 2010).

Malgré le lien statistique observé entre température, humidité relative et risque d'émergence des épidémies à Nouméa (Focks et al., 1993a), la modélisation mécanistique de l'influence de ces variables climatiques sur les paramètres du cycle de vie d'*Aedes aegypti* qui peuvent influencer la transmission de la dengue n'a pas permis de découvrir le mécanisme par lequel les variables climatiques et le risque d'émergence des épidémies de dengue à Nouméa sont liés. Ces résultats illustrent bien la difficulté à lier les phénomènes de transmission virale survenant au niveau individuel aux observations statistiques faites sur des populations entières.

6.1.2. Structure spatiale de la dengue et facteurs influençants

6.1.2.a. Echelle régionale et lien avec les dynamiques nationales ou territoriales

A l'échelle du Pacifique, les disparités entre les réseaux de surveillance des différents pays ne permettent pas de comparer les niveaux d'incidence des pays entre eux. Nous avons donc observé uniquement les disparités spatiales du risque d'émergence épidémique, tel que comptabilisé par le nombre d'épidémies survenues dans chacun des vingt PTIP étudiés. A cette échelle régionale, nous avons mis en évidence que les schémas de propagation sont différents d'une vague épidémique à l'autre, que la structure spatiale des pays touchés par une épidémie varie d'une année à l'autre, et varie d'une vague épidémique à l'autre, et que la structure spatiale du nombre d'épidémies par pays ou de l'évolution de la fréquence des épidémies au sein de chacun des pays ne présente aucun gradient spatial. Malgré le phénomène de

diffusion régionale d'un même sérotype à l'ensemble des PTIP (A-Nuegoonpipat et al., 2004; Cao-Lormeau et al., 2011; Steel et al., 2010), il ne semble pas y avoir de facteur environnemental structurant le risque d'émergence épidémique dans l'espace à cette échelle-là. Nous pensons donc que les cartes annuelles observées sont la résultante de l'influence locale de facteurs propres à chacun des pays sur la dynamique inter-annuelle de la dengue dans chacun des pays. Il serait très certainement utile, à l'image de ce qui a été fait en Nouvelle-Calédonie (Descloux et al., 2012), d'identifier dans chacun des pays les facteurs associés à un risque accru d'émergence épidémique, et d'en effectuer une carte. Cela pourrait certainement aider à identifier les zones où la température joue un rôle important, et aider à mieux anticiper le risque d'émergence dans les zones tempérées comme l'Europe, en identifiant éventuellement des valeurs seuils de température en-deçà ou au-delà desquelles le risque d'émergence est limité ou favorisé.

6.1.2.b. Echelles territoriales et locales, et lien avec la dynamique à l'échelle du Pacifique

A l'échelle du territoire, la circulation virale moyenne au cours des années épidémiques est hétérogène. Les taux d'incidence sont plus élevés dans les communes présentant une température moyenne élevée et une densité moyenne de personnes par foyer élevée. Le mode de vie tribal, avec des personnes vivant en communauté, passant une grande partie du temps hors de la maison, et ne prenant pas forcément de précautions pour se prémunir des piqures de moustique est également associé à un risque de transmission élevé. D'ici à 2100, une augmentation de 3 °C de la température moyenne, comme prévu par les modèles du GIEC, pourrait faire doubler le nombre de cas de dengue en moyenne. Les niveaux de précipitations dans une commune donnée n'influencent pas le risque de transmission virale.

A l'échelle de Nouméa, la circulation virale se fait sous la forme de cluster de cas, dont le rayon est de ~ 1,5 km. En moyenne la circulation virale est également hétérogène. L'analyse des facteurs de risque a mis en évidence qu'habiter dans une maison accroît légèrement le risque de transmission. Cependant aucun des facteurs étudiés (indicateurs socio-économiques, indicateurs de densité de végétation, indicateurs de densité de population et indicateurs des types de bâti) ne permet de prédire les taux d'incidence de manière précise, suggérant que les facteurs étudiés n'ont qu'une influence faible sur la variabilité spatiale des taux d'incidence. Une enquête de terrain menée en parallèle de l'étude de modélisation a permis de mettre en évidence que les facteurs de risque d'une transmission accrue sont des facteurs très locaux, liés au comportement des humains (mise en place de mesures de protection individuelle, entretien des jardins), difficilement accessibles par des études de modélisation.

Les facteurs intégrés à l'échelle du territoire et à l'échelle locale n'étant pas exactement les mêmes, la comparaison des résultats est difficile. Par exemple, la température joue un rôle primordial dans l'hétérogénéité spatiale de la circulation virale sur le territoire. Bien que le climat sur Nouméa soit probablement moins contrasté qu'à l'échelle du territoire, l'influence des micro-climats sur

l'hétérogénéité spatiale du risque de transmission à Nouméa n'a pu être étudiée, faute de données climatiques à haute-résolution. A l'échelle du territoire, puisque la deuxième variable à avoir été identifiée comme facteur de risque est le nombre de personnes par foyer, il semble que les densités locales de population, et particulièrement les phénomènes de transmission à l'intérieur d'un même foyer aient une influence importante sur les taux d'incidence, suggérant que les structures spatiales observées sur le territoire résultent de l'agrégation de foyers de transmission très locaux. Il est alors surprenant que l'association statistique entre densité de population et intensité de la transmission virale soit inexistante à l'échelle locale de Nouméa. Enfin, à l'échelle du territoire, les variables socio-économiques sont des facteurs importants de variabilité spatiale de la transmission virale, tandis qu'à l'échelle de Nouméa, ces mêmes variables ne sont pas associées à un risque accru de transmission. Ceci peut être expliqué par le fait qu'à Nouméa, la mixité socio-économique est importante au sein des îlots, et les variables socio-économiques ne reflètent pas un mode de vie particulier, tandis que sur le territoire les différentes populations de Nouvelle-Calédonie sont spatialement séparées, le paysage socio-économique est beaucoup plus contrasté qu'à Nouméa, et les variables socio-économiques doivent être interprétées comme représentatives d'une mode de vie global, et non pas en termes de niveaux purement économiques.

Enfin, un résultat important mis en évidence dans cette thèse : que ce soit à l'échelle territoriale ou à l'échelle locale, et quels que soient les facteurs associés à un risque de transmission accru à un endroit donné, lors de la ré-émergence d'un même sérotype dans les trois années qui suivent une épidémie la dengue circule préférentiellement dans les zones peu impactées les années précédentes, mettant en évidence le rôle important de l'immunité de population sur la variabilité inter-annuelle du risque spatial de transmission virale. L'immunité de population semble donc un facteur primordial de la dynamique spatiale de la dengue, que ce soit par son impact sur la dynamique inter-annuelle dans les pays d'Asie (Cummings et al., 2005; Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006), ou son impact sur les structures spatiales à des échelles locales (Salje et al., 2012) ou territoriales comme mis en évidence dans cette thèse. Dans ce contexte, l'introduction d'un vaccin dans les années à venir, même s'il ne protège pas 100 % de la population (Capeding et al., 2014), pourrait avoir des conséquences importantes sur la dynamique spatiale des pays dans lesquels il serait introduit.

6.2. Implications opérationnelles

Sur un plan opérationnel, cette thèse a permis :

- D'identifier les zones à risque à chacune des échelles d'étude ;
- D'émettre des hypothèses sur les facteurs propres au Pacifique ou à la Nouvelle-Calédonie favorisant la circulation virale dans le temps ou dans l'espace, et donc d'émettre des recommandations sur les mesures de lutte qui auront potentiellement le plus d'impact dans la réduction du risque épidémique ;
- De fournir un indicateur annuel du risque d'occurrence des épidémies en Nouvelle-Calédonie.

Dans le Pacifique, du fait de la disparité entre les différents systèmes de surveillance, il est difficile d'identifier si certains pays sont plus touchés que d'autres. La Polynésie Française et la Nouvelle-Calédonie ont été touchées par l'ensemble des vagues épidémiques survenues depuis 1971. Sur six des huit vagues épidémiques, la Polynésie Française a été touchée entre un et trois ans avant la Nouvelle-Calédonie. L'indicateur annuel du risque d'occurrence d'épidémies en Nouvelle-Calédonie basé sur le climat (Descloux et al., 2012) doit donc être interprété à la lumière du contexte régional : le risque épidémique est accru en Nouvelle-Calédonie si un nouveau sérotype est apparu en Polynésie Française les années précédentes.

En Nouvelle-Calédonie, nous avons identifié des conditions climatiques précises favorisant l'émergence d'une épidémie une année donnée, à savoir lorsqu'il y a plus de 11 jours où la température maximale dépasse 32 °C en janvier, février, mars, ou lorsqu'il y a plus de 10 jours où l'humidité relative dépasse 95 % en janvier (Descloux et al., 2012). Un modèle prédictif a également été construit et intégré au système d'évaluation annuel du risque épidémique par les autorités sanitaires. Ce modèle prédictif a une bonne valeur prédictive positive (79 %), mais sa faible valeur prédictive négative (65 %) doit conduire les autorités à interpréter un risque faible de manière prudente.

Une fois une épidémie lancée sur le territoire, quelle que soit l'échelle d'étude (territoriale ou locale), la structure spatiale de la dengue est hétérogène. A l'échelle du territoire, les zones les plus impactées par la dengue en moyenne sont les communes situées sur le Nord-Est du territoire (Hienghène, Ouégoa, Pouébo). Les facteurs favorisant des taux d'incidence élevés sont une température moyenne élevée, un nombre de personnes par foyer élevé et un mode de vie tribal. A l'échelle de Nouméa, les zones les plus fortement touchées sont Ducos et Rivière Salée, ainsi que quelques quartiers du Grand Nouméa comme Koutio et Robinson. Bien que notre étude sur Nouméa ne permette pas de conclure de manière définitive sur les raisons de cette hétérogénéité à l'échelle locale, il semblerait que des facteurs liés au comportement individuel et influençant l'environnement très local du moustique soient en cause (entretien des jardins, mises en place de mesures de protection personnelle). Que ce soit à l'échelle du territoire ou de Nouméa, il semble donc que les taux d'incidence puissent baisser de manière significative si l'ensemble des habitants met en place des mesures personnelles de protection

individuelle. Ce résultat devrait conduire les autorités à favoriser des campagnes de communication axées sur la responsabilité individuelle de chacun. L'importance du lien entre nombre de personnes par foyer et taux d'incidence de la dengue dans les communes du territoire suggère également que la mise en place de mesures de protection pour les membres d'un foyer dans lequel une personne est infectée pourrait avoir un impact dans la réduction des taux d'incidence.

Les schémas spatiaux observés varient d'une année à l'autre, probablement sous l'influence de l'immunité de population : quelle que soit l'échelle d'étude (territoire, Nouméa), les zones les plus touchés lors de l'introduction d'un sérotype sont celle les moins touchées lors de la ré-émergence du même sérotype dans les 4 à 5 années qui suivent, et vice-versa. Les autorités sanitaires des zones peu impactées une année donnée doivent donc accroitre leur vigilance les années qui suivent.

A des échelles locales, plusieurs études ont montré un impact majeur des lieux hautement fréquentés sur l'augmentation des taux d'incidence (Honório et al., 2009; Stoddard et al., 2013; Vazquez-Prokopec et al., 2010). A Nouméa, il est possible que la zone industrielle de Ducos, une zone hautement fréquentée par l'ensemble de la population, possédant un environnement favorisant la présence de gîtes larvaires (nombreuses zones d'entrepôt de matériel, casses automobiles, etc.), et présentant des taux d'incidence élevés joue un rôle considérable sur la diffusion de l'épidémie au sein de la ville. La quantification précise de ce lien est cependant difficile à effectuer rapidement, car elle nécessiterait l'incorporation dans les études de modélisation de données longues à récolter, comme des données sur les mouvements humains ou des données géographiques sur l'emplacement des commerces à risque. En attendant, au vu du résultat des études publiées dans d'autres villes, il est possible qu'un accroissement de l'effort de lutte ciblé sur la zone industrielle de Ducos ait un impact sur les taux d'incidence sur l'ensemble du Grand Nouméa.

En ce qui concerne la lutte anti-vectorielle, nous n'avons malheureusement pas pu, au cours de cette thèse, tester son efficacité dans la réduction du risque de contracter la dengue à l'échelle du territoire, du fait de l'absence de données sur les mesures réellement implémentées dans chacune des communes. A l'échelle de Nouméa, la lutte est homogène sur l'ensemble de la ville, et il paraît difficile de mener à bien une étude de recherche à proprement parler, une étude de l'efficacité de la LAV devant nécessairement passer par la comparaison des taux d'incidence entre une zone traitée et une zone « témoin » (*i.e.* dans laquelle il n'y a pas d'épandage), ce qui poserait un problème éthique pour les habitants de l'éventuelle zone témoin. Sur le plan temporel, la lutte anti-vectorielle, et notamment la mise en place de campagnes de prévention à partir de 2003, a eu un effet drastique sur la réduction de la densité vectorielle à Nouméa. Cependant, cette réduction n'a pas empêché la survenue d'épidémies. L'expérience passée en Amérique (Gubler, 2004) et à Singapour (Ooi et al., 2006), où la baisse des densités vectorielles a permis de ralentir la transmission de la dengue, du moins de manière transitoire, montre que les indices entomologiques doivent atteindre un seuil extrêmement bas avant qu'une

épidémie ne puisse survenir. Les efforts pour y parvenir sont considérables, et difficiles à maintenir dans le temps : la dengue a re-émergé dans ces deux pays. A Nouméa, l'influence de la LAV sur la sévérité des épidémies est difficilement quantifiable, et il est possible que, bien qu'elle n'empêche pas la survenue des épidémies, la LAV permette de réguler les taux incidence hebdomadaires, et permette d'éviter une surcharge trop importante des hôpitaux.

6.3. Perspectives

6.3.1. Facteurs associés à la dynamique spatiale des épidémies dans le Pacifique

6.3.1.a. Dynamique inter-annuelle

A l'échelle du Pacifique, il est raisonnable de penser que les 4 sérotypes de dengue sont introduits chaque année à partir des pays d'Asie. Les raisons de l'émergence d'un seul sérotype dominant sur les 4 introduits sont inconnues et peuvent inclure un lien épidémiologique avec les pays endémiques d'Asie (Zhang et al., 2005) et d'Amérique Latine (Guzman and Kouri, 2003), des différences génétiques résultant de phénomènes de compétition évolutive entre les différents virus (Zhang et al., 2005; Anderson and Rico-Hesse, 2006; Lee et al., 2010; Ty Hang et al., 2010; Duong et al., 2013a, 2013b), et des facteurs d'immunité de population (Cummings et al., 2005; Ferguson et al., 1999; Hay et al., 2000; Wearing and Rohani, 2006).

D'autre part, les facteurs liés à la variabilité inter-annuelle du nombre de pays touchés par une épidémie sont inconnus. Comme pour les structures spatiales, cette variabilité peut résulter de la diffusion régionale du virus et de l'agrégation des dynamiques inter-annuelles de chacun des pays, dont les facteurs de variabilité seraient locaux. Mais il est également possible que des phénomènes immunitaires soient à l'origine de la période de 5 à 7 ans des cycles de transmission observés dans le Pacifique, en particulier étant donnée la taille restreinte de la population de certains PTIP (plus de la moitié des PTIP ont moins de 90 000 habitants).

Pour répondre à ces deux questions, il conviendrait dans un premier temps de confirmer l'introduction annuelle des 4 sérotypes dans le Pacifique. Pour résoudre cette question, une collaboration avec la CPS, qui recueille les alertes de chacun des pays, s'impose. La **construction d'un modèle de dynamique des populations** incluant les 4 sérotypes permettrait d'évaluer l'impact des phénomènes immunitaires sur la dynamique inter-annuelle. Enfin, des **études phylogénétiques** des différentes lignées virales co-circulant dans le Pacifique permettraient peut-être également d'apporter un éclairage sur les phénomènes d'émergence d'un seul des 4 sérotypes.

Enfin, **des études de modélisation à l'échelle territoriale** ou nationale dans les PTIP autres que la Nouvelle-Calédonie permettraient d'identifier les facteurs associés à la dynamique inter-annuelle de

l'émergence épidémique dans chacun des pays et apporterait certainement un éclairage important. Notamment, si de telles études étaient réalisées, la cartographie des facteurs principaux impliqués dans le risque annuel d'émergence épidémique permettrait certainement d'émettre des hypothèses sur les mécanismes sous-jacents et de guider la construction de modèles mécanistiques.

6.3.1.b. Diffusion régionale

La diffusion régionale du virus existe (A-Nuegoonpipat et al., 2004; Cao-Lormeau et al., 2011; Dupont-Rouzeyrol et al., 2014; Steel et al., 2010), cependant les schémas de propagation diffèrent d'une vague épidémique à l'autre. Il est vraisemblable que la diffusion se fasse par l'intermédiaire des mouvements humains. Afin d'identifier des connexions privilégiées entre pays, il conviendrait dans un premier temps de décrire précisément l'histoire de la propagation de la dengue dans le Pacifique pour chacune des vagues épidémiques. Cela peut être effectué au moyen **d'analyses phylogénétiques**, comme déjà décrit pour la vague due au sérotype 2 dans les années 1970 (Steel et al., 2010), couplé à **l'analyse des données de surveillance**. La description des processus de propagation par l'analyse des données de surveillance ne donnera vraisemblablement de bons résultats que si les données sont mensuelles. Si des séries temporelles complètes mensuelles ne sont pas disponibles, il faudrait au moins connaître la date de déclaration du premier cas dans chacun des pays. Si ces analyses ne mettent en évidence aucun schéma particulier, les facteurs associés à la diffusion régionale seront certainement à aller chercher parmi les évènements ponctuels créant des mouvements de population internationaux, comme l'organisation d'évènements sportifs ou culturels de grande envergure.

6.3.2. Facteurs associés à la dynamique spatiale des épidémies en Nouvelle-Calédonie

Nous avons vu qu'à Nouméa, le risque annuel d'émergence épidémique est très corrélé au risque annuel d'émergence épidémique sur le territoire. Ce risque est sous le contrôle de conditions climatiques précises. Dans les autres communes, l'intensité de la circulation virale est également sous le contrôle des températures. Par contre, au cours des années épidémiques, toutes les communes ne sont pas forcément atteintes par la dengue. Une question subsiste donc : le risque d'émergence d'un foyer épidémique dans les communes autres que Nouméa est-il lié aux conditions climatiques locales, ou bien est-il régi par un phénomène de diffusion à partir de la commune de Nouméa ? Il conviendrait pour cela de définir ce qu'est un foyer épidémique dans chacune des communes, ce qui permettrait d'identifier à la fois les années pour lesquelles chaque commune a été affectée, ainsi que la semaine de démarrage du foyer épidémique dans chacune des communes. Une fois les séries temporelles de foyers épidémiques définies **dans chacune des communes**, il conviendrait d'**étudier les conditions climatiques associées temporellement à l'émergence de ces foyers** ; ces études temporelles dans chacune des

communes pourraient être réalisées soit sur un plan inter-annuel (pour pouvoir le comparer aux résultats obtenus à Nouméa (Descloux et al., 2012)), soit sur un plan saisonnier. Les études inter-annuelles seront moins robustes sur le plan statistique par rapport à l'étude effectuée à Nouméa, car les cas de dengue n'ont été géo-référencés à l'échelle des communes qu'à partir de 1995, et les séries temporelles ne comprennent donc que 18 années au total, dont « seulement » 7 années épidémiques. Sur le plan saisonnier, l'identification éventuelle des conditions climatiques favorables à l'émergence d'un foyer épidémique dans chacune des communes permettrait d'anticiper le risque en se basant sur une surveillance météorologique. En début de thèse, nous avons mené des analyses préliminaires dans ce sens. Cependant, à cette date, nous n'avions accès qu'aux données météorologiques de 24 stations du territoire, ce qui était insuffisant pour mener ces études. Ces études nécessiteraient l'obtention de séries temporelles longues à haute résolution spatiale. Pour ces études, en plus des facteurs climatiques, il faudrait tenir compte du caractère contagieux de la maldie, et prendre en compte les **phénomènes de diffusion virale sur le territoire par l'intermédiaire des mouvements humains** (Reiner et al., 2014; Stoddard et al., 2009; Teurlai et al., 2012). Les données sur les mouvements humains pourraient être récoltées au moyen de GPS (Vazquez-Prokopec et al., 2009), au moyen de données anonymes de téléphonie mobile (González et al., 2008), ou encore auprès des gendarmeries ou des sociétés de transport public en Nouvelle-Calédonie.

A l'échelle de Nouméa, du fait du caractère saisonnier des épidémies, les données ont été agrégées annuellement, et l'indicateur de risque créé est un indicateur annuel, calculé en décembre pour prédire le risque épidémique de l'année suivante. Or en 2013, la saisonnalité de l'épidémie de dengue a été altérée : cette épidémie a débuté au mois de septembre 2012. Il conviendrait donc **d'améliorer la résolution temporelle du modèle de risque d'émergence épidémique** pour obtenir un modèle capable de prédire le risque en temps réel. Une étude dans ce sens est sur le point d'être finalisée (Mangeas *et al.*, en préparation). La méthodologie employée pour la construction d'un tel modèle pourrait également être utilisée dans les pays où la saisonnalité des épidémies est moins marquée, comme en Polynésie Française.

Enfin, faute de données, le facteur manquant que nous n'avons pu étudier dans cette thèse est la **lutte anti-vectorielle**. Si des données informatisées étaient disponibles, elles pourraient être intégrées aux modèles statistiques, afin d'évaluer l'efficacité des actions de lutte.

6.3.3. Mécanismes impliqués dans la dynamique spatiale

En ce qui concerne le risque annuel d'émergence épidémique à Nouméa, il conviendrait d'**améliorer le modèle de capacité vectorielle construit**. Notamment, au vu de l'analyse de sensibilité, la mortalité des moustiques semble être un paramètre primordial du risque d'émergence. Dans le modèle construit, nous avons utilisé les paramètres de Focks (Focks et al., 1993b), à savoir un taux de mortalité constant, indépendant de l'âge des moustiques, et indépendant de la température pour des températures

comprises entre 0 °C et 40 °C. Or récemment, une revue des études de l'influence de la température sur la mortalité des moustiques a montré un lien complexe entre la température et la mortalité des *Aedes aegypti* (Brady et al., 2013). D'autres études ont également mis en évidence l'importance des variations journalières de température sur la mortalité des moustiques et sur la transmission de la dengue (Carrington et al., 2013; Lambrechts et al., 2011). Il conviendrait d'intégrer la complexité du lien entre température et mortalité dans le modèle construit.

Etant donnée l'influence potentielle des facteurs immunitaires sur la dynamique des structures spatiales de la dengue en Nouvelle-Calédonie, il serait également intéressant de **développer un modèle mécanistique** non plus basé uniquement sur la capacité vectorielle des moustiques, mais également sur les paramètres liés à l'hôte. Un modèle complet **de dynamique des populations** (Anderson and May, 1999) permettrait de modéliser non seulement le risque d'émergence épidémique, mais également la sévérité des épidémies, tout en modélisant explicitement les phénomènes immunitaires liés à l'hôte. Dans un premier temps, ce modèle pourrait être construit de manière dynamique à Nouméa uniquement, puis spatialisé pour permettre la modélisation des taux d'incidence dans chacune des communes. La construction d'un tel modèle permettrait d'intégrer les phénomènes immunitaires à l'évaluation annuelle du risque d'émergence, et permettrait d'anticiper la sévérité des épidémies. Ce modèle est d'autant plus intéressant qu'en Nouvelle-Calédonie, l'épidémie de 2013, a été la plus importante en termes de nombre de cas, et qu'elle a été causée par le virus DENV-1 qui à la suite de l'épidémie majeure de 2003, circulait à bas bruit sur le territoire depuis plus de 10 ans. Cette émergence d'une épidémie de grande ampleur alors que la population devait être en partie immunisée par les épidémies précédentes, et alors que le virus dominant dans le Pacifique était le virus DENV-4 pose la question de la vitesse à laquelle la population calédonienne s'immunise, et de la vitesse à laquelle la population d'hôtes susceptibles se renouvelle. Un modèle de dynamique des populations permettrait de répondre à ces interrogations.

6.3.4. Estimation des paramètres de transmission en Nouvelle-Calédonie

La compréhension de l'épidémiologie de la dengue ainsi que la construction de modèles mécanistiques nécessitent l'estimation précise d'un certain nombre de paramètres du cycle de transmission.

6.3.4.a. Nombre de cas introduits par an

D'après les enquêtes réalisées par la DASS-NC, il y aurait en Nouvelle-Calédonie entre 10 et 30 cas de dengue introduits chaque année sur le territoire. Cependant, ces estimations se basent sur la détection des cas cliniques par le réseau de surveillance, et sont donc probablement sous-estimées du fait de la forte proportion de formes inapparentes de dengue. Dans une étude réalisée aux Pays-Bas (Baaten et al.,

2011), un échantillon de voyageurs à destination des zones endémiques pour la dengue a été testé de manière sérologique avant le départ et à leur retour. Cette étude a permis d'estimer de manière précise la proportion de voyageurs ayant subi une séroconversion suite à une infection par le virus de la dengue, que cette infection ait été apparente ou inapparente. Une étude similaire menée en Nouvelle-Calédonie permettrait certainement d'avoir une image plus précise des cas introduits sur le territoire.

6.3.4.b. Proportion de formes asymptomatiques

Au cours des épidémies d'ampleur majeure (1995, 2003, 2009), environ 10 000 cas de dengue ont été détectés cliniquement par le réseau de surveillance. Ce chiffre ne correspond pas au nombre de personnes réellement infectées par le virus de la dengue sur le territoire, d'une part parce qu'une partie des cas cliniques ne sont pas détectés par le réseau de surveillance, d'autre part parce que de nombreuses infections sont inapparentes, et ne peuvent donc être comptabilisées par un réseau de surveillance clinique. Les différentes études à travers le monde ont montré que le ratio du nombre de formes cliniques par rapport aux formes inapparentes varie de 1:2 à 1:9 (Baaten et al., 2011; Burke et al., 1988; Dussart et al., 2012; Endy et al., 2002; Grange et al., 2014; Mammen Jr et al., 2008; Thai et al., 2011; Yap et al., 2013). En fonction de la proportion de formes inapparentes en Nouvelle-Calédonie, le nombre réel de cas pourrait donc varier de 20 000 si la proportion de formes inapparentes est de 50 %, à 90 000 si cette proportion s'élève à 90 %. La Nouvelle-Calédonie comptant 250 000 habitants au total, la proportion de formes inapparentes peut donc avoir un impact majeur sur le temps nécessaire à la construction de l'immunité de population.

L'estimation de la proportion de formes inapparentes est cependant longue et coûteuse, car elle nécessite la mise en place d'études longitudinales prospectives permettant le suivi de cohortes durant plusieurs années de suite. Il faudrait suivre ces cohortes suffisamment longtemps pour qu'une épidémie survienne au cours de l'étude, c'est-à-dire pendant 4 à 5 ans au moins.

6.3.4.c. Virémie des patients et dose infectieuse

Trois paramètres inconnus en Nouvelle-Calédonie sont la durée de la virémie des patients, les titres viraux moyens, ainsi que la dose virale infectieuse minimale nécessaire à l'infection d'un moustique au cours d'un repas sanguin. Nous avons vu, grâce à l'analyse de sensibilité du modèle de capacité vectorielle, que la durée de la virémie des patients est un paramètre ayant une influence importante sur le cycle de transmission. Des études réalisées en partenariat avec le CHT pourraient être menées afin d'évaluer les paramètres de virémie, et des études en partenariat avec l'URE-EnM de l'IPNC pourraient être menées pour évaluer les doses infectieuses nécessaires à l'infection des moustiques vecteurs.

6.3.4.d. Compétence vectorielle

La compétence vectorielle d'*Aedes aegypti* pour les différentes arboviroses circulant en Nouvelle-Calédonie n'est pas connue. Grâce à un projet financé par le Fonds Pacifique (AeDenPac), des tests de compétence vectorielle d'*Aedes aegypti* pour la dengue sont actuellement en cours à l'IPNC. Il serait intéressant d'évaluer également la compétence vectorielle d'*Aedes aegypti* pour les autres *Flavivirus* circulant en Nouvelle-Calédonie, à savoir le virus du Zika et le virus du chikungunya. En effet, la compétence vectorielle d'*Aedes aegypti* a été reconnue être plus importante pour le virus de la dengue que pour celui du chikungunya (Black et al., 2002; Lambrechts et al., 2010). Les différences de compétence vectorielle pour chacun des *Flavivirus* permettraient peut-être d'expliquer pourquoi les virus de la dengue et du Zika provoquent des épidémies majeures en Nouvelle-Calédonie (plusieurs milliers de cas lors des années épidémiques), tandis que le virus du chikungunya ne se propage que très faiblement (moins de 100 cas par an lors des années au cours desquelles la circulation virale est avérée). En effet la compétence vectorielle a été liée à des interactions génétiques entre vecteur et virus (Lambrechts, 2011). Or le virus de la dengue et celui du Zika sont très proches génétiquement, tandis que le virus du chikungunya est plus éloigné (Gaunt et al., 2001). Les différences génétiques entre les *Flavivirus* circulant en Nouvelle-Calédonie pourraient donc être à l'origine de compétences vectorielles très différentes, avec un impact majeur sur les dynamiques de transmission.

6.3.4.e. Mortalité des *Aedes aegypti*

Nous avons vu, grâce à l'analyse de sensibilité du modèle de compétence vectorielle que la mortalité est un paramètre pouvant influencer le risque d'émergence épidémique de manière quantitativement importante. Or les taux de mortalité d'*Aedes aegypti* en Nouvelle-Calédonie sont inconnus. Il serait utile d'estimer ces taux de mortalité, et notamment d'estimer la variation de ces taux de mortalité en fonction de la température. Cependant ces manipulations sont difficiles à réaliser, car il n'y a pas, sur le territoire, de laboratoire d'entomologie médicale équipé de chambres thermiques séparées permettant de réaliser les mesures de mortalité à différentes températures en même temps. A défaut des données de terrain, les équations de modélisation de la mortalité d'*Aedes aegypti* développées par Brady (Brady et al., 2013) et basées sur les données expérimentales de plus de 400 études pourraient être utilisées.

6.3.4.f. Renouvellement de la population

La Nouvelle-Calédonie, possède une particularité par rapport aux autres îles du Pacifique : de par son lien avec la France et de par la présence actuelle de trois exploitations minières employant chacune plusieurs milliers de personnes étrangères ou métropolitaines, une partie de la population n'est pas fixe, s'installant sur le territoire quelques années avant de repartir, assurant ainsi chaque année le renouvellement d'une partie de la population susceptible vis-à-vis de la dengue. Sur le plan

épidémiologique, il serait intéressant de pouvoir quantifier de manière précise ces mouvements de population.

6.3.5. Intégration des échelles

Etant donnée la diffusion régionale du virus, le risque annuel d'émergence d'une épidémie dans chacun des pays du Pacifique dépend de la situation épidémiologique régionale. Inversement, la circulation prolongée d'un même sérotype dans le Pacifique peut être expliqué par la structure spatiale des épidémies dans chacun des pays. En effet, nous avons vu par exemple qu'en Nouvelle-Calédonie, lors de la ré-émergence du virus DENV-1 en 2004 et 2008, la dengue circulait essentiellement dans des zones n'ayant pas été touchées auparavant. La circulation spatiale de la dengue au sein d'un pays est donc liée au risque de ré-émergence d'un même sérotype.

En Nouvelle-Calédonie, il conviendrait de quantifier l'impact de la situation épidémiologique régionale sur le risque annuel d'émergence épidémique, et de l'intégrer au modèle de prédiction du risque, afin d'améliorer les capacités de prédiction de ce modèle, qui est utilisé de manière opérationnelle par les autorités sanitaires.

6.3.6. Autres arboviroses

Enfin, dans cette thèse, nous nous sommes limités à l'étude de l'épidémiologie de la dengue. Cependant, ces dernières années ont été marquées par une augmentation significative de la circulation d'autres arboviroses dans le Pacifique.

En particulier, le virus Zika, identifié pour la première fois dans la forêt Zika en Ouganda dans les années 40 (Haddow et al., 1964) a causé la première épidémie avérée en 2007 à Yap (Duffy et al., 2009), une île du Pacifique faisant partie des Etats Fédérés de Micronésie. En 2013, une seconde épidémie d'une ampleur sans précédent a éclaté en Polynésie Française (Cao-Lormeau et al., 2014b), au cours de laquelle 28 000 cas cliniques ont été enregistrés (Musso et al., 2014). Le virus Zika s'est ensuite rapidement propagé dans le Pacifique avec des épidémies aux îles Cook, à l'île de Pâques, et en Nouvelle-Calédonie en 2014 (Musso et al., 2014). Ce virus étant transmis par le même moustique que la dengue, il serait extrêmement intéressant de comparer les dynamiques spatiales des deux virus. Les différences éventuelles pourraient provenir soit de différences de compétence vectorielle, soit de différences entre les phénomènes immunitaires induits. Les facteurs influençant les deux cycles de transmission de manière identique sont les facteurs environnementaux et leur influence sur le cycle de vie des moustiques, les facteurs liés au comportement humain (mode de vie, mouvements humains, et utilisation de mesures de protection personnelle), ainsi que les modalités de lutte anti-vectorielle.

Le virus du Chikungunya a causé des épidémies majeures dans les années 2005-2006 dans l'Océan Indien et a été détecté dans le Pacifique pour la première fois en 2011 en Nouvelle-Calédonie (33 cas) (Dupont-Rouzeyrol et al., 2012). Depuis il est détecté régulièrement dans le Pacifique : en Papouasie Nouvelle-Guinée en Juin 2012, en Nouvelle-Calédonie en avril 2013, à Yap en août 2013 (Roth et al., 2014), et de nouveau en Nouvelle-Calédonie en mars 2014. Le virus en cause dans le Pacifique ne présente pas la mutation qui avait rendu le vecteur *Aedes albopictus* particulièrement compétent dans l'Océan Indien (Dupont-Rouzeyrol et al., 2012), mais sa circulation accrue dans le Pacifique représente une menace rélle.

En Nouvelle-Calédonie, en 2014, pour la première fois les trois arbovirus (Dengue, Chikungunya et Zika) ont circulé en même temps, avec cependant des différences notables de transmission (seulement quelques cas pour le chikungunya contre plusieurs centaines voire des milliers de cas pour la dengue et le Zika). Les différences de circulation peuvent être dues à des différences de compétence vectorielle, *Aedes albopictus*, le vecteur principal du Chikungunya étant absent du territoire. La co-circulation des trois arbovirus pose des défis à relever en termes de surveillance clinique, et il est possible que des compétitions entre les espèces virales modifient la dynamique spatiale de chacun des virus dans le futur.

CONCLUSION

Dans ce travail de thèse, nous avons étudié les facteurs déterminant la dynamique spatiale de la dengue à trois échelles différentes : l'échelle régionale du Pacifique, l'échelle nationale de la Nouvelle-Calédonie et enfin l'échelle locale de Nouméa.

Dans le Pacifique, la dengue survient sous la forme de vagues épidémiques, avec l'émergence et la diffusion régionale d'un sérotype dominant tous les 5 à 7 ans. Les facteurs déterminant la dynamique des épidémies au sein d'un pays donné semblent propres à chacun des pays. En Nouvelle-Calédonie, des épidémies émergent tous les 5 à 7 ans, dues le plus souvent au sérotype dominant dans le Pacifique. Entre 10 et 30 cas de dengue sont introduits tous les ans sur le territoire. Cependant le développement d'une épidémie dépend fortement des conditions météorologiques locales, et un modèle prédictif du risque épidémique a pu être développé et intégré dans le système d'alerte des autorités sanitaires. Le développement d'un modèle déterministe n'a pas permis de savoir par quel mécanisme exact les facteurs météorologiques influencent l'émergence d'une épidémie. Lors des années épidémiques, les communes de Nouvelle-Calédonie sont touchées de manière hétérogène, sous l'influence de la température moyenne, des densités locales de population ainsi que de facteurs socio-économiques liés aux modes de vie des différentes populations. Des campagnes de sensibilisation axées sur la responsabilité individuelle pourraient peut-être permettre de diminuer l'incidence de la dengue dans certaines communes. A Nouméa, le fait d'habiter en maison ainsi que le fait d'être né en Nouvelle-Calédonie sont des facteurs de risque, mais ils n'expliquent pas entièrement l'hétérogénéité de la circulation virale entre les différents quartiers. Il semble que les mesures de protection individuelle soient un élément important de la réduction du risque épidémique. Aucun lien statistique n'a pu être mis en évidence entre les facteurs environnementaux influant le cycle de vie du moustique et les taux d'incidence. Si ces liens existent, leur mise en évidence nécessite des études à une résolution spatiale extrêmement fine, et l'utilisation d'enquêtes pourrait s'avérer bénéfique. Que ce soit à l'échelle du territoire ou à l'échelle plus locale de Nouméa, l'immunité de population semble jouer un rôle déterminant dans la variation des schémas spatiaux d'incidence d'une année à l'autre.

Les pistes de recherche future sont :

- La continuation du travail de modélisation statistique, pour effectuer les études non réalisées dans cette thèse ;
- La construction d'un modèle déterministe de type SIR pour tester les différentes hypothèses émises au cours de cette thèse concernant les mécanismes pouvant expliquer certains schémas observés, et notamment le rôle de l'immunité de population ;
- Des études virologiques afin de caractériser les agents pathogènes responsables d'évènements se démarquant des autres sur le plan épidémiologique (*i.e.* co-circulation de différents sérotypes lorsque

le sérotype 3 domine dans le Pacifique ou encore disparition du sérotype 4 de la Nouvelle-Calédonie au profit du sérotype 1 en 2010).

Nous avons vu dans la revue bibliographique qu'il existait une multitude d'études parfois contradictoires sur l'épidémiologie de la dengue. Nous pensons que le classement des différentes études en fonction de la cible de modélisation (émergence d'une épidémie ou sévérité de l'épidémie) et de l'échelle étudiée (spatiale ou temporelle, et en précisant s'il s'agit d'études spatiales régionales, nationales ou locales, ou d'études temporelles inter-annuelles ou intra-annuelles) permettrait de clarifier nos connaissances à ce sujet. Nous espérons que ce travail de thèse, en apportant des résultats approfondis dans une région du monde peu étudiée jusqu'à présent permettra aux chercheurs et à l'ensemble des personnes impliquées dans la lutte contre la dengue de construire une image plus précise de l'épidémiologie de cette maladie.

Outre les résultats obtenus aux trois échelles, cette thèse a permis de développer une approche générique de l'analyse des données spatio-temporelles de surveillance en épidémiologie, approche pouvant être utilisée dans d'autres contexte épidémiologiques, et consistant en deux étapes : i) Etude approfondie de la variabilité spatio-temporelle des données épidémiologiques et ii) Etude des facteurs de risque, en commençant par des études uniquement spatiales ou temporelles, réalisées sur des moyennes, et en finissant éventuellement par l'étude des évènements inhabituels. En particulier dans le cadre de l'étude des schémas spatio-temporels récurrents, nous espérons que le transfert de la technique d'analyse EOF dans le domaine de l'épidémiologie pourra être utile à l'étude d'autres maladies. Les méthodes utilisées ici pourraient être utilisées dans d'autres contextes géographiques ou épidémiologiques.

Aaskov, J., Buzacott, K., Field, E., Lowry, K., Berlioz-Arthaud, A., and Holmes, E.C. (2007). Multiple recombinant dengue type 1 viruses in an isolate from a dengue patient. J. Gen. Virol. *88*, 3334–3340.

Adams, B., and Kapan, D.D. (2009). Man Bites Mosquito: Understanding the Contribution of Human Movement to Vector-Borne Disease Dynamics. PLoS ONE *4*, e6763.

Adams, B., Holmes, E.C., Zhang, C., Mammen, M.P., Nimmannitya, S., Kalayanarooj, S., and Boots, M. (2006). Cross-protective immunity can account for the alternating epidemic pattern of dengue virus serotypes circulating in Bangkok. Proc. Natl. Acad. Sci. *103*, 14234–14239.

Ahmad, O.B., Boschi-Pinto, C., Lopez, A.D., Murray, C.J.L., Lozano, R., and Inoue, M. (2001). Age standardisation of rates: a new WHO standard (World Health Organization).

Aldstadt, J., Yoon, I.-K., Tannitisupawong, D., Jarman, R.G., Thomas, S.J., Gibbons, R.V., Uppapong, A., Iamsirithaworn, S., Rothman, A.L., Scott, T.W., et al. (2012). Space-time analysis of hospitalised dengue patients in rural Thailand reveals important temporal intervals in the pattern of dengue virus transmission: Space-time dengue transmission intervals. Trop. Med. Int. Health *17*, 1076–1085.

Anderson, J.R., and Rico-Hesse, R. (2006). Aedes Aegypti Vectorial Capacity Is Determined by the Infecting Genotype of Dengue Virus. Am. J. Trop. Med. Hyg. *75*, 886–892.

Anderson, R.M., and May, R.M. (1999). Infectious diseases of humans: dynamics and control (Bristol, Avon: Oxford University Press).

Anonymous Program for Climate Model Diagnosis and Intercomparison.

A-Nuegoonpipat, A., Berlioz-Arthaud, A., Chow, V., Endy, T., Lowry, K., Mai, L.Q., Ninh, T.U., Pyke, A., Reid, M., Reynes, J.-M., et al. (2004). Sustained transmission of dengue virus type 1 in the Pacific due to repeated introductions of different Asian strains. Virology *329*, 505–512.

Arias-Goeta, C., Mousson, L., Rougeon, F., and Failloux, A.-B. (2013). Dissemination and Transmission of the E1-226V Variant of Chikungunya Virus in Aedes albopictus Are Controlled at the Midgut Barrier Level. PLoS ONE *8*, e57548.

Arrivillaga, J., and Barrera, R. (2004). Food as a limiting factor for Aedes aegypti in water-storage containers. J. Vector Ecol. *29*, 11–20.

Aziz, S., Ngui, R., Lim, Y. a. L., Sholehah, I., Nur Farhana, J., Azizan, A.S., and Wan Yusoff, W.S. (2012). Spatial pattern of 2009 dengue distribution in Kuala Lumpur using GIS application. Trop. Biomed. *29*, 113–120.

Baaten, G.G.G., Sonder, G.J.B., Zaaijer, H.L., van Gool, T., Kint, J.A.P.C.M., and van den Hoek, A. (2011). Travel-related Dengue Virus Infection, the Netherlands, 2006–2007. Emerg. Infect. Dis. *17*, 821–828.

Banu, S., Hu, W., Hurst, C., and Tong, S. (2011). Dengue transmission in the Asia-Pacific region: impact of climate change and socio-environmental factors: Dengue transmission in the Asia-Pacific region. Trop. Med. Int. Health *16*, 598–607.

Banu, S., Hu, W., Guo, Y., Hurst, C., and Tong, S. (2014). Projecting the impact of climate change on dengue transmission in Dhaka, Bangladesh. Environ. Int. *63*, 137–142.

Baqar, S., Hayes, C.G., and Ahmed, T. (1980). The effect of larval rearing conditions and adult age on the susceptibility of Culex tritaeniorhynchus to infection with West Nile virus. Mosq News *40(2)*, 165–171.

Barbazan, P., Tuntaprasart, W., Souris, M., Demoraes, F., Nitatpattana, N., Boonyuan, W., and Gonzalez, J.-P. (2008). Assessment of a new strategy, based on *Aedes aegypti* (L.) pupal productivity, for the surveillance and control of dengue transmission in Thailand. Ann. Trop. Med. Parasitol. *102*, 161–171.

Barmak, D.H., Dorso, C.O., Otero, M., and Solari, H.G. (2011). Dengue epidemics and human mobility. ArXiv11023869 Phys. Q-Bio.

Barthel, A., Gourinat, A.-C., Cazorla, C., Joubert, C., Dupont-Rouzeyrol, M., and Descloux, E. (2013). Breast Milk as a Possible Route of Vertical Transmission of Dengue Virus? Clin. Infect. Dis. *57*, 415–417.

Bar-Zeev, M. (1958). The Effect of Temperature on the Growth Rate and Survival of the Immature Stages of Aëdes aegypti (L.). Bull. Entomol. Res. *49*, 157–163.

Beckett, C.G., Kosasih, H., Faisal, I., Nurhayati, Tan, R., Widjaja, S., Listiyaningsih, E., Ma'roeff, C., Wuryadi, S., Bangs, M.J., et al. (2005). Early detection of dengue infections using cluster sampling around index cases. Am. J. Trop. Med. Hyg. *72*, 777–782.

Belkin, J.N. (1962). The mosquitoes of the South Pacific (Diptera, Culicidae) (Berkeley and Los Angeles: University of California Press).

Bendali, F. (2006). La conception et la mise en oeuvre de programmes d'épidémiosurveillance efficaces dans les pays d'Afrique subsaharienne. Rev. Sci. Tech. Int. Off. Epizoot. *25*, 199–209.

Benedict, M.Q., Levine, R.S., Hawley, W.A., and Lounbios, L.P. (2007). Spread of the Tiger: Global Risk of Invasion by the Mosquito Aedes albopictus. Vector Borne Zoonotic Dis. Larchmt. N *7*, 76–85.

Bénichou, P., and Le Breton, O. (1987). Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques (Use of topography on mapping of statistical rainfall fields). La Météorologie *7*, 23–34.

Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., et al. (2013). The global distribution and burden of dengue. Nature *496*, 504–507.

Bivand, R.S., Pebesma, E., and Gómez-Rubio, V. (2013). Applied Spatial Data Analysis with R (Springer New York).

Bjørnstad, O.N., and Falck, W. (2001). Nonparametric spatial covariance functions: Estimation and testing. Environ. Ecol. Stat. *8*, 53–70.

Bjørnstad, O.N., Ims, R.A., and Lambin, X. (1999). Spatial population dynamics: analyzing patterns and processes of population synchrony. Trends Ecol. Evol. *14*, 427–432.

Black, W.C., Bennett, K.E., Gorrochotegui-Escalante, N., Barillas-Mury, C.V., Fernandez-Salas, I., de Lourdes Muñoz, M., Farfan-Ale, J.A., Olson, K.E., and Beaty, B.J. (2002). Flavivirus susceptibility in Aedes aegypti. Arch. Med. Res. *33*, 379–388.

Bohra, A., and Andrianasolo, H. (2001). Application of GIS in Modeling Dengue Risk Based on Sociocultural Data: Case of Jalore, Rajasthan, India.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., and White, J.-S.S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. *24*, 127–135.

Bonvallot, J., Gay, J.-C., and Habert, E. (2013). Atlas de la Nouvelle-Calédonie (Marseille: IRD Orstom).

Bosio, C.F., Beaty, B.J., and Black, W.C. (1998). Quantitative genetics of vector competence for dengue-2 virus in Aedes aegypti. Am. J. Trop. Med. Hyg. *59*, 965–970.

Bosio, C.F., Fulton, R.E., Salasek, M.L., Beaty, B.J., and Black, W.C. (2000). Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito Aedes aegypti. Genetics *156*, 687–698.

Bouldouyre, M.A., Baumann, F., Berlioz-Arthaud, A., Chungue, E., and Lacassin, F. (2006). Factors of severity at admission during an epidemic of dengue 1 in New Caledonia (South Pacific) in 2003. Scand. J. Infect. Dis. *38*, 675–681.

Brady, O.J., Johansson, M.A., Guerra, C.A., Bhatt, S., Golding, N., Pigott, D.M., Delatte, H., Grech, M.G., Leisnham, P.T., Maciel-de-Freitas, R., et al. (2013). Modelling adult Aedes aegypti and Aedes albopictus survival at different temperatures in laboratory and field settings. Parasit. Vectors *6*, 351.

Brandt, W.E., McCown, J.M., Gentry, M.K., and Russell, P.K. (1982). Infection enhancement of dengue type 2 virus in the U-937 human monocyte cell line by antibodies to flavivirus cross-reactive determinants. Infect. Immun. *36*, 1036–1041.

Bretherton, C.S., Smith, C., and Wallace, J.M. (1992). An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. J. Clim. *5*, 541–560.

Britton, T. (2001). Epidemics in heterogeneous communities: estimation of R0 and secure vaccination coverage. J. R. Stat. Soc. Ser. B Stat. Methodol. *63*, 705–715.

Brown, J.L., Wilkinson, R., Davidson, R.N., Wall, R., Lloyd, G., Howells, J., and Pasvol, G. (1996). Rapis diagnosis and determination of duration of viraemia in dengue fever using a reverse transcriptase polymerase chain reaction. Trans R Soc Trop Med Hyg *90*, 140–143.

Burke, D.S., Nisalak, A., Johnson, D.E., and Scott, R.M. (1988). A prospective study of dengue infections in Bangkok. Am. J. Trop. Med. Hyg. *38*, 172–180.

Cáceres, C., Yung, V., Araya, P., Tognarelli, J., Villagra, E., Vera, L., and Fernández, J. (2008). Complete nucleotide sequence analysis of a Dengue-1 virus isolated on Easter Island, Chile. Arch. Virol. *153*, 1967–1970.

Canyon, D.V., Hii, J.L.K., and MŸller, R. (1999). Adaptation of Aedes aegypti (Diptera: Culicidae) oviposition behavior in response to humidity and diet. J. Insect Physiol. *45*, 959–964.

Cao-Lormeau, V.-M., Roche, C., Aubry, M., Teissier, A., Lastere, S., Daudens, E., Mallet, H.-P., Musso, D., and Aaskov, J. (2011). Recent Emergence of Dengue Virus Serotype 4 in French Polynesia Results from Multiple Introductions from Other South Pacific Islands. PLoS ONE *6*, e29555.

Cao-Lormeau, V.-M., Roche, C., Musso, D., Mallet, H.-P., Dalipanda, T., Dofai, A., Nogareda, F., Nilles, E.J., and Aaskov, J. (2014a). Dengue Virus Type 3, South Pacific Islands, 2013. Emerg. Infect. Dis. *20*, 1034–1036.

Cao-Lormeau, V.-M., Roche, C., Teissier, A., Robin, E., Berry, A.-L., Mallet, H.-P., Sall, A.A., and Musso, D. (2014b). Zika Virus, French Polynesia, South Pacific, 2013. Emerg. Infect. Dis. *20*, 1084–1086.

Capeding, M.R., Tran, N.H., Hadinegoro, S.R.S., Ismail, H.I.H.M., Chotpitayasunondh, T., Chua, M.N., Luong, C.Q., Rusmil, K., Wirawan, D.N., Nallusamy, R., et al. (2014). Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial. The Lancet.

Carbajo, A.E., Cardo, M.V., Vezzani, D., and others (2012). Is temperature the main cause of dengue rise in non-endemic countries? The case of Argentina. Int J Health Geogr *11*, 26.

Carrington, L.B., Seifert, S.N., Willits, N.H., Lambrechts, L., and Scott, T.W. (2013). Large diurnal temperature fluctuations negatively influence Aedes aegypti (Diptera: Culicidae) life-history traits. J. Med. Entomol. *50*, 43–51.

Cauchemez, S., Boelle, P.-Y., Donnelly, C.A., Ferguson, N.M., Thomas, G., Leung, G.M., Hedley, A.J., Anderson, R.M., and Valleron, A.-J. (2006). Real-time Estimates in Early Detection of SARS. Emerg. Infect. Dis. *12*, 110–113.

Cavarero, V., Peltier, A., Aubail, X., Leroy, A., Dubuisson, B., Jourdain, S., Ganachaud, A., Gibelin, A.-L., Lefèvre, J., Menkes, C., et al. (2012). Les évolutions passées et futures du climat de la Nouvelle-Calédonie. La Météorologie *77*, 13–21.

Cazelles, B., and Hales, S. (2006). Infectious diseases, climate influences, and nonstationarity. PLoS Med. *3*, e328.

Cazelles, B., Chavez, M., McMichael, A.J., and Hales, S. (2005). Nonstationary influence of El Ni–o on the synchronous dengue epidemics in Thailand. PLOS Med. *2*, 313–318.

Cazelles, B., Chavez, M., Constantin de Magny, G., Guegan, J.-F., and Hales, S. (2007). Time-dependent spectral analysis of epidemiological time-series with wavelets. J. R. Soc. Interface *4*, 625–636.

Chowell, G., Torre, C.A., Munayco-Escate, C., SuáRez-Ognio, L., LóPez-Cruz, R., Hyman, J.M., and Castillo-Chavez, C. (2008). Spatial and temporal dynamics of dengue fever in Peru: 1994–2006. Epidemiol. Infect. *136*, 1667.

Chungue, E., Deubel, V., Cassar, O., Laille, M., and Martin, P.M.V. (1993). Molecular epidemiology of dengue 3 viruses and genetic relatedness among dengue 3 strains isolated from patients with mild or severe form of dengue fever in French Polynesia. J Gen Virol *74*, 2765–2770.

Cluzel, D., Aitchison, J.C., and Picard, C. (2001). Tectonic accretion and underplating of mafic terranes in the Late Eocene intraoceanic fore-arc of New Caledonia (Southwest Pacific): geodynamic implications. Tectonophysics *340*, 23–59.

Collot, J., Geli, L., Lafoy, Y., Vially, R., Cluzel, D., Klingelhoefer, F., and Nouzé, H. (2008). Tectonic history of northern New Caledonia Basin from deep offshore seismic reflection: Relation to late Eocene obduction in New Caledonia, southwest Pacific. Tectonics *27*, TC6006.

Colwell, R., Epstein, P., Gubler, D., Hall, M., Reiter, P., Shukla, J., Sprigg, W., Takafuji, E., and Trtanj, J. (1998). Global climate change and infectious diseases. Emerg. Infect. Dis. *4*, 451–452.

Cortes, C., and Vapnik, V. (1995). Support-vector networks. Mach. Learn. *20*, 273–297.

Costa, J.V., Donalisio, M.R., and Silveira, L.V. de A. (2013). Spatial distribution of dengue incidence and socio-environmental conditions in Campinas, São Paulo State, Brazil, 2007. Cad. Saúde Pública *29*, 1522–1532.

Costagliola, D., Flahault, A., Galinec, D., Garnerin, P., Menares, J., and Valleron, A.-J. (1991). A routine tool for detection and assessment of epidemics of influenza-like syndromes in France. Am. J. Public Health *81*, 97–99.

Couret, J., Dotson, E., and Benedict, M.Q. (2014). Temperature, Larval Diet, and Density Effects on Development Rate and Survival of Aedes aegypti (Diptera: Culicidae). PLoS ONE *9*, e87468.

Coutinhoa, F.A.B., Burattinia, M.N., Lopeza, L.F., and Massada, E. (2006). Threshold Conditions for a Non-Autonomous Epidemic System Describing the Population Dynamics of Dengue. Bull. Math. Biol. *68*, 2263–2282.

Cox, J., Brown, H.E., and Rico-Hesse, R. (2011). Variation in Vector Competence for Dengue Viruses Does Not Depend on Mosquito Midgut Binding Affinity. PLoS Negl Trop Dis *5*, e1172.

Cullen, J., Chitprarop, U., Doberstyn, E., and Sombatwattanangkul, K. (1984). Epidemiological early warning system for malaria control in northern Thailand. Bull. World Health Organ. *62*, 107–114.

Cummings, D.A.T., Irizarry, R.A., Huang, N.E., Endy, T.P., Nisalak, A., Ungchusak, K., and Burke, D.S. (2004). Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature *427*, 344–347.

Cummings, D.A.T., Schwartz, I.B., Billings, L., Shaw, L.B., and Burke, D.S. (2005). Dynamic effects of antibody-dependent enhancement on the fitness of viruses. Proc. Natl. Acad. Sci. U. S. A. *102*, 15259–15264.

Cuong, H.Q., Hien, N.T., Duong, T.N., Phong, T.V., Cam, N.N., Farrar, J., Nam, V.S., Thai, K.T.D., and Horby, P. (2011). Quantifying the Emergence of Dengue in Hanoi, Vietnam: 1998–2009. PLoS Negl. Trop. Dis. *5*, e1322.

Cuong, H.Q., Vu, N.T., Cazelles, B., Boni, M.F., Thai, K.T.D., Rabaa, M.A., Quang, L.C., Simmons, C.P., Huu, T.N., and Anders, K.L. (2013). Spatiotemporal Dynamics of Dengue Epidemics, Southern Vietnam. Emerg. Infect. Dis. *19*, 945–953.

DASS-NC (2010). Situation sanitaire en Nouvelle-Calédonie. Mémento 2009 (Direction des affaires sanitaires et sociales (DASS)).

DASS-NC (2011). Situation sanitaire en Nouvelle-Calédonie. 2011 (Direction des affaires sanitaires et sociales (DASS)).

Dégallier, N. (2011). Capacidade vetorial : importancia na modelagem da transmissao de dengue. In Entomologia medica e veterinaria : segunda ediçao, C.B. Marcondes, ed. (Sao Paulo: Atheneu), pp. 37–44.

Degallier, N., Vilarinhos, P. d. T.R., Carvalho, M. d. S.L. d., Knox, M.B., and Caetano, J. (2000). People's knowledge and practice about dengue, its vectors and control means in Brasilia (DF), Brazil: its relevance with entomological factors. J Amer Mosq Cont Assoc *16 (2)*, 114–123.

Degallier, N., Favier, C., Boulanger, J.-P., Menkes, C.E., and Oliveira, C. (2005). Une nouvelle méthode d'estimation du taux de reproduction des maladies (R 0): application à l'étude des épidémies de dengue dans le District Fédéral (Brésil). Environ. Risques Santé *4*, 131–135.

Degallier, N., Favier, C., Menkes, C.E., Lengaigne, M., Ramalho, W.M., Souza, R., Servain, J., and Boulanger, J.-P. (2010). Toward an early warning system for dengue prevention: modeling climate impact on dengue transmission. Clim. Change *98*, 581–592.

Degallier, N., Servain, J., Lucio, P.S., Hannart, A., Durand, B., de Souza, R.N., and Ribeiro, Z.M. (2012). The influence of local environment on the aging and mortality of Aedes aegypti (L.): case study in Fortaleza-CE, Brazil. J. Vector Ecol. J. Soc. Vector Ecol. *37*, 428–441.

Delatte, H., Dehecq, J.S., Thiria, J., Domerg, C., Paupy, C., and Fontenille, D. (2008). Geographic distribution and Developmental Sites of Aedes albopictus (Diptera : Culicidae) During a chinkungunya Epidemic Event. Vector-Borne Zoonotic Dis. *8*, 25–34.

Descloux, E. (2011). Dynamique moléculaire et épidémique du virus de la dengue dans différents écosystèmes (Aix Marseille 2).

Descloux, E., Mangeas, M., Menkes, C.E., Lengaigne, M., Leroy, A., Tehei, T., Guillaumot, L., Teurlai, M., Gourinat, A.-C., Benzler, J., et al. (2012). Climate-Based Models for Understanding and Forecasting Dengue Epidemics. PLoS Negl Trop Dis *6*, e1470.

Dibo, M.R., Chierotti, A.P., Ferrari, M.S., Mendonça, A.L., and Chiaravalloti Neto, F. (2008). Study of the relationship between Aedes (Stegomyia) aegypti egg and adult densities, dengue fever and climate in Mirassol, state of São Paulo, Brazil. Mem. Inst. Oswaldo Cruz *103*, 554–560.

Dietz, K. (1993). The estimation of the basic reproduction number for infectious diseases. Stat. Methods Med. Res. *2*, 23–41.

Dormann, C.F., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., et al. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography *30*, 609–628.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography *36*, 27–46.

Duffy, M.R., Chen, T.-H., Hancock, W.T., Powers, A.M., Kool, J.L., Lanciotti, R.S., Pretrick, M., Marfel, M., Holzbauer, S., Dubray, C., et al. (2009). Zika Virus Outbreak on Yap Island, Federated States of Micronesia. N. Engl. J. Med. *360*, 2536–2543.

Duong, V., Ly, S., Lorn Try, P., Tuiskunen, A., Ong, S., Chroeung, N., Lundkvist, A., Leparc-Goffart, I., Deubel, V., Vong, S., et al. (2011). Clinical and Virological Factors Influencing the Performance of a NS1 Antigen-Capture Assay and Potential Use as a Marker of Dengue Disease Severity. PLoS Negl Trop Dis *5*, e1244.

Duong, V., Henn, M.R., Simmons, C., Ngan, C., Y, B., Gavotte, L., Viari, A., Ong, S., Huy, R., Lennon, N.J., et al. (2013a). Complex dynamic of dengue virus serotypes 2 and 3 in Cambodia following series of climate disasters. Infect. Genet. Evol. *15*, 77–86.

Duong, V., Simmons, C., Gavotte, L., Viari, A., Ong, S., Chantha, N., Lennon, N.J., Birren, B.W., Vong, S., Farrar, J.J., et al. (2013b). Genetic diversity and lineage dynamic of dengue virus serotype 1 (DENV-1) in Cambodia. Infect. Genet. Evol. *15*, 59–68.

Dupont-Rouzeyrol, M., Caro, V., Guillaumot, L., Vazeille, M., D'Ortenzio, E., Thiberge, J.-M., Baroux, N., Gourinat, A.-C., Grandadam, M., and Failloux, A.-B. (2012). Chikungunya virus and the mosquito vector Aedes aegypti in New Caledonia (South Pacific Region). Vector Borne Zoonotic Dis. Larchmt. N *12*, 1036–1041.

Dupont-Rouzeyrol, M., Aubry, M., O'Connor, O., Roche, C., Gourinat, A.-C., Guigon, A., Pyke, A., Grangeon, J.-P., Nilles, E., Chanteau, S., et al. (2014). Epidemiological and molecular features of dengue virus type-1 in New Caledonia, South Pacific, 2001–2013. Virol. J. *11*, 61.

Dussart, P., Baril, L., Petit, L., Beniguel, L., Quang, L.C., Ly, S., Azevedo, R. do S.S., Meynard, J.-B., Vong, S., Chartier, L., et al. (2012). Clinical and Virological Study of Dengue Cases and the Members of Their Households: The Multinational DENFRAME Project. PLoS Negl Trop Dis *6*, e1482.

Effler, P.V., Pang, L., Kitsutani, P., Vorndam, V., Nakata, M., Ayers, T., Elm, J., Tom, T., Reiter, P., Rigau-Perez, J.G., et al. (2005). Dengue fever, Hawaii, 2001–2002. Emerg Infect Dis *11*, 742–749.

Elliott, P., and Wartenberg, D. (2004). Spatial Epidemiology: Current Approaches and Future Challenges. Environ. Health Perspect. *112*, 998–1006.

Elliott, P., Wakefield, J., Best, N., Briggs, D., and (Sous la direction de) (2001). Spatial Epidemiology: Methods and Applications (Oxford; New York: Oxford University Press).

Elston, D.A., Moss, R., Boulinier, T., Arrowsmith, C., and Lambin, X. (2001). Analysis of aggregation, a worked example: numbers of ticks on red grouse chicks. Parasitology *122*, 563–569.

Endy, T.P., Chunsuttiwat, S., Nisalak, A., Libraty, D.H., Green, S., Rothman, A.L., Vaughn, D.W., and Ennis, F.A. (2002). Epidemiology of Inapparent and Symptomatic Acute Dengue Virus Infection: A Prospective Study of Primary School Children in Kamphaeng Phet, Thailand. Am. J. Epidemiol. *156*, 40–51.

Endy, T.P., Anderson, K.B., Nisalak, A., Yoon, I.-K., Green, S., Rothman, A.L., Thomas, S.J., Jarman, R.G., Libraty, D.H., and Gibbons, R.V. (2011). Determinants of Inapparent and Symptomatic Dengue Infection in a Prospective Study of Primary School Children in Kamphaeng Phet, Thailand. PLoS Negl Trop Dis *5*, e975.

Escoffier, B., and Pagès, J. (2008). Analyses factorielles simples et multiples. Objectifs, méthodes et interprétation (Paris: Dunod).

Espinoza Gómez, F., Hernández Suárez, C.M., and Coll Cárdenas, R. (2001). Factores que modifican los índices larvarios de Aedes aegypti en Colima, México. Rev. Panam. Salud Pública *10*, 6–12.

Estallo, E.L., Más, G., Vergara-Cid, C., Lanfri, M.A., Ludueña-Almeida, F., Scavuzzo, C.M., Introini, M.V., Zaidenberg, M., and Almirón, W.R. (2013). Spatial Patterns of High Aedes aegypti Oviposition Activity in Northwestern Argentina. PLoS ONE *8*, e54167.

Eubank, S., Guclu, H., Anil Kumar, V.S., Marathe, M.V., Srinivasan, A., Toroczkai, Z., and Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. Nature *429*, 180–184.

Fansiri, T., Fontaine, A., Diancourt, L., Caro, V., Thaisomboonsuk, B., Richardson, J.H., Jarman, R.G., Ponlawat, A., and Lambrechts, L. (2013). Genetic Mapping of Specific Interactions between Aedes aegypti Mosquitoes and Dengue Viruses. PLoS Genet *9*, e1003621.

Farrington, C.P., Andrews, N.J., Beale, A.D., and Catchpole, M.A. (1996). A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease. J. R. Stat. Soc. Ser. A Stat. Soc. *159*, 547.

Farrington, C.P., Kanaan, M.N., and Gay, N.J. (2001). Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. J. R. Stat. Soc. Ser. C Appl. Stat. *50*, 251–292.

Fauran, P. (1996). Nouveaux aspects ŽpidŽmiologiques de la dengue. Bull. SociŽtŽ Pathol. Exot. *89*, 163–165.

Favier, C., Degallier, N., Dubois, M.A., Boulanger, J.-P., Menkes, C.E., and Torres, L. (2005). Dengue epidemic modeling: stakes and pitfalls. Asia Pac. Biotech News *9*, 1191–1194.

Favier, C., Degallier, N., Vilarinhos, P. d. T.R., Carvalho, M. d. S.L. d., Yoshizawa, M.A.C., and Knox, M.B. (2006a). Effects of climate and different management strategies on Aedes aegypti breeding sites: a longitudinal survey in Brasilia (DF, Brazil). Trop. Med. Int. Health *11*, 1104–1118.

Favier, C., Degallier, N., Lima, J.R.C., Rosa-Freitas, M.G., Luitgards-Moura, J.F., Mondet, B., Oliveira, C., Weimann, E.T. d. S., and Tsouris, P. (2006b). Early determination of the reproductive number for vector-borne diseases: the case of Dengue in Brazil. Trop. Med. Int. Health *11*, 332–340.

Ferguson, N., Anderson, R., and Gupta, S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. Proc. Natl. Acad. Sci. *96*, 790–794.

Ferrari, M.J., Bjørnstad, O.N., and Dobson, A.P. (2005). Estimation and inference of R0 of an infectious pathogen by a removal method. Math. Biosci. *198*, 14–26.

Fine, P.E.M. (2003). The interval between successive cases of an infectious disease. Am. J. Epidemiol. *158*, 1039–1047.

Fischer, S., Alem, I.S., Majo, M.S.D., Campos, R.E., and Schweigmann, N. (2011). Cold Season Mortality and Hatching Behavior of Aedes aegypti L. (Diptera: Culicidae) Eggs in Buenos Aires City, Argentina. J. Vector Ecol. *36*, 94–99.

Focks, D.A., Haile, D.G., Daniels, E., and Mount, G.A. (1993a). Dynamic life table model for Aedes aegypti (Diptera: Culicidae): analysis of the literature and model development. J. Med. Entomol. *30(6)*, 1003–1017.

Focks, D.A., Haile, D.G., Daniels, E., and Mount, G.A. (1993b). Dynamic life table model for Aedes aegypti (Diptera: Culicidae): simulation results and validation. J. Med. Entomol. *30(6)*, 1018–1028.

Focks, D.A., Daniels, E., Haile, D.G., and Keesling, J.E. (1995). A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. Am. J. Trop. Med. Hyg. *53 (5)*, 489–506.

Focks, D.A., Brenner, R.J., Hayes, J., and Daniels, E. (2000). Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. Am J Trop Med Hyg *62*, 11–18.

Fouque, F., Carinci, R., Gaborit, P., Issaly, J., Bicout, D.J., and Sabatier, P. (2006). Aedes aegypti survival and dengue transmission patterns in French Guiana. J. Vector Ecol. *31*, 390–399.

Gagnon, A.S., Bush, A.B., and Smoyer-Tomic, K.E. (2001). Dengue epidemics and the El Niño southern oscillation. Clim. Res. *19*, 35–43.

Garrett-Jones, C. (1964). Prognosis for Interruption of Malaria Transmission Through Assessment of the Mosquito's Vectorial Capacity. Nature *204*, 1173–1175.

Gatrell, A.C., Bailey, T.C., Diggle, P.J., and Rowlingson, B.S. (1996). Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. Trans. Inst. Br. Geogr. *21*, 256–274.

Gaunt, M.W., Sall, A.A., de Lamballerie, X., Falconar, A.K., Dzhivanian, T.I., and Gould, E.A. (2001). Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J. Gen. Virol. *82*, 1867–1876.

Getis, A., Morrison, A.C., Gray, K., and Scott, T.W. (2003). Characteristics of the spatial pattern of the dengue vector, Aedes aegypti, in Iquitos, Peru. Am. J. Trop. Med. Hyg. *69*, 494–505.

Gharbi, M., Quenel, P., Gustave, J., Cassadou, S., Ruche, G.L., Girdary, L., and Marrama, L. (2011). Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. BMC Infect. Dis. *11*, 166.

Girod, R., Gaborit, P., Marrama, L., Etienne, M., Ramdini, C., Rakotoarivony, I., Dollin, C., Carinci, R., Issaly, J., Dusfour, I., et al. (2011). Viewpoint: High susceptibility to Chikungunya virus of Aedes aegypti from the French West Indies and French Guiana: Chikungunya virus in French West Indies. Trop. Med. Int. Health *16*, 134–139.

Githeko, A.K., Lindsay, S.W., Confalonieri, U.E., and Patz, J.A. (2000). Climate change and vector-borne diseases: a regional analysis. Bull. World Health Organ. *78*, 1136–1147.

González, M.C., Hidalgo, C.A., and Barabási, A.-L. (2008). Understanding individual human mobility patterns. Nature *453*, 779–782.

Grange, L., Simon-Loriere, E., Sakuntabhai, A., Gresh, L., Paul, R., and Harris, E. (2014). Epidemiological risk factors associated with high global frequency of inapparent dengue virus infections. Microb. Immunol. *5*, 280.

Grenfell, B.T., Bjørnstad, O.N., and Kappey, J. (2001). Travelling waves and spatial hierarchies in measles epidemics. Nature *414*, 716–723.

Grimstad, P.R., and Haramis, L.D. (1984). Aedes triseriatus (Diptera: Culicidae) and La Crosse virus. III. Enhanced oral transmission by nutritionÐdeprived mosquitoes. J Med Ent *21(3)*, 249–256.

Grimstad, P.R., and Walker, E.D. (1991). Aedes triseriatus (Diptera: Culicidae) and La Crosse virus. IV. Nutritional deprivation of larvae affects the adult barriers to infection and transmission. J. Med. Entomol. *28(3)*, 378–386.

Gubler, D.J. (1998). Dengue and dengue hemorrhagic fever. Clin. Microbiol. Rev. *11*, 480–496.

Gubler, D.J. (2004). The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? Comp. Immunol. Microbiol. Infect. Dis. *27*, 319–330.

Gubler, D.J., Suharyono, W., Tan, R., Abidin, M., and Sie, A. (1981). Viraemia in patients with naturally acquired dengue infection. Bull. World Health Organ. *59*, 623–630.

Gubler, D.J., Reiter, P., Ebi, K.L., Yap, W., Nasci, R., and Patz, J.A. (2001). Climate variability and change in the United States: potential impacts on vector- and rodent-borne diseases. Environ. Health Perspect. *109 (supplement 2)*, 223–233.

Guillaumot, L. (2005). Arboviruses and their vectors in the Pacific--status report. Pac. Health Dialog *12*, 45–52.

Guzman, M.G., and Kouri, G. (2003). Dengue and dengue hemorrhagic fever in the Americas: lessons and challenges. J. Clin. Virol. *27*, 1–13.

Guzmán, M.G., Kourí, G., Valdés, L., Bravo, J., Vázquez, S., and Halstead, S.B. (2002). Enhanced severity of secondary dengue-2 infections: death rates in 1981 and 1997 Cuban outbreaks. Rev. Panam. Salud Pública Pan Am. J. Public Health *11*, 223–227.

Guzman, M.G., Halstead, S.B., Artsob, H., Buchy, P., Farrar, J., Gubler, D.J., Hunsperger, E., Kroeger, A., Margolis, H.S., Martínez, E., et al. (2010). Dengue: a continuing global threat. Nat. Rev. Microbiol. *8*, S7–S16.

Haddow, A.J., Williams, M.C., Woodall, J.P., Simpson, D.I.H., and Goma, L.K.H. (1964). Twelve isolations of Zika virus from Aedes (Stegomyia) africanus (Theobald) taken in and above a Uganda forest. Bull. World Health Organ. *31*, 57–69.

Hales, S., Weinstein, P., and Woodward, A. (1996). Dengue fever epidemics in the South Pacific: driven by El Nino Southern Oscillation? The Lancet *348*, 1664–1665.

Hales, S., Weinstein, P., Souares, Y., and Woodward, A. (1999). El Nino and the dynamics of vector borne disease transmission. Environ. Health Perspect. *107*.

Hales, S., De Wet, N., Maindonald, J., and Woodward, A. (2002). Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. The Lancet *360*, 830–834.

Halstead, S.B. (1970). Observations related to pathogenesis of dengue hemorrhagic fever. VI. Hypotheses and discussion. Yale J Biol Med *42*, 350–362.

Halstead, S.B. (2008). Dengue Virus–Mosquito Interactions. Annu. Rev. Entomol. *53*, 273–291.

Hand, D.J. (2001). Measuring Diagnostic Accuracy of Statistical Prediction Rules. Stat. Neerlandica *55*, 3–16.

Hanley, J.A., and McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology *143*, 29–36.

Hannachi, A., Jolliffe, I.T., and Stephenson, D.B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. Int. J. Climatol. *27*, 1119–1152.

Harrington, L.C., Scott, T.W., Lerdthusnee, K., Coleman, R.C., Costero, A., Clark, G.G., Jones, J.J., Kitthawee, S., Kittayapong, P., Sithiprasana, R., et al. (2005). Dispersal of the dengue vector Aedes aegypti within and between rural communities. Am.J.trop.Med.Hyg. *72*, 209–220.

Harrington, L.C., Françoisevermeylen, null, Jones, J.J., Kitthawee, S., Sithiprasasna, R., Edman, J.D., and Scott, T.W. (2008). Age-dependent survival of the dengue vector Aedes aegypti (Diptera: Culicidae) demonstrated by simultaneous release-recapture of different age cohorts. J. Med. Entomol. *45*, 307–313.

Hay, S.I., Myers, M.F., Burke, D.S., Vaughn, D.W., Endy, T., Ananda, N., Shanks, G.D., Snow, R.W., and Rogers, D.J. (2000). Etiology of interepidemic periods of mosquito-borne disease. Proc. Natl. Acad. Sci. *97*, 9335–9339.

Heads, M. (2008). Panbiogeography of New Caledonia, south-west Pacific: basal angiosperms on basement terranes, ultramafic endemics inherited from volcanic island arcs and old taxa endemic to young islands. J. Biogeogr. *35*, 2153–2175.

Heisterkamp, S.H., Dekkers, A.L.M., and Heijne, J.C.M. (2006). Automated detection of infectious disease outbreaks: hierarchical time series models. Stat. Med. *25*, 4179–4196.

Held, L., Höhle, M., and Hofmann, M. (2005). A statistical framework for the analysis of multivariate infectious disease surveillance counts. Stat. Model. *5*, 187–199.

Liu-Helmersson, J., Stenlund, H., Wilder-Smith, A., and Rocklöv, J. (2014). Vectorial Capacity of Aedes aegypti: Effects of Temperature and Implications for Global Dengue Epidemic Potential. PLoS ONE *9*, e89783.

Hethcote, H.W. (1989). Three Basic Epidemiological Models. In Applied Mathematical Ecology, S.A. Levin, T.G. Hallam, and L.J. Gross, eds. (Springer Berlin Heidelberg), pp. 119–144.

Hii, Y.L., Rocklöv, J., Ng, N., Tang, C.S., Pang, F.Y., and Sauerborn, R. (2009). Climate variability and increase in intensity and magnitude of dengue incidence in Singapore. Glob. Health Action *2*.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. *25*, 1965–1978.

Hoffmann, A.A., Montgomery, B.L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P.H., Muzzi, F., Greenfield, M., Durkan, M., Leong, Y.S., Dong, Y., et al. (2011). Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. Nature *476*, 454–457.

Höhle, M., and Paul, M. (2008). Count data regression charts for the monitoring of surveillance time series. Comput. Stat. Data Anal. *52*, 4357–4368.

Höhle, M., Paul, M., and Held, L. (2009). Statistical approaches to the monitoring and surveillance of infectious diseases for veterinary public health. Prev. Vet. Med. *91*, 2–10.

Holmes, E.C., and Burch, S.S. (2000). The causes and consequences of genetic variation in dengue virus. Trends Microbiol. *8*, 74–77.

Holmes, E.C., and Twiddy, S.S. (2003). The origin, emergence and evolutionary genetics of dengue virus. Infect. Genet. Evol. *3*, 19–28.

Holmes, E.C., Worobey, M., and Rambaut, A. (1999). Phylogenetic evidence for recombination in dengue virus. Mol. Biol. Evol. *16*, 405–409.

Honório, N.A., Nogueira, R.M.R., Codeço, C.T., Carvalho, M.S., Cruz, O.G., de Avelar Figueiredo Mafra Magalhães, M., de Araújo, J.M.G., de Araújo, E.S.M., Gomes, M.Q., Pinheiro, L.S., et al. (2009). Spatial Evaluation and Modeling of Dengue Seroprevalence and Vector Density in Rio de Janeiro, Brazil. PLoS Negl. Trop. Dis. *3*, e545.

Hopp, M.J., and Foley, J.A. (2001). Global-scale relationships between climate and the Dengue fever vector, Aedes aegypti. Clim. Change *48*, 441–463.

Hopp, M.J., and Foley, J.A. (2003). Worldwide fluctuations in dengue fever cases related to climate variability. Clim. Res. *25*, 85–94.

Horel, J.D. (1981). A Rotated Principal Component Analysis of the Interannual Variability of the Northern Hemisphere 500 mb Height Field. Mon. Weather Rev. *109*, 2080–2092.

Hu, W., Clements, A., Williams, G., Tong, S., and Mengersen, K. (2011). Spatial Patterns and Socioecological Drivers of Dengue Fever Transmission in Queensland, Australia. Environ. Health Perspect. *120*, 260–266.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., and Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci. *454*, 903–995.

Huang, X., Clements, A.C., Williams, G., Milinovich, G., and Hu, W. (2013a). A threshold analysis of dengue transmission in terms of weather variables and imported dengue cases in Australia. Emerg. Microbes Infect. *2*, e87.

Huang, X., Williams, G., Clements, A.C.A., and Hu, W. (2013b). Imported Dengue Cases, Weather Variation and Autochthonous Dengue Incidence in Cairns, Australia. PLoS ONE *8*, e81887.

Hugo, L.E., Jeffery, J.A.L., Trewin, B.J., Wockner, L.F., Thi Yen, N., Le, N.H., Nghia, L.T., Hine, E., Ryan, P.A., and Kay, B.H. (2014). Adult Survivorship of the Dengue Mosquito Aedes aegypti Varies Seasonally in Central Vietnam. PLoS Negl Trop Dis *8*, e2669.

Ben-Hur, A., and Weston, J. (2010). A User's Guide to Support Vector Machines. In Data Mining Techniques for the Life Sciences, O. Carugo, and F. Eisenhaber, eds. (Totowa, NJ: Humana Press), pp. 223–239.

Huy, R., Buchy, P., Conan, A., Ngan, C., Ong, S., Ali, R., Duong, V., Yit, S., Ung, S., Te, V., et al. (2010). National dengue surveillance in Cambodia 1980-2008: epidemiological and virological trends and the impact of vector control. Bull. World Health Organ. *88*, 650–657.

Institut de la statistique et des études économiques (ISEE) (1989). Recensement de la population - 1989.

Institut de la statistique et des études économiques (ISEE) (1996). Recensement de la population - 1996.

Institut de la statistique et des études économiques (ISEE) (2004). Recensement de la population - 2004.

Institut de la statistique et des études économiques (ISEE) (2009). Recensement de la population - 2009.

IPNC (1996). Rapport technique - Année 1996 (Nouméa: Institut Pasteur de Nouvelle-Calédonie).

IPNC (1999). Rapport technique - Année 1999 (Nouméa: Institut Pasteur de Nouvelle-Calédonie).

IPNC (2003). Rapport Technique - Année 2003 (Nouméa: Institut Pasteur de Nouvelle-Calédonie).

Jansen, C.C., and Beebe, N.W. (2010). The dengue vector Aedes aegypti: what comes next. Microbes Infect. *12*, 272–279.

Jetten, T.H., and Focks, D.A. (1997). Potential changes in the distribution of dengue transmission under climate warming. Am. J. Trop. Med. Hyg. *57 (3)*, 285–297.

Johansson, M.A., Dominici, F., and Glass, G.E. (2009a). Local and Global Effects of Climate on Dengue Transmission in Puerto Rico. PLoS Negl Trop Dis *3*, e382.

Johansson, M.A., Cummings, D.A.T., and Glass, G.E. (2009b). Multiyear Climate Variability and Dengue—El Niño Southern Oscillation, Weather, and Dengue Incidence in Puerto Rico, Mexico, and Thailand: A Longitudinal Data Analysis. PLoS Med. *6*, e1000168.

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., and Daszak, P. (2008). Global trends in emerging infectious diseases. Nature *451*, 990–993.

JORF n°121 du 27 mai 1998 page 8039 Accord sur la Nouvelle-Calédonie signé à Nouméa le 5 mai 1998.

Khormi, H.M., Kumar, L., and Elzahrany, R.A. (2011). Modeling spatio-temporal risk changes in the incidence of dengue fever in Saudi Arabia: a geographical information system case study. Geospatial Health *6*, 77–84.

Kiedrzynski, T., Souarès, Y., and Stewart, T. (1998). Dengue in the Pacific: an updated story. Pac. Health Dialog *5*, 129–136.

Kliks, S.C., Nimmanitya, S., Nisalak, A., and Burke, D.S. (1988). Evidence that maternal dengue antibodies are important in the development of dengue hemorrhagic fever in infants. Am. J. Trop. Med. Hyg. *38*, 411–419.

Koelle, K., and Pascual, M. (2004). Disentangling Extrinsic from Intrinsic Factors in Disease Dynamics: A Nonlinear Time Series Approach with an Application to Cholera. Am. Nat. *163*, 901–913.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In IJCAI, pp. 1137–1145.

Kool, J.L., Paterson, B., Pavlin, B.I., Durrheim, D., Musto, J., and Kolbe, A. (2012). Pacific-wide simplified syndromic surveillance for early warning of outbreaks. Glob. Public Health *7*, 670–681.

Kouri, G.P., Guzmán, M.G., Bravo, J.R., and Triana, C. (1989). Dengue haemorrhagic fever/dengue shock syndrome: lessons from the Cuban epidemic, 1981. Bull. World Health Organ. *67*, 375.

Kuhn, R.J., Zhang, W., Rossmann, M.G., Pletnev, S.V., Corver, J., Lenches, E., Jones, C.T., Mukhopadhyay, S., Chipman, P.R., Strauss, E.G., et al. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell *108*, 717–725.

Kuno, G. (1995). Review of the factors modulating dengue transmission. Epidemiol. Rev. *17 (2)*, 321–335.

Kyle, J.L., and Harris, E. (2008). Global Spread and Persistence of Dengue. Annu. Rev. Microbiol. *62*, 71–92.

Lambrechts, L. (2011). Quantitative genetics of Aedes aegypti vector competence for dengue viruses: towards a new paradigm? Trends Parasitol. *27*, 111–114.

Lambrechts, L., Scott, T.W., and Gubler, D.J. (2010). Consequences of the Expanding Global Distribution of Aedes albopictus for Dengue Virus Transmission. PLoS Negl. Trop. Dis. *4*, e646.

Lambrechts, L., Paaijmans, K.P., Fansiri, T., Carrington, L.B., Kramer, L.D., Thomas, M.B., and Scott, T.W. (2011). Impact of daily temperature fluctuations on dengue virus transmission by Aedes aegypti. Proc. Natl. Acad. Sci. *108*, 7460–7465.

Lee, K.-S., Lai, Y.-L., Lo, S., Barkham, T., Aw, P., Ooi, P.-L., Tai, J.-C., Hibberd, M., Johansson, P., Khoo, S.-P., et al. (2010). Dengue Virus Surveillance for Early Warning, Singapore. Emerg. Infect. Dis. *16*, 847–849.

Lefèvre, J., Marchesiello, P., Jourdain, N.C., Menkes, C., and Leroy, A. (2010). Weather regimes and orographic circulation around New Caledonia. Mar. Pollut. Bull. *61*, 413–431.

Leitmeyer, K.C., Vaughn, D.W., Watts, D.M., Salas, R., Villalobos de Chacon, I., Ramos, C., and Rico-Hesse, R. (1999). Dengue virus structural differences that correlate with pathogenesis. J. Virol. *73*, 4738–4747.

Leo, Y.-S., Thein, T.L., Fisher, D.A., Low, J.G., Oh, H.M., Narayanan, R.L., Gan, V.C., Lee, V.J., and Lye, D.C. (2011). Confirmed adult dengue deaths in Singapore: 5-year multi-center retrospective study. BMC Infect. Dis. *11*, 123.

Leport, C., Guégan, J.-F., and (sous la direction de) (2011). Les maladies infectieuses émergentes état de la situation et perspectives ([Paris]: la Documentation française).

Leroy, A. (2011). Prévisibilité des épidémies de dengue par Arpège à échéance saisonnière - Prévisibilité des fréquences d'occurrence de jours chauds et jours pluvieux (Météo France).

Li, D., Liu, W., Guigon, A., Mostyn, C., Grant, R., and Aaskov, J. (2010). Rapid Displacement of Dengue Virus Type 1 by Type 4, Pacific Region, 2007–2009. Emerg. Infect. Dis. *16*, 123–125.

Liebman, K.A., Stoddard, S.T., Morrison, A.C., Rocha, C., Minnick, S., Sihuincha, M., Russell, K.L., Olson, J.G., Blair, P.J., Watts, D.M., et al. (2012). Spatial Dimensions of Dengue Virus Transmission across Interepidemic and Epidemic Periods in Iquitos, Peru (1999–2003). PLoS Negl. Trop. Dis. *6*, e1472.

Liew, C., and Curtis, C.F. (2004). Horizontal and vertical dispersal of dengue vector mosquitoes, Aedes aegypti and Aedes albopictus, in Singapore. Med. Vet. Entomol. *18*.

Lorenz, E.N. (1956). Empirical Orthogonal Functions and Statistical Weather Prediction (Massachusetts Institute of Technology, Department of meteorology).

Lowe, R., Bailey, T.C., Stephenson, D.B., Graham, R.J., Coelho, C.A.S., Sá Carvalho, M., and Barcellos, C. (2011). Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil. Comput. Geosci. *37*, 371–381.

Lowe, R., Bailey, T.C., Stephenson, D.B., Jupp, T.E., Graham, R.J., Barcellos, C., and Carvalho, M.S. (2013). The development of an early warning system for climate-sensitive disease risk with a focus on dengue epidemics in Southeast Brazil. Stat. Med. *32*, 864–883.

Lu, L., Lin, H., Tian, L., Yang, W., Sun, J., and Liu, Q. (2009). Time series analysis of dengue fever and weather in Guangzhou, China. BMC Public Health *9*, 395.

MacDonald, G. (1957). The Epidemiology and Control of Malaria (Oxford University Press).

Maciel-de-Freitas, R., Codeco, C.T., and Lourenço-de-Oliveira, R. (2007). Body size-associated survival and dispersal rates of Aedes aegypti in Rio de Janeiro. Med. Vet. Entomol. *21*, 284–292.

Madden, R.A., and Julian, P.R. (1971). Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific. J. Atmospheric Sci. *28*, 702–708.

Mammen Jr, M.P., Pimgate, C., Koenraadt, C.J., Rothman, A.L., Aldstadt, J., Nisalak, A., Jarman, R.G., Jones, J.W., Srikiatkhachorn, A., Ypil-Butac, C.A., et al. (2008). Spatial and temporal clustering of dengue virus transmission in Thai villages. PLoS Med. *5*, e205.

Massad, E., Coutinho, F.A.B., Burattini, M.N., and Amaku, M. (2009). Estimation of R_0 from the initial phase of an outbreak of a vector-borne infection. Trop. Med. Int. Health.

McMichael, A.J., Woodruff, R.E., and Hales, S. (2006). Climate change and human health: present and future risks. The Lancet *367*, 859–869.

Météo France (2014). Formulaire de demande de données.

Mohammed, A., and Chadee, D.D. (2011). Effects of different temperature regimens on the development of Aedes aegypti (L.) (Diptera: Culicidae) mosquitoes. Acta Trop. *119*, 38–43.

Monath, T.P. (1994). Dengue: the risk to developed and developing countries. Proc. Natl. Acad. Sci. *91*, 2395–2400.

Morin, C.W., Comrie, A.C., and Ernst, K. (2013). Climate and Dengue Transmission: Evidence and Implications. Environ. Health Perspect. *121*, 1264–1272.

Morrison, A.C., Getis, A., Santigo, M., Rigau-Perez, J.G., and Reiter, P. (1998). Exploratory space-time analysis of reported denue cases during an outbreak in Florida, Puerto Rico, 1991-1992. Am. J. Trop. Med. Hyg. *58*, 287–298.

Muir, L.E., and Kay, B.H. (1998). Aedes aegypti survival and dispersal estimated by mark-release-recapture in northern Australia. Am. J. Trop. Med. Hyg. *58 (3)*, 277–282.

Muñoz-Jordán, J.L., Sánchez-Burgos, G.G., Laurent-Rolle, M., and García-Sastre, A. (2003). Inhibition of interferon signaling by dengue virus. Proc. Natl. Acad. Sci. *100*, 14333–14338.

Murgue, B., Roche, C., Chungue, E., and Deparis, X. (2000). Prospective study of the duration and magnitude of viraemia in children hospitalised during the 1996-1997 dengue-2 outbreak in French Polynesia. J. Med. Virol. *60*, 432–438.

Musso, D., Nilles, E.J., and Cao-Lormeau, V.-M. (2014). Rapid spread of emerging Zika virus in the Pacific area. Clin. Microbiol. Infect. n/a – n/a.

Naish, S., Dale, P., Mackenzie, J.S., McBride, J., Mengersen, K., and Tong, S. (2014). Climate change and dengue: a critical and systematic review of quantitative modelling approaches. BMC Infect. Dis. *14*, 167.

Nakhapakorn, K., and Tripathi, N.K. (2005). An information value based analysis of physical and climatic factors affecting dengue fever and dengue haemorrhagic fever incidence. Int. J. Health Geogr. *4*, 13.

Narahashi, T. (1982). Cellular and molecular mechanisms of action of insecticides: neurophysiological approach. Neurobehav. Toxicol. Teratol. *4*, 753–758.

Narahashi, T. (1996). Neuronal ion channels as the target sites of insecticides. Pharmacol. Toxicol. *79*, 1–14.

Nasci, R.S., and Mitchell, C.J. (1994). Larval diet, adult size, and susceptibility of Aedes aegypti (Diptera: Culicidae) to infection with Ross River virus. J Med Ent *31(1)*, 123–126.

North, G.R., Bell, T.L., Cahalan, R.F., and Moeng, F.J. (1982). Sampling Errors in the Estimation of Empirical Orthogonal Functions. Mon. Weather Rev. *110*, 699–706.

Ooi, E.-E., Goh, K.-T., and Gubler, D.J. (2006). Dengue Prevention and 35 Years of Vector Control in Singapore. Emerg. Infect. Dis. *12*, 887–893.

Ortíz Bultó, P.L., Rodríguez, A.P., Valencia, A.R., Vega, N.L., Díaz, M., and Carrera, A.P. (2006). Assessment of Human Health Vulnerability to Climate Variability and Change in Cuba. Environ. Health Perspect.

Pant, C.P., and Yasuno, M. (1973). Field studies on the gonotrophic cycle of Aedes aegypti in Bangkok, Thailand. J Med Ent *10(2)*, 219–223.

Patz, J.A., Martens, W.J., Focks, D.A., and Jetten, T.H. (1998). Dengue fever epidemic potential as projected by general circulation models of global climate change. Environ. Health Perspect. *106*, 147–153.

Penna, F., and Lucia, M. (2004). Ecological Study of Rio de Janeiro City DEN-3 Epidemic, 2001-2002.

Perry, W.J. (1950). The Mosquitoes and Mosquito-Borne Diseases on New Caledonia, an Historic Account: 1885–1946. Am. J. Trop. Med. Hyg. *s1-30*, 103–114.

Pessanha, J.E.M.P., Caiaffa, W.T., Almeida, M.C. de M., Brandão, S.T., and Proietti, F.A. (2012). Diffusion Pattern and Hotspot Detection of Dengue in Belo Horizonte, Minas Gerais, Brazil. J. Trop. Med. *2012*, 1–11.

Pham, H.V., Doan, H.T., Phan, T.T., and Minh, N.N.T. (2011). Ecological factors associated with dengue fever in a central highlands Province, Vietnam. BMC Infect. Dis. *11*, 172.

Pinto, E., Coelho, M., Oliver, L., and Massad, E. (2011). The influence of climate variables on dengue in Singapore. Int. J. Environ. Health Res. *21*, 415–426.

Plaut, G., and Vautard, R. (1994). Spells of Low-Frequency Oscillations and Weather Regimes in the Northern Hemisphere. J. Atmospheric Sci. *51*, 210–236.

Pongsiri, A., Ponlawat, A., Thaisomboonsuk, B., Jarman, R.G., Scott, T.W., and Lambrechts, L. (2014). Differential Susceptibility of Two Field Aedes aegypti Populations to a Low Infectious Dose of Dengue Virus. PLoS ONE *9*, e92971.

Pontes, R.J., Freeman, J., Oliveira-Lima, J.W., Hodgson, J.C., and Spielman, A. (2000). Vector densities that potentiate dengue outbreaks in a Brazilian city. Am. J. Trop. Med. Hyg. *62*, 378–383.

Pumpuni, C.B., Knepler, J., and Craig, G.B. (1992). Influence of temperature and larval nutrition on the diapause inducing photoperiod of Aedes albopictus. J. Am. Mosq. Control Assoc. *8*, 223–227.

Rabaa, M.A., Klungthong, C., Yoon, I.-K., Holmes, E.C., Chinnawirotpisan, P., Thaisomboonsuk, B., Srikiatkhachorn, A., Rothman, A.L., Tannitisupawong, D., Aldstadt, J., et al. (2013). Frequent In-Migration and Highly Focal Transmission of Dengue Viruses among Children in Kamphaeng Phet, Thailand. PLoS Negl. Trop. Dis. *7*, e1990.

Racloz, V., Ramsey, R., Tong, S., and Hu, W. (2012). Surveillance of Dengue Fever Virus: A Review of Epidemiological Models and Early Warning Systems. PLoS Negl. Trop. Dis. *6*, e1648.

Rasmusson, E.M., and Carpenter, T.H. (1982). Variations in Tropical Sea Surface Temperature and Surface Wind Fields Associated with the Southern Oscillation/El Niño. Mon. Weather Rev. *110*, 354–384.

Rasmusson, E.M., Arkin, P.A., Chen, W.-Y., and Jalickee, J.B. (1981). Biennial Variations in Surface Temperature over the United States as Revealed by Singular Decomposition. Mon. Weather Rev. *109*, 587–598.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reich, N.G., Shrestha, S., King, A.A., Rohani, P., Lessler, J., Kalayanarooj, S., Yoon, I.-K., Gibbons, R.V., Burke, D.S., and Cummings, D.A.T. (2013). Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. J. R. Soc. Interface *10*, 20130414–20130414.

Reiner, R.C., Stoddard, S.T., and Scott, T.W. (2014). Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. Epidemics *6*, 30–36.

Reiter, P. (2001). Climate change and mosquito-borne disease. Environ. Health Perspect. *109*, 141–161.

Reiter, P., Lathrop, S., Bunning, M., Biggerstaff, B.J., Singer, D., Tiwari, T., Baber, L., Amador, M., Thirion, J., Hayes, J., et al. (2003). Texas lifestyle limits transmission of dengue virus. Emerg. Infect. Dis. *9*, 86–89.

Richman, M.B. (1981). Obliquely Rotated Principal Components: An Improved Meteorological Map Typing Technique? J. Appl. Meteorol. *20*, 1145–1159.

Rico-Hesse, R. (1990). Molecular evolution and distribution of Dengue viruses type 1 and 2 in nature. Virology *174*, 479–793.

Rico-Hesse, R. (2003). Microevolution and virulence of dengue viruses. Adv. Virus Res. *59*, 315–341.

Rico-Hesse, R., Harrison, L.M., Salas, R.A., Tovar, D., Nisalak, A., Ramos, C., Boshell, J., de Mesa, M.T.R., Nogueira, R.M.R., and Rosa, A.T. da (1997). Origins of Dengue Type 2 Viruses Associated with Increased Pathogenicity in the Americas. Virology *230*, 244–251.

Rigau-Pérez, J.G., Clark, G.G., Gubler, D.J., Reiter, P., Sanders, E.J., and Vance Vorndam, A. (1998). Dengue and dengue haemorrhagic fever. The Lancet *352*, 971–977.

Roger, F., Thonnat, J., Hendrikx, P., and Domenech, J. (2004). Les systèmes de suivi et de surveillance des maladies et le rôle des acteurs de santé animale publics et privés : l'expérience de l'Afrique. Rev. Sci. Tech. Int. Off. Epizoot. *23*, 137–145.

Rohani, P. (2009). The Link between Dengue Incidence and El Niño Southern Oscillation. PLoS Med *6*, e1000185.

Rosen, L. (1996). Dengue hemorrhagic fever. Bull. Société Pathol. Exot. *89*, 91–94.

Rosen, L., and Gubler, D. (1974). The use of mosquitoes to detect and propagate Dengue virus. Am. J. Trop. Med. Hyg. *23(6)*, 1153–1160.

Rosen, L., Shroyer, D.A., Tesh, R.B., Freier, J.E., and Lien, J.C. (1983). Transovarial transmission of Dengue viruses by mosquitoes: Aedes albopictus and Aedes aegypti. Am. J. Trop. Med. Hyg. *32(5)*, 1108–1119.

Rossi, G., Lampugnani, L., and Marchi, M. (1999). An approximate CUSUM procedure for surveillance of health events. Stat. Med. *18*, 2111–2122.

Roth, A., Hoy, D., Horwood, P.F., Ropa, B., Hancock, T., Guillaumot, L., Rickart, K., Frison, P., Pavlin, B., and Souares, Y. (2014). Preparedness for Threat of Chikungunya in the Pacific. Emerg. Infect. Dis. *20*.

Rueda, L.M., Patel, K.J., Axtell, R.C., and Stinner, R.E. (1990). Temperature-dependent development and survival rates of Culex quinquefasciatus and Aedes aegypti (Diptera: Culicidae). J. Med. Entomol. *27(5)*, 892–898.

Sabchareon, A., Wallace, D., Sirivichayakul, C., Limkittikul, K., Chanthavanich, P., Suvannadabba, S., Jiwariyavej, V., Dulyachai, W., Pengsaa, K., Wartel, T.A., et al. (2012). Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. The Lancet *380*, 1559–1567.

Sabin, A.B. (1952). Research on Dengue during World War II. Am. J. Trop. Med. Hyg. *1*, 30–50.

Salje, H., Lessler, J., Endy, T.P., Curriero, F.C., Gibbons, R.V., Nisalak, A., Nimmannitya, S., Kalayanarooj, S., Jarman, R.G., Thomas, S.J., et al. (2012). Revealing the microscale spatial signature of dengue transmission and immunity in an urban population. Proc. Natl. Acad. Sci. *109*, 9535–9538.

Sanchez, L., Cortinas, J., Pelaez, O., Gutierrez, H., Concepción, D., and Van der Stuyft, P. (2010). Breteau Index threshold levels indicating risk for dengue transmission in areas with low Aedes infestation. Trop. Med. Int. Health *15*, 173–175.

Sarfraz, M.S., Tripathi, N.K., Tipdecho, T., Thongbu, T., Kerdthong, P., and Souris, M. (2012). Analyzing the spatio-temporal relationship between dengue vector larval density and land-use using factor analysis and spatial ring mapping. BMC Public Health *12*, 853.

Scherrer, B. (2008). Biostatistique : Volume 1 (Montréal: Gaëtan Morin).

Scherrer, B., Legendre, P., and Beltangady, M. (2009). Biostatistique : Volume 2 (Montréal: Gaetan Morin).

Schneider, J.R., Mori, A., Romero-Severson, J., Chadee, D.D., and Severson, D.W. (2007). Investigations of dengue-2 susceptibility and body size among Aedes aegypti populations. Med. Vet. Entomol. *21*, 370–376.

Scott, T.W., Chow, E., Strickman, D., Kittayapong, P., Wirtz, R.A., Lorenz, L., and Edman, J.D. (1993). Blood-feeding patterns of Aedes aegypti (Diptera: Culicidae) collected in a rural Thai village. J Med Ent *30*, 922–927.

Scott, T.W., Amerasinghe, P.H., Morrison, A.C., Lorenz, L.H., Clark, G.G., Strickman, D., Kittayapong, P., and Edman, J.D. (2000). Longitudinal studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: blood feeding frequency. J Med Ent *37*, 89–101.

Seidahmed, O.M.E., Hassan, S.A., Soghaier, M.A., Siam, H.A.M., Ahmed, F.T.A., Elkarsany, M.M., and Sulaiman, S.M. (2012). Spatial and Temporal Patterns of Dengue Transmission along a Red Sea Coastline: A Longitudinal Entomological and Serological Survey in Port Sudan City. PLoS Negl. Trop. Dis. *6*, e1821.

Serfling, R.E. (1963). Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths. Public Health Rep. 1896-1970 *78*, 494.

Shang, C.-S., Fang, C.-T., Liu, C.-M., Wen, T.-H., Tsai, K.-H., and King, C.-C. (2010). The Role of Imported Cases and Favorable Meteorological Conditions in the Onset of Dengue Epidemics. PLoS Negl. Trop. Dis. *4*, e775.

Shi, R.X., Wang, J.F., Xu, C.D., Lai, S.J., and Yang, W.Z. (2014). Spatiotemporal pattern of hand–foot–mouth disease in China: an analysis of empirical orthogonal functions. Public Health *128*, 367–375.

Shroyer, D.A. (1990). Vertical maintenance of dengue-1 virus in sequential generations of Aedes albopictus. J Amer Mosq Cont Ass *6(2)*, 312–314.

Sinclair, M.R. (1996). A Climatology of Anticyclones and Blocking for the Southern Hemisphere. Mon. Weather Rev. *124*, 245–264.

Singh, N., Kiedrzynski, T., Lepers, C., and Benyon, E.K. (2005). Dengue in the Pacific-an update of the current situation. Pac. Health Dialog *12*, 111–119.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., and Powers, J.G. (2005). A description of the advanced research WRF version 2 (DTIC Document).

Smith, D.L., Perkins, T.A., Reiner, R.C., Barker, C.M., Niu, T., Chaves, L.F., Ellis, A.M., George, D.B., Menach, A.L., Pulliam, J.R.C., et al. (2014a). Recasting the theory of mosquito-borne pathogen transmission dynamics and control. Trans. R. Soc. Trop. Med. Hyg. *108*, 185–197.

Smith, K.R., Woodward, A., and (sous la direction de) (2014b). Assessment Report n° 5. Working Group II. Climate Change 2014: Impacts, Adaptation, and Vulnerability (Intergovernmental Panel on Climate Change).

Spiegel, J.M., Bonet, M., Ibarra, A.-M., Pagliccia, N., Ouellette, V., and Yassi, A. (2007). Social and environmental determinants of Aedes aegypti infestation in Central Havana: results of a case-control study nested in an integrated dengue surveillance programme in Cuba: Social and environmental determinants of Aedes aegypti infestation. Trop. Med. Int. Health *12*, 503–510.

Steel, A., Gubler, D.J., and Bennett, S.N. (2010). Natural attenuation of dengue virus type-2 after a series of island outbreaks: A retrospective phylogenetic study of events in the South Pacific three decades ago. Virology *405*, 505–512.

Steinhauer, D.A., and Holland, J.J. (1987). Rapid Evolution of RNA Viruses. Annu. Rev. Microbiol. *41*, 409–431.

Stewart Ibarra, A.M., Ryan, S.J., Beltrán, E., Mejía, R., Silva, M., and Muñoz, Á. (2013). Dengue Vector Dynamics (Aedes aegypti) Influenced by Climate and Social Factors in Ecuador: Implications for Targeted Control. PLoS ONE *8*, e78263.

Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Paz Soldan, V., Kochel, T.J., Kitron, U., Elder, J.P., and Scott, T.W. (2009). The Role of Human Movement in the Transmission of Vector-Borne Pathogens. PLoS Negl. Trop. Dis. *3*, e481.

Stoddard, S.T., Forshey, B.M., Morrison, A.C., Paz-Soldan, V.A., Vazquez-Prokopec, G.M., Astete, H., Reiner, R.C., Vilcarromero, S., Elder, J.P., Halsey, E.S., et al. (2013). House-to-house human movement drives dengue virus transmission. Proc. Natl. Acad. Sci. *110*, 994–999.

Stroup, D.F., Williamson, G.D., Herndon, J.L., and Karon, J.M. (1989). Detection of aberrations in the occurrence of notifiable diseases surveillance data. Stat. Med. *8*, 323–329.

Styer, L.M., Carey, J.R., Wang, J.-L., and Scott, T.W. (2007). Mosquitoes do senesce: departure from the paradigm of constant mortality. Am. J. Trop. Med. Hyg. *76*, 111–117.

Sumanochitrapon, W., Strickman, D., Sithiprasana, R., Kittayapong, P., and innis, B.L. (1998). Effect of size and geographic origin of Aedes aegypti on oral infection with dengue-2 virus. Am. J. Trop. Med. Hyg. *58*, 283–286.

Takahashi, M. (1976). The effects of environmental and physiological conditions of Culex tritaeniorhynchus on the pattern of transmission of Japanese encephalitis virus. J Med Ent *13(3)*, 275–284.

Talbalaghi, A., Moutailler, S., Vazeille, M., and Failloux, A.-B. (2010). Are Aedes albopictus or other mosquito species from northern Italy competent to sustain new arboviral outbreaks? Med. Vet. Entomol. *24*, 83–87.

Tambyah, P.A., Koay, E.S.C., Poon, M.L.M., Lin, R.V.T.P., and Ong, B.K.C. (2008). Dengue Hemorrhagic Fever Transmitted by Blood Transfusion. N. Engl. J. Med. *359*, 1526–1527.

Taylor, K.E., Stouffer, R.J., and Meehl, G.A. (2012). An Overview of CMIP5 and the Experiment Design. Bull. Am. Meteorol. Soc. *93*, 485–498.

Teng, H.-J., and Apperson, C.S. (2000). Development and Survival of Immature Aedes albopictus and Aedes triseriatus (Diptera: Culicidae) in the Laboratory: Effects of Density, Food, and Competition on Response to Temperature. J. Med. Entomol. *37*, 40–52.

Teurlai, M., Huy, R., Cazelles, B., Duboz, R., Baehr, C., and Vong, S. (2012). Can Human Movements Explain Heterogeneous Propagation of Dengue Fever in Cambodia? PLoS Negl Trop Dis *6*, e1957.

Thai, K.T.D., Cazelles, B., Nguyen, N.V., Vo, L.T., Boni, M.F., Farrar, J., Simmons, C.P., van Doorn, H.R., and de Vries, P.J. (2010). Dengue Dynamics in Binh Thuan Province, Southern Vietnam: Periodicity, Synchronicity and Climate Variability. PLoS Negl. Trop. Dis. *4*, e747.

Thai, K.T.D., Nishiura, H., Hoang, P.L., Tran, N.T.T., Phan, G.T., Le, H.Q., Tran, B.Q., Nguyen, N.V., and de Vries, P.J. (2011). Age-Specificity of Clinical Dengue during Primary and Secondary Infections. PLoS Negl Trop Dis *5*, e1180.

Thammapalo, S., Chongsuwiwatwong, V., Geater, A., Lim, A., and Choomalee, K. (2005). Socio-demographic and environmental factors associated with Aedes breeding places in Phuket, Thailand. Southeast Asian J. Trop. Med. Public Health *36*, 426–433.

Tipayamongkholgul, M., and Lisakulruk, S. (2011). Socio-geographical factors in vulnerability to dengue in Thai villages: a spatial regression analysis. Geospatial Health *5*, 191–198.

Toma, B., Dufour, B., Bénet, J.-J., Sanaa, M., Shaw, A., and Moutou, F. (2001). Epidémiologie appliquée à la lutte collective contre les maladies animales transmissibles majeures (AEEMA).

Toma, B., André-Fontaine, G., Artois, M., Augustin, J.C., Bastian, S., Bénet, J.-J., Cerf, O., Haddad, N., Lacheretz, A., Picavet, D.P., et al. (2004). Les zoonoses infectieuses, Polycopié des Unités de maladies contagieuses des Ecoles vétérinaires françaises.

Torrence, C., and Compo, G.P. (1998). A Practical Guide to Wavelet Analysis. Bull. Am. Meteorol. Soc. *79*, 61–78.

Tran, A., Deparis, X., Dussart, P., Morvan, J., Rabarisn, P., Remy, F., Polidori, L., and Gardon, J. (2004). Dengue spatial and temporal patterns, French Guiana, 2001. Emerg. Infect. Dis. *10*, 615–621.

Tran, A., Ponçon, N., Toty, C., Linard, C., Guis, H., Ferré, J.-B., Lo Seen, D., Roger, F., de la Rocque, S., Fontenille, D., et al. (2008). Using remote sensing to map larval and adult populations of Anopheles hyrcanus (Diptera: Culicidae) a potential malaria vector in Southern France. Int. J. Health Geogr. *7*, 9.

Trent, D.W., Manske, C.L., Fox, G.E., Chu, M.C., Kliks, S.C., and Monath, T.P. (1990). The molecular epidemiology of dengue viruses. Genetic variation and microevolution Applied Virology Research, vol.2. E. Kurstak, R.G. Marusyk, F.A. Murphy, and M.H.V. van Regenmortel, eds. (Plenum Publishing Corporation), pp. 293–315.

Turbelin, C., Pelat, C., Boëlle, P.-Y., Lévy-Bruhl, D., Carrat, F., Blanchon, T., Hanslik, T., and others (2009). Early estimates of 2009 pandemic influenza A (H1N1) virus activity in general practice in France: incidence of influenza-like illness and age distribution of reported cases. Eurosurveillance *14*.

Twiddy, S.S., Woelk, C.H., and Holmes, E.C. (2002). Phylogenetic evidence for adaptive evolution of dengue viruses in nature. J. Gen. Virol. *83*, 1679–1689.

Ty Hang, V.T., Holmes, E.C., Veasna, D., Quy, N.T., Tinh Hien, T., Quail, M., Churcher, C., Parkhill, J., Cardosa, J., Farrar, J., et al. (2010). Emergence of the Asian 1 Genotype of Dengue Virus Serotype 2 in Viet Nam: In Vivo Fitness Advantage and Lineage Replacement in South-East Asia. PLoS Negl Trop Dis *4*, e757.

Vapnik, V. (1998). Statistical learning theory (New York: Wiley).

Vaughn, D.W., Green, S., Kalayanarooj, S., Innis, B.L., Nimmannitya, S., Suntayakorn, S., Endy, T.P., Raengsakulrach, B., Rothman, A.L., Ennis, F.A., et al. (2000). Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J. Infect. Dis. *181*, 2–9.

Vazeille, M., Moutailler, S., Pages, F., Jarjaval, F., and Failloux, A.-B. (2008). Introduction of *Aedes albopictus* in Gabon: what consequences for dengue and chikungunya transmission? Trop. Med. Int. Health *13*, 1176–1179.

Vazeille, M., Mousson, L., Martin, E., and Failloux, A.-B. (2010). Orally Co-Infected Aedes albopictus from La Reunion Island, Indian Ocean, Can Deliver Both Dengue and Chikungunya Infectious Viral Particles in Their Saliva. PLoS Negl. Trop. Dis. *4*, e706.

Vazquez-Prokopec, G.M., Stoddard, S.T., Paz-Soldan, V., Morrison, A.C., Elder, J.P., Kochel, T.J., Scott, T.W., and Kitron, U. (2009). Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. Int. J. Health Geogr. *8*, 68.

Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., and Ritchie, S.A. (2010). Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. PLoS Negl. Trop. Dis. *4*, e920.

Vincent, D.G. (1994). The South Pacific Convergence Zone (SPCZ): A Review. Mon. Weather Rev. *122*, 1949–1970.

Vincent, E.M., Lengaigne, M., Menkes, C.E., Jourdain, N.C., Marchesiello, P., and Madec, G. (2011). Interannual variability of the South Pacific Convergence Zone and implications for tropical cyclone genesis. Clim. Dyn. *36*, 1881–1896.

Vong, S., Khieu, V., Glass, O., Ly, S., Duong, V., Huy, R., Ngan, C., Wichmann, O., Letson, G.W., Margolis, H.S., et al. (2010). Dengue Incidence in Urban and Rural Cambodia: Results from Population-Based Active Fever Surveillance, 2006–2008. PLoS Negl. Trop. Dis. *4*, e903.

Vynnycky, E., Trindall, A., and Mangtani, P. (2007). Estimates of the reproduction numbers of Spanish influenza using morbidity data. Int. J. Epidemiol. *36*, 881–889.

Wallace, J.M. (1972). Empirical Orthogonal Representation of Time Series in the Frequency Domain. Part II: Application to the Study of Tropical Wave Disturbances. J. Appl. Meteorol. *11*, 893–900.

Wallace, J.M., and Dickinson, R.E. (1972). Empirical Orthogonal Representation of Time Series in the Frequency Domain. Part I: Theoretical Considerations. J. Appl. Meteorol. *11*, 887–892.

Wallinga, J., and Teunis, P. (2004). Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. Am. J. Epidemiol. *160*, 509–516.

Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. *58*, 236–244.

Ward, M.P., Maftei, D., Apostu, C., and Suru, A. (2009). Estimation of the basic reproductive number (R0) for epidemic, highly pathogenic avian influenza subtype H5N1 spread. Epidemiol. Infect. *137*, 219–226.

Watts, D.M., Burke, D.S., Harrison, B.A., Whitmire, R.E., and Nisalak, A. (1987). Effect of temperature on the vector efficiency of Aedes aegypti for dengue 2 virus. Am. J. Trop. Med. Hyg. *36(1)*, 143–152.

Weare, B.C., and Nasstrom, J.S. (1982). Examples of Extended Empirical Orthogonal Function Analyses. Mon. Weather Rev. *110*, 481–485.

Wearing, H.J., and Rohani, P. (2006). Ecological and immunological determinants of dengue epidemics. Proc. Natl. Acad. Sci. *103*, 11802–11807.

WHO (1980). L' éradication mondiale de la variole : rapport final de la Commission mondiale pour la Certification de l' Eradication de la Variole, Genève, décembre 1979. The global eradication of smallpox : final report of the Global Commission for the Certification of Smallpox Eradication, Geneva, December 1979.

WHO (2009). Dengue guidelines for diagnosis, treatment, prevention, and control. (Geneva: TDR : World Health Organization).

Woolhouse, M.E.J., and Gowtage-Sequeria, S. (2005). Host Range and Emerging and Reemerging Pathogens. Emerg. Infect. Dis. *11*, 1842–1847.

Worobey, M., Rambaut, A., and Holmes, E.C. (1999). Widespread intra-serotype recombination in natural populations of dengue virus. Proc. Natl. Acad. Sci. *96*, 7352–7357.

Wu, P.-C., Lay, J.-G., Guo, H.-R., Lin, C.-Y., Lung, S.-C., and Su, H.-J. (2009). Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan. Sci. Total Environ. *407*, 2224–2233.

Yang, H.M., Macoris, M.L.G., Galvani, K.C., Andrighetti, M.T.M., and Wanderley, D.M.V. (2009). Assessing the effects of temperature on the population of Aedes aegypti, the vector of dengue. Epidemiol. Infect. *137*, 1188–1202.

Yap, G., Li, C., Mutalib, A., Lai, Y.-L., and Ng, L.-C. (2013). High Rates of Inapparent Dengue in Older Adults in Singapore. Am. J. Trop. Med. Hyg. *88*, 1065–1069.

Ye, Y.H., Ng, T.S., Frentiu, F.D., Walker, T., van den Hurk, A.F., O'Neill, S.L., Beebe, N.W., and McGraw, E.A. (2014). Comparative susceptibility of mosquito populations in North Queensland, Australia to oral infection with dengue virus. Am. J. Trop. Med. Hyg. *90*, 422–430.

Yoon, I.-K., Getis, A., Aldstadt, J., Rothman, A.L., Tannitisupawong, D., Koenraadt, C.J.M., Fansiri, T., Jones, J.W., Morrison, A.C., Jarman, R.G., et al. (2012). Fine Scale Spatiotemporal Clustering of Dengue Virus Transmission in Children and Aedes aegypti in Rural Thai Villages. PLoS Negl Trop Dis *6*, e1730.

Yu, H.-L., Yang, S.-J., Yen, H.-J., and Christakos, G. (2011). A spatio-temporal climate-based model of early dengue fever warning in southern Taiwan. Stoch. Environ. Res. Risk Assess. *25*, 485–494.

Zhang, C., Mammen, M.P., Chinnawirotpisan, P., Klungthong, C., Rodpradit, P., Monkongdee, P., Nimmannitya, S., Kalayanarooj, S., and Holmes, E.C. (2005). Clade Replacements in Dengue Virus Serotypes 1 and 3 Are Associated with Changing Serotype Prevalence. J. Virol. *79*, 15123–15130.

Annexe 1 : Glossaire

Epidémiologie

Prévalence	Nombre total de cas ou de foyers d'une maladie dans une population déterminée au cours d'une période donnée ou à instant donné.
Incidence	Nombre total de cas ou de foyers **nouveaux** d'une maladie dans une population déterminée au cours d'une période donnée.
Taux de prévalence	Nombre total de cas ou de foyers d'une maladie dans une population déterminée au cours d'une période donnée ou à instant donné divisé par le nombre d'individus soumis au risque.
Taux d'incidence	Nombre total de cas ou de foyers **nouveaux** d'une maladie dans une population déterminée au cours d'une période donnée divisé par le nombre d'individus soumis au risque.
Taux de morbidité	Nombre d'individus malades divisé par le nombre d'individus dans la population à risque au cours d'une période donnée
Taux de mortalité	Nombre d'individus décédés divisé par le nombre d'individus dans la population à risque au cours d'une période donnée
Taux de létalité	Nombre d'individus décédés divisé par le nombre d'individus malades au cours d'une période donnée
Test virologique	Test diagnostic détectant les particules virales
Test sérologique	Test diagnostic détectant les anticorps dirigés contre un agent pathogène
Titre	Concentration d'un élément au sein d'une solution (exemple : titre virémique : concentration de virus dans le sang)
Virémie	Charge virale dans le sang d'un hôte
Séroconversion	Nom donné à l'augmentation de la concentration sanguine d'anticorps dirigés contre un agent pathogène, signe d'une infection passée due à cet agent pathogène.
Infection	Moment où un agent pathogène entre dans l'organisme de l'hôte
Infecté	Etat d'un organisme porteur d'un agent pathogène. Un organisme infecté peut être symptomatique ou asymptomatique, infectieux ou non-infectieux
Infectieux	Etat d'un organisme infecté capable de transmettre l'agent pathogène dont il est porteur
Phase clinique	Phase de la maladie définie par la présence d'un ou plusieurs symptômes
Période d'incubation	Période entre l'infection d'un hôte par un agent pathogène et l'apparition des premiers signes cliniques
Epidémie	Forme épidémiologique d'une maladie, pour laquelle l'incidence subit de fortes variations au cours de périodes successives, partant d'un niveau faible, voire nul, pour augmenter considérablement (accélération) avant de se stabiliser et de décroître pour parfois s'annuler
Pandémie	Se dit d'une épidémie se propageant sur des très grandes distances et atteignant successivement plusieurs continents
Endémie	Forme épidémiologique d'une maladie, pour laquelle l'incidence (ou le taux d'incidence) reste assez stable au cours de périodes successives (vitesse de transmission régulière quel que soit le niveau). On parle d'endémie à bas bruit si le niveau d'incidence est faible, à haut bruit s'il est fort.
Hyper-endémie	Pour la dengue, forme épidémiologique de la maladie pour laquelle plusieurs des 4 sérotypes circulent simultanément au même endroit.
Zoonose	Maladies et infections qui se transmettent naturellement des animaux vertébrés à l'homme et vice-versa
Maladie infectieuse émergente	Définition aux USA : maladie de nature infectieuse dont l'incidence chez les humains a augmenté au cours des vingt dernières années, ou menace d'augmenter dans un futur proche. Définition au Royaume Uni : entité clinique d'origine infectieuse nouvellement identifiée ou pathologie infectieuse connue dont l'incidence a augmenté dans un endroit donné ou dans un groupe de population donné. Définition en France : maladie infectieuse – ou présumée infectieuse – inattendue (en référence aux propriétés intrinsèques ou à la connaissance que l'on a de la biologie de son agent responsable), touchant l'homme, l'animal ou les deux.

Sensibilité (Se)	Capacité d'un test diagnostic à détecter des évènements positifs. Exemple : proportion d'individus dont le test est positif parmi les individus malades. Cette notion peut être étendue aux modèles : capacité d'un modèle à détecter un évènement positif. Par exemple, proportion d'années pour lesquelles un modèle de risque a prédit un risque élevé parmi les années ayant réellement fait l'objet d'un évènement épidémique majeur.
Spécificité (Sp)	Capacité d'un test diagnostic à détecter des évènements négatifs. Exemple : proportion d'individus dont le test est négatif parmi les individus sains. Cette notion peut être étendue aux modèles : capacité d'un modèle à détecter un évènement négatif. Par exemple, proportion d'années pour lesquelles un modèle de risque a prédit un risque faible parmi les années n'ayant pas enregistré d'évènement épidémique majeur.
Valeur Prédictive Positive (VPP)	Probabilité de la survenue d'un évènement positif sachant qu'un test (ou un modèle) a donné un résultat positif (exemple : probabilité qu'une personne soit malade sachant qu'un test diagnostic a donné un résultat positif, probabilité de survenue d'une épidémie sachant qu'un modèle a prédit la survenue). Correspond à la confiance que l'on a dans le résultat positif d'un test. La valeur prédictive positive d'un test est d'autant plus élevée que la prévalence de la maladie est forte
Valeur Prédictive Négative (VPN)	Probabilité de la survenue d'un évènement négatif sachant qu'un test (ou un modèle) a donné un résultat négatif (exemple : probabilité qu'une personne soit saine sachant qu'un test diagnostic a donné un résultat négatif, (1-probabilité de survenue d'une épidémie sachant qu'un modèle a prédit une absence d'épidémie). Correspond à la confiance que l'on a dans le résultat négatif d'un test. La valeur prédictive négative d'un test est d'autant plus élevée que la prévalence de la maladie est faible
Cas index	Premier cas apparu lors d'une épidémie. Ce n'est pas forcément le premier cas détecté, et une enquête permettant de remonter la chaîne de transmission à partir des premiers cas déclarés (appelée « enquête amont ») est souvent nécessaire.

Entomologie

Compétence vectorielle	Aptitude intrinsèque d'un arthropode vecteur à assurer la transmission biologique et active d'un agent infectieux
Capacité vectorielle	Nombre moyen quotidien de piqures infectantes à partir d'un cas infecté. Elle dépend De la densité vectorielle, de la préférence trophique du vecteur, de sa compétence vectorielle, de la durée du cycle gonotrophique et de la durée d'incubation extrinsèque
Gîte	Lieu de ponte des vecteurs adultes
Imago	Nom donné à la forme adulte de certaines espèces animales invertébrées
Stades pré-imaginaux	Pour les insectes, stades de développement compris entre la ponte des œufs et l'émergence de l'imago (développement larvaire, pupation, développement des nymphes)
Cycle gonotrophique	Cycle de nutrition et de reproduction des moustiques adultes, comprenant la reproduction, le repas sanguin et la ponte.
Indice maison	Pourcentage de maisons infestées (avec gîtes positifs) par une espèce donnée de vecteur
Indice de Breteau	Densité de gîtes positifs par maison
Indice de Breteau pondéré	Indice de Breteau, pondéré par la productivité des gîtes
Indice de Productivité d'Adultes (IPA)	Densité de larves de stade 4 et de nymphes par maison
Indice Pondoir Piège Collant	Nombre de femelles *Ae. aegypti* récoltées par piège
Durée d'incubation extrinsèque	Durée séparant le moment où le moustique s'infecte et le moment où il devient infectieux. La durée d'incubation extrinsèque est étroitement liée à la compétence vectorielle et varie en fonction de la température

Statistiques et modélisation

Variable qualitative	Variable aléatoire dont les valeurs possibles ne sont pas quantifiables par des nombres (exemple : couleur des yeux). Les valeurs possibles prises par une variable aléatoire qualitative sont appelées « modalités »
Variable quantitative	Variable aléatoire dont les valeurs possibles sont des nombres représentant une quantité
Modèle de régression	Modèle statistique analytique permettant de quantifier le lien entre une variable réponse quantitative et une ou plusieurs variables explicatives. Certains modèles de régression permettent également d'effectuer des classifications, comme les modèles logistiques.
Modèle de classification	Modèle statistique analytique permettant de quantifier le lien entre une variable réponse qualitative et une ou plusieurs variables explicatives
Apprentissage automatique	Méthode de modélisation dans laquelle les modèles sont construits de manière automatique au moyen d'une approche essai/erreur /validation.
Modèles supervisés	Modèles pour lesquels le jeu de données à modéliser est constitué de plusieurs variables explicatives ainsi que d'une ou plusieurs variables à expliquer. L'objectif des modèles supervisés est de prédire la valeur de la variable à expliquer à partir des variables explicatives (prédicteurs). Les modèles supervisés peuvent être des modèles de régression ou de classification en fonction de la nature de la variable à expliquer
Modèles non supervisés	Modèles pour lesquels le jeu de données à modéliser est constitué de plusieurs variables servant à caractériser des individus statistiques (pas de variable à expliquer). L'objectif des modèles non supervisés est de regrouper les individus en fonction de caractéristiques communes décrites par les variables
Modèles type SIR	Modèles de dynamique de population utilisés en épidémiologie pour décrire la transmission des maladies d'un hôte à un autre en épidémiologie au moyen d'équations différentielles. Il existe de nombreuses variantes en fonction de l'épidémiologie de la maladie (mode de transmission, durée d'évolution….).

Annexe 2 : Formulaires de déclaration des cas de dengue en Nouvelle-Calédonie

Médecin ou biologiste déclarant (tampon)	*Si notification par un biologiste*
Nom : _____	**Nom :** _____
Établissement / service : _____	Établissement / service : _____
Adresse : _____	Adresse : _____
Tél/Fax : _____	Tél/Fax : _____
Signature :	**Signature :**

DASS Direction des Affaires Sanitaires et Sociales

INSTITUT PASTEUR DE NOUVELLE-CALÉDONIE

Dengue

IMPORTANT : cette maladie justifie une intervention urgente locale, nationale ou internationale. Vous devez la signaler par tout moyen approprié (téléphone, fax…) au médecin inspecteur de la DASS-NC. Cette fiche peut être utilisée pour notifier les cas isolés et les cas groupés. Elle doit être complétée par le déclarant en fonction des informations dont il dispose au moment de la notification, et par la DASS-NC en fonction des données de l'enquête effectuée.

Nom : _____ Prénom : _____ Sexe : M ☐ F ☐ Date de naissance (jj/mm/aaaa) : _____

Adresse exacte : _____ Téléphone : _____

Quartier : _____ Tribu : _____ Commune domicile : _____ Date de notification : _____

LA MENTION D'UNE ADRESSE EST INDISPENSABLE POUR L'INITIATION RAPIDE DE LA LUTTE PERIFOCALE

Autres adresses exactes (professionnelle, scolaire,…) : _____

_____ Téléphone : _____

Contexte épidémiologique :

Cas isolé ☐ Contexte épidémique ☐ précisez _____

Voyage hors NC récent : Non ☐ Oui ☐ où ? _____

Date de retour : _____

Notion d'antécédents de Dengue : Non ☐ Oui ☐ Année : _____ Lieu : _____ Type : _____

Ne sait pas ☐

Clinique :

Date d'apparition des signes : _____ Fièvre : Non ☐ Oui ☐

Signes présents :

Début brutal	☐	Courbatures, myalgies	☐	Céphalées	☐
Nausées/vomissements	☐	Douleurs rétro-orbitaires	☐	Eruption	☐
Signes hémorragiques	☐	Précisez (purpura, épistaxis…) :			
Signes de Choc	☐	Précisez (hypotension, pouls filant…) :			

Autres symptômes (à préciser) : _____

Evolution :

Guérison	Oui ☐	Non ☐	Ne sait pas ☐	
Hospitalisation	Oui ☐	Non ☐	Ne sait pas ☐	Si oui, date d'admission : _____
				date de sortie : _____
Décès	Oui ☐	Non ☐	Ne sait pas ☐	Si oui, date du décès : _____

Prélèvements effectués (minimum 1 tube sec) :

Date de prélèvement : _____ **soit J** ___ *(J1 correspond au premier jour des signes cliniques)*

Précisez : Prélèvement initial ☐ 2ème prélèvement ☐

CADRE À REMPLIR PAR LE LABORATOIRE

Analyses réalisées :

Profil précoce :	ARN Viral (PCR)	(J1 à J5)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
Profil tardif :	Antigène NS1 strip	(J1 à J5)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
	Antigène NS1 ELISA	(J2 à J8)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
	IgM Dengue	(> J4)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐

Hématologie : GB : _____ Plaq : _____ Hte : ___,__

Conclusion LABO : Cas cas suspect ☐ cas probable ☐ cas confirmé ☐ DH ☐ DSC ☐

Virus Type 1 ☐ **Type 2** ☐ **Type 3** ☐ **Type 4** ☐ **Non typé** ☐

Maladie à déclaration obligatoire (délibération 425 du 26 novembre 2008) - Cette fiche fait l'objet d'un traitement informatique automatisé déclaré à la CNIL.
Conformément à la loi du 6 janvier 1978, le patient ou le médecin déclarant ont le droit d'accès et de rectification immédiat au fichier par demande directe au médecin inspecteur de la DASS-NC,
5 rue du Général Galliéni, BP M 9885.1 NOUMEA Cedex. Fiche à transmettre au médecin inspecteur de la DASS-NC, 5 rue du Général Galliéni, BP M 9885.1 NOUMEA Cedex - Tél : 24 37 00 - Fax : 05 11 33 - Numéro vert : 05 11 03

Médecin ou biologiste déclarant (tampon)	***Si notification par un biologiste***
Nom :	**Nom :**
Établissement / service :	Établissement / service :
Adresse :	Adresse :
Tél/Fax :	Tél/Fax :
Signature :	**Signature :**

DASS
Direction des Affaires Sanitaires et Sociales

INSTITUT PASTEUR DE NOUVELLE-CALÉDONIE

Dengue
Réseau sentinelle

IMPORTANT : cette maladie justifie une intervention urgente locale, nationale ou internationale. Vous devez la signaler par tout moyen approprié (téléphone, fax…) au médecin inspecteur de la DASS-NC. Cette fiche peut être utilisée pour notifier les cas isolés et les cas groupés. Elle doit être complétée par le déclarant en fonction des informations dont il dispose au moment de la notification, et par la DASS-NC en fonction des données de l'enquête effectuée.

Nom : _____ Prénom : _____ Sexe : M ☐ F ☐ Date de naissance (jj/mm/aaaa) : __/__/____

Adresse exacte : _____ Téléphone : _____

Quartier : _____ Tribu : _____ Commune domicile : _____ Date de notification : __/__/____

LA MENTION D'UNE ADRESSE EST INDISPENSABLE POUR L'INITIATION RAPIDE DE LA LUTTE PERIFOCALE

Autres adresses exactes (professionnelle, scolaire,…) : _____

_____ Téléphone : _____

Contexte épidémiologique :

Cas isolé ☐ Contexte épidémique ☐ précisez _____

Voyage hors NC récent : Non ☐ Oui ☐ où ? _____

Date de retour : __/__/____

Notion d'antécédents de Dengue : Non ☐ Oui ☐ Année : ____ Lieu : _____ Type : ____

Ne sait pas ☐

Clinique :

Date d'apparition des signes : __/__/____ Fièvre : Non ☐ Oui ☐

Signes présents :

Début brutal	☐	Courbatures, myalgies	☐	Céphalées	☐
Nausées/vomissements	☐	Douleurs rétro-orbitaires	☐	Eruption	☐
Signes hémorragiques	☐	Précisez (purpura, épistaxis…) :			
Signes de Choc	☐	Précisez (hypotension, pouls filant…) :			

Autres symptômes (à préciser) : _____

Evolution :

Guérison	Oui ☐	Non ☐	Ne sait pas ☐	
Hospitalisation	Oui ☐	Non ☐	Ne sait pas ☐	Si oui, date d'admission : __/__/____
				date de sortie : __/__/____
Décès	Oui ☐	Non ☐	Ne sait pas ☐	Si oui, date du décès : __/__/____

Prélèvements effectués (minimum 1 tube sec) :

Date de prélèvement : __/__/____ **soit J** ____ *(J1 correspond au premier jour des signes cliniques)*

Précisez : Prélèvement initial ☐ 2ème prélèvement ☐

CADRE À REMPLIR PAR LE LABORATOIRE

Analyses réalisées :

Profil précoce :	ARN Viral (PCR)	(J1 à J5)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
Profil tardif :	Antigène NS1 strip	(J1 à J5)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
	Antigène NS1 ELISA	(J2 à J8)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐
	IgM Dengue	(> J4)	Résultat :	Positif ☐	Négatif ☐	Douteux ☐

Hématologie : GB : _____ Plaq : _____ Hte : ____,__

Conclusion LABO : Cas cas suspect ☐ cas probable ☐ cas confirmé ☐ DH ☐ DSC ☐

Virus **Type 1** ☐ **Type 2** ☐ **Type 3** ☐ **Type 4** ☐ **Non typé** ☐

Maladie à déclaration obligatoire (délibération 423 du 26 novembre 2008) - Cette fiche fait l'objet d'un traitement informatique automatisé déclaré à la CNIL.
Conformément à la loi du 6 janvier 1978, le patient ou le médecin déclarant ont le droit d'accès et de rectification immédiat au fichier par demande directe au médecin inspecteur de la DASS-NC.
5 rue du Général Galliéni, BP M4 98851 NOUMEA Cedex. Fiche à transmettre au médecin inspecteur de la DASS-NC, 5 rue du Général Galliéni, BP M4 98851 NOUMEA Cedex - Tél : 24 37 00 - Fax. 05 11 33 - Numéro vert. 05 11 03

Annexe 3 : Liste des champs disponibles dans les bases de données mises à disposition par la DASS-NC

Base de données	Nom du champ	Description et codage
1995 - 2005	Adresse	Adresse du lieu d'habitation du patient. Non standardisé.
	Quartier	Référencés uniquement sur 3 communes : Nouméa, Dumbéa, Mont-Dore. Les codes sont ceux utilisés par l'ISEE.
	Communeder	Commune de résidence des patients. Les codes sont ceux utilisés par l'ISEE (ils sont différents des codes postaux utilisés par l'OPT). Code à 2 chiffres (à homogénéiser avec la base 2008-2011).
	Proresi	Province de résidence du patient (N=Nord, S=Sud, I=Iles).
	Etablissem	Code de l'établissement déclarant le cas.
	Provincens	Province où travaille le médecin ou l'établissement ayant déclaré. T = CHT
	Sexe	1 = homme 2 = femme
	Ethnie	"Mélanésien", "européen", "asiatique",… correspondent aux noms utilisés par l'ISEE.
	Agemois	Age en mois
	Agean	Age en années
	Dateconsul	Date de consultation
	Reseau	Y = le médecin appartient au réseau sentinelle N = le médecin déclarant n'appartient pas au réseau sentinelle
	Declartme	Déclaration du médecin : Y = le médecin a déclaré lui-même le cas (fax, téléphone…) ; N = les informations ont été obtenues autrement (via les laboratoires comme l'IPNC)
	ConfirmeIP	Faux= suspicion clinique uniquement, non confirmée biologiquement Vrai = le cas été confirmé biologiquement par l'IPNC A homogénéiser avec les champs "cas_clin" et "type" de la base 2008-2011
	OMS	Code de la maladie en CIM10 (Classification Internationale des Maladies) de l'OMS
	Serotype	Sérotype identifié. D'après Nancy Klingler, l'information n'est fiable qu'à partir de fin 1996. En 1995, tous les cas ont été assimilés comme étant dus au sérotype 3.
2006 - 2007	Adresse	Adresse du lieu d'habitation des patients. Non standardisé.
	Quartier	Référencés uniquement sur 3 communes : Nouméa, Dumbéa, Mont-Dore. Les codes sont ceux utilisés par l'ISEE.
	Lotissement	?
	Communeder	Commune de résidence des patients. Les codes sont ceux de l'ISEE (ils sont différents des codes postaux utilisés par l'OPT)
	Proresi	Province de résidence (N=Nord, S=Sud, I=Iles)
	Etablissem	Code de l'établissement déclarant le cas
	Prescripte	?
	Declar	?
	Nordre	?
	Provincens	Province où travaille le médecin ou l'établissement ayant déclaré. T = CHT
	Sexe	1 = homme 2 = femme
	Ethnie	"Mélanésien", "européen", "asiatique",… correspondent aux codes de l'ISEE
	Datenaiss	Date de naissance; à homogénéiser avec les deux champs suivant pour n'en conserver qu'un (âge au moment des symptômes par exemple)
	Agemois	Age en mois. Pour les < 1 an. A homogénéiser avec date de naissance
	Agean	Age en années (>1 an). A homogénéiser avec date de naissance
	Date1erSS	Date des premiers signes cliniques

	Dateconsul	Date de consultation
	Datedeclar	Date à laquelle le médecin a déclaré. Permet de mesurer la réactivité des médecins (utile uniquement pour l'évaluation du réseau de surveillance)
	Datesortie	Date de sortie de l'hôpital pour ceux qui ont été hospitalisés
	ConfirmeIP	Faux= suspicion clinique uniquement (pas de test labo) Vrai = Cas clinique dont le test labo est + (confirmé) A homogénéiser avec les champs "cas_clin" et "type" de la base 2008
	Serotype	Sérotype indiqué que pour les patients dont un échantillon a été envoyé à l'IPNC. 9 = inconnu (non testé)
	NumIPNC	Numéro du patient dans la base de données de l'IPNC
	Remarque	Destination d'un voyage récent pour certains cas. Problème : comme cette information n'est pas récoltée de manière systématique, il n'est pas possible de savoir si les patients pour lesquels rien n'est marqué ont fait un voyage récent ou non, et donc les quelques données notées sont inutilisables.
	Typecas	Pas de donnée
	Voyage	Pas de donné
	Reseau	Le médecin appartient-t-il au réseau sentinelle ?
	Declartme	Déclaration du médecin : Y = le médecin a déclaré lui-même le cas (fax, téléphone..;) ; N = les infos ont été obtenues autrement (via les labos comme l'IPNC)
	Decede	Issue de la maladie
	OMS	Code de la maladie en CIM10 (Classification Internationale des Maladies) de l'OMS
	Intervmun	???
2008 – 2012	Adresse	Nous ne l'avons pas dans la feuille Excel qu'ils nous ont donné, mais ils doivent l'avoir quelque part, nécessairement, pour la lutte?
	Numero	Identifiant dans la base
	cas_clin	VRAI = cas suspect, clinique, non confirmé FAUX = cas confirmé par sérologie +/- typage A homogénéiser avec le champ "confirmIP" des bases 1995-2005 et 2006-2007
	Date	Date de prélèvement
	sexe	1 = homme 2 = femme
	datenaiss	Date de naissance. Age au moment des symptômes à calculer
	age	Age (en années). 0 pour les moins d'un an
	code_com	Code commune de l'ISEE. Différent de celui utilisé par l'OPT pour le courrier. Code à 5 chiffres, il faut extraire les deux derniers pour retrouver le code des bases précédentes.
	code_quart	Référencés uniquement sur 3 communes : Nouméa, Dumbéa, Mont-Dore. Pour les autres communes, cf géo-référencement à l'échelle de la tribu. Les codes quartiers utilisés dans cette base sont les codes identifiant des polygones du *shapefile* utilisé par la DASS-NC pour la création de cartes.
	tribu_t	Nom de la tribu (ou du quartier pour les trois communes du Grand Nouméa. Les noms sont standardisés et sont les même que ceux utilisé par l'ISEE.
	tribu	Numéro ISEE de la tribu, quand connu
	Voyage_recup	Les noms apparaissant pour les voyages sont standardisés. Par contre : "-" ou " " : on ne sait pas si cela correspond à "pas de voyage" ou "non renseigné sur la fiche". Donc on n'a aucun moyen d'être sûr que ces patients n'ont pas voyagé dans le passé récent. Donc difficile de séparer cas autochtone de cas importé (on ne peut avoir qu'une idée globale en allant regarder les adresses, pour voir sis plusieurs cas sont survenus au même endroit, ou à deux maisons à côté.)
	Voyage	
	Type	1: cas possible (cas clinique) 2: cas probable (cas clinique + une sérologie positive, mais non confirmé en NS1 ou en PCR)

		3: cas confirmé (cas clinique positif en PCR ou bien positifs en IgM 2 fois de suite) A homogénéiser avec les champs "cas_clin" de la base 2008-2011, et le champ "confirmIP" des bases d'avant
	Serotype	1, 2, 3 ou 4. Pas de code pour le sérotype inconnu
	Prel_J	??
	NS1_Strip	1 : positif 2: négatif 3: douteux 4: positif faible
	NS1_Elisa	1 : positif 2: négatif 3: douteux 4: positif faible
	PCR	1 : positif 2: négatif 3: douteux 4: positif faible
	IgM_Pan_Pen	1 : positif 2: négatif 3: douteux 4: positif faible
	Clin	Pas de données
	Plaq	?? J'imagine la numération des plaquettes dans le sang
	Clin_mya	VRAI = Présence de myalgie
	Clin_fie	VRAI = Présence de fièvre
	Clin_cep	VRAI = Présence de céphalée
	Clin_nau	VRAI = Présence de nausée
	Clin_thr	VRAI = Présence de ?? thrombocytopénie ?
	Clin_eru	VRAI = Présence de ?? éruption cutanée ?
	Clin_gen	VRAI = Présence de ??
	Clin_epi	VRAI = Présence de ??
	Clin_hem	VRAI = Présence de ?? signes hémorragiques ?
	Clin_pur	VRAI = Présence de purpura
	Epid	Compilation des colonnes précédentes par Luc pour savoir si il y a eu présence de signes hémorragiques ou non
	Hemo	
	Hemo_final	
	Detail	
	Com_lib	Libellé de la commune
1995 – 2012 (base complète homogénéisée)	Adresse	Adresse du lieu d'habitation du patient. Non standardisé
	Quartier_2008	Référencés uniquement sur 3 communes : Nouméa, Dumbéa, Mont-Dore. Les codes quartiers utilisés dans cette base sont les codes identifiant des polygones du *shapefile* utilisé par la DASS-NC pour la création de cartes.
	Communeder	Code commune de l'ISEE. Différent de celui utilisé par l'OPT pour le courrier. 99 : commune de résidence inconnue
	Sexe	1 = homme 2 = femme
	Agean	Age (en années). 0 pour les moins d'un an
	Dateconsul	Date de consultation
	Cas_clin	VRAI = cas suspect, clinique, non confirmé FAUX = cas confirmé par sérologie +/- typage
	Type	1: cas possible (cas clinique)

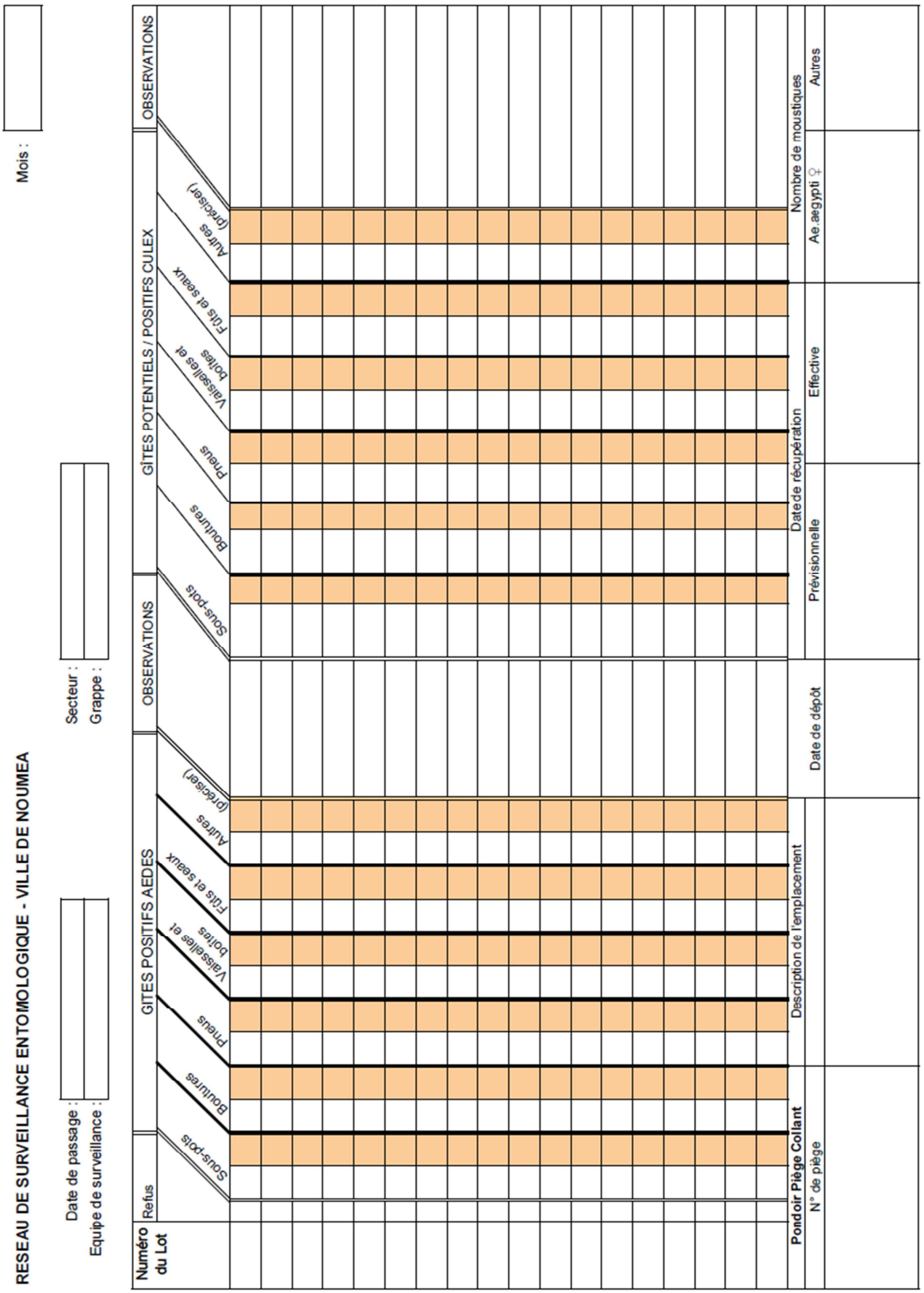
		2: cas probable (cas clinique + une sérologie positivie, mais non confirmé en NS1 ou en PCR) 3: cas confirmé (cas clinique positif en PCR ou bien positifs en IgM 2 fois de suite)
	Serotype	1, 2, 3 ou 4
	Voyage	
	Adresse_2	Adresse du lieu d'habitation du patient. Non standardisé
	Tribu	Numéro ISEE de la tribu, quand connu
	Tribu_t	Nom de la tribu (ou du quartier pour les trois communes du Grand Nouméa. Les noms sont standardisés et sont les même que ceux utilisé par l'ISEE.

Annexe 4 : Liste des codes quartier utilisés dans les différents fichiers

Nom du quartier	Code bases 1995-2005 et 2006-2007	Code bases 2008-2012 et 1995-2012 homogénéisée	Code « shapefile » de l'IRD
Koutio	502	502	4
Auteuil	503	503	5
Coeur de Ville	501	501	95
Dumbéa sur mer	504	504	96
Plaine Adam	0	505	2
Nakutakoin	509	506	97
Les Koghis	508	507	3
Plaine de Koé	511	508	1
Katiramona Sud	507	509	71
Nondoué - La Couvelée	510	510	94
Dzumac Montagne des Sources	0	512	72
Yahoué	1709	1701	6
Pont des Français	1705	1702	7
Conception	1702	1703	8
Robinson	1706	1704	9
Boulari	1701	1705	10
Saint Michel	1701	1706	11
Saint Louis	1707	1707	12
La Coulée	1703	1708	13
La Lembi-Mouirange	0	1709	17
Vallon Dore	1708	1710	14
Mont Dore Sud	0	1711	15
Plum	1704	1712	16
Ile Ouen	1704	1713	19
Le Grand Sud	508	1714	18
Val Plaisance	1829	1801	30
Anse Vata	1802	1802	52
Baie des Citrons	1804	1803	51
NGéa	1816	1804	57
Motor Pool	1815	1805	55
Receiving	1815	1806	56
Orphelinat	1819	1807	54
Trianon	1828	1808	50
Faubourg Blanchot	1808	1809	37
Artillerie	1803	1810	48
Quartier Latin	1825	1811	49
Vallée des Colons	1830	1812	36
Magenta	1812	1813	38
Ouémo	1820	1814	60
Aérodrome	1801	1815	46
Portes de Fer	1824	1816	43
Haut Magenta	1809	1817	39
Vallée du Génie	1831	1818	59
Centre Ville	1805	1819	42
Vallée du Tir	1832	1820	40
Doniambo	1806	1821	61
Nouville	1810	1822	68
Montagne Coupée	1813	1823	45
Montravel	1814	1824	44
PK4	1821	1825	47
PK6	1822	1826	41

Tina	1822	1827	32
Normandie	1817	1828	70
PK7	1823	1829	31
Rivière Salée	1826	1830	69
ZI Ducos	0	1831	66
Ducos	1807	1832	62
Logicoop	1811	1833	64
Tindu	1827	1834	53
Numbo	1818	1835	63
Koumourou	0	1836	67
Kaméré	1811	1837	65
Ile Sainte Marie	0	1890	29
Ilot Uéré	0	1891	34
Ilot Brun	0	1893	58
Ilot Sainte Marie	0	1894	35

Annexe 5 : Formulaire utilisé pour la collecte des données entomologiques à Nouméa



Les colonnes blanches sont faites pour noter les informations concernant les larves de stade 4, et les colonnes colorées les informations concernant les nymphes.

Annexe 6 : Article 2 : étude des processus de propagation de la dengue au Cambodge par la technique d'analyse en ondelette

Can Human Movements Explain Heterogeneous Propagation of Dengue Fever in Cambodia?

Magali Teurlai[1,2], Rekol Huy[3], Bernard Cazelles[4,5], Raphaël Duboz[6], Christophe Baehr[7,8], Sirenda Vong[1]*

1 Epidemiology and Public Health Unit, Institut Pasteur du Cambodge, Phnom Penh, Cambodia, **2** IRD UMR LOCEAN, UMR ESPACE-DEV, New-Caledonia, France, **3** National Dengue Control Program, National Centre for Parasitology, Entomology and Malaria Control, Ministry of Health, Phnom Penh, Cambodia, **4** Ecologie & Evolution, UMR 7625, CNRS-UPMC-ENS, Paris, France, **5** UMMISCO UMI 209 IRD - UPMC, Bondy, France, **6** CIRAD UPR Agirs, Montpellier, France, **7** Météo France, CNRM, Toulouse, France, **8** CNRS, GAME URA 1357, Toulouse, France

Abstract

Background: Determining the factors underlying the long-range spatial spread of infectious diseases is a key issue regarding their control. Dengue is the most important arboviral disease worldwide and a major public health problem in tropical areas. However the determinants shaping its dynamics at a national scale remain poorly understood. Here we describe the spatial-temporal pattern of propagation of annual epidemics in Cambodia and discuss the role that human movements play in the observed pattern.

Methods and Findings: We used wavelet phase analysis to analyse time-series data of 105,598 hospitalized cases reported between 2002 and 2008 in the 135 (/180) most populous districts in Cambodia. We reveal spatial heterogeneity in the propagation of the annual epidemic. Each year, epidemics are highly synchronous over a large geographic area along the busiest national road of the country whereas travelling waves emanate from a few rural areas and move slowly along the Mekong River at a speed of ~11 km per week (95% confidence interval 3–18 km per week) towards the capital, Phnom Penh.

Conclusions: We suggest human movements – using roads as a surrogate – play a major role in the spread of dengue fever at a national scale. These findings constitute a new starting point in the understanding of the processes driving dengue spread.

Citation: Teurlai M, Huy R, Cazelles B, Duboz R, Baehr C, et al. (2012) Can Human Movements Explain Heterogeneous Propagation of Dengue Fever in Cambodia? PLoS Negl Trop Dis 6(12): e1957. doi:10.1371/journal.pntd.0001957

Editor: Duane J. Gubler, Duke University-National University of Singapore, Singapore

Received May 13, 2012; **Accepted** October 29, 2012; **Published** December 6, 2012

Copyright: © 2012 Teurlai et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: MT was supported financially by the Fondation Pierre Ledoux-Jeunesse Internationale and the International Division of Institut Pasteur. BC is supported by the FP7 Cooperation Programme from the European Community, DenFREE (FP7-HEALTH-2011-282378). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vongs@wpro.who.int

Introduction

Cambodia is a low-income tropical country, hyper-endemic for all four serotypes of dengue infection. As such, dengue epidemics occur every year during the rainy season and result in considerable morbidity and economic burden. The basis for these recurrent epidemics is an increased vector activity during the rainy season and complex interactions between hosts and viruses with short lived cross-protective immunity [1–3]. In the absence of a vaccine, control is limited to vector control measures. Regarding dengue dynamics, locally, dengue outbreaks are explosive and tend to be focal, perhaps reflecting the limited dispersal of the vector, which visits few houses in a life-time, and have a limited flight range [4,5]. At an international scale, human movement is known to be a major factor responsible for the virus transportation among big urban centres [6]. At a national scale, little is known about the spatial propagation of the disease. In many endemic countries, urban centres are thought to act as a reservoir of the virus from where it can spread to the rest of the country [6–9]. In Thailand [7] and more recently in Southern Vietnam [8], researchers demonstrated the existence of a travelling wave either within the 3-year periodic mode or the annual mode of oscillation. However, the underlying factors responsible for these waves remain unknown [1,7–10]. Hypotheses include immunological host-virus interactions, differences in virus virulence, or heterogeneity of the spatial distribution of the host population [1]. Synchronisation of cases has also been observed in the 3 year periodic band in Thailand [7]. Mechanisms responsible for higher synchronicity could include climatic forcing [8,10].

The objective of this study was to determine whether there was a spatial-temporal pattern of dengue propagation in Cambodia repeating year after year. The characterisation of such patterns is important to understand the forces driving dengue spatial spread and aid better control and logical allocation of public health resources.

Methods

Study area

Cambodia is a low-income country located in South-East Asia, divided into 24 provinces and 180 districts covering 181,035 km^2. Out of the 13.4 million people in 2008, more than 80% live in

Author Summary

Dengue fever is a mosquito borne viral infection. It has become a major public health problem during the past decades: only 9 countries were affected in the 1970s; dengue is now endemic in more than 100 countries. In the absence of any vaccine or specific treatment, control of dengue fever is currently limited to vector control measures, which are difficult to implement and hardly sustainable, especially in low income countries. To implement efficient control measures, it is crucial to understand the dynamics of propagation of the disease and the key factors underlying these dynamics. In this study, data from 8-year national surveillance in Cambodia were analysed. Dengue fever follows a recurrent pattern of propagation at the national scale. The annual epidemics originate from a few rural areas identified in this work. This study also suggests additional evidence for the role of human movement in the spatial dynamics of the disease, which should be accounted for in control measures. These results differ from the current knowledge about dengue dynamics and are therefore of interest for future research.

rural areas; 1.3 million people (9.9%) live in Phnom Penh, the capital city. Cambodia is lagging behind other countries in South-East Asia in terms of economic or demographic development. For example, unlike Thailand, the demographic transition has not occurred yet. The South-West and the North of the country (see grey districts in Figure 1) are mountainous regions with low population density (mean district density of 18 people per km^2). The rest of the country is composed of flat plains with few cities (mainly Phnom Penh, Kampong Cham, Siem Reap and Battambang, see Figure 1) scattered among rural areas. The weather is warm all year long, and the climate is dominated by the annual monsoon cycle, with a dry season (December to April) alternating with a wet season (May to November). Climatic variation from one area to another is limited. Temperature is homogeneous across the country and ranges annually from 21°C to 35°C. As a result, mosquitoes can be active all year long when considering only temperature. Rainfall or water availability are more likely to be the factors limiting the vector's activity.

The data

Cambodian National surveillance recorded 109,332 dengue cases during 2002–2008, a period over which the reporting process was stable. Cases were reported passively from public hospitals and actively from 5 major sentinel hospitals located in the cities of Siem Reap (1 hospital), Kampong Cham (1 hospital) and Phnom Penh (3 hospitals). Cases were clinically diagnosed using the 1997 World Health Organization (WHO) case definitions, allowing clinical and paraclinical (haematocrit and platelets count) distinction between classic dengue fever, dengue haemorrhagic fever and dengue shock syndrome [2]. Of note, the new WHO case definition was only introduced in 2009 [11]. To increase specificity, only severe cases (e.g. dengue haemorrhagic fever) affecting children less than 16 years old were recorded in the database. Assuming that epidemic patterns of dengue would be stochastic in low population density areas, we excluded the 45 (/180) districts with less than 20 people per km^2 from the analysis, dismissing 3298 declared cases (3.03% of the total declared cases). Since patients' districts of residence were recorded, we calculated weekly incidence rates for each of the 135 remaining districts. Denominators for incidence rates were interpolated linearly using the 1998 and 2008 national censuses. Based on age distribution

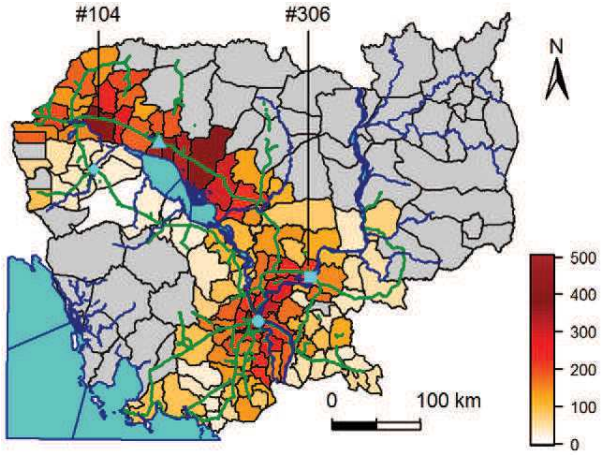


Figure 1. Map of mean annual dengue fever incidence rates in districts of Cambodia. Mean annual incidence rates (in number of cases declared per 100,000 people per year) are calculated over 2002 to 2008, for districts with more than 20 people per km^2. Cambodia is surrounded by the Indian Ocean (bottom left), Thailand (West), Lao (North) and Vietnam (East and South-East). Phnom Penh, the capital, is represented by a circle, Siem Reap by a triangle, Kampong Cham by a square and Battambang by a lozenge. Blue lines represent the Mekong River, going north to south, and the Tonle Sap River linking the Tonle Sap central Lake to the Mekong River. Green lines represent national roads. Grey districts have less than 20 people per km^2.
doi:10.1371/journal.pntd.0001957.g001

similarities between provinces, and on similarities between 1998 and 2008 age structures, we assumed that age structure was homogeneous over the country and have not standardised incidence rates according to age. As national surveillance data were made available to be utilised for such temporal and spatial analyses and have been routinely published nationally and once internationally [2] no specific approval was requested from the Ministry of Health's National Ethics Committee. Moreover, data that were provided by the National Programme were anonymised prior to transfer to the Institut Pasteur du Cambodge. We subsequently randomly assigned new codes to each record and deleted the previous ones to unlink the present dataset to the national database.

Temporal analysis

Time series of dengue incidence were square-root transformed to stabilise the variance, subsequently centred to zero-mean and normalised to unit variance. We then performed a wavelet transform and used wavelet phase analysis to describe dynamic patterns of dengue. This spectral method is well-suited for the analysis of non-stationary time-series such as epidemic curves. It is particularly useful to filter the data in any given frequency band and extract the phase of any given periodic component [10,12–14] (see Figure 2 for an illustration of the method).

The wavelet transform was done using R software [15] and functions translated from Cazelles' Matlab toolbox. We used a Morlet wavelet as the mother wavelet. All equations used and vocabulary relative to wavelet analysis are detailed in [10,12–14].

The wavelet coefficients corresponding to a period ranging from 0.8 to 1.2 year were used to reconstruct filtered time series corresponding to the annual epidemic in each district. These filtered time series, called "annual component of incidence", are illustrated in Figure 2B.

The phases of the epidemic have been computed in the periodic band 0.8–1.2 year (see equation (5) in [10]), thus obtaining, for

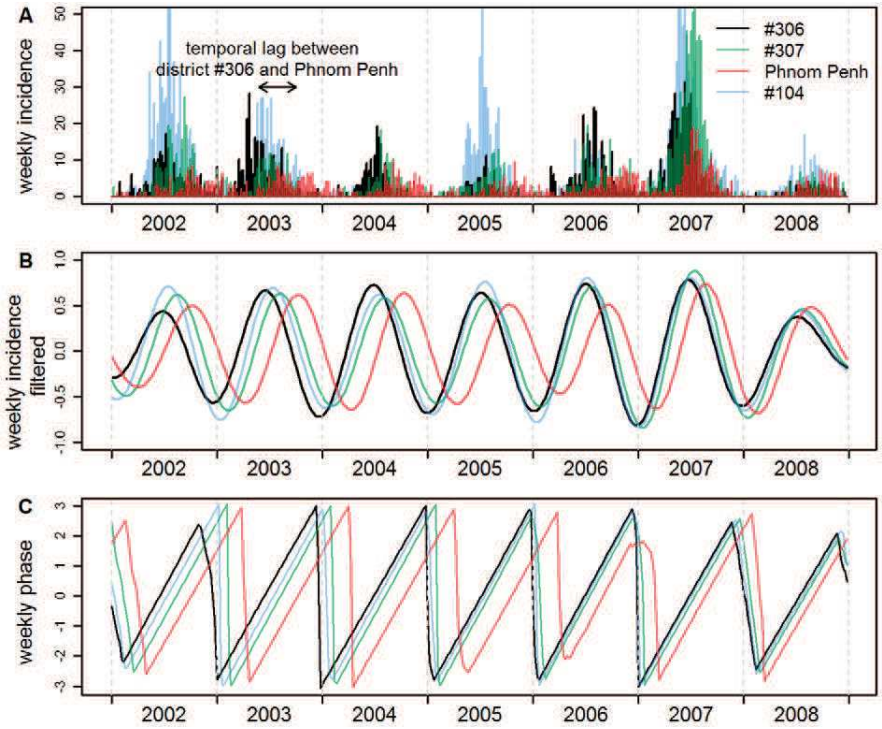


Figure 2. Weekly raw incidence rates, filtered incidence and phase of four districts in Cambodia. The four districts are: district #306 (black), a rural district located around Kampong Cham; Phnom Penh (red); District #307 (green), a rural district located mid-way between #306 and Phnom Penh; District #104 (blue). (A) Weekly raw incidence rates (in number of cases declared per 100,000 people per week). (B) Annual component of incidence, obtained by filtering raw weekly incidence in the 0.8–1.2 year periodic band using wavelet analysis. (C) Phase of the annual component of incidence, computed in the 0.8–1.2 periodic band using wavelet analysis (see Methods).
doi:10.1371/journal.pntd.0001957.g002

each district, a single time series of the phase of the cyclic annual component of the epidemic (Figure 2C). This phase can be viewed as the timing of the epidemics, and is almost not influenced by the intrinsic value of incidence. For a given week, if the seasonal epidemic occurs at the same time in two districts, the two annual components have the same phase. Another example in Figure 2, district #306 (a rural district located around Kampong Cham) has an advanced phase compared to Phnom Penh: the epidemic in Phnom Penh is lagging behind the one in district #306.

By calculating phase differences, one can then determine in which order districts are affected by the annual epidemic. This ranking allowed us to identify districts hit early by the seasonal epidemic. Time series of the temporal lag between seasonal epidemics at different locations were thus estimated from phase differences, according to equations (6)–(8) in [10]. This temporal lag is proportional to the phase difference.

In some districts identified as early districts, there were not enough cases reported to be confident that there was significant dengue transmission ongoing, given that cases are declared on a clinical basis. Therefore, for each district, we identified epidemic years using the national epidemic threshold [2,16]. The national weekly threshold [2] for an epidemic was calculated as two standard deviations above a three weeks sliding mean computed over the weekly national incidence rates of the five past available non epidemic years [16] (from 2002 to 2006 in our study). We excluded from the analysis time periods from districts with low incidence, only including epidemic years defined as districts with a weekly incidence rate from January to December above the national threshold for an epidemic during two consecutive weeks. Subsequent analyses were restricted to the years with detected epidemics.

Spatial analysis of dengue propagation

For a given year, to evaluate the speed at which dengue epidemic propagates along a given geographical axis comprised of I districts, we performed a regression analysis: $Y_i = \beta 0 + \beta 1\ X1_i + \varepsilon_i$, with i in [1, I], Y_i the annual mean of the temporal lag between district i and a district located along the axis and selected as a time reference to compute temporal lags, and $X1_i$ the corresponding distance as the crow flies, in km, between the centres of district i and the reference district. The speed of dengue propagation along the axis was estimated as the inverse of the regression slope $\beta 1$.

To compare the dynamic pattern along J different axes, we performed, each year, an analysis of covariance [17]: $Y_{ij} = \beta 0_j + \beta 1_j\ X1_{ij} + \varepsilon_{ij}$, with i in [1, I_j], j in [1, J], Y_{ij} the annual mean of the temporal lag between district i of axis j and a district located along the axis j and chosen as a time reference, $X1_{ij}$ the corresponding distance separating the centres of districts i and the reference district on axis j, and $\beta 0_j$ and $\beta 1_j$ the intercept and slope of the regression line of axis j. We will call "p-value of interaction" the p-value of the hypothesis that the model slopes $\beta 1_j$ are all equal, or in other words, that the speed of propagation is the same along different axes. If the model gave a p-value of interaction below 0.05, the propagation was considered heterogeneous along the different axes, and separate regressions were performed for each axis.

As we wanted to know whether the results were consistent from year to year, we performed this analysis each year.

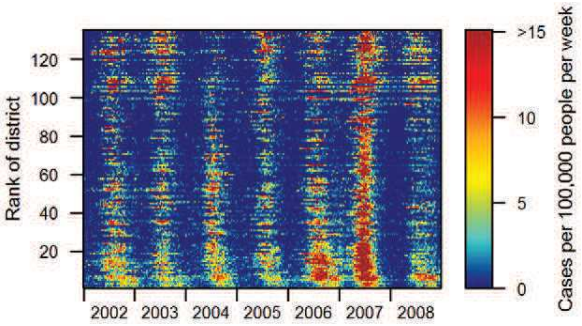


Figure 3. Apparent dengue haemorrhagic fever weekly incidence rates in the 135 most populous districts of Cambodia. Weekly incidence rates (cases per 100,000 people per week) were computed in each of the 135 districts where population density is higher than 20 people per km^2 in Cambodia. Districts are ranked by increasing distance to Phnom Penh from bottom to top.
doi:10.1371/journal.pntd.0001957.g003

All analyses and figures were performed using R software [15].

All confidence intervals (C.I.) provided were calculated using classical methods for calculating a confidence interval around a mean, unless stated otherwise.

Results

Figure 3 shows the weekly incidence in the 135 (/180) most populous districts (average district population of 78,000), where 105,598 of the 109,332 reported patients reside. Annual epidemics do not appear synchronous, peaking at different times of the year in different districts (between May and October). Unexpectedly, in Phnom Penh, the capital, where the virus circulates during the dry season [2], the annual epidemic lags behind the one in other districts indicating that Phnom Penh is not the starting point. Inspection of weekly and mean annual incidence maps over the whole period (not shown here, but partly visible in Figure 1, Movie S1, and Figure S1) showed two possible geographic axes seminal in the propagation of dengue: the national road between Kampong Cham and Siem Reap – the busiest one in Cambodia –, and the Mekong River.

To characterise and compare the spatio-temporal pattern of dengue incidence along those two axes and determine the focal starting areas of the annual epidemic, we performed a phase analysis using a wavelet approach (see Methods).

The examination of wavelet power spectra (Figure S2) revealed the annual seasonal component of incidence is the most powerful in 117/135 districts, meaning that in Cambodia, the seasonal cycle of dengue incidence time series has more power than the inter-annual cycle. A significant 2–3 years periodic component was detected in some districts, but our time series were too short in time to study it.

Figure 4 represents, in a time-space domain, the phase of the annual component of incidence filtered in the 0.8–1.2 year periodic band for each district of the two geographic features identified previously (see Movie S2 for maps of these phases in all 135 districts). White parts represent years when no epidemic was detected in the district (see Methods). The phase represents the timing of the epidemic in each district. For a given week, if all districts have the same phase (same colour), the epidemics occur at the same time.

First, the pattern of spatial synchronicity observed is consistent from year to year. To test this result, each year, we ranked the 135 districts according to their phase during the epidemic period

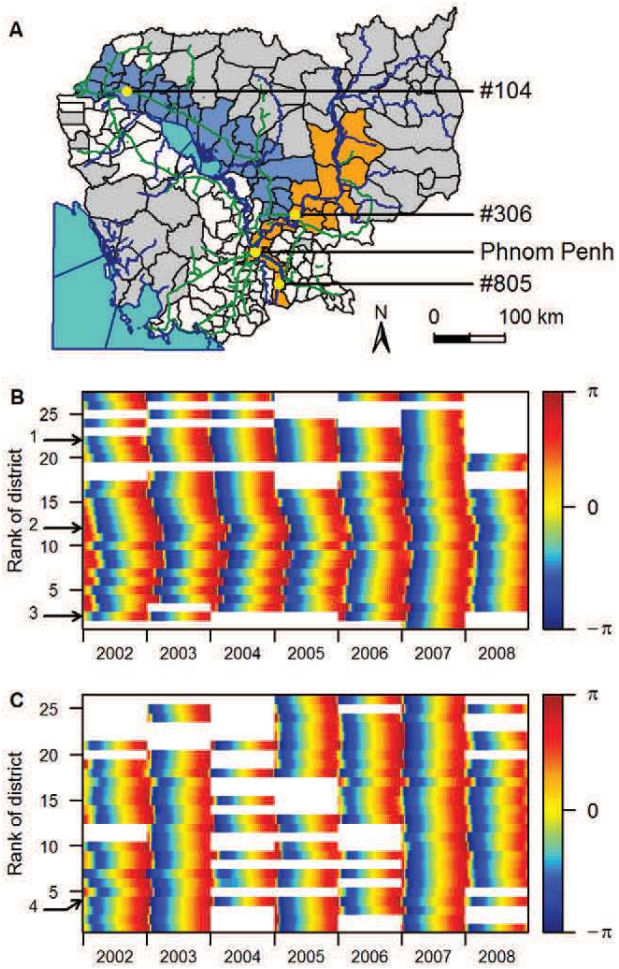


Figure 4. Phases of the annual component of incidence for districts located along two geographic axes. Phases are computed in the 0.8–1.2 year periodic band. (A) Map of the two geographic areas chosen: the national road in blue, and the Mekong River in orange. (B) Phase of districts along the Mekong River (orange in Figure 4A), presented from the most southerly to the most northerly from bottom to top. (C) Phase of districts along the national road (blue in Figure 4A), presented from West to East from bottom to top. The arrows indicate districts: 1, #306; 2, Phnom Penh; 3, #805 (Figure 4B); 4, #104 (Figure 4C).
doi:10.1371/journal.pntd.0001957.g004

(weeks 13 to 39), from the district where the epidemic has the most advanced phase to the one with the highest phase delay. We then performed a Spearman correlation test on these ranks, comparing ranks from year n with ranks of years n+1. Except in 2007, correlation coefficients were all higher than 0.38 and significant (all p-values$<$0.03), inferring that districts are affected in a similar chronological order year after year. In 2007, the order in which districts were affected by the epidemic was not significantly correlated to the order of the previous year.

This ranking also allowed us, each year, to identify districts where the annual epidemic appears early. We have identified three starting points, all located in rural areas: district #306 and a few rural districts around (district #306 being the most early of them), district #104 and 2 other districts around, and, some years, district #805, located along the Mekong River, South to Phnom Penh, along the Vietnamese border (see Figure 4). These districts are very similar to other rural districts included in the analysis,

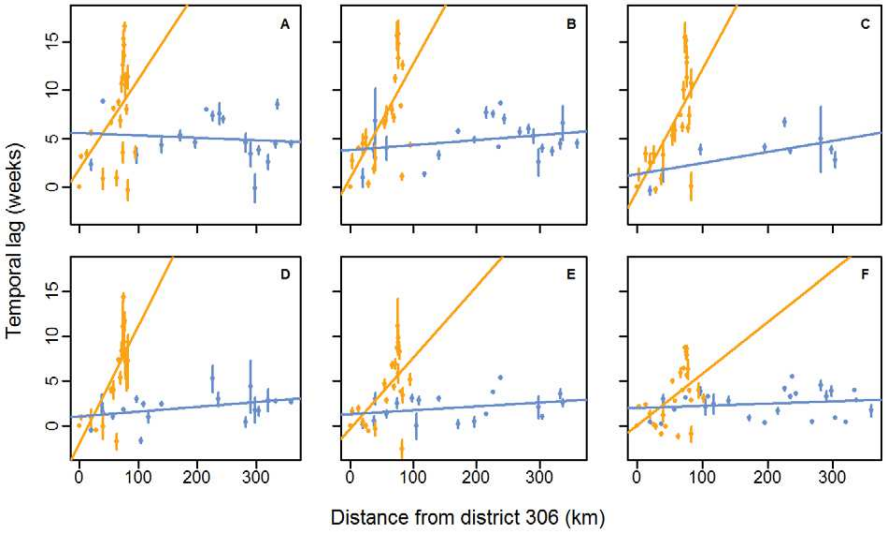


Figure 5. Scatterplot of mean annual temporal lags between epidemics against distances between districts. Temporal lags between epidemics and distances are computed relative to district #306. The lines show the linear regressions between the mean annual temporal lag of the annual epidemic in each district and the distance for 2002 (A), 2003 (B), 2004 (C), 2005 (D), 2006 (E) and 2007 (F). Colours represent the geographic localisation of each district, according to Figure 4A. The number of districts included in the analysis changes every year, according to whether an epidemic occurred in the district (Table 1). Error bars represent the 95% C.I. associated with the mean. Normality and homoscedasticity of residuals were confirmed using the Shapiro-Wilks and the Bartlett tests respectively (alpha level of 0.05).
doi:10.1371/journal.pntd.0001957.g005

composed of a flat flood plain, with a mean population density of 155 people per km^2. They are all located away from the three urban centres where sentinel hospitals involved in active surveillance are. They are consistently the same, year after year.

Secondly, Figure 4 demonstrates that the propagation of dengue is heterogeneous. Annual epidemics are highly synchronous along the national road linking Kampong Cham to Siem Reap whereas an oriented propagation emanates from a rural area located around Kampong Cham (district #306) and ends in Phnom Penh 11 weeks later (95% C.I. 4–18 weeks). This was obscured when districts were classified independently of the geographic area, taking into account only the distance to Phnom Penh (Figures S3, S4 and S5). Some years, another oriented propagation is seen from a rural area near the Vietnamese border (district #805) towards Phnom Penh (Figure 4B).

To evaluate the speed of propagation along each axis, we calculated the temporal lag of the seasonal pattern in each district relative to the district #306 [10] (see Methods). This district was chosen as a time reference point to calculate time lags because it is located at the intersection of the Mekong River and the national road areas, which was convenient to present results of both geographic areas on the same figure (Figure 5). Moreover, based on the ranking of districts using their phase, district #306 was the one with the most number of districts lagging behind over the study period (132, 134, 131, 118, 124 and 104 districts/134 respectively from 2002 to 2007). The results of the analysis of covariance performed on the mean annual lag to compare the speed of propagation of the epidemic in each geographic axis (Figure 4A) are presented in Table 1. Results show that regardless of the year, the evolution of the temporal lag between epidemics according to the distance differs significantly (α-level of 0.05) from one area to the other. We therefore performed separate regressions for each of the two axes (Figure 5). Propagation is always faster along the national road than along the Mekong River. The speed of propagation of the travelling wave along the Mekong River, evaluated by the inverse of the regression slope, is estimated to be 11 km per week (95% C.I. 3–18 km per week) over the study period. Along the national road, the speed is quasi-instantaneous with respect to our time domain sampling rate.

To test the role Phnom Penh plays in this heterogeneity, we ran the same analysis, but excluding Phnom Penh districts. The propagation remained significantly heterogeneous at an alpha-level of 5%, except in 2002 and 2007, and remained significantly heterogeneous at an alpha-level of 10% in 2002 and 2007 (results not shown).

The removal of districts with low population density has no effect at all on results shown. The exclusion of low incidence years did not modify the two conclusions arrived at in the paper (onset of the seasonal epidemic in highly localised rural areas, and heterogeneous propagation in the country), but removed the potential bias linked with the very low number of cases during the non epidemic years (see Figures S6 and S7 that show the same as Figures 4 and 5 when all years are included).

To sum up, our results show that the seasonal epidemic consistently starts in the same three rural districts in Cambodia. Then the propagation is not homogeneous in the country. In districts located along the busiest road, dengue epidemics appear simultaneously and early (with all districts being hit in less than a month), whereas districts located along the Mekong River get hit by the seasonal epidemic later, with the delay being proportional to the distance to district #306.

In 2007 an exceptional epidemic occurred in Cambodia with DENV-3 as the dominant etiological agent, and a four fold increase in weekly incidence rate compared to the four previous years (Figures 2 and 3). During this epidemic, the wave in the one-year periodic mode travelled from Kampong Cham to Phnom Penh in only 7 weeks, at a speed of 17 km per week, a higher synchronicity reflecting a more rapid propagation during this peculiar event. The sequence in which districts were affected by the seasonal variation was also modified, as shown by non-significant Spearman rank coefficients when correlations were calculated between any year and the year 2007.

Table 1. Results of the analysis of covariance and linear regressions (see methods).

	Covariance analysis		Regression (''Mekong River'')			Regression (''national road'')	
year	n[1]	p-value of interaction	n[1]	p-value[2]	1/slope estimate[3]	n[1]	p-value[2]
2002	40	0.005	22	0.02	11	18	0.70
2003	44	<0.001	22	0.001	9	22	0.22
2004	31	<0.001	22	<0.001	8	9	0.07
2005	38	<0.001	19	<0.001	8	19	0.09
2006	40	<0.001	22	0.004	13	18	0.22
2007	52	0.001	26	0.006	17	26	0.378

[1]Number of districts where an annual epidemic occurred according to the national threshold, along the Mekong River, the national road (Figure 4A), or, for the covariance analysis, both.
[2]P-value of the regression slope estimate.
[3]Inverse of the estimate of the regression slope β1, in km per week (see methods).
doi:10.1371/journal.pntd.0001957.t001

Discussion

A rural onset

Regarding the onset of the national epidemic, there is a tendency for the dengue season to start in few rural districts more often than in any other district moving from rural areas towards urban centres, with, for example, annual epidemics in districts #306 and #104 (Figure 4A) leading epidemics in the rest of the country by 3 weeks on average over the study period (95% C.I. 2–4 weeks). In addition, recent prospective cohort data showed that rural areas were affected by dengue to the same degree as urban areas or, as during the 2007 epidemic, at even higher incidence rates [18]. This finding is not consistent with the common thought that the most populated areas spread the disease [7–9] or known mechanisms underlying travelling waves [7,13]. One plausible explanation for a rural origin of the spread, also supported by a recent study in Vietnam [19], is that more than 80% of rural Cambodians do not have access to public water supply, and store their drinking water in big jars that have been identified as major breeding sites for *Aedes* mosquitoes in Cambodia [2]. This result (onset in rural areas) has strong implications regarding the control of dengue in low-income countries.

Factors that could explain the heterogeneous pattern of propagation

To our knowledge, it is the first time that a recurrent heterogeneous pattern of propagation of dengue is revealed in surveillance data at a national scale. This pattern of propagation could come from actual transmission of dengue viruses between districts, via infected mosquitoes or humans, or correspond merely to spatial differences in the emergence of the epidemic, differences due either to spatial differences in local (within-district) dynamics, or to forcing by extrinsic factors such as climate (Moran effect). Given the limited dispersal range of *Aedes spp*. vectors and the high synchronicity of the epidemics over a 400 km wide area along the national road, it is unlikely that viral transmission from one place to another via infected mosquitoes can account for the pattern observed.

Climate is quite homogeneous in Cambodia, with a narrow temperature range across the country, and the monsoon striking the country from South to North within a month only, around April-May. Climate, by influencing the vector's life cycle, is clearly driving the dengue seasonality observed in Cambodia. It could easily explain the high synchronicity of epidemics observed along the national road if this synchronisation was observed country-wide. However, two of our observations preclude climatic forcing to be the underlying factor for the spatial-temporal pattern observed. First, climate is more homogeneous in the country compared to the observed spatial-temporal pattern of dengue fever. The existence of a very slow oriented dengue propagation along the Mekong River, going North to South, with the epidemic in some districts lagging more than 3 months behind the one in districts that are only less than 200 kilometres in distance cannot be accounted for by climatic differences between districts. Secondly, the epidemic starts as early as in March in the three districts identified as starting points, whereas the monsoon only begins end of April/beginning of May.

We believe the heterogeneous propagation observed is related to the heterogeneity of traffic on the roads of the country, where traffic allows the movement of human carriers or transported mosquitoes infected with dengue. The national road –the busiest country road connecting the two main economic cities within 4–5 hours – probably explains synchronicity. By contrast, the existence of dirt roads along the entire Mekong River, on which traffic and population movement between smaller villages are slower and more difficult than along the national road supports our hypothesis. Despite its epidemiological relevance, our understanding of the relationship between human movement and pathogen transmission remains limited [20–22]. The importance of the relationship between human movement and pathogen transmission in explaining the spatio-temporal dynamics of dengue incidence or other diseases has been increasingly pointed-out during the last decade [5,8,9,23–25] and explored mainly through a theoretical modelling approach [22,26,27].

2007, a peculiar year

The higher synchronisation and higher incidence levels across the country in 2007 argues for an association with higher net reproduction ratios of infection due to lower herd immunity when a new serotype is introduced [25,28]. This hypothesis is supported by the fact that serotype 3 invaded Cambodia at the end of year 2006, replacing serotype 2, in place since 1999, as the major circulating virus [2]. The fact that heterogeneity of propagation remained despite this serotype change supports the hypothesis that another factor – such as human movement – plays an important role in the dynamics.

Limitations

Major limitations of our analysis include underreporting inherent to passive surveillance data and potential selection bias

leaning towards underreporting from urban centres compared with rural areas. However, underreporting would only affect the amplitude of the epidemics in each district and therefore have little effects on the study of synchronism when using the wavelet phase analysis approach [12]. Secondly, it is unlikely that rural Cambodians were over-represented as most hospitals and those that recruit dengue patients through active surveillance are free of charge and located in urban areas (three cities). One could also assume that collecting data from both an active and a passive surveillance system could affect the timing of detection of the epidemics, with active surveillance sites more likely to reveal a small increase in incidence levels earlier than passive surveillance sites. This would lead to a bias in the analysis. However, after a thorough analysis of the surveillance system (Institut Pasteur Cambodia's not published report), we believe that severe cases are scarcely missed even by the passive surveillance system; we also believe that our results are not affected by this bias, given the fact that districts affected early in the epidemic are located in rural areas, kilometres away from the urban sentinel sites. Another common limitation when analysing surveillance data is the introduction of spatial or temporal biases due to difficulties in standardising surveillance systems in time and space, especially in low income countries. We thus, on purpose, excluded any results or comments that would rely on spatial differences in incidence levels only. We found the wavelet approach very robust to these biases when exploring spatial-temporal patterns: unlike many other temporal methods, the level of incidence in a given district does not influence the calculation of time lags between epidemics.

One could think that not standardising incidence according to age could impair the results obtained. But age standardisation only affects the results by modifying incidence levels, and as our results do not rely on differences in incidence levels, this is not a real concern.

Perspectives

Our findings and speculations require further research for additional evidence. Firstly, reasons as to why the three areas identified as being hit early by the epidemic repeatedly year after year (districts #306, #104 and #805) are unclear. This warrants further filed investigations to identify specific factors that trigger the epidemic in these settings.

Secondly, the data we analysed here can only reveal the spatio-temporal dynamics and help make hypotheses on underlying factors. Given the results of this study, and particularly the heterogeneity of dengue spatial dynamics, surveillance data on the spatial distribution of the serotypes (and genotypes if possible) of the co-circulating dengue viruses would help validate (or not) our hypotheses on dengue dynamics in Cambodia, using independent data.

Lastly, given the potential benefit in term of disease control from demonstrating the efficient role of humans' movement in dengue spatial transmission, one might consider further investigations in this direction, collecting data and using dynamic models for instance.

Supporting Information

Figure S1 Maps of weekly incidence rates in Cambodia. Maps show weeks 5, 9, 11, 14, 18, 22, 25, 28, 31, 35, 38, 42 and 50 of year 2002 in Cambodian districts (in number of cases declared per 100,000 people per week).
(PDF)

Figure S2 Map of Cambodian districts with their identifying numbers and wavelet power spectra of the weekly incidence rates of districts with more than 20 people per km^2. Power is colour coded, blue indicating low power. Dashed lines represent the limit of the cone of influence where data are modified by edge effects. The black lines show the 0.05 alpha-level of significance computed based on 500 bootstrapped series [10].
(PDF)

Figure S3 Phases of the annual component of incidence for districts located along the ''Mekong'' axis (orange in figure 4A). Phases are computed in the 0.8–1.2 year periodic band. Districts are ranked by increasing distance to Phnom Penh, from bottom to top. The arrows indicate: 1, district #306; 2, Phnom Penh; 3, district #805.
(TIF)

Figure S4 Phases of the annual component of incidence for districts located along the ''national road'' axis (blue in figure 4A). Phases are computed in the 0.8–1.2 year periodic band. Districts are ranked by increasing distance to Phnom Penh, from bottom to top. The arrow indicates district #104.
(TIF)

Figure S5 Phases of the annual component of incidence for districts with more than 20 people per km^2. Phases are computed in the 0.8–1.2 year periodic band. Districts are ranked by increasing distance to Phnom Penh, from bottom to top. The arrows indicate: 1, Phnom Penh; 2, District #306; 3, district #805; 4, district #104 (see figure 4A).
(TIF)

Figure S6 Phases of the annual component of incidence for districts located along two geographic axes (see Figure 4A for a map of the two geographic areas chosen). Phases are computed in the 0.8–1.2 year periodic band. Figure shows the same as Figure 4B and 4C, but with all years included. (A) Phase of districts along the Mekong River (orange in Figure 4A), presented from the most southerly to the most northerly from bottom to top. (B) Phase of districts along the national road (blue in Figure 4A), presented from West to East from bottom to top. The arrows indicate districts: 1, #306; 2, Phnom Penh (Figure S6A) and #104 (Figure S6B).
(TIF)

Figure S7 Scatterplot of mean annual temporal lags between epidemics against distances between districts. Temporal lags between epidemics and distances are computed relative to district #306. The lines show the linear regressions between the mean annual temporal lag of the annual epidemic in each district and the distance for 2002 (A), 2003 (B), 2004 (C), 2005 (D), 2006 (E) and 2007 (F). Colours represent the geographic localisation of each district, according to Figure 4A. Error bars represent the 95% C.I. associated with the mean. Figure shows the same as Figure 5, but with all years included. Each year, the number of units included in the analysis is 27 for the Mekong axis, and 26 for the national road axis.
(TIF)

Movie S1 Evolution of weekly incidence rates in Cambodia. Incidence rates are shown on maps of Cambodian districts from January 2002 to December 2008.
(MOV)

Movie S2 Evolution the weekly phases of the annual component of incidence. Phases are computed in the 0.8–1.2 year periodic band and shown on maps of Cambodian districts from January 2002 to December 2008.
(MOV)

Acknowledgments

We thank Tristan Rouyer for providing an R toolbox (translated from Cazelles' Matlab toolbox) for performing wavelet analysis, Kévin Cazelles, Christophe Menkès and Morgan Mangeas for their help with parts of the analysis, and Yann Quéau for useful comments.

Author Contributions

Conceived and designed the experiments: MT BC SV. Analyzed the data: MT RH BC RD CB SV. Wrote the paper: MT RH BC RD CB SV. Provided the data: RH. Designed the toolbox to perform wavelet analysis: BC.

References

1. Wearing HJ, Rohani P (2006) Ecological and immunological determinants of dengue epidemics. Proc Natl Acad Sci U S A 103: 11802–11807.
2. Huy R, Buchy P, Conan A, Ngan C, Ong S, et al. (2010) National dengue surveillance in Cambodia 1980–2008: epidemiological and virological trends and the impact of vector control. Bull World Health Organ doi:10.2471/BLT.09.073908.
3. Hay SI, Myers MF, Burke DS, Vaughn DW, Endy T, et al. (2000) Etiology of interepidemic periods of mosquito-borne disease. Proc Natl Acad Sci U S A 97: 9335–9339.
4. Mammen MP Jr, Pimgate C, Koenraadt CJM, Rothman AL, Aldstadt J, et al. (2008) Spatial and temporal clustering of dengue virus transmission in Thai villages. PLoS Med doi:10.1371/journal.pmed.0050205.
5. Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, et al. (2005) Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. Am J Trop Med Hyg 72: 209–220.
6. Chan KL (1985) Singapore's dengue hemorrhagic fever control programme: a case study on the successful control of *Aedes aegypti* and *Aedes albopictus* using mainly environmental measures as a part of integrated vector control. Tokyo: southeast Asian Medical Information Centre.
7. Cummings DAT, Irizarry RA, Huang NE, Endy TP, Nisalak A, et al. (2004) Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature 427: 344–347.
8. Thai KTD, Cazelles B, Nguyen NV, Vo LT, Boni MF, et al. (2010) Dengue dynamics in Binh Thuan province, southern Vietnam: periodicity, synchronicity and climate variability. PLoS Negl Trop Dis doi:10.1371/journal.pntd.0000747.
9. Rabaa MA, Hang VTT, Wills B, Farrar J, Simmons CP, et al. (2010) Phylogeography of recently emerged DENV-2 in southern Viet Nam. PLoS Negl Trop Dis doi:10.1371/journal.pntd.0000766.
10. Cazelles B, Chavez M, McMichael AJ, Hales S (2005) Nonstationary influence of el Niño on the synchronous dengue epidemics in Thailand. PLoS Med doi:10.1371/journal.pmed.0020106.
11. WHO (2009) Dengue guidelines for diagnosis, treatment, prevention and control. Available: http://whqlibdoc.who.int/publications/2009/9789241547871_eng.pdf. Accessed 2012 Apr 15.
12. Torrence C, Compo GP (1998) A practical guide to wavelet analysis. Bull Am Meteorol Soc 79: 61–78.
13. Grenfell BT, Bjornstad ON, Kappey J (2001) Travelling waves and spatial hierarchies in measles epidemics. Nature 414: 716–723.
14. Cazelles B, Chavez M, Constantin de Magny G, Guégan JF, Hales S (2007) Time-dependent spectral analysis of epidemiological time-series with wavelets. J R Soc Interface 4: 625–636.
15. R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available: http://www.R-project.org. Accessed 2012 Nov 7.
16. Stroup DF, Williamson GD, Hernon JL (1989) Detection of aberrations in the occurrence of notifiable diseases surveillance data. Stat Med 8: 323–329.
17. Zuur AF, Ieno EN, Smith GM (2007) Analysing Ecological Data. New York: Springer. 685p.
18. Vong S, Khieu V, Glass O, Ly S, Duong V, et al. (2010) Dengue incidence in urban and rural Cambodia: results from population-based active fever surveillance, 2006–2008. PLoS Negl Trop Dis doi:10.1371/journal.pntd.0000903.
19. Schmidt WP, Suzuki M, Thiem VD, White RG, Tsuzuki A, et al. (2011) Population density, water supply and the risk of dengue fever in Vietnam: cohort study and spatial analysis. PLoS Med doi:10.1371/journal.pmed.1001082.
20. Vazquez-Prokopec GM, Stoddard ST, Paz-Soldan V, Morrison AC, Elder JP, et al. (2009) Usefulness of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. International Journal of Health Geographics 8 doi:10.1186/1476-072X-8-68.
21. Gonzales MC, Hidalgo CA, Barabasi AL (2008) Understanding individual human mobility patterns. Nature 453: 779–782.
22. Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Paz-Soldan V, Kochel TJ, et al. (2009) The role of human movement in the transmission of vector-borne pathogens. PLoS Negl Trop Dis doi:10.1371/journal.pntd.0000481.
23. Morrison AC, Getis A, Santiago M, Rigau-Perez JG, Reiter P (1998) Exploratory space-time analysis of reported dengue cases during an outbreak in Florida, Puerto Rico, 1991–1992. Am J Trop Med Hyg 58: 287–298.
24. Honorio NA, Ribeiro-Nogueira RM, Torres Codeço C, Carvalho MS, Gonçalves-Cruz O, et al. (2009) Spatial evaluation and modeling of dengue seroprevalence and vector density in Rio de Janeiro, Brazil. PLoS Negl Trop Dis doi:10.1371/journal.pntd.0000545.
25. Viboud C, Bjornstad ON, Smith DL, Simonsen L, Miller MA, et al. (2006) Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 312: 447–451.
26. Adams B, Kapan DD (2009) Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics. PLoS One doi:10.1371/journal.pone.0006763.
27. Barmak DH, Dorso CO, Otero M, Solari HG (2011) Dengue epidemics and human mobility. Phys Rev E, 84: 011901 doi: 10.1103/PhysRevE.84.011901.
28. Longini IM, Fine PEM, Thacker SB (1986) Predicting the global spread of new infectious agents. Am J Epidemiol 123: 383–391.

Annexe 7 : Equations utilisées pour l'analyse en ondelette

La méthode, introduite par Morlet et Meyer, est basée sur l'utilisation d'une famille de fonctions dérivées d'une fonction unique d'ondelette appelée ondelette mère, $\Psi(t)$. La famille de fonctions $\Psi_{a,\tau}(t)$, dérivées de cette ondelette mère, peut être exprimée en fonction de 2 paramètres permettant, l'un une localisation dans le temps (τ), l'autre une localisation dans le domaine fréquentiel (a) (Cazelles et al., 2005, 2007; Torrence and Compo, 1998):

$$\Psi_{a,\tau}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-\tau}{a}\right) \quad (1)$$

Le paramètre *a* est un paramètre d'échelle, et est lié aux périodes (inverse de la fréquence) de Fourier par un lien qui dépend de l'ondelette mère choisie.

Le choix des échelles d'analyse est déterminé par la longueur du signal, ainsi que sa fréquence d'échantillonnage. L'échelle analysable minimale a_0 correspond à la fréquence de Shannon, qui est le double de la fréquence d'échantillonnage (ici 1 /2 semaines). L'échelle maximale analysable est :

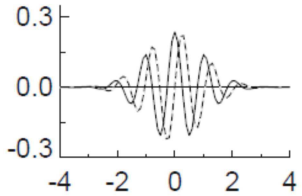
$$a_{\max} = a_0 . 2^{J\delta j} \quad (2)$$

avec

$$J = \delta j^{-1} . \log_2\left(\frac{N\delta t}{a_0}\right) \quad (3)$$

où δj représente la résolution dans le domaine fréquentiel (fixée ici à 1/50), *N* le nombre d'observations, δt le pas d'échantillonnage en années (i.e. 1/52 pour des données hebdomadaires, 1/1 pour des données annuelles et 1/12 pour des données mensuelles).

Pour l'analyse de nombreux signaux, l'ondelette de Morlet est utilisée. Elle est définie comme:



$$\Psi(t) = \pi^{-1/4}\exp(-i\,2\pi f_o\,t)\exp\left(\frac{-t^2}{2}\right) \quad (4)$$

Elle est de forme complexe. Sur la figure, la ligne pleine représente la partie réelle et la ligne pointillée la partie imaginaire.

Pour l'ondelette de Morlet et en prenant un paramètre f_0 (fréquence centrale de l'ondelette mère) égal à 2π, l'échelle d'analyse a est équivalente à la période p de Fourier et à l'inverse de la fréquence f correspondante.

La transformée d'ondelette d'un signal $x(t)$ pour une échelle a donnée au temps τ est définie comme la convolution entre le signal et la forme conjuguée de l'ondelette :

$$W_x(a,\tau) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} x(t)\Psi^*\left(\frac{t-\tau}{a}\right)dt = \int_{-\infty}^{\infty} x(t)\Psi^*_{a,\tau}(t)dt \qquad (5)$$

En répétant cette transformation du signal pour un ensemble d'échelles d'analyse a et différentes localisations dans le temps τ, on obtient un ensemble de coefficients $W_x(a,\tau)$ appelés coefficients d'ondelette.

Lorsque l'ondelette mère choisie a une forme complexe (comme c'est le cas pour l'ondelette de Morlet), ce coefficient est aussi un nombre complexe, pouvant être écrit sous la forme $|W_x(a,\tau)|\ e^{i\,\theta(a,\tau)}$, où $\theta(a,\tau)$ représente la phase du signal au temps τ pour l'échelle a, et peut être calculée par :

$$\Phi_x(f,\tau) = \tan^{-1}\frac{\Im(W_x(f,\tau))}{\Re(W_x(f,\tau))} \qquad (6)$$

La puissance du signal $x(t)$ dans la fréquence $f = 1/a$ au temps τ est donnée par $|W_x(f,\tau)|^2$. La représentation de cette puissance dans diverses fréquences et à différents temps τ est appelée spectre d'ondelette.

Enfin, le signal $x(t, f_1, f_2)$ correspondant filtré dans la bande de fréquence $[f_1, f_2]$ peut-être reconstruit à partir des coefficients d'ondelette.

$$x(t) = \frac{\delta j.\delta t^{1/2}}{C_\delta.\Psi_0(0)}.\sum_{a=1/f1}^{1/f2}\frac{\Re(W_x(f,\tau))}{a^{1/2}} \qquad (7)$$

où $C_\delta = 0.776$ et $\Psi_0(0) = \pi^{-1/4}$ pour l'ondelette de Morlet avec une fréquence centrale de 2π.

L'analyse en ondelette d'un signal $x(t)$ permet donc :

- d'estimer la répartition de la puissance totale (ou variance) d'un signal entre diverses fréquences au cours du temps ;
- de filtrer la composante d'un signal correspondant à une fréquence donnée en n'opérant la reconstruction de signal qu'à partir des coefficients correspondant à cette fréquence ;

– pour une composante de fréquence donnée, d'extraire sa phase à chaque instant.

L'association statistique entre deux signaux *x(t)* et *y(t)* (comme par exemple les courbes épidémiques de 2 zones géographiques distinctes) peut être quantifiée en calculant la cohérence d'ondelette :

$$R_{x,y}(f,\tau)=\frac{\left|<W_{x,y}(f,\tau)>\right|^2}{\left|<W_x(f,\tau)>\right|^2\left|<W_y(f,\tau)>\right|^2} \tag{8}$$

où $W_x(f,\tau)$ et $W_y(f,\tau)$ représentent les coefficients d'ondelette des signaux *x(t)* et *y(t)* respectivement, $W_{x,y}(f,\tau)=W_x(f,\tau)\,W_y^*(f,\tau)$ le coefficient croisé d'ondelette (l'astérisque signifie que l'on utilise la forme conjuguée de $W_y(f,\tau)$), et les parenthèses angulaires représentent un lissage dans le temps et dans les fréquences.

Lorsque l'ondelette mère est de forme complexe, le coefficient croisé peut aussi s'écrire :

$$W_{x,y}(f,\tau)=\left|W_{x,y}(f,\tau)\right|e^{i\left(\theta_y(f,\tau)-\theta_x(f,\tau)\right)} \tag{9}$$

où $\theta_y(f,\tau)$ et $\theta_x(f,\tau)$ représentent respectivement la phase des signaux y(t) et x(t) au temps τ pour la fréquence f. Il est alors possible d'estimer la différence de phase entre 2 signaux par :

$$\Phi_{x,y}(f,\tau)=\tan^{-1}\frac{\Im\left(W_{x,y}(f,\tau)\right)}{\Re\left(W_{x,y}(f,\tau)\right)} \tag{10}$$

Dans une bande de fréquence $[f_1,f_2]$ (ou de période) donnée, le déphasage entre 2 signaux rapporté à 2π est égal au décalage temporel $\Delta T(\tau)$ entre les 2 signaux rapporté à la période d'analyse considérée :

$$\Delta T(\tau)=\frac{\Phi_{x,y}(f,\tau)}{2\pi F(\tau)} \tag{11}$$

avec $F(\tau)$ la fréquence instantanée définie dans une bande de fréquence $[f_1,f_2]$:

$$F(\tau)=\frac{\int_{f_1}^{f_2} f\left|W_{x,y}(f,\tau)\right|df}{\int_{f_1}^{f_2}\left|W_{x,y}(f,\tau)\right|df} \tag{12}$$

Le signe de la différence de phase permet de savoir quel signal est en avance sur l'autre.

Le cône d'influence représente la région du spectre influencée par les effets de bords. Cette zone dépend uniquement de l'échelle d'analyse. Pour l'ondelette de Morlet, l'étendue de ce cône pour une échelle *a* donnée est $\sqrt{2}.a$.

Afin de diminuer les effets de bord, il est recommandé, pour les séries de longueur finie, de rajouter des zéros en début et fin de série ("padding") jusqu'à obtenir une série dont la longueur est égale à la première puissance de 2 supérieure au nombre d'observations.

Annexe 8 : Questionnaire utilisé dans l'enquête spatiale des facteurs de risque socio-économiques de la dengue à Nouméa

Nom de l'enquêteur					
Quartier | date de l'enquête					
Code maison					
Prénom de l'enquêté					
Adresse complète et téléphone (pour compléter réponses si besoin par la suite)					
Q	***Enquête auprès du foyer***	**réponses**	*lire toutes les réponses*	**Codes**	*encercler le choix*
1	Depuis quelle année habitez-vous ici ?				
2	Quel est votre appartenance communautaire si vous êtes Calédonien ?	Kanak	Européenne	Wallis/Futuna	Tahitienne
		Indonésienne	Vietnamienne	Ni-Vanuatu	Autre
3	Lieu de naissance si vous n'êtes pas Calédonien				
4	Nb personnes dans le foyer				
5	Enfants (<18 ans) dans le foyer				
6	Personnes > 60ans dans le foyer				
7	Adultes avec un emploi dans le foyer				
8	Si oui, dans quel quartier travaillez-vous ?	*personne*			
9	Vous faites-vous piquer sur votre lieu de travail ?	*personne*	oui	non	nsp
10	Personnes restant à la maison		Toute la journée	La moitié du temps	Parfois
11	Quels sont les endroits que vous fréquentez régulièrement où il y a du monde? ? Nakamal ; parcs à jeux ; bars/resto ; cinéma ; magasins ou autres lieux publics	*personne*			
12	Vous faites-vous piquer dans ces lieux ?		oui	non	nsp
13	Y a-t-il des membres du foyer qui ont eu la dengue ?	oui et combien ?		non	nsp
14	Quel est votre mode de déplacement (voiture ,2 roues, bus, vélo…)				
Lutter contre les moustiques dans le foyer					
15	Où trouvez-vous des moustiques ? (Lire les réponses possibles : jardin, terrasses ; chambres, pièces de vie ; garage)				
16	Que faites-vous pour lutter contre les moustiques ?				
17	Est-ce efficace?		oui	non	nsp

18	Quand vous faites-vous le plus piquer? (plusieurs réponses possibles)	matin	oui	non	nsp
		midi	oui	non	nsp
		soir	oui	non	nsp
		toute la journée	oui	non	nsp
		la nuit	oui	non	nsp
19	Vous faîtes vous piquer plutôt à	l'intérieur	oui	non	nsp
Perceptions et connaissances					
20	Etes-vous informé quand il y a une épidémie de dengue ?		oui	non	nsp
21	Vous protégez-vous plus lorsqu'il y a une épidémie de dengue ?		oui	non	nsp
22	Selon vous, quelle est la mesure de lutte contre les moustiques la plus efficace ?				

Q	***Observation du milieu de vie***	***Réponses***	***Codes***	***encercler le choix***	***encercler le choix***
23	Type de construction	Indépendante	oui	non	nsp
		Mitoyenne	oui	non	nsp
		Collective	oui	non	nsp
		Villa partagée	Haut	Bas	nsp
24	Maison	Sur pilotis	oui	non	nsp
		Sans étage	oui	non	nsp
		Avec 1 étage	oui	non	nsp
		Plusieurs étages	oui	non	nsp
		Climatisation	oui	non	nsp
		Moustiquaires aux fenêtres	oui	non	nsp
25	Gouttières	Aucune	oui	non	nsp
		Bon état	oui	non	nsp
		Mauvais état	oui	non	nsp
26	Jardin	Aéré	oui	non	nsp
		Dense	oui	non	nsp
		Entretenu	oui	non	nsp
		Peu entretenu	oui	non	nsp
27	Présence de haie(s)	Aérée	oui	non	nsp
		Dense	oui	non	nsp
28	Plantes	Grands arbres (>5m ou plus grand que la maison) ?	oui	non	nsp
		fruitiers	oui	non	nsp
		Serre (en ombrière)	oui	non	nsp
		<10 plantes en pot avec soucoupe	oui	non	nsp

		10-30 plantes en pot avec soucoupe	oui	non	nsp
		>30 plantes en pot avec soucoupe	oui	non	nsp
29	Animaux	Chiens, chats	oui	non	nsp
		Oiseaux, volailles	oui	non	nsp
		Autres:	oui	non	nsp
30	Objets divers	Stockage d'encombrants	oui	non	nsp
		Pneus	oui	non	nsp
		Bateau	oui	non	nsp
		Objets divers abandonnés	oui	non	nsp
30	Etat de la maison		Maison en bon état	Maison en état moyen	Maison en état médiocre/mauvais
31	Etat des abords de la maison		Abords impeccables	Abords moyens	Abords à l'abandon
32	Ombrage		Moins de 25% ombragé ; pas de grands arbres	Entre 25 et 50% ombragé ombragé ; peu d'arbres ou de de buissons	Terrain fortement ombragé ; arbres et buissons
Q	***Observation de la rue***	***Réponses***	***Codes***	***encercler le choix***	
33	Rue large	2 voies	oui	non	nsp
34	Rue étroite	1 voie	oui	non	nsp
35	Terrain abandonnés	nb :		non	nsp
36	Maisons abandonnées	nb :		non	nsp