



# World Health Organization

Modélisation mathématique pour la biologie

Etudes épidémiologiques

---

## Modélisation de la propagation de la peste, le choléra, la dengue et la grippe saisonnière

---

*Auteurs :*

Mlle Imane BENHAYOUN

M. Ahmed EL AAMRANI

M. Augustin GIROS

M. Naoufal LAYAD

*Encadrants :*

Pr. Véronique LETORT

Dr. Jean LETORT

Version du  
28 janvier 2018



# Remerciements

L'ensemble des membres du groupe voudrait commencer par remercier très chaleureusement l'Organisation Mondiale de la Santé (WHO) pour nous avoir offert la possibilité de réaliser ce travail en nous offrant un programme à la fois riche et concret. Ce programme a la mention d'avoir mobilisé toute l'Organisation - le Siège, les Régions et les pays -, faisant ainsi de la santé mondiale le principal objectif de l'Organisation ainsi que de ses partenaires. Nous tenons également à souligner l'effort qu'a mis l'OMS pour répondre à nos besoins et surtout aux besoins des pays atteints de ces épidémies.

Nous tenons également à remercier l'Ecole CentraleSupélec pour avoir mis à notre disposition un éventail d'outils pluridisciplinaires pour amener à bien notre projet. Nous avons donc pu tirer profit de l'enseignement qu'elle nous a offerte, ainsi que l'échange avec des scientifiques et des chercheurs de haut niveau dans le domaine de la biologie et des mathématiques appliquées.

On aimerait donc tout particulièrement remercier Pr. Véronique Letort Le Chevalier pour son encadrement tout au long de la réalisation de cette mission, ainsi que ses nombreuses remarques critiques visant à améliorer la qualité de notre travail.



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 La dengue</b>	<b>3</b>
1.1 Etude biologique, médicale et épidémiologique . . . . .	3
1.2 Description du modèle et simulations . . . . .	5
1.2.1 Modèle SIR-SI . . . . .	5
1.3 Analyse du modèle . . . . .	10
1.4 Estimation des paramètres . . . . .	12
1.5 Discussion, perspectives . . . . .	14
<b>2 Le choléra</b>	<b>15</b>
2.1 Etude biologique, médicale et épidémiologique . . . . .	15
2.2 Modélisation du choléra : Modèle de Codeço . . . . .	17
2.3 Simulation . . . . .	20
2.4 Estimation . . . . .	20
2.5 Analyse du modèle de Codeço . . . . .	21
2.6 Challenge de l'OMS . . . . .	23
2.7 Prise en compte des conditions météorologiques . . . . .	23
2.8 Discussion, perspectives . . . . .	24
<b>3 La grippe saisonnière</b>	<b>27</b>
3.1 Etude biologique, médicale et épidémiologique . . . . .	27
3.2 Modélisation de la maladie . . . . .	30
3.3 Simulation . . . . .	32

---

3.4	Analyse de sensibilité . . . . .	34
3.5	Estimation . . . . .	35
3.6	Discussion et perspectives . . . . .	38
<b>4</b>	<b>La peste bubonique</b>	<b>39</b>
4.1	Etude biologique,médicale et épidémiologique . . . . .	39
4.1.1	Présentation de la maladie . . . . .	39
4.1.2	Signes et symptômes . . . . .	40
4.1.3	Comment gérer cette épidémie? . . . . .	40
4.2	Description du modèle et simulations . . . . .	41
4.2.1	Introduction au modèle . . . . .	41
4.2.2	Description du modèle et formalisme mathématique . . . . .	42
4.2.3	Simulations . . . . .	44
4.3	Analyse du modèle . . . . .	47
4.3.1	Etude de la sensibilité . . . . .	47
4.4	Estimation . . . . .	49
4.4.1	Méthodes utilisées en estimation . . . . .	49
4.4.2	Difficultés et problèmes liés à l'estimation . . . . .	51
4.5	Discussion et perspectives . . . . .	52
4.5.1	Proposition d'un nouveau système d'équations . . . . .	52
4.5.2	Le choix de la fonction de température . . . . .	52
4.5.3	Conclusion sur le modèle . . . . .	52
	<b>Conclusion</b>	<b>53</b>

# Table des figures

1.1	La moustique 'aedes aegypti' . . . . .	3
1.2	Schéma du modèle compartimental du modèle SIR-SI . . . . .	6
1.3	Définitions et valeurs de paramètres basiques utilisés dans les simulations .	7
1.4	Validation (1) de l'implémentation du modèle SIR-SI . . . . .	8
1.5	Validation (2) de l'implémentation du modèle SIR-SI . . . . .	8
1.6	Simulation du jeu de paramètres basiques . . . . .	9
1.7	L'influence de la vaccination sur le modèle SIR-SI . . . . .	9
1.8	Méthode de Morris pour le modèle SIR-SI . . . . .	10
1.9	Méthode de Sobol pour le modèle SIR-SI . . . . .	11
1.10	Résultat d'approximation uni-populationnelle (1) . . . . .	12
1.11	Résultats d'approximation uni-populationnelle (2) . . . . .	13
1.12	Résultats d'approximation multi-populationnelle . . . . .	13
2.1	Bacille 'Vibrio Cholerae' dans l'eau . . . . .	15
2.2	Diagramme compartimental du modèle de Codeço [Mraidi 2014] . . . . .	18
2.3	Liste de paramètres pour 3 communautés distinctes (Modèle de Codeço) .	18
2.4	L'évolution d'une situation épidémique (A) et d'une situation endémique (B) pour les 4 variables, (a) Susceptibles, (b) Infectés, (c) Guéris, (d) Concen- tration en bactérie dans le milieu aquatique . . . . .	19
2.5	Simulation (1) du modèle de Codeço - Evolution de $I(t)$ . . . . .	20
2.6	Données de simulation (1) bruitées . . . . .	20
2.7	Courbe de simulation (1) approximée . . . . .	21
2.8	Paramètres estimés de la simulation (1) . . . . .	21
2.9	Analyse de Sobol appliquée au modèle de Codeço . . . . .	22

2.10	Indices de Sobol pour le modèle de Codeço . . . . .	22
2.11	Modèle de Codeço avec conditions météorologiques . . . . .	23
2.12	Modèle de Codeço avec conditions météorologiques . . . . .	24
2.13	Profil de la Température au Nékem (2018) . . . . .	25
3.1	Évolution hebdomadaire des taux de consultations pour syndrome grippal par groupe d'âge, saison 2016- 2017, France métropolitaine [Bourdillon 2017]	30
3.2	Diagramme compartimental de la grippe saisonnière . . . . .	31
3.3	Simulation (1) du modèle de la grippe saisonnière . . . . .	32
3.4	Simulation (1)-suite du modèle de la grippe saisonnière . . . . .	33
3.5	Méthode de Morris appliquée à la grippe saisonnière . . . . .	34
3.6	Résultats de la méthode de Morris . . . . .	34
3.7	Données simulées après bruitage . . . . .	35
3.8	Estimation sur les données simulées . . . . .	36
3.9	Résultats de l'estimation sur les données simulées . . . . .	37
3.10	Résultats de l'estimation (2) sur les données simulées . . . . .	37
4.1	La bactérie <i>Yersinia pestis</i> responsable de la peste. . . . .	39
4.2	Le nombre des décès hebdomadaires dus à la peste à Bombay . . . . .	41
4.3	Le processus de la transmission de la peste bubonique. . . . .	42
4.4	Le profil de $\theta(t)$ à Bombay entre 01/1897 et 12/1911. . . . .	44
4.5	L'évolution du nombre des humains infectés (en rouge) et des rats infectés (en bleu), des rats immunisés (en noir) et des puces libres (en marron) . .	45
4.6	L'évolution de $H(t)$ (en rouge) et de $I(t)$ (en bleu) . . . . .	46
4.7	L'évolution de $R(t)$ (en noir) et $F(t)$ (en marron) . . . . .	46
4.8	L'étude de sensibilité du modèle selon la méthode de Morris . . . . .	48
4.9	Estimation du paramètre $a$ dans le modèle de la peste . . . . .	49
4.10	Estimation du paramètre $\pi_0$ dans le modèle de la peste . . . . .	50
4.11	Estimation des paramètres avec algorithme génétique . . . . .	51



# Introduction

Durant la soixante-septième session du Comité régional européen de l'OMS qui s'est déroulée en Hongrie entre le 11 et le 14 septembre 2017, notre équipe a eu l'immense joie de rencontrer le Docteur Tedros Adhanom Ghebreyesus. Encore fraîchement élu à la tête de l'Organisation Mondiale de la Santé, Docteur Tedros est le premier africain à occuper ce poste. Le mandat de son prédécesseur, Docteur Margaret Chan, a été largement critiqué pour avoir réagi trop lentement face à l'épidémie d'Ebola en Afrique de l'Ouest. Cette dernière a été accusée d'avoir manqué d'anticipation concernant la gravité de l'épidémie qui a débutée en décembre 2013, et qui a fini par tuer plus de 11 000 personnes.

Ancien ministre éthiopien de la santé et des affaires étrangères, et ancien président du conseil d'administration du Fonds Mondial de lutte contre le sida, la tuberculose et le paludisme, Docteur Tedros met parmi ses objectifs de campagne la nécessité de faire progresser la couverture sanitaire universelle et de veiller à ce que l'OMS réagisse rapidement et efficacement aux épidémies et aux situations d'urgence.

Durant cette soixante-septième session à Budapest, Dr Tedros nous a raconté son enfance, et plus particulièrement l'injustice profonde qu'il avait ressentie quand son frère de cinq ans est mort en raison de l'épidémie de rougeole. La disparité des soins entre le monde développé et le monde en développement continue d'être totalement inacceptable. Et c'est là que nous devons agir à priori.

Conscient de la nécessité d'apporter des solutions aux épidémies les plus fréquentes aux quatre coins du monde, et prêt à faire pression aux gouvernements pour y arriver, Dr Tedros nous a personnellement confié l'étude de 4 épidémies largement répandues qui sont le choléra, la peste bubonique, la dengue, et la grippe saisonnière, que nous cherchons à modéliser dans ce rapport.



# Chapitre 1

## La dengue

### 1.1 Etude biologique, médicale et épidémiologique

La dengue est une infection virale, le moustique vecteur est principalement ‘*Aedes aegypti*’, et plus rarement ‘*Aedes albopictus*’. Un moustique infecté peut transmettre le virus toute la durée de sa vie. L’homme permet aussi de transmettre la maladie au sein des moustiques, et aussi aux hommes au maximum 12 jours après l’infection. Historiquement, le mot "dengue" proviendrait de l’espagnol « dengüero » qui signifie « maniéré » pour décrire l’articulation anormale chez les personnes infectées.



FIGURE 1.1 – La moustique ‘*aedes aegypti*’

Une personne atteinte de dengue présente une forte fièvre, des maux de têtes, des douleurs musculaires et articulaires, de la fatigue, de nausées et de vomissements, et parfois d’une éruption cutanée ressemblant à celle de la rougeole. Une dengue sévère, elle, se manifeste par une baisse du nombre de plaquettes. La période d’incubation est entre 4 et 10 jours, tandis que la période infectieuse dure entre 2 et 7 jours. [WHO 2018b]

Géographiquement, on ne possède pas de description de la maladie avant la découverte du continent américain, où on pourrait rattacher les premières observations qui ressemblent à la dengue telle que décrite aujourd'hui. En Europe, la dengue est apparue en 1784, plus précisément en Andalousie à l'Espagne. La dernière grande épidémie de dengue qui a frappé l'Europe est survenu à Athènes, en Grèce, en 1927-1928. Celle-ci a atteint des chiffres records avec 1 millions de malades et un millier de décès durant l'été 1928.

La maladie de dengue a actuellement une incidence relativement élevée par rapport aux autres épidémies, et c'est la raison pour laquelle on la qualifie d'épidémie « ré-émergente ». Environ la moitié de la population mondiale est susceptible d'être infectée. Actuellement, on estime à 390 millions le nombre de personnes infectées annuellement (ou 284-528 millions pour un intervalle de confiance à 95%), dont seulement un quart présente des manifestations cliniques.

On estime aussi à 500 000 le nombre de personnes atteintes de dengue sévère, avec un taux de mortalité de 2.5%. La dengue sévère s'explique par une fuite plasmatique, une accumulation liquidienne, une détresse respiratoire, des hémorragies profuses ou une insuffisance organique. La dengue sévère a été pour la première fois observée aux années 1950, aux Philippines et en Thaïlande, et continue aujourd'hui d'être davantage répandu dans les pays de l'Asie et de l'Amérique du Sud.

Il existe quatre sérotypes ou catégories distinctes pour les souches du virus de la dengue. En conséquence, un individu est susceptible d'être infecté par chacun de ses 4 sérotypes au cours de sa vie. On distingue également la dengue classique de la dengue hémorragique. La dengue hémorragique se manifeste comme l'indique son nom par des hémorragies multiples, notamment gastro-intestinales, cutanées et cérébrales. Elle est souvent observée chez des enfants de moins de 15 ans, et présente un état de choc, qui peut provoquer la mort en cas d'absence de perfusion. Cette mortalité infantile, à cause de la dengue sévère, s'observe tout particulièrement dans certains pays d'Asie et d'Amérique latine.

Pour examiner la maladie, un diagnostic virologique, précis et rapide est donc nécessaire. Jusqu'à ce jour-ci, il n'existe pas de traitement spécifique pour la dengue, et un diagnostic précoce et l'accès aux soins adaptées permet de faire baisser le taux de mortalité en dessous du seuil de 1%. La prévention et la maîtrise de cette épidémie repose sur des mesures efficaces de lutte antivectorielle, dont l'objectif est de réduire une éventuelle chaîne de transmission vectorielle et d'éviter ainsi une épidémie.

## 1.2 Description du modèle et simulations

La réémergence de la dengue et sa large diffusion pendant les dernières décennies constitue un problème majeur pour l'Organisation Mondiale de la Santé et les autorités sanitaires dans les pays en développement. Avec plus de 2,5 millions de personnes susceptibles, l'épidémie nécessite un large effort pour comprendre sa dynamique, évaluer son impact et trouver des moyens de contrôle. Plusieurs auteurs ont pointé la dengue comme un problème sanitaire majeur pour le 20ème et 21ème siècle. D'où la nécessité d'une recherche active et efficace dans le domaine.

Différents modèles mathématiques ont donc été proposés, ils utilisent pour la plupart une dynamique compartimentale cherchant à modéliser l'interaction homme-moustique. Un modèle général avec une population de taille constante et faisant face à un seul virus a été examiné par Esteva et Varga [Esteva 1998]. Ces modèles ont aussi cherché à évaluer l'impact des applications d'insecticides à très faible volume (ULV) sur les épidémies de dengue [Newton 1992].

Les auteurs ont également proposés un modèle avec une population humaine censée croître de façon exponentielle et avoir un taux d'infection constant [Esteva 1999], un autre avec deux sérotypes de virus et une population humaine variable [Esteva 2000], sans oublier l'étude de l'impact de la transmission verticale et l'alimentation interrompue sur la dynamique de la maladie [Esteva 2003]. Et en supposant que les épidémies de dengue sont fortement influencées par la quantité de pluie parmi d'autres facteurs environnementaux, un modèle avec une population vectorielle variable a été aussi considéré [Feng 1997].

Selon le rapport du groupe de travail scientifique sur les insectes et la santé humaine [Den 2002], l'idée de deux virus de dengue coexistant dans la même épidémie est controversé, les modèles mathématiques avec une population humaine constante ( $N_h$ ) et deux virus différents agissant à des intervalles de temps séparés ont été considérés par les auteurs dans des articles préexistants [Derouich 2003]. Sur cette base, le présent document traite d'une population humaine variable  $N_h(t)$ .

### 1.2.1 Modèle SIR-SI

Pour pouvoir considérer le modèle  $SIR - SI$  que nous détaillerons plus tard, on a été amené à considérer les hypothèses suivantes :

#### Hypothèses du modèle

1. Nous considérons un modèle compartimental, c'est-à-dire que chaque population est divisée en classes, et que chaque individu d'une population passe d'une classe à une autre avec un taux approprié. (S pour susceptibles, I pour infectés, R pour guéris).

[Feng 1997]

2. Plus d'un type de virus de la dengue ne peut exister à la fois. En effet, une infection par un des 4 stéréotypes produit une immunité temporaire aux autres stéréotypes. [Feng 1997] D'où la simulation d'une épidémie à la fois.
3. Un individu infecté peut être guéri. En effet, une infection par un des 4 stéréotypes peut produire en cas de survie une immunité permanente à ce même stéréotype. [Feng 1997] D'où l'existence d'un compartiment R (guéris) pour la population humaine.
4. On remarque aussi qu'un moustique infecté n'a pas d'immunité virale pour la dengue. Il peut donc transmettre le virus toute la durée de sa vie. [Feng 1997] D'où l'absence d'un compartiment R (guéris) pour la population des moustiques.
5. Pour des raisons de clarté et de simplicité, on suppose que la période de latence n'est pas cruciale pour l'interaction Susceptible-Infectés [Derouich 2003]. D'où l'absence d'un compartiment 'Exposed' pour la population humaine et des moustiques.

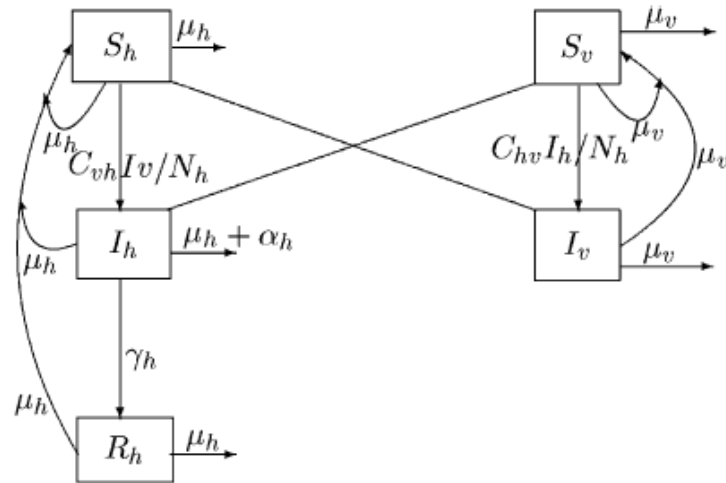


FIGURE 1.2 – Schéma du modèle compartimental du modèle SIR-SI

La durée de vie de l'homme est prise égale à 25 000 jours (68,5 ans), et celui du vecteur est de 4 jours.

### Paramètres de base

La définition et la valeur des paramètres utilisés dans ce modèle sont données ci-dessous :

Name of the parameter	Notation	Base value
Transmission probability of vector to human	$p_{hv}$	0.75
Transmission probability of human to vector	$p_{vh}$	0.75
Bites per susceptible mosquito per day	$b_s$	0.5
Bites per infectious mosquito per day	$b_i$	1.0
Effective contact rate, human to vector	$C_{hv}$	0.375
Effective contact rate, vector to human	$C_{vh}$	0.75
Human life span	$\frac{1}{\mu_h}$	25000 days
Vector life span	$\frac{1}{\mu_v}$	4 days
Host infection duration	$\frac{1}{\mu_h + \gamma_h}$	3 days

FIGURE 1.3 – Définitions et valeurs de paramètres basiques utilisés dans les simulations

## Equations du modèle

D'après (fig. 1.2), on peut écrire les équations de notre modèle :

$$\frac{dS_h}{dt} = \mu_h N_h - \left( \mu_h + p + C_{vh} \frac{I_v}{N_h} \right) S_h \quad (1.1)$$

$$\frac{dI_h}{dt} = \left( C_{vh} \frac{I_v}{N_h} \right) S_h - (\mu_h + \gamma_h) I_h \quad (1.2)$$

$$\frac{dR_h}{dt} = p S_h + \gamma_h I_h - \mu_h R_h \quad (1.3)$$

$$\frac{dS_v}{dt} = \mu_v N_v - \left( \mu_v + C_{hv} \frac{I_h}{N_h} \right) S_v \quad (1.4)$$

$$\frac{dI_v}{dt} = \left( C_{hv} \frac{I_h}{N_h} \right) S_v - \mu_v I_v \quad (1.5)$$

sous la contrainte :

$$S_h + I_h + R_h = N_h \quad (1.6)$$

$$S_v + I_v = N_v \quad (1.7)$$

Les paramètres  $\mu_h, \mu_v, C_{vh}, C_{hv}, \gamma_h$  sont données en  $jours^{-1}$

## Validation du modèle

Si on prend le jeu de données suivant :

$$\mu_h = \mu_v = C_{vh} = C_{hv} = \gamma_h = 0 \text{ et } p = 0$$

Analytiquement, on devrait avoir des variables constantes donc si on prend à  $t=0$  :

$$S_h(0) = 10000, I_h(0) = 1, R_h(0) = 0, S_v(0) = 50000, I_v(0) = 1$$

On a alors en traçant l'évolution de  $I_h$  et  $I_v$  en fonction du nombre de jours :

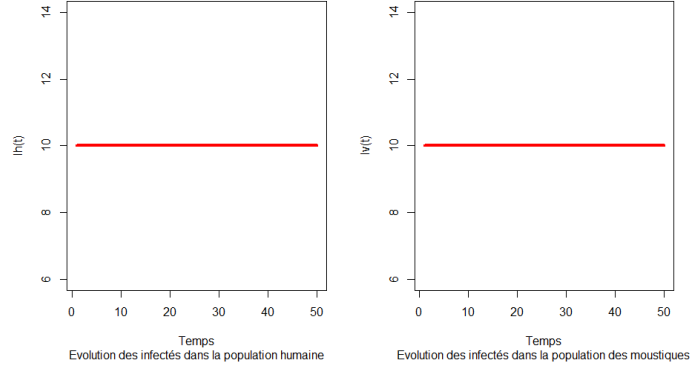


FIGURE 1.4 – Validation (1) de l'implémentation du modèle SIR-SI

Egalement, si on prend le jeu de données qui avec la même population initiale que l'exemple précédent :

$$\mu_h = \mu_v = C_{vh} = C_{hv} = 0, p = 0, \gamma_h = 0.3$$

Ainsi, on a :

$$\frac{dI_h}{dt} = -\gamma_h I_h \quad \text{et} \quad \frac{dI_v}{dt} = 0$$

Donc,

$$I_h(t) = I_h(0) \exp(-\gamma_h t)$$

On a alors en traçant l'évolution de  $I_h$  et  $I_v$  en fonction du nombre de jours :

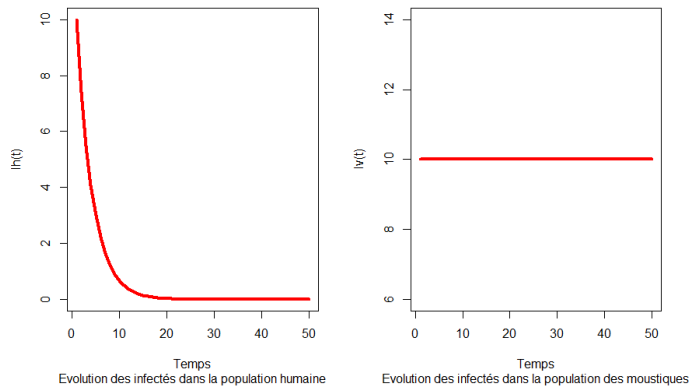


FIGURE 1.5 – Validation (2) de l'implémentation du modèle SIR-SI

Les 2 cas asymptotiques du modèle valident bien l'implémentation.



Ensuite, on peut prendre le jeu de paramètres basiques (fig. 2.2), et simuler notre modèle SIR-SI représentatif de l'épidémie de la dengue :

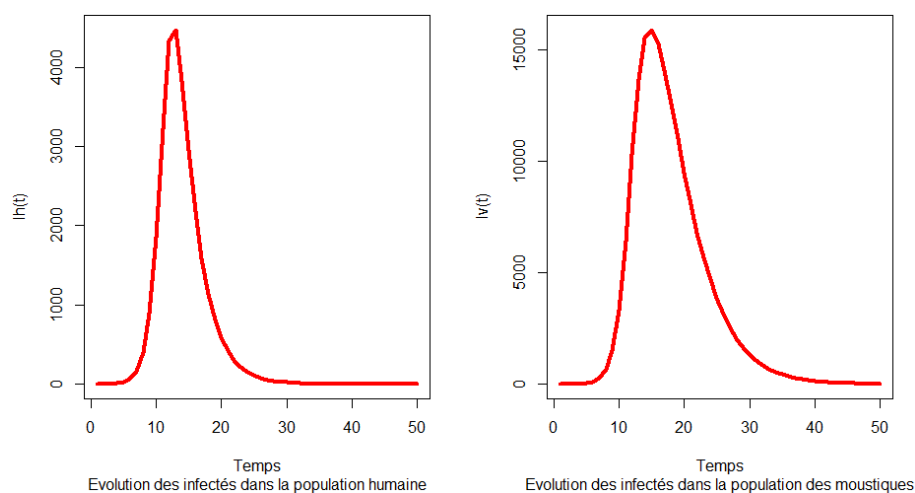


FIGURE 1.6 – Simulation du jeu de paramètres basiques

On peut aussi simuler l'influence d'une intervention par exemple la vaccination :

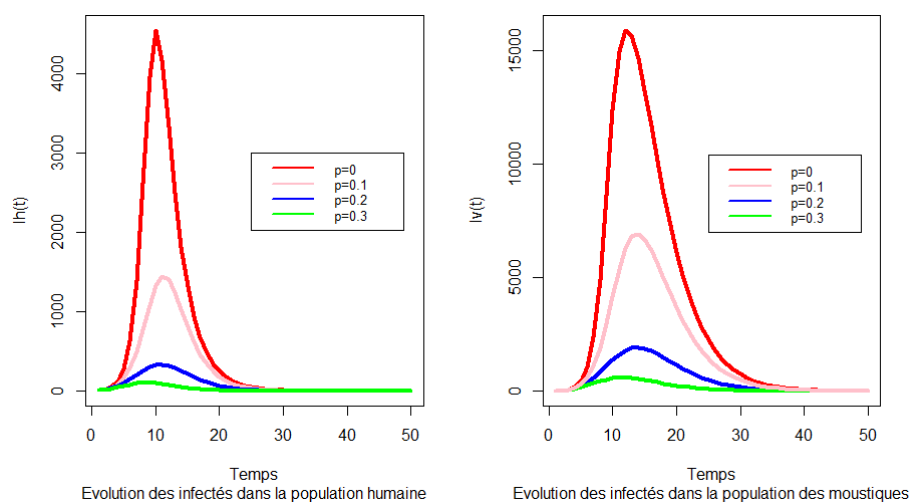


FIGURE 1.7 – L'influence de la vaccination sur le modèle SIR-SI

On voit bien que plus le taux de vaccination augmente de 0 à 1, plus l'épidémie de la dengue est contrôlée. Remarquons, qu'un taux de 0.3 permet pratiquement d'éradiquer l'épidémie. Ce taux est relativement faible par rapport aux taux observés chez d'autres épidémies. La vaccination est donc très attrayante pour répondre à cette épidémie.

### 1.3 Analyse du modèle

Dans un premier temps, nous avons conduit une analyse de la sensibilité globale du modèle selon la méthode de Morris pour identifier les facteurs influents et classer les facteurs selon leur influence. [Iooss. 2011] Malgré la difficulté pour séparer les interactions des effets non linéaires, cette méthode s'avère relativement efficace en résultats et économe en nombre d'appels au modèle. Une fois appliquée à notre modèle, l'analyse de Morris nous donne :

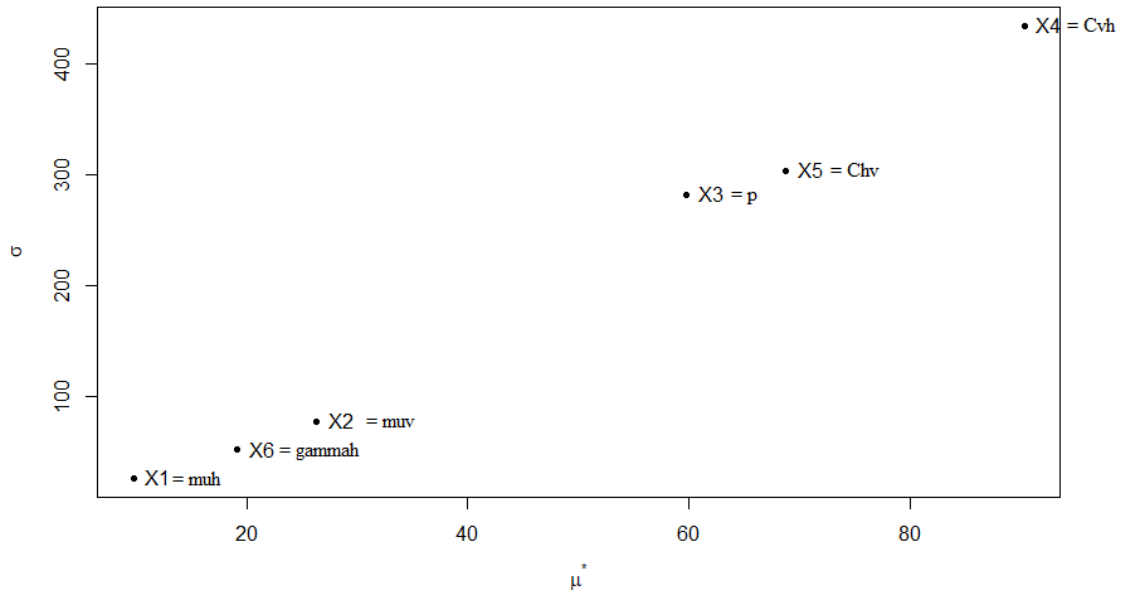


FIGURE 1.8 – Méthode de Morris pour le modèle SIR-SI

Sur ce graphique,  $\mu^*$  est la moyenne de la valeur absolue des effets élémentaires et  $\sigma$  l'écart type des effets élémentaires. La première évalue l'impact total de la variable tandis que la seconde mesure l'importance des interactions [Iooss. 2011]. Cette méthode est avant tout une méthode de criblage qui permet d'identifier les entrées importantes du modèle.

En réitérant le calcul plusieurs fois, on remarque que l'algorithme converge plutôt bien, avec les paramètres  $\mu_h, \mu_v, \gamma_h$  qui ont dans la majorité des cas un poids dans  $I_h(t)$  nettement négligeable devant les paramètres  $C_{hv}, C_{vh}$  et  $p$ . En effet, on montre que les paramètres qui sont près de l'origine ont un poids dans le modèle plus faible que ceux qui s'en écartent dans la diagonale. [Iooss. 2011]

Pour des problèmes d'estimation de ces paramètres, on peut donc fixer  $\mu_h, \mu_v, \gamma_h$  autour de valeurs usuelles et chercher à estimer  $C_{hv}, C_{vh}$  avec  $p$  le taux de vaccination étant choisi à priori. On remarque également que d'après cette analyse de sensibilité, le taux de vaccination a un effet non négligeable sur le modèle par rapport aux autres variables, d'où l'intérêt de creuser cette possibilité.

On peut également appliquer l'analyse de sensibilité globale basée sur la variance par méthode de Sobol. Travaillant dans un cadre probabiliste, elle décompose la variance de la sortie du modèle ou du système en fractions qui peuvent être attribuées à des entrées ou à des ensembles d'entrées. [Sobol 2001]

Ici encore, deux indices de sensibilité sont calculés : l'indice  $S_i$  qui mesure l'effet du premier ordre de la variable  $i$  sur  $I_h(t)$  ("main effect"). et l'indice  $S_{\tau_i}$  qui mesure l'effet total de la variable  $i$  sur  $I_h(t)$ . Ces indices sont calculés à l'aide d'approximation par Monte-Carlo de la décomposition fonctionnelle de la variance ("total effect") [Sobol 2001]

Une fois appliquée sur notre modèle, la méthode de Sobol nous donne :

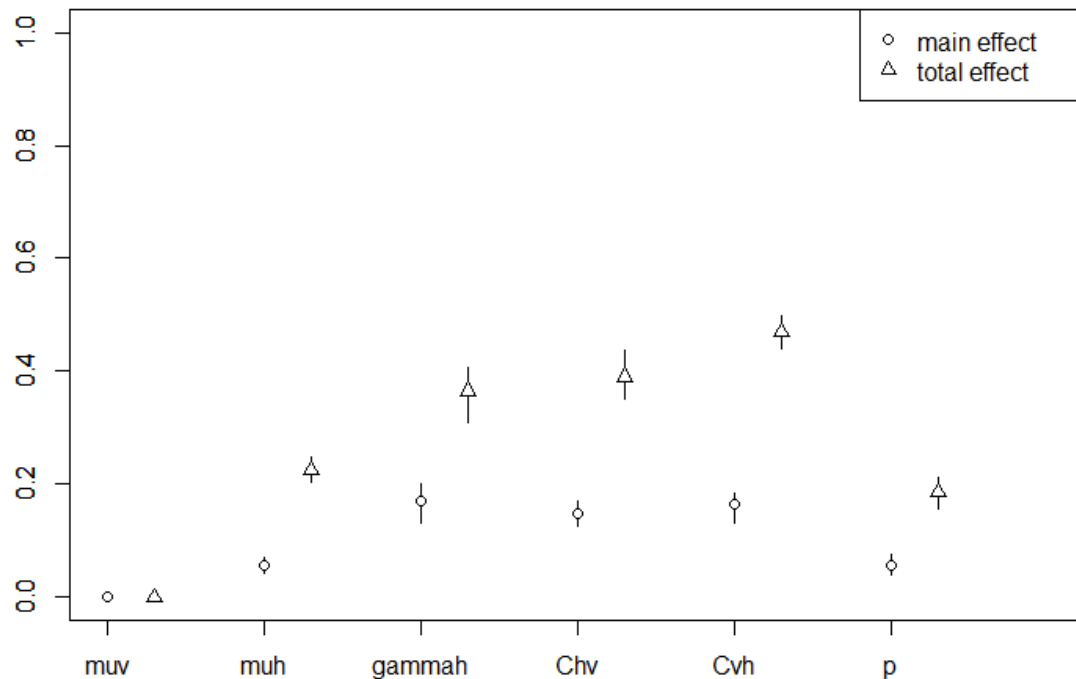


FIGURE 1.9 – Méthode de Sobol pour le modèle SIR-SI

On remarque donc qu'encore une fois  $C_{vh}$  et  $C_{hv}$  ont les effets de premier ordre et les effets globaux les plus importants. Il faut donc impérativement chercher à les estimer dans notre modèle puisqu'une petite variation de ces derniers entraîne une grande variation dans  $I_h(t)$ . En parallèle,  $\mu_h$  et  $\mu_v$  ont des effets très négligeables, et peuvent donc être fixées à des valeurs plausibles en fonction du cas étudié. Quant au taux de vaccination  $p$ , même s'il a un effet moindre ici par rapport à  $\gamma_h$ , il a le mérite d'être fixé et choisi judicieusement par les autorités pour freiner l'épidémie.

## 1.4 Estimation des paramètres

Suite à l'analyse de sensibilité, nous avons donc opté pour l'estimation des paramètres, dans un premier temps à fixer  $\mu_h, \mu_v, \gamma_h$  et aussi  $p = 0$  et dans un deuxième temps à estimer  $C_{vh}$  et  $C_{hv}$  qui représentent le taux de transmission homme à moustique et inversement. Cette estimation a été faite à l'aide de l'algorithme du gradient, qui opère une régression non-linéaire par force brute.

En effet, il cherche à minimiser la fonction d'erreur, qui est par défaut, l'erreur quadratique moyenne, en corrigeant chaque point par des termes en gradient. Ces termes sont pondérés par des coefficients d'apprentissage  $\alpha$  ou "learning rate".

Nous avons donc bruité (par un bruit gaussien) nos données simulées, et nous avons cherché à approximer nos paramètres pour essayer de reproduire uniquement dans un premier temps la courbe  $I_h(t)$ . On prendra ici la simulation avec les données basiques (fig. 1.6). Si on approxime sur l'ensemble de l'intervalle  $[0, 50 \text{ jours}]$ , nous avons :

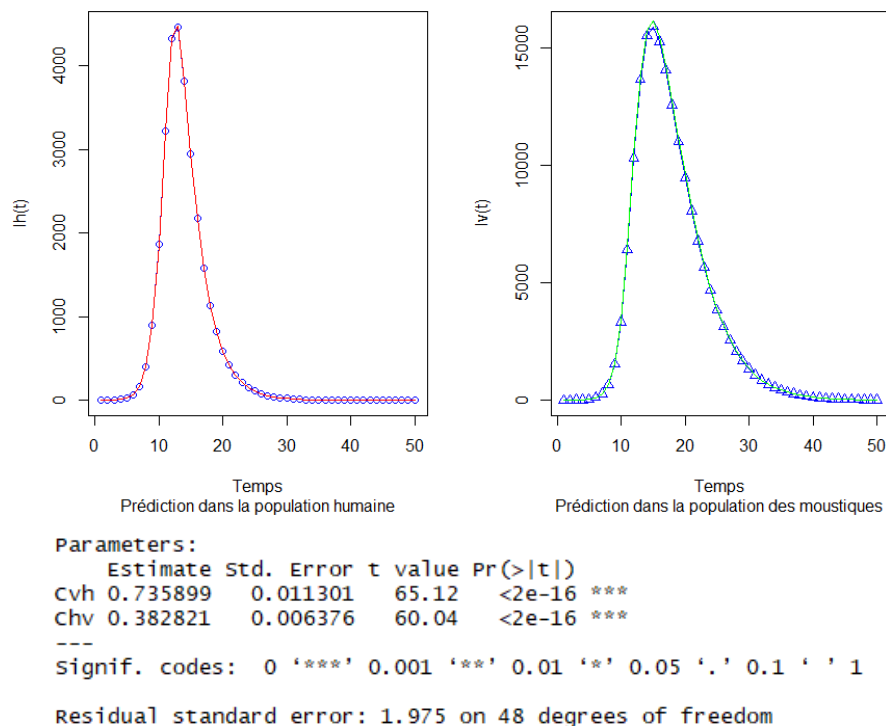


FIGURE 1.10 – Résultat d'approximation uni-populationnelle (1)

L'approximation dans ce cas est judicieuse, puisque les paramètres réels sont  $C_{vh} = 0.75$  et  $C_{hv} = 0.375$ . On arrive donc à bien approximer les 2 courbes. Cela peut s'expliquer par le fait lorsqu'on a les données d'une épidémie du début jusqu'à la fin, les degrés de liberté du système diminuent.

Par contre, si on approxime  $I_h(t)$  uniquement sur l'intervalle  $[0, 5 \text{ jours}]$ , on a :

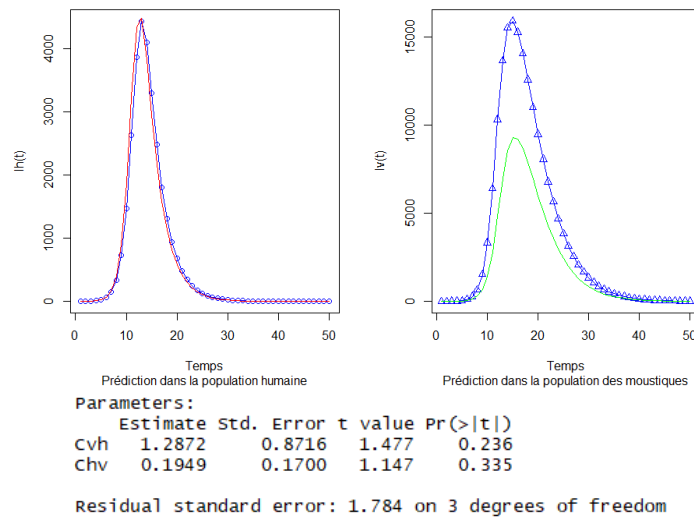


FIGURE 1.11 – Résultats d'approximation uni-populationnelle (2)

On remarque qu'avec ce modèle d'approximation, les paramètres estimés ne sont pas plausibles et on n'arrive pas à prédire l'évolution de l'épidémie pour les deux populations sur le long terme..

Ce qu'on peut donc essayer de faire c'est approximer cette fois-ci  $I_h(t)$  et  $I_v(t)$ , et on a pour une approximation sur  $[0, 5 \text{ jours}]$  :

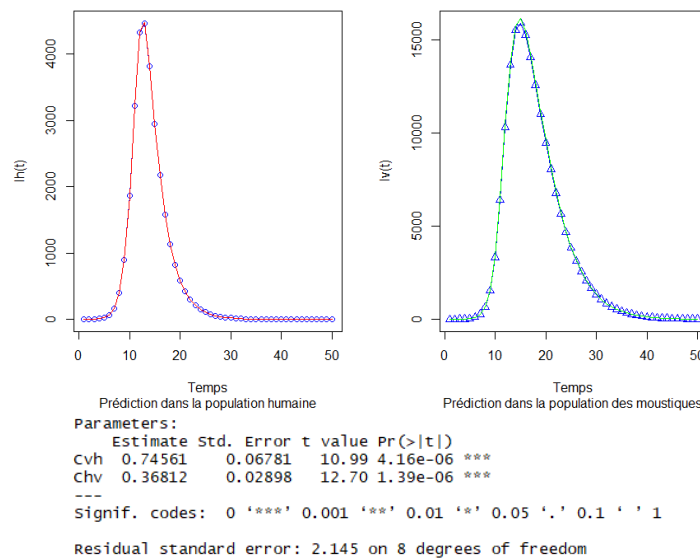


FIGURE 1.12 – Résultats d'approximation multi-populationnelle

Les résultats cette-fois-ci sont cohérents et appartiennent à l'intervalle souhaité centré

en  $(C_{hv}, C_{vh}) = (0.75, 0.375)$ . On peut donc affirmer que :

Dans ce modèle, il est nécessaire pour pouvoir prédire l'évolution de l'épidémie, de connaître  $I_h(t)$  et  $I_v(t)$ .

De surcroît, comme on approxime sur les 2 fonctions, il faudra normaliser les données pour que l'erreur quadratique moyenne soit autant influencée par la population des hommes infectés que par la population des moustiques infectées.

## 1.5 Discussion, perspectives

Ce modèle est, en effet, un premier modèle qui nous a permis de voir les différents paramètres influençants dans le modèle. Il nous a aussi permis de voir les contraintes liées à la prédiction d'une épidémie de la dengue avec un modèle SIR-SI puisqu'il faut suivre en parallèle les 2 populations.

Cependant, il est très limité parcequ'il étudie une seule épidémie de la dengue à la fois au court terme. En effet, une infection par la dengue est souvent suivie d'une infection par une dengue plus sévère, appelée la dengue hémorragique, qu'il faudra aussi modéliser dans le moyen terme.

De plus, même si nous avons souligné la nécessité d'une vaccination, celle-ci n'est pas du tout simple à mettre en oeuvre, puisqu'elle nécessite une vaccination aux 4 stéréotypes du virus. Un vaccin contre la dengue a été commercialisé par Sanofi Pasteur courant 2016. Les résultats affichés en Asie et en Amérique latine sont décourageants, ce qui a poussé le géant de l'industrie pharmaceutique à retirer son vaccin, et indemniser les gouvernements concernés.

Egalement durant ce comité régional européen de l'OMS, nous avons aussi eu le plaisir d'échanger avec Mme Boatto S., chercheur du département de Mathématiques dans l'Université Fédérale de Rio de Janeiro. Cet échange nous a permis de nous rendre compte que dans des milieux urbains très denses, le modèle SIR-SI présente des limites. Et des recherches montrent que dans ce cas, il n'est pas primordial de suivre la population des moustiques.

En effet, dans un milieu urbain, les interactions diffèrent et il faudra donc les adapter en fonction des différentes sous-populations et des différents quartiers. D'où l'idée de modéliser l'épidémie de la dengue par un modèle SIR avec interactions environnementales (modèle SIR-E). [Stolerman 2015]

# Chapitre 2

## Le choléra

### 2.1 Etude biologique, médicale et épidémiologique

Le choléra est une infection diarrhéique aiguë provoquée par l'ingestion d'aliments ou d'eau contaminés par le bacille *Vibrio cholerae*. Le choléra reste à l'échelle mondiale une menace pour la santé publique et un indicateur de l'absence d'équité et de l'insuffisance du développement social. Selon les estimations, il y a chaque année 1,3 à 4 millions de cas de choléra, et 21 000 à 143 000 décès dus à la maladie dans le monde.



FIGURE 2.1 – Bacille 'Vibrio Cholerae' dans l'eau

La grande incertitude sur les données épidémiologiques de la maladie est due à deux choses :

- Les pays atteints sont pour la plupart en développement et ne sont pas encore capable de collecter de manière précise des données épidémiologiques
- Certain pays ont tendance à estimer à la baisse certains chiffres pour ne pas nuire à l'attractivité de leur pays

## Symptômes

Le choléra est une maladie extrêmement virulente qui peut provoquer une diarrhée aqueuse aiguë et sévère. Les symptômes apparaissent entre 12 heures et 5 jours après l'ingestion d'aliments ou d'eau contaminée. Touchant les enfants comme les adultes, la maladie peut tuer en l'espace de quelques heures si aucun traitement n'est administré.

La plupart des sujets infectés par *V. cholerae* ne manifestent aucun symptôme, bien que le bacille soit présent dans leurs selles pendant 1 à 10 jours après l'infection et soit éliminé dans l'environnement, où il peut potentiellement infecter d'autres personnes.

Pour les personnes qui manifestent des symptômes, ceux-ci restent bénins à modérés dans la majorité des cas, tandis que chez une minorité, une diarrhée aqueuse aiguë, s'accompagnant de déshydratation sévère, se développe. En l'absence de traitement, elle peut entraîner la mort.

Le taux de létalité du choléra est très variable en fonction du développement des pays : de l'ordre de 2% dans les pays développés et jusqu'à 50% dans les pays les plus pauvres.

## Historique

Au XIX<sup>e</sup> siècle, le choléra s'est répandu dans le monde entier à partir de son réservoir original, dans le delta du Gange en Inde. Les 6 pandémies qui ont eu lieu par la suite ont tué des millions de personnes sur tous les continents. La pandémie actuelle (la septième) a démarré en Asie du Sud, en 1961, a atteint l'Afrique en 1971, puis les Amériques en 1991. Le choléra est désormais endémique (présent en permanence) dans de nombreux pays.

## Traitement

Pour traiter le choléra de la manière la plus efficace possible, il faut administrer au patient de sels de réhydratation orale (SRO). Les doses vont d'un sachet dissout dans 1L d'eau potable jusqu'à 6L de solution par perfusion

## Objectif

L'objectif de l'OMS est de faire baisser de 90% le nombre de décès dûs au choléra. Afin d'atteindre cet objectif, plusieurs moyens peuvent être mis en œuvre :

- Campagne de vaccination massive
- Plan d'assainissement et traitement des eaux
- Campagne de sensibilisation des populations dans les pays en développement
- Traitement des patients

[WHO 2018a]



## 2.2 Modélisation du choléra : Modèle de Codeço

De nombreux modèles compartimentaux existent dans la littérature pour décrire le fonctionnement du choléra. Ces modèles sont de complexité extrêmement variable et ne savent pas à quoi s'attendre pour le challenge annuel de l'OMS, nous avons décidé de choisir un juste milieu avec un modèle complet rendant compte des facteurs extérieurs dont l'importance est cruciale pour le choléra, à savoir le non assainissement des eaux.

Le modèle choisi est donc le modèle de Codeço [Guegan 2006] que nous allons maintenant décrire.

Le modèle de Codeço est régi par le système d'équation différentiel suivant :

$$\frac{dS}{dt} = -a\lambda(B)S \quad (2.1)$$

$$\frac{dI}{dt} = a\lambda(B)S - rI \quad (2.2)$$

$$\frac{dB}{dt} = B(n_b - m_b) + eI \quad (2.3)$$

avec  $\lambda(B) = \frac{B}{K+B}$

Il est décrit par un jeu de trois variables :

- $S$  : nombre de personnes susceptibles d'attraper la maladie
- $I$  : nombre de personnes infectées par la maladie
- $B$  : concentration en bactérie dans l'eau (en nombre de cellules par mL)

Et 5 paramètres :

- $a$  : taux d'exposition à l'eau contaminée en  $\text{jour}^{-1}$
- $K$  : concentration en *V.cholerae* dans l'eau dont l'ingestion donne le choléra avec 5% de chance
- $r$  : taux de guérison du choléra
- $l = n_b - m_b$  : différence entre le taux de croissance et le taux de mort de la bactérie
- $e$  : contribution de chaque personne infectée à la pollution de l'environnement (en nombre de cellules par mL par jour par personne)

H correspond à la population étudiée.

Il s'agit en fait de la représentation mathématique d'un modèle compartimental SIR-E représenté ci-après :

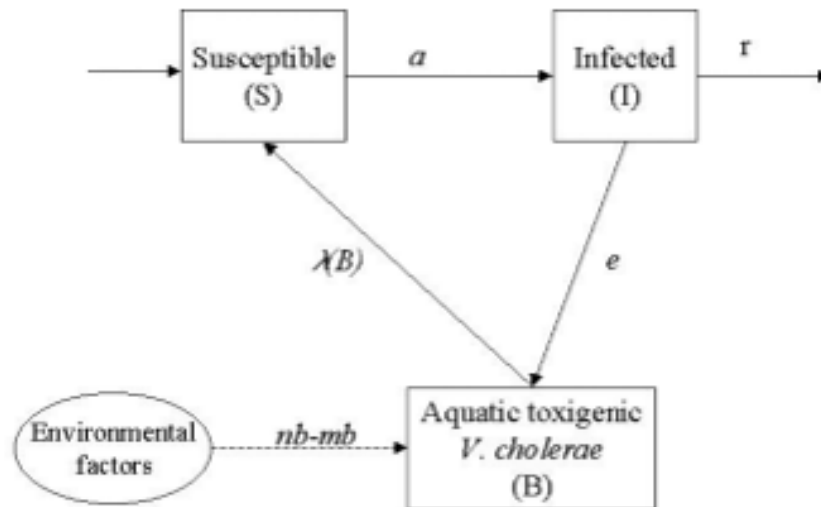


FIGURE 2.2 – Diagramme compartimental du modèle de Codeço [Mraidi 2014]

La première action a donc été d’implémenter ce modèle sur ordinateur et de le tester afin d’estimer les paramètres.

Pour ce faire, nous avons utilisé les données issues de l’article [Guegan 2006] d’où nous avons tiré le modèle de Codeço afin de pouvoir tester, valider et améliorer nos algorithmes par la suite. L’article fournissait une table des paramètres avec toutes les valeurs standards d’une situation réelle observée au Bangladesh :

Parameters	Community 1	Community 2	Community 3
Expected cholera dynamics	cholera-free	epidemic	endemic
$H$	10,000 ind.	10,000 ind.	10,000 ind.
$n$	0.0001	0.0001	0.001
$a$	0.5	1	1 %
$K$	$10^6$ cells/ml	$10^6$ cells/ml	$10^6$ cells/ml
$\tau$	0.2	0.2	0.2
$n_b - m_b$	-0.33	-0.33	-0.33
$e$	10 cells/ml	10 cells/ml	10 cells/ml

FIGURE 2.3 – Liste de paramètres pour 3 communautés distinctes (Modèle de Codeço)

Il fournissait également des simulations des différentes variables en fonction de la situation épidémique ou endémique du choléra. Ensuite, nous avons donc cherché à confirmer d’une part la valeur des paramètres et d’autre part l’allure des graphes en implémentant sur R les algorithmes de simulation et d’estimation.

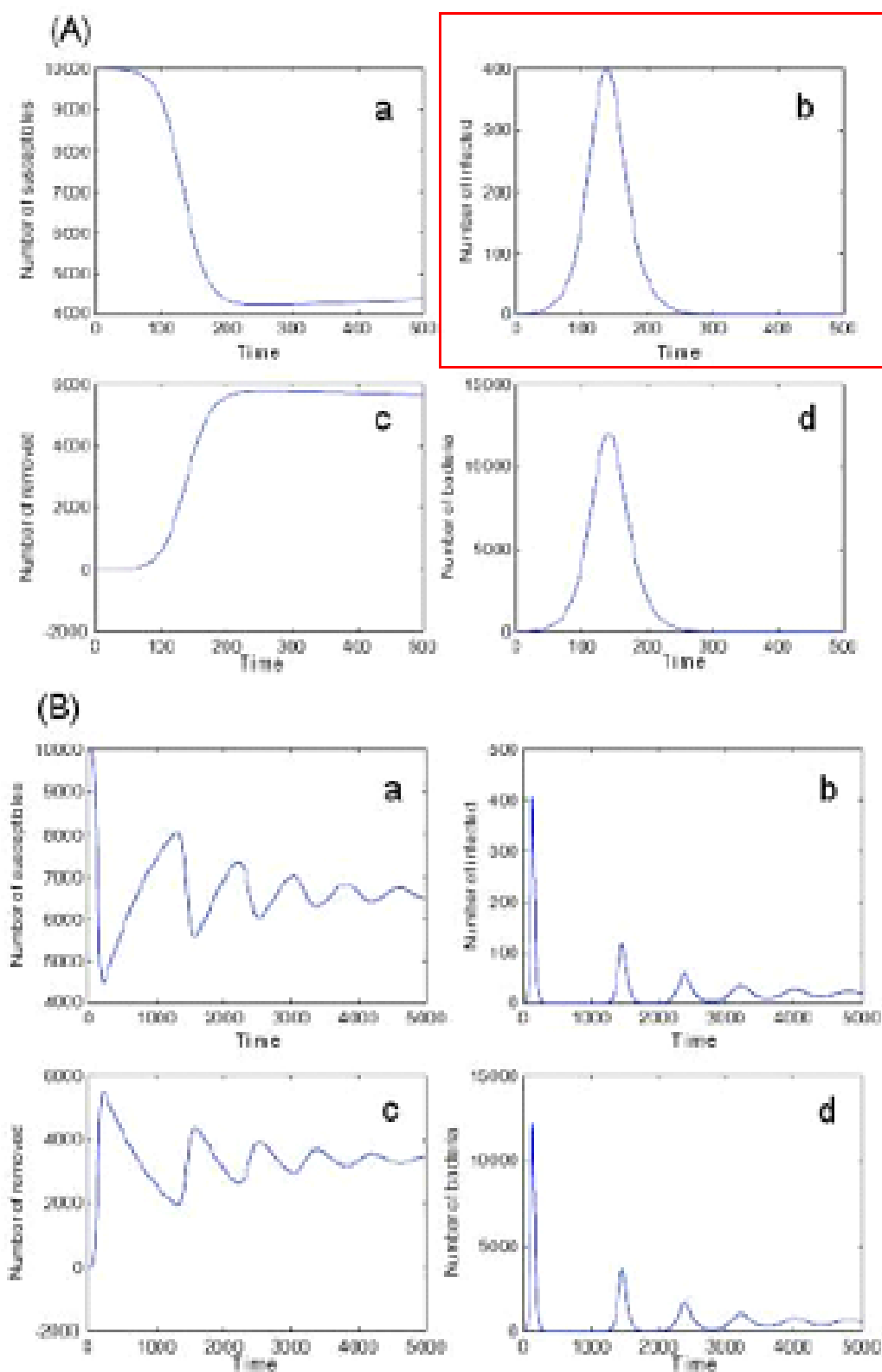


FIGURE 2.4 – L'évolution d'une situation épidémique (A) et d'une situation endémique (B) pour les 4 variables, (a) Susecptibles, (b) Infectés, (c) Guéris, (d) Concentration en bactérie dans le milieu aquatique

## 2.3 Simulation

On se place dans le cas d'une situation épidémique (car proche du futur challenge) pour les infectés. C'est le graphique A.b encadré en rouge ci-dessus.

En reprenant les paramètres proposés dans l'article, la simulation sur R fournit une allure comparable à la figure A.b.

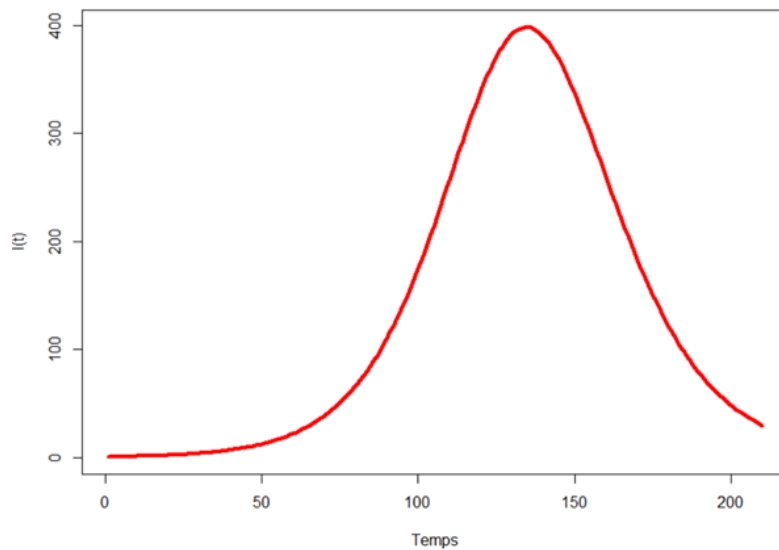


FIGURE 2.5 – Simulation (1) du modèle de Codeço - Evolution de  $I(t)$

## 2.4 Estimation

Afin de tester l'estimation des paramètres, on bruite nos données à l'aide d'un bruit gaussien centré et on affiche les valeurs obtenues après bruitage qu'on a aussi tronqué pour avoir plus de lisibilité. Ensuite, on essaye de retrouver les paramètres de bases.

t	I
1	1
30	10
60	20
90	95
120	320
150	390
180	92
210	33

FIGURE 2.6 – Données de simulation (1) bruitées

Après estimation des paramètres, on obtient une allure satisfaisante de la courbe :

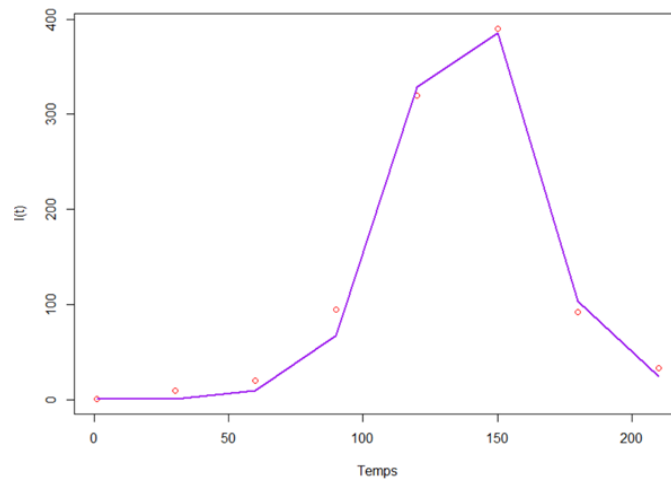


FIGURE 2.7 – Courbe de simulation (1) approximée

Et les valeurs des paramètres estimés sont :

```
Parameters:
  Estimate Std. Error t value Pr(>|t|)
a 1.749e+00 8.444e+01 0.021 0.9848
K 1.411e+06 1.177e+11 0.000 1.0000
r 3.528e-01 8.440e-02 4.180 0.0249 *
l 2.254e-02 1.785e-01 0.126 0.9075
e 3.062e+00 2.554e+05 0.000 1.0000
```

FIGURE 2.8 – Paramètres estimés de la simulation (1)

On retrouve bien des valeurs proches de celles fournies dans la table de l'article. [Guegan 2006]  
L'estimation est donc satisfaisante.

## 2.5 Analyse du modèle de Codeço

Afin de préparer au mieux le futur challenge, nous avons également mené une analyse de sensibilité afin de déterminer quels paramètres auraient le plus d'influence sur le modèle.

Voici ce que l'on peut observer suite à cette analyse dont le script est disponible en pièce jointe. La méthode utilisée est celle de Sobol.

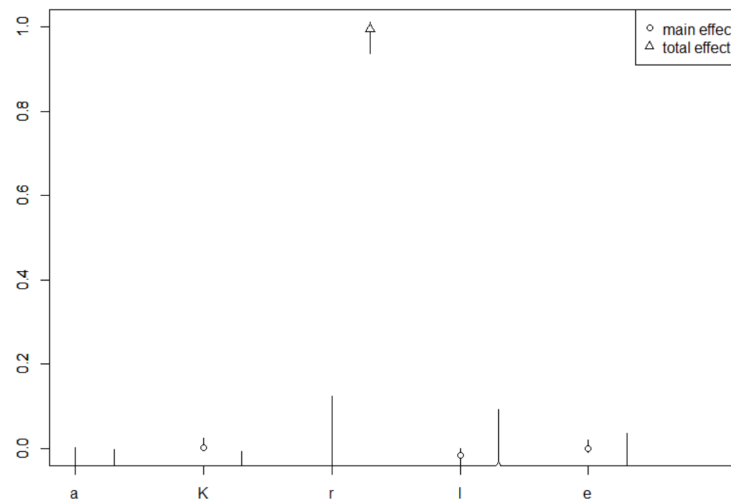


FIGURE 2.9 – Analyse de Sobol appliquée au modèle de Codeço

Les résultats sont résumés dans le tableau ci-après :

```
Call:
sobol2007(model = simulation, X1 = X1, X2 = X2, nboot = 30)

Model runs: 700

First order indices:
      original      bias std. error  min. c.i.  max. c.i.
a  7.382089e-05  3.318227e-01  9.548296e-01 -4.191957e+00  2.023326e-03
K -2.512713e-07 -1.602920e-03  5.226024e-03 -1.134156e-03  2.450564e-02
r  6.106073e-02  1.226173e+02  3.815692e+02 -1.585145e+03  1.249778e-01
l  5.565194e-06  1.537179e-02  5.687684e-02 -3.074803e-01  1.100857e-05
e  6.284824e-07 -4.328344e-04  4.070662e-03 -8.563752e-03  1.962455e-02

Total indices:
      original      bias std. error  min. c.i.  max. c.i.
a -1.626788e-03  0.13263163  0.23434837 -0.7640476 -0.001455786
K -8.540491e-03  0.12793225  0.19366506 -0.6321240 -0.008119285
r  1.011079e+00  0.01644580  0.01880356  0.9366462  1.011931393
l -2.419017e-03  0.04181089  0.10635053 -0.4851880  0.092245625
e -1.568945e-05  0.05325048  0.13030253 -0.4847700  0.035029246
> |
```

FIGURE 2.10 – Indices de Sobol pour le modèle de Codeço

On peut observer que le paramètre  $r$  comporte un effet significatif sur le nombre d'infectés. On aura donc tendance à faire varier ce paramètre pour les futures estimations et être particulièrement attentif à sa valeur. Les effets des autres paramètres sont cependant négligeables.

## Prédiction

Afin de préparer au mieux le challenge on a également codé une fonction permettant de renvoyer la valeur  $I_{pred}(J+k)$  correspondant au nombre d'infecté prévu  $k$  jour après le dernier jour de la base de données

## 2.6 Challenge de l'OMS

Le mercredi 17 janvier 2018, date du début du challenge de la modélisation, organisé par le comité régional européen de l'OMS, nous apprenons la situation préoccupante du Néyem qui rencontre une propagation du choléra dans 2 camps de réfugiés "Khoraz" et "Murkhazi". Certains éléments de la situation sont rassurants et d'autres le sont moins.

### Point positif :

- Un modèle SIR-E semble bien correspondre à la situation et la variable B apparaît dans le jeu de données même si peu de données sont disponibles.

### Point négatif :

- Certains facteurs comme le taux de scolarisation ou encore la météo semblent très importants dans la réalité de la situation tandis que nous n'avions pas prévu de les prendre en compte avec le modèle de Codeço.

Le challenge ayant été évolutif nous n'avons malheureusement pas gardé chaque script à chaque étape du challenge. On va tout de même tenter de montrer l'évolution du modèle et l'adaptation que nous avons faite du modèle de Codeço.

Lors de la première phase, faute de temps nous avons utilisé simplement le modèle de Codeço qui avait fourni des résultats corrects mais n'avait pas du tout prévu le regain du nombre d'infectés dans le camp de "Khoraz". C'est à partir de la phase deux, que nous avons décidé de prendre en compte les conditions météorologiques.

## 2.7 Prise en compte des conditions météorologiques

Nous avons tenté de décrire l'influence des conditions météo sur le modèle par l'intermédiaire d'une analogie avec les modèle prenant en compte la saisonnalité (la grippe par exemple). Pour cela, nous avons introduit deux nouveaux paramètres et agit sur les facteurs extérieurs en modulant la concentration B en bactérie dans le fleuve par une fonction trigonométrique ad hoc. Les deux paramètres correspondent à l'amplitude et la période de cette fonction trigonométrique, et ont été estimés à première vue par l'observation des données météorologiques fournies en phase 2. Le nouveau système d'équation devient alors :

```
deriv <- function(t,y,p)
{
  with(as.list(c(p, y)), {
    dS <- -a*(B/(K+B))*S*tan((1+m*sin((2*pi/g)*t)))
    dI <- a*(B/(K+B))*S*tan((1+m*sin((2*pi/g)*t)))-r*I
    dB <- -B*I+e*I
    list(c(dS,dI,dB))
  })
}
```

FIGURE 2.11 – Modèle de Codeço avec conditions météorologiques

Le modèle fonctionnait très bien pour prédire les données de la phase 2 (malheureusement le script avec l’affichage de la courbe on été modifiés et la version originale perdu donc seul le script de la phase 3 reste disponible).

Pour la phase 3, l’estimation des paramètres nous donne l’allure suivant :

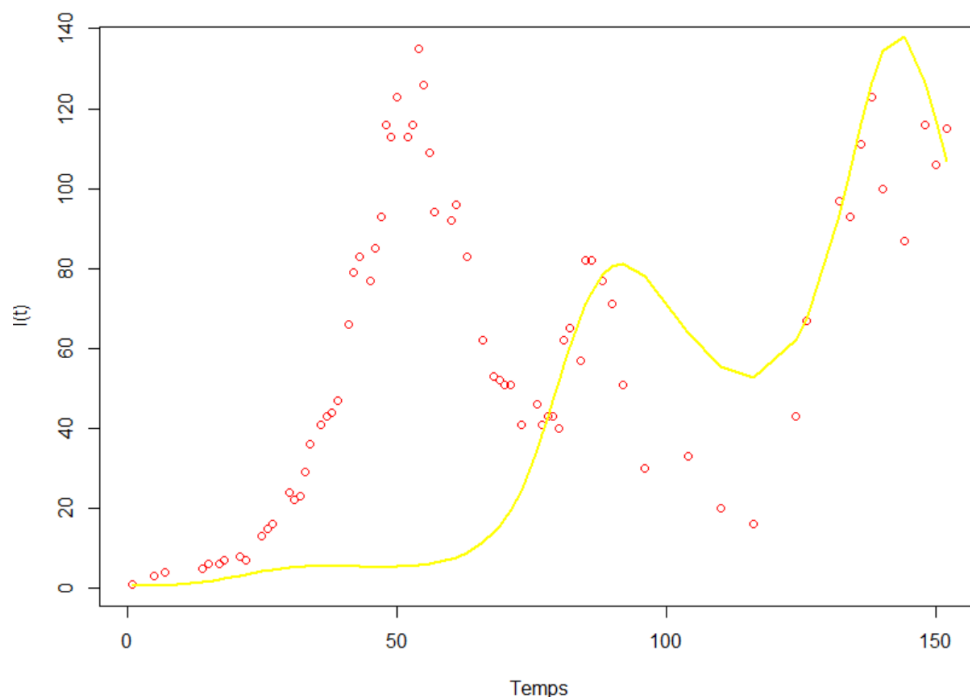


FIGURE 2.12 – Modèle de Codeço avec conditions météorologiques

## 2.8 Discussion, perspectives

### Prise en compte de la météo

Comment on peut le constater sur le graphique ci-dessus, le modèle est très perfectible mais nous avons néanmoins pu rendre compte de certains regains dûs à la météo. Plus de temps nous aurait certainement permis d’affiner les paramètres et d’obtenir un modèle d’autant plus satisfaisant.

En particulier une réflexion plus poussée sur le choix de la fonction trigonométrique rendant au mieux compte des conditions météorologiques aurait sans doute fait la différence.



Les données météorologiques fournies prenaient la forme suivante :

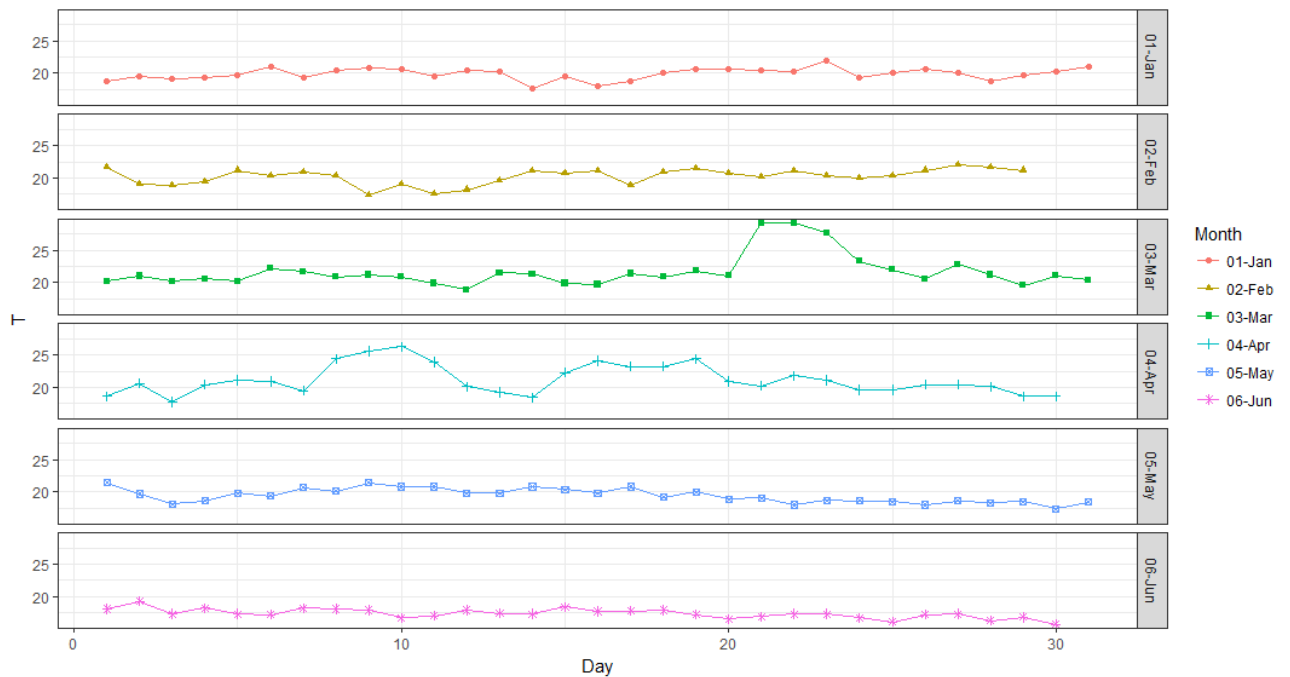


FIGURE 2.13 – Profil de la Température au Nékem (2018)

Une amélioration du modèle serait donc de trouver une fonction correspondant de façon plus précise à la variation des températures sur la période du challenge.

### Autres évolutions possibles

Afin d'améliorer le modèle, on aurait également pu :

- Prendre en compte encore plus de paramètres tels que le taux de scolarisation, la pluie, la température, un paramètre lié au fleuve qui relierait les populations infectées des deux camps
- Faire l'étude de la simulation des paramètres en utilisant les équations différentielles en B car on avait des informations à chaque phase de l'épidémie sur la concentration en bactérie dans le fleuve
- Utiliser des modèles à retard où le temps d'immunité est pris en compte (le temps apparaît dans les équations différentiels)
- Mieux appréhender la variable  $C$  correspondant aux déclarations de cas symptomatiques



# Chapitre 3

## La grippe saisonnière

### 3.1 Etude biologique, médicale et épidémiologique

La grippe saisonnière est une infection virale aiguë provoquée par un virus grippal. Il existe 3 types de grippe saisonnière – A, B et C. La grippe serait apparue chez les oiseaux il y a environ 6 000 ans, la grippe humaine serait née vers -2500 en Chine avec le développement de la domestication des oiseaux, notamment des canards qui constituent le réservoir important des gènes viraux.

D’après l’OMS, elle est responsable dans le monde d’une morbidité élevée et de 250 000 à 500 000 décès par an, essentiellement de jeunes enfants et des personnes âgées. [WHO 2018c] En France, la grippe touche chaque année entre 2 et 8 millions de personnes selon le Groupe d’expertise et d’information sur la grippe et provoque entre 1500 et 2000 morts, essentiellement chez les personnes de plus de 65 ans.

Dans les pays en développement l’impact de la grippe est assez mal connu. Sous les tropiques, il est admis que les flambées épidémiques tendent à avoir des taux d’atteinte et de mortalité élevés. Le coût sanitaire et social annuel de la grippe est considérable : 460 millions d’euros en France pour une épidémie moyenne.

### Symptômes

Les symptômes apparaissent de 1 à 4 jours après la contamination. L’infection dure généralement une semaine et se caractérise par l’apparition brutale d’une forte fièvre, de douleurs musculaires, de maux de tête, d’une sensation de profond malaise, d’une toux sèche, d’une gorge irritée et d’une rhinite.

La plupart des sujets atteints guérissent en une à deux semaines sans traitement médical. La grippe fait courir des risques sérieux aux plus jeunes, aux personnes âgées et aux ma-

lades souffrant de certaines pathologies (pneumopathies, sida, diabète, cancer, problèmes cardiaques ou rénaux, notamment), chez qui elle peut provoquer de graves complications, des pathologies concomitantes, une pneumonie et même la mort. [Pasteur 2018]

## Transmission et prévention

### Une transmission par voie respiratoire

Les virus grippaux pénètrent dans l'organisme par voie respiratoire, au niveau du rhinopharynx. Ils se transmettent facilement d'une personne à l'autre par voie aérienne, au moyen des microgouttelettes et des particules excrétées par les sujets infectés lorsqu'ils parlent, toussent ou éternuent.

Les sujets atteints deviennent contagieux un jour avant l'apparition des premiers symptômes et le restent pendant 7 jours. La maladie se propage rapidement, en particulier quand il y a de fortes concentrations de populations (transports en commun, collectivités).

Les virus de la grippe survivent plus longtemps à l'extérieur de l'organisme lorsque les températures chutent, c'est la raison pour laquelle les épidémies saisonnières surviennent en hiver dans les climats tempérés. [Pasteur 2018]

### Quelques mesures de prévention en cas d'épidémie

En cas d'épidémie déclarée, des mesures de protection individuelle, comme le port d'un masque, s'imposent pour éviter d'être infecté ou d'infecter soi-même les personnes de son entourage proche. La contamination peut aussi avoir lieu lorsqu'une personne touche une surface contaminée - bien que les virus grippaux ne survivent que quelques heures sur une surface inerte -, puis porte sa main à proximité du nez, d'où l'importance du lavage des mains réguliers en situation épidémique. [Pasteur 2018]

## Traitement

### Les antiviraux classiques

On dispose depuis plusieurs années de molécules qui agissent en empêchant la réplication du virus. Ces dernières sont à l'origine de traitements appelés antiviraux, qui diffèrent selon la pharmacocinétique, les effets secondaires qu'ils produisent, les voies d'administration, les tranches d'âge ciblées, les posologies et les coûts. Cependant, les virus grippaux récents ayant développé une résistance à ces antiviraux, leur utilisation est aujourd'hui limitée.

Pendant plusieurs années, l'amantadine et la rimantadine ont été les seuls antiviraux. Même s'ils présentent l'avantage d'être relativement peu onéreux, ces deux médicaments

ne sont efficaces que contre les virus de type A et peuvent avoir de graves effets secondaires (délires et convulsions survenant surtout à forte dose chez les personnes âgées). Lorsqu'ils sont utilisés à plus faible dose, pour la prophylaxie en cas de pandémie, ces effets sont beaucoup moins susceptibles de survenir. [Pasteur 2018]

### **Les inhibiteurs de la neuraminidase**

Ces dernières années, pour palier le développement de résistances des virus grippaux aux molécules existantes, une nouvelle classe d'antiviraux, les inhibiteurs de la neuraminidase (protéine de surface des virus grippaux), a fait son apparition. À l'image du zanamivir (Relenza®) et de l'oseltamivir (Tamiflu®), ces derniers entraînent moins d'effets secondaires que les antiviraux classiques, même si le zanamivir peut exacerber l'asthme et d'autres affections respiratoires chroniques. Ces médicaments sont cependant onéreux, ce qui limite leur utilisation dans un grand nombre de pays. [Pasteur 2018]

## **Vaccination**

### **Une vaccination annuelle efficace**

Les modifications génétiques constantes des virus grippaux imposent d'ajuster chaque année la composition du vaccin pour y introduire les souches les plus récentes en circulation. Si l'efficacité du vaccin dépend avant tout de l'âge et de l'état immunitaire du sujet vacciné, le degré de similitude entre les souches vaccinales et les virus en circulation entre donc également en ligne de compte.

Chez les personnes âgées, la vaccination diminue la morbidité de 60 % et la mortalité de 70 à 80 %. Chez les adultes en bonne santé, le vaccin est très efficace (diminution de la morbidité de 70 à 90 %) et limite nettement les conséquences sanitaires et économiques d'une pandémie.

### **Mise au point des vaccins**

Le réseau mondial de l'OMS pour la surveillance de la grippe établit chaque année la composition du vaccin. Il rassemble les Centres Collaborateurs OMS du monde entier et il est chargé de contrôler les virus en circulation chez l'homme et d'identifier rapidement les nouvelles souches. Sur la base des informations recueillies par le réseau, l'OMS recommande chaque année la composition du vaccin pour qu'il soit efficace contre les trois souches les plus récentes en circulation. [Pasteur 2018]

## 3.2 Modélisation de la maladie

### Choix du modèle

De nombreux modèles par compartiments existent dans la littérature pour caractériser la propagation de la grippe en général et saisonnière en particulier. Pour modéliser la grippe saisonnière, et puisque les enfants sont plus affectés par la grippe que les autres, nous avons décidé de diviser la population en 2 groupes. Un premier pour les enfants de moins de 15 ans et le deuxième pour les plus de 15 ans. Cette décision de répartition se base en partie sur le graphe suivant :

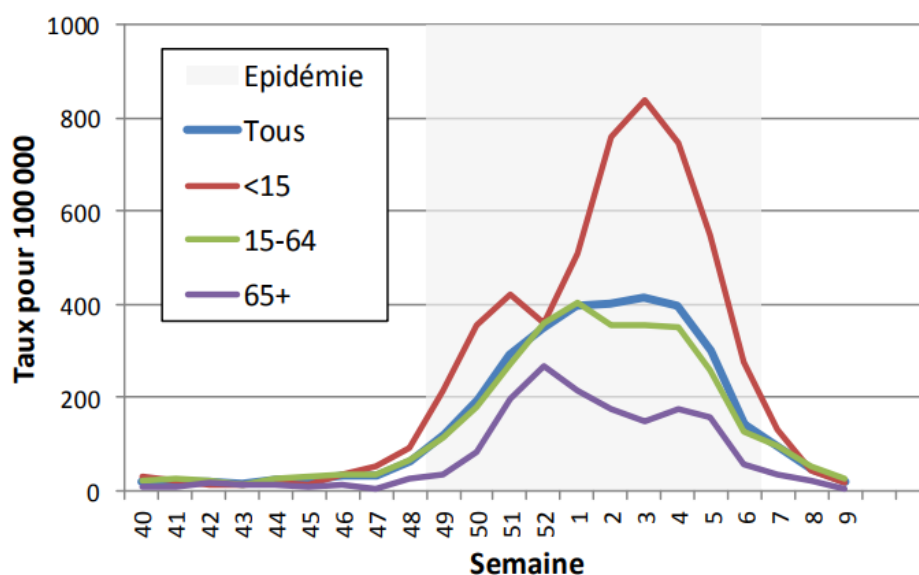


FIGURE 3.1 – Évolution hebdomadaire des taux de consultations pour syndrome grippal par groupe d'âge, saison 2016- 2017, France métropolitaine [Bourdillon 2017]

### Modèle retenu

Comme le virus de la grippe est facilement transférable, il nous a paru judicieux d'avoir un compartiment E (exposé) dans notre modélisation. Notre modélisation ne prend en compte qu'une seule année, nous avons donc fait l'hypothèse qu'une fois une personne guérit de la maladie, elle est immunisée (puisque une fois guérie, la personne est effectivement immunisée pour une année). Pour simplifier notre problème, nous partons d'une situation où personne n'est immunisé contre la maladie. Tout cela nous mène à choisir une modélisation en compartiments double SEIR. Le modèle double SEIR est régi par le système

d'équations différentielles suivant [Laguzet 2015] :

$$\frac{dS_j}{dt} = -\lambda_j S_j(t) \quad (3.1)$$

$$\frac{dE_j}{dt} = \lambda_j S_j(t) - \delta E_j(t) \quad (3.2)$$

$$\frac{dI_j}{dt} = \delta E_j(t) - \gamma I_j(t) \quad (3.3)$$

$$\frac{dR_j}{dt} = \gamma I_j(t) \quad (3.4)$$

avec  $\lambda_j = a_j (\beta_A I_A(t) + \beta_B I_B(t))$  le taux de contamination

$$S_j(0) = S_0^j, E_j(0) = E_0^j, I_j(0) = I_0^j \text{ et } R_j(0) = R_0^j$$

Notre modèle est décrit par un jeu de quatre variables

- S : nombre de personnes susceptibles d'attraper le virus (les non immunisées)
- E : nombre de personnes exposées à la maladie (dans un environnement infecté)
- I : nombre de personnes infectées par le virus
- R : nombre de personnes guéries et donc immunisées)

Et 5 paramètres :

- $\delta$  : Le temps d'incubation du virus, il est le même pour les deux tranches d'âge
- $\gamma$  : Le temps de guérison. Dans la littérature, il est souvent mentionné que le temps de guérison est presque le même pour tous les âges, donc nous en avons fait l'hypothèse.
- $\beta_A$  : taux d'exposition ou de contact pour les personnes de moins de 15 ans
- $a_a, a_b$  : capacité à contracter la maladie. Si  $a_a > a_b$  : cela indique que les personnes de la classe A attrapent plus la maladie que les personnes de la classe B à situation identique.

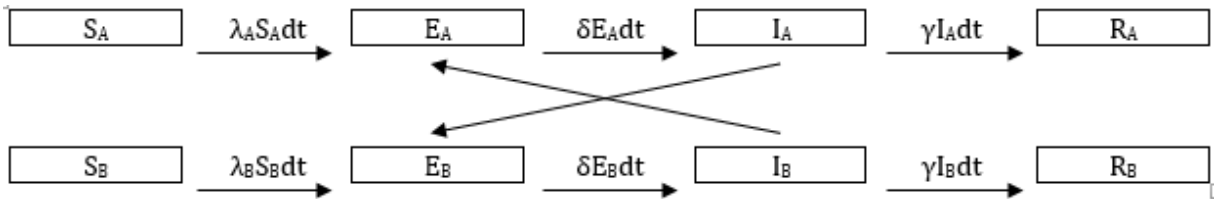


FIGURE 3.2 – Diagramme compartimental de la grippe saisonnière

La première action a donc été d'implémenter en machine ce modèle afin d'estimer les paramètres plus tard. Pour ce faire, nous avons vérifié l'exactitude de notre modèle en le testant avec les valeurs fournies par l'article scientifique, pour vérifier que notre algorithme retourne bien les mêmes courbes que celles de l'article.

### 3.3 Simulation

On se place dans le cas d'une situation épidémique comme dans l'article. En reprenant les paramètres proposés, et les valeurs initiales suivantes :  $S_A = 0.82, E_A = 0, I_A = 0.06, R_A = 0, S_B = 0.11, E_B = 0, I_B = 0.01, R_B = 0$ , la simulation sur R fournit une allure comparable aux figures de l'article. (Cf Script en pièce jointe « Modèle double SEIR ») :  
 Nb : Les unités sont en pourcentage (pour ne pas avoir à fixer le nombre de population)

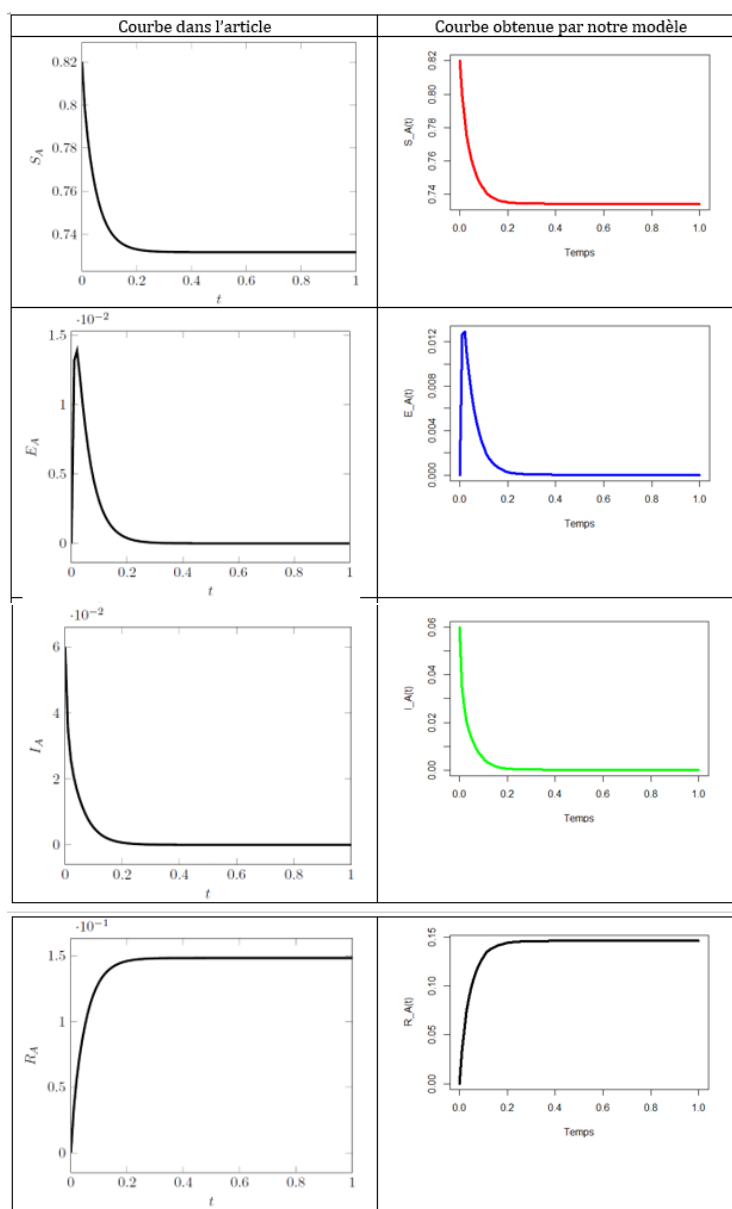


FIGURE 3.3 – Simulation (1) du modèle de la grippe saisonnière



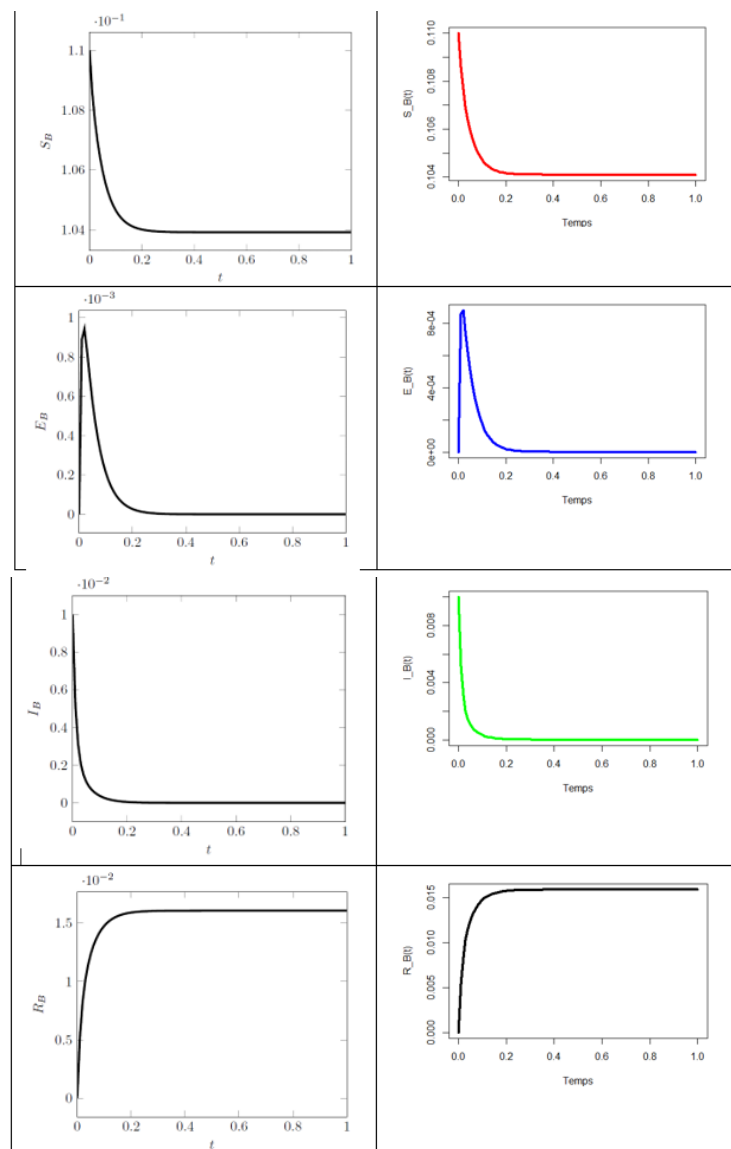


FIGURE 3.4 – Simulation (1)-suite du modèle de la grippe saisonnière

Nous remarquons que seule l'allure de la courbe d'infectés change entre les 2 groupes, cela est en accord avec le fait que les coefficients  $\delta$  et  $\gamma$  soient les mêmes pour les 2 groupes.

### 3.4 Analyse de sensibilité

Afin de pouvoir mieux estimer nos paramètres, nous avons mené une analyse de sensibilité afin de déterminer quels paramètres aurait le plus d'influence sur le modèle. La méthode utilisée est celle de Morris. Voici ce qu'on peut donc observer :

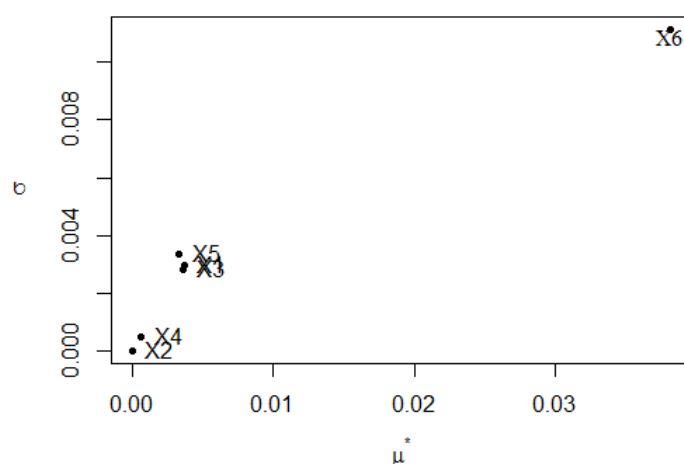


FIGURE 3.5 – Méthode de Morris appliquée à la grippe saisonnière

Les valeurs exactes étant :

```
call:
morris(model = simulationM, factors = 6, r = 50, design = list(type =
"oat", levels = 10, grid.jump = 1), binf = 0, bsup = 1)

Model runs: 350
```

	mu	mu.star	sigma
x1	3.448840e-03	3.448840e-03	2.515408e-03
x2	4.597467e-06	4.597467e-06	8.034107e-06
x3	2.979079e-03	2.979079e-03	2.745665e-03
x4	4.705678e-04	4.705678e-04	4.149203e-04
x5	4.227396e-03	4.227396e-03	4.000116e-03
x6	-3.916917e-02	3.916917e-02	1.097028e-02

FIGURE 3.6 – Résultats de la méthode de Morris

Nous remarquons que le paramètre le plus important ici est  $X6 = \gamma$ , le facteur de guérison. Ceci paraît logique puisqu'une fois guérie, la personne ne retombe plus malade. Si on prend une très grande valeur, on remarque que tout se passe plus rapidement, le processus SEIR converge en moins de temps. Gamma est alors un facteur dit « absorbant » et c'est pourquoi il est prédominant.

## 3.5 Estimation

Afin de tester l'estimation des paramètres, nous avons pris le modèle simulé, que nous avons bruité progressivement puis essayé de retrouver le signal de base :

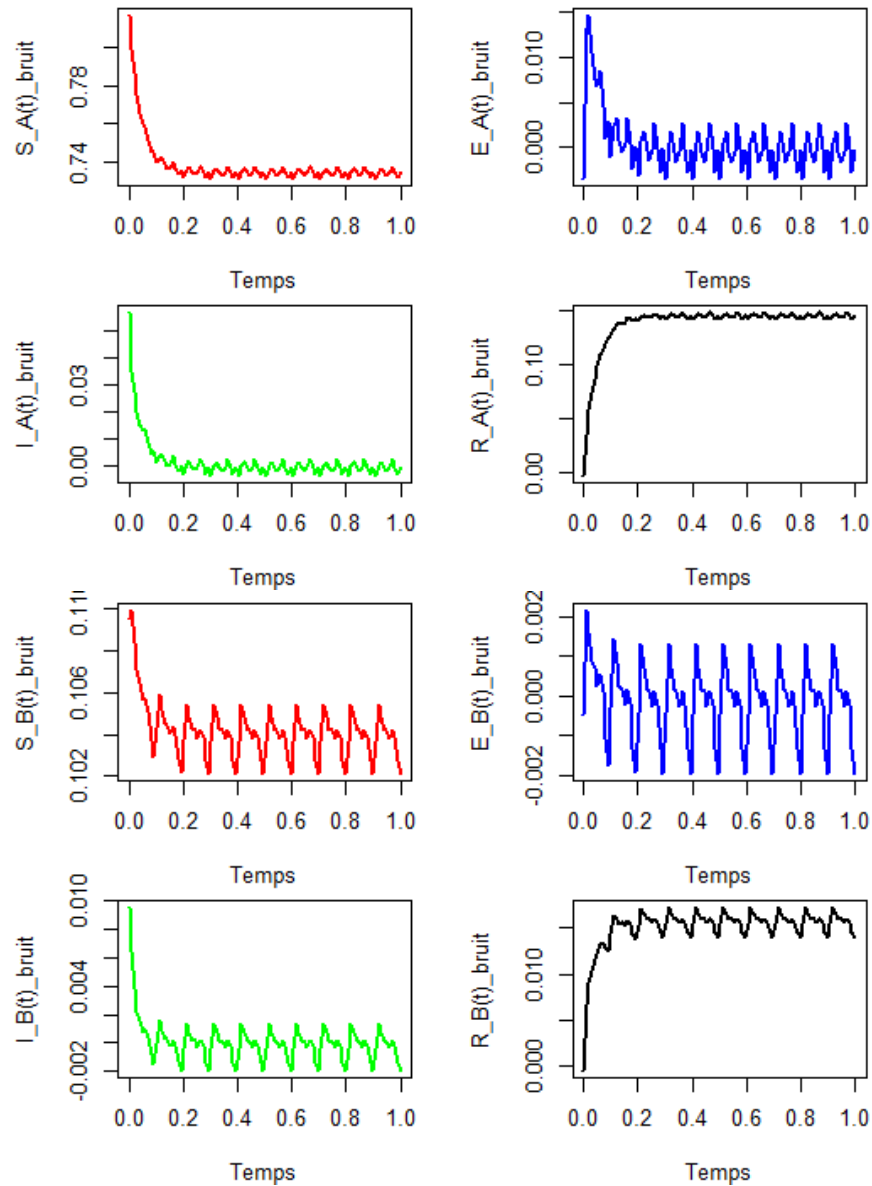


FIGURE 3.7 – Données simulées après bruitage

Après estimation avec les 6 paramètres qui varient et à l'aide de la méthode du gradient, nous obtenons bien les courbes souhaitées :

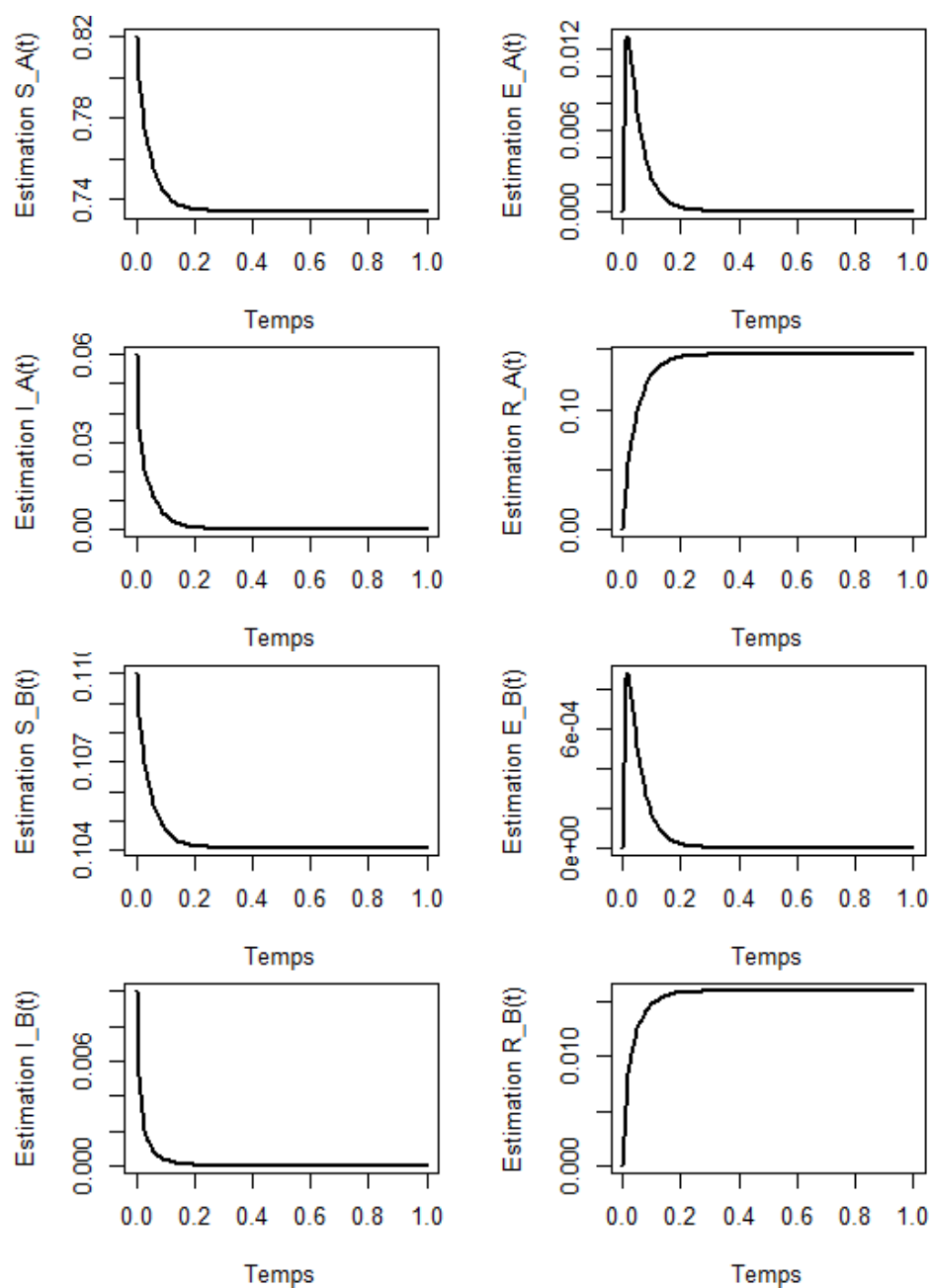


FIGURE 3.8 – Estimation sur les données simulées

Avec un jeu de paramètres estimées cohérent :

```
Parameters:
      Estimate Std. Error t value Pr(>|t|)
a_A      1.155e+00  1.031e+03   0.001   0.999
a_B      5.774e-01  5.154e+02   0.001   0.999
beta_A    4.603e+01  4.110e+04   0.001   0.999
beta_B    2.302e+01  2.055e+04   0.001   0.999
delta_SEIR 1.000e+02  3.470e+00  28.818 <2e-16 ***
gamma     7.400e+01  1.003e+00  73.787 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.001053 on 794 degrees of freedom
```

FIGURE 3.9 – Résultats de l'estimation sur les données simulées

Nous remarquons qu'il y a une différence par rapport aux paramètres de l'article dans  $a$  et  $\beta$ . Ceci peut s'expliquer par le fait que la modélisation utilise  $\lambda_j = f(a_j * \beta_j)$  et non pas  $a$  et  $\beta$  directement. Cela fait que ce qui compte est leur produit et non pas chacun indépendamment. Nous pouvons facilement vérifier que les produits dans l'article et dans notre modélisation sont les mêmes.

Pour challenger encore plus notre modèle, nous avons augmenté abusivement les paramètres. Avec un jeu de paramètres initiaux 2 fois plus grand, le modèle converge toujours :

```
Parameters:
      Estimate Std. Error t value Pr(>|t|)
a_A      2.655e+02  2.943e+05   0.001   0.999
a_B      1.328e+02  1.471e+05   0.001   0.999
beta_A    2.002e-01  2.219e+02   0.001   0.999
beta_B    1.001e-01  1.109e+02   0.001   0.999
delta_SEIR 1.000e+02  3.339e+00  29.952 <2e-16 ***
gamma     7.400e+01  9.666e-01  76.553 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.001015 on 794 degrees of freedom
```

FIGURE 3.10 – Résultats de l'estimation (2) sur les données simulées

Après avoir estimé nos paramètres, nous avons pu prédire le comportement de l'épidémie grâce à une interpolation des données que l'on avait déjà. Le script de cette prédiction est dans le code. Comme la grippe n'est pas le virus choisi pour le challenge, nous nous sommes contentés de prédire une seule valeur.

## 3.6 Discussion et perspectives

La grippe saisonnière est l'un des virus les plus communs pour les humains qui mute rapidement ; un même individu peut être infecté chaque année. Nous aurions donc pu prendre en compte ce retour à l'état S à partir de R. Une autre amélioration envisageable aurait été de répartir les personnes atteintes en 3 groupes, ce qui aurait été plus efficace, mais aurait causé plus de redondance dans le code (raison pour laquelle nous l'avons évité). Trois groupes (0-15 ans, 15-65 ans, et 65ans et plus) auraient été une bonne répartition. Pour affiner notre modèle, nous aurions aussi pu :

- Prendre le temps en considération dans nos équations différentielles (hypothèse d'un retard)
- Prendre des valeurs différentes de  $\beta$  et  $\gamma$  selon les différents groupes
- Prendre en compte un paramètre de mortalité, car même s'il est faible, la grippe reste un virus qui peut mener à la mort

# Chapitre 4

## La peste bubonique

### 4.1 Etude biologique, médicale et épidémiologique

#### 4.1.1 Présentation de la maladie

La peste est une zoonose bactérienne, due à *Yersinia pestis*, que l'on trouve habituellement chez les petits mammifères et les puces qui les parasitent. Elle se transmet d'un animal à l'autre par les puces.



FIGURE 4.1 – La bactérie *Yersinia pestis* responsable de la peste.

L'être humain peut être infecté :

- par les piqûres de puces vecteurs infectées
- par contact non protégé avec des liquides corporels infectieux
- par inhalation de gouttelettes respiratoires/de petites particules venant d'un patient atteint de peste pulmonaire.

Cette maladie est très grave chez l'homme, notamment pour les formes septicémique (infection systémique causée par la circulation d'une bactérie dans le sang) et pulmonaire, avec un taux de létalité de 30% à 100% en l'absence de traitement. Sauf si elle est traitée

rapidement, la peste pulmonaire est toujours mortelle ; elle est particulièrement contagieuse et elle peut déclencher de graves épidémies par transmission d'une personne à l'autre du fait des gouttelettes en suspension dans l'air. De 2010 à 2015, on a enregistré 3248 cas de peste dans le monde, dont 584 mortels. [WHO 2018d]

Au cours de l'histoire, la peste a été responsable de pandémies étendues avec une forte mortalité. Connue sous le nom de « peste noire » au XIV<sup>e</sup> siècle, elle a provoqué plus de 50 millions de morts en Europe et exterminé plus du tiers de la population européenne. En particulier, la France a perdu plus de 40 % de sa population pendant cette période.

En tant que zoonose, elle est présente sur tous les continents, sauf l'Océanie. Il y a un risque de peste humaine dès qu'il y a une coexistence d'un foyer naturel (bactérie, réservoir animal et vecteur) et d'une population humaine.

### 4.1.2 Signes et symptômes

Les sujets infectés présentent en général un état fébrile aigu et d'autres symptômes systémiques non spécifiques après une période d'incubation de 1 à 7 jours (fièvre brutale, frissons, céphalées, douleurs corporelles, état de faiblesse, vomissements et nausées).

Selon la voie d'infection, on observe deux principales formes de peste :

-**La peste bubonique** est la forme la plus courante, due à la piqure d'une puce infectée. Il en résulte une inflammation du ganglion avec tension douloureuse des tissus : c'est ce qu'on appelle le « bubon ». Au stade avancé, les ganglions enflammés finissent par s'ulcérer et suppurer. La transmission interhumaine de la peste bubonique est rare. La peste bubonique peut progresser aux poumons ; on parle alors de peste pulmonaire.

-**La peste pulmonaire** est la forme la plus virulente et aussi une forme rare de la maladie. La durée d'incubation peut n'être que de 24 heures. Le sujet infecté peut transmettre la maladie à autrui par des gouttelettes respiratoires. Si elle n'est pas diagnostiquée et traitée de manière précoce, cette forme peut être mortelle.

### 4.1.3 Comment gérer cette épidémie ?

Il y a plusieurs mesures à prendre pour bien gérer ce genre de flambées épidémiques, que l'Organisation Mondiale de la Santé résume dans ce qui suit :

- Bien localiser la source de la maladie. Désinfecter.
- Isoler les cas infectés, les surveiller. Assurer une bonne protection aux agents de santé.
- Assurer et offrir un traitement efficace. De nos jours, la peste peut être traitée facilement à l'aide d'antibiotiques et prévenue grâce à l'application de précautions standards.
- Etudier les échantillons prélevés avec soins



## 4.2 Description du modèle et simulations

### 4.2.1 Introduction au modèle

On va s'intéresser dans la suite au modèle de **Kermack et McKendrick** qui a cherché à modéliser la propagation de la peste bubonique à Bombay au début du XX<sup>ème</sup> siècle. En outre, les rapports publiés sur l'épidémie de la peste mettent en lumière un aspect de saisonnalité qui influe fortement sur la répartition des infectés et des morts entre 1897 et 1911. Pour toutes ces raisons, on a décidé d'étudier un modèle saisonnier pour la peste à Bombay. Notre objectif était d'adopter un modèle qui s'adapte parfaitement à la réalité.

Voici l'allure de courbe de répartition des décès dus à la peste bubonique à Bombay, entre 1897 et 1911.

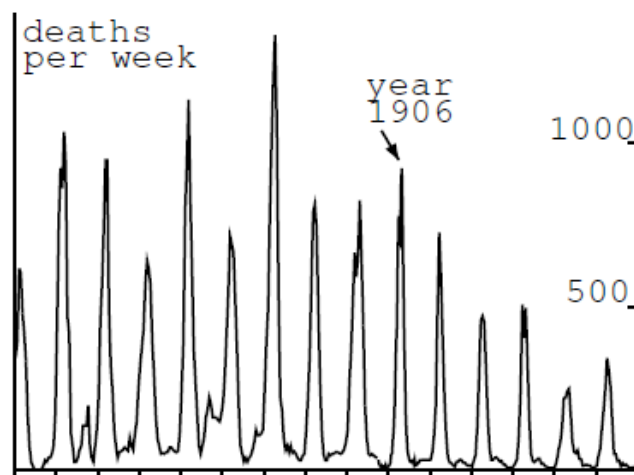


FIGURE 4.2 – Le nombre des décès hebdomadaires dus à la peste à Bombay

Une comparaison avec les données statistiques météorologiques a montré que l'épidémie ne pouvait pas se maintenir quand la température est au-dessous de  $26^{\circ}\text{C}$  ( $80^{\circ}\text{F}$ ). Les bacilles de la peste sont très sensibles à la température. Et les puces peuvent rester infectieuses bien plus longtemps lorsqu'il fait froid que lorsqu'il fait chaud. En outre, la saisonnalité de la présence des puces chez les rats et la fertilité des rats, dues à la variation de la température, sont à prendre en compte.

Avec ce qui précède, on peut donc affirmer que la variation des populations des rats et des humains est due à un facteur saisonnier touchant les bacilles et les puces.

### 4.2.2 Description du modèle et formalisme mathématique

Nous proposons, comme modèle, incluant cet aspect saisonnier, un modèle légèrement complexe, faisant appel à trois populations (puces, rats et humains), mais plus réaliste. On utilisera les notations standards (S,I,R).

- $S(t)$  le nombre des rats susceptibles
- $I(t)$  le nombre des rats infectés
- $R(t)$  le nombre des rats immunisés
- $P(t) = S(t) + I(t) + R(t)$  le nombre total des rats en vie
- $F(t)$  le nombre des puces infectées vivant librement, c'est à dire qui ne sont ni fixées sur un rat, ni sur un humain
- $H(t)$  le nombre d'humains infectés
- $D(t)$  l'incidence des décès dûs à la peste chez les humains

Notons qu'on ne tient pas en compte les humains et les puces susceptibles, car on considère qu'ils sont en léger excès. A titre d'exemple, annuellement on observe 10 000 décès dûs à la peste chez les humains dans une population d'un million. Un million est donc largement supérieur à 10 000.

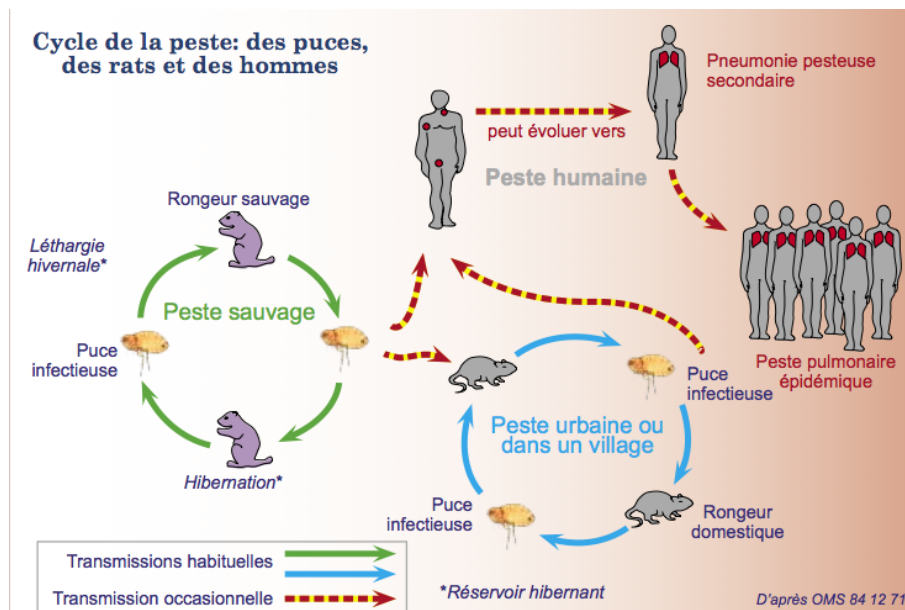


FIGURE 4.3 – Le processus de la transmission de la peste bubonique.

Voici dans cette figure 3.3, le processus de transmission de la peste bubonique. Le but de la modélisation est donc de chercher à représenter les différentes interactions entre l'homme et les agents responsables de l'infectin. L'ensemble des équations qui permettent de modéliser cet épidémie que nous avons cherché à étudier et qui sont largement considérés

[Bacaër 2012], sont :

$$\frac{dS}{dt} = b(P) - mS - c(1 - \omega)\pi(\theta(t))\frac{S}{P}F + \epsilon m' I \quad (4.1)$$

$$\frac{dI}{dt} = c(1 - \omega)\pi(\theta(t))\frac{S}{P}F + m' I \quad (4.2)$$

$$\frac{dR}{dt} = \epsilon' m' I - mR \quad (4.3)$$

$$\frac{dF}{dt} = f\pi'(1 - \epsilon - \epsilon')m' I - cF \quad (4.4)$$

$$\frac{dH}{dt} = c\omega\pi(\theta(t))F - aH \quad (4.5)$$

$$D(t) = \sigma aH(t) \quad (4.6)$$

Le tableau ci-dessous nous donne la signification et la valeur numérique pour chacun des paramètres utilisés pour décrire notre modèle :

Paramètre	Signification	Valeur numérique
$b(P)$	fertilité des rats $b(P) = rP/(1 + P/K)$	$r=0.4/\text{mois}$ $K=50\ 000$
$1/m$	espérance de vie des rats	$m=0.03/\text{mois}$
$1/c$	temps pour que les puces trouvent un hôte	$c=30/\text{mois}$
$\omega$	fraction des puces qui trouvent un hôte humain	$\omega = 0.02$
$\pi(\theta)$	probabilité de transmission du puce au rat ou à l'homme	$\pi_0(\frac{3}{4} - \frac{1}{4}\tanh(\theta - 80))$ avec $\pi_0 = 0.9$
$\theta(t)$	température en F avec $50\text{ F} = 10\text{ C}$	(voir courbe)
$1/m'$	durée de la peste chez les rats	$m' = 3/\text{mois}$
$\epsilon$	fraction des rats guérissants sans immunité	$\epsilon=1$
$\epsilon'$	fraction des rats guérissants avec immunité	$\epsilon' = 0.1$
$f$	nombre moyen des puces par rat	$f=4$
$\pi'$	probabilité de transmission du rat à la puce	$\pi' = \pi_0$
$1/a$	la durée de la peste chez les humains	$a=4/\text{mois}$
$\sigma$	mortalité	$\sigma = 0.9$

TABLE 4.1 – Tableau des paramètres du modèle de la peste bubonique

Voici quelques commentaires sur le modèle adopté et sur les paramètres :

- Ce modèle convient pour la peste bubonique uniquement, car la peste pulmonaire peut se transmettre entre les humains.
- La peste bubonique ne peut se transmettre entre les humains en l'absence des puces.
- La fraction  $\omega$  des puces libres qui trouvent un hôte humain dépend des conditions sanitaires. Ce paramètre est à ajuster, en remarquant que le nombre moyen de décès est quasi-proportionnel à  $\omega$ .

- La probabilité de transmission  $\pi(\theta) = \pi_0(0.75 - 0.25\tanh(\theta - 80))$  et  $\pi_0 = 0.9$  modélise la dépendance de la transmission à la température du développement des bacilles dans l'estomac des puces.
- La température  $\theta(t)$  constatée à Bombay, est proche d'une fonction périodique.

**On fait l'hypothèse simplificatrice que la température (fig. 4.11) est une vraie fonction périodique du temps, avec une période égale à 12 mois. Il faut noter que ce choix périodique va influencer les tailles des pics des infectés et des morts.**

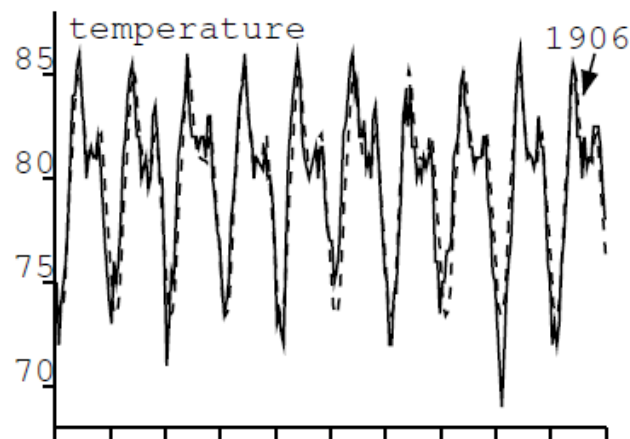


FIGURE 4.4 – Le profil de  $\theta(t)$  à Bombay entre 01/1897 et 12/1911.

- Le nombre de puces trouvées sur un rat, dépend de la température, on a pris  $f$  constant égale à la moyenne  $f=4$ .
- Le rat étant un animal à sang chaud, on a considéré que la probabilité de transmission de la peste du rat à la puce ne dépend pas de la température et on fixe  $\pi' = \pi_0$ .

### 4.2.3 Simulations

La courbe des humains infectés a la même allure que la courbe des morts car  $D(t) = \sigma a H(t)$ , donc on a essayé plusieurs profils de température qui sont périodiques, pour avoir des simulations cohérentes avec la réalité (**Figure 2.2**)

On présentera dans la suite les résultats de simulations obtenus en résolvant le système des équations différentielles, en utilisant la fonction "ode" implémentée en R. On travaille sur un intervalle de temps de 120 mois ( 10 ans), car la propagation de la peste bubonique à Bombay qu'on étudie, a eu lieu entre 1897 et 1911.

Les courbes d'évolution des populations dans le cas d'une fonction de température tel que  $\theta(t) = 80 + 10\cos(2\pi t/12)$  en F, avec un vecteur d'états initial ( $S(0) = 499, I(0) = 1, R(0) = 0, F(0) = 100, H(0) = 0$ ) et Tmax=120 mois sont données par :

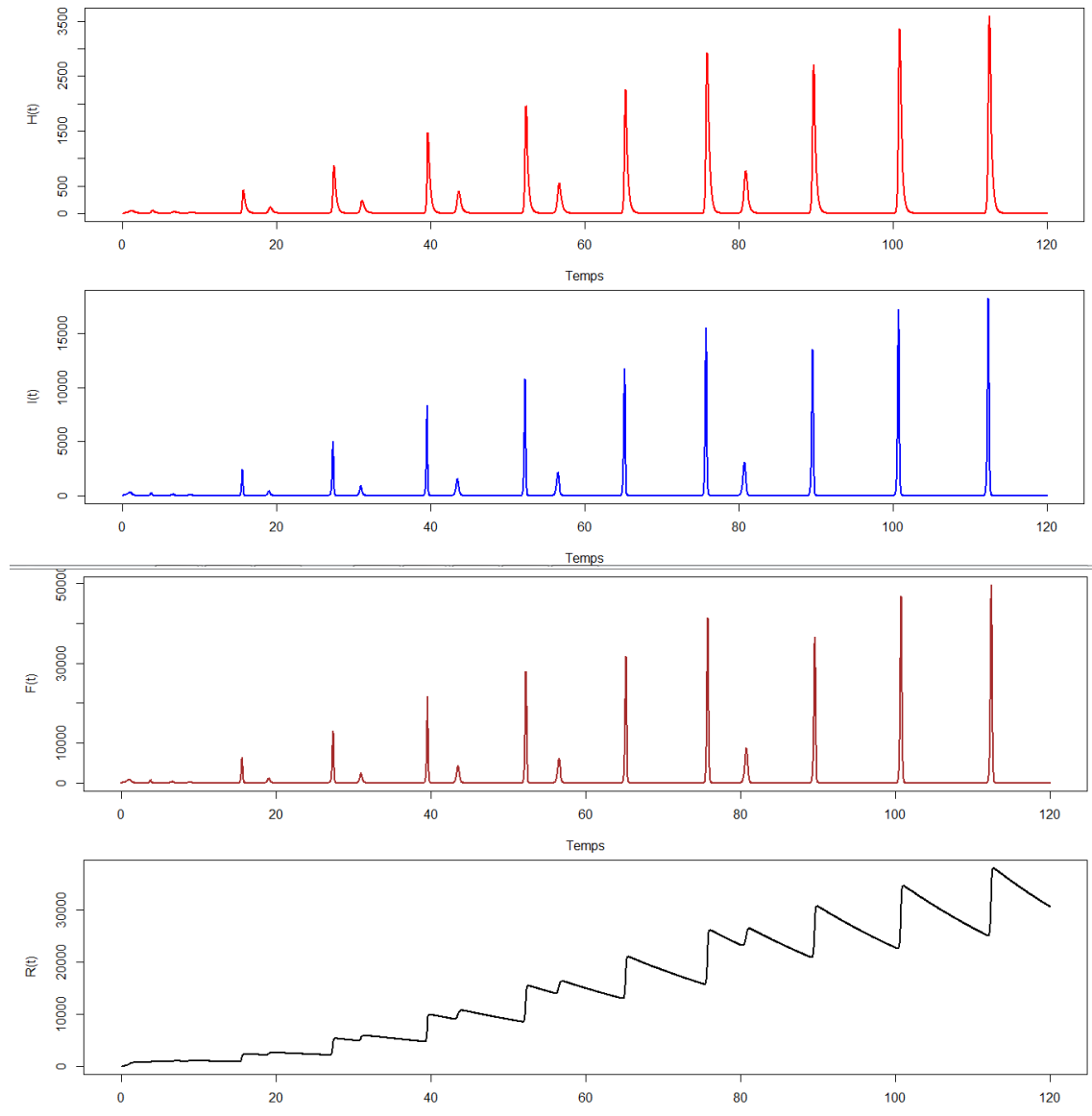
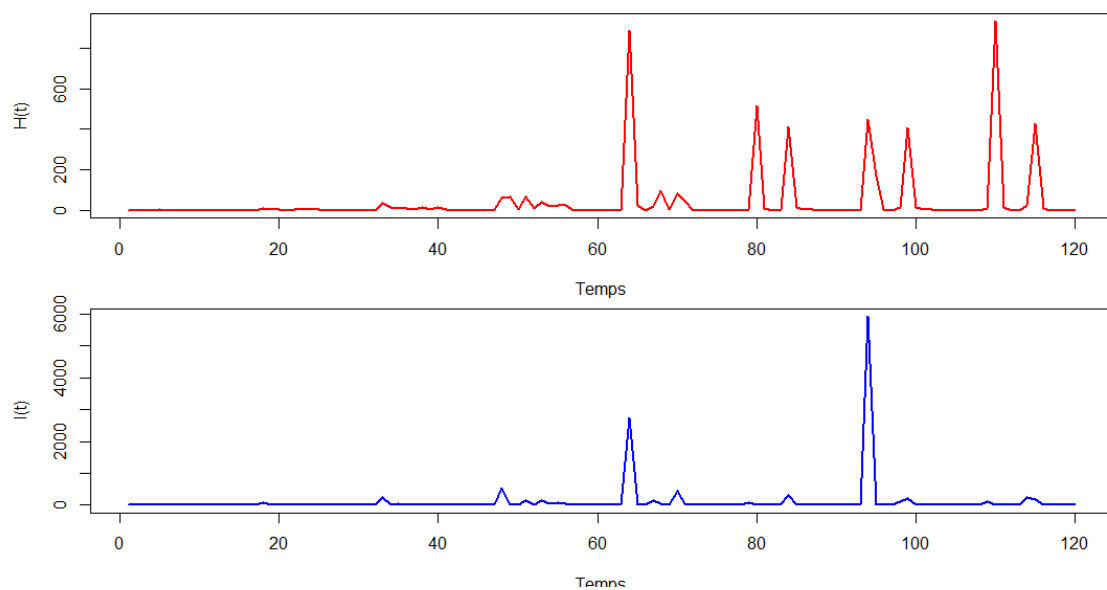
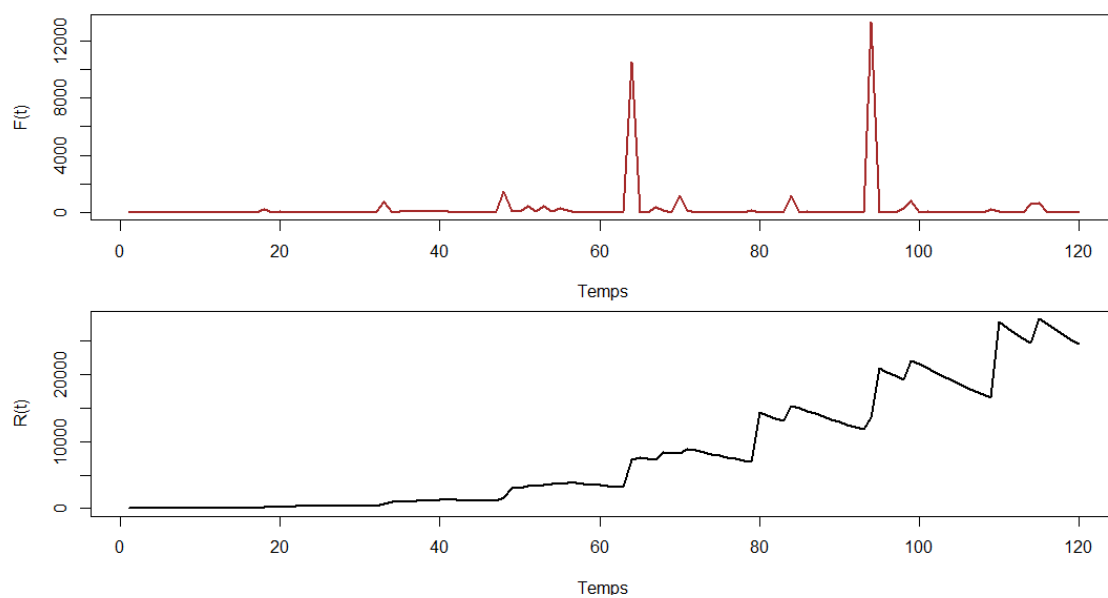


FIGURE 4.5 – L'évolution du nombre des humains infectés (en rouge) et des rats infectés (en bleu), des rats immunisés (en noir) et des puces libres (en marron)

Les courbes d'évolution des population dans le cas d'une fonction de température  $\theta(t) = 80 + 4\cos(2\pi t/12) + (4/9)\cos(6\pi t/12) + (4/25)\cos(10\pi t/12)$  en F. Dans ce cas, on prend le fondamental et les 2 premiers harmoniques d'un signal triangulaire de période 12 mois, avec un vecteur d'états initial ( $S(0) = 499, I(0) = 1, R(0) = 0, F(0) = 100, H(0) = 0$ ) et Tmax=120 mois. Les simulations nous donnent :

FIGURE 4.6 – L'évolution de  $H(t)$  (en rouge) et de  $I(t)$  (en bleu)FIGURE 4.7 – L'évolution de  $R(t)$  (en noir) et  $F(t)$  (en marron)

Le choix de la fonction de température influence fortement l'évolution des différentes populations, on a essayé plusieurs fonctions de températures pour avoir une simulation proche de la réalité, mais quelques unes d'entre elles compliquent l'estimation des paramètres du problème dans la suite.

D'après les deux courbes des humains infectés obtenues, avec les deux profils de température, on constate que la taille des premiers pics épidémiques est douteuse, et on s'est posé la question pourquoi la taille des pics va varier. Mais en cherchant dans la bibliographie, on a pu interpréter cette variation des tailles de pics. Les rapports des morts et des infetés humains à Bombay, montrent qu'il y avait une baisse remarquable de la population à Bombay( de 850 000 à 437 000 entre décembre 1896 et février 1897) [Bacaër 2012]. Les gens avaient fui Bombay pour échapper à la peste. Cette migration a certainement diminué la taille des premiers pics épidémiques, la population retournera à son niveau normal quelques mois après les premiers pics.

Les figures de l'évolution des rats infectés montrent qu'il y a oscillation des populations des rats, comparée à la situation sans maladie, la peste a diminué considérablement la population totale des rats. La population des rats susceptibles est aussi très réduite pendant la saison de la peste de février à avril mais commence à croître à partir d'avril lorsque des températures plus élevées réduisent la transmission. [Bacaër 2012]

## 4.3 Analyse du modèle

Le modèle qu'on a introduit dans la partie précédente, fait intervenir 14 paramètres, qui vont influencer l'évolution des différentes populations, de manières différentes, donc nous sommes obligés de choisir les bons paramètres à estimer, afin de pouvoir prédire le mieux possible l'évolution des populations, à partir des données présentes. Pour toutes ces raisons, on a décidé de réaliser une étude de sensibilité globale du modèle.

### 4.3.1 Etude de la sensibilité

Nous nous sommes intéressés, dans cette partie à l'étude de sensibilité globale du modèle introduit. Et grâce à la Librairie "Sensitivity" implémentée en R, on a pu obtenir une carte montrant les paramètres auxquels notre modèle est sensible. Voici ci-dessous la carte de sensibilité obtenue en utilisant la méthode de Morris :

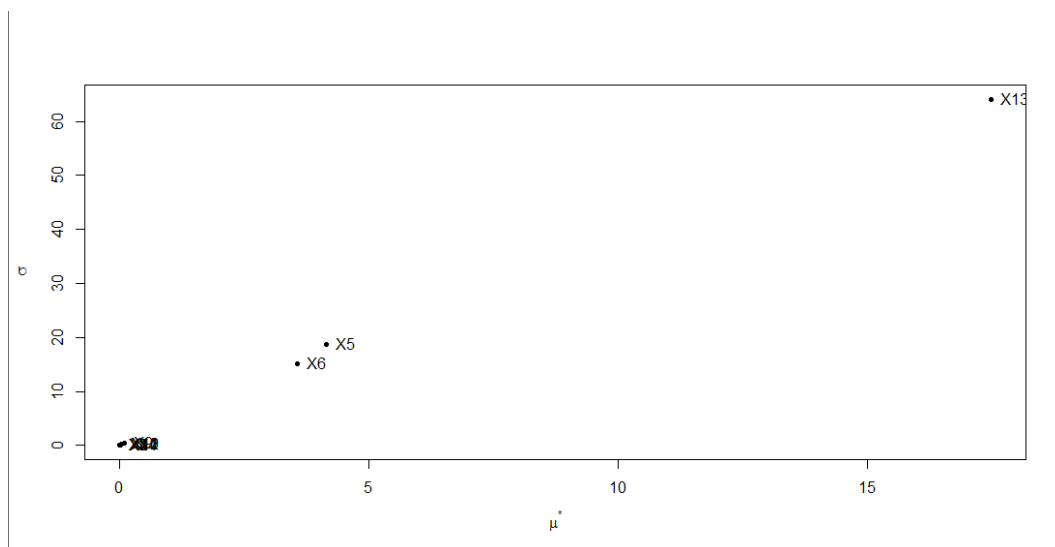


FIGURE 4.8 – L'étude de sensibilité du modèle selon la méthode de Morris

D'après les résultats de sensibilité obtenus par la méthode de Morris, on constate que notre modèle est fortement influencé par la fréquence de la peste chez les hommes ( $X13 = a$ ). Dans le modèle de base, on a pris  $a$  égale à 4/mois. Cette forte influence est tout à fait naturelle, vu que le nombre des humains infectés dépend nécessairement de la durée de la peste chez eux. En outre, on constate que le modèle est lié aussi à la proportion des puces libres qui trouvent un hôte humain ( $X5 = \omega$ ). Cette grandeur était fixée, dans le modèle de base, à  $\omega = 2\%$ . Ce choix était arbitraire, afin d'ajuster les courbes obtenues par la simulation aux courbes statistiques réelles. Cette sensibilité à  $\omega$  peut être expliquée par le fait que lorsque la proportion des puces libres qui trouvent un hôte humain augmente, même le nombre des humains infectés va augmenter. Cette affirmation apparaît plus clairement est clair dans l'équation suivante :

$$\frac{dH}{dt} = c\omega\pi(\theta(t))F - aH$$

En plus, notre modèle est aussi sensible au choix de ( $X6 = \pi_0$ ). Cette grandeur représente l'amplitude de la probabilité de transmission d'une puce à un rat ou à l'homme (qui dépend aussi du profil de température adopté). Elle a été ajustée par les concepteurs du modèle pour que les résultats soient proches de la réalité, et elle a été fixée  $\pi_0 = 0.9$ . Cette sensibilité peut être interprétée par le fait que l'augmentation de cette probabilité, va engendrer l'augmentation de la probabilité de transmission de la bactérie responsable de la peste. Enfin, on constate que le modèle est moins sensible aux autres paramètres. Certes, leur choix est indispensable pour avoir les résultats souhaités, mais le modèle n'est pas très sensible à leurs variations.



## 4.4 Estimation

Afin de pouvoir prédire l'évolution des différentes populations, et en particulier le nombre des infectés chez les population humaine à partir des données statistiques disponibles jusqu'au jour d'aujourd'hui, on doit pouvoir estimer les paramètres du modèle introduit précédemment. Pour cela, on a décidé d'estimer nos paramètres en se basant sur des données simulées, comme celles présentées dans la partie des simulations, en rajoutant un bruit normal centré et réduit aux données. Mais vu que le nombre des infectés humains oscillent, et est proche de 0 pour plusieurs périodes, on a décidé de rajouter un bruit positif, afin d'avoir des données simulées bruitées réalistes, en utilisant la syntaxe suivante :

$$\text{Bruit} < -\text{sapply}(\text{rnorm}(Tmax, mean = 0, sd = 1), \text{function}(x)\text{max}(x, 0))$$

### 4.4.1 Méthodes utilisées en estimation

#### Estimation avec nls2

Avant d'effectuer l'étude de sensibilité du modèle, on a essayé d'estimer chaque paramètre tout seul, et les résultats n'étaient pas toujours satisfaisants. En utilisant la fonction "nls2" implémentée en R, on a pu estimer 2 paramètres, chacun tout seul, avec des précisions qui varient d'un paramètre à l'autre.

Après l'effctuation de l'étude de sensibilité, nous avons décidé d'estimer les paramètres  $a$ ,  $\pi_0$  et  $\omega$ .

On a obtenu, grâce à cette méthode d'etimation,  $a = 3.24$  alors que la valeur réelle était  $a = 4$ , avec un écart type égale à 0.9, et un point de départ  $a_0 = 2$ . Voici la courbe obtenue pour cette estimation :

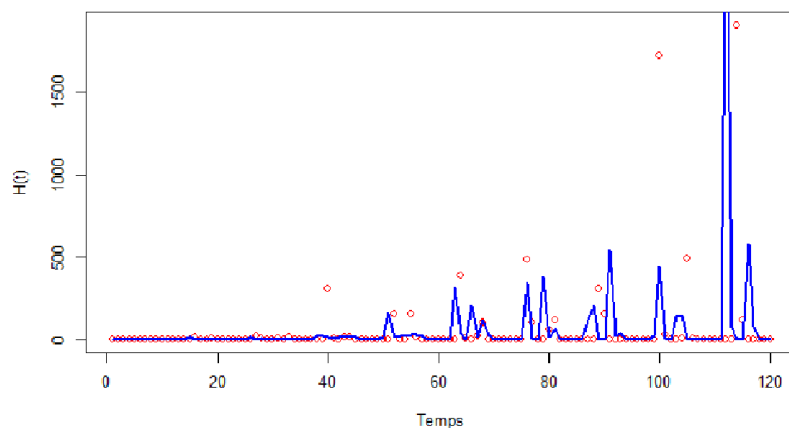


FIGURE 4.9 – Estimation du paramètre  $a$  dans le modèle de la peste

De même, on a essayé d'estimer le paramètre  $\pi_0$  avec la même méthode. Cette fois-ci, on a trouvé un  $\pi_0$  estimé égale à 1, sachant que la valeur réelle  $\pi_0 = 0.9$ , avec un écart type égale à 5.2, et un point de déprat  $\pi_0 = 0.6$ , cet écart type est dû au fait qu'on a limité l'intervalle de l'optimisation à  $[0,1]$  vu que  $\pi_0$  est une probabilité de transmission. Voici la courbe obtenue pour cette estimation :

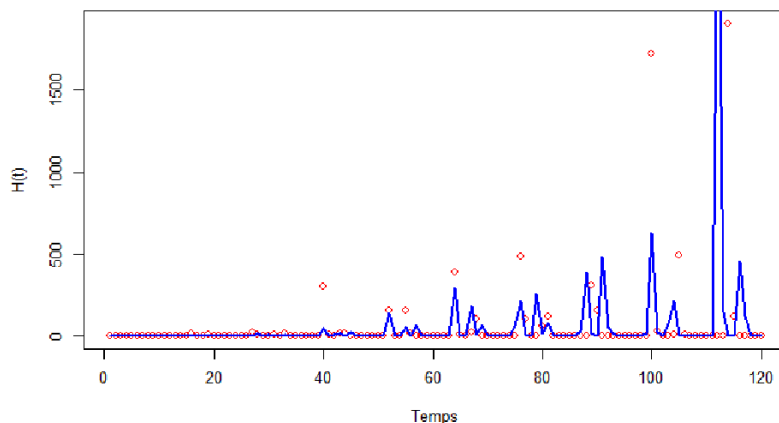


FIGURE 4.10 – Estimation du paramètre  $\pi_0$  dans le modèle de la peste

Cependant, nous n'avons pas pu estimer correctement plus qu'un paramètre à la fois si on choisit des points de départ loins de la vraie valeur, et si on ne délimite pas l'intervalle sur lequel la fonction "nls2" cherche les paramètres estimés. Tous ces problèmes sont dûs aux problèmes de convergence de l'algorithme d'optimisation "port" et la forte sensibilité des paramètres.

Pour ces raisons, on a essayé un autre algorithme d'optimisation qui est l'algorithme génétique implémenté en R, sous la librairie GA.

### Estimation avec l'algorithme génétique

L'utilisation de cet algorithme nécessite l'introduction de la fonction à maximiser. On a donc choisi de maximiser l'opposé de la fonction de la moyenne des carrés des résidus, ce qui revient donc à minimiser cete fonction. L'algorithme utilisé pour implémenter cette méthode d'estimation des trois paramètres ( $a$ ,  $\omega$  et  $\pi_0$ ) basée sur l'algorithme génétique peut être consulté sur le script joint du document.

Même en utilisant cette méthode, si on choisit un domaine de recherche quelconque, on n'arrive pas à estimer le triplet ( $a$ ,  $\omega$  et  $\pi_0$ ), et on reçoit des messages d'erreurs. Pour cela, on choisit un domaine de recherche restreint. En outre, on a constaté que si le nombre de points n'est pas suffisant, l'optimisation par cet algorithme coince. On a donc augmenté le nombre de points en utilisant `seq(1, Tmax, by = 0.1)` au lieu de `1 : Tmax`. Et on obtient par cette méthode les résultats d'estimation suivants :

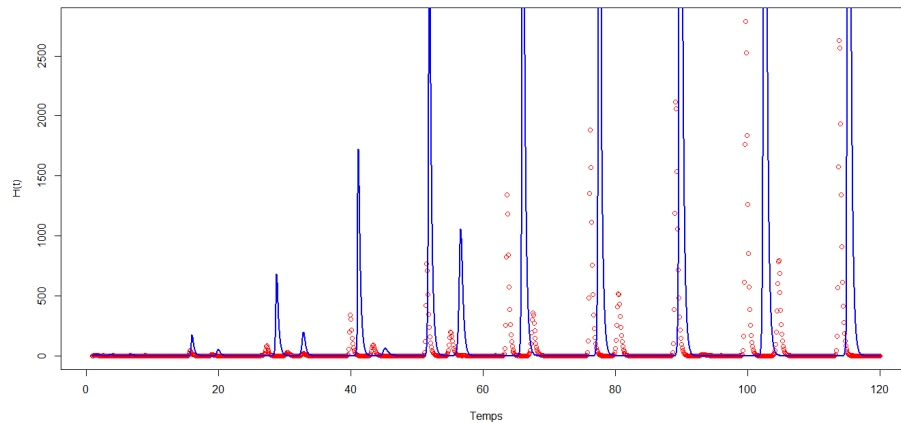


FIGURE 4.11 – Estimation des paramètres avec algorithme génétique

Ces résultats correspondent à des paramètres estimés ( $\mathbf{a}=3.74$ ,  $\omega=0.05$ ,  $\pi_0=0.81$ ) alors que les vraies valeurs des données non bruitées sont ( $\mathbf{a}=4$ ,  $\omega=0.9$  et  $\pi_0=0.9$ ).

#### 4.4.2 Difficultés et problèmes liés à l'estimation

Notons que ces difficultés sont en majeure partie du fait que nous avons essayé de prédire l'évolution de la peste non pas sur une ou deux années, mais sur 10 ans!

**La sensibilité :** Comme déjà évoqué dans la partie précédente, l'estimation de tous les paramètres avec la fonction "nls2" n'aboutissait pas. Car, on essayait d'estimer des vecteurs de paramètres de tailles différentes, et de sensibilités différentes.

**Le choix du point de départ pour l'optimisation :** Pour cette même fonction, l'estimation n'aboutissait pas pour tous les points de départ, de ce fait on était obligé de donner un point de départ proche des valeurs souhaitées, ce qui est faisable dans notre cas, et qui sera difficile si on avait des données réelles.

**Le choix du domaine de recherche :** Egalement, le choix d'un domaine de recherche aléatoire pour l'algorithme génétique ne permet pas de trouver les bons paramètres souhaités. On était donc obligé de fixer un domaine restreint de recherche des paramètres.

**Le choix du nombre de points :** Afin de résoudre le systèmes des équations différentielles et de pouvoir estimer avec précision les paramètres, on devait choisir le nombre de points suffisamment grand, tout en restant dans la limite raisonnable du temps d'exécution.

**La largeur des pics de la courbe des infectés :** Les pics de la courbe d'évolution des infectés humains ont une petite largeur. De ce fait, sur la quasi-totalité du temps, le nombre des infectés est nul. Ainsi, on n'avait pas assez d'information sur les infectés pour pouvoir bien estimer les paramètres souhaités.

## 4.5 Discussion et perspectives

### 4.5.1 Proposition d'un nouveau système d'équations

L'étude de sensibilité a montré que le nombre des infectés humains introduit n'est pas sensible à la variation de plusieurs paramètres, en particulier  $1/c$ . L'explication est simple c'est le temps nécessaire pour que les puces trouvent un hôte, ce temps moyen est égale à  $1/30$  mois. Donc, c'est l'échelle de temps la plus courte du modèle. De ce fait, on peut s'attendre à ce que  $\frac{dF}{dt} = 0$  (Régime quasi-stationnaire), et donc on aura  $f\pi'(1 - \epsilon - \epsilon')m'I = cF$ , et en remplaçant  $cF$  par son expression, on obtiendra le nouveau système d'équations ne faisant pas intervenir  $c$  :

$$\frac{dS}{dt} = b(P) - mS - (1 - \omega)\pi(\theta(t))\frac{S}{P}f\pi'(1 - \epsilon - \epsilon')m'I + \epsilon m'I \quad (4.7)$$

$$\frac{dI}{dt} = (1 - \omega)\pi(\theta(t))\frac{S}{P}f\pi'(1 - \epsilon - \epsilon')m'I + m'I \quad (4.8)$$

$$\frac{dR}{dt} = \epsilon' m'I - mR \quad (4.9)$$

$$\frac{dH}{dt} = \omega\pi(\theta(t))f\pi'(1 - \epsilon - \epsilon')m'I - aH \quad (4.10)$$

$$D(t) = \sigma aH(t) \quad (4.11)$$

### 4.5.2 Le choix de la fonction de température

Comme on a déjà vu, le choix de la fonction de température joue un rôle primordial dans l'allure de l'évolution des différentes populations. Les différents résultats présentés dans ce chapitre résultent du choix d'une fonction de température  $\theta(t) = 80 + 10\cos(2\pi t/12)enF$ . Donc pour avoir des résultats qui collent parfaitement avec la réalité, il faut proposer une fonction de température qui décrit parfaitement la figure (4.4).

### 4.5.3 Conclusion sur le modèle

Le modèle initial de **Kermack et McKendrick** ne tenait pas compte de la saisonnalité. Le choix d'inclure l'aspect périodique dans le modèle peut se révéler pertinent pour certaines études contemporaines de modélisation des épidémies. Mais en même temps la saisonnalité de cette épidémie de peste en Inde et son interprétation en terme de seuil de température ne semble pas cohérente avec d'autres données statistiques concernant des flambées de pestes dans d'autres pays. Le modèle mathématique dépend donc du pays considéré. Dans ce rapport, nous avons essayé de modéliser la peste de Bombay du début du XX<sup>ème</sup> siècle, puisqu'elle s'agit de l'une des figures les plus reproduites dans les livres sur la modélisation mathématique en épidémiologie.

## Conclusion et perspectives

A l'issue de ce rapport, nous avons donc l'honneur de pouvoir présenter ces résultats au comité directeur de l'OMS, et en particulier au président de l'organisation Docteur Tedros Adhanom Ghebreyesus, particulièrement concerné par la cause épidémiologique. Ce rapport résume les connaissances biologiques et épidémiologiques de certaines épidémies qui continuent à frapper le monde d'aujourd'hui, à savoir : **la dengue, le choléra, la grippe saisonnière, et la peste bubonique**, ainsi qu'une étude globale des modèles mathématiques et des simulations numériques décrivant les modes de propagation de ces épidémies au sein des populations.

Nous nous sommes intéressés à ces études mathématiques afin de pouvoir prédire l'évolution de ces maladies, à partir des données statistiques disponibles, et de permettre à l'OMS et aux pays concernés de prendre les mesures nécessaires pour s'opposer aux flambées épidémiques. Ces études mathématiques ne sont pas parfaites. Pour ce fait, nous avons à l'issue de nos modèles discuté leurs limites et les pistes d'améliorations possibles. De surcroît, ces études ne sont pas uniquement valables aux épidémies déjà étudiées, mais peuvent être généralisées à d'autres épidémies présentant des similitudes.

Cette mission a permis à chaque membre de l'équipe de grandir, de renforcer ces connaissances biologiques et épidémiologiques, ainsi que d'améliorer ces compétences quantitatives en simulation, en estimation, et en prédiction. Ces compétences sont devenues aujourd'hui indispensables dans le monde de la biologie en général. La cause sur laquelle nous avons travaillé et nous continuerons à travailler est de pouvoir prédire ces épidémies avec une grande précision sur une longue échelle de temps.



# Bibliographie

- [Bacaër 2012] N. Bacaër. *Le modèle de Kermack et McKendrick pour la peste à Bombay et la reproductivité nette d'un type avec saisonnalité*. Journal of Mathematical Biology, Springer Verlag (Germany), pages 403–422, 2012.
- [Bourdillon 2017] F. Bourdillon. *Grippe, bulletin hebdomadaire*. Santé publique France, vol. Semaine 9, 08 Mars 2017.
- [Den 2002] *Report of the Scientific Working Group on Insect Vectors and Human Health*. TDR/SWG/VEC/ 03.1, Geneva, 2002.
- [Derouich 2003] M. Derouich, A. Boutayeb et E. H. Twizell. *A model of dengue fever*. Biomed. Eng. Online 2, page 523–544, 2003.
- [Esteva 1998] L. Esteva et C. Vargas. *Analysis of a dengue disease transmission model*. Math. Biosci. 150 (2), page 131–151, 1998.
- [Esteva 1999] L. Esteva et C. Vargas. *Analysis of a dengue disease transmission model with variable human population*. Math. Biosci. 38, page 220–240, 1999.
- [Esteva 2000] L. Esteva et C. Vargas. *Influence of vertical and mechanical transmission on the dynamics of dengue disease*. Math. Biosci. 167 (1), page 51–64, 2000.
- [Esteva 2003] L. Esteva et C. Vargas. *Coexistence of different serotypes of dengue virus*. J. Math. Biol. 46, page 31–47, 2003.
- [Feng 1997] Z. Feng et V. Vealsco-Hernandez. *Competitive exclusion in a vector-host model for the dengue fever*. J. Math. Biol. 35, page 523–544, 1997.
- [Guegan 2006] J.F. Guegan. *Quantification des risques d'émergence d'épidémies à choléra dans le bassin méditerranéen en relation avec le changement climatique*. 2006. CHOLCLIM.
- [Iooss. 2011] B. Iooss. *Revue sur l'analyse de sensibilité globale de modèles numériques*. Journal de la Société Française de Statistique, 2011.
- [Laguzet 2015] L. Laguzet. *Modélisation mathématique et numérique des comportements sociaux en milieu incertains. Application à l'épidémiologie*. Mathématiques [math]. Université Paris-Dauphine, 2015.

- [Mraidi 2014] R. Mraidi. *Modélisation et contrôle de la transmission du virus de la maladie de Newcastle dans les élevages aviaires familiaux de Madagascar*. École Doctorale Sciences Technologies Santé Université de La Réunion, 2014.
- [Newton 1992] E. A. Newton et A. Reiter. *A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics*. Am. J. Trop. Med. Hyg. 47, pages 709–720, 1992.
- [Pasteur 2018] Institut Pasteur. *Grippe*, 2018. <https://www.pasteur.fr/fr/centre-medical/fiches-maladies/grippe>.
- [Sobol 2001] I.M. Sobol. *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*. MATH COMPUT SIMULAT, pages 271–280, 2001.
- [Stolerman 2015] L. M. Stolerman, D. Coombs et S. Boatto. *Sir-network model and its application to dengue fever*. SIAM Journal on Applied Mathematics, vol. 75, page 2581–2609, 2015.
- [WHO 2018a] WHO. *Cholera*, 2018. <http://www.who.int/cholera/en/>.
- [WHO 2018b] WHO. *Dengue fever*, 2018. <http://www.who.int/topics/dengue/>.
- [WHO 2018c] WHO. *Grippe saisonnière*, 2018. <http://www.who.int/mediacentre/factsheets/fs211>.
- [WHO 2018d] WHO. *Peste*, 2018. <http://www.who.int/mediacentre/factsheets/fs267/fr/>.