

École polytechnique de Louvain

Data science for large networks : theory, algorithmics and applications

Author: **Hugues Herve KAMTO**

Supervisor: **Jean-Charles DELVENNE**

Readers: **Pierre SCHAUS, Niko SPEYBROECK**

Academic year 2020–2021

Master [120] en science des données, orientation technologie de l'information

Table des matières

0	Introduction générale	1
1	Modèles compartimentaux	5
1.1	Modèles à deux compartiments : $SI(S)$	6
1.2	Modèles à trois compartiments : $SIR(S)$	10
1.3	Modèles à quatre compartiments : $SEIR(S)$	14
2	Modèle de Kermack-McKendrick	17
2.1	Proportionnalité du modèle SIR	19
2.2	Taux de reproduction basique : R_0	19
2.3	Taux de croissance exponentielle du modèle SIR	21
2.4	Pic de l'épidémie	22
2.5	Taille finale d'une épidémie	24
2.6	Illustration : la grande peste à Eyam	26
2.7	Vaccination et immunisation collective	28
3	Processus de branchement (Branching process)	31
3.1	Processus de branchement de Galton-Watson	31
3.2	Probabilité d'extinction	33
3.3	Relation avec le modèle SIR de Kermack-McKendrick	34
4	Modèle de Reed-Frost	36
4.1	Graphe aléatoire	36
4.1.1	Graphe aléatoire Erdos-Rényi	36
4.1.2	Modèle d'attachement préférentiel	37
4.1.3	Small-world	38
4.2	Composante connexe géante	39
4.3	Modèles épidémiques sur les réseaux : Reed-Frost	41
4.3.1	Description du modèle	42
4.3.2	Probabilité de transmission	44
4.3.3	Taux de reproduction de base R_0	44
4.3.4	Taille finale de l'épidémie	44

Table des matières

4.4	Modèle de prévention : principe de percolation	46
4.4.1	Principe de percolation	46
4.4.2	Effet de percolation sur taux de reproduction R_0	47
5	Comparaison Kermack-McKendrick vs Reed-Frost	49
6	Taux de reproduction effectif : R_{eff}	55
6.1	Modèle par compartiment	58
6.2	Modèle par graphe	58
6.3	Simulation	59
6.4	Discussion	65
7	Bulles	70
7.1	Contexte	70
7.2	Modélisation	71
7.3	Mise en oeuvre sur données à grande échelle	75
8	Conclusion générale	79
	Bibliographie	81
	Annexes	84
A	Kermack-McKendrick vs Reed-Frost : cas de plusieurs sources	85
B	Taux de reproduction effectif : cas de plusieurs sources	87
C	Effet des bulles : cas avec plusieurs sources	88
D	Résultat d'estimation de R_{eff} : modèle Corri-Ferguson	90

Dédicaces

Je dédie ce mémoire à mon frère feu Tatap Meudjo Honoré brutalement arraché à la vie le 04 aout 2021.

Remerciements

Un mémoire se prépare avant tout dans une faculté et c'est à cette faculté qui est l'école polytechnique de l'université de Louvain que vont mes premiers remerciements. Je suis particulièrement reconnaissant pour la qualité de la formation et des moyens mis à disposition pour la réalisation de ce mémoire.

Ma gratitude va à mon promoteur, le professeur Jean-Charles Delvenne, pour son suivi tout au long de mes travaux. La confiance qu'il m'a faite, sa haute disponibilité et ses réponses toujours promptes à mes questions m'ont donné la motivation et le plaisir à découvrir la recherche. En plus de vos qualités d'enseignants, je vous remercie pour vos qualités humaines.

Je remercie tout particulièrement les professeurs Pierre Schaus et Niko Speybroeck qui me font l'honneur d'être les lecteurs de ce modeste travail.

Enfin, je tiens à remercier ma famille d'accueil en Belgique et mon ami Justin Davaux pour toutes leurs hospitalités, leurs encouragements et leurs soutiens durant mes études.

Abstract

Ce mémoire vise à comparer les modèles compartimentaux avec ceux par graphes pour la simulation de l'épidémie spécialement en ce qui concerne le calcul du taux de reproduction effectif R_{eff} qui n'est pas toujours très détaillé dans la littérature, ceci est d'une grande importance, car il permet de déterminer et d'évaluer la situation de contrôlabilité d'une épidémie en cours de propagation, fournissant ainsi des informations aux décideurs leur permettant d'adopter ou adapter leurs politiques restrictives ou de relâchement comme nous l'avons vécu avec les prises de décisions gouvernementales pour le cas de la covid-19 et dans la façon donc les médias expliquent l'épidémie aux gens. Or le calcul de ce facteur à un instant t de l'évolution d'une épidémie n'est pas sans ambiguïté, et notamment il peut être calculé différemment selon le modèle avec lequel l'épidémie est capturée.

Alors d'abord on va comparer les modèles à compartiments à ceux multi-agents ou par graphe lorsque celui-ci est aussi homogène donc plus proche possible de l'hypothèse d'homogénéité qui régit les systèmes à compartiments. Et l'on constate que la manière naturelle de calculer le R_{eff} dans un système à compartiments et la manière de calculer le R_{eff} dans un système à graphes homogènes basés sur le degré moyen coïncident raisonnablement tout au long de la propagation de l'épidémie dû aux hypothèses d'homogénéités. Nous constatons que, plus le graphe social est hétérogène, plus la différence entre les taux de reproduction effectifs se présentent ceci dès les premiers stades de l'épidémie.

Ces constatations sont montrées de façon théorique grâce à la théorie de la percolation et de façon numérique sur les graphes synthétiques et sur les graphes réels. Dans les simulations numériques, nous avons effectué des simulations avec une seule source d'infection et celle avec plusieurs sources et à la fin, nous formulons des recommandations pour le calcul du R_{eff} avec des mesures de précautions, bien qu'il existe une différence de calcul du R_{eff} dans les deux modèles, celle-ci ne semble pas dramatique une fois la propagation de l'épidémie passée sous contrôle.

Concernant les bulles, nous effectuons une modélisation très simplifiée de ces dernières dans les réseaux aléatoires de contact et constatons qu'elles ont un impact majeur dans la propagation de l'épidémie, leurs introductions dans un réseau social permettent dans un premier temps de freiner la vitesse de propagation du virus en augmentant son temps de propagation et dans un second temps de réduire le nombre maximal de personnes atteint par la maladie ceci quelle que soit la nature du réseau.

Abstract

Toutefois, nous remarquons que leurs impacts précédents n'ont pas la même intensité dans tous les réseaux, ainsi plus le réseau social sera hétérogène plus les impacts des bulles seront drastiques, ceci dû à la présence dans ces derniers des noeuds spéciaux (super-spreaders) qui sont complètement neutralisés par l'introduction des bulles dans les réseaux. Nos constatations sont faites de façon numérique sur les réseaux aléatoires synthétiques d'Erdős-Rényi et Barabási-Albert et enfin sur un réseau de contacts aléatoire disponible en ligne.

Notons que les modélisations effectuées dans le cadre de notre travail ne correspondent pas forcément à la réalité, car elles ne tiennent pas compte de certains paramètres réels. Mais toutefois, les résultats obtenus servent d'avertissement aux éventuels risques encourus lors des prises de décisions concernant les propagations des épidémies.

Mots clés : propagation d'épidémie, modèle épidémiologique par compartiments, modèle épidémiologique sur graphe, bulle sociale, taux de reproduction de base, taux de reproduction effectif, graphes homogènes, graphes hétérogènes.

Chapitre 0

Introduction générale

Les problèmes d'analyse de la transmission d'infections dans une population sont des problèmes qui remontent depuis la nuit des temps, mais qui restent toujours d'actualité à ce jour. Certaines maladies infectieuses conduisant à une épidémie déciment de façon catastrophique des vies humaines, provoquant en plus des pertes en vie humaine, des conséquences économiques, sociaux et politiques [1], comme nous démontre la crise sanitaire du coronavirus qui se vit actuellement dans chaque pays du monde.

Lors de l'apparition d'une épidémie, une multitude de questions sont directement posées par les professionnels de santé publique, parmi lesquelles : quel est son potentiel épidémique ? Doit-on s'attendre à une épidémie de grande ampleur ou simplement à juste quelques cas d'infections du virus ? Si l'épidémie est d'une grande ampleur combien de temps disposons-nous pour intervenir ? Sur quel temps s'étendra l'épidémie c'est-à-dire pendant combien de temps l'épidémie risque de rester présente et quelle est la proportion de personne de la population totale qu'elle touchera ? Quelles seront ses conséquences sociales notamment celles sur le système de santé ? [1, 2]

Il devient donc primordial de comprendre à la fois le nombre probable d'infections de la maladie et son évolution dans le temps afin d'apporter un ensemble de réponses à toutes ces questions qui relèvent à la fois de la santé publique et du système de soins de santé.

L'avènement des mathématiques comme moyen de modélisation d'épidémie remonte depuis le 20^{ème} siècle grâce à Bernoulli en 1760, qui mettait en oeuvre le premier modèle mathématique pour modéliser l'infection de la variole [3]. Ces modélisations une fois faites fournissent de nombreuses informations importantes concernant l'épidémiologie des maladies infectieuses.

L'apparition des nouvelles maladies infectieuses a permis d'améliorer le modèle défini par Bernoulli.

Chapitre 0. Introduction générale

Plusieurs approches mathématiques ont été dès lors développées afin d'étudier tant l'évolution d'une épidémie que les stratégies et mécanismes à mettre en oeuvre pour ralentir sa progression, la stopper et/ou éventuellement définitivement l'éradiquer.

Il existe deux principaux types de modèles : les premiers qualifiés de modèles déterministes qui ne tiennent pas compte du caractère aléatoire de la transmission de la maladie et les seconds qualifiés de modèles stochastiques qui contrairement au premier tiennent compte du caractère aléatoire de la transmission de la maladie [3].

Dans le premier groupe, l'approche consiste généralement à l'utilisation des modèles épidémiques compartimentaux, en règle générale, la population d'intérêt est subdivisée en un petit nombre de compartiments en fonction de l'état de l'infection et les flux entre ces compartiments sont décrits par un ensemble de faible dimension d'équations différentielles ordinaires. Il existe une multitude de modèles basés sur les compartiments et donc le plus populaire est le modèle susceptible-infecté-récupéré, introduit en 1927 moins d'une décennie après la pandémie de grippe de 1918 [4], sa popularité est due en partie à sa simplicité, qui permet aux modélisateurs d'approcher le comportement de la maladie en estimant un petit nombre de paramètres. Ce modèle vise à prédire le nombre d'individus sensibles à l'infection, infectés ou guéris de l'infection à un moment donné.

Le développement des réseaux aléatoires statique [5, 6, 7] permettant de modéliser les contacts et relations humaines ont débuté en 1956 par les chercheurs Erdos & Renyi, fournissent une base solide au second groupe de modélisation épidémiologique.

Bien que plus compliqués à analyser que ceux déterministes, ces modèles offrent un moyen naturel de décrire une population et ses interactions, les sommets constituant le réseau représentent les individus et les arêtes les interactions entre les individus qui pourraient potentiellement conduire à la transmission de l'infection. Le modèle Reed-Frost [8] développé Lowell Reed et Wade Frost est le plus populaire dans cette catégorie, toutefois ces modèles nécessitent d'imposer plusieurs hypothèses simplificatrices pour une meilleure analyse.

Les professionnels de santé publique disposent ainsi d'un ensemble de modèles aidant à mieux comprendre une épidémie, se basant sur les réponses fournies par ces modèles des décisions et stratégies sont proposées et mises en oeuvre afin d'essayer de limiter le nombre d'infections, de stopper la propagation du virus, de l'éradiquer ou de le maintenir sous un certain seuil de contrôlabilité.

Ces modèles servant donc de base de décisions, il est donc naturel de se poser les questions suivantes : fournissent-ils les mêmes résultats pour la modélisation d'une même épidémie, si oui sous quelle(s) condition(s). Le fait d'utiliser un modèle plutôt qu'un autre

pour la modélisation d’une épidémie a-t-elle des conséquences dramatiques en cas d’erreur ? Quelle stratégie mettre en oeuvre et en fonction de quel modèle ? Quels sont les risques encourus si on se trompait de modèle pour définir une stratégie ? Dans ce mémoire, nous nous intéressons principalement à la comparaison des modèles compartimentaux plus particulièrement celui *SIR* de Kermack–McKendrick avec ceux multi-agents ou par graphe principalement celui Reed-Frost pour la simulation de l’épidémie spécialement en ce qui concerne le calcul du taux de reproduction de base R_0 et le taux reproduction effectif R_{eff} qui sont d’une importance capitale dans la caractérisation et l’évolution d’une épidémie et qui malheureusement ne sont pas toujours très détaillée dans la littérature et qui pourtant sont beaucoup évoqués dans les médias.

Le mémoire est organisé comme suit :

Dans le chapitre un, nous présenterons l’évolution historique des différents modèles par compartiments sans prise en compte de certains aspects tels que le taux de mortalité et taux de natalité. Pour chacun d’eux, nous l’étudierons tout en observant leurs avantages et inconvénients pour la modélisation de certaines maladies infectieuses.

Dans le chapitre deux, nous considérerons le modèle par compartiments *SIR* Kermack-McKendrick à population constante proposé par les chercheurs Kermack & McKendrick pour expliquer l’évolution de l’épidémie de peste à Bombay en Inde [4], ce modèle est notre choix idéal, car il est un très bon compromis entre simplicité de modélisation et efficacité des résultats dans la réalité. Nous l’étudierons cette fois-ci, par rapport au chapitre un de façon détaillée et modéliserons une méthode de prévention et de lutte contre les épidémies qui est la vaccination.

Dans le chapitre trois, nous introduirons le processus de branchement à travers le modèle de Galton Watson qui nous permettra par la suite d’approximer le comportement du début d’une épidémie, de déduire ses conditions et probabilité d’extinction et enfin de prédire la proportion de personnes touchée à la fin de celle-ci.

Dans le chapitre quatre, après une brève introduction sur les différents réseaux de contacts aléatoire homogènes et hétérogènes, nous présenterons le modèle Reed–Frost qui est un modèle de propagation de maladies infectieuses sur un réseau aléatoire et qui par analogie à celui présenté au chapitre deux est un modèle *SIR*. Grâce à la théorie de la percolation, nous modéliserons également la même technique de prévention de propagation de maladies infectieuses que celle du chapitre deux et enfin nous étudierons le concept de composante géante de graphe qui nous servira par la suite pour la détermination de la taille finale d’une épidémie.

Chapitre 0. Introduction générale

Dans le chapitre cinq, nous effectuerons grâce aux simulations une comparaison des résultats tant avec vaccination que sans vaccination des relations théoriques obtenues lors des études du modèle par compartiments de Kermack-McKendrick au chapitre deux et celui du modèle multi-agents Reed-Frost lors du chapitre trois.

Le chapitre six quant à lui, présentera les différentes méthodes de calcul du taux de reproduction effectif lors de la propagation d'une épidémie dans une population tant pour des modèles par compartiments et que ceux basés sur les graphes. Une première comparaison est effectuée entre ces deux méthodes, ensuite nous introduisons la méthode d'Anne Corri et Neil Ferguson pour l'estimation du taux de reproduction effectif d'une épidémie et enfin nous comparons ses résultats avec ceux des deux différentes méthodes de calcul de taux de reproduction effectif.

Enfin dans le chapitre sept, nous introduirons le concept de bulle sociale, nous le modéliserons pour le cas des modèles épidémiologiques multi-agents et nous analyserons ses effets dans le processus de propagation épidémique. Des simulations seront effectuées pour différentes tailles de bulles et ceci sur plusieurs réseaux aléatoires de contacts très célèbres et très étudiés dans la littérature que sont les réseaux d'Erdős-Rényi et les réseaux de Barabási-Albert, afin de valider nos résultats obtenus, nous utilisons un graphe social aléatoire disponible en ligne.

Chapitre 1

Modèles compartimentaux

En Épidémiologie, les modèles compartimentaux sont l'un des premiers modèles mathématiques à avoir été utilisés pour modéliser la propagation d'une épidémie dans une population [3]. Ils sont fondés sur une subdivision de la population totale en plusieurs classes d'individus indiscernables appelées compartiments. Ce modèle consiste principalement à diviser une population en plusieurs compartiments d'individus ou groupes d'individus interagissant entre eux.

Cette approche dont le principal avantage est la simplicité de sa conception est utilisée pour modéliser de très nombreuses maladies et continue d'être un sujet de recherche actif [9], continuant ainsi de jouer un rôle crucial en épidémiologie en permettant d'analyser mathématiquement les déroulements d'épidémies en fonction de certains facteurs.

Leurs études permettent d'apprendre beaucoup sur les comportements de base des systèmes épidémiologiques et aident au raisonnement lorsque l'on doit faire face à des modèles épidémiologiques plus complexes.

Les modèles à compartiments utilisés dans ce mémoire supposent que l'effectif N de la population est constant, ainsi ils ne prennent pas en compte les naissances et les décès. Bien que des modèles de même nature plus complexes puissent prendre compte ces facteurs, cette hypothèse est plutôt raisonnable si on modélise par exemple, la propagation d'une épidémie au sein d'une population qui ne reçoit pas de nouvel individu pendant le temps de l'étude de l'épidémie. Enfin ces modèles supposent aussi une homogénéité de la population étudiée, c'est-à-dire que tout individu non contaminé court le même risque d'être exposé à la maladie qu'un individu déjà contaminé.

Dans ce chapitre nous présentons un certain nombre de modèles compartimentaux très répandus, et nous étudions précisément plus en détails le modèle de Kermack-McKendrick.

Notons que l'ensemble de ce chapitre a été inspiré du livre [2] et des articles [10, 11, 12].

Les grands classiques

La modélisation des épidémies par des modèles compartimentaux sont les plus répandus dans la littérature. Le nombre de compartiments étant l'élément caractéristique du modèle. On distingue principalement les modèles à deux, trois, quatre compartiments donc nous les présenterons dans la suite. toutefois, notons qu'il existe également des modèles particuliers à plus de quatre compartiments développés par des chercheurs pour modéliser certaines épidémies [9].

La construction de ces différents modèles, suppose certaines hypothèses qui sont :

- La première étant l'homogénéité de la population, c'est-à-dire que tous les individus ont la même probabilité de contracter l'infection et de la perdre s'il est infecté, et donc tous les individus d'un même compartiment sont identiques.
- Le nombre total d'individus est constant, pas de variation de la population totale au cours de l'étude.
- La maladie se transmet uniquement d'un individu à l'autre.

1.1 Modèles à deux compartiments : $SI(S)$

Présentation du modèle

En 1906 apparaît le premier modèle dynamique de Hamer. C'est un modèle épidémique simple où l'on considère que la population étudiée N peut être décomposée en deux classes ou compartiments :

- Les individus sains et susceptibles d'être infectés S .
- Les individus infectés ou contagieux I .

L'infection se propage par contact direct entre un individu de la classe des susceptibles avec un individu de la classe des infectés ou contagieux. Donc, il y a d'autant plus de nouveaux cas, qu'il y a plus de susceptibles S et plus d'infectés I pour les contaminer, avec un facteur k de proportionnalité appelé taux d'infection.

Chapitre 1. Modèles compartimentaux

Dans ce modèle, un individu lorsqu'il est infecté, devient infectieux et le reste jusqu'à la fin de sa vie. Cette hypothèse est raisonnable pour beaucoup de maladies dans les premières étapes de l'infection. On suppose que la population est fermée, c'est-à-dire, pour tout $t \in \mathbb{R}^+$:

$$S(t) + I(t) = N$$

Où N est constant et correspond à la taille de la population totale. Le modèle de SI est alors décrit par le système différentiel suivant :

$$\begin{cases} \frac{dS}{dt} &= -f(S(t), I(t)) \\ \frac{dI}{dt} &= f(S(t), I(t)) \end{cases}$$

Où $f(S, I)$ est le taux avec lequel la contamination se produit.

Étude du modèle :

Il est clair que f est une fonction croissante de S et I , et le modèle le plus simple est le cas où :

$$f(S, I) = \beta SI$$

Dans lequel β exprime à la fois que tous les contacts possibles n'ont pas nécessairement lieu, et que ceux-ci ne sont pas toujours à l'origine d'un nouveau cas, c'est-à-dire qu'un contact entre un susceptible et un infecté n'entraîne pas forcément une contamination. La figure 1.1 montre les différents compartiments du modèle SI ainsi que le diagramme de transmission de la maladie.

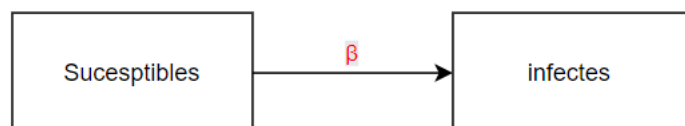


FIG. 1.1 : Représentation des compartiments du modèle SI

On a alors le système différentiel :

$$\frac{dS}{dt} = -\beta S(t)I(t) \quad (1.1)$$

$$\frac{dI}{dt} = \beta S(t)I(t) \quad (1.2)$$

Comme supposé dans les hypothèses de départ plus haut, la population totale est donc constante, c'est-à-dire à chaque instant t nous avons :

$$S(t) + I(t) = N$$

Donc

$$S(t) = N - I(t)$$

Ainsi en remplaçant $S(t)$ par sa valeur dans l'équation 1.1, on obtient :

$$\frac{dI}{dt} = -\beta(N - I(t))I(t)$$

En termes de modélisation, on traduit l'apparition d'une épidémie par l'introduction d'un individu infecté dans une population saine à l'instant initial $t = 0$, c'est-à-dire que :

$$I(0) = 1$$

La résolution de cette équation nous donne la relation :

$$I(t) = \frac{N}{1 + (N - 1) \exp(-\beta N t)}$$

La dérivée de l'équation précédente par rapport à t s'écrit :

$$\frac{dI(t)}{dt} = \beta \frac{(N - 1)N^2 \exp(-\beta N t)^2}{(N - 1) \exp(-\beta N t)}$$

La courbe de l'évolution de cette fonction dans le temps fournit la courbe épidémiologique.

La figure 1.2 illustre l'évolution de l'épidémie dans un modèle à deux compartiments S et I sur une période de 150 jours.

Chapitre 1. Modèles compartimentaux

Les paramètres choisis pour cette simulation sont tirés de la source [2]. La population initiale est constituée de 1000 individus sains et 1 individu infecté. Le taux d'infection β est de 0.2.

Nous remarquons que la population passe progressivement du compartiment S vers le compartiment I .

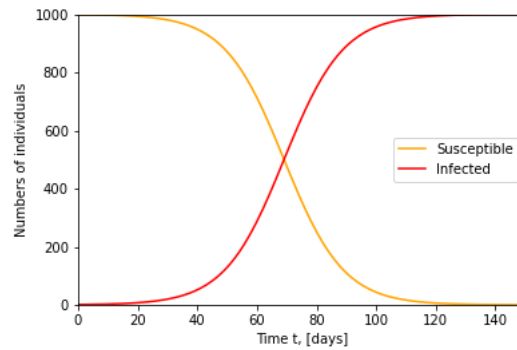


FIG. 1.2 : Simulation du modèle SI pour $N = 1000, I_0 = 1, \beta = 0.2$

Le modèle SIS est une modification du modèle SI , il considère les mêmes compartiments que dans le modèle SI , il permet un retour du compartiment des infectés I au compartiment des susceptibles S . Il s'ensuit que ces individus peuvent être à nouveau infectés.

Un modèle SIS peut être utilisé pour la modélisation et la propagation d'une épidémie telle que le rhume [10].

La figure 1.3 représente les différents compartiments des modèles SIS ainsi que les différentes interactions entre compartiments de ce dernier.

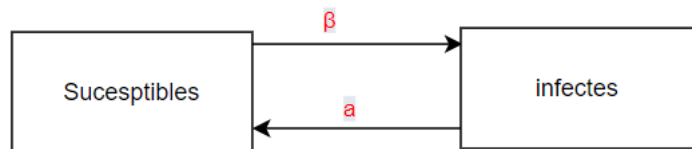


FIG. 1.3 : Représentation des compartiments du modèle SIS

Le modèle SIS est alors décrit par le système différentiel suivant :

$$\begin{aligned}\frac{dS}{dt} &= -\beta S(t)I(t) + aI(t) \\ \frac{dI}{dt} &= \beta S(t)I(t) - aI(t)\end{aligned}$$

Où β et a sont des réels strictement positifs.

Le taux d'infection β est construit de la même manière que pour le modèle SI . Tandis que le taux de guérison a qui est en quelque sorte la probabilité de retour du compartiment I au compartiment S correspond au taux de transfert entre I et S . Ce taux est directement proportionnel au nombre d'individus infectés.

Notons que ces deux modèles simples, SI et SIS , ne sont pas assez réalistes, car en général lorsque l'on guérit d'une maladie, on acquiert généralement une immunité temporaire ou définitive face à celle-ci.

1.2 Modèles à trois compartiments : $SIR(S)$

Les modèles à trois compartiments dérivent principalement de ceux à deux compartiments vus précédemment, la principale différence étant la présence d'un troisième compartiment R : le compartiment pour les individus rétablis qui sont alors immunisés.

Avec ce modèle nous pouvons modéliser des épidémies contre lesquelles il est possible d'être immunisé de façon temporelle, c'est-à-dire qu'un individu rétabli qui perd son immunité sera considéré comme un individu sain pouvant être à nouveau infecté ou définitive c'est-à-dire qu'un individu guéri le restera jusqu'à la fin de l'épidémie.

Présentation du modèle :

Généralement connu sous le nom de modèle de Kermack–McKendrick, il peut être utilisé pour étudier de nombreuses maladies telles que la rougeole, les oreillons.

Nous le présentons dans cette section, mais nous l'étudierons de façon plus détaillée dans le chapitre suivant.

Ce modèle est composé de trois groupes d'individus distincts :

- Un groupe d'individus sains S susceptibles d'être infectés.
- Un groupe d'individus infectés I .
- Un groupe d'individus rétablis R , donc immunisés ne pouvant plus être contaminés.

Ce modèle ne contient que deux paramètres que sont β le taux d'infection et γ taux guérison, dont les valeurs sont déterminées à partir des données observées. Les hypothèses dans ce modèle sont les mêmes que ceux à deux compartiments vus précédemment, la population totale est donc supposée constante, c'est-à-dire qu'à chaque instant t , l'équation suivante est vérifiée :

$$S(t) + I(t) + R(t) = N$$

la figure 1.5 décrit les différents mouvements entre les différents compartiments du modèle *SIR* correspondant à l'infection des personnes saines au rétablissement de personnes infectées.

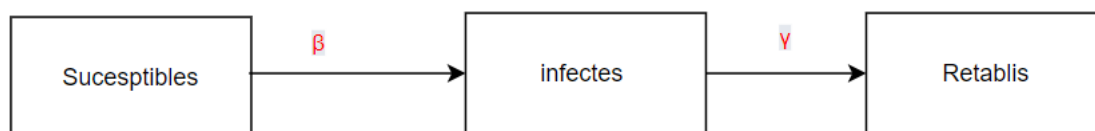


FIG. 1.4 : Représentation des compartiments du modèle *SIR*

Le modèle de Kermack-Kendrick se formule alors par le système différentiel suivant :

$$\begin{cases} \frac{dS}{dt} &= -\beta S(t)I(t) \\ \frac{dI}{dt} &= \beta S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t) \end{cases}$$

Ainsi la variation du nombre de personnes contaminées I tient compte de deux termes. Le premier est identique à celui du modèle *SI* et le second pourrait faire l'objet de plusieurs options. L'hypothèse généralement retenue est que, dans l'intervalle de temps dt , une proportion fixe d'individus cesse d'être considérée comme infectée. Cette proportion concerne en principe des individus guéris et immunisés en tout état de cause sortis du processus d'infection. Cette proportion désignée par le paramètre γ peut valoir zéro et l'on est ramené au cas du modèle *SI*. Il peut également être égal à 1, auquel cas tous les infectés sont sortis du processus durant l'intervalle de temps dt . Dans la plupart des cas, ce paramètre γ varie entre ces deux valeurs.

Étude du modèle :

Le modèle *SIR* est appelé modèle déterministe général, son étude permet d'établir le théorème du seuil.

Son système d'équation définit précédemment est muni de la condition initiale :

$$S(0) = S_0,$$

$$I(0) = I_0,$$

$$R(0) = 0$$

Telle que $I_0 \approx 0$ ainsi $S_0 \approx N$. Cela revient à considérer une population de susceptibles dans laquelle on introduit un petit nombre d'infectieux. Puisque la population totale N est constante par hypothèse, c'est-à-dire que $S(t) + I(t) + R(t) = N$, le système d'équation se réduit donc à l'étude de :

$$\begin{aligned}\frac{dS}{dt} &= -\beta S(t)I(t) \\ \frac{dI}{dt} &= \beta S(t)I(t) - \gamma I(t)\end{aligned}$$

considéré sur l'ensemble suivant :

$$\Delta = \{0 \leq S \leq N, 0 \leq I \leq N, 0 \leq S + I \leq N\}$$

En divisant la deuxième équation du système par γI , on obtient :

$$\frac{I(t)}{\gamma I(t)} = \frac{\beta S(t)}{\gamma} - 1 \quad (1.3)$$

Alors, pour $\frac{\beta S(t)}{\gamma} > 1$, chaque individu infecté contaminera plus d'un individu susceptible et la maladie se propagera à un nombre toujours grandissant d'individus. Il en sera ainsi jusqu'à ce que le nombre de susceptibles $S(t)$, soit tel que $\frac{\beta S(t)}{\gamma} < 1$. Nous y reviendrons plus en détails dans le chapitre suivant.

La figure 1.5 illustre la dynamique de propagation d'une épidémie dans une population pour le cas d'un modèle à trois compartiments S , I et R sur une période t de 150 jours.

Les paramètres choisis pour cette simulation sont également tirés de la source [2]. La population totale initiale est de 1000 individus sains donc 1 est infecté et 0 individu guéri. Le taux d'infection β est de 0.2 et le taux de retrait γ de 0.06.

Nous constatons que le nombre de personnes contaminées augmente jusqu'à un nombre maximum et ensuite commence à décroître vers 0. Le nombre d'individus susceptibles à la fin tend vers 0 et la population passe progressivement vers le compartiment rétabli R .

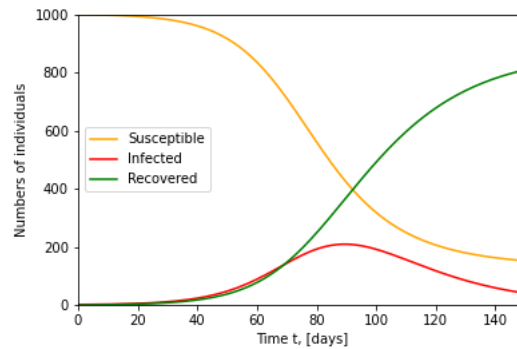


FIG. 1.5 : Simulation du modèle SIR pour $N = 1000, I_0 = 1, R_0 = 0, \beta = 0.1, \gamma = 0.04$

Le modèle $SIRS$ est une combinaison des modèles SIS et SIR . Il comporte toujours trois compartiments distincts comme le modèle SIR . Néanmoins, cette modélisation permet aux individus guéris de redevenir susceptibles et donc potentiellement recontaminés. La Figure 1.6 représente la dynamique de ce modèle

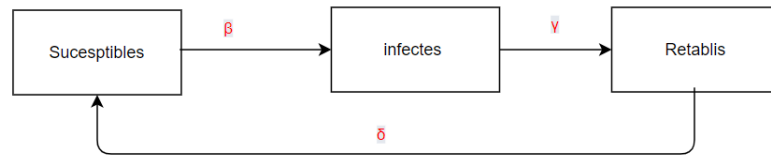


FIG. 1.6 : Représentation des compartiments du modèle $SIRS$

Le modèle de $SIRS$ se formule alors par le système différentiel suivant :

$$\begin{cases} \frac{dS}{dt} = -\beta S(t)I(t) + \sigma R(t) \\ \frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt} = \gamma I(t) - \sigma R(t) \end{cases}$$

Où β , γ et σ sont des réels positifs. σ représente le taux de transfert entre le compartiment des individus rétablis et celui des individus susceptibles. Il est proportionnel au nombre d'individus rétablis et donc représente la part d'individus guéris qui vont perdre leur immunité au cours d'une unité de temps.

1.3 Modèles à quatre compartiments : $SEIR(S)$

Présentation du modèle :

Le modèle $SEIR$ est utilisé pour les maladies pour lesquelles la période pendant laquelle l'individu est infecté, mais n'est pas encore infectieux est longue.

Pendant la phase d'incubation, un individu susceptible qui vient d'être infecté par le virus est dans un état latent, c'est-à-dire qu'il est infecté, mais ne peut pas encore transmettre la maladie à une autre personne susceptible. Pendant cette phase, les premiers symptômes de la maladie apparaissent. Puis vient une période d'infection, pendant laquelle l'individu initialement infecté devient infectieux, c'est-à-dire qu'il est infecté et qu'il peut transmettre la maladie à une autre personne susceptible. L'individu présente également les symptômes de la maladie durant cette période.

Enfin apparaît la phase de guérison qui marque la fin de la maladie. L'individu peut être guéri, décédé ou simplement retiré. Ainsi un modèle comprenant les états susceptibles, exposés, infectés et retirés semble être conforme à la dynamique de plusieurs maladies dont nous pouvons citer entre autres la plus récente, la covid-19.

Le modèle susceptibles-exposés-infectés-rétablis est représenté par le schéma de transmission de la figure suivante :



FIG. 1.7 : Représentation des compartiments du modèle $SIER$

Étude du modèle :

Soit β la probabilité de passage entre le compartiment des susceptibles et le compartiment des exposés, λ la probabilité de passage entre le compartiment des exposés et le compartiment des infectés et γ la probabilité de passage entre le compartiment des infectés et le compartiment des rétablis.

Le système régissant le modèle $SEIR$ s'écrit donc de la manière suivante :

$$\begin{cases} \frac{dS}{dt} &= -\beta S(t)I(t) \\ \frac{dE}{dt} &= -\beta S(t)I(t) - \lambda E(t) \\ \frac{dI}{dt} &= \lambda E(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t) \end{cases}$$

Comme dans les modèles vu précédemment, la population totale est supposée constante, donc à chaque instant t on a :

$$N = S(t) + E(t) + I(t) + R(t)$$

Dans le cadre d'un modèle *SEIR* déterministe, les probabilités de passage entre les compartiments (β , λ et γ) sont fixées.

La figure 1.8 représente l'évolution des populations dans les compartiments S , E , I et R lors de la propagation d'une épidémie. Les paramètres choisis pour cette simulation sont également tirés de la source [2].

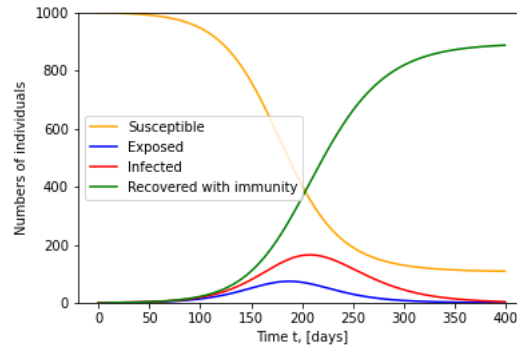


FIG. 1.8 : Simulation du modèle *SEIR* pour $N = 1000$, $E_0 = 0$, $I_0 = 1$, $R_0 = 0$, $\beta = 0.1$, $\gamma = 0.04$, $\lambda = 0.1$

Nous remarquons que le temps nécessaire au déclenchement de l'épidémie est plus élevé que dans le modèle *SIR*. Au cours du temps, le nombre de susceptibles décroît alors que le nombre de rétablis ou guéris croît. Le nombre d'exposés et d'infectieux croît jusqu'à un maximum puis commence à décroître. À la fin de l'épidémie, l'ensemble de la population est dans le compartiment guéri et le nombre d'individus dans les autres compartiments tend vers zéro.

Chapitre 1. Modèles compartimentaux

Comparativement au modèle *SIRS*, le modèle *SEIRS* est une combinaison des modèles *SIS* et *SEIR*. Il comporte quatre compartiments distincts comme le modèle *SEIR*. Néanmoins, cette modélisation permet aux individus rétablis de redevenir susceptibles d'être infectés par un individu infectieux et donc de recommencer le processus de migration à travers différents compartiments. Ainsi La Figure 1.9 représente la dynamique de ce modèle.

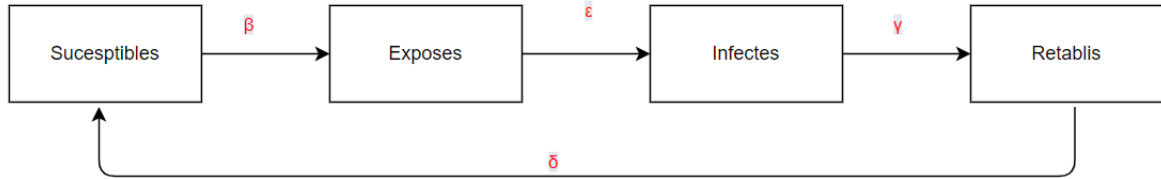


FIG. 1.9 : Représentation des compartiments du modèle *SEIRS*

Ainsi son système d'équation différentielle est le suivant :

$$\begin{cases} \frac{dS}{dt} = -\beta S(t)I(t) + \sigma R(t) \\ \frac{dE}{dt} = -\beta S(t)I(t) - \lambda E(t) \\ \frac{dI}{dt} = \lambda E(t) - \gamma I(t) \\ \frac{dR}{dt} = \gamma I(t) - \sigma R(t) \end{cases}$$

Les modèles *SEIR* ou *SEIRS* comparativement aux modèles *SI(S)* et *SIR(S)* vus précédemment sont plus proches de la réalité, mais ont pour inconvénient d'être plus complexes à mettre en oeuvre que ces derniers.

De ce fait nous choisissons donc pour la suite de notre travail le modèle *SIR* comme notre modèle par compartiments, car il présente l'avantage d'être un bon compromis entre simplicité et réalisme. Nous le décrirons plus en détails dans le chapitre suivant.

Chapitre 2

Modèle de Kermack-McKendrick

Proposé [4] à Londres et Cambridge en 1927 par deux chercheurs Kermack médecin écossais spécialisé en épidémiologie et McKendrick écossais lui aussi et mathématicien pour expliquer l'évolution de l'épidémie de peste à Bombay en Inde en 1905-1906. L'objectif visé par leurs travaux était de comprendre pourquoi l'épidémie n'avait pas touché toute la population entière.

Ce modèle est un très bon compromis entre simplicité et efficacité, utilisé en épidémiologie pour calculer le nombre de personnes sensibles, infectées et guéries dans une population frappée par une épidémie. Il est particulièrement approprié à utiliser sous les hypothèses suivantes [13] :

- La population est fixe.
- La seule façon pour une personne de quitter le groupe sensible est d'être infectée. La seule façon pour une personne de quitter le groupe infecté est de se remettre de la maladie. Une fois qu'une personne s'est rétablie, elle a reçu l'immunité et ceci de façon définitive.
- L'âge, le sexe, le statut social et la race n'affectent pas la probabilité d'être infecté.
- Il n'y a pas d'immunité héréditaire. Les membres de la population se mélangent de façon homogène et ont les mêmes interactions entre eux au même degré.
- Les taux naturels de natalité et de mortalité sont inclus.
- Toutes les naissances se font exclusivement dans la classe sensible.
- Le taux de mortalité est égal pour les membres des trois classes, et on suppose que les taux de natalité et de mortalité sont égaux de telle sorte que la population totale soit stationnaire tout au long de l'étude.

Chapitre 2. Modèle de Kermack-McKendrick

Le modèle de Kermack-McKendrick s'écrit facilement à l'aide d'équations différentielles ordinaires qui décrivent l'évolution temporelle du nombre d'individus dans chaque compartiment SIR . Ainsi de façon descriptive, le processus est le suivant :

Un individu est d'abord sain, donc présent dans le compartiment S ; il est ensuite infecté I avec une probabilité β c'est-à-dire qu'il quitte le compartiment des susceptibles vers celui des infectés ; enfin, il guérit avec une probabilité γ , donc passe du compartiment des infectés vers le compartiment des guéris ou rétablis R , β représente le taux de contamination tandis que γ représente le taux de guérison.

À chaque instant t les valeurs $S(t)$, $I(t)$ et $R(t)$ représentent respectivement le nombre d'individus susceptibles, infectés et guéris ou rétablis.

Le système d'équations différentielles ordinaire du modèle SIR est donc le suivant [4] :

$$\begin{cases} \frac{dS}{dt} &= -\beta S(t)I(t) \\ \frac{dI}{dt} &= \beta S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t) \end{cases}$$

C'est un modèle qui ne contient que deux paramètres : β le taux d'infection et γ le taux de retrait, dont les valeurs sont principalement déterminées à partir des données observées lors d'une épidémie.

Notons que comme par hypothèse la population totale est constante, à chaque instant t l'équation suivante est toujours vérifiée :

$$S(t) + I(t) + R(t) = N \quad (2.1)$$

Prenons comme conditions initiales du système d'équation précédent :

$$S(0) = S_0$$

$$I(0) = I_0$$

les valeurs S_0 et I_0 sont strictement positives.

L'hypothèse sur le nombre de personnes guéries peut être faite de façon suivante n'ayant pas d'influence sur le comportement du modèle.

$$R(0) = 0$$

Les solutions analytiques des équations du modèle par compartiments de Kermack-McKendrick sont impossible à trouver, grâce aux conditions initiales précédemment définies un ensemble de propriétés peuvent être extraites de ce système comme nous le ferons dans les prochaines sections.

2.1 Proportionnalité du modèle SIR

Le modèle SIR est un modèle continu, les valeurs à chaque instant t dans ses compartiments sont un ensemble de valeurs décimales difficilement compréhensibles. Dans la plupart des modélisations en vu d'une interprétation, il est plus pertinent de parler de proportion d'individus susceptible, infectée ou guérie que de nombre d'individus.

En termes de proportionnalité, les équations des systèmes SIR vue précédemment se réécrivent en divisant chacune des variables S, I, R du système d'équation par la population totale N , ainsi nous obtenons le système d'équations suivant :

$$\begin{cases} \frac{ds}{dt} &= -\beta s(t)i(t) \\ \frac{di}{dt} &= \beta s(t)i(t) - \gamma i(t) \\ \frac{dr}{dt} &= \gamma i(t) \end{cases}$$

Où les variables, $s(t)$, $i(t)$ et $r(t)$ représentent, respectivement, la proportion d'individus susceptible, infectée et rétablie dans la population et β et γ restant inchangés.

2.2 Taux de reproduction basique : R_0

L'une des propriétés les plus importantes de la modélisation mathématique d'une épidémie est le taux de reproduction de base [14] aussi appelé taux de reproductivité généralement noté R_0 dans la littérature [15].

En effet, il correspond au nombre moyen d'individus qu'une personne contagieuse pourra à son tour infecter tant qu'elle sera contagieuse. Son importance en épidémiologie vient du fait qu'il définit une quantité appelé seuil, permettant de déterminer si une épidémie aura lieux ou non.

Définissons ce seuil pour le modèle SIR , réécrivons la seconde équation $\frac{dI}{dt}$ de notre système sous la forme suivante en mettant $I(t)$ en facteur, nous obtenons :

$$\frac{dI(t)}{dt} = I(t)(\beta S(t) - \gamma) \quad (2.2)$$

Cette équation 2.2 nous permet de déduire la condition nécessaire au déclenchement de l'épidémie. Ainsi :

Si $\frac{dI(t)}{dt} < 0$ c'est-à-dire $S(t) \leq \frac{\gamma}{\beta}$, le nombre d'individus infectés va diminuer au fur et à mesure et l'épidémie va progressivement s'arrêter et disparaître.

Sinon le nombre d'individus infectés va augmenter au fur et à mesure, ce qui signifie que l'épidémie débute.

Nous pouvons réécrire notre condition $S(t) \leq \frac{\gamma}{\beta}$ en :

$$S(t) \leq \frac{\gamma}{\beta} \Rightarrow \frac{\beta}{\gamma} S(t) \leq 1 \quad (2.3)$$

Au temps initial $t = 0$, la plus grande proportion de la population se trouve dans le compartiment susceptible, cela se traduit par la relation suivante :

$$S_0 \approx N \quad (2.4)$$

Le taux de reproduction R_0 est déduit de l'équation 2.3 en remplaçant la variable $S(t)$ par sa valeur approximative établit précédemment :

$$R_0 = \frac{\beta}{\gamma} N \quad (2.5)$$

Ainsi,

- Si $R_0 \leq 1$ aucune chance d'avoir une épidémie
- Si par contre $R_0 > 1$, il y aura forcément épidémie.

Le tableau ci-dessous tiré de la référence [16] donne quelques valeurs de R_0 pour quelques maladies :

Polio	Grippe	VIH SIDA	Rougeole	Oreillons
5 – 7	2 – 3	2 – 5	12 – 18	4 – 7

En plus de l'apport de l'information sur la condition de déclenchement d'une épidémie, le taux de reproduction de base permet également d'évaluer la gravité d'une épidémie.

Ainsi, plus le taux de reproduction basique R_0 d'une maladie sera élevé, plus il y aura d'individus infectés comme nous le ferons dans la section finale de l'épidémie.

2.3 Taux de croissance exponentielle du modèle *SIR*

Les paramètres caractérisant une épidémie sont généralement difficile à estimer. les paramètres tels que la taille de la population N sont précisément connus auprès des instituts nationaux de statistique de chaque pays ou chaque région, la valeur du taux de guérison γ directement estimable grâce aux données des hôpitaux, mais par contre le taux de contamination β reste difficile à estimer, car il dépend fortement de la maladie et du comportement de la population face à celle-ci.

Pour remédier à ce problème, les chercheurs ont recours au taux de croissance exponentielle [2], permettant de déterminer les premiers instants de l'épidémie et donc d'approximer le taux de contamination β pour la population étudiée.

Nous avons vu qu'au temps initial $t = 0$, la population est en très grande partie constituée d'individus appartenant au compartiment des susceptibles, ce qui nous a permis d'établir l'approximation suivante : $S_0 \approx N$. Ainsi l'équation $\frac{dI}{dt}$ peut être approximée aux premiers instants par :

$$\begin{aligned}\frac{dI(t)}{dt} &\approx \beta N I(t) - \gamma I(t) \\ &= I(t)(\beta N - \gamma) \\ &= I(t)\gamma(R_0 - 1)\end{aligned}$$

Ainsi :

$$\frac{dI(t)}{I(t)} = \gamma(R_0 - 1)dt$$

La solution de cette équation permet d'écrire $I(t)$ sous la forme d'une fonction exponentielle définit de façon suivante :

$$I(t) = I_0 \exp^{\gamma(R_0 - 1)t} \quad (2.6)$$

Posons K comme le coefficient en t dans la fonction *exp* de la relation précédente donc :

$$K = \gamma(R_0 - 1)$$

Ainsi, K définit le taux de croissance exponentielle du modèle et peut donc être déterminé dès le début de la maladie. Une fois déterminé, il permettra donc d'approximer la valeur

du taux de contamination de la maladie β .

Puisse que :

$$K = \gamma(R_0 - 1)$$

$$\Rightarrow K = \beta N - \gamma$$

$$\Rightarrow \beta = \frac{K + \gamma}{N}$$

Des erreurs présentes dans les données et/ou valeurs manquantes sur l'ensemble des données recueillies peuvent éloigner cette valeur approximative de β de celle réelle.

2.4 Pic de l'épidémie

Lorsqu'une épidémie survient, le nombre d'individus infectés va progressivement augmenter au fil du temps jusqu'à atteindre un nombre maximal. Une fois ce nombre atteint elle va ensuite commencer à diminuer au fur et à mesure jusqu'à tendre vers la valeur de 0.

C'est cette valeur maximale de I qui est appelée pic de l'épidémie ou pic épidémique. Il est communément noté I_{max} et il est atteint lorsque $\frac{dI}{dt} = 0$. Ce qui arrive d'après la deuxième équation du système d'équation différentielle ordinaire du modèle SIR lorsque :

$$S = \frac{\gamma}{\beta} \quad (2.7)$$

En divisant la seconde équation du système par la première, nous obtenons :

$$\frac{\frac{dI(t)}{dt}}{\frac{dS(t)}{dt}} = \frac{dI}{dS}$$

$$\frac{dI}{dS} = \frac{\beta SI - \gamma I}{-\beta SI}$$

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta S}$$

L'intégration de cette dernière par rapport à S donne :

$$I = -S + \frac{\gamma}{\beta} \ln S + \text{constante} \quad (2.8)$$

Chapitre 2. Modèle de Kermack-McKendrick

Les conditions initiales notamment celle $I(0) = I_0$ nous permettent de terminer la constante, qui vaut grâce à l'équation précédente :

$$constante = I_0 + S_0 - \frac{\gamma}{\beta} \ln S_0$$

Les équations 2.7 et 2.8 permettent ainsi d'écrire la valeur maximale du nombre d'individus d'infectés atteint par I avant de commencer à diminuer.

$$I_{max} = -\frac{\gamma}{\beta} + \frac{\gamma}{\beta} \ln\left(\frac{\gamma}{\beta}\right) + I_0 + S_0 - \frac{\gamma}{\beta} \ln(S_0) \quad (2.9)$$

Nous remarquons que la valeur I_{max} dépend du taux de reproduction de base R_0 vu dans la section précédente, Ainsi le pic de l'épidémie sera d'autant plus grand que le taux de reproduction de base le sera comme l'illustre la figure 2.1 suivante.

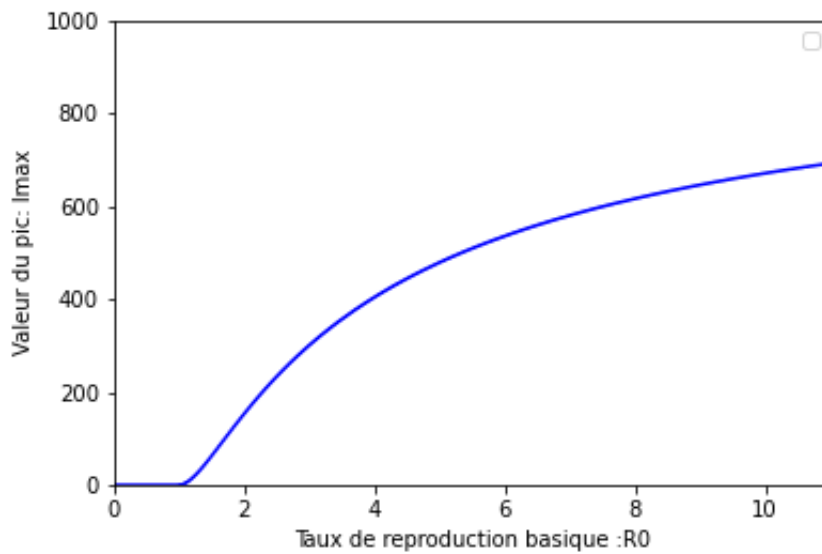


FIG. 2.1 : Pic épidémique I_{max} en fonction du R_0 pour le modèle SIR

2.5 Taille finale d'une épidémie

Dans cette section, nous nous intéressons à la taille finale de l'épidémie, c'est-à-dire le nombre de personnes total touché par une épidémie lorsque celle-ci se termine, nous essayons de prédire l'ampleur d'une épidémie connaissant son taux de reproduction R_0 . En effet, dans la suite, nous démontrerons que la proportion d'individus susceptibles n'ayant pas été touchée par une épidémie donnée dans une population donnée se stabilise à la fin de l'épidémie.

Pour ce fait, nous introduisons deux nouvelles variables S_∞ représentant le nombre d'individus n'ayant pas été infectés par l'épidémie, une fois celle-ci terminée.

$$S_\infty = \lim_{t \rightarrow +\infty} S(t)$$

Et R_∞ le nombre total de personnes touchées et donc guéries à la fin de l'épidémie.

$$R_\infty = \lim_{t \rightarrow +\infty} R(t)$$

L'équation $\frac{dS(t)}{dt}$ s'écrit :

$$\frac{dS(t)}{dt} = -\beta S(t)I(t)$$

Comme le nombre d'individus susceptibles $S(t)$ est toujours strictement positif, effectuons la division de l'équation précédente par $S(t)$ et réarrangeons le résultat obtenu, nous obtenons donc :

$$\frac{dS(t)}{S(t)} = -\beta I(t)dt$$

Intégrons ensuite par rapport à t , dans l'intervalle $[0, +\infty]$:

$$\ln S(t) \Big|_0^\infty = -\beta \int_0^\infty I(t) dt$$

Rappelons que $S(0) = S_0$ et $S_\infty = \lim_{x \rightarrow +\infty} S(t)$.

Ainsi la relation précédente devient :

$$\ln S_\infty - \ln S_0 = -\beta \int_0^\infty I(t) dt$$

$\Rightarrow$

$$\ln \frac{S_\infty}{S_0} = -\beta \int_0^\infty I(t) dt \quad (2.10)$$

Effectuons une somme des deux premières équations du système d'équation *SIR* vu plus haut, c'est-à-dire : $\frac{dS}{dt} + \frac{dI}{dt}$, nous obtenons :

$$\frac{dSt}{dt} + \frac{dIt}{dt} = -\gamma I(t)$$

En intégrant par rapport à t dans l'intervalle $[0, +\infty]$:

$$S(t) + I(t) \Big|_0^\infty = -\gamma \int_0^\infty I(t) dt$$

Initialement $I(0) = I_0$ et définissons I_∞ tel que :

$$I_\infty = \lim_{t \rightarrow +\infty} I(t)$$

Ainsi :

$$S(t) + I(t) \Big|_0^\infty = S_\infty + I_\infty - S_0 - I_0$$

Nous savons qu'à la fin de l'épidémie, aucune infection n'a plus lieu donc le nombre d'infectés à cet instant vaut zéro, c'est-à-dire que $I_\infty = 0$.

Au temps $t = 0$, c'est-à-dire au tout début de l'épidémie, nous savons que $S_0 + I_0 = N$ car $R_0 = 0$ puisque nous n'avons pas d'individus guéris au temps initial.

Nous obtenons donc :

$$S(t) + I(t) \Big|_0^\infty = S_\infty - N$$

Finalement :

$$S_\infty - N = -\gamma \int_0^\infty I(t) dt \quad (2.11)$$

Les deux équations 2.10 et 2.11 combinées conduisent à la relation suivante :

$$\begin{aligned} \ln \frac{S_0}{S_\infty} &= \frac{\beta}{\gamma} \left(N - S_\infty \right) \\ &= \frac{\beta}{\gamma} N \left(1 - \frac{S_\infty}{N} \right) \end{aligned}$$

Chapitre 2. Modèle de Kermack-McKendrick

Comme vu dans la section précédente, nous savons que le taux de reproduction $R_0 = \frac{\beta}{\gamma}N$, en remplaçant dans l'équation précédente, on obtient :

$$\ln \frac{S_0}{S_\infty} = R_0 \left(1 - \frac{S_\infty}{N} \right) \quad (2.12)$$

L'équation 2.12 montre que $S_\infty > 0$ à la fin de l'épidémie et elle est bien dépendante du taux de reproduction de base R_0 .

En termes de proportion, c'est-à-dire que $S_0 \approx 1$ et $N = 1$, on obtient :

$$S_\infty = e^{-R_0(1-S_\infty)} \quad (2.13)$$

Ainsi par déduction, la taille de l'épidémie R_∞ est donnée par la relation suivante sachant que $I_\infty=0$.

$$R_\infty = N - S_\infty \quad (2.14)$$

En termes de proportion :

$$1 - R_\infty = e^{-R_0 R_\infty} \quad (2.15)$$

2.6 Illustration : la grande peste à Eyam

Dans cette section, nous illustrons le fonctionnement du modèle *SIR* en nous basant sur des données réelles tirées des références [17, 2].

Le village d'Eyam près de Sheffield, en Angleterre, a subi une épidémie de peste bubonique en 1665 – 1666 dont la source est généralement considérée comme la grande peste de Londres. Comme des enregistrements détaillés étaient conservés et que la communauté était persuadée de se mettre en quarantaine pour essayer d'empêcher la propagation de la maladie à d'autres communautés, la maladie d'Eyam a été utilisée comme étude de cas pour la modélisation *SIR* [2].

En mesurant le temps en mois avec une population initiale de 7 infectieux, 254 susceptibles au départ et une population finale de 83 . Les valeurs des susceptibles et les infectieux à Eyam sont données pour différentes dates, en commençant par $S(0) = 254$, $I(0) = 7$, comme indiqué dans le tableau suivant :

Date	Susceptibles	infectés
19 juin	254	7
3/4 juillet	235	14.5
19 juillet	201	22
3/4 août	153.5	29
19 août	121	21
3/4 septembre	108	8
19 septembre	97	8
4/5 octobre	inconnu	inconnu
20 octobre	83	0

TAB. 2.1 : Données sur la peste d'Eyam

Le temps durant lequel une personne est infectée et contagieuse, a été estimé à 11 jours. L'épidémie se déroule entre les mois allant de juin à octobre en considérant le mois comme unité de temps, cette estimation de 11 jours est de 0.3666 mois. Cette information permet de déterminer ainsi le paramètre γ du modèle *SIR* qui correspond au taux de guérison.

$$\gamma = \frac{1}{0.3666} = 2.72$$

Déterminons grâce à la relation (2.12), le taux de reproduction basique R_0 de cette épidémie.

Sachant que $S(0) = 254$, $S_\infty = 83$, $I(0) = 7$ donc $N = 261$

$$\ln \frac{S_0}{S_\infty} = R_0 \left(1 - \frac{S_\infty}{N} \right)$$

$$R_0 = \ln \frac{S_0}{S_\infty} \frac{1}{\left(1 - \frac{S_\infty}{N} \right)}$$

$$R_0 = 1.64$$

Ainsi le taux de contamination β déduit grâce au taux de reproduction R_0 vaut 0.0177. Maintenant que nous avons les paramètres β et γ du modèle, La figure 2.2 reprend les résultats de la simulation du modèle *SIR* sur les données de la grande peste à Eyam.

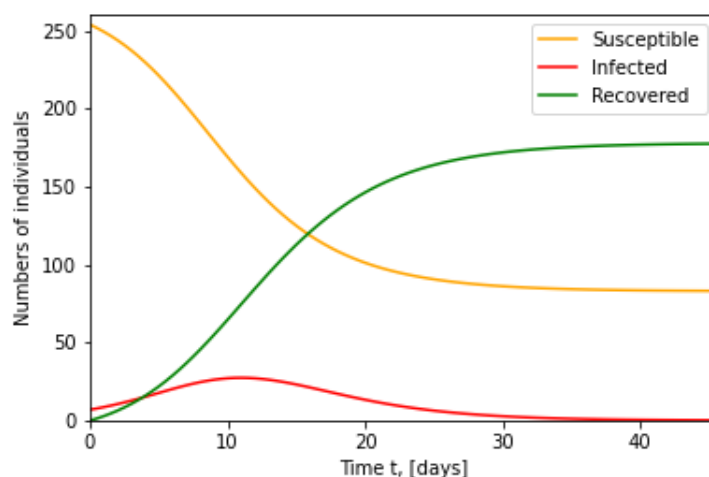


FIG. 2.2 : Simulation du modèle *SIR* pour la grande peste à Eyam.

2.7 Vaccination et immunisation collective

Dans cette section nous nous intéressons à l'une des méthodes généralement utilisées pour éradiquer une épidémie : la vaccination [18]

Selon les hypothèses, plusieurs méthodes existent pour la modélisation de cette technique dans les modèles épidémiologiques par compartiments.

Dans ce mémoire, nous supposons dans un premier temps que les individus vaccinés le sont avant l'apparition de la maladie, ensuite qu'un individu une fois vacciné acquiert une immunité à vie donc définitive face à la maladie, donc de ce fait ne peut plus être infectés, ni propager la maladie.

Ce qui n'est pas toujours le cas en réalité, car un individu peut ne pas souffrir des symptômes grâce à un vaccin, mais tout de même être porteur et transmetteur. C'est l'une des difficultés qui se pose pour la covid-19 notamment.

Et donc, grâce aux hypothèses effectuées, un moyen simple de la modéliser serait de réduire la taille totale de la population d'une fraction p de la population protégée contre l'infection. Donc cette fraction p est la proportion de la population sensible qui est immunisée avant le déclenchement d'une épidémie.

Chapitre 2. Modèle de Kermack-McKendrick

Comme le montre la figure 2.3, la vaccination dans ce modèle équivaut à un transfert d'une proportion de la population susceptible vers le compartiment récupéré : ainsi donc un individu vacciné ne peut ni infecter ni être infecté.

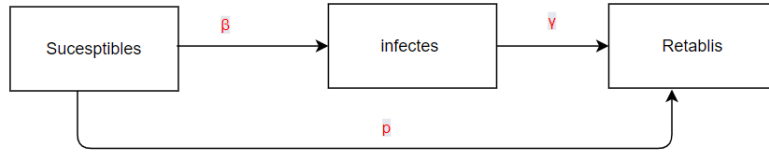


FIG. 2.3 : Modèle SIR avec vaccination

En supposant qu'une proportion p de la population est vaccinée nous obtenons donc ainsi le système suivant :

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta(1-p)SI \\ \frac{dI(t)}{dt} &= \beta(1-p)SI - \gamma I\end{aligned}\tag{2.16}$$

Comme la nouvelle population des susceptibles est maintenant de $(1-p)S$, par conséquent le nouveau taux de reproduction basique R'_0 est tel que :

$$R'_0 = (1-p)R_0$$

Comme nous l'avons vu dans les sections précédentes, les valeurs du taux de reproduction basique permettent de déterminer si oui ou non il y aura épidémie. Ainsi, pour éviter que l'épidémie explose donc ait lieu, il faut que R'_0 soit inférieur à 1, c'est-à-dire il faudrait choisir une proportion critique p_c appelée seuil critique de vaccination de telle sorte que le taux de reproduction de base vérifie cette relation : $R'_0 < 1$.

$$(1-p_c)R_0 < 1$$

Soit :

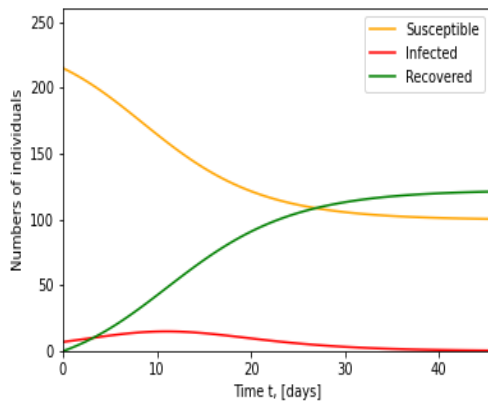
$$p_c > 1 - \frac{1}{R_0}$$

Ce résultat nous démontre ainsi qu'il n'est pas nécessaire de vacciner toute une population comme on pourrait le croire pour éradiquer une épidémie, mais plutôt de vacciner une proportion suffisante p_c de la population susceptible permettant de ramener le nombre de

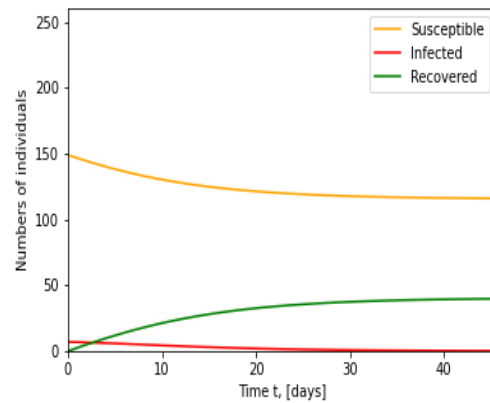
Chapitre 2. Modèle de Kermack-McKendrick

personnes susceptibles en dessous du seuil d'épidémie. C'est ce que les chercheurs appellent l'immunité de groupe.

Les figures 2.4a et 2.4b illustrent des simulations pour la grande peste à Eyam, avec des proportions de vaccinations respectives de 15% et 40% correspondant au seuil de vaccination critique pour cette épidémie.



(a) Proportion de vaccination $p = 15\%$



(b) Seuil de vaccination $p_c = 40\%$

FIG. 2.4 : Simulation du modelé *SIR* avec pourcentage de vaccination et immunité acquise

Chapitre 3

Processus de branchement (Branching process)

Introduit en 1873 par Sir Galton-Watson, dont l'objectif visé était l'étude de la statistique liée à l'extinction des patronymes dans les familles royales en Angleterre [19], où le nom de famille se transmettait uniquement de père à fils à cette période.

Dans la section suivante, nous présenterons quelques aspects de la théorie du processus de branchement du modèle de Galton-Watson en temps discret et son lien avec la propagation d'une épidémie dans une population.

3.1 Processus de branchement de Galton-Watson

Le processus de branchement de Galton-Watson est un modèle stochastique simple utilisé pour l'étude de la propagation des espèces au fil des générations. Elle repose sur l'hypothèse que les individus de la population donnent naissance à un certain nombre de progénitures (enfants) indépendamment les uns des autres et tous avec la même distribution [20].

Plus précisément, le modèle peut être décrit comme suit :

- Le processus débute avec un seul individu, souvent appelé génération 0 de la population.
- L'unique individu de la génération 0 vit exactement une unité de temps, donne naissance à un nombre aléatoire $Z \in N_0$ d'enfants tel que $E(Z) \in [0, \infty]$ et disparaît. Ces enfants constituent la génération 1 de la population.
- Chacun des individus de cette première génération donne également naissance à des enfants qui lui sont propres, tous avec la même distribution que Z et indépendamment

Chapitre 3. Processus de branchement (Branching process)

de tous les autres individus de la population. Ceux-ci constituent la deuxième génération de la population.

- Cette génération donne à son tour naissance à une troisième génération d'individus selon les mêmes règles que la génération précédente et le processus se poursuit ainsi de suite.
- Par conséquent, le nombre d'enfants de chaque individu Z est une variable aléatoire indépendante et distribuée de façon identique.

La figure 3.1 illustre un exemple de déroulement d'un processus de branchement de Galton Watson pour une population jusqu'à la troisième génération.

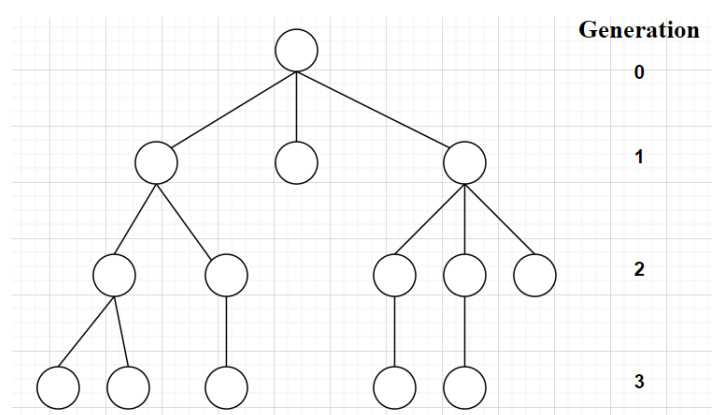


FIG. 3.1 : Processus de branchement de Galton Watson pour $n=3$

Soit X_n le nombre total d'individus nés au temps $n \in \mathbb{N}$, X_n s'interprète comme la taille de la génération au temps n .

Ainsi le processus de branchement est constitué de l'ensemble des : $\{X_0, X_1, \dots, X_n\}$ généré au temps 0 jusqu'au temps n .

L'état du processus de branchement au temps n est X_n , X_n ne peut prendre que des valeurs strictement supérieures à 0 satisfaisant toujours la relation de récurrence suivante :

$$X_n = \sum_{i=1}^{X_{n-1}} Z_{n,i} \quad (3.1)$$

Où

- La valeur X_0 toujours égale à 1 .

Chapitre 3. Processus de branchement (Branching process)

- $Z_{n,i}$ est la variable aléatoire définissant le nombre d'enfants du i^{eme} individu de la génération $n - 1$ à la génération n

Comme le processus de naissance d'enfants est indépendant les uns des autres, les individus X_n constituent ainsi une somme de X_{n-1} variables aléatoires indépendante.

Nous pouvons donc écrire sa fonction de génération [21] par la relation suivante :

$$G_{X_n}(s) = G_{X_{n-1}}(G_Z(s)) \quad (3.2)$$

Où $(G_Z(s))$ est la fonction de densité de probabilité [22] de la distribution des enfants du processus de Galton-Watson.

En appliquant ce résultat à plusieurs reprises, nous obtenons :

$$G_{X_n}(s) = G_Z(G_Z(\dots(G_Z(s))\dots))$$

$$G_{X_n}(s) = G_Z^{(n)}(s)$$

Ainsi, La fonction génératrice pour le nombre total d'individus dans la génération n est l'itération n fois de la fonction génératrice des probabilités de la distribution d'enfants.

3.2 Probabilité d'extinction

Dans cette section, nous essayons de trouver la probabilité que le processus de Galton-Watson s'éteigne ; Soit S_n cette probabilité.

Posons :

$$S_n = \text{probabilite}(X_n = 0)$$

Ainsi S_n s'écrit de façon suivante :

$$S_n = \sum_{k=0}^{\infty} P(X_n = 0 | X_1 = k) P(X_1 = k)$$

$$S_n = \sum_{k=0}^{\infty} S_{n-1}^k p_k$$

$$S_n = G_Z(S_{n-1})$$

Ainsi la valeur limite de S_n lorsque $n \rightarrow +\infty$, S_{∞} satisfait l'équation du point fixe [23] :

$$S_\infty = G_Z(S_\infty)$$

La résolution de l'équation précédente revient à rechercher les intersections entre les courbes des équations $y = s$ et $y = G_z(s)$.

Ainsi la probabilité d'extinction éventuelle du processus de Galton-Watson est donnée par la plus petite racine non négative α de l'équation $s = G_z(s)$. Et cette probabilité sera comme l'indique la figure 3.2 :

- S_n égale à 1 si le nombre moyen de descendant(R_0) est inférieur à 1
- S_n inférieur à 1 donc égale à α si le nombre moyen de descendant(R_0) est supérieur à 1

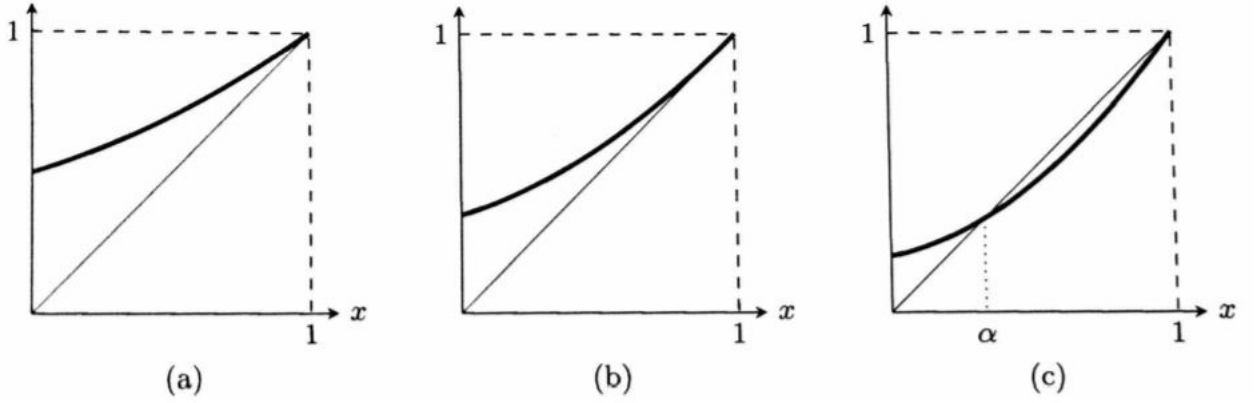


Figure by : Fabrice Grela, ENS Rennes

FIG. 3.2 : Probabilité d'extinction cas : (a) $R_0 < 1$, (b) $R_0 = 1$ et (c) $R_0 > 1$

3.3 Relation avec le modèle SIR de Kermack-McKendrick

Dans le chapitre précédent, nous avons vu que le modèle par compartiments SIR de Kermack-McKendrick était caractérisé par un taux d'infection β et un taux de guérison γ . Pour une infection donnée, les infections secondaires sont produites au taux $\frac{\beta S}{N} \approx \beta$ suivant un processus de poisson de taux β .

Si la durée de l'infection est fixée à $\frac{1}{\gamma}$, le nombre d'infections secondaires serait une distribution de Poisson avec une moyenne $\frac{\beta}{\gamma} = R_0$ de fonction génératrice [20] :

$$G_z(s) = \frac{1}{1 + R_0(1-s)}$$

Chapitre 3. Processus de branchement (Branching process)

La résolution de $G_Z(s) = s$ donne une probabilité d'extinction telle que $s = 1$ ou $s = 1/R_0$ dépendamment de la valeur R_0 .

Si $R_0 < 1$, $s = 1$ est la solution ; Et par contre si $R_0 > 1$ la solution est $s = 1/R_0$.

Ainsi commençant par un infectieux dans un modèle par compartiments SIR de Kermack-McKendrick :

- Si $R_0 < 1$, le processus épidémique s'éteint avec probabilité 1.
- Si $R_0 > 1$, le processus épidémique s'éteint avec probabilité $\frac{1}{R_0}$.
- Si $R_0 = 1$, on ne peut rien dire, l'un des deux cas précédents peut se produire.

Chapitre 4

Modèle de Reed–Frost

Les données réelles des réseaux de contacts humains au cours des dernières années ont fourni de solides arguments en faveur de l'utilisation des réseaux dans la modélisation de l'épidémiologie des maladies, dans cette partie nous nous intéressons à l'étude de la modélisation des épidémies dans des graphes qui contrairement au modèle par compartiments considère chaque individu individuellement d'où sa qualification de modèle multi-agents.

4.1 Graphe aléatoire

Un graphe G est constitué de deux ensembles d'informations : un ensemble de n sommets (noeuds) $S = (s_1 \dots s_N)$, et un ensemble de m arêtes, $A = (a_1 \dots a_m)$, entre des paires de sommets.

Chaque arête peut être écrite comme une paire de sommets, $a_k = (s_i, s_j)$. Nous nous intéressons dans ce mémoire qu'aux graphes non orientés, c'est-à-dire des graphes qui n'ont pas de boucles et ne comprennent pas plus d'une arête non dirigée entre une paire de sommets. Il existe plusieurs modèles de graphes aléatoires bien connus dans la littérature donc les principaux sont :

4.1.1 Graphe aléatoire Erdos-Rényi

C'est le premier et le plus connu des modèles de graphe aléatoire publié en 1956 par deux chercheurs Paul Erdős et Alfred Rényi [24].

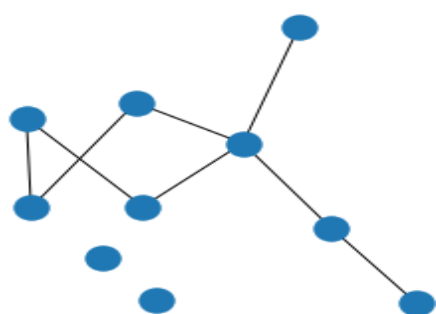
Ce modèle a le moins de structures possibles, c'est un graphe constitué de n individus ou sommets et m arêtes qui suppose que chaque paire d'individus est connectée les uns aux autres, indépendamment, avec une probabilité p égale $\frac{\mu}{n-1}$.

C'est-à-dire que l'existence de chaque arête est indépendante de celle des autres et que chaque arête possède une probabilité p d'exister. Ce modèle est caractérisé par un seul

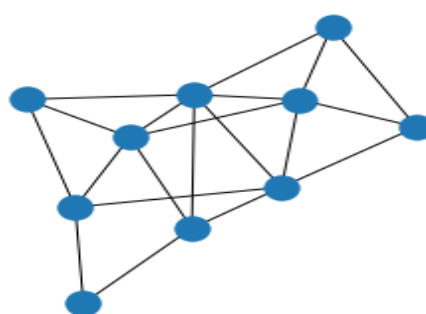
paramètre μ étant le degré moyen des individus, et donc dans ce graphe, un individu a au plus $n - 1$ connexions possibles, chacune étant présente avec une probabilité $\frac{\mu}{n-1}$, donc le nombre de voisins de tout individu a une distribution binomiale : $P \approx \text{Bin}(n - 1, \frac{\mu}{n})$, la moyenne exacte est en fait $\mu(1 - \frac{1}{n})$.

Lorsque n tend ∞ , cette distribution tend vers la distribution de Poisson de moyenne μ , donc $P \approx \text{poisson}(\mu)$.

Les figures 4.1a et 4.1b sont des illustrations possibles de deux graphes aléatoires Erdős-Rényi possédant 10 sommets pour deux valeurs de probabilité respective p égale 0.2 et 0.5



(a) Graphe aléatoire Erdős-Rényi $p = 0.2$



(b) Graphe aléatoire Erdős-Rényi $p = 0.5$

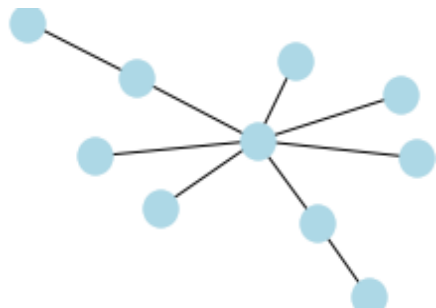
4.1.2 Modèle d'attachement préférentiel

Le modèle d'attachement préférentiel est le second modèle de graphe aléatoire proposé par Barabási et Albert en 1999 pour expliquer les phénomènes de regroupements [5].

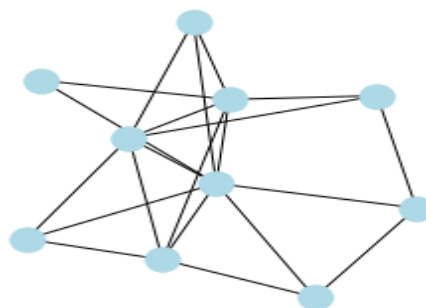
L'idée décrite par ces auteurs est que dans un graphe dynamique, un sommet ne se connecte pas aux autres sommets de façon aléatoire, mais avec une probabilité plus forte vers les sommets les plus connectés, dont plus un sommet présente une connectivité importante par rapport aux autres, plus il attirera de nouveaux sommets. Le principe de construction de ce graphe est le suivant :

Le graphe commence par un premier graphe complètement connecté de m_0 sommets. De nouveaux sommets sont ajoutés au graphe, chaque sommet est connecté à $m \leq m_0$ sommets existants, avec une probabilité proportionnelle au nombre de liens que ces sommets existants ont déjà. Les sommets fortement liés appelés hubs ont tendance à accumuler rapidement encore plus de liens. Tandis que les sommets avec seulement quelques liens sont peu susceptibles d'être choisi comme destination pour un nouveau lien. Les nouveaux sommets ont donc une préférence pour s'attacher aux sommets déjà fortement liés.

Les figures 4.2a et 4.2b sont des illustrations possibles de deux graphes aléatoires de Barabási-Albert comportant 10 sommets pour deux valeurs de m_0 égale 1 et 3.



(a) Graphe aléatoire Barabási-Albert $m_0 = 1$



(b) Graphe aléatoire Barabási-Albert $m_0 = 3$

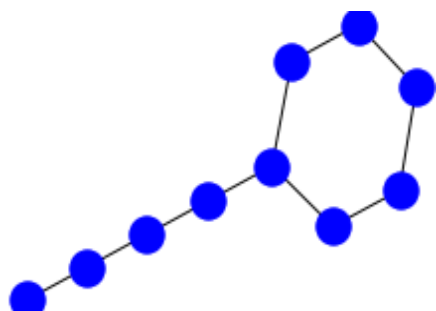
4.1.3 Small-world

Bien qu'il en existe encore d'autres modèles de graphe aléatoire, le modèle small-world est le dernier modèle que nous présenterons.

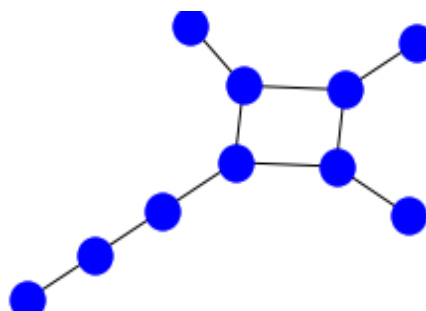
Ce modèle de graphe a été proposé en 1998 par Watts et Strogatz [6]. Dans ce modèle, tous les sommets sont étiquetés et placés sur une ligne, qui est constituée d'un cercle, le modèle à deux paramètres, k et p , le premier étant un entier positif et le second un nombre compris entre 0 et 1 typiquement proche de 0.

Chaque sommet est d'abord connecté à ses k voisins les plus proches des deux côtés ensuite, chaque arête, indépendamment et avec une probabilité p , connecté l'un de ses sommets.

Les figures 4.3a et 4.3b sont des illustrations possibles de deux graphes aléatoires Small-world comportant 10 sommets avec paramètres k et p .



(a) Small-world : $n = 10$, $k = 2$, $p = 0.2$



(b) Small-world : $n = 10$, $k = 2$, $p = 0.5$

Dans les sections suivantes, nous nous intéresserons au réseau aléatoire Erdős-Rényi qui est un modèle où tous les noeuds jouent plus ou moins un rôle analogue et dont la plupart

des noeuds possèdent un nombre de connexions proche du degré moyen. Il est donc qualifié de modèle homogène.

4.2 Composante connexe géante

Un graphe G non orienté est constitué d'un ensemble de sommets S et d'un ensemble d'arêtes A noté $G = (S, A)$.

Le graphe G sera dit connexe si et seulement si quels que soient les sommets u et v de S , il existe une chaîne reliant u à v [25].

On définit donc une composante connexe d'un graphe G comme un sous ensemble $S' = \{s_1, \dots, s_j\} \in S$ tel que pour tout $s_i, s_k \in S'$, il existe un chemin dans G de s_i à s_k .

Comme montré par les auteurs dans [25] tout graphe peut être décomposé en un unique ensemble de composantes connexes et le nombre de composants connexe d'un graphe est supérieur ou égal à $n - |E|$.

Soit C une composante connexe de G , la taille de C est égale à la somme totale des sommets constituant la composante. Ainsi la composante C sera dite géante si sa taille de $\frac{|C|}{n}$ ne tend pas vers 0 lorsque n devient grand.

En pratique, pour le cas particulier d'un graphe aléatoire d'Erdős–Rényi on définit une fonction de probabilité p telle que $p = p(n)$, et pour chaque valeur de n , on génère un graphe de même type (Erdős–Rényi) selon la loi $p(n)$, et on s'intéresse à chaque fois au rapport $\frac{|C|}{n}$. S'il est > 0 avec probabilité de 1, cela traduit l'existence d'une composante géante celle-ci existera forcément si le produit de $np > 1$ [25].

Notons également qu'un sommet isolé forme à lui seul une composante connexe.

Taille de la composante géante : cas d'un graphe Erdős–Rényi :

Considérons un graphe aléatoire d'Erdős–Rényi $G(n, p)$ et supposons qu'il admet une composante géante dont l'ordre est $O(n)$.

Définissons u comme la fréquence des sommets n'appartenant pas à la composante géante, c'est-à-dire que u est la probabilité qu'un sommet choisi au hasard n'appartienne pas à la composante géante.

Calculons cette probabilité de la manière suivante :[4]

Choisissons n'importe quel autre sommet il peut être soit dans la composante géante ou non. Pour que le nœud d'origine n'appartienne pas à la composante géante, ces deux éléments ne doivent pas être connectés, probabilité $1 - p$ ou être connectés, mais ce dernier sommet

Chapitre 4. Modèle de Reed–Frost

n'est pas dans la composante géante, probabilité pu . Il y a $n - 1$ sommets à vérifier, donc

$$u = (1 - p + pu)^{n-1} \quad (4.1)$$

Soit $\lambda > 0$, le degré moyen du graphe $G(n, p)$, donc :

$$p = \frac{\lambda}{n}$$

En remplaçant p par sa valeur dans l'équation 4.1 on obtient :

$$u = \left(1 - \frac{\lambda}{n}(1 - u)\right)^{n-1}$$

Introduisons la fonction logarithme des deux côtés de l'équation précédente :

$$\log u = (n - 1) \log \left(1 - \frac{\lambda}{n}(1 - u)\right)$$

Nous savons que $\log(1 + x) \approx x \ \forall x$ petit, ainsi :

$$\log u \approx -\frac{\lambda(n-1)}{n}(1 - u)$$

$$\log u \approx -\lambda(1 - u)$$

$$u = e^{-\lambda(1-u)}$$

Ainsi nous déduisons que la fréquence des sommets dans la composante géante $v = 1 - u$ est :

$$1 - v = e^{-\lambda v} \quad (4.2)$$

L'équation 4.2 a toujours pour solution $v = 0$, mais cette solution n'est pas unique pour toutes les valeurs possibles de λ .

La figure 4.4 permet d'analyser graphiquement le nombre de solutions de l'équation $1 - v = e^{-\lambda v}$. Pour cela considérons les deux courbes des fonctions définies par $f(v) = 1 - e^{-\lambda v}$ et $g(v) = v$, leurs intersections donnent les solutions (racines) de l'équation (4.2).

Si $v \geq 0$ alors il est possible d'avoir une solution triviale $v = 0$ dans les cas $\lambda < 1$, $\lambda = 1$ et deux solutions si $\lambda > 1$ comme l'indique l'analyse graphique ci-dessous.

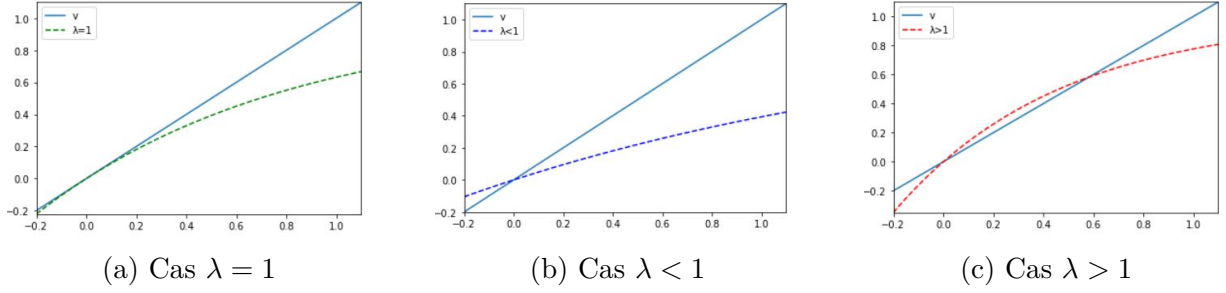


FIG. 4.4 : Analyse graphique du nombre de solutions de l'équation $1 - v = e^{-\lambda v}$

En résumé :

- Si $\lambda < 1$, pas de composante géante dans le graphe.
- Si $\lambda > 1$, une composante géante existe et contient une fréquence v de sommets de $G(n, p)$.

Ainsi nous établissons une relation entre l'existence d'une composante géante dans un graphe Erdős–Rényi et le processus de branchement de Galton-Watson. La présence ou non d'une composante géante dans un graphe Erdős–Rényi de degré moyen λ peut être approximée à un processus de Galton-Watson de moyenne λ , la fréquence de sommets constituant la composante géante est donc égale à la probabilité de survie dans un processus de Galton-Watson.

4.3 Modèles épidémiques sur les réseaux : Reed-Frost

Dans cette section, nous décrivons le modèle épidémique de Reed-Frost. C'est un modèle épidémiologique *SIR* simple, formulé par Lowell Reed et Wade Frost en 1928 permettant de décrire l'évolution d'une maladie contagieuse dans une population fermée au fil du temps [26].

Par opposition aux modèles par compartiments qui sont des modèles déterministes, c'est-à-dire des modèles dans lequel une entrée donnée introduite dans le modèle produit toujours la même sortie. Ainsi un modèle déterministe est un modèle où la sortie est entièrement déterminée par les paramètres et les conditions initiales.

Reed-Frost est un modèle stochastique, c'est-à-dire un modèle dans lequel il existe un élément aléatoire tel que pour une entrée donnée au modèle, le résultat prend une gamme de valeurs de sorties possibles. Ainsi cela traduit donc que le résultat des opérations du modèle n'est pas uniquement déterminé par l'entrée donnée.

Notons également le modèle Reed-Frost est un modèle à temps discret, c'est-à-dire qu'il introduit une relation entre des variables d'entrée et des variables de sortie, dans lequel

ces variables ne peuvent avoir qu'un nombre fini de valeurs, par opposition à un système continu tel que les modèles compartimentaux, dans lequel entre deux valeurs de variables, d'entrée ou de sortie, on peut toujours supposer une valeur intermédiaire.

De nombreux facteurs peuvent influencer la propagation d'une maladie infectieuse, par conséquent, quelques hypothèses de base sont faites sur le modèle Reed-Frost afin de créer un modèle cohérent [27].

Le modèle Reed-Frost repose donc sur 5 hypothèses [27] :

- L'infection ne peut se propager que par un contact adéquat avec des personnes infectées.
- Les personnes sensibles deviennent infectées après un contact adéquat avec une personne infectée au cours d'une période donnée. Ils ne seront alors infectieux que pour la période suivante, avant de devenir totalement immunisés.
- L'infection est introduite dans une population fermée, c'est-à-dire que la population totale est fixée sans que personne n'entre ou ne quitte le groupe d'individus.
- Les individus ont une probabilité fixe de contact adéquat avec tout autre individu au cours d'une période donnée.
- Les conditions ci-dessus restent constantes pendant toute la durée de l'épidémie.

4.3.1 Description du modèle

Dans la section précédente, nous avons décrit plusieurs modèles de graphes aléatoires. Dans la section actuelle, nous supposons qu'un graphe nous est donné et appelons-le G , nous décrivons le modèle épidémique en temps discret Reed-Frost sur le graphe non orienté statique donné.

Le modèle Reed-Frost est un modèle SIR , où les individus sont initialement susceptibles (S), ensuite s'ils sont infectés, ils deviennent infectieux (I) et après un certain temps se rétablissent et deviennent immunisés (R).

Supposons que la probabilité de transmission entre un individu susceptible et individu contaminé est donnée par p , l'application du modèle sur le graphe G se décrit comme suit : Initialement, à la génération $n = 0$, un cas index choisi au hasard est infectieux et le reste de la population/du réseau est susceptible. Les individus qui sont infectieux dans la génération n , infectent chaque voisin susceptible dans G , indépendamment avec une probabilité p , puis se rétablissent, notons que lors du déroulement du processus, rien ne se passe avec les voisins immunisés ou infectieux. Ceux qui sont infectés par au moins un infectieux deviennent infectieux à la génération $n + 1$, tandis que les infectieux de la génération n deviennent immunisés. Le processus de propagation de l'épidémie se poursuit ainsi jusqu'à

Chapitre 4. Modèle de Reed–Frost

la première génération où aucune nouvelle infection ne survient.

La figure 4.5 donne l’illustration d’une exécution du modèle Reed-Frost sur un réseau d’Erdős–Rényi donc la population totale est $N = 10000$ et le degré moyen $\mu_D = 8$ et la probabilité de transmission d’une infection entre un susceptible et un infectieux est $p = 0.3$.

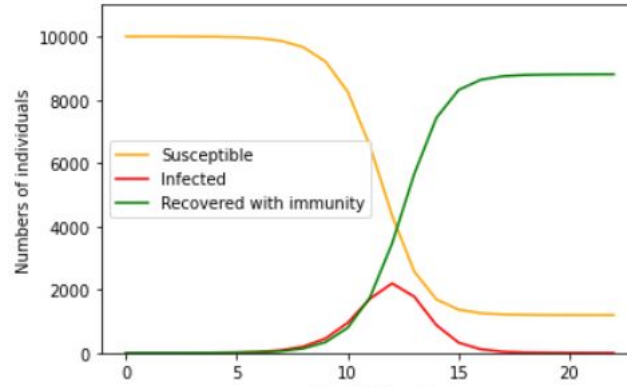


FIG. 4.5 : Une exécution possible du modèle Reed-Frost sur un graphe Erdős–Rényi de $N = 10000$ noeuds, de degré moyen $\mu_D = 8$ et une probabilité de transmission $p = 0.3$.

Notons que le modèle Reed-Frost étant un modèle stochastique comme dit plus haut, celui-ci doit être exécuté plusieurs fois pour générer des tendances de comportement, comme l’illustre la figure suivante :

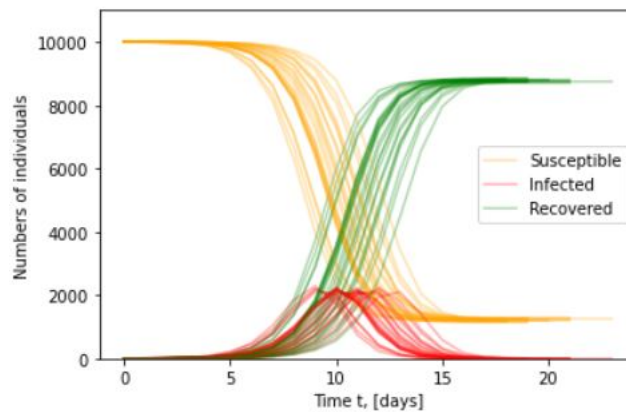


FIG. 4.6 : 30 exécutions du modèle Reed-Frost sur un graphe Erdős–Rényi de $N = 10000$ noeuds, de degré moyen $\mu_D = 8$ et une probabilité de transmission $p = 0.3$.

4.3.2 Probabilité de transmission

Comme dit précédemment, le modèle Reed-Frost est un modèle *SIR* par conséquent caractérisé par deux paramètres que sont le taux de contamination β et le taux de guérison γ , défini sur un graphe G , la probabilité de transmission pour ce modèle est donnée par la relation suivante [28] :

$$p = \frac{\beta}{\beta + \gamma} \quad (4.3)$$

4.3.3 Taux de reproduction de base R_0

Le taux de reproduction de base comme nous l'avons vu dans les chapitres précédents est un paramètre critique en épidémiologie, car permet de définir les conditions de seuil pour une éventualité de déclenchement d'une épidémie étant donnée une maladie son taux de contamination β et son taux de récupération γ , R_0 définit donc les conditions permettant de déterminer si une maladie peut ou non envahir une population. Il est défini comme le nombre de cas secondaires générés par un individu infectieux dans une grande population sensible.

Dans un modèle Reed-Frost sur réseau aléatoire plus précisément sur un graphe Erdos-Rényi de distribution de degré de moyenne μ_D , R_0 se définit comme le produit de la probabilité de transmission par le degré moyen du graphe :

$$R_0 = \frac{\beta}{\beta + \gamma} * \mu_D \quad (4.4)$$

Cependant, dans les modèles plus complexes, il n'est pas toujours évident de le définir.

4.3.4 Taille finale de l'épidémie

La taille finale d'une épidémie est la proportion de la population initialement susceptible qui a été touchée par la maladie lorsque l'épidémie se termine.

Notons que les individus infectés au déclenchement d'une l'épidémie (au temps $t = 0$), c'est-à-dire les cas d'index ne sont généralement pas comptés dans la taille finale de l'épidémie.

L'approximation du modèle Reed-Frost de taux de reproduction de base R_0 par un processus de branchement de Galton-Watson de moyenne R_0 permet d'exprimer la taille de l'épidémie.

L'approche est la suivante [27] :

Chapitre 4. Modèle de Reed–Frost

Notons τ , la proportion qui n'échappe pas à l'épidémie et donc $1 - \tau$ la proportion inverse, c'est-à-dire la proportion initialement susceptible qui échappe à l'épidémie à la fin de celle-ci.

Supposons que $\frac{I_0}{n} \approx \mu$ lorsque n tend vers ∞ donc μ mesure la proportion asymptotique d'individus initialement infectée présente dans la population et considérons la probabilité qu'un individu susceptible donné échappe à l'infection.

On sait que le nombre moyen d'individus infecté par une personne infectée est R_0 , sachant qu'on a approximativement n susceptibles au départ alors la probabilité qu'un individu soit contaminé est R_0/n son inverse $1 - R_0/n$ donne la probabilité que l'individu échappe à l'infection d'un infectieux.

Pour échapper à toute l'épidémie, elle doit échapper à l'infection de tous les $n\mu$ individus initialement infectés et de tous les $n\tau$ individus finalement infectés dans la population initialement susceptible. Asymptotiquement (n tend vers ∞), la proportion qui échappe à l'épidémie devrait être égale à la probabilité qu'un susceptible donné échappe à l'infection. Ainsi, la taille finale est donnée par la solution de l'équation :

$$1 - \tau = e^{-R_0(\tau + \mu)}$$

Dans la plupart des situations, le nombre d'infectieux initial I_0 est très petit devant la population totale n donc on peut écrire l'approximation $\frac{I_0}{n} \approx \mu \approx 0$, donc on a donc finalement :

$$1 - \tau = e^{-R_0\tau} \tag{4.5}$$

Cette équation ne peut pas être résolue explicitement, mais des solutions numériques sont facilement obtenues pour des valeurs données de R_0 .

La figure 4.7, décrit l'évolution de la taille de l'épidémie en fonction de R_0 .

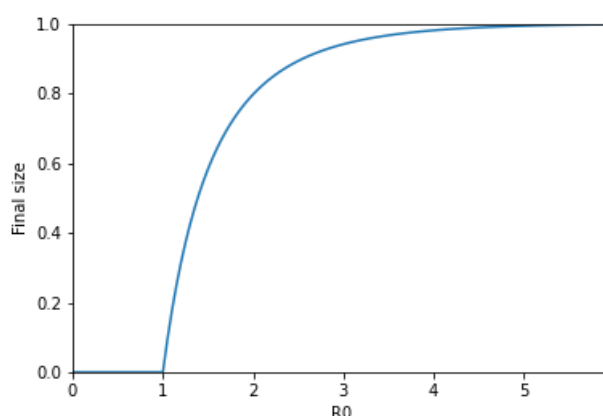


FIG. 4.7 : Courbe τ en fonction de R_0

4.4 Modèle de prévention : principe de percolation

4.4.1 Principe de percolation

La notion de percolation a été longtemps étudiée en physique et en mathématique, c'est un processus physique qui décrit pour un système, une transition d'un état vers un autre [29].

Dans cette section, nous établirons une équivalence entre l'application de cette théorie dans les graphes et le processus de vaccination décrit dans le chapitre de modèle de Kermack-McKendrick.

Rappelons toutes fois que, les hypothèses formulées pour cette technique dans ce mémoire étaient les suivantes : Les individus sont vaccinés avant l'apparition de la maladie. Une fois un individu vacciné, il acquiert une immunité à vie donc définitive face à la maladie, et de ce fait ne peut plus être infecté et ni propager la maladie.

Soit $G = (S, A)$ un graphe, donc S est l'ensemble des sommets et A l'ensemble des arêtes des graphes, en mathématique la percolation du graphe G peut se mettre en oeuvre par deux méthodes qui sont [29] : La percolation par lien qui est l'étude des phénomènes de percolation sur les arêtes des graphes et la percolation par site quant à elle, est l'étude des phénomènes de percolation sur les sommets des graphes.

Nous nous intéressons ici à la percolation par site.

Notre principe consistera à choisir aléatoirement et uniformément une proportion p de sommets à supprimer dans le graphe, sans tenir compte des propriétés des sommets tel que son rôle, le degré et d'en étudier les effets sur le taux de reproduction du nouveau graphe

G_r obtenue constitué de la proportion $1 - p$ de sommets restants.

4.4.2 Effet de percolation sur taux de reproduction R_0

Soit G un graphe d'Erdős–Rényi, supposons qu'une proportion p de sommets choisit aléatoirement et uniformément est supprimée. Les sommets étant uniformément choisis pour la suppression, chaque suppression de sommet est donc un essai de Bernoulli avec une probabilité de succès p [28].

En supposant que le graphe G a une distribution de degré P_k c'est-à-dire que la probabilité qu'un sommet ait un degré k est p_k .

Sa fonction génératrice des probabilités est la série entière [28] :

$$G(x) = \sum_{k=0}^{\infty} x^k p_k$$

La suppression de sommet dans le graphe est un essai de Bernoulli avec une probabilité de succès p et une probabilité d'échec $1 - p$, ainsi la fonction génératrice de probabilité est :

$$G_r(x) = G(p + x(1 - p))$$

Et le degré moyen du graphe percolé :

$$\mu_{D_r} = E'(1) = \frac{d}{dx} G_r(1)$$

$$\mu_{D_r} = (1 - p)G'(1) = (1 - p)\mu_D$$

Où μ_D est le degré moyen du graphe avant la percolation.

Ainsi, le taux de reproduction de base R_0 en fonction de p est :

$$R_0(p) = \frac{\beta}{\beta + \gamma} \mu_{D_r}$$

En remplaçant μ_{D_r} par sa valeur obtenue en (4.6)

$$R_0(p) = (1-p)^{\frac{\beta}{\beta+\gamma}} \mu_D$$

Finalement :

$$R_0(p) = (1-p)R_0 \quad (4.6)$$

L'équation 4.6 montre que $R_0(p)$ une fonction décroissante de p , en conclusion la percolation de sommets dans un graphe de façon uniformément aléatoire réduit le R_0 initial.

La figure 4.8 décrit l'évolution de $R_0(p)$ en fonction de p pour une valeur initiale de $R_0 = 2$.

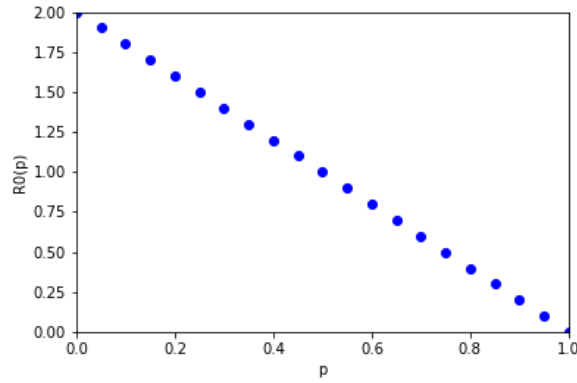


FIG. 4.8 : Variation de R_0 en fonction de p

Ainsi grâce à la relation de l'équation 4.6, nous pouvons écrire la couverture vaccinale critique pour ce cas.

La couverture vaccinale critique décrivant la proportion de sommets à supprimer pour stopper la propagation de l'épidémie dans le réseau. Nous avons vu précédemment que l'épidémie arrêta de se propager lorsque son taux de reproduction de base R_0 devenait plus petit que 1.

Donc :

$$R_0(p) < 1$$

$\Rightarrow$

$$(1-p)R_0 < 1$$

La relation précédente implique de trouver la première valeur de p notée p_c telle que :

$$p_c > 1 - \frac{1}{R_0}$$

Chapitre 5

Comparaison Kermack-McKendrick vs Reed-Frost

Dans cette section nous comparons le modèle *SIR* de Kermack-McKendrick par compartiments déterministe à celui du modèle de Reed-Frost stochastique sur graphe Erdős-Rényi ayant le même taux de reproduction de base R_0 .

Données : comme données, considérons une épidémie caractérisée par un taux de contamination de 0.5 et une durée d'infection de 5 jours. Les paramètres pour les modèles sont les suivants :

- Pour le modèle par compartiments *SIR*, nous avons $\beta = 0.5$, $\gamma = 0.2$ et $R_0 = 2.5$.
- Pour le modèle Reed-Frost stochastique sur graphe, nous avons comme probabilité de transmission $p = \frac{\beta}{\beta+\gamma} = 0.7142$, un graphe Erdős-Rényi de degré moyen $\mu_D = 3.5$ donc un $R_0 = 2.5$, égale à celui par compartiment.
- Nos deux modèles sont initialisés de la manière suivante :
La population totale est $N = 10000$ individus (sommets), au départ un seul individu (sommet) est infecté $I(0) = 1$ et aucun individu (sommet) n'est guéri $R(0) = 0$.

Les figures 5.1a, 5.1b décrivent respectivement l'évolution de $S(t)$, $I(t)$ et $R(t)$ au fil du temps, pour le modèle par compartiments et pour le modèle Reed-Frost dans le graphe Erdős-Rényi.

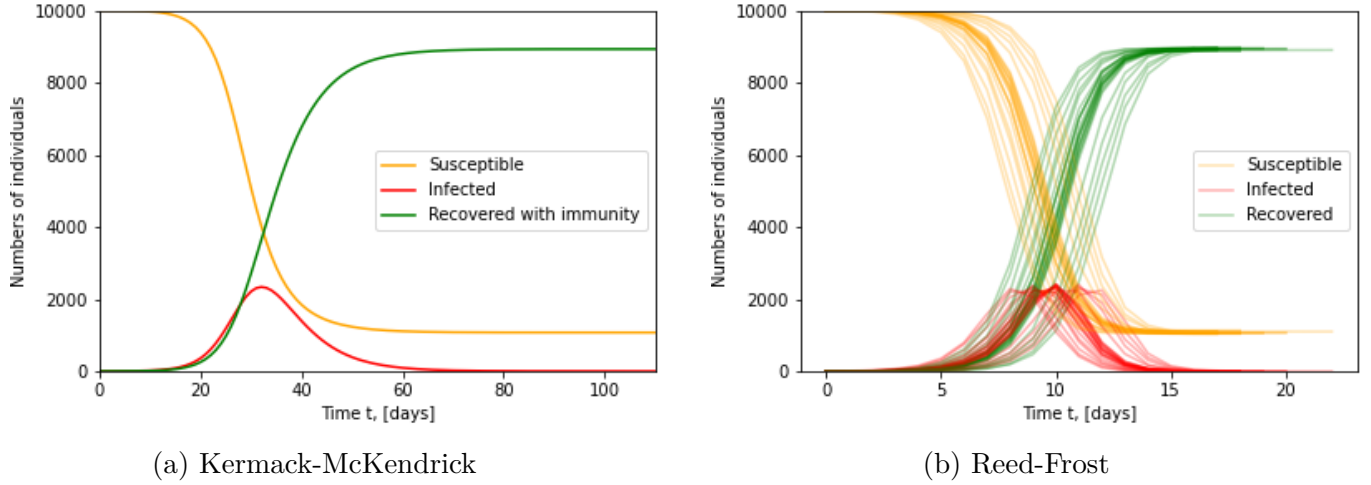


FIG. 5.1 : Évolution de $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles

Bien qu'ayant des temps de propagation différents, l'évolution des paramètres $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles semblent similaires même s'il faut noter que les valeurs finales sont très légèrement différentes.

Le décalage entre les différentes exécutions dans la figure 5.1b est dû à la stochasticité du modèle Reed-Frost expliquée plus haut.

Afin de comparer l'évolution $R(t)$ des deux modèles par compartiments de Kermack-McKendrick et celui de Reed-Frost, un ajustement de l'échelle du temps de propagation est nécessaire. Le modèle de Reed-Frost est un modèle discret, les individus infectés se rétablissent après une unité de temps, hors le temps moyen de rétablissement $\frac{1}{\gamma}$ vaut 5 pour l'épidémie décrit en données.

Ainsi nous ajustons l'échelle du temps du modèle Reed-Frost en le multipliant par le facteur 5.

La figure 5.2 est une comparaison de l'évolution de $R(t)$ dans les deux modèles Reed-Frost et Kermack-McKendrick après ajustement du temps.

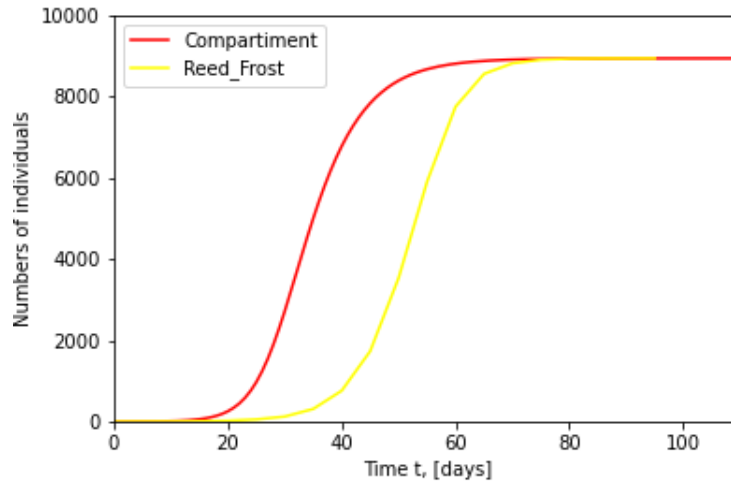


FIG. 5.2 : Comparaison $R(t)$ dans les modèles épidémiques de Reed-Frost et Kermack-McKendrick

Nous remarquons que la taille finale de l'épidémie R_∞ dans les deux modèles est sensiblement la même.

Les équations définissant la taille finale de l'épidémie (proportion de personnes touchées à la fin d'une épidémie) pour les deux modèles SIR par compartiments et pour Reed-Frost sont identiques, elles ne dépendent que du taux de reproduction de base R_0 . Pour rappel ces équations établies aux chapitres 2 et 3 sont les suivantes :

Pour le modèle par compartiments de Kermack-McKendrick :

$$1 - R_\infty = e^{-R_0 R_\infty}$$

Pour le modèle multi-agents de Reed-Frost :

$$1 - \tau = e^{-R_0 \tau}$$

La figure 5.3 donne la solution de ces deux équations obtenue numériquement par simulation pour $R_0 = 2.5$. Les proportions théoriques des personnes touchées à la fin de l'épidémie dans les deux modèles sont les mêmes 0.8926 donc environ 89.26%.

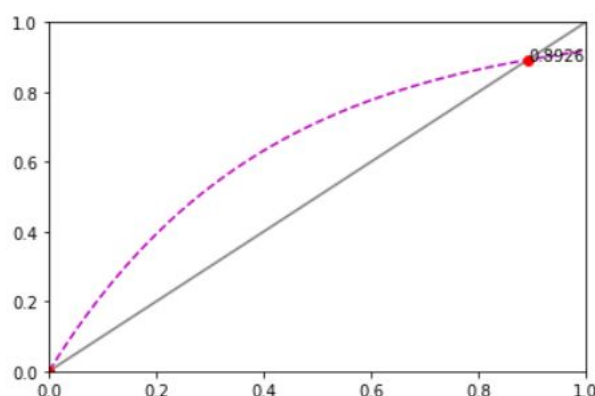


FIG. 5.3 : Résolution numérique de l'équation de proportion touchée dans le modèle multi-agents de Reed-Frost et celui par compartiments de Kermack-McKendrick par méthode de point fixe.

En Conclusion, grâce à la figure 5.2 on observe que la proportion des personnes touchée à la fin de l'épidémie obtenue par simulation de l'épidémie est approximativement la même dans les deux modèles ; De plus cette proportion est égale à celle fournie théoriquement 89.26%.

Supposons qu'un pourcentage de $p = 25\%$ de la population initiale est vacciné et donc immunisé contre cette maladie

Les figures 5.4a, 5.4b décrivent respectivement l'évolution de $S(t)$, $I(t)$ et $R(t)$ au fil du temps, pour le modèle par compartiments de Kermack-McKendrick et pour le modèle de Reed-Frost dans le graphe Erdős-Rényi percolé uniformément et aléatoirement de 25% de ses sommets.

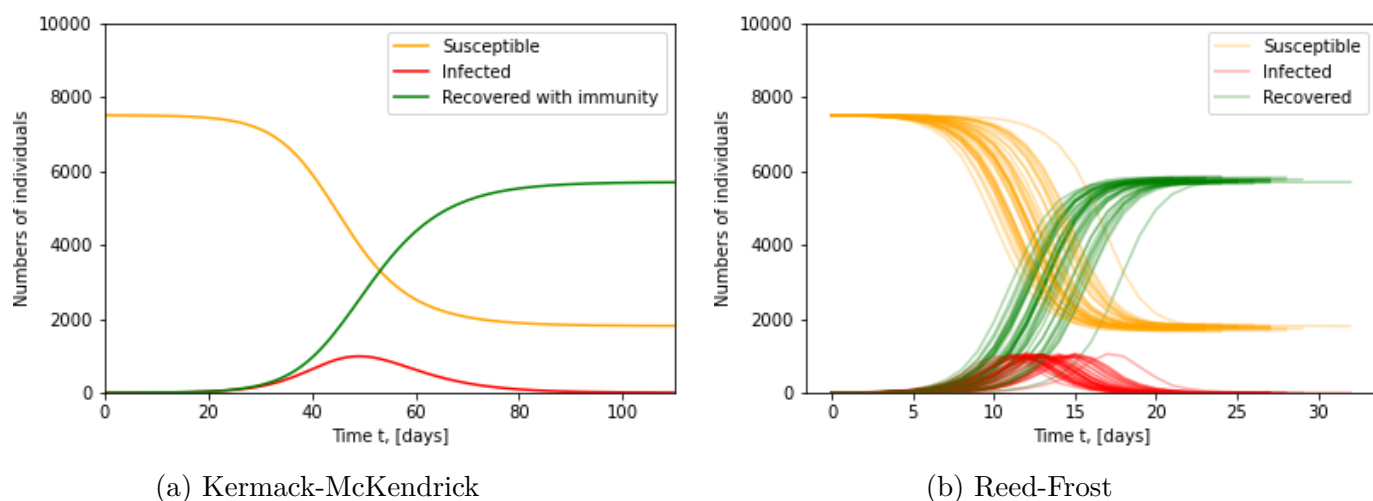


FIG. 5.4 : Évolution de $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles

Comme dans les cas précédents, bien qu'ayant des temps de propagation différents, l'évolution des paramètres $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles évoluent de façon similaire même s'il faut noter que les valeurs finales sont légèrement différentes.

La figure 5.5 est une comparaison de l'évolution de $R(t)$ au fil du temps dans les deux modèles avec l'ajustement de l'échelle du temps pour le modèle de Reed-Frost comme décrit précédemment.

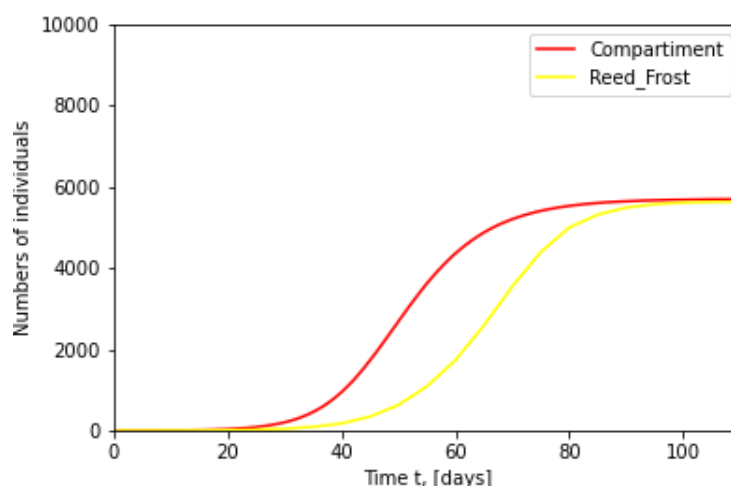


FIG. 5.5 : Comparaison $R(t)$ dans les modèles épidémiques de Reed-Frost et Kermack-McKendrick

Chapitre 5. Comparaison Kermack-McKendrick vs Reed-Frost

Nous remarquons que la taille de l'épidémie R_∞ dans les deux modèles est la même.

Comme établi dans les chapitres 2 et 3, sous les hypothèses décrites précédemment la vaccination d'un pourcentage p de la population initiale fait passer le taux de reproduction basique de R_0 à $(1 - p)R_0$, dans le cas présent de 2.5 à 1.875.

La figure 5.6 donne la solution de ces deux équations obtenue numériquement par simulation pour $R_0 = 1.875$. La proportion théorique des personnes touchée à la fin de l'épidémie dans les deux modèles est la même 0.7590 donc environ 75.9%

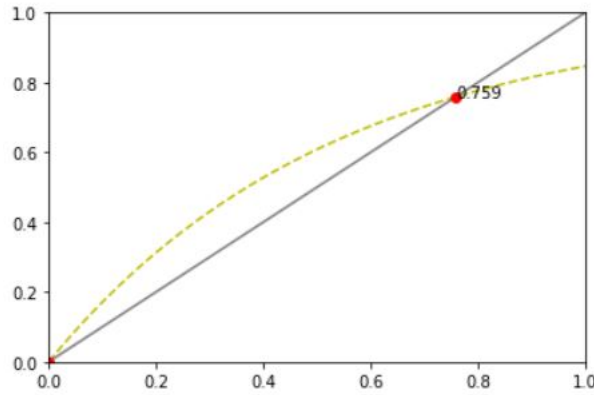


FIG. 5.6 : Résolution numérique de l'équation de proportion touchée dans le modèle multi-agents de Reed-Frost et celui par compartiments de Kermack-McKendrick par méthode de point fixe.

Ainsi on observe que la proportion des personnes touchée à la fin de l'épidémie obtenue par simulation de l'épidémie (figure 5.5) est approximativement la même dans les deux modèles ; De plus cette proportion est égale à celle fournie théoriquement.

Des résultats similaires sont obtenus en supposant qu'au temps initial le nombre de cas d'index est supérieur à 1, précisément pour $I_0 = 300$ (voir annexe A pour résultats de simulation).

En Conclusion, le modèle par compartiments de Kermack-McKendrick et celui de Reed-Frost sur un graphe aléatoire d'Erdős-Rényi prédisent la même taille d'épidémie, la mise en place d'une mesure de prévention plus précisément la vaccination prédisent également la même taille d'épidémie à la fin.

Chapitre 6

Taux de reproduction effectif : R_{eff}

Le taux de reproduction basique R_0 comme vu dans les chapitres précédents est l'un des facteurs clé dans l'étude des systèmes de modélisation des maladies infectieuses en épidémiologie, il correspond en effet au nombre moyen de personnes qu'une personne infectée va contaminer dans une population totalement susceptible.

La valeur de ce paramètre fournit une information permettant de simuler des exemples d'évolution du nombre d'individus contaminés dans une population si rien n'est fait. La figure 6.1 illustre une séquence possible de l'évolution du nombre de contaminés pour une maladie infectieuse donc le taux de reproduction de base R_0 vaut 2 et ayant débuté avec un cas d'index : p_1 .

Elle décrit la séquence suivante, le cas d'index va contaminer en moyenne R_0 personnes, donc p_2 et p_3 .

p_2 et p_3 vont à leurs tours respectivement contaminer R_0 personnes en moyenne, soit p_4 et p_7 pour p_1 et p_5 et p_6 pour p_3 , le processus de contamination se répandra ainsi de suite.

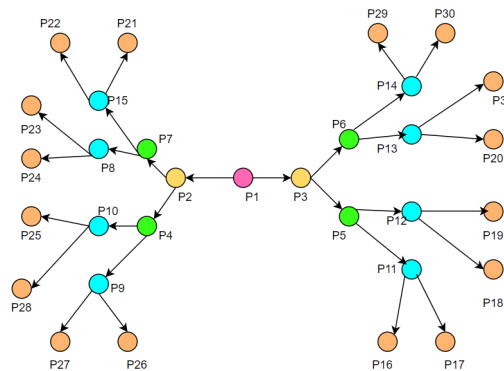


FIG. 6.1 : Un exemple de processus de contamination d'une épidémie dans une population pour $R_0 = 2$

Chapitre 6. Taux de reproduction effectif : R_{eff}

Ce paramètre permet de déterminer si oui ou non une épidémie aura lieu, d'où son intérêt immédiat en épidémiologie.

Si R_0 est supérieur à 1 chaque individu infecté va à son tour se reproduire en infectant plus d'une personne dans la population, ce qui permettra à la maladie de se développer et de se répandre dans la population, conséquence l'épidémie a lieu.

Dans le cas contraire, c'est-à-dire si R_0 est inférieur à 1, un individu infecté infectera à son tour en moyenne moins d'une personne dans la population ce qui ne permettra pas à la maladie de se répandre. Par conséquent, pas d'épidémie.

Il existe un cas neutre définissant le seuil entre épidémie ou pas d'épidémie, c'est le cas lorsque $R_0 = 1$. En plus de ce rôle de seuil, le taux de reproduction de base permet également de classer les maladies par potentiel épidémique permettant de comparer différentes maladies entre elles, il en découle que, plus le taux de reproduction d'une maladie sera élevé plus le potentiel épidémique de cette dernière sera grand comme nous le montre la figure 6.2 suivante :

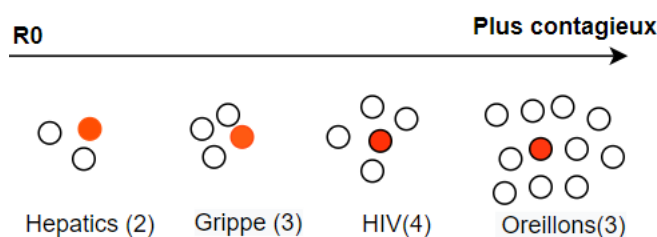


FIG. 6.2 : Classification des maladies par potentiel épidémique

De plus, il est d'une interprétation fort simple, c'est un paramètre calculé généralement au tout début de l'épidémie, il varie selon le modèle utilisé pour capturer l'épidémie.

Dans le cas d'un modèle *SIR* par compartiments comme nous l'avons vu plus haut que celui-ci est fourni par la relation suivante :

$$R_0 = \beta * \frac{1}{\gamma}$$

Où β est le taux de contamination de la maladie, ce dernier s'écrit comme produit de deux facteurs p et c , où p et c représentent respectivement la probabilité de transmission de la maladie lors d'un contact et le nombre de contacts par jours.

Ainsi le taux de reproduction de base R_0 s'écrit de la façon suivante :

$$R_0 = p * c * D \quad (6.1)$$

Avec D représentant la durée infectieuse, définit par $D = \frac{1}{\gamma}$

Cette relation 6.1 nous montre que le taux de contamination basique R_0 dépend de trois facteurs :

- La probabilité de transmettre l'infection lors d'un contact entre une personne susceptible et une personne infectée.
- La fréquence des contacts entre personnes.
- Et de la durée de la contagion après l'infection.

R_0 sera d'autant plus important que l'un de ces trois facteurs sera élevé. Par exemple, dans le cadre de la crise sanitaire liée à la covid-19 en Belgique, le R_0 au tout début de l'épidémie a été estimé à 3.79 dans leurs recherches par les membres de l'institut namurois des systèmes complexes [naXys](#) [3] et à 0.6 pendant la période de confinement total en mai 2020 par Yves Van Laethem, porte-parole interfédéral de la lutte contre la covid-19, lors de la conférence de presse quotidienne du centre de crise publiée par la RTBF dans son article du lundi 04 mai 2020 à 15h50 disponible [ici](#).

Dans le cas des modèles par graphe, où p est la probabilité de transmission de l'infection entre un sommet infecté un sommet susceptible et P_k la distribution des degrés du graphe G , le taux de reproduction basique est donnée par la relation suivante :

$$R_0 = p * \mu_D \quad (6.2)$$

Où μ_D est degré moyen du graphe G .

Tout comme dans le cas de la modélisation par compartiment, nous voyons que le modèle dépend de plusieurs facteurs, notamment :

- La probabilité de transmettre l'infection lors d'un contact entre un sommet susceptible et un sommet infecté.
- La fréquence des contacts entre sommets.

Si le taux de reproduction de base R_0 caractérise l'évolution de l'épidémie dans une population au tout début de l'épidémie lorsque la grande majorité de la population est sensible ou susceptible, le taux de reproduction effectif R_{eff} parfois noté R_t décrit l'évolution de l'épidémie sous les contraintes en vigueur telle que la vaccination, soit par immunité individuelle suite à une infection, soit au fur et à mesure que les gens meurent.

Il représente le nombre de personnes dans une population qui peuvent être infectées par

un individu malade à un moment donné. Ce taux peut être régulièrement recalculé afin d'estimer l'évolution de l'épidémie en fonction des mesures qui ont été prises. Si ce taux de reproduction est inférieur à 1, cela veut dire que chaque malade va contaminer moins d'une personne susceptible dans la population et l'épidémie va progressivement disparaître. Mais s'il est supérieur à 1, chaque malade infectera plus d'une personne susceptible et l'épidémie se poursuivra.

Notons donc que le but des autorités est naturellement de faire baisser le taux R_{eff} et de le maintenir en dessous de 1 afin de contenir, stopper la progression de l'épidémie.

Bien que facilement interprétable, cette valeur est très souvent difficile à exprimer dans la réalité, à cause de l'absence des données caractérisant la maladie telle que la probabilité de transmission de la maladie à chaque contact ou alors le nombre de contacts que chaque personne rencontre par jour.

C'est notamment le cas pour l'épidémie de la covid-19 qui selon le journal le Monde dans son article publié par Assma Maad, Adrien Sénécat et Maxime Vaudano le 15 mars 2021 et mis à jour le 16 mars 2021 estimait qu'à cette date, il restait compliqué d'évaluer les probabilités de transmission de la covid-19 vu la multitude de facteurs qui entrent en jeu.

6.1 Modèle par compartiment

Dans les modèles par approche par compartiment, le taux de reproduction effectif peut être défini par le produit entre les trois éléments suivants : le taux de reproduction de base, la fraction restante de la population susceptible d'être infectée et la probabilité qu'un individu infecté ne soit pas isolé.

Mathématiquement cette relation se traduit par la relation suivante :

$$R_t = R_0 * (1 - p) * s \quad (6.3)$$

Avec s la fraction de la population susceptible d'être infectée à l'instant t , p la probabilité qu'un individu infecté soit isolé, R_0 taux de reproduction de base.

6.2 Modèle par graphe

Dans les modèles multi-agents par approche par réseaux de contacts aléatoire ou graphes, nous définissons le taux de reproduction effectif à un instant t comme le produit de la probabilité de transmission et du degré moyen du graphe des susceptibles à l'instant t .

Grâce à la théorie de percolation vu plus haut, nous redéfinissons le graphe G à chaque

instant, notre principe est donc le suivant :

À l'instant $t + 1$, le graphe est percolé des p noeuds guéris, donc infectés au temps t , c'est-à-dire que les noeuds contaminés au temps t et guéris au temps $t + 1$ sont supprimés, nous obtenons donc grâce à cette opération un graphe temporaire noté G' .

Ensuite, nous trouvons le degré moyen du graphe temporaire G' et nous le multiplions par la probabilité de transmission p , pour obtenir le taux effectif R_{eff} , le processus est repris ainsi de suite jusqu'à épuisement des infectées dans le graphe.

Ce principe se traduit mathématiquement par la relation suivante :

$$R_t = p * \mu(t)$$

Avec $\mu(t)$ le degré moyen du graphe G à l'instant t .

Notons que, au temps initial $t = 0$ le taux de reproduction effectif noté R_{eff} ou encore R_t est égale au taux de reproduction de base R_0 dans les deux relations précédentes.

6.3 Simulation

Dans cette section, nous simulons numériquement l'évolution de R_{eff} dans les modèles par compartiments et les modèles par graphes. Notons que dans le cas des modèles par graphes, plusieurs types de graphes sont utilisés : les graphes homogènes et hétérogènes.

Un modèle de graphe est dit homogène, si tous ses sommets jouent à peu près le même rôle en terme mathématique cela signifie qu'ils ont tous approximativement le même nombre de voisins ou d'arêtes.

C'est l'exemple des modèles de graphes aléatoires d'Erdős-Rényi étudiés dans les chapitres précédents, dans lequel la plupart des sommets du graphe de ce type ont un nombre de connexions proche de la valeur du degré moyen du graphe.

Par contre, un modèle de graphe sera dit hétérogène, s'il existe certains de ses sommets qui jouent des rôles particuliers, on dit qu'ils possèdent en fait des statuts particuliers. C'est l'exemple du graphe Barabási-Albert étudié plus haut, où nous avons des sommets appelés hubs qui possèdent un nombre de connexions bien plus grand que le degré moyen du graphe.

Dans la vie réelle, on constate que beaucoup d'individus sont en relation avec peu de

personnes donc possèdent peu de connexions alors que certains autres sont en relation avec un très grand nombre donc possèdent un très grand nombre de connexions. Par conséquent, les réseaux de contacts sociaux sont hétérogènes. Toutes ces raisons ont motivé nos choix pour ces trois catégories de graphes lors de nos simulations.

Données :

Hypothèses :

Nous supposons qu'une maladie pouvant causer une épidémie survient dans une population ayant les caractéristiques suivantes :

- La population totale est de 17903 personnes.
- Chaque individu est en relation ou en contact en moyenne avec 20 individus.
- La probabilité de transmission de cette maladie à chaque contact est de 0.04
- La durée moyenne de contagion d'une personne ayant contracté la maladie est de 5 jours.

Pour le modèle par compartiments de Kermack-McKendrick :

- Le taux de contamination β est donné par le produit entre le nombre de contacts et la probabilité de transmission à chaque contact donc : $20 * 0.04 = 0.8$.
- Le taux de reproduction de base R_0 est donné par le produit entre taux de contamination et la durée moyenne de contagion $\beta * D$, soit $20 * 0.04 * 5 = 4$.

Pour le modèle basé sur les graphes de Reed-Frost :

- Le graphe comporte 17903 sommets.
- Le degré moyen du graphe μ_D est égale au nombre de contacts moyen de chaque individu, soit $\mu_D = 20$.
- La probabilité de contamination est donné par le produit de la durée de contagion D par la probabilité de transmission, soit $5 * 0.04 = 0.2$
- Le taux de reproduction de base R_0 est donné par le produit du degré moyen du graphe et la probabilité de contamination, soit $20 * 0.2 * 5 = 4$.

Chapitre 6. Taux de reproduction effectif : R_{eff}

Sous hypothèses et données décrites précédemment, la figure 6.3 décrit l'évolution du R_t dans le modèle *SIR* déterministe par compartiments de Kermack-McKendrick :

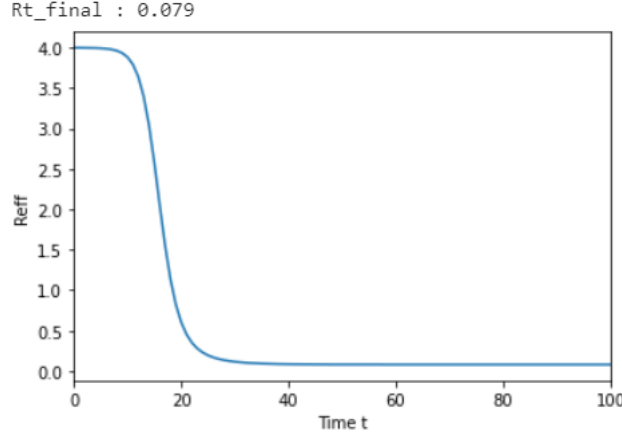


FIG. 6.3 : Évolution du taux effectif cas du modèle de Kermack-McKendrick

Les figures 6.4a et 6.4b décrivent la distribution des degrés dans le graphe aléatoire d'Erdős-Rényi et l'évolution du R_t pour le modèle multi-agents *SIR* Reed-Frost pour une simulation de 200 exécutions.

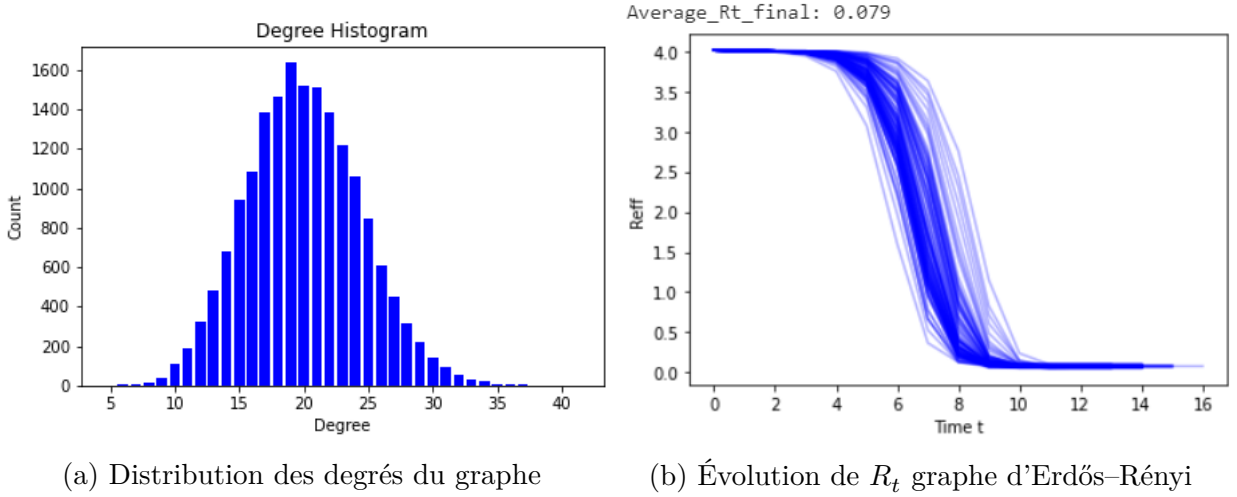


FIG. 6.4 : Simulation pour un modèle de Reed-Frost sur le graphe d'Erdős-Rényi

Les figures 6.5a et 6.5b décrivent la distribution des degrés dans le graphe aléatoire de Barabási-Albert et l'évolution du R_t pour le modèle multi-agents *SIR* Reed-Frost pour une simulation de 200 exécutions.

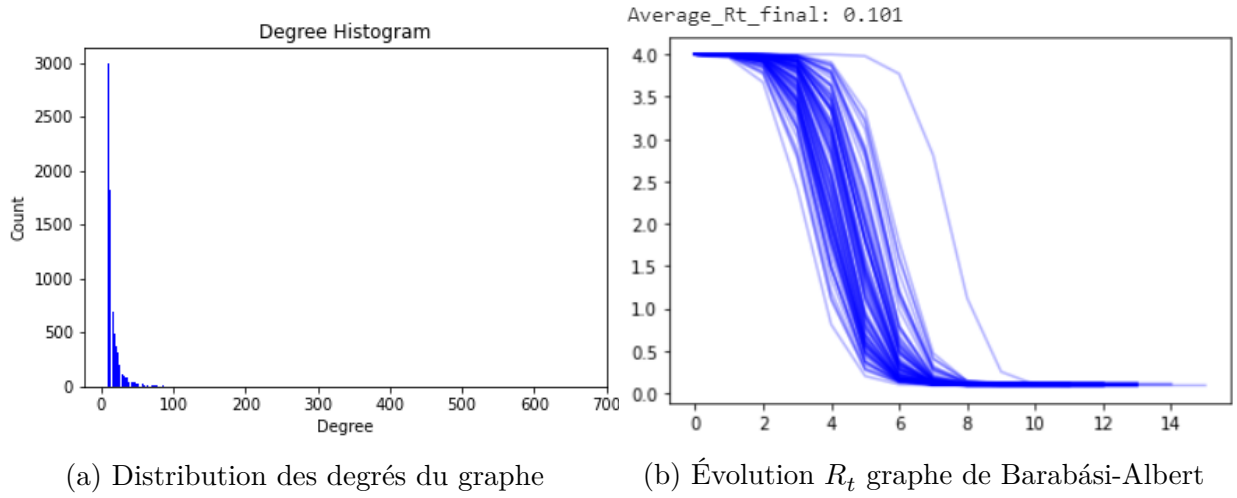


FIG. 6.5 : Simulation pour un modèle de Reed-Frost sur le graphe de Barabási-Albert

La figure 6.6a décrit la distribution des degrés dans le graphe social réel [30] de même caractéristiques (taille de population et degré moyen) que les deux modèles de graphes d'Erdős-Rényi et Barabási-Albert précédents.

La figure 6.6b quant à elle décrit l'évolution du R_t dans le graphe réel pour le modèle multi-agents *SIR* Reed-Frost pour une simulation de 200 exécutions.

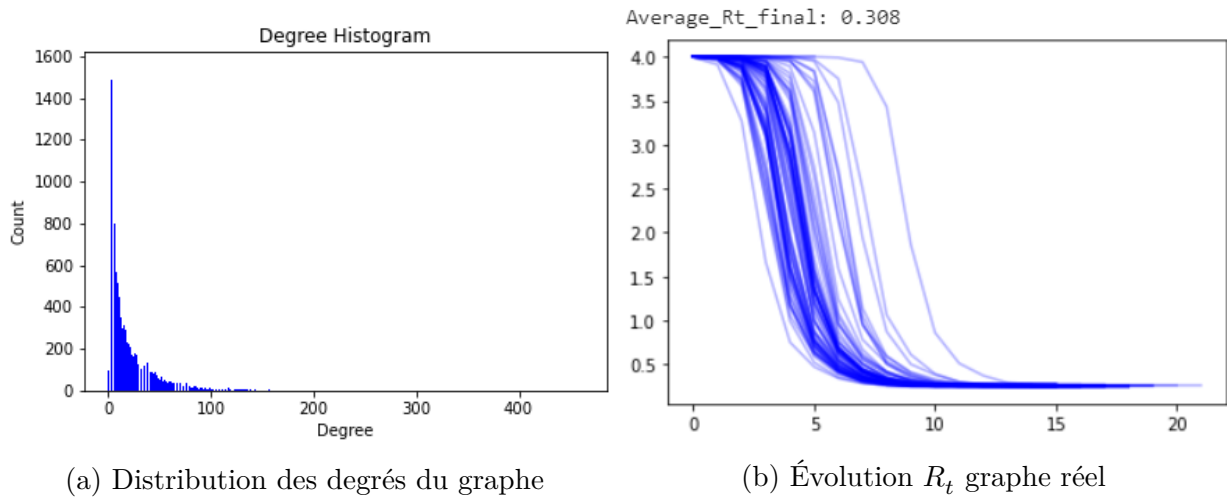


FIG. 6.6 : Simulation pour un modèle Reed-Frost sur le graphe réel

Comme nous l'avons mentionné plus haut, nous utilisons la théorie de la percolation des graphes, afin d'obtenir le degré moyen du graphe à chaque instant ; De ce fait, intéressons-nous maintenant à la configuration de nos trois différents graphes à la fin de l'épidémie.

Les figures que nous présenterons dans la suite sont des résultats obtenus pour une unique exécution de chaque modèle.

En premier, les figures 6.8a et 6.8b décrivent la distribution des degrés et la structure du graphe aléatoire du modèle Erdős–Rényi à la fin de l'épidémie.

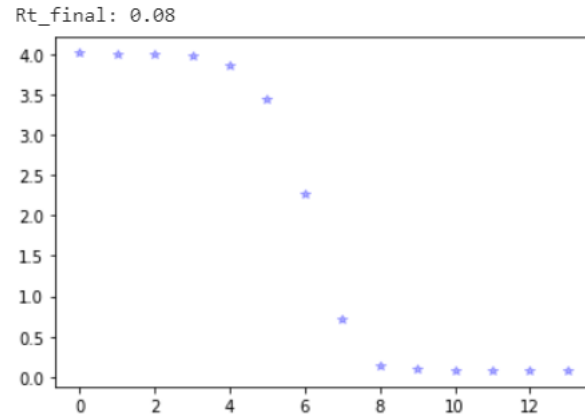
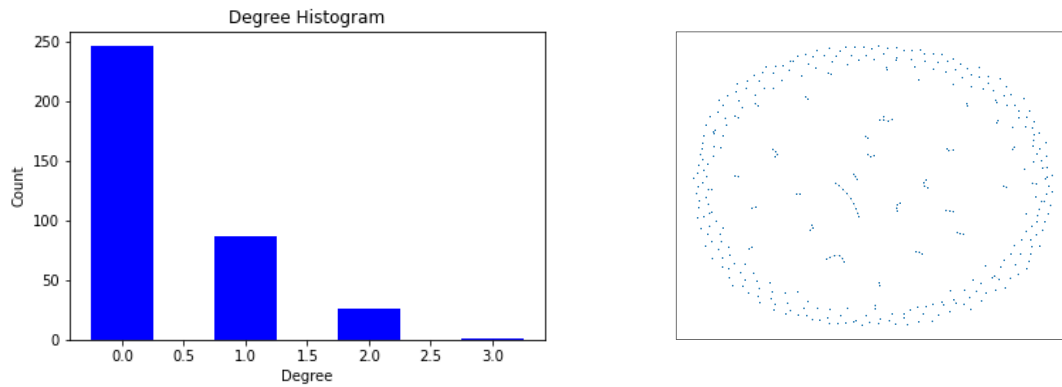


FIG. 6.7 : Estimation du taux effectif R_{eff} par le modèle multi-agents Reed-Frost sur le réseau d'Erdős–Rényi



(a) Distribution des degrés à la fin de l'épidémie (b) Structure du graphe à la fin de l'épidémie

FIG. 6.8 : Simulation pour un modèle Reed-Frost sur réseau Erdős–Rényi

En second, les figures 6.10a et 6.10b décrivent la distribution des degrés et la structure du graphe aléatoire du modèle Barabási–Albert à la fin de l'épidémie.

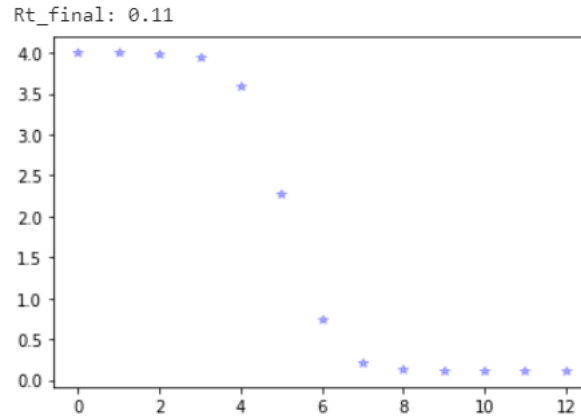
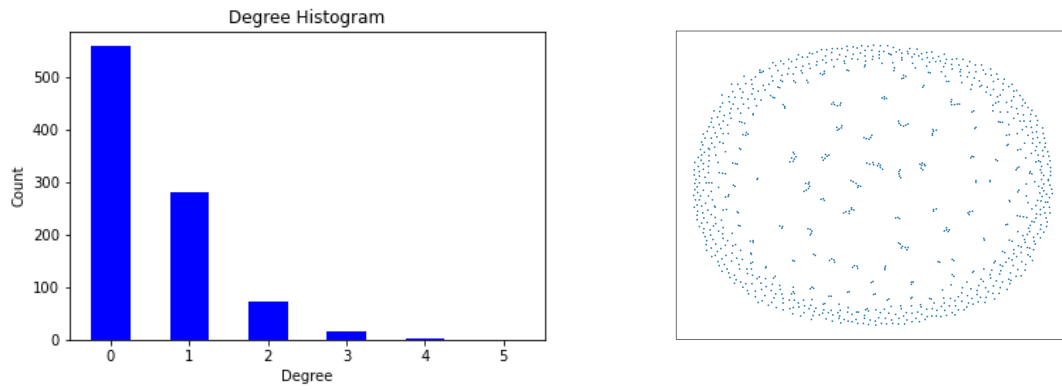


FIG. 6.9 : Estimation du taux effectif R_{eff} par le modèle multi-agents Reed-Frost sur le réseau de Barabási-Albert



(a) Distribution des degrés à la fin de l'épidémie (b) Structure du graphe à la fin de l'épidémie

FIG. 6.10 : Simulation pour un modèle Reed-Frost sur réseau Barabási-Albert

Enfin, les figures 6.12a et 6.12b décrivent la distribution des degrés et la structure du graphe aléatoire réel à la fin de l'épidémie.

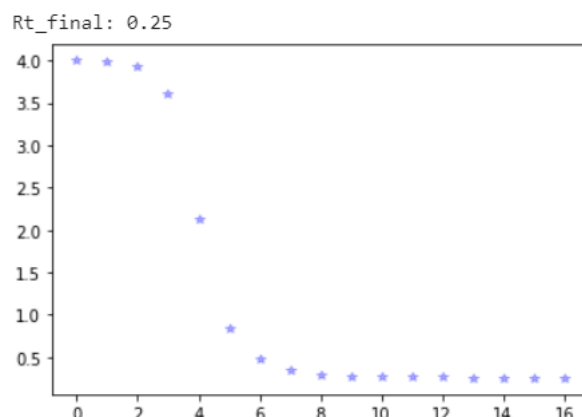
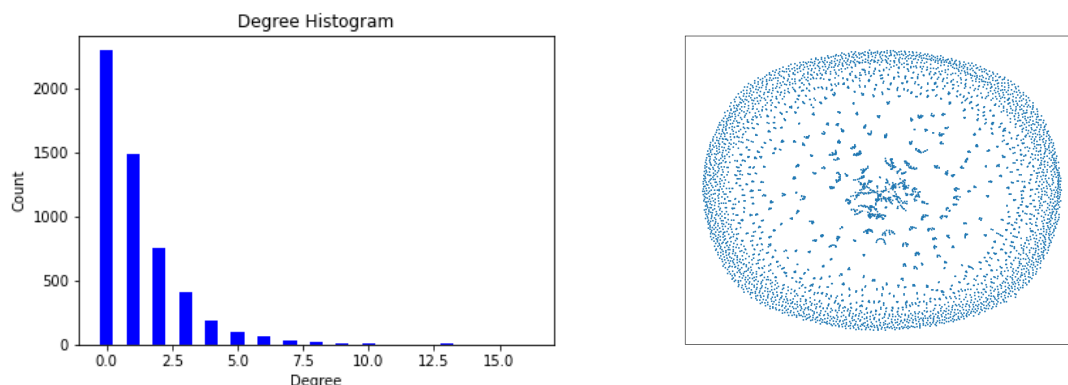


FIG. 6.11 : Estimation du taux effectif R_{eff} par le modèle multi-agents Reed-Frost sur le réseau réel



(a) Distribution des degrés à la fin de l'épidémie (b) Structure du graphe à la fin de l'épidémie

FIG. 6.12 : Simulation pour un modèle Reed-Frost sur le réseau aléatoire réel

6.4 Discussion

Nous constatons que la manière de calculer le taux effectif dans les systèmes par compartiments ici celui de Kermack-McKendrick et la manière de calcul le taux effectif dans les modèles basés sur les graphes ici le modèle de Reed-Frost coïncident de façon remarquable du début jusqu'à la fin de l'épidémie lorsque le graphe social est homogène.

Nous constatons que cette hypothèse d'homogénéité est conservée jusqu'à la fin de l'épidémie où le graphe a été décomposé en ensemble de très petites composantes connexes par phénomène percolation.

Dans le cas où le réseau social est hétérogène, le calcul du taux effectif coïncide raisonnablement à celui du calcul dans les modèles par compartiments au tout début de l'épidémie,

mais pas à la fin. Où malgré la décomposition du graphe par percolation au fur et mesure que l'épidémie progresse, l'hétérogénéité du graphe est de plus en plus accentuée. On remarque donc l'existence d'une composante géante à la fin de l'épidémie.

Nous nous sommes également intéressés au calcul du taux effectif dans les deux modèles avec plusieurs sources de contamination au temps initial $t = 0$, en considérant que 1% de la population totale soit infectée (voir Annexe B). Malgré une décroissance rapide du taux effectif dans les deux modèles considérés par rapport à ceux avec une seule source de contamination, les conclusions émises précédemment restent valables.

Ce travail a été essentiellement basé sur des hypothèses assez simples et les conclusions trouvées permettent de se faire une idée sur quel modèle utiliser pour l'évaluation du taux de reproduction effectif à chaque instant t dans la réalité. Ainsi, pour une meilleure évaluation du taux de reproduction effectif R_t celui-ci doit se faire sur des modèles basés sur graphes et principalement ceux hétérogènes. Toutefois, une évaluation de ce paramètre par des modèles par compartiments ou par graphe social homogène par manque de données sur la population ne serait pas une catastrophe car les deux modèles correspondent au début de l'épidémie et une fois le taux de reproduction effectif passé sous le seuil de 1, même si les deux modèles sont différents on est sûr que l'épidémie est sous contrôle et donc s'arrêtera.

Ne disposant pas toujours de toutes les données, notamment sur l'interaction d'une population à l'échelle d'un pays, plusieurs méthodes basées sur des données d'incidences ont été proposées pour estimer le taux de base effectif au cours d'une épidémie, l'une de ces méthodes est celle développée en septembre 2013 par Anne Cori et Neil Ferguson [31] chercheurs britanniques spécialisés en maladies infectieuses de l'Impérial College de Londres. Elle est principalement basée sur l'exploitation des données de surveillance épidémiologique notamment celle des apparitions des nouveaux cas d'infections.

Cette méthode suppose une population homogène et exploite grâce au processus de branchement de Galton-Watson les informations sur les séries chronologiques d'incidence de la maladie. Elles intègre l'incertitude dans la distribution de l'intervalle en série, c'est-à-dire le temps entre l'apparition des symptômes dans un cas primaire et l'apparition des symptômes dans les cas secondaires [32, 31]. À travers ce modèle, ils ont développé un estimateur bayésien qui tient compte du caractère aléatoire de l'apparition des infections ainsi que de la variation de l'intervalle sériel.

Cette méthode très populaire a été utilisée pour estimer le taux effectif sur les données de certaines épidémies historiques tel que grippe pandémique, le syndrome respiratoire aigu sévère, la rougeole et la variole et est aujourd'hui utilisée comme méthode d'estimation du taux de reproduction effectif pour le cas de la covid-19 en France.

Ce modèle est implémenté et fourni d'une part en version Microsoft Excel, d'autre part

dans un package R sous le nom d'EpiEstim et enfin via une interface Web disponible [ici](#) [33], que nous utiliserons par la suite.

Les figures suivantes représentent l'estimation de R_{eff} obtenue à travers l'interface web précédent(EpiEstim) pour l'épidémie spécifiée en données d'entrée. Les données d'incidences correspondent à la liste des nombres d'infections générés à chaque instant t par les simulations des différents modèles :

Cas du modèle par compartiments de Kermack-McKendrick

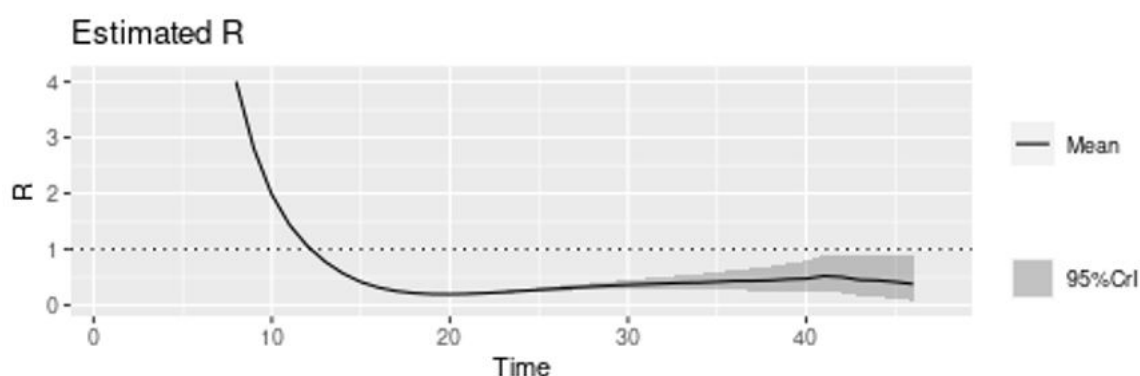


FIG. 6.13 : Estimation du R_{eff} par la méthode Corri-Ferguson sur les données incidents du nombre de nouveaux infectés à chaque période de temps fournies par le modèle par compartiments de Kermack-McKendrick : À comparer avec la Fig.6.3

On remarque que le R_{eff} dans le modèle par compartiments de Kermack-McKendrick débute à 4 et se termine dans un intervalle de confiance à 95% entre $[0.071, 0.905]$. (Voir tableau 1 annexe D).

Cas du modèle multi-agents : réseau Erdős-Rényi

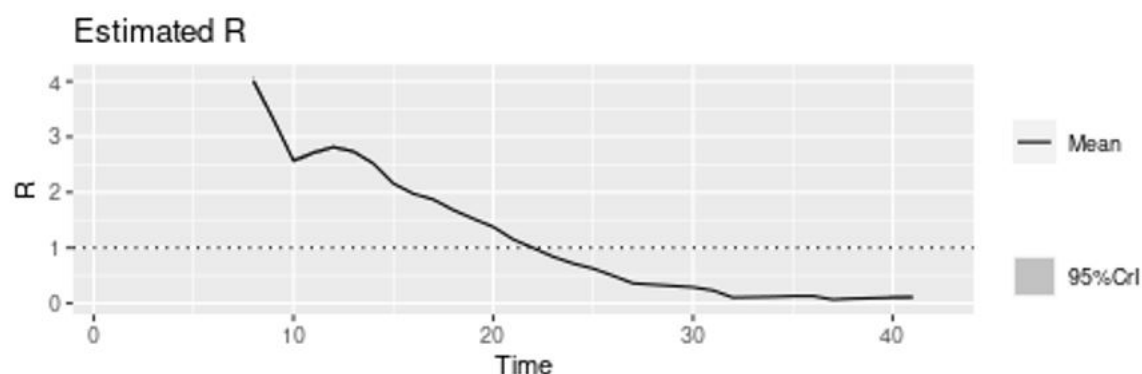


FIG. 6.14 : Estimation du R_{eff} par la méthode Corri-Ferguson sur les données incidents du nombre de nouveaux infectés à chaque période de temps fournies par le modèle multi-agents Reed-Frost sur le réseau Erdős-Rényi : À comparer avec la Fig.6.7

Dans le cas du réseau Erdős-Rényi, on remarque que le R_{eff} débute à 4 et se termine dans un intervalle de confiance à 95% entre $[0.069, 0.153]$. (Voir tableau 2 annexe D).

Cas du modèle multi-agents : réseau Barabási-Albert

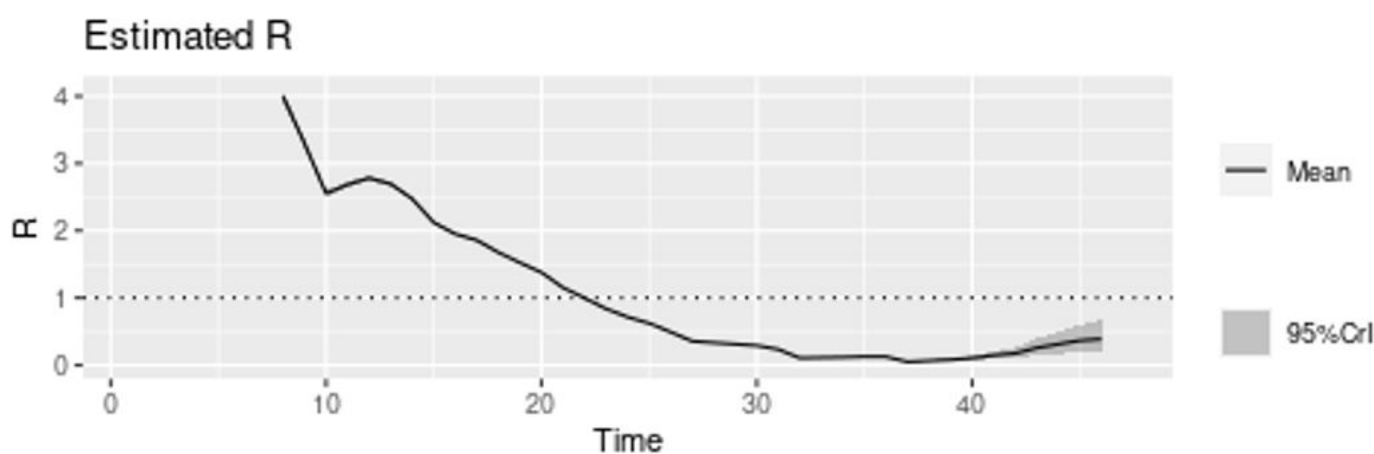


FIG. 6.15 : Estimation du R_{eff} par la méthode Corri-Ferguson sur les données incidents du nombre de nouveaux infectés à chaque période de temps fournies par le modèle multi-agents Reed-Frost sur le réseau Barabási-Albert : À comparer avec la Fig.6.9

Ici dans le réseau Barabási-Albert, on remarque que le R_{eff} débute à 4 et se termine dans un intervalle de confiance à 95% entre $[0.176, 0.667]$. (Voir tableau 3 annexe D).

Cas du modèle multi-agents : réseau réel

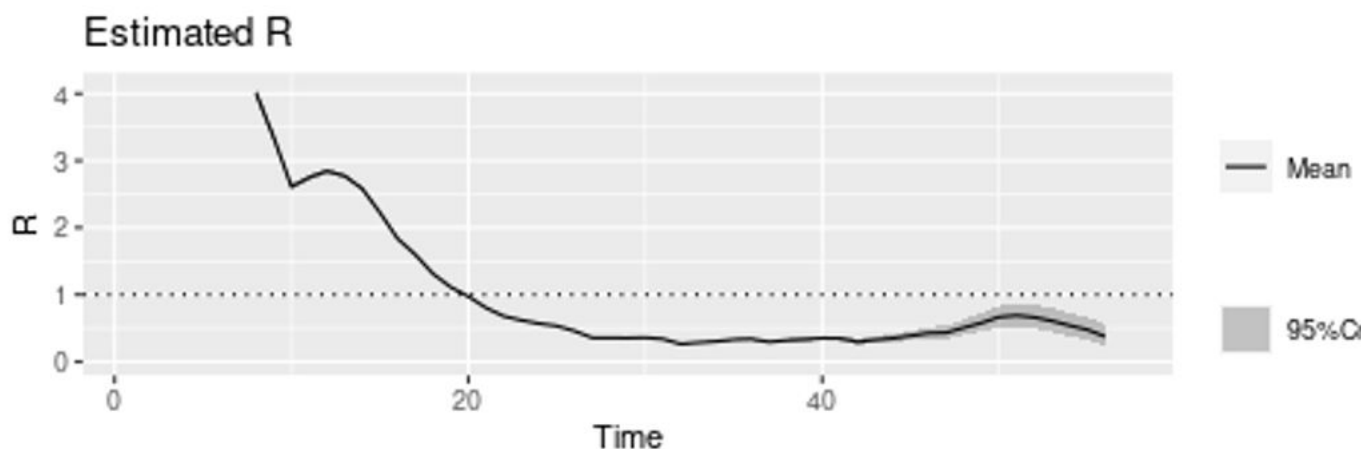


FIG. 6.16 : Estimation du R_{eff} par la méthode Corri-Ferguson sur les données incidents du nombre de nouveaux infectés à chaque période de temps fournies par le modèle multi-agents Reed-Frost sur le réseau réel : À comparer avec la Fig.6.11

Enfin, dans le réseau réel on remarque que le R_{eff} débute à 4 et se termine dans un intervalle de confiance à 95% entre $[0.229, 0.557]$. (Voir tableau 4 annexe D).

On constate par comparaison des figures précédentes que, l'estimation de l'évolution du R_{eff} par la méthode d'Anne Cori et Neil Ferguson :

- Fournit des résultats à peu près similaires aux estimations du R_{eff} effectuées plus haut dans tous les modèles pour ce qui est du début de l'épidémie.
- Pour ce qui est de la valeur de l'estimation du R_{eff} à la fin, la méthode d'Anne Cori et Neil Ferguson fournit un intervalle de confiance à 95% du R_{eff} contenant la valeur finale de l'estimation du R_{eff} dans les modèles précédents exception faite sur le modèle Reed-Frost sur le réseau de Barabási-Albert.
- Pour le cas particulier du modèle multi-agent Reed-Frost, on remarque des dissemblances sur certains points dans la variation de la courbe.

Toutefois, remarquons que le modèle d'Anne Cori et Neil Ferguson fournit via l'interface web est difficile à mettre en oeuvre, car il nécessite en plus des paramètres de réglage statistique, des hypothèses paramétriques sur la distribution des nouveaux cas d'infections.

Chapitre 7

Bulles

7.1 Contexte

Les épidémies sont des situations extrêmement dangereuses, nous avons vu dans les chapitres précédents un ensemble de paramètres permettant de catégoriser et de classer les épidémies dont le plus important est le nombre de reproduction de base. L'étude montre que pour certaines épidémies si rien n'est fait, c'est-à-dire si aucune mesure n'est prise, l'épidémie affectera une proportion de plus de 90% de la population totale, ceci est une situation probablement inacceptable.

Les stratégies sont donc ainsi développées par des professionnels de santé publique pour essayer de freiner, stopper et ou définitivement éliminer la propagation du virus dans une population. Ces stratégies visent généralement à rechercher l'atteinte d'un seuil communément appelé l'immunité de groupe ou immunité collective, qui représente en fait le pourcentage d'une population donnée qui doit être immunisée contre une maladie à partir duquel un sujet infecté introduit dans cette population va transmettre l'infection à moins d'une personne en moyenne.

Nous avons vu que cette proportion est donnée par la relation $1 - \frac{1}{R_0}$ et donc fortement dépendant du taux de reproduction de base R_0 de la maladie. De ce fait plus le taux de reproduction de base de la maladie sera élevé, plus cette proportion sera grande.

Deux stratégies sont généralement utilisées afin d'atteindre cette proportion :

- La première consiste à laisser librement le virus se propager dans la population jusqu'à ce que la proportion de personnes nécessaire contaminée et guérie (donc immunisée contre le virus) pour que le virus ne circule plus soit atteint. Il est clair que cette stratégie ne

peut être utilisée uniquement pour les maladies donc la guérison entraîne l'immunité, et pour des raisons d'éthique, ne peut pas être utilisée dans les cas épidémies qualifiées de dangereuses, c'est-à-dire pour des épidémies ayant une probabilité de décès élevé lorsqu'on contracte la maladie et donc on a recours à la seconde stratégie qui est celle de la vaccination.

- La vaccination est une approche éthique et scientifiquement correct pour protéger les individus contre une infection, le succès de cette technique dans l'éradication des épidémies n'est plus à démontrer. Cette technique permet de diminuer la population susceptible à une infection, et donc implicitement de réduire le nombre de personnes que le virus pourra infecter. La question fondamentale dans cette approche étant toujours la recherche du pourcentage de personnes à vacciner pour que le virus d'une épidémie ne puisse plus circuler ou se propager dans une population. Cette proportion correspond comme nous l'avons vu au seuil de l'immunité de groupe.

Ainsi le seuil d'immunité collective peut être atteint soit de manière naturelle, soit grâce à la vaccination.

Dans certaines situations, comme l'apparition des nouvelles épidémies ou des variants d'une épidémie déjà existante, le développement et la validation d'un vaccin pour l'éradication de l'épidémie ou d'un variant est généralement très long, il faut donc trouver autre chose pour atteindre les objectifs cités plus haut. Des nouvelles stratégies ont donc vu le jour, donc l'une d'elle est le concept de bulles sociales. Ce concept est une technique assez nouvelle ayant fait son apparition lors de la crise sanitaire que vit actuellement le monde avec l'épidémie de la covid-19, une étude récente à évaluer son impact sur les stratégies de déconfinement en Belgique démontrant son efficacité dans la réduction du taux d'hospitalisation en Belgique [34].

Dans les prochaines sections, nous définirons un principe de mise en oeuvre principalement dans les modèles multi-agents, c'est-à-dire pour des modélisations épidémiologiques par graphe principalement à l'aide des réseaux aléatoires de contacts et comparons son impact dans les différentes situations de réseaux aléatoires de contacts précédemment étudiés.

Remarquons que l'ensemble des stratégies mises en oeuvre par les politiques de santé publique visent principalement à rendre l'épidémie contrôlable, c'est-à-dire à ramener son taux de reproduction effectif sous la barre de 1.

7.2 Modélisation

Le concept de bulle sociale est une approche nouvelle ayant fait apparition dans certains pays telle que la Belgique et l'Angleterre pour faire barrage à l'épidémie de la covid-19.

Chapitre 7. Bulles

C'est un concept donc l'approche est assez simple, elle consiste à la création d'une sphère (bulle) contenant un ensemble de personnes avec qui on est en relation. Cette bulle est en principalement caractérisée par sa taille, qui représente le nombre total de personnes en relation dans la bulle.

De façon plus théorique, dans un système de modélisation épidémiologique par graphe, supposons que nous disposons d'un réseau social G , constitué de N sommets ou noeuds représentant des personnes, de M arêtes entre ces N sommets représentant les relations, supposons que nous ignorons les paramètres individuels de chaque noeud du réseau principalement le nombre de voisins et définissons comme départ le degré moyen du réseau social comme le nombre moyen de voisin pour chaque noeud choisi au hasard dans le graphe.

Dans ce réseau social ainsi défini, mettons en oeuvre les bulles sociales de taille k .

Notre principe est donc le suivant :

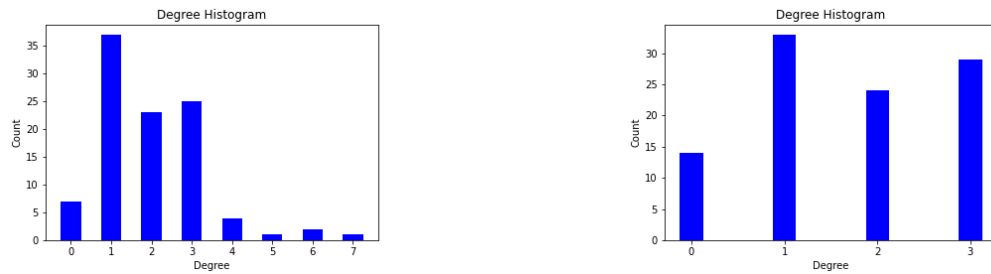
- Nous supprimons entièrement dans un premier temps, toutes les arêtes de notre réseau social.
- Ensuite nous parcourons l'ensemble des noeuds constituant notre réseau, et pour chacun de ses noeuds, nous appliquons la procédure suivante :
 1. Nous déterminons pour le noeud courant la taille de liste des noeuds non saturés (c'est-à-dire ne disposant pas encore d'un nombre total de k arêtes actives) avec lequel il est en contact ou relation (voisin), notons le v .
 2. Si cette taille est égale à 0 alors nous passons au noeud suivant.
 3. Sinon nous déterminons pour ce noeud combien de contact il lui reste pour être saturé, notons ce nombre n .
 4. Enfin, nous activons aléatoirement pour ce noeud n arêtes correspondant aux n nouveaux voisins choisis parmi les v voisins non saturés déterminés en 1.

À la fin de cette procédure, nous obtenons donc un nouveau réseau social, où chaque personne ne peut être en relation qu'avec un maximum de k de personnes.

Les figures suivantes sont des résultats de la distribution de degré de noeuds avant et après la mise en oeuvre des bulles sur deux réseaux. Sur chaque réseau, nous utilisons la même taille de bulle qui est 3.

Chapitre 7. Bulles

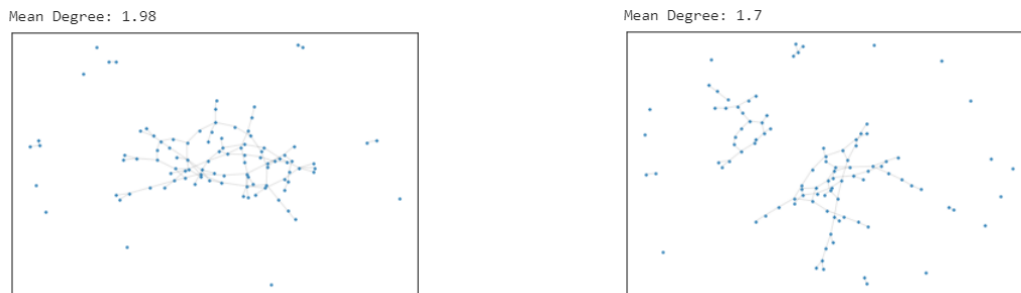
Les figures 7.1a et 7.1b décrivent respectivement la distribution des degrés pour un réseau Erdős–Rényi comportant de 100 noeuds.



(a) Distribution avant introduction des bulles (b) Distribution après introduction des bulles

FIG. 7.1 : Distribution des degrés avant et après l'introduction des bulles dans le réseau Erdős–Rényi.

Les figures 7.2a et 7.2b décrivent le réseau avant et après la mise en place des bulles de taille 3. Nous remarquons que ce procédé fait passer le degré moyen initial du réseau de 2 à 0.98.

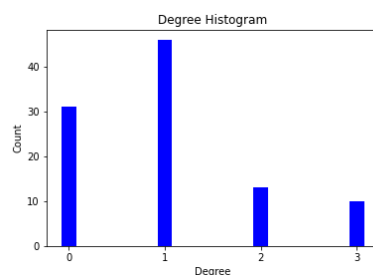
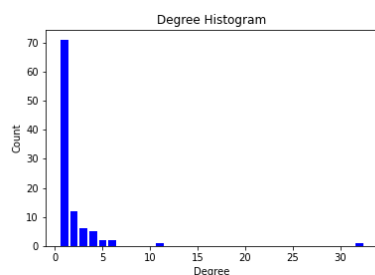


(a) Réseau social de départ (b) Après introduction de bulle de taille 3

FIG. 7.2 : Représentation du réseau aléatoire Erdős–Rényi avant et après l'introduction des bulles dans le réseau.

Les figures 7.3a et 7.3b décrivent également respectivement la distribution de degré d'un réseau aléatoire de Barabási–Albert G , comportant le même nombre de noeuds et de même degré moyen que le réseau aléatoire Erdős–Rényi précédent.

Chapitre 7. Bulles

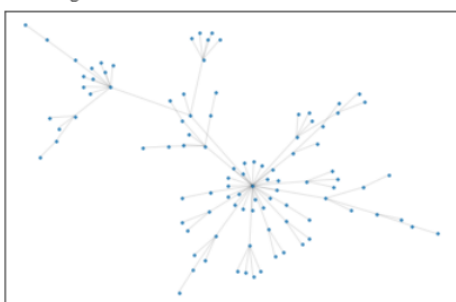


(a) Distribution avant introduction des bulles (b) Distribution après introduction des bulles

FIG. 7.3 : Distribution des degrés avant et après l'introduction des bulles dans le réseau Barabási-Albert.

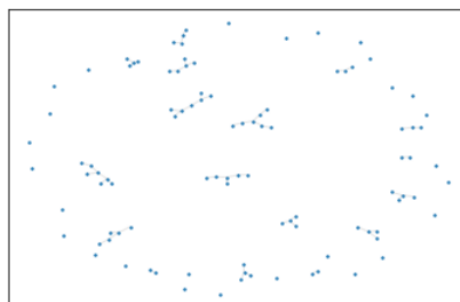
L'effet de la mise en place de bulle sociale de taille 3 dans le réseau initial fait passer le degré moyen de ce dernier de 2 à 1.1 comme nous le montre les figures 7.4a et 7.4b.

Mean Degree: 1.98



(a) Réseau social de départ

Mean Degree: 1.1

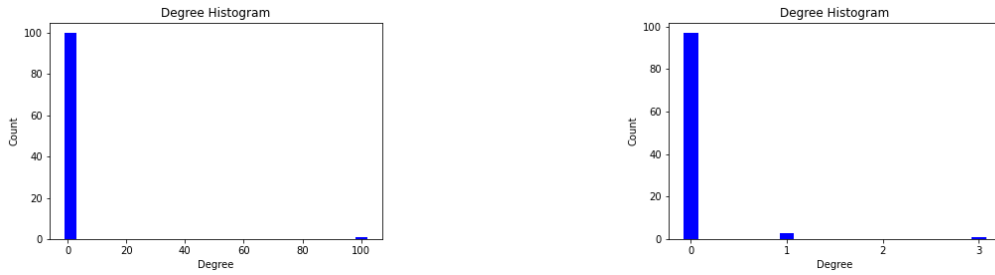


(b) Après introduction de bulle de taille 3

FIG. 7.4 : Représentation du réseau aléatoire Barabási-Albert avant et après l'introduction des bulles dans le réseau.

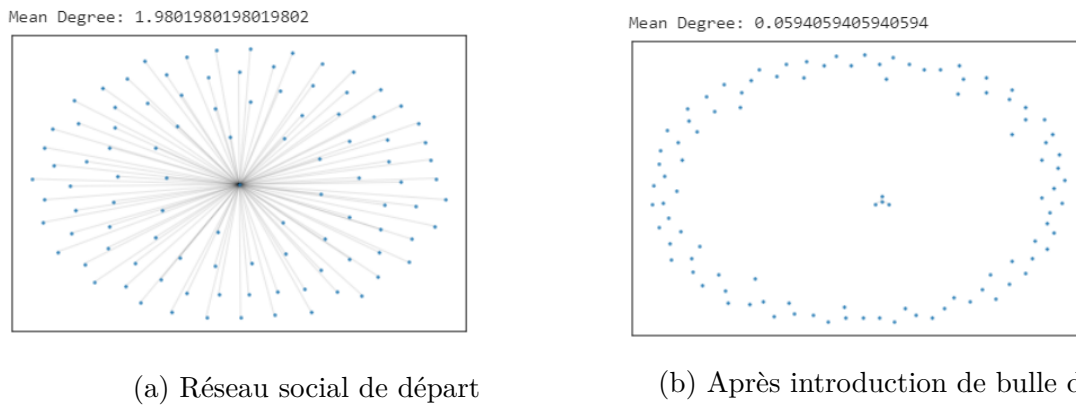
Cet effet est encore plus drastique, comme nous pouvons le constater par illustration à la figure 7.6b. Le réseau dans est un réseau de Barabási-Albert de même caractéristiques que ceux précédents et comportant un noeud central (figure 7.6a).

L'introduction des bulles sociales de taille 3 sur un tel réseau fait varier le degré moyen initial de 2 à un degré moyen de 0.059.



(a) Distribution avant introduction des bulles (b) Distribution après introduction des bulles

FIG. 7.5 : Distribution des degrés avant et après l'introduction des bulles dans le réseau aléatoire Barabási-Albert avec hub.



(a) Réseau social de départ

(b) Après introduction de bulle de taille 3

FIG. 7.6 : Représentation du réseau aléatoire Barabási-Albert avec hub central avant et après l'introduction des bulles dans le réseau.

7.3 Mise en oeuvre sur données à grande échelle

Dans cette section, nous vérifions les résultats obtenus plus haut sur des réseaux de contacts à grande échelle.

Dans la même logique que celle faite plus haut, nous mettons en oeuvre des bulles sociales de taille k dans nos différents réseaux de contacts. Cette fois-ci nous nous intéressons à des réseaux de contacts de grande taille plus représentatif du monde réel.

Reprenons les données du chapitre 6 comme données pour nos simulations, c'est-à-dire : Supposons qu'une maladie pouvant cause une épidémie arrive dans une population ayant les caractéristiques suivantes :

- La population totale est de 17 903 personnes.
- Chaque individu pris au hasard dans cette population possède en moyenne 20 individus

Chapitre 7. Bulles

avec qui il est en contact.

- La probabilité de transmission de cette maladie à chaque contact est de 0.2.

Modélisons donc la population décrite plus haut par 3 réseaux aléatoires dont 2 synthétiques (Erdős–Rényi et Barabási–Albert) et le dernier issu du monde réel [30], possédant les mêmes caractéristiques en nombre de noeuds et degré moyen.

Introduisons dans ces réseaux des bulles sociales de taille $k = 8$ et intéressons-nous à ses effets.

Les figures 7.7a et 7.7b décrivent respectivement la variation du nombre de personnes infectées par le virus avant et après l'introduction des bulles de taille 8 pour le cas du réseau Erdős–Rényi.

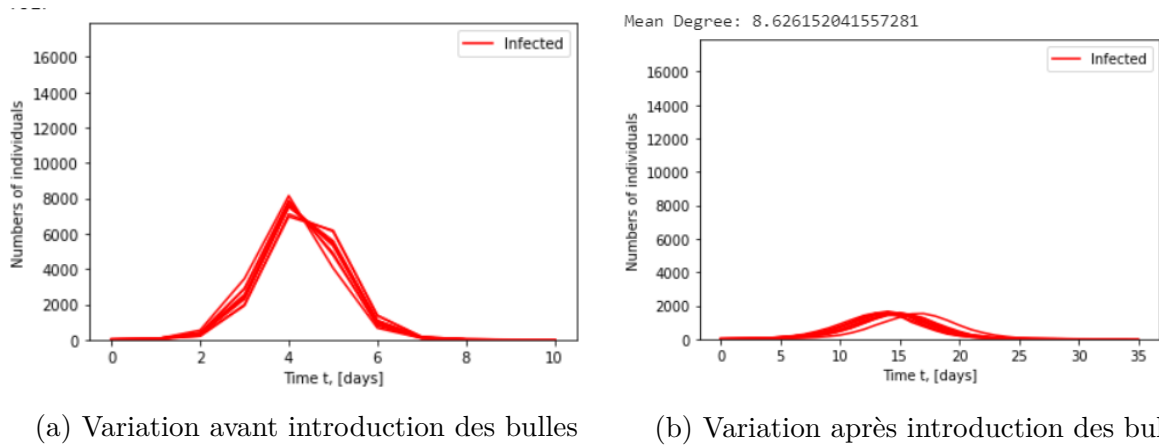


FIG. 7.7 : Variation du nombre de personnes infectées avant et après l'introduction des bulles dans le réseau aléatoire d'Erdős–Rényi

Les figures 7.8a et 7.8b décrivent respectivement la variation du nombre de personnes infectées par le virus avant et après l'introduction des bulles de taille 8 pour le cas du réseau Barabási–Albert.

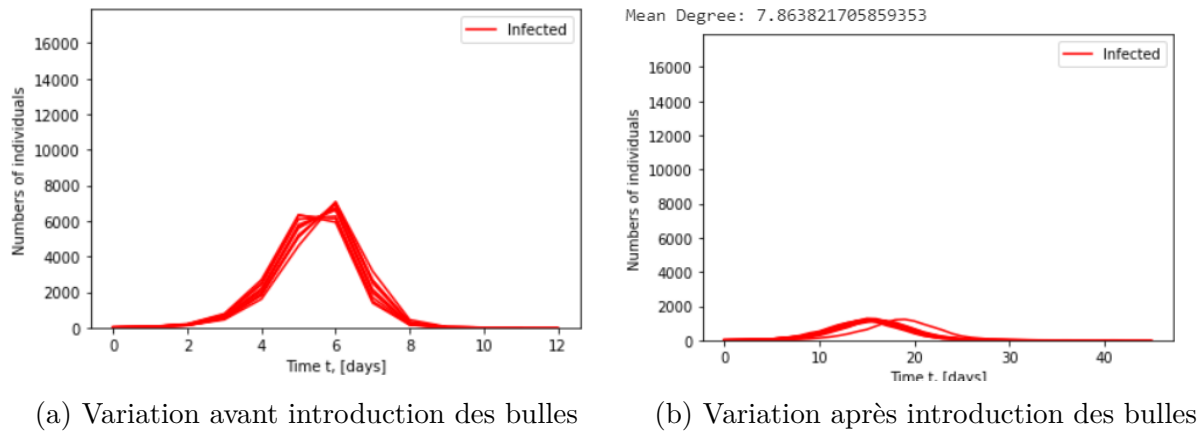


FIG. 7.8 : Variation du nombre de personnes infectées avant et après l'introduction des bulles dans le réseau aléatoire de Barabási-Albert

Les figures 7.9a et 7.9b décrivent respectivement la variation du nombre de personnes infectées par le virus avant et après l'introduction des bulles de taille 8 pour le cas du réseau réel.

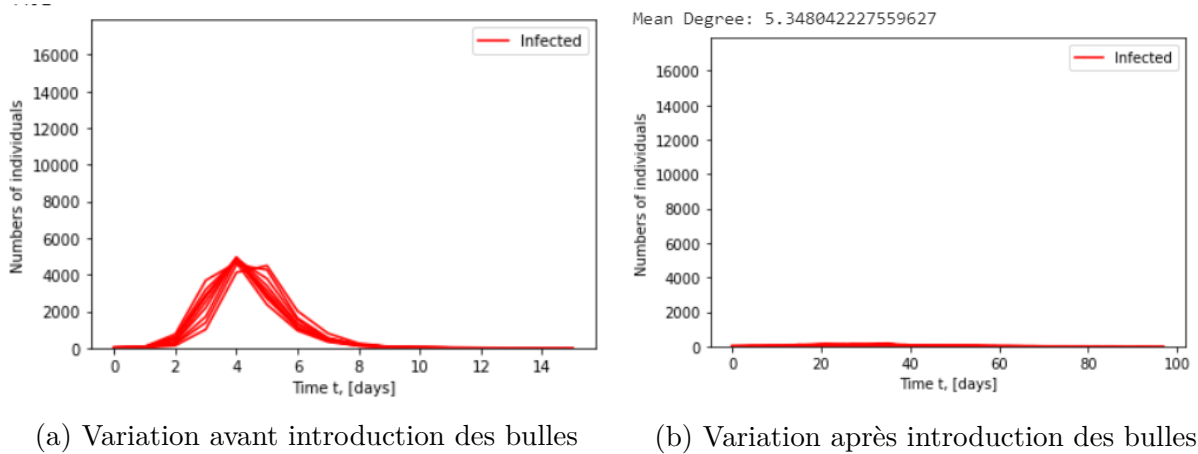


FIG. 7.9 : Variation du nombre de personnes infectées avant et après l'introduction des bulles dans le réseau aléatoire réel

Nous remarquons en premier que l'infection se propage beaucoup plus rapidement, et en second que le nombre maximum de personnes atteint est plus grand dans les réseaux homogènes qu'hétérogènes. Ainsi plus le réseau est hétérogène, plus le temps de propagation est long et plus bas est le nombre de personnes maximum atteint par la maladie.

L'introduction des bulles permet dans un premier temps de freiner la vitesse de propagation de la maladie en augmentant le temps de propagation de l'épidémie, et enfin de

Chapitre 7. Bulles

réduire également le nombre de personnes maximum atteint par la maladie ceci quel que soit le réseau.

Remarquons toutefois que l'intensité de l'impact n'est pas la même dans tous les réseaux, plus le réseau est hétérogène plus l'intensité de l'impact des bulles est drastique comme l'indique la figure 7.9b. L'introduction des bulles dans le réseau stoppe la propagation du virus.

En conclusion, les bulles sont plus efficaces dans les réseaux aléatoires hétérogènes que ceux homogènes, ceci est expliqué par la présence des super-spreaders dans les réseaux hétérogènes qui sont complètement neutralisés par l'introduction des bulles dans ces réseaux.

Enfin pour une certaine taille de bulle k en dessous d'un seuil, l'efficacité des bulles est autant similaire dans les réseaux homogènes qu'hétérogènes, dans le cas de nos expériences en dessous de 7 comme taille de bulle, l'épidémie est stoppée dans tous les trois réseaux.

Toutes ses conclusions restent également valables pour les cas de propagation sur des réseaux avec plusieurs sources d'épidémie au temps $t = 0$ (voir annexe C).

Chapitre 8

Conclusion générale

Notre objectif dans ce mémoire était double, d'une part le travail visait à comparer les modèles basés sur la segmentation de la population en compartiments d'états avec les modèles multi-agents basés sur la représentation de chaque individu de la population comme une seule entité avec un ensemble d'interactions entre eux. Ces deux modèles sont précisément le modèle Reed-Frost et le modèle de Kermack-McKendrick, cette comparaison concerne l'évolution du taux de reproduction de base à chaque instant couramment appelé taux de reproduction effectif R_{eff} . D'autre part, le travail visait à évaluer l'impact des bulles sociales dans la propagation des épidémies notamment dans le cas des modèles épidémiologiques basés sur des réseaux d'agents.

Après une revue de la littérature portant sur les différents modèles par compartiments tous caractérisés par le nombre de compartiments constituant le modèle, une étude plus détaillée du modèle de Kermack-McKendrick, qui est un modèle présentant un bon compromis entre simplicité de modélisation et efficacité des résultats dans la réalité. En vue d'une comparaison, nous avons étudié le modèle stochastique de Reed-Frost qui est un modèle de propagation des maladies infectieuses sur graphes.

Nous constatons que pour une même épidémie capturée avec ces deux modèles, ils prédisent approximativement la même proportion de personnes touchée par la maladie à la fin de l'épidémie, particulièrement lorsque le réseau est homogène à l'exemple d'un réseau d'Erdős-Rényi. Cette constatation est faite de façon théorique grâce aux concepts de processus de branchement de Galton-Watson, de composante géante et aux résultats des simulations.

La théorie de la percolation dans les graphes nous a permis d'établir une équivalence entre le principe de vaccination dans le modèle par compartiment et celui dans le modèle multi-agents.

Ensuite, les deux différentes méthodes de calcul du taux effectif lors de la propagation d'une

Chapitre 8. Conclusion générale

épidémie dans une population pour des modèles par compartiments et ceux basés sur les graphes ont été établies. Grâce aux simulations, nous constatons que les deux méthodes coïncident raisonnablement tout au long du processus de propagation de l'épidémie dans la population lorsque le graphe social est une fois de plus homogène, dans le cas où celui-ci est hétérogène une différence dès les premiers stades de l'épidémie apparaît et cette différence s'accroît avec le degré d'hétérogénéité du réseau ; Nous validons ces résultats par un graphe plus hétérogène.

Ces deux différentes méthodes pour l'estimation du taux effectif dans les modèles multi-agents et les modèles par compartiments sont comparées à la méthode d'estimation du taux effectif très populaire d'Anne Cori et Neil Ferguson. Elles sont sensiblement égales au début et à la fin de l'épidémie même si la variation au cours du temps semble à peu près différentes. Enfin, pour les modèles multi-agents basés sur les graphes, nous modélisons et introduisons les bulles sociales dans le réseau, nous constatons que la mise en place des bulles sociales permettent de ralentir la progression d'une épidémie quelle que soit la nature du réseau.

Mais l'intensité de l'impact de la mise en place des bulles sociales varie selon la taille des bulles et la structure des interactions entre les individus de la population. D'une part cette intensité est proportionnelle au niveau d'hétérogénéité de la structure du réseau, de ce fait plus la structure du réseau sera hétérogène plus l'intensité de l'impact sera grande et d'autre part inversement proportionnelle à la taille de bulles c'est-à-dire plus la taille des bulles sera petite plus l'intensité de l'impact sera grande.

Toutes les simulations ont été effectuées avec une seule source de maladie au temps initial qui est le cas le plus fréquent dans la réalité et avec plusieurs sources de maladie, à la fin nous formulons les recommandations suivantes :

La mise en place des bulles sociales dans une population afin de lutter efficacement contre la propagation d'une maladie infectieuse dans celle-ci devrait viser à identifier et à limiter le nombre d'interactions des super-spreaders présents dans cette population.

Pour ce qui est du modèle à mettre en oeuvre pour capturer une épidémie en vue d'estimer l'évolution du taux effectif au fil du temps, le modèle peut être basé sur des graphes, ou sur des compartiments celui-ci ne changera pas l'évolution de ce facteur dans le temps, le tout étant que le modèle choisi puisse capturer efficacement l'épidémie.

Comme mentionné lors de nos discussions dans les chapitres précédents, les résultats et conclusions de ce travail proviennent des modélisations et d'applications des modèles sur des hypothèses assez simples ne reflétant pas forcément la réalité. Toutefois, ils permettent de se faire une idée sur des dangers éventuels encourus lors des choix de prise de décision en cas d'épidémie. Par conséquent, des perspectives d'améliorations de ce travail pourraient être faites en considérant les graphes sociaux dynamiques comme modélisation de population.

Bibliographie

- [1] Pierre-Yves Boëlle. La modélisation des épidémies de maladies émergentes : les exemples du chikungunya et de la pandémie grippale. *Annales des Mines - Responsabilité et environnement*, N° 51(3) :49, 2008.
- [2] Fred Brauer and Carlos Castillo-Chavez. Epidemic models. In *Texts in Applied Mathematics*, pages 345–409. Springer New York, nov 2011.
- [3] Pierre-Charles Pradier. Histoire des modèles pandémiques. *Revue d'économie financière*, N° 139-140(3) :19–29, feb 2021.
- [4] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics—i. *Bulletin of Mathematical Biology*, 53(1-2) :33–55, mar 1991.
- [5] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439) :509–512, oct 1999.
- [6] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684) :440–442, jun 1998.
- [7] Tom Britton. Epidemic models on social networks—with inference. *Statistica Neerlandica*, 74(3) :222–241, feb 2020.
- [8] Philip D. O'Neill. Perfect simulation for reed-frost epidemic models. *Statistics and Computing*, 13(1) :37–44, 2003.
- [9] Hangyul M. Chung, Megan M. Cartwright, David M. Bortz, Trachette L. Jackson, and John G. Younger. DYNAMICAL SYSTEM ANALYSIS OF STAPHYLOCOCCUS EPIDERMIDIS BLOODSTREAM INFECTION. *Shock*, 30(5) :518–526, nov 2008.
- [10] Fred Brauer, Pauline van den Driessche, and Jianhong Wu, editors. *Mathematical Epidemiology*. Springer Berlin Heidelberg, 2008.
- [11] G. Bastin. *Lectures on Mathematical Modelling of Biological Systems*. August 22, 2018.

Bibliographie

- [12] EL BERRAI. *Contribution à l'analyse déterministe du modèle SIR*. PhD thesis, Université Hassan I I-Mohammedia-Casablan, 2014.
- [13] A. G. McKendrick William Ogilvy Kermack. Contributions to the mathematical theory of epidemics. II. —the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 138(834) :55–83, oct 1932.
- [14] Kamal R Mahtani Jeffrey K Aronson, Jon Brassey. An introduction to viral reproduction numbers, r_0 and r_e . Technical report, On behalf of the Oxford COVID-19 Evidence Service Team Centre for Evidence-Based Medicine, Nuffield Department of Primary Care Health Sciences, University of Oxford, 2020.
- [15] Fiona M Guerra, Shelly Bolotin, Gillian Lim, Jane Heffernan, Shelley L Deeks, Ye Li, and Natasha S Crowcroft. The basic reproduction number (r_0) of measles : a systematic review. *The Lancet Infectious Diseases*, 17(12) :e420–e428, dec 2017.
- [16] https://en.wikipedia.org/wiki/Basic_reproduction_number.
- [17] <https://petiteshistoiresdessciences.com/2020/04/23/la-modelisation-epidemiologique-2-3-1906-la- peste-de-bombay>. April 23, 2020.
- [18] Joydip Dhar Sudipa Chauhan, Om Prakash Misra. Stability Analysis OF SIR Model with Vaccination. *American Journal of Computational and Applied Mathematics*, 4(1) :17–23, nov 2014.
- [19] K. B. Athreya and P. E. Ney. T. e. harris and branching processes. *The Annals of Probability*, 39(2), mar 2011.
- [20] Christine Jacob. Branching processes : Their role in epidemiology. *International Journal of Environmental Research and Public Health*, 7(3) :1186–1204, mar 2010.
- [21] Mogens Fosgerau, Daniel McFadden, and Michel Bierlaire. Choice probability generating functions. *Journal of Choice Modelling*, 8:1–18, sep 2013.
- [22] James C. Fu and Y. M. Chang. On probability generating functions for waiting time distributions of compound patterns in a sequence of multistate trials. *Journal of Applied Probability*, 39(1) :70–80, mar 2002.
- [23] J.P. Demailly. , *Analyse numérique et équations différentielles*. Presses universitaires de Grenoble, 1996.
- [24] P. Erdős and A. Rényi. On the evolution of random graphs. In *The Structure and Dynamics of Networks*, pages 38–82. Princeton University Press, dec 2011.

Bibliographie

- [25] Fan Chung and Linyuan Lu. The volume of the giant component of a random graph with given expected degrees. *SIAM Journal on Discrete Mathematics*, 20(2) :395–411, jan 2006.
- [26] PHILIP E SARTWELL. MEMOIR ON THE REED-FROST EPIDEMIC THEORY. *American Journal of Epidemiology*, 103(2) :138–140, feb 1976.
- [27] ABBEY H. An examination of the reed-frost theory of epidemics. *Hum Biol* 24(3) :201–33, 1952.
- [28] Dean Koch, Reinhard Illner, and Junling Ma. Edge removal in random contact networks and the basic reproduction number. *Journal of Mathematical Biology*, 67(2) :217–238, may 2012.
- [29] Pablo Crotti. Une introduction à la théorie de la percolation, Printemps 2009.
- [30] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.
- [31] Anne Cori, Neil M. Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9) :1505–1512, sep 2013.
- [32] Francisco Arroyo-Marioli, Francisco Bullano, Simas Kucinskas, and Carlos Rondón-Moreno. Tracking r of COVID-19: A new real-time estimation using the kalman filter. *PLOS ONE*, 16(1) :e0244474, jan 2021.
- [33] <https://shiny.dide.imperial.ac.uk/epiestim/>.
- [34] Lander Willem, Steven Abrams, Pieter J. K. Libin, Pietro Coletti, Elise Kuylen, Oana Petrof, Signe Møgelmoose, James Wambua, Sereina A. Herzog, Christel Faes, Philippe Beutels, and Niel Hens. The impact of contact tracing and household bubbles on deconfinement strategies for COVID-19. *Nature Communications*, 12(1), mar 2021.

Annexes

Annexe A

Kermack-McKendrick vs Reed-Frost : cas de plusieurs sources

Ici, nous comparons le modèle déterministe de Kermack-McKendrick et celui stochastique sur graphe Reed-Frost avec plusieurs sources de contaminées au temps initial : Nous supposons que 3% de la population sont infectées au départ de l'épidémie donc $I_0 = 300$.

Les figures suivantes décrivent l'évolution de l'épidémie pour les états de susceptibles, infectés et guéris.

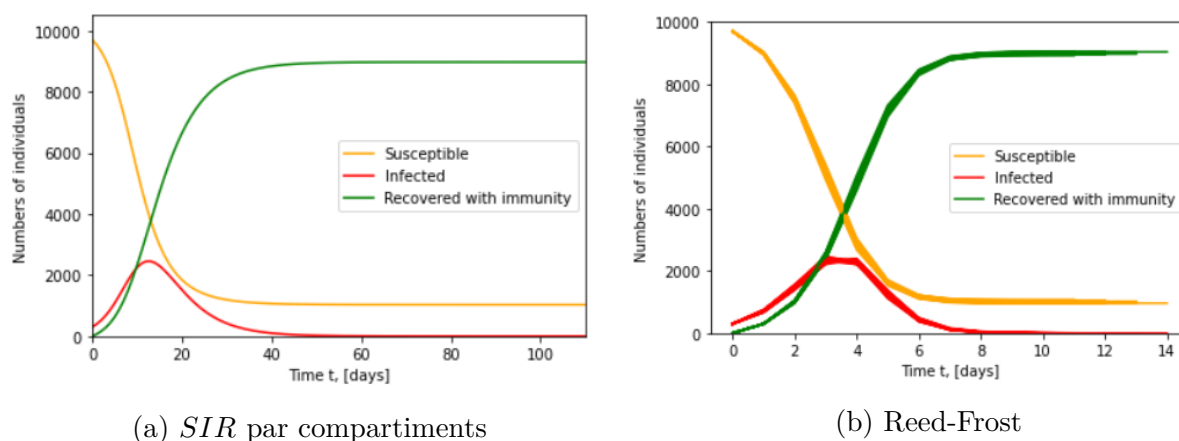
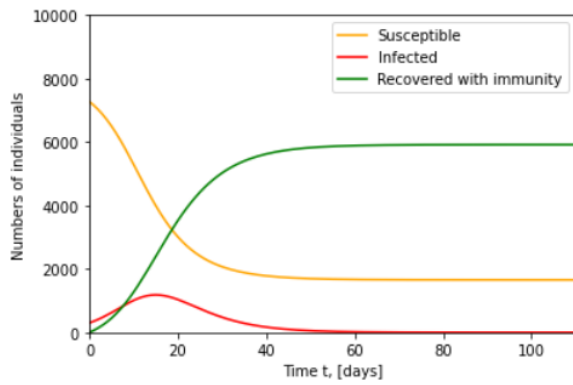


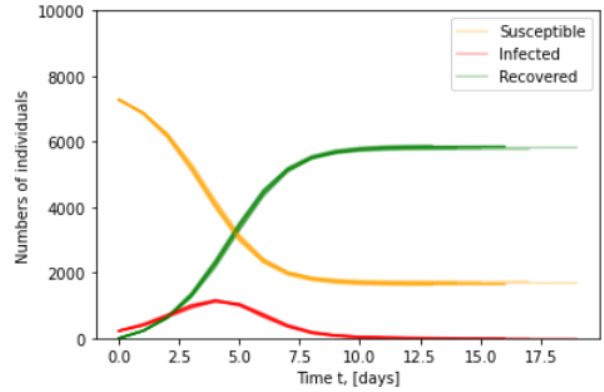
FIG. A.1 : Évolution de $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles

Annexe A. Kermack-McKendrick vs Reed-Frost : cas de plusieurs sources

Avec 25% de la population vaccinée : Les figures suivantes décrivent l'évolution des l'épidémie pour les états de susceptibles, infectés et guéris après vaccination de 25% de la population initiale.



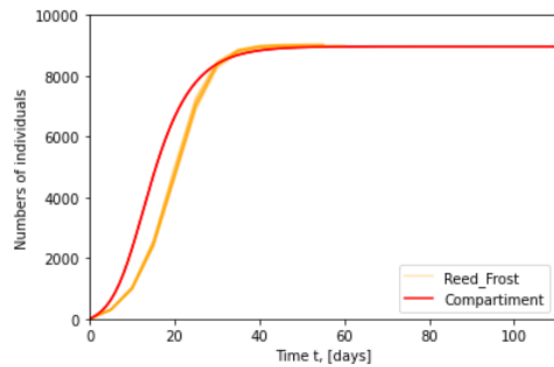
(a) *SIR* par compartiments



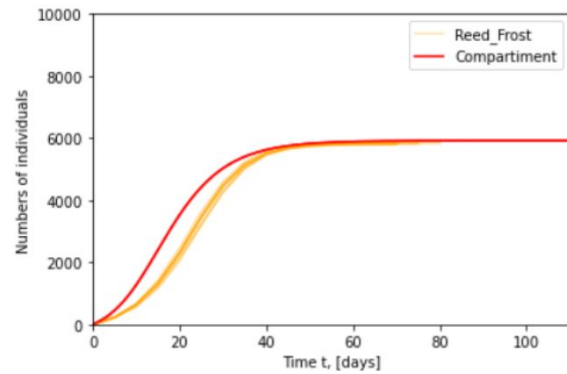
(b) Reed-Frost

FIG. A.2 : Évolution de $S(t)$, $I(t)$ et $R(t)$ dans les deux modèles

Les figures suivantes fournissent une comparaison de la taille de l'épidémie dans les deux modèles pour le cas sans vaccination et le cas avec vaccination.



(a) Sans vaccination



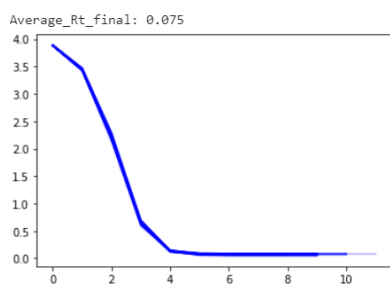
(b) Avec 25% de population vaccinée

FIG. A.3 : Comparaison de la taille finale de l'épidémie dans les deux modèles

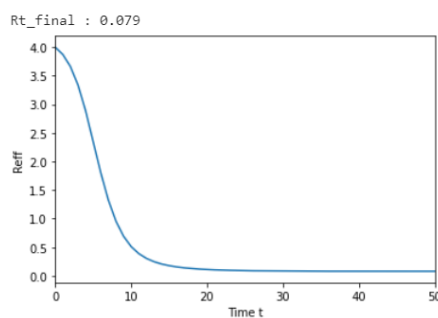
Annexe B

Taux de reproduction effectif : cas de plusieurs sources

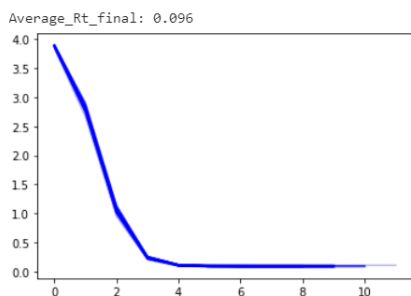
Nous reprenons ici les simulations de l'évolution du taux reproduction effectif pour les modèles par graphe de Reed-Frost et les modèles par compartiments de Kermack-McKendrick avec plusieurs personnes ($I_0 = 537$) contaminées au temps $t = 0$.



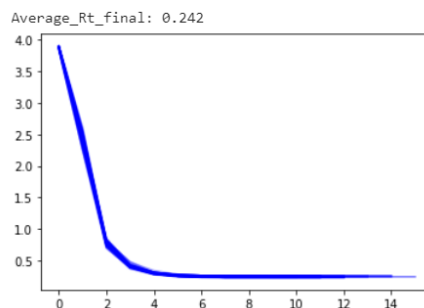
(a) Réseau Erdős-Rényi



(b) Modèle Kermack-McKendrick



(a) Réseau Barabási-Albert



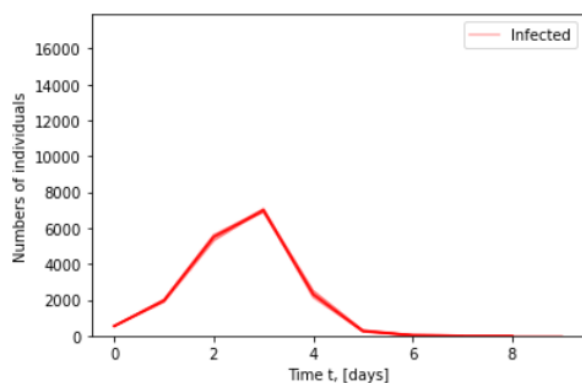
(b) Réseau réel

Annexe C

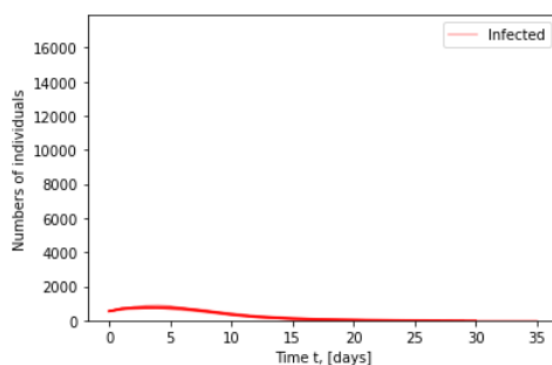
Effet des bulles : cas avec plusieurs sources

Les graphes suivants représentent l'évolution du nombre de contaminés dans une population avant et après la mise en place des bulles de taille 8 identique à ceux utiliser au chapitre portant sur les bulles.

La différence étant faite sur le nombre de cas contaminés au départ, ici $I_0 = 537$



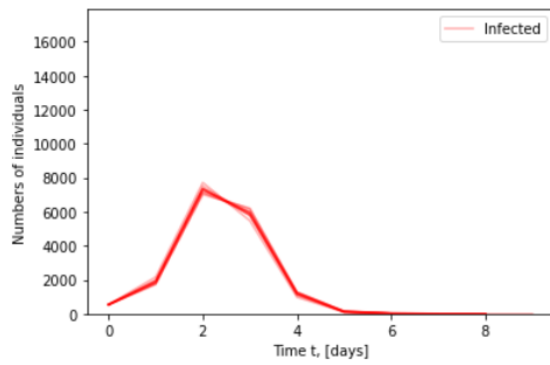
(a) Réseau sans Bulles



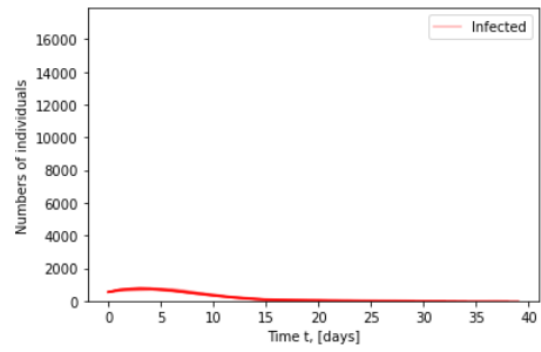
(b) Réseau avec Bulles

FIG. C.1 : Effet des bulles dans le réseau d'Erdős-Rényi

Annexe C. Effet des bulles : cas avec plusieurs sources

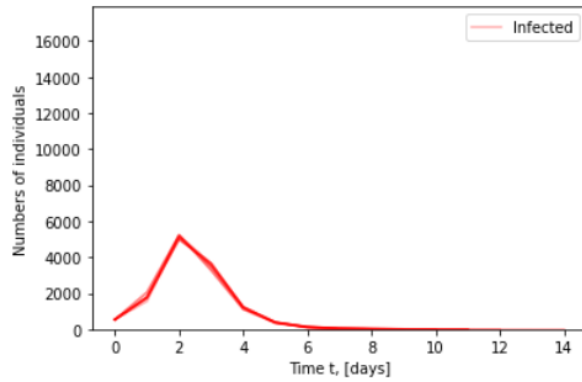


(a) Réseau sans Bulles

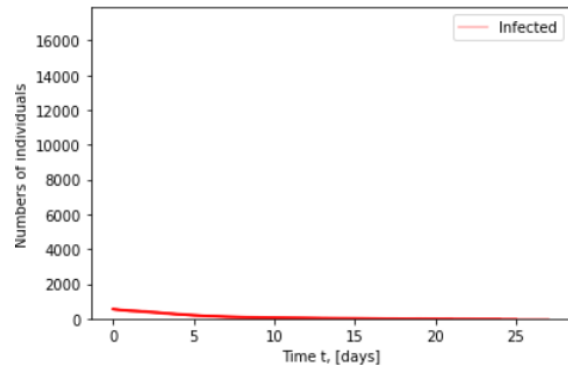


(b) Réseau avec Bulles

FIG. C.2 : Effet des bulles dans le réseau de Barabási-Albert



(a) Réseau sans Bulles



(b) Réseau avec Bulles

FIG. C.3 : Effet des Bulles dans le graphe réel

Annexe D

Résultat d'estimation de R_{eff} : modèle Corri-Ferguson

Les quatre tableaux suivants décrivent respectivement les résultats retournés par le modèle d'estimation du R_{eff} d'Anne Cori pour le modèle par compartiments *SIR* (tableau 1), le modèle par graphe de Reed-Frost pour les graphes : Erdős-Rényi (tableau 2), Barabási-Albert (tableau 3) et le graphe réel (tableau 4).

Mean(R)	Std(R)	Quantile.0.02	Quantile.0.05	Quantile.0.25	Median(R)	Quantile.0.95	Quantile.0.97
3.997	0.037	3.924	3.935	3.971	3.997	4.058	4.07
2.796	0.025	2.748	2.755	2.779	2.796	2.837	2.845
1.98	0.017	1.946	1.952	1.969	1.98	2.009	2.015
1.435	0.013	1.41	1.414	1.426	1.435	1.456	1.46
1.057	0.01	1.038	1.041	1.05	1.057	1.074	1.077
0.777	0.008	0.761	0.763	0.771	0.777	0.79	0.793
0.566	0.007	0.553	0.555	0.561	0.566	0.577	0.579
0.414	0.006	0.403	0.405	0.41	0.414	0.424	0.426
0.313	0.005	0.303	0.304	0.309	0.313	0.322	0.323
0.25	0.005	0.24	0.241	0.246	0.25	0.258	0.26
0.214	0.005	0.204	0.205	0.21	0.214	0.222	0.224
0.197	0.006	0.186	0.188	0.193	0.197	0.207	0.209
0.194	0.007	0.181	0.183	0.19	0.194	0.205	0.208
0.201	0.008	0.185	0.188	0.195	0.201	0.214	0.217
0.214	0.01	0.195	0.198	0.207	0.214	0.231	0.234
0.232	0.012	0.209	0.212	0.224	0.232	0.253	0.257
0.253	0.015	0.223	0.228	0.242	0.252	0.278	0.284
0.273	0.019	0.237	0.243	0.26	0.273	0.305	0.311
0.295	0.023	0.251	0.258	0.279	0.294	0.333	0.341
0.314	0.027	0.262	0.27	0.295	0.313	0.36	0.37
0.332	0.032	0.271	0.28	0.309	0.331	0.387	0.398
0.349	0.038	0.279	0.289	0.323	0.348	0.414	0.427
0.364	0.044	0.283	0.295	0.333	0.362	0.439	0.454
0.375	0.05	0.283	0.296	0.34	0.372	0.461	0.479
0.387	0.057	0.283	0.298	0.347	0.384	0.485	0.506
0.399	0.065	0.282	0.298	0.353	0.395	0.511	0.535
0.403	0.072	0.274	0.292	0.352	0.399	0.529	0.557
0.417	0.082	0.273	0.292	0.36	0.412	0.561	0.592
0.434	0.092	0.272	0.294	0.368	0.427	0.596	0.633
0.434	0.102	0.257	0.281	0.362	0.426	0.615	0.657
0.442	0.114	0.247	0.272	0.36	0.432	0.644	0.692
0.464	0.129	0.247	0.274	0.372	0.452	0.694	0.748
0.473	0.143	0.236	0.265	0.371	0.459	0.73	0.791
0.52	0.165	0.25	0.282	0.402	0.503	0.817	0.889
0.503	0.178	0.217	0.25	0.374	0.482	0.826	0.907
0.45	0.184	0.165	0.196	0.317	0.426	0.789	0.876
0.442	0.198	0.144	0.174	0.298	0.413	0.81	0.906
0.415	0.208	0.113	0.142	0.263	0.381	0.805	0.91
0.378	0.218	0.071	0.103	0.218	0.337	0.793	0.905

Mean(R)	Std(R)	Quantile.0.02	Quantile.0.05	Quantile.0.25	Median(R)	Quantile.0.95	Quantile.0.97
4.002	0.046	3.912	3.926	3.971	4.002	4.079	4.094
3.309	0.035	3.24	3.251	3.285	3.309	3.367	3.378
2.568	0.026	2.518	2.526	2.55	2.568	2.61	2.618
2.711	0.022	2.668	2.675	2.696	2.711	2.748	2.755
2.814	0.02	2.775	2.781	2.801	2.814	2.848	2.854
2.735	0.018	2.7	2.705	2.723	2.735	2.765	2.771
2.517	0.016	2.487	2.492	2.507	2.517	2.543	2.548
2.156	0.012	2.132	2.136	2.148	2.156	2.177	2.181
1.973	0.011	1.952	1.955	1.966	1.973	1.99	1.993
1.87	0.009	1.851	1.854	1.863	1.87	1.885	1.888
1.679	0.008	1.663	1.666	1.673	1.679	1.692	1.695
1.522	0.007	1.508	1.51	1.517	1.522	1.534	1.536
1.376	0.006	1.363	1.365	1.371	1.376	1.386	1.388
1.149	0.006	1.138	1.14	1.145	1.149	1.158	1.16
0.994	0.005	0.985	0.986	0.991	0.994	1.002	1.004
0.837	0.004	0.829	0.83	0.834	0.837	0.844	0.846
0.715	0.004	0.707	0.709	0.713	0.715	0.722	0.723
0.624	0.004	0.617	0.618	0.622	0.624	0.63	0.632
0.494	0.003	0.487	0.488	0.491	0.494	0.499	0.501
0.358	0.003	0.352	0.353	0.356	0.358	0.363	0.364
0.334	0.003	0.328	0.329	0.332	0.334	0.339	0.34
0.311	0.003	0.304	0.305	0.308	0.311	0.316	0.317
0.287	0.004	0.281	0.282	0.285	0.287	0.293	0.295
0.23	0.004	0.223	0.224	0.228	0.23	0.236	0.237
0.101	0.003	0.096	0.097	0.099	0.101	0.106	0.107
0.108	0.003	0.102	0.103	0.106	0.108	0.114	0.115
0.115	0.004	0.107	0.108	0.112	0.115	0.121	0.123
0.126	0.005	0.116	0.117	0.122	0.126	0.134	0.136
0.129	0.007	0.116	0.118	0.124	0.129	0.14	0.143
0.065	0.007	0.052	0.054	0.06	0.064	0.076	0.078
0.082	0.009	0.065	0.067	0.075	0.081	0.097	0.1
0.092	0.012	0.07	0.073	0.083	0.091	0.112	0.116
0.102	0.015	0.074	0.078	0.092	0.102	0.129	0.135
0.107	0.021	0.069	0.075	0.092	0.106	0.145	0.153

Mean(R)	Std(R)	Quantile.0.02	Quantile.0.05	Quantile.0.25	Median(R)	Quantile.0.95	Quantile.0.97
3.997	0.046	3.907	3.922	3.966	3.997	4.073	4.088
3.295	0.035	3.227	3.238	3.272	3.295	3.353	3.364
2.554	0.025	2.504	2.512	2.536	2.553	2.595	2.603
2.684	0.022	2.641	2.648	2.67	2.684	2.721	2.728
2.781	0.02	2.742	2.748	2.768	2.781	2.814	2.82
2.695	0.018	2.66	2.665	2.683	2.695	2.724	2.73
2.472	0.015	2.443	2.447	2.462	2.472	2.498	2.502
2.119	0.012	2.095	2.099	2.111	2.119	2.139	2.143
1.95	0.01	1.93	1.933	1.943	1.95	1.967	1.97
1.857	0.009	1.839	1.842	1.851	1.857	1.873	1.875
1.676	0.008	1.66	1.662	1.67	1.676	1.689	1.692
1.523	0.007	1.509	1.511	1.518	1.523	1.535	1.537
1.378	0.006	1.366	1.368	1.374	1.378	1.389	1.391
1.152	0.006	1.141	1.143	1.148	1.152	1.161	1.162
0.997	0.005	0.987	0.989	0.994	0.997	1.005	1.007
0.837	0.004	0.829	0.83	0.834	0.837	0.845	0.846
0.714	0.004	0.706	0.707	0.711	0.714	0.721	0.722
0.622	0.004	0.614	0.616	0.619	0.622	0.628	0.629
0.491	0.003	0.484	0.486	0.489	0.491	0.497	0.498
0.355	0.003	0.349	0.35	0.353	0.355	0.36	0.361
0.332	0.003	0.326	0.327	0.33	0.332	0.338	0.339
0.311	0.003	0.304	0.305	0.309	0.311	0.316	0.317
0.29	0.004	0.283	0.284	0.287	0.29	0.296	0.297
0.234	0.004	0.226	0.227	0.231	0.234	0.24	0.241
0.105	0.003	0.099	0.1	0.103	0.105	0.11	0.11
0.111	0.003	0.104	0.105	0.109	0.111	0.117	0.118
0.116	0.004	0.108	0.109	0.113	0.116	0.123	0.124
0.124	0.005	0.114	0.116	0.121	0.124	0.133	0.135
0.123	0.007	0.11	0.112	0.118	0.123	0.134	0.137
0.048	0.006	0.037	0.039	0.044	0.048	0.058	0.06
0.064	0.008	0.049	0.051	0.058	0.064	0.078	0.081
0.079	0.011	0.059	0.062	0.071	0.078	0.098	0.102
0.104	0.016	0.075	0.079	0.093	0.103	0.131	0.137
0.144	0.026	0.097	0.104	0.126	0.143	0.19	0.201
0.174	0.041	0.103	0.113	0.145	0.171	0.247	0.263
0.256	0.064	0.146	0.16	0.21	0.25	0.369	0.395
0.31	0.083	0.17	0.188	0.251	0.303	0.458	0.493
0.365	0.105	0.188	0.21	0.289	0.355	0.553	0.598
0.387	0.129	0.177	0.202	0.294	0.372	0.62	0.677

Mean(R)	Std(R)	Quantile.0.02	Quantile.0.05	Quantile.0.25	Median(R)	Quantile.0.95	Quantile.0.97
4.007	0.047	3.915	3.93	3.975	4.007	4.085	4.1
3.342	0.036	3.271	3.283	3.318	3.342	3.402	3.414
2.611	0.027	2.559	2.567	2.593	2.611	2.655	2.663
2.749	0.023	2.703	2.711	2.733	2.748	2.787	2.794
2.844	0.021	2.803	2.81	2.83	2.844	2.878	2.885
2.773	0.019	2.736	2.742	2.76	2.773	2.803	2.809
2.581	0.016	2.55	2.555	2.57	2.581	2.608	2.613
2.224	0.013	2.198	2.202	2.215	2.224	2.246	2.25
1.835	0.011	1.815	1.818	1.828	1.835	1.853	1.856
1.597	0.009	1.58	1.583	1.591	1.597	1.612	1.615
1.312	0.007	1.297	1.3	1.307	1.312	1.324	1.326
1.112	0.006	1.1	1.102	1.108	1.112	1.123	1.125
0.973	0.006	0.961	0.963	0.969	0.973	0.982	0.984
0.803	0.005	0.793	0.795	0.8	0.803	0.811	0.813
0.679	0.005	0.669	0.671	0.675	0.679	0.686	0.688
0.618	0.005	0.609	0.61	0.615	0.618	0.625	0.627
0.569	0.004	0.56	0.562	0.566	0.569	0.576	0.578
0.533	0.005	0.525	0.526	0.53	0.533	0.541	0.542
0.459	0.004	0.45	0.451	0.456	0.459	0.466	0.467
0.357	0.004	0.348	0.35	0.354	0.357	0.364	0.365
0.355	0.005	0.346	0.348	0.352	0.355	0.362	0.364
0.353	0.005	0.344	0.345	0.35	0.353	0.361	0.363
0.359	0.006	0.348	0.35	0.355	0.359	0.368	0.37
0.338	0.006	0.326	0.328	0.334	0.338	0.348	0.35
0.264	0.006	0.253	0.254	0.26	0.264	0.275	0.277
0.285	0.007	0.271	0.274	0.28	0.285	0.297	0.299
0.302	0.008	0.287	0.289	0.297	0.302	0.316	0.318
0.33	0.01	0.311	0.314	0.323	0.33	0.346	0.349
0.338	0.011	0.316	0.319	0.33	0.338	0.357	0.361
0.294	0.012	0.27	0.274	0.286	0.294	0.315	0.319
0.319	0.014	0.291	0.296	0.309	0.319	0.343	0.348
0.332	0.016	0.301	0.306	0.321	0.332	0.36	0.365
0.352	0.019	0.316	0.321	0.339	0.351	0.383	0.389
0.349	0.022	0.308	0.314	0.334	0.348	0.385	0.393
0.296	0.023	0.253	0.259	0.28	0.295	0.335	0.342
0.326	0.027	0.276	0.283	0.307	0.325	0.371	0.38
0.35	0.03	0.293	0.302	0.329	0.349	0.402	0.412
0.387	0.036	0.319	0.329	0.362	0.386	0.449	0.461
0.427	0.044	0.345	0.357	0.396	0.425	0.501	0.517
0.433	0.05	0.34	0.353	0.398	0.431	0.518	0.537
0.506	0.06	0.396	0.412	0.465	0.504	0.608	0.63
0.577	0.069	0.45	0.469	0.529	0.575	0.695	0.72
0.667	0.081	0.518	0.54	0.611	0.664	0.805	0.835
0.69	0.09	0.525	0.549	0.628	0.686	0.844	0.877
0.662	0.094	0.491	0.516	0.596	0.657	0.823	0.857
0.608	0.093	0.44	0.464	0.543	0.603	0.768	0.802
0.539	0.09	0.377	0.4	0.476	0.534	0.694	0.728
0.477	0.089	0.319	0.341	0.414	0.471	0.631	0.665
0.376	0.084	0.229	0.249	0.316	0.369	0.523	0.557

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl