

DEPARTEMENT DES MATHEMATIQUES

**Master Mathématique et Application au Calcul Scientifique
(MACS)**

MEMOIRE DE FIN D'ETUDES

**Pour l'obtention du Diplôme de Master Sciences et Techniques
(MST)**

**Les mathématiques des épidémies et
modélisation de la pandémie de COVID-19**

Réalisé par : MEDOUAR Fouad

Encadré par : Pr. EL KHOUKHI Fatima

Soutenu 16 juillet 2021 Devant le jury composé de :

- | | |
|---------------------------|---------------------|
| - Pr. AMMOR Ouafae | FST de Fès |
| - Pr. EL BARAKA Azzeddine | FST de Fès |
| - Pr. ELGOUMY Badreddine | INSA EUROMED de Fès |
| - Pr. EL KHOUKHI Fatima | FLSH de Meknès |
| - Pr. EZZAKI Fatima | FST de Fès |
| - Pr. HILALI Abdelmajid | FST de Fès |

Année Universitaire 2020 / 2021

FACULTE DES SCIENCES ET TECHNIQUES DE FES

☒ B.P. 2202 – Route d'Imouzzer – FES

Dédicaces

Je dédie ce modeste travail à :

A ma mère, qui m'a donnée la vie, la tendresse et le courage pour réussir.

A mon père, qui présente pour moi le symbole de l'amour et du respect.

A mon frère.

A tous mes Professeurs pour leur soutien au cours de toutes mes années d'études.

A tous ceux qui sont proches de mon coeur, et dont je n'ai pas cité le nom.

Remerciements

Je remercie mon Dieu qui m'a donné la volonté, la patience, et surtout la santé durant toutes mes années d'étude.

Mes profonds remerciements à mes premiers fans, mes parents pour leur soutien quotidien infaillible, merci à leur enthousiasme débordant qui a été pour moi pilier fondateur de mon action, sans eux je n'aurais jamais pu réaliser ce travail.

J'exprime ma profonde gratitude à mon encadrante Pr. EL KHOUKHI Fatima pour son soutien inoubliable et je la remercie encore pour le sujet qu'il m'a proposé et pour son suivi permanent enrichi de beaucoup d'encouragement, ses remarques et suggestions sans lesquelles ce mémoire n'aurait pas lieu.

Je tiens aussi à remercier infiniment les membres de jury, les Professeurs Mme. AMMOR Ouafae, Mr. EL BARAKA Azzeddine, Mr. ELGOUMY Badreddine, Mme. EZZAKI Fatima et Mr. HILALI Abdelmajid, pour le privilège qu'ils m'ont accordé en acceptant d'évaluer ce travail et par la suite d'apporter leurs précieux jugements, qui seront certainement, des outils efficaces et pertinents pour l'amélioration de mon projet de fin d'études Master.

Par ailleurs, puisque l'occasion se présente ici, je remercie également les autres membres de ma famille, tout spécialement mon frère, mes soeurs ainsi que mon ami Mohamed, qui m'ont toujours soutenu, même à distance.

Je tiens enfin à remercier tous ceux qui ont contribué d'une façon ou d'une autre à la réalisation de ce travail.

Table des matières

INTRODUCTION GENERALE	9
1 Modélisation mathématique en épidémiologie	11
1.1 Introduction	11
1.2 Branches de l'épidémiologie et ses champs d'application	13
1.3 Analyse statistique et modélisation : modélisation statistique	14
1.3.1 Analyse statistique	16
1.3.2 Analyse statistique et Recherche opérationnelle	17
1.3.3 Application en épidémiologie : aide à la décision sanitaire	17
1.4 Quelques modèles mathématiques en épidémiologie	19
1.4.1 Modèle de Bernoulli	19
1.4.2 Modèle logistique de Verhulst	22
1.4.3 Les modèles compartimentaux déterministes	23
1.4.3.1 Modèle déterministe simple de Hamer	23
1.4.3.2 Modèle déterministe général de Kermack et MC Kendrick	25
1.4.3.3 Autres extensions du modèle SIR	27
1.4.4 Modèle de Ross	27
1.4.5 Les modèles stochastiques	29
1.4.5.1 Modèle classique de Reed-Frost	29
1.4.5.2 Modèle SI Stochastique	30
1.5 CONCLUSION	31
2 Modélisation mathématique du COVID-19	32
2.1 Introduction	32
2.2 Paramètres de la maladie	33
2.2.1 Taux de reproduction	33
2.2.1.1 Techniques pour le Calcul du taux de reproduction	34

2.2.2	Période d'incubation	37
2.2.3	Temps de génération, Intervalle sériel	38
2.2.4	Taux de létalité	38
2.2.4.1	Calcul du ratio de létalité réel (IFR)	39
2.2.4.2	Calcul du ratio de létalité apparent (CFR)	39
2.2.4.3	Calcul du taux de létalité apparent (CFR) pendant une épidémie	40
2.3	Quelques modèles mathématiques appliqués au COVID-19	43
2.3.1	Modèles déterministes	43
2.3.2	Modèles stochastiques	48
2.4	CONCLUSION	49
3	Etude de la propagation de la pandémie COVID-19 : Le cas du MAROC	50
3.1	Introduction	50
3.2	La situation épidémique au Maroc	51
3.3	Simulation 1 : Modèle logistique	51
3.3.1	Estimation des paramètres	51
3.3.2	Résultats et Discussion	53
3.4	Simulation 2 : Modèle SIR	57
3.4.1	Estimation des paramètres	57
3.4.2	Résultats et Discussion	58
3.5	CONCLUSION	62
	CONCLUSION GENERALE	63
	Bibliographie	66

Table des figures

1.1	Diagramme de transmission du modèle SI.	24
1.2	Modèle SI - Evolution des populations S et I au cours du temps.	24
1.3	Diagramme de transmission du modèle SIR.	25
1.4	Modèle SIR - Evolution des populations S, I et R au cours du temps.	26
1.5	Modèle de Ross.	28
2.1	Quelques estimations du taux de reproduction du COVID-19 dans plusieurs pays au début de l'épidémie	37
2.2	Le taux de létalité pour le COVID-19 en Chine dans la période entre le 1 ^{er} janvier et 20 février 2020	42
2.3	Variation du taux de létalité au Maroc entre 14 Mars 2020 et 25 Avril 2021	42
2.4	Diagramme de transmission du modèle SEAIR.	46
2.5	Diagramme de transmission du modèle SIQ.	48
2.6	Comparaison du comportement déterministe et stochastique avec les cas quotidiens de COVID-19 signalés au Maroc.	49
3.1	Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 9 juin 2020.	54
3.2	Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 31 Octobre 2020.	55
3.3	Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 28 Janvier 2021.	56
3.4	Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars au 31 Mars 2021.	57
3.5	Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 9 juin 2020.	59
3.6	Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 31 Octobre 2020.	60

3.7	Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 28 Janvier 2021.	61
3.8	Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 31 Mars 2021.	62

Liste des tableaux

2.1	Période d'incubation pour différentes maladies.	38
2.2	Le taux de létalité pour le COVID-19 en Chine dans la période entre le 1 ^{er} janvier et 20 février 2020	41
3.1	Les résultats des paramètres estimés du modèle logistique pour différentes périodes à partir de 12 Mars 2020 jusqu'à 31 Mars 2021.	53
3.2	Les résultats des paramètres estimés du modèle SIR pour différentes périodes à partir de 12 Mars 2020 jusqu'à 31 Mars 2021.	59
3.3	Quelques résultats supplémentaires concernant la fin de l'épidémie au Maroc. .	62

INTRODUCTION GENERALE

Généralement, les modèles mathématiques présentent une traduction simplifiée de la réalité formalisant un phénomène complexe, ils permettent d'en étudier ses différents paramètres et les relations qui existaient entre eux. Ainsi, en générant et testant des hypothèses, ils conduisent à une première compréhension des systèmes étudiés. Toute modélisation est, de par sa nature, une abstraction et a surtout vocation à susciter la réflexion, et éventuellement, à éclairer la prise de décision.

En dehors de leur intérêt descriptif et de génération d'hypothèses, ces modèles mathématiques sont aussi, voir surtout, utilisés pour anticiper des événements dans des situations très diverses, et particulièrement présents depuis longtemps dans l'épidémiologie des maladies transmissibles, à travers les différents modèles d'équations différentielles ou probabilistes. Ils permettent par exemple d'anticiper la taille d'une épidémie, comme celle du COVID-19 (CoronaVirus Infectious Disease 2019).

Depuis le début de la pandémie, une multitude de données sont accumulées chaque jour, permettant de modéliser l'évolution des nouveaux cas de COVID-19 et d'en prédire la propagation. La modélisation mathématique a permis, dans le cadre de la crise sanitaire actuelle, de définir les bons indicateurs de suivi l'épidémie, de faire des prédictions concernant sa propagation et son évolution et d'appuyer la prise de décision publique. Par conséquent, le modèle mathématique est un outil indispensable au service de la lutte contre la pandémie COVID-19, au niveau International et aussi au niveau du Maroc. Notre objectif dans le présent mémoire, est doublé. D'une part, nous visons dresser un état des lieux de la connaissance et une synthèse bibliographique portant sur la modélisation mathématique en épidémiologie puis pour le COVID-19, en présentant un ensemble de modèles mathématiques bien connus dans la littérature du domaine. Et d'autre part, mener une étude expérimentale comparative entre deux modèles mathématiques classiques, à savoir le modèle de croissance logistique et le modèle SIR, à base de la simulation numérique en utilisant le logiciel MATLAB, pour étudier la propagation de COVID-19 au Maroc et estimer quelques paramètres associés à l'évolution de cette maladie au niveau national, suivant différentes périodes à partir du début de l'épidémie jusqu'à 31 Mars 2021.

Le manuscrit est divisé en trois chapitres :

- Dans le premier chapitre, nous donnons un aperçu sur la théorie et les champs d'applications de l'épidémiologie mathématique, aussi bien que la modélisation statistique et les principaux paramètres utilisés dans la modélisation mathématique des épidémies. De plus, nous présentons un ensemble de modèles classiques les plus connus dans la littérature de l'épidémiologie mathématique.
- Dans le deuxième chapitre, nous présentons la pandémie actuelle COVID-19, son mécanisme de transmission et ses paramètres, puis nous mettons l'accent sur quelques travaux portant sur sa modélisation mathématique.
- Dans le troisième chapitre, nous exposons l'étude expérimentale menée et nous discutons et interprétons les résultats obtenus, en testant sur deux modèles mathématiques bien connus modélisant la pandémie actuelle du COVID-19, le cas du Maroc.

Modélisation mathématique en épidémiologie

1.1 Introduction

L'épidémiologie est une branche interdisciplinaire de la science, elle utilise différents concepts mathématiques afin d'étudier des phénomènes biologiques, notamment les épidémies [1]. Il s'agit d'une science ancienne, qui s'est développée avec les grandes épidémies telles que la variole, la grippe espagnole, etc... et aussi avec les possibilités de simulation numérique offertes par les progrès informatiques. L'Organisation Mondiale de Santé (OMS) définit l'épidémiologie comme étant « l'étude de la distribution et les déterminants des états de santé et des maladies dans les populations humaines ainsi que les influences qui déterminent cette distribution ».

Le terme « épidémiologie » est un mot d'origine grecque : « épi » veut dire « au sujet de », « démos » veut dire « population » et « logos » veut dire « étude ». On peut donc traduire ce terme comme voulant dire « étude de ce qui touche une population ». Nous définissons aujourd'hui l'épidémiologie comme l'étude :

- De la répartition de maladies ou états de santé dans une population donnée,
- des déterminants qui influent sur l'occurrence d'une maladie,
- des interventions, ce qui va pouvoir guider les politiques de santé publique.

On distingue généralement trois termes, *épidémie*, *endémie* et *pandémie*, cette distinction se fait essentiellement au niveau de la proportion de personnes atteintes par l'agent infectieux :

- *Une épidémie* désigne une augmentation inattendue du nombre de cas d'une maladie dans une zone géographique spécifique. La fièvre jaune, la variole, la rougeole sont des exemples d'épidémies. Notamment, une maladie épidémique ne doit pas nécessairement être contagieuse. Par exemple, la fièvre du Nil occidental et l'augmentation rapide des

taux d'obésité sont également considérées comme des épidémies. En termes plus généraux, les épidémies peuvent faire référence à une maladie avec des taux nettement supérieurs à la fréquence attendue dans une communauté ou une région.

- *Une endémie* est une épidémie qui est constamment présente mais limitée à une région particulière. Cela rend sa propagation prévisible. Le paludisme, par exemple, est considéré comme endémique dans certains pays et régions.
- *Une pandémie* est déclarée par l'OMS, lorsque la croissance d'une maladie est exponentielle. Cela signifie que le taux de croissance monte en flèche et que les cas augmentent chaque jour plus que la veille. En étant déclaré pandémie, le virus n'a rien à voir avec l'immunité de la population ou la gravité de la maladie. Cela signifie qu'un virus couvre une vaste zone, affectant plusieurs pays et populations

Ainsi la modélisation mathématique d'une épidémie consiste à en créer une représentation mathématique, dont le rôle du modélisateur est d'estimer les valeurs des paramètres qui feront correspondre au mieux « l'épidémie modélisée » à l'épidémie réelle, elle suppose donc des « allers-retours » entre le calcul mathématique et les données issues du terrain. La modélisation épidémiologique vise plusieurs objectifs :

- Repérage des mécanismes par lesquels les maladies se propagent ;
- Évaluation de l'incidence réelle de l'épidémie dans la population, ce qui permet de mieux estimer le niveau d'immunisation de la population ;
- Prévision des besoins en ressources hospitalières pour adapter les mesures sanitaires telles que le confinement, la fermeture des écoles, commerces etc ...
- Estimation de l'intérêt des différentes mesures dans la limitation de la propagation de l'épidémie.

Cette étape de modélisation mathématique de l'épidémiologie des maladies transmissibles constitue un véritable outil de santé publique. Elle permet de mettre à l'épreuve, sans perdre du temps ou du frais, les mesures de lutte contre la propagation du maladie qui sont envisagées. Le modèle n'est néanmoins pas la réalité et n'est pas supposé la reproduire en totalité. Il doit donc reproduire au mieux les caractéristiques du phénomène étudié en fonction des objectifs fixés pour le cadre de l'étude.

L'étape de modélisation est l'étape la plus délicate, la plus longue et souvent la plus périlleuse. En effet, il faut essayer de bien comprendre le problème initial (problème réel) pour tenter de proposer un modèle adapté. La première tentative proposée ne répond que très rarement aux attentes, s'en suit alors plusieurs modifications, jusqu'à parvenir à un modèle final qui regroupe et traduit un maximum de contraintes que le phénomène réel, d'ou la notion « allers-retours » entre calculs mathématiques et données réels. Si cette étape est négligée, si les contraintes ne sont pas bien posées, on aboutit alors à une formulation mathématique qui ne correspond pas au problème étudié. La résolution du problème mathématique fourni alors une

solution non adéquate au problème concret. Enfin, si le problème est bien posé, l'étape suivante consiste alors à résoudre ce problème, c'est-à-dire à analyser le modèle dans le but de comprendre, de prédire et d'agir.

La modélisation est aujourd'hui un domaine de recherche à part entière, dont les capacités ont été décuplées par la formidable puissance de calcul permise par les ordinateurs actuels. Dans le contexte de la pandémie de COVID-19 (CoronaVirus Infectious Disease 2019), les épidémiologistes développent, testent et ajustent des modèles pour simuler la propagation de cette maladie infectieuse ; il s'agit de mieux la comprendre et d'optimiser les interventions visant à la contrôler.

1.2 Branches de l'épidémiologie et ses champs d'application

On distingue principalement trois branches de l'épidémiologie, descriptive, étiologique et évaluative :

- L'épidémiologie descriptive dont l'objectif est d'identifier les problèmes de santé ; elle a pour but l'étude de la fréquence et de la répartition dans le temps et l'espace des états de santé de groupes d'individus.
- L'épidémiologie explicative (ou analytique ou étiologique) dont l'objectif est de connaître leurs causes ;
- L'épidémiologie évaluative dont l'objectif est d'évaluer les solutions apportées à ces problèmes (actions de santé).

Dans chacune des trois branches, la récupération des données s'appuie sur une méthodologie rigoureuse, c'est-à-dire sur l'élaboration d'un protocole d'enquête qui doit respecter des règles strictes. Le principal critère qui permet de classer les différents types d'enquêtes est le moment où sont recueillies les données. Au niveau pratique, on distingue deux champs d'application :

- L'épidémiologie de population ou « classique » qui concerne les études épidémiologiques en population générale. Elles ont pour objectif l'élaboration de stratégies en santé publique. Et sont fondées sur la description et la mesure des phénomènes de santé dans une population
- L'épidémiologie clinique ou « recherche clinique », où, il s'agit des études épidémiologiques sur des populations de patients en vue d'améliorer les connaissances et techniques médicales en vue de leurs applications à l'activité clinique courante.(permettant l'amélioration des connaissances et stratégies médicales, ainsi que l'évaluation de leurs applications dans le cadre d'activités cliniques).

1.3 Analyse statistique et modélisation : modélisation statistique

La modélisation statistique consiste essentiellement à définir des outils appropriés pour modéliser des données observées, en tenant compte de leur nature aléatoire. Les méthodes de modélisation statistique sont, en fait, très nombreuses [2][3].

La croissance considérable des masses de données enregistrées dans différents secteurs (internet, biologie à haut débit, marketing...), le besoin d'exploiter ces données sur le plan statistique, ainsi que les outils modernes de calcul ont donné naissance ces dernières années à de nombreux modèles statistiques, de plus en plus sophistiquées en temps calcul. Nous citons principalement :

- ▷ *Le modèle linéaire (gaussien) de base* : À la fois le plus simple, le plus ancien et le plus connu des modèles statistiques, il englobe essentiellement la régression linéaire, l'analyse de variance et l'analyse de covariance. Dans ce modèle, les variables explicatives ne sont pas aléatoires. Pour pouvoir être exploité pleinement, ce modèle nécessite l'hypothèse de normalité des erreurs, donc de la variable à expliquer.
- ▷ *Le modèle linéaire généralisé* : Il généralise le précédent à deux niveaux : d'une part, la loi des erreurs, donc de la variable réponse, n'est plus nécessairement gaussienne, mais doit appartenir à l'une des lois de la famille exponentielle ; d'autre part, la liaison linéaire entre l'espérance de la variable réponse et les variables explicatives se fait à travers une fonction particulière appelée fonction lien.
- ▷ *Les modèles non linéaires* : De façon très générale, il s'agit de modèles permettant d'expliquer la variable réponse (aléatoire) au moyen des variables explicatives (non aléatoires dans les modèles usuels), à travers une fonction quelconque, inconnue (on est donc en dehors du cadre du modèle linéaire généralisé).
- ▷ *Les modèles mixtes* : On désigne sous ce terme les modèles permettant d'expliquer la variable aléatoire réponse au moyen de diverses variables explicatives, certaines étant aléatoires (on parle en général de facteurs à effets aléatoires) et intervenant dans la modélisation de la variance du modèle, d'autres ne l'étant pas (on parle de facteurs à effets fixes) et intervenant dans la modélisation de la moyenne. On trouve ainsi des modèles linéaires gaussiens mixtes, des modèles linéaires généralisés mixtes et des modèles non linéaires mixtes.
- ▷ *Les modèles pour données répétées* : On appelle données répétées, ou données longitudinales, des données observées au cours du temps sur les mêmes individus (en général, il s'agit de personnes ou d'animaux suivis dans le cadre d'une expérimentation médicale ou biologique). De façon claire, il est nécessaire de prendre en compte dans ces modèles

une certaine dépendance entre les observations faites sur un même individu à différents instants.

- ▷ *Les modèles pour séries chronologiques* : Les séries chronologiques sont des observations, au cours du temps, d'une certaine grandeur représentant un phénomène économique, social ou autre. Ces données répétées et séries chronologiques ont en commun de rendre compte de l'évolution au cours du temps d'un phénomène donné, on notera que ces deux types de données ne sont pas réellement de même nature (dans une série chronologique, ce sont rarement des personnes ou des animaux que l'on observe).
- ▷ *L'analyse discriminante et la classification* : S'il est plus courant d'utiliser ces méthodes dans un contexte d'exploration des données plutôt que dans un contexte de modélisation, l'analyse discriminante et la classification peuvent tout de même être utilisées dans la phase de recherche d'un modèle permettant d'ajuster au mieux les données considérées. C'est en particulier le cas lorsque la variable réponse du modèle envisagé est de nature qualitative.
- ▷ *Les modèles par arbre binaire de régression et de classification* : Ces méthodes (plus connues sous le nom de CART, pour Classification And Regression Trees) consistent à découper une population en deux parties, en fonction de celle des variables explicatives et du découpage en deux de l'ensemble de ses valeurs ou modalités qui expliquent au mieux la variable réponse. On recommence ensuite sur chaque sous-population ainsi obtenue, ce qui permet de définir, de proche en proche, un arbre binaire et de classer les variables explicatives selon l'importance de leur liaison avec la variable réponse (on parle d'arbre de régression en présence d'une variable réponse quantitative et d'arbre de classification en présence d'une variable réponse qualitative).

Au niveau des modèles décrits précédemment, il y a presque toujours une variable privilégiée, en général appelée variable à expliquer, ou variable réponse, et notée Y (il s'agit d'une variable aléatoire). Le but est alors de construire un modèle permettant d'expliquer "au mieux" cette variable Y en fonction de variables explicatives observées sur le même échantillon [4].

Dans la pratique, un modèle statistique n'est réellement opérationnel que si l'on précise la loi de probabilité P_i (cas des variables aléatoire indépendantes et identiquement distribuées) ou les lois P (cas seulement indépendant). Après avoir ainsi précisé la loi de probabilité (ou la famille de lois de probabilité) du modèle, il reste d'abord à faire des tests, d'une part pour essayer de simplifier le modèle retenu, d'autre part pour tester la signification de ce dernier, ensuite à en estimer les paramètres. C'est tout ce travail, choix de la loi de probabilité ou de la famille des lois, tests, choix du modèle, estimation des paramètres du modèle retenu, validation du modèle qui constitue la modélisation statistique.

1.3.1 Analyse statistique

L'analyse statistique est une composante de l'analyse des données. Dans le contexte de l'informatique décisionnelle, l'analyse statistique implique la collecte et l'examen de tous les échantillons de données tirés d'un jeu de données. Dans le domaine des statistiques, un échantillon est une sélection représentative d'une population. On distingue quatre types d'analyse statistique [5] :

- ▷ *L'analyse descriptive* : répertorie et résume les valeurs de chaque variable dans un ensemble de données. Par exemple, si les répondants au sondage ont attribué une note de 1 à 10 à une question donnée, une analyse descriptive peut indiquer le nombre et le pourcentage de répondants pour chaque note, les notes moyenne et médiane, le mode ou la note la plus courante, ainsi qu'une mesure de tendance centrale telle que l'écart type. L'analyse descriptive permet de se familiariser avec un ensemble de données et à identifier les problèmes liés aux données, tels que les répondants qui n'ont fourni aucune évaluation.
- ▷ *L'analyse exploratoire* : Après la compréhension des données identifiées, la prochaine étape consiste à rechercher les relations entre les éléments de données. C'est appelé l'analyse exploratoire des données, et se concentre généralement sur corrélations entre variables. Par exemple, un ensemble de données montre une corrélation extrêmement élevée entre le nombre de caries d'un enfant et la taille de son vocabulaire. Toutefois, cela ne signifie pas que si on autorise cet enfant à avoir plus de caries, son vocabulaire augmentera également. Il peut y avoir d'autres facteurs qui déterminent les résultats, tels que l'âge, que nous n'avons pas dans notre jeu de données.
- ▷ *L'analyse inférentielle* : Pour développer l'indice de confiance des consommateurs, le Conference Board par exemple n'interroge pas chaque consommateur sur sa confiance dans l'économie. Il utilise une analyse inférentielle à tirer des conclusions sur les consommateurs sur la base de données provenant d'un échantillon plus petit de la population. Il est important de comprendre la méthode d'échantillonnage utilisée dans l'analyse inférentielle, car il est souvent possible de tirer des conclusions très différentes d'un même ensemble de données en sélectionnant différents échantillons. Comme beaucoup d'analyses inférentielles, l'indice de confiance du consommateur sélectionne un échantillon aléatoire de son ensemble de données afin que le résultat soit approximativement le même, quel que soit l'échantillon que nous choisissons.
- ▷ *L'analyse prédictive* : L'objectif de l'analyse prédictive est d'utiliser des données, prédire un résultat inconnu, et ensuite prendre des mesures en fonction de cette prédiction. Par exemple, les compagnies d'assurance utilisent des données telles que le sexe, l'âge, l'état matrimonial et le pointage de crédit pour prédire quels sont les clients les plus susceptibles d'avoir un accident. Ensuite, ils augmentent les taux d'assurance pour les

clients appartenant aux groupes à haut risque. Les analystes développent des modèles prédictifs en formant le modèle sur une partie de l'ensemble de données où le résultat est connu, puis en appliquant le modèle aux données restantes lorsque le résultat est inconnu.

L'analyse statistique peut être décomposée en cinq étapes [6] :

- ▷ Décrire la nature des données à analyser ;
- ▷ Explorer la relation entre les données et la population correspondante ;
- ▷ Créer un modèle pour synthétiser les relations entre les données et la population ;
- ▷ Prouver (ou refuser) la validité du modèle ;
- ▷ Se servir de l'analytique prédictif pour dérouler des scénarios qui guideront les actions à venir.

1.3.2 Analyse statistique et Recherche opérationnelle

La recherche opérationnelle intervient lorsqu'il s'agit de prendre des décisions. Les conséquences de ces décisions peuvent être plus ou moins bonnes. Le principal objet de la recherche opérationnelle consiste précisément à fournir des données qui permettent de retenir la meilleure décision, c'est-à-dire celle dont les effets seront les plus favorables. Mais il faut un critère de choix, ce qui signifie qu'on doit pouvoir, de façon objective, préférer une décision à une autre. Ce critère de choix n'est valable que si les conséquences probables de chacune des décisions peuvent être constatées numériquement, quantitativement. Mais il est alors nécessaire que les influences particulières qui affectent le problème et la solution puissent aussi être exprimées sous une forme quantitative.

L'emploi de la statistique est par conséquent indispensable, car dans toute recherche opérationnelle, il faut que les données numériques requises soient rassemblées selon les méthodes statistiques. La recherche opérationnelle est donc finalement impensable sans statistique. Grâce à celle-ci, les données quantitatives pour la formulation du problème peuvent être réunies. Ces données sont ensuite transformées en solutions à l'aide de méthodes scientifiques déterminées. Ces solutions sont ensuite comparées entre elles quant à l'objet à atteindre. L'enchaînement de ces données quantitatives par des méthodes de solution scientifiques constitue le modèle de recherche opérationnelle [7] [8].

1.3.3 Application en épidémiologie : aide à la décision sanitaire

L'épidémiologie descriptive comme déjà défini a pour objet de décrire les variations dans le temps, l'espace ou à travers différents groupes sociaux de l'état de santé et des facteurs

de risques de la population, à partir de données obtenues exhaustivement ou grâce à des échantillons de la population. La connaissance de l'état de santé d'une population s'exprime au moyen de données numériques appelées indicateurs de santé qui sont des variables quantitatives représentant certaines composantes de l'état de santé [9] [10].

Quand l'échantillon est constitué au hasard, il s'agit d'un échantillon aléatoire. Seul un échantillon aléatoire peut être réellement représentatif de la population ciblée, c'est-à-dire que les résultats obtenus sur cet échantillon peuvent être extrapolés à l'ensemble de la population. La démarche essentielle consiste à établir des taux en rapportant le nombre de personnes présentant une condition donnée à l'effectif de la population concernée, et cela pendant une certaine période de temps. Ces taux permettent d'étudier la fréquence des problèmes en fonction des caractéristiques des personnes (sexe, âge, profession...), de leur répartition géographique, de leur évolution dans le temps. Il s'agit ainsi de produire des connaissances, afin d'apprécier le niveau de santé de la population et de ses différents sous-groupes, de repérer des problèmes de santé émergents (surveillance épidémiologique), de formuler des hypothèses qui serviront ensuite à concevoir des enquêtes de type analytique, d'aider à choisir les moyens à mettre en oeuvre pour améliorer l'état de santé, notamment en termes d'actions de prévention.

L'épidémiologie descriptive est un outil au service de la politique de santé en général, qui s'applique en particulier au champ de la santé au travail. Il est plus souvent nécessaire de concevoir et de conduire des enquêtes épidémiologiques spécifiques, on parle d'enquêtes :

- Transversales : il s'agit d'investigations de courte durée destinées à appréhender les phénomènes présents au moment de l'enquête, c'est une photographie à un moment donné ;
- Prospectives : elles consistent à suivre un groupe de sujet afin d'étudier les phénomènes qui les affectent au cours du temps ;
- Rétrospectives : elles étudient des phénomènes antérieurs au moment de l'enquête.

En milieu de travail par exemple, la surveillance épidémiologique consiste à effectuer, de façon systématique et régulière, le recueil, le traitement statistique et l'interprétation de données démographiques, de données sur l'état de santé et d'informations décrivant l'activité professionnelle. Il s'agit de choisir des indicateurs de l'état de santé d'une population de travailleurs dans une entreprise, un secteur ou plusieurs secteurs d'activités, constituant une cohorte, et de les suivre pour détecter des phénomènes de santé « anormaux ». La surveillance épidémiologique est avant tout un outil d'alerte et a un objectif interventionnel. Elle peut, en identifiant des problèmes, donner lieu à des études épidémiologiques qui en préciseront l'ampleur et en rechercheront les causes et permettre la mise en place de mesures de maîtrise des risques. En particulier, un suivi épidémiologique peut conduire à l'investigation d'un « cluster », ou excès supposé de cas de maladies dans une population [11].

1.4 Quelques modèles mathématiques en épidémiologie

La modélisation mathématique des maladies infectieuses est un outil pour étudier la façon dont les maladies se propagent, prévoir la trajectoire future d'une éclosion et aider à orienter la planification de la santé publique et la lutte contre les maladies infectieuses. Les modèles utilisent des équations mathématiques pour estimer le nombre de cas de maladie qui pourraient survenir dans les semaines ou les mois à venir. Ils aident les chercheurs à simuler des possibilités réelles dans un environnement virtuel. Bien que les modèles ne puissent pas prédire ce qui se passera, ils peuvent nous aider à comprendre ce qui pourrait se passer dans certains scénarios. Cela peut nous aider à planifier et à agir pour obtenir le meilleur résultat possible.

Il est évident que toute modélisation est, de par sa nature, une abstraction et a surtout vocation à susciter la réflexion, et éventuellement, à éclairer la prise de décision. Il est important de préciser qu'aucun modèle ne peut appréhender toute la complexité de la propagation de cette maladie. La prédiction de l'évolution de la propagation du virus reste délicate alors que de nombreuses incertitudes subsistent à son sujet. Les modèles mathématiques ne prennent pas en compte tous les paramètres dont l'épidémie dépend mais ils donnent une idée sur la propagation dans le futur proche. Ils ciblent des populations humaines en interaction, leur fréquence, leur distribution dans le temps (et dans l'espace), et proposent des prévisions sur ces densités de populations, qu'elles soient saines, infectées ou guéries.

Un modèle épidémiologique se fonde sur deux concepts : les compartiments et les règles. Les compartiments divisent la population en divers états possibles par rapport à la maladie. Tandis que les règles spécifient la proportion des individus passant d'un compartiment à un autre [12] [13].

L'épidémie actuelle de coronavirus remet l'épidémiologie au premier plan. Une partie de cette discipline consiste à modéliser l'évolution de l'épidémie grâce à des outils mathématiques. Ces outils peuvent consister en un système d'équations aux dérivées partielles, de la théorie des graphes, des probabilités ou encore la science des données et analyse statistique. Dans le cas d'une représentation mathématique à l'aide d'équations différentielles, si ces dernières deviennent trop compliquées, des simulations apportent un éclairage sur le comportement de la maladie.

1.4.1 Modèle de Bernoulli

L'histoire de la modélisation mathématique en épidémiologie a commencé le 30 Avril 1760 par les travaux du mathématicien et physicien Daniel Bernoulli ¹, dans cette étude Bernoulli a

1. **Daniel Bernoulli** : (1700-1782) mathématicien et physicien et l'un des trois membres les plus importants de cette famille, fils du médecin et mathématicien Jean Bernoulli (1667-1748) et neveu de Jacques Bernoulli (1654-1705) connu pour ses travaux en probabilité.

développé un modèle mathématique pour l'estimation des avantages de l'inoculation contre la variole afin de diminuer la mortalité dans la population française.

Pour construire son modèle, Bernoulli a adopté plusieurs hypothèses, ces hypothèses sont tirées de [14].

- Un individu infecté par la variole a une probabilité p de mourir et une probabilité $1 - p$ de survivre, sans prendre en considération le facteur de l'âge des individus ;
- Un individu a une probabilité q d'être infecté dans l'année, et cela indépendamment de son âge (i.e. la probabilité qu'un individu soit infecté pendant le petit intervalle de temps dx entre l'âge x et l'âge $x + dx$ est $q \cdot dx$);
- Lorsqu'un individu survit après avoir été infecté par la variole, il est immunisé pour le reste de sa vie, (ils ont été immunisés);
- La probabilité qu'un individu meurt dans un petit intervalle de temps dx entre l'âge x et l'âge $x + dx$ est $m(x) \cdot dx$, où $m(x)$ est le taux de mortalité naturelle à l'âge x .

Considérant une population constituée de P_0 individus nés la même année, et soit :

- $S(x)$ le nombre d'individus qui sont encore en vie à l'âge x sans avoir été infecté par la variole ;
- $R(x)$ le nombre d'individus qui sont encore en vie à l'âge x et immunisés contre la variole ;
- $P(x) = S(x) + R(x)$ la population totale vivante à l'âge x .

Bernoulli a écrit alors qu'entre l'âge x et l'âge $x + dx$, chaque individu n'ayant jamais été infecté par la variole (susceptible), a une probabilité $q \cdot dx$ d'attraper la variole et une probabilité $m(x) \cdot dx$ de mourir d'une autre cause. Donc la variation par rapport au temps du nombre d'individus susceptibles est donnée comme suit :

$$\frac{dS}{dx} = -qS(x) - m(x)S(x) \quad (1.1)$$

Durant cette période, le nombre d'individus qui survivent en devenant immunisés est $(1 - p)qS(x) \cdot dx$. D'autre part, il y'a $m(x)R(x)$ individus qui sont déjà immunisés qui meurent naturellement, ce qui conduit à une seconde équation différentielle :

$$\frac{dR}{dx} = q(1 - p)S(x) - m(x)R(x) \quad (1.2)$$

En additionnant les deux équations précédentes, on obtient la variation de la population totale :

$$\frac{dP}{dx} = -pqS(x) - m(x)P(x) \quad (1.3)$$

On obtient que la fraction d'individus qui à l'âge x est encore susceptible d'attraper la variole parmi toute la population est :

$$\frac{S(x)}{P(x)} = \frac{1}{(1 - p)e^{qx} + p} \quad (1.4)$$

Pour obtenir la formule 1.4, Bernoulli a éliminé $m(x)$ dans les équations 1.1 et 1.3 :

$$-m(x) = q + \frac{1}{S} \frac{dS}{dx} = pq \frac{S}{P} + \frac{1}{P} \frac{dP}{dx}$$

ce qui donne :

$$\frac{1}{P} \frac{dS}{dx} - \frac{S}{P^2} \frac{dP}{dx} = -q \frac{S}{P} + pq \left[\frac{S}{P} \right]^2$$

On remarque que le membre de gauche est la dérivée de la fonction $f(x) = S(x)/P(x)$, donc

$$\frac{df}{dx} = -qf + pqf^2 \quad (1.5)$$

Divisons cette équation par f^2 et posons $g(x) = 1/f(x)$, on obtient que $dg/dx = qg - pq$ et $g(0) = 1/f(0) = 1$. Soit $h(x) = g(x) - p$, on obtient $dh/dx = qh$, cette équation admettant comme solution $h(x) = h(0)e^{qx} = (1 - p)e^{qx}$. Finalement on aura $g(x) = (1 - p)e^{qx} + p$ et $f(x) = 1/g(x)$ ce qui assure le résultat 1.4

Bernoulli utilise la table de mortalité de Halley (1656-1742) construite pour la ville de Breslau (1693). Dans cette table Halley décrit le nombre de survivants au début de l'année x ($x = 1, 2, \dots$) d'une cohorte d'individus. Les valeurs des paramètres choisis par Bernoulli dans son modèle sont :

- La probabilité de mourir de la variole $p = 1/8$;
- La probabilité annuelle d'attraper la variole $q = 1/8$. Cette probabilité a été estimée de sorte que le nombre total des morts dus à la variole représente 1/13 de tous les morts, ce qui correspond à la proportion constatée dans plusieurs villes d'Europe.

Avec la formule 1.4 et les valeurs de $P(x)$ données dans le tableau de Halley, on peut calculer le nombre d'individus $S(x)$ encore vivants à l'âge x sans avoir été infectés. On détermine alors le nombre d'individus encore vivants à l'âge x qui ont eu la variole et qui ont survécu : $R(x) = P(x) - S(x)$. Enfin, on détermine pour chaque âge x le nombre de morts dus à la variole entre l'âge x et l'âge $x + 1$. Il s'agit de l'intégrale : $pq \int_x^{x+1} S(t) dt$ qui peut être estimée par la formule $pq(S(x) + S(x + 1))/2$ (formule du trapèze). Finalement, il en déduit que sur les 1300 nouveau-nés, 101 sont destinés à mourir de la variole, ce qui fait 1/13 comme escompté.

Bernoulli, considère ensuite le cas où la variole serait inoculée de manière inoffensive à toute la population dès la naissance. La variole serait éradiquée et la question qui se pose est quel serait le gain en espérance de vie. Ainsi, partant du même nombre P_0 à la naissance et en notant $P^*(x)$ le nombre d'individus encore vivants à l'âge x en l'absence de variole (la variation du nombre d'individus dépend juste du taux de mortalité naturel), dans ce cas le nombre total d'individus vivants à l'âge x :

$$\frac{dP^*(x)}{dx} = -m(x)P^*(x) \quad (1.6)$$

Ainsi Bernoulli montre que :

$$P^*(x) = \frac{P(x)}{1 - p + pe^{-qx}} \quad (1.7)$$

En éliminant comme précédemment $m(x)$ entre les équations 1.3 et 1.6, Bernoulli a obtenu après réarrangement :

$$\frac{1}{P^*} \frac{dP}{dx} - \frac{P}{P^{*2}} \frac{dP^*}{dx} = -pq \frac{S}{P} \frac{P}{P^*}$$

Posons $h(x) = P(x)/P^*(x)$, en utilisant la formule 1.4, en multipliant le numérateur et le dénominateur par e^{-qx} , on obtient :

$$\frac{1}{h} \frac{dh}{dx} = -pq \frac{e^{-qx}}{1 - p + pe^{-qx}},$$

qui est équivalent à $\frac{d}{dx} \ln h = \frac{d}{dx} \ln (1 - p + pe^{-qx})$. avec $h(0) = 1$, donc $h(x) = 1 - p + pe^{-qx}$, ce qui montre l'équation 1.7.

Pour comparer $P(x)$ et $P^*(x)$, Bernoulli utilise l'espérance moyen de vie à la naissance donné par la formule :

$$E(P) = \frac{1}{P_0} \int_0^{\infty} P(x) dx \quad (1.8)$$

avec la variole, et la même expression avec $P^*(x)$ à la place de $P(x)$ en l'absence de variole. Il a trouvé finalement que l'espérance de vie E avec la variole est de 26 ans et 7 mois. Sans la variole, il obtient une espérance de vie E^* de 29 ans et 8 mois. Donc l'inoculation à la naissance permettrait de gagner plus de 3 ans d'espérance de vie.

1.4.2 Modèle logistique de Verhulst

En dynamique des populations, le modèle de Verhulst est un modèle de croissance proposé par Pierre François Verhulst² en 1840. Le modèle est donné par l'équation différentielle suivante :

$$\frac{dC}{dt} = rC \left(1 - \frac{C}{K} \right) \quad (1.9)$$

Où $C(t)$ désigne la population étudiée à l'instant t , r désigne le taux de croissance. Lorsque la population $C(t)$ est petite par rapport au paramètre K , on obtient l'équation approximative :

$$\frac{dC}{dt} \simeq rC \quad (1.10)$$

dont la solution est $C(t) \simeq C(0)e^{rt}$. Le taux de croissance diminue à mesure que $C(t)$ se rapproche de K . Il deviendrait même négatif si $C(t)$ pouvait dépasser K . Pour obtenir l'expression exacte de la solution de l'équation 1.9, on procéderait comme Daniel Bernoulli pour l'équation 1.5.

Divisons l'équation 1.9 par C^2 et posons $p = 1/C$, on obtient $dp/dt = -rp + r/K$. Avec $q = p - 1/K$, on obtient $dq/dt = -rq$ et $q(t) = q(0)e^{-rt} = (1/C(0) - 1/K)e^{-rt}$. On peut donc en déduire $p(t)$ et $C(t)$. Enfin nous obtenons après réarrangement :

$$C(t) = \frac{C(0)e^{rt}}{1 + C(0)(e^{rt} - 1)/K} \quad (1.11)$$

2. **Pierre François Verhulst** : (né à Bruxelles le 28 octobre 1804 - mort le 15 février 1849 dans cette même ville) est un mathématicien belge.

La population totale augmente progressivement de $P(0)$ à l'instant $t = 0$ jusqu'à la limite K , qui n'est atteint que lorsque $t \rightarrow +\infty$. Sans donner les valeurs qu'il a utilisées pour les paramètres r et K , Verhulst a comparé son résultat avec des données concernant la population de la France entre 1817 et 1831, ainsi que d'autres pays en Europe, L'ajustement s'est avéré assez bon [14].

1.4.3 Les modèles compartimentaux déterministes

Les modèles déterministes permettent de classer les individus en différents sous-groupes. Ils sont considérés parmi les modèles les plus utilisés pour décrire le comportement des maladies infectieuses. Ainsi les modèles déterministes tentent de décrire et d'expliquer ce qui se passe sur la moyenne à l'échelle de la population. Ils intègrent des grandes populations. En général ces modèles sont relativement faciles à mettre en place, puisqu'ils nécessitent moins de données. Un autre aspect important vient du fait qu'actuellement les logiciels informatiques exploitant ce type de modèles sont largement disponibles et conviviaux.

Les modèles compartimentaux sont utilisés pour faciliter les calculs de probabilité de contagion. Ces modèles divisent la population en classes épidémiologiques. Elles permettent d'estimer comment le nombre d'individus dans chaque compartiment varie au cours du temps. Par abus de notation, la lettre utilisée pour représenter un compartiment est également employée pour représenter le nombre d'individus dans le compartiment.

1.4.3.1 Modèle déterministe simple de Hamer

Ce modèle est présenté la première fois par William Heaton Hamer³ en 1906, appelé aussi modèle (SI). Il s'agit du premier modèle déterministe simple où l'on considère que la population totale N peut être décomposée en deux catégories : les individus susceptibles d'être infectés S , et les individus infectés I qui sont atteints et qui sont donc contagieux. Pour illustrer son modèle, Hamer posait au départ les hypothèses suivantes [15] :

- Il n'y a ni décès ni guérison (taille de population constante), et pour toute instant t on a :
 $N = S(t) + I(t)$;
- Un individu, lorsqu'il est infecté, devient infectieux et il reste infectieux jusqu'à la fin de sa vie ;
- L'infection s'établit par contact direct entre les individus infectés et les individus susceptibles.

Ce modèle est alors décrit par le système d'équations différentielles suivant :

$$\begin{cases} \frac{dS}{dt} = -\beta SI \\ \frac{dI}{dt} = \beta SI \end{cases} \quad (1.12)$$

3. **William Heaton Hamer** : William Heaton Hamer était un étudiant de l'université de Cambridge. Après avoir obtenu son diplôme en médecine, il entra au département médical de Londres et devient médecin en 1911.

où β est le taux d'infection ou taux de contagion qui désigne 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 nouvel cas (un contact n'entraînant pas forcément une contamination). La figure 1.1 représente le diagramme de transmission de la maladie.

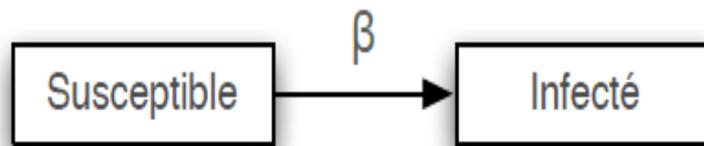


FIGURE 1.1 – Diagramme de transmission du modèle SI.

D'autre part la population totale est constante et $N = S(t) + I(t)$, on obtient :

$$\begin{cases} \frac{dS}{dt} = -\beta S(N - S) \\ \frac{dI}{dt} = \beta I(N - I) \end{cases} \quad (1.13)$$

Il s'agit d'un système d'équations différentielles de type logistique introduite par Verhulst afin de décrire un modèle d'évolution des populations, admettant comme solution :

$$I(t) = \frac{N}{1 + (N - 1) \exp(-\beta N t)} \quad (1.14)$$

Avec $I(0) = 1$ (l'apparition d'une pandémie se traduit par l'introduction d'un individu infecté dans une population saine à l'instant initial). Cette solution est représentée par une sigmoïde appelée courbe épidémique et qui correspond au nombre total de cas à chaque instant. La figure 1.2 représente l'évolution des populations S et I pendant une durée de temps de temps en 70 jours.

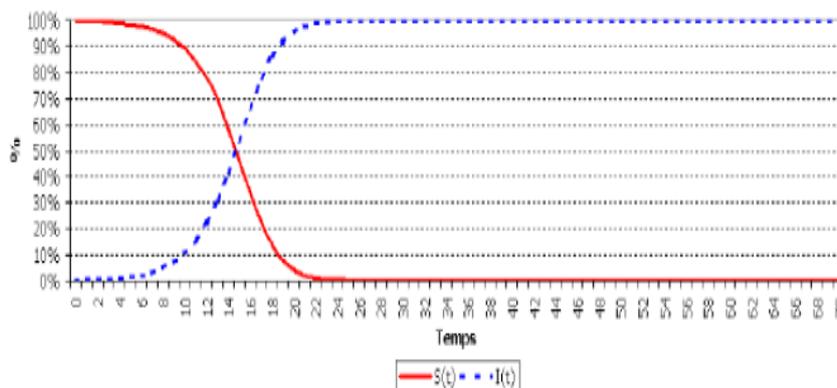


FIGURE 1.2 – Modèle SI - Evolution des populations S et I au cours du temps.

La population compte 500 individus sains à l'instant $t = 0$, dans laquelle on introduit un individu infecté. On remarque que la population va progressivement passer de l'état S à l'état I . Ce modèle est trop simpliste, en fait il ne correspond pas à la réalité en épidémiologie, d'où la nécessité de subdiviser la population totale en plus de compartiments.

1.4.3.2 Modèle déterministe général de Kermack et MC Kendrick

Au début du XX^e siècle, William Ogilvy Kermack⁴ et Anderson Gray McKendrick⁵ ont publié un modèle simple de la propagation des épidémies par contact direct, appelé aussi modèle SIR . Par la suite, ils ont confronté leur modèle avec les données de la diffusion de la peste de Bombay entre 1905 et 1906 [15]. Ils partagent la population en trois catégories, les susceptibles S , les infectés I et les retirés R , ce modèle ne contient que deux paramètres : β le taux d'infection et le taux de retrait (appelé aussi taux de guérison) γ , dont les valeurs sont estimées à partir des données observées. La taille de la population totale est supposée constante comme dans le modèle de Hamer : $N = S(t) + I(t) + R(t)$. La figure 1.3 montre le diagramme de transmission de la maladie en utilisant le modèle SIR . Ce modèle est formulé par le système d'équations différentielles suivantes :

$$\begin{cases} \frac{dS(t)}{dt} = -\beta S(t)I(t) \\ \frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) \\ \frac{dR(t)}{dt} = \gamma I(t) \end{cases} \quad (1.15)$$

La variation du nombre d'individus infectés I dépend de deux termes, le premier est identique à celui du modèle de Hamer et le second désigne la proportion des individus infectés qui vont guérir. En principe, ce rapport concerne des individus guéris et immunisés, ou bien isolés, ou encore décédés, c'est à dire en tout état de cause en dehors du processus d'infection. Elle est désignée par le paramètre γ peut être nulle et dans ce cas on se ramène au cas du modèle simple de Hamer. Il peut aussi être égal à 1 (tous les individus infectés sortent du processus pendant l'intervalle de temps dt), mais généralement, il prend une valeur intermédiaire.



FIGURE 1.3 – Diagramme de transmission du modèle SIR .

4. **William Ogilvy Kermack** : (1898 - 1970) était un biochimiste. Il a réalisé des études mathématiques sur la propagation des épidémies et a établi des liens entre les facteurs environnementaux et certaines maladies.

5. **Anderson Gray McKendrick** : (1876 - 1943) était un médecin militaire et épidémiologiste qui a été le pionnier de l'utilisation de méthodes mathématiques en épidémiologie.

L'étude du modèle précédent permet d'établir le « théorème du seuil ». Le système 1.15 est muni des conditions initiales : $S(0) = S_0 = N, I(0) = I_0 = 0, R(0) = 0$. Cela revient à considérer une population constituée des individus susceptibles dans laquelle on introduit un petit nombre d'infectieux. En divisant la deuxième équation du système (1.12) par γI , on obtient l'équation :

$$\frac{I'(t)}{\gamma I(t)} = \frac{\beta S(t)}{\gamma} - 1 \quad (1.16)$$

Puisque $\gamma I(t) > 0$, alors le signe de $I'(t)$ dépend du signe de $\frac{\beta S(t)}{\gamma} - 1$, pour $\frac{\beta S(t)}{\gamma} > 1$, chaque personne infecté va contaminer plus d'un individu susceptible, et donc la maladie se propagera à un nombre toujours grandissant d'individus. Donc il en sera ainsi jusqu'à ce que le nombre de personnes susceptibles $S(t)$ soit tel que $\frac{\beta S(t)}{\gamma} < 1$. Le rapport $\frac{\beta}{\gamma}$ peut alors s'interpréter comme étant le nombre de contacts pouvant transmettre la maladie par les personnes infectés durant toute leur période de contagion, d'ou, en multipliant par le nombre des individus susceptibles à chaque instant on obtient le nombre des nouveaux cas causés par un seul individu contagieux. L'équation 1.16 met donc en évidence l'importance des conditions initiales dans ce genre de modèles, puisque, si $\frac{\beta S_0}{\gamma} > 1$, il y aura forcément une épidémie alors que dans le cas contraire, seuls quelques individus seront infectés avant que la diffusion de la maladie s'arrête d'elle-même.

Le taux $R_0 = \frac{\beta N}{\gamma}$ est appelé *taux de reproduction de base*, c'est le nombre de cas secondaires produits par un individu infectieux au cours de sa période d'infectiosité, dans une population entièrement constituée de susceptibles.

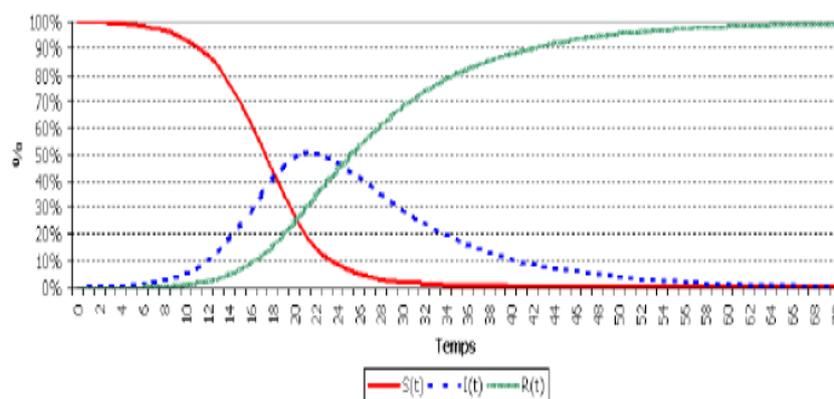


FIGURE 1.4 – Modèle SIR - Evolution des populations S, I et R au cours du temps.

La figure 1.4 représente l'évolution des populations S, I et R avec les mêmes hypothèses que l'exemple précédent. On constate que le nombre d'infectés croît jusqu'à un maximum (pic épidémique) pour diminuer ensuite. Le nombre d'individus susceptibles tend vers zéro, et la population passe progressivement dans l'état retiré.

1.4.3.3 Autres extensions du modèle SIR

Le modèle SIR était la base de plusieurs modèles mathématiques en épidémiologie, il est régulièrement utilisé comme par exemple dans la grippe A en 2009. Ce modèle a subi des changements par l'ajout des extensions (par exemple de nouveaux compartiments) au modèle standard. Ces différentes extensions ont tous les mêmes principes de base déjà énoncés mais elles comprennent des degrés de complexité variés. Nous présentons ci-dessous quelques extensions du modèle SIR avec leur interprétation [16] :

- *SIRS* : Il s'agit du modèle SIR avec immunisation temporaire (les individus du compartiment R réintègrent après un certain délai le compartiment des susceptibles) ;
- *SEIR/SEIRS* : Introduction d'une nouvelle classe Exposé " E " (Individus infectés mais ne peuvent pas encore transmettre la maladie aux personnes susceptibles). *SEIR* : immunité permanente, *SEIRS* : immunité temporaire ;
- *MSIR* : Modèle SIR avec immunité passive (un individu est né avec une immunité passive transféré naturellement par la mère) ;
- *MSEIR/MSEIRS* : Modèle avec immunité passive et période de latente (avec le facteur d'immunité passive, on ajoute la période de latence : *MSEIR* : immunité permanente, *MSEIRS* : immunité temporaire).

1.4.4 Modèle de Ross

Le paludisme est une maladie infectieuse transmise par la piqûre d'un moustique de type anophèle, lui-même contaminé par un humain infecté. C'est en 1897 que Ronald Ross⁶ prouva l'hypothèse émise en 1884 affirmant que le paludisme était transmis par les anophèles.

Même si ses travaux ont été fortement reconnus dans la communauté médicale, les idées de Ross sur le contrôle du paludisme par une lutte anti-vectorielle étaient quant à elles fortement contestées. En effet même s'il était d'accord avec l'idée qu'il était impossible d'éradiquer totalement les moustiques d'une région donnée, il était en total désaccord avec l'idée selon laquelle la transmission du paludisme allait continuer tant qu'il resterait des moustiques. Dans le but de construire un argumentaire qualitatif mais aussi quantitatif, Ross bâtit en 1911 le premier modèle mathématique de la transmission du paludisme dénommé modèle SIS. Ses résultats étaient ainsi résumés dans ce qu'il appelait **le théorème du moustique**. L'idée principale de ce théorème est qu'en dessous d'un certain seuil critique R_0 pour la population des moustiques, le paludisme disparaîtrait de lui-même [17]. Dans son modèle, Ross subdivise chacune des deux populations hôtes en deux compartiments fictifs [18] :

- $S_h(t)$: population humaine susceptible ;

6. **Ronald Ross** : (1857-1932) médecin britannique qui a reçu le prix Nobel de physiologie en 1902 pour ses travaux sur la transmission du paludisme.

- $I_h(t)$: population humaine infectieuse ;
- $S_m(t)$: population de moustiques susceptibles ;
- $I_m(t)$: population de moustiques infectieux ;

Pour illustrer son modèle, Ross s'est basé sur les hypothèses suivantes :

- Les deux populations (humains et moustiques) sont supposées constantes.
- Compte tenu de sa durée de vie relativement courte, un moustique infecté n'a pas le temps de guérir.
- Un humain susceptible devient infectieux à l'issue d'une piqûre d'un moustique infectieux et redevient susceptible après guérison ; et un moustique sain devient infectieux après piqûre d'un humain infectieux. Ces hypothèses sont illustrées dans la figure 1.5.

Remarque 1.4.1. *les flèches en pointillés dans la figure indiquent le sens de l'infection ainsi que flèches pleines représentent le passage d'une classe à l'autre.*

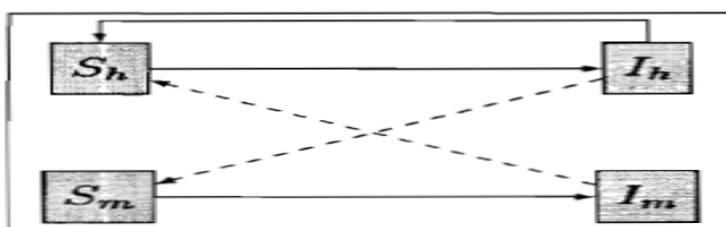


FIGURE 1.5 – Modèle de Ross.

Paramètres du modèle :

- a : nombre moyen de piqure fait par les moustiques par unité de temps ;
- b_{hm} : probabilité qu'une piqûre par un moustique sain sur un humain infectieux transfère l'infection au moustique ;
- b_{mh} : probabilité qu'une piqûre par un moustique infectieux sur un humain susceptible soit infectante ;
- r : probabilité qu'un humain infecté guérisse du paludisme ;
- q_m : taux de mortalité d'un moustique ;

Faisant le changement de variables suivant :

$$x = \frac{I_h}{N_h}; \quad y = \frac{I_m}{N_m}; \quad m = \frac{N_m}{N_h}$$

Avec N_m (resp N_h) représente la population humaine (resp de moustiques) totale.

On obtient donc le système d'équations suivantes décrivant ce modèle :

$$\begin{cases} \dot{x} = mab_{hm}y(1-x) - rx \\ \dot{y} = ab_{mh}x(1-y) - q_my \end{cases} \quad (1.17)$$

Pour illustrer sa théorie, Ross a recherché des valeurs numériques raisonnables pour les paramètres de son modèle. Il a supposé que :

- $q_m = \log(3)/10$ par jour (la mortalité des moustiques est telle que seulement un tiers d'entre eux sont encore en vie après dix jours);
- $r = \log(2)/90$ par jours (la moitié des individus sont toujours infectées après trois mois);
- $a = \log(8/7)$;
- $b_{mh} = 1/3$ (un tiers de tous les moustiques infectés qui sont infectieux), ainsi $b_{hm} = 1/4$.

En tenant compte de la dynamique des humains et des moustiques dans les groupes constitués, il obtient après calcul, le seuil critique $R_0 = \frac{a^2 b_{mh} b_{hm} N_m}{q_m r N_h}$. Il formule alors le théorème du moustique comme suit :

- Si $R_0 < 1$, le paludisme disparaîtra;
- Si $R_0 > 1$, la maladie persiste.

Ross venait ainsi de montrer qu'il n'est pas nécessaire d'éradiquer toute la population de moustiques pour faire disparaître le paludisme, C'est précisément ce que Ross voulait montrer avec son modèle. À ce moment-là, il peut y avoir une lutte anti-vectorielle, conduisant à la disparition du paludisme dans certaines régions.

Remarque 1.4.2. *Le modèle de Ross décrit précédemment est un modèle de type SI pour les deux populations (humains et moustiques). Il existe d'autres types de compartimentage, tels que les modèles SIR / SI, SEIR / SEI, etc... [18].*

1.4.5 Les modèles stochastiques

Les modèles stochastiques sont beaucoup plus conformes. Ils tiennent compte de la nature aléatoire de tout les événements de transmission et de développement d'une infection.

Ils se basent notamment sur les probabilités de variation de risques d'exposition lors d'une épidémie. Ils fournissent beaucoup plus de possibilités pour une modélisation à l'échelle individuelle, dans lesquels chaque individu joue un rôle important dans le modèle.

1.4.5.1 Modèle classique de Reed-Frost

Le modèle Reed-Frost est Le premier modèle stochastique à temps discret proposé en 1928 par Lowell Reed⁷ et Wade Hampton Frost⁸. Le modèle s'appuie sur quelques hypothèses de base [17] :

- La population est considérée homogène (tous les individus ont la même probabilité d'être infecté);

7. **Lowell Reed** : Lowell Jacob Reed (1886 - 1966) était le 7e président de l'Université Johns Hopkins à Baltimore, Maryland. né à Berlin.

8. **Wade Hampton Frost** : (1886 - 1966) Wade Hampton Frost (1880 ? 1938) est né à Marshall, en Virginia. Il était le fils d'un médecin de campagne. Il a obtenu son doctorat en médecine en 1903 à l'Université de Virginie.

- L'infection se transmet d'un individu infecté à un autre susceptible par contact physique ;
- Chaque individu contaminé devient infectieux après une unité de temps ;
- Chaque individu infectieux, guéri après une unité de temps, cette guérison lui confère une immunité définitive ;

On désigne par S_n, I_n, R_n les nombres respectifs d'individus susceptibles, infectés et remis (guéris) à l'instant n . Pour chaque instant on calcul le nombre d'individus infectieux de l'étape $n + 1$ en fonction du nombre d'individus infectieux à l'instant n . Un individu infectieux à l'étape $n + 1$ est donc un susceptible de l'instant n qui a rencontré un infectieux à l'instant n . Soit p la probabilité qu'un susceptible rencontre un individu infectieux, la probabilité pour un susceptible de ne pas rencontrer les I_n infectieux est donc le produit :

$$\overbrace{(1-p)(1-p)\cdots(1-p)}^{I_n \text{ facteurs}} = (1-p)^{I_n} \quad (1.18)$$

La probabilité qu'un susceptible soit contaminé par un infectieux à l'instant n , et donc de devenir infectieux à l'instant $n + 1$ est $1 - (1-p)^{I_n}$. En multipliant cette proportion par le nombre des individus susceptibles à l'instant n , on obtient la proportion des nouveaux infectés I_{n+1} :

$$(1 - (1-p)^{I_n}) S_n \quad (1.19)$$

Ainsi le nouveau nombre des susceptibles est égal au nombre des susceptibles à l'instant n moins le nouveau nombre des individus infectés. On obtient donc le système suivant représentant le nombre des cas à chaque instant :

$$\begin{cases} S_{n+1} = S_n - (1 - (1-p)^{I_n}) S_n = (1-p)^{I_n} S_n \\ I_{n+1} = (1 - (1-p)^{I_n}) S_n \\ R_{n+1} = R_n + I_n \end{cases} \quad (1.20)$$

1.4.5.2 Modèle SI Stochastique

Comme dans le modèle SI déterministe détaillé précédemment, S représente les susceptibles et I les infectieux, et β le taux de contact entre les susceptibles et les infectieux. Le nombre de nouveau cas est proportionnel au nombre de contacts entre les individus susceptibles et ceux qui sont infectieux. La population totale est N , et on a donc :

$$I = N - S$$

On introduit un individu contagieux dans la population, et la taille totale devient $N + 1$. Notons $P_S(t)$ la probabilité d'avoir S susceptibles à l'instant t . L'événement " avoir S susceptibles à l'instant $t + dt$ " peut se produire de deux façons : Soit qu'il y'a $S + 1$ personnes

susceptibles à l'instant $t + dt$, ou bien qu'il y'a juste S personnes susceptibles en même instant. donc la probabilité $P_S(t)$ d'avoir S personnes susceptibles à l'instant $t + dt$ est :

$$P_{S(t+dt)} = (P_{(S+1)(t+dt)}\beta(S+1)(N-S) + P_{S(t+dt)}(1 - \beta(S+1)(N-S)))dt \quad (1.21)$$

Cette équation n'est pas élémentaire à intégrer et elle conduit à des formules qui sont, même numériquement, très difficiles à manipuler. Seuls quelques exemples avec N relativement petit ont été calculé. D'une manière similaire on obtient une équation représentant le modèle SIR, voir [19].

1.5 CONCLUSION

Dans ce chapitre, nous avons défini la notion de l'épidémiologie aussi bien que la modélisation statistique. Puis, nous avons présenté quelques modèles mathématiques bien connus en épidémiologie comme le modèle de Bernoulli ainsi que d'autres types de modèles, à savoir les modèles déterministes et stochastiques.

Modélisation mathématique du COVID-19

2.1 Introduction

La pandémie COVID-19 qui fait référence à « CoronaVirus Infectious Disease 2019 » est une maladie provoquée par le nouveau virus SARS-CoV-2 qui se caractérise par une forte transmissibilité et une létalité relativement peu élevée qui lui permet de se propager longtemps. Cette pandémie pose plusieurs questions aussi bien sur son analyse statistique, sa modélisation potentielle et l'identification des facteurs, sur lesquels il faut agir pour enrayer sa propagation. Le COVID-19 est une maladie respiratoire pouvant être mortelle chez les patients fragilisés par l'âge ou une autre maladie chronique. Elle se transmet par contact rapproché avec des personnes infectées. La maladie pourrait aussi être transmise par des patients asymptomatiques mais les données scientifiques manquent pour en attester avec certitude.

En décembre 2019, des responsables de Wuhan, en Chine, ont signalé pour la première fois le premier cas humain de COVID-19. La maladie a été causée par le nouveau coronavirus qui a été ensuite connu sous le nom de SARS-CoV-2. Bien que certains des premiers cas connus soient liés au marché de gros de produits alimentaires à Wuhan. De nombreux premiers patients étaient des employés du marché ou bien des visiteurs réguliers du marché. Les échantillons environnementaux collectés en décembre 2019 étaient positifs pour le SARS-CoV-2, indiquant en outre que le marché de Wuhan était la cause de l'épidémie ou y avait joué un rôle. On distingue les trois termes qui sont généralement confondus :

- **Coronavirus** : Les coronavirus est une grande famille de virus, divisés en quatre genres : α -coronavirus, β -coronavirus, γ -coronavirus et δ -coronavirus. À l'heure actuelle, seuls sept types de coronavirus sont considérés comme pathogènes chez l'humain, dont quatre provoquent généralement des symptômes de rhume ou de grippe légers chez les personnes ayant une forte immunité, à savoir deux α -coronavirus (HCoV-NL63, HCoV-

229E) et deux β -coronavirus (HCoV-OC43, HCoV-HKU1). Les infections causées par les trois autres coronavirus peuvent entraîner des infections graves et mortelles. Les deux premiers sont des virus qui sont déjà connus : le virus responsable du syndrome respiratoire aigu sévère (SARS-CoV) apparu en Chine, et l'origine de l'épidémie de 2002-2003, et le virus responsable du syndrome respiratoire (MERS-CoV) en 2012. Le troisième est le SARS-CoV-2 le nouveau β -coronavirus isolé en janvier 2020.

- **Le virus** : Le SARS-CoV-2 est le nom officiel du nouveau coronavirus identifié, le nom a été proposé par la Comité international de taxonomie des virus, l'organisme en charge de la classification des virus.
- **La maladie** : Le COVID-19 fait référence à la «maladie» qui est causée par le virus SARS-CoV-2 de la famille des coronavirus. Cette maladie infectieuse est une maladie zoonotique (Une zoonose est une maladie infectieuse transmissible d'un animal à l'homme. Les zoonoses peuvent être transmis directement, ou indirectement par la consommation de produits animaux (lait, viande...)). La maladie est apparue à Wuhan, dans la province du Hubei, en Chine en décembre 2019. Il s'est répandu rapidement, se propageant d'abord en Chine, puis provoquant une épidémie mondiale à l'étranger.

Dès l'enregistrement du premier cas confirmé de contamination de sa population par le COVID-19, officiellement le 2 Mars 2020, le Maroc a su promptement capitaliser les leçons issues de l'expérience internationale de lutte contre ce virus particulièrement virulent et contagieux, en prenant résolument les mesures les plus éprouvées pour bloquer la propagation de la pandémie sur son territoire.

Outre la fermeture immédiate des frontières et des établissements d'enseignement et de formation et la mobilisation de tout le potentiel d'accueil de ses structures hospitalières et sanitaires, il a instauré une radicale et emblématique opération de confinement de la population et décrété l'état d'urgence sanitaire pour en faire respecter les règles avec l'appui d'une campagne intense de communication en faveur des pratiques de distanciation physique, de gestes barrières et d'autres dispositions individuelles ou collectives d'auto-protection.

2.2 Paramètres de la maladie

2.2.1 Taux de reproduction

Le taux de reproduction nous permet de suivre la progression de l'épidémie et de détecter les changements dans la propagation du virus. Cet indicateur doit être interprété en fonction du contexte de transmission. Par exemple, si l'infection est bien contrôlée (c'est-à-dire quelques cas) et qu'une éclosion est observée, le taux de reproduction augmentera. Cependant, si le nombre de cas est déjà élevé, une valeur élevée du taux de reproduction sera encore plus in-

quiétante. C'est un indicateur épidémiologique primordial, mais il est important de l'interpréter en rapport avec d'autres données épidémiologiques. On dispose de deux facteurs à distinguer :

- *Taux de reproduction de base R_0* : C'est le nombre moyen d'infections secondaires transmises lorsqu'un individu infectieux est introduit dans une population entièrement constituée par des individus sains, c'est-à-dire au début d'une possible épidémie. A titre indicatif, pour la grippe espagnole de 1918, le R_0 était estimé entre 2 et 3 ; Pour les gripes de 1957 et 1968, R_0 était de 1,5-3 ; pour la tuberculose, R_0 est de 10 ; pour la rougeole, il est de 12-18.

Ce paramètre permet de définir un seuil appelé seuil d'«immunité collective», défini par :

$$S = 1 - \frac{1}{R_0} \quad (2.1)$$

au-dessus duquel la population dispose d'une immunité suffisante pour faire tendre à disparaître tout foyer épidémique. Cette proportion peut être comprise comme la proportion minimale de la population à vacciner, de sorte que l'infection ne dégénère pas en épidémie [1]. Plus que le taux de reproduction est élevé, plus le seuil d'immunité collective est élevé. Ainsi, pour un virus tel que celui de la rougeole, qui a un R_0 supérieur à 10, il faut immuniser plus de 90% de la population pour se prémunir d'une épidémie.

- *Taux de reproduction effectif R_t* : C'est le nombre moyen d'infections secondaires produites par une personne infectée dans une population constituée de personnes susceptibles d'acquérir le virus. Un taux de reproduction supérieur à 1 désigne que l'épidémie est en croissance. Par exemple, un R_t de 3 implique qu'une personne infectée transmettra le virus à trois personnes (en moyenne) durant la période où elle est infectieuse. Si le taux est près de 1, la transmission et le nombre de cas demeurent stables. Un taux inférieur à 1 indique que l'épidémie est bien contrôlée. Supposons qu'à un certain instant pendant une épidémie des mesures contre la propagation de la maladie sont appliquées, le taux de reproduction effectif peut être donné comme suit : $R_t = (1 - c)R_0$ où c désigne l'effet des mesures prises. Dans la littérature, les auteurs désignent parfois le R_t comme étant le taux de reproduction de base.

2.2.1.1 Techniques pour le Calcul du taux de reproduction

- ▷ **En utilisant le temps de génération [20]** : La première étape consiste à calculer le taux d'accroissement (s'exprime souvent sous forme d'un pourcentage : par exemple un taux de croissance de 20% par jour signifie que la population est multipliée par 1,2 chaque jour) de la courbe d'incidence (courbe associée aux nombres de nouveaux cas détectés chaque jour). Le taux de croissance est en général noté r et au passage, ce paramètre nous permet de calculer le temps de doublement de l'épidémie T_d par la formule :

$$T_d = \frac{\ln(2)}{r} \quad (2.2)$$

Si le temps de doublement est de 4, alors en 4 jours une personne infectée en donnera 2 autres cas, qui 4 jours après en vont donner chacun 2, et ainsi de suite. Après 40 jours, donc 10 doublements, on aura eu 2^{10} infections.

Le taux de reproduction de base est lié au produit entre le taux de croissance r et le temps de génération T_g (voir ci-dessous) par la formule :

$$R_0 \approx r \times T_g + 1 \quad (2.3)$$

- ▷ **En utilisant la fonction de survie [21]** : Considérons une large population et soit $F(a)$ la probabilité qu'un individu nouvellement infecté reste contagieux pendant le temps a . Soit $b(a)$ le nombre moyen de des infectés qu'un individu infectieux produira par unité de temps pendant toute la durée a . Ainsi le nombre de reproduction de base est donné par la formule :

$$R_0 = \int_0^{\infty} b(a)F(a)da \quad (2.4)$$

Cette méthode peut aussi être étendue pour étudier des modèles dans lesquels une série d'états sont impliqués dans la "reproduction" d'un individu infecté. Par exemple la modélisation de la malaria. Un humain infecté peut transmettre le virus à un moustique, qui peut de sa part infecter d'autres humains. Ce cycle doit être pris en considération dans le calcul du taux de reproduction. En général, si seulement deux états infectieux distincts sont impliqués dans un cycle d'infection, $F(a)$ peut être définie comme étant la probabilité qu'un individu dans l'état 1 à l'instant zéro infecte un individu dans l'état 2 durant la période a . De même, $b(a)$ est le nombre moyen de nouveaux individus dans l'état 1 produit par un individu qui a été dans l'état 2 pendant le temps a . Dans le cas de la malaria, $F(a)$ pourrait être la probabilité qu'un humain infecté au temps zéro infecte un moustique qui va rester en vie durant la période a . Plus concrètement, $F(a)$ serait l'intégrale :

$$F(a) = \int_0^a p(t) \times q(t) \times r(t)dt \quad (2.5)$$

avec : $p(t)$ est la probabilité qu'un humain infecté à l'instant 0 est encore vivant à l'instant t , $q(t)$: la probabilité qu'un humain infecté pendant le temps t infecte un moustique et $r(t)$ la probabilité qu'un moustique infecté vive jusqu'à l'âge a . tandis que $b(a)$ serait tout simplement le nombre moyen d'humains infectés par un moustique infecté pendant le temps a .

- ▷ **En utilisant la matrice « Next generation » [21]** : C'est une méthode introduite la première fois en 1990, c'est une méthode générale pour le calcul du R_0 dans le cas des

modèles avec des compartiments, englobant les situations dans lesquelles la population est divisée en classes discrètes et disjointes. Dans cette méthode, R_0 est défini comme le rayon spectral (la plus grande valeur propre) d'une certaine matrice appelée «l'opérateur de la prochaine génération». La formation de cet opérateur consiste à déterminer deux compartiments, infectés et non infectés, à partir du modèle. Supposons qu'il existe n compartiments dont m sont infectés. On ordonne les compartiments de sorte que les m premiers correspondent à des compartiments d'infectés. On définit le vecteur $\bar{x} = x_i$, avec $i = 1, \dots, n$ où x_i est la proportion de la population dans le compartiment i . Soit $F_i(\bar{x})$ le taux d'apparition de nouveaux cas infectés dans le compartiment i . Posons $V_i(\bar{x}) = V_i^-(\bar{x}) - V_i^+(\bar{x})$ où V_i^+ désigne le taux de transmission des individus dans le i^{eme} compartiment, et V_i^- est le taux de transfert des individus hors du i^{eme} compartiment. Donc $F_i(\bar{x}) - V_i(\bar{x})$ donne le taux de variation de x_i . Notez que F_i ne doit inclure que les cas récemment apparues, et pas les termes qui décrivent le transfert d'individus infectieux entre les compartiments infectés. Supposons que les F_i et les V_i satisfaisaient les conditions décrites dans [22] : On obtient la matrice du prochaine génération FV^{-1} , avec F et V sont définis comme suit :

$$F = \left[\frac{\partial F_i(x_0)}{\partial x_j} \right]_{i,j=1\dots m} \quad \text{et} \quad V = \left[\frac{\partial V_i(x_0)}{\partial x_j} \right]_{i,j=1\dots m} \quad (2.6)$$

avec x_0 désigne l'équilibre sans maladie (le point auquel aucune maladie n'est présente dans la population). Ainsi le taux de reproduction de base R_0 est défini comme étant le rayon spectral de la matrice FV^{-1} . A titre d'exemple, considérons le modèle SEIR avec 4 compartiments :

$$\begin{aligned} \dot{S} &= -\beta SI + \lambda - \mu S \\ \dot{E} &= \beta SI - (\mu + k)E \\ \dot{I} &= kE - (\gamma + \mu)I \\ \dot{R} &= \gamma I - \mu R \end{aligned} \quad (2.7)$$

où μ est le taux de mortalité naturel is, β le taux d'infection, k est le taux auquel un individu en période de latence devient infectieux et γ est le taux de guérison, et λ désigne le taux de natalité. Puisque nous sommes préoccupés seulement par les populations qui propagent l'infection, nous n'avons besoin que des classes exposées E et infectées I .

$$\left. \begin{aligned} \dot{E} &= \beta SI - (\mu + k)E, \\ \dot{I} &= kE - (\gamma + \mu)I. \end{aligned} \right\} \quad (2.8)$$

pour le système 2.8 on a :

$$F = \begin{pmatrix} 0 & \beta\lambda/\mu \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad V = \begin{pmatrix} \mu + k & 0 \\ -k & \gamma + \mu \end{pmatrix} \quad (2.9)$$

ce qui donne :

$$R_0 = \frac{k\beta\lambda}{\mu(\mu + k)(\mu + \gamma)} \quad (2.10)$$

La figure 2.1 représente les valeurs du taux de reproduction pour différents pays au début de la pandémie du COVID-19.

Pays	R_0	Date
	1,4-2,5	23 janvier
USA	2,3-3,3	27 janvier
Suisse	2,2 [1,4-3,8]	25 janvier
Canada	6,47 [5,71-7,23]	24 janvier
UK	3,11 [2,39-4,13]	27 janvier
USA	2	26 janvier
Chine	2,2 [1,4-3,9]	29 janvier
Chine	Preprint : 3,30-5,47 2,56 [2,49-2,63]	31 janvier
UK/OMS	2,6 [1,5-3,5]	25 janvier
Chine	2,68 [2,47-2,86]	31 janvier

FIGURE 2.1 – Quelques estimations du taux de reproduction du COVID-19 dans plusieurs pays au début de l'épidémie

2.2.2 Période d'incubation

La période d'incubation fait référence au temps écoulé entre l'infection et le début d'apparition des symptômes de la maladie. Ce facteur fournit des données précieuses sur la dynamique du transmission de la maladie et aide à mettre en oeuvre les mesures de contrôle appropriées, telles que la recherche des contacts et l'estimation des temps d'isolement à domicile pour les personnes qui ont été infectées par le pathogène ou ont été en contact avec une personne infectée. En 13 janvier 2021 l'institut Pasteur a estimé la période d'incubation du sars-cov 2 au Maroc à 5,5 jours. Pour le MERS-CoV, la période d'incubation moyenne était de 5 jours ; Dans la grippe saisonnière, la période d'incubation est de 2 jours. Pour le cas de la grippe porcine (virus de la grippe H1N1), la période d'incubation est de 1 à 4 jours [23]. Il faut distinguer la période d'incubation à la période de latence qui est un autre facteur épidémiologique désignant la durée entre l'infection et le début de la phase infectieuse. En général la période d'incubation est plus longue que la durée de latence. Le tableau 2.1 donne les estimations du période d'incubation pour différentes maladies.

Maladie	Période d'incubation
Choléra	entre 0 et 3 jours
Ébola	entre 2 et 21 jours
Covid-19 (Coronavirus SARS-CoV-2)	entre 2 et 14 jours
SARS	entre 1 et 10 jours
Variole	entre 7 et 17 jours
VIH	2 à 3 semaines à plusieurs mois, voire plus

TABLE 2.1 – Période d'incubation pour différentes maladies.

2.2.3 Temps de génération, Intervalle sériel

Le temps de génération correspond à la durée nécessaire pour qu'une personne récemment infectée en infecte une autre. Il s'agit d'une question qui a alimenté bien des discussions, notamment en épidémiologie. Il existe plusieurs points de vue sur sa définition, son utilité et les approches permettant de l'estimer. C'est un paramètre primordiale dans la dynamique d'une maladie transmissible, en effet la mise en place des mesures pour la maîtrise de la transmission d'une maladie tes que l'identification des cas, isolement et quarantaine sera plus facile pour une maladie avec une longue période de génération. Par exemple deux maladies ayant un taux de reproduction similaire égal à 2, mais un intervalle de génération différent (respectivement 3 et 9 jours), conduisent à un nombre de cas cumulés sur une période de 9 jours de 14 nouveaux cas pour la première maladie et de 2 pour la seconde. Le temps de génération peut être calculé par la formule suivante :

$$T_g = \text{période d'incubation} + \frac{\text{période de contagiosité}}{2} \quad (2.11)$$

Généralement, la plupart des études épidémiologiques se réfèrent à l'intervalle sériel comme la durée entre l'apparition des premiers symptômes chez une personne infecté et l'apparition des symptômes chez un cas secondaire infecté par ce premier cas. Certaines études de modélisation définissent l'intervalle sériel comme étant le temps de génération ; cependant, le temps de génération dépend d'événements non observables (c.-à-d. la date du début de l'infectiosité : la capacité d'infecter les autres), mais pour la plupart des maladies infectieuses comme le cas de la grippe, l'infectiosité peut commencer avant l'apparition des symptômes.

2.2.4 Taux de létalité

L'une des caractéristiques importantes des maladies infectieuses, en particulier des maladies infectieuses causées par de nouveaux agents pathogènes tels que le SARS-CoV-2, dont la gravité est la mesure ultime de sa capacité létale. Le taux de létalité peut nous servir à mesurer la gravité de la maladie, à identifier les personnes à risque et à évaluer la qualité des soins de santé.

Deux indicateurs sont utilisés pour évaluer le pourcentage de personnes infectées qui meurent de la maladie. Le premier est le taux de létalité réel IFR (Infection fatality rate), qui est utilisé pour estimer la proportion de décès parmi toutes les personnes infectées. Le second est le taux de létalité apparent CFR (Case fatality rate), qui est utilisé pour estimer la proportion de décès parmi les cas confirmés. Afin de mesurer avec précision l'IFR, on doit avoir une compréhension globale du nombre d'infections et du nombre de décès dus à la maladie [24].

Comme pour la plupart des maladies infectieuses, le niveau de transmission de la COVID-19 est souvent sous-estimé car un grand nombre de personnes infectées ne sont pas détectées, soit parce qu'ils ne présentent aucun symptôme, ou qu'ils ne présentent que des symptômes bénins, alors n'allez pas dans les établissements médicaux. Aussi les personnes négligées ou mal desservies de la population sont moins susceptibles d'avoir accès aux soins de santé. Le sous-dépistage des cas peut être exacerbé lors d'une épidémie, lorsque la capacité de dépistage est restreinte ou limitée aux personnes atteintes de formes graves de la maladie et aux groupes à risque prioritaires (soignants dispensant des soins de première ligne, personnes âgées, personnes présentant des comorbidités).

Les différences de mortalité entre les différents pays sont des importants indicateurs du risque relatif de décès, lesquels orientent les décisions politiques concernant l'affectation des ressources médicales limitées pendant la pandémie de COVID-19 [24].

2.2.4.1 Calcul du ratio de létalité réel (IFR)

La gravité effective d'une maladie peut être décrite par le taux réel de létalité (infection fatality rate) :

$$IFR = \frac{\text{nombre de décès dus à la maladie}}{\text{nombre des individus infectés}} \times 100 \quad (2.12)$$

Une méthode pour calculer le nombre réel de personnes infectées consiste à effectuer un contrôle sérologique sur un échantillon aléatoire représentatif de la population afin de détecter l'exposition à un agent pathogène. De nombreuses enquêtes sérologiques de ce type sont étaient menées dans le monde.

2.2.4.2 Calcul du ratio de létalité apparent (CFR)

Le taux de létalité apparent CFR (Case fatality rate) est la proportion de personnes chez qui une maladie a été diagnostiquée et qui y succombent ; il s'agit donc d'une mesure de la gravité pour les cas détectés :

$$CFR = \frac{\text{nombre de décès dus à la maladie}}{\text{nombre total de cas détectés}} \times 100 \quad (2.13)$$

Les taux de létalité apparents fiables susceptibles d'être utilisés dans le but d'évaluer la létalité d'une épidémie que les mesures de santé publique mises en oeuvre, ces derniers sont

généralement obtenus à la fin d'une flambée épidémique une fois que tous les cas ont été soit décédés ou rétablis. Toutefois, ce calcul peut ne pas être valable durant une épidémie, puisqu'il repose sur les deux hypothèses :

- Au début d'une épidémie, la surveillance a tendance à se concentrer davantage sur les patients symptomatiques, de sorte que les cas plus légers et asymptomatiques sont rarement détectés, conduisant à une surestimation du taux de létalité apparente. À mesure que les tests et la détection active des cas augmentent, cette surestimation peut diminuer.

- Lors d'une épidémie, certains des cas évolutifs déjà détectés peuvent ensuite décéder, ce qui entraîne une sous-estimation de la valeur du CFR, en effet, avec le COVID-19, il y en a beaucoup qui sont malades et qui peuvent mourir à cause de ce virus mais qui vont être répertoriés comme étant morts de quelque chose d'autre. Ce n'est pas un problème une fois qu'une épidémie est terminée, car le nombre total de décès sera connu et nous pourrons l'utiliser pour calculer le CFR. Mais lors d'une épidémie, nous devons être prudents dans la manière d'interpréter le CFR car l'issue (guérison ou décès) d'un grand nombre de cas est encore inconnue.

C'est ce qui s'est passé lors de l'épidémie de SARS-CoV en 2003 : le CFR était estimé entre 3% et 5% au cours des premiers stades de l'épidémie, mais était passé à environ 10% à la fin. Ce n'est pas seulement un problème pour les statisticiens ; cela a eu de réelles conséquences négatives sur la compréhension de l'épidémie. Les faibles chiffres publiés initialement ont entraîné une sous-estimation de la gravité de l'épidémie, et la variation du CFR au fil du temps a donné une fausse impression que le SARS devenait de plus en plus mortel avec le temps.

2.2.4.3 Calcul du taux de létalité apparent (CFR) pendant une épidémie

Le taux de létalité apparent des cas (CFR) calculé à l'aide de la formule (3.3) pendant une épidémie est une estimation conditionnelle, et cette valeur varie en fonction de la notification tardive du cas ou du décès. Par conséquent, la valeur estimée de ce facteur change considérablement au cours de l'épidémie et, à mesure que la maladie progresse, elle est résolue et tend à stabiliser la valeur finale estimée. La solution simple pour réduire le préjudice causé par les retards dans la résolution des cas pendant l'épidémie est de limiter l'analyse aux cas résolus ; la nouvelle formule est donnée comme suit :

$$CFR = \frac{\text{nombre de décès dus à la maladie}}{\text{nombre de décès dus à la maladie} + \text{nombre des individus rétablis}} \times 100 \quad (2.14)$$

Cependant, cette méthode ne peut pas éliminer tous les écarts associés aux retards de notification. Par exemple, les différences dans le temps requis pour résoudre les cas peuvent biaiser cette estimation. Si les patients atteints de la maladie meurent plus rapidement qu'ils ne guérissent, dans ce cas le taux de létalité apparent (CFR) peut être surestimé, sinon, il peut être sous-estimé. Le tableau 2.2 montre l'impact de quelques facteurs sur le taux de létalité au début de l'épidémie actuelle du COVID-19 en Chine.

Facteurs	Valeurs	CFR (%)
Age	0 – 9	–
	10 – 19	0.2
	20 – 29	0.2
	30 – 39	0.2
	40 – 49	0.4
	50 – 59	1.3
	60 – 69	3.6
	70 – 79	8.0
	> 79	14.8
Sexe	Homme	2.8
	Femme	1.7
Maladie Préexistante	Maladie cardiovasculaire	10.5
	Diabète	7.3
	Maladie respiratoire chronique	6.3
	Hypertension	6.0
	Cancer	5.6
	Aucune maladie préexistante	0.9

TABLE 2.2 – Le taux de létalité pour le COVID-19 en Chine dans la période entre le 1^{er} janvier et 20 février 2020

Le taux de létalité est souvent confondu avec un autre facteur épidémiologique, le «taux de mortalité», ce dernier mesure la probabilité qu'un individu de la population meure à cause de la maladie; pas seulement parmi ceux qui sont infectés ou qui sont confirmés comme infectés. Il est calculé en divisant le nombre de décès dus à la maladie par le nombre total des individus dans la population. Par exemple, s'il y avait 10 décès dans une population de 1 000, le taux de mortalité serait de 1%, même si par exemple seul 100 personnes avaient reçu un diagnostic de la maladie. Cette différence est importante; par exemple durant la pandémie de grippe espagnole en 1918, une estimation était donnée est que cette pandémie a tué 50 millions de personnes. Cela aurait représenté 2,7% de la population mondiale. Ce qui signifie que le taux de mortalité était de 2,7%. Mais cette proportion était souvent déclarés comme étant le taux de létalité, ce qui est faux, car le virus n'a pas infecté tout le monde. Si le taux de mortalité était de 2,7%, alors le taux de létalité était beaucoup plus élevé.

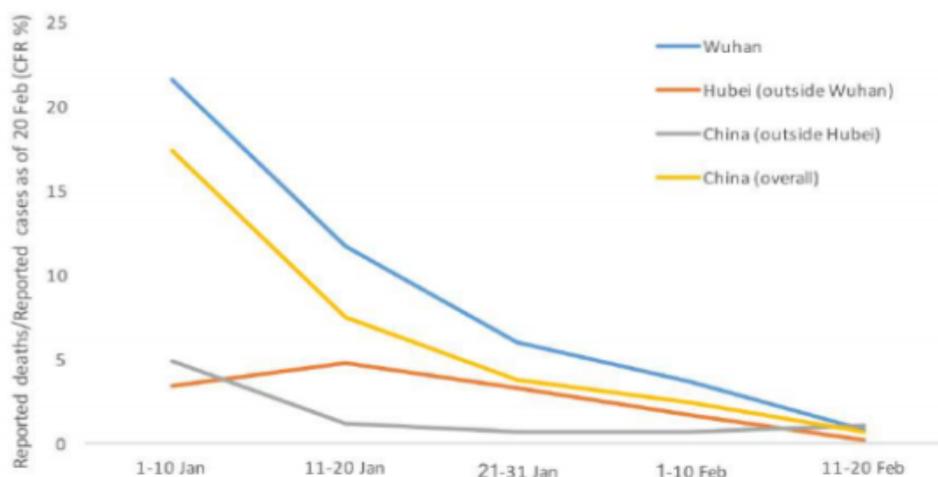


FIGURE 2.2 – Le taux de létalité pour le COVID-19 en Chine dans la période entre le 1^{er} janvier et 20 février 2020

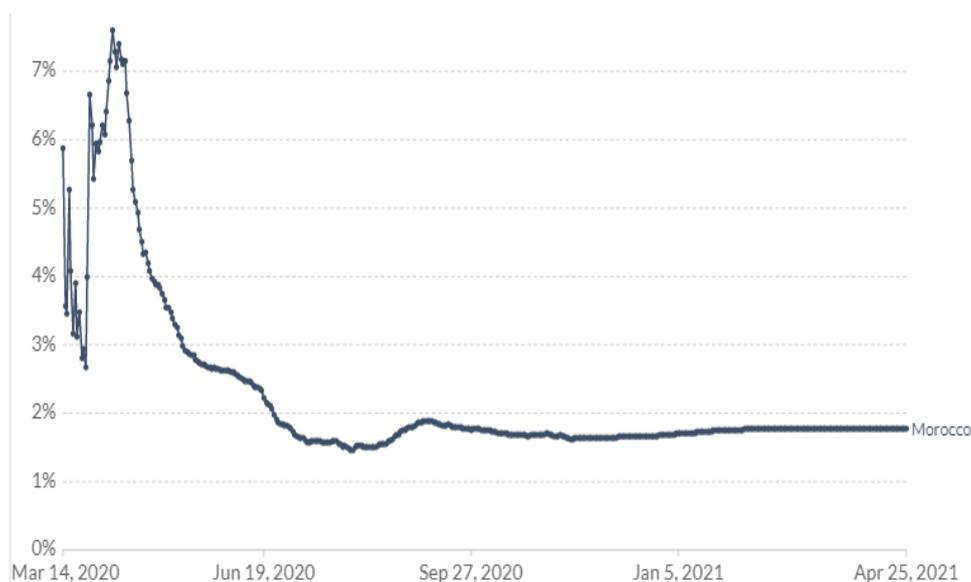


FIGURE 2.3 – Variation du taux de létalité au Maroc entre 14 Mars 2020 et 25 Avril 2021

La figure 2.2 montre les valeurs du CFR pour le COVID-19 dans plusieurs villes en Chine à partir du début janvier au 20 février 2020. On constate que dans les premiers stades de l'épidémie, le CFR était beaucoup plus élevé : 17,3% de la Chine dans son ensemble (en jaune) et plus de 20% au centre de l'épidémie, à Wuhan (en bleu). Mais dans les semaines qui ont suivi, le CFR a diminué, atteignant aussi bas que 0,7% pour les patients qui ont présenté des symptômes pour la première fois après le 1^{er} février. L'OMS affirme que c'est parce que «les mesures de soins ont évolué au cours de l'épidémie». On remarque également que le CFR était différent en fonction des endroits. Au 1^{er} février, le CFR à Wuhan était toujours de 5,8% alors qu'il était de

0,7% dans le reste de la Chine. La figure 2.3 donne la variation du taux de létalité au Maroc tout au long de l'épidémie de COVID-19.

2.3 Quelques modèles mathématiques appliqués au COVID-19

Le champ de l'épidémiologie mathématique a connu, au cours de la pandémie de COVID-19, une production doublée d'une médiatisation exceptionnelle de ses travaux. Alors même que les données et les connaissances sur la maladie émergente étaient parcellaires, une grande diversité de modèles a été développée et appliquée dans des délais inédits, dans l'objectif d'estimer le nombre de reproduction, la date de début de l'épidémie ou l'incidence cumulée, mais aussi afin d'explorer différents scénarios d'interventions non pharmaceutiques. Il convient cependant de noter qu'un modèle est une approximation de la réalité ; que sa pertinence et son utilité sont directement fonction des choix qui ont été faits lors de son élaboration ; que ces choix sont guidés par les questions, énoncées a priori, auxquelles le modèle est conçu pour répondre. Une grande vigilance s'impose donc lors de l'utilisation d'un modèle par d'autres que ceux qui l'ont élaboré.

2.3.1 Modèles déterministes

- ▷ Ebraheem *et al.* [25] ont développé un modèle SIR pour montrer l'effet de l'incorporation de deux périodes dans le modèle SIR simple. Ces délais correspondent à la période d'incubation et de période de guérison telle qu'elle apparaît dans COVID-19. Dans un premier temps, on considère le modèle SIR classique :

$$\begin{aligned} \frac{dS}{dt} &= \frac{-\beta S(t)I(t)}{N} \\ \frac{dI}{dt} &= \frac{\beta S(t)I(t)}{N} - \gamma I(t) - \alpha I(t) \\ \frac{dR}{dt} &= \gamma I(t) \end{aligned} \quad (2.15)$$

En introduisons la période d'incubation " τ_1 " et la période de guérison " τ_2 " on obtient le système :

$$\begin{aligned} \frac{dS}{dt} &= -\beta S(t - \tau_1) I(t - \tau_1) \\ \frac{dI}{dt} &= \beta S(t - \tau_1) I(t - \tau_1) - \gamma I(t - \tau_2) - \alpha I(t) \\ \frac{dR}{dt} &= \gamma I(t - \tau_2) \end{aligned} \quad (2.16)$$

Dans le système d'équations, un individu susceptible est supposé interagir avec un individu infectieux et ne se déplace vers le compartiment infecté qu'après un certain temps "période d'incubation". La période d'incubation " τ_1 " n'est valable que lors du passage du compartiment des susceptibles au compartiment des infectés. De même, " τ_2 " est la période pendant laquelle un individu infecté passe du compartiment infecté au compartiment des personnes guéris. Les variations des paramètres du modèle constituent différents cas correspondant à différentes situations.

Les variations sont introduites en fonction de ce qui est observé dans la littérature pour prédire quelles sont les actions possibles du future. Le modèle est bien adapté aux données pandémiques pour quatre pays (Allemagne, Italie, Koweït et Oman). L'étude montre que l'allongement de la période de guérison entraînera une double exposition et/ou une grande propagation de la maladie. Bien que le taux de mortalité ne représente qu'une fraction mineure du nombre total de cas infectés, il est important de l'inclure pour un nombre de reproduction plus précis.

- ▷ Yafia [26] a présenté un modèle mathématique SIR pour étudier la dynamique de la propagation du COVID-19. Le système d'équations différentielles représentant le modèle est donné comme suit :

$$\begin{cases} \frac{dS}{dt} = \Lambda - \beta(1 - \alpha)(1 - \delta)SI - \alpha S \\ \frac{dI}{dt} = \beta(1 - \alpha)(1 - \delta)SI - \gamma I - \delta I \\ \frac{dR}{dt} = \gamma I + \alpha S + \delta I \\ N = S + I + R = \text{constante si } \Lambda = 0 \\ S(0) > 0, I(0) > 0, R(0) > 0 \end{cases} \quad (2.17)$$

où Λ est le taux de naissance, α et δ sont respectivement le taux de confinement et d'isolation, β est le taux de transmission et γ désigne le taux de guérison. Dans ce travail l'auteur a pu déterminer le rôle du confinement des individus sains et l'isolement des personnes infectieuses, en fonction du nombre de reproduction de base R_0 . L'idée étant d'étudier la propagation du virus durant trois périodes, début de la pandémie (2-8 Mars), période intermédiaire (8-19 Mars) et une période après confinement. Pendant la première période, le taux de transmission était faible, ce qui explique une croissance modérée du nombre des infectés, ainsi pendant la deuxième période le taux de transmission était plus grand, ce qui a conduit à une augmentation du nombre de personnes infectées. Dans la troisième période avec le confinement et l'isolation, le taux de transmission devient plus petit. L'expression du nombre de reproduction de base pendant la troisième période est donné par : $R_0 = \frac{\Lambda\beta(1-\alpha)(1-\delta)}{\alpha(\gamma+\delta)}$, à partir de cette expression si on augmente les taux (α et δ), le R_0 devient plus petit que 1 sinon il sera plus grand que 1. Finalement, le modèle décrit précédemment a été utilisé pour effectuer des prévisions de la propagation du COVID-19 au Maroc à partir du 12 Mars 2020.

- ▷ Serhani *et al.* [27] ont construit un modèle mathématique décrivant la transmission du

COVID-19 au Maroc. Le modèle est donné par par le système suivant :

$$\left\{ \begin{array}{l} \dot{S}(t) = -\beta S(A + \alpha I) + \xi Q, \\ \dot{I}(t) = \theta Q + \delta A - (\mu + d_I) I, \\ \dot{A}(t) = \beta SA - (\delta + \lambda) A, \\ \dot{Q}(t) = \alpha \beta SI - (\xi + \theta) Q, \\ \dot{R}(t) = \mu I + \lambda A, \\ \dot{D}(t) = d_I I, \\ \text{avec les condition initials} \\ S(0) = S_0, I(0) = I_0, A(0) = A_0, Q(0) = Q_0, R(0) = R_0, D(0) = D_0 \end{array} \right. \quad (2.18)$$

Où " S ", " I " et " R ", sont les compartiments du modèle SIR simple, " A " sont les infectés asymptomatiques, " Q " la fraction de la population confinés et " D " sont les morts. Il est établi par une analyse théorique et illustré par des simulations que le niveau de confinement est très important pour empêcher la propagation de la maladie en l'absence de vaccin. Plusieurs scénarios sont testés avec différentes valeurs du coefficient de contrôle du confinement C_0 . Dans cette étude une relation entre le nombre de reproduction de base R_0 et C_0 a été établie, montrant qu'un confinement à domicile non convenablement pratiqué peut conduire à une persistance de la pandémie au-delà de la période prévue. Ainsi dans l'absence des informations crédibles complètes sur l'évolution réelle de la pandémie, en particulier pour les personnes infectées asymptomatiques, il est très difficile d'estimer les différents paramètres du modèle.

- ▷ Alizon *et al.*[28] ont développé un modèle SEAIR pour étudier la pandémie du COVID-19. Dans ce travail, le nombre de reproduction de base R_0 est calculé en fonction des paramètres du modèle. Ainsi les auteurs ont effectué des simulations d'intervention à partir des valeurs des paramètres par défaut. Malgré sa simplicité, le modèle permet de tester une stratégie de suppression de l'épidémie (contrôle fort mais de courte durée) avec la stratégie d'atténuation (contrôle plus léger mais durable). Le système d'équations

différentielle représentant le modèle est le suivant :

$$\begin{aligned}
 \frac{dS}{dt} &= -\lambda S \\
 \frac{dE_1}{dt} &= p\lambda S - \epsilon E_1 + p\nu \\
 \frac{dA_1}{dt} &= \epsilon E_1 - \sigma A_1 \\
 \frac{dI_1}{dt} &= \sigma A_1 - \gamma_1 I_1 \\
 \frac{dR_1}{dt} &= \gamma_1 I_1 \\
 \frac{dE_2}{dt} &= (1-p)\lambda S - \epsilon E_2 + (1-p)\nu \\
 \frac{dA_2}{dt} &= \epsilon E_2 - \sigma A_2 \\
 \frac{dI_2}{dt} &= \sigma A_2 - (\gamma_2 + \alpha) I_2 \\
 \frac{dR_2}{dt} &= \gamma_2 I_2 \\
 \frac{dM}{dt} &= \alpha I_2
 \end{aligned}
 \tag{2.19}$$

dont le schéma associé est donné dans la figure suivante :

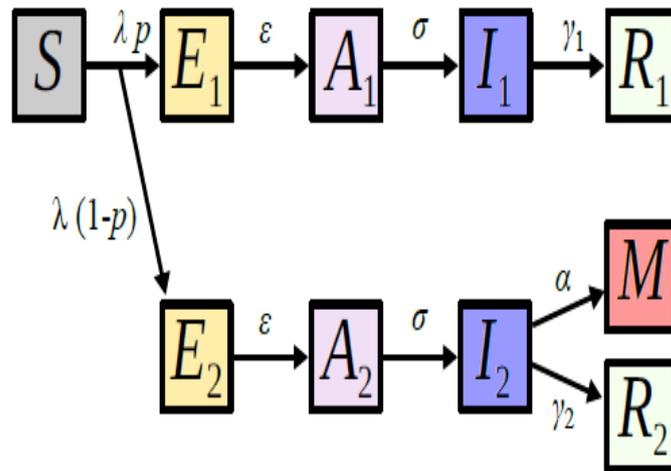


FIGURE 2.4 – Diagramme de transmission du modèle SEAIR.

Avec "S" sont les susceptibles, "E" : les infectés mais non infectieux, "A" : les infectés asymptomatiques, "I" : les infectés avec des symptômes, "R" sont les guéris et "M" sont les morts. Notons que les compartiments portant l'indice 2 sont les individus qui ont des cas plus sévères.

- ▷ Zhihui *et al.*[29] ont construit un modèle épidémique de type SEIR pour le COVID-19 intégrant l'impact des contacts entre les individus susceptibles et les infectés asymptomatiques/symptomatiques est présenté. Le modèle est donné par le système suivant :

$$\begin{cases} \dot{S}(t) = A - \frac{\beta_1(1-d)SI_A}{1+\beta_1(1-d)h_1I_A} - \frac{\beta_2(1-d)SI}{1+\beta_2(1-d)h_2I} - \mu S \\ \dot{E}(t) = \frac{\beta_1(1-d)SI_A}{1+\beta_1(1-d)h_1I_A} + \frac{\beta_2(1-d)SI}{1+\beta_2(1-d)h_2I} - \mu E - \mu_1 E - \gamma_{11} E - \gamma_{12} E \\ \dot{I}_A(t) = \gamma_{11} E - \mu I_A - \mu_{21} I_A - \gamma_{21} I_A, \\ \dot{I}(t) = \gamma_{12} E - \mu I - \mu_{22} I - \gamma_{22} I \\ \dot{R}(t) = \mu_1 E + \gamma_{21} I_A + \gamma_{22} I - \mu R \end{cases} \quad (2.20)$$

Le compartiment I_A désigne les personnes infectés asymptomatiques. Les principales résultats de ce travail montrent que les contacts entre les individus infectés sains et les personnes asymptomatiques/symptomatiques joue un rôle important dans le contrôle du COVID-19, et la valeur seuil correspondante, qui peut induire l'extinction du COVID-19, est obtenue par une analyse mathématique basée sur l'équation différentielle ordinaire et la simulation numérique basée sur les données publiées. De plus, les résultats révèlent que le temps passé à l'extérieur (sans respect des mesures du confinement) des individus infectés symptomatiques/asymptomatiques n'a aucune influence sur le contrôle théorique du COVID-19. Ainsi cette étude donne quelques méthodes théoriques pour contrôler le COVID-19 et quelques suggestions raisonnables pour les décideurs politiques. Cependant, elle utilise la distance de contact moyenne dans les processus de modélisation et ne considère pas les effets temporels et spatiaux sur la distance de contact. D'autre part, ce modèle considère la transmission respiratoire et la transmission par contact uniquement dans les processus de modélisation, et ne prend pas en compte la transmission par aérosol, ce qui peut conduire à une sous estimation du dangerosité du maladie.

- ▷ Bhadauria *et al.* [30] ont présenté un modèle mathématique non linéaire (modèle SIQ), pour étudier l'impact du confinement sur la propagation du COVID-19. Le modèle est donné par le système suivant :

$$\begin{cases} \frac{dS}{dt} = A - \frac{\beta SI}{\alpha + \eta I} - \mu S + (1 - \theta)m \\ \frac{dI}{dt} = \frac{(1-k)\beta SI}{\alpha + \eta I} - \mu I - \delta_1 I - \sigma I - \alpha_1 I + \theta m \\ \frac{dQ}{dt} = \frac{k\beta SI}{\alpha + \eta I} + \sigma I - \mu Q - \delta_2 Q - \alpha_2 Q \end{cases} \quad (2.21)$$

La dynamique de transmission du virus est représenté par la figure 2.5 :

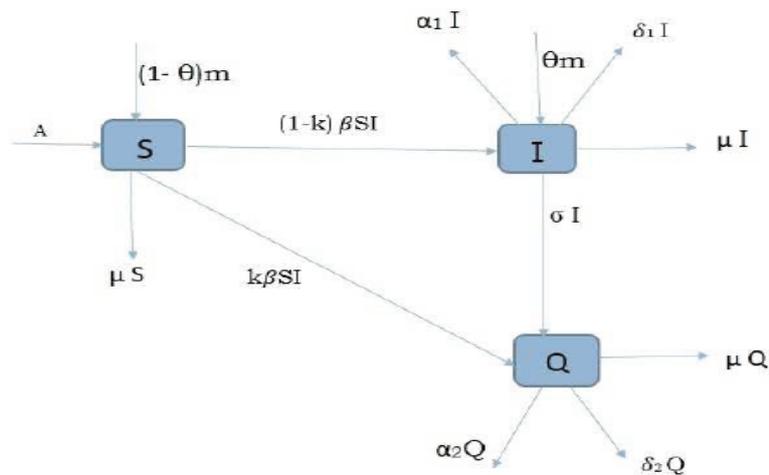


FIGURE 2.5 – Diagramme de transmission du modèle SIQ.

(Le compartiment "Q" désigne la fraction des individus confinés). L'étude du modèle proposé est divisé en deux phases ; état de confinement complet et un confinement partiel. Une analyse qualitative du modèle dans les deux cas est effectuée et le taux de reproduction de base est déterminé par l'approche du nouvelle génération. De plus, pour justifier les résultats analytiques, une simulation numérique est effectuée et des graphiques sont tracés. L'étude montre que la population infectieuse augmente significativement avec l'augmentation du nombre d'immigrés en cas d'état de confinement partiel. Mais afin d'éradiquer complètement la maladie, le confinement est la mesure la plus importante à suivre par le gouvernement. Une analyse de sensibilité du nombre de reproduction de base est effectuée. Il est observé à partir de cette analyse que le taux de transmission du virus est le facteur qui influence le plus le nombre de reproduction de base. L'étude révèle aussi qu'un confinement complet est obligatoire pour réduire l'infection à zéro et que le contrôle de l'immigration de la population dans le système est une étape exceptionnelle pour contrôler la propagation de la maladie. Pourtant, le confinement complet n'est pas la solution ultime car il peut provoquer une crise financière dans un proche avenir et il n'est donc pas possible de mettre en oeuvre un confinement complet.

2.3.2 Modèles stochastiques

- ▷ Mahrouf *et al.* [31], ont proposé un nouveau modèle déterministe et son modèle stochastique correspondant pour décrire le comportement du COVID-19 au Maroc. Ces modèles nous fournissent l'évolution et la prédiction des individus à surveiller, à savoir, les cas diagnostiqués positifs, qui peuvent aider à examiner l'efficacité des mesures mises en oeuvre au Maroc, et les différentes formes développées, qui peuvent quantifier la capacité du système de santé publique ainsi que le nombre de décès. Dans un premier lieu, l'étude montre que ces modèles sont mathématiquement et biologiquement bien posés

en prouvant l'existence et l'unicité des solutions. Ainsi le nombre de reproduction de base est calculé en utilisant le principe du "next generation", son expression est donné par :

$$\mathcal{R}_0 = \frac{\beta\epsilon(1-u)}{(1-\alpha)(\eta_s + \mu_s) + \alpha'}$$

Deuxièmement, l'extinction de la maladie a été étudié en fonction du R_0 , les résultats obtenus montrent que si $R_0 < 1$, il y aura une extinction de la maladie, sinon la maladie persiste. Finalement, quelques simulations numériques ont été réalisées pour recueillir des informations afin de pouvoir lutter contre la propagation du virus. La figure 2.6 donne la variation du nombre des infectés ainsi que les cas signalés au Maroc obtenus par les deux modèles.

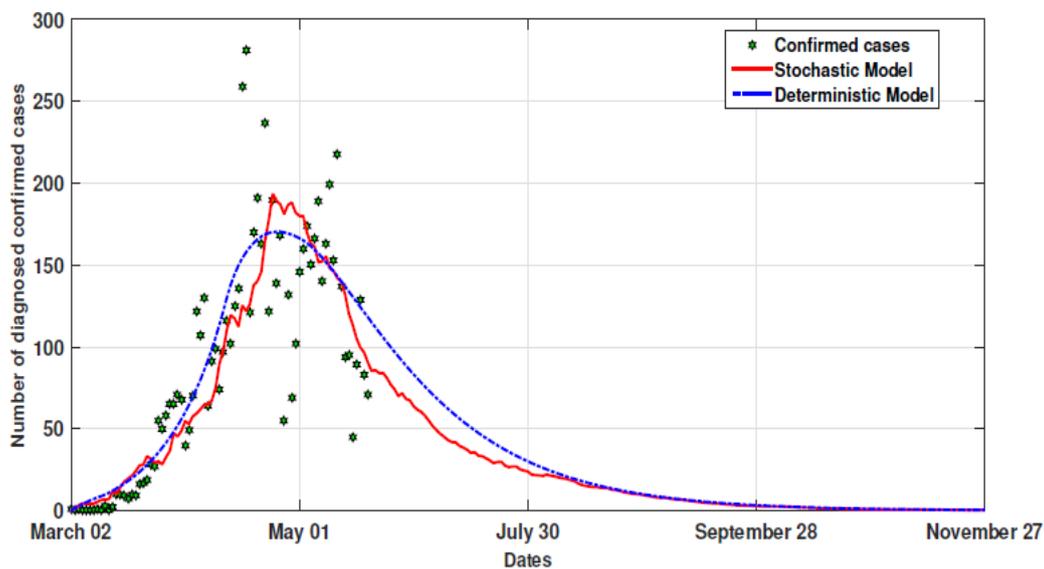


FIGURE 2.6 – Comparaison du comportement déterministe et stochastique avec les cas quotidiens de COVID-19 signalés au Maroc.

On remarque que le comportement du modèle stochastique présente certaines particularités contrairement au modèle déterministe, à savoir que l'amplitude de son pic est plus élevée et la convergence vers l'éradication est plus rapide.

2.4 CONCLUSION

Dans ce chapitre, nous avons donné quelques définitions à propos de la pandémie de Coronavirus. Par la suite, nous avons rappelé quelques paramètres épidémiologiques permettant la description de la propagation de cette maladie. Dans la dernière partie, nous avons donné une revue de la littérature sur les modèles mathématiques modélisant la pandémie du COVID-19.

Etude de la propagation de la pandémie COVID-19 : Le cas du MAROC

3.1 Introduction

La pandémie de COVID-19 est toujours en cours dans plusieurs pays, y compris le Maroc. Cela présente un défi pour les autorités médicales et de santé publique. Cette épidémie doit être bien contrôlée, c'est pourquoi la première étape à franchir est d'obtenir des informations précises sur la prédiction de la durée finale de cette épidémie. Bien que le processus de prédiction sera plus difficile car il est le résultat d'une série complexe d'actions et d'interactions entre une population et un transmetteur de maladie. En outre, d'autres facteurs influencent les prédictions comme l'effet du confinement, les politiques de dépistage des suspects et les vaccinations, car ils ont un impact sur la circulation du virus dans le future. Tous les pays ont pris des mesures pour contrôler la propagation de l'épidémie, vue que sa fin n'est pas connue.

Dans le cadre de la lutte contre ce virus, les modèles mathématiques nous aident à comprendre sa propagation et permettent d'évaluer les effets des interventions entreprises pour la limiter, visant aplatir la courbe épidémique dans le but de la maintenir sous la capacité du système de santé. La simulation numérique permet par la suite, d'explorer la gamme de comportements générés par les équations qui composent le modèle étudié. Il s'agit de tester et d'évaluer des stratégies d'action. En effet, en épidémiologie, un modèle de propagation peut participer à la définition de stratégies de vaccination ou d'isolement des malades.

Au niveau de ce chapitre, nous visons mener une étude expérimentale à base de la simulation numérique, en utilisant le logiciel MATLAB, pour étudier la propagation de COVID-19 au Maroc. Nous proposons une étude comparative entre deux modèles mathématiques, à savoir le modèle de croissance logistique et le modèle SIR simple, en les testant sur des jeux de données réels. L'objectif de notre étude, étant de prédire l'évolution temporelle de la pandémie au Maroc pour différentes périodes à partir de 12 Mars 2020 jusqu'à 31 Mars 2021, en se basant sur des

indicateurs épidémiologiques comme la date du pic épidémique (date dans laquelle le nombre de cas quotidien atteint sa valeur maximale) et la taille finale de l'épidémie (le nombre final des cas infectés).

3.2 La situation épidémique au Maroc

Devant l'évolution rapide de la situation épidémiologique internationale du COVID-19, l'OMS l'a décrétée « Urgence de Santé Publique de Portée internationale » le 30 janvier 2020, puis pandémie le 12 mars 2020. Dès l'alerte initiale relative à ce nouveau virus, le Maroc a entamé le processus de préparation pour y faire face et a développé son plan national de veille et de riposte au COVID-19 (voir le Portail Officiel du Coronavirus au Maroc :

"<http://www.covidmaroc.ma/Pages/Accueilfr.aspx>", le site web de coronavirus-statistiques :

"[https://www.coronavirus-statistiques.com/stats-pays/coronavirus-nombre-de-cas-maroc/La situation épidémique au Maroc](https://www.coronavirus-statistiques.com/stats-pays/coronavirus-nombre-de-cas-maroc/La-situation-epidémique-au-Maroc)" et aussi le site Web « Coronavirus pandemic :daily updated re-search and data » "<https://ourworldindata.org/coronavirus>").

Au niveau de l'étude expérimentale présenté dans ce chapitre, la situation épidémique au Maroc a été analysée à l'aide des données portant sur les cas infectés cumulés collectés auprès du site web "<https://ourworldindata.org/coronavirus>" [32]. Nous présentons une étude comparative entre les deux modèles mathématiques « le modèle de croissance logistique et le modèle SIR simple ». A travers cette étude, nous visons prédire l'évolution temporelle de la pandémie au Maroc pour différentes périodes à partir de 12 Mars 2020 jusqu'à fin Mars 2021 pour étendre l'analyse de l'épidémie.

Nous notons que les tests numériques menés au niveau de notre étude, ont été réalisés sur un PC (Intel Core i7-4600U with Intel HD Graphics 4400 (2.1 GHz Up to 3.3 GHz with Intel Turbo Boost Technology, 4 MB cache, 2 cores)), sous Window 7, en utilisant le logiciel MATLAB R2016b.

3.3 Simulation 1 : Modèle logistique

3.3.1 Estimation des paramètres

D'après la description du modèle logistique présentée dans le chapitre 1, nous rappelons que :

La variation des cas cumulés des infectés pendant une épidémie est donnée par l'équation [33] :

$$\frac{dC}{dt} = rC \left(1 - \frac{C}{K} \right) \quad (3.1)$$

où $r > 0$ le taux d'infection, et $K > 0$ la taille finale de l'épidémie. D'après la description du modèle logistique dans le chapitre 1.

La solution de l'équation 3.1 est donnée par :

$$C(t) = \frac{K}{1 + A \exp(-rt)} \quad (3.2)$$

avec $A = \frac{K-C_0}{C_0}$, supposons que $K \gg C_0$, on obtient donc $A \gg 1$, d'après 3.2 on a :

$$C(t) = \frac{K e^{rt}}{e^{rt} + A} = \frac{C_0}{1 - C_0/K} \frac{e^{rt}}{1 + e^{rt}/A} \approx C_0 e^{rt} \quad (3.3)$$

Cela signifie que tant que la valeur de C est encore faible, la population augmentera de façon exponentielle. De plus, si cette valeur se rapproche de la valeur K ($C \rightarrow K$), alors le taux de croissance de la population est proche de zéro ($dC/dt \rightarrow 0$) ce qui entraînera un arrêt progressif de la croissance de la population.

Le taux de croissance $\frac{dC}{dt}$ atteint son maximum lorsque $\frac{d^2C}{dt^2} = 0$, de cette condition, et en utilisant l'équation 3.2 nous obtenons que le pic de taux de croissance se produit à l'instant :

$$t_p = \frac{\ln A}{r}$$

à ce moment, le nombre de cas est :

$$C_p = \frac{K}{2}$$

et le taux de croissance est :

$$\left(\frac{dC}{dt} \right)_p = \frac{rK}{4}$$

L'équation 3.2 est caractérisé par les deux limitations suivantes :

$$\lim_{t \rightarrow +\infty} C(t) = K$$

et

$$\lim_{t \rightarrow -\infty} C(t) = 0$$

Pour répondre à la question sur le temps de doublement T_d , c'est-à-dire, la durée de temps nécessaire pour doubler le nombre de cas, on résolve l'équation suivante :

$$C(t + T_d) = 2C(t)$$

on obtient :

$$T_d = \frac{\ln 2}{r} - \frac{1}{r} \ln \left(e^{-rt} - \frac{1}{A} \right) - t$$

Lorsque $t \rightarrow t_p = \frac{\ln A}{r}$, i.e. $C \rightarrow K/2$, alors $t_d \rightarrow \infty$. Pour $C \geq K/2$ le temps de doublement perd son sens.

Les paramètres r , A et K peuvent être estimés à partir de la population $C(t)$ en trois dates différentes mais également espacées. C_{k-2m} , C_{k-m} , et C_k avec $k \approx n$ et $m \approx n/2$, où n représente le nombre total de jours.

$$K = \frac{C_{k-m} (C_{k-2m}C_{k-m} - 2C_{k-2m}C_k + C_{k-m}C_k)}{C_{k-m}^2 - C_kC_{k-2m}} \quad (3.4)$$

$$A = \frac{(C_k - C_{k-m})(C_{k-m} - C_{k-2m})}{C_{k-m}^2 - C_kC_{k-2m}} \left[\frac{C_k (C_{k-m} - C_{k-2m})}{C_{k-2m} (C_k - C_{k-m})} \right]^{(k-m)/m} \quad (3.5)$$

$$r = \frac{1}{m} \ln \left[\frac{C_k (C_{k-m} - C_{k-2m})}{C_{k-2m} (C_k - C_{k-m})} \right] \quad (3.6)$$

3.3.2 Résultats et Discussion

La simulation et la mise en oeuvre pratique du modèle logistique sera faite en appliquant le code MATLAB réalisé par Milan Batista [34]. Sur la base des résultats de prédiction des cas épidémiques au Maroc, les résultats d'estimation de ce modèle montrent que le nombre total de cas confirmés de COVID-19 est d'environ 8248, 437.303, 519526 et 499.872 (voir le tableau 3.1). De plus, les résultats d'estimation du taux d'infection r sont de 0,086, 0,029, 0,031 et 0,032 (1/jour), il est indiqué que les cas infectés au Maroc ont diminué. Ensuite, les résultats estimés du pic épidémique calculé à partir des premiers cas infectés par COVID-19 sont respectivement le 28 avril 2020, le 2 novembre 2020 pour les deux premières périodes, et 4 novembre 2020 pour les deux dernières périodes de temps considérés. La première courbe dans les figures 3.1, 3.2, 3.3 et 3.4 représente les résultats de prédiction du modèle de logistique pour les cas infectés cumulés, ces prédictions sont comparés aux données réelles, tandis que les deux lignes rouges entre les lignes noires représentent la limite d'erreur de prédiction. La ligne verte en pointillé prédit la taille totale des cas infectés dans la phase finale de l'épidémie. La ligne rouge croise la courbe des cas cumulés pour prédire le pic de l'épidémie.

Période	Pic épidémique t_p	Taux d'infection r	La taille estimée de K	Paramètre A
12 Mars-9 juin	28 Avril 2020	0.086	8248	61.147
12 Mars-31 Octobre	2 Novembre 2020	0.029	437303	892.56
12 Mars- 28 Janvier 2021	4 Novembre 2020	0.031	519526	1639.1
12 Mars- 31 Mars 2021	4 Novembre 2020	0.032	499827	2158.5

TABLE 3.1 – Les résultats des paramètres estimés du modèle logistique pour différentes périodes à partir de 12 Mars 2020 jusqu'à 31 Mars 2021.

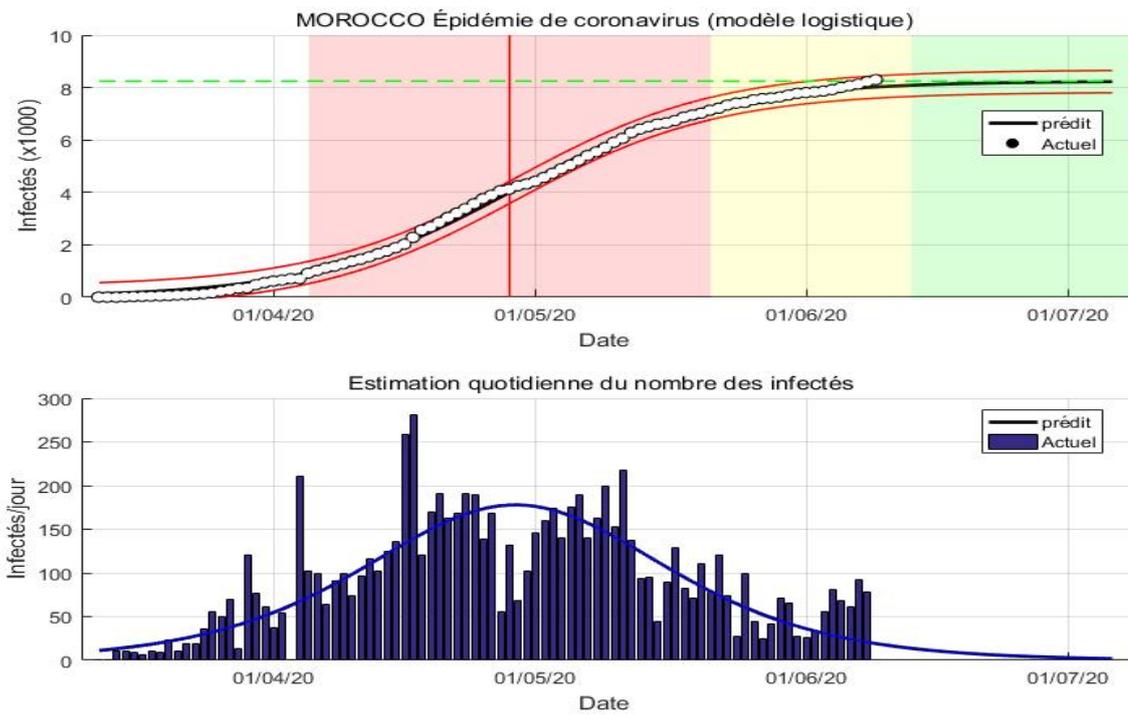


FIGURE 3.1 – Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 9 juin 2020.

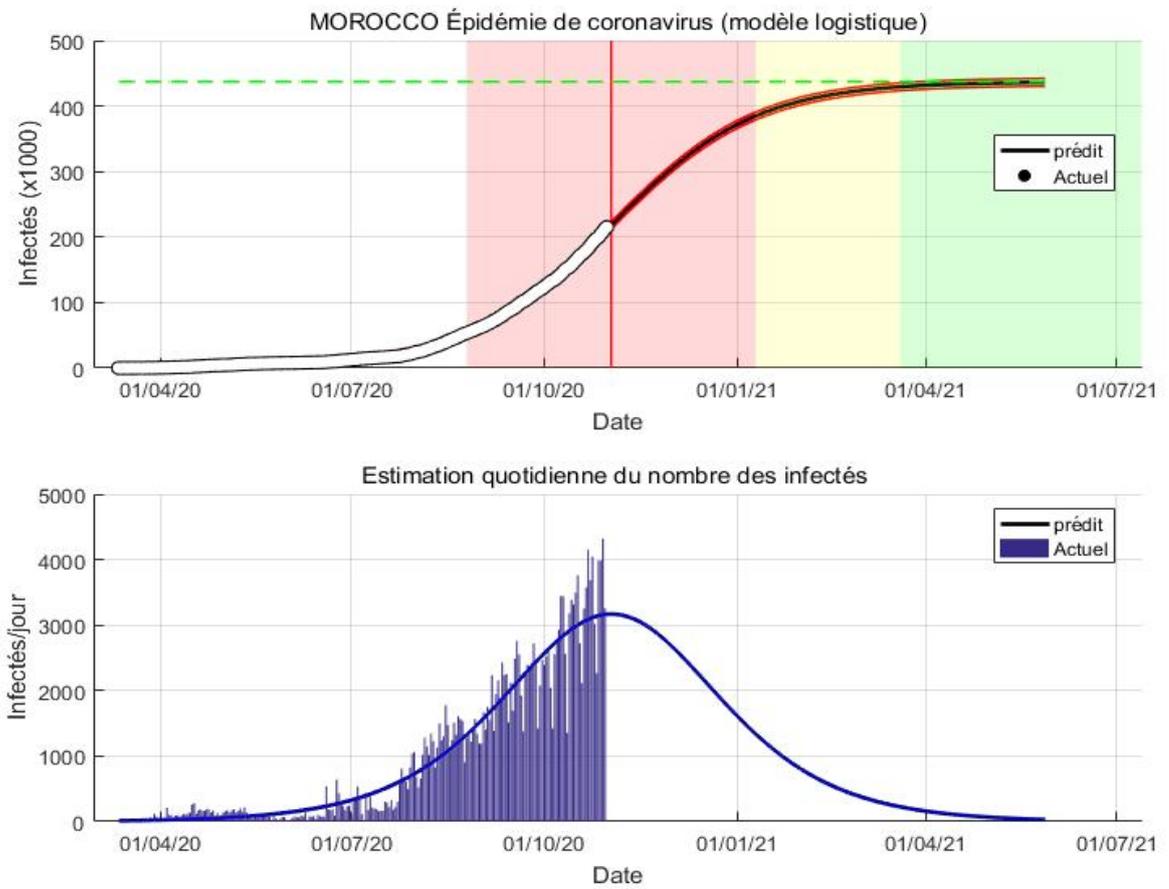


FIGURE 3.2 – Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 31 Octobre 2020.

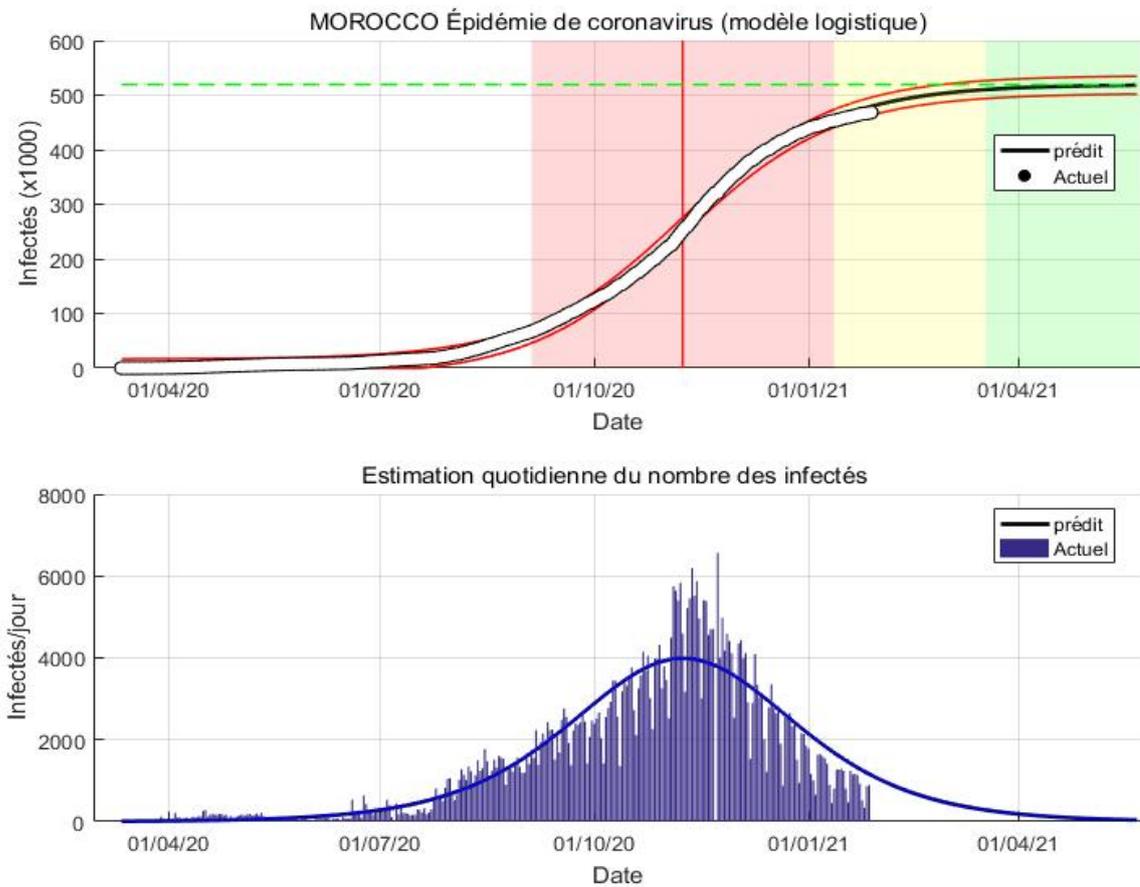


FIGURE 3.3 – Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 28 Janvier 2021.

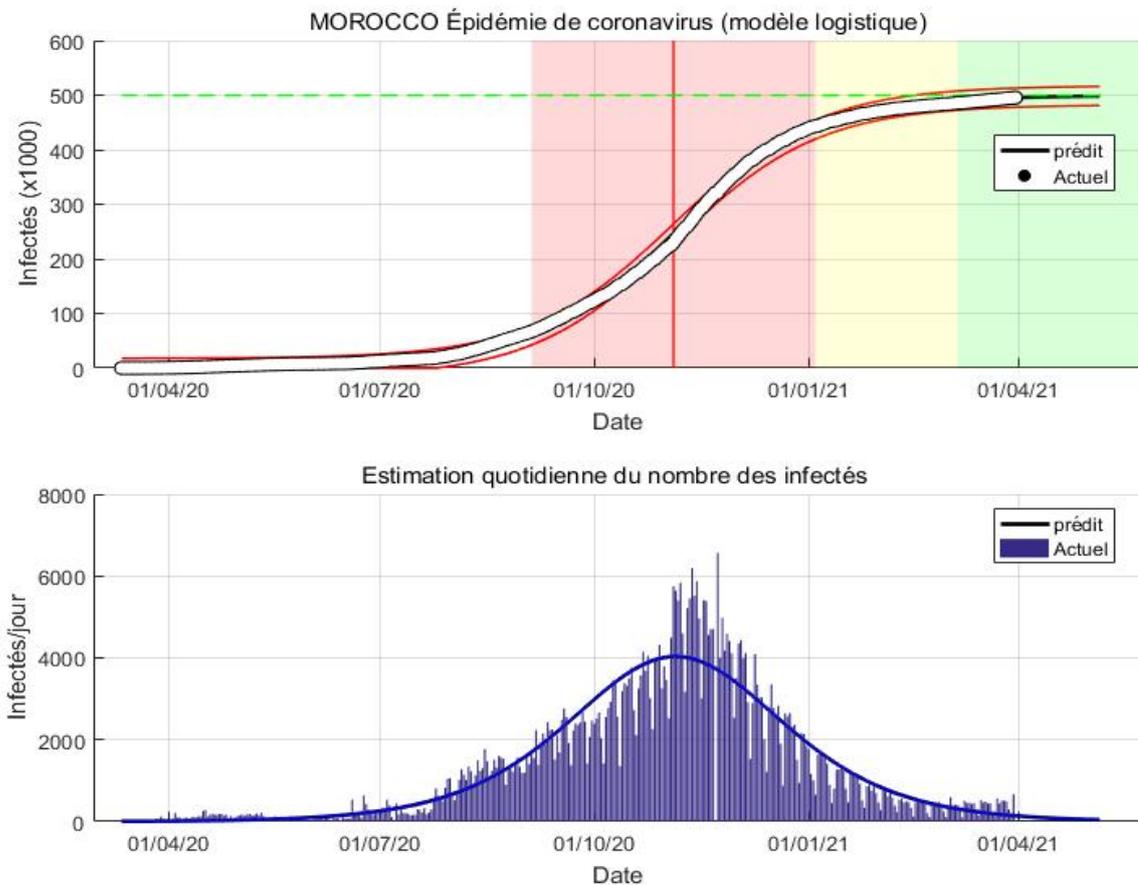


FIGURE 3.4 – Les résultats de prédiction du modèle logistique pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars au 31 Mars 2021.

3.4 Simulation 2 : Modèle SIR

3.4.1 Estimation des paramètres

Le modèle SIR décrit au premier chapitre est l'un des modèles compartimentés les plus simples. Le modèle se compose de trois compartiments : S pour le nombre d'individus susceptibles, I pour le nombre d'infectieux et R pour le nombre d'individus guéris ou décédés. Ce modèle est raisonnablement prédictif pour les maladies infectieuses qui se transmettent d'humain à humain, et où la guérison confère une résistance durable, comme le Covid-19. Le système SIR peut être exprimé par l'ensemble d'équations différentielles ordinaires suivants :

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \quad (3.7)$$

$$\frac{dI}{dt} = -\frac{\beta SI}{N} - \gamma I \quad (3.8)$$

$$\frac{dR}{dt} = \gamma I \quad (3.9)$$

avec $N = S + I + R$, et β et γ correspondent au taux de transmission et de guérison respectivement. Deux observations peuvent être faites :

- ▷ Notons tout d'abord qu'à partir de l'équation $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ il s'ensuit que $S(t) + I(t) + R(t) = N$. Notez que cette relation implique qu'il suffit d'étudier l'équation pour deux des trois variables.
- ▷ Deuxièmement, la dynamique de la classe infectieuse dépend du nombre de reproduction de base R_0 . Ce ratio signifie biologiquement que la maladie s'éteint s'il est inférieur ou égal à 1, alors que si ce nombre est supérieur à 1 la maladie restera dans la population.

Les conditions initiales sont :

$R(0) = R_0, S(0) = S_0, I(0) = I_0$. L'élimination de I des équations 3.7 et 3.9 donne

$$S = S_0 \exp\left[-\frac{\beta}{\gamma N}(R - R_0)\right]$$

A partir de cette équation, notons que lorsque $t \rightarrow \infty$, le nombre de susceptibles S_∞ :

$$S_\infty = S_0 \exp\left[-\frac{\beta}{\gamma N}(R_\infty - R_0)\right] \quad (3.10)$$

où R_∞ est le nombre final d'individus guéris. Le nombre final d'individus infectés étant censé être 0, rappelons que $N = S + I + R$, par conséquent $N = S_\infty + R_\infty$. A partir de ceci et 3.10, l'équation pour R_∞ est :

$$R_\infty = N - S_0 \exp\left[-\frac{\beta}{\gamma N}(R_\infty - R_0)\right]$$

Pour utiliser le modèle, les paramètres β et γ , et les valeurs initiales S_0 et I_0 sont estimés à partir des données disponibles. Nous pouvons estimer les paramètres et les valeurs initiales en minimisant la différence entre le nombre réel et prévu de cas.

3.4.2 Résultats et Discussion

La simulation et la mise en oeuvre pratique du modèle SIR se fera suivant la description [35] et en utilisant le code MATLAB de Milan Batista [36]. Pour estimer les paramètres, on utilise la fonction *fminsearch* de MATLAB. Pour résoudre l'équation du modèle, on utilise la fonction *ode45* de MATLAB.

Les résultats du modèle SIR montrent que le nombre final de cas confirmés estimés de COVID-19 pour les quatre périodes considérés est d'environ 9201, 644.078, 525.006 et 503.172 pour les quatre périodes (voir tableau 3.2). Ainsi, pendant la période du confinement le taux de reproduction était de 0,48, après il a dépassé la valeur 1 (1,14), et puis sa valeur est diminué jusqu'à et 0,93, ce qui montre que l'épidémie est en cours de diminution. Ensuite, les résultats estimés du pic épidémique calculé à partir des premiers cas infectés par COVID-19 sont respectivement le 27 avril 2020, le 24 novembre 2020, 7 novembre 2020 et le 5 novembre 2020 pour les quatre périodes de temps considérées.

Période	Pic épidémique t_p	Taux de transmission β	Taux de guérison γ	Taux de reproduction	Taille finale
12 Mars-9 juin	27 Avril 2020	0.147	0.066	0.48	9201
12 Mars-31 Octobre	24 Novembre 2020	0.1	0.073	1.14	644078
12 Mars- 28 Janvier 2021	7 Novembre 2020	0.291	0.26	0.91	525006
12 Mars- 31 Mars 2021	5 Novembre 2020	0.491	0.459	0.93	503172

TABLE 3.2 – Les résultats des paramètres estimés du modèle SIR pour différentes périodes à partir de 12 Mars 2020 jusqu'à 31 Mars 2021.

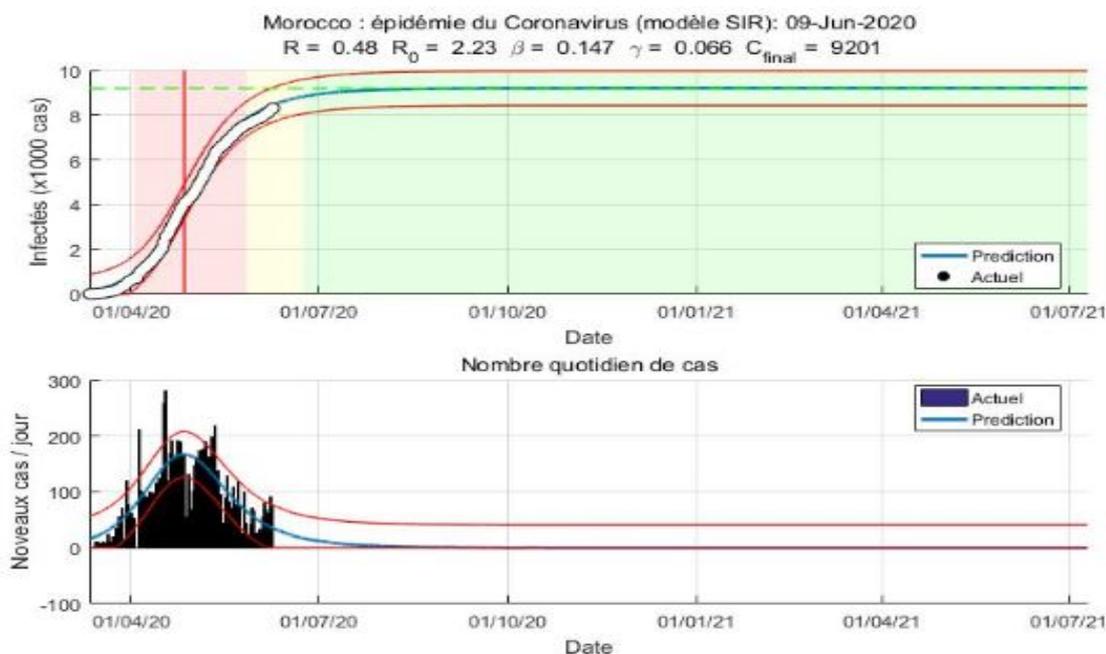


FIGURE 3.5 – Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 9 juin 2020.

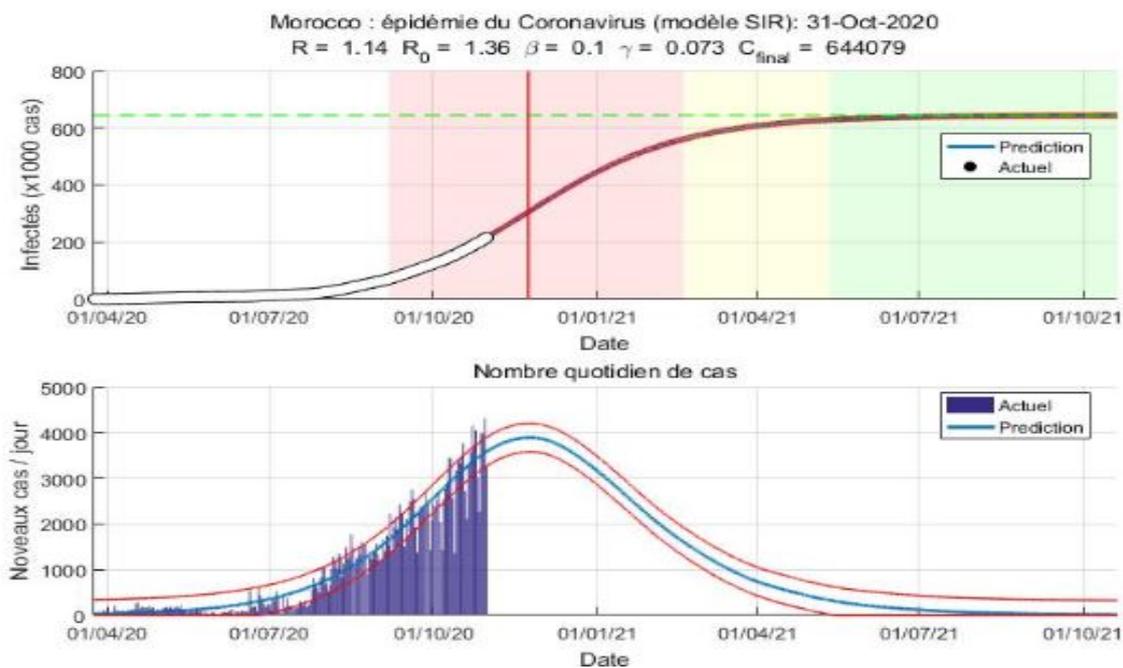


FIGURE 3.6 – Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 mars au 31 Octobre 2020.

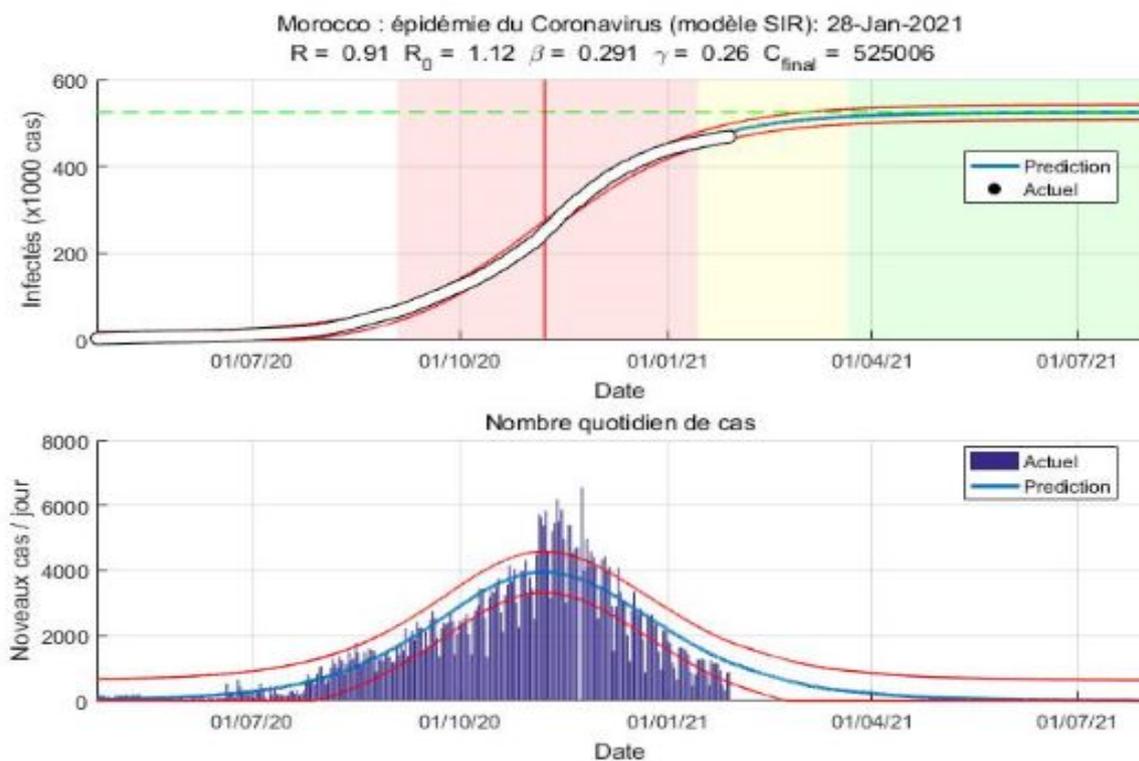


FIGURE 3.7 – Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 28 Janvier 2021.

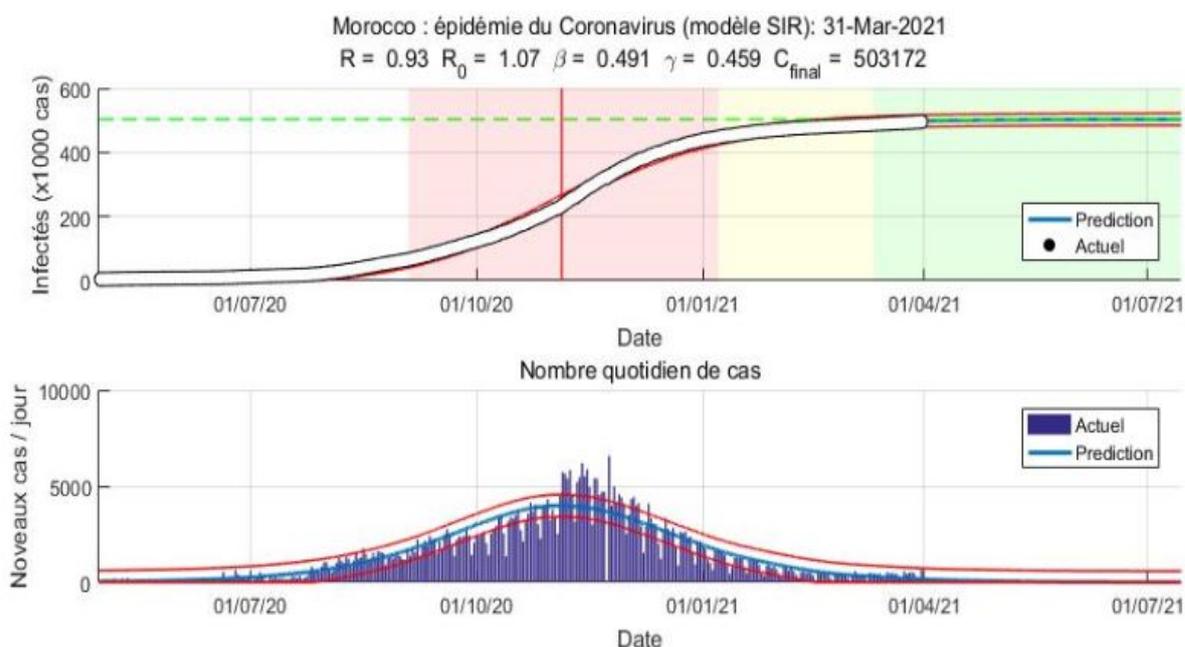


FIGURE 3.8 – Les résultats de prédiction du modèle SIR pour les cas infectés cumulés et quotidiens au Maroc dans la période du 12 Mars 2020 au 31 Mars 2021.

Le tableau 3.3 représente les estimations de la date du fin de l'épidémie au Maroc obtenus par le modèle SIR pour les quatre périodes.

Période	Fin de l'épidémie (5 cas)	Fin de l'épidémie (1 cas)	Durée total de l'épidémie (jour)
12 Mars-9 juin	29 Septembre 2020	4 Novembre 2020	237
12 Mars-31 Octobre	29 Mai 2022	13 Aout 2022	883
12 Mars- 28 Janvier 2021	12 Décembre 2021	6 février 2022	695
12 Mars- 31 Mars 2021	11 Novembre 2021	2 Janvier 2022	661

TABLE 3.3 – Quelques résultats supplémentaires concernant la fin de l'épidémie au Maroc.

3.5 CONCLUSION

Dans ce chapitre, nous avons estimé quelques données associées à la propagation du COVID-19 au Maroc suivant 4 périodes du début de l'épidémie jusqu'à le mois Mars 2021. Cette simulation a été effectuée en utilisant deux types de modèles mathématiques en épidémiologie (le modèle logistique et le modèle SIR), pour chaque modèle les estimations des paramètres sont données. Finalement pour le modèle SIR une prédiction de la durée de l'épidémie est aussi estimée.

CONCLUSION GENERALE

La modélisation mathématique est de plus en plus utilisée en épidémiologie comme un outil essentiel d'aide à la décision sanitaire. Dans ce travail, nous avons présenté d'une part, des outils mathématiques pour étudier théoriquement et numériquement certains phénomènes épidémiologiques à savoir la diffusion et la propagation d'une épidémie. En effet, au niveau des deux premiers chapitres de ce mémoire, en plus de la théorie et les indicateurs représentatifs en épidémiologie et la modélisation statistique, nous avons rappelé le principe et principaux paramètres d'un ensemble de modèles mathématiques bien connus pour la modélisation mathématique en épidémiologie et aussi pour le COVID-19. Et d'autre part, nous nous sommes intéressés à la modélisation et l'étude de l'évolution du COVID-19 au Maroc, en utilisant deux approches logistiques et déterministes. Plus précisément, ces deux modèles mathématiques ont été utilisés pour simuler numériquement la propagation du COVID-19 au Maroc. C'est l'objectif du troisième chapitre.

Comme perspectives futures de notre travail, nous proposons de :

- Etudier la complexité du problème et étendre la période d'étude pour simuler numériquement la propagation de la pandémie du COVID-19 au Maroc, durant des périodes à partir du 1er Avril 2021 ;
- Faire une simulation numérique pour d'autres modèles mathématiques déterministes et dynamiques modélisant le COVID-19, et à base d'autres indicateurs et paramètres ;
- Proposer un modèle de programmation mathématique pour modéliser le COVID-19, aussi bien qu'une modélisation graphique ;
- Utiliser les outils de la Recherche Opérationnelle pour aide à la décision sanitaire.

Bibliographie

- [1] Dedet J.P. Les épidémies - 2e édition de la peste noire à la covid-19. Dunod, 2021.
- [2] F. Bertrand, E. Claeys, and M. Maumy-Bertrand. *Modélisation statistique par la pratique avec R : Cours et exercices corrigés*. Dunod, 2019.
- [3] N. Marie. *Introduction à la modélisation probabiliste et statistique*. Ellipses, 2018.
- [4] WikiStat. Introduction à la modélisation statistique, <http://wikistat.fr>.
- [5] Quatre types d'analyse de données. <https://fr.jobs-job.com/four-types-of-data-analysis-12196>.
- [6] La Rédaction TechTarget. Analyse statistique. <https://whatis.techtarget.com/fr/definition/Analyse-statistique>.
- [7] P. Billeter. Recherche opérationnelle et statistique. *Revue économique et sociale : bulletin de la Société d'Etudes Economiques et Sociales*, 1961.
- [8] R. Faure, B. Lemaire, and C. Picouveau. *Précis de recherche opérationnelle-7e éd. : Méthodes et exercices d'application*. Dunod, 2014.
- [9] T. Ancelle. *Statistique Epidémiologie, SCIENCES FONDAMENTALES*. Ellipses, 2002.
- [10] P. Vallin and D. Vanderpooten. *Aide à la décision : une approche par les cas*. Ellipses Paris, 2000.
- [11] E. Bourgkard, V. Demange, and C. Aubry. L'épidémiologie en santé au travail. Technical report, 2004.
- [12] F. Brauer, C. Castillo-Chavez, and Z. Feng. *Mathematical models in epidemiology*, volume 32. Springer, 2019.
- [13] M. Martcheva. *An introduction to mathematical epidemiology*, volume 61. Springer, 2015.
- [14] N. Bacaër. *A short history of mathematical population dynamics*. Springer Science & Business Media, 2011.
- [15] D. Moulay. *Modélisation et analyse mathématique de systèmes dynamiques en épidémiologie. Application au cas du Chikungunya*. PhD thesis, Université du Havre, 2011.

- [16] F. Younsi. Mise en place d'un système d'information décisionnel pour le suivi et la prévention des épidémies. *PhD diss. University of Lyon*, 2016.
- [17] C. Olivier Sigata. Modèles Stochastiques En Épidémiologie : Cas des modèles SIR. mémoire de Master de l'Université Félix Houphouët Boigny, 2014.
- [18] O. Koutou. Modélisation mathématique et contrôle optimal de la dynamique de transmission du paludisme. DEA de l'Université Polytechnique de Bobo-Dioulasso, 2017.
- [19] I. EL Berrai. Contribution à l'analyse déterministe et stochastique du modèle épidémiologique sir. 2016.
- [20] Groupe de modélisation de l'équipe ETE. Nombre de reproduction du covid-19.
- [21] Jane M. Heffernan, R. Smith, and L. Wahl. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4) :281–293, 2005.
- [22] P. Van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2) :29–48, 2002.
- [23] Dutta Sanchari S. Période d'incubation covid-19. 30 août 2009. [https://www.news-medical.net/health/Coronavirus-Incubation-Period-\(French\).aspx](https://www.news-medical.net/health/Coronavirus-Incubation-Period-(French).aspx).
- [24] World Health Organization et al. Estimating mortality from covid-19 : scientific brief, 4 august 2020. Technical report, World Health Organization, 2020.
- [25] H. K Ebraheem, N. Alkhateeb, H. Badran, and E. Sultan. Delayed dynamics of sir model for covid-19. *Open Journal of Modelling and Simulation*, 9(2) :146–158, 2021.
- [26] R. Yafia. Modélisation et dynamique en épidémiologie, covid19 avec confinement et isolation-application au cas du maroc. 2020.
- [27] M. Serhani and H. Labbardi. Mathematical modeling of covid-19 spreading with asymptomatic infected and interacting peoples. *Journal of Applied Mathematics and Computing*, 66(1) :1–20, 2021.
- [28] S. Alizon, B. Reyné, and C. Selinger. *Modélisation de l'épidémie de COVID-19 : modèle SEAIR*. PhD thesis, 2020.
- [29] Z. Ma, S. Wang, X. Lin, X. Li, X. Han, H. Wang, and H. Liu. Modelling for the covid-19 with the contacting distance. 2021.
- [30] A. Bhadauria, R. Pathak, and M. Chaudhary. A siq mathematical model on covid-19 investigating the lockdown effect. *Infectious Disease Modelling*, 6 :244–257, 2021.
- [31] M. Mahrouf, A. Boukhouima, H. Zine, EM. Lotfi, DFM. Torres, and N. Yousfi. Modeling and forecasting of covid-19 spreading by delayed stochastic differential equations. *axioms* 2021, 10, 18. 2021.

- [32] H. Ritchie, E. Ortiz-Ospina, D. Beltekian, E. Mathieu, J. Hasell, B. Macdonald, C. Giattino, C. Appel, L. Rodés-Guirao, and M. Roser. Coronavirus pandemic (covid-19). *Our World in Data*, 2020. <https://ourworldindata.org/coronavirus>.
- [33] M. Batista. Estimation of the final size of the coronavirus epidemic by the sir model. *Online paper, ResearchGate*, pages 1–12, 2020.
- [34] M. Batista. fitvirus. 2020. <https://www.mathworks.com/matlabcentral/fileexchange/74411-fitvirus>), MATLAB Central File Exchange. Retrieved June 22, 2021.
- [35] M. Batista. Estimation of the final size of the coronavirus epidemic by the logistic model, updated on march 25, 2020. pages 1–15, 2020.
- [36] M. Batista. fitviruscv19. 2020. <https://www.mathworks.com/matlabcentral/fileexchange/74658-fitviruscv19> MATLAB Central File Exchange. Retrieved June 22, 2021.