

Pandæsim: un simulateur stochastique de propagation épidémique

Biology 2020, 9(9), 299

Patrick Amar

(traduction post-éditée par N. Bacaër, suggestions d'amélioration : nicolas.bacaer@ird.fr)



Résumé simple

Afin d'étudier l'efficacité des contre-mesures utilisées contre la pandémie de Covid-19 à l'échelle d'un pays, nous avons conçu un modèle et développé un programme de simulation efficace basé sur un cadre de simulation stochastique discret bien connu avec une localisation spatiale standard à gros grains. Notre approche particulière nous permet également de mettre en œuvre des résolutions continues déterministes d'un même modèle. Nous l'avons appliqué à l'épidémie de Covid-19 en France où des contre-mesures de confinement ont été utilisées. Avec la méthode discrète stochastique, nous avons trouvé de bonnes corrélations entre les résultats de la simulation et les statistiques recueillies auprès des hôpitaux. En revanche, l'approche déterministe continue conduit à des résultats très différents. Nous avons proposé une explication basée sur le fait que les effets de discrétisation sont élevés pour les petites valeurs, mais faible pour les grandes valeurs. Lorsque nous ajoutons la stochasticité, cela peut expliquer les différences de comportement de ces deux approches. Ce système est un outil de plus pour étudier différentes contre-mesures aux épidémies, du confinement à la distanciation sociale, ainsi que les effets de la vaccination de masse. Il pourrait être amélioré en incluant la possibilité de réinfection individuelle.

Résumé

De nombreuses méthodes ont été utilisées pour modéliser la propagation de l'épidémie. Elles comprennent des systèmes d'équations différentielles ordinaires pour des environnements globalement homogènes et des systèmes d'équations aux dérivées partielles pour prendre en compte la localisation spatiale et l'inhomogénéité. Des systèmes d'équations différentielles stochastiques ont été utilisés pour modéliser la stochasticité inhérente aux processus de propagation épidémique. Dans notre étude de cas, nous voulions modéliser le nombre d'individus dans différents états de la maladie et leur localisation dans le pays. Parmi les nombreuses méthodes existantes, nous avons utilisé notre propre variante de l'algorithme stochastique bien connu de Gillespie, ainsi que la méthode des sous-volumes pour prendre en compte la localisation spatiale. Notre algorithme nous permet de passer facilement de la simulation discrète stochastique à la résolution déterministe continue utilisant des valeurs moyennes. Nous avons appliqué nos approches à l'étude de l'épidémie de Covid-19 en France. La version discrète stochastique de Pandæsim a montré de très bonnes corrélations entre les résultats de la simulation et les statistiques recueillies auprès des hôpitaux, à la fois au jour le jour et sur les chiffres globaux, y compris les effets du confinement. De plus, nous avons mis en évidence des différences de comportement intéressantes entre les méthodes continue et discrète qui peuvent survenir dans certaines conditions particulières.

1. Introduction

La France a été touchée par l'épidémie de SRAS-CoV-2 probablement début janvier 2020, le premier cas ayant été signalé le 24 janvier [1], et est entrée en confinement le 17 mars 2020 [2]. En réponse à la réduction attendue du nombre de cas, le gouvernement français a assoupli les restrictions de confinement le 11 mai 2020 et les a de nouveau assouplies le 25 mai (sauf en Ile-de-France, où la densité de population est très élevée). Ces mesures ont été prises pour arrêter la croissance exponentielle du nombre de cas, comme observé précédemment en Chine [3, 4].

Le nombre de reproduction de base R_0 nous indique le nombre moyen de nouvelles infections causées par un individu infectieux et décrit la croissance exponentielle de l'épidémie [5]. Si R_0 est supérieur à 1, l'épidémie se

propagera; sinon, quand R_0 est inférieur à 1, la maladie disparaîtra progressivement [6]. Par rapport au R_0 de H1N1 (1,25) [7] le nombre de reproduction de Covid-19 indique une transmission potentielle terrible. Le R_0 a été estimé à 2,2 [8], 3,8 [9] et 2,68 [10, 11] par de nombreuses sources de recherche différentes dans le monde. L'Organisation mondiale de la santé (OMS) a publié une estimation R_0 de 1,4 à 2,5 [12].

De nombreuses approches ont déjà été utilisées pour modéliser l'épidémie de Covid-19 à l'aide de modèles de compartiments et d'équations différentielles ordinaires déterministes (EDO) [13, 14] ainsi que pour estimer les effets des mesures de lutte sur la dynamique de l'épidémie [15]. Ces approches particulières donnent de bons résultats, mais elles ne prennent pas en compte la nature stochastique ou les aspects spatiaux du mécanisme de propagation. Cependant, les équations différentielles stochastiques (EDS) ont été utilisées avec succès pour s'attaquer aux aspects stochastiques de la propagation épidémique [16, 17, 18, 19]. Plus récemment, des modèles épidémiques multirégionaux utilisant des modèles discrets et continus, prenant en compte l'efficacité du contrôle des mouvements ont été publiés [20, 21], ainsi que des modèles EDS multirégionaux [22]. Des modèles stochastiques basés sur l'épidémiologie économique ont été appliqués à l'épidémie de Covid-19, par exemple en Corée du Sud, pour déterminer le stock optimal de vaccins et l'efficacité de la distanciation sociale [23]. Des approches utilisant des systèmes basés sur des agents ont également été utilisées pour modéliser à la fois les caractéristiques stochastiques et spatiales de la propagation épidémique [24, 25]. Dans les méthodes basées sur des agents, le nombre d'instructions machine nécessaires pour chaque pas de temps, par rapport à la taille des données (complexité algorithmique), est au mieux proportionnel au nombre d'agents. Ceux qui utilisent un agent par individu peuvent avoir besoin d'une puissance de calcul élevée lorsqu'ils sont utilisés sur de grandes populations. Ces approches sont souvent appliquées à des zones plus petites (villes principalement) que l'ensemble du pays, et / ou utilisent un agent pour modéliser un ensemble d'individus (100 dans [24]).

Les méthodes centrées sur la population ont une complexité algorithmique qui ne dépend pas de la taille de la population, mais du nombre de règles considérées à chaque itération (par exemple, le nombre de réactions pour les systèmes de biochimie). Lorsqu'elles sont utilisées sur de grandes populations, ces méthodes sont beaucoup plus efficaces que les méthodes centrées sur l'entité, mais elles ne prennent pas en compte la localisation spatiale. Nous avons adopté ici un modèle hybride dérivé de la méthode des sous-volumes qui ajoute des capacités de localisation spatiale à gros grains à l'algorithme de simulation stochastique standard (ASS) utilisé, par exemple, dans le domaine de la biochimie. Pour augmenter l'efficacité du calcul, nous avons également utilisé une variante originale [26] de l'algorithme de Gillespie avec saute-tau [27, 28] qui adapte automatiquement la proportion d'aléa par rapport au calcul moyen, à chaque pas de temps. Notre implémentation nous permet de passer facilement de cette variante stochastique de ASS à un solveur continu déterministe (SCD), et donc de comparer les deux méthodes.

Pour tester notre approche, nous l'avons appliquée à l'épidémie de SRAS-CoV-2 en France où des données pertinentes [29, 30] ont été mises à disposition pendant toute la durée de l'épidémie. La plupart des paramètres de simulation que nous avons utilisés ont été obtenus à partir de statistiques rassemblées dans la littérature, comme la proportion de cas nécessitant une hospitalisation et la proportion de formes sévères parmi eux [31, 32] nécessitant des lits en USI (unité de soins intensifs). Le nombre d'individus infectieux et leurs localisations au début de l'épidémie ont été déduits des données statistiques mises à disposition par le gouvernement français et de la littérature [33, 34, 35]. Nous avons utilisé notre outil de simulation pour vérifier les effets des mesures de contrôle sur la dynamique de l'épidémie et comparé les résultats aux données statistiques réelles. Nous avons centré notre étude des impacts de l'épidémie uniquement sur la partie de la population qui se déplace au quotidien: travailleurs, élèves, étudiants, retraités, etc. Les personnes en maison de retraite n'étaient pas prises en compte car leur environnement et leur mode de vie sont très différents.

2. Matériels et méthodes

2.1. Aperçu

À partir d'un état initial connu, nous avons voulu calculer un échantillon stochastique de l'évolution dans le temps du nombre de personnes à chaque état de la maladie. Une transition entre de tels états est souvent décrite par un ensemble de règles probabilistes, ou par un automate stochastique. La propagation de l'épidémie peut être modélisée comme un processus markovien en ce sens que le nombre de personnes dans chaque état au temps $t+\Delta t$ ne dépend que des nombres au temps t (et d'autres variables qui ne dépendent pas de t). Dans la plupart des cas, il n'est pas possible de trouver une solution analytique qui donne ces nombres en fonction du temps. Heureusement, il existe des méthodes numériques itératives. L'une d'elles est l'algorithme de Gillespie, fréquemment utilisé pour trouver les évolutions des quantités d'espèces chimiques $S(t) = \{s_1(t), \dots, s_n(t)\}$ qui peuvent réagir selon les règles

chimiques $R = \{r_1, \dots, r_m\}$ et leur cinétique $K = \{k_1, \dots, k_m\}$. À partir de la valeur initiale $S(0)$ des n espèces, l'algorithme calcule les valeurs au temps $t > 0$ en répétant le processus suivant:

- Avec les quantités $S(t)$, les règles et leur cinétique, calculer stochastiquement à quel moment chaque réaction est déclenchée $\{t_1, \dots, t_m\}$.
- Soit r_i la prochaine réaction: $t_i = \inf\{t_1, \dots, t_m\}$.
- Appliquer r_i , c'est-à-dire mettre à jour le vecteur $S(t_i)$ en diminuant les quantités de substrats de r_i et augmenter les quantités de ses produits.
- Mettre à jour l'heure: $t \leftarrow t_i$.

Cet algorithme donne une trajectoire stochastique exacte du système, mais peut être lent lorsque certaines réactions sont rapides. Ces réactions rapides seront souvent déclenchées, de sorte que l'incrément de temps à chaque itération sera petit et le nombre d'itérations par seconde élevé. Pour diminuer le temps de calcul, la méthode saute-tau utilise un pas de temps fixe, τ . A chaque itération, le nombre de fois où chaque réaction est déclenchée pendant l'intervalle de temps τ est estimée stochastiquement sur la base des quantités au temps t . Cette méthode donne une approximation de la trajectoire stochastique du système, qui est précise si τ est petit. La valeur de τ doit être choisie suffisamment grande pour minimiser le nombre d'itérations par seconde, mais pas trop grand pour obtenir une bonne précision. L'algorithme utilisé dans Pandæsim, une variante de la méthode Gillespie saute-tau, est détaillé à la fin de cette section.

Les méthodes centrées sur la population telles que celles présentées ici partagent la même contrainte: les entités évoluant dans l'environnement sont considérées comme réparties de manière homogène dans l'environnement. En d'autres termes, la localisation spatiale n'est pas prise en compte. Les approches centrées sur l'entité, qui calculent le comportement de chaque individu à chaque pas de temps, prennent en compte la localisation spatiale de chaque individu, mais nécessitent beaucoup plus de puissance de calcul. Pour ajouter une localisation spatiale à grain grossier à notre modèle, nous avons partitionné le territoire en sous-régions où un exemple d'ASS centré sur la population est exécutée. Ces exemples utilisent le même pas de temps et sont synchronisées. Les interactions entre les sous-régions sont modélisées en prenant des échantillons stochastiques d'individus qui voyagent entre les sous-régions. Cela se fait à une échelle de temps plus élevée car de tels déplacements sont moins fréquents que les déplacements à l'intérieur de la sous-région d'origine. La plupart des personnes qui voyagent retournent dans leur sous-région d'origine après une période de temps variable. Ainsi, la population de chaque sous-région reste à peu près la même, bien que les gens entrent et sortent de la sous-région. Si cela n'est pas pris en compte dans le modèle, la population de chaque sous-région peut avoir tendance à devenir la même avec le temps. Nous décrivons dans la section suivante comment cette contrainte est implémentée dans notre modèle.

2.2. Modèle Pandæsim

Le territoire étudié est divisé en deux niveaux d'organisation géographique: région et sous-région. Une région contient au moins deux sous-régions, une sous-région n'appartient qu'à une seule région et tout le territoire est couvert (partition). Dans notre étude de cas, la France, le premier niveau est la région administrative, chacune contenant de deux à une douzaine de départements. La France compte 13 régions et 96 départements. Bien entendu, cela peut s'appliquer à n'importe quelle partition d'un territoire. Par exemple, en Angleterre, nous pourrions utiliser les neuf régions pour le premier niveau et les 46 comtés cérémoniels et le Grand Londres pour le deuxième niveau.

La population est divisée en quatre tranches d'âge: 0 à 25 ans, 26 à 50 ans, 51 à 75 ans et plus de 76 ans [36, 37, 38]. Chacune de ces quatre sous-populations a ses propres valeurs pour les paramètres de population (immunité aux infections, taux de déplacement, etc.). Nous avons utilisé un exemple d'un processus de simulation centré sur la population pour chaque sous-région, avec un pas de temps d'une heure. La simulation du niveau supérieur (région) utilise un pas de temps plus grand, un jour, et traite principalement les personnes qui se rendent dans une autre sous-région. Ainsi, la répartition de la population est supposée homogène à l'intérieur de chaque sous-région, mais peut être hétérogène au niveau de la région et donc au niveau de l'ensemble du territoire. Selon l'âge, et à l'exception des personnes malades ou hospitalisées, chaque jour, les personnes ont la probabilité de se déplacer de leur domicile vers un autre lieu appartenant soit à la même sous-région (déplacements locaux), soit à une autre région (déplacements éloignés). Ces probabilités font partie des paramètres de population mentionnés précédemment. Bien entendu, les mesures de contrôle de type quarantaine interdisent tout type de déplacement local ou éloigné; les gens doivent rester dans leurs sous-régions respectives.

Le nombre de personnes de chaque tranche d'âge quittant leur sous-région d'origine est un échantillon stochastique (ou valeur moyenne pour le solveur continu déterministe) d'un pourcentage de la population de cette sous-région. Pour les déplacements locaux, ils sont dispersés en fonction de la population relative de chaque sous-région appartenant à leur région. Les sous-régions les plus peuplées attirent davantage de voyageurs. Pour les voyages à distance, les gens vont de leur région d'origine aux sous-régions les plus peuplées des autres régions, où se trouvent les aéroports et les gares. La même méthode est utilisée pour répartir les voyageurs en fonction des populations relatives de leurs sous-régions de destination. Cette façon de calculer combien d'individus voyagent et où ils vont est un moyen simple de maintenir constante la densité de population de chaque sous-région.

Le modèle centré sur la population de sous-région est une variante du modèle sensible, exposé, infectieux et retiré largement utilisé. Nous avons ajouté deux états: hospitalisé et décédé. Les états exposés et infectieux ont des significations légèrement différentes dans notre modèle; ils ont été renommés asymptomatiques et malades (Figure 1). Contrairement aux personnes malades, qui présentent des symptômes de la maladie, les personnes récemment infectées sont des hôtes asymptomatiques, mais les deux sont infectieuses. Les patients hospitalisés sont également contagieux, mais dans une moindre mesure parce qu'ils sont confinés à l'intérieur de l'hôpital. Les trois flèches en pointillés rouges sur la figure indiquent les sources et cibles potentielles de l'infection. Nous avons supposé que les personnes en état de guérison sont immunisées contre le virus et ne peuvent donc pas être réinfectées [39].

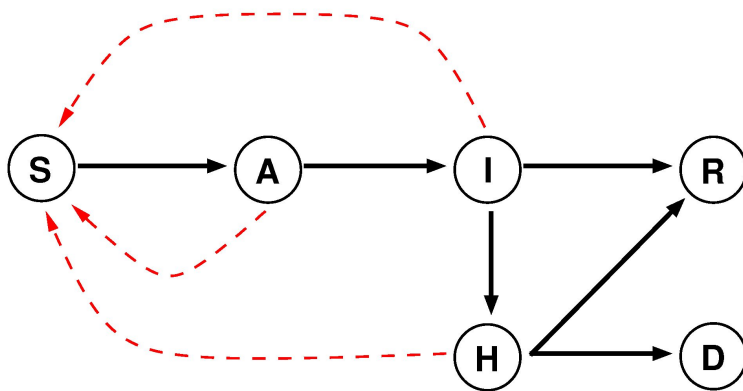


Figure 1. Graphique d'état de l'évolution d'une infection virale. Les états sont: sensible (S), asymptomatique (A), malade (I), hospitalisé (H), guéri (R) et décédé (D). Les flèches noires indiquent les transitions entre les états, et les flèches rouges en pointillé indiquent les infections possibles.

2.3. Données et paramètres de simulation

Une période d'incubation d'environ cinq à six jours avant l'apparition des premiers symptômes a été observée [40, 41]. En conséquence, dans notre modèle, les personnes asymptomatiques sont subdivisées en six sous-catégories selon le nombre de jours depuis la contamination. Une grande majorité des cas, environ 80%, présentent une forme bénigne de la maladie qui n'est probablement même pas signalée. Les autres cas nécessitent une hospitalisation, et parmi eux, de 5% [31] à plus de 15% [32] présentent des formes sévères dans lesquelles les patients doivent être admis en USI. La durée de la maladie, après la période d'incubation, dépend de l'âge du patient et de la gravité de la forme de la maladie. Dans notre modèle, il a été fixé à un maximum de 15 jours, et nous avons donc subdivisé les personnes malades (resp. Hospitalisées) en 15 sous-catégories au plus en fonction du nombre de jours depuis l'apparition des premiers symptômes (resp. hospitalisation). Les personnes atteintes d'infections légères se rétablissent après une période de temps variable stochastiquement (7 à 15 jours) qui dépend de leur âge. La forme sévère de la maladie est (stochastiquement) létale selon un taux variant également avec l'âge du patient. Le solveur déterministe utilise des valeurs moyennes fixes. Tous ces taux, probabilités et durées moyennes sont des paramètres du modèle.

2.4. Algorithme d'évolution

Comme mentionné précédemment, l'algorithme de simulation utilise un pas de temps d'une heure. Il calcule principalement de manière stochastique le vecteur d'état: $V(t) = \{ S(t), E_1(t), \dots, E_5(t), i_1(t), \dots, i_{15}(t), H_1(t), \dots, H_{15}(t), R(t), D(t) \}$, c'est-à-dire le nombre de personnes dans chaque état et sous-catégorie, à chaque pas de temps. Il existe quatre vecteurs d'état, un pour chaque tranche d'âge. Bien entendu, ces quatre vecteurs ne sont pas indépendants puisque quel que soit leur âge, les personnes contagieuses peuvent infecter des personnes sensibles

quel que soit leur âge. Fondamentalement, à partir de la valeur du vecteur d'état au temps t , le processus calcule la nouvelle valeur du vecteur d'état au temps $t + \tau$ (ici $\tau = 1h$). Ainsi, à partir d'une valeur initiale connue du vecteur d'état au temps $t = 0$, nous pouvons obtenir sa valeur à tout moment ($t = t_{fin} > 0$) en répétant ce processus jusqu'à ce que t_{fin} soit atteint, ou jusqu'à ce qu'une valeur spécifique du vecteur d'état soit atteinte. Pandæsim arrête automatiquement la simulation lorsqu'il n'y a plus de personnes infectieuses.

Notre modèle suppose que les gens ont des routines quotidiennes uniformes. Sans mesures spécifiques, l'horaire quotidien commence à 8 heures du matin pour le travail (ou école, université, etc.) avec l'utilisation des transports en commun pendant une heure. Vient ensuite rester au travail trois heures, suivi d'une pause de deux heures à midi, quatre heures l'après-midi au travail, une autre heure dans les transports en commun pour rentrer à la maison et les 13 heures restantes à la maison. Nous avons défini quatre environnements possibles, chacun ayant sa probabilité de contagion: domicile, transports en commun, lieu de travail et restaurant. Ces paramètres ont des valeurs par défaut qui reflètent les concentrations locales de personnes: très faible à la maison, plus élevée au travail et au restaurant et beaucoup plus élevée dans les transports en commun. Pour réduire le nombre de paramètres, nous avons utilisé la même valeur pour le lieu de travail et le restaurant.

De nombreux types de mesures peuvent être utilisés pour ralentir la propagation de l'épidémie; nous avons mis en œuvre deux exemples de telles mesures:

- Quarantaine douce: les gens n'utilisent pas du tout les transports en commun et ne vont pas au restaurant pendant la pause de midi.
- Quarantaine complète: Cela correspond à ce qui s'est réellement passé en France; les gens étaient confinés chez eux sauf une heure de marche par jour dans les zones à faible densité de population (les parcs publics, les forêts, etc. étaient interdits). Encore une fois, pour réduire le nombre de paramètres, nous avons supposé que la probabilité de contagion lors de la balade était la même qu'au travail. Cela nous a également permis de prendre en compte les simulations faites pour obtenir de la nourriture dans des endroits plus peuplés comme les épiceries ou les supermarchés.

À partir d'un état initial (nombre de personnes contagieuses dans chaque sous-région), l'algorithme de simulation itère le processus suivant à chaque pas de temps jusqu'à ce que l'épidémie se termine ou que la durée maximale de la simulation soit atteinte (par défaut à 720 jours).

- Premièrement, le taux d'infection au temps t , $I_{rt}(t)$, est calculé comme le produit du taux d'infection quotidien global, $G_{dri}(t)$, par le facteur d'infection du lieu actuel (domicile, lieu de travail, transports en commun) $L_{inf}(t)$. Ce taux d'infection $I_{rt}(t)$ est utilisé de la même manière que la propension dans l'ASS standard.

$$I_{rt}(t) = G_{dri}(t) L_{inf}(t). \quad (1)$$

- Ensuite, pour chacune des quatre tranches d'âge, le solveur continu déterministe calcule le nombre moyen d'individus de cet âge qui passeront de l'état sensible à l'état asymptomatique, $AvNew_{asympt}$, en tant que produit de la population dans cet état et du taux d'infection au temps t :

$$AvNew_{asympt}[age](t) = population_{susceptible}[age](t) I_{rt}(t) \quad (2)$$

Le solveur discret stochastique (SDS) calcule des nombres entiers stochastiques de telle sorte que, sur le long terme, ils atteindront les mêmes valeurs que le solveur continu. Même lorsque la population est un nombre entier d'individus, ce produit, $AvNew_{asympt}$, est généralement un nombre à virgule flottante car le taux d'infection est lui-même un nombre à virgule flottante. Ce nombre a une partie entière (≥ 0) et une partie fractionnaire (entre 0 et 1). Le nombre (discret) de nouveaux hôtes asymptomatiques est alors calculé comme la partie entière du nombre moyen, plus 1 si un nombre aléatoire uniforme pris dans l'intervalle $[0...1]$ est en dessous de la partie fractionnaire:

$$New_{asympt}[age](t) = \lfloor AvNew_{asympt}[age](t) \rfloor + 1 \text{ si } rnd \leq \text{Frac}(AvNew_{asympt}[age](t)) \\ 0 \text{ autrement} \quad (3)$$

Comme la différence est de 0,5 sur la moyenne, plus la valeur est élevée, plus l'impact relatif de cette discrétisation stochastique est faible et le résultat équivaut à une approche discrète moyennée. A l'inverse,

plus la valeur est basse, plus la discrétisation stochastique devient importante. Ce mécanisme permet au simulateur de choisir automatiquement la meilleure stratégie pour s'adapter à la plage de valeurs de la population [26].

- Enfin, lorsque l'heure actuelle indique le début d'un nouveau jour, $t \equiv 0 \pmod{24}$, les individus de chaque état restent dans le même état mais décalés d'un jour, ou passent à un autre état. Toutes les transitions d'états sont calculées de manière stochastique par le SDS (ou de manière déterministe par le SCD) en utilisant la méthode décrite précédemment.
 - La population dans l'état asymptotique qui a en moyenne atteint la limite des 5/6 jours est déplacée au premier jour de l'état malade.
 - Selon le paramètre de durée de la maladie par tranche d'âge, une partie de la population en état de maladie est déplacée vers l'état hospitalisé ou rétabli. Les autres sont restés malades un jour de plus.
 - Selon le paramètre de gravité de la maladie par tranche d'âge, une proportion de la population à l'état hospitalisé est déplacée vers l'état décédé ou guéri. Les autres restent à l'état hospitalisé un jour de plus.

Le taux d'infection quotidien global est alors simplement calculé en multipliant la constante de propagation du virus, K_{prop} , par la proportion de la population contagieuse totale:

$$G_{dri}(t) = K_{prop} \frac{\sum_{age} pop_{contagieuse}(t)}{pop_{tot}} \quad (4)$$

En ajustant les résultats de la simulation après le début du confinement aux données recueillies à partir des statistiques hospitalières, nous avons trouvé empiriquement une bonne estimation de K_{prop} pour le SARS-CoV-2 à 0,75. Nous pensons qu'en utilisant Pandæsim pour modéliser un autre type d'épidémie, seule cette constante, ainsi que les paramètres de gravité, doivent être modifiés.

3. Résultats

Nous avons appliqué notre outil de simulation à l'épidémie de SRAS-CoV-2 en France. Nous avons utilisé les partitions de région et de département du pays pour les régions et sous-régions de notre modèle. La plupart des paramètres que nous avons utilisés sont issus de la littérature et des données statistiques mises à disposition par le gouvernement français. Quelques autres ont été obtenus empiriquement, principalement le nombre de personnes contagieuses dans chaque région au début des simulations, et la constante de propagation du SRAS-CoV-2. Les valeurs par âge du pourcentage de létalité [42], de la durée de la maladie et du pourcentage de voyageurs locaux et éloignés sont présentées sur le tableau A2, les différents taux de contamination sur le tableau A3 et le nombre initial de personnes contagieuses dans chaque département dans le tableau A1 à l'annexe A.

Afin de tester notre algorithme centré sur la population, nous avons d'abord effectué des simulations sans contre-mesures et sans possibilité de déplacement, local ou distant. Ces simulations ont été exécutées en utilisant successivement le solveur discret stochastique et le solveur continu déterministe. Lorsque le nombre initial de personnes contagieuses était relativement élevé, par exemple, dans la sous-région du Val-de-Marne (180), les résultats pour les deux solveurs étaient presque identiques: 5207 décès pour la moyenne de 1000 simulations stochastiques et 5204 décès pour une simulation déterministe (Figure 2 et Figure 3). L'écart type pour ces 1000 essais est passé de ≈ 2 au début des simulations (avec quelques dizaines de décès) à ≈ 41 au pic de l'infection (quelques milliers de décès), puis ≈ 5 à la fin. Les mêmes types de résultats sont apparus pour les personnes malades avec la valeur maximale de l'écart type de ≈ 2300 atteint au 90e jour, avec 137 381 personnes malades.

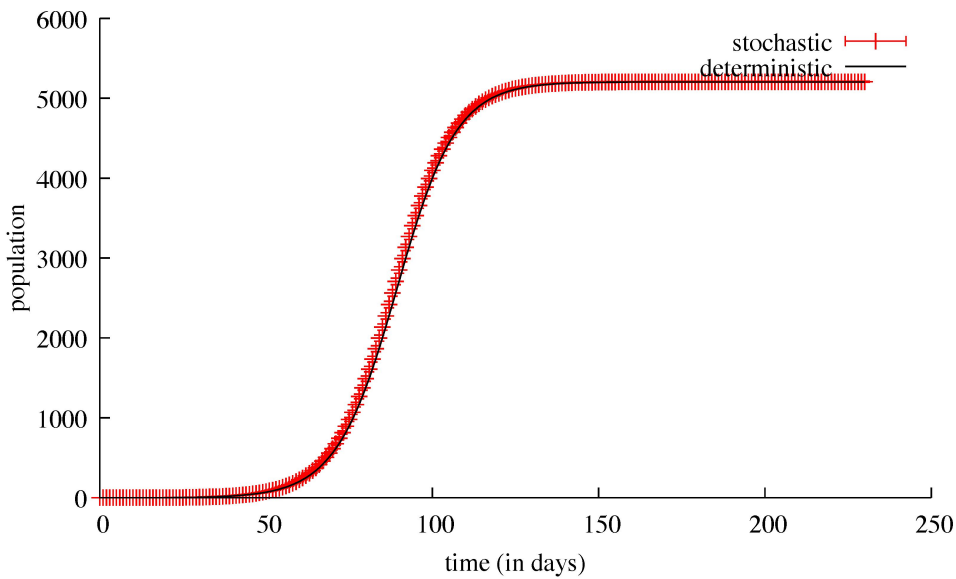
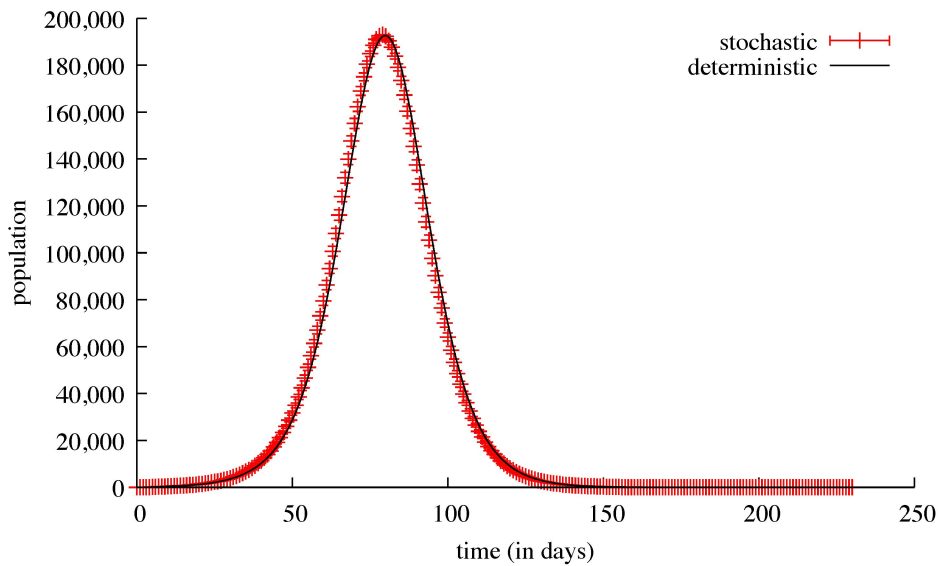


Figure 2. Résultats des simulations de la sous-région du Val-de-Marne, sans possibilité de déplacement à l'extérieur ou à l'intérieur de cette sous-région. Les résultats de la résolution continue déterministe sont représentés par une courbe noire. Les moyennes et écarts types de 1000 simulations discrètes stochastiques du même modèle sont représentés par des barres rouges. La vue du dessus montre le nombre de personnes malades, tandis que la vue du bas montre le nombre cumulé de décès.

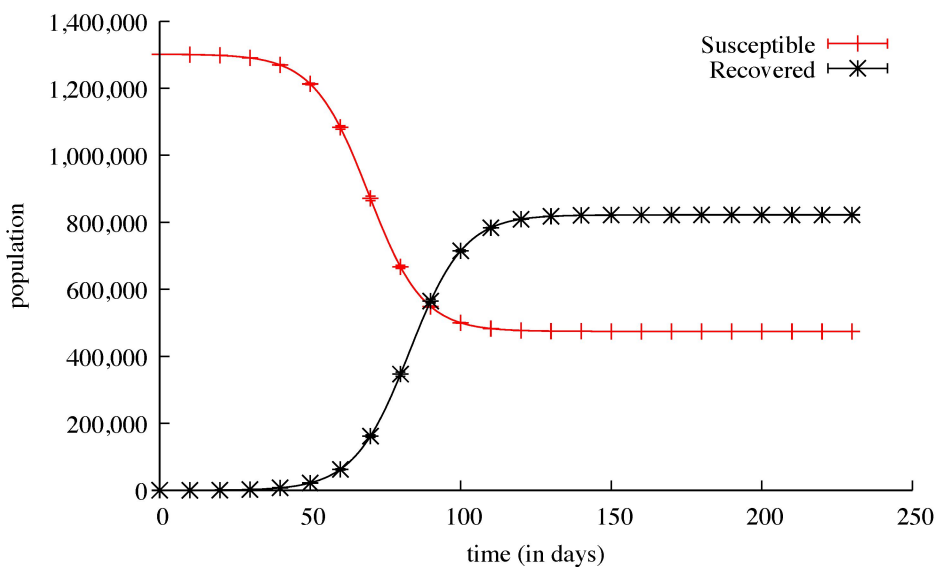


Figure 3. Les moyennes et les écarts types de 1000 simulations discrètes stochastiques du même modèle. La population sensible est représentée en rouge, la population guérie en noir, les deux avec des barres d'erreur tous les 10 jours.

En revanche, lorsque le nombre initial de personnes contagieuses était faible, comme dans le Loiret (2), le SCD n'a pas trouvé de décès, alors que 1000 essais du SDS ont montré deux comportements distincts; 127 de ces essais ont montré les mêmes résultats que le SCD, aucun décès à la fin de l'épidémie. Les 873 autres essais ont pris une autre direction menant à 4499 décès en moyenne avec un écart type de ≈ 264 (figure 4). Les raisons de cette apparente incohérence seront expliquées dans la section discussion.

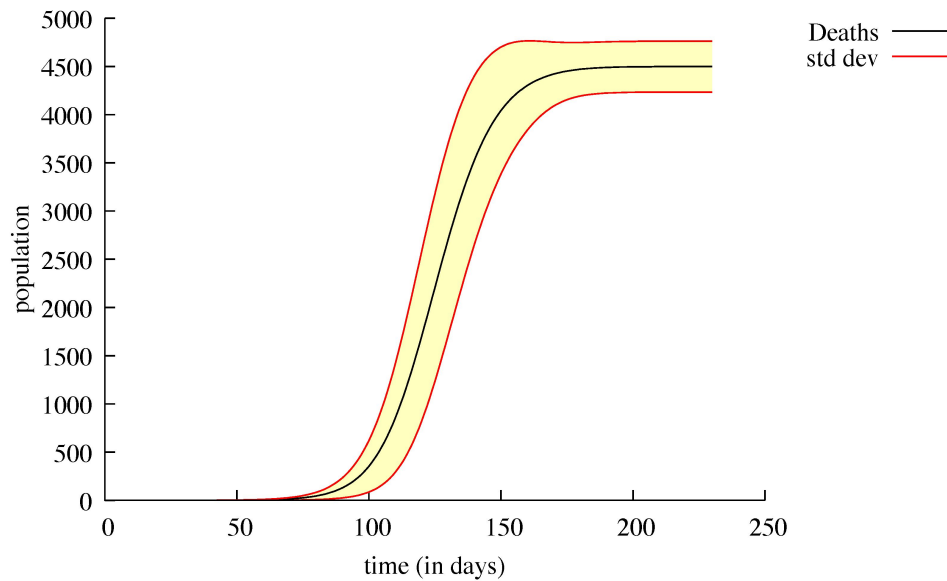
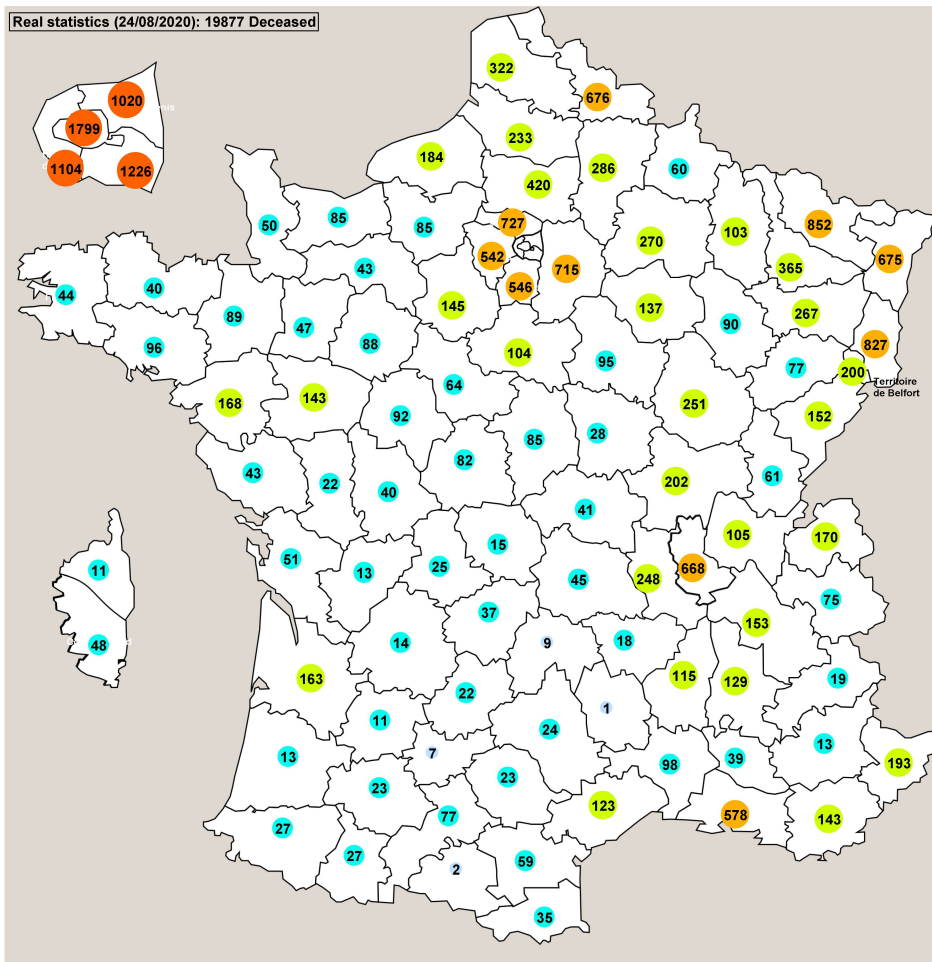
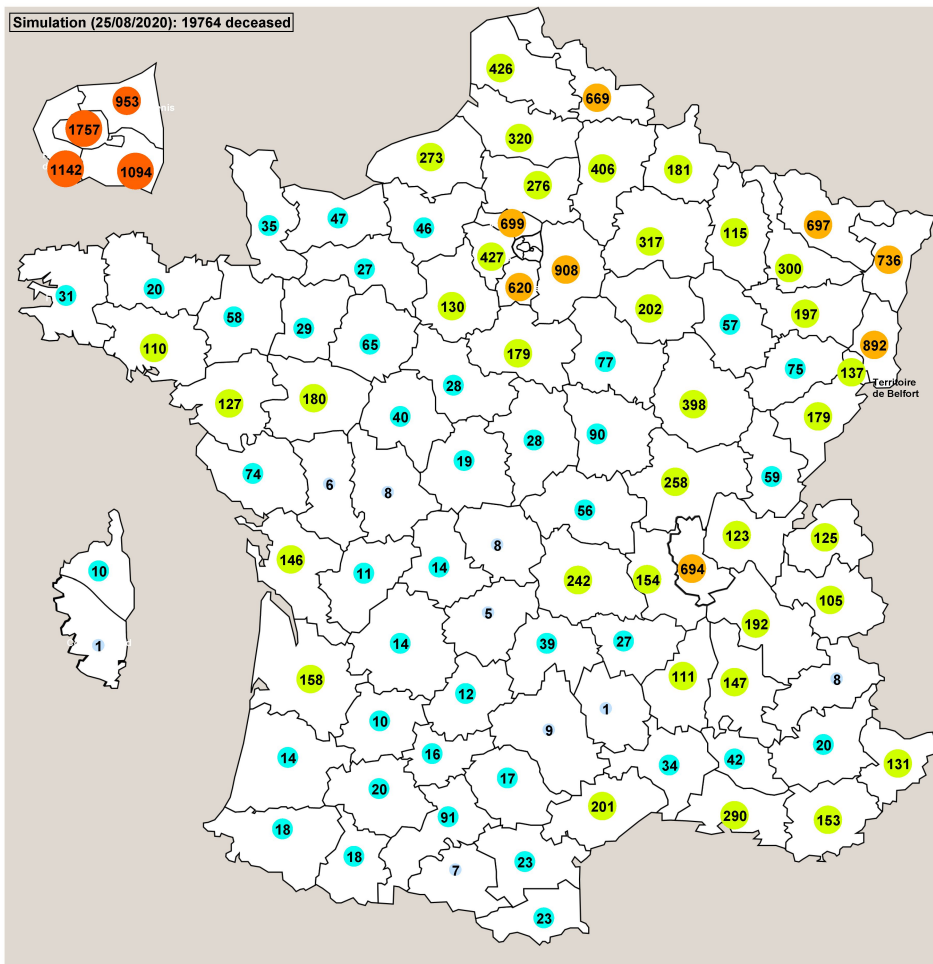


Figure 4. Nombre de décès issus des 873 simulations (plus de 1000) de la sous-région du Loiret, sans possibilité de déplacement à l'extérieur ou à l'intérieur de cette sous-région. La moyenne est représentée en noir; l'écart type est la zone jaune entourée par les lignes rouges.

En utilisant la contre-mesure appliquée en France (confinement), les simulations nous ont montré rétrospectivement que la date probable à laquelle il y avait un total de 897 personnes contagieuses en France (début des simulations) était approximativement fin janvier 2020. Ceci est en corrélation avec la période de temps lorsque la première personne décédée a été signalée (24 janvier). La vue de la fenêtre principale de Pandaesim représentée sur la figure 5 montre le nombre réel de personnes décédées dans chaque département. La carte illustrée à la figure 6 affiche les valeurs moyennes de 500 exécutions d'une simulation stochastique. Les résultats globaux sont très proches, 19 877 pour les statistiques réelles et 19 764 pour la valeur moyenne des simulations. Les résultats département par département sont également assez proches, sauf pour quelques départements, mais les ordres de grandeur sont plus ou moins identiques.





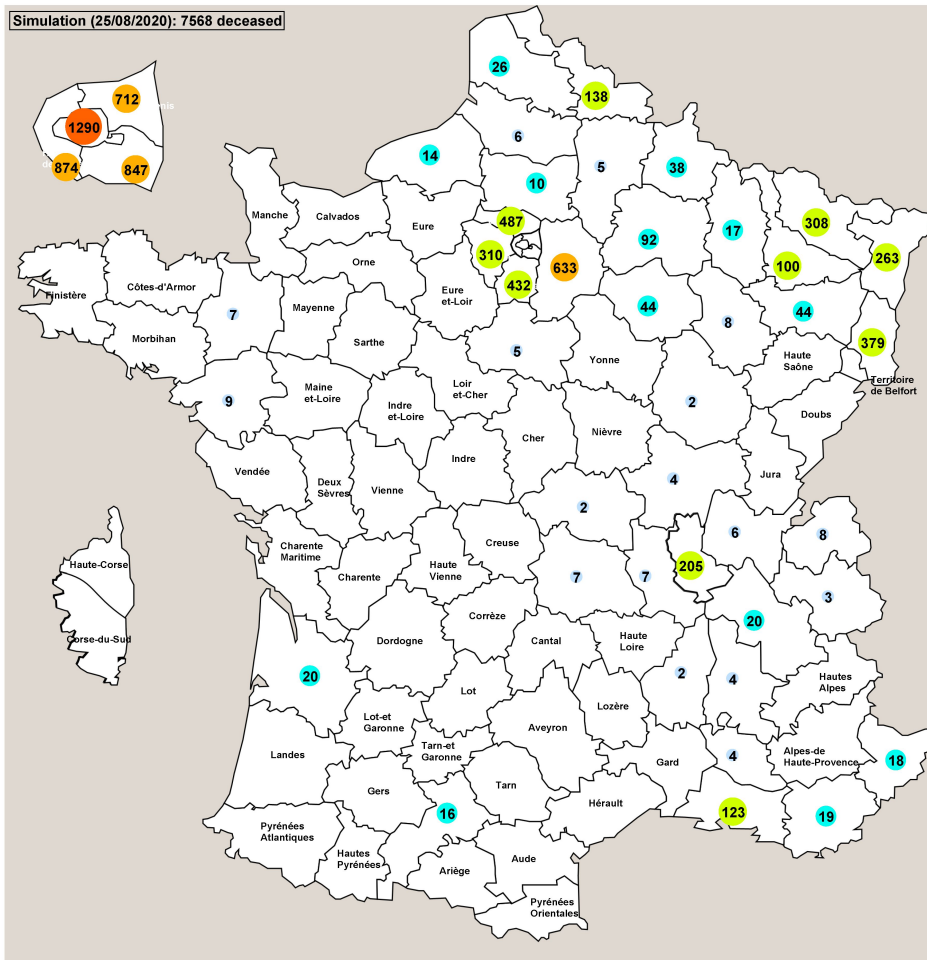


Figure 7. Résolution continue déterministe du modèle en utilisant les mêmes valeurs de paramètres que celles des simulations stochastiques illustrées à la figure 6 . Lorsque le nombre de décès est égal à 0, le nom du département est affiché à la place.

4. Discussion

Nous avons développé un modèle hybride et un programme de simulation dérivé de modèles standards et de techniques de simulation largement utilisés dans les domaines de la propagation épidémique et de la biochimie. Notre approche a utilisé une variante originale de l'ASS de Gillespie avec tau-saut, où l'algorithme interne peut être facilement commuté du discret stochastique au continu déterministe. Cela nous a permis de comparer ces deux méthodes de simulation. Pour tester notre approche, nous l'avons appliquée à l'épidémie de SRAS-CoV-2 en France, pour laquelle des données pertinentes étaient disponibles.

Nous avons également testé les conséquences et l'efficacité de la contre-mesure de confinement appliquée en France pendant 55 jours. Afin de gagner en localisation spatiale mais avec un algorithme efficace centré sur la population où la population était supposée être homogène, nous avons partitionné le territoire en unités relativement petites pour lesquelles un exemple de la simulation centrée sur la population a été exécutée. Les mouvements de populations entre ces unités ont été pris en compte à une échelle plus élevée, avec un pas de temps plus grand.

Nous avons d'abord testé un exemple de notre algorithme centré sur la population, où aucune contre-mesure n'a été utilisée. En utilisant chaque méthode (SDS et SCD) avec les mêmes valeurs de paramètres, nous avons comparé les résultats dans deux situations différentes: (i) avec un nombre modérément élevé, et (ii) avec un nombre très faible de personnes initialement contagieuses. Lorsque les chiffres étaient relativement élevés, les résultats des deux méthodes étaient très similaires. Cela n'était pas surprenant car à chaque pas de temps la valeur absolue de l'incrément calculé par chaque méthode doit être significativement supérieure à 1, et l'arrondi stochastique à l'entier inférieur ou supérieur ne peut pas être relativement très éloigné de la valeur en virgule flottante calculée par la méthode continue . Cependant, lorsque les nombres sont faibles, la valeur absolue ajoutée au prochain pas de temps n'est qu'un peu supérieure à 0, et donc l'arrondi stochastique à 0 ou à 1 change radicalement la trajectoire future. Ceci est particulièrement important dans ce cas précis où les populations connaissent une croissance exponentielle. Cela peut ressembler à un comportement chaotique car une petite

différence dans les conditions initiales peut conduire à des futurs très différents, mais lorsque les chiffres augmentent, l'importance de cet effet de changement est atténuée.

Nous avons utilisé de nombreux lots de simulations avec au départ seulement deux individus contagieux dans la sous-région. Les résultats de 100, 200, 500 et 1000 simulations ont montré à peu près les mêmes proportions de cas, $\approx 12\%$, se terminant sans aucun décès, tandis que le reste du lot a convergé vers environ 4500 décès. Le même modèle utilisant le SCD ne montre aucun décès. Nous pensons que ce comportement est une conséquence d'une bifurcation due à la non-linéarité élevée du système. Lorsque le nombre d'individus contagieux est inférieur à un certain seuil, la contagion a tendance à s'estomper, mais si ce nombre dépasse le seuil, il y a une sorte de rétroaction positive qui l'augmente jusqu'à ce qu'une partie suffisamment importante de la population totale soit éliminée. Si nous supposons que le nombre initial d'individus contagieux dans notre exemple (2) est inférieur au seuil, le résultat affiché par le SCD est donc correct. En raison à la fois de ses incréments discrets et de son comportement stochastique, le SDS peut parfois calculer une trajectoire qui dépasse le seuil et bascule dans l'autre sens.

Afin d'approfondir l'étude de ce phénomène de bifurcation, nous avons tenté de trouver la valeur approximative du seuil. Nous avons d'abord utilisé le SCD avec le nombre initial d'individus contagieux variant de 1 à 20. Aucun décès n'a été trouvé jusqu'à 15; puis 38 décès de 16 à 18; et 4508 décès pour 19 ans et plus. Ensuite, nous avons fait les mêmes tests avec 200 simulations SDS, en comptant le nombre de simulations menant à zéro décès, et dans l'autre cas, le nombre moyen de décès. Avec au départ 1 à 5 individus contagieux, le nombre de simulations sans décès est passé de 70 à 2; avec six individus et plus initialement contagieux, plus aucune simulation ne conduit à zéro décès. Pour toutes les simulations ne conduisant pas à zéro décès, le nombre moyen de décès était de ≈ 4514 . Le seuil pour le SDS est quelque part en dessous de 5. Comme prévu, cette valeur est très faible.

Ensuite, nous avons testé l'ensemble du simulateur avec tous les processus centrés sur la population, fonctionnant indépendamment pendant 24 pas dans chaque sous-région, puis synchronisés en échangeant une partie de chaque population de manière stochastique ou déterministe. Encore une fois, selon le type de solveur choisi et pour les raisons évoquées précédemment, les résultats étaient différents mais pas trop. Le nombre de personnes voyageant d'une sous-région donnée étant une (petite) fraction de la population totale de cette sous-région, les conséquences en termes de propagation de l'infection sont très dépendantes de la valeur elle-même: inférieure à 1, elle est amplifiée par le traitement stochastique, ou bien lissée avec le calcul continu.

Les résultats globaux et les résultats locaux des sous-régions se sont avérés très similaires en utilisant les deux méthodes. Cela peut s'expliquer par le fait que les sous-régions à faible population contagieuse initiale «bénéficient» de la migration de personnes contagieuses en provenance de sous-régions plus peuplées, et comme aucune contre-mesure n'est appliquée, le nombre de personnes contagieuses augmente rapidement au-dessus du seuil. La principale différence apparaît dans la forme des courbes globales: le solveur déterministe a montré une plus grande dépendance à l'effet de propagation (Figure 8). Étant donné que les sous-régions de dates avaient leurs pics de contamination étaient très différentes, l'effet de propagation était plus lent.

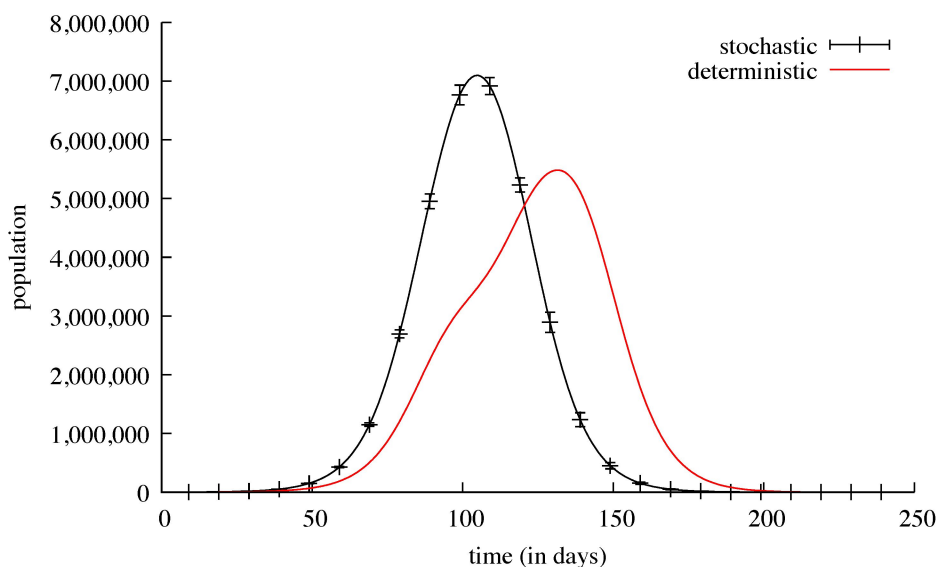


Figure 8. La courbe noire montre le nombre quotidien de personnes malades dans le pays. Il s'agit de la moyenne de 1000 exécutions d'une simulation stochastique tracée avec des barres tous les 10 jours indiquant l'écart type. La courbe rouge est une résolution continue déterministe du même modèle dans les mêmes conditions.

Bien que le nombre global de décès soit sensiblement le même (379 336 pour le SCD, 383 454 pour le SDS), la pente de la courbe obtenue avec le SDS est plus raide que celle obtenue avec le SCD (Figure 9). Cela peut être expliqué par la séquentialité relative des pics d'infection montrés par le solveur continu, alors qu'avec le solveur stochastique, tous les pics sont presque simultanés et donc le résultat est plus élevé.

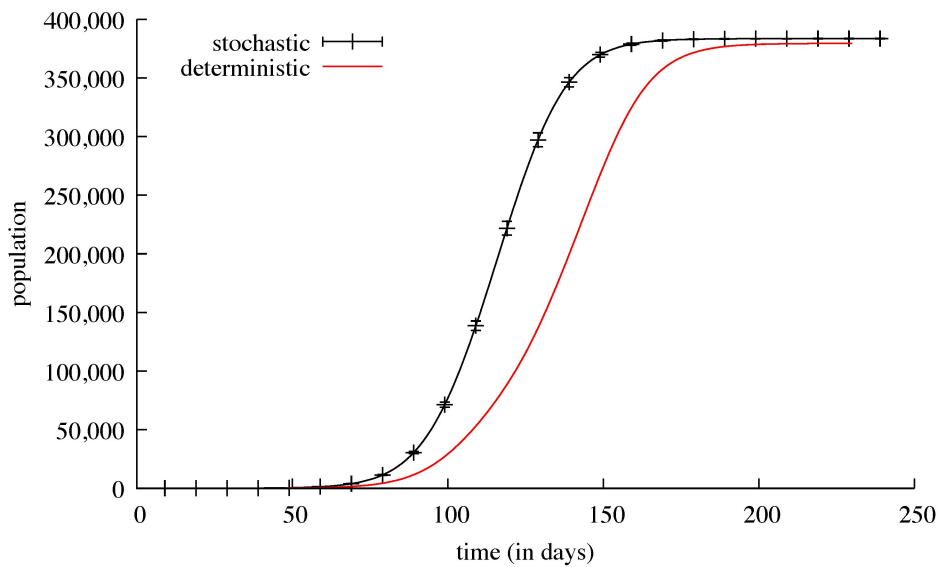


Figure 9. La courbe noire montre le nombre cumulé de décès dans tout le pays. C'est la moyenne de 1000 exécutions d'une simulation stochastique tracée avec des barres tous les 10 jours indiquant l'écart type. La courbe rouge est une résolution continue déterministe du même modèle dans les mêmes conditions.

Pour notre dernier test, nous avons paramétré le simulateur avec l'équivalent de la contre-mesure de confinement utilisée en France. L'effet de cette contre-mesure était de diminuer le nombre de personnes contagieuses, et alors que le SDS donnait des résultats qui correspondent aux statistiques réelles (Figure 5), le SCD n'a pas bien fonctionné principalement parce que le nombre initial de personnes contagieuses était trop faible pour être pris en compte (Figure 7). Plus de la moitié des départements n'ont fait état d'aucun décès et le nombre total de décès a donc été largement sous-estimé. Nous supposons que si nous partons d'un état initial où il y a suffisamment de personnes contagieuses dans la plupart des sous-régions, il est très probable que le SCD produira des résultats fiables.

5. Conclusions

Cette étude nous a permis de comparer deux méthodes différentes pour obtenir la trajectoire d'un système complexe. Au début, nous étions convaincus qu'ils donneraient des résultats très similaires, mais les faits nous ont prouvé que nous avons tort. Les raisons qui ont causé l'incohérence du comportement de l'algorithme discret stochastique d'une part et de l'algorithme continu déterministe d'autre part, nous amènent à être plus confiants dans l'approche stochastique pour la simulation de ce modèle de propagation épidémique particulier. Plus généralement, avec ce type de modèle, une phase de croissance exponentielle est très sensible à toute variation, même minime, des valeurs initiales, et aux artefacts, ou erreurs de calcul, et peut donc parfois présenter des comportements chaotiques.

Néanmoins, cette approche hybride, mélange d'un processus efficace centré sur la population jouant le rôle d'agent dans un système multi-agents, semble très prometteuse. Les résultats des simulations stochastiques étaient très similaires aux statistiques réelles recueillies à partir des données hospitalières. Les travaux futurs pourraient inclure des améliorations au simulateur telles que la mise en œuvre d'autres types de contre-mesures, l'utilisation de méthodes plus précises pour modéliser le comportement des individus et l'utilisation de différents types de sous-régions pour refléter leur diversité. Dans cette étude, nous n'avons supposé aucune réinfection possible, de sorte que l'épidémie s'est effectivement arrêtée après un certain temps. Bien que simplifiant le modèle, cette hypothèse interdit la possibilité de modéliser d'autres vagues d'infection. Des publications récentes ont discuté des conséquences de différents scénarios de transmission,⁴³]. Une perspective intéressante serait d'inclure dans notre modèle une probabilité de réinfection afin de tester l'efficacité des contre-mesures.

Le financement

Cette recherche n'a reçu aucun financement externe.

Remerciements

Un grand merci à Martin Davy de Sys2Diag, pour la première version de la boîte de dialogue des paramètres et la collecte d'informations sur le SARS-CoV-2.

Les conflits d'intérêts

Les auteurs ne déclarent aucun conflit d'intérêt.

Abréviations

Les abréviations suivantes sont utilisées dans ce manuscrit:

USI Unité de soins intensifs

ASS Algorithme de simulation stochastique

EDO Équations différentielles ordinaires

EDS Equations différentielles stochastiques

OMS Organisation Mondiale de la Santé

SCD Solveur continu déterministe

SDS Solveur discret stochastique

Annexe A

Annexe A.1.

Afin d'ajuster les résultats de la simulation aux statistiques réelles, nous avons estimé le nombre d'hôtes asymptomatiques dans chaque sous-région (départements) au début des simulations (tableau A1).

Tableau A1. Nombre initial de personnes contagieuses. <:figcaption>

Région	Départements							
Ile de France	Paris	65	Val-de-Marne	180	Val-d'Oise	70	Yvelines	20
	Seine-Saint-Denis	140						
Hauts de France	Somme	10	Nord	40	Oise	10	Aisne	10
Normandie	Seine-Maritime	10						
Bretagne	Morbihan	2	Ille-et-Vilaine	5				
Pays de Loire	Loire-Atlantique	8	Maine-et-Loire	8				
Centre	Eure-et-Loir	6	Loiret	2				
Aquitaine	Charente-Maritime	2	Gironde	6				
Occitanie	Haute-Garonne	2	Hérault	10	Gard	1		
Corse	Corse-du-Sud	2						
Provence Alpes Cote d'Azur	Bouches-du-Rhône	40						
Auvergne Rhone Alpes	Rhône	60	Ardèche	3	Drôme	3	Loire	4
Bourgogne Franche-Comté	Territoire de Belfort	3	Côte-d'Or	30	Doubs	2		
Grand Est	Meuse	1	Moselle	70	Bas-Rhin	20	Haut-Rhin	60

Valeurs par âge du pourcentage de létalité (extrapolées à partir de [42]), de la durée de la maladie et du pourcentage de voyageurs locaux et éloignés (tableau A2).

Tableau A2. Paramètres de la population.

Âge	Létalité%	Durée de la maladie (jours)	Voyage local%	Voyage à distance%
0 à 25	dix	7	5	1
26–50	15	8	6	1
51–75	20	dix	6	1
76+	55	14	0,5	0,4

Taux de contamination selon la localisation, pourcentage de patients hospitalisés pouvant infecter des personnes en voie de guérison et proportion de forme sévère de la maladie (tableau A3)

Tableau A3. Taux de contagion et paramètres globaux.

Emplacement	Taux %
Accueil	0,02
Lieu de travail	2
Transport public	4
Patients contagieux%	1
Forme sévère%	20
Constante de propagation	0,75

Tableau A4. Population de chaque tranche d'âge par région (source: INSEE, 1er janvier 2020).

Région	0 à 25	26–50	51–75	76+
Ile de France	3,164,218	4.177.466	2 982 661	683 650
Hauts de France	1 597 206	1 846 011	1 477 121	418 273
Normandie	917 615	808 834	926 890	318 070
Bretagne	733 777	868 726	1026 010	300 503
Pays de Loire	1 052 858	958 431	1 150 747	268 259
Centre	652 748	617 614	717 309	256 969
Aquitaine	1 483 728	1 435 736	1 869 051	661 315
Occitanie	1 471 676	1 527 461	1 873 453	594 186
Corse	69 362	96 471	91 685	40 289
Provence Alpes Côte d'Azur	1 270 520	1 185 877	1 562 820	473 619
Auvergne Rhone Alpes	2 152 246	2 272 047	2 202 878	693 612
Bourgogne Franche-Comté	572 106	719 857	811 137	291 031
Grand Est	1 403 834	1 556 258	1 567 738	446 914

Références

1. Bernard-Stoecklin, S.; Rolland, P.; Silue, Y.; Mailles, A.; Campese, C.; Simondon, A.; Mechain, M.; Meurice, L.; Nguyen, M.; Bassi, C.; et al. First cases of coronavirus disease 2019 (Covid-19) in France: Surveillance, investigations and control measures. *Eurosurveillance* 2020, 25, 2000094.

2. Décret no 2020-260 du 16 Mars 2020 Portant Réglementation des Déplacements dans le Cadre de la Lutte Contre la Propagation du Virus Covid-19. Legifrance. <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000041728476&categorieLien=id>
3. Kraemer, M.U.G.; Yang, C.-H.; Gutierrez, B.; Wu, C.-H.; Klein, B.; Pigott, D.M.; du Plessis, L.; Faria, N.R.; Li, R.; Hanage, W.P.; et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020, 368, 493–497.
4. Tian, H.; Liu, Y.; Li, Y.; Wu, C.-H.; Chen, B.; Kraemer, M.U.G.; Li, B.; Cai, J.; Xu, B.; Yang, Q.; et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 2020, 368, 638–642.
5. Diekmann, O.; Heesterbeek, J.A.; Metz, J.A. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 1990, 28, 365–382.
6. Zhao, S.; Lin, Q.; Ran, J.; Musa, S.S.; Yang, G.; Wang, W.; Lou, Y.; Gao, D.; Yang, L.; He, D.; et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int. J. Infect. Dis.* 2020, 92, 214–217.
7. Roberts, M.G.; Nishiura, H. Early estimation of the reproduction number in the presence of imported cases: Pandemic influenza H1N1-2009 in New Zealand. *PLoS ONE* 2011, 6.
8. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.; Lau, E.H.; Wong, J.Y.; et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* 2020.
9. Read, J.M.; Bridgen, J.R.E.; Cummings, D.A.T.; Ho, A.; Jewell, C.P. Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. *medRxiv* 2020.
10. Wu, J.T.; Leung, K.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *Lancet* 2020, 395, 689–697.
11. Tindale, L.; Coombe, M.; Stockdale, J.E.; Garlock, E.; Lau, W.Y.V.; Saraswat, M.; Lee, Y.-H.B.; Zhang, L.; Chen, D.; Wallinga, J.; et al. Transmission interval estimates suggest pre-symptomatic spread of COVID-19. *medRxiv* 2020.
12. Nature. Coronavirus Latest: Scientists Scramble to Study Virus Samples. Available online: <https://www.nature.com/articles/d41586-020-00154-w>
13. Fang, Y.; Nie, Y.; Penny, M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *J. Med. Virol.* 2020.
14. Tang, B.; Xia, F.; Tang, S.; Bragazzi, N.L.; Li, Q.; Sun, X.; Liang, J.; Xiao, Y.; Wu, J. The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China. *Int. J. Infect. Dis. Ijid Off. Publ. Int. Soc. Infect. Dis.* 2020, 95, 288–293.
15. Prem, K.; Liu, Y.; Russell, T.W.; Kucharski, A.J.; Eggo, R.M.; Davies, N.; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group; Jit, M.; Klepac, P. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *Lancet Public Health* 2020, 5, e261–e270.
16. Ji, C.; Jiang, D.; Shi, N. The Behavior of an SIR Epidemic Model with Stochastic Perturbation. *Stoch. Anal. Appl.* 2012, 30, 755–773.
17. Jiang, D.; Ji, C.; Shi, N.; Yu, J. The long time behavior of DI SIR epidemic model with stochastic perturbation. *J. Math. Anal. Appl.* 2010, 372, 162–180.
18. Cai, Y.; Kang, Y.; Banerjee, M.; Wang, W. A stochastic SIRS epidemic model with infectious force under intervention strategies. *J. Differ. Equ.* 2015, 259, 7463–7502.
19. Gray, A.; Greenhalgh, D.; Hu, L.; Mao, X.; Pan, J. A stochastic differential equation SIS epidemic model. *SIAM J. Appl. Math.* 2011, 71, 876–902.
20. Zakary, O.; Rachik, M.; Elmouki, I. A multi-regional epidemic model for controlling the spread of Ebola: Awareness, treatment, and travel-blocking optimal control approaches. *Math. Methods Appl. Sci.* 2017, 40, 1265–1279.
21. Abouelkheir, I.; El Kihal, F.; Rachik, M.; Zakary, O.; Elmouki, I. A multi-regions SIRS discrete epidemic model with a travel-blocking vicinity optimal control approach on cells. *Br. J. Math. Comput. Sci.* 2017, 20, 1–16.
22. El Kihal, F.; Abouelkheir, I.; Rachik, M.; Elmouki, I. Role of Media and Effects of Infodemics and Escapes in the Spatial Spread of Epidemics: A Stochastic Multi-Region Model with Optimal Control Approach. *Mathematics* 2019, 7, 304.
23. Park, H.; Kim, S.H. A Study on Herd Immunity of COVID-19 in South Korea: Using a Stochastic Economic-Epidemiological Model. *Environ. Resour. Econ.* 2020, 76, 665–670.

24. Hackl, J.; Dubernet, T. Epidemic Spreading in Urban Areas Using Agent-Based Transportation Models. *Future Internet* 2019, 11, 92.
25. Hunter, E.; Mac Namee, B.; Kelleher, J. An open-data-driven agent-based model to simulate infectious disease outbreaks. *PLoS ONE* 2019, 14, e0211245.
26. Amar, P.; Paulevé, L. HSIM: An hybrid stochastic simulation system for systems biology. In *Proceedings of the Third International Workshop on Static Analysis and Systems Biology*, Deauville, France, 10 September 2012; pp. 3–21.
27. Gillespie, D.T. A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *J. Comput. Phys.* 1976, 22, 403–434.
28. Rathinam, M.; Petzold, L.R.; Cao, Y.; Gillespie, D.T. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chem. Phys.* 2003, 119, 12784–12794.
29. GÉEDOS—Géo Données en Santé Publiques. Santé Publique France. <https://geodes.santepubliquefrance.fr/>
30. French Government Website. Info Coronavirus Covid 19. <https://www.gouvernement.fr/info-coronavirus/carte-et-donnees>
31. Guan, W.-J.; Ni, Z.-Y.; Hu, Y.; Liang, W.-H.; Ou, C.-Q.; He, J.-X.; Liu, L.; Shan, H.; Lei, C.-L.; Hui, D.S.; et al. Clinical characteristics of coronavirus disease 2019 in china. *N. Engl. J. Med.* 2020, 382, 1708–1720.
32. Grasselli, G.; Pesenti, A.; Cecconi, M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. *JAMA* 2020.
33. Fontanet, A.; Tondeur, L.; Madec, Y.; Grant, R.; Besombes, C.; Jolly, N.; Pellerin, S.F.; Ungeheuer, M.-N.; Cailleau, I.; Kuhmel, L.; et al. Cluster of COVID-19 in northern France: A retrospective closed cohort study. *medRxiv* 2020.
34. Sali, H.; Kiem, C.T.; Lefrancq, N.; Courtejoie, N.; Bosetti, P.; Paireau, J.; Andronico, A.; Hoze, N.; Richet, J.; Dubost, C.-L.; et al. Estimating the burden of SARS-CoV-2 in France. *Sciences* 2020.
35. Béraud, G.; Kazmierczak, S.; Beutels, P.; Levy-Bruhl, D.; Lenne, X.; Mielcarek, N.; Yazdanpanah, Y.; Boëlle, P.-Y.; Hens, N.; Dervaux, B. The French connection: The first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS ONE* 2015, 10, e0133203.
36. Russell, T.W.; Hellewell, J.; Jarvis, C.I.; van Zandvoort, K.; Abbott, S.; Ratnayake, R.; Flasche, S.; Eggo, R.M.; Edmunds, W.J.; Kucharski, A.J. Cmmid Covid-Working Group, Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. *Euro Surveill.* 2020, 25, 2000256.
37. Mizumoto, K.; Kagaya, K.; Zarebski, A.; Chowell, G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill.* 2020, 25.
38. Verity, R.; Okell, L.C.; Dorigatti, I.; Winskill, P.; Whittaker, C.; Imai, N.; Cuomo-Dannenburg, G.; Thompson, H.; Walker, P.G.; Fu, H.; et al. Estimates of the severity of coronavirus disease 2019: A model-based analysis. *Lancet Infect. Dis.* 2020.
39. Bao, L.; Deng, W.; Gao, H.; Xiao, C.; Liu, J.; Xue, J.; Lv, Q.; Liu, J.; Yu, P.; Xu, Y.; et al. Reinfection could not occur in SARS-CoV-2 infected rhesus macaques. *bioRxiv* 2020.
40. Lauer, S.A.; Grantz, K.H.; Bi, Q.; Jones, F.K.; Zheng, Q.; Meredith, H.R.; Azman, A.S.; Reich, N.G.; Lessler, J. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* 2020, 172, 577–582.
41. Du, Z.; Xu, X.; Wu, Y.; Wang, L.; Cowling, B.J.; Meyers, L.A. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg. Infect. Dis.* 2020.
42. Worldometer. 2020. Available online: <https://www.worldometers.info/coronavirus> (accessed on 24 August 2020).
43. Kissler, S.M.; Tedijanto, C.; Goldstein, E.; Grad, Y.H.; Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 2020, 368, 860–868.