

$$W(\alpha, x, \xi) \Rightarrow \max$$

Séminaire Mathématique de Béjaïa (LaMOS)

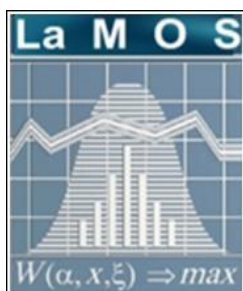
Revue Spécialisée de Mathématiques et Applications

*Les mathématiques,
la santé (Covid 19, ...)
et autres domaines
d'application*

Scientificamerica.com



Karima Adel – Aissanou et Djamil Aïssani



Publication et Edition

LaMOS Research Unit
University of Béjaïa

<http://www.lamos.org>

Volume 18
Béjaïa 2020

ISSN : 1112-9433

E – ISSN : 2800 - 1214

partie I- Les mathématiques et la santé (Covid 19,...)

partie II- Autres applications

partie III- Processus de publication



LaMOS

MODÉLISATION ET OPTIMISATION DES SYSTÈMES

ACCUEIL

VOLUMES DE LA REVUE

LAMOS



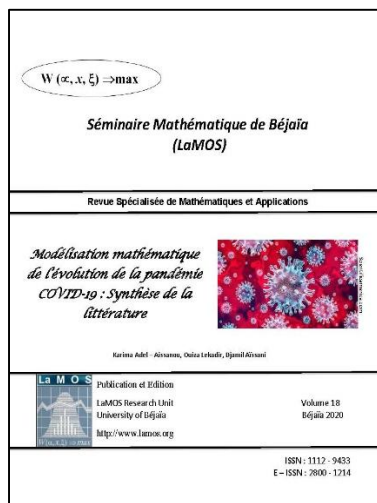
Revue spécialisée de Mathématiques et Applications

ISSN : 1112 – 9433. E – ISSN : 2800 - 1214

La revue «Séminaire Mathématique de Béjaïa (LaMOS)» est éditée par l'Unité de Recherche LaMOS – Université de Béjaïa.

University of Bejaia, Targua Ouzamour, 06000, Algeria

E-Mail: lamos_bejaia@hotmail.com



Editor (Directeur de la publication) :

Pr Djamil Aïssani, Director of the Research Unit **LaMOS**

Editorial Board :

Pr Mohamed Saïd Radjef, Chef de la Division, ADRO

Pr Smail Adjabi, Chef de la Division, MS2P

Pr Bibi Mohand Ouamer, Chef de l'équipe OCO

Pr Louiza Bouallouche, Chef de l'équipe EPSIRT

Pr Redouane Laggoune, Chef de l'équipe MSF

Dr Rabah Medjoudj, Chef de l'équipe FSE2

Pr Mohamed Boualem, Chef de l'équipe SR2

Pr Karim Abbas, Chef de l'équipe MAPPS

Dr Fazia Rahmoune, Chef du Département RO

Dr Kamal Kabyl, Responsable du Séminaire SMB

Fondation : année 2002. Dépôt légal : 934 - 2003

Présentation :

La revue «**Séminaire Mathématique de Béjaïa (LaMOS)**» est une revue spécialisée de mathématiques et applications. Fondée en 2002, elle paraît une fois par an.

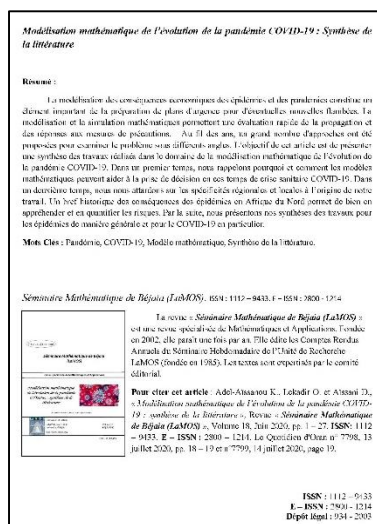
Elle édite les Comptes Rendus annuels du Séminaire Hebdomadaire de l'Unité de Recherche **LaMOS**. Elle peut également éditer des Actes de congrès, ou bien des monographies.

Depuis sa parution, près de 200 textes ont été édités. Ainsi, cette année 2020, sont parus un numéro spécial sur les «**Les mathématiques, la santé (Covid 19, ...) et autres domaines d'application**» (165 pages) (Volume 18), ainsi que la monographie «**Approximations dans les Modèles Stochastiques**» (245 pages). Les textes sont revus par les membres du comité éditorial.

ISSN : 1112 – 9433. E – ISSN : 2800 - 1214

Copyright : Séminaire Mathématique de Béjaïa (LaMOS)

Date : 2018



Modélisation mathématique de l'évolution de la pandémie COVID-19 : Synthèse de la littérature

Résumé :

La modélisation des conséquences épidémiologiques et des pandémies constitue un domaine innovant de la préparation de plans d'urgence pour diversifier les nouvelles flambées. La modélisation et la simulation mathématiques permettent une évaluation rapide de la propagation et des systèmes aux moyens de prévention. Au fil des ans, un grand nombre d'approches ont été proposées pour résoudre le problème sous différents angles. L'objectif de cet article est de présenter une synthèse des travaux réalisés dans le domaine de la modélisation mathématique de l'évolution de la pandémie COVID-19. Dans un premier temps, nous rappelons pourquoi et comment les modèles mathématiques peuvent aider à la prise de décision en cas de crise sanitaire COVID-19. Dans un deuxième temps, nous nous concentrons sur les aspects clés épidémiologiques et mathématiques de notre travail. Un bref historique des conséquences des épidémies en Afrique du Nord permet de bien en appréhender et en quantifier les risques. Par la suite, nous présentons nos systèmes des tests pour les épidémies de manière générale et pour la COVID-19 en particulier.

Mots Clés : Pandémie, COVID-19, Modèle mathématique, Synthèse de la littérature.

Séminaire Mathématique de Béjaïa (LaMOS), ISSN: 1112-9433, E-ISSN: 2800-1214

La revue «**Séminaire Mathématique de Béjaïa (LaMOS)**» est une revue spécialisée de Mathématiques et Applications. Fondée en 2002, elle paraît une fois par an. Elle édite les Comptes Rendus annuels du Séminaire Hebdomadaire de l'Unité de Recherche LaMOS (fondée en 1985). Les textes sont revus par le comité éditorial.

Pour citer cet article : Adel-Aissaoui K., Tekadit O. et Aïssani D., «**Modélisation mathématique de l'évolution de la pandémie COVID-19 : synthèse de la littérature**», Revue «**Séminaire Mathématique de Béjaïa (LaMOS)**», Volume 18, Juin 2020, pp. 1 – 27, ISSN: 1112 – 9433, E – ISSN : 2800 – 1214, Le Quotidien d'Oran n° 7798, 13 juillet 2020, pp. 18 – 19 et n° 7799, 14 juillet 2020, page 19.

ISSN : 1112 - 9433
E – ISSN : 2800 - 1214
Dépôt légal : 934 - 2003

Revue : Séminaire Mathématique de Béjaia (LaMOS).

Editeurs : Unité de recherche LaMOS (Modélisation et Optimisation des Systèmes).

Adresse : Université de Béjaia, Campus Targa Ouzamour, 06000 (Algérie).

Tel/Fax : (213) 34 81 37 09.

E-mail : lamos_bejaia@hotmail.com

<http://www.lamos.org>

©Publications de l'Unité de recherche LaMOS, 2020.

Tous droits de traduction, de reproduction
et d'adaptation réservés pour tous les pays.

ISSN : 1112 - 9433

E - ISSN : 2800 - 1214

Dépôt légal : 934 - 2003

. Préface

Voici un nouveau numéro (Volume 18) de la revue spécialisée "Revue Séminaire Mathématique de Béjaia (LaMOS)". Il s'inscrit dans l'évolution et l'enrichissement continu des orientations du journal. Il n'existe pas de limites pour l'application des mathématiques, il n'est plus non plus à ses débuts.

Ce numéro comprend trois parties : La première partie est intitulée : "Les mathématiques et les systèmes de santé". Elle concerne cinq applications des mathématiques, parfois combinée à l'informatique, dans un domaine particulier qui est la santé. C'est un choix original, qui est motivé dans cette période post COVID, au besoin de répondre rationnellement à des questions réelles. En effet, les mathématiciens se sont fortement impliqués dans la lutte contre la pandémie.

De fait, le premier article (de cette partie) concerne une synthèse complète de la littérature sur la "Modélisation mathématique de l'évolution de la pandémie COVID-19 " [2]. Il est suivi par un autre article qui traite " l'impact de la crise sanitaire sur le trafic portuaire au niveau de l'EPB (Entreprise Portuaire de Béjaia)". Précisons ici que plusieurs travaux réalisés au niveau de l'Unité de Recherche LaMOS concerne la Covid-19. C'est le cas de la thèse de Zaitri M. A. [1]. Dans laquelle, on peut suivre en détail la construction d'un modèle mathématique de contrôle optimal dédié à l'évolution d'une pandémie et les solutions proposées ainsi qu'une application à la COVID-19. Il en est de même de l'article de Demouche M. et al. " Diagnosis of the Covid-19 with Deep Learning " [3].

Les trois autres articles de cette première partie abordent la "Simulation du mouvement des malades au niveau des urgences à l'hôpital Khelil Amrane Béjaia, Algérie", "l'Application des méthodes d'apprentissage dans la prédiction du diabète de Type 2" et la " Conception d'un système multi-agents d'Aide à la décision pour une Gestion Efficace des Services d'Urgences : cas de l'hôpital de Kherrata".

La deuxième partie illustre cinq travaux sur l'application des mathématiques à l'évaluation des performances et à la théorie des jeux. Les deux premiers articles concernent les : "Jeux Bayésiens" et "l' Equilibre de Berge : entre situationnisme et optimalité de Pareto". Quant aux trois articles suivants, ils abordent les questions sur l'évaluation des performances des réseaux de communication : "Routing and control of congestion in télécommunication networks", " Mo-

dèle analytique pour l'évaluation des performances d'un réseaux pair à pair (cas du Bittorent)" et ensuite "Amélioration de la norme IEEE 802.11^e mode EDCA sous un canal non idéal".

La troisième partie développe un article sur le processus de soumission : "Overview and Writing".

Ce numéro plaira spécialement aux personnes ayant un goût pour les mathématiques appliquées, mais aussi à ceux qui cherchent des réponses liées à la santé et l'informatique. On y trouvera des travaux originaux et de qualité. La présentation est exceptionnellement claire et les motivations sont données avec un grand soin, signe de l'expérience de chacun des auteurs.

Références

1. Zaitri M. A., *Minimisation d'un problème linéaire quadratique de contrôle optimale avec des contraintes de phase initiales et terminales suivi d'une application à un modèle épidémiologique de type Covid-19*, Thèse de doctorat en sciences, filière mathématiques appliquées, option Modélisation mathématiques et techniques de Décision, université de Béjaïa, 2023.
2. Adel-Aissanou K., Lekadir O. et Aissani D., *Modèles mathématiques et évolution de la pandémie Covid-19 : synthèse de la littérature*, Quotidien d'Oran, Lundi 13. juillet 2020, pp. 18 et 19,
3. Demouche M., Baadache A., Aissani D. and Zerrouki T., *"Diagnosis of the Covid-19 with Deep Learning"*, International Journal of Medical Engineering and Informatics, 2023 (Accepted).

Table des matières

Préface	3
<hr/>	
partie I Les mathématiques et les systèmes de santé	
<hr/>	
1 Modélisation mathématique de l'évolution de la pandémie covid-19 : synthèse de la littérature	
<i>Adel-Aissanou K., Lekadir O. et Aïssani D.</i>	9
2 L'Impact de la Crise Sanitaire sur le Trafic Portuaire de Béjaia(EPB)	
<i>Ramdani A., Ghiles L., Saadi N. et Aïssani D.</i>	29
3 Modélisation et simulation du mouvement des malades au niveau de l'hôpital Khelil Amrane Béjaia, Algérie.	
<i>Oudina S., Baalla M., Adel-Aissanou K. et Aïssani D.</i>	53
4 Application des Méthodes d'Apprentissage dans la Prédiction du Diabète de Type 2.	
<i>Mehidi D., Medjoudj S., Adel-Aissanou K. et Aïssani D.</i>	65
5 Conception d'un Système Multi-Agents d'Aide à la Décision pour une Gestion Efficace des Services d'Urgence : Cas de l'Hôpital de Kherrata	
<i>Boukredera D., Adel-Aissanou K., Abdoune K. et Bouriahi L.</i>	85
<hr/>	
partie II Autres domaines d'application	
<hr/>	
I - Théorie des jeux	105
6 Jeux Bayésiens et applications	
<i>Bouibed K.</i>	107
7 Equilibre de Berge : Entre Situationnisme et Optimalité de Pareto	
<i>Kadi S., Zaidi H., Tazdaït T. et Aïssani D.</i>	111
II- Evaluation des performances des réseaux de communication	127

8 Routing and Control of Congestion in Telecommunication Networks

Touati N., Hadji M., Laib F. et Aïssani D. 129

9 Modèles analytiques pour l'évaluation des performances d'un réseau pair-à-pair (cas du Bittorent)

Arkoub D., Krouri Y., Aïssani D. et Amad M. 137

10 Amélioration de la norme IEEE 802.11^e mode EDCA sous un canal non idéal

Bachiri L., Abdi N., Aroul K. et Aïssani D. 151

partie III Processus de publication

11 Publishing process : Overview and writing

Barèche A. 165

Index 171

Les mathématiques et les systèmes de santé

Modélisation mathématique de l'évolution de la pandémie covid-19 : synthèse de la littérature

Adel-Aissanou K., Lekadir O. et Aïssani D.

Research Unit **LaMOS** (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria
lamos_bejaia@hotmail.com ,

Résumé La modélisation des conséquences économiques des épidémies et des pandémies constitue un élément important de la préparation de plans d'urgence pour d'éventuelles nouvelles flambées. La modélisation et la simulation mathématiques permettent une évaluation rapide de la propagation et des réponses aux mesures de précautions. Au fil des ans, un grand nombre d'approches ont été proposées pour examiner le problème sous différents angles. L'objectif de cet article est de présenter une synthèse des travaux réalisés dans le domaine de la modélisation mathématique de l'évolution de la pandémie COVID-19. Dans un premier temps, nous rappelons pourquoi et comment les modèles mathématiques peuvent aider à la prise de décision en ces temps de crise sanitaire COVID-19. Dans un deuxième temps, nous nous attardons sur les spécificités régionales et locales à l'origine de notre travail. Un bref historique des conséquences des épidémies en Afrique du Nord permet de bien en appréhender et en quantifier les risques. Par la suite, nous présentons nos synthèses des travaux pour les épidémies de manière générale et pour le COVID-19 en particulier.

Mots clés : Pandémie, COVID-19, Modèle mathématique, Synthèse de la littérature.

1.1 Introduction : Pourquoi la modélisation mathématique d'une épidémie ?

Depuis le début de l'humanité, l'homme a utilisé les mathématiques pour essayer de résoudre certains problèmes auxquels il était confronté. En même temps, la contribution à la résolution des défis de la société et d'autres disciplines a été source d'inspiration pour de nombreux développements mathématiques à travers l'histoire. La crise sanitaire actuelle générée par COVID-19 ne fait pas exception et prouve une fois de plus la contribution de cette discipline au bien-être social. [5].

Chaque année, des millions de personnes dans le monde meurent de maladies infectieuses telles que la rougeole, le paludisme, la tuberculose, le VIH. Bien qu'il existe de nombreux facteurs de complication, des modèles mathématiques simples peuvent donner un aperçu de la dynamique des épidémies de maladies et aider les responsables à prendre des décisions concernant les politiques de santé publique. [25].

Une épidémie est une explosion soudaine d'une maladie au sein d'une population donnée et une pandémie est une épidémie qui se propage à travers les régions, qui peuvent être des pays, des continents ou le monde entier. Ces épidémies ou pandémies causent de graves problèmes de développement économique et social qui entraînent un déséquilibre dans la survie des humains

et de plusieurs autres espèces vivantes. [18].

Depuis décembre 2019, de nombreux cas inexplicables de pneumonie avec toux, dyspnée, fatigue et fièvre comme principaux symptômes sont survenus à Wuhan, en Chine en peu de temps [16, 31]. Les autorités sanitaires chinoises et les CDC ont rapidement identifié l'agent pathogène de ces cas comme un nouveau type de coronavirus, que l'OMS - Organisation Mondiale de la Santé - a nommé COVID-19 le 10 janvier 2020 (World Health Organization) [33]. A partir de là, de nombreuses équipes de recherche ont abordé différentes études tournant autour des questions suivantes : Comment les modèles mathématiques peuvent-ils aider à la prise de décision en ces temps de crise sanitaire ? Et comment parviennent-ils à intégrer les informations et incertitudes en cours ? [10]. Rappelons qu'un modèle est une abstraction (une simplification) de la réalité. Le rôle de la modélisation mathématique d'une épidémie n'est donc pas de reproduire la réalité mais d'essayer de mieux la comprendre, d'évaluer des hypothèses pour indiquer une tendance.

Crepey P. & Khalatbari A. constatent que les épidémies, tout comme de nombreuses lois naturelles, commencent toutes selon une dynamique qui obéit à une fonction mathématique qui "tend vers l'infini". C'est la fonction exponentielle. Concrètement, le nombre d'infectés augmente de plus en plus vite au cours du temps. Ce modèle grossier du départ de l'épidémie ne tient compte ni des données de la transmission, ni des mesures sanitaires. Par la suite, on doit essayer de voir comment se transmet la maladie. C'est ce qu'on appelle le déroulé (ou bien l'historique) de la maladie : période d'incubation, contagiosité, etc... Ces éléments permettent de concevoir un modèle de base. Celui qui a été élaboré par Crepey P & Khalatbari A est le modèle SIR, où toute la population peut être partagée en trois états : sains (S), infectés (I), rétablis (R). La dynamique de la maladie dépend des échanges entre ces trois boîtes, aussi appelés "compartiments" [10].

Précisons ici que ce modèle SIR à trois compartiments ne représente pas le COVID-19 car les spécificités de ce virus nécessitent des modèles plus complexes. Ainsi, il faut avoir des éléments sur la durée de l'incubation. Il en est de même pour la présence de personnes asymptomatiques après leur infection. L'équipe de Crespey P. a d'abord utilisé un modèle à 4 boîtes, appelé SEIR en y ajoutant une boîte supplémentaire, E qui désigne les personnes ayant été exposées mais qui ne présentent pas encore de symptômes. Cette équipe poursuit ses investigations en profondeur sur un modèle à 8 compartiments qui sera capable de tenir compte des personnes qui restent asymptomatiques après avoir été infectées, de celles qui commencent à être contagieuses quelques jours avant l'apparition des symptômes, des personnes hospitalisées et enfin des personnes contaminées qui demeurent confinées au sein de leur famille [10].

L'objectif de cet article est de présenter une synthèse des travaux réalisés dans le domaine de la modélisation mathématique de l'évolution de la pandémie COVID-19. Dans un premier

temps, nous rappelons pourquoi et comment les modèles mathématiques peuvent aider à la prise de décision en ces temps de crise sanitaire COVID-19. Dans un deuxième temps, nous nous attardons sur les spécificités régionales et locales à l'origine de notre travail. Un bref historique des conséquences des épidémies en Afrique du Nord permet de bien en appréhender et en quantifier les risques. Par la suite, nous présentons nos synthèses des travaux de modélisation mathématiques réalisés pour les épidémies de manière générale et pour le COVID-19 en particulier.

1.2 A l'origine de cette synthèse de la littérature sur la modélisation mathématique des épidémies

Depuis une vingtaine d'années, l'Unité de Recherche LaMOS (Modélisation et Optimisation des Systèmes - <http://www.lamos.org>) travaille occasionnellement avec les structures de santé de Béjaïa (au début, l'Hôpital Khellil Amrane, puis, après leurs création, avec la Faculté de médecine et le CHU) pour aborder des problèmes d'évaluation de performance des systèmes de santé (voir par exemple [2], [11]). Ces travaux se font dans le cadre d'un groupe pluridisciplinaire. Au début de la crise du COVID-19 en Algérie, nous avons été en contact avec la cellule de crise de la Wilaya de Béjaïa. Il s'agissait d'aborder la question de la modélisation mathématique de l'évolution de l'épidémie à l'échelle locale. A titre d'exemple, il s'agissait de pouvoir répondre à des questions de type : au bout de combien de jours les capacités d'accueil des services spécialisés (et notamment de celui de réanimation) seront-elles atteintes ? A ce niveau, une récolte des données est essentielle. Un questionnaire avait été élaboré pour chaque patient (début des symptômes, date du test, date de l'hospitalisation, date du résultat du test, date du début du traitement, date de la dégradation de l'état de santé, ...). La fiabilité des résultats dépend de la rigueur des personnels de santé chargés du suivi (et donc de la récolte des données). A cette époque, le laboratoire COVID-19 de l'Université de Béjaïa n'avait pas encore été envisagé et la wilaya dépendaient des performances de l'Institut Pasteur d'Alger.

Certes, la courbe de l'augmentation exponentielle de la morbidité et de la mortalité algérienne est nettement réduite comparée à celles des pays occidentaux. D'un autre côté, les données observées de la COVID-19 à ce jour sont très en dessous des chiffres prévus par les modélisations mathématiques. Le Pr Hamdi Cherif M [14] estime que *"le challenge pour sortir de ce tunnel est la mise en œuvre d'un plan de déconfinement qui doit d'abord tenir compte des points faibles des confinements antérieurs et les spécificités locales et régionales, avec l'espoir que notre écosystème diminuera la virulence et la contagion du virus "* [14]. Et il ajoute *" qu'il est temps d'aller sur le terrain pour des études critiques des différentes situations épidémiologiques en vue de mettre en place des stratégies locales, spécifiques et adaptées pour chaque wilaya ou région "* [14]. En effet, l'éclosion des foyers épidémiques (clusters) de la COVID-19 constitue aujourd'hui le défi pour les autorités sanitaires alors que la première phase de déconfinement est enclenchée. Le bulletin épidémiologique de l'INPS (Institut National de Santé

Publique) du 04 juin 2020 portant sur les cas cumulés confirmés positifs par PCR note qu'à cette date, le taux d'incidence nationale était de 23.08 cas pour 100 000 habitants. Mais on doit tenir compte de certaines spécificités locales et régionales. Ainsi, pour la Wilaya de Béjaïa, c'est la région de Kherrata qui a été la plus touchée. Or elle est frontalière avec la Wilaya de Sétif dont l'incidence est de loin supérieure au taux national (34.64).

1.3 Historique des épidémies en Afrique du Nord

A l'instar de ce qui se passe partout sur la planète, l'Afrique du Nord et l'espace méditerranéen n'ont pas été épargnés par les catastrophes naturelles : séismes, inondations, épidémies. Certaines ont été ponctuelles, d'autres ont dévasté des régions entières, ont fait le tour de la Méditerranée, en détruisant une part importante de la population, vidant des pays peuplés, désorganisant des empires, modifiant les rapports de force. Suivant leur origine et leurs circonstances, leurs effets ont été parfois immédiats, parfois différés, courts ou longs, directs ou indirects [20]. Déjà dans la seconde moitié du III^e siècle, Arnobe, à propos des épidémies qui consomment le genre humain, affirme : "*consultez les livres d'annales rédigés en différentes langues ; vous apprendrez que des pays entiers ont été souvent dépeuplés, vidés de leurs habitants*" [3].

Les textes anciens mentionnent sous le nom de pestes toutes sortes de maladies à forme épidémique, au pronostic grave, sinon fatal. D'un point de vue médical, toutes n'étaient pas des pestes à proprement parler. Les descriptions sont souvent trop vagues pour qu'on puisse en saisir la nature exacte, et guère plus leur ampleur et leur impact. Laporte J P souligne que "suivant les cas, les épidémies avaient une influence directe et indirecte considérable dans tous les domaines économique, démographique, social, politique, militaire, religieux, artistique, sanitaire et médical". À court terme, elles interrompaient la circulation des hommes, et donc celle des denrées, entraînaient, elles aussi, la famine [20].

A cette époque, les éléments de diagnostic étaient limités, mais le tableau clinique était très important, avec une série de symptômes identifiés : vomissements, charbon, hémorragie, symptômes nerveux, crachements de sang de la forme pneumonique, La mortalité pouvait atteindre jusqu'à plus de 50% de la population suivant les épisodes locaux. La guerre et les échanges favorisaient l'extension de la peste, en affaiblissant la résistance des populations.

Biraben J. N., considéré comme étant le grand spécialiste de l'histoire de la peste, analyse le cas de la peste justinienne qui apparut en 541 dans le port égyptien de Péluse, venant semble-t-il d'Ethiopie [7]. De là, elle s'étendit sur l'Egypte, jusqu'à Alexandrie, la Palestine, la Syrie. Quelques mois plus tard, elle est apparue à Constantinople, où elle dura quatre mois. Elle fut ensuite diffusée rapidement dans certains ports méditerranéens par la voie maritime. L'intérieur des pays était moins touché, car la population était moins dense. Biraben J. N.

affirme que ce sont les soldats de Justinien envoyés de Constantinople à Carthage pour faire campagne contre les Maures qui y auraient apporté la peste. Il parle d'une diminution de 50% de la population de Constantinople (40 % de morts, 10 % de fuites) [7].

Si on se projette maintenant au présent, on constate que l'Algérie a connu ces dernières décennies une incroyable série de catastrophes : en plus des calamités dites "naturelles" (inondation de Bab-El-Oued en novembre 2001, plus de 800 morts ; séisme de Boumerdès du 21 mai 2003 : 2 300 morts, autant de disparus et 11 000 blessés) une succession d'épidémie de botulisme, tuberculose (2001), typhoïde (2002), rougeole, méningite (2002), peste (juin 2003), dans la banlieue d'Oran... Abid L rappelle que cette ville a justement servi de cadre à Albert Camus pour son roman mythique, "*la Peste*" où dans la conclusion, prémonitoire (de ce roman) Camus fait dire à son héros : "*Et (Rieux) savait ce qu'on peut lire dans les livres que le bacille de la peste ne meure ni ne disparaît jamais, qu'il peut rester, pendant des dizaines d'années, endormi dans les meubles et le linge, qu'il attend patiemment dans les chambres, les caves... Et que, peut-être, le jour viendrait où, pour le malheur et l'enseignement des hommes, la peste réveillerait ses rats et les enverrait mourir dans une cité heureuse*" [1].

1.4 Historique de l'évolution du covid-19

Actuellement, des personnes du monde entier ont été touchées par la maladie du coronavirus 2019 (COVID-19), qui est la cinquième pandémie après celle de la grippe de 1918. À partir de maintenant, nous pouvons retracer le premier rapport et la flambée subséquente d'un groupe de nouveaux cas de pneumonie humaine dans la ville de Wuhan, en Chine, en décembre 2019. La première date d'apparition des symptômes était le 1er décembre 2019 [22], [16], [36]. Le 11 mars 2020, l'augmentation rapide du nombre de cas en dehors de la Chine a conduit le docteur Tedros Adhanom Ghebreyesus, directeur général de l'OMS, à annoncer que la flambée pouvait être qualifiée de pandémie.

À la mi-mars 2020, la région européenne de l'OMS était devenue l'épicentre de l'épidémie, notifiant plus de 40 % des cas confirmés dans le monde. Au 28 avril 2020, la région concourait à 63 % de la mortalité mondiale due au virus.

Le premier cas de COVID-19 en Afrique est apparu en février 2020 en Egypte. 53 pays africains sur 54 sont désormais touchés par le coronavirus. Depuis, plusieurs études trop alarmistes ont exploré la piste d'une "catastrophe africaine", qui, près de trois mois après le début de la crise, n'a pas eu lieu. Alors que l'Europe et les Etats-Unis ont dépassé le million de cas, le nombre officiel de contaminations en Afrique s'élève, à la mi-mai, à 75 530 cas enregistrés, selon le Centre pour la prévention et le contrôle des maladies de l'Union africaine (CDC Afrique). le premier cas confirmé de COVID-19 aux USA est annoncé en janvier 2020. Les États-Unis et le Brésil sont les pays où la pandémie de COVID-19 progresse le plus rapidement.

1.5 Les modèles mathématiques modélisant les pandémies

Siettos C I & Russo L [32] ont présenté et discuté les principales approches utilisées pour la surveillance et la modélisation de la dynamique des maladies infectieuses. Ils ont présenté les concepts de base qui expliquent leur mise en œuvre et leur pratique et, pour chaque catégorie, ils ont donné une liste annotée d'œuvres représentatives. La FIGURE 1.1 donne un aperçu des modèles mathématiques pour les maladies infectieuses

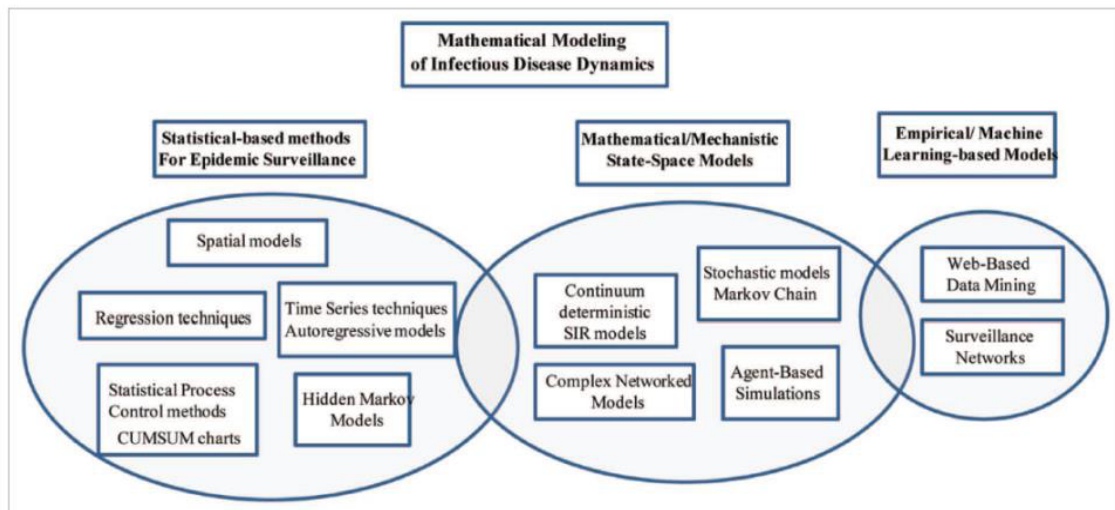


FIGURE 1.1: Un aperçu des modèles mathématiques pour les maladies infectieuses. Source : Siettos C.I.& Russo L. [32].

Au fil des ans, un grand nombre d'approches ont été proposées pour examiner le problème sous différents angles. Ceux-ci englobent trois catégories générales (voir FIGURE 1.1) :

- (1) les méthodes statistiques de surveillance des épidémies et d'identification des schémas spatiaux dans les épidémies réelles,
- (2) les modèles mathématiques dans le cadre de systèmes dynamiques (également appelés modèles de l'espace d'état) utilisés pour prévoir l'évolution d'une propagation épidémique "hypothétique" ou en cours, et
- (3) apprentissage automatique / méthodes expertes pour prévoir l'évolution d'une épidémie en cours. Pour ces trois catégories, il existe à nouveau des approches différentes tissant une littérature vaste et diversifiée.

1.6 Synthèse des travaux sur la modélisation mathématique des épidémies

En 1760, Bernoulli D a développé un modèle mathématique d'une épidémie de variole appelée à l'époque "la petite variole" pour analyser la mortalité causée par cette maladie [6]. Son modèle est connu comme étant le premier modèle mathématique de l'histoire de l'épidémiologie [4], [34].

Dans l'article [35], les auteurs donnent une bonne revue de littérature concernant la modélisation mathématique de la grippe saisonnière. Ils ont modélisé la propagation d'une épidémie de grippe dans la population d'Oran, en Algérie. En particulier, ils ont construit un modèle SEIR-SW (Susceptible-Exposed-Infected-Removed dans un réseau Small World) pour identifier les principales caractéristiques de la transmission épidémique et son évolution dans le temps.

Dans [17], les auteurs ont précisé dans quelle mesure les modèles mathématiques de transmission de maladies fournissent des prédictions fiables. Ils ont conclu qu'afin d'utiliser efficacement les modèles mathématiques pour étudier la dynamique de la propagation d'une pandémie, y compris les stratégies de contrôle possibles, il faut être sûr que les valeurs utilisées pour les différents paramètres du modèle correspondent à la réalité. Bien que certains paramètres puissent être déterminés sur la base de connaissances antérieures, d'autres paramètres doivent être estimés en ajustant le modèle aux données disponibles.

Un avantage important de l'utilisation de modèles est que la représentation mathématique des processus biologiques permet la transparence et la précision des hypothèses épidémiologiques, nous permettant ainsi de tester notre compréhension de l'épidémiologie de la maladie en comparant les résultats des modèles et les modèles observés [24].

D'autres chercheurs ont utilisé la modélisation mathématique dans la prédiction de la propagation des épidémies. Dans [19], les auteurs ont développé un modèle mathématique pour la dynamique de transmission de l'épidémie de grippe A / H1N1 de 2009 en République de Corée. Dans [8], les auteurs donnent en détail une modélisation mathématique des épidémies et comment interpréter les résultats. Les auteurs du papier [12] se sont préoccupés des conséquences économiques qui auraient pu se produire si la maladie d'Ebola s'était propagée à certains pays en développement de la région Asie-Pacifique. Leurs estimations ont été effectuées avec deux modèles liés : un modèle de transmission des maladies et un modèle multirégional d'équilibre général calculable dynamique (ECG) de l'économie mondiale. Dans ([26], l'auteur a donné une modélisation mathématique de la propagation de l'épidémie de SIDA. Les auteurs dans [23] ont présenté un modèle mathématique de transmission de la grippe porcine avec la vaccination de nouveaux nés. Dans [9], les auteurs ont passé en revue les progrès récents dans la modélisation et la caractérisation des schémas de croissance épidémique précoce à partir des données sur les flambées de maladies infectieuses, et ont étudié les types de formulations mathématiques qui sont les plus utiles pour saisir une diversité de profils de croissance épidémique précoce, allant de la dynamique de croissance sub-exponentielle à la dynamique de croissance exponentielle. Plus précisément, ils ont examiné des modèles mathématiques qui intègrent des détails spatiaux ou des structures de mélange de population réalistes, y compris des modèles de métapopulation, des modèles de réseaux individuels et des modèles simples de type SIR qui intègrent les effets des changements de comportement réactifs ou de mélange inhomogène. Dans [15], les auteurs

ont effectué une revue systématique des modèles mathématiques estimant l'impact épidémiologique des futurs vaccins antituberculeux humains. Avec une méthode rigoureuse, ils ont affirmé que les évaluations économiques ont révélé que les vaccins antituberculeux étaient extrêmement rentables, en particulier lorsqu'ils étaient destinés aux adolescents/adultes. Les auteurs de l'article [32] ont présenté et discuté les principales approches utilisées pour la surveillance et la modélisation de la dynamique des maladies infectieuses. Ils ont présenté les concepts de base qui sous-tendent leur mise en œuvre et leur pratique et pour chaque catégorie, ils ont donné une liste annotée d'œuvres représentatives. Dike, C.O., Zainuddin, Z.M., & Dike, I.J. [12] proposent la technique de mise en file d'attente comme une approche mathématique prometteuse et efficace pour l'étude de la transmission et du contrôle de l'épidémie d'Ebola. De fait, les déductions de l'application de cette théorie (de files d'attente) ont été utilisées pour analyser l'épidémie d'Ebola en Guinée en 2014.

1.7 Synthèse des travaux sur la modélisation mathématique du covid-19

Dans son étude, Nesteruk I. a utilisé le modèle SIR (Susceptible-Infectious-Removed) populaire pour obtenir des valeurs optimales pour les paramètres du modèle en utilisant une approche statistique et a donc prédit le nombre de personnes infectées du COVID-19, sensibles et retirées en fonction du temps [28]. Okhuuse V A a montré que l'équilibre sans maladie (E_0) pour le coronavirus COVID-19 ne satisfait pas aux critères d'une stabilité asymptotique locale ou globale [29]. Cela est dû au fait qu'en tant que pandémie déclarée par l'OMS (2020), le coronavirus COVID-19 n'a pas encore de vaccin curatif et des mesures de précaution sont recommandées par le biais de procédures de quarantaine et d'observatoires. Dans [30], les auteurs ont obtenu des prévisions avec une méthode d'itération simple qui n'a besoin que des valeurs quotidiennes des cas confirmés en entrée. La méthode prend en compte les récupérations et les décès attendus, et elle détermine les taux de croissance quotidiens maximaux autorisés qui s'éloignent de l'augmentation exponentielle vers des nombres stables et en baisse. Les prévisions montrent que les taux de croissance quotidiens devraient être maintenus au moins en dessous de 5% si nous souhaitons voir des plateaux prochainement - malheureusement loin de la réalité dans la plupart des pays à ce jour. Les auteurs ont fourni un exécutable ainsi que le code source pour une application simple de la méthode sur les données d'autres pays. Li Y et al. [21] ont établi un modèle dynamique des maladies infectieuses et un modèle des séries chronologiques pour prédire la tendance et la prévision à court terme de la transmission du COVID-19, ce qui a favorisé l'intervention et la prévention du COVID-19 par les départements à tous les niveaux en Chine continentale et gagner plus de temps pour les essais cliniques. Pour comprendre comment maximiser l'efficacité de politiques de quarantaine tout en atténuant les dommages économiques, Goldsztejn U et al. [13] ont développé un modèle mathématique pour évaluer l'impact des mesures de quarantaine sur les décès, les hospitalisations et la production économique. Ils ont utilisé le modèle SEIR pour simuler la propagation de COVID-19. Ils ont intégré les effets des politiques de quarantaine et ont segmenté la population en fonction des risques sanitaires et de la vulnérabilité économique. Ils ont simulé un scénario de base qui laisse les

restrictions de quarantaine actuelles en place. Ensuite, ils ont simulé un assouplissement rapide des restrictions de quarantaine pour les non-seniors peu de temps après la maîtrise de l'épidémie. Enfin, ils ont assoupli progressivement les restrictions de quarantaine pour les jeunes, tout en maintenant des restrictions extrêmement strictes pour les seniors. Ils ont obtenu des résultats satisfaisants. Le comportement des processus de branchement repose sur un paramètre crucial appelé taux moyen de reproduction (ou taux moyen de contamination ou taux moyen d'infection) et sa comparaison à la valeur critique 1. Mourid T [27] a étudié ces modèles pour comprendre l'importance de la valeur critique 1 sur le comportement du modèle et ses applications en épidémiologie, en particulier dans la propagation du virus COVID-19.

1.8 Conclusion

La propagation du COVID-19 dans la population est affectée par la complexité de nombreux facteurs. Il est difficile d'établir un modèle mathématique de propagation dynamique avec des paramètres à estimer et d'obtenir des résultats de simulation assez précis. Pour mieux comprendre la tendance de transmission de COVID-19, nous avons passé en revue les travaux utilisant la modélisation mathématique pour étudier le comportement des pandémies en général, ensuite ceux qui ont traité le COVID-19 dans le but de mettre en place une méthodologie de travail pour une application sur les données en rapport avec la pandémie COVID-19 dans la wilaya de Béjaia, Algérie.

Références

1. Abid L., (2003). *"Résurgence de la peste en Algérie : la sonnette d'alarme"*, *santémaghreb.com*.
2. Adjabi S., Abderrahmani H., Dehas N. et Aïssani D., (2002). *"Movement's Simulation of the patient in the Khellil Amrane's Hospital"*. In Proceedings of the International Congress **ICORD 2001**, B-e-D, South Africa, pp. 1 and others.
3. Arnobius, (1953). *"Arnobii Adversus nationes : libri VII"*. Ed Concetto Marchesi, Milan.
4. Bacaër N., (2009). *"Histoires de mathématiques et de populations"*. Ed. Cassini, Paris.
5. Bcamath, (2020). *"Mathematical-research-takes-up-the-challenge-of-predicting-the-evolution-of-the-covid-19-epidemic-and-its-socio-economic-impact"*. In <https://wp.bcamath.org/news>.
6. Bernoulli D., (1760). *"Essai d'une nouvelle analyse de la mortalité causée par la petite variole et les avantages de l'inoculation pour la prévenir"*. Histoire de l'académie royale des sciences, Paris.
7. Biraben J. N., (1976). *"Les Hommes et la peste en France et dans les pays européens et méditerranéens (Civilisations et sociétés)"*, Ed. École des hautes études en sciences sociales, France.
8. Choisy M., Guégan J. F. et Rohani P., (2007). "Mathematical Modeling of Infectious Diseases Dynamics", *In Encyclopedia of Infectious Diseases : Modern Methodologies*, Ed John Wiley & Sons, Inc. (sous la direction de Tibayrenc M).
9. Chowell G., Sattenspiel L., Bansal S. et Viboud C., (2016). "Mathematical models to characterize early epidemic growth : A Review", *In Physics of Life Reviews ;* Vol. 18 : 66-97.
10. Crepey P. et Khalatbari A., (2020). *"Coronavirus : les maths pour mieux comprendre l'épidémie"*, in Sciences et Avenir.
11. Dehas N., Adjabi S., et Aïssani D. et Abderahmani H., (2006). *"Evaluation des Performances d'un système de santé : cas de l'hôpital Khellil Amrane (Béjaia)"*. In Actes du Congrès International **GISEH'06 (Gestion et Ingénierie des Systèmes Hospitaliers)**, Luxembourg.
12. Dike C.O., Zainuddin Z.M., et Dike I.J. (2018). Mathematical Models for Mitigating Ebola Virus Disease Transmission : A Review. *Advanced Science Letters*, 24, 3536-3543.
13. Goldsztejn U., Schwartzman D. et Nehorai A., (2020). "Public policy and economic dynamics of COVID-19 spread : a mathematical modeling study", In MedRxiv.
14. Hamdi Cherif M., (2020). *Interview*, Journal Liberté, 08 juin 2020, p. 5.

15. Harris R. C., Sumner T., Knight G. M. et White R. G., (2016). "Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines", *In Human Vaccines et Immunotherapeutics*, Vol. 12(11), 2813-2832.
16. Huang C., Wang Y., Li X., Rem L., Zhao J., Hu Y., et al., (2020). "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China". *The Lancet*.
17. Huppert A. et Katriel G., (2013). "Mathematical modelling and prediction in infectious disease epidemiology", *Clinical Microbiology and Infection*, European Society of Clinical Microbiology and Infectious Diseases, CMI, 19, 999-1005.
18. Jonnalagadda J. M. et Gaddam K., (2016). "Epidemic Analysis and Mathematical Modelling of H1N1(A) with Vaccination". Preprints 2016. 2016080230.
19. Kim S., Lee J. et Jung E., (2017). "Mathematical model of transmission dynamics and optimal control strategies for 2009 A/H1N1 influenza in the Republic of Korea", *In Journal of Theoretical Biology*, 412, 74-85.
20. Laporte J. P., (2016). "Peuplement et catastrophes naturelles dans l'Afrique du Nord ancienne". In Actes du Colloque International "Le Peuplement du Maghreb antique et médiéval", Sousse, Tunisie, pp. 117 - 149.
21. Li Y., Wang B., Peng R., Zhou C., Zhan Y., Liu Z., et al., (2020). "Mathematical Modeling and Epidemic Prediction of COVID-19 and Its Significance to Epidemic Prevention and Control Measures". *In Annals of Infectious Disease and Epidemiology*. Vol. 5(1) : 1052.
22. Liu Y-C. et al., (2020). "COVID-19 : The first documented coronavirus pandemic in history", *Biomedical Journal*.
23. Ljeoma O. J., Chioma I. S., Sunday U. E. et Andrew O., (2015). "Mathematical model of the transmission dynamics of swine flu with vaccination of non newborns", *In International Journal of Mathematics Sciences et Engineering Applications. (IJMSEA)*, Vol. 9 (1), 17 pages.
24. May R. M., (2004). "Uses and abuses of mathematics in biology". *In Science* ; 303 : 790-793.
25. Medlock J., (2002). "Mathematical Modeling of Epidemics". University of Washington. Applied Mathematics Department.
26. Merahi M., (2009). "Processus stochastique de population, applications à la théorie stochastique des épidémies : cas du SIDA", thèse de doctorat en sciences, Université de Mentouri de Constantine, Algérie.
27. Mourid T., (2020). "Modèles probabilistes en épidémiologie et le Covid-19", *Quotidien d'Oran* juin 2020, p. 5.
28. Nesteruk I., (2020). "Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China". In MedRxiv.
29. Okhuuse V. A., (2020). "Mathematical predictions for COVID-19 as a global pandemic", In MedRxiv.
30. Perc M., Gorišek Miksic' N., Slavinec M. et Stožer A., (2020). "Forecasting COVID-19". In *Frontiers in Physics*. Vol. 8(127).
31. Shen M., Peng Z. et Xiao Y., (2020). "Modeling the epidemic trend of the 2019 novel coronavirus outbreak in China". In BioRxiv.
32. Siettos C.I. et Russo L., (2013). "Mathematical modeling of infectious disease dynamics". *In Virulence*. Vol. 4(4), 295-306.
33. World Health Organization (WHO), (2020). Coronavirus. :10
34. Younsi F.Z., (2016). "Mise en place d'un Système d'Information Décisionnel pour le suivi et la prévention des épidémies", thèse de doctorat en sciences, université d'Oran, Algérie.
35. Younsi F.Z., Bounnekar A., Hamdadou D. et Boussaid O., (2015). "SEIR-SW, Simulation Model of Influenza Spread Based on the Small World Network". *In Tsinghua Science & Technology*, 460-473.
36. Zhu N., Zhang D., Wang W., Li X., Yang B., Song J., et al, (2019). "A novel coronavirus from patients with pneumonia in China, 2019". *In the New England Journal of Medicine*, 382(8) :727-733.

Contexte du travail

- *Correspondances*
- *Journal Liberté du 27 juin 2020*
- *El Watan du 9 juillet 2020*
- *Le quotidien d'Oran des 13 et 14 juillet 2020*

I - Message du Directeur du LaMOS

Sam 27/06/2020 10 :10

À : bouda_a@yahoo.fr; Kassa Idjdarene; KATI Djamel Edine; hamid_602001@yahoo.fr; TLIBA Souhil

Synthèse_Modèles_Mathématiques_Covid_19_Adel_Aissani_Lekadir_Juin_2020.pdf
953 Ko

Chers Collègues,

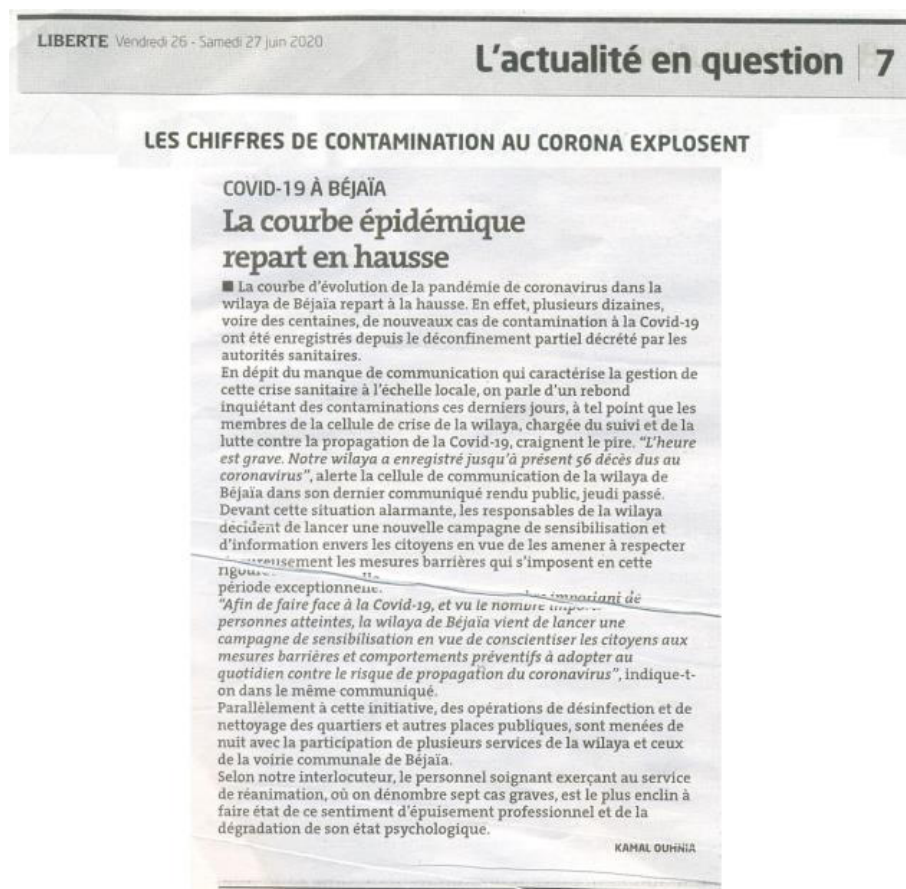
COVID_19. Veuillez trouver ci-joint une synthèse des études commandées ces dernières semaines à des mathématiciens par les pouvoirs publics (de plusieurs pays) chargés de la préparation des plans d'urgence.

Bien à vous

Pr Djamil Aïssani

Directeur de l'Unité de Recherche LaMOS

<http://www.lamos.org>



II - Réponses :

1° Vice Recteur

Bonjour Monsieur Aissani, Merci beaucoup. Bien reçu.

Bien à vous. **Dr KATI Djamel Edine**

Vice-recteur chargé des relations extérieures, de la coopération, de l'animation, de la communication et des manifestations scientifiques.

2° Recteur

Le samedi 27 juin 2020 à 14 :44 :25 UTC+2, ahmed BOUDA <bouda_a@yahoo.fr> a écrit :

À : Vous

Cc :

- Kassa Idjdarene ;
- KATI Djamel Edine ;
- hamid_602001@yahoo.fr ;
- TLIBA Souhil

Merci beaucoup Professeur Aissani pour ce travail. Les collègues de la Faculté de Médecine n'hésiteront pas sans doute à exploiter ces résultats et à vous demander d'éventuelles clarifications.

Bien cordialement

3° Doyen Faculté des Sciences Exactes

Nos félicitations cher collègue Professeur Aissani. Vous avez tous les encouragements de notre part à continuer dans ce projet, qui honore la science en générale et l'université de Bejaia en particulier. Cordiales salutations

4° Doyen Faculté de Médecine

Cher Djamil, Bonjour,

c'est une excellente approche. bravo à toi et à ton équipe.

étant médecin, je suis convaincu que les mathématiques et informatiques apporteront beaucoup de solutions au domaine de la santé.

je suis prêt à mobiliser une équipe pour travailler avec la tienne. comme tu le sais, nous avons une cohorte de malades hospitalisés pour la covid 19. nous pouvons aussi exploiter les données du call center initié par la faculté de médecine. ci-joint, un tableau excel (les noms ont été

supprimés pour la confidentialité).

encore félicitations.

Souhil.

..... **Pr. Souhil TLIBA. MD, PhD ;**
Chef de Service de Neurochirurgie, Chu de Bejaia.
Doyen de la Faculté de Médecine de l'Université de Bejaia
Directeur de Laboratoire de Recherche " Génie Biologique des Cancers "
Tel/Fax : 00213 34 12 32 97/00213 34 22 19 23 Cellulaire : 00213 661 60 97 57
Email : tliba.souh@gmail.com

HB

El Watan

LE QUOTIDIEN INDÉPENDANT - Jeudi 9 juillet 2020

N°9069 - Vingt-huitième année - Prix : Algérie : 30 DA. France : 1 €. USA : 2,15 \$. ISSN : 1111-0333 - http://www.elwatan.com

El Watan - Jeudi 9 juillet 2020 - 2

L'ACTUALITÉ

SITUATION SANITAIRE À BÉJAÏA

Bus et plages bondés et vente de moutons anarchique malgré la gravité de la situation

● Le virus continue à se propager dans un environnement des plus favorables et dans le giron d'une insouciance populaire.

Lundi dernier, Béjaïa a déploré sa 64^e victime emportée par la Covid-19. Au cours des dernières 24 heures, elle a compté 33 nouveaux cas positifs et 20 autres qui attendent les résultats des analyses. Le virus continue à se propager dans un environnement des plus favorables et dans le giron d'une insouciance populaire. Pendant que, dans la ville, les bus circulent bondés d'usagers, sur les plages, les baigneurs sont nombreux à y trouver leur bonheur, avons-nous constaté. Certains viennent des wilayas limitrophes, dont Sétif. Le flux n'est pas celui des masses des précédentes saisons estivales, mais force est de constater que malgré la crise sanitaire, on loue cabanons et appartements aux estivants. Dernièrement, une enquête épidémiologique menée sur des cas testés positifs a révélé, avons-nous appris à la DSP, que des contaminés sont les membres d'une même famille sétifienne qui a campé au village touristique de Capritour. Les foyers de contamination sont multiples et actifs. A l'approche de l'Aïd du sacrifice, plusieurs points de vente de moutons sont apparus anarchiquement dans la ville de Béjaïa, entraînant des rües d'enfants et d'adultes tout autour des enclos. Ce sont autant de facteurs qui promettent de faire perdurer la crise.

Depuis le début de l'épidémie, on a dépassé le chiffre de 1400 contaminés dans la wilaya. Durant les deux dernières semaines, il ne se passe pas un jour sans qu'il y ait 35 cas positifs en moyenne, en piquant parfois jusqu'à la limite des 45 cas. Idris Khodja Hadj, le DSP de la wilaya, est formel : le relâchement dans les gestes barrières est «suicidaire». «La responsabilité est individuelle et collective», nous dit-il. Un «sérieux problème de confiance a creusé un fossé entre citoyens et autorités et empêche de faire entendre raison sur la gravité de la situation, voire sur l'existence de l'épidémie. L'hôpital d'Amizour a vécu une

scène de tension, lorsque de nombreux citoyens sont venus récupérer la dépouille d'un des leurs dont ils ont refusé de croire qu'il est décédé pour cause de Covid-19. Il a fallu beaucoup d'efforts pour calmer ces esprits. Dans le même état d'esprit, les autorités ont «dû intervenir pour confiner de force des personnes chez elles», à en croire le DSP. Le déni de la réalité bloque aussi les enquêtes épidémiologiques, qui sont lancées pour chaque cas. «Les gens refusent de coopérer en donnant, par exemple, de fausses adresses», affirme le DSP.

«On a réussi à faire fléchir l'épidémie. Dans la deuxième quinzaine du mois de mai, les chiffres ont baissé. A l'hôpital de Kherrata, il n'y avait que deux malades, avec la remontée, il y a aujourd'hui 75 malades. En trois semaines, le nombre des contaminés a doublé», déplore-t-il. Les chiffres montent au moment où les malades ne consultent pas tous à temps. Selon notre interlocuteur, «au début de l'épidémie, on avait moins de cas de forme sévère. De plus en plus de personnes consultent tardivement. Elles ne vont consulter qu'au stade terminal».

Les statistiques officielles montrent que 70% des malades décédés ont plus de 65 ans. «La majorité des contaminés sont des gens actifs qui contaminent les personnes âgées», constate Idris Khodja Hadj. L'équation est implacable : nos jeunes condamnent nos vieux.

En juin dernier, le P Nouassria, chef de service des maladies infectieuses au CHU de Béjaïa, avait annoncé que 56% des cas atteints par le Sars-CoV-2 sont âgés entre 26 et 55 ans. La situation est telle, que la pression continue à s'exercer sur les hôpitaux, notamment à Kherrata et Béjaïa, ce qui compromet la reprise de certaines activités médicales suspendues ou reléguées au second plan. «On allait reprendre certaines activités, mais avec cette reprise, nous les avons mises en veilleuse», nous apprend Idris Khodja Hadj. «On

ira vers l'hospitalisation des cas les plus graves seulement», informe-t-il encore.

NOCES INTERDITES

Hier, le wali a interdit aux maires d'établir les actes de mariage. La fermeture des salles de fêtes n'a pas empêché que des couples convolent en justes noces avec DJ et foule. Il y a une semaine, trois cortèges de mariage ont sillonné les rues de la ville dans la même journée. Au même moment, «le personnel médical est à la limite du surmenage», fait remarquer Idris Khodja Hadj. Au surmenage s'ajoute le danger de la contamination. 70 d'entre eux ont eu le virus, dont 7 étaient, à la fin de la semaine dernière, pris en charge à l'école coranique de Kherrata, réquisitionnée comme structure sanitaire annexe de l'hôpital de la ville. «Il y a des médecins et des infirmiers qui ont repris leur service après s'être rétablis d'une contamination. Ils n'ont jamais faibli ou failli», témoigne le DSP. La situation a obligé certains à demeurer, par mesure de prévention, éloignés de leurs familles. Ils sont 770 à loger dans 11 hôtels, la majorité à Béjaïa-ville. 13 autres hôtels restent prêts à héberger le personnel médical et paramédical dans la wilaya. Leur utilisation est tributaire de l'évolution de l'épidémie qui a, cependant, dicté de recourir à des infrastructures extra hospitalières, comme l'Auberge des jeunes de Kherrata où sont hospitalisés des malades du nouveau coronavirus. Au Château de la comtesse d'Aokas se trouvent 23 autres malades. En revanche, les hôpitaux non encore achevés de Tazmalt et de Oued Ghir

ne pourront pas servir. Pour le premier, après résiliation de contrat, une nouvelle entreprise sera installée pour l'achèvement des travaux au bout de cinq mois, et il est attendu d'un autre entrepreneur de finir les VRD. Tandis que pour Oued Ghir, les chambres ne sont pas adaptées pour une prise en charge en réanimation, contrairement à l'EPH de Souk El Terine, où l'on est en phase de réception des équipements. «S'il y a une extrême nécessité, on l'occupera, il nous arrange, parce qu'il y a l'oxygène», assure le DSP. Avec l'apport des donateurs et la contribution du privé, la wilaya dispose de 30 lits de réanimation. «En termes d'espace et de lits, c'est gérable. Les appareils sont là», assure encore le DSP. Lundi dernier, neuf lits de réanimation étaient occupés à l'hôpital Khellil Amrane, selon le surveillant médical Hafid Boudrahem. On devait dégager un nouvel espace au niveau de l'ancien service de pédiatrie. Selon le DSP, les deux dernières semaines, il y a eu constamment 10 malades en réanimation. Si le problème matériel demeure gérable, celui des moyens humains risque de l'être moins. L'urgence dictera de pomper dans le personnel infirmier des polycliniques pour satisfaire les besoins du CHU. Ce qui ne manquera pas de provoquer une pression sur le personnel restant dans ces structures sanitaires de proximité. La demande a été en tout cas faite par le DSP, qui a aussi reçu la liste des médecins de la CNAS et de la DOU engageables en cas de besoin. «Il faut que les gens nous aident à tuer ce virus», implore Idris Khodja Hadj, le DSP de Béjaïa.

K. Medjdoub

DÉBAT

Modèles mathématiques et évolution de la pandémie Covid-19 : synthèse de la littérature

Par Karima Adel-Aïssanou*, Ouiza Lekadir* et Djamil Aïssani*

1re partie

La modélisation des conséquences économiques des épidémies et des pandémies constitue un élément important de la préparation de plans d'urgence pour d'éventuelles nouvelles flambées. La modélisation et la simulation mathématiques permettent une évaluation rapide de la propagation et des réponses aux mesures de précautions. Au fil des ans, un grand nombre d'approches ont été proposées pour examiner le problème sous différents angles. L'objectif de cet article est de présenter une synthèse des travaux réalisés dans le domaine de la modélisation mathématique de l'évolution de la pandémie Covid-19. Dans un premier temps, nous rappelons pourquoi et comment les modèles mathématiques peuvent aider à la prise de décision en ces temps de crise sanitaire Covid-19. Dans un deuxième temps, nous nous attardons sur les spécificités régionales et locales à l'origine de notre travail. Un bref historique des conséquences des épidémies en Afrique du Nord permet de bien en appréhender et en quantifier les risques. Par la suite, nous présentons nos synthèses des travaux pour les épidémies de manière générale et pour le Covid-19 en particulier.

Mots clés : pandémie, Covid-19, modèle mathématique, synthèse de la littérature.

INTRODUCTION : POURQUOI LA MODÉLISATION MATHÉMATIQUE D'UNE ÉPIDÉMIE ?

Depuis le début de l'humanité, l'homme a utilisé les mathématiques pour essayer de résoudre certains problèmes auxquels il était confronté. En même temps, la contribution à la résolution des défis de la société et d'autres disciplines a été source d'inspiration pour de nombreux développements mathématiques à travers l'histoire. La crise sanitaire actuelle générée par le Covid-19 ne fait pas exception et prouve une fois de plus la contribution de cette discipline au bien-être social. [5].

Chaque année, des millions de personnes dans le monde meurent de maladies infectieuses telles que la rougeole, le paludisme, la tuberculose, le VIH. Bien qu'il existe de nombreux facteurs de complication, des modèles mathématiques simples peuvent donner un aperçu de la dynamique des épidémies de maladies et aider les responsables à prendre des décisions concernant les politiques de santé publique. [25].

Une épidémie est une explosion soudaine d'une maladie au sein d'une population donnée et une pandémie est une épidémie qui se propage à travers les régions, qui peuvent être des pays, des continents ou le monde entier. Ces épidémies ou pandémies causent de graves problèmes de développement économique et social qui entraînent un déséquilibre dans la survie des humains et de plusieurs autres espèces vivantes. [18].

Depuis décembre 2019, de nombreux cas inexplicables de pneumonie avec toux, dyspnée, fatigue et fièvre comme principaux symptômes sont survenus à Wuhan, en Chine, en peu de temps [16, 31]. Les autorités sanitaires chinoises et les CDC ont rapidement identifié l'agent pathogène de ces cas comme un nouveau type de coronavirus, que l'OMS (Organisation mondiale de la santé) a nommé Covid-19 le 10 janvier 2020 (World Health Or-

ganization) [33]. A partir de là, de nombreuses équipes de recherche ont abordé différentes études tournant autour des questions suivantes : comment les modèles mathématiques peuvent-ils aider à la prise de décision en ces temps de crise sanitaire ? Et comment parviennent-ils à intégrer les informations et incertitudes en cours ? [10]. Rappelons qu'un modèle est une abstraction (une simplification) de la réalité. Le rôle de la modélisation mathématique d'une épidémie n'est donc pas de reproduire la réalité mais d'essayer de mieux la comprendre, d'évaluer des hypothèses pour indiquer une tendance. Crepey P & Khalatbari A. constatent que les épidémies, tout comme de nombreuses lois naturelles, commencent toutes selon une dynamique qui obéit à une fonction mathématique qui «tend vers l'infini». C'est la fonction exponentielle. Concrètement, le nombre d'infectés augmente de plus en plus vite au cours du temps. Ce modèle grossier du départ de l'épidémie ne tient compte ni des données de la transmission, ni des mesures sanitaires. Par la suite, on doit essayer de voir comment se transmet la maladie. C'est ce qu'on appelle le déroulé (ou bien l'historique) de la maladie : période d'incubation, contagiosité, etc. Ces éléments permettent de concevoir un modèle de base. Celui qui a été élaboré par Crepey P & Khalatbari A. est le modèle SIR, où toute la population peut être partagée en trois états : sains (S), infectés (I), rétablis (R). La dynamique de la maladie dépend des échanges entre ces trois boîtes, aussi appelées «compartiments» [10]. Précisons ici que ce modèle SIR à trois compartiments ne représente pas le Covid-19 car les spécificités de ce virus nécessitent des modèles plus complexes. Ainsi, il faut avoir des éléments sur la durée de l'incubation. Il en est de même pour la présence de personnes asymptomatiques après leur infection. L'équipe de Crepey P a d'abord utilisé un modèle à 4 boîtes, appelé SEIR en y ajoutant une boîte supplémentaire E qui désigne les personnes ayant été exposées mais qui ne présentent pas encore de symptômes. Cette équipe poursuit ses investigations en profondeur sur un modèle à 8 compartiments qui sera capable de tenir compte des personnes qui restent asymptomatiques après avoir été infectées, de celles qui commencent à être contagieuses quelques jours avant l'apparition des symptômes, des personnes hospitalisées et enfin des personnes contaminées qui demeurent confinées au sein de leurs familles [10].

L'objectif de cet article est de présenter une synthèse des travaux réalisés dans le domaine de la modélisation mathématique de l'évolution de la pandémie Covid-19. Dans un premier temps, nous rappelons pourquoi et comment les modèles mathématiques peuvent aider à la prise de décision en ces temps de crise sanitaire Covid-19. Dans un deuxième temps, nous nous attardons sur les spécificités régionales et locales à l'origine de notre travail. Un bref historique des conséquences des épidémies en Afrique du Nord permet de bien en appréhender et en quantifier les risques. Par la suite, nous présentons nos synthèses des travaux de modélisation mathématiques réalisés pour les épidémies de manière générale et pour le Covid-19 en particulier.

Suite en page 19

* Unité de Recherche LaMOS (Modélisation et optimisation des systèmes), Université de Béjaïa

Modèles mathématiques et évolution de la pandémie Covid-19 : synthèse de la littérature

Suite de la page 18

2- A L'ORIGINE DE CETTE SYNTHÈSE SUR LA MODÉLISATION MATHÉMATIQUE DES ÉPIDÉMIES

Depuis une vingtaine d'années, l'Unité de recherche LaMOS (Modélisation et optimisation des systèmes – <http://www.lamos.org>) travaille occasionnellement avec les structures de santé de Béjaïa (au début, l'hôpital Khellil Amrane, puis, après leur création, avec la faculté de médecine et le CHU) pour aborder des problèmes d'évaluation de performance des systèmes de santé (voir par exemple [2, 11]). Ces travaux se font dans le cadre d'un groupe pluridisciplinaire. Au début de la crise du Covid-19 en Algérie, nous avons été en contact avec la cellule de crise de la wilaya de Béjaïa. Il s'agissait d'aborder la question de la modélisation mathématique de l'évolution de l'épidémie à l'échelle locale. A titre d'exemple, il s'agissait de pouvoir répondre à des questions de type : au bout de combien de jours les capacités d'accueil des services spécialisés (et notamment de celui de réanimation) seront-elles atteintes ? A ce niveau, une récolte des données est essentielle. Un questionnaire avait été élaboré pour chaque patient (début des symptômes, date du test, date de l'hospitalisation, date du résultat du test, date du début du traitement, date de la dégradation de l'état de santé...). La fiabilité des résultats dépend de la rigueur des personnels de santé chargés du suivi (et donc de la récolte des données). A cette époque, le laboratoire Covid-19 de l'Université de Béjaïa n'avait pas encore été envisagé et la wilaya dépendait des performances de l'Institut Pasteur d'Alger.

Certes, la courbe de l'augmentation exponentielle de la morbidité et de la mortalité algérienne est nettement réduite comparée à celles des pays occidentaux. D'un autre côté, les données observées de la Covid-19 à ce jour sont très en dessous des chiffres prévus par les modélisations mathématiques. Le Pr Hamdi Cherif M. [14] estime que « le challenge pour sortir de ce tunnel est la mise en œuvre d'un plan de déconfinement qui doit d'abord tenir compte des points faibles des confinements antérieurs et les spécificités locales et régionales, avec l'espoir que notre écosystème diminuera la virulence et la contagion du virus » [14]. Et il ajoute « qu'il est temps d'aller sur le terrain pour des études critiques des différentes situations épidémiologiques en vue de mettre en place des stratégies locales, spécifiques et adaptées pour chaque wilaya ou région » [14]. En effet, l'éclatement des foyers épidémiques (clusters) de la Covid-19 constitue aujourd'hui le défi pour les autorités sanitaires alors que la première phase de déconfinement est enclenchée. Le bulletin épidémiologique de l'INPS (Institut national de santé publique) du 04 juin 2020 portant sur les cas cumulés confirmés positifs par PCR note qu'à cette date, le taux d'incidence nationale était de 23.08 cas pour 100.000 habitants. Mais on doit tenir compte de certaines spécificités locales et régionales. Ainsi, pour la wilaya de Béjaïa, c'est la région de Kherrata qui a été la plus touchée. Or elle est frontalière avec la wilaya de Sétif dont l'incidence est de loin supérieure au taux national (34.64).

3- HISTORIQUE DES ÉPIDÉMIES EN AFRIQUE DU NORD

A l'instar de ce qui se passe partout sur la planète, l'Afrique du Nord et l'espace méditerranéen n'ont pas été épargnés par les catastrophes naturelles : séismes, inondations, épidémies. Certaines ont été ponctuelles, d'autres ont dévasté des régions entières, ont fait le tour de la Méditerranée, en détruisant une part importante de la population, vidant des pays peuplés, désorganisant des empires, modifiant les rapports de force. Suivant leur origine et

leurs circonstances, leurs effets ont été parfois immédiats, parfois différés, courts ou longs, directs ou indirects [20]. Déjà dans la seconde moitié du IIe siècle, Arnobe, à propos des épidémies qui consomment le genre humain, affirme : « Consultez les livres d'annales rédigés en différentes langues, vous apprendrez que des pays entiers ont été souvent dépeuplés, vidés de leurs habitants » [3].

Les textes anciens mentionnent sous le nom de pestes toutes sortes de maladies à forme épidémique, au pronostic grave, sinon fatal. D'un point de vue médical, toutes n'étaient pas des pestes à proprement parler. Les descriptions sont souvent trop vagues pour qu'on puisse en saisir la nature exacte, et guère plus leur ampleur et leur impact. Laporte J P souligne que « suivant les cas, les épidémies avaient une influence directe et indirecte considérable dans tous les domaines économique, démographique, social, politique, militaire, religieux, artistique, sanitaire et médical ». A court terme, elles interrompaient la circulation des hommes, et donc celle des denrées, entraînaient, elles aussi, la famine [20].

A cette époque, les éléments de diagnostic étaient limités, mais le tableau clinique était très important, avec une série de symptômes identifiés : vomissements, charbon, hémorragie, symptômes nerveux, crachements de sang de la forme pneumonique... La mortalité pouvait atteindre jusqu'à plus de 50% de la population suivant les épisodes locaux. La guerre et les échanges favorisaient l'extension de la peste, en affaiblissant la résistance des populations.

Biraben J. N., considéré comme étant le grand spécialiste de l'histoire de la peste, analyse le cas de la peste justinienne qui apparut en 541 dans le port égyptien de Péluse, venant semble-t-il d'Éthiopie [7]. De là, elle s'étendit sur l'Égypte jusqu'à Alexandrie, la Palestine, la Syrie. Quelques mois plus tard, elle est apparue à Constantinople, où elle dura quatre mois. Elle fut ensuite diffusée rapidement dans certains ports méditerranéens par la voie maritime. L'intérieur des pays était moins touché, car la population était moins dense. Biraben J. N. affirme que ce sont les soldats de Justinien envoyés de Constantinople à Carthage pour faire campagne contre les Maures qui y auraient apporté la peste. Il parle d'une diminution de 50% de la population de Constantinople (40% de morts, 10% de fuites) [7].

Si on se projette maintenant au présent, on constate que l'Algérie a connu ces dernières décennies une incroyable série de catastrophes : en plus des calamités dites « naturelles » (inondation de Bab-El-Oued en novembre 2001, plus de 800 morts; séisme de Boumerdes du 21 mai 2003 : 2.300 morts, autant de disparus et 11.000 blessés) une succession d'épidémies de botulisme, tuberculose (2001), typhoïde (2002), rougeole, méningite (2002), peste (juin 2003), dans la banlieue d'Oran... Abid L. rappelle que cette ville a justement servi de cadre à Albert Camus pour son roman mythique, « La Peste » où dans la conclusion, prémonitrice (de ce roman) Camus fait

dire à son héros : « Et (Rieux) savait ce qu'on peut lire dans les livres que le bacille de la peste ne meurt ni ne disparaît jamais, qu'il peut rester, pendant des dizaines d'années, endormi dans les meubles et le linge, qu'il attend patiemment dans les chambres, les caves... Et que, peut-être, le jour viendrait où, pour le malheur et l'enseignement des hommes, la peste réveillerait ses rats et les enverrait mourir dans une cité heureuse » [1].

4- HISTORIQUE DE L'ÉVOLUTION DU COVID-19

Actuellement, des personnes du monde entier ont été touchées par la maladie du coronavirus 2019 (Covid-19), qui est la cinquième pandémie après celle de la grippe de 1918. À partir de maintenant, nous pouvons retracer le premier rapport et la flambée subséquente d'un groupe de nouveaux cas de pneumonie humaine dans la ville de Wuhan, en Chine, en décembre 2019. La première date d'apparition des symptômes était le 1^{er} décembre 2019 [22, 16, 36]. Le 11 mars 2020, l'augmentation rapide du nombre de cas en dehors de la Chine a conduit le docteur Tedros Adhanom Ghebreyesus, directeur général de l'OMS, à annoncer que la flambée pouvait être qualifiée de pandémie.

À la mi-mars 2020, la région européenne de l'OMS était devenue l'épicentre de l'épidémie, notifiant plus de 40% des cas confirmés dans le monde. Au 28 avril 2020, la région concourait à 63% de la mortalité mondiale due au virus.

Le premier cas de Covid-19 en Afrique est apparu en février 2020 en Égypte. 53 pays africains sur 54 sont désormais touchés par le coronavirus. Depuis, plusieurs études trop alarmistes ont exploré la piste d'une « catastrophe africaine », qui, près de trois mois après le début de la crise, n'a pas eu lieu. Alors que l'Europe et les États-Unis ont dépassé le million de cas, le nombre officiel de contaminations en Afrique s'élève, à la mi-mai, à 75.530 cas enregistrés, selon le Centre pour la prévention et le contrôle des maladies de l'Union africaine (CDC Afrique).

Le premier cas confirmé de Covid-19 aux USA est annoncé en janvier 2020. Les États-Unis et le Brésil sont les pays où la pandémie de Covid-19 progresse le plus rapidement.

5- LES MODÈLES MATHÉMATIQUES MODÉLISANT LES PANDÉMIES

Siettos C I & Russo L [32] ont présenté et discuté les principales approches utilisées pour la surveillance et la modélisation de la dynamique des maladies infectieuses. Ils ont présenté les concepts de base qui expliquent leur mise en œuvre et leur pratique et, pour chaque catégorie, ils ont donné une liste annotée d'œuvres représentatives. La figure 1 donne un aperçu des modèles mathématiques pour les maladies infectieuses

Au fil des ans, un grand nombre d'approches ont été proposées pour examiner le problème sous différents angles. Ceux-ci englobent trois catégories générales (voir figure 1) : (1) les méthodes statistiques de surveillance des épidémies et d'identification des schémas spatiaux dans les épidémies réelles, (2) les modèles mathématiques dans le cadre de systèmes dynamiques (également appelés modèles de l'espace d'état) utilisés pour prévoir l'évolution d'une propagation épidémique « hypothétique » ou en cours, et (3) apprentissage automatique / méthodes expertes pour prévoir l'évolution d'une épidémie en cours. Pour ces trois catégories, il existe à nouveau des approches différentes tissant une littérature vaste et diversifiée.

6- SYNTHÈSE DES TRAVAUX SUR LA MODÉLISATION MATHÉMATIQUE DES ÉPIDÉMIES

En 1760, Bernoulli D. a développé un modèle mathématique d'une épidémie de variole appelée à l'époque « la petite variole » pour analyser la mortalité causée par cette maladie [6]. Son modèle est connu comme étant le premier modèle mathématique de l'histoire de l'épidémiologie [4, 34].

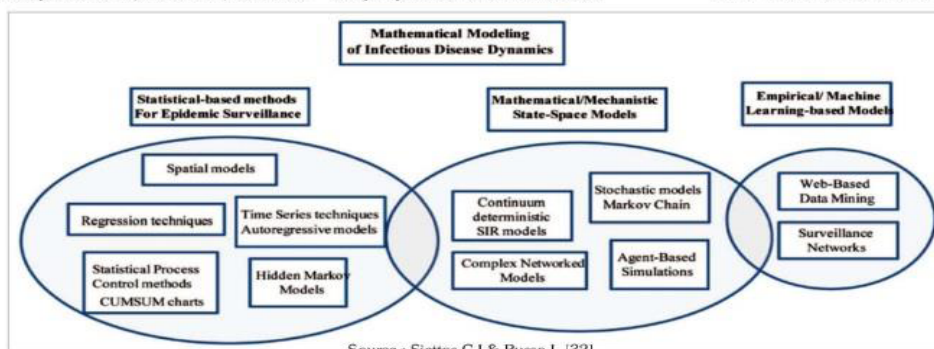
Dans l'article [35], les auteurs donnent une bonne revue de littérature concernant la modélisation mathématique de la grippe saisonnière. Ils ont modélisé la propagation d'une épidémie de grippe dans la population d'Oran, en Algérie. En particulier, ils ont construit un modèle SEIR-SW (Susceptible-Exposed-Infected-Removed dans un réseau Small World) pour identifier les principales caractéristiques de la transmission épidémique et son évolution dans le temps.

Dans [17], les auteurs ont précisé dans quelle mesure les modèles mathématiques de transmission de maladies fournissent des prédictions fiables. Ils ont conclu qu'afin d'utiliser efficacement les modèles mathématiques pour étudier la dynamique de la propagation d'une pandémie, y compris les stratégies de contrôle possibles, il faut être sûr que les valeurs utilisées pour les différents paramètres du modèle correspondent à la réalité. Bien que certains paramètres puissent être déterminés sur la base de connaissances antérieures, d'autres paramètres doivent être estimés en ajustant le modèle aux données disponibles.

Un avantage important de l'utilisation de modèles est que la représentation mathématique des processus biologiques permet la transparence et la précision des hypothèses épidémiologiques, nous permettant ainsi de tester notre compréhension de l'épidémiologie de la maladie en comparant les résultats des modèles et les modèles observés [24].

D'autres chercheurs ont utilisé la modélisation mathématique dans la prédiction de la propagation des épidémies.

A suivre...
Karima Adel-Aïssanou,
Ouiza Lekadir et Djamil Aïssani



Source : Siettos C I & Russo L [32].

Le Quotidien

Edition Nationale d'Information D'ORAN

DÉBAT

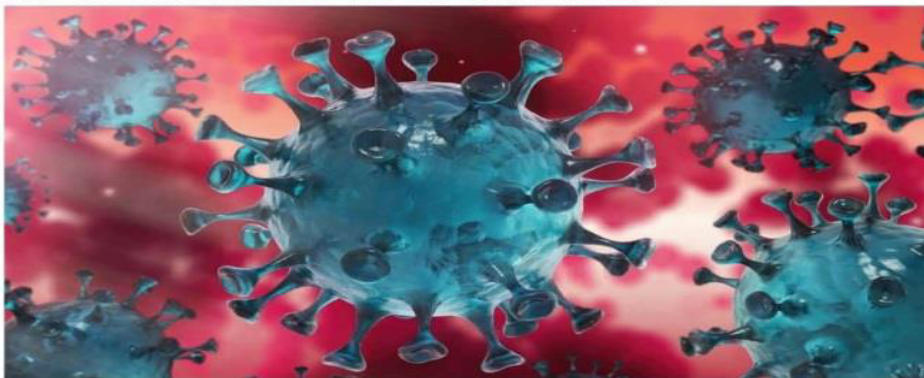
Le Quotidien d'Oran
Mardi 14 juillet 2020 19

Modèles mathématiques et évolution de la pandémie Covid-19 : synthèse de la littérature

Par Karima Adel-Aissanou*,
Ouiza Lekadir* et Djamil Aissani*

Suite et fin

Dans [19], les auteurs ont développé un modèle mathématique pour la dynamique de transmission de l'épidémie de grippe A / H1N1 de 2009 en République de Corée. Dans [8], les auteurs donnent en détail une modélisation mathématique des épidémies et comment interpréter les résultats. Les auteurs du papier [12] se sont préoccupés des conséquences économiques qui auraient pu se produire si la maladie d'Ebola s'était propagée à certains pays en développement de la région Asie-Pacifique. Leurs estimations ont été effectuées avec deux modèles liés : un modèle de transmission des maladies et un modèle multirégional d'équilibre général calculable dynamique (EGC) de l'économie mondiale. Dans [26], l'auteur a donné une modélisation mathématique de la propagation de l'épidémie de sida. Les auteurs dans [23] ont présenté un modèle mathématique de transmission de la grippe porcine avec la vaccination de nouveaux-nés. Dans [9], les auteurs ont passé en revue les progrès récents dans la modélisation et la caractérisation des schémas de croissance épidémique précoce à partir des données sur les flambées de maladies infectieuses, et ont étudié les types de formulations mathématiques qui sont les plus utiles pour saisir une diversité de profils de croissance épidémique précoce, allant de la dynamique de croissance sub-exponentielle à la dynamique de croissance exponentielle. Plus précisément, ils ont examiné des modèles mathématiques qui intègrent des détails spatiaux ou des structures de mélange de population réalistes, y compris des modèles de métapopulation, des modèles de réseaux individuels et des modèles simples de type SIR qui intègrent les effets des changements de comportement réactifs ou de mélange inhomogène. Dans [15], les auteurs ont effectué une revue systématique des modèles mathématiques estimant l'impact épidémiologique des futurs vaccins antituberculeux humains. Avec une méthode rigoureuse, ils ont affirmé que les évaluations économiques ont révélé que les vaccins antituberculeux étaient extrêmement rentables, en particulier lorsqu'ils étaient destinés aux adolescents/adultes. Les auteurs de l'article [32] ont présenté et discuté les principales approches utilisées pour la surveillance et la modélisation de la dynamique des maladies infectieuses. Ils ont présenté les concepts de base qui soutiennent leur mise en œuvre et leur pratique et pour chaque catégorie, ils ont donné une liste annotée d'œuvres représentatives. Dike,



C.O., Zainuddin, Z.M., & Dike, I.J. [12] proposent la technique de mise en file d'attente comme une approche mathématique prometteuse et efficace pour l'étude de la transmission et du contrôle de l'épidémie d'Ebola. De fait, les déductions de l'application de cette théorie (de files d'attente) ont été utilisées pour analyser l'épidémie d'Ebola en Guinée en 2014.

7. SYNTHÈSE DES TRAVAUX SUR LA MODÉLISATION MATHÉMATIQUE DU COVID-19

Dans son étude, Nesteruk I. a utilisé le modèle SIR (Susceptible-Infectieux-Removed) populaire pour obtenir des valeurs optimales pour les paramètres du modèle en utilisant une approche statistique et a donc prédit le nombre de personnes infectées du Covid-19, sensibles et retirées en fonction du temps [28]. Okhuse V A a montré que l'équilibre sans maladie ($S_{0,0}$) pour le coronavirus Covid-19 ne satisfait pas aux critères d'une stabilité asymptotique locale ou globale [29]. Cela est dû au fait qu'en tant que pandémie déclarée par l'OMS (2020), le coronavirus Covid-19 n'a pas encore de vaccin curatif et des mesures de prévention sont recommandées par l'OMS [2020], les auteurs ont obtenu des prévisions avec une méthode d'itération simple qui n'a besoin que des valeurs quotidiennes des cas confirmés en entrée. La méthode prend en compte les récupérations et les décès attendus, et elle détermine les taux de croissance quotidiens maximaux autorisés qui s'éloi-

gnent de l'augmentation exponentielle vers des nombres stables et en baisse. Les prévisions montrent que les taux de croissance quotidiens devraient être maintenus au moins en dessous de 5% si nous souhaitons voir des plateaux prochainement - malheureusement loin de la réalité dans la plupart des pays à ce jour. Les auteurs ont fourni un exécutable ainsi que le code source pour une application simple de la méthode sur les données d'autres pays. Li Y et al. [21] ont établi un modèle dynamique des maladies infectieuses et un modèle des séries chronologiques pour prédire la tendance et la prévision à court terme de la transmission du Covid-19, ce qui a favorisé l'intervention et la prévention du Covid-19 par les départements à tous les niveaux en Chine continentale et gagner plus de temps pour les essais cliniques. Pour comprendre comment maximiser l'efficacité de politiques de quarantaine tout en atténuant les dommages économiques, Goldsztejn U et al. [13] ont développé un modèle mathématique pour évaluer l'impact des mesures de quarantaine sur les décès, les hospitalisations et la production économique. Ils ont utilisé le modèle SEIR pour simuler la propagation de Covid-19. Ils ont intégré les effets des politiques de quarantaine et ont segmenté la population en fonction des risques sanitaires et de la vulnérabilité économique. Ils ont simulé un scénario de base qui laisse les restrictions de quarantaine actuelles en place. Ensuite, ils ont simulé un assouplissement rapide des restrictions de quarantaine pour les non seniors peu de temps après la maîtrise de l'épidémie. Enfin, ils ont

assoupli progressivement les restrictions de quarantaine pour les jeunes, tout en maintenant des restrictions extrêmement strictes pour les seniors. Ils ont obtenu des résultats satisfaisants. Le comportement des processus de branchement repose sur un paramètre crucial appelé taux moyen de reproduction (ou taux moyen de contamination ou taux moyen d'infection) et sa comparaison à la valeur critique T . Mourid T [27] a étudié ces modèles pour comprendre l'importance de la valeur critique T sur le comportement du modèle et ses applications en épidémiologie, en particulier dans la propagation du virus Covid-19.

CONCLUSION

La propagation du Covid-19 dans la population est affectée par la complexité de nombreux facteurs. Il est difficile d'établir un modèle mathématique de propagation dynamique avec des paramètres à estimer et d'obtenir des résultats de simulation assez précis. Pour mieux comprendre la tendance de transmission de Covid-19, nous avons passé en revue les travaux utilisant la modélisation mathématique pour étudier le comportement des pandémies en général, ensuite ceux qui ont traité le Covid-19 dans le but de mettre en place une méthodologie de travail pour une application sur les données en rapport avec la pandémie Covid-19 dans la wilaya de Béjaïa, Algérie.

* Unité de Recherche LaMOS (Modélisation et optimisation des systèmes), Université de Béjaïa

Références bibliographiques

- [1] Abid L. (2003). «Résumé de la peste en Algérie : la sonnette d'alarme ». *santé.émagreb.com*.
- [2] Adjabi S., Abderahmani H., Dehas N. & Aissani D. (2002). «Movement's Simulation of the patient in the Khellil Anrane's Hospital». In Proceedings of the International Congress ICORD 2001, B-e-D, South Africa, pp. 1 and others.
- [3] Arnobius. (1953). «*Arnobii Adversus nationes libri VII*». Ed. Concetta Marchese, Milan.
- [4] Bacsin N. (2009). «*Histoires de mathématiques et de populations*». Ed. Cassini, Paris.
- [5] Bcamath. (2020). «Mathematical research takes up the challenge of predicting the evolution of the Covid-19 epidemic and its socio-economic impact». In <https://www.bcamath.org/news>.
- [6] Bernoulli D., (1760). «Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et les avantages de l'oculation pour la prévention». Histoire de l'Académie royale des sciences, Paris.
- [7] Binben J N. (1976). «*Les Hommes et la peste en France et dans les pays européens et méditerranéens (Civilisations et sociétés)*». Ed. Ecole des hautes études en sciences sociales, France.
- [8] Chioisy M., Guégan J F & Rohani P. (2007). «Mathematical Modeling of Infectious Diseases Dynamics». In *Encyclopedia of Infectious Diseases: Modern Methodologies*, Ed. John Wiley & Sons, Inc. (sous la direction de Tibaeync M).
- [9] Chowell G., Sattenspiel L., Bansal S & Viboud C. (2016). «Mathematical models to characterize entry epidemic growth: A Review». In *Physics of Life Reviews*, Vol. 18, 66-97.
- [10] Crepey P & Khalafbari A. (2020). «Coronavirus : les maths pour mieux comprendre l'épidémie». In Sciences et Avenir.
- [11] Dehas N., Adjabi S. et Aissani D. & Abderahmani H. (2006). «Évaluation des Performances d'un système de santé : cas de l'hôpital Khellil Anrane (Béjaïa)». In Actes du Congrès International GISEH'06 (Gestion et Ingénierie des Systèmes Hospitaliers), Luxembourg.
- [12] Dike, C.O., Zainuddin, Z.M., & Dike, I.J. (2018). Mathematical Models for Mitigating Ebola Virus Disease Transmission: A Review. *Advanced Science Letters*, 24, 3536-3543.
- [13] Goldsztejn U., Schwartzman D & Nehorai A. (2020). «Public policy and economic dynamics of COVID-19 spread: a mathematical modeling study». In *MedRxiv*.
- [14] Hamdi Cherif M. (2020). Interview, *Journal Liberté*, 08 juin 2020, p. 5.
- [15] Harris R C., Sumner T, Knight G M & White R G. (2016). «Systematic review of mathematical models exploring the epidemiological impact of future TB vaccines». In *Human Vaccines & Immunotherapeutics*, Vol. 12(11), 2813-2832.
- [16] Huang C., Wang Y, Li X, Ren L, Zhao J, Hu Y, et al., (2020). «Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China». *The Lancet*.
- [17] Huppert A & Katriel G. (2013). «Mathematical modelling and prediction in infectious disease epidemiology». *Clinical Microbiology and Infections*, European Society of Clinical Microbiology and Infectious Diseases, CMJ, 19, 999-1005.
- [18] Jomnagadda J M & Gaddam K. (2016). «Epidemic Analysis and Mathematical Modelling of H1N1(A) with Vaccination». *Preprints 2016*, 20160802250.
- [19] Kim S., Lee J & Jung E. (2017). «Mathematical model of transmission dynamics and optimal control strategies for 2009 A/H1N1 influenza in the Republic of Korea». In *Journal of Theoretical Biology*, 412, 74-85.
- [20] Laporte J P. (2016). «Peuplement et onatophores nativistes dans l'Afrique du Nord ancienne». In Actes du Colloque International «Le Peuplement du Maghreb antique et médiéval », Souise, Tunisie, pp. 117 - 149.
- [21] Li Y, Wang B, Peng R, Zhou C, Zhan Y, Liu Z, et al., (2020). «Mathematical Modeling and Epidemic Prediction of Covid-19 and Its Significance to Epidemic Prevention and Control Measures». In *Annals of Infectious Disease and Epidemiology*, Vol. 5(1): 1052.
- [22] Liu Y-C et al., (2020). «Covid-19: The first documented coronavirus pandemic in history». *Biomedical Journal*.
- [23] Ljeoma O J, Chioma I S, Sunday U E & Andrew O. (2015). «Mathematical model of the transmission dynamics of swine flu with vaccination of non newborns». In *International Journal of Mathematics Sciences & Engineering Applications*. (IJMS&EA), Vol. 9 (11), 17 pages.
- [24] May R M. (2004). «Uses and abuses of mathematics in biology». In *Science*, 303: 790-793.
- [25] Medlock J. (2002). «Mathematical Modeling of Epidemics». University of Washington, Applied Mathematics Department.
- [26] Merahi M. (2009). «Processus stochastique de population, applications à la théorie stochastique des épidémies : cas du SIDA», thèse de doctorat en sciences, Université de Moutouria de Constantine.
- [27] Mourid T. (2020). «Modèles probabilistes en épidémiologie et le Covid-19». *Le Quotidien d'Oran* juin 2020, p. 5.
- [28] Nesteruk I. (2020). «Statistics based predictions of coronavirus 2019-nCoV spreading in mainland China». In *MedRxiv*.
- [29] Okhuse V A. (2020). «Mathematical predictions for COVID-19 as a global pandemic». In *MedRxiv*.
- [30] Perc M, Goršek Milotic N, Slawinec M & Stojar A. (2020). Forecasting Covid-19. In *Frontiers in Physics*, Vol. 8(127).
- [31] Shen M, Peng Z & Xiao Y. (2020). «Modeling the epidemic trend of the 2019 novel coronavirus outbreak in China». In *MedRxiv*.
- [32] Siettos C I & Russo L. (2013). «Mathematical modeling of infectious disease dynamics». In *Virusland*, Vol. 4(4), 295-306.
- [33] World Health Organization (WHO). (2020). *Coronavirus:10*
- [34] Younsi F Z. (2016). «*Étude en place d'un Système d'Information Décisionnel pour le suivi et la prévention des épidémies*», thèse de doctorat en sciences, université d'Oran, Algérie.
- [35] Younsi F Z., Bounekar A., Hamdadou D & Boussaid O. (2015). «SER-SW, Simulation Model of Influenza Spread Based on the Small World Network». In *Tarbiyah Science & Technology*, 460-473.
- [36] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al., (2019). «A novel coronavirus from patients with pneumonia in China, 2019». In *The New England Journal of Medicine*, 382(8) :727-733.



LA WILAYA RECENSE 1 176 MALADES ET 75 DÉCÈS CONFIRMÉS PAR LA PCR

Flambée exponentielle depuis le déconfinement à Béjaïa

Depuis l'apparition de l'épidémie de Covid-19 dans la wilaya de Béjaïa et jusqu'à avant-hier, la région a enregistré, selon le directeur de la santé de la wilaya, pas moins de 1176 malades et 75 décès dus au coronavirus. Un bilan officiel confirmé par la PCR sur les 2342 cas suspects. L'épidémie a connu une flambée exponentielle depuis le début du déconfinement, nous a déclaré Idris Khodja Hadj, DSP de la wilaya de Béjaïa.

Une flambée inquiétante, soutient-il, due au relâchement de la population dans le respect strict des gestes barrières. "Toutes les communes de la wilaya sont presque touchées par la propagation du virus depuis le déconfinement. La situation est inquiétante dans 17 communes de la wilaya", signale notre interlocuteur. Il s'agit de toutes les communes de la côte est, du chef-lieu de wilaya et de certaines communes de la vallée de la Soummam.

Pour étayer ses propos, le DSP signale qu'"avant le déconfinement, on enregistrait deux hospitalisations par jour,

mais maintenant, nous sommes à 80 hospitalisations". "Avant, c'était des clusters professionnels. Maintenant, c'est communautaire. Le virus actuellement circule dans la société du fait au relâchement de la population", regrette Idris Khodja Hadj. Une croissance de la courbe des contaminés et des décès dans la région qui a contraint, signale notre interlocuteur, sa direction à prendre des mesures palliatives pour l'hospitalisation des malades Covid-19 par la réquisition des structures comme, à titre d'exemple, le château de la Comtesse d'Aokas et l'Auberge de jeunes de Kherrata, en sus de l'ouverture d'un autre étage au CHU Khelil-Amrane, dédié au Covid-19. Des mesures palliatives prises du fait du débordement des structures sanitaires. Devant cette flambée de l'épidémie, Le DSP ne cache pas que le personnel de la santé est au bord de la dépression. "La pression s'exerce au quotidien sur le personnel de la santé. La situation est telle que nous sommes obligés de faire travailler des infirmiers stagiaires à l'école paramédicale d'Aokas

pour renforcer les CHU", signale le DSP à cet effet, et qui précise, au passage, que 53 malades et 2 décès Covid-19 sont enregistrés dans les rangs du personnel sanitaire.

"Actuellement, nous avons les moyens pour contenir la situation, mais si la population ne se ressaisit pas, nous risquons d'être dépassés", avertit solennellement le DSP. Ce dernier appelle au sursaut de la population pour le respect des mesures barrières et ne pas croire au schisme de certains. "Le virus circule et le danger est permanent", déclare avec force Khodja Hadj à l'adresse de la population. Et d'ajouter: "La responsabilité est collective et individuelle." Une réunion de la cellule de crise de wilaya s'est tenue, hier, au siège de la wilaya pour étudier la situation préoccupante de l'épidémie de coronavirus et les nouvelles mesures à prendre dans la région par le wali. Il n'est pas exclu, selon certains observateurs, que le wali instaure un nouveau confinement dans certaines communes très touchées par le rebond de cette épidémie.

L. OUBIRA

L'Impact de la Crise Sanitaire sur le Trafic Portuaire de Béjaïa(EPB)

Ramdani A., Ghiles L., Saadi N. et Aïssani D.

Research Unit **LaMOS** (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria
lamos_bejaia@hotmail.com ,

Résumé L'application des méthodes de la recherche opérationnelle au niveau des entreprises est indispensable pour améliorer les différents processus de service et production, c'est à cet effet qu'il est plus nécessaire pour une entreprise d'établir une planification future dans toutes les filières et les secteurs, un préalable indispensable pour réduire les risques ou encore annexer plus de bénéfice, et cela se fait notamment avec les prévisions. Dans ce travail, nous avons étudié un phénomène clé de la prévision relevé durant notre stage au sein de l'EPB. L'objectif de notre étude est de traiter l'évolution de trafic des hydrocarbures et le trafic des conteneurs durant l'ère de la crise Covid- 19 et de prévoir les quantités de tonnage des hydrocarbures qui seront engendrées pour l'année 2021. Des modèles mathématiques sont formulés sur la base de la modélisation statistique et linéaire qui reflètent mieux les caractéristiques des systèmes étudiés. Dans un premier temps, nous avons modélisé les deux séries chronologiques en faisant appel à la méthode de Box et Jenkins et aux calculs des prévisions. Dans un deuxième temps, avec le principe de modélisation linéaire, nous avons établi un modèle qui sera ensuite appliqué pour le calcul des prévisions. La comparaison entre les résultats obtenus et les valeurs réelles permet d'apprécier l'impact de la crise sanitaire sur les deux terminaux.

Mots clés : Entreprise Portuaire de Bejaia ; Modélisation statistique ; Prévision ; Covid-19.

2.1 Introduction

La statistique est une science qui permet l'exploitation d'une masse d'informations provenant de l'observation d'un phénomène. Elle concerne un ensemble de méthodes mathématiques indispensables à l'analyse de grands ensembles de données et dont l'objectif est de fournir les éléments nécessaires à l'appréciation d'une situation et l'aide à la décision. Il existe, de nos jours, un ensemble de méthodes rigoureuses permettant d'analyser et de traiter des données concernant un phénomène et réaliser une étude qui permettra de comprendre son évolution et effectuer des prévisions, notamment sur les séries temporelles et régression linéaire. L'étude des séries temporelles correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Cette étude est appliquée dans plusieurs domaines, en météorologie, en économie, etc. Elle consiste à appliquer les méthodes de prévision. On distingue notamment, deux méthodes qui sont souvent utilisées, la méthode de Box et Jenkins (1976) et le lissage exponentiel. La première est très populaire grâce à sa simplicité et ses résultats fiables. La deuxième méthode a l'avantage de fournir une prévision "bon marché" et parfois très satisfaisante, mais les deux méthodes sont parfois loin d'être les mieux adaptées. L'analyse de régression peut être définie comme la recherche de la relation stochastique qui lie deux ou plusieurs variables. Son champ d'application recouvre de multiples domaines, parmi lesquels on peut citer la physique, l'astronomie, la biologie, la chimie, la médecine, la géographie, la sociologie, l'histoire, l'économie, la linguistique et le droit. La régression est

l'une des méthodes les plus connues et les plus appliquées en statistique pour l'analyse de données quantitatives. Elle est utilisée pour établir une liaison entre une variable quantitative et une ou plusieurs autres variables quantitatives, sous la forme d'un modèle. Si on s'intéresse à la relation entre deux variables, on parlera de régression simple en exprimant une variable en fonction de l'autre. Si la relation porte entre une variable et plusieurs autres variables, on parlera de régression multiple. La régression linéaire simple et multiple font une classe particulière de modèles de régression. Le but est d'expliquer une variable dépendante, appelée variable endogène par une ou plusieurs variables explicatives, dites exogènes, à travers une fonction affine. Afin d'avoir une idée précise et claire de la manière de la mise en œuvre des outils statistiques, nous avons effectué un stage pratique à la direction exploitation de L'EPB ,au niveau du service statistique, où nous avons pu mettre en pratique nos connaissances théoriques acquises dans la matière. Durant notre stage pratique nous nous sommes intéressés aux données statistiques des deux terminaux : le terminal hydrocarbure et le terminal à conteneur, qui figurent parmi les principaux terminaux du port, pour voir l'impact de la pandémie Covid- 19 sur ces deux potentiels, et cela sur la base de modélisations et des prévisions, en faisant appel aux deux méthodes ; la méthode de box et Jenkins et la méthode de la régression linéaire.

2.2 Position du problème

Le nombre des navires qui ont accosté au port de Béjaïa a atteint 1.573 navires contre 1.974 navires en 2019, cette baisse est induite, explique l'entreprise publique, "par la diminution du mouvement de la navigation durant cette crise sanitaire dont nous subissons les effets depuis mars 2020, notamment après la fermeture des frontières, la baisse de la production mondiale et des échanges commerciaux internationaux".

Le terminal à conteneur, comme le terminal hydrocarbure, sont les terminaux les plus actifs dans le mouvement de la navigation de port de Béjaïa, les quantités de tonnage qui s'engendrent par le trafic des hydrocarbures connaissent une variation très élevée avant la crise sanitaire, bien précis, avant mars 2020, mais depuis, ces quantités ont connues un contraste remarquable, ce qui a causé des perturbations sur la gestion de cette marchandise à l'heure actuelle, et des difficultés de planification future pour le terminal hydrocarbure, en raison du manque d'informations présentes, et aussi l'incertitude de l'information future.

Pour opter à une planification adéquate, la question qui se pose est de savoir ; sur la base des prévisions, quelles sont les quantités de tonnage qui seront engendrées par le trafic des hydrocarbures dans cette crise sanitaire pour l'année 2021 ?, et quel est l'impact de la crise sanitaire sur le trafic à conteneur et le trafic des hydrocarbures durant l'année 2020 ?

2.3 Modélisation et résolution

Cette section concerne en premier temps l'étude statistique descriptive des données recueillies au niveau de EPB. Ensuite nous réalisons une étude prévisionnelle par la méthode

de Box et Jenkins sur la quantité en tonne engendrée par le trafic des hydrocarbures .

La deuxième partie de cette section est consacrée à l'application numérique d'une régression linéaire multiple, avec les données qu'on a pu récupérer au sein de service statistique de l'EPB qui concerne le nombre des navires traités (les navires qui transportent les conteneurs) durant l'année 2019.

Le logiciel utilisé pour la réalisation de nos applications dans ce chapitre est le logiciel R.

2.3.1 Description des données

Lors de notre stage effectué au sein de l'entreprise EPB, nous avons recueilli les données concernant les quantités en tonne engendrées par le trafic des hydrocarbures de 2014 jusqu'à 2019 et le nombre de navires traités durant l'année 2019. Les données collectées sont représentées dans les tableaux ci-dessous.

Collecte de données

Les données dont nous disposons proviennent des annuaires statistiques établis par le service d'exploitation de la Direction de l'EPB, qui regroupe les données des tonnages engendré par le trafic des hydrocarbures de 2014 jusqu'à 2019. Après la collecte de ces données, nous avons regroupé ces données dans un tableau de 12 mois sur 6 ans. Dans le but de prévoir le tonnage qui sera engendré pour l'année 2020, nous appliquons la méthode de Box et Jenkins.

Mois \ Année	2019	2018	2017	2016	2015	2014
jan	1.2	1.6	1.9	1.9	2	2.8
Fév	2.5	2.5	2.6	2	1.7	2.5
Mar	1.3	1.3	1.3	1.5	2.8	1.8
Avr	2	1.8	1.5	1.3	1	1.115
Mai	1	1.5	1.3	1.4	1.3	1.4
Jun	1.2	1.6	1.2	1.16	1.21	1.2
Jui	1.3	1.5	1.2	1.2	1.3	1.3
Aou	1.24	1.32	1.5	1.3	1.5	1.5
Sept	1.5	1.2	1.6	1.7	1.2	1.7
Oct	2	1.3	1.2	1.5	1.7	1.9
Nov	1.9	2.5	1.5	2.5	1.8	1.9
Déc	2	2	2	2	2.5	1

TABLE 2.1: Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

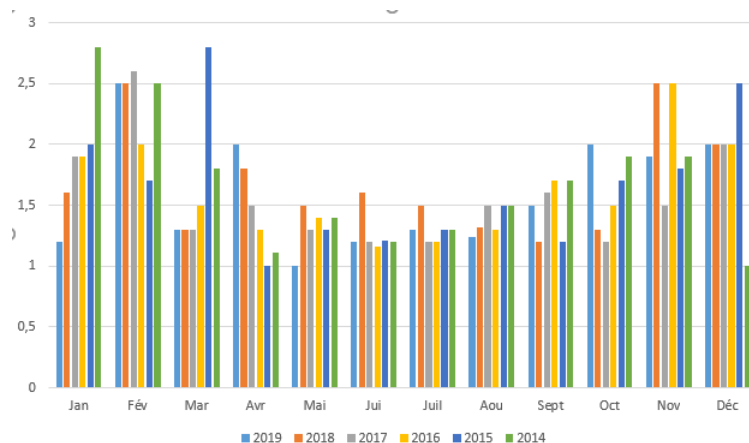


FIGURE 2.1: Tonnage engendré par le trafic des hydrocarbures de 2014 à 2019

D'après la figure (1), les valeurs de tonnage mensuels de 2014 jusqu'à 2019, c'est-à-dire avant la crise de Covid- 19, se varient dans un intervalle de 1 million de tonnes jusqu'à 2 million de tonnes en moyenne par mois, on remarque qu'elle atteint la valeur 2.8 million de tonnes en mars 2019, dans l'ensemble, si on varie la quantité totale qui s'engendre par année on trouve 18 million de tonnes en moyenne chaque ans.

Mois,Année	2020	2019	2018	2017	2016	2015
Jan	2	1.2	1.6	1.9	1.9	2
Févr	1.7	2.5	2.5	2.6	2	1.7
Mar	1.2	1.3	1.3	1.3	1.5	2.8
Avr	1.15	2	1.8	1.5	1.3	1
Mai	1.2	1	1.5	1.3	1.4	1.3
Jui	1.1	1.2	1.6	1.2	1.16	1.214
Juil	1.2	1.3	1.5	1.2	1.2	1.3
Aou	1.2	1.244	1.32	1.5	1.3	1.5
Sept	1.1	1.5	1.2	1.6	1.7	1.2
Oct	1.4	2	1.3	1.2	1.5	1.7
Nov	2	1.9	2.5	1.5	2.5	1.8
Déc	1	2	2	2	2	2.5

TABLE 2.2: Les quantités de tonnage engendrées par le trafic des hydrocarbures (2015-2020)

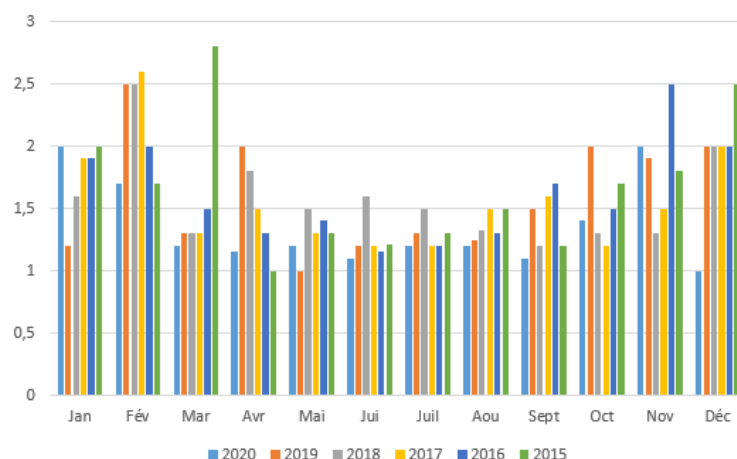


FIGURE 2.2: Tonnage engendré par le trafic des hydrocarbures (2015-2020)

D'après la figure (2), on constate que les valeurs de tonnage mensuels de 2015 jusqu'à 2019 se variaient dans un intervalle de 1 million de tonnes jusqu'à 2 millions de tonnes en moyenne par mois, bien avant la crise sanitaire, en revanche, pour l'année 2020, on remarque une baisse dans les quantités de tonnage. Cependant, on a enregistré une quantité de 15.7 millions de tonnes pour l'année 2020.

Temps(mois)	Importation(EVP)	Exportation(EVP)	Nombrenavires	Livraison
Jan	7080	950	25	6225
Fév	6200	1500	18	6030
Mar	7400	800	25	6100
Avr	6500	800	20	6500
Mai	7200	500	22	6420
Jui	5600	600	23	6325
Juil	6500	1000	26	4090
Aou	7100	750	27	6380
Sep	5300	650	21	6020
Oct	6500	756	19	5400
Nov	6060	850	25	5990
Déc	6550	990	22	7080

TABLE 2.3: Les nombre de navires (porte-conteneurs) traités durant l'année 2019

2.3.2 Prévision par de la méthode de Box-Jenkins(2020)

Dans cet exemple, on s'intéresse à effectuer une prévision sur le tonnage qui sera engendré par le trafic des hydrocarbures pour l'années 2020 à partir des données sur la période 2014-2019 (tableau 1). Pour cela on applique la méthode de Box-Jenkins dont le principe est d'étudier systématiquement une série chronologique à partir de leurs caractéristiques. L'objectif est de chercher dans la famille ARIMA et SARIMA le modèle le plus adapté pour représenter le phénomène étudié.

2.3.3 Présentation graphique des données

Afin de tracer le graphe de la série originale, il faut introduire les données sous forme d'un vecteur dans le logiciel R puis faire appel aux commandes suivantes qui permettent de le rendre sous forme d'une série chronologique.

```
> y = c(1.2, 2.5, 1.3, 2, 1, 1.2, 1.3, 1.244, 1.5, 2, 1.9, 2, 1.6, 2.5, 1.3, 1.8, 1.5, 1.6, 1.5, 1.32, 1.2, 1.3, 2.5, 2, 1.9, 2, 1.5, 1.3, 1.4, 1.116, 1.2, 1.3, 1.7, 1.5, 2.5, 2, 2, 1.7, 2.8, 1, 1.3, 1.214, 1.3, 1.5, 1.2, 1.7, 1.8, 2.5, 2.8, 2.5, 1.8, 1.115, 1.4, 1.2, 1.3, 1.5, 1.7, 1.9, 1.9, 1)
```

```
> serie=ts(y,start=c(2014,1),end=c(2019,12),frequency=12)
```

La commande qui nous permet de visualiser le graphe de la série originale est :

```
> plot(serie)
```

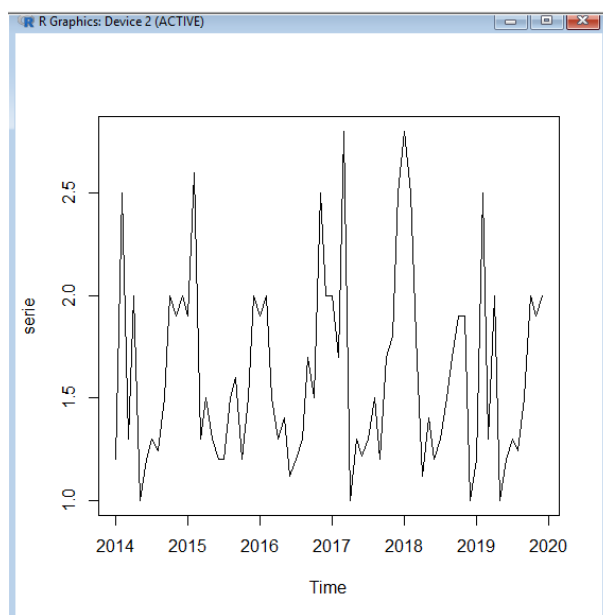


FIGURE 2.3: Le graphe de la série originale

Le graphe de la série originale figure(3) montre que celle-ci comporte une tendance et une composante saisonnière, pour illustrer cela nous allons décomposer la série originale en ses différentes composantes .Nous utilisons les commandes :

```
> domp=decompose(serie,type=c("additive","multiplicative"),filter=NULL)
```

```
> plot(domp)
```

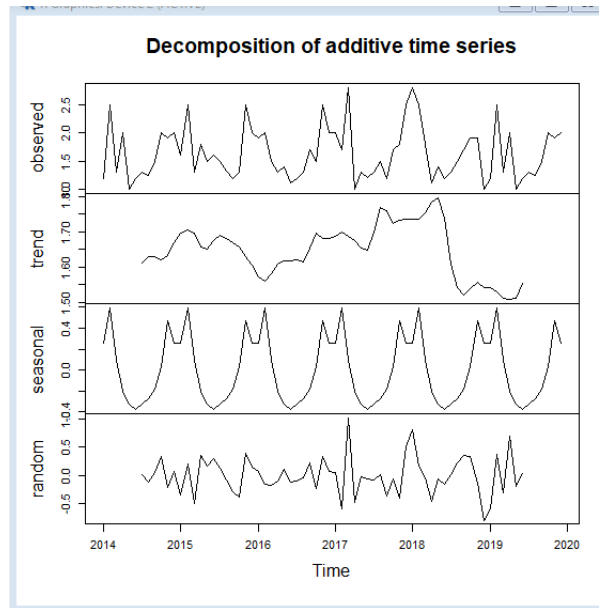


FIGURE 2.4: Graphe de la série et ses différentes composantes.

2.3.4 Analyse préliminaire

Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série originale sont :

- > acf(serie)
- > pacf(serie)

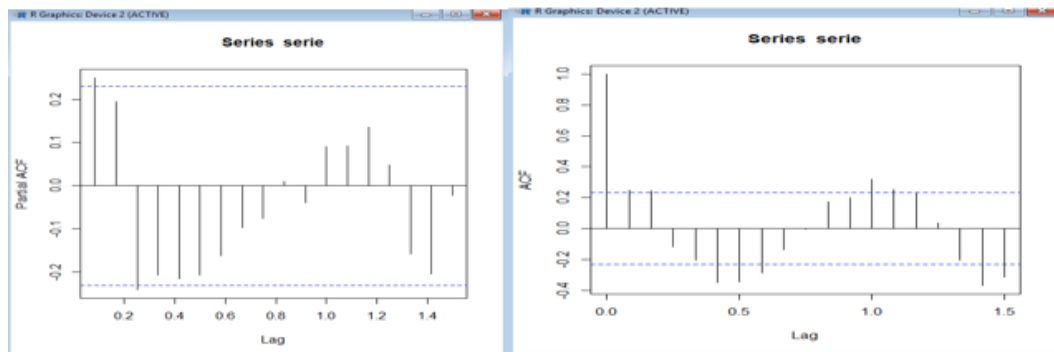


FIGURE 2.5: Corrélogramme et corrélogramme partiel de la série originale

2.3.5 Stationnarité

L'examen des deux graphes ACF et PACF nous suggèrent l'existence d'une tendance donc il sera nécessaire de la différencier, ainsi qu'une composante saisonnière, donc une différenciation d'ordre 12 sera nécessaire pour enlever cette saisonnalité.

L'instruction suivante nous permet d'affaiblir la variance de la série. On utilise la transformation en logarithme pour rendre la série stationnaire, en appliquant une différence première ordinaire et une différence saisonnière d'ordre 12. Les commandes qui nous permettent d'obtenir la série

chronologique differencier sont les suivantes :

```
> dift=diff(serie)
> difs=diff(dift,frequency=12)
> plot(difs)
```

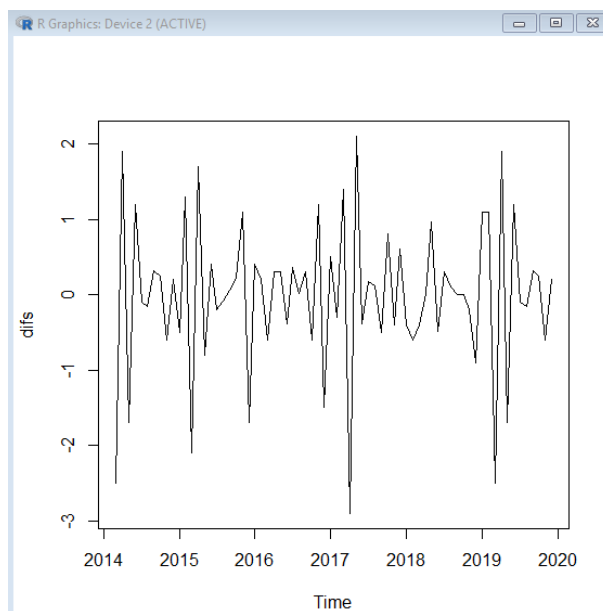


FIGURE 2.6: Graphe de la série chronologique differencier

2.3.6 Identification du modèle

D'après le graphe de la série originale, on constate que cette série est sujet a une saisonalite d'ordre 12 (12 mois) et une tendance. Donc le modèle adéquat est $SARIMA(p, d, q)(P, D, Q)_{(s=12)}$. En appliquant à la série originale un filtre ordinaire d'ordre 1 ($d = 1$), et un filtre saisonnier d'ordre 1 ($D = 1$), on a obtenu une série sans tendance ni saisonnalité.

Après avoir tester différents modèles, nous avons opté pour le modèle $SARIMA(1, 1, 0)(0, 1, 1)_{12}$. Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série transformée sont :

```
> acf(difs)
> pacf(difs)
```

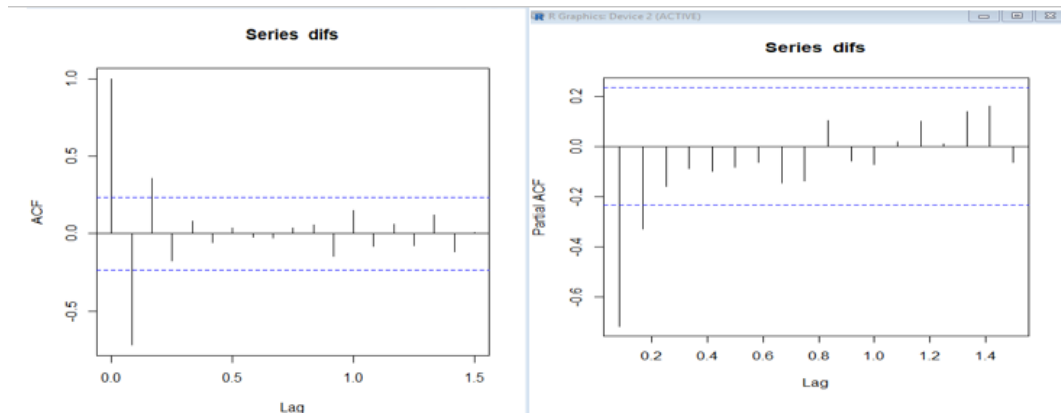



FIGURE 2.7: Corrélogramme et corrélogramme partiel de la série différencier

2.3.7 Estimation des paramètres du modèle

Après avoir opté pour un modèle SARIMA(1, 1, 0)(0, 1, 1)₁₂ la commande :
`> sarima=arima (s, order = c(1, 1, 0), seasonal = list (order = c(0, 1, 1), period = 12)),`
 permet d'effectuer l'ajustement de la série et d'obtenir les résultats suivants : Les estimateurs des paramètres du modèle sont :

$$\hat{\phi}_1 = -0.6213.$$

$$\hat{\theta}_s = -0.5808.$$

Les écarts type des paramètres estimés sont :

$$\hat{\sigma}(\hat{\phi}_1) = 0.1273.$$

$$\hat{\sigma}(\hat{\theta}_s) = 0.2269.$$

Nous remarquons que tous les estimateurs vérifient les conditions de stationnarité : $|\hat{\phi}| < 1$, $|\hat{\theta}_s| < 1$ et d'inversibilité.

Le modèle s'écrit :

$$(1 - \hat{\phi}_1 B)(1 - B)(1 - B^{12})y_t = (1 - \hat{\theta}_s B^{12})\epsilon_t$$

Alors

$$(1 + 0.6213B)(1 + B^{12})y_t = (1 + 0.5808B^{12})\epsilon_t.$$

2.3.8 Validation du modèle

Tests sur les coefficients du modèle

La valeur critique lue sur la table de student est $T(72, 0.025) = 2,39$

1. Test de $H_0 : \phi_1 = 0^n$ contre $H_1 : \phi_1 \neq 0^n$ Ce test est basé sur la statistique $T_s = \frac{|\hat{\phi}_1|}{\hat{\sigma}_{\phi_1}} \longrightarrow t(n-2, \alpha/2)$

sa réalisation $t_1 = \frac{0.6213}{0.1273} = 4.88 > 2.39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre ϕ_1 est significativement différent de zéro.

2. $H_0 : \theta_s = 0$ contre $H_1 : \theta_s \neq 0$ Ce test est basé sur la statistique :

$$T_s = \frac{|\hat{\theta}_s|}{\hat{\sigma}_{\theta_s}} \longrightarrow t(n-2, \alpha/2)$$

sa réalisation $t_1 = \frac{0.5808}{0.2269} = 2.63 > 2.39$, Alors on rejette l'hypothèse H_0 et par conséquent on accepte l'hypothèse H_1 donc le paramètre θ_s est significativement différent de zéro

Test sur les résidus

Ce test est basé sur la statistique suivante :

$$Q = n(n+2) \sum_{i=1}^n \frac{p_i^2}{n-i} \longrightarrow \chi_{n-k}^2$$

Test : $H_0 : \rho_1 = \rho_2 = \dots = 0, i = 1, \dots, n$ contre $H_1 : \exists i \in 1 \dots, n; \rho_i \neq 0^n$,

où ρ_i désignent les coefficients d'autocorrélations théoriques,

M est la partie entière de : $\text{Min}(n/2, 3\sqrt{n})$ donc $n = 25$.

L'instruction :

> `Box.test(residu1,lag=23,type="Ljung")` nous donne la valeur empirique de la statistique de Box et Ljung, $Q^* = 13,6550$. Au niveau de signification $\alpha = 0.05$ la valeur critique lue sur la table de khi-deux est :

$$\chi_{M-k}^2(\alpha) = \chi_{25-2-1}^2(0,05) = \chi_{22}^2(0,05) = 26,245$$

Comme $Q^* = 13,6550 < \chi_{22}^2(0,05) = 26,245$, alors on accepte l'hypothèse H_0 , par conséquent les résidus forment un bruit blanc, le corrélogramme et le corrélogramme partiel des résidus nous montrent clairement que tous les pics sont à l'intérieure de l'enveloppe. Les commandes sous R qui nous permettent de les représenter sont :

```
> residu=residuals(sarima)
> residu1=ts(residu,start=2010,frequency=12)
> acf(residu1)
> pacf(residu1)
```

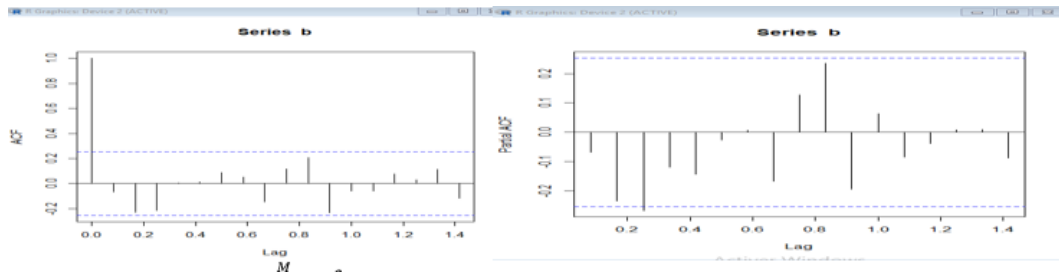


FIGURE 2.8: Corrélogramme et corrélogramme partiel des résidus.

2.3.9 Prédiction

Les commandes qui nous permettent d'obtenir les prévisions pour l'année 2020 sont :

```
> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
```

```
|
> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
Time Series:
Start = 61
End = 72
Frequency = 1
[1] 1.479052 1.357856 1.677086 1.565544 1.667962 1.573922 1.660269 1.580985 1.653783 1.586940 1.648316 1.591961
```

Le tableau ci-dessus nous montre les résultats de la prédiction de tonnage engendré par le trafic des hydrocarbures (unité : million tonnes) obtenus avec le logiciel R est les valeurs réelles de l'année 2020 :

mois (année 2020) \ année	Valeurs prévues	Valeurs réelles
jan	1.479052	2
Fév	1.357856	1.7
Mar	1.677086	1.2
Avr	1.565544	1.15
Mai	1.667962	1.2
Jun	1.573922	1.1
Jui	1.660269	1.2
Aou	1.653783	1.2
Sept	1.580985	1.1
Oct	1.653783	1.4
Nov	1.648316	2
Déc	1.591961	1

TABLE 2.4: Les résultats de la prévision de tonnage engendré par le trafic des hydrocarbures pour 2020

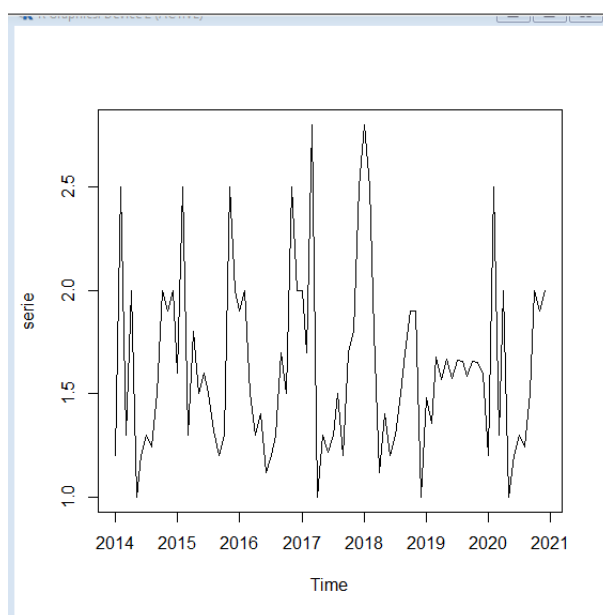


FIGURE 2.9: Graphe de la série et la prévision

2.3.10 Interprétation des résultats

D'après le tableau, les prévisions données par Box-Jenkins concernant les quantités de tonnage, montrent qu'il y a un grand écart avec les valeurs réelles, en revanche, si on fait la somme des quantités prévues pour l'année 2020 on obtient une quantité de 19.11 millions de tonnes, sachant que qu'à cette année le port a enregistré au total 15.7 millions de tonnes, cela s'explique qu'il y a plusieurs facteurs influents, parmi ces facteurs la crise de Corona virus qui a causé un grand recul dans le secteur de l'import et l'export.

cette prévision effectué nous éclaircit sur l'effet de cette crise sur trafic des hydrocarbures, on remarque qu'une perte de 4 million de tonnes induite de ce recul, sinon une quantité de 19.11 million de tonnes sera recueillie si les opérations import-expot n'ont été pas touchées par la crise sanitaire.

2.4 Prédiction par la méthode de Box-Jenkins (2021)

Dans cet exemple, nous allons prédire les quantités de tonnage engendrées par le trafic des hydrocarbures pour 2021. Nous utilisons les données réelles de 2020.

2.4.1 Présentation graphique des données

Afin de tracer le graphe de la série originale, il faut introduire les données sous forme d'un vecteur dans le logiciel R puis appliquer les commandes suivantes qui permettent de le rendre sous forme d'une série chronologique.

```
> m = c(2, 1.7, 1.2, 1.15, 1.2, 1.1, 1.2, 1.2, 1.1, 1.4, 2, 1, 1.9, 2.6, 1.3, 1.5, 1.3, 1.2, 1.2, 1.5, 1.6, 1.2, 1.5, 2, 1.9, 2, 1.3, 1.4, 1.116, 1.2, 1.3, 1.7, 1.5, 2.5, 2, 2, 1.7, 2.8, 1, 1.3, 1.214, 1.3, 1.5, 1.2, 1.7, 1.8, 2.5, 2.8, 2.5, 1.8, 1.115, 1, 1.2, 1.3, 1.5, 1.7, 1.9, 1.9, 1)
```

```
> serie = ts(m, start = c(2015, 1), end = c(2020, 12), frequency = 12)
```

La commande qui nous permet de visualiser le graphe de la série originale est :

```
> plot(serie)
```

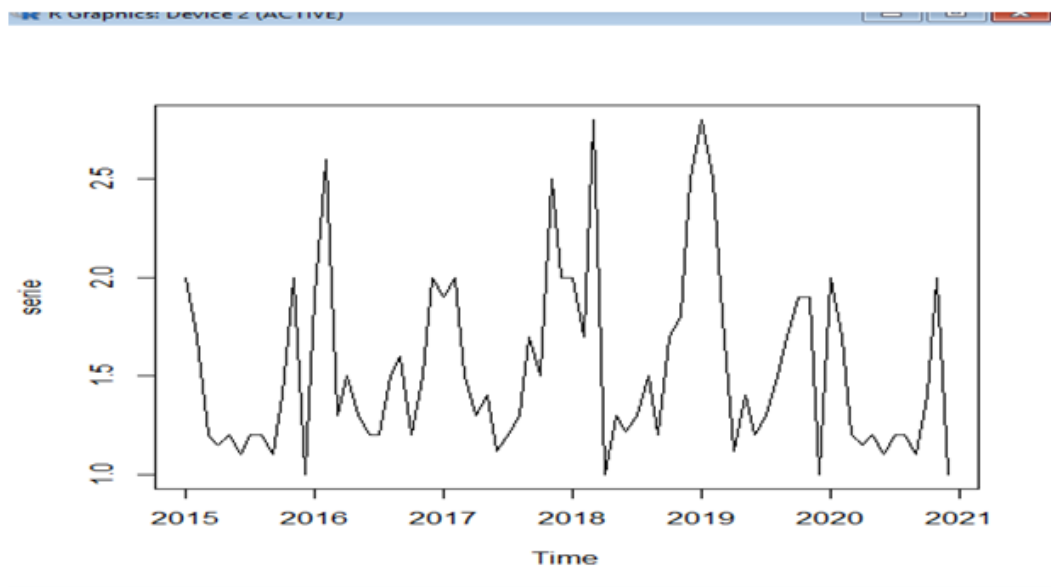


FIGURE 2.10: Graphe de la série originale

Les commandes qui nous permettent de visualiser respectivement le corrélogramme et le corrélogramme partiel de la série originale sont :

```
> acf(serie)
```

```
> pacf(serie)
```

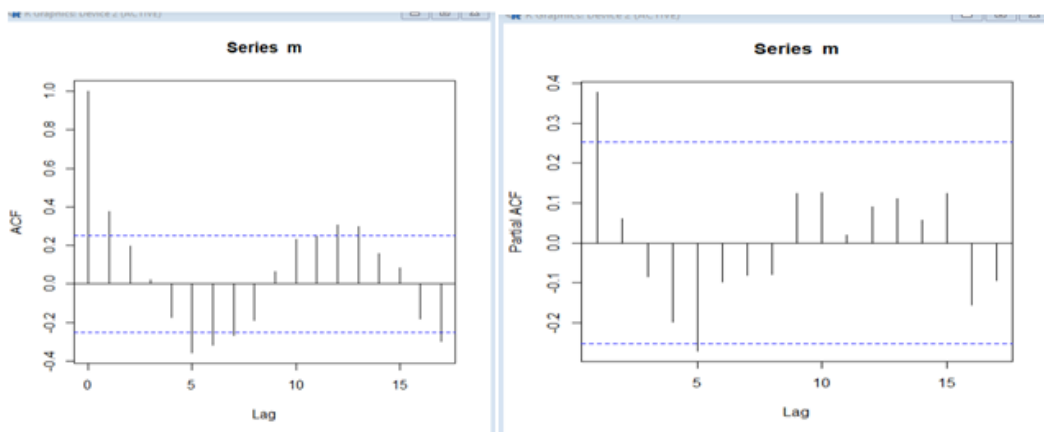


FIGURE 2.11: Corrélogramme et corrélogramme partiel de la série originale

2.4.2 Stationnarité

D'après les deux graphes de L'ACF et le PACF, la série que nous disposons est une série non stationnaire. Pour la rendre stationnaire, on doit éliminer sa saisonnalité et sa tendance, pour se faire nous utilisons la commande.

```
> adf.test(m,alternative="stationary")
```

```
> library("tseries")
Registered S3 method overwritten by 'quantmod':
  method      from
as.zoo.data.frame zoo

'tseries' version: 0.10-48

'tseries' is a package for time series analysis and computational
finance.

See 'library(help="tseries")' for details.

> adf.test(m,alternative="stationary")

Augmented Dickey-Fuller Test

data: m
Dickey-Fuller = -4.1215, Lag order = 3, p-value = 0.01038
alternative hypothesis: stationary
```

on remarque que la propriété de la stationnarité est vérifiée,

2.4.3 Identification du modèle

Nous avons une série chronologique stationnaire, les modèles qui représente ce type de série sont AR(p), MA(q) et ARMA(p,q), après avoir testé différents modèles sous R, nous avons opté pour le modèle ARIMA(1 ; 0 ; 1) qui s'écrit aussi ARMA(1,1). Les commandes qui nous permettent d'identifier le modèle sont :

```
> mydata.arima111=arima(y,order=c(1,0,1))
> mydata.arima111
```

```
> arima(m,order=c(1,0,0))

Call:
arima(x = m, order = c(1, 0, 0))

Coefficients:
      ar1  intercept
      0.3874      1.575
s.e.  0.1202      0.088

sigma^2 estimated as 0.178:  log likelihood = -33.44,  aic = 72.87
> arima(m,order=c(1,0,1))

Call:
arima(x = m, order = c(1, 0, 1))

Coefficients:
      ar1      ma1  intercept
      0.4543 -0.0790      1.5753
s.e.  0.2330      0.2492      0.0907

sigma^2 estimated as 0.1777:  log likelihood = -33.39,  aic = 74.78
> arima(m,order=c(1,0,2))

Call:
arima(x = m, order = c(1, 0, 2))

Coefficients:
      ar1      ma1      ma2  intercept
      0.3542  0.0017  0.1129      1.5757
s.e.  0.3214  0.3282  0.1581      0.0924

sigma^2 estimated as 0.1763:  log likelihood = -33.16,  aic = 76.33
> arima(m,order=c(1,0,3))

Call:
arima(x = m, order = c(1, 0, 3))
```

Cette technique est basée sur le critère d'AIC (Akaike Information Criterion),qui est calculée comme suit :

$$AIC = \ln(SCR/n) + (2p/n)$$

Tel que :

Ln : logarithme népérien.

SCR : la somme des carrés des résidus du modèle.

n : nombre d'observation.

P : le nombre de variable explicatives.

Premièrement on calcule l'AIC pour le modèle global et aussi pour chaque variable, puis nous éliminons les variables dont leurs AIC est petit, jusqu'à ce qu'on obtient le modèle approprié.

2.4.4 Prévision

Les commandes qui nous permettent d'obtenir les prévisions pour l'année 2021 sont :

```
> mydata.pred1 = predict(mydata.arima111,n.ahead = 12)
> mydata.pred1pred
```

```
> mydata.pred1=predict(mydata.arima111,n.ahead=12)
> mydata.pred1$pred
Time Series:
Start = 61
End = 72
Frequency = 1
[1] 1.369764 1.481894 1.532839 1.555986 1.566503 1.571281 1.573452 1.574438
[9] 1.574886 1.575090 1.575182 1.575225
```

Le tableau ci-dessous représente les résultats de la prévision de tonnage des hydrocarbures pour l'année 2021 :

Mois(année2021)	Valeurs prévues
Jan	1.369764
Fév	1.481894
Mar	1.532839
Avr	1.555986
Mai	1.566503
Jui	1.571281
Juil	1.573452
Aou	1.574438
Sep	1.574886
Oct	1.575090
Nov	1.575182
Déc	1.575225

TABLE 2.5: Les résultats de la prévision de tonnage des hydrocarbures pour l'année 2021

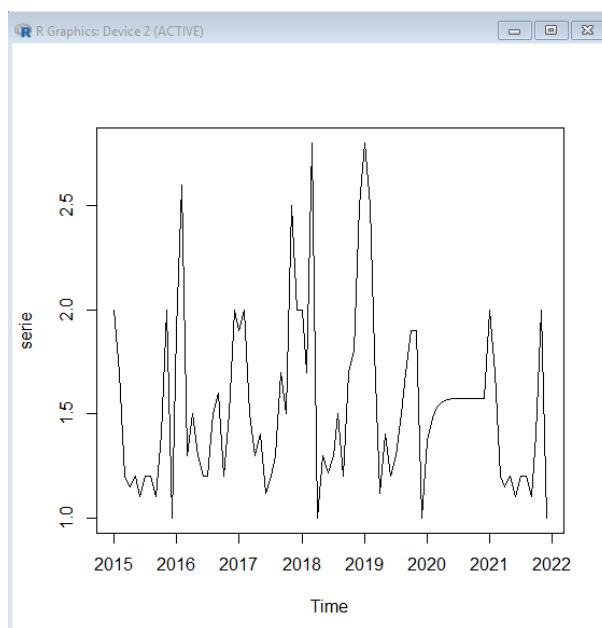


FIGURE 2.12: Graphe de la série et la prévision

2.4.5 Interprétation des résultats

D'après le tableau, les prévisions données par Box-Jenkins concernant les quantités de tonnage, nous donne une vision approximative pour l'année 2021.

Cependant, selon les résultats obtenus, le port de Béjaia enregistrera une quantité de 18.5 million de tonnes qui sera engendrée par le trafic des hydrocarbures en 2021, cela assure une hausse considérable de cette dernière par rapport à la quantité enregistrée en 2020.

Le premier exemples que nous avons traité nous fournis des informations importantes sur les quantités de tonnage des hydrocarbures pour les deux années 2020, cependant, les résultats obtenus nous confirme l'exactitude de la prévision en comparant les valeurs obtenues avec les valeurs réelles, cela nous confirme l'idée de la réalisation d'une prévision pour l'année 2021 dans le deuxième exemple.

L'étude réalisée dans le deuxième exemple nous donne une vision approximative sur les quantités de tonnages qui seront engendrées par le trafic des hydrocarbures en 2021, les résultats obtenus de cette prévision assurent une hausse des tonnages, ce qui permet aux responsables du port de Béjaia d'opter à une planification futur dans les secteurs concernés.

2.5 Prévision par la méthode de régression linéaire

Cette partie est consacrée à l'application numérique d'une régression linéaire multiple, avec les données qu'on a pu récupérer au sein de service statistique de l'EPB qui concerne le nombre des navires traités (les navires qui transportent les conteneurs) durant l'année 2019 (tableau 4.3). Nous avons essayé d'expliquer le nombre de navires traités en fonction de certain nombre de facteurs (Exportation, importation, temps,livraison) ensuite réaliser une prévision pour l'année 2020, tout en appliquant une régression linéaire multiple.

2.5.1 Modèle de régression

Pour formuler le modèle adéquat pour les choix des variables à expliquer, nous utilisons la méthode d'élimination (Back-Ward), cette méthode permet de sélectionner parmi les variables suggérées précédemment, celles les plus appropriées à prendre en considération pour notre modèle. Nous commençons par la création du tableau d'observation dans R..

```

> Importation=c(7080,6200,7400,6500,7200,5600,6500,7100,5300,650,6060,6550)
> Exportation=c(950,1500,800,800,500,600,1000,750,650,756,850,990)
> NbrNavire=c(25,18,25,20,22,23,26,27,21,19,25,22)
> Livraison=c(6025,6030,6100,6500,6420,6325,4090,6980,6020,5400,5990,7090)
> Temps=c(1:12)
> D=data.frame(Importation,Exportation,Livraison,NbrNavire,Temp)
> D
  Importation Exportation Livraison NbrNavire Temps
1         7080          950       6025         25     1
2         6200         1500       6030         18     2
3         7400          800       6100         25     3
4         6500          800       6500         20     4
5         7200          500       6420         22     5
6         5600          600       6325         23     6
7         6500         1000       4090         26     7
8         7100          750       6980         27     8
9         5300          650       6020         21     9
10          650          756       5400         19    10
11         6060          850       5990         25    11
12         6550          990       7090         22    12
~ |

```

Maintenant sous R nous allons faire appel la fonction `step()` qui sert à appliquer la méthode d'élimination comme suit :

```

> step(lm(NbrNavire~Importation+Exportation+Livraison+Temp,data=D),direction="backward")
Start: AIC=25.13
NbrNavire ~ Importation + Exportation + Livraison + Temps

      Df Sum of Sq  RSS   AIC
- Temps      1    7.359 49.689 25.050
<none>                 42.330 25.127
- Exportation 1     8.205 50.535 25.253
- Livraison   1    12.944 55.274 26.329
- Importation 1    42.826 85.156 31.515

Step: AIC=25.05
NbrNavire ~ Importation + Exportation + Livraison

      Df Sum of Sq  RSS   AIC
<none>                 49.689 25.050
- Livraison   1    10.188 59.877 25.289
- Exportation 1    12.191 61.880 25.683
- Importation 1    35.488 85.177 29.518

Call:
lm(formula = NbrNavire ~ Importation + Exportation + Livraison,
    data = D)

Coefficients:
(Intercept)  Importation  Exportation  Livraison
 28.107265     0.001059    -0.004241    -0.001339

```

Cette technique est basée sur le critère d'AIC (Akaike Information Criterion), utilisée dans l'exemple précédent. Premièrement on calcule l'AIC pour le modèle global et aussi pour chaque variable, puis nous éliminons les variables dont leurs AIC est petit, jusqu'à ce qu'on obtient le modèle approprié.

Dans notre cas l'AIC de modèle global = 25 :13, on compare l'AIC pour chaque variable, et on sélectionne l'AIC le plus petit. On remarque que l'AIC le plus petit est AIC = 25 :050, qui correspond à l'AIC de la variable temps, on le compare avec l'AIC global on trouve que $AIC(\text{temps}) < AIC \text{ global}$, donc on supprime la variable temps. D'où, les variables retenues sont : Importation, Exportation, Livraison.

Le modèles s'écrit :

$$NbrVessel = \beta_0 + \beta_1 Importation + \beta_2 Exportation + \beta_3 Livraison + \epsilon$$

D'après la fonction step() function, on obtient les estimateurs :

$$\hat{\beta}_0 = 28,1, \quad \hat{\beta}_1 = 0,001, \quad \hat{\beta}_2 = -0,004, \quad \hat{\beta}_3 = -0,001$$

alors

$$Nbr\hat{V}essel = 28,1 + (0,001)Import + (-0,004)Export + (-0,001)Livraison.$$

2.5.2 Tests statistiques de Student et Fisher

Pour effectuer les tests de Student et Fisher, nous utilisons la fonction summary() dans R.

```
> summary(lm(NbrNavire~Importation+Exportation+Livraison+Temps,data=D))

Call:
lm(formula = NbrNavire ~ Importation + Exportation + Livraison +
    Temps, data = D)

Residuals:
    Min       1Q   Median       3Q      Max
-2.6024 -1.7405  0.2027  0.9863  3.4892

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 25.5968373  7.2411876   3.535  0.00953 **
Importation  0.0012916  0.0004854   2.661  0.03241 *
Exportation -0.0035536  0.0030509  -1.165  0.28225
Livraison   -0.0015299  0.0010457  -1.463  0.18686
Temps       0.2609059  0.2365165   1.103  0.30645
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.459 on 7 degrees of freedom
Multiple R-squared:  0.5411,    Adjusted R-squared:  0.2789
F-statistic: 2.064 on 4 and 7 DF,  p-value: 0.1894
```

Le coefficient de détermination $R^2 = 0,54$, nous avons donc un bon ajustement linéaire.

2.5.3 Test global de Fisher

Nous testons les hypothèses suivantes :

$$H_0 : \beta_0 = 0; \hat{\epsilon}a_1 = 0; \hat{\epsilon}a_2 = 0; \beta_3 = 0;$$

contre

$$H_1 : \beta_0 \neq 0; \beta_1 \neq 0; \beta_2 \neq 0; \beta_3 \neq 0.$$

Ce test est basé sur la statistique de Fisher (statistique F) F^* :

Nous avons $F^* = 3,535$, au niveau de signification de $\alpha = 0,05$ la valeur sur la table de Fisher est $F(4, 9, 0,05) = 2,60$, comme $F^* > F(4, 9, 0,05)$ nous rejetons H_0 , donc le modèle est globalement significatif.

2.5.4 Test étudiant individuel

Au seuil de signification $\alpha = 0,05$ nous avons :

Pour $\hat{\beta}_0$

$$H_0 : \hat{\beta}_0 = 0;$$

Contre

$$H_1 : \hat{\beta}_0 \neq 0$$

$$t(0.05; 9) = 3.535 > T_{\hat{\beta}_0} = 0.00953.$$

Conclusion

Nous acceptons l'hypothèse H_1 .

Pour $\hat{\beta}_1$

$$H_0 : \hat{\beta}_1 = 0;$$

Contre

$$H_1 : \hat{\beta}_1 \neq 0.$$

$$t(0.05; 9) = 2.661 > T_{\hat{\beta}_1} = 0.03241$$

Conclusion

Nous acceptons l'hypothèse H_1 .

Pour $\hat{\beta}_2$

$$H_0 : \hat{\beta}_2 = 0;$$

Contre

$$H_1 : \hat{\beta}_2 \neq 0.$$

$$t(0.05; 9) = -1.165 > T_{\hat{\beta}_2} = 0.28225.$$

Conclusion

Nous rejetons l'hypothèse H_0 .

Pour $\hat{\beta}_3$

$$H_0 : \hat{\beta}_3 = 0;$$

Contre

$$H_1 : \hat{\beta}_3 \neq 0.$$

$$t(0.05; 9) = -1.463 > T_{\hat{\beta}_3} = 0.18686.$$

Conclusion

Nous acceptons l'hypothèse H_1 .

2.5.5 Intervalle de confiance

Nous calculons l'intervalle de confiance pour les quatre coefficients avec la fonction confint() sous R. On obtient :

L'intervalle de confiance pour $\hat{\beta}_0$: $IC_0 = [1.204171e + 01; 44.172817408]$;

L'intervalle de confiance pour $\hat{\beta}_1$: $IC_1 = [3, 736097 \exp(-5); 0.002080983]$;

L'intervalle de confiance pour $\hat{\beta}_2$: $IC_2 = [-1.122047 \exp(-2); 0.002739299]$;

L'intervalle de confiance pour $\hat{\beta}_3$: $IC_3 = [-3, 748486 \exp(-3); 0.001071472]$.

Nous avons : $\hat{\beta}_0 \in IC_0$, $\hat{\beta}_1 \in IC_1$, $\hat{\beta}_2 \in IC_2$, $\hat{\beta}_3 \in IC_3$.

Ensuite, le modèle est validé.

2.5.6 Prévision

Afin de faire une prévision du nombre de navires qui seront traités pour l'année 2020, nous utilisons la fonction predict() sous R.

Ci-dessus, les résultats approximatifs des prévisions :

Le tableau suivant présente les résultats de la prévision et le nombre réel de navires traités en 2020 :

Mois	Nombre de navires prévus en 2020	Nombre de navires réel en 2020
Jan	23	22
Fév	20	22
Mar	24	25
Avr	23	23
Mai	25	27
Jui	23	20
Juil	25	19
Aou	23	29
Sep	23	25
Oct	18	21
Nov	23	24
Déc	21	23

TABLE 2.6: Nombre des navires réels et prévus pour l'année 2020

2.5.7 Interprétation des résultats

D'après les résultats obtenus, nous constatons que le nombres de navires prévus et réels sont approximativement analogues durant tous les mois, alors nous déduisons que dans l'ensemble, si nous comparons le nombre de navires traités dans les deux années, nous déduisons que le trafic des conteneurs au niveau du port de Béjaia face a la crise sanitaire est stable par rapport à l'année précédente.

2.5.8 Conclusion

Dans cette partie, nous avons procédé dans un premier temps à l'application de la méthode Box et Jenkins, étant donné que la première série n'est pas stationnaire (elle comporte une ten-

dance et une saisonnalité), il a été question d'une différenciation d'ordre 1 et saisonnière d'ordre 12. Après la stationnarisation de la série, et à partir des graphiques des fonctions d'autocorrélation et d'autocorrélation partielle, nous avons sélectionné le modèle SARIM A(1 ; 1 ; 0)(0 ; 1 ; 1)₁₂. La validation du modèle s'est faite grâce au test de student sur les coefficients et le test de Ljung-Box sur les résidus. Ainsi, nous avons pu obtenir des prévisions pour l'année 2020. Puis nous avons élaboré un modèle de série chronologique avec la même méthode pour effectuer une prévision pour l'année 2021. Dans un deuxième temps, nous avons modélisé le nombre de navires traités en 2019 avec la méthode de la régression linéaire multiple pour effectuer une prévision pour 2020.

2.6 Conclusion générale

À la lumière de notre travail, nous avons présenté d'une part, une revue statistique d'une série chronologique qui représente, dans notre cas, les quantités de tonnage engendrées par le trafic des hydrocarbures par mois sur six ans, d'une autre part, un modèle de régression linéaire multiple. Comme indiqué dans l'introduction, le but de cette étude sur cette série chronologique, et de ce modèle de régression linéaire est la prévision, qui consiste à utiliser le meilleur des modèles estimés pour donner les valeurs estimées à des instants futurs. En premier lieu, nous avons présenté les méthodes statistiques de la prévision telle que la méthode de Box et Jenkins et de la régression linéaire multiple. En deuxième lieu, grâce au logiciel R, nous avons appliqué la méthode de Box et Jenkins pour prévoir les quantités de tonnage de l'année 2020. Dans le but d'effectuer une comparaison avec les valeurs réelles et voir l'impact de la pandémie Covid-19 sur le trafic hydrocarbure, nous avons élaboré un modèle de série chronologique avec la même méthode pour effectuer une prévision pour l'année 2021. Ensuite, nous avons adopté la méthode de la régression linéaire multiple pour modéliser le nombre de navires traités en 2019 avec certains facteurs (livraison, importation, exportation, temps), et réaliser une prévision pour l'année 2020. Les résultats obtenus montrent que pour le trafic hydrocarbure, la quantité annuelle enregistrée en 2020 semble être inférieure à la quantité prévue, cela éclaire l'influence de la pandémie Covid-19 sur les opérations import-export de ce terminal. La prévision pour l'année 2021 permet de nous enrichir sur la vision future et envisager une planification assez adéquate pour les années à venir. L'étude réalisée sur le trafic des conteneurs montre que la crise sanitaire n'impacte pas sur les mesures de performance, au vu de la comparaison du nombre réel des navires traités en 2020 avec le nombre de navires obtenu par les prévisions.

2.7 Bibliographie

- [1] S. Adjabi. "Cours des méthodes statistiques de prévisions". Département de Recherche Opérationnelle, Université de Bejaia, 2013-2014.
- [4] F. Avram. "Séries temporelles : Régression et modélisation ARIMA (p, d, q)". Université de Pau, 2012.

- [5] A. Ahmed, Contribution à l'économétrie dans des séries temporelles à valeurs entières, Université Charle de Gaulls Lille 3 (2016).
- [6] G. Arnauad (2012/2013). Régression linéaire. Université Rennes 2 Master de statistique.
- [7] A. Antoniadis, J. Berruyer ,R. Carmona (1992). Régression non linéaire et applications, Economica.
- [8] J. Bergonzini et C. Duby (1995). Analyse et planification des expériences. Les dispositifs en blocs, ed. Masson, Paris, Milan, Barcelone, p 353.
- [10] G. Box and G.M. Jenkins. "Time series analysis : forecasting and control". 2nd edition, Holden- Day, San Francisco, 1976.
- [11] M. David et J.S. Michaud. "La prévision approche empirique d'une méthode statistique". Université de Lyon, Messon, 1983.
- [12] J. Daudin, J. c. duby ; s. robin et p. trécourt. "Analyse des série chronologiques".Université de Bordeaux, 1996.
- [13] R. Giraud et N. Chaix (1989). Econométrie. Presses Universitaires de France (PUF).
- [14] D. C. Hoaglin, and R. E. Welsch, (1978) : The hat matrix in regression and anova. The American Statistician.
- [16] V. Fourcassié et C. Jost, (2012). Introduction aux modeles linéaires généraux (General linear model - GLM). Cours Modules Statistiques Master 2 NCC.
- [17] G. Melard. "Méthodes de prévision à cours terme." Université de Bruxelles, 1990.
- [19] Y. Tillé, (2011). Résumé du Cours de Modèles de Régression. Institut de statistique, Université de Neuchâtel.

Modélisation et simulation du mouvement des malades au niveau de l'hôpital Khelil Amrane Béjaïa, Algérie.

Oudina S., Baalla M., Adel-Aissanou K. et Aïssani D.

Research Unit **LaMOS** (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria
lamos_bejaia@hotmail.com , ak_adel@yahoo.fr

Résumé L'objectif de ce travail est d'évaluer les performances actuelles du service des urgences au niveau de l'Hôpital Khelil Amrane (Béjaïa), puis d'évaluer ces mêmes performances après modification de quelques paramètres du système. Nous avons modélisé le mouvement des malades de service des urgences au niveau de l'Hôpital Khellil Amrane par un réseau de files d'attente. Pour évaluer les performances de notre système, nous avons procédé par deux étapes : l'ajustement des lois qui régissent le système avec le test de Kolmogorov-Smirnov et la simulation à l'aide du l'environnement Matlab.

Une analyse sensitive a permis d'évaluer les performances du système en cas de changement de quelques paramètres (affectation d'un autre médecin à la salle de consultation, variation du nombre de lits affectés au niveau des autre salles,...).

Mots clés : Hôpital Khelil Amrane (Béjaïa), Service des urgences, Modélisation Mathématique, Réseaux de file d'attente, Ajustement, Simulation.

3.1 Introduction

L'hôpital joue un rôle considérable dans la société, grâce à plusieurs facteurs : le progrès des sciences médicales, la concentration du personnel qualifié et des équipements spécialisés. Le système hospitalier doit être rendu accessible à toutes les classes sociales et couvrir l'étendue géographique où vit la population.

Les hôpitaux qui doivent, de par leur statut, répondre à des missions de service public, d'enseignement et de recherche sont confrontés à la nécessité d'optimiser leurs moyens matériels et humains en respectant leur fonction première : accueillir et soigner des malades en assurant la qualité et la sécurité des soins prodigués [1].

Les hôpitaux en Algérie n'ont pas encore acquit un mode de gestion efficace. La situation économique et sociale est liée aux problèmes rencontrés par les hôpitaux publics. La majorité des citoyens algériens continuent à se soigner dans les structures publiques, en raison de la détérioration de la situation sociale et à l'augmentation des prix des prestations de soins dans les structures privées. Ceci a engendré un flux très élevé de malades entraînant un certain nombre de problèmes, à savoir : temps d'attente très important dans les salles de soin (qui peut aggraver l'état de santé des malades), surcharge de quelques services d'hospitalisation et des personnels (c'est-à-dire que le nombre de lits et l'effectif du personnel ne sont pas suffisants pour couvrir toutes les demandes de soins).

Devant cette situation, les services d'accueil des urgences montrent de plus en plus un besoin

d'être assistés par des outils d'aide à la décision afin de mieux rentabiliser leur structure en assurant au patient un service rendu adapté et de qualité.

C'est dans ce contexte que nous avons réalisé ce travail qui consiste à mesurer les performances actuelles de l'Hôpital Khelil Amrane (Béjaia) et de voir l'impact d'un événement imprévu sur ces performances. Nous avons alors modélisé le mouvement des malades au niveau du service des urgences par un réseau de files d'attente. Pour évaluer les performances de notre modèle, nous avons eu besoin de connaître les différentes lois qui gouvernent le système. A cet effet, nous avons procédé par étapes :

- Collecte des données concernant les services des urgences ;
- Traitement de ces données ;
- Ajustement des différentes lois du système en se basant sur les données ;
- Simulation et évaluation des performances

3.2 La modélisation du mouvement des malades : état de l'art

Le système de santé est l'un des systèmes le plus important de la civilisation moderne. Dans les systèmes de santé, les services des urgences sont les plus complexes. Ils fonctionnent 24 heures sur 24, 7 jours sur 7. Ils font l'objet d'une demande croissante et sont généralement surchargés, leurs activités n'étant pas linéaire, elle varie continuellement. Les patients arrivent selon un processus stochastique aux urgences avec différents problèmes de santé et différents niveaux de gravité. Donc, certaines décisions vitales doivent être prises avec précision pour éviter de perdre du temps et aussi de fournir un mauvais traitement aux patients. Les managers hospitaliers ont besoin d'outils d'aide à la décision leur permettant de prendre en compte l'ensemble des acteurs de leurs services et les aléas qui les impactent en vue d'améliorer leurs fonctionnements et d'assurer une meilleure qualité de service aux patients. De nombreux travaux (génie logiciel, recherche opérationnelle...) ont été menés ces dernières années pour les aider dans cette démarche, en essayant, la plupart du temps, d'adapter les méthodes et outils ayant fait leurs preuves d'efficacité dans les systèmes industriels. Ces travaux montrent l'importance d'une bonne connaissance du système étudié et d'une bonne formalisation de cette connaissance, grâce notamment à la modélisation, pour pouvoir concevoir des outils d'aide à la décision adaptés. Ils montrent également les limites de l'adaptation de ces méthodes et outils aux systèmes hospitaliers. Dans ce sens, les auteurs de [4] abordent le problème de la minimisation du temps de cycle de parcours d'un patient au sein du service d'urgence de l'hôpital Habib Bourguiba Sfax (Tunisie) dont le but d'améliorer la performance de ce service. Ils ont commencé par la modélisation du processus de passage des patients. Celle-ci permet de mieux identifier les dysfonctionnements et problèmes rencontrés. Ensuite, ils ont proposé de passer à l'étape d'analyse par l'application des deux techniques complémentaires : la simulation et les réseaux de files d'attente. Ainsi, la simulation trouve son intérêt lorsqu'elle est utilisée pour modéliser les flux de patients complexes et pour tester les scénarios résultants du changement de certains paramètres. Dans cette recherche, les auteurs ont choisi la simulation des flux de patients en utilisant le logiciel Witness. Pour valider les résultats obtenus par ce logiciel, ils ont

proposé d'appliquer les réseaux de files d'attente. Les résultats obtenus par ces deux techniques étaient cohérents. Ils indiquent que les patients passent un temps d'attente plus important chez le médecin généraliste qu'au passage des patients chez le médecin spécialiste. De ce fait, ils ont utilisé le diagramme en arbre et la méthode AHP floue pour déterminer les indicateurs de performance sur lesquels ils ont agi et lancé à nouveau le modèle Witness. Ils ont terminé par l'application de la méthode PROMETHEE II floue [6] pour choisir les actions d'amélioration. Wang Tao [8] a travaillé sur l'organisation et pilotage des services sur le trajet des urgences. L'auteur a proposé des solutions appropriées aux services d'urgence permettant d'améliorer la prise en charge des urgences dans une vision globale intégrant l'ensemble des interactions entre les différents acteurs et entre les différents flux de patients. Notamment, son travail se focalise sur la gestion des lits d'hospitalisation en aval des urgences. un modèle mathématique est développé pour optimiser la planification des lits en tenant compte des flux des patients d'urgence et programmés. Il a étudié l'impact de la connaissance des disponibilités en lits d'aval au service d'urgence, et la planification des admissions en aval des urgences par le biais du couplage de méthodes de simulation et de recherche opérationnelle. Dans [7], les auteurs ont spécifié et ont développé une plate-forme de modélisation et de simulation dédiée aux systèmes hospitaliers, appelée medPRO (medical Process-Resource-Organisation), accompagnée d'une méthodologie d'analyse adaptée au domaine médical. Cette plate-forme est construite autour d'un cadre de modélisation et d'un guide méthodologique conçus pour un système hospitalier particulier. UML (Unified Modelling Language) a été choisi pour la modélisation de ces systèmes. Plusieurs points de vue sont proposés : Processus (vue centrée sur le patient), Ressource (comportement des ressources humaines et matérielles), et Organisation (relation entre les intervenants). Le comportement dynamique du modèle est spécifié grâce à une classe dédiée de réseaux de Petri, appelés réseaux de Petri de Santé : un algorithme de simulation à événements discrets a également été développé pour les réseaux de Petri. Une large partie de ce travail est dédiée au système de décision, qui est utilisé (i) pour appliquer des méthodes de planification et d'ordonnancement issues du génie industriel à des systèmes hospitaliers, et (ii) pour piloter le déroulement de la simulation en temps réel au travers d'une approche hybride hiérarchique/hétéroarchique. Trois études de cas sont également présentées pour montrer l'efficacité de la plateforme medPRO : ils se sont intéressés à l'unité neuro-vasculaire, à la pharmacie et au bloc opératoire du CHU de Saint-Etienne (France). Plusieurs outils d'optimisation spécifiques ont été développés et inclus dans la plate-forme. Colette Mercé et al. [1] ont détaillé la notion de "parcours du patient" qui constitue un concept nouveau qui dépasse en matière d'organisation et de rationalisation les approches traditionnelles. Gérer l'agenda du malade implique de recueillir le maximum d'informations sur le malade avant son entrée à l'hôpital et au début de son séjour puis tout au long de sa présence, de manière à effectuer une programmation prévisionnelle des soins et à raccourcir ainsi la durée de l'hospitalisation. Ce travail s'intéresse à l'analyse et à la modélisation du parcours du patient dans un contexte opérationnel. Ceci permet d'une part de mettre en relation les différents processus identifiés et d'autre part de positionner les différents centres de décisions impliqués. Il est alors possible d'envisager la planification

des activités associées à ce parcours afin de maîtriser au mieux son déroulement. Dans le but de modéliser et de simuler l'Unité d'Hémobiologie du Laboratoire Central de l'hôpital militaire régional universitaire d'Oran (HMRUO) en Algérie, et en utilisant la méthodologie ASCI (Analyse, Spécification, Conception et Implémentation), les auteurs de [5] ont procédé à l'évaluation du taux d'utilisation des salles et des médecins de cette unité. Pour cela, ils ont utilisé l'outil ARIS pour spécifier le modèle de connaissance ainsi que SIMULA, Witness et QNAP2 pour le passage d'un modèle de connaissance vers des modèles d'action. Un modèle intermédiaire de réseau de files d'attente a été employé. Nabila Dehas et al [2] ont fait une première étude sur la modélisation du mouvement des malades au urgences de l'hôpital Khelil Amrane béjaia (Algérie) selon l'organisation de 1999. D'autres travaux liés à la gestion hospitalière sont détaillés dans [3].

3.3 Mouvement des malades aux urgences de l'hôpital Khelil Amrane

Généralement, tous les malades avec leurs accompagnateurs, sont reçus au niveau de la salle d'attente, ou ils attendent leurs tours de passage. Les cas jugés graves par l'infirmier de service seront prioritaires et introduits directement dans la salle de consultation.

Après consultation et diagnostic, Le médecin prendra la décision appropriée.

Trois cas de figure peuvent se présenter :

- Le malade nécessite seulement un traitement médical, et juste une ordonnance médicale sera délivrée.
- Le malade est blessé ou fracturé, il est orienté vers les salles de soins ou de plâtre puis, sortira après traitement.
- Le malade présentant des signes de gravités est orienté par le médecin à la salle d'observation pour établissement de bilans standards et de radios, afin de la traiter ou de déterminer sa pathologie.

La décision de sa sortie, de son hospitalisation, ou de son évacuation sera prise après analyse du bilan médical par le médecin.

Le schéma 3.2 illustre les différentes étapes du mouvement des malades au niveau du service des urgences.

3.4 Collecte des données

Pour identifier les paramètres de notre modèle, nous avons effectué une collecte de données nécessaire à notre travail. Pour ce faire, avons chercher les informations sur une durée de 10 jours à raison 8h par jour.

Nous avons utilisé deux méthodes pour la collecte des données :

1. Enregistrer manuellement les entrées et les sorties des malades au niveau de la salle de consultation au service des urgences

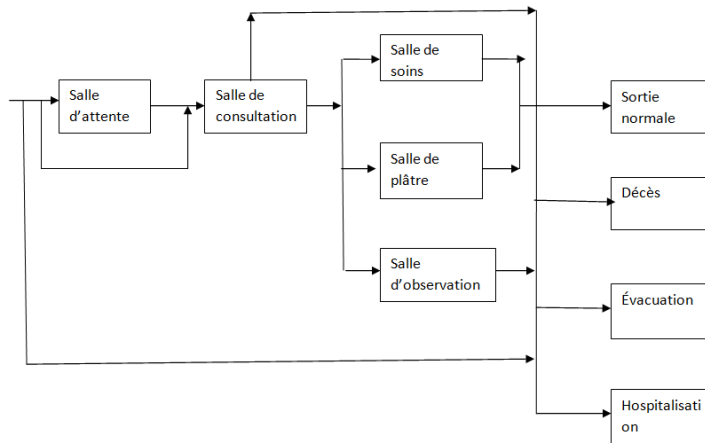


FIGURE 3.1: Mouvements des malades au niveau du service des urgences.

- Nous avons utilisé le registre des arrivées des malades de la salle de la consultation et les temps des entrées et des sorties au niveau de la salle d'observation ,la salle de soins, la salle de plâtre.

3.5 Modélisation du mouvement des malades

Le mouvement des malades au niveau de l'Hôpital Khelil Amrane peut être modélisé par un réseau de files d'attente. Les malades représentent les arrivées dans le système. Ils se distinguent par leurs classes, les serveurs du système sont les salles de consultation, la salle des soins, la salle de plâtre ou la salle d'observation.

3.5.1 Description des arrivées

Classes des malades

Les malades arrivés à l'Hôpital sont répartis en 02 classes :

C1 : Les malades jugés très grave.

C2 : Les malades moins graves.

Les données nous ont permis de calculer les proportions de chaque classe. Le tableau 3.1 résume les résultats obtenus.

Classe	C1	C2
Fréquence	0.34	0.66
Fréquence cumulée	0.34	1

TABLE 3.1: Fréquence des arrivées à la salle de consultation par classe.

Identification du flot des arrivées

Pour déterminer le flot des arrivées, nous avons utilisé 02 types de données :

- Le premier type de données utilisées s'étalent de la période allant du 1^{er} Avril 2016 au 30 Avril 2016, prélevées sur les registres tenus au niveau du service des urgences.
- Le deuxième type de données sont collectées au niveau de la salle d'attente sur une période de 03 jours.

Soient :

X_1 : le nombre des malades arrivés à une date donnée.

X_2 : représente l'intervalle de temps, en heure, séparant deux arrivées.

Les résultats obtenus après l'ajustement des variables X_1 et X_2 par le test de Kolmogorov-Smirnov sont présentés sur le tableau 3.2 :

Variable	Lois	KS Calcul	KS théor	Taille d'échant	Paramètre
X_1	Poisson	0.15426	0.23527	48	m=89.63889
X_2	Poisson	0.12536	0.290	30	m=89.63889

TABLE 3.2: Ajustement des arrivées .

Présentation des résultats

L'ajustement de la variable X_1 par une loi de Poisson donne une statistique de Kolmogorov-Smirnov égale à 0.15426. cette valeur est inférieure à la valeur tabulée (0.18763), au niveau $\alpha=0.01$. L'ajustement de la variable X_2 par une loi de Poisson donne une statistique de Kolmogorov-Smirnov égale à 0.12536. cette valeur est inférieure à la valeur tabulée (0.240), au niveau $\alpha=0.01$.

3.5.2 Description des serveurs

Aprésent, nous allons définir les serveurs de chaque salle du service des urgences.

Type de serveur

Le service des urgences est composé de 04 stations : la salle de soins, la salle de plâtre, la salle d'observation et salle de consultation. Tous les malades passent par la salle de consultation.

A la sortie de cette salle, certains malades sont orientés vers d'autres salles. On sépare 04 classes de la manière suivante :

C_3 : Elle représente les malades qui sont affectés à la salle de soins.

C_4 : Elle représente les malades qui sont affectés à la salle de plâtre.

C_5 : Elle représente les malades qui sont affectés à la salle d'observation.

$C6$: Elle représente les malades sortant après consultation.

Les fréquences associées à ces classes sont regroupées dans le tableau 3.3 :

Classe	$C3$	$C4$	$C5$	$C6$
Fréquence	0.06	0.21	0.34	0.39
Fréquence cumulée	0.06	0.27	0.61	1

TABLE 3.3: Les fréquences associées à chaque classes.

A la sortie d'un malade de la salle d'observation, soit il est hospitalisé, soit il quitte le système, selon les fréquences cumulées citées dans le tableau 3.4

Classe	Orientation	Hospitalisation
Fréquence	0.16	0.84
Fréquence cumulée	0.16	1

TABLE 3.4: Fréquences d'orientation des malades.

3.5.3 Discipline de service

La discipline de service de chaque serveur est la suivante :

- Pour la salle de consultation, la discipline de service entre les malades de la classe $C1$ et ceux de classe $C2$ est la priorité relative.
- Pour les autres salles, les malades sont servis selon la discipline FIFO.
- Le nombre de serveurs associés à chaque salle du service des urgences est présenté par le tableau 3.5

Salle	Nombre de serveurs
Consultation	02 médecins
Soins	03 paillasse
Plâtre	01 paillasse
Observation	15 lits

TABLE 3.5: Les serveurs associés à chaque salle de service des urgences.

3.5.4 Identification de la loi de la durée de service

Les durées de service données en minute, pour la salle de consultation, la salle de soins et la salle de plâtre. Les durées de service données en jour, pour la salle d'observation et sont calculées par la formule 3.1

$$D_i = S_i - E_i \quad (3.1)$$

avec

E_i : temps d'entrée d'un malade i ;

S_i : temps de sortie d'un malade i ;

On considère

$Y1$: la durée de service dans la salle de consultation.

$Y2$: la durée de service dans la salle de plâtre.

$Y3$: la durée de service dans la salle de soins.

$Y4$: la durée de service dans la salle d'observation.

Les lois des durées de service sont ajustées à l'aide du test de Kolmogorov-Smirnov. Les résultats obtenus sont représentés sur le tableau 4.1

Variable	Lois	KS Calcul	KS théor	Taille d'échant	Paramètre
Y1	Exponentielle	0.1619	0.221814904	54	m=0.05337362
Y2	Exponentielle	0.25898	0.290	30	m=0.05628518
Y3	Exponentielle	0.31999	0.356	20	m= 0.06116208
Y4	Exponentielle	0.14643	0.270	35	m=0.1320755

TABLE 3.6: Ajustement des durées des salles.

3.5.5 Interprétation

L'ajustement des variable $Y1$, $Y2$, $Y3$ et $Y4$ par une loi Exponentielle donne des statistiques inférieures aux valeurs tabulées de Kolmogorov-Smirnov, au niveau $\alpha=0.01$. On accepte alors l'ajustement des durées de service des salles par une loi exponentielle.

3.6 Le modèle obtenu

Nous avons modélisé le système par un réseau de files d'attente.

- La première file correspond à la salle de consultation. Elle est de type $M/M/2$.
- La deuxième file correspond à la salle de soins. Elle est de type $M/M/3$.
- La troisième file correspond à la salle de plâtre. Elle est de type $M/M/1$.
- La quatrième file correspond à la salle d'observation. Elle est de type $M/M/15$.

3.7 Simulation

Dans cette partie, nous constituons un algorithme sur l'environnement Matlab pour évaluer les performances du système en cas de changement de quelques paramètres : le taux d'arrivée, le taux de service et les fréquences des malades qui sont affectés de la salle de consultation vers

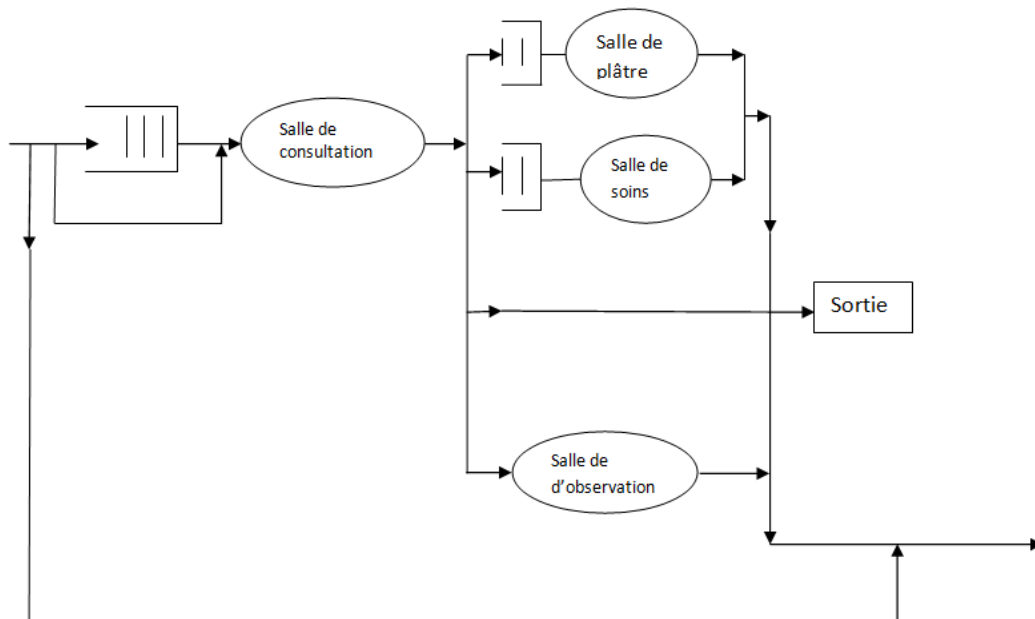


FIGURE 3.2: Modèle analytique.

les autres salles.

Cet algorithme se représente comme suit :

Etape01 : lire $(\lambda, \mu, s$ et $p1, p2, p3)$;

Etape02 : calculer les caractéristiques du système de files d'attente suivant :

si $M/M/1$ **alors**

$$L_s = \frac{\rho}{1 - \rho}$$

$$W_s = \frac{1}{\mu - \lambda}$$

$$L_q = \frac{\rho^2}{1 - \rho}$$

$$W_q = \frac{\rho}{\mu(1 - \rho)}$$

$$\rho = \frac{\lambda}{\mu}$$

sinon $M/M/s$

avec P_0 est la probabilité que la station soit vide.

$$p_0 = \left[\sum_{n=0}^s \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^{s+1}}{s!(s - \lambda/\mu)} \right]^{-1}$$

$$L_s = L_q + \frac{\lambda}{\mu}$$

$$L_q = \left(\frac{\lambda}{\mu}\right)^s \frac{1}{s!} \frac{\rho}{(1 - \rho)^2} p_0$$

$$\rho = \frac{\lambda}{s\mu}$$

$$W_s = W_q + \frac{1}{\mu}$$

$$W_q = \frac{L_q}{\lambda}$$

Les tableaux 4.4,3.8, 3.9 et 3.10 résument les résultats obtenus par notre algorithme

Les données	Les entrés	Les résultats	Les sorties
s	2	L_s	0.0139
λ	15	Lq_1	0.0071
μ	20	Lq_2	0.0018
p_1	0,2	W_s	0.1006
P_2	0,35	Wq_1	4.7217e-004
p_3	0,1	Wq_2	1.2221e-004

TABLE 3.7: Les résultats de simulation de service des urgences si $\lambda = 15$ et $\mu = 20$.

Les données	Les entrés	Les résultats	Les sorties
s	2	L_s	0.0322
λ	10	Lq_1	0.0142
μ	12	Lq_2	0.0041
p_1	0,5	W_s	0.1685
P_2	0,1	Wq_1	0.0014
p_3	0,2	Wq_2	4.1335e-004

TABLE 3.8: Les résultats de simulation de service des urgences si $\lambda = 10$ et $\mu = 12$.

Les données	Les entrés	Les résultats	Les sorties
S	2	L_s	0.0172
λ	20	Lq_1	0.0110
μ	18	Lq_2	7.6153e-014
p_1	0,1	W_s	0.1117
P_2	0,3	Wq_1	5.4924e-004
p_3	0,5	Wq_2	3.8077e-015

TABLE 3.9: Les résultats de simulation de service des urgences si $\lambda = 20$ et $\mu = 18$.

Tableau 4 :

Sachant que :

λ : le nombre des malades arrivés par heure.

μ : le nombre des malades sortis par heure.

p_1 : La fréquence des malades qui sont affectés vers la salle de soins.

P_2 : La fréquence des malades qui sont affectés vers la salle de plâtre.

Les données	Les entrés	Les résultats	Les sorties
s	2	L_s	-0.0621
λ	35	Lq_1	-0.0786
μ	12	Lq_2	0.0026
p_1	0,1	W_s	0.1645
P_2	0,2	Wq_1	-0.0022
p_3	0,15	Wq_2	7.4614e-005

TABLE 3.10: Les résultats de simulation de service des urgences si $\lambda = 35$ et $\mu = 12$.

p_3 : La fréquence des malades qui sont affectés vers la salle d'observation.

L_s : Le nombre moyen de clients dans le système.

Lq_1 : Le nombre moyen de clients dans la salle de consultation.

Lq_2 : Le nombre moyen de clients dans les autres salles.

W_s : Temps moyen de séjour d'un client dans le système.

Wq_1 : Temps moyen d'attente d'un client dans la salle de consultation.

Wq_2 : Temps moyen d'attente d'un client dans les autres salles.

3.7.1 Discussion

Les résultats présentés dans les tableaux 4.4 et 3.8 montrent que lorsque le taux λ des arrivées aux urgences est inférieur au temps moyen de service μ ($\lambda < \mu$), le système est stable. C'est à dire que les capacités de l'hôpital à l'heure de notre étude peuvent répondre aux besoins des citoyens. Par contre, et d'après les tableaux 3.9 et 3.10, quand μ ($\lambda > \mu$), le système est insatiable. Pratiquement, cette situation correspond aux cas où le flux des malades augmente à cause d'un imprévu (accident, épidémie, malaies saisonnière...). C'est une situation qui exige que les dirigeants du service des urgences doivent réagir au plus vite en trouvant des solutions immédiates.

3.8 Conclusion

Dans ce travail, nous avons présenté une méthodologie de modélisation du mouvement des malades au niveau du service des urgences de l'Hôpital Khellil Amrane de Béjaia, en vue de déterminer les besoins théoriques de chaque salle de service. Le mouvement des malades (au niveau du service des urgences) a été modélisé par un réseau de file d'attente. Le traitement statistique des données a permis de déterminer les lois régissant le modèle. Il révèle que le processus d'arrivées des malades est Poissonien, et que la loi de chaque salle associée au service des urgences est exponentielle. L'ajustement de ces lois a été validé par le test de Kolmogorov-Smirnov.

Pour voir les limites du service d'urgence tel qu'il est actuellement, nous avons réalisé des simulations pour différentes valeurs du taux des arrivées des patients et du temps moyen de service. Cette procédure montre l'impact d'une situation inhabituelle sur la performance du système.

Pour trouver des solutions, nous avons formulé des recommandations aux dirigeants du service des urgences. Ces recommandations sont :

- Définir un plan d’actions en cas d’une situation sanitaires inhabituelle, par exemple : faire appel à d’autres médecins urgentistes ou mobiliser les étudiants internistes pour une bonne prise en charge, mobiliser des lits dans d’autres services pour les patients nécessitant une prise en charge prolongée, orienter les patients peu urgents vers d’autres structures de santé.
- Mettre en place un système d’information pour la collecte des données et effectuer des prévisions pour préméditer la prise de décision.
- Modifié le réseau de file d’attente donné dans ce travail pour voir l’impact de l’ajout d’un médecin ou de lits sur le fonctionnement du service des urgences.

Plusieurs travaux de recherches peuvent compléter ce travail : La modélisation et la simulation des autres services de l’hôpital, l’utilisation d’autres outils de modélisation et de simulation, l’étude de la planification des services de l’hôpital.

Références

1. Colette, M., Gérard, F., Daniel, D. (2004). 'Analyse et modélisation du parcours du patient'. *Logistique et Management*. pp 263-44.
2. Dehas, N., Aissani, A., Adjabi, S., Abedrahmani, H. (2006). 'Evaluation des Performances d'un Système de Santé : Cas de l'Hôpital Khellil Amrane (Béjaïa)', GISEH'06, septembre 2006 Luxembourg .
3. El oualidi, M., Saadi, J., El hiki, L., Artiba, A., Bellabdaoui, A. (2010). 'Modélisation et simulation du flux des patients au service des urgences. Cas de l'hôpital Ibn Rochd à Casablanca (Maroc)'. Conférence Francophone Gestion et Ingénierie des Systèmes Hospitaliers. Maroc.
4. Jlassi, J. (2009). 'Amélioration de la performance par la modélisation des flux logistiques des patients dans un service d'urgence hospitalier'. *Thèse de Doctorat*, Sciences de l'ingénieur [physics]. Université Paris VIII Vincennes-Saint Denis.
5. Khaled, B., Alain, T. (2010). 'Modélisation et simulation de l'unité d'hémobiologie du laboratoire central de l'Hmruo d'Oran-Algérie', *In proceeding 8^{ème} Conférence Internationale de MOdélisation et SIMulation - MOSIM'10, 10 au 12 mai 2010*. pp 974-50.
6. Martel, A. (1979). 'Technique et application de la recherche opérationnelle', Edition Gieten Morin.
7. Vincent, A. (2008). 'Modélisation, analyse et pilotage de flux en milieu hospitalier à l'aide d'UML et des réseaux de Petri'. *Thèse de Doctorat*, Génie Industriel. Université de Lille.
8. Wang, T. (2008). 'Organisation et pilotage des services sur le trajet des urgences'. *Thèse de Doctorat*, Informatique. INSA de Lyon.

Application des Méthodes d'Apprentissage dans la Prédiction du Diabète de Type 2.

Mehidi D., Medjoudj S., Adel-Aissanou K. et Aïssani D.

Research Unit **LaMOS** (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria
lamos_bejaia@hotmail.com , ak_adel@yahoo.fr

Résumé L'utilisation des systèmes experts et les techniques dites intelligentes en diagnostic médical ne cesse d'augmenter graduellement. L'apprentissage automatique est une méthode parmi d'autres utilisées dans le diagnostic médical. Cet article présente une approche d'apprentissage supervisé basé sur les algorithmes K-Plus Proche Voisin (K-PPV), SVM, Naïve Bayes, Radom Forest, Décision Tee classifieur (Arbre de décision), pour reconnaître les personnes susceptibles de développer un diabète en utilisant deux bases de données différentes, à savoir, celles du CHU et du cabinet privé du Dr Djamel MEHIDI. Les performances des classifieurs ont été comparées en fonction du taux de précision, temps d'exécution. Les plus hauts taux de classification obtenus par l'application de Radom Forest et Naïve Bayes sont respectivement 86% et 85%, en appliquant l'approche 10-folds cross-validation.

Mots clés : Prédiction du diabète, K-Plus Proche Voisin (K-PPV), SVM, Naïve Bayes, Radom Forest, Arbre de décision.

4.1 Introduction

Le diabète mellitus, ou diabète sucré, aussi appelé simplement diabète est l'une des maladies majeures de notre monde moderne. En 2019, la Fédération Internationale pour le Diabète a estimé que 463 millions d'adultes dans le monde sont diabétiques et que la maladie est directement à l'origine de 4.2 millions de morts [7]. La prévalence de la maladie dans le monde est prédite d'évoluer de 9.3% à 10.9% (700 millions) d'ici 2045. Voir la figure 4.1.



Source : International Diabetes Federation

FIGURE 4.1: Le Diabète dans le monde

En Algérie, sa prévalence en 2022 est estimée à 14.4%, soit près de 4.5 millions de personnes. C'est pourquoi, il y a un véritable besoin de sensibilisation et de prévention de cette

maladie, encore trop ignorée à ce jour.

Nous entendons chaque jour le terme intelligence artificielle, l'IA définie, plus généralement, comme la capacité d'une machine capable d'agir, par elle-même ou sous le contrôle de l'homme à reproduire des actions ou des fonctions qui sont habituellement celles des êtres humains. Aujourd'hui, nous la retrouvons dans nos machines informatiques, objets connectés, applications, réseaux sociaux, transports, et dans le secteur médical. L'application de l'IA à la médecine offre une perspective essentielle à l'essor de ces nouvelles technologies, qu'il s'agisse de renforcer le lien entre patients et médecins, de poser des diagnostics plus rapides et plus précis, ou encore d'optimiser la création de nouveaux traitements. L'innovation a pour objectif de combattre la mort et la maladie. Quoi de plus noble ?

L'IA permet aux médecins de gagner du temps en laissant la machine analyser elle-même les données et fournir des estimations. Le but à plus long terme : réussir à prédire de nombreuses maladies, afin que les médecins puissent intervenir le plus tôt possible.

Le diagnostic médical est un processus de classification. L'utilisation de l'informatique pour la réalisation de cette classification devient de plus en plus fréquente. Même si la décision de l'expert est le facteur le plus important lors du diagnostic, les systèmes de classification fournissent une aide substantielle, car elles réduisent les erreurs dues à la fatigue et le temps nécessaire pour le diagnostic. Actuellement, la plupart des hôpitaux modernes sont bien équipés avec des dispositifs de collecte de données. L'augmentation du volume de données entraîne des difficultés à extraire des informations utiles pour l'aide à la décision. Les méthodes traditionnelles d'analyse de données peuvent être complètes par l'utilisation de méthodes dites intelligentes.

L'apprentissage est la capacité de s'améliorer avec l'expérience, de se rappeler les décisions antérieures et les résultats afin de faire de meilleurs choix à l'avenir dans des situations similaires. L'apprentissage automatique est une discipline de l'intelligence artificielle. L'apprentissage automatique cherche à trouver le moyen de construire des programmes informatiques qui s'améliorent automatiquement avec l'expérience. L'une des métaphores utilisées dans le domaine de l'apprentissage automatique et qui considère la résolution de problèmes comme un type d'apprentissage qui, une fois le problème résolu, est capable de reconnaître la problématique et réagir en utilisant la stratégie apprise.

Dans ce travail, nous nous intéresserons à la prédiction du diabète type 2 qui est un dysfonctionnement du système de régulation de la glycémie. Beaucoup de travaux ont été menés afin d'effectuer la classification ou le diagnostic du diabète. Hung-Chun Lin et al. ont obtenu une précision de 62.8% en utilisant un ensemble de cellules mémoires qui ont subi un apprentissage leur permettant d'effectuer une classification et cela grâce à la méthode K-plus proche voisin, dans la classification du diabète. Une autre méthode a utilisé un réseau de neurones pour

effectuer la classification du diabète. L'apprentissage du réseau de neurones a été effectué pour chaque patient sur une période de 24h et le test a été effectué sur une autre période de 24h. L'erreur de prédiction sur les patients a été plus faible que 10mg/dl . Keller et al ont proposé en 1985 la classification du diabète qui est effectuée par l'algorithme K-ppv flou considérée comme étant l'une des plus importantes méthodes parmi les algorithmes non-paramétriques. Les auteurs ont proposé le classifieur K-ppv flou. Pour cela, l'algorithme K-ppv flou alloue au vecteur de données un degré d'appartenance à une classe donnée. Cet algorithme a permis d'améliorer le taux de classification de 6.42% par rapport aux autres algorithmes classiques.

La méthode appliquée dans ce travail pour la prédiction du diabète de type 2 est l'application d'algorithmes à apprentissage supervisé. Notre méthode est une évolution d'algorithmes étudiés. Nous effectuons un apprentissage supervisé sur la base de données du CHU de Bejaia et du cabinet médical de Dr Mehidi. En plus, nous améliorons le meilleur algorithme qui donnera comme résultat une classification des patients qui peuvent être susceptibles d'être atteint du diabète et qui a comme but l'aide à la décision du médecin en termes de temps d'examination et de taux de précision du résultat.

4.2 Travaux antérieurs

Dans [8], Les auteurs ont proposé un modèle de prédiction du diabète de type1 basé sur les réseaux de neurones, ce modèle est testé sur 22 patients (12 femmes et 10 hommes). Dont 14 sont diabétiques et 8 en bonne santé. Parmi les 22, 16 patients ont suivi un traitement à injection d'insuline et 6 avec une pompe. L'évolution du niveau de glucose a été enregistrée pour chaque patient, toutes les 5 minutes pendant 3 jours. Chaque patient a eu une vie normale avec repas et activités autant à la maison qu'au bureau. Trois cas ont été pris en compte : un patient avec un traitement classique à base d'injection (P1), un patient avec un traitement basé sur la pompe à insuline (P2) et une personne en bonne santé (P3). Pour P1, une grande variabilité du niveau de glucose a été mesurée. Pour P2 la variabilité est plus réduite mais encore importante par rapport à la personne en bonne santé. L'apprentissage du réseau de neurones a été effectué pour chaque patient sur une période de 24h et le test a été effectué sur une autre période de 24h. L'erreur de prédiction sur les 3 patients a été plus faible que 10mg/dl . Hung-Chun Lin et al [9] ont utilisé l'algorithme AIRS2 (Artificial Immune Recognition System) qui est inspiré du système immunitaire pour la prédiction du diabète du type 2 avec ensemble de cellules mémoires qui ont subit un apprentissage leur permettant d'effectuer une classification et cela grâce à l'algorithme AIRS. Les auteurs ont combiné une hybridation de deux algorithmes, le premier est l'AIRS2 qui est une évolution de l'algorithme AIRS. Le deuxième algorithme est le K-plus proche voisin flou. L'algorithme AIRS2 effectuera un apprentissage supervisé sur la base de données du diabète et donnera comme résultat un ensemble de cellules mémoires. Le K-plus proche voisin effectuera la classification de la base de test en utilisant l'ensemble de cellules mémoires généré dans la phase d'apprentissage. Keller et al [10], ont proposé en 1985, le classifieur K-ppv flou pour palier aux limitations du K-ppv. L'algorithme K-ppv flou

alloue au vecteur de données un degré d'appartenance à une classe donnée. Le principe est d'allouer le degré d'appartenance à une classe en fonction de la distance du vecteur de ces K-ppv et de l'appartenance de ces voisins à la classe. la classification est effectuée usuellement par l'algorithme K-ppv. L'algorithme a été remplacé par le K-ppv flou et est l'une des plus importantes méthodes parmi les algorithmes non paramétriques, qu'il ne spécifie pas le degré d'appartenance du vecteur à la classe qui lui a été attribué.

4.3 Notre proposition

Les algorithmes d'apprentissage automatique peuvent nous aider à détecter l'apparition du diabète. La détection précoce du diabète peut réduire les risques pour la santé du patient. Les médecins, les patients et les proches du patient peuvent bénéficier des résultats de la prédiction. Dans les milieux cliniques à faibles ressources, il est nécessaire de prédire l'état du patient après l'admission pour répartir les ressources de manière appropriée. Dans l'étude qui suit, l'objectif principal est d'appliquer ces différents algorithmes de classification sur des données concernant le risques de développer un diabète de type 2 récoltés au niveau de deux infrastructures différentes dans le but d'aboutir à un résultat d'aide à la prédiction de ce type de diabète.

En effet, avant de se focaliser sur les outils, nous avons procédé aux points suivant :

1. Collecter les données, et les rendre propre (renseigner les valeurs vides, supprimer et remplacer les valeurs aberrantes).
2. Sélectionner les variables pertinentes (table de corrélation).
3. Application de la fonction `train_test_split` pour découper nos données en données d'apprentissage et test.
4. Apprentissage et évaluation, en appliquant la technique cross validation.
5. Evaluation de performance de chaque algorithme.
6. Optimisation de l'algorithme RandomForest (chercher les bonnes valeurs pour chaque paramètre) via la fonction `RandomizedSearchCV`.
7. Exportation du model (meilleur algorithme + prétraitement) sous format binaire avec la librairie `pickle`.
8. Développement d'une application web en python (à base du framework Flask) qui permet aux visiteurs de saisir leurs données et de prédire s'ils sont susceptibles de développer un diabète ou pas.

La prévalence du diabète de type 2 est sous estimée car cette anomalie glycémique asymptotique peut évoluer de façon insidieuse et silencieuse pendant de nombreuses années avant que le diagnostic ne soit porté, le travail mentionné est utile aux médecins comme outil d'aide à la décision.

4.3.1 Méthodes

Pour répondre à cette recherche comparative, nous utiliserons les deux bases de données issues du Centre Hospitalo-Universitaire de Bejaia et une du cabinet médical du Dr Djamel MEHIDI qui contiennent respectivement des informations de 268 et 500 patients tous genres confondus (femmes et hommes). Elles seront utilisées dans la classification des algorithmes d'apprentissage supervisé tel que Naives Bayes, le Plus Proche Voisin et SVM.

4.3.2 Outils utilisés

Python : Est un langage de programmation puissant et facile à apprendre. Il dispose de structures de données de haut niveau et permet une approche simple mais efficace de la programmation orientée objet. Parce que sa syntaxe est élégante, que son typage est dynamique et qu'il est facile à interpréter, Python est un langage idéal pour l'écriture de scripts et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plateformes. L'interpréteur Python et sa vaste bibliothèque standard sont disponibles librement, sous forme de sources ou de binaires, pour toutes les plateformes majeures depuis le site Internet <https://www.python.org/> et peuvent être librement redistribués. Ce même site distribue et pointe vers des modules, des programmes et des outils tiers. Enfin, il constitue une source de documentation. L'interpréteur Python peut être facilement étendu par de nouvelles fonctions et types de données implémentés en C ou C++ (ou tout autre langage appellable depuis le C). Python est également adapté comme langage d'extension pour personnaliser des applications [1].

Scikit-Learn : Est une bibliothèque qui fournit une gamme d'algorithmes d'apprentissage supervisés et non supervisés via une interface cohérente en Python. La vision de la bibliothèque est un niveau de robustesse et de support requis pour une utilisation dans les systèmes de production. Cela signifie qu'il faut se concentrer sur des préoccupations telles que la simplicité d'utilisation, la qualité du code, la collaboration, la documentation et les performances [2].

Pandas : [3] Est une bibliothèque écrite pour le langage de programmation Python permettant la manipulation et l'analyse des données. Elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques. Pandas est un logiciel libre sous licence. Les principales structures de données sont :

- Les séries (pour stocker des données selon une dimension - grandeur en fonction d'un index).
- Les DataFrames : pour manipuler des données aisément et efficacement avec des index pouvant être des chaînes de caractères (stocker des données selon 2 dimensions - lignes et colonnes).
- Les Panels pour représenter des données selon 3 dimensions.
- Format de lecture et écriture des données structurées en mémoire depuis et vers différents formats : fichiers CSV, fichiers textuels, fichier du tableur Microsoft Excel, base de données SQL.

Spyder : [4] Est un environnement de développement pour Python, libre et multiplateforme (Windows, Mac OS, GNU/Linux), il intègre de nombreuses bibliothèques d'usage scientifique.

Spyder a un ensemble unique de fonctionnalités - multiplateforme, open-source, écrit en Python et disponible sous une licence non-copyleft. Spyder est extensible avec des plugins, comprend le support d'outils interactifs pour l'inspection des données et incorpore des instruments d'assurance de la qualité et d'introspection spécifiques au code Python.

Il offre une combinaison unique de fonctionnalités avancées d'édition, d'analyse, de débogage et de profilage d'un outil de développement complet avec l'exploration de données, l'exécution interactive, l'inspection approfondie et les capacités de visualisation d'un package scientifique.

Python-Seaborn (Visualisation de données statistiques pour python) : Seaborn est une bibliothèque pour créer des graphiques statistiques attrayants et informatifs en Python[5].

4.3.3 Les données

C'est la totalité des éléments pris en considération, sur lesquels nous utiliserons l'apprentissage automatique dans le but d'appliquer la classification de l'apprentissage supervisé dans la prédiction du diabète de type 2.

La collecte de données a été effectuée au niveau de deux infrastructures différentes ; l'une au Centre Hospitalo Universitaire de Bejaia (CHU), au service de médecine interne dirigé par le professeur BOUALI, la seconde au niveau du cabinet médical privé de Dr Djamel MEHIDI, où nous avons effectué une collecte de données respectivement de 267 et 500 patients. Les problèmes d'apprentissage sont énoncés sous forme de données. Ces séries caractérisant une série d'instances du phénomène à apprendre, que l'on nomme patient. Chaque patient P est constitué d'une description D et d'une sortie S

$$P = (D, S)$$

Où :

$$D \in X = \{\text{Genre, \hat{A}ge, Groupe sanguin, Pression artérielle, IMC, Antecedants familiaux, Sportif, Fumeur}\}$$

$$S \in Y = \{0, 1\}$$

On nomme respectivement les ensembles X et Y l'espace d'entrée et l'espace de sortie. La description $X = \{x_1, x_2, \dots, x_n\}$ de l'ensemble des données s'avère être un vecteur de valeurs réelles de dimension n , où n est le nombre des attributs de patients. Selon le phénomène représenté, un attribut peut être numérique (par exemple, âge) ou de type libraire (par exemple, le genre a été déterminé arbitrairement par la valeur 0 équivaut à « homme » et 1 équivaut à « femme »). La sortie Y du patient détermine sa classe d'appartenance. Les problèmes étudiés dans ces études seront tous des problèmes de classification binaire, c'est-à-dire des problèmes de classification à seulement deux classes. Pour chaque problème, on détermine arbitrairement que les patients appartenant à une classe sont des patients malade « 0 » et que les autres patients ne sont pas malade « 1 ».

Durant la phase de collecte des données, nous avons constaté que certaines informations étaient manquantes notamment dans l'âge et qu'il existait des valeurs aberrantes particulièrement

dans la pression artérielle. C'est pourquoi nous avons dû effectuer un prétraitement avant de commencer notre étude. Les ensembles de données contiennent 10 variables particulières qui ont été considérées comme des facteurs à risque élevés de développer un diabète. Cette base de données contient des patients de tout genre et âge confondus. Les valeurs sont de différents types ; le tableau 4.1 la description de notre ensemble de données :

Paramètre physiologique	Description	Valeurs	Analyse de données
Genre	Homme ou Femme	0 ou 1	Homme :397(M) Femme : 371 (F)
Âge	Age de la personne		Âge 5 à 20 ans : 30 Âge 21 à 35 ans : 131 Âge 36 à 50 ans : 142 Âge 51 à 78 ans : 97
Groupe sanguin	Groupe sanguin	A ou B ou AB ou O	A : 242 B : 144 AB : 164 O : 218
IMC	$Imc = ((taille) * 2 / poids)$	Bon ou Mauvais	
Activité physique	Oui Non	1 ou 0 ou Non renseigné	Oui : 81(1) Non : 137(0) Non renseigné : (551)
Pression artérielle	Optimale Normale Elevée	≥ 14 < 14	767 4
Fumeur	Oui Non	1 ou 0 ou Non renseigné	Oui : 234(1) Non : 534(0) Non renseigné : (1)
Antécédents familiaux	Oui Non	1 ou 0 ou Non renseigné	Oui : 231 Non : 227 Non renseigné :310
Outcome	Si une personne est diabétique ou pas	0 ou 1	Non : 335(1) Oui : 432(0)

TABLE 4.1: Analyse des données collectées

4.3.4 Préparation et nettoyage des données

Afin de sélectionner les variables les plus significatives pour les modèles, nous commençons par épurer les données : élimination de certaines valeurs, traitement des valeurs manquantes, détection des valeurs aberrantes et bien d'autres types d'incohérences qui peuvent gêner l'analyse.

La figure 4.2 représente une description des données obtenues

description :	Age	Pression_arterielle	Insuline	IMC	Glucose \
count	764.000000	768.000000	768.000000	768.000000	768.000000
mean	33.267016	70.928385	79.799479	31.992578	122.572917
std	11.782850	32.322266	115.244002	7.884160	32.458296
min	21.000000	0.000000	0.000000	0.000000	0.000000
25%	24.000000	64.000000	0.000000	27.300000	100.000000
50%	29.000000	72.000000	30.500000	32.000000	118.500000
75%	41.000000	80.000000	127.250000	36.600000	144.000000
max	81.000000	594.000000	846.000000	67.100000	199.000000

	Sportif	Fumeur	Malade
count	768.000000	768.000000	768.000000
mean	0.397135	0.304688	0.436198
std	0.489623	0.460575	0.496236
min	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000
75%	1.000000	1.000000	1.000000

FIGURE 4.2: Description des deux bases de données

D'après la représentation des données, on constate que :

- Des effectifs très grands (768 patients) ;
- La moyenne de chaque donnée est calculée ;
- Les bornes min et max sont affichées ;
- Les quartiles 25% 50% 75% sont donnés.

Pendant la récolte des données, nous avons constaté que certaines valeurs d'âge étaient manquantes et que des valeurs aberrantes de pression artérielle existaient, ce qui poserait des problèmes pendant la phase d'apprentissage. Nous avons remédié à cette situation en utilisant 2 méthodes :

– **Remplacer les valeurs manquantes d'âge**

Pour ce faire, nous avons créé une fonction calculer moyenne qui compte le nombre de ligne vide ou égale à zéro de la colonne âge, une fois le nombre de case trouvé ; la moyenne globale d'âge des patients sera calculée (résultat retourné : moyenne âge = 33 ans) puis introduite dans les cases vides ou égale à zéro.

– **Remplacer les valeurs aberrantes de la pression artérielle**

Nous avons fait appel à la méthode *data.loc* dans le but de repérer les valeurs supérieures à 140 mm hg, le résultat est donné dans la figure4.3

N° Patient	Pression artérielle
13.0	260
19.0	470
24.0	594
767.0	270

FIGURE 4.3: Données aberrantes de la pression artérielle

Une fois le résultat retourné, nous les avons remplacés par des 0.0 puis calculer la moyenne globale de cette dernière : *moyenne_Pression_arterielle*, une fois le résultat retourné nous les avons remplacé par cette nouvelle valeur.

4.3.5 Choix des attributs

Afin de sélectionner les variables les plus significatives pour les modèles, nous avons réalisé une étude des corrélations sur notre base de données. cette étude est nécessaire afin de répertorier les différentes relations entre ces données. La corrélation étudie l'intensité de la liaison qui peut exister entre ces variables [6]. Le coefficient de corrélation entre deux variables aléatoires réelles X et Y ayant chacune une variance est définie par r

$$r = \frac{Cov(X, Y)}{\sigma_x \sigma_y}$$

Où $Cov(X, Y)$ désigne la covariance des variables X et Y et σ_x, σ_y leurs écarts types respectifs. Le tableau donné dans la figure 4.4, représente les relations entre les différentes variables :

	Genre	Age	Groupe_sanguin	Pression_artérielle	IMC	Antécédants_familiaux	Sportif	Fumeur	Malade
Genre	1.000000	0.013412	0.008795	-0.039266	0.001274	0.093322	0.035481	-0.022866	0.111236
Age	0.013412	1.000000	-0.063793	-0.159273	0.038377	0.074763	-0.017677	0.012488	0.263138
Groupe_sanguin	0.008795	-0.063793	1.000000	0.016537	0.008306	-0.016168	-0.024722	-0.004888	-0.054919
Pression_artérielle	-0.039266	-0.159273	0.016537	1.000000	-0.136239	-0.041651	0.000341	-0.046338	-0.086456
IMC	0.001274	0.038377	0.008306	-0.136239	1.000000	0.147906	0.038119	-0.007347	0.273021
Antécédants_familiaux	0.093322	0.074763	-0.016168	-0.041651	0.147906	1.000000	0.028286	-0.025816	0.489008
Sportif	0.035481	-0.017677	-0.024722	0.000341	0.038119	0.028286	1.000000	-0.022720	0.042712
Fumeur	-0.022866	0.012488	-0.004888	-0.046338	-0.007347	-0.025816	-0.022720	1.000000	-0.023219
Malade	0.111236	0.263138	-0.054919	-0.086456	0.273021	0.489008	0.042712	-0.023219	1.000000

FIGURE 4.4: Table de corrélation des deux bases de données

– Interprétation du tableau

Il existe une forte corrélation entre les différentes variables telle que ; malade et genre, malade et âge, malade et IMC, malade et antécédents familiaux, malade et sportif ;

Il existe une faible corrélation entre les différentes variables telle que ; malade et groupe sanguin, malade et pression artérielle, malade et fumeur.

– Sélection de variables

Cette table de corrélation préposée pour le problème de sélection des variables à hauts risques de développer un diabète peut être décomposée en deux classes ; une classe à forte corrélation qui comporte des variables à hauts risques de développer un diabète : genre, âge, IMC, antécédents familiaux et sportif, une autre classe à faible corrélation qui comporte à son tour des variables à faible risque : pression artérielle, groupe sanguin ainsi que fumeur. Dans ce qui suit, dans notre système d'apprentissage, nous utiliserons les variables ayant une forte corrélation entre elles car celles ayant une faible corrélation n'ont pas un grand impact dans la prédiction du diabète.

4.3.6 Apprentissage

Notre but est une aide à la décision dans la prédiction de développement d'un diabète chez une personne. Nous nous intéresserons à la classification dans l'apprentissage supervisé, il est essentiel dans tout programme d'apprentissage d'avoir un mécanisme qui sépare l'ensemble des données en deux groupes. Afin d'assurer ce mécanisme, nous aurons recours à la méthode de Cross Validation.

Data-train :(ensemble d'apprentissage) qui consiste à faire un apprentissage des données sur elle-même. Elle comprend la partie apprentissage et la partie validation.

Data-test : (ensemble test) qui consiste à tester les données restantes dans la Data-train, pour le test final de notre modèle.

Apprentissage et test

Dans un premier temps, nous avons partagé nos deux différentes bases de données comme suit

- Cas 1 : base de données du cabinet médical.
- Cas 2 : base de données du CHU.

Dans cette partie, la fonction *train_test_split()* prend en paramètres nos variables d'entrées (X) des deux cas, leurs variables de sortie (Y) et un pourcentage de découpage '*test_size= 0.2*'. Voir figure4.5

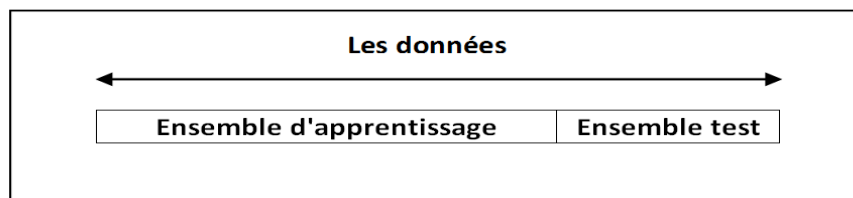


FIGURE 4.5: Répartition des données en un ensemble d'apprentissage et un ensemble test

Et elles retournent 8 variables ; quatre dans le cas 1 et quatre autres dans le cas 2 qui sont : X_{train} , y_{train} pour l'apprentissage et X_{test} , y_{test} pour le test final.

Validation (validation croisée)

Afin d'éviter le sur apprentissage, nous avons opté pour la validation croisée qui est similaire à *train_test_split()* , mais qui consiste à diviser nos données en k sous-ensembles différents. Nous utilisons $k - 1$ sous-ensembles pour entrainer nos données et laisser le dernier sous-ensemble comme données de test. Nous faisons ensuite la moyenne du modèle par rapport à chacun des sous-ensembles, puis finalisons notre modèle. Après cela, nous le testons par rapport à l'ensemble de test. La figure4.6 détaille ce schéma.

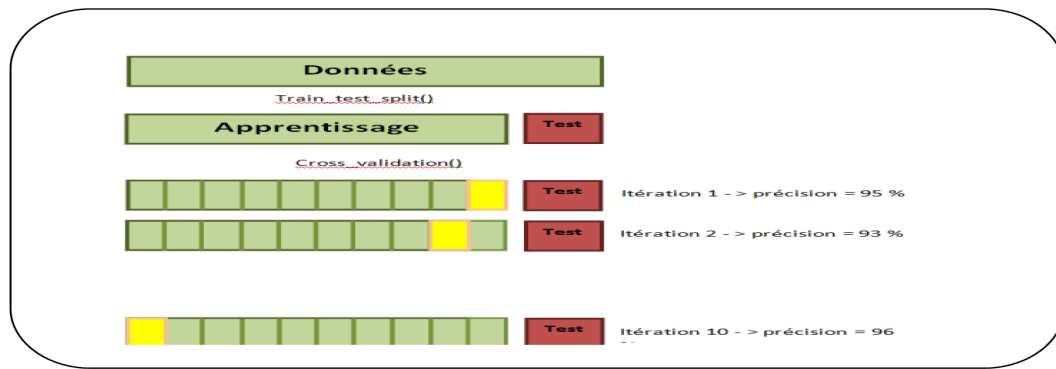


FIGURE 4.6: Schéma récapitulatif de la Validation croisée

Afin de déterminer la valeur améliorée des paramètres de l'algorithme d'apprentissage, nous avons mené plusieurs exécutions en faisant varier les valeurs de ces paramètres. Pour chaque expérience, la classification a été effectuée avec les mêmes ensembles obtenus en utilisant cinq méthodes différentes

1. Random Forest
2. K-plus proche voisin
3. Arbre de décision
4. SVM
5. Naives Bayes.

Random Forest (La forêt d'arbres aléatoires)

Les forêts d'arbres décisionnels[11] (ou forêts aléatoires de l'anglais *random forest classifier*) ont été premièrement proposées par Ho en 1995 [12] et ont été formellement proposées en 2001 par Leo Breiman[13] et Adele Cutler[14]. Elles font partie des techniques d'apprentissage automatique. Cet algorithme combine les concepts de sous-espaces aléatoires et de bagging. L'algorithme des forêts d'arbres décisionnels effectue un apprentissage sur de multiples arbres de décision entraînés sur des sous-ensembles de données légèrement différents.

Dans cette section, nous formerons notre classificateur de forêts d'arbres aléatoires en utilisant notre base de données *data_train*. Le résultat de précision de chaque itérations est donné dans la figure4.7 et a figure4.8

– **Cas 1 :**

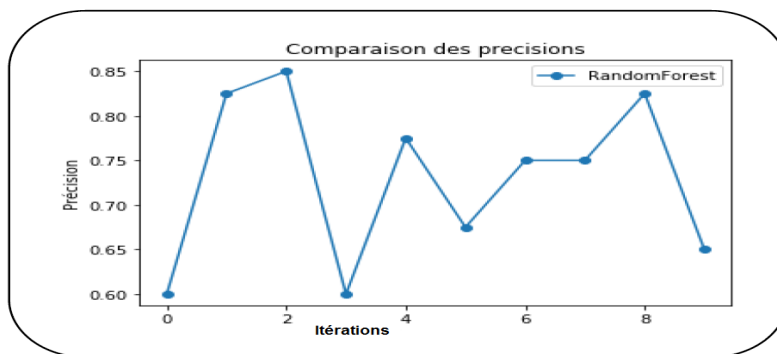


FIGURE 4.7: Représentation graphique de Random Forest : Cas1

Le graphe représente le taux de précision en fonction du nombre d’itérations. Nous remarquons que la meilleure itération est l’itération 2 avec un taux de précision de 85%.

– Cas 2 :

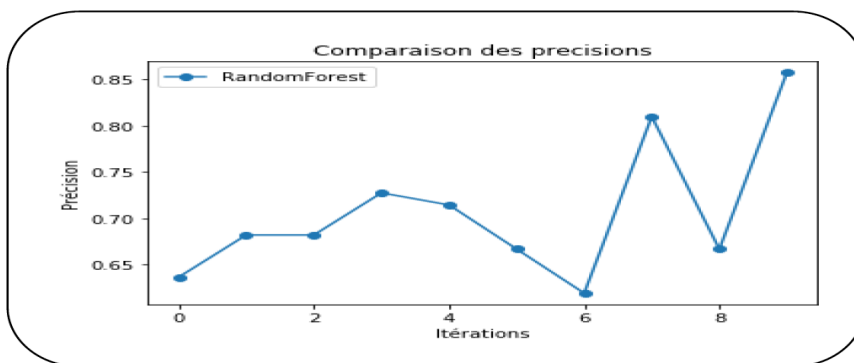


FIGURE 4.8: Représentation graphique de Random Forest : Cas2

Le graphe représente le taux de précision en fonction du nombre d’itérations. La meilleure itération est l’itération 9 avec un taux de précision de 85%.

Decision Tree Classifier (Arbre de décision)

L’apprentissage par arbre de décision désigne une méthode basée sur l’utilisation d’un arbre de décision comme modèle prédictif. On l’utilise notamment en fouille de données et en apprentissage automatique. Dans ces structures d’arbre, les feuilles représentent les valeurs de la variable-cible et les embranchements correspondent à des combinaisons de variables d’entrée qui mènent à ces valeurs. En analyse de décision, un arbre de décision peut être utilisé pour représenter de manière explicite les décisions réalisées et les processus qui les amènent. En apprentissage et en fouille de données, un arbre de décision décrit les données mais pas les décisions elles-mêmes, l’arbre serait utilisé comme point de départ au processus de décision.

C’est une technique d’apprentissage supervisé : on utilise un ensemble de données pour lesquelles on connaît la valeur de la variable-cible afin de construire l’arbre (données dites éti-

quetées), puis on extrapole les résultats à l'ensemble des données de test. Les arbres de décision font partie des algorithmes les plus populaires en apprentissage automatique [15, 16].

– Cas 1 :

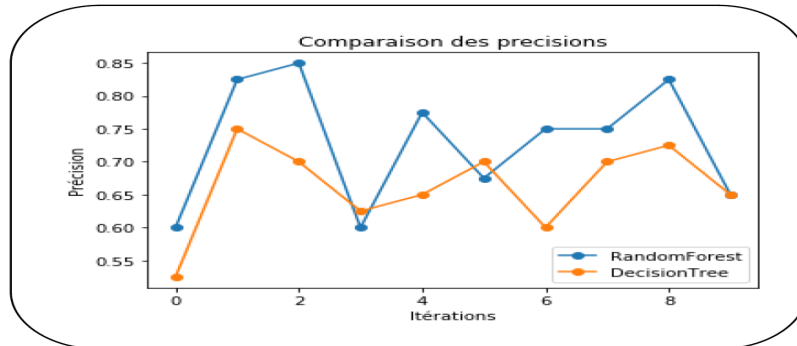


FIGURE 4.9: Comparaison graphique entre le l’algorithme Random Forest et Decision Tree Classifier : Cas1

Le graphe donnée dans la figure4.9 représente le taux de précision en fonction du nombre d’itérations pour les deux méthodes : Random Forest et Decision Tree Classifier. Nous remarquons qu’à nouveau la meilleure itération est celle de Random Forest avec un taux de précision de presque 85% comparé à celui de Decision Tree Classifier qui est de 75%.

– Cas 2 :

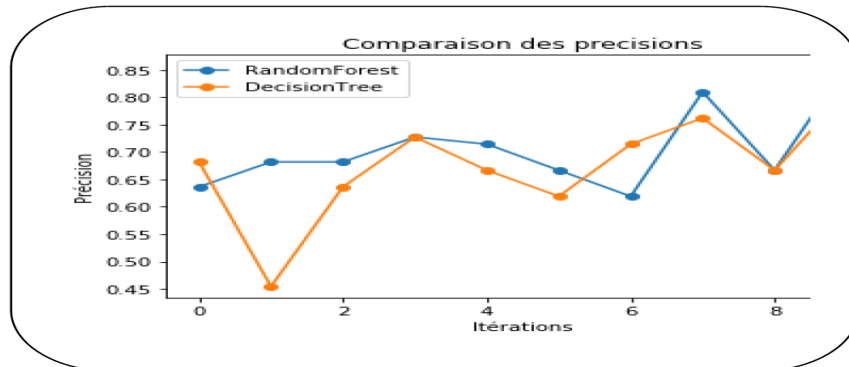


FIGURE 4.10: Comparaison graphique entre le l’algorithme Random Forest et Decision Tree Classifier : Cas2

Le graphe donnée dans la figure4.10 représente le taux de précision en fonction du nombre d’itérations pour les deux méthodes : Random Forest et Decision Tree Classifier. Nous remarquons qu’à nouveau la meilleure itération est celle de RandomForest avec un taux de précision de presque 85% comparé à celui de Decision Tree Classifier qui est de 80%.

K-Nearest Neighbors(K-plus proche voisin)

Dans ce cadre, on dispose d’une base de données d’apprentissage constituée de N couples "entrée-sortie". Pour estimer la sortie associée à une nouvelle entrée x , la méthode des k plus

proches voisins consiste à prendre en compte (de façon identique) les k échantillons d'apprentissage dont l'entrée est la plus proche de la nouvelle entrée x , selon une distance à définir. Puisque cet algorithme est basé sur la distance, la normalisation peut améliorer sa précision [17, 18].

– Cas 1 :

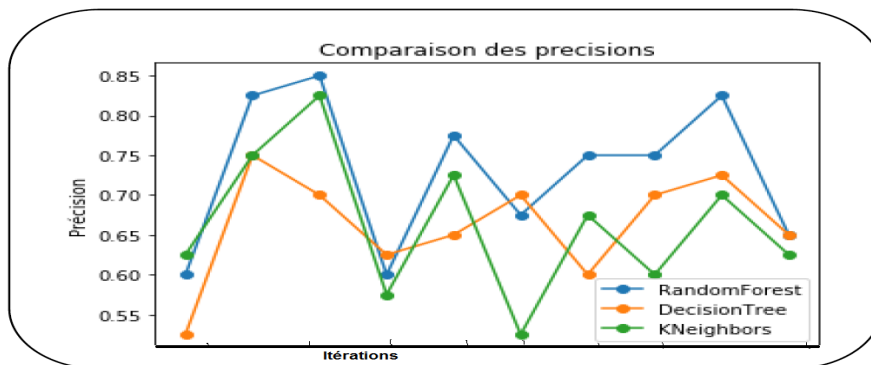


FIGURE 4.11: Représentation graphique des itérations en fonction du taux de précision : Cas1

Le graphe donnée dans la figure4.11 représente le taux de précision en fonction du nombre d'itérations pour les méthodes : Random Forest, Decision Tree Classifier et la méthode du K -plus proche voisin. L'algorithme Random Forest ayant encore le meilleur taux de précisions contrairement à l'algorithme K -plus proche voisin avec un taux de précision de 82.5%.

– Cas 2 :

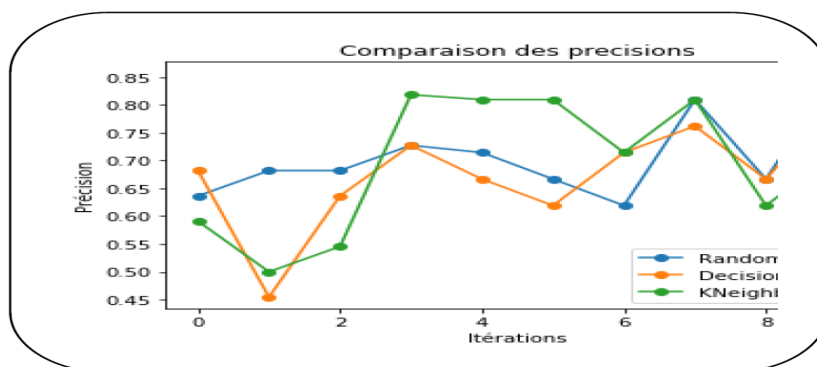


FIGURE 4.12: Représentation graphique des itérations en fonction du taux de précision : Cas2

Le graphe donnée dans la figure4.11 représente le taux de précision en fonction du nombre d'itérations pour les méthodes : Random Forest, Decision Tree Classifier et la méthode du K -plus proche voisin pour le cas 2. L'algorithme Random Forest ayant encore le meilleur taux de précisions précédé par l'algorithme K -plus proche voisin avec un taux de précision de 81% à l'itération 3.

Support Vector Machine (Machines à vecteurs de support SVM)

Les machines à vecteurs de support ou séparateurs à vaste marge (en anglais Support Vector Machine, SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de classification et de régression. Les SVM sont une généralisation des classifieurs linéaires. Les séparateurs à vastes marges reposent sur deux idées clés : la notion de marge maximale et la notion de fonction noyau. Ces deux notions existaient depuis plusieurs années avant qu'elles ne soient mises en commun pour construire les SVM [19].

– Cas 1 :

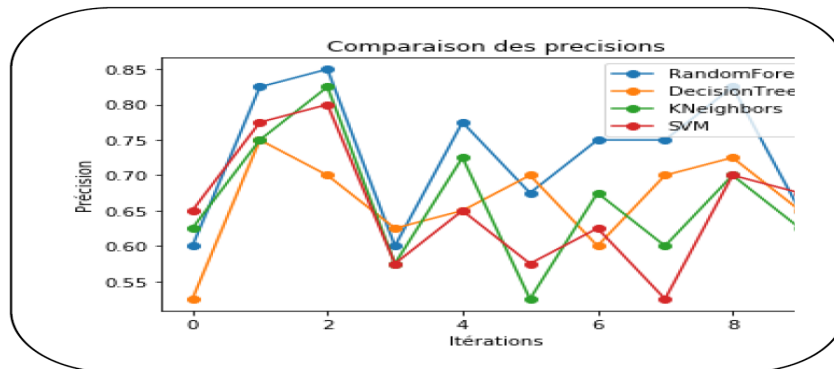


FIGURE 4.13: Représentation graphique des itérations en fonction du taux de précision : Cas1

Le graphe donnée dans la figure4.13 représente le taux de précision en fonction du nombre d'itérations pour les méthodes : Random Forest, Decision Tree Classifier, la méthode de *K*-plus proche voisin et SVM. La meilleure itération est toujours celle de l'algorithme RandomForest avec un taux de précision qui dépasse les 80% contrairement aux 3 autres qui ont tous le même taux de précisions le plus faible qui est égale à 52,5%.

– Cas 2 :

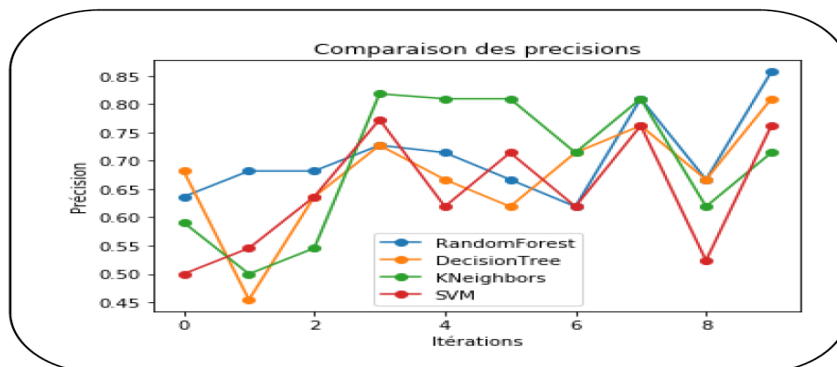


FIGURE 4.14: Représentation graphique des itérations en fonction du taux de précision : Cas2

Le graphe donnée dans la figure4.14 représente le taux de précision en fonction du nombre d'itérations pour les méthodes : Random Forest, Decision Tree Classifier, la méthode de

K-plus proche voisin pour le cas 2. L’algorithme Random Forest ayant le meilleur taux de précisions comparé à SVM qui a le taux de précision le plus faible à l’itération 1.

Naives Bayes(Classification naïve bayésienne GNV)

La classification naïve bayésienne est un type de classification bayésienne probabiliste simple basée sur le théorème de Bayes avec une forte indépendance (dite naïve) des hypothèses. Elle met en œuvre un classifieur bayésien naïf, ou classifieur naïf de Bayes, appartenant à la famille des classifieurs linéaires. Malgré leur modèle de conception « naïf » et ses hypothèses de base extrêmement simplistes, les classifieurs bayésiens naïfs ont fait preuve d’une efficacité plus que suffisante dans beaucoup de situations réelles complexes. En 2004, un article a montré qu’il existe des raisons théoriques derrière cette efficacité inattendue [20].

– Cas 1 :

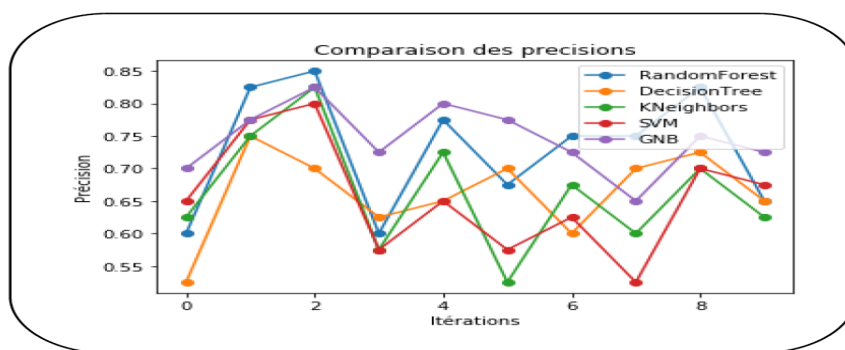


FIGURE 4.15: Représentation graphique des itérations en fonction du taux de précision : Cas 1

Le graphe donnée dans la figure4.15 représente le taux de précision en fonction du nombre d’itérations entres les quatre précédants algorithmes et celui de Naives Bayes pour le cas 1. Comme les précédentes comparaisons celui de Random Forest garde le meilleur taux de précisions sois 85%.

– Cas 2 :

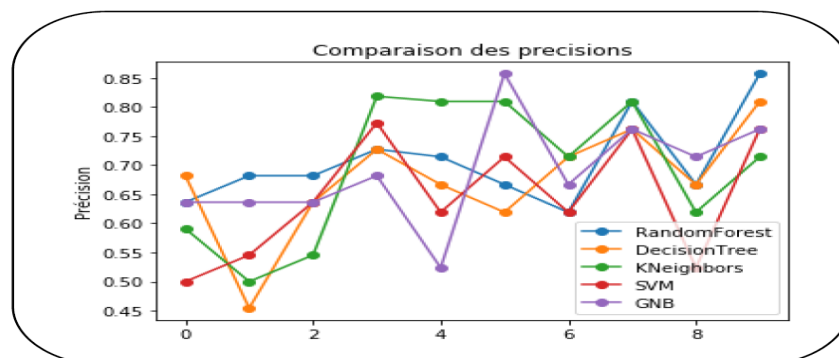


FIGURE 4.16: Représentation graphique des itérations en fonction du taux de précision : Cas2

Le graphe donnée dans la figure 4.14 représente le taux de précision en fonction du nombre d'itérations pour les quatre précédents algorithmes et celui de Naives Bayes pour le cas 1. Contrairement aux autres traitements, l'algorithme Naives Bayes est le meilleur algorithme avec un taux de précision de 86% à l'itération 5.

4.4 Discussion

Parmi les algorithmes que nous avons testé dans les deux précédents cas (les deux ensembles de données), à savoir celui du cabinet médical et du CHU, nous avons constaté que les algorithmes Naïves Bayes et Random Forest ont le plus grand taux de précision avec 86% pour le premier et 85% pour le deuxième.

Afin de choisir un des deux et obtenir des prédictions encore meilleures, trois techniques classiques sont envisageables :

1. Collecte de données : augmentez le nombre d'exemples d'apprentissage ;
2. Traitement des variables : ajoutez d'autres variables et un meilleur traitement des entités ;
3. Réglage des paramètres du modèle : envisagez d'autres valeurs pour les paramètres d'apprentissage utilisés par l'algorithme.

Nous avons opté pour le réglage des paramètres. Pour cela, nous ferons appel à de la bibliothèque *Scikit-learn* qui nous propose la fonction *RandomizedSearchCV* qui consiste à tester un ensemble de valeurs prédéfinies par l'utilisateur pour chaque paramètre, présenté sous forme de liste. La fonction retourne l'ensemble des paramètres du modèle avec les bonnes valeurs qui correspondent au meilleur taux de précision. Ces valeurs ont été utilisées dans l'apprentissage de nos modèles Random Forest et Naives Bayes une deuxième fois, et nous avons constaté un gain de précision de 16% pour l'algorithme Random Forest soit 87%, contrairement à celui de Naives Bayes où il n'y a eu aucune amélioration.

```
Précision de random Forest est : 0.8660606060606
Le resultat précédent est : 0.85060606060606
Le gain de precision est : 16.393939393939396 %
```

Ce modèle est sauvegardé sous format binaire avec la librairie Pickle, pour l'utiliser dans notre application web.

4.5 Application de prédiction web

Le résultat de notre travail, nous a permis de développer une application dans le but de permettre à différents individus de prédire si oui ou non ils peuvent être susceptibles de développer

un diabète de type 2. La page d'accueil ne nécessite pas une authentification. Elle permet à l'utilisateur d'introduire ses données personnelles, à savoir : son genre, l'âge, IMC, antécédents familiaux, groupe sanguin ainsi que la sédentarité. (Voir la figure 4.17)



FIGURE 4.17: Page d'accueil de l'application.

Pour l'implémentation du meilleur algorithme qui s'est avéré être celui de Random Forest, nous utiliserons dans notre interface principale un formulaire avec huit champs :

- Taille, Poids, Age de type float ;
- Pression artérielle et groupe sanguin de type sélection ;
- Genre, Sportif et sportif type booléen

Afin de traiter la requête : tous les champs doivent être remplis, sinon un message d'erreur s'affiche. (Voir la figure 4.18)

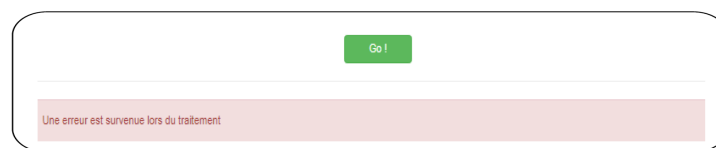


FIGURE 4.18: Message d'erreur.

Une fois, que la requête est validé, les informations du formulaire sont récupérées. Notre application fait appel au modèle (Random Forest) : sauvegarder en format binaire et en lui passant les données saisis par l'utilisateur. Notre modèle va nous retourner un résultat 0 pour non-malade et 1 pour malade. En plus des informations supplémentaires, le temps d'exécution et la précision. (Voir la figure 4.19)

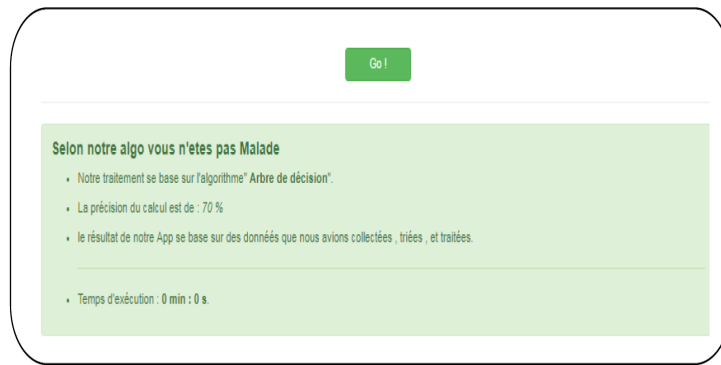


FIGURE 4.19: Résultat final de l'application.

4.6 Conclusion

Dans ce travail, nous avons mené une étude critique des travaux proposés sur la prédiction du diabète type 2. Pour ce faire, nous avons commencé par la définition des critères d'évaluation des différentes solutions existantes. Ensuite, nous avons fait une comparaison des travaux passés en revue, dans laquelle nous avons repris l'essentiel des avantages et inconvénients des travaux proposés. Pour finir par proposer notre méthode qui est une amélioration des travaux étudiés. Nous avons présenté une méthode pour la prédiction du diabète basée sur l'application des algorithmes d'apprentissage automatique supervisé. Nous avons testé les algorithmes d'apprentissage supervisé sur les deux différentes bases de données, à savoir celle du CHU de Béjaia et du cabinet privé du Dr Djamel MEHIDI. Par la suite, nous avons amélioré le meilleur algorithme d'apprentissage, à savoir RamdanForest, en termes de taux de précision, temps d'exécution. Le taux de classification obtenu avec notre méthode est parmi les meilleurs résultats obtenus pour la classification du diabète de type 2, par rapport aux autres algorithmes des travaux de l'état de l'art, soit un taux de précision de 86%.

En termes de perspectives, la prédiction du diabète à l'aide des méthodes d'apprentissage peut être élargie en utilisant les méthodes de base de connaissance pour augmenter l'interopérabilité du diagnostic.

Références

1. D. Hellmann, "The Python Standard Library by Example", Addison-Wesley, 2017.
2. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, "Scikit-learn : Machine learning in Python ", Journal of Machine Learning Research, vol. 12, pp. 2825 – 2830, ? 2011.
3. A. C. Müller and S. Guido, "Introduction to Machine Learning with Python", O'Reilly, 2016.
4. <https://www.spyder-ide.org/>
5. <https://software.opensuse.org/package/python-seaborn?locale=fr>
6. http://grasland.script.univ-paris-diderot.fr/STAT98/stat98_6/stat98_6.htm
7. I. D. Federation. Atlas du diabete de la fid neuvième édition 2019, 2019. URL <https://www.federationdesdiabetiques.org/>.
8. L. N. De Castro and F.J. Von Zuben, "Learning and Optimization Using the Clonal Selection Principle," IEEE Transactions on Evolutionary Computation ; Special Issue on Artificial Immune Systems, 2001.

9. D. Goodman, L. Boggess, and A. Watkins, "An Investigation into the Source of Power for AIRS ; an Artificial Immune Classification System," Proceedings of the International Joint Conference on Neural Networks (IJCNN'03)., pp. 1678 - 1683. 2003.
10. J.M. Keller, M.R. Gray, and J.A. Givens, "A fuzzy k-nearest neighbor algorithm," IEEE Transactions on Systems, Man and Cybernetics, SMC-15, pp.580 – 585,1985.
11. R. Nisbet, J. Elder, and G. Miner, "Handbook for Statistical Analysis And Data Mining", Academic Press, Page 247, Edition 2009.
12. Ho, T. Kam, " Random Decision Forests ", Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, ? 14-16 august 1995, pp. 278-282, 1995.
13. L. Breiman, "Random Forests", Machine Learning, vol. 45, no 1, ? pp. 5–32, 2001.
14. A. Liaw, "Documentation for R package randomForest" [archive], 16 octobre 2012, 2012.
15. X. Wu, V. Kumar, J. R. Quinlan and J. Ghosh, "Top 10 algorithms in data mining", Knowledge and Information Systems, vol. 14, no 1, ? pp. 1–37, 2008.
16. S. M. Pirayonesi and T. E. El-Diraby, "Data Analytics in Asset Management : Cost-Effective Prediction of the Pavement Condition Index", Journal of Infrastructure Systems, vol. 26, no 1, pp. 04019036, 2020.
17. S. M Pirayonesi and T. E. El-Diraby, "Role of Data Analytics in Infrastructure Asset Management : Overcoming Data Size and Quality Problems", Journal of Transportation Engineering, Part B : Pavements, vol.146, no 2, 2020.
18. T. Hastie, R. Tibshirani and J. Friedman, "The elements of statistical learning : data mining, inference, and prediction" , Springer, 2001.
19. B. Boser, I. Guyon, and V. Vapnik, "Pattern recognition system using support vectors". US Patent 5,649,068, 1997.
20. D. Pedro and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss", Machine Learning, vol. 29, pp. 103–137, 1997.

Conception d'un Système Multi-Agents d'Aide à la Décision pour une Gestion Efficace des Services d'Urgence : Cas de l'Hôpital de Kherrata

Boukredera D., Adel-Aissanou K., Abdoune K. et Bouriahi L.

ak_adel@yahoo.fr , djamila.boukredera@univ-bejaia.dz

Résumé Le surpeuplement des services d'urgence est un problème permanent et récurrent pour le personnel hospitalier, les administrateurs de santé, les décideurs politiques et les patients. Face à la demande croissante des patients pour ces services, aux restrictions budgétaires et au manque du personnel de santé, les administrateurs sont à la recherche de solutions pratiques et applicables afin de réduire le temps d'attente des patients et fournir une qualité de service acceptable. Dans cette optique, nous proposons un modèle décisionnel basé sur les systèmes multi-agents (SMA) dont le but est d'apporter une aide précieuse aux administrateurs en exploitant les avantages qu'offrent le paradigme agents qui est très préconisé pour modéliser des systèmes dynamiques et complexes où les entités qui le composent sont souvent appelés à coopérer, collaborer voire négocier pour atteindre un objectif tracé au préalable. Le modèle SMA proposé est un modèle de simulation automatique dont le but est de modéliser l'activité des urgences et d'analyser par la suite les effets de la réorganisation des ressources dans ces services sans apporter de changements potentiellement coûteux. Afin d'assurer une représentation fidèle du système réel, un réseau de files d'attente a été utilisé pour modéliser le parcours des patients dans le service d'urgence. Comme cas pratique, nous nous sommes intéressés au service des urgences de l'hôpital de Kherrata qui dispose d'une structure et d'une organisation qui ne peuvent plus fournir des prestations de soins adaptées aux besoins des patients et à leur satisfaction. Notre modèle est implémenté et testé dans la plateforme de simulation en libre accès NetLogo.

Mots clés : Services d'Urgence, Système Multi-Agents, Systemes d'Aide à la Décision, Systèmes de Files d'Attente, NetLogo.

5.1 Introduction

Les patients se rendent toujours dans un service d'urgence (SU) pour y recevoir des soins médicaux immédiats. Cependant, ils sont généralement confrontés à une très longue attente causée par différents paramètres notamment le manque de lits, de personnel de santé et la procédure de gestion non adéquate de ce type de services.

Améliorer la qualité et l'efficacité des soins de santé a été depuis longtemps l'enjeu majeur des structures hospitalières, et plus particulièrement les services des urgences. L'objectif est d'assurer une bonne prise en charge du patient tout en minimisant le temps d'attente et de séjour dans le service d'urgence. Ce temps et la qualité des soins sont les principales mesures de la satisfaction du patient. Ce problème est très complexe vu que les services d'urgence sont caractérisés par des activités non linéaires et en continuelle variation. C'est une organisation complexe avec des processus dynamiques et flexibles et de nombreux acteurs autonomes impliqués. De plus, les patients qui arrivent aléatoirement aux urgences présentent différents problèmes de santé avec différents niveaux de gravité. Les gestionnaires sont alors appelés à prendre des décisions souvent vitales nécessitant précision et rapidité afin d'éviter d'éventuels mauvais diagnostics ou traitements qui pourraient mener au décès du patient.

Les services d'urgence passent souvent par des phases critiques où le surpeuplement et la

manque de moyens matériel et humain entraîne inévitablement la dégradation des conditions de travail, l'allongement des temps d'attente et de séjour des patients ainsi que le risque de détérioration de leurs états de santé. Ceci provoque alors une baisse de la qualité des soins, l'insatisfaction des patients et l'inquiétude des gestionnaires face à de telles situations. Néanmoins, un service d'urgence aussi critique doit être satisfait avec la meilleure qualité possible le plus rapidement possible. Une solution simple et évidente à ce problème serait d'augmenter la capacité des urgences définie par la taille de l'établissement de santé et le personnel disponible, qui comprend des médecins, des infirmières, du personnel d'admission et de service. Toutefois, cette solution aussi simple n'est pas la meilleure approche et pourrait être irréalisable. Les responsables des systèmes de santé devraient alors trouver une meilleure réorganisation du service qui vise à maximiser l'utilisation des ressources de santé, afin de minimiser le temps d'attente des patients et d'augmenter leur satisfaction.

Cependant, les administrateurs des services d'urgence ne disposent généralement pas d'un outil d'aide à la décision pratique pour évaluer l'impact d'une modification du flux de travail menant à ces situations critiques. Dans cet article, nous proposons de venir en aide à ces gestionnaires en leur proposant un système d'aide à la décision basé sur un modèle multi-agents. C'est un modèle de simulation qui modélise les différentes phases du flux de travail des urgences, telles que le triage, l'examen du patient, son placement sous observation, la réanimation, etc.

Ce système d'aide à la décision a pour but d'aider les gestionnaires des urgences à répondre à la fois à des questions du type "Que se passerait-il si..." leur permettant de prendre les décisions les plus éclairées possibles, et à des questions plus complexes impliquant l'optimisation du système. Les gestionnaires des urgences seraient donc capables de mettre en place des stratégies et des directives de gestion pour réduire les délais de prise en charge du patient, améliorer le rendement de l'équipe soignante et améliorer, par conséquent, le service rendu au patient et sa satisfaction. Par ailleurs, nous proposons de modéliser le parcours du patient dans le service des urgences depuis son arrivée jusqu'à sa sortie par un réseau de files d'attente où les aspects aléatoires seront pris en considération.

Notre contribution dans ce travail de recherche est double :

- Développer un modèle de simulation basée sur des agents afin de permettre une exploration libre de la performance des urgences sous différents paramètres dans les situations de pré-saturation et de saturation du service d'urgence ;
- Etudier et analyser les performances du système en variant les paramètres notamment le nombre de lits disponibles et/ou l'ajout de médecins supplémentaires au moment opportun.

L'analyse de ces cas simulés est présentée pour illustrer comment les changements de ces paramètres ont un impact sur le temps de séjour du patient ainsi que les autres mesures de performance critiques.

Notre modèle a été implémenté dans la plateforme Netlogo qui est un environnement de modélisation programmable à base d'agents. Afin de procéder à des simulations avec des données

réelles, nous nous sommes intéressés au service des urgences de l'hôpital de Kherrata d'où nous avons collecté toutes les données nécessaires à l'évaluation des performances de notre modèle. Les indicateurs clés de performance considérés dans cette étude sont le délai moyen de séjour d'un patient au SU et le nombre moyen des patients évacués vers d'autres établissements. Les résultats de simulation obtenus sont très prometteurs et dénotent de l'efficacité du modèle proposé. Ils indiquent que la simulation basée sur des agents peut être utilisée par les administrateurs des urgences pour planifier les changements, localiser les points de blocage et étudier les relations non linéaires entre les différentes phases du flux de travail des urgences.

Le reste de cet article est organisé comme suit :

5.2 Travaux Antérieurs

La modélisation et la simulation axées sur les agents (Agent Based Modeling and Simulation, ABM&S) est une nouvelle approche de modélisation des systèmes complexes qui est efficace et bien utilisée grâce aux nombreux avantages qu'elles présentent notamment un niveau de détail accru dans les expériences basées sur la simulation, un processus d'apprentissage transparent, et la possibilité de contrôler et de modifier facilement le comportement individuel [10, 3].

Le principe de base de ABM&S n'est pas de concevoir des agents pour résoudre un problème donné mais plutôt d'agréger des règles simples de comportements individuels dans un environnement partagé pour expliquer un comportement collectif complexe émergent au niveau du groupe et qui n'est pas spécifiquement codé par le modélisateur et qui ne peut être prédit ou expliqué par les règles au niveau de l'agent. En fait, ABM&S a le potentiel de révéler un tout plus grand que la somme de ses parties [9]. C'est une sorte de modèle à micro-échelle [4] qui simule les opérations et les interactions simultanées de plusieurs agents dans une tentative de recréer et de prédire l'apparition de phénomènes émergents complexes. Un autre avantage important est que cette méthodologie permet de construire des modèles en l'absence de connaissances sur les interdépendances globales.

ABM&S est très utilisée dans divers domaines scientifiques et plus particulièrement en économie, biologie, les sciences sociales et le secteur de la santé. Bien que la simulation et la modélisation dans les établissements de santé ne soient pas nouvelles, la modélisation basée sur les agents dans ces environnements est relativement nouvelle et d'actualité. C'est dans ce contexte que s'inscrit notre travail qui se focalise principalement sur la gestion des services d'urgence dans un établissement hospitalier. Dans qui suit, nous allons présenter une revue de littérature qui se limitera aux travaux scientifiques qui ont trait à ABM&S dans le secteur de la santé.

Dans [13], l'auteur a proposé une simulation à base d'agents qui modélise diverses phases du flux de travail du service d'urgence, tels que le triage, l'examen par l'infirmier, l'examen

par le résident, l'examen du médecin traitant, etc. L'objectif de ce papier est de développer et de tester un modèle de simulation basé sur des agents pour étudier les performances sous différents contextes du processus de triage et de la procédure de radiologie. Les auteurs ont montré l'impact des changements du processus de triage et du processus de radiologie sur le temps de traitement des patients et les autres mesures de performance critiques de tout le système.

Les auteurs de [16], ont présenté une modélisation et simulation à base d'agents dans un service d'urgences. Ils ont proposé une nouvelle approche, où tous les agents du personnel vont participer au processus décisionnel pour réaffecter les ressources au service des urgences, en fonction de leurs observations dans leurs sections respectives. Grâce à leur approche auto-organisationnelle, les auteurs ont pu améliorer les indicateurs clés de performance de leur système.

Dans [8], les auteurs ont proposé une approche alternative où les médecins et leurs assistants aux urgences sont modélisés comme des pseudo-agents en interaction dans une simulation à événements discrets (DES). Cette approche est alors comparée avec l'approche traditionnelle en ignorant ces interactions. Les résultats obtenus ont montré qu'avec la modélisation des interactions, il y a une meilleure exploitation des médecins et leurs assistants. Cependant, Le temps d'interaction entre le médecin et l'assistant entraîne une augmentation de la durée du séjour aux urgences et des attentes plus longues pour les lits.

Dans [7], l'auteur a montré comment les procédures et les rôles des urgences peuvent être représentés à l'aide d'organisations d'agents, et comment les différentes valeurs et habitudes des membres du personnel peuvent être représentées par des pratiques sociales. L'auteur a développé un modèle simple d'une salle d'urgence dans la plateforme de simulation, où il a généré des contraintes organisationnelles en se basant les journaux d'événements collectés dans un hôpital réel.

les auteurs de [1], ont proposé d'utiliser la méthode algorithmique de recherche exhaustive (recherche par force brute) afin d'aider les chefs des services des urgences à améliorer le fonctionnement de ses services et de trouver la configuration optimale du service et du personnel.

Dans [12], les auteurs ont développé un modèle de simulation multi-méthodes qui émule le flux réel de patients dans le service des urgences dans des conditions de pandémie COVID-19. Deux types d'agents ont été principalement modélisés : les patients et les chambres. l'étude recommande plusieurs phénomènes à étudier et à inclure dans les futurs modèles de simulation, tels qu'un plus grand nombre d'agents (médecins, infirmières, lits), des retards dus aux interac-

tions avec d'autres services de l'hôpital et des changements dans le temps de traitement en cas d'occupation plus importante.

5.3 Modèle Conceptuel du Service des Urgences

Un service d'urgence typique peut être vu comme l'interaction de plusieurs éléments tels que des médecins, des infirmiers, des techniciens, des réceptionnistes, des lits, des dispositifs médicaux qui sont tous interconnectés par des flux d'informations et des processus à savoir l'enregistrement, le triage, le diagnostic, la sortie, etc. L'objectif de toute cette organisation est de fournir des diagnostics, des traitements et des informations.

Le service d'urgence étudié dans ce travail est celui de l'établissement hospitalier de Kherrata où nous avons réalisé une étude prospective afin de collecter toutes les informations nécessaires concernant les différents processus et les données historiques du service. Grâce à l'analyse de ces données, nous avons pu déterminer la nature des distributions statistiques suivies par chaque processus et leurs paramètres correspondants.

5.3.1 Description des Processus du Service des Urgences

Le service des urgences de l'hôpital de Kherrata est composé de six zones principales : la zone d'attente, la zone de consultation, la zone de triage, la zone de réanimation, la zone d'observation et la zone de soins. Un patient arrive au service d'urgence de deux manières : par ses propres moyens ou par une ambulance. Dans le premier cas, les patients entrent dans la zone de triage où ils sont classés par l'agent réceptionniste selon la gravité de leurs états. Nous distinguons trois niveaux de gravité : gravement urgent, urgent et non urgent. Les patients gravement urgents sont plus prioritaires que les autres patients et sont directement admis dans la salle de consultation si l'un des médecins est disponible, sinon ils rejoignent la salle d'attente. De même pour les patients urgents qui sont plus prioritaires que les patients non urgents.

Quant aux patients qui arrivent aux urgences par ambulance, leur triage se fait dans l'ambulance où nous distinguons deux types de patients : patients de type 1 qui regroupent toutes les personnes ayant été victimes d'un accident et les patients de type 2 qui sont tous les autres patients évacués pour une autre cause que l'accident. Les patients de type 1 ou 2 sont directement admis dans l'une des trois zones de traitement, à savoir la zone d'observation, la zone de soins ou la zone de réanimation. Dans le cas où il n'y a plus de lits disponibles dans les deux premières zones, les patients rejoignent la salle d'attente correspondante. Cependant, pour les patients nécessitant une réanimation, s'il n'y a plus de lits disponibles ils seront évacués vers d'autres établissements hospitaliers. Les patients dans les salles d'attente sont traités selon la politique premier venu premier servi dépendant du niveau d'acuité de ces patients.

Comme le montre la Figure 5.1, les patients qui arrivent par leurs propres moyens et après consultation peuvent quitter l'hôpital si leurs états de santé ne nécessitent pas de soins

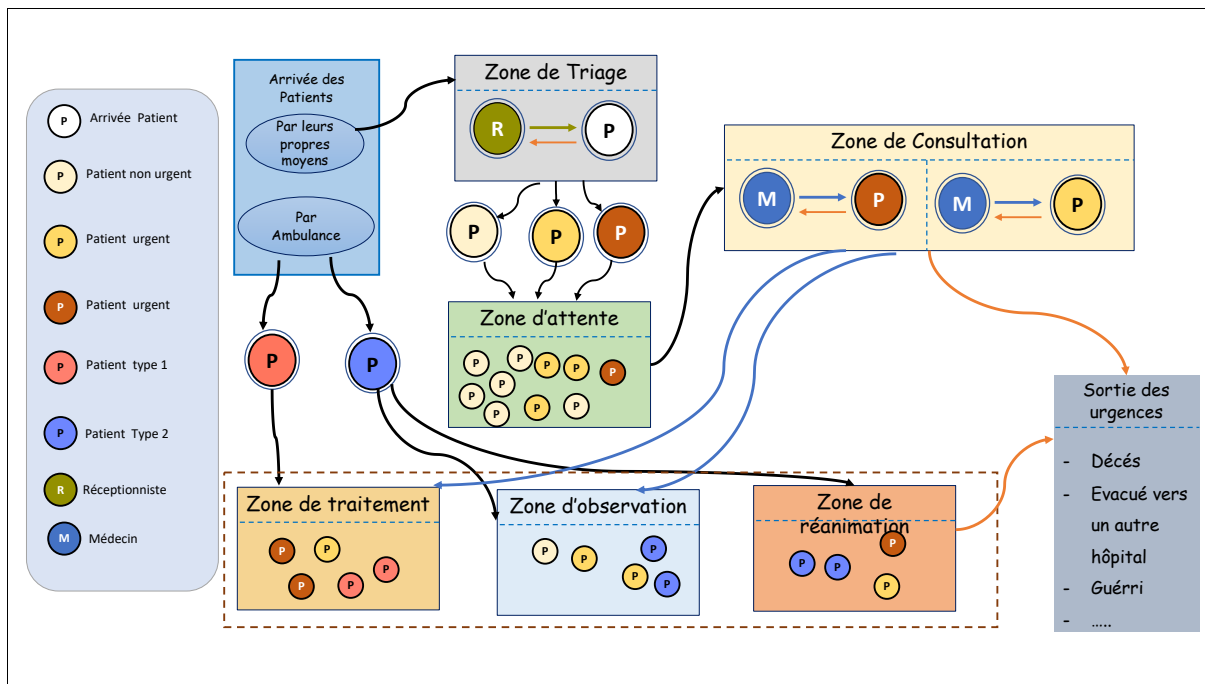


FIGURE 5.1: Flux des patients dans le service d'urgence de Kherrata

supplémentaires ou rejoindre l'une des trois zones de traitement selon la gravité de leurs cas diagnostiqués par le médecin. Il existe différentes issues possibles pour tout patient qui entre dans le SU : admission à l'hôpital, retour à la maison, transfert vers un autre hôpital ou décès.

Il est intéressant de noter que le processus décrit par la Figure 5.1 peut être vu naturellement comme un réseau de files d'attente avec un routage probabiliste.

5.3.2 Modèle de Files d'Attente

Le Service d'urgence peut être considéré comme un système où chaque patient qui franchit la porte est un inconnu qui pourrait suivre un chemin différent dicté de manière fonctionnellement non déterministe. En théorie, deux patients qui rentrent dans ce système n'auront pas deux chemins identiques. Par conséquent, du point de vue d'un patient, il reçoit un service ou attend que des ressources (personnel de santé ou ressources physiques tels que des lits ou des équipements de test) soient disponibles. Ce système tel qu'il est décrit n'est rien d'autre qu'un système d'attente qui se définit par l'arrivée séquentielle d'un flux d'événements appelés "clients" réclamant un service. Une file est naturellement créée dans le cas où l'offre est inférieure à la demande. La théorie considère généralement le temps séparant l'arrivée des clients et la durée de service comme deux variables aléatoires [11]. Après être servi, le client rejoint une autre file d'attente ou quitte le système. Notons que les systèmes d'attente sont largement utilisés dans le domaine de l'industrie et de la santé notamment les services des urgences. Pour plus de détails, le lecteur peut se référer à la synthèse bibliographique présentée dans [2, 5].

Dans le contexte des services d'urgence, la même méthodologie est appliquée pour modéliser les files d'attente de nombreuses ressources différentes qu'elles soient physiques ou humaines. Le SU est donc vu comme un réseau de files d'attente où le client correspond au patient et le serveur au réceptionniste, au médecin ou aux lits d'hospitalisation. Dans ce qui suit, nous décrivons les caractéristiques de notre modèle de files d'attente :

- **Processus d'arrivée des patients** : Les arrivées sont aléatoires, individuelles et indépendantes qui suivent un processus de poisson de taux λ représentant le nombre moyen de patients qui arrivent pendant une unité de temps. Nous avons pris l'heure comme unité de temps.
- **Processus de service** : Il s'agit du temps de service, représentant la période entre deux services consécutifs, et la discipline adoptée de la file d'attente.
- **Temps de service** : La durée de service de taux μ représente le nombre moyen de clients servis pendant une unité de temps, donc $1/\mu$ est le temps moyen de service d'un patient. Il s'agit donc d'une file d'attente avec un processus de Markov en entrée et en sortie.
- **Discipline de la file** : Il s'agit d'une file avec une capacité infinie de patients. La discipline de service qui se base sur la priorité des patients est imposée sachant que les patients sont répartis en plusieurs classes (gravement-urgents, urgents, non-urgents) qui définissent ainsi des ordres de priorité différents (plus prioritaire, prioritaire, non prioritaire). Cependant, l'arrivée d'un patient prioritaire pendant le service d'un patient non prioritaire ne permet pas l'interruption de ce dernier ; c'est une politique non préemptive.

A chaque étape du processus, les patients peuvent décider de poursuivre le traitement ou de quitter les urgences sans être traités. Par ailleurs, pour les patients les plus prioritaires, le modèle suppose que ces patients ne quitteraient pas le service d'urgence sans recevoir de soins vu la gravité de leur état.

La Figure 5.2 montre le réseau de files d'attente qui décrit le service d'urgence de l'hôpital de Kherrata. Les différentes moyennes λ ont été calculées sur la base des données collectées mais n'ont pas fait l'objet d'ajustements. Nous distinguons les moyennes suivantes (voir Figure 5.2) :

- λ_1 : représente le nombre moyen des patients qui arrivent avec leurs propres moyens / heure.
- λ_2 : représente le nombre moyen des patients qui arrivent par ambulance et seront transférer vers la zone des soins / heure.
- λ_3 : représente le nombre moyen des patients qui arrivent par ambulance et seront transférer vers la zone d'observation / heure.
- λ_4 : représente le nombre moyen des patients qui arrivent par ambulance et seront transférer vers la zone de réanimation / heure.

Les temps moyens de service ont été aussi calculés et peuvent être résumés comme suit :

- S_1 : représente le temps moyen du service de triage ;

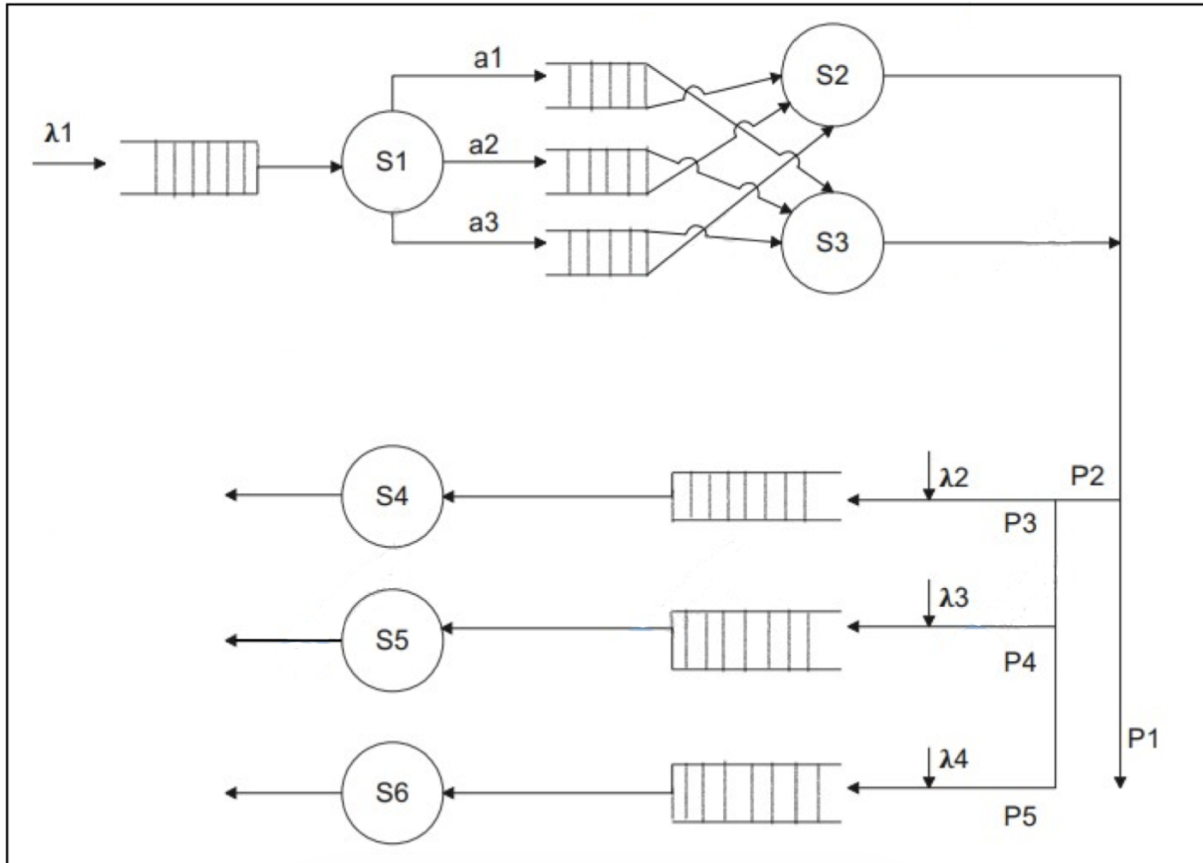


FIGURE 5.2: modèle de files d'attente des urgences de l'hôpital de Kherrata.

- $S2$: représente le temps moyen du service de consultation par le médecin 1 ;
- $S3$: représente le temps moyen du service de consultation par le médecin 2 ;
- $S4$: représente le temps moyen du service dans la zone de soins ;
- $S5$: représente le temps moyen du service dans la zone d'observation ;
- $S6$: représente le temps moyen du service dans la zone de réanimation.

Dans la Figure 5.2, $a1, a2$ et $a3$ représentent respectivement les probabilités qu'un patient soit classé dans la catégorie gravement-urgent, urgent ou non-urgent après l'étape de triage. Cependant, après la consultation, les patients sont soit libérés (avec une probabilité P) ou admis pour des traitements supplémentaires (avec une probabilité $P2$). Ces derniers peuvent être transférés vers la zone des soins (avec une probabilité $P3$), ou vers la zone d'observation (avec une probabilité $P4$) ou vers la zone de réanimation (avec une probabilité $P5$).

5.3.3 Description du Modèle Agents

Le modèle du service d'urgence défini dans ce travail est un modèle basé purement sur des agents. Ce sont des entités régies par des règles définissant le comportement des agents individuels qui peuplent le système. Grâce aux informations obtenues lors des entretiens menés avec le personnel médical et paramédical des urgences de l'hôpital de Kherrata, et après avoir défini le modèle conceptuel du service (voir Section 3), nous avons distingué trois types d'agents :

l'agent *Patient*, l'agent *Médecin*, l'agent *Réceptionniste* et l'agent *Chef de service*. Dans ce qui suit, nous détaillerons le rôle et le comportement de chacun ainsi que les interactions entre eux.

- **Agent Patient** : C'est un agent passif car sa progression dans le service d'urgence dépend largement de la disponibilité et de l'efficacité des prestataires de soins. Durant tout le processus dans les urgences, le patient alterne entre deux états : recevoir un traitement ou attendre. Le changement d'état se fait suite à son interaction avec un agent prestataire de soins. A cet effet, un agent patient qui arrive avec ses propres moyens interagit avec l'agent réceptionniste qui assure le processus de triage et classe ce patient dans l'une des trois catégories précédemment citées (non-urgent, urgent et gravement-urgents). Par contre, un agent patient qui arrive par ambulance est directement admis dans l'une des trois zones (traitement, observation, réanimation). Selon la disponibilité d'un médecin traitant et/ou d'un lit d'hospitalisation, l'agent patient peut se retrouver soit dans l'état en cours de traitement ou en attente de traitement.
- **Agent Réceptionniste** : L'agent réceptionniste est chargé d'accueillir les patients, d'interagir avec eux et de procéder à leur triage en les classant selon leurs états d'acuité dans l'une des trois catégories de priorité suscitées. Cet agent communique avec l'agent chef de service afin de lui fournir, en temps réel, toutes les informations nécessaires lui permettant de prendre des décisions bien éclairées notamment dans des situations critiques.
- **Agent Médecin** : Il est chargé de fournir un traitement aux patients, d'interagir avec eux et avec le personnel médical. Il est appelé à se déplacer aux différentes zones afin de prendre en charge tous les types de patients.
- **Agent Chef de service** : Il est responsable de la gestion du service d'urgence. Il interagit avec les autres agents du système afin de collecter les informations nécessaires lui permettant de gérer au mieux le service et de prendre les meilleures décisions en cas de problèmes. Son objectif est d'améliorer le fonctionnement du service en considérant les bons aménagements et en déterminant la configuration optimale du personnel.

Comme nous pouvons le remarquer, les modèles d'agents ne sont qu'un aspect d'un système de modélisation et de simulation basé sur les agents. Par ailleurs, le modèle d'interaction qui relie tous les agents pour former un système dynamique est un autre aspect. En effet, grâce à la communication entre les différents agents, l'information est véhiculée en temps réel aux agents décideurs leur permettant de prédire des situations critiques de pré-saturation voire de saturation du service. Ces informations sont très importantes lors du processus de prise de décisions.

5.4 Modèle d'Implémentation et Collecte de Données

Le modèle que nous avons proposé a été implémenté dans l'environnement de simulation NetLogo [14] qui est un langage de programmation et un environnement de modélisation intégrée pour le développement des systèmes multi-agents. Il est utilisé dans l'enseignement et dans le domaine professionnel sans nécessiter des connaissances en programmation pour modéliser des phénomènes liés [15, 6]. C'est un un modèle de simulation flexible qui peut être

facilement modifié et complété.

Notre objectif est d'implémenter le modèle proposé en reproduisant aussi fidèlement que possible la structure et le comportement du système réel. Pour ce faire, nous avons effectué un stage de deux mois au service d'urgence de l'hôpital de Kherrata afin de collecter minutieusement toutes les données nécessaires qui nous permettront, d'une part, de déterminer les données en entrée de notre système de simulation et d'autre part d'identifier les paramètres clés du modèle qui représentent les caractéristiques à ajuster lors des scénarios de simulation.

La Figure 5.3 montre une capture d'écran de notre modèle de simulation implémenté dans Netlogo.

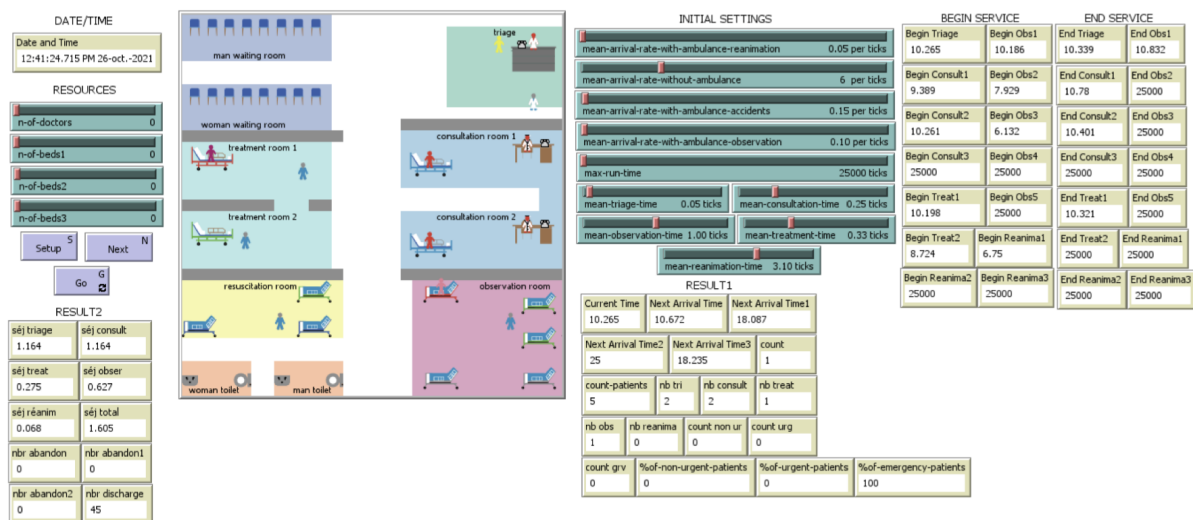


FIGURE 5.3: Implémentation de notre modèle de simulation dans Netlogo.

Collecte des données : La collecte de données représente la source de connaissances pour une meilleure compréhension du système complexe. Nous avons essayé de collecter autant de données que possible pour couvrir le comportement du système réel. Nous nous sommes intéressés particulièrement aux données concernant le nombre de patients arrivés au service des urgences, le taux d'arrivée des patients, les horaires de travail des médecins et des infirmiers, la durée de chaque traitement, etc.

Nous avons obtenu les données qui couvrent des informations sur notre système entre le 1er Mars 2021 et le 31 Juin 2021, en observant directement le système, en faisant des discussions et des entretiens avec le personnel du service d'urgence ainsi qu'à partir des registres des patients (registres de consultation, les registres d'accidents, etc.). La Figure 5.4 montre la base de données obtenue et enregistrée dans un classeur Excel à quatre feuilles représentant les 4 mois étudiés. Dans une autre feuille "TOTAL" on a calculé les différents λ de notre système.

Par ailleurs, nous avons aussi collecté les données concernant les taux de services et la durée de traitement de chaque zone, comme le montre la Figure 5.5.

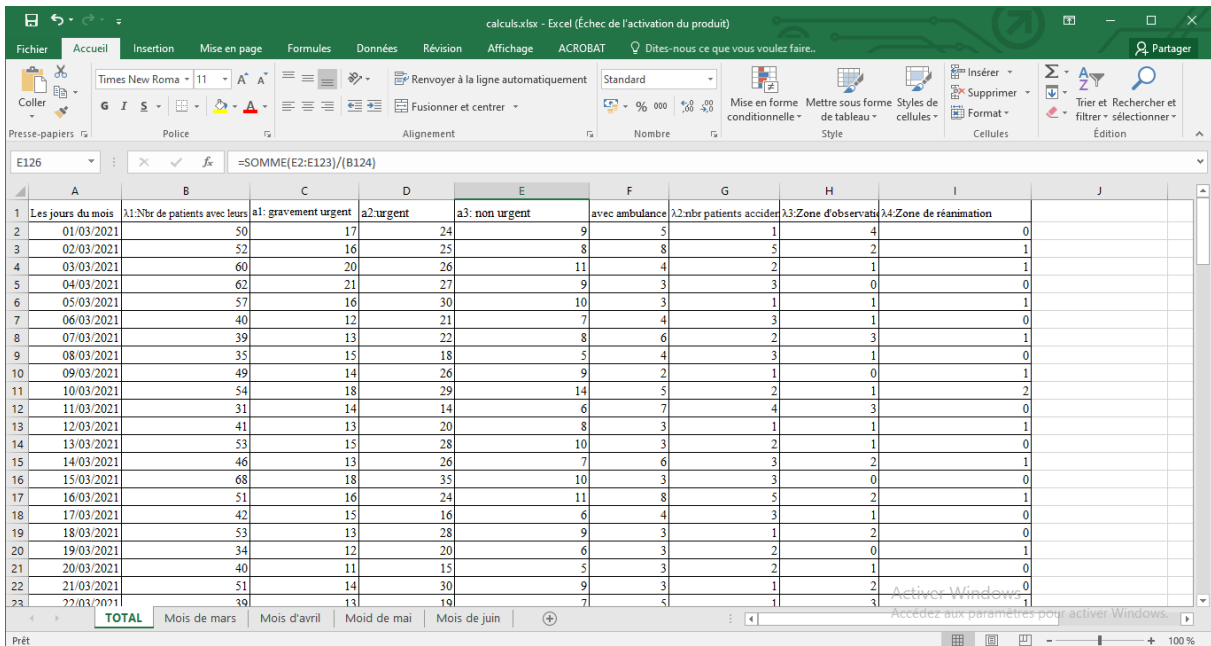


FIGURE 5.4: Données collectées sur les taux d'arrivées des patients.

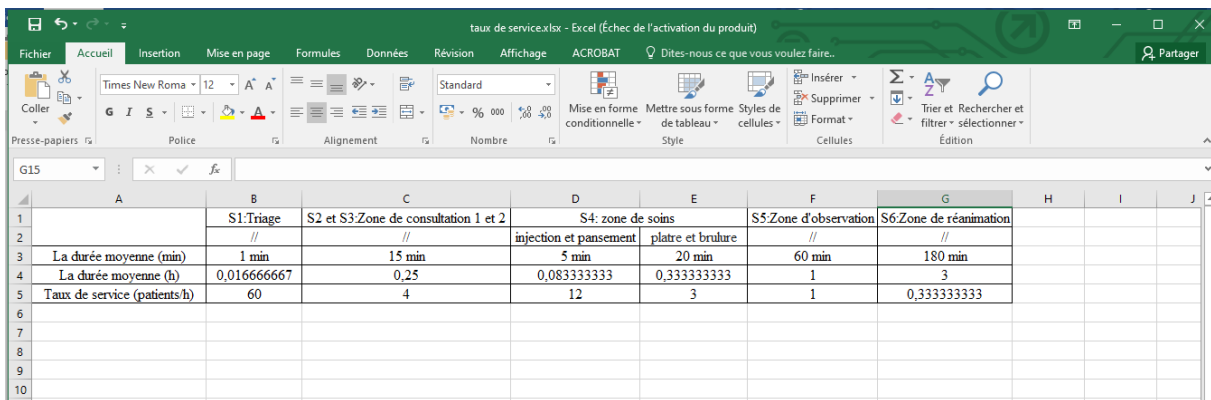


FIGURE 5.5: Données collectées sur les taux données de services des différentes zones.

A partir des données recueillies ainsi que les informations données par le personnel de l'hôpital, nous avons constaté que : 95,5% des patients arrivant au service des urgences avec leurs propres moyens, les 4.5% restant pour les patients qui arrivent soit par ambulances ou pompiers. Parmi les patients qui arrivent au S.U, 27% sont les patients gravement urgents, 54% sont les patients urgents, et 19% sont les patients non urgents.

5.5 Evaluation et Résultats

La vérification et la validation du modèle de simulation sont des étapes qui permettent d'évaluer sa robustesse et de vérifier sa capacité à reproduire les données observées dans le monde réel avant son utilisation pour le processus d'amélioration et d'optimisation. À cette fin, nous avons vérifié auprès des administrateurs que le modèle de simulation reflète la réalité avec un haut niveau de précision. Ensuite, nous nous sommes référés à trois indicateurs de perfor-

mances simulées pour nous permettre de comparer la solution actuelle avec les sorties obtenues après ajustement des paramètres du système.

5.5.1 Indicateurs Clés de Performance

Les résultats de simulation seront comparés avec les données du système réel afin d'évaluer la performance de notre système, en se basant sur les indicateurs clés de performance. Notre principal objectif est d'augmenter le nombre de patients traités ce qui revient à diminuer le temps moyen de séjour d'un patient ainsi que le nombre moyen des évacuations vers d'autres établissements de santé. Les indicateurs clés de performance de notre modèle sont alors :

- Délai moyen de séjour d'un patient au service.
- Nombre moyen des patients évacués vers d'autres établissements de santé.

5.5.2 Paramètres du Modèle

Dans notre modèle de simulation, les paramètres de notre système sont les différentes variables représentées par des curseurs Netlogo dans l'interface utilisateur graphique. Nous distinguons deux types de paramètres :

- Les paramètres concernant les temps moyens d'arrivée des différents types de patients ainsi que les temps moyens de service et le temps maximal pour la simulation. Les valeurs de ces paramètres seront variées de sorte à simuler les différents scénarios notamment lorsque le service d'urgence travaille dans des conditions normales (flux modéré des patients) ou des scénarios de crise ou de saturation, c'est-à-dire lorsque le service ne peut plus répondre à la demande (flux trop élevé des patients et/ou un manque de ressources). Ce type de paramètres est illustré par la Figure 5.6.
- Les paramètres concernant les ressources dont la variation des valeurs peut influencer positivement ou négativement la performance du système sous certaines contraintes opérationnelles. Dans notre cas, il s'agit du nombre de médecins en service et le nombre de lits d'hospitalisation disponibles pour les différentes zones du système. La Figure 5.7 montre ces paramètres comme des curseurs Netlogo dans l'interface graphique de l'utilisateur.

5.5.3 Amélioration de la Performance du Système

Afin d'apporter une aide précieuse aux décideurs notamment en situation de crise, nous allons procéder à la simulation de plusieurs scénarios où les résultats obtenus seront discutés par rapport aux différentes valeurs des paramètres et aux indicateurs de performance préalablement définis.

Méthodologie de Prise de Décision

Les outils d'aide à la décision sont utilisés dans différentes études liées à la simulation dans les services des urgences. La proposition de prise de décision par le chef de service a été utilisée pour réorganiser les urgences en fonction de différents critères. Dans ce processus, le nombre de médecins change selon le besoin mais leurs fonctions restent inchangées.

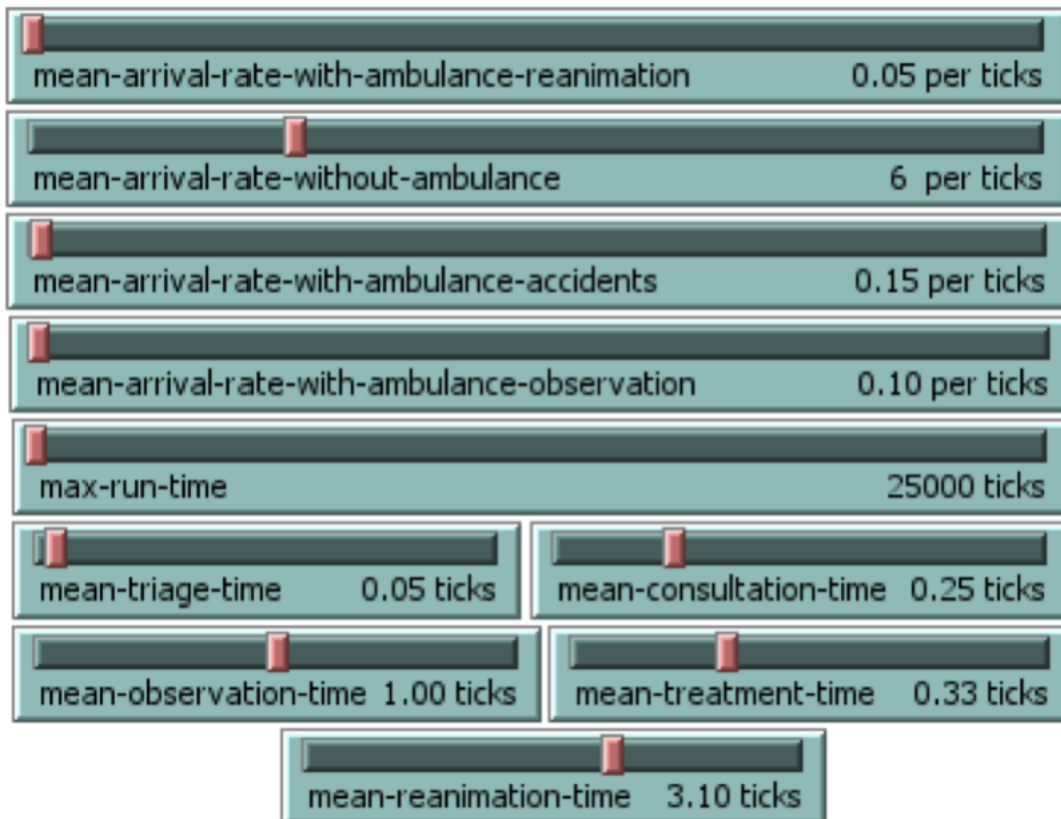


FIGURE 5.6: Paramètres concernant les différents temps moyens du système modélisé.

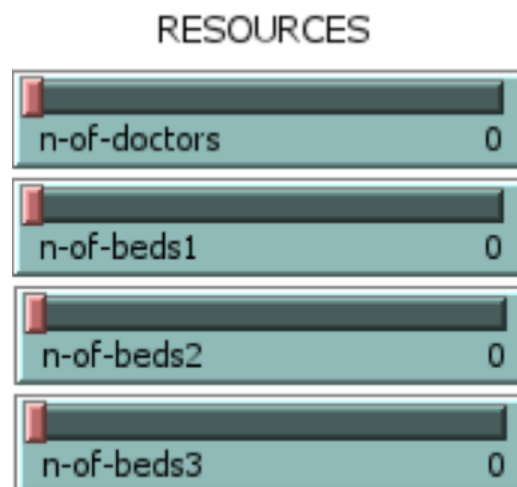


FIGURE 5.7: Paramètres concernant le nombre de ressources disponibles.

Communication entre les Agents du système

L'une des caractéristiques de la simulation à base d'agents (ABS) est la possibilité de faire communiquer les agents. Dans notre modèle, nous avons utilisé la communication 1-à-1, entre le chef de service et la réceptionniste pendant le processus de prise de décision. En cas de pré-saturation des urgences, la réceptionniste informe le chef de service en lui envoyant un message "attention pré-saturation", afin qu'il puisse prendre une décision pour améliorer la performance du système. Le chef de service peut décider soit de faire appel à un médecin disponible d'un

autre service, ou bien de faire évacuer les patients non urgents. En cas de saturation des urgences, la réceptionniste informe le chef de service en lui envoyant un message "attention saturation", afin qu'il puisse prendre une décision pour améliorer la performance du système. Le chef de service peut décider soit de faire appel à un médecin disponible d'un autre service ou de faire une évacuation totale des patients.

Les Scénarios Proposés

Afin d'évaluer les performances de notre système, en se basant sur les indicateurs cités précédemment, de nombreuses exécutions de simulation selon différents scénarios ont été conduites.

- **Scénario 1** : en cas de pré-saturation des urgences, l'agent réceptionniste envoie un message pour informer l'agent chef de service, qui devrait alors prendre une décision. Il va décider soit :
 - De faire appel à un troisième médecin disponible d'un autre service, ou
 - D'envoyer un message à la réceptionniste et lui demander d'évacuer les patients non urgents et de n'accepter dans le service que les patients gravement urgents et urgents.
- **Scénario 2** : en cas de saturation des urgences, le chef de service va décider de faire une évacuation totale des patients, en envoyant un message à la réceptionniste et lui demander de n'accepter aucun patient. Il peut faire aussi appel à un troisième médecin disponible d'un autre service pour la consultation.
- **Scénario 3** : Dans ce scénario, nous proposons de rajouter un quatrième lit dans la zone de réanimation, dans le but de réduire le nombre moyen des patients évacuées.
- **Scénario 4** : Dans ce scénario, nous proposons de combiner les trois scénarios précédents afin de réduire la durée de séjour moyenne ainsi que le nombre de patients évacuer.

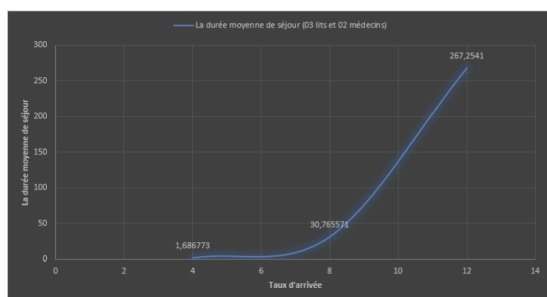
5.5.4 Résultats de la Simulation

Pour recueillir les résultats statistiques à partir de notre modèle, nous avons utilisé un outil de Netlogo appelé *BehaviorSpace*, cet outil exécutera automatiquement le modèle plusieurs fois avec plusieurs paramètres différents et collectera les résultats dans un fichier *csv*. Nous avons fait plusieurs expériences différentes, pour chaque valeur de λ (4, 8 et 12), dans chaque expérience nous avons recueilli les résultats de 50 répétitions.

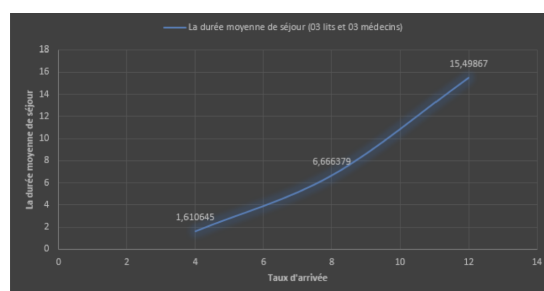
Durée Moyenne de Séjour

Les statistiques obtenues nous permettent de vérifier que la durée moyenne de séjour des patients augmente avec l'augmentation des taux d'arrivées des patients avec leurs propres moyens (λ). La Figure 5.8 illustre les résultats de la simulation de la durée moyenne de séjour d'un patient dans le service d'urgence de Kherrata avec différentes configurations. Les Figures 5.8a, 5.8b et 5.8c illustrent respectivement les résultats avec les configurations : (03 lits, 02 médecins), (03 lits, 03 médecins) et (04 lits, 03 médecins).

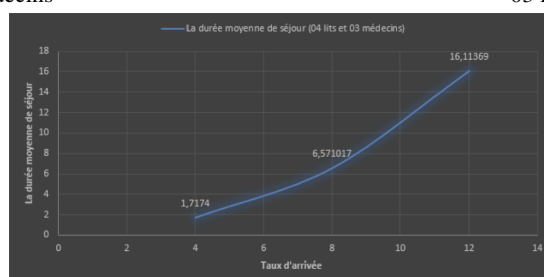
D'après les graphes de la Figure 5.8, nous remarquons que la durée moyenne de séjour des patients dépend amplement du taux d'arrivée λ . En effet, quand $\lambda = 12$ la durée moyenne de



03 lits et 02 médecins



03 lits et 03 médecins



04 lits et 03 médecins

FIGURE 5.8: Durée moyenne de séjour d'un patient dans le SU de Kherrata.

séjour est plus élevé que lorsque $\lambda = 8$. De même, cette durée est d'autant plus élevée avec $\lambda = 8$ que quand $\lambda = 4$.

Nous déduisons ainsi, que la durée moyenne de séjour des patients augmente avec l'augmentation des taux d'arrivées.

Nombre Moyen des Patients Évacués vers d'Autres Établissements Hospitaliers

Dans cette section, Nous allons présenter les résultats de la simulation obtenus pour le nombre moyen des patients évacués vers d'autres établissements hospitaliers. Cette situation survient dans le cas de pré-saturation et de saturation du service d'urgence.

durant les différentes simulations, nous avons fait varier les taux d'arrivée en affectant à λ trois valeurs distinctes, à savoir 4, 8 et 12. Le but est de reproduire les différentes situations que pourrait vivre le service d'urgence. De plus, nous avons fait varier les paramètres ressources du système, à savoir le nombre de lits et de médecins en service afin de simuler les différents scénarios proposés.

La Figure 5.9 illustre les résultats obtenus du nombre moyen des évacuations selon différentes configurations. Dans les Figures 5.9a, 5.9b et 5.9c tracent respectivement ces résultats selon les configurations (03 lits, 02 médecins), (04 lits, 02 médecins) et (04 lits, 03 médecins). Comme nous pouvons le remarquer, le nombre moyen des évacuations augmente avec l'augmentation du taux d'arrivée des patients. En effet, quand $\lambda = 12$ le nombre moyen des évacuations est plus élevé que lorsque $\lambda = 8$. De même, ce nombre est d'autant plus élevé avec $\lambda = 8$ que quand $\lambda = 4$.

Par ailleurs, nous remarquons que le nombre de lits d'hospitalisation disponibles ainsi que le nombre de médecins en service ont une grande influence sur les résultats obtenus. En

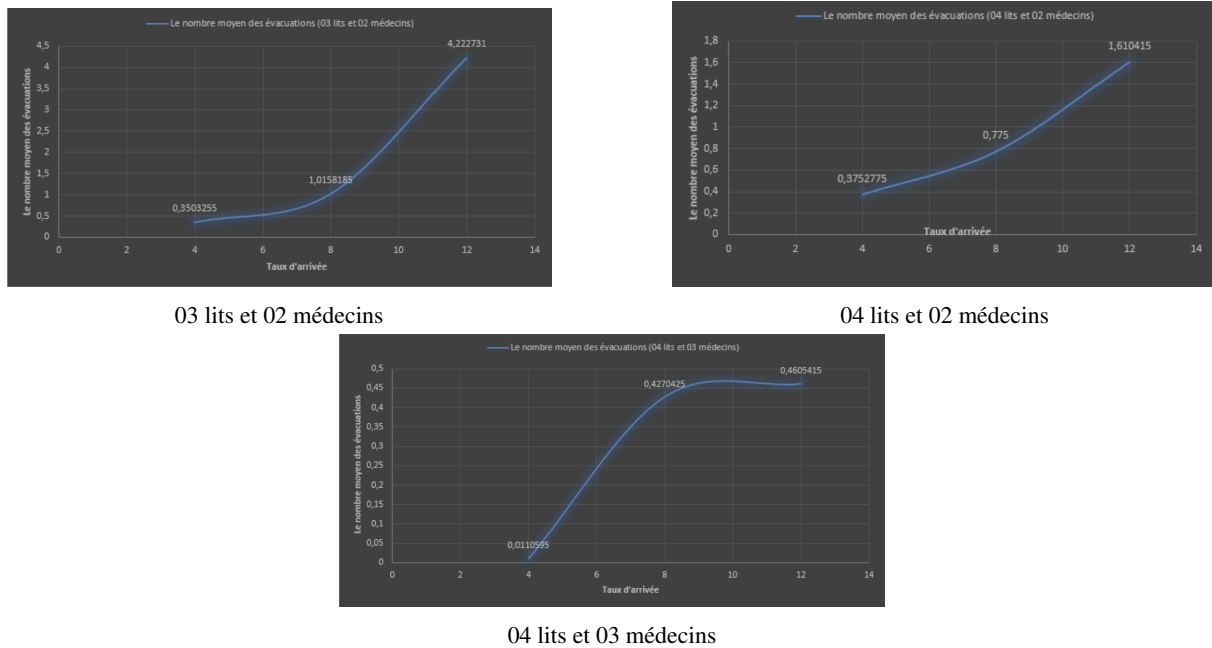


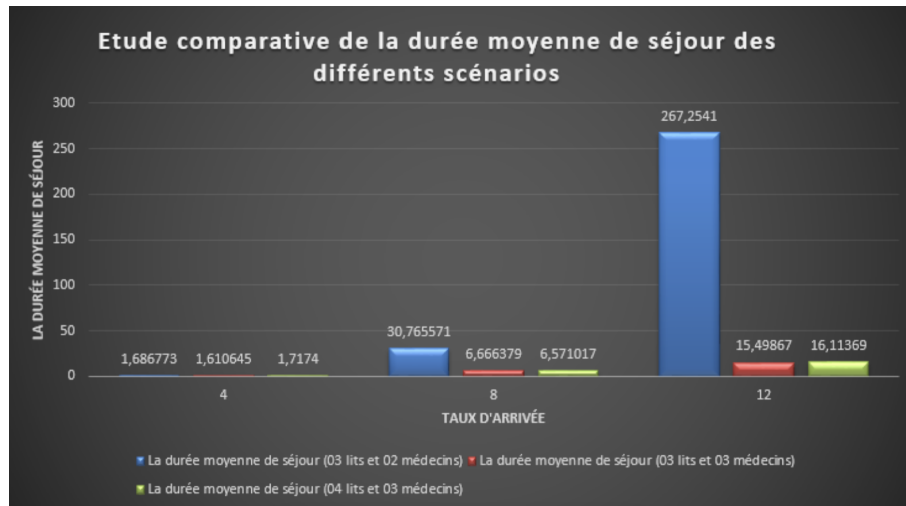
FIGURE 5.9: Nombre moyen des patients évacués vers d'autres établissements hospitaliers.

effet, en ajoutant un lit de réanimation, le nombre moyen des patients évacués diminue. Par contre, l'ajout d'un troisième médecin de manière temporaire n'affecte pas significativement le nombre moyen des évacuations.

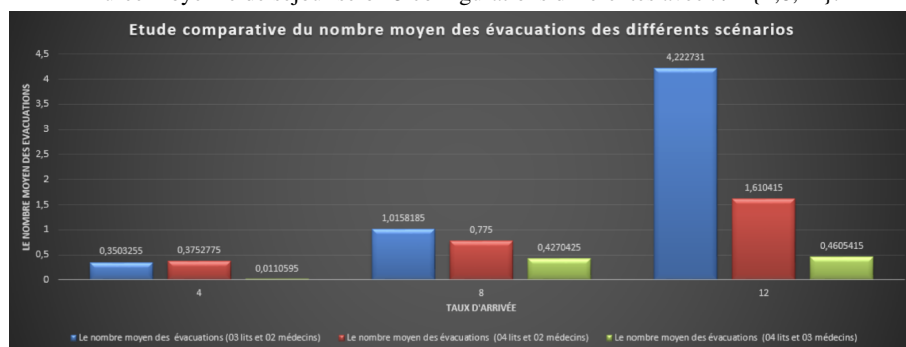
- **Etude Comparative** : Nous avons effectué une étude comparative afin de tester la performance des scénarios proposés. Pour ce faire, nous avons effectué des comparaisons des résultats de simulation concernant la durée moyenne de séjour et le nombre moyen des évacuations en fonction du taux d'arrivée. La Figure 5.10 illustre les résultats obtenus de la simulation des différents scénarios en faisant varier la valeur du taux d'arrivée λ concernant le temps de séjour (voir Figure 5.10a) et le nombre moyen des évacuations des patients (voir Figure 5.10b). Nous avons noté que les scénarios 1 et 2 ont contribué d'une manière significative à la diminution du temps moyen de séjour des patients, surtout lorsque λ dépasse 8 patients par heures. Cependant, nous avons remarqué qu'après l'exécution du scénario 3, le nombre moyen des évacuations a diminué par l'ajout d'un lit de réanimation. La combinaison des trois scénarios dans le scénario 4 a donné de meilleurs résultats à propos de la durée moyenne de séjour et le nombre moyen des évacuations.

5.6 Conclusion

Dans ce travail, nous avons proposé un système d'aide à la décision pour les gestionnaires des urgences de l'hôpital de Kherrata. L'objectif est d'améliorer le fonctionnement de leur service ainsi que la qualité et l'efficacité des soins fournis. Pour ce faire, nous avons proposé une approche basée sur la combinaison de la modélisation et la simulation à base d'agents (ABMS). Notre modèle de simulation à base d'agents a été implémenté et vérifié dans un environnement



Durée moyenne de séjour selon 3 configurations différentes avec $\lambda = \{4,8,12\}$.



Nombre moyen des évacuations selon 3 configurations différentes avec $\lambda = \{4,8,12\}$.

FIGURE 5.10: Résultats des scénarios proposés.

de modélisation programmable multi-agent, NetLogo. Les expériences ont été réalisées à l'aide des simulations parallèles avec différents paramètres et deux indices de performance différents à savoir le temps moyen de séjour et le nombre moyen des évacuations.

Les résultats des simulation des différents scénarios que nous avons proposés, ont montré que le temps de séjour moyen et le nombre moyen des évacuations ont été considérablement améliorées. Notons que la simulation à base d'agents proposée et la méthode de prise de décision peuvent être facilement mises en œuvre dans n'importe quel service d'urgence après quelques adaptations.

Comme perspective, nous proposons d'améliorer notre outil d'aide à la décision en exploitant les avantages qu'offrent les SMA, pour modéliser des entités pouvant interagir et communiquer dans un environnement bien défini, offrant la possibilité d'étudier la dynamique des systèmes complexes.

Références

1. Eduardo Cabrera, Manel Taboada, Ma Luisa Iglesias, Francisco Epelde, and Emilio Luque. Optimization of healthcare emergency departments by agent-based simulation. *Procedia computer science*, 4 :1880–1889, 2011.
2. Amir Elalouf and Guy Wachtel. Queuing problems in emergency departments : a review of practical approaches and research methodologies. In *Operations Research Forum*, volume 3, pages 1–46. Springer, 2022.

3. Marcia R Friesen and Robert D McLeod. A survey of agent-based modeling of hospital environments. IEEE Access, 2 :227–233, 2014.
4. Leif Gustafsson and Mikael Sternad. Consistent micro, macro and state-based population modelling. Mathematical biosciences, 225(2) :94–107, 2010.
5. Joshua W Joseph. Queuing theory and modeling emergency department resource utilization. Emergency Medicine Clinics, 38(3) :563–572, 2020.
6. Daniel Kornhauser, William Rand, and Uri Wilensky. Visualization tools for agent-based modeling in netlogo. Proceedings of Agent2007, pages 15–17, 2007.
7. John Bruntse Larsen, J Villadsen, N Carstens, CK Holst, N Gierasimczuk, M Rosendahl, and J Dix. Hospital Staff Planning with Multi-Agent Goals. PhD thesis, PhD thesis, Department of Applied Mathematics and Computer Science..., 2019.
8. Morgan E Lim, Andrew Worster, Ron Goeree, and Jean-Éric Tarride. Simulating an emergency department : the importance of modeling the interactions between physicians and delegates in a discrete event simulation. BMC medical informatics and decision making, 13(1) :1–11, 2013.
9. Charles M Macal. Everything you need to know about agent-based modelling and simulation. Journal of Simulation, 10(2) :144–156, 2016.
10. Muaz Niazi and Amir Hussain. Agent-based computing from multi-agent systems to agent-based models : a visual survey. Scientometrics, 89(2) :479–499, 2011.
11. Philippe Robert. Réseaux et files d’attente : méthodes probabilistes, volume 35. Springer Science & Business Media, 2000.
12. Gaute Terning, Eric Christian Brun, and Idriss El-Thalji. Modeling patient flow in an emergency department under covid-19 pandemic conditions : A hybrid modeling approach. In Healthcare, volume 10, page 840. Multidisciplinary Digital Publishing Institute, 2022.
13. Lu Wang. An agent-based simulation for workflow in emergency department. In 2009 Systems and Information Engineering Design Symposium, pages 19–23. IEEE, 2009.
14. U Wilensky. Netlogo< http ://ccl. northwestern. edu/netlogo/>, center for connected learning and computer-based modeling, northwestern university. Evanston, IL, 1999.
15. Uri Wilensky and William Rand. An introduction to agent-based modeling : modeling natural, social, and engineered complex systems with NetLogo. Mit Press, 2015.
16. M Yousefi and RPM Ferreira. An agent-based simulation combined with group decision-making technique for improving the performance of an emergency department. Brazilian journal of medical and biological research, 50, 2017.

Deuxième partie

Autres domaines d'application

I - Théorie des jeux

Jeux Bayésiens et applications

Bouibed K.

Research Unit LaMOS (Modeling and Optimization of Systems) and Faculty of Exact Sciences, Bejaia University, Bejaia 06000, Algeria

Résumé Jeux Bayésiens (également connus sous le nom des jeux à informations incomplètes) sont des modèles de situations de décisions interactives où les décideurs (joueurs) n'ont que des informations partielles sur les données de jeu et sur les autres joueurs. Ce travail présente quelques applications des jeux Bayésiens dans les réseaux informatiques, notamment l'interaction entre les fournisseurs de services (les serveurs) et les clients inconnus et la modélisation de problèmes de sécurité dans les réseaux Ad Hoc.

Mots clés : Jeux Bayésiens, Jeux de sécurités, Reseaux Ad Hoc.

6.1 Définition d'un jeu Bayésien

Definition 1 [2] *Un jeu statique Bayésien est donné par :*

$$(J_B) = \langle I, (A_i)_i, (\Theta_i)_i, (f_i)_i, P \rangle, \quad (6.1)$$

- avec $I = \{1, \dots, i, \dots, n\}$ est l'ensemble des joueurs,
- Θ_i est l'ensemble de types du joueur i , le type $\theta_i \in \Theta_i$ (appelé aussi caractéristique, état d'information,...) résume l'information dont dispose le joueur i avant de jouer. $\theta = (\theta_1, \dots, \theta_i, \dots, \theta_n)$ de $\Theta = \prod_{i=1}^n \Theta_i$ représente donc l'information totale dont disposeraient les joueurs s'ils la mettaient en commun.
- Une distribution de probabilité a priori P sur Θ .
- L'ensemble des actions A_i , la fonction d'utilité (gain) f_i définie sur $A \times \Theta$ avec $f_i(a, \theta)$ est le niveau d'utilité atteint par le joueur i si les actions $a = (a_1, \dots, a_i, \dots, a_n)$ ont été choisies et si les types sont $\theta = (\theta_1, \dots, \theta_i, \dots, \theta_n)$.

Un jeu Bayésien est la donnée d'une famille des jeux paramétrés par les types des joueurs et une distribution a priori sur ces types lors du jeu. Une valeur précise de θ est réalisée mais en général les joueurs ne savent pas exactement laquelle.

Ces jeux tiennent leur nom de l'application de la règle de Bayes. Ainsi, les jeux à information incomplète sont de ce fait transformés en des jeux à information imparfaite.

6.1.1 Déroulement du jeu selon Harsanyi

Harsanyi a considéré que la caractéristique θ_i est la réalisation d'une variable aléatoire et représente le résultat du choix d'un nouvel agent (joueur) appelé Nature.

D'abord, la nature fait un tirage de tous les types $\theta_1, \dots, \theta_i, \dots, \theta_n$ selon la distribution commune

$P(\theta_1, \dots, \theta_i, \dots, \theta_n)$ et le type tiré θ_i est communiqué au joueur i de manière confidentielle. Ensuite, chaque joueur choisit son action en fonction de son type et de la probabilité.

Avec cette approche on définit un jeu Bayésien qui se déroule comme suit :

1. La nature procède au tirage aléatoire du vecteur $\theta = (\theta_1, \dots, \theta_i, \dots, \theta_n)$ par la fonction de distribution jointe à partir de laquelle le tirage est effectué. $P(\theta)$ est une connaissance commune à tous les joueurs ;
2. La nature révèle le type θ_i au seul joueur i . En connaissant P et θ_i , chaque joueur i peut en déduire des croyances $P_i(\theta_{-i} / \theta_i)$ sur la valeur θ_{-i} du type des autres joueurs en faisant appel à la règle de Bayes :

$$P_i(\theta_{-i} / \theta_i) = \frac{P(\theta_{-i}, \theta_i)}{P(\theta_i)} = \frac{P(\theta_{-i}, \theta_i)}{\sum_{\theta_{-i}} P(\theta_{-i}, \theta_i)}, \quad (P(\theta_i) \neq 0);$$

3. Chaque joueur choisit une action de son espace A_i , ce choix se fait en fonction du type θ_i ;
4. Chaque joueur i reçoit un paiement qui dépend non seulement du profil de stratégies pures $x = (x_1, \dots, x_i, \dots, x_n)$ mais également de sa caractéristique privée $\theta_i \in \Theta_i$ elle prend donc la forme $f_i(x_i, x_{-i}, \theta_i)$.

6.2 Application des jeux Bayésiens à la sécurité des réseaux informatiques

L'interaction entre le fournisseur de services (le serveur) et un client inconnu est modélisée par un jeu Bayésien dynamique à deux joueurs. En considérant deux types de clients, les clients malveillants et les clients normaux. Le jeu Bayésien proposé (appelé jeu de sécurité (JS)) à deux joueurs : un serveur s et un client c [1].

On considère deux types ($\theta_c = 0$, désigne un client régulier et $\theta_c = 1$, désigne un client malveillant où un attaquant) pour les clients de serveur, le type de serveur est toujours $\theta_s = 0$, le joueur s a deux stratégies (défendre et ne pas défendre) et le joueur c a aussi deux stratégies (attaquer et ne pas attaquer). La nature des communications entre le serveur et le client est répétée et jouée aux temps $t = 0, 1, 2, \dots, T$. Différents paquets sont envoyés par le client au serveur et vice versa. Chaque paquet et les réponses peuvent être considérés comme une étape du jeu de sécurité (JS). Les gains des deux joueurs sont donnés comme suit

(a) Le joueur c est régulier

		Joueur c	
		Ne pas attaquer	Attaquer
Joueur s	Ne pas défendre	(0, 0)	(0, 0)
	Défendre	$(-\alpha, -\beta)$	$(-\alpha, -\beta)$

(b) Le joueur c est un attaquant

		Joueur c	
		Ne pas attaquer	Attaquer
Joueur s	Ne pas défendre	$(-G', G' - \tau')$	$(-G, G - \tau)$
	Défendre	$(g' - G' - \alpha, G' - g' - \tau')$	$(g - G - \alpha, G - g - \tau)$

Avec α est le coût de la stratégie défendre, on quantifier la dégradation de service par β . τ, τ' les coûts de l'attaquant lorsque il joue respectivement attaquer et ne pas attaquer avec $\tau' \geq \tau$.

G (resp G') est l'information acquise par l'attaquant lorsque il joue attaquer (resp. ne pas attaquer) avec $G \geq G'$, g (resp g') la prévention de la fuite de l'information lorsque l'attaquant joue la stratégie attaquer (resp. ne pas attaquer) avec $g \geq g'$.

L'objectif de l'analyse du jeu de sécurité (JS) est de proposer une probabilité optimale de jouer la stratégie défendre pour le serveur où il est incertain sur le type de son client.

6.3 Jeux Bayésiens pour la détection d'intrusion dans les réseaux mobile Ad Hoc

On considère un réseau Ad Hoc avec N noeud. Chaque noeud défenseur est équipé d'un système de détection d'intrusion (IDS). Le joueur i est un noeud attaquant et le joueur j est un noeud défenseur. Le joueur i a des informations privées sur son type, soit régulier, noté par $\theta_i = 0$, ou malveillant, noté par $\theta_i = 1$. Le défenseur j est du type régulier $\theta_j = 0$. Le type du défenseur j est une connaissance commune des deux joueurs. Le joueur i du type malveillant a deux stratégies pures : attaquer et ne pas attaquer. Le joueur i du type régulier a une stratégie pure : ne pas attaquer. Le joueur j a deux stratégies pures : surveiller et ne pas surveiller [3]. Les gains des deux joueurs sont donnés comme suit

(a) Le joueur i est malveillant

		joueur j	
		Surveiller	Ne pas surveiller
joueur i	Attaquer	$((1 - 2\alpha)w - c_a, (2\alpha - 1)w - c_s)$	$(w - c_a, -w)$
	Ne pas attaquer	$(0, -\beta w - c_s)$	$(0,0)$

(b) Le joueur i est régulier

		joueur j	
		Surveiller	Ne pas surveiller
joueur i	ne pas attaquer	$(0, -\beta w - c_s)$	$(0,0)$

– où w la valeur de sécurité du défenseur j , $-w$ représente une perte de sécurité dont la valeur est équivalente à un degré de dommages.

- α représente le taux de détection de l'IDS, β représente la fausse alarme de l'IDS avec $\alpha, \beta \in [0, 1]$.
- Les coûts des attaques et des surveillances sont indiqués respectivement par c_a et c_s , où $c_a, c_s > 0$ avec $w > c_a, c_s$.
- Soit μ_0 la croyance du joueur j sur le type malveillant du joueur i .

Supposons que μ_0 est une connaissance commune, si la croyance du défenseur j sur le joueur i malveillant est assez élevée : $\mu_0 > \frac{\beta w + c_s}{(2\alpha + \beta)w}$, le jeu Bayésien statique n'a pas d'équilibre en stratégies pures. Mais il possède un équilibre en stratégies mixte ($(p^*$, si i est malveillant, ne pas attaquer si i régulier), q^* , μ_0), avec $p^* = \frac{\beta w + c_s}{(2\alpha + \beta)w\mu_0}$ est la probabilité que le noeud malveillant i joue la stratégie attaquer et $q^* = \frac{w - c_a}{2\alpha}$ est la probabilité que le noeud défenseur j joue surveiller. Si la croyance du défenseur j sur le joueur i malveillant est très faible $\mu_0 < \frac{\beta w + c_s}{(2\alpha + \beta)w}$, un équilibre en stratégies pures existe ((attaquer si malveillant, ne pas attaquer si régulière, ne pas surveiller, μ_0)).

L'avantage d'utiliser un modèle du jeu Bayésien statique est que, au lieu d'appliquer une surveillance permanente IDS, le défenseur peut mettre en oeuvre une stratégie selon sa solution d'équilibre Bayésien qui maximise son gain espéré, mais un inconvénient possible dans la pratique est la difficulté d'attribuer des probabilités a priori (μ_0) précises sur le type du joueur i . Pour remédier à ce problème, le modèle a été étendu à un jeu Bayésien dynamique, où le défenseur mit à jour ses croyances selon l'évolution du jeu, c'est à dire le défenseur j mit à jour ses croyances sur les types de son adversaire i à la fin de chaque étape du jeu en calculant ses croyances a posteriori en tenant compte de l'histoire du profil d'actions du joueur i .

Références

1. S. Farhang, M.H. Manshaei, M.N. Esfahani and Q. Zhu : A dynamic Bayesian security game framework for strategic defense mechanism design. R. Poovendran and W. Saad (Eds.) : *Game Sec 2014*, LNCS 8840 319–328, (2014).
2. J.C. Harsanyi : Game with incomplete information played by "Bayesian" players, I. The basic model. *Management science* 14 :317-334, (1967).
3. Y. Liu, C. Comaniciu and H. Man : A Bayesian game approach for intrusion detection in wireless Ad Hoc networks. *Game Nets'06*, October14, 2006, Pisa, Italy. Copyright ACM 1-59593-507-X/06/10, (2006).

Equilibre de Berge : Entre Situationnisme et Optimalité de Pareto

Kadi S., Zaidi H., Tazdaït T. et Aïssani D.

a - Département de Recherche Opérationnelle, université de Béjaïa, Algérie.

b - Centre CIREP Paris, France.

c - Unité de recherche LaMOS, université de Béjaïa, Algérie.

Résumé Dans ce travail, nous exposons l'intérêt pratique de l'équilibre de Berge en parlant du situationnisme lors d'un jeu, où nous montrons, à la Gauthier, quand est-ce que les individus ont intérêt à choisir la règle de comportement de Berge et quand pourront-ils choisir de se comporter selon Nash. Nous présentons également l'intérêt de l'efficacité au sens de Pareto dans le choix de la règle de comportement à suivre, en exhibant trois différentes situations de jeu.

Mots clés : Équilibre de Berge, équilibre de Nash, équilibre de Berge- Vaisman, efficacité de Pareto, situationnisme dans le jeu.

Abstract

In this work, we expose the practical interest of Berge equilibrium by speaking about the situational gaming, where we show, à la Gauthier', when individuals should follow the Berge rule behavior and when should they behave according to Nash. We also present the interest of Pareto efficiency in making choices about the rule-following behavior, by exhibiting three different game situations.

Key words : Berge equilibrium and equilibrium, Berge- Vais man equilibrium, Pareto efficiency, situational gaming.

Introduction

Berge a introduit pour la première fois l'équilibre de Berge en 1957 pour un jeu sous forme normale basé sur la notion d'équilibre d'une structure de coalition R respectivement à un ensemble de coalitions S . Cet équilibre n'a reçu aucune attention de la part des chercheurs de la théorie des jeux pendant deux décennies.

L'équilibre de Nash peut conduire chaque individu à une situation de non regret, même si elle ne peut pas lui garantir un gain optimal. Cependant, les conditions et les théorèmes d'existence de l'équilibre de Nash ne sont pas constructifs dans le sens où ils ne donnent pas une méthode pour exhiber un équilibre dans un jeu quelconque.

Dans le cas où le jeu n'admet pas d'équilibre de Nash, ou bien si on a une multitude d'équilibres de Nash, alors, on fait appel à l'équilibre de Berge, qui est considéré comme une solution alternative dans ces deux situations précédentes, où le gain de chaque joueur est maximisé par les autres joueurs.

L'objectif de ce travail est l'étude de l'équilibre de Berge sur un aspect particulier, en parlant du situationnisme et en donnant quelques exemples de situations où cet équilibre est Pareto-optimal.

La première section est consacrée à la définition d'un équilibre de Berge tel que Berge l'a introduit en 1957, puis on définit l'équilibre simple de Berge ainsi que l'équilibre de Berge individuellement Rationnel.

Dans la deuxième section, nous présenterons une synthèse bibliographique sur les différents travaux qui ont été élaborés à propos de ces équilibres.

La dernière section portera sur l'intérêt pratique de l'équilibre de Berge, en parlant du situationnisme. On décrit la prise de décision dans une vision de situation qui suppose que les règles de comportement des individus varient en fonction des situations rencontrées. On montrera également à la Gauthier qu'un joueur pourrait être mieux loti en jouant la règle de comportement en Berge dans quelques situations de jeu. Comme on s'intéressera à montrer l'intérêt que porte la notion Pareto-optimale dans le choix de la règle de comportement à suivre lors de différentes situations de jeux. Pour cela, on présentera trois exemples de situations illustrant que le choix de se comporter selon Berge, tout en étant individuellement rationnel, fournit un résultat optimal de Pareto. Ceci pouvant bien être une raison qui inciterait les individus à opter pour la règle de Berge.

1 Équilibre de Berge

Dans les jeux stratégiques, on se contente de l'énumération de toutes les stratégies avec les issues et les gains qu'elles engendrent. Ce sont les jeux qui se déroulent en un seul coup et où les joueurs interviennent simultanément. On les appelle les jeux sous forme normale.

$$\langle I, \{X_i\}_{i \in I}, \{f_i\}_{i \in I} \rangle; \quad (1)$$

où :

$I = \{1, 2, \dots, N\}$: est l'ensemble des joueurs, supposé fini ou infini dénombrable.

X_i : est l'ensemble des stratégies du joueur i , $i \in I$. ($X_i \subset \mathbb{R}^n$).

$X = \prod_{i \in I} X_i$: est l'ensemble des profils de stratégies du jeu (issues du jeu).

$f_i : X \rightarrow \mathbb{R}$: est la fonction de gain du joueur i , $i \in I$.

On note \mathcal{C} l'ensemble de toutes les coalitions (c-à-d : des sous ensembles non-vides de I) [12].

Pour chaque coalition $R \in \mathcal{C}$, on note par :

$-R$: la coalition complémentaire de R ,

avec :

$$-R = \{i \in I, \text{ tel que } i \notin R\}.$$

Si R est réduit à un singleton $\{i\}$, alors on note par $-i$ l'ensemble $-R$.

On note également, $X_R = \prod_{i \in R} X_i$: l'ensemble des profils de stratégies des joueurs dans la coalition R .

Si $\{K_i\}_{i \in \{1, \dots, s\}} \subset \mathbb{N}$ est une partition de l'ensemble des joueurs I , alors, chaque profil de stratégies $x = (x_1, \dots, x_N) \in X$ peut être écrit comme suit :

$$x = (x_{K_1}, x_{K_2}, \dots, x_{K_s}), \text{ où } x_{K_i} \in X_{K_i} = \prod_{j \in K_i} X_j.$$

Définition 1.1 (*Équilibre de Berge*)

Considérons le jeu (1). Soit $R = \{R_i\}_{i \in M} \subset \mathcal{C}$ une partition (structure de coalition) de I et $S = \{S_i\}_{i \in M}$ un ensemble de sous-ensembles de I .

Une stratégie réalisable $\bar{x} \in X$ est un point d'équilibre pour l'ensemble R relatif à l'ensemble S ou bien **un équilibre de Berge** pour le jeu (1), si :

$$f_{r_m}(\bar{x}) \geq f_{r_m}(x_{S_m}, \bar{x}_{-S_m}),$$

$\forall m \in M, M \subset \mathbb{N};$

$\forall r_m \in R_m \text{ et } \forall x_{S_m} \in X_{S_m}.$

Cela signifie qu'à l'équilibre de Berge, les joueurs dans la coalition S_m jouent un profil de stratégies qui maximise le gain des joueurs dans la coalition R_m , tout en ignorant ou négligeant leurs propres gains (quand $S_m \cap R_m = \emptyset$).

Cette situation fait que l'équilibre de Berge paraît insensé et irrationnel. En fait, les gains

des joueurs dans S_m sont pris en compte par quelques autres joueurs. En effet, soit $j \in S_m$, puisque la famille des coalitions R est une partition de l'ensemble des joueurs I , alors il existe un certain $p \in M$ tel que $j \in R_p$. Selon la définition de l'équilibre de Berge, les joueurs de la coalition S_p correspondante maximisent les fonctions de gain des joueurs dans R_p dont la fonction de gain du $j^{\text{ème}}$ joueur.

Il apparaît qu'à l'équilibre de Berge, globalement, chaque joueur maximise le gain d'au moins un autre joueur, en retour son gain est maximisé par au moins un autre joueur.

En général, le problème de rationalité individuelle peut paraître quand $S_m \cap R_m = \emptyset$, car dans ce cas les joueurs dans R_m ne maximisent pas leurs propres fonctions de gain.

Définition 1.2 (*Équilibre simple de Berge*)

Un profil de stratégie $\bar{x} \in X$ est un *équilibre simple de Berge* (au sens de Zhukovskii) pour le jeu (1), si :

$$f_i(\bar{x}) \geq f_i(x_{-i}, \bar{x}_i), \quad \forall i \in I \text{ et } \forall x_{-i} \in X_{-i}. \quad (2)$$

Cette définition signifie que quand un joueur $i \in I$ joue sa stratégie \bar{x}_i de l'équilibre simple de Berge, il ne peut pas obtenir un gain maximum sauf si les autres joueurs $-i$ jouent volontairement (ou obligatoirement) leur stratégie \bar{x}_{-i} de l'équilibre de Berge \bar{x} . Autrement dit, si au moins un des joueurs de la coalition $-i$ dévie de sa stratégie d'équilibre, le gain du joueur i dans le profil de stratégies résultant est inférieur ou égal à son profit $f_i(\bar{x})$.

•Cas particulier :

En posant $M = I$, $R_i = \{i\}$ et $S_i = -i$, $\forall i \in I$, on voit bien que l'équilibre simple de Berge est un cas particulier de l'équilibre de Berge.

Il est important de noter que pour quelques structures de coalitions R et ensembles de coalitions S , l'équilibre de Berge peut ne pas être individuellement rationnel tel que le stipule Vaisman pour un équilibre simple de Berge (voir définition 1.3). Alors, pour chaque équilibre de Berge, il est nécessaire d'introduire la notion de rationalité individuelle à sa définition ou bien choisir uniquement les équilibres de Berge qui sont individuellement rationnels dans le processus de résolution du jeu.

Définition 1.3 (*Équilibre de Berge-Vaisman*)

On dit qu'un profil de stratégies $\bar{x} \in X$ est un *équilibre de Berge-Vaisman* (ou bien un équilibre de Berge individuellement rationnel) pour le jeu (1), si :

1. $\forall i \in I, \forall y_{-i} \in X_{-i}, f_i(\bar{x}) \geq f_i(y_{-i}, \bar{x}_i)$,
2. $\forall i \in I$,

$$\alpha_i = \sup_{x_i \in X_i} \inf_{y_{-i} \in X_{-i}} f_i(y_{-i}, x_i) \leq f_i(\bar{x}). \quad (3)$$

La première condition de la définition signifie que l'équilibre de Berge-Vaisman est un équilibre simple de Berge. La deuxième condition stipule que le profil de stratégies \bar{x} est individuellement rationnel. Autrement dit, pour chaque joueur $i \in I$, l'équilibre de Berge-Vaisman \bar{x} fournit un gain qui est supérieur ou égal à son seuil de sécurité, noté α_i . On dit alors que l'équilibre de Berge-Vaisman est individuellement rationnel.

Il y a deux principales raisons qui ont motivé l'introduction de l'équilibre de Berge-Vaisman comme solution alternative de l'équilibre de Nash. La première est l'absence du concept de solution (en stratégies pures) dans les jeux où il n'y a pas d'équilibre de Nash, la seconde réside dans la difficulté de choisir un équilibre de Nash dans les jeux où il y a plus d'un équilibre de Nash.

L'équilibre de Berge-Vaisman peut être utilisé dans l'étude de plusieurs modèles non-coopératifs, plus particulièrement les jeux de coalition. De plus, contrairement à l'équilibre de Nash, ce concept permet d'atteindre des profils de stratégies coopératifs. En effet, avec cet équilibre il n'est pas nécessaire d'introduire des suppositions comportementales afin d'obtenir des profils de stratégies coopératifs. Par conséquent, il devient possible d'atteindre la coopération dans un contexte non-coopératif. Cette propriété est très importante pour les jeux tels que le dilemme du prisonnier.

•Cas particulier :

Dans le cas où $M = I$, $R_i = \{i\}$ et $S = \{i\}$, $\forall i \in I$, alors l'équilibre de Berge est un équilibre de Nash.

Remarque 1.1 *Un équilibre de Berge, qui est aussi un équilibre de Nash, est appelé **équilibre de Berge-Nash** ou **équilibre B-Nash**.*

De façon similaire, on peut définir l'équilibre simple de Berge-Nash, et l'équilibre de Berge-Vaisman-Nash.

2 Synthèse bibliographique

L'équilibre de Berge, du concept d'équilibre non-coopératif, a été développé, en 1957, par le mathématicien français Claude Berge. Il fait ressortir la coopération comme issue d'équilibre dans le dilemme du prisonnier. Ce concept est passé totalement inaperçu dans la littérature alors même que Berge l'avait introduit pour pallier aux situations de multiplicité d'équilibres de Nash et aux situations caractérisées par l'absence de cet équilibre.

L'équilibre de Berge est donc resté à l'écart pendant deux décennies. Dans les années 80, Zhukovskii et son groupe de chercheurs ont commencé à étudier un cas spécial de cet équilibre, qu'on appellera "**équilibre simple de Berge**" (voir définition 1.2) [14].

En 1988, Radjef a étudié l'existence d'un équilibre de Berge pour des jeux différentiels à N personnes. Il a donc ainsi fourni un théorème qui a été une référence de base dans plusieurs travaux qui lui ont succédé [13].

Abalo et Kostreva [1] s'intéressent aux jeux non-coopératifs et au concept de solution qui est l'équilibre de Berge relié à l'équilibre de Nash. Pour cela, ils ont fourni deux théorèmes d'existence (théorème 2.2 et théorème 2.3) qui donnent des conditions suffisantes pour l'existence d'un équilibre de Berge qui est aussi un équilibre de Nash (équilibre de Berge-Nash). Plus tard [3], ils généralisent leurs résultats et fournissent également deux théorèmes d'existence (théorème 3.1 et théorème 3.2) qui constituent aussi des conditions suffisantes pour l'existence, cette fois-ci, d'un équilibre de Berge. Ces théorèmes ont été démontrés pour des jeux où le nombre de joueurs est fini et aussi dans le cas d'une infinité de joueurs. Puis, ils présentent des résultats contribuant à une nouvelle solution au problème de sélection d'un équilibre de Nash approprié [4].

En 2007, Nessah, Larbani et Tazdaït ont présenté un article fournissant une condition nécessaire et suffisante pour l'existence d'un équilibre de Berge au sens de Zhukovskii [10]. Les auteurs ont repris les théorèmes d'Abalo et Kostreva (théorème 3.1 et théorème 3.2 [3], théorème 2 et théorème 3 [4]) et ont montré, par un contre exemple, qu'au fait, les conditions de ces théorèmes n'étaient pas suffisantes. Donc, ils ont proposé une condition qui pallie à ce problème.

Ces mêmes auteurs se sont intéressés également aux conditions d'existence pour un équilibre de Berge-Vaisman pour un jeu sous forme normale. Ils fournissent aussi deux méthodes de calcul des équilibres de Berge-Vaisman : la première calcule les équilibres de Berge-Vaisman et la seconde ceux de Berge-Vaisman qui sont aussi des équilibres de Nash (équilibres de Berge-Vaisman-Nash) [12].

En 2008, Larbani et Nessah se sont intéressés au problème d'existence des équilibres de Berge et ceux de Berge-Nash [9].

Au cours de cette année (2011), plusieurs auteurs ont publié quelques ouvrages. Parmi eux, on trouve un article intéressant [5] où sont démontrés quelques résultats de base des équilibres de Berge et leurs relations avec les équilibres de Nash, en fournissant une méthode directe qui trouve les équilibres de Berge pour des jeux à N -personnes. Les auteurs ont exploré quelques exemples spécifiques. Ils ont expliqué comment l'équilibre de Berge fournit un modèle contraignant de coopération dans les dilemmes sociaux.

Puis, on trouve également l'article [6], qui fournit une étape vers la manière de traiter les préférences morales en se basant sur des règles de comportement. Les auteurs décrivent, ainsi, la prise de décision dans une vision de situation (de jeu) qui suppose que les règles de comportement des individus varient en fonction des situations rencontrées. De ce fait, ils

proposent une règle de comportement, complémentaire à celle de Nash, qui est la règle de Berge.

Au cours de ce mois d'avril, est paru l'ouvrage [7], où les auteurs s'intéressent également à la prise de décision dépendant, à la fois, des individus et de la situation dans laquelle se déroule le jeu. Étant donné que l'action est gouvernée par plusieurs règles de comportement, ils se sont basés sur celle de Berge, susceptible de compléter la règle de Nash. Ils définissent, alors, des conditions d'existence pour les équilibres de Berge, Berge individuellement rationnels et Berge-Nash. Pour le cas général avec un nombre infini dénombrable de joueurs, ils utilisent le théorème de l'égalité g -maximum pour dériver des conditions générales suffisantes pour l'existence de ces équilibres. Ils définissent également des procédures qui les calculent.

3 Situationnisme et optimalité au sens de Pareto

Le situationnisme traduit l'idée que les individus adaptent leur comportement aux situations auxquelles ils sont confrontés. Donc, ils coopèrent dans les situations de coopération et ils jouent selon leur propre intérêt dans les situations de concurrence.

3.1 Équilibre de Berge entre utilitarisme et situation de jeu

Considérons un individu qui joue suivant la règle de Berge en espérant que l'autre joueur fera de même, tout en sachant que cette règle de comportement conduit à un meilleur résultat. Autrement dit, on pose que quelque soit la règle de comportement choisie, les individus sont incités (motivés) par des motivations utilitaristes.

Ils jouent suivant une règle de comportement orientée vers un égoïsme ou bien vers un soutien mutuel, mais bien sûr, motivés par le succès.

La règle de comportement en Berge et l'équilibre de Berge donnent des résultats particuliers dans les situations de jeux à somme nulle. Afin d'illustrer cela, considérons un jeu à somme nulle et à deux joueurs, avec :

f_1 : la fonction d'utilité du premier joueur, et

f_2 : la fonction de perte du second joueur.

Donc, $f_2 = -f_1$

Sachant que (\bar{x}_1, \bar{x}_2) est un équilibre de Berge si :

$$\begin{cases} \forall x_2 \in X_2, & f_1(\bar{x}_1, x_2) \leq f_1(\bar{x}_1, \bar{x}_2), \\ \forall x_1 \in X_1 & f_2(x_1, \bar{x}_2) \leq f_2(\bar{x}_1, \bar{x}_2). \end{cases} \quad (4)$$

Étant donné que $f_2 = -f_1$, le premier joueur obtient son gain maximum si le second accepte de jouer sa stratégie de Berge \bar{x}_2 qui consiste à la maximisation de sa perte (c-à-d :

$$\max_{x_2} f_1(\bar{x}_1, x_2) = \max_{x_2} [-f_2(\bar{x}_1, x_2)].$$

De manière similaire, le second joueur obtient sa perte minimale si le premier accepte de jouer sa stratégie de Berge \bar{x}_1 de manière à minimiser son utilité ($\min_{x_1} f_2(x_1, \bar{x}_2) = \min_{x_1} [-f_1(x_1, \bar{x}_2)]$).

Ceci est un comportement sacrificiel qui n'est pas compatible avec un traitement utilitariste. Plus généralement, les situations de compétition, telles que les jeux compétitifs, les duopoles de Bertrand ou Cournot, ne concordent pas avec le soutien mutuel, qui nécessite l'atteinte d'un but exclusif mutuel.

Dans de telles situations, on peut supposer que les individus adoptent un comportement de type Nash. Cependant, la règle de Berge est souvent appropriée dans les situations de coopération où les agents doivent coordonner leurs activités afin d'atteindre des résultats satisfaisants.

Mais dans tous les cas, et comme suggéré par une preuve expérimentale, il n'y a pas de règle précise et une partie de l'approche de situation est l'environnement dans lequel les joueurs communiquent, qui affecte significativement la règle de comportement qu'ils choisissent de suivre.

3.2 Situationnisme selon Gauthier

En 1986, Gauthier a été le premier qui a effectué une analyse des jeux suivant une vision de situation. L'approche de Gauthier est spécialement pertinente dans le sens où elle capte le besoin d'une association entre les règles de comportement et les situations de jeu. L'approche de Gauthier nous dit, en terme d'utilitarisme, si un jeu quelconque doit se jouer suivant la règle de Berge ou bien Nash.

Gauthier considère que les individus prennent leurs dispositions avant d'interagir. Les dispositions sont définies comme étant des règles de comportement pouvant varier suivant la situation de jeu.

En se basant sur le dilemme du prisonnier, Gauthier suppose que les individus peuvent adopter une des dispositions suivantes :

1. "*Maximisation directe*" : une règle de comportement selon laquelle ils cherchent à maximiser leurs profits étant données les stratégies de ceux avec lesquels ils interagissent.
2. "*Maximisation contraignante*" : une règle de comportement suivant laquelle ils cherchent, en quelque sorte, à maximiser leurs gains, pas suivant les stratégies mais selon les profits de ceux avec lesquels ils interagissent.

Étant données ces deux possibilités, le résultat de Gauthier est que les joueurs choisissent la maximisation contraignante dans le dilemme du prisonnier.

En utilisant cette approche, les individus sont donc face à deux dispositions qui s'offrent à eux : adopter la règle de comportement en Berge ou bien en Nash. L'exemple suivant montre, en quelque sorte, la manière de procéder [6].

Exemple 3.1 Soit le jeu de la confiance suivant :

		P_2	
		<i>Honorer</i>	<i>Exploiter</i>
P_1	<i>Confiance</i>	X, X	Y, Z
	<i>Méfiance</i>	$0, 0$	$0, 0$

$$Y < 0 < X < Z.$$

L'équilibre de Nash est unique, donné par (*Méfiance*, *Exploiter*) et l'équilibre de Berge est aussi unique, donné par (*Confiance*, *Honorer*).

En supposant que les joueurs sont couplés aléatoirement. Ils ont le choix entre deux dispositions basées sur la règle de comportement en Nash (RN) ou bien la règle de comportement en Berge (RB).

On appelle le joueur RN, le premier type, et le joueur RB, le second type. Les joueurs suivent RB quand ils s'attendent à ce que leurs adversaires feront de même, et le choix d'appartenir à un type plutôt qu'à un autre dépend des gains relatifs espérés.

La probabilité de jouer en tant que type 1 ou type 2 est la même. On note par α la part des joueurs RB dans la population.

• D'abord, on suppose qu'avant le début du jeu, les joueurs savent à la fois leur propre type et le type de leurs adversaires. Les gains RN et RB espérés sont donc respectivement :

$$\left\{ \begin{array}{l} g_e(RN) = 0, \\ g_e(RB) = \alpha X. \end{array} \right. \quad (5)$$

– Pour $\alpha > 0$:

Le gain espéré pour les joueurs RB est strictement supérieur au gain espéré pour les joueurs RN, $g_e(RB) > g_e(RN)$.

– Pour $\alpha = 0$:

Les gains espérés pour les deux types de joueurs sont les mêmes, $g_e(RB) = g_e(RN)$.

Être un joueur RB améliore le bien être dès qu'il y a d'autres joueurs RB dans la population avec lesquels les relations de confiance peuvent être établies. Autrement dit, s'il y a un unique joueur RB il jouera comme s'il était un joueur RN.

• En relaxant la supposition que les individus peuvent identifier le type de leurs adversaires, le résultat demeure vigoureux. Pour voir cela, on suppose qu'à la fois les joueurs RB et RN manquent à identifier ceux avec lesquels ils interagissent.

Soit β la probabilité que les joueurs RB s'identifient les uns les autres et θ la probabilité qu'ils manquent à identifier les joueurs RN.

Les gains espérés des joueurs RB et RN sont respectivement :

$$\begin{cases} g_e(RB) = \alpha\beta X + (1 - \alpha)\frac{\theta Y}{2}, \\ g_e(RN) = \frac{\alpha\theta Z}{2}. \end{cases} \quad (6)$$

et un individu choisit d'être un joueur RB si :

$$\frac{\beta}{\theta} > \frac{\alpha Z - (1 - \alpha)Y}{2\alpha X}. \quad (7)$$

Si la condition (7) est vérifiée, les joueurs adoptent la règle de Berge même s'ils peuvent interagir sans le savoir avec des joueurs RN.

Remarque 3.1 *Quand la proportion des joueurs RB croît, $\frac{\beta}{\theta}$ croît également, en diminuant le risque de prendre des joueurs RN pour des joueurs RB.*

Remarque 3.2 *Quand le gain relatif à la coopération croît, la condition (7) devient moins contraignante rendant, ainsi, la règle de Berge plus attrayante.*

On déduit que les individus seront plus probables d'adopter la règle de Berge quand l'ampleur du dilemme social est importante et quand leur environnement social n'est pas aussi égoïste.

Il existe d'autres situations de jeu pouvant mener au résultat contraire, faisant de la règle de Nash la plus préférée, telles que les situations de jeux à somme nulle, la plupart des situations de compétition et même dans les jeux de coopération où les intérêts ne sont pas conflictuels. Considérons cet exemple :

Exemple 3.2 *Soit le jeu suivant :*

		P_2	
		C	D
P_1	C	$3, 3$	$4, 1$
	D	$1, 4$	$2, 2$

Il y a un unique équilibre de Nash en (C, C) et un unique équilibre de Berge en (D, D) .

En laissant les individus choisir leur disposition avant le début du jeu, on peut déduire que la meilleure option pour les joueurs est de jouer en mode Nash.

• Afin d'illustrer cela, d'abord, on considère que les individus savent le type de leurs adversaires. On aura alors :

$$\begin{cases} g_e(RN) = 3, \\ g_e(RB) = 3 - \alpha. \end{cases} \quad (8)$$

et pour $\alpha > 0$, il y a intérêt à choisir une disposition RN.

• Maintenant, supposons que les individus ne connaissent pas le type de leurs adversaires. On a alors :

$$\begin{cases} g_e(RN) = \alpha\theta + 3, \\ g_e(RB) = \alpha(2\theta - \beta) + 3 - 2\theta. \end{cases} \quad (9)$$

On déduit que : $g_e(RN) > g_e(RB)$, ce qui nous mène à l'inégalité suivante :

$$\frac{\beta}{\theta} > \frac{(\alpha - 2)}{\alpha}.$$

Ceci est toujours vrai et les individus choisiront toujours de jouer en mode Nash.

Généralement, et en prenant en compte la vision rationnelle révisionniste suivant laquelle les individus suivent des règles de conduite qui font que leur vie va mieux, cette approche de disposition peut être appliquée pour une situation de jeu quelconque qui indique quand est-ce que l'équilibre de Berge pourrait être appliqué et quand est-ce que l'équilibre de Nash est meilleur.

Le choix de chacun d'eux se rapporte à la structure d'utilité du jeu et à l'appréciation probabiliste, des joueurs, subjective du type des adversaires avec lesquels ils interagissent.

3.3 Équilibre de Berge et optimalité de Pareto

Parmi les raisons qui peuvent inciter un individu à choisir d'adopter une règle de comportement en Berge plutôt qu'en Nash, on peut penser à l'efficacité au sens de Pareto. Sachant qu'un équilibre de Berge ne conduit pas toujours à un optimum de Pareto, mais une fois cette condition vérifiée, on saura qu'il n'y a pas un meilleur choix que l'équilibre de Berge.

Dans cette section, nous prendrons soin d'énoncer quelques exemples de situations où l'équilibre de Berge est bien Pareto-optimal.

Exemple 3.3 Dilemme du prisonnier :

Dans cet exemple, on reprend le dilemme du prisonnier, où on considère les gains comme étant des variables : R, S, T, P ; avec $T > R > P > S$.

		P_2	
		C	D
P_1	C	R, R	S, T
	D	T, S	P, P

La stratégie D est la stratégie dominante, et le résultat (D, D) est l'équilibre de Nash en stratégies pures. Tandis que le résultat (C, C) est l'équilibre de Berge en stratégies pures de ce jeu. En effet, d'après la définition 1.2 :

$$f_1(C, C) > f_1(C, D), \text{ et}$$

$$f_2(C, C) > f_2(D, C).$$

Ce qui n'est pas vérifié pour le profil de stratégies (D, D) , car :

$$f_1(D, D) < f_1(D, C), \text{ et}$$

$$f_2(D, D) < f_2(C, D).$$

Si on défini α_i comme étant le niveau maximum de sécurité du joueur i , avec :

$$\alpha_i = \sup_{x_i \in X_i} \inf_{x_{-i} \in X_{-i}} f_i(x_i, x_{-i}) \leq f_i(\bar{x}), \quad i \in I, \quad x_{-i} \in X_{-i}. \quad (10)$$

Si l'inégalité (10) est vraie pour chaque joueur $i \in I$, alors l'équilibre de Berge fournit un gain qui ne doit pas être inférieur au *maximin*.

En considérant le dilemme du prisonnier, on trouve :

$$\alpha_1 = \max\{\min_{x_2} f_1(C, x_2), \min_{x_2} f_1(D, x_2)\} = \max\{S, P\} = P < R, \text{ et}$$

$$\alpha_2 = \max\{\min_{x_1} f_2(x_1, C), \min_{x_1} f_2(x_1, D)\} = \max\{S, P\} = P < R.$$

On conclut donc que le profil de stratégies (C, C) est individuellement rationnel.

Nous avons également que l'équilibre de Berge, dans cette situation, est Pareto-optimal. Ce qui inciterait les joueurs à choisir de jouer suivant la règle de comportement en Berge plutôt qu'en Nash.

Exemple 3.4 Jeu de la poule mouillée :

Nous reprendrons le même tableau précédent, sauf que dans ce cas on considère que $T > R > S > P$.

Ce jeu admet deux équilibres de Nash, en stratégies pures, qui sont (C, D) et (D, C) . L'unique équilibre de Berge pour ce jeu est (C, C) . En effet :

$$f_1(C, C) > f_1(C, D), \text{ et}$$

$$f_2(C, C) > f_2(D, C).$$

Ce qui n'est pas vérifié pour les deux profils de stratégies (C, D) et (D, C) , puisque :

$$f_1(C, D) < f_1(C, C), \text{ et}$$

$$f_2(C, D) > f_2(D, D).$$

Et encore :

$$f_1(D, C) > f_1(D, D), \text{ et}$$

$$f_2(D, C) < f_2(C, C).$$

En calculant le *maxmin* de chaque joueur, on trouve :

$$\alpha_1 = \max\{\min_{x_2} f_1(C, x_2), \min_{x_2} f_1(D, x_2)\} = \max\{S, P\} = S < R, \text{ et}$$

$$\alpha_2 = \max\{\min_{x_1} f_2(x_1, C), \min_{x_1} f_2(x_1, D)\} = \max\{S, P\} = S < R.$$

Ce qui nous permet de déduire que l'équilibre (C, C) est individuellement rationnel.

L'équilibre de Berge, dans cette situation, est également optimal de Pareto. Ceci pourrait bien encourager les individus à adopter une règle de comportement en Berge plutôt qu'une autre.

Exemple 3.5 *Jeu de la confiance* :

En reprenant le tableau suivant, où $Y < 0 < X < Z$:

		P_2	
		<i>Honorer</i>	<i>Exploiter</i>
P_1	<i>Confiance</i>	X, X	Y, Z
	<i>Méfiance</i>	$0, 0$	$0, 0$

Sachant que l'équilibre de Nash est unique (*Méfiance, Exploiter*) et l'équilibre de Berge est aussi unique (*Confiance, Honorer*), car :

$$f_1(\text{Confiance}, \text{Honorer}) > f_1(\text{Confiance}, \text{Exploiter}), \text{ et}$$

$$f_2(\text{Confiance}, \text{Honorer}) > f_2(\text{Méfiance}, \text{Honorer}).$$

Ce qui n'est pas vrai pour le profil de stratégies (*Méfiance, Exploiter*), puisque :

$$f_1(\text{Méfiance}, \text{Exploiter}) \geq f_1(\text{Méfiance}, \text{Honorer}), \text{ et}$$

$$f_2(\text{Méfiance}, \text{Exploiter}) < f_2(\text{Confiance}, \text{Exploiter}).$$

Le calcul du *maxmin* donne :

$$\alpha_1 = \max\{\min_{x_2} f_1(\text{Confiance}, x_2), \min_{x_2} f_1(\text{Méfiance}, x_2)\} = \max\{Y, 0\} = Y < X, \text{ et}$$

$$\alpha_2 = \max\{\min_{x_1} f_2(x_1, \text{Honorer}), \min_{x_1} f_2(x_1, \text{Exploiter})\} = \max\{0, 0\} = 0 < X.$$

Donc, l'équilibre de Berge est individuellement rationnel. Cet équilibre est également Pareto-optimal, ce qui ramènerait les individus à choisir d'opter pour la règle de comporte-

ment de Berge.

Dans ces trois exemples précédents, on a présenté trois situations de jeux où les équilibres de Berge individuellement rationnels sont optimaux de Pareto. Ainsi, quand les individus choisissent d'adopter la règle de comportement en Berge, tout en étant individuellement rationnels, ils seront mieux lotis qu'en choisissant de se comporter en Nash, et vue l'efficacité de Pareto, cela leur permet de savoir qu'il n'y a pas un autre résultat meilleur que celui fourni en optant pour la règle de Berge.

Conclusion

L'objectif assigné à ce travail a été l'étude de l'équilibre de Berge sur un aspect particulier, en parlant du situationnisme et en donnant quelques exemples de situations où cet équilibre est Pareto-optimal.

Nous avons tout d'abord rappelé la définition d'un équilibre de Berge tel que Berge l'a introduit en 1957, puis on a défini l'équilibre simple de Berge ainsi que l'équilibre de Berge individuellement Rationnel.

Par la suite, nous avons présenté une synthèse bibliographique sur les différents travaux qui ont été abordés à propos de ces équilibres.

Nous nous sommes enfin intéressés à l'intérêt pratique de l'équilibre de Berge en parlant du situationnisme. Nous avons décrit la prise de décision dans une vision de situation qui suppose que les règles de comportement des individus varient en fonction des situations rencontrées. Nous avons aussi montré à la Gauthier qu'un joueur pourrait être mieux loti en jouant la règle de comportement en Berge dans quelques situations de jeu. Comme on s'est également intéressé à montrer l'intérêt que porte la notion Pareto-optimale dans le choix de la règle de comportement à suivre lors de différentes situations de jeux. Pour cela, on a présenté trois exemples de situations qui illustrent que le choix de se comporter selon Berge, tout en étant individuellement rationnel, fournit un résultat optimal de Pareto.

Ce qui nous laisse penser à généraliser ce concept pour avoir comme perspective de trouver des conditions générales sous lesquelles l'équilibre de Berge, qui est individuellement rationnel, soit Pareto-optimal.

Dans le but de nous permettre d'identifier les classes de jeux où les équilibres de Berge, individuellement rationnels, sont optimaux de Pareto. Car, ceci pourrait bien aider les joueurs dans le processus de prise de décision de la règle de comportement à adopter.

Références

- [1] K.Y. Abalo and M.M. Kostreva. *Fixed points, Nash games and their organizations*. Topological Methods in Nonlinear Analysis 8(1996) 205-215.
- [2] K.Y. Abalo and M.M. Kostreva. *Equi-well-posed games*. J. Optim. Theory Appl. 89(1996) 89-99.
- [3] K.Y. Abalo and M.M. Kostreva. *Some existence theorems of Nash and Berge equilibria*. Applied Mathematics Letters. 17(2004) 569-573.
- [4] K.Y. Abalo and M.M. Kostreva. *Berge equilibrium : Some recent results from fixed point theorems*. Appl. Math. Comput. 169(2005) 624-238.
- [5] A.M. Colman, T.W. Körner, O. Musy and T. Tazdaït. *Mutual support in games : Some properties of Berge equilibria*. Journal of Mathematical Psychologie (2011), doi :10.1016/j.jmp.2011.02.001.
- [6] P. Courtois, R. Nessah and T. Tazdaït. *How to play the games ? Nash versus Berge behavior rules*. Document de recherche, LAMETA, UMR, DR n° 2011-05. 2011.
- [7] P. Courtois, M. Larbani, R. Nessah and T. Tazdaït. *Situational gaming : On Berge equilibria*. Document de travail, IESEG School of Management. Lille, 2011.
- [8] D.P. Gauthier. *Morals by agreement*. Oxford, Oxford University Press. 1986.
- [9] M. Larbani and R. Nessah. *A note on the existence of Berge and Berge-Nash equilibria*. Mathematical Social Sciences 55 (2008) 258-271.
- [10] M. Larbani, R. Nessah and T. Tazdaït. *A note on Berge equilibrium*. Appl. Math. Lett. 20(2007) 926-932.
- [11] M. Larbani, R. Nessah and T. Tazdaït. *Strong Berge and Pareto equilibrium existence for a noncooperative game*. L'Archive Hal, halshs-00271464, version 1, 09 Apr 2008.
- [12] M. Larbani, R. Nessah and T. Tazdaït. *On Berge equilibria*. IESEG Working Paper Series 2010-Eco-16.
- [13] M. S. Radjef. *Sur l'existence d'un équilibre de Berge pour un jeu différentiel à N personnes*. Cahier Mathématiques. Université d'Oran, 1988, n°.1, 89-93.
- [14] V. I. Zhukovskii. *Some problems of non-antagonistic differential games*. In P. Kenderov (Ed.), Mathematical methods in operations research. Sofia : Bulgarian Academy of Sciences, 1985, 103-195.

II- Evaluation des performances des réseaux de communication

Routing and Control of Congestion in Telecommunication Networks

Touati N., Hadji M., Laib F. et Aïssani D.

makhadji@yahoo.fr, lamos_bejaia@hotmail.com

Abstract

A telecommunication network is composed of a set auto-switches and of a router which are connected by telecommunication circuits which ensure the network transmission function. We will first study the problem that all networks, whether of telecommunication or not, are confronted with. This problem is the congestion. Secondly, we will treat another problem that the solution can help us to solve the problem of congestion. This problem is called routing. To clarify our problematic, we suppose that we have a well-defined connection, so :

- What can we offer as solution to congestion problem in the network ?
- Which routing technique one should use to optimize information transmission ?
- What are performances of studied system ?

Key words : Telecommunication network, congestion, routing, performances.

A telecommunication network is composed of a set auto-switches and of a router which are connected by telecommunication circuits which ensure the network transmission function. We will first study the problem that all networks, whether of telecommunication or not, are confronted with. This problem is the congestion. Secondly, we will treat another problem that the solution can help us to solve the problem of congestion. This problem is called routing. To clarify our problematic, we suppose that we have a well-defined connection, so :

- What can we offer as solution to congestion problem in the network ?
- Which routing technique one should use to optimize information transmission ?
- What are performances of studied system ?

The routing problem is defined to be the responsibility of deciding on which output line a packet should be redirected. Once the route is defined many routing processes are proposed, the best process is the one which gives : simplicity, robustness and accuracy. The congestion is defined as a traffic jam observed at a node of the network, as a result of a burst of information packets arrivals.

To solve these two problems, we will start by defining a variant of the congestion problem. This variant is defined by a basic principle which consists in placing a breakthrough bucket ahead of every emission source. The breakthrough bucket can be represented by a queue with a limited capacity, and which functions according to a FIFO (First In, First Out) discipline. When the source emits information packets randomly, the role of the breakthrough bucket consists in regulating the burst arrival flow, and let it pass with a constant rate. This method presents a drawback, which is that the breakthrough bucket can not allow a constant service rate to pass if it is greater than the one given by queue capacity. To face this problem, we have thought about using another variant more efficient and which has for mission to regulate the burst of arrivals, as well as to transmit them to destinations. We have applied this variant on the network that we will give.

The variant of token breakthrough bucket was modelled as shown in the next figure :

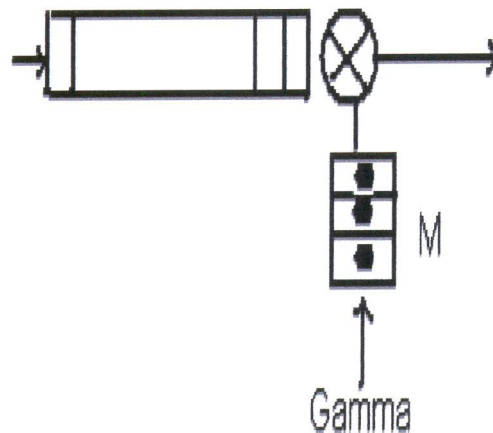


FIG. 1 - Token Bucket

We have two queues. The breakthrough bucket one, in which the arriving information packets will be put. The second queue is a pool of tokens of size

M, with a token arrival rate of γ . Functioning principle is divided into two cases : firstly, if we have clients in the breakthrough bucket queue, and no tokens in the pool, then the arriving clients will be put in the first queue and wait tokens arrival. Secondly, if we have tokens and no arriving clients, then arriving tokens will be stored in their pool if the capacity M is not already reached. The only case, in which we can have information packets at the output, is when there are, at once, at least a client and at least a token. In this case, the first client in the breakthrough queue take a token and pass to the router, this one will send it to its destination. This principle allow to a burst of data to pass to the router queue, to join the destination network.

We will now try to find the optimal parameters of the break through bucket (M, γ) which ensure an optimal control of congestion. Determining the optimal parameters, will allow us to reach wanted QoS (Quality of Service).

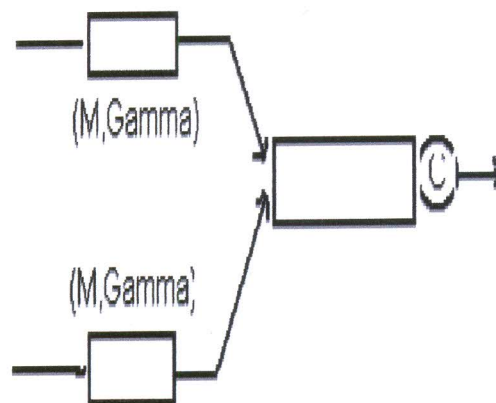


FIG. 2 = Connection with N token buckets

Modelling this system has given the following non linear mathematical

problem :

$$\begin{cases} \text{Min } \sum_{i=1}^n \gamma_i, \\ e^{-\theta_i M_i} \leq \zeta, & ; \\ M_1 + M_2 + \dots + M_n = B, & ; \\ m_i \leq \gamma_i \leq r_i, & ; \\ \theta_i = \frac{r_i(\theta_i - m_i)\alpha_i}{(r_i - \gamma_i)(r_i - m_i)\gamma_i}, & i=1, \dots, N, \end{cases}$$

Resolving this problem using the K-K-T theorem has given as results :

$$M^* = \frac{B}{K}$$

$$\gamma^* = \frac{-a_2 + \sqrt{a_2^2 - 4a_1a_3}}{2a_2}$$

with

$$a_1 = r - m$$

$$a_2 = \frac{\alpha r B}{-K \zeta} = r(r - m)$$

$$a_3 = (m \alpha B r) / K \log(\zeta)$$

Results of the variant we have chosen, were obtained by applying it to a test network, this network is shown by the figure below :

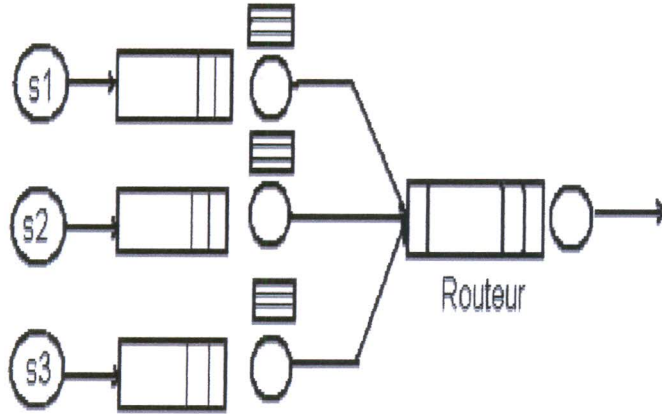


FIG. 3 - Networks Test

Emission sources S1, S2 and S3 are supposed :

- S1 emits information packets with an exponential mode ;
- S2 emits information packets periodically ;
- S3 is considered of type ON / OFF.

To visualize our results we have conceived a software that we called SIMTOK , that can show results of a simulation (see fig 4 and 5).

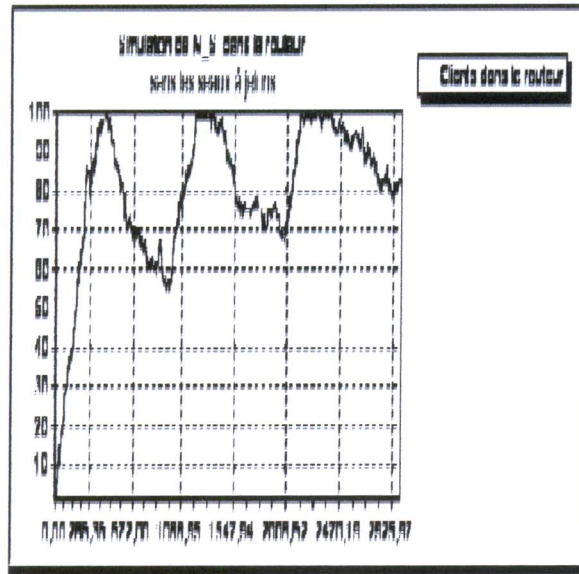


FIG. 4 -

The part reserved for routing using the distance vector in our software, allows us to give procedures used to update routing tables if need be. We will suppose that at a given moment, an update request is issued by a node of the network, or by the entire network, the network status at this moment is given by the next figure :

So, we obtain the following results : Every line of the matrix 1 and 2 represents a distance vector of the node given by the line index. Every box $[i,j]$ of the matrix 1 gives the first node in the shortest path between node i and node j . As for matrix 2, every box $[i,j]$ gives the cost of this path.

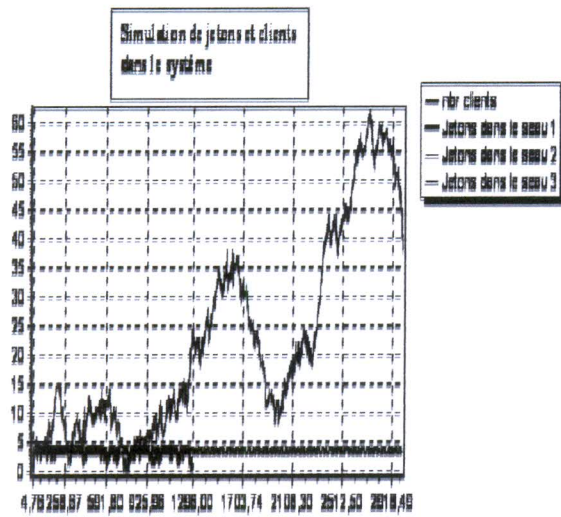


FIG. 5 =

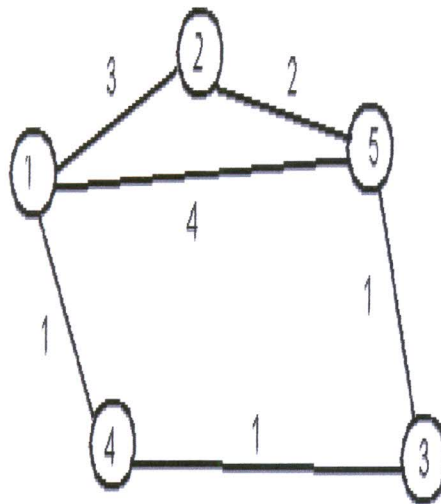


FIG. 6 = Network Test

	1	2	3	4	5
1	0	2	4	4	4
2	1	0	5	1	5
3	4	5	0	4	5
4	1	1	3	0	3
5	3	2	3	3	0

TAB. 1 = Matrix 1

	1	2	3	4	5
1	0	3	2	1	3
2	3	0	3	4	2
3	2	3	0	1	1
4	1	4	1	0	2
5	3	2	1	2	0

TAB. 2 = Matrix 2

Modèles analytiques pour l'évaluation des performances d'un réseau pair-à-pair (cas du Bittorent)

Arkoub D., Krouri Y., Aïssani D. et Amad M.

lamos_bejaia@hotmail.com

Résumé Ce travail explore une problématique commune à tous les réseaux informatiques : l'évaluation des performances d'un réseau ; une étape cruciale pour l'étude de ses caractéristiques.

Notre étude est orientée vers les réseaux Pair-à-Pair qui, en quelques années, ont offert un tout autre paradigme de conception des réseaux. En effet, l'architecture client-serveur est abandonnée au profit d'un réseau de machines, qui cumulent chacune les deux fonctions de client et de serveur.

Nous avons modélisé une situation réelle des réseaux P2P (le Bittorent) en ayant opté pour les réseaux de files d'attente. Ce modèle analytique se base sur des facteurs de performance tels : le nombre moyen de fichiers dans le système, le temps moyen d'attente pour le transfert de données nous avons eu recours à la simulation pour valider nos résultats analytiques.

Mots clés : Systèmes informatiques, Réseaux Pair-à-Pair, Evaluation de Performances, Réseaux de Files d'attente, Modélisation, Simulation, Bittorent.

9.1 Introduction

Le succès d'Internet n'est plus à prouver, cependant un problème pointe à l'horizon : la quantité de données à distribuer sur Internet est en perpétuelle croissance. La partie du problème la plus importante est sans doute la distribution de gros fichiers (souvent du contenu multi-média-vidéo, musique . . .) à un grand nombre d'utilisateurs.

Eu fait, lorsqu'on y regarde de plus près, le problème ne vient pas de la quantité de données mais plutôt du mode de distribution. Eu effet, la méthode traditionnelle pour distribuer

du contenu est de le mettre sur un serveur web. Les utilisateurs voulant télécharger un fichier à partir d'un de ces serveurs doivent se partager la bande passante de celui-ci. Plus le fichier est populaire, plus la charge est importante sur le serveur.

On assiste, ces dernières années, dans le monde de l'informatique, au développement important d'un type de réseaux informatiques : les réseaux Pair-à-Pair. Son essor est tel qu'il est aujourd'hui l'un des principaux acteurs de l'Internet.

L'architecture classique client-serveur est abandonnée au profit d'une conception différente du partage des ressources. En effet dans un réseau Pair-à-Pair, chaque client est un serveur, il participe activement au réseau. Il peut ainsi collaborer à l'élaboration d'une tâche collective.

Dans le présent travail, nous nous focaliserons sur des facteurs de performance principaux des réseaux dans le but de les comprendre et ainsi de pouvoir envisager des solutions pour les améliorer.

Nous nous sommes intéressés au modèle Bittorrent(très populaire pour le partage des fichiers multimédia) en nous basant sur les réseaux de files d'attente pour en évaluer les performances.

2 Description du Bittorrent

Bittorrent[9] est un protocole de partage de fichiers. En fait, c'est un système peer to peer comme eDonkey. Les fichiers sont échangés par morceaux entre les utilisateurs. Ils sont identifiés par une URL (*ex : <http://www.levien.com/mandrake9.1.torrent>*) afin de s'adapter parfaitement à la navigation Internet. Pour réunir les informations et les utilisateurs de BitTorrent, il existe des serveurs centraux, appelés tracker. Ces trackers permettent de coordonner le téléchargement des pièces de chaque fichier aux utilisateurs du programme. Tous les utilisateurs participent au partage en recevant (*download*), mais aussi en envoyant (*upload*) des parties du fichier. Ce fonctionnement permet d'optimiser le transfert des données et de limiter la charge sur le fichier source.

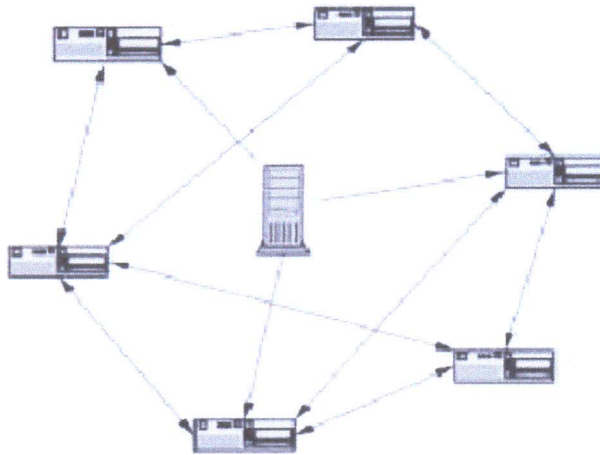


FIGURE 1 – Topologie de Bittorrent

Les caractéristiques du Bittorrent sont :

1. Son objectif étant uniquement la distribution de contenu (*le plus efficacement possible et en limitant la bande passante du fournisseur initial*) et non pas la localisation de ce dernier, BitTorrent ne contient pas de fonctionnalité pour trouver un fichier, comme dans les logiciels peer to peer classiques.
2. A la différence de beaucoup d'autres protocoles peer to peer, les clients ne font pas partie d'un réseau global (*comprenant tous les utilisateurs du protocole*), mais sont plutôt regroupés par fichier : chaque fichier distribué verra son propre groupe d'utilisateurs. Chacun de ces groupes sera organisé autour d'un composant central, le tracker, qui sert en quelque sorte d'annuaire dynamique de clients intéressés par la distribution d'un seul fichier.
3. Il traite les gros fichiers comme un certain nombre de morceaux et chaque morceau peut être téléchargé d'une source différente.
4. Il contient divers mécanismes pour assurer son efficacité, dont notamment un mécanisme qui récompense les clients qui envoient des données aux autres.

3 Description du modèle de réseau de files d'attente (réseau de Jackson)

L'étude des modèles de phénomènes d'attente s'avère souvent complexe, car le flot des arrivées dans une station est la superposition des flots de requêtes en provenance des autres noeuds du réseau. Ce flot n'est pas poissonnien et les caractéristiques du réseau sont difficiles à déterminer. Il existe cependant une classe de modèles pour lesquels on arrive à obtenir une solution simple : ce sont les réseaux exponentiels pour lesquels la distribution de probabilité des états s'obtient sous forme de produit. Cette propriété remarquable des réseaux exponentiels est une conséquence du théorème de Burke selon lequel les flots des départs sont poissonniens.

En 1957, Jackson a décrit une classe générale de réseaux qui admettent une modélisation markovienne. Ces réseaux possèdent une propriété remarquable permettant de simplifier leur étude en les étudiant station par station. Soit le réseau de Jackson constitué de N files comme indiqué sur la figure suivante.

Ce réseau ne possède que des serveurs exponentiels de taux μ_i , $i = 1, \dots, N$. Les

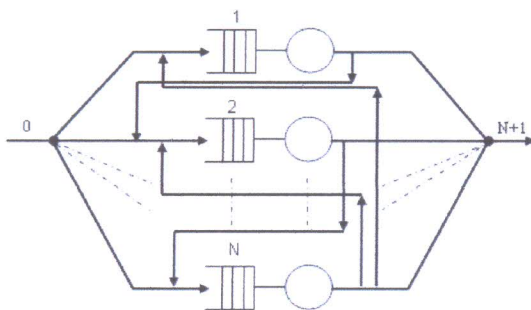


FIGURE 2 – Réseau de Jackson ouvert

clients sont acheminés dans le réseau suivant des probabilités de routage p_{ij} (*probabilité en sortant de la file i pour aller vers la file j*). Le flux total des clients arrivant dans le système forme un processus de Poisson de paramètre λ avec une seule classe de clients, une capacité de stockage illimitée à toutes les stations, une discipline de service FIFO pour toutes les files. Les réseaux de Jackson possèdent la propriété d'avoir une distribution stationnaire en forme produit pour les probabilités conjointes des longueurs des files d'attente.

Les réseaux de Jackson ouverts Un réseau de Jackson ouvert est un réseau formé par l'interconnexion de n files de type $-/M/m$ (que l'on suppose numéroté de 1 à n) où les clients arrivent selon un processus de Poisson et se déplacent en suivant des règles de routage markoviennes. Plus précisément les clients arrivent de l'extérieur du système selon des processus de Poisson indépendants. Le taux d'arrivée dans la file i étant constant et égal à λ_i ($1 \leq i \leq n$), chaque station i du réseau est régie par une discipline FIFO possède un nombre fini m_i de serveur et fourni des traitements dont les durées sont indépendantes et identiquement distribuées selon une loi exponentielle de paramètre μ_i après avoir complété son service dans une file i . Un client est envoyé dans une file j avec la probabilité p_{ij} et quitte le système avec la probabilité p_{i0} .

L'état d'un réseau de Jackson à l'instant t est donnée par le vecteur $X(t) = (X_1(t), \dots, X_n(t))$ ou $X_i(t)$ est le nombre de clients présents dans la file i à l'instant t .

- Le taux d'arrivée λ_i dans la file i est donné par :

$$\lambda_i = \lambda p_{0i} + \sum_{j=0}^n \lambda_j p_{ji}$$

- La condition de stabilité du système est $\lambda_i < \mu_i$, $i=1,2,\dots$

4 Application à Bittorent

A la différence de beaucoup d'autres protocoles peer to peer, dans Bittorent, les clients (*nœuds*) ne font pas partie d'un réseau global (*comprenant tous les utilisateurs du réseau*), mais sont plutôt regroupés par fichier : chaque fichier distribué verra son propre groupe d'utilisateurs. Chacun de ces groupes sera organisé autour d'un composant central, le tracker, qui sert en quelque sorte d'annuaire dynamique de clients intéressés par le téléchargement de ce fichier. Après avoir été organisés sous une topologie donnée, les nœuds peuvent commencer le téléchargement des paquets, sachant qu'un client peut être associé à plusieurs téléchargements à la fois.

Nous considérons, pour notre étude l'exemple ci-dessous, un réseau Bittorent constitué de sept nœuds collaborant au téléchargement d'un fichier (*sous forme d'un ensemble de paquets de même taille*) reçu par le nœud 1, qui le fait transférer aux autres nœuds intéressés. Nous allons considérer que les nœuds qui reçoivent les paquets du nœud 1 acceptent de collaborer en les transmettant aux autres. Donc, on dit qu'un paquet ne circule plus dans le réseau quand il est transféré à tous les nœuds. Il est alors intéressant de suivre le transfert des paquets entre les nœuds jusqu'à ce qu'ils atteignent le dernier nœud à savoir le nœud 7,

qui se trouve dans la dernière position. Le graphe formé par les nœuds montre l'existence de trois chemins possibles entre le nœud 1 et le nœud 7 illustrés dans la figure ci-dessous :

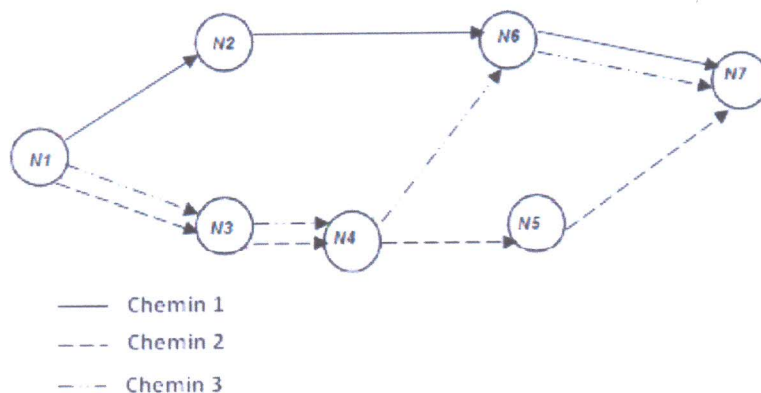


FIGURE 3 – Graphe formé par les noeuds

Le choix du chemin pour chaque paquet se fait d'une manière aléatoire. Il est donc nécessaire de prendre en considération les trois chemins possibles qui forment trois réseaux de files d'attente, où chaque nœud constitue un système d'attente.

La matrice des probabilités de transmission des paquets entre les noeuds s'écrit comme suit :

$$p_{ij} = \begin{pmatrix} - & 1 & 6/8 & 6/8 & 6/8 & 4/8 & 4/8 & 4/8 \\ 4/8 & - & 2/8 & 2/8 & 0 & 0 & 0 & 0 \\ 6/8 & 0 & - & 0 & 0 & 0 & 2/8 & 0 \\ 6/8 & 0 & 0 & - & 2/8 & 0 & 0 & 0 \\ 4/8 & 0 & 0 & 0 & - & 2/8 & 2/8 & 0 \\ 6/8 & 0 & 0 & 0 & 0 & - & 0 & 2/8 \\ 6/8 & 0 & 0 & 0 & 0 & 0 & - & 2/8 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & - \end{pmatrix}$$

4.1 Calcul des principaux facteurs de performance pour les chemins possibles entre le noeud 1 et le noeud 7

Chemin 1 : Noeud 1 → Noeud 2 → Noeud 6 → Noeud 7

- L_{c1} = le nombre moyen de paquets dans le réseau constitué par les noeuds de chemin 1

$$= L_1 + L_2 + L_6 + L_7 = 24$$

avec L_1, L_2, L_6, L_7 est le nombre moyen de paquets dans les noeud 1,2,6,7

- L_{qc1} = le nombre moyen de paquets en attente d'être servis dans les noeuds de chemin 1

$$= L_{q1} + L_{q2} + L_{q6} + L_{q7} = 12$$

avec $L_{q1}, L_{q2}, L_{q6}, L_{q7}$ le nombre moyen de paquets en attente dans chacun des noeuds 1,2,6,7

- W_{c1} = est le temps moyen passé (séjour) dans l'ensemble des noeuds de chemin 1

$$= W_1 + W_2 + W_6 + W_7 = 1.91 \mu s$$

avec W_1, W_2, W_6, W_7 le temps moyen de séjour dans chacun des noeuds 1,2,6,7

- W_{qc1} = le temps moyen passé en attente dans l'ensemble des noeuds du chemin 1

$$= W_{q1} + W_{q2} + W_{q6} + W_{q7} = 1.52 \mu s$$

avec $W_{q1}, W_{q2}, W_{q6}, W_{q7}$ le temps moyen passé en attente dans chacun des noeuds 1,2,6,7

Chemin 2 : Noeud 1 → Noeud 3 → Noeud 4 → Noeud 5 → Noeud 7

- L_{c2} = le nombre moyen de paquets dans le réseau constitué par les noeuds du chemin 2

$$= L_1 + L_3 + L_4 + L_5 + L_7 = 21$$

avec L_1, L_3, L_4, L_5, L_7 le nombre moyen de paquets dans chacun des noeuds 1,3,4,5,7

- L_{qc2} = le nombre moyen de paquets en attente d'être servis dans les noeuds de chemin 2

$$= L_{q1} + L_{q3} + L_{q4} + L_{q5} + L_{q7} = 16$$

avec $L_{q1}, L_{q3}, L_{q4}, L_{q5}, L_{q7}$ le nombre moyen de paquets en attente dans le noeud 1,3,4,5,7

- W_{c2} = est le temps moyen passé (séjour) dans l'ensemble des noeuds de chemin 2

$$= W_1 + W_3 + W_4 + W_5 + W_7 = 2.53 \mu s$$

- avec W_1, W_3, W_4, W_5, W_7 le temps moyen de séjour dans chacun des noeuds 1,3,4,5,7
- W_{qc2} = est le temps moyen passé en attente dans l'ensemble des noeuds de chemin 2

$$=W_{q1} + W_{q3} + W_{q4} + W_{q5} + W_{q7}=2.08\mu s$$

avec $W_{q1}, W_{q3}, W_{q4}, W_{q5}, W_{q7}$ le temps moyen passé en attente dans chacun des noeuds 1,3,4,5,7

Chemin 3 : Noeud 1 → Noeud 3 → Noeud 4 → Noeud 6 → Noeud 7

- L_{c2} =le nombre moyen de paquets dans le réseau constitué par les noeuds de chemin 2

$$=L_1 + L_3 + L_4 + L_6 + L_7 = 20$$

avec L_1, L_3, L_4, L_6, L_7 le nombre moyen de paquet dans chacun des noeuds 1,3,4,6,7

- L_{qc3} =le nombre moyen de paquets en attente d'être servis dans les noeuds de chemin 3

$$= L_{q1} + L_{q3} + L_{q4} + L_{q6} + L_{q7}=15$$

avec $L_{q1}, L_{q3}, L_{q4}, L_{q6}, L_{q7}$ le nombre moyen de paquets en attente dans les noeuds 1,3,4,6,7

- W_{c3} =est le temps moyen passé (séjour) dans l'ensemble des noeuds de chemin 2

$$=W_1 + W_3 + W_4 + W_6 + W_7=2.41\mu s$$

avec W_1, W_3, W_4, W_6, W_7 le temps moyen de séjour dans chacun des noeuds 1,3,4,6,7

- W_{qc3} = est le temps moyen passé en attente dans l'ensemble des noeuds de chemin 2

$$=W_{q1} + W_{q3} + W_{q4} + W_{q6} + W_{q7}=1.92\mu s$$

avec $W_{q1}, W_{q3}, W_{q4}, W_{q6}, W_{q7}$ le temps moyen passé en attente dans chacun des noeuds 1,3,4,6,7

Dans cet exemple d'application, nous avons tenu compte du fait que les noeuds acceptent de collaborer en envoyant leurs paquets reçus à d'autres clients intéressés. Ce qui est un cas idéal dans la phase de téléchargement.

On prend à titre d'exemple le noeud 6 qui reçoit les paquets des noeuds 2 et 4. Cela augmente le nombre de paquets reçus par celui-ci (qui équivaut à 4 paquets/us).

Dans le cas où, un seul nœud accepte de partager avec le nœud 6, le nombre de paquets reçus se réduira à 2 paquets/us.

A partir des résultats analytiques sur le temps de séjour moyen du paquet dans le système qui se traduit par le temps moyen qu'il faut pour qu'un nœud reçoive un paquet. On remarque que plus le nœud est loin de la source plus le temps nécessaire pour son téléchargement augmente. Cela est dû au passage de ce paquet par des nœuds intermédiaires.

A titre d'exemple, les nœuds 4 et 2. 2 est directement connecté à la source, 4 lui est à une distance de longueur deux de la source. Les temps moyens de téléchargement d'un paquet pour 2 et 4 sont respectivement : $1.025 \text{ paquet}/\mu s$ et $1.75 \text{ paquets}/\mu s$.

La vitesse de téléchargement d'un paquet par un noeud dépend de sa capacité et de celle du noeud auquel il est connecté. Si celle-ci est grande, le temps de téléchargement se réduit. A titre d'illustration les noeuds 6 et 5 téléchargent les paquets à partir du noeud 4. Les capacités de ces noeuds sont dans l'ordre $\mu_6 = 11, \mu_5 = 8, \mu_4 = 8$. On a $W_6 = 0.83\mu s, W_5 = 1\mu s$, on voit bien que le temps de téléchargement par le noeud 6 est plus petit que par le noeud 5 même s'ils téléchargent à partir de la même source. Cela est dû au fait que $\mu_6 > \mu_5$.

5 Simulation du modèle analytique

Dans cette section nous avons eu recours à la simulation comme l'illustrent les graphes suivants :

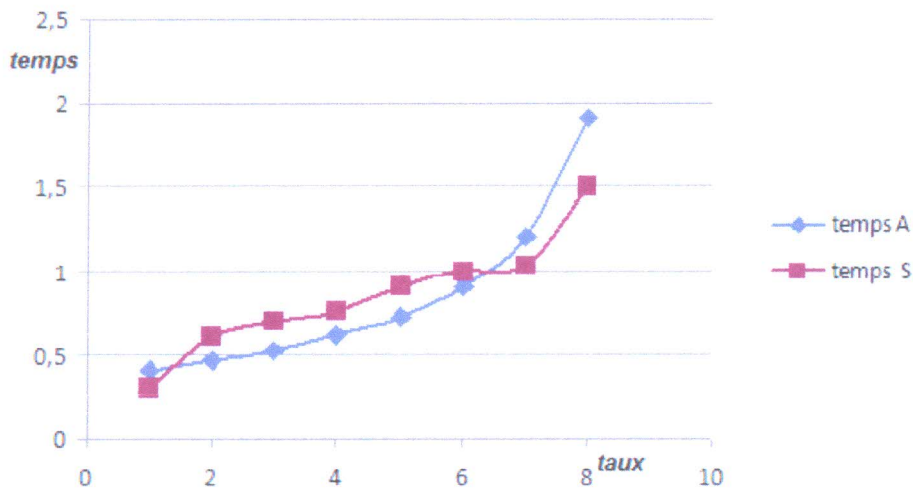


FIGURE 4 – La variation des temps moyens de séjour en fonction de $\lambda(\text{chemin1})$

Ce graphique représente la variation du temps de séjour d'un paquet dans le chemin 1 en fonction du taux d'arrivée variant entre 1 et 8 dans le réseau.

On voit à partir de ce graphique que le temps de séjour d'un paquet s'accroît avec l'augmentation du flot entrant ce qui montre que plus il y a de paquets dans le réseau plus le temps de téléchargement s'accroît.

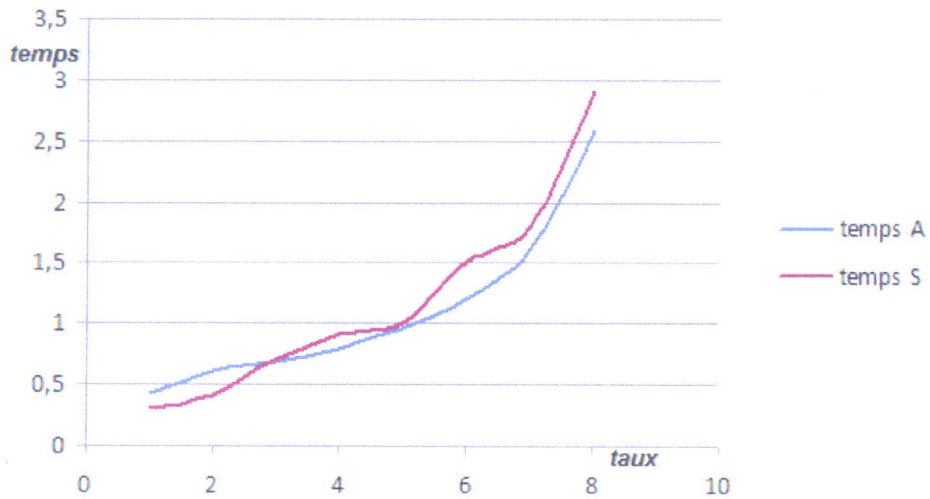


FIGURE 5 – La variation des temps moyens de séjour en fonction de $\lambda(\text{chemin2})$

Ce graphe représente la variation du temps de séjour d'un paquet dans le chemin 2 en fonction du taux d'arrivée variant entre 1 et 8 dans le réseau.

On voit à partir de ce graphe que le temps de séjour d'un paquet s'accroît avec l'augmentation du flot entrant ce qui montre que plus il y a de paquets dans le réseau plus le temps de téléchargement s'accroît.

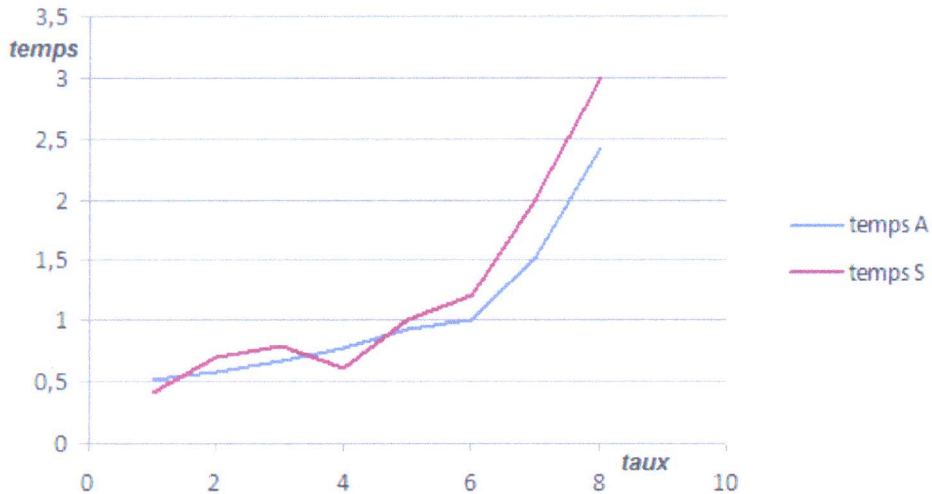


FIGURE 6 – La variation des temps moyens de séjour en fonction de $\lambda(\text{chemin3})$

Ce graphe représente la variation du temps de séjour d'un paquet dans le chemin 3 en fonction du taux d'arrivée variant entre 1 et 8 dans le réseau.

On voit à partir de ce graphe que le temps de séjour d'un paquet s'accroît avec l'augmentation du flot entrant ce qui montre que plus il y a de paquets dans le réseau plus le temps de téléchargement s'accroît.

Conclusion Dans cette partie, nous nous sommes tournés vers une autre approche pour l'évaluation des performances du réseau étudié plus haut : la simulation.

Les résultats obtenus étant proches de ceux obtenus dans le chapitre précédent, cela nous permet d'affirmer que la méthode analytique proposée modélise correctement le réseau Bittorrent pris en exemple.

Références Bibliographiques

- [1] A.Ruegg, Processus Sthocastiques avec application aux phénomènes d'attente, Presses polytechniques romandes, 1989.
- [2] J-F Hêche, T M. Liebling, D de Werra, Recherche opérationnelle pour l'ingénieur, Presses polytechniques romandes, 2003
- [3] F.Valois, Réseaux de files d'attentes à forme produit, 2012.
- [4] M.Amad, Performance Et Optimisation des architectures P2P pour les applications à grande échelle, Univ de Béjaia (thèse de Doctorat), 2012.
- [5] F. Lesueur, Autorité de certification distribuée pour des réseaux pair-à-pair structurés, Univ de Rennes, 2010
- [6] S.Ktar, Interconexion et routage dans les réseaux P2P, Univ du Havre (thèse de Doctorat), 2010.
- [7] C.Duvallet, Architectures et protocoles des réseaux p2p, Univ du Havre, 2008.
- [8] G.MENTEN, bittorrent chaîne manquante des protocole p2p, <http://bitconjurer.org.bittorrent>.
- [9] Bittorrent <http://bitconjurer.org.bittorrent>.
- [10] R.Lo Cigo, Simulation and analysis of "interleave of protocol" fordistributing video streaming in p2p, Univ de Torino (thèse de Doctorat), 2007.
- [11] Kazaa. <http://www.kazaa.com/>.
- [12] Napster. <http://www.napster.com/>.
- [13] K.R.Krishna, B.Sikdar, A Queuing Model for Evaluating the Transfer Latency of Peer-to-Peer Systems, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, 2010.
- [14] J.Siebert, Impact du comportement des utilisateurs dans les réseau p2p, modélisation et simulation multi-agents, Univ Paris-Sud (thèse de Doctorat), 2011.
- [15] L.G.ERICE, Réseau p2p hiérarchique : design et application, Ecole nationale supérieure des télécommunications,2004.
- [16] L.HAN, Z.LEI, Modeling structured peer to peer systems, The Journal Of China Universities Of Posts and Telecommunications, 2006.
- [17] A.Monsellato, A stochastic model for distributed real time streamingEvolution of the network structure, Univ de Pavia Italie, 2010.
- [18] J.Kurose, Z.Ge, R.F.Daniel, J.Sharad, T.Don, Modeling Peer-Peer File Sharing Systems, Univ du Massachusetts, 2003.
- [19] M. Amad, Découverte et Localisation De Services en Mode P2P, Univ de Béjaïa (thèse

de Magistère), 2005.

[20] K. Xu, Modeling and Analysis Of P2P Live Video Streaming, Univ de l'Illinois, 2009.

Amélioration de la norme IEEE 802.11^e mode EDCA sous un canal non idéal

Bachiri L., Abdi N., Aroul K. et Aïssani D.

Research Unit LaMOS (Modeling and Optimization of Systems)
lamos_bejaia@hotmail.com, Bachiri.lina@gmail.com, abdi--nass89@hotmail.com et milina_87@hotmail.fr

Résumé La norme IEEE 802.11e est la version améliorée de standard IEEE 802.11^e pour fournir la qualité de service dans les réseaux sans fil. Dans IEEE 802.11^e, la principale méthode d'accès dans la couche MAC est la méthode amélioré d'accès au canal distribué (EDCA), qui fournit un ordre de priorité en différenciant l'espace d'arbitrage Inter Frame (AIFS), les valeurs de la taille minimale et maximale de la fenêtre de contention et la limite d'opportunité de transmission (TXOPlimit). Dans ce travail, nous proposons une modification du réseau IEEE 802.11^e EDCA sous un canal bruité appelé A-EDCA (Amélioration de EDCA), pour réduire les retransmissions inutiles dans le cas de l'échec de transmission dû à l'erreur dans le canal. Puis, nous présentons un modèle analytique pour l'analyse des performances de A-EDCA avec la transmission en rafales sous un canal bruité et des conditions d'un trafic non saturé. Ce modèle capte toutes les caractéristiques de la QoS, à savoir l'AIFS, la taille minimale de la fenêtre de contention, la taille maximale de la fenêtre de contention, la collision virtuelle et le TXOPlimit. Nous comparons les performances de EDCA avec les performances de A-EDCA.

Key words: IEEE 802.11^e, EDCA, canal bruité, trafic non saturé, débit de transmission, chaîne de Markov

10.1 Introduction

La norme IEEE 802.11 est actuellement l'une des technologies les plus populaires d'accès au canal. La fonction de coordination distribuée (DCF) est l'un des mécanismes d'accès aux canaux dans IEEE 802.11 DCF ne supporte aucun type d'accès prioritaire au support sans fil. En conséquence, DCF fournit seulement le service meilleur effort, et il n'y a pas de mécanisme pour fournir un meilleur service pour le trafic en temps réel. Comme une solution à ce problème, le 'Groupe de travail IEEE 802.11 E' a proposé une nouvelle norme de la couche MAC, à savoir l'IEEE 802.11^e[24]. Dans IEEE 802.11^e, EDCA est le mécanisme fondamental.

De nombreux efforts de recherche ont été menés pour étudier les performances de la norme IEEE 802.11^e EDCA par la méthode analytique ou par simulation. La plupart d'entre elles supposent la condition de canal idéal, ce qui signifie que les corruptions de paquets sont uniquement dûes aux collisions. Dans ce qui suit, nous fournissons un bref résumé des modèles proposés pour les fonctions DCF et EDCA dans la littérature.

Bianchi [1], a proposé un modèle analytique simple pour calculer le débit normalisé. Pour conduire un retard de saturation Ziouva et Antonakopoulous [25] ont amélioré le modèle de Bianchi. Dans [15], les auteurs ont proposé une nouvelle méthode de temporisation pour améliorer la performance de la norme IEEE 802.11 DCF. Dans [3], les auteurs fournissent une amélioration de modèle pour IEEE 802.11 sous un canal bruité.

Les auteurs dans [2], ont proposé un nouveau modèle analytique pour analyser la performance de EDCA. Dans [4], les auteurs modifient le modèle de Ziouva et Antonakopoulous [25] et étendent le modèle pour soutenir EDCA. Dans [5, 6], les auteurs ont analysé IEEE 802.11^e EDCA sous des condition de non saturation. Dans [7, 8], les auteurs ont étudié l'efficacité des transmissions en rafale avec accusés de réception de bloc. Dans [9, 10, 11], les auteurs ont proposé un modèle analytique pour étudier l'effet des transmissions en rafale. Dans [12], un modèle

analytique est proposé pour évaluer la performance de 802.11^e dans le cas saturé. Les auteurs dans [13] analysent le comportement temporel de la fonction EDCA, quand il est utilisé pour supporter le trafic en temps réel. Dans [14], les auteurs évaluent, par des simulations, les avantages découlant de l'adoption de la norme IEEE 802.11^e dans la prestation multimédia sur les réseaux locaux sans fil. Dans [3], les auteurs fournissent une étendue pour l'analyse analytique de IEEE 802.11 sous le canal bruité. Dans [16, 17], les auteurs étudient l'influence de l'erreur de transmission sur la performance de la couche MAC. Dans [19], les auteurs analysent mathématiquement la performance de débit et de retard du programme ACK de IEEE 802.11^e dans un environnement de canal erroné. Dans [20], Basé sur les travaux précédents, notre proposition résume le modèle de Xiaos [12] et le modèle de Kongs [2], les auteurs ont proposé un modèle basé sur le formalisme de chaîne de Markov pour l'analyse de EDCA. Le modèle développé reflète les procédures de backoff et d'accès en tenant compte du blocage de temporisateur de backoff qui inclut le temps externe de collision, le temps de transmission, les paramètres AIFS et CW. Dans [21], les auteurs de ce document, Ont d'abord développé un modèle analytique pour dériver l'évaluation moyenne de retard et de perte pour 802.11^e sur un canal non idéal. Comme deuxième contribution, un ajustement pour le nombre maximum de retransmissions est proposé, pour maintenir les exigences de qualité de service pour le trafic en temps réel. Dans [22], les auteurs proposent un modèle analytique pour EDCA sous un canal idéal. Dans [23], les auteurs ont étudié l'effet de la limite de transmission d'opportunité sur le temps de transmission.

Dans cet article, nous proposons une amélioration de la norme IEEE 802.11^e EDCA, appelé A-EDCA (Amélioration de EDCA) afin de réduire les retransmissions inutiles dans le cas des erreurs de transmission. Puis, nous présentons un modèle analytique pour l'analyse de la performance de A-EDCA sous un canal bruité et dans des conditions non saturées.

Le reste de cet article est organisé comme suit : Dans la section 2, nous donnons une brève introduction à la fonction IEEE 802.11^e EDCA. Nous proposons une amélioration de EDCA dans la section 3. Ensuite, nous proposons un modèle analytique pour A-EDCA dans la section 4. Les résultats numériques sont donnés dans la section 5. Dans cette section, les performances de A-EDCA sont comparées aux performances de EDCA. Enfin, la section 6 conclut cet article.

10.2 Aperçu général des DCF et EDCA

10.2.1 Distributed coordination function (DCF)

DCF est un mécanisme d'accès distribué à la couche MAC proposé dans la norme IEEE 802.11 [28]. L'accès au support est contrôlé par l'utilisation d'espaces intertrames, ou IFS (Inter-Fram Spacing), qui correspondent à l'intervalle de temps entre la transmission de deux trames. Les intervalles IFS sont des périodes d'inactivité sur le support de transmission. une station qui veut émettre une trame de données, doit écouter le canal durant un intervalle de temps appelé DIFS, plus un temps additionnel de backoff. La station peut transmettre sa trame lorsque le canal reste inactif pendant toute cette période. Le temps de backoff est uniformément choisi dans l'intervalle (0, CW), CW est initialement réglé à CW_{min} et doublé après chaque transmission non réussie jusqu'à ce qu'une valeur maximale de la taille de la fenêtre de contention, CW_{max} est atteinte. Pour chaque slot time, le compteur de backoff est décrémenté par un. La station transmet une trame quand son temps backoff atteint zéro. Pour chaque réception réussie d'une trame, la station réceptrice envoie une trame d'acquittement à l'émetteur après un SIFS. La source retransmet la trame si elle ne reçoit pas l'ACK dans un ACKTimeout spécifié. DCF définit aussi un mécanisme optionnel appelé RTS/CTS (Request-to-Send/Clear-to-Send) dans le but de résoudre le problème des stations cachées.

10.2.2 Enhanced distributed channel access (EDCA)

EDCA représente une extension de la méthode d'accès DCF par l'introduction de la différenciation de services qui est basée sur l'utilisation de trois paramètres : AIFS, les fenêtres de contention minimale et maximale (CW_{min} et CW_{max}) et $TXOP_{limit}$. EDCA définit quatre AC où chaque AC correspond à un niveau de priorité et pouvant être utilisé par un type de trafic comme indiqué dans la table 1. la valeur prioritaire d'utilisateur est définie dans les spécifications de IEEE 802.11D [27]. EDCA définit les quatre AC suivante pour chaque station : AC_{VO} (voix) avec une priorité plus haute, AC_{VI} (vidéo), AC_{BE} (best effort) et AC_{BK} (Background) avec la priorité plus basse. Chaque AC dans une station se comporte comme une station virtuelle. Lorsqu'une collision se produit entre les différents AC_i au sein de la même station, l'AC de priorité supérieure peut transmettre, tandis que la plus faible priorité AC souffre d'une collision virtuelle.

TABLE 10.1: Relation entre les priorités et les catégories d'accès dans EDCA

priority	User priority	802.11D designation	Access Category 802.11 ^e	Designation
lowest	1	BK	AC_{BK}	Background
	2	_	AC_{BK}	Background
	0	BE	AC_{BE}	Best effort
	3	EE	AC_{BE}	Best effort
	4	CL	AC_{VI}	Video
	5	VI	AC_{VI}	Video
	6	VO	AC_{VO}	Voice
highest	7	NC	AC_{VO}	Voice

10.3 Amélioration de EDCA sous un canal bruité

Dans IEEE 802.11^e EDCA, une AC ne fait pas de différence entre une collision et une erreur de transmission. Le comportement d'une AC après une collision ou une erreur de transmission est le même, elle double la fenêtre de contention. Quand l'AC double sa fenêtre de contention, c'est pour éviter les collisions. Donc si le paquet est perdu à cause des erreurs de transmission, le doublement de la fenêtre de contention n'est pas nécessaire. Pour améliorer EDCA, nous proposons de traiter les erreurs de transmission avant les collisions. Donc notre solution est basée sur la séparation des erreurs de transmission et les collisions. Dans le cas des erreurs de transmission, on ne double pas la fenêtre de contention par contre dans les collisions, on double la fenêtre de contention, ce qu'on a nommé A-EDCA (Amélioration de EDCA). Dans ce qui suit, nous allons présenter le modèle analytique de A-EDCA.

10.4 Modèle analytique

Dans cette section, nous proposons une chaîne de Markov à temps discret à quatre dimensions pour A-EDCA sachant qu'on a utilisé un modèle proposé dans [29] avec l'introduction des erreurs de transmission. Nous supposons un nombre fixe de stations N et chaque station dispose de quatre AC_i , capable d'estimer les métriques de performances dans les conditions de non saturation. L'analyse est effectuée sous un canal non idéal et dans les conditions de non saturation.

10.4.1 Description détaillée de la chaîne de Markov à temps discret

Dans ce modèle, le temps est divisé en slot et chaque état représente une AC_i , ($i = 0, 1, 2, 3$) dans un slot de temps de même longueur. À la fin de chaque slot de temps, il y a un déclenchement d'un évènement qui représente une transition à un autre état [26]. Dans notre modèle, on a modélisé une seule AC , parce que chaque AC est une station virtuelle et elle a sa propre procédure de backoff.

Dans cette chaîne, nous définissons : $s(i, t)$ est le processus stochastique représentant l'étape j du backoff, où $0 \leq j \leq m$ et m est l'étape maximum du backoff

$p(i, t)$ est le processus stochastique représentant l'état de l' AC_i .

$b(i, t)$ est le processus stochastique qui désigne la valeur compteur de backoff à l'instant t et qui est uniformément choisi dans l'intervalle $[0, W_{i,j}]$ où $W_{i,j}$ dépend de la retransmission de Backoff et satisfait la formule suivante : $W_{i,j+1} = 2W_{i,j} + 1$

Le processus $(s(i,t); b(i,t); p(i,t))$ constitue une chaîne de Markov discrète à quatre dimensions sous l'hypothèse que la probabilité de transition P_i soit constante.

À chaque intervalle de temps, l'état de chaque AC_i est décrit par (d, i, j, k) , où

$d = B$: l' AC_i décrémente son intervalle de temps de backoff,

$d = T$: l' AC_i transmet son paquet ,

et $d = E$: pour l'erreur de transmission.

i est juste un indice se tenant pour la classe prioritaire, j représente l'étape de backoff et prend les valeurs $(0, 1, \dots, m)$ et k représente le délai d'attente et prend des valeurs $(0, 1, \dots, W_{i,j} - 1)$.

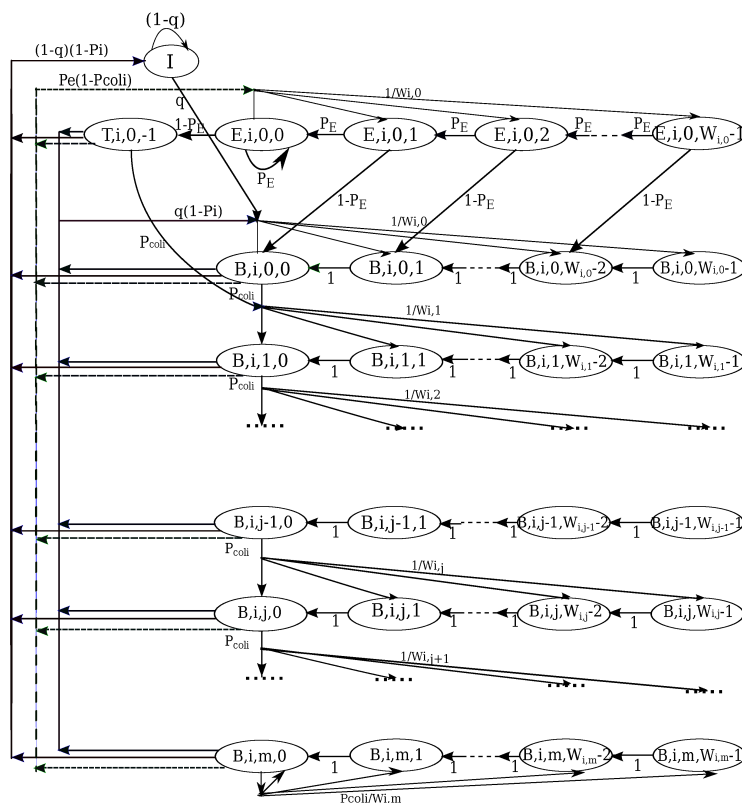


FIGURE 10.1: Modèle de chaîne de Markov.

Probabilités de transitions

Les probabilités de transition dans ce modèle sont décrites comme suit :

$$\left\{ \begin{array}{l}
 p\{B, i, j, k|B, i, j, (k+1)\} = 1, \quad 0 \leq j \leq m, 0 \leq k \leq (w_{i,j} - 1) \\
 p\{B, i, 0, k|B, i, j, 0\} = q(1 - P_i)/w_{i,0}, \quad 0 \leq j \leq m, 0 \leq k \leq (w_{i,0} - 1) \\
 p\{B, i, j, k|B, i, j - 1, 0\} = P_{coli}/w_{i,j}, \quad 0 \leq j \leq m, 0 \leq k \leq (w_{i,j} - 1) \\
 p\{B, i, m, k|B, i, m, 0\} = P_{coli}/w_{i,m}, \quad 0 \leq k \leq (w_{i,m} - 1) \\
 p\{B, i, 0, k|I\} = q/w_{i,0}, \quad 0 \leq k \leq (w_{i,0} - 1) \\
 p\{B, i, 1, k|T, i, 0, -1\} = P_{coli}/w_{i,1}, \quad 0 \leq k \leq (w_{i,1} - 1) \\
 p\{E, i, 0, k|E, i, 0, (k+1)\} = P_E, \quad 0 \leq k \leq (w_{i,0} - 1) \\
 p\{E, i, 0, k|T, i, 0, -1\} = P_e(1 - P_{coli}), \quad 0 \leq k \leq (w_{i,0} - 1) \\
 p\{E, i, 0, k|B, i, j, 0\} = P_e(1 - P_{coli}), \quad 0 \leq j \leq m, 0 \leq k \leq (w_{i,j} - 1) \\
 p\{E, i, 0, 0|E, i, 0, 0\} = P_E, \\
 p\{I|T, i, 0, -1\} = (1 - q)(1 - P_i), \quad 0 \leq k \leq (w_{i,j} - 1) \\
 p\{I|B, i, j, 0\} = (1 - q)(1 - P_i), \\
 p\{T, i, 0, -1|E, i, 0, 0\} = 1 - P_E, \\
 p\{I|I\} = 1 - q,
 \end{array} \right. \quad (10.1)$$

Probabilités des états stationnaires

Soit $b_{i,j,k}$ la probabilité stationnaire de l'état (i, j, k)

$$\left\{ \begin{array}{l}
 b_{B,i,j,0} = P_{coli}^{j-1} (P_{coli} b_{B,i,j,0} + P_{coli} (1 - P_E) b_{E,i,0,0}), \quad 0 \leq j \leq m; \\
 b_{B,i,m,0} = \frac{P_{coli}^{m-1}}{1 - P_{coli}} (P_{coli} b_{B,i,j,0} + P_{coli} (1 - P_E) b_{E,i,0,0}), \\
 \sum_{j \geq 0} b_{B,i,j,0} = \frac{b_{B,i,j,0} + P_{coli} (1 - P_E) b_{E,i,0,0}}{1 - P_{coli}},
 \end{array} \right. \quad (10.2)$$

$$b(I) = \frac{b_{E,i,0,0}}{q} \left[\frac{(1 - q)(1 - P_i)(1 - P_E)^{2W_{i,0}}}{P_e(1 - P_{coli})(1 - P_E^{W_{i,0}})} \right] \quad (10.3)$$

En utilisant la condition de normalisation, on obtient :

$$1 = \sum_{j=0}^m \sum_{k=0}^{W_{i,j}-1} b_{B,i,k,0} + \sum_{k=0}^{W_{i,0}-1} b_{E,i,k,0} + b_{T,i,0-1} + b_I \quad (10.4)$$

$$b_{E,0,0,0} = \left[1 + \frac{W_{i,0}(W_{i,0} + 1)}{2} \left[\frac{(1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})} \right] + \left[\frac{W_{i,0} P_{coli} (1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})} \right] \right. \\
 \left. \left[\frac{1}{2(1 - P_{coli})} \left[1 + \frac{2W_{i,0}(1 - P_{coli} - P_{coli}(2P_{coli})^{m-1})}{(1 - 2P_{coli})} \right] \right] + \right. \\
 \left. \frac{(1 - q)(1 - P_i)(1 - P_E)^2 W_{i,0}}{q P_e (1 - P_{coli})(1 - P_E^{W_{i,0}})} \right]^{-1} \quad (10.5)$$

• Probabilité de transmission des catégories d'accès τ_i :

La probabilité qu'une AC_i accède au canal dans un slot de temps choisi aléatoirement est donné par :

$$\begin{aligned}
\tau_i &= \sum_{j \geq 0} b_{B,i,0,0} + (1 - P_E)b_{E,i,0,0} \\
&= b_{E,i,j,0} \left[\frac{W_{i,0}(1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})(1 - P_{coli})} \right] \\
&= \left[1 + \frac{W_{i,0}(W_{i,0} + 1)}{2} \left[\frac{(1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})} \right] + \left[\frac{W_{i,0}P_{coli}(1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})} \right] \right. \\
&\quad \left. \left[\frac{1}{2(1 - P_{coli})} \left[1 + \frac{2W_{i,0}(1 - P_{coli} - P_{coli}(2P_{coli})^{m-1})}{(1 - 2P_{coli})} \right] \right] + \right. \\
&\quad \left. \frac{(1 - q)(1 - P_i)(1 - P_E)^2 W_{i,0}}{qP_e(1 - p_{coli})(1 - P_E^{W_{i,0}})} \right]^{-1} \left[\frac{W_{i,0}(1 - P_E)^2}{P_e(1 - P_E^{W_{i,0}})(1 - P_{coli})} \right] \quad (10.6)
\end{aligned}$$

• **Probabilité de transmission global τ :**

La probabilité que la station accède au canal est :

$$\tau = 1 - \prod_{i=0}^3 (1 - \tau_i) \quad (10.7)$$

• **Probabilité d'erreur P_e :**

$$P_e = 1 - (1 - BER)^{paquet}$$

• **Probabilité d'erreur P_E :**

La probabilité qu'il y ait une erreur de transmission sur le canal est exprimée en :

$$P_E = (N - 1)\tau(1 - \tau)^{(N-2)}P_e \quad (10.8)$$

• **Probabilité d'échec P_i :**

La probabilité d'échec de transmission est exprimée par :

$$P_i = 1 - (1 - p_{coli})(1 - P_e) \quad (10.9)$$

• **Probabilité de collision :**

Une collision se produit quand une station transmet et une autre parmi $(N - 1)$ autres stations essaye aussi de transmettre. Donc, P_{coli} s'exprime comme suit :

$$P_{coli} = 1 - (1 - \tau^{N-1}) \prod_{i>i'} (1 - \tau_{i'}) P_e \quad (10.10)$$

• **Probabilité qu'il y ait un paquet dans la file est donnée par :**

$$q = 1 - e^{-\lambda T} \quad (10.11)$$

10.4.2 Débit normalisé

Le débit normalisé de chaque AC_i est calculé comme le rapport de temps occupé par la transmission de données utiles sur l'intervalle de temps entre deux transmissions successives [3]. Le débit de chaque AC_i , S_i est exprimé en :

$$S_i = \frac{P_{tr}P_{si}(1 - P_e)E[P]}{(1 - P_{tr})\delta + P_{tr} \sum_{i=0}^3 P_{si}(1 - P_e)T_s + P_{tr}(1 - \sum_{i=0}^3 P_{si})T_c + P_{tr} \sum_{i=0}^3 P_{si}P_eT_e} \quad (10.12)$$

• $E[P]$ est la taille de la charge, T_s est le temps moyen d'une transmission réussite d'une seul trame, T_{TXOP_i} est le temps moyen d'une transmission d'un rafale de trames, T_c est le temps de collision et T_e est le temps d'erreur.

T_s , T_c et T_e en mode de base peuvent être écrites comme suit :

$$\begin{cases} T_s = AIFS + H + P + SIFS + ASK \\ T_{TXOP_i} = Nb_i T_s \\ T_c = H + P + SIFS \\ T_e = AIFS + RTS + SIFS + CTS + H + P + SIFS + ACK_{timeout} \end{cases} \quad (10.13)$$

• $Nb_i = TXOP_i / T_s$, Nb_i est le nombre de paquets transmis dans un rafale pour chaque AC_i .

• **Probabilité de transmission pour N stations :**

Soit P_{tr} la probabilité d'avoir au moins une transmission (réussie ou échouée) dans un slot de temps donné, donc P_{tr} peut être obtenue par

$$P_{tr} = 1 - (1 - \tau)^N \quad (10.14)$$

• **Probabilité de transmission avec succès**

Une AC_i peut être transmise avec succès seulement quand aucune autre AC de priorité plus élevé dans la même station et aucune autre station des $(N - 1)$ restantes ne transmet. Par conséquent, la probabilité conditionnelle de transmission avec succès P_{si} est définie par :

$$P_{si} = \frac{N \cdot \tau_i \cdot (1 - \tau)^{(N - 1)} \cdot \prod_{i' > i} (1 - \tau_{i'})}{P_{tr}} \quad (10.15)$$

10.5 Résultats et discussion

Dans cette section, nous présentons les résultats numériques de notre modèle. Les résultats analytiques suivants ont été obtenus à partir des valeurs des paramètres indiqués dans la table2.

TABLE 10.2: Paramètres de la couche PHY et de la couche MAC de IEEE 802.11^e

Paramètres	Valeur numérique
Délai de propagation	64 μs
Taille de la charge	1000bits
L'entête physique	192 bits
L'entête Mac	272 bits
Trame RTS	L'entête physique + 160 bits
Trame CTS	L'entête physique + 112 bits
Trame ACK	L'entête physique + 112 bits
CTS – timeout	DIFS+CTS
ACK – timeout	DIFS+ACK
Taux de données	11 Mbps
Taux de base	2 Mbps
Slot de temps	20 μs
SIFS	1 slot de temps
AIFS[AC3]	7 slot de temps
AIFS[AC2]	3 slot de temps
AIFS[AC1]	2 slot de temps
AIFS[AC0]	2 slot de temps
CW[AC3]	{7,15,31,63}
CW[AC2]	{15,31,63,127}
CW[AC1]	{31,63,127,255}
CW[AC0]	{31,63,127,255}

La figure (2) représente la variation du débit normalisé en fonction du nombre de stations. Dans la figure (a), on remarque une dégradation des débits normalisés par rapport à l'augmentation du nombre de stations. La décroissance du débit normalisé qui dépend des catégories d'accès est expliquée par le fait que plus il y a des stations dans le réseau, plus la collision est élevée. La différenciation avec le TXOPlimit, donne plus de priorité pour AC3 qui atteint son maximum presque à 0.45 bit/s, car elle est prioritaire pour transmettre plusieurs trames en rafales plutôt que les autres (AC1, AC2, AC0), qui sont autorisées à transmettre un petit nombre de trames.

Tandis que la figure (b) montre la diminution des débits normalisés par rapport à l'augmentation du nombre de stations des catégories AC0, AC1, AC2 et une dégradation légère pour AC3 qui ne dépasse pas 0.4 bit/s sans utiliser le TXOPlimit.

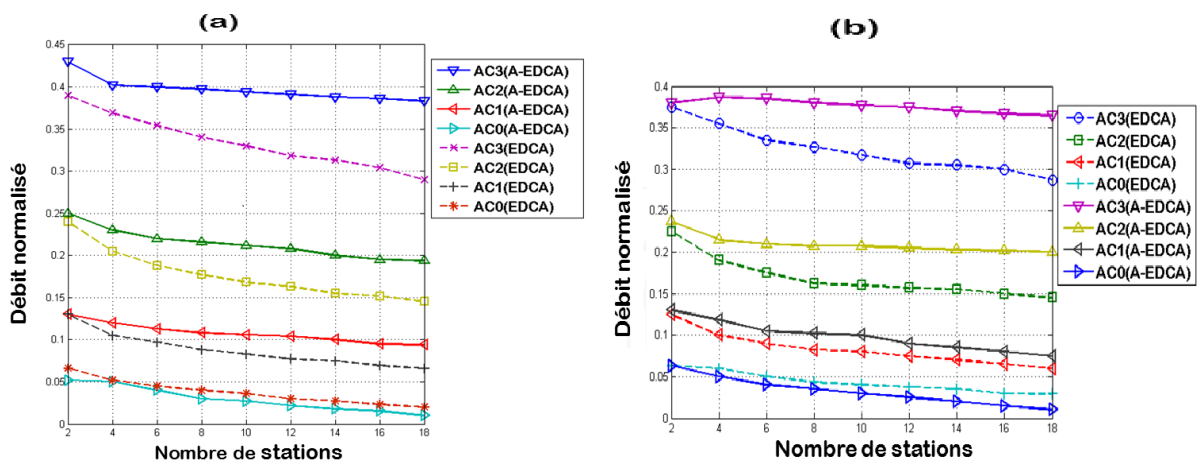


FIGURE 10.2: Débit normalisé en fonction du nombre de station.

La figure (3) représente la variation de débit total par rapport à l'augmentation du nombre de stations. On remarque toujours une diminution du débit par rapport à l'augmentation du nombre de station, ceci est expliquée par le fait que plus il y a des stations dans le réseau, plus la collision est élevée et il peut atteindre une valeur maximale égale presque à 0.9 bit/s en utilisant le TXOPlimit comme c'est illustré dans la figure (c) et qui ne dépasse pas 0.8 bit/s sans utilisé le TXOPlimit comme l'illustre la figure (d). La figure (4) représente la variation du

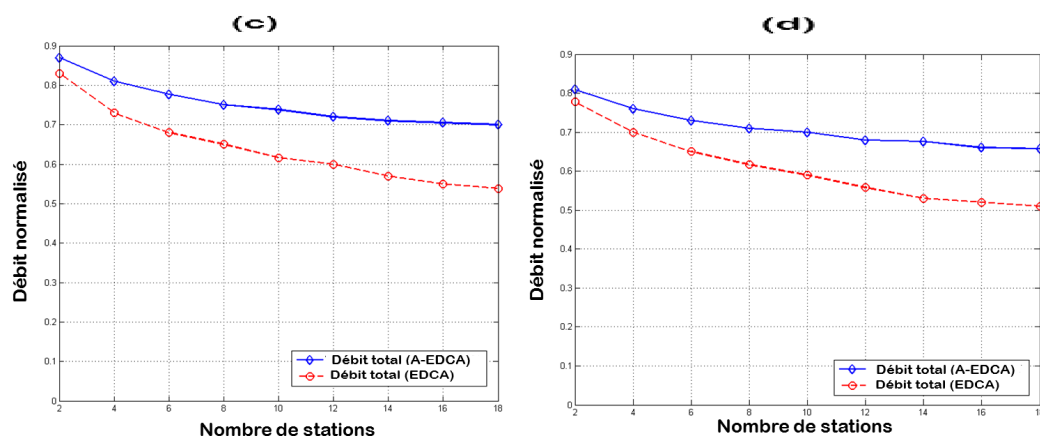


FIGURE 10.3: Débit total en fonction du nombre de station.

débit normalisé de chaque catégorie d'accès en fonction du taux d'erreur par bit (BER). Cette figure montre que quand le taux d'erreur par bit augmente, le débit normalisé disponible pour chaque AC diminue légèrement jusqu'à qu'il atteint la valeur égale à 10^{-5} , car le taux d'erreur est négligeable et lorsque le taux d'erreur par bit (BER) est supérieure à 10^{-5} , le débit normalisé de toutes les catégories d'accès divergent vers la même valeur (qui est égale à 0). Ceci peut être expliqué par la probabilité d'erreur qui est très grande.

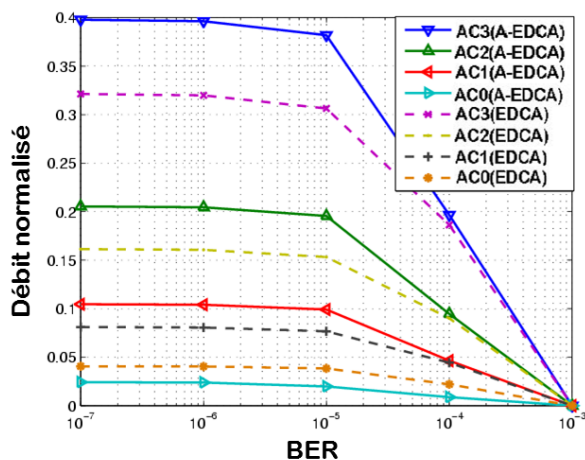


FIGURE 10.4: Débit normalisé en fonction du taux d'erreur par bit (BER).

10.6 Conclusion

EDCA est le mécanisme fondamental d'accès au support dans le protocole IEEE 802.11^e. Dans cet article, Pour améliorer les performances de la norme 802.11^e, nous avons proposé un nouveau traitement pour les erreurs de transmissions appelé A-EDCA. Nous avons présenté un modèle analytique pour évaluer la performance de A-EDCA sous un canal non idéal et dans des conditions non saturées. Les résultats obtenus montrent que A-EDCA est meilleur performant que EDCA de base. Nous concluons que la méthode modifiée proposée donne une meilleure performance de débit que le IEEE 802.11^e EDCA existant actuellement.

Références

1. G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function, IEEE Journal on selected Areas Communications, vol. 18, pp. 535-547, (March 2000).
2. Z. kong, D.H.K. Tsang, B. Bensaou, and D. Guo, Performance analysis of IEEE 802.11^e contention-based channel access, IEEE Journal selected Areas Communications, vol. 22, pp. 2095-2106, (December 2004).
3. D. Senthilkumar, A. Krishnan, Enhancement to IEEE 802.11 Distributed Coordination Function to Reduce Packet Retransmissions Under imperfect channel condition, Wireless Pers Commun, vol. 65, pp. 929-953, (2012).
4. Y.C. Yao and al., The Performance Evaluation of IEEE 802.11^e for QoS Support in Wireless LANs, Wireless Personal Communications, vol. 69, pp. 413-424, (April 2013).
5. P.E Engelstad, O.N. Osterbo, Analysis of the total delay of the IEEE 802.11^e EDCA and 802.11 DCF, in :Proceedings of IEEE International Conference on Communication (ICC'06), vol. 2, pp. 552-559, (2006).
6. K. Kosek-Szott, M. Natkanies, A. R. Pach, A simple but accurate throughput model for IEEE 802.11 EDCA in saturation and non-saturation conditions, Compture Networks, vol.55, pp. 622-635, (2011)
7. I. Tinnirello and S. Choi, Efficiency analysis of burst transmissions with block ACK in contention-based 802.11^e WLANs, in Proceedings of IEEE ICC05, Seoul, korea, vol. 5, pp. 3455-3460, (May 2005).
8. T. Li, Q. Ni, T. Turletti, and Y. Xiao, Performance Analysis of the IEEE 802.11^e Block ACK Scheme in a Noisy Channel, in Proceedings IEEE BroadNets05, (October 2005).
9. F. Peng, H.M. Alnuweiri, and V.C.M. Leung, Analysis of Burst Transmission in IEEE 802.11^e Wireless LANs, in Proceedings of IEEE ICC06, Istambul, vol. 2, pp. 535-539, (June 2006).

10. G. Min, J. Hu, M. E. Woodward, Modiling and analysis of TXOP differentiation in infrastructure-based WLANs, *Computer Networks*, vol.55, pp. 2545-2557, (2011).
11. J. Hu, G. Min, M. E. Woodward, Performance analysis of the TXOP burst transmission scheme in single-hop ad hoc networks with unbalanced stations, *Computer Communications*, vol. 34, pp. 1593-1603, (2011).
12. Y. Xiao, Enhanced DCF of IEEE 802.11^e to support QoS, in : *Proc IEEE WCNC03*, New Orleans, Louisiana, vol. 2, pp. 1291-1296, (March 2003).
13. R. Moraes, P. Portugal, F. Vasques, Simulation analysis of the IEEE 802.11^e EDCA protocol for an industrially-relevant real-time communication scenario, *IEEE* (2006).
14. A. Andreadis, G. Benelli, R. Zambon, Evaluation of QoS Support for Multimedia Traffics in IEEE 802.11^e, University of Siena, Department of Information Engineering, Siena, Italy.
15. H. Minooei, H. Nojumi, Performance evaluation of a new backoff method for IEEE 802.11, *Computer Communications*, vol.30, 3698-3704, (2007).
16. S. Choudhury, J.D. Gibson, Throughput optimization for wireless LAN in the presence of packet error rate constraints. *IEEE Communications Letters*, vol.12, pp.11-12, (2008).
17. F. Daneshgaran, M. Laddomada, F. Mesiti, M. Mondin, Unsaturated throughput analysis of IEEE 802.11 in presence of non ideal transmission channel and capture effects, *IEEE Transactions on Wireless Communications*, vol.7, pp.1276-1286, (2008).
18. D. Malone, K. Duffy, D. Leith. Modeling the 802.11 distributed coordination function in nonsaturated heterogeneous conditions. *IEEE/ACM Transactions on Networking*, vol. 15, pp. 159-172, (2007).
19. H. Lee, I. Tinnirello, J. Yu, S. Choi, A performance analysis of block ACK scheme for IEEE 802.11^e networks, *Computer Networks*, vol. 54, pp. 2468-2481, (2010).
20. S. Pan, J. Wu, Throughput analysis of IEEE 802.11^e EDCA under heterogeneous traffic, *Computer Communications*, vol. 32, pp. 935-942, (2009).
21. M. Varposhti, N. Movahhedinia, Supporting QoS in IEEE 802.11^e wireless LANs over fading channel, *Computer Communications*, vol. 32, pp. 985-991, (2009).
22. N. Chendeb, Y. Ghamri-Doudane, B. El-Hassan, Effect of transmission opportunity limit on transmission time modeling in 802.11^e, in : *IPOM*, San Jos, California, (2007).
23. N. Chendeb Taher, Y. Ghamri Doudane, B. El-Hassan, N. Agoulmine, An accurate analytical model for 802.11^e EDCA under different traffic conditions with contention free bursting, *J. Comput. Syst. Networks Commun*, (2011).
24. Wireless medium access control (MAC) and physical layer (PHY) specifications : Medium access control (MAC) enhancements for quality of service (QoS), *IEEE Std.802.11^e/Draft 5.0*, (July 2003).
25. E. Ziouva, T. Antonakopoulos, CSMA/CA performance under high traffic conditions : Throughput and delay analysis, *Compter Communications*, vol. 25, pp. 313-321, (2002).
26. S.W. Pan, J.S. Wu, Throughput analysis of IEEE 802.11^e EDCA under heterogeneous traffic, *Computer Communications*, vol. 32, pp. 935-942, (2009).
27. IEEE 802.11 WG, IEEE 802.1d, Part 3 : Mac bridges, *ANSI/IEEE Std. 802.1D*, 1998 edition, (1998).
28. Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, *ISO/IEC IEEE 802.11 Standard*, (1999).
29. L. BACHIRI, D. AISSANI et L. BOUALLOUCHE-MEDJKOUNE, Delay analysis of IEEE 802.11^e EDCA network under fading channel, *The First International Symposium on Informatics and its Applications ISIA'2014*, Université de M'Sila, (February 2014).

Troisième partie

Processus de publication

11

Publishing process : Overview and writing

Barèche A.

aicha_bareche@yahoo.fr

Résumé The main objective of this work is to give an overview on the publishing process and to explain its different stages.

Keywords : Peer review, Manuscript, Scientific writing, Scientific journal, Submission.

11.1 Introduction

Publishing process has been known since the 17th century. One of the earliest research journals is the **Philosophical Transactions of the Royal Society**. This process has known a great evolution over years and centuries :

- First, scientific writing follows a rigid structure - a format developed over hundreds of years, which is considered to be the best way for communicating scientific findings to the broader research community.
- Peer review is a central concept for most academic publishing ; other scholars in a field must find a work sufficiently high in quality for it to merit publication. The origins of routine peer review for submissions dates to 1752 when the **Royal Society of London** took over official responsibility for the **Philosophical Transactions**. However there were some earlier examples.
- Not long ago, journals required three to five hard copies of a submission, perhaps along with a disk copy. Now, more and more journals are accepting submissions on-line rather than through the mail.

11.2 Submission process

The submission process can be summarized in the following figure. For more details, see [1, 3, 4, 6].

11.3 Writing a manuscript

11.3.1 Structure

Scientific writing follows a rigid structure - a format developed over hundreds of years, which is considered to be the best way for communicating scientific findings to the broader research community (see [2, 5]). Most disciplines use the format outlined in the grid below.

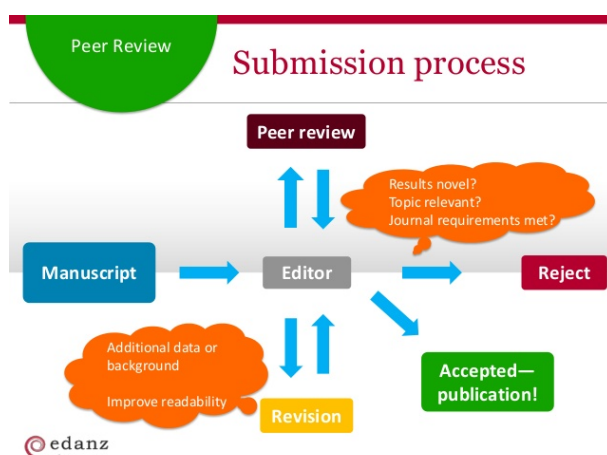


FIGURE 11.1: Submission process

Though the headings are standard for most journals, there is some variation, so it is essential to read the Guide for Authors of the journal you intend to submit your manuscript to before you start writing.

This format has the advantage that it enables the article to be read on several levels. Some people will just look at the title, others will read only the title and abstract, and those who want a deeper understanding of the research will read most, if not all, of the article.

TABLE 11.1: Structure of a manuscript

Section	Purpose
Title	Reflects content, entices reader
Author	Ensures recognition of the researcher(s)
Abstract	Summarizes the research and the conclusions
Keywords	Ensures the article is correctly identified in abstracting and indexing services
Body text	
Introduction	Puts the work into context
Methods	Explains how the data were collected
Results	Describes what was discovered
Discussion & Conclusions	Explores the implications of the findings
Acknowledgements	Ensures those who helped with the research are recognized
References	Ensures previously published work is recognized
Supplementary material	Provides online additions to the article, such as raw data, video and audio

11.3.2 Presentation

How well a manuscript is written also depends on style, language, your mastery of English, formatting, the illustrations and graphs you choose... in other words, the presentation of your material (see [2, 5]).

Style and language

When writing your manuscript, bear in mind that your chosen journal probably has a specific style. If you can write in that style, your chances of getting accepted will increase. There is also

an underlying style for writing a scientific manuscript. The objective is to report your findings and conclusions clearly, and as concisely as possible. Some pointers :

- Try to avoid embellishment with unnecessary words or phrases - keep it simple.
- Make sure transitions are succinct, with one section naturally flowing into another. Although advance organizers can be helpful, unneeded repetition should be avoided.
- Use the active voice wherever possible.
- Tenses are important : For known facts and hypotheses, use the present tense. When you refer to experiments you have conducted, use the past tense.
- Take care about punctuation, (page, formula, table and figure)-numbering, ...

Références

1. Burnham JC. (1990) The evolution of editorial peer review. *JAMA*. 263 : 1323-1329.
2. Davis, M. (2005) *Scientific Papers and Presentations*, 2nd Edition, Academic Press.
3. Gannon F. (2001) The essential role of peer review. *EMBO Reports*. 2 : 743.
4. Kronick DA. (1990) Peer review in 18th century scientific journalism. *JAMA*. 263 : 1321-1322.
5. Lagendijk, A. (2008) *Survival Guide for Scientists ; Writing-Presentation*, Amsterdam University Press.
6. Pressley, M. (in press). Overcoming rejection. In S. B. Wepner & L. Gambrell (Eds.), *Writing for literacy publications : Top ten guidelines*. Newark, DE : International Reading Association.

Index des Auteurs

Index

Aïssani D., 9, 28, 52, 65, 111, 128, 136, 151

Abdi N., 151

Abdoune K., 85

Adel-Aïssanou K., 9, 52, 65, 85

Amad M., 136

Arkoub D., 136

Aroul K., 151

Baalla M., 52

Bachiri L., 151

Barèche A., 165

Bouibed K., 105

Boukredera D., 85

Bouriahi L., 85

Ghiles L., 28

Hadji M., 128

Kadi S., 111

Krouri Y., 136

Laib F., 128

Lekadir O., 9

Medjoudj S., 65

Mehidi D., 65

Oudina S., 52

Ramdani A., 28

Saadi N., 28

Tazdaït T., 111

Touati N., 128

Zaidi H., 111

LaMOS Edition, 2020.

Les mathématiques, la santé (Covid 19,...) et autres domaines d'application

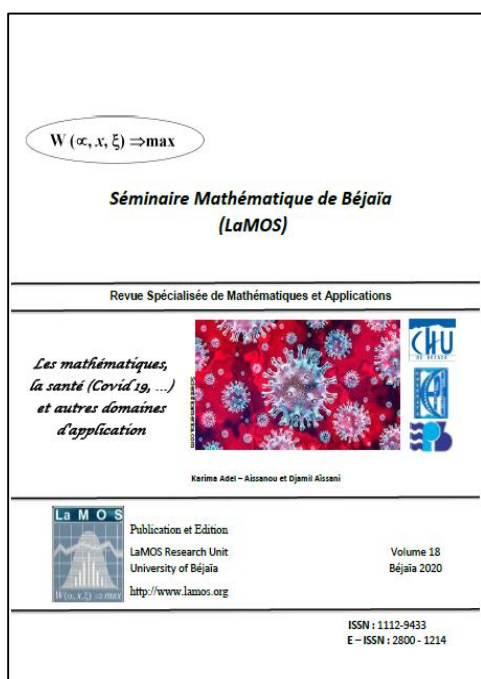
Résumé :

Ce numéro comprend trois parties: La première partie est intitulée: « *Les mathématiques et la santé (Covid-19, ...)* ». Elle concerne cinq applications des mathématiques, parfois combinée à l'informatique, dans un domaine particulier qui est la santé. De fait, le premier article concerne une synthèse complète de la littérature sur la "Modélisation mathématique de l'évolution de la pandémie COVID-19 ». Il est suivi par un autre article qui traite de « l'impact de la crise sanitaire sur le trafic portuaire au niveau de l'EPB (Entreprise Portuaire de Béjaïa)". Les trois autres articles de cette première partie abordent la "Simulation du mouvement des malades au niveau des urgences à l'hôpital Khelil Amrane 'Béjaïa)", l'"Application des méthodes d'apprentissage dans la prédiction du diabète de Type 2" et la "Conception d'un système multi-agents d'Aide à la décision pour une Gestion Efficace des Services d'Urgences: cas de l'hôpital de Kherrata".

La deuxième partie illustre cinq travaux sur l'application des mathématiques à la théorie des jeux et à l'évaluation des performances aux réseaux de communication. Quant à la troisième partie, elle développe un article sur le processus de soumission: "Overview and Writing".

Mots Clés: Modèle mathématique, Santé, COVID-19, Gestion des urgences, Diabètes de Type2, Jeux Bayésiens, Equilibre de Berge, réseaux de communication.

Séminaire Mathématique de Béjaïa (LaMOS). ISSN : 1112 – 9433. E – ISSN :2800 – 1214



La revue « *Séminaire Mathématique de Béjaïa (LaMOS)* » est une revue spécialisée de Mathématiques et Applications. Fondée en 2002, elle paraît une fois par an. Elle édite les Comptes Rendus Annuels du Séminaire Hebdomadaire de l'Unité de Recherche LaMOS (fondée en 1985). Elle peut également éditer des Actes de congrès, ou bien des monographies. Les textes sont expertisés par le comité éditorial. Parmi les auteurs de ce volume, citons les professeurs Tazdait T. (CIRED-Paris), Laïb F. (Cevital-Alger), Amad M. (Université de Bouira) et Boukredera D. (Université de Béjaïa).

Pour citer ce numéro : Adel-Aïssanou K. et Aïssani D., « Les mathématiques, la santé (Covid 19,...) et autres domaines d'application », Revue « *Séminaire Mathématique de Béjaïa (LaMOS)* », Volume 18, Juin 2020, 165 pages. ISSN:1112 – 9433.E – ISSN : 2800 – 1214.

ISSN : 1112 – 9433
E – ISSN : 2800 -1214
Dépôt légal : 934 - 2003