



HAL
open science

Modélisation et prévision des flux quotidiens des patients aux urgences hospitalières en utilisant l'analyse de séries chronologiques

Farid Kadri, Sondès Chaabane, Fouzi Harrou, Christian Tahon

► **To cite this version:**

Farid Kadri, Sondès Chaabane, Fouzi Harrou, Christian Tahon. Modélisation et prévision des flux quotidiens des patients aux urgences hospitalières en utilisant l'analyse de séries chronologiques. 7ème conférence de Gestion et Ingénierie des Systèmes Hospitaliers (GISEH), Jul 2014, Liège, Belgique. pp.1-8. hal-01023299v1

HAL Id: hal-01023299

<https://hal.science/hal-01023299v1>

Submitted on 11 Jul 2014 (v1), last revised 21 Jul 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modélisation et prévision des flux quotidiens des patients aux urgences hospitalières en utilisant l'analyse de séries chronologiques

KADRI Farid^{1,2}, CHAABANE Sondès^{1,2}, HARROU Fouzi³, TAHON Christian^{1,2}

¹Univ. Lille Nord de France, F-59000 Lille, France

²UVHC, TEMPO lab., "Production, Service, Information" Team, F-59313, Valenciennes

³Texas A&M University at Qatar, Chemical Engineering Program, Doha, Qatar

Résumé : La gestion efficace des flux de patients aux services d'urgence est devenue une question urgente pour la plupart des administrations hospitalières. Pour gérer cet afflux de patients, les services d'urgences nécessitent des ressources humaines et matérielles importantes, cependant, limitées. Dans ces conditions, le personnel médical se voit confronté très fréquemment à des situations de tension. Pour faire face à ces situations, les services d'urgences (SU) n'ont pas d'autre choix que de s'adapter. Dans ce cadre, la capacité de prévoir avec exactitude les exigences des services d'urgence a des implications considérables pour les hôpitaux pour améliorer l'allocation des ressources et la planification stratégique. L'objectif de cette étude est de développer un modèle de prévision du nombre de présences quotidiennes au service des urgences de l'hôpital de Lille, France. Le modèle ARIMA (Autoregressive Intergated Moving Average) a été appliqué aux arrivées quotidiennes des patients aux services d'urgences pédiatriques du CHRU de Lille. L'étude démontre comment l'analyse des séries chronologiques peut être un outil utile et facilement disponible pour la prévision, au moins à court terme, de la demande pour les services d'urgences hospitalières.

Mots clés : Services d'urgences, Situations de tension, séries chronologiques, Modèle ARIMA, Prévision.

Abstract: Effective management of patient flow in emergency departments has become a pressing issue for most hospital administrations. To handle the influx of patients, emergency services require significant human and material resources, however, limited. Under these conditions, the medical staff frequently confronted with situations of tension. To deal with these situations, emergency departments (ED) have no choice but to adapt. In this context, the ability to accurately predict the requirements of emergency services has significant implications for hospitals to improve resource allocation and strategic planning. The objective of this study is to develop a predictive model of the number of daily attendance at the emergency department of the hospital in Lille, France. The ARIMA (Autoregressive Moving Average Integrated) was applied to the daily arrivals of patients with pediatric emergency services of Lille University Hospital. The study shows how the time-series analysis can be a useful and easily available tool for predicting, at least in the short term, demand for hospital emergency departments.

Key-words: Emergency department, Strain situation, Time series, ARIMA, Forecasting.

1 Introduction

Au cours des dernières décennies, en raison de la demande croissante aux soins médicaux d'urgences, la gestion des services d'urgences de l'hôpital est devenue plus importante qu'avant. Aux Etats-Unis, entre 1993 et 2003, les visites aux services d'urgences ont été augmentées de 26% tandis que le nombre des services

d'urgences a été réduit de 9% (Kellermann, 2006). En France, à partir de la fin des années 1980, le nombre de passages dans les services d'urgences a augmenté très rapidement, leur croissance annuelle était de 43% en moyenne entre 1990 et 1998 (Baubeau et al., 2000), et de 4,5% entre 1999 et 2002. Aussi, le nombre de passages aux urgences a doublé entre 1990 et 2004, et il est passé de 7 millions à 14 millions (Cours des comptes, 2007).

Pour faire face à cet afflux de patients, les services d'urgences nécessitent d'importantes ressources humaines et matérielles (Kadri et al., 2014a). Malheureusement, ces ressources sont limitées. La prévision des admissions quotidiennes aux urgences hospitalières fournit des informations utiles qui peuvent être utilisées dans l'allocation et la planification des ressources (Kadri et al., 2013). La prévision permet aussi l'amélioration de l'efficacité des services d'urgences en réduisant le nombre de patients dans les files d'attente et en augmentant le nombre de passage (Kadri et al., 2014b).

L'objectif de ce travail est de proposer un modèle statistique basé sur des séries chronologiques. Ce modèle est utilisé pour prédire les admissions quotidiennes au service d'urgences pédiatriques du CHRU de Lille, France. La section 2 présente un état de l'art des admissions aux urgences hospitalières et la technique de modélisation par le processus ARIMA. La section 3 présente les données utilisées dans le cas d'étude. Dans la section 4, un modèle ARIMA a été utilisé pour prédire les admissions quotidiennes aux urgences pédiatriques de Lille. Enfin, la section 5 examine les principaux points abordés dans ce travail et conclut l'étude.

2 Prévision des admissions et modèles de prévision

2.1 Prévision des admissions aux urgences hospitalières

La revue de la littérature montre que l'analyse des séries chronologiques a été largement appliquée dans le domaine hospitalier afin de prévoir les arrivées de patients (jours, mois, années) et la durée du séjour dans les unités d'hospitalisation. Tandberg and Qualls, (1994) ont utilisé trois méthodes statistiques pour prédire le nombre d'admission à l'hôpital universitaire au Nouveau-Mexique. Rotstein et al., (1997) ont développé un modèle de régression linéaire pour la prévision des arrivées des patients aux urgences hospitalières de l'hôpital en Israël. Abdel-Aal and Mangoud, (1998) ont utilisé deux méthodes d'analyse de séries chronologiques (ARIMA et l'approche ad hoc) pour modéliser les admissions mensuelles à la clinique de Roi Fayçal, Al - Khobar, en Arabie Saoudite. (Jones et al., 2002) ont utilisé la méthode de Box-Jenkins (ARMA) pour prévoir le nombre d'admissions quotidiennes et le nombre de lits occupés aux urgences de l'hôpital de Bromley au Royaume-Uni. Earnest et al., (2005) ont utilisé des modèles ARMA pour prévoir l'occupation des couloirs de l'hôpital due à des infections SRAS. De Gooijer and Hyndman, (2006) ont passé en revue 25 années (période 1982-2005) de recherche sur la prévision par séries chronologiques. Les auteurs ont divisé les modèles de séries chronologiques en modèles linéaires et non linéaires tout en comparant la force et la faiblesse de chaque type de modèle. Champion et al., (2007) ont utilisé deux méthodes statistiques : i) un lissage exponentiel et ii) la méthode de Box- Jenkins pour prévoir le nombre d'admission chaque mois pour la période 2000 à 2005 aux urgences de l'hôpital régional de Victoria en Australie. Sun et al., (2009) ont utilisé les séries chronologiques pour prévoir les fréquentations quotidiennes des services d'urgences. Les auteurs ont conclu que l'analyse des séries chronologiques peut fournir un outil utile et facile pour prédire la charge de travail des services d'urgences qui peut être utilisé dans la planification des ressources.

2.2 Modèles de prévision

Les séries chronologiques sont des données sur la même unité d'observation à de multiples instants du temps. L'étude d'une série chronologique permet d'analyser, de décrire et d'expliquer un phénomène au cours du

temps et d'en tirer des conséquences pour la prise de décision. L'un des objectifs principaux de l'étude des séries chronologiques est la prédiction qui consiste à prévoir les valeurs futures de la série à partir de ses valeurs observées.

Les trois modèles de séries chronologiques univariées qui sont largement appliqués dans la littérature pour modéliser différents phénomènes sont i) le processus autorégressif (AR), ii) le processus de moyenne mobile (MA), et iii) le modèle ARMA (autorégressif et moyenne mobile) qui mixte les deux processus.

Les processus autorégressifs construits à partir de l'idée que l'observation à l'instant t s'explique linéairement par les observations précédentes. Le modèle AR est intuitivement intéressant car il décrit comment une observation dépend directement d'une ou plusieurs mesures précédentes, plus un bruit blanc. On dit que (Y_t) est un processus autorégressif d'ordre p (centré) s'il s'écrit :

$$Y_t = \sum_{i=1}^p a_i Y_{t-i} + \varepsilon_t \quad (1)$$

a_j est un paramètre non saisonnier, et ε_t est un bruit blanc $\varepsilon_t \sim N(0, \sigma^2)$.

Le modèle de moyenne mobile MA(q) est une fonction linéaire où les valeurs de la série (Y_t) s'expriment par la combinaison linéaire de l'erreur aléatoire qui a entaché les q essais précédents. Le modèle moyenne mobile (Moving Average) d'ordre q répond à l'équation :

$$Y_t = \varepsilon_t + \sum_{j=1}^q b_j \varepsilon_{t-j} \quad (2)$$

où b_j est un paramètre non saisonnier du modèle MA(q).

Le modèle mixte ou ARMA(p,q) définit des processus sous la forme d'une récurrence autorégressive avec un second membre de type moyenne mobile. Le modèle ARMA a d'abord été présenté par Box et Jenkins (Box and Jenkins, 1976). Un processus autorégressif moyenne mobile d'ordres p et q , ARMA(p,q), est de la forme :

$$Y_t = \sum_{i=1}^p a_i Y_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t \quad (3)$$

Le modèle ARMA est un outil d'analyse des séries temporelles très accessible aussi bien au niveau des contraintes méthodologiques qu'au niveau des modèles mathématiques utilisées qui sont des équations stochastiques linéaires peu complexes. Ce modèle est popularisé dans les travaux de Box and Jenkins, (1976) et utilisé essentiellement dans l'identification de la structure de la série temporelle et la prévision des valeurs futures d'un phénomène.

Si la série chronologique n'est pas stationnaire sur sa tendance, on a recours à un modèle de type ARIMA (p,d,q) où le d est le degré de la courbe de tendance ou encore le degré d'intégration de la série. Le modèle ARIMA a été présenté comme l'approche la plus efficace pour modéliser une grande variété de séries temporelles (Champion et al., 2007; Earnest et al., 2005; Jones et al., 2002; Makridakis et al., 1998). Quand la série présente une saisonnalité ou contient des fluctuations saisonnières, le modèle SARIMA peut être utilisé pour mesurer l'effet saisonnier.

3 Cas d'étude

Nous nous intéressons au service d'urgences pédiatriques (SUP) du centre hospitalier régional universitaire (CHRU) de Lille. Les données utilisées dans cette étude représentent les admissions journalières au service pour l'année 2012. Le SUP reçoit 24 heures sur 24, 7 jours sur 7, toute personne se présentant spontanément ou amenée par des ambulances ou par le SMUR (service mobile d'urgence et de réanimation), ou par les pompiers, en attente de soins dans les différents services médicaux et chirurgicaux.

La figure 1 présente le nombre d'arrivée par jour au SUP entre Janvier et Décembre 2012. Cette figure montre que la courbe des arrivées des patients n'est orientée ni vers le haut ni vers le bas sur la durée de l'ensemble de données. Ainsi, il apparaît que cette série chronologique n'a pas de tendance à long terme, ni de saisonnalité. La stationnarité de cette série (figure 1) est confirmée par le test de Phillips-Perron ($p.value = 0,01 < 5\%$) (Banerjee, 1993; Perron, 1993). Puisque la série est stationnaire alors l'ordre de terme de différenciation dans le modèle final est de 0 ($d = 0$). Cela nous conduit à proposer le modèle ARIMA ($p, 0, q$) ou ARMA (p, q).

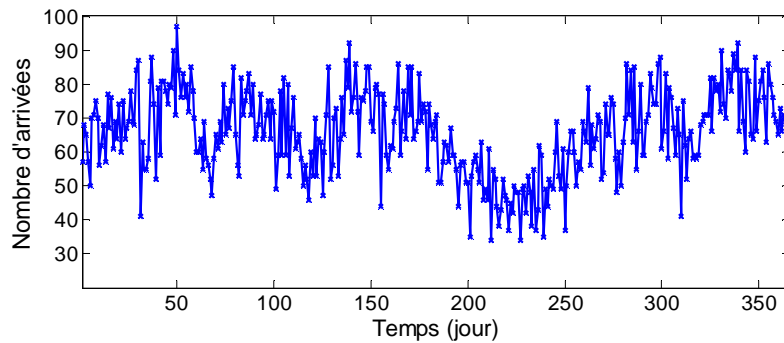


Figure 1. Nombre d'arrivée par jour au SUP de CHRU-Lille (Janvier-Décembre 2012)

La figure 2 représente la fonction d'auto-corrélation (ACF) de la série, on peut remarquer que l'ACF ne représente pas de périodicité. Aussi, l'ACF décroît exponentiellement jusqu'à zéro en augmentant les décalages. On remarque la présence d'une forte dépendance à court terme dans la série chronologique. Cette dernière, décrit la structure de corrélation d'ordre inférieur d'une série. C'est la propriété du modèle ARMA appelé aussi mémoire courte.

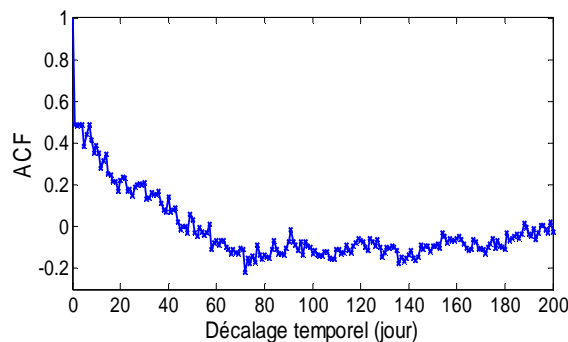


Figure 2. ACF de la série.

4 Modèle ARIMA proposé et prévision

La présente section est consacrée à la modélisation du nombre d'admission quotidienne au SUP de l'hôpital de Lille, France. L'élaboration d'un modèle statistique descriptif de l'évolution de flux de patients aux SU se fait à partir d'un échantillon de données dit échantillon d'apprentissage. Dans notre cas d'étude les données d'apprentissage représentent le nombre d'arrivée par jour au SUP entre Janvier et Novembre 2012. Ces données ont été centrées et réduites (de moyenne nulle et de variance unitaire), et ensuite utilisées pour développer un modèle descriptif.

Pour trouver le meilleur modèle ARIMA, qui expliquent le mieux les données, nous avons utilisé la méthodologie de Box and Jenkins, (1976) pour estimer les paramètres qui caractérisent la série de données. Les ordres (p,q) du modèle sont fixés par l'analyse de la fonction d'auto-corrélation simple (ACF) et partielle (PACF) de la série chronologique. Le modèle ARIMA $(p,0,q)$ le mieux adapté pour les admissions quotidiennes au SUP du CHRU de Lille est un ARIMA $(2,0,1)$, qui est un modèle autorégressif et moyenne mobile stationnaire et non saisonniers. L'estimation des paramètres de l'ARIMA $(2,0,1)$ peut être réalisée par plusieurs méthodes. Dans cette étude, les paramètres des modèles ARIMA ont été estimés en utilisant la méthode du maximum de vraisemblance pour l'ensemble de données (Box and Jenkins, 1976; Shumway and Stoffer, 2011). Les paramètres estimés pour le modèle ARIMA $(2,0,1)$ sont : $a_1 = 0.22$, $a_2 = 0.01$, et $b = 0.23$. En utilisant l'opérateur de retard B ($B^m Y_t = Y_{t-m}$), l'équation mathématique de modèle ARIMA $(2,0,1)$ pour les arrivées quotidiennes au SUP est donnée comme suit :

$$(1 - 0.22B - 0.01B^2)Y_t = \mu + (1 - 0.23B)\varepsilon_t \quad (4)$$

4.1 Validation du modèle sélectionné

La figure 3. (a) montre la série brute (données observées) avec la série ajustée par le modèle ARIMA $(2,0,1)$ entre Janvier et Novembre 2012. Les diagrammes de dispersion des données observées par rapport aux données prédites sont présentés dans la figure 3. (b). Cette figure montre que les points sont bien distribués le long de la ligne diagonale et il n'ya aucune indication d'une courbure ou d'autres anomalies.

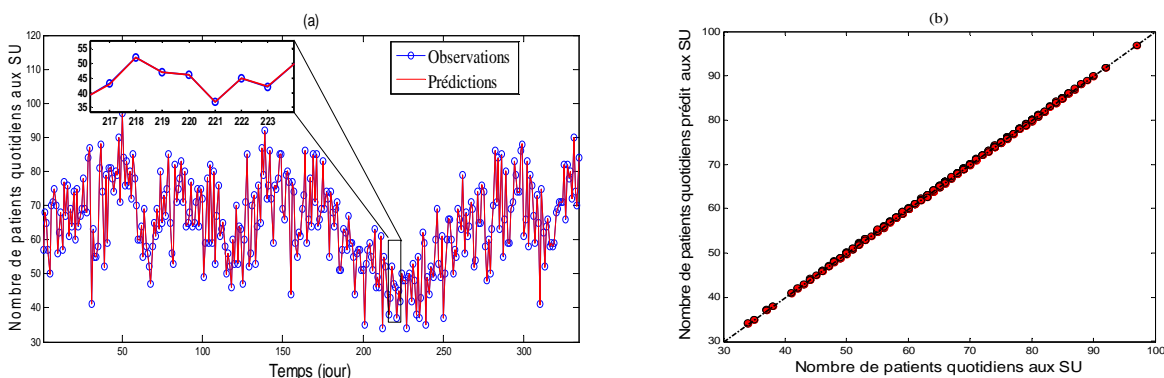


Figure 3 : Série brute et les données ajustées par le modèle ARIMA $(2,0,1)$, (Janvier-Novembre 2012).

Il existe plusieurs approches dans la littérature pour évaluer la capacité de prédiction d'un modèle (Makridakis et al., 1998; Mayer and Butler, 1993; Windhorst and Johansson, 1999). Dans cette étude, on a utilisé EQM (Erreur Quadratique moyenne) et le coefficient de détermination ou R^2 pour évaluer le modèle sélectionné. La qualité de l'ajustement de modèle ARIMA sélectionné a été évaluée en premier en termes de critère EQM qui représente l'écart moyen des valeurs prédites par rapport à celles observées, il est défini comme suit :

$$EQM = \sqrt{\frac{\sum(\hat{Y}-Y)^2}{n}} \quad (5)$$

où Y est la valeur mesurée, \hat{Y} représente les valeurs prédites correspondantes, et n est le nombre de données. En outre, nous évaluons la qualité de l'ajustement en utilisant le coefficient de détermination, ou R^2 , qui est défini comme suit :

$$R^2 = 1 - \frac{SSR}{SSY} \quad (6)$$

où SSR est la somme des carrés des résidus, il est également connu comme la somme des carrés des erreurs de prédiction. SSY est le modèle théorique de référence où pour chaque réponse expérimentale (données observées) une valeur constante est calculée comme la moyenne de la réponse expérimentale. Si la valeur de $R^2 = 1$, elle indique que le modèle s'adapte parfaitement aux données. Le degré d'ajustement, diminue au fur et à mesure que cette statistique se rapproche de zéro.

Pour le modèle sélectionné, la valeur de $R^2 = 0.99$ et la valeur de l'EQM = 0.141. La haute valeur de R^2 et la faible valeur de l'EQM du modèle sélectionné, signifie que le modèle choisi explique bien les données observées.

4.2. Analyse des résidus

L'analyse des résidus est l'étape la plus importante dans la validation du modèle. Elle consiste à évaluer comment les résidus sont distribués. Les résidus d'un bon modèle présentent diverses propriétés : normalité, linéarité, homoscedasticité, et indépendance. Tout d'abord il est nécessaire de vérifier si les résidus sont distribués selon une loi normale. Dans cette étude, la normalité des résidus a été vérifiée par l'analyse de la droite de Henry et de l'histogramme et illustrée par les figures 4. (a) et 4. (b), respectivement. Ces figures montrent que l'hypothèse d'une distribution normale pour les résidus est satisfaisante. L'indépendance des résidus (plus précisément, l'absence d'autocorrélation) a également été étudiée. Si les résidus respectent cette hypothèse, alors l'ACF empirique des résidus ne recèle pas de valeurs systématiques significativement différentes de 0 pour des retards différents de 0. La figure 5 qui présente l'ACF empirique des résidus, montre que les résidus sont approximativement non corrélés. Comme les résidus sont normalement distribués et non corrélés, on en déduit que le modèle correspond à la réalité.

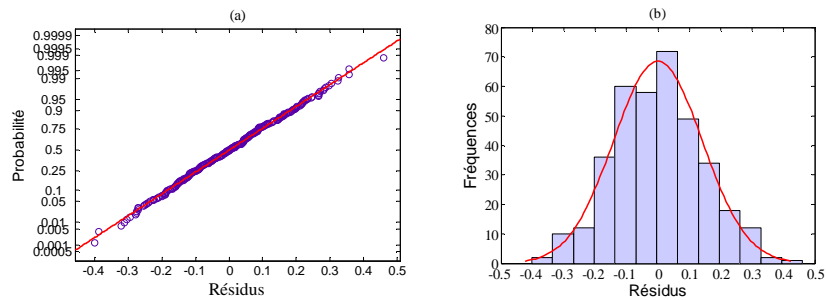


Figure 4 : Test de distribution gaussienne a) la ligne de Henry b) histogramme.

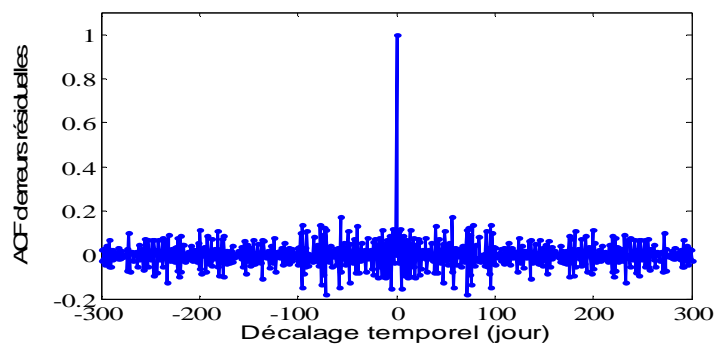


Figure 5 : ACF d'erreurs résiduelles.

4.3 Prédiction

Une fois que le modèle a été ajusté aux données, la prédiction de valeurs futures de la série chronologique peut être effectuée. Dans cette partie, le modèle ARIMA(2,0,1) est examiné pour sa capacité prédictive. Le modèle est comparé avec les valeurs de données réelles, qui n'ont pas été incluses dans la procédure d'ajustement du modèle. La série et les prévisions de nombre d'admissions au SUP à un horizon (H) de sept jours sont présentées dans la figure 6. La comparaison entre les données prédites et les données observées est présentée dans le tableau 1.

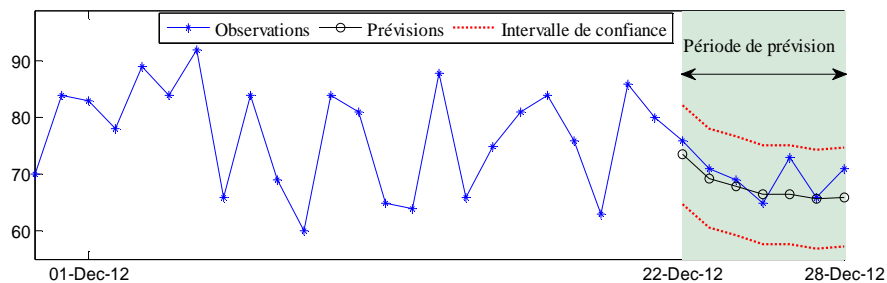


Figure 6: Prédiction à un horizon $h = 7$ jours

Tableau 1: La comparaison entre les données observées et les données prédites

Horizon H	Observation	Prevision	Erreur
1	76	73,52	2,47
2	71	69,28	1,71
3	69	67,93	1,06
4	65	66,41	1,41
5	73	66,41	6,58
6	66	65,64	0,35
7	71	65,98	5,01

En se basant sur le tableau 1, on peut noter que l'erreur de prévision moyenne pour les arrivées aux SUP est de 2,66 (3 patients), avec un taux d'erreur de 3,79%. Les résultats indiquent que le modèle de prévision proposé fournit une description acceptable des admissions quotidiennes aux urgences pédiatriques de l'hôpital de Lille.

5 Conclusion

Dans cet article, nous avons proposé un modèle de prévision des admissions quotidiennes aux urgences pédiatriques du CHRU de Lille. Ce modèle est basé sur l'analyse des séries chronologiques univariées. Les résultats indiquent que le modèle proposé offre une bonne description et une performance de prévision acceptable du nombre d'admissions aux urgences pédiatriques du CHRU de Lille.

Cette étude montre que les méthodes de prévision sont utiles dans l'analyse des systèmes de gestion d'hospitalière. La prévision précise de la demande quotidienne aux soins d'urgences hospitalières facilitera la planification en temps opportun des ressources humaines et matérielles au sein de ces établissements de santé. La suite de nos travaux consiste à améliorer le modèle de prévision en utilisant l'analyse multivariée qui prend en compte d'autres variables explicatives telles que des mesures météorologiques et les événements épidémiques.

Références

- Abdel-Aal, R.E., Mangoud, A.M., 1998. Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis. *Comput. Methods Programs Biomed.* 56, 235–247.
- Banerjee, A., 1993. *Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data.* Oxford University Press.
- Baubeau, D., Deville, A., Joubert, M., 2000. Les passages aux urgences de 1990 à 1998: une demande croissante de soins non programmés (No. 72).
- Box, G.E.P., Jenkins, G.M., 1976. *Time series analysis: forecasting and control.* Holden-Day.
- Champion, R., Kinsman, L.D., Lee, G.A., Masman, K.A., May, E.A., Mills, T.M., Taylor, M.D., Thomas, P.R., Williams, R.J., 2007. Forecasting emergency department presentations. *Aust. Health Rev. Publ. Aust. Hosp. Assoc.* 31, 83–90.
- Cours des comptes, 2007. Les urgences médicales, constats et évolution récente, rapport public annuel - 08 février 2007.
- Earnest, A., Chen, M.I., Ng, D., Sin, L.Y., 2005. Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore. *BMC Health Serv. Res.* 5, 36. doi:10.1186/1472-6963-5-36
- Gooijer, J.G.D., Hyndman, R.J., 2006. 25 years of time series forecasting. *Int. J. Forecast.* 22, 443–473.
- Jones, S.A., Joy, M.P., Pearson, J., 2002. Forecasting demand of emergency care. *Health Care Manag. Sci.* 5, 297–305.
- Kadri, F., Harrou, F., Chaabane, S., Tahon, C., 2014a. Time series modelling and forecasting of emergency department overcrowding. *Journal of Medical Systems*, (DOI: 10.1007/s10916-014-0107-0)
- Kadri, F., Chaabane, S., Tahon, C., 2014b. A simulation-based decision support system to prevent and predict strain situations in emergency department systems. *Simulation Modelling Practice and Theory* 42, 32–52.
- Kadri, F., Pach, C., Chaabane, S., Berger, T., Trentesaux, D., Tahon, C., Sallez, Y., 2013. Modelling and management of the strain situations in hospital systems using an ORCA approach, *IEEE IESM*, 28-30 October. RABAT - MOROCCO, 1-9.
- Kellermann, A.L., 2006. Crisis in the Emergency Department. *N. Engl. J. Med.* 355, 1300–1303.
- Makridakis, S.G., Wheelwright, S.C., Hyndman, R.J., 1998. *Forecasting: Methods and Applications*, 3rd Edition. ed.
- Mayer, D.G., Butler, D.G., 1993. Statistical validation. *Ecol. Model.* 68, 21–32. doi:10.1016/0304-3800(93)90105-2
- Perron, phillips, 1993. Racines unitaires en Macroéconomie: le cas d'une variable in Macroéconomie: développements récents, Malgrange P. et Salvat-Bronsard L. eds. *Economica* 327–358.
- Rotstein, Z., Wilf-Miron, R., Lavi, B., Shahar, A., Gabbay, U., Noy, S., 1997. The dynamics of patient visits to a public hospital ED: a statistical model. *Am. J. Emerg. Med.* 15, 596–599.
- Shumway, R.H., Stoffer, D., 2011. *Time series analysis and its applications With R Examples*, Springer Texts in Statistics. ed. New York.
- Sun, Y., Heng, B., Seow, Y., Seow, E., 2009. Forecasting daily attendances at an emergency department to aid resource planning. *BMC Emerg. Med.* 9, 1. doi:10.1186/1471-227X-9-1
- Tandberg, D., Qualls, C., 1994. Time series forecasts of emergency department patient volume, length of stay, and acuity. *Ann. Emerg. Med.* 23, 299–306.
- Windhorst, U., Johansson, H., 1999. *Modern techniques in neuroscience research*, 1 st. ed. Springer, New York.