Monitoring patient flow in a hospital emergency department: ARMA-based nonparametric GLRT scheme

Fouzi Harrou
Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology, Saudi Arabia

Farid Kadri
Aeroline and Customer Services, Agence, Sopra Steria Group, France

Ying Sun
Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology, Saudi Arabia

Sofiane Khadraoui
Department of Electrical Engineering, University of Sharjah, United Arab Emirates

Abstract
Overcrowding in emergency departments (EDs) is a primary concern for hospital administration. They aim to efficiently manage patient demands and reducing stress in the ED. Detection of abnormal ED demands (patient flows) in hospital systems aids ED managers to obtain appropriate decisions by optimally allocating the available resources following patient attendance. This paper presents a monitoring strategy that provides an early alert in an ED when an abnormally high patient influx occurs. Anomaly detection using this strategy involves the amalgamation of autoregressive-moving-average (ARMA) time series models with the generalized likelihood ratio (GLR) chart. A nonparametric procedure based on kernel density estimation is employed to determine the detection threshold of the ARMA-GLR chart. The developed ARMA-based GLR has been validated through practical data from the ED at Lille Hospital, France. Then, the ARMA-based GLR method’s performance was compared to that of other commonly used charts, including a Shewhart chart and an exponentially weighted moving average chart; it proved more accurate.

Corresponding author:
Fouzi Harrou, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.
Email: fouzi.harrou@kaust.edu.sa
Keywords
hospital emergency department, patient flow, overcrowding, time series models, hypothesis testing, strain situation

Introduction
Management efficiency in production systems is generally evaluated by its ability to achieve the desired performance using limited human, material, and financial resources. Therefore, management plans that maximize production systems for improved resource allocation and more strategic decision making are in high demand.1–7

The effective management of healthcare production systems can become a matter of life and death when decisions are made about allocation of resources in emergency departments (EDs) where patient flow is unpredictable.8 Efficiently managing EDs is essential to ensure better resource utilization, manage patient flows, and avoid overcrowding situations. In the US, between 1993 and 2003 the EDs demands increased by an average 26% while the number of EDs decreased about 9%.2,9 In France the EDs demands was doubled between 1990 and 2014,7 and this demand continues to grow.

To handle the unpredictable patient flows, the EDs need significant resources to prevent putting the medical and the paramedical staff under strain that greatly complicates their work, but these resources are limited.7 A previous study showed that patient influx can generate stressful situation that affect the performance of staff in ED.4,7,8 Early detection of abnormally high patient demands in ED could prevent overcrowding, which will reduce stress, maintain quality, and allow for the efficient allocation of resources. In this paper, we develop a statistical method to enhance detection of abnormally high patient arrivals at EDs with the aim of providing an early mechanism warning of the increased demand for care.

Generally, anomaly detection is one of the vital components for systems monitoring when anomalies are first identified and then isolated to ensure that they can be managed appropriately.10,11 Statistical process control (SPC) is considered as an efficient tool for monitoring production systems. SPC was successfully applied in various domains, including manufacturing,12 healthcare,13 and medical research. There are several monitoring charts in the literature including the Shewhart chart, the exponentially weighted moving average (EWMA) chart, and the generalized likelihood ratio (GLR) chart.14,15 The GLR-based monitoring approach has been shown attractive features in detecting abnormalities in time-series data,16 which make it advantageous for use in several applications.17–19

An essential feature is the GLR approach simplicity and the detection threshold is the only design parameter that requires to be defined.17 Essentially, the general assumption for univariate monitoring charts such as, GLR-based chart, is that the observations are uncorrelated or statistically independent over time. However, the real data often show serial dependency. Many production systems, such as hospital systems, yield data that have correlation between successive observations. This autocorrelation is expected to affect the performance of conventional SPC charts,20 violating the assumption of independence; monitoring charts developed under the assumption of independence are compromised by this violation.20 This paper focuses on the development of monitoring system that combined time-series models and a GLR monitoring chart to detect abnormal ED demands. However, it is worthwhile to notice that GLR test is typically designed based on the assumption that the monitored process and its derived variables follow a predefined probability distribution, generally Gaussian. Unfortunately, such assumption is mostly violated in practice. Indeed, the established detection thresholds become invalid, which alters systematically their detection performance. To overcome this issue, non-parametric chart can be designed to accurately set the detection thresholds when no prior knowledge about probability distribution are available. Here, a nonparametric procedure based on kernel density estimation is employed to determine the threshold of the ARMA-based GLR chart, which extends the flexibility of the developed method. To verify the proposed monitoring chart’s effectiveness, we compare the obtained results to those of state of the art approaches.
The following section briefly reviews the ARMA time-series modeling. In the third section, we introduce the GLR hypothesis testing and its application. The fourth section presents the developed ARMA-based GLR approach. The fifth section presents the experimental results, and the sixth section concludes with a discussion.

**Time-series modeling and ED demands**

In the few last decades, the increasing use of time series data has been motivated by the increasing of the data produced by production systems. Times series methods are widely applied in several application domains. Indeed, Time series approaches are commonly used to capture correlations in production processes. Stationary processes are considered as one of the most important class of time series models, which assume that the process remains in equilibrium around a constant mean. These models can provide a framework for seeking statistical control when monitoring auto-correlated processes. If we fit a good model that describes well the auto-correlated process, then we can approximate the uncorrelated process and monitored it via statistical hypothesis testing.

Time series methods are widely applied in various domains because of their ease of use, implementation, and interpretation in modeling and forecasting future variables. In hospital systems, in particular EDs, time-series models have been used to predict ED demands. For example, Carvalho-Silva et al. applied a seasonal ARIMA model to forecast ED visits at the Braga Hospital in Portugal. Rodríguez and Hernández presented a method based on time-series models to forecast hourly patient arrival to Tenerife ED, Spain, over 6 years (1997–2002). Kadri et al. proposed an approach based on a multivariate ARIMA model for modeling ED visits in Lille hospital France. Champion et al. used exponential smoothing and the Box-Jenkins approach for predicting the monthly demands for ED of a hospital in Victoria, Australia, from 2000 to 2005. Barişi analyzed the electromyographic (EMG) signals using an adaptive ARMA. Sun et al. predicted ED demands by fitting ARMA models. They concluded that the time-series models represent a good tool for predicting ED demands and workload of medical staff. Aboagye-Sarfo et al. developed multivariate vector-ARMA forecast models for predicting emergency department (ED) demand in Western Australia. They collected data of the number of monthly visiting patients for 7 years (2006–2012) from public hospital ED in Western Australia.

**Autoregressive and moving average models (p,q)**

The main steps to build a successful model are (i) the data analysis and pre-treatment, which usually involves features of studied system, (ii) the model selection, and (iii) the model estimation that determines the optimal parameter values of the selected model before validating the selected model.

ARMA model provide a sophisticate and analytical models. It combines elements of moving average and auto-regression to fit stationary time series data. One attractive feature of an ARMA model is that it includes fewer parameters than a moving average or autoregressive process standalone. The ARMA model of order (p; q) also denoted as ARMA(p; q) is written as:

\[(1 - \alpha_1 B - \ldots - \alpha_p B^p) y_t = (1 - \beta_1 B - \ldots - \beta_q B^q) \xi_t,\]

where \(y_t \in \mathbb{R}\), \(\xi_t\) is zero-mean Gaussian noise \(\xi_t \sim \mathcal{N}(0, \sigma^2)\), \(\alpha_i\) are the autoregressive (AR) coefficients, \(\beta_i\) are the moving average (MA) coefficients, and \(By_t = y_{t-1}\) is the backshift operator. The main steps required to determine the best model for Box-Jenkins methodology are schematically presented in Figure 1.
The most suitable models chosen are those that provide the best predictions. To this end, the data is split into training and validation data. The ARMA model is first trained using training data and evaluated via the validation data. The quality of the model is accomplished by analyzing the residuals. Several statistical indicators to assess the model quality including $R^2$, root mean square error (RMSE), mean absolute percentage error (MAPE), and graphical plots can be used to select the best model. In this study, $R^2$, MAPE, and RMSE are use to verify the accuracy of the selected model.

$$R^2 = 1 - \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{n} (y_t - mean(Y))^2},$$

(2)
\[
RMSE = \sqrt{\frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n}},
\]

\[
MAPE = 100\left[\frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{\sum_{t=1}^{n} |y_t|}\right] / n
\]

where \( y_t \) is the number of PED visits, \( \hat{y}_t \) is its corresponding predicted values, and \( n \) is the number of data points. \( R^2 \), also called the coefficient of determination, is a statistical metric that quantifies the ratio of the explained variance by the constructed model to the total variance in the dataset.\(^{37}\) It is always between 0 and 1; the higher the value, the better the model. For instance, \( R^2 = 0.99 \) means that the model captures 99% of the variance in the data, and only 1% is in the residuals. Of course, a more accurate model is characterized by higher \( R^2 \) and lower \( RMSE \) and \( MAPE \) values.

**GLR-based detector**

Within the hypotheses testing framework, the GLR test represents an effective parametric hypothesis test that can be applied for detecting anomalies and outliers in time series data.\(^{38}\) It gained a special attention in variety of applications, such as pollution monitoring,\(^{18}\) road traffic monitoring,\(^{39,40}\) imaging analysis,\(^{41,42}\) electronic systems,\(^{43}\) and train positioning via satellite.\(^{44}\) Essentially, the essence of the GLR test is to decide among a null hypothesis \( \mathcal{H}_0 \) and the alternative one \( \mathcal{H}_1 \) by using only the observed data. Toward this end, we suppose that an observation vector \( Y = [y_1, y_2, \ldots, y_n] \in \mathbb{R}^n \) is produced by one of the following Gaussian distributions:

\[
\begin{align*}
\mathcal{H}_0 &= \{ Y \sim \mathcal{N}(0, \sigma^2 I_n) \} \\
\mathcal{H}_1 &= \{ Y \sim \mathcal{N}(\theta \neq 0, \sigma^2 I_n) \}.
\end{align*}
\]

Here, \( \theta \) refers to the value of the fault, and \( \sigma^2 > 0 \) denotes the variance. Here, \( \mathcal{H}_0 \) represents the normal situation, and \( \mathcal{H}_1 \) describes the situation with potential abnormalities. It is worth to point out that the GLR detector discriminates \( \mathcal{H}_0 \) from \( \mathcal{H}_1 \) by comparing the charting statistics, \( \mathcal{L}(Y) \), to the detection limit, \( h(\alpha) \).

\[
\delta(Y) = \begin{cases} 
\mathcal{H}_0 & \text{if } \mathcal{L}(Y) = 2\log \frac{\sup_{\theta \in \mathbb{R}^n} f_{\theta}(Y)}{f_{\theta=0}(Y)} < h(\alpha) \\
\mathcal{H}_1 & \text{else}
\end{cases}
\]

Consequently, the GLR detection statistic, \( \mathcal{L}(Y) \), is defined as

\[
\mathcal{L}(Y) = 2\log \sup_{\theta} \left\{ \exp \left\{ -\frac{\|Y - \theta\|^2}{2\sigma^2} \right\} / \exp \left\{ -\frac{\|Y\|^2}{2\sigma^2} \right\} \right\},
\]

where \( f_{\theta}(Y) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \|Y - \theta\|^2 \right\} \) is the pdf of \( Y \). Then, (7) can be expressed as
The estimation of $\theta$ is expressed as $\hat{\theta} = \arg \min_{\theta} \| Y - \theta \|_2^2 = Y$. Concerning the likelihood ratio, $\mathcal{L}(Y)$, it is re-written as,

$$\mathcal{L}(Y) = \frac{1}{\sigma^2} \left\{ \| Y - \hat{\theta} \|_2^2 + \| Y \|_2^2 \right\}. \quad (8)$$

It is worth noticing that the detection threshold $h(\alpha)$, is defined such that the desired probability of false alarms is a priori fixed.

$$\mathbb{P}_0 \left( \mathcal{L}(Y) \geq h(\alpha) \right) = \alpha. \quad (9)$$

In sum, an anomaly is flagged using the GLR detector if the GLR charting statistic is above the detection threshold, $h(\alpha)$.

**Anomaly detection using an ARMA-based GLR test**

The aim of this study is to detect abnormally high patient demands in ED using based on ED visits time-series data. To this end, we first design an adequate ARMA model to describe the nominal patient flow in ED using data when the ED is operating at a normal state. The reference ARMA model will be used for monitoring the possible abnormal demand for ED services. In other words, the ARMA model is designed based on anomaly-free data and then adopted for monitoring new data. In short, it is used to generate residuals for anomaly detection (Figure 2). Indeed, residuals, which are the difference between the measured and estimated data obtained from the selected ARMA model, are adopted as an indicator to detect ED demand abnormality. The residuals offer pertinent information about the presence of abnormal conditions. Residuals reflect the operating conditions of the inspected ED, in the absence of a stressful situation, the residuals are close to zero, while residuals become larger in the case of abnormally high patient demands in ED. The GLR-based monitoring approach is applied residuals to check the concordance between the patient influx and the predicted influx form the ARMA model (Figure 2). An abnormal event is reported if the residuals surpass the GLR detection threshold. Of course, the general procedure of the developed strategy for monitoring patient flow in a hospital emergency department is depicted in Figure 2.

**KDE-based GLR detection threshold**

When the underlying process is with non-normal or unknown distribution, KDE can effectively estimate the probability distribution of such observations. To extend the flexibility of the proposed monitoring approach, the detection threshold of the GLR test is computed in nonparametric way using KDE. Assume that $e = [e_1, e_2, \ldots, e_n]$ are residuals generated from ARMA model. For any real values of $e$, KDE is applied to nonparametrically estimates the probability density function (PDF) of the residuals by the following formula:\textsuperscript{45,46}

$$\hat{f}(e) = \frac{1}{nw} \sum_{i=1}^{n} K(\frac{e - e_i}{w}) \quad (10)$$
where $e_i$ refers to residuals, $n$ is the number of samples. $K$ is the kernel function and $w$ denotes the kernel bandwidth parameter. Among a lot of kernel functions, the Gaussian kernel is usually used:

$$K(e) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{e^2}{2}\right).$$

(11)

$w$ represents the smoothing bandwidth factor that determines the probability estimation quality. Small values of $w$ can lead to quite a noisy estimation and under-smoothed PDFs. On the other
hand, large values over-smooth them. Various procedures have been developed in the literature to compute the adequate parameter $w$. Silverman’s rule of thumb calculates an optimal $w$ based on the assumption that the data is normally distributed. This rule is usually used if the data is unimodal and close to normal distribution. When a large amount of data is available, and the data is far away from normal or multimodal, the approach called Improved Sheather Jones can be used to make the approximation of data distribution. As shown in, for $n$ observations with the standard deviation $\sigma$, the optimal choice of $w$ (i.e. that minimizes the mean integrated squared error) can be obtained by the following expression:

$$w = \left( \frac{4\sigma^5}{3n} \right)^{1/3} \approx 1.06\sigma n^{-1/5}. \quad (12)$$

For fitting long-tailed distributions and outliers, a robust estimation of $\sigma$ could be computed as $\hat{\sigma} = \text{median}(|X_i - \hat{X}|) / 0.6745$, where $\hat{X}$ represents the median of the data. After estimating the distribution of the monitoring statistics based on KDE, the detection threshold is computed by the $100\cdot(1 - \alpha)$th quantile of $f(e)$. The detection threshold, $h(\alpha)$, can then be obtained from the PDF of the residuals, $e$, for a given probability false alarm, $\alpha$ as follows:

$$\int_{h(\alpha)}^{\infty} f(e)de = 1 - \alpha$$

$$\int_{-\infty}^{h(\alpha)} f(e)de = \alpha$$

The detection threshold, $h(\alpha)$, is computed as the $(1 - \alpha)$th quantile of the distribution $\hat{F}(e)$ of the residuals.

$$h(\alpha) = F(e)^{-1}(1 - \alpha). \quad (15)$$

where $F^{-1}$ denotes the inverse CDF (the quantile function) of the residuals based on anomaly-free data.

**Case study**

This study is accomplished using data for arrivals to the PED in CHRU-Lille. The CHRU-Lille hospital assists around four million inhabitants in Nord-Pas-de-Calais in France, and its PED admits on average 23,900 patients per year. Also, the PED shares with other hospital departments access to its resources (e.g. clinical laboratory, scanner, and X-rays). The daily attendance to the PED from January 2012 to December 2012 is used in this study. The second co-author has full access to these data following a collaboration with the PED in CHRU-Lille during a HOST research project.

**Data analysis**

In this study, a time series model has been adopted to construct a descriptive model for patient attendance at the PED dataset. Firstly, we perform a descriptive study of the data to summarize their main characteristic features. A key preliminary step in the time-series analysis is plotting the original observations over time to get simple descriptive measures of the principal characteristics of the series. This
enables to preliminarily evaluate data patterns and behavior over time and may reveal relevant features of the series, such as trends, outliers, discontinuity, and seasonality. Figure 3 indicates that daily visits are almost stable during the period of the data set. We can observe the absence of a long-term trend or seasonality in this time-series data (Figure 3). This can be confirmed with the Phillips-Perron test with a *p*-value of 0.01 less than 5%. Prior to model construction, the data is scaled to zero mean-centered with unit variance.

The catalog of descriptive statistics for the daily patient arrival at the PED presented in Table 1 shows the highest attendance on Sundays and Mondays (69 patients). But in general, the average number of patients arriving at the PED is well distributed on all days of the week.

Figure 4 shows the autocorrelation function (ACF) of the daily PED visits data given in Figure 3. Indeed, the plot of the sample ACF, also called a correlogram, often helps to interpret a set of autocorrelation coefficients and to construct models for time-series data. From Figure 4, we observe the presence of an important short-term dependence (short-memory) in data. To model short-term correlations in a time series, the ARMA model can be used.

### Modeling of PED patient attendances

An ARMA time series model structure is quite enough to model the patient visits at the PED. To estimate the parameters of the selected ARMA model, the method of maximum likelihood
estimates is used. The best-fitted model was ARMA(1,1). The maximum likelihood estimates and their standard errors for this best-fit model ARMA(1,1) are presented in Table 2.

Figure 5 shows the observed and predicted daily attendance obtained from the selected ARMA model. According to Figure 5, the observed data are well-adjusted by the selected model.

Furthermore, goodness-of-fit results for the daily attendances at PED data studied are summarized in Table 3. Results in Table 3 show that the selected ARMA (1, 1) model with the selected parameters describes well the daily attendances data. The $r^2 = 0.99$ and the low RMSE of 0.11 and with low MAPE of 0.17% confirms the good predictive quality of the ARMA model.

Analysis of the residuals. The second step of modeling is aimed at analyzing and evaluating the residuals of the selected model. Figure 7 shows the plot of the residual of the selected ARMA model for daily attendance at the PED. Form this plot, it can be seen a random pattern and small variations around the zero mean, indicating support for the suitability of the selected ARMA model. Figure 8 depicts the scatter plots of residuals versus estimated values computed using the ARMA model. The Figure 8 displays a random pattern showing a good fit for the designed ARMA (1, 1) model.
Also, to show visually the normality of the residuals, histogram, and \(Q-Q\) plot of ARMA' residuals are presented in Figure 9(a) and (b). Figure 9(a) and (b) shows that the residuals are normally distributed. ACF of residual errors shown in Figure 10 clearly shows the absence of autocorrelation in residuals.

Figure 11 shows the \(p\)-values for the Ljung-Box statistics for lags till 20. The Ljung-Box statistics test if the residuals are uncorrelated or not.\(^{34}\) It can be seen that all \(p\)-values surpass the dashed blue line at 0.05, which means that the model has a reasonable quality of prediction. These results are consistent with the model assumption that the residuals are independent and identically distributed.

**Figure 5.** Observed and predicted daily attendance. The blue line and circles represent the actual observed patient visits to the PED included in the training sets, while the red line indicates model fit.

**Figure 6.** Scatter graph of estimated and recorded PED visits from Jan to Dec 2012.
Detection results

This section investigates and evaluates the performance of three residual-based monitoring charts for identifying abrupt or gradual increases in visits at the PED. The three charts are the GLR chart, a Shewhart individuals chart, and the EWMA chart applied to the residuals obtained from the ARMA model. The GLR chart was discussed in detail in Section 3. The details about the basic idea of the Shewhart and EWMA charts can be found.\textsuperscript{13,14,18,51,52} Two statistical metrics are used to judge

\begin{table}[H]
\centering
\caption{Evaluation of the ARMA model.}
\begin{tabular}{|l|c|}
\hline
Parameter & Values \\
\hline
$R^2$ & 0.99 \\
RMSE & 0.11 \\
MAPE & 0.17 \\
\hline
\end{tabular}
\end{table}

\textbf{Figure 7.} Residuals of the designed ARMA (1,1) model based on training data.

\textbf{Figure 8.} Graphing procedures used to evaluate equality of the variance: Predicted values versus residuals for selected ARMA model.
the detection efficiency of the three ARMA-based charts: the false detection rate (FAR) and the miss detection rate (MDR).

Here, we assess the ARMA-GLR monitoring chart’s detection capability using the time-series data of daily attendance recorded in the PED from January 2012 to December 2012. After building the ARMA model, two scenarios are investigated; one with an abrupt increase of patient flow (case A) and the second with a gradual increase in patient flow (case B). For comparison, we conducted the same tests using Shewhart and EWMA charts.

**Case A: Abrupt abnormal situation.** Herein, the abilities of the ARMA-GLR monitoring approach for detecting abrupt abnormal situation are demonstrated. Two examples ($A_1$ and $A_2$) are presented here to illustrate the capability of this control chart to identify an abrupt abnormal situation.
In the case study $A_1$, we simulate an abrupt increase of ED visits by adding a small bias to the test data, $Y$, between data points 100–108. The amplitude of the introduced bias is equal to 30% of the total variation in the testing data $Y$. This could describe a sudden augmentation of the number of arrivals. The results of Shewhart and EWMA monitoring charts are depicted in the top and middle of Figure 12, respectively. The top frame illustrates that the I-chart Shewhart statistic fails to detect a few anomalies, and the middle frame shows that the control limit of EWMA chart has been overpassed, indicating the occurrence of abnormal situation.

This is because the Shewhart individual chart takes a decision using the actual observation alone, while the EWMA charting incorporates all observations from past samples in addition to the present observation in the decision process. The parameters of the EWMA were chosen as the following: smoothing parameter, $\lambda = 0.25$, control limit width, $L = 3$. The bottom frame illustrates the results of the ARMA-GLR method to accurately detect this abnormal situation by starting a violation of the confidence limit without any false alarms. The GLR detection threshold was found to be $h(\alpha) = 0.46$ for a false alarm ($\alpha$) of 5%.

In the case study ($A_2$), we inject a moderate bias level (15%) to the testing data between data points 100–108. Results of the I-chart Shewhart scheme show that no anomaly was identified making this chart unsuccessful (Figure 13 top frame). The EWMA chart in the middle panel of Figure 13 detects a great portion of this anomaly with some false alarms. Essentially, this chart gets its advantage from the memory effect of the EWMA chart compared to the Shewhart individuals chart. However, the GLR best detected the given abnormal situation by showing a violation of confidence limits without any false alarms (Figure 13, bottom frame).

Table 4 confirms that the ARMA-GLR chart outperforms the ARMA-based Shewhart and EWMA schemes by achieving the best performance in terms of low FAR and MDR values.

**Case B: Gradual abnormal situation.** In this case study, a gradual abnormal patient arrivals is simulated by adding a ramp change with a slope of 0.1 to testing data between 300 and 365 sample intervals. Monitoring results of Shewhart and EWMA statistics are plotted in the top and middle panels of the Figure 14, respectively. They increase linearly from the 300th sample, exceeding the control limits at the 328th sample (Shewhart) and 320th sample (EWMA). Thus, although the EWMA method detects
the anomaly earlier than the Shewhart statistic, without regions of missed detection, it does produce some missed detections. However, the EWMA chart may detect slow drifts of mean faster than the Shewhart chart. The bottom panel of Figure 14 shows that the GLR statistic exceeds the violation

Figure 12. Monitoring results of the Shewhart-chart statistic (top), EWMA with $(L, \lambda) = (3, 0.2)$ (Plot (middle)), and GLR (bottom) decision statistics of an abnormal situation (Case $A_1$, example 1).
Figure 13. Monitoring results of the Shewhart I-chart statistic (top), EWMA with \((L, \lambda) = (3, 0.2)\) (Plot (middle)), and GLR (bottom) decision statistics of an abnormal situation (Case \(A_1\), example 2).

Table 4. FAR and MDR for each monitoring chart.

| Chart    | Case A – Example 1 |         | Case A – Example 2 |         |
|----------|-------------------|---------|-------------------|---------|
|          | FAR   | MDR    | FAR   | MDR    |
| Shewhart | 0     | 50     | 0     | 100    |
| EWMA     | 10    | 0.83   | 20    | 0.83   |
| GLR test | 0     | 0      | 0     | 0      |
threshold at the 307th sample, therefore, missing fewer detection that either of other charts, testifying to its superior performance. Here, the GLR approach begins to flag the abnormal increase of patient flow earlier than the other charts (EWMA and Shewhart). Detecting abnormal situations at an early stage is crucial to managing patient flow, avoiding strain situations, and optimizing ED resources.

**Figure 14.** Monitoring of the Shewhart I-chart statistic (top), EWMA with \((L, \lambda) = (3, 0.2)\) (Plot (middle)), and GLR (bottom) decision statistics in the presence of gradual abnormal situation (Case B).
The speed of detection provided by the GLR approach helps foster reactive control and gives ED managers time to decision making. The number of days after detection and before reaching strain situation. Note that as the magnitude of the strain situation is not theoretically defined, it is not easy to compute the exact time for early detection. However, the GLR approach can sense the small gradual increase after few observations from the introduction of the gradual anomaly.

In summary, these results show that the GLR chart, with a nonparametric detection threshold, outperforms the others by adequately detecting anomalies. We have shown that the GLR control chart is preferable for detecting a wide range of shift sizes, because control charts require specification of several chart design parameters by the user, making their design more complicated. Compared to other solutions, it presents the highest detection rate and reduces considerably the false alarms.

**Conclusion**

This study reports the development of an anomaly detection approach for monitoring abnormal daily demands at PED of Lille Hospital, France. This methodology combines an ARMA model with GLR test. The selected time series model was an ARMA (1,1) model, this latter was used as the model framework of the developed model. To monitor abnormal patient demands at the PED, we applied the GLR test to the residuals obtained from the selected ARMA model. Two case studies has been used to validate the proposed approach. Results demonstrated that the proposed methodology (ARMA-based GLR) shows a good performance compared to the other tested methods (Shewhart and EWMA monitoring charts).

The detection of strain situations in hospital emergency departments allows improving the care activity control and avoiding stressful conditions such that patient need and internal resources can be best managed. The early detection of these abnormal situations will also help the ED staff to prepare for absorbing the unusual rise in daily patient attendances at the ED without reducing the quality care of these latter. Therefore, the information provided by the presented anomaly-detection methodology can be used to optimize the performance of EDs.

To further improve patient flow management at hospital EDs, we plan to develop an approach to forecast daily ED visits by incorporating explanatory variables, such as meteorological measurements (e.g. temperature) and epidemic events. Toward this end, the ARMA models with external inputs could be used. Indeed, an accurate forecast of patient flow provides relevant information about the future trends of ED demands, facilitating the planning of nursing rosters and staff allocation within the ED and assisting in bed occupancy prediction. Further, it will be interesting to combine forecasting and anomaly detection and investigate whether improved forecasting can enhance anomaly detection. This is a topic for future research.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**Ethical approval**

The authors declare that this study does not involve human subjects, human materials, and human data. The authors confirm here that all datasets used in this work do not need any bio-ethical approval since they have already been anonymized by the PDE before making them available to the authors.
**References**

1. Zhang Y, Luo L, Zhang F, et al. Emergency patient flow forecasting in the radiology department. *Health Informatics J* 2020; 20(4): 2362–2374.
2. Kellermann AL. Crisis in the emergency department. *N Eng J Med* 2006; 355(13): 1300–1303.
3. Zeroual A, Harrou F, Dairi A, et al. Deep learning methods for forecasting COVID-19 time-series data: A comparative study. *Chaos Solitons Fractals* 2020; 140: 110121.
4. Harrou F, Kadri F, Chaabane S, et al. Improved principal component analysis for anomaly detection: Application to an emergency department. *Comput Indus Eng* 2015; 88: 63–77.
5. Kolker A. Process modeling of emergency department patient flow: Effect of patient length of stay on ED diversion. *J Med Syst* 2008; 32(5): 389–401.
6. Kadri F, Pach C, Chaabane S, et al. Modelling and management of strain situations in hospital systems using an orca approach. In: *Industrial engineering and systems management (IESM), proceedings of 2013 international conference on*, 2013, Agdal, Morocco, pp.1–9, IEEE.
7. Kadri F, Harrou F, Chaabane S, et al. Time series modelling and forecasting of emergency department overcrowding. *J Med Syst* 2014; 38(9): 1–20.
8. Kadri F, Chaabane S and Tahon C. A simulation-based decision support system to prevent and predict strain situations in emergency department systems. *Simul Model Pract Theor* 2014; 42: 32–52.
9. Warden G, Griffin R, Erickson S, et al. *Hospital-based emergency care: at the breaking point*. Washington, DC: National Academy of Science, 2006.
10. Isermann R. Model-based fault-detection and diagnosis: status and applications. *Ann Rev Control* 2005; 29: 71–85.
11. Yin S, Zhu X and Jing C. Fault detection based on a robust one class support vector machine. *Neurocomput* 2014; 145: 263–268.
12. Romagnoli J and Palazoglu A. *Introduction to process control*. CRC press, 2005. Boca Raton.
13. Kadri F, Harrou F, Chaabane S, et al. Seasonal arma-based spc charts for anomaly detection: Application to emergency department systems. *Neurocomput* 2016; 173: 2102–2114.
14. Montgomery DC. *Introduction to statistical quality control*. New York: John Wiley& Sons, 2005.
15. Bissell D. *Statistical methods for SPC and TQM*, volume 26. CRC Press, 1994. London.
16. Harrou F, Fillatre L and Nikiforov I. Anomaly detection/detectability for a linear model with a bounded nuisance parameter. *Ann Rev Control* 2014; 38(1): 32–44.
17. Basseville M and Nikiforov I. *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs, 1993. United States.
18. Harrou F, Fillatre L, Bobbia M, et al. Statistical detection of abnormal ozone measurements based on constrained generalized likelihood ratio test. In: *IEEE 52nd annual conference on decision and control (CDC)*, 2013, Firenze, Italy, pp.4997–5002, IEEE.
19. Harrou F, Nounou M and Nounou H. Enhanced monitoring using PCA-based GLR fault detection and multiscale filtering. In *IEEE symposium on computational intelligence in control and automation (CICA)*, 2013, Singapore, pp.1–8, IEEE.
20. Apley D and Shi J. The glrt for statistical process control of autocorrelated processes. *IIE Trans* 1999; 31(12): 1123–1134.
21. Brockwell P and Davis R. *Time series: theory and methods*. Springer Science & Business Media, 2013. New York.
22. Shumway R and Stoffer D. *Time series analysis and its applications*. Springer Science & Business Media, 2013. Berlin, Heidelberg.
23. Lin W. Modeling and forecasting hospital patient movements: univariate and multiple time series approaches. *Int J Forecast* 1989; 5(2): 195–208.
24. Jones S, Evans R, Allen T, et al. A multivariate time series approach to modeling and forecasting demand in the emergency department. *J Biomed Inform* 2009; 42(1): 123–139.
25. Afilal M, Yalaoui F, Dugardin F, et al. Forecasting the emergency department patients flow. *J Med Syst* 2016; 40(7): 1–18.
26. Abdel-Aal R and Mangoud A. Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis. *Comput Method Programs Biomed* 1998; 56(3): 235–247.
27. Carvalho-Silva M, Monteiro MTT, de Sá-Soares F, et al. Assessment of forecasting models for patients arrival at emergency department. *Oper Res for Health Care* 2018; 18: 112–118.
28. Martín R and Cáceres H. A method for ascertaining the seasonal pattern of hospital emergency department visits. *Revista Española de Salud Pública* 2005; 79(1): 5–15.
29. Kadri F, Harrou F and Sun Y. A multivariate time series approach to forecasting daily attendances at hospital emergency department. In: *2017 IEEE symposium series on computational intelligence (SSCI)*, 2017, pp.1–6, Honolulu, HI, USA, IEEE.
30. Champion R, Kinsman L, Lee G, et al. Forecasting emergency department presentations. *Aust Health Rev* 2007; 31(1): 83–90.
31. Barişçi N. The adaptive ARMA analysis of EMG signals. *J Med Syst* 2008; 32(1): 43–50.
32. Sun Y, Heng B, Seow Y, et al. Forecasting daily attendances at an emergency department to aid resource planning. *BMC Emerg Med* 2009; 9(1): 1.
33. Aboagye-Sarfo P, Mai Q, Sanfilippo FM, et al. A comparison of multivariate and univariate time series approaches to modelling and forecasting emergency department demand in western australia. *J Biomed Inform* 2015; 57: 62–73.
34. Box G and Jenkins G. *Time series analysis: forecasting and control*, volume 734. John Wiley & Sons, 2011. Hoboken, New Jersey.
35. Jones S, Thomas A, Evans R, et al. Forecasting daily patient volumes in the emergency department. *Acad Emerg Med* 2008; 15(2): 159–170.
36. Makridakis S, Wheelwright S and Hyndman R. *Forecasting methods and applications*. John Wiley & Sons, 2008. New York NY USA.
37. Piñeiro G, Perelman S, Guerschman JP, et al. How to evaluate models: observed vs. predicted or predicted vs. observed? *Ecol Model* 2008; 216(3-4): 316–322.
38. Harrou F, Sun Y, Hering AS, et al. *Statistical process monitoring using advanced data-driven and deep learning approaches: theory and practical applications*. Elsevier, 2020. San Diego.
39. Harrou F, Zeroual A and Sun Y. Traffic congestion detection based on hybrid observer and GLR test. In: *2018 Annual American control conference (ACC)*, 2018, Milwaukee, WI, USA, pp.604–609, IEEE.
40. Zeroual A, Harrou F and Sun Y. Road traffic density estimation and congestion detection with a hybrid observer-based strategy. *Sustain Cities and Soc* 2019; 46: 101411.
41. Newey M, Benitz G and Kogon S. A generalized likelihood ratio test for sar ccd. In: *Signals, systems and computers (ASILOMAR), 2012 conference record of the forty sixth asilomar conference on*, 2012, Pacific Grove, CA, USA, pp.1727–1730, IEEE.
42. Boisgontier H, Noblet V, Heitz F, et al. Generalized likelihood ratio tests for change detection in diffusion tensor images: Application to multiple sclerosis. *Med Image Anal* 2012; 16(1): 325–338.
43. Nunzi E, Carbone P and Tavella P. Fault detection in atomic clock frequency standards affected by mean and variance changes and by an additive periodic component: the glrt approach. In *Instrumentation and measurement technology conference proceedings, 2008. IMTC 2008*, 2008, Victoria, BC, Canada, pp.1594–1597, IEEE.
44. Harrou F, Fillatre L and Nikiforov I. Anomaly detection with bounded nuisance parameters and safe train navigation. In *2008 10th International conference on control, automation, robotics and vision*, 2008, Hanoi, Vietnam, pp.1021–1026, IEEE.
45. Martin E and Morris A. Non-parametric confidence bounds for process performance monitoring charts. *J Process Control* 1996; 6(6): 349–358.
46. Harrou F, Zeroual A and Sun Y. Traffic congestion monitoring using an improved kNN strategy. *Measurement* 2020; 156: 107534.
47. Sheather SJ and Jones MC. A reliable data-based bandwidth selection method for kernel density estimation. *J Royal Stat Soc: Series B (Methodological)* 1991; 53(3): 683–690.
48. Silverman BW. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986. London.
49. Mugdadi AR and Ahmad IA. A bandwidth selection for kernel density estimation of functions of random variables. *Comput Stat Data Anal* 2004; 47(1): 49–62.
50. Bowman AW and Azzalini A. *Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations*, volume 18. OUP Oxford, 1997. Oxford University Press.
51. Harrou F and Nounou M. Monitoring linear antenna arrays using an exponentially weighted moving average-based fault detection scheme. *Syst Sci Control Eng: An Open Access J* 2014; 2(1): 433–443.
52. Lucas J and Saccucci M. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 1990; 32(1): 1–12.
53. Harrou F, Sun Y and Madakyaru M. Kullback-leibler distance-based enhanced detection of incipient anomalies. *J Loss Prev Process Indus* 2016; 44: 73–87.