Seasonal ARMA-based SPC charts for anomaly detection: Application to emergency department systems

Farid Kadri [a,*], Fouzi Harrou [c], Sondès Chaabane [b], Ying Sun [c], Christian Tahon [b]

[a] PIMM Laboratory, UMR CNRS 800, Arts et Métiers ParisTech, Paris, France
[b] LAMIH Research Center UMR CNRS 8201, UVHC, Le Mont Houy, F-59313 Valenciennes, France
[c] CEMSE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

ABSTRACT

Monitoring complex production systems is primordial to ensure management, reliability and safety as well as maintaining the desired product quality. Early detection of emergent abnormal behaviour in monitored systems allows pre-emptive action to prevent more serious consequences, to improve system operations and to reduce manufacturing and/or service costs. This study reports the design of a new methodology for the detection of abnormal situations based on the integration of time-series analysis models and statistical process control (SPC) tools for the joint development of a monitoring system to help supervising of the behaviour of emergency department services (EDs). The monitoring system developed is able to provide early alerts in the event of abnormal situations. The seasonal autoregressive moving average (SARMA)-based exponentially weighted moving average (EWMA) anomaly detection scheme proposed was successfully applied to the practical data collected from the database of the paediatric emergency department (PED) at Lille regional hospital centre, France. The method developed utilizes SARMA as a modelling framework and EWMA for anomaly detection. The EWMA control chart is applied to the uncorrelated residuals obtained from the SARMA model. The detection results of the EWMA chart are compared with two other commonly applied residual-based tests: a Shewhart indivi-duals chart and a Cumulative Sum (CUSUM) control chart.

1. Introduction

1.1. The state of the art

Demand for enhanced production system safety is increasing to maintain safe and reliable system operations. In other words, reliable detection and diagnosis of anomalies is an important task to preserve high quality products as well as system safety. Since abnormalities in complex production systems are difficult to avoid and may result in serious degradations, it is important to be able to detect potential anomalies in the monitored system as early as possible. Early anomaly-detection can therefore provide significant improvements in the reduction of operational and maintenance costs, system down-time, and lead to increased levels of safety, which is of ever growing importance. Anomaly-detection is essential for proper and safe operation of various production systems, and it has recently become even more important than ever before. For example, anomaly-detection in hospital systems, particularly emergency departments (EDs) are among the most growing areas of concern for many countries. In particular, monitoring of patient flow in EDs is an urgent and critical issue for many hospital administrations in France and worldwide. This is due to the fact that emergency departments are: i) a key component of healthcare systems that should provide immediate care and vital lifesaving services, ii) the growing demand for emergency medical care and the reduction of the emergency departments and iii) the most overcrowding component. Of course, early detection of abnormal situations plays an important role regarding system management, reliability, and safety (e.g. manufacturing industries, shop floor, hospital systems…).

The primary objective of anomaly-detection is to identify any abnormal event indicating a distance from the process behaviour compared to its nominal behaviour [1]. In other words, anomaly-detection is a binary decision making process; either the system is under normal, or abnormal operating conditions. Abnormal situations occur when systems deviate significantly (outside the allowed range) from their normal regime during online operation [2,3]. Generally, two different kinds of anomalies can be distinguished by their form, anomalies may develop abruptly

---

(stepwise) or incipiently (driftwise) [1]. These types of anomalies are widely encountered in practice. For example, in an EDs slow or gradual anomalies usually indicate a slow increasing demand or patient flow, while abrupt anomalies, are characterized by sudden increasing demands (patient flow). The need for monitoring techniques that can accurately and quickly detect abnormal situations (abrupt or incipient anomalies) has greatly attracted the attention of researchers and engineers. Over the last decades, there has been resurgent interest in monitoring approaches for safe system operations. Generally, anomaly-detection techniques can be classified into three main categories: data-based or model-free techniques, model-based techniques, and knowledge-based techniques [4]. Fault detection using knowledge-based techniques is usually a heuristic process [5].The main limitation of these methods is that they are more suitable to small-scale systems having a small number of variables, and thus may not be appropriate to monitor complex processes. Furthermore, data-based anomaly-detection methods rely on the availability of historical data of the inspected system under normal operation mode [5–7]. These data are first used to build an empirical model, which is then used to detect faults in future data. However, the applicability of data-based methods mainly depends on the availability of quantity and quality of the data [8–12]. On the other hand, model-based anomaly detection approaches are based on the comparison of measured process variables with information retrieved from an explicit mathematical model, usually derived from a basic understanding of the inspected process under normal operating conditions [1,13,14]. These methods have received more attention from engineers and practitioners because of their mathematical and systematic characteristics.

1.2. Motivation

Statistical process control (SPC) have been extensively employed to successfully monitor various production systems. Control charts are one of the most frequently used procedures in SPC, and have been widely used as a monitoring tool in quality engineering [15]. When the goal is to monitor only one process variable, univariate control charts, such as Shewhart [16], cumulative summation (CUSUM) [17], and EMWA [18] can be used . The majority of SPC techniques have been designed to monitor uncorrelated variables. However, in real life the data collected in time are often autocorrelated. Under such conditions, conventional SPC control charts (Shewhart, CUSUM and EWMA charts) are ineffective. It is pointed that the autocorrelation affects the false alarm rate and the shift detection power [2]. False alarm rates are often high when SPC control charts are used to monitor highly auto-correlated processes [2]. Time-series models are commonly utilized for modelling auto-correlated processes [19]. If the right time-series model is available for describing an autocorrelated process, the residuals can be approximately uncorrelated and monitored using traditional SPC charts. The overarching goal of this paper is to exploit the advantages of univariate time-series models to describe the autocorrelation and then to apply SPC techniques to the uncorrelated residuals of the appropriate time-series model by successfully combining time-series analysis with SPC schemes.

1.3. Contribution

As mentioned above, the traditional statistical monitoring charts, which assume that data are uncorrelated, suffer some major shortcomings when this assumption is violated. Therefore, alternative mechanisms or schemes have to be employed to overcome the limitations. When a time series model is available for capturing the autocorrelation in the process data, the residuals which are the difference between the process measurements and the output of the time series model, can be approximately uncorrelated and monitored by conventional SPC charts. Therefore, the main objective of this paper is to combine the advantages of the univariate EWMA monitoring scheme and time series modelling to enhance their performances and widen their applicability in practice. Such a choice is mainly motivated by the greater ability of the EWMA monitoring scheme to detect small changes in the process mean. In such a framework, the EWMA monitoring scheme is applied to the uncorrelated residuals of the appropriate time series model. More specifically, this work focuses on the integration of time-series modelling and SPC tools for the joint development of a monitoring system to help supervise the behaviour of ED systems. Of course, this study reports the development of a seasonal autoregressive moving average (SARMA) model to describe daily patient arrivals at the paediatric emergency department (PED) in Lille, France. The proposed SARMA-based EWMA anomaly detection strategy is successfully applied to detect signs of abnormal situations caused by daily patient arrivals at the PED in Lille, France.

The remainder of this paper is organized as follows. Section 2 briefly reviews the conventional SPC-based monitoring techniques. Section 3 describes shortly the state of the art in time-series modelling in EDs. Section 4 outlines the methodology used in this work for the early detection of abnormal situations in production systems. Section 5 presents the application of the proposed methodology in the early detection of abnormal situations at the PED in Lille regional hospital centre, France. The data set used in the case study is described. The detection results of the proposed SARMA-based EWMA anomaly detection methodology are compared with those of the Shewhart I-chart and CUSUM control chart. Finally, Section 6 reviews the main points discussed in this work and concludes the study.

2. Statistical process control techniques

This section presents the basic principles of statistical process control charts and how they can be used in anomaly-detection.

2.1. SPC overview

The aim of SPC is to monitor a process to detect abnormal behaviour. SPC applications have been widely used in various production systems for monitoring processes and improving product quality [1,20]. Statistical controls charts (also referred to as monitoring charts) are one of the most commonly used tools in SPC and have been extensively used in quality engineering as a monitoring tool to detect the presence of possible anomalies in the mean or variance of process measurements. Control charts play a crucial role in detecting whether a process is still working under normal operating conditions (usually termed in-control) or not (out-of-control). In other words, a control chart is a picture of a process over time that helps to identify the state of the monitored process, i.e., it is running satisfactorily or not [15]. Within such a framework, numerous control charts have been developed to monitor process variable means over time, including: i) Shewhart charts [16], ii) CUSUM charts [17], and iii) EWMA charts [21]. For many decades, the primary utilization of control charts has focused on industrial quality control applications. Nowadays, the use of control charts has been extended to many areas, including healthcare systems [22,23], economics, informatics, and the environment [24–26]. The main characteristics of the control charts (monitored statistic, lower control limit (LCL) and upper control limit (UCL) for alerting, and the output), are summarized in Table 1 [15].

Table 1
Characteristics of the three main univariate control charts.

	Monitored statistic	LCL/UCL	Output
Shewhart	$Shewhart_t = y_t$	LCL, UCL $= \mu_0 \mp L\sigma_0$ μ_0 and σ_0 are the mean and standard deviation of anomaly-free data, respectively.	$S_t = if \; [Y_t > UCL \; and\backslash or \; Y_t < LCL]$
CUSUM	$C_t^+ = \max [0, C_{t-1}^+ + y_t - (\mu_{t.C} + k)]$ where $C_0^+ = 0$, k, $\mu_{t.c}$ and C_t^+ are the slack variable, the in-control mean and the upper CUSUM statistics, respectively. Typically, k is selected to be half of the expected shift in mean.	The threshold UCL$=H$, Usually the reasonable estimate of H is $H = 5 \times \sigma$.	$C_t = if [C_t^+ > H]$ when the C_t^+ statistic exceeds a threshold H, the process is considered to be out of control.
EWMA	$z_t = \lambda y_t + (1-\lambda) z_{t-1}$	LCL, UCL $= \mu_0 \mp L\sigma_0 \sqrt{\frac{\lambda}{(2-\lambda)}}$ where λ is a weighted parameter	$E_t = if[z_t > UCL \; and\backslash or \; z_t < LCL]$
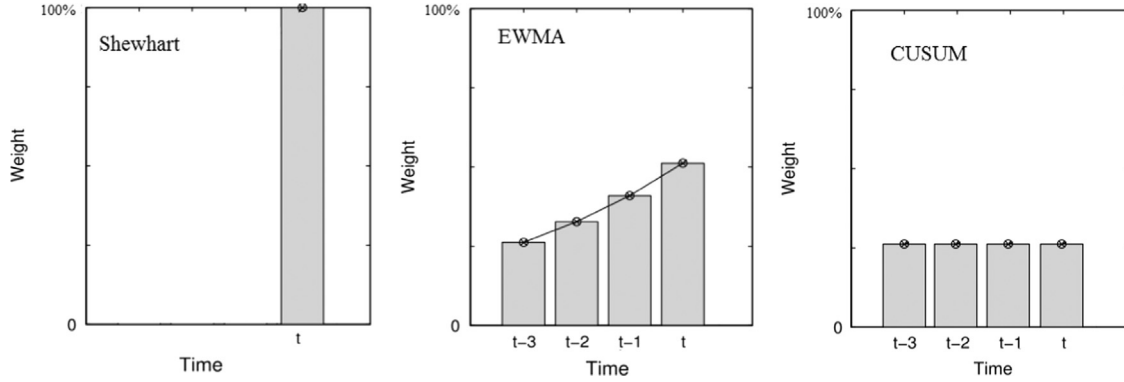
Fig. 1. Univariate process monitoring charts.

Shewhart control charts, were the first method proposed for testing the hypothesis and can be effectively used to detect large shifts in the process mean [27]. However, a key disadvantage of Shewhart charts is that they only use the last data sample from an inspected process and do not retain any memory of the previous data (see Fig. 1). Consequently, these charts are relatively insensitive to small or moderate changes in the process variable mean [15]. Of course, this potentially makes Shewhart control charts less useful in monitoring processes when anomalies relating to small or moderate process mean shifts are of interest. The ability to detect smaller parameter shifts can be improved by using a chart based on a statistic that incorporates information from past samples in addition to current samples. These shortcomings motivate the use of other alternatives, such as EWMA and CUSUM charts, which are better suited to detecting smaller shifts in process mean [15]. This is because the CUSUM and EWMA charts take into account the information contain in the entire process history (see Fig. 1). The CUSUM chart gives equal weights to the entire process history observations when it accumulates all useful information in the historical data. However, since EWMA uses a weighted average of all past and current observations, it is a lot less sensitive to violating the normality assumption than CUSUM charts [28]. Also, CUSUM is relatively slow to respond to large shifts. Therefore, EWMA-based charts are an appropriate monitoring scheme to be adopted when dealing with individual observations [15]. According to the literature, EWMA is one of the most frequently used control charts for monitoring auto-correlated processes because of its flexibility and sensitivity to small shifts [29], which is presented in more details next. In the application presented later in Section 5, the Shewhart and CUSUM charts are used as a benchmark for anomaly-detection using time series modeling.

2.2. EWMA control scheme-based monitoring

In EWMA control schemes, the moving average is calculated by multiplying the historical observations by a weight that decays exponentially over time [15]. The properties of this control scheme and its use in process monitoring have been studied in [18]. The EWMA decision statistic, z_t, is described by the following recursive formula:

$$z_t = \lambda y_t + (1-\lambda) z_{t-1} \tag{1}$$

where y_t is the value of the monitored variable at time t ($t=1,\dots,n$), and λ is the smoothing parameter lying between 0 and 1 (i.e. $0 < \lambda \leq 1$) that determines the temporal memory of the EWMA decision function. The initial value z_0 is set as equal to the in-control process mean. The smoothing parameter determines the depth of memory of the EWMA, i.e. how much memory is retained in the averages. Larger λ values retain less temporal memory, i.e. give more weight to recent data and less weight to older data. However, a small λ value gives more weight to older data. In other words, this coefficient λ acts as a forgetting factor. Generally, in quality control, a smaller λ value leads to quicker detection of smaller shifts [30]. Indeed, λ should be adjusted to an appropriate value according to the characteristics of the process monitored. The value of λ is usually set between 0.2 and 0.3 [18]. The standard deviation of z_t is defined as $\sigma_{z_t} = \sigma_0 \sqrt{\frac{\lambda}{(2-\lambda)}\left[1-(1-\lambda)^{2t}\right]}$, where σ_0 is the standard deviation of the anomaly-free or preliminary data set. The EWMA control scheme declares an anomaly when the value of z_t does not lie within the interval between an upper, UCL, and lower, LCL, control limits. The upper and lower control limits are usually set as: UCL/LCL $= \mu_0 \pm L\sigma_{z_t}$. where L is the width of the control limits (which determines the confidence limits, usually specified as 3 for a false alarm rate of 0.3%).

It is well known that most traditional univariate SPC charts rely on the prior assumption that the process data are not correlated.

Such an assumption is not valid in the case of EDs, where patient flow data are autocorrelated. For such autocorrelated processes, conventional univariate SPC charts, such as Shewhart chart, CUSUM chart and EWMA chart, generally become ineffective and impractical. Indeed, monitoring highly autocorrelated processes based on conventional SPC control charts usually results in high false alarm rates [2,13], or even in faults being undetected. Such a limitation makes conventional SPC techniques inadequate to successfully monitor patient flow in EDs. A possible solution to avoid such limitations to monitor the behaviour of EDS is to use time-series modelling. More specifically, time-series modeling of the correlation present in patient flow variables is advantageous in the sense that residuals resulting from the time-series model are approximately uncorrelated and can therefore be monitored using conventional SPC charts. More details about time-series modelling in EDs are presented next.

3. Time series modeling and EDs

3.1. Time series analysis

A time series is a sequence of measurements indexed over time (t). One of the main objectives of time-series analysis is to identify the nature of the phenomenon represented by the sequence of observations. The literature review shows that time-series analysis has been largely applied for modelling and analysis in various fields such as electrical energy [31], the environment [32], and health care establishments [33].

The three univariate time-series models which have been widely applied in the literature are autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) processes. Adding non-stationary models to the mix leads to the ARIMA model (autoregressive integrated moving average), popularized in the work by [34]. When a time series exhibits seasonality, a SARIMA model can be used to measure the seasonal effect or eliminate seasonality. To take into account explanatory variables (meteorological measurements (temperature, wind direction...), epidemics...) the autoregressive moving average model with external inputs (ARMAX) and its variants (ARX, ARIMAX, SARIMAX...) is suited [35,36]; The results presented in the literature show that ARMA models and their variants [34] provide a very accessible time-series analytical tool both in terms of methodological constraints and the level of mathematical models used, which are less complex linear stochastic equations. In this study, univariate time-series analysis was used to describe the variation in demand for emergency care in an ED system.

3.2. Time-series modelling

The ARMA model of order (p, q) denoted auto-regressive moving average, ARMA (p,q), is written as:

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{2}$$

where:

y_t is the variable to be predicted using previous samples of the time series

φ_i, θ_j are non seasonal parameters.

ε_t, is a sequence of i.i.d. (independent and identically distributed) terms which have zero mean.

Using the backshift operator, B, defined as $B^k y_t = y_{t-k}$, we can re-write Eq. (2) as follows:

$$\varphi_p(B)y_t = \theta_q(B)\varepsilon_t,$$
$$\varphi_p(B) = (1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)$$
$$\theta_q(B) = \left(1 - \theta_q B - \theta_q B^2 - \ldots - \theta_q B^q\right) \tag{3}$$

In many cases, ARMA model is a very useful method to fit time-series data. However, when a time series exhibits seasonality, SARMA models that exploit the correlation between the data at successive periods of time are used. The generalized form of a SARMA model can be written as [37]:

$$\boldsymbol{\varphi_p(B)\Phi_p B^s y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t}$$
$$\varphi_p(B) = (1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)$$
$$\Phi_P B^s = \left(1 - \Phi_1 B - \Phi_2 B^2 - \ldots - \Phi_q B^P\right)$$
$$\theta_q(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right) \tag{4}$$
$$\Theta_Q(B^s) = \left(1 - \Theta_1 B - \Theta_2 B^2 - \ldots - \Theta_Q B^Q\right)$$

where, ∇ is the differencing operator ($\nabla Y_t = y_t - y_{t-1}$), $\Phi_P B^s$ is the polynomial of Q seasonal AR coefficients, and $\Theta_Q(B^s)$ is the polynomial of Q seasonal MA coefficients.

Such a model is denoted SARMA (p, q)(P, Q)S, in which (p,q) represents the non-seasonal part,(P, Q) represents the seasonal part, and S represents the length of the seasonality.

3.3. Time series and emergency department admissions

In the case of emergency department admissions, increasing attention has been accorded to time-series models to predict patient arrivals in hospital establishments [33,38]. The literature review shows that time-series analysis has been largely applied in the hospital sector to forecast patient arrivals, length of stay, and for projecting the utilization of inpatient days. In this regard, some of the recent research efforts made by previous researchers deserve mentioning. Jones et al. [39] used Box–Jenkins models to forecast the number of daily emergency admissions and the number of beds occupied by emergency admissions on a daily basis at Bromley Hospitals NHS Trust in the United Kingdom. Rodríguez and Hernández [40] studied ED patterns in a hospital in Tenerife, Spain. Their analysis was based on a time series of the number of ED presentations every hour over a 6-year period (1997–2002). Alkan and Kiymik [41] used autoregressive (AR) and Welch spectral estimation methods to analyze the electro-encephalogram (EEG) signals. The parameters of autoregressive (AR) method were estimated by using Yule–Walker. The results demonstrate superior performance of the covariance methods over Yule–walker AR and Welch methods. Earnest et al., [42] used ARIMA models to forecast ward occupancy, due to SARS infection, for up to three days in advance. The authors used seasonal decomposition methods to obtain detailed estimates of the effects of the hour of day, the day of week, and the week of the year on attendances. Gooijer and Hyndman [43] reviewed the past 25 years of research into time-series forecasting covering the period 1982–2005. The authors divided the time-series models into linear and nonlinear models. They discussed the strengths and weaknesses of each method. Barişçi [44] applied an adaptive auto regressive-moving average (A-ARMA) to analyze the electromyographic (EMG) signals. The author found that A-ARMA present good models with few parameters and their ability to, with small p and q, represent a very rich set of stationary time series in a parsimonious way. Jones [38] evaluated the use of four statistical methods (Box–Jenkins, time-series regression, exponential smoothing, and artificial neural network models) to predict daily

ED patient volumes over 27 months (from 2005 to 2007) in three different hospital EDs in Utah and southern Idaho (USA). The authors compared four models with a benchmark multiple linear regression model already available in the emergency medicine literature. They claim that regression-based models incorporating calendar variables account for site-specific special-day effects, and allow for residual autocorrelation providing a more appropriate, informative, and consistently accurate approach for forecasting daily ED patient volumes. Jones [45] studied the temporal relationships between the demands for key resources in the emergency department (ED) and the inpatient hospital in 2006, and developed multivariate forecasting models. The authors compared the ability of the models they developed (multivariate models) to provide out-of-sample forecasts of ED census and the demands for diagnostic resources with a univariate benchmark model. Sun et al., [46] used a time-series method to predict daily emergency department attendances. The authors applied an autoregressive integrated moving average (ARIMA) method separately to three acuity categories and total patient attendances at the ED of an acute care regional general hospital from July 2005 to March 2008. They concluded that time-series analysis provides a useful, readily available tool for predicting emergency department workload that can be used for staff roster and resource planning. Nouira and Trabelsi [47] presented a hybrid system (called Online Multi-Agent Monitoring System) to solve the patient's critical state monitoring problems in real time. The medical monitoring system combines multi-agent approaches with time-series models (ARIMA).

In the next section, the SPC charts will be integrated with time-series modelling to extend their anomaly-detection abilities for the case where a data are correlated in a medical context.

4. Time-series analysis model-based SPC control chart methodology for abnormal situation detection

The detection of abnormal situation can be accomplished using statistical analysis and control chart methods. To this end, a time-series analysis model was integrated with SPC charts to develop a methodology for the monitoring these situations. The methodology adopted is outlined in the following steps, as shown in Fig. 2.

Step 1: Construction of the reference descriptive model. The reference model is obtained by identifying the most relevant time-series model. This step is composed of three main sub-steps presented as follows:

1. Collection of the training data set (anomaly free data) representative of a normal situation. The training data set corresponds to the normal operating conditions of the inspected system. This is necessary for the construction of the reference model.
2. Data pre-treatment, which consists in the analysis (trends, seasonality in the time series) and scaling of the data used to develop the system model to zero mean and unit variance.
3. Build the reference descriptive model using the training data set defined which represents normal situations. The objective is to find and validate the best descriptive model based on the time-series analysis of the training data.

Step 2: Computation of the control chart limits. In this second step, based on the residuals obtained from the reference model, the upper control limit (UCL) and the lower control limit (LCL) of the control chart(s) can be computed.

Step 3: Detection of anomalies. This step is composed of five main sub-steps:

1. Definition of the testing data set (anomaly data) that may possibly contain abnormal situations. The testing data set corresponds to abnormal operations in the inspected system.
2. Data pre-treatment, which consists in the analysis and the scaling of the testing data with the mean and standard deviation.
3. Generation of the residual vector (ε_t), which is the difference between the measurements and the output of the reference time-series model constructed in step 1.
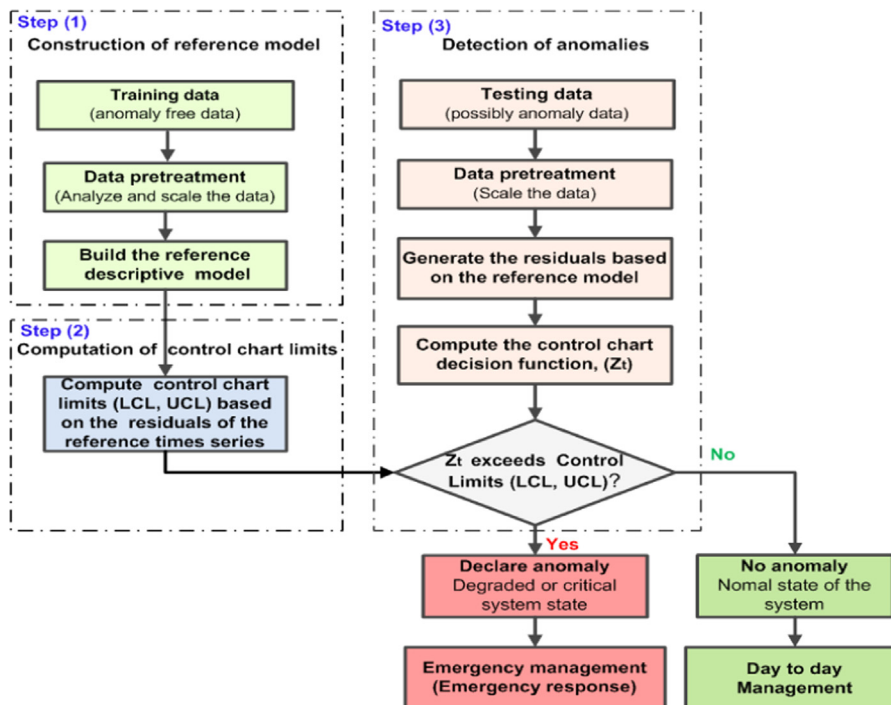

Fig. 2. Flowchart of the proposed abnormal situation detection methodology.

4. Computation of the control chart decision function, z_t. The computation of z_t for individual observations in the case of an EWMA control chart is summarized in the next subsection.
5. Verification for abnormal situations. If the control chart decision function (z_t) exceeds the control limits (UCL, LCL) an anomaly is declared. In this case, emergency management (corrective actions and emergency measures) is triggered in order to attempt to return to an acceptable operating state. If no anomaly is detected (z_t does not exceed the control limits), the day-to-day management is maintained.

To detect abnormal situations, the difference between the measurements and the reference model (residuals) may be used as an indicator. The residuals are close to zero when the behaviour of the monitored system is normal. However, when an abnormal situation occurs, the residuals deviate significantly from zero indicating the presence of a new situation that is significantly distinguishable from the normal situation.

As mentioned in Section 2.2, the design of the EWMA chart is flexible since it allows practitioners to select the right parameters to achieve the best possible performance. Also, the EWMA chart has the best detection time performance and its control limits can be easily determined. This control chart was thus chosen for the abnormal situation detection problem in the case study (Section 5). Details regarding the time-series modelling and the EWMA control scheme are discussed in the following two subsections.

5. Case study: the paediatric emergency department (PED)

5.1. Abnormal situation in an emergency department (ED)

The emergency department is the most crowded hospital department. To deal with the influx of patients, emergency departments require significant human and material resources, and a high degree of coordination among human and material elements. Unfortunately, these resources are limited. This patient influx often leads to strain situations in these establishments [48]. Consequently, ED managers need to continuously monitor ED behaviour, as well as the strain situations often generated by the influx of patients, using detection and control strategies. These control strategies may help ED managers to control problems related to the growing demand for emergency medical care, to prepare for the increase in care activities and to optimize internal resources.

From a patient flow viewpoint, an abnormal situation in an emergency department (ED) is defined as a disequilibrium that may trigger a strain phenomenon, in which threshold value(s) are exceeded, between [48,49]:

- The demand for care (patient flow), and
- The care production capacity (supply).

The main factors identified that may affect this equilibrium are:

- Factors that influence the number of inputs (patient flow): seasonal epidemics (in winter: influenza, colds, gastroenteritis, bronchiolitis, etc.; in summer: trauma), health crises, cumulative causes such as the aging population…,
- Factors that influence the care production capacity: e.g. care system performance, number and competence of medical staff (capacity, experience feedback, availability of physicians downstream), and internal and external transfer capacity (availability of care services downstream).

Evaluating these situations imperatively requires relevant indicators [48]. For example, the number of patient arrivals at the ED is an important indicator in the modelling and management of these situations in such establishments.

The objective of this case study is to detect early abnormal situations generated by the influx of patients at the paediatric emergency department (PED) in Lille regional hospital centre, France.

5.2. SARMA-based EWMA methodology for detecting abnormal patient arrivals at the PED

The best way to manage abnormal situations caused by the influx of patients at the PED is to detect them as soon as possible to allow the PED managers to prepare for the increase in care activities and to have the flexibility (enough time) to implement corrective actions. In this sub-section, based on Fig. 2, the SARMA model is integrated with the EWMA chart to develop a methodology for the early detection of the occurrence of these situations in a PED. The main steps of the SARMA-based EWMA strategy are summarized in Table 2.

5.2.1. Training data set

This retrospective study was conducted utilizing a data set extracted from the database of the PED from November 2011 to March 2012 (epidemic period). The data used in this paper are the time series of hourly patient attendances at the PED. Before attempting to use these data, it was critical to conduct a preliminary descriptive analysis of the latter.

5.2.2. Data pre-treatment

Fig. 3 presents the actual number of patient arrivals per hour for each day of the epidemic period. According to this figure, the number of patients arriving at the PED varied considerably according to the hour of day. It also shows that the maximum arrival rate (A_r) at the PED was $A_r = 13$ patients/hour, which representsmore than onepatientper 5 min.

Fig. 4 illustrates the box plots of the mean hourly patient arrivals on every day of the week. This figure confirms that a pattern exists over the day; the green vertical dotted lines denote the transition between successive days.

The data presented in Fig. 4 were used as the anomaly-free training data and represent typical hourly patient arrivals at the PED. These data were used to develop a descriptive reference

Table 2
Steps of the proposed abnormal situation detection strategy.

| Step	Action
1.	**Given** • A training anomaly-free data set that represents the normal state of the PED and a testing data set (possibly anomaly data), • The parameters of the EWMA control scheme: smoothing parameter λ and the control limit width L,
2.	**Data preprocessing** • Scale the data to zero mean and unit variance,
3.	**Build the SARMA model using the training anomaly-free data** • Compute the residuals from the selected SARMA model, • Compute the EWMA control limits based on the residuals of the reference model
4.	**Test the new data** • Scale the new data with the mean and standard deviation obtained from the training data, • Generate the residuals based on the reference SARMA model, • Compute the EWMA decision function, z_t
5.	**Check for anomalies** • Declare a strain situation when the EWMA decision function, z_t, exceeds the control limits.
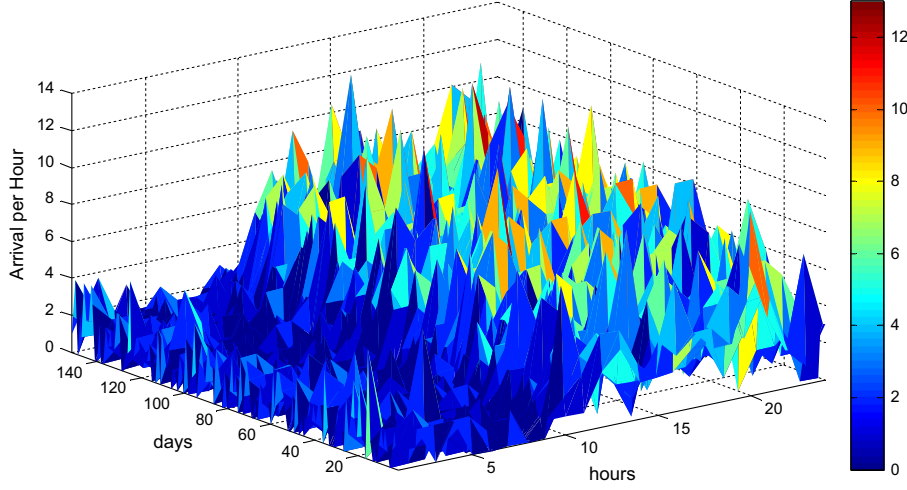
Fig. 3. Actual patient arrivals per hour for each day from November 2011 to March 2012.
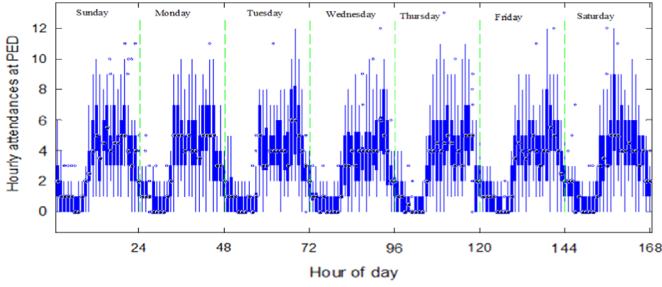

Fig. 4. Box plots of the number of mean hourly patient arrivals at the PED from November 2011 to March 2012.

model representing the nominal behaviour of the inspected PED, which was then used to detect abnormal patient arrivals in future data.

The data series of weekday patient arrival patterns at the PED over a 24-hour period and the corresponding auto-correlation function (ACF) are shown in Figs. 5a and 6b, respectively.

According to Fig. 5a, it can be seen that the pattern of patient arrivals per hour at the PED follows a typical cycle; the vertical dashed lines denote the transition between successive days. This figure presents seasonality at regular 24-h intervals. Fig. 5b shows a clear periodicity of the ACF every 24 h (period between adjacent peaks in the ACF). These data, which are shown in Fig. 6a, were scaled to be zero mean with a unit variance, and then used to develop a model using the following equation:

$$\left(\overline{Y} = \frac{Y - \mu_y}{\sigma_y}\right) \tag{5}$$

where σ_y is the standard deviation and μ_y is the time-series mean.

5.2.3. Build the reference SARMA model

Successful modeling relies generally on four main steps: data analysis and pretreatment, model selection, model estimation and validation. Data analysis and pretreatment usually consists to extract main characteristic features of studied system from a subset of training data, which therefore help us for the model selection step. Model estimation stage aims at determining optimal parameter values of the selected model that perfectly describes the training data. The last step consists in validating the selected model and checking the accuracy of the model to fit the data. The construction and validation of the best reference model of hourly patient attendances at the PED is presented in this sub-section.

5.2.3.1. Model estimation and selection. From the time-series plots depicted in Fig. 5, it is apparent that the time series is stationary and has seasonality ($s=24$). This led us to study a seasonal auto-regression moving average model SARMA(p, q) × (P, Q)$_{24}$. In this study, the parameters of the SARMA model were estimated using the maximum likelihood method for each data set [34]. The best-fit model for total attendances was SARMA($1,1$) × ($1,0$)$_{24}$.The mathematical equation of the SARMA model for hourly patient attendances at the PED is presented in Eq. (6), and the estimated parameters with their sampling standard deviation are shown in Table 3.

$$\left(1 - 0.98B^{24}\right)(1 - 0.77B)Y_t = (1 - 0.015B)\varepsilon_t \tag{6}$$

5.2.3.2. Model validation. To evaluate the goodness or the accuracy of the model to fit the data, various statistics have been developed [50]. In this study, mean-absolute-percentage-error (MAPE) and R^2 (known as the coefficient of determination) have been used to assess the goodness of the constructed SARMA model. Table 4 presents the statistical validation measures applied to the time series. As shown in Table 4, the high R^2 and the low *MAPE* of the SARMA reference model show that the latter is very representative of the observed time series.

5.2.3.3. Residual analysis. When conducting a residual analysis, a scatter plot of the residuals is one of the most commonly used plots to detect nonlinearity and outliers. Generally, if the points in a residual plot are randomly dispersed around zero, a linear regression model is appropriate for the data, if not a non-linear model is more appropriate. Fig. 6 shows the residuals of the selected SARMA model.

The verification of the normality of the residuals using Henry's line and the histogram for the selected SARMA model is illustrated in Fig. 7 (left) and (right), respectively. Fig. 7 shows that the assumption of a normal distribution for the residuals appears to be reasonable.

The independence of the residuals (absence of autocorrelation) was then verified. The residuals were assumed to be non-auto correlated. To determine whether residuals are non-auto correlated, the ACF of the residuals is examined. If the assumption is satisfied, the ACF of the residuals should be large for any non-zero lag. If the ACF is significantly different from zero, this implies that there is dependence between observations. The ACF of the residuals for the selected SARMA model presented in Fig. 8 indicates that the residuals were not significantly different from a white
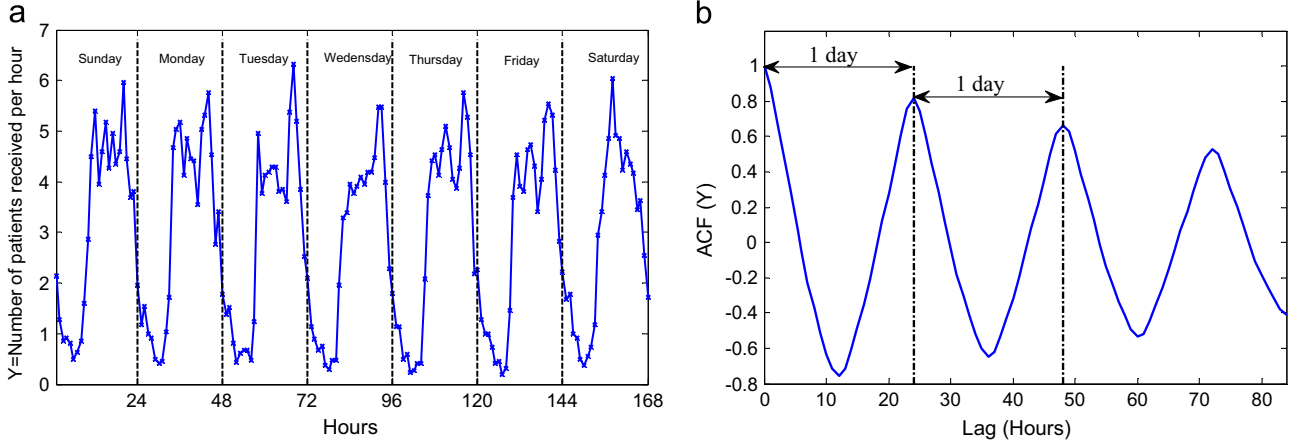
Fig. 5. (a) Number of patients received per hour of day, and (b) ACF of hourly patient attendances at PED.
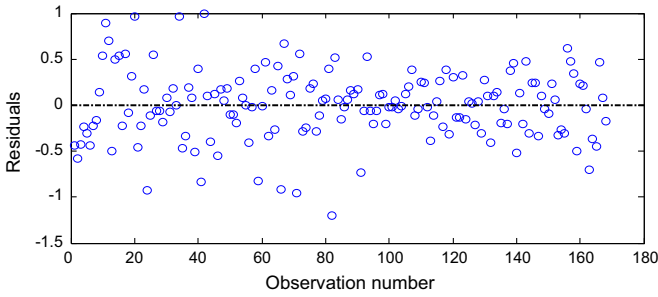

Fig. 6. Residuals of the selected reference SARMA model.

Table 3
Model parameters.

Parameters	Θ	φ	θ
Values	0.98 ± 0.012	0.77 ± 0.022	0.015 ± 0.02

Table 4
Statistical validation measures applied to the reference model.

Parameters	SARMA$(1, 1) \times (1, 0)_{24}$
R^2	0.89
MAPE	3.15%

noise series. According to Fig. 8, the residuals are approximately uncorrelated. As the residuals are normally distributed and uncorrelated, it can be deduced that the constructed SARMA model fits the data well. Once developed, SARMA models can be combined with the EWMA monitoring chart for detecting abnormal patient arrivals at the PED.

5.2.4. Detection results

In this sub-section, the performance of the proposed SARMA-based EWMA anomaly detection algorithm is assessed and compared to that of the SARMA-based Shewhart and CUSUM methods through its utilization to monitor the occurrence of abnormal situations generated by abnormal patient arrivals at the PED.

In the absence of testing data from the PED database containing real anomalies, two different anomalies were simulated in this study to assess the performance of the proposed anomaly

detection strategy. Indeed, we simulated anomalies whose magnitude is chosen according to guidelines provided by an expert of EDs. These simulated case studies were very helpful because they enabled a theoretical comparison between the various monitoring charts as the simulated anomalies in these case studies were known. To do so, before pre-processing the data, additive anomalies were injected into the raw data. Two different types of anomalies were simulated: abrupt anomalies and gradual anomalies. The two different types of anomaly are summarized in Table 5 and shown in Fig. 9.

Two case studies are presented in this section.

- In the first case study (case A), an abrupt increase in patient arrivals at the PED was considered as a result of a sudden increase in the number of patient arrivals,
- In the second case study (case B) it was assumed that the number of patient arrivals gradually increased and the abnormal situation slowly developed; this abnormal situation was termed gradual abnormal situation.

In this study, the performances of studied monitoring charts are validated by missed detection and false alarms rates. False alarm and missed detection rates are the most important detection performance criteria, from a practical point of view. A false alarm is an indication of an anomaly, when no anomaly has occurred. Missed detection denotes that a faulty data are misclassified as faultless data. Missed detection rate (MDR) is the number faulty data that does not exceed the control limits (missed detection) over the total number of faulty data.

$$\text{MDR} = \frac{missed\ detection}{faulty\ data}\%.$$

The FAR is the number of false alarms over the total number of faultless data and is defined as:

$$\text{FAR} = \frac{\text{false alarms}}{faultless\ data}\%.$$

The smaller values of the miss detection and false alarm rates show the better performance of the corresponding monitoring chart.

5.2.4.1. Case A: abrupt abnormal patient arrivals.
In this first case, the problem of detecting abrupt abnormal patient arrivals was investigated. The testing data set was firstly scaled using the mean and standard deviation of the anomaly-free training data. The SARMA model built using the anomaly-free training data was used in this case to detect possible anomalies using unseen testing data. Two sub-cases are presented here to illustrate the ability of the
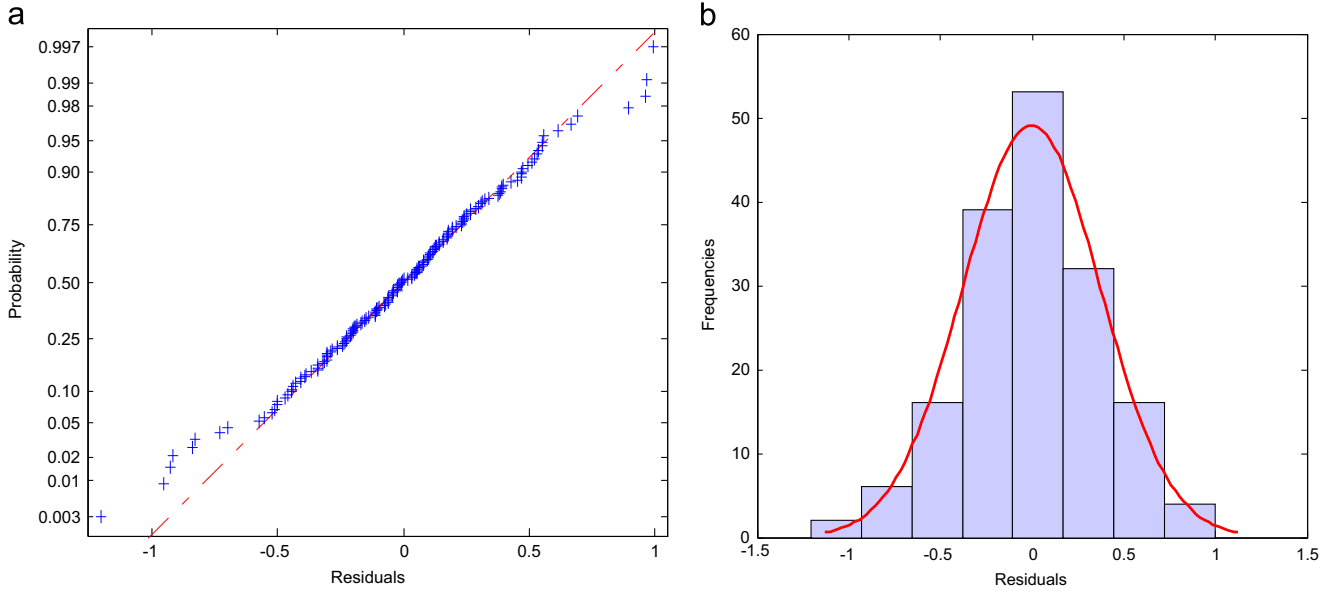
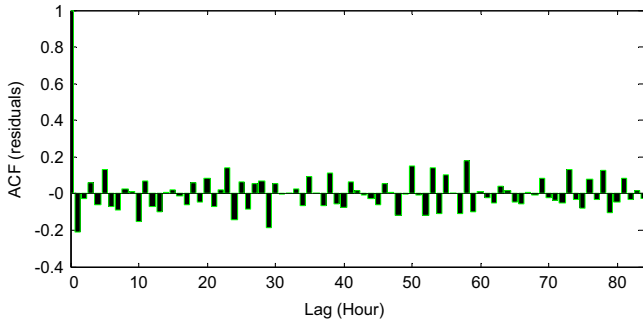**Fig. 7.** Gaussian distribution test: Henry's line (left); histogram (right).

Fig. 8. ACF of residual errors of the selected SARMA model.

Table 5
Anomaly types tested in the simulation study.

Cases	Anomaly type	Description	Parameters
Case A	Abrupt anomaly	$y_t = y_t + b$	b: the rate of increase (%)
Case B	Gradual anomaly	$y_t = y_t + s(k - k_s)$	S: slope, and k_s: starting point of the gradual anomaly

proposed SARMA-based EWMA control scheme to detect abrupt abnormal patient arrivals (Table 6).

In this first case study $\mathbf{A}_1$, the anomaly-detection abilities of SARMA-based EWMA anomaly-detection chart and the SARMA-based CUSUM and Shewhart charts are assessed in the presence of abrupt abnormal patient arrivals. The Shewhart I-chart statistic based on the testing data is shown in Fig. 10. The blue shaded area represents the zone where the simulated abnormal patient arrivals were produced. The dotted lines represent the lower and upper control limits of the chart, which are calculated using the training anomaly-free data and a 95% confidence level. Fig. 10 shows that the anomaly is readily detectable by the Shewhart I-chart, even though some observations in the faulty region do not exceed the Shewhart I-chart confidence limits (i.e., with several missed detections). The false alarm rate and missed detection rate are 0 and 50% respectively in this case. This is maybedue to the fact that the Shewhart monitoring chart evaluates the process perfor-mance based on the observed data at each individual time point; it

is less efficient in detecting relatively small and moderate anomalies. The Shewhart individual chart is indirectly considered in this study for comparison because it is a special case of the EWMA chart when $\lambda = 1$.

To overcome the limitation of Shewhart chart in detecting moderate or small anomalies, the CUSUM-based monitoring chart try to make use of all observed data available at the current time point, including those observed at the current time point and all historical data that are observed before the current time point, to evaluate the process performance. The results of the SARMA-based CUSUM anomaly-detection chart plotted in Fig. 11 clearly show the violation of the confidence limits and thus the ability of this anomaly detection algorithm to detect this anomalous behaviour. This highlights the importance of taking into account the infor-mation in the historical data in decision making procedure. However, Fig. 11 also shows that SARMA-based CUSUM chart is resulted in several false alarms. The false alarm rate is 0.30%. From Fig. 11, we can see that once the abnormal situation ends, the CUSUM chart takes time to return to the nominal state. In fact, using these features leads to poor performance in terms of false alarms. This is expected because of the nature of the CUSUM testing data set algorithms whereby when the rate of change of the CUSUM statistic is negative, the accumulator is reset to zero. This, in turn, accelerates the return to the statistical in-control state. In the chart, the reference value k and the decision limit h are chosen to be 0.25 and 2.044, respectively, so that the ARL0 value of the chart is 200 (i.e., false alarm rate $\alpha = 0.5\%$).

Fig. 12 depicts the monitoring results of the SARMA-based EWMA anomaly detection chart based on the testing data. The parameters of the EWMA are chosen to be: the smoothing para-meter $\lambda = 0.3$, and the control limit width $L = 3$. The dashed hor-izontal lines represent a 95% confidence limit used to identify the possible anomalies. Fig. 12 clearly shows the violation of the con-fidence limits and thus the ability of the proposed SARMA-based EWMA monitoring chart to detect this abnormal situation correctly (substantial patient flow) without any false alarms. The anomaly was clearly detected in the SARMA-based EWMA monitoring chart and its detection rate was also highest among the studied charts. Overall, the memory control charts, CUSUM and EWMA charts, had fewer missed detection than the Shewhart chart.

In the second example, a moderate bias level, 30% of the total variation in the testing data, was injected between samples 112–120.
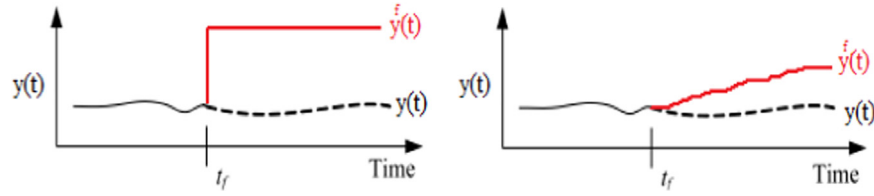
Fig. 9. Anomaly profiles: abrupt anomaly (left); gradual anomaly (right).

Table 6
Case A: abrupt abnormal patient arrivals.

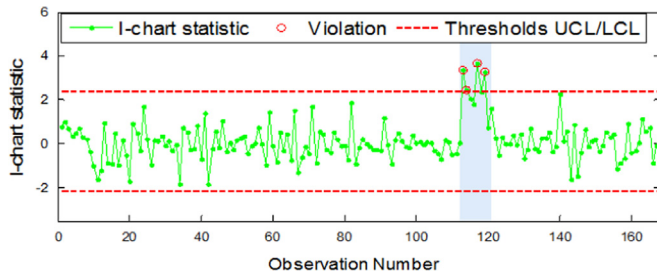| Case	Description of abnormality
Case A_1	An additive anomaly was introduced in Y from sample numbers 112 to 120. This anomaly is represented by a constant bias in amplitude equal to 70% of the total variation in Y
Case A_2	A 30% of the total variation in the testing data was injected between samples 112 to 120

Fig. 10. Time evolution of the ShewhartI-chart statistic in the presence of an abnormal situation (Case A_1).
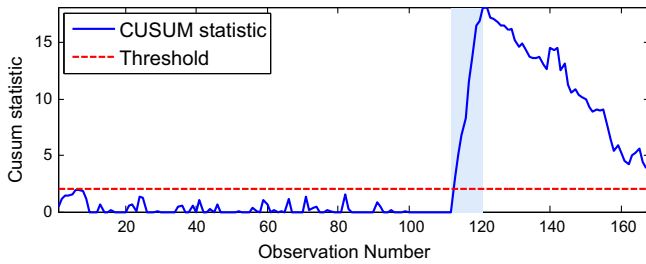

Fig. 11. Time evolution of the CUSUM statistic in the presence of an abnormal situation (Case A_1).
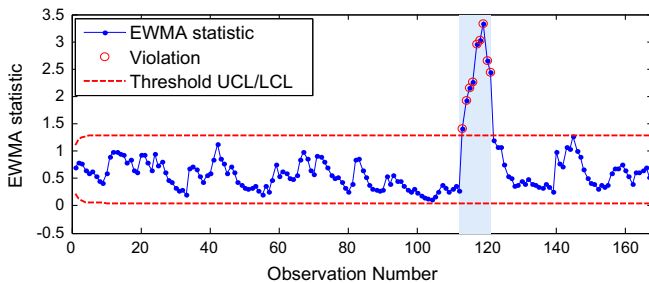

Fig. 12. Time evolution of the EWMA statistic in the presence of an abnormal situation (Case A_1).

The monitoring results through the Shewhart I-chart monitoring statistic are shown in Fig. 13, which indicates that no anomaly has been detected through this testing. This result shows that the Shewhart I-chart is completely unable to detect this moderate simulated anomaly. This is because this monitoring chart only takes into account the information provided by the present data samples in the
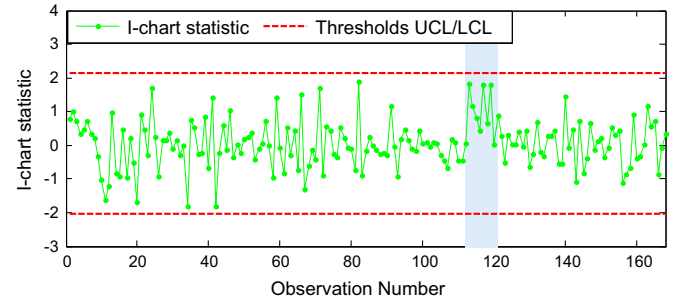

Fig. 13. Time evolution of the I-chart statistic in the presence of an abnormal situation (Case A_2).
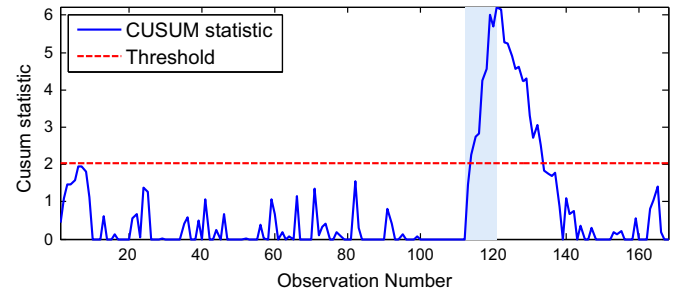

Fig. 14. Time evolution of the CUSUM statistic in the presence of an abnormal situation (Case A_2).

decision making process, which makes this monitoring chart not very powerful in detecting small or moderate anomalies. Indeed, the history data obtained before actual point contain useful information for process monitoring; but, the Shewhart chart ignores such information at all. For this reason, they would not be effective for detecting abrupt anomalies, especially in cases when the shifts are relatively small. As shown by these results, the Shewhart statistic exhibits low capability for incipient anomaly detection.

The performances of the SARMA-based CUSUM chart are demonstrated in Fig. 14. This figure clearly shows the violation of the confidence limits and thus the ability of the proposed SARMA-based CUSUM anomaly detection algorithm to detect this strain situation correctly (substantial patient flow), but with several false alarms. From Fig. 14 it can be noticed that the CUSUM anomaly detection chart is again superior to Shewhart chart. The performance of the SARMA-based EWMA anomaly detection method, on the other hand, (shown in Fig. 15), shows that it could detect this anomaly without any false alarms. CUSUM keeps reporting false alarms, and EWMA again performs well also for the given anomalous behaviour. For the EWMA chart, the chosen smoothing parameter and control limit width were $\lambda = 0.25$ and $L = 3$, respectively. Based on the above results, it is clear that SARMA-based EWMA monitoring chart demonstrates satisfactory results when compared to Shewhart and CUSUM charts.

To evaluate the results of the monitoring charts, an examination of the percentage of false alarms and missed detections would be essential. The performances of these three residual-based
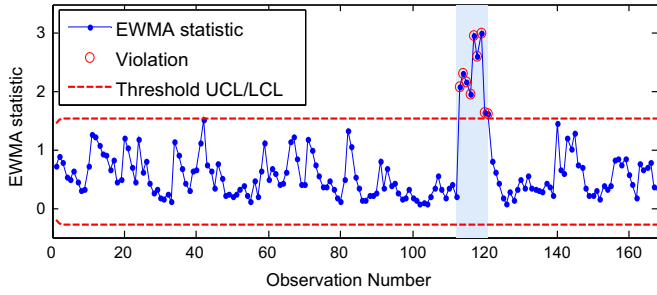
Fig. 15. Time evolution of the EWMA statistic in the presence of an abnormal situation (Case A$_2$).

Table 7
Percentage of false alarms and missed detections for all statistical charts.

Abrupt anomaly	Case A1		Case A1	
Statistical chart	False alarms	Missed detections	False alarms	Missed detection
Shewhart	0	0.5	0	0
CUSUM	0.3	0	0.08	0.12
EWMA	0	0	0	0

monitoring charts are summarized in Table 7. The proposed SARMA-based EWMA control chart has a less number of false alarms and missed detection that SARMA-based Shewhart and CUSUM charts. The results confirm the efficiency of proposed method especially in term of false alarms.

5.2.4.2. Case B: gradual abnormal patient arrivals. A slowly increasing (or ramping) abnormal situation usually indicates a slow increase in the number of patient arrivals at the ED over a long period of time. The aim of this case study was to assess the potential of the proposed SARMA-based EWMA control scheme in detecting slow or gradual abnormal patient arrivals. Towards this end, a slow increase in ED arrivals with a slope of 0.1 was added to the test data starting at sample number 85 of the simulated testing data. In other words, the input variable y_t was linearly increased from sample 85 to sample number 95 of the testing data by adding $0.1(k-85)$ to the y_t value of each sample in this range, where k is the sample number. The result of the Shewhart I-chart algorithm is shown in Fig. 16. According to this figure, there is no anomaly. Thus, this anomalous behaviour remained undetectable when the I-chart was applied. From this case study, it can be seen that the Shewhart chart is not sensitive to either the gradual abnormal patient arrivals or the abrupt abnormal patient arrivals in cases when such anomalies are small. As mentioned above, the major reason for this limitation of the Shewhart charts is that they use observed data at a single time point alone to evaluate the process performance at that time point. On the other hand, the results of SARMA-based EWMA and CUSUM anomaly-detection algorithms for the considered gradual abnormal patient arrivals are shown in Figs. 17 and 18. This figure shows that the SARMA-based EWMA and CUSUM statistics was able to successfully detect this abnormal situation by exceeding the threshold value. The SARMA-based EWMA and CUSUM statistics were gradually increased as the demand slowly increased, and began to violate the control limits when the abnormal situation became sufficiently large to be detected by the given model. The MEWMA and MCUSUM charts begins to increase linearly from the 85th sample and and exceed the control limits at the 91th sample (see Figs. 17 and 18). The proposed method produces fewer false alarms chart to detect this gradual anomaly than of the SARMA-based CUSUM chart.
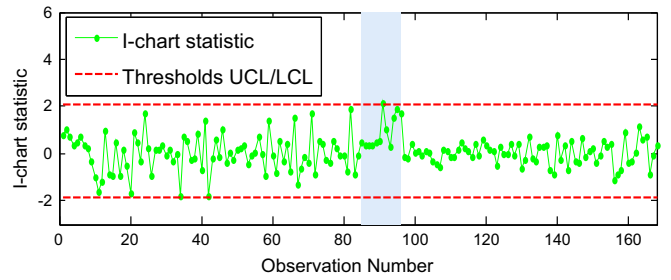

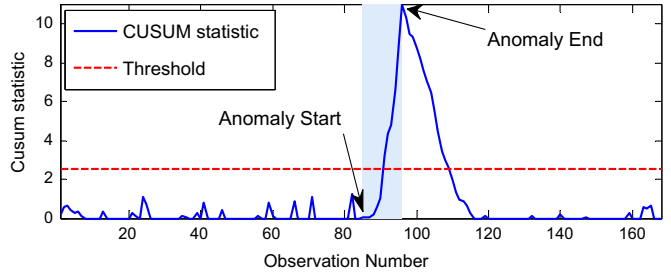**Fig. 16.** Time evolution of the I-chart statistic in the presence of a gradual abnormal situation.

Fig. 17. Time evolution of the CUSUM statistic in the presence of a gradual abnormal situation.
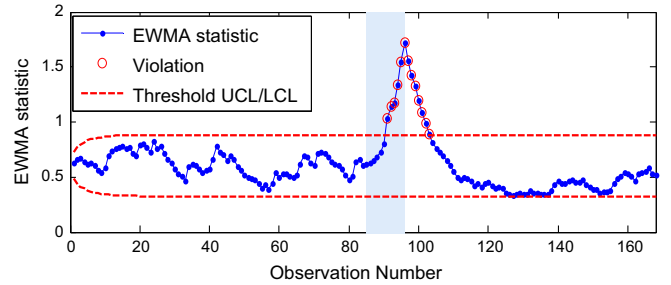

Fig. 18. Time evolution of the EWMA statistic in the presence of a gradual abnormal situation.

In summary, Shewhart monitoring chart, which is based on the uncorrelated residuals obtained from the selected SARMA model, only consider the last data point, making them less efficient in detecting relatively moderate and small anomalies. For this reason, moderate and small anomalies may lead it to report high numbers of missed detection. Alternatively, CUSUM and EWMA monitoring charts consider not only the last data point, but the entire past data. It makes them more sensitive than Shewhart chart to both gradual and incipient anomalies. It has been shown that the performances of univariate monitoring charts can be drastically improved by taking into account the information in the historical data in decision making procedure. CUSUM and EWMA charts are effective in the detection of small anomalies, while Shewhart charts are quick at detecting large shifts. The main difference between CUSUM and EWMA is that CUSUM gives assigns equal weights to all previous observations whereas EWMA weight all available observations based on its importance in characterizing the process, exponentially decaying weights assigned to the historical data. The proposed SARMA-based EWMA anomaly detection chart showed satisfactory anomaly-detection abilities for all anomaly types that are tested in this study compared with the CUSUM and Shewhart chart. The simulation results demonstrate an improved ability of the proposed SAARMA-based EWMA monitoring method over the CUSUM and Shewhart charts, especially for anomalies with small magnitudes.

6. Conclusion

In this paper, a statistical methodology combining time-series models with SPC control techniques was presented for the early detection of abnormal situations that may occur in EDs.

Firstly, this article reports the development of a time-series model to describe hourly patient arrivals at the PED. Based on the data analysed from November 2011 to March 2012, the reference model with the best fit was a seasonal stationary auto-regression moving average model (SARMA). The results indicate that the proposed model provides an acceptable description of PED admissions.

Secondly, an anomaly detection procedure based on the reference SARMA model developed was proposed for the early detection of abnormal situations that may be caused by an influx of patients at the PED. SARMA was used as the modelling framework of the SARMA-based EWMA anomaly detection methodology. Then, the EWMA control scheme was applied to the residuals to detect anomalies when the data did not fit the reference SARMA model. Satisfactory detection results were obtained using the proposed method. To assess the fault detection abilities of the proposed SARMA-based EWMA anomaly detection method, two case studies were considered, one involving an abrupt abnormal situation caused by a sudden increase in the number of patient arrivals, the second one involving a gradual abnormal situation caused by a gradual increase in patient arrivals at the PED. The simulation results of all the cases clearly show the effectiveness of the proposed SARMA-based EWMA anomaly detection methodology over conventional methods, such as Shewhart and CUSUM charts, in detecting anomalies in hourly patient arrivals at the PED, especially small or moderate anomalies.

Early detection of abnormal situations in EDs is beneficial to foster reactive control thus enabling these situations to be avoided and/or the consequences reduced. The latter is important not only for anticipating and managing patients and internal resources in EDs, but also to predict sufficient hospitalization capacity downstream so as to be able to absorb the occasional increase in patient flow without reducing the quality of patient care. Therefore, by exploiting the information provided by this anomaly detection scheme, the performance of the ED can be optimized. The approach proposed and lessons learned from this study may assist other EDs in undertaking their own analyses to aid management and planning.

The next step of the study is to integrate the developed methodology in a decision support system (DSS).This will be a step toward real-time monitoring of ED behaviour aimedat preventing and predicting abnormal situations and examining the relationship between the indicators of these situations and the corresponding corrective actions (action plans).

References

[1] R. Isermann, Fault-Diagnosis Systems - An Introduction from Fault Detection to Fault Tolerance, Springer, 2006.

[2] P. Qiu, Introduction to Statistical Process Control, CRC Press, 2013.

[3] Z. Ge, Z. Song, Multivariate Statistical Process Control: Process Monitoring Methods and Applications, Springer Sci. Bus. Media, 2012.

[4] S.J. Qin, Survey on data-driven industrial process monitoring and diagnosis,", Annu. Rev. Control 36 (2) (2012) 220–234.

[5] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, K. Yin, A review of process fault detection and diagnosis: Part III: process history based methods, Comput. Chem. Eng. 27 (3) (2003) 327–346.

[6] S. Yin, Z. Xiangping, K. Okyay, Improved PLS focused on key-performance-indicator-related fault diagnosis, IEEE Trans. Ind. Electron. 62 (3) (2015) 1651–1658.

[7] S. Yin, X.D. Steven, X. Xiaochen, L. Hao, A review on basic data-driven approaches for industrial process monitoring, IEEE Trans. Ind. Electron. 11 (2014) 6418–6428.

[8] J. Dong, K. Zhang, Y. Huang, G. Li, K. Peng, Adaptive total PLS based quality-relevant process monitoring with application to the Tennessee Eastman process, Neurocomputing 154 (2015) 77–85.

[9] C. Jing, J. Hou, SVM and PCA based fault classification approaches for complicated industrial process, Neurocomputing 167 (2015) 636–642.

[10] Y. Pan, L. Hongyi, Z. Qi, Fault detection for interval type-2 fuzzy systems with sensor nonlinearities, Neurocomputing 145 (2014) 488–494.

[11] F. Harrou, M.N. Nounou, H.N. Nounou, M. Madakyaru, Statistical fault detection using PCA-based GLR hypothesis testing, J. Loss Prev. Process Ind 26 (1) (2013) 129–139.

[12] F. Harrou, F. Kadri, S. Chaabane, C. Tahon, Y. Sun, Improved Principal Component Analysis for Anomaly Detection: Application to an Emergency Department, Comput. Ind. Eng. 88 (2015) 63–77.

[13] R. Razavi-Far, H. Davilu, V. Palade, C. Lucas, Model-based fault detection and isolation of a steam generator using neuro-fuzzy networks, Neurocomputing 72 (13) (2009) 2939–2951.

[14] F. Harrou, L. Fillatre, I. Nikiforov, Anomaly detection/detectability for a linear model with a bounded nuisance parameter, Annu. Rev. Control 38 (1) (2014) 32–44.

[15] D.C. Montgomery, Introduction to Statistical Quality Control, Wiley, Hoboken, NJ, 2012.

[16] W.A. Shewhart, Economic control of quality of manufactured product, D. Van Nostrand Company, Inc., New York, 1931.

[17] E.S. Page, Cumulative sum charts, Technometrics 3 (1) (1961) 1–9.

[18] J.S. Hunter, The exponentially weighted moving average, J. Qual. Technol. 18 (4) (1986) 203–210.

[19] P.J. Brockwell, R.A. Davis, Time Series: Theory and Methods, Springer, 2009.

[20] D. Biau, L. Dessolle, and R. Porcher, Applications and Experiences of Quality Control to Surgical and Interventional Procedures, in Applications and Experiences of Quality Control, ISBN 978-953-307-236-4, Published: April 26, 2011.

[21] S. Crowder, M. Hamilton, An EWMA for monitoring a process standard deviation, J. Qual. Technol. 24 (1) (1992) 12–21.

[22] J.C. Benneyan, The design, selection, and performance of statistical control charts for healthcare process improvement, Int. J. Six Sigma Compet. Advant. 4 (4) (2008) 209–239.

[23] W.H. Woodall, B.M. Adams, J.C. Benneyan, The use of control charts in healthcare, In: F.W. Faltin, R.S. Kenett, F. Ruggeri (Eds.), Statistical Methods in Healthcare, John Wiley & Sons, Ltd, 2012, pp. 251–267.

[24] M. Frisén, Optimal Sequential Surveillance for Finance, Public Health, and Other Areas, Seq. Anal. 28 (3) (2009) 310–337.

[25] F. Harrou, L. Fillatre, M. Bobbia, I. Nikiforov, Statistical detection of abnormal ozone measurements based on constrained generalized likelihood ratio test, In: Proceedings of 52th IEEE Conference on Decision and Control, Florence, Italy., Florence, Italy, December 10-13 2013, pp. 10–13.

[26] Y. Park. A Statistical Process Control Approach For Network Intrusion Detection, 2005, Georgia Institute of Technology, USA.

[27] S.W. Human, S. Chakraborti, C.F. Smit, Shewhart-type control charts for variation in phase I data analysis, Comput. Stat. Data Anal. 54 (4) (2010) 863–874.

[28] D.M. Hawkins, D.H. Olwell. Cumulative Sum Charts and Charting for Quality Improvement, 1998, Springer, ISBN: 978-1-4612-7245-8 (Print) 978-1-4612-1686-5 236-4, Published: April 26, 2011.

[29] N. Boudaoud, Z. Cherfi, A comparative study of CUSUM and EWMA charts: detection of incipient drifts in a mutlivariate framework, Qual. Eng 17 (4) (2005) 703–709.

[30] S. Rabhu, G. Runger, Designing a multivariate ewma control chart, J. Qual. Technol. 29 (1) (1997) 8–15.

[31] Y. Wang, J. Wang, G. Zhao, Y. Dong, Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: a case study of China, Energy Policy 48 (2012) 284–294.

[32] L. Jian, Y. Zhao, Y.P. Zhu, M.B. Zhang, D. Bertolatti, An application of ARIMA model to predict submicron particle concentrations from meteorological factors at a busy roadside in Hangzhou, China, Sci. Total Environ. 426 (2012) 336–345.

[33] F. Kadri, S. Chaabane, F. Harrou, C. Tahon, Time series modelling and forecasting of emergency department overcrowding, J. Med. Syst. 38 (107) (2014) 1–20.

[34] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, 1976.

[35] G.E.P. Box, G.C. Tiao, Intervention analysis with applications to economic and environmental problems, J. Am. Stat. Assoc. 70 (349) (1975) 70.

[36] C. Lim, M. McAleer, J.C.H. Min, ARMAX modelling of international tourism demand, Math. Comput. Simul. 79 (9) (2009) 2879–2888.

[37] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, Time Series Analysis: Forecasting and Control, 4th, John Wiley, Hoboken, N.J, 2008.

[38] S.S. Jones, A. Thomas, R.S. Evans, S.J. Welch, P.J. Haug, G.L. Snow, Forecasting daily patient volumes in the emergency department, Acad. Emerg. Med. Off. J. Soc. Acad. Emerg. Med. 15 (2) (2008) 159–170.

[39] S.A. Jones, M.P. Joy, J. Pearson, Forecasting demand of emergency care, Health Care Manag. Sci 5 (4) (2002) 297–305.

[40] G. Martín Rodríguez, J.J.Cáceres Hernández, A method for ascertaining the seasonal pattern of hospital emergency department visits, Rev. Esp. Salud Pública 79 (1) (2005) 5–15.

[41] A. Alkan, M.K. Kiymik, Comparison of AR and Welch methods in epileptic seizure detection, J. Med. Syst. 30 (6) (2006) 413–419.

[42] A. Earnest, M.I. Chen, D. Ng, L.Y. Sin, 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 (1) (2005) 1–8.

[43] J.G.D. Gooijer, R.J. Hyndman, 25 years of time series forecasting, Int. J. Forecast 22 (2006) 443–473.

[44] N. Barişçi, The adaptive ARMA analysis of EMG signals, J. Med. Syst. 32 (1) (2008) 43–50.

[45] S.S. Jones, R.S. Evans, T.L. Allen, A. Thomas, P.J. Haug, S.J. Welch, G.L. Snow, "A multivariate time series approach to modeling and forecasting demand in the emergency department, J. Biomed. Inform. 42 (1) (2009) 123–139.

[46] Y. Sun, B. Heng, Y. Seow, E. Seow, Forecasting daily attendances at an emergency department to aid resource planning, BMC Emerg. Med 9 (1) (2009) 1–9, http://dx.doi.org/10.1186/1471-227X-9-1.

[47] K. Nouira, A. Trabelsi, Intelligent Monitoring System for Intensive Care Units, J. Med. Syst. 36 (4) (2012) 2309–2318.

[48] F. Kadri, S. Chaabane, C. Tahon, A simulation-based decision support system to prevent and predict strain situations in emergency department systems, Simul. Model. Pract. Theory 42 (2014) 32–52.

[49] F. Kadri, C. Pach, S. Chaabane, T. Berger, D. Trentesaux, C. Tahon, Y. Sallez, Modelling and management of the strain situations in hospital systems using un ORCA approach, IEEE IESM," Rabat - Morocco, 28–30 October 2013, p. 10.

[50] D.G. Mayer, D.G. Butler, Statistical validation, Ecol. Model. 68 (1–2) (1993) 21–32.
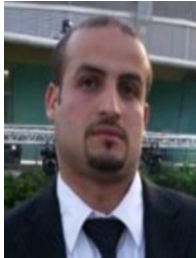
Farid Kadri received the Ing. Diploma in industrial maintenance and security from the Industrial Maintenance and Security Institute (IMSI), University of Oran, Algeria, in 2008 and the M.Sc. degree in Systems Optimization and Security in 2009 from the University of Technology of Troyes, France. In 2014, he received Ph.D. degree in Automation and Computer Engineering from the University of Valenciennes and Hainaut-Cambresis (UVHC), France. From 2014 to 2015, he was an assistant professor at the IUT of Valenciennes. Since 2015, he is an assistant professor at the Arts et Métiers ParisTech, Paris, France. His current research interests include modelling and simulation, statistical decision theory and its applications, fault detection and monitoring, decision support systems, resilience and crisis management.

Fouzi Harrou received the Dipl.-Ing in Telecommunications from Abou Bekr Belkaid University, Algeria, in 2004 and the M.Sc. degree in Telecommunications and Networking in 2006 from the University of Paris VI, France. In 2010, he received the Ph.D. degree in Systems Optimization and Security from the University of Technology of Troyes (UTT), France, and was an Assistant Professor at the UTT, from 2009 to 2010. In 2010, he was an Assistant Professor at the Institute of Automotive and Transport Engineering at Nevers, France. From 2011 to 2012, he was Postdoctoral Research Associate at the Systems Modelling and Dependability Laboratory, UTT. From 2012 to 2014, he was an Assistant Research Scientist, in Chemical Engineering Department at the Texas A&M University at Qatar, Doha, Qatar. Since 2015, he is Postdoctoral Fellow in the Division of Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) at King Abdullah University of Science and Technology (KAUST). His current research interests include statistical decision theory and its applications, fault detection and signal processing, and Spatio-temporal statistics with environmental applications.

Dr. Sondes Chaabane received her engineering degree in computer science engineering (1999) from ENSI, Tunisia and her Ph.D. (2004) from INSA of Lyon, France. She pursued her postdoctoral research (2005) at Catholic University of Mons (FUCAM), Belgium. She is currently full assistant professor at the University of Valenciennes and Hainaut-Cambrésis (UVHC, France) and researcher in LAMIH Laboratory. She is in charge of modeling and Simulation of production systems, operation research, planning and scheduling, performance assessment, and maintenance courses at the ENSIAME Engineering School at this university. Her research activities are based on manufacturing and hospital management problems such as scheduling under uncertainty, optimization and simulation process, strain situations modeling and piloting. Her scientific research includes: human and material resources sizing, planning and scheduling, logistics, hospital management.

Christian Tahon has an ESTP Engineering School degree. He is qualified in Applied Physics, he obtained a PhD (Design of a control system for a centrifuge machine for rolling mills) in1984 and a HDR (Decision support system for production systems management) in 1991 at the University of Valenciennes. He is currently a full professor (France) in the LAMIH - CNRS UMR 8201 laboratory at the University of Valenciennes. He is involved in many research and scientific associations and is a referee for scientific journals. His teaching and research interests focus on the areas of design of decision support systems, design of control systems, design and management of complex services systems.

Ying Sun is an Assistant Professor of Statistics in the division of Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) at King Abdullah University of Science and Technology (KAUST) in Saudi Arabia. She joined KAUST in June 2014 after one-year service as an assistant professor in the Department of Statistics at the Ohio State University, USA. At KAUST, she leads a multidisciplinary research group on environmental statistics, dedicated to developing statistical models and methods for space-time data to solve important environmental problems. Prof. Sun received her Ph.D. degree in Statistics from Texas A&M University in 2011, and was a postdoctorate researcher in the research network of Statistics in the Atmospheric and Oceanic Sciences (STATMOS), affiliated with the University of Chicago and the Statistical and Applied Mathematical Sciences Institute (SAMSI). She demonstrated excellence in research and teaching, published research papers in top statistical journals as well as subject matter journals, won multiple best paper awards from the American Statistical Association and the Transportation Research Board National Academies. Her research interests include spatio-temporal statistics with environmental applications, computational methods for large datasets, uncertainty quantification and visualization, functional data analysis, robust statistics, statistics of extremes.