Int. J. Production Economics 141 (2013) 377–387

Contents lists available at SciVerse ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Minimal Euclidean distance chart based on support vector regression for monitoring mean shifts of auto-correlated processes

Shichang Du [a,*], Jun Lv [b]

[a] Department of Industrial Engineering and Logistics Engineering, School of Mechanical Engineering, Shanghai Jiaotong University, 800 Dongchuan Road, Shanghai 200240, PR China

[b] School of Business, East China Normal University, Shanghai 200241, PR China

ARTICLE INFO

Article history:
Received 15 September 2011
Accepted 28 August 2012
Available online 15 September 2012

Keywords:
Minimal Euclidean distance
Support vector regression
Control chart
Statistical process control

ABSTRACT

Though traditional control charts have been widely used as effective tools in statistical process control (SPC), they are not applicable in many industrial applications where the process variables are highly auto-correlated. In this study, one new minimal Euclidean distance (MED) based monitoring approach is proposed for enhancing the monitoring mean shifts of auto-correlated processes. Support vector regression (SVR) is used to predict the values of a variable in time series. Through calculating minimal Euclidean distance (MED) values over time series, a novel MED chart is developed for monitoring mean shifts, and it can provide a comprehensive and quantitative assessment for the current process state. The performance of the proposed MED control chart is evaluated based on average run length (ARL). Simulation experiments are conducted and one industrial case is illustrated to validate the effectiveness of the developed MED control chart. The analysis results indicate that the developed MED control chart is more effective than other control charts for small process mean shifts in auto-correlated processes, and it can be used as a promising tool for SPC.

1. Introduction

The ability to monitor and reduce process variation for cost reduction and quality improvement in industrial processes plays a critical role in the success of one enterprise in today's globally competitive marketplace (Montgomery, 2001; Wu et al., 2007; Du et al., 2008; Lee et al., 2012). Control charts have been widely used as effective tools in statistical process control (SPC) for monitoring the process variation in industrial applications. In particular, control charts for monitoring independent observations have been extensively investigated and applied (Alt, 1984; Mason et al., 1995; Aparisi and Haro, 2003; Yang and Rahim, 2005; Torng et al., 2009; Magalhãesa et al., 2009; Wu et al., 2009; Du and Xi, 2010; Ou et al., 2011; Costa and Machado, 2011). With tremendous growth of advanced automatic data inspection and measurement techniques during the past few years, the process variables are being collected automatically at higher rates, therefore, for many industrial applications, a basic statistical assumption of independence is often violated, i.e. the data collected at regular time intervals from the processes is serially auto-correlated (Montgomery and Friedman, 1989; Cook and Chiu, 1998).

Several attempts have been made to extend traditional SPC techniques to deal with auto-correlated processes. One of the most interesting approaches to SPC for auto-correlated processes was proposed by Alwan and Roberts (1988). They introduced two charts, which they referred to as the common-cause control chart (CCC) and the special-cause control chart (SCC). CCC is a plot of forecasted values that are determined by fitting the correlated process with an autoregressive moving average model (ARMA), and SCC is a traditional Shewhart chart of the residuals. Their work attracted further investigation on time series modeling techniques application for monitoring correlated processes (Montgomery and Friedman, 1989; Montgomery and Mastrangelo, 1991; Wardell et al., 1992, 1994; Schmid, 1997; Adams and Tseng, 1998; Timmer et al.,1998; Jiang et al., 2000; Wright et al., 2001; Orlando et al., 2002; Kalgonda and Kulkarni, 2004). The time series based control charts approaches essentially involve fitting an adequate time series model to the correlated process data and applying a traditional control chart to the stream of residuals from the time series model. All these control chart approaches have been shown to improve the monitoring performance in the presence of auto-correlation. However, these time series modeling techniques require that a strict model has been identified for the time series of process observations before residuals can be obtained (Hwarng, 2004), and their performance is not very good for monitoring small shifts (Wardell et al., 1994), and they require one to have some skill in time series analysis (Box et al., 1994). Therefore, some other control charts based on residual have been developed for enhancing the performance of monitoring correlated processes (Testik, 2005; Pan and Jarrett, 2007).

* Corresponding author. Tel.: +86 137 644 79720; fax: +86 21 644 80001.
E-mail address: lovbin@sjtu.edu.cn (S. Du).

0925-5273/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijpe.2012.09.002

Recently, some researchers tried to find alternative methods that allow less restrictive assumptions, more flexibility and adaptability to real data situations. Examples of such techniques are machine learning methods such as neural network (NN) and support vector machine (SVM). These techniques allow learning the specific structure directly from the data and can be applied without forcing any assumptions. Some authors have proposed NN approaches as effective tools for monitoring auto-correlated processes (Cook and Chiu, 1998; Cook et al., 2001; Zobel et al., 2004; Hwarng, 2005; Pacella and Semeraro, 2007; Jamal et al., 2007; Du and Xi, 2011).

Support vector machine (SVM) has recently become a new generation learning system based on recent advances on statistical learning theory for solving a variety of learning, classification and prediction problems (Cortes and Vapnik, 1995; Gunn, 1998; Cristianini and Shawe-Taylor, 2000; Deng and Yeh, 2011). SVMs calculate a separating hyperplane that maximizes the margin between data classes to produce good generalization abilities. The main difference between NNs and SVMs is in their risk minimization (Gunn, 1998). In case of SVMs, structural risk minimization principle is used to minimize an upper bound based on an expected risk, whereas in NNs, traditional empirical risk minimization is used to minimize the error in the training of data. The difference in risk minimization leads to a better generalization performance for SVMs than NNs (Gunn, 1998). Support vector regression (SVR) is an important extension of SVM and is a regression method by introduction of an alternative loss function (Vapnik, 1998).

The applications of SVM to monitor the process variation are spare. Chinnam (2002) demonstrated that SVM can be extremely effective in minimizing both type I and type II errors for detecting shifts in the auto-correlated processes, and performed as well or better than traditional Shewhart control charts and other machine learning methods. Sun and Tsung (2003) and Kumar et al. (2006) developed one kernel-distance based K-chart using support vector for monitoring the independent observations. Issam and Mohamed (2008) presented SVR based cumulative sum (CUSUM) control chart for auto-correlated process. In their approach, SVR is firstly calculated and then CUSUM is applied to the stream of residuals from SVR. Therefore, two methods including SVR and CUSUM need to be calculated in their approach.

In this paper, one new minimal Euclidean distance (MED) based control chart is developed as a promising tool for monitoring auto-correlated processes. SVR is used to predict the values of a variable in time series. By using MED, the quantization error is provided for quantifying the deviation degree of current process with in-control process state space. Depending on how far away the current process is deviating from the in-control process state, a quantitative assessment index can be obtained by calculation of the MED of the new measurement data to the SVR trained by in-control process datasets.

The rest of this paper is organized as follows. The SVR theory is reviewed briefly in Section 2. The methodology using the proposed MED chart for monitoring auto-correlated processes is developed in Section 3. Experiments and performance analysis of the MED chart are conducted based on average run length (ARL) in Section 4. Further analysis of MED chart and one industrial application is illustrated to validate the effectiveness of the MED chart in Sections 5 and 6, respectively. A procedure for applying the MED chart into real processes is presented in Section 7. Finally, the conclusions and future work are given in Section 8.

2. Brief review of SVR

This section provides a brief review of SVR formulation. Consider a training data set $\{(\mathbf{x}_1,y_1),(\mathbf{x}_2,y_2),\ldots,(\mathbf{x}_N,y_N)\}$ $i=1,2,\ldots,N$, where N is the total number of training vectors, $\mathbf{x}_i \in R^d \subset R$ is the ith d-dimensional input vector, and $y_i \in \{1,-1\}$ is known target. The standard form of support vector regression (Vapnik, 1998) is:

$$\min_{\mathbf{w},b,\xi,\xi*} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{i=1}^{N}\xi_i+C\sum_{i=1}^{N}\xi_i^* \tag{1}$$

$$\text{subject to } \begin{cases} \mathbf{w}^T\phi(x_i)+b-y_i \le \varepsilon+\xi_i \\ y_i-\mathbf{w}^T\phi(x_i)-b \le \varepsilon+\xi_i^* \\ \xi_i,\xi_i^* \ge 0 \end{cases} \tag{2}$$

where $\mathbf{w}$ is the vector of hyperplane coefficients, defining a direction perpendicular to the hyperplane, the index i labels the N training cases, ξ_i and ξ_i^* are slack variables, measuring the degree of misclassification of the sample $\mathbf{x}_i$, C is the error penalty factor, penalizing the non-zero ξ_i, the bias b is a scalar, representing the bias of the hyperplane, the parameter ε is the ε-insensitive loss, and the map function ϕ is a non-linear transformation to map the input vectors into a high-dimensional feature space.

Eq. (2) corresponds to dealing with a called ε-insensitive loss function $|\xi|_\varepsilon$, which is one of most important loss functions (Vapnik, 1998). After transforming the problem to the dual form using Lagrangian transformation and applying optimality conditions, the following optimization problem is obtained:

$$\min_{\alpha,\alpha*} \ \frac{1}{2}(\alpha-\alpha^*)^T Q(\alpha-\alpha^*)+\varepsilon\sum_{i=1}^{N}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{N}y_i(\alpha_i-\alpha_i^*) \tag{3}$$

$$\text{subject to } \sum_{i=1}^{N}(\alpha_i-\alpha_i^*)=0, \quad 0\le\alpha_i,\ \alpha_i^*\le C \tag{4}$$

where $Q_{ij}=K(x_i,x_j)=\phi(x_i)^T\phi(x_j)$, and α_i and α_i^* are the Lagrange multipliers.

The estimated approximate function $\hat{f}$ is:

$$\hat{f}_{w,b}(x)=\sum_{i=1}^{N}(\hat{\alpha}_i-\hat{\alpha}_i^*)K(x_i,x)+b \tag{5}$$

where $K(x_i,x)$ is kernel function.

Any function that satisfies Mercer's theorem (Cristianini and Shawe-Taylor, 2000) can be used as a kernel function. Some popular SVM kernel functions include:

$$\text{linear function}: \quad K(\mathbf{x}_i,\mathbf{y}_j)=\mathbf{x}_i\times\mathbf{y}_j \tag{6}$$

Gaussian radial basis function (GRBF):

$$K(\mathbf{x}_i,\ \mathbf{y}_j)=\exp(-\gamma||\mathbf{x}_i-\mathbf{y}_j||^2) \tag{7}$$

polynomial function with degree d:

$$K(\mathbf{x}_i,\mathbf{y}_j)=((\mathbf{x}_i\times\mathbf{y}_j)+\tau)^d \quad \tau>0 \tag{8}$$

3. Methodology

Among various out-of-control conditions, this study is concerned with process mean shifts, which are defined as unanticipated sudden shifts in process mean vector. The primary possible causes for the mean shifts result from the introduction of new workers, machines or methods, a change in the measurement method or standard, etc.

3.1. MED chart based on SVR

In general, the in-control operation datasets are relatively easier to acquire, but it is hard to obtain lots of out-of-control datasets. Out-of-control state detection can be implemented

based on the residual between the in-control state and current state. Large residual indicates that the current process is out of the in-control operation state.

Muller et al. (1997) introduced SVR in time series prediction and found that it provides excellent results comparing to NN. Thissena et al. (2003) also showed that for complex time series structures SVR outperforms ARMA and NN models. Therefore, this study uses SVR to predict the value of a variable, and then the residual is calculated for quantifying the deviation degree of the current process with the in-control operation state. At first, SVR is trained with the in-control process data, and then the residual vector is obtained by calculating the difference between the predicted values and the observed values of the current process variable. For in-control operation state, the mean vector of the residuals is zero, whereas, for out-of-control state, a shift in the mean vector is reflected in the mean of the residual vector. If the difference between the input vector and the in-control vector is larger than a predetermined control limit, the current process is probably out-of-control.

According to the distance between the current process state and the in-control process state, a quantitative assessment index can be obtained by calculating the MED of the measurement data of current state to the SVR trained by training data sets from in-control process. The MED value is calculated by equation:

$$\mathrm{MED} = ||M - \mathrm{SVR}_{in\text{-}control}|| \tag{9}$$

where $||\bullet||$ presents minimal Euclidean distance, M means the data within the moving window (i.e., the moving window vector), and $\mathrm{SVR}_{in\text{-}control}$ is the predicted value of SVR trained by the data sets from in-control state.

From the out-of-control monitoring of view, the distance between the $\mathrm{SVR}_{in\text{-}control}$ and the input vector actually represents how far away the input vector deviates from the in-control process state. An extremely high MED value, i.e., exceeding the control limit value, means that the input vector belongs to an out-of-control class. Therefore, the process changes can be quantified and visualized by following the trends of MED. These MED values over time series form an MED chart that is used for monitoring processes.

3.2. Selection of kernel function

The prediction performance of SVR varies depending on the choice of the kernel function and its parameters. Therefore, one suitable kernel function should be first selected for implementing SVR. The choice of kernel functions is highly problem-dependent and it is the important factor in SVR applications. There are usually two non-linear kernel functions for nonlinear SVM including the GRBF and the polynomial function (see Eqs. (7) and (8)).

To take advantage of merits of two non-linear kernel functions, one hybrid kernel function $K_{hybrid}(x_i,x_j)$ combining GRBF and polynomial kernel in a generalized kernel form is presented in this study.

$$K_{hybrid}(x_i,x_j) = \beta \times \exp(-\gamma||\mathbf{x}_i - \mathbf{y}_j||^2) + (1-\beta)((\mathbf{x}_i \times \mathbf{y}_j)+1)^d \tag{10}$$

The parameter γ and d represents kernel width and degree, and β is coefficient. Since K_{hybrid} still satisfy Mercer's condition, it can be used as a kernel function for SVR. For instance, one case of the kernel of degree 2, width 0.5, and β 0.5 is presented:

$$K_{hybrid} = 0.5 \times \exp(-2||x_1 - x_1' + x_2 - x_2'||^2) + 0.5(1 + 2x_1x_1' + 2x_2x_2' + 2x_1x_2x_1'x_2' + (x_1x_1')^2 + (x_2x_2')^2)$$

3.3. Optimization of kernel function parameters

For $K_{hybrid}(x_i,x_j)$ kernel, three parameters, including penalty parameter C, kernel width γ and degree d, need to be optimized in order to obtain the best prediction result. It is widely reported that the particle swarm optimization (PSO) algorithm is very easy to implement and has fewer parameters to adjust when compared to other evolutionary algorithms (Eberhart and Kennedy, 1995). In this study, PSO technique is used to obtain an optimal subset of parameters.

It is an iterative process in which the change in weights of a particle at the beginning of an iteration are calculated using Eq. (11) and new position of every particle is found using Eq. (12).

$$v_i(j+1) = w_p v_i(j) + c_1 r_1(p_{id} - x_i(j)) + c_2 r_2(p_{gd} - x_i(j)) \tag{11}$$

$$x_i(j+1) = x_i(j) + v_i(j+1) \tag{12}$$

where j is the current step number, w_p is the inertia weight, c_1 and c_2 are the acceleration constants, r_1 and r_2 are two random numbers in the range [0,1], $x_i(j)$ is the current position of the particle, p_{id} is the best one of the solutions this particle has reached, p_{gd} is the best one of the solutions all the particles have reached.

3.4. Data presentation of auto-correlated process

How to represent the data is very critical for the MED chart. Auto-regressive processes of lag 1 (i.e. AR(1)) are very popular in many processes. Auto-regression is a form of regression as the value of a particular variable is dependent on the part variable of itself at varying time lags. An AR(1) to model process mean shifts can be represented by:

$$X(t) = \mu + \Phi \times (X(t-1) - \mu) + \upsilon(t) \tag{13}$$

where $X(t)$ and $X(t-1)$ are the values of time series at time t and $t-1$, respectively, μ is the mean of the data series, Φ is the auto-regressive coefficient ranged within $[-1,1]$, and $\upsilon(t)$ is a normal independently distributed error.

If the process mean is shifted δ at time point t, the process can be represented by

$$X(t) = \delta + X(t-1) \tag{14}$$

The time when a step shift occurs is called the point of shift. Thus, δ is the expected difference between the shifted value of $X(t)$ and non-shifted value of $X(t-1)$. The model represented in Eqs. (13) and (14) are consistent with the AR(1) model considered in Zhang (1998), the ARMA (1,1) model with $\theta_1 = 0$ in Wardell et al. (1992, 1994) and Jiang et al. (2000).

3.5. Generation of data sets

As for training and testing datasets of the MED chart, a moving window method with n observation points over a time series stream is used for data presentation. The window size greatly affects the monitoring performance of MED chart. A small window size is able to detect out-of-control signals more quickly, but might result in a short in-control ARL_0, which is equivalent to a high type I error. A large window size is able to potentially provide more required data for identification of out-of-control signals, but reduce the detection efficiency by increasing the time required to detect these signals (i.e. longer out-of-control ARL_1 or high type II error). Therefore, a compromise size to balance type I and type II errors needs to be found. Preliminary investigations have been conducted to choose a suitable size of moving window (discussed in Section 4.2.2). A window size of 16 was selected in

this study since there would be no significant improvement beyond these values during training stage.

Only in-control moving window vectors are used for training the SVR. Once an SVR (i.e. $SVR_{in\text{-}control}$) is created and trained, the measurement window vectors are fed to the $SVR_{in\text{-}control}$ to output the MED values. According to these *MDE* values, the state of current processes is observed and determined. The process starts in in-control state, and then is shifted δ away from the current process mean. Using Eq. (13), the test moving window vectors for each shift were generated. After the first shifted point enters into the moving window, the MED value will increase gradually as more shifted points enter into the moving window. These MED values of all the windows are plotted over time series stream, which form the MED chart. This datasets generation approach is considered as being more practical because in the real-world problems out-of-control signal often occurs after a period in which the process is in-control and the starting point of the out-of-control signal is usually unknown.

Shift signals are added to the auto-correlated process model to provide out-of-control signals (see Eq. (14)). Process shifts $\delta=0.0,0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0$ were considered (when δ equals 0, the process is in control). Each δ value is used in turn during the simulation of the out-of-control status (when δ equals 0, the process is in control) and has the same chance to be picked up. Without loss of generality, suppose that in-control process variable has mean μ zero and variance σ^2 one. The same number of moving window vectors was generated for each parameter Φ.

3.6. Evaluating normality

The assumption of normality for quality characteristic variable usually needs to be checked. Plots are always useful tools in data processing. For univariate process, special plots called *Q–Q* plots can be used to assess the assumption of normality. When the points lie very nearly along a straight line, the normality assumption remains tenable. Normality is suspect if the points deviate from straight line.

In practice, the normality of data set may not be a viable assumption. One important way to handle this situation is to transform the non-normal data to near normality. Transformation is a re-expression of the data set in different units. For instance, when a histogram of positive observations exhibits a long right hand tail, transforming the observations by taking their logarithms or square roots will often markedly improve the symmetry about the mean and the approximation to a normal distribution.

Different methods for transformations have been developed, which can be divided into several classes according to transformed scale, such as "square roots transformation", "logit transformation", "Fisher's z-transformation", and "power transformation". Among these transformations, the "power transformation" is a very useful transformation method. Box and Cox (1964) modified the transformation method,

$$x^{(\lambda)}=\begin{cases}\frac{x^{\lambda}-1}{\lambda} & \lambda\neq 0\\ \ln x & \lambda=0\end{cases} \tag{15}$$

which is continuous in λ for $x>0$.

Given the observations $x_1,x_2,\ldots,x_n$, the solution for the choice of an appropriate power λ is the solution that maximized the expression

$$\ell(\lambda)=-\frac{n}{2}\ln\left[\frac{1}{n}\sum_{j=1}^{n}(x_j^{(\lambda)}-\bar{x}^{(\lambda)})^2\right]+(\lambda-1)\sum_{j=1}^{n}\ln x_j \tag{16}$$

$x_j^{(\lambda)}$ is defined in (4) and

$$\bar{x}^{(\lambda)}=\frac{1}{n}\sum_{j=1}^{n}x_j^{(\lambda)}=\frac{1}{n}\sum_{j=1}^{n}\left(\frac{x_j^{\lambda}-1}{\lambda}\right) \tag{17}$$

is the arithmetic average of the transformed observations. The first term in (16) is, apart from a constant, the logarithm of a normal likelihood function, after maximizing it with respect to the population mean and variance parameters.

The calculation of $\ell(\lambda)$ for many values of λ is an easy task for computer. It is helpful to have a graph of $\ell(\lambda)$ versus λ, as well as a tabular display of the pairs $(\lambda,\ell(\lambda))$, in order to study the behavior near the maximizing value $\hat{\lambda}$. For instance, if either $\lambda=0$(logarithm) or $\lambda=0.5$(square root) is near $\hat{\lambda}$, one of these may be preferred because of its simplicity.

Through using this procedure, the non-normal data set can be transformed to near normality.

3.7. Control limit for MED chart

After the measurement window vector of current process is fed to the $SVR_{in\text{-}control}$ trained on in-control datasets, an MED value is obtained that represents the deviation degree of current process with the in-control process state. When a shift δ occurs in a process, the MED chart would identify that there is a shift as soon as possible, which can be realized by computing the MED value to determine the moving window at which a shift is indicated. Therefore, a control limit γ_0 should be provided to check whether the MED value is larger than this control limit. Namely, when an MED value exceeds the control limit γ_0, the process could be in out-of-control state. Therefore, for each performance evaluation, the system would read the moving window one by one until the control limit γ_0 is reached. For instance, if the control limit γ_0 is exceeded at moving window 32, it is record that the out-of-control ARL_1 is 15 (i.e., $32-17=15$, where 17 means the previous in-control 16 window vectors). Therefore, the proposed approach provides a convenient way to calculate the ARL. The setting of the control limit γ_0 is based on consideration of the balance between type I and type II errors.

4. Simulation and performance analysis

Some simulation experiments have been done to verify the effectiveness of the proposed MED control chart. The influences of the key factors of the MED chart on monitoring performance are analyzed. The analysis can help us to obtain the suitable parameter setting to improve the performance of the MED control chart. In each of the following tests, only one factor is varied while the remaining ones are kept constant. Average run length (ARL) is one widely accepted evaluation index to evaluate the efficacy of monitoring and classifying out-of-control processes, which is also used to evaluate the performance of MED charts in this study. ARL includes two indices, in-control ARL_0 and out-of-control ARL_1.

4.1. Parameters of SVR

The performance of learning in SVR is influenced by its relative parameters. The relative parameters of SVR having better performances during training in this study are summarized as follows:

(1) *Loss functi*on: ε-insensitive loss function is used.
(2) *Kernel function*: $K_{hybrid}(x_i,x_j)$ function in Eq. (10) is used, and $\beta=0.5$.
(3) *The inertia weights*: minimum weight w_{min} is 0.2 and maximum weight w_{max} is 0.8.

(4) *The speed of particles*: v_{max} equals 4 and v_{min} equals -4.
(5) *The initial velocities*: the initial velocities of the initial particles are generated at random in the range $[-4,4]$.
(6) *The study factors*: c_1 equals 2 and c_2 equals 2.

4.2. Performance analysis

4.2.1. Performance analysis to the number of training examples

The number of training datasets influences the performance of a monitoring system significantly. An SVR must be trained with sufficient training examples in order to obtain good performance. Insufficient training data sets might not be able to represent the whole input–output behavior.

The datasets of different sizes with $\Phi=0.475$ and $\gamma_0=8.4598$ are generated to test the MED chart. The performance of the MED chart is evaluated as the number of training examples is increased incrementally. Table 1 presents the test results of the MED chart trained by different examples. From these results, some conclusions can be drawn: (1) one can find that increasing the training examples can improve the performance of the MED chart up to a certain level of accuracy (performance limits). This could be explained by the fact that with enough large training sets there is a better chance of true representation of a problem space. However, any further increase of the training size after reaching such limits will not greatly improve the performance of the MED chart. Moreover, the larger training set results in higher time cost of training. Table 1 shows that no discernible improvement in performance can be obtained when using more than 500 training examples. Thus, training set comprising 500 examples was used for training SVR. (2) Out-of-control signals with small magnitudes require more representation of the density distribution by larger training data sets. (3) Out-of-control signals with large shift magnitudes require smaller training examples as they are easier to identify. This is due to the fact that out-of-control signals with large shift magnitudes have stronger pattern features that separate easier itself with other random in-control patterns.

4.2.2. Performance analysis to different window sizes

Table 2 shows the out-of-control ARLs for different sizes ranged between 8 and 20 observations. The results show that increasing the window size results in improved performance of the MED chart for detecting small shift magnitudes $\delta<2.0\sigma$. In contrast, increasing the window size results in decreased performance of the MED chart for detecting large shift magnitudes $\delta>2.0\sigma$. For example, the MED chart with window size 12 has better performance than that of the MED chart with window size 8 when $\delta=0.5\sigma$, 1.0σ, 2.0σ, and poorer performance when $\delta=3.0\sigma$. The reasons why the size of window results in different performance for detecting small and large magnitude shifts, are following: the large window can accommodate more shifted points with low shift magnitudes, so that more out-of-control pattern features can be exposed to the MED charts and the MED value of the out-of-control window vector is easier to exceed the control limit γ_0. Therefore, for detecting small shifts, the MED chart with large window has better performance in comparison with the MED chart with small window. Moreover, an out-of-control signal with large magnitude shift can be detected before shifted observation points occupy the whole moving window. Thus, for detecting large shifts, the moving window with small size decreases the percentage of the number of random elements in a window vector, in contrast, increases the percentage of the number of out-of-control elements in the window vector. Thus, this makes the out-of-control window vector be distinctively different from the random in-control window vector. Therefore, the window size is one of the key factors that influents significantly the performance of the MED chart, and should be set to balance the out-of-control ARLs of small shifts and large shifts. Fig. 2 also shows that the window size of 16 has the best performance comparing to other window sizes. Thus, the window size of 16 was used for training SVR.

Table 1
Performance analysis of the MED chart to the number of training examples.

	Training examples				
δ	100	300	500	700	900
0.5	73.62	71.91	67.26	67.20	67.24
1.0	27.83	23.66	20.38	20.40	20.35
2.0	13.64	11.43	6.04	6.02	6.01
3.0	5.68	5.69	5.65	5.64	5.61

Table 2
Performance analysis of the MED chart to the window sizes.

	Window size			
δ	8	12	16	20
0.5	70.58	70.25	67.26	67.20
1.0	21.82	21.17	20.38	20.40
2.0	10.89	9.84	6.04	9.87
3.0	5.66	5.68	5.65	5.69

4.2.3. Performance analysis to different kernel function and parameters

Fifteen runs (500 training examples for each run) with $\Phi=0.475$ and $\gamma_0=8.4598$ have been performed to find optimal parameters of K_{hybrid} kernel function. PSO was applied to find the best combination of the parameters. The related parameters C and γ were varied in the arbitrarily range [1,1000], and [0,1] so as to cover high and small regularization of the MED chart, and fat as well as thin kernels, respectively. The degree d was varied in the range [2,5] with integer value in order to span polynomials with low and high flexibility. The detailed results about the optimal values of the SVR parameters (i.e., the penalty parameter C, kernel width γ and degree d) are shown in Table 3. From Table 3, it is observed that the combination with $C=387.4$, $\gamma=0.0075$, and $d=3$ generally has the best performance in the thirteenth run.

The proposed PSO optimization approach successfully finds the global optimum just with 100 iterations. This result was repeated in multiple runs of the program. Values of the parameters C and γ in five different runs of the program with 150 iterations are presented in Fig. 1. In each different run of program, PSO first generates random values for C, γ and d, and then it searches for better values of them that produce better ARL. Usually, after 50 iterations, the approach converges to be best parameters values of the K_{hybrid} kernel function.

One experiment has been carried out to compare the performance of K_{hybrid} kernel function with the commonly used GRBF (see Eq. (7)) and polynomial function (see Eq. (8)). The same parameters (i.e. $C=387.4$, $\gamma=0.0075$, and $d=3$) having best performance in Table 3 are used to fairly compare the performance of three different kernel functions. Table 4 shows the out-of-control ARLs for different kernel function. Based on the simulations results, it is found that the MED chart with K_{hybrid} kernel function has better monitoring capability than those from the other two kernels.

4.3. Comparison with other control charts

4.3.1. Comparison with statistics-based control charts

Comparisons between the MED chart and some classic statistics-based control charts are shown in Table 5. The compared control

Table 3
The out-of-control ARL_1 for different parameters of K_{hybrid} kernel function.

Run	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
C	296.7	428.3	338.4	305.7	452.9	412.1	289.4	324.7	401.5	348.2	421.3	397.1	287.4	292.8	412.9
γ	0.0109	0.0093	0.0078	0.0097	0.0042	0.0097	0.0108	0.0103	0.0099	0.0089	0.0085	0.0108	0.0075	0.0084	0.0049
d	2	4	4	3	5	5	2	2	5	4	4	5	3	2	3
$ARL_1(\delta = 0.5)$	73.51	75.08	68.54	75.18	69.47	77.18	74.04	70.15	78.49	74.84	71.054	76.49	67.25	75.94	66.08
$ARL_1(\delta = 1.0)$	25.94	26.41	23.47	24.98	23.41	26.09	27.84	25.67	24.18	24.18	24.27	29.14	20.38	24.08	25.47
$ARL_1(\delta = 2.0)$	8.90	11.87	6.18	10.78	6.12	8.15	9.08	7.04	5.48	6.41	7.64	8.63	6.04	6.04	7.99
$ARL_1(\delta = 3.0)$	7.49	6.18	5.69	6.49	5.91	6.18	7.48	6.84	6.15	5.94	6.78	6.08	5.65	7.22	6.88

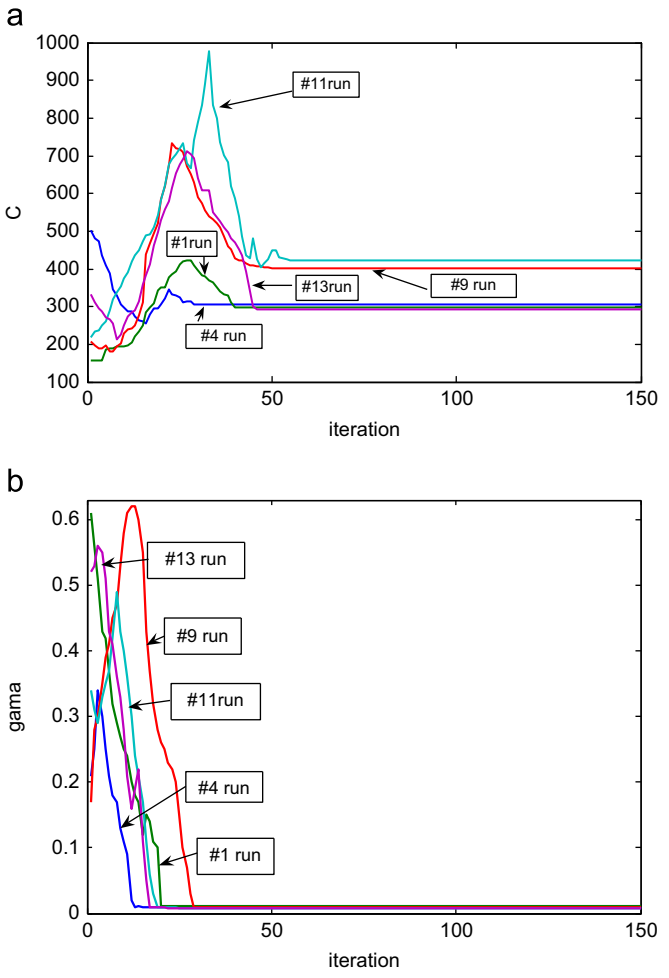

Fig. 1. The parameters C and γ of K_{hybrid} kernel function for five different runs. (a) The parameter C of K_{hybrid} kernel function and (b) The parameter γ of K_{hybrid} kernel function.

Table 4
Performance analysis of the MED chart to the kernel functions.

δ	Training examples		
	GRBF	Polynomial	Hybrid kernel
0.5	71.33	72.04	67.26
1.0	21.34	23.89	20.38
2.0	6.25	6.12	6.04
3.0	5.70	5.68	5.65

charts are SCC (Wardell et al., 1994), $\overline{X}$ (Wardell et al., 1994), exponentially weighted moving average (EWMA) (Wardell et al., 1994), exponentially weighted moving average charts for stationary processes (EWMAST) (Zhang, 1998) and ARMA (Jiang et al., 2000) commonly used as monitoring tools for auto-correlated process shifts. The results indicate that the MED chart outperform the other control charts for small shifts ($0 \le \delta \le 2$). The SCC and $\overline{X}$ charts perform better for large shifts because they use only a current single observation, whereas the MED chart uses 16 continual observation pointes for an observation vector to generate one MED value. For example, when the first shift point enters the moving window, the other 15 observation pointes are in an in-control state, and immediate deviation cannot be observed by the MED chart. Only when more shift pointes or larger shifts enter the moving windows can be MED chart give an out-of-control alarm.

Table 5
Comparisons between the MED chart and some statistics-based charts based on ARL.

Φ	γ_0	δ	MED	SCC[a]	$\bar{X}$[a]	EWMA[a]	EWMAST[b]	ARMA[c]
0.00	3.4587	0.00	371.26	370.38	370.40	369.00	370.40	N/A
		0.50	17.41	152.22	152.22	28.19	28.19	N/A
		1.00	10.23	43.89	43.89	9.73	9.73	N/A
		2.00	4.25	6.30	6.30	4.18	4.18	N/A
		3.00	2.73	2.00	2.00	2.76	2.76	N/A
		4.00	1.06	1.07	1.24	2.65	2.71	2.81
		5.00	1.00	1.00	1.08	1.17	1.68	1.57
		6.00	1.00	1.00	1.00	1.00	1.00	1.00
0.475	8.4598	0.00	375.87	370.38	365.34	376.53	373.66	370.00
		0.50	67.26	253.13	166.77	70.05	77.91	65.60
		1.00	20.38	117.96	51.05	20.69	22.45	20.30
		2.00	6.04	22.64	8.69	7.16	6.06	6.61
		3.00	5.65	4.02	2.50	4.28	3.35	3.67
		4.00	1.21	1.25	1.38	2.19	2.02	2.09
		5.00	1.11	1.09	1.26	1.07	1.99	1.66
		6.00	1.00	1.00	1.00	1.00	1.00	1.00
0.95	21.0157	0.00	371.39	370.38	369.15	365.16	368.31	370.00
		0.50	134.67	330.95	259.73	245.67	222.07	226.00
		1.00	23.64	138.84	118.92	107.83	105.12	102.00
		2.00	13.61	1.08	22.44	27.79	23.33	25.80
		3.00	2.90	1.00	1.43	10.01	7.41	8.65
		4.00	1.04	1.05	1.40	2.47	2.05	2.44
		5.00	1.00	1.00	1.27	1.64	1.90	1.08
		6.00	1.00	1.00	1.00	1.00	1.00	1.00

[a] Results ($0 \le \delta \le 3$) taken from Wardell et al. (1994).
[b] Results ($0 \le \delta \le 3$) taken from Zhang (1998).
[c] Results ($0 \le \delta \le 3$) taken from Jiang et al. (2000).

4.3.2. Comparison with NN-based control scheme and SVR-based CUSUM

For further comparison, the performances of the MED chart, the back-propagation network (BPN) based monitoring scheme (Hwarng, 2004), and SVR-based CUSUM (Issam and Mohamed, 2008) are shown in Table 6. Table 6 presents the best results of ARL and SRL (the standard deviation of ARL) of three control schemes. For comparison, the algorithm 1 in the work of Issam and Mohamed (2008) was applied to obtain the ARL results of SVR-based CUSUM chart, and the best results of SVR-based CUSUM chart are presented in Table 6. Each ARL was computed for different Φ and δ. The MED chart outperforms the SVR-based CUSUM chart (Issam and Mohamed, 2008) for small shifts ($0 \le \delta \le 3$). For example, when Φ equals 0.0, 0.25, 0.5, 0.75, 0.95 and δ equals 3.0, the ARL_1 of the MED chart and SVR-based CUSUM chart are 2.73 and 2.89, 4.02 and 4.35, 5.38 and 5.24, 5.99 and 6.01, and 2.90 and 3.00, respectively. For large mean shifts ($4 \le \delta \le 6$), the MED chart shows a performance similar to that of SVR-based CUSUM chart. For example, when Φ equals 0.0, 0.25, 0.5, 0.75, 0.95 and δ equals 4.0, the ARL_1 of the MED chart and SVR-based CUSUM chart are 1.06 and 1.04, 1.29 and 1.37, 2.16 and 2.38, 1.76 and 1.98, and 1.04 and 1.06, respectively. Moreover, all of the charts (in Tables 5 and 6) are capable of detecting the large mean shifts ($\delta = 5,6$) immediately (the ARL_1 are less than two).

5. Further analysis of the MED chart

In order to further analyze the performance of the MED chart, one representative MED chart for auto-correlated process is plotted. Fig. 2 is the MED chart of an auto-correlated process with auto-correlation $\Phi = 0.95$ and shift $\delta = 2.0\sigma$. In the MED chart, the window number means the moving window position over the time series stream and MED indicated the distance value between the input vector and the corresponding $SVR_{in\text{-}control}$. The starting shift occurs at window number 30. The red circles on the curved lines mean that the MED values of the moving window vectors exceed the control limit.

Table 6
Comparisons between the MED chart and NN-based scheme and SCC based on ARL.

Φ	γ_0	δ	MED		BPN[a]		SVR-based CUSUM[b]	
			ARL	SRL	ARL	SRL	ARL	SRL
0.00	3.4587	0.00	371.26	370.21	372.96	370.10	371.79	369.43
		0.50	17.41	28.91	25.38	17.93	20.33	48.02
		1.00	10.23	5.69	8.29	5.76	12.07	21.35
		2.00	4.25	2.09	2.47	1.51	4.26	2.14
		3.00	2.73	0.75	1.29	0.61	2.89	2.38
		4.00	1.06	0.51	1.10	0.98	1.04	1.02
		5.00	1.00	0.00	1.00	0.00	1.00	0.00
		6.00	1.00	0.00	1.00	0.00	1.00	0.00
0.25	5.9871	0.00	370.89	378.86	371.32	385.70	370.42	384.01
		0.50	35.12	38.60	32.46	24.96	37.09	39.45
		1.00	13.28	13.09	11.87	9.03	14.05	15.98
		2.00	4.24	2.06	3.39	2.19	5.64	2.46
		3.00	4.02	1.49	1.63	0.99	4.35	1.98
		4.00	1.29	1.28	1.60	0.84	1.37	1.41
		5.00	1.00	0.00	1.17	0.65	1.03	1.00
		6.00	1.00	0.00	1.00	0.00	1.00	0.00
0.50	9.0125	0.00	373.52	372.88	371.35	373.57	377.38	378.41
		0.50	53.21	43.09	52.07	45.74	57.05	45.17
		1.00	15.09	15.64	16.74	12.88	16.29	18.33
		2.00	6.69	3.19	4.84	3.26	6.22	6.81
		3.00	5.38	1.96	2.22	1.52	5.24	4.26
		4.00	2.16	1.88	2.89	1.48	2.38	1.98
		5.00	1.05	1.04	1.05	0.59	1.03	0.00
		6.00	1.00	0.00	1.00	0.00	1.00	0.00
0.75	12.1134	0.00	371.61	370.14	370.60	368.36	371.88	370.45
		0.50	90.98	85.39	91.72	94.81	95.32	94.89
		1.00	18.09	31.27	35.42	32.71	21.87	20.48
		2.00	9.76	4.09	8.95	9.23	9.84	9.78
		3.00	5.99	3.19	3.52	3.28	6.01	6.23
		4.00	1.76	2.08	2.04	2.14	1.98	2.10
		5.00	1.00	0.00	1.12	0.87	1.02	1.70
		6.00	1.00	0.00	1.00	0.00	1.00	0.00
0.95	21.0157	0.00	371.39	374.01	370.37	374.11	371.38	369.45
		0.50	134.67	131.45	152.09	148.81	145.42	151.61
		1.00	23.64	60.38	77.00	69.16	21.04	36.31
		2.00	13.61	21.16	32.07	27.57	15.21	7.21
		3.00	2.90	1.24	10.17	1.00	3.00	3.25
		4.00	1.04	0.65	2.54	2.87	1.06	1.88
		5.00	1.00	0.00	1.08	0.84	1.00	0.00
		6.00	1.00	0.00	1.00	0.00	1.00	0.00

[a] Results ($0 \le \delta \le 3$) taken from Wardell et al. (1994).
[b] Calculation according to algorithm 1 taken from Issam and Mohamed (2008).

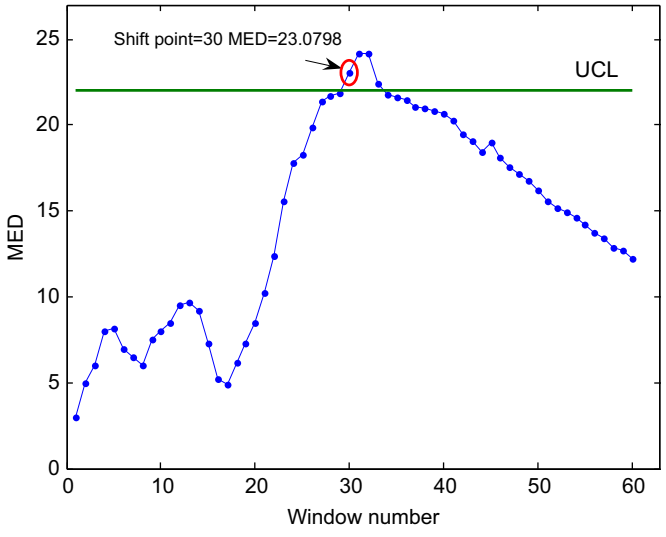

Fig. 2. MED chart of an auto-correlated process with shift $\delta = 2.0\sigma$ and $\Phi = 0.95$.

Some important features can be observed from the MED chart. Firstly, three obvious states are shown in the MED chart: (1) in-control state (the first 17 windows number); (2) fast upward trend state (when the first shift point enters the 18th moving window number); (3) deviation state (when the first shift point exceeds the control limit). Secondly, an upward trend can be observed when a shift point enters the moving window at window number 18, and then the MED values increases gradually as more shift points enter. Thirdly, the MED values of the third state are obviously larger than those of the first state, which indicates that the deviation of the out-of-control state is completely presented. Fourthly, according to the feature of the fast upward trend state, the starting points of detected shift could be observed. The starting point of shift is usually relevant to the time at which the out-of-control condition occurs, and so can be an important clue used by quality practitioners to identify the root causes. These important features provide important process information such as three obvious process sates, MED values and starting points, etc., which is very useful information for quality practitioners to monitor the process and identify the root causes.

6. Case study

To further validate the effectiveness of the developed MED chart, its real application in an aircraft horizontal stabilizer assembly process for monitoring and identification of process mean shifts is illustrated. Most components of an aircraft have a very high demand of high quality, reliability, and low variation. Aircraft horizontal stabilizer is one of these most important components and has very strict dimensional specification. The assembly processes can directly affect the reliability and performance of the final aircraft. However, the complexity of aircraft horizontal stabilizer assembly processes makes it very difficult to efficiently detect the mean shifts depend solely on the operator's experience. The immediate detecting the out-of-control conditions in aircraft horizontal stabilizer assembly processes can greatly narrow down the set of the possible root causes, facilitating rapid analysis and corrective action by quality engineers.

Fig. 3 shows the computer-aided design (CAD) model of aircraft horizontal stabilizer, which is mainly composed of two beams (front beam and back beam), two rear spars (front rear spar and back rear spar), and several webs. There are a total of sixteen measurement points (eight measurement points on front beam and back beam, respectively) numbered in the CAD model. Among these sixteen measurement points, two measurement points are most important (the fourth measure point on front beam and back beam) and the characteristics are called key product characteristics (KPCs), which are the center balance points and directly affect the final dimensional quality and product reliability of aircraft horizontal stabilizer. These two measurement points are observed and monitored based on the developed MED control chart. Since the front beam and back beam are assembled independently, the two measurement points of KPCs on front beam and back beam are monitored independently.

Fig. 4 is the fixture layout in horizontal stabilizer assembly process. Firstly, the rear spars are loaded and located into fixture

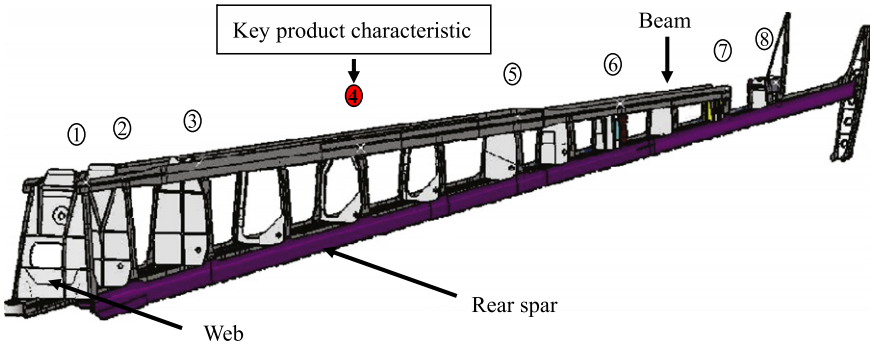

Fig. 3. The CAD model of aircraft horizontal stabilizer and key product characteristic.

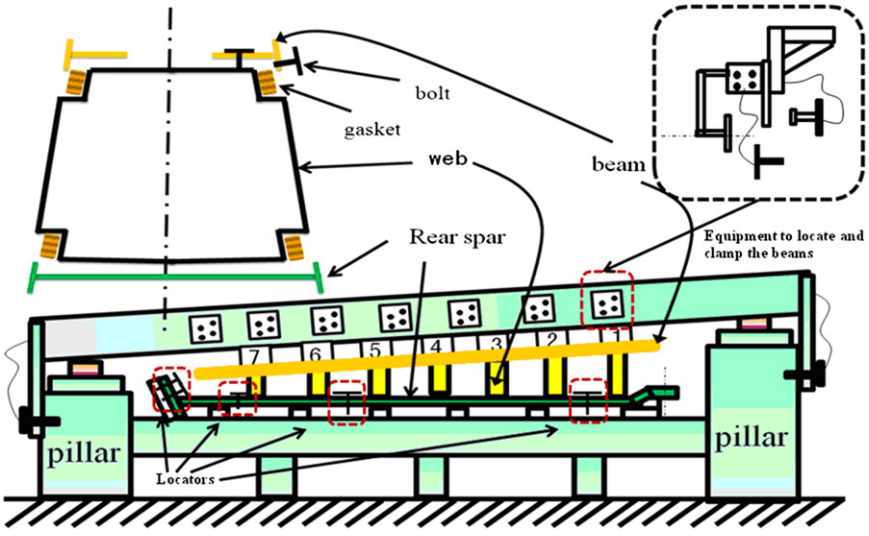

Fig. 4. Fixture scheme in aircraft horizontal stabilizer assembly process.

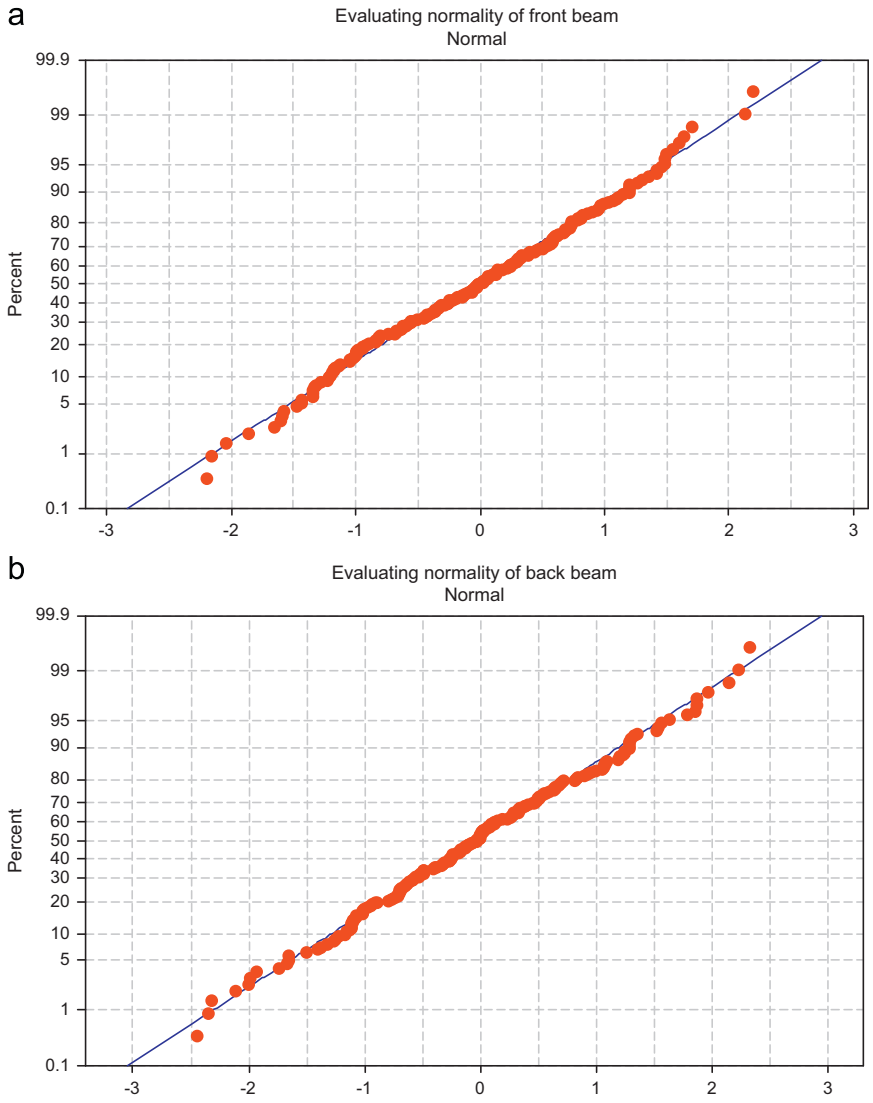

Fig. 5. Evaluating normality of the dataset from the horizontal stabilizer assembly process. (a) Evaluating normality of front beam and (b) Evaluating normality of back beam.

system, and is completed by assembling two rear spars into one rear spar component. Then the webs are loaded, located, clamped and assembled into the rear spar component. Finally, two beams are located and assembled. The two KPCs are formed on the aircraft horizontal stabilizer.

The 185 sample points from real-world assembly processes have been collected. The fitted time series model for the data series is an AR(1) model with a positive auto-correlation:

$$X(t) = 0.6728 \times X(t-1) + \upsilon_t \quad \upsilon_t \sim N(0, 0.09413)$$

The generated data set in simulation experiments of Section 4 follows the normal distribution through the computer programs, however, the normality of the collected data set from real-world processes needs to be checked. Fig. 5 presents the normality evaluation results of the data set from horizontal stabilizer assembly process. From these plots, it is shown that the data set quite follows normal distribution.

The control limit $\gamma_0 = 7.6905$ was used for this case. Fig. 6 presents the result of the MED charts for monitoring the beam and back beam KPCs. From Fig. 6 the shift was detected at window number 93 for front beam and at window number 127 for back beam, respectively. The reason why the process mean shifted is that the position of the fixture locator deviates from the nominal design due to the worn-out locators and the excessive looseness of the locator. Actually, the mean shift ($\delta_1 = 1.2140$) occurred at window number 81 for front beam and the mean shift ($\delta_1 = 1.3827$) occurred at window number 117 for back beam. Therefore, the out-of-control ARL_1 of the MED chart was 13 for front beam and 11 for back beam. This indicates that the developed MED chart exhibits a strong ability to monitor the mean shift in the real aircraft horizontal stabilizer assembly process.

7. A procedure for applying the MED chart into real processes

A general procedure for applying the MED charts in auto-correlated processes to monitor mean shifts is proposed. The parameters of SVR having better performance in Section 4.1 are used. In order to obtain a balance between type I and type II errors, it is needed to first decide the desired ARL_0, and then to create the required balance by adjustment of the control limit. This adjustment is addressed by Steps 3 and 4.

Step 1: collect time series measurements for the variable when the process is in control.

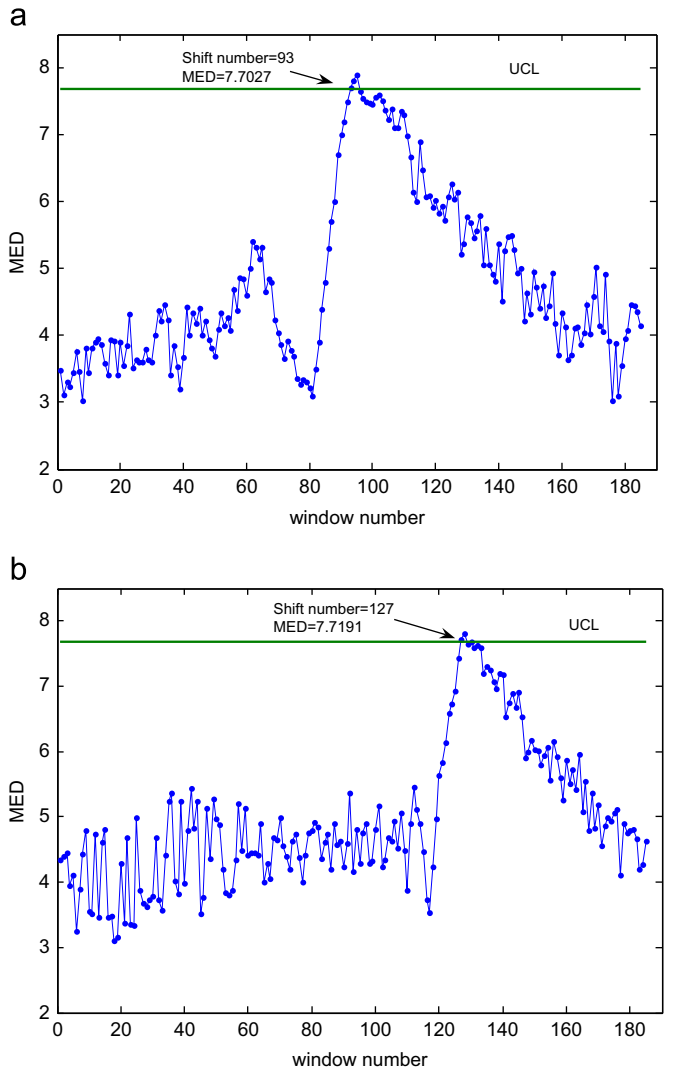

Fig. 6. Monitoring the horizontal stabilizer assembly process using the MED chart. (a) Monitoring the front beam using the MED chart and (b) Monitoring the back beam using the MED chart.

Step 2: judge the normality of the data set. If the data set does not follow the normal distribution, transforming data set to near normality according to the steps in Section 3.6.
Step 3: training SVR using in-control datasets with training size (400–600), window size (14–18) and 150 iterations.
Step 4: adjust the control limit of the MED chart by against test moving window vectors over time series to maintain the desired ARL_0.
Step 5: if ARL_1 is not achievable, one or more of the factors mentioned in Step 2 should be adjusted until the required ARL_0 is achieved and the balance is obtained. After completed the above four steps, the construction of an MED chart is completed.
Step 6: the monitoring begins with the first in-control 16 points for the process variable (when the size window is 16). This step is considered as being more practical since in the real world problems out-of-control signal often occurs after a period in which the process is in control and the starting point of the out-of-control signal is usually unknown.
Step 7: input the moving window vector to SVR, and then the output MED values are generated.
Step 8: if the MED value is beyond the control limit of the MED chart, an alarm will be triggered and go to step (8). Otherwise, the time point of sampling increases by 1, and return to step (5).
Step 9: calculate the out-of-control ARL_1, and take actions to correct the process.

8. Conclusions and future work

Monitoring the mean shifts in auto-correlated processes has been a challenging task for traditional SPC techniques. This study has proposed one new monitoring approach based on SVR for recognizing the mean shifts of auto-correlated processes. The monitoring approach is capable of providing a comprehensible and quantitative assessment value for current process state through calculating the MED values. Based on these MED values over time series, a novel MED control chart is developed to monitor the mean shifts of auto-correlated processes.

Some important details of the construction of the MED chart are discussed and analyzed using the simulation experiments. The simulation results indicate that the MED chart shows the improved performance, which outperforms those of some statistics-based charts and the NN-based control scheme for small process mean shifts in auto-correlated processes. Moreover, the influences of some key parameters of SVR upon its performance are analyzed, which aims to find the suitable parameters for constructing the MED chart. The effectiveness of the MED chart is further validated through the datasets from one case, and a general procedure for using the MED chart in the industrial applications is also proposed. Some merits of the proposed MED chart are concluded as follows: the MED chart possesses high performance for detecting the mean shifts of auto-correlated processes immediately based on ARL. The best time to identify a shift in a process is immediately after the shift has occurred since it becomes more difficult to identify shifts as more observations are taken. Early detection makes the root cause of the signal easier to identify, whereas causes of shifts that occurred in the distant past are difficult to identify. Increasing the probability of early detection would result in the fastest rate of continuous quality improvement. Furthermore, just like CUSUM chart, the MED chart also provides some important process information including quantitative assessment values, the starting shift points and the whole tendency state of process state, which are very important for quality engineers to identify the root causes as soon as possible.

There are several possible directions for future research. Firstly, further work can focus on finding out the better optimization algorithms to find an optimal subset of the parameters of SVR. Secondly, further investigation can be done to apply the proposed MED chart into monitoring the covariance changes, which is also important in SPC. Thirdly, in the future more industry cases need be collected in order to further validate and improve the MED chart.

Acknowledgments

The authors greatly acknowledge the editor and the referees for their valuable comments and suggestions that have led to a substantial improvement of the paper. This work was supported by the National Natural Science Foundation of China (Grant nos. 50905114 and 51275558), National Science and Technology Support Program of China (Grant no. 2012BAF06B03), the research fund of State Key Lab of MSV of China (Grant no. MSV-

2010-13), and the research fund of Aerospace Advanced Technology Research Center founded by Shanghai Aerospace Technology Research Institute and Shanghai Jiao Tong University (Grant no. 12GFZ-JJ08-011).

References

Adams, B.M., Tseng, I.T., 1998. Robustness of forecast-based monitoring schemes. Journal of Quality Technology 30 (4), 328–339.

Alt, F.B., 1984. Multivariate quality control. The Encyclopedia of Statistical Sciences, 110–122.

Alwan, L.C., Roberts, H.V., 1988. Time-series modeling for detecting level shifts of autocorrelated processes. Journal of Business and Economics Statistics 6, 87–96.

Aparisi, F., Haro, C.L., 2003. A comparison of T^2 control charts with variable sampling schemes as opposed to MEWMA chart. International Journal of Production Economics 41, 2169–2182.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society (B) 26, 211–252.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 1994. Time Series Analysis: Forecasting and Control. Prentice-Hall, Englewood Cliffs, NJ.

Cook, D.F., Chiu, C.C., 1998. Using radial basis function neural networks to recognize shifts in correlated manufacturing process parameters. IIE Transaction 30 (3), 227–234.

Chinnam, R.B., 2002. Support vector machines for recognizing shifts in correlated and other manufacturing processes. International Journal of Production Research 40, 4449–4466.

Cristianini, N., Shawe-Taylor, J., 2000. A Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, Cambridge, UK.

Cook, D.F., Zobel, C.W., Nottingham, Q.J., 2001. Utilization neural networks for the recognition of variance shifts in correlated manufacturing process parameters. International Journal of Production Research 39 (17), 3881–3887.

Costa, A.F.B., Machado, M.A.G., 2011. Variable parameter and double sampling $\overline{X}$ charts in the presence of correlation: the Markov chain approach. International Journal of Production Economics 130, 224–229.

Cortes, C., Vapnik, V., 1995. Support vector network. Machine Learning 20, 273–297.

Deng, S., Yeh, T.H., 2011. Using least squares support vector machines for the airframe structures manufacturing cost estimation. International Journal of Production Economics 131, 701–708.

Du, S., Xi, L., 2010. An integrated system for on-line intelligent monitoring and identifying process variability and its application. International Journal of Computer Integrated Manufacturing 23 (6), 529–542.

Du, S., Xi, L., 2011. Fault diagnosis in assembly processes based on engineering-driven rules and PSOSAEN algorithm. Computers and Industrial Engineering 60, 77–88.

Du, S., Xi, L., Ni, J., Pan, E., Liu, C.R., 2008. Product lifecycle-oriented quality and productivity improvement based on stream of variation methodology. Computers in Industry 59 (2–3), 180–192.

De Magalhãesa, M.S., Costa, A.F.B., Moura Neto, F.D., 2009. A hierarchy of adaptive $\overline{X}$ control charts. International Journal of Production Economics 119, 271–283.

Eberhart, C.R., Kennedy, J., 1995. Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks. Piscataway, NJ, pp. 1942–1948.

Gunn, S.R., 1998. Support Vector Machines for Classification and Regression. Technical Report. University of Southampton.

Hwarng, H.B., 2004. Detecting process mean shift in the presence of autocorrelation: a neural-network based monitoring scheme. International Journal of Production Research 42, 573–595.

Hwarng, H.B., 2005. Simultaneous identification of mean shift and correlation change in AR(1) processes. International Journal of Production Research 43 (2), 1761–1783.

Issam, B.K., Mohamed, L., 2008. Support vector regression based residual MCUSUM control chart for autocorrelated process. Applied Mathematics and Computation 201, 565–574.

Jamal, J., Seyed, T., Zkhavan, N., Babak, A., 2007. Artificial neural networks in applying MCUSUM residuals charts for AR(1) processes. Applied Mathematics and Computation 189, 1889–1901.

Jiang, W., Tsui, K.L., Woodall, W.H., 2000. A new SPC monitoring method: the ARMA chart. Technometrics 42 (4), 399–410.

Kalgonda, A.A., Kulkarni, S.R., 2004. Multivariate quality process control for autocorrelated processes. Journal of Applied Statistics 31, 317–327.

Kumar, S., Choudhary, A.K., Kumar, M., Shankar, R., Tiwari, M.K., 2006. Kernel distance-based robust support vector methods and its application in developing a robust K-chart. International Journal of Production Research 44, 77–96.

Lee, P.H., Torng, C.C., Liao, L.F., 2012. An economic design of combined double sampling and variable sampling interval $\overline{X}$ control chart. International Journal of Production Economics 138, 102–106.

Mason, R.L., Tracy, N.D., Yong, J.C., 1995. Decomposition of it T^2 for multivariate control chart interpretation. Journal of Quality Technology 27 (2), 140–143.

Montgomery, D.C., 2001. Statistical Quality Control, 4th ed.. Wiley, New York.

Montgomery, D.C., Friedman, D.J., 1989. Statistical Process Control in Computer Integrated Manufacturing. Dekker, New York.

Montgomery, D.C., Mastrangelo, C.M., 1991. Some statistical process control methods for autocorrelated data. Journal of Quality Technology 23, 179–193.

Muller, K.R., Smola, A., Ratsch, G., Scholkopf, B., Kohlmorgen, J., Vapnik, V., 1997. Predicting time series with support vector machines. In: Gerstner, W., Germond, A., Hasler, M., Nicoud, J.D. (Eds.), Artificial Neural Networks ICANN 97, Lecture Notes in Computer Science, vol. 1327, pp. 999–1006.

Orlando, O., Atienza, L.C., Tang, B.W., 2002. A CUSUM scheme for autocorrelated observations. Journal of Quality Technology 34, 187–199.

Ou, Y., Wu, Z., Goh, T.N., 2011. A new SPRT chart for monitoring process mean and variance. International Journal of Production Economics 132, 303–314.

Pacella, M., Semeraro, Q., 2007. Using recurrent neural networks to detect changes in autocorrelated processes for quality monitoring. Computers and Industrial Engineering 52 (4), 502–520.

Pan, X., Jarrett, J., 2007. Using vector autoregressive residuals to monitor multivariate processes in the presence of serial correlation. International Journal of Production Economics 106, 204–216.

Schmid, W., 1997. On EWMA charts for time series. Frontiers in Statistical Quality Control 5, 115–137.

Sun, R., Tsung, F., 2003. A kernel-distance-based multivariate control chart using support vector methods. International Journal of Production Research 41, 2975–2989.

Testik, M.C., 2005. Model inadequacy and residuals control charts for autocorrelated processes. Quality and Reliability Engineering International 21, 115–130.

Thissena, U., Brakela, R.V., Weijerb, A.P., Melssena, W.J., Buydens, L.M.C., 2003. Using support vector machines for time series prediction. Chemometrics and Intelligent Laboratory System 69, 35–49.

Timmer, D.H., Pignateillo, J.J.R., Longnecker, M., 1998. The development and evaluation of CUSUM-based control charts for an AR(1) process. IIE Transactions 30 (6), 525–534.

Torng, C.C., Lee, P.H., Liao, N.Y., 2009. An economic-statistical design of double sampling control chart. International Journal of Production Economics 120, 495–500.

Vapnik, V., 1998. Statistical Learning Theory. Wiley, New York, NY.

Wardell, D., Moskowita, H., Plante, R., 1992. Control charts in presence of data correlation. Management Science 38, 1084–1105.

Wardell, D., Moskowita, H., Plante, R., 1994. Run-length distribution of special cause control charts of correlation processes. Techmetrics 36 (1), 3–17.

Wright, C.M., Booth, D.E., Hu, M.Y., 2001. Joint estimation: SPC method for shirt-run autocorrelated data. Journal of Quality Technology 33 (3), 365–378.

Wu, Z., Shamsuzzaman, M., Wang, Q., 2007. The cost minimization and manpower deployment to SPC in a multistage manufacturing system. International Journal of Production Economics 106, 275–287.

Wu, Z., Khoo, M.B.C., Shu, L., Jiang, W., 2009. An np control chart for monitoring the mean of a variable based on an attribute inspection. International Journal of Production Economics 121, 141–147.

Yang, S.F., Rahim, M., 2005. Economic statistical process control for multivariate quality characteristics under Weibull shock model. International Journal of Production Economics 98, 215–226.

Zhang, N.F., 1998. A statistical control chart for stationary process data. Technometrics 40 (1), 24–38.

Zobel, C.W., Cook, D.F., Nottingham, Q.J., 2004. An augmented neural network classification approach to detecting mean shifts in correlated manufacturing process parameters. International Journal of Production Research 42 (4), 741–758.