A regression control chart for autocorrelated processes

Article *in* International Journal of Industrial and Systems Engineering · January 2014

DOI: 10.1504/IJISE.2014.058838

2 authors:

Aslan deniz Karaoglan
Balikesir University

Mirac Bayhan
Dokuz Eylul University

A regression control chart for autocorrelated processes

Aslan Deniz Karaoglan*

Department of Industrial Engineering,
Balikesir University,
Cagis Campus, 10145, Balikesir-Turkey
Fax: +902666121257
E-mail: deniz@balikesir.edu.tr
*Corresponding author

Gunhan Mirac Bayhan

Department of Industrial Engineering,
Dokuz Eylul University,
Tinaztepe Campus, 35160, Izmir-Turkey
Fax: +902323017210
E-mail: mirac.bayhan@deu.edu.tr

Abstract: In this study, we present a new regression control chart which is able to detect the mean shift in a production process. This chart is designed for autocorrelated process observations having a linearly increasing trend. Existing approaches may individually cope with autocorrelated and trending data. The proposed chart requires the identification of trend stationary first order autoregressive (trend AR(1)) model as a suitable time series model for process observations. For a wide range of possible shifts and autocorrelation coefficients, performance of the proposed chart is evaluated by simulation experiments. Average correct signal rate and average run length are used as performance criteria.

Keywords: quality control; statistical process control; SPC; autocorrelation; linear trend; trend AR(1).

Reference to this paper should be made as follows: Karaoglan, A.D. and Bayhan, G.M. (2014) 'A regression control chart for autocorrelated processes', *Int. J. Industrial and Systems Engineering*, Vol. 16, No. 2, pp.238–256.

Biographical notes: Aslan Deniz Karaoglan received a Diploma degree in Industrial Engineering from Gazi University (Turkey) and PhD in Industrial Engineering from Dokuz Eylul University (Turkey). His research interests are statistical process control, design of experiments, neural networks, scheduling, and Petri nets.

Gunhan Mirac Bayhan received a Diploma degree in Industrial Engineering and a PhD in Computer Science from Ege University (Turkey). Her research interests are neural network based optimisation, statistical quality control, analytic and numerical modelling, time series analysis and forecasting, and Petri nets.

1 Introduction

1.1 General overview

Control charts are statistical process control tools used to determine whether a process is in-control. Since the first control chart has been proposed by Shewhart (1931), lots of charts have been developed and then improved to be used for different process data. In its basic form, a Shewhart control chart compares process observations with a pair of control limits. For developing a Shewhart chart two assumptions are to be made:

1 process data are independently distributed

2 the distribution function underlying process data is normal.

The most frequently reported effect on control charts of violating such assumptions is the erroneous assignment of the control limits. Alwan and Roberts (1988) showed that about 85% of a sample of 235 control chart applications displayed incorrect control limits (Pacella and Semeraro, 2007). More than half of these displacements were due to violation of the independence assumption. Misplacement of control limits was due to serial correlation (i.e., autocorrelation) in the data. However, many processes such as those found in refinery operations, smelting operations, wood product manufacturing, waste-water processing and the operation of nuclear reactors have been shown to have autocorrelated observations.

In the literature three general approaches are recommended for autocorrelated data:

1 fit ARIMA model to data and then apply traditional control charts such as Shewhart, CUSUM, EWMA to process residuals

2 monitor the autocorrelated observations by modifying the standard control limits to account for the autocorrelation

3 eliminate the autocorrelation by using an engineering controller (Montgomery, 1997; Testik, 2005).

On the other hand, if independent process data exhibit an underlying trend due to systemic causes, usually control charts based on ordinary least squares (OLS) regression are used for monitoring and control. In chemical processes linear trend often occurs because of settling or separation of the components of a mixture. They can also result from human causes, such as operator fatigue or the presence of supervision. The traditional Shewhart control chart with horizontal control limits and a centerline with a slope of zero have been proven unreliable when systemic trend exists in process data. A device useful for monitoring and analysing processes with trend is the regression control chart. Rather than using standard Shewhart charts, practitioners typically implement regression based control charts to monitor a process with systemic trend (Utley and May, 2008).

1.2 Literature review

A regression based control chart which is the combination of the conventional control chart and regression analysis was first proposed by Mandel (1969). Mandel (1969) used regression control chart to monitor the variety of postal management problems. This chart is designed to control a varying (rather than a constant) average, and assumes that the values of the dependent variable are linearly (causally) related with the values of the independent variable. However, since Mandel's regression control chart was developed for independent data, it is not an effective tool for monitoring process shift in correlated process observations.

Mandel also devised a simplification of the regression control chart. The simplification functioned as a residual chart because the values that were plotted on it were the residuals from the regression analysis (Utley and May, 2009). Zhang (1984) adopted Mandel's idea of a residual control chart for statistical process control data in the cause selecting chart (CSC). The CSC is constructed for an outgoing quality characteristic only after it has been adjusted for the effect of incoming quality characteristic. Hawkins (1991) developed a procedure called regression adjustment. The scheme essentially consists of plotting univariate control charts of the residuals from each variable obtained when that variable is regressed on all the others (Montgomery, 2009). A very important application of regression adjustment occurs when the process has a distinct hierarchy of variables, such as a set of input process variables and a set of output variables. Sometimes this situation is called a cascade process. If the proper set of variables is included in the regression model, the residuals from the model will typically be uncorrelated, even though the output variable exhibited correlation. The regression adjustment procedure has many possible applications in chemical and process plants where there are often cascade processes with several inputs but only a few outputs, and where many of the variables are highly autocorrelated at low lags (Montgomery, 2009). Two years later, Hawkins (1993a) applied regression control chart to cascade processes and cited CSC as a particularly useful methodology for controlling quality in cascade processes. If linear regression is used to model a cascade process, then the values plotted on the cause selecting control chart are actually the standardised residuals from the regression relationship (Sulek et al., 2006). In the same year, Wade and Woodall (1993) reviewed the concepts of the CSC and examined the relationship between the cause-selecting chart and multivariate Hotelling T^2 chart. In their opinion, the cause-selecting approach is an improvement over the use of separate Shewhart control charts for each of two related quality characteristics. A review of the literature on control charts for multivariate quality control (MQC) is given by Lowry and Montgomery (1995), by discussing principal components and regression adjustment of variables in MQC. Haworth (1996) used a multiple regression control chart to manage software maintenance. A quality control tool was developed for managers of complex software maintenance processes that can be modelled with a multiple regression model. Kalagonda and Kulkarni (2003) proposed a diagnostic procedure called '*D*-technique' to detect the nature of shift. For this purpose, two sets of regression equations, each consisting of regression of a variable on the remaining variables, are used to characterise the 'structure' of the 'in-control' process and that of the 'current' process. To determine the sources responsible for an out of control state, it is shown that it is enough to compare these two structures using the dummy variable multiple regression equation. In the same year, Omura and Steffe (2003) constructed Mandel's regression control chart for apparent viscosity and average shear rate data. According to the authors, no standardised test existed to objectively assess flow behaviour of fluid foods with large particulates. Therefore, to monitor the process data using a regression control chart could be useful for quality control. In the following year; Shu et al. (2004) studied the run-length performance of EWMAREG (EWMA chart for regression residuals) and SheREG

(Shewhart chart for regression residuals) with estimated parameters of regression equation, and used these charts for monitoring multistage processes where process data usually follow a multivariate normal distribution. The authors also studied the run length performance of regression control charts. However, Zhang (1984) and Wade and Woodal (1993) considered the CSC with sample size one, while the studies about construction of cause-selecting charts with sample size greater than one are discussed by Yang (2005) for joint $\overline{x}$ and $\overline{e}$ cause-selecting charts. Yang and Yang (2005) considered the problem of monitoring the mean of a quality characteristic x on the first process step, and the mean of a quality characteristic y on the second process, in which the observations x can be modeled as an ARMA model and observation y can be modelled as a transfer function of x since the state of the second process is dependent on the state of the first process. In the following year, Yang and Yang (2006) addressed the $\overline{x} - s^2$ and $\overline{e} - s_e^2$ charts for two dependent process steps with over-adjusted means and variances. Sulek et al. (2006) examined the CSC as a methodology to monitor and identify potential problem areas in an actual cascade service process. The authors utilised the CSC, a type of a regression based control chart, as an appropriate methodology for analysing the performance of a downstream stage in a multistage process by controlling the effect of performance in the upstream stage. Yang and Su (2007) constructed an adaptive sampling interval Z_x control chart to monitor the quality variable produced by the first process step, and used the adaptive sampling interval Z_e control chart to monitor the specific quality variable produced by the second process step. Asadzadeh et al. (2008) reviewed CSC for monitoring and diagnosing multistage processes. The following year, Asadzadeh et al. (2009) proposed a robust CSC to monitor multistage processes where outliers are presented in historical dataset. In the same year, Yang and Chen (2009) constructed the variable sampling interval (VSI) $Z_{\overline{x}} - Z_{s^2}$ and $Z_{\overline{e}} - Z_{s_{e^2}}$ control charts in order to effectively monitor the quality variable produced by the first process step with incorrect adjustment and the quality variable produced by the second process step with incorrect adjustment, respectively. When the residual terms are not normally distributed, an alternative method for estimating the regression line is needed. One alternative method is the least absolute value (LAV) regression model. In contrast to the OLS approach, which minimises the sum of the squared residuals, the LAV model minimises the sum of the absolute values of the residuals. Utley and May (2009) proposed a control chart methodology for residual control charts that is based on LAV regression. Tsai and Chen (2009) monitored the stencil printing process using a modified regression residual control chart.

Jarrett and Pan (2009) suggested multivariate methods for the construction of quality control charts for the control and improvement of output of manufacturing processes. Ben Khediri et al. (2010) proposed support vector regression which is a non-parametric method to construct several control charts that allow monitoring of multivariate non-linear autocorrelated processes. Capizzi and Masarotto (2011) develop a control chart which integrates the least angle regression algorithm with a multivariate exponentially weighted moving average. They combined a variable selection method with a multivariate control chart to detect changes in both the mean and variability of a multidimensional process with Gaussian errors. Yu and Liu (2011) proposed a logistic regression (LR)-based process monitoring model for enhancing the monitoring of processes and developed logistic regression probability chart (LRProb chart). Kim et al.

(2012) attempted to integrate state-of-the-art data mining algorithms (artificial neural networks, support vector regression, and multivariate adaptive regression splines) with SPC techniques to achieve efficient monitoring in multivariate and autocorrelated processes. The residuals of data mining models were utilised to construct multivariate cumulative sum control charts (MCUSUM) to monitor the process mean.

In addition to autocorrelated or trended observations, many industrial processes give such data that exhibit both trend and autocorrelation among adjacent observations. This observation has been the motivation for the present work on developing a new regression control chart (NRC chart) that cope with autocorrelated observations in which observation values increase with respect to time. The NRC chart requires identification of trend AR(1) model as a suitable time series model for observations. As it is known, a traditional residual chart considers only the current sample when determining the status of a process and hence does not provide any pattern-related information. By using NRC we are able to monitor current samples of given autocorrelated and trending process directly and to observe the progress of the process.

The rest of the paper is organised as follows. The next section describes a step-by-step procedure to facilitate practitioners in establishing the NRC chart. In Section 3, an illustrative example is given. Performance of the chart is evaluated in Section 4. Conclusions are pointed out in Section 5.

2 Establishing the NRC chart

An autoregressive process of lag 1, AR(1), is the representative model for autocorrelated processes. In an AR(1) process, the current observation is correlated with its previous observation. Past studies emphasise the role of AR(1) processes in process control (Guh, 2008). An AR(1) model can be expressed as follows:

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t \tag{1}$$

where t is the time of sampling, x_t is the sample value at time t, ξ is the constant, ϕ is the autoregressive coefficient ($-1 < \phi < 1$), and ε_t is the independent random error term (common cause variation) at time t following $M\left(0, \sigma_\varepsilon^2\right)$. Let autocorrelated process observations with an increasing linear trend (trend AR(1) process) (X_t) be represented by:

$$X_t = x_t + dt \tag{2}$$

where d is the trend slope and t is the time step (or observation number), and autocorrelated and trending process observations (X_t) with a mean shift be depicted by:

$$Z_t = X_t + \delta_\mu \tag{3}$$

where δ_μ is the magnitude of upward mean shift. In this study, our aim is to test for an upward shift in the mean of $\{Z_t\}$ by using the NRC chart. If a trend and autocorrelation are diagnosed in the data, then the NRC chart is constructed in the following five steps:

Step 1 Fit a simple linear regression model to the data and calculate the standard deviation of the process observations.

The centre line of the NRC chart is a simple regression line as is in a conventional regression chart and formularised as $\hat{Z}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \varepsilon_t$ where Z_t is the tth observation, β_0 (intercept) and β_1 (slope) refer to unknown coefficients, and their values are determined by the method of least squares. ε_t is assumed to be a random variable with mean zero. We let $\hat{\beta}_0$ and $\hat{\beta}_1$ denote the least squares estimates for β_0 and β_1 respectively,

$$\hat{\beta}_0 = \overline{Z} - \hat{\beta}_1 \overline{T} \tag{4}$$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{N}(Z_t - \overline{Z})(t - \overline{T})}{\sum_{t=1}^{N}(t - \overline{T})^2} \tag{5}$$

$$\overline{Z} = \frac{\sum_{t=1}^{N} Z_t}{N} \quad \text{and} \quad \overline{T} = \frac{\sum_{t=1}^{N} t}{N} \tag{6}$$

where N represents the sample size and $\overline{Z}$ the mean value of Z_t. The standard deviation of the process σ is estimated by

$$\hat{\sigma} = \sqrt{\frac{\sum_{t=1}^{N}(Z_t - \overline{Z})^2}{N-1}} \tag{7}$$

Step 2 Calculate the NRC variation parameter (NRCVP for short) to represent the variation in the sample data.

$$NRCVP = \sqrt{\hat{\sigma}_e \left(\frac{1}{1-\phi^2} \right)} \tag{8}$$

where

$$\hat{\sigma}_e = \sqrt{\frac{\sum_{t=1}^{N}(e_t - \overline{e})^2}{N-1}} \tag{9}$$

is the standard deviation of $\{e_t\}$ which is the difference between expected and observed values of Z_t,

$$e_t = Z_t - \phi Z_{t-1}, \tag{10}$$

and $\overline{e}$ is the sample mean of $\{e_t\}$

$$\overline{e} = \frac{\sum_{t=1}^{N} e_t}{N} \tag{11}$$

Step 3 Calculate the time dependent EWMAARZ parameter to widen the control limits up. The aim is to obtain an acceptable false alarm rate.

$$\text{EWMAARZ} = \sqrt{\frac{\sum_{t=1}^{N} \hat{\sigma}_{EWMA(t)}}{N}} \tag{12}$$

where

$$\hat{\sigma}_{EWMA(t)} = \alpha e_t^2 + (1-\alpha)\hat{\sigma}_{EWMA(t-1)} \tag{13}$$

and $\hat{\sigma}_{EWMA(1)} = \hat{\sigma}_e$. During the experiments we conducted, we observed that the NRC chart gives better performance for $\alpha = 0.80$. $\hat{\sigma}_{EWMA}$ is the estimated smoothed standard deviation of Z_t and shows similar characteristics with the smoothed variance in moving centerline EWMA (MCEWMA) chart.

As can be seen in equation (12), EWMAARZ is affected by the process residuals, and implicitly by the autoregressive parameter ϕ. As depicted in equation (10), residuals get larger for $\phi < 0$ and vice versa.

Step 4 Calculate $\bar{C}^+ \left(\bar{C}^-\right)$ to widen (narrow) the upper control limit if there is an upward (downward) shift in the mean of the process.

There is an upward (downward) shift in the process mean if $\bar{C}^+ \left(\bar{C}^-\right)$ continues to get larger (smaller).

$$\bar{C}^+ = \frac{\sum_{t=1}^{N} C^+_{NRC(t)}}{N} \tag{14}$$

$$\bar{C}^- = \left| \frac{\sum_{t=1}^{N} C^-_{NRC(t)}}{N} \right| \tag{15}$$

where

$$\begin{aligned} &C^+_{NRC(1)} = 0, C^-_{NRC(1)} = 0, C^+_{NRC(t)} = \left\{\max[0, Z_t - (\hat{M}_o^t + k)]\right\}, \\ &C^-_{NRC(t)} = \left\{\min[0, Z_t - (\hat{M}_o^t - k)]\right\} \end{aligned} \tag{16}$$

where

$$\hat{M}_o^t = \phi^2 \frac{D_0}{2} + \hat{\beta}_1 t \qquad t = 1, 2, \ldots, N \tag{17}$$

denotes the target varying mean for process observations at each time step t. In the CUSUM control chart, deviations from the constant target mean (μ_0) are used to calculate accumulating deviations. In the present paper, with a similar purpose, we use $\hat{M}_o^t$ as a reference value for the process observations. The distance between each process observation and $\hat{M}_o^t$ is calculated in each time step *t*. On the other hand, if we consider the conventional regression control chart that is proposed by Mandel, the centre line formulated by $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 t$ also represents the target varying mean. Intercepts show difference between $\hat{M}_o^t$ and the target varying mean in the Mandel's regression control chart.

The intercept in $\hat{M}_o^t$ is formulated as $\phi^2 \dfrac{D_o}{2}$. Calculations for D_o are given in Table 1. For positive autocorrelation, because of the nature of the process, a relatively large observation at the previous time step tends to be followed by another large value at the current time step. NRC adjusts its control limits' width with respect to the autocorrelated process observations by using $\bar{C}^+$ and $\bar{C}^-$. For this purpose, NRC takes into consideration the sign of autocorrelation and the magnitude of autocorrelation coefficient. The functional role of $\bar{C}^+$ and $\bar{C}^-$, that are the means of calculated $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$ values, will be given in detail at the following pages. As mentioned before $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$ represent derivations from the target varying mean at each time step *t*. By using $\bar{C}^+$ and $\bar{C}^-$ we decrease the width of the control limits. In other words, $\bar{C}^+$ and $\bar{C}^-$ have decreasing effect on the width between upper and lower control limits. For weak autocorrelation cases we expect the control limits to get narrower when it is compared with the strong autocorrelation. So the calculated values for $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$ are expected to be smaller for strong autocorrelation to get larger control limits. As it can be observed from the formulations of $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$ that are given in equation (16), this can be provided by determining large $\hat{M}_o^t$ to get smaller $C^+_{NRC(t)}$ values for strong autocorrelation and vice versa. By including the square of autocorrelation coefficient in the intercept, it is aimed to have larger (respectively smaller) target varying mean when observations are strongly autocorrelated (respectively weakly autocorrelated) in order to increase the correct signal rate.

While calculating $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$, we use a slack value *k* to prevent the inclusion of small deviations from the process mean. In the relevant literature, *k* is often chosen as a halfway between the target mean and the out-of-control value of the mean (Montgomery, 1997). It is important to select the right value for *k* since a large value of *k* will allow for large shifts in the mean without detection, while a small value of *k* will increase the frequency of false alarms. For a conventional CUSUM chart, *k* is selected to be equal to 0.5 σ. During our comprehensive experiments, we observed that the NRC gives better performance for $k = \dfrac{NRCVP}{6}$.

The reason for calculating $C^+_{NRC(t)}$ and $C^-_{NRC(t)}$ is similar to that of C_t^+ (upper cumulative sum) and C_t^- (lower cumulative sum) statistics in CUSUM chart (Montgomery, 1997). The basic purpose of a CUSUM chart is to track the distance between the actual data point and the grand mean. By keeping a cumulative sum of these distances, it can be determined if there is a change in the process mean. But, because $C^{+}_{NRC(t)}$ and $C^{-}_{NRC(t)}$ display some distinct characteristics from C^{+}_{t} and C^{-}_{t}, they are time dependent, and are not affected from their previous values, and we select the minimum value while calculating the $C^{-}_{NRC(t)}$ (as depicted in Equation (16)), not the maximum value as in the calculation of C^{-}_{t} (Montgomery, 1997).

Step 5 Calculate the other parameters required for constructing the NRC.

These parameters and the formulas used for calculations are given in Tables 1 through 4. As can be seen in Table 1 and Table 2, to calculate B, B_2, D_0 and B_3, B_4, L, D_1, D_3, D_4 and D_5, distinct formulas are employed depending on the sign of $\hat{\beta}_0$ and of ϕ. On the other hand, signs of both $\hat{\beta}_0$ and determine which formulas will be used for D_2^{+} and D_2^{-}. As can be seen in Table 3, if $\phi > 0$, while deciding which formula will be used for D_2, magnitude of $\hat{\beta}_0$ is compared with $\phi NRCVP$. According to our experiments, the NRC gives better performance in terms of false alarm rate with the parameters given in Tables 1 through 4, and the false alarm probability is lower when $\hat{\beta}_0$ and $\hat{\beta}_1$ have the same sign than when these parameters have opposite signs. Also big values of C_t^{+} increase the correct alarm rate for upward shift and vice versa. The magnitude and the sign of $\hat{\beta}_0$ directly affect the control chart's performance. If the data have positive autocorrelation, unless the shift size is not changed, the performance gets better (worse) as the magnitude of positive $\hat{\beta}_0$ (negative $\hat{\beta}_0$) gets bigger, and vice versa for negative autocorrelation. It is because the signs of $\hat{\beta}_0$ and ϕ affect the performance of the NRC chart that Tables 1 through 4 are arranged in respect of the signs of these parameters. Because the NRC chart has several parameters, the design procedure seems to be complicated. However, as can be seen in Tables 1 – 4, specifying values for some parameters can decrease the number of NRC parameters and reduce the calculation complexity. Tables 1–4 show these special cases of the NRC chart.

Table 1 Parameter calculations according to the sign of $\hat{\beta}_0$

If $\hat{\beta}_0 < 0$	**If $\hat{\beta}_0 \geq 0$**
$B_2^{+} = 0 \quad B_2^{-} = 0$	$B_2^{+} = 1 \quad B_2^{-} = 1$
$B^{+} = \hat{\beta}_o + \left(\frac{EWMAARZ}{\overline{e}}\right)^2$	$B^{+} = 0 \quad B^{-} = 0$
$B^{-} = -\hat{\beta}_o + \left(\frac{EWMAARZ}{\overline{e}}\right)^2$	
$D_0 = NRCVP$	$D_0 = NRCVP$

Table 2 Parameter calculations according to the sign of ϕ

If $\phi > 0$	If $\phi < 0$
$B_3^+ = 1 \quad B_3^- = 1$	$B_3^+ = \dfrac{\hat{\beta}_1}{\lvert\phi\rvert\lvert\hat{\beta}_0\rvert} \quad B_3^- = \dfrac{\hat{\beta}_1}{\lvert\phi\rvert\lvert\hat{\beta}_0\rvert}$
$L = 3.0$	$L = 1.5$
$D_1 = 0 \quad D_3 = 1 \quad D_4 = 1 \quad D_5 = 1$	$D_1 = \dfrac{1}{\lvert\phi\rvert} \quad D_3 = 3L\phi^2 \quad D_4 = -1.5 \quad D_5 = \lvert\phi\rvert$
$B_4^+ = 1 \quad B_4^- = 1$	$B_4^+ = \lvert\phi\rvert \quad B_4^- = \lvert\phi\rvert$

Table 3 Parameter calculations for D_2 when $\phi > 0$

$\hat{\beta}_0 > 0$	$\hat{\beta}_0 > \phi NRCVP$	$D_2^+ = \dfrac{\hat{\beta}_0}{\phi(L/2)} + \dfrac{\hat{\beta}_0}{L(NRCVP)}(L/2)\dfrac{\bar{C}^-}{\hat{\beta}_1}$
		$D_2^- = \dfrac{\hat{\beta}_0}{\phi(L/2)} + \dfrac{\hat{\beta}_0}{L(NRCVP)}(L/2)\dfrac{\bar{C}^-}{\hat{\beta}_1} + 3\bar{C}^-\phi^2$
	$\hat{\beta}_0 \le \phi NRCVP$	$D_2^+ = \phi\hat{\beta}_0 + \phi^2 3L\bar{C}^- - (1/3)\bar{C}^+(1/\phi)$
		$D_2^- = \phi\hat{\beta}_0 + \phi^2 3L\bar{C}^- + (1/3)\bar{C}^+(1/\phi)$
$\hat{\beta}_0 \le 0$		$D_2^+ = \phi 2L(\bar{C}^- - \hat{\beta}_0)$
		$D_2^- = \phi 2L(\bar{C}^+ - \hat{\beta}_0)$

Table 4 Parameter calculations for D_2 when $\phi < 0$

$\hat{\beta}_0 > 0$	$D_2^+ = (-1/3)\dfrac{\hat{\beta}_0}{\phi^2} + (L)\left(\dfrac{\hat{\beta}_0}{L(NRCVP)}\right)^2\left(\dfrac{\bar{C}^-\hat{\beta}_1}{\phi^2}\right) + \dfrac{\bar{C}^+}{\lvert\phi\rvert} - \dfrac{\bar{C}^-}{\phi^2}$
	$D_2^- = \dfrac{\hat{\beta}_0}{\lvert\phi\rvert} + (L)\left(\dfrac{\hat{\beta}_0}{L(NRCVP)}\right)^2\left(\dfrac{\bar{C}^-\hat{\beta}_1}{\phi^2}\right) + 2\dfrac{\bar{C}^-}{\lvert\phi\rvert} - L\dfrac{\bar{C}^+}{\lvert\phi\rvert}$
$\hat{\beta}_0 \le 0$	$D_2^+ = \phi 2L(\bar{C}^- - \hat{\beta}_0)$
	$D_2^- = \phi 2L(\bar{C}^+ - \hat{\beta}_0)$

Step 6 Calculate control limits and the center line.

Control limits and the centre line of the NRC chart are regression lines as given below.

Center line:

$$CL_t = \hat{\beta}_0 + \hat{\beta}_1 t \tag{18}$$

Upper control limit (UCL):

$$UCL_t = Y_{UCLpre_t} + L(NRCVP)\lvert\phi\rvert B_4^+ + D_2^+ - \frac{D_3 D_5^2 \bar{C}^+}{\phi^2} \tag{19}$$

where

$$Y_{UCLpre_t} = B^+ + B_2^+\hat{\beta}_0 + D_1\left|\phi\right|\hat{\beta}_1\hat{\sigma} + B_3^+\left|\phi\right| EWMAARZ + \hat{\beta}_1 t \tag{20}$$

Lower control limit (LCL):

$$LCL_t = Y_{LCLpre_t} - L(NRCVP)\left|\phi\right| B_4^- - D_2^- + \frac{D_4 D_5^2 \bar{C}^-}{\phi^2} \tag{21}$$

where

$$Y_{LCLpre_t} = -B^- - B_2^-\hat{\beta}_0 - D_1\left|\phi\right|\hat{\beta}_1\hat{\sigma} - B_3^-\left|\phi\right| EWMAARZ + \hat{\beta}_1 t \tag{22}$$

We use the simple linear regression equation with intercept ($\hat{\beta}_0$) and slope ($\hat{\beta}_1$) to represent the centre line of the NRC. We also noted that if a relatively low observation from the autocorrelated process at the previous time step tends to be followed by another low value at the current time step, and a relatively large observation at the previous time step tends to be followed by another large value at the current time step, then this type of behaviour is indicative of positive autocorrelation. Naturally, the direct contrary is indicative of negative autocorrelation. So the pattern on control chart varies according to the sign of the autocorrelation. To adjust control limits of the NRC and consequently to provide a high correct signal rate, the calculations show disparities according to the combinations of the signs of intercept ($\hat{\beta}_0$) and autocorrelation coefficient (ϕ). For positive autocorrelation wider control limits are needed. The control limits' width should be narrower as the magnitude of positive autocorrelation coefficient decreases. For negative autocorrelation, control limits' width should be narrower with respect to the case of positive autocorrelation, and when strong negative autocorrelation exists it should be larger compared with the limits for weak negative autocorrelation. Control limits of the NRC are also affected from the magnitude and the sign of $\hat{\beta}_0$.

To adjust continuously the distance between the centre line and upper control limit due to the variations in observations that stem from the effect of autocorrelation, the parameters β_4^+ and L are employed. For a negative autocorrelation, with the effect of β_4^+, control limits get narrower while the autocorrelation decreases. If there is positive autocorrelation between process observations, then β_4^+ has no effect on control limits. Parameters D_2^+, D_3 and D_5 are also used to reflect the effect of deviations of observations from target varying mean with the combined effect of autocorrelation. The effect of D_2^+ on upper control limit varies according to the signs of $\hat{\beta}_0$ and ϕ. For a positive strong autocorrelation, if the sign of $\hat{\beta}_0$ is positive then the upper control limit will be wider than when $\hat{\beta}_0$ is negative. This effect begins to turn in direct contradiction with the decreasing autocorrelation between process observations. If there is a strong negative autocorrelation and if the sign of $\hat{\beta}_0$ is positive then the upper control limit will be narrower than when $\hat{\beta}_0$ is negative. The same effect continues for the decreasing negative autocorrelation from strong to weak with less impact. D_3 and D_5 decrease the width of upper control limit for negative autocorrelation since they have no impact on it for positive autocorrelation cases. Another parameter that is used for determining the width of upper control limit is $Y_{UCL\,\mathrm{Pr}\,e_t}$, which changes with respect to t. For $Y_{UCL\,\mathrm{Pr}\,e_t}$, by using parameters β^+ and β_2^+, we take into consideration the effect of smoothed standard deviation of the shifted process that depends on exponentially weighted residuals (EWMAARZ for NRC). The sign of $\hat{\beta}_0$ affects the width of the control limits. This effect is reflected in the calculations of $Y_{UCL\,\mathrm{Pr}\,e_t}$ by β^+ and β_2^+. By considering the magnitude of $\hat{\beta}_0$ and exponentially weighted residuals, control charts' upper limit gets narrower for negative $\hat{\beta}_0$ values while it gets wider for positive values. By using D_1 and B_3^+, combined effect of autocorrelation and exponentially weighted residuals are added to the mathematical formulation of upper control limit. D_1 and B_3^+ show disparities according to the sign of the autocorrelation coefficient ϕ. $Y_{UCL\,\mathrm{Pr}\,e_t}$ has an effect on determining the width of the upper control limit by reflecting the combined effect of exponentially weighted residuals as regards the signs of $\hat{\beta}_0$ and ϕ. The same approaches are employed in the calculations of the lower limit of the NRC given in equations 21–22.

3 Design implications

Figure 1 and Figure 2 display the NRC chart for 3.0σ and 0.5σ mean shifts, respectively. In these figures while the dashed line represents the shifted process, unshifted process is indicated by a solid line. To illustrate how this chart signals, we computerised the design procedure of the chart with MATLAB 7.4.0, and applied it to a sample of 500 observations generated using equation (3). The design of the chart for these sample data was completed in 0.64s (less than 1s) of CPU time on a personal computer (AMD turion, 1.79 GHz, 2.87 GB RAM). To model assignable causes, a shift is added in the mean of Z_t in equation (3) starting at observation 51. The parameter values employed for building the chart, degree of serial correlation, magnitudes of the mean shifts added to the 51st observation, and run length results (the average number of points before an out-of-control signal is observed) are listed in Table 5. As can be seen in Figures 1–2, the chart gives out-of-control signals at time steps 11 and 19 after the mean shift occurs.

Figure 1 The NRC chart for the 3.0σ shift (see online version for colours)

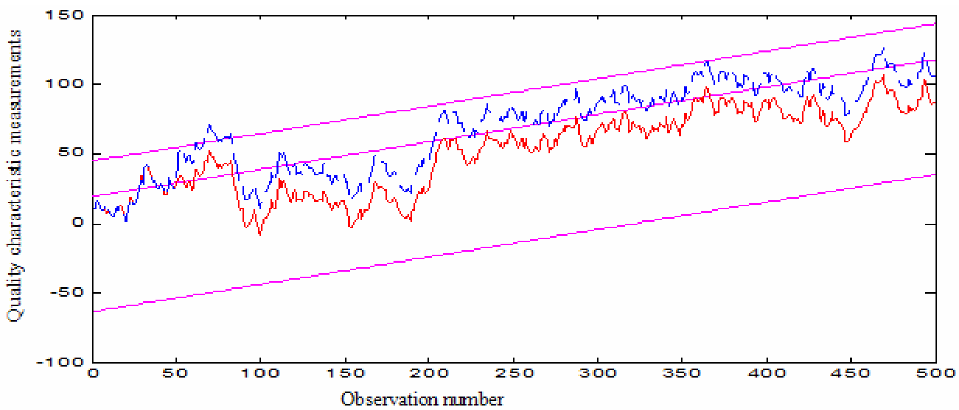

Figure 2 The NRC chart for the 0.5σ shift (see online version for colours)

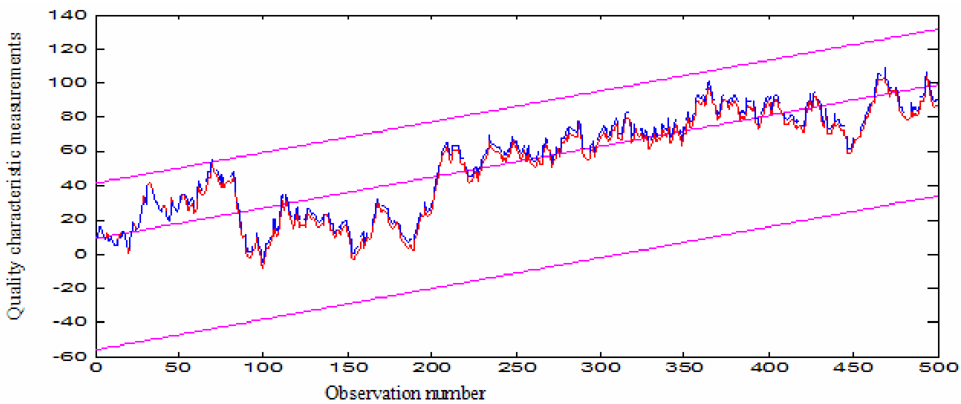

Table 5 Parameter values and the run length results for the illustrative example

Parameter	*0.5σ*	*3.0σ*
ϕ	0.95	0.95
L	3.0	3.0
$\hat{\beta}_1$	0.1768	0.1976
$\hat{\beta}_0$	6.9341	19.0294
X_1	10.2000	10.2000
$\bar{e}$	2.7054	3.6066
NRCVP	6.6163	6.7144
$\bar{C}^+$	7.5301	15.9814
$\bar{C}^-$	3.9357	0.7747
EWMAARZ	5.0509	5.6830
Run length	19	11

4 Performance of the NRC chart

In this section, we evaluate the average correct signal rate and the simulated average run length (ARL) performance of the NRC chart using the following design parameters $\xi = 0$, $X_1 = 10$, $\varepsilon \sim N(0,4)$, $N = 500$, and $d = 0.2$.

To investigate the performance, we generated data sets using equation (3), as we did in Section 3, and employing a wide range of possible shifts and autocorrelation coefficients. Each data set involves 500 observations. A shift is added in the mean of Z_t in (3) starting at observation 51. The considered shift magnitudes and autocorrelation coefficients are δ_μ = 0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, and ϕ = 0.95, 0.475, –0.475, –0.95, respectively. For the sake of simplicity, we classified shift magnitudes in three groups as small (δ_μ = 0.5, 1.0), moderate (δ_μ = 1.5, 2.0), large (δ_μ = 2.5, 3.0), and autocorrelation coefficients as weak (ϕ = 0.475, –0.475) and strong (ϕ = 0.95, –095). For each dataset 1,000 simulation replications are performed. Simulation results are explained in detail below.

If the test statistic of a shifted process does not fall between the control limits or the test statistic of an unshifted process falls between the control limits it is said that the control chart's signal is correct (Montgomery, 1997). In this context, we computed the average correct signal rates for several shift-autocorrelation combinations, which are displayed in Table 6. As can be seen from this table, signals of the chart are thoroughly accurate for all shift magnitudes in the presence of strong and weak negative autocorrelation. Its correct signal rate is also very good for large shift-positive autocorrelation combinations.

Table 6 Average correct signal rate

ϕ / δ_μ	*(0.95)*	*(0.475)*	*(–0.475)*	*(–0.95)*
0.0	0.8310	0.9350	1.0000	0.9990
0.5	0.3475	0.5833	1.0000	0.9965
1.0	0.5101	0.7547	1.0000	0.9981
1.5	0.6159	0.8105	1.0000	0.9984
2.0	0.6876	0.8132	1.0000	0.9986
2.5	0.7388	0.8803	1.0000	0.9988
3.0	0.7675	0.9180	1.0000	0.9989

The simulated ARL performance of the NRC chart is shown in Table 7. In this table, ARL for δ_μ = 0.0 indicates ARL_0, in-control performance of the chart. We can see that the chart has large in-control ARL but small out-of-control ARL. That is, when the process has no mean shift the ARL is very large, and when a mean shift occurs the ARL decreases to indicate the occurrence of the mean shift quickly (Winkel and Zhang, 2004; Zhang, 2000). Consequently, we can say that the proposed chart has an ARL performance of what a desirable chart should have. About the overall ARL performance of the NRC chart we can say that it performs well for large shifts, and shows its best ARL performance for negative autocorrelation cases. For small and moderate shifts its ARL performance is good for strong autocorrelation.

For the residuals of given dataset, the simulated ARLs were also estimated for widely used residual charts; X, CUSUM and EWMA charts. For the X chart and EWMA chart 3-sigma and 2.5-sigma control limits are used respectively. Smoothing constant (λ) is selected as 0.05 for EWMA chart. For the CUSUM chart, the decision interval with h = 4.77 and k = 0.5 was used (for details see Lucas and Saccucci, 1990; Hawkins, 1993b; Montgomery, 1997; Montgomery, 2009).

Table 7 ARL performance of the NRC

ϕ	*Mean shift*	*EWMA residual L=2.5 λ = 0.05*	*X residual L = 3*	*CUSUM residual k = 0.5 h = 4.77*	*NRC*
0.95	0.0	343.9	370.5	372.6	388.3
	0.5	290.3	335.9	308.1	46.6
	1.0	204.1	260.2	200.2	34.3
	1.5	110.7	135.2	122.4	13.7
	2.0	66.7	12.9	76.8	9.6
	2.5	44.2	3.5	50.8	3.2
	3.0	29.4	1.3	35.5	1.2
0.475	0.0	363.2	370.5	371.2	411.4
	0.5	117.6	248.8	119.2	212.6
	1.0	29.5	113.7	32.9	182.1
	1.5	11.4	43.6	11.9	41.7
	2.0	7.6	22.7	7.5	36.4
	2.5	5.4	4.6	5.5	8.1
	3.0	3.6	4.1	3.8	3.6
–0.475	0.0	379.0	370.5	368.4	449.0
	0.5	17.9	78.6	18.3	449.0
	1.0	6.8	12.2	7.1	165.3
	1.5	6.1	6.1	6.3	17.9
	2.0	3.8	3.7	4.0	4.4
	2.5	2.9	1.9	3.6	1.7
	3.0	2.1	1.4	2.3	1.1
–0.95	0.0	368.3	370.5	375.5	449.0
	0.5	11.2	6.7	12.3	57.5
	1.0	4.2	1.9	4.7	22.9
	1.5	2.9	1.8	3.1	12.6
	2.0	2.4	1.7	2.6	7.1
	2.5	2.1	1.2	2.3	3.6
	3.0	1.8	1.0	2.2	1.1

From Table 7, it is clear that when the process is strongly and positively autocorrelated NRC outperforms the residual charts for all mean shifts. For weak positive autocorrelation ARL performance of the NRC is well and outperforms the residual charts for large mean shifts. For small to moderate shifts ARL performance of NRC is competitive with the results of X residual chart. For weak negative autocorrelation NRC outperforms the residual charts for large mean shifts and its ARL performance is competitive with these charts for moderate mean shifts. When the mean shifts are large and the process is negatively and strongly autocorrelated the ARL performance of NRC is good and competitive with the residual charts. X, CUSUM and EWMA residual charts take into account the residuals from the current sample data while determining the status of a process and hence do not provide any pattern-related information. However, by using the NRC, practitioners will be able to monitor current samples of an autocorrelated and trending process directly and to observe the progress of the process.

5 Conclusions

In this article, we propose a new regression chart (the NRC chart) for detecting shifts in the mean of production processes. This chart can handle data in which observations are both autocorrelated and their values linearly increase with respect to time. The chart requires the identification of trend AR(1) model as a suitable time series model for process observations. This is one of the highlights of this paper while it is also the limitation of this study at the same time.

The average correct signal rate and ARL performance of the chart were investigated by simulation approach. Based on the results of simulation, it is safe to say that the NRC is a considerably powerful chart. As is known, no single control chart will give optimal performance across a wide variety of situations (Lu and Reynolds, 1999). In this sense, we tried to explain when the proposed control chart performs well for several types of autocorrelation structures and shift magnitudes.

As it is known, a traditional residual chart takes into account only the current sample when determining the status of a process and hence does not provide any pattern-related information. By using the NRC practitioners will be able to monitor current samples of an autocorrelated and trending process directly and to observe the progress of the process. The NRC can be easily computerised and directly applied to the original data. This is the contribution of NRC. This study could be extended for autocorrelated data with decreasing trend. For future research, comparing the performance of the NRC with alternative approaches to modelling combined effects would be of great interest to us.

Additionally statistical process control (SPC) and automatic process control (APC) are two important methods that have been used for improving product quality and process productivity. SPC is used for process monitoring, while APC is used for process adjustment. SPC reduces process variability by detecting and eliminating special causes of process variation, while APC reduces variability by adjusting the process to keep the product variability on target. Both the methods were initially thought to be in conflict with each other, but in recent years, many researchers have shown their interest in integrating SPC and APC techniques to reduce total variability of the process (Akram et al., 2012). As future research integrating NRC with APC techniques may be searched by researchers.

Acknowledgements

The authors would gratefully like to thank to the editor and the anonymous referees whose valuable suggestions lead to an improved organisation of this paper.

References

Akram, M.A., Saif, A.A. and Rahim, M.A. (2012) 'Quality monitoring and process adjustment by integrating SPC and APC: a review', *International Journal of Industrial and Systems Engineering*, Vol. 11, No. 4, pp.375–405.

Alwan, L.C. and Roberts, H.V. (1988) 'Time series modeling for statistical process control', *Journal of Business and Economic Statistics*, Vol. 6, No. 1, pp.86–95.

Asadzadeh, S., Aghaie, A. and Shahriari, H. (2009) 'Monitoring dependent process steps using robust cause-selecting control charts', *Quality and Reliability Engineering International*, Vol. 25, No. 7, pp.851–874.

Asadzadeh, S., Aghaie, A. and Yang, S.F. (2008) 'Monitoring and diagnosing multistage processes: a review of cause selecting control charts', *Journal of Industrial and System Engineering*, Vol. 2, No. 3, pp.215–236.

Ben Khediri, I., Weihs, C. and Limam, M. (2010) 'Support vector regression control charts for multivariate nonlinear autocorrelated processes', *Chemometrics and Intelligent Laboratory Systems*, Vol. 103, No. 1, pp.76–81.

Capizzi, G. and Masarotto, G. (2011) 'A least angle regression control chart for multidimensional data', *Technometrics*, Vol. 53, No. 3, pp.285–296.

Guh, R.S. (2008) 'Real-time recognition of control chart patterns in autocorrelated processes using a learning vector quantization network-based approach', *International Journal of Production Research*, Vol. 46, No. 14, pp.3959–3991.

Hawkins, D.M. (1991) 'Multivariate quality control based on regression-adjusted variables', *Technometrics*, Vol. 33, No. 1, pp.61–75.

Hawkins, D.M. (1993a) 'Regression adjustment for variables in multivariate quality-control', *Journal of Quality Technology*, Vol. 25, No. 3, pp.170–182.

Hawkins, D.M. (1993b) 'Cumulative sum control charting: an underutilized SPC tool', *Quality Engineering*, Vol. 5, No. 3, pp.463–477.

Haworth, D.A. (1996) 'Regression control charts to manage software maintenance', *Journal of Software Maintenance and Evolution-Research and Practice*, Vol. 8, No. 1, pp.35–48.

Jarrett, J.E. and Pan, X. (2009) 'Multivariate process control charts and their use in monitoring output quality: a perspective', *International Journal of Industrial and Systems Engineering*, Vol. 4, No. 5, pp.471–482.

Kalagonda, A.A. and Kulkarni, S.R. (2003) 'Diagnosis of multivariate control chart signal based on dummy variable regression technique', *Communications in Statistics-Theory and Methods,* Vol. 32, No. 8, pp.1665–1684.

Kim, S.B., Jitpitaklert, W., Park, S.K. and Hwang, S.J. (2012) 'Data mining model-based control charts for multivariate and autocorrelated processes', *Expert Systems with Applications*, Vol. 39, No. 2, pp.2073–2081.

Lowry, C.A. and Montgomery, D.C. (1995) 'A review of multivariate control charts', *IEE Transactions*, Vol. 27, No. 6, pp.800–810.

Lu, C.W. and Reynolds, M.R. (1999) 'Control charts for monitoring the mean and variance of autocorrelated processes', *Journal of Quality Technology*, Vol. 31, No. 3, pp.259–274.

Lucas, J. and Saccucci, M.S. (1990) 'Exponentially weighted moving average control schemes, properties and enhancements', *Technometrics*, Vol. 32, No. 1, pp.1–12.

Mandel, J. (1969) 'The regression control chart', *Journal of Quality Technology*, Vol. 1, No. 1, pp.1–6.

Montgomery, D.C. (1997) *Introduction to Statistical Quality Control*, 3rd ed., John Wiley & Sons Inc., New York.

Montgomery, D.C. (2009) *Statistical Quality Control: A Modern Introduction*, 6th ed., Hoboken, John Wiley & Sons Inc., New Jersey.

Omura, A.P. and Steffe, J.F. (2003) 'Mixer viscometry to characterize fluid foods with large particulates', *Journal of Food Process Engineering*, Vol. 26, No. 5, pp.435–445.

Pacella, M. and Semeraro, Q. (2007) 'Using recurrent neural networks to detect changes in autocorrelated processes for quality monitoring', *Computers & Industrial Engineering*, Vol. 52, No. 4, pp.502–520.

Shewhart, W.A. (1931) *Economic Control of Quality of Manufactured Product*, American Society for Quality Control, Milwaukee, Wisconsin.

Shu, L., Tsung, F. and Tsui, K.L. (2004) 'Run-length performance of regression control charts with estimated parameters', *Journal of Quality Technology*, Vol. 36, No. 3, pp.280–293.

Sulek, J.M., Marucheck, A. and Lind, M.R. (2006) 'Measuring performance in multi-stage service operations: an application of cause selecting control charts', *Journal of Operations Management*, Vol. 24, No. 5, pp.711–727.

Testik, M.C. (2005) 'Model inadequacy and residuals control charts for autocorrelated processes', *Quality and Reliability Engineering International*, Vol. 21, No. 2, pp.115–130.

Tsai, T.N. and Chen, L.H. (2009) 'Monitoring of the stencil printing process using a modified regression residual control chart: an empirical study', *International Journal of Industrial Engineering – Theory, Applications and Practice*, Vol. 16, No. 4, pp.248–259.

Utley, J.S. and May, J.G. (2008) 'Residual control charts and LAV regression', *39th Annual Meeting of the Decision Sciences Institute*, pp.3521–3526.

Utley, J.S. and May, J.G. (2009) 'Monitoring service quality with residual control charts', *Managing Service Quality*, Vol. 19, No. 2, pp.162–178.

Wade, M.R. and Woodall, W.H. (1993) 'A review and analysis of cause selecting control charts', *Journal of Quality Technology*, Vol. 25, No. 3, pp.161–170.

Winkel, P. and Zhang, N.F. (2004) 'Serial correlation of quality control data-on the use of proper control charts', *Scandinavian Journal of Clinical & Laboratory Investigation*, Vol. 64, No. 3, pp.195–204.

Yang, C.M. and Yang, S.F. (2006) 'Optimal control policy for dependent process steps with over-adjusted means and variances', *International Journal of Advanced Manufacturing Technology*, Vol. 29, Nos. 7–8, pp.758–765.

Yang, S.F. (2005) 'Dependent processes control for over adjusted process means', *International Journal of Advanced Manufacturing Technology*, Vol. 26, Nos. 1–2, pp.109–116.

Yang, S.F. and Chen, W.Y. (2009) 'Controlling over-adjusted process means and variances using VSI cause selecting control charts', *Expert Systems with Applications*, Vol. 36, No. 3, pp.7170–7182.

Yang, S.F. and Su, H.C. (2007) 'Adaptive sampling interval cause-selecting control charts', *International Journal of Advanced Manufacturing Technology*, Vol. 31, Nos. 11–12, pp.1169–1180.

Yang, S.F. and Yang, C.M. (2005) 'Effects of imprecise measurement on the two dependent processes control for the autocorrelated observations', *International Journal of Advanced Manufacturing Technology*, Vol. 26, No. 5, pp.623–630.

Yu, J.B. and Liu, J.P. (2011) 'LRProb control chart based on logistic regression for monitoring mean shifts of auto-correlated manufacturing processes', *International Journal of Production Research*, Vol. 49, No. 8, pp.2301–2326.

Zhang, G.X. (1984) 'A new type of control charts and theory of diagnosis with control charts', *World Quality Congress Transactions*, American Society for Quality Control, Milwaukee, WI, pp.175–185.

Zhang, N.F. (2000) 'Statistical control charts for monitoring the mean of a stationary process', *Journal of Statistical Computation and Simulation*, Vol. 66, No. 3, pp.249–258.

Abbreviations and notation

The abbreviations and notations used in this paper are as follows:

ξ	Constant of AR(1) process
ϕ	Autoregressive coefficient
ε	Random error term
σ_e	Standard deviation of residuals
d	Trend slope of trend AR(1) process
Z_t	Shifted trend AR(1) process variable
δ_μ	Magnitude of upward mean shift
e_t	Residual (difference between expected and observed values of Z_t)
N	Sample size
$\hat{\beta}_0$	Intercept in a simple linear regression model
$\hat{\beta}_1$	Slope in a simple linear regression model
α	Smoothing constant
σ	Standard deviation of a sample from trend AR(1) process
$\hat{M}_o^t$	Estimated target value for process mean at time t
k	Reference value (allowance, or the slack value)
$\bar{C}^+$	Mean of deviation above $\hat{M}_o^t$
$\bar{C}^-$	Mean of deviations below $\hat{M}_o^t$
UCL_t	Upper control limit for proposed chart at time t
LCL_t	Lower control limit for proposed chart at time t