Special Issue Article

Quality and Reliability Engineering International

(wileyonlinelibrary.com) DOI: 10.1002/qre.2032

Published online in Wiley Online Library

A Multivariate Control Chart for Autocorrelated Tool Wear Processes

Keith Harris,[a*†] Kostas Triantafyllopoulos,[a] Eleanor Stillman[a] and Thomas McLeay[b]

Full automation of metal cutting processes has been a long held goal of the manufacturing industry. One key obstacle to achieving this ambition has been the inability to monitor completely the condition of the cutting tool in real time, as premature tool breakage and heavy tool wear can result in substantial costs through damage to the machinery and increasing the risk of non-conforming items that have to be scrapped or reworked. Instead, the condition of the tool has to be indirectly monitored using modern sensor technology that measures the acoustic emission, sound, spindle power and vibration of the tool during a cut. An online monitoring procedure for such data is proposed. Firstly, the standard deviation is extracted from each sensor signal to summarise the state of the tool after each cut. Secondly, a multivariate autoregressive state space model is specified for estimating the joint effects and cross-correlation of the sensor variables in Phase I. Then we apply a distribution-free monitoring scheme to the model residuals in Phase II, based on binomial type statistics. The proposed methodology is illustrated using a case study of titanium alloy milling (a machining process used in the manufacture of aircraft landing gears) from the Advanced Manufacturing Research Centre in Sheffield, UK, and is demonstrated to outperform alternative residual control charts in this application. **© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.**

Keywords: statistical process control (SPC); sensor data; tool condition monitoring; nonparametric control charts; multivariate autoregressive state space models

1. Introduction

There is a global drive towards increased productivity in the advanced manufacturing sector in order to meet the growing demands of lower cost targets, increased volume and increased process capability. In particular, metal machining remains the most important of all basic manufacturing processes. This is demonstrated by the fact that in 2015 the UK aerospace industry alone had a turnover of £29.2bn,[1] and machining processes are believed to have contributed to at least 30% of the products made in this trade. Productivity could be substantially improved via the automation of decision-making tasks previously taken, usually very conservatively, by the operator of the machine. This is now theoretically possible because of the development of modern sensor technology that enables the automatic collection of data related to the performance of the metal machining process. However, the large amounts of potentially noisy data generated by such sensors pose substantial challenges for process monitoring and control, and thus, new statistical algorithms need to be formulated in order to achieve the goal of automation.

In this paper, we focus on one particularly important limitation to machining productivity, tool wear, and develop a statistical process control (SPC) technique that could potentially be used to automate the decision as to when to replace the tool. This problem is challenging for several reasons. Firstly, in order to monitor in detail the condition of the tool, multiple sensors can be employed, so multivariate methods will be needed. Secondly, the machining process is continuously observed in real time for the duration of each cut of the metal workpiece, which can yield many thousands or even millions of observations per cut, because of the high-sampling frequency of the sensors. In our approach, important features correlated with tool wear are extracted from the sensor signals and are monitored using multivariate control charts. Finally, the extracted multivariate time series of sensor features is non-stationary and exhibits local variation, which is not attributed to out-of-control behaviour. In order to control such a process, we fit a multivariate autoregressive state space model to the observed data in Phase I, the period where the tool can be safely assumed to be far from worn out, and in Phase II, we apply a novel distribution-free monitoring scheme to the fitted model's residuals.

[a] *School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, UK*
[b] *The University of Sheffield Advanced Manufacturing Research Centre (AMRC) with Boeing, Advanced Manufacturing Park, Wallis Way, Catcliffe, Rotherham, S60 5TZ, UK*
**Correspondence to: Keith Harris, School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield, S3 7RH, UK.*
[†] *E-mail: K.J.Harris@sheffield.ac.uk*
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The remainder of this paper is organised as follows: Section 2 discusses our specific case study of interest, tool condition monitoring for titanium alloy milling. Section 3 explains our monitoring procedure based on fitting a multivariate autoregressive state space model in Phase I and applying binomial-type nonparametric control charts to the model residuals in Phase II. Section 4 includes an estimation of the average run length (ARL) for this chart, obtained via simulation for various mean-shift scenarios. Section 5 illustrates our methodology using two different types of sensor data from a tool wear experiment. Finally, Section 6 provides some conclusions and avenues for further research.

2. Case study

Our case study involves tool condition monitoring for metal cutting processes. In manufacturing, it is important to monitor tool wear, as tool breakages can result in unscheduled machine downtime in a factory, whilst excessive wear can cause poor quality and the

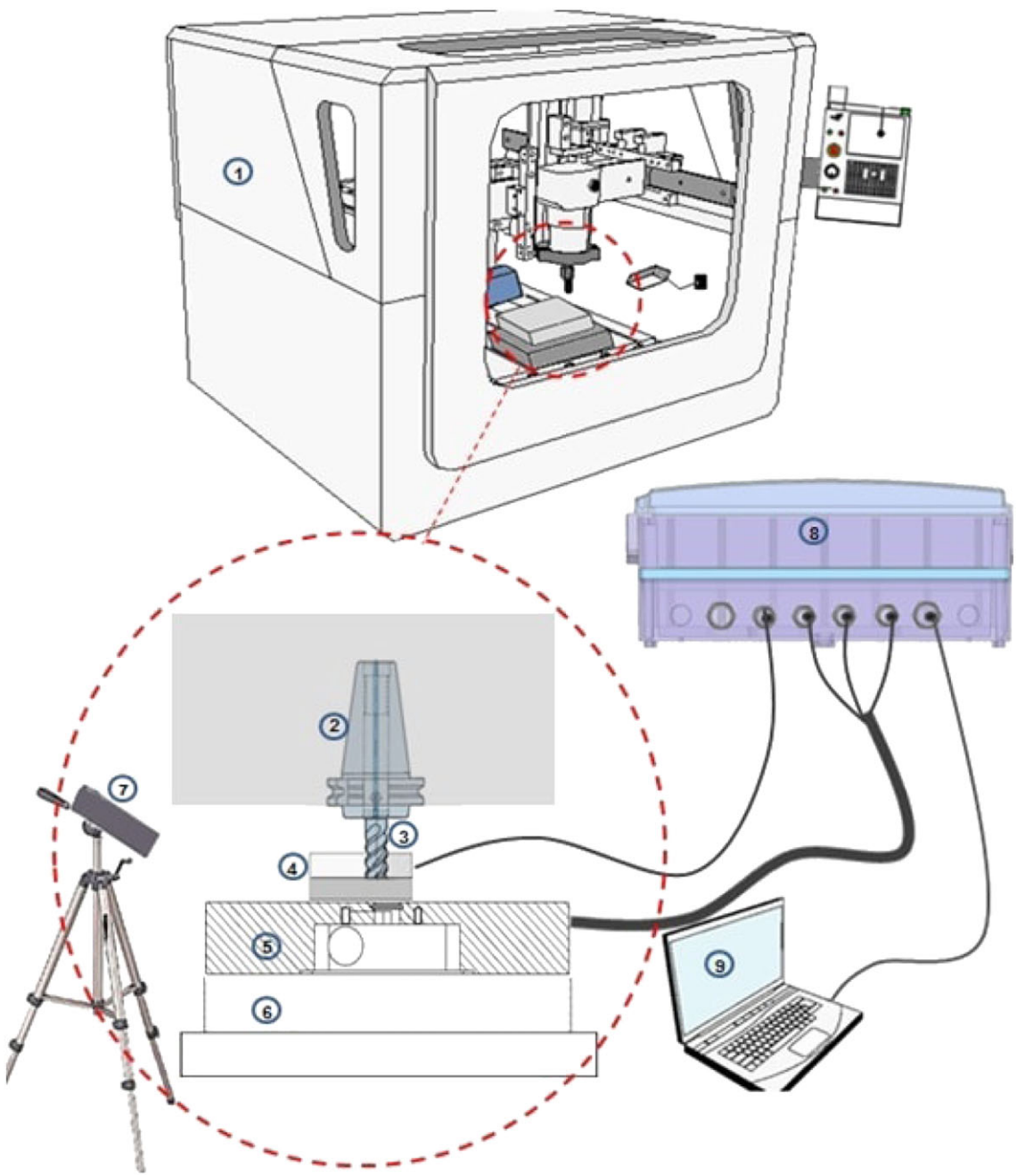

Figure 1. Experimental set-up. Labels correspond to: (1) Mori Seiki NMV8000 milling machine; (2) Sandvik hydraulic collet; (3) an end mill tool; (4) the workpiece: a titanium alloy plate; (5) a sensor plate containing a triaxial accelerometer, a high frequency single axis accelerometer oriented along the *z*-axis and an acoustic emission sensor; (6) Kistler dynamometer plate; (7) thermal camera vision acquisition system; (8) data acquisition box; and (9) laptop.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

scrapping of the finished part for failing to meet specifications, both of which can result in significant financial loss.[2] Clearly, there is a trade-off, as changing tools too soon would drive up tooling costs unnecessarily.

The specific machining operation of interest in our case study is 'milling' where a metal workpiece is fed past a rotating cylindrical tool with multiple cutting edges, which are called flutes or teeth.[3] In our study, the material of interest is Ti-6Al-4V, a titanium alloy consisting of roughly 90% titanium, 6% aluminium and 4% vanadium. Titanium alloy milling is of commercial interest as this process is used to produce high-strength landing gears for aircraft.

Tool wear is typically monitored indirectly by using sensors to measure, continuously, process variables like acoustic emission, sound, vibration and spindle power that are influenced by the state of the cutting tool. A time series of observations is thus obtained for each sensor during each cut of the workpiece. Because of the high frequency sampling rate, these time series (or signals) usually have many thousands or even millions of observations. An illustration of the experimental set-up used in our case study is shown in Figure 1. Notice that a sensor plate containing accelerometers and an acoustic emission sensor is clamped onto the workpiece being cut whilst also being connected to a data acquisition box that stores the sensor information. The signals detected by these sensors are then processed to extract features (the mean, standard deviation, root mean square, skewness and kurtosis) that might be correlated with tool wear. It is common in the machining literature then to develop decision-making support systems from these extracted features using machine learning techniques such as neural networks, fuzzy logic and genetic algorithms.[4] It is possible to measure tool wear directly by using a microscope to take images of the tool (Figure 2), but this would be impractical in a production setting. Moreover, automated quantification of tool wear from such an image would also be very challenging.

In our tool wear experiment, we extracted the mean, standard deviation, root mean square, variance, skewness and kurtosis of each of our eight sensor signals (one acoustic emission sensor, two microphones, a high frequency single axis accelerometer oriented along the *z*-axis, a triaxial accelerometer and a spindle power sensor) after each cut. The standard deviation data is shown in Figures 3 and 4. Many factors might be associated with tool condition, but in this study, we focus our attention on the correspondence between

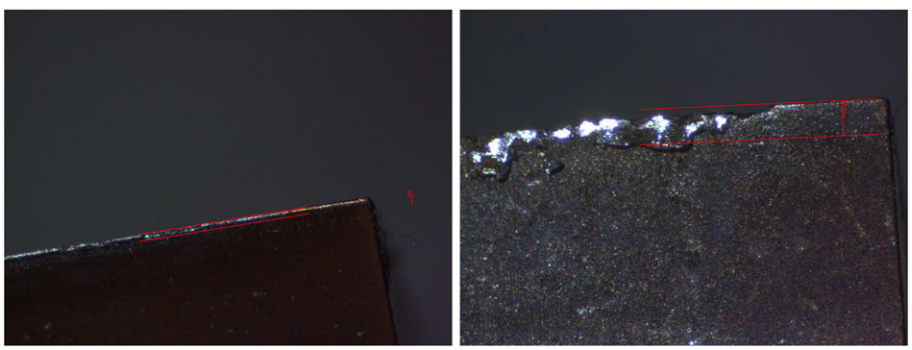

Figure 2. An image from a tool with neligible wear (left hand side) and from a worn out tool (right hand side). The images are on different scales, the left image showing around 2–3 mm of the flute length and the right image showing about 6 mm.

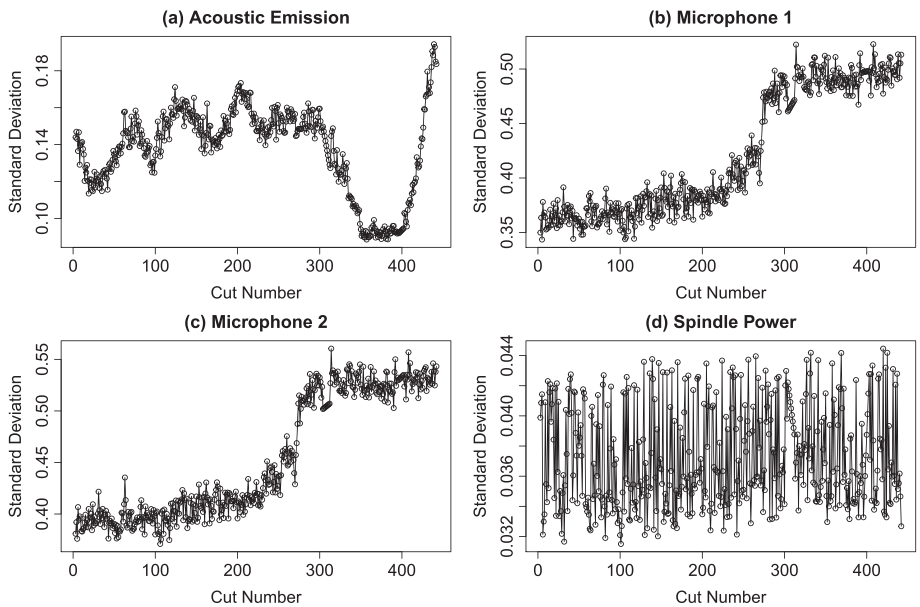

Figure 3. The standard deviation over time of the acoustic emission sensor, (a), two microphones, (b) and (c), and the spindle power sensor, (d).

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

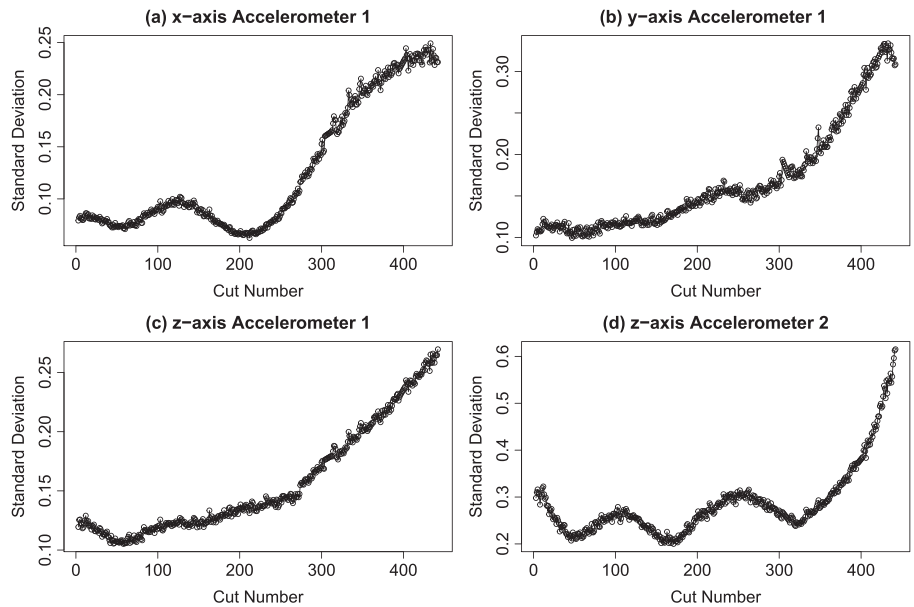

Figure 4. The standard deviation over time of the triaxial accelerometer signals, (a), (b) and (c), and the high frequency single axis accelerometer, (d).

increased variability in the microphone and accelerometer signals, as measured by their standard deviation and the tool's deterioration and eventual failure. This decision was motivated by the particularly strong co-evolution between these sensors' signal variability and tool wear; other features did not exhibit such a clear association with tool degradation.

As we see from Figures 3 and 4, there are trends or cycles in the sensor signals' standard deviation, which is expected and considered acceptable, as long as they are within a certain range. However, it is not known what the exact threshold is for unacceptable levels of standard deviation of each sensor, and this needs to be inferred from the data. The data also clearly exhibits local variation, which is not attributed to out-of-control behaviour. One needs to exclude such behaviour as a worrying sign and to include it in the expected dynamics of the system. Thus, unlike many applications of SPC, in the case of tool wear, we do not want to detect the process as being out of control as soon as possible, because we want to maximise tool usage before premature replacement. For example, looking at the triaxial accelerometer data of Figure 4, we would not want to flag out of control before the 250*th* cut, because the longer term deterioration in standard deviation associated with substantial tool wear does not happen until after this point.

Thus far, relatively little work has been done on tool condition monitoring using SPC techniques. This can partially be explained by the assumptions of independence and stability of the traditional Shewhart control chart being violated for tool wear data and more sophisticated techniques needing to be developed. Xie *et al.* [5] used double exponential smoothing to forecast a univariate tool wear process and applied a standard three-sigma control chart to the residuals, which was supplemented by an additional line that specified the maximum tool wear allowed before the tool had to be replaced. Multivariate SPC approaches have received scant attention so far.

3. Monitoring procedure

It is well known that monitoring time series data frequently poses appreciable difficulties, generally requiring modification of standard procedures in order to accommodate non-stationary behaviour.[6–8] Knoth and Schmid [8] discuss two types of charts for time series data: modified charts and residual charts. The former chart is suitable for weakly stationary time series in the presence of autocorrelation (which is responsible for the adjustments needed as compared with independent and identically distributed data). The residual chart is primarily proposed for data that exhibits structural dynamics, for which the in-control state is difficult to define. Such data are described in two examples of Jiang *et al.*[9]

The data we consider in this paper exhibit clear non-stationarity (see, for example, Figures 3(b), 3(c) and 4(a)-(c)), for which the residuals can offer a practical guide to specifying deviations from the norm and hence to defining an in-control state. This requires a two-phase control scheme: in Phase I, a retrospective analysis is conducted to understand the process better, identify an appropriate model, estimate its parameters and establish control limits for the residuals; in Phase II, the process is monitored using the fitted model and the control limits established in Phase I.[10] It should be ensured that the chosen model fits the data in Phase I well, so that it provides a good representation of the process and hence any deviations of the data from it – in Phase II – will signify deviations from target in the process. Adopting this approach, we fit an appropriate multivariate time series model to the observed data in Phase I and then devise a control procedure based on the residuals from this fitted model.

Suppose that observations $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are collected over time and that $\{\mathbf{y}_t\}$ forms an S-dimensional multivariate time series, so that $\mathbf{y}_i$ and $\mathbf{y}_j$ are serially correlated, for $i \neq j$; in other words, $\{\mathbf{y}_t\}$ exhibits autocorrelation. The model we choose to fit in Phase I is a

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

linear state space model, that is,

$$\mathbf{y}_t = H\mathbf{z}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \Sigma), \tag{1}$$

$$\mathbf{z}_t = F\mathbf{z}_{t-1} + \boldsymbol{\zeta}_t, \quad \boldsymbol{\zeta}_t \sim N(\mathbf{0}, \Omega), \tag{2}$$

with initial state $\mathbf{z}_0 \sim N(\mathbf{m}_0, P_0)$. It should be noted that the error terms $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\zeta}_t$ are assumed to be individually and mutually independent. In Equations (1) and (2), H is the design matrix, F is the transition matrix, Σ is the observation covariance matrix and Ω is the transition covariance matrix, and they are all of dimensions $S \times S$. The topic of estimation of these components in state space models has attracted a host of publications, including approximate and simulation-based methods.[11–13] In our case, the model parameters $H, F, \Sigma, \Omega, \mathbf{m}_0$ and P_0 were estimated in Phase I by maximum likelihood. Using predictive error decomposition, the log-likelihood function based on a data frame $\mathcal{D}_n = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ is

$$\begin{aligned}
\ell(H, F, \Sigma, \Omega, \mathbf{m}_0, P_0; \mathcal{D}_n) &= \log p(\mathbf{y}_1, \ldots, \mathbf{y}_n \mid H, F, \Sigma, \Omega, \mathbf{m}_0, P_0) \\
&= \log \prod_{t=1}^{n} p(\mathbf{y}_t \mid \mathcal{D}_{t-1}, H, F, \Sigma, \Omega, \mathbf{m}_0, P_0) \\
&= -\frac{nS}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} |Q_t| - \frac{1}{2} \sum_{t=1}^{n} (\mathbf{y}_t - \mathbf{f}_t)^T Q_t^{-1} (\mathbf{y}_t - \mathbf{f}_t),
\end{aligned}$$

where $\mathbf{f}_t$ and Q_t are the mean vector and the covariance matrix of $\mathbf{y}_t$, given $\mathcal{D}_{t-1}$ and $H, F, \Sigma, \Omega, \mathbf{m}_0$ and P_0, which are calculated by the Kalman filter.[12] In the previous log-likelihood, $\mathbf{f}_t$ and Q_t are implicitly conditional on $H, F, \Sigma, \Omega, \mathbf{m}_0$ and P_0. The maximisation of this log-likelihood can be performed either by adopting the EM algorithm or by direct maximisation procedures.[11,13] In this paper, we adopt the latter using the `dse` package in `R`.[14–16]

Model (1)–(2) describes the dynamics of the process. It identifies a moving mean, which is what is expected locally (the norm), and deviations from it signify an out-of-control state. We calculate the residuals, which measure these deviations. A control strategy based on these residuals is now proposed. It should be noted that we define the residuals at each time point t in Phase II to be the difference in the one-step ahead forecasts made at time $t-1$ using the state space model fitted in Phase I (i.e. the model hyperparameters are fixed at the values estimated in Phase I) and the corresponding true value. We thus obtain a time series of residual vectors $\{\mathbf{e}_t\}$, where $\mathbf{e}_t = [e_{1,t}, \ldots, e_{S,t}]^T$ and $t = 1, \ldots, N$.

Our method differs from Pan and Jarrett's,[17] because we propose a distribution-free monitoring scheme based on binomial type-statistics to be applied to the set of residuals $\{\mathbf{e}_t\}$, rather than Hotelling's T^2 chart. The justification for using the Hotelling chart is that $\mathbf{e}_t$ is approximately independent and identically distributed as a multivariate normal, but as this is not exactly true, a distribution-free approach is preferred. Our distribution-free approach is similar to the one adopted by Bersimis and Triantafyllopoulos [18] for monitoring air pollution; an overview of the earlier research into nonparametric monitoring methods can be found in the papers by Chakraborti *et al.* [19,20] and the references therein. As we are primarily concerned about increased variability in the sensor signals, we convert the vector of residuals $\mathbf{e}_t$ into binary vectors:

$$u_{i,t} = \begin{cases} 1, & \text{if } e_{i,t} > 0, \\ 0, & \text{if } e_{i,t} \leqslant 0, \end{cases} \tag{3}$$

where $i = 1, \ldots, S$ and $\mathbf{u}_t = [u_{1,t}, \ldots, u_{S,t}]^T$. This construction dichotomises the individual residuals $e_{i,t}$ into two classes: one for positive residuals, which corresponds to increased sensor signal variability and hence increased tool wear, and the other with negative residuals, which corresponds to lower than expected sensor signal variability and suggests that tool wear has not significantly increased since the last cut.

For the overall control of the wear, we may use the statistic $C(r, t)$, which represents the number of residuals from time $t-r$ to time t that are greater than 0. Specifically, $C(r, t)$ is defined as the following cumulative sum:

$$C(r, t) = \sum_{i=1}^{S} \sum_{j=t-r}^{t} u_{i,j}. \tag{4}$$

This statistic is based on the idea that if the process is in-control, the residuals must be randomly distributed above and below zero, whilst if the process is out of control, which would correspond to the tool becoming worn out, then there would be an extreme number of positive residuals. This statistic under the null hypothesis that the process is in-control follows a binomial distribution with probability of success equal to 0.5 and number of trials equal to $S(r+1)$. For ease of implementation and comparison, we propose to use the normal approximation to the binomial distribution (because we advocate choosing r to be large enough for the Central Limit Theorem to apply), and thus, we define the statistic $C'(r, t) = (2C(r, t) - S(r+1))/\sqrt{S(r+1)}$, which follows the standard normal distribution $N(0, 1)$. Using this statistic, a classical rule is to issue an alarm when the observed value of $C'(r, t)$ is larger than 3.

If desired, our control scheme can be modified by adding a dead band where small positive residuals are also deemed to be of no concern. The incorporation of a dead band scheme in the control chart is motivated by similar dead band approaches in feedback

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

adjustment charts.[21] Whilst Box *et al.* [21] are using a dead band to trim out adjustments of low value (in particular when there are adjustment costs), our motivation is slightly different. We propose to implement the scheme at the individual residual level, by identifying those residuals, which exceed some threshold k. In this case, we convert the vector of residuals $\mathbf{e}_t$ into binary vectors as follows:

$$u_{i,t} = \begin{cases} 1, & \text{if } e_{i,t} > k, \\ 0, & \text{if } e_{i,t} \leqslant k, \end{cases} \tag{5}$$

where $k > 0$. For example, we could choose $k = 0.5\hat{\tau}$, where $\hat{\tau}$ is the estimated standard deviation of the residuals in Phase I. With a dead band scheme as mentioned earlier, residuals, which are positive but relatively small, will be coded as zeros and will not contribute to $C(r,t)$; only those residuals that are deemed as 'significant' (or having the potential to indicate aberrant process behaviour) enter $C(r,t)$. As a result, one would expect to issue fewer signals using an appropriate dead band threshold, and this would lead to more flexible and robust control. For example, the chart can be made suitable for detecting small shifts in the residuals when k is low, or to focus only on larger shifts when k is larger. Thus, the dead band scheme provides versatility in the sense that it can allow monitoring of small or large shifts depending on the requirements of the specific application. The value of k should be chosen according to process specifications and model performance in Phase I.

For the purposes of comparison, we will also apply to the residuals two other standard multivariate control charts, the Hotelling T^2 chart and the multivariate exponentially weighted moving average (MEWMA) chart.[22,23] For the MEWMA chart, a value for the smoothing parameter λ has to be chosen. In general, small values of λ are used to detect small shifts in the process mean quickly, whilst larger values of λ are used for fast detection of large shifts.[24] We chose to set $\lambda = 0.2$ as a compromise between these two extremes. The limits for the MEWMA chart were obtained using the `spc` package in `R`.[25]

4. Simulation study

This section investigates the ARL performance of our proposed nonparametric control chart and compares its performance with the Hotelling T^2 and MEWMA charts under the assumption that the residuals are independent and identically distributed as a multivariate normal in Phase I. For the sake of simplicity, we consider only the bivariate case ($S = 2$).

A Monte Carlo simulation was conducted to obtain the ARL values for each chart for various mean shifts in the residuals. It was assumed that the process is in the in-control state for 200 observations and that the in-control process mean vector and covariance matrix were $\boldsymbol{\mu}_0 = 0$ and $\Sigma = I$; therefore, all initial values of the data were generated from the $N(0, I)$ distribution.

For comparison purposes, we considered seven different shifts in the mean vector $\boldsymbol{\mu}_0$, five immediate shifts of various sizes and two where the mean vector slowly drifts away from its starting position. It should be noted that the covariance matrix remained unchanged from that assumed in the in-control state. The size of the mean shift is determined by the distance of the out-of-control-mean, $\boldsymbol{\mu}_1$, from the in-control mean $\boldsymbol{\mu}_0$, and can be measured by the non-centrality parameter $D = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^T \Sigma^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)}$.[26] For the immediate shifts, we considered scenarios where $D = 0, 0.5, 1, 2$ and 3. We also investigated two other scenarios of more relevance to our practical examples of Section 5 in which the residuals in Phase II drift upwards away from zero (see, for example, Figure 13): one where both components of the mean increased gradually and the size of the shift was 0.5 after 50 observations, and another where only the first component increased, but more rapidly, resulting in a shift size of 0.5 after 40 observations. 5,000 observations were simulated from the out-of-control state, i.e., from $N(\boldsymbol{\mu}_1, I)$. For each scenario, 10 000 data sets were simulated to calculate the ARLs of the three methods.

To aid comparability, the control limits of both the Hotelling and MEWMA charts were tuned so that their in-control ARLs, ARL_0, were very close to 200. Similarly, the length of the window in the nonparametric chart, $r + 1$, was chosen to be 13, to achieve an in-control ARL slightly above 200. Because of the discrete nature of the statistic $C(r,t)$, it is impossible to achieve an in-control ARL of exactly 200 for small or moderate r.

The results of the simulation study are shown in Table I. We see that the nonparametric chart has the lowest out-of-control ARL for two of the six out-of-control scenarios, the immediate shift of size 0.5 and the gradual shift where both components drift off target. The MEWMA chart performs slightly better than the nonparametric chart when the immediate shift is of size 1 and the gradual

Table I. ARLs of the nonparametric, Hotelling and MEWMA ($\lambda = 0.2$) charts for various mean shifts.

$\boldsymbol{\mu}_1$	D	Nonparametric	Hotelling	MEWMA ($\lambda = 0.2$)
$(0,0)^T$	0	236.08	199.09	200.09
$(0.4,0.3)^T$	0.5	26.77	118.57	37.79
$(0.6,0.8)^T$	1	10.15	42.59	9.62
$(0,2)^T$	2	11.79	7.16	2.90
$(\sqrt{1.8}, \sqrt{7.2})^T$	3	4.49	2.19	1.63
$(0.008i, 0.006i)^T$	0.5 for $i = 50$	43.47	78.70	51.98
$(0.0125i, 0)^T$	0.5 for $i = 40$	48.66	69.97	45.26

Here, $\boldsymbol{\mu}_1$ denotes the out-of-control-mean vector, D is the non-centrality parameter and $i = 1, \ldots, 5000$ is the observation number from the out-of-control state.
ARL, average run length; MEWMA, multivariate exponentially weighted moving average.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

shift where only one of the two components increases. In all four of these scenarios, the Hotelling chart performs poorly. The nonparametric control chart is clearly outperformed by both the Hotelling and MEWMA charts in the two cases where there is a large immediate shift ($D = 2$ and 3). This is unsurprising, as the cumulative nature of the nonparametric statistic $C(r, t)$ limits quick detection. Indeed, the earliest it can signal out-of-control is after four observations and only when all eight residuals are positive. However, it should be stressed that the nonparametric chart performs best, or competitively with the MEWMA chart, for the scenarios that most closely resemble our practical examples where the residuals gradually drift off target or there is a small shift that warns of a bigger one ahead (Section 5).

5. Application

5.1. Microphone data

We first apply our methodology to the microphone data (Figure 3 (b) and (c)); we observe that the data is strongly autocorrelated, which is obvious from the local variation of the data and could be explored formally by plotting the sample autocorrelation functions (not shown). Indeed, these plots indicate that the data from both microphones are non-stationary time-series, with the level (or time-windowed mean) gradually growing for about the first 272 cuts, and then there is a sudden shift upwards over the next five cuts, at which point the data stabilises to an approximately constant (or very slowly increasing) level. For this data set, we consider Phase I to be up to time $t = 200$, during which the time series model of Equations (1) and (2) is estimated, and Phase II is from time $t = 201 - 425$. The decision to establish the in-control situation as the first 200 cuts was based on the expert opinion of our engineering collaborators, who deemed that this was an appropriate period over which to consider the tool as still being in good condition.

The state space model (1)–(2) is used with $S = 2$, for the bivariate observation vector $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})^T$ or

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{pmatrix} \begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},$$
$$\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} = \begin{pmatrix} f_{1,1} & f_{1,2} \\ f_{2,1} & f_{2,2} \end{pmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \zeta_{1,t} \\ \zeta_{2,t} \end{pmatrix},$$

where $h_{i,j}$ is the ij-th element of H, $f_{i,j}$ is the ij-th element of F, $z_t = (z_{1,t}, z_{2,t})^T$ and the error terms are defined likewise. We fit this model in Phase I, and we estimate the model components H, F as well as the error covariance matrices $\Sigma = (\sigma_{i,j})$, $\Omega = (\omega_{i,j})$ and $\boldsymbol{m}_0, P_0$ (see also the discussion in Section 3); later we show the estimated values of $h_{i,j}$, $f_{i,j}$, $\sigma_{i,j}$ and $\omega_{i,j}$, for $i, j = 1, 2$:

$$\hat{h}_{1,1} = 0.834, \quad \hat{h}_{1,2} = 0.033, \quad \hat{h}_{2,1} = 0.010, \quad \hat{h}_{2,2} = 0.996$$
$$\hat{f}_{1,1} = 1.225, \quad \hat{f}_{1,2} = -0.243, \quad \hat{f}_{2,1} = 1.135, \quad \hat{f}_{2,2} = -0.223$$
$$\hat{\sigma}_{1,1} = 0.020, \quad \hat{\sigma}_{1,2} = -0.023, \quad \hat{\sigma}_{2,2} = 0.996,$$
$$\hat{\omega}_{1,1} = 0.119, \quad \hat{\omega}_{1,2} = -0.093, \quad \hat{\omega}_{2,2} = 0.110$$

Some comments are in order. The eigenvalues of $\hat{H}$ are 0.998 and 0.832, and those of $\hat{F}$ are 0.999 and 0.003. Because one eigenvalue of $\hat{H}$ and one eigenvalue of $\hat{F}$ are so close to 1, we conclude that the model is very nearly non-stationary. This is confirmed by the plots of Figure 3 (panels (b) and (c)) where for Phase I (time points 1–200); we see a relatively local departure from stationarity, in particular as compared with the variation in Phase II. The matrix components $\hat{H}$ and $\hat{F}$ pick this up well, and the state space model fitted in Phase I resembles a multivariate local level model.[13] The estimated observation correlation coefficient (in Σ) is quite close to 0 (-0.163), but

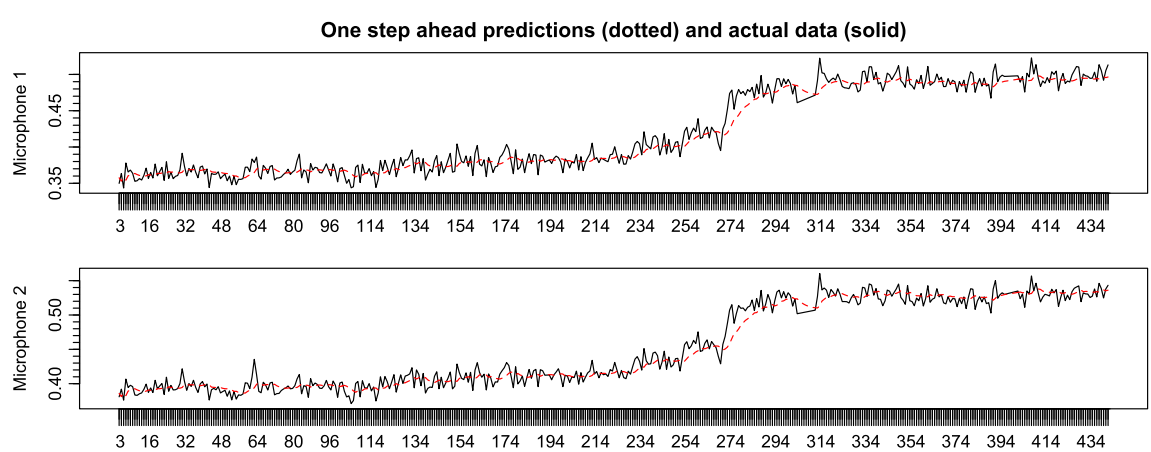

Figure 5. The standard deviations over time of microphones 1 and 2 are plotted in the upper and lower panels, respectively. The red dashed lines represent the one-step ahead forecasts.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

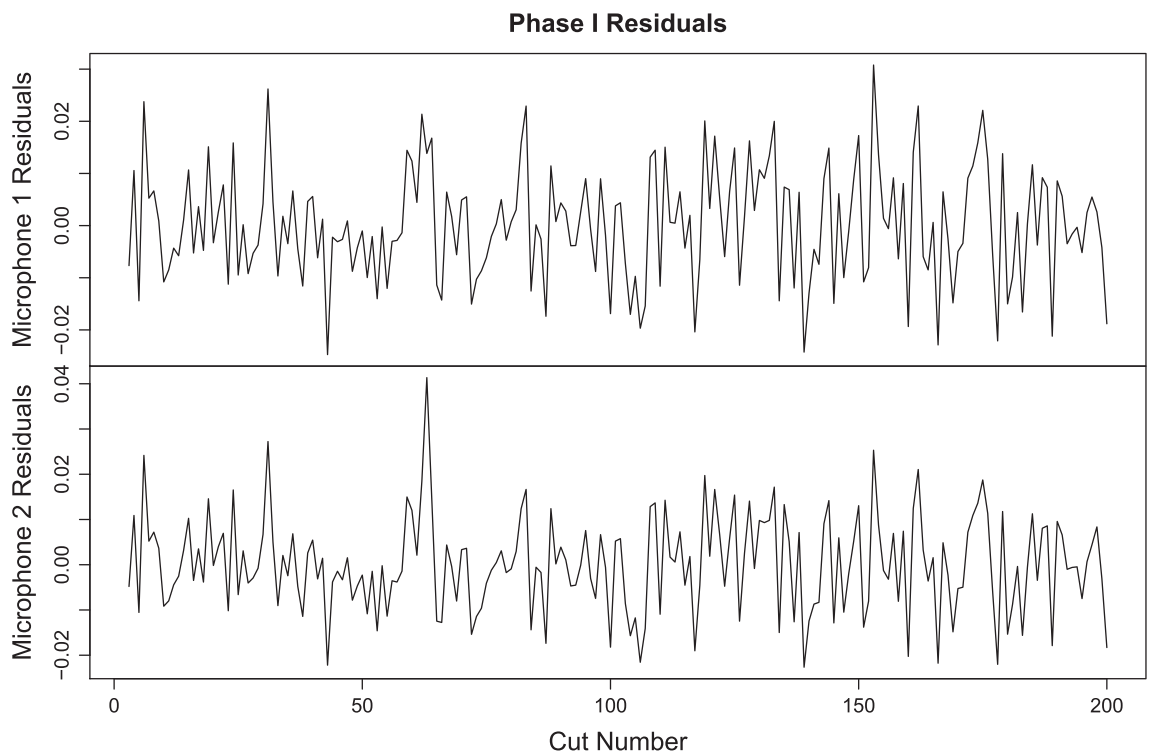

Figure 6. Time plots of the residuals in Phase I for the microphone data.

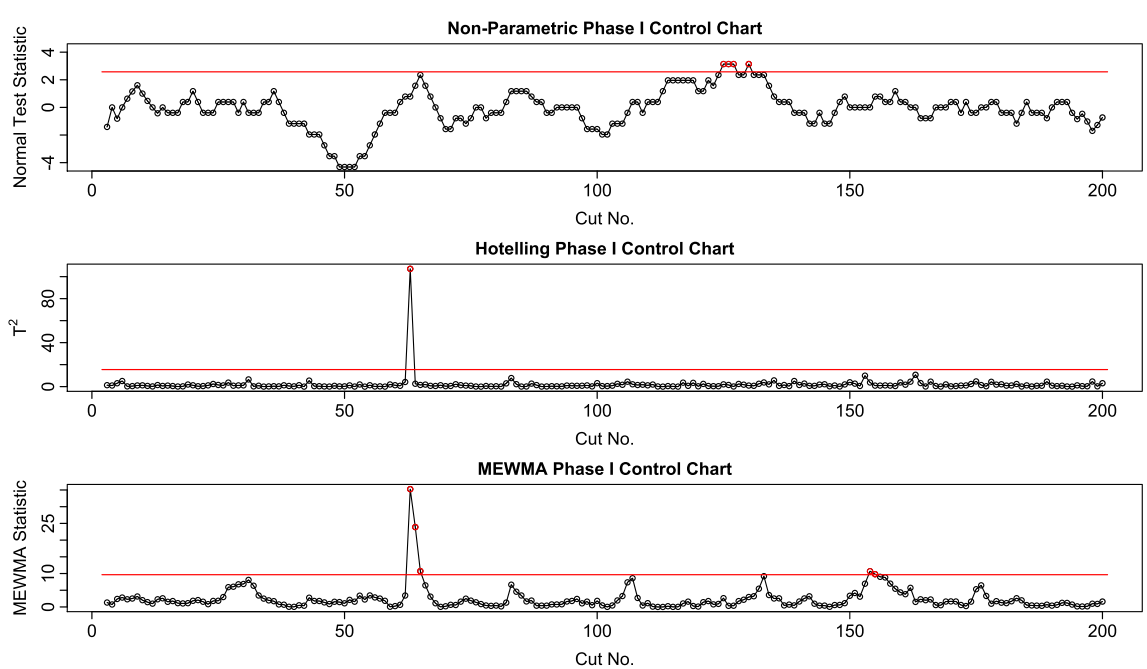

Figure 7. The nonparametric, Hotelling and multivariate exponentially weighted moving average (MEWMA) Phase I control charts for the microphone data.

the state correlation coefficient is −0.813. This implies a strong correlation between the states, whilst given the states the correlation of the data is low. In our model, most of the correlation dependence of the observations is captured by the dynamic structure of the state vectors.

The one-step ahead forecasts are shown overlaid on the data in Figure 5, and we see that the model does a good job of capturing the gradually increasing level of the data in Phase I. The residuals in Phase I are shown in Figure 6, and we see no obvious pattern in the two plots. The autocorrelation functions of the residuals (not shown) indicate that there is no evidence against the assumption that the residuals are independently distributed. Figure 7 shows the three control charts in Phase I, our nonparametric chart with $r = 12$,

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

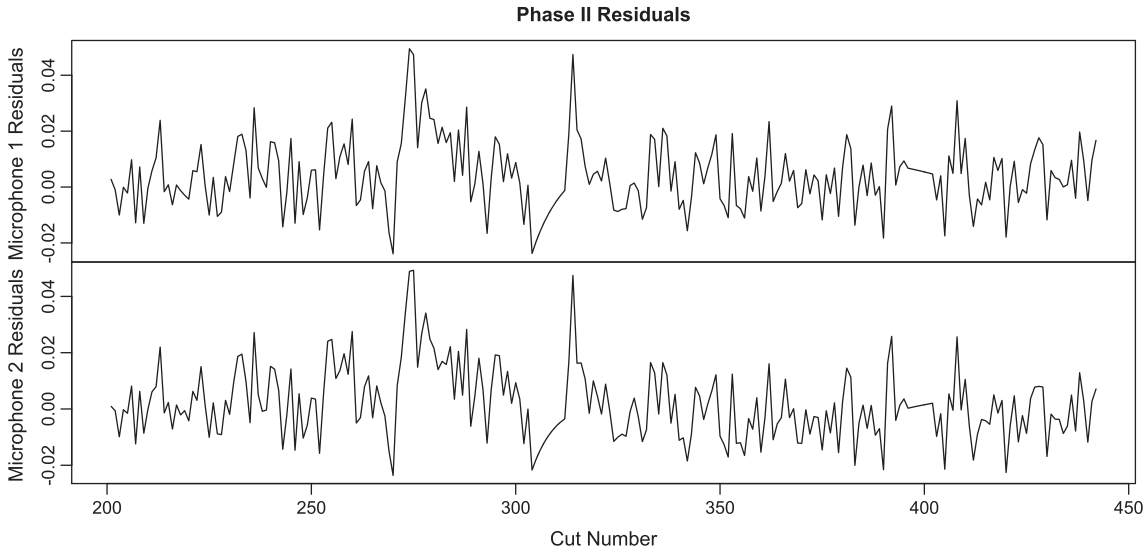

Figure 8. Time plots of the residuals in Phase II for the microphone data.

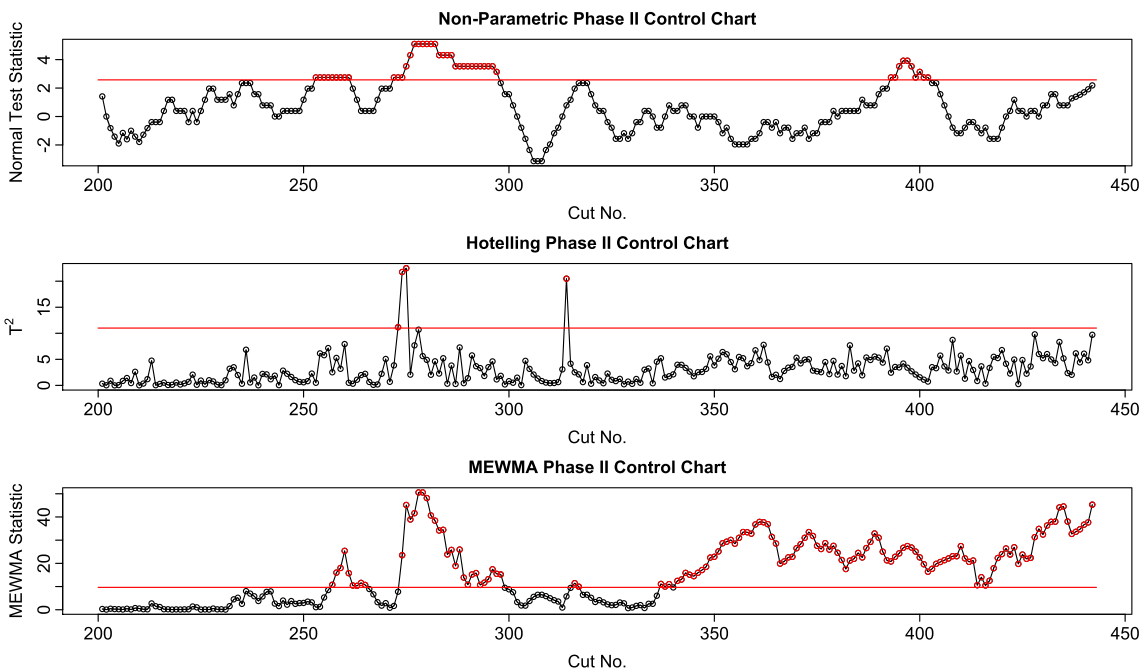

Figure 9. The nonparametric, Hotelling and multivariate exponentially weighted moving average (MEWMA) Phase II control charts for the microphone data.

the Hotelling chart and the MEWMA chart with $\lambda = 0.2$. The Hotelling and MEWMA charts were obtained using the `MSQC` package in `R`.[27] The nonparametric chart signals out of control 4 times at the 0.5% significance level. However, all these observations fall just above the control limit, suggesting that the assumption that there is an equal split of positive and negative residuals over time is not too badly violated. The Hotelling chart flags one observation as being clearly out of control, whilst the MEWMA chart flags five observations, although three of those points are just over the limits. This suggests that the model fits the Phase I data well apart from one or two outliers.

Proceeding now into a discussion of Phase II, we see from Figure 5 that the one-step ahead forecasts of the model fitted in Phase I are very slow to adapt to the sudden upward shift in the level of both time series at time $t = 273$, leading to a long run of positive residuals, the first few of which are very large. This can also be seen from the plots of the residuals in Phase II shown in Figure 8. The control charts in Phase II are shown in Figure 9, whilst the times of the first out-of-control signals for different significance levels are shown in Table II. We see that, for this example, the Hotelling chart performs well, first signalling that the process is out of control at the 0.5% significance level at time $t = 273$, coinciding with the sudden shift in the level of the data. The MEWMA chart arguably performs even better, as it signals at time $t = 257$, detecting a smaller upwards shift before the big shift takes place. Our nonparametric chart first

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

Table II. Time of first out-of-control signal in Phase II (microphone data).

Chart	Significance	First out-of-control cut
Non-parametric	0.01	235
Hotelling T^2	0.01	273
MEWMA ($\lambda = 0.2$)	0.01	256
Non-parametric	0.005	253
Hotelling T^2	0.005	273
MEWMA ($\lambda = 0.2$)	0.005	257
Non-parametric	0.001	275
Hotelling T^2	0.001	274
MEWMA ($\lambda = 0.2$)	0.001	258

MEWMA, multivariate exponentially weighted moving average.

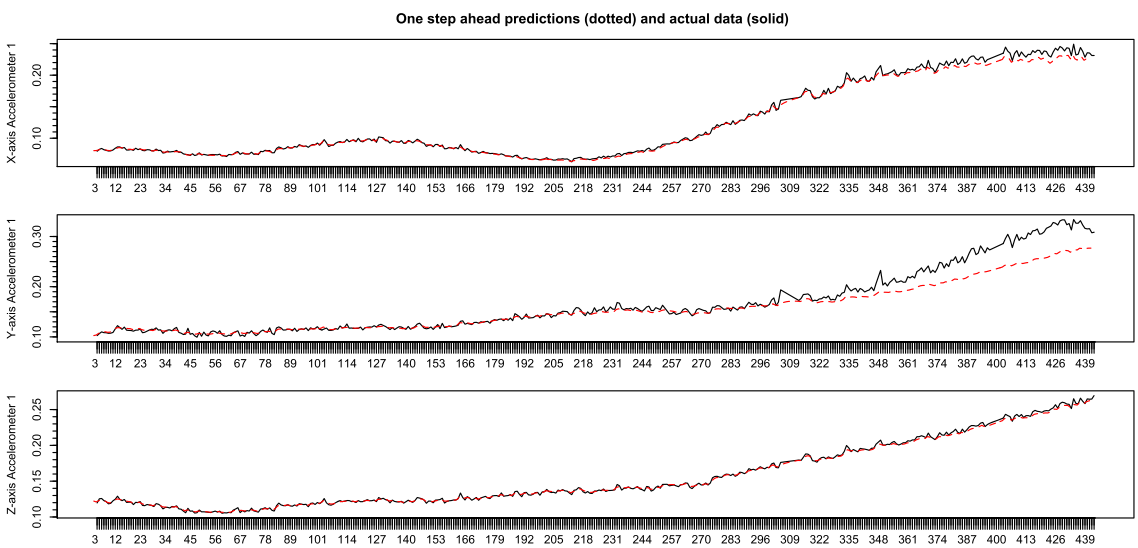

Figure 10. The standard deviations over time of the *x*-axis, *y*-axis and *z*-axis triaxial accelerometer are plotted in the upper, middle and lower panels, respectively. The red dashed lines represent the one-step ahead forecasts.

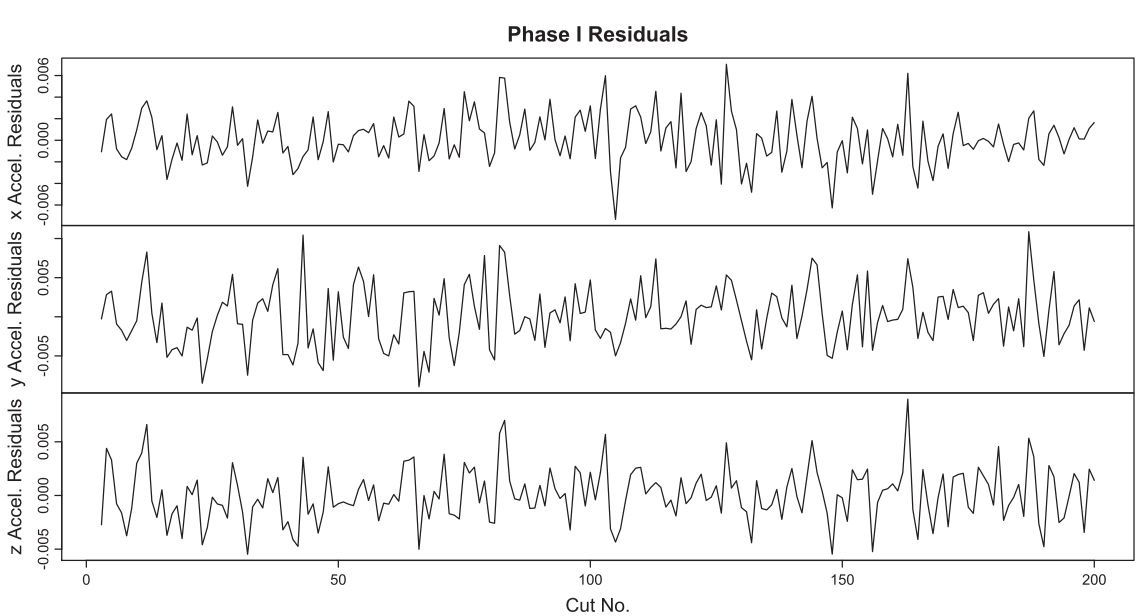

Figure 11. Time plots of the residuals in Phase I for the triaxial accelerometer data.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

signals out of control at the 0.5% significance level at time $t = 253$, thus providing an even earlier warning of the process becoming out of control. It is unsurprising that both the Hotelling and MEWMA charts also perform well in this case because both are good at immediately detecting sudden shifts in the mean vector of a process. However, we will see that both charts perform very poorly in the case of the triaxial accelerometer data of Figure 4, signalling out of control far too soon.

5.2. *Triaxial accelerometer data*

We now apply our methodology to the triaxial accelerometer data in Figure 4 (panels (a)–(c)). Once again, we see that the individual time series are strongly autocorrelated, but here, there is not a sudden shift in the level of these series. Let $y_{1,t}$ denote the x-axis accelerometer observation, $y_{2,t}$ the y-axis accelerometer observation and $y_{3,t}$ the z-axis accelerometer observation at time t; we then form the trivariate observation vector $\boldsymbol{y}_t = (y_{1,t}, y_{2,t}, y_{3,t})^T$. For the x-axis accelerometer data ($y_{1,t}$), we see slowly cyclical behaviour for

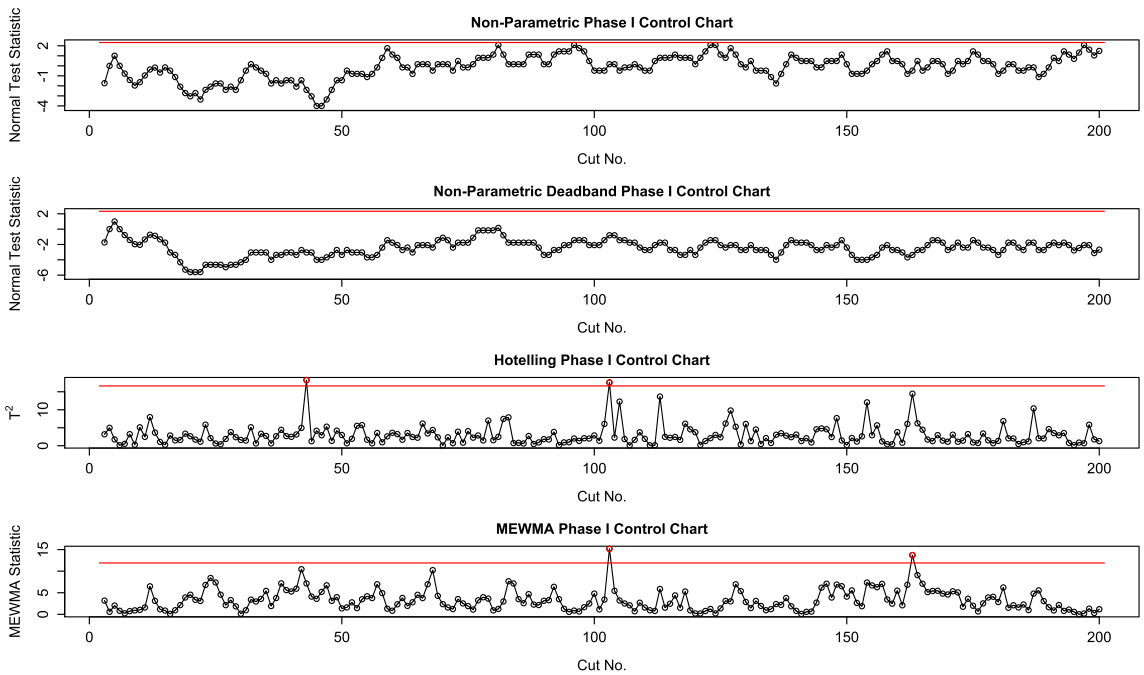

Figure 12. The nonparametric, nonparametric with dead band, Hotelling and multivariate exponentially weighted moving average (MEWMA) Phase I control charts for the triaxial accelerometer data.

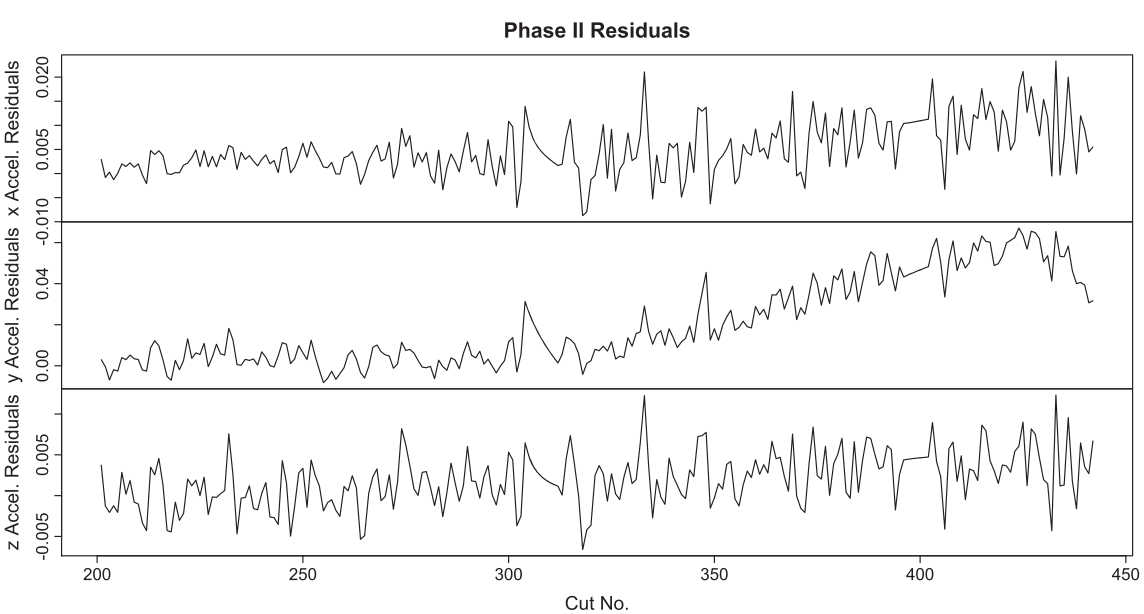

Figure 13. Time plots of the residuals in Phase II for the triaxial accelerometer data.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

the first 250 or so cuts, at which point the cycle breaks down, and we see persistent growth in the level of the data for the remainder of the experiment, whilst for the y and z-axis accelerometer data ($y_{2,t}$ and $y_{3,t}$), we see approximately linear growth for about the first 270 cuts, and then there is a sudden shift in the gradient of the data that persists for the next 140 or more cuts. We again consider Phase I to be up to time $t = 200$ and Phase II to be from time $t = 201 - 425$. We fit model (1)–(2), with $S = 3$. The matrices H, F, Σ, Ω and the initial state vector mean and covariance matrix m_0 and P_0 are estimated by maximum likelihood in Phase I as described in Section 3 and as previously demonstrated for the microphone data in Section 5.1. Here, for space economy, we do not show the values of the maximum likelihood estimates of these model components.

Like the microphone data, we see from the plot of the one-step ahead forecasts in Figure 10 that the linear state-space model does a good job of capturing the slowly adapting level of the triaxial accelerometer data in Phase I. Once again, there is no obvious pattern in the Phase I residuals, as seen by their plot in Figure 11. There is also no evidence against the assumption that the residuals are independently distributed, because all their sample autocorrelations fall within or very close to within the 95% confidence bounds (not shown). In this example, we show four control charts in Phase I (Figure 12): our nonparametric chart with $r = 12$, the Hotelling chart and the MEWMA chart with $\lambda = 0.2$ as before, and also the modified nonparametric chart with $r = 12$ and a dead band of $k = 0.5\hat{\tau}$. These control charts also suggest that the state-space model fits the Phase I data satisfactorily, as neither nonparametric chart signals that the process is out of control at any point, whilst the Hotelling and MEWMA charts only flag three points combined (out of a total of 200) as potentially being out-of-control at the 0.5% significance level, all of which fall just outside the control limits.

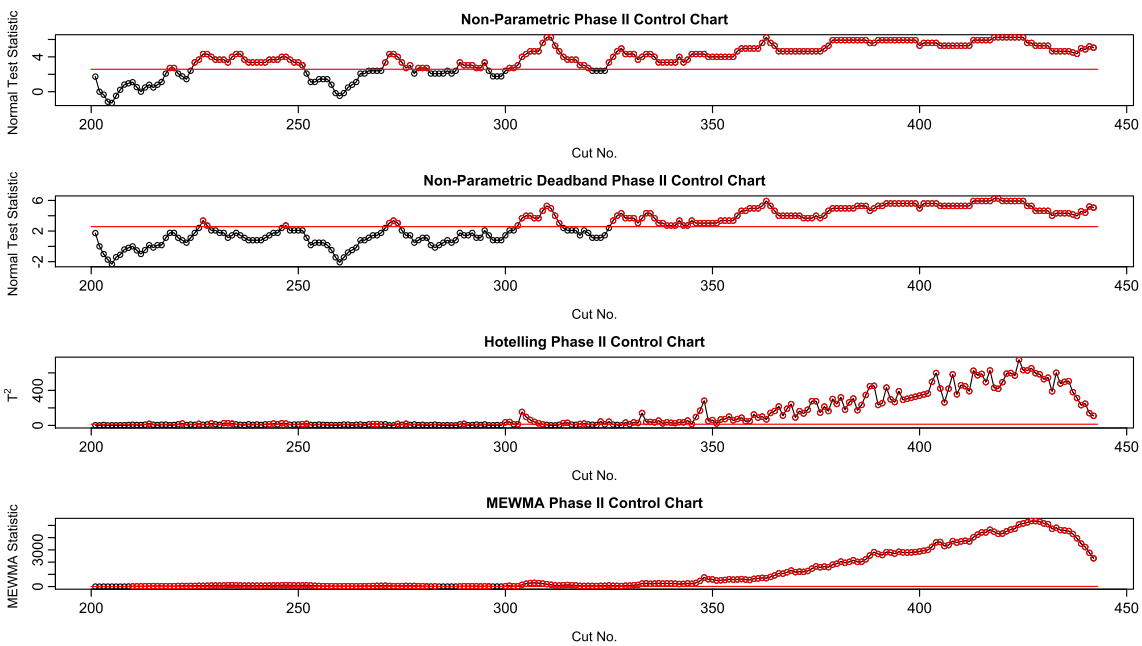

Figure 14. The nonparametric, nonparametric with dead band, Hotelling and multivariate exponentially weighted moving average (MEWMA) Phase II control charts for the triaxial accelerometer data.

Table III. Time of first out-of-control signal in Phase II (triaxial accelerometer data).

Chart	Significance	First out-of-control cut
Non-parametric	0.01	219
Non-parametric deadband	0.01	226
Hotelling T^2	0.01	214
MEWMA ($\lambda = 0.2$)	0.01	210
Non-parametric	0.005	219
Non-parametric deadband	0.005	227
Hotelling T^2	0.005	214
MEWMA ($\lambda = 0.2$)	0.005	210
Non-parametric	0.00025	226
Non-parametric deadband	0.00025	304
Hotelling T^2	0.00025	222
MEWMA ($\lambda = 0.2$)	0.00025	211

MEWMA, multivariate exponentially weighted moving average.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

We will now discuss our Phase II results. Firstly, we see from Figure 10 that after about time $t = 300$ the one-step ahead forecasts start becoming very poor, consistently under-predicting the *y*-axis accelerometer's standard deviation by a growing amount, and also a tendency to under-predict the *x* and *z* axes' standard deviation too. This leads to a clear positive trend in the residuals for the *y*-axis from time $t = 300$ to $t = 420$, as shown by the plots of the residuals in Phase II of Figure 13. We also see from Figure 13 that the variability of the *x*-axis residuals noticeably increases around time $t = 300$. The Phase II control charts are shown in Figure 14, whilst the times of the first out-of-control signals are presented in Table III. Here, both the MEWMA and Hotelling charts signal that the process is out of control at the 0.025% significance level very quickly; the former at time $t = 211$ and the latter at time $t = 222$. Thus, both these charts perform poorly, signalling out of control far too soon, as the process clearly does not become out of control until much later. Our nonparametric chart performs slightly better, first signalling out of control at the 0.025% significance level at time $t = 226$, but this also seems too soon. Arguably, in this case, the modified nonparametric chart with a dead band of $k = 0.5\hat{\tau}$ performs the best, as it does not signal out of control at the 0.025% significance level until time $t = 304$, by which point the process is clearly deteriorating.

6. Conclusion

In this paper, a method was proposed for indirectly monitoring the condition of a tool using acoustic emission, microphone, vibration and spindle power sensors. The high-frequency sampling rate of these sensors means that large amounts of data is amassed in the short space of time it takes to cut a workpiece, particularly acoustic emission data, which is typically sampled at a rate of over one million observations per second. From these sensor signals, features correlated with tool wear can be extracted and used to monitor the condition of the tool after each cut of the workpiece. In our case study of milling a titanium alloy, the standard deviation of each sensor signal was felt to be the most appropriate feature to monitor. The resulting multivariate time series is non-stationary, and it is thus important to develop a monitoring procedure that can distinguish between local variation in the data that is not attributable to out-of-control behaviour and longer term changes that are associated with tool wear.

A multivariate linear state space model is fitted in Phase I to describe the expected dynamics (i.e. the norm) of the process when it is in control. In Phase II, the differences between the one-step ahead forecasts from the state space model fitted in Phase I and their corresponding true values are calculated, as these residuals measure the deviations from the norm. A nonparametric control chart for monitoring these residuals based on binomial type-statistics is proposed and compared with applying standard multivariate charts (like Hotelling's T^2 and the MEWMA) to the residuals. Our simulation study demonstrated that the nonparametric chart performs best, or competitively with the MEWMA chart, for the scenarios that most closely resembled our practical examples where the residuals gradually drift off target or there is a small shift in the mean that warns of a bigger one ahead.The standard charts for residual monitoring performed well when applied to the microphone data, as there was a sudden shift in the level of this process, but performed very poorly when applied to the triaxial accelerometer data, signalling out of control far too soon, because the level of the process increases much more slowly. The proposed nonparametric control chart also performed well on the microphone data, whilst performing much better on the triaxial accelerometer data than the standard charts when it was modified by adding a dead band of $k = 0.5\hat{\tau}$.

An alternative nonparametric approach to the one proposed here would be to develop a control chart based on runs type rules, which is another possibility that may perform well on the triaxial accelerometer data and is an area of ongoing research. The next step would be to develop feedback adjustment mechanisms based on the sensor data for taking corrective action to prolong the life of the tool. For example, if the sound and vibration were too high, we might consider reducing the cutting speed or feed rate of the machining process. The end goal is to develop a methodological framework to enable machining processes to be fully automated and efficiently optimised.

Acknowledgements

This research was funded by the EPSRC grant EP/K031406/1 entitled 'Simulation Tools for Automated and Robust Manufacturing' and carried out in conjunction with The University of Sheffield Advanced Manufacturing Research Centre (AMRC). In particular, we would like to thank Sam Turner and Salvador Sumohano Verdeja of the AMRC for their useful discussions and comments in the process of doing this research.

References

1. UK Aerospace Outlook 2015 report, ADS Group Limited. (Available from: https://www.adsgroup.org.uk/uk-aerospace-outlook-2015/)[Last accessed on May 4th].
2. Karandikar J, McLeay T, Turner S, Schmitz T. Tool wear monitoring using naïve Bayes classifiers. *The International Journal of Advanced Manufacturing Technology* 2015; **77**:1613–1626.
3. Groover MP. Fundamentals of Modern Manufacturing: Materials, Processes, and Systems 5th ed. Wiley: Hoboken, New Jersey, 2012.
4. Tetia R, Jemielniak K, O'Donnell G, Dornfeld D. Advanced monitoring of machining operations. *CIRP Annals - Manufacturing Technology* 2010; **59**: 717–739.
5. Xie M, Goh TN, Wiklund H, Tang XY. Statistical monitoring and control of tool wear processes. *International Journal of Reliability, Quality & Safety Engineering* 2000; **7**:331–340.
6. Bagshaw M, Johnson RA. Sequential procedures for detecting parameter changes in a time-series model. *Journal of the American Statistical Association* 1977; **72**:593–597.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.

7. Alwan LC, Roberts HV. Time series modeling for statistical process control. *Journal of Business and Economic Statistics* 1988; **6**(1):87–95.
8. Knoth S, Schmid W. Control charts for time series: A review, in Frontiers in Statistical Quality Control 7. Physica-Verlag: Heidelberg, 2004; 210–236.
9. Jiang W, Tsui KL, Woodall WH. A new SPC monitoring method: the ARMA chart. *Technometrics* 2000; **42**:399–410.
10. Pan JN, Chen ST. Monitoring long-memory air quality data using ARFIMA model. *Environmetrics* 2008; **19**:209–219.
11. Shumway RH, Stoffer DS. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis* 1982; **3**: 253–264.
12. West M, Harrison J. Bayesian Forecasting and Dynamic Models 2nd ed. Springer: New York, 1997.
13. Shumway RH, Stoffer DS. Time Series Analysis and its Applications: With R Examples 3rd ed. Springer: New York, 2010.
14. Gilbert PD. Combining VAR estimation and state space model reduction for simple good predictions. *Journal of Forecasting* 1995; **14**(3):229–250.
15. Gilbert PD. Brief user's guide: Dynamic systems estimation, 2009. (Available from: https://cran.r-project.org/web/packages/dse/vignettes/Guide.pdf)[Last accessed on May 4th, 2016], originally published.
16. R Core Team. R: A language and environment for statistical computing, R Foundation for Statistical Computing: Vienna, Austria, 2015. (Available from: https://www.R-project.org/)[Last accessed on May 4th].
17. Pan X, Jarrett J. Applying state space to SPC: monitoring multivariate time series. *Journal of Applied Statistics* 2004; **31**(4):397–418.
18. Bersimis S, Triantafyllopoulos K. Dynamic non-parametric monitoring of air-pollution, University of Piraeus, 2016. Submitted for publication in *Quality and Reliability Engineering International*.
19. Chakraborti S, van der Laan P, Bakir ST. Nonparametric control charts: an overview and some results. *Journal of Quality Technology* 2001; **33**(3): 304–315.
20. Chakraborti S, van der Laan P, van de Wiel MA. A class of distribution-free control charts. *Journal of the Royal Statistical Society. Series C* 2004; **53**(3):443–462.
21. Box GEP, Luceno A, del Carmen Paniagua Quinones M. Statistical Control by Monitoring and Adjustment 2nd ed. Wiley: New York, 2009.
22. Lowry CA, Montgomery DC. A review of multivariate control charts. *IIE Transactions* 1995; **27**(6):800–810.
23. Bersimis S, Psarakis S, Panaretos J. Multivariate statistical process control charts: an overview. *Quality and Reliability Engineering International* 2007; **23**(5):517–543.
24. Khoo MBC. Increasing the sensitivity of multivariate EWMA control chart. *Quality Engineering* 2003; **16**(1):75–85.
25. Knoth S. spc: Statistical Process Control - collection of some useful functions, 2016. (Available from: https://cran.r-project.org/web/packages/spc/) R package version 0.5.3. Last accessed on May 4th, 2016, originally published.
26. Noorossana R, Vaghefi SJM. Effect of autocorrelation on performance of the MCUSUM control chart. *Quality and Reliability Engineering International* 2006; **22**:191–197.
27. Santos-Fernández E. Multivariate Statistical Quality Control using R. Springer: New York City, 2012. http://www.springer.com/gb/book/9781461454526.

Authors' biographies

Keith Harris received the BSc degree in mathematics and statistics from the University of York, York, UK, in 2002, and the MSc and PhD degrees in statistics from the University of Sheffield, Sheffield, UK, in 2003 and 2008, respectively. From 2008 to 2013, he was a PDRA at the University of Glasgow, first working on a project entitled "Classifiers in Medicine and Biology (Advancing Machine Learning Methodology for New Classes of Prediction Problems)" in the Department of Computing Science, and later on a Unilever funded project in the School of Engineering that focused on developing new statistical methods for the analysis of genomics data from microbial communities. He is currently a PDRA in the School of Mathematics and Statistics at the University of Sheffield working on an EPSRC funded project called "Simulation Tools for Automated and Robust Manufacturing."

Kostas Triantafyllopoulos is a Senior Lecturer at the School of Mathematics and Statistics of the University of Sheffield. He holds a PhD in Statistics from the University of Warwick. Prior to Sheffield, he has worked as a Research Associate at the University of Bristol and as a Lecturer at the University of Newcastle upon Tyne. His research interests include time series analysis and statistical process control. He has published widely and is involved in research grants including the Nuffield Foundation and Engineering and Physical Sciences Research Council (UK). He has wide teaching experience in statistics and is the supervisor of seven doctoral students and 30 masters students.

Eleanor Stillman took BSc and PhD degrees at Sheffield before taking up her staff position there. She is now a Senior Lecturer in Probability & Statistics and a Chartered Statistician. Her main research interests lie in the practical application of statistics to geology and materials science. A long term concern has been the modelling of particle size, with investigations into its relationship with sediment transport processes and strength of composite materials. Other projects include the design of resistant glazes and the modelling, monitoring and control of manufacturing processes. She is also active in statistical education and ethical review.

Thomas McLeay has 10 years of experience in manufacturing research. He spent 5 years developing machining processes for high-performance disc machining, now in production at the Rolls-Royce UK Discs site. In 2010, he set up the Process Monitoring & Control Research Group, a group focused on next generation manufacturing solutions that apply informatics, novel sensing and intelligent control to a range of manufacturing projects. He is now head of Core Research in the Advanced Manufacturing Research Centre's Machining Group at the University of Sheffield and is a principal- and co-investigator in IUK, EPSRC and HVM Catapult funded projects with a total project value above £10m.

© 2016 The Authors Quality and Reliability Engineering International Published by John Wiley & Sons Ltd.