CHAPTER 8

Statistical Process Control for Geometric Specifications

Bianca Maria Colosimo and Quirico Semeraro
Department of Mechanical Engineering, Politecnico di Milano, Milano, Italy

Massimo Pacella
Dipartimento di Ingegneria dell'Innovazione, University of Salento, Lecce, Italy

INTRODUCTION

In the area of statistical process control (SPC), approaches for *profile monitoring* proposed in the literature (Woodall et al. 2004; Woodall 2007) can be classified with reference to the modeling technique considered (linear or nonlinear regression, nonparametric regression or approaches for multivariate data reduction as the principal/independent component analysis) as well as to the type of application faced (i.e., calibration study, process signal, or geometric specification monitoring).

With reference to the type of application faced, most of the studies on profile monitoring deal with calibration studies (Stover and Brill 1998; Kang and Albin 2000; Kim et al. 2003; Mahmoud and Woodall 2004; Chang and Gan 2006; Gupta et al. 2006; Zou et al. 2006; Mahmoud et al. 2007) where the profiles that have to be monitored are straight lines. A second stream of application concerns monitoring of signals coming out from sensored machines. For example, Jin and Shi (1999, 2001) and Ding et al. (2006) referred to profiles representing force and torque signals collected from online sensors on a press in a stamping process. The third stream of applications concerns the use of profile monitoring for quality control of *geometric specifications* (Colosimo and Pacella 2007, 2010; Colosimo et al. 2008, 2010). The

Statistical Analysis of Profile Monitoring, First Edition. Edited by Rassoul Noorossana, Abbas Saghaei and Amirhossein Amiri.

present chapter specifically deals with this kind of applications of profile monitoring, which is significant in the manufacturing field.

In manufacturing of mechanical components, technical drawings show an increasing number of geometric specifications (aka geometric tolerances). This is mainly due to the deeper interest in relating the characteristics of a geometric feature (i.e., a machined profile or surface) to the functional properties of the produced item. A geometric specification is a product specification concerning a geometric feature, i.e., a shape in a two-dimensional (2D) or in the three-dimensional (3D) space. This specification can concern simple (e.g., roundness and cylindricity) or even complex shapes (free-forms) and is designed for constraining the deviation of the actual feature from the ideal one, since this deviation can affect the functional properties of the produced item.

A manufacturing process may experience changes due to the material machined, improper setup, errors of the operator, wear or sudden changes of the machine conditions, etc. Usually, these changes may cause deteriorated process performance, i.e., the process may produce an increased number of nonconforming geometric features, i.e., profiles and surfaces that deviate excessively from the ideal or nominal shape. In this framework, a statistical process monitoring procedure should quickly detect, by issuing an alarm, any deviation of the machined shape from the usual pattern obtained when the process is in its in-control state.

The widespread use of modern measurement systems in manufacturing such as coordinate measuring machines (CMMs), which decisively reduce the complexity and the time required for checking geometric specifications, also contribute to the increasing important role that profiles and surfaces play in modern manufacturing of mechanical components. These machines, which can be controlled either manually by an operator or automatically by a computer, are the most general-purpose devices nowadays available in the industry for measuring tasks. Measurements are defined by a probe attached to a moving axis of the machine. This probe (which can be mechanical, optical, or laser, among others) assesses the position of a target in a given reference system. By precisely recording the coordinates of the targets, points are generated which can then be analyzed for describing the geometric feature. Machined profiles and surfaces can be thought as functional data if a subset of the coordinates, which are used to describe the position of each sampled point, can be represented as function of one or more independent variables. When the independent variables are fixed, they may act as a counter of the sampled point.

Consider a 2D circular profile described in a polar reference system. The radius can be modeled as a function of the angular position. For example, Figure 8.1 depicts a circular profile, which consists of a set of $P = 36$ points located on equally spaced angular positions. The independent variable is the angular position $\theta_p = (p-1)(2\pi/P)$, where $p = 1, \ldots, P$ and $\theta_p \in [0, 2\pi]$ radiants (alternatively, θ_p may be expressed in degrees as it is in Figure 8.1), while the dependent one is the radius $y(\theta_p)$. The circular profile in the polar diagram of Figure 8.1 can be equivalently represented as in Figure 8.2.

Similarly, consider a 3D cylindrical surface described in a cylindrical reference system, the surface can be modeled by representing the radius as a function of both the

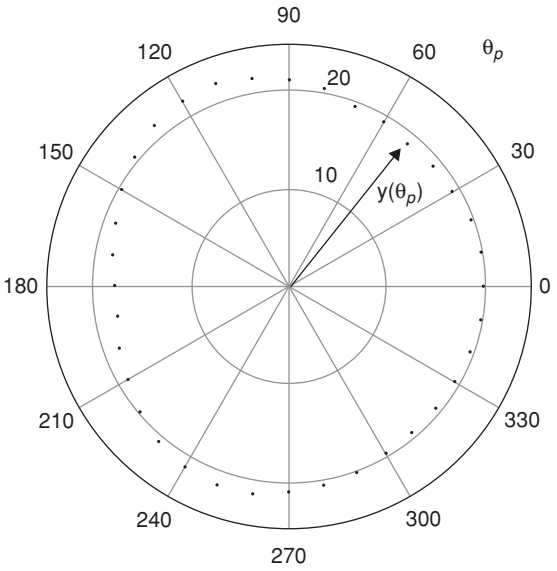

Figure 8.1 An example of circular profile in a polar reference system (angular positions in degrees); the arrow shows the measure at a given angular position (6th point, angular position 50°).

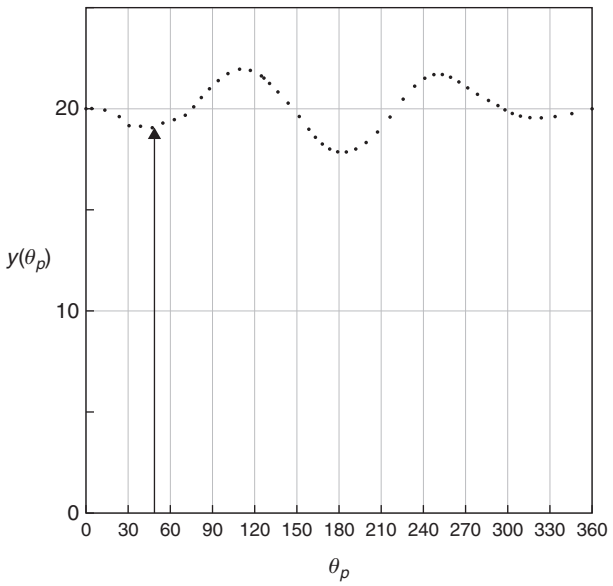

Figure 8.2 An alternative representation of the circular profile depicted in Figure 8.1; the arrow shows the measure at a given angular position (6th point, angular position 50°).

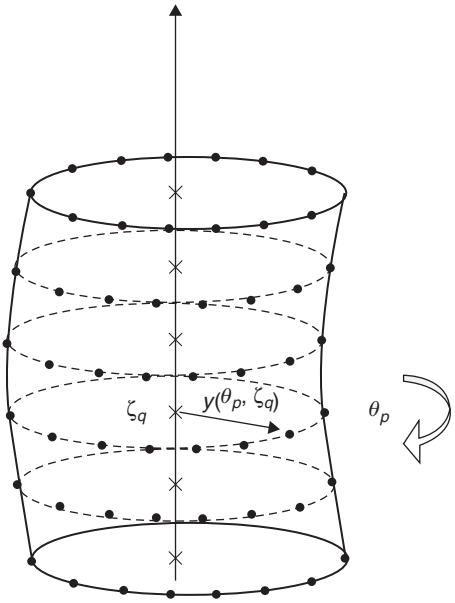

Figure 8.3 An example of cylindrical surface in a cylindrical reference system (function of the angular and vertical position); the arrow shows the measure at a given angular and vertical position.

angular and vertical position. Figure 8.3 depicts a cylindrical surface, which consists of a discrete set of points sampled on fixed and deterministic angular positions, $\theta_p = (p-1)(2\pi/P)$ where $p = 1, \ldots, P$ and $\theta_p \in [0, 2\pi]$ (θ_p may be expressed in degrees), as well as at fixed and deterministic vertical positions $\zeta_q = 2\frac{q-1}{Q-1} - 1$ where $q = 1, \ldots, Q$ and $\zeta_q \in [-1, 1]$. ζ_q is a normalized variable with reference to the vertical position of the point. In Figure 8.3 the independent variables are θ_p and ζ_q, while the dependent one is $y\left(\theta_p, \zeta_q\right)$.

Colosimo and Pacella (2007) and Colosimo et al. (2008) dealt with 2D roundness profiles obtained by lathe-turning, showing that both principal component analysis (PCA) and spatial autoregressive regression (SARX) models can be used for modeling and then monitoring the geometric profile. By combining these models with control charting, the authors showed how out-of-controls of the manufactured profile can be easily and quickly detected. Subsequently, Colosimo and Pacella (2010) compared performance of different approaches (as the previous two model-based approaches as well as a different simpler approach) to outline scenarios in which a specific approach outperforms the others in geometric error monitoring. Eventually, Colosimo et al. (2010), who used as case study the cylindricity of lathe-turned items, showed that 3D surfaces can be modeled as functional data as well and surface monitoring can be considered as a generalization of profile monitoring. The SARX model for 2D roundness profiles in (Colosimo et al. 2008) was exploited as starting reference to identify the parametric models of the large-scale pattern characterizing all the 3D cylindrical surfaces. The model was further extended to represent the spatial

correlation characterizing adjacent points on each machined surface, i.e., a parametric model of cylindrical surfaces able to represent both the large-scale and the small-scale (spatial correlation) characteristics of the machined surface.

The present chapter summarizes results obtained in the aforementioned papers in order to show how profile monitoring can be effectively used for monitoring geometric features. A manufacturing reference case study, which is presented in Subsections 8.1.1 and 8.1.2, is used for this aim. The remainder of the chapter is structured as follows. In Section 8.1, examples of geometric feature concerning circularity are introduced. Section 8.2 presents two analytical methods, i.e., regression and PCA, for modeling the functional relationship representing the manufactured shape and for checking whether or not the estimated model is stable over time. Section 8.3 shows how traditional approaches for SPC can be used for monitoring geometric form errors. Given their easiness of use, these approaches are viable solutions for the industrial practice. In Section 8.4, a comparison study is presented in order to allow one to select a specific approach in a given production scenario. Section 8.5 discusses a possible generalization of the approaches for modeling 3D functional data and how profile monitoring can be generalized to surface monitoring. Conclusions and final remarks are given in Section 8.6.

8.1 EXAMPLES OF GEOMETRIC FEATURE CONCERNING CIRCULARITY

8.1.1 An Example of 2D Geometric Feature Concerning Circularity

Assume we collect a group of N 2D profiles, where each profile consists of P measurements observed at a fixed set of locations. Let $y_n(p)$ denote the dependent variable measured at a specific location of index p on the nth profile, ($p = 1, \ldots, P$ and $n = 1, \ldots, N$). In the case of circular profiles, each consisting of P points sampled on a fixed set of equally spaced positions, the dependent variable $y_n(p)$ is the radial deviation from the nominal radius measured at the angular positions $\theta_p = (p-1)\frac{2\pi}{P}$.

The real case study presented by Colosimo et al. (2008) is used as reference throughout this chapter for 2D profiles. In particular, the case study refers to circular profiles measured on $N = 100$ items produced by lathe-turning starting from C20 carbon steel cylinders (original diameter 30 mm, final diameter 26 mm). One circular profile was measured for each item, at a given distance from the spindle, by using a CMM. Each circular profile consisted of $P = 748$ equally spaced measurements of the radius, where each measurement corresponds to a given angular position. The original measurements were scaled by subtracting the least-squares estimate of the radius and centered at the least-squares estimate of the centre. Hence, each point in a circular profile can be seen as the radial deviation from the mean radius observed at a given angle and this is why observations on a profile can be either positive or negative. A graphical representation of the whole experimental data set, in which the independent variable is the location index on the part and the dependent one is the

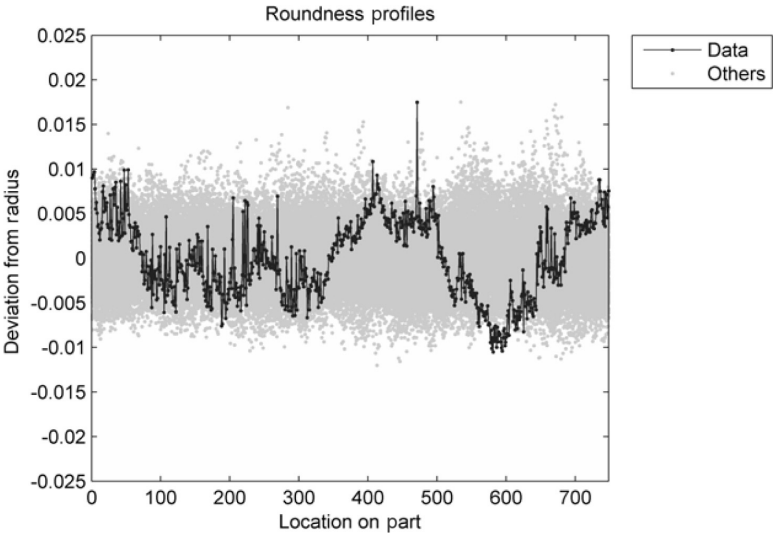

Figure 8.4 100 roundness profiles of 748 points each (one out of the 100 profiles in bold line); vertical axis scale (mm); the independent variable is the location index on the part (from 1 to 748).

deviation from the nominal radius in mm, is shown in Figure 8.4 (where 1 out of the 100 profiles is in bold line).

From Figure 8.4 it seems that no systematic pattern is characterizing the roundness profiles of the reference case study. This appearance hides a common problem of shape analysis, which consists in feature registration or alignment. In fact, the profiles shown in Figure 8.4 are actually misaligned because of the random contact angle of the turning process. Figure 8.5 shows the aligned roundness profiles after that a proper registration procedure is implemented on the data. The goal of the registration procedure implemented is to minimize the phase delay caused by the random contact angle (Colosimo and Pacella 2007). From a visual inspection of Figure 8.5, it can be easily observed that the roundness profiles share a common shape (pattern), i.e., the turning process leaves a specific *manufacturing signature* on the machined components (Colosimo et al. 2008).

For a circular feature, the geometric form error of the actual profile from the ideal geometry is represented by a synthetic variable. In the case of a circular profile, such a value is referred to as *out-of-roundness* (OOR). The OOR is computed as the difference between the maximum and the minimum radial distances of the actual geometric feature with respect to a predetermined center, which is the center of the so-called *substitute geometry*. Geometrically, this corresponds to find two concentric circles, one circumscribing and one inscribing the profile sampled on the manufactured feature. The OOR is then estimated by the width of the annulus determined by these two concentric circles (see Figure 8.6).

Several algorithms and methods can be implemented in order to determine the common center of the circumscribed circle and of the inscribed one. Previous studies

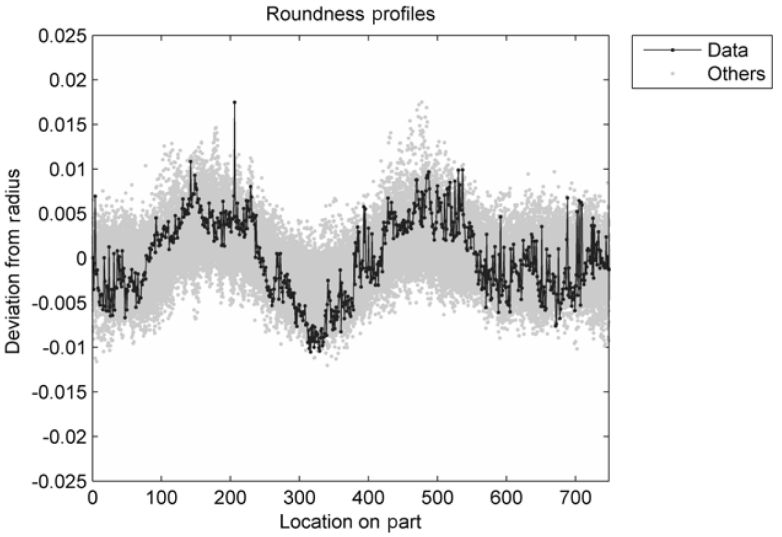

Figure 8.5 Hundred roundness profiles of 748 points each; aligned data (one out of the 100 profiles in bold); vertical axis scale (mm); the independent variable is the location index on the part (from 1 to 748).

(Colosimo et al. 2008) showed that no difference in profile monitoring results can be found if the minimum zone (MZ) rather than the least-squares (LS) approach is used for computing the form error. This is why the LS algorithm is considered as reference from now on.

The LS algorithm (i.e., the widely used procedure for form tolerance evaluation) minimizes the sum of squared deviations of measured points from the fitted feature.

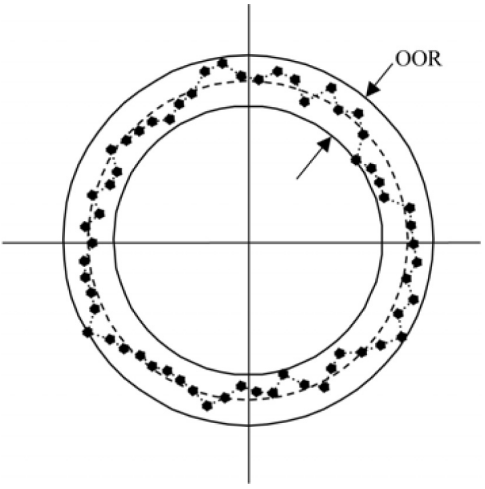

Figure 8.6 A schematic representation of the OOR value for a circular profile; the dashed line represents the substitute geometry (Adapted from Colosimo et al. 2008.)

The LS method associates one substitute feature to measurement points (e.g., one circle for roundness) and calculates the maximum peak-to-valley distance of the measurement points from the substitute feature (Moroni and Pacella 2008).

Usually, a circular profile is considered as conforming to requirements when the OOR value is lower than the geometric specification (i.e., the tolerance on the technical drawing). However, it should be noted that many different circular profiles, which can induce different functional properties of the machined items, can be characterized by the same OOR value resulting in important differences in characteristics of the profile, for example, in the assembly precision, (Cho and Tu 2002).

8.1.2 An Example of 3D Geometric Feature Concerning Cylindricity

Assume we collect a group of N 3D surfaces, where each surface consists of $P \times Q$ measurements observed at a fixed grid of locations. Let $y_n(p, q)$ denote the dependent variable measured at a specific location of indices p and q on the nth surface ($p = 1, \ldots, P, q = 1, \ldots, Q$ and $n = 1, \ldots, N$). In the case of cylindricity, $y_n(p, q)$ is the radial deviation from the nominal cylinder, which is expressed in a cylindrical reference system, i.e., at angular positions $\theta_p = (p - 1)(2\pi / P)$ and vertical positions $\zeta_q = 2\frac{q-1}{Q-1} - 1$ ($\theta_p \in [0, 2\pi]$ and $\zeta_q \in [-1, 1]$).

As previously described, the case study that consists of $N = 100$ items produced by lathe-turning starting from C20 carbon steel cylinders is also used as a reference for 3D surfaces. In particular, one cylindrical surface was measured for each item produced by lathe-turning using a CMM. A set of $P = 68$ generatrices was sampled on each surface and a set of $Q = 61$ points was measured on each of the 68 generatrices. Therefore, a set of $P \times Q = 4148$ points, equally distributed on the cylindrical surface, was eventually available for each produced item.

Among different geometric specifications, cylindricity plays an important role when functionality relates to rotation or coupling of mechanical components (e.g., shafts and holes). According to the standards, the cylindricity is the condition of a surface of revolution in which all the points of the surface are equidistant from a common axis. The deviation of the actual shape from the ideal one is specified by the cylindricity error, which is the minimum radial distance between two coaxial cylinders that contain among them the actual surface. This error is often referred to as *out-of-cylindricity* (OOC). According to the standard, the OOC error has to be computed in two steps. The first step consists in computing the reference cylinder, which is fitted on the actual data by using one of several possible criteria (e.g., LS, MZ). Then, the deviations of the actual shape from the reference cylinder is computed and possibly modeled in the 3D space.

8.2 CONTROL CHARTS FOR PROFILE MONITORING

Many approaches proposed in the literature for profile monitoring focus on the case of linear profiles only. However, when geometric features are of interest, approaches focusing on nonlinear profiles should be more appropriate (Jin and Shi 1999, 2001;

Young et al. 1999; Walker and Wright 2002; Ding et al. 2006; Williams et al. 2007; Jensen et al. 2008; Zou et al. 2008; Zhang and Albin 2009).

Another important issue related to profile monitoring when geometric features are of interest is the autocorrelation of data. Walker and Wright (2002) outlined that "autocorrelation is frequently present in data that are observed within small intervals of time or space. Indeed, the data analyzed exhibited a significant amount of positive autocorrelation" (Walker and Wright 2002, p. 124). Despite the specific mention to the autocorrelation problem, the approach proposed by Walker and Wright (2002) was based on independent data, as most of the other approaches focused on profiles monitoring. The only exception to this general rule is the papers by Jensen et al. (2008) and Colosimo et al. (2008, 2010). In particular, while the first paper dealt with autocorrelated profile data, the second ones focused on spatially autocorrelated data and are more suitable for modeling geometric form features (i.e., profiles and surfaces measured on machined items). As a matter of fact, when data refer to geometric form errors, measurements are often spatially correlated because they are obtained in similar conditions of the machining process and related to similar (local) properties of the machined material. Spatial autocorrelation, which allows one to represent contiguity in space, is different from temporal autocorrelation (time-series models). In the case of profiles, contiguity in space implies that the dependency among data on a profile is bidirectional (i.e., a given point may be correlated to points located on its left and on its right, regardless of the specific direction), while time-series models are suitable to represent just a one-direction dependency (i.e., past data influence future ones) (Whittle 1954). Second, spatial models allow one to represent a specific type of relationship among points observed in closed profiles (e.g., roundness profile). In fact, when data on a closed or circuit profile are sequentially numbered (by defining an arbitrary starting point), observations at the beginning and at the end of the profile are spatially correlated. Colosimo et al. (2008, 2010) presented an approach for modeling roundness profiles based on fitting a SARX model (Cressie 1993). The regression model presented in these papers is summarized in Subsection 8.2.1.

8.2.1 The Regression-Based Approach

With reference to the case study presented in Subsection 8.1.1, assume we organize data into column vectors $\mathbf{y}_n' = \left[y_n(1) \ldots y_n(p) \ldots y_n(P) \right]$, the general SARX model can be written in matrix notation as follows:

$$\begin{aligned} \mathbf{y}_n &= \mathbf{X}\mathbf{b}_n + \boldsymbol{\upsilon}_n, \\ (\mathbf{I} - \mathbf{R}_n)\,\boldsymbol{\upsilon}_n &= \boldsymbol{\varepsilon}_n, \\ \mathbf{R}_n &= \sum_{s=1}^{S} a_{sn}\mathbf{W}^{(s)}. \end{aligned} \tag{8.1}$$

The first expression in Equation (8.1) describes the $P \times 1$ vector of response for the nth profile $\mathbf{y}_n$ as formed by a *large-scale* and a *small-scale* component (Cressie 1993). The large-scale component is given by $\mathbf{X}\mathbf{b}_n$ where $\mathbf{X}$ is a $P \times R$ matrix of R regressor

variables that are assumed to be known and constant and $\mathbf{b}'_n = \left[b_{1n} \cdots b_{rn} \cdots b_{Rn} \right]$ is the $R \times 1$ vector of regression parameters that are normally distributed with mean $\boldsymbol{\beta}' = \left[\beta_1 \cdots \beta_r \cdots \beta_R \right]$ and covariance matrix $\mathbf{B}$ ($\mathbf{b}_n \sim N(\boldsymbol{\beta}, \mathbf{B})$).

The small-scale component is the $P \times 1$ vector of error terms $\boldsymbol{\upsilon}_n$ in Equation (8.1). Error terms are assumed to be spatially correlated and are represented as a spatial autoregressive process (SAR) of order S. The SAR (S) model is given by the last two expressions of Equation (8.1), where $\mathbf{I}$ is the $P \times P$ identity matrix, $\boldsymbol{\varepsilon}_n$ is a $P \times 1$ vector of independently and normally distributed residuals $\left(\boldsymbol{\varepsilon}_n \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)\right)$, and $\mathbf{a}'_n = \left[a_{1n} \cdots a_{sn} \cdots a_{Sn} \right]$ is the vector of coefficients of the SAR (S) model for the nth profile, which is assumed to be normally distributed with mean $\boldsymbol{\alpha}' = \left[\alpha_1 \cdots \alpha_s \cdots \alpha_S \right]$ and covariance matrix $\mathbf{A}$ ($\mathbf{a}_n \sim N(\boldsymbol{\alpha}, \mathbf{A})$).

The $P \times P$ matrix $\mathbf{W}^{(s)}$ of elements $w^{(s)}(p_1, p_2)$ ($p_1, p_2 = 1, \ldots, P$) represents the core of the small-scale model, since it is the sth order adjacent matrix ($s = 1, \ldots, S$). $w^{(1)}(p_1, p_2)$ is set equal to 1 if the p_1th point is the neighbor of the p_2th, and 0 otherwise. Analogously, the element of a second-order adjacency matrix, $w^{(2)}(p_1, p_2)$, is set equal to 1 if the p_1th point is a neighbor of the original first-generation neighbors of the p_2th point, and so on. By definition, all the adjacency matrices are binary (i.e., they can have only 0 and 1 elements) and symmetric matrices whose diagonal elements are equal to 0 (Cressie 1993).

For each profile, two vectors of coefficients $\mathbf{b}_n$ and $\mathbf{a}_n$ are considered. In order to let the model have the most general form, we further assume that these two vectors could also be correlated, i.e., $\text{cov}(\mathbf{b}_n, \mathbf{a}_n) = \mathbf{D}$ (Colosimo and Pacella 2010). In other words, with reference to the parametric model structure given in Equation (8.1), we merge the two vectors characterizing the observed pattern into a C-length single coefficient vector related to the nth profile (where $C = R + S$):

$$\begin{aligned} \mathbf{c}'_n &= \left[\mathbf{b}'_n \quad \mathbf{a}'_n\right] = \left[b_{1n} \cdots b_{rn} \cdots b_{Rn}\, a_{1n} \cdots a_{sn} \cdots a_{Sn}\right], \\ \mathbf{c}_n &\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \text{ where } \boldsymbol{\mu}' = \left[\boldsymbol{\beta}' \quad \boldsymbol{\alpha}'\right], \quad \boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{B} & \mathbf{D} \\ \mathbf{D}' & \mathbf{A} \end{bmatrix}, \end{aligned} \tag{8.2}$$

the model for the nth profile shown in Equation (8.1) is completely defined by a SARX model that requires one to estimate the $C = R + S$ parameters, which are the components of the vector $\mathbf{c}_n$, and the residuals' variance σ^2.

Let $\hat{\mathbf{c}}'_n = \left[\hat{\mathbf{b}}'_n \; \hat{\mathbf{a}}'_n \right]$ represent the vector of C parameters' estimates for the jth profile. A T^2 control chart can be designed with reference to the statistics

$$T_n^2 = (\hat{\mathbf{c}}_n - \boldsymbol{\mu})' \, \boldsymbol{\Sigma}^{-1} (\hat{\mathbf{c}}_n - \boldsymbol{\mu}), \tag{8.3}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix of the coefficients (estimated from a set of in-control profiles), respectively. Williams et al. (2006) studied the performance of different control limits to be used for the T_n^2. When the number of samples N is at least twice the number of parameters estimated ($C + [C(C + 1)/2]$), the following upper control limit (UCL) can be used:

$$UCL = \chi^2_{\alpha, C}, \tag{8.4}$$

where $\chi^2_{\alpha,C}$ is the $100\,(1-\alpha)$ percentile of the chi-square distribution with C degrees of freedom. The $P \times 1$ vector of estimated residuals associated with the nth profile can be described as follows:

$$\mathbf{e}_n = (\mathbf{I} - \mathbf{R}_n)\left(\mathbf{y}_n - \mathbf{X}\hat{\mathbf{b}}_n\right), \tag{8.5}$$

where $\mathbf{R}_n = \sum_{s=1}^{S} \hat{a}_{sn}\mathbf{W}^{(s)}$. The estimated variance of residuals $\hat{\sigma}_n^2$ is given by

$$\hat{\sigma}_n^2 = \frac{\mathbf{e}'_n\mathbf{e}_n}{P-1}. \tag{8.6}$$

In order to monitor the residual variance, a traditional Shewhart-type control chart can be used. Let σ^2 denote the actual residuals' variance, the upper and lower control limits can be computed as follows:

$$\begin{aligned} UCL &= \frac{\sigma^2}{P-1}\chi^2_{\alpha/2,\,P-1}, \\ CL &= \sigma^2, \\ LCL &= \frac{\sigma^2}{P-1}\chi^2_{1-(\alpha/2),\,P-1}, \end{aligned} \tag{8.7}$$

where $\chi^2_{\alpha/2,\,P-1}$ and $\chi^2_{1-(\alpha/2),\,P-1}$ are respectively the upper and lower $\alpha/2$ percentage points of the chi-square distribution with $P-1$ degrees of freedom associated with the residuals (Montgomery 2004).

8.2.1.1 Application of the Regression-Based Approach

The SARX model in Equation (8.1) can be fitted to the data of the case study in previous subsection 8.1.1. In the case of circular profiles, the large-scale component in Equation (8.1) can be expressed as linear combination of *harmonics* (Cho and Tu 2001). Therefore, the element of the rth column and nth row of $\mathbf{X}$ in Equation (8.1), say $x_r\,(p)$, can be expressed either as $x_r\,(p) = \cos\left(h\theta_p\right)$ or $x_r\,(p) = \sin\left(h\theta_p\right)$, i.e., as a sinusoidal function of frequency equal to $h\,(2\pi/P)$ rad/sample. $h \in \{1, \ldots, P/2\}$ is an index, which represents the undulations per revolution (upr) of the sinusoidal function.

Two harmonics were selected for modeling the radial deviations in the actual case study, namely the second and the third one ($h = 2, 3$). Indeed, the process signature was mainly affected by ovality and triangularity. The oval contour was possibly due to a bi-lobe error motion affecting the spindle's lathe or to eccentricity caused by an improper setup, while the tri-lobe pattern was due to a similar error motion of the spindle. Hence, matrix $\mathbf{X}$ in Equation (8.1) has $R = 4$ columns. The pth row of matrix $\mathbf{X}$ is equal to $\left[\cos\left(2\theta_p\right)\ \sin\left(2\theta_p\right)\ \cos\left(3\theta_p\right)\ \sin\left(3\theta_p\right)\right]$.

As for the small-scale component of the actual roundness profiles, the vector error terms $\boldsymbol{\upsilon}_n$ in Equation (8.1) was modeled as a SAR model of order 2 ($S = 2$) using the algorithm implemented in the spatial econometrics toolbox (LeSage 1999).

Table 8.1 Parameters for the Distribution of SARX Coefficients $\mathbf{c}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for the Actual Roundness Data of the Reference Case Study

$$\boldsymbol{\beta}' = \begin{bmatrix} -0.0341 & 0.0313 & 0.0080 & -0.0322 \end{bmatrix}$$

$$\boldsymbol{\alpha}' = \begin{bmatrix} 0.3021 & 0.2819 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 4.0646 & -2.0200 & 0.6540 & 0.2652 \\ -2.0200 & 3.8961 & 1.4851 & 0.0614 \\ 0.6540 & 1.4851 & 2.2346 & -0.1074 \\ 0.2652 & 0.0614 & -0.1074 & 3.1214 \end{bmatrix} \times 10^{-4}$$

$$\mathbf{D} = \begin{bmatrix} -0.8844 & -2.4101 \\ -1.2123 & 1.9568 \\ -1.1844 & 0.5958 \\ -1.4993 & -3.7224 \end{bmatrix} \times 10^{-4}$$

$$\mathbf{A} = \begin{bmatrix} 38.0199 & 15.8999 \\ 15.8999 & 43.2491 \end{bmatrix} \times 10^{-4}$$

Given the identified model, each profile of index n was associated to a vector of $C = 4 + 2$ parameters $\mathbf{c}_n = \begin{bmatrix} b_{1n} & b_{2n} & b_{3n} & b_{4n} & a_{1n} & a_{2n} \end{bmatrix}$. We observed that the C-length parameter vector $\mathbf{c}_n$ changes from profile to profile according to a C-variate normal distribution of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\mathbf{c}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (the Shapiro–Wilk normality test for multivariate distribution was used, with a p-value equal to 0.03121). Table 8.1 summarizes the vector $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\beta}' & \boldsymbol{\alpha}' \end{bmatrix}$ and the matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{B} & \mathbf{D} \\ \mathbf{D}' & \mathbf{A} \end{bmatrix}$ estimated on the actual data.

Hence, by considering the values of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in Table 8.1, instances of the C-length parameter vector $\mathbf{c}_n$ can be simulated on a computer (the command "mvnrnd" of MATLAB was used in our work). By combining instances of the C-length parameter vector $\mathbf{c}_n$ to instances of the P-length vector of residuals ε_n the SARX model in Equation (8.1) can be used to simulate on a computer realistic roundness profiles (Colosimo and Pacella 2010). In particular, the residuals ε_n were simulated as independently and normally distributed, with zero mean and common standard deviation σ (the command "randn" of MATLAB was used in our work). The numerical value $\sigma = 9.2244 \cdot 10^{-4}$ mm was estimated from actual data and used for the simulations of roundness profiles in our work.

As for example, we considered a set of $N = 100$ roundness profiles obtained from computer simulation. Two control charts (i.e., a multivariate control chart on the estimated coefficients $\hat{\mathbf{c}}_n'$ and a univariate control chart on the estimated residuals' variance $\hat{\sigma}_n^2$) were designed on the simulated data. Assuming a nominal false-alarm probability $\alpha' = 1\%$, the Type I error probability of each chart was set equal to $\alpha = 1 - \sqrt{1 - \alpha'} = 0.5012\%$. Given that $C = 6$ and $N = 100$ in our case, the condition suggested in Williams et al. (2006), i.e., $N > 2(C + [C(C + 1)/2])$, to apply the asymptotic control limit given in Equation (8.4) holds. Therefore, the control limit of the T^2 control chart is given by $UCL = \chi^2_{0.005012,6} = 18.5416$. Figure 8.7 depicts the T^2 control chart which indicates no out-of-control profiles.

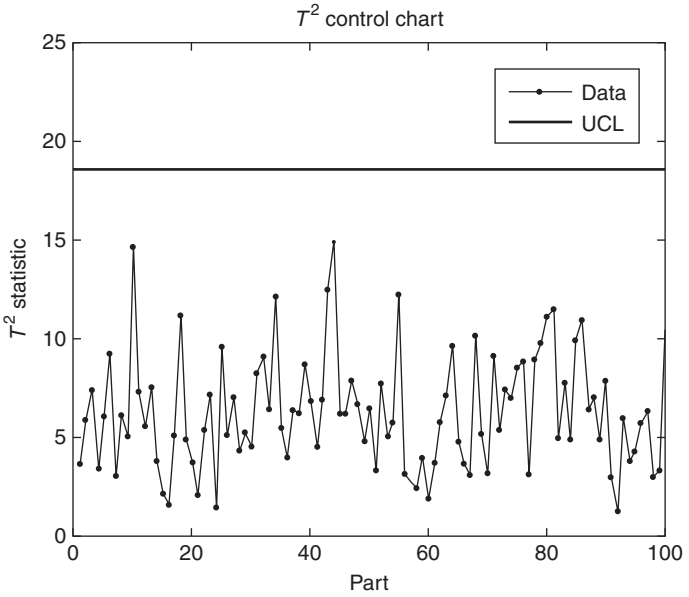

Figure 8.7 T^2 control chart for Phase I of 100 roundness profiles (simulated data).

Similarly, in order to monitor the residuals' variance, a control chart was implemented on the set of 100 simulated roundness profiles. The control limits in Equation (8.7) were used for this control chart. An estimate of the residuals' variance equal to 8.4717×10^{-7}, which was obtained as mean of the residuals' variance for each of the $N = 100$ simulated profile, was considered for the limits. The numerical values $UCL = 9.7539 \times 10^{-7}$ and $LCL = 7.2935 \times 10^{-7}$ were obtained. Figure 8.8 depicts the σ^2 control chart which indicates no out-of-control profiles.

8.2.2 The PCA-Based Approach

Ramsay and Silverman (2005) presented an extension of PCA to functional data, i.e., an approach that allows one to find a set of orthonormal functions (also called functional principal components—PCs), so that the original data can be approximated in terms of a linear combination of these basic functions. In particular, Ramsay and Silverman (2005) showed that, in the case of equally spaced observations, the easiest way to compute the PCs consists of modeling the curve data sampled at regular intervals as a multivariate vector, and performing a traditional PCA on the set of samples collected over different curves.

When the PCA outlines a set of significant PCs to be retained, the coefficients (or loadings) defining these significant PCs can be interpreted as eigenfunctions (also called empirical orthogonal functions). These eigenfunctions do not have a parametric expression and are empirical, since they are obtained by the data at hand. A rough sketch of how PCA works on profile data is discussed below.

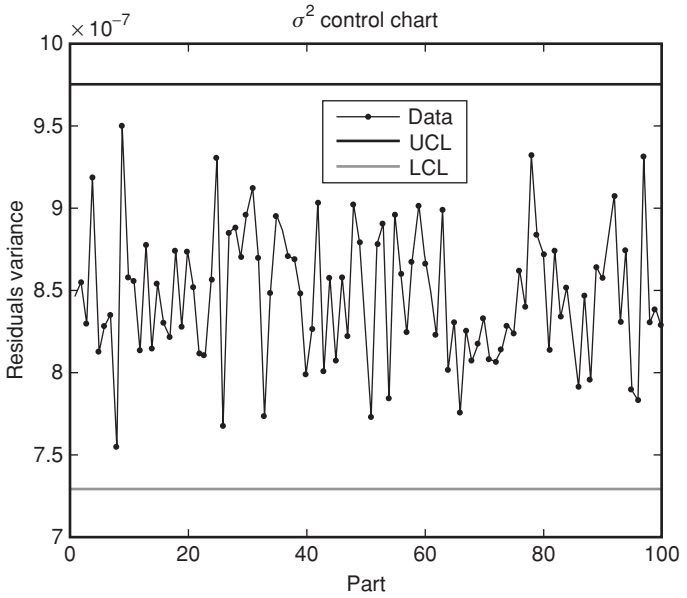

Figure 8.8 Shewhart control chart for monitoring the residual variance of Phase I roundness profiles (simulated data).

Assume we organize N vectors of $P \times 1$ profile data $\mathbf{y}_p$ into an $N \times P$ matrix $\mathbf{Y}$ whose nth row is the transpose of the nth data vector $\mathbf{y}_n$. PCA consists of performing a spectral decomposition of the covariance matrix of $\mathbf{Y}$. The covariance matrix describes the variability of the data observed at each location with respect to the mean value observed at the same location in all the profiles.

Let $\mathbf{S}$ be the covariance matrix, i.e., $\mathbf{S} = [1/(N-1)] \sum_{n=1}^{N} (\mathbf{y}_n - \bar{\mathbf{y}})(\mathbf{y}_n - \bar{\mathbf{y}})'$, where $\bar{\mathbf{y}} = (1/N) \sum_{n=1}^{N} \mathbf{y}_n$ is the sample mean profile, the spectral decomposition consists of finding the $P \times P$ matrices $\mathbf{U}$ and $\mathbf{L}$ that satisfy the following relationship:

$$\mathbf{U}'\mathbf{S}\mathbf{U} = \mathbf{L}, \tag{8.8}$$

where $\mathbf{L}$ is a diagonal matrix that contains the eigenvalues of $\mathbf{S}$ (say l_p), while $\mathbf{U}$ is an orthonormal matrix whose pth column $\mathbf{u}_p$ is the pth eigenvector of $\mathbf{S}$ (the so-called loadings).

With reference to the nth profile $\mathbf{y}_n$, denote by $\mathbf{z}_n$ the vector

$$\mathbf{z}_n = \mathbf{U}'(\mathbf{y}_n - \bar{\mathbf{y}}) = \left[z_{n1} \ \ldots \ z_{np} \ \ldots \ z_{nP} \right]', \tag{8.9}$$

where z_{np} are the so-called scores. Each profile can then be expressed as a linear combination of loadings $\mathbf{u}_n$, where the weights of the linear combination are the scores z_{np}

$$\mathbf{y}_n = \bar{\mathbf{y}} + z_{n1}\mathbf{u}_1 + \cdots + z_{np}\mathbf{u}_p + \cdots + z_{nP}\mathbf{u}_P. \tag{8.10}$$

Since the PCs are statistically uncorrelated and each PC has variance equal to the corresponding eigenvalue (l_p), we can rank the PCs according to the associated eigenvalue and decide to retain just the most important ones (i.e., the ones which are associated with larger variances). Different approaches can be used to select the proper set of PCs (Jolliffe 2002; Jackson 2003). For instance, cross-validation can be effectively used to choose the number M of significant PCs (Colosimo and Pacella 2007). When a subset M of the whole number of P PCs is retained ($M < P$), the original profile can be estimated as follows:

$$\hat{\mathbf{y}}_{n(M)} = \bar{\mathbf{y}} + z_{n1}\mathbf{u}_1 + \cdots + z_{nm}\mathbf{u}_m + \cdots + z_{nM}\mathbf{u}_M. \tag{8.11}$$

Similar to the regression-based approach, also in the case of PCA a T^2 control chart can be used for monitoring the vector of the first M retained PCs. In this case, the T^2 statistic is given by (Jackson 2003)

$$T_n^2 = \frac{z_{n1}^2}{l_1} + \cdots + \frac{z_{n2}^2}{l_m} + \cdots + \frac{z_{nM}^2}{l_M}. \tag{8.12}$$

If the profile changes in a direction orthogonal to that of the first M PCs, previous T^2 control chart will not be able to issue an alarm. For this reason, another control chart based on the Q statistic (sometimes referred to as the squared prediction error or SPE control chart) also has to be used (Jackson 2003). Given the estimate in Equation (8.11), the Q statistic can be computed as the sum of the squared errors obtained by reconstructing each observation by the first M PCs:

$$Q_n = \left(\mathbf{y}_n - \hat{\mathbf{y}}_{n(M)}\right)' \left(\mathbf{y}_n - \hat{\mathbf{y}}_{n(M)}\right). \tag{8.13}$$

The upper control limit of the T^2 statistics in Equation (8.12) can be computed as follows (Williams et al. 2006).

$$UCL = \chi^2_{\alpha,M}, \tag{8.14}$$

where $\chi^2_{\alpha,M}$ is the $100\,(1-\alpha)$ percentile of the chi-square distribution with M degrees of freedom. With reference to the Q statistic, according to Nomikos and MacGregor (1995) the upper control limit can be computed as

$$UCL = g\chi^2_{\alpha,h}, \tag{8.15}$$

where g and h can be estimated as $\hat{g} = \hat{\sigma}^2_Q/(2\bar{Q})$, $\hat{h} = 2\bar{Q}^2/\hat{\sigma}^2_Q$ while $\bar{Q}$ and $\hat{\sigma}^2_Q$ are the sample mean and the sample variance obtained by computing the Q statistics via Equation (8.13) for the set of the N profiles.

8.2.2.1 Application of the PCA-Based Approach

With reference to the aforementioned data set of $N = 100$ simulated roundness profiles, we used a cross-validation approach in order to determine the number of

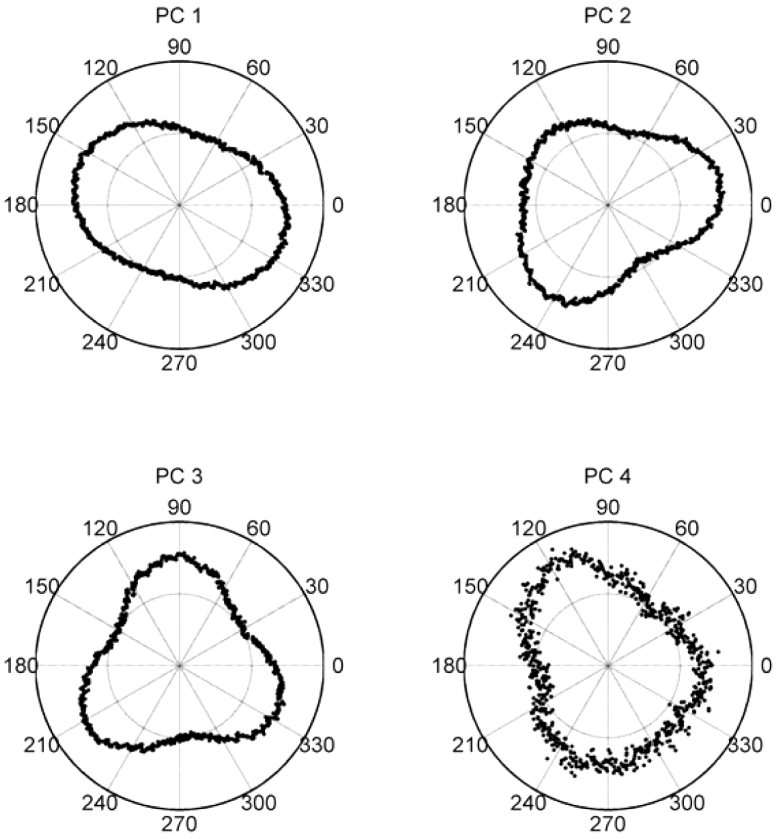

Figure 8.9 First four eigenfunctions related to the four retained PCs (computed on simulated data).

significant PCs to be retained (Colosimo and Pacella 2007). With reference to the simulated roundness profiles the number of significant PCs is equal to $M = 4$.

Once the PCA has been performed, the retained PCs should be interpreted to gain more insight into the systematic pattern characterizing the machined profiles. To this aim, each eigenfunction $\mathbf{u}_p$ (i.e., the coefficients of each eigenvector also known as loadings) can be graphically presented as a function of the location. Following this practice, Figure 8.9 depicts the polar diagrams of the first four eigenfunctions $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$, which are related to the four retained PCs. The first PC, which describes the most important component of variability, presents a bi-lobe form. This qualitative observation indicates that the main variability around the mean profile is due to a periodic function characterized by a frequency of 2 upr. The second PC is a mixture of a bi-lobe and a tri-lobe contour. This mixture is obtained by combining two periodic functions, namely, a 2 and a 3 upr harmonic. The third PC presents a tri-lobe contour, while the fourth PC is again a mixture of a bi-lobe and a tri-lobe contour.

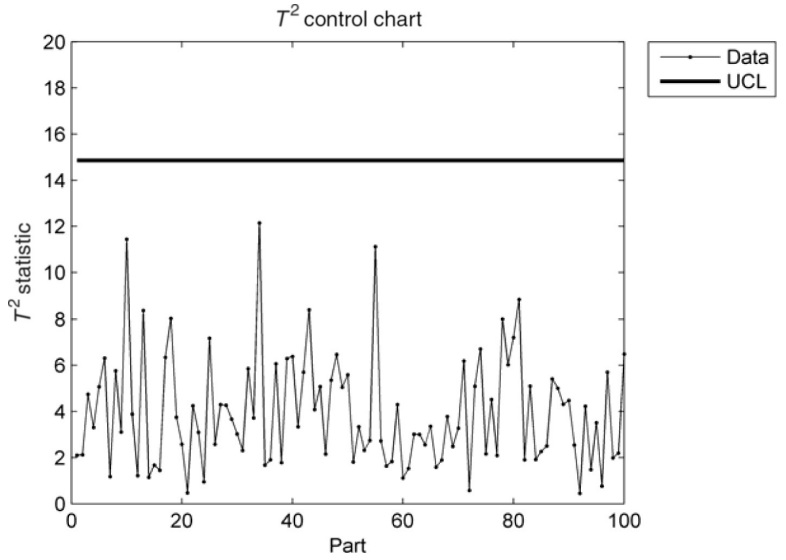

Figure 8.10 T^2 control chart for Phase I of 100 roundness profiles (simulated data).

Given the PCA model, a multivariate control chart on the scores $(z_{n1}, z_{n2}, z_{n3}, z_{n4})$ and a univariate control chart on the Q statistic have to be designed. Assuming a nominal false-alarm probability $\alpha' = 1\%$, the Type I error probability of each chart is set equal to $\alpha = 1 - \sqrt{1-\alpha'} = 0.5012\%$. The control limit of the T^2 control chart in Equation (8.14) is $UCL = \chi^2_{0.005012,4} = 14.8546$. The UCL of the Q control chart is in Equation (8.15) and is about equal to 1.325×10^{-3}.

Figure 8.10 depicts the PCA-based T^2 control chart, which indicates no out-of-control conditions for the simulated profiles. Similarly, Figure 8.11 depicts the Q control chart, which indicates two out-of-control conditions for the 52nd and 54th profile, where the Q statistic resulted slightly above the UCL. As we dealt with simulated in-control profiles no special causes existed to explain these out-of-controls. Therefore, they can be considered false alarms released by the Q control chart and probably due to the different structure of the error terms' correlation, which indeed changes from profile to profile.

8.3 SIMPLE APPROACHES FOR MONITORING MANUFACTURED PROFILES: THE INDUSTRIAL PRACTICE

8.3.1 Control Chart on Geometric Errors

When quality of a manufactured product is related to a geometric specification (represented by a tolerance value), the Shewhart's control chart can still be used for process monitoring. In fact, the information related to a geometric feature must be summarized in just one synthetic variable to decide whether the machined item

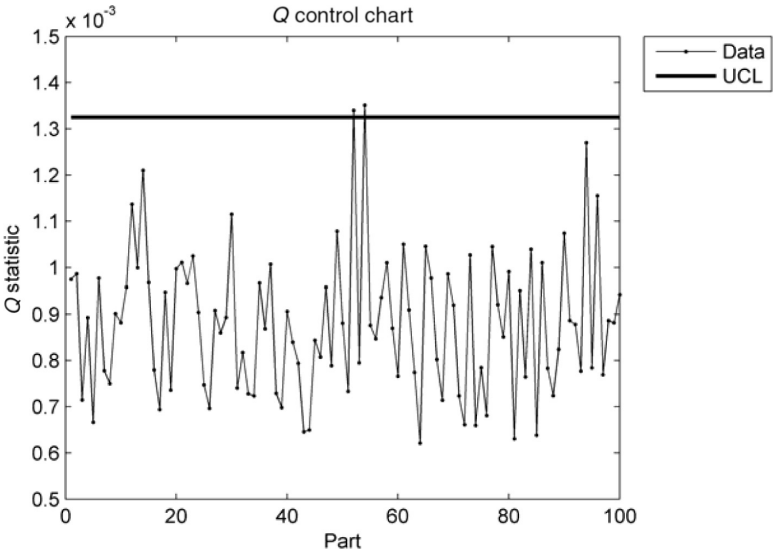

Figure 8.11 Q control chart for Phase I of 100 roundness profiles (simulated data).

has to be scrapped/reworked or can be considered conforming to the requirement. In the case of a circular profile, the OOR value is computed. Then, the circular feature is considered conforming to the requirement if the corresponding OOR is not greater than the tolerance value.

Since the OOR values have to be computed for quality inspection of manufactured items, these values can also be exploited for process monitoring. Indeed, the simplest approach for process monitoring consists of designing a Shewhart's control chart of the OOR values. In order to exemplify this approach, the set of 100 roundness profiles previously obtained from simulation is considered as reference. Table 8.1 reports the OOR values (LS method) obtained for each item. The implementation developed by NPL (the UK's National Measurement Laboratory) and based on the Least Squares Geometric Elements (LSGE) library for MATLAB was considered (Forbes 1994).

With reference to the 100 samples in Phase I, summarized in Table 8.2, an individual's control charts on the OOR values was designed. The Anderson–Darling test was implemented in order to detect any departures from normality for the distribution function of the 100 OOR values. From numerical computation on the data reported in Table 8.1, the set of 100 OORs resulted to be normally distributed (p-value equal to 0.559). Therefore, a control chart for individuals was designed for the sequence of sample OOR values (with a nominal false-alarm probability $\alpha = 1\%$). This control chart, in which no out-of-control alarms are released, is showed in Figure 8.12.

8.3.2 The Location Control Chart

A different approach, aimed at combining simplicity with the need of keeping all the information of data observed at each location of the machined feature is

Table 8.2 OOR Values (100 Simulated Profiles) on the basis of the LS Algorithm; Scale OOR (mm)

	OOR		OOR		OOR		OOR		OOR
1	0.0143	21	0.0125	41	0.0150	61	0.0154	81	0.0137
2	0.0164	22	0.0099	42	0.0127	62	0.0163	82	0.0128
3	0.0137	23	0.0144	43	0.0169	63	0.0142	83	0.0074
4	0.0155	24	0.0135	44	0.0097	64	0.0129	84	0.0126
5	0.0127	25	0.0180	45	0.0126	65	0.0146	85	0.0087
6	0.0139	26	0.0114	46	0.0106	66	0.0100	86	0.0101
7	0.0147	27	0.0133	47	0.0169	67	0.0118	87	0.0129
8	0.0069	28	0.0075	48	0.0119	68	0.0144	88	0.0137
9	0.0158	29	0.0138	49	0.0179	69	0.0142	89	0.0177
10	0.0086	30	0.0110	50	0.0112	70	0.0175	90	0.0176
11	0.0114	31	0.0147	51	0.0150	71	0.0120	91	0.0119
12	0.0111	32	0.0142	52	0.0131	72	0.0112	92	0.0116
13	0.0169	33	0.0148	53	0.0096	73	0.0169	93	0.0139
14	0.0098	34	0.0190	54	0.0118	74	0.0060	94	0.0102
15	0.0137	35	0.0124	55	0.0126	75	0.0134	95	0.0144
16	0.0100	36	0.0106	56	0.0092	76	0.0123	96	0.0112
17	0.0089	37	0.0132	57	0.0125	77	0.0132	97	0.0181
18	0.0179	38	0.0124	58	0.0099	78	0.0198	98	0.0135
19	0.0128	39	0.0120	59	0.0136	79	0.0120	99	0.0127
20	0.0088	40	0.0142	60	0.0106	80	0.0110	100	0.0153

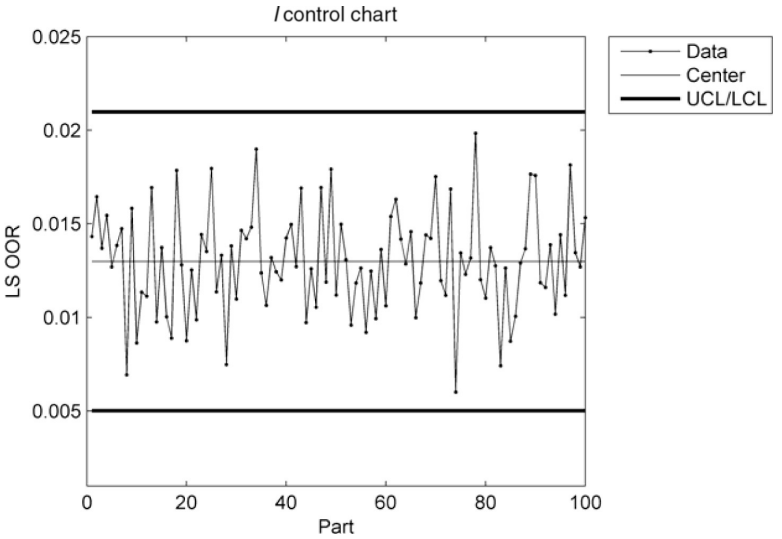

Figure 8.12 Individuals control charts of the transformed OOR values based on the LS algorithm; vertical axis scale (mm).

the *location control chart*. This approach, which was presented in Boeing (1998, pp. 89–92), consists of applying a Shewhart's control chart separately to each data point observed at a given location of the part. The rationale behind this approach is that, if the observed shape is in-control, the data observed at that specific location should stay within that interval with a given probability. On the other hand, when the process goes out-of-control, it is likely that the control interval will be violated at one or more locations.

In order to design the location control chart, the first step consists of identifying the center of each interval, i.e., the systematic pattern of the in-control shape. This reference is usually estimated as the average of all the in-control data observed at each location. Starting from the mean shape, the location control chart (LOC CC) consists of computing the upper and lower control limits at each location, by using the usual approach that place the limits at $\pm K$ standard deviations from the sample mean. According to this method, an alarm is issued when at least one point, in the whole set of data observed, exceeds the control limits.

Due to its inner simplicity, this chart can be easily applied in industrial practice (and in fact its origin is in Boeing with reference to applications in which numerous measurements of the same variable, e.g., a dimension such as thickness, are made at several locations on each manufactured part). However, since the control limits used at each location depend on the responses at that specific position only, the main disadvantage with this method is that the multivariate structure of data is completely ignored. The only form of relationship between control intervals at each location is a constraint on the false alarm, as discussed in Subsection 8.3.2.1.

8.3.2.1 Design of the Location Control Chart

The control limits for location of index p are as follows:

$$\begin{aligned} UCL(p) &= \bar{y}(p) + Z_{\alpha/2} s(p), \\ CL(p) &= \bar{y}(p), \\ LCL(p) &= \bar{y}(p) - Z_{\alpha/2} s(p), \end{aligned} \tag{8.16}$$

where $\bar{y}(p) = \frac{1}{N} \sum_{n=1}^{N} y_n(p)$ and $s(p) = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} [y_n(p) - \bar{y}(p)]^2}$ are, respectively, the sample mean and the sample standard deviation of data observed at location p, while $Z_{\alpha/2}$ represents the $(1 - \alpha/2)$ percentile of the standard normal distribution. The constant $K = Z_{\alpha/2}$ of the control band is computed as a function of the required false alarm rate. However, given that P dependent control rules are simultaneously applied, the percentile of the standardized normal distribution used to compute K should be corrected. To this aim, the Bonferroni's rule for dependent events is used to attain an actual false alarm rate not greater than a predefined value. Let α' denote the upper bound of the false alarm rate the value $\alpha = \alpha'/P$ is used for designing the P control limits in Equation (8.16).

It is worth noting that different procedures can also be used, such as the Simes' modified Bonferroni procedure (Simes 1986). Colosimo and Pacella (2010) showed

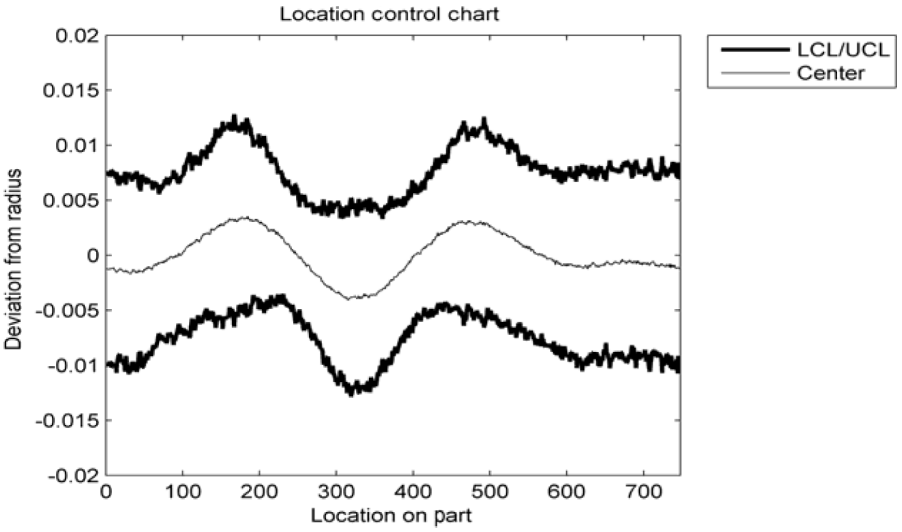

Figure 8.13 LOC CC (748 locations) with reference to the 100 simulated profiles; actual false alarm rate not greater than 0.01; vertical axis scale (mm).

that when compared with the standard Bonferroni's method, the Simes' procedure does not produce significant effects on the false alarm rate of the location control chart. Furthermore, since the Simes' procedure does not allow the graphical representation of the control region as the Bonferroni's procedure does, this latter method is the one considered in practice.

Figure 8.13 shows the average profile and the control band of the simulated roundness profiles. The Bonferroni's rule for dependent events is used to attain an actual false alarm rate not greater than a predefined value. In particular, assuming the standard value $\alpha' = 0.01$ as upper bound for the Type I error probability, the value $\alpha = 1.3369 \times 10^{-5}$ (i.e., $Z_{\alpha/2} = 4.354$) is used for designing the 748 control limits in Equation (8.16).

Figure 8.14 shows one out of the 100 profiles depicted against the control limits of the location control chart. The profile is plotted against this control region with the advantage of allowing a simple identification of the locations where problems arise. In this specific case, the profile is considered in-control. Indeed, from the visual inspection it appears that the profile behavior is close to the average common profile and there is no apparent discrepancy in the shape of the profile when compared to the center line of the location control chart.

8.4 PERFORMANCE COMPARISON

In order to allow practitioners to select a specific approach for monitoring functional data in a given situation, the present section provides a numerical comparison based on simulation of (1) the LOC CC; (2) the individuals' control chart of the geometrical

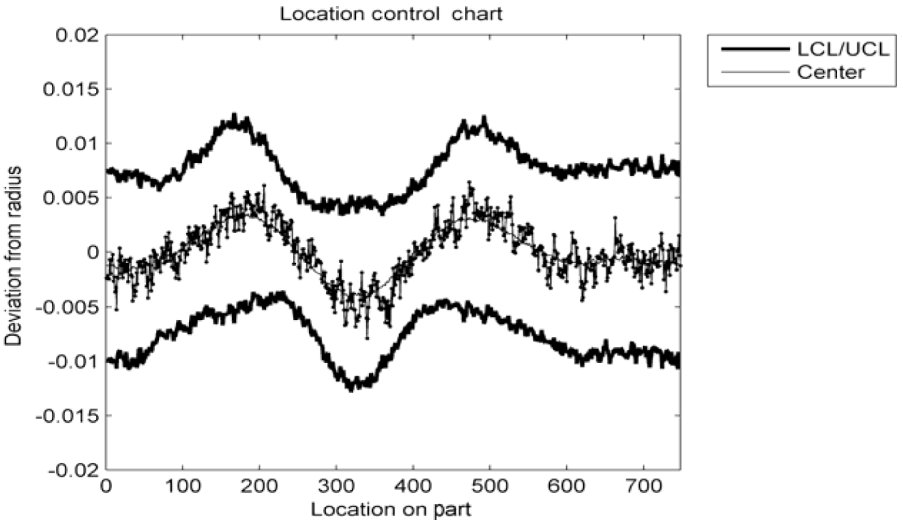

Figure 8.14 One out of the 100 profiles of the reference case study depicted against the control limits of the location control chart; vertical axis scale (mm).

errors (OOR CC); (3) the regression-based approach (REG CC); and (4) the PCA-based approach (PCA CC).

The comparison study is based on two simulation scenarios, which are obtained from the real case study of roundness profiles obtained by lathe-turning. The simulation scenarios were designed for representing two different, although realistic, productive situations. Performance is measured both as the ability to obtain a predefined false alarm rate in the design phase of the control chart (Phase I) as well as to detect unusual patterns in the functional data during the operating phase (Phase II).

For Phase I analysis, only groups of uncontaminated process samples were considered in this study. In-control profiles were obtained by simulation, while the performance of the competing methods were compared in terms of the probability of obtaining at least one statistic outside the control limits when performing control charting using the set of simulated samples. Control chart parameters estimated from these Phase-I samples were not used in the subsequent Phase-II comparison study, where performance comparison was based on the ideal assumption that the in-control parameters are known. Indeed, computer simulation was used to obtain a large data set of in-control profiles in order to estimate as closely as possible control chart's parameters.

Basically, the objective in this section is to investigate situations where each approach should be preferred to the others, thus to provide some guidelines for implementing profile monitoring in actual applications.

8.4.1 Production Scenarios Under Study

In this section, the focus is on two different production scenarios. On the one hand, the first production scenario mimics the real case study where the C-length parameter

vector $\mathbf{c}_n$ changes from profile to profile according to a C-variate normal distribution of mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\mathbf{c}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The actual values of the mean vector $\boldsymbol{\mu}$ and of the covariance matrix $\boldsymbol{\Sigma}$ used for the simulation of this scenario are summarized in Table 8.1.

On the other hand, the second scenario is obtained by perturbing the SARX model in Equation (8.1) with reference to the variability that characterizes the C-length parameter vector $\mathbf{c}_n$. In particular, a null matrix is considered as covariance matrix ($\boldsymbol{\Sigma} = \mathbf{0}$) for the C-variate normal distribution of model parameters. Hence, a fixed vector of parameters is used for each instance ($\mathbf{c}_n = \boldsymbol{\mu}$).

The production scenario where $\mathbf{c}_n$ may change from profile to profile is referred to as *random-effect* model. The production scenario, in which parameters $\mathbf{c}_n$ do not change, is referred to as *fixed-effect* model. It is worth noting that random-effect and fixed-effect are referred to both the large-scale and small-scale variability components in Equation (8.1).

A fixed-effect model is usually assumed in traditional approaches for profile monitoring (Woodall et al. 2004). In manufacturing, assuming that the input material is very stable and homogeneous, profile-to-profile variability of the small-scale component (spatial structure) can probably be neglected. Furthermore, assuming a process more stable and/or more controlled, also the profile-to-profile variability in the large-scale component (harmonic structure) tends to vanish. In other words, while the first scenario with random effects corresponds to a common machining process, in which natural causes of variability affects the variability in both the parameters and residuals of the model, the second scenario with fixed effects can be considered representative of a process more stable and/or more controlled in which natural causes can affect only the residuals of the model.

8.4.2 Phase I: Performance Comparison

In Phase I, a set of N process samples is analyzed in order to both evaluate the stability of the process and estimate the in-control state's parameters. During Phase I, the performances of the competing control charts are compared in terms of the probability of deciding whether or not the process is stable. This is the probability of obtaining at least one statistic outside the control limits when performing control charting using the set of N process samples.

Given a desired false alarm or Type I error rate α', assumed equal to $\alpha' = 1\%$ in the following, the competing approaches can be compared in terms of the probability of actually achieving this nominal value.

To compare the performance of the alternative approaches, we considered 10000 replicates of Phase I control charting for different values of the number of profiles used in this phase, specifically $N = 50$, 100, 150, and 200. We recorded the false alarm rate in each replicate (computed as the number of out-of-control signals divided by N) and then we stored the average false alarm rate obtained in the whole set of 10000 replicates.

Table 8.3 summarizes the actual Type I error rates for each of the competing methods in the two simulation scenarios considered in our work. For the two production

Table 8.3 Phase I Simulation Results; Average type I Error Rate and Corresponding Standard Deviations within Brackets

	Random-Effect Model				Fixed-Effect Model			
N	LOC CC	OOR CC	REG CC	PCA CC	LOC CC	OOR CC	REG CC	PCA CC
50	0.070% (0.004%)	0.815% (0.013%)	1.048% (0.014%)	1.210% (0.015%)	0.048% (0.003%)	1.124% (0.015%)	1.049% (0.014%)	0.935% (0.014%)
100	0.283% (0.005%)	0.833% (0.009%)	1.037% (0.010%)	1.495% (0.012%)	0.208% (0.005%)	1.167% (0.011%)	1.013% (0.010%)	1.022% (0.010%)
150	0.391% (0.005%)	0.836% (0.007%)	1.044% (0.008%)	1.562% (0.010%)	0.306% (0.005%)	1.167% (0.009%)	0.985% (0.008%)	1.048% (0.008%)
200	0.454% (0.005%)	0.828% (0.006%)	1.031% (0.007%)	1.607% (0.009%)	0.353% (0.004%)	1.186% (0.008%)	0.988% (0.007%)	1.065% (0.007%)

scenarios under study it can be observed that the false alarm rates produced by the LOC CC are always lower than the nominal value (1%), although the gap between actual and nominal rate decreases as the number N of profiles used in Phase I increases. This is an expected result as the Bonferroni's inequality allows one to set only an upper bound on the actual false alarm rate. Note that a reduced false alarm rate could seem an advantage at first sight. However, it means that control limits are too far from the center line, thus, resulting in the ineffective detection of out-of-control profiles when they arise.

The false alarm rates produced by the OOR CC are smaller than the nominal value in the case of random-effect model, while they are greater than the nominal value in the case of fixed-effect model. These results can be mainly ascribed to a slight departure from normality observed for the OOR values, where the departure depends on the specific scenario (random-effects or fixed-effects). It is also worth noting that, in the case of the individuals control chart on the geometrical errors, the false alarm rate of the control chart is not influenced by the number N of profiles used in Phase I.

In contrast to the previous two approaches, the actual false alarm rates obtained by the regression-based approaches are much closer to the nominal value. In fact, as shown in Table 8.3, in spite of both the specific scenario and the dimension N, there is no practical difference between the actual false alarm rate produced by the REG CC and the nominal value.

The PCA CC is a little more sensitive than the REG CC to the production scenario, since it gives a false alarm closer to the nominal value in the case of fixed-effect model but not in the case of random-effect. With reference to this last scenario, the PCA CC consists of two control charts. A multivariate control chart based on the first four retained PCs and a univariate control chart for monitoring the residuals (the Q chart). The excessive false alarm rate produced is mainly due to the Q chart. As a matter of fact, Colosimo and Pacella (2010) showed that the upper control limit for the Q statistic proposed by Nomikos and MacGregor (1995) is substantially adequate when

an independently distributed error term is considered in the model, as well as in the case of correlated errors with fixed effects in the small-scale component of the model. However, such a limit is not suitable in the case of correlated errors with random effects in the small-scale component of the model.

Finally, with reference to the Phase I simulation results reported in Table 8.3, it is fair to conclude that regression-based control charts can easily be designed, since the actual false alarm rate can be achieved in practice by using an analytic computation of the control limits. Similar conclusions can also be drawn with reference to the PCA CC , but only in the specific production scenario of fixed effects. On the contrary, the design of an LOC CC can be a difficult task, since the actual false alarm rate can be different from the expected rate. In the case of the individuals control chart on the geometrical errors, we observed that a departure from normality may also make difficult the design of this control chart. However, as we are monitoring non-negative values (the OOR values), if a proper power transformation (Box–Cox transformation) can be identified, we expect that the design of this individuals control chart should be simplified.

8.4.3 Phase II: Performance Comparison

The objective in Phase II is to quickly detect any change in the process from its in-control state. The monitoring approaches are compared in terms of the average run length (ARL), where the run length is defined as the number of samples taken until an out-of-control signal is issued.

In order to evaluate performance in Phase II, occurrences of assignable causes are simulated by a total of three out-of-control conditions. These out-of-controls are simulated by spindle-motion errors (Cho and Tu 2001), which are modeled by introducing a spurious harmonic in the baseline model of roundness profile data. Each condition is then characterized by a parameter directly proportional to the severity of the out-of-control introduced in the baseline model. In particular, denoting by $y_n(p)$ the measurement of index $p = 1, \ldots, P$ on the profile of index $n = 1, 2, \ldots$, the out-of-controls are simulated according to the following three models:

half-frequency spindle-motion error, which is modeled as:

$$y_n(p) + \sqrt{\frac{2}{P}}\delta_1 \sin\left(\frac{1}{2}\theta_p\right), \tag{8.17}$$

where δ_1 is the size of the shift ($\delta_1 = 0.1, 0.15, 0.2, 0.25$);

bi-lobe out-of-control, which is simulated by incrementing the amplitude of the second harmonic in the baseline model as follows:

$$y_n(p) + \sqrt{\frac{2}{P}}\delta_2 \left[c_{1n}\cos\left(2\theta_p\right) + c_{2n}\sin\left(2\theta_p\right)\right], \tag{8.18}$$

where δ_2 is the increasing factor ($\delta_2 = 0.1, 0.2, 0.3, 0.4$);

tri-lobe out-of-control, which is simulated by incrementing the amplitude of the third harmonic with respect to the baseline model as follows:

$$y_n(p) + \sqrt{\frac{2}{P}}\delta_3 \left[c_{3n}\cos\left(3\theta_p\right) + c_{4n}\sin\left(3\theta_p\right)\right], \tag{8.19}$$

where δ_3 is the increasing factor ($\delta_3 = 0.1,\ 0.2,\ 0.3,\ 0.4$).

Data obtained under these Phase II models are also scaled (by subtracting the least squares estimation of the radius) and centered (on the least squares estimation of the center) before applying the profile monitoring method. In fact, we are assuming that centering and scaling are standard steps applied to data when the focus is on the geometrical errors (Cho and Tu 2001).

Performance comparison is based on the ideal assumption that the in-control parameters for each competing method are known. Indeed, computer simulation is used in our work to obtain a large dataset of in-control profiles in order to estimate as closely as possible the parameters of each method. For the two production scenarios, all simulations were conducted by first tuning each competing approach in order to achieve the same in-control ARL value of about 100. Hence, the performance in Phase II are related to the ability of detecting out-of-control profiles, given that all the approaches are designed to achieve the same false-alarm probability approximately equal to 1%.

Tables 8.4 and 8.5 summarize the simulation results for the two production scenarios under study. In particular, Table 8.4 refers to the case of random-effect model, while Table 8.5 refers to the case of fixed-effect model. Each table reports the ARLs estimated by computing a set of 1000 run lengths, given new profiles simulated

Table 8.4 Phase II Simulation Results for the Production Scenario with Random-Effect Model; ARLs and Corresponding Standard Deviations within Brackets (1000 trials)

	Delta	LOC CC	OOR CC	REG CC	PCA CC
Half-frequency	0.1	73.26 (2.18)	98.15 (3.12)	93.18 (2.83)	80.04 (2.56)
	0.15	50.53 (1.57)	97.89 (3.06)	78.68 (2.63)	61.07 (1.90)
	0.2	35.61 (1.16)	85.00 (2.71)	68.49 (2.16)	44.43 (1.37)
	0.25	22.49 (0.72)	63.51 (2.01)	48.50 (1.56)	29.05 (0.89)
Bi-lobe	0.1	64.08 (1.92)	72.39 (2.32)	76.41 (2.42)	64.29 (1.89)
	0.2	36.07 (1.16)	44.66 (1.42)	47.48 (1.60)	38.03 (1.22)
	0.3	24.01 (0.77)	25.69 (0.81)	29.76 (0.92)	21.21 (0.67)
	0.4	14.98 (0.45)	18.40 (0.57)	16.93 (0.53)	12.61 (0.38)
Tri-lobe	0.1	70.44 (2.16)	88.29 (2.80)	72.51 (2.27)	70.08 (2.19)
	0.2	43.66 (1.44)	67.87 (2.18)	47.72 (1.50)	37.55 (1.17)
	0.3	30.33 (0.94)	46.82 (1.49)	27.57 (0.85)	21.70 (0.67)
	0.4	19.43 (0.60)	33.81 (1.04)	17.39 (0.54)	12.60 (0.40)

Table 8.5 Phase II Simulation Results for the Production Scenario with Fixed-Effect Model; Actual ARLs and Corresponding Standard Deviations within Brackets (1000 trials)

	Delta	LOC CC	OOR CC	REG CC	PCA CC
Half-frequency	0.1	33.31 (1.00)	42.47 (1.36)	1.27 (0.02)	6.00 (0.17)
	0.15	12.26 (0.38)	19.71 (0.59)	1.00 (0.00)	1.51 (0.03)
	0.2	4.81 (0.13)	11.21 (0.33)	1.00 (0.00)	1.03 (0.01)
	0.25	2.37 (0.06)	5.83 (0.17)	1.00 (0.00)	1.00 (0.00)
Bi-lobe	0.1	72.11 (2.13)	58.40 (1.86)	6.75 (0.19)	45.89 (1.40)
	0.2	37.70 (1.21)	21.72 (0.69)	1.20 (0.02)	7.02 (0.20)
	0.3	21.47 (0.63)	8.87 (0.27)	1.00 (0.00)	1.71 (0.04)
	0.4	10.63 (0.32)	4.22 (0.12)	1.00 (0.00)	1.05 (0.01)
Tri-lobe	0.1	86.85 (2.72)	82.43 (2.71)	16.70 (0.50)	64.30 (1.94)
	0.2	60.26 (1.92)	38.99 (1.24)	2.44 (0.06)	20.18 (0.57)
	0.3	33.59 (1.04)	20.37 (0.67)	1.11 (0.01)	5.20 (0.15)
	0.4	21.56 (0.66)	10.42 (0.32)	1.01 (0.00)	1.87 (0.04)

according to a specific out-of-control model. Standard deviations estimated for such mean values are also included within brackets.

Since in industrial applications the analyst is not expected to know a priori what kind of out-of-control condition will affect the production process and how severe it will be, we consider a measure of the overall performance for each of the five competing approaches in each production scenario. To do this, we consider the mean ARL values for each competing approach in signaling a generic out-of-control condition of any severity for that production scenario. We are assuming that all out-of-controls previously considered are equally probable and that the analyst knows the model of monitored functional data (this is plausible when a retrospective phase of control charting has been accomplished).

Figures 8.15 and 8.16 graphically depicts the 95% Bonferroni's confidence intervals of the overall ARLs presented by the four competing approaches in each production scenarios considered in our study. A discussion on the performance comparisons is detailed in the following two subsections for the production scenarios with random-effect model and fixed-effect model, respectively.

8.4.3.1 Production Scenario with Random-Effect Model

From the results reported in Table 8.4, graphically summarized in Figure 8.15, it can be observed that the OOR CC presents the lowest power of detection when compared to the competing methods.

Similarly, the REG CC has a small power of detection in signaling out-of-controls. This may be mainly ascribed to the variability in the regression parameters that characterizes this production scenario with random effects for the baseline model. As a matter of fact, it should be noted that for the out-of-control conditions considered in our study (half-frequency, bi-lobe and tri-lobe), the majority of alarms released

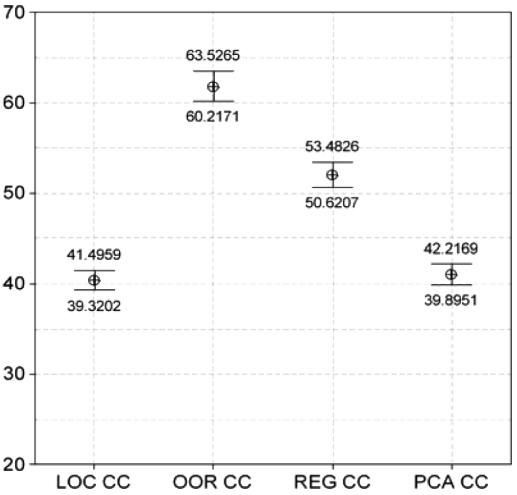

Figure 8.15 95% Bonferroni's confidence intervals of the overall ARL for the competing approaches for the production scenario with random-effect model.

by the REG CC are produced by the multivariate control chart on the vector of fitted parameters. The extra variability in the regression parameters, which naturally characterizes the vector of fitted parameters in this scenario, causes a lower detection power of the regression-based control charts, in particular of the multivariate control chart.

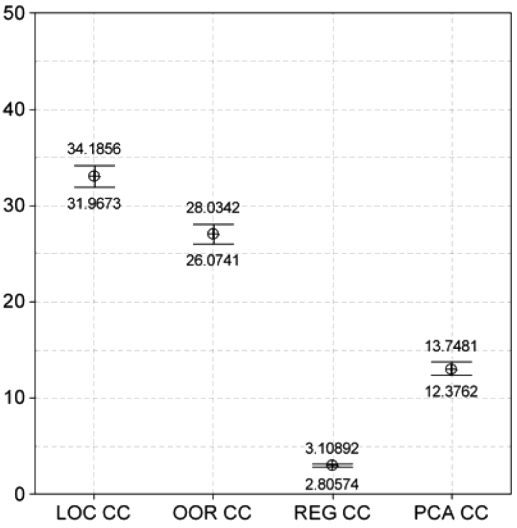

Figure 8.16 95% Bonferroni's confidence intervals of the overall ARL for the competing approaches for the production scenario with fixed-effect model.

On the other hand, it can be noted that the PCA CC presents, in many cases, outperforming performance when compared with the regression-based approach. As previously mentioned, the PCA CC consists of a multivariate control chart based on the first four retained PCs and of a univariate control chart for monitoring the residuals variability (the Q control chart).

Furthermore, from Table 8.4 it can be also noted that the LOC CC presents comparable performance to that observed for the PCA CC . In a few of cases, especially in the case of half-frequency out-of-control, surprisingly the LOC CC outperforms the PCA CC. This result shows that the simple LOC CC can be considered a valuable alternative to parametric methods for profile monitoring, at least in a production scenario with random effects.

8.4.3.2 Production Scenario with Fixed-Effect Model

In the case of a fixed-effect model as reference production scenario (results in Table 8.5 and Figure 8.16), each of the competing approach presents better performance in signaling any kind of out-of-control condition when compared with the production scenario with random-effect model. This can be easily explained by observing that no extra variability is naturally affecting the in-control profiles.

The REG CC outperforms all of the other competing methods, even if in a few cases, especially when a high severity of the out-of-control condition is considered, the regression-based and the PCA CC may have comparable performance.

Note that for the out-of-control conditions considered in our study (half-frequency, bi-lobe, and tri-lobe), the majority of alarms released by the REG CC are produced by the multivariate control chart on the vector of fitted parameters, while the PCA CC consists of a univariate control chart only (a Q control chart). As a matter of fact, when PCA is performed in the case of a fixed-effect model, no significant PCs are identified (Colosimo and Pacella 2007) as the PCA is performed after data centering and this first step consists of subtracting the mean pattern (described by the fixed-effect model) from each profile data. Thus, PCA is performed just on the error terms and hence no significant PC is correctly reported. Also, note that in this case the Q statistic is given by the sum of the squared difference between data observed at each location and the average profile at that location.

It can be noted that the LOC CC presents the lowest overall power of detection when compared to the competing methods. Also the OOR control chart does not present a better performance than those that characterize the model-based control chats (regression-based and PCA-based).

8.5 MOVING FROM 2D PROFILES TO 3D SURFACES

In the case of cylindrical features, Henke et al. (1999) presented an analytical model for aiding the interpretation of the relationship between manufacturing processes and the typical deviation from the ideal cylinder. The approach combines Chebyshev polynomials and sinusoidal functions for representing axial errors and circular errors, respectively. Subsequently, a similar approach was proposed by Zhang et al. (2005),

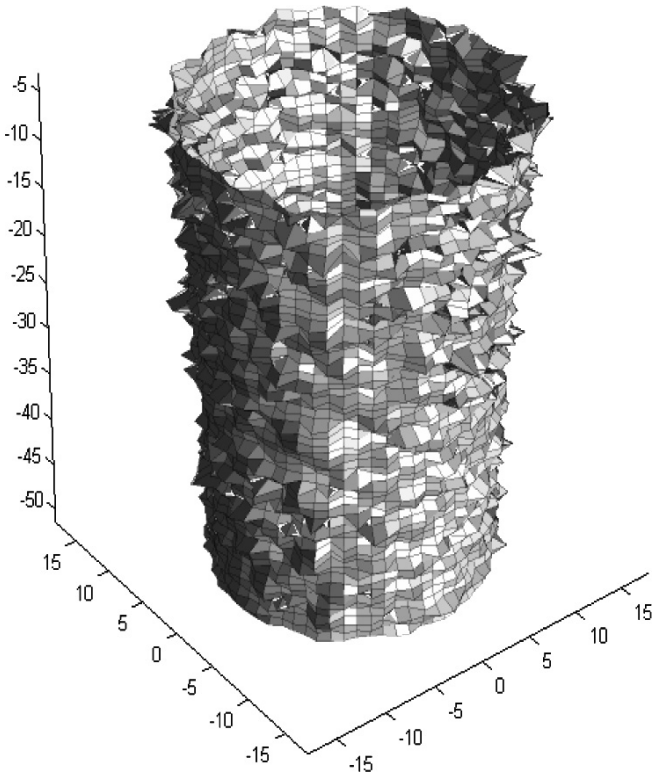

Figure 8.17 The actual shape of one out of the 100 machined cylinders.

who used the Legendre polynomials instead of the Chebyshev ones to describe the deviation along the cylinder axis, while keeping the sinusoidal functions for describing the cross-section form errors. Regression models presented in the literature on cylindrical specifications can be considered as starting reference to identify the parametric models of the large-scale pattern characterizing all the cylindrical surfaces machined. Figure 8.17 shows the shape of one real cylinder obtained by lathe-turning (Colosimo et al. 2010) while Figure 8.18 shows some typical form errors associated to machined surfaces (Henke et al. 1999 and Zhang et al. 2005).

Similarly to what was done for 2D profiles related with circularity, a parametric model for the machined cylinder can be obtained by merging a large-scale pattern

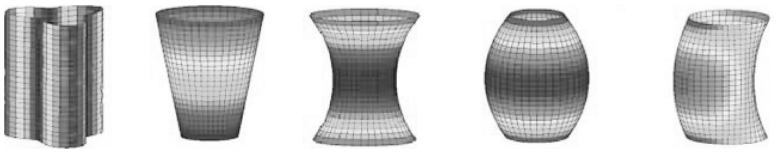

Figure 8.18 Typical form errors in manufactured cylindrical surfaces (from left to right: three-lobed, taper, hourglass, barrel, banana).

Table 8.6 The Regressor Functions as a Function of the Location Indices (ζ_q, θ_p) Shown in Figure 8.3

	Order of the Chebyshev Polynomial.	Order of the Periodic Component	Regressor
1	0	2	$T_0\left(\zeta_q\right)\cos\left(2\theta_p\right)$
2	0	2	$T_0\left(\zeta_q\right)\sin\left(2\theta_p\right)$
3	0	3	$T_0\left(\zeta_q\right)\cos\left(3\theta_p\right)$
4	0	3	$T_0\left(\zeta_q\right)\sin\left(3\theta_p\right)$
5	1	0	$T_1\left(\zeta_q\right)$
6	1	2	$T_1\left(\zeta_q\right)\cos\left(2\theta_p\right)$
7	1	2	$T_1\left(\zeta_q\right)\sin\left(2\theta_p\right)$
8	1	3	$T_1\left(\zeta_q\right)\cos\left(3\theta_p\right)$
9	1	3	$T_1\left(\zeta_q\right)\sin\left(3\theta_p\right)$
10	2	0	$T_2\left(\zeta_q\right)$
11	2	1	$T_2\left(\zeta_q\right)\cos\left(\theta_p\right)$
12	2	1	$T_2\left(\zeta_q\right)\sin\left(\theta_p\right)$
13	2	2	$T_2\left(\zeta_q\right)\cos\left(2\theta_p\right)$
14	2	2	$T_2\left(\zeta_q\right)\sin\left(2\theta_p\right)$
15	2	3	$T_2\left(\zeta_q\right)\cos\left(3\theta_p\right)$
16	2	3	$T_2\left(\zeta_q\right)\sin\left(3\theta_p\right)$

with a small-scale one. The large-scale component can be expressed as a combination of functions such as $T_i\left(\zeta_q\right)\cos\left(h\theta_p\right)$ or $T_i\left(\zeta_q\right)\sin\left(h\theta_p\right)$ i.e., as a combination of sinusoidal functions with frequency equal to h upr, multiplied by a second-type Chebyshev polynomial of order $i (i = 0, 1, 2)$, i.e., $T_0\left(\zeta_q\right) = 1$, $T_1\left(\zeta_q\right) = 2\zeta_q$, $T_2\left(\zeta_q\right) = 4\zeta_q^2 - 1$, where $\zeta_q = 2\frac{q-1}{Q-1} - 1$.

From Table 8.6, which summarizes the expression of the regressor functions, it can be observed that some of the possible combinations of the Chebyshev polynomials and the periodic functions are not included as possible regressor functions. As examples, $T_0\left(\zeta_q\right)$, representing the (least-square) cylinder radius, or $T_0\left(\zeta_q\right)\cos\left(\theta_p\right)$ and $T_0\left(\zeta_q\right)\sin\left(\theta_p\right)$, associated to the translation of the (least-square) cylinder axis, are not included in this table. Indeed, the large-scale component describes the deviation of the observed data from the ideal cylinder. By definition of form error, only the deviation from a perfect cylinder has to be modeled and this is why polynomials aimed at representing the cylinder radius or position can not be included in the cylindiricity form error model.

As for the small-scale component, i.e., the error terms, they can be considered to be uncorrelated and normally distributed. However, if the assumption of uncorrelated residuals is rejected, models able to deal with spatial correlation structures should be entailed. Colosimo et al. (2010) used approaches taken from spatial statistics to deal with the issue of possibly correlated errors. In particular, they assumed that the grid of equally spaced points observed on the machined surface is kept fixed, as it were implemented on a coordinate measuring machine that automatically performs the required measurement path. Given this fixed grid of measurement locations, each

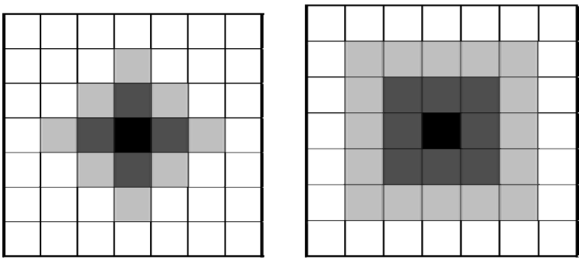

Figure 8.19 Plot of the neighbors of the rook-based (on the left) and the queen-based (on the right) contiguity: the first-order and second-order neighbors of the center point (shown in black) are represented in darker and lighter gray, respectively.

point can be related to a different set of adjacent points. Therefore, a hypothesis on the adjacency, contiguity, or spatial weight matrices $\mathbf{W}^{(s)}$ $(s = 1, \ldots, S)$ has to be firstly considered.

Two traditional ways for defining these matrices are presented in the literature on spatial statistics (Cressie 1993). The first is the rook-based contiguity, where neighbors share a common border (Figure 8.19 on the left). The second is the queen-based contiguity, which defines neighbors as locations that share either a border or a vertex in their boundaries (Figure 8.19 on the right). In practice, a spatial weights matrix is rather arbitrarily selected, especially when there is no formal theoretical foundation for the extent of spatial interaction. Despite of the specific type of contiguity structure (rook or queen-based), a SAR model of proper order can be used as reference model for the noise term of the surfaces.

8.5.1 Quality Control Charts for Surface Monitoring

In order to detect out-of-control cylindrical surfaces, two alternative approaches can be considered. The first approach is the simplest one and consists in computing the OOC value associated to each machined surface and in monitoring it with an individuals control chart.

Similarly to control charts suggested in the literature on profile monitoring, the second approach consists of a multivariate control chart for monitoring the vector of estimated parameters and a univariate control chart aimed at monitoring the estimated residuals variance.

In the paper of Colosimo et al. (2010) the method for monitoring profile data developed by Colosimo et al. (2008) was extended to 3D forms (i.e., surfaces). The proposed approach combined a regression model with spatial correlated noise to univariate and multivariate control charting and was applied to real data related to cylindricity surfaces obtained by lathe-turning. The simulation study showed that the results found in (Colosimo et al. 2008) for roundness profiles are confirmed when surfaces are machined instead of profiles. In particular, the main results were that approaches based on monitoring the coefficients of the parametric model that describes the systematic pattern of machined surface outperforms the industrial practice, which

is based on monitoring a synthetic indicator of form error (such as the OOCs). Furthermore, the simulation study showed that the extra effort required for modeling the spatial correlation can be worthy just when the small-scale (e.g., correlation structure and/or noise variance) characteristics of the machined surface are of interest.

8.6 CONCLUDING REMARKS

Data collected by measuring equipment can be modeled as functional data, where the quality outcome (dependent variable) is actually a function of one or more location variables (independent variables). In this framework, the present chapter has introduced different approaches for profile monitoring of geometric features. Both parametric models and nonparametric methods have been considered.

With the former kind of approaches, the in-control shape of the profiles is summarized by a parametric model, while profile monitoring is based on monitoring the parameters of this model (regression-based and PCA CC). On the other hand, with the latter kind of approaches, one can monitor the discrepancies between observed profiles and a baseline profile established using historical data. In this chapter, the LOC CC has been considered to this aim. A Shewhart's control chart on the geometrical errors has been also included in the comparison study, as this kind of approach is still the most representative of industrial practice.

By comparing the Phase I performance, which refers to the ability to obtain a predefined false alarm rate when uncontaminated process samples are considered, we can conclude that the extra effort required by an REG CC is worthwhile. Indeed, a regression-based control chart can easily be designed, since the actual false alarm rate can be achieved in practice using an analytic computation of the control limits, despite the specific production scenario considered and the number of historical data available. On the contrary, the actual false alarm rate presented by the other approaches for profile monitoring can be very different from that expected.

By comparing the overall performance in Phase II, which refers to the ability to signal a generic out-of-control condition of any severity for any type of production scenario (with random-effect and fixed-effect model), we can conclude that the extra effort required by the regression-based and PCA CC is also worthwhile. In fact, both the regression-based and PCA CC are more effective in signaling a generic change in the functional data. However, while the PCA CC shows superior robustness to change of the productive scenario, the REG CC may produce poor performance in signaling out-of-controls in the case of a production scenario with a random-effects model.

Furthermore, although the LOC CC is less effective in Phase II than the parametric approaches, the performance observed using this method is comparable to those produced by the competing approaches in some cases. Given the simplicity of the location control chart, its use in practice can be justified in these production situations, even if attention should be paid to properly designing (Phase I) this tool. On the other hand, the OOR CC is not suitable for profile monitoring and hence this approach is not recommended for actual applications.

Finally, in this chapter the method for monitoring 2D profile data has been also extended to 3D forms (i.e., surfaces). In this direction, there is some work (Colosimo et al. 2010) and we think that profile monitoring research will focus more and more on this topic in the near future. Although a specific geometry was used in (Colosimo et al. 2010), any 3D shape can be monitored with a similar approach, given that the regression terms and the spatial weight matrices are appropriately selected.

ACKNOWLEDGMENTS

This research work was partially funded by the Italian Ministry of University and Research (MIUR). The authors also thank the Machining Tool and Production Systems Research Laboratory (MUSP) in Piacenza (Italy) for providing the actual real case studies.

REFERENCES

Boeing Commercial Airplane Group, Quality Assurance Department, 1998. Advanced quality system tools, AQS D1-9000-1. <http:// www.boeingsuppliers.com/supplier/d1-9000-1.pdf> [Accessed March 2011].

Chang, T. C. and Gan, F. F. (2006) Monitoring linearity of measurement gauges. *Journal of Statistical Computation and Simulation*, **76**(10), 889–911.

Cho, N. W. and Tu, J. F. (2001) Roundness modeling of machined parts for tolerance analysis. *Precision Engineering*, **25**(1), 35–47.

Cho, N. W. and Tu, J. F. (2002) Quantitative circularity tolerance analysis and design for 2D precision assemblies. *International Journal of Machine Tools & Manufacture*, **42**(13), 1391–1401.

Colosimo, B. M. and Pacella, M. (2007) On the use of principal component analysis to identify systematic patterns in roundness profiles. *Quality and Reliability Engineering International*, **23**(6), 707–725.

Colosimo, B. M., Pacella, M., and Semeraro, Q. (2008) Statistical process control for geometric specifications: On the monitoring of roundness profiles. *Journal of Quality Technology*, **40**(1), 1–18.

Colosimo, B. M., Mammarella, F., and Petrò, S. (2010) Quality control of manufactured surfaces. In: *Frontiers of Statistical Quality Control 9*, edited by H. J. Lenz and P. T. Wilrich. Springer, New York.

Colosimo, B. M. and Pacella, M. (2010) A comparison study of control charts for statistical monitoring of functional data. *International Journal of Production Research*, **48**(6), 1575–1601.

Cressie, N. A. C. (1993) *Statistics for Spatial Data*, Revised edition. John Wiley & Sons, Inc., New York.

Ding, Y., Zeng, L., and Zhou, S. (2006) Phase I analysis for monitoring nonlinear profiles in 755 manufacturing processes. *Journal of Quality Technology*, **38**(3), 199–216.

Forbes, B. (1994) *Algorithms for Geometrical Tolerance Assessment. Report DITC 228/94*. National Physical Laboratory, UK.

Gupta, S., Montgomery, D. C., and Woodall, W. H. (2006) Performance evaluation of two methods for online monitoring of linear calibration profiles. *International Journal of Production Research*, **44**(10), 1927–1942.

Henke, R. P., Summerhays, K. D., Baldwin, J. M., Cassou, R. M., and Brown, C. W. (1999) Methods for evaluation of systematic geometric deviations in machined parts and their relationships to process variables. *Precision Engineering*, **23**(4), 273–292.

Jackson, J. E. (2003) *A User's Guide to Principal Components*. John Wiley & Sons, Inc., New York.

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2008) Monitoring correlation within linear profiles using mixed models. *Journal of Quality Technology*, **40**(2), 167–183.

Jin, J. and Shi, J. (1999) Feature-preserving data compression of stamping tonnage information using wavelets. *Technometrics*, **41**, 327–339.

Jin, J. and Shi, J. (2001) Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *Journal of Intelligent Manufacturing*, **12**, 257–268.

Jolliffe, I. T. (2002) *Principal Component Analysis*, 2nd Edition. Springer, Springer Series in Statistics, New York.

Kang, L. and Albin, S. L. (2000) On-line monitoring when the process yields a linear profile. *Journal of Quality Technology*, **32**, 418–426.

Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003) On the monitoring of linear profiles. *Journal of Quality Technology*, **35**, 317–328.

LeSage, J. P. (1999) *The Theory and Practice of Spatial Econometrics*. Available at http://www.spatialeconometrics.com/.

Mahmoud, M. A. and Woodall, W. H. (2004) Phase I analysis of linear profiles with calibration applications. *Technometrics*, **46**, 377–391.

Mahmoud, M. A., Parker, P. A., Woodall, W. H., and Hawkins, D. M. (2007) A change point method for linear profile data. *Quality & Reliability Engineering International*, **23**(2), 247–268.

Montgomery, D. C. (2004) *Introduction to Statistical Quality Control*, 5th Edition. John Wiley & Sons, Inc., New York.

Moroni, G. and Pacella, M. (2008) An approach based on process signature modeling for roundness evaluation of manufactured items. *Journal of Computing and Information Science in Engineering*, **8**(2), 021003-1–021003-10.

Nomikos, P. and MacGregor, J. F. (1995) Multivariate SPC charts for monitoring batch processes. *Technometrics*, **37**, 41–59.

Ramsay, J. O. and Silverman, B. W. (2005) *Functional Data Analysis*, 2nd Edition. Springer, Springer Series in Statistics, New York.

Simes, R. J. (1986) An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, **73**(3), 751–754.

Stover, F. S. and Brill, R. V. (1998) Statistical quality control applied to ion chromatography calibrations. *Journal of Chromatography, A*, **804**, 37–43.

Walker, E. and Wright, S. P. (2002) Comparing curves using additive models. *Journal of Quality Technology*, **34**, 118–129.

Whittle, P. (1954) On stationary processes in the plane. *Biometrika*, **41**, 434–449.

Williams, J. D., Woodall, W. H., Birch, J. B., and Sullivan, J. H. (2006) Distribution of Hotelling's T^2 statistic based on the successive differences estimator. *Journal of Quality Technology*, **38**, 217–229.

Williams, J. D., Woodall, W. H., and Birch, J. B. (2007) Statistical monitoring of nonlinear product and process quality profiles. *Quality and Reliability Engineering International*, **23**(8), 925–941.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004) Using control charts to monitor process and product quality profiles. *Journal of Quality Technology*, **36**(3), 309–320.

Woodall, W. H. (2007) Current research on profile monitoring. *Produção*, **17**(3), 420–425.

Young, T. M., Winistorfer, P. M., and Wang, S. (1999) Multivariate control charts of MDF and OSB vertical density profile attributes. *Forest Products Journal*, **49**, 79–86.

Zhang, H. and Albin, S. (2009) Detecting outliers in complex profiles using a 2 control chart method. *IIE Transactions*, **41**(4), 335–345.

Zhang, X. D., Zhang, C., Wang, B., and Feng, S. C. (2005) Unified functional tolerancing approach for precision cylindrical components. *International Journal of Production Research*, **43**(1), 25–47.

Zou, C. L., Zhang, Y. J., and Wang, Z. J. (2006) A control chart based on a change-point model for monitoring linear profiles. *IIE Transactions*, **38**(12), 1093–1103.

Zou, C. L., Tsung, F. G., and Wang, Z. J. (2008) Monitoring profiles based on nonparametric regression methods. *Technometrics*, **50**(4), 512–526.