

Statistical performance of tests for factor effects on the shape of objects with application in manufacturing

HUSSAM ALSHRAIDEH¹ and ENRIQUE DEL CASTILLO^{2,*}

¹*Department of Industrial and Manufacturing Engineering, Jordan University of Science and Technology, Irbid 22110, Jordan*

²*Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA 16802, USA*

E-mail: exd13@psu.edu

Received February 2010 and accepted February 2012

This article considers experiments in manufacturing where the response of interest is the geometric shape of a manufactured part and the goal is to determine whether the process settings varied during the experiment affect the resulting shape of the part. An approach in practice to determine factor effects is to estimate the form error of the part—if a standard definition of the form error of interest exists—and conduct an analysis of variance (ANOVA) on the form errors. Instead, we study the performance of several statistical shape analysis techniques to analyze this class of experiments. Simulated shape data were used to perform power comparisons for two- and three-dimensional shapes. The ANOVA on the form errors was found to have a poor performance in detecting mean shape differences in circular and cylindrical shapes. Procrustes-based tests such as an ANOVA test due to Goodall and a recently proposed ANOVA permutation test provide the highest power to detect differences in the mean shape. These tests can also be applied to parts produced in “free form” manufacturing, where no standard definition of form error exists, provided that correspondent points exist on each part.

Keywords: Statistical shape analysis, form error, permutation tests, free-form manufacturing, cylindricity, circularity

1. Introduction

We consider manufacturing experiments where the end purpose is to optimize the geometric shape of the parts produced, and the first step toward this goal is to determine whether any of the process conditions varied in the experiment affect the resulting shape of the parts. This is a situation similar to that in the classic Design of Experiments (DOE) approach, with the additional feature that the response of the process is the complete part geometry. Our interest, then, is *not* part inspection, which is typically performed based on the tolerance form errors, provided that a standard definition of the particular form error of interest exists, but—just as in classic DOE—our aim is process characterization and optimization of the shape of a part as a function of the process parameters. This is an activity typically performed at the R&D or engineering design stage, prior to regular production. In this article, we review and contrast statistical methods used to detect factor effects in the mean shape of an object. As it will be shown, statistical shape analysis methods, recently proposed for use in manufacturing (Del Castillo and Colosimo, 2011),

provide a powerful approach to this type of problem, especially in cases where a standard definition of the form error is non-existent (e.g., “free form” manufacturing).

For an instance of a designed experiment of the kind we wish to analyze, Del Castillo and Colosimo (2011) reported an experiment in lathe-turning of titanium alloy parts in which the depth of cut and cutting speed were varied according to a factorial design. Measurements of the two-dimensional shapes were obtained using a Coordinate Measuring Machine (CMM). The goal of the experiment was to determine the effect of depth and cutting speed on the circularity of the parts and to determine the best settings of these factors to achieve the most circular parts. A standard analysis of the effect of these factors on the circularity (or cylindricity) of the parts can be conducted through an analysis of variance (ANOVA) of the form error for roundness (or cylindricity) of each part. This error is frequently calculated in tolerancing practice as the smallest difference between the radii of the two coaxial circles (or cylinders) that enclose all of the measurements in a part (Krulikowski, 1996; Henzold, 2006); please refer to Fig. 1(a)). Provided that a standard definition of form error exists, an ANOVA can then be conducted on the form errors to analyze the impact the factors have on the mean shape of the parts (for an interesting paper on this approach, see Colosimo and Pacella (2011)).

*Corresponding author

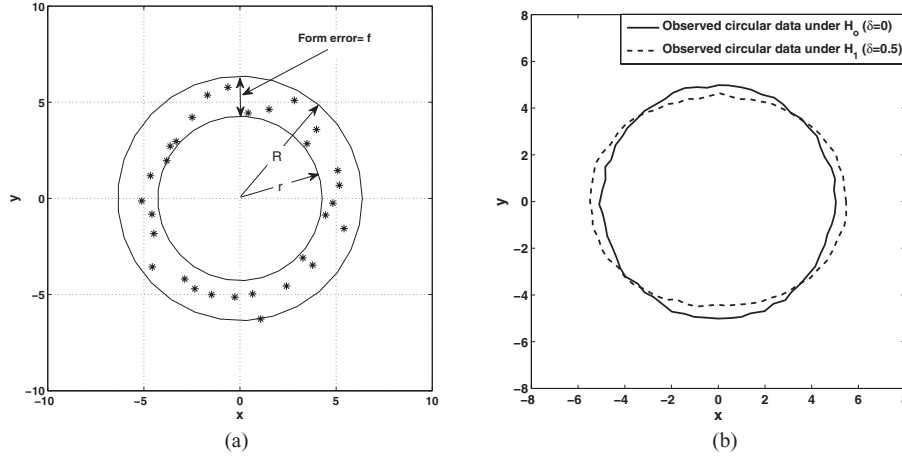


Fig. 1. (a) Circularity form error calculation and (b) circular and bilobed shapes with parameters $\sigma = 0.05$ and $r = 5$.

As an alternative to the use of the form errors and a standard ANOVA test, or for parts with complex geometry for which no standard form error definition exists (i.e., free-form manufactured parts), we can use Statistical Shape Analysis (SSA) techniques to perform the analysis of the experiment. This was suggested by Del Castillo and Colosimo (2011), who provided some limited evidence that a specific SSA technique, a Procrustes-based permutation ANOVA, has significantly more power to detect differences in the (two-dimensional) roundness of lathe-turned parts than the traditional ANOVA method based on the roundness form errors. SSA techniques have been widely used over the last 15 years in applications other than manufacturing; for example, in the analysis of shapes of objects of paleontological, biological, or geological interest or for text recognition in computer vision. For an introduction to SSA with emphasis on these non-manufacturing applications, see the book by Dryden and Mardia (1998). For more mathematical treatments, see the books by Small (1996) and Kendall *et al.* (1999).

The general problem can be stated as follows. Suppose we wish to compare the mean shapes of a groups of n parts each. The groups of parts relate to different levels of a factor varied in a manufacturing experiment; e.g., depth of cut. Let \mathbf{X}_{ij} be a $k \times m$ matrix of coordinate measurements (with dimension $m = 2$ or 3) representing each part j in group i , which we will refer to as a *configuration matrix*. In practice, the coordinates are usually gathered via a CMM or laser scanners (see Barcenas and Griffin (2001)). A labeled coordinate that corresponds between parts is called a *landmark* in the SSA literature. In this article, we assume that all configurations have correspondent landmarks (a review of matching algorithms for finding such correspondence can be found in Davies *et al.* (2008)). Assume the model:

$$\mathbf{X}_{ij} = \boldsymbol{\mu}_i + \mathbf{E}_{ij} \quad i = 1, 2, \dots, a \quad j = 1, 2, \dots, n, \quad (1)$$

where $\boldsymbol{\mu}_i = \boldsymbol{\mu} + \boldsymbol{\tau}_i$, with dimension $k \times m$, represent the mean shape at the i th level and \mathbf{E}_{ij} is a $k \times m$ matrix of

errors such that $\text{vec}(\mathbf{E}_{ij}) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is a $km \times km$ covariance matrix and $\text{vec}(\cdot)$ is the operator that concatenates the columns of a matrix into one vector. This is just a one-way ANOVA on a matrix response. In this case, it is of interest to test the significance of the factor levels effect; that is,

$$H_0 : \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \dots = \boldsymbol{\tau}_a = \mathbf{0}$$

versus the alternative

$$H_1 : \text{at least a } \boldsymbol{\tau}_i \neq \mathbf{0}.$$

If H_0 is true, all of the a groups have the same mean shape $\boldsymbol{\mu}$ (to be precise, this should be denoted by $[\boldsymbol{\mu}]$, a set equal to an equivalence relation modulo rotations, since rotations are arbitrary; see Small (1996) and Section 2.1 in this article. To simplify notation, we will simply use $\boldsymbol{\mu}$ to refer to the mean shape hereafter.) Analogous models and hypotheses can be set for two-way experimental layouts for shape responses. We point out that a multivariate ANOVA (MANOVA) test cannot be applied, given that in manufacturing the number of measurements k is typically very large compared to the sample size (number of parts). The goal is to study the performance of tests for H_0 in the case where the shapes are two- or three-dimensional.

The rest of this article is organized as follows. Model assumptions are explained in Section 2. Section 3 reviews the statistical tests for shape differences with which we will be concerned. None of the limited existing studies about the performance of these tests refers to shapes with a geometry of interest in manufacturing (for a review of these existing performance studies and further comparisons of the tests studied in this article for mean shapes *not* frequently occurring in manufacturing, see Alshraideh (2011)). In Section 4 we study the performance of the shape difference tests for two shapes that frequently occur in manufactured parts, namely, two-dimensional circular and three-dimensional cylindrical shapes. The performance comparisons in Section 4 deal with both the *isotropic* error structure case—i.e., the case where the variances at each landmark and

each dimension are the same—that is, $\Sigma = \sigma^2 \mathbf{I}_{km}$ —and the *anisotropic* (or *non-isotropic*) error structure. The performance of these tests under non-normal errors is also investigated. The article concludes with a summary and recommendations in Section 5.

2. Model assumptions

2.1. Kendall's shape space

The tests contrasted in this article originate from the work by Kendall (1984) on the statistical analysis of similarity shapes based on landmarks. Kendall's representation of shape is based on the aforementioned configurations $\mathbf{X} \in \mathbb{R}^{k \times m}$ from which the effects of similarity transformations (translations, rotations, and dilatations or changes of scale) are filtered out. The shape of an object is thus the geometric information that remains after we filter out the effect of similarity transformations. In this theory the so-called *pre-shapes* are unit-scaled and centered configurations (denoted by $\mathbf{Z} \in \mathbb{R}^{(k-1) \times m}$) that can be understood as lying on the surface of a hypersphere of unit radius, the preshape space S_m^k . The shape space Σ_m^k is obtained from the quotient $S_m^k / SO(m)$ where $SO(m)$ is the special orthogonal group; i.e., the space of all rotation matrices excluding reflections. All of the elements along an orbit on S_m^k have the same shape; i.e., the shape of an object is an equivalence class (a set) defined in the preshape space under rotations.

The mapping from preshape to shape space is a Riemannian immersion, where orbits (or fibers) on S_m^k map into points on Σ_m^k . The shape space is in general a non-linear Riemannian manifold (for a thorough mathematical exploration of these spaces, see Kendall *et al.* (1999)), and the statistical analysis of shapes is hence related to statistical analysis on manifolds, a field of increasing interest in statistics and computer vision. In the case where the dimension m is equal to 3, this manifold contains singularities since there can exist rotations that leave the preshape unchanged. This happens when the landmarks are all collinear in some direction. Since this will not be the case in manufacturing data, we safely assume non-degenerate configurations.

The case where the shapes greatly differ (or vary) requires us to consider the curvature of Kendall's shape space, since this contradicts the Euclidean space assumed by Pythagoras' theorem, which is the basis of ANOVA and principal component analysis. Techniques for this case have recently been proposed by Huckemann *et al.* (2010a, 2010b). We make the assumption in what follows that the variability of the shapes and the distance between the mean shapes is small, a plausible assumption in discrete-part manufacturing. As emphasized by Dryden and Mardia (1998), a local linear approximation to shape space, as implied in the ANOVA techniques contrasted in this article, suffices if shapes vary little; i.e., when data are concentrated.

2.2. Assumed model

The coordinates of the different observed configuration matrices \mathbf{X}_{ij} in Equation (1) may be oriented or located in space differently or may have different scales. Some type of alignment or *registration* pre-processing is necessary before conducting a test for mean shape difference. The most widely used shape registration method in SSA is the Generalized Procrustes Algorithm (GPA), a method well known in multivariate statistics used to compare data matrices, developed by Gower (1975) and Ten Berge (1977). If β , Γ , and γ are a scaling factor, an orthogonal rotation matrix, and a translation vector, respectively, then the usual model assumed to generate the data is the so-called perturbation model (see, e.g., Dryden and Mardia (1998)):

$$\mathbf{X} = \beta(\mu + \mathbf{E})\Gamma + \mathbf{1}_k \gamma^T \quad (2)$$

where $\mathbf{1}_k$ is a $k \times 1$ vector of ones and matrix \mathbf{E} is as before. The perturbation model indicates that the observed configurations are the result of similarity transformations applied to the mean shape μ after it is perturbed with additive noise \mathbf{E} . GPA estimates the mean shape μ from a sample of n objects that may have different scales, orientations, and locations in space and, as a by-product, registers (aligns) the objects in the sample. GPA minimizes the sum of squares of all pairwise full Procrustes distances, $d_F(\mathbf{X}_i, \mathbf{X}_j)$, where:

$$d_F^2(\mathbf{X}_i, \mathbf{X}_j) = \min_{\beta_i, \Gamma_i, \gamma_i} \|\beta_i \mathbf{X}_i \Gamma_i + \mathbf{1}_k \gamma_i' - (\beta_j \mathbf{X}_j \Gamma_j + \mathbf{1}_k \gamma_j')\|^2.$$

This registration method assumes isotropic variances for all objects. A modified GPA algorithm called GPA(Σ) (Goodall, 1991) has been suggested to find the mean shape when the error structure is non-isotropic. The tests in Section 3 all assume both a small variance σ^2 and a small difference between mean effects; thus, a linear approximation to Kendall's shape space is adequate.

Goodall (1991) suggested to decompose the variance-covariance matrix of the errors Σ into two parts, the $k \times k$ landmarks covariance matrix Σ_K and the $m \times m$ dimension covariance matrix Σ_D , which allows us to model different variability along each axis or dimension. Then Σ is the Kronecker product of Σ_K and Σ_D : $\Sigma = \Sigma_K \otimes \Sigma_D$. We will adopt this separable covariance structure in some of the simulated data tests in Section 4. In Section 4.3, an *anisotropic* covariance model (Ecker and Gelfand, 1999) will be considered as a variation of the covariance matrix Σ_K . The model uses an $m \times m$ positive definite matrix \mathbf{B} to account for the unequal correlations and can be written as

$$(\Sigma_K)_{ij} = \sigma^2 \exp(-\phi(\mathbf{d}_{ij}' \mathbf{B} \mathbf{d}_{ij})^{1/2}) + \tau^2, \quad \sigma^2 > 0, \phi > 0, \tau^2 > 0, \quad (3)$$

where $(\Sigma_K)_{ij}$ is the ij th element of Σ_K and \mathbf{d}_{ij} is an $m \times 1$ vector of Euclidean distances along each of the m dimensions between landmarks i and j . This model provides geometric anisotropy since the correlation between any two

landmarks depends on the separation vector \mathbf{d}_{ij} rather than merely on its length.

3. Tests for detecting a difference in mean shape

We now review the tests that will be evaluated in Section 4.

3.1. ANOVA F test and ANOVA permutation test on mean shape differences

Consider the one-way ANOVA model (1). The expectation of \mathbf{X}_{ij} can be written as

$$E[\mathbf{X}_{ij}] = \boldsymbol{\mu} + \boldsymbol{\tau}_i, \quad i = 1, \dots, a \quad j = 1, \dots, n, \quad (4)$$

where $\boldsymbol{\tau}_i$ is the factor effect at level i . Let $\bar{\mathbf{X}}_{i\bullet}$ be the sample mean of the registered configurations of the i th group, and let $\bar{\mathbf{X}}_{\bullet\bullet}$ be the grand sample mean of all registered configurations. Goodall (1991) showed that the statistic

$$F_0 = n(n-1)a \frac{\sum_{i=1}^a d_F^2(\bar{\mathbf{X}}_{i\bullet}, \bar{\mathbf{X}}_{\bullet\bullet})}{(a-1) \sum_{i=1}^a \sum_{j=1}^n d_F^2(\mathbf{X}_{ij}, \bar{\mathbf{X}}_{i\bullet})} \quad (5)$$

follows approximately an F distribution with $(a-1)M$ and $a(n-1)M$ degrees of freedom for small σ , where $M = (k-1)m - 1 - m(m-1)/2$.

A more robust alternative is to test H_0 using a permutation test (see, for example, Edgington (1995)). Del Castillo and Colosimo (2011) extended Goodall's one-way ANOVA to a two-way ANOVA test for shapes. They also proposed a two-way ANOVA permutation test for shape responses and presented methods to visualize the effects of the factors and their interaction on the shape of the object.

3.2. Distance-based tests

Lele and Richtsmeier (1991) proposed a two-sample test called EDMA-I (for Euclidean Distance Matrix Analysis) for determining differences in mean shape. The EDMA-I test statistic is defined as: $T = \max \mathbf{D}_{ij}(E(\mathbf{X}), E(\mathbf{Y})) / \min \mathbf{D}_{ij}(E(\mathbf{X}), E(\mathbf{Y}))$ where \mathbf{D} is the *average form difference matrix* of the two configurations \mathbf{X} and \mathbf{Y} and $E(\cdot)$ is the expectation operator. See Lele and Richtsmeier (1991) for more details. Lele and Cole (1996) proposed an alternative test, EDMA-II, for shape differences that does not require equal variance-covariance matrices. The EDMA-II test statistic is defined as: $Z = \min(\mathbf{S}_1 - \mathbf{S}_2)$ or $\max(\mathbf{S}_1 - \mathbf{S}_2)$, whichever is most different from zero, where \mathbf{S}_1 and \mathbf{S}_2 are the *mean shape matrices* (Lele and Cole, 1996).

4. Performance analysis for shapes with a geometry of interest in manufacturing

4.1. Detection of differences in mean shape for two-dimensional circular parts

Given the pre-eminence of lathe and drilling processes in discrete-part practice, parts with a circular shape are of utmost importance in manufacturing. Process settings such as cutting speed, feed rate, or some type of machine conditioning may have an effect on the circularity of these parts. A simulation study was conducted to determine the performance of the tests considered in the previous section for determining differences in mean shape for circular shapes. In this case, a standard ANOVA on the (circularity) form errors is possible and was included in the comparisons as well. International standards exist for measuring form errors in circular (or cylindrical) geometries. These are frequently calculated in tolerancing practice as the smallest difference between the radii of the two coaxial circles (or cylinders) that enclose all the measurements in a part (Krulikowski (1996), Henzold (2006); also see Fig. 1(a)). A standard one-way ANOVA is then used to test for the equality of mean form error.

Simulated circular shape data that are less circular as the level of a single factor increases was generated for these comparisons. As the value of the factor changes from low to high, a second harmonic with amplitude δ was added to the circular configurations to simulate bilobed shapes (a common problem in lathe machining; see Fig. 1(b)). Amplitude values were chosen such that $\delta = w\sigma/r$, where σ is the noise (error) standard deviation, r is the radius of the true circle (set equal to five), and w is the “non-circularity” parameter $w = \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$ we desire the tests to detect. Figure 1(b) shows the mean shapes for two values of δ .

In all tests shown in this and later sections, the basic simulation setup was as follows: 20 configurations per sample were simulated and each circle consisted of $k = 64$ landmarks. One hundred replications and 100 permutations (for tests that require them) were conducted. In all cases, the perfect circle with respect to which we wish to detect deviations had a radius equal to $r = 5$. In all tests in this and later sections, the test size α was fixed at 0.05.

The simulation results for the case of normally distributed errors ($N(0, 0.05^2)$ and independent and identically distributed errors were added at each coordinate) are shown in Fig. 2. For this case, the power curves for the ANOVA F -test and ANOVA permutation test for shapes and the standard ANOVA on the form error were found previously by Del Castillo and Colosimo (2011). Comparing the curves, it is seen that the ANOVA permutation test for shape responses possesses power as good as the ANOVA F -test, whereas the EDMA-I test has very low power for this type of shapes. The standard ANOVA on the form errors performed considerably worse than the ANOVA F -test and

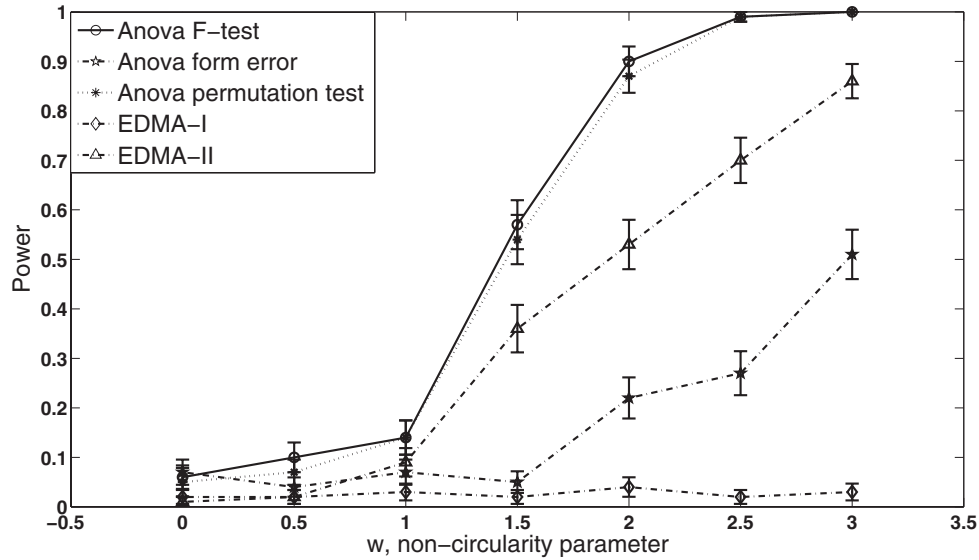


Fig. 2. Power curves using two-dimensional circular shapes. The graph shows the probability of detecting a non-circular shape using the different tests in Section 3, parameterized as a function of the non-circularity parameter w , which makes the mean shape more ellipsoidal (see text).

the ANOVA permutation test for shape responses, clearly indicating the superiority of the latter over the former.

4.2. Detection of differences in mean for three-dimensional cylindrical parts

The tests of the previous sections were also compared for three-dimensional (3D) objects. A common instance of a 3D shape of interest in manufacturing is a cylinder (see, for example, Traband *et al.* (2004)). For example, in lathe machining, different problems in the process can result in non-cylindrical shapes in the form of a barrel, a banana, etc. (see Colosimo *et al.* (2007)).

We consider the barrel shape in our performance study of tests for shape difference (see Fig. 3) as a deviation from a cylindrical shape one wishes to detect. The study consisted of two samples: the first sample contains parts with

a cylindrical mean shape of radius r and height h ; the parts in the second sample have a barrel mean shape with radius $r_1 = r$ at both the top and the bottom and radius $r_2 > r_1$ at the middle of the cylinder, which has height h . In analogy to the circularity case, we refer to the difference $\delta = r_2 - r_1$ as the amplitude. The change from r_1 to r_2 along the height was considered to be a sinewave with amplitude δ and a period of $2h$. The amplitude δ can be calculated as $\delta = w\sigma/r_1$ where w is a “non-cylindricity” parameter, in analogy to the non-circularity parameter described in Section 4.1. Figure 3 shows the mean shape of the two samples when $\delta = \{0, 1\}$.

Twenty ($n = 20$) parts were simulated at each value of the non-cylindricity parameter $w = \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$. A total of 320 landmarks per configuration were simulated, with 64 landmark circles generated at heights $\{0, h/4, h/2, 3h/4, h\}$. Independent and identically

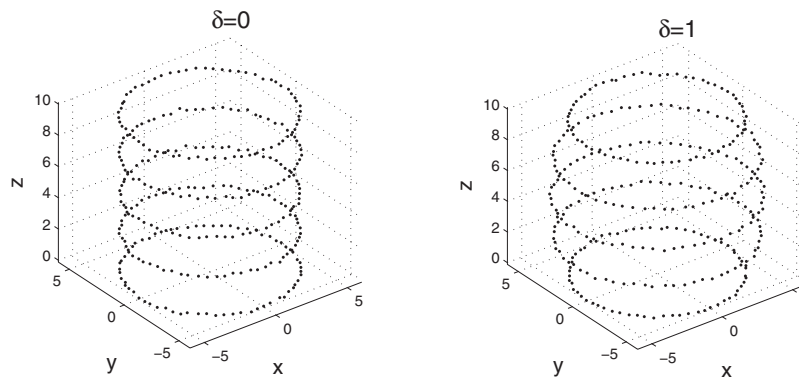


Fig. 3. Sketch of two different mean shapes of cylindrical parts with $r = 5$ and $h = 10$. Left: perfect cylinder ($\delta = 0$), right: “barrel” ($\delta = 1$).

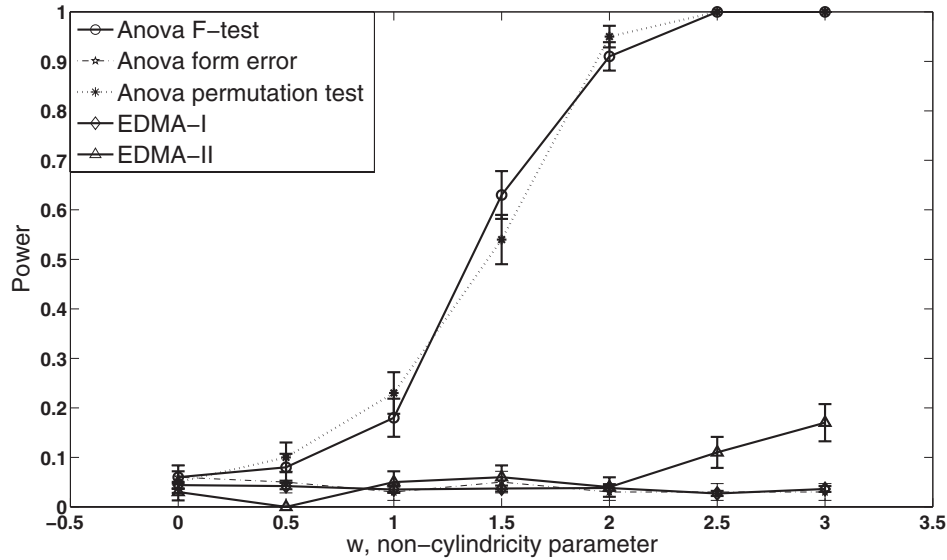


Fig. 4. Statistical power for detecting a difference in mean shape for cylinders. The figure shows the probability of detecting a non-cylindrical shape using the different tests in Section 3, parameterized as a function of the non-cylindricity parameter w , which makes the mean shape more “barrel” shaped (see text).

distributed $N(0, 0.05^2)$ errors were added to each coordinate. The power performance results are shown in Fig. 4. Comparing the power curves it can be said that the ANOVA F -test and the ANOVA permutation test are the most powerful and control Type-I error well, whereas the EDMA-I, EDMA-II, and a standard ANOVA on the cylindricity form error (computed analogously to the circularity error by using two concentric cylinders; see Krulikowski (1996)) all have very low power.

4.3. Performance under non-isotropic variance

To assess behavior under non-isotropic errors, the covariance model described by Equation (3) was used with $\sigma = \phi = \tau = 0.05$. Matrix \mathbf{B} was chosen to include non-isotropic correlations (see Ecker and Gelfand (1999)). For the two-dimensional (2D) circular shapes and the 3D cylindrical shapes, \mathbf{B} was set at

$$\mathbf{B} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1.5 \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1.5 & 0.5 \\ 0.5 & 0.5 & 2 \end{pmatrix},$$

respectively. Correlated $N(\mathbf{0}, \Sigma)$ errors with covariance matrix $\Sigma = \Sigma_K \otimes \Sigma_D$ with $\Sigma_D = \mathbf{I}$ were generated and added to each landmark. Power results for all five tests of interest are shown in Figs. 5 and 6.

Figure 6 shows that when unequal variances or correlations are present in the error structure, the ANOVA F -test and the ANOVA permutation test on shapes lose some of

their power. However, these tests still have a higher power than the distance-based and the ANOVA on the form errors. It is also noticeable that these two tests appear to possess higher power for the 3D case than the 2D (circular) case (5). The difference in power is actually due to a larger number of landmarks ($k = 320$) in the cylindrical case than in the circular case, where $k = 64$. This was confirmed by Alshraideh (2011), who conducted power comparisons with increasing number of landmarks under different non-isotropic models. For correlation model (3), increasing the number of landmarks decreases the correlations between landmarks since these are a function of the inter-landmark distances, and this benefits the ANOVA tests on shapes since they assume independence.

4.4. Performance under non-normal errors

The results in previous sections show how the two ANOVA tests for shapes provide higher power than all other tests discussed in Section 3. However, the ANOVA tests that use an F statistic are based on the assumption of normally distributed errors, an assumption that may not hold in practice. Here we consider the robustness of all of the ANOVA tests described in Section 3 (ANOVA F -test and ANOVA permutation test for shapes and a standard ANOVA on the form errors) under two cases of non-normal errors: a uniform distribution, and a t -distribution with five degrees of freedom. A simulation study similar to the one in Section 4.1 was run except that the added errors were distributed as uniform (a, b) or $t_5(0, \sigma^2)$. The values of a and b were chosen such that the mean is zero and the variance is σ^2 . This yields $b = -a$, $a = -\sqrt{3}\sigma$, and

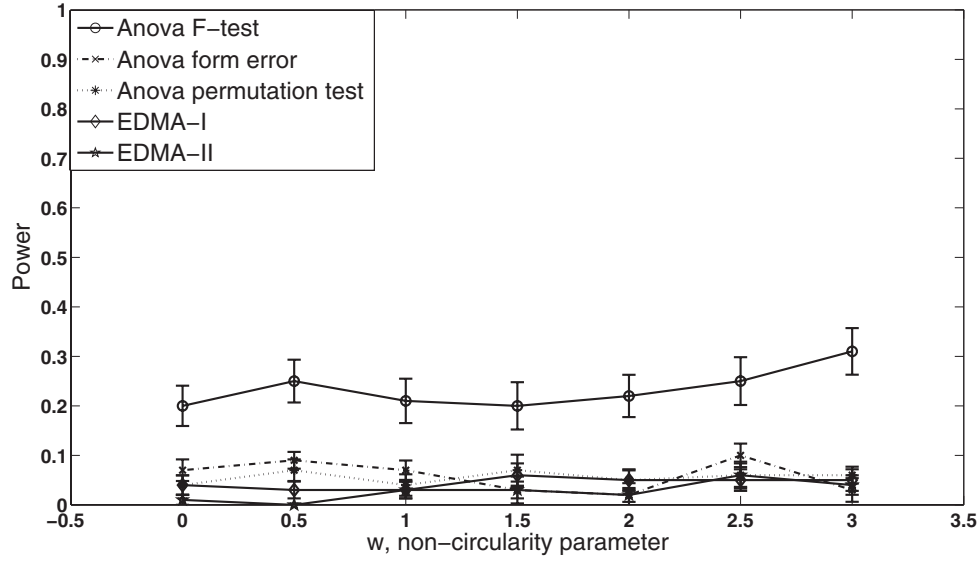


Fig. 5. Statistical power for detecting changes in 2D circular shapes under exponential non-isotropic errors as described by Equation (3).

$b = \sqrt{3}\sigma$. The fitted kernel density functions for the distributions of form errors under the ideal case of normal errors ($\text{vec}(\mathbf{E}) \sim N(\mathbf{0}, 0.05^2 \mathbf{I}_{km})$), the case of uniform errors ($\text{vec}(\mathbf{E}) \sim \text{uniform}((-0.05\sqrt{3})\mathbf{1}_{km}, (0.05\sqrt{3})\mathbf{1}_{km})$), and the case of t -distributed errors ($\text{vec}(\mathbf{E}) \sim t_5(\mathbf{0}, 0.05^2 \mathbf{I}_{km})$) are shown in Fig. 7. This figure shows how the variance of the distribution of the form errors increases as the tails in the distribution of the errors become thicker.

The estimated power curves for this case are shown in Figs. 8 and 9 (here, results for $k = 320$ landmarks are reported). It is seen that the ANOVA F -test and the

ANOVA permutation test for shape responses still show the best relative performance. However, the ANOVA F -test does not have the advertised Type-I error rate of 5%. The standard ANOVA on the circularity form errors has a good power performance under uniform errors but shows a poor performance under t -distributed errors. This is not a surprising result since the circularity form error is a range statistic and hence it will be more sensitive to a distribution with tails that can generate large errors than to a bounded distribution of the errors.

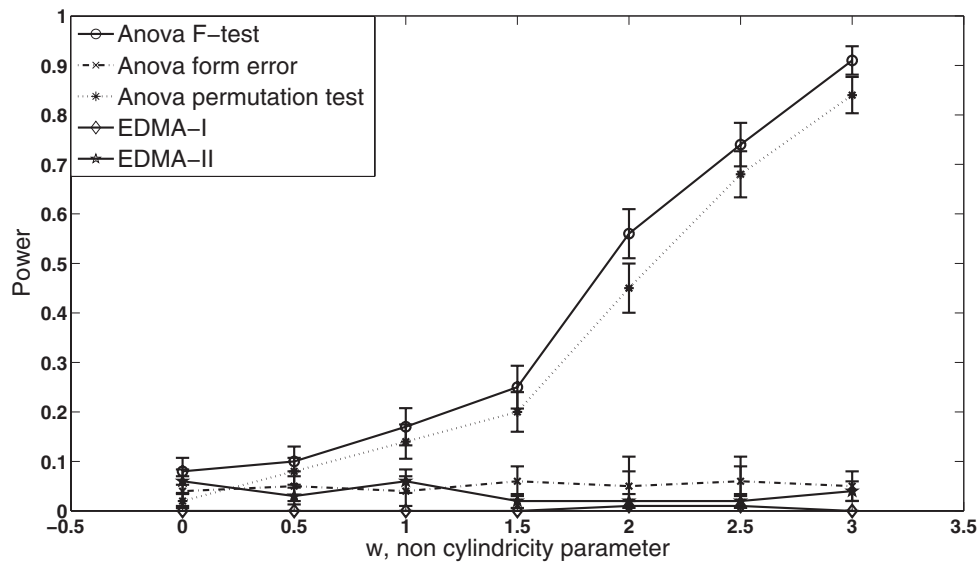


Fig. 6. Statistical power for detecting changes in 3D cylindrical shapes under exponential non-isotropic errors as described by Equation (3).

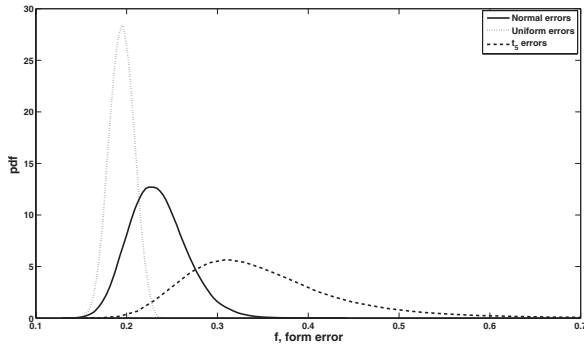


Fig. 7. The fitted kernel density functions from 100 000 simulated circularity form errors ($k = 64$, isotropic errors).

4.5. More reasons for using a shape test instead of the form errors to analyze a shape response experiment

The previous simulation results clearly indicate that the ANOVA tests for shape responses are the most powerful to detect effects and, in particular, the ANOVA permutation test for shapes is also the most robust under a variety of circumstances for both circular and cylindrical shapes. An additional reason why the SSA tests reviewed in Section 3 are preferable over using a standard ANOVA on the circularity (or cylindricity) form errors is that by using the latter, a change in shape can go undetected when in fact the underlying shape changes considerably. This can happen when the form error remains constant, yet a change in shape has occurred that, if detected, provides process

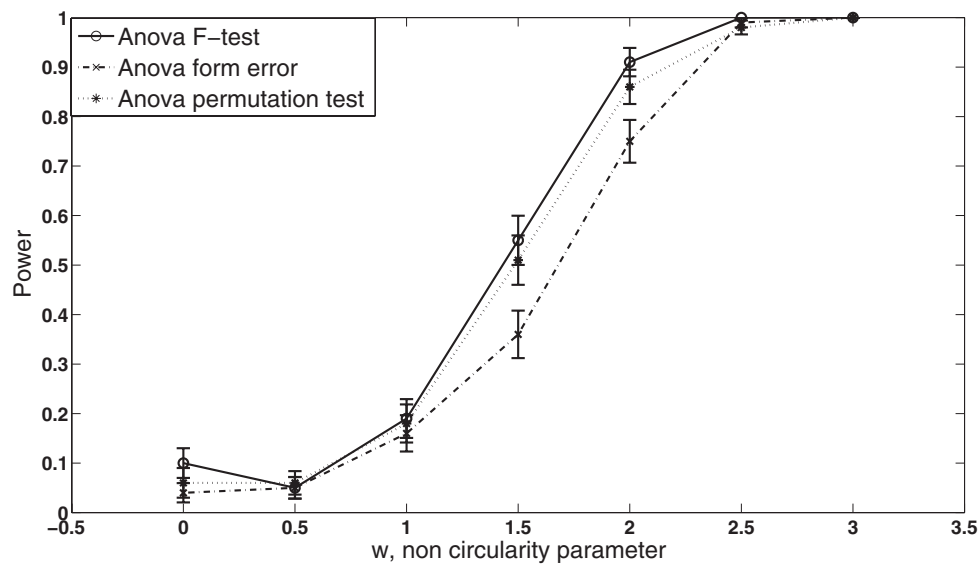


Fig. 8. Statistical power for detecting changes in 2D circular shapes under uniform errors (see text for simulation details).

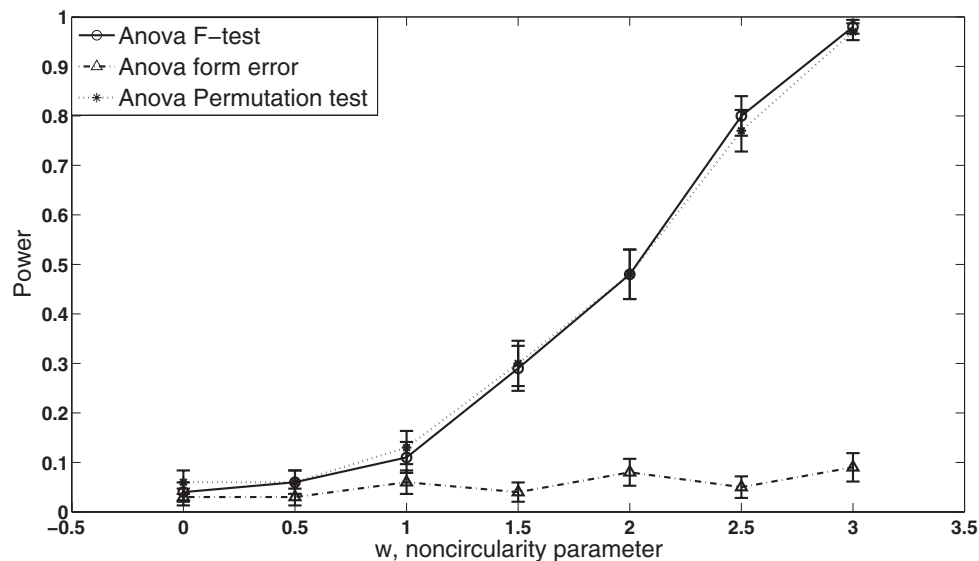


Fig. 9. Statistical power for detecting changes in 2D circular shapes $t_5(0, 0.05^2)$ errors (see text for simulation details).

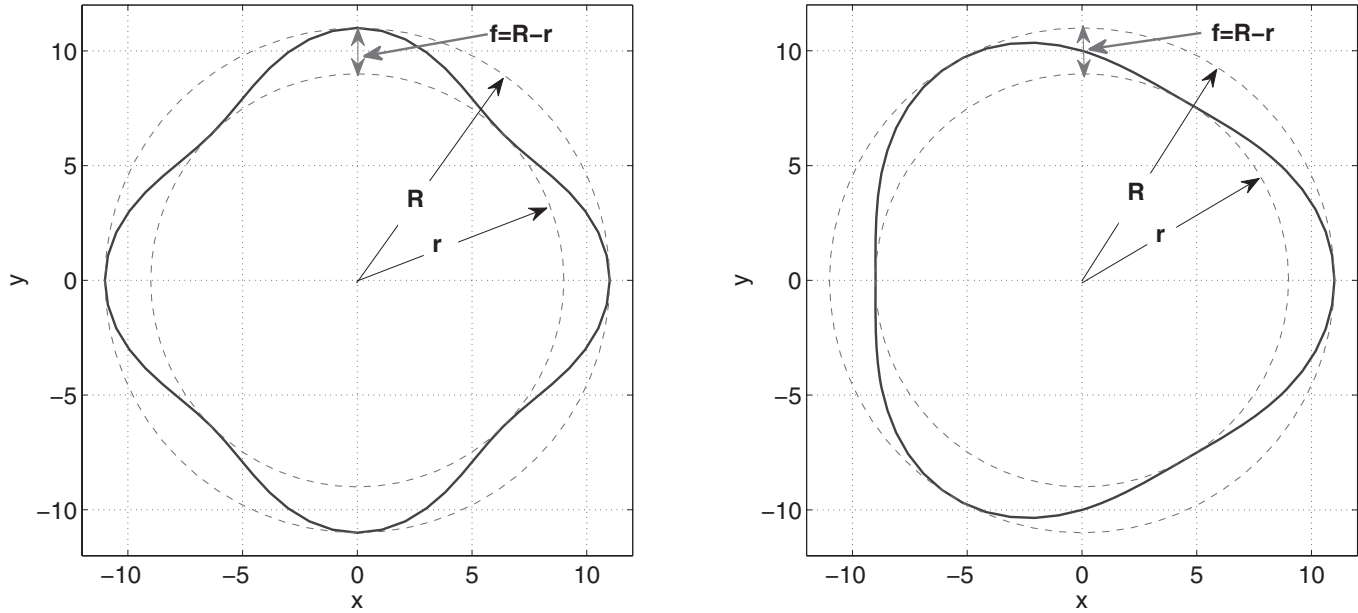


Fig. 10. Form error at two levels of a factor in an experimental study where the response of interest is the machined part cross section.

understanding. For illustration, assume a one-factor, two-level lathe-turning experiment where the response of interest is the shape of the cross section of the machined part. Furthermore, assume that the mean shapes obtained at the two different levels of the controllable factor are as shown in Fig. 10. These two mean shapes are commonly found in lathe-turning due to the irregular (excentric) rotational movement of the machine chuck so they are realistic; a process engineer would certainly like to detect a shape change, and hence a process change, such as that indicated in the figure. However, the circularity form errors for these two mean shapes are exactly the same and therefore the experimental factor has no significant effect on average; a standard ANOVA on the circularity error will fail to detect any change in mean shape. Applying instead SSA tools such as Goodall's F -test for shapes or the ANOVA permutation test for shapes, the difference in mean shapes will be easily detected. The reason for this behavior is simple: a standard ANOVA on the circularity form error is based on a single summary statistic of all shape data that is similar to a range that neglects a considerable amount of information and hence is very inefficient. In contrast, the test statistics used by the SSA test use all of the information available in the shapes and hence are more powerful than a standard ANOVA on the form error.

Another important reason why the SSA tests for shape responses and, in particular, the ANOVA permutation test for shapes should be preferred in practice, is simply that for some geometric features no standard definition of form error may exist, so performing a standard ANOVA on the form error is not even possible. Alshraideh (2011) performed power analyses for shapes of arbitrary geometry (e.g., triangles) where no standard definitions of form error

exist and found similar conclusions regarding the power and robustness of the tests studied in the present article.

5. Conclusions and summary of results

This article has reviewed and contrasted methods aimed at experimental situations where the object is to optimize the shape of a part or product. This is a task frequently done at the R&D or engineering design stage; thus, our goal has not been that of process inspection of the parts, where form errors are used to determine conformance to tolerances. The performance analyses conducted in this article consider two shapes that are common in manufacturing applications, namely, 2D circles and 3D cylinders. The analysis was conducted for three different error assumptions: isotropic errors, non-isotropic errors, and non-normal errors. A summary of the results found, related to the statistical power and Type-I error rate of each test, is as follows:

1. *Circles (2D) and cylinders (3D) under isotropic errors:* Both the ANOVA F -test and the ANOVA permutation test for shapes showed the best power performance with Type-I error rates close to nominal.
2. *Circles (2D) and cylinders (3D) under non-isotropic errors:* The ANOVA permutation test for shapes gives the best power performance among all the tests considered. The ANOVA F -test is powerful but provides a higher than advertised Type-I error rate.
3. *Circles under non-normal errors:* Again, the ANOVA permutation test showed the best power performance and Type-I error rate control. Similar to the case of non-isotropic errors, the ANOVA F -test showed similar

power to that for the ANOVA permutation test but with a higher than nominal Type-I error rate. In contrast, the ANOVA on the form errors provides good power performance only in case distribution of the errors is bounded since it is based on a range statistic.

The following conclusions and recommendations can be made from these results.

1. Based on the simulation results it can be concluded that the ANOVA permutation test for shapes in Del Castillo and Colosimo (2011) has the best overall power and best control of the Type-I error rate and hence it is the test we recommend.
2. The EDMA-I and EDMA-II tests exhibited very low power in three dimensions under both isotropic and non-isotropic errors. We cannot generalize our results to all 3D shapes since some other studies have shown that these tests do have good power in detecting shape differences for some geometrical shapes (however, this seems to be true only for shapes with very few landmarks; see Alshraideh (2011)).
3. Neither the ANOVA on the form errors nor the EDMA-I test is recommended for analyzing shape response experiments since they exhibited very low power in most of the cases considered. Our results show that EDMA-I has the true designed Type-I error under isotropic errors but exhibits low power.

Additional power performance studies for the tests in Section 3, for other error structures and mean shapes, including the case of interaction effects, are available in Alshraideh (2011). Further work should consider also the recent results by Dryden *et al.* (2008), who show how shape spaces are homeomorphic to the space of cross-products of preshapes (\mathbf{Z}/\mathbf{Z}). This has the advantage of reducing the dimensionality of the problem (we work with $m \times m$ matrices), but it is an open question how much power tests developed on this space have to detect changes in the original configuration space.

Acknowledgement

This work was partially supported by NSF grant CMI-0825786.

References

- Alshraideh, H.A. (2011) Analysis and optimization of profile and geometric shape response experiments. Ph.D. Thesis, Department of Industrial & Manufacturing Engineering, Penn State University, University Park, PA.
- Barcenas, C.C. and Griffin, P.M. (2001) Geometric tolerance verification using superquadrics. *IIE Transactions*, **33**, 1109–1120.

- Colosimo, B.M., Mammarella, F. and Petro, S. (2007) Quality control of manufactured of the surfaces, in *Proceedings of the Ninth International Workshop on Intelligent Statistical Quality Control*.
- Colosimo, B.M. and Pacella, M. (2011) Analyzing the effect of process parameters on the shape of 3D profiles. *Journal of Quality Technology*, **43**(3), 169–195.
- Davies, R., Twining, C. and Taylor, C. (2008) *Statistical Models of Shape, Optimisation and Evaluation*, Springer-Verlag, London, UK.
- Del Castillo, E. and Colosimo, B.M. (2011) Statistical shape analysis of experiments for manufacturing processes. *Technometrics*, **53**(1), 1–15.
- Dryden, I.L., Kume, A., Le, H.L. and Wood, A.T.A. (2008) A multi-dimensional scaling approach to shape analysis. *Biometrika*, **95**, 779–798.
- Dryden, I.L. and Mardia, K.V. (1998) *Statistical Shape Analysis*, Wiley, Chichester, UK.
- Ecker, M.D. and Gelfand, A.E. (1999) Bayesian modeling and inference for geometrically anisotropic spatial data. *Mathematical Geology*, **31**, 67–83.
- Edgington, E.S. (1995) *Randomization Tests*, Marcel Dekker, New York, NY.
- Goodall, C. (1991) Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society, Series B*, **53**(2), 285–339.
- Gower, J.C. (1975) Generalized Procrustes analysis. *Psychometrika*, **40**(1), 33–51.
- Henzold, G. (2006) *Geometrical Dimensioning and Tolerancing for Design, Manufacturing, and Inspection*, second edition, Butterworth-Heinemann, Oxford, UK.
- Huckemann, S., Hotz, T. and Munk, A. (2010a) Intrinsic MANOVA for Riemannian manifolds with an application to Kendall's space of planar shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**(4), 593–603.
- Huckemann, S., Hotz, T. and Munk, A. (2010b) Intrinsic shape analysis: geodesic principal component analysis for Riemannian manifolds modulo Lie group actions (with discussion). *Statistica Sinica*, **20**(1), 1–100.
- Kendall, D.G. (1984) Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, **16**, 81–121.
- Kendall, D.G., Barden, D., Carne, T.K. and Le, H. (1999) *Shape and Shape Theory*, Wiley, Chichester, UK.
- Krulikowski, A. (1996) *Geometric Dimensioning & Tolerancing*, Effective Training Inc., Westland, MI.
- Lele, S.R. and Cole, T.M. (1996) A new test for shape differences when variance-covariance matrices are unequal. *Journal of Human Evolution*, **31**, 193–212.
- Lele, S.R. and Richtsmeier, J.T. (1991) Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology*, **86**, 415–427.
- Small, C.G. (1996) *The Statistical Theory of Shape*, Springer, New York, NY.
- Ten Berge, J.M.F. (1977) Orthogonal Procrustes rotation for two or more matrices. *Psychometrika*, **42**(2), 267–276.
- Traband, M., Medeiros, D. and Chandra, M. (2004) A statistical approach to tolerance evaluation for circles and cylinders. *IIE Transactions*, **36**(8), 777–785.

Biographies

Hussam Alshraideh is an Assistant Professor in the Department of Industrial Engineering, Jordan University of Science and Technology in Irbid, Jordan. He holds a Ph.D. degree in Industrial & Manufacturing Engineering from Penn State University and an M.S. in Industrial Engineering from Arizona State University.

Enrique del Castillo is a Distinguished Professor in the Department of Industrial & Manufacturing Engineering, with a joint appointment as Professor in the Department of Statistics at Penn State University. He is the author of over 90 refereed papers in engineering statistics methods that have appeared in journals such as *Technometrics*, *The American Statistician*, *Biometrics*, *Journal of Statistical Planning and Inference*, *Journal of Applied Statistics*, *IEEE Transactions on Semiconductor Manufacturing*, *Journal of Quality Technology*, and *IIE Transactions*, among others. He

is the author of the book *Process Optimization, a Statistical Approach* (Springer, 2007) and of the book *Statistical Process Adjustment for Quality Control* (Wiley, 2002). He is a past editor-in-chief (2006–2009) of the *Journal of Quality Technology*. A former recipient of an NSF CAREER grant and a former Fulbright Scholar, he has held visiting professorship positions at the Universities of Tilburg (The Netherlands), Navarra (Spain), Politecnico di Milano (Italy), and the National University of Singapore.