XVIII ICIEOM 2012
International Conference on Industrial Engineering and Operations Management
Guimarães - Portugal
July 9 - 11, 2012

Simulated Analysis for Multivariate GR&R study

Rogério S. Peruchi [*], Anderson P. Paiva [*], Pedro P. Balestrassi [*], Alberto Garcia-Diaz [+]

[*] Institute of Industrial Engineering and Management, Federal University of Itajubá, 37500-093 Itajubá, MG, Brazil
[+] Department of Industrial & Information Engineering, University of Tennessee, 37996-0700 Knoxville, TN, USA

Email: rogerioperuchi@unifei.edu.br, andersonppaiva@unifei.edu.br, pedro@unifei.edu.br, agd@utk.edu

Abstract

This article explores the analysis of measurement system with correlated characteristics through the study of repeatability and reproducibility. The main contribution of this research is to propose a method for multivariate analysis of a measurement system by considering the weighted principal components (WPC). To prove its efficiency, we generate simulated data with different correlation structures for measurement systems that are acceptable, marginal, and unacceptable. The proposed method is compared with classical univariate and multivariate methods. It was observed that, compared to the other methods, the WPC was more robust in estimating the assessment indices of a multivariate measurement system.

Keywords: measurement system analysis, repeatability and reproducibility, principal component analysis, correlated quality characteristics.

1 Introduction

Quality improvement projects are often characterized by their objective to reduce variability and achieve zero-defect production. If a product fails to conform, analysts generally attribute the failure to the process. The analysts then act to improve process capability. In some instances, however, there may be nothing wrong with the process capability. Yet the measurement error, when compared to the variability of the process, remains unacceptable (Al-Refaie & Bata, 2010). Hence, before a team of analysts tries to improve a process, they should investigate both the variability of the measurement process as well as the variability of the manufacturing process. In manufacturing, a measurement system is not used to producing the exact dimension of a part. It provides measurements that, due to errors (random and systematic), vary from the true value (AIAG, 2010). In any activity involving measurements, some of the observed variability is due to the product itself, σ_P^2, while the remainder is due to measurement error or variability in the measurement system, σ_{MS}^2 (Li & Al-Refaie, 2008; Senol, 2004; Woodall & Borror, 2008).

In measurement system analysis (MSA), the study used to measure the components of variation is called Gage Repeatability and Reproducibility (GR&R). Repeatability is the variation in measurements obtained with one measuring instrument when used several times by an appraiser while measuring the identical characteristic on the same part. Reproducibility is the variation in the average of measurements made by different appraisers using the same gage when measuring a characteristic on one part (Awad et al., 2009; Burdick et al., 2003; Erdmann et al., 2010; Polini & Turchetta, 2004). GR&R aims to determine that a measurement system's variability is less than that of the monitored process (Al-Refaie & Bata, 2010).

In assessing measurement systems that measure multiple characteristics, univariate approach may prove unsatisfactory, as might the strategy of prioritizing the CTQ. In such systems, an analysis must consider the correlation structure of the CTQs, a task more suited to multivariate methods. For GR&R studies, the literature reveals few multivariate studies (Wang & Chien, 2010). This article deals with a multivariate analysis of a measurement system through studies of repeatability and reproducibility of the measurement process. Its main contribution is to propose a new method for multivariate analysis of a measurement system based on principal component analysis. The new method Weighted Principal Components (WPC) ponders the principal component scores by their eigenvalues. To prove the efficiency of the method, simulated data are generated with different correlation structures so as to measure systems that are unacceptable, marginal, and acceptable. The results obtained by WPC will be compared to those obtained by Principal Component Analysis (PCA) and Multivariate Analysis of Variance (MANOVA) methods. The simulation study concludes that the proposed method is more robust than MANOVA (Majeske, 2008) and PCA (Wang & Chien, 2010).

The remainder of this paper is structured as follows. Section 2 shows a brief review about how to obtain principal component scores and how to evaluate a measurement system using the WPC method proposed by the authors. In Section 3, a simulation study is conducted to evaluate the performance of the multivariate methods for different correlation structures and for unacceptable, marginal, and acceptable measurement systems. Finally, Section 4 presents the main findings involving the analysis using the multivariate methods MANOVA, PCA, and WPC.

2 Multivariate GR&R study based on PCA and WPC

According to Wang & Chien (2010), to deal with multiple CTQs in a GR&R study, PCA is an alternative method to the MANOVA proposed by Majeske (2008). Principal component analysis is one of the most widely applied tools used to summarize the common patterns of variation among variables. Furthermore, this statistical technique is also able to retain significant information in the first axes of the PCs, since the variation associated with experimental error, measurement error, rounding error is summarized in the last axes of PCs (Paiva et al., 2007). Principal components are algebraically a linear combination ℓ of q random variables $\mathbf{CTQ_1, CTQ_2, \ldots, CTQ_q}$. Geometrically these combinations represent a new coordinate system obtained during the rotation of an original system (Johnson & Wichern, 2002; Paiva et al., 2008). The coordinates of the axes now have the variables $\mathbf{CTQ_1, CTQ_2, \ldots, CTQ_q}$ and represent the direction of maximum. The principal components are uncorrelated and depend only on the variance-covariance matrix **Σ** (or the correlation matrix **R**) of variables $\mathbf{CTQ_1, CTQ_2, \ldots, CTQ_q}$ and their development does not require the assumption of multivariate normality. Since the matrix has pairs of eigenvalues-eigenvectors $(\boldsymbol{\lambda}_1, \mathbf{e}_1), (\boldsymbol{\lambda}_2, \mathbf{e}_2), \ldots, \geq (\boldsymbol{\lambda}_q, \mathbf{e}_q)$, where $\boldsymbol{\lambda}_1 \geq \boldsymbol{\lambda}_2 \geq \ldots \geq \boldsymbol{\lambda}_q \geq \mathbf{0}$, then the i^{th} principal component is:

$$\mathbf{PC_i} = \mathbf{e_i^T Y} = \mathbf{e_{1i} Y_1} + \mathbf{e_{2i} Y_2} + \ldots + \mathbf{e_{qi} Y_q} \quad i = 1, 2, \ldots, q \tag{1}$$

The i^{th} principal component can be obtained according to:

$$\begin{aligned} &Maximize: \quad Var\left[\mathbf{e_i^T CTQ}\right] \\ &Subject\ to: \quad \mathbf{e_i^T e_i} = 1 \\ &\qquad\qquad Cov\left[\mathbf{e_i^T CTQ, e_k^T CTQ}\right] = \mathbf{0}, \quad k < i \end{aligned} \tag{2}$$

An original set of variables can be replaced by an uncorrelated linear combination called principal component scores. Considering **Z**, the matrix of standardized data, and **E**, the matrix of eigenvectors of the multivariate set, each principal component scores can then be obtained from (He et al., 2011; Johnson & Wichern, 2002):

$$\mathbf{PC_{score}} = \mathbf{Z^T E} = \begin{bmatrix} \left(\frac{y_{11} - \bar{y}_1}{\sqrt{s_{11}}}\right) & \left(\frac{y_{12} - \bar{y}_2}{\sqrt{s_{22}}}\right) & \cdots & \left(\frac{y_{1q} - \bar{y}_q}{\sqrt{s_{qq}}}\right) \\ \left(\frac{y_{21} - \bar{y}_1}{\sqrt{s_{11}}}\right) & \left(\frac{y_{22} - \bar{y}_2}{\sqrt{s_{22}}}\right) & \cdots & \left(\frac{y_{2q} - \bar{y}_q}{\sqrt{s_{qq}}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{y_{n1} - \bar{y}_1}{\sqrt{s_{11}}}\right) & \left(\frac{y_{n2} - \bar{y}_2}{\sqrt{s_{22}}}\right) & \cdots & \left(\frac{y_{nq} - \bar{y}_q}{\sqrt{s_{qq}}}\right) \end{bmatrix}^{\mathbf{T}} \times \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1q} \\ e_{21} & e_{22} & \cdots & e_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1q} & e_{2q} & \cdots & e_{qq} \end{bmatrix} \tag{3}$$

Eq. (4) represents a complete model for a multivariate GR&R study with *q* quality characteristics, *p* parts, *o* operators, and *r* replicates that can be analyzed by PCA. This model is similar to the univariate model. The original responses, however, are replaced by the principal component scores.

$$\mathbf{PC_q} = \boldsymbol{\mu} + \boldsymbol{\alpha}_\mathbf{i} + \boldsymbol{\beta}_\mathbf{j} + (\boldsymbol{\alpha\beta})_\mathbf{ij} + \boldsymbol{\varepsilon}_\mathbf{ijk} \quad \forall i = 1,2,\ldots,p;\ j = 1,2,\ldots,o;\ k = 1,2,\ldots,r \tag{4}$$

The variable μ is a constant and α_i, β_j, $(\alpha\beta)_{ij}$, ε_{ijk} are independent normal random variables with zero mean and variance, σ_α^2, σ_β^2, $\sigma_{\alpha\beta}^2$, and σ_ε^2, for parts (process), operators, part*operator interaction, and error term, respectively. In their analysis of measurement systems, Wang and Chien (2010) compared the PCA with two other methods. However, the authors conducted the analysis separately for each principal component. This methodology may be inappropriate because the individual analysis of each component can provide different interpretations. When responses have very high correlations (*%PC₁>95%*), analysis of the first principal component explains reasonably well the variability of the measurement system. When correlations between the responses are not very high, however, it becomes necessary to analyze more than one principal component. Indeed, the first principal component alone cannot explain the whole data set. Therefore, we propose in this article, a method of a multivariate GR&R study using weighted principal components. In this case, the response of the model is principal components scores weighted by their respective eigenvalues. This proposal is based on the work of Paiva et al. (2010), who used a technique of multi-objective optimization based on a weighting of the principal components. They used the technique to study a process of welding with a multiple set of responses moderately correlated. Therefore, the proposed model is given by:

$$\mathbf{WPC} = \boldsymbol{\mu} + \boldsymbol{\alpha}_\mathbf{i} + \boldsymbol{\beta}_\mathbf{j} + (\boldsymbol{\alpha\beta})_\mathbf{ij} + \boldsymbol{\varepsilon}_\mathbf{ijk} \quad \forall i = 1,2,\ldots,p;\ j = 1,2,\ldots,o;\ k = 1,2,\ldots,r \tag{5}$$

where:

$$\mathbf{WPC} = \sum_{\mathbf{i=1}}^{\mathbf{q}} \left[\boldsymbol{\lambda}_\mathbf{i}(\mathbf{PC}_\mathbf{i})\right] = \lambda_1 \mathbf{PC}_\mathbf{1} + \lambda_2 \mathbf{PC}_\mathbf{2} + \ldots + \lambda_q \mathbf{PC}_\mathbf{q} \tag{6}$$

That is, the response used in model (5) is the result of a pondering of the principal components by their eigenvalues, according to Eq. (6). The variable μ is a constant and α_i, β_j, $(\alpha\beta)_{ij}$, ε_{ijk} are independent normal random variables with zero mean and variance, $\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2, \sigma_\varepsilon^2$, respectively. In Johnson and Wichern (2002), it appears that there are a variety of rules to estimating the appropriate number of non-trivial PCA axes (PC scores) that can be taken to represent the original data set. However, due to the weighting of the principal components by their eigenvalues, all principal components can be included in the model. The components with eigenvalue of greater importance will, in the model, be weighted more, and whole information will be included in the study. The components of variance in model (5) can be translated into GR&R notation by:

$$\hat{\sigma}_P^2 = \hat{\sigma}_\alpha^2 = \frac{MSP - MSPO}{or} \tag{7}$$

$$\hat{\sigma}_{repeatability}^2 = \hat{\sigma}_\varepsilon^2 = MSE \tag{8}$$

$$\hat{\sigma}_{reproducibility}^2 = \hat{\sigma}_\beta^2 + \hat{\sigma}_{\alpha\beta}^2 = \frac{MSO - MSPO}{pr} + \frac{MSPO - MSE}{r} \tag{9}$$

$$\hat{\sigma}_{MS}^2 = \hat{\sigma}_{repeatability}^2 + \hat{\sigma}_{reproducibility}^2 \tag{10}$$

$$\hat{\sigma}_T^2 = \hat{\sigma}_P^2 + \hat{\sigma}_{MS}^2 \tag{11}$$

MSP, MSO, MSPO, and *MSE* are, respectively, the mean square for the factors part, operator, interaction, and the error term. If the interaction effect is not significant, the complete model can be reduced to Eq. (12) and its components of variance can be translated into a GR&R notation by Eqs. (13)-(17).

$$\mathbf{WPC} = \boldsymbol{\mu} + \boldsymbol{\alpha}_\mathbf{i} + \boldsymbol{\beta}_\mathbf{j} + \boldsymbol{\varepsilon}_\mathbf{ijk} \forall i = 1,2,\ldots,p;\ j = 1,2,\ldots,o;\ k = 1,2,\ldots,r \tag{12}$$

$$\hat{\sigma}_P^2 = \frac{MSP - MSE}{or} \tag{13}$$

$$\hat{\sigma}^2_{repeatability} = \hat{\sigma}^2_{\varepsilon} = MSE \quad (14)$$

$$\hat{\sigma}^2_{reproducibility} = \hat{\sigma}^2_{\beta} = \frac{MSO - MSE}{pr} \quad (15)$$

$$\hat{\sigma}^2_{MS} = \hat{\sigma}^2_{repeatability} + \hat{\sigma}^2_{reproducibility} \quad (16)$$

$$\hat{\sigma}^2_{T} = \hat{\sigma}^2_{P} + \hat{\sigma}^2_{MS} \quad (17)$$

The assessment index of measurement system used to compare the performance of the methods is %R&R (repeatability and reproducibility percentage) in Eq. (18). The acceptance criterion for the measurement system of the multivariate $\%R\&R_m$ index is the same for the univariate *%R&R* index (Majeske, 2008; Wang & Chien, 2010). If $\%R\&R_m$ is less than 10%, the measurement system is considered acceptable. If $\%R\&R_m$ is between 10% and 30%, the measurement system is considered marginal—acceptable depending on the application, the cost of the measurement device, the cost of repair and other factors. If, according to the index, the measurement system exceeds 30%, then it is considered unacceptable and should be improved (AIAG, 2010; Al-Refaie & Bata, 2010; Burdick et al., 2003; Woodall & Borror, 2008).

$$\%R \,\&\, R_m = \left(\frac{\sigma_{MS}}{\sigma_T}\right)100\% \quad (18)$$

3 Simulation

The purpose of this simulation is to evaluate almost all possible situations in multivariate analysis of a measurement system and to compare the results achieved through multivariate methods. Simulated data will be generated for measurement systems that are unacceptable (*%R&R>30%*), marginal (*10%<%R&R<30%*), and acceptable (*%R&R<10%*), as well as correlations that are very low ($\%PC_1 \leq 65\%$), low ($65\% < \%PC_1 \leq 75\%$), medium ($75\% < \%PC_1 \leq 85\%$), high ($85\% < \%PC_1 \leq 95\%$), and very high ($\%PC_1 > 95\%$), a total of 15 scenarios and 1800 simulated measurements. $\%PC_1$ is the result obtained from $\lambda_1 / \sum_{i=1}^{q} \lambda_i$. Figure 1 exemplifies two extreme cases, very high correlations (Figure 1a) and very low (Figure 1b) between two response variables CTQ_2 and CTQ_4. It is observed that the greater the similarity pattern of change of factor levels (parts and operators), the greater the correlation between the characteristics. Also, if you set very different mean values for two operators measuring the same part, the analysis of variance indicates statistically significant differences between operators and/or interaction term (part*operator). Figure 2 shows a flowchart that details how to obtain the simulated data. According to this flowchart, the simulated data were generated from the information in Table 1, according to the same amount of parts and operators in Majeske (2008), *p=5*, *o=2* and *r=3*. The data for the 15 simulated scenarios can be found in http://www.pedro.unifei.edu.br/download/tables.rar.

This simulation study will focus only on the comparison of multivariate methods. To understand how to calculate univariate *%R&R* index using ANOVA method, see AIAG (2010). The calculation of multivariate $\%R\&R_m$ index through MANOVA and PCA methods can be found in Majeske (2008) and Wang and Chien (2010), respectively. Thus, the $\%R\&R_m$ and *%R&R* indices were calculated to compare the multivariate methods. For each scenario we tried to obtain close *%R&R* index values for CTQ_1, CTQ_2, CTQ_3, and CTQ_4. It is expected that the indices obtained by multivariate methods are close to those obtained by ANOVA method. The criterion used in this work to determine if the estimated multivariate index, $\%R\&R_m$, is correct is based on confidence intervals for mean calculated from data obtained by ANOVA method. The lower (LLCI) and upper (ULCI) limits of the confidence intervals are calculated using Eqs. (19) and (20):

$$LLCI = \overline{CTQ} - t_{N-1,\,\alpha/2}\left(s/\sqrt{N}\right) \quad (19)$$

$$ULCI = \overline{CTQ} + t_{N-1,\,\alpha/2}\left(s/\sqrt{N}\right) \quad (20)$$

where $\overline{CTQ}$ is the mean of *%R&R* between CTQ_1, CTQ_2, CTQ_3, and CTQ_4; *s* is the standard deviation; *N* is the sample size and $t_{N-1,\alpha}$ is the $(1-\alpha)100^{th}$ percentile of a *t* distribution with *(N-1)* degrees of freedom.

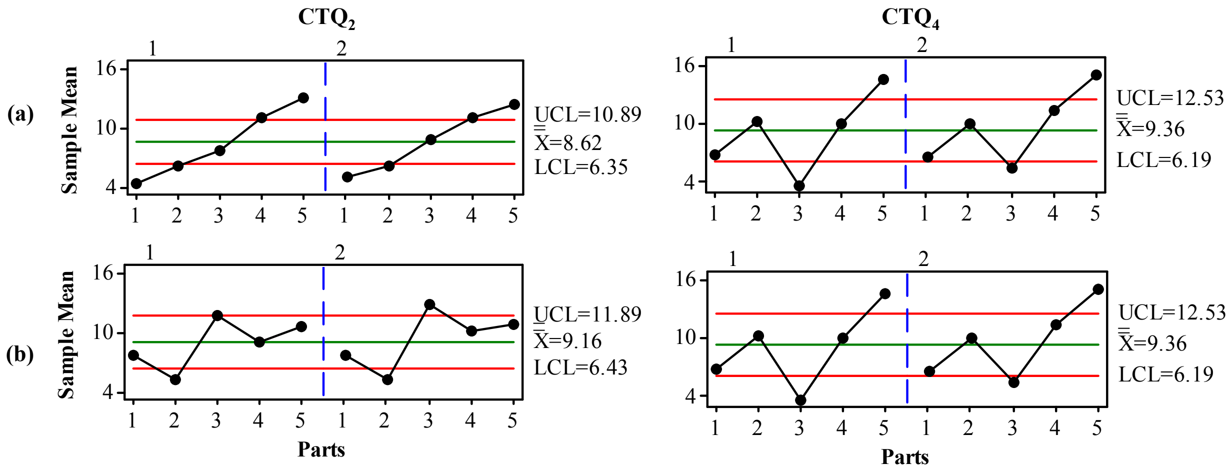

Figure 1: Xbar charts by operators: (a) example of a very high correlation (0.999) between the responses CTQ_2 and CTQ_4; (b) example of a very low correlation (0.088) between the responses CTQ_2 and CTQ_4

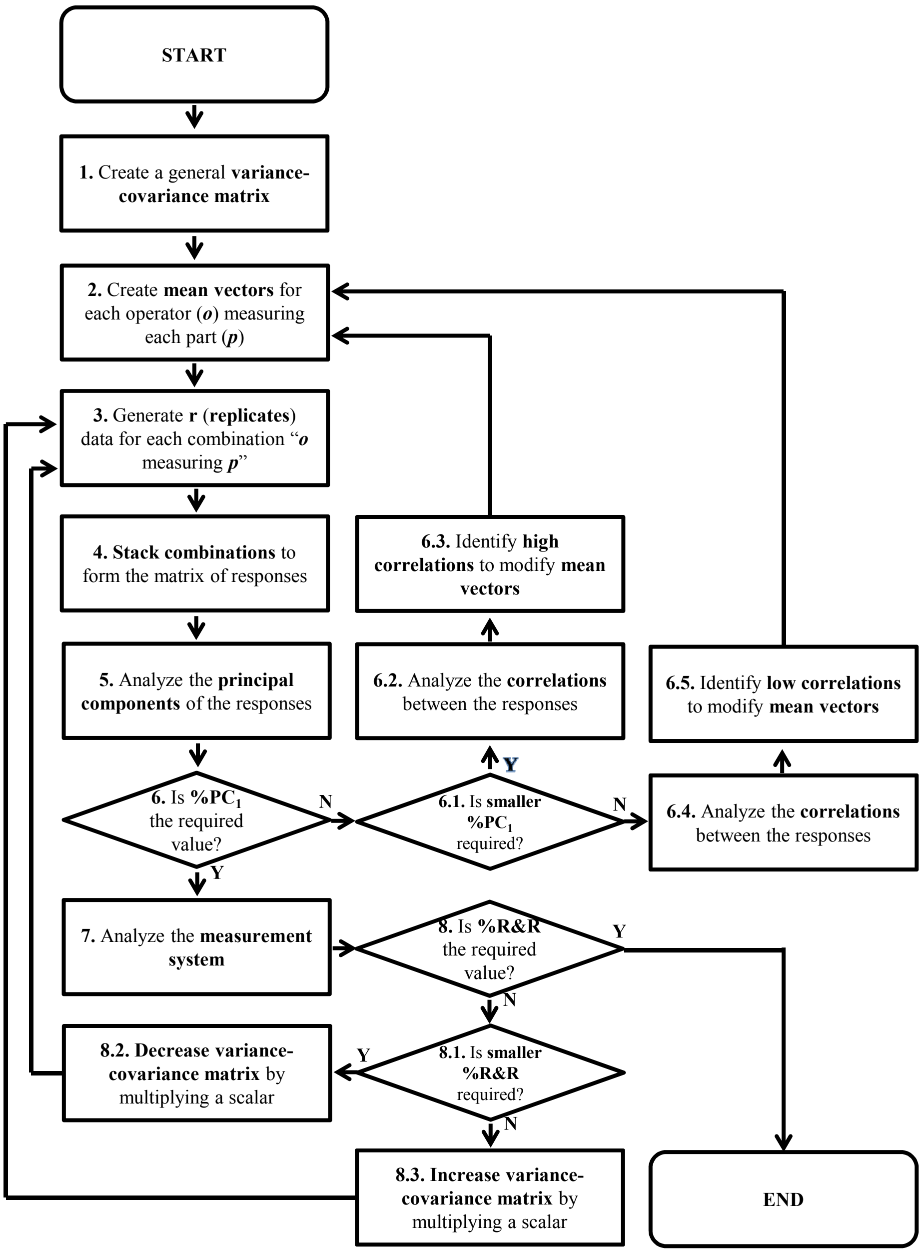

Figure 2: Detailed flowchart explaining how to obtain simulated data to a multivariate GR&R study

Table 1: Mean vectors and variance-covariance matrices used to generate simulated data with different correlations and measurement systems (MS)

Scenarios	Mean vectors										Variance-covariance matrix
	P_1O_1	P_2O_1	P_3O_1	P_4O_1	P_5O_1	P_1O_2	P_2O_2	P_3O_2	P_4O_2	P_5O_2	
1	4.00	8.00	6.00	10.00	5.00	4.10	8.10	5.90	9.90	4.90	1.10 1.27 1.39 1.50
Very Low corr.	8.00	6.00	13.00	9.00	11.00	7.90	6.10	12.90	9.10	10.90	1.27 1.50 1.63 1.76
Unacceptable MS	9.00	10.00	13.00	16.00	7.00	9.10	10.10	12.90	15.90	7.10	1.39 1.63 1.80 1.92
	7.00	11.00	5.00	10.00	15.00	7.10	10.90	5.10	10.10	15.10	1.50 1.76 1.92 2.10
2	4.00	8.00	6.00	10.00	5.00	4.10	8.10	5.90	9.90	4.90	1.10 1.27 1.39 1.50
Low corr.	8.00	7.00	9.00	12.00	11.00	7.90	6.90	9.10	12.10	10.90	1.27 1.50 1.63 1.76
Unacceptable MS	9.00	10.00	7.00	13.00	15.00	9.10	10.10	6.90	13.10	14.90	1.39 1.63 1.80 1.92
	7.00	13.00	11.00	14.00	17.00	7.10	13.10	11.10	13.90	16.90	1.50 1.76 1.92 2.10
3	9.00	7.00	5.00	12.00	10.00	9.01	6.99	5.01	12.01	9.99	1.10 1.27 1.39 1.50
Medium corr.	8.00	7.00	9.00	12.00	11.00	7.99	6.99	9.01	12.01	10.99	1.27 1.50 1.63 1.76
Unacceptable MS	9.00	10.00	7.00	13.00	15.00	9.01	10.01	6.99	13.01	14.99	1.39 1.63 1.80 1.92
	7.00	13.00	9.00	17.00	14.00	7.01	13.01	8.99	16.99	14.01	1.50 1.76 1.92 2.10
4	6,00	4,00	8,00	10,00	12,00	6,01	4,01	7,99	9,99	12,01	1.50 1.58 1.63 1.67
High corr.	3,00	6,00	9,00	11,00	15,00	3,01	6,01	9,01	10,99	14,99	1.58 1.70 1.73 1.78
Unaccpetable MS	6,00	8,00	11,00	15,00	13,00	6,01	8,01	11,10	15,10	13,10	1.63 1.73 1.80 1.83
	8,00	10,00	12,00	16,00	14,00	7,99	10,01	12,01	16,01	14,01	1.67 1.78 1.83 1.90
5	4.00	6.00	8.00	10.00	12.00	4.01	6.01	7.99	9.99	12.01	1.10 1.27 1.39 1.50
Very high corr.	5.00	7.00	9.00	11.00	13.00	5.01	7.01	9.01	10.99	12.99	1.27 1.50 1.63 1.76
Unacceptable MS	6.00	8.00	10.00	12.00	14.00	6.01	8.01	9.99	11.99	13.99	1.39 1.63 1.80 1.92
	8.00	10.00	12.00	14.00	16.00	7.99	10.01	12.01	14.01	15.99	1.50 1.76 1.92 2.10
6	4.00	8.00	6.00	10.00	5.00	4.10	8.10	5.90	9.90	4.90	0.22 0.25 0.28 0.30
Very Low corr.	8.00	6.00	13.00	9.00	11.00	7.90	6.10	12.90	9.10	10.90	0.25 0.30 0.33 0.35
Marginal MS	5.00	8.00	9.00	14.00	12.00	5.10	8.10	8.90	13.90	12.10	0.28 0.33 0.36 0.38
	7.00	13.00	5.00	10.00	17.00	7.10	13.10	5.10	10.10	16.90	0.30 0.35 0.38 0.42
7	6.00	8.00	4.00	11.00	10.00	6.10	8.10	3.90	10.90	9.90	0.22 0.25 0.28 0.30
Low corr.	8.00	7.00	9.00	12.00	11.00	7.90	6.90	9.10	12.10	10.90	0.25 0.30 0.33 0.35
Marginal MS	7.00	13.00	10.00	11.00	15.00	7.10	13.10	9.90	11.10	14.90	0.28 0.33 0.36 0.38
	9.00	11.00	14.00	13.00	17.00	9.10	10.90	14.10	12.90	16.90	0.30 0.35 0.38 0.42
8	9.00	7.00	5.00	12.00	10.00	9.01	6.99	5.01	12.01	9.99	0.22 0.25 0.28 0.30
Medium corr.	8.00	7.00	9.00	12.00	11.00	7.99	6.99	9.01	12.01	10.99	0.25 0.30 0.33 0.35
Marginal MS	9.00	10.00	7.00	13.00	15.00	9.01	10.01	6.99	13.01	14.99	0.28 0.33 0.36 0.38
	7.00	13.00	9.00	17.00	14.00	7.01	13.01	8.99	16.99	14.01	0.30 0.35 0.38 0.42
9	6,00	4,00	8,00	10,00	12,00	6,01	4,01	7,99	9,99	12,01	0.22 0.25 0.28 0.30
High corr.	3,00	6,00	9,00	11,00	15,00	3,01	6,01	9,01	10,99	14,99	0.25 0.30 0.33 0.35
Marginal MS	6,00	8,00	11,00	15,00	13,00	6,01	8,01	11,10	15,10	13,10	0.28 0.33 0.36 0.38
	8,00	10,00	12,00	16,00	14,00	7,99	10,01	12,01	16,01	14,01	0.30 0.35 0.38 0.42
10	4.00	6.00	8.00	10.00	12.00	4.01	6.01	7.99	9.99	12.01	0.22 0.25 0.28 0.30
Very high corr.	5.00	7.00	9.00	11.00	13.00	5.01	7.01	9.01	10.99	12.99	0.25 0.30 0.33 0.35
Marginal MS	6.00	8.00	10.00	12.00	14.00	6.01	8.01	9.99	11.99	13.99	0.28 0.33 0.36 0.38
	8.00	10.00	12.00	14.00	16.00	7.99	10.01	12.01	14.01	15.99	0.30 0.35 0.38 0.42
11	4.00	8.00	6.00	10.00	5.00	4.10	8.10	5.90	9.90	4.90	0.04 0.04 0.05 0.05
Very Low corr.	8.00	6.00	13.00	9.00	11.00	7.90	6.10	12.90	9.10	10.90	0.04 0.05 0.05 0.06
Acceptable MS	5.00	8.00	9.00	14.00	12.00	5.10	8.10	8.90	13.90	12.10	0.05 0.06 0.06 0.06
	7.00	13.00	5.00	10.00	17.00	7.10	13.10	5.10	10.10	16.90	0.05 0.06 0.06 0.07
12	6.00	8.00	4.00	11.00	10.00	6.01	8.01	3.99	10.99	9.99	0.04 0.04 0.05 0.05
Low corr.	7.00	5.00	9.00	13.00	11.00	6.99	4.99	9.01	13.01	10.99	0.04 0.05 0.05 0.06
Acceptable MS	7.00	13.00	10.00	11.00	15.00	7.01	13.01	9.99	11.01	14.99	0.05 0.06 0.06 0.06
	6.00	10.00	14.00	12.00	17.00	6.01	9.99	14.01	12.01	16.99	0.05 0.06 0.06 0.07
13	9.00	7.00	5.00	12.00	10.00	9.01	6.99	5.01	12.01	9.99	0.04 0.04 0.05 0.05
Medium corr.	8.00	7.00	9.00	12.00	11.00	7.99	6.99	9.01	12.01	10.99	0.04 0.05 0.05 0.06
Acceptable MS	9.00	10.00	7.00	13.00	15.00	9.01	10.01	6.99	13.01	14.99	0.05 0.06 0.06 0.06
	7.00	13.00	9.00	17.00	14.00	7.01	13.01	8.99	16.99	14.01	0.05 0.06 0.06 0.07
14	6,00	4,00	8,00	10,00	12,00	6,01	4,01	7,99	9,99	12,01	0.04 0.04 0.05 0.05
High corr.	3,00	6,00	9,00	11,00	15,00	3,01	6,01	9,01	10,99	14,99	0.04 0.05 0.05 0.06
Acceptable MS	6,00	8,00	11,00	15,00	13,00	6,01	8,01	11,10	15,10	13,10	0.05 0.06 0.06 0.06
	8,00	10,00	12,00	16,00	14,00	7,99	10,01	12,01	16,01	14,01	0.05 0.06 0.06 0.07
15	4.00	6.00	8.00	10.00	12.00	4.01	6.01	7.99	9.99	12.01	0.04 0.04 0.05 0.05
Very high corr.	5.00	7.00	9.00	11.00	13.00	5.01	7.01	9.01	10.99	12.99	0.04 0.05 0.05 0.06
Acceptable MS	6.00	8.00	10.00	12.00	14.00	6.01	8.01	9.99	11.99	13.99	0.05 0.06 0.06 0.06
	8.00	10.00	12.00	14.00	16.00	7.99	10.01	12.01	14.01	15.99	0.05 0.06 0.06 0.07

Table 2 presents the results of calculations of the *%R&R* index as well as the mean value and the 95% confidence interval, obtained by ANOVA method. Table 3 show the results of calculations of the mean value, 95% confidence interval and *%R&R*$_m$ index, obtained by PCA, MANOVA and WPC methods. The analysis and comparison will be performed in two ways: intra- and inter-methods. The intra-method analysis will provide an overview of the methods' performance to estimate the *%R&R*$_m$ index. The inter-method analysis will seek to justify the methods' deviations of estimates of the *%R&R*$_m$ index from the confidence intervals.

Table 2: Results for calculations of the *%R&R* index, mean and 95% confidence interval.

SCENARIO			UNIVARIATE (*%R&R*)				MEAN CI		
S	**MS**	**Correlation**	**CTQ_1**	**CTQ_2**	**CTQ_3**	**CTQ_4**	**Mean**	**LLCI**	**ULCI**
S1	Unacceptable	Very Low	49.9	39.3	38.3	34.1	40.42	29.69	51.14
S2		Low	42.2	55.5	44.3	39.8	45.44	34.42	56.47
S3		Medium	40.8	52.4	42.6	36.9	43.18	32.63	53.72
S4		High	45.3	33.2	41.2	47.8	41.86	31.70	52.03
S5		Very High	31.1	34.9	37.8	41.1	36.21	29.45	42.97
S6	Marginal	Very Low	15.8	14.1	13.7	10.2	13.48	9.75	17.21
S7		Low	18.6	27.2	21.3	24.1	22.82	16.95	28.69
S8		Medium	15.5	23.7	17.0	14.6	17.69	11.16	24.21
S9		High	13.2	10.3	13.6	16.9	13.50	9.19	17.80
S10		Very High	15.2	19.0	19.7	20.9	18.70	14.80	22.59
S11	Acceptable	Very Low	8.4	6.3	4.9	5.3	6.22	3.67	8.77
S12		Low	5.6	4.6	6.7	5.4	5.54	4.15	6.92
S13		Medium	6.2	9.6	6.6	5.9	7.07	4.37	9.76
S14		High	5.7	4.5	5.9	7.3	5.84	4.00	7.69
S15		Very High	6.5	7.6	8.6	9.2	7.95	6.07	9.83

Table 3: Results for calculations of the mean, 95% confidence interval and *%R&R*$_m$.

	MEAN CI			MULTIVARIATE (*%R&R*$_m$)					
S	**Mean**	**LLCI**	**ULCI**	**PC_1**	**PC_2**	**PC_3**	**PC_4**	**MANOVA**	**WPC**
S1	40.42	29.69	51.14	52.24 (55.8)	19.55 (29.1)	15.32 (14.1)	**42.08 (1.0)**	10.78	**39.71**
S2	45.44	34.42	56.47	**53.84 (70.7)**	10.48 (17.5)	20.91 (8.1)	19.32 (3.7)	13.30	**52.84**
S3	43.18	32.63	53.72	**47.79 (79.4)**	11.21 (9.4)	17.65 (7.5)	9.64 (3.7)	11.32	**47.87**
S4	41.86	31.70	52.03	**44.38 (88.3)**	8.94 (8.4)	29.62 (3.1)	75.92 (0.2)	28.15	**44.03**
S5	36.21	29.45	42.97	**36.10 (99.7)**	97.05 (0.2)	100.00 (0.0)	100.00 (0.0)	64.09	**36.11**
S6	13.48	9.75	17.21	18.46 (44.9)	6.92 (29.9)	2.79 (23.5)	32.43 (1.7)	4.97	**12.65**
S7	22.82	16.95	28.69	**24.97 (66.2)**	2.77 (17.8)	8.37 (15.9)	46.89 (0.2)	10.04	**26.98**
S8	17.69	11.16	24.21	**19.71 (79.8)**	6.01 (9.2)	**11.78 (7.5)**	7.19 (3.5)	5.40	**19.86**
S9	13.50	9.19	17.80	**14.17 (89.8)**	6.01 (7.6)	**9.99 (2.5)**	54.07 (0.1)	**14.31**	**14.00**
S10	18.70	14.80	22.59	**18.63 (99.9)**	93.87 (0.1)	100.00 (0.0)	100.00 (0.0)	47.23	**18.63**
S11	6.22	3.67	8.77	**6.41 (45.5)**	**6.39 (35.9)**	**4.59 (17.4)**	15.83 (1.1)	**4.08**	**4.10**
S12	5.54	4.15	6.92	**6.69 (67.6)**	1.15 (18.3)	2.25 (13.8)	**5.33 (0.4)**	2.01	**6.89**
S13	7.07	4.37	9.76	**7.87 (79.7)**	1.95 (8.9)	**5.75 (7.8)**	3.08 (3.6)	2.28	**8.04**
S14	5.84	4.00	7.69	**5.99 (90.3)**	3.34 (7.1)	3.96 (2.6)	33.24 (0.1)	**7.22**	**5.91**
S15	7.95	6.07	9.83	**7.92 (100.0)**	96.68 (0.0)	100.00 (0.0)	100.00 (0.0)	39.35	**7.92**

In the intra-method analysis is verified that, in the estimation of the *%R&R*$_m$ index, the WPC was more robust than the MANOVA and PCA. The MANOVA method was able to estimate the multivariate index within the confidence interval only in scenarios S9, S11 and S14. For PCA method, Wang and Chien (2010) evaluated the principal components that had a cumulative percentage of explanation at least 95% for the original variables. Thus, the PCA method was capable to estimate the multivariate index within the confidence interval only in scenarios S5, S10, S11 and S15. As seen in Table 10, WPC estimated *%R&R*$_m$ index within the confidence interval for all 15 scenarios evaluated.

For the inter-method analysis by PCA, Table 3 presents the analysis of measurement systems simulated to the four principal components. The values in parentheses show the explanation percentage of variability of the CTQs for each principal component. It is recommended evaluate the scores of principal components that represent at least 95% of cumulative variability for CTQs. Therefore, scenarios with correlation structure:

- Very low, low and medium: PC_1, PC_2 and PC_3 were analyzed;
- High: PC_1 and PC_2 were analyzed;
- Very high: only PC_1 was analyzed.

In S5, S10 and S15, only PC_1 was analyzed, the results showed that PCA was capable to estimate $\%R\&R_m$ within the confidence interval. For other scenarios, PC_1 was insufficient to provide a reasonable explanation of variability of the CTQs. Thus, when other principal components were analyzed, $\%R\&R_m$ index was estimated outside the confidence interval (except for S11). In short, when the correlation structure between the CTQs requires that other principal components are analyzed, besides PC_1, the PCA method may fail (see Table 3).

For the inter-method analysis by MANOVA, Table 4 presents how the $\%R\&R_m$ index was estimated for the 15 simulated scenarios. It verifies that this method was capable to estimate the multivariate index within of the confidence interval only in S9, S11 and S14. This index was obtained by MANOVA using geometric mean of $\sqrt{\lambda_{MS}/\lambda_T}$ according to the amount of quality characteristics. This simulation study dealt with four characteristics. Therefore, four eigenvalues of the $\hat{\Sigma}_{MS}$ and $\hat{\Sigma}_T$ matrices were extracted. If the individual relationship $\sqrt{\lambda_{MS}/\lambda_T}$ for each pair of eigenvalues, 1, 2, 3 and 4, in $\hat{\Sigma}_{MS}$ and $\hat{\Sigma}_T$, provide different interpretations, the $\%R\&R_m$ index estimated by MANOVA may not represent well the performance of the measurement system. This is due to the fact that the geometric mean provides the same degree of importance in the analysis of each pair of eigenvalues. However, it is known that the first eigenvalues have a greater percentage of explaining the measured phenomenon greater than the last eigenvalues. Therefore, the need is confirmed that some form of weighting for the calculation of this index should be used.

Table 4: $\%R\&R_m$ index for the inter-method analysis by MANOVA

S	Mean	LLCI	ULCI	$\sqrt{\frac{\lambda_{MS_1}}{\lambda_{T_1}}}$	$\sqrt{\frac{\lambda_{MS_2}}{\lambda_{T_2}}}$	$\sqrt{\frac{\lambda_{MS_3}}{\lambda_{T_3}}}$	$\sqrt{\frac{\lambda_{MS_4}}{\lambda_{T_4}}}$	$\left(\prod_{i=1}^{4}\sqrt{\frac{\lambda_{MS_i}}{\lambda_{T_i}}}\right)^{1/4}$
S1	40.42	29.69	51.14	**49.07 (61.2)**	3.29 (27.4)	5.13 (10.5)	16.34 (0.9)	10.78
S2	45.44	34.42	56.47	**45.90 (74.6)**	9.35 (13.9)	8.32 (7.7)	8.76 (3.9)	13.30
S3	43.18	32.63	53.72	**45.97 (80.4)**	5.81 (10.8)	7.80 (5.2)	7.87 (3.6)	11.32
S4	41.86	31.70	52.03	**43.30 (87.3)**	52.70 (8.6)	14.66 (3.1)	18.78 (1.1)	28.15
S5	36.21	29.45	42.97	**37.75 (99.8)**	51.77 (0.2)	98.97 (0.0)	92.13 (0.0)	64.09
S6	13.48	9.75	17.21	**15.94 (53.4)**	2.14 (29.0)	2.40 (16.3)	7.43 (1.3)	4.97
S7	22.82	16.95	28.69	**27.39 (67.7)**	4.44 (18.3)	3.37 (13.9)	**24.73 (0.2)**	10.04
S8	17.69	11.16	24.21	**18.15 (81.2)**	3.00 (11.0)	4.45 (4.6)	3.51 (3.2)	5.40
S9	13.50	9.19	17.80	**14.50 (90.4)**	**14.24 (6.9)**	**10.03 (2.6)**	20.27 (0.1)	**14.31**
S10	18.70	14.80	22.59	**16.92 (99.9)**	56.75 (0.0)	70.15 (0.0)	73.86 (0.0)	47.23
S11	6.22	3.67	8.77	**7.27 (50.4)**	2.80 (34.5)	2.21 (34.5)	**6.14 (14.2)**	**4.08**
S12	5.54	4.15	6.92	**6.23 (71.5)**	0.95 (15.5)	0.73 (12.7)	3.80 (0.3)	2.01
S13	7.07	4.37	9.76	**7.11 (81.5)**	1.34 (10.6)	1.69 (4.6)	1.68 (3.3)	2.28
S14	5.84	4.00	7.69	**6.59 (90.8)**	**7.39 (6.4)**	**4.10 (2.8)**	13.63 (0.1)	**7.22**
S15	7.95	6.07	9.83	**7.78 (100.0)**	38.19 (0.0)	98.44 (0.0)	82.04 (0.0)	39.35

In the inter-method analysis by WPC, Table 3 shows that the $\%R\&R_m$ index was estimated within the confidence interval for the 15 simulated scenarios. WPC was more robust than PCA and MANOVA because it overcame some shortcomings of these methods. For PCA, when PC_1 is insufficient to explain the whole variability of CTQs, other principal components can provide evaluations for the measurement system outside the confidence interval. MANOVA provides a general interpretation for the measurement system, however the strategy of using geometric mean was not satisfactory. In WPC, the strategy of weighting the scores of principal components by their eigenvalues proved to be sufficient to correct the shortcomings previously mentioned. Moreover, the simulation study showed that the higher correlations between the CTQs, the closer to the mean value will be the estimates of $\%R\&R_m$ using the WPC method.

4 Conclusion

This article addressed the multivariate analysis of measurement systems through studies of repeatability and reproducibility of the measurement process. The main contribution of this paper is its proposal for a new method for multivariate analysis of the measurement system by weighting the principal components. To prove the efficiency of the method, simulated data were generated with different correlation structures for measurement systems considered acceptable, marginal, and unacceptable. The results obtained by WPC method were compared to those obtained by the multivariate methods (PCA and MANOVA). Moreover, statistical analysis provided the following conclusions:

- PCA method may fail when correlation structure between CTQs is not considered sufficiently high and more than the first principal component is required to be analyzed;
- MANOVA method uses geometric mean to estimate multivariate index for evaluating the measurement system. This approach may be incorrect when the relations $\sqrt{\lambda_{MS}/\lambda_T}$ for each *q* pair of eigenvalues provide significant difference for their calculations;
- Taking all situations of this simulation study into account, WPC method showed more robust than PCA and MANOVA method. WPC was able to overcome shortcomings such as: to provide an single assessment for all CTQs in multivariate GR&R study; to estimate the multivariate $\%R\&R_m$ index inside the confidence interval even when the correlation structure of CTQs is considered low; and to provide a strategy of weighting that guarantee greater importance for principal components most statistically significant to estimate the $\%R\&R_m$ index.

5 Acknowledgments

The authors would like to thank FAPEMIG, CAPES and CNPq for their support in this research.

References

AIAG. (2010). *Measurement Systems Analysis: Reference Manual*, fourth ed. Automotive Industry Action Group, Detroit, MI.

Al-Refaie, A., & Bata, N. (2010). Evaluating measurement and process capabilities by GR&R with four quality measures. *Measurement,* 43, 842-851.

Awad, M., Erdmann, T.P., Shanshal, Y., & Barth, B. (2009). A measurement system analysis approach for hard-to-repeat events. *Quality Engineering,* 21, 300-305.

Burdick, R.K., Borror, C.M., & Montgomery, D.C., 2003. A review of methods for measurement systems capability analysis. *Journal of Quality Technology,* 35, 342-354.

Dejaegher, B., Jimidar, M., De Smet, M., Cockaerts, P., Smeyers-Verbeke, J., & Vander Heyden, Y., 2006. Improving method capability of a drug substance HPLC assay. *Journal of Pharmaceutical and Biomedical Analysis* 42, 155–170.

Deldossi, L., & Zappa, D., 2011. Measurement uncertainty with nested mixed effect models. *Quality and Reliability Engineering International* 27, 373-379.

Erdmann, T.P., Does, R.J.M.M., & Bisgaard, S., 2010. Quality quandaries: a gage R&R study in a hospital. *Quality Engineering* 22, 46-53.

He, S.G., Wang, G.A., & Cook, D.F., 2011. Multivariate measurement system analysis in multisite testing: An online technique using principal component analysis. *Expert Systems with Applications* 38, 14602-14608.

Johnson, R.A., & Wichern, D., 2002. *Applied Multivariate Statistical Analysis*. fifth ed. Prentice-Hall, New Jersey.

Kaija, K., Pekkanen, V., Mäntysalo, M., Koskinen, S., Niittynen, J., Halonen, E., & Mansikkamäki, P., 2010. Inkjetting dielectric layer for electronic applications. *Microelectronic Engineering* 87, 1984–1991.

Li, M.H.C., & Al-Refaie, A., 2008. Improving wooden parts' quality by adopting DMAIC procedure. *Quality and Reliability Engineering International* 24, 351-360.

Lyu, J., & Chen, M.N., 2008. Gauge capability studies for attribute data, *Quality and Reliability Engineering International* 24, 71-82.

Majeske, K.D., 2008. Approval criteria for multivariate measurement systems. *Journal of Quality Technology* 40, 140-154.

Paiva, A.P., Costa, S.C., Paiva, E.J., Balestrassi, P.P., & Ferreira, J.R., 2010. Multi-objective optimization of pulsed gas metal arc welding process based on weighted principal component scores. *International Journal of Advanced Manufacturing Technology* 50, 113-125.

Paiva, A.P., Ferreira, J.R., & Balestrassi, P.P., 2007. A multivariate hybrid approach applied to AISI 52100 hardened steel turning optimization. *Journal of Materials Processing Technology* 189, 35-26.

Paiva, A.P., Paiva, E.J., Ferreira, J.R., Balestrassi, P.P., & Costa, S.C., 2008. A multivariate mean square error optimization of AISI hardened steel turning. *International Journal of Advanced Manufacturing Technology* 43, 631-643.

Polini, & W., Turchetta, S., 2004. Test protocol for micro-geometric wear of sintered diamond tools. Wear 257, 246-256.

Senol, S., 2004. Measurement system analysis using designed experiments with minimum α-β Risks and n. *Measurement* 36, 131-141.

Wang, F.K., & Chien, T.W., 2010. Process-oriented basis representation for a multivariate gauge study. *Computers and Industrial Engineering* 58, 143-150.

Wang, F.K., & Yang, C.W., 2007. Applying principal component analysis to a GR&R study. *Journal of the Chinese Institute of Industrial Engineering* 24, 182-189.

Woodall, W.H., & Borror, C.M., 2008. Some relationships between gage R&R criteria. *Quality and Reliability Engineering International* 24, 99-106.